Fundamental limits of remote estimation under communication constraints

Aditya Mahajan
McGill University

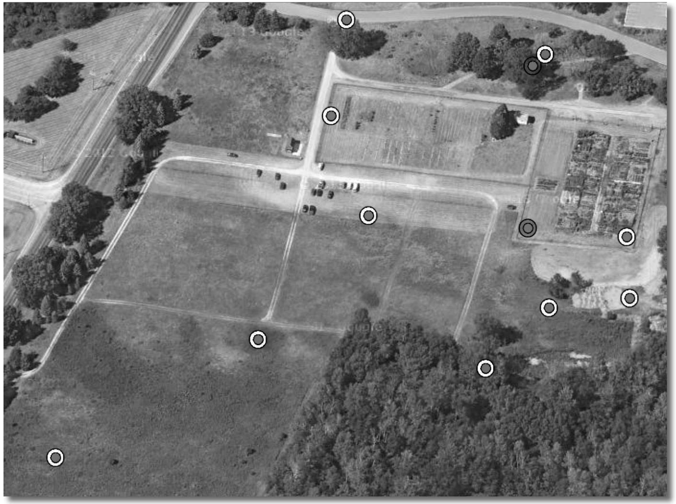
Joint work with Jhelum Chakravorty

Modeling and Optimization in Mobile, Ad-Hoc Wireless Networks (WiOpt)
11 May, 2016

There is a need to revisit estimation theory to take network resources into account.

Many applications require:

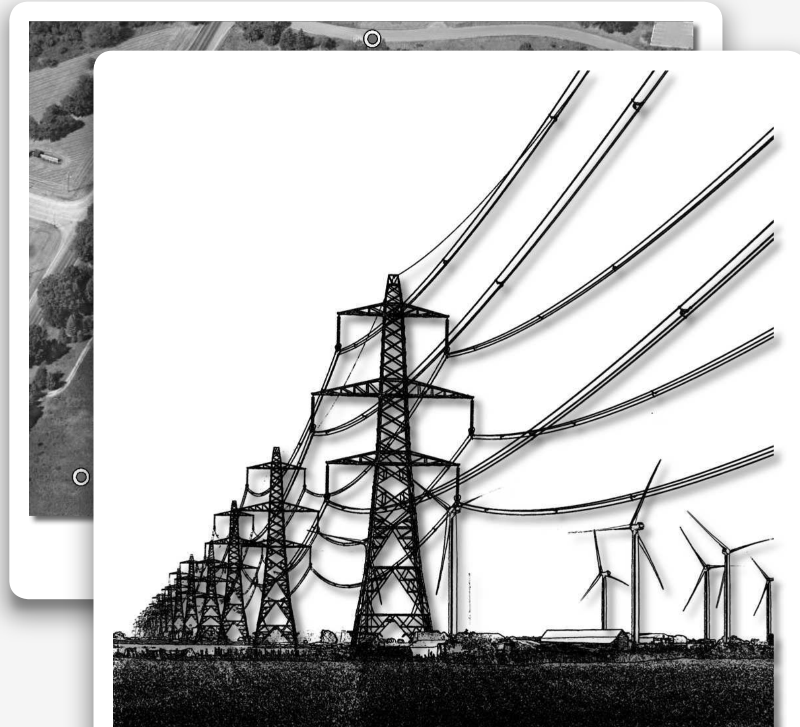
- Sequential transmission of data
- Zero- (or finite-) delay reconstruction



Sensor Networks

Many applications require:

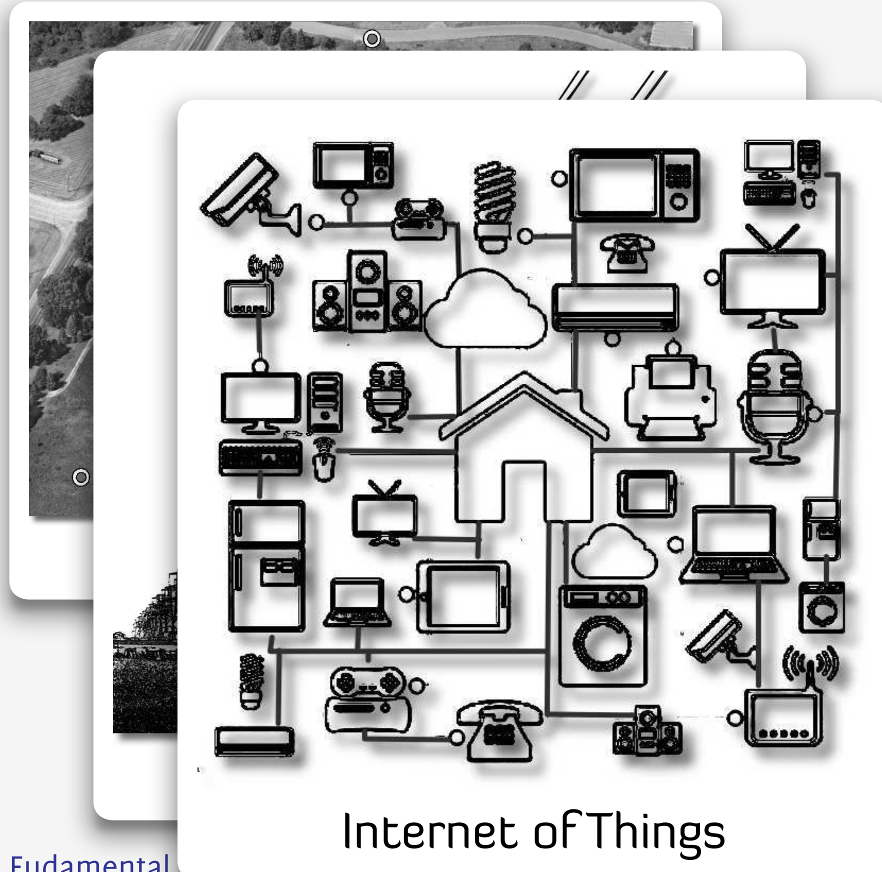
- Sequential transmission of data
- Zero- (or finite-) delay reconstruction



Smart Grids

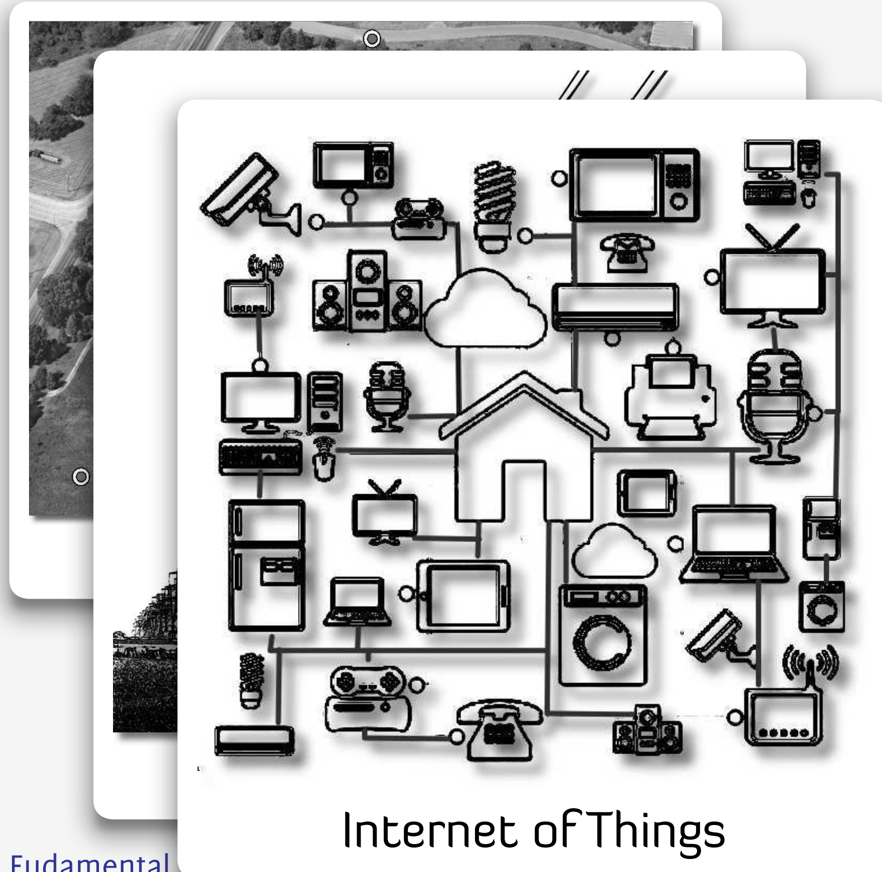
Many applications require:

- Sequential transmission of data
- Zero- (or finite-) delay reconstruction



Many applications require:

- Sequential transmission of data
- Zero- (or finite-) delay reconstruction



Many applications require:

- Sequential transmission of data
- Zero- (or finite-) delay reconstruction

Salient features

- Sensing is cheap
- Transmission is expensive
- Size of data-packet is not critical

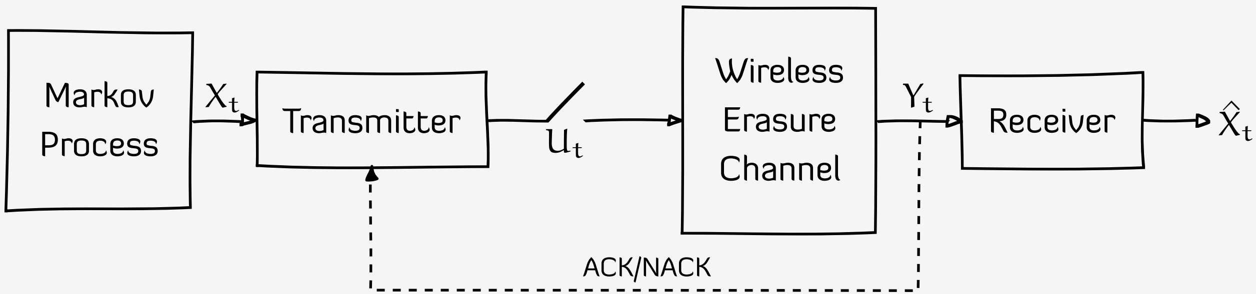
Many applications require:

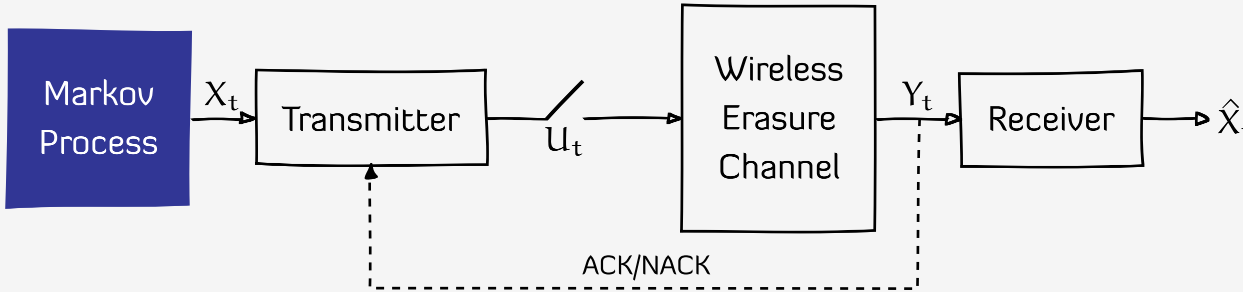
- Sequential transmission of data
- Zero- (or finite-) delay reconstruction

Salient features

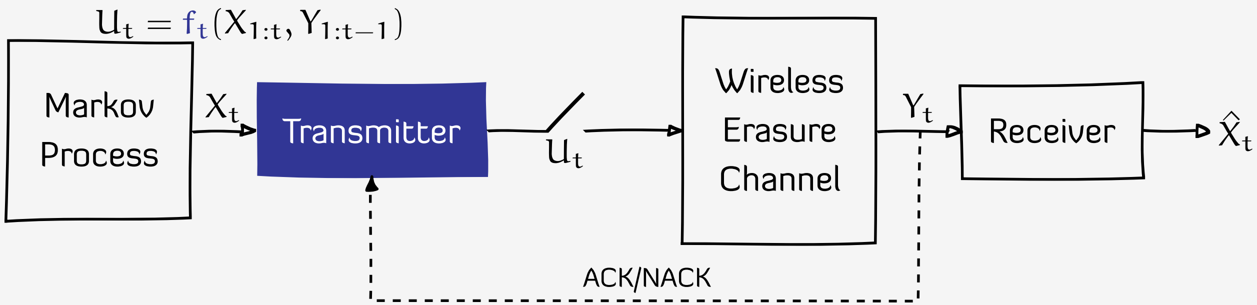
- Sensing is cheap
- Transmission is expensive
- Size of data-packet is not critical

Analyze a stylized model and evaluate fundamental trade-offs


Markov
Process
X_t
Transmitter
U_t
Wireless
Erasure
Channel
Y_t
Receiver
$\hat{X}_t$
ACK/NACK



- First order time-homogeneous Markov process
- The transmitter decides whether or not to transmit the current state
- The transmitted symbol is sent over an erasure channel (with acknowledgments)
- The receiver generates an estimate based on received symbol

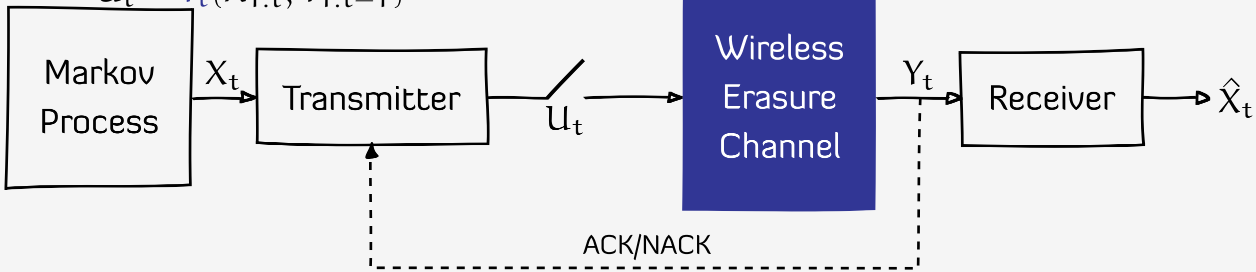


- First order time-homogeneous Markov process
- The transmitter decides whether or not to transmit the current state
- The transmitted symbol is sent over an erasure channel (with acknowledgments)
- The receiver generates an estimate based on received symbol

$$Y_t = \begin{cases} X_t, & \text{if } U_t H_t = 1 \\ \mathfrak{E}, & \text{if } U_t H_t = 0 \end{cases}$$

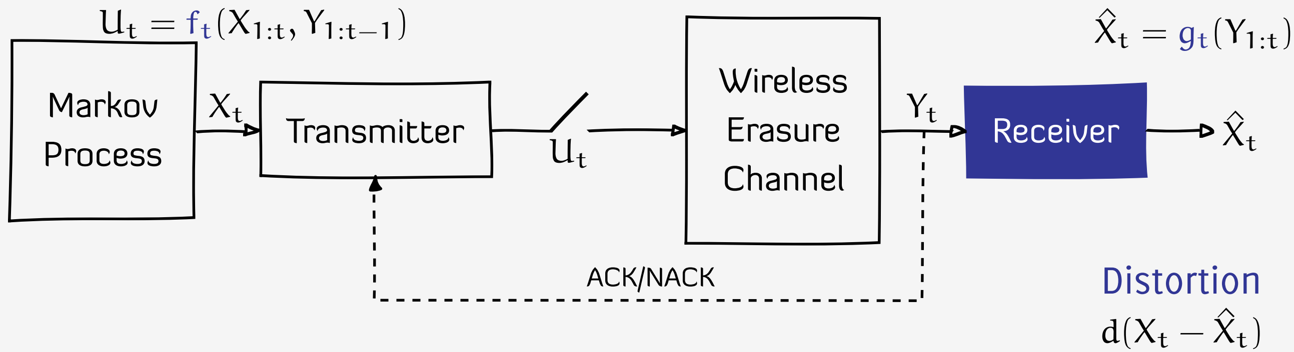
$$\mathbb{P}(H_t = 0) = \varepsilon$$

$$U_t = f_t(X_{1:t}, Y_{1:t-1})$$



- First order time-homogeneous Markov process
- The transmitter decides whether or not to transmit the current state
- The transmitted symbol is sent over an erasure channel (with acknowledgments)
- The receiver generates an estimate based on received symbol

