The Rate-Distortion Function for Sampled Cyclostationary Gaussian Processes with Memory and with Bounded Processing Delay

Zikun Tan[1] **Ron Dabora**[1] **H. Vincent Poor**[2]

[1]Department of Electrical & Computer Engineering
Ben-Gurion University of the Negev, Be'er Sheva, Israel

[2]Department of Electrical & Computer Engineering
Princeton University, Princeton, NJ, USA

Ann Arbor, MI, USA

June 24th, 2025

2025 IEEE International Symposium on Information Theory

Outline

- Man-made signals are typically generated via **repetitive** operations, resulting in **cyclostationary** statistics.

  - A **wide-sense cyclostationary (WSCS)** process has a *periodic mean* and a *periodic autocorrelation function (AF)* with the *same* period:

  - A real **continuous-time (CT)** random process $X(t)$, $t \in \mathcal{R}$, is called **WSCS** if $\exists \mathbf{T_c} \in \mathcal{R}^{++}$, s.t. $\forall\ \lambda \in \mathcal{R}$,

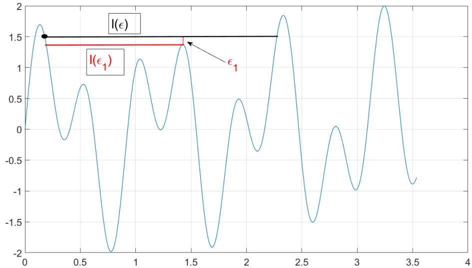
$$m_X(t) \triangleq \mathbb{E}\{X(t)\} = m_X(t + \mathbf{T_c}),$$
$$c_X(t, \lambda) \triangleq \mathbb{E}\{X(t) \cdot X(t + \lambda)\} = c_X(t + \mathbf{T_c}, \lambda).$$

  - Similarly, we define a **discrete-time (DT) WSCS** process.

- For example, *baseband OFDM signals* are **CT WSCS**.

- A real DT deterministic function $f[i]$, $i \in \mathcal{Z}$, is called **almost periodic**, if
  - $\forall \epsilon \in \mathcal{R}^{++}$, $\exists l_\epsilon \in \mathcal{N}^{+}$, which satisfies:
  - $\forall \alpha \in \mathcal{Z}$, $\exists \Delta \in [\alpha, \alpha + l_\epsilon)$, s.t.

$$\sup_{i \in \mathcal{Z}} |f[i + \Delta] - f[i]| < \epsilon.$$



- A real **zero-mean** DT random process $X[i]$, $i \in \mathcal{Z}$, is called **wide-sense almost cyclostationary (WSACS)** if $c_X[i, \Delta]$ is *almost periodic* in $i \in \mathcal{Z}$.

Sampling a CT WSCS process

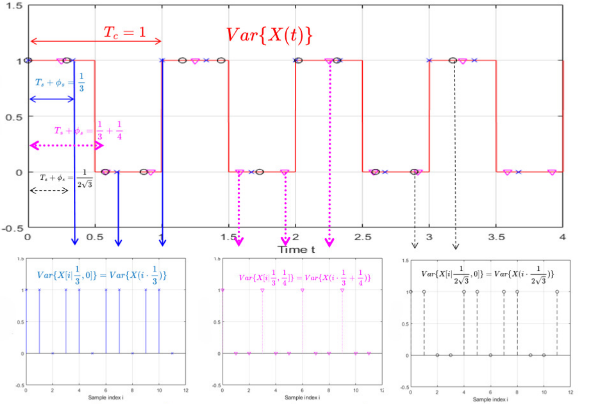
- Given a **CT WSCS source process** $X_c(t)$ with the **period of statistics** of T_c.
- Sample $X_c(t)$ using the **interval** $T_s(\epsilon) \triangleq \frac{T_c}{p+\epsilon}$, where $p \in \mathcal{N}^+$ and $\epsilon \in (0, 1)$, and the **initial sampling phase** $\phi_s \in [0, T_c)$

$$\longrightarrow \text{DT sampled process:}$$

$$X_\epsilon^{\phi_s}[i] \triangleq X_c\big(T_s(\epsilon) \cdot i + \phi_s\big).$$

  - **Synchronous sampling**: $\epsilon \in \mathcal{Q} \longrightarrow X_\epsilon^{\phi_s}[i]$ is a **DT WSCS process**;
  - **Asynchronous sampling**: $\epsilon \notin \mathcal{Q} \longrightarrow X_\epsilon^{\phi_s}[i]$ is a **DT WSACS process**.

Sampling a CT WSCS process



Scenario 1: synchronous sampling, zero initial sampling phase

Scenario 2: synchronous sampling, non-zero initial sampling phase

Scenario 3: asynchronous sampling, zero initial sampling phase

$\underbrace{\text{Sampling a } \textbf{CT WSCS} \text{ process}}_{\textbf{NOT}\ \textit{necessarily}\ \text{a}\ \textbf{DT WSCS}\ \text{process.}}$ vs. $\underbrace{\text{Sampling a } \textbf{CT stationary} \text{ process}}_{\textit{Necessarily}\ \text{a}\ \textbf{DT stationary}\ \text{process.}}$

- Denote with $\mathcal{X}$ the **source alphabet** and with $\hat{\mathcal{X}}$ the **reconstruction alphabet**.
- A **lossy source code** (M, l) consists of
  - a **message set** of **size** M;
  - an **encoder**: $f_l(\cdot) : \mathcal{X}^l \mapsto \{0, 1, 2, \ldots, M-1\}$;
  - a **decoder**: $g_l(\cdot) : \{0, 1, 2, \ldots, M-1\} \mapsto \hat{\mathcal{X}}^l$.
- The **squared-error distortion function** between a **source sequence** $\{x_i\}_{i=0}^{l-1} \equiv x^l$ and a **reconstruction sequence** $\{\hat{x}_i\}_{i=0}^{l-1} \equiv \hat{x}^l$ is given as

$$d_{se}\left(x^l, \hat{x}^l\right) \triangleq \frac{1}{l} \sum_{i=0}^{l-1} (x_i - \hat{x}_i)^2 \equiv \overline{d}\left(x^l, \hat{x}^l\right).$$

- A **rate-distortion pair** (R, D) is **achievable**, if there exists a **lossy source code** $(2^{lR}, l)$ satisfying

$$\limsup_{l \to \infty} \mathbb{E}\left\{\overline{d}\left(\mathbf{X}^l, g_l\left(f_l(\mathbf{X}^l)\right)\right)\right\} \leq D.$$

- The **rate-distortion function (RDF)** is the **infimum** of all **code rates** R, for which the **rate-distortion pair** (R, D) is **achievable**.

- **Standard rate-distortion analysis** relies on the **stationarity** and the **ergodicity** (or the **information-stability**) of the source.

- However, DT WSACS processes are **nonstationary** and **nonergodic**.

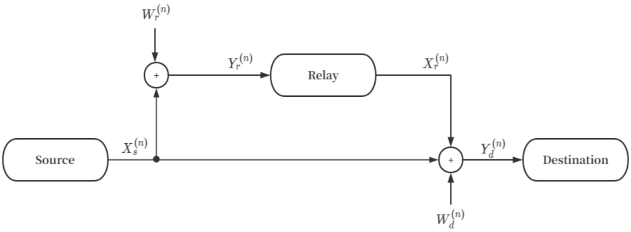
  $\Rightarrow$ Conventional information-theoretic arguments are **inapplicable**.

- Suitable frameworks for analyzing **information-unstable processes**:
  - **Asymptotically mean-stationary (AMS) processes** [Gray: 2011];
  - **Information-spectrum framework** [Han: 2010].

- We use the **information-spectrum framework** in our study.

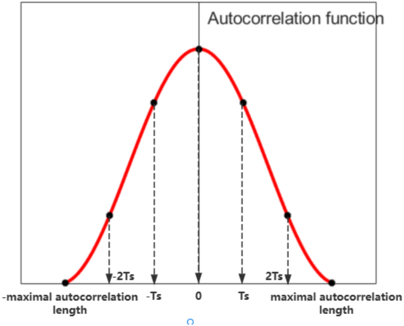
Outline

Compress-and-forward relay networks



- The (*Gaussian*) signal received at the **relay** is **sampled**,
  - before being **compressed** and **forwarded** to the destination.
- Practically, the **ratio** between the **actual sampling interval** at the relay and the **actual period of statistics** of the received signal is **irrational**, due to
  - **physical separation** of the **relay oscillator** and the **source oscillator**;
  - **inherent variability** of oscillators' frequencies (i.e., **jitter**).

- To reduce the **loss of information**, the **sampling interval** at the relay is typically *smaller* than the **maximal autocorrelation length** of the received signal.



Therefore, the sampled received signal at the relay is a **DT WSACS Gaussian process with memory.**

- [Kipnis et al.: *IEEE TIT* 2018] characterized the **distortion-rate function (DRF)** of **DT WSCS Gaussian processes**,
  - by transforming the **DT scalar WSCS process** into a **DT vector stationary process**.

- [Kipnis et al.: *IEEE TIT* 2018] characterized the **DRF** of **DT WSCS Gaussian processes**,
  - by transforming the **DT scalar WSCS process** into a **DT vector stationary process**.
- [Abakasanga et al.: *Entropy* 2020] characterized the **RDF** of **DT WSACS Gaussian memoryless processes**,
  - using the **information-spectrum framework**.

- [Kipnis et al.: *IEEE TIT* 2018] characterized the **DRF** of **DT WSCS Gaussian processes**,
  - by transforming the **DT scalar WSCS process** into a **DT vector stationary process**.
- [Abakasanga et al.: *Entropy* 2020] characterized the **RDF** of **DT WSACS Gaussian memoryless processes**,
  - using the **information-spectrum framework**.
- The **capacity** of **additive DT WSACS Gaussian channels** was characterized using the **information-spectrum framework**,
  - in [Shlezinger et al.: *IEEE TCOM* 2020] for *memoryless noise*;
  - in [Dabora and Abakasanga: *IEEE TIT* 2023] for *noise with memory*.

- [Kipnis et al.: *IEEE TIT* 2018] characterized the **DRF** of **DT WSCS Gaussian processes**,
  - by transforming the **DT scalar WSCS process** into a **DT vector stationary process**.
- [Abakasanga et al.: *Entropy* 2020] characterized the **RDF** of **DT WSACS Gaussian memoryless processes**,
  - using the **information-spectrum framework**.
- The **capacity** of **additive DT WSACS Gaussian channels** was characterized using the **information-spectrum framework**,
  - in [Shlezinger et al.: *IEEE TCOM* 2020] for *memoryless noise*;
  - in [Dabora and Abakasanga: *IEEE TIT* 2023] for *noise with memory*.
- [Tan et al.: *IEEE ISIT* 2024] characterized the **RDF** for **DT WSACS Gaussian processes with memory**, **without delay** between processing of **consecutive sequences**,
  - using the **information-spectrum framework**.

- [Kipnis et al.: *IEEE TIT* 2018] characterized the **DRF** of **DT WSCS Gaussian processes**,
  - by transforming the **DT scalar WSCS process** into a **DT vector stationary process**.
- [Abakasanga et al.: *Entropy* 2020] characterized the **RDF** of **DT WSACS Gaussian memoryless processes**,
  - using the **information-spectrum framework**.
- The **capacity** of **additive DT WSACS Gaussian channels** was characterized using the **information-spectrum framework**,
  - in [Shlezinger et al.: *IEEE TCOM* 2020] for *memoryless noise*;
  - in [Dabora and Abakasanga: *IEEE TIT* 2023] for *noise with memory*.
- [Tan et al.: *IEEE ISIT* 2024] characterized the **RDF** for **DT WSACS Gaussian processes with memory**, **without delay** between processing of **consecutive sequences**,
  - using the **information-spectrum framework**.

Our task: RDF characterization for compressing DT WSACS Gaussian processes with memory, with finite and bounded delay between processing of consecutive sequences.

Outline

- Consider a sequence of real random variables (RVs) $\{X_i\}_{i=0}^{\infty}$:
  - The **limit superior in probability** is given as

$$\text{p-}\limsup_{i\to\infty} X_i \triangleq \inf\left\{\alpha \in \mathcal{R} \,\middle|\, \lim_{i\to\infty} \Pr\{X_i > \alpha\} = 0\right\};$$

  - Given a **common probability measure** P, if

$$\lim_{u\to\infty} \sup_{i\geq 0} \int_{|X_i|\geq u} |X_i| \mathrm{d}P = 0,$$

  then the sequence is said to be **uniformly integrable**.

- Consider a real DT **source process** $X[i]$, $i \in \mathcal{N}$, and a **reconstruction process** $\hat{X}[i]$:
  - Denote a **block of l symbols** $\{X[i]\}_{i=0}^{l-1} \equiv \mathbf{X}^l$;
  - Denote a **block of l reconstruction symbols** $\{\hat{X}[i]\}_{i=0}^{l-1} \equiv \hat{\mathbf{X}}^l$.
  - Let $f_{\hat{\mathbf{X}}^l|\mathbf{X}^l}(\hat{\mathbf{x}}^l|\mathbf{x}^l)$ denote the **conditional probability density function (PDF)** of $\hat{\mathbf{X}}^l$ given $\mathbf{X}^l$.
  - Let $f_{\hat{\mathbf{X}}^l}(\hat{\mathbf{x}}^l)$ denote the **PDF** of $\hat{\mathbf{X}}^l$.
  - The **mutual information density rate** between $\mathbf{X}^l$ and $\hat{\mathbf{X}}^l$ is defined as

$$i(\mathbf{X}^l; \hat{\mathbf{X}}^l) \triangleq \frac{1}{l} \log \frac{f_{\hat{\mathbf{X}}^l|\mathbf{X}^l}(\hat{\mathbf{x}}^l|\mathbf{x}^l)}{f_{\hat{\mathbf{X}}^l}(\hat{\mathbf{x}}^l)}.$$

  - The **spectral sup-mutual information rate** between $\mathbf{X}^\infty$ and $\hat{\mathbf{X}}^\infty$, $\bar{\mathrm{I}}(\mathbf{X}^\infty, \hat{\mathbf{X}}^\infty)$, is defined as

$$\bar{\mathrm{I}}(\mathbf{X}^\infty, \hat{\mathbf{X}}^\infty) \triangleq \text{p-}\limsup_{l\to\infty} i(\mathbf{X}^l; \hat{\mathbf{X}}^l).$$

Problem Formulation and Main Result

Source Sequence Generation Model

- Consider a real **CT WSCS Gaussian** source process $X_c(t)$, s.t.
  - $\mathbb{E}\{X_c(t)\} = 0$.
  - AF $c_{x_c}(t, \lambda)$ is **periodic** in t with **period** T_c, **bounded** and **uniformly continuous** in both t and λ.
  - AF has **finite-memory** with maximal autocorrelation length λ_c,
    - i.e., $c_{x_c}(t, \lambda) \triangleq \mathbb{E}\{X_c(t) \cdot X_c(t+\lambda)\} = 0$, $\forall |\lambda| > \lambda_c$.

- $X_c(t)$ is *sampled* with **sampling interval** $T_s(\epsilon) \leq \lambda_c$ and **initial sampling phase** $\phi_s \in [0, T_c)$ $\longrightarrow$ DT sampled process:

$$X_\epsilon^{\phi_s}[i] \triangleq X_c(T_s(\epsilon) \cdot i + \phi_s),$$

$T_s(\epsilon) \triangleq \frac{T_c}{p+\epsilon}$, where $p \in \mathcal{N}^+$ and $\epsilon \notin \mathcal{Q}$, $\epsilon \in (0,1)$ is the **sampling mismatch**.

- **Finite and bounded delay** (i.e., **guard interval**) between processing of consecutive sequences is **allowed**.

  $\Rightarrow$ All sequences are **statistically independent**.

sequence 0	guard interval	• • •	guard interval	sequence n-1	guard interval	sequence n	guard interval	• • •

1. Let the **source process** $X_c(t)$ be sampled with **interval** $T_s(\epsilon_{\mathbf{n}}) \triangleq \frac{T_c}{p+\epsilon_{\mathbf{n}}}$, $\epsilon_{\mathbf{n}} \triangleq \frac{\lfloor n \cdot \epsilon \rfloor}{n} \in \mathcal{Q}$, $n \in \mathcal{N}^+$, and with an **initial sampling phase** $\phi_{\mathbf{s}} \in [0, T_c)$:

$$X_{\epsilon_{\mathbf{n}}}^{\phi_s}[i] \triangleq X_c\big(i \cdot T_s(\epsilon_{\mathbf{n}}) + \phi_{\mathbf{s}}\big).$$

$X_{\epsilon_{\mathbf{n}}}^{\phi_s}[i]$ is a **DT WSCS process** with

- maximal autocorrelation length $\tau_c \triangleq \left\lceil \frac{(p+1)\cdot\lambda_c}{T_c} \right\rceil \geq \left\lceil \frac{(p+\epsilon_{\mathbf{n}})\cdot\lambda_c}{T_c} \right\rceil$;
- period of statistics $p_n \triangleq p \cdot n + \lfloor n \cdot \epsilon \rfloor$.

2. Transform $X_{\epsilon_{\mathbf{n}}}^{\phi_s}[i]$ into a **p_n-dimensional stationary process**:

$$\left(\mathbf{X}_{\epsilon_{\mathbf{n}},\phi_s}^{p_n}[i]\right)_m \triangleq X_{\epsilon_{\mathbf{n}}}^{\phi_s}[i \cdot p_n + m], \quad m = 0, 1, \ldots, p_n - 1.$$

- **Autocorrelation matrix**

$$\mathsf{C}_{\mathbf{X}_{\epsilon_{\mathbf{n}},\phi_s}^{p_n}}[\Delta] \triangleq \mathbb{E}\left\{\mathbf{X}_{\epsilon_{\mathbf{n}},\phi_s}^{p_n}[i] \cdot \left(\mathbf{X}_{\epsilon_{\mathbf{n}},\phi_s}^{p_n}[i+\Delta]\right)^T\right\}, \quad \Delta \in \mathcal{Z};$$

- **Power spectral density (PSD) matrix**

$$\mathsf{S}_{\mathbf{X}_{\epsilon_{\mathbf{n}},\phi_s}^{p_n}}(f) \triangleq \sum_{\Delta \in \mathcal{Z}} \mathsf{C}_{\mathbf{X}_{\epsilon_{\mathbf{n}},\phi_s}^{p_n}}[\Delta] \cdot e^{-j2\pi f}, \quad -\frac{1}{2} \leq f \leq \frac{1}{2}.$$

- Let $\underline{\lambda_{\epsilon_{\mathbf{n}},\phi_s,m}^{p_n}(f)}$, $0 \leq m \leq p_n - 1$, denote the **eigenvalues** of $\mathsf{S}_{\mathbf{X}_{\epsilon_{\mathbf{n}},\phi_s}^{p_n}}(f)$.

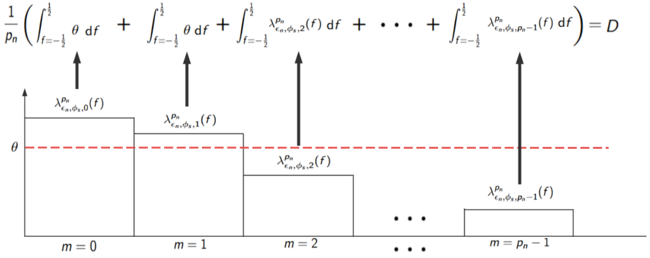
3. Given the **distortion constraint** D, the **RDF** for compressing $X_{\epsilon_n}^{\phi_s}[i]$ is [Kipnis et al.: *IEEE TIT* 2018]

$$R_{\epsilon_n}^{\phi_s}(D) = \frac{1}{2p_n}\sum_{m=0}^{p_n-1}\int_{f=-\frac{1}{2}}^{\frac{1}{2}}\left(\log\left(\frac{\lambda_{\epsilon_n,\phi_s,m}^{p_n}(f)}{\theta}\right)\right)^{+} df,$$

where θ is selected s.t.

$$D = \frac{1}{p_n}\sum_{m=0}^{p_n-1}\int_{f=-\frac{1}{2}}^{\frac{1}{2}}\min\left\{\lambda_{\epsilon_n,\phi_s,m}^{p_n}(f),\theta\right\}df.$$

- **Reverse waterfilling**



RDF characterization for $X_{\epsilon}^{\phi_s}[i]$:

1. Define $R_{\epsilon_{\mathbf{n}}}(D) \triangleq \min_{\phi_s \in [0, T_c)} R_{\epsilon_{\mathbf{n}}}^{\phi_s}(D)$.

2. If the **AF** of $X_c(t)$ satisfies

$$\min_{0 \le t < T_c} \left\{ c_{X_c}(t, 0) - 2 \cdot \tau_c \cdot \max_{|\lambda| > \frac{T_c}{p+1}} \left\{ \left| c_{X_c}(t, \lambda) \right| \right\} \right\} \geq \gamma_c > 0,$$

3. for the **distortion constraint** $D \leq \gamma_c$,

4. when a **finite and bounded delay** between processed sequences up to $\tau_c \cdot T_s(\epsilon) + T_c$ in CT is allowed,

the **RDF** for compressing $X_{\epsilon}^{\phi_s}[i]$ is given as

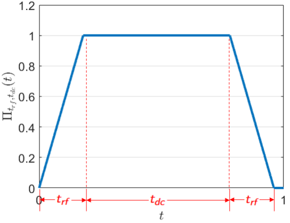
$$R_{\epsilon}(D) = \limsup_{n \to \infty} R_{\epsilon_{\mathbf{n}}}(D).$$

Outline

- **Gaussian asymptotically wide-sense stationary (AWSS) vector processes** [Gutiérrez-Gutiérrez et al.: *Entropy* 2018]
  - Such processes have **a constant mean** and their $n \times n$ **autocorrelation matrices** are **asymptotically block Toeplitz**.
  - **DT WSACS processes** *CANNOT* be transformed into **AWSS vector processes**.

- **AMS processes** [Gray: 2011]
  - **DT WSACS processes** are **AMS** [Gardner: *J. Sound Vib.* 2023].
  - For an **AMS process** with **reference letter**, its **DRF** is *equivalent* to the **DRF** of its **associated stationary source**.
  - In our scenario, we were *NOT* able to evaluate the AMS **stationary mean**, i.e., **probability measure** of a **Gaussian mixture** as the number of summands goes to **infinity**.

Setup of CT AF:



1. Consider a **periodic pulse function** $\Pi_{\mathrm{t_{rf}},\mathrm{t_{dc}}}(t)$, $t_{rf} = 0.01$, $t_{dc} \in [0, 0.98]$ and **period** 1:
2. Set **period of AF** of $X_c(t)$ as $\mathbf{T_c} = 5$ µsec.
3. Set **normalized initial sampling phase** as $\phi \triangleq \frac{\phi_s}{T_c} \in [0, 1)$.
4. The variance function of $X_c(t)$ is given as

$$c_{X_c}(t, 0) \triangleq 2 + 8 \cdot \Pi_{\mathrm{t_{rf}},\mathrm{t_{dc}}}\left(\frac{t}{T_c} - \phi\right).$$

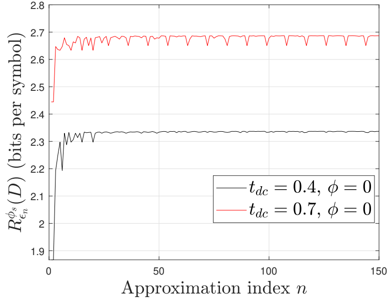
5. Set **maximal autocorrelation length** as $\lambda_c = 4$ µsec. Denote the **indicator function** as $\mathbf{1}(\cdot)$. For $\lambda \geq 0$, the AF is given as

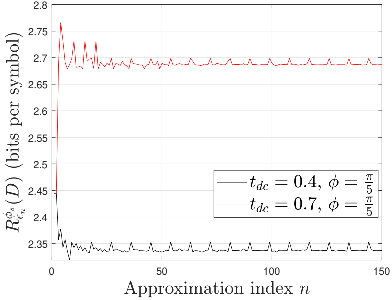
$$c_{X_c}(t, \lambda) \triangleq \mathbf{1}\Big(\lambda \in [0, \lambda_c]\Big) \cdot e^{-\lambda \cdot 10^{6.1}} \cdot c_{X_c}(t, 0).$$

6. For $\lambda < 0$, $c_{X_c}(t, \lambda) = c_{X_c}(t + \lambda, -\lambda)$.

The RDF $R_{\epsilon_{\mathbf{n}}}(D)$ w.r.t. Approximation Index n

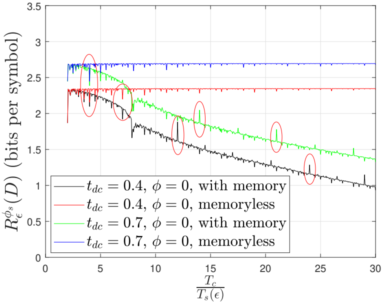
The **RDF** $R_{\epsilon_{\mathbf{n}}}(D)$ w.r.t. **approximation index** n for $\epsilon = \frac{\pi}{7}$:

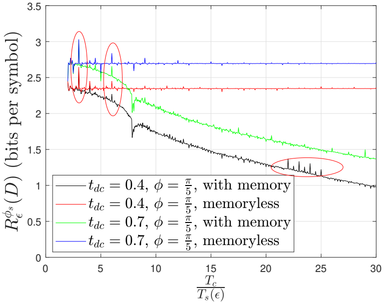




- When t_{dc} is **higher**, $R_{\epsilon_{\mathbf{n}}}(D)$ is **higher**.
  - Larger **variance** $\longrightarrow$ larger **dynamic range** $\longrightarrow$ more **bits per symbol**.
- When n is **large enough**, $R_{\epsilon_{\mathbf{n}}}(D)$ becomes **periodically stable**.
  - Larger n $\longrightarrow$ larger **period of statistics** p_n $\longrightarrow$ shifting of sampling instances **negligible**.
- $R_{\epsilon_{\mathbf{n}}}(D)$ does *NOT* converge.
  - **Asynchronism** of sampling $\longrightarrow$ **nonstationarity**.

The **RDF** w.r.t. **sampling frequency**:





- RDF for **processes with memory** is generally *smaller* than RDF for **memoryless processes**.
- **Small sampling rate variations** could lead to **large compression rate variations**.

Practical insight: In design of communications systems, assume **asynchronous sampling**.

Outline

Summary

- We defined the problem of **RDF characterization** for **DT WSACS Gaussian processes with memory**,
  - with **finite and bounded processing delay** between consecutive sequences.
- This problem arises in practical sampling scenarios:
  - e.g., **compress-and-forward relay networks**.
- We presented the **RDF result** for compressing **DT WSACS Gaussian processes with memory** in our defined setup,
  - as a **limit superior** of a **sequence** of **computable RDFs**:

$$R_{\epsilon}(D) = \limsup_{n\to\infty} R_{\epsilon_{\mathbf{n}}}(D).$$

  - **Useful** for numerical evaluations.

Summary

- We discussed **additional** frameworks for RDF analysis and notions implied from **numerical evaluations**.
  - **Gaussian AWSS vector processes** and **AMS processes** are *NOT* useful for RDF analysis in our problem.

- **Important insight**: Assume **asynchronous sampling** in **practical communications systems design**: