Capacity Computation and Coding for Input-Constrained Channels

V. Arvind Rameshwar

(guided by) Navin Kashyap

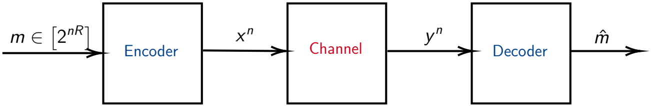
Indian Institute of Science, Bangalore

PhD Defence

November 2023

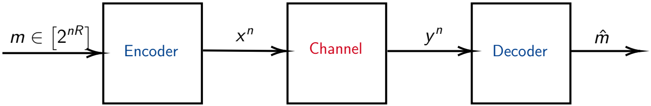
The big picture

- Suppose that we wish to transmit a message m over a noisy medium (or channel):

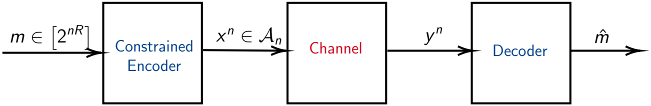


The big picture

- Suppose that we wish to transmit a message m over a noisy medium (or channel):

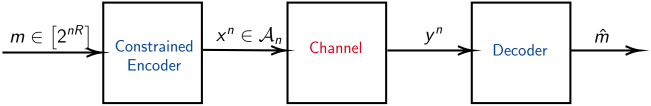


- Suppose also that we require the input x^n to the channel to be "constrained" to lie in $\mathcal{A}_n \subseteq \{0,1\}^n$:



The big picture

- Suppose also that we require the input x^n to the channel to be "constrained" to lie in $\mathcal{A}_n \subseteq \{0,1\}^n$:



- Broad questions for this thesis:

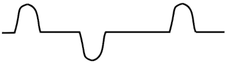
Q1) What do good coding schemes over such channels look like?

Q2) What is the largest rate of reliable codes over such channels, or what is the channel capacity?

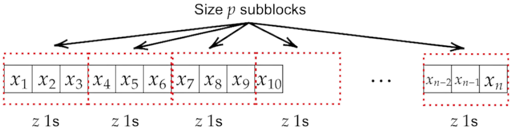
Sample input constraints

- Runlength-limited (RLL) constraints: Help alleviate ISI in magneto-optical recording

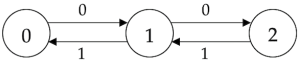
. . . 0 1 0 0 0 1 0 0 0 0 0 1 0 0 . . . $\longleftrightarrow$

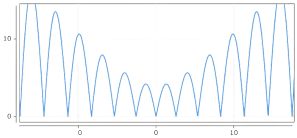


- Subblock composition constraints: Maintain receiver battery levels in energy-harvesting communication



- Charge constraints: Ensure spectral nulls (DC-freeness) in frequency spectrum

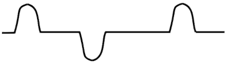




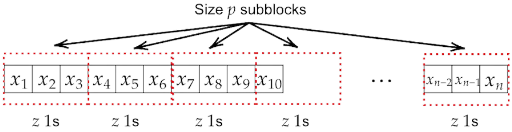
Sample input constraints

- Runlength-limited (RLL) constraints: Help alleviate ISI in magneto-optical recording

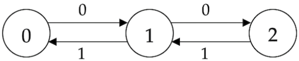
. . . 0 1 0 0 0 1 0 0 0 0 0 1 0 0 . . . $\longleftrightarrow$

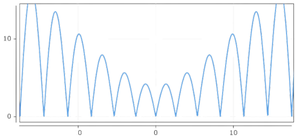


- Subblock composition constraints: Maintain receiver battery levels in energy-harvesting communication



- Charge constraints: Ensure spectral nulls (DC-freeness) in frequency spectrum

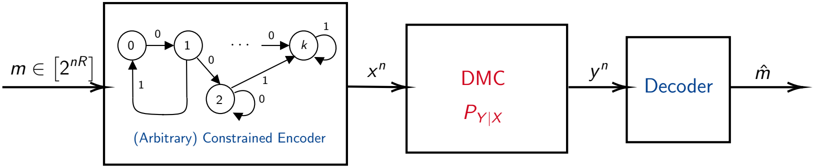




Channel models

Our focus will be on two classes of channels:

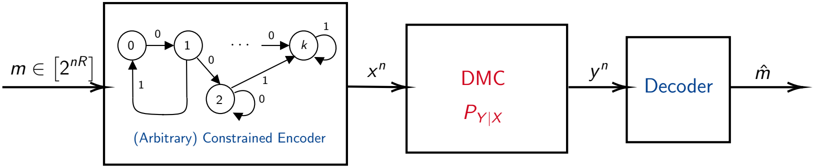
1. Input-constrained discrete memoryless channels (DMCs)



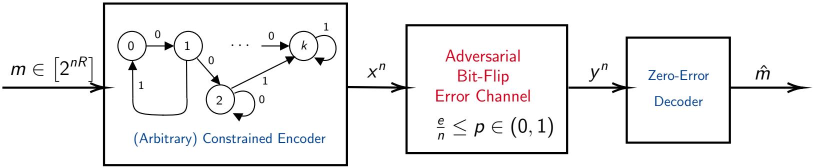
Channel models

Our focus will be on two classes of channels:

1. Input-constrained discrete memoryless channels (DMCs)



2. Input-constrained adversarial noise channels



We deliberate not about ends, but about means. For a doctor does not deliberate whether he shall heal, nor an orator whether he shall persuade [...] they consider how it will be achieved and by what means this will be achieved, until they come to the first cause [...] and what is last in the order of analysis seems to be the first in the order of becoming.

Aristotle, Nicomachean Ethics, Book III, 3, 1112b

*We deliberate not about ends, but about means. For a doctor does not deliberate whether he shall heal, nor an orator whether he shall persuade [. . .] they consider how it will be achieved and by what means this will be achieved, until they come to the first cause [. . .] and **what is last in the order of analysis seems to be the first in the order of becoming**.*

Aristotle, Nicomachean Ethics, Book III, 3, 1112b

Q2) What is the largest rate at which reliable information transfer can happen over such channels, or what is the channel capacity?

Q1) What do good coding schemes over such channels look like?

Summary of contributions

- Bounds on the capacities of input-constrained memoryless channels:
  1. Simple, single-letter lower bounds for input-driven finite-state channels
  2. A stochastic approximation algorithm for the binary erasure channel (BEC) with a no-consecutive-ones input constraint
  3. Upper bounds on the capacity of the (d, ∞)-RLL input-constrained BEC via an explicit characterization of feedback capacity

Summary of contributions

- Bounds on the capacities of input-constrained memoryless channels:
  1. Simple, single-letter lower bounds for input-driven finite-state channels
  2. A stochastic approximation algorithm for the binary erasure channel (BEC) with a no-consecutive-ones input constraint
  3. Upper bounds on the capacity of the (d, ∞)-RLL input-constrained BEC via an explicit characterization of feedback capacity
- Coding schemes for input-constrained memoryless symmetric channels:
  1. Coding schemes using (d, ∞)-RLL constrained subcodes of Reed-Muller (RM) codes
  2. A Fourier-analytic perspective on the sizes of constrained subcodes of general linear codes

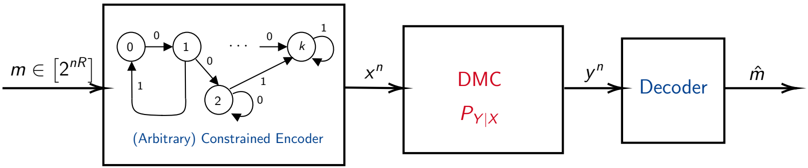
Summary of contributions

- Bounds on the capacities of input-constrained memoryless channels:
  1. Simple, single-letter lower bounds for input-driven finite-state channels
  2. A stochastic approximation algorithm for the binary erasure channel (BEC) with a no-consecutive-ones input constraint
  3. Upper bounds on the capacity of the (d, ∞)-RLL input-constrained BEC via an explicit characterization of feedback capacity
- Coding schemes for input-constrained memoryless symmetric channels:
  1. Coding schemes using (d, ∞)-RLL constrained subcodes of Reed-Muller (RM) codes
  2. A Fourier-analytic perspective on the sizes of constrained subcodes of general linear codes
- Codes for input-constrained adversarial error channels:
  1. Upper bounds on sizes of codes via a version of Delsarte's linear program

