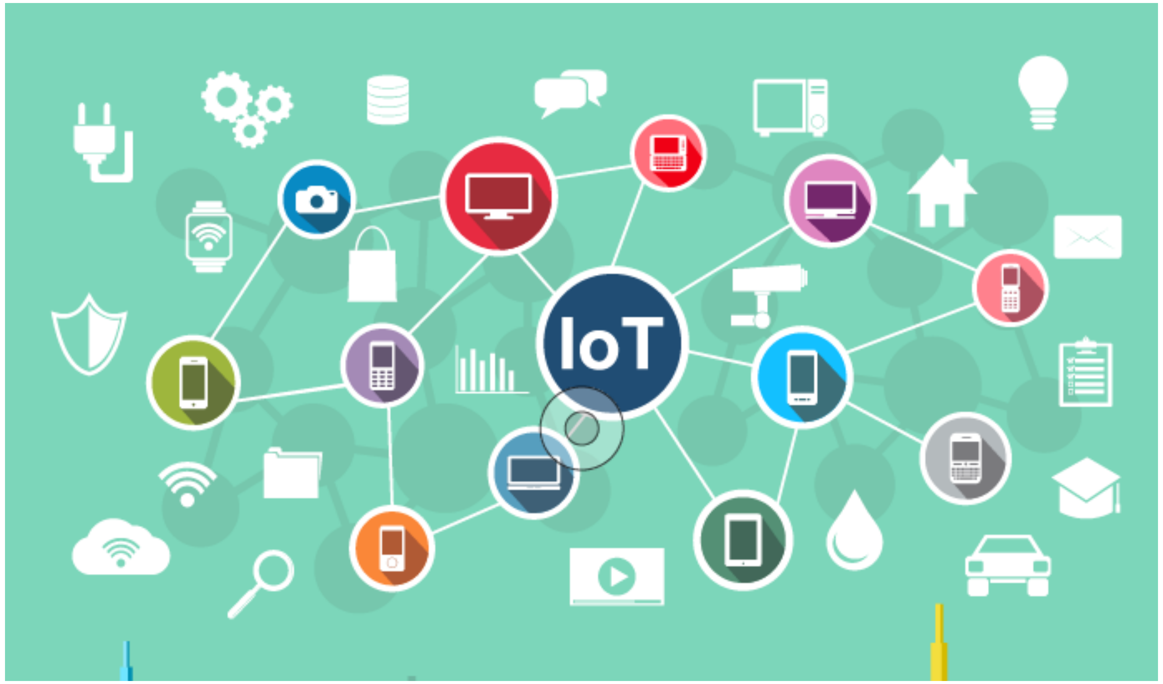
Optimal Causal Rate-constrained sampling for a class of continuous Markov processes

Nian Guo, Victoria Kostina

California Institute of Technology

ISIT, June 2020

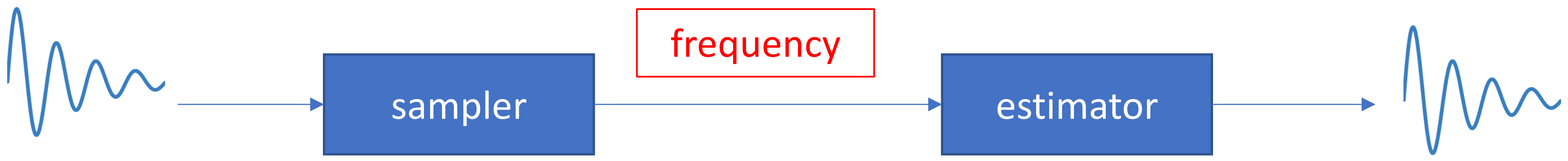
Background and Motivation



- Communication: limited resource.
- Delay: undesirable.
- Interest: Minimize distortion.

