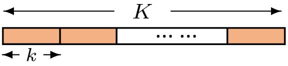
1. Concatentated Codes

Gadiel Seroussi

October 17, 2022

Concatenated codes: review

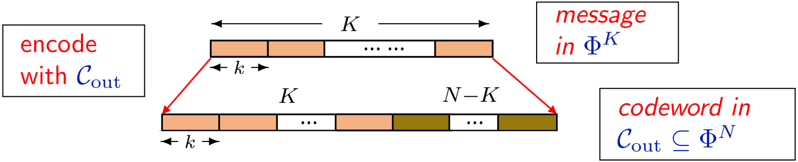
- Let $\mathcal{C}_{\text{in}}$ be an $[n, k, d]$ code over $F = \mathbb{F}_q$ (the *inner code*), and let $\mathcal{C}_{\text{out}}$ be an $[N, K, D]$ code over $\Phi = \mathbb{F}_{q^k}$ (the *outer code*).
  - We focus only on *linear* codes.
- Represent Φ as vectors in F^k using a fixed basis of Φ over F



message in Φ^K

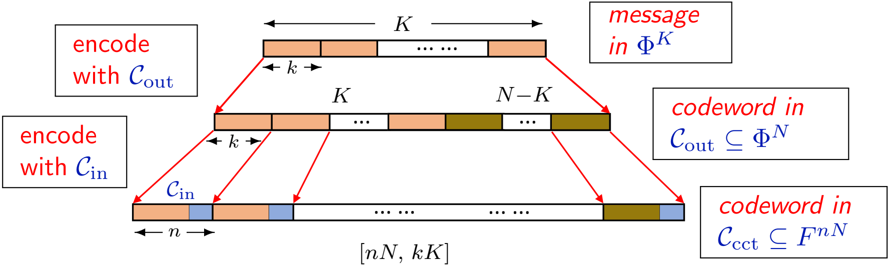
Concatenated codes: review

- Let $\mathcal{C}_{\text{in}}$ be an $[n, k, d]$ code over $F = \mathbb{F}_q$ (the *inner code*), and let $\mathcal{C}_{\text{out}}$ be an $[N, K, D]$ code over $\Phi = \mathbb{F}_{q^k}$ (the *outer code*).
  - We focus only on *linear* codes.
- Represent Φ as vectors in F^k using a fixed basis of Φ over F



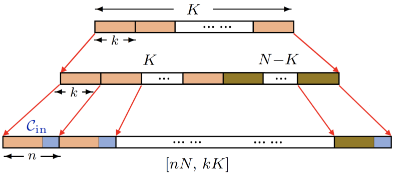
Concatenated codes: review

- Let $\mathcal{C}_{\text{in}}$ be an $[n, k, d]$ code over $F = \mathbb{F}_q$ (the *inner code*), and let $\mathcal{C}_{\text{out}}$ be an $[N, K, D]$ code over $\Phi = \mathbb{F}_{q^k}$ (the *outer code*).
  - We focus only on *linear* codes.
- Represent Φ as vectors in F^k using a fixed basis of Φ over F



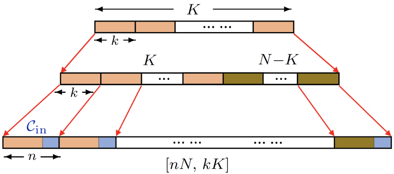
- A *concatenated code* $\mathcal{C}_{\text{cct}}$ is constructed by replacing each $\mathbb{F}^k$-symbol in $\mathcal{C}_{\text{out}}$ by its mapping to $\mathbb{F}^n$ according to $\mathcal{C}_{\text{in}}$.

Properties and variants

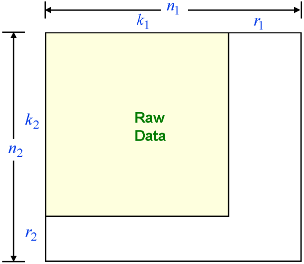


- $\mathcal{C}_{\text{cct}}$ has parameters $[nN, kK, \geq dD]$ over F
- $\mathcal{C}_{\text{out}}$ is typically taken to be a GRS code.
- Variants:
  - Use a different inner code
    $\mathcal{C}_{\text{in}}^{(j)},\ j = 1, 2, \ldots, N$
    for each coordinate of $\mathcal{C}_{\text{out}}$.

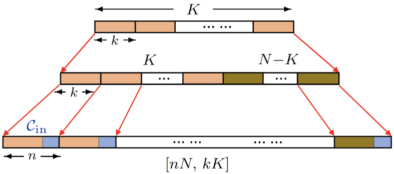
Properties and variants



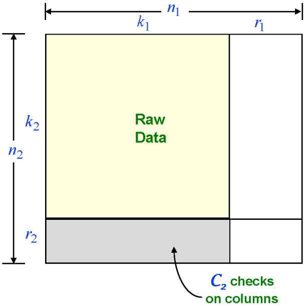
- $\mathcal{C}_{\text{cct}}$ has parameters $[nN, kK, \geq dD]$ over F
- $\mathcal{C}_{\text{out}}$ is typically taken to be a GRS code.
- Variants:
  - Use a different inner code $\mathcal{C}_{\text{in}}^{(j)},\ j = 1, 2, \ldots, N$ for each coordinate of $\mathcal{C}_{\text{out}}$.
  - *Product code*: given $\mathcal{C}_1 : [n_1, k_1]$ and $\mathcal{C}_2 : [n_2, k_2]$, a codeword in the product code $\mathcal{C}_1 \times \mathcal{C}_2$ is shown in the figure.



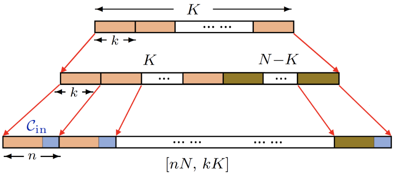
Properties and variants



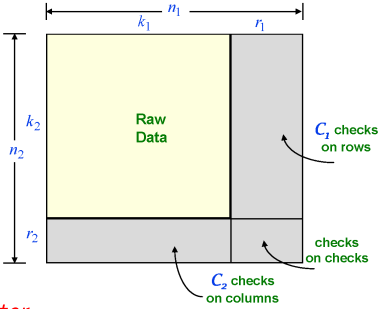
- $\mathcal{C}_{\text{cct}}$ has parameters $[nN, kK, \geq dD]$ over F
- $\mathcal{C}_{\text{out}}$ is typically taken to be a GRS code.
- Variants:
  - Use a different inner code $\mathcal{C}_{\text{in}}^{(j)},\ j = 1, 2, \ldots, N$ for each coordinate of $\mathcal{C}_{\text{out}}$.
  - *Product code*: given $\mathcal{C}_1 : [n_1, k_1]$ and $\mathcal{C}_2 : [n_2, k_2]$, a codeword in the product code $\mathcal{C}_1 \times \mathcal{C}_2$ is shown in the figure.



Properties and variants



- $\mathcal{C}_{\text{cct}}$ has parameters $[nN, kK, \geq dD]$ over F
- $\mathcal{C}_{\text{out}}$ is typically taken to be a GRS code.
- Variants:
  - Use a different inner code $\mathcal{C}_{\text{in}}^{(j)}$, $j = 1, 2, \ldots, N$ for each coordinate of $\mathcal{C}_{\text{out}}$.
  - *Product code*: given $\mathcal{C}_1 : [n_1, k_1]$ and $\mathcal{C}_2 : [n_2, k_2]$, a codeword in the product code $\mathcal{C}_1 \times \mathcal{C}_2$ is shown in the figure.



More on this later.

Review: Some notation and properties

- Volume of Hamming sphere of radius t in F^n, $F = GF(q)$.