Summary of contributions

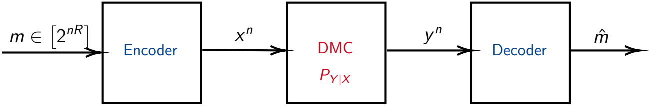
- Bounds on the capacities of input-constrained memoryless channels:
  1. Simple, single-letter lower bounds for input-driven finite-state channels
  2. A stochastic approximation algorithm for the binary erasure channel (BEC) with a no-consecutive-ones input constraint
  3. Upper bounds on the capacity of the (d, ∞)-RLL input-constrained BEC via an explicit characterization of feedback capacity
- Coding schemes for input-constrained memoryless symmetric channels:
  1. Coding schemes using (d, ∞)-RLL constrained subcodes of Reed-Muller (RM) codes
  2. A Fourier-analytic perspective on the sizes of constrained subcodes of general linear codes
- Codes for input-constrained adversarial error channels:
  1. Upper bounds on sizes of codes via a version of Delsarte's linear program

Estimating the capacities of input-constrained memoryless channels



Revisiting the channel model

- For an *unconstrained* DMC,



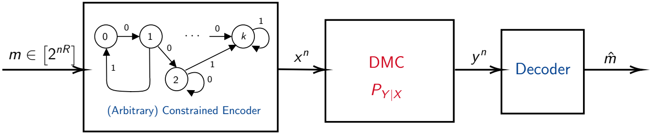
Theorem (Shannon (1948))

The capacity of an unconstrained DMC is

$$C = \max_{\{P(x)\}} I_P(X;Y). \qquad \text{[Single-letter expression!]}$$

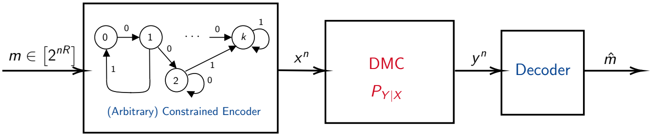
Revisiting the channel model

- We now (re-)introduce our input constraints, represented by (labelled, directed) graphs:



Revisiting the channel model

- We now (re-)introduce our input constraints, represented by (labelled, directed) graphs:

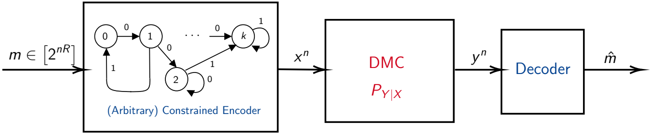


- For the rest of this talk, we assume that the initial "state" of the constrained encoder is fixed and made known to both the encoder and decoder.

- Such channels are a special class of input-driven finite-state channels (FSCs), with a known initial state.

Revisiting the channel model

- We now (re-)introduce our input constraints, represented by (labelled, directed) graphs:



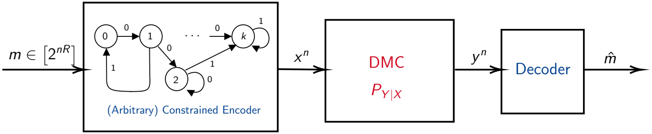
Theorem (Blackwell, Breiman, Thomasian (1958) and Gallager (1968))
The capacity of an input-driven FSC with a fixed, known, initial state s_0 is given by

$$C = \lim_{n\to\infty} \max_{\{P(x^n|s_0)\}} \frac{1}{n} I_P(X^n; Y^n \mid s_0).$$

[Multi-letter expression!]

Revisiting the channel model

- We now (re-)introduce our input constraints, represented by (labelled, directed) graphs:



Theorem (Blackwell, Breiman, Thomasian (1958) and Gallager (1968))

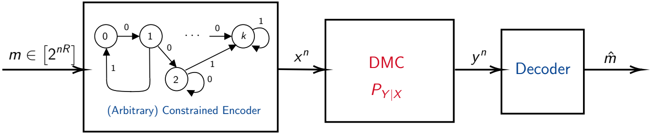
The capacity of an input-driven FSC with a fixed, known, initial state s_0 is given by

$$C = \lim_{n\to\infty} \max_{\{P(x^n|s_0)\}} \frac{1}{n} I_P(X^n; Y^n \mid s_0). \qquad \text{[Multi-letter expression!]}$$

- Explicitly solving for C for general channels is a wide-open problem.
- Evaluating info. rate using simple (Markovian) inputs $\equiv$ Computing entropy rate of a Hidden Markov Process [Hard!]

Revisiting the channel model

- We now (re-)introduce our input constraints, represented by (labelled, directed) graphs:



Theorem (Blackwell, Breiman, Thomasian (1958) and Gallager (1968))

The capacity of an input-driven FSC with a fixed, known, initial state s_0 is given by

$$C = \lim_{n\to\infty} \max_{\{P(x^n|s_0)\}} \frac{1}{n} I_P(X^n; Y^n \mid s_0). \qquad \text{[Multi-letter expression!]}$$

Refining Q2): Can we derive good bounds on the capacities of input-constrained DMCs?

A simple lower bound

- For the broader class of input-driven FSCs (with s_0 known), observe that for a fixed P,

$$I_P(X^n; Y^n \mid s_0) \geq \sum_{t=1}^{N} I(X_t; Y_t \mid X^{t-1}, s_0).$$

A simple lower bound

- For the broader class of input-driven FSCs (with s_0 known), observe that for a fixed P,

$$I_P(X^n; Y^n \mid s_0) \geq \sum_{t=1}^{N} I(X_t; Y_t \mid X^{t-1}, s_0).$$

- Hence,

$$\begin{aligned}
C &= \lim_{n\to\infty} \max_{\{P(x_t|x^{t-1},s_0)\}_{t=1}^{n}} \frac{1}{n} I(X^n; Y^n \mid s_0) \\
&\geq \lim_{n\to\infty} \max_{\{P(x_t|x^{t-1},s_0)\}_{t=1}^{n}} \frac{1}{n} \sum_{t=1}^{n} I(X_t; Y_t \mid X^{t-1}, s_0) \\
&= \sup_{\{P(x_t|x^{t-1})\}_{t\geq 1}} \liminf_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} I(X_t; Y_t \mid X^{t-1}, s_0) \quad \text{[Permuter et al. (2008)]} \\
&\geq \sup_{\{Q(x|s)\in\mathcal{P}\}} I_Q(X; Y|S),
\end{aligned}$$

where $\mathcal{P}$ is the class of input distributions inducing Markov chains on the states of the constraint with an aperiodic, closed, communicating class containing s_0.

Applying the lower bound

A recurring motif: the (d, ∞)-runlength limited (RLL) constraint

A binary sequence is said to satisfy the (d, ∞)-RLL constraint if there exist at least d 0s between every pair of successive 1s.

Applying the lower bound

A recurring motif: the (d, ∞)-runlength limited (RLL) constraint

> A binary sequence is said to satisfy the (d, ∞)-RLL constraint if there exist at least d 0s between every pair of successive 1s.

- For the $(2, \infty)$-RLL constraint,

  1 0 0 0 1 0 0 0 0 1 0 0 1 ✓

  1 0 0 1 0 1 0 0 0 1 0 0 1 X

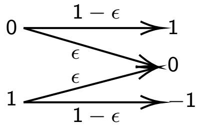
- $(1, \infty)$-RLL $\equiv$ no-consecutive-ones.
- In magneto-optical recording systems, the (d, ∞)-RLL constraint alleviates intersymbol interference (ISI).

Applying the lower bound

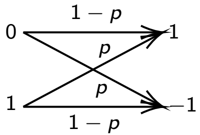
A recurring motif: the (d, ∞)-runlength limited (RLL) constraint

> A binary sequence is said to satisfy the (d, ∞)-RLL constraint if there exist at least d 0s between every pair of successive 1s.

