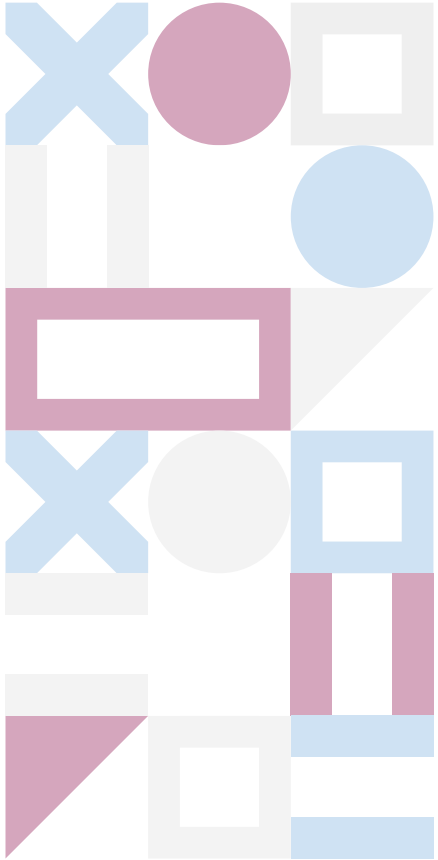


Introduccion a los códigos de distancia mínima y exponentes de error.

Jose Alejandro Montaña

CONTENIDO

INTRODUCCIÓN

- Recuento Historico.
- Nociones Básicas sobre Códigos.

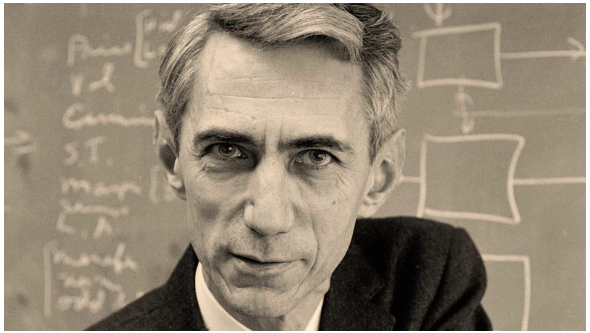
CODIGOS DE DISTANCIA MINIMA

- Limitaciones en estos códigos (Cotas).
- Conjetura de Goppa.

Exponentes de Error

- Cotas de Chernoff & Chebyshev.
- Calculo de Exponentes de Error

Teoría de corrección de errores



C. Shannon (1916-2001)

His work

- Proporciona una medida exacta del contenido de información en la salida de una fuente aleatoria en términos de su entropía (1948).
- Su teoría está basada en el manejo de errores de tipo estocásticos/probabilísticos.

R. Hamming (1915 - 1998)

Paper más reconocido

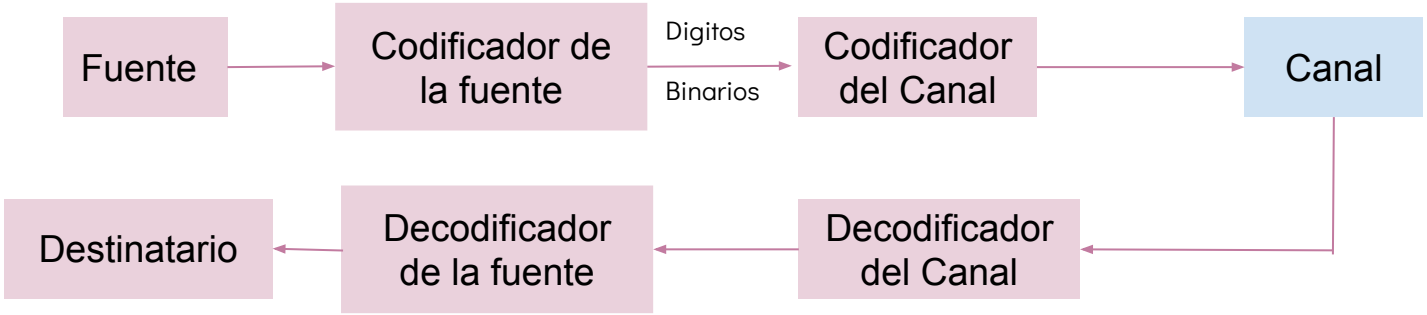
- Proporciona una visión de corrección de errores desde el punto de vista combinatorio/geométrico (1950).
- Su acercamiento al problema es más adecuado para atacar errores de tipo caso extremo/adversario.