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

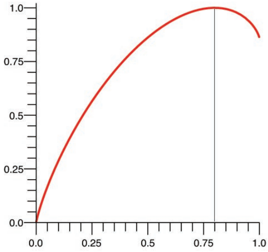
- *Symmetric q-ary entropy function*
  $\mathsf{H}_q : [0,1] \to [0,1]$

  $$\mathsf{H}_q(x) = -x \log_q x - (1-x)\log_q(1-x) + x\log_q(q-1)\,.$$

- Bounds on $V_q(n,t)$

$$\frac{1}{\sqrt{8t(1-(t/n))}} \cdot q^{n\mathsf{H}_q(t/n)} \le V_q(n,t) \le q^{n\mathsf{H}_q(t/n)}\,.$$

$\mathsf{H}_q(x)$, $q=5$

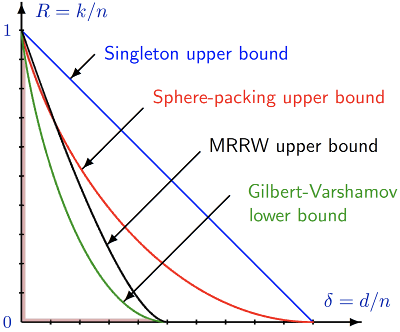


x

Asymptotically good codes (in the min. distance sense)

We seek a sequence of linear codes $\{\mathcal{C}_i : [n_i, k_i, d_i]\}_{i=1}^{\infty}$, with $n_i \xrightarrow{i\to\infty} \infty$, such that

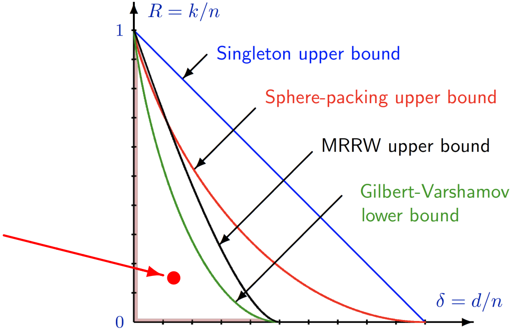
- with $R_i = k_i/n_i$, $\liminf_{i\to\infty} R_i > 0$ *rate bounded away from zero*,
- with $\delta_i = d_i/n_i$, $\lim \inf_{i\to\infty} \delta_i > 0$ *relative distance bounded away from zero*,
- $\mathcal{C}_i$ can be *constructed* in time polynomial in n_i,
- $\mathcal{C}_i$ can be *encoded* and *decoded* in time polynomial in n_i.



Asymptotically good codes (in the min. distance sense)

We seek a sequence of linear codes $\{\mathcal{C}_i : [n_i, k_i, d_i]\}_{i=1}^{\infty}$, with $n_i \xrightarrow{i\to\infty} \infty$, such that

- with $R_i = k_i/n_i$, $\liminf_{i\to\infty} R_i > 0$ *rate bounded away from zero*,
- with $\delta_i = d_i/n_i$, $\lim \inf_{i\to\infty} \delta_i > 0$ *relative distance bounded away from zero*,
- $\mathcal{C}_i$ can be *constructed* in time polynomial in n_i,
- $\mathcal{C}_i$ can be *encoded* and *decoded* in time polynomial in n_i.



Construction of good concatenated codes (i)

- Consider a finite field $F = GF(q)$, its extension $\Phi = GF(q^k)$, and an element $\beta \in \Phi$. The map
$$x \mapsto \beta \cdot x \,,$$
acting on elements of Φ, is a linear transformation over F.

- Given a basis $\Omega = (\omega_1\, \omega_2\, \ldots\, \omega_k)$ of Φ over F, this map is represented by a $k \times k$ matrix $M(\beta)$, such that if $y = \beta x$, $x \in \Phi$, then
$$\mathbf{y} = M(\beta) \cdot \mathbf{x} \,,$$
where $\mathbf{x}$ and $\mathbf{y}$ are (column) vector representations of x and y, respectively, with respect to the basis Ω, i.e., $x = \Omega \cdot \mathbf{x}$ and $y = \Omega \cdot \mathbf{y}$.

- Consider the code $\mathcal{C}(\beta)$ generated by
$$G_\beta = \left[I_{k\times k} \;\middle|\; M(\beta)^T \,\right] \,.$$
$\mathcal{C}(\beta)$ is an $[n = 2k, k, d]$ code over F.

The Wozencraft code ensemble

Definition

The *Wozencraft* $[2k,k]$ *code ensemble over* F is the set

$$\mathcal{W}_F(2k,k) = \{\, \mathcal{C}(\beta) \,:\, \beta \in \Phi \,\}$$

- All nonzero codewords in $\mathcal{C}(\beta)$ are of the form $[\mathbf{a} \,|\, \mathbf{b}]$ with $b/a = \beta$ ($a \neq 0$).

The Wozencraft code ensemble

Definition

The *Wozencraft* $[2k, k]$ *code ensemble over* F is the set

$$\mathcal{W}_F(2k, k) = \{\, \mathcal{C}(\beta) \,:\, \beta \in \Phi \,\}$$

- All nonzero codewords in $\mathcal{C}(\beta)$ are of the form $[\mathbf{a} \,|\, \mathbf{b}]$ with $b/a = \beta$ ($a \neq 0$).
- The definition of Wozencraft codes is extended to cover lengths n, $k < n \leq 2k$ by defining the $[n, k]$ code $\mathcal{C}_{\beta,n}$ as

$$\mathcal{C}_{\beta,n} = \{\, (c_1\, c_2 \,\ldots\, c_n) \,:\, (c_1\, c_2 \,\ldots, c_n, \ldots\, c_{2k}) \in \mathcal{C}(\beta) \,\}\, , k < n \leq 2k \, .$$

Definition

The *Wozencraft* $[n, k]$ *code ensemble over* F is the set

$$\mathcal{W}_F(n, k) = \{\mathcal{C}_{\beta,n} \,:\, \beta \in \Phi \,\} \, .$$

Properties of Wozencraft codes (i)

Lemma

Every nonzero word $\mathbf{c} \in F^n$ belongs to at most q^{2k-n} codes in $\mathcal{W}_F(n,k)$.

Proof.

For $n = 2k$, a nonzero word $\mathbf{c} = [\mathbf{a} \,|\, \mathbf{b}]$ can belong only to $\mathcal{C}(\beta)$ for $\beta = b/a$ ($a \neq 0$), or none if $a = 0$. When $2k > n$, $\mathbf{c}$ can be completed in q^{2k-n} ways into a word of length $2k$. Each such completion belongs to at most one code $\mathcal{C}(\beta)$. Hence, there are at most q^{2k-n} values β such that $\mathbf{c} \in \mathcal{C}_{\beta,n}$.

Properties of Wozencraft codes (ii)

- What can we say about minimum distance of Wozencraft codes? For example, $\mathcal{C}(0)$ contains the word $(1\,0\,\ldots\,0\,0\,0\,\ldots\,0)$ (*bad*). However,

Proposition

The number of codes in $\mathcal{W}_F(n,k)$ with minimum distance less than a given integer d is at most $q^{2k-n}(V_q(n,d-1)-1)$.

Proof.

There are $V_q(n,d-1)-1$ nonzero words of weight less than d in F^n. By the Lemma, each such word belongs to at most q^{2k-n} codes in $\mathcal{W}_F(n,k)$.

Justesen codes

- Let k and n be positive integers such that $k < n \leq 2k$, and write, for convenience, $\Phi = \{\beta_1, \beta_2, \ldots, \beta_{q^k}\}$.
- Let $\mathcal{E}_j$ denote an encoder for $\mathcal{C}_{\beta_j, n}$, and d_j its minimum distance.
- Let $\mathcal{C}_{\text{out}}$ be a $[N, K, D]$ extended GRS code with $N = q^k$, $K = \lceil RN \rceil$ for some given $R \in (0, 1]$, and $D = N - K + 1 > (1 - R)N$.

Justesen codes

- Let k and n be positive integers such that $k < n \leq 2k$, and write, for convenience, $\Phi = \{\beta_1, \beta_2, \ldots, \beta_{q^k}\}$.
- Let $\mathcal{E}_j$ denote an encoder for $\mathcal{C}_{\beta_j,n}$, and d_j its minimum distance.
- Let $\mathcal{C}_{\text{out}}$ be a $[N, K, D]$ extended GRS code with $N = q^k$, $K = \lceil RN \rceil$ for some given $R \in (0, 1]$, and $D = N - K + 1 > (1 - R)N$.

Definition

The *Justesen code* $\mathcal{C}_{\mathrm{J}}$ is defined as follows

$$\mathcal{C}_{\mathrm{J}} = \left\{ \left(\mathcal{E}_1(\mathbf{z}_1) \,|\, \mathcal{E}_2(\mathbf{z}_2) \,|\, \ldots \,|\, \mathcal{E}_N(\mathbf{z}_N) \right) \;:\; (\mathbf{z}_1\, \mathbf{z}_2\, \ldots\, \mathbf{z}_N) \in \mathcal{C}_{\text{out}} \right\}.$$

- Like a concatenated code, but with a different inner code in each coordinate.
- As with concatenated codes, the parameters are $[nN, kK]$. How about the minimum distance D_{J}? It will not be of the form dD, because there is no fixed d for the inner codes.