We illustrate lower bounds over (d, ∞)-RLL input-constrained BECs and BSCs:



Binary Erasure Channel (BEC)

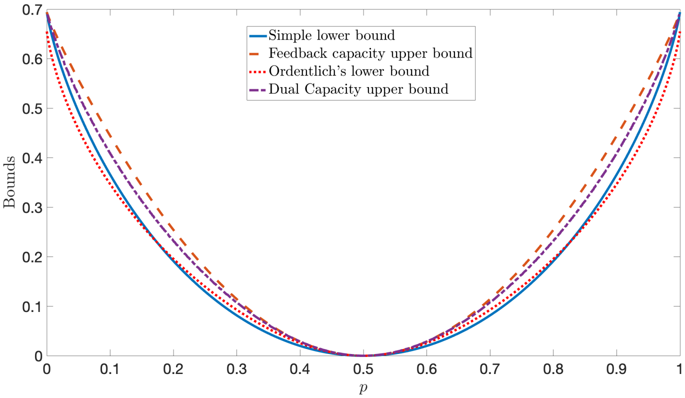


Binary Symmetric Channel (BSC)

Plots and Comparisons: BSC

For the (d, ∞)-RLL input-constrained BSC,

$$C_d(p) \geq \max_{a \in [0,1]} \frac{h_b(ap + \bar{a}\bar{p}) - h_b(p)}{ad + 1}.$$ [Zehavi and Wolf (1988)]

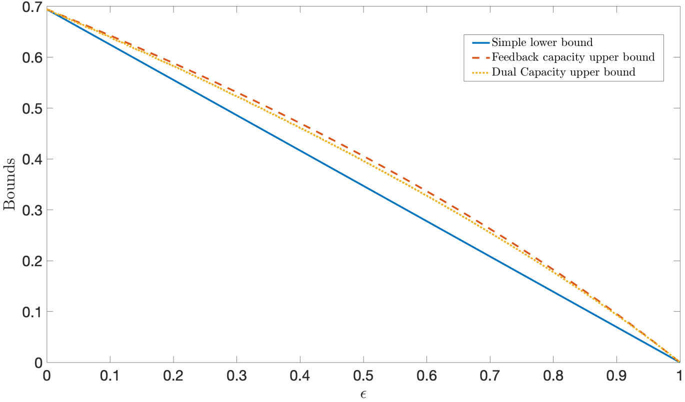


Plots for the $(1, \infty)$-RLL input-constrained BSC

Plots and Comparisons

For the (d, ∞)-RLL input-constrained BEC,

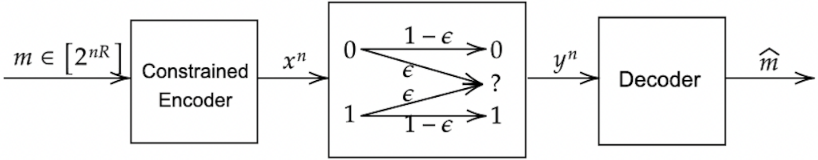
$$C_d(\epsilon) \geq \underbrace{\kappa_d}_{\text{noiseless capacity}} \cdot (1 - \epsilon). \quad [\text{Li and Han (2018)}]$$



Plots for the $(1, \infty)$-RLL input-constrained BEC [$\kappa_1 \approx 0.694$]

Improving the lower bound for a special case

We shall now work towards deriving better bounds for the $(1, \infty)$-RLL input-constrained BEC.



- The channel is the BEC.
- The codewords satisfy the $(1, \infty)$-RLL input constraint.

Key ideas

- We restrict the input process $(X_i)_{i\geq 1}$ to (again) be first-order Markov and ergodic, with $X_0 \sim \pi$ (stat. dist.). Then,

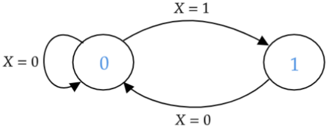
$$
\begin{aligned}
C &= \lim_{N\to\infty} \max_{\{P(x^N)\}} \frac{1}{N} I(X^N; Y^N) \\
&\geq \sup_{\{P(x|x^-)\}} \lim_{N\to\infty} \frac{1}{N} I(X^N; Y^N) =: \underbrace{C_f}_{\text{first-order cap.}} .
\end{aligned}
$$

Key ideas

- We restrict the input process $(X_i)_{i\geq 1}$ to (again) be first-order Markov and ergodic, with $X_0 \sim \pi$ (stat. dist.). Then,

$$C = \lim_{N\to\infty} \max_{\{P(x^N)\}} \frac{1}{N} I(X^N; Y^N)$$

$$\geq \sup_{\{P(x|x^-)\}} \lim_{N\to\infty} \frac{1}{N} I(X^N; Y^N) =: \underbrace{C_f}_{\text{first-order cap.}} .$$



is the input Markov process

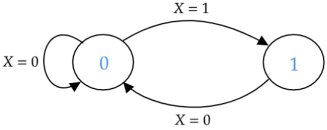
$$Q = \begin{bmatrix} 1-a & a \\ 1 & 0 \end{bmatrix}$$

is the transition probability matrix.

Key ideas

- We restrict the input process $(X_i)_{i\geq 1}$ to (again) be first-order Markov and ergodic, with $X_0 \sim \pi$ (stat. dist.). Then,

$$C = \lim_{N\to\infty} \max_{\{P(x^N)\}} \frac{1}{N} I(X^N; Y^N)$$
$$\geq \sup_{\{P(x|x^-)\}} \lim_{N\to\infty} \frac{1}{N} I(X^N; Y^N) =: \underbrace{C_f}_{\text{first-order cap.}} .$$



is the input Markov process

$$Q = \begin{bmatrix} 1-a & a \\ 1 & 0 \end{bmatrix}$$

is the transition probability matrix.

> Given such an input distribution, we identify an associated Markov process $(L_i, \tilde{X}_i, Y_i)_{i\geq 1}$, which succintly encapsulates the information in Y^N.

A useful theorem

Theorem

The first order capacity C_f is given by (see also [Li and Han (2018)])

$$C_f(\epsilon) = (1-\epsilon) \cdot \max_{a \in (0,1)} \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left[\mathsf{H}\left(a Q^{(L_i - 1)}(\tilde{X}_i, 0) \right) \right].$$

A useful theorem

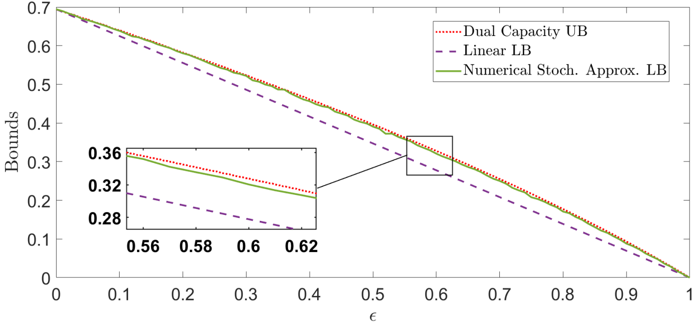
Theorem

The first order capacity C_f is given by (see also [Li and Han (2018)])

$$C_f(\epsilon) = (1-\epsilon) \cdot \max_{a \in (0,1)} \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left[\mathrm{H}\left(a Q^{(L_i - 1)}(\tilde{X}_i, 0) \right) \right].$$

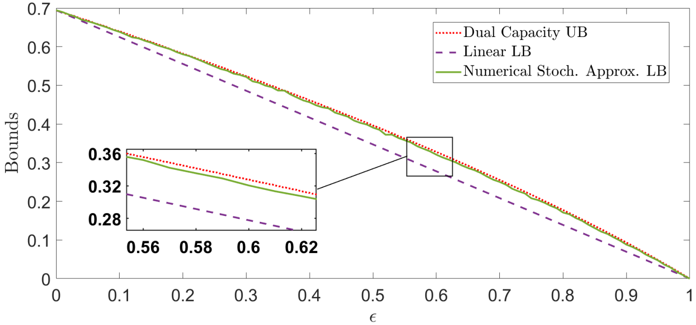
We then devise a two-timescale stochastic approximation algorithm for approximately computing C_f.

Plots and Comparisons



The simple linear lower bound equals $\kappa_1(1-\epsilon)$, where $\kappa_1 \approx 0.694$.

Plots and Comparisons

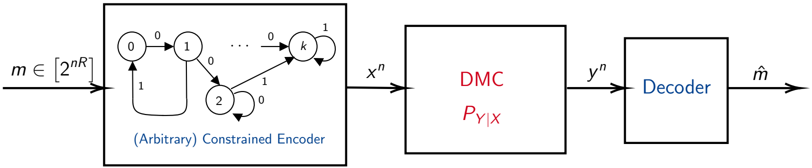


The simple linear lower bound equals $\kappa_1(1-\epsilon)$, where $\kappa_1 \approx 0.694$.

