Bellman meets Shannon: On Dynamic Programming and Capacity Computation

V. Arvind Rameshwar

Department of Electrical Communication Engineering
Indian Institute of Science, Bengaluru

Networks Seminar Series, June 2020

FSCs and Capacity

FSCs and Capacity

- Discrete Finite-State Channels (FSCs) are used to model:

FSCs and Capacity

- Discrete Finite-State Channels (FSCs) are used to model:
  - Inter-Symbol Interference in Magnetic and Optical Recording [e.g., Immink, Siegel, Wolf, '98]
  - Inter-Cell Interference in NAND Flash Memories [e.g., Li, Kavčić, Han, '16]
  - Fading in Mobile Radio Channels [e.g., Semmar, Lecours, Chouinard, Ahern, '91]

FSCs and Capacity

- Discrete Finite-State Channels (FSCs) are used to model:
    - Inter-Symbol Interference in Magnetic and Optical Recording [e.g., Immink, Siegel, Wolf, '98]
    - Inter-Cell Interference in NAND Flash Memories [e.g., Li, Kavčić, Han, '16]
    - Fading in Mobile Radio Channels [e.g., Semmar, Lecours, Chouinard, Ahern, '91]
- Computing Capacities of FSCs

FSCs and Capacity

- Discrete Finite-State Channels (FSCs) are used to model:
  - Inter-Symbol Interference in Magnetic and Optical Recording [e.g., Immink, Siegel, Wolf, '98]
  - Inter-Cell Interference in NAND Flash Memories [e.g., Li, Kavčić, Han, '16]
  - Fading in Mobile Radio Channels [e.g., Semmar, Lecours, Chouinard, Ahern, '91]
- Computing Capacities of FSCs $\rightarrow$ Capacity-achieving coding schemes

Hurdles and Goals

Hurdles and Goals

Without Feedback:

- Computing Capacity is equivalent to computing $H(\mathcal{Y})$
- Hard, for even simple input-constrained DMCs with Markovian input

With Feedback:

- Pros: Simple recipe for computing FB capacity
- Cons: Computationally intractable

Hurdles and Goals

Without Feedback:

- Computing Capacity is equivalent to computing $H(\mathcal{Y})$
- Hard, for even simple input-constrained DMCs with Markovian input

With Feedback:

- Pros: Simple recipe for computing FB capacity
- Cons: Computationally intractable

- Twin Goals:

Hurdles and Goals

Without Feedback:

- Computing Capacity is equivalent to computing $H(\mathcal{Y})$
- Hard, for even simple input-constrained DMCs with Markovian input

With Feedback:

- Pros: Simple recipe for computing FB capacity
- Cons: Computationally intractable

- Twin Goals:
  1. Get good bounds on capacity, without feedback
  2. Numerically evaluate feedback capacity to suggest simple coding schemes

Detour: An Introduction to DP

Detour: An Introduction to DP

Detour: An Introduction to DP

- States: $z_t \in \mathcal{Z}$

Detour: An Introduction to DP

- States: $z_t \in \mathcal{Z}$
- Action: $u_t \in \mathcal{U}$;
  Policy: $\pi = \{u_1, u_2, \ldots\}$

Detour: An Introduction to DP

- States: $z_t \in \mathcal{Z}$
- Action: $u_t \in \mathcal{U}$;
  Policy: $\pi = \{u_1, u_2, \ldots\}$
- Disturbance:
  $w_t \sim P_W(\cdot|z_{t-1}, u_t)$

Detour: An Introduction to DP

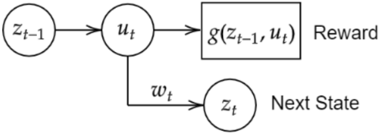
- States: $z_t \in \mathcal{Z}$
- Action: $u_t \in \mathcal{U}$;
  Policy: $\pi = \{u_1, u_2, \ldots\}$
- Disturbance:
  $w_t \sim P_W(\cdot|z_{t-1}, u_t)$
- Dynamics:
  $z_t = F(z_{t-1}, u_t, w_t)$

Detour: An Introduction to DP

- States: $z_t \in \mathcal{Z}$
- Action: $u_t \in \mathcal{U}$;
  Policy: $\pi = \{u_1, u_2, \ldots\}$
- Disturbance:
  $w_t \sim P_W(\cdot | z_{t-1}, u_t)$
- Dynamics:
  $z_t = F(z_{t-1}, u_t, w_t)$
- Reward: $r_t = g(z_{t-1}, u_t)$

