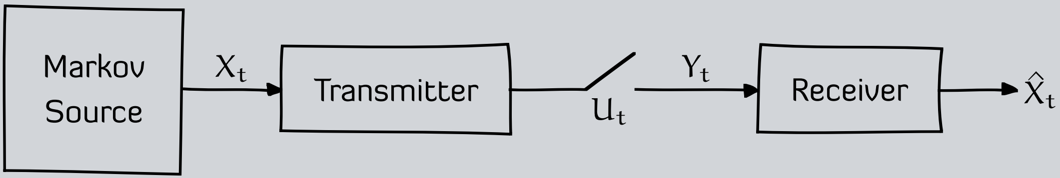
Distortion-transmission trade-off in real-time transmission of Markov sources

Jhelum Chakravorty and Aditya Mahajan

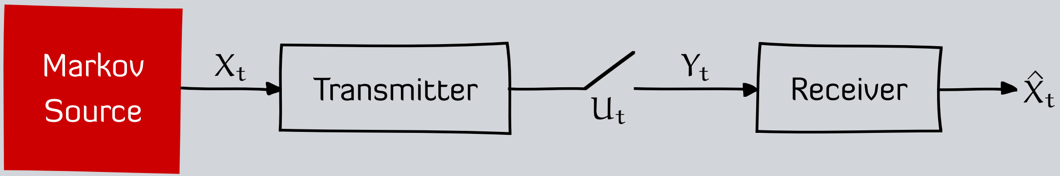
McGill University

IEEE Information Theory Workshop (ITW)
28 April, 2015

The system model



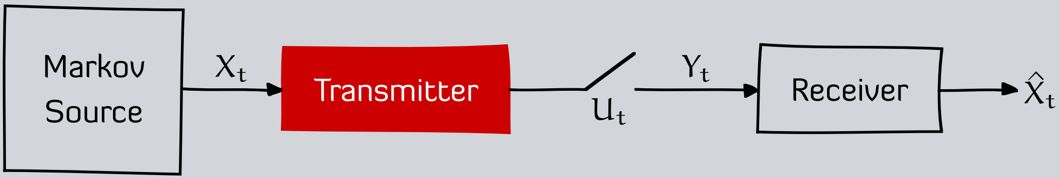
The system model



$$X_{t+1} = X_t + W_t$$

The system model

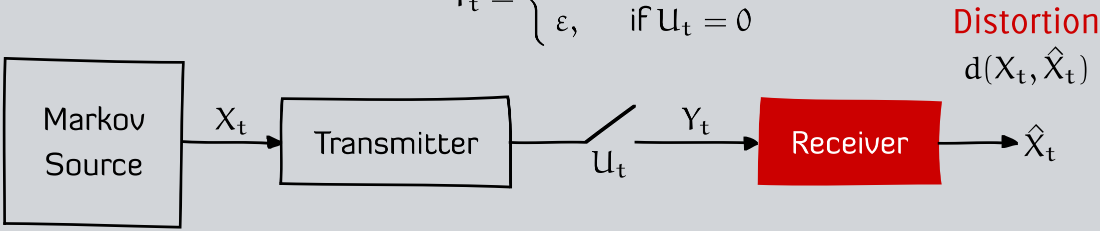
$$Y_t = \begin{cases} X_t, & \text{if } U_t = 1 \\ \varepsilon, & \text{if } U_t = 0 \end{cases}$$



$$X_{t+1} = X_t + W_t \qquad U_t = f_t(X_{1:t}, U_{1:t-1})$$

The system model

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



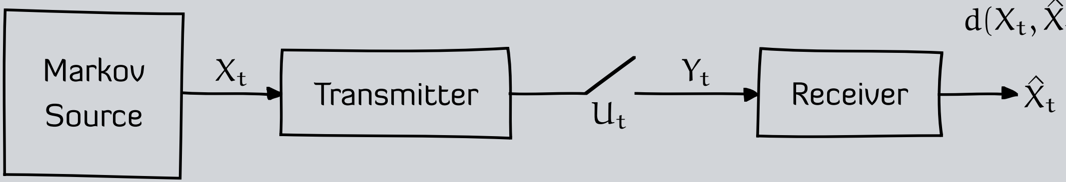
$$X_{t+1} = X_t + W_t \qquad U_t = f_t(X_{1:t}, U_{1:t-1}) \qquad \hat{X}_t = g_t(Y_{1:t})$$

The system model

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

Distortion

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



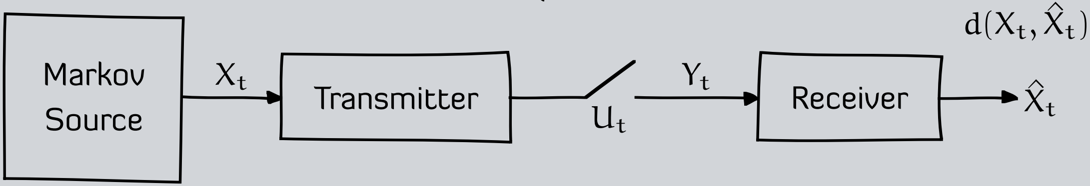
$X_{t+1} = X_t + W_t$ $\quad U_t = f_t(X_{1:t}, U_{1:t-1})$ $\quad \hat{X}_t = g_t(Y_{1:t})$

Communication Strategies

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

The system model

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



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

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

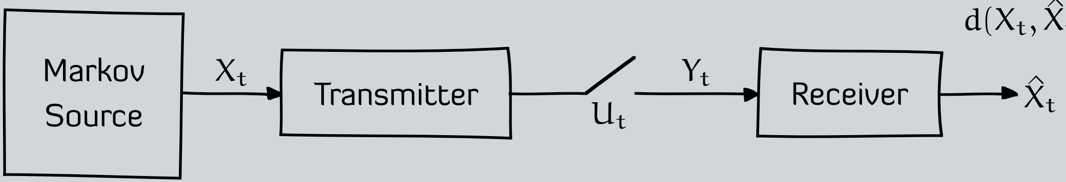
$$D_\beta(f, g) = (1 - \beta)\, \mathbb{E}_0^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t d(X_t - \hat{X}_t)\Big]; \qquad N_\beta(f, g) = (1 - \beta)\, \mathbb{E}_0^{(f,g)}\Big[\sum_{t=0}^{\infty} \beta^t U_t\Big]$$

The system model

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

Distortion

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



$X_{t+1} = X_t + W_t$ $\quad U_t = f_t(X_{1:t}, U_{1:t-1})$ $\quad \hat{X}_t = g_t(Y_{1:t})$

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

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

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

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

Optimization problems

Costly communication

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

Constrained communication

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

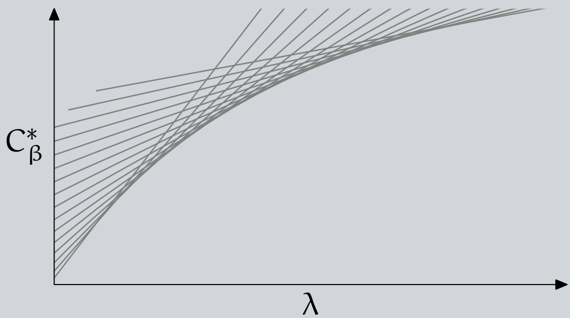
Optimization problems

Costly communication

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

Constrained communication

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



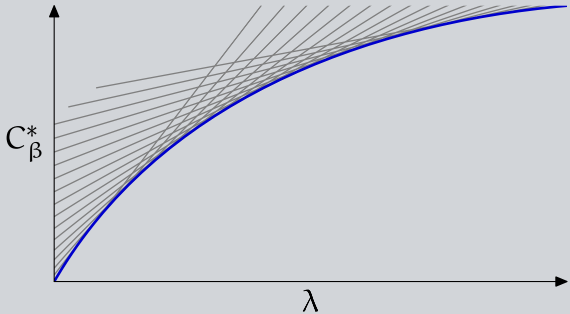
Optimization problems

Costly communication

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

Constrained communication

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



C^*_β is cts, inc, and concave

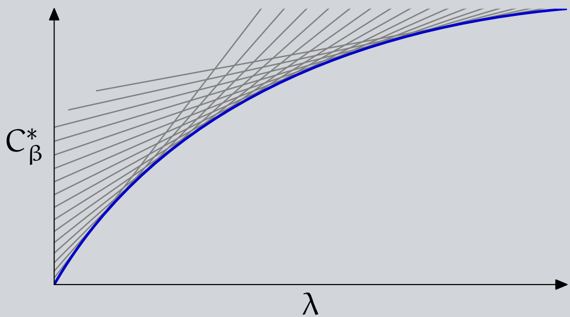
Optimization problems

Costly communication

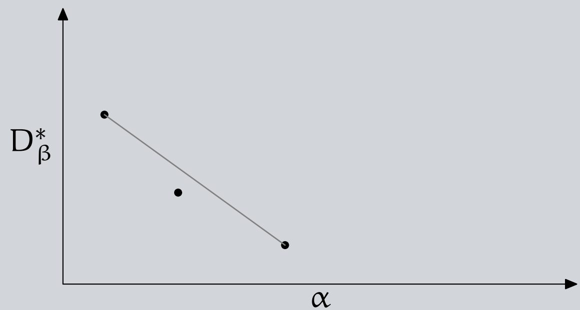
For any $\lambda \in \mathbb{R}_{>0}$, $\quad C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) \coloneqq \inf_{(f,g)} \{D_\beta(f,g) + \lambda N_\beta(f,g)\}$