The minimum distance of the Justesen code

- A codeword $\mathbf{c}_{\min} \in \mathcal{C}_{\mathrm{J}}$ of minimum weight has at least D nonzero sub-blocks $\mathcal{E}_j(\mathbf{z}_j)$.
- By the previous proposition, for every positive integer d, we have
$$\mathsf{D}_{\mathrm{J}} = \mathsf{wt}(\mathbf{c}_{\min}) > d \cdot \left(D - q^{2k-n} V_q(n, d-1)\right) . \qquad (\blacklozenge)$$

The minimum distance of the Justesen code

- A codeword $\mathbf{c}_{\min} \in \mathcal{C}_{\mathrm{J}}$ of minimum weight has at least D nonzero sub-blocks $\mathcal{E}_j(\mathbf{z}_j)$.
- By the previous proposition, for every positive integer d, we have
$$\mathsf{D}_{\mathrm{J}} = \mathsf{wt}(\mathbf{c}_{\min}) > d \cdot \left(D - q^{2k-n} V_q(n, d-1)\right) . \qquad (\blacklozenge)$$
- **Example:** $q = 2,\ n = 2k$. Let $\theta \in (0,1)$ be such that $n\theta$ is an integer and $H_2(\theta) = \frac{1}{2} - \epsilon$, with $\epsilon \in (0, \frac{1}{2})$. Choose $d = n\theta + 1$. Then, we have
$$\begin{aligned} \mathsf{D}_{\mathrm{J}} &> d\big(D - V_q(n, d-1)\big) > n\theta \left(N(1-R) - 2^{nH(\theta)}\right) \\ &= n\theta \left(N(1-R) - 2^{2k(\frac{1}{2}-\epsilon)}\right) = n\theta \left(N(1-R) - 2^{k-n\epsilon}\right) \\ &= nN\theta\big(1 - R - o(1)\big) . \quad (\text{recall } N = 2^k) \end{aligned}$$

  Therefore, $\mathcal{C}_{\mathrm{J}}$ has rate $R_{\mathrm{J}} = \frac{1}{2} R > 0$ and relative distance $\delta_{\mathrm{J}} = \frac{\mathsf{D}_{\mathrm{J}}}{nN} = \theta(1-R) - o(1) > 0$.

The minimum distance of the Justesen code

- A codeword $\mathbf{c}_{\min} \in \mathcal{C}_\mathrm{J}$ of minimum weight has at least D nonzero sub-blocks $\mathcal{E}_j(\mathbf{z}_j)$.
- By the previous proposition, for every positive integer d, we have

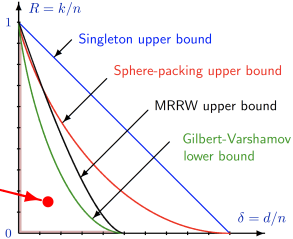
$$\mathsf{D}_\mathrm{J} = \mathsf{wt}(\mathbf{c}_{\min}) > d \cdot \left(D - q^{2k-n} V_q(n, d-1)\right) . \qquad (\blacklozenge)$$

- **Example:** $q = 2,\ n = 2k$. Let $\theta \in (0,1)$ be such that $n\theta$ is an integer and $H_2(\theta) = \frac{1}{2} - \epsilon$, with $\epsilon \in (0, \frac{1}{2})$. Choose $d = n\theta + 1$. Then, we have

$$\begin{aligned} \mathsf{D}_\mathrm{J} &> d\big(D - V_q(n, d-1)\big) > n\theta\left(N(1-R) - 2^{nH(\theta)}\right) \\ &= n\theta\left(N(1-R) - 2^{2k(\frac{1}{2}-\epsilon)}\right) = n\theta\left(N(1-R) - 2^{k-n\epsilon}\right) \\ &= nN\theta\big(1 - R - o(1)\big) . \qquad (\text{recall } N = 2^k) \end{aligned}$$

Therefore, $\mathcal{C}_\mathrm{J}$ has rate $R_\mathrm{J} = \frac{1}{2}R > 0$ and relative distance $\delta_\mathrm{J} = \frac{\mathsf{D}_\mathrm{J}}{nN} = \theta(1-R) - o(1) > 0$.

We've got constructive, asymptotically good codes!



Justesen code: general case asymptotics

$$\mathsf{D}_{\mathrm{J}} = \mathsf{wt}(\mathbf{c}_{\min}) > d \cdot \left(D - q^{2k-n} V_q(n, d-1)\right) . \qquad (\blacklozenge)$$

- To study the asymptotic trade-off R_{J} vs. δ_{J} in the general case, write $r = k/n$, and let θ be a real number (function) satisfying

$$\theta = \mathsf{H}_q^{-1}(1 - r - \epsilon(n)) ,$$

where

$$\lim_{n\to\infty} \epsilon(n) = 0 \quad \text{and} \quad \lim_{n\to\infty} n\,\epsilon(n) = \infty \quad \text{(e.g., } \epsilon(n) = \log n/n\text{).}$$

Justesen code: general case asymptotics

$$\mathsf{D}_{\mathrm{J}} = \mathsf{wt}(\mathbf{c}_{\min}) > d \cdot \left(D - q^{2k-n} V_q(n, d-1)\right) . \qquad (\blacklozenge)$$

- To study the asymptotic trade-off R_{J} vs. δ_{J} in the general case, write $r = k/n$, and let θ be a real number (function) satisfying
$$\theta = \mathsf{H}_q^{-1}(1 - r - \epsilon(n)) ,$$
where
$$\lim_{n\to\infty} \epsilon(n) = 0 \quad \text{and} \quad \lim_{n\to\infty} n\,\epsilon(n) = \infty \quad (\text{e.g., } \epsilon(n) = \log n/n).$$

- Selecting $d = \lceil \theta n \rceil$ in ($\blacklozenge$), and recalling that $V_q(n,t) \le q^{n\mathsf{H}_q(t/n)}$ and $N = q^k = q^{rn}$, we obtain
$$\begin{aligned} \mathsf{D}_{\mathrm{J}} &> \theta n \cdot \left((1-R)N - q^{(2r-1)n} \cdot q^{n\mathsf{H}_q(\theta)}\right) \\ &= \theta n N \left((1-R) - q^{n(r-1+\mathsf{H}_q(\theta))}\right) = \theta n N \left((1-R) - q^{n\epsilon(n)}\right) \\ &\Longrightarrow \ \delta_{\mathrm{J}} = \frac{\mathsf{D}_{\mathrm{J}}}{nN} > \theta\,(1 - R - o(1)) . \end{aligned}$$

Justesen code: general case asymptotics

$$\mathsf{D}_{\mathrm{J}} = \mathsf{wt}(\mathbf{c}_{\min}) > d \cdot \left(D - q^{2k-n} V_q(n, d-1) \right) . \qquad (\blacklozenge)$$

- To study the asymptotic trade-off R_{J} vs. δ_{J} in the general case, write $r = k/n$, and let θ be a real number (function) satisfying

  $$\theta = \mathsf{H}_q^{-1}(1 - r - \epsilon(n)) \,,$$

  where

  $$\lim_{n\to\infty} \epsilon(n) = 0 \quad \text{and} \quad \lim_{n\to\infty} n\,\epsilon(n) = \infty \quad \text{(e.g., } \epsilon(n) = \log n/n\text{)}.$$

- Selecting $d = \lceil \theta n \rceil$ in ($\blacklozenge$), and recalling that $V_q(n,t) \le q^{n\mathsf{H}_q(t/n)}$ and $N = q^k = q^{rn}$, we obtain

  $$\begin{aligned} \mathsf{D}_{\mathrm{J}} &> \theta n \cdot \left((1{-}R)N - q^{(2r-1)n} \cdot q^{n\mathsf{H}_q(\theta)} \right) \\ &= \theta n N \left((1{-}R) - q^{n(r-1+\mathsf{H}_q(\theta))} \right) = \theta n N \left((1{-}R) - q^{n\epsilon(n)} \right) \end{aligned}$$

  $$\Longrightarrow \quad \delta_{\mathrm{J}} = \frac{\mathsf{D}_{\mathrm{J}}}{nN} > \theta\,(1 - R - o(1)) \,.$$

- For the rate R_{J} of $\mathcal{C}_{\mathrm{J}}$, we have

  $$R_{\mathrm{J}} \ge rR = \left(1 - \mathsf{H}_q(\theta) - \epsilon(n)\right)R = \left(1 - \mathsf{H}_q(\theta) - o(1)\right)R \,.$$