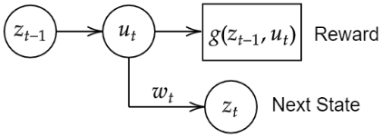
Detour: An Introduction to DP

- States: $z_t \in \mathcal{Z}$
- Action: $u_t \in \mathcal{U}$;
  Policy: $\pi = \{u_1, u_2, \ldots\}$
- Disturbance:
  $w_t \sim P_W(\cdot | z_{t-1}, u_t)$
- Dynamics:
  $z_t = F(z_{t-1}, u_t, w_t)$
- Reward: $r_t = g(z_{t-1}, u_t)$



Detour: An Introduction to DP

- States: $z_t \in \mathcal{Z}$
- Action: $u_t \in \mathcal{U}$;
  Policy: $\pi = \{u_1, u_2, \ldots\}$
- Disturbance:
  $w_t \sim P_W(\cdot|z_{t-1}, u_t)$
- Dynamics:
  $z_t = F(z_{t-1}, u_t, w_t)$
- Reward: $r_t = g(z_{t-1}, u_t)$



$$\text{Calculate } \rho^* = \sup_{\pi} \liminf_{N\to\infty} \frac{1}{N} \mathbb{E}_{\pi} \left[\sum_{t=1}^{N} g(Z_{t-1}, U_t) \right]$$

The Bellman Equation

The Bellman Equation

Theorem (Bellman Equation)

If $\rho \in \mathbb{R}$ *and a bounded function* $h : \mathcal{Z} \to \mathbb{R}$ *satisfies for all* $z \in \mathcal{Z}$,

$$\rho + h(z) = \sup_{u \in \mathcal{U}} \left[g(z,u) + \sum_{w} P(w|z,u) h(F(z,u,w)) \right],$$

then $\rho = \rho^*$.

The Bellman Equation

Theorem (Bellman Equation)

If $\rho \in \mathbb{R}$ and a bounded function $h : \mathcal{Z} \to \mathbb{R}$ satisfies for all $z \in \mathcal{Z}$,

$$\rho + h(z) = \sup_{u \in \mathcal{U}} \left[g(z,u) + \sum_{w} P(w|z,u) h(F(z,u,w)) \right],$$

then $\rho = \rho^*$.

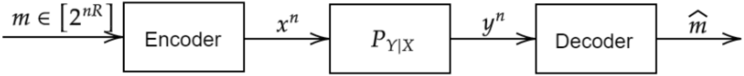
Also helps identify optimal stationary policy

FSCs Without Feedback

The Setup

The Setup

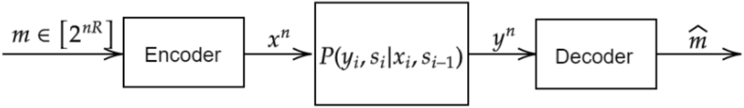
- DMC:



$$P(y^n|x^n) = \prod_{i=1}^{n} P(y_i|x_i)$$

The Setup

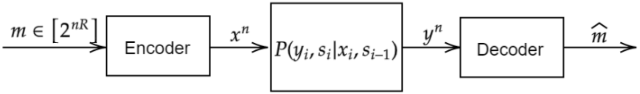
- Generic FSC:



$$P(y^n, s^n|x^n, s_0) = \prod_{i=1}^{n} P(y_i, s_i|s_{i-1}, x_i)$$

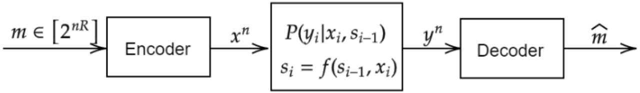
The Setup

- Generic FSC:



$$P(y^n, s^n|x^n, s_0) = \prod_{i=1}^{n} P(y_i, s_i|x_i, s_{i-1})$$

- Input-Driven FSC:

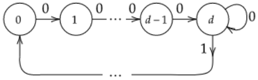


$$P(y^n, s^n|x^n, s_0) = \prod_{i=1}^{n} P(y_i|x_i, s_{i-1}) \mathbb{1}\{s_i = f(s_{i-1}, x_i)\}$$

Examples

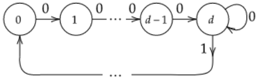
Examples

- Input-Constrained DMCs:
  - (d, ∞)-RLL input constraint:

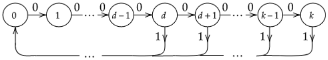


Examples

- Input-Constrained DMCs:
  - (d, ∞)-RLL input constraint:



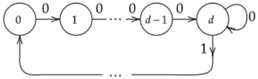
  - (d, k)-RLL input constraint:



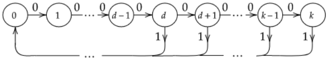
Examples

- Other Channels:

- Input-Constrained DMCs:
  - (d, ∞)-RLL input constraint:

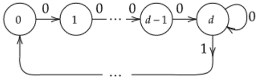


  - (d, k)-RLL input constraint:

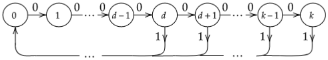


Examples

- Input-Constrained DMCs:
  - (d, ∞)-RLL input constraint:



  - (d, k)-RLL input constraint:

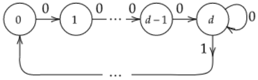


- Other Channels:
  - Flash-Memory Channel $(101 \to 111$ w.p. $\epsilon)$:

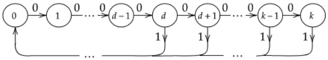
$$s_i = (x_i, x_{i-1})$$

Examples

- Input-Constrained DMCs:
  - (d, ∞)-RLL input constraint:



  - (d, k)-RLL input constraint:



- Other Channels:
  - Flash-Memory Channel ($101 \to 111$ w.p. ϵ):

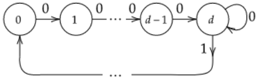
$$s_i = (x_i, x_{i-1})$$

- ISI: $y_i = \sum_{k=0}^{m} h_k x_{i-k} + z_i$

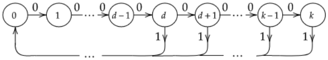
$$s_i = (x_{i-m}, \ldots, x_{i-1})$$

Examples

- Input-Constrained DMCs:
  - (d, ∞)-RLL input constraint:



  - (d, k)-RLL input constraint:



- Other Channels:
  - Flash-Memory Channel ($101 \to 111$ w.p. ϵ):

$$s_i = (x_i, x_{i-1})$$

- ISI: $y_i = \sum_{k=0}^{m} h_k x_{i-k} + z_i$

$$s_i = (x_{i-m}, \ldots, x_{i-1})$$

Assumption: s_0 known to encoder and decoder

Summary of Results

Summary of Results

$$C = \lim_{N\to\infty} \max_{Q(x^N|s_0)} \frac{1}{N} I(X^N; Y^N|s_0)$$

Summary of Results

$$C = \lim_{N \to \infty} \max_{Q(x^N|s_0)} \frac{1}{N} I(X^N; Y^N|s_0)$$

Upper Bounds

- FB Capacity [Sabag et al., '16, '18]:
  1. $(1, \infty)$-RLL input-constrained BEC
  2. $(1, \infty)$-RLL input-constrained BIBO
- Dual-Capacity [Thangaraj, '17]
- Dual-Capacity + DP [Huleihel et al., '19]

Summary of Results

$$C = \lim_{N\to\infty} \max_{Q(x^N|s_0)} \frac{1}{N} I(X^N; Y^N|s_0)$$

Upper Bounds

- FB Capacity [Sabag et al., '16, '18]:
  1. $(1,\infty)$-RLL input-constrained BEC
  2. $(1,\infty)$-RLL input-constrained BIBO
- Dual-Capacity [Thangaraj, '17]
- Dual-Capacity + DP [Huleihel et al., '19]

Lower Bounds

- Simulation-Based:
  1. M-C [Arnold et al., '06]
  2. GBA [Vontobel et al., '08]
  3. Stoch. Approx. [Han, '15]
- Analytical:
  1. Asymptotics of BSC, BEC [Han and Marcus, '09] [Li and Han, '18]
  2. Markov inputs for BSC [Zehavi and Wolf, '88]
  3. General input-driven (this paper)

Key Ideas

Key Ideas

Recall:

⋆ Reverse Directed Information:

$$I(Y^N \to X^N | s_0) = \sum_{t=1}^{N} I(Y^t; X_t | X^{t-1}, s_0)$$

⋆ (Delayed) Forward Directed Information:

$$I(X^{N-1} \to Y^N | s_0) = \sum_{t=1}^{N} I(X^{t-1}; Y_t | Y^{t-1}, s_0)$$

Ideas (Contd. . .)

Ideas (Contd. . .)

Theorem (Conservation Law (Massey, '05))

$$I(X^N; Y^N|s_0) = I(Y^N \to X^N|s_0) + I(X^{N-1} \to Y^N|s_0)$$
$$\geq I(Y^N \to X^N|s_0)$$

Ideas (Contd. . .)

