Distortion transmission function for transmitting Markov processes under communication constraints

Aditya Mahajan
McGill University

Joint work with Jhelum Chakravorty

Information Theory and Application Workshop
2 Feb, 2016

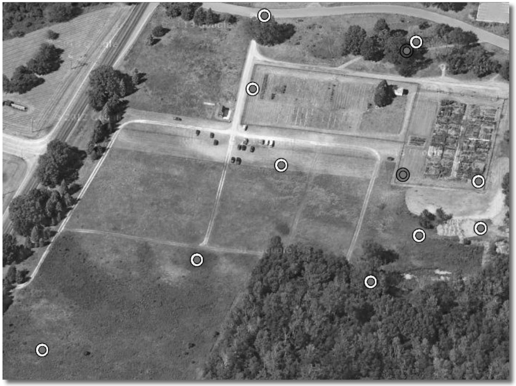
What is an analogue of rate-distortion function **(or** distortion-rate function**) in networks?**

Motivation

Many applications require:

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

Motivation

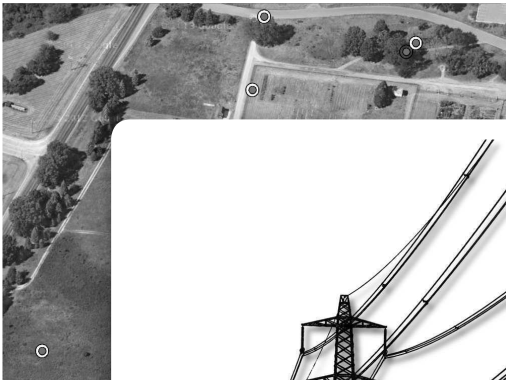


Sensor Networks

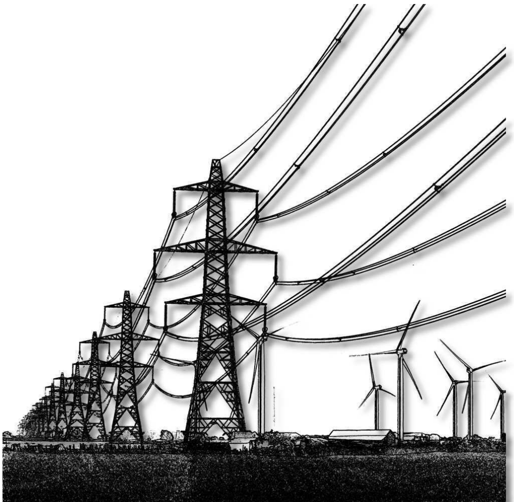
Many applications require:

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

Motivation



S

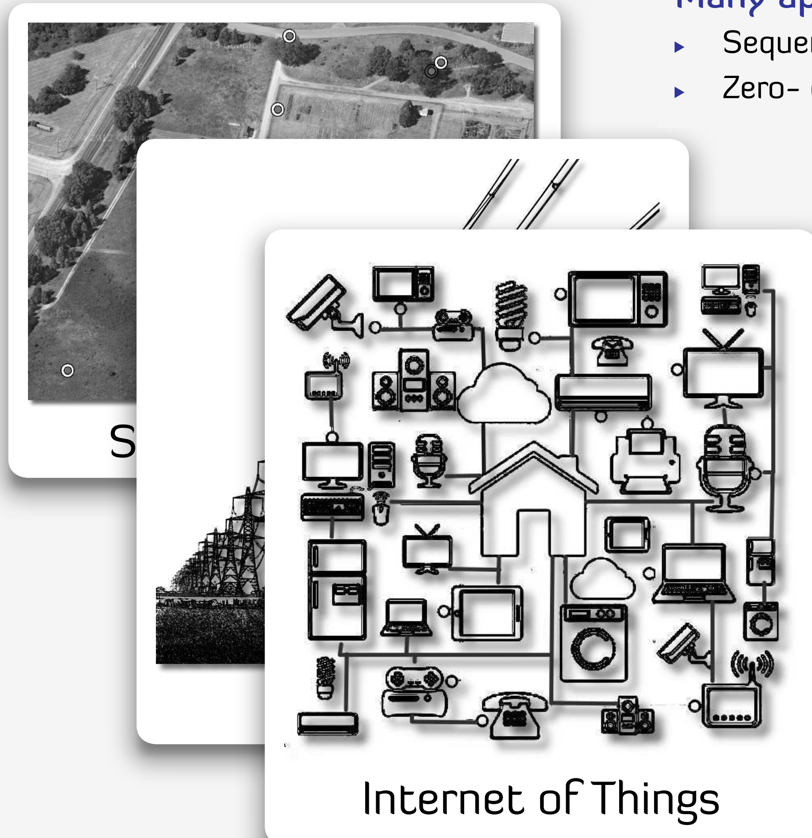


Smart Grids

Many applications require:

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

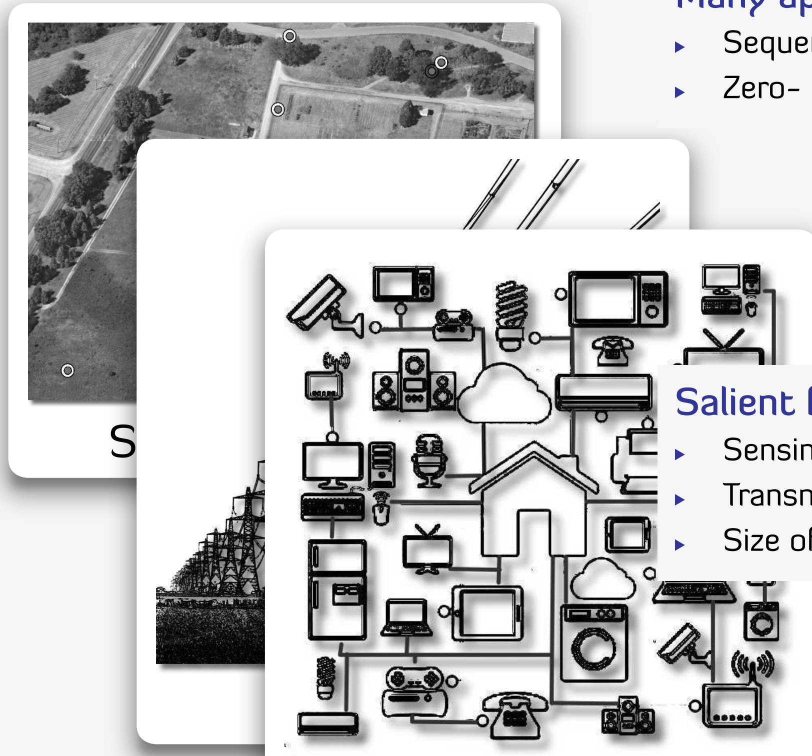
Motivation



Many applications require:

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

Motivation



Internet of Things

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

Motivation

Many applications require:

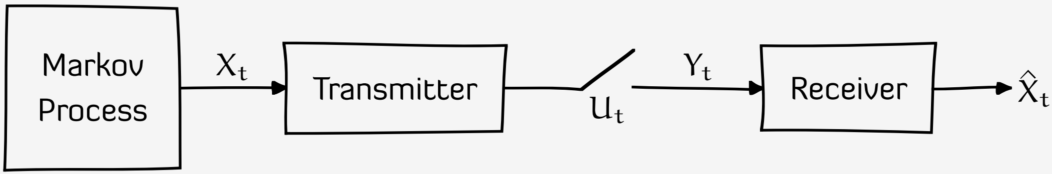
- Sequential transmission of data
- Zero- (or finite-) delay reconstruction

Salient features

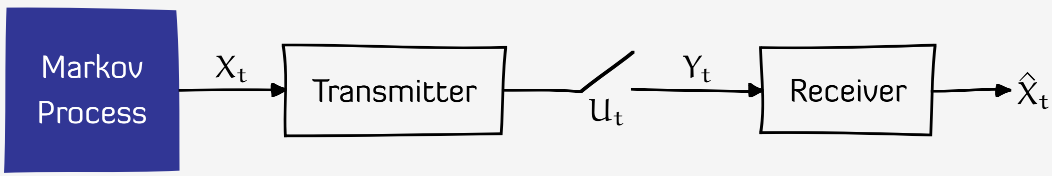
- Sensing is cheap
- Transmission is expensive
- Size of data-packet is not critical

