Ringraziamenti

Desidero ringraziare sentitamente i professori Sergio Benedetto e Roberto Garello per la proficua e stimolante collaborazione. I suggerimenti del professor Garello, sia per quanto riguarda l'implementazione dell'algoritmo di decodifica dei codici LDPC, che per lo studio delle costellazioni non binarie, sono stati davvero preziosi.

Ringrazio inoltre l'ingegner Alberto Perotti per il sopporto fornito durante la mia permanenza presso l'Istituto Superiore Mario Boella, e la dottoressa Gabriella Bosco per il paziente aiuto datomi con il codice per le simulazioni.

Indice

Introduzione **1**

1 Il teorema di Shannon **5**
1.1 Canali senza memoria con ingresso discreto 5
1.2 Codici a blocco e probabilità di errore 8
1.3 La decodifica ML e il Gallager bound 10
1.4 Dimostrazione del teorema di Shannon con media sul random coding ensemble 12

2 Canali G–simmetrici e codici a blocco G–lineari **17**
2.1 Canali G–simmetrici 17
2.2 Costellazioni geometricamente uniformi e canale gaussiano additivo 19
2.3 Spettri di distanze e stime della probabilità di errore dei codici a blocco su G su canali G–simmetrici 25
2.4 Codici G–lineari 30

3 Ensemble classici di codici $\mathbb{Z}_m$–lineari su canali $\mathbb{Z}_m$–simmetrici **33**
3.1 Il teorema inverso di codifica per codici $\mathbb{Z}_m$–liberi su canali $\mathbb{Z}_m$–simmetrici 34
3.2 Ensemble di codici immagine di omomorfismi di $\mathbb{Z}_m$–moduli . 36
3.3 Ensemble di codici nucleo di omomorfismi di $\mathbb{Z}_m$–moduli . . . 39
3.4 $\mathbb{Z}_m$–capacità di un canale $\mathbb{Z}_m$–simmetrico 43
3.5 La costellazione 2^r–PSK con gruppo generatore $\mathbb{Z}_{2^r}$ 46
3.6 Conclusioni 53
3.6.1 Generalizzazione al caso di uscite continue 53
3.6.2 Costellazione 3–PSK×2–PAM 54
3.6.3 Costellazione $(2^r, h)$–PSK 56

4 I codici a bassa densità su $\mathbb{Z}_m$ con decodifica ML **57**
4.1 Costruzione dell'ensemble dei codici a bassa densità 57
4.2 Una stima dall'alto alla probabilità media di errore di ensemble arbitrari di codici $\mathbb{Z}_m$–lineari 59
4.3 Una prima stima di $\mathbb{P}(\mathbf{x} \in \mathcal{C})$ 64
4.4 Una nuova stima di $\mathbb{P}(\mathbf{x} \in \mathcal{C})$ 71
4.5 La probabilità di errore dei codici a bassa densità su $\mathbb{Z}_m$ . . . 75
4.6 "Expurgation is possible" 87

Conclusioni **92**

A **93**
A.1 Notazioni di teoria dei gruppi 93
A.2 Il metodo dei tipi 94
A.3 Aritmetica 96

Introduzione

La teoria della trasmissione numerica su un canale disturbato, nata negli anni '40 con i primi lavori di Shannon [24], e sviluppatasi poi negli anni '50 e '60, ha vissuto nell'ultimo decennio una rapida evoluzione, a partire dall'introduzione nel 1993 degli schemi noti come *turbo codici* [5]. Tali schemi si basano sull'uso di codificatori convoluzionali concatenati, in serie o in parallelo, e di un algoritmo di decodifica iterativo subottimo, che garantisce un eccellente compromesso tra *prestazioni*, in termini di probabilità di errore in funzione del rapporto tra potenza del segnale trasmesso e del rumore, e *complessità*, in termini di numero di operazioni di macchina necessarie per la decodifica. L'introduzione dei turbo codici ha costituito un enorme passo in avanti per la teoria dell'informazione perché per la prima volta si sono raggiunti risultati molto vicini ai limiti indicati dalla teoria di Shannon. Da un punto di vista pratico, le ottime prestazioni dei turbo codici e la loro bassa complessità ne hanno permesso l'introduzione come standard in numerose importanti applicazioni quali ad esempio la comunicazione cellulare di terza generazione (UMTS).

Nella seconda metà degli anni '90 poi, in seguito al grande interesse nato nella comunità scientifico–ingegneristica intorno ai turbo codici, fu riscoperta un'altra famiglia di codici altrettanto efficiente, quella dei *codici a bassa densità* (LDPC). I codici a bassa densità erano stati introdotti da Gallager nella sua tesi di dottorato del 1960 [13], poi sostanzialmente dimenticati (anche perché la tecnologia di allora non permetteva di effettuare simulazioni che ne mostrassero l'effettiva bontà), quindi riscoperti nel 1995, vedi [20], [26] [34]. Anche i codici LDPC sono decodificati con un algoritmo iterativo subottimo, noto con il nome di 'belief propagation' [29], che sfrutta la sparsità del grafo di Tanner associato a tali codici per limitare la complessità computazionale: è stato dimostrato, nel contesto dei grafi fattoriali, che i due algoritmi di decodifica dei codici LDPC e dei codici turbo sono istanze di un medesimo, più generale algoritmo (vedi [17]). I codici a bassa densità con decodifica belief propagation hanno mostrato di avere prestazioni paragonabili, e in alcuni casi migliori di quelle dei turbo codici (si veda ad esempio [8]). L'ottimo

compromesso tra performance e complessità che garantiscono ne ha favorito l'applicazione pratica: tra le più importanti, la prima è costituita dai nuovi standard per il broadcasting video digitale via satellite.

Finora, sia i turbo codici che i codici LDPC sono stati usati soprattutto con modulazioni binarie, al fine di ottenere grandi *guadagni di codifica*. Il numero sempre crescente di applicazioni caratterizzate da congestione della banda di frequenze disponibili, motiva la ricerca di schemi di trasmissione che coniughino tali guadagni di codifica con un'elevata *efficienza spettrale*. Questo ha incentivato i tentativi di progettare codici turbo e a bassa densità congiuntamente a schemi di modulazione di alto livello non binari.

Esistono sostanzialmente due approcci a tale problema: uno che possiamo chiamare *pragmatico* e un altro *unificato*. Nell'approccio pragmatico si procede progettando e ottimizzando indipendentemente un codice binario e una modulazione di alto ordine; si procede poi ad ottimizzare la mappatura dei blocchi di bit sui segnali della modulazione. Il principale vantaggio di questo approccio è costituito dalla possibilità di applicarlo a trasmissioni nelle quali il rate deve poter essere variato in tempo reale in funzione del canale; gli inconvenienti sono che i margini di ottimizzazione congiunta sono scarsi (ridotti alla sola mappatura dei bit sui segnali della modulazione), e che le prestazioni possono essere stimate essenzialmente solo per mezzo di simulazioni numeriche, non fornendo la loro analisi teorica risultati apprezzabili. Nell'approccio unificato, invece, i codici sono progettati e ottimizzati tenendo conto della costellazione scelta: questa soluzione presenta una minore flessibilità nelle applicazioni, ma consente di avere più ampi margini di ottimizzazione.

Molte modulazioni importanti sono basate su *costellazioni* simmetriche che ammettono struttura di gruppo compatibile con la distanza euclidea. In questo caso è naturale considerare codici che abbiano struttura di gruppo compatibile con quella della costellazione [12], [18]. In particolare i codici geometricamente uniformi godono della 'Uniform Error Property', i.e. l'indipendenza della parola trasmessa, proprietà che ne semplifica considerevolmente lo studio teorico, rendendo effettivamente possibile determinarne le prestazioni per via analitica.

Uno degli esempi fondamentali che consideriamo in questa tesi è la costellazione m–PSK, che ammette struttura di gruppo $\mathbb{Z}_m$. I codici lineari su $\mathbb{Z}_m$ hanno naturalmente struttura di modulo e si prestano a considerazioni simili a quelle impiegate nel caso dei codici su campi. Nel caso binario è ben noto che i codici lineari permettono di raggiungere la capacità di Shannon sul canale gaussiano additivo con modulazione 2–PAM [15]. Uno dei risultati principali di questa tesi è l'estensione di tale affermazione al caso del canale gaussiano additivo con costellazione 2^r–PSK: si tratta di un risultato non

banale dal momento che, come verrà mostrato, altre costellazioni presentano ostruzioni algebriche all'uso di codici $\mathbb{Z}_m$–lineari.

Vengono poi analizzate le prestazioni dei codici a bassa densità su $\mathbb{Z}_m$. Tale analisi è effettuata supponendo di utilizzare la decodifica ottima, quella di massima verosiglianza (ML), seguendo la linea di [21], [23], [4]. Come già detto, nella pratica non si usa questo tipo di decodifica perché troppo onerosa dal punto di vista computazionale, ma quella subottima dell'algoritmo belief propagation. Il vantaggio dell'analisi dei codici LDPC con decodifica ML è che questa è notevolmente più semplice: essa fornisce delle stime dall'alto delle prestazioni con decodifica belief propagation e permette inoltre di individuare criteri per l'analisi della struttura intrinseca di tali codici (girth, etc.). Il risultato principale dell'analisi dei codici LDPC su $\mathbb{Z}_m$ svolta in questa tesi è che, per delle opportune scelte dei parametri, tali codici, così come quelli binari (si vedano [13], [21]) raggiungono le prestazioni dei codici $\mathbb{Z}_m$–lineari. Nel caso del canale gaussiano aditivo con modulazione 2^r–PSK, possiamo concludere quindi, grazie al risultato ottenuto per i codici lineari cui accennavamo prima, che i codici a bassa densità raggiungono capacità.

Nel primo capitolo richiamiamo alcuni risultati della teoria classica di Shannon con l'intento di introdurne gli strumenti fondamentali, strumenti che vengono utilizzati nei capitoli seguenti: in particolare l'idea chiave di mediare su *ensemble* di codici piuttosto che analizzare le prestazioni di un singolo codice, per poi fare affermazioni di tipo probabilistico su tali ensemble.

Nel secondo capitolo si introduce una classe di canali simmetrici, famiglia che comprende il canale gaussiano additivo con gli ingressi vincolati a stare su una costellazione geometricamente uniforme e alcune sue opportune quantizzazioni. Si ricavano poi due stime per le prestazioni di singoli codici a blocco su tali canali: la prima è il classico Battacharyya bound, mentre l'altra fa uso delle tecniche di randomizzazione per codici deterministici recentemente introdotte in [25].

Nel terzo capitolo vengono studiate le prestazioni di ensemble classici di codici $\mathbb{Z}_m$–lineari su canali $\mathbb{Z}_m$–simmetrici. Viene introdotto un nuovo parametro di soglia per la comunicazione affidabile con codici $\mathbb{Z}_m$–lineari, la $\mathbb{Z}_m$–*capacità* del canale: si mostra come per alcune costellazioni tale parametro coincida con la capacità classica, mentre per altre ne è strettamente inferiore. Le prime costellazioni non presentano alcuna ostruzione algebrica all'uso di codici a blocco $\mathbb{Z}_m$–liberi (tra queste la 2^r–PSK), mentre le seconde sì (tra queste la 2–PAM$\times$3–PSK).

Nel quarto capitolo vengono introdotti i codici a bassa densità su $\mathbb{Z}_m$ a partire dal loro grafo di Tanner. Si ricava, facendo uso di tecniche simili a quelle di [23] e [4], una stima della probabilità di errore di ensemble arbitrari di codici a blocco $\mathbb{Z}_m$–lineari in funzione degli spettri medi di distanze del-

l'ensemble, stima studiata ad hoc per gli ensemble di codici a bassa densità. Dallo studio di tali spettri medi di distanze si ottengono l'esatto andamento asintotico della probabilità media di errore degli ensemble di codici LDPC, andamenti che sono polinomialmente decrescenti a 0, quindi non esponenziali come quelli degli ensemble di codici $\mathbb{Z}_m$–lineari. Infine, si applicano tecniche di *expurgation* per ottenere andamenti asintotici arbitrariamente vicini a quelli degli ensemble classici $\mathbb{Z}_m$–lineari.

Capitolo 1

Il teorema di Shannon

1.1 Canali senza memoria con ingresso discreto

Fissato un insieme finito $\mathcal{X}$, indichiamo con $\mathcal{P}(\mathcal{X})$ l'insieme delle distribuzioni di probabilità su $\mathcal{X}$. $\mathcal{P}(\mathcal{X})$ è un sottoinsieme convesso e compatto di $\mathbb{R}^{|\mathcal{X}|}$. In questa tesi useremo, fissata arbitrariamente una base $a > 0$, le notazioni $\log(x) := \log_a(x)$ e $\exp(x) := a^x$. L'*entropia* di una distribuzione $p \in \mathcal{P}(\mathcal{X})$ è data da

$$H(p) := -\sum_{x \in \mathcal{X}} p(x) \log p(x).$$

Se la base del log è 2 l'unità di misura dell'entropia è il bit (binary digit), se è il numero di Nepero e l'unità di misura è il nat. La funzione

$$H : \mathcal{P}(\mathcal{X}) \to [0, \log |\mathcal{X}|]$$

è continua e concava.

Se X è una variabile aleatoria su $\mathcal{X}$ di distribuzione $p_X \in \mathcal{P}(\mathcal{X})$, chiamiamo entropia di X e la indichiamo con $H(X)$ il valore $H(p_X)$. Intuitivamente $H(X)$ è una misura dell'incertezza della variabile aleatoria X: $H(X)$ assume il suo valore massimo $\log |\mathcal{X}|$ quando X è distribuita uniformemente su $\mathcal{X}$ (massima incertezza), mentre $H(X)$ si annulla quando X è costante (nessuna incertezza). Infine, se $x \in [0, 1]$, poniamo $H(x)$ pari all'entropia di una variabile aleatoria di distribuzione Bernoulli di parametro x. Tale abuso di notazione non causerà problemi quando sarà chiara la natura dell'argomento di H.

Date due variabili aleatorie discrete X e Y su $\mathcal{X}$ e $\mathcal{Y}$ rispettivamente, con $\mathcal{X}$ finito e $\mathcal{Y}$ finito o numerabile, di distribuzione congiunta $p_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$,

definiamo l'entropia di Y condizionata a X come

$$H(Y|X) := \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p_{XY}(x,y) \log p_{Y|X}(x|y) \ ,$$

dove $p_{Y|X}$ è la distribuzione di X condizionata a Y. Definiamo poi la *mutua informazione* di X e Y

$$I(X;Y) := H(Y) - H(Y|X) \ \ .$$

Introduciamo ora il concetto di *canale discreto senza memoria* (DMC). Formalmente un DMC con insieme degli ingressi finito $\mathcal{X}$, e insieme delle uscite $\mathcal{Y}$ finito o numerabile, è una famiglia di distribuzioni di probabilità di transizione

$$\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}} \ \ .$$

Intuitivamente $W(y|x)$ rappresenta la probabilità che in uscita del canale si abbia il valore y, dato che in ingresso c'è il valore x. La proprietà di assenza di memoria si estrinseca nel fatto che l'utilizzo del canale per N volte consecutive è descritto dalle probabilità di transizione

$$W_N(\mathbf{y}|\mathbf{x}) := \prod_{i=1}^{N} W(y_i|x_i) \ ;$$

in altri termini l'i–esima uscita è influenzata solo dall'i–esimo ingresso, e tale dipendenza è invariante nel tempo, cioè dalla posizione i nella sequenza.

Sia X una variabile aleatoria di distribuzione p_X. A partire da X e dal canale di probabilità di transizione W si può definire una variabile aleatoria Y su $\mathcal{Y}$ descritta congiuntamente ad X dalla distribuzione $p_{XY} \in \mathcal{P}(\mathcal{X}\times\mathcal{Y})$ data da

$$p_{XY}(x,y) = P_X(x)W(y|x) \ .$$

Si noti che la distribuzione marginale di Y è quindi data da

$$p_Y(y) = \sum_{x\in\mathcal{X}} p_X(x)W(y|x) \ .$$

Con questo modello la mutua informazione di X e Y è funzione solo della distribuzione in ingresso $p_X \in \mathcal{P}(\mathcal{X})$ e del canale $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}}$:

$$I(X;Y) := \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} p_X(x)W(y|x) \log \left(\frac{W(y|x)}{\sum\limits_{z\in\mathcal{X}} p_X(z)W(y|z)} \right) \ .$$

Per ogni canale discreto senza memoria $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}}$ fissato,

$$\mathcal{P}(\mathcal{X}) \ni p_X \mapsto I(X;Y) \in [0, log|\mathcal{X}|]$$

è una funzione continua e concava. Il massimo valore della mutua informazione $I(X;Y)$ al variare di p_X in $\mathcal{P}(\mathcal{X})$ dunque esiste poiché si tratta di una funzione continua definita su un insieme compatto. Tale valore, funzione del canale, viene chiamato *capacità* e indicato con

$$C = C(\{W\}) := \max_{p_X\in\mathcal{P}(\mathcal{X})} I(X;Y) \ .$$

Il teorema di Shannon, che enunceremo formalmente nel paragrafo 1.2, caratterizza la capacità di un canale senza memoria come un limite fondamentale alla quantità di informazione che si può trasmettere in modo affidabile sul canale stesso.

Possiamo generalizzare al caso di insieme delle uscite continuo $\mathcal{Y} = \mathbb{R}^n$, sempre con insieme degli ingressi $\mathcal{X}$ finito. Sia $\mathcal{P}(\mathbb{R}^n)$ l'insieme delle densità di probabilità su $\mathbb{R}^n$. Un canale senza memoria con ingressi discreti $\mathcal{X}$ e uscite continue $\mathbb{R}^n$ una famiglia di densità di probabilità su $\mathbb{R}^n$

$$\{W(\cdot|x) \in \mathcal{P}(\mathbb{R}^n)\}_{x\in\mathcal{X}} \ .$$

Siano ora X una variabile aleatoria a valori in $\mathcal{X}$ di distribuzione $p_X \in \mathcal{P}(\mathcal{X})$, e Y una variabile aleatoria a valori in $\mathbb{R}^n$ distribuita congiuntamente a X secondo la legge $p_{XY} \in \mathcal{P}(\mathcal{X} \times \mathbb{R}^n)$

$$p_{XY}(A, x) = p_X(x) \int_A W(y|x)\mathrm{d}y, \quad A \in \mathcal{B}(\mathbb{R}^n), x \in \mathcal{X} \ ,$$

dove $\mathcal{B}(\mathbb{R}^n)$ indica l'insieme dei Boreliani di $\mathbb{R}^n$. La mutua informazione di X e Y si definisce come

$$I(X;Y) := \sum_{x\in\mathcal{X}} p_X(x) \int_{\mathbb{R}^n} W(y|x) \log \left(\frac{W(y|x)}{\sum\limits_{z\in\mathcal{X}} p_X(z)W(y|z)} \right) \mathrm{d}y \ .$$

Quando tale valore è ben definito per ogni $p_X \in \mathcal{P}(\mathcal{X})$, la funzione

$$\mathcal{P}(\mathcal{X}) \ni p_X \mapsto I(X;Y) \in \mathbb{R}$$

è a valori positivi, convessa e continua. Analogamente al caso di $\mathcal{Y}$ discreto, si definisce la capacità del canale $\{W(\cdot|x) \in \mathcal{P}(\mathbb{R}^n)\}_{x\in\mathcal{X}}$ come

$$C := \max_{p_X\in\mathcal{P}(\mathcal{X})} I(X;Y) \ .$$

Si può mostrare (vedi [15], [9]) che la capacità di un canale senza memoria ad ingresso discreto e uscita continua può essere espressa come l'estremo superiore della capacità dei DMC che si ottengono quantizzando l'uscita. In particolare si supponga che, per ogni $x \in \mathcal{X}$, $W(\cdot|x)$ sia una densità di probabilità su $\mathbb{R}^n$ tale che $W(\cdot|x)\log(W(\cdot|x))$ sia integrabile secondo Riemann. Per ogni $\Delta > 0$, si consideri il canale discreto senza memoria

$$\{W_\Delta(\cdot|x) \in \mathcal{P}(\mathbb{Z}^n)\}_{x\in\mathcal{X}}$$

definito da

$$W_\Delta(\mathbf{i}|x) := \int_{I_{\Delta,\mathbf{i}}} W(\mathrm{d}y|x),$$

dove

$$I_{\Delta,\mathbf{i}} := ((i_1-1)\Delta, i_1\Delta] \times \ldots \times ((i_n-1)\Delta, i_n\Delta].$$

Sia C_Δ la capacità di tale canale. Allora

$$C = \lim_{\Delta\to 0} C_\Delta \ .$$

1.2 Codici a blocco e probabilità di errore

Supporremo, a partire da questo momento, che esista uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ sul quale siano ben definite tutte le variabili aleatorie che consideremo nel seguito.

Fissiamo un canale di trasmissione senza memoria ad ingresso discreto $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}}$.

Un *codice a blocco* $\mathcal{C}$ di lunghezza N e cardinalità M su $\mathcal{X}$ è un elemento di $(\mathcal{X}^N)^M$, i.e.

$$\mathcal{C} = (\mathbf{x}_1, \ldots, \mathbf{x}_M), \quad \mathbf{x}_j \in \mathcal{X}^N \ .$$

Il *rate* di $\mathcal{C}$ è definito come il rapporto tra il logaritmo della sua cardinalità e la sua lunghezza

$$R := \frac{\log M}{N} \ .$$

Come si vede, un codice stato definito come una M–upla ordinata di parole in $\mathcal{X}^N$; sono possibili dunque ripetizioni della stessa parola. Definiamo allora il *supporto* di un codice a blocco $\mathcal{C}$ come

$$\operatorname{supp}\mathcal{C} = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\} \ .$$

Diciamo inoltre che una parola di informazione $i \in \{1, \ldots, M\}$ è *degenere* se $\mathbf{x}_j$ appare più di una volta nel codice $\mathcal{C}$, i.e. se

$$\exists j \in \{1, \ldots, M\} \setminus \{i\} \ \text{ t.c. } \mathbf{x}_j = \mathbf{x}_i \ .$$

Un codice a blocco di cardinalità M si dice *non degenere* se tutte le parole di informazione $i \in \{1, \dots M\}$ sono non degeneri; si osservi che per un codice non degenere $\mathcal{C}$ si ha

$$|\operatorname{supp} \mathcal{C}| = M \ .$$

Sia dato un codice a blocco $\mathcal{C} \in (\mathcal{X}^N)^M$. Sia $\mathbf{U}$ una variabile aleatoria di distribuzione uniforme sull'insieme $\{1, \dots, M\}$; $\mathbf{U}$ rappresenta la parola di informazione da trasmettere. Associamo ad $\mathbf{U}$ in modo deterministico la variabile aleatoria $\mathbf{X} = \mathbf{x}_{\mathbf{U}}$ a valori in $\mathcal{X}^N$. Da X e dal canale si ottiene al solito la variabile aleatoria $\mathbf{Y}$ su $\mathcal{Y}^N$. La descrizione probabilistica della coppia $(\mathbf{U}, \mathbf{Y})$ è data dalla distribuzione $p_{U,\mathbf{Y}} \in \mathcal{P}(\{1, \dots, M\} \times \mathcal{Y}^N)$ definita da

$$p_{\mathbf{U},\mathbf{Y}}(j, \mathbf{y}) = \frac{1}{M} W_N(\mathbf{y}|\mathbf{x}_j) \ .$$

Un *decodificatore* per il codice $\mathcal{C}$ è una mappa

$$\Phi_{\mathcal{C}} : \mathcal{Y}^N \to \{0, \dots, M\} \ .$$

Il valore $\Phi_{\mathcal{C}}(\mathbf{y})$ rappresenta la stima da parte del ricevitore della parola di informazione $\mathbf{U}$ trasmessa, a partire dall'osservazione dell'uscita $\mathbf{y}$ del canale; ammettiamo che il decodificatore possa assumere il valore 0 in corrispondenza di un segnale di mancata decodifica.

La *probabilità di errore* della coppia codice–decodificatore $P(e|\mathcal{C}, \Phi_{\mathcal{C}})$ viene definita come la probabilità dell'evento $\Phi_{\mathcal{C}}(\mathbf{Y}) \neq \mathbf{U}$, i.e.

$$P(e|\mathcal{C}, \Phi_{\mathcal{C}}) := \mathbb{P}\left(\{\Phi_{\mathcal{C}}(\mathbf{Y}) \neq \mathbf{U}\}\right) = \frac{1}{M} \sum_{i=1}^{M} P(e|\mathcal{C}, \Phi_{\mathcal{C}}, i) \ ,$$

dove

$$P(e|\mathcal{C}, \Phi_{\mathcal{C}}, i) := \mathbb{P}(\Phi_{\mathcal{C}}(\mathbf{Y}) \neq \mathbf{U}|\mathbf{U} = i) = \mathbb{P}(\Phi_{\mathcal{C}} \neq i|\mathbf{U} = i) \ .$$

Usiamo la notazione $P(e|\mathcal{C}, \Phi_{\mathcal{C}})$ piuttosto che $\mathbb{P}(e|\mathcal{C}, \Phi_{\mathcal{C}})$ per enfatizzare il fatto che la probabilità di errore è funzione deterministica della coppia $(\mathcal{C}, \Phi_{\mathcal{C}})$, una volta fissato il canale di trasmissione.

La probabilità di errore costituisce il parametro fondamentale di misura delle prestazioni di un sistema di trasmissione.

Possiamo ora enunciare una versione del teorema di Shannon. Tale teorema afferma che la trasmissione affidabile, i.e. con probabilità di errore che possa essere resa arbitrariamente piccola, è possibile solo a rate minori della capacità del canale.

Teorema 1 (Shannon)
Sia fissato un canale senza memoria con ingresso discreto $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}}$ *di capacità* C*. Allora:*

- *per ogni* R *tale che* $0 \leq R < C$*, e per ogni* $\varepsilon > 0$*, esistono un codice a blocco* $\mathcal{C}$ *su* $\mathcal{X}$ *di rate* $\hat{R} \geq R$ *ed un decodificatore* $\Phi_{\mathcal{C}}$ *tali che*

$$P(e|\mathcal{C}, \Phi_{\mathcal{C}}) < \varepsilon \ ;$$

- *per ogni* R *tale che* $R > C$*, esiste una costante* $K > 0$ *tale che qualunque codice a blocco* $\mathcal{C}$ *su* $\mathcal{X}$ *di rate* R*, con qualsiasi decodificatore* $\Phi_{\mathcal{C}}$*, abbia probabilità di errore limitata dal basso da* K*:*

$$P(e|\mathcal{C}, \Phi_{\mathcal{C}}) > K \ .$$

La dimostrazione della seconda affermazione –nota come teorema inverso di codifica del canale– si ottiene piuttosto agevolmente con considerazioni di teoria dell'informazione e si può trovare ad esempio in [9]. Nel seguito ci soffermeremo sulla dimostrazione della parte diretta in quanto le tecniche dimostrative risulteranno fondamentali per ottenere i risultati di questa tesi.

1.3 La decodifica ML e il Gallager bound

Fissati un canale discreto senza memoria $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}}$ ed un codice a blocco $\mathcal{C}$ su $\mathcal{X}$ di lunghezza N e cardinalità M, ci proponiamo di trovare il decodificatore $\Phi_{\mathcal{C}} : \mathcal{C}^N \to \{0, 1, \ldots, M\}$ che minimizzi la probabilità di errore $P(e|\mathcal{C}, \Phi_{\mathcal{C}})$.

Usiamo, fissato un insieme A, la notazione $\mathbb{1}_A(\mathbf{x})$ per denotarne la funzione indicatrice. Possiamo esprimere la probabilità di errore della coppia $(\mathcal{C}, \Phi_{\mathcal{C}})$ come

$$\begin{aligned} P(e|\mathcal{C}) &= \tfrac{1}{M} \sum_{m=1}^{M} \sum_{y\in\mathcal{Y}} W_N(\mathbf{y}|\mathbf{x}_m) \mathbb{1}_{\{\mathbf{z}:\Phi_{\mathcal{C}}(\mathbf{z})\neq m\}}(\mathbf{y}) \\ &= 1 - \tfrac{1}{M} \sum_{m=1}^{M} \sum_{y\in\mathcal{Y}} W_N(\mathbf{y}|\mathbf{x}_m) \mathbb{1}_{\{\mathbf{z}:\Phi_{\mathcal{C}}(\mathbf{z})=m\}}(\mathbf{y}) \ . \end{aligned}$$

Per ogni $\mathbf{y} \in \mathcal{Y}^N$ si ha che, indicando con $\delta_{i,j}$ la delta di Kronecker,

$$\operatorname*{argmax}_{j\in\{0,1,\ldots M\}} \sum_{m=1}^{M} W_N(\mathbf{y}|\mathbf{x}_m)\delta_{j,m} = \operatorname*{argmax}_{j\in\{1,\ldots M\}} W_N(\mathbf{y}|\mathbf{x}_j) \ .$$

Definiamo il *decodificatore a massima verosimiglianza* (ML) come

$$\Phi_{ML}(\mathbf{y}) = \begin{cases} m & \text{se} \quad W_N(\mathbf{y}|\mathbf{x}_m) > W_N(\mathbf{y}|\mathbf{x}_n) \ , \ \forall n \neq m \\ 0 & \text{se} \quad \nexists \text{ tale } m \end{cases} \ . \tag{1.1}$$

Si è scelto dunque di risolvere i casi dubbi, in cui il massimo della funzione di verosimiglianza

$$i \longmapsto W(\mathbf{y}|\mathbf{x}_i)$$

non è unico in $\{1, \dots, M\}$, con un messaggio di mancata decodifica. Tale scelta è leggermente subottima, in quanto nei suddetti casi si potrebbe decodificare in uno dei punti di massimo della funzione di verosimiglianza, arbitrariamente scelto. Tuttavia la definizione (1.1) ci permetterà di semplificare le analisi successive, senza pregiudicarne i risultati. Si osservi che dalla (1.1) segue che, se $i \in \{1, \dots, M\}$ è una parola di informazione degenere, allora

$$P(e|\mathcal{C}, \Phi_{ML}, i) = 1 \ .$$

Dunque i codici degeneri non godono di buone probabilità di errore; tuttavia conviene considerare anche i codici degeneri per ragioni che saranno chiare nel seguito. Si osservi inoltre che per i codici non degeneri la probabilità di errore con decodifica ML dipende solo dal supporto del codice e non dal suo ordinamento.

D'ora in poi prenderemo sempre in considerazione, quando non diversamente specificato, il caso di decodifica di massima verosimiglianza definita nella (1.1). Useremo al posto di $P(e|\mathcal{C}, \Phi_{ML})$ e $P(e|\mathcal{C}, \Phi_{ML}, i)$ le notazioni più compatte $P(e|\mathcal{C})$ e $P(e|\mathcal{C}, i)$ rispettivamente, per indicare la probabilità di errore del codice $\mathcal{C}$ con decodifica ML.

La stima seguente ha una semplice dimostrazione, ma risulta fondamentale per la dimostrazione del teorema di Shannon.

Lemma 2 (Gallager bound)
Siano dati un canale senza memoria $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathcal{X}}$ e un codice a blocco $\mathcal{C}$ su $\mathcal{X}$ di lunghezza N e cardinalità M. Per ogni $\rho \geq 0$, $1 \leq m \leq M$, la probabilità di errore di $\mathcal{C}$ condizionata alla trasmissione dell' m–esima parola, con decodifica ML, soddisfa

$$P(e|\mathcal{C}, m) \leq \sum_{\mathbf{y} \in \mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_m)^{\frac{1}{1+\rho}} \Big(\sum_{m' \neq m} W_N(\mathbf{y}|\mathbf{x}_{m'})^{\frac{1}{1+\rho}} \Big)^{\rho} \ . \tag{1.2}$$

Dimostrazione
Introduciamo le regioni di decisione

$$D_m = \left\{ y \in \mathcal{Y}^N \ : \ W_N(\mathbf{y}|\mathbf{x}_m) > W_N(\mathbf{y}|\mathbf{x}_{m'}), \forall m' \neq m) \right\}, \quad 1 \leq m \leq N \ .$$

Per $\lambda > 0$ arbitrario e $1 \leq m \leq M$, si definiscano gli insiemi

$$\Lambda_m := \left\{ \mathbf{y} \in \mathcal{Y}^N \ : \ \sum_{m' \neq m} \left(\frac{W_N(\mathbf{y}|\mathbf{x}_{m'})}{W_N(\mathbf{y}|\mathbf{x}_m)} \right)^{\lambda} \geq 1 \right\}.$$

Si verifica che $(D_m)^c \subseteq \Lambda_m$. Quindi, per ogni $\rho \geq 0$,

$$\begin{aligned}
\mathbb{P}(e|\mathcal{C}, m) &= \sum_{\mathbf{y} \in \mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_m) \mathbb{1}_{(D_m)^c}(\mathbf{y}) \leq \\
&\leq \sum_{\mathbf{y} \in \mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_m) \mathbb{1}_{\Lambda_m}(\mathbf{y}) \leq \\
&\leq \sum_{\mathbf{y} \in \mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_m) \left(\sum_{m' \neq m} \left(\frac{W_N(\mathbf{y}|\mathbf{x}_{m'})}{W_N(\mathbf{y}|\mathbf{x}_m)} \right)^{\lambda} \right)^{\rho} = \\
&= \sum_{\mathbf{y} \in \mathcal{Y}^N} W(\mathbf{y}|\mathbf{x}_m)^{1-\lambda\rho} \left(\sum_{m' \neq m} W_N(\mathbf{y}|\mathbf{x}_{m'})^{\lambda} \right)^{\rho},
\end{aligned}$$

e, scegliendo $\lambda = 1/(1+\rho)$, si ottiene la (1.2). ■

Il Gallager bound costituisce una generalizzazione dello union bound, che si ritrova prendendo $\rho = 1$ nella (1.2).

1.4 Dimostrazione del teorema di Shannon con media sul random coding ensemble

Vi sono molte diverse dimostrazioni del teorema di Shannon. Praticamente tutte si fondano su uno stesso principio, che è esattamente l'idea base di Shannon: l'introduzione del concetto di *ensemble* di codici. Invece di analizzare la probabilità di errore di un singolo codice, compito reso difficile dal fatto che la sua cardinalità cresce esponenzialmente con N una volta fissato il rate R, se ne studia la media su un opportuno insieme di codici; tale analisi risulta spesso più facile perchè la statistica su un ensemble permette di mediare quantità intrattabili per un singolo codice.

Esponiamo ora la dimostrazione della prima parte del teorema 1, così come si trova nei testi classici [15] e [33]; il che ci permetterà di formalizzare una serie di risultati che saranno utili per le estensioni successive.

Fissato un canale discreto senza memoria $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathcal{X}}$, sia, per ogni $0 \leq \rho \leq 1$, $p \in \mathcal{P}(\mathcal{X})$

$$E_0(\rho, p) := -log \left(\sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} p(x) W(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} \right).$$

Possiamo ora definire l'*esponente di errore* del canale

$$E : [0, \log |\mathcal{X}|] \to \mathbb{R}$$

$$E(R) := \max_{0 \le \rho \le 1} \left\{ \max_{p \in \mathcal{P}(\mathcal{X})} \{E_0(\rho, p)\} - \rho R \right\} . \tag{1.3}$$

Teorema 3 (Gallager)
Per ogni DMC l'esponente di errore $E(R)$ è una funzione continua, decrescente, convessa, strettamente positiva per $R \in [0, C)$, e $E(C) = 0$. Inoltre $E(R)$ dipende con continuità dal canale $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathcal{X}}$.

Dimostrazione
Per la dimostrazione della prima parte dell'enunciato si rimanda ai testi classici [15] e [33]. Per quanto riguarda la dipendenza continua da W, basta osservare che

$$(\{W\}, p, \rho) \mapsto E_0(\{W\}, \rho, p) - \rho R$$

è uniformemente continua perché continua sull'insieme compatto $(\mathcal{P}(\mathcal{Y}))^{|\mathcal{X}|} \times \mathcal{P}(\mathcal{X}) \times [0, 1]$. Da questo segue con semplici ragionamenti la continuità di $E(R)$ come funzione di $\{W\}$. ■

Introduciamo il random coding ensemble $\mathcal{E}_{RC}$. Fissati $R \in [0, \log |\mathcal{X}|]$ ed $N \in \mathbb{N}$, poniamo

$$M = \lceil \exp(NR) \rceil \ .$$

Sia

$$\mathcal{C} = (\mathbf{X}_1, \ldots, \mathbf{X}_M)$$

un vettore di M variabili aleatorie su $\mathcal{X}^N$ indipendenti tra loro, indipendenti dalla parola di informazione $\mathbf{U}$ e dal canale, e uniformemente distribuite con legge

$$p_N(\mathbf{x}) = \prod_{i=1}^{N} p(x_i),$$

dove $p \in \mathcal{P}(\mathcal{X})$ qualsiasi. Queste variabili aleatorie inducono naturalmente una struttura probabilistica sull'insieme $(\mathcal{X}^N)^M$ dei codici a blocco su $\mathcal{X}$ di lunghezza N e cardinalità M. Questo spazio di probabilità si indica con $\mathcal{E}_{RC}(N, R)$ ed è noto come *random coding ensemble*.