Rate-minimum distance trade-off for the Justesen code

$$\delta_{\mathrm{J}} > \theta\,(1 - R - o(1)), \qquad R_{\mathrm{J}} \ge \big(1 - \mathsf{H}_q(\theta) - o(1)\big)\,R\,.$$

We can maximize the rate over θ, for a given δ_J (setting $R \approx 1 - \frac{\delta}{\theta}$).
Notice, however, that the rates of the Wozencraft codes must be in the interval $[\frac{1}{2}, 1)$, so we must have $\theta \le \mathsf{H}_q^{-1}(\frac{1}{2})$. [$q{=}2:\ \theta \le \theta_0 \approx 0.1100$]

Rate-minimum distance trade-off for the Justesen code

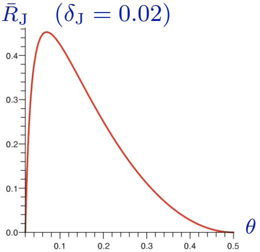
$$\delta_{\mathrm{J}} > \theta\,(1-R-o(1)), \qquad R_{\mathrm{J}} \ge \big(1-\mathsf{H}_q(\theta)-o(1)\big)\,R\,.$$

We can maximize the rate over θ, for a given δ_J (setting $R \approx 1-\frac{\delta}{\theta}$).
Notice, however, that the rates of the Wozencraft codes must be in the interval $[\frac{1}{2},1)$, so we must have $\theta \le \mathsf{H}_q^{-1}(\frac{1}{2})$. $[q{=}2:\ \theta \le \theta_0 \approx 0.1100]$

- We obtain the lower bound $R_{\mathrm{J}} \ge \bar{R}_{\mathrm{J}}(\delta,q) - o(1)$

$$\bar{R}_{\mathrm{J}}(\delta,q) = \max_{\theta\in[\delta,\mathsf{H}_q^{-1}(\frac{1}{2})]} \Big(1-\mathsf{H}_q(\theta)\Big)\left(1-\frac{\delta}{\theta}\right)$$

(note that for $\delta = \mathsf{H}_q^{-1}(\frac{1}{2})$ we get $R_{\mathrm{J}}(\delta,q)=0$).

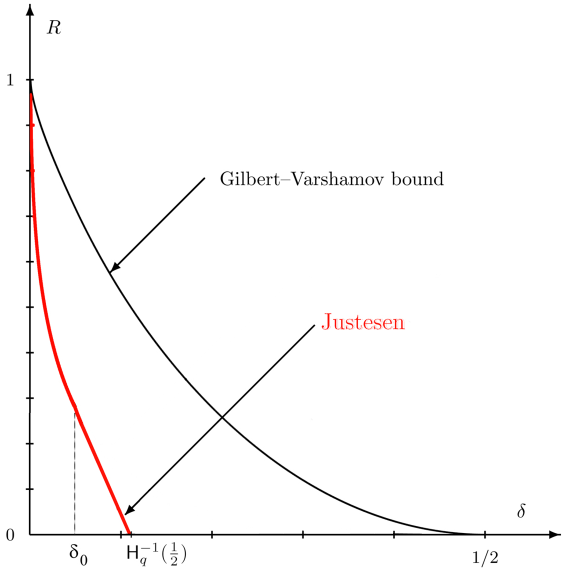


Example

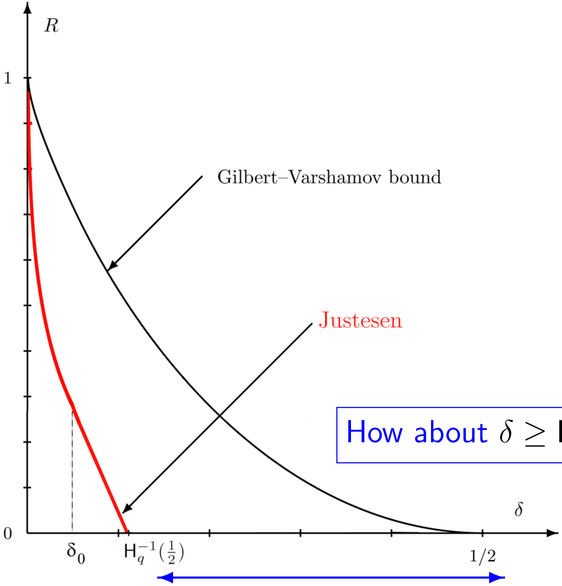
For $q=2$ we find, numerically, $\delta_0(2) \approx 0.0439$, $\bar{R}_{\mathrm{J}}(\delta_0(2),2) \approx 0.3005\,.$

- When $\delta > \delta_0(q)$, the maximum is obtained at $\theta = \mathsf{H}_q^{-1}(\frac{1}{2})$, and the bound becomes $\bar{R}_{\mathrm{J}}(\delta,q) = \frac{1}{2}\left(1-\frac{\delta}{\mathsf{H}_q^{-1}(\frac{1}{2})}\right)$, a straight line.

Justesen codes—Asymptotics



Justesen codes—Asymptotics



How about $\delta \geq \mathsf{H}_q^{-1}(\frac{1}{2})$?

Gilbert-Varshamov revisited

Theorem (Asymptotic Gilbert-Varshamov bound)

Let $F = GF(q)$ and n and nr be positive integers with $r \in [0, 1]$. There exist a linear $[n, nr, \geq \delta n]$ code $\mathcal{C}_{\mathrm{GV}}$ over F with

$$\delta = \mathsf{H}_q^{-1}(1 - r)\,.$$

Gilbert-Varshamov revisited

Theorem (Asymptotic Gilbert-Varshamov bound)

Let $F = GF(q)$ and n and nr be positive integers with $r \in [0,1]$. There exist a linear $[n, nr, \geq \delta n]$ code $\mathcal{C}_{\mathrm{GV}}$ over F with

$$\delta = \mathsf{H}_q^{-1}(1-r)\,.$$

The code $\mathcal{C}_{\mathrm{GV}}$ is constructed by building a parity-check matrix $H = [\mathbf{h}_1\,\mathbf{h}_2\,\ldots\mathbf{h}_i\,\ldots]$ column by column, according to the following rule:

Choose $\mathbf{h}_{i+1}$ among columns that are not linear combinations of $\lceil \delta n \rceil - 2$ columns from $\{\mathbf{h}_1\,\mathbf{h}_2\,\ldots\mathbf{h}_i\}$.

The number of the linear combinations to check is $O(V_q(n-1, \lceil \delta n \rceil - 2)) = q^{n(\mathsf{H}_q(\delta) - o(1))} \implies$ construction of H takes time exponential in n for each fixed δ.

Exponential in n is polynomial in q^{rn}.

Construction of good concatenated codes (ii)

- We use $\mathcal{C}_{\mathrm{GV}}$ as the inner code $\mathcal{C}_{\mathrm{in}}$, concatenated with an $[N = q^{rn}, K, D]$ extended primitive GRS code over $\Phi = GF(q^{rn})$ as $\mathcal{C}_{\mathrm{out}}$. Here, $K = RN$ and $D > (1-R)N$ for some real $R \in (0,1)$.

- The parameters of $\mathcal{C}_{\mathrm{cct}}$ are given by

$$\begin{aligned} n_{\mathrm{cct}} &= nN = nq^{n(1-\mathsf{H}_q(\delta))}, \\ k_{\mathrm{cct}} &= (1-\mathsf{H}_q(\delta))R \cdot nN, \\ d_{\mathrm{cct}} &\geq \delta(1-R) \cdot nN\,. \end{aligned}$$

- The length of $\mathcal{C}_{\mathrm{cct}}$ can be arbitrarily large.
- The rate and relative minimum distance satisfy

$$\begin{aligned} R_{\mathrm{cct}} &= (1-\mathsf{H}_q(\delta))R\,, \\ \delta_{\mathrm{cct}} &\geq \delta(1-R)\,. \end{aligned}$$

Construction of good concatenated codes (ii)

- We use $\mathcal{C}_{\mathrm{GV}}$ as the inner code $\mathcal{C}_{\mathrm{in}}$, concatenated with an $[N = q^{rn}, K, D]$ extended primitive GRS code over $\Phi = GF(q^{rn})$ as $\mathcal{C}_{\mathrm{out}}$. Here, $K = RN$ and $D > (1-R)N$ for some real $R \in (0,1)$.

