Fundamental limits of distributed tracking

Victoria Kostina

Babak Hassibi

Caltech

ISIT 2020
Control

Causality

Information

Delays
Control

Separation between control and estimation

Causal Estimation (Tracking)
Distributed tracking

**Goal**  Design (encoders, decoder) that operate at the minimum sum rate $\sum_{k=1}^{K} R_k$ and achieve distortion

$$\frac{1}{t} \sum_{i=1}^{t} \mathbb{E} \left[ d(X_i, \hat{X}_i) \right] \leq d$$
Distributed tracking

- The source $X_1, X_2, \ldots$ has memory of the past.
- Encoder(s) and decoder have memory of the past.
- $\hat{X}_i$ is independent of the future given the past (causal enc / dec).
Two parts

1. General coding theorem.
2. Explicit fundamental tradeoff in the Gaussian case.
Class of sources

- no restrictions on the source alphabet
Class of observation channels

Acts on each component of the source vector independently, but may have memory of the past

\[ P_{Y_i^k | X_i, \text{past } X's, \text{past } Y^k's} = P_i^{\otimes n} \]

\[ i = 1, 2, \ldots, t \]
$t = 1$: Classical CEO problem

Chief Estimation Officer, or Chief Executive Officer

- Introduced (discrete-alphabet source) by Berger-Zhang-Viswanathan’96.
- Viswanathan-Berger’97: an achievability bound on the rate-distortion dimension of the symmetric Gaussian CEO problem.
- Oohama’98: sum rate - distortion region for the symmetric Gaussian CEO problem.
- Prabhakaran-Tse-Ramanchandran’04: full Gaussian CEO rate region.
- Chen-Zhang-Berger-Wicker’04: water-filling representation of the minimum sum rate.
- Chen-Berger’08, Behroozi-Soleymani’09: successive coding schemes.
- Wagner-Tavildar-Viswanath’08: found the rate region of the distributed Gaussian lossy compression problem by coupling it to the CEO problem.
- Wang-Chen-Wu’10: a simple converse on the sum rate.
- Courtade-Weissman’14: the rate-distortion regions of the distributed source coding and the CEO problem under logarithmic loss.
Sum rate - distortion function

Definition

An \((R, d, t)\) causal CEO code: \(K\) causal encoders operating at sum rate \(R\) and a causal decoder that achieve distortion \(d\) over time horizon \(t\).

Definition

The rate-distortion pair \((R, d)\) is achievable at time horizon \(t\) in the CEO problem if \(\forall \gamma > 0, \exists n_0 \in \mathbb{N}\) such that \(\forall n \geq n_0\), an \((R, d + \gamma, t)\) exists. The causal sum rate - distortion function is defined as follows:

\[
R_{\text{CEO}_t}(d) \triangleq \inf \left\{ R : (R, d) \text{ is achievable at time horizon } t \right\}
\]
Notation

- $X_{[t]} \triangleq (X_1, \ldots, X_t)$
- $U^k_{[t]} \triangleq (U^k_1, \ldots, U^k_t)$ (codewords by $k$-th encoder up to time $t$)
- $U^K_{[t]} \triangleq \begin{bmatrix} U^1_{[t]} \\ \vdots \\ U^K_{[t]} \end{bmatrix}$
- Massey’s directed information (’90)
  \[ I(X_{[t]} \to \hat{X}_{[t]}) \triangleq \sum_{i=1}^{t} I(X_i; \hat{X}_i | \hat{X}_{i-1}) \]
- Causally conditional (Kramer, ’98) probability kernel
  \[ P_{Y_{[t]} | X_{[t]}} \triangleq \prod_{i=1}^{t} P_{Y_i | Y_{i-1}, X_i} \]
Main coding theorem

\[ R_t \text{CEO}(d) = \inf_{P_{U[K][t]} \parallel Y[K][t]} \prod_{k=1}^{K} P_{U[k][t]} \parallel Y[k][t], \]

\[ I \left( Y[K][t] \rightarrow U[K][t] \right), \]

\[ \frac{1}{t} \sum_{i=1}^{t} \mathbb{E}[d(X_i, \hat{X}_i)] \leq d \]
Converse

- Data processing
- Chain rule
- Standard single-letterization argument
Achievability

Step \( i \):

\[
\begin{align*}
\text{ENC} & \quad \text{ENC} \\
\downarrow & \quad \downarrow & \quad \downarrow \\
U^k & \quad U^k \quad U^k 
\end{align*}
\]

Coded SI

need to consider steps \( \{1, 2, \ldots, t\} \) jointly

Structured multiterminal source coding problem with \( K \cdot t \) encoders and \( t \) decoders
Causal nonasymptotic Berger-Tung bound

- We make use of the achievability proof technique developed by Yassaee, Aref, Gohari, 2013.
- It uses a stochastic likelihood coder.
- The technique extends naturally to causal coding problems.
- We extend Yassaee et al. ’13 to $K > 2$ by proposing a novel decoder that falls into the class of generalized likelihood decoders (Merhav ’17).
Two parts

1. General coding theorem.
2. Explicit fundamental tradeoff in the Gaussian case.
Gauss-Markov source: a simple source with memory

\[ X_1, \{V_i\}_{i=1}^{\infty} \sim \mathcal{N}(0, \sigma^2 I_n) \text{ i.i.d.} \]

\[ X_{i+1} = aX_i + V_i \]

- \( a = 0 \): i.i.d. Gaussian source
- \( |a| < 1 \): asymptotically stationary source
- \( |a| \geq 1 \): nonstationary source
Gauss-Markov source: a simple source with memory

\[ X_{i+1} = aX_i + V_i \]


- V. Kostina and B. Hassibi, "Rate-cost tradeoffs in scalar LQG control and tracking with side information", *Allerton*, 2018.