- Aunque cada forma de abordar el problema tiene un propósito distinto y tienen limitaciones distintas, estas 2 construcciones comparten herramientas y técnicas en común.

Segundo teorema de Shannon (Noisy-Channel coding theorem)

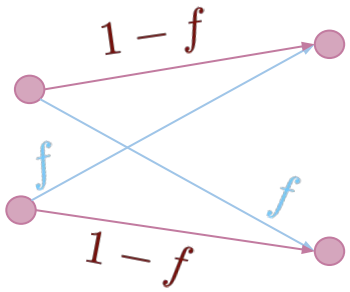
"Si la entropía de la fuente es menor que su capacidad de canal, entonces bajo ciertas condiciones, la información puede ser transmitida con una fiabilidad arbitraria."

- Cada canal tiene asociado un número real, llamado capacidad, el cual cuantifica la tasa a la cual es posible establecer un comunicación fiable sobre el canal.
- Si la información es transmitida a una tasa R, queremos decir que $k = NR$ bits del mensaje están siendo transmitidos en N usos del canal.



- La probabilidad de error está acotada por la capacidad del canal, la cual aumenta con R

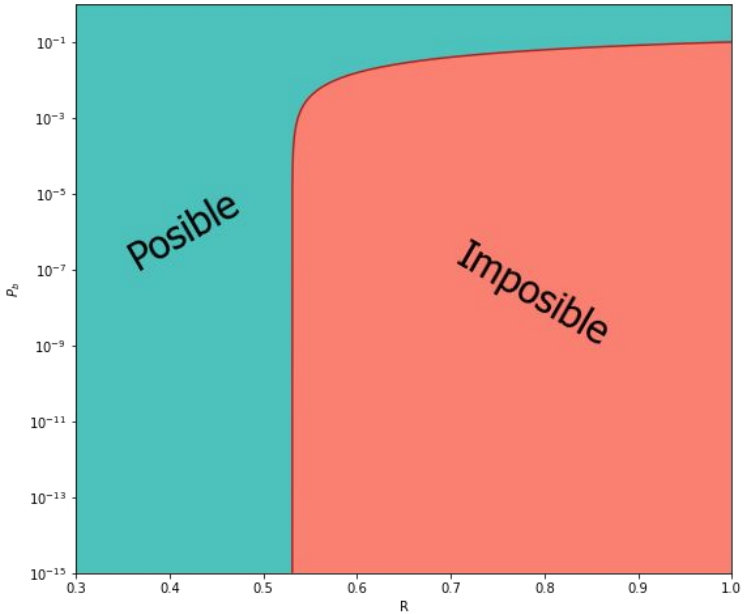
Canal Binario



$$R = \frac{C}{1 - H_2(P_e)}$$

$$C(f) = 1 - H_2(f) = 1 - \left[f \log_2 \frac{1}{f} + (1 - f) \log_2 \frac{1}{1 - f} \right]$$