- The parameters of $\mathcal{C}_{\mathrm{cct}}$ are given by

$$\begin{aligned} n_{\mathrm{cct}} &= nN = nq^{n(1-\mathsf{H}_q(\delta))}, \\ k_{\mathrm{cct}} &= (1-\mathsf{H}_q(\delta))R \cdot nN, \\ d_{\mathrm{cct}} &\geq \delta(1-R) \cdot nN\,. \end{aligned}$$

- The length of $\mathcal{C}_{\mathrm{cct}}$ can be arbitrarily large.
- The rate and relative minimum distance satisfy

$$\begin{aligned} R_{\mathrm{cct}} &= (1-\mathsf{H}_q(\delta))R\,, \\ \delta_{\mathrm{cct}} &\geq \delta(1-R)\,. \end{aligned}$$

- Given a designed relative minimum distance $\delta_{\mathrm{cct}} \in (0, 1-q^{-1})$, we can maximize R_{cct} over δ and R, subject to $\delta(1-R) \leq \delta_{\mathrm{cct}}$. This yields

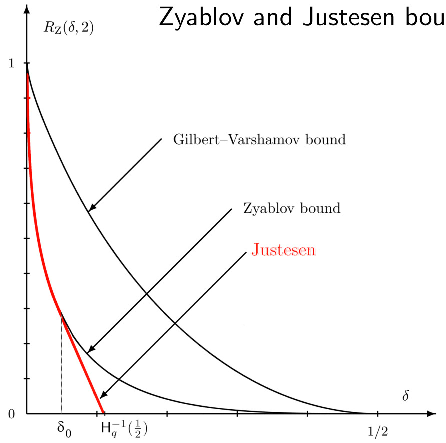
Zyablov bound

$$R_{\mathrm{cct}} \geq R_Z(\delta_{\mathrm{cct}}, q) = \max_{\delta \in [\delta_{\mathrm{cct}}, 1-(1/q)]} \Big(1 - \mathsf{H}_q(\delta)\Big)\left(1 - \frac{\delta_{\mathrm{cct}}}{\delta}\right).$$

The Zyablov bound

$$R_{\mathrm{cct}} \geq R_Z(\delta_{\mathrm{cct}}, q) = \max_{\delta \in [\delta_{\mathrm{cct}}, 1-(1/q)]} \Big(1 - \mathsf{H}_q(\delta)\Big)\left(1 - \frac{\delta_{\mathrm{cct}}}{\delta}\right).$$

Zyablov and Justesen bounds for $q = 2$

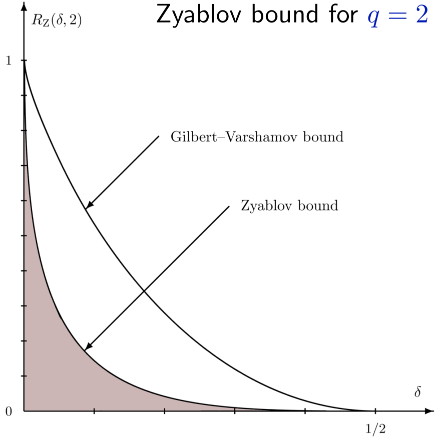


The bounds coincide for $\delta \leq \delta_0$

The Zyablov bound

$$R_{\mathrm{cct}} \geq R_Z(\delta_{\mathrm{cct}}, q) = \max_{\delta \in [\delta_{\mathrm{cct}}, 1-(1/q)]} \Big(1 - \mathsf{H}_q(\delta)\Big)\left(1 - \frac{\delta_{\mathrm{cct}}}{\delta}\right).$$

Zyablov bound for $q = 2$



- The Zyablov bound is inferior to the GV bound.
- However, a generator matrix for a code $\mathcal{C}_{\mathrm{cct}}$ achieving the bound can be constructed in time *polynomial* in n_{cct}.
  - A parity check matrix for $\mathcal{C}_{\mathrm{GV}}$ can be constructed in time $O\left(V_q(n-1, \lceil \delta n \rceil - 2)\right) = O({n_{\mathrm{cct}}}^{(1/r)-1})$, where $r = 1 - \mathsf{H}_q(\delta)$.
  - A matrix for the GRS code is also easily built.

Decoding of concatenated codes

Minimum distance is dD. Can we decode up to $\lfloor (dD-1)/2 \rfloor$ errors?

- Suppose that a codeword

$$\mathbf{c} = (\mathbf{c}_1 \,|\, \mathbf{c}_2 \,|\, \ldots \,|\, \mathbf{c}_N) \in \mathcal{C}_{\text{cct}}$$

was transmitted through a noisy channel, and

$$\mathbf{y} = (\mathbf{y}_1 \,|\, \mathbf{y}_2 \,|\, \ldots \,|\, \mathbf{y}_N) \in F^{nN}$$

was received, where $\mathbf{y}_j \in F^n, \; j = 1, 2, \ldots, N$, and assume $\mathsf{d}(\mathbf{y}, \mathbf{c}) < dD/2$ (*as words in* F^{nN}).

- Suppose also that we have a *nearest codeword decoder* $\mathcal{D}_{\text{in}}$ for $\mathcal{C}_{\text{in}}$.
- Let

$$\hat{\mathbf{c}}_j = \mathcal{D}_{\text{in}}(\mathbf{y}_j)\,, \quad \text{and} \quad \hat{\mathbf{z}}_j = \mathcal{E}_{\text{in}}^{-1}(\hat{\mathbf{c}}_j)\,, \;\; j = 1, 2, \ldots, N\,.$$

Decoding of concatenated codes (ii)

- The following decoding strategy is parametrized by $\mu \in \{1, 2, \ldots, \lceil d/2 \rceil\}$.

  - Compute $\mathbf{x} = \mathbf{x}(\mu) = (\, \mathbf{x}_1\, \mathbf{x}_2\, \ldots\, \mathbf{x}_N \,) \in (\, \Phi \cup \{?\} \,)^N$, with

  $$\mathbf{x}_j = \begin{cases} \hat{\mathbf{z}}_j & \text{if } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ ? & \text{(erasure) otherwise} \end{cases} \qquad (\dagger\dagger)$$

  - Use an errors+erasures decoder $\mathcal{D}_{\text{out}}$ for $\mathcal{C}_{\text{out}}$ on $\mathbf{x}$, obtaining a decoded word $\hat{\mathbf{c}} \in \mathcal{C}_{\text{out}}$, or a FAIL indicator.

The parameter μ is a threshold that $\mathcal{D}_{\text{in}}$ utilizes to determine whether to attempt correction of a corrupted codeword or declare it erased.

Decoding of concatenated codes (ii)

- The following decoding strategy is parametrized by $\mu \in \{1, 2, \ldots, \lceil d/2 \rceil\}$.

  - Compute $\mathbf{x} = \mathbf{x}(\mu) = (\,\mathbf{x}_1\,\mathbf{x}_2\,\ldots\,\mathbf{x}_N\,) \in (\,\Phi \cup \{?\}\,)^N$, with

$$\mathbf{x}_j = \begin{cases} \hat{\mathbf{z}}_j & \text{if } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ ? & \text{(erasure) otherwise} \end{cases} \qquad (\dagger\dagger)$$

  - Use an errors+erasures decoder $\mathcal{D}_{\text{out}}$ for $\mathcal{C}_{\text{out}}$ on $\mathbf{x}$, obtaining a decoded word $\hat{\mathbf{c}} \in \mathcal{C}_{\text{out}}$, or a FAIL indicator.

  The parameter μ is a threshold that $\mathcal{D}_{\text{in}}$ utilizes to determine whether to attempt correction of a corrupted codeword or declare it erased.

- Let ρ_μ and τ_μ denote, respectively, the number of erasures in $\mathbf{x}(\mu)$ and of non-erased locations j where $x_j \neq \mathcal{E}_{\text{in}}^{-1}(\mathbf{c}_j)$.
  $\mathcal{D}_{\text{out}}$ will reconstruct the original codeword $\mathbf{c} \in \mathcal{C}_{\text{out}}$ if

$$2\tau_\mu + \rho_\mu < D\,. \qquad (*)$$

  We will prove that there exists $\mu \in \{1, 2, \ldots, \lceil d/2 \rceil\}$ such that $(*)$ holds whenever the total number of errors is $T \leq \lfloor (dD-1)/2 \rfloor$.

Decoding of concatenated codes (iii)

- Define, for $\mu \in \{1, 2, \ldots, \lceil d/2 \rceil\}$,

$$\chi_j(\mu) = \begin{cases} 0 & \text{if } \hat{\mathbf{c}}_j = \mathbf{c}_j \text{ and } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ 1 & \text{if } \hat{\mathbf{c}}_j \neq \mathbf{c}_j \text{ and } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ \frac{1}{2} & \text{if } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) \geq \mu \end{cases}.$$

A "decoding penalty" for the jth block given the threshold μ.