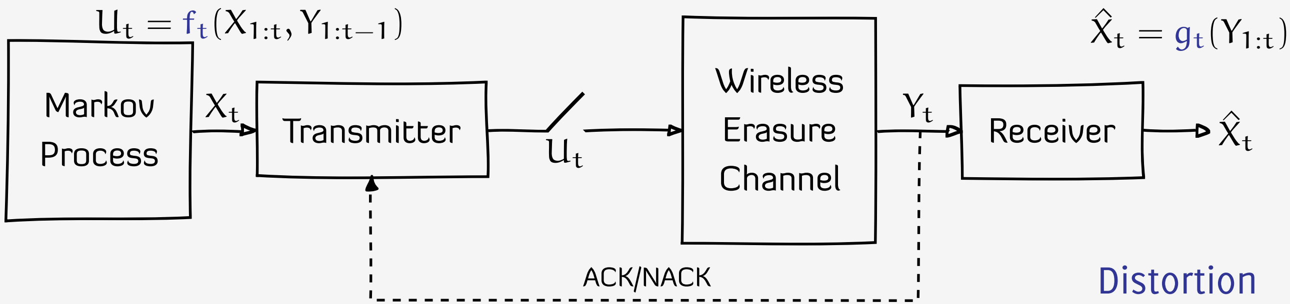
$$\mathbb{P}(H_t = 0) = \varepsilon \qquad Y_t = \begin{cases} X_t, & \text{if } U_t H_t = 1 \\ \mathfrak{E}, & \text{if } U_t H_t = 0 \end{cases}$$



- First order time-homogeneous Markov process
- The transmitter decides whether or not to transmit the current state
- The transmitted symbol is sent over an erasure channel (with acknowledgments)
- The receiver generates an estimate based on received symbol

$$Y_t = \begin{cases} X_t, & \text{if } U_t H_t = 1 \\ \mathfrak{E}, & \text{if } U_t H_t = 0 \end{cases}$$

$$\mathbb{P}(H_t = 0) = \varepsilon$$



Distortion

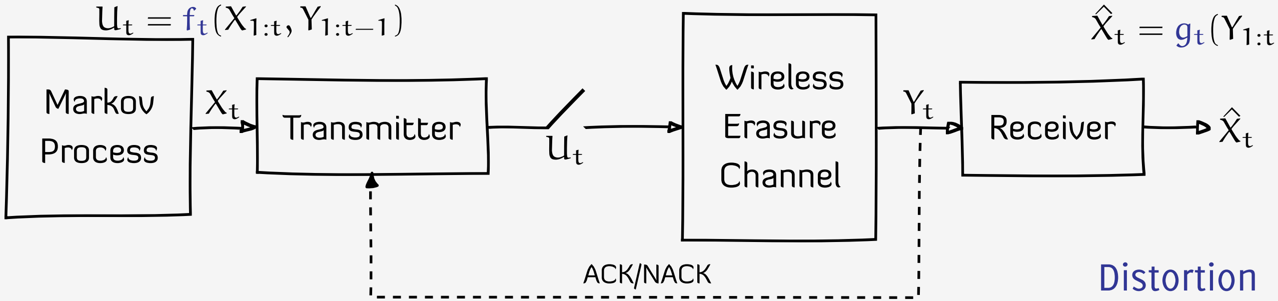
$d(X_t - \hat{X}_t)$

Communication Strategies

- Transmission strategy $f = \{f_t\}_{t=0}^{\infty}$.
- Estimation strategy $g = \{g_t\}_{t=0}^{\infty}$.

$$Y_t = \begin{cases} X_t, & \text{if } U_t H_t = 1 \\ \mathfrak{E}, & \text{if } U_t H_t = 0 \end{cases}$$

$$\mathbb{P}(H_t = 0) = \varepsilon$$



Distortion

$d(X_t - \hat{X}_t)$

1. **Discounted setup, $\beta \in (0, 1)$**

$$D_\beta(f, g) = (1-\beta)\, \mathbb{E}_0^{(f,g)} \left[\sum_{t=0}^{\infty} \beta^t d(X_t - \hat{X}_t) \right]; \qquad N_\beta(f, g) = (1-\beta)\, \mathbb{E}_0^{(f,g)} \left[\sum_{t=0}^{\infty} \beta^t U_t \right]$$

2. **Average cost setup, $\beta = 1$**

$$D_1(f, g) = \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}_0^{(f,g)} \left[\sum_{t=0}^{T-1} d(X_t - \hat{X}_t) \right]; \qquad N_1(f, g) = \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}_0^{(f,g)} \left[\sum_{t=0}^{T-1} U_t \right]$$

Optimization problems

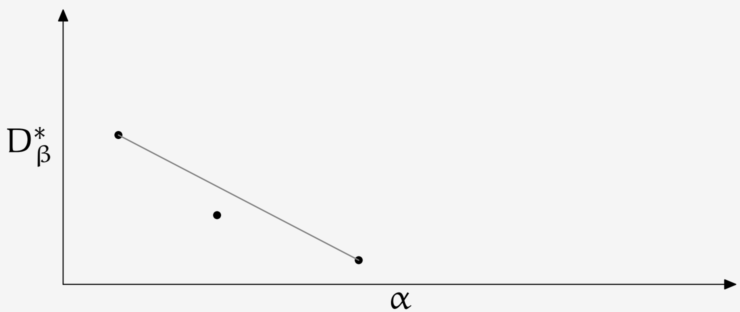
Constrained communication

$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \right\}$$

Optimization problems

Constrained communication

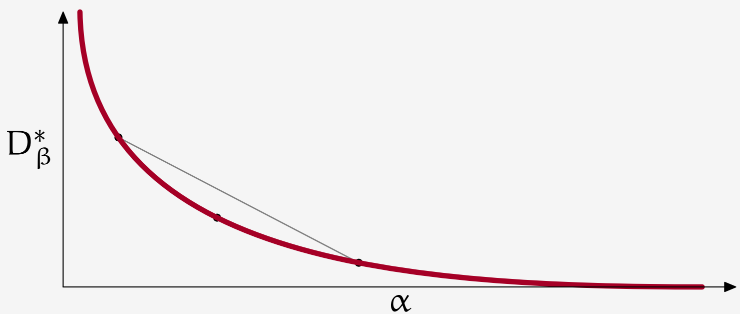
$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \}$$



Optimization problems

Constrained communication

$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \}$$



D^*_β is cts, dec, and convex

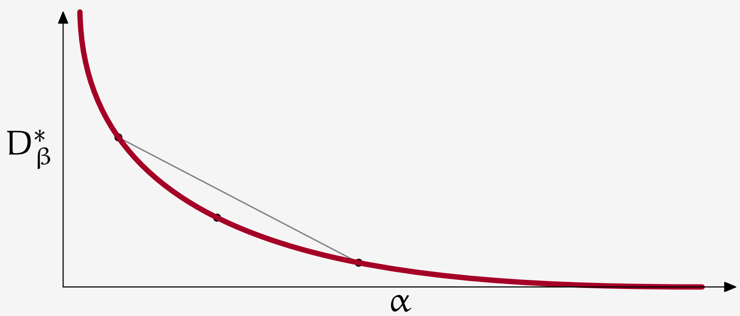
Optimization problems

Constrained communication

$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \right\}$$

Costly communication (Lagrange relaxation)

$$\text{For } \lambda \in \mathbb{R}_{>0}, \quad C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) + \lambda N_\beta(f,g) \right\}$$



D^*_β is cts, dec, and convex

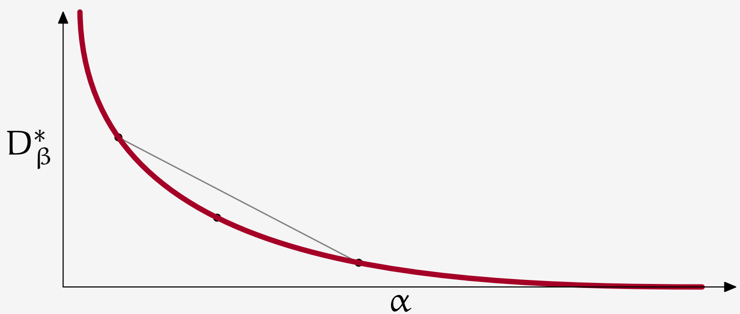
Optimization problems

Constrained communication

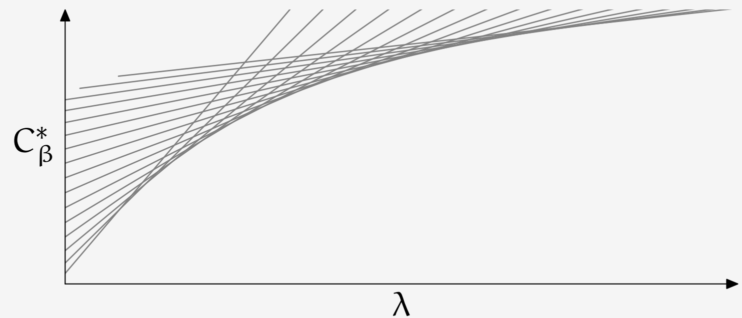
$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \right\}$$

Costly communication (Lagrange relaxation)

$$\text{For } \lambda \in \mathbb{R}_{>0}, \quad C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) + \lambda N_\beta(f,g) \right\}$$



D^*_β is cts, dec, and convex



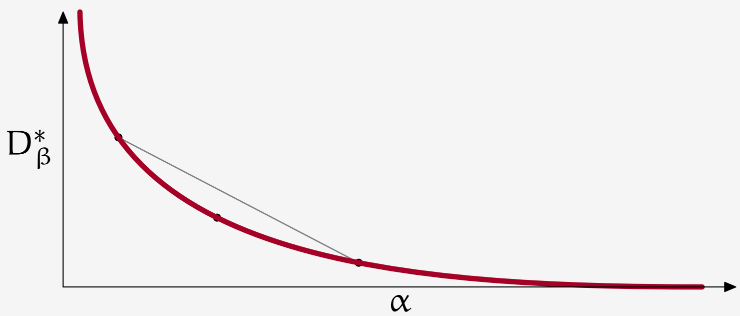
Optimization problems

Constrained communication

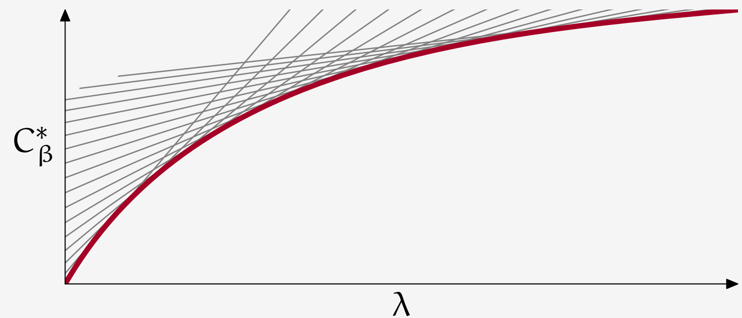
$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \right\}$$

Costly communication (Lagrange relaxation)

$$\text{For } \lambda \in \mathbb{R}_{>0}, \quad C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) + \lambda N_\beta(f,g) \right\}$$



D^*_β is cts, dec, and convex



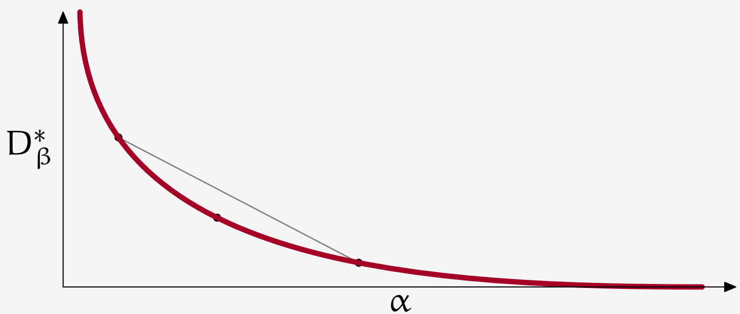
C^*_β is cts, inc, and concave

Optimization problems

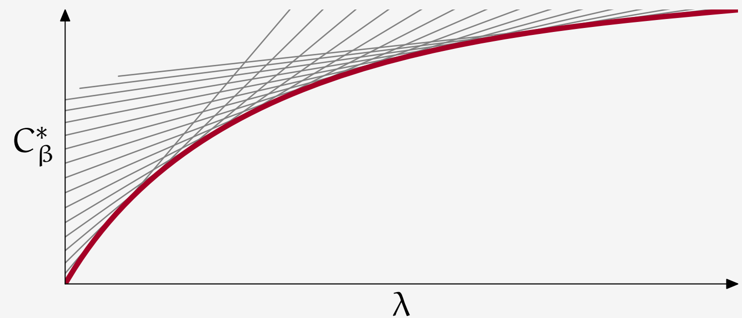
Constrained communication

Costly

Our result: Provide computable expressions for these trade-offs and identify optimal strategies that achieve them.



D^*_β is cts, dec, and convex



C^*_β is cts, inc, and concave

Comparison to Information Theory

- Costly communication is analogous to communication under power constraint.
- Constrained communication is analogous to distortion-rate function.

  So, we call it distortion-transmission function.

- Due to zero-delay reconstruction, information theoretic approaches do not apply.

Comparison to Information Theory

- Costly communication is analogous to communication under power constraint.
- Constrained communication is analogous to distortion-rate function.
  So, we call it distortion-transmission function.
- Due to zero-delay reconstruction, information theoretic approaches do not apply.

Previous work on remote-state estimation

- [Marshak 1954] Static (one-shot) problem with arbitrary source distribution
- [Kushner 1964] Off-line choice of measurement times
- [Åstrom Bernhardsson 2002] Lebesque sampling (or event-based sampling)

Comparison to Information Theory

- Costly communication is analogous to communication under power constraint.
- Constrained communication is analogous to distortion-rate function.
  So, we call it distortion-transmission function.
- Due to zero-delay reconstruction, information theoretic approaches do not apply.

Previous work on remote-state estimation

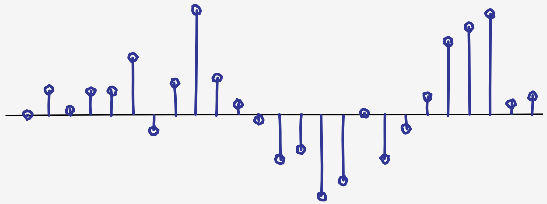
- [Marshak 1954] Static (one-shot) problem with arbitrary source distribution
- [Kushner 1964] Off-line choice of measurement times
- [Åstrom Bernhardsson 2002] Lebesque sampling (or event-based sampling)

Other related work

- Event-based estimation . . .
- Censoring censors . . .
- Sensor sleep scheduling . . .
- Age of Information . . .