Definiamo le *probabilità medie di errore* sull'ensemble

$$\overline{P(e|m)} := \mathbb{E}_{\mathcal{C}}[P(e|m, \mathcal{C})], \quad \overline{P(e)} := \mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C})].$$

Si osservi che, poiché le parole di codice sono realizzazioni di variabili aleatorie indipendenti e identicamente distribuite, $\overline{P(e|m)}$ non dipende da m, e quindi

$$\overline{P(e)} = \mathbb{E}_{\mathcal{C}}\big[\mathbb{E}_{\mathbf{U}} P(e|\mathbf{U}, \mathcal{C})\big] = \mathbb{E}_{\mathcal{C}}\big[P(e|\mathcal{C}, 1)\big] = \overline{P(e|1)}.$$

Teorema 4
Siano dati un canale senza memoria con ingresso discreto $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}}$ *di esponente di errore* $E(R)$*, e un* $R \in [0, \log|\mathcal{X}|]$*. Si supponga di utilizzare decodifica ML. Allora, per ogni* $N \in \mathbb{N}$*, la probabilità media di errore del random coding ensemble* $\mathcal{E}_{RC}(N, R)$ *soddisfa*

$$\overline{P(e)} \leq \exp(-NE(R)). \tag{1.4}$$

Dimostrazione
Utilizzando nell'ordine il Gallager bound per $0 \leq \rho \leq 1$ fissato, la disuguaglianza di Jensen, l'indipendenza delle $\mathbf{X}_i$ (si noti come sia sufficiente l'indipendenza a coppie) e la loro equidistribuzione, si ottiene

$$\begin{aligned}
\overline{P(e)} &= \mathbb{E}_{\mathcal{C}}[P(e|1,\mathcal{C})] \\
&\leq \mathbb{E}_{\mathbf{X}_1}\ldots\mathbb{E}_{\mathbf{X}_M}\left(\sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{X}_1)^{\frac{1}{1+\rho}}\Big(\sum_{m=2}^{M} W_N(\mathbf{y}|\mathbf{X}_m)^{\frac{1}{1+\rho}}\Big)^{\rho}\right) \\
&\leq \sum_{\mathbf{y}\in\mathcal{Y}^N}\mathbb{E}_{\mathbf{X}_1}[W_N(\mathbf{y}|\mathbf{X}_1)^{\frac{1}{1+\rho}}]\left(\sum_{m=2}^{M}\mathbb{E}_{\mathbf{X}_m}[W_N(\mathbf{y}|\mathbf{X}_m)^{\frac{1}{1+\rho}}]\right)^{\rho} \\
&= (M-1)^{\rho}\sum_{\mathbf{y}\in\mathcal{Y}^N}\Big(\mathbb{E}_{\mathbf{X}_1}[W_N(\mathbf{y}|\mathbf{X}_1)^{\frac{1}{1+\rho}}]\Big)^{1+\rho} \\
&= (M-1)^{\rho}\left(\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}} p(x)W(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right)^{N} \\
&\leq \exp(-N\left[E_0(\rho,p)-\rho R\right]).
\end{aligned}$$

Per il penultimo passaggio è stata usata l'ipotesi che il canale sia senza memoria e che i simboli di ciascuna parola di codice siano scelti i.i.d.. Dall'arbitrarietà di p in $\mathcal{P}(\mathcal{X})$ e di ρ in $[0,1]$ segue la (1.4). ■

Dal teorema precedente segue immediatamente la parte diretta del teorema di Shannon. Consideriamo una successione di random coding ensemble di rate R fissato

$$(\mathcal{E}_{RC}(R,N))_{N\in\mathbb{N}}\ ,$$

indipendenti tra loro. Facciamo osservare che l'esistenza di uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ sul quale sono ben definiti questi ensemble è un fatto non ovvio (si veda [16] Cap.10) la cui dimostrazione esula dagli scopi di questa tesi.

Poiché per ogni N il minimo della probabilità di errore sull'ensemble $\mathcal{E}_{RC}(N, R)$ è minore o uguale della media, ne segue che esiste una successione

di codici a blocco di rate R

$$(\mathcal{C}_N)_N := \left(\operatorname*{argmin}_{\mathcal{C} \in (\mathcal{X}^N)^M} P(e|\mathcal{C}) \right)_N$$

tale che

$$P(e|\mathcal{C}_N) \leq \exp(-NE(R)). \tag{1.5}$$

Poichè, per il Teorema 3, $E(R) > 0$ per ogni $R < C$, la (1.5) garantisce che $P(e|\mathcal{C}_N)$ converga a zero esponenzialmente per $N \to +\infty$.

Si osservi che il random coding ensemble $\mathcal{E}_{RC}(R, N)$ contiene tutti i codici a blocco di rate R e lunghezza N, compresi quelli degeneri. Per eliminare i codici degeneri, infatti, sarebbe stato necessario introdurre delle dipendenze probabilistiche tra le $\mathbf{X}_i$; questo avrebbe reso molto più complessa l'analisi successiva.

Dalla (1.4) si possono trarre anche conclusioni ben più forti della (1.5). Ad esempio, poichè $P(e|\mathcal{C})$ è una variabile aleatoria a valori non negativi il cui valore atteso è finito, possiamo applicare la disuguaglianza di Markov. Arbitrariamente fissato un valore $\varepsilon > 0$, si definiscano gli insiemi

$$A_N^\varepsilon := \left\{ \mathcal{C} \in (\mathcal{X}^N)^M \text{ t.c. } P(e|\mathcal{C}) \geq \exp(-N[E(R) - \varepsilon]) \right\} .$$

Abbiamo allora che

$$\begin{aligned} \mathbb{P}(A_N^\varepsilon) &= \mathbb{P}\big(P(e|\mathcal{C}) \geq \exp(-N[E(R) - \varepsilon])\big) \\ &\leq \mathbb{P}\big(P(e|\mathcal{C}) \geq \exp(N\varepsilon)\overline{P(e)}\big) \\ &\leq \exp(-N\varepsilon) \quad , \end{aligned}$$

e quindi

$$\sum_N \mathbb{P}(A_N^\varepsilon) \leq \sum_N \exp(-N\varepsilon) < +\infty \ .$$

Il teorema di Borel–Cantelli permette di concludere quindi che

$$\mathbb{P}\Big(\limsup_{N \to +\infty} A_N^\varepsilon \Big) = 0 \ ;$$

dunque, fissato un rate $R < C$, quasi ogni successione $(\mathcal{C}_N)_N$ di codici a blocco realizzazioni dei random coding ensemble $\mathcal{E}_{RC}(R, N)$ ha probabilità di errore esponenzialmente decrescente a 0 per N sufficientemente grande, con esponente arbitrariamente vicino all'esponente di errore del canale $E(R)$.

Quindi la teoria di Shannon indica anche, in un certo senso, come costruire un codice a blocco di probabilità di errore fissata, con decodifica di massima verosimiglianza. Il problema è che tale codice viene generato privo, almeno

a priori, di una qualsiasi struttura. Ciò comporta che la sua decodifica ML sia un problema ad alta complessità computazionale; in effetti è possibile mostrare (vedi [6]) che si tratta di un problema NP–hard.

Come si è appena visto, quindi, lo studio del comportamento asintotico della probabilità media di errore su un ensemble di codici sia da un punto di vista teorico per affermare l'esistenza di un codice di prestazioni arbitrariamente buone, sia da un punto di vista pratico perchè indica un algoritmo, seppur non deterministico, per la generazione con una probabilità fissata di codici di prestazioni arbitrariamente buone. In questa tesi il concetto di ensemble verrà largamente usato; introduciamo dunque delle notazioni.

Definizione 1
Fissati un insieme finito $\mathcal{X}$, un intero positivo N ed un valore $R \in [0, \log|\mathcal{X}|]$, un ensemble *$\mathcal{E}(R, N)$ è un insieme di codici a blocco di lunghezza N e rate maggiore o uguale a R, con una struttura probabilistica. Dato un canale senza memoria di ingressi $\mathcal{X}$, indichiamo il valor medio della probabilità di errore dell'ensemble $\mathcal{E}(R, N)$ con decodifica ML su tale canale con*

$$\overline{P(e)} := \mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C})] \ .$$

Fissati un canale senza memoria di ingressi $\mathcal{X}$ ed un valore $R \in [0, \log|\mathcal{X}|]$, considereremo successioni di ensemble indipendenti di codici a blocco di rate maggiore o uguale a R

$$(\mathcal{E}(N, R))_{N\in\mathbb{N}} \ .$$

Diremo che una successione di ensemble indipendenti di codici a blocco su $\mathcal{X}$ di rate maggiore o uguale a R è *molto buona* per tale canale se la successione delle probabilità medie di errore sugli ensemble tende a zero, i.e.

$$\mathbb{E}_{\mathcal{E}(N,R)}[P(e|\mathcal{C})] \to 0, \quad N \to +\infty \ .$$

Diremo invece che una successione di ensemble indipendenti $(\mathcal{E}(N, R))_{N\in\mathbb{N}}$ è *buona* se esiste $\varepsilon > 0$ tale che per ogni canale senza memoria di ingressi $\mathcal{X}$ di capacità

$$C \geq \log|\mathcal{X}| - \varepsilon$$

la successione delle probabilità medie di errore sugli ensemble tende a zero, i.e.

$$\mathbb{E}_{\mathcal{E}(N,R)}[P(e|\mathcal{C})] \to 0, \quad N \to +\infty \ .$$

Si osservi la differenza sostanziale che intercorre tra le due definizioni appena date: per parlare di una famiglia di ensemble molto buona si fissa un canale ed un rate sotto la sua capacità, mentre per definire una famiglia buona si fissa soltanto il rate e si lascia variare il canale.

Capitolo 2

Canali G–simmetrici e codici a blocco G–lineari

Nel capitolo precedente i canali senza memoria sono stati studiati indipendentemente dall'eventuale struttura algebrica degli insiemi di ingresso $\mathcal{X}$ e di uscita $\mathcal{Y}$. In questo capitolo viene introdotta una struttura algebrica su tali insiemi e si studia una classe di canali simmetrici analizzando le principali proprietà di cui tali canali godono. Si mostra poi il contesto tipico in cui nascono i canali G–simmetrici, cioè gli schemi di modulazione su costellazioni di segnali geometricamente uniformi usate in ingresso di un canale gaussiano additivo. Infine si introduce la classe dei codici G–lineari e se ne propongono stime per la probabilità di errore su canali G–simmetrici.

2.1 Canali G–simmetrici

Introduciamo ora un concetto di simmetria per i canali senza memoria.

Definizione 2
Dati due insiemi finiti $\mathcal{X}$ e $\mathcal{Y}$ e un gruppo G, un canale senza memoria $\{W(\cdot|x)\}_{x\in\mathcal{X}}$ si dice debolmente G–simmetrico *se*

- *G agisce transitivamente su $\mathcal{X}$;*
- *G agisce su $\mathcal{Y}$;*
- $W(gy|gx) = W(y|x), \quad \forall g \in G, x \in \mathcal{X}, y \in \mathcal{Y}.$

Se inoltre l'azione di G su $\mathcal{X}$ è semplicemente transitiva, il canale si dice G–simmetrico.

In seguito, per un canale G–simmetrico identificheremo sempre l'insieme degli ingressi $\mathcal{X}$ con G.

Si osservi inoltre che, fissato un G–insieme $\mathcal{Y}$, l'insieme dei canali G–simmetrici di uscite $\mathcal{Y}$ è può essere messo in corrispondenza biunivoca con l'insieme $\mathcal{P}(\mathcal{Y})$ delle distribuzioni di probabilità su $\mathcal{Y}$, tramite l'applicazione

$$\mathcal{P}(\mathcal{Y}) \ni w \stackrel{\Psi}{\longmapsto} \{W(\cdot|g) := w(g^{-1}(\cdot)) \in \mathcal{P}(\mathcal{Y})\}_{g\in G} \ .$$

In effetti è immediato verificare che $\Psi(w)$ è un canale G–simmetrico per ogni $w \in \mathcal{P}(\mathcal{Y})$, e che Ψ è iniettiva. Inoltre ogni canale G–simmetrico $\{W(\cdot|g)\}_{g\in G}$ si può porre

$$\{W(\cdot|g)\}_{g\in G} = \Psi(W(\cdot|g_0)) \in \mathcal{P}(\mathcal{Y}) \ ;$$

dunque Ψ è suriettiva.

Proposizione 5
La capacità C e l'esponente di errore $E(R)$ di un canale senza memoria debolmente G–simmetrico $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathcal{X}}$ si ottengono con distribuzione in ingresso uniforme.

Dimostrazione
1.Cominciamo col mostrare che $H(Y|X)$ è indipendente dalla distribuzione in ingresso $q \in \mathcal{P}(\mathcal{X})$. Sia $x_0 \in \mathcal{X}$ qualsiasi;

$$\begin{aligned} H(Y|X) &= -\sum_{x\in\mathcal{X}} q(x) \sum_{y\in\mathcal{Y}} W(y|x)\log(W(y|x)) \\ &= -\tfrac{1}{|Stab(x_0)|} \sum_{g\in G} q(gx_0) \sum_{y\in\mathcal{Y}} W(y|gx_0)\log(W(y|gx_0)) \\ &= -\tfrac{1}{|Stab(x_0)|} \sum_{g\in G} q(gx_0) \sum_{y\in\mathcal{Y}} W(g^{-1}y|x_0)\log(W(g^{-1}y|x_0)) \\ &= -\sum_{y\in\mathcal{Y}} W(y|x_0)\log(W(y|x_0)). \end{aligned}$$

Si ha quindi che

$$\operatorname*{argmax}_{p_X\in\mathcal{P}(\mathcal{X})}\{H(Y) - H(Y|X)\} = \operatorname*{argmax}_{p_X\in\mathcal{P}(\mathcal{X})}\{H(Y)\} := \operatorname*{argmax}_{p_X\in\mathcal{P}(\mathcal{X})}\{f(p_X)\} \ .$$

La funzione f è invariante rispetto a G, nel senso che $f(p) = f(gp)$ per ogni $p \in \mathcal{P}(\mathcal{X})$, $g \in G$ (con la convenzione $gp(x) = p(gx)$); infatti

$$\begin{aligned} f(gp) &= -\sum_{y\in\mathcal{Y}} (\sum_{x\in\mathcal{X}} p(gx)W(y|x))\log(\sum_{x\in\mathcal{X}} p(gx)W(y|x)) \\ &= -\sum_{y\in\mathcal{Y}} (\sum_{x\in\mathcal{X}} p(gx)W(gy|gx))\log(\sum_{x\in\mathcal{X}} p(gx)W(gy|gx)) \\ &= -\sum_{y\in\mathcal{Y}} (\sum_{x\in\mathcal{X}} p(x)W(gy|x))\log(\sum_{x\in\mathcal{X}} p(x)W(gy|x)) \\ &= -\sum_{y\in\mathcal{Y}} (\sum_{x\in\mathcal{X}} p(x)W(y|x))\log(\sum_{x\in\mathcal{X}} p(x)W(y|x)) = f(p). \end{aligned} \tag{2.1}$$

Indichiamo con $u_{\mathcal{X}} \in \mathcal{P}(\mathcal{X})$ la distribuzione uniforme su $\mathcal{X}$. Dalla (2.1) e dalla concavità di f su $\mathcal{P}(\mathcal{X})$, segue che, per ogni $p \in \mathcal{P}(\mathcal{X})$ fissata,

$$f(u_{\mathcal{X}}) = f\Big(\frac{1}{|G|}\sum_{g\in G} gp\Big) \geq \frac{1}{|G|}\sum_{g\in G} f(gp) = f(p)$$

e dunque

$$\operatorname*{argmax}_{p\in\mathcal{P}(\mathcal{X})} f(p) = u_{\mathcal{X}}.$$

2. Per ogni $\rho \in [0,1]$ fissato,

$$\operatorname*{argmax}_{p\in\mathcal{P}(\mathcal{X})}(E_0(\rho,p)) = \operatorname*{argmin}_{p\in\mathcal{P}(\mathcal{X})}(\exp(-E_0(\rho,p))) := \operatorname*{argmin}_{p\in\mathcal{P}(\mathcal{X})}(f(p)).$$

La funzione f è convessa su $\mathcal{P}(\mathcal{X})$ e G–invariante (si vede con ragionamento analogo alla(2.1)). Quindi, per ogni $p \in \mathcal{P}(\mathcal{X})$,

$$f(u_{\mathcal{X}}) = f\Big(\frac{1}{|G|}\sum_{g\in G} gp\Big) \geq \frac{1}{|G|}\sum_{g\in G} f(gp) = f(p)$$

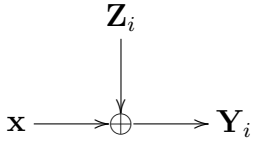
e dunque

$$\operatorname*{argmin}_{p\in\mathcal{P}(\mathcal{X})} f(p) = u_{\mathcal{X}}.$$

■

2.2 Costellazioni geometricamente uniformi e canale gaussiano additivo

La comunicazione digitale avviene nella realtà convertendo successioni di simboli di informazione in sequenze di forme d'onda che poi vengono effettivamente trasmesse su un canale fisico. Tale operazione di conversione viene chiamata modulazione; le forme d'onda utilizzate dal modulatore possono essere convenientemente rappresentate come punti di uno spazio euclideo di dimensione finita. Chiamiamo *costellazione* n–dimensionale un sottoinsieme finito S di $\mathbb{R}^n$ che genera $\mathbb{R}^n$; S rappresenta l'insieme dei segnali utilizzati dal modulatore. Il modello di canale che consideriamo è il *canale gaussiano additivo* (AWGN). Si tratta di un canale senza memoria a tempo discreto con uscite $\mathbf{Y}_i \in \mathbb{R}^n$ al tempo i, dove $\mathbf{Y}_i$ è la somma dell'ingresso $\mathbf{x}$ e di un rumore $\mathbf{Z}_i$.



Il rumore $(\mathbf{Z}_i)_i$ è una successione di variabili aleatorie indipendenti tra loro e da $\mathbf{X}_i$, e di identica distribuzione normale n–variata di media $\mathbf{0}$ e matrice di covarianza diagonale $\frac{N_0}{2}\mathrm{I}_n$, dove $N_0 \geq 0$ è un valore associato all'intesità del rumore, i.e.

$$\mathbf{Y}_i = \mathbf{X}_i + \mathbf{Z}_i, \qquad \mathbf{Z}_i \sim \mathcal{N}(\mathbf{0}, \tfrac{N_0}{2}\mathrm{I}_n) \ . \tag{2.2}$$

Il canale gaussiano additivo è un buon modello per una vasta gamma di canali di telecomunicazione usati nella pratica, incluse le connessioni radio e via satellite. La validità di tale modello si giustifica da un punto di vista fisico nella maniera seguente. Il rumore additivo può essere dovuto a varie cause che supponiamo in prima approssimazione indipendenti tra loro; inoltre è possibile supporre che tali sorgenti di rumore siano indipendenti tra una trasmissione e l'altra. Il teorema centrale del limite assicura allora che l'effetto cumulato di un gran numero di tali rumori possa essere correttamente approssimato con una variabile aleatoria di distribuzione normale.

Introduciamo ora il concetto di costellazione geometricamente uniforme. Useremo la notazione standard $\mathrm{Iso}(\mathbb{R}^n)$ per indicare il gruppo delle simmetrie di $\mathbb{R}^n$, i.e. il gruppo delle trasformazioni $T : \mathbb{R}^n \to \mathbb{R}^n$ che lasciano invariata la distanza euclidea:

$$|T(\mathbf{x}_1) - T(\mathbf{x}_2)| = |\mathbf{x}_1 - \mathbf{x}_2|, \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n.$$

Sia S una costellazione di $\mathbb{R}^n$. Il gruppo di simmetria di S è il sottogruppo di $\mathrm{Iso}(\mathbb{R}^n)$ delle isometrie che lasciano invariato S; tale sottogruppo viene indicato con $\Gamma(S)$. S si dice geometricamente uniforme se l'azione di $\Gamma(S)$ su S è transitiva. In altri termini possiamo dare la definizione seguente.

Definizione 3
Una costellazione $S \subset \mathbb{R}^n$ *si dice* geometricamente uniforme (GU) *se*

$$\forall s_1, s_2 \in S \quad \exists g \in \Gamma(S) \ \textit{t.c.} \ gs_1 = s_2 \ .$$

Sia S una costellazione geometricamente uniforme; ci si può chiedere se esiste un sottogruppo di $\Gamma(S)$ la cui azione su S sia semplicemente transitiva, cioè tale che S possa essere espressa come orbita di un punto $s_0 \in S$ sotto l'azione di G. In caso affermativo tale sottogruppo viene si dice gruppo generatore di S.

Definizione 4
$G \leq \Gamma(S)$ *è un* gruppo generatore *di* S *se*

$$\forall s_1, s_2 \in S, \quad \exists \ ! \ g \in G \quad \textit{t.c.} \quad gs_1 = s_2.$$

Figura 2.1: La costellazione 2–PAM con labeling $\mathbb{Z}_2$

Non tutte le costellazioni geometricamente uniformi ammettono gruppo generatore: sono stati costruiti dei controesempi espliciti (si veda [32]). In genere tuttavia prenderemo in considerazione solo costellazioni dotate di gruppo generatore. Quando una costellazione geometricamente uniforme S ammette gruppo generatore G, è possibile mettere in corrispondenza biunivoca G ed S ed introdurre in tal modo una struttura di gruppo su S. Si fissi infatti un punto arbitrario $s_0 \in S$; definiamo un labeling $\mu_{s_0} : G \to S$ come

$$\mu_{s_0}(g) = gs_0.$$

Allora S è un gruppo rispetto alla legge di composizione

$$s_1 \cdot s_2 := \mu_{x_0}\left(\mu_{x_0}^{-1}(s_1) \cdot \mu_{x_0}^{-1}(s_2)\right).$$

Mostriamo ora alcuni esempi di costellazioni geometricamente uniformi e di gruppi generatori.

Esempio 1
Un primo semplice esempio è costituito dalla costellazione monodimensionale 2–PAM (acronimo di Pulse Amplitude Modulation):

$$2 - \mathrm{PAM} = \{-L, L\}\ ,$$

dove $L > 0$ fissato. Evidentemente

$$\Gamma\big(2 - \mathrm{PAM}\big) = \{1, r\}$$

dove r è la riflessione intorno all'origine e 1 l'identità. Tale gruppo genera 2–PAM ed è chiaramente isomorfo al campo binario $\mathbb{Z}_2$. □

Esempio 2
La costellazione m–PSK (acronimo di Phase Shift Keying) è definita come un insieme di m punti di $\mathbb{R}^2$ equispaziati sulla circonferenza di raggio $L > 0$ fissato:

$$m - \mathrm{PSK} := \left\{ Le^{\frac{2\pi k}{m} i} \mid k = 0, \dots m-1 \right\}\ .$$

Il suo gruppo di simmetria è

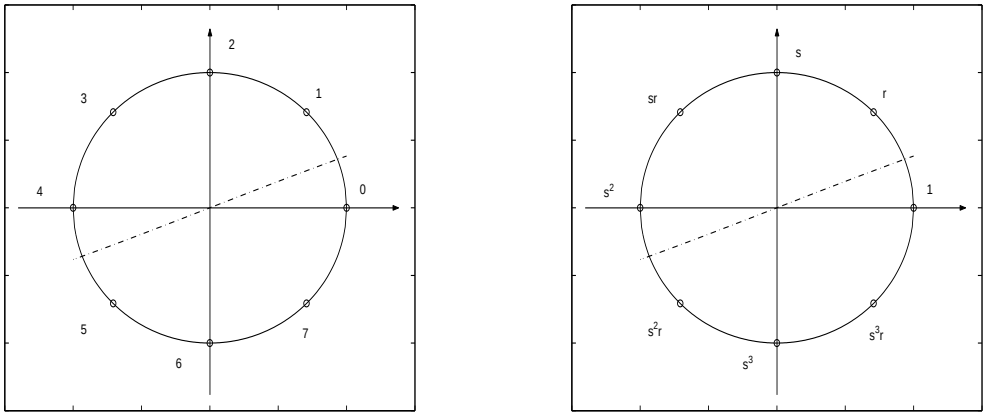


Figura 2.2: La costellazione 8–PSK con labeling $\mathbb{Z}_8$ e D_4

$$\Gamma(m-\text{PSK}) = R \ltimes V_m \ ,$$

dove R è il gruppo generato dalla riflessione intorno ad un qualsiasi asse passante per l'origine ed inclinato di un angolo multiplo dispari di $\frac{2\pi}{2m}$ rispetto all'asse delle ascisse, V_m è gruppo delle rotazioni intorno all'origine di un angolo multiplo di $\frac{2\pi}{m}$, e $\ltimes$ indica il prodotto semidiretto di due gruppi. Questo gruppo è isomorfo al gruppo diedrale D_m.

Un gruppo generatore per l'm–PSK è dunque $G = V_m$; tale gruppo è isomorfo a $\mathbb{Z}_m$.

Se m è pari esiste un altro gruppo generatore dell'm–PSK, $G' = R \ltimes V_{m/2}$. G' è isomorfo al gruppo diedrale $D_{m/2}$ che è non abeliano.

Nella figura 2.2 è riportata a titolo di esempio la costellazione 8–PSK con i due possibili labeling $\mathbb{Z}_8$ e D_4. □

Esempio 3
Fissati un numero pari m, e due parametri $L, h \in (0, +\infty)$, consideriamo la costellazione tridimensionale

$$(m,h)-\text{PSK}:=\left\{ \left(Le^{\frac{2\pi i}{m}(2j+k)}, (-1)^j Lh \right) \;\middle|\; k=0,\ldots,\tfrac{m}{2}-1,\ j=0,1 \right\}$$

ottenuta come unione di due costellazioni $\frac{m}{2}$–PSK di raggio L, poste su due piani paralleli Π_1 e Π_2, entrambi di distanza $Lh > 0$ dall'origine (figura 2.3) e ruotate di un angolo $\alpha = \frac{2\pi}{m}$ tra loro. Sia Π_0 il piano passante per l'origine parallelo a Π_1 e Π_2. Si osservi che al limite per $h \to 0$, tenendo costante $L > 0$, tale costellazione degenera nella m–PSK giacente su Π_0.

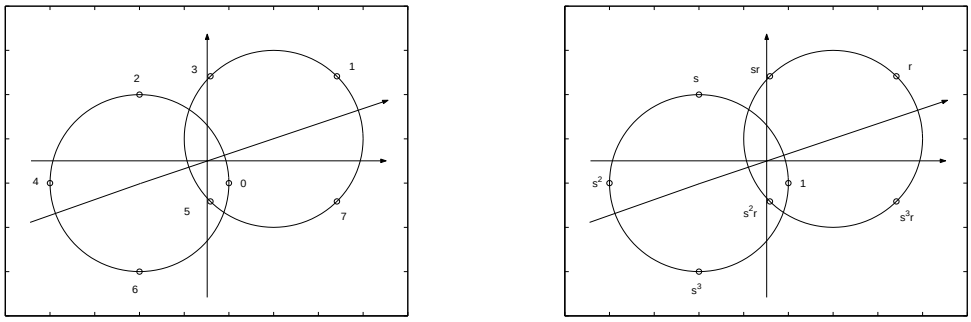


Figura 2.3: La costellazione $(8, h)$–PSK con labeling $\mathbb{Z}_8$ e D_4

Il gruppo di simmetria dell'(m, h)–PSK è

$$\Gamma\big((m, h) - \mathrm{PSK}\big) = R \ltimes V_m$$

dove R è il gruppo generato dalla riflessione intorno all'asse giacente nel piano Π_0, passante per l'origine e inclinato di un angolo pari a $\frac{2\pi}{2m}$ rispetto all'asse delle ascisse.

Un gruppo generatore per S è quello generato da

$$t := rs,$$

dove s è la rotazione di un angolo $\frac{2\pi}{m}$ intorno all'asse normale ai due piani Π_1 e Π_2 passante per l'origine, e r è la riflessione rispetto al piano Π_0. Si verifica facilmente che tale gruppo è isomorfo a $\mathbb{Z}_m$.

Un altro gruppo generatore dell'$(m, h) - \mathrm{PSK}$ è $G' = R \ltimes V_{m/2}$. G' è isomorfo al gruppo diedrale $D_{m/2}$. □

Esempio 4
Fissato un numero dispari l e due parametri $L, h \in (0, +\infty)$, consideriamo la costellazione tridimensionale

$$(l - \mathrm{PSK}) \times (2 - \mathrm{PAM}) := \left\{ \big(Le^{\frac{2\pi i}{l}k}, (-1)^j Lh\big),\ k = 0, \ldots, l,\ j = 0, 1 \right\},$$

ottenuta a partire da una costellazione l–PSK giacente su un piano Π_1 distante h dall'origine, riflettendola rispetto ad un piano Π_0 parallelo a Π_1 e passante per l'origine. Il gruppo di simmetria di tale costellazione è quello generato da r, riflessione rispetto al piano Π_0, e da s, rotazione di un angolo $\frac{2\pi}{l}$ intorno all'asse ortogonale a Π_0 e passante per l'origine. Tale gruppo è anche generatore di l–PSK×2–PAM, ed è isomorfo a $\mathbb{Z}_2 \times \mathbb{Z}_l$ e dunque, poiché l è dispari, a $\mathbb{Z}_{2l}$. □

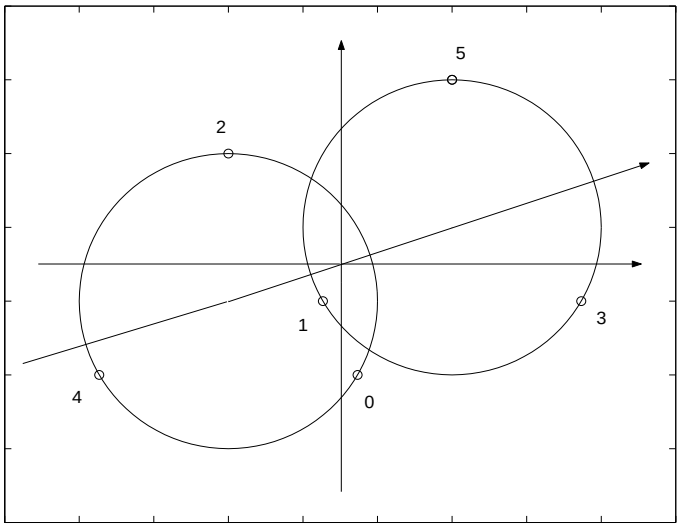


Figura 2.4: La costellazione 3–PSK×2–PAM con labeling $\mathbb{Z}_6$

Si supponga di utilizzare i segnali di una costellazione n–dimensionale geometricamente uniforme S in ingresso ad un canale gaussiano additivo. Il canale così ottenuto è di tipo discrete–input ($\mathcal{X} = S$) continuous–output ($\mathcal{Y} = \mathbb{R}^n$). Chiameremo il canale

$$\left\{ W(y|x) := \frac{1}{(\pi N_0)^{n/2}} e^{-||y-x||^2/N_0} \in \mathcal{P}(\mathbb{R}^n) \right\}_{x \in S}$$

canale gaussiano additivo di tipo S.

Sia ora $S' \subseteq \mathbb{R}^n$ una costellazione geometricamente uniforme tale che G sia un sottogruppo del gruppo di simmetria di S'. Si ipotizzi di utilizzare i segnali di S su un canale gaussiano additivo e di operare una quantizzazione delle uscite su S' nella maniera seguente. Introduciamo le *regioni di Voronoi* di S'

$$V_s := \{ r \in \mathbb{R}^n \ \text{ t.c. } \ ||r - s|| \leq ||r - s^*|| \ , \ \forall s^* \in S' \}, \qquad s \in S';$$

(non sono una partizione di $\mathbb{R}^n$, ma la loro intersezione ha misura nulla). Chiameremo *canale gaussiano additivo di tipo* (S, S') il canale discreto senza memoria

$$\left\{ W(y|x) := \int_{V_y} \frac{1}{(\sqrt{\pi N_0})^n} e^{-||t-x||^2/(N_0)} \mathrm{d}t \right\}_{x \in S} . \tag{2.3}$$

Proposizione 6
Sia $S \subset \mathbb{R}^n$ una costellazione geometricamente uniforme di gruppo generatore G, e sia $S' \in \mathbb{R}^n$ un'altra costellazione geometricamente uniforme tale che

$$G \leq \Gamma(S').$$

Allora i canali gaussiani additivi di tipo S e di tipo (S, S') sono G–simmetrici.

Dimostrazione
Iniziamo con il canale gaussiano di tipo S. Si ha che $G \leq \mathrm{Iso}(R^n)$, quindi per ogni $g \in G$

$$W(gy|gx) = \frac{1}{(\sqrt{\pi N_0})^n} e^{-||gx-gy||^2} = \frac{1}{(\sqrt{\pi N_0})^n} e^{-||x-y||^2} = W(y|x).$$

Passiamo al canale gaussiano di tipo (S, S'). G agisce isometricamente su S e S'. Le regioni di Voronoi soddisfano

$$V_{gs} = \{r \in \mathbb{R}^n \ \text{ t.c. } \ ||r - gs|| \leq ||r - gs^*|| \ , \ \forall s^* \in S'\} =$$

$$= \{r \in \mathbb{R}^n \ \text{ t.c. } \ ||g^{-1}r - s|| \leq ||g^{-1}r - s^*|| \ , \ \forall s^* \in S'\} = gV_s \ ,$$

e quindi, per ogni $g \in G$,$x \in S$,$y \in S'$,

$$\begin{aligned} W(gy|gx) &= \frac{1}{(\sqrt{2\pi\sigma^2})^n} \int_{V_{gx}} e^{-||gy-t||^2/(2\sigma^2)} \mathrm{d}t = \\ &= \frac{1}{(\sqrt{2\pi\sigma^2})^n} \int_{V_x} e^{-||gy-gt||^2/(2\sigma^2)} \mathrm{d}t = \\ &= \frac{1}{(\sqrt{2\pi\sigma^2})^n} \int_{V_x} e^{-||y-t||^2/(2\sigma^2)} \mathrm{d}t = W(y|x) \ . \end{aligned} \tag{2.4}$$

■

2.3 Spettri di distanze e stime della probabilità di errore dei codici a blocco su G su canali G–simmetrici

Come abbiamo visto, l'insieme degli ingressi di un canale G–simmetrico può essere identificato con lo stesso gruppo G. Questo giustifica lo studio ed il progetto di codici a blocco su gruppi G. Un codice a blocco su G di lunghezza N e cardinalità M è una M–upla di elementi del gruppo prodotto G^N. Proprio grazie al fatto che G^N è un gruppo, è possibile definire l'inverso di una parola $\mathbf{x}_j \in \mathcal{C}$ e il prodotto di due parole $\mathbf{x}_j, \mathbf{x}_k \in \mathcal{C}$ come rispettivamente

$$\mathbf{x}_j^{-1} \in G^N, \qquad \mathbf{x}_j \mathbf{x}_k \in G^N.$$

Tipicamente, un codice a blocco $\mathcal{C}$ su G non è chiuso rispetto a tali operazioni, i.e. non si ha

$$\mathbf{x}_j^{-1} \in \mathcal{C} \ , \quad \mathbf{x}_1\mathbf{x}_2 \in \mathcal{C} \ , \quad \forall \mathbf{x}_j, \mathbf{x}_k \in \mathcal{C}.$$

Questo è una prerogativa dei codici G–lineari, che saranno oggetto del prossimo paragrafo. Qui introduciamo invece la nozione di *spettro di distanze* di un codice a blocco su G, non necessariamente G–lineare.

Richiamiamo alcune delle notazioni del metodo dei tipi introdotte in appendice. Per ogni $\mathbf{x} \in G^N$, il *tipo* di $\mathbf{x}$ è la distribuzione $\boldsymbol{\theta}(\mathbf{x}) \in \mathcal{P}(G)$ definita da

$$\theta_a(\mathbf{x}) := \frac{1}{N}|\{j : x_j = a\}| \ .$$

Il sottoinsieme di $\mathcal{P}(G)$ di tutti i possibili tipi di elementi di G si indica con $\mathcal{P}_N(G)$. Fissato $\boldsymbol{\theta} \in \mathcal{P}_N(G)$, indichiamo con $\mathcal{T}_{\boldsymbol{\theta}}^N$ il sottoinsieme di G^N delle N–uple il cui tipo è pari a $\boldsymbol{\theta}$.

Fissiamo ora un codice a blocco $\mathcal{C} \in (G^N)^M$, un tipo $\boldsymbol{\theta} \in \mathcal{P}_N(G)$, e una parola di informazione $m = 1, \ldots, M$. Indichiamo con

$$S^r(\boldsymbol{\theta}|\mathcal{C}, m)$$

(rispettivamente $S^l(\boldsymbol{\theta}|\mathcal{C}, m)$) il numero di parole $\mathbf{x}_j$ di $\mathcal{C}$, con $j \neq m$, tali che $\mathbf{x}_m^{-1}\mathbf{x}_j \in \mathcal{T}_{\boldsymbol{\theta}}^N$ (rispettivamente $\mathbf{x}_j\mathbf{x}_m^{-1} \in \mathcal{T}_{\boldsymbol{\theta}}^N$), i.e.

$$S^r(\boldsymbol{\theta}|\mathcal{C}, m) := \sum_{j \neq m} \delta_{\{\mathbf{x}_m^{-1}\mathbf{x}_j \in \mathcal{T}_{\boldsymbol{\theta}}^N\}}, \qquad S^l(\boldsymbol{\theta}|\mathcal{C}, m) := \sum_{j \neq m} \delta_{\{\mathbf{x}_j\mathbf{x}_m^{-1} \in \mathcal{T}_{\boldsymbol{\theta}}^N\}} \ ,$$

dove stiamo usando la notazione, per $A \subseteq G^N$,

$$\delta_{\{\mathbf{x} \in A\}} := \begin{cases} 1 \text{ se } \mathbf{x} \in A \\ 0 \text{ se } \mathbf{x} \notin A \end{cases}$$

Si osservi come, se G è abeliano, allora $S^r(\boldsymbol{\theta}|\mathcal{C}, m) = S^l(\boldsymbol{\theta}|\mathcal{C}, m) := S(\boldsymbol{\theta}|\mathcal{C}, m)$.