Moving to Q1): With some knowledge about capacity estimates, we shall proceed to constructing error-correcting constrained codes.

Coding Schemes Over Input-Constrained Channels

Part I: Input-Constrained Memoryless Channels



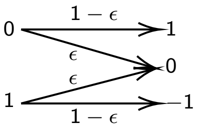
Background on BMS channels

In this talk, we restrict our attention to binary-input memoryless symmetric (BMS) channels:

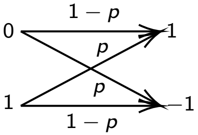
$$Y = (-1)^X \cdot Z,$$

for noise Z independent of X.

Examples:



Binary Erasure Channel (BEC)



Binary Symmetric Channel (BSC)

BMS channels and linear codes

- Suppose that we were to use a linear code $\mathcal{C}$ over the BMS channel.
- Under (optimal) block-MAP decoding, the block-error probabilities are independent of the codeword transmitted.
- Hence, constrained subcodes of $\mathcal{C}$ have the same (average) error probabilities as $\mathcal{C}$ itself!

BMS channels and linear codes

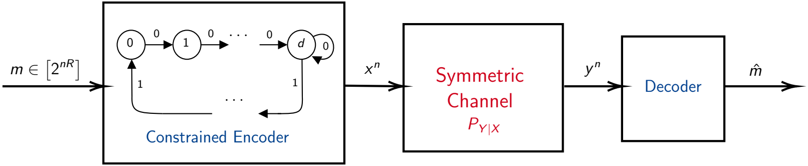
- Suppose that we were to use a linear code $\mathcal{C}$ over the BMS channel.
- Under (optimal) block-MAP decoding, the block-error probabilities are independent of the codeword transmitted.
- Hence, constrained subcodes of $\mathcal{C}$ have the same (average) error probabilities as $\mathcal{C}$ itself!

Our approach: Use constrained subcodes of capacity-achieving codes.

[cf. Abbe & Sandon (2023), Arikan (2009), . . .]

A recurring task: Compute/estimate the rates of constrained subcodes of linear codes.

Designing Coding Schemes Over (d,∞)-RLL Input-Constrained $\underbrace{\text{BMS}}_{\text{binary-input memoryless symmetric}}$ Channels



Selected results

Theorem

For any $R \in (0,1)$, there exists an explicit sequence of (d,∞)-RLL linear subcodes $\left\{\mathcal{C}_m^{(d)}\right\}_{m\geq 1}$ of a sequence of RM codes of rate R such that

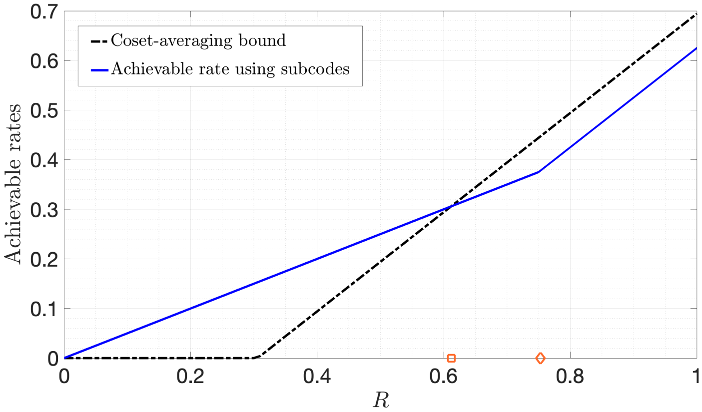
$$\mathit{rate}\left(\mathcal{C}_m^{(d)}\right) \xrightarrow{m\to\infty} R \cdot 2^{-\lceil \log_2(d+1)\rceil}.$$

Theorem

For any $R \in (0,1)$, there exists a sequence of $(1,\infty)$-RLL subcodes $\left\{\hat{\mathcal{C}}_m^{(d)}\right\}_{m\geq 1}$ of a sequence of RM codes of rate R such that

$$\mathit{rate}\left(\hat{\mathcal{C}}_m^{(d)}\right) \xrightarrow{m\to\infty} \max\left(0, R - \frac{3}{8}\right).$$

Plots and Comparisons - I



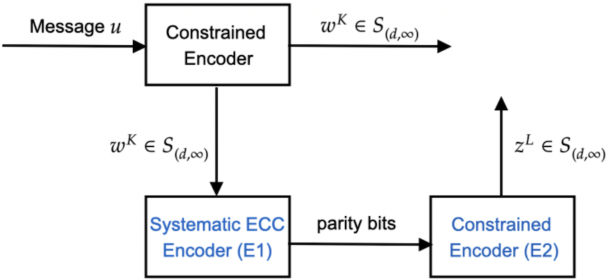
Plot comparing the achievable rates using $(1,\infty)$-RLL RM subcodes with the coset-averaging lower bound that is approximately $R - 0.3058$, of [Patapoutian and Kumar (1992)]

Selected results: a concatenated coding scheme

We adopt the "reverse concatenation" strategy of [Bliss (1981)] and [Mansuripur (1991)] that is commonly used to limit error propagation during decoding of constrained codes.

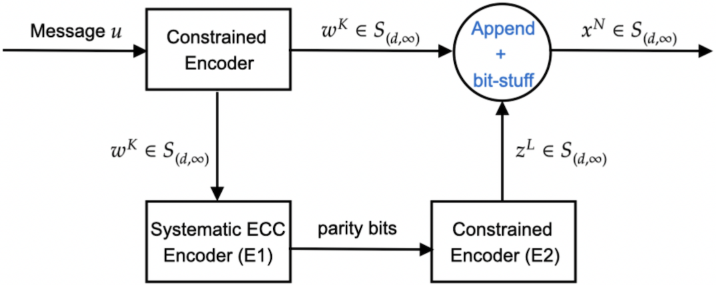
Selected results: a concatenated coding scheme

Encoding (the Bliss scheme):



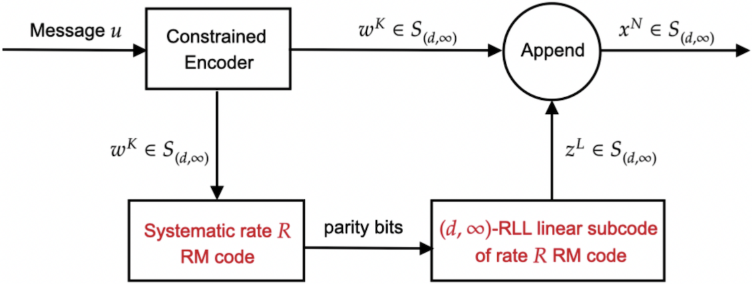
Selected results: a concatenated coding scheme

Encoding (the Bliss scheme):



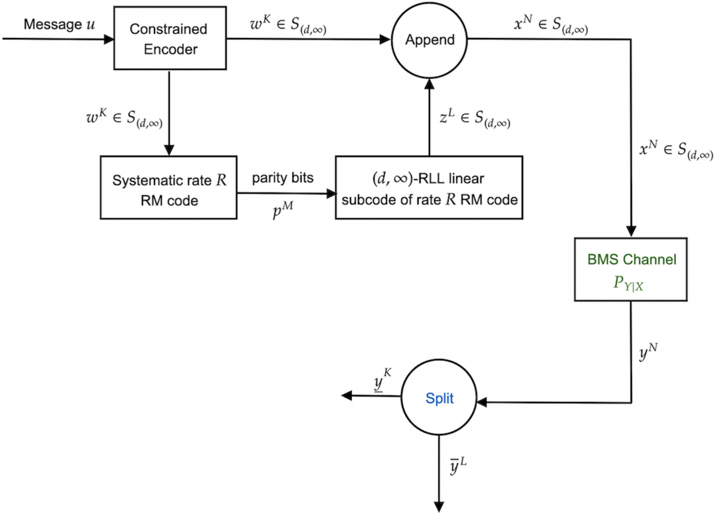
Selected results: a concatenated coding scheme

Encoding:



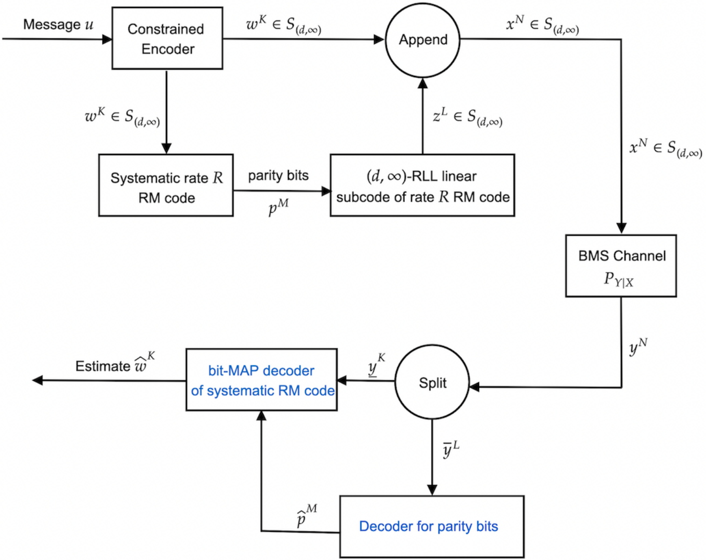
Selected results: a concatenated coding scheme

Encoding+Decoding:



Selected results: a concatenated coding scheme

Encoding+Decoding:



Coding theorem

Let C be the capacity of the unconstrained BMS channel.

Slight modifications to the previous concatenated coding scheme yield the following coding theorem:

Coding theorem

Let C be the capacity of the unconstrained BMS channel.

Slight modifications to the previous concatenated coding scheme yield the following coding theorem:

Theorem

For any $R \in (0, C)$, there exists a sequence of (d, ∞)-RLL constrained concatenated codes $\{\mathcal{C}_m^{conc}\}_{m \geq 1}$ that achieves a rate lower bound given by

$$\liminf_{m \to \infty} rate(\mathcal{C}_m^{conc}) \geq \frac{\kappa_d \cdot R^2 \cdot 2^{-\lceil \log_2(d+1) \rceil}}{R^2 \cdot 2^{-\lceil \log_2(d+1) \rceil} + 1 - R + \epsilon},$$

over (d, ∞)-RLL input-constrained BMS channels, where $\epsilon > 0$ can be arbitrarily small.

Plots and Comparisons - II

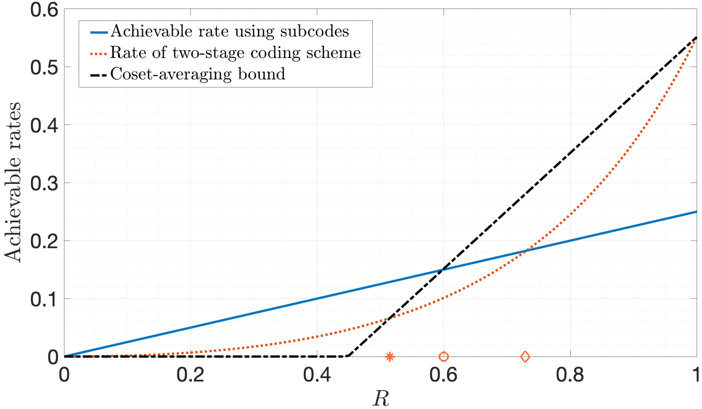


Figure: Plot comparing the achievable rates using $(2, \infty)$-RLL linear RM subcodes with the lower bound via the probabilistic method and the rate achieved by the concatenated coding scheme

Plots and Comparisons - II

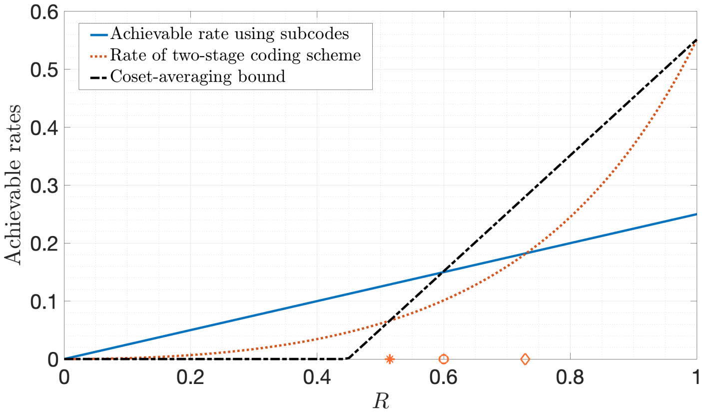
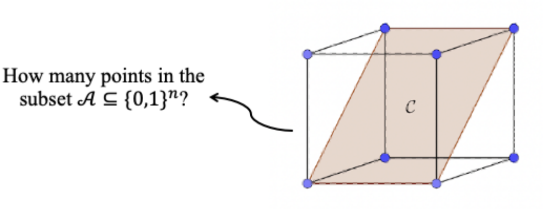


Figure: Plot comparing the achievable rates using $(2, \infty)$-RLL linear RM subcodes with the lower bound via the probabilistic method and the rate achieved by the concatenated coding scheme

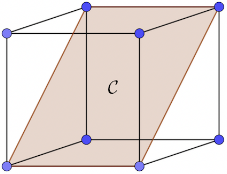
Continuing with Q1): How do we identify constrained subcodes of general linear codes, for arbitrary constraints?

Counting constrained codewords in general linear codes



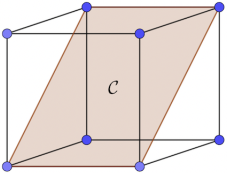
The problem

- Motivated by the previous section, we now consider the problem of computing rates of (arbitrarily-)constrained codewords in linear codes $\mathcal{C}$.



The problem

- Motivated by the previous section, we now consider the problem of computing rates of (arbitrarily-)constrained codewords in linear codes $\mathcal{C}$.

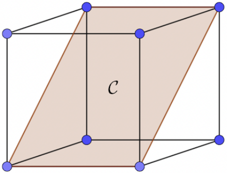


- The problem: Given a set of constrained codewords $\mathcal{A} \subseteq \mathbb{F}_2^n$, we would like to gain insight into

$$N(\mathcal{C};\mathcal{A}) = \sum_{\boldsymbol{x}\in\mathcal{C}} \mathbb{1}\{\boldsymbol{x}\in\mathcal{A}\} = \sum_{\boldsymbol{x}\in\{0,1\}^n} \mathbb{1}\{\boldsymbol{x}\in\mathcal{A}\}\cdot\mathbb{1}\{\boldsymbol{x}\in\mathcal{C}\}.$$

The problem

- Motivated by the previous section, we now consider the problem of computing rates of (arbitrarily-)constrained codewords in linear codes $\mathcal{C}$.



- The problem: Given a set of constrained codewords $\mathcal{A} \subseteq \mathbb{F}_2^n$, we would like to gain insight into

$$N(\mathcal{C}; \mathcal{A}) = \sum_{\boldsymbol{x} \in \mathcal{C}} \mathbb{1}\{\boldsymbol{x} \in \mathcal{A}\} = \sum_{\boldsymbol{x} \in \{0,1\}^n} \mathbb{1}\{\boldsymbol{x} \in \mathcal{A}\} \cdot \mathbb{1}\{\boldsymbol{x} \in \mathcal{C}\}.$$

This looks like an inner product between Boolean functions!

