

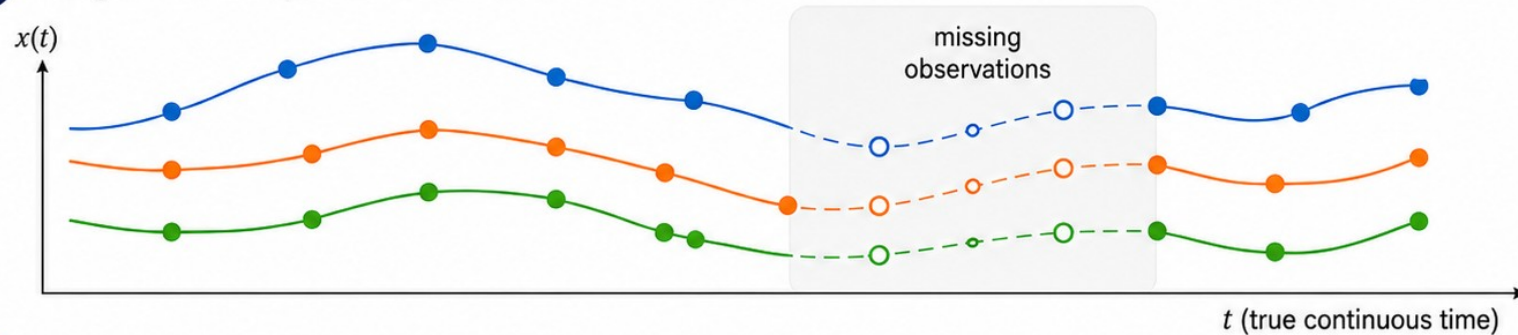
Latent Laplace Diffusion for Irregular Multivariate Time Series

Zinuo You, Jin Zheng, John Cartlidge,
Financial Engineering Lab
University of Bristol, Bristol, UK

