van_der_Schaar \LAB

Time Series Diffusion in the Frequency Domain

Jonathan Crabbé†, Nicolas Huynh†, Jan Stanczuk & Mihaela van der Schaar

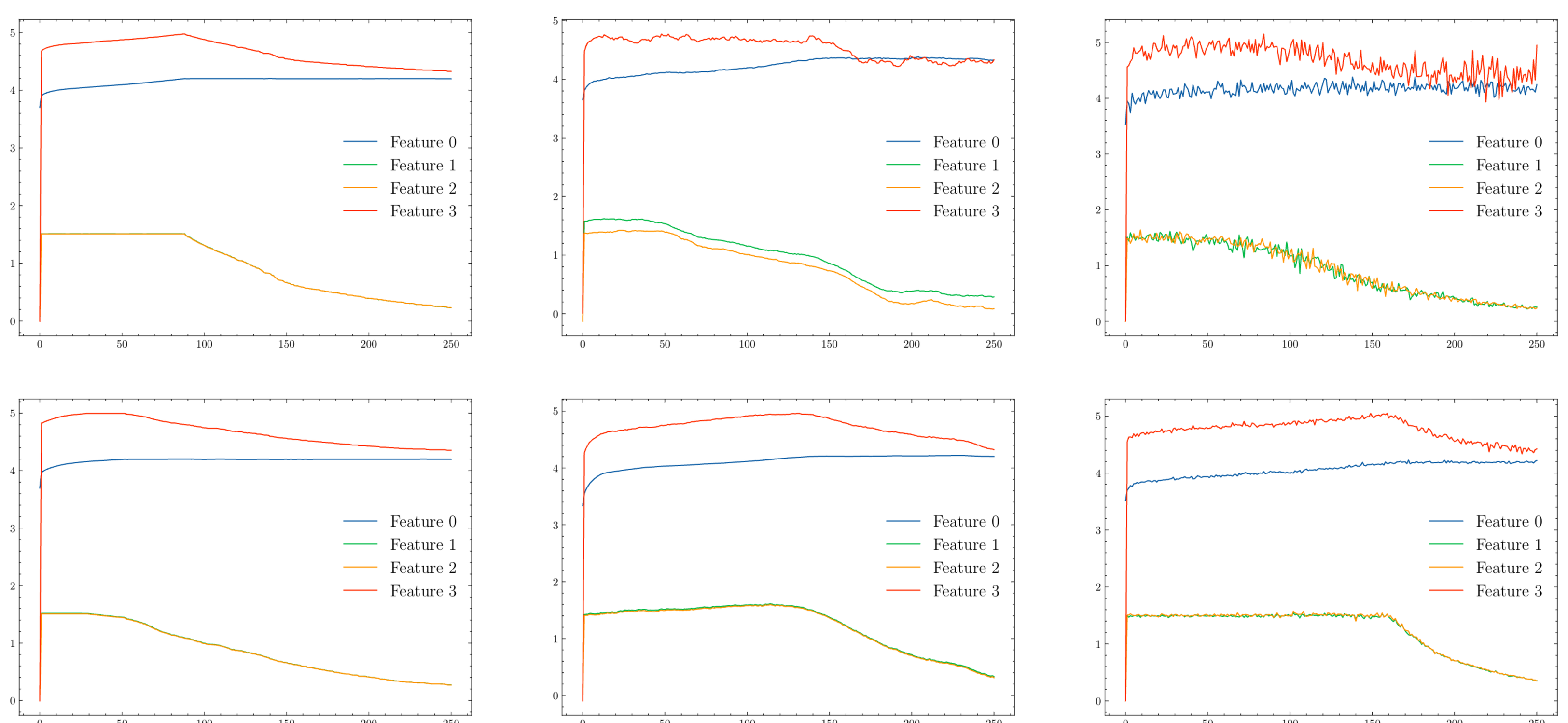
UNIVERSITY OF CAMBRIDGE

1. Motivation

Fourier analysis is a key tool in signal processing. Representing time series in the **frequency domain** is instrumental in applications like compression.

We explore the synergies between **Fourier analysis** and **diffusion models**. We compare models trained with **time** and **frequency** time series representations.

We find that representing time series in the **frequency domain tends to be preferable**. We explain this phenomenon by noting that many datasets are **localized** in the frequency domain.



