Fundamental limits of remote-estimation under communication constraints

Aditya Mahajan
McGill University

Joint work with Jhelum Chakravorty

Mathematical Cybernetics: Hybrid, Stochastic and Decentralized Systems
Carleton University, Ottawa, 28–29 May 2015
Motivation

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

Sensor Networks
Motivation

Many applications require:
- Sequential transmission of data
- Zero- (or finite-) delay reconstruction

Smart Grids
Motivation

Many applications require:
- Sequential transmission of data
- Zero- (or finite-) delay reconstruction

Internet of Things
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
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
A completely solved example of a “simple” decentralized system with non-classical information structure
The system model

- Markov Process
- Transmitter
- Receiver
- Estimation under communication constraints—(Mahajan and Chakravorty)
The system model

Markov Process

Transmitter

Receiver

$X_t$ $U_t$ $Y_t$ $\hat{X}_t$
The system model

The system model consists of a Markov Process, a Transmitter, a Receiver, and an estimated signal $\hat{X}_t$. The equations are:

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

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

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

Distortion

\[ d(X_t - \hat{X}_t) \]

Estimation under communication constraints–(Mahajan and Chakravorty)

- Markov Process
- Transmitter
- Receiver

\[ U_t = f_t(X_{1:t}, U_{1:t-1}) \quad \hat{X}_t = g_t(Y_{1:t}) \]
The system model

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

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

Communication Strategies

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

Estimation under communication constraints—(Mahajan and Chakravorty)
The system model

\[ Y_t = \begin{cases} X_t, & \text{if } U_t = 1 \\ \epsilon, & \text{if } U_t = 0 \end{cases} \]

\[ U_t = f_t(X_{1:t}, U_{1:t-1}) \]

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

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

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

2. Average cost setup, \( \beta = 1 \)

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

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

Constrained communication
For $\alpha \in (0, 1)$, $D_\beta^*(\alpha) := \inf_{(f, g)} \{ D_\beta(f, g) : N_\beta(f, g) \leq \alpha \}$
Optimization problems

Costly communication

For $\lambda \in \mathbb{R}_{> 0}$, 

$$C^*_\beta(\lambda) = C_\beta(f^*, g^*; \lambda) := \inf_{(f, g)} \{D_\beta(f, g) + \lambda N_\beta(f, g)\}$$

Constrained communication

For $\alpha \in (0, 1)$, 

$$D^*_\beta(\alpha) := \inf_{(f, g)} \{D_\beta(f, g) : N_\beta(f, g) \leq \alpha\}$$
Optimization problems

Costly communication
For $\lambda \in \mathbb{R}_{>0}$, $\mathcal{C}^*_\beta (\lambda) = \mathcal{C}_\beta (f^*, g^*; \lambda) := \inf_{(f,g)} \{ D_\beta (f, g) + \lambda N_\beta (f, g) \}$

Constrained communication
For $\alpha \in (0, 1)$, $\mathcal{D}^*_\beta (\alpha) := \inf_{(f,g)} \{ D_\beta (f, g) : N_\beta (f, g) \leq \alpha \}$

$\mathcal{C}^*_\beta$ is cts, inc, and concave

Estimation under communication constraints—(Mahajan and Chakravorty)
Optimization problems

Costly communication

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

Constrained communication

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

$C^*_\beta$ is cts, inc, and concave
Optimization problems

Costly communication
For $\lambda \in \mathbb{R}_{>0}$, $C^*_\beta(\lambda) = C(\lambda) = \inf_{(f, g)} \{D(\lambda) + \lambda N_\beta(f, g)\}$

Constrained communication
For $\alpha \in (0, 1)$, $D^*_\beta(\alpha) := \inf_{(f, g)} \{D(\lambda) : \lambda N_\beta(f, g) \leq \alpha\}$

$C^*_\beta$ is cts, inc, and concave

$D^*_\beta$ is cts, dec, and convex
Optimization problems

Costly communication

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

Constrained communication

For $\alpha \in (0, 1)$, 
\[ D^*_\beta (\alpha) := \inf_{(f, g)} \{D_\beta (f, g) : N_\beta (f, g) \leq \alpha\} \]

$C^*_\beta$ is cts, inc, and concave

$D^*_\beta$ is cts, dec, and convex

Distortion-transmission function

Estimation under communication constraints—(Mahajan and Chakravorty)
Optimization problems

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

Constrained communication
For $\alpha \in (0, 1)$, 
\[ D^*_\beta(\alpha) := \inf_{(f, g)} \{ D_\beta(f, g) : N_\beta(f, g) \leq \alpha \} \]

Distortion-transmission function
We provide explicit computable expressions for both curves.
\( X_{t+1} = X_t + W_t, \quad W_t \sim \mathcal{N}(0, 1) \)
\( X_{t+1} = X_t + W_t, \quad W_t \sim \mathcal{N}(0, 1) \)
Periodic transmission strategy
Periodic transmission strategy

Error process
Periodic transmission strategy

Error process

\[ D = 0.69 \quad N \approx 1/3 \]
An alternative strategy
An alternative strategy

Error process
An alternative strategy

Error process

\[ D = 0.24 \quad N \approx 1/3 \]
Distortion-transmission function

Periodic transmission strategy
Threshold based strategy

Estimation under communication constraints–(Mahajan and Chakravorty)
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 Gaussian 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 Gaussian 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
Dealing with non-classical information structure

Classical info. struct.
Dealing with non-classical information structure

Classical info. struct.