Presentiamo ora due stime alla probabilità di errore di codici a blocco su G su canali G–simmetrici in funzione del loro spettro di distanze. Tali stime sono valide senza ipotesi di G–linearità del codice, e per G non necessariamente abeliano. La prima è un risultato classico noto come Bhattacharyya bound.

Lemma 7 (Bhattacharyya bound)
Siano dati un canale senza memoria G–simmetrico $\{W(\cdot|g) \in \mathcal{P}(\mathcal{Y})\}_{g \in G}$, ed un codice a blocco $\mathcal{C}$ su G di lunghezza N e cardinalità M

$$\mathcal{C} = (\mathbf{x}_1, \ldots, \mathbf{x}_M) \ .$$

La probabilità di errore con decodifica ML di $\mathcal{C}$, condizionata alla trasmissione della m–esima parola, soddisfa

$$P(e|\mathcal{C},m) \leq \sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)} S^x(\boldsymbol{\theta}|\mathcal{C},m)\mathbf{D}^{N\boldsymbol{\theta}} \tag{2.5}$$

con $x \in \{r,l\}$, e dove $\boldsymbol{D} = (D_g)_{g\in G}$ è il parametro di Bhattacharyya del canale

$$D_g := \sum_{y\in\mathcal{Y}} \sqrt{W(y|g)W(y|g_0)} \tag{2.6}$$

Dimostrazione
Dimostriamo per $x = r$. Richiamiamo la definizione dell'evento di errore condizionato alla trasmissione di $\mathbf{x}_m$

$$\mathbf{e}_m := \{\mathbf{y} \in \mathcal{Y} \text{ t.c. } \exists k \neq m : W_N(\mathbf{y}|\mathbf{x}_k) \geq W_N(\mathbf{y}|\mathbf{x}_m)\};$$

Per ogni $\mathbf{y} \in \mathbf{e}_m$ si ha

$$\sum_{n\neq m} \sqrt{\frac{W_N(\mathbf{y}|\mathbf{x}_n)}{W_N(\mathbf{y}|\mathbf{x}_m)}} \geq 1$$

e quindi

$$P(e|\mathcal{C},m) = \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_m)\mathbb{1}_{\mathbf{e}_m}(\mathbf{y}) \leq \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_m) \sum_{n\neq m} \sqrt{\frac{W_N(\mathbf{y}|\mathbf{x}_n)}{W_N(\mathbf{y}|\mathbf{x}_m)}}.$$

Sfruttando la G–simmetria del canale otteniamo e indicando con $\mathbf{g_0}$ l'elemento neutro di G^N, i.e. $\mathbf{g_0} = (g_0,\dots,g_0)$,

$$\begin{aligned}
P(e|\mathcal{C},m) &\leq \sum_{n\neq m}\sum_{\mathbf{y}\in\mathcal{Y}^N} \sqrt{W_N(\mathbf{y}|\mathbf{x}_n)W_N(\mathbf{y}|\mathbf{x}_m)} = \\
&= \sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)} \left(\sum_{n\neq m} \delta_{\{\mathbf{x}_m^{-1}\mathbf{x}_n \in \mathcal{T}_{\boldsymbol{\theta}}^N\}} \sum_{\mathbf{y}\in\mathcal{Y}^N} \sqrt{W_N(\mathbf{y}|\mathbf{x}_m)W_N(\mathbf{y}|\mathbf{x}_n)} \right) \\
&= \sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)} \left(\sum_{n\neq m} \delta_{\{\mathbf{x}_m^{-1}\mathbf{x}_n \in \mathcal{T}_{\boldsymbol{\theta}}^N\}} \sum_{\mathbf{y}\in\mathcal{Y}^N} \sqrt{W_N(\mathbf{x}_m^{-1}\mathbf{y}|\mathbf{g_0})W_N(\mathbf{x}_m^{-1}\mathbf{y}|\mathbf{x}_m^{-1}\mathbf{x}_n)} \right) \\
&= \sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)} \left(\sum_{n\neq m} \delta_{\{\mathbf{x}_m^{-1}\mathbf{x}_n \in \mathcal{T}_{\boldsymbol{\theta}}^N\}} \sum_{\mathbf{y}\in\mathcal{Y}^N} \sqrt{W_N(\mathbf{y}|\mathbf{g_0})W_N(\mathbf{y}|\mathbf{x}_m^{-1}\mathbf{x}_n)} \right) \\
&= \sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)} S^r(\boldsymbol{\theta}|\mathcal{C},m) \prod_{g\in G} \left(\sum_{y\in\mathcal{Y}} \sqrt{W(y|g_0)W_N(y|g)} \right)^{\theta_g N}
\end{aligned}$$

■

Passiamo alla seconda stima, che giocherà un ruolo determinante nel seguito e che è una conseguenza del Gallager bound. Le tecniche usate per la sua dimostrazione sono state introdotte in [25].

Lemma 8
Siano dati un canale G–simmetrico $\{W(\cdot|g) \in \mathcal{P}(\mathcal{Y})\}_{g\in G}$, si supponga di utilizzare la decodifica ML. Sia

$$\mathcal{C} = (\mathbf{x}_1, \ldots, \mathbf{x}_M)$$

un codice a blocco su G di lunghezza N e cardinalità M.

La probabilità di errore di $\mathcal{C}$, condizionata alla trasmissione della m–esima parola, soddisfa

$$P(e|\mathcal{C}, m) \leq \sum_{\mathbf{v}\in G^N} \tfrac{1}{|G|^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(\sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)} S^x(\boldsymbol{\theta}|\mathcal{C}, m) \tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} W_N(\mathbf{y}|\mathbf{v}\mathbf{x})^{\frac{1}{1+\rho}} \right)^{\rho}. \tag{2.7}$$

con $x \in \{r, l\}$.

Dimostrazione
Possiamo supporre senza perdita di generalità $m = 1$. Generiamo dal codice $\mathcal{C}$ l'ensemble $\mathcal{E}_{\mathcal{C}} = (\{\tilde{\mathcal{C}}\}, \mathbb{P})$ nel modo seguente:

- $\mathcal{C}' = (\mathbf{c}'_1{=}\mathbf{x}_1, \mathbf{c}'_2 = \mathbf{x}_{\Pi(2)}, \ldots, \mathbf{c}'_M = \mathbf{x}_{\Pi(M)})$, con Π uniformemente distribuita sul gruppo S_{M-1} delle permutazioni di $\{2, \ldots, M\}$;
- $\mathcal{C}'' = (\mathbf{c}''_1 = \Lambda\mathbf{c}'_1, \ldots, \mathbf{c}''_M = \Lambda\mathbf{c}'_M)$, con Λ uniformemente distribuita sul gruppo S_N delle permutazioni di $\{1, \ldots, N\}$, indipendente da Π;
- $\tilde{\mathcal{C}} = (\mathbf{c}_1, \ldots, \mathbf{c}_M) = (\mathbf{g}\mathbf{c}_1, \ldots, \mathbf{g}\mathbf{c}_M) = \mathbf{g}\mathcal{C}''$, con $\mathbf{g}$ uniformemente distribuito su G^N, indipendente da Π e Λ.

Abbiamo che

$$\begin{aligned} \mathbb{P}(\mathbf{c}_1 = \mathbf{x}) &= \sum_{\mathbf{y}\in G^N} \mathbb{P}(\mathbf{c}_1 = \mathbf{x}|\mathbf{c}''_1 = \mathbf{y})\mathbb{P}(\mathbf{c}''_1 = \mathbf{y}) \\ &= \sum_{\mathbf{y}\in G^N} \mathbb{P}(\mathbf{g}\mathbf{y} = \mathbf{x}|\mathbf{c}''_1 = \mathbf{y})\mathbb{P}(\mathbf{c}''_1 = \mathbf{y}) \\ &= \mathbb{P}(\mathbf{g} = \mathbf{y}^{-1}\mathbf{x}) \sum_{\mathbf{y}\in G^N} \mathbb{P}(\mathbf{c}''_1 = \mathbf{y}) = \frac{1}{|G|^N}. \end{aligned} \tag{2.8}$$

Inoltre, per ogni $2 \leq l \leq M$,

$$\mathbb{P}(\mathbf{c}'_l = \mathbf{c}'_1\mathbf{x}) = \frac{1}{(M-1)} \sum_{n=2}^{M} \delta_{\{\mathbf{x}_n = \mathbf{x}_1\mathbf{x}\}}.$$

Per ogni $\mathbf{x} \in G^N$ indichiamo con $Stab(\mathbf{x})$ lo stabilizzatore di $\mathbf{x}$ in S_N, i.e. il sottogruppo di S_N che lascia invariata $\mathbf{x}$; la cardinalità di $Stab(\mathbf{x})$ è pari a

$$(\boldsymbol{\theta}(\mathbf{x}))! := \prod_{g\in G} \theta_g(\mathbf{x})!.$$

Sfruttando la G–simmetria del canale otteniamo

$$\begin{aligned}
\mathbb{P}(\mathbf{c}_l = \mathbf{vx}|\mathbf{c}_1 = \mathbf{v}) &= \mathbb{P}(\mathbf{c}_1^{-1}\mathbf{c}_l = \mathbf{x}|\mathbf{c}_1 = \mathbf{v}) \\
&= \mathbb{P}((\mathbf{c}_1'')^{-1}\mathbf{g}^{-1}\mathbf{g}\mathbf{c}_l'' = \mathbf{x}|\mathbf{c}_1 = \mathbf{v}) \\
&= \mathbb{P}((\mathbf{c}_1'')^{-1}\mathbf{c}_l'' = \mathbf{x}) \\
&= \sum_{\lambda \in\ S_N} \tfrac{1}{N!}\mathbb{P}((\mathbf{c}_1')^{-1}\mathbf{c}_l' = \lambda\mathbf{x}) \\
&= \tfrac{1}{N!}\sum_{\lambda \in\ S_N} \mathbb{P}(\mathbf{c}_l' = \mathbf{x}_1\lambda\mathbf{x}) \\
&= \tfrac{1}{N!}\sum_{\mathbf{y}\in\mathcal{T}^N_{\boldsymbol{\theta}(\mathbf{x})}} \sum_{\lambda\in Stab(\mathbf{y})} \mathbb{P}(\mathbf{c}_l' = \mathbf{x}_1\mathbf{y}) \\
&= \tfrac{1}{N!}\tfrac{1}{M-1}(N\boldsymbol{\theta}(\mathbf{x}))! \sum_{\mathbf{y}\in\mathcal{T}^N_{\boldsymbol{\theta}(\mathbf{x})}} \sum_{n=2}^{M} \delta_{\{\mathbf{x}_n = \mathbf{x}_1\mathbf{y}\}} \\
&= \tfrac{1}{M-1}\tbinom{N}{N\boldsymbol{\theta}}^{-1} S^r(\boldsymbol{\theta}(\mathbf{x})|\mathcal{C}, 1)\ .
\end{aligned} \tag{2.9}$$

Il fatto che il canale considerato sia senza memoria e che si usi decodifica ML garantisce che per ogni $\mathcal{C}''$ così generato valga

$$P(e|\mathcal{C}'', 1) = P(e|\mathcal{C}, 1),$$

mentre la proprietà di G–simmetria assicura che per ogni $\tilde{\mathcal{C}}$

$$P(e|\tilde{\mathcal{C}}, 1) = P(e|\mathcal{C}'', 1),$$

e quindi

$$P(e|\mathcal{C}, 1) = \mathbb{E}_{\tilde{\mathcal{C}}}[P(e|\tilde{\mathcal{C}}, 1)].$$

Applichiamo il Gallager bound a ciascun codice $\tilde{\mathcal{C}}$, mediamo sull'ensemble $\mathcal{E}_{\mathcal{C}}$ e applichiamo la (2.8):

$$\begin{aligned}
P(e|\mathcal{C}, 1) &\leq \mathbb{E}_{\tilde{\mathcal{C}}}\left[\sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{c}_1)\left(\sum_{n=2}^{M}\left(\frac{W_N(\mathbf{y}|\mathbf{c}_n)}{W_N(\mathbf{y}|\mathbf{c}_1)}\right)^{\frac{1}{1+\rho}}\right)^{\rho}\right] \\
&= \sum_{\mathbf{v}\in G^N}\frac{1}{|G|^N}\sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}}\mathbb{E}_{\tilde{\mathcal{C}}}\left[\left(\sum_{n=2}^{M} W_N(\mathbf{y}|\mathbf{c}_n)^{\frac{1}{1+\rho}}\right)^{\rho}\middle|\ \mathbf{c}_1 = \mathbf{v}\right].
\end{aligned} \tag{2.10}$$

Il valore atteso condizionato nella precedente può essere stimato applicando la

disuguaglianza di Jensen e quindi la (2.9):

$$\begin{aligned}\mathbb{E}_{\tilde{\mathcal{C}}}\left[\left(\sum_{n=2}^{M} W_N(\mathbf{y}|\mathbf{c}_n)^{\frac{1}{1+\rho}}\right)^{\rho}\middle|\,\mathbf{c}_1=\mathbf{v}\right] &\le \left(\sum_{n=2}^{M}\mathbb{E}_{\tilde{\mathcal{C}}}\left[W_N(\mathbf{y}|\mathbf{c}_n)^{\frac{1}{1+\rho}}\middle|\,\mathbf{c}_1=\mathbf{v}\right]\right)^{\rho}\\ &= \left(\sum_{n=2}^{M}\sum_{\mathbf{x}\in G^N} W_N(\mathbf{y}|\mathbf{vx})^{\frac{1}{1+\rho}}\mathbb{P}(\mathbf{c}_n=\mathbf{vx}|\mathbf{c}_1=\mathbf{v})\right)^{\rho}\\ &= \left(\sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)}\sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}} S^r(\boldsymbol{\theta}(\mathbf{x})|\mathcal{C},1)\tbinom{N}{N\boldsymbol{\theta}(\mathbf{x})}^{-1} W_N(\mathbf{y}|\mathbf{vx})^{\frac{1}{1+\rho}}\right)^{\rho}\\ &= \left(\sum_{\boldsymbol{\theta}\in\mathcal{P}_N(G)} S^r(\boldsymbol{\theta}|\mathcal{C},1)\tbinom{N}{N\boldsymbol{\theta}}^{-1}\sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}} W_N(\mathbf{y}|\mathbf{vx})^{\frac{1}{1+\rho}}\right)^{\rho}.\end{aligned} \tag{2.11}$$

Dalla (2.10) e dalla (2.11) si ottiene la (2.7) con $x = r$. Per $x = l$ si ripete lo stesso ragionamento con un'unica modifica nella definizione dell'ensemble $\mathcal{E}_{\mathcal{C}}$: si considera $\tilde{\mathcal{C}} = \mathcal{C}\mathbf{g}$ invece che $\tilde{\mathcal{C}} = \mathbf{g}\mathcal{C}$. ■

2.4 Codici G–lineari

Introduciamo finalmente la classe dei codici a blocco G–lineari.

Definizione 5
Un codice a blocco $\mathcal{C} = (\mathbf{x}_1, \ldots, \mathbf{x}_M) \subseteq (G^N)^M$ si dice G–lineare se

- *il suo supporto* $\operatorname{supp}\mathcal{C} = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$ *è un sottogruppo di G^N;*
- *la cardinalità dell'insieme $\{i$ t.c. $\mathbf{x}_i = \mathbf{x}\}$ è costante al variare di $\mathbf{x}$ in* $\operatorname{supp}\mathcal{C}$.

Dalla definizione qui sopra segue immediatamente che ogni codice G–lineare contiene almeno una parola uguale all'elemento neutro $\mathbf{g}_0$ di G^N. Si può sempre supporre quindi che $\mathbf{x}_1 = \mathbf{g}_0$, cosa che faremo sempre d'ora in poi quando non diversamente specificato. Quindi, se $\mathcal{C}$ è G–lineare, gli spettri $S^r(\boldsymbol{\theta}|\mathcal{C}, m)$ e $S^l(\boldsymbol{\theta}|\mathcal{C}, m)$ di distanze da una parola $\mathbf{x}_m \in \mathcal{C}$, non dipendono da m. Infatti

$$S^r(\boldsymbol{\theta}|\mathcal{C}, m) = \sum_{j\neq m}\delta_{\{\mathbf{x}_m^{-1}\mathbf{x}_j\in\mathcal{T}_{\boldsymbol{\theta}}^N\}} = \sum_{k\neq 1}\delta_{\{\mathbf{x}_k\in\mathcal{T}_{\boldsymbol{\theta}}^N\}} = S^r(\boldsymbol{\theta}|\mathcal{C}, 1).$$

Anche la probabilità di errore dei codici G–lineari gode di una proprietà analoga, nota come *Uniform Error Property*. Tale proprietà verrà largamente usata nei capitoli successivi. La Uniform Error Property è essenziale per rendere possibile l'analisi teorica delle prestazioni dei codici su gruppi, analisi che risulterebbe altrimenti impraticabile: è questo il motivo che porta a studiare i codici G–lineari.

Proprietà 9 (Uniform Error Property)
La probabilità di errore di un codice G–lineare $\mathcal{C}$ su un canale G–simmetrico con decodifica ML è indipendente dalla parola trasmessa, i.e.

$$P(e|\mathcal{C}) = P(e|\mathcal{C}, m), \qquad \forall m = 1, \ldots, M \ .$$

Dimostrazione
Se $\mathcal{C}$ è degenere il risultato segue immediatamente dal fatto che

$$P(e|\mathcal{C}, m) = 1 \ , \qquad m = 1, \ldots, M \ .$$

Supponiamo ora che $\mathcal{C}$ non sia degenere. Allora

$$\begin{aligned} P(e|\mathcal{C}, m) &= 1 - \sum_{\mathbf{y} \in \mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_m) \delta_{\{W_N(\mathbf{y}|\mathbf{x}_m) > W_N(\mathbf{y}|\mathbf{x}_j), \forall j \neq m\}} \\ &= 1 - \sum_{\mathbf{y} \in \mathcal{Y}^N} W_N(\mathbf{x}_m^{-1}\mathbf{y}|\mathbf{g}_0) \delta_{\{W_N(\mathbf{x}_m^{-1}\mathbf{y}|\mathbf{g}_0) > W_N(\mathbf{x}_m^{-1}\mathbf{y}|\mathbf{x}_m^{-1}\mathbf{x}_j), \forall j \neq m\}} \\ &= 1 - \sum_{\mathbf{y} \in \mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{x}_1) \delta_{\{W_N(\mathbf{y}|\mathbf{x}_1) > W_N(\mathbf{y}|\mathbf{x}_i), \forall i \neq 1\}} \\ &= P(e|\mathcal{C}, 1). \end{aligned}$$

■

Chiudiamo il capitolo mostrando che nel caso particolare in cui S è la costellazione 2–PAM introdotta nell'Esempio 1 con gruppo generatore $G \simeq \mathbb{Z}_2$ e $\mathcal{C}$ è un codice a blocco $\mathbb{Z}_2$–lineare i due lemmi del paragrafo precedente forniscono le note stime per la probabilità di errore dei codici binari lineari. Per tali codici usiamo la notazione

$$S(w|\mathcal{C}) := S\left(\left(1-\tfrac{w}{N}, \tfrac{w}{N}\right)|\mathcal{C}\right)$$

per indicare il numero di parole di $\mathcal{C}$ di peso di Hamming w. Consideriamo i due casi di decodifica hard e di decodifica soft. Iniziamo dalla decodifica hard: il canale gaussiano additivo di tipo (S, S) è in questo caso un canale binario simmetrico di probabilità di transizione

$$p = \frac{1}{2} \operatorname{erfc}\left(\frac{L}{\sqrt{N_0}}\right) \ .$$

Per tale canale il parametro di Battacharyya è

$$D = 2\sqrt{p(1-p)} \ \ ;$$

la (2.5) dunque diventa

$$P(e|\mathcal{C}) \leq \sum_{w=1}^{N} S(w|\mathcal{C}) [2\sqrt{p(1-p)}]^w \ . \tag{2.12}$$

Passiamo alla decodifica soft. Il canale gaussiano additivo di tipo S ha in questo caso parametro di Battacharyya

$$D = \frac{1}{\sqrt{\pi N_0}} \int_{\mathbb{R}} e^{\frac{1}{2}\left(-\frac{(y-L)^2}{N_0} - \frac{(y+L)^2}{N_0}\right)} = e^{-\frac{L^2}{N_0}} \ ;$$

la (2.5) dunque diventa

$$P(e|\mathcal{C}) \leq \sum_{w=1}^{N} S(w|\mathcal{C})[\exp(-L^2/N_0)]^w \ . \tag{2.13}$$

Capitolo 3

Ensemble classici di codici $\mathbb{Z}_m$–lineari su canali $\mathbb{Z}_m$–simmetrici

I canali $\mathbb{Z}_m$–simmetrici saranno oggetto di questo capitolo. Le prestazioni dell'ensemble dei codici a blocco su $\mathbb{Z}_m$ con decodifica a massima verosimiglianza possono essere analizzate con gli strumenti della teoria classica di Shannon introdotti nel primo capitolo, senza fare intervenire la struttura algebrica di $\mathbb{Z}_m$. In particolare si mostra che la probabilità media di errore del random coding ensemble $\mathcal{E}_{RC}$ ha andamento esponenzialmente decrescente a zero per ogni rate $R < C_m$ con esponente pari a $-NE_m(R)$, dove C_m ed E_m sono rispettivamente la capacità e l'esponente di errore del canale $\mathbb{Z}_m$–simmetrico considerato.

Ora verranno invece analizzate le prestazioni dei codici $\mathbb{Z}_m$–lineari su canali $\mathbb{Z}_m$–simmetrici; in particolare sarà studiato il comportamento di due ensemble classici di codici $\mathbb{Z}_m$–lineari. Tale analisi si baserà sulla struttura algebrica di $\mathbb{Z}_m$, e in particolare sul fatto che $\mathbb{Z}_m$ è un anello ad ideali principali.

Verrà introdotto il concetto di *sottocanale* di un canale $\mathbb{Z}_m$–simmetrico e si definiranno due nuove quantità, la $\mathbb{Z}_m$–capacità $\hat{C}_m$ e il $\mathbb{Z}_m$–esponente di errore $\hat{E}_m(R)$. Si mostrerà poi come tali quantità, che sono minori o uguali di C_m ed $E_m(R)$ rispettivamente, caratterizzano il comportamento degli ensemble di codici $\mathbb{Z}_m$–lineari.

Infine, ci si concentrerà sulla $\mathbb{Z}_m$–capacità e verranno studiate dettagliatamente due famiglie di canali gaussiani additivi di tipo (S, S'), con S costellazione geometricamente uniforme di gruppo generatore $\mathbb{Z}_m$; per una di queste si vedrà che $\hat{C}_m < C_m$, cioè la restrizione ai codici $\mathbb{Z}_m$–lineari provoca una effettiva perdita di capacità. Per la famiglia di costellazioni 2^r–PSK si

mostrerà invece che $\hat{C}_m = C_m$ e quindi l'uso di codici $\mathbb{Z}_{2^r}$–lineari congiuntamente a tali costellazioni non provoca perdita di capacità. Si tratta di un risultato originale e di notevole interesse applicativo, dal momento che tali costellazioni sono tra le più usate nella pratica.

3.1 Il teorema inverso di codifica per codici $\mathbb{Z}_m$–liberi su canali $\mathbb{Z}_m$–simmetrici

Siano K e N due interi positivi tali che $K \leq N$. Sia ϕ un omomorfismo di $\mathbb{Z}_m^K$ in $\mathbb{Z}_m^N$. A partire da ϕ definiamo il codice $\mathbb{Z}_m$–lineare

$$\mathcal{C}_\phi := (\mathbf{x_u} = \phi\mathbf{u})_{\mathbf{u}\in\mathbb{Z}_m^K} \in \left(\mathbb{Z}_m^N\right)^{m^K}. \tag{3.1}$$

Si osservi che si è scelto per comodità di tenere come insieme degli indici $\mathbb{Z}_m^K$ invece che $\{1, \ldots, m^K\}$. Il rate di $\mathcal{C}$ è pari a

$$R = \frac{1}{N}\log(m^K) = \frac{K}{N}\log m.$$

Si noti che se m è un numero primo, si ottengono, al variare di ϕ tra gli omomorfismi iniettivi di $\mathbb{Z}_m^K$ in $\mathbb{Z}_m^N$, tutti i codici lineari non degeneri con supporto di dimensione K; se ϕ non è iniettivo, si ottengono invece codici con supporto di dimensione più piccola ma tutti degeneri. Nel caso in cui m non è primo la situazione è analoga: al variare di ϕ tra gli omomorfismi iniettivi si ottengono tutti i codici lineari non degeneri il cui supporto è uno $\mathbb{Z}_m$-modulo libero di dimensione K e come prima si ottengono, se ϕ non è iniettiva, codici con supporto più piccolo ma tutti degeneri.

Fissato $\phi \in \mathrm{Hom}(\mathbb{Z}_m^K, \mathbb{Z}_m^N)$, per ogni $l \mid m$ consideriamo l'omomorfismo $\phi_l : \frac{m}{l}\mathbb{Z}_m^K \to \mathbb{Z}_m^N$ che si ottiene restringendo il dominio di ϕ a $\frac{m}{l}\mathbb{Z}_m^K$: l'immagine di ϕ_l è un sottogruppo di $\frac{m}{l}\mathbb{Z}_m^N$, dunque si può pensare ϕ_l come elemento di $\mathrm{Hom}(\frac{m}{l}\mathbb{Z}_m^K, \frac{m}{l}\mathbb{Z}_m^N)$. Consideriamo il codice $\mathcal{C}_{\phi_l}$ su $\frac{m}{l}\mathbb{Z}_m$:

$$\mathcal{C}_{\phi_l} = (\mathbf{x_u} = \phi\mathbf{u})_{\mathbf{u}\in\frac{m}{l}Z_m^K} \in \left(\frac{m}{l}\mathbb{Z}_m^N\right)^{l^K}. \tag{3.2}$$

Il rate di $\mathcal{C}_{\phi_l}$ è

$$R_l = \frac{1}{N}\log(l^K) = \frac{K}{N}\log l = \frac{\log l}{\log m}R.$$

Consideriamo ora, per ogni $l \mid m$, con $l > 1$, il sottocanale 'l–esimo'

$$\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\frac{m}{l}\mathbb{Z}_m}$$

ottenuto restringendo l'insieme degli ingressi da $\mathbb{Z}_m$ al suo sottogruppo $\frac{m}{l}\mathbb{Z}_m$. Si verifica immediatamente che ciascuno di tali sottocanali è $\frac{m}{l}\mathbb{Z}_m$–simmetrico. Indichiamo con C_l la capacità dell'l–esimo sottocanale. Il teorema inverso di codifica di canale ci permette di concludere che condizione necessaria perchè la probabilità di errore di $\mathcal{C}_{\phi_l}$ sull'l–esimo sottocanale possa essere resa arbitrariamente piccola è che

$$R_l = \frac{\log l}{\log m} R < C_l \ .$$

La probabilità di errore del codice iniziale $\mathcal{C}_\phi$ è maggiore o uguale a quella di ciascun suo sottocodice, quindi

$$P(e|\mathcal{C}) \geq P(e|\mathcal{C}^{(l)}) \ , \quad \forall l > 1 \text{ t.c. } l \mid m \ .$$

Possiamo dunque concludere che condizione necessaria perchè la probabilità di errore di $\mathcal{C}$ possa essere resa arbitrariamente piccola è che

$$R < \min_{\substack{l>1: \\ l|m}} \frac{\log m}{\log l} C_l \ . \tag{3.3}$$

La (3.3) motiva la definizione seguente.

Definizione 6
Dato un canale $\mathbb{Z}_m$-simmetrico $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathbb{Z}_m}$ la sua $\mathbb{Z}_m$-capacità *è*

$$\hat{C}_m := \min_{\substack{l>1: \\ l|m}} \frac{\log m}{\log l} C_l \ .$$

Si osservi che evidentemente $\hat{C}_m \leq C_m$. Possiamo riassumere le conclusioni di questo paragrafo nell'enunciato seguente.

Teorema 10
Sia dato un canale $\mathbb{Z}_m$–simmetrico $W \in \mathcal{P}(\mathcal{Y})$ di $\mathbb{Z}_m$–capacità $\hat{C}_m$. Allora, per ogni $R > \hat{C}_m$ esiste $K > 0$ tale che ogni codice a blocco $\mathcal{C}_\phi$ di tipo (3.2) di rate R con un qualsiasi decodificatore abbia probabilità di errore

$$P(e|\mathcal{C}) \geq K \ .$$

Sottolineiamo il fatto che il teorema precedente riguarda tutti i codici $\mathbb{Z}_m$–lineari che possono essere definiti come nella forma (3.2): come abbiamo osservato prima, questi non sono tutti i possibili codici $\mathbb{Z}_m$–lineari.

Facciamo notare inoltre come nel caso in cui m sia un numero primo non ci siano sottocanali da considerare e che dunque

$$\hat{C}_m = C_m \ .$$

Se invece m non è primo, a priori $\hat{C}_m$ potrebbe essere un numero strettamente inferiore alla capacità di Shannon C_m (ritorneremo su questo problema nel paragrafo 3.3). Nel prossimo paragrafo dimostreremo invece alcuni risultati che mostreranno una sorta di controparte diretta del risultato precedente e cioè che effettivamente, utilizzando codici $\mathbb{Z}_m$–lineari, si può arrivare a trasmettere a qualunque rate R inferiore a $\hat{C}_m$.

3.2 Ensemble di codici immagine di omomorfismi di $\mathbb{Z}_m$–moduli

In questo capitolo verrà dimostrata la parte diretta del teorema di Shannon per codici $\mathbb{Z}_m$–lineari su un canale $\mathbb{Z}_m$–simmetrico.

Iniziamo ad introdurre delle notazioni. Si consideri un canale $\mathbb{Z}_m$–simmetrico fissato $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x\in\mathbb{Z}_m}$. Per ogni $l \mid m$, $l > 1$, definiamo l'esponente di errore dell'l–esimo sottocanale come

$$E_l(R) := \max_{0\le\rho\le 1} \left\{ \max_{p\in\mathcal{P}(\frac{m}{l}\mathbb{Z}_m)} E_0(\rho, p) - \rho R \right\} = \max_{0\le\rho\le 1} \left\{ E_0(\rho, u_{\frac{m}{l}\mathbb{Z}_m}) - \rho R \right\} .$$

Definiamo poi il $\mathbb{Z}_m$–esponente di errore del canale nella maniera seguente:

$$\hat{E}_m(R) := \min_{\substack{l>1: \\ l|m}} \left\{ E_l\left(\frac{\log l}{\log m} R\right) \right\} \ , \qquad R \in [0, \log m].$$

Dal Teorema 3 segue immediatamente che $\hat{E}_m(R) > 0$ se e solo se $R < \hat{C}_m$.

Fissiamo ora tre interi positivi L, K, N tali che

$$N = K + L \ .$$

In questo e nel prossimo ricaviamo delle stime della probabilità media di errore di due ensemble classici di codici $\mathbb{Z}_m$–lineari: quello dei codici immagine di omomorfismi

$$\phi : \mathbb{Z}_m^K \to \mathbb{Z}_m^N$$

e quello dei codici nucleo di omomorfismi

$$\phi : \mathbb{Z}_m^N \to \mathbb{Z}_m^L \ .$$

Iniziamo con l'ensemble dei codici immagine. Fissati arbitrariamente $N \in \mathbb{N}$ e $R \in [0 \log m]$, poniamo

$$K := \left\lceil \tfrac{R}{\log m} N \right\rceil .$$

A partire da ogni ϕ in $\mathrm{Hom}(\mathbb{Z}_m^K, \mathbb{Z}_m^N)$ definiamo come nel paragrafo precedente il codice $\mathbb{Z}_m$–lineare

$$\mathcal{C}_\phi := \left(\mathbf{x_u} = \phi\mathbf{u} \in \mathbb{Z}_m^N\right)_{\mathbf{u} \in \mathbb{Z}_m^K} .$$

Sia ora Φ una variabile aleatoria uniformemente distribuita su $\mathrm{Hom}(\mathbb{Z}_m^K, \mathbb{Z}_m^N)$. Φ induce naturalmente una struttura probabilistica sull'insieme $(\mathbb{Z}_m^N)^{m^K}$: indichiamo questo ensemble con

$$\mathcal{E}_{\,\mathrm{Im}}\,(N, R) \ .$$

Si noti che vi è una corrispondenza biunivoca tra le $\phi \in \mathrm{Hom}(\mathbb{Z}_m^K, \mathbb{Z}_m^N)$ iniettive e i codici non degeneri dell'ensemble $\mathcal{E}_{\,\mathrm{Im}}\,(N, R)$, mentre ai codici degeneri corrispondono più di una possibile $\phi \in \mathrm{Hom}(\mathbb{Z}_m^K, \mathbb{Z}_m^N)$. Ovviamente questi codici degeneri non sono di interesse e possono solo peggiorare le performance dell'ensemble. Tuttavia, come era accaduto per il random coding ensemble, è molto più semplice studiare l'ensemble tutto intero che restringerci a priori agli omomorfismi ϕ iniettivi.

Teorema 11
Per ogni $N \in \mathbb{N}$, $R \in [0, \log m]$, la probabilità media di errore dell'ensemble $\mathcal{E}_{Im}(N, R)$ con decodifica ML su un canale $\mathbb{Z}_m$–simmetrico soddisfa

$$\overline{P(e)} \leq \sum_{\substack{l|m \\ l>1}} \exp(-N E_l(R_l)) \tag{3.4}$$

dove, per ogni $l \,|\, m$, $l>1$, $E_l(R)$ è l'esponente di errore dell'l–esimo sottocanale e $R_l := \frac{K}{N} \log l$ il suo rate di utilizzo.