$f = 0.1$ $C = 0.53$



Ejemplo (Código de repetición)

$$01011 \rightarrow 000\ 111\ 000\ 111\ 111$$

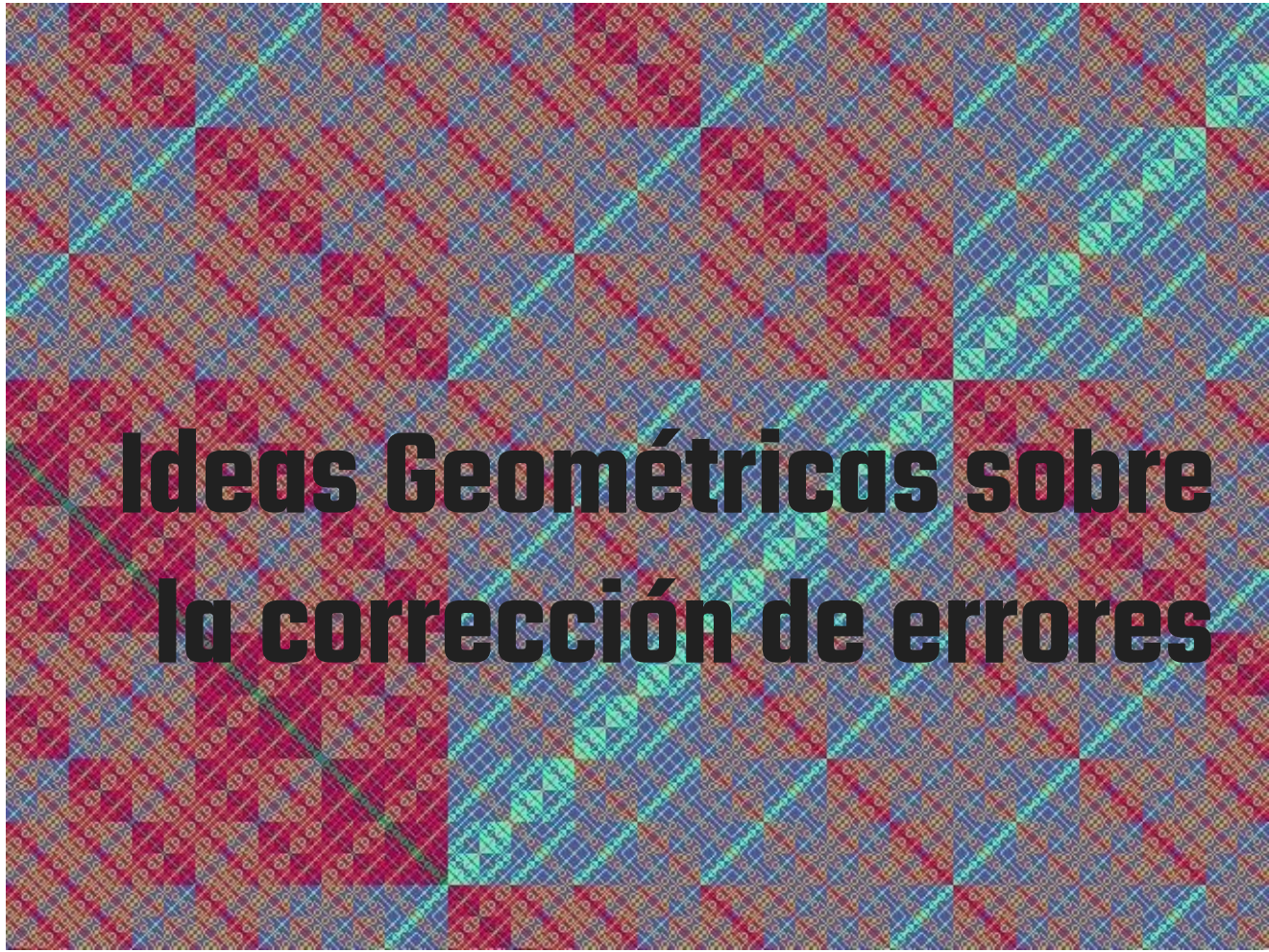
$$R = \frac{1}{3} \quad \Rightarrow \quad P_e = 0.03$$