\[ f_t : X_t, Y_{1:t-1} \rightarrow U_t \]

\[ g_t : Y_{1:t-1}, Y_t \rightarrow \hat{X}_t \]
Dealing with non-classical information structure

Non-Classical info. struct.

\[ f_t \quad X_t, Y_{1:t-1} \quad u_t \]

\[ g_t \quad Y_{1:t-1}, Y_t \quad \hat{X}_t \]
Dealing with non-classical information structure

Non-Classical info. struct.

\[ f_t \quad X_t, Y_{1:t-1} \quad u_t \]

\[ g_t \quad Y_{1:t-1}, Y_t \quad \hat{X}_t \]
Dealing with non-classical information structure

Belongs to the class of **tractable non-classical information structures** (called **partial-history sharing**) identified in [Mahajan–Nayyar–Teneketzis 2013]

Common info $C_t := \bigcap_{s \geq t} \bigcap_{i=1}^{n} I_s^i$

Local info $L_t^i := I_t^i \setminus C_t$

$g(C, L) = \psi(C)(L)$

---


Estimation under communication constraints–(Mahajan and Chakravorty)
Dealing with non-classical information structure

Belongs to the class of tractable non-classical information structures (called partial-history sharing) identified in [Mahajan-Nayyar-Teneketzis 2013]

\begin{align*}
\text{Common info } & C_t := \bigcap_{s \geq t} \bigcap_{i=1}^{n} I^i_s \\
\text{Local info } & L^i_t := I^i_t \setminus C_t \\
g(C, L) = \psi(C)(L)
\end{align*}

\[ f_t \quad X_t, Y_{1:t-1} \quad U_t \]

\[ g_t \quad Y_{1:t-1}, Y_t \quad \hat{X}_t \]

\[ h_t \quad Y_{1:t-1} \quad (\varphi_t, Y_t) \]

\[ \varphi_t \quad X_t \quad U_t \]

\[ \gamma_t \quad Y_t \quad \hat{X}_t \]

\[ \text{Equiv.} \]

Dealing with non-classical information structure

The coordinated system is a centralized (i.e., single-agent) partially observed system

Belongs to the class of tractable non-classical information structures (called partial-history sharing) identified in [Mahajan-Nayyar-Teneketzis 2013]

$\Psi = \cap \Psi \cap \bigcap_{i=\mathcal{I}} I_i$

Local info $L_i \triangleq I_i \Psi \subseteq C \Psi$

$\Psi(C, L) = \psi(C)(L)$

The coordinated system is a centralized (i.e., single-agent) partially observed system

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

$\varphi_t \triangleq X_t, U_t$

$Y_{1:t-1}, (\varphi_t, Y_t)$

$\hat{X}_t$ is a centralized (i.e., single-agent) partially observed system

$Y_t, \hat{X}_t$

$\gamma_t \triangleq Y_t, \hat{X}_t$


Estimation under communication constraints–(Mahajan and Chakravorty)
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

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

Estimation under communication constraints–(Mahajan and Chakravorty)
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 \( t = T, \ldots, 0 \)
\[
V_t(\xi) = \min_{\hat{x} \in \mathcal{X}} \mathbb{E}[d(X_t - \hat{x}) + W_{t+1}(\Pi_{t+1}) \mid \Xi_t = \xi],
\]
\[
W_t(\pi) = \min_{\varphi: \mathcal{X} \rightarrow \{0, 1\}} \mathbb{E}[\lambda \varphi(X_t) + V_t(\Xi_t) \mid \Pi_t = \pi, \varphi_t = \varphi].
\]

Estimation under communication constraints–(Mahajan and Chakravorty)
Information states and dynamic program

**Information states**

- **Pre-transmission belief**: 
  \[ \Pi_{\psi}(x) = \mathbb{P}(X_{\psi} = x | Y_{k:}\psi - k) \]

- **Post-transmission belief**: 
  \[ \Xi_{\psi}(x) = \mathbb{P}(X_{\psi} = x | Y_{k:}\psi) \]

**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 \( t = T, \ldots, 0 \)

\[ V_{t}(\xi) = \min_{\hat{x} \in X} \mathbb{E}[d(X_{t} - \hat{x}) + W_{t+1}(\Pi_{t+1}) | \Xi_{t} = \xi], \]

\[ W_{t}(\pi) = \min_{\varphi: X \rightarrow \{0, 1\}} \mathbb{E}[\lambda \varphi(X_{t}) + V_{t}(\Xi_{t}) | \Pi_{t} = \pi, \varphi_{t} = \varphi]. \]

"Standard" POMDP. Optimal strategies can be computed numerically.
Can we use the DP to say something more about the optimal strategy?
Simplifying modeling assumptions

Markov process \[ X_{t+1} = X_t + W_t \]
# Simplifying modeling assumptions

**Markov process**

\[ X_{t+1} = X_t + W_t \]

**Markov chain setup**

State spaces \( X_t, W_t \in \mathbb{Z} \)

**Guass-Markov setup**

State spaces \( X_t, W_t \in \mathbb{R} \)
## Simplifying modeling assumptions

<table>
<thead>
<tr>
<th>Markov process</th>
<th>$X_{t+1} = X_t + W_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markov chain setup</strong></td>
<td>$X_t, W_t \in \mathbb{Z}$</td>
</tr>
<tr>
<td>Noise distribution</td>
<td>Unimodal and symmetric</td>
</tr>
<tr>
<td>$p_e = p_{-e} \geq p_{e+1}$</td>
<td>Zero-mean Gaussian</td>
</tr>
<tr>
<td><strong>Guass-Markov setup</strong></td>
<td>$X_t, W_t \in \mathbb{R}$</td>
</tr>
</tbody>
</table>