Dimostrazione
La linearità dei codici dell'ensemble considerato permette di applicare la uniform error property a ciascun $\mathcal{C}_\phi$, e quindi si ha

$$P(e|\mathcal{C}_\Phi) = P(e|\mathcal{C}_\Phi, \mathbf{0}).$$

Consideriamo la partizione di $\mathbb{Z}_m^K$ in sottoinsiemi di elementi dello stesso ordine

$$\begin{aligned} \mathbb{Z}_m^K &= \bigcup_{l|m} H_{K,l}, \\ H_{K,l} \ &:= \ \{\mathbf{u} \in \mathbb{Z}_m^K \text{ t.c. } \mathrm{MCD}(u_j, m) = \tfrac{m}{l}\} \subsetneq \tfrac{m}{l}\mathbb{Z}_m^K \ . \end{aligned} \tag{3.5}$$

Una stima union bound permette di scrivere, per ogni $\phi \in Hom(\mathbb{Z}_m^K, \mathbb{Z}_m^N)$

$$P(e|\mathcal{C}_\phi, \mathbf{0}) \leq \sum_{\substack{l|m \\ l>1}} P(e|\mathcal{C}_\phi^{(l)}, \mathbf{0}) \tag{3.6}$$

dove $\mathcal{C}_\phi^{(l)}$ è il codice ottenuto restringendo il dominio di ϕ da tutto $\mathbb{Z}_m^K$ all'insieme

$$\{\mathbf{0}\} \cup H_{K,l} \ .$$

Applichiamo il Lemma 8 a ciascun codice $\mathcal{C}_\phi^{(l)}$ (si osservi che tali codici non sono lineari a differenza di $\mathcal{C}_\phi$), e mediamo su Φ:

$$\begin{aligned}
&\overline{P(e|\mathcal{C}_\Phi^{(l)}, \mathbf{0})} \leq \\
&\leq \mathbb{E}_\Phi\left[\sum_{\mathbf{v}\in\mathbb{Z}_m^N} \tfrac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(\sum_{\boldsymbol{\theta}\in\mathcal{P}_N(\mathbb{Z}_m)} S(\boldsymbol{\theta}|\mathcal{C}_\Phi^{(l)}, \mathbf{0}) \tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}}\right)^\rho\right] \\
&\leq \sum_{\mathbf{v}\in G^N} \tfrac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(\sum_{\boldsymbol{\theta}\in\mathcal{P}_N(\mathbb{Z}_m)} \overline{S(\boldsymbol{\theta}|\mathcal{C}_\Phi^{(l)}, \mathbf{0})} \tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}}\right)^\rho
\end{aligned} \tag{3.7}$$

Direttamente dalla definizione di spettro di distanze abbiamo

$$S_{\mathbf{0}}(\boldsymbol{\theta}|\mathcal{C}_\phi^{(l)}) = \sum_{\mathbf{u}\in H_{K,l}} \delta_{\{\phi\mathbf{u}\in\mathcal{T}_{\boldsymbol{\theta}}^N\}}.$$

Inoltre, per ogni $\mathbf{u} \in H_{K,l}$, $\Phi\mathbf{u}$ è distribuita uniformemente su $\frac{m}{l}\mathbb{Z}_m^N$ (per un'idea della dimostrazione si veda l'Appendice), quindi

$$\overline{S_{\mathbf{0}}(\boldsymbol{\theta}|\mathcal{C}_\Phi^{(l)})} = \mathbb{E}_\Phi\Big[\sum_{\mathbf{u}\in H_{K,l}} \mathbb{1}_{\{\Phi\mathbf{u}\in\mathcal{T}_{\boldsymbol{\theta}}^N\}}\Big] = \sum_{\mathbf{u}\in H_{K,l}} \mathbb{P}(\Phi\mathbf{u}\in\mathcal{T}_{\boldsymbol{\theta}}^N) = \left(\frac{1}{l}\right)^N \tbinom{N}{N\boldsymbol{\theta}} \delta_{\{\boldsymbol{\theta}\in\mathcal{P}_N(\frac{m}{l}\mathbb{Z}_m)\}} \ . \tag{3.8}$$

Fissiamo ora un insieme $\Omega_l \subseteq \mathbb{Z}_m^N$ di cardinalità $(\frac{m}{l})^N$ contenente un elemento per ciascuna classe laterale di $\frac{m}{l}\mathbb{Z}_m^N$. Sostituendo la (3.8) nella (3.7) si ottiene

$$\begin{aligned}
&\overline{P(e|\mathcal{C}^{(l)}, \mathbf{0})} \leq \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \tfrac{1}{m^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(|H_{K,l}| \left(\tfrac{1}{l}\right)^N \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}}\right)^\rho \\
&= |H_{K,l}|^\rho \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{v}\in\Omega_l} \left(\tfrac{l}{m}\right)^N \sum_{\mathbf{w}\in\frac{m}{l}\mathbb{Z}_m^N} \left(\tfrac{1}{l}\right)^N W_N(\mathbf{y}|\mathbf{v}+\mathbf{w})^{\frac{1}{1+\rho}} \left(\left(\tfrac{1}{l}\right)^N \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}}\right)^\rho \\
&= |H_{K,l}|^\rho \sum_{\mathbf{v}\in\Omega_l} \left(\tfrac{l}{m}\right)^N \sum_{\mathbf{y}\in\mathcal{Y}^N} \left(\left(\tfrac{1}{l}\right)^N \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}}\right)^{1+\rho} .
\end{aligned} \tag{3.9}$$

Poiché il canale è $\frac{m}{l}\mathbb{Z}_m$–simmetrico, si ha che, per ogni $\mathbf{v} \in \Omega_l$,

$$\begin{aligned}
&\sum_{\mathbf{y}\in\mathcal{Y}^N} \Big(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \left(\tfrac{1}{l}\right)^N W_N(\mathbf{y}|\mathbf{x}+\mathbf{v})^{\frac{1}{1+\rho}} \Big)^{1+\rho} \\
&= \sum_{\mathbf{y}\in\mathcal{Y}^N} \Big(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \left(\tfrac{1}{l}\right)^N W_N((-\mathbf{v})\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}} \Big)^{1+\rho} \\
&= \sum_{\mathbf{y}\in\mathcal{Y}^N} \Big(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \left(\tfrac{1}{l}\right)^N W_N(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}} \Big)^{1+\rho}.
\end{aligned}$$

Inoltre, dall'inclusione $H_{K,l} \subseteq \frac{m}{l}\mathbb{Z}_m^N$, segue

$$|H_{K,l}| \leq l^K.$$

Abbiamo quindi

$$\begin{aligned}
\overline{P(e|\mathcal{C}^{(l)},\mathbf{0})} &\leq l^{K\rho} \sum_{\mathbf{y}\in\mathcal{Y}^N} \left(\left(\tfrac{1}{l}\right)^N \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} W_N(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}} \right)^{1+\rho} \\
&= l^{K\rho} \left[\sum_{y\in\mathcal{Y}} \left(\sum_{x\in\frac{m}{l}\mathbb{Z}_m} \tfrac{1}{l} W_N(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} \right]^N \\
&= \exp\left(-N \left[E_0(\mathbf{u}_{\frac{m}{l}\mathbb{Z}_m}, \rho) - \rho R_l \right] \right).
\end{aligned} \tag{3.10}$$

Per l'arbitrarietà di $\rho \in [0,1]$, e poichè, essendo il sottocanale l–esimo è $\frac{m}{l}\mathbb{Z}_m$–simmetrico, $E_l(R)$ si ottiene con distribuzione uniforme su $\frac{m}{l}\mathbb{Z}_m$, si ha infine

$$\overline{P(e|\mathcal{C}^{(l)},\mathbf{0})} \leq \exp(-NE_l(R_l)); \tag{3.11}$$

sostituendo nella (3.6) si ottiene la (3.12). ■

È chiaro, per le considerazioni svolte precedentemente, che la stima (3.12) continua a valere considerando l'ensemble dei codici a blocco non degeneri il cui supporto sia un $\mathbb{Z}_m$–modulo libero di dimensione esattamente K, con la probabilità uniforme.

3.3 Ensemble di codici nucleo di omomorfismi di $\mathbb{Z}_m$–moduli

Passiamo ora all'ensemble dei codici definiti come nucleo di omomorfismi. A partire da ogni $\phi \in \mathrm{Hom}(\mathbb{Z}_m^N, \mathbb{Z}_m^L)$ definiamo il codice $\mathbb{Z}_m$–lineare

$$\mathcal{C}_{\mathrm{Ker}\,\phi} = (\mathbf{x}_1, \ldots, \mathbf{x}_M)$$

come un qualsiasi ordinamento di $\operatorname{Ker}\phi$ tale che $\mathbf{x}_1 = \mathbf{0}$. È ovviamente sempre possibile definire un tale ordinamento dal momento che $\mathbf{0} \in \operatorname{Ker}\phi$ per ogni omomorfismo ϕ. Si noti che per costruzione $\mathcal{C}_{\operatorname{Ker}\phi}$ è sicuramente non degenere.

Fissati $R \in [0, \log m]$ e $N \in \mathbb{N}$, poniamo

$$L := \left\lfloor \left(1 - \frac{R}{\log m}\right) N \right\rfloor \ .$$

Sia ora Φ una variabile aleatoria distribuita uniformemente su $\operatorname{Hom}(\mathbb{Z}_m^N, \mathbb{Z}_m^L)$. Φ induce una struttura probabilistica sull'insieme $\bigcup_{M=m^{N-L}}^{m^N} (\mathbb{Z}_m^N)^M$ dei codici a blocco su $\mathbb{Z}_m$ di lunghezza N e rate maggiore o uguale a R. Indichiamo con

$$\mathcal{E}_{\operatorname{Ker}}(N, R)$$

questo ensemble.

Si osservi che i codici di questo ensemble hanno rate maggiore o uguale a $R_0 = \frac{N-L}{N} \log m$: in effetti si ha che il rate è uguale a R_0 se e soltanto se $\phi \in \operatorname{Hom}(\mathbb{Z}_m^N, \mathbb{Z}_m^L)$ è suriettiva. Si noti che, se m è un numero primo, questo ensemble contiene tutti i possibili codici lineari (prendendo tutte le possibili permutazioni) di dimensione maggiore o uguale a $K = N - L$. Nel caso di m non primo la caratterizzazione dei codici che si ottengono è più complessa: se ϕ è suriettiva allora $\operatorname{Ker}\phi$ è uno $\mathbb{Z}_m$–modulo libero di rango esattamente K, ed in effetti tutti gli $\mathbb{Z}_m$–modulo liberi si possono ottenere in questo modo. Invece, non si può dare una caratterizzazione intrinseca dei codici che si ottengono con $\phi \in \operatorname{Hom}(\mathbb{Z}_m^N, \mathbb{Z}_m^L)$ non suriettive, se non dire che sono codici di rate strettamente maggiore di R. Si osservi che ad uno stesso codice corrispondono diverse mappe ϕ; si può dimostrare tuttavia che a tutti i codici $\mathbb{Z}_m$–liberi di rango K corrisponde sempre uno stesso numero di $\phi \in \operatorname{Hom}(\mathbb{Z}_m^N, \mathbb{Z}_m^L)$, così che questi codici pesano in egual misura all'interno dell'ensemble. Inoltre si potrebbe far vedere che l'insieme delle ϕ non suriettive ha una probabilità che va a 0 per $N \to +\infty$, esponenzialmente in N. Tutto questo suggerisce che questo ensemble prestazioni medie molto simili a quelle dei codici immagine di omomorfismi. Così è, come si mostra direttamente nel risultato seguente. Il motivo che induce a studiarlo è che da esso nascono gli ensemble dei codici a bassa densità, che introdurremo più avanti in questa tesi.

Teorema 12
Per ogni $N \in \mathbb{N}$, $R \in [0, \log m]$ la probabilità media di errore dell'ensemble $\mathcal{E}_{\operatorname{Ker}}(N, R)$ con decodifica ML su un canale $\mathbb{Z}_m$–simmetrico soddisfa

$$\overline{P(e)} \leq \sum_{\substack{l|m \\ l>1}} \exp(-N E_l(R_l)) \ , \tag{3.12}$$

dove, per ogni $l|m$, $l>1$, $E_l(R)$ è l'esponente di errore dell'l–esimo sottocanale e $R_l = \frac{N-L}{N}\log l$ il suo rate di utilizzo.

Dimostrazione
La uniform error property permette di scrivere, per ogni codice $\mathcal{C}_{\mathrm{Ker}\,\phi}$ dell'ensemble,

$$P(e|\mathcal{C}_{\mathrm{Ker}\,\phi}) = P(e|\mathcal{C}_{\mathrm{Ker}\,\phi}, 1).$$

Partizioniamo $\mathbb{Z}_m^N$ in classi di elementi dello stesso ordine:

$$\mathbb{Z}_m^N = \bigcup_{l|m} H_{N,l}$$
$$H_{N,l} := \{\mathbf{x} \in \mathbb{Z}_m^N \text{ t.c. } \mathrm{MCD}(x_j, m) = \tfrac{m}{l}\} \subsetneq \tfrac{m}{l}\mathbb{Z}_m^N \ .$$

Sia, per ogni $l \mid m$, $l > 1$, $\mathcal{C}^{(l)}_{\mathrm{Ker}\,\phi}$ un codice di supporto $\{\mathbf{0}\} \cup (\mathrm{Ker}\,\phi \cap H_{N,l})$ tale che $\mathbf{x}_1 = \mathbf{0}$.

Sia ora $\mathbf{x} \in H_{N,l}$ una N–upla fissata. Allora $\Phi\mathbf{x}$ è una variabile aleatoria di distribuzione uniforme su $\frac{m}{l}\mathbb{Z}_m^N$ (si veda l'Appendice), e quindi in particolare

$$\mathbb{P}(\Phi\mathbf{x} = \mathbf{0}) = \left(\frac{1}{l}\right)^L \ . \tag{3.13}$$

Definiamo $J_l^N \subset \mathcal{P}_N(\mathbb{Z}_m)$ come l'insieme dei tipi delle parole di $H_{N,l}$. Dalla (3.13) segue che, per ogni $\boldsymbol{\theta} \in J_l^N$, si ha

$$\overline{S(\boldsymbol{\theta}|1)} = \binom{N}{N\boldsymbol{\theta}} \left(\frac{1}{l}\right)^L$$

e quindi, per ogni $\boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m)$,

$$\overline{S\big(\boldsymbol{\theta}|\mathcal{C}^{(l)}_{\ker\Phi}, 1\big)} = \binom{N}{N\boldsymbol{\theta}} \left(\frac{1}{l}\right)^L \delta_{\{\boldsymbol{\theta}\in J_l^N\}}.$$

Applichiamo il Lemma 8 a ciascun codice, mediamo sull'ensemble, e usiamo Jensen

$$\begin{aligned}
&\overline{P(e|\mathcal{C}^{(l)}, 1)} \leq \\
&\leq \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \tfrac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(\sum_{\boldsymbol{\theta}\in\mathcal{P}_N(\mathbb{Z}_m)} \overline{S(\boldsymbol{\theta}|\mathcal{C}^{(l)}_{\ker\Phi}, 1)} \tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}} \right)^\rho \\
&\leq \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \tfrac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(\big(\tfrac{1}{l}\big)^L \sum_{\boldsymbol{\theta}\in J_l^N} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}} \right)^\rho \\
&= l^{K\rho} \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \tfrac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(\big(\tfrac{1}{l}\big)^N \sum_{\mathbf{x}\in H_{l,N}} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}} \right)^\rho \\
&\leq l^{K\rho} \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \tfrac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \left(\big(\tfrac{1}{l}\big)^N \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} W_N(\mathbf{y}|\mathbf{v}+\mathbf{x})^{\frac{1}{1+\rho}} \right)^\rho .
\end{aligned} \tag{3.14}$$

Di qui in poi tutto segue come nella dimostrazione del Teorema 11. ■

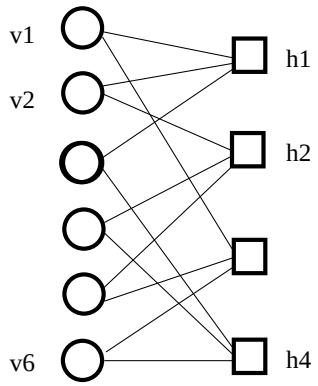


Figura 3.1: Un grafo di Tanner

Introduciamo ora il *grafo di Tanner* di un codice $\mathcal{C}_{\mathrm{Ker}\,\phi}$. È un fatto noto che il gruppo $\mathrm{Hom}(\mathbb{Z}_m^N, \mathbb{Z}_m^L)$ degli omomorfismi di $\mathbb{Z}_m$–moduli liberi è isomorfo a gruppo $\mathbb{Z}_m^{L\times N}$ delle matrici a valori in $\mathbb{Z}_m$ di dimensioni L per N. Sia dunque $H \in \mathbb{Z}_m^{L\times N}$ la matrice corrispondente a $\mathrm{Ker}\,\phi$; H è detta la *matrice di parità* del codice $\mathcal{C}_{\mathrm{Ker}\,\phi}$. Associamo ad H il grafo bipartito

$$\mathcal{G} = (\mathcal{N} \cup \mathcal{M}, E)$$

ed il labeling

$$c : E \to \mathbb{Z}_m$$

definiti nel modo seguente. Siano $\mathcal{N} = \{v_1, \ldots, v_N\}$, $\mathcal{M} = \{h_1, \ldots, h_L\}$. ogni coppia $(v_n, h_m) \in \mathcal{N} \times \mathcal{M}$ appartiene ad E se e solo $H_{m,n} \neq 0$: poniamo inoltre

$$c((v_n, h_m)) := H_{m,n} \ .$$

Viceversa, sia $\mathcal{C} = (\mathcal{N} \cup \mathcal{M}, E)$, $h : E \to \mathbb{Z}_m$, un grafo bipartito etichettato. Associamo a tale grafo l'omomorfismo $\phi \in \mathrm{Hom}(\mathbb{Z}_m^N, \mathbb{Z}_m^L)$ definito da

$$(\Phi(\mathbf{x}))_m = \sum_{(v_n,h_m)\in E} x_n c((v_n, h_m)) \ , \qquad i = 1, \ldots, L \ .$$

In seguito chiameremo $\mathcal{N}$ insieme dei variable nodes, e $\mathcal{M}$ insieme dei check nodes. Ciascun variable node corrisponde ad un simbolo della parola di codice, ciascun check node ad uno vincoli sui simboli della parola che definiscono il codice. Ogni vincolo $h_m \in \mathcal{M}$ coinvolge un certo insieme di variable nodes che indichiamo con $\mathcal{N}(m)$ ed è soddisfatto se la somma dei simboli coinvolti si annulla. A sua volta ogni variable node $v_n \in \mathcal{N}$ partecipa ad un certo insieme di check, che indichiamo con $\mathcal{M}(n)$. Il grafo di Tanner ci servirà a definire i codici a bassa densità nel prossimo capitolo.

3.4 $\mathbb{Z}_m$–capacità di un canale $\mathbb{Z}_m$–simmetrico

Nei paragrafi precedenti sono stati introdotti la $\mathbb{Z}_m$–capacità $\hat{C}_m$ e il $\mathbb{Z}_m$–esponente $\hat{E}_m(R)$ di un canale $\mathbb{Z}_m$–simmetrico. È stato mostrato inoltre che $\hat{C}_m$ è l'effettiva soglia per la comunicazione affidabile. Infatti, come conseguenza diretta dei teoremi 10, 11, e 12, si ha il seguente enunciato.

Corollario 13
Sia dato un canale senza memoria $\mathbb{Z}_m$–discreto di $\mathbb{Z}_m$–capacità $\hat{C}_m$. Allora

- *per ogni $R < \hat{C}_m$, per ogni $\varepsilon > 0$, esiste un codice a blocco $\mathbb{Z}_m$–lineare $\mathcal{C}$ di rate maggiore o uguale a R la cui probabilità di errore con decodifica ML soddisfa*
$$P(e|\mathcal{C}) < \varepsilon \ .$$
- *per ogni $R \geq \hat{C}_m$, esiste $K > 0$ indipendente da N tale che ogni codice a blocco $\mathbb{Z}_m$–libero $\mathcal{C}$ di rate R, con un qualsiasi decodificatore $\Phi_{\mathcal{C}}$, ha probabilità di errore*
$$P(e|\mathcal{C}) \geq K \ .$$

A questo punto ci si pone naturalmente la domanda se $\hat{C}_m$ sia minore strettamente oppure uguale a C_m, cioè se la restrizione dall'insieme dei codici a blocco a quello dei codici $\mathbb{Z}_m$–liberi comporti o meno una perdita di capacità.

Se m è un numero primo, allora $\mathbb{Z}_m$ non ha sottogruppi propri, quindi l'unico sottocanale da considerare è lo stesso canale. Dunque in questo caso si ha
$$\hat{C}_m = C, \qquad \hat{E}_m(R) = E_m(R).$$
In effetti, quando m è primo $\mathbb{Z}_m$ è un campo; che i codici lineari su campi finiti raggiungano capacità su un canale simmetrico è un risultato ben noto nella letteratura (si vedano ad esempio [15] e [33]).

Quando m non è primo $\mathbb{Z}_m$ ha dei sottogruppi propri, quindi un canale $\mathbb{Z}_m$–lineare ha degli effettivi sottocanali. Per scoprire se tali sottocanali provochino una perdita di capacità per i codici $\mathbb{Z}_m$–lineari, i.e. se
$$\exists \ \ l \mid m, \ l \neq 1 \quad \text{t.c.} \qquad (\log m)C_l < (\log l)C_m,$$
è necessario studiare il caso specifico. Per quanto riguarda i canali gaussiani additivi di tipo S o di tipo (S, S') la risposta dipende direttamente dalla geometria della costellazione S. Nel prossimo paragrafo verrà dimostrato che, in un caso di notevole interesse applicativo, la costellazione $S = 2^r$–PSK con gruppo generatore $G = \mathbb{Z}_{2^r}$, non si ha perdita di capacità, i.e.
$$\hat{C}_{2^r} = C_{2^r} \ .$$

Si tratta di un risultato tutt'altro che ovvio, dal momento che non è vero per una qualunque costellazione geometricamente uniforme con un qualsiasi gruppo generatore. Nel seguito verrà presentato infatti un esempio di canale gaussiano additivo di tipo (S, S) che presenta un'ostruzione algebrica per i codici $\mathbb{Z}_m$–lineari.

Esempio 5
Consideriamo la costellazione tridimensionale

$$S = (3 - \mathrm{PSK}) \times (2 - \mathrm{PAM})$$

con gruppo generatore $\mathbb{Z}_6 \simeq \mathbb{Z}_2 \times \mathbb{Z}_3$, introdotta nell'Esempio 4 e riportata nella figura 2.4 del capitolo precedente. Fissiamo $L = 1$ e facciamo variare $h > 0$. Consideriamo il canale gaussiano additivo di tipo (S, S), i.e. il canale senza memoria $\mathbb{Z}_6$–simmetrico

$$\{W(\cdot|x) \in \mathcal{P}(\mathbb{Z}_6)\}_{x \in \mathbb{Z}_6}$$

ottenuto come quantizzazione di un canale AWGN di ingressi S su S stessa in uscita. Tale canale è descritto dalla distribuzione $W(\cdot|0) \in \mathcal{P}(\mathbb{Z}_6)$ definita da

$$W(y|0) = (\mathbf{w}_6)_y \ ,$$

dove, direttamente dalla definizione delle regioni di Voronoi di S, si trova

$$\mathbf{w}_6 = \Big((1-p)(1-2q), pq, (1-p)q, p(1-2q), (1-p)q, pq\Big) \ ,$$

$$p = p(h) = \frac{1}{\sqrt{N_0 \pi}} \int_h^{+\infty} e^{(-x^2/N_0)} \mathrm{d}x$$

$$q = q(N_0) = \frac{1}{N_0 \pi} \iint\limits_{\substack{x \geq 0 \\ y \geq 1 - \frac{1}{\sqrt{3}} x}} e^{(-(x^2+y^2)/N_0)} \mathrm{d}x \mathrm{d}y \ .$$

Sia C_6 la capacità del canale $\mathbb{Z}_6$–simmetrico, e siano C_2 e C_3 dei suoi due sottocanali. Si verifica allora che

$$C_6 = C_2 + C_3 \ . \tag{3.15}$$

Definiamo infatti

$$\mathbf{w}_2 = \big(1 - p, p\big) \in \mathcal{P}(3\mathbb{Z}_6) \ ,$$
$$\mathbf{w}_3 = \big(1 - 2q, q, q\big) \in \mathcal{P}(2\mathbb{Z}_6) \ .$$

Possiamo calcolare esplicitamente

$$\begin{array}{l} C_6 = H(Y) - H(Y|X) = \log 6 - H(\mathbf{w}_6) \\ C_3 = H(Y) - H(Y|X) = \log 3 + H(\mathbf{w}_2) - H(\mathbf{w}_6) \\ C_3 = H(Y) - H(Y|X) = \log 2 + H(\mathbf{w}_3) - H(\mathbf{w}_6) \quad . \end{array} \tag{3.16}$$

Ma

$$\begin{aligned} H(\mathbf{w}_6) &= -(1-p)(1-2q)\log\big((1-p)(1-2q)\big) - pq\log\big(pq\big) \\ &\quad -(1-p)q\log\big((1-p)q\big) - p(1-2q)\log\big(p(1-2q)\big) \\ &\quad -(1-p)q\log\big((1-p)q\big) - pq\log\big(pq\big) \\ &= (1-p)H(\mathbf{w}_3) - (1-p)\log(1-p) + pH(\mathbf{w}_3) - p\log p \\ &= H(\mathbf{w}_3) + H(\mathbf{w}_2) \quad . \end{aligned}$$

Dalla (3.16) segue dunque la (3.15). Perché la $\mathbb{Z}_6$–capacità del canale coincida con la sua capacità si deve avere quindi

$$\begin{cases} (\log 2 + \log 3)C_2 \geq \log 2(C_3 + C_2) \\ (\log 2 + \log 3)C_3 \geq \log 3(C_3 + C_2) \end{cases}$$

cioè

$$C_3 = C_2 \quad ,$$

e quindi

$$\log 2 - H(\mathbf{w}_2) = \log 3 - H(\mathbf{w}_3) \quad . \tag{3.17}$$

Ma la (3.17) equivale alla seguente equazione non lineare in h

$$H\left(\frac{1}{2}\operatorname{erfc}\left(\sqrt{h/N_0}\right)\right) = H(\mathbf{w}_3(N_0)) - \log\frac{3}{2} \; . \tag{3.18}$$

Si osservi che H è una funzione continua e strettamente crescente nell'intervallo $[0, \frac{1}{2}]$ e assume valori in $[0, \log 2]$, mentre erfc è continua e strettamente decrescente su $(0, +\infty)$, e assume valori in $(0, 1)$. Dunque, per ogni $N_0 > 0$ fissato, il membro sinistro dell'equazione (3.18) è decrescente in h sull'intervallo $(0, +\infty)$, tende a $\log 2$ per $h \to 0$, e a 0 per $h \to +\infty$. La (3.18) è risolubile solo per ogni N_0 tale che

$$\log 3/2 < H(\mathbf{w}_3(N_0)) < \log 3 \; ; \tag{3.19}$$

inoltre in tal caso la soluzione è unica. La (3.19) equivale a richiedere che sia

$$N_0 \in (0, N_0^*) \; ,$$

dove N_0^* è un valore di soglia definito come l'unica soluzione di

$$H(\mathbf{w}_3(N_0)) = \log 3/2 \ . \tag{3.20}$$

In effetti, si può mostrare che il membro sinistro della (3.20) è strettamente crescente in N_0 sull'intervallo $(0, +\infty)$, e che tende a 0 per $N_0 \to 0$ e a $\log 3$ per $N_0 \to +\infty$; dunque N_0^* è ben definito.

In conclusione, possiamo affermare che i codici liberi su $\mathbb{Z}_6$ usati con la costellazione 3–PSK×2–PAM non permettono tipicamente (i.e. tranne che per valori di N_0 minori di N_0^*, fissando h che soddisfi la (3.18)) di raggiungere la capacità del canale gaussiano additivo di tipo

$$\Big(3-\mathrm{PSK} \times 2-\mathrm{PAM} \ \ , \ \ 3-\mathrm{PSK} \times 2-\mathrm{PAM}\Big) \ . \qquad \square$$

3.5 La costellazione 2^r–PSK con gruppo generatore $\mathbb{Z}_{2^r}$

Restringiamoci al caso di costellazione $S_r = 2^r$–PSK con gruppo generatore $G = \mathbb{Z}_{2^r}$. Useremo in questo paragrafo delle notazioni diverse da quelle adottate finora: in particolare indichiamo con $C_{q,r}$ la capacità del canale gaussiano additivo di tipo

$$(2^q - \mathrm{PSK}, 2^r - \mathrm{PSK}).$$

Consideriamo il vettore delle probabilità di transizione

$$\mathbf{w} := (w_0, \ldots, w_{2^r-1}), \qquad w_j := W(j|0),$$

e, per ogni $0 \le q \le r$ e $i \in \mathbb{Z}_{2^r}$, introduciamo le notazioni seguenti

$$\Lambda_{i,q} := i + 2^{r-q}\mathbb{Z}_{2^r}, \quad \lambda_{i,q} := \sum_{j \in \Lambda_{i,q}} w_j, \quad \underline{\omega}_{i,q} = (w_i, w_{i+2^{r-q},\ldots})$$

$$\underline{\lambda}_q = (\lambda_{0,q}, \lambda_{1,q}, \ldots, \lambda_{2^{r-q}-1,q}), \quad H_{q,r} := H(\underline{\lambda}_q).$$

Poniamo inoltre, per $0 \le q \le r-1$ fissato,

$$i^* = i + 2^{r-q-1} \ ;$$

si osservi come $\Lambda_{i,q} \cup \Lambda_{i^*,q} = \Lambda_{i,q+1}$, e quindi $\lambda_{i,q} + \lambda_{i^*,q} = \lambda_{i,q+1}$.

Esempio 6
Per chiarire le notazioni consideriamo il caso $r=3$, i.e. $G=\mathbb{Z}_8$. Per $q=0$ si ha

$$i^* = i+4,$$
$$\Lambda_{i,0} = \{i\}, \quad \underline{\omega}_{i,0} = (w_i),$$
$$\lambda_0 = \mathbf{w}, \quad H_{0,3} = H(\mathbf{w}).$$

Per $q=1$ si ha

$$i^* = i+2,$$
$$\Lambda_{i,1} = \{i, i+4\}, \quad \Lambda_{i^*,1} = \{i+2, i+6\}$$
$$\underline{\omega}_{i,1} = (w_i, w_{i+4}), \quad \underline{\omega}_{i^*,1} = (w_{i+2}, w_{i+6})$$
$$\underline{\lambda}_1 = (w_0+w_4, w_1+w_5, w_2+w_6, w_3+w_7)$$

Per $q=2$ si ha

$$i^* = i+1,$$
$$\Lambda_{i,2} = \{i, i+2, i+4, i+6\}, \quad \Lambda_{i^*,2} = \{i+1, i+3, i+5, i+7\}$$
$$\underline{\omega}_{i,2} = (w_i, w_{i+2}, w_{i+4}, w_{i+6}), \quad \underline{\omega}_{i^*,2} = (w_{i+1}, w_{i+3}, w_{i+5}, w_{i+7})$$
$$\underline{\lambda}_2 = (w_0+w_2+w_4+w_6, w_1+w_3+w_5+w_7)$$

Infine, per $q=3$ si ha

$$\Lambda_{i,3} = \mathbb{Z}_8, \quad \underline{\omega}_{i,3} = \mathbf{w}$$
$$\underline{\lambda}_3 = (1), \quad H_{3,3} = 0.$$

□

Lemma 14

$$C_{q,r} = q + H_{q,r} - H_{0,r} \tag{3.21}$$

Dimostrazione
Per la Proposizione 5, $C_{q,r}$ è ottenuta con distribuzione uniforme su $2^{r-q}\mathbb{Z}_{2^r}$. La distribuzione corrispondente sulle uscite è data da

$$p_Y(j) = \frac{1}{2^q}\lambda_{j,q}, \quad \forall j \in \mathbb{Z}_{2^r},$$

e dunque ha entropia $H(p_Y) = q + H_{q,r}$. L'entropia condizionata è invece data da:

$$H(Y|X) = H(\mathbf{w}) = H_{0,r}$$

Da questo segue la (3.21). ■

Mentre fino a questo punto l'ipotesi sulla geometria della costellazione non è ancora stata usata, è nella dimostrazione del risultato seguente che questa gioca un ruolo fondamentale nel determinare l'ordinamento del vettore **w** delle probabilità di transizione. Mostriamo dapprima qual è tale ordinamento nell'esempio della costellazione 8–PSK.

Esempio 7 (segue)
Nel caso $r = 3$ si verifica che

$$w_0 > w_7 = w_1 > w_6 = w_2 > w_5 = w_3 > w_4, \tag{3.22}$$

e quindi c'è una alternanza tra w di indice pari (P) e di indice dispari (D) del tipo

$$w_P \geq w_D \geq w_D \geq w_P \geq w_P \geq w_D \geq w_D \geq w_P. \qquad \square$$

L'ordinamento di **w** nel caso r generico presenta una struttura simile.

Lemma 15
Per ogni $1 \leq q \leq r-1$ *e* $0 \leq i \leq 2^{r-q-1}-1$, *le componenti del vettore* $\underline{\omega}_{i,q+1}$ *soddisfano uno dei due ordinamenti seguenti:*

$$\begin{cases} w_i & \geq & w_{i^*+2^{r-q}(2^q-1)} & \geq & w_{i^*} & \geq & w_{i+2^{r-q}(2^q-1)} & \geq \\ w_{i+2^{r-q}} & \geq & w_{i^*+2^{r-q}(2^q-2)} & \geq & w_{i^*+2^{r-q}} & \geq & w_{i+2^{r-q}(2^q-2)} & \geq \\ \vdots & & \vdots & & \vdots & & \vdots & \\ w_{i+2^{r-q}(2^{q-1}-1)} & \geq & w_{i^*+2^{r-1}} & \geq & w_{i^*+2^{r-q}(2^{q-1}-1)} & \geq & w_{i+2^{r-1}} & \end{cases}$$

$$\begin{cases} w_{i^*+2^{r-q}(2^q-1)} & \geq & w_i & \geq & w_{i+2^{r-q}(2^q-1)} & \geq & w_{i^*} & \geq \\ w_{i^*+2^{r-q}(2^q-2)} & \geq & w_{i+2^{r-q}} & \geq & w_{i+2^{r-q}(2^q-2)} & \geq & w_{i^*+2^{r-q}} & \geq \\ \vdots & & \vdots & & \vdots & & \vdots & \\ w_{i^*+2^{r-1}} & \geq & w_{i+2^{r-q}(2^{q-1}-1)} & \geq & w_{i+2^{r-1}} & \geq & w_{i^*+2^{r-q}(2^{q-1}-1)} & \end{cases} \tag{3.23}$$

Dimostrazione
Dalla definizione delle regioni di Voronoi di $S' = 2^r$–PSK

$$V_s := \{r \in \mathbb{R}^n \ \text{ t.c. } \ ||r-s||_{\mathbb{R}^n} \leq ||r-s^*||_{\mathbb{R}^n} \ , \ \forall s^* \in S'\}, \qquad s \in S' \ ,$$

e dal fatto che la densità di una gaussiana n–variata di valore atteso s e matrice di convarianza $N_0 Id_n$ sia funzione decrescente della distanza euclidea dal punto s, segue che l'ordinamento decrescente delle probabilità di transizione **w** coincide con quello crescente delle distanze euclidee dal punto 0 della costellazione. Definiamo

$$\phi = 2\pi \frac{i}{2^r}, \quad \theta = 2\pi \frac{1}{2^{q+1}}$$

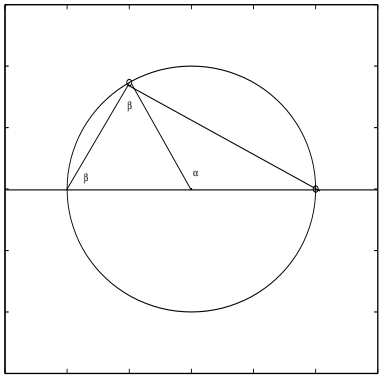


Figura 3.2: La relazione tra la fase α e la distanza dal punto 0

e consideriamo l'insieme delle fasi dei segnali di $\Lambda_{i,q+1}$, i.e.

$$\Omega_{i,q+1} = \{\phi + l\theta | l = 0, \ldots, 2^{q+1} - 1\}.$$

Con semplici ragionamenti geometrici (vedi figura 3.2) si verifica che il segmento congiungente i segnali di fase 0 e α rispettivamente ha lunghezza proporzionale a $\sin(\beta)$, dove $\beta := \frac{1}{2}(\pi - |\pi - \alpha|)$.

Quindi, posto

$$f(\alpha) := \pi - |\pi - \alpha|,$$

ci si riduce a trovare l'ordinamento crescente di $f(\Omega_{i,q+1})$. Poichè

$$f(\phi + l\theta) = \begin{cases} \phi + l\theta & 0 \leq l \leq 2^q \\ (2^q - l)\theta - \phi & 2^q \leq l \leq 2^{q+1} - 1 \end{cases}$$

si ha che

$$f(\Omega_{i,q+1}) = \bigcup_{k=0}^{2^{q-1}-1} \Gamma^k$$

con

$$\Gamma^0 = \{\phi, \theta - \phi, \theta + \phi, 2\theta - \phi\}, \quad \Gamma^k = 2k\theta + \Gamma^0.$$

Si noti che, per $0 \leq i \leq \lfloor 2^{r-q-2} \rfloor$, si ha

$$0 \leq \phi \leq \frac{1}{2}\theta$$

e quindi

$$\phi \leq \theta - \phi \leq \theta + \phi \leq 2\theta - \phi;$$

ne segue l'ordinamento (i).
Per $\lfloor 2^{r-q-2} \rfloor + 1 \leq i \leq 2^{r-q-1} - 1$, si ha invece

$$\frac{1}{2}\theta \leq \phi \leq \theta$$

e quindi

$$\theta - \phi \leq \phi \leq 2\theta - \phi \leq \theta + \phi;$$

ne segue l'ordinamento (ii). ■

Dato $1 \leq \alpha \leq q$, partizioniamo $\Lambda_{i,q}$ come

$$\Lambda_{i,q} = \Delta_{i,q}^{\alpha,+} \cup \Delta_{i,q}^{\alpha,-}$$

nel modo seguente:

- si divida $\Lambda_{i,q}$ in $2^{q-\alpha}$ 2^{α}–uple scegliendo di volta in volta i 2^{α} indici di $\Lambda_{i,q}$ corrispondenti alle probabilità di transizione w maggiori;
- di ogni 2^{α}–upla si assegnino i $2^{\alpha-1}$ indici corrispondenti a probabilità di transizione w maggiori a $\Delta_{i,q}^{\alpha,+}$, gli altri a $\Delta_{i,q}^{\alpha,-}$.

Definiamo inoltre

$$\delta_{i,q}^{\alpha+} = \sum_{j \in \Delta_{i,q}^{\alpha,+}} w_j \ , \qquad \delta_{i,q}^{\alpha-} = \sum_{j \in \Delta_{i,q}^{\alpha,-}} w_j.$$

Si osservi che specificare per ogni $1 \leq \alpha \leq q$ l'appartenenza a $\Delta_{i,q}^{\alpha,+}$ o a $\Delta_{i,q}^{\alpha,-}$ individua univocamente un elemento di $\Lambda_{i,q}$, i.e.

$$\left| \bigcap_{\alpha=1}^{q} \Delta_{i,q}^{\alpha,s_\alpha} \right| = 1, \qquad \forall (s_1, \ldots, s_q) \in \{+,-\}^q. \tag{3.24}$$

Esempio 8 (segue)
Per $r = 3$, $q = 2$, $\alpha = 1$, l'algoritmo precedente produce le partizioni seguenti:

$$\Lambda_{0,2} = \Delta_{0,2}^{1,+} \cup \Delta_{0,2}^{1,-} = \{w_0, w_2\} \cup \{w_4, w_6\},$$

$$\Lambda_{1,2} = \Delta_{1,2}^{1,+} \cup \Delta_{1,2}^{1,-} = \{w_1, w_7\} \cup \{w_3, w_5\}.$$

Dall'ordinamento (3.22) segue che

$$\delta_{0,2}^{1,+} + \delta_{1,2}^{1,+} = w_0 + w_2 + w_1 + w_7 \geq w_0 + w_2 + w_6 + w_1 = \lambda_{0,2}, \tag{3.25}$$

$$\delta_{0,2}^{1,+} + \delta_{1,2}^{1,+} = w_0 + w_2 + w_1 + w_7 \geq w_3 + w_5 + w_1 + w_7 = \lambda_{1,2}. \tag{3.26}$$

La (3.25) e la (3.26) implicano in particolare che

$$H\left(\delta_{0,2}^{1,+} + \delta_{1,2}^{1,+}, \delta_{0,2}^{1,-} + \delta_{1,2}^{1,-}\right) \leq H\left(\lambda_{0,2}, \lambda_{1,2}\right). \qquad \square$$

Questo risultato si generalizza come segue.