- It is readily verified that

$$2\tau_\mu + \rho_\mu = 2 \sum_{j=1}^{N} \chi_j(\mu)\,.$$

Decoding of concatenated codes (iii)

- Define, for $\mu \in \{1, 2, \ldots, \lceil d/2 \rceil\}$,

$$\chi_j(\mu) = \begin{cases} 0 & \text{if } \hat{\mathbf{c}}_j = \mathbf{c}_j \text{ and } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ 1 & \text{if } \hat{\mathbf{c}}_j \neq \mathbf{c}_j \text{ and } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ \frac{1}{2} & \text{if } \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) \geq \mu \end{cases} .$$

A "decoding penalty" for the jth block given the threshold μ.

- It is readily verified that

$$2\tau_\mu + \rho_\mu = 2 \sum_{j=1}^{N} \chi_j(\mu) .$$

- Define the probability measure on μ

$$\mathsf{P}_\mu (\mu = x) = \begin{cases} 2/d & \text{if } x \in \{1, 2, \ldots, \lfloor d/2 \rfloor\} \\ 1/d & \text{if } d \text{ is odd and } x = \lceil d/2 \rceil \end{cases} .$$

Decoding of concatenated codes (iv)

Lemma

For every $j \in \{1, 2, \ldots, N\}$,

$$\mathsf{E}_{\mu}\left\{\chi_j(\mu)\right\} \leq \frac{\mathsf{d}(\mathbf{y}_j, \mathbf{c}_j)}{d} .$$

Decoding of concatenated codes (iv)

$$\chi_j(\mu) = \begin{cases} 0 & \hat{\mathbf{c}}_j = \mathbf{c}_j,\ \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ 1 & \hat{\mathbf{c}}_j \neq \mathbf{c}_j,\ \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) < \mu \\ \frac{1}{2} & \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) \geq \mu \end{cases}, \qquad \mathsf{P}_\mu(x) = \begin{cases} 2/d & 1 \leq x \leq \lfloor d/2 \rfloor \\ 1/d & d \text{ odd},\ x = \lceil d/2 \rceil \end{cases},$$

$$1 \leq \mu \leq \lceil d/2 \rceil.$$

Proof.

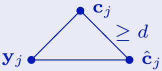
Case 1: $\hat{\mathbf{c}}_j = \mathbf{c}_j$ or $w_j \triangleq \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) \geq d/2$. Here, $\chi_j(\mu)$ takes either the value 0 (when $\mu > w_j$) or $\frac{1}{2}$ (when $\mu \leq w_j$), never the value 1. We have

$$\mathsf{E}_\mu\left\{\chi_j(\mu)\right\} = \tfrac{1}{2}\mathsf{P}_\mu\left\{\mu \leq w_j\right\} \leq \frac{w_j}{d} \overset{\text{def}}{=} \frac{\mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j)}{d} \overset{\text{MLD}}{\leq} \frac{\mathsf{d}(\mathbf{y}_j, \mathbf{c}_j)}{d}.$$

Case 2: $\hat{\mathbf{c}}_j \neq \mathbf{c}_j$ and $w_j < d/2$. Here, $\chi_j(\mu)$ takes the value 1 (when $\mu > w_j$), or $\frac{1}{2}$ (when $\mu \leq w_j$), never the value 0. We have

$$\mathsf{E}_\mu\left\{\chi_j(\mu)\right\} = \tfrac{1}{2}\mathsf{P}_\mu\left\{\mu \leq w_j\right\} + \mathsf{P}_\mu\left\{\mu > w_j\right\} = 1 - \tfrac{1}{2}\,\overbrace{\mathsf{P}_\mu\left\{\mu \leq w_j\right\}}^{w_j < \frac{d}{2} \Rightarrow \text{all} = \frac{2}{d}}$$

$$= 1 - \frac{w_j}{d} = \frac{d - \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j)}{d} \overset{\text{triangle}}{\leq} \frac{\mathsf{d}(\mathbf{y}_j, \mathbf{c}_j)}{d}.$$



Decoding of concatenated codes (iv)

Lemma

For every $j \in \{1, 2, \ldots, N\}$,

$$\mathsf{E}_\mu \{\chi_j(\mu)\} \le \frac{\mathsf{d}(\mathbf{y}_j, \mathbf{c}_j)}{d} .$$

Proof.

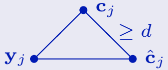
Case 1: $\hat{\mathbf{c}}_j = \mathbf{c}_j$ or $w_j \triangleq \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j) \ge d/2$. Here, $\chi_j(\mu)$ takes either the value 0 (when $\mu > w_j$) or $\frac{1}{2}$ (when $\mu \le w_j$), never the value 1. We have

$$\mathsf{E}_\mu \{\chi_j(\mu)\} = \tfrac{1}{2}\mathsf{P}_\mu \{\mu \le w_j\} \le \frac{w_j}{d} \overset{\text{def}}{=} \frac{\mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j)}{d} \overset{\text{MLD}}{\le} \frac{\mathsf{d}(\mathbf{y}_j, \mathbf{c}_j)}{d} .$$

Case 2: $\hat{\mathbf{c}}_j \ne \mathbf{c}_j$ and $w_j < d/2$. Here, $\chi_j(\mu)$ takes the value 1 (when $\mu > w_j$), or $\frac{1}{2}$ (when $\mu \le w_j$), never the value 0. We have

$$\mathsf{E}_\mu \{\chi_j(\mu)\} = \tfrac{1}{2}\mathsf{P}_\mu \{\mu \le w_j\} + \mathsf{P}_\mu \{\mu > w_j\} = 1 - \tfrac{1}{2} \overbrace{\mathsf{P}_\mu \{\mu \le w_j\}}^{w_j < \frac{d}{2} \Rightarrow \text{ all} = \frac{2}{d}}$$

$$= 1 - \frac{w_j}{d} = \frac{d - \mathsf{d}(\mathbf{y}_j, \hat{\mathbf{c}}_j)}{d} \overset{\text{triangle}}{\le} \frac{\mathsf{d}(\mathbf{y}_j, \mathbf{c}_j)}{d} .$$



Decoding of concatenated codes (v)

Theorem

There exists $\mu \in \{1, 2, \ldots, \lceil d/2 \rceil\}$ *such that* $2\tau_\mu + \rho_\mu < D$.

Decoding of concatenated codes (v)

Theorem

There exists $\mu\in\{1,2,\ldots,\lceil d/2\rceil\}$ *such that* $2\tau_\mu+\rho_\mu<D$.

Proof.

Taking expected values of both sides of $2\tau_\mu+\rho_\mu=2\sum_{j=1}^{N}\chi_j(\mu)$ we obtain

$$\mathsf{E}_\mu\left\{2\tau_\mu+\rho_\mu\right\}=2\sum_{j=1}^{N}\mathsf{E}_\mu\left\{\chi_j(\mu)\right\}\ .$$

By the Lemma, we have

$$2\sum_{j=1}^{N}\mathsf{E}_\mu\left\{\chi_j(\mu)\right\}\le\frac{2}{d}\sum_{j=1}^{N}\mathsf{d}(\mathbf{y}_j,\mathbf{c}_j)=\frac{2\,\mathsf{d}(\mathbf{y},\mathbf{c})}{d}<D\ .$$

Combining the last two equations we obtain

$$\mathsf{E}_\mu\left\{2\tau_\mu+\rho_\mu\right\}<D\ .$$

$\Rightarrow$ There must be at least one $\mu\in\{1,2,\ldots,\lceil d/2\rceil\}$ for which $2\tau_\mu+\rho_\mu<D$.

Forney's Generalized Minimum Distance Decoder (GMD)

Input: received word $\mathbf{y} = (\mathbf{y}_1 \,|\, \mathbf{y}_2 \,|\, \ldots \,|\, \mathbf{y}_N) \in F^{nN}$.
Output: codeword $\mathbf{c} \in \mathcal{C}_{\mathrm{cct}}$ or a decoding-failure indicator FAIL.