A brief refresher on Fourier analysis on $\mathbb{F}_2^n$

- Given any function $f : \{0,1\}^n \to \mathbb{R}$ and a vector $\boldsymbol{s} = (s_1, \ldots, s_n) \in \{0,1\}^n$, the Fourier coefficient of f at $\boldsymbol{s}$ is

$$\widehat{f}(\boldsymbol{s}) := \frac{1}{2^n} \sum_{\boldsymbol{x} \in \{0,1\}^n} f(\boldsymbol{x}) \cdot (-1)^{\boldsymbol{x} \cdot \boldsymbol{s}}.$$

- An inner product $\langle f, g \rangle$ between $f, g : \{0,1\}^n \to \mathbb{R}$ can be defined as

$$\langle f, g \rangle := \frac{1}{2^n} \sum_{\boldsymbol{x} \in \{0,1\}^n} f(\boldsymbol{x}) g(\boldsymbol{x}).$$

Theorem (Plancherel's Theorem)

For any $f, g \in \{0,1\}^n \to \mathbb{R}$, we have that

$$\langle f, g \rangle = \sum_{\boldsymbol{s} \in \{0,1\}^n} \widehat{f}(\boldsymbol{s}) \widehat{g}(\boldsymbol{s}).$$

Workhorse

- Observe that

$$N(\mathcal{C};\mathcal{A}) = 2^n \cdot \sum_{\boldsymbol{s}\in\{0,1\}^n} \widehat{\mathbb{1}_{\mathcal{A}}}(\boldsymbol{s}) \cdot \widehat{\mathbb{1}_{\mathcal{C}}}(\boldsymbol{s}).$$

Workhorse

- Observe that

$$N(\mathcal{C};\mathcal{A}) = 2^n \cdot \sum_{\boldsymbol{s}\in\{0,1\}^n} \widehat{\mathbb{1}_{\mathcal{A}}}(\boldsymbol{s}) \cdot \widehat{\mathbb{1}_{\mathcal{C}}}(\boldsymbol{s}).$$

- For linear codes $\mathcal{C}$, it is easy to show that

$$\widehat{\mathbb{1}_{\mathcal{C}}}(\boldsymbol{s}) = \frac{|\mathcal{C}|}{2^n} \cdot \mathbb{1}_{\mathcal{C}^\perp}(\boldsymbol{s}).$$

- Hence,

$$N(\mathcal{C};\mathcal{A}) = |\mathcal{C}| \cdot \sum_{\boldsymbol{s}\in\mathcal{C}^\perp} \widehat{\mathbb{1}_{\mathcal{A}}}(\boldsymbol{s}).$$

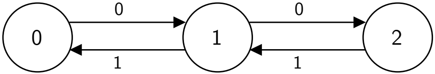
Workhorse

- Observe that
$$N(\mathcal{C};\mathcal{A}) = 2^n \cdot \sum_{\boldsymbol{s}\in\{0,1\}^n} \widehat{\mathbb{1}_{\mathcal{A}}}(\boldsymbol{s}) \cdot \widehat{\mathbb{1}_{\mathcal{C}}}(\boldsymbol{s}).$$
- For linear codes $\mathcal{C}$, it is easy to show that
$$\widehat{\mathbb{1}_{\mathcal{C}}}(\boldsymbol{s}) = \frac{|\mathcal{C}|}{2^n} \cdot \mathbb{1}_{\mathcal{C}^\perp}(\boldsymbol{s}).$$
- Hence,
$$N(\mathcal{C};\mathcal{A}) = |\mathcal{C}| \cdot \sum_{\boldsymbol{s}\in\mathcal{C}^\perp} \widehat{\mathbb{1}_{\mathcal{A}}}(\boldsymbol{s}).$$

1. If $\dim(\mathcal{C}) \gg n/2$, then we can employ our insight to count in a low-dimensional space!
2. For many constraints of interest, the Fourier transform above is analytically/numerically computable!

Example 1: 2-charge constraint

- We consider a spectral null constraint, whose sequences in $\{+1, -1\}^n$ have a null at zero frequency.
- We let S_2 denote those sequences in $\{0, 1\}^n$ that can be mapped to 2-charge constrained sequences via the map $x \mapsto (-1)^x$, for $x \in \{0, 1\}$.



Sequences in S_2 can be read off the labels of paths here.

Computation of Fourier coefficients and consequences

Theorem

There exists a vector space V such that for $\boldsymbol{s} \in V$,

$$\widehat{\mathbb{1}_{S_2}}(\boldsymbol{s}) = 2^{\lfloor \frac{n}{2} \rfloor - n} \cdot (-1)^{\gamma(\boldsymbol{s})},$$

where $\gamma : \{0,1\}^n \to \{0,1\}$. Further, for $\boldsymbol{s} \notin V$, we have $\widehat{\mathbb{1}_{S_2}}(\boldsymbol{s}) = 0$.

Computation of Fourier coefficients and consequences

Theorem

There exists a vector space V such that for $\boldsymbol{s} \in V$,

$$\widehat{\mathbb{1}_{S_2}}(\boldsymbol{s}) = 2^{\lfloor \frac{n}{2} \rfloor - n} \cdot (-1)^{\gamma(\boldsymbol{s})},$$

where $\gamma : \{0,1\}^n \to \{0,1\}$. Further, for $\boldsymbol{s} \notin V$, we have $\widehat{\mathbb{1}_{S_2}}(\boldsymbol{s}) = 0$.

We use this theorem to construct a sequence $\{\mathcal{C}^{(n)}\}_{n \geq 1}$ of linear codes of rate R such that the rate of their S_2-constrained subcodes obeys

$$\liminf_{n \to \infty} \text{rate}\left(\mathcal{C}_2^{(n)}\right) > R - \frac{1}{2}.$$

We thus obtain rates better than the coset-averaging bound of [Patapoutian and Kumar (1992)], using subcodes!

Computation of Fourier coefficients and consequences

Theorem

There exists a vector space V such that for $\boldsymbol{s} \in V$,

$$\widehat{\mathbb{1}_{S_2}}(\boldsymbol{s}) = 2^{\lfloor \frac{n}{2} \rfloor - n} \cdot (-1)^{\gamma(\boldsymbol{s})},$$

where $\gamma : \{0,1\}^n \to \{0,1\}$. Further, for $\boldsymbol{s} \notin V$, we have $\widehat{\mathbb{1}_{S_2}}(\boldsymbol{s}) = 0$.

We can also use the theorem to count S_2-constrained codewords in well-known linear codes:

(m,r)	$(5,3)$	$(6,4)$	$(7,5)$	$(8,6)$
$N(\mathrm{RM}(m,r); S_2)$	2048	6.711×10^7	1.441×10^{17}	1.329×10^{36}

Some sample numerical values for high rate RM codes

Example 2: (d, ∞)-RLL constraint

- Recall:

  (d, ∞)-RLL $\equiv$ at least d 0s b/w successive 1s

- Let S^d denote the set of constrained sequences and $\widehat{\mathbb{1}_{S^d}}^{(n)}$ denote the Fourier transform at blocklength $n \geq 1$.

Example 2: (d,∞)-RLL constraint

- Recall:

$$(d,\infty)\text{-RLL} \equiv \text{at least } d \text{ 0s b/w successive 1s}$$

- Let S^d denote the set of constrained sequences and $\widehat{\mathbb{1}_{S^d}}^{(n)}$ denote the Fourier transform at blocklength $n \geq 1$.

Theorem

For $n \geq d+2$ and for $\mathbf{s} = (s_1, \ldots, s_n) \in \{0,1\}^n$,

$$\widehat{\mathbb{1}_{S^d}}^{(n)}(\mathbf{s}) = 2^{-1} \cdot \widehat{\mathbb{1}_{S^d}}^{(n-1)}(s_2^n) + (-1)^{s_1} \cdot 2^{-(d+1)} \cdot \widehat{\mathbb{1}_{S^d}}^{(n-d-1)}\left(s_{d+2}^n\right).$$

Example 2: (d, ∞)-RLL constraint

- Recall:

$$(d, \infty)\text{-RLL} \equiv \text{at least } d \text{ 0s b/w successive 1s}$$

- Let S^d denote the set of constrained sequences and $\widehat{\mathbb{1}_{S^d}}^{(n)}$ denote the Fourier transform at blocklength $n \geq 1$.

Theorem

For $n \geq d+2$ and for $\mathbf{s} = (s_1, \ldots, s_n) \in \{0,1\}^n$,

$$\widehat{\mathbb{1}_{S^d}}^{(n)}(\mathbf{s}) = 2^{-1} \cdot \widehat{\mathbb{1}_{S^d}}^{(n-1)}(s_2^n) + (-1)^{s_1} \cdot 2^{-(d+1)} \cdot \widehat{\mathbb{1}_{S^d}}^{(n-d-1)}\left(s_{d+2}^n\right).$$

The recursive procedure arising from the above theorem is faster for computing Fourier transforms than the Fast Walsh-Hadamard Transform!

Example 2: (d, ∞)-RLL constraint

- Recall:

$$(d, \infty)\text{-RLL} \equiv \text{at least } d \text{ 0s b/w successive 1s}$$

- Let S^d denote the set of constrained sequences and $\widehat{\mathbb{1}_{S^d}}^{(n)}$ denote the Fourier transform at blocklength $n \geq 1$.