Lemma 16
Per ogni $1 \leq \alpha \leq q \leq r-1$, $0 \leq i \leq 2^{r-q-1}$,

$$H\left(\frac{\delta_{i,q}^{\alpha+}+\delta_{i^*,q}^{\alpha+}}{\lambda_{i,q+1}}, \frac{\delta_{i,q}^{\alpha-}+\delta_{i^*,q}^{\alpha-}}{\lambda_{i,q+1}}\right) \leq H\left(\frac{\lambda_{i,q}}{\lambda_{i,q+1}}, \frac{\lambda_{i^*,q}}{\lambda_{i,q+1}}\right). \tag{3.27}$$

Dimostrazione
Facciamo vedere che

$$\begin{array}{c}\delta_{i,q}^{\alpha+}+\delta_{i^*,q}^{\alpha+} \geq \lambda_{i,q} \\ \delta_{i,q}^{\alpha+}+\delta_{i^*,q}^{\alpha+} \geq \lambda_{i^*,q}.\end{array}$$

Per costruzione abbiamo

$$\delta_{i,q}^{1,+} \leq \delta_{i,q}^{2,+} \leq \ldots \leq \delta_{i,q}^{q,+}$$

e quindi è sufficiente mostrare che

$$\begin{array}{c}\delta_{i,q}^{1,+}+\delta_{i^*,q}^{1,+} \geq \lambda_{i,q} \\ \delta_{i,q}^{1,+}+\delta_{i^*,q}^{1,+} \geq \lambda_{i^*,q}.\end{array} \tag{3.28}$$

Ma da entrambi gli ordinamenti (3.23i) e (3.23ii) segue che

$$\Delta_{i,q}^{1,+} = \{w_{i+j2^{r-q}} | 0 \leq j \leq 2^{q-1}-1\}, \quad \Delta_{i^*,q}^{1,+} = \{w_{i^*+(2^{q-1}-j)2^{r-q}} | 0 \leq j \leq 2^{q-1}\}$$

e che

$$w_{i+j2^{r-q}} \geq w_{i^*+j2^{r-q}}, \quad w_{i^*+(2^{q-1}-j)2^{r-q}} \geq w_{i+(2^{q-1}-j)2^{r-q}}, \quad \forall 0 \leq j \leq 2^{q-1}-1;$$

quindi

$$\begin{aligned}\delta_{i,q}^{1,+}+\delta_{i^*,q}^{1,+} &= \sum_{j=0}^{2^{q-1}-1}(w_{i+j2^{r-q}}+w_{i^*+(2^{q-1}-i^*)2^{r-q}}) \geq \\ &\geq \sum_{j=0}^{2^{q-1}-1}(w_{i+j2^{r-q}}+w_{i+(2^{q-1}-j)2^{r-q}}) = \lambda_{i,q}\end{aligned}$$

e, analogamente,

$$\delta_{i,q}^{1,+}+\delta_{i^*,q}^{1,+} \geq \sum_{j=0}^{2^{q-1}-1}(w_{i^*+(2^{q-1}-j)2^{r-q}}+w_{i^*+j2^{r-q}}) = \lambda_{i^*,q}$$

■

Possiamo finalmente dimostrare il seguente risultato.

Teorema 17
Per ogni $1 \leq q \leq r-1$,

$$qC_{q+1,r} \leq (q+1)C_{q,r} \tag{3.29}$$

Dimostrazione
Per la (3.21), la (3.29) è equivalente a

$$H_{0,r} + qH_{q+1,r} \leq (q+1)H_{q,r}. \tag{3.30}$$

Ma

$$H_{0,r} = H_{q,r} + \sum_{i=0}^{2^{r-q}-1} \lambda_{i,q} H\left(\frac{\omega_{i,q}}{\lambda_{i,q}}\right) \tag{3.31}$$

$$H_{q,r} = H_{q+1,r} + \sum_{i=0}^{2^{r-q-1}-1} \lambda_{i,q+1} H\left(\frac{\lambda_{i,q}}{\lambda_{i,q+1}}, \frac{\lambda_{i^*,q}}{\lambda_{i,q+1}}\right) \tag{3.32}$$

e quindi è sufficiente mostrare che, per ogni $i = 0, 1, \ldots, 2^{r-q-1} - 1$,

$$q\lambda_{i,q+1} H\left(\frac{\lambda_{i,q}}{\lambda_{i,q+1}}, \frac{\lambda_{i^*,q}}{\lambda_{i,q+1}}\right) \geq \lambda_{i,q} H\left(\frac{\omega_{i,q}}{\lambda_{i,q}}\right) + \lambda_{i^*,q} H\left(\frac{\omega_{i^*,q}}{\lambda_{i^*,q}}\right). \tag{3.33}$$

Dalla (3.24) segue in particolare che

$$H\left(\frac{\omega_{i,q}}{\lambda_{i,q}}\right) \leq \sum_{\alpha=1}^{q} H\left(\frac{\delta_{i,q}^{\alpha+}}{\lambda_{i,q+1}}, \frac{\delta_{i,q}^{\alpha-}}{\lambda_{i,q+1}}\right). \tag{3.34}$$

Quindi, utilizzando nell'ordine la (3.34), la concavità di H e la (3.27), si ha

$$\begin{aligned}
&\frac{\lambda_{i,q}}{\lambda_{i,q+1}} H\left(\frac{\omega_{i,q}}{\lambda_{i,q}}\right) + \frac{\lambda_{i^*,q}}{\lambda_{i,q+1}} H\left(\frac{\omega_{i^*,q}}{\lambda_{i,q}}\right) \leq \\
&\leq \sum_{\alpha=1}^{q} \left[\frac{\lambda_{i,q}}{\lambda_{i,q+1}} H\left(\frac{\delta_{i,q}^{\alpha,+}}{\lambda_{i,q}}, \frac{\delta_{i,q}^{\alpha,-}}{\lambda_{i,q}}\right) + \frac{\lambda_{i^*,q}}{\lambda_{i,q+1}} H\left(\frac{\delta_{i^*,q}^{\alpha,+}}{\lambda_{i^*,q}}, \frac{\delta_{i^*,q}^{\alpha,-}}{\lambda_{i^*,q}}\right)\right] \leq \\
&\leq \sum_{\alpha=1}^{q} H\left(\frac{\delta_{i,q}^{\alpha+} + \delta_{i^*,q}^{\alpha+}}{\lambda_{i,q+1}}, \frac{\delta_{i,q}^{\alpha-} + \delta_{i^*,q}^{\alpha-}}{\lambda_{i,q+1}}\right) \leq \\
&\leq \sum_{\alpha=1}^{q} H\left(\frac{\lambda_{i,q}}{\lambda_{i,q+1}}, \frac{\lambda_{i^*,q}}{\lambda_{i,q+1}}\right) = qH\left(\frac{\lambda_{i,q}}{\lambda_{i,q+1}}, \frac{\lambda_{i^*,q}}{\lambda_{i,q+1}}\right)
\end{aligned} \tag{3.35}$$

■

Dal teorema si ottiene direttamente il risultato seguente.

Corollario 18
Per ogni $1 \leq q_1 \leq q_2 \leq r$

$$C_{q_2,r} \leq \frac{q_2}{q_1} C_{q_1,r} \tag{3.36}$$

■

Si osservi ora che, poiché le dimostrazioni del Lemma 15, e quindi del Lemma 16 e del Teorema 3.29, sfruttano soltanto l'ordinamento delle distanze euclidee dal punto 0 della costellazione, essi restano validi anche per ogni altra costellazione con lo stesso ordinamento, per la quale, dunque, resta valido il Corollario 18.

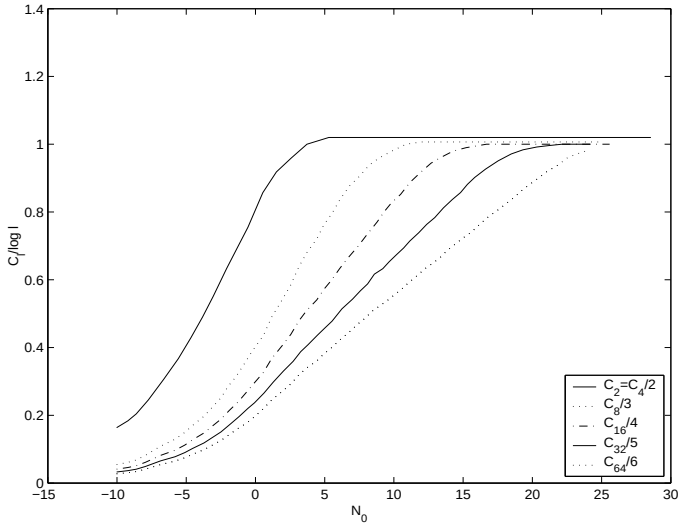


Figura 3.3: Capacità normalizzata (i.e. in [Sym/Ch.Use]) per alcuni canali AWGN di tipo m–PSK

3.6 Conclusioni

Concludiamo riassumendo il principale risultato ottenuto in questo capitolo, che costituisce una risposta –seppur parziale– alla questione sollevata da Loeliger in [18] a pag.1680.

Teorema 19
I codici $\mathbb{Z}_{2^r}$–lineari raggiungono la capacità di un qualsiasi canale gaussiano additivo di tipo (2^r–PSK,2^{r+q}–PSK).

Dimostrazione
Applicando il Teorema 11 ed il Teorema 12, ed il Corollario 18 possiamo concludere che, per ogni $R < C$ le successioni di ensemble di codici $\mathbb{Z}_{2^r}$–lineari $(\mathcal{E}_{\text{Im}}(R,N))_N$ e $(\mathcal{E}_{\text{Ker}}(R,N))_N$ sono molto buone per il canale. Da questo segue che esiste una successione di codici $\mathbb{Z}_{2^r}$–lineari di probabilità di errore tendente a zero, con decodifica ML. ■

3.6.1 Generalizzazione al caso di uscite continue

Se si considera il limite per $q \to +\infty$, si ottiene immediatamente il medesimo risultato per il canale a ingresso discreto e uscita continua che si ottiene dal canale gaussiano additivo di tipo 2^r–PSK proiettandone le uscite su S_1, i.e. la circonferenza di raggio 1. Per i canale gaussiano additivo di tipo 2^r–PSK, i.e. quello di uscita $\mathbb{R}^2$ senza proiezione su S_1, non abbiamo un analogo risultato

teorico. La capacità di un qualsiasi canale gaussiano additivo di tipo S è pari a

$$C_m = \log m - \frac{1}{m\sqrt{\pi}^N} \sum_{x=0}^{m-1} \int_{\mathbb{R}^n} e^{-|\mathbf{t}|^2} \log \left(\sum_{i=0}^{m-1} e^{\left(-2\frac{\mathbf{t}\cdot(s_x - s_i)}{\sqrt{N_0}} - \frac{|s_x - s_i|^2}{N_0} \right)} \right) \mathrm{d}\mathbf{t} \tag{3.37}$$

I risultati di un'integrazione numerica della (3.37) (ripresi da [22]) nel caso dei canali AWGN di tipo m–PSK, con $m = 2, 4, 8, 16, 32, 64$, sono riportati nella figura 3.3, normalizzati rispetto a $\log m$. Tale grafico, oltre a semplici considerazioni circa l'ampiezza della banda usata da tali costellazioni, a parità di rapporto segnale–rumore, ci induce ad avanzare la seguente congettura.

Congettura Sia C_{2^r} la capacità del canale gaussiano additivo di tipo 2^r–PSK. Allora, per ogni $1 \leq r \leq s$

$$rC_{2^s} \leq sC_{2^r} \ . \qquad \square$$

Se tale congettura fosse confermata, allora avremmo dimostrato che i codici $\mathbb{Z}_{2^r}$–lineari raggiungono la capacità di un qualsiasi canale gaussiano additivo di tipo 2^r–PSK. Questo completerebbe la risposta alla questione posta da Loeliger per le costellazioni 2^r-PSK.

3.6.2 Costellazione 3–PSK×2–PAM

Ritorniamo alla costellazione 3–PSK×2–PAM. Abbiamo mostrato nell'esempio 5 che, tranne che per particolari valori di N_0 e h, tale costellazione presenta delle ostruzioni algebriche, e quindi i codici $\mathbb{Z}_6$–liberi non raggiungono capacità. Tuttavia è possibile raggiungere la capacità con codici $\mathbb{Z}_6$–lineari non liberi, nella maniera seguente. Abbiamo visto che:

$$C_6 = C_2 + C_3 \ .$$

Per ogni $R \in [0, C_6)$, poniamo

$$R_2 := \frac{C_2}{C_6} R \ , \qquad R_3 := \frac{C_3}{C_6} R \ ;$$

segue immediatamente che

$$R_2 < C_2 \ , \qquad R_3 < C_3 \ .$$

Possiamo quindi applicare il Teorema 12, separatamente ai canali gaussiani additivi di tipo (2–PAM,2–PAM) e (3–PSK,3–PSK) e concludere (si

osservi che 2 e 3 sono primi quindi non ci sono sottocanali) che la successione di ensemble di codici $\mathbb{Z}_2$–lineari di rate maggiore o uguale a R_2, $(\mathcal{E}^{(2)}_{\mathrm{Ker}}(R_2,N))_N$, e quella di codici $\mathbb{Z}_3$–lineari di rate maggiore o uguale a R_3, $(\mathcal{E}^{(3)}_{\mathrm{Ker}}(R_3,N))_N$, indipendenti tra loro, hanno probabilità medie di errore che soddisfano rispettivamente

$$\overline{P(e)}^{(2)} \leq \exp(NE_2(R_2)) \ , \qquad \overline{P(e)}^{(3)} \leq \exp(NE_3(R_3)) \ ,$$

dove $E^{(2)}(R)$ e $E^{(3)}(R)$ sono gli esponenti di errore dei due canali AWGN di tipo (2–PAM,2–PAM) e (3–PSK,3–PSK), da non confondersi con $E_2(R)$ ed $E_3(R)$, esponenti di errore dei sottocanali di tipo (2–PAM,2–PAM$\times$3–PSK) e (3–PSK,2–PAM$\times$3–PSK).

A partire da $\mathcal{E}^{(2)}_{\mathrm{Ker}}(R,N)$ e $\mathcal{E}^{(3)}_{\mathrm{Ker}}(R,N)$ definiamo ora l'ensemble prodotto

$$\mathcal{E}^{(6)}(R,N) := \mathcal{E}^{(3)}_{\mathrm{Ker}}(R_3,N) \times \mathcal{E}^{(2)}_{\mathrm{Ker}}(R_2,N) \ ,$$

dei codici $\mathcal{C}^{(6)} = \mathcal{C}^{(3)} \times \mathcal{C}^{(2)}$ su $\mathbb{Z}_6$ di rate maggiore o uguale a $R = R_2 + R_3$, con la probabilità

$$\mathbb{P}(\mathcal{C}^{(6)}) := \mathbb{P}(\mathcal{C}^{(2)})\mathbb{P}(\mathcal{C}^{(3)}) \ .$$

I codici $\mathcal{C}^{(6)}$ così definiti sono $\mathbb{Z}_6$–lineari, ma generalmente non $\mathbb{Z}_6$–liberi. La probabilità di errore di ciascuno di essi soddisfa

$$1 - P(e|\mathcal{C}^{(6)}) = (1 - P(e|\mathcal{C}^{(2)}))(1 - P(e|\mathcal{C}^{(3)}))$$

e quindi, mediando sull'ensemble $\mathcal{E}^{(6)}(R,N)$, si ha

$$1 - \overline{P(e)}^{(6)} = (1 - \overline{P(e)}^{(2)})(1 - \overline{P(e)}^{(3)}) \ ,$$

da cui segue

$$\begin{aligned} \overline{P(e)}^{(6)} &= \overline{P(e)}^{(2)} + \overline{P(e)}^{(3)} - \overline{P(e)}^{(2)}\overline{P(e)}^{(3)} \\ &\leq \overline{P(e)}^{(2)} + \overline{P(e)}^{(3)} \\ &\leq \exp(NE_2(R_2)) + \exp(NE_2(R_2)) \ . \end{aligned}$$

La probabilità media di errore dell'ensemble $\mathcal{E}^{(6)}(R,N)$, dunque, tende a 0 esponenzialmente in N.

Possiamo concludere che i codici lineari, a differenza di quelli liberi, su $\mathbb{Z}_6$ raggiungono la capacità del canale gaussiano di tipo (3–PSK$\times$2–PAM, 3–PSK$\times$2–PAM).

3.6.3 Costellazione $(2^r, h)$–PSK

Consideriamo ora la famiglia di costellazioni tridimensionali (m, h)–PSK introdotta nell'Esempio 3 del capitolo precedente. Abbiamo già osservato che al limite per $h \to 0$ tali costellazioni degenerano nella m–PSK. È ragionevole dunque aspettarsi che le prestazioni dei codici $\mathbb{Z}_m$–lineari su un canale gaussiano additivo di tipo (m, h)–PSK siano simili a quelle su un canale di tipo m–PSK, per valori di h sufficientemente piccoli. In effetti possiamo dimostrare l'enunciato seguente.

Proposizione 20
Dati q e r in $\mathbb{N}$, e $h \in \mathbb{R}$ tale che

$$0 < h \leq \sqrt{\sin^2 \frac{2\pi}{2^r} - \sin^2(\frac{\pi}{2^r})} \tag{3.38}$$

i codici 2^q–lineari raggiungono la capacità del canale gaussiano additivo di tipo

$$\Big((2^q, h) - PSK, (2^r, h) - PSK\Big) \ .$$

Dimostrazione
La (3.38) implica che l'ordinamento delle distanze euclidee dal punto 0 dei punti della costellazione $(2^r, h)$–PSK sia lo stesso di quello della costellazione 2^r–PSK. Allora, come si è osservato in conclusione del paragrafo 3.5, vale in questo caso il Corollario 18. Si ripete quindi la dimostrazione del Teorema 19. ■

Per h soddisfacente la (3.38) dunque, non abbiamo ostruzioni algebriche alle prestazioni dei codici $\mathbb{Z}_m$–liberi. Per valori di h molto grandi, invece, ci sembra ragionevole supporre che si abbia

$$\hat{C}_m < C_m \ ,$$

e che quindi i codici $\mathbb{Z}_m$–liberi non raggiungano la capacità del canale gaussiano quantizzato su (m, h)–PSK. Non solo, ma in questo caso non si può ripetere lo stesso ragionamento fatto per i codici $\mathbb{Z}_6$–lineari non liberi sulla costellazione 2–PAM×3–PSK, poichè (m, h)–PSK non è il prodotto di due costellazioni ortogonali tra loro. Pensiamo che, in questo caso, nemmeno i codici $\mathbb{Z}_m$–lineari non liberi possano raggiungere la capacità. L'unica possibilità che rimarrebbe, qualora le nostre supposizioni venissero confermate, è quella di usare l'altro gruppo generatore della costellazione (m, h)–PSK, il gruppo diedrale $D_{m/2}$. Tutto questo motiva il progetto di studiare, in futuro, codici lineari su gruppi non abeliani.

Capitolo 4

I codici a bassa densità su $\mathbb{Z}_m$ con decodifica ML

I codici a bassa densità (codici LDPC) sono stati introdotti per la prima volta nel caso binario da Gallager nel 1960 nella sua tesi di PhD [13] e sono poi stati studiati da diversi autori ([21], [20], [26], [30], [31]) e generalizzati su campi finiti non binari in [11]. Noi useremo l'approccio proposto da [4] per cercare un'estensione al caso di $\mathbb{Z}_m$. Nel seguente paragrafo, daremo la definizione precisa dellensemble dei codici LDPC. Obiettivo del resto del capitolo è uno studio fondamentale delle proprietà medie dell'ensemble dei codici a bassa densità e del loro comportamento asintotico. I risultati fondamentali sono contenuti nei teoremi 30, 31, e 37 che riguardano il comportamento della probabilità media di errore dei codici dell'ensemble. Arriveremo ai risultati sopra in vari passi: l'idea fondamentale è utilizzare ancora la stima del Lemma 8 come è stato fatto per l'analisi delle prestazioni degli ensemble classici studiati nel Capitolo 3. Tuttavia, come vedremo, in questo nuovo ensemble la stima dello spettro medio delle distanze risulta molto più complicata che nel caso classico e richiede uno spezzamento di $\mathbb{Z}_m^N$ più raffinato di quello utilizzato nella dimostrazione del Teorema 12.

4.1 Costruzione dell'ensemble dei codici a bassa densità

Introduciamo in questo paragrafo l'ensemble di codici a bassa densità, che sarà oggetto di studio di questo capitolo, a partire dal loro grafo di Tanner: questo approccio è dovuto a [19]. È possibile considerare grafi di Tanner regolari, in cui tutti i variable nodes hanno lo stesso grado e così pure i check nodes, oppure irregolari, nei quali la distribuzione dei gradi dei nodi

di ciascuna delle due classi viene opportunamente scelta. È stato dimostrato che i codici LDPC binari con grafo di Tanner irregolare hanno migliori prestazioni con decodifica iterativa –si vedano [19], [30] e [31]. In questo capitolo prenderemo in considerazione soltanto codici LDPC con grafo di Tanner regolare.

Fissati tre interi positivi c, d e N tali che

$$L := \frac{c}{d} N \in \mathbb{N} \, ,$$

e una $\pi \in S_{Nc}$, permutazione dell'insieme $\{1, \ldots, Nc\}$, si definisca il grafo bipartito (c, d) regolare

$$\mathcal{G}_\pi = (V = \mathcal{N} \cup \mathcal{M}, E_\pi)$$

dove

$$\mathcal{N} = \{v_1, \ldots, v_N\}, \qquad \mathcal{M} = \{h_1, \ldots, h_L\} \, ,$$
$$E_\pi = \big((v_{\lceil i/c \rceil}, h_{\lceil \pi(i)/d \rceil}), \; i = 1, \ldots, Nc \big) \in (\mathcal{N} \times \mathcal{M})^{Ld} \, .$$

Si osservi che si ammette la presenza di archi paralleli, i.e. che collegano una stessa coppia di nodi. Definiamo poi, per ogni $1 \leq n \leq N$, il vicinato di v_n come

$$\mathcal{M}(n) = \left(h_{\lceil \frac{\pi((c-1)n+1)}{d} \rceil}, \ldots, h_{\lceil \frac{\pi(cn)}{d} \rceil} \right) \in \mathcal{M}^c .$$

Ad ogni grafo bipartito $\mathcal{G}_\pi$ corrisponde un codice a blocco $\mathcal{C}_\pi$ di lunghezza N su $\mathbb{Z}_m$ che sia un ordinamento del nucleo dell'omomorfismo $\phi_\pi : \mathbb{Z}_m^N \to \mathbb{Z}_m^L$ definito da

$$(\Phi(\mathbf{x}))_m = \sum_{(v_n, h_m) \in E_\pi} x_n \, , \qquad i = 1, \ldots, L \, ;$$

ordinamento arbitrario con l'unico vincolo che $\mathbf{x}_1 = 0$. Osserviamo che per ogni $\pi \in S_{Nc}$ il codice $\mathcal{C}_\pi$ è non degenere, $\mathbb{Z}_m$–lineare, e ha rate R maggiore o uguale al rate di progetto R_0, che definiamo come

$$R_0 := \left(1 - \frac{c}{d} \right) \log m \, .$$

Sia ora Π una variabile aleatoria distribuita uniformemente sul gruppo di permutazioni S_{Nc}. Π induce naturalmente una struttura probabilistica sull'insieme dei codici a blocco $\mathbb{Z}_m$–lineari di lunghezza N e rate maggiore o uguale a R_0. Questo spazio di probabilità è l'ensemble dei codici a bassa densità (c, d)–regolari di lunghezza N che studieremo nel presente capitolo, e si indica con $\mathcal{E}_{LDPC}(c, d, N)$.

4.2 Una stima dall'alto alla probabilità media di errore di ensemble arbitrari di codici $\mathbb{Z}_m$–lineari

Il risultato che presentiamo viene dimostrato con un ragionamento simile a quello delle dimostrazioni dei teoremi 11 e 12 del capitolo precedente. La novità fondamentale consiste nel separare le probabilità di transizione verso parole che sono 'quasi' in un sottogruppo di $\mathbb{Z}_m^N$ o in una sua classe laterale. Per spiegare il significato del 'quasi' c'è bisogno di introdurre delle ulteriori notazioni. Per ogni $\boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m)$, e per ogni $l \mid m$, si definisca $\boldsymbol{\theta}^{(l)} \in \mathcal{P}_N(\mathbb{Z}_m/\frac{m}{l}\mathbb{Z}_m)$ nel modo seguente

$$\theta_i^{(l)} := \sum_{n \in \frac{m}{l}\mathbb{Z}_m + i} \theta_n \qquad i \in \mathbb{Z}_m \Big/ \tfrac{m}{l}\mathbb{Z}_m \quad . \tag{4.1}$$

Fissato $\delta \in [0,1]$, per ogni $l \mid m$ indichiamo con $J_{\delta,l,i}^N$ il sottoinsieme di $\mathcal{P}_N(\mathbb{Z}_m)$ costituito dai tipi delle sequenze $\mathbf{x}$ che sono 'quasi' in $(i + \frac{m}{l}\mathbb{Z}_m)^N$ e non in una classe più piccola:

$$J_{\delta,l,i}^N := \left\{ \boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m) \text{ t.c. } \theta_i^{(l)} \geq 1-\delta \text{ e } \forall h \,|\, l, h < l, \forall j \in \frac{m}{l}\mathbb{Z}_m + i,\ \theta_j^{(l)} \geq \delta \right\}.$$

Poniamo poi

$$J_{\delta,l}^N = \bigcup_{0 \leq i \leq m/l-1} J_{\delta,l,i}^N.$$

Al variare di $l \mid m$, gli insiemi $J_{\delta,l}^N$ costituiscono un ricoprimento di $\mathcal{P}_N(\mathbb{Z}_m)$, i.e.

$$P_N(\mathbb{Z}_m) = \bigcup_{l|m} J_{\delta,l}^N; \tag{4.2}$$

si osservi che non si tratta tuttavia di un'unione disgiunta. Infine, definiamo

$$\mathcal{T}_{\delta,l,i}^N := \mathcal{T}_{J_{\delta,l,i}^N}^N, \quad \mathcal{T}_{\delta,l}^N := \mathcal{T}_{J_{\delta,l}}^N$$

cioè $\mathcal{T}_{\delta,l,i}^N$ è l'insieme delle parole il cui tipo sta in $J_{\delta,l,i}^N$. Si osservi che possiamo scrivere gli elementi di $\mathcal{T}_{\delta,l,i}^N$ come perturbazione di elementi di $\left(\frac{m}{l}\mathbb{Z}_m + i\right)^N$ con elementi di $\mathcal{T}_{\delta,1,0}^N$ e quindi abbiamo in particolare

$$\mathcal{T}_{\delta,l}^N \subseteq \frac{m}{l}\mathbb{Z}_m^N + \mathcal{T}_{\delta,1}^N. \tag{4.3}$$

Il teorema che stiamo per dimostrare fornisce una stima delle prestazioni di ensemble arbitrari di codici $\mathbb{Z}_m$–lineari il cui rate sia maggiore o uguale

ad un valore R fissato. Dati $R \in [0, \log m]$ e $N \in \mathbb{N}$, definiamo K e L come segue:

$$K := \left\lceil N \left(\frac{R}{\log m} \right) \right\rceil \ , \qquad L := N - K \ .$$

Si osservi come, per R fissato, si ha che

$$\lim_{N \to +\infty} \frac{K}{N} = \frac{R}{\log m} \ .$$

Introduciamo inoltre dei termini di perturbazione, $f(\delta)$ e α_l, funzione rispettivamente del parametro $\delta \geq 0$, e degli spettri medi di distanze $\overline{S(\boldsymbol{\theta})}$ dell'ensemble $\mathcal{E}$:

$$f(\delta) := \delta \log(m-1) + H(\delta) + \log \frac{1-\delta}{1-2\delta} \ , \tag{4.4}$$

$$\alpha_l := \max_{\boldsymbol{\theta} \in J^N_{\delta,l}} \frac{\overline{S(\boldsymbol{\theta})}}{\binom{N}{N\boldsymbol{\theta}} \left(\frac{1}{l}\right)^L} \ . \tag{4.5}$$

Teorema 21
Sia dato un canale $\mathbb{Z}_m$–simmetrico $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathbb{Z}_m}$; si supponga di utilizzare decodifica ML. Fissati arbitrariamente $\delta \in [0,1]$ e $R \in [0, \log m]$, la probabilità media di errore $\overline{P(e)}$ di un ensemble arbitrario $\mathcal{E}(R,N)$ di codici $\mathbb{Z}_m$–lineari di lunghezza N e rate maggiore o uguale a R soddisfa

$$\overline{P(e)} \leq \sum_{\boldsymbol{\theta} \in J^N_{\delta,1}} \overline{S(\boldsymbol{\theta})} \mathbf{D}^{N\boldsymbol{\theta}} + m \sum_{\substack{l|m \\ l>1}} \exp\left(-N \left[E_l \left(R_l + \frac{\log \alpha_l}{N} \right) - f(\delta) \right] \right) \tag{4.6}$$

dove $\mathbf{D}$ è il parametro di Bhattacharyya del canale, i.e.

$$D_i = \sum_{y \in \mathcal{Y}} \sqrt{W(y|0) W(y|i)}, \qquad i \in \mathbb{Z}_m.$$

mentre, per ogni $l \mid m$, $l > 1$, E_l e R_l sono rispettivamente l'esponente di errore ed il minimo rate di utilizzo del sottocanale l–esimo:

$$R_l := \frac{K}{N} \log l \ , \qquad E_l := \max_{0 \leq \rho \leq 1} \left\{ \max_{\mathbf{q} \in \mathcal{P}(\frac{l}{m}\mathbb{Z}_m)} \{E_0(q,\rho)\} - \rho R \right\} \ .$$

Dimostrazione
Le ipotesi di $\mathbb{Z}_m$–linearità dei codici dell'ensemble, e di $\mathbb{Z}_m$–simmetria del canale consentono di scrivere

$$\overline{P(e)} = \mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C})] = \mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C},1)] = \mathbb{E}_{\mathcal{C}} \left[\sum_{\mathbf{y} \in e_1} W_N(\mathbf{y}|\mathbf{0}) \right] \tag{4.7}$$

dove $\mathbf{e}_1$ è l'evento di errore avendo trasmesso la parola $\mathbf{x}_1 = \mathbf{0}$, i.e.

$$e_1 = \{\mathbf{y} \in \mathcal{Y}^N \text{ t.c. } \exists m \geq 2 : W_N(\mathbf{y}|\mathbf{x}_m) \geq W_N(\mathbf{y}|\mathbf{0})\}.$$

Definiamo

$$e_1^{(l)} := \{\mathbf{y} \in \mathcal{Y}^N \text{ t.c. } \exists m \geq 2 : \mathbf{x}_m \in \mathcal{T}_{\delta,l}^N \text{ e } W_N(\mathbf{y}|\mathbf{x}_m) \geq W_N(\mathbf{y}|\mathbf{0})\};$$

dalla (4.2) segue che

$$e_1 \subseteq \bigcup_{l|m} e_1^{(l)} \ .$$

Uno union bound permette quindi di stimare

$$P(e|\mathcal{C},1) \leq \sum_{l|m} \Big(\sum_{\mathbf{y}\in e_1^{(l)}} W_N(\mathbf{y}|\mathbf{0}) \Big) \ . \tag{4.8}$$

Definiamo, per ogni $l \mid m$ il codice $\mathcal{C}^{(l)}$ che consiste esattamente della parola $\mathbf{0}$ e delle parole di $\mathcal{C}$ che stanno in $\mathcal{T}_{\delta,l}^N$, con lo stesso ordinamento con cui appaiono in $\mathcal{C}$, cioè in particolare con $\mathbf{x}_1^{(l)} = \mathbf{0}$. Si osservi che tali codici $\mathcal{C}^{(l)}$ tipicamente non sono $\mathbb{Z}_m$–lineari, benché $\mathcal{C}$ lo sia. Possiamo allora interpretare ciascun addendo della sommatoria più esterna nella (4.8) come la probabilità di errore di $\mathcal{C}^{(l)}$ condizionata alla trasmissione della prima parola

$$P(e|\mathcal{C}^{(l)},1) = \sum_{\mathbf{y}\in e_1^{(l)}} W_N(\mathbf{y}|\mathbf{0}) \ .$$

Stimiamo separatamente tali probabilità di errore. Per $P(e|\mathcal{C}^{(1)})$ usiamo lo union–Bhattacharyya bound:

$$P(e|\mathcal{C}^{(1)},1) \leq \sum_{\boldsymbol{\theta}\in\mathcal{P}_N(\mathbb{Z}_m)} S(\boldsymbol{\theta}|\mathcal{C}^{(1)},1)\mathbf{D}^{N\boldsymbol{\theta}} = \sum_{\boldsymbol{\theta}\in J_{\delta,1}^N} S(\boldsymbol{\theta}|\mathcal{C})\mathbf{D}^{N\boldsymbol{\theta}}. \tag{4.9}$$

Per $l > 1$, applichiamo invece il Lemma 8 del Capitolo 2 a ciascun codice $\mathcal{C}^{(l)}$ con un $\rho_l \in [0,1]$ arbitrariamente scelto; si ha

$$\begin{aligned}
&P(e|\mathcal{C}^{(l)},1) \leq \\
&\leq \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \frac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho_l}} \left(\sum_{\boldsymbol{\theta}\in\mathcal{P}_N(\mathbb{Z}_m)} S(\boldsymbol{\theta}|\mathcal{C}^{(l)},1)\tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}} W_N(\mathbf{y}|\mathbf{x}+\mathbf{v})^{\frac{1}{1+\rho_l}} \right)^{\rho_l} \\
&= \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \frac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho_l}} \left(\sum_{\boldsymbol{\theta}\in J_{\delta,l}^N} S(\boldsymbol{\theta}|\mathcal{C})\tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}} W_N(\mathbf{y}|\mathbf{x}+\mathbf{v})^{\frac{1}{1+\rho_l}} \right)^{\rho_l} .
\end{aligned}$$

Mediando sull'ensemble $\mathcal{E}(R,N)$ e usando la disuguaglianza di Jensen otteniamo

$$\begin{aligned}
&\mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C}^{(l)},1)] \leq \\
&\leq \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \frac{1}{m^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho_l}} \left(\sum_{\boldsymbol{\theta}\in J_{\delta,l}^N} \overline{S(\boldsymbol{\theta})}\tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} W_N(\mathbf{y}|\mathbf{x}+\mathbf{v})^{\frac{1}{1+\rho_l}} \right)^{\rho_l}
\end{aligned} \tag{4.10}$$

Ricordando la definizione di α_l e usando la (4.3), possiamo stimare

$$\begin{aligned}
&\sum_{\boldsymbol{\theta}\in J_{\delta,l}^N} \overline{S(\boldsymbol{\theta})}\tbinom{N}{N\boldsymbol{\theta}}^{-1} \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} W_N(\mathbf{y}|\mathbf{x}+\mathbf{v})^{\frac{1}{1+\rho_l}} \leq \\
&\leq \left(\tfrac{1}{l}\right)^L \alpha_l \sum_{\mathbf{x}\in\mathcal{T}_{\delta,l}^N} W_N(\mathbf{y}|\mathbf{x}+\mathbf{v})^{\frac{1}{1+\rho_l}} \\
&\leq \left(\tfrac{1}{l}\right)^L \alpha_l \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} W_N(\mathbf{y}|\mathbf{x}+\mathbf{v}+\mathbf{t})^{\frac{1}{1+\rho_l}}
\end{aligned} \tag{4.11}$$

Fissiamo un insieme $\Omega_l \subseteq \mathbb{Z}_m^N$ di cardinalità $(\frac{m}{l})^N$ contenente un elemento per ciascuna classe laterale di $\frac{m}{l}\mathbb{Z}_m^N$. Sostituendo la (4.11) nella (4.10), applicando la disuguaglianza di Jensen, e sfruttando la $\mathbb{Z}_m$–simmetria del canale, otteniamo

$$\begin{aligned}
&\mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C}^{(l)},1] \leq \\
&\leq \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{v}\in\mathbb{Z}_m^N} \tfrac{1}{m^N} W_N(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho_l}} \left(\alpha_l \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^L} W_N(\mathbf{y}|\mathbf{x}+\mathbf{v}+\mathbf{t})^{\frac{1}{1+\rho_l}}\right)^{\rho_l} \\
&\leq l^{K\rho_l}\alpha_l^{\rho_l} \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{z}\in\Omega_l} \left(\tfrac{l}{m}\right)^N \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{w}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N(\mathbf{y}|\mathbf{z}+\mathbf{w})^{\frac{1}{1+\rho_l}} \\
&\qquad\qquad \left(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N(\mathbf{y}|\mathbf{z}+\mathbf{w}+\mathbf{x}+\mathbf{t})^{\frac{1}{1+\rho_l}}\right)^{\rho_l} \\
&= l^{K\rho_l}\alpha_l^{\rho_l} \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{z}\in\Omega_l} \left(\tfrac{l}{m}\right)^N \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{w}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N(\mathbf{y}|\mathbf{z}+\mathbf{w})^{\frac{1}{1+\rho_l}} \\
&\qquad\qquad \left(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N(\mathbf{y}|\mathbf{z}+\mathbf{x}+\mathbf{t})^{\frac{1}{1+\rho_l}}\right)^{\rho_l} \\
&= l^{K\rho_l}\alpha_l^{\rho_l} \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{z}\in\Omega_l} \left(\tfrac{l}{m}\right)^N \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{w}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N((-\mathbf{z})\mathbf{y}|\mathbf{w})^{\frac{1}{1+\rho_l}} \\
&\qquad\qquad \left(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N((-\mathbf{z})\mathbf{y}|\mathbf{x}+\mathbf{t})^{\frac{1}{1+\rho_l}}\right)^{\rho_l} \\
&= l^{K\rho_l}\alpha_l^{\rho_l} \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{w}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N(\mathbf{y}|\mathbf{w})^{\frac{1}{1+\rho_l}} \left(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N(\mathbf{y}|\mathbf{x}+\mathbf{t})^{\frac{1}{1+\rho_l}}\right)^{\rho_l} \\
&= l^{K\rho_l}\alpha_l^{\rho_l} \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} \sum_{\mathbf{w}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N(\mathbf{y}|\mathbf{w})^{\frac{1}{1+\rho_l}} \left(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \tfrac{1}{l^N} W_N((-\mathbf{t})\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho_l}}\right)^{\rho_l} \\
&= l^{K\rho_l}\alpha_l^{\rho_l} \sum_{\mathbf{t}\in\mathcal{T}_{\delta,1}^N} \sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})A((-\mathbf{t})\mathbf{y})^{\rho_l}
\end{aligned} \tag{4.12}$$

avendo posto

$$A(\mathbf{y}) := \sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N} \frac{1}{l^N} W_N(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho_l}}$$

.