Theorem (Conservation Law (Massey, '05))

$$I(X^N; Y^N|s_0) = I(Y^N \to X^N|s_0) + I(X^{N-1} \to Y^N|s_0)$$
$$\geq I(Y^N \to X^N|s_0)$$

Can we massage this LB into a computable expression?

Ideas (Contd...)

Theorem (Conservation Law (Massey, '05))

$$I(X^N; Y^N|s_0) = I(Y^N \to X^N|s_0) + I(X^{N-1} \to Y^N|s_0)$$
$$\geq I(Y^N \to X^N|s_0)$$

Can we massage this LB into a computable expression?

Yes!

Steps Towards DP Formulation

Steps Towards DP Formulation

- **Step 1**:

Steps Towards DP Formulation

- **Step 1**:

  $$I(X^N; Y^N|s_0) \geq I(Y^N \to X^N|s_0)$$

Steps Towards DP Formulation

- **Step 1**:

$$I(X^N; Y^N|s_0) \geq I(Y^N \to X^N|s_0)$$
$$= \sum_{t=1}^{N} I(Y^t; X_t|X^{t-1}, s_0)$$

Steps Towards DP Formulation

- **Step 1**:

$$I(X^N; Y^N|s_0) \geq I(Y^N \to X^N|s_0)$$
$$\geq \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0) \quad \text{[Input-driven channel]}$$

Steps Towards DP Formulation

- **Step 1**:

$$I(X^N; Y^N|s_0) \geq I(Y^N \to X^N|s_0)$$
$$\geq \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0) \quad \text{[Input-driven channel]}$$

- **Step 2**:

Steps Towards DP Formulation

- **Step 1**:

$$I(X^N; Y^N|s_0) \geq I(Y^N \to X^N|s_0)$$
$$\geq \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0) \quad \text{[Input-driven channel]}$$

- **Step 2**:

$$C = \lim_{N\to\infty} \max_{Q(x^N|s_0)} \frac{1}{N} I(X^N; Y^N|s_0)$$

Steps Towards DP Formulation

- **Step 1**:

$$I(X^N; Y^N|s_0) \geq I(Y^N \to X^N|s_0)$$

$$\geq \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0) \quad \text{[Input-driven channel]}$$

- **Step 2**:

$$C = \lim_{N\to\infty} \max_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^{N}} \frac{1}{N} I(X^N; Y^N|s_0)$$

Steps Towards DP Formulation

- **Step 1**:

$$I(X^N; Y^N|s_0) \geq I(Y^N \rightarrow X^N|s_0)$$
$$\geq \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0) \quad \text{[Input-driven channel]}$$

- **Step 2**:

$$C = \lim_{N\to\infty} \max_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^N} \frac{1}{N} I(X^N; Y^N|s_0)$$
$$\geq \lim_{N\to\infty} \max_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^N} \frac{1}{N} \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0)$$

Steps Towards DP Formulation (Contd. . .)

Steps Towards DP Formulation (Contd. . .)

- **Step 3**:

Steps Towards DP Formulation (Contd. . .)

- **Step 3**:

$$C \geq \lim_{N\to\infty} \max_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^{N}} \frac{1}{N}\sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0)$$

Steps Towards DP Formulation (Contd. . .)

- **Step 3**:

$$C \geq \lim_{N\to\infty} \max_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^N} \frac{1}{N}\sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0)$$

$$= \sup_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^N} \liminf_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0) \text{ [P. et al., '08]}$$

Steps Towards DP Formulation (Contd. . .)

- **Step 3**:

$$C \geq \lim_{N\to\infty} \max_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^N} \frac{1}{N}\sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0)$$

$$= \sup_{\{Q(x_t|x^{t-1},s_0)\}_{t=1}^N} \liminf_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0) \text{ [P. et al., '08]}$$

$$\geq \sup_{\{Q(x_t|s_{t-1})\}_{t=1}^N} \liminf_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t|X^{t-1}, s_0)$$

Lower Bound as a DP Problem

Lower Bound as a DP Problem

$$C \geq \sup_{\{Q(x_t|s_{t-1})\}} \liminf_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t | X^{t-1}, s_0)$$

Lower Bound as a DP Problem

$$C \geq \sup_{\{Q(x_t|s_{t-1})\}} \liminf_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t | X^{t-1}, s_0)$$