Analyze a stylized model and evaluate fundamental trade-offs

The system model

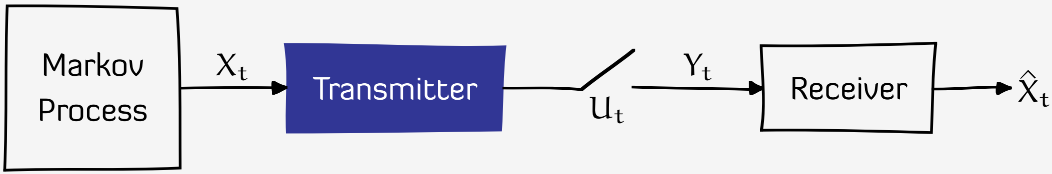


The system model

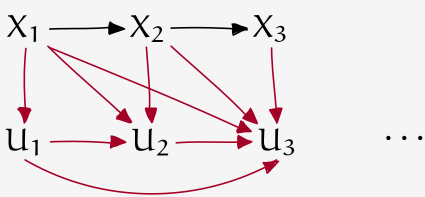


The system model

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



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

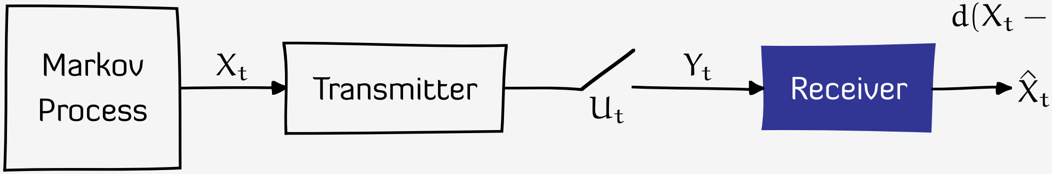


The system model

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

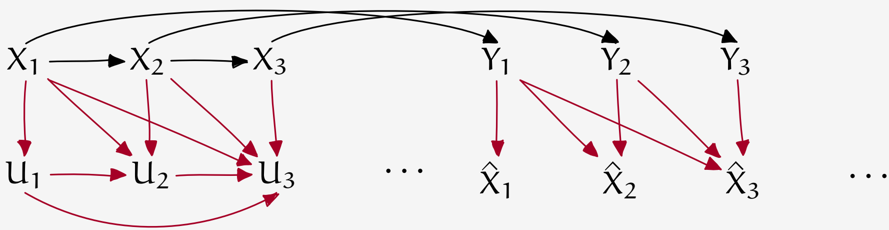
Distortion

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



$u_t = f_t(X_{1:t}, u_{1:t-1})$

$\hat{X}_t = g_t(Y_{1:t})$

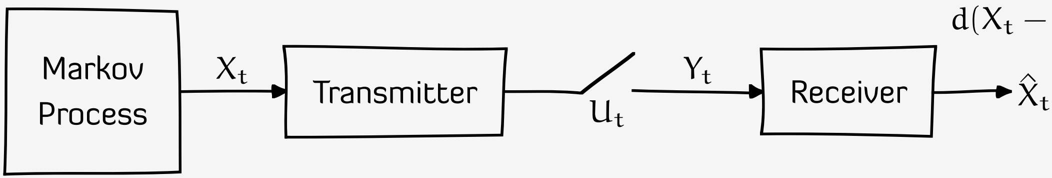


The system model

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

Distortion

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



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

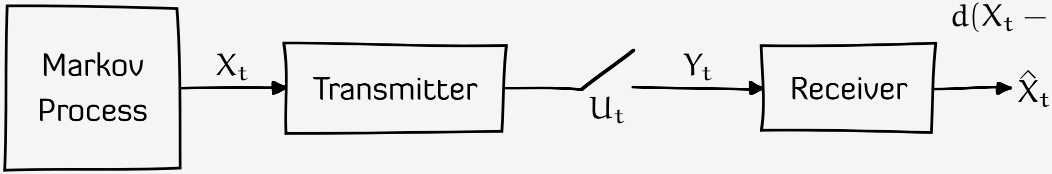
Communication Strategies

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

The system model

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

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



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

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

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

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

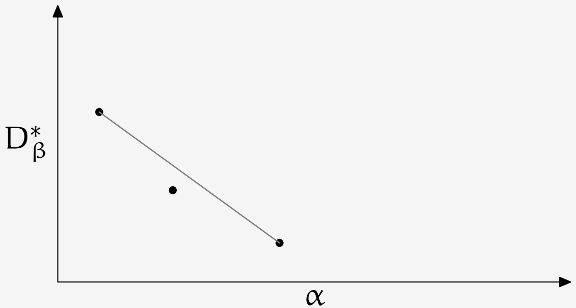
Constrained communication

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

Optimization problems

Constrained communication

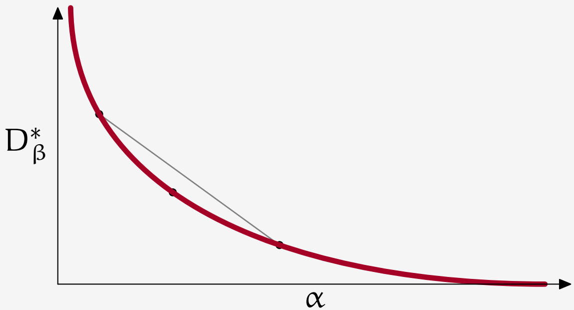
$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \left\{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \right\}$$



Optimization problems

Constrained communication

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



D^*_β is cts, dec, and convex

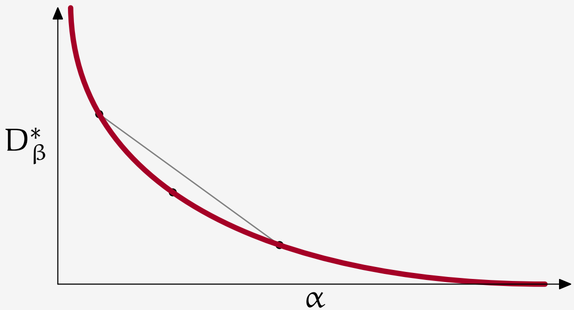
Optimization problems

Constrained communication

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

Costly communication (Lagrange relaxation)

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



D^*_β is cts, dec, and convex

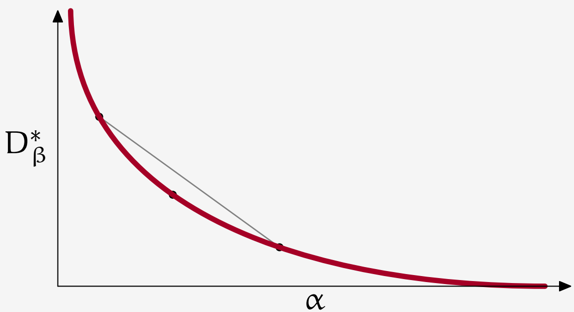
Optimization problems

Constrained communication

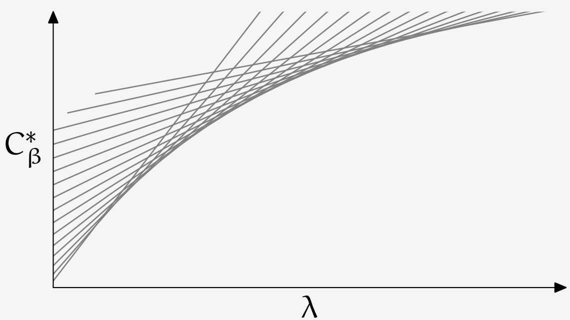
$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \{ D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha \}$$

Costly communication (Lagrange relaxation)

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



D^*_β is cts, dec, and convex



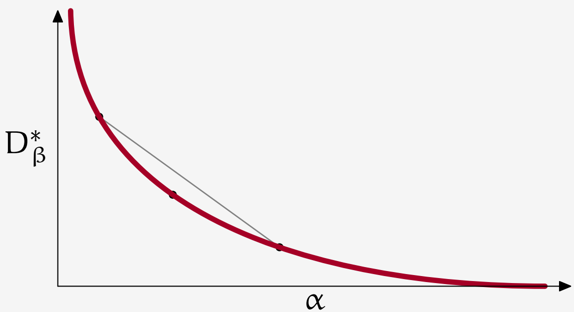
Optimization problems

Constrained communication

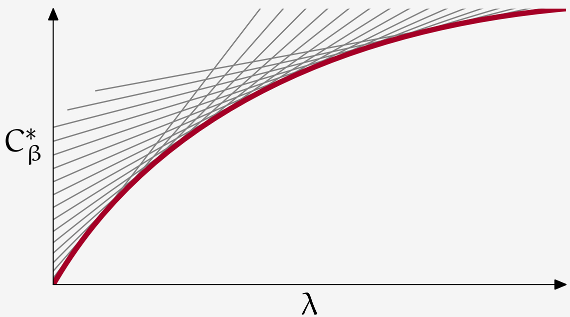
$$\text{For } \alpha \in (0,1), \quad D^*_\beta(\alpha) \coloneqq \inf_{(f,g)} \{D_\beta(f,g) : N_\beta(f,g) \leqslant \alpha\}$$

Costly communication (Lagrange relaxation)

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



D^*_β is cts, dec, and convex



C^*_β is cts, inc, and concave

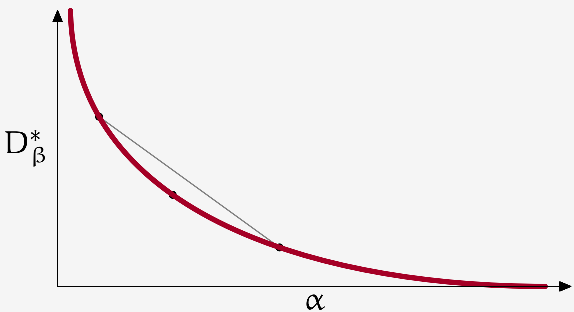
Optimization problems

Constrained communication

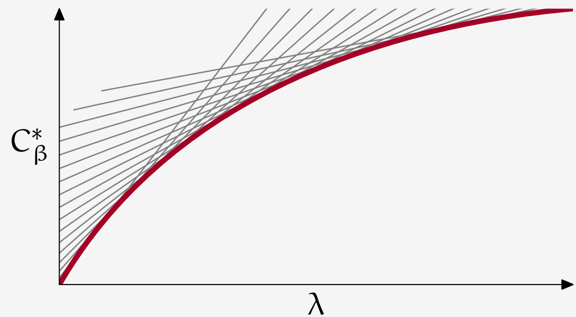
C

$\lambda N_\beta(f, g)\}$

Our result: Provide computable expressions for these curves and identify 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.
- Distortion-transmission is analogous to distortion-rate function.
- 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 function.
- 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 function.
- 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.

Other related work

- Event-based control . . .
- Censoring censors . . .
- Sensor sleep scheduling

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)

- [Imer-Başar 2010]

  i.i.d. Gaussian source with fixed number of transmissions
- [Rabi, Moustakides, Baras 2012]

  Gauss-Markov source with fixed number of transmissions
- [Lipsa-Martins 2011, Molin-Hirche 2009]

  Gauss-Markov source with communication cost
- [Nayyar-Başar-Teneketzis-Veeravalli 2013]

  Discrete-Markov source with communication cost (and energy harvesting)

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)

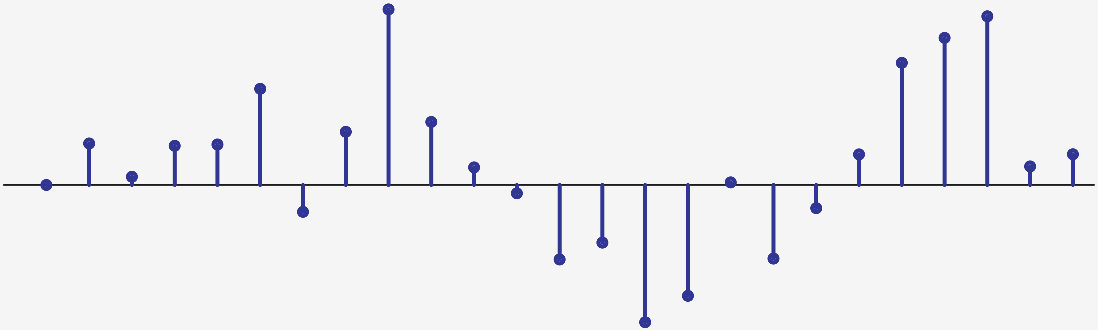
- [Imer-Başar 2010]
  i.i.d. Gaussian source with fixed number of transmissions
- [Rabi, Moustakides, Baras 2012]
  Gauss-Markov source with fixed number of transmissions
- [Lipsa-Martins 2011, Molin-Hirche 2009]
  Gauss-Markov source with communication cost
- [Nayyar-Başar-Teneketzis-Veeravalli 2013]
  Discrete-Markov source with communication cost (and energy harvesting)

We build on these results to identify the distortion-transmission function.

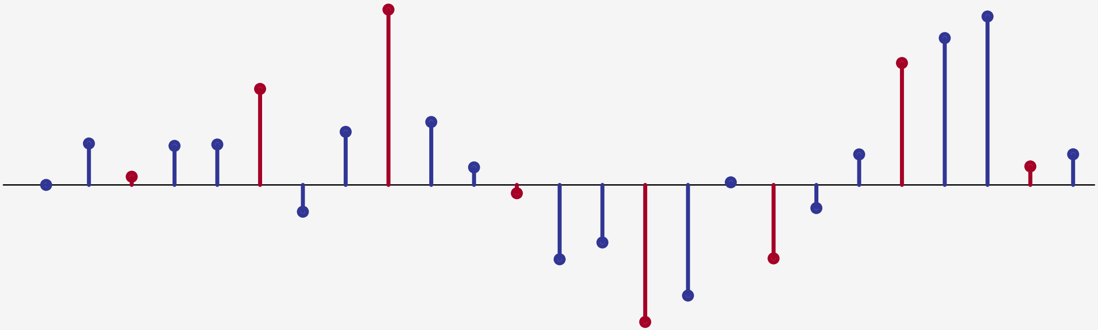
An illustrative example

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

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

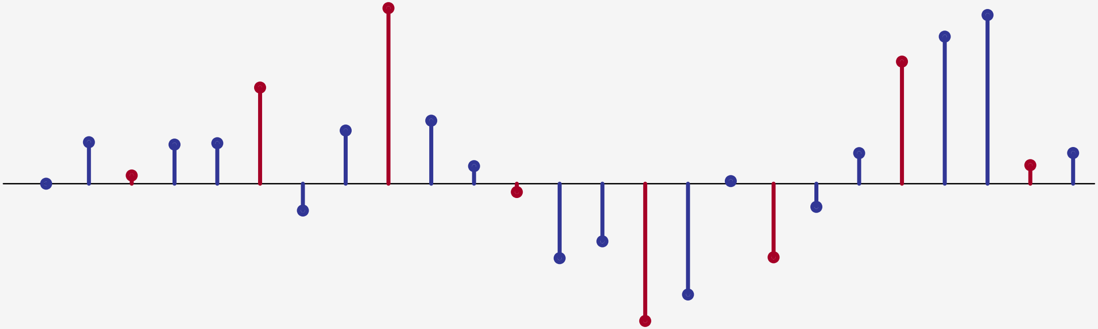


Periodic transmission strategy

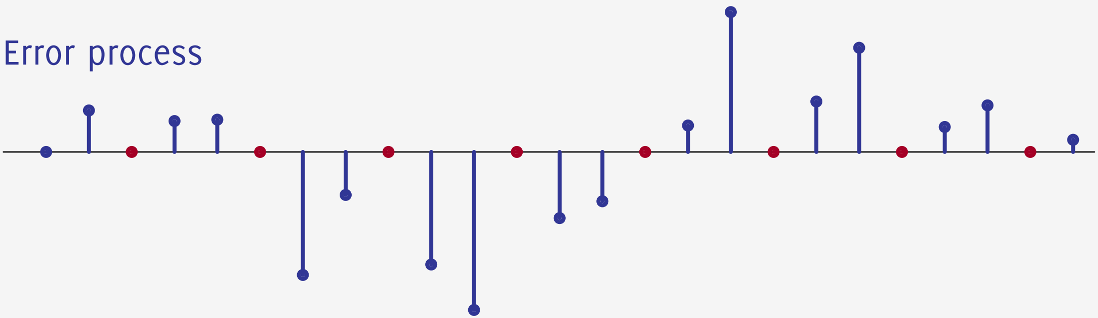


Distortion transmission function–(Mahajan and Chakravorty)

Periodic transmission strategy

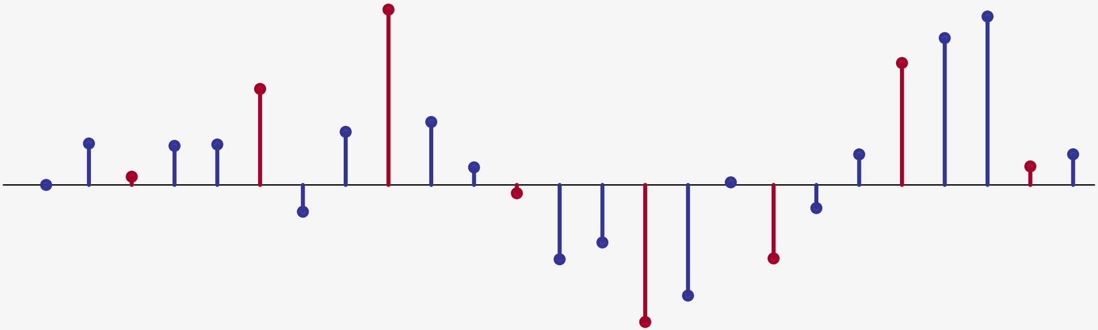


Error process

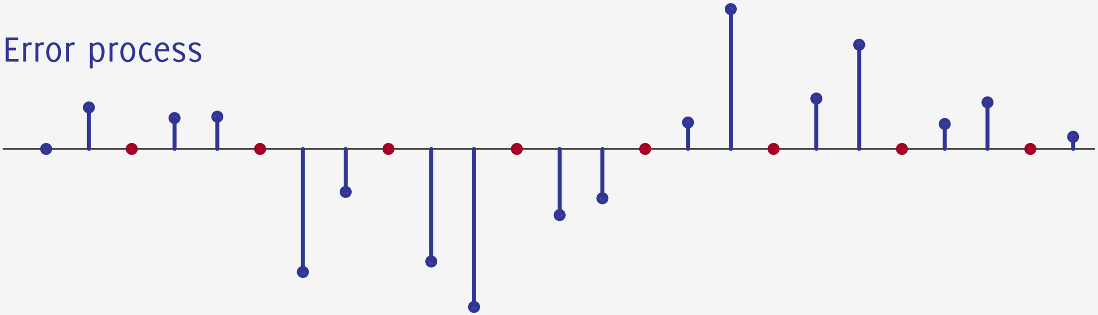


Distortion transmission function–(Mahajan and Chakravorty)

Periodic transmission strategy



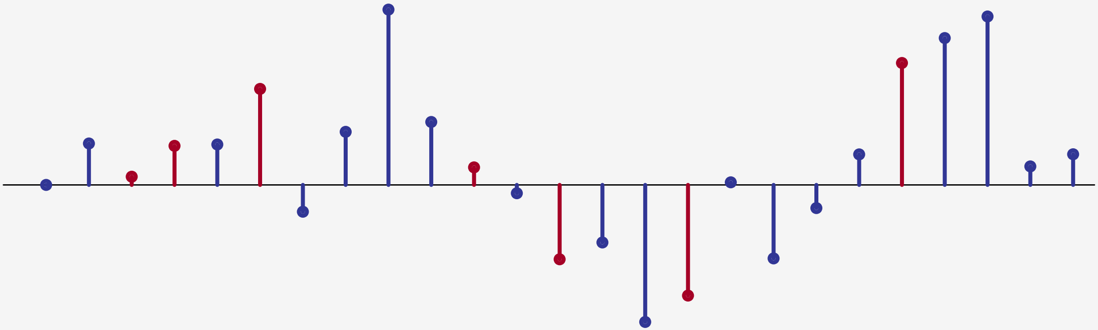
Error process



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

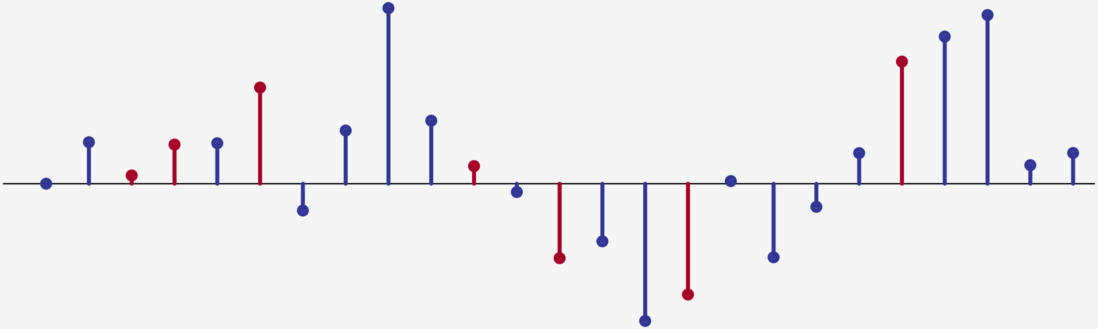
Distortion transmission function–(Mahajan and Chakravorty)