1. For $j = 1, 2, \ldots, N$ do:
   - apply a nearest-codeword decoder for $\mathcal{C}_{\mathrm{in}}$ to $\mathbf{y}_j$ to produce $\hat{\mathbf{c}}_j \in \mathcal{C}_{\mathrm{in}}$, corresponding to $\mathbf{z}_j = \mathcal{E}_{\mathrm{in}}^{-1}(\hat{\mathbf{c}}_j) \in \Phi$.
2. For $\mu = 1, 2, \ldots, \lceil d/2 \rceil$ do:
   a. let $\mathbf{x}(\mu) = (\mathbf{x}_1 \ \mathbf{x}_2 \ \ldots \ \mathbf{x}_N)\, (\Phi \cup \{?\})^N$ be as defined in $(\dagger\dagger)$, and let $\rho_\mu \leftarrow |\{j \ : \ \mathbf{x}_j = ?\}|$ // number of erasures in $\mathbf{x}(\mu)$
   b. apply an error-erasure decoder for $\mathcal{C}_{\mathrm{out}}$ to recover ρ_μ erasures and correct up to $\tau_\mu = \lfloor \frac{1}{2}(D-1-\rho_\mu) \rfloor$ errors in $\mathbf{x}$, producing either a codeword
   $$(\mathbf{z}_1\, \mathbf{z}_2\, \ldots\, \mathbf{z}_N) \in \mathcal{C}_{\mathrm{out}}\,, \qquad \text{or} \qquad \mathrm{FAIL}\,;$$
   c. if decoding is successful in Step b then do:
      i. let $\mathbf{c} \leftarrow (\mathcal{E}_{\mathrm{in}}(\mathbf{z}_1)\, \mathcal{E}_{\mathrm{in}}(\mathbf{z}_2)\, \ldots\, \mathcal{E}_{\mathrm{in}}(\mathbf{z}_N))$;
      ii. if $\mathrm{d}(\mathbf{y}, \mathbf{c}) < dD/2$ then output $\mathbf{c}$ and exit.
3. If no codeword $\mathbf{c}$ was produced in Step c then return FAIL.

GMD complexity

- **Step 1**: Brute-force search for closest codeword takes $O(n|\Phi|) = O(nq^k)$. When $N \approx q^k$ (e.g. primitive RS codes), this is $O(nN)$ per block $\mathbf{y}_j$, or overall $O(nN^2)$.
- **Step 2**: Assuming a GRS code is used, Step 2b has complexity $O(ND) = O(N^2)$. Overall for Step 2: $O(dN^2) = O(nN^2)$.
  $\Longrightarrow$ overall complexity is $O(nN^2)$.

GMD complexity

- **Step 1**: Brute-force search for closest codeword takes $O(n|\Phi|) = O(nq^k)$. When $N \approx q^k$ (e.g. primitive RS codes), this is $O(nN)$ per block $\mathbf{y}_j$, or overall $O(nN^2)$.
- **Step 2**: Assuming a GRS code is used, Step 2b has complexity $O(ND) = O(N^2)$. Overall for Step 2: $O(dN^2) = O(nN^2)$.
  $\implies$ overall complexity is $O(nN^2)$.
- Improvements:
  - Step 1 can be done with a syndrome look-up table of size $O(nq^{n-k})$ in time $O(nN)$.
  - Further speed-up is possible by noticing that since $\mu < d/2$, only $(d-1)/2$ decoding is required for $\mathcal{C}_{\text{in}}$. $\mathcal{C}_{\text{in}}$ can be chosen as a code with an efficient decoding algorithm (e.g., Hamming, Golay, BCH, alternant).
  - Step 2 for GRS codes can be accelerated to $(n^3 N \log^2 N \log\log N)$.

Concatenated codes that attain channel capacity

- We will show that it is possible to approach the capacity of the q-ary symmetric channel (QSC) with linear concatenated codes that
  - can be constructed, deterministically, in polynomial time
  - can be encoded and decoded in polynomial time
  - achieve an exponentially decaying probability of decoding error

Concatenated codes that attain channel capacity

- We will show that it is possible to approach the capacity of the q-ary symmetric channel (QSC) with linear concatenated codes that
  - can be constructed, deterministically, in polynomial time
  - can be encoded and decoded in polynomial time
  - achieve an exponentially decaying probability of decoding error

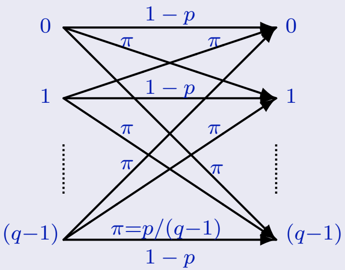
  First, we need to review the basics of *channel capacity*.

Concatenated codes that attain channel capacity

- We will show that it is possible to approach the capacity of the q-ary symmetric channel (QSC) with linear concatenated codes that
  - can be constructed, deterministically, in polynomial time
  - can be encoded and decoded in polynomial time
  - achieve an exponentially decaying probability of decoding error

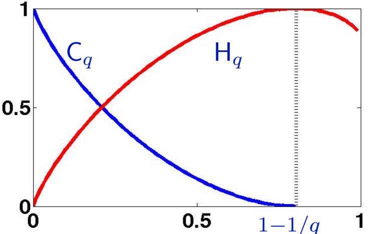
First, we need to review the basics of *channel capacity*.

q-ary symmetric channel (QSC)



For the QSC(p), we have

$$\mathsf{C}_q(p) \triangleq 1 - \mathsf{H}_q(p)\,.$$



Channel capacity—a (very brief) review: Converse theorem

Theorem (Shannon Converse Coding Theorem for the q-ary symmetric channel)

Let $\mathcal{C}$ be an (n, q^{nR}) code over F where n and nR are integers such that $\mathsf{C}_q(p) < R \leq 1$, and let $\mathcal{D} : F^n \to \mathcal{C} \cup \{\text{'E'}\}$ be a decoder for $\mathcal{C}$ over a q-ary symmetric channel with cross-over probability p. Then the decoding error probability P_{err} of $\mathcal{D}$ satisfies

$$P_{\text{err}} \geq 1 - q^{-n(\mathsf{D}_q(\theta_q(R)\|p) - o(1))} ,$$

where

$$\theta_q(R) = \mathsf{H}_q^{-1}(1-R) .$$

- $\mathsf{D}_q(\theta|p) \triangleq \theta \log_q(\frac{\theta}{p}) + (1-\theta)\log_q(\frac{1-\theta}{1-p})$ is the *(information) divergence* or *Kullback-Liebler distance* of θ with respect to p (positive if $\theta \neq p$).
- Notice that $\theta_q(R) = \mathsf{C}_q^{-1}(R)$.
- The theorem says that communication *is impossible at rates above the channel capacity.*

Channel capacity—a (very brief) review: Coding theorem

Theorem (Shannon Coding Theorem for *linear codes* over the q-ary symmetric channel)

Let n and nR be integers such that $R < 1 - \mathsf{H}_q(p)$ and let $\overline{P_{\mathrm{err}}(\mathcal{C})}$ denote the average of $P_{\mathrm{err}}(\mathcal{C})$, under MLD, over all linear $[n, nR]$ codes $\mathcal{C}$ over F. Then,

$$\overline{P_{\mathrm{err}}(\mathcal{C})} < 2q^{-nE_q(p,R)} ,$$

where

$$E_q(p,R) = 1 - \mathsf{H}_q(\theta_q^*(p,R)) - R$$

and

$$\theta_q^*(p,R) = \frac{\log_q(1-p) + 1 - R}{\log_q(1-p) - \log_q(p/(q-1))} .$$

Corollary

For every $\rho \in (0,1]$, all but a fraction at most ρ of the linear $[n, nR]$ codes $\mathcal{C}$ over F satisfy

$$P_{\mathrm{err}}(\mathcal{C}) < (1/\rho) \cdot 2q^{-nE_q(p,R)} .$$

The construction

- We construct a linear concatenated code $\mathcal{C}_{\mathrm{cct}}$.
- Choose, as inner code $\mathcal{C}_{\mathrm{in}}$, an $[n, nr]$ code $\mathcal{C}$ over $F = GF(q)$, with

  $$r < 1 - \mathsf{H}_q(p)$$

  and such that the decoding error probability satisfies

  $$P_{\mathrm{err}}(\mathcal{C}) < 4q^{-nE_q(p,r)}$$

  where $E_q(p, r)$ is the exponent in Shannon's Coding Theorem for linear codes (we sacrifice some error probability to allow for computations with reduced precision—linear in n).
- Let $N_0(n, r, q)$ denote an upper bound on the number of operations over F required to construct $\mathcal{C}$.

The construction (ii)

- Use, as outer code, a linear concatenated code $\mathcal{C}_{\mathrm{out}}$ of length N over $\Phi = GF(q^{rn})$, where $N \geq \max\{N_0, q^{rn}\}$. Furthermore, let $\mathcal{C}_{\mathrm{out}}$ attain Zyablov's bound, and assume its minimum distances satisfies $D_{\mathrm{out}} \geq \lceil \delta N \rceil$ for some real parameter $\delta \in [0, 1]$ (we will determine a relation between r and δ later on).