Si queremos algo cuya probabilidad de error sea del orden de $P_e \sim 10^{-15}$

$$N \approx 60$$

Probabilidad de error con N repeticiones:

$$p_{\mathrm{e}} = \sum_{n=(N+1)/2}^{N} \binom{N}{n} f^n (1-f)^{N-n}$$


Ideas Geométricas sobre
la corrección de errores

Conceptos Básicos

Distancia de Hamming:

La distancia de Hamming entre dos secuencias "x" y "y" de la misma longitud sobre un alfabeto $\sum$ se denota por $\Delta(x,y)$ y se define como el número de posiciones en las cuales 2 cadenas difieren.

$$\Delta(x,y) = |\{i|x_i \neq y_i\}|$$

La distancia de Hamming fraccionaria está definida como

$$\delta(x,y) = \frac{\Delta(x,y)}{N}$$

Peso de Hamming:

El peso de Hamming de una cadena x sobre un alfabeto $\sum$ está definido como el número de no ceros en la cadena. De manera formal, el peso de Hamming de una cadena está definido como

$$\mathcal{W}(x) = |\{i|x_i \neq 0\}|$$

Note que $\mathcal{W}(x-y) = \Delta(x,y)$

Dada una cadena $x \in \Sigma^N$, la bola de Hamming de radio r alrededor de x es el conjunto

$$\{y \in \Sigma^N | \Delta(x,y) \leq r\}$$

Código:

Un código o bloque de corrección de error C de longitud N sobre un alfabeto finito Σes un subconjunto de Σ^N. Los elementos de C se denominan “palabras clave” (Code words) en C. El conjunto de todas las palabras clave se conoce también como “Code book”.

Si $|\Sigma| = q$, decimos que C es un código q-ario. La longitud N de las palabras clave de C se le llama el bloque de C.

Tasa del Código:

La tasa de un código $C \subseteq \Sigma^N$ denotada por $R(C)$ se define como

$$R(C) = \frac{\log |C|}{N \log |\Sigma|}$$

De esta forma, $R(C)$ es la cantidad de información no redundante por bit en la palabra clave de C.

Vea que si un código q-ario es de dimensión ℓ tendrá q^ℓ palabras claves

Distancia mínima:

Se define como la distancia de Hamming mínima entre 2 palabras clave de C.

$$\Delta(C) = \min_{\substack{c_1, c_2 \in C \\ c_1 \neq c_2}} \Delta(c_1, c_2)$$

Notación:

Un código q-ario de longitud N y dimensión k se entenderá se escribe como $[N, k]_q$. Si el código cumple además que tiene distancia mínima d se escribirá como

$$[N, k, d]_q$$

“Buenos Códigos”

Nos interesamos por los códigos q-arios de longitud de bloque N y distancia mínima d.

- ¿Cuál es la longitud máxima del código que se puede alcanzar?
- ¿Existe alguna forma de alcanzar la tasa de Shannon con estos códigos?

Cotas

- Cota de Hamming (Sphere Packing bound).
- Cota de Gilvert Varshamov.

Cota de Hamming

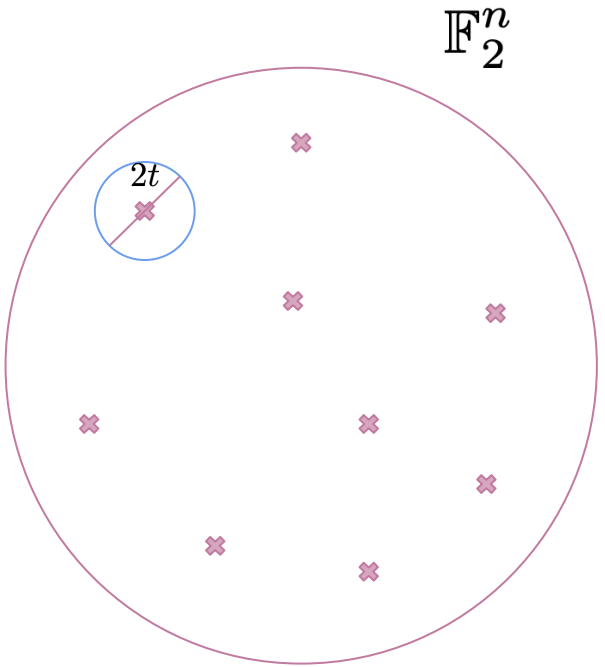
Considere el código $[n, k, d]$, nos preguntamos entonces por cuántos vectores hay dentro de una esfera de Hamming de radio t