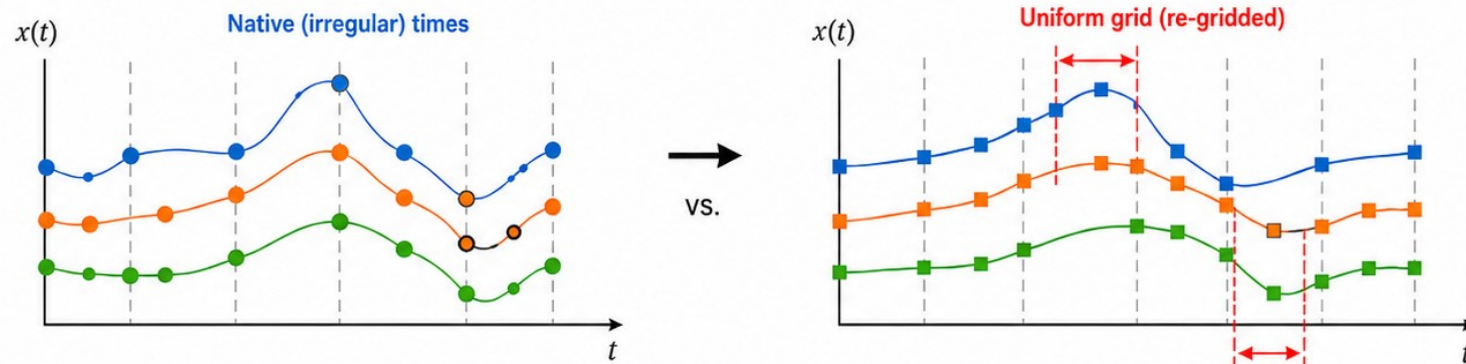
Email: zinuo.you@bristol.ac.uk

BACKGROUND

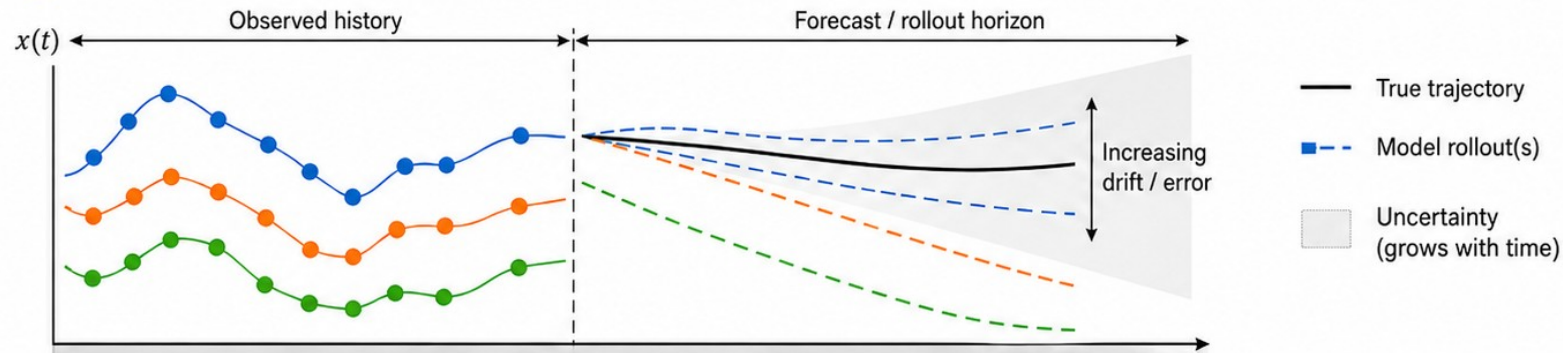
1 Irregular / incomplete observations



2 Re-gridding distorts dynamics

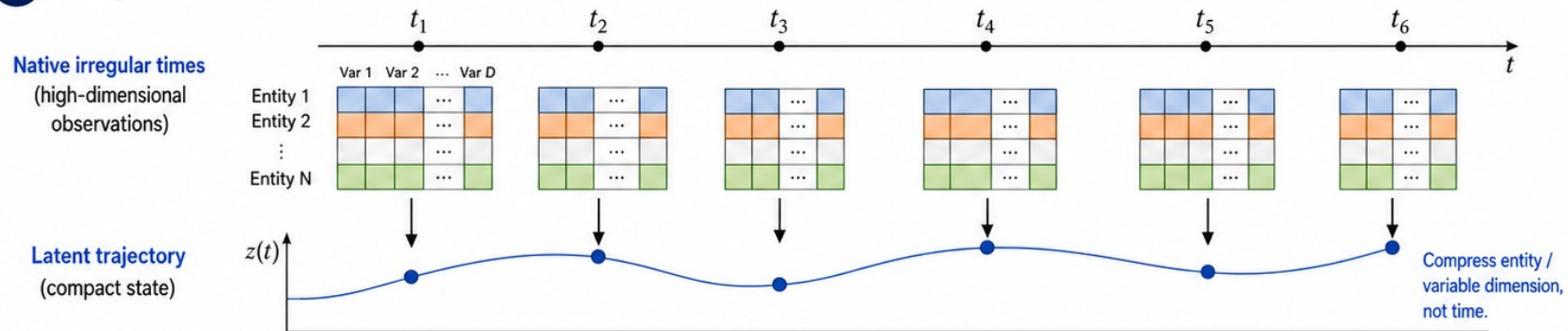


3 Long continuous-time rollouts can drift / diverge

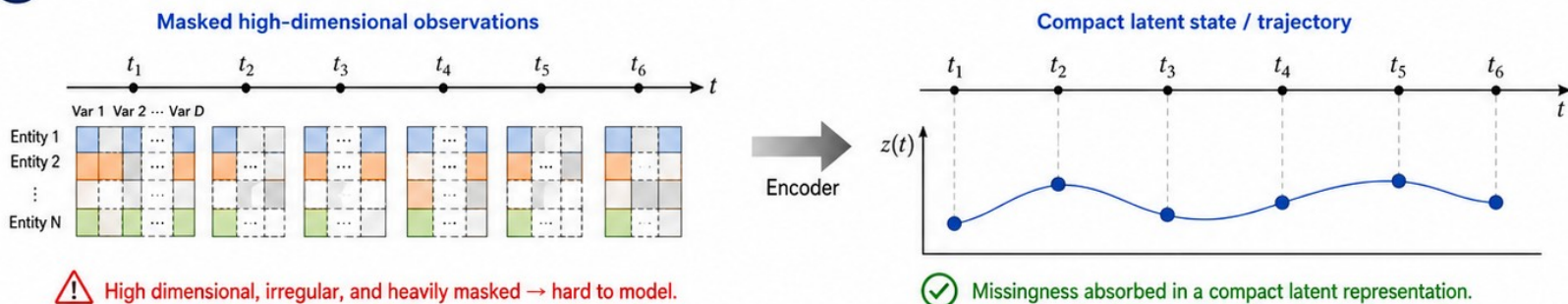


LATENT TRAJECTORY

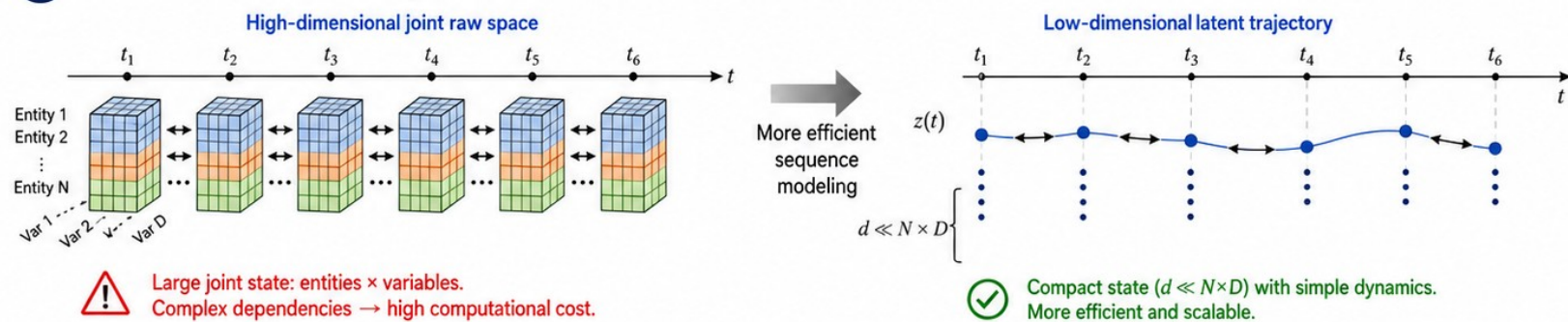
1 Respect native times.



2 Sparse / masked observations are difficult in raw space.



3 Joint raw-space modeling is inefficient.



LONG-TERM STABILITY & EQUIVALENT PARAMETERIZATION & RENEWAL-AVERAGING ON GAPS

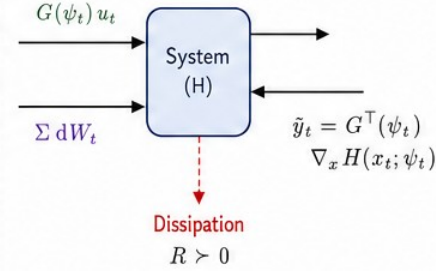
Stochastic port-Hamiltonian SDE

Auxiliary Hamiltonian state x_t in \mathbb{R}^{d_z} evolves as

$$dx_t = \underbrace{[(J - R)\nabla_x H(x_t; \psi_t)]}_{\text{deterministic port-Hamiltonian drift}} dt + \underbrace{G(\psi_t)u_t}_{\text{external input}} dt + \underbrace{\Sigma dW_t}_{\text{stochastic variation}}$$

Structure:

- $J^\top = -J$ (skew-symmetric interconnection)
- $R \succ 0$ (dissipation)
- $G(\psi_t)$ input matrix
- ΣdW_t stochastic variation



Hamiltonian / port variables

- $H(x_t; \psi_t)$ is the context-conditioned stored energy.
- $\nabla_x H$ is the energy gradient.
- ψ_t is history-dependent and piecewise constant between observation updates.