Estimation under communication constraints—(Mahajan and Chakravorty)
## Simplifying modeling assumptions

<table>
<thead>
<tr>
<th>Markov process</th>
<th>$X_{t+1} = X_t + W_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markov chain setup</strong></td>
<td></td>
</tr>
<tr>
<td>State spaces</td>
<td>$X_t, W_t \in \mathbb{Z}$</td>
</tr>
<tr>
<td>Noise distribution</td>
<td>Unimodal and symmetric</td>
</tr>
<tr>
<td>$p_e = p_{-e} \geq p_{e+1}$</td>
<td></td>
</tr>
<tr>
<td>Distortion</td>
<td>Even and increasing $d(e) = d(-e) \leq d(e+1)$</td>
</tr>
<tr>
<td><strong>Guass-Markov setup</strong></td>
<td></td>
</tr>
<tr>
<td>State spaces</td>
<td>$X_t, W_t \in \mathbb{R}$</td>
</tr>
<tr>
<td>Noise distribution</td>
<td>Zero-mean Gaussian $\varphi_\sigma(\cdot)$</td>
</tr>
<tr>
<td>Distortion</td>
<td>Mean-squared $d(e) =</td>
</tr>
</tbody>
</table>

![Unimodal and symmetric distribution](image1)

![Even and increasing distortion](image2)

Estimation under communication constraints—(Mahajan and Chakravorty)
Step 1: Structure of optimal strategies

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

Step 3: Optimal costly comm.

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  Optimal costly comm.

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  Optimal costly comm.

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  Optimal costly comm.

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** Optimal costly comm.

**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** Optimal costly comm.

**Step 4** Distortion-transmission trade-off

\[
D_\beta^{(k+2)} \quad D_\beta^{(k+1)} \quad D_\beta^{(k)}
\]

\[
\lambda^{(k)} \quad \lambda^{(k+1)}
\]

\[
D_\beta^*(\alpha) \quad (N_\beta^{(k)}, D_\beta^{(k)}) \quad (N_\beta^{(k+1)}, D_\beta^{(k+1)})
\]

\[
0 \quad \alpha \quad \alpha_c \quad 1
\]
Almost uniform and unimodal (ASU) distribution about $\alpha$

$\pi_\alpha \geq \pi_{\alpha+1} \geq \pi_{\alpha-1} \geq \pi_{\alpha+2} \geq \cdots$
Preliminaries

Almost uniform and unimodal (ASU) distribution about $\alpha$

$\pi_{a} \geq \pi_{a+1} \geq \pi_{a-1} \geq \pi_{a+2} \geq \cdots$

ASU Rearrangement

Estimation under communication constraints–(Mahajan and Chakravorty)
Preliminaries

Almost uniform and unimodal (ASU) distribution about $\alpha$

$$\pi_\alpha \geq \pi_{\alpha+1} \geq \pi_{\alpha-1} \geq \pi_{\alpha+2} \geq \cdots$$

ASU Rearrangement

Majorization

$$\pi \succeq \xi \text{ iff }$$

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

Invariant to permutations.

Estimation under communication constraints–(Mahajan and Chakravorty)

[LM11, NBTV13]
Step 1 Properties of the value functions [LM11, NBTV13]

Definition $\xi \triangleright \tilde{\xi}$ if $\xi \geq \tilde{\xi}$ and $\tilde{\xi}$ is ASU about some point $a$.
Properties of the value functions

Definition

\( \xi \succ \tilde{\xi} \) if \( \xi \geq \tilde{\xi} \) and \( \tilde{\xi} \) is ASU about some point \( a \).

Lemma

Similar to Schur-concavity

- If \( \xi \succ \tilde{\xi} \) then \( W_t(\xi) \geq W_t(\tilde{\xi}) \).
- If \( \pi \succ \tilde{\pi} \) then \( V_t(\pi) \geq V_t(\tilde{\pi}) \).
**Step 1** Properties of the value functions

Definition

\( \xi \triangleright \tilde{\xi} \) if \( \xi \succeq \tilde{\xi} \) and \( \tilde{\xi} \) is ASU about some point \( a \)

Lemma

- If \( \xi \triangleright \tilde{\xi} \) then \( W_t(\xi) \geq W_t(\tilde{\xi}) \).
- If \( \pi \triangleright \tilde{\pi} \) then \( V_t(\pi) \geq V_t(\tilde{\pi}) \).

Lemma (Arg min of \( W \))

If \( \xi \) is ASU about \( a \) then \( a \) is the arg min of

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

[LM11, NBTV13]

Similar to Schur-concavity
Step 1 Properties of the value functions [LM11, NBTV13]

Definition
\[ \xi \succ \tilde{\xi}, \text{if } \xi \succeq \tilde{\xi}, \text{and } \tilde{\xi} \text{ is ASU about some point } a \]

Lemma
- If \( \xi \succ \tilde{\xi} \) then \( W_t(\xi) \geq W_t(\tilde{\xi}) \).
- If \( \pi \succ \tilde{\pi} \) then \( V_t(\pi) \geq V_t(\tilde{\pi}) \).