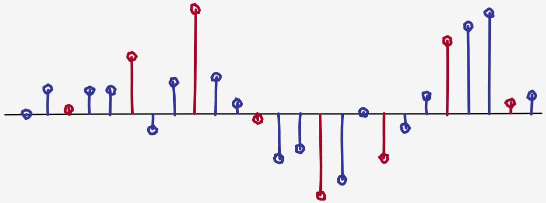
An illustrative example

$X_{t+1} = X_t + W_t$, $W_t \sim \mathcal{N}(0, 1)$. Perfect channel



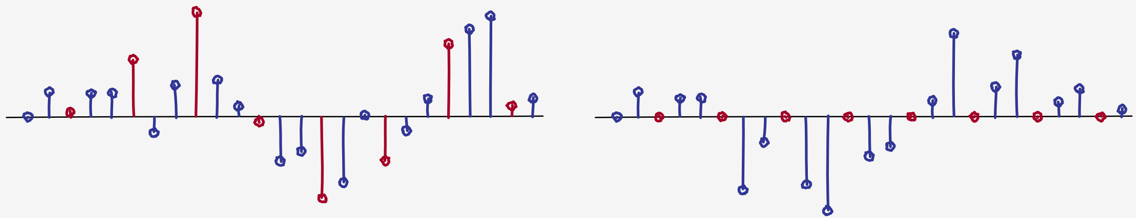
$X_{t+1} = X_t + W_t$, $W_t \sim \mathcal{N}(0, 1)$. Perfect channel

Periodic Transmission Strategy



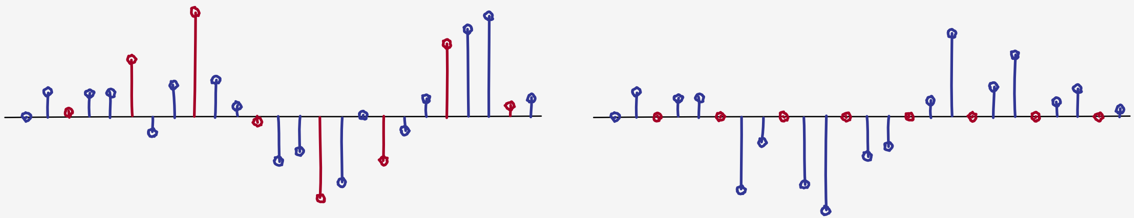
$X_{t+1} = X_t + W_t$, $W_t \sim \mathcal{N}(0,1)$. Perfect channel

Periodic Transmission Strategy



$X_{t+1} = X_t + W_t$, $W_t \sim \mathcal{N}(0, 1)$. Perfect channel

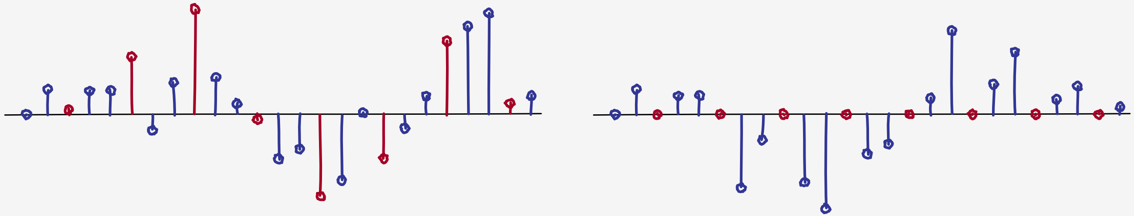
Periodic Transmission Strategy



$D = 0.67$
$N \approx 1/3$

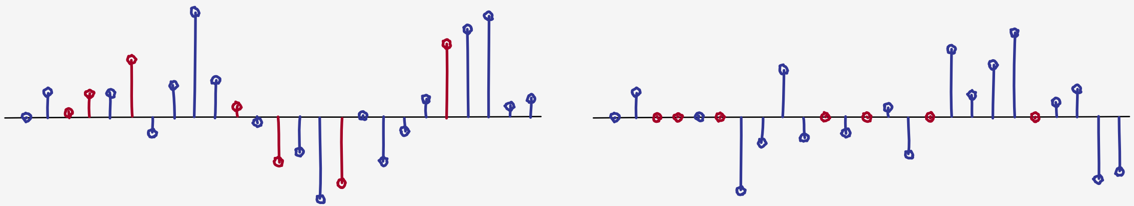
$X_{t+1} = X_t + W_t$, $W_t \sim \mathcal{N}(0,1)$. Perfect channel

Periodic Transmission Strategy



$D = 0.67$
$N \approx 1/3$

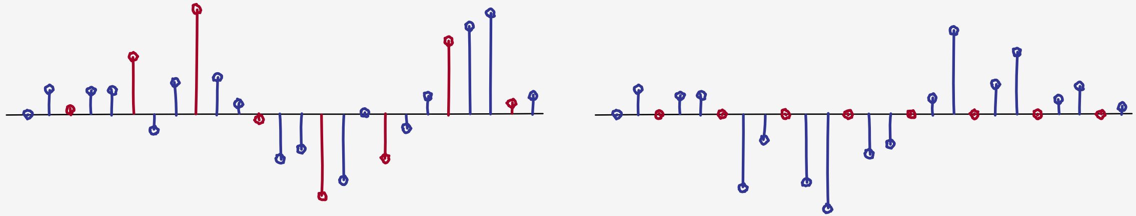
Randomized Transmission Strategy



$D = 2.00$
$N \approx 1/3$

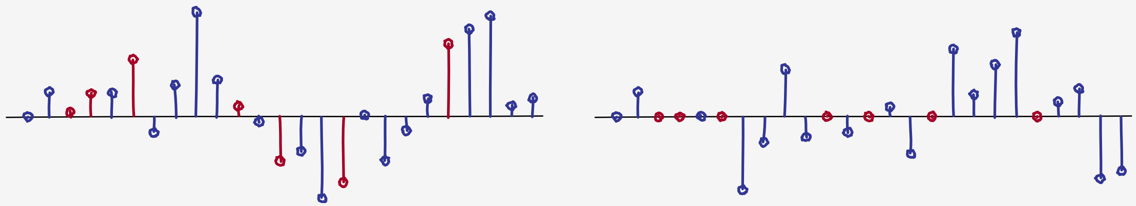
$X_{t+1} = X_t + W_t$, $W_t \sim \mathcal{N}(0,1)$. Perfect channel

Periodic Transmission Strategy



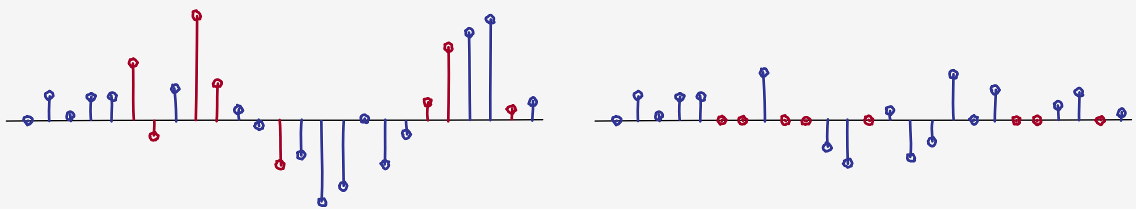
$D = 0.67$
$N \approx 1/3$

Randomized Transmission Strategy



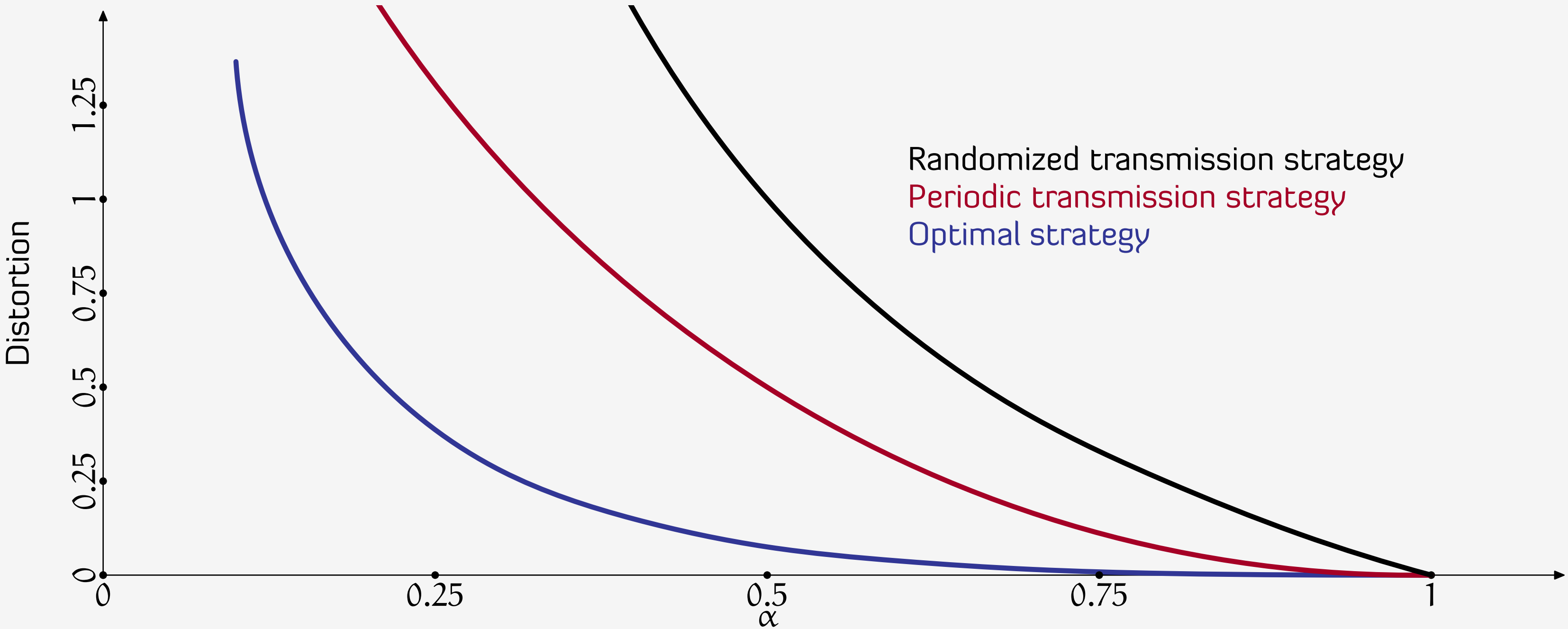
$D = 2.00$
$N \approx 1/3$

Optimal Transmission Strategy



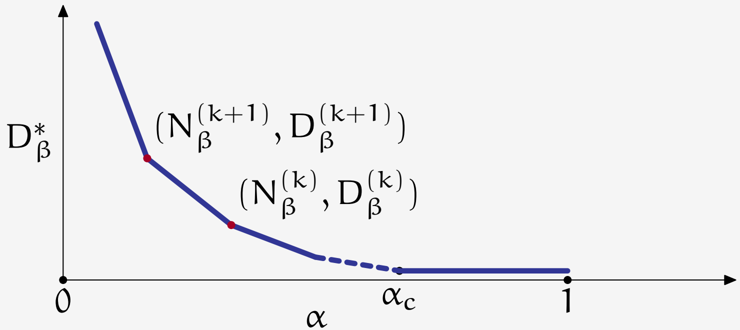
$D = 0.24$
$N \approx 1/3$

Distortion-transmission trade-off: Perfect channel



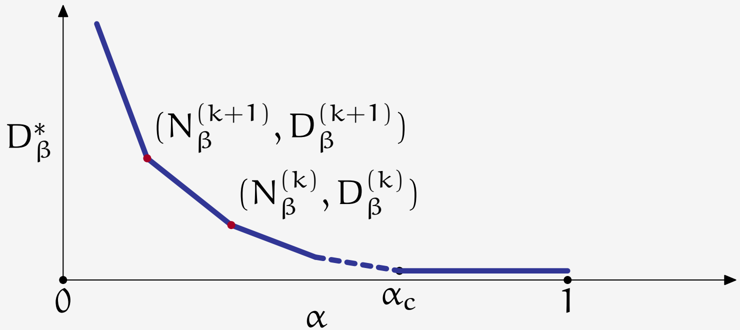
Main results

Distortion transmission function for discrete auto-regressive sources



D^*_β is PWL, dec, and convex

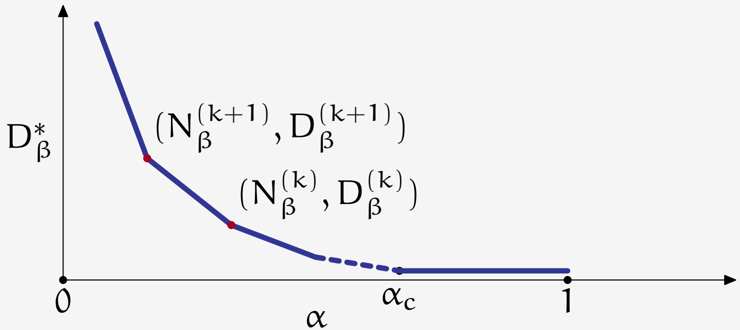
Distortion transmission function for discrete auto-regressive sources



D^*_β is PWL, dec, and convex

How to compute $D^*_\beta(\alpha)$

Distortion transmission function for discrete auto-regressive sources



D^*_β is PWL, dec, and convex

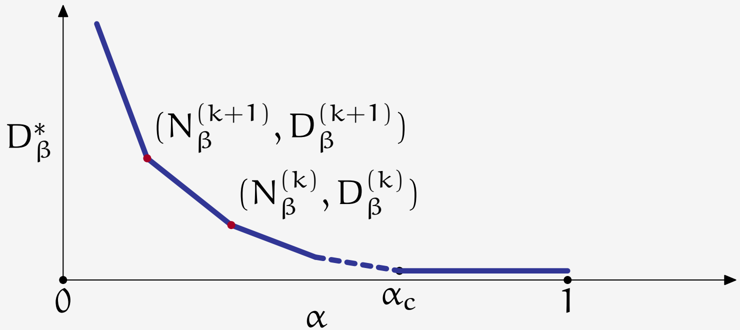
How to compute $D^*_\beta(\alpha)$

- Compute $L^{(k)}_\beta = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)} \odot d.$

$$M^{(k)}_\beta = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)}.$$

- Then $D^{(k)}_\beta = \dfrac{L^{(k)}(0)}{M^{(k)}_\beta(0)}$ and $N^{(k)}_\beta = \dfrac{1}{M^{(k)}_\beta(0)} - (1-\beta)$

Distortion transmission function for discrete auto-regressive sources



D^*_β is PWL, dec, and convex

Optimal transmission strategy

- Find k^* such that $\alpha \in (N^{(k^*+1)}_\beta, N^{(k^*)}_\beta]$.
- Compute θ^* such that
  $$\theta^* N^{(k)}_\beta + (1-\theta^*) N^{(k+1)}_\beta = \alpha$$
- If $|X_t - a\hat{X}_{t-1}| > k^*(\alpha)$, transmit.
- If $|X_t - a\hat{X}_{t-1}| = k^*(\alpha)$, transmit w.p. θ^*.
- Else, do not transmit.

Optimal estimation strategy

$$\hat{X}_t = \begin{cases} Y_t, & \text{if } Y_t \neq \mathfrak{E} \\ a\hat{X}_{t-1}, & \text{if } Y_t = \mathfrak{E} \end{cases}$$

How to compute $D^*_\beta(\alpha)$

- Compute $L^{(k)}_\beta = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)} \odot d$.

  $$M^{(k)}_\beta = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)}.$$

- Then $D^{(k)}_\beta = \dfrac{L^{(k)}(0)}{M^{(k)}_\beta(0)}$ and $N^{(k)}_\beta = \dfrac{1}{M^{(k)}_\beta(0)} - (1-\beta)$

Identify strategies that achieve the optimal trade-off

Simple and intuitive threshold based strategies are optimal.

Provide computable expressions for distortion-transmission function

Based on simple matrix calculations for discrete Markov processes

Based on solving Fredholm integral equations for continuous processes

Identify strategies that achieve the optimal trade-off

Simple and intuitive threshold based strategies are optimal.

Provide computable expressions for distortion-transmission function

Based on simple matrix calculations for discrete Markov processes

Based on solving Fredholm integral equations for continuous processes

Beautiful example of stochastics and optimization

Decentralized stochastic control (or team theory) and POMDPs

Stochastic orders and majorization

Markov chain analysis, stopping times, and renewal theory

Constrained MDPs and Lagrangian relaxations

What’s the conceptual difficulty?

Static (one-shot) problem

Static (one-shot) problem

$\mathcal{S} \subset \mathcal{X}$ is the silence set

Static (one-shot) problem

$\mathcal{S} \subset \mathcal{X}$ is the silence set

$\hat{x}$ is the estimate when no packet is received

Static (one-shot) problem

$\mathcal{S} \subset \mathcal{X}$ is the silence set

$\hat{x}$ is the estimate when no packet is received

Cost when $x \in \mathcal{S}$

$d(x - \hat{x})$

Static (one-shot) problem

$\mathcal{S} \subset \mathcal{X}$ is the silence set

$\hat{x}$ is the estimate when no packet is received

Cost when $x \in \mathcal{S}$

$d(x - \hat{x})$

Cost when $x \notin \mathcal{S}$

$\lambda + \varepsilon d(x - \hat{x})$

Static (one-shot) problem

$\mathcal{S} \subset \mathcal{X}$ is the silence set

$\hat{x}$ is the estimate when no packet is received

Cost when $x \in \mathcal{S}$

$d(x - \hat{x})$

Cost when $x \notin \mathcal{S}$

$\lambda + \varepsilon d(x - \hat{x})$

Total expected cost

$$c(\hat{x}, \mathcal{S}) \coloneqq \lambda \, \mathbb{P}(X \notin \mathcal{S}) + \varepsilon \sum_{x \notin \mathcal{S}} \mathbb{P}(X = x) d(x - \hat{x}) + \sum_{x \in \mathcal{S}} \mathbb{P}(X = x) d(x - \hat{x})$$

Static (one-shot) problem

$\mathcal{S} \subset \mathcal{X}$ is the silence set

$\hat{x}$ is the estimate when no packet is received

Cost when $x \in \mathcal{S}$

$d(x - \hat{x})$

Cost when $x \notin \mathcal{S}$

$\lambda + \varepsilon d(x - \hat{x})$

Total expected cost

$$c(\hat{x}, \mathcal{S}) \coloneqq \lambda\, \mathbb{P}(X \notin \mathcal{S}) + \varepsilon \sum_{x \notin \mathcal{S}} \mathbb{P}(X = x) d(x - \hat{x}) + \sum_{x \in \mathcal{S}} \mathbb{P}(X = x) d(x - \hat{x})$$

> Choose $(\hat{x}, \mathcal{S})$ to minimize $c(\hat{x}, \mathcal{S})$.
>
> Set-valued (or combinatorial) optimization problem.

