On the Rate-Distortion Function for Sampled Cyclostationary Gaussian Processes with Memory

Zikun Tan[1] **Ron Dabora**[1,2] **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

July 8, 2024

2024 IEEE International Symposium on Information Theory

Outline

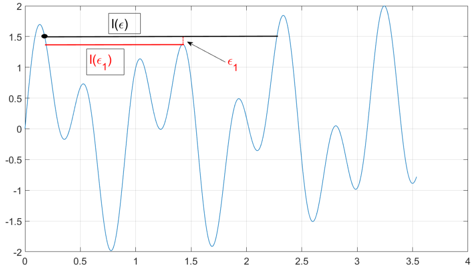
Outline

- 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.
  - For example, *baseband OFDM signals* are **continuous-time (CT) WSCS**.
- However, in many communications scenarios, the associated random processes are *assumed* to be *stationary*.
- Applying stationary signal processing algorithms in cyclostationary signal processing often leads to errors [Gardner: *IEEE TCOM* 1987].

- A real CT random process $X(t)$, $t \in \mathcal{R}$, is called **WSCS** if both its *mean* $m_X(t)$ and its *AF* $c_X(t, \lambda)$ are *periodic* in *time* t with some *period* $T_c \in \mathcal{R}^{++}$ for *any lag* $\lambda \in \mathcal{R}$, i.e.,
  - $m_X(t) \triangleq \mathbb{E}\{X(t)\} = m_X(t + T_c)$,
  - $c_X(t, \lambda) \triangleq \mathbb{E}\{X(t) \cdot X(t + \lambda)\} = c_X(t + T_c, \lambda)$.
- A real discrete-time (DT) random process $X[i]$, $i \in \mathcal{Z}$, is called **WSCS** if both its *mean* $m_X[i]$ and its *AF* $c_X[i, \Delta]$ are *periodic* in *time* i with some *period* $N_c \in \mathcal{N}^+$ for *any lag* $\Delta \in \mathcal{Z}$, i.e.,
  - $m_X[i] \triangleq \mathbb{E}\{X[i]\} = m_X[i + N_c]$,
  - $c_X[i, \Delta] \triangleq \mathbb{E}\{X[i] \cdot X[i + \Delta]\} = c_X[i + N_c, \Delta]$.

- A real DT deterministic function $f[i]$, $i \in \mathcal{Z}$, is called **almost periodic**, if for any $\epsilon \in \mathcal{R}^{++}$, there exists an *associated number* $l_\epsilon \in \mathcal{N}^+$ which satisfies that for any $\alpha \in \mathcal{Z}$, there exists $\Delta \in [\alpha, \alpha + l_\epsilon)$, such that

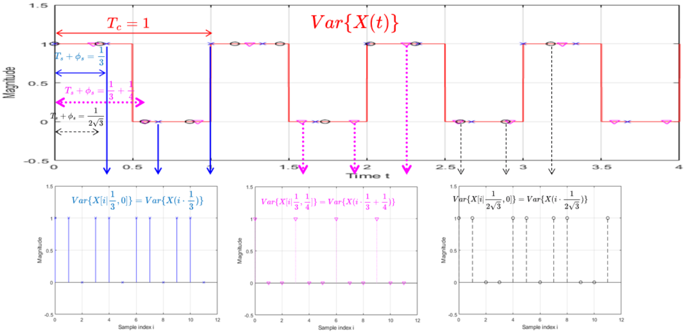
$$\sup_{i \in \mathcal{Z}} |f(i + \Delta) - f(i)| < \epsilon.$$



- A real DT random process $X[i]$, $i \in \mathcal{Z}$, is called **wide-sense almost cyclostationary (WSACS)** if both its *mean* $m_X[i]$ and its *AF* $c_X[i, \Delta]$ are *almost periodic* in *time* i *for any lag* $\Delta \in \mathcal{Z}$.