Training Samples | Frequency Diffusion Samples | Time Diffusion Samples

2. Background Knowledge

Diffusion models. In the SDE formulation, diffusion is a stochastic process that gradually noises samples from the distribution of interest through $t \in [0, T]$

$$d\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, t)dt + g(t)d\boldsymbol{w}$$

In the limit $t \to T$, samples are well-described by an isotropic Gaussian. These can be transported back to the initial distribution through the reverse process:

$$d\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{x}, t)dt + g(t)d\hat{\boldsymbol{w}}$$

where the inversion is performed with the score $\boldsymbol{s}(\boldsymbol{x}, t) \coloneqq \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x})$ through $\boldsymbol{b}(\boldsymbol{x}, t) = \boldsymbol{f}(\boldsymbol{x}, t) - g^2(t)\, \boldsymbol{s}(\boldsymbol{x}, t)$. This score is learned by a model by minimizing

$$\mathcal{L}_{SM}(\boldsymbol{s}_{\boldsymbol{\theta}}, \boldsymbol{s}_{t|0}, \boldsymbol{x}, t) \coloneqq \|\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x}, t) - \boldsymbol{s}_{t|0}(\boldsymbol{x}, t)\|$$

Fourier analysis. If we consider a time series $\boldsymbol{x} = (\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}) \in \mathbb{R}^{N \cdot M}$, the discrete Fourier transform (DFT) is defined as

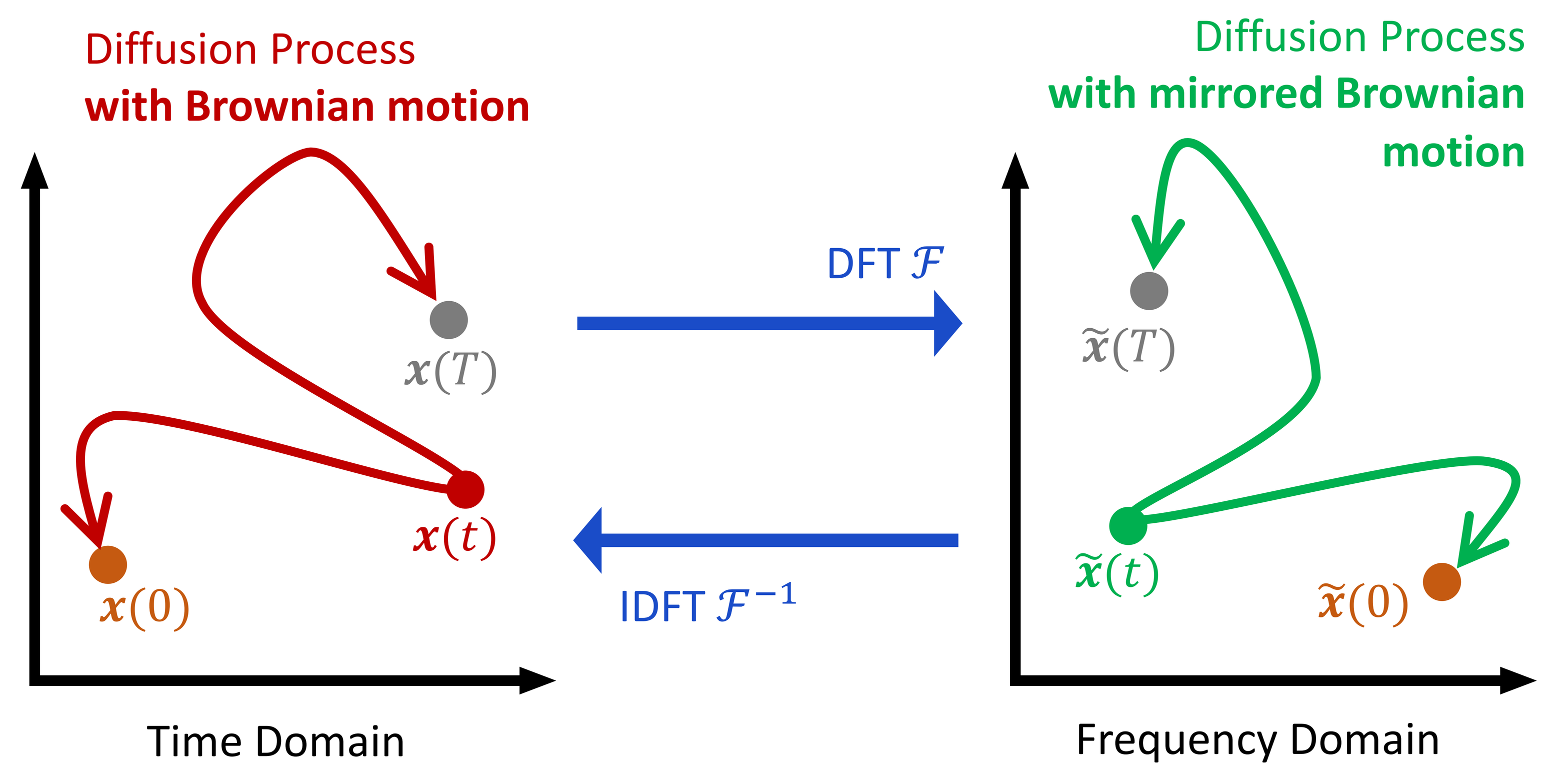
$$\tilde{\boldsymbol{x}}_\kappa \coloneqq \frac{1}{\sqrt{N}} \sum_{\tau=0}^{N} \boldsymbol{x}_\tau \, e^{-i \cdot \kappa \cdot 2\pi \cdot \frac{\tau}{N}}$$