Dynamic problem

$\mathcal{S}_1^1 \subset \mathcal{X}$ is the silence set

$\hat{x}_1$ is the estimate when no packet is received

Dynamic problem

$\mathcal{S}_1^1 \subset \mathcal{X}$ is the silence set

$\hat{x}_1$ is the estimate when no packet is received

If a packet is received $(U_t H_t = 1)$

$\mathcal{S}_2^1(x_1) \subset \mathcal{X}$ is the silence set

$\hat{x}_2^1$ is the estimate when no packet is received

Dynamic problem

$\mathcal{S}_1^1 \subset \mathcal{X}$ is the silence set

$\hat{x}_1$ is the estimate when no packet is received

If a packet is received $(U_t H_t = 1)$

$\mathcal{S}_2^1(x_1) \subset \mathcal{X}$ is the silence set

$\hat{x}_2^1$ is the estimate when no packet is received

If a packet is not received $(U_t H_t = 0)$

$\mathcal{S}_2^0(\mathcal{S}_1^1) \subset \mathcal{X}$ is the silence set

$\hat{x}_2^0(\mathcal{S}_1^1)$ is the estimate when no packet is recei

Dynamic problem

$\mathcal{S}_1^1 \subset \mathcal{X}$ is the silence set

$\hat{x}_1$ is the estimate when no packet is received

If a packet is received $(U_t H_t = 1)$

$\mathcal{S}_2^1(x_1) \subset \mathcal{X}$ is the silence set

$\hat{x}_2^1$ is the estimate when no packet is received

If a packet is not received $(U_t H_t = 0)$

$\mathcal{S}_2^0(\mathcal{S}_1^1) \subset \mathcal{X}$ is the silence set

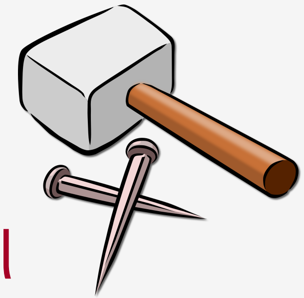
$\hat{x}_2^0(\mathcal{S}_1^1)$ is the estimate when no packet is recei

Sequential optimization problem where the optimization problem at each step is a set-valued optimization problem that depends on a history of previously chosen sets!.

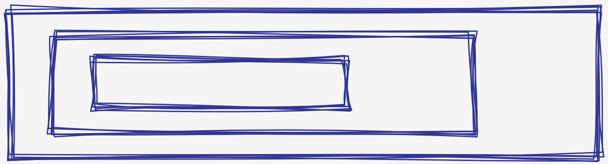
Exhaustive search complexity: $\left(|\mathcal{X}| 2^{|\mathcal{X}|}\right)^{(2^{|\mathcal{X}|})^T}$

So how do we start?

Decentralized stochastic control

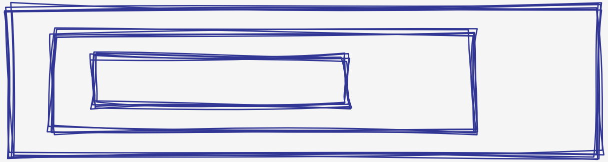


Dealing with non-classical information structure



Classical info. struct.

Dealing with non-classical information structure

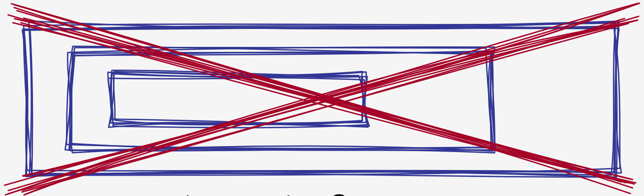


Classical info. struct.

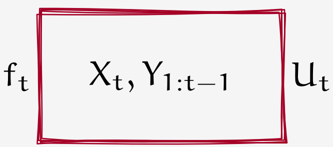
f_t : $X_t, Y_{1:t-1}$ → U_t

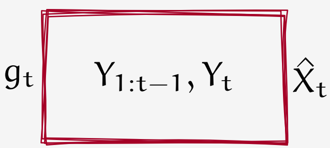
g_t : $Y_{1:t-1}, Y_t$ → $\hat{X}_t$

Dealing with non-classical information structure

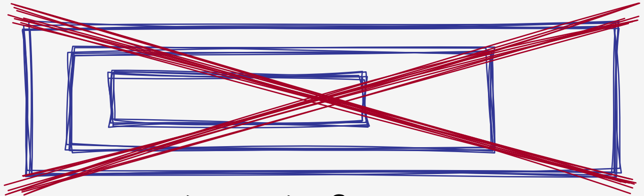


Non-Classical info. struct.





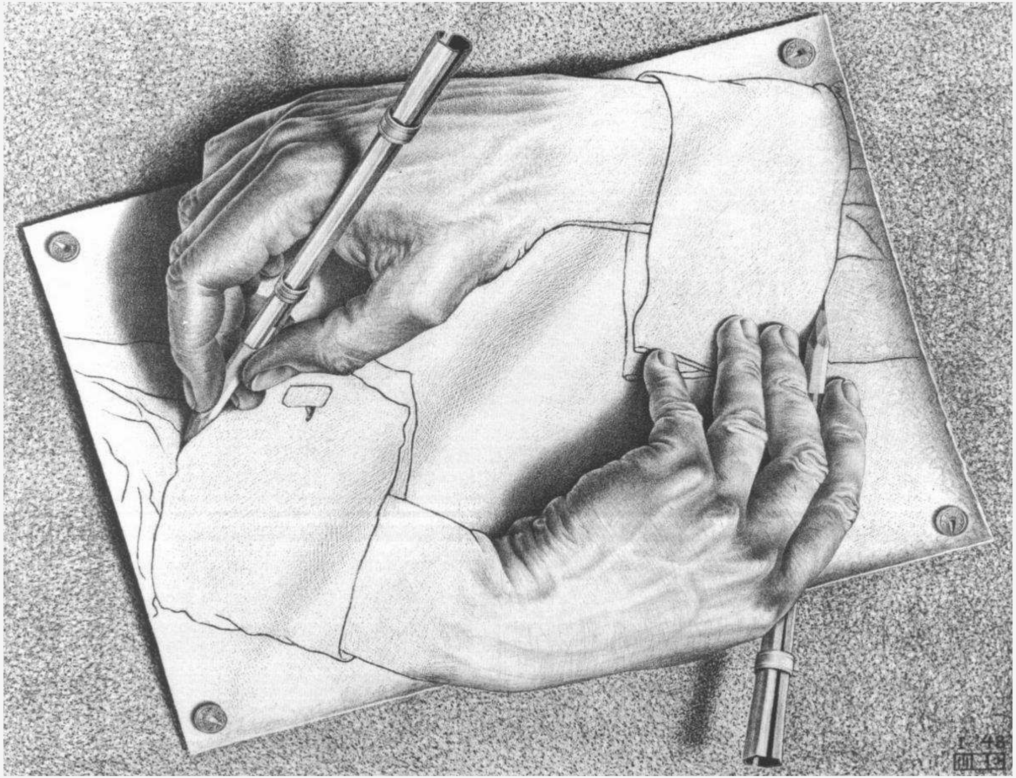
Dealing with non-classical information structure



Non-Classical info. struct.

f_t | $X_t, Y_{1:t-1}$ | U_t

g_t | $Y_{1:t-1}, Y_t$ | $\hat{X}_t$



The common information approach (Nayyar, Mahajan, Teneketzis 2013)

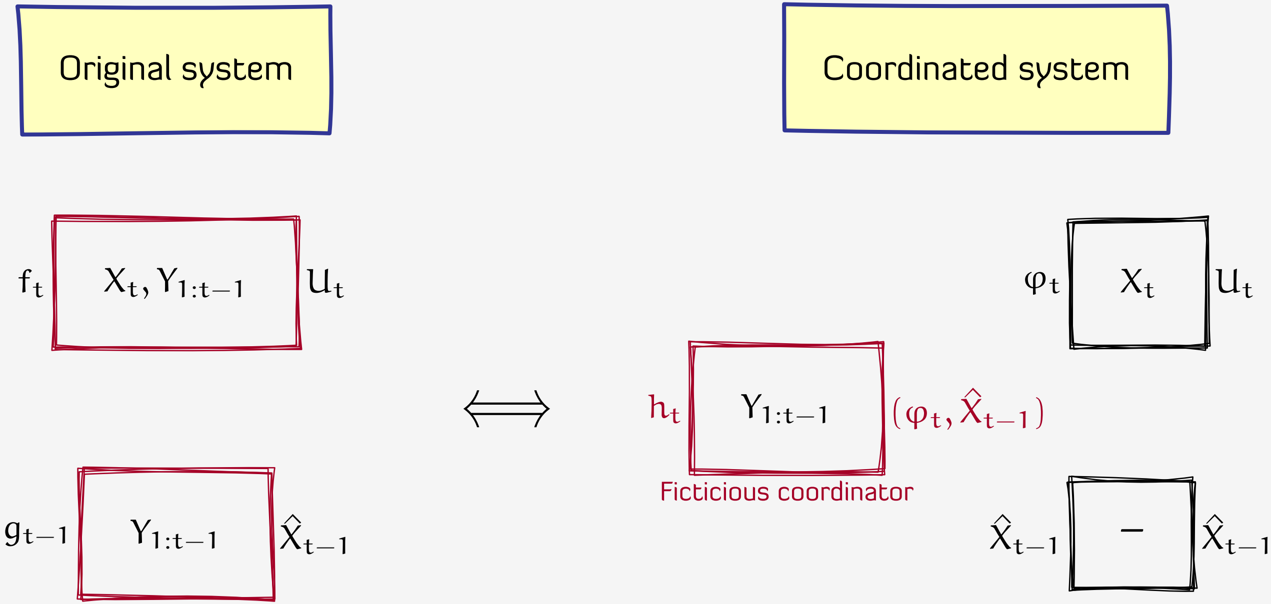
Original system

f_t | $X_t, Y_{1:t-1}$ | U_t

g_{t-1} | $Y_{1:t-1}$ | $\hat{X}_{t-1}$

▷ Nayyar, Mahajan and Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," IEEE TAC 2013.

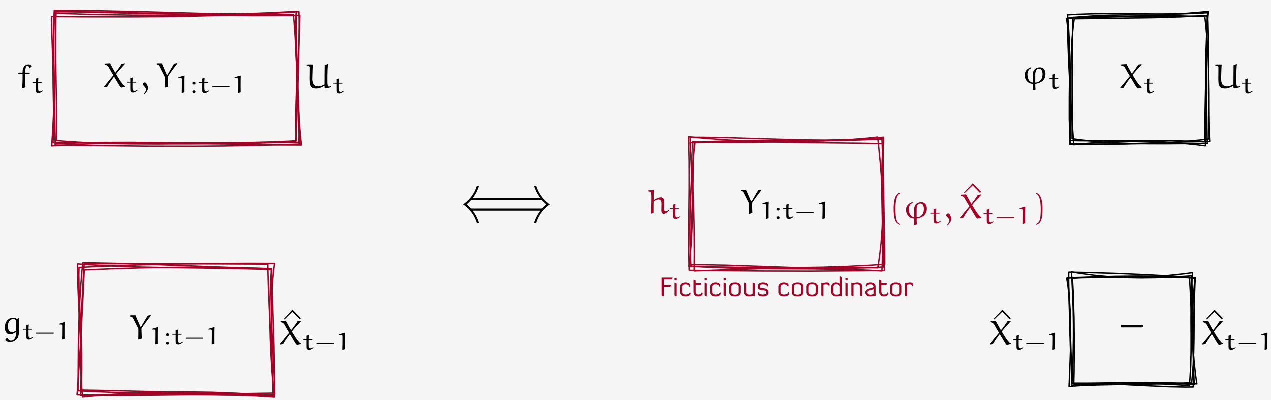
The common information approach (Nayyar, Mahajan, Teneketzis 2013)



▷ Nayyar, Mahajan and Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," IEEE TAC 2013.

The common information approach (Nayyar, Mahajan, Teneketzis 2013)

Original system | **Coordinated system**



- The coordinated system is equivalent to the original system.

$$f_t(x, y_{1:t-1}) = h^1_t(y_{1:t-1})(x).$$

- The coordinated system is centralized. Belief state $\mathbb{P}(X_t \mid Y_{1:t-1})$.

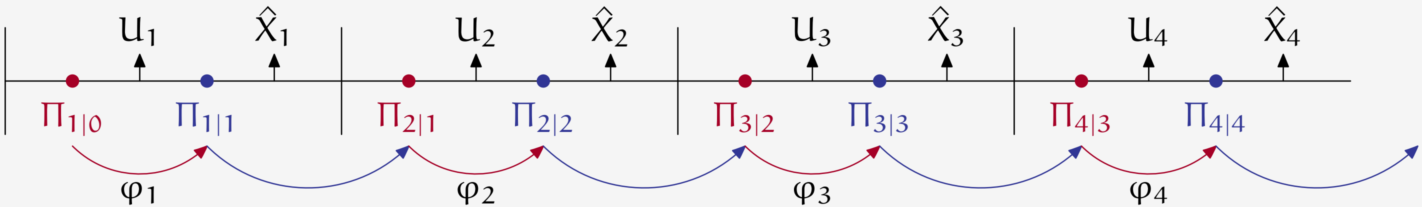
- Nayyar, Mahajan and Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," IEEE TAC 2013.

Information states and dynamic program

Information states

Pre-transmission belief : $\Pi_{t|t-1}(x) = \mathbb{P}(X_t = x \mid Y_{1:t-1})$.

Post-transmission belief : $\Pi_{t|t}(x) = \mathbb{P}(X_t = x \mid Y_{1:t})$.

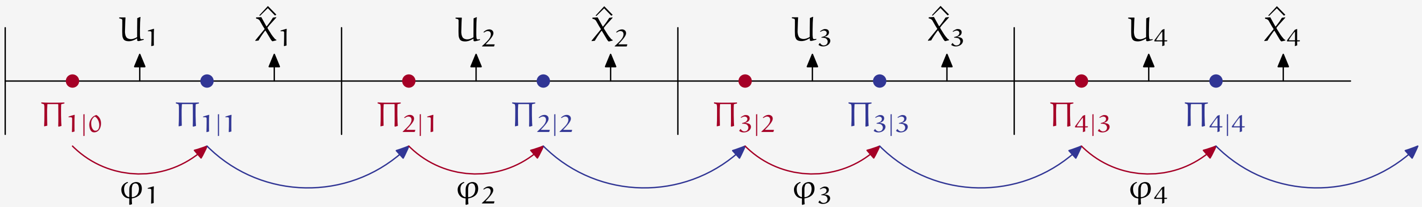


Information states and dynamic program

Information states

Pre-transmission belief : $\Pi_{t|t-1}(x) = \mathbb{P}(X_t = x \mid Y_{1:t-1})$.

Post-transmission belief : $\Pi_{t|t}(x) = \mathbb{P}(X_t = x \mid Y_{1:t})$.



Structural results

There is no loss of optimality in using

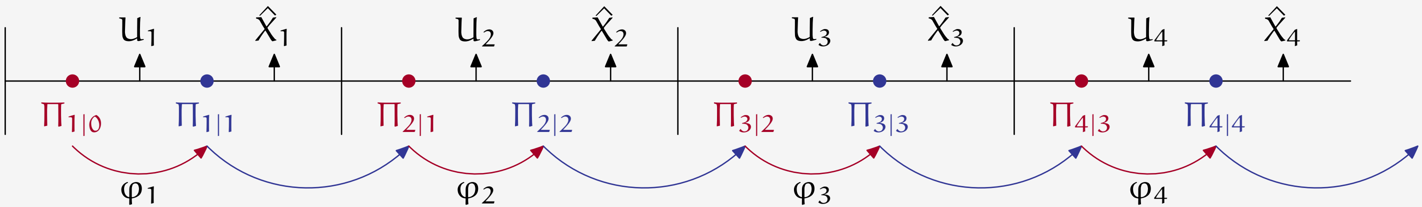
$$U_t = f_t(X_t, \Pi_{t|t-1}) \quad \text{and} \quad \hat{X}_t = g_t(\Pi_{t|t}).$$

Information states and dynamic program

Information states

Pre-transmission belief : $\Pi_{t|t-1}(x) = \mathbb{P}(X_t = x \mid Y_{1:t-1})$.

Post-transmission belief : $\Pi_{t|t}(x) = \mathbb{P}(X_t = x \mid Y_{1:t})$.