Background

Statistics of Sampled CT WSCS Processes



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

- Sampling a **CT WSCS** process $\longrightarrow$ *NOT necessarily* a **DT WSCS** process.
  - The sampled CT WSCS processes are **nonstationary**.
- In contrast, sampling a **CT stationary** process $\longrightarrow$ *necessarily* a **DT stationary** process.

- A **lossy source code** (M, l), where M is the **size** of the **message set** and l is the **blocklength**, consists of an **encoder** $f_l(\cdot)$ and a **decoder** $g_l(\cdot)$.
- The **squared-error distortion function** between the source symbol x and its reconstruction symbol $\hat{x}$ is

$$d_{se}(x, \hat{x}) \triangleq (x - \hat{x})^2$$

- For a block of l source symbols $\{X_i\}_{i=0}^{l-1} \equiv \mathbf{X}^l$, a **rate-distortion pair** (R, D) is **achievable**, if there exists a **lossy source code** $(2^{lR}, l)$ satisfying

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

- 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**.
  - Conventional information-theoretic arguments are *inapplicable*.

- We employ the **information-spectrum framework** [Han: 2010], which applies to **information-unstable processes**.
  - For a sequence of real random variables (RVs) $\{X_i\}_{i=0}^{\infty}$, its **limit superior in probability** is

$$\text{p-}\limsup_{i\to\infty} X_i \triangleq \inf\left\{\alpha \in \mathcal{R} \,\middle|\, \lim_{i\to\infty} \Pr\{X_i > \alpha\} = 0\right\} \triangleq \alpha_0.$$

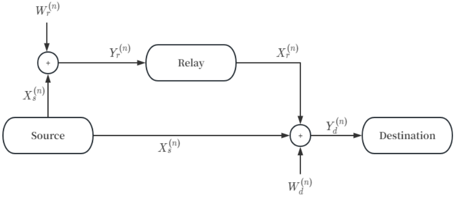
  - A sequence of real continuous RVs $\{X_i\}_{i=0}^{\infty}$ satisfying

$$\lim_{u\to\infty} \sup_{i\geq 0} \int_{x:|x|\geq u} f_{X_i}(x)|x|\mathrm{d}x = 0$$

    is said to be **uniformly integrable**.

Outline

- **Compress-and-forward relay networks**:



- The (*Gaussian*) signal received at the **relay** needs to be **sampled**, before being **compressed** and **forwarded** to the destination.
- Usually, the **ratio** between the **sampling interval** at the relay and the **period of statistics** of the received signal is **irrational**, due to:
  - The **physical separation** of the **relay oscillator** and the **source oscillator**.
  - **Inherent variability** of oscillators' frequencies.
- 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**.

Outline

- **Previous related works:**
  - The work [Kipnis et al.: *IEEE TIT* 2018] characterized the **distortion-rate function (DRF)** of **DT WSCS Gaussian processes with memory**, using the **decimated component decomposition (DCD)** approach.
  - The work [Abakasanga et al.: *Entropy* 2020] characterized the **RDF** of **DT WSACS Gaussian memoryless processes**, using the **information-spectrum framework**.
  - The work [Dabora and Abakasanga: *IEEE TIT* 2023] characterized the **capacity** of **additive DT WSACS Gaussian channels with memory**, using the **information-spectrum framework**.

Our task: **RDF characterization for compressing DT WSACS Gaussian processes with memory**.

Problem Formulation

Source Message Generation Model

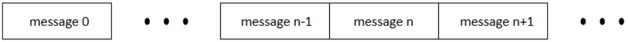
- In our scenario, we have a real CT source process $X_c(t)$, such that
  - it is **zero-mean**, **WSCS** and **Gaussian** with the period of statistics T_c.
  - its AF $c_{x_c}(t, \lambda)$ is **bounded** and **uniformly continuous** in both t and λ.
  - it has a **finite-memory** with the maximal autocorrelation 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 the **sampling interval** $T_s(\epsilon) = \frac{T_c}{p+\epsilon}$ and the **initial sampling phase** $\phi_s \in [0, T_c) \longrightarrow$ DT sampled process

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

  where
  - $p \in \mathcal{N}^+$, $\epsilon \notin \mathcal{Q}$ and $\epsilon \in (0, 1)$,
  - $T_s(\epsilon) \leq \lambda_c$.