Constrained communication

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



C^*_β is cts, inc, and concave



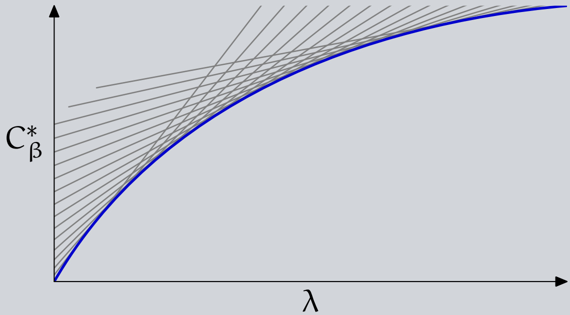
Optimization problems

Costly communication

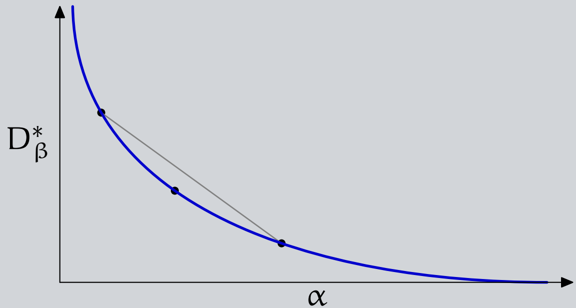
For any $\lambda \in \mathbb{R}_{>0}$, $\quad C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) \coloneqq \inf_{(f,g)} \{D_\beta(f,g) + \lambda N_\beta(f,g)\}$

Constrained communication

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



C^*_β is cts, inc, and concave



D^*_β is cts, dec, and convex

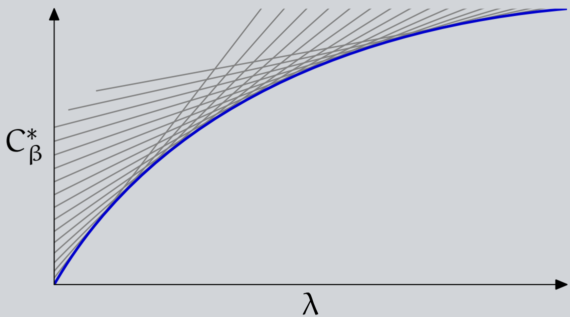
Optimization problems

Costly communication

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

Constrained communication

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



C^*_β is cts, inc, and concave

Distortion-transmission trade-off

D^*_β is cts, dec, and convex

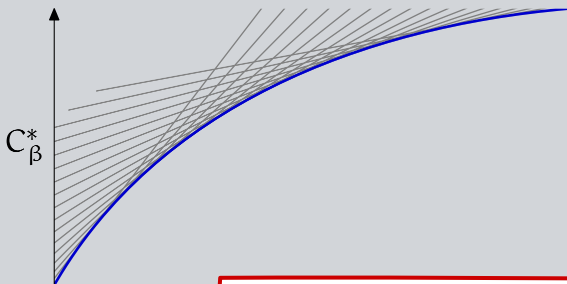
Optimization problems

Costly communication

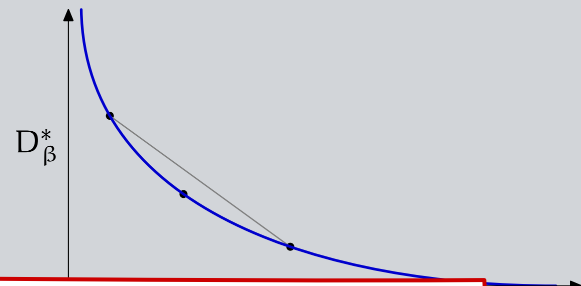
For any $\lambda \in \mathbb{R}_{>0}$, $\quad C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) \coloneqq \inf_{(f,g)} \{D_\beta(f,g) + \lambda N_\beta(f,g)\}$

Constrained communication

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



Distortion-transmission trade-off



C^*_β is c … ex

We provide explicit computable expressions for both curves

Comparison to Information Theory

- Costly communication is analogous to communication under power constraint.
- Distortion-transmission is analogous to distortion-rate trade-off.

Comparison to Information Theory

- Costly communication is analogous to communication under power constraint.
- Distortion-transmission is analogous to distortion-rate trade-off.
- The source reconstruction must be done in **real-time** (or with zero delay).

Comparison to Information Theory

- Costly communication is analogous to communication under power constraint.
- Distortion-transmission is analogous to distortion-rate trade-off.
- The source reconstruction must be done in **real-time** (or with zero delay).

Comparison to real-time communication

- Special case of the real-time communication model
  [Witsenhausen 1979, Walrand-Varaiya 1983, Teneketzis 2006, Teneketzis-Mahajan 2009 . . .].
- Existing results in the literature establish structure of optimal coding strategies and a dynamic program to identify optimal strategies.
- The resultant dynamic programs correspond to decentralized control problem and are hard to solve.

Comparison to Information Theory

- Costly communication is analogous to communication under power constraint.
- Distortion-transmission is analogous to distortion-rate trade-off.
- The source reconstruction must be done in **real-time** (or with zero delay).

Comparison to real-time communication

- Special case of the real-time communication model [Witsenhausen 1979, Walrand-Varaiya 1983, Teneketzis 2006, Teneketzis-Mahajan 2009 . . .].
- Existing results in the literature establish structure of optimal coding strategies and a dynamic program to identify optimal strategies.
- The resultant dynamic programs correspond to decentralized control problem and are hard to solve.

Our approach

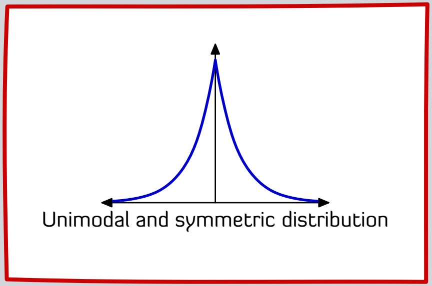
- Previous results have established the structure of optimal strategies.
- Exploit the structural results to explicitly identify optimal strategies.

Modeling assumptions

	Markov chain setup	Guass-Markov setup
State spaces	$X_t, W_t \in \mathbb{Z}$	$X_t, W_t \in \mathbb{R}$

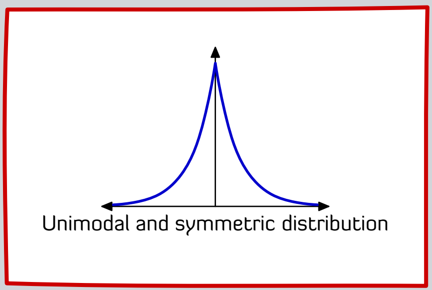
Modeling assumptions

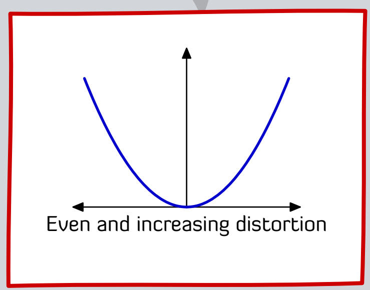
	Markov chain setup	Guass-Markov setup
State spaces	$X_t, W_t \in \mathbb{Z}$	$X_t, W_t \in \mathbb{R}$
Noise distribution	Unimodal and symmetric $p_e = p_{-e} \geqslant p_{e+1}$	Zero-mean Gaussian $\varphi_\sigma(\cdot)$



Modeling assumptions

	Markov chain setup	Guass-Markov setup
State spaces	$X_t, W_t \in \mathbb{Z}$	$X_t, W_t \in \mathbb{R}$
Noise distribution	Unimodal and symmetric $p_e = p_{-e} \geqslant p_{e+1}$	Zero-mean Gaussian $\varphi_\sigma(\cdot)$
Distortion	Even and increasing $d(e) = d(-e) \geqslant d(e+1)$	Mean-squared $d(e) = \lvert e \rvert^2$





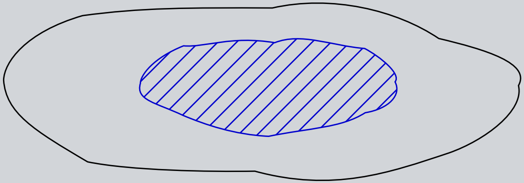
Step 1 Structure of optimal strategies

Step 2 Performance of arbitrary threshold strategies $f^{(k)}$

Step 3 Values of λ for which $f^{(k)}$ is optimal

Step 4 Distortion-transmission trade-off

Step 1 Structure of optimal strategies



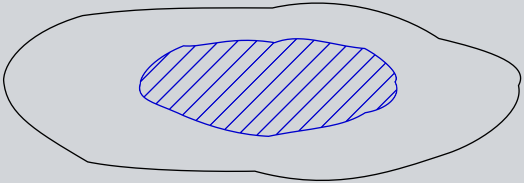
Search space of strategies (f, g)

Step 2 Performance of arbitrary threshold strategies $f^{(k)}$

Step 3 Values of λ for which $f^{(k)}$ is optimal

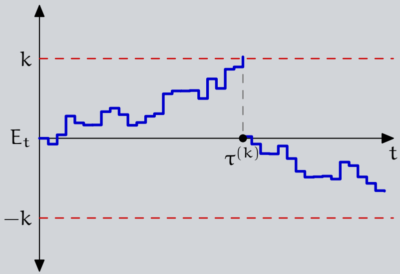
Step 4 Distortion-transmission trade-off

Step 1 Structure of optimal strategies



Search space of strategies (f, g)

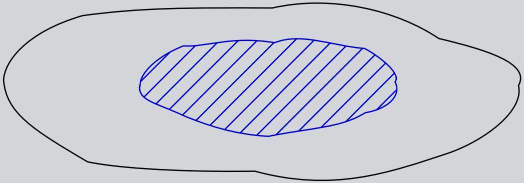
Step 2 Performance of arbitrary threshold strategies $f^{(k)}$



Step 3 Values of λ for which $f^{(k)}$ is optimal

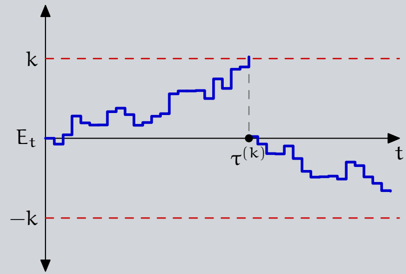
Step 4 Distortion-transmission trade-off

Step 1 Structure of optimal strategies

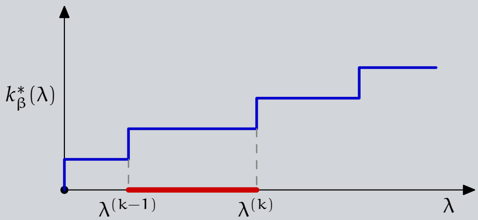


Search space of strategies (f, g)

Step 2 Performance of arbitrary threshold strategies $f^{(k)}$

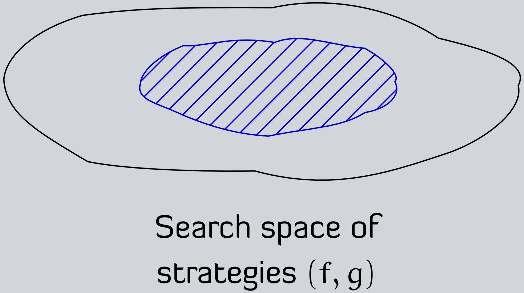


Step 3 Values of λ for which $f^{(k)}$ is optimal

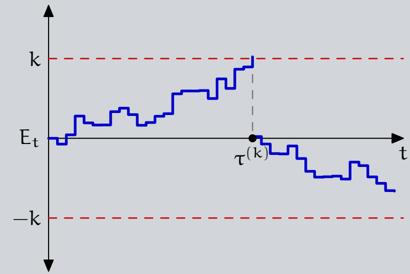


Step 4 Distortion-transmission trade-off

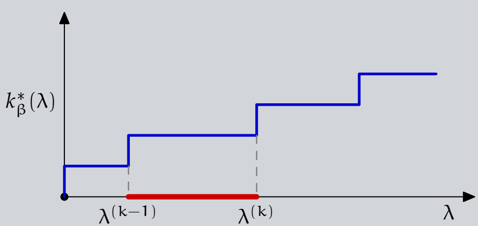
Step 1 Structure of optimal strategies



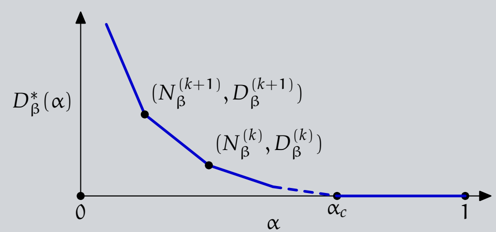
Step 2 Performance of arbitrary threshold strategies $f^{(k)}$



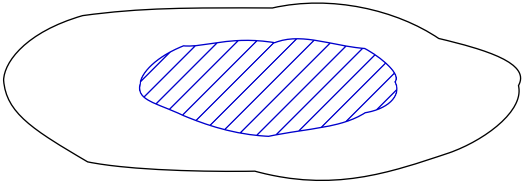
Step 3 Values of λ for which $f^{(k)}$ is optimal



Step 4 Distortion-transmission trade-off

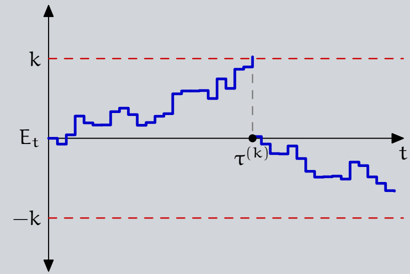


Step 1 Structure of optimal strategies

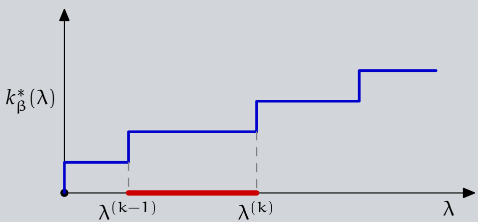


Search space of strategies (f, g)

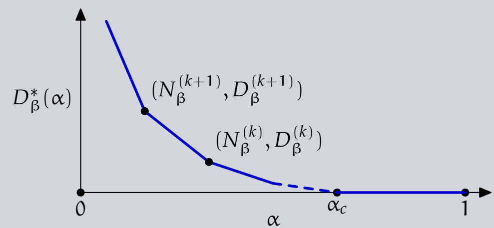
Step 2 Performance of arbitrary threshold strategies $f^{(k)}$



Step 3 Values of λ for which $f^{(k)}$ is optimal



Step 4 Distortion-transmission trade-off



Step 1 Structure of optimal strategies

Model the communication system as decentralized stochastic control

- Two decision makers: transmitter and receiver. Non-nested information.
- Common-information approach [Nayyar-Mahajan-Teneketzis 2013]

  Equivalent centralized problem from the point of view of a coordinator.

  Choose code functions at each step (rather than actions).

Step 1 Structure of optimal strategies

Model the communication system as decentralized stochastic control

- Two decision makers: transmitter and receiver. Non-nested information.
- **Common-information approach** [Nayyar-Mahajan-Teneketzis 2013]

  Equivalent centralized problem from the point of view of a coordinator.

  Choose code functions at each step (rather than actions).

Previous results

- **Guass-Markov setup** [Lipsa-Martins 2009 and 2011, Molin-Hirche 2009]
- **Markov-chain setup** [Nayyar-Başar-Teneketzis-Veeravalli 2013]

Step 1 Structure of optimal strategies

Model the communication system as decentralized stochastic control

- Two decision makers: transmitter and receiver. Non-nested information.
- **Common-information approach** [Nayyar-Mahajan-Teneketzis 2013]

  Equivalent centralized problem from the point of view of a coordinator.

  Choose code functions at each step (rather than actions).

Previous results

- **Guass-Markov setup** [Lipsa-Martins 2009 and 2011, Molin-Hirche 2009]
- **Markov-chain setup** [Nayyar-Başar-Teneketzis-Veeravalli 2013]

  Proof idea: Majorization-based partial order on belief states.

  Prove that $\pi \succeq_m \varphi \implies V(\pi) \geqslant V(\varphi)$.

Step 1 Structure of optimal estimator (Nayyar et al, 2013)

Transmitted Process Let Z_t denote the most recently transmitted value of the Markov source.

$$Z_0 = 0 \quad \text{and} \quad Z_t = \begin{cases} X_t & \text{if } U_t = 1; \\ Z_{t-1} & \text{if } U_t = 0. \end{cases}$$

The estimator can keep track of Z_t as follows:

$$Z_0 = 0 \quad \text{and} \quad Z_t = \begin{cases} Y_t & \text{if } Y_t \neq \varepsilon; \\ Z_{t-1} & \text{if } Y_t = \varepsilon. \end{cases}$$

Step 1 Structure of optimal estimator (Nayyar et al, 2013)

Transmitted Process Let Z_t denote the most recently transmitted value of the Markov source.

$$Z_0 = 0 \quad \text{and} \quad Z_t = \begin{cases} X_t & \text{if } U_t = 1; \\ Z_{t-1} & \text{if } U_t = 0. \end{cases}$$

The estimator can keep track of Z_t as follows:

$$Z_0 = 0 \quad \text{and} \quad Z_t = \begin{cases} Y_t & \text{if } Y_t \neq \varepsilon; \\ Z_{t-1} & \text{if } Y_t = \varepsilon. \end{cases}$$

Theorem 1 The process $\{Z_t\}_{t=0}^{\infty}$ is a sufficient statistic at the estimator and an optimal estimation strategy is given by

$$\hat{X}_t = g_t^*(Z_t) = Z_t \qquad (\star)$$

Remark

- The optimal estimation strategy is time-homogeneous and can be specified in closed form.

Step 1 Structure of optimal transmitter (Nayyar et al)

Error process Let $E_t = X_t - Z_{t-1}$ denote the error process. $\{E_t\}_{t=0}^{\infty}$ is a controlled Markov process where

$$E_0 = 0 \quad \text{and} \quad \mathbb{P}(E_{t+1} = n \mid E_t = e, U_t = u) = \begin{cases} p_{|e-n|}, & \text{if } u = 0; \\ p_n, & \text{if } u = 1. \end{cases}$$

Step 1 Structure of optimal transmitter (Nayyar et al)

Error process Let $E_t = X_t - Z_{t-1}$ denote the error process. $\{E_t\}_{t=0}^{\infty}$ is a controlled Markov process where

$$E_0 = 0 \quad \text{and} \quad \mathbb{P}(E_{t+1} = n \mid E_t = e, U_t = u) = \begin{cases} p_{|e-n|}, & \text{if } u = 0; \\ p_n, & \text{if } u = 1. \end{cases}$$

Theorem 2 When the estimation strategy is of the form ($\star$), then $\{E_t\}_{t=0}^{\infty}$ is a sufficient statistic at the transmitter.

Furthermore, an optimal transmission strategy is characterized by a time-varying threshold $\{k_t\}_{t=0}^{\infty}$, i.e.,

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

Step 1 Main idea

Restrict attention to time-homogeneous estimation strategies of the form

$$\hat{X}_t = g^*_t(Z_t) = Z_t.$$

Consider the problem of finding the best-response transmission strategy.

Under appropriate technical conditions, the best-response strategy is time-homogeneous.

Step 1 Main idea

Restrict attention to time-homogeneous estimation strategies of the form

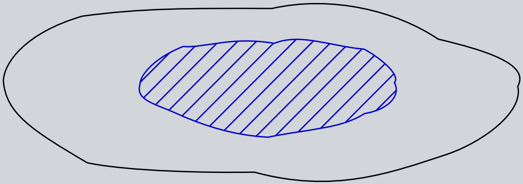
$$\hat{X}_t = g_t^*(Z_t) = Z_t.$$

Consider the problem of finding the best-response transmission strategy.

Under appropriate technical conditions, the best-response strategy is time-homogeneous.

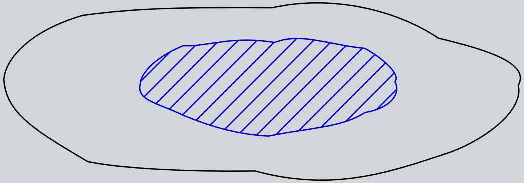
Find the best treshold-based strategy within the class $\mathcal{F} = \{f^{(k)} : k \in \mathbb{Z}_{\geqslant 0}\}$ where

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



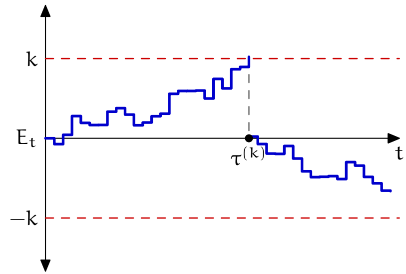
Search space of strategies (f, g)

Step 1 Structure of optimal strategies

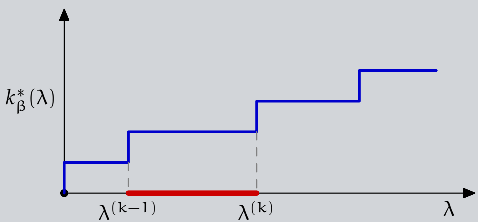


Search space of strategies (f, g)

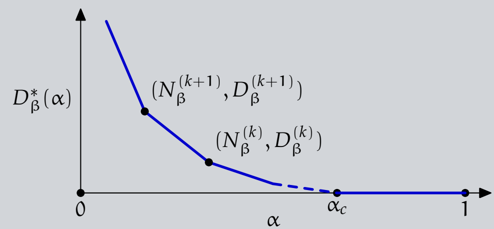
Step 2 Performance of arbitrary threshold strategies $f^{(k)}$



Step 3 Values of λ for which $f^{(k)}$ is optimal



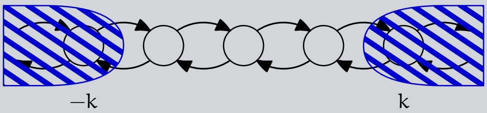
Step 4 Distortion-transmission trade-off



Step 2 Performance of threshold strategies

Consider a threshold-based strategy

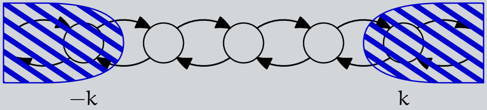
$$f^{(k)}(e) = \begin{cases} 1 & \text{if } |e| \geqslant k \\ 0 & \text{otherwise} \end{cases}$$



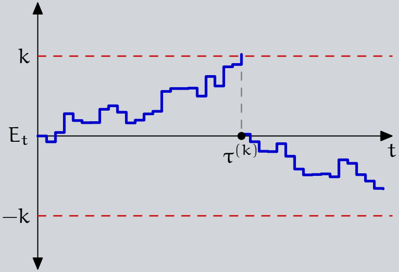
Step 2 Performance of threshold strategies

Consider a threshold-based strategy

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



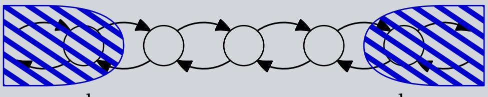
Let $\tau^{(k)}$ denote the stopping time of first transmission (starting at $E_0 = 0$).



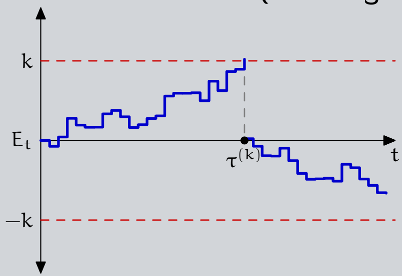
Step 2 Performance of threshold strategies

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 transmission (starting at $E_0 = 0$).



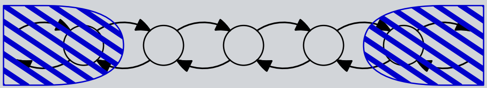
Define

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

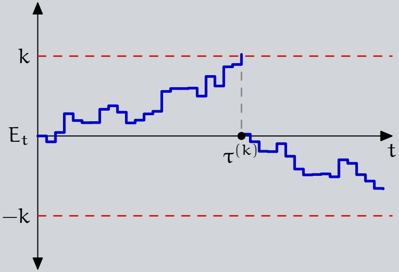
Step 2 Performance of threshold strategies

Consider a threshold-based strategy

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



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



Define

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

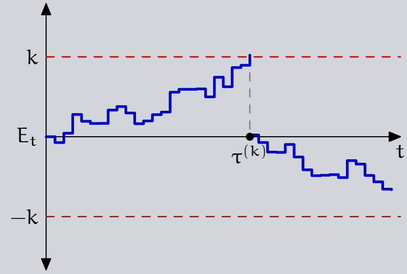
Proposition $\{E_t\}_{t=0}^{\infty}$ is a regenerative process and by renewal theory, we have that

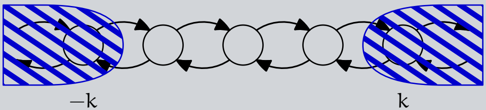
$$D_\beta^{(k)} \coloneqq D_\beta(f^{(k)}, g^*) = \frac{L_\beta^{(k)}}{M_\beta^{(k)}} \quad \text{and} \quad N_\beta^{(k)} \coloneqq N_\beta(f^{(k)}, g^*) = \frac{1}{M_\beta^{(k)}} - (1-\beta).$$

Step 2 Computing $D_\beta^{(k)}$ and $N_\beta^{(k)}$

Notation

- $\mathcal{S}^{(k)} = \{-(k-1), \dots, k-1\}$.
- $[P^{(k)}]_{ij} = p_{|i-j|}$, for $i, j \in \mathcal{S}^{(k)}$.
- $[d^{(k)}]_i = d(i)$, for $i \in \mathcal{S}^{(k)}$.
- $[\mathbf{1}^{(k)}]_i = 1$, for $i \in \mathcal{S}^{(k)}$.

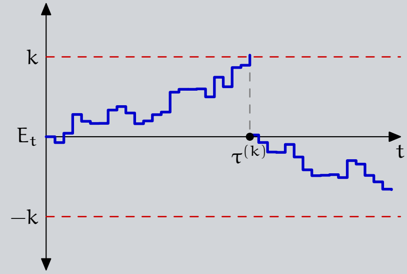




Step 2 Computing $D_\beta^{(k)}$ and $N_\beta^{(k)}$

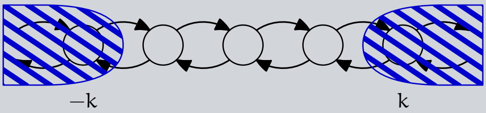
Notation

- $\mathcal{S}^{(k)} = \{-(k-1), \dots, k-1\}$.
- $[P^{(k)}]_{ij} = p_{|i-j|}$, for $i, j \in \mathcal{S}^{(k)}$.
- $[d^{(k)}]_i = d(i)$, for $i \in \mathcal{S}^{(k)}$.
- $[\mathbf{1}^{(k)}]_i = 1$, for $i \in \mathcal{S}^{(k)}$.

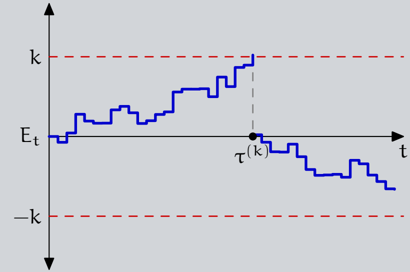


Proposition

- $L_\beta^{(k)} = \big[[I - \beta P^{(k)}]^{-1} d^{(k)}\big]_0$.
- $M_\beta^{(k)} = \big[[I - \beta P^{(k)}]^{-1} \mathbf{1}^{(k)}\big]_0$.



Step 2 Computing $D_\beta^{(k)}$ and $N_\beta^{(k)}$

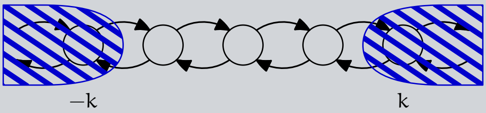


Notation

- $\mathcal{S}^{(k)} = \{-(k-1), \dots, k-1\}$.
- $[P^{(k)}]_{ij} = p_{|i-j|}$, for $i, j \in \mathcal{S}^{(k)}$.
- $[d^{(k)}]_i = d(i)$, for $i \in \mathcal{S}^{(k)}$.
- $[\mathbf{1}^{(k)}]_i = 1$, for $i \in \mathcal{S}^{(k)}$.

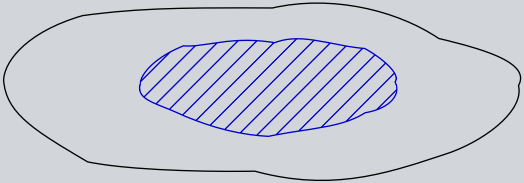
Proposition

- $L_\beta^{(k)} = \big[[I - \beta P^{(k)}]^{-1} d^{(k)}\big]_0$.
- $M_\beta^{(k)} = \big[[I - \beta P^{(k)}]^{-1} \mathbf{1}^{(k)}\big]_0$.



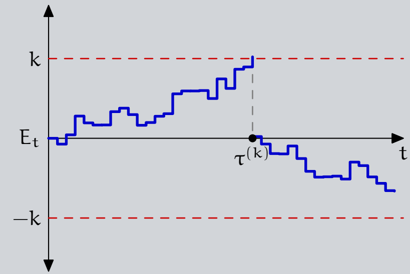
$D_\beta^{(k)}$ and $N_\beta^{(k)}$ can be computed using these expressions.

Step 1 Structure of optimal strategies

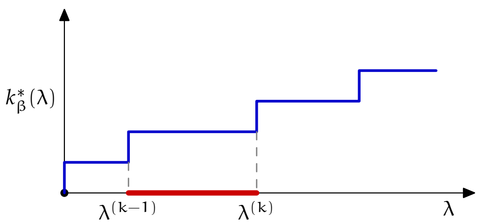


Search space of strategies (f, g)

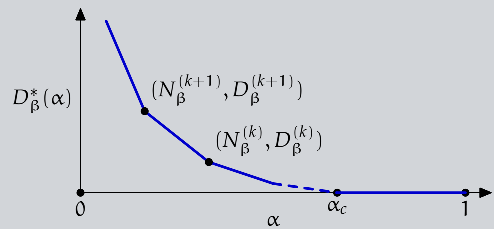
Step 2 Performance of arbitrary threshold strategies $f^{(k)}$



Step 3 Values of λ for which $f^{(k)}$ is optimal



Step 4 Distortion-transmission trade-off



Step 3 Properties of optimal thresholds

Monotonicity $L_\beta^{(k+1)} > L_\beta^{(k)}$ and $M_\beta^{(k+1)} > M_\beta^{(k)}$ Depends on unimodularity of noise

Step 3 Properties of optimal thresholds

Monotonicity $L_\beta^{(k+1)} > L_\beta^{(k)}$ and $M_\beta^{(k+1)} > M_\beta^{(k)}$

Implication:

$D_\beta^{(k+1)} \geqslant D_\beta^{(k)}$ and $N_\beta^{(k+1)} < N_\beta^{(k)}$

Use DP and monotonicity of Bellman operator

Step 3 Properties of optimal thresholds

Monotonicity $L_\beta^{(k+1)} > L_\beta^{(k)}$ and $M_\beta^{(k+1)} > M_\beta^{(k)}$

Implication:

$$D_\beta^{(k+1)} \geqslant D_\beta^{(k)} \quad \text{and} \quad N_\beta^{(k+1)} < N_\beta^{(k)}$$

Submodularity $C_\beta^{(k)}(\lambda) \coloneqq D_\beta^{(k)} + \lambda N_\beta^{(k)}$ is submodular in (k, λ).

Step 3 Properties of optimal thresholds

Monotonicity $L_\beta^{(k+1)} > L_\beta^{(k)}$ and $M_\beta^{(k+1)} > M_\beta^{(k)}$

Implication:

$$D_\beta^{(k+1)} \geqslant D_\beta^{(k)} \quad \text{and} \quad N_\beta^{(k+1)} < N_\beta^{(k)}$$

Submodularity $C_\beta^{(k)}(\lambda) \coloneqq D_\beta^{(k)} + \lambda N_\beta^{(k)}$ is submodular in (k, λ).

Proof: $C_\beta^{(k+1)}(\lambda) - C_\beta^{(k)}(\lambda) = D_\beta^{(k+1)}(\lambda) - D_\beta^{(k)}(\lambda) - \lambda\underbrace{(N_\beta^{(k)}(\lambda) - N_\beta^{(k+1)}(\lambda))}_{\geqslant 0}$.

Step 3 Properties of optimal thresholds

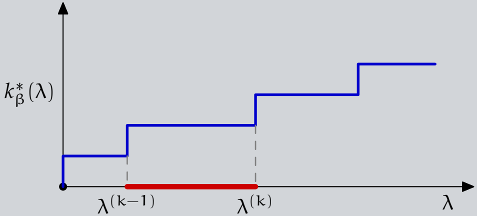
Monotonicity $L_\beta^{(k+1)} > L_\beta^{(k)}$ and $M_\beta^{(k+1)} > M_\beta^{(k)}$

Implication:

$$D_\beta^{(k+1)} \geqslant D_\beta^{(k)} \quad \text{and} \quad N_\beta^{(k+1)} < N_\beta^{(k)}$$

Submodularity $C_\beta^{(k)}(\lambda) \coloneqq D_\beta^{(k)} + \lambda N_\beta^{(k)}$ is submodular in (k, λ).

Proposition $k_\beta^*(\lambda) \coloneqq \arg\min_{k \in \mathbb{Z}_{\geqslant 0}} C_\beta^{(k)}(\lambda)$ is increasing in λ.



Step 3 Properties of optimal thresholds

Monotonicity $L_\beta^{(k+1)} > L_\beta^{(k)}$ and $M_\beta^{(k+1)} > M_\beta^{(k)}$

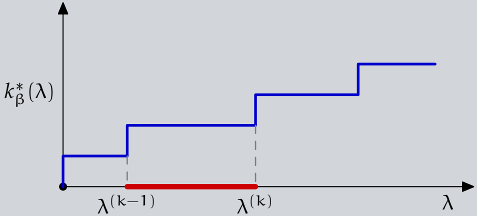
Implication:

$D_\beta^{(k+1)} \geqslant D_\beta^{(k)}$ and $N_\beta^{(k+1)} < N_\beta^{(k)}$

Submodularity $C_\beta^{(k)}(\lambda) \coloneqq D_\beta^{(k)} + \lambda N_\beta^{(k)}$ is submodular in (k, λ).

Proposition $k_\beta^*(\lambda) \coloneqq \arg\min_{k \in \mathbb{Z}_{\geqslant 0}} C_\beta^{(k)}(\lambda)$ is increasing in λ.

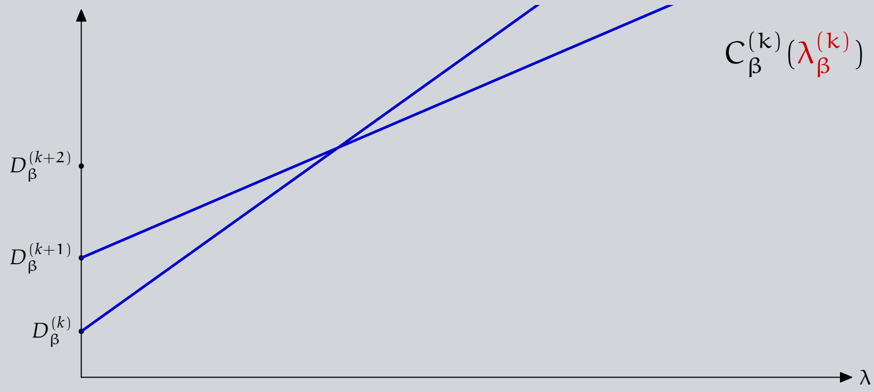
Define $\Lambda_\beta^{(k)} \coloneqq \{\lambda \in \mathbb{R}_{\geqslant 0} : k_\beta^*(\lambda) = k\} = [\lambda_\beta^{(k-1)}, \lambda_\beta^{(k)}]$.



Step 3 Optimal costly communication

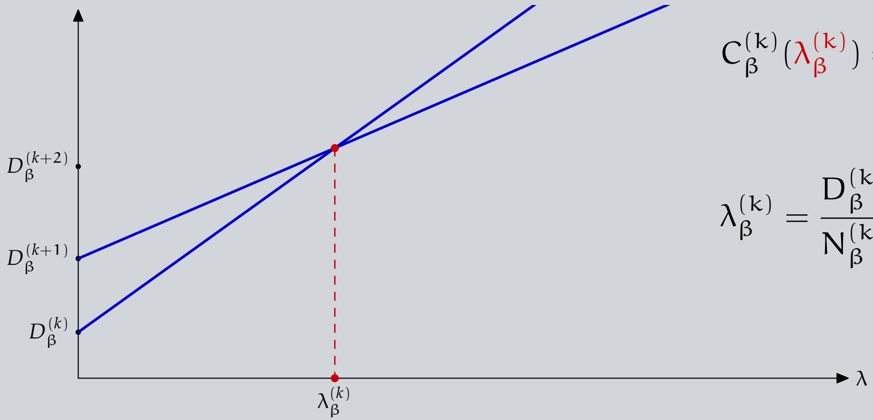
$$C_\beta^{(k)}(\lambda_\beta^{(k)}) = C_\beta^{(k+1)}(\lambda_\beta^{(k)})$$

Step 3 Optimal costly communication



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

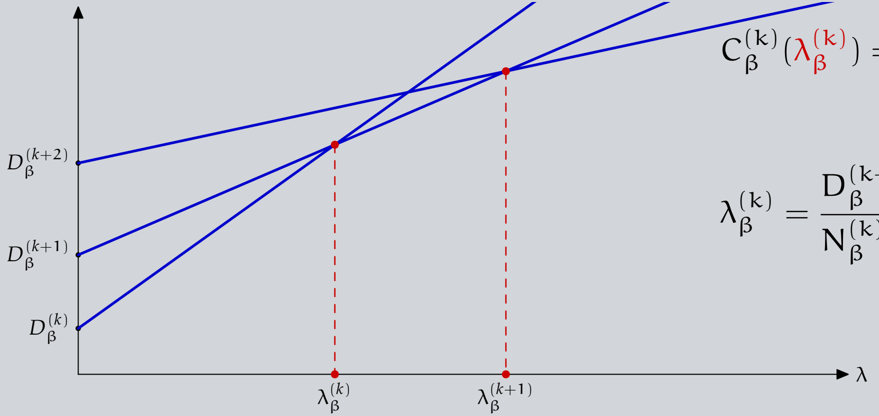
Step 3 Optimal costly communication



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

$$\lambda_\beta^{(k)} = \frac{D_\beta^{(k+1)} - D_\beta^{(k)}}{N_\beta^{(k)} - N_\beta^{(k+1)}}$$

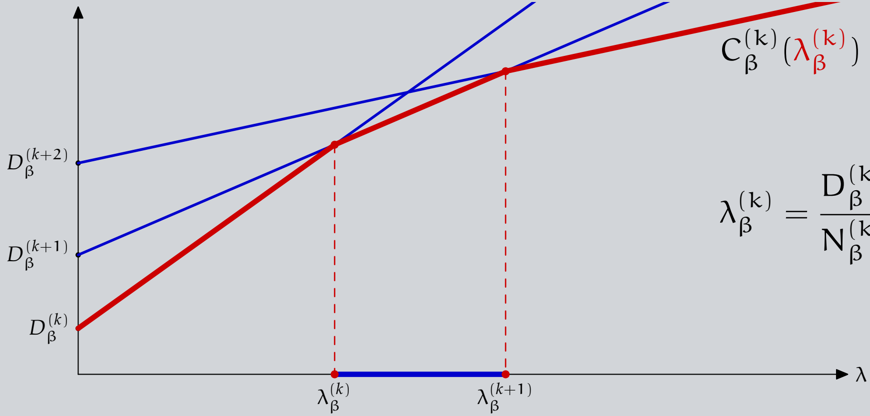
Step 3 Optimal costly communication



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

$$\lambda_\beta^{(k)} = \frac{D_\beta^{(k+1)} - D_\beta^{(k)}}{N_\beta^{(k)} - N_\beta^{(k+1)}}$$

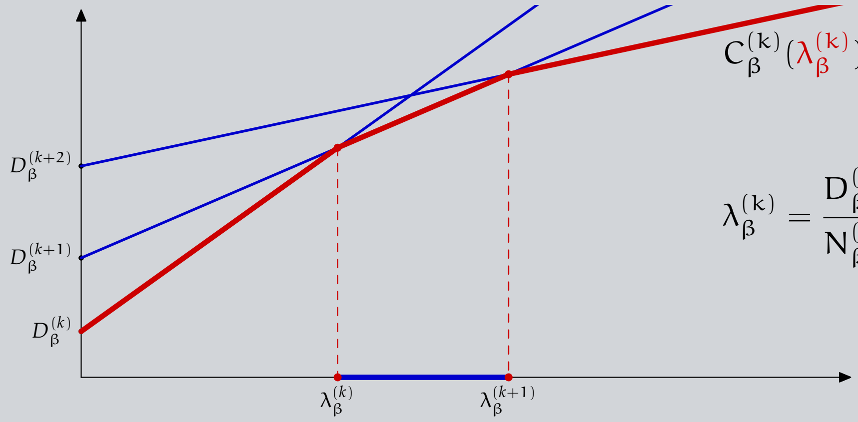
Step 3 Optimal costly communication



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

$$\lambda_\beta^{(k)} = \frac{D_\beta^{(k+1)} - D_\beta^{(k)}}{N_\beta^{(k)} - N_\beta^{(k+1)}}$$

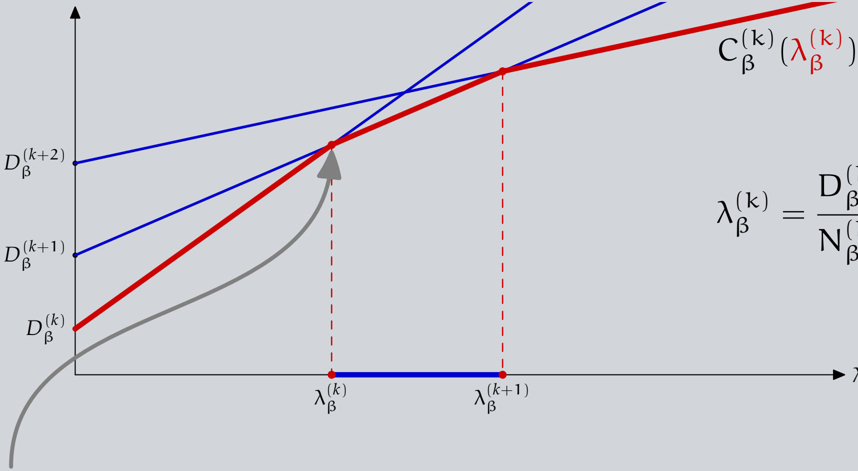
Step 3 Optimal costly communication



Theorem

- For all $\lambda \in (\lambda_\beta^{(k)}, \lambda_\beta^{(k+1)}]$ the threshold strategy $f^{(k+1)}$ is optimal.
- $C_\beta^*(\lambda) = \min_{k \in \mathbb{Z}_{\geqslant 0}} C_\beta^{(k)}$ is piecewise linear, continuous, concave, and increasing function of λ.

Step 3 Optimal costly communication



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

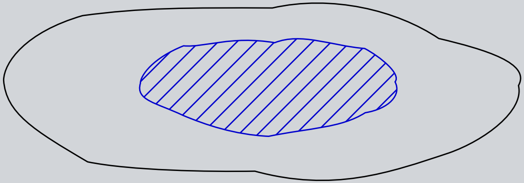
$$\lambda_\beta^{(k)} = \frac{D_\beta^{(k+1)} - D_\beta^{(k)}}{N_\beta^{(k)} - N_\beta^{(k+1)}}$$

$(\lambda_\beta^{(k)}, D_\beta^{(k)} + \lambda_\beta^{(k)} N_\beta^{(k)})$

Theorem

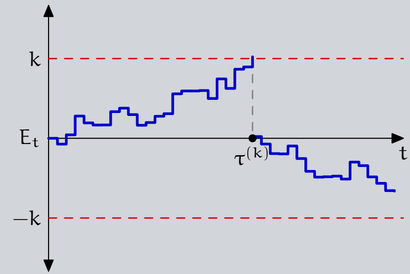
- For all $\lambda \in (\lambda_\beta^{(k)}, \lambda_\beta^{(k+1)}]$ the threshold strategy $f^{(k+1)}$ is optimal.
- $C_\beta^*(\lambda) = \min_{k \in \mathbb{Z}_{\geqslant 0}} C_\beta^{(k)}$ is piecewise linear, continuous, concave, and increasing function of λ.

Step 1 Structure of optimal strategies

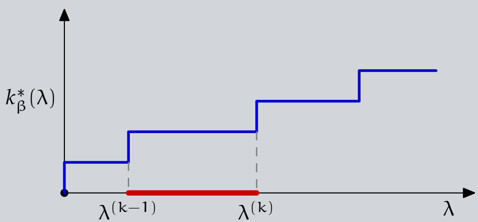


Search space of strategies (f, g)

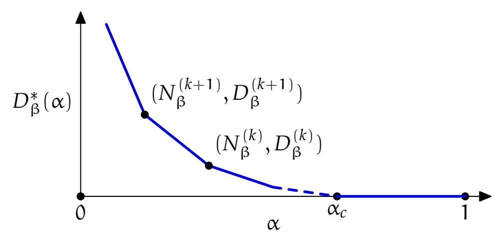
Step 2 Performance of arbitrary threshold strategies $f^{(k)}$



Step 3 Values of λ for which $f^{(k)}$ is optimal



Step 4 Distortion-transmission trade-off



Step 4 Distortion-transmission trade-off

Sufficient conditions for constrained optimality

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

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

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for $C_\beta(f, g; \lambda^\circ)$.

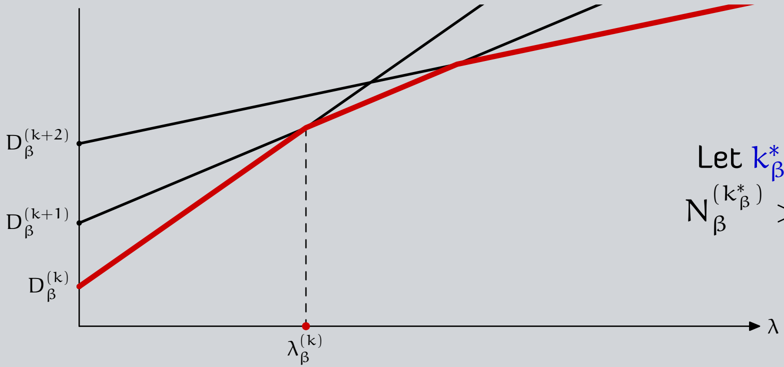
Step 4 Distortion-transmission trade-off

Sufficient conditions for constrained optimality

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

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

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for $C_\beta(f, g; \lambda^\circ)$.



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

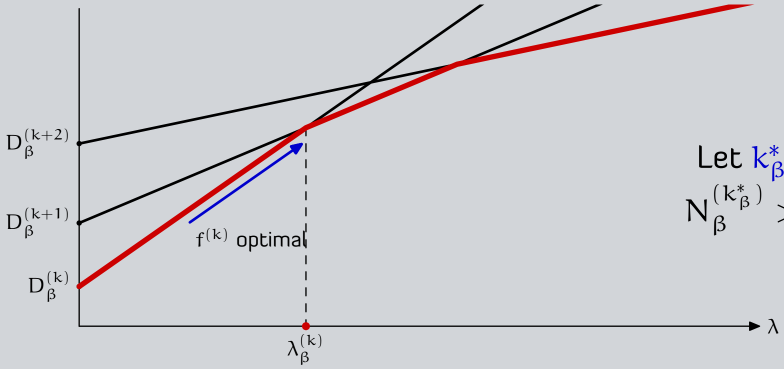
Step 4 Distortion-transmission trade-off

Sufficient conditions for constrained optimality

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

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

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for $C_\beta(f, g; \lambda^\circ)$.



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

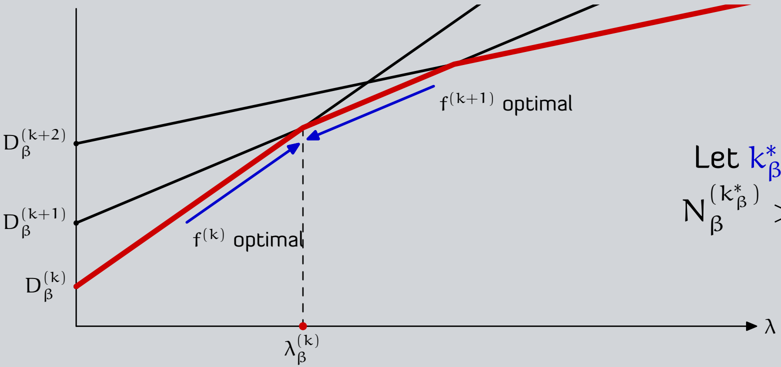
Step 4 Distortion-transmission trade-off

Sufficient conditions for constrained optimality

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

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

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for $C_\beta(f, g; \lambda^\circ)$.



Let k^*_β be such that $N_\beta^{(k^*_\beta)} > \alpha > N_\beta^{(k^*_\beta+1)}$

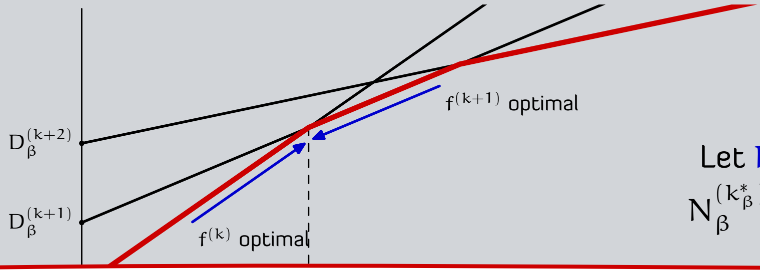
Step 4 Distortion-transmission trade-off

Sufficient conditions for constrained optimality

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

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

(C2) There exists $\lambda^\circ \geqslant 0$ such that (f°, g°) is optimal for $C_\beta(f, g; \lambda^\circ)$.



Let k^*_β be such that
$N_\beta^{(k^*_\beta)} > \alpha > N_\beta^{(k^*_\beta+1)}$

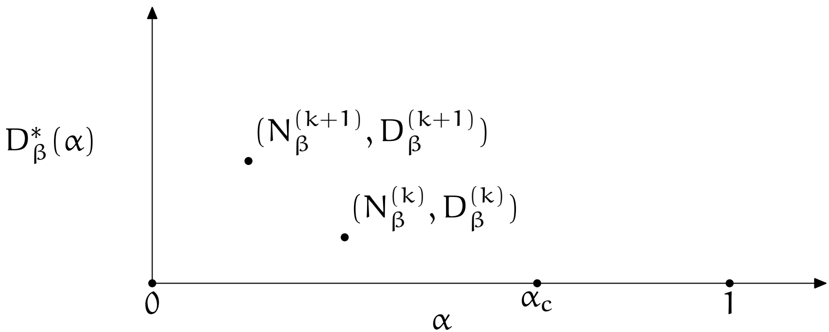
Bernoulli randomized strategy $(\theta^*, f^{(k)}, f^{k+1})$ is optimal where

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

Step 4 Distortion-transmission trade-off

Sufficient conditions for constrained optimality

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



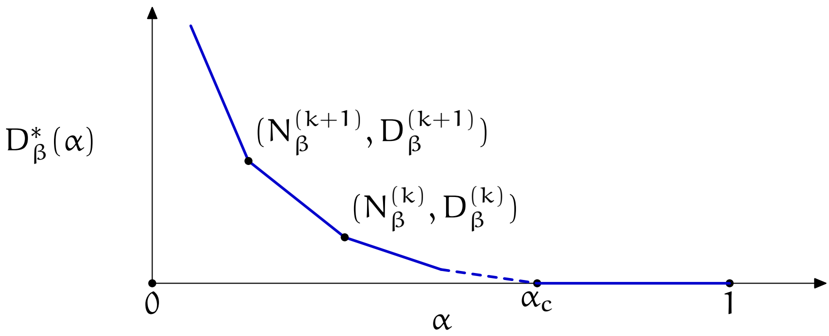
Bernoulli randomized strategy $(\theta^*, f^{(k)}, f^{k+1})$ is optimal where

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

Step 4 Distortion-transmission trade-off

Sufficient conditions for constrained optimality

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



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

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

Step 4 Features of optimal strategy

Optimal strategy

$$f^*(e) = \begin{cases} 1 & \text{if } |e| > k^*_\beta \\ 1 & \text{w.p. } \theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{w.p. } 1-\theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{if } |e| < k^*_\beta \end{cases}$$

Step 4 Features of optimal strategy

Optimal strategy

$$f^*(e) = \begin{cases} 1 & \text{if } |e| > k^*_\beta \\ 1 & \text{w.p. } \theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{w.p. } 1 - \theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{if } |e| < k^*_\beta \end{cases}$$

Randomized action
at a single state

Step 4 Features of optimal strategy

Optimal strategy

$$f^*(e) = \begin{cases} 1 & \text{if } |e| > k^*_\beta \\ 1 & \text{w.p. } \theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{w.p. } 1-\theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{if } |e| < k^*_\beta \end{cases}$$

Randomized action
at a single state

Deterministic implementation

Time-sharing strategies

- Assume $\theta^* = a/(a+b)$.
- Choose strategy $f^{(k^*)}$ for a visits to state zero and strategy $f^{(k^*+1)}$ for b visits to state zero and so on.

Step 4 Features of optimal strategy

Optimal strategy

$$f^*(e) = \begin{cases} 1 & \text{if } |e| > k^*_\beta \\ 1 & \text{w.p. } \theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{w.p. } 1-\theta^* \text{ if } |e| = k^*_\beta \\ 0 & \text{if } |e| < k^*_\beta \end{cases}$$

Randomized action at a single state

Deterministic implementation

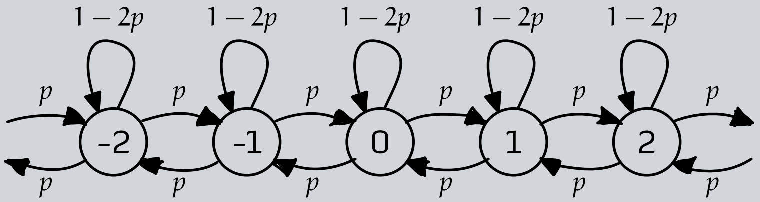
Time-sharing strategies

- Assume $\theta^* = a/(a+b)$.
- Choose strategy $f^{(k^*)}$ for a visits to state zero and strategy $f^{(k^*+1)}$ for b visits to state zero and so on.

Steering strategies:

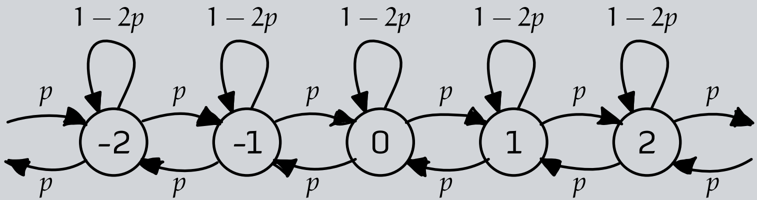
- Let a^i_t denote the number of times action i is chosen in the past.
- At states $\{-k^*, k^*\}$ choose an action that steers the empirical frequency closer to the desired randomization probability.

An example: Symmetric birth-death Markov Chain



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

An example: Symmetric birth-death Markov Chain



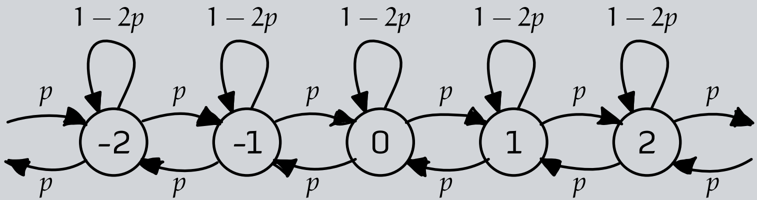
Discounted cost Let $K_\beta = -2 - (1-\beta)/\beta p$ and $m_\beta = \cosh^{-1}(-K_\beta/2)$.

$$D_\beta^{(k)} = \frac{\sinh(k m_\beta) - k \sinh(m_\beta)}{2\sinh^2(k m_\beta/2)\sinh(m_\beta)}$$

$$N_\beta^{(k)} = \frac{2\beta p \sinh^2(m_\beta/2)\cosh(k m_\beta)}{\sinh^2(k m_\beta/2)} - (1-\beta)$$

Average cost $D_1^{(k)} = \dfrac{k^2-1}{3k}$ and $N_1^{(k)} = \dfrac{2p}{k^2}$

An example: Symmetric birth-death Markov Chain



Discounted cost Let $K_\beta = -2 - (1-\beta)/\beta p$ and $m_\beta = \cosh^{-1}(-K_\beta/2)$.

$$D_\beta^{(k)} = \frac{\sinh(k m_\beta) - k \sinh(m_\beta)}{2\sinh^2(k m_\beta/2)\sinh(m_\beta)}$$

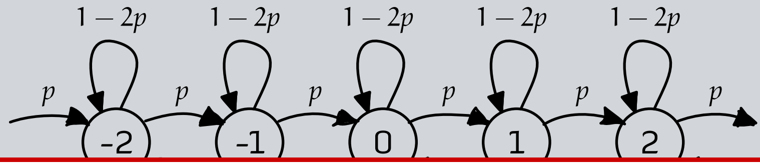
$$N_\beta^{(k)} = \frac{2\beta p \sinh^2(m_\beta/2)\cosh(k m_\beta)}{\sinh^2(k m_\beta/2)} - (1-\beta)$$

$\lambda_\beta^{(k)}$ can be computed in terms of $D_\beta^{(k)}$ and $N_\beta^{(k)}$.

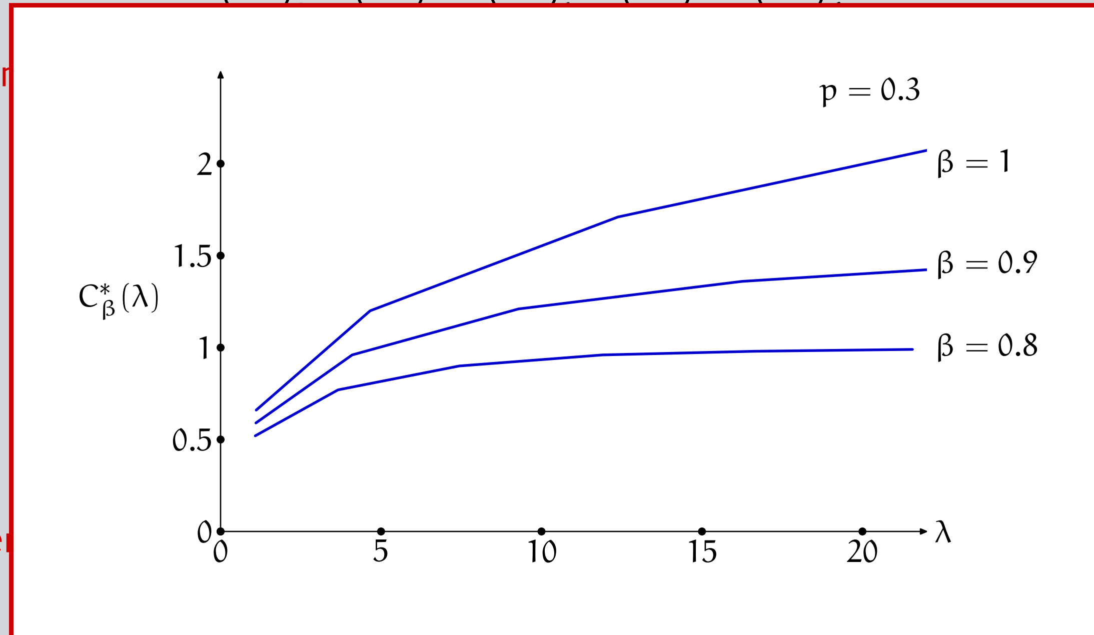
Average cost $D_1^{(k)} = \dfrac{k^2-1}{3k}$ and $N_1^{(k)} = \dfrac{2p}{k^2}$

$$\lambda_1^{(k)} = \frac{k(k+1)(k^2+k+1)}{6p(2k+1)}$$

An example: Symmetric birth-death Markov Chain

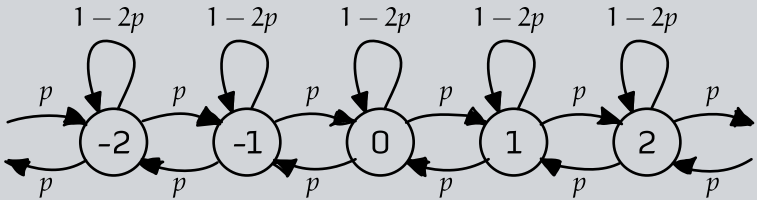


Discour



Aver

An example: Symmetric birth-death Markov Chain



Discounted cost Let $K_\beta = -2 - (1-\beta)/\beta p$ and $m_\beta = \cosh^{-1}(-K_\beta/2)$.