Special case between context updates

If ψ_t is fixed, $u_t = 0$, and $\Sigma = 0$, then

$$\frac{d}{dt} \mathbb{E}[H] = -\mathbb{E}[\nabla_x H^\top R \nabla_x H] \leq 0$$

Expected energy is non-increasing between context updates.

Expected energy balance (Itô)

$$\frac{d}{dt} \mathbb{E}[H] = \underbrace{\mathbb{E}[(\nabla_\psi H)^\top \dot{\psi}_t]}_{\text{context update term}} - \underbrace{\mathbb{E}[\nabla_x H^\top R \nabla_x H]}_{\text{dissipation } (\leq 0)} + \underbrace{\mathbb{E}[\tilde{y}_t^\top u_t]}_{\text{input power}} + \underbrace{\frac{1}{2} \mathbb{E}[\text{tr}(\Sigma \Sigma^\top \nabla_x^2 H)]}_{\text{Itô correction / noise}}$$

- $\nabla_x^2 H(x_t)$: Hessian of H
- $\text{tr}(\cdot)$: trace operator
- This is the instantaneous energy balance and does not depend on the sampling grid.

Between observation updates, $\dot{\psi}_t = 0$. In the absence of external inputs and noise, the expected energy decreases due to dissipation.

Long-horizon stability intuition

Local mean dynamics

$$\mathbb{E}[\delta x_t] = e^{A(t-t_0)} \mathbb{E}[\delta x_{t_0}] + \int_{t_0}^t e^{A(t-r)} B \delta u_r dr$$

$$A = (J - R) \nabla_x^2 H(\bar{x}_{t_0}; \psi_{t_0}), \quad B \approx G(\psi_{t_0})$$

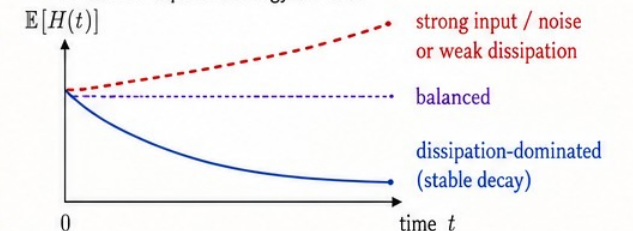
Forecasting view

For $t > t_0$, future u_t is unobserved. The history summary sets the initial state / residues, and generation follows autonomous transients.

Stable modal bias

poles: $-\rho_k \pm i\omega_k$, with $\rho_k > 0$
 \Rightarrow exponentially decaying mean modes

Intuition: expected energy vs. time



Takeaway: stochastic port-Hamiltonian structure provides a stability-inducing bias. Dissipation discourages energy growth, and enforcing $\rho_k > 0$ yields stable long-horizon mean dynamics consistent with the paper.

1 Local linearization and Laplace-domain view

$$d\delta x_t = A \delta x_t dt + B \delta u_t dt + \Sigma_{t_0} dW_t$$

$$A = (J - R) \nabla_x^2 H(\tilde{x}_{t_0}; \psi_{t_0}), \quad B \approx G(\psi_{t_0})$$

$$\mathbb{E}[\delta x_t] = e^{A(t-t_0)} \mathbb{E}[\delta x_{t_0}] + \int_{t_0}^t e^{A(t-r)} B \delta u_r dr$$

$$\mathbb{E}[\delta y_t] = C e^{A(t-t_0)} \mathbb{E}[\delta x_{t_0}] + (g * \delta u)(t)$$

$$g(t) = C e^{At} B, \quad G(s) = \mathcal{L}\{g\}(s) = C(sI - A)^{-1} B$$



For forecasting, future $\delta u(t)$ is unobserved, so generation focuses on autonomous transients from the history-conditioned initial state.

2 Stable modal parameterization

$$G(s) = \sum_{k=1}^K \frac{\omega_k c_k b_k^T}{s^2 + 2\rho_k s + (\rho_k^2 + \omega_k^2)}$$

$$c_k, b_k \in \mathbb{R}^{d_z}, \quad \rho_k, \omega_k > 0$$

Real state-space block for one mode k

$$A_k = \begin{bmatrix} -\rho_k & -\omega_k \\ \omega_k & -\rho_k \end{bmatrix}, \quad B_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix} b_k^T, \quad C_k = c_k \begin{bmatrix} -1 & 0 \end{bmatrix}$$

$$\text{eig}(A_k) = -\rho_k \pm i\omega_k$$

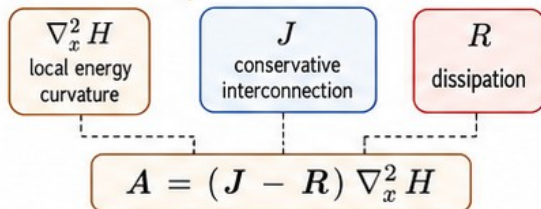
Block assembly over K modes

$$A = \text{blkdiag}(A_1, \dots, A_K), \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_K \end{bmatrix}, \quad C = [C_1 \ \dots \ C_K]$$