Lemma (Arg min of \( W \))
If \( \xi \) is ASU about \( a \) then \( a \) is the arg min of
\[
V_t(\xi) = \min_{\hat{\chi} \in \hat{X}} \mathbb{E}[d(X_t - \hat{\chi}) + W_{t+1}(\Pi_{t+1}) | \Xi_t = \xi],
\]

Lemma (Arg min of \( V \))
If \( \pi \) is ASU about \( a \) then the arg min of
\[
W_t(\pi) = \min_{\varphi: X \to \{0, 1\}} \mathbb{E}[\lambda \varphi(X_t) + V_t(\Xi_t) | \Pi_t = \pi, \varphi_t = \varphi]
\]
is of the form
\[
\varphi(x) = \begin{cases} 
1, & \text{if } |x - a| > k(\pi) \\
0, & \text{if } |x - a| < k(\pi) \\
q_+, & \text{if } x - a = k(\pi) \\
q_-, & \text{if } x - a = -k(\pi)
\end{cases}
\]
Step 1 Structure of optimal strategies

\[ \pi_1 \text{ is ASU about } z_0 \]
Step 1  Structure of optimal strategies

\[ \pi_1 \text{ is ASU about } z_0 \]

Is \( |x_1 - z_0| > k_1 \)?
Step 1 Structure of optimal strategies

\[ z_0 \]

\( \pi_1 \) is ASU about \( z_0 \)

Is \( |x_1 - z_0| > k_1 ? \)

\textbf{NO.} \( u_1 = \epsilon, z_1 = z_0 \)

\[ z_1 \]

\( \xi_1 \) is ASU about \( z_1 \)

Estimation under communication constraints–(Mahajan and Chakravorty)
**Step 1** Structure of optimal strategies

\[ \pi_1 \text{ is ASU about } z_0 \]

Is \(|x_1 - z_0| > k_1\)?

**YES.** \(u_1 = 1, z_1 = x_1\)

**NO.** \(u_1 = \varepsilon, z_1 = z_0\)
Step 1 Structure of optimal strategies

π₁ is ASU about z₀

Is \(|x₁ - z₀| > k₁\)?

**YES.** \(u₁ = 1, z₁ = x₁\)  
**NO.** \(u₁ = \varepsilon, z₁ = z₀\)

ξ₁ is ASU about z₁

In both cases: \(\hat{x}_1 = z₁\)
Step 1 Structure of optimal strategies

\[ z_0 \]

\( \pi_1 \) is ASU about \( z_0 \)

Is \( |x_1 - z_0| > k_1 \)？

YES. \( u_1 = 1, z_1 = x_1 \)

NO. \( u_1 = \varepsilon, z_1 = z_0 \)

\( \xi_1 \) is ASU about \( z_1 \)

\( \hat{x}_1 = z_1 \)

In both cases: \( \hat{x}_1 = z_1 \)

Estimation under communication constraints–(Mahajan and Chakravorty)
Step 1 Structure of optimal strategies

\[ t = 2 \]

\[ X_2 = X_1 + W_1 \implies \pi_1 = \xi_1 \ast p \]

\( \pi_1 \) is ASU about \( z_0 \)

\[ |x_1 - z_0| > k_1 \]

**YES.** \( u_1 = 1, z_1 = x_1 \)

**NO.** \( u_1 = \epsilon, z_1 = z_0 \)

\( \xi_1 \) is ASU about \( z_1 \)

In both cases: \( \hat{x}_1 = z_1 \)

Estimation under communication constraints—(Mahajan and Chakravorty)
Step 1 Structure of optimal estimator

Transmitted Process: Let $Z_t$ denote the most recently transmitted value of the Markov process.
Step 1 Structure of optimal estimator

Transmitted Process

Let $Z_t$ denote the most recently transmitted value of the Markov process.

Lemma

$\Xi_t$ is ASU about $Z_t$
Step 1  **Structure of optimal estimator**  [LM11, NBTV13]

**Transmitted Process**

Let $Z_t$ denote the most recently transmitted value of the Markov process.

**Lemma**

$\Xi_t$ is ASU about $Z_t$

**Theorem**

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

**Remark**

The optimal estimation strategy is *time-homogeneous* and can be specified in closed form.
Step 1 Structure of optimal transmitter

Lemma $\Pi_t$ is ASU about $Z_{t-1}$

$Z_{t-1}$
Step 1  Structure of optimal transmitter  

**Lemma**

$\Pi_t$ is ASU about $Z_{t-1}$

**Theorem**

$$U_t = f_t(X_t, \Pi_t) = \begin{cases} 1, & \text{if } |X_t - E_t| \geq k_t \\ 0, & \text{if } |X_t - E_t| < k_t \end{cases}$$

[LM11, NBTV13]
Step 1 Structure of optimal transmitter [LM11, NBTV13]

Lemma \( \Pi_t \) is ASU about \( Z_{t-1} \)

Theorem \( U_t = f_t(X_t, \Pi_t) = \begin{cases} 1, & \text{if } |X_t - E_t| \geq k_t \\ 0, & \text{if } |X_t - E_t| < k_t \end{cases} \)

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

Remark The optimal transmission strategy is a function of the error process.
The results extend to infinite horizon setup under appropriate regularity conditions.

Time-homogeneous threshold-based strategies are optimal.
How do we find the optimal threshold-based strategy?
Step 1: Structure of optimal strategies

Search space of strategies \((f, g)\)

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

Step 3: Optimal costly comm.