Theorem

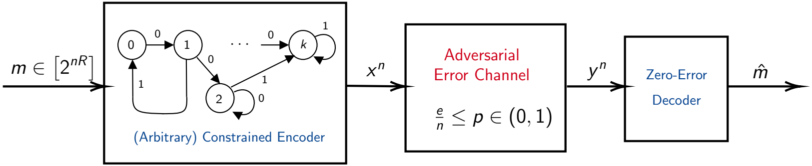
For $n \geq d+2$ and for $\mathbf{s} = (s_1, \ldots, s_n) \in \{0,1\}^n$,

$$\widehat{\mathbb{1}_{S^d}}^{(n)}(\mathbf{s}) = 2^{-1} \cdot \widehat{\mathbb{1}_{S^d}}^{(n-1)}(s_2^n) + (-1)^{s_1} \cdot 2^{-(d+1)} \cdot \widehat{\mathbb{1}_{S^d}}^{(n-d-1)}\left(s_{d+2}^n\right).$$

Similar recurrence relations can also be proved for the flash memory ("no-101") constraint and a version of the even constraint, which requires that the length of every run of 0s be even.

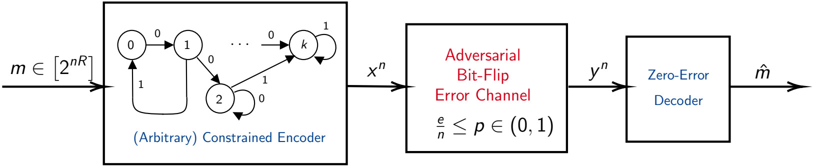
Coding Schemes Over Input-Constrained Channels

Part II: Input-Constrained Adversarial Error Channels



The problem

Consider the transmission of constrained sequences over a combinatorial bit-flip error channel:



- The error-correcting capability of a constrained code over such a channel is determined by its minimum Hamming distance.

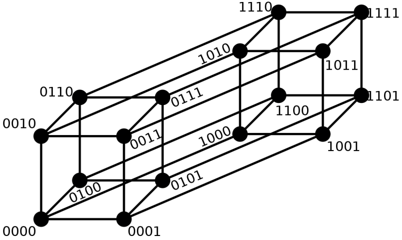
Min. Hamming dist. is d $\iff$ Bounded dist. decoder can correct $\approx d/2$ errors and detect $\approx d$ errors

What is the largest size of a t-error correcting constrained code?

A more formal description

Fix a blocklength $n \geq 1$. Suppose that we are given a constraint represented by the set $\mathcal{A} \subseteq \{0,1\}^n$ of constrained sequences.

What is the size of the largest subset of $\mathcal{A}$ having minimum Hamming distance at least d?

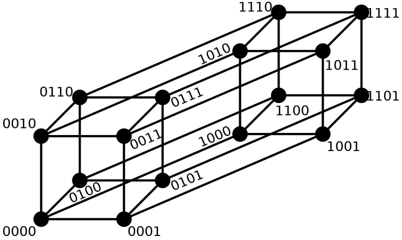


The 4-dimensional binary Hamming space

A more formal description

Fix a blocklength $n \geq 1$. Suppose that we are given a constraint represented by the set $\mathcal{A} \subseteq \{0,1\}^n$ of constrained sequences.

What is the size of the largest subset of $\mathcal{A}$ having minimum Hamming distance at least d?



The 4-dimensional binary Hamming space

Let us call the size of a largest collection $A(n, d; \mathcal{A})$.

When $\mathcal{A} = \{0,1\}^n$, we call this size simply $A(n, d)$.

Flashback: Delsarte's LP (for unconstrained systems)

Consider the LP $\mathrm{Del}(n, d)$:

$$\begin{aligned}
\text{maximize} \quad & \sum_{w=0}^{n} a_w \\
\text{subj. to} \quad & a_w \geq 0, \ \text{for all } w \in [0:n], \\
& \sum_{j=0}^{n} a_j \cdot K_w(j) \geq 0, \ \text{for all } w \in [0:n], \\
& a_w = 0, \ \text{for } w \in [1:d-1], \\
& a_0 = 1,
\end{aligned}$$

where $K_w = K_w^{(n)}$ is the w^{th}-Krawtchouk polynomial at length n. Here,

$$K_w(j) = \sum_{\ell=0}^{w} (-1)^{\ell} \binom{j}{\ell} \binom{n-j}{w-\ell}.$$

Flashback: Delsarte's LP (for unconstrained systems)

Consider the LP Del(n, d):

$$\begin{aligned}
\text{maximize} \quad & \sum_{w=0}^{n} a_w \\
\text{subj. to} \quad & a_w \geq 0, \ \text{for all } w \in [0:n], \\
& \sum_{j=0}^{n} a_j \cdot K_w(j) \geq 0, \ \text{for all } w \in [0:n], \\
& a_w = 0, \ \text{for } w \in [1:d-1], \\
& a_0 = 1,
\end{aligned}$$

- From [Delsarte (1973)]: a feasible solution is the distance distribution $(b_w : 0 \leq w \leq n)$ of *any* binary length-n code $\mathcal{C}$ of minimum distance at least d.
- Objective value of solution is $|\mathcal{C}|$.
- Hence, $A(n, d) \leq \mathrm{OPT}(\mathrm{Del}(n, d))$.

Our LP for constrained systems

Fix $\mathcal{A} \subseteq \{0,1\}^n$. Consider the LP $\overline{\mathsf{Del}}(n, d; \mathcal{A})$:

Our LP for constrained systems

Fix $\mathcal{A} \subseteq \{0,1\}^n$. Consider the LP $\overline{\mathrm{Del}}(n, d; \mathcal{A})$:

$$\underset{f\colon \{0,1\}^n \to \mathbb{R}}{\text{maximize}} \quad \sum_{\boldsymbol{x} \in \{0,1\}^n} f(\boldsymbol{x})$$

subject to:

$$f(\boldsymbol{x}) \geq 0, \ \forall \ \boldsymbol{x} \in \{0,1\}^n,$$

$$\widehat{f}(\boldsymbol{s}) \geq 0, \ \forall \ \boldsymbol{s} \in \{0,1\}^n,$$

$$f(\boldsymbol{x}) = 0, \ \text{if } 1 \leq w(\boldsymbol{x}) \leq d-1,$$

$$f(0^n) \leq \mathrm{OPT}(\mathrm{Del}(n,d)), \tag{C1}$$

$$f(\boldsymbol{x}) \leq 2^n \cdot (\mathbb{1}_{\mathcal{A}} \star \mathbb{1}_{\mathcal{A}})(\boldsymbol{x}), \ \forall \ \boldsymbol{x} \in \{0,1\}^n. \tag{C2}$$

Our LP for constrained systems

Fix $\mathcal{A} \subseteq \{0,1\}^n$. Consider the LP $\overline{\mathsf{Del}}(n, d; \mathcal{A})$:

$$\underset{f\colon \{0,1\}^n \to \mathbb{R}}{\text{maximize}} \quad \sum_{\boldsymbol{x} \in \{0,1\}^n} f(\boldsymbol{x})$$

subject to:

$$f(\boldsymbol{x}) \geq 0, \ \forall\ \boldsymbol{x} \in \{0,1\}^n,$$

$$\widehat{f}(\boldsymbol{s}) \geq 0, \ \forall\ \boldsymbol{s} \in \{0,1\}^n,$$

$$f(\boldsymbol{x}) = 0, \ \text{if } 1 \leq w(\boldsymbol{x}) \leq d-1,$$

$$f(0^n) \leq \mathrm{OPT}(\mathsf{Del}(n,d)), \tag{C1}$$

$$f(\boldsymbol{x}) \leq 2^n \cdot (\mathbb{1}_{\mathcal{A}} \star \mathbb{1}_{\mathcal{A}})(\boldsymbol{x}), \ \forall\ \boldsymbol{x} \in \{0,1\}^n. \tag{C2}$$