Difiere en 1 posición

$$1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{t}$$

Palabra clave

Difiere en 2 posiciones

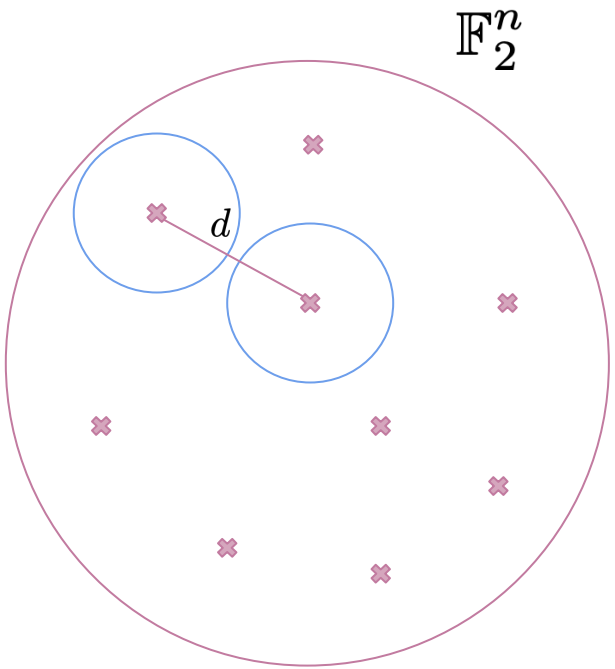


Nos preguntamos entonces qué tan grande puede ser t

Cota de Hamming

Considere el código $[n, k, d]$, nos preguntamos entonces por cuántos vectores hay dentro de una esfera de Hamming de radio t

$$t \leq \left\lfloor \frac{d-1}{2} \right\rfloor$$



Dado que hay 2^k palabras clave y 2^n posibles palabras en el código, se tiene que

$$2^k \left(1 + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{\lfloor \frac{d-1}{2} \rfloor}\right) \leq 2^n$$

Generalizando este razonamiento se tiene que

$$|C| \cdot V_q(n, \lfloor (d-1)/2 \rfloor) \leq q^n$$

con

$$V_q(n, t) = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$$

Codigos Perfectos

Se dice que un código es perfecto si la igualdad en la cota de Hamming se cumple.

Ejemplo

- El código de repetición $[n, 1, n]$, pues se tiene que

$$V_2(n,(n-1)/2) = \sum_{i=0}^{(n-1)/2} \binom{n}{i} = \frac{1}{2} \cdot \sum_{i=0}^{n} \binom{n}{i} = 2^{n-1}$$

- El código de Hamming $[2^r - 1, 2^r - r - 1, 3]$.

n debe ser impar para que sea un código perfecto.

Cota de Gilbert-Varshamov

Mientras la cota de Hamming nos proporciona Una condición suficiente para la existencia de un código lineal con ciertos parámetros

Teorema

Sean n,k y d parametros tales que

$$V_q(n-1, d-2) < q^{n-k}$$

Entonces existe un código lineal $[n, k]$ con distancia mínima de al menos d

Casi todos los códigos lineales tienen parámetros tales que cumplen la cota de Gilbert-Varshamov

Superar la cota de Gilbert-Varshamov es en general una tarea muy complicada

Tsfasman-Vlăduţ-Zink

https://doi.org/10.1002/mana.19821090103

Muestran una mejora asintótica para alfabetos de tamaño $q \geq 49$

Para $q < 46$ a la fecha se desconoce de una mejora a esta cota.