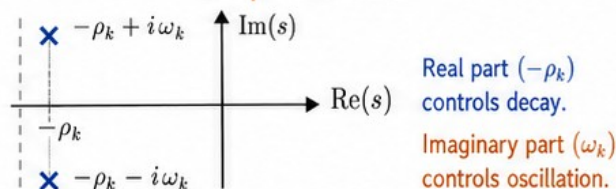
$$\rho_k > 0 \Rightarrow A \text{ is Hurwitz} \Rightarrow \text{exponentially decaying mean modes.}$$

3 Why complex poles are an equivalent representation

A. Physics source



B. Spectral effect



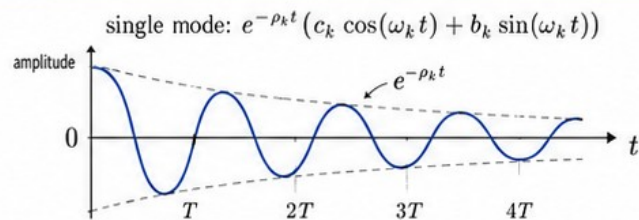
C. Equivalent realization

- The Hessian and (J, R) determine the local Jacobian A .
- The resolvent $C(sI - A)^{-1}B$ has a modal / partial-fraction expansion.
- Thus learned poles and residues summarize the same local linear dynamics without explicitly computing the full Hessian or Jacobian.

$$(\nabla^2 H, J, R) \rightarrow A \rightarrow G(s) \rightarrow \{\rho_k, \omega_k, c_k, b_k\}$$

4 Physics view and intuition

- ω_k reflects conservative oscillation frequencies induced by J acting on local curvature.
- ρ_k captures dissipative decay rates driven by R .
- Residues b_k, c_k encode how the history-conditioned latent state projects onto the modal basis.



$$\tilde{z}_0(\tilde{t}_r) = \sum_{k=1}^K e^{-\hat{\rho}_k \tilde{t}_r} [\hat{c}_k \cos(\hat{\omega}_k \tilde{t}_r) + \hat{b}_k \sin(\hat{\omega}_k \tilde{t}_r)]$$

Closed-form evaluation at all query times enables horizon-wide generation without physical-time stepping.

Takeaway: the learned complex poles are not arbitrary—they are a stable modal realization of the local port-Hamiltonian mean dynamics, equivalently capturing how local energy curvature and dissipation shape latent trajectories.

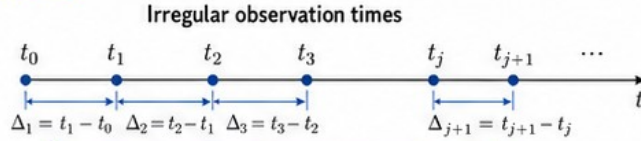
1 Continuous-time mode under irregular sampling

Single continuous-time mode (complex pole):

$$s_k = -\rho_k + i\omega_k$$

Real 2×2 block for one mode:

$$A_k = \begin{bmatrix} -\rho_k & -\omega_k \\ \omega_k & -\rho_k \end{bmatrix}$$



Irregular gaps change how a continuous-time mode appears in event index j .

Event-to-event update (complex coordinate):

$$\zeta_{j+1}^{(k)} = e^{s_k \Delta_{j+1}} \zeta_j^{(k)}$$

Equivalent update (real 2D coordinate):

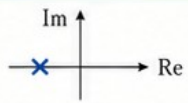
$$\xi_{j+1}^{(k)} = e^{A_k \Delta_{j+1}} \xi_j^{(k)}$$

The discrete (event-indexed) dynamics depend on the actual realization of gaps Δ_j . Even with the same intrinsic pole s_k , different gap patterns yield different apparent behavior.

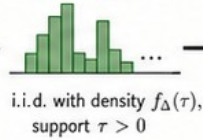
2 Renewal averaging gives an effective event-domain pole

Continuous-time pole

$$s_k = -\rho_k + i\omega_k$$



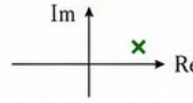
Random gaps Δ



Renewal averaging ($\mathbb{E}[\cdot]$)

Effective event-domain pole

$$\bar{s}_k = -\bar{\rho}_k + i\bar{\omega}_k$$



Complex coordinate (mean evolution under i.i.d. gaps):

$$\mathbb{E}[\zeta_j^{(k)}] = (\mathbb{E}[e^{s_k \Delta}])^j \zeta_0^{(k)} = \lambda_k^j \zeta_0^{(k)} = e^{\bar{s}_k j} \zeta_0^{(k)}$$

$$\lambda_k = \mathbb{E}[e^{s_k \Delta}], \quad \bar{s}_k = \log(\lambda_k) = -\bar{\rho}_k + i\bar{\omega}_k$$

Real coordinate (mean evolution):

$$\mathbb{E}[\xi_j^{(k)}] = \Phi_k^j \xi_0^{(k)}, \quad \Phi_k := \mathbb{E}[e^{A_k \Delta}]$$

If the principal matrix logarithm exists,

$$\Phi_k = e^{\bar{A}_k}, \quad \text{so } \mathbb{E}[\xi_j^{(k)}] = e^{\bar{A}_k j} \xi_0^{(k)}$$

The event-domain pole depends jointly on intrinsic dynamics and the gap distribution.