$$D_\beta^{(k)} = \frac{\sinh(k m_\beta) - k \sinh(m_\beta)}{2\sinh^2(k m_\beta/2)\sinh(m_\beta)}$$

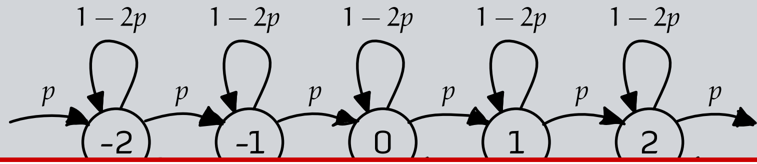
$$N_\beta^{(k)} = \frac{2\beta p \sinh^2(m_\beta/2)\cosh(k m_\beta)}{\sinh^2(k m_\beta/2)} - (1-\beta)$$

$$k_\beta^* = \sup\left\{k \in \mathbb{Z}_{\geqslant 0} : \frac{\sinh^2(m_\beta/2)\cosh(k m_\beta)}{\sinh^2(k m_\beta/2)} \geqslant \frac{1+\alpha-\beta}{2\beta p}\right\}$$

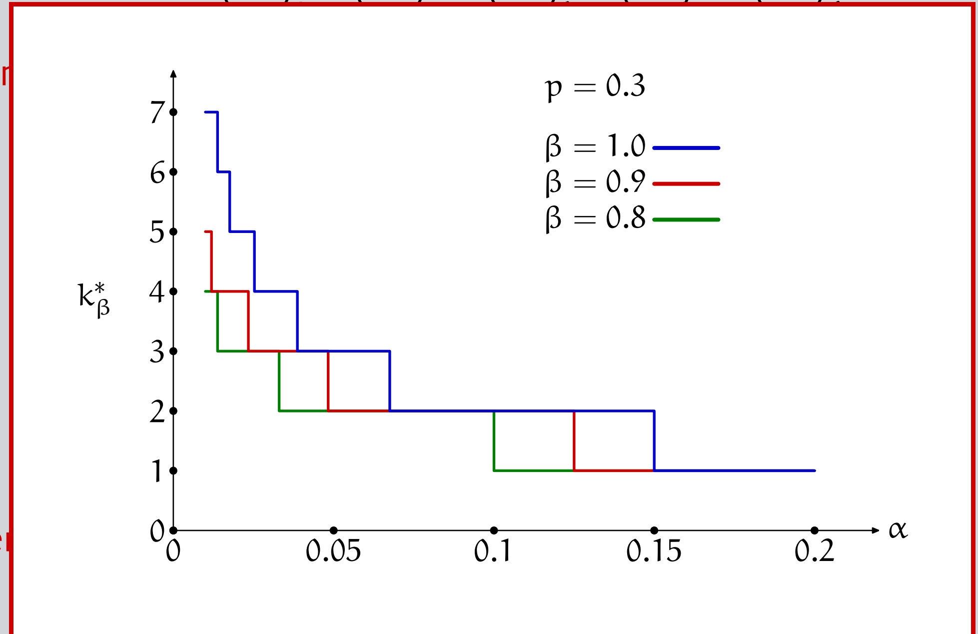
Average cost $D_1^{(k)} = \dfrac{k^2-1}{3k}$ and $N_1^{(k)} = \dfrac{2p}{k^2}$

$$k_1^* = \left\lfloor \sqrt{\frac{2p}{\alpha}} \right\rfloor$$

An example: Symmetric birth-death Markov Chain

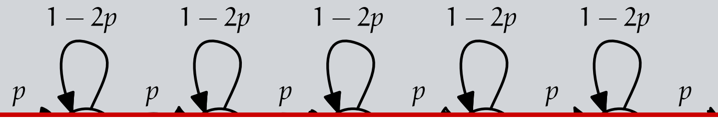


Discour



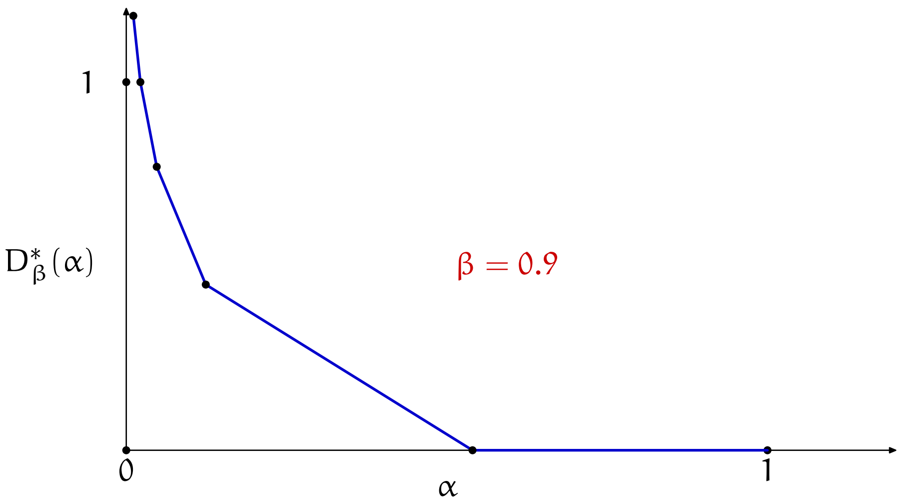
Aver

An example: Symmetric birth-death Markov Chain



Discour

Aver



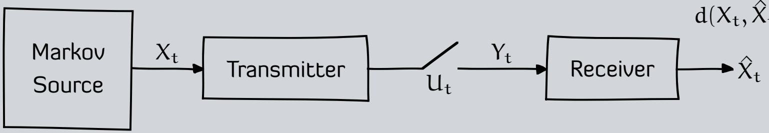
Summary

The system model

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

Distortion

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



$X_{t+1} = X_t + W_t$ $\quad U_t = f_t(X_{1:t}, U_{1:t-1})$ $\quad \hat{X}_t = g_t(Y_{1:t})$

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

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

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

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

Summary

The system model

X_{t-}

1. Disc

$D_\beta($

2. Ave

$D_1($

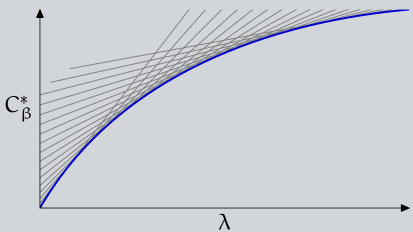
Optimization problems

Costly communication

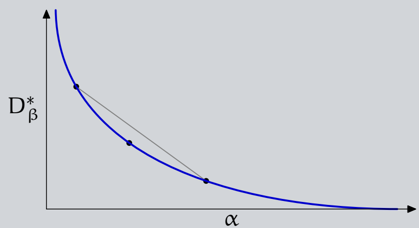
For any $\lambda \in \mathbb{R}_{>0}$, $\quad C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) \coloneqq \inf_{(f,g)} \{ D_\beta(f,g) + \lambda N_\beta(f,g) \}$

Constrained communication

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



C^*_β is cts, inc, and concave



D^*_β is cts, dec, and convex

Summary

The system model

Optimization problems

Cos

Con

X_{t}

1. Disc

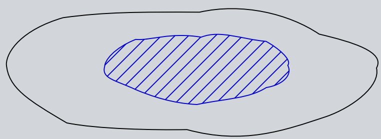
$D_\beta($

2. Ave

C^*_β

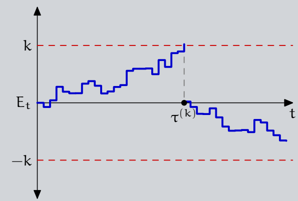
$D_1($

Step 1 Structure of optimal strategies

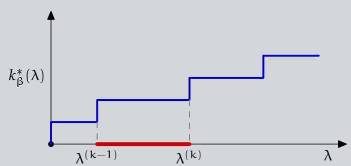


Search space of strategies (f, g)

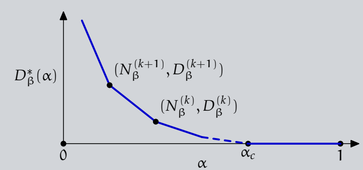
Step 2 Performance of arbitrary threshold strategies $f^{(k)}$



Step 3 Values of λ for which $f^{(k)}$ is optimal



Step 4 Distortion-transmission trade-off



Conclusion

Analyze a fundamental trade-off in real-time communication

Conclusion

Analyze a fundamental trade-off in real-time communication

Possible generalizations (where the proposed approach may work)

- Symmetric finite sources
- Erasure channels

Conclusion

Analyze a fundamental trade-off in real-time communication

Possible generalizations (where the proposed approach may work)

- Symmetric finite sources
- Erasure channels

More realistic models (where the proposed approach will fail)

- Non-symmetric sources
- Rate constraints (affect of quantization)
- Network delays

Conclusion

Analyze a fundamental trade-off in real-time communication

Possible generalizations (where the proposed approach may work)

- Symmetric finite sources
- Erasure channels

More realistic models (where the proposed approach will fail)

- Non-symmetric sources
- Rate constraints (affect of quantization)
- Network delays

Full version to be posted on arxiv soon.