Conjetura de Goppa

La versión binaria de la cota de Gilbert-Varshamov es una cota asintoticamente exacta.

Trabajos posteriores [T. Jiang, A. Vardy (2005)], buscan mejorar esta cota para n y d pequeños, se ha encontrado una mejora asintótica pero el caso general aún sigue siendo una pregunta abierta

https://arxiv.org/abs/math/0404325v1

"Estamos más interesados en eventos atípicos que producen errores que en aquellos típicos en los que no se producen."

Desigualdad de Chernoff

$$\Pr(X \geq a) \leq \frac{\mathrm{E}(X)}{a}$$

Desigualdad de Markov

$$\Pr(X \geq a) = \Pr\left(e^{t \cdot X} \geq e^{t \cdot a}\right) \leq \frac{\mathrm{E}\left[e^{t \cdot X}\right]}{e^{t \cdot a}}$$

$$\Pr(X \leq a) = \Pr\left(e^{-tX} \geq e^{-ta}\right) \leq \frac{\mathrm{E}\left[e^{-t \cdot X}\right]}{e^{-t \cdot a}}$$

Códigos aleatorios Binarios

Un código binario C de longitud N y tasa R bits por segundo es un conjunto de $M = 2^{NR}$ N-tuplas $\boldsymbol{x}_i, 0 \leq i \leq M-1$.

Estudiamos entonces el comportamiento de un ensamble de códigos aleatorios (RCE) sobre el canal simétrico binario (BSC). Estamos interesados en el caso en que cada palabra clave es escogida de forma aleatoria y con igual probabilidad de tomar los valores 1 o 0.

La probabilidad de que una palabra aleatoria $\boldsymbol{x}_i$ de longitud N diste $d = N\delta$ de una palabra $\boldsymbol{b}$

$$\Pr\left\{d_H\left(\boldsymbol{x}_i, \boldsymbol{b}\right) = d\right\} = \binom{N}{d}\left(\frac{1}{2}\right)^d\left(\frac{1}{2}\right)^{N-d}$$

$$\doteq 2^{-N(1-\mathcal{H}(\delta))} = 2^{-ND\left(\delta\|\frac{1}{2}\right)}$$

Donde

$$D\left(p\|\frac{1}{2}\right) = 1 - \mathcal{H}(p) \quad \text{y} \quad D(p\|q) = p\log\frac{p}{q} + (1-p)\log\frac{1-p}{1-q}$$

Note que en este RCE 2 distancias son variables aleatorias independientes, siempre y cuando $\{i,j\} = \{i',j'\}$ or $\{i,j\} = \{j',i'\}$

Consideramos el conjunto de pares palabras clave en C tales que distan d

$$S_{\mathcal{C}}(d) = \sum_{i=0}^{M-1} \sum_{j=0}^{i-1} \Phi\left\{d_H\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = d\right\}$$

donde $\Phi\left\{d_H\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = d\right\}$ es el indicador de que ocurre el evento entre las llaves

$$\mathrm{E}S_{\mathcal{C}}(d) = \begin{pmatrix} M \\ 2 \end{pmatrix} \mathrm{E}\Phi \doteq 2^{N(2R-1+\mathcal{H}(\delta))}$$

La distancia mínima $d_{\min}(\mathcal{C})$ del código es igual al mínimo d, tal que

$$S_{\mathcal{C}}(d) \neq 0$$

Un hecho bien conocido es que la distancia mínima relativa de un código aleatorio se aproxima a $\delta_{\mathrm{GV}}(2R)$ as $N \to \infty$

Donde $\delta_{\mathrm{GV}}(R)$ está definida como la raiz de

$$\mathcal{H}(\delta) = 1 - R$$

Es natural dar los resultados en términos de la distancia relativa de Gilbert-Varshamov