Structural results

There is no loss of optimality in using

$$U_t = f_t(X_t, \Pi_{t|t-1}) \quad \text{and} \quad \hat{X}_t = g_t(\Pi_{t|t}).$$

Dynamic Program

$$V_{T+1|T}(\pi) = 0, \quad \text{and for } t = T, \dots, 0$$

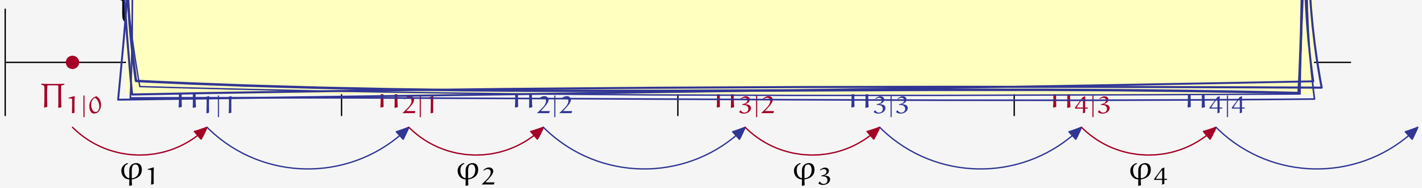
$$V_{t|t}(\pi) = \min_{\hat{x} \in \mathcal{X}} \mathbb{E}[d(X_t - \hat{x}) + V_{t+1|t}(\Pi_{t+1}) \mid \Pi_{t|t} = \pi],$$

$$V_{t|t-1}(\pi) = \min_{\varphi : \mathcal{X} \to \{0,1\}} \mathbb{E}[\lambda\varphi(X_t) + V_{t|t}(\Pi_{t|t}) \mid \Pi_{t|t-1} = \pi, \varphi_t = \varphi].$$

Information states and dynamic program

Information s

"Standard" POMDP. Optimal strategies can be computed numerically (at least, in principle).



Structural results

There is no loss of optimality in using

$$U_t = f_t(X_t, \Pi_{t|t-1}) \quad \text{and} \quad \hat{X}_t = g_t(\Pi_{t|t}).$$

Dynamic Program

$$V_{T+1|T}(\pi) = 0, \quad \text{and for } t = T, \dots, 0$$

$$V_{t|t}(\pi) = \min_{\hat{x} \in \mathcal{X}} \mathbb{E}[d(X_t - \hat{x}) + V_{t+1|t}(\Pi_{t+1}) \mid \Pi_{t|t} = \pi],$$

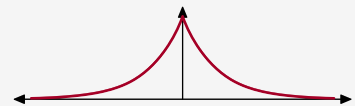
$$V_{t|t-1}(\pi) = \min_{\varphi : \mathcal{X} \to \{0,1\}} \mathbb{E}[\lambda \varphi(X_t) + V_{t|t}(\Pi_{t|t}) \mid \Pi_{t|t-1} = \pi, \varphi_t = \varphi].$$

Can we use the DP to say something more about the optimal strategy?

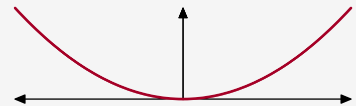
Simplifying modeling assumptions

Markov process $X_{t+1} = aX_t + W_t$

- Discrete state process: $X_t,\ a,\ W_t \in \mathbb{Z}$
- Continuous state process: $X_t,\ a,\ W_t \in \mathbb{R}$



Noise Distribution Unimodal and symmetric

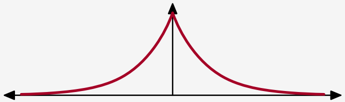


Distortion function Even and increasing

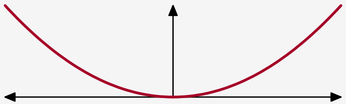
Simplifying modeling assumptions

Markov process $X_{t+1} = aX_t + W_t$

- Discrete state process: $X_t, a, W_t \in \mathbb{Z}$
- Continuous state process: $X_t, a, W_t \in \mathbb{R}$



Noise Distribution — Unimodal and symmetric



Distortion function — Even and increasing

Proof outline

Step 1 Show that threshold-based strategies are optimal

Step 2 Find performance of arbitrary threshold based strategies

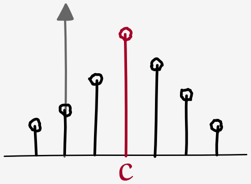
Step 3 Solution to the costly communication problem

Step 4 Solution to the constrained communication problem

Preliminaries

[Hajek Mitzel Yang 2008, Lipsa Martins 2011, Nayyar et. al. 2013]

Almost uniform and unimodal (ASU) distribution about c

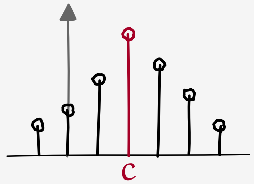


$$\pi_c \geqslant \pi_{c+1} \geqslant \pi_{c-1} \geqslant \pi_{c+2} \geqslant \cdots$$

Preliminaries

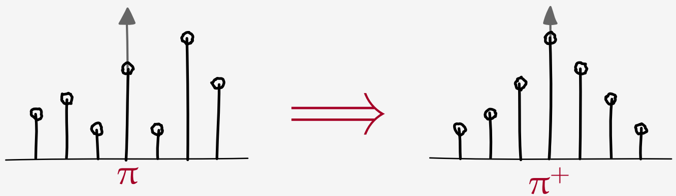
[Hajek Mitzel Yang 2008, Lipsa Martins 2011, Nayyar et. al. 2013]

Almost uniform and unimodal (ASU) distribution about c



$$\pi_c \geqslant \pi_{c+1} \geqslant \pi_{c-1} \geqslant \pi_{c+2} \geqslant \cdots$$

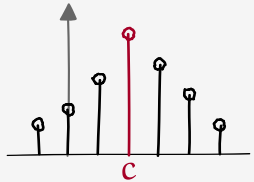
ASU Rearrangement



Preliminaries

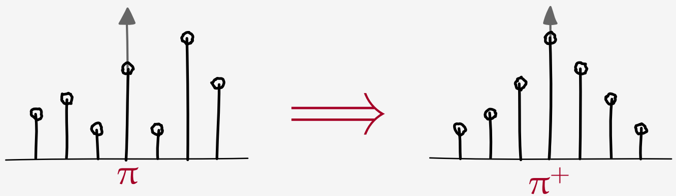
[Hajek Mitzel Yang 2008, Lipsa Martins 2011, Nayyar et. al. 2013]

Almost uniform and unimodal (ASU) distribution about c



$$\pi_c \geqslant \pi_{c+1} \geqslant \pi_{c-1} \geqslant \pi_{c+2} \geqslant \cdots$$

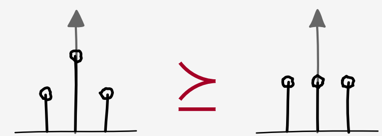
ASU Rearrangement



Majorization

$\pi \succeq \xi$ iff

$$\sum_{i=-n}^{n} \pi_i^+ \geqslant \sum_{i=-n}^{n} \xi_i^+ \quad \text{and} \quad \sum_{i=-n}^{n+1} \pi_i^+ \geqslant \sum_{i=-n}^{n+1} \xi_i^+$$



Invariant to permutations.

Step 1 Properties of the value function

Backward induction argument

- **Value function is "almost" Schur-concave**:
  If $\xi \succeq \pi$ and π is ASU, then $V_{t|t-1}(\xi) \geqslant V_{t|t-1}(\pi)$ and $V_{t|t}(\xi) \geqslant V_{t|t}(\pi)$

Step 1 Properties of the value function

Backward induction argument

- **Value function is "almost" Schur-concave**:
  If $\xi \succeq \pi$ and π is ASU, then $V_{t|t-1}(\xi) \geqslant V_{t|t-1}(\pi)$ and $V_{t|t}(\xi) \geqslant V_{t|t}(\pi)$

- **Optimal estimation strategy**:
  If π is ASU about c, then c is the arg min of

$$V_{t|t}(\pi) = \min_{\hat{x} \in \mathcal{X}} \mathbb{E}[d(X_t - \hat{x}) + V_{t+1|t}(\Pi_{t+1|t}) \mid \Pi_{t|t} = \pi],$$

- **Optimal transmission strategy**:
  If π is ASU about c, then the arg min of

$$V_{t|t-1}(\pi) = \min_{\varphi : \mathcal{X} \to \{0,1\}} \mathbb{E}[\lambda \varphi(X_t) + V_{t|t}(\Pi_{t|t}) \mid \Pi_{t|t} = \pi, \varphi_t = \varphi]$$

  is of the threshold form in $|x - ac|$.

Step 1 Properties of the value function

Define

Oblivious estimation process

$$Z_t = \begin{cases} X_t, & \text{if } Y_t \neq \mathfrak{E} \\ aZ_{t-1}, & \text{if } Y_t = \mathfrak{E} \end{cases}$$

Error process

$$E_t = X_t - aZ_{t-1}$$

Step 1 Properties of the value function

Define

Oblivious estimation process

$$Z_t = \begin{cases} X_t, & \text{if } Y_t \neq \mathfrak{E} \\ aZ_{t-1}, & \text{if } Y_t = \mathfrak{E} \end{cases}$$

Error process

$$E_t = X_t - aZ_{t-1}$$

Forward induction argument

- $\Pi_{t|t-1}$ is ASU around Z_{t-1}
- $\Pi_{t|t}$ is ASU around Z_t

Step 1 Properties of the value function

Define

Oblivious estimation process

$$Z_t = \begin{cases} X_t, & \text{if } Y_t \neq \mathfrak{E} \\ aZ_{t-1}, & \text{if } Y_t = \mathfrak{E} \end{cases}$$

Error process

$$E_t = X_t - aZ_{t-1}$$

Forward induction argument

- $\Pi_{t|t-1}$ is ASU around Z_{t-1}
- $\Pi_{t|t}$ is ASU around Z_t

Structure of optimal strategies

- **Optimal transmitter**: There exists thresholds $\{k_t\}_{t\geqslant 0}$ such that

$$U_t = f_t^*(E_t) = \begin{cases} 1 & \text{if } |E_t| \geqslant k_t \\ 0 & \text{if } |E_t| < k_t \end{cases}$$

- **Optimal estimator**: $\hat{X}_t = g_t^*(Z_t) = Z_t$

Some comments

The result is non-intuitive

- The transmitter does not try to send information through timing information.
- The estimation strategy is the same to the one for intermittent observations and **does not depend on the choice of the threshold**

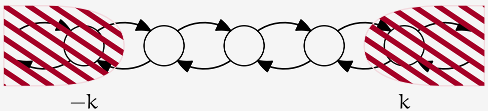
For infinite-horizon setup time-homogeneous threshold-based strategies are optimal.

How do we find the optimal threshold-based strategy?

Step 2 Performance of threshold-based strategies

Consider a threshold-based strategy

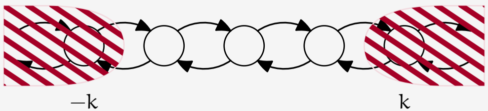
$$f^{(k)}(e) = \begin{cases} 1 & \text{if } |e| \geq k \\ 0 & \text{otherwise} \end{cases}$$



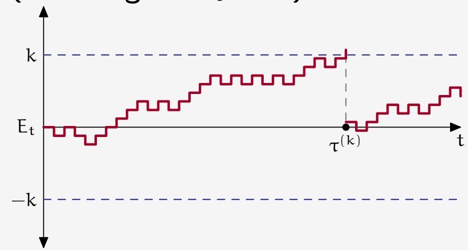
Step 2 Performance of threshold-based strategies

Consider a threshold-based strategy

$$f^{(k)}(e) = \begin{cases} 1 & \text{if } |e| \geq k \\ 0 & \text{otherwise} \end{cases}$$



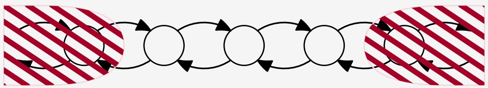
Let $\tau^{(k)}$ denote the stopping time of first reception (starting at $E_0 = 0$).



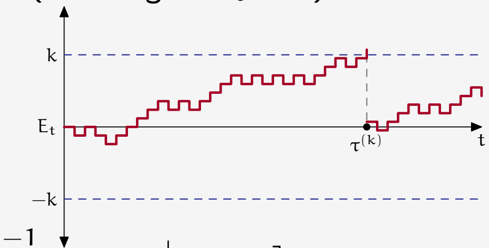
Step 2 Performance of threshold-based strategies

Consider a threshold-based strategy

$$f^{(k)}(e) = \begin{cases} 1 & \text{if } |e| \geq k \\ 0 & \text{otherwise} \end{cases}$$



Let $\tau^{(k)}$ denote the stopping time of first reception (starting at $E_0 = 0$).



Define

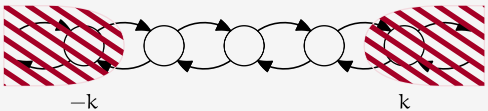
$$L_\beta^{(k)}(e) = (1-\beta)\,\mathbb{E}\Bigg[\sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_t) \Bigg| E_0 = e\Bigg].$$

$$M_\beta^{(k)}(e) = (1-\beta)\,\mathbb{E}\Bigg[\sum_{t=0}^{\tau^{(k)}-1} \beta^t \Bigg| E_0 = e\Bigg].$$

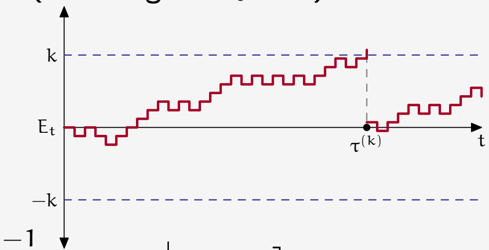
Step 2 Performance of threshold-based strategies

Consider a threshold-based strategy

$$f^{(k)}(e) = \begin{cases} 1 & \text{if } |e| \geq k \\ 0 & \text{otherwise} \end{cases}$$



Let $\tau^{(k)}$ denote the stopping time of first reception (starting at $E_0 = 0$).



Define

$$L^{(k)}_\beta(e) = (1-\beta)\,\mathbb{E}\Big[\sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_t) \Big| E_0 = e\Big].$$

$$M^{(k)}_\beta(e) = (1-\beta)\,\mathbb{E}\Big[\sum_{t=0}^{\tau^{(k)}-1} \beta^t \Big| E_0 = e\Big].$$

Proposition $\{E_t\}_{t=0}^{\infty}$ is a regenerative process. By renewal theory,

$$D^{(k)}_\beta \coloneqq D_\beta(f^{(k)}, g^*) = \frac{L^{(k)}_\beta(0)}{M^{(k)}_\beta(0)} \quad \text{and} \quad N^{(k)}_\beta \coloneqq N_\beta(f^{(k)}, g^*) = \frac{1}{M^{(k)}_\beta(0)} - (1-\beta).$$

Step 2 Performance of threshold-based strategies

Consider … :eption

$f^{(k)}(e$

Computing $L_\beta^{(k)}$ and $M_\beta^{(k)}$ is sufficient to compute the performance of $f^{(k)}$ (i.e., to compute $D_\beta^{(k)}$ and $N_\beta^{(k)}$).