3 Stability and effect of gap variability

Mean stability is preserved

$$|\lambda_k| = |\mathbb{E}[e^{s_k \Delta}]| \leq \mathbb{E}[|e^{s_k \Delta}|] \leq \mathbb{E}[e^{\text{Re}(s_k) \Delta}] \leq 1$$

$$\implies \text{Re}(\bar{s}_k) = \log |\lambda_k| \leq 0$$

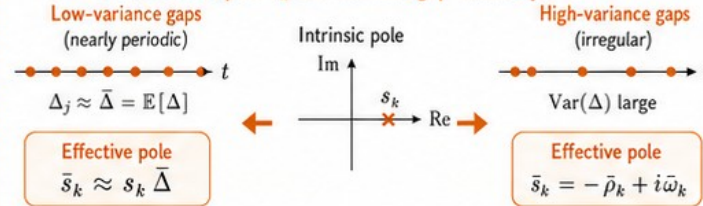
Random sampling preserves mean stability when $\text{Re}(s_k) < 0$ and $\mathbb{P}(\Delta > 0) > 0$.

Second-order approximation (small-to-moderate variability)

$$\bar{s}_k \approx s_k \mathbb{E}[\Delta] + \frac{1}{2} s_k^2 \text{Var}(\Delta)$$

Gap mean and gap variance alter effective damping and phase.

Same intrinsic pole s_k but different gap variability



Irregular sampling itself can make event-domain decay appear stronger (more negative $\text{Re}(\bar{s}_k)$) or weaker (less negative).

4 Why the summarizer must encode times and gap patterns

Raw irregular history

Values y :	y_0	y_1	—	y_2	y_2	—	y_4	y_5	—	y_6	...
Mask m :	1	1	0	1	1	0	1	1	0	1	...
Times t :	t_0	t_1		t_2	t_3		t_4	t_5		t_6	...
Gaps Δt (uneven):	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	...				

Gap-aware summarizer (history encoder)

Token types / cues the summarizer should encode

- 1 Timestamps / Δt
actual timing of events
- 2 Gap variability
irregularity pattern / statistics
- 3 Masks / observed values
which values are observed

Continuous-time modal parameters (goal of inference)

$$(\hat{\rho}_k, \hat{\omega}_k, \hat{c}_k, \hat{b}_k)$$

$$k = 1, \dots, K$$

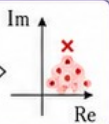
Effective event-domain poles entangle intrinsic poles with sampling statistics

$$\lambda_k = \mathbb{E}[e^{s_k \Delta}], \quad \bar{s}_k = \log(\lambda_k)$$

Therefore, the model must condition on timing information to infer gap-robust continuous-time parameters $(\hat{\rho}_k, \hat{\omega}_k, \hat{c}_k, \hat{b}_k)$.

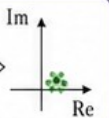
WITHOUT timing / gap encoding

Confuses intrinsic dynamics with sampling artifacts.



WITH gap-aware summarization

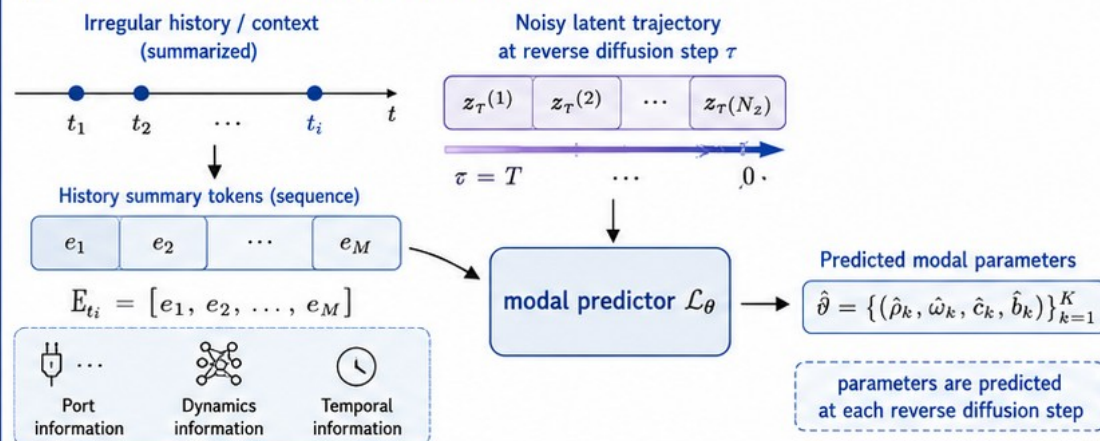
Better recovery of consistent continuous-time poles.



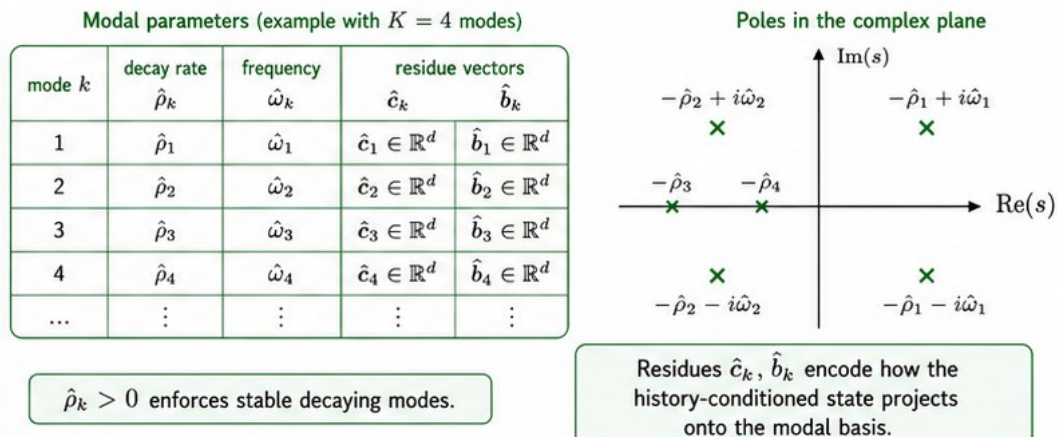
Takeaway: Renewal averaging shows that irregular gaps reshape the observed event-domain dynamics.