Teorema

Para $0 \leq R < \frac{1}{2}$ y cualquier $\varepsilon > 0$, la probabilidad de que un código de longitud N y tasa R tomado de un RCE tenga distancia relativa menor que $\delta_{\mathrm{GV}}(2R) - \varepsilon$ va a cero de forma exponencial si $d = N\delta$

$$\delta_{\mathrm{GV}}(2R) + \varepsilon \leq \delta \leq 1 - \delta_{\mathrm{GV}}(2R) - \varepsilon$$

Entonces la probabilidad de que el número de pares de palabras clave que disten d, satisface que

$S_{\mathcal{C}}(d) \doteq 2^{N(2R-1+\mathcal{H}(\delta))}$ va a 1, cuando $N \to \infty$

Prueba

Tomamos

$$\frac{d}{N} \to \delta \leq \delta_{\mathrm{GV}}(2R) - \varepsilon$$

Entonces

$$\Pr\{S_{\mathcal{C}}(d) \geq 1\} \leq \mathrm{E}S_{\mathcal{C}}(d) \doteq 2^{-N(1-\mathcal{H}(\delta)-2R)} \to 0$$

Pero en el caso de que

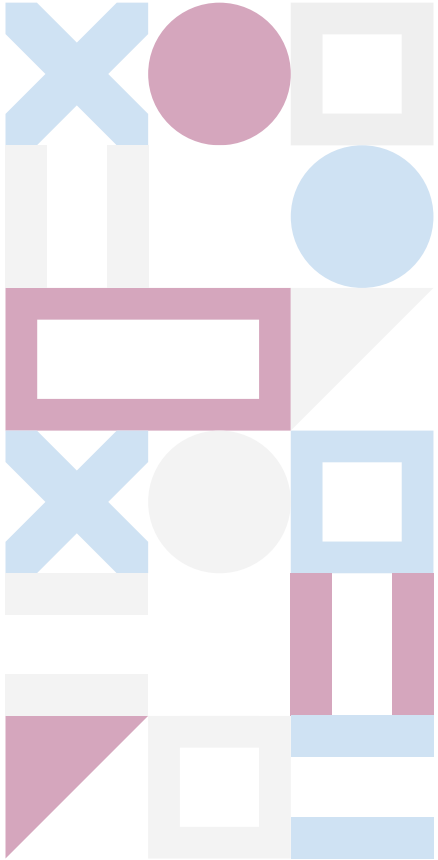
$$\delta_{\mathrm{GV}}(2R) + \varepsilon < \delta < 1 - \delta_{\mathrm{GV}}(2R) - \varepsilon$$

En promedio el número de pares que distan d es exponencialmente grande

$$\Pr\left\{|S_{\mathcal{C}}(d) - \mathrm{E}S_{\mathcal{C}}(d)| \geq \binom{M}{2}\alpha\right\} \leq \frac{\mathrm{E}\Phi}{\binom{M}{2}\alpha^2}$$

$$\Pr\left\{|S_{\mathcal{C}}(d) - \mathrm{E}S_{\mathcal{C}}(d)| > \binom{M}{2}\alpha\right\}$$
$$\leq \frac{2\mathrm{E}\Phi}{M(M-1)\alpha^2} \doteq 2^{-N(2R-1+\mathcal{H}(\delta)-2\Delta)}$$

Con probabilidad $1 - 2^{-\Omega(N)}$ con $N \to \infty$, la distancia relativa mínima de un código tomado de RCE es aproximadamente $\delta_{\mathrm{GV}}(2R)$ para $0 \leq R \leq \frac{1}{2}$ y 0 para $\frac{1}{2} \leq R \leq 1$



¿Cómo se pueden usar estos códigos de distancia mínima para poder estudiar un ensamble de estados fermiónicos?

Gracias!

¿Preguntas?

$$R \leq 1 - \frac{q}{q-1}\delta \rightarrow \text{ Plotkin}$$

$$R \leq 1 - H_q(\delta) \rightarrow \text{ G.V}$$

$$R \leq 1 - H_q(\frac{\delta}{2}) \rightarrow \text{ Hamming}$$