Step 4: Distortion-transmission trade-off
### Step 2 Performance of threshold strategies

Consider a **threshold-based** strategy

\[
 f^{(k)}(e) = \begin{cases} 
 1 & \text{if } |e| \geq k \\
 0 & \text{otherwise}
\end{cases}
\]
Step 2 Performance of threshold strategies

Consider a threshold-based strategy

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

Let \( \tau^{(k)} \) denote the stopping time of first transmission (starting at \( E_0 = 0 \)).
Consider a **threshold-based strategy**

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

**Define**

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

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

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

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

Define

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

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

Proposition

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

$$D^{(k)}_\beta := D_\beta(f^{(k)}, g^*) = \frac{L^{(k)}_\beta(0)}{M^{(k)}_\beta(0)}$$

and

$$N^{(k)}_\beta := N_\beta(f^{(k)}, g^*) = \frac{1}{M^{(k)}_\beta(0)} - (1 - \beta).$$
**Step 2 Performance of threshold strategies**

Consider a threshold-based strategy $f^{(k)}$. Let $\tau^{(k)}$ denote the stopping time of first transmission (starting at $E_0 = 0$).

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

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

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

**Proposition**

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

$$D^{(k)}_\beta := D_\beta(f^{(k)}, g^*) = \frac{L^{(k)}_\beta(0)}{M^{(k)}_\beta(0)}$$

and

$$N^{(k)}_\beta := N_\beta(f^{(k)}, g^*) = \frac{1}{M^{(k)}_\beta(0)} - (1 - \beta).$$
Step 2  Computing $L^{(k)}_{\beta}$ and $M^{(k)}_{\beta}$

Markov chain setup

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

$$M^{(k)}_{\beta}(e) = 1 + \beta \sum_{n=-k}^{k} p_{n-e} M^{(k)}_{\beta}(n)$$
Step 2 Computing $L_{\beta}^{(k)}$ and $M_{\beta}^{(k)}$

Markov chain 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

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

$$M_{\beta}^{(k)} = \left[ [I - \beta P^{(k)}]^{-1} 1^{(k)} \right].$$
**Step 2** Computing $L^{(k)}_{\beta}$ and $M^{(k)}_{\beta}$

Markov chain setup

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

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

Proposition

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

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

Gauss-Markov setup

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

$$M^{(k)}_{\beta}(e) = 1 + \beta \int_{-k}^{k} \varphi(n-e) M^{(k)}_{\beta}(n) \, dn$$
Step 2  Computing $L^{(k)}_\beta$ and $M^{(k)}_\beta$

Markov chain setup

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

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

Proposition

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

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

Gauss-Markov setup

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

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

Fredholm Integral Equations of the 2nd kind.

Solutions exist and are unique.
**Step 2  Computing $L^{(k)}_{\beta}$ and $M^{(k)}_{\beta}$**

$D^{(k)}_{\beta}$ and $N^{(k)}_{\beta}$ can be computed using these expressions.

**Proposition**

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

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

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

**Gauss-Markov setup**

\[
L^{(k)}_{\beta}(e) = d(e) + \beta \int_{-k}^{k} \varphi(n - e) L^{(k)}_{\beta}(n) dn
\]

\[
M^{(k)}_{\beta}(e) = 1 + \beta \int_{-k}^{k} \varphi(n - e) M^{(k)}_{\beta}(n) dn
\]

Fredholm Integral Equations of the 2nd kind.
Solutions exist and are unique.
We found the performance of a generic threshold-based strategy. How does this lead to identifying an optimal strategy?
**Step 1**  Structure of optimal strategies

Search space of strategies \((f, g)\)

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

**Step 3**  Optimal costly comm.

**Step 4**  Distortion-transmission trade-off
Properties of optimal thresholds

Monotonicity

\[ L^{(k+1)}_{\beta} > L^{(k)}_{\beta} \quad \text{and} \quad M^{(k+1)}_{\beta} > M^{(k)}_{\beta} \]

Depends on unimodularity of noise
Step 3  Properties of optimal thresholds

**Monotonicity**

\[ L^{(k+1)}_{\beta} > L^{(k)}_{\beta} \quad \text{and} \quad M^{(k+1)}_{\beta} > M^{(k)}_{\beta} \]

**Implication:**

\[ D^{(k+1)}_{\beta} \geq D^{(k)}_{\beta} \quad \text{and} \quad N^{(k+1)}_{\beta} < N^{(k)}_{\beta} \]

Use DP and monotonicity of Bellman operator
Step 3 Properties of optimal thresholds

Monotonicity

\[ L^{(k+1)}_{\beta} > L^{(k)}_{\beta} \text{ and } M^{(k+1)}_{\beta} > M^{(k)}_{\beta} \]

Implication:

\[ D^{(k+1)}_{\beta} \geq D^{(k)}_{\beta} \text{ and } N^{(k+1)}_{\beta} < N^{(k)}_{\beta} \]

Submodularity

\[ C^{(k)}_{\beta}(\lambda) := D^{(k)}_{\beta} + \lambda N^{(k)}_{\beta} \text{ is submodular in } (k, \lambda). \]
Step 3 Properties of optimal thresholds

Monotonicity

\[ L^{(k+1)}_β > L^{(k)}_β \quad \text{and} \quad M^{(k+1)}_β > M^{(k)}_β \]

Implication:

\[ D^{(k+1)}_β \geq D^{(k)}_β \quad \text{and} \quad N^{(k+1)}_β < N^{(k)}_β \]

Submodularity

\[ C^{(k)}_β (λ) := D^{(k)}_β + λN^{(k)}_β \text{ is submodular in } (k, λ). \]

Proposition

\[ k^*_β (λ) := \arg \min_{k\in\mathbb{Z}_{\geq 0}} C^{(k)}_β (λ) \text{ is increasing in } λ. \]
### Step 3 Properties of optimal thresholds

**Monotonicity**

\[ L^{(k+1)}_\beta > L^{(k)}_\beta \quad \text{and} \quad M^{(k+1)}_\beta > M^{(k)}_\beta \]

**Implication:**

\[ D^{(k+1)}_\beta \geq D^{(k)}_\beta \quad \text{and} \quad N^{(k+1)}_\beta < N^{(k)}_\beta \]

**Submodularity**

\[ C^{(k)}_\beta (\lambda) := D^{(k)}_\beta + \lambda N^{(k)}_\beta \] is submodular in \((k, \lambda)\).