This is why the history summarizer must encode timestamps and gap patterns: to disentangle sampling effects from the intrinsic continuous-time latent dynamics.

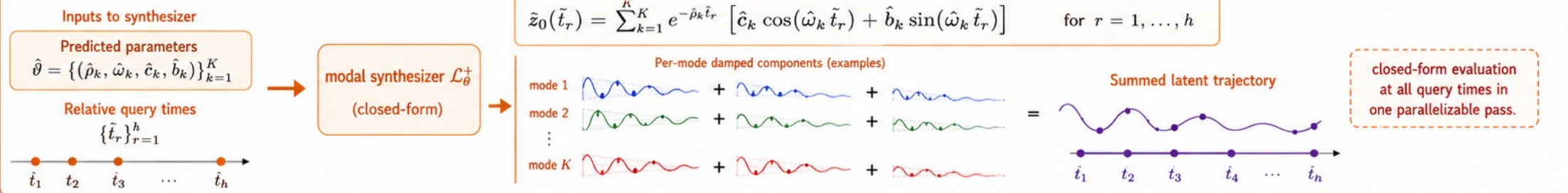
1 Conditioned modal parameter prediction



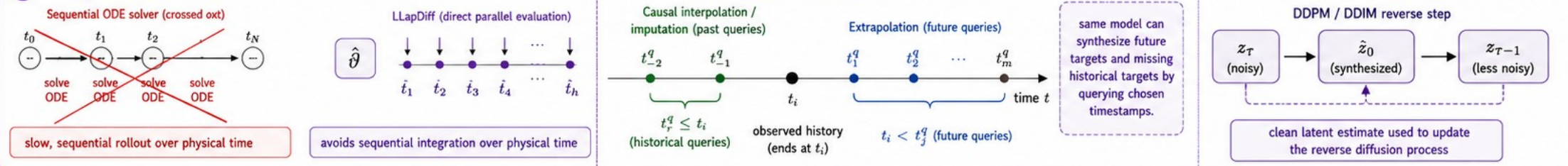
2 What the predictor outputs



3 Analytical horizon-wise synthesis

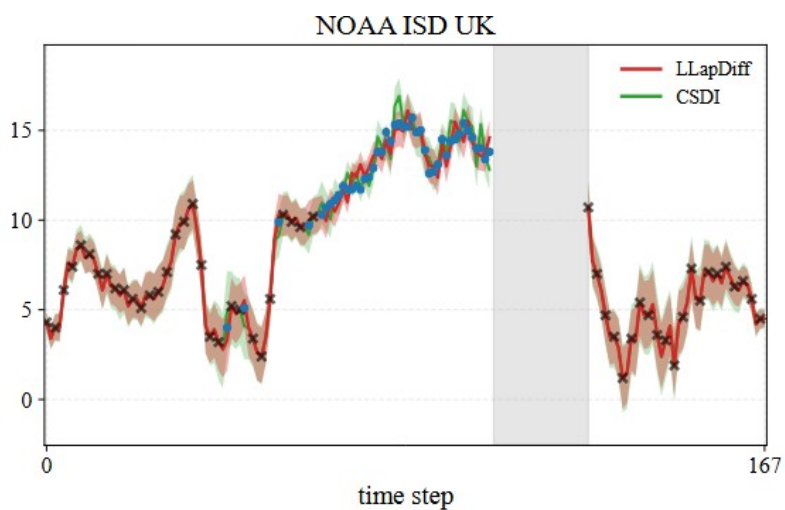
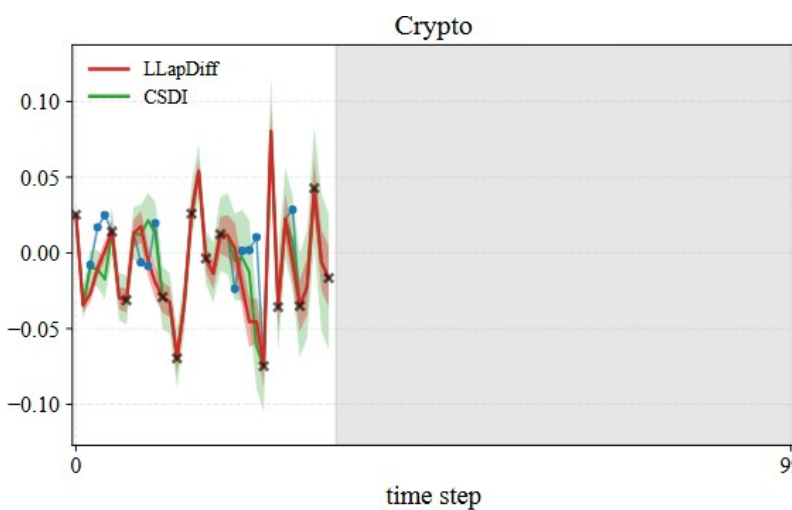
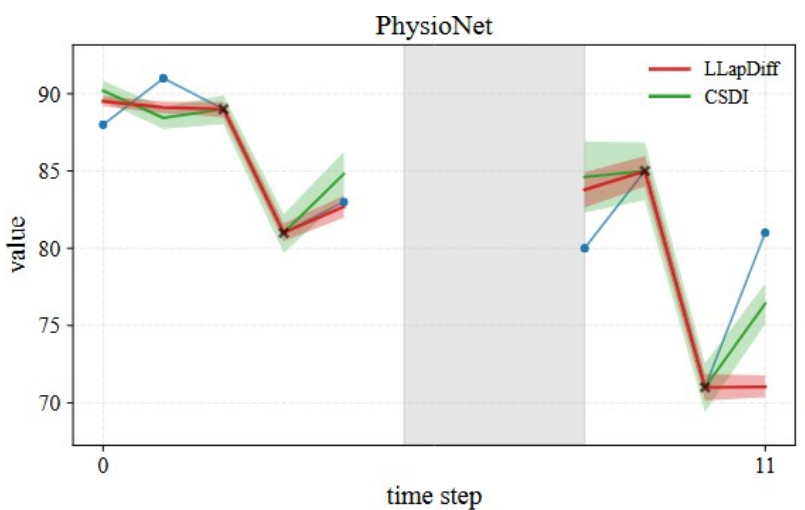