This transform is invertible, as the original time series can be reconstructed as

$$\boldsymbol{x}_\tau = \frac{1}{\sqrt{N}} \sum_{\kappa=0}^{N} \tilde{\boldsymbol{x}}_\kappa \, e^{i \cdot \kappa \cdot 2\pi \cdot \frac{\kappa}{N}}$$

3. Time and Frequency Duality

What does a diffusion process performed in the time domain look like in the frequency domain? We prove that this is still a diffusion process with an important nuance: Brownian motions are replaced by **mirrored Brownian motions**, which have **redundant components**.



$$d\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, t)dt + \boldsymbol{g}(t)d\boldsymbol{w} \qquad d\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, t)dt + \boldsymbol{g}(t)d\tilde{\boldsymbol{v}}$$

$$d\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{x}, t)dt + \boldsymbol{g}(t)d\hat{\boldsymbol{w}} \qquad d\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{b}}(\tilde{\boldsymbol{x}}, t)dt + \boldsymbol{g}(t)d\tilde{\boldsymbol{v}}'$$

$$\tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, t) = \mathcal{F}[\boldsymbol{f}(\mathcal{F}^{-1}[\tilde{\boldsymbol{x}}], t)]$$
$$\tilde{\boldsymbol{b}}(\tilde{\boldsymbol{x}}, t) = \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{x}}, t) - g^2(t)\Lambda^2 \tilde{\boldsymbol{s}}(\tilde{\boldsymbol{x}}, t)$$

Hence, performing a diffusion process in either domain implies a **dual diffusion process** domain in the other domain.

How about the score-matching objective? Again, we show that we can associate to any score learned in either domain an **auxiliary score** in the other domain. Those two scores share the **same score matching objective**:

$$\boldsymbol{s}'_{\boldsymbol{\theta}}(\boldsymbol{x}, t) = \mathcal{F}^{-1}[\ \tilde{\boldsymbol{s}}_{\boldsymbol{\theta}}(\mathcal{F}[\boldsymbol{x}], t)] \qquad \tilde{\boldsymbol{s}}_{\boldsymbol{\theta}}(\tilde{\boldsymbol{x}}, t)$$

$$\mathcal{L}_{SM}(\boldsymbol{s}_{\boldsymbol{\theta}}, \boldsymbol{s}_{t|0}, \boldsymbol{x}, t) = \mathcal{L}_{SM}(\tilde{\boldsymbol{s}}_{\boldsymbol{\theta}}, \tilde{\boldsymbol{s}}_{t|0}, \tilde{\boldsymbol{x}}, t)$$

Time and frequency diffusion appear similar. However, we observe consistently **lower sliced Wasserstein distances** with the true distribution **for samples of the frequency models**. What is happening?