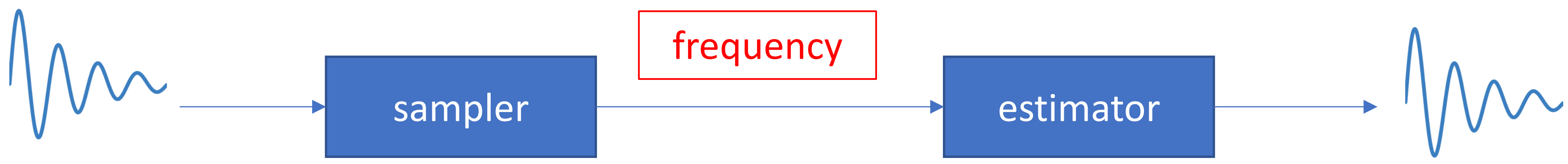
Remote estimation

Background and motivation



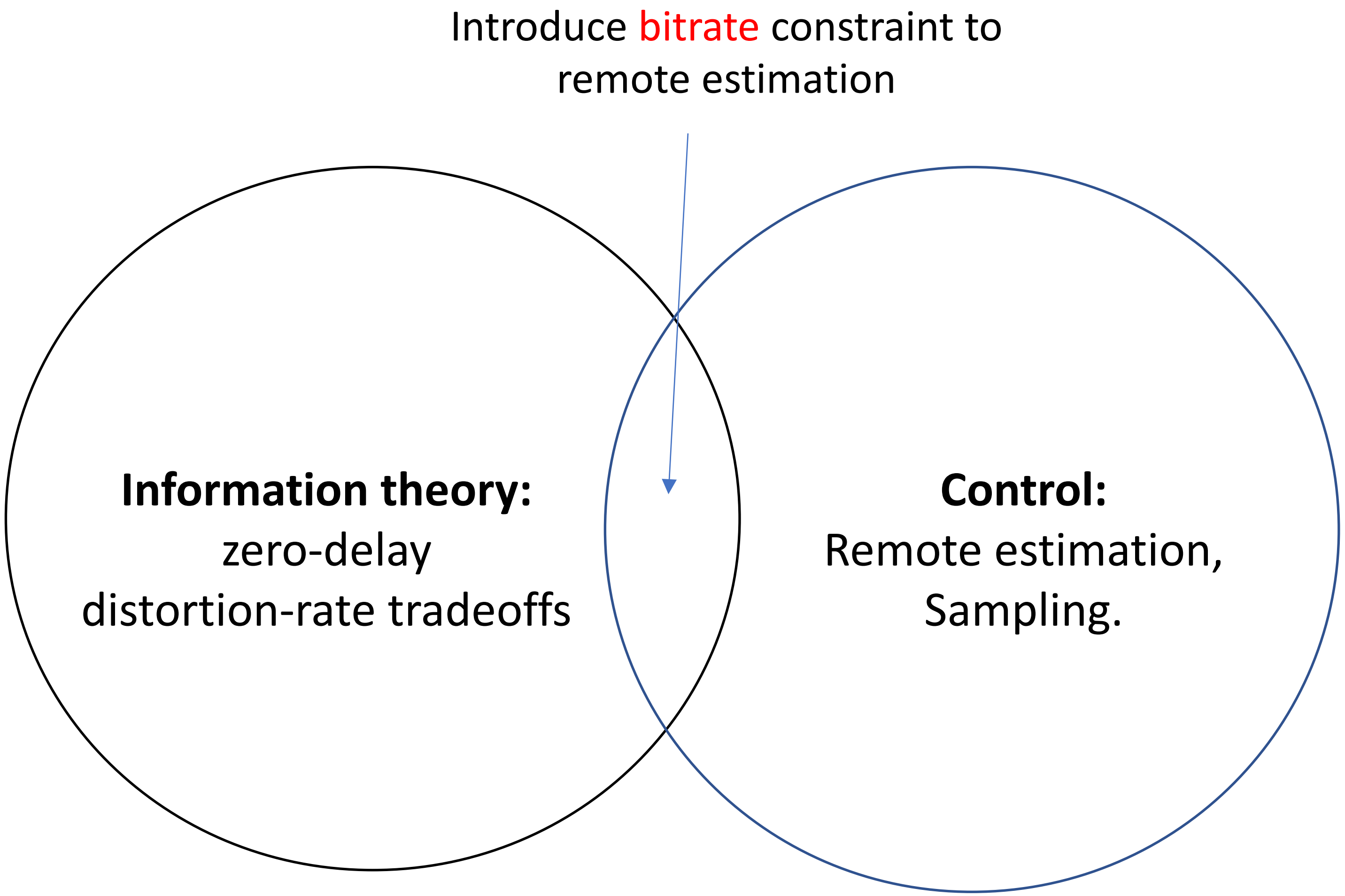
- Discrete-time processes:
  O. C. Imer and T. Başar, 2010, G. M. Lipsa and N. C. Martins, 2011, J. Wu, Q. Jia, K. H. Johansson and L. Shi, 2013, J. Chakravorty and A. Mahajan, 2017 and A. Molin and S. Hirche, 2017.
- Continuous-time processes:
  M. Rabi, G. V. Moustakides, and J. S. Baras, 2012, K. Nar and T. Başar, 2014, Y. Sun, Y. Polyanskiy and E. Uysal-Biyikoglu, 2017, and T. Z. Ornee and Y. Sun, 2019.