- The source message generation is **continuous**, i.e., the delay in any two consecutive source messages is **NOT** allowed.



- Given a real DT process $X[i]$, $i \in \mathcal{N}$, a **block of l symbols** $\{X[i]\}_{i=0}^{l-1} \equiv \mathbf{X}^l$ and its **block of l reconstruction symbols** $\{\hat{X}[i]\}_{i=0}^{l-1} \equiv \hat{\mathbf{X}}^l$.
  - Let $f_{\hat{\mathbf{X}}^l|\mathbf{X}^l}(\hat{\mathbf{X}}^l|\mathbf{X}^l)$ and $f_{\hat{\mathbf{X}}^l}(\hat{\mathbf{X}}^l)$ denote the **probability density functions (PDFs)** of $\hat{\mathbf{X}}^l$ given $\mathbf{X}^l$ and of $\hat{\mathbf{X}}^l$, respectively.
  - The **mutual information density rate** between $\mathbf{X}^l$ and $\hat{\mathbf{X}}^l$ is

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

  - Then, the **spectral sup-mutual information rate** $\bar{\mathrm{I}}(\mathbf{X}^l, \hat{\mathbf{X}}^l)$ between $\mathbf{X}^l$ and $\hat{\mathbf{X}}^l$ is

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

- For an **arbitrary** DT process $X[i]$, $i \in \mathcal{N}$, if for a **block of l symbols** $\{X[i]\}_{i=0}^{l-1} \equiv \mathbf{X}^l$, there exists a **deterministic reference word** $\{r_i\}_{i=0}^{l-1} \equiv \mathbf{r}^l$ which makes the **sequence of RVs** $\{\overline{d}(\mathbf{X}^l, \mathbf{r}^l)\}_{l=0}^{\infty}$ **uniformly integrable**.
    - Denote the **block of l reconstruction symbols** $\{\hat{X}_i\}_{i=0}^{l-1} \equiv \hat{\mathbf{X}}^l$ and the **joint cumulative distribution function (CDF)** of $\mathbf{X}^l$ and $\hat{\mathbf{X}}^l$ as $F_{\mathbf{X}^l, \hat{\mathbf{X}}^l}$.
    - Then, the **RDF** for compressing $X[i]$ using the **fixed-length coding** and the **average distortion criterion** is [Han: 2010]

$$R(D) = \inf_{\substack{F_{\mathbf{X}^l, \hat{\mathbf{X}}^l}: \\ \limsup_{l \to \infty} \mathbb{E}\{\overline{d}(\mathbf{X}^l, \hat{\mathbf{X}}^l)\} \leq D}} \bar{\mathrm{I}}(\mathbf{X}^l, \hat{\mathbf{X}}^l).$$

Outline

- **The RDF for compressing DT WSACS Gaussian processes with memory**:
  - Denote a **block of l symbols** $\{X_\epsilon^{\phi_s}[i]\}_{i=0}^{l-1} \equiv \mathbf{X}_{\epsilon,\phi_s}^l$, its **block of l reconstruction symbols** $\{\hat{X}_\epsilon^{\phi_s}[i]\}_{i=0}^{l-1} \equiv \hat{\mathbf{X}}_{\epsilon,\phi_s}^l$, and
$$\mathbf{S}_{\epsilon,\phi_s}^l \triangleq \mathbf{X}_{\epsilon,\phi_s}^l - \hat{\mathbf{X}}_{\epsilon,\phi_s}^l.$$
  - Define the set
$$\mathcal{C}_{\mathbf{S}_{\epsilon,\phi_s}^l} \triangleq \{\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l} \in \mathcal{R}^{l\times l} \,|\, \frac{1}{l}\mathrm{tr}\{\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l}\} \le D, \mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l} \succ 0, \mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l} = (\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l})^T, \mathsf{C}_{\mathbf{X}_{\epsilon,\phi_s}^l} \succeq \mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l}\}.$$
  - Then the **RDF** $R_\epsilon(D)$ for compressing $X_\epsilon^{\phi_s}[i]$ is
$$R_\epsilon(D) = \frac{1}{T_c}\int_{\phi_s=0}^{T_c} R_\epsilon^{\phi_s}(D)\mathrm{d}\phi_s,$$
where
$$R_\epsilon^{\phi_s}(D) \triangleq \limsup_{l\to\infty} \min_{\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l} \in \mathcal{C}_{\mathbf{S}_{\epsilon,\phi_s}^l}} \frac{1}{2l}\log\left(\frac{\det(\mathsf{C}_{\mathbf{X}_{\epsilon,\phi_s}^l})}{\det(\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^l})}\right),$$
in which $\mathbf{S}_{\epsilon,\phi_s}^l$ is a **Gaussian vector**.
  - **Corollary (RDF with the fixed code rate)**
    - If the **code rate** has to be **fixed**, the **RDF** is
$$R_\epsilon(D) = \min_{\phi_s\in[0,T_c)} R_\epsilon^{\phi_s}(D).$$

- **Converse part**:
  1. Consider a *specific* **sequence of lossy source codes** $(\tilde{M}(l,D), l)$ defined by the **encoder** $\tilde{f}_l(\cdot)$ and the **decoder** $\tilde{g}_l(\cdot)$, such that
  $$\limsup_{l\to\infty} \mathbb{E}\{\overline{d}_{se}(\mathbf{X}^l_{\epsilon,\phi_s}, \hat{\tilde{\mathbf{X}}}^l_{\epsilon,\phi_s})\} \le D,$$
  where $\hat{\tilde{\mathbf{X}}}^l_{\epsilon,\phi_s} = \tilde{g}_l\left(\tilde{f}_l(\mathbf{X}^l_{\epsilon,\phi_s})\right)$.
  2. By [Kostina2013], we obtain
  $$\log \tilde{M}(l,D) \ge \inf_{\substack{f(\hat{\mathbf{X}}^l_{\epsilon,\phi_s}|\mathbf{X}^l_{\epsilon,\phi_s}): \\ \mathbb{E}\{\overline{d}_{se}(\mathbf{X}^l_{\epsilon,\phi_s}, \hat{\mathbf{X}}^l_{\epsilon,\phi_s})\} \le \mathbb{E}\{\overline{d}_{se}(\mathbf{X}^l_{\epsilon,\phi_s}, \hat{\tilde{\mathbf{X}}}^l_{\epsilon,\phi_s})\}}} I(\mathbf{X}^l_{\epsilon,\phi_s}; \hat{\mathbf{X}}^l_{\epsilon,\phi_s}).$$
  3. By **asymptotic properties** of the **limit superior** and the condition $\limsup_{l\to\infty} \mathbb{E}\{\overline{d}_{se}(\mathbf{X}^l_{\epsilon,\phi_s}, \hat{\tilde{\mathbf{X}}}^l_{\epsilon,\phi_s})\} \le D$, we obtain
  $$R^{\phi_s}_{\epsilon}(D) \ge \limsup_{l\to\infty} \inf_{\substack{f(\hat{\mathbf{X}}^l_{\epsilon,\phi_s}|\mathbf{X}^l_{\epsilon,\phi_s}): \\ \mathbb{E}\{\overline{d}_{se}(\mathbf{X}^l_{\epsilon,\phi_s}, \hat{\mathbf{X}}^l_{\epsilon,\phi_s})\} \le D}} \frac{1}{l} I(\mathbf{X}^l_{\epsilon,\phi_s}; \hat{\mathbf{X}}^l_{\epsilon,\phi_s}).$$
  4. Minimizing $\frac{1}{l} I(\mathbf{X}^l_{\epsilon,\phi_s}; \hat{\mathbf{X}}^l_{\epsilon,\phi_s})$, the **optimal** $\mathbf{S}^l_{\epsilon,\phi_s}$ is found to be **Gaussian**, therefore
  $$R^{\phi_s}_{\epsilon}(D) \ge \limsup_{l\to\infty} \min_{\mathsf{C}_{\mathbf{S}^l_{\epsilon,\phi_s}} \in \mathcal{C}_{\mathbf{S}^l_{\epsilon,\phi_s}}} \frac{1}{2l} \log\left(\frac{\det(\mathsf{C}_{\mathbf{X}^l_{\epsilon,\phi_s}})}{\det(\mathsf{C}_{\mathbf{S}^l_{\epsilon,\phi_s}})}\right).$$