**Proposition**

\[ k^*_\beta (\lambda) := \arg \min_{k \in \mathbb{Z}_{\geq 0}} C^{(k)}_\beta (\lambda) \] is increasing in \(\lambda\).

Thus, optimal threshold increases with increase in \(\lambda\).
Characterizing the optimal threshold for a given communication cost is tricky.

Instead, we will characterize the optimal communication cost for a given threshold.
Define $\Lambda^{(k)}_\beta := \{ \lambda \in \mathbb{R}_{\geq 0} : k^*_\beta (\lambda) = k \} = [\lambda^{(k-1)}_\beta, \lambda^{(k)}_\beta]$.

$C^{(k)}_\beta (\lambda^{(k)}_\beta ) = C^{(k+1)}_\beta (\lambda^{(k)}_\beta )$
Define $\Lambda^{(k)}_\beta := \{ \lambda \in \mathbb{R}_{\geq 0} : k^*_\beta (\lambda) = k \}$

$= [\lambda^{(k-1)}_\beta, \lambda^{(k)}_\beta]$.

$C^{(k)}_\beta (\lambda^{(k)}_\beta) = C^{(k+1)}_\beta (\lambda^{(k)}_\beta)$
Step 3  Optimal costly communication: Markov chain

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

$$C_{\beta}^{(k)} (\lambda_{\beta}^{(k)}) = C_{\beta}^{(k+1)} (\lambda_{\beta}^{(k)})$$
Step 3  Optimal costly communication: Markov chain

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

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

Estimation under communication constraints–(Mahajan and Chakravorty)
Define $\Lambda^{(k)}_\beta := \{\lambda \in \mathbb{R}_{\geq 0} : k^*_\beta(\lambda) = k\}$

$= [\lambda^{(k-1)}_\beta, \lambda^{(k)}_\beta]$.

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

$C^{(k)}(\lambda^{(k)}_\beta) = C^{(k+1)}(\lambda^{(k)}_\beta)$
Step 3
Optimal costly communication: Markov chain

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

$$C_{\beta}^*(\lambda) = \min_{k \in \mathbb{Z}_{\geq 0}} C_{\beta}^{(k)}$$

is piecewise linear, continuous, concave, and increasing function of $\lambda$. 

Estimation under communication constraints–(Mahajan and Chakravorty)
Step 3: Optimal costly communication: Markov chain

\[ \lambda^{(k)}_\beta = \frac{D^{(k+1)}_\beta - D^{(k)}_\beta}{N^{(k)}_\beta - N^{(k+1)}_\beta} \]

Theorem

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

\[ C^*_\beta(\lambda) = \min_{k \in \mathbb{Z}_{\geq 0}} C^{(k)}_\beta \]

is piecewise linear, continuous, concave, and increasing function of \( \lambda \).
Example
Symmetric birth-death Markov chain ($p = 0.3$)

\[ C^*_\beta(\lambda) \]

\[ \beta = 1.0 \]
\[ \beta = 0.95 \]
\[ \beta = 0.9 \]
Step 3  Optimal costly communication: Gauss-Markov

Lemma

$D^{(k)}_{\beta}$ is increasing in $k$ and $N^{(k)}_{\beta}$ is decreasing in $k$.

$D^{(k)}_{\beta}$ and $N^{(k)}_{\beta}$ are differentiable in $k$. 
Step 3 Optimal costly communication: Gauss-Markov

Lemma

- $D^{(k)}_\beta$ is increasing in $k$ and $N^{(k)}_\beta$ is decreasing in $k$.
- $D^{(k)}_\beta$ and $N^{(k)}_\beta$ are differentiable in $k$.

Theorem

- Strategy $f^{(k)}$ is optimal for $\lambda = -\frac{\partial_k D^{(k)}_\beta}{\partial_k N^{(k)}_\beta}$

$C^*_\beta(\lambda) = \min_{k \in \mathbb{R} > 0} C^{(k)}_\beta$ is continuous, concave, and increasing function of $\lambda$. 
Optimal costly communication: Gauss-Markov

**Lemma**

$D^{(k)}_{\beta}$ is increasing in $k$ and $N^{(k)}_{\beta}$ is decreasing in $k$.

$D^{(k)}_{\beta}$ and $N^{(k)}_{\beta}$ are differentiable in $k$.

**Theorem**

Strategy $f^{(k)}$ is optimal for $\lambda = -\frac{\partial_k D^{(k)}_{\beta}}{\partial_k N^{(k)}_{\beta}}$

$C^*_\beta (\lambda) = \min_{k \in \mathbb{R}_{\geq 0}} C^{(k)}_{\beta}$ is continuous, concave, and increasing function of $\lambda$.

**Scaling with variance $\sigma^2$**

$C^*_{\beta, \sigma}(\lambda) = \sigma^2 C^*_{\beta, 1} \left( \frac{\lambda}{\sigma^2} \right)$
Step 3  Optimal costly communication: Gauss-Markov

Lemma  \( D_\beta^{(k)} \) is increasing in \( k \) and \( N_\beta^{(k)} \) is decreasing in \( k \).

\( D_\beta^{(k)} \) and \( N_\beta^{(k)} \) are differentiable in \( k \).

Theorem  Strategy \( f^{(k)} \) is optimal for \( \lambda = -\frac{\partial_k D_\beta^{(k)}}{\partial_k N_\beta^{(k)}} \)

\( C_\beta^*(\lambda) = \min_{k \in \mathbb{R}_{\geq 0}} C_\beta^{(k)} \) is continuous, concave, and increasing function of \( \lambda \).

Scaling with variance \( \sigma^2 \)

\( C_{\beta, \sigma}^*(\lambda) = \sigma^2 C_{\beta, 1}^* \left( \frac{\lambda}{\sigma^2} \right) \)

Computation  Use bisection search to find \( k \) such that \( \lambda = -\frac{\partial_k D_\beta^{(k)}}{\partial_k N_\beta^{(k)}} \)
Example Gauss-Markov with $\sigma^2 = 1$
Step 1  Structure of optimal strategies

Search space of strategies \((f, g)\)

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

Step 3  Optimal costly comm.

Step 4  Distortion-transmission trade-off
Sufficient conditions for constrained optimality
A strategy \((f^\circ, g^\circ)\) is optimal for the constrained communication problem if

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

\[(C2) \text{ There exists } \lambda^\circ \geq 0 \text{ such that } (f^\circ, g^\circ) \text{ is optimal for } C_\beta(f, g; \lambda^\circ).\]
Step 4 Distortion-transmission trade-off: Markov chain

Sufficient conditions for constrained optimality

A strategy \((f^\circ, g^\circ)\) is optimal for the constrained communication problem if

\[(C1) \quad N_\beta(f^\circ, g^\circ) = \alpha\]

\[(C2) \quad \text{There exists } \lambda^\circ \geq 0 \text{ such that } (f^\circ, g^\circ) \text{ is optimal for } C_\beta(f, g; \lambda^\circ).\]

Let \(k^\ast_\beta\) be such that

\[N^{(k^\ast_\beta)}_\beta > \alpha > N^{(k^\ast_\beta + 1)}_\beta\]
Sufficient conditions for constrained optimality

A strategy \((f^o, g^o)\) is optimal for the constrained communication problem if

\[(C1) \quad N_\beta(f^o, g^o) = \alpha\]

\[(C2) \quad \text{There exists } \lambda^o \geq 0 \text{ such that } (f^o, g^o) \text{ is optimal for } C_\beta(f, g; \lambda^o).\]

Let \(k^*_\beta\) be such that

\[N_\beta^{(k^*_\beta)} > \alpha > N_\beta^{(k^*_\beta + 1)}\]
Step 4  Distortion-transmission trade-off: Markov chain

Sufficient conditions for constrained optimality

A strategy \((f^o, g^o)\) is optimal for the constrained communication problem if

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

(C2) There exists \(\lambda^o \geq 0\) such that \((f^o, g^o)\) is optimal for \(C_\beta(f, g; \lambda^o)\).

Let \(k_\beta^*\) be such that
\[N_\beta^{(k_\beta^*)} > \alpha > N_\beta^{(k_\beta^* + 1)}\]
**Step 4**  
**Distortion-transmission trade-off: Markov chain**

**Sufficient conditions for constrained optimality**  
A strategy \((f^\circ, g^\circ)\) is optimal for the constrained communication problem if

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

\[(C2) \text{ There exists } \lambda^\circ \geq 0 \text{ such that } (f^\circ, g^\circ) \text{ is optimal for } C_\beta(f, g; \lambda^\circ).\]

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

\[\theta^* N_\beta^{(k)} + (1 - \theta^*) N_\beta^{(k+1)} = \alpha\]
Let $k^*$ be such that $N_{\beta}^{k^*} \beta \geq \alpha > N_{\beta}^{k^*} \beta$. 

**Step 4** Distortion-transmission trade-off: Markov chain 

Sufficient conditions for constrained optimality 

A strategy $(f(\cdot), g(\cdot))$ is optimal for the constrained communication problem if 

1. $N(\beta^*, f(\cdot), g(\cdot)) = \alpha$ 
2. There exists $\lambda(\cdot)$ such that $(f(\cdot), g(\cdot))$ is optimal for $C(\beta, f, g; \lambda(\cdot))$.

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: Markov chain

Sufficient conditions for constrained optimality

A strategy \((f \circ \lambda, g \circ \lambda)\) is optimal for the constrained communication problem if

(C1) \(N_{\beta}(f \circ \lambda, g \circ \lambda) = \alpha\)

(C2) There exists \(\lambda \circ \lambda_0\) such that \((f \circ \lambda, g \circ \lambda)\) is optimal for \(C_{\beta}(f, g; \lambda \circ \lambda_0)\).

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

\[
\theta^* N^{(k)}_{\beta} + (1 - \theta^*) N^{(k+1)}_{\beta} = \alpha
\]
Example: Symmetric birth-death Markov chain ($p = 0.3$)

Symmetric birth-death Markov chain ($\alpha = \lambda \beta$)

Parameter values:
- $\alpha = 0.9$
- $\alpha = 0.95$
- $\alpha = 1.0$

Graph showing the function $D^*_\beta(\alpha)$ with different values of $\beta$. The graph illustrates how the function changes with varying $\alpha$ and $\beta$. The y-axis represents $D^*_\beta(\alpha)$, and the x-axis represents $\alpha$. The curves for different $\beta$ values are shown, with $\beta = 0.9$, $\beta = 0.95$, and $\beta = 1.0$. The graph highlights the impact of $\beta$ on the function's behavior.
Step 4: Distortion-transmission trade-off: Gauss-Markov