Define

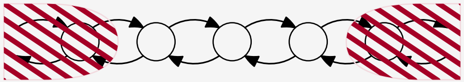
$$L_\beta^{(k)}(e) = (1-\beta)\,\mathbb{E}\Big[\sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_t) \Big| E_0 = e\Big].$$

$$M_\beta^{(k)}(e) = (1-\beta)\,\mathbb{E}\Big[\sum_{t=0}^{\tau^{(k)}-1} \beta^t \Big| E_0 = e\Big].$$

Proposition $\{E_t\}_{t=0}^\infty$ is a regenerative process. By renewal theory,

$$D_\beta^{(k)} := D_\beta(f^{(k)}, g^*) = \frac{L_\beta^{(k)}(0)}{M_\beta^{(k)}(0)} \quad \text{and} \quad N_\beta^{(k)} := N_\beta(f^{(k)}, g^*) = \frac{1}{M_\beta^{(k)}(0)} - (1-\beta).$$

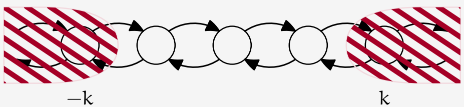
Step 2 Computing $L_\beta^{(k)}$ and $M_\beta^{(k)}$



$$L_\beta^{(k)}(e) = \begin{cases} d(e) + \beta \sum_{n\in\mathbb{Z}} p_{n-e} L_\beta^{(k)}(n), & \text{if } |e| < k \\ \varepsilon\Big[d(e) + \beta \sum_{n\in\mathbb{Z}} p_{n-ae} L_\beta^{(k)}(n)\Big], & \text{if } |e| \geqslant k \end{cases}$$

$$M_\beta^{(k)}(e) = \begin{cases} 1 + \beta \sum_{n\in\mathbb{Z}} p_{n-e} M_\beta^{(k)}(n), & \text{if } |e| < k \\ \varepsilon\Big[1 + \beta \sum_{n\in\mathbb{Z}} p_{n-ae} M_\beta^{(k)}(n)\Big], & \text{if } |e| \geqslant k \end{cases}$$

Step 2 Computing $L_\beta^{(k)}$ and $M_\beta^{(k)}$

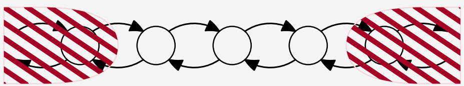


$$L_\beta^{(k)}(e) = \begin{cases} d(e) + \beta \sum_{n\in\mathbb{Z}} p_{n-e} L_\beta^{(k)}(n), & \text{if } |e| < k \\ \varepsilon\Big[d(e) + \beta \sum_{n\in\mathbb{Z}} p_{n-ae} L_\beta^{(k)}(n)\Big], & \text{if } |e| \geqslant k \end{cases}$$

$$M_\beta^{(k)}(e) = \begin{cases} 1 + \beta \sum_{n\in\mathbb{Z}} p_{n-e} M_\beta^{(k)}(n), & \text{if } |e| < k \\ \varepsilon\Big[1 + \beta \sum_{n\in\mathbb{Z}} p_{n-ae} M_\beta^{(k)}(n)\Big], & \text{if } |e| \geqslant k \end{cases}$$

$$\begin{bmatrix} \vdots \\ L_\beta^{(k)}(-2) \\ L_\beta^{(k)}(-1) \\ L_\beta^{(k)}(0) \\ L_\beta^{(k)}(1) \\ L_\beta^{(k)}(2) \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \varepsilon d(-2) \\ d(-1) \\ d(0) \\ d(1) \\ \varepsilon d(2) \\ \vdots \end{bmatrix} + \beta \begin{bmatrix} \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \cdots & \varepsilon p_1 & \varepsilon p_2 & \varepsilon p_3 & \varepsilon p_4 & \varepsilon p_5 & \cdots \\ \cdots & p_0 & p_1 & p_2 & p_3 & p_4 & \cdots \\ \cdots & p_{-1} & p_0 & p_1 & p_2 & p_3 & \cdots \\ \cdots & p_{-2} & p_{-1} & p_0 & p_1 & p_2 & \cdots \\ \cdots & \varepsilon p_{-3} & \varepsilon p_{-2} & \varepsilon p_{-1} & \varepsilon p_0 & \varepsilon p_1 & \cdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \end{bmatrix} \begin{bmatrix} \vdots \\ L_\beta^{(k)}(-2) \\ L_\beta^{(k)}(-1) \\ L_\beta^{(k)}(0) \\ L_\beta^{(k)}(1) \\ L_\beta^{(k)}(2) \\ \vdots \end{bmatrix}$$

Step 2 Computing $L_\beta^{(k)}$ and $M_\beta^{(k)}$



$$L_\beta^{(k)}(e) = \begin{cases} d(e) + \beta \sum_{n\in\mathbb{Z}} p_{n-e} L_\beta^{(k)}(n), & \text{if } |e| < k \\ \varepsilon\left[d(e) + \beta \sum_{n\in\mathbb{Z}} p_{n-ae} L_\beta^{(k)}(n)\right], & \text{if } |e| \geqslant k \end{cases}$$

$$M_\beta^{(k)}(e) = \begin{cases} 1 + \beta \sum_{n\in\mathbb{Z}} p_{n-e} M_\beta^{(k)}(n), & \text{if } |e| < k \\ \varepsilon\left[1 + \beta \sum_{n\in\mathbb{Z}} p_{n-ae} M_\beta^{(k)}(n)\right], & \text{if } |e| \geqslant k \end{cases}$$

Proposition

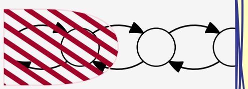
$$L_\beta^{(k)} = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)} \odot d,$$

$h^{(k)} \odot P$ is substochastic.

$$M_\beta^{(k)} = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)}.$$

Step 2 Computing $L_\beta^{(k)}$ and $M_\beta^{(k)}$

$$\begin{cases} d(e) + \beta \sum p_{n-e} L_\beta^{(k)}(n), & \text{if } |e| < k \end{cases}$$



$D_\beta^{(k)}$ and $N_\beta^{(k)}$ can be computed using these expressions.

$$M_\beta^{(k)}(e) = \begin{cases} 1 + \beta \sum_{n\in\mathbb{Z}} p_{n-e} M_\beta^{(k)}(n), & \text{if } |e| < k \\ \varepsilon\Big[1 + \beta \sum_{n\in\mathbb{Z}} p_{n-ae} M_\beta^{(k)}(n)\Big], & \text{if } |e| \geqslant k \end{cases}$$

Proposition $L_\beta^{(k)} = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)} \odot d$, $\quad h^{(k)} \odot P$ is substochastic.

$M_\beta^{(k)} = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)}$.

Step 3 Solution to costly optimization problem

Proposition

- $C^{(k)}_\beta(\lambda) \coloneqq D^{(k)}_\beta + \lambda N^{(k)}_\beta$ is submodular in (k, λ).
- Hence, $k^*_\beta(\lambda) \coloneqq \arg\min_{k \geq 0} C^{(k)}_\beta(\lambda)$ is increasing in λ

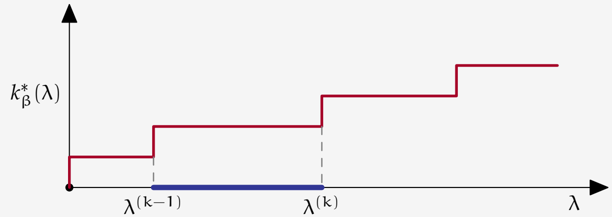
Step 3 Solution to costly optimization problem

Proposition

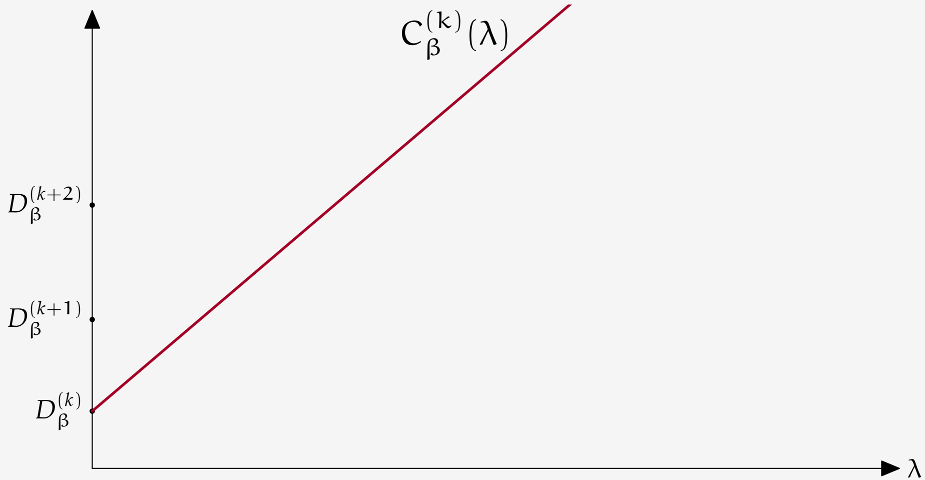
- $C^{(k)}_\beta(\lambda) \coloneqq D^{(k)}_\beta + \lambda N^{(k)}_\beta$ is submodular in (k, λ).
- Hence, $k^*_\beta(\lambda) \coloneqq \arg\min_{k \geqslant 0} C^{(k)}_\beta(\lambda)$ is increasing in λ

Define $$\Lambda^{(k)}_\beta \coloneqq \{\lambda \in \mathbb{R}_{\geqslant 0} : k^*_\beta(\lambda) = k\} = [\lambda^{(k-1)}_\beta, \lambda^{(k)}_\beta].$$

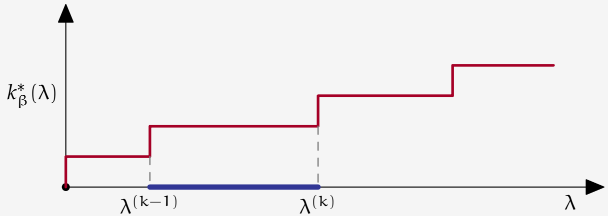
$$C^{(k)}_\beta(\lambda^{(k)}_\beta) = C^{(k+1)}_\beta(\lambda^{(k)}_\beta)$$



Step 3 Solution to costly optimization problem

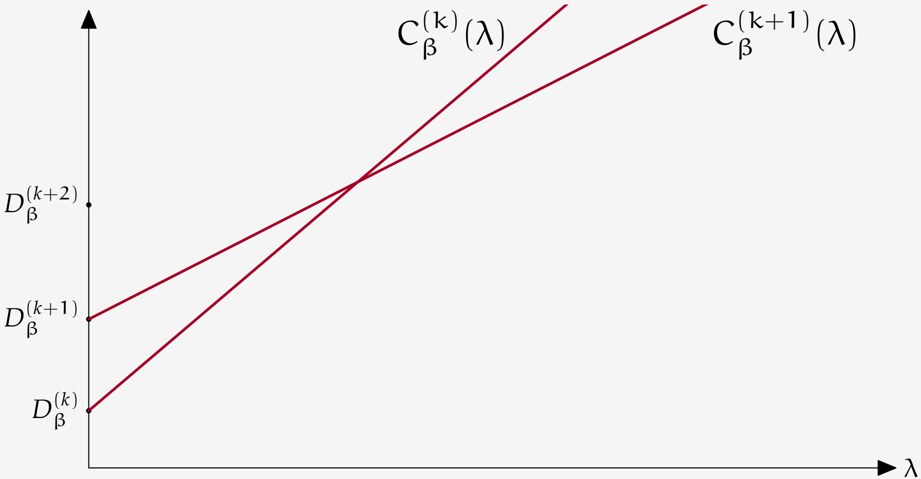


Define $\Lambda_\beta^{(k)} \coloneqq \{\lambda \in \mathbb{R}_{\geq 0} : k^*_\beta(\lambda) = k\}$
$= [\lambda_\beta^{(k-1)}, \lambda_\beta^{(k)}]$.



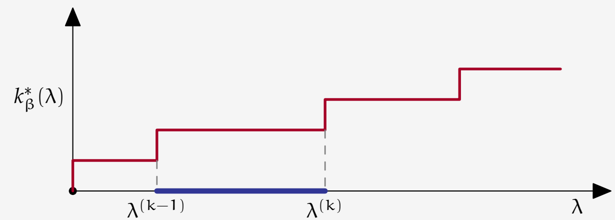
$$C_\beta^{(k)}(\lambda_\beta^{(k)}) = C_\beta^{(k+1)}(\lambda_\beta^{(k)})$$

Step 3 Solution to costly optimization problem

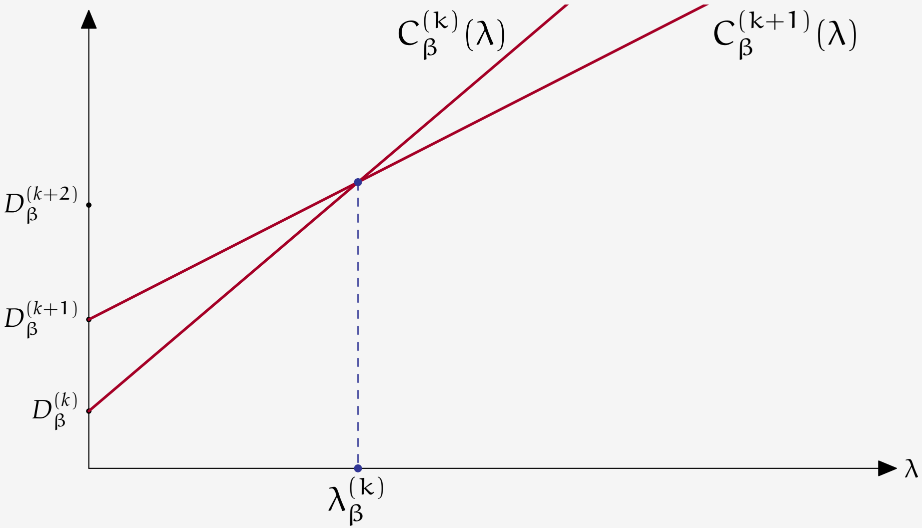


Define $\Lambda_\beta^{(k)} \coloneqq \{\lambda \in \mathbb{R}_{\geqslant 0} : k_\beta^*(\lambda) = k\}$
$= [\lambda_\beta^{(k-1)}, \lambda_\beta^{(k)}]$.

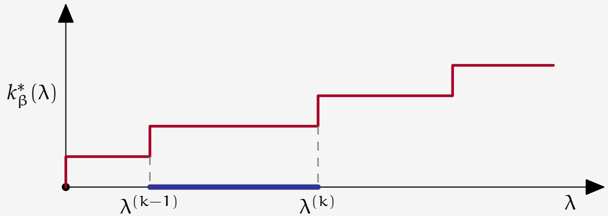
$$C_\beta^{(k)}(\lambda_\beta^{(k)}) = C_\beta^{(k+1)}(\lambda_\beta^{(k)})$$



Step 3 Solution to costly optimization problem

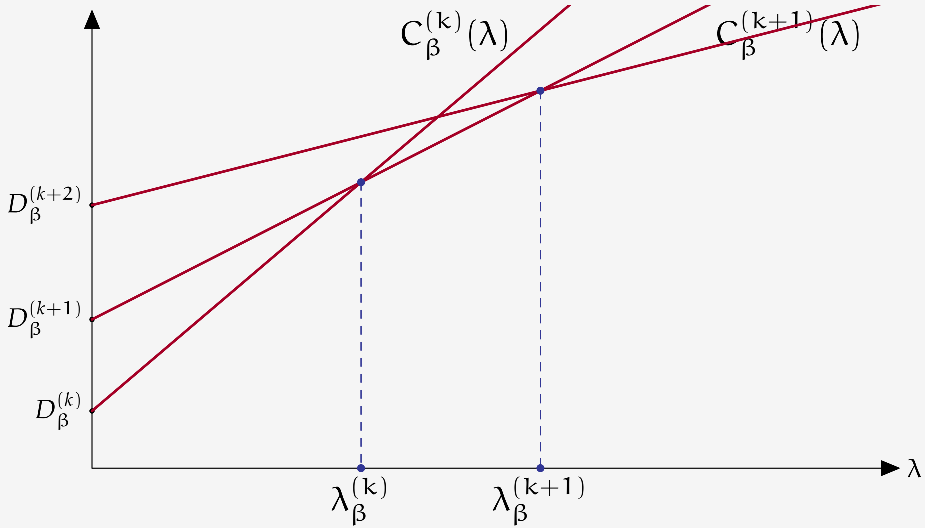


Define $\Lambda_\beta^{(k)} := \{\lambda \in \mathbb{R}_{\geqslant 0} : k_\beta^*(\lambda) = k\}$
$= [\lambda_\beta^{(k-1)}, \lambda_\beta^{(k)}]$.