Background and motivation



Digital communication: quantization matters!

Bitrate constraint


Introduce bitrate constraint to
remote estimation
Information theory:
zero-delay
distortion-rate tradeoffs
Control:
Remote estimation,
Sampling.

- Notation.

$$x = \{x_1, x_2, \ldots\},\ x^i \triangleq \{x_1, x_2, \ldots, x_i\}$$

System Model

- **Input**: $\{X_t\}_{t=0}^{T}$

System Model

- **Encoder**: $0 \le \tau_1 \le \tau_2 \le \cdots \le \tau_N \le T$ (codeword-generating time)
  - causal sampling policy: set-valued $\{\mathcal{A}_t\}_{t=0}^T$,

$$\tau_{i+1} = \inf\{t \ge \tau_i, X_t \notin \mathcal{A}_t\}.$$

$U_1, U_2, \ldots, U_N \in \mathbb{Z}_+$ (codeword)

- causal compressing policy: $\mathbb{Z}_+$-valued $\{f_t\}_{t=0}^T$,

$$U_i = f_{\tau_i}$$

System Model

- **Channel**: noiseless, delay free
- **Decoding policy**: optimal

$$\hat{X}_t = \mathbb{E}[X_t | U^i, \tau^i, t < \tau_{i+1}], \;\; t \in [\tau_i, \tau_{i+1})$$

- Free timing information.
  - Time synchronization
  - τ_i can be random, but...

System Model

$$\xrightarrow{X_t} \boxed{\text{encoder}} \xrightarrow{(U_i,\tau_i)} \boxed{\text{channel}} \xrightarrow{(U_i,\tau_i)} \boxed{\text{decoder}} \xrightarrow{\hat{X}_t}$$

- **Rate constraint**:

total number of codewords (random)

$$\frac{1}{T}\mathbb{E}\left[\sum_{i=1}^{N} \ell(U_i)\right] \le R \text{ (bits per sec)}.$$

codeword length

- **Distortion measure**:

$$\frac{1}{T}\mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right] \le d. \qquad \hat{X}_t = \mathbb{E}[X_t | U^i, \tau^i, t < \tau_{i+1}],\ t \in [\tau_i, \tau_{i+1})$$

Penalizes delay

- Real-time estimation
- Causal information

Distortion-rate tradeoffs

- (R, d, T) **causal rate-constrained code** :
  a pair of enc-dec policies that satisfies (R, d) within time T.