The construction (ii)

- Use, as outer code, a linear concatenated code $\mathcal{C}_{\mathrm{out}}$ of length N over $\Phi = GF(q^{rn})$, where $N \geq \max\{N_0, q^{rn}\}$. Furthermore, let $\mathcal{C}_{\mathrm{out}}$ attain Zyablov's bound, and assume its minimum distances satisfies $D_{\mathrm{out}} \geq \lceil \delta N \rceil$ for some real parameter $\delta \in [0, 1]$ (we will determine a relation between r and δ later on).
- Given δ, the rate R of $\mathcal{C}_{\mathrm{out}}$ is lower-bounded by

$$R \geq R_Z(\delta, q^{rn})$$

(the choice of δ will be such that R is close to 1).

The construction (ii)

- Use, as outer code, a linear concatenated code $\mathcal{C}_{\mathrm{out}}$ of length N over $\Phi = GF(q^{rn})$, where $N \geq \max\{N_0, q^{rn}\}$. Furthermore, let $\mathcal{C}_{\mathrm{out}}$ attain Zyablov's bound, and assume its minimum distances satisfies $D_{\mathrm{out}} \geq \lceil \delta N \rceil$ for some real parameter $\delta \in [0, 1]$ (we will determine a relation between r and δ later on).
- Given δ, the rate R of $\mathcal{C}_{\mathrm{out}}$ is lower-bounded by

  $$R \geq R_Z(\delta, q^{rn})$$

  (the choice of δ will be such that R is close to 1).
- We will skip the analysis of the encoding complexity, and go directly to decoding (more interesting). The gory details are in Roth (2005).
- Let

  $$\mathbf{y} = (\mathbf{y}_1 \,|\, \mathbf{y}_2 \,|\, \ldots \,|\, \mathbf{y}_N) \in F^{nN}$$

  be the received word, where each $\mathbf{y}_j \in F^n$.

The construction (iii)

Decoder for $\mathcal{C}_{\text{cct}}$

1. Apply a nearest-codeword decoder for $\mathcal{C}_{\text{in}} = \mathcal{C}$ to each sub-block $\mathbf{y}_j$ to produce a codeword $\hat{\mathbf{c}}_j$ of $\mathcal{C}$.
2. Apply a GMD decoder for $\mathcal{C}_{\text{out}}$ to correct up to $\lceil \delta N/2 \rceil$ errors in the word
$$(\mathcal{E}^{-1}(\hat{\mathbf{c}}_1) \,|\, \mathcal{E}^{-1}(\hat{\mathbf{c}}_2) \,|\, \ldots \,|\, \mathcal{E}^{-1}(\hat{\mathbf{c}}_N)) \in \Phi^N$$
(note that $\lceil \delta N/2 \rceil - 1 = \lfloor (\lceil \delta N \rceil - 1)/2 \rfloor$, and recall that $D_{\text{out}} \geq \lceil \delta N \rceil$).

- As before, the decoding operation can be done in time $O((nN)^2)$ (actually less).

Bounding error probability and rate

- **Error probability:** Decoding will fail only if $\tau = \lceil \delta N/2 \rceil$ or more of the sub-blocks $\mathbf{y}_j$ are decoded incorrectly by a nearest-codeword decoder for $\mathcal{C}$. For each sub-block, this probability is $P = P_{\text{err}}(\mathcal{C})$. Hence, recalling that the channel is memoryless, we can write

$$\boxed{P < 4q^{-nE_q(p,r)}}$$

$$\begin{aligned}
P_{\text{err}}(\mathcal{C}_{\text{cct}}) &\le \sum_{i=\tau}^{N} \binom{N}{i} P^i (1-P)^{N-i} \\
&\le \sum_{i=\tau}^{N} \binom{N}{i} P^i \le P^\tau \sum_{i=\tau}^{N} \binom{N}{i} \\
&\le 2^N \cdot P^\tau \le 2^N \cdot P^{N\delta/2} < 4^N q^{-NnE_q(p,r)\delta/2} \\
&\le q^{-nN(E_q(p,r)\delta/2 - o(1))} . \qquad (*)
\end{aligned}$$

Bounding error probability and rate (ii)

- **Rate:** It can be shown that

$$R_Z(\delta, q^{rn}) = (1-\sqrt{\delta})^2 - o(1)/r\,,$$

and, therefore,

$$R_{\rm cct} \ge rR \ge r\cdot(1-\sqrt{\delta})^2 - o(1)\,. \qquad (**)$$

Bounding error probability and rate (ii)

- **Rate:** It can be shown that

$$R_Z(\delta, q^{rn}) = (1-\sqrt{\delta})^2 - o(1)/r\,,$$

and, therefore,

$$R_{\mathrm{cct}} \geq rR \geq r \cdot (1-\sqrt{\delta})^2 - o(1)\,. \qquad (**)$$

- **Error** (from previous slide):

$$P_{\mathrm{err}}(\mathcal{C}_{\mathrm{cct}}) \leq q^{-nN(E_q(p,r)\delta/2 - o(1))}\,. \qquad (*)$$

Bounding error probability and rate (ii)

- **Rate:** It can be shown that

$$R_Z(\delta, q^{rn}) = (1-\sqrt{\delta})^2 - o(1)/r\,,$$

and, therefore,

$$R_{\text{cct}} \geq rR \geq r\cdot(1-\sqrt{\delta})^2 - o(1)\,. \qquad (**)$$

- **Error** (from previous slide):

$$P_{\text{err}}(\mathcal{C}_{\text{cct}}) \leq q^{-nN(E_q(p,r)\delta/2 - o(1))}\,. \qquad (*)$$

- Given a designed rate $\mathcal{R} < 1 - \mathsf{H}_q(p)$, we select the rate r of $\mathcal{C}_{\text{in}}$ so that $\mathcal{R} \leq r \leq 1 - \mathsf{H}_q(p)$ and set δ to

$$\delta = (1-\sqrt{\mathcal{R}/r})^2\,.$$

Therefore, from $(**)$, we have $R_{\text{cct}} \geq \mathcal{R} - o(1)$, while the error exponent in $(*)$ satisfies

$$-\frac{\log_q P_{\text{err}}(\mathcal{C}_{\text{cct}})}{nN} \geq \tfrac{1}{2}E_q(p,r)(1-\sqrt{\mathcal{R}/r})^2 - o(1)\,.$$

Bounding error probability and rate (iii)

- By maximizing over r we obtain

$$-\frac{\log_q P_{\mathrm{err}}(\mathcal{C}_{\mathrm{cct}})}{nN} \geq E_q^*(p, \mathcal{R}) - o(1)\,,$$

where

$$E_q^*(p, \mathcal{R}) = \max_{\mathcal{R} \leq r \leq 1 - \mathsf{H}_q(p)} \tfrac{1}{2} E_q(p, r)(1 - \sqrt{\mathcal{R}/r})^2\ .$$

In particular, $E_q^*(p, \mathcal{R}) > 0$ whenever $\mathcal{R} < 1 - \mathsf{H}_q(p)$.

Summary

Theorem

Let $F = \mathrm{GF}(q)$ and fix a crossover probability $p \in [0, 1-(1/q))$ of a QSC. For every $\mathcal{R} < 1 - \mathsf{H}_q(p)$ there exists an infinite sequence of linear concatenated codes $\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_i, \cdots$ over F such that the following holds.

1. *Each code $\mathcal{C}_i$ is a linear $[n_i, k_i]$ code over F and the values n_i and k_i can be computed from $\mathcal{R}$, q, and i in time complexity that is polynomially large in the length of the bit representations of $\mathcal{R}$, q, i, and n_i.*
2. *The code rates k_i/n_i satisfy*
$$\liminf_{i\to\infty} \frac{k_i}{n_i} \geq \mathcal{R} \ .$$
3. *A generator matrix of $\mathcal{C}_i$ can be constructed in time $O(n_i^2)$, and can be used to encode also in time $O(n_i^2)$.*
4. *There is a decoder for $\mathcal{C}_i$ whose time complexity is $O(n_i^2)$ and its decoding error probability $P_{\mathrm{err}}(\mathcal{C}_i)$ satisfies*
$$-\liminf_{i\to\infty} \frac{1}{n_i} \log_q P_{\mathrm{err}}(\mathcal{C}_i) \geq E_q^*(p, \mathcal{R}) > 0 \ .$$