Randomized strategy

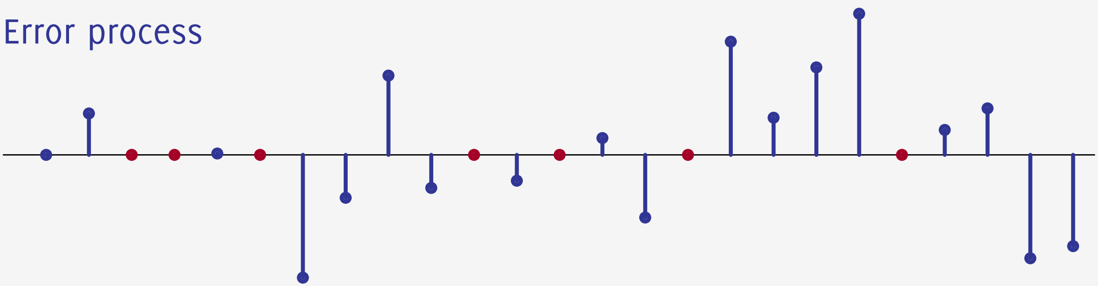


Distortion transmission function–(Mahajan and Chakravorty)

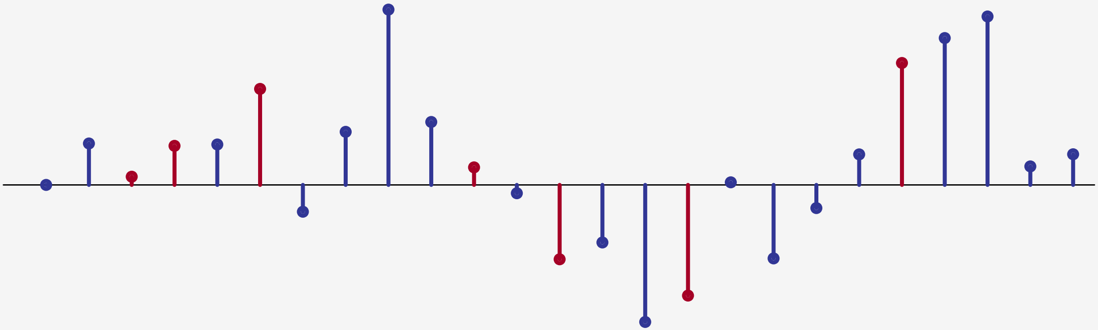
Randomized strategy



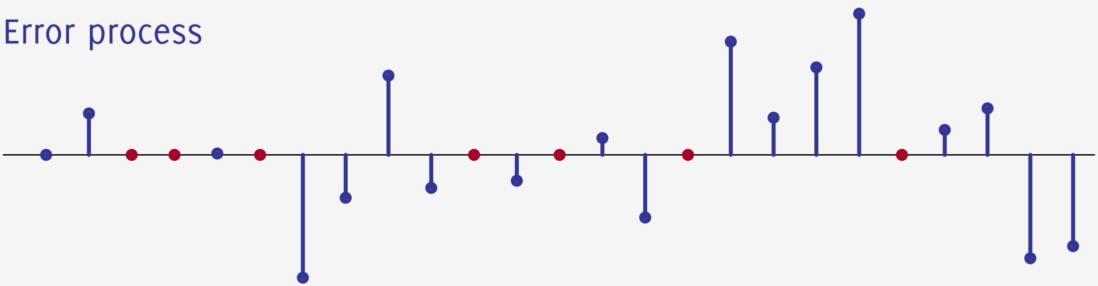
Error process



Randomized strategy

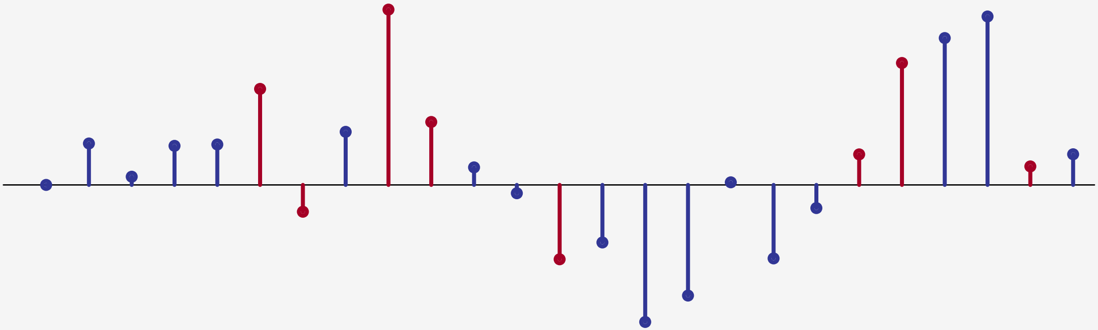


Error process



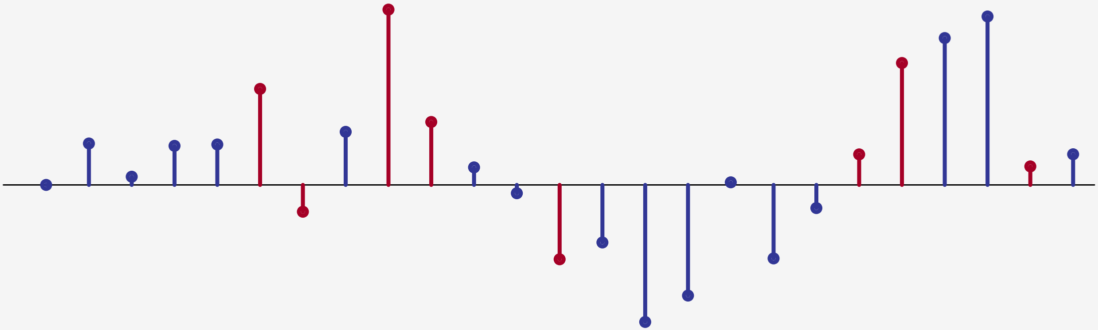
$D = 2.00 \quad N \approx 1/3$

The optimal strategy

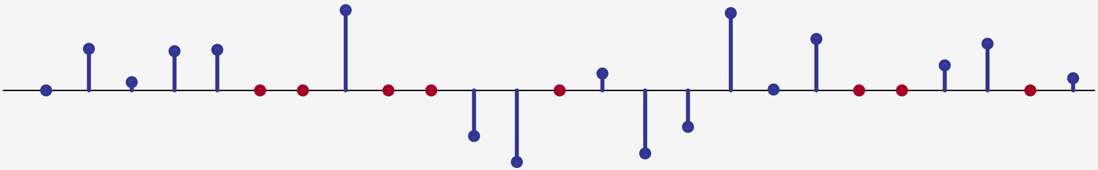


Distortion transmission function–(Mahajan and Chakravorty)

The optimal strategy

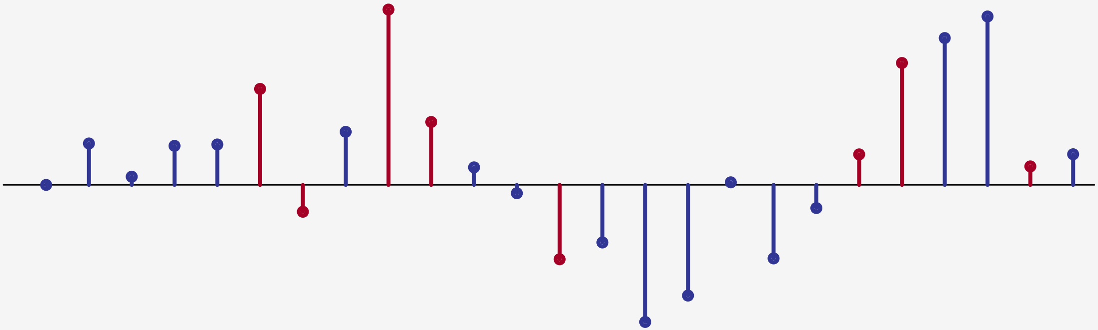


Error process

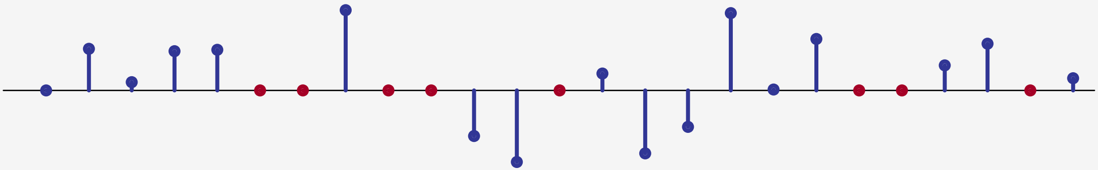


Distortion transmission function–(Mahajan and Chakravorty)