- **Distortion-rate function**:
  $D(R) \triangleq \inf\{d : \exists (R, d, T) \text{ causal rate-constrained code}\}$

- **Goal**: find the enc policy that achieves the optimal tradeoff between R and d.

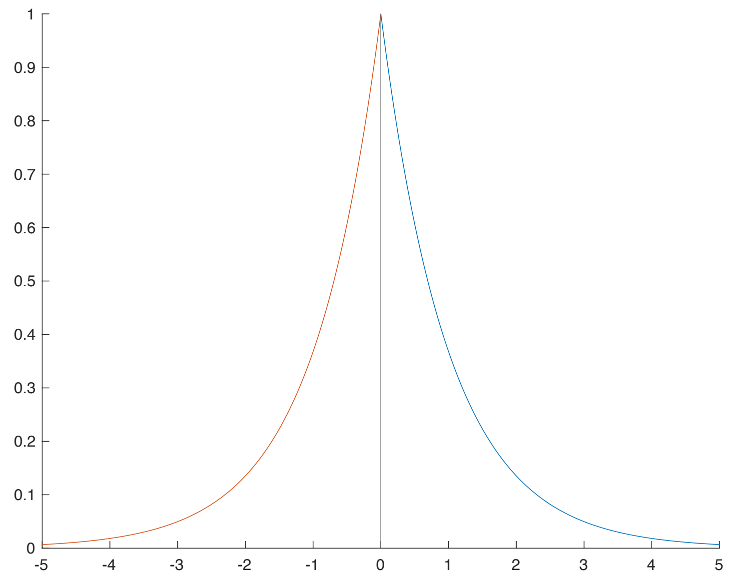
Highlights of the problem setting

- Distortion measure: penalize delay
- Causality: constraining the enc-dec to use causal information, $D(R) \uparrow$
  Less delay.
- Free timing information

Regularity conditions and assumptions

- Regularity conditions on the input process:
  - Strong Markov property
  - Continuous paths
  - Mean-square residual error: $X_t - \mathbb{E}[X_t | X_\tau, \tau]$ even, quasi-concave, etc.

- Examples:
  - Wiener process
  - Ornstein-Uhlenbeck processes
  - Lévy processes



Regularity conditions and assumptions

- Assumptions on the causal sampling policies:
  - $$\mathbb{E}[\tau_{i+1} - \tau_i] < \infty,\ \ i = 0, 1, \ldots,$$
    $$\mathbb{E}\left[\int_{\tau_i}^{\tau_{i+1}} (X_t - \mathbb{E}[X_t | \{X_{\tau_j}\}_{j=1}^{i}, \tau^i, t < \tau_{i+1}])^2 dt\right] < \infty,\ \ i = 0, 1, \ldots$$
  - For all $i = 0, 1, \ldots,$ $f_{\tau_{i+1}|\tau_i}$ exists,
    $\mathcal{A}_t$ is almost surely continuous in t on each of the intervals $[\tau_i, \tau_{i+1})$.

Definition: The Sign-of-innovation (SOI) code

The SOI code for $\{X_t\}_{t=0}^{T}$:

- Encoder
  1. Symmetric threshold sampling policy:

$$\tau_{i+1} = \inf\{t \geq \tau_i : X_t - \mathbb{E}[X_t | X_{\tau_i}, \tau_i] \notin (-a(t, \tau_i, i), a(t, \tau_i, i))\}.$$