Si ha

$$\begin{aligned}
&\sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})A((-\mathbf{t})\mathbf{y})^{\rho} + \sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})A(\mathbf{t}\mathbf{y})^{\rho} \\
&= \sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})A((-\mathbf{t})\mathbf{y})^{\rho} + \sum_{\mathbf{y}\in\mathcal{Y}^N} A((-\mathbf{t})\mathbf{y})A(\mathbf{y})^{\rho} \\
&\leq \sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})^{\rho+1} + \sum_{\mathbf{y}\in\mathcal{Y}^N} A((-\mathbf{t})\mathbf{y})^{\rho+1} \\
&= 2\sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})^{\rho+1}
\end{aligned}$$

dove si è usata la disuguaglianza, valida per ogni $x, y, \rho \geq 0$,

$$xy^{\rho} + x^{\rho}y \leq x^{\rho+1} + y^{\rho+1} \quad .$$

Si osservi che $\mathcal{T}^N_{\delta,1} = -\mathcal{T}^N_{\delta,1}$ e che la cardinalità di $\mathcal{T}^N_{\delta,1}$ può essere sovrastimata così

$$\begin{aligned}
|\mathcal{T}^N_{\delta,1}| &\leq m\left(\sum_{w=0}^{\lfloor N\delta\rfloor}\tbinom{N}{w}(m-1)^w\right) \\
&\leq m\tbinom{N}{N\delta}(m-1)^{N\delta}\sum_{w=0}^{\lfloor N\delta\rfloor}\left(\tfrac{\delta}{1-\delta}\right)^w \\
&\leq m\exp(N[\delta\log(m-1) + H(\delta) + \log\tfrac{1-\delta}{1-2\delta}]) \\
&= m\exp(Nf(\delta)) \quad ,
\end{aligned}$$

dove $f(\delta)$ è stata definita nella (4.4). Si ha quindi

$$\begin{aligned}
&\sum_{\mathbf{t}\in\mathcal{T}^N_{\delta,1}}\sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})A((-\mathbf{t})\mathbf{y})^{\rho_l} = \\
&= \frac{1}{2}\Big(\sum_{\mathbf{t}\in\mathcal{T}^N_{\delta,1}}\sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})A((-\mathbf{t})\mathbf{y})^{\rho_l} + \sum_{\mathbf{t}\in\mathcal{T}^N_{\delta,1}}\sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})A(\mathbf{t}\mathbf{y})^{\rho_l}\Big) \\
&\leq \sum_{\mathbf{t}\in\mathcal{T}^N_{\delta,1}}\sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})^{1+\rho_l} \\
&\leq m\exp(Nf(\delta))\sum_{\mathbf{y}\in\mathcal{Y}^N} A(\mathbf{y})^{1+\rho_l} \quad .
\end{aligned}$$

Sostituendo nella (4.12), fattorizzando $W_N(\mathbf{y}|\mathbf{x})$ e applicando la proprietà distributiva si ottiene

$$\begin{aligned}
\mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C}^{(l)},1)] &\leq m\exp(Nf(\delta))l^{K\rho_l}\alpha_l^{\rho_l}\sum_{\mathbf{y}\in\mathcal{Y}^N}\left(\sum_{\mathbf{x}\in\frac{m}{l}\mathbb{Z}_m^N}\tfrac{1}{l^N}W_N(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho_l}}\right)^{1+\rho_l} \\
&= l^{K\rho_l}\alpha_l^{\rho_l}m\exp(Nf(\delta))\left[\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\frac{m}{l}\mathbb{Z}_m}\left(\tfrac{1}{l}\right)W(y|x)^{\frac{1}{1+\rho_l}}\right)^{1+\rho_l}\right]^N \\
&= m\exp\left(-N\left[E_0(\mathbf{u}_{\frac{m}{l}\mathbb{Z}_m},\rho_l) - \rho\big(R_l + \tfrac{\log\alpha_l}{N}\big) - f(\delta)\right]\right) \quad ,
\end{aligned} \tag{4.13}$$

dove $\mathbf{u}_{\frac{m}{l}\mathbb{Z}_m}$ è la distribuzione uniforme su $\frac{m}{l}\mathbb{Z}_m$. Infine, ricordando che l'esponente di errore di un canale $\frac{m}{l}\mathbb{Z}_m$–simmetrico si ottiene con distribuzione uniforme in ingresso, i.e. per ogni $0 \leq \rho \leq 1$

$$\max_{p \in \mathcal{P}(\frac{m}{l}\mathbb{Z}_m)} E_0(p, \rho) = E_0(\mathbf{u}_{\frac{m}{l}\mathbb{Z}_m}, \rho) \quad ,$$

ottimizzando su ρ_l abbiamo

$$\mathbb{E}_{\mathcal{C}}[P(e|\mathcal{C}^{(l)}, 1)] \leq m \exp\left(-N\left[E_l\left(R_l + \frac{\log \alpha_l}{N}\right) - f(\delta)\right]\right) \quad . \tag{4.14}$$

Sostituendo la (4.14) e la (4.9) nella (4.8) si ottiene la (4.6). ■

Il teorema 21 riconduce l'analisi delle prestazioni medie di un ensemble $\mathcal{E}(R, N)$ di codici $\mathbb{Z}_m$–lineari di lunghezza N alla stima del suo spettro medio di distanze $\overline{S(\boldsymbol{\theta})}$, al variare di $\boldsymbol{\theta}$ in $J^N_{\delta,l}$. Tali stime per gli ensemble di codici a bassa densità $\mathcal{E}_{LDPC}(c, d, N)$ introdotti nel paragrafo precedente saranno oggetto dei prossimi due paragrafi.

4.3 Una prima stima di $\mathbb{P}(\mathbf{x} \in \mathcal{C})$

I codici a bassa densità, così come tutti i codici definiti come ordinamento del nucleo di un omomorfismo, sono non degeneri per costruzione. Quindi,

$$S(\boldsymbol{\theta}(\mathbf{0})|\mathcal{C}_\pi) = 0$$

per ogni $\mathcal{C}_\pi$ nell'ensemble $\mathcal{E}_{LDPC}(c, d, N)$. Inoltre, tutti gli altri tipi $\boldsymbol{\theta} \in \mathcal{P}(\mathbb{Z}_m)$, si ha

$$\overline{S(\boldsymbol{\theta})} = \mathbb{E}_\Pi\left[S(\boldsymbol{\theta}|\mathcal{C}_\Pi)\right] = \mathbb{E}_\Pi\Big[\sum_{\mathbf{x} \in \mathcal{T}^N_{\boldsymbol{\theta}}} \mathbb{1}_{\{\mathbf{x} \in \mathrm{supp}(\mathcal{C}_\Pi)\}}\Big] = \sum_{\mathbf{x} \in \mathcal{T}^N_{\boldsymbol{\theta}}} \mathbb{P}_\Pi(\mathbf{x} \in \mathrm{supp}(\mathcal{C}_\Pi)) \quad .$$

La notazione $\mathbb{P}_\Pi(\mathbf{x} \in \mathrm{supp}\,\mathcal{C}_\Pi)$ pone enfasi sul fatto che la variabile rispetto alla quale si calcola la probabilità è Π e non $\mathbf{x}$. Nel seguito tuttavia, quando non ci sarà pericolo di confusione, useremo, per indicare tale quantità, la notazione più semplice

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \quad .$$

Dunque $\mathbb{P}(\mathbf{x} \in \mathcal{C})$ è una funzione deterministica di $\mathbf{x} \in \mathbb{Z}^N_m$. Come vedremo, per la simmetria insita nella definizione stessa degli ensemble di codici

a bassa densità $\mathcal{E}_{LDPC}(c,d,N)$ data nel paragrafo 4.1, $\mathbb{P}(\mathbf{x}\in\mathcal{C})$ è funzione solo del tipo $\boldsymbol{\theta}(\mathbf{x})$, così che

$$\overline{S(\boldsymbol{\theta}|\mathcal{C})} = \sum_{\mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N} \mathbb{P}(\mathbf{x}\in\mathcal{C}) = |\mathcal{T}_{\boldsymbol{\theta}}^N| = \binom{N}{N\boldsymbol{\theta}}\mathbb{P}(\mathbf{x}\in\mathcal{C}) \ , \qquad \mathbf{x}\in\mathcal{T}_{\boldsymbol{\theta}}^N \ .$$

Nel seguito proponiamo due stime di $\mathbb{P}(\mathbf{x}\in\mathcal{C})$ che possono essere viste come una sorta di estensione delle stime in [4]. La prima di tali stime, che presentiamo in questo paragrafo, sarà utilizzata per $\mathbf{x}\in\mathcal{T}_{\delta,l}^N$ con $l>1$. L'altra invece, presentata nel prossimo paragrafo, servirà quando $l=1$.

L'enunciato seguente mostra come $\mathbb{P}(\mathbf{x}\in\mathcal{C})$ sia funzione solo di $\boldsymbol{\theta}(\mathbf{x})$.

Lemma 22
Per ogni $\mathbf{x}\in\mathbb{Z}_m^N$, la probabilità che un codice $\mathcal{C}$ dell'ensemble $\mathcal{E}_{LDPC}(c,d,N)$ contenga $\mathbf{x}$ è

$$\mathbb{P}(\mathbf{x}\in\mathcal{C}) = \binom{Nc}{Nc\ \boldsymbol{\theta}(\mathbf{x})}^{-1} \sum_{\substack{(\boldsymbol{\omega}^{(1)},\dots,\boldsymbol{\omega}^{(L)})\in(\Gamma_d)^L \\ t.c.\ \sum_{i=1}^{L}\boldsymbol{\omega}^{(i)}=L\boldsymbol{\theta}(\mathbf{x})}} \prod_{i=1}^{L}\binom{d}{d\boldsymbol{\omega}^{(i)}}$$

dove

$$\Gamma_d := \left\{ \boldsymbol{\omega}\in\mathcal{P}_d(\mathbb{Z}_m)\ t.c.\ m \mid d\sum_{k=1}^{d} k\omega_k \right\} .$$

Dimostrazione
Si osservi come prima cosa che una parola $\mathbf{x}\in\mathbb{Z}_m^d$ soddisfa l'equazione

$$\sum_{i=1}^{d} x_i = 0 \ (\bmod m) \tag{4.15}$$

se e solo se il suo tipo $\boldsymbol{\theta}(\mathbf{x})\in\mathcal{P}_d(\mathbb{Z}_m)$ è tale che

$$m \mid d\sum_{i=0}^{m} i\theta_i(\boldsymbol{x}),$$

i.e. $\boldsymbol{\theta}(\mathbf{x})\in\Gamma_d$. Inoltre, il numero di soluzioni della (4.15) di tipo $\boldsymbol{\omega}\in\Gamma_d$ è chiaramente pari a

$$\binom{d}{d\boldsymbol{\omega}} .$$

Sia ora $\mathbf{x}\in\mathbb{Z}_m^{Nc}$, e consideriamo il sistema di equazioni

$$\sum_{i=(j-1)d+1}^{jd} x_i = 0 \pmod m \quad \forall j=1,\dots L \quad . \tag{4.16}$$

Si osservi che si tratta di un sistema di equazioni mutamente esclusive e congiuntamente esaustive su $\mathbb{Z}_m^{Nc}$, nel senso che ogni componente dell'incognita $\mathbf{x}$ partecipa ad una e una sola equazione del sistema. Ne segue che $\mathbf{x} \in \mathbb{Z}_m^{Nc}$ è soluzione di (4.16) se e solo se è possibile scriverla come concatenazione di L soluzioni della (4.15), i.e.

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \vdots \\ \mathbf{x}^{(L)} \end{pmatrix}$$

con $\mathbf{x}^{(j)} \in \mathbb{Z}_m^d$ soluzione della (4.15) per ogni $j = 1, \ldots, L$. Per le considerazioni appena svolte si ha che, per ogni tipo $\boldsymbol{\theta} \in \mathcal{P}_{Nc}(\mathbb{Z}_m)$, il numero di soluzioni di tipo $\boldsymbol{\theta}$ del sistema (4.16) è dato da

$$W_{\boldsymbol{\theta}} = \sum_{\substack{(\boldsymbol{\omega}^{(1)},\ldots,\boldsymbol{\omega}^{(L)}) \in (\Gamma_d)^L \\ \text{t.c. } \sum_{i=1}^{L} \omega^{(i)} = L\boldsymbol{\theta}}} \prod_{i=1}^{L} \binom{d}{d\boldsymbol{\omega}^{(i)}} \quad . \tag{4.17}$$

Sia ora $\mathbf{x} \in \mathbb{Z}_m^N$ una N–upla fissata; definiamo $\mathbf{x}^c \in \mathbb{Z}_m^{Nc}$ come la replica di $\mathbf{x}$ per c volte:

$$x_i^c := x_{\lceil i/c \rceil} \qquad i = 1, \ldots, Nc.$$

Si ossservi che il tipo $\boldsymbol{\theta}(\mathbf{x}^c) \in \mathcal{P}_{Nc}(\mathbb{Z}_m)$ coincide con il tipo $\boldsymbol{\theta}(\mathbf{x}) \in \mathcal{P}_N(\mathbb{Z}_m)$. Possiamo quindi esprimere la cardinalità dello stabilizzatore di $\mathbf{x}^c$ nel gruppo delle permutazioni S_{Nc} (definendo l'azione di $\pi \in S_{Nc}$ su $\mathbf{x} \in \mathbb{Z}_m^{Nc}$ come $(\pi\mathbf{x})_i = \mathbf{x}_{(\pi i)}$) nella maniera segunte:

$$|Stab(\mathbf{x}^c)| = \prod_{i=1}^{m} \big(Nc\ \theta_i(\mathbf{x}^c)\big)! = \prod_{i=1}^{m} \big(Nc\ \theta_i(\mathbf{x})\big)! \quad .$$

In conclusione, fissata una N–upla $\mathbf{x} \in \mathbb{Z}_m^N$ la probabilità che $\mathbf{x}$ sia una parola di codice dell'ensemble $\mathcal{E}_{LDPC}(c, d, N)$ è pari a

$$\begin{aligned}
\mathbb{P}(\mathbf{x} \in \mathcal{C}) &= \mathbb{P}_\Pi \left(\sum_{i:\lceil \Pi(i)/d \rceil = j} x_i^c = 0 \pmod m,\ j = 1, \ldots, L \right) \\
&= \mathbb{P}_\Pi \left(\sum_{i=(j-1)d+1}^{jd} \big(\Pi^{-1}\mathbf{x}^c\big)_i = 0 \pmod m,\ j = 1, \ldots, L \right) \\
&= \sum_{\pi \in S_{NC}} \tfrac{1}{(Nc)!} \delta_{\{\pi\mathbf{x}^c \text{ è soluzione di (4.16)}\}} \\
&= \frac{1}{(Nc)!} |Stab(\mathbf{x}^c)| W_{\boldsymbol{\theta}(\mathbf{x})} \\
&= W_{\boldsymbol{\theta}(\mathbf{x})} \tbinom{Nc}{Nc\boldsymbol{\theta}(\mathbf{x})}^{-1} .
\end{aligned}$$

■

Per stimare i coefficienti $W_{\boldsymbol{\theta}}$ introdotti nella dimostrazione precedente, è utile ricorrere ai cosiddetti multinomi generatori associati. Premettiamo una notazione: date due N–uple $\mathbf{x}, \mathbf{z}$ useremo la notazione

$$\mathbf{x}^{\mathbf{z}} := x_1^{z_1} \dots x_n^{z_n}$$

con la convenzione $0^0 = 1$. Definiamo ora

$$W_{\boldsymbol{\theta}} := \sum_{\substack{(\boldsymbol{\omega}^{(1)},\dots,\boldsymbol{\omega}^{(L)}) \in (\Gamma_d)^L \\ \text{t.c. } \sum_{i=1}^{L} \omega^{(i)} = L\boldsymbol{\theta}}} \prod_{i=1}^{L} \binom{d}{d\boldsymbol{\omega}^{(i)}} \quad .$$

Sia $\xi \in \mathbb{C}$ una radice m–esima primitiva dell'unità (p.e. $\xi = \exp(i\frac{2\pi}{m})$). Per ogni $0 \le j \le m-1$ definiamo i multinomi

$$g_j(\mathbf{z}) := (z_0 + \xi^j z_1 + \xi^{2j} z_2 + \dots + \xi^{(m-1)j} z_{m-1})^d = \sum_{\boldsymbol{\theta} \in \mathcal{P}_d(\mathbb{Z}_m)} \binom{d}{d\boldsymbol{\theta}} \mathbf{z}^{d\boldsymbol{\theta}} \xi^{jd \sum_{i=0}^{m-1} i\theta_i}$$

e la loro media

$$\alpha(\mathbf{z}) := \sum_{j=0}^{m-1} \frac{1}{m} g_j(\mathbf{z}) \quad .$$

Vale il risultato seguente.

Lemma 23

$$W(\mathbf{z}) = (\alpha(\mathbf{z}))^L \quad . \tag{4.18}$$

Dimostrazione
Si verifica che

$$1 + \xi^k + \xi^{2k} + \dots + \xi^{(m-1)k} = \begin{cases} m & \text{se } m \mid k \\ 0 & \text{se } m \nmid k \end{cases} \tag{4.19}$$

Infatti, se $k \mid m$, allora $\xi^{ka} = 1$ per ogni $1 \le a \le m-1$. Se invece $k \nmid m$, allora $\xi^k \neq 1$, quindi abbiamo

$$1 + \xi^k + \xi^{2k} + \dots + \xi^{(m-1)k} = \frac{\xi^{mk} - 1}{\xi^k - 1} = 0 \quad .$$

Da questo segue che

$$\alpha(\mathbf{z}) = \sum_{j=0}^{m-1} \frac{1}{m} g_j(\mathbf{z}) = \sum_{\boldsymbol{\theta} \in \mathcal{P}_d(\mathbb{Z}_m)} \binom{d}{d\boldsymbol{\theta}} \mathbf{z}^{d\boldsymbol{\theta}} \left(\sum_{j=0}^{m-1} \frac{1}{m} \xi^{dj \sum_{i=0}^{m-1} i\theta_i} \right) = \sum_{\substack{\boldsymbol{\theta} \in \mathcal{P}_d(\mathbb{Z}_m): \\ m \mid d \sum_i i\theta_i}} \binom{d}{d\boldsymbol{\theta}} \mathbf{z}^{d\boldsymbol{\theta}} \quad .$$

Ma allora dalla forma (4.17) segue subito la tesi. ■

Sia ora $\mathbf{x} \in \mathbb{Z}_m^N$ una parola fissata; forniamo una stima dall'alto della probabilità che $\mathbf{x}$ sia una parola di codice dell'ensemble $\mathcal{E}_{LDPC}(c, d, N)$.

Lemma 24
Per ogni $\mathbf{x} \in \mathbb{Z}_m^N$,

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \leq (cN+1)^m (A(\boldsymbol{\theta}(\mathbf{x})))^L \tag{4.20}$$

dove

$$A(\boldsymbol{\theta}) := \frac{1}{m}\left(\sum_{l|m} \varphi\left(\frac{m}{l}\right) \left[1 - (1-\lambda) \sum_{\substack{n,k=0 \\ n \neq k}}^{m/l-1} \theta_n^{(l)} \theta_k^{(l)} \right]^{d/2} \right), \tag{4.21}$$

$$\lambda := \cos\left(\frac{2\pi}{m}\right) < 1 \ ,$$

e $\varphi : \mathbb{N} \to \mathbb{N}$ *è la funzione di Eulero, i.e. per ogni* $n \in \mathbb{N}$ $\varphi(n)$ *è pari al numero di interi positivi relativamente primi con* n, *e* $\boldsymbol{\theta}^{(l)} \in \mathcal{P}_N(\mathbb{Z}_m/\frac{m}{l}\mathbb{Z}_m)$ *sono definiti nella (4.1).*

Dimostrazione
Segue dal Lemma 23 che, per ogni $\boldsymbol{\theta} \in \mathcal{P}_{Nc}(\mathbb{Z}_m)$, si ha :

$$W_{\boldsymbol{\theta}} \mathbf{z}^{Nc\boldsymbol{\theta}} \leq W(\mathbf{z}) = (\alpha(\mathbf{z}))^L \ , \tag{4.22}$$

qualunque sia $\mathbf{z} \in (\mathbb{R}^+)^m$. Segue allora dal Lemma 22 che, dalla (4.22) scegliendo $\mathbf{z} = \boldsymbol{\theta}(\mathbf{x})$, e usando la solita stima (A.2) per i coefficienti binomiali che vale:

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \leq \frac{(\alpha(\boldsymbol{\theta}(\mathbf{x})))^L}{\boldsymbol{\theta}^{Nc\boldsymbol{\theta}}} \frac{(Nc+1)^m}{\exp(NcH(\boldsymbol{\theta}(\mathbf{x})))} = (\alpha(\boldsymbol{\theta}(\mathbf{x})))^L (Nc+1)^m. \tag{4.23}$$

Resta da stimare

$$\alpha(\boldsymbol{\theta}) = \frac{1}{m} \sum_{j=0}^{m-1} g_j(\boldsymbol{\theta}), \qquad g_j(\boldsymbol{\theta}) = \left(\sum_{n=0}^{m-1} \theta_n \xi^{jn} \right)^d \ . \tag{4.24}$$

Posto

$$\mathrm{MCD}(m, j) = l \ ,$$

e ricordando la definizione (4.1) di $\boldsymbol{\theta}^{(l)}$, possiamo stimare $g_j(\boldsymbol{\theta})$ nel modo seguente:

$$\begin{aligned} g_j(\boldsymbol{\theta}) &= \left(\sum_{n=0}^{m/l-1} \theta_n^{(l)} \xi^{jn} \right)^d \\ &= \left(\sum_{n=0}^{m/l-1} (\theta_n^{(l)})^2 + \sum_{\substack{n,\,k=0 \\ n \neq k}}^{m/l-1} \theta_n^{(l)} \theta_k^{(l)} \Re(\xi^{j(n-k)}) \right)^{d/2} \\ &\leq \left(\sum_{n=0}^{m/l-1} \theta_n^{(l)})^2 + \lambda \sum_{\substack{n,\,k=0 \\ n \neq k}}^{m/l-1} \theta_n^{(l)} \theta_k^{(l)} \right)^{d/2} = \\ &= \left(\left(\sum_{n=0}^{m/l-1} (\theta_n^{(l)}) \right)^2 - (1-\lambda) \sum_{\substack{n,\,k=0 \\ n \neq k}}^{m/l-1} \theta_n^{(l)} \theta_k^{(l)} \right)^{d/2} = \\ &= \left(1 - (1-\lambda) \sum_{\substack{n,\,k=0 \\ n \neq k}}^{m/l-1} \theta_n^{(l)} \theta_k^{(l)} \right)^{d/2}. \end{aligned} \tag{4.25}$$

Si not ora che i due insiemi

$$\{0 \leq j \leq m-1 \text{ t.c. } \mathrm{MCD}(m,j) = l\}\,, \quad \{0 \leq j \leq m/l - 1 \text{ t.c. } \mathrm{MCD}(m,j) = 1\}$$

sono in corrispondenza biunivoca tramite la mappa $j \mapsto j/l$. Sostituendo la (4.25) nella (4.24), e quindi nella (4.23), si ottiene la (4.20). ■

Osserviamo che, fissato $\boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m)$, per ogni $l \mid m$, il termine

$$\sum_{\substack{n,k=0 \\ n \neq k}}^{m/l-1} \theta_n^{(l)} \theta_k^{(l)}$$

si annulla se e solo se

$$\mathrm{supp}(\boldsymbol{\theta}^{(l)}) = \{i\}$$

per qualche $0 \leq i < m/l$. Questo avviene se e solo se

$$\mathrm{supp}(\boldsymbol{\theta}) \subseteq \frac{m}{l}\mathbb{Z}_m + i.$$

Presentiamo i due risultati seguenti, che costituiscono un'applicazione del Lemma 24.

Corollario 25
Fissato n tale che $0 \le n \le N$, per ogni $\mathbf{x} \in \mathbb{Z}_m^N$ tale che

$$\theta_a(\mathbf{x}) = 1 - \frac{n}{N}$$

per qualche $a \in \mathbb{Z}_m$, si ha che

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \le (cN+1)^m \frac{1}{m^L} \left(m - \varphi(m) + \varphi(m) \left[1 - (1-\lambda) \frac{n}{N} \left(1 - \frac{n}{N} \right) \right]^{d/2} \right)^L .$$

Dimostrazione
Stimiamo i termini che compaiono nel membro sinistro della (4.21) nella maniera seguente:

$$\sum_{\substack{n,k=0 \\ n \neq k}}^{m/l-1} \theta_n^{(l)} \theta_k^{(l)} \ge 0 \ , \quad l > 1;$$

$$\sum_{\substack{n,k=0 \\ n \neq k}}^{m-1} \theta_n^{(1)} \theta_k^{(1)} \ge \theta_a(1-\theta_a) = \frac{n}{N} \left(1 - \frac{n}{N} \right) .$$

Applicando il lemma precedente e queste stime si ottiene la tesi. ■

Corollario 26
Fissato $\delta \in [0, \frac{1}{2}]$, per ogni $\mathbf{x} \in \mathcal{T}_{\delta,l}^N$ si ha

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \le (cN+1)^m \left(\frac{1}{l} + (1 - \frac{1}{l}) B(\delta) \right)^L , \tag{4.26}$$

dove

$$B(\delta) := \left[1 - (1-\lambda)\delta(1-\delta) \right]^{d/2} .$$

Dimostrazione
Se $\boldsymbol{\theta} \in J_{\delta,l,i}^N$, allora, per ogni $h \mid m$ tale che $l \nmid h$ si ha

$$\delta \le \theta_i^{(h)} \le 1 - \delta .$$

Infatti, posto $a = MCD(l,h)$, $l \nmid h$ implica $a < l$ e quindi, direttamente dalla definizione di $J_{\delta,l,i}^N$ segue che

$$\theta_i^{(h)} = \sum_{k \in \frac{m}{h}\mathbb{Z}_m + i} \theta_k \ge \sum_{k \in \frac{m}{a}\mathbb{Z}_m + i} \theta_k = \theta_i^{(a)} \ge \delta .$$

Inoltre, se $l \nmid m$, allora esiste $j \in \frac{m}{l}\mathbb{Z}_m$ tale che $j \notin \frac{m}{h}\mathbb{Z}_m$, e quindi

$$\theta_i^h = \sum_{n \in i + \frac{m}{h}\mathbb{Z}_m} \theta_n \leq 1 - \theta_{i+j} \leq 1 - \delta \quad .$$

Dunque, per ogni $h \mid m$ tale che $l \nmid h$

$$\sum_{\substack{n,k=0 \\ n \neq k}}^{m/h-1} \theta_n^{(h)} \theta_k^{(h)} \geq \theta_i^{(h)} \sum_{\substack{k=0 \\ k \neq i}}^{m/h-1} \theta_k^{(h)} = \theta_i^{(h)}(1 - \theta_i^{(h)}) \geq \min_{\delta \leq x \leq 1-\delta} x(1-x) = \delta(1-\delta)$$

Abbiamo quindi, per ogni $\boldsymbol{\theta} \in J_{\delta,l}^N$,

$$A(\boldsymbol{\theta}) \quad \leq \quad \frac{1}{m}\left(\sum_{h : l|h|m} \varphi(\tfrac{m}{h}) + \sum_{\substack{h|m: \\ l \nmid h}} [1 - (1-\lambda)\delta(1-\delta)]^{d/2} \right)$$

Possiamo applicare la formula (A.4) riportata in Appendice, e ottenere

$$\sum_{h : l|h|m} \varphi\left(\frac{m}{h}\right) = \sum_{h | \frac{m}{l}} \varphi\left(\frac{m}{lh}\right) = \frac{m}{l} \ ,$$

e quindi

$$A(\boldsymbol{\theta}) \leq \frac{1}{m}\left(\frac{m}{l} + \left(m - \frac{m}{l}\right) B(\delta)\right) = \frac{1}{l}\left(1 + \left(1 - \frac{1}{l}\right) B(\delta)\right).$$

Infine, la (4.26) segue dal lemma precedente. ■

4.4 Una nuova stima di $\mathbb{P}(\mathbf{x} \in \mathcal{C})$

Si osservi come la stima (4.26) risulti del tutto inutile se $l = 1$. In effetti, in questo caso possiamo ottenere una stima migliore. Fissato δ abbastanza piccolo, studiamo quindi $\mathbb{P}(\mathbf{x} \in \mathcal{C})$ per le N–uple $\mathbf{x} \in \mathcal{T}_{\delta,1}^N$, i.e. le parole con quasi tutti i simboli uguali. Ovviamente la N-upla di soli '0' è in ogni codice dell'ensemble. Il risultato seguente riguarda le N-uple con un'alta percentuale di '0'.

Lemma 27
Fissato $\boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m)$ tale che

$$1 - \frac{2}{d} \leq \theta_0 < 1, \tag{4.27}$$

per ogni $\mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N$ *si ha che*

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \leq \binom{L}{\lfloor \frac{cw}{2} \rfloor} \left(\frac{wc}{2L}\right)^{wc} \tag{4.28}$$

dove $w := (1-\theta_0)N$ *è il numero di elementi non nulli di* $\mathbf{x}$.

Dimostrazione
$\mathbf{x}^c$ ha wc elementi non nulli. Perchè $\mathbf{x}$ sia in $\mathcal{C}$ è necessario che ciascuno degli L check sommi nessuno oppure almeno due di tali elementi. Quindi, definita $\mathbf{y} \in \{0,1\}^{Nc}$ come

$$y_i := \begin{cases} 0 & \text{se } (\mathbf{x}^c)_i = 0 \\ 1 & \text{se } (\mathbf{x}^c)_i \neq 0 \end{cases} \qquad i = 1, \ldots, Nc \ ,$$

si ha

$$\begin{aligned} \mathbb{P}(\mathbf{x} \in \mathcal{C}) &= \mathbb{P}_\Pi \left(\sum_{i=(j-1)d+1}^{jd} (\Pi \mathbf{x}^c)_i = 0, \ 1 \leq j \leq L \right) \\ &\leq \mathbb{P}_\Pi \left(\sum_{i=(j-1)d+1}^{jd} y_{\Pi(i)} \neq 1, \ 1 \leq j \leq L \right) \\ &= \frac{|Stab(\mathbf{y})|}{(Nc)!} |A| \quad , \end{aligned} \tag{4.29}$$

dove

$$A := \left\{ \mathbf{y} \in \{0,1\}^{Nc} \text{ t.c. } w_H(\mathbf{y}) = wc \quad \text{e} \quad \sum_{i=(j-1)d+1}^{jd} y_i \neq 1, \ 1 \leq j \leq L \right\}.$$

Ma $|Stab(\mathbf{y})| = (wc)!$, mentre per stimare $|A|$ ragioniamo come segue. Ad ogni $\mathbf{y} \in \{0,1\}^{Nc}$ corrisponde una matrice $B \in \{0,1\}^{d \times L}$ definita da

$$B_{i,j} := y_{d(j-1)+i}.$$

Ovviamente B e $\mathbf{y}$ hanno lo stesso numero di elementi non nulli. Inoltre $\mathbf{y} \in A$ se e solo se tutte le colonne di B hanno un numero di elementi non nulli diverso da 1. Il numero di colonne di B con almeno due elementi non nulli deve essere minore o uguale a $\lfloor \frac{wc}{2} \rfloor$; si osservi che l'ipotesi (4.27) implica che $\lfloor \frac{wc}{2} \rfloor \leq L$. Il numero di sottoinsiemi di cardinalità $\lfloor \frac{wc}{2} \rfloor$ dell'insieme delle colonne di B è pari a $\binom{L}{\lfloor \frac{wc}{2} \rfloor}$. Sia C uno di tali sottoinsiemi fissato; il numero di assegnazioni di wc elementi non nulli alle colonne di C è al più $\binom{\lfloor \frac{wc}{2} \rfloor d}{wc}$. Dunque il numero di matrici $B \in \{0,1\}^{d \times L}$ di peso wc tali che ogni colonna contenga un numero di elementi diverso da 1 è minore o uguale a $\binom{L}{\lfloor \frac{wc}{2} \rfloor}\binom{\lfloor \frac{wc}{2} \rfloor d}{wc}$ e quindi

$$|A| \leq \binom{L}{\lfloor \frac{wc}{2} \rfloor}\binom{\lfloor \frac{wc}{2} \rfloor d}{wc} \ .$$

Sostituendo nella (4.29) si ottiene, ricordando che $Ld = Nc$,

$$\begin{aligned}
\mathbb{P}(\mathbf{x} \in \mathcal{C}) &\leq \binom{L}{\lfloor \frac{wc}{2} \rfloor} \binom{\lfloor \frac{wc}{2} \rfloor d}{wc} \binom{Ld}{wc}^{-1} \\
&= \binom{L}{\lfloor \frac{wc}{2} \rfloor} \frac{(Ld - wc)!}{(Ld)!} \frac{(\lfloor \frac{wc}{2} \rfloor d)!}{(\lfloor \frac{wc}{2} \rfloor d - wc)!} \\
&= \binom{L}{\lfloor \frac{wc}{2} \rfloor} \frac{(\lfloor \frac{wc}{2} \rfloor d)(\lfloor \frac{wc}{2} \rfloor d - 1) \dots (\lfloor \frac{wc}{2} \rfloor d - wc + 1)}{(Ld)(Ld-1) \dots (Ld - wc + 1)} \frac{(\lfloor \frac{wc}{2} \rfloor d)!}{(\lfloor \frac{wc}{2} \rfloor d - wc)!} \\
&\leq \binom{L}{\lfloor \frac{wc}{2} \rfloor} \left(\frac{wc}{2L} \right)^{wc} .
\end{aligned}$$

■

Per quanto riguarda le parole di peso 1, i.e. con un solo elemento non nullo, proponiamo la seguente stima, più raffinata della precedente e necessaria –come vedremo nel seguito– per determinare l'esatto andamento asintotico della probabilità media di errore degli ensemble $\mathcal{E}_{LDPC}$.

Lemma 28
Per ogni $\mathbf{x} \in \mathbb{Z}_m^N$ *tale che* $\theta_0(\mathbf{x}) = 1 - \frac{1}{N}$,

- *se* $\mathrm{MCD}(c, m) = 1$ *allora*

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) = 0 \ ;$$

- *se* $\mathrm{MCD}(c, m) > 1$ *allora*

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \leq \binom{L}{c/a} \left(\frac{c}{aL} \right)^c ,$$

dove

$$a := \mathrm{mfpc}(m, c) \ , \tag{4.30}$$

è il più piccolo fattore primo comune a m *e* c.

Dimostrazione
Sia $\mathbf{x} \in \mathbb{Z}_m^N$ definito da

$$x_n = b \ , \qquad x_i = 0 \ , i \neq n \tag{4.31}$$

per qualche $1 \leq n \leq N$, $1 \leq b \leq m - 1$. Allora la somma degli L checks su $\mathbf{x}$ è pari alla somma degli elementi di $\mathbf{x}^c$, i.e.

$$\sum_{j=1}^{L} \sum_{i=(j-1)d+1}^{jd} (\pi \mathbf{x}^c)_i = \sum_{i=1}^{Nc} (\pi \mathbf{x}^c)_i = bc \ .$$

Perché $\mathbf{x}$ sia in $\mathcal{C}$ è necessario che

$$m \mid bc \ . \tag{4.32}$$

D'altra parte, se $\mathrm{MCD}(c, m) = 1$, allora necessariamente la (4.32) implica che $m \mid b$ e questo è impossibile in quanto $1 \leq b \leq m - 1$. Questo dimostra il primo punto.