Sufficient conditions for constrained optimality

A strategy \((f^\circ, g^\circ)\) is optimal for the constrained communication problem if

\[(C1) \quad N_\beta(f^\circ, g^\circ) = \alpha\]

\[(C2) \quad \text{There exists } \lambda^\circ \geq 0 \text{ such that } (f^\circ, g^\circ) \text{ is optimal for } C_\beta(f, g; \lambda^\circ).\]
**Step 4**\[\text{Distortion-transmission trade-off: Gauss-Markov}\]

**Sufficient conditions for constrained optimality**

A strategy \((f^\circ, g^\circ)\) is optimal for the constrained communication problem if

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

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

**Theorem**

There exists a \(k^*_\beta(\alpha)\) such that \(N_\beta^{(k^*_\beta(\alpha))} = \alpha\). Therefore,

\[D^*_\beta(\alpha) = D_\beta^{(k^*_\beta(\alpha))}\]
Sufficient conditions for constrained optimality

A strategy \((f^\circ, g^\circ)\) is optimal for the constrained communication problem if

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

\[(C2) \text{ There exists } \lambda^\circ \geq 0 \text{ such that } (f^\circ, g^\circ) \text{ is optimal for } C_\beta(f, g; \lambda^\circ). \]

**Theorem**

There exists a \(k^*_\beta(\alpha)\) such that \(N_{\beta}^{(k^*_\beta(\alpha))} = \alpha\). Therefore,

\[D^*_\beta(\alpha) = D^{(k^*_\beta(\alpha))}_\beta\]

**Scaling with variance \(\sigma^2\)**

\[D^*_{\beta, \sigma}(\alpha) = \sigma^2 D^*_{\beta, 1}(\alpha).\]
### Step 4: Distortion-transmission trade-off: Gauss-Markov

**Sufficient conditions for constrained optimality**

A strategy \((f^o, g^o)\) is optimal for the constrained communication problem if

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

(C2) There exists \(\lambda^o \geq 0\) such that \((f^o, g^o)\) is optimal for \(C_\beta (f, g; \lambda^o)\).

**Theorem**

There exists a \(k^*_\beta (\alpha)\) such that \(N_{\beta}^{(k^*_\beta (\alpha))} = \alpha\). Therefore,

\[
D_\beta^* (\alpha) = D_{\beta}^{(k^*_\beta (\alpha))}
\]

**Scaling with variance \(\sigma^2\)**

\[
D_{\beta, \sigma}^* (\alpha) = \sigma^2 D_{\beta, 1}^* (\alpha).
\]

**Computation**

Use bisection search to find \(k\) such that \(N_{\beta}^{(k)} = \alpha\).
Example Gauss-Markov with $\sigma^2 = 1$

![Graph showing $D^*_\beta(\alpha)$ for different values of $\beta$.]
Conclusion

Analyze fundamental limits of estimation under communication constraints
Conclusion

Analyze fundamental limits of estimation under communication constraints

Possible generalizations to more realistic models

- Packet drops
- Rate constraints (effect of quantization)
- Network delays
Conclusion

Analyze fundamental limits of estimation under communication constraints

Possible generalizations to more realistic models
- Packet drops
- Rate constraints (effect of quantization)
- Network delays

A simple non-trivial “toy-problem” for decentralized control
- Decentralized control is full of difficult problems and negative results.
- It is important to identify “easy” problems and positive results.
Conclusion

Analyze fundamental limits of estimation under communication constraints

Possible generalizations to more realistic models
- Packet drops
- Rate constraints (effect of quantization)
- Network delays

A simple non-trivial “toy-problem” for decentralized control
- Decentralized control is full of difficult problems and negative results.
- It is important to identify “easy” problems and positive results.

The system model

\[ Y_t = \begin{cases} X_t, & \text{if } U_t = 1 \\ \xi_t, & \text{if } U_t = 0 \end{cases} \]

Distortion communication

\[ D_\beta(f,g) = (1-\beta) \mathbb{E}[\|f(X) - g(Y)\|^p] \]

Constrained communication

\[ D_\beta^*(\alpha) = \inf_{f,g} \{D_\beta(f,g) : N_\beta(f,g) \leq \alpha \} \]

Estimation under communication constraints–(Mahajan and Chakravorty)

Step 1 Structure of optimal strategies

\[ \pi_1 \text{ is ASU about } z_0 \]

YES. \( u_1 = 1, z_1 = 1 \)

NO. \( u_1 = 0, z_1 = 0 \)

etc. . .

Step 2 Performance of threshold strategies

Consider a threshold-based strategy

\[ \tau_k^f = \begin{cases} 1, & \text{if } |e| \geq k \\ 0, & \text{otherwise} \end{cases} \]

Define

\[ L_k^f(e) = (1-\beta) \mathbb{E} \left[ \sum_{t=0}^{\tau_k^f-1} \beta^t |E_t| \right] \]

\[ M_k^f(e) = (1-\beta) \mathbb{E} \left[ \sum_{t=0}^{\tau_k^f-1} \beta^t |E_t| \right] \]

Step 3 Optimal costly communication: Markov chain

\[ A_{\beta}^k = \frac{N_\beta^k - D_\beta^{(k+1)}}{N_\beta^k - D_\beta^{(k+1)}} \]

Sufficient conditions for constrained optimality

A strategy \( \theta^* \) is optimal if

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

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

\[ C_\beta^k | \lambda(k) = \alpha \]

Estimation under communication constraints–(Mahajan and Chakravorty)