- **Achievability part**:

1. Denote the **optimal block of l reconstruction symbols** as $\{\hat{X}_{\epsilon,\phi_s}^{l,*}\}_{i=0}^{l-1} \equiv \hat{\mathbf{X}}_{\epsilon,\phi_s}^{l,*}$ and the **mutual information density rate** between $\mathbf{X}_{\epsilon,\phi_s}^{l}$ and $\hat{\mathbf{X}}_{\epsilon,\phi_s}^{l,*}$ as $\underline{Z_l(F_{X_\epsilon,\hat{X}_\epsilon^*}^{\phi_s})}$.

   The **mutual information density** $V_{\epsilon,l}^{\phi_s}$ between $\mathbf{X}_{\epsilon,\phi_s}^{l}$ and $\hat{\mathbf{X}}_{\epsilon,\phi_s}^{l,*}$ is

$$V_{\epsilon,l}^{\phi_s} = l \cdot \underline{Z_l(F_{X_\epsilon,\hat{X}_\epsilon^*}^{\phi_s})} \stackrel{\text{dist.}}{=} \underbrace{\frac{1}{2}\log\left(\frac{\det(\mathsf{C}_{\mathbf{X}_{\epsilon,\phi_s}^{l}})}{\det(\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^{l,*}})}\right)}_{\text{a constant}} + \underbrace{\frac{1}{2}\log(e)\cdot \underline{\tilde{V}_{\epsilon,\phi_s}^{l}}}_{\text{a RV}}.$$

2. Rewrite $\underline{\tilde{V}_{\epsilon,\phi_s}^{l}} \stackrel{\text{dist.}}{=} \sum_{i=0}^{2l-\tilde{l}_{\epsilon,l}-1} \lambda_{\epsilon,i}^{\phi_s} \cdot \underline{\underline{(\tilde{\gamma}_{\epsilon,i}^{\phi_s})^2}}$, where $2l-\tilde{l}_{\epsilon,l}-1 \in \mathcal{N}$, $\lambda_{\epsilon,i}^{\phi_s} \in \mathcal{R}$ and $\underline{\underline{(\tilde{\gamma}_{\epsilon,i}^{\phi_s})^2}}$ are **mutually independent central chi-square RVs** with **single degrees of freedom**.

3. Calculate $\mathbb{E}\{\underline{\tilde{V}_{\epsilon,\phi_s}^{l}}\}$ and $\mathrm{Var}\{\underline{\tilde{V}_{\epsilon,\phi_s}^{l}}\}$, then $\mathbb{E}\{\underline{Z_l(F_{X_\epsilon,\hat{X}_\epsilon^*}^{\phi_s})}\}$ and $\mathrm{Var}\{\underline{Z_l(F_{X_\epsilon,\hat{X}_\epsilon^*}^{\phi_s})}\}$ are obtained.