Passiamo al caso in cui c e m non sono primi tra loro; sia $e = MCD(c, m)$. Allora

$$m \mid bc \Leftrightarrow \frac{m}{e} \mid \frac{c}{e}\, b \Leftrightarrow \frac{m}{e} \mid b \ ,$$

dove l'ultima equivalenza deriva dal fatto che $\frac{c}{e}$ e $\frac{m}{e}$ sono primi tra loro. Poniamo

$$b = \frac{m}{e} f \ ,$$

dove $1 \leq f \leq e - 1$. Si noti ora che il numero di elementi l non nulli che un check deve sommare per essere soddisfatto deve essere tale che $m \mid lb$, cioè $m \mid efl$. Questo implica che

$$\mathrm{MCD}(e, l) > 1 \ . \tag{4.33}$$

Infatti, per ogni l tale che $MCD(e, l) = 1$, si ha

$$e \nmid fl \ ,$$

e quindi

$$m \nmid \frac{m}{e} fl \ .$$

Il minimo l per il quale vale la (4.33) è proprio a definito in (4.30). Bisogna dunque studiare la probabilità che ogni check sommi nessuno oppure almeno a elementi non nulli di $\mathbf{x}^c$. Tale probabilità si stima con gli stessi ragionamenti fatti nella dimostrazione precedente con $a = 2$; si ottiene dunque

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \leq \binom{L}{c/a} \left(\frac{c}{aL}\right)^c \ . \qquad \blacksquare$$

Passiamo ora a considerare le N-uple con un'alta percentuale di elementi uguali tra loro e diversi da '0'.

Lemma 29
Fissato $a \in \mathbb{Z}_m \setminus \{0\}$, per ogni $\mathbf{x} \in \mathbb{Z}_m^N$ tale che

$$1 - \frac{1}{d} < \theta_a(\mathbf{x}) < 1, \tag{4.34}$$

si ha che

$$\mathbb{P}(\mathbf{x} \in \mathcal{C}) \leq \binom{L}{\lfloor \frac{cw}{2} \rfloor} \left(\frac{wc}{2L}\right)^{wc} \tag{4.35}$$

dove

$$w = (1 - \theta_a(\mathbf{x}))N$$

è il numero di elementi diversi da a di $\mathbf{x}$.

Dimostrazione
Sia $\mathbf{a} \in \mathbb{Z}_m$, una N–upla di tutti simboli a. La somma di d elementi qualsiasi di $\mathbf{x}$ è pari a da, quindi ciascuno degli L checks è soddisfatto se e solo se $m \mid da$. Dunque $\mathbf{a} \in \mathcal{C}$ se e solo se $m \mid da$.

Sia ora $\mathbf{x} \in \mathbb{Z}_m^N$ soddisfacente la (4.34). Consideriamo separatamente i due casi $m \mid da$ e $m \nmid da$.

Se $m \mid da$, allora la N–upla $\mathbf{a}$ è sicuramente in $\mathcal{C}$. Quindi, poichè $\mathcal{C}$ è $\mathbb{Z}_m$–lineare,

$$\mathbf{x} \in \mathcal{C} \Leftrightarrow \mathbf{x} - \mathbf{a} \in \mathcal{C}.$$

Ma il tipo della N-upla $\mathbf{x} - \mathbf{a}$ soddisfa le ipotesi del lemma 27, applicando il quale si ottiene la (4.28).

Passiamo al caso $m \nmid da$. La (4.34) implica che $\mathbf{x}^c$ abbia almeno $L(d-1)+c$ elementi uguali ad a e quindi necessariamente almeno uno degli L checks somma d di tali elementi e non è soddisfatto poiché $m \nmid da$. ■

4.5 La probabilità di errore dei codici a bassa densità su $\mathbb{Z}_m$

Usando le stime ricavate nei paragrafi precedenti possiamo finalmente dimostrare due risultati sul comportamento asintotico della probabilità media di errore degli ensemble di codici low-density su $\mathbb{Z}_m$ che generalizzano quelli presenti in letteratura per il caso binario. Il primo enunciato afferma che la successione di ensemble $(\mathcal{E}_{LDPC}(c,d,N))_N$ è molto buona se si possono scegliere arbitrariamente c e d fissato il loro rapporto. Il secondo afferma che, per c e d fissati, la successione di ensemble $\mathcal{E}_{LDPC}(c,d,N)$ è buona.

Teorema 30
Si consideri un canale $\mathbb{Z}_m$–simmetrico fissato di $\mathbb{Z}_m$–capacità $\hat{C}$; si supponga di utilizzare decodifica ML. Sia assegnato un rate di progetto R tale che

$$0 < R < \hat{C} \quad .$$

Allora esiste $d_0 \in \mathbb{N}$ tale che, per ogni scelta di c e d in $\mathbb{N}$ soddisfacente

$$d \geq d_0 \ , \tag{4.36}$$

$$c \geq 3 \ , \qquad \frac{R}{\log m} \leq 1 - \frac{c}{d} < \frac{\hat{C}}{\log m} \quad , \tag{4.37}$$

la probabilità media di errore degli ensemble $\mathcal{E}_{LDPC}(c, d, N)$ *soddisfa le seguenti stime asintotiche:*

- *se* $\mathrm{MCD}(c, m) = 1$,

$$\overline{P(e)} = O(N^{2-c}), \quad N \to +\infty; \tag{4.38}$$

- *se* $\mathrm{MCD}(c, m) > 1$,

$$\overline{P(e)} = O(N^{1-\frac{a-1}{a}c}), \quad N \to +\infty, \tag{4.39}$$

dove

$$a := \mathrm{mfpc}(c, m) \quad .$$

Dimostrazione
Applichiamo il teorema 21 all'ensemble di codici low–density $\mathcal{E}_{LDPC}(c, d, N)$, dove c e d verranno fissati più avanti soddisfacenti (4.37), così come il parametro δ:

$$\overline{P(e)} \leq \sum_{\boldsymbol{\theta} \in J^N_{\delta,1}} \overline{S(\boldsymbol{\theta})} \mathbf{D}^{N\boldsymbol{\theta}} + m \sum_{\substack{l|m \\ l>1}} \exp\left(-N \left[E_l\left(R_l + \frac{\log \alpha_l}{N} \right) - f(\delta) \right] \right) \quad , \tag{4.40}$$

$$R_l = \left(1 - \frac{c}{d}\right) \log l \ .$$

Riscriviamo la (4.40) nella maniera seguente

$$\overline{P(e)} \leq P_I + P_{II} + P_{III} + P_{IV} \quad , \tag{4.41}$$

dove

$$P_I := \sum_{a=1}^{m-1} \sum_{\substack{\boldsymbol{\theta} \in \mathbb{P}_N(\mathbb{Z}_m): \\ 1-\delta \leq \theta_a \leq 1}} \overline{S(\boldsymbol{\theta})} \mathbf{D}^{N\boldsymbol{\theta}}$$

$$P_{II} := \sum_{\substack{\boldsymbol{\theta} \in \mathbb{P}_N(\mathbb{Z}_m): \\ 1-\beta \leq \theta_0 < 1}} \overline{S(\boldsymbol{\theta})} \mathbf{D}^{N\boldsymbol{\theta}}$$

$$P_{III} := \sum_{\substack{\boldsymbol{\theta} \in \mathbb{P}_N(\mathbb{Z}_m): \\ 1-\delta \leq \theta_0 < 1-\beta}} \overline{S(\boldsymbol{\theta})} \mathbf{D}^{N\boldsymbol{\theta}}$$

$$P_{IV} := m \sum_{\substack{l \mid m \\ l > 1}} \exp\left(-N \left[E_l\left(R_l + \frac{\log \alpha_l}{N} \right) - f(\delta) \right] \right) \quad , \tag{4.42}$$

$$\beta := \frac{2}{d} e^{-12} \exp\left(- \left[\frac{2}{c-2} \log\left((m-1)\frac{d}{2} \right) \right] \right) \quad . \tag{4.43}$$

Stimiamo separatamente i termini P_I, P_{II}, P_{III} e P_{IV}. Iniziamo con P_I.

$$\begin{aligned} P_I &\leq \sum_{a=1}^{m-1} \sum_{w=0}^{\lfloor N\delta \rfloor} (m-1)^w \tbinom{N}{w} D_a^{N-w} \\ &\leq \sum_{a=1}^{m-1} \frac{1}{1 - \frac{\delta D_a}{(1-\delta)(m-1)}} \left((m-1)^\delta D_a^{1-\delta} \right)^N \quad . \end{aligned}$$

Dunque, condizione sufficiente affinché P_I abbia andamento esponenzialmente decrescente a 0 in N, è che

$$(m-1)^\delta D_a^{1-\delta} < 1, \quad a = 1, \ldots, m-1 \quad . \tag{4.44}$$

Per quanto riguarda P_{II}, poiché $\beta \leq \frac{2}{d}$ la (4.27) è soddisfatta, e possiamo quindi applicare il lemma 27, oltre al lemma 28:

$$\begin{aligned} &\sum_{\substack{\boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m): \\ 1-\beta \leq \theta_0 < 1}} \overline{S(\boldsymbol{\theta})} \mathbf{D}^{N\boldsymbol{\theta}} \\ &= \sum_{\substack{\boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m): \\ \theta_0 = 1 - \frac{1}{N}}} N \mathbb{P}(\mathbf{x} \in \mathcal{C} | \mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N) + \sum_{\substack{\boldsymbol{\theta} \in \mathcal{P}_N(\mathbb{Z}_m): \\ 1-\beta \leq \theta_0 < 1 - \frac{1}{N}}} \tbinom{N}{N\boldsymbol{\theta}} \mathbb{P}(\mathbf{x} \in \mathcal{C} | \mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N) \\ &= h(N) + \sum_{l=2}^{\beta N} \tbinom{N}{l} (m-1)^l \tbinom{L}{\lfloor \frac{cl}{2} \rfloor} \left(\frac{lc}{2L} \right)^{lc} \\ &= h(N) + \sum_{l=2}^{\beta N} g(l), \end{aligned}$$

dove si è posto

$$h(N) := \begin{cases} N \tbinom{L}{c/a} \left(\frac{c}{aL} \right)^c = O(N^{1 - \frac{a-1}{a} c}) & \text{se MCD}(c, m) > 1 \\ 0 & \text{se MCD}(c, m) = 1 \end{cases}$$

$$a = \text{mfpc}(m, c)$$

$$g(l) := \binom{N}{l} (m-1)^l \binom{L}{\lfloor \frac{lc}{2} \rfloor} \left(\frac{lc}{2L} \right)^{lc} \quad .$$

Si ha che

$$\begin{aligned} \frac{g(l+2)}{g(l)} &= (m-1)^2 \frac{\tbinom{N}{l+2}}{\tbinom{N}{l}} \frac{\tbinom{L}{\lfloor \frac{lc}{2} \rfloor}}{\tbinom{L}{\lfloor \frac{lc}{2} \rfloor + c}} \frac{\left(\frac{(l+2)c}{2L} \right)^{(l+2)c}}{\left(\frac{lc}{2L} \right)^{lc}} \leq \\ &\leq (m-1)^2 \left(\frac{N-l}{l} \right)^2 \left(\frac{L - \lfloor \frac{lc}{2} \rfloor}{\lfloor \frac{lc}{2} \rfloor} \right)^c \left(1 + \tfrac{2}{l} \right)^{(l+2)c} \leq \\ &\leq \left(\tfrac{d}{2} \right)^c (m-1)^2 e^{4c} \beta^{c-2} \quad . \end{aligned}$$

Da come è stato definito β nella (4.43) segue immediatamente che, comunque si scelgano c e d,

$$\alpha := \left((m-1)e^{2c}\frac{d}{2}\right)^2 \left(\beta\frac{d}{2}\right)^{c-2} < 1 \quad .$$

Abbiamo così che

$$\begin{aligned}\sum_{l=2}^{N\beta} g(l) &\leq \sum_{l=1}^{+\infty} g(l) = \\ &= (g(2)+g(3)) \sum_{l=0}^{+\infty} \alpha^n \\ &= \tfrac{1}{1-\alpha}\,(g(2)+g(3))\end{aligned}$$

e quindi

$$\begin{aligned}\sum_{l=2}^{N\beta} g(l) &\leq \tfrac{1}{1-\alpha}\left(\tbinom{N}{2}(m-1)^2\tbinom{L}{c}\left(\tfrac{c}{L}\right)^{2c} + \tbinom{N}{3}(m-1)^3\tbinom{L}{\lfloor\frac{3c}{2}\rfloor}\left(\tfrac{3c}{2L}\right)^{3c}\right) \\ &= O\left(N^{2-c}\right).\end{aligned}$$

Possiamo concludere quindi che, se m e c sono primi tra loro, allora

$$P_{II} = O(N^{2-c}); \tag{4.45}$$

altrimenti

$$P_{II} = O(N^{1-\frac{a-1}{a}c}) \ . \tag{4.46}$$

Passiamo a stimare P_{III}. Applicando il Lemma 25 si ha

$$\begin{aligned}&\sum_{\substack{\boldsymbol{\theta}\in\mathcal{P}_N(\mathbb{Z}_m):\\ 1-\delta\leq\theta_0<1-\beta}} \overline{S(\boldsymbol{\theta})}\mathbf{D}^{N\boldsymbol{\theta}} \leq \sum_{\substack{\boldsymbol{\theta}\in\mathcal{P}_N(\mathbb{Z}_m):\\ 1-\delta\leq\theta_0<1-\beta}} \overline{S(\boldsymbol{\theta})} \\ &= \sum_{n=\lfloor N\beta\rfloor}^{\lfloor N\delta\rfloor} \sum_{\substack{\mathbf{x}\in\mathbb{Z}_m^N:\\ \theta_0(\mathbf{x})=1-n/N}} \mathbb{P}(\mathbf{x}\in\mathcal{C}) \\ &\leq \frac{N}{m^L} \max_{N\beta\leq n\leq N\delta} \left\{\tbinom{N}{n}(m-1)^n(cN+1)^m\Big(m-\varphi(m)+\varphi(m)\big[1-(1-\lambda)\tfrac{n}{N}(1-\tfrac{n}{N})\big]^{d/2}\Big)^L\right\}.\end{aligned}$$

Dunque,

$$\begin{aligned}&\frac{\log P_{III}}{N} \leq \frac{\log N(cN+1)^m}{N} \\ &+ \max_{\beta\leq t\leq\delta}\left\{\tilde{f}(\delta) + \tfrac{c}{d}\left[\log\left(m-\varphi(m)+\varphi(m)\left[1-(1-\lambda)t(1-t)\right]^{d/2}\right) - \log m\right]\right\}\end{aligned}$$

dove

$$\hat{f}(\delta) := H(\delta) + \delta\log(m-1) \quad .$$

Condizione sufficiente affinché P_{III} abbia andamento esponenzialmente decrescente a 0 in N, per N sufficientemente grande, è che, supposto $\beta \leq \delta \leq \frac{1}{2}$, si abbia

$$\tilde{f}(\delta) + \frac{c}{d}\left[\log\left(m - \varphi(m) + \varphi(m)\left[1 - \frac{1-\lambda}{2}\beta\right]^{d/2}\right)\right] - \log m < 0 \quad . \qquad (4.47)$$

Utilizzando la disuguaglianza

$$1 - t \leq e^{-t} \qquad ,$$

e ricordando la definizione di β, abbiamo che condizione sufficiente per la (4.47) è

$$\tilde{f}(\delta) + \frac{c}{d}\left[\log\left(m - \varphi(m) + \varphi(m)e^{\left[-\frac{\lambda}{2e^{12}}\exp\left(-\left[\frac{2\log((m-1)d/2)}{c-2}\right]\right)\right]}\right)\right] < 0 \quad . \qquad (4.48)$$

Stimiamo ora P_{IV}. Poiché si è supposto

$$1 - c/d < \hat{C}/\log m \ ,$$

si ha che

$$R_l < C_l \ , \qquad l \mid m,\ l > 1,$$

e quindi esiste una costante strettamente positiva A tale che

$$E_l(R_l) \geq A \ , \qquad l \mid m,\ l > 1.$$

Utilizzando la (4.26) si può stimare

$$\alpha_l \leq (cN+1)^m \left(1 + (l-1)B(\delta)\right)^L$$

e quindi

$$\frac{\log \alpha_l}{N} \leq \frac{c}{d}\log\left(1 + (l-1)B(\delta)\right) + \frac{m}{N}\log(cN+1).$$

Dunque, condizione sufficiente affinché P_{IV} abbia andamento definitivamente decrescente a zero esponenzialmente in N è che

$$E_l\left(R_l + \frac{c}{d}\log(1 + (l-1)B(\delta))\right) - f(\delta) > 0 \qquad \forall l \mid m, l > 1. \qquad (4.49)$$

Si scelgano ora $c^*, d^* \in \mathbb{N}$ soddisfacenti la (4.37). Si scelga $\delta > 0$ tale che siano soddisfatte contemporaneamente la (4.44), la (4.48) e la seguente

$$E_l(R_l) + f(\delta) > A/2 \ , \qquad \forall l \mid m, l > 1 \quad .$$

Si osservi che è sempre possibile trovare un tale δ poiché

$$\lim_{\delta \to 0} f(\delta) = \lim_{\delta \to 0} \tilde{f}(\delta) = 0 \quad .$$

e per ogni $a = 1, \ldots, m-1$

$$\lim_{\delta \to 0} (m-1)^{\delta} D_a^{1-\delta} = D_a < 0 \ .$$

Fissiamo un siffatto $\delta > 0$. Si osservi che potrebbe accadere, con la scelta fatta di c^* e d^*, che $\beta > \delta$. Poniamo ora $d = d^* k$ e $c = c^* k$. Si osservi che, con tale scelta, il primo membro della (4.48) è decrescente in k, mentre

$$\lim_{k \to +\infty} B(\delta) = 0. \tag{4.50}$$

Dalla (4.50), dalla continuità di E_l, e dal fatto che β è infinitesimo per $d \to +\infty$, segue che esiste k_0 tale che, per ogni $k \geq k_0$, la (4.49) è soddisfatta, e così pure la (4.48). In tal modo sia σ_3 che P_{II} hanno andamento definitivamente decrescente a zero esponenzialmente in N.

La (4.38) e la (4.39) si ottengono così osservando che l'unico addendo nella (4.41) di andamento asintotico non esponenziale è P_{II}. ■

Il teorema precedente assicura che, per un canale $\mathbb{Z}_m$–simmetrico fissato, per ogni rate di progetto R sotto la sua $\mathbb{Z}_m$–capacità, è sempre possibile trovare una successione di ensemble di codici low–density su $\mathbb{Z}_m$ di rate maggiore o uguale a R che abbiano probailità di errore tendente a 0 in N, scegliendo opportunamente i valori di c e d. Il risultato seguente afferma invece come, fissati c e d, sia sempre possibile ottenere probabilità di errore tendente a 0 in N per successioni di ensemble $\mathcal{E}_{LDPC}(c, d, N)$, migliorando il canale, cioè diminuendo l'entropia delle probabilità di transizione: questa operazione si effettua nella pratica aumentando il rapporto tra la potenza media del segnale e la potenza del rumore.

Sia $\mathcal{Y}$ un $\mathbb{Z}_m$–insieme finito. Per le considerazioni svolte nel paragrafo 2.1, l'insieme dei DMC $\mathbb{Z}_m$–simmetrici di uscite $\mathcal{Y}$ è in corrispondenza biunivoca con $\mathcal{P}(\mathcal{Y})$ attraverso la relazione

$$\mathcal{P}(\mathcal{Y}) \ni W \mapsto \{W(y|x) := W((-x)y) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathbb{Z}_m} \ .$$

Consideriamo lo spazio metrico $\mathcal{P}(\mathcal{Y})$ con la norma L^1. Ricordiamo che l'applicazione $W \mapsto E(R, W)$ è continua rispetto a tale metrica (Teorema 3), e così pure la funzione entropia $W \mapsto H(W)$.

Teorema 31
Siano dati due interi c e d tali che $d > c \geq 3$. Allora esiste $\varepsilon > 0$ tale che la probabilità media di errore degli ensemble $\mathcal{E}_{LDPC}(c, d, N)$ con decodifica ML su ogni canale $\mathbb{Z}_m$–simmetrico $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathbb{Z}_m}$, la cui distribuzione di probabilità di transizione $W = W(\cdot|0) \in \mathcal{P}(\mathcal{Y})$ abbia entropia

$$H(W) \leq \varepsilon \ ,$$

soddisfa le stime asintotiche (4.38) e (4.39).

Dimostrazione
Consideriamo un canale $\mathbb{Z}_m$–simmetrico di probabilità di transizione $W^* \in \mathcal{P}(\mathcal{Y})$ tale che

$$W^*(y^*) = 1 \ , \qquad W(y) = 0 \ , \ \ \forall y \in \mathcal{Y} \setminus \{y^*\} \ , \tag{4.51}$$

per un $y^* \in \mathcal{Y}$ fissato. Si evidentemente $H(W^*) = 0$; ciascun sottocanale $\{W^*(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \frac{m}{l}\mathbb{Z}_m}$ ha dunque capacità

$$C_l^* = \log l \quad ,$$

ed esponente di errore

$$E_l^*(R) = \log l - R \quad .$$

Si ripetano i medesimi ragionamenti della dimostrazione del Teorema 30, fino alla (4.49). Ragioniamo poi nella maniera seguente. Si fissi un $\delta > 0$ tale che, per ogni $l \mid m$, $l > 1$, siano soddisfatte contemporaneamente la (4.44), la (4.48) e la (4.49). Poichè $E_l(R_l, W)$ dipende con continuità da W, esiste un intorno di W^* in $\mathcal{P}(\mathbb{Z}_m)$ di raggio $\tilde{\varepsilon} > 0$, $B(W, \tilde{\varepsilon})$, sul quale la (4.49) è ancora soddisfatta (e ovviamente lo sono anche la (4.44) e la (4.48) che non dipendono da W). Per ogni $W \in B(W^*, \tilde{\varepsilon})$ abbiamo dunque che valgono le stime asintotiche (4.38) e (4.39). Ma, come detto, la funzione entropia H è continua su $\mathcal{P}(\mathcal{Y})$, e si annulla esattamente sull'insieme $\{W^* \ , \ y^* \in \mathcal{Y}\}$. Esiste dunque ε tale che

$$||W - W^*||_1 \geq \tilde{\varepsilon} \ , \ \ \forall y^* \in \mathcal{Y} \ \Rightarrow \ H(W) \geq \varepsilon \ ,$$

da cui segue la tesi. ■

Mostriamo ora che la probabilità media di errore sull'ensemble dei codici low–density ha esattamente l'andamento asintotico individuato dai teoremi precedenti; non c'è quindi una convergenza a 0 esponenzialmente in N, cosa che invece abbiamo visto accadere nel capitolo precedente per gli ensemble classici di codici $\mathbb{Z}_m$–lineari. Gli ordini di convergenza a zero presentati nei due teoremi precedenti non sono quindi dovuti ad una stima grossolana, ma piuttosto ad una debolezza dell'ensemble $\mathcal{E}_{LDPC}(c, d)$ così come è stato definito. Questa debolezza è dovuta alla presenza di una piccola percentuale di codici con alta probabilità di errore; si tratta di quei codici che contengono parole con un alto numero di zeri, cioè dei codici di distanza minima molto bassa. Nel paragrafo successivo l'ensemble $\mathcal{E}_{LDPC}(c, d)$ verrà modificato con l'eliminazione di tali codici in modo tale da ottenere probabilità di errore esponenzialmente decrescente in N.

Premettiamo una considerazione. Introduciamo delle notazioni: dato un canale $\mathbb{Z}_m$–simmetrico $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathbb{Z}_m}$, indichiamo con

$$W(0 \to a) := \sum_{\substack{y \in \mathcal{Y}: \\ W(y|a) \geq W(y|0)}} W(y|0)$$

$$W_{min} = \min_{a \in \mathbb{Z}_m \setminus \{0\}} W(0 \to a) \quad .$$

Diciamo che $\{W(\cdot|x) \in \mathcal{P}(\mathcal{Y})\}_{x \in \mathbb{Z}_m}$ è deterministico se

$$W_{\min} = 0 \quad .$$

Su un canale deterministico, la probabilità di errore con decodifica ML di qualsiasi codice a blocco non degenere è 0. Non è possibile dunque trovare delle stime dal basso alla probabilità di errore di ensemble di codici su tali canali. Introduciamo il concetto di distanza minima di un codice $\mathbb{Z}_m$–lineare; si tratta di un parametro molto importante per lo studio delle prestazioni con decodifica ML. Per ogni $\mathbf{x} \in \mathbb{Z}_m^N$ indichiamo con $w(\mathbf{x})$ il suo peso di Hamming, i.e.

$$w(\mathbf{x}) := N(1 - \theta_0(\mathbf{x})).$$

Definiamo poi la distanza minima di un codice $\mathbb{Z}_m$–lineare non degenere $\mathcal{C}$ come il minimo peso delle sue parole non nulle:

$$d_{\min}(\mathcal{C}) := \min_{\substack{\mathbf{x} \in \mathcal{C}: \\ \mathbf{x} \neq 0}} w(\mathbf{x}).$$

Lemma 32
Sia dato un canale $\mathbb{Z}_m$–simmetrico. Sia $\mathcal{E}(R, N)$ un ensemble di codici a blocco di lunghezza N. Allora, per ogni $n \in \mathbb{N}$ si ha

$$\overline{P(e)} \geq W_{\min}^n \mathbb{P}(d_{\min} \leq n). \tag{4.52}$$

Dimostrazione
Sia $\mathcal{C}$ un codice a blocco lineare di lunghezza $N \geq n$ e distanza minima $d_{min} = n$. La probabilità di errore di $\mathcal{C}$ è maggiore o uguale di una quantità strettamente positiva indipendente da N. Infatti, sia $\mathbf{x}_{\min}$ una parola di $\mathcal{C}$ di peso n, e siano $i_1, \ldots, i_n$ le posizioni dei suoi elementi non nulli.

$$\begin{aligned} P(e|\mathcal{C}) &= P(e|\mathcal{C}, 1) \\ &= \sum_{\mathbf{y} \in \mathcal{Y}} W_N(\mathbf{y}|\mathbf{0}) \\ &\geq \prod_{j=1}^{n} W(0 \to x_{i_j}) \\ &\geq W_{min}^n \quad . \end{aligned}$$

La probabilità media di errore su ciascun ensemble $\mathcal{E}_N$ soddisfa dunque

$$\begin{aligned} \overline{P(e)} &= \sum_{\substack{\mathcal{C} \text{ t.c.} \\ d_{\min}\mathcal{C} \leq n}} P(e|\mathcal{C})\mathbb{P}(\mathcal{C}) + \sum_{\substack{\mathcal{C} \text{ t.c.} \\ d_{\min}\mathcal{C} > n}} P(e|\mathcal{C})\mathbb{P}(\mathcal{C}) \\ &\geq \sum_{\substack{\mathcal{C} \text{ t.c.} \\ d_{\min}(\mathcal{C}) \leq n}} P(e|\mathcal{C})\mathbb{P}(\mathcal{C}) \\ &\geq \mathbb{P}(d_{\min} \leq n) W_{min}^n \, . \end{aligned}$$

■

Si osservi che il Lemma 32 è del tutto inutile se il canale considerato è deterministico. Se invece il canale è non deterministico, tale lemma garantisce che per ogni $n \in \mathbb{N}$ fissato si abbia

$$\mathbb{P}(d_{min} \leq n) = O(\overline{P(e)}) \quad N \to +\infty \quad ,$$

cioè che l'andamento asintotico di $\overline{P(e)}$ possa essere stimato dal basso da quello di $\mathbb{P}(d_{min} \leq n)$.

Teorema 33
Fissati N, $3 \leq c < d$ ed un canale $\mathbb{Z}_m$–simmetrico non deterministico, esiste una costante $K > 0$ tale che la probabilità media di errore dell'ensemble $\mathcal{E}_{LDPC}(c, d, N)$ soddisfa

$$\overline{P(e)} \geq KN^{2-c} \quad . \tag{4.53}$$

Dimostrazione
Stimiamo la probabilità che $d_{min} \leq 2$ per poi applicare la (4.52).

Ragioniamo sul grafo di Tanner $\mathcal{G} = (\mathcal{N} \cup \mathcal{M}, \mathcal{E})$. Fissati due variable nodes $v_i, v_j \in \mathcal{N}$ con $i < j$, sia

$$E_{i,j} := \{\mathcal{M}(i) = \mathcal{M}(j)\}$$

l'evento 'v_i, v_j hanno vicinato identico'. Indichiamo con $p' = \mathbb{P}(E_{i,j})$. Se si verifica E_{ij}, allora per ogni $a \in \mathbb{Z}_m$ la N–upla

$$\mathbf{x} \in \mathbb{Z}_m^N \text{ t.c.} \qquad x_i = a,\ x_j = -a,\ x_n = 0\ \forall k \notin \{i, j\}$$

è sicuramente una parola di codice. Dunque, se si verifica l'evento $\bigcup_{\substack{i,j=1 \\ i>j}}^{N} E_{i,j}$, allora sicuramente $d_{\min}(\mathcal{C}) \leq 2$, e quindi

$$\mathbb{P}(d_{\min} \leq 2) \geq \mathbb{P}\left(\bigcup_{\substack{i,j=1 \\ i>j}}^{N} E_{i,j} \right).$$

Sia F_i l'evento 'esistono più di $d/2$ archi uscenti dal nodo v_i ed entranti in uno stesso $h_k \in \mathcal{M}$'. Poniamo $q = \mathbb{P}(F_i)$. Scelti $a \geq \lceil \frac{d}{2} \rceil$ archi uscenti da v_i, la probabilità che essi entrino tutti in un $h_k \in \mathcal{M}$ fissato è pari a $\binom{d}{a}\binom{dL}{a}^{-1}$. Con una stima di tipo union bound otteniamo dunque

$$q \leq L \sum_{a=\lceil \frac{d}{2} \rceil}^{c} \binom{c}{a}\binom{d}{a}\binom{dL}{a}^{-1} \to 0, \quad N \to +\infty \quad .$$

La probabilità di $E_{i,j}$ condizionata al verificarsi di F_i è nulla perché se sia v_i che v_j avessero ciascuno più di $d/2$ archi incidenti su uno stesso h_k, allora il numero

di archi entranti in h_k sarebbe maggiore di d. Stimiamo ora la probabilità di $E_{i,j}$, condizionata al fatto che non si sia verificato l'evento F_i. Fissato $\mathcal{M}(i) \in F_i^c$, tra tutte le possibili $\binom{(N-1)c}{c}$ scelte di $\mathcal{M}(j)$ ce n'è almeno una tale che $\mathcal{M}(j) = \mathcal{M}(i)$, e dunque

$$\mathbb{P}(E_{i,j}|F_i^c) \geq \binom{(N-1)c}{c}^{-1} \geq \frac{1}{(c(N-1))^c} \quad .$$

Abbiamo quindi che

$$p' = \mathbb{P}(E_{i,j}) = \mathbb{P}(E_{i,j}|F_i^c)(1-q) \geq K' \frac{1}{N^c}. \quad , \tag{4.54}$$

con $K' > 0$.

Fissati tre diversi variable nodes $v_i, v_j, v_k \in \mathcal{N}$ con $i < j < k$, sia ora

$$E_{i,j,k} = \{\mathcal{M}(i) = \mathcal{M}(j) = \mathcal{M}(k)\}$$

l'evento 'v_i ,v_j e v_k hanno vicinato identico' e sia p'' la sua probabilità. Una volta fissato arbitrariamente $\mathcal{M}(i)$, tra le possibili $\binom{(N-1)c}{c}$ scelte vincolate di $\mathcal{M}(j)$, ce ne sono al più $(d-1)^c c!$ tali che $\mathcal{M}(j) = \mathcal{M}(i)$. Una volta fissati $\mathcal{M}(i)$ e $\mathcal{M}(j)$ tali che $\mathcal{M}(i) = \mathcal{M}(j)$, tra le possibili $\binom{(N-2)c}{c}$ scelte vincolate di $\mathcal{M}(k)$, ce ne sono al più $(d-2)^c c!$ tali che $\mathcal{M}(j) = \mathcal{M}(i) = \mathcal{M}(k)$. Si ha dunque la seguente stima

$$p'' = \mathbb{P}(E_{i,j,k}) \leq (d-1)^c (d-2)^c (c!)^2 \binom{(N-1)c}{c}^{-1} \binom{(N-2)c}{c}^{-1} \leq K'' N^{-2c} \quad , \tag{4.55}$$

con $K'' > 0$. Analogamente, fissati quattro diversi variable nodes $v_i, v_j, v_k, v_l \in \mathcal{N}$, con $i < j < k < l$, sia

$$E_{i,j,k,l} = \{\mathcal{M}(i) = \mathcal{M}(j) = \mathcal{M}(k) = \mathcal{M}(l)\}$$

l'evento 'v_i ,v_j, v_k e v_l hanno vicinato identico' e sia p''' la sua probabilità. Con ragionamenti simili a quelli che hanno condotto alla (4.55), si ottiene che esiste $K''' > 0$ tale che

$$p''' = \mathbb{P}(E_{i,j,k,l}) \leq (d-1)^c (d-2)^c (d-3)^c (c!)^3 \binom{(N-1)c}{3c}^{-1} \leq K''' N^{-3c} \quad . \tag{4.56}$$

La probabilità dell'evento $\bigcup_{\substack{i,j=1 \\ i>j}}^{N} E_{\{i,j\}}$ può essere stimata dal basso con una stima union/intersection bound. Si osservi che l'intersezione di due eventi distinti $E_{i,j}$ e $E_{k,l}$ è un evento di tipo $E_{i,j,k}$ oppure $E_{i,j,l}$ se rispettivamente $i = k$ oppure

$j = l$, oppure di tipo $E_{i,j,k,l}$ se $i \neq k$ e $j \neq l$. Abbiamo dunque

$$\begin{aligned}
\overline{P(e)} &\geq W_{\min}^2 \mathbb{P}(d_{\min} \leq 2) \\
&\geq W_{\min}^2 \mathbb{P}\Big(\bigcup_{\substack{i,j=1 \\ i>j}}^{N} E_{\{i,j\}} \Big) \\
&\geq W_{\min}^2 \left(\sum_{\substack{i,j=1 \\ i>j}}^{N} \mathbb{P}(E_{\{i,j\}}) - \sum_{\substack{i,j,k=1 \\ i>j>k}}^{N} \mathbb{P}(E_{\{i,j,k\}}) - \sum_{\substack{i,j,k,l=1 \\ i>j>k>l}}^{N} \mathbb{P}(E_{\{i,j,k,l\}}) \right) \\
&\geq W_{\min}^2 \left(\tbinom{N}{2} p' - \tbinom{N}{3} p'' - \tbinom{N}{4} p''' \right) \\
&\geq W_{\min}^2 \left(K' N^{(2-c)} - K'' N^{(3-2c)} - K''' N^{(4-3c)} \right) \\
&\geq K N^{(2-c)} \quad ,
\end{aligned}$$

dove $K > 0$ segue dal fatto che, per ipotesi, $W_{\min} > 0$. ■

Possiamo dimostrare qualcosa di più se facciamo l'ipotesi che m e c non sono primi tra loro.

Teorema 34
Si supponga che c e m non siano primi tra loro, i.e. $MCD(m, c) > 1$; si definisca

$$a := \mathrm{mfpc}(c, m) \quad .$$

Allora

$$\overline{P(e)} \geq K N^{1 - \frac{a-1}{a} c} \quad .$$

Dimostrazione
Studiamo la probabilità di avere $d_{\min} = 1$, per poter applicare poi la (4.52).

Sia $\mathcal{C}$ un codice il cui grafo di Tanner contenga un variable node $v_n \in \mathcal{N}$ il cui vicinato $\mathcal{M}(n)$ ha la caratteristica seguente: la molteplicità di ogni $h_m \in \mathcal{M}$ in $\mathcal{M}(n)$ è un multiplo di a. Questo corrisponde ad avere gli archi uscenti da v_n paralleli a gruppi di a. Un codice $\mathcal{C}$ il cui grafo di Tanner ha questa proprietà, ha distanza minima uguale a 1. Infatti la N–upla $\mathbf{x} \in \mathbb{Z}_m^N$, definita da

$$x_n = \frac{m}{a}, \qquad x_i = 0 \ \forall i \neq n$$

è sicuramente in $\mathcal{C}$.

Stimiamo dunque la probabilià che $\mathcal{C}$ abbia un siffatto grafo di Tanner. Fissato un variable node $v_n \in \mathcal{N}$, sia E_n l'evento 'v_n ha un numero x di archi verso ciascun

check node $h \in \mathcal{M}$ tale che x sia un multiplo di a'. Poniamo $p := \mathbb{P}(E_n)$. Possiamo stimare p come segue: ci sono

$$\binom{Ld}{c}$$

possibili vicinati $\mathcal{M}(n)$ di v_n. Di questi, ce ne sono almeno $\binom{L}{c/a}$ tali che tutti i check nodes in $\mathcal{M}(n)$ abbiano molteplicità esattamente a. Dunque

$$p \geq \frac{\binom{L}{c/a}}{\binom{Ld}{c}} \geq K' N^{-\frac{a-1}{a}c} \quad ,$$

dove $K' > 0$, costante in N.

Fissati due variable nodes distinti $v_i, v_j \in \mathcal{N}$, sia $E_{ij} := E_i \cap E_j$, e sia $q = \mathbb{P}(E_{ij})$. Sia F_{ij} l'evento 'l'unione dei vicinati di v_i e v_j, i.e. $\mathcal{M}(n) \cup \mathcal{M}(k)$ ha solo elementi di molteplicità multiplo di a'. Si ha evidentemente che $F_{ij} \subseteq E_{ij}$. Si può quindi stimare q dal basso con $\mathbb{P}(F_{ij})$ nel modo seguente. Il numero totale di possibili unioni di due vicinati è

$$\binom{Nc}{2c}.$$

Se $\mathcal{M}(n) \cup \mathcal{M}(k)$ ha solo elementi di molteplicità multiplo di a, ci sono al più $2c/a$ check nodes raggiunti da archi partenti da v_n e v_k, quindi il numero di unioni di vicinati in F_{ij} è al più

$$\binom{L}{2c/a}\binom{d2c/a}{2c}.$$

Abbiamo quindi che

$$q \leq K'' N^{-2c/a}$$

con $K'' > 0$ costante in N.