Definition: The Sign-of-innovation (SOI) code

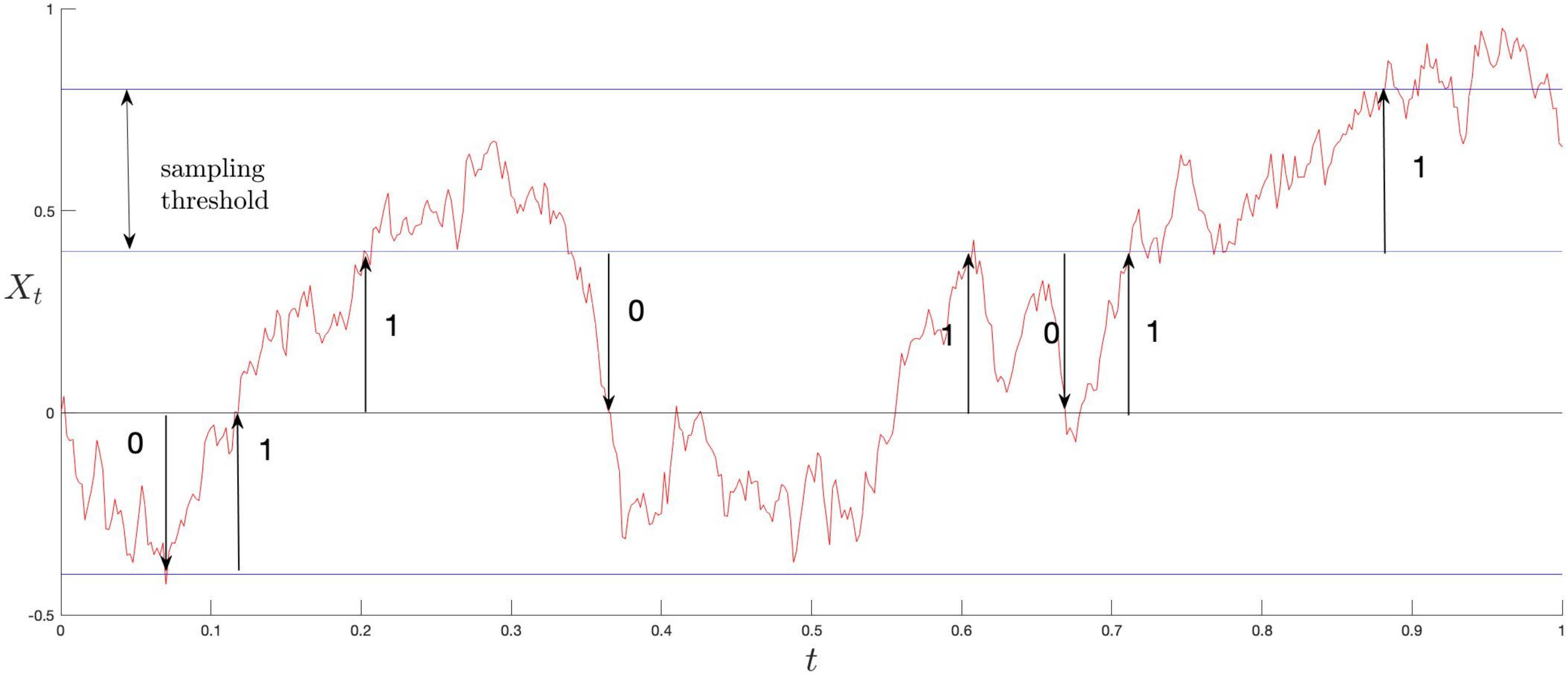
The SOI code for $\{X_t\}_{t=0}^{T}$:

- Encoder
  1. Symmetric threshold sampling policy:

$$\tau_{i+1} = \inf\{t \geq \tau_i : X_t - \mathbb{E}[X_t | X_{\tau_i}, \tau_i] \notin (-a(t, \tau_i, i), a(t, \tau_i, i))\}.$$

  2. SOI compressor:

$$U_i = \begin{cases} 1 & X_{\tau_i} - \mathbb{E}[X_{\tau_i} | X_{\tau_{i-1}}, \tau_{i-1}] = a(\tau_i, \tau_{i-1}, i-1) \\ 0 & X_{\tau_i} - \mathbb{E}[X_{\tau_i} | X_{\tau_{i-1}}, \tau_{i-1}] = -a(\tau_i, \tau_{i-1}, i-1). \end{cases}$$


sampling
threshold
X_t
t
1
0.5
0
-0.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
1
0
1
0
1
1