The optimal strategy

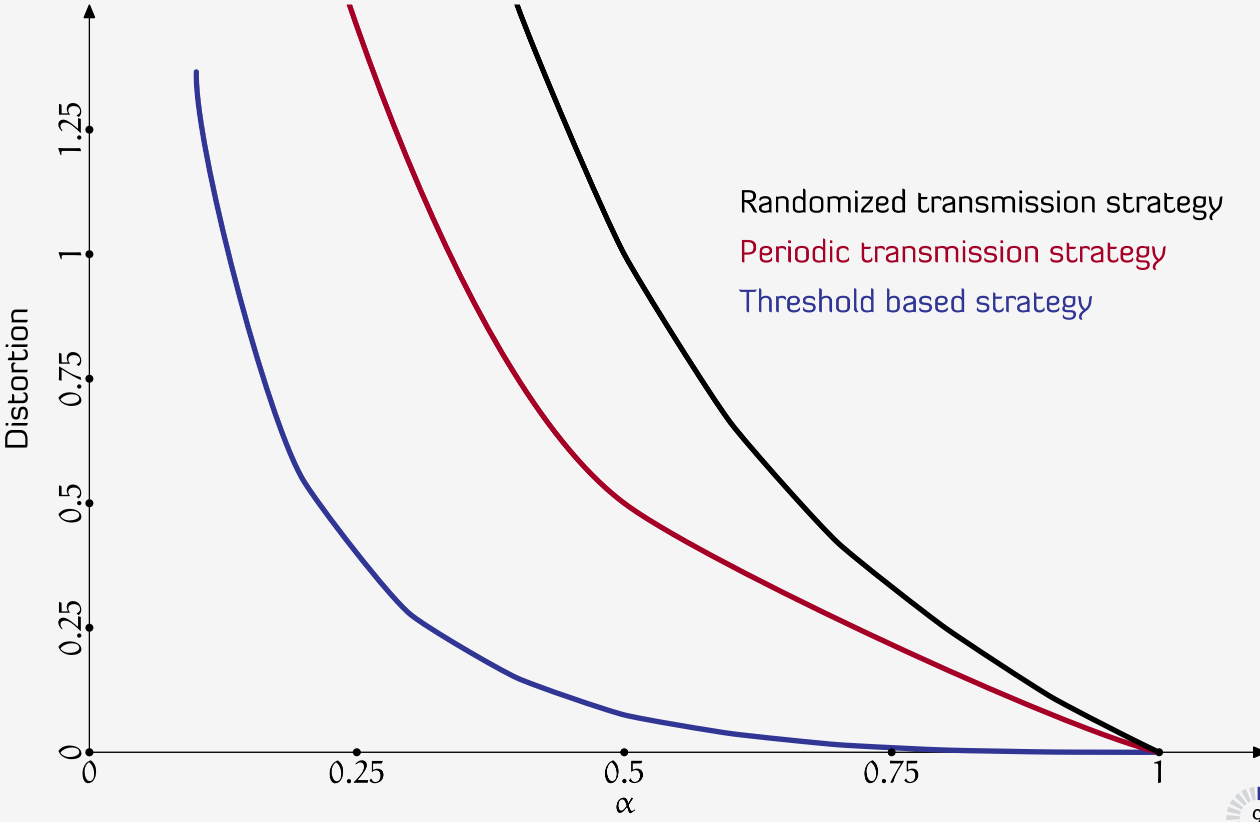


Error process



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

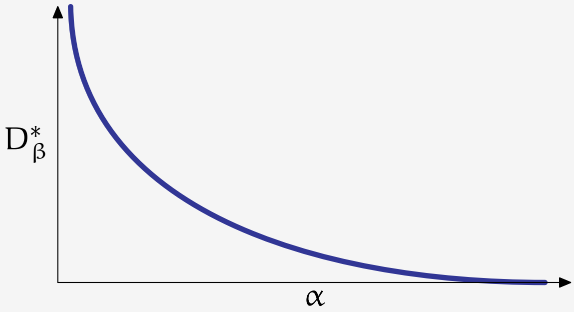
Distortion-transmission function



Distortion transmission function–(Mahajan and Chakravorty)

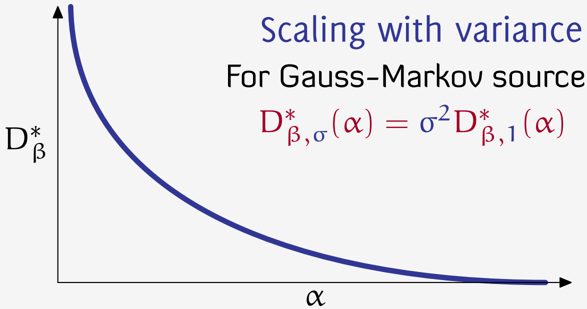
Main results

Distortion transmission function for continuous AR sources



D^*_β is cts, dec, and convex

Distortion transmission function for continuous AR sources



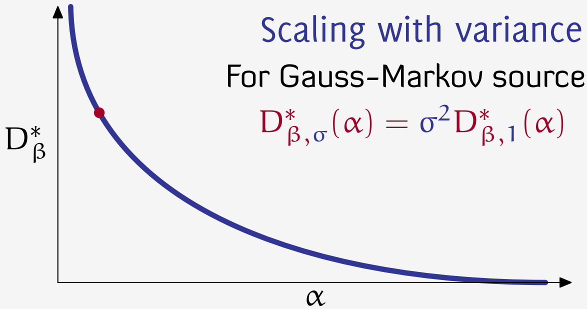
Scaling with variance

For Gauss-Markov sources:

$$D^*_{\beta,\sigma}(\alpha) = \sigma^2 D^*_{\beta,1}(\alpha)$$

D^*_β is cts, dec, and convex

Distortion transmission function for continuous AR sources



D^*_β is cts, dec, and convex

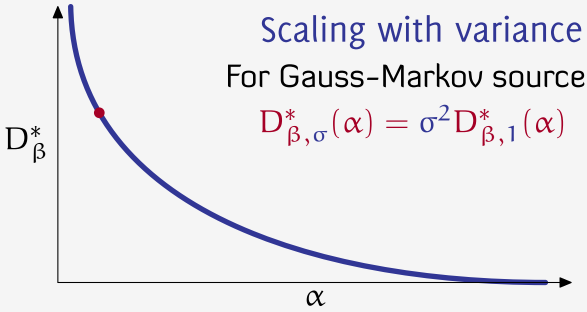
Scaling with variance

For Gauss-Markov sources:

$$D^*_{\beta,\sigma}(\alpha) = \sigma^2 D^*_{\beta,1}(\alpha)$$

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

Distortion transmission function for continuous AR sources



D^*_β is cts, dec, and convex

Scaling with variance

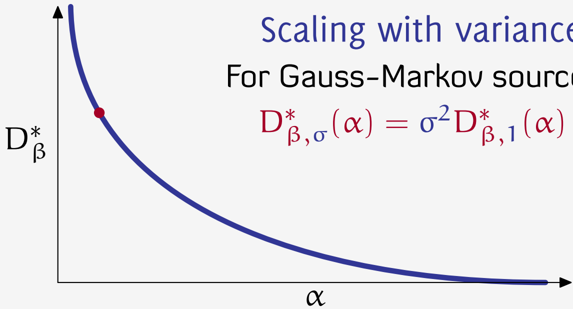
For Gauss-Markov sources:

$$D^*_{\beta,\sigma}(\alpha) = \sigma^2 D^*_{\beta,1}(\alpha)$$

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

- Find $k^*(\alpha)$ such that $M_\beta^{(k^*)}(0) = \frac{1}{\alpha}$, where $M_\beta^{(k)}(e) = 1 + \beta \int_{-k}^{k} \varphi(w - ae) M^{(k)}(w) dw$
- Compute $L_\beta^{(k*)}(0)$ where $L_\beta^{(k)}(e) = d(e) + \beta \int_{-k}^{k} \varphi(w - ae) L^{(k)}(w) dw$
- Then $D^*_\beta(\alpha) = \dfrac{L_\beta^{k^*}(0)}{M_\beta^{k^*}(0)}$

Distortion transmission function for continuous AR sources



D^*_β is cts, dec, and convex

Scaling with variance

For Gauss-Markov sources:

$D^*_{\beta,\sigma}(\alpha) = \sigma^2 D^*_{\beta,1}(\alpha)$

Optimal transmission strategy

Transmit when $|X_t - a\hat{X}_t| > k^*(\alpha)$

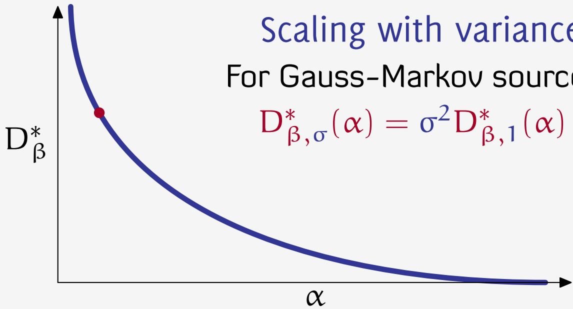
Optimal estimation strategy

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

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

- Find $k^*(\alpha)$ such that $M^{(k^*)}_\beta(0) = \frac{1}{\alpha}$, where $M^{(k)}_\beta(e) = 1 + \beta \int_{-k}^{k} \varphi(w - ae) M^{(k)}(w) dw$
- Compute $L^{(k*)}_\beta(0)$ where $L^{(k)}_\beta(e) = d(e) + \beta \int_{-k}^{k} \varphi(w - ae) L^{(k)}(w) dw$
- Then $D^*_\beta(\alpha) = \dfrac{L^{k^*}_\beta(0)}{M^{k^*}_\beta(0)}$

Distortion transmission function for continuous AR sources



Scaling with variance

For Gauss-Markov sources:

$D^*_{\beta,\sigma}(\alpha) = \sigma^2 D^*_{\beta,1}(\alpha)$

D^*_β is cts, dec, and convex

Optimal transmission strategy

Transmit when $|X_t - a\hat{X}_t| > k^*(\alpha)$

Optimal estimation strategy

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

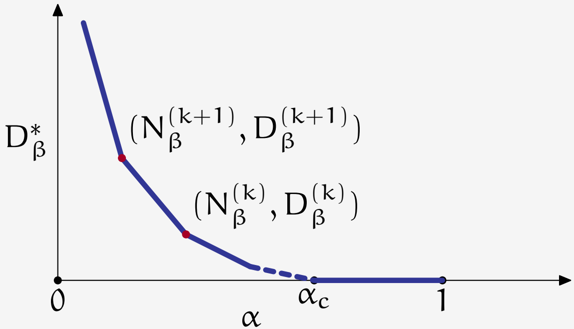
Salient features

- 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

$)dw$

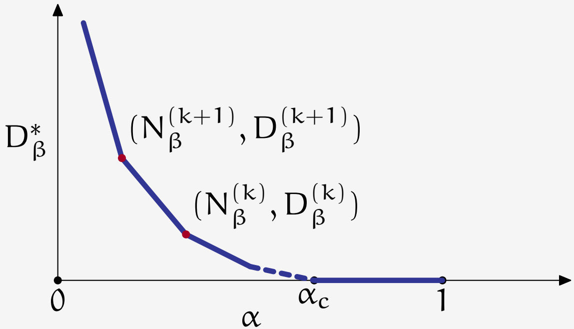
$M_\beta(0)$

Distortion transmission function for discrete AR sources



D^*_β is PWL, dec, and convex

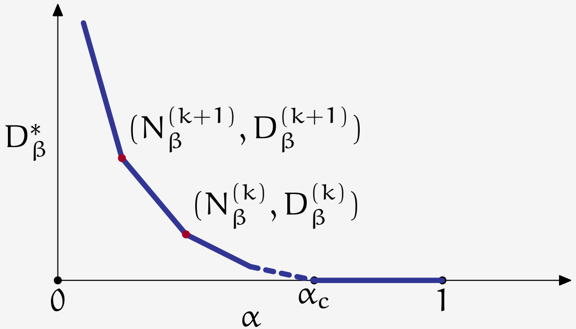
Distortion transmission function for discrete AR sources



D^*_β is PWL, dec, and convex

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

Distortion transmission function for discrete AR sources



D^*_β is PWL, dec, and convex

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

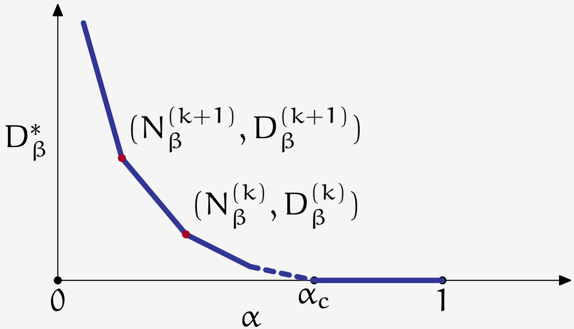
- Compute $L_\beta^{(k)} = \big[[I - \beta P^{(k)}]^{-1} d^{(k)}\big].$

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

- Then

$$D_\beta^{(k)} = \frac{L^{(k)}(0)}{M_\beta^{(k)}(0)} \quad \text{and} \quad N_\beta^{(k)} = \frac{1}{M_\beta^{(k)}(0)} - (1-\beta)$$

Distortion transmission function for discrete AR sources



D^*_β is PWL, dec, and convex

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

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

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

- Then

$$D^{(k)}_\beta = \frac{L^{(k)}(0)}{M^{(k)}_\beta(0)} \quad \text{and} \quad N^{(k)}_\beta = \frac{1}{M^{(k)}_\beta(0)} - (1 - \beta)$$

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 $|X_t - a\hat{X}_t| > k^*(\alpha)$, transmit.
- If $|X_t - a\hat{X}_t| = k^*(\alpha)$, transmit w.p. θ^*.
- Else, do not transmit.

Optimal estimation strategy

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

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 solving Fredholm integral equations for continuous processes

Based on simple matrix calculations for discrete 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 solving Fredholm integral equations for continuous processes

Based on simple matrix calculations for discrete processes

Beautiful example of stochastics and optimization

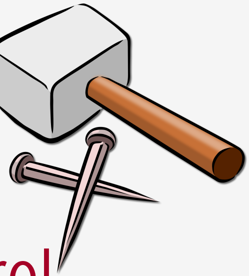
Decentralized stochastic control and POMDPs

Stochastic orders and majorization

Markov chain analysis, stopping times, and renewal theory

Constrained MDPs and Lagrangian relaxations

So how do we start?



Decentralized stochastic control

The common information approach

Original system

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

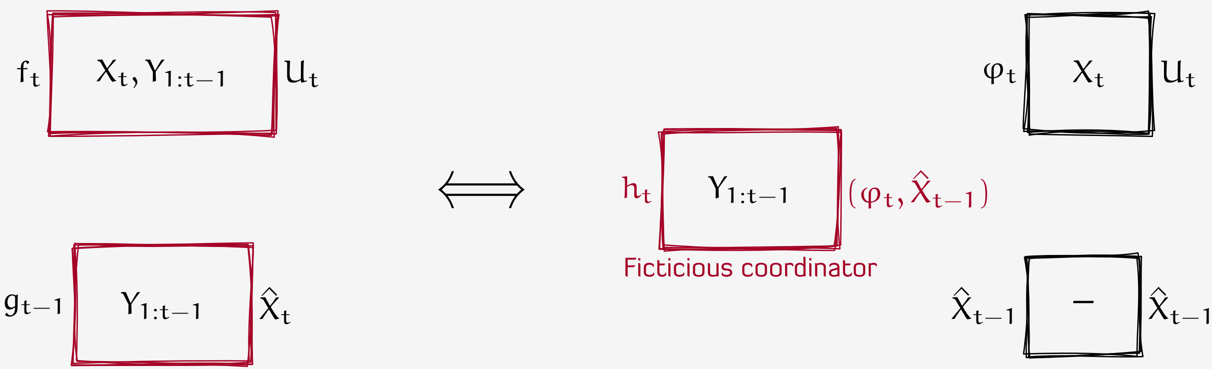
g_{t-1} $\boxed{Y_{1:t-1}}$ $\hat{X}_t$

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

The common information approach

Original system

Coordinated system

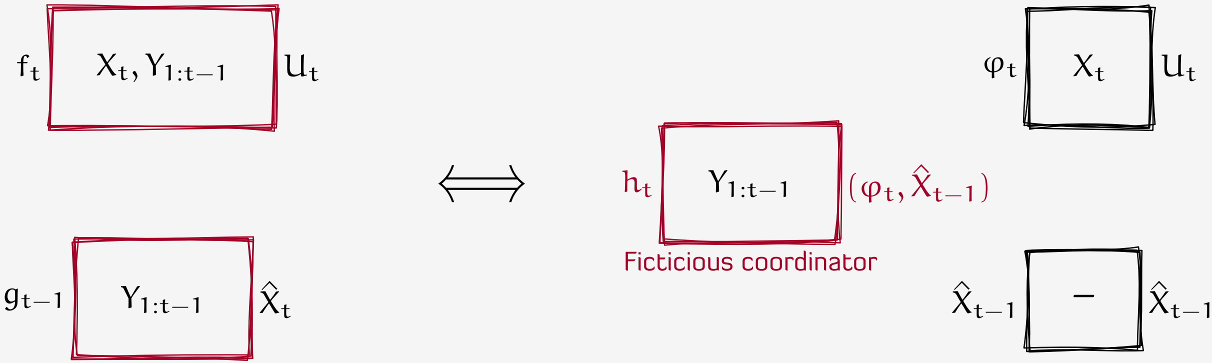


- Nayyar, Mahajan and Teneketzis, “Decentralized stochastic control with partial history sharing: A common information approach,” IEEE TAC 2013.

The common information approach

Original system

Coordinated system



- 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 $\mathbb{P}(X_t \mid Y_{1:t-1})$.

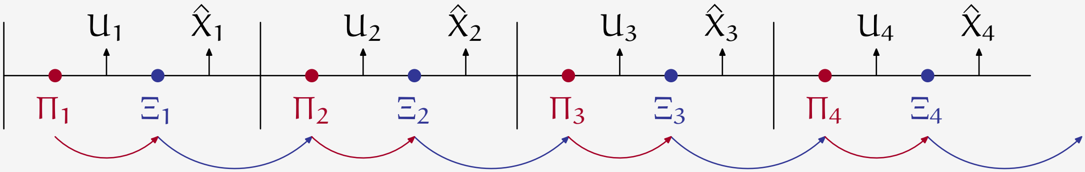
▸ Nayyar, Mahajan and Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," IEEE TAC 2013.