	DP Notation	Our Instance
State	z_{t-1}	$P(s_{t-1}\|x^{t-1})$
Action	u_t	$Q(x_t\|s_{t-1})$
Disturbance	w_t	x_t
Reward	$g(z_{t-1}, u_t)$	$I(X_t, S_{t-1}; Y_t\|x^{t-1}, s_0)$

Remarks on Simplifying the DP Approach

Remarks on Simplifying the DP Approach

- It is possible to show that under some mild conditions,

$$C \geq \sup_{\{Q(x|s) \in \mathcal{P}\}} I(X;Y|S), \text{ where}$$

$$\mathcal{P} \triangleq \Big\{\{Q(x|s)\} : \text{M.C. on } \mathcal{S} \text{ induced by } Q \text{ is aperiodic and has a unique stationary distribution}\Big\}.$$

Remarks on Simplifying the DP Approach

- It is possible to show that under some mild conditions,

$$C \geq \sup_{\{Q(x|s) \in \mathcal{P}\}} I(X;Y|S), \text{ where}$$

$$\mathcal{P} \triangleq \Big\{ \{Q(x|s)\} : \text{M.C. on } \mathcal{S} \text{ induced by } Q \text{ is aperiodic and has a unique stationary distribution} \Big\}.$$

- Derived lower bound can be achieved directly by random coding, where the inputs are generated according to $Q^*(x_i|s_{i-1})$ (where $Q^* \in \mathcal{P}$ maximizes $I(X;Y|S)$), and the decoder performs ML decoding.

Remarks on the BEC(ϵ)

Remarks on the BEC(ϵ)

- For the BEC(ϵ) with an input constraint presented by an irreducible deterministic graph on $\mathcal{S}$,

$$\begin{aligned} I(X;Y|S) &= H(X|S) - H(X|Y,S) \\ &= H(X|S) - \epsilon H(X|Y=?,S) \\ &= H(X|S)(1-\epsilon). \end{aligned}$$

  Hence, $\sup_{\{Q(x|s)\in\mathcal{P}\}} I(X;Y|S) = C_0(1-\epsilon)$, where C_0 is the noiseless capacity of the input-constraint.

Application: (d, ∞)-RLL Input-Constrained BEC(ϵ)

Application: (d,∞)-RLL Input-Constrained BEC(ϵ)

Theorem

The capacity of the (d,∞)-RLL input-constrained binary erasure channel with erasure probability ϵ satisfies

$$C_{d,\infty}^{BEC(\epsilon)} \geq C_{d,\infty} \cdot (1-\epsilon),$$

where $C_{d,\infty} = \max_{a\in[0,1]} \frac{h_b(a)}{ad+1}$ is the (noiseless) capacity of the (d,∞)-RLL constraint.

Application: (d, k)-RLL Input-Constrained BEC(ϵ)

Application: (d,k)-RLL Input-Constrained BEC(ϵ)

Theorem

The capacity of the (d,k)-RLL input-constrained $(k<\infty)$ binary erasure channel with erasure probability ϵ satisfies

$$C_{d,k}^{BEC(\epsilon)} \geq C_{d,k}\cdot(1-\epsilon),$$

where $C_{d,k} = \max_{a_d,\ldots,a_{k-1}} \frac{\sum_{i=d}^{k-1} h_b(a_i)\prod_{j=d}^{i-1}(1-a_j)}{d+1+\sum_{i=d}^{k-1}\prod_{j=d}^{i}(1-a_j)}$ *is the (noiseless) capacity of the (d,k)-RLL constraint.*

Plots: $(1, \infty)$-RLL Input-Constrained BEC(ϵ)

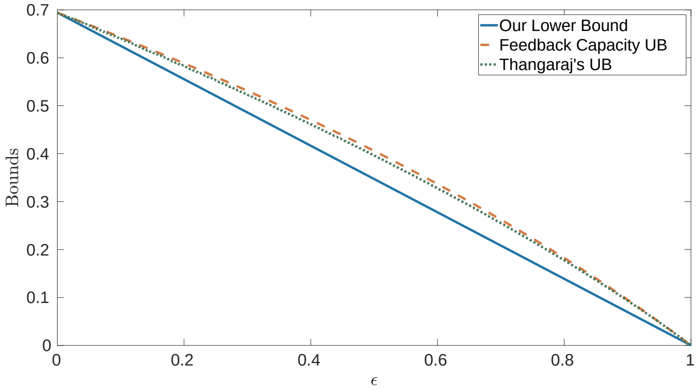


Figure: Comparison of our lower bound with the feedback capacity and dual-capacity upper bounds.