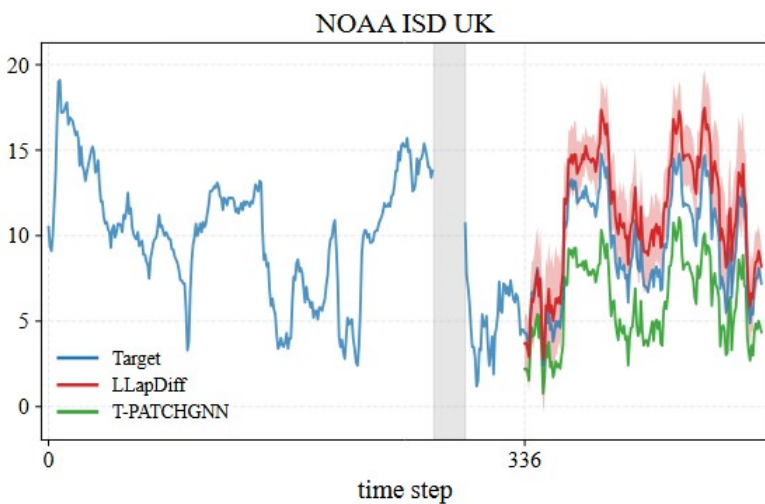
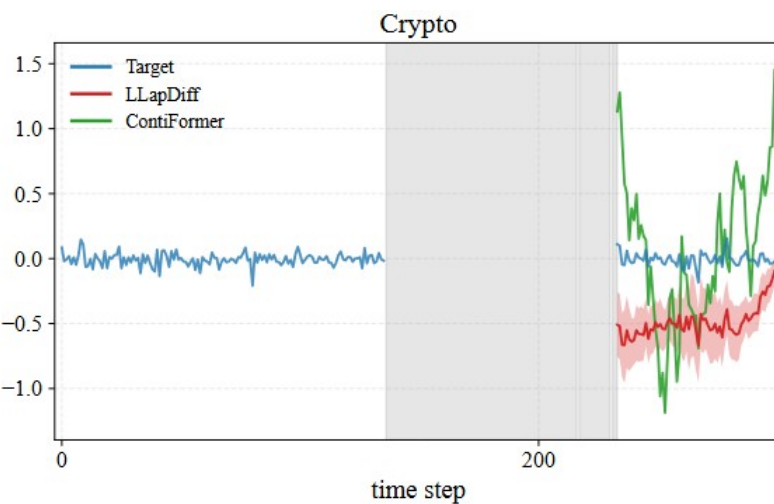
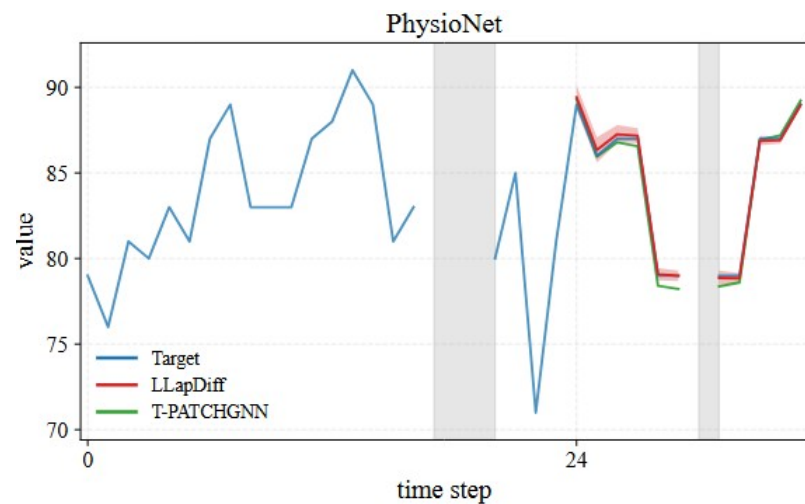


4 Why this is useful



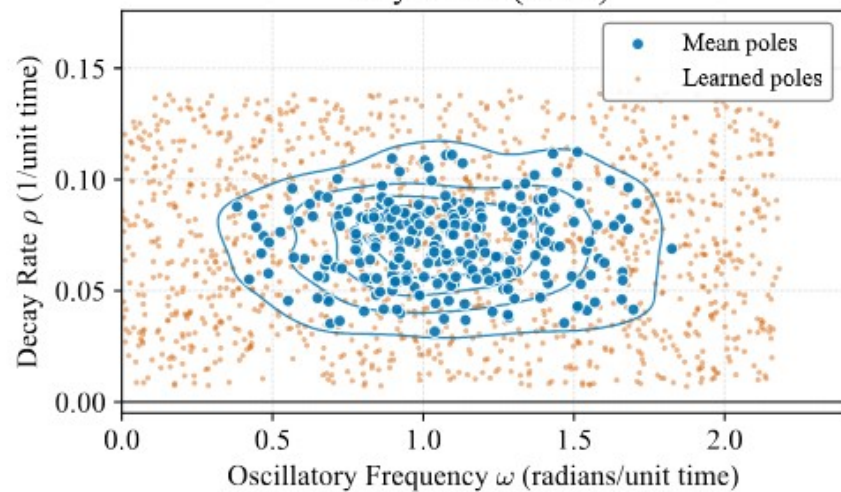
Takeaway: LLapDiff predicts stable modal parameters from the noisy latent and history summary, then synthesizes the entire latent trajectory analytically at arbitrary query times—enabling horizon-wide generation without physical-time stepping and supporting both extrapolation and causal interpolation.