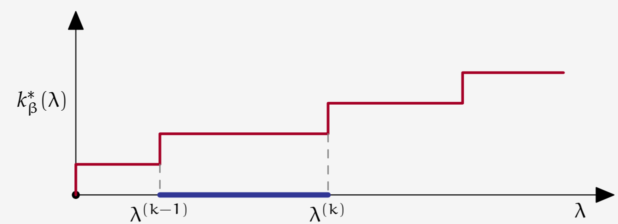
$$C_\beta^{(k)}(\lambda_\beta^{(k)}) = C_\beta^{(k+1)}(\lambda_\beta^{(k)})$$

Step 3 Solution to costly optimization problem

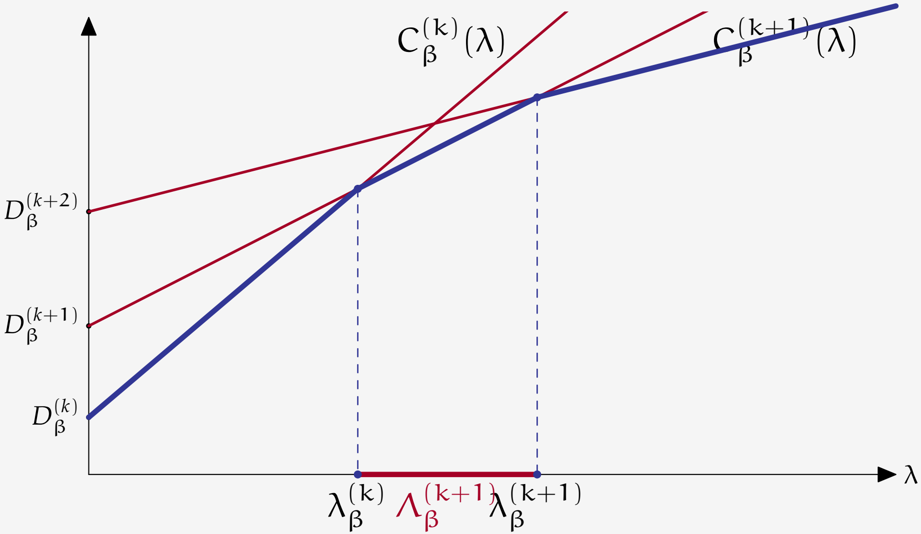


Define $\Lambda_\beta^{(k)}$
$$\Lambda_\beta^{(k)} \coloneqq \{\lambda \in \mathbb{R}_{\geqslant 0} : k^*_\beta(\lambda) = k\} = [\lambda_\beta^{(k-1)}, \lambda_\beta^{(k)}].$$

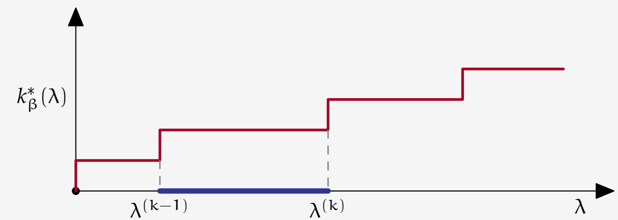
$$C_\beta^{(k)}(\lambda_\beta^{(k)}) = C_\beta^{(k+1)}(\lambda_\beta^{(k)})$$



Step 3 Solution to costly optimization problem

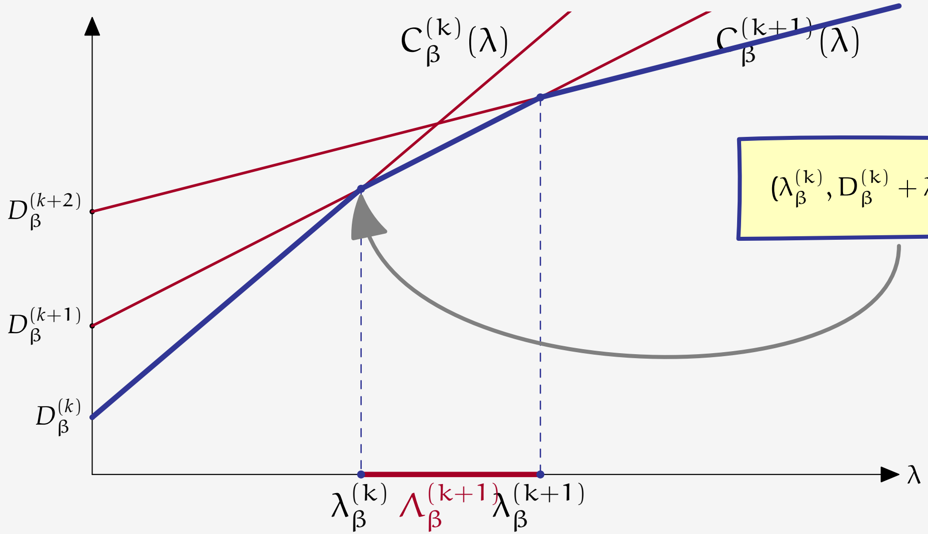


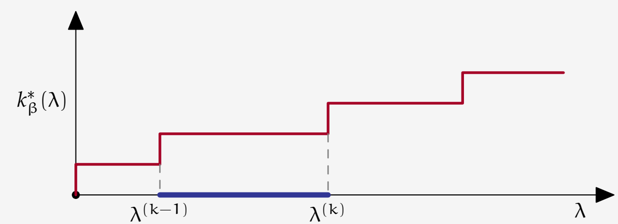
Define $\Lambda_\beta^{(k)} \coloneqq \{\lambda \in \mathbb{R}_{\geqslant 0} : k^*_\beta(\lambda) = k\}$
$= [\lambda_\beta^{(k-1)}, \lambda_\beta^{(k)}]$.



$$C_\beta^{(k)}(\lambda_\beta^{(k)}) = C_\beta^{(k+1)}(\lambda_\beta^{(k)})$$

Step 3 Solution to costly optimization problem

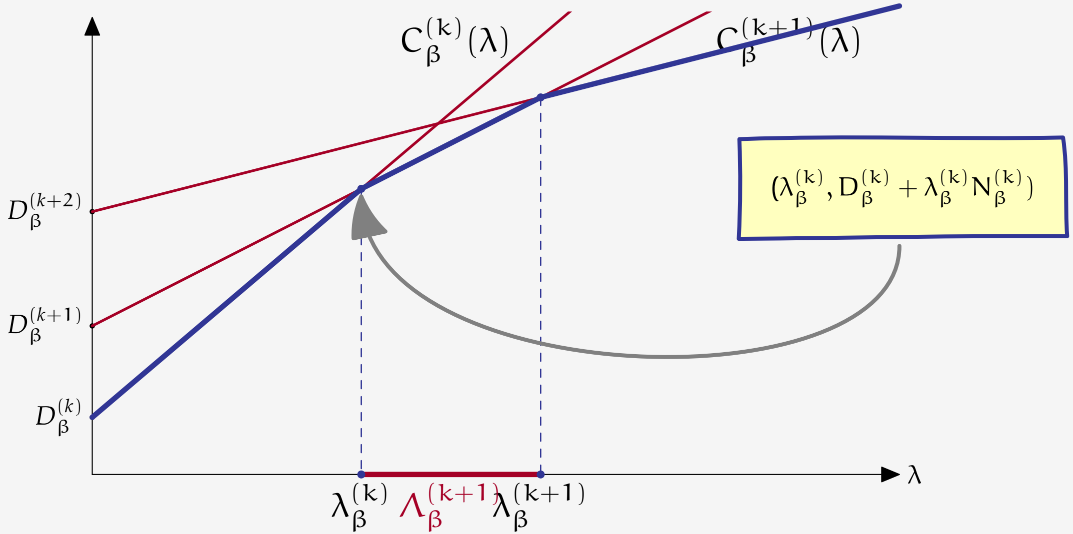




Define $\Lambda_\beta^{(k)} \coloneqq \{\lambda \in \mathbb{R}_{\geq 0} : k_\beta^*(\lambda) = k\}$
$= [\lambda_\beta^{(k-1)}, \lambda_\beta^{(k)}].$

$$C_\beta^{(k)}(\lambda_\beta^{(k)}) = C_\beta^{(k+1)}(\lambda_\beta^{(k)})$$

Step 3 Solution to costly optimization problem



Theorem

Strategy $f^{(k+1)}$ is optimal for $\lambda \in (\lambda_\beta^{(k)}, \lambda_\beta^{(k+1)}]$.

$C_\beta^*(\lambda) = \min_{k \in \mathbb{Z}_{\geqslant 0}} C_\beta^{(k)}$ is piecewise linear, continuous, concave, and increasing function of λ.

Step 4 Solution to constrained communication problem

Sufficient condition for optimality

A strategy (f°, g°) is optimal for the constrained problem if

(C1) $N_\beta(f^\circ, g^\circ) = \alpha$

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for the Lagrange relaxation with parameter λ°.

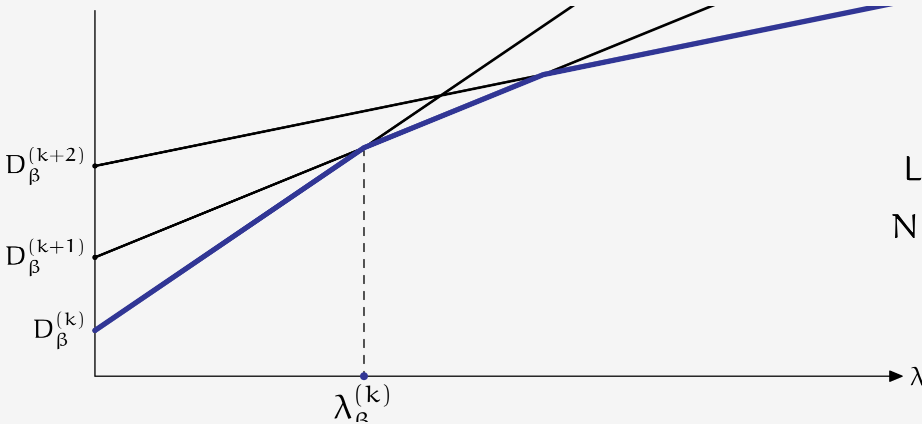
Step 4 Solution to constrained communication problem

Sufficient condition for optimality

A strategy (f°, g°) is optimal for the constrained problem if

(C1) $N_\beta(f^\circ, g^\circ) = \alpha$

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for the Lagrange relaxation with parameter λ°.



Let k_β^* be such that
$N_\beta^{(k_\beta^*)} > \alpha > N_\beta^{(k_\beta^*+1)}$

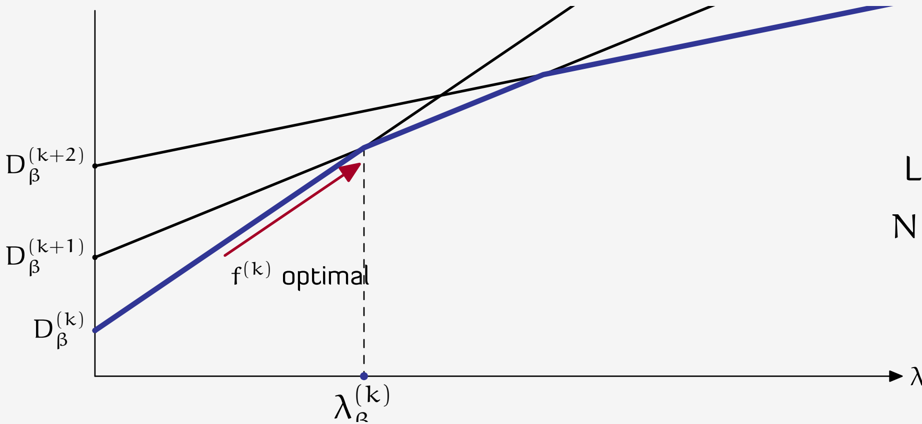
Step 4 Solution to constrained communication problem

Sufficient condition for optimality

A strategy (f°, g°) is optimal for the constrained problem if

(C1) $N_\beta(f^\circ, g^\circ) = \alpha$

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for the Lagrange relaxation with parameter λ°.



Let k_β^* be such that
$N_\beta^{(k_\beta^*)} > \alpha > N_\beta^{(k_\beta^*+1)}$

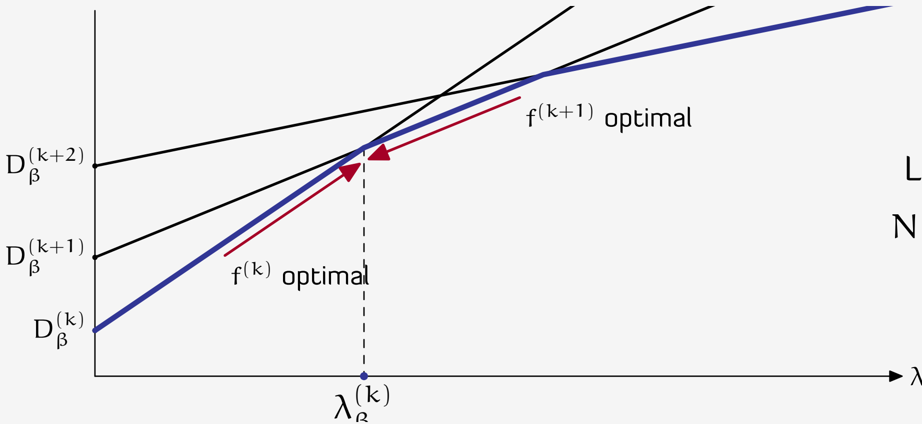
Step 4 Solution to constrained communication problem

Sufficient condition for optimality

A strategy (f°, g°) is optimal for the constrained problem if

(C1) $N_\beta(f^\circ, g^\circ) = \alpha$

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for the Lagrange relaxation with parameter λ°.



Let k_β^* be such that
$N_\beta^{(k_\beta^*)} > \alpha > N_\beta^{(k_\beta^*+1)}$

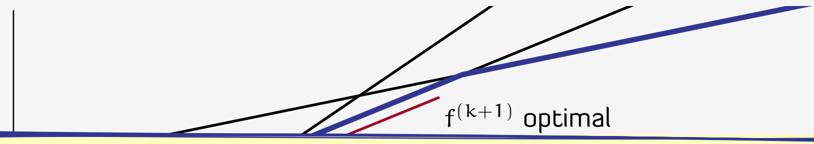
Step 4 Solution to constrained communication problem

Sufficient condition for optimality

A strategy (f°, g°) is optimal for the constrained problem if

(C1) $N_\beta(f^\circ, g^\circ) = \alpha$

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for the Lagrange relaxation with parameter λ°.



such that

$> N_\beta^{(k^*_\beta+1)}$

Randomized strategy $(\theta^*, f^{(k)}, f^{k+1})$ is optimal where

$$\theta^* N_\beta^{(k)} + (1-\theta^*) N_\beta^{(k+1)} = \alpha$$

Step 4 Solution to constrained communication problem

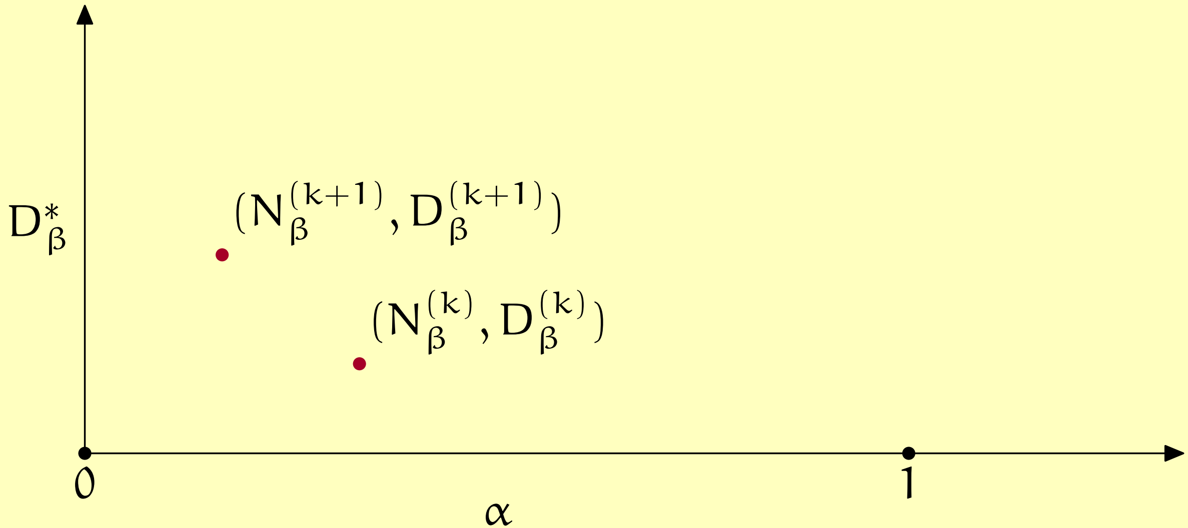
Sufficient condition for optimality

A strate

(C1

(C2 rameter λ°.