Application: (d, k)-RLL Input-Constrained BSC(p)

Application: (d, k)-RLL Input-Constrained BSC(p)

Theorem

The capacity of the (d, k)-RLL input-constrained $(k < \infty)$ binary symmetric channel with cross-over probability p satisfies

$$C_{d,k}^{BSC(p)} \geq \max_{a_d,\ldots,a_{k-1}\in[0,1]} \frac{\sum_{i=d}^{k-1} \left(h_b(a_i p + \bar{a}_i \bar{p}) - h_b(p)\right) \prod_{j=d}^{i-1}(1-a_j)}{d+1+\sum_{i=d}^{k-1}\prod_{j=d}^{i}(1-a_j)},$$

where $\bar{\alpha} := 1 - \alpha$.

Plots: $(0, k)$-RLL Input-Constrained BSC(p)

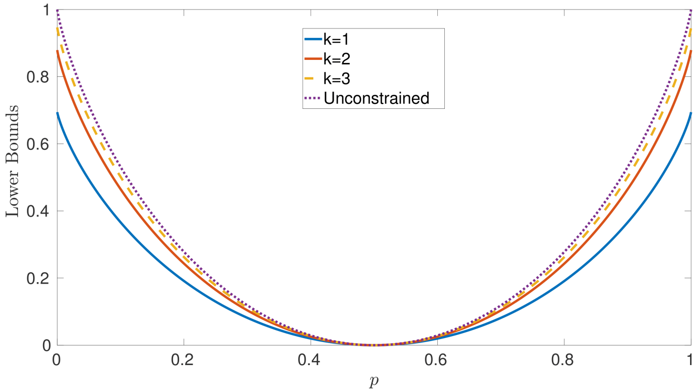


Figure: Our lower bounds for the $(0, 1)$, $(0, 2)$ and $(0, 3)$-RLL input-constrained BSC(p).

Application: (d, ∞)-RLL Input-Constrained BSC(p)

Application: (d, ∞)-RLL Input-Constrained BSC(p)

Theorem

The capacity of the (d, ∞)-RLL input-constrained binary symmetric channel with cross-over probability p obeys

$$C_{d,\infty}^{BSC(p)} \geq \max_{a \in [0,1]} \frac{h_b(ap + \bar{a}\bar{p}) - h_b(p)}{ad + 1},$$

where $\bar{\alpha} := 1 - \alpha$.

Plots: (d, ∞)-RLL Input-Constrained BSC(p)

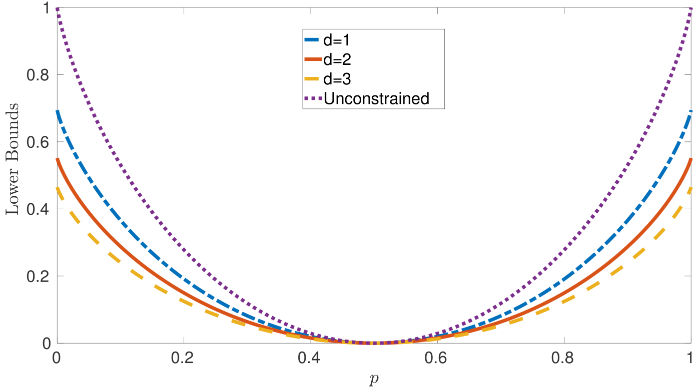


Figure: Our lower bounds for the $(1, \infty)$, $(2, \infty)$, $(3, \infty)$-RLL input-constrained BSC(p).

Plots: $(1, \infty)$-RLL Input-Constrained BSC(p)

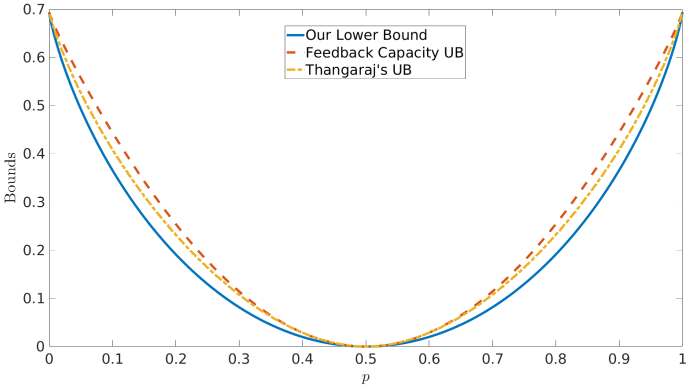


Figure: Comparison of our lower bound with the feedback capacity and dual-capacity upper bounds.