QUALITATIVE ILLUSTRATIONS

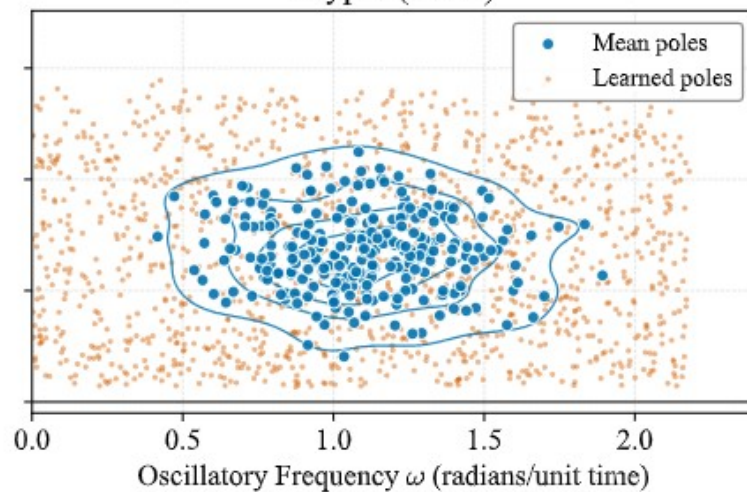


LEARNED POLES DYNAMICS

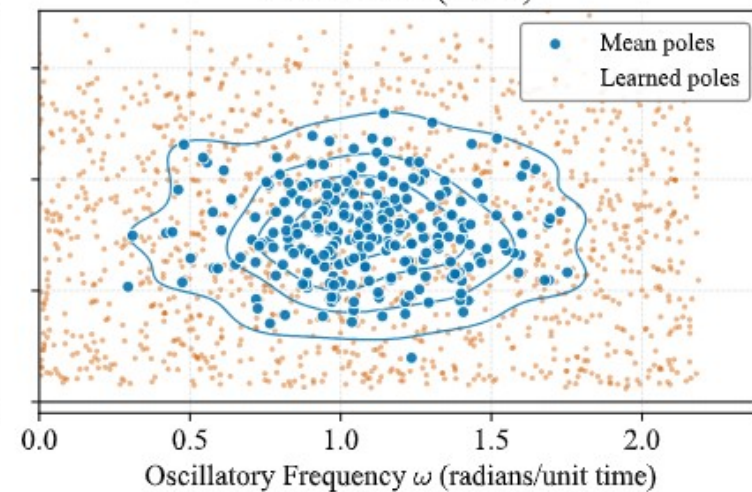
PhysioNet (Cond)



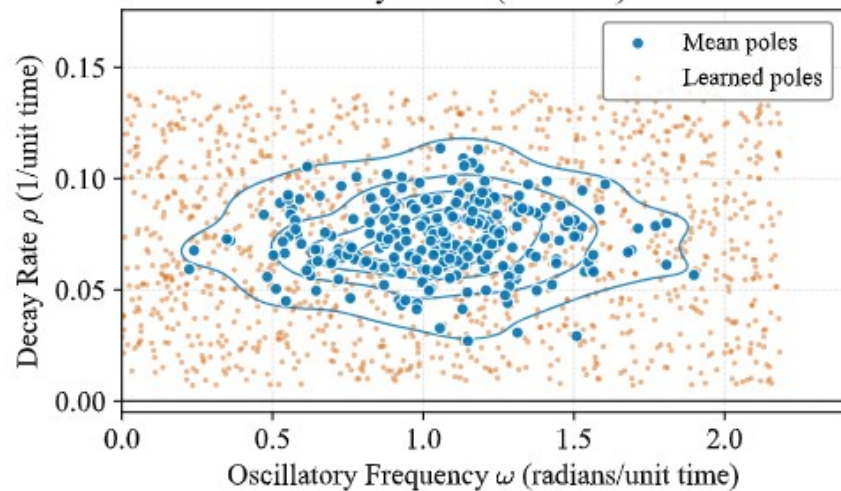
Crypto (Cond)



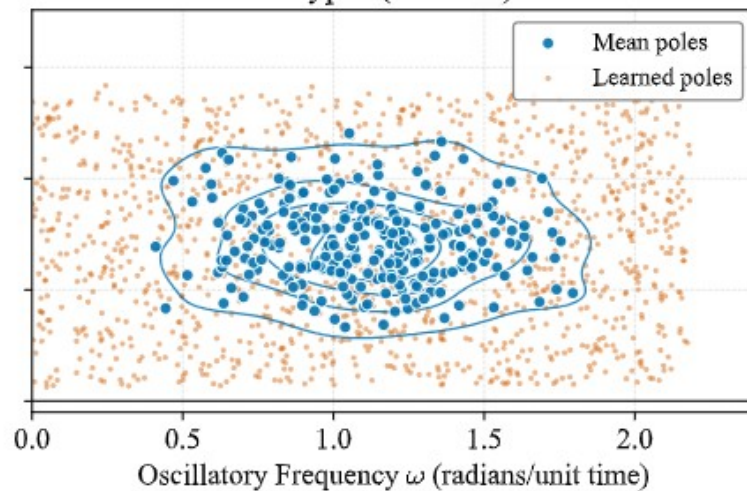
NOAA UK (Cond)