- Here, $f = 2^n \cdot (\mathbb{1}_{\mathcal{C} \cap \mathcal{A}} \star \mathbb{1}_{\mathcal{C} \cap \mathcal{A}})$ is feasible, with

$$\mathrm{val}(f) = |\mathcal{C} \cap \mathcal{A}|^2.$$

Our LP for constrained systems

Fix $\mathcal{A} \subseteq \{0,1\}^n$. Consider the LP $\overline{\mathrm{Del}}(n, d; \mathcal{A})$:

$$\underset{f\colon \{0,1\}^n \to \mathbb{R}}{\text{maximize}} \quad \sum_{\boldsymbol{x} \in \{0,1\}^n} f(\boldsymbol{x})$$

subject to:

$$f(\boldsymbol{x}) \geq 0, \ \forall \ \boldsymbol{x} \in \{0,1\}^n,$$

$$\widehat{f}(\boldsymbol{s}) \geq 0, \ \forall \ \boldsymbol{s} \in \{0,1\}^n,$$

$$f(\boldsymbol{x}) = 0, \ \text{if } 1 \leq w(\boldsymbol{x}) \leq d-1,$$

$$f(0^n) \leq \mathrm{OPT}(\mathrm{Del}(n,d)), \tag{C1}$$

$$f(\boldsymbol{x}) \leq 2^n \cdot (\mathbb{1}_{\mathcal{A}} \star \mathbb{1}_{\mathcal{A}})(\boldsymbol{x}), \ \forall \ \boldsymbol{x} \in \{0,1\}^n. \tag{C2}$$

Theorem

$$A(n, d; \mathcal{A}) \leq (\mathrm{OPT}(\overline{\mathrm{Del}}(n, d; \mathcal{A})))^{1/2}.$$

Example 1: (d, ∞)-RLL constraint

Let us run $\overline{\mathrm{Del}}(n, d; S^1)$ with $n = 10$[1]:

d	$\overline{\mathrm{Del}}(n, d; S^1)$	$\mathrm{GenSph}(n, d; S^1)$	$\mathrm{Del}(n, d)$
3	74.762	111	85.333
4	42.048	111	42.667
5	12	63	12
6	6	63	6
7	3.2	26	3.2

[1] The upper bounds can be rounded down to yield integral bounds on code sizes.

Example 1: (d, ∞)-RLL constraint

... and $\overline{\mathrm{Del}}(n, d; S^2)$ with $n = 10$[1]:

d	$\overline{\mathrm{Del}}(n, d; S^2)$	$\mathrm{GenSph}(n, d; S^2)$	$\mathrm{Del}(n, d)$
3	32.075	46.5	85.333
4	21.721	46.5	42.667
5	7.856	34	12
6	4.899	34	6
7	2.529	19	3.2

[1] The upper bounds can be rounded down to yield integral bounds on code sizes.

Example 1: (d, ∞)-RLL constraint

… and $\overline{\mathrm{Del}}(n, d; S^2)$ with $n = 10$[1]:

d	$\overline{\mathrm{Del}}(n, d; S^2)$	$\mathrm{GenSph}(n, d; S^2)$	$\mathrm{Del}(n, d)$
3	32.075	46.5	85.333
4	21.721	46.5	42.667
5	7.856	34	12
6	4.899	34	6
7	2.529	19	3.2

1. We have that

$$(\mathrm{OPT}(\overline{\mathrm{Del}}(n, d; \mathcal{A})))^{1/2} \leq \min\{\mathrm{OPT}(\mathrm{Del}(n, d)), |\mathcal{A}|\}.$$

2. The computability of the Fourier transforms $\widehat{\mathbb{1}_{\mathcal{A}}}$ can be harnessed to compute $\mathbb{1}_{\mathcal{A}} \star \mathbb{1}_{\mathcal{A}}(\boldsymbol{x})$.
3. For many constraints, the LP can be symmetrized to yield non-trivial savings in complexity.

[1] The upper bounds can be rounded down to yield integral bounds on code sizes.

Open questions for further research

Open questions

- Is it possible to prove (analytically) that the capacity of a (d, ∞)-RLL input-constrained BSC(p) obeys

$$C_d(p) \geq \kappa_d(1 - h_b(p)) \text{ ?} \quad [\text{Wolf's Conjecture (1988)}]$$

- Can one derive asymptotic upper bounds on the rate-distance tradeoff for constrained codes, using our LP formulation?

- Can we use data-driven methods to construct good codes for other channels with memory? [2]

[2] Our recent work on sampling-based methods for approximately computing the sizes of "small" subcodes of RM codes is at https://arxiv.org/abs/2309.08907.

List of publications

Journal

1. V. A. R. and Navin Kashyap, "Estimating the sizes of binary error-correcting constrained codes," accepted to the IEEE Journal on Selected Areas in Information Theory, May 2023.
2. V. A. R. and Navin Kashyap, "Coding schemes based on Reed-Muller codes for (d, ∞)-RLL input-constrained channels," in IEEE Transactions on Information Theory, vol. 69, no. 11, pp. 7003-7024, Nov. 2023.

List of publications

Conference

1. V. A. R. and Navin Kashyap, "A version of Delsarte's linear program for constrained systems," accepted to the 2023 IEEE International Symposium on Information Theory (ISIT). **Recipient of a Jack Keil Wolf ISIT Student Paper Award**.
2. V. A. R. and Navin Kashyap, "Counting constrained codewords in binary linear codes via Fourier expansions," accepted to the 2023 IEEE International Symposium on Information Theory (ISIT).
3. V. A. R. and Navin Kashyap, "Linear runlength-limited subcodes of Reed-Muller codes and coding schemes for input-constrained BMS channels," 2022 IEEE Information Theory Workshop (ITW), Nov. 2022.
4. V. A. R. and Navin Kashyap, "A feedback capacity-achieving coding scheme for the (d, ∞)-RLL input-constrained binary erasure channel," 2022 IEEE International Conference on Signal Processing and Communications (SPCOM), Jul. 2022. **Recipient of a Best Student Paper Award**.
5. V. A. R. and Navin Kashyap, "On the performance of Reed-Muller codes Over (d, ∞)-RLL input-constrained BMS channels," 2022 IEEE International Symposium on Information Theory (ISIT), Espoo, Finland, Jun. 2022.

List of publications

Conference

6. V. A. R. and Navin Kashyap, “Numerically computable lower bounds on the capacity of the $(1, \infty)$-RLL input-constrained binary erasure channel,” 2021 National Conference on Communications (NCC), Jul. 2021. **Recipient of a Best Paper Award**.
7. V. A. R. and Navin Kashyap, “Bounds on the feedback capacity of the (d, ∞)-RLL input constrained binary erasure channel,” in 2021 IEEE International Symposium on Information Theory (ISIT), Jul. 2021.
8. V. A. R., Aryabhatt M. Reghu, and Navin Kashyap, “On the capacity of the flash memory channel with feedback,” in 2020 International Symposium on Information Theory and its Applications (ISITA2020), Kapolei, USA, Oct. 2020.
9. V. A. R. and Navin Kashyap, “Computable lower bounds for capacities of input-driven finite-state channels,” in 2020 IEEE International Symposium on Information Theory (ISIT 2020), Los Angeles, California, USA, Jun. 2020.

Tools employed

Mostly Standard

1. Information theory
2. Error-control coding
3. Dynamic programming and Markov decision processes

Tools employed

Mostly Standard

1. Information theory
2. Error-control coding
3. Dynamic programming and Markov decision processes

Mostly Non-Standard

A. Stochastic approximation
B. Fourier analysis
C. Sampling from high-dimensional spaces [outside this thesis]

Tools employed

Mostly Standard

1. Information theory
2. Error-control coding
3. Dynamic programming and Markov decision processes

Mostly Non-Standard

A. Stochastic approximation
B. Fourier analysis
C. Sampling from high-dimensional spaces [outside this thesis]

"*Lending perspective via cross-disciplinary connections*"

Acknowledgements


VIEW OF
RESORTS WORLD
SENTOSA
圣淘沙名胜世界
リゾート・ワールド・
セントーサ
ARBORA
dusk


SAMSUNG

Prof. Henry Pfister (Duke U.), for the stimulating discussions

My UG research advisers, Prof. Lalitha Vadlamani of IIIT-H (L) and Prof. Prajakta Nimbhorkar of CMI (R), for initiating me into research

Labmates and Veena ma'am, for the bright and lively atmosphere in the lab

Thank You!