Further Work

Further Work

- Deriving efficient coding schemes that meet the lower bounds.

Further Work

- Deriving efficient coding schemes that meet the lower bounds.
- Improving our lower bound by estimating delayed forward directed information at the input distribution that achieves our lower bound.

Further Work

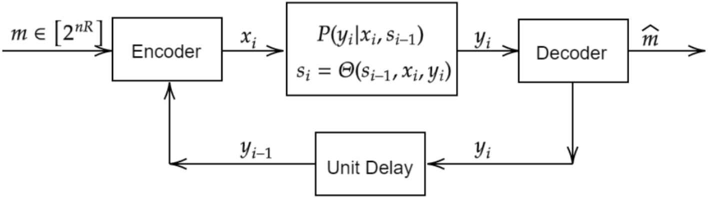
- Deriving efficient coding schemes that meet the lower bounds.
- Improving our lower bound by estimating delayed forward directed information at the input distribution that achieves our lower bound.
- Extending our lower bound to 2D input-constrained DMCs.

FSCs With Feedback

The Setup

The Setup

- Unifilar FSC with feedback:



$$P(y^n, s^n|x^n, s_0) = \prod_{i=1}^{n} P(y_i|x_i, s_{i-1})\mathbb{1}\{s_i = \Theta(s_{i-1}, x_i, y_i)\},$$

where the inputs $x_i \sim P(x_i|x^{i-1}, y^{i-1}, s_0)$.

NAND Flash Memory Channel with ICI

NAND Flash Memory Channel with ICI

- In the 1D ICI model,
  high voltages in neighbour cells $\stackrel{\text{w.p. } \epsilon}{\rightarrow}$ high voltage in victim cell

NAND Flash Memory Channel with ICI

- In the 1D ICI model, high voltages in neighbour cells $\stackrel{\text{w.p. } \epsilon}{\rightarrow}$ high voltage in victim cell
- Channel model ([Li, Han, Siegel, '19]):

$$P(Y^n = y^n \mid X_{-1}^n = x_{-1}^n) = \prod_{i=1}^{n} P(Y_i = y_i \mid X_{i-2}^i = x_{i-2}^i), \text{ where}$$

$$P(Y_i = 1 \mid X_{i-2}^i = x_{i-2}^i) = \begin{cases} 1, & \text{if } x_{i-1} = 1, \\ 0, & \text{if } x_{i-1} = 0, \text{ and} \\ & \quad x_{i-2}^i \neq (1,0,1), \\ \epsilon, & \text{if } x_{i-2}^i = (1,0,1). \end{cases}$$

NAND Flash Memory Channel with ICI

- In the 1D ICI model, high voltages in neighbour cells $\stackrel{\text{w.p. } \epsilon}{\rightarrow}$ high voltage in victim cell
- Channel model ([Li, Han, Siegel, '19]):

$$P(Y^n = y^n \mid X_{-1}^n = x_{-1}^n) = \prod_{i=1}^{n} P(Y_i = y_i \mid X_{i-2}^i = x_{i-2}^i), \text{ where}$$

$$P(Y_i = 1 \mid X_{i-2}^i = x_{i-2}^i) = \begin{cases} 1, & \text{if } x_{i-1} = 1, \\ 0, & \text{if } x_{i-1} = 0, \text{ and} \\ & \quad x_{i-2}^i \neq (1,0,1), \\ \epsilon, & \text{if } x_{i-2}^i = (1,0,1). \end{cases}$$

- The flash memory channel is connected and unifilar.

Feedback Capacity Expression

Feedback Capacity Expression

Theorem (Permuter, Cuff, Van Roy, Weissman, '08)

The feedback capacity of a connected unifilar FSC when the initial state, s_0, is known at the encoder and decoder can be expressed as:

$$C^{fb} = \sup_{\{P(x_t|s_{t-1},y^{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t \mid Y^{t-1}),$$

where $P(x_t \mid y^{t-1}, x^{t-1}, s^{t-1}) = P(x_t \mid s_{t-1}, y^{t-1})$*, for* $t \geq 1$.

Feedback Capacity Expression

Theorem (Permuter, Cuff, Van Roy, Weissman, '08)

The feedback capacity of a connected unifilar FSC when the initial state, s_0, is known at the encoder and decoder can be expressed as:

$$C^{fb} = \sup_{\{P(x_t|s_{t-1},y^{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t \mid Y^{t-1}),$$

where $P(x_t \mid y^{t-1}, x^{t-1}, s^{t-1}) = P(x_t \mid s_{t-1}, y^{t-1})$, *for* $t \geq 1$.