Utilizzando lo union/intersection bound abbiamo la seguente stima

$$\begin{aligned}
\overline{P(e)} &\geq W_{\min}\mathbb{P}(d_{\min} = 1) \\
&\geq W_{\min}\mathbb{P}\Big(\bigcup_{n=1}^{N} E_n\Big) \\
&\geq W_{\min}\Big(\sum_{n=1}^{N}\mathbb{P}(E_n) - \sum_{i=1}^{N}\sum_{j=i+1}^{N}\mathbb{P}(E_{ij})\Big) \\
&= W_{\min}\left(Np - \tbinom{N}{2}q^2\right) \\
&\geq W_{\min}\left(K' N^{1-(a-1)c/a} + K'' N^{2(1-c/a)}\right) \\
&\geq K N^{1-\frac{a-1}{a}c} \quad ,
\end{aligned}$$

dove $K > 0$ segue dal fatto che, per ipotesi, $W_{\min} > 0$. ■

Abbiamo trovato dunque gli esatti andamenti asintotici di $\overline{P(e)}$.

Corollario 35

- *Se* $MCD(m, c) > 1$*, allora*

$$\overline{P(e)} \asymp N^{1-\frac{a-1}{a}c}$$

  dove $a = \mathrm{mfpc}(m, c)$*;*

- *se* a *e* m *sono primi tra loro, allora*

$$\overline{P(e)} \asymp N^{2-c}.$$

4.6 "Expurgation is possible"

Il comportamento asintotico della probabilità media di errore degli ensemble di codici a bassa densità è dominato, come si è mostrato nel paragrafo precedente, da termini polinomiali in N. Quello che succede è che un insieme di codici di probabilità tendente a 0 ha prestazioni molto peggiori di quelle del resto dei codici, che hanno invece comportamento esponenzialmente decrescente. Scopo di questo ultimo paragrafo è mettere in luce questo fenomeno.

L'idea fondamentale delle tecniche di *expurgation* è quella di definire, a partire dall'ensemble $\mathcal{E}_{LDPC}(c, d, N)$, un nuovo ensemble di codici low–density $\mathcal{E}^*_{LDPC}(c, d, N)$, eliminando dal precedente i codici di prestazioni peggiori e rinormalizzando la probabilità. Dalle dimostrazioni dei precedenti teoremi si evince che il termine decrescente non esponenzialmente a zero è dovuto alle parole in $\mathcal{T}^N_{\delta,1,0}$, i.e. quelle di peso basso. Nel seguito mostreremo che, fissato $\delta > 0$ sufficientemente piccolo, la probabilità un codice $\mathcal{C}$ nell'ensemble $\mathcal{E}_{LDPC}(c, d, N)$ contenga parole in $\mathcal{T}^N_{\delta,1}$ tende a zero per N tendente a $+\infty$, quindi l'eliminazione dei codici che contengono tali parole dall'ensemble aumenta la probabilità degli altri di un fattore che è arbitrariamente vicino a 1 per N sufficientemente grande. Questo ci permetterà di dimostrare poi che le prestazioni degli ensemble espurgati di codici a bassa densità $\mathcal{E}^x_{LDPC}(c, d)$, hanno prestazioni asintotiche che possono essere rese arbitrariamente vicine a quelle degli ensemble classici di codici $\mathbb{Z}_m$–lineari, ricavate nel capitolo 3. Questo risultato contraddice l'affermazione 'expurgation is impossible' contenuta in [4] a pag. 437. In effetti, gli autori di questo articolo alludevano all'impossibilità di espurgare gli ensemble codici a bassa densità per ottenere i medesimi andamenti asintotici, in termini di esponente di errore, del random coding ensemble. Ma questo non accade, come abbiamo mostrato nel capitolo 3, nemmeno per l'ensemble $\mathcal{E}_{\mathrm{Ker}}$ dei codici nucleo di omomorfismi. È con i comportamenti asintotici di questi ensemble, dettati dalla (3.12) del

Teorema 12, che bisogna confrontare gli andamenti asintotici della probabilità media di errore degli ensemble $\mathcal{E}_{LDPC}(c,d,N)$. Mostreremo nel seguito che, con probabilità tendente a 1, le prestazioni medie degli ensemble di codici a bassa densità sono arbitrariamente vicine a quelle degli ensemble classici di codici $\mathbb{Z}_m$–lineari.

Introduciamo le notazioni seguenti. Fissati c,d,N sia $\mathcal{E}_{LDPC}(c,d,N)$ l'ensemble di codici a bassa densità non espurgato, i.e. quello definito nel paragrafo 4.1 e studiato finora in questo capitolo. Nel seguito continueremo ad usare le notazioni $\mathbb{P}(\cdot)$, $\mathbb{E}[\cdot]$, $\overline{P(e)}$ per indicare rispettivamente la probabilità, il valore atteso e la probabilità media di errore rispetto all'ensemble $\mathcal{E}_{LDPC}(c,d,N)$. Dato $\delta \in [0,1]$, definiamo l'evento Ex 'il codice contiene parole in $\mathcal{T}_{\delta,1}^N$ oltre alla parola $\mathbf{0}$', i.e.

$$Ex := \{\mathcal{C} : \mathcal{C} \cap \mathcal{T}_{\delta,1}^N \neq \{\mathbf{0}\}\} \ .$$

Definiamo l'ensemble $\mathcal{E}_{LDPC}^x(c,d,N)$ come l'ensemble $\mathcal{E}_{LDPC}(c,d,N)$ condizionato all'evento Ex^c, i.e. ponendo per ogni codice a blocco $\mathcal{C}$ su $\mathbb{Z}_m$ di lunghezza N e rate maggiore o uguale a $(1-\frac{c}{d})\log m$

$$\mathbb{P}^x(\mathcal{C}) := \frac{\mathbb{P}(\mathcal{C})}{1-\mathbb{P}(Ex)} \mathbb{1}_{(Ex)^c}(\mathcal{C}) \ .$$

Useremo le notazioni $\mathbb{P}^x(\cdot)$, $\mathbb{E}^x[\cdot]$, $\overline{P(e)}^x$ per indicare rispettivamente la probabilità, il valore atteso e la probabilità media di errore rispetto all'ensemble $\mathcal{E}_{LDPC}^x(c,d,N)$. Si osservi che

$$\mathbb{P}^x(\mathcal{C}) \leq \frac{\mathbb{P}(\mathcal{C})}{1-\mathbb{P}(Ex)} \ , \quad \forall \mathcal{C} \ .$$

Lemma 36
Dati $c,d \in \mathbb{N}$ tali che

$$3 \leq c < d, \quad \mathrm{MCD}(d,m) = 1 \ ,$$

si considerino gli ensemble di codici $\mathcal{E}_{LDPC}(c,d,N)$. Allora esiste $\gamma > 0$ tale che, per ogni $\delta < \gamma$

$$\mathbb{P}(\mathcal{C} \cap \mathcal{T}_{\delta,1}^N \neq \{\mathbf{0}\}) \to 0 \ .$$

Dimostrazione
Applichiamo una stima union bound, ricordando che $\mathbb{P}(\mathbf{x} \in \mathcal{C})$ è funzione solo del tipo $\boldsymbol{\theta}(\mathbf{x})$:

$$\mathbb{P}(\mathcal{C} \cap \mathcal{T}_{\delta,1}^N \neq \{\mathbf{0}\}) \leq \sum_{\boldsymbol{\theta} \in J_{\delta,1}^N \setminus \{\boldsymbol{\theta}(\mathbf{0})\}} \sum_{\mathbf{x} \in \mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N} \mathbb{P}(\mathbf{x} \in \mathcal{C}) = P_I + P_{II} + P_{III}$$

dove,

$$P_I := \sum_{a\in\mathbb{Z}_m} \sum_{\substack{\mathbf{x}\in\mathbb{Z}_m^N:\\ 1-\beta<\theta_a(\mathbf{x})<1}} \mathbb{P}(\mathbf{x}\in\mathcal{C})$$

$$P_{II} := \sum_{a\in\mathbb{Z}_m} \sum_{\substack{\mathbf{x}\in\mathbb{Z}_m^N:\\ 1-\delta\le\theta_a(\mathbf{x})\le1-\beta}} \mathbb{P}(\mathbf{x}\in\mathcal{C})$$

$$P_{III} := \sum_{a\in\mathbb{Z}_m\setminus\{0\}} \mathbb{P}(\mathbf{a}\in\mathcal{C}) \ ,$$

$$\beta := \frac{1}{d}e^{-12}\exp\left(-\left[\frac{2}{c-2}\log\left((m-1)\frac{d}{2}\right)\right]\right) \quad .$$

Usando le stime di $\mathbb{P}(\mathbf{x}\in\mathcal{C})$ fornite dal Lemma 27 e dal Lemma 29 (si osservi che le ipotesi (4.27) e (4.34) sono soddisfatte grazie al fatto che $\beta \le \frac{1}{d}$ per ogni scelta di c e d) e ripetendo i ragionamenti della dimostrazione del teorema 30, si ottiene che

$$\lim_{N\to+\infty} P_I = 0 \quad .$$

Usando il Corollario 25, e ripetendo i medesimi ragionamenti della dimostrazione del Teorema 30, si ottiene che

$$\tilde{f}(\delta)+\frac{c}{d}\left[\log\left(m-\varphi(m)+\varphi(m)\exp\left[-\frac{\lambda}{2e^{12}}\exp\left(-\left[\frac{2\log\left((m-1)d/2\right)}{c-2}\right]\right)\right]\right)\right] < 0 \tag{4.57}$$

è condizione sufficiente affinché P_{II} abbia andamento esponenzialmente decrescente a 0 in N, per N sufficientemente grande.

Infine, dal fatto che, come già osservato nella dimostrazione del Lemma 29, si ha

$$\mathbf{a}\in\mathcal{C} \Leftrightarrow m \mid da \ ,$$

possiamo concludere che $P_{III} = 0$ ogni volta che d è primo con m.

Definiamo $\gamma(c,d)$ come

$$\gamma(c,d) := \sup\{\delta\in[0,\frac{1}{2}) \text{ t.c. } (4.57)\} \ ; \tag{4.58}$$

poiché $\tilde{f}(\delta)\to 0$ per $\delta\to 0$, esiste $\delta > 0$ tale che la (4.57) sia soddisfatta. Ne segue che $\gamma > 0$. ■

Teorema 37
Si consideri un canale $\mathbb{Z}_m$–simmetrico fissato di $\mathbb{Z}_m$–capacità $\hat{C}$; si supponga di utilizzare decodifica ML. Sia assegnato un rate di progetto R tale che

$$0 \le R < \hat{C} \quad .$$

Per ogni $\epsilon > 0$, esistono $\delta > 0$ e $c, d \in \mathbb{N}$ tali che

$$3 \leq c < d \ , \tag{4.59}$$

gli ensemble δ–espurgati $\mathcal{E}^x_{LDPC}(c, d, N)$ abbiano probabilità media di errore su un canale $\mathbb{Z}_m$–simmetrico soddisfacente la stima

$$\overline{P(e)}^x \leq m \sum_{\substack{l|m \\ l>1}} \exp(-N\left[E_l(R_l) - \epsilon\right]) \ , \tag{4.60}$$

dove $E_l(R)$ sono gli esponenti di errore dei sottocanali e $R_l := \frac{\log l}{\log m} R$ i loro rate di utilizzo.

Dimostrazione
Poiché ciascun $E_l(R)$ è funzione continua di R, esiste $\tilde{\varepsilon} > 0$ tale che

$$E_l(R_l + \tilde{\varepsilon}) > E_l(R_l) - \frac{\varepsilon}{2}$$

per ogni $l \mid m$, $l > 1$. È sempre possibile trovare c^* e d^* soddisfacenti la (4.59) e tali che

$$\frac{R_0}{\log m} \leq 1 - \frac{c^*}{d^*} < \frac{R_0}{\log m} + \frac{\tilde{\varepsilon}}{2} \ .$$

Applichiamo ora il bound (4.6) con $\delta < \gamma$ che verrà specificato in seguito; si ottiene

$$\overline{P(e)}^x \leq P_I^x + P_{II}^x$$

dove

$$P_I^x := \sum_{a \in \mathbb{Z}_m} \sum_{\substack{\boldsymbol{\theta} \in \mathbb{P}_N(\mathbb{Z}_m): \\ 1-\delta \leq \theta_a \leq 1}} \overline{S(\boldsymbol{\theta})}^x \mathbf{D}^{N\boldsymbol{\theta}} = 0 \ ;$$

$$P_{II}^x := m \sum_{\substack{l \mid m \\ l > 1}} \exp\left(-N\left[E_l\left(\log l\left(1 - \frac{c^*}{d^*}\right) + \frac{\log \alpha_l^x}{N}\right) - f(\delta)\right]\right) \ . \tag{4.61}$$

Ma

$$\begin{aligned} \overline{S(\boldsymbol{\theta})}^x &= \sum_{\mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N} \mathbb{P}^x(\mathbf{x} \in \mathcal{C}) \\ &\leq \frac{1}{1 - \mathbb{P}(\mathcal{C} \cap \mathcal{T}_{\delta,1}^N \neq \{\mathbf{0}\})} \sum_{\mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N} \mathbb{P}(\mathbf{x} \in \mathcal{C}) \\ &= p(N)\overline{S(\boldsymbol{\theta})} \ , \end{aligned}$$

dove, per il Lemma 36,

$$p(N) := \frac{1}{1 - \mathbb{P}(\mathcal{C} \cap \mathcal{T}_{\delta,1}^N \neq \{\mathbf{0}\})} \longrightarrow 1 \ , \quad N \to +\infty \ .$$

Si ha dunque che

$$\begin{aligned}\alpha_l^x &= \max_{\boldsymbol{\theta}\in J_{\delta,l}^N} \frac{\overline{S(\boldsymbol{\theta})}^x}{(\frac{1}{l})^L} \\ &\leq \max_{\boldsymbol{\theta}\in J_{\delta,l}^N} \frac{\overline{S(\boldsymbol{\theta})}p(N)}{(\frac{1}{l})^L} \\ &\leq p(N)(cN+1)^m \frac{\left(\frac{1}{l}+(1+\frac{1}{l})B(\delta)\right)^L}{(\frac{1}{l})^L}\end{aligned}$$

e quindi

$$\frac{\log \alpha_l^x}{N} \leq \frac{c}{d}\log(1+(l-1)B(\delta)) + o(1)\ , \qquad N \to +\infty.$$

Scegliamo dunque $\delta > 0$ tale che

$$f(\delta) < \frac{\varepsilon}{2}\ .$$

Poniamo poi $d = d^*$, $c = c^*$. Poiché $B(\delta) \to 0$ per $d \to +\infty$, si ha che esistono $N_0, k_0 \in \mathbb{N}$ tali che, per ogni $k \geq k_0$ e $N \geq N_0$,

$$\frac{\log \alpha_l^x}{N} < \frac{\tilde{\varepsilon}}{2}\ .$$

Si osservi che, con tale scelta, il primo membro della (4.57) è decrescente in k, e quindi la (4.57) continua ad essere soddisfatta. Se scegliamo k come il più piccolo intero primo con m e maggiore o uguale a k_0, abbiamo dunque che

$$\begin{aligned}E_l\left(\log l\left(1-\frac{c}{d}\right)+\frac{\log \alpha_l^x}{N}\right) - f(\delta) &\geq E_l\left(R_l + \frac{\tilde{\varepsilon}}{2} + \frac{\tilde{\varepsilon}}{2}\right) - \frac{\varepsilon}{2} \\ &\geq E_l(R_l) - \varepsilon\end{aligned}$$

e, sostituendo nella (4.61), otteniamo la tesi. ■

Conclusioni

In questa tesi abbiamo studiato le prestazioni di alcuni ensemble di codici a blocco $\mathbb{Z}_m$–lineari per modulazioni non binarie.

Uno dei nostri risultati fondamentali è il Teorema 19 del Capitolo 3 dove mostra che nel caso di modulazione m–PSK co $m = 2^r$ tali codici permettono di raggiungere capacità. In generale per altre costellazioni con tali codici si può raggiungere soltanto una capacità inferiore.

Si sono poi considerati codici a bassa densità su $\mathbb{Z}_m$ e abbiamo ottenuto, per ensemble di questi, stime delle probabilità medie di errore: Teoremi 30, 31 e 37. Questi risultati mostrano come in particolare con tali codici si raggiungano le stesse prestazioni che dell'ensemble di tutti i lineari.

Alcuni fondamentali problemi che rimangono aperti sono i seguenti.

- I risultati del Capitolo 3 mostrano come, per per alcune costellazioni geometricamente uniformi che ammettono gruppo generatore $\mathbb{Z}_m$, gli $\mathbb{Z}_m$–moduli liberi sono sufficienti per raggiungere capacità. Per altre sono necessari codici non liberi; per altre ancora neppure i moduli non liberi sembrano essere sufficienti. Rimane un problema aperto studiare quali siano le costellazioni per le quali i codici lineari su $\mathbb{Z}_m$ permettano di raggiungere capacità, e se in ogni caso una costellazione GU ammetta sempre codici GU che permettano di raggiungere capacità.

- I codici a bassa densità studiati sono di tipo molto semplice in quanto generati da matrici di parità con soli 0 e 1: rimane da chiarire se l'utilizzo degli altri elementi invertibili di $\mathbb{Z}_m$ possa migliorare sensibilmente le prestazioni di questi codici. Inoltre vanno considerati codici irregolari, cioè con grafo di Tanner con gradi dei vertici variabili, come è stato fatto nel caso binario.

- Sarà importante infine individuare criteri di progetto per i codici a bassa densità che tengano anche conto dell'algoritmo di decodifica utilizzato. Per questo sarà importante affiancare allo studio teorico opportune simulazioni.

Appendice A

A.1 Notazioni di teoria dei gruppi

Richiamiamo alcuni termini e notazioni di teoria dei gruppi usati in questa tesi. Un gruppo è un insieme G con una legge di composizione associativa, dotato di un elemento identico g_0, e tale che ogni elemento $g \in G$ possieda un inverso.

I gruppi G che prenderemo in considerazione in questa tesi sono gruppi di trasformazioni di un insieme $\mathcal{X}$, i.e. gli elementi di G sono funzioni invertibili di $\mathcal{X}$ in sé, l'operazione di gruppo è la composizione di funzioni, e l'elemento neutro, che denotiamo con g_0 è la funzione identità.

Sia G è un gruppo di trasformazioni di un insieme $\mathcal{X}$: si dice in questo caso che $\mathcal{X}$ è un G–insieme. Fissato un $x \in \mathcal{X}$, l'orbita di x sotto G è l'insieme

$$O(x) = \{gx, g \in G\} \ ;$$

$\mathcal{X}$ può essere partizionato in orbite disgiunte. Fissato $x \in \mathcal{X}$, lo stabilizzatore di x in G è il sottogruppo di G delle trasformazioni che lasciano invariato x:

$$Stab(x) = \{g \in G \text{ t.c. } gx = x\} \ .$$

Quando $\mathcal{X} = O(x)$ per qualche $x \in \mathcal{X}$, si dice che l'azione di G su $\mathcal{X}$ è transitiva; se inoltre $Stab(\mathcal{X}) = \{g_0\}$ per ogni $x \in \mathcal{X}$ l'azione di G su $\mathcal{X}$ è detta semplicemente transitiva.

A.2 Il metodo dei tipi

Richiamiamo alcune notazioni del metodo dei tipi introdotte in [10]. Sia $\mathcal{X}$ un insieme finito. Data una sequenza $\mathbf{x} \in \mathcal{X}^N$, il tipo di $\mathbf{x}$ è la distribuzione di probabilità $\boldsymbol{\theta}(\mathbf{x}) \in \mathcal{P}(\mathcal{X})$ definita ponendo, per ogni $a \in \mathcal{X}$, $\theta_a(\mathbf{x})$ pari alla frequenza relativa di a in $\mathbf{x}$, i.e.

$$\theta_a(\mathbf{x}) = \frac{1}{N}|\{j : x_j = a\}|.$$

Il sottoinsieme di $\mathcal{P}(\mathcal{X})$ che comprende tutte i possibili tipi delle parole $\mathbf{x} \in \mathcal{X}^N$ viene indicato con $\mathcal{P}_N(\mathcal{X})$. $\mathcal{P}_N(\mathcal{X})$ è un insieme finito: la sua cardinalità è pari al numero di soluzioni $(y_1, \ldots, y_{|\mathcal{X}|}) \in \mathbb{N}^{|\mathcal{X}|}$ dell'equazione

$$y_1 + \ldots + y_{|\mathcal{X}|} = N \ ;$$

dunque si ha $|\mathcal{P}_N(\mathcal{X})| = \binom{N+|\mathcal{X}|-1}{|\mathcal{X}|-1}$. Inoltre $|\mathcal{P}_N(\mathcal{X})|$ può essere sovrastimato con una funzione polinomiale di N notando semplicemente che, per ogni $a \in \mathcal{X}$, θ_a può assumere al più $N+1$ valori; abbiamo

$$|\mathcal{P}_N(\mathcal{X})| = \binom{N + |\mathcal{X}| - 1}{|\mathcal{X}| - 1} \leq (N+1)^{|\mathcal{X}|}. \tag{A.1}$$

Fissato un tipo $\boldsymbol{\theta} \in \mathcal{P}_N(\mathcal{X})$, indichiamo con $\mathcal{T}^N_{\boldsymbol{\theta}}$ il sottoinsieme delle sequenze di $\mathcal{X}^N$ il cui tipo è $\boldsymbol{\theta}$, i.e.

$$\mathcal{T}^N_{\boldsymbol{\theta}} = \{\mathbf{x} \in \mathcal{X}^N \text{ t.c. } \boldsymbol{\theta}(\mathbf{x}) = \boldsymbol{\theta}\}.$$

L'insieme dei $\mathcal{T}^N_{\boldsymbol{\theta}}$ al variare di $\theta \in \mathcal{P}_N(\mathbb{Z}_m)$ costituisce una partizione di $\mathbb{Z}^N_m$. Si verifica che la cardinalità di $\mathcal{T}^N_{\boldsymbol{\theta}}$ è pari a $\binom{N}{N\boldsymbol{\theta}}$; la seguente proposizione indica che $\binom{N}{N\boldsymbol{\theta}}$ ha andamento asintoticamente esponenziale in N, con esponente pari ad $H(\boldsymbol{\theta})$.

Proposizione 38

$$\frac{1}{|\mathcal{P}_N(\mathcal{X})|} \exp(NH(\boldsymbol{\theta})) \leq \binom{N}{N\boldsymbol{\theta}} = |\mathcal{T}^N_{\boldsymbol{\theta}}| \leq \exp(NH(\boldsymbol{\theta})) \ . \tag{A.2}$$

Dimostrazione
Esistono essenzialmente due dimostrazioni di questo enunciato. La prima si basa sull'uso delle formule di Stirling per l'approssimazione del coefficiente fattoriale. La seconda, che è quella che proponiamo, si basa su ragionamenti probabilistici. Abbiamo detto che un tipo $\boldsymbol{\theta} \in \mathcal{P}_N(\mathcal{X})$ è una particolare distribuzione di probabilità su $\mathcal{X}$. Consideriamo allora una variabile aleatoria $\mathbf{X}$ su $\mathcal{X}^N$ di distribuzione

$$P_{\mathbf{X}}(\mathbf{x}) := \theta_{x_1} \ldots \theta_{x_N}.$$

Iniziamo dalla stima dall'alto. Calcoliamo la probabilità che $\mathbf{X}$ sia in $\mathcal{T}_{\boldsymbol{\theta}}^N$.

$$\begin{aligned} 1 &\geq \mathbb{P}(\mathbf{X} \in \mathcal{T}_{\boldsymbol{\theta}}^N) \\ &= \sum_{\mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N} P_{\mathbf{X}}(\mathbf{x}) \\ &= \sum_{\mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N} \theta_{x_1} \dots \theta_{x_N} \\ &= \sum_{\mathbf{x} \in \mathcal{T}_{\boldsymbol{\theta}}^N} \prod_{a \in \mathcal{X}} \theta_a^{N\theta_a} \\ &= |\mathcal{T}_{\boldsymbol{\theta}}^N| \exp(-NH(\boldsymbol{\theta})) \quad , \end{aligned}$$

da cui segue immediatamente $|\mathcal{T}_{\boldsymbol{\theta}}^N| \leq \exp(-NH(\boldsymbol{\theta}))$.

Per quanto riguarda la stima dal basso, mostriamo prima che, per ogni tipo $\hat{\boldsymbol{\theta}} \in \mathcal{P}_N(\mathcal{X})$ si ha

$$\mathbb{P}(\mathbf{X} \in \mathcal{T}_{\boldsymbol{\theta}}^N) \geq \mathbb{P}(\mathbf{X} \in \mathcal{T}_{\hat{\boldsymbol{\theta}}}^N) \ . \tag{A.3}$$

Infatti,

$$\begin{aligned} \frac{\mathbb{P}(\mathbf{X} \in \mathcal{T}_{\boldsymbol{\theta}}^N)}{\mathbb{P}(\mathbf{X} \in \mathcal{T}_{\hat{\boldsymbol{\theta}}}^N)} &= \frac{\binom{N}{N\boldsymbol{\theta}} \prod_{a \in \mathcal{X}} \theta_a^{N\theta_a}}{\binom{N}{N\hat{\boldsymbol{\theta}}} \prod_{a \in \mathcal{X}} \theta_a^{N\hat{\theta_a}}} \\ &= \prod_{a \in \mathcal{X}} \frac{(N\hat{\theta_a})!}{(N\theta_a)!} \theta_a^{N(\theta_a - \hat{\theta}_a)} \ . \end{aligned}$$

Usando la disuguaglianza $\frac{m!}{n!} \geq n^{m-n}$, valida per ogni $n, m \in \mathbb{N}$, si ottiene

$$\begin{aligned} \frac{\mathbb{P}(\mathbf{X} \in \mathcal{T}_{\boldsymbol{\theta}}^N)}{\mathbb{P}(\mathbf{X} \in \mathcal{T}_{\hat{\boldsymbol{\theta}}}^N)} &\geq \prod_{a \in \mathcal{X}} (N\theta_a)^{N\hat{\theta}_a - N\theta_a} \theta_a^{N(\theta_a - \hat{\theta}_a)} \\ &= \prod_{a \in \mathcal{X}} N^{N(\hat{\theta}_a - \theta_a)} \\ &= N^{N\sum_{a \in \mathcal{X}} \hat{\theta}_a - N\sum_{a \in \mathcal{X}} \theta_a} \\ &= 1 \end{aligned}$$

da cui segue la (A.3). Ma allora si ha che, usando la (A.1) e la (A.3)

$$\begin{aligned} 1 &= \sum_{\hat{\boldsymbol{\theta}} \in \mathcal{P}_N(\mathcal{X})} \mathbb{P}(\mathbf{X} \in \mathcal{T}_{\hat{\boldsymbol{\theta}}}^N) \\ &\leq |\mathcal{P}_N(\mathcal{X})| \max_{\hat{\boldsymbol{\theta}} \in \mathcal{P}_N(\mathcal{X})} \mathbb{P}(\mathbf{X} \in \mathcal{T}_{\hat{\boldsymbol{\theta}}}^N) \\ &\leq (N+1)^{|\mathcal{X}|} \binom{N}{N\boldsymbol{\theta}} \exp(-NH(\boldsymbol{\theta})) \end{aligned}$$

da cui segue la seconda parte della (A.2). ■

Infine, dato un insieme di distribuzioni $J \subseteq \mathcal{P}(\mathcal{X})$, indichiamo con $\mathcal{T}_J^N$ l'unione dei $\mathcal{T}_{\boldsymbol{\theta}}^N$ al variare di $\boldsymbol{\theta} \in J \cap \mathcal{P}_N(\mathcal{X})$, i.e.

$$\mathcal{T}_J^N := \{\mathbf{x} \in \mathcal{X}^N \text{ t.c. } \boldsymbol{\theta}(\mathbf{x}) \in J\};$$

chiaramente la cardinalità di T_J^N cresce esponenzialmente in N con esponente pari a

$$\sup_{\boldsymbol{\theta}\in J} H(\boldsymbol{\theta}).$$

A.3 Aritmetica

Introduciamo alcune notazioni usate nel capitolo 4.

Dati $n_1, n_2 \in \mathbb{N}$, indichiamo con $\mathrm{MCD}(n_1, n_2)$ il loro massimo comun divisore; se $\mathrm{MCD}(n_1, n_2) > 1$, cioè se n_1 ed n_2 non sono primi tra loro, indichiamo con $\mathrm{mfpc}(n_1, n_2)$ il minimo fattore primo comune di n_1 ed n_2.

Definiamo poi la funzione

$$\varphi : \mathbb{N} \to \mathbb{N}$$

chiamata funzione φ di Eulero e definita ponendo, per ogni $n \in \mathbb{N}$, $\varphi(n)$ pari al numero di interi minori o uguali a n primi con n. Vale l'enunciato seguente (di veda [1] per la dimostrazione).

Proposizione 39

$$\varphi(n) = \sum_{l|n} \varphi(l) \tag{A.4}$$

□

Infine, dimostriamo un risultato usato nelle dimostrazioni dei Teoremi 11 e 12 del Capitolo 3.

Proposizione 40
Sia $\mathbf{x} \in \mathbb{Z}_m^N$ *tale che*

$$\mathrm{MCD}(x_i, m) = l$$

e sia $\mathbf{Y}$ *una variabile aleatoria distribuita uniformemente su* $\mathbb{Z}_m^N$*. Allora la variabile aleatoria*

$$H := \mathbf{x} \cdot \mathbf{Y}$$

è uniformemente distribuita su $l\mathbb{Z}_m^N$*.*

Dimostrazione
Osserviamo, innanzitutto, che è sufficiente dimostrare con l'enunciato $l = 1$. Infatti, se

$$\mathrm{MCD}(x_j, m) = l \ ,$$

allora

$$\mathrm{MCD}\left(\frac{x_j}{l}, \frac{m}{l}\right) = 1 \ ,$$

e se $\frac{1}{l}\mathbf{x}\cdot\mathbf{Y}$ è distribuita uniformemente su $\mathbb{Z}_{\frac{m}{l}}^N$ allora $\mathbf{x}\cdot\mathbf{Y}$ è distribuita uniformemente su $l\mathbb{Z}_m^N$.

Supponiamo ora

$$m = p^r \ ,$$

con p primo. L'ipotesi MCD$(x_j, m) = 1$ implica in questo caso che esistea $j \in \{1, \ldots, N\}$ tale che $x_j = 1$. Ma allora, dalla formula della probabilità totale segue che, per ogni $a \in \mathbb{Z}_{p^r}$,

$$\begin{aligned}
\mathbb{P}(H = a) &= \sum_{b\in\mathbb{Z}_{p^r}} \mathbb{P}(H - Y_j = b | Y_j = a - b)\mathbb{P}(Y_j = a - b) \\
&= \frac{1}{p^r} \sum_{b\in\mathbb{Z}_{p^r}} \mathbb{P}(H - Y_j = b | Y_j = a - b) \\
&= \frac{1}{p^r} \ .
\end{aligned}$$

Per l'estensione al caso $m \in \mathbb{N}$ generico, si usa il teorema di decomposizione di Kronecker:

$$\mathbb{Z}_m \simeq \mathbb{Z}_{p_1^{r_1}} \times \ldots \times \mathbb{Z}_{p_n^{r_n}} \ ,$$

dove $m = p_1^{r_1} \ldots p_n^{r_n}$. ■

Bibliografia

[1] G. E. Andrews, *Number Theory*, Dover Books on Advanced Mathematics, New York, 1994.

[2] M. Artin, *Algebra*, Bollati Boringhieri, Torino, 1997.

[3] S. Benedetto, E. Biglieri, *Principles of Digital Transmission With Wireless Applications*, Kluwer Academic / Plenum Publishers, New York, 1999.

[4] A. Bennatan, D. Burshetein , "On The Application of LDPC Codes to Arbitrary Discrete Memoryless Channels", *IEEE Trans. Inf. Theory*, vol. 50, pp.417-438, Mar. 2004.

[5] C. Berrou, A. Glavieux, P. Thitimajshima, "Near Shannon limit error–correcting coding and decoding: Turbo–codes", *Proc. 1993 IEEE Trans. Int. Conf. on Comm.*, Geneva, Switzerland, pp.1064-1070, 1993.

[6] E. R. Berlekamp, R. J. McEliece, H. C. A. Van Tilborg, "On the Inherent Intractability of Certain Coding Problems", *IEEE Trans. Inf. Theory*, vol. IT-24, no 3, pp.384-386, May 1978.

[7] G. Caire, E. Biglieri, "Linear Block Codes Over Cyclic Groups", *IEEE Trans. Inf. Theory*, vol. 41, NO 5, pp.1246-1256, Sep. 1995.

[8] S.- Y. Chung, G.D. Forney, Jr, T. J. Richardson, R. Urbanke, "On The Design of Low–Density Parity–Check Codes within 0.0045 dB of the Shannon Limit", *IEEE Comm. Letters*, vol. 5, n. 2, pp.58-60, Feb. 2001

[9] T. M. , *Elements of Information Theory*, vol. 1, John Wiley and Sons, New York, 1991.

[10] I. Csiszar, "Te Method of Types", *IEEE Trans. Inf. Theory*, vol. 44, pp.2505-2523, Oct. 1998.

[11] M. C. Davey, D. J. C. MacKay, "Low density parity check over GF(q)", *IEEE Comm. Letters*, 2 (6), pp. 165-167, 1998.

[12] G. D. Forney, Jr., "Geometrically Uniform Codes", *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241-1260, 1991.

[13] R. G. Gallager, *Low Density Parity Check Codes*, MIT Press, Cambridge MA, 1963.

[14] R. G. Gallager, "A Simple Derivation of the Coding Theorem and Some Applications", *IEEE Trans. Inform. Theory*, IT-11, pagg.3-18, 1965.

[15] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

[16] J. Jacod, P. Protter, *Probability Essentials*, Springer, Berlin, 2000.

[17] F. Kschischang, B. Frey, H.-A. Loeliger, "Factor Graphs and the Sum–Product Algorithm", *IEEE Trans. Inform. Theory*, vol. 47, pp. 418-519, Feb. 2001.

[18] H.-A. Loeliger, "Signal Sets Matched To Groups", *IEEE Trans. Inform. Theory*, vol. 37, n. 6, pp. 1675-1679, Nov. 1991.

[19] M.-G. Luby, M. Mitzenmacher, M.-A. Shokrollahi, D.-A. Spielman, "Improved Low–Density Parity–Check Codes Using Irregular Graphs and Belief–Propagation", *IEEE Trans. Inform. Theory*, vol. 47, n. 2, pp. 585-598, Feb. 2001.

[20] D.J.C. MacKay, R,M. Neal, "Good codes based on very sparse matrices", *Criptography and Coding*, 5th IMA Conf., LNCS 1025, ed. by C. Boyd, pp. 100-111, Springer, 1995.

[21] D.J.C. MacKay, "Good Error Correcting Codes Based On Very Sparse Matrices", *IEEE Trans. Inf. Theory*, vol. 45, pp.399-431, Mar. 1999.

[22] P. E. McIllree, "Channel Capacity Calculations for M–ary N–Dimensional Signals Sets", *M. E. I. thesis*, Univ. of South Australia, Feb. 1995.

[23] G. Miller e D. Burshetein , "Bounds on the Maximum Likelihood Decoding Error Probability of Low–Density Parity–Check Codes", *IEEE Trans. Inf. Theory*, vol. 47, pp.2696-2710, Nov. 2001.

[24] C. E. Shannon, "A Mathematical Theory of Communication", *Bell Sys. Thec. Journal*, vol. 27, pp. 379-423, pp. 623-656, Jul. Oct. 1948.

[25] N. Shulman, M. Feder "Random Coding Techniques for Nonrandom Codes", *IEEE Trans. Inform. Theory*, vol. 45, NO.6, pp. 2001-2004, 1999.

[26] D.A. Spielman, "Linear-time encodable and decodable error–correcting codes", *IEEE Trans. Inform. Theory*, vol. 42, NO.6, pp. 1723-1731, Sept. 1996.

[27] D. Sridhara, T. E. Fuja "Low density parity check codes over groups and rings", *ITW2002*, Bangalore, India, Oct 20-25, 2002.

[28] D. Slepian, "On neighbor distances and symmetry in group codes", *IEEE Trans. Inform. Theory*, vol. 17, pp. 630-632, Set 1971.

[29] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Network of Plausible Inference*, Morgan Kaufmann, San Matteo, 1988.

[30] T. Richardson, R. Urbanke, "The Capacity of Low–density Parity Check Codes Under Message–Passing Decoding", *IEEE Trans. Inform. Theory*, vol. 47, pp. 599-618, Feb 2001.

[31] T. Richardson, A. Shokrollahi, R. Urbanke, "Design of Capacity–Approaching Irregular Low–density Parity Check Codes", *IEEE Trans. Inform. Theory*, vol. 47, pp. 619-637, Feb 2001.

[32] D. Slepian, "On neighbor distances and symmetry in group codes", *IEEE Trans. Inform. Theory*, vol. 17, pp. 630-632, Set 1971.

[33] A.J. Viterbi, J. K. Omura, *Principles of Digital Communicaton and Coding*, McGraw–Hill, New York, 1979.

[34] D. Wiberg, "Codes and Decoding on General Graphs", *Ph. D. thesis*, Linkoping Univ. , S-581 83, Linkoping, Sweden, 1996.