Causal Gaussian CEO problem

\[ X_{i+1} = aX_i + V_i \]
\[ V_i \sim \mathcal{N}(0, \sigma_V^2 I) \]

\[ W_i^1 \sim \mathcal{N}(0, \sigma_{W_1}^2 I) \]
\[ Y_i^1 \]

\[ W_i^2 \sim \mathcal{N}(0, \sigma_{W_2}^2 I) \]
\[ Y_i^2 \]

\[ \ldots \]

\[ W_i^K \sim \mathcal{N}(0, \sigma_{W_K}^2 I) \]
\[ Y_i^K \]

\[ \hat{X}_i \]
Notation: causal MSE estimation

- \( \sigma^2_{X \parallel Y^k} \triangleq \frac{1}{n} \lim_{t \to \infty} \mathbb{E} \left[ \left| X_t - \mathbb{E} \left[ X_t \mid Y^k_t \right] \right|^2 \right] \)
  observations by \( k \)-th encoder up to time \( t \)

- \( \sigma^2_{X \parallel Y^K} \triangleq \frac{1}{n} \lim_{t \to \infty} \mathbb{E} \left[ \left( X_t - \mathbb{E} \left[ X_t \mid Y^K_t \right] \right)^2 \right] \)
  observations by all encoders up to time \( t \)
Separation of the causal CEO encoders

\[ 
Y_{i1} \xrightarrow{\text{ESTIMATOR 1}} \tilde{X}_{i1} \xrightarrow{\text{Q 1}} R_1 \\
Y_{i2} \xrightarrow{\text{ESTIMATOR 2}} \tilde{X}_{i2} \xrightarrow{\text{Q 2}} R_2 \\
\vdots \\
Y_{ik} \xrightarrow{\text{ESTIMATOR K}} \tilde{X}_{ik} \xrightarrow{\text{Q K}} R_K \\
\tilde{X}_i \xrightarrow{\text{DEC}} \hat{X}_i \\
\]

\[ \tilde{X}_i^k \triangleq \mathbb{E} \left[ X_i | Y_{i[i]}^k \right] \]
Gaussian setting: Main Theorem

The causal sum rate - distortion function is given by

$$\lim_{t \to \infty} R_{CEO_t}(d) = \frac{1}{2} \log \frac{\bar{d}}{d} + \min_{\{d_k\}_{k=1}^{K}} \sum_{k=1}^{K} \frac{1}{2} \log \frac{\bar{d}_k - \sigma^2_{X \| Y^k}}{d_k - \sigma^2_{X \| Y^k} \bar{d}_k} d_k,$$

where

$$\bar{d} \triangleq a^2 d + \sigma^2_V,$$

$$\bar{d}_k \triangleq a^2 d_k + \sigma^2_V,$$

and the minimum is over \( \{d_k\}_{k=1}^{K} \) s.t.

$$\frac{1}{d} \leq \frac{1}{\sigma^2_{X \| Y^K}} - \sum_{k=1}^{K} \left( \frac{1}{\sigma^2_{X \| Y^k}} - \frac{1}{d_k} \right),$$

$$d_k \geq \sigma^2_{X \| Y^k}.$$
Achievability

- Evaluate the directed M.I. in the Main Coding Theorem with

\[ U^{k*}_i = \bar{X}^k_i + Z^k_i \]

- This corresponds to quantizing the innovations of the process \( \bar{X}^k_i \) using random coding and binning with Gaussian codebooks.
Causally conditioned directed information (Kramer, ’98)

$$I(X_t \rightarrow \hat{X}_t \, \| \, Z_t) \triangleq \sum_{i=1}^{t} I(X_i; \hat{X}_i \, \| \, \hat{X}_{i-1}, Z_i)$$
Converse

$$R_{CEO, t}(d) \geq \min \left\{ I \left( X_{[t]} \rightarrow U_{[t]}^{[K]} \right) + \sum_{k=1}^{K} I \left( \bar{X}_{[t]}^k \rightarrow U_{[t]}^k \| X_{[t]} \right) \right\}$$

causal rate-distortion function (Gorbunov-Pinsker’74)

causal rate-distortion function with side information (Kostina-Hassibi ’18)

An ingredient:

- Optimal combining of independent Kalman filters:

$$\bar{X}_i = \sum_{k=1}^{K} \frac{\sigma_X^2}{\sigma_X^2 \| Y^{[K]} \}} \bar{X}_i^k ,$$
Causal coding problems are both practically *important* and theoretically *tractable*