Definition: The Sign-of-innovation (SOI) code

The SOI code for $\{X_t\}_{t=0}^{T}$:

- Decoder: at time τ_i,

$$X_{\tau_i} = \underbrace{(2U_i - 1)a(\tau_i, \tau_{i-1}, i-1)}_{X_{\tau_i} - \mathbb{E}[X_{\tau_i} | X_{\tau_{i-1}}, \tau_{i-1}]} + \mathbb{E}[X_{\tau_i} | X_{\tau_{i-1}}, \tau_{i-1}],$$

$$\hat{X}_t = \mathbb{E}[X_t | X_{\tau_i}, \tau_i], \;\; t \in [\tau_i, \tau_{i+1}).$$

General form: $\hat{X}_t = \mathbb{E}[X_t | U^i, \tau^i, t < \tau_{i+1}]$

Optimal rate-constrained code

- Theorem (Structural).
  Optimal rate-constrained code is the SOI code.

- For **time-homogeneous** continuous Markov processes satisfying regularity conditions, in the infinite time horizon:

$$\tau_{i+1} = \inf\{t \geq \tau_i : X_t - \mathbb{E}[X_t | X_{\tau_i}, \tau_i] \notin (-a'(t-\tau_i), a'(t-\tau_i))\}.$$

Optimal rate-constrained code

- Calculate $a'(t-\tau_i)$

$$\min_{\substack{\{a'(t)\}_{t\geq 0}:\\ \mathbb{E}[\tau_1]=\frac{1}{R}}} \frac{\mathbb{E}\left[\int_0^{\tau_1}(X_t-\mathbb{E}[X_t]^2)dt\right]}{\mathbb{E}[\tau_1]}$$

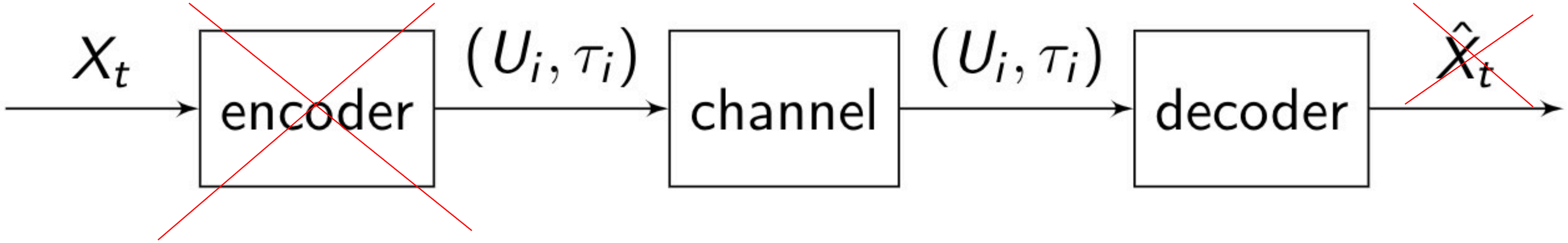
- **Corollary:**
  - Wiener process: $a=\sqrt{\frac{1}{R}},\ D(R)=\frac{1}{6R}$.
  - OU process: $a=\sqrt{R_1^{-1}\left(\frac{1}{R}\right)},\ D(R)=R\cdot R_2\left(R_1^{-1}\left(\frac{1}{R}\right)\right)$.

- $a(t, \tau_i, i)$: optimal frequency-constrained sampling policy of $\{X_t\}_{t=0}^T$.

$$\Downarrow$$

- What is the frequency-constrained setting?
- What is the optimal frequency-constrained sampling policy?
- Why does that sampling policy + the SOI compressor form the optimal rate-constrained code?

Distortion-frequency tradeoffs



- ~~Encoder~~ **Sampler**: transmit $\mathbb{R}$-valued samples $U_i = X_{\tau_i}$

- **Decoding policy**: $\bar{X}_t = \mathbb{E}[X_t | \{X_{\tau_j}, \tau_j\}_{j=1}^{i}, t < \tau_{i+1}]$.