Was also shown to be amenable to formulation as the optimal average reward of an infinte-horizon average-reward DP problem

DP Formulation

DP Formulation

$$C^{\text{fb}} = \sup_{\{P(x_t|s_{t-1},y^{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t \mid Y^{t-1})$$

DP Formulation

$$C^{\text{fb}} = \sup_{\{P(x_t|s_{t-1},y^{t-1})\}_{t\geq 1}} \liminf_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} I(X_t, S_{t-1}; Y_t \mid Y^{t-1})$$

	DP Notation	Our Instance
State	z_{t-1}	$P(s_{t-1}\|y^{t-1})$
Action	u_t	$P(x_t\|s_{t-1}, y^{t-1})$
Disturbance	w_t	y_t
Reward	$g(z_{t-1}, u_t)$	$I(X_t, S_{t-1}; Y_t\|y^{t-1})$

Numerical Evaluation

Numerical Evaluation

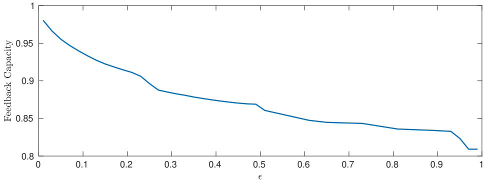


Figure: Feedback capacity as a function of ϵ, obtained by numerical evaluation of the DP problem.

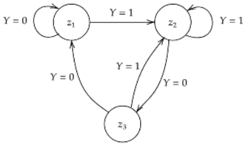


Figure: Graph of DP states under optimal policy.

Numerical evaluations are close to best known upper bounds from [Sabag et al., '16]

Deterministic Flash Memory Channel ($\epsilon = 1$)

Deterministic Flash Memory Channel ($\epsilon = 1$)

- Easy observation: $C^{\text{fb}}(1) = C(1)$, since the encoder knows y_t as a function of x_{t-2}^{t}, as soon as x_t is determined.

Deterministic Flash Memory Channel ($\epsilon = 1$)

- Easy observation: $C^{\text{fb}}(1) = C(1)$, since the encoder knows y_t as a function of x_{t-2}^t, as soon as x_t is determined.
- From numerical evaluations of single-letter upper bound from [Sabag et al., '16], $C^{\text{fb}}(1) \lessapprox 0.8114$.

Deterministic Flash Memory Channel ($\epsilon = 1$)

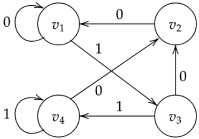
- Easy observation: $C^{\text{fb}}(1) = C(1)$, since the encoder knows y_t as a function of x_{t-2}^t, as soon as x_t is determined.
- From numerical evaluations of single-letter upper bound from [Sabag et al., '16], $C^{\text{fb}}(1) \lessapprox 0.8114$.

Can we design a simple coding scheme that meets this upper bound?

Constrained Coding Scheme

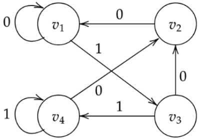
Constrained Coding Scheme

- Consider the constraint, $S_{\text{no-101}}$, that forbids the contiguous 101 pattern, with deterministic presentation (having adjacency matrix G) shown below:



Constrained Coding Scheme

- Consider the constraint, $S_{\text{no-101}}$, that forbids the contiguous 101 pattern, with deterministic presentation (having adjacency matrix G) shown below:

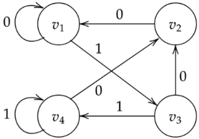


- By standard results, noiseless capacity,

$$\begin{aligned} \text{cap}(S_{\text{no-101}}) &= \log(\lambda_G) \quad [\lambda_G \to \text{largest eigenvalue of } G] \\ &\approx 0.8114. \end{aligned}$$

Constrained Coding Scheme

- Consider the constraint, $S_{\text{no-101}}$, that forbids the contiguous 101 pattern, with deterministic presentation (having adjacency matrix G) shown below:



- By standard results, noiseless capacity,

$$\begin{aligned} \text{cap}(S_{\text{no-101}}) &= \log(\lambda_G) \quad [\lambda_G \rightarrow \text{largest eigenvalue of } G] \\ &\approx 0.8114. \end{aligned}$$

- Hence, $C^{\text{fb}}(1) = C(1) \approx 0.8114$.

Thank You!