R



D^*_β is PWL, dec, and convex

Step 4 Solution to constrained communication problem

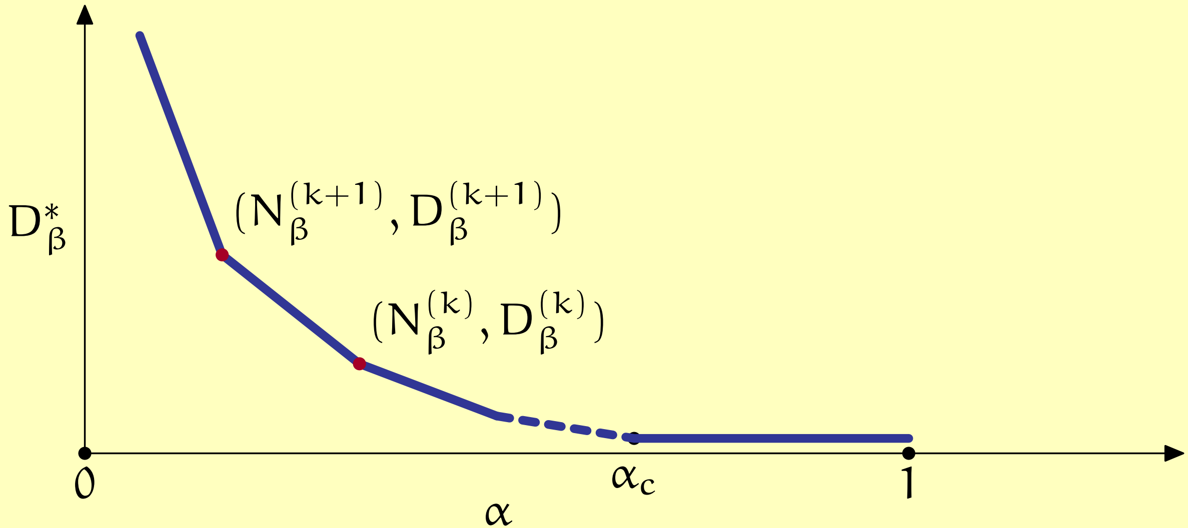
Sufficient condition for optimality

A strate

(C1

(C2 rameter λ°.

R

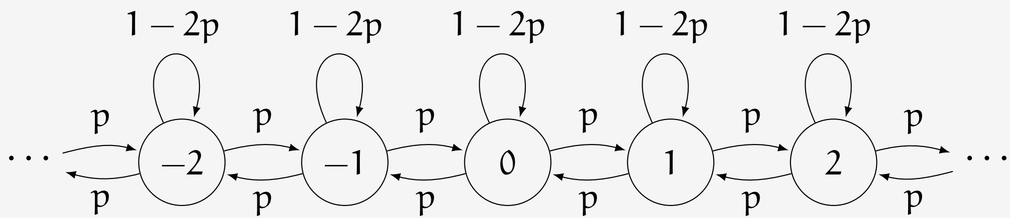


D^*_β is PWL, dec, and convex

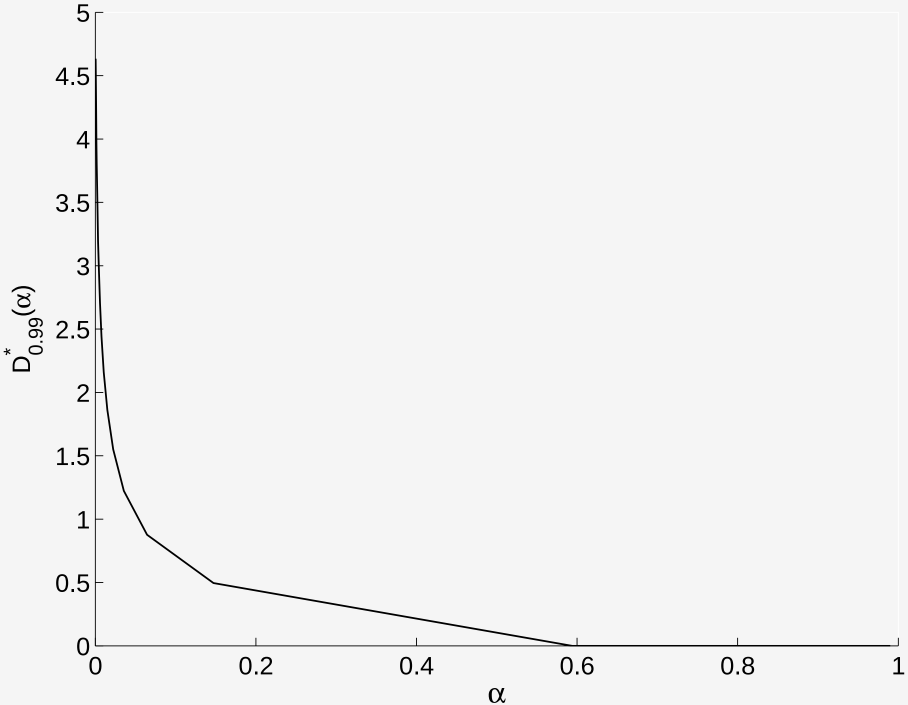
Example

Example Symmetric birth-death Markov chain

$$p_n = \begin{cases} p, & \text{if } |n| = 1; \\ 1 - 2p, & \text{if } n = 0; \\ 0, & \text{otherwise,} \end{cases} \qquad \text{where } p \in (0, \tfrac{1}{3}), \qquad d(e) = |e|$$

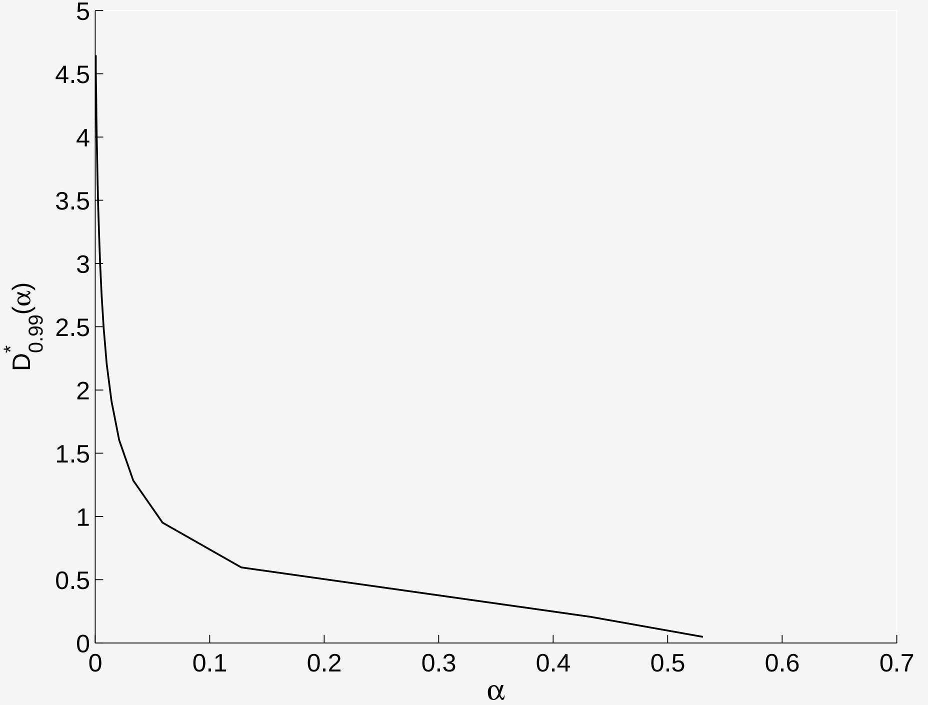


Example Symmetric birth-death Markov chain $(p = 0.3)$



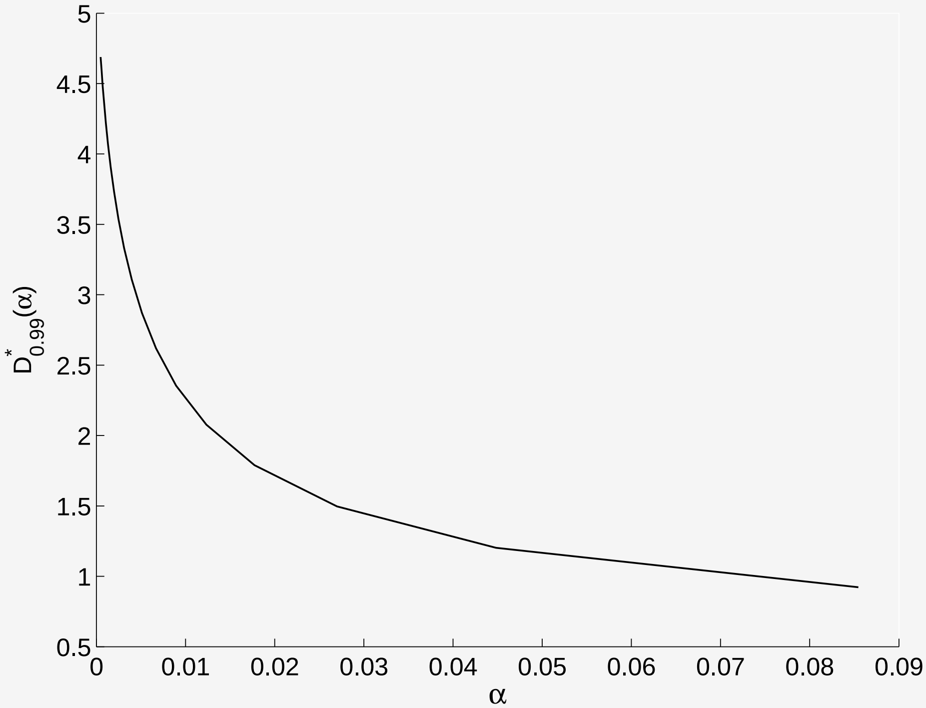
$\varepsilon = 0$

Example Symmetric birth-death Markov chain $(p = 0.3)$



$\varepsilon = 0.3$

Example Symmetric birth-death Markov chain $(p = 0.3)$

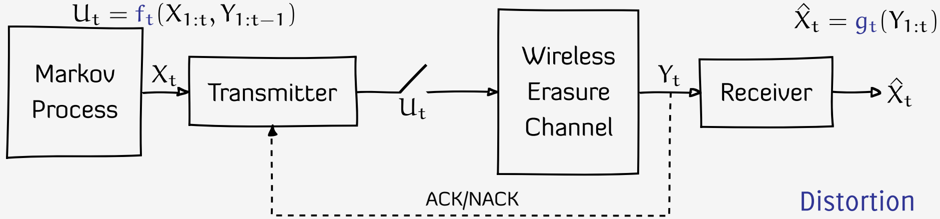


$$\varepsilon = 0.7$$

Summary

$$Y_t = \begin{cases} X_t, & \text{if } U_t H_t = 1 \\ \mathfrak{E}, & \text{if } U_t H_t = 0 \end{cases} \qquad \mathbb{P}(H_t = 0) = \varepsilon$$

$$U_t = f_t(X_{1:t}, Y_{1:t-1}) \qquad \hat{X}_t = g_t(Y_{1:t})$$



Distortion
$d(X_t - \hat{X}_t)$

1. Discounted setup, $\beta \in (0, 1)$

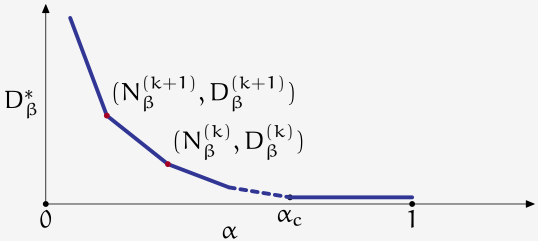
$$D_\beta(f, g) = (1-\beta)\,\mathbb{E}_0^{(f,g)}\left[\sum_{t=0}^{\infty} \beta^t d(X_t - \hat{X}_t)\right]; \qquad N_\beta(f, g) = (1-\beta)\,\mathbb{E}_0^{(f,g)}\left[\sum_{t=0}^{\infty} \beta^t U_t\right]$$

2. Average cost setup, $\beta = 1$

$$D_1(f, g) = \limsup_{T\to\infty} \frac{1}{T}\,\mathbb{E}_0^{(f,g)}\left[\sum_{t=0}^{T-1} d(X_t - \hat{X}_t)\right]; \qquad N_1(f, g) = \limsup_{T\to\infty} \frac{1}{T}\,\mathbb{E}_0^{(f,g)}\left[\sum_{t=0}^{T-1} U_t\right]$$

Summary

Distortion transmission function for discrete auto-regressive sources



D^*_β is PWL, dec, and convex

1.

Optimal transmission strategy

- Find k^* such that $\alpha \in (N_\beta^{(k^*+1)}, N_\beta^{(k^*)}]$.
- Compute θ^* such that
  $\theta^* N_\beta^{(k)} + (1-\theta^*) N_\beta^{(k+1)} = \alpha$
- If $|X_t - a\hat{X}_{t-1}| > k^*(\alpha)$, transmit.
- If $|X_t - a\hat{X}_{t-1}| = k^*(\alpha)$, transmit w.p. θ^*.
- Else, do not transmit.

Optimal estimation strategy

$$\hat{X}_t = \begin{cases} Y_t, & \text{if } Y_t \neq \mathfrak{E} \\ a\hat{X}_{t-1}, & \text{if } Y_t = \mathfrak{E} \end{cases}$$

2.

How to compute $D^*_\beta(\alpha)$

- Compute $L_\beta^{(k)} = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)} \odot d.$

$$M_\beta^{(k)} = [I - \beta h^{(k)} \odot P]^{-1} h^{(k)}.$$

- Then $D_\beta^{(k)} = \dfrac{L^{(k)}(0)}{M_\beta^{(k)}(0)}$ and $N_\beta^{(k)} = \dfrac{1}{M_\beta^{(k)}(0)} - (1-\beta)$

Fud

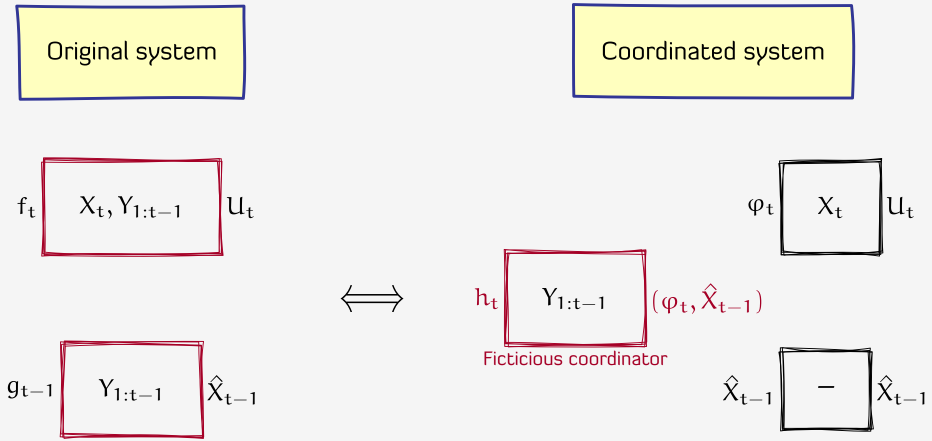
Summary

Distortion transmission function for discrete auto-regressive sources

The common information approach (Nayyar, Mahajan, Teneketzis 2013)

1.

2.



- The coordinated system is equivalent to the original system.

$$f_t(x, y_{1:t-1}) = h_t^1(y_{1:t-1})(x).$$

- The coordinated system is centralized. Belief state $\mathbb{P}(X_t \mid Y_{1:t-1})$.
- Nayyar, Mahajan and Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," IEEE TAC 2013.

Summary

Distortion transmission function for discrete auto-regressive sources

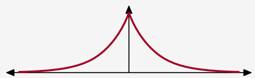
The common information approach (Nayyar, Mahajan, Teneketzis 2013)

Simplifying modeling assumptions

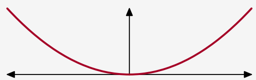
Markov process $X_{t+1} = aX_t + W_t$

- Discrete state process: $X_t, a, W_t \in \mathbb{Z}$
- Continuous state process: $X_t, a, W_t \in \mathbb{R}$

Noise Distribution Unimodal and symmetric



Distortion function Even and increasing



Proof outline

Step 1 Show that threshold-based strategies are optimal

Step 2 Find performance of arbitrary threshold based strategies

Step 3 Solution to the costly communication problem

Step 4 Solution to the constrained communication problem

Concluding Remarks

Presented results for discounted cost and countable state space

The results also apply to

- Long-term average setup (using the vanishing discount approach)
- Continuous state space (use Fredholm integral equations to compute $L_\beta^{(k)}$ and $M_\beta^{(k)}$)

Concluding Remarks

Presented results for discounted cost and countable state space

The results also apply to

- Long-term average setup (using the vanishing discount approach)
- Continuous state space (use Fredholm integral equations to compute $L_\beta^{(k)}$ and $M_\beta^{(k)}$)

Results are derived under idealized assumptions

Future directions

- Power or rate control . . .
- Scheduling multiple sources . . .
- Model network delays . . .

Concluding Remarks

Presented results for discounted cost and countable state space

The results also apply to

- Long-term average setup (using the vanishing discount approach)
- Continuous state space (use Fredholm integral equations to compute $L_\beta^{(k)}$ and $M_\beta^{(k)}$)

Results are derived under idealized assumptions

Future directions

- Power or rate control . . .
- Scheduling multiple sources . . .
- Model network delays . . .

Full version available at arXiv:1505.04829.