- ~~Rate constraint~~ **Frequency constraint**:

$$\frac{\mathbb{E}[N]}{T} \leq F \text{ (samples per sec)}$$

- **Distortion measure**: $\frac{1}{T}\mathbb{E}\left[\int_0^T (X_t - \bar{X}_t)^2 dt\right] \leq d$.

Tradeoffs

- (F, d, T) **causal freq-constrained code**

- **Distortion-frequency function**:
  $\underline{D}(F) \triangleq \inf\{d : \exists\ (F, d, T) \text{ causal freq-constrained code}\}$.

Optimal freq.-constrained sampling policy

- Theorem

For a class of continuous Markov processes satisfying regularity conditions:

$$\tau_{i+1} = \inf\{t \geq \tau_i : X_t - \mathbb{E}[X_t | X_{\tau_i}, \tau_i] \notin (-a(t, \tau_i, i), a(t, \tau_i, i))\}.$$

Optimal freq.-constrained sampling policy

- Proof idea:

For any causal sampling policy $\{\mathcal{A}_t\}_{t=0}^{T}$, there exists a symmetric threshold sampling policy,

Sampling policy

$\{\mathcal{A}_t\}_{t=0}^{T}$

$$\boxed{\mathbb{E}[N_T]} = \boxed{\mathbb{E}[N'_T]},$$

$$\boxed{\mathbb{E}\left[\sum_{i=0}^{N_T}\int_{\tau_i}^{\tau_{i+1}}(X_t-\bar{X}_t)^2\,dt\right]} \geq \boxed{\mathbb{E}\left[\sum_{i=0}^{N'_T}\int_{\tau'_i}^{\tau'_{i+1}}(X_t-\bar{X}'_t)^2\,dt\right]}$$

Symmetric threshold policy

- Majorization $\rightarrow$ inequality
- Real induction $\rightarrow$ the statement holds on an interval.

Optimal freq.-constrained sampling policy

- What's new?
  - A wider class of continuous-time stochastic processes.
    - Wiener process (M. Rabi et al., 2012, Nar and Başar, 2014, Sun et al. 2017)
    - Ornstein-Uhlenbeck processes (M. Rabi et al., 2012, Ornee and Sun, 2019.)
  - Prior literature ignored the implied knowledge that "the next sample has not arrived" in the decoding policy.

General form: $\mathbb{E}[X_t|\{X_{\tau_j}, \tau_j\}_{j=1}^{i}, t < \tau_{i+1}]$.

Confirms the conjecture.

$$\bar{X}_t = \mathbb{E}[X_t|X_{\tau_i}, \tau_i], \ t \in [\tau_i, \tau_{i+1}).$$

- What is the frequency-constrained setting?
- What is the optimal frequency-constrained sampling policy?
- Why does that sampling policy + the SOI compressor form the optimal rate-constrained code?

From freq. constraint to rate constraint

- Converse:
  Distortion-rate function $(R) \geq$ Distortion-freq. function $(F = R)$.
- Achievability:
  $$\underbrace{\text{Optimal freq.-const. sampling} + \text{SOI compressor}}_{\text{achieves the lower bound}}$$

Summary

- Introduce a bitrate constraint to remote estimation problems.
- Optimal frequency-constrained sampling policy is a symmetric threshold sampling policy.
- Optimal rate-constrained code is an SOI code.

Open problems

- Random channel delay?
  - The SOI code is optimal for the Wiener process + fixed delay. (Guo and Kostina, 2019).
- Noisy channel?
  - E.g. BEC, BSC, AWGN.
- Multi-dimensional stochastic processes?
  - Freq-constrained, multi-dimensional Wiener (Nar and Başar, 2014)
- Partially observed stochastic process?
  - Kalman-filter related.
- Compression of the stopping times $\tau_1, \tau_2, \ldots$?