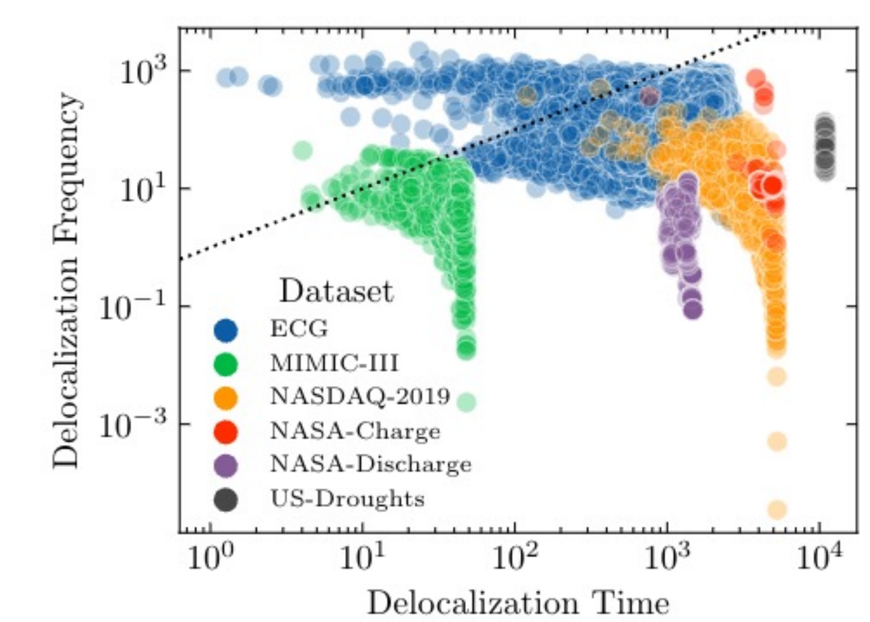
Dataset	*Metric Domain*	*Diffusion Domain*	
		Frequency	Time
ECG	Frequency	**0.012** ± **0.000**	0.020 ± 0.000
	Time	**0.015** ± **0.000**	0.021 ± 0.000
MIMIC-III	Frequency	**0.144** ± **0.004**	0.206 ± 0.006
	Time	**0.152** ± **0.004**	0.211 ± 0.006
NASDAQ-2019	Frequency	**45.812** ± **2.096**	64.056 ± 3.040
	Time	**43.602** ± **2.044**	60.512 ± 2.960
NASA-Charge	Frequency	**0.211** ± **0.008**	0.27 ± 0.006
	Time	**0.229** ± **0.008**	0.316 ± 0.008
NASA-Discharge	Frequency	**1.999** ± **0.084**	2.974 ± 0.134
	Time	**2.028** ± **0.082**	2.942 ± 0.134
US-Droughts	Frequency	**0.633** ± **0.018**	2.849 ± 0.090
	Time	**0.738** ± **0.020**	2.913 ± 0.092

4. The Uncertainty Principle Strikes Back

The **uncertainty principle** stipulates that a time series cannot be localized in the time and the frequency domain simultaneously. The localization in either domain can be quantified through the delocalization metrics

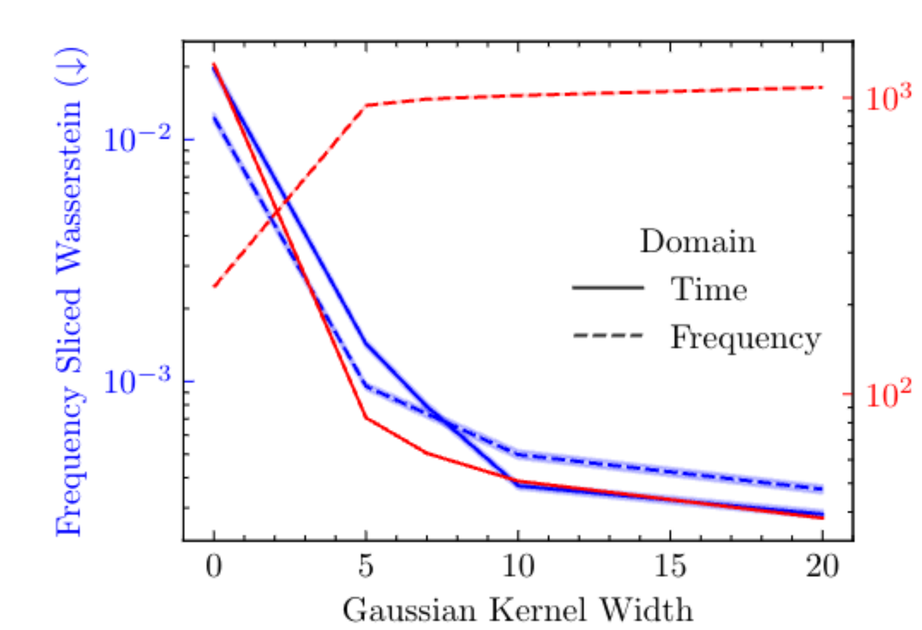
$$\Delta_{\text{time}}(\boldsymbol{x}) \coloneqq \min_\tau \frac{1}{\|\boldsymbol{x}\|^2} \sum_{\tau'} d_{cyc}(\tau, \tau') \|\boldsymbol{x}_{\tau'}\|^2$$
$$\Delta_{\text{freq}}(\tilde{\boldsymbol{x}}) \coloneqq \min_\kappa \frac{1}{\|\tilde{\boldsymbol{x}}\|^2} \sum_{\kappa'} d_{cyc}(\kappa, \kappa') \|\tilde{\boldsymbol{x}}_{\kappa'}\|^2$$

Measuring these on various time series reveals that time series are **substantially more localized** (i.e., less delocalized) **in the frequency domain**.



This suggests that a partial **explanations for the superiority** of the frequency diffusion models is the localization of time series in the frequency domain.

To verify this explanation, we synthetically delocalize time series in the frequency domain with spectral Gaussian convolutions. This intervention leads time diffusion model to outperform. This (partially) **confirms the explanation**.



5. More Information

Paper

SCAN ME

Code

SCAN ME