Information states and dynamic program

Information states

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

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

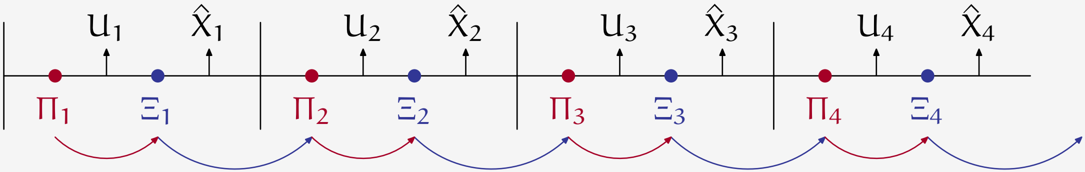


Information states and dynamic program

Information states

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

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



Structural results

There is no loss of optimality in using

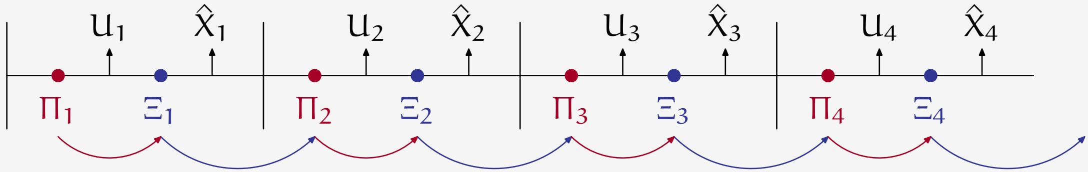
$$U_t = f_t(X_t, \Pi_t) \quad \text{and} \quad \hat{X}_t = g_t(\Xi_t).$$

Information states and dynamic program

Information states

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

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



Structural results

There is no loss of optimality in using

$$U_t = f_t(X_t, \Pi_t) \quad \text{and} \quad \hat{X}_t = g_t(\Xi_t).$$

Dynamic Program

$$W_{T+1}(\pi) = 0$$

and for $t = T, \dots, 0$

$$V_t(\xi) = \min_{\hat{x} \in \mathcal{X}} \mathbb{E}[d(X_t - \hat{x}) + W_{t+1}(\Pi_{t+1}) \mid \Xi_t = \xi],$$

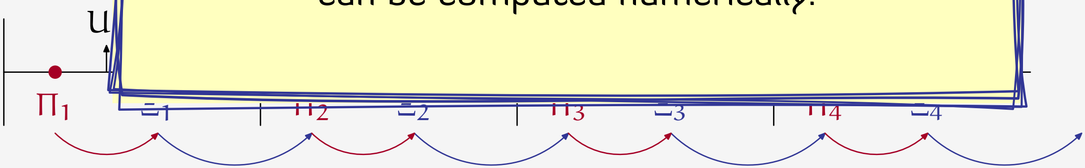
$$W_t(\pi) = \min_{\varphi : \mathcal{X} \to \{0,1\}} \mathbb{E}[\lambda \varphi(X_t) + V_t(\Xi_t) \mid \Pi_t = \pi, \varphi_t = \varphi].$$

Information states and dynamic program

Informa … $x \mid Y_{1:t-1})$.

… $x \mid Y_{1:t})$.

"Standard" POMDP. Optimal strategies can be computed numerically.



Structural results

There is no loss of optimality in using

$$U_t = f_t(X_t, \Pi_t) \quad \text{and} \quad \hat{X}_t = g_t(\Xi_t).$$

Dynamic Program

$$W_{T+1}(\pi) = 0$$

and for $t = T, \dots, 0$

$$V_t(\xi) = \min_{\hat{x} \in \mathcal{X}} \mathbb{E}[d(X_t - \hat{x}) + W_{t+1}(\Pi_{t+1}) \mid \Xi_t = \xi],$$

$$W_t(\pi) = \min_{\varphi : \mathcal{X} \to \{0,1\}} \mathbb{E}[\lambda \varphi(X_t) + V_t(\Xi_t) \mid \Pi_t = \pi, \varphi_t = \varphi].$$

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 $\varphi(\cdot)$

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 $\varphi(\cdot)$

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 Use results from constrained optimization

Step 1 Threshold-based strategies are optimal

[Lipsa Martins 2011, Nayyar et. al. 2013]

Oblivious estimation process

$$Z_t = \begin{cases} X_t & \text{if } U_t = 1 \text{ (or } Y_t \neq \varepsilon) \\ aZ_{t-1} & \text{if } U_t = 0 \text{ (or } Y_t = \varepsilon) \end{cases}$$

Error process

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

Step 1 Threshold-based strategies are optimal

[Lipsa Martins 2011, Nayyar et. al. 2013]

Oblivious estimation process

$$Z_t = \begin{cases} X_t & \text{if } U_t = 1 \text{ (or } Y_t \neq \varepsilon) \\ aZ_{t-1} & \text{if } U_t = 0 \text{ (or } Y_t = \varepsilon) \end{cases}$$

Error process

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

Optimal estimator

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

Optimal transmitter

There exists thresholds $\{k_t\}_{t \geqslant 0}$ such that

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

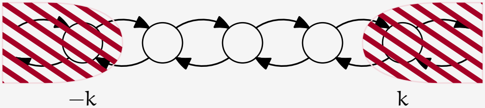
For infinite-horizon setup time-homogeneous threshold-based strategies are optimal.

How do we find the optimal threshold-based strategy?

Step 2 Performance of threshold-based strategies

Consider a threshold-based strategy

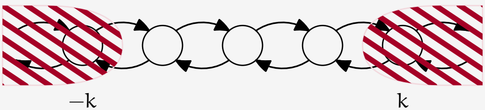
$$f^{(k)}(e) = \begin{cases} 1 & \text{if } |e| \geq k \\ 0 & \text{otherwise} \end{cases}$$



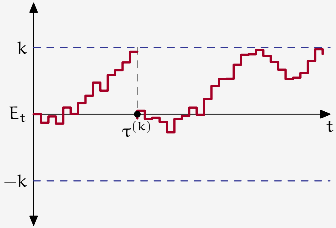
Step 2 Performance of threshold-based strategies

Consider a threshold-based strategy

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



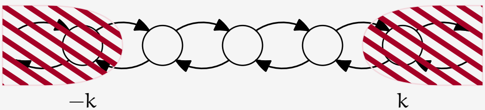
Let $\tau^{(k)}$ denote the stopping time of first transmission (starting at $E_0 = 0$).



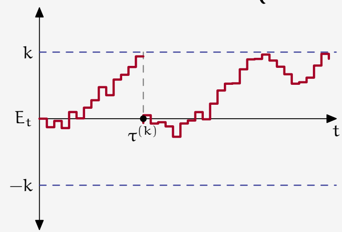
Performance of threshold-based strategies

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 transmission (starting at $E_0 = 0$).



Define

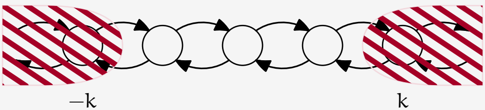
$$L_\beta^{(k)}(e) = (1-\beta)\,\mathbb{E}\left[\sum_{t=0}^{\tau^{(k)}-1} \beta^t d(E_t) \middle| E_0 = e\right].$$

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

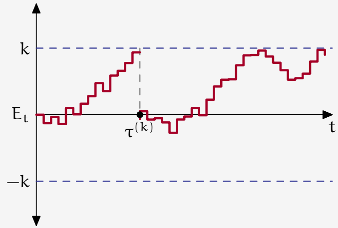
Step 2 Performance of threshold-based strategies

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 transmission (starting at $E_0 = 0$).



Define

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

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

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

Conside … of

… 0)).

$f^{(k)}$

$-k$

Computing $L^{(k)}_\beta$ and $M^{(k)}_\beta$ is sufficient
to compute the performance of $f^{(k)}$
(i.e., to compute $D^{(k)}_\beta$ and $N^{(k)}_\beta$).

Define

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

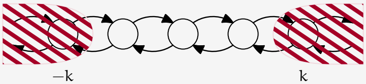
$$M^{(k)}_\beta(e) = (1-\beta)\,\mathbb{E}\Big[\sum_{t=0}^{\tau^{(k)}-1} \beta^t \Big| E_0 = e\Big].$$

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

Discrete state setup

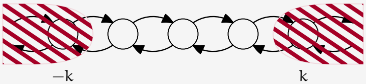


$$L_\beta^{(k)}(e) = d(e) + \beta \sum_{n=-k}^{k} p_{n-e} L_\beta^{(k)}(n)$$

$$M_\beta^{(k)}(e) = 1 + \beta \sum_{n=-k}^{k} p_{n-e} M_\beta^{(k)}(n)$$

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

Discrete state setup



$$L_\beta^{(k)}(e) = d(e) + \beta \sum_{n=-k}^{k} p_{n-e} L_\beta^{(k)}(n)$$

$$M_\beta^{(k)}(e) = 1 + \beta \sum_{n=-k}^{k} p_{n-e} M_\beta^{(k)}(n)$$

Proposition

$$L_\beta^{(k)} = \left[[I - \beta P^{(k)}]^{-1} d^{(k)}\right].$$

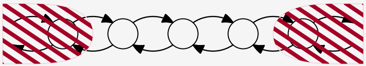
$P^{(k)}$ is substochastic.

$$M_\beta^{(k)} = \left[[I - \beta P^{(k)}]^{-1} \mathbf{1}^{(k)}\right].$$

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

Discrete state setup

$$L_\beta^{(k)}(e) = d(e) + \beta \sum_{n=-k}^{k} p_{n-e} L_\beta^{(k)}(n)$$



$$M_\beta^{(k)}(e) = 1 + \beta \sum_{n=-k}^{k} p_{n-e} M_\beta^{(k)}(n)$$

Proposition

$$L_\beta^{(k)} = \left[[I - \beta P^{(k)}]^{-1} d^{(k)} \right].$$

$P^{(k)}$ is substochastic.

$$M_\beta^{(k)} = \left[[I - \beta P^{(k)}]^{-1} \mathbf{1}^{(k)} \right].$$

Continuous state setup

$$L_\beta^{(k)}(e) = d(e) + \beta \int_{-k}^{k} \varphi(n-e) L_\beta^{(k)}(n) dn$$

$$M_\beta^{(k)}(e) = 1 + \beta \int_{-k}^{k} \varphi(n-e) M_\beta^{(k)}(n) dn$$

Fredholm Integral Equations of the 2nd kind.

Solutions exist and are unique and easy to compute.

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

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

Proposition

$$L_\beta^{(k)} = \left[[I - \beta P^{(k)}]^{-1} d^{(k)}\right].$$

$P^{(k)}$ is substochastic.

$$M_\beta^{(k)} = \left[[I - \beta P^{(k)}]^{-1} \mathbf{1}^{(k)}\right].$$

Continuous state setup

$$L_\beta^{(k)}(e) = d(e) + \beta \int_{-k}^{k} \varphi(n-e) L_\beta^{(k)}(n)\,dn$$

$$M_\beta^{(k)}(e) = 1 + \beta \int_{-k}^{k} \varphi(n-e) M_\beta^{(k)}(n)\,dn$$

Fredholm Integral Equations of the 2nd kind.