4. By the **Chebyshev inequality** and the **limit-superior in probability**, we obtain

$$R_\epsilon^{\phi_s}(D) \le \limsup_{l\to\infty} \min_{\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^{l}} \in \mathcal{C}_{\mathbf{S}_{\epsilon,\phi_s}^{l}}} \frac{1}{2l}\log\left(\frac{\det(\mathsf{C}_{\mathbf{X}_{\epsilon,\phi_s}^{l}})}{\det(\mathsf{C}_{\mathbf{S}_{\epsilon,\phi_s}^{l}})}\right).$$

- Combining the **converse** and the **achievability**, we obtain

$$R_{\epsilon}^{\phi_s}(D) = \limsup_{l\to\infty} \min_{\mathsf{C}_{\mathbf{S}^l_{\epsilon,\phi_s}} \in \mathcal{C}_{\mathbf{S}^l_{\epsilon,\phi_s}}} \frac{1}{2l} \log\left(\frac{\det(\mathsf{C}_{\mathbf{X}^l_{\epsilon,\phi_s}})}{\det(\mathsf{C}_{\mathbf{S}^l_{\epsilon,\phi_s}})}\right).$$

- Considering the **continuousness** of the source message generation and the **equidistribution theorem** [Kuipers and Niederreiter: 1974], [Coppel: 2009],
  - as the number of source messages goes to **infinity**, the **sampling phase** of each source message becomes *asymptotically* **uniformly distributed** over $[0, T_c)$.
  - Therefore, the **RDF** for compressing $X_{\epsilon}^{\phi_s}[i]$ is

$$R_{\epsilon}(D) = \frac{1}{T_c}\int_{\phi_s=0}^{T_c} R_{\epsilon}^{\phi_s}(D)\mathrm{d}\phi_s.$$

- **Remark 1**:
  - Intuitively, as $\epsilon \notin \mathcal{Q}$, the AF of $X_{\epsilon}^{\phi_s}[i]$ **never repeats itself** and *asymptotically* **exhibits all values** of $c_{X_c}(t, \lambda)$ as the **blocklength** l goes to **infinity**.
  - Therefore, we conjecture that $R_{\epsilon}^{\phi_s}(D)$ should **NOT** be affected by the **initial sampling phase** $\phi_s \in [0, T_c)$, therefore

$$R_{\epsilon}^{\phi_s}(D) = R_{\epsilon}^{0}(D).$$

- **Remark 2**:
  - When the sampling is **synchronous** ($\epsilon \in \mathcal{Q}$), the **RDF** is **affected** by the **initial sampling phase** $\phi_s \in [0, T_c)$, namely

$$R_{\epsilon}^{sync}(D) = R_{\epsilon}^{\phi_s}(D),$$

    which goes back to the result in [Kipnis et al.: *IEEE TIT* 2018].
  - In this case, if the **RDF** has to be **fixed** and **minimized**, we have

$$R_{\epsilon}^{sync}(D) = \min_{\phi_s \in [0, T_c)} R_{\epsilon}^{\phi_s}(D).$$

Outline

Summary

- We defined our problem, which is to characterize the **RDF for compressing DT WSACS Gaussian processes with memory**, where the source message generation is **continuous**.
- We reviewed the application of the **information-spectrum framework**, in the context of the RDF characterization for compressing an **arbitrary** DT process.
- We presented the result of the **RDF for compressing DT WSACS Gaussian processes with memory**, in which the **delay** between any two consecutive source messages is **NOT allowed**.
- In the future, we plan to prove that the term $R_{\epsilon}^{\phi_s}(D)$ is **irrelevant** with ϕ_s for **asynchronous sampling** and the **RDF for compressing DT WSACS Gaussian processes with memory**, where the **delay** between any two consecutive source messages is **allowed**.