Solutions exist and are unique and easy to compute.

Step 3 Solution to constrained optimization 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°.

We find the choice of thresholds such that these conditions are satisfied.

Step 3 Solution to constrained optimization problem

Sufficient condition for optimality

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

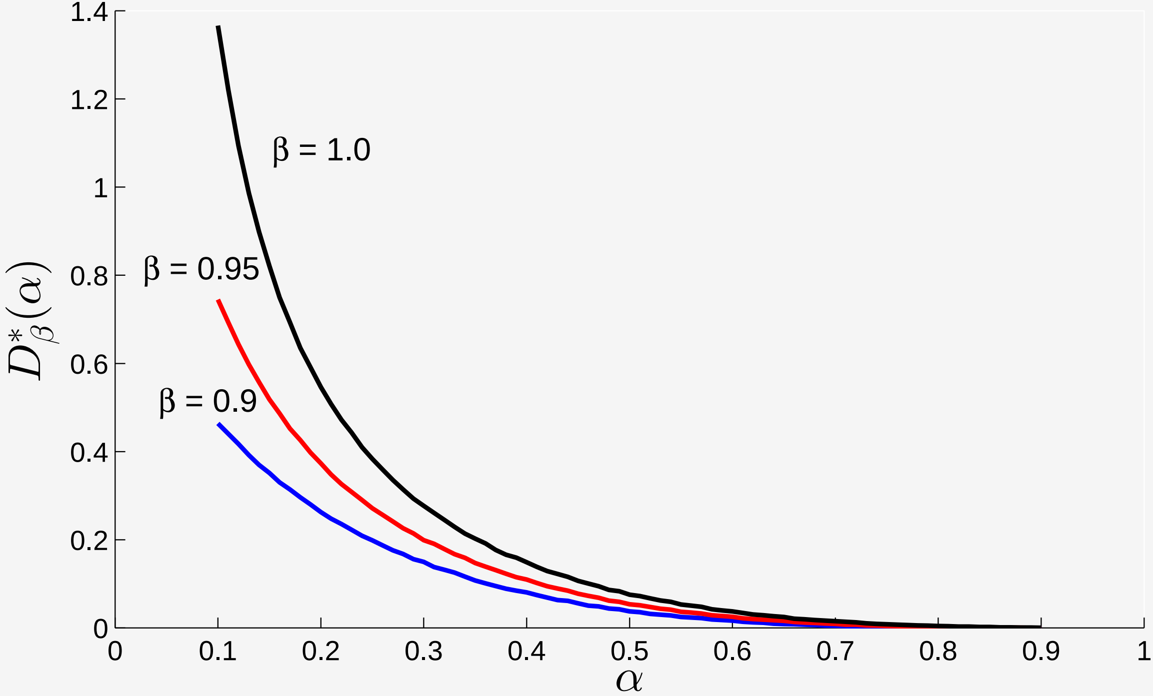
(C1) $N_\beta(f^\circ, g^\circ) = \alpha$

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

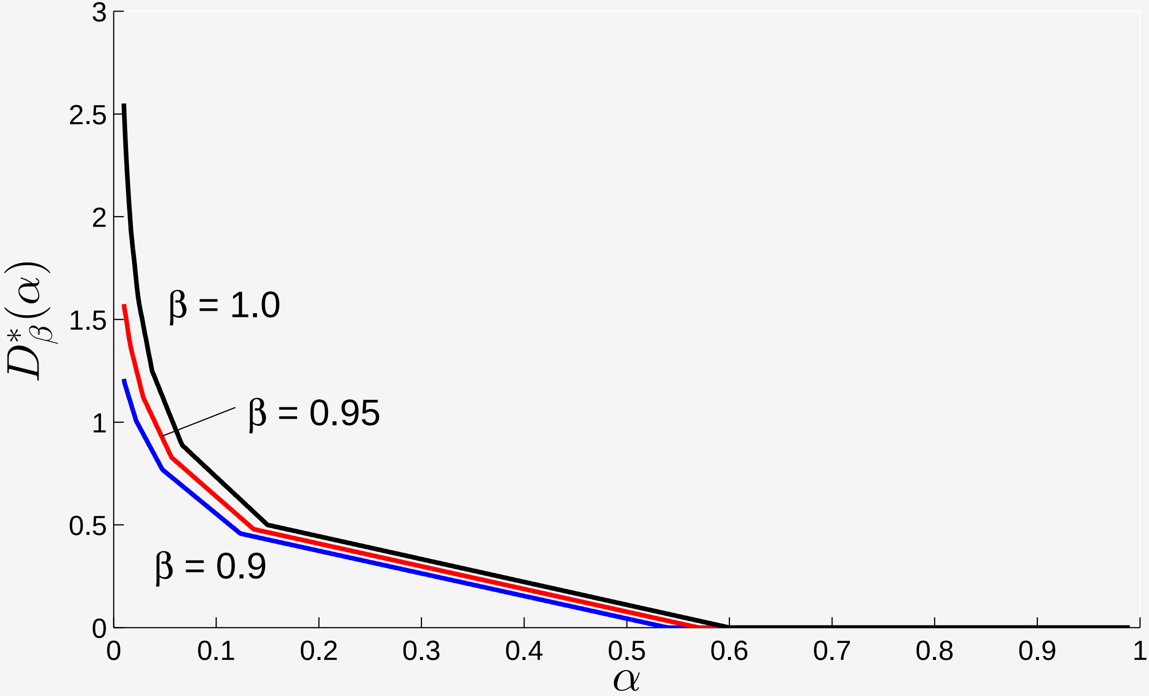
We find the choice of thresholds such that these conditions are satisfied.

Examples

Example Gauss-Markov with $\sigma^2 = 1$



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



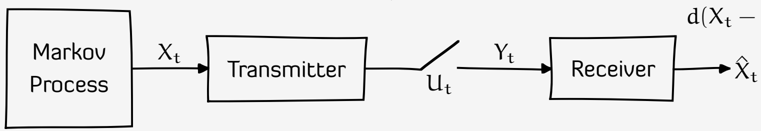
Summary

The system model

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

Distortion

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



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

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

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

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

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

Distortion transmission function–(Mahajan and Chakravorty) 2

Summary

The system model

1. Disc

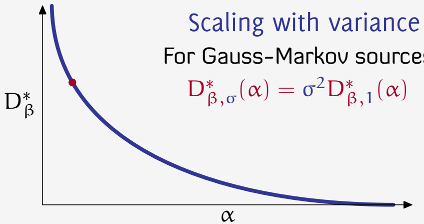
$D_\beta($

2. Ave

$D_1($

Distortio

Distortion transmission function for continuous AR sources



Scaling with variance

For Gauss-Markov sources:

$D^*_{\beta,\sigma}(\alpha) = \sigma^2 D^*_{\beta,1}(\alpha)$

D^*_β is cts, dec, and convex

Optimal transmission strategy

Transmit when $|X_t - a\hat{X}_t| > k^*(\alpha)$

Optimal estimation strategy

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

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

- Find $k^*(\alpha)$ such that $M^{(k^*)}_\beta(0) = \dfrac{1}{\alpha}$, where $M^{(k)}_\beta(e) = 1 + \beta \displaystyle\int_{-k}^{k} \varphi(w - ae) M^{(k)}(w)\,dw$
- Compute $L^{(k*)}_\beta(0)$ where $L^{(k)}_\beta(e) = d(e) + \beta \displaystyle\int_{-k}^{k} \varphi(w - ae) L^{(k)}(w)\,dw$
- Then $D^*_\beta(\alpha) = \dfrac{L^{k^*}_\beta(0)}{M^{k^*}_\beta(0)}$

Summary

The system model

Distortion transmission function for continuous AR sources

Scaling with variance

D^*_β

1. Disc

$D_\beta($

H

▸ F

2. Ave

$D_1($

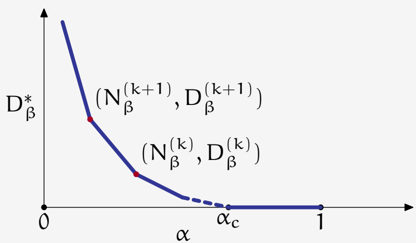
▸ C

▸ T

Distortio

Distortio

Distortion transmission function for discrete AR sources



D^*_β is PWL, dec, and convex

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

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

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

- Then

$$D^{(k)}_\beta = \frac{L^{(k)}(0)}{M^{(k)}_\beta(0)} \quad \text{and} \quad N^{(k)}_\beta = \frac{1}{M^{(k)}_\beta(0)} - (1 - \beta)$$

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 $|X_t - a\hat{X}_t| > k^*(\alpha)$, transmit.
- If $|X_t - a\hat{X}_t| = k^*(\alpha)$, transmit w.p. θ^*.
- Else, do not transmit.

Optimal estimation strategy

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

Conclusion

Analyze fundamental limits of estimation
under communication constraints

Conclusion

Analyze fundamental limits of estimation under communication constraints

Results are derived under idealized assumptions

- No rate constraint
- Noiseless transmission

Possible generalizations to more realistic models

- Noisy source observations
- Rate constraints (effect of quantization)
- Packet drops
- Network delays

Conclusion

Analyze fundamental limits of estimation under communication constraints

Results are derived under idealized assumptions

- No rate constraint
- Noiseless transmission

Possible generalizations to more realistic models

- Noisy source observations
- Rate constraints (effect of quantization)
- Packet drops
- Network delays

Current work: Interactive communication

Conclusion

Analyze fundamental limits of estimation under communication constraints

Results are derived under idealized assumptions

- No rate constraint
- Noiseless transmission

Possible generalizations to more realistic models

- Noisy source observations
- Rate constraints (effect of quantization)
- Packet drops
- Network delays

Current work: Interactive communication

Full version available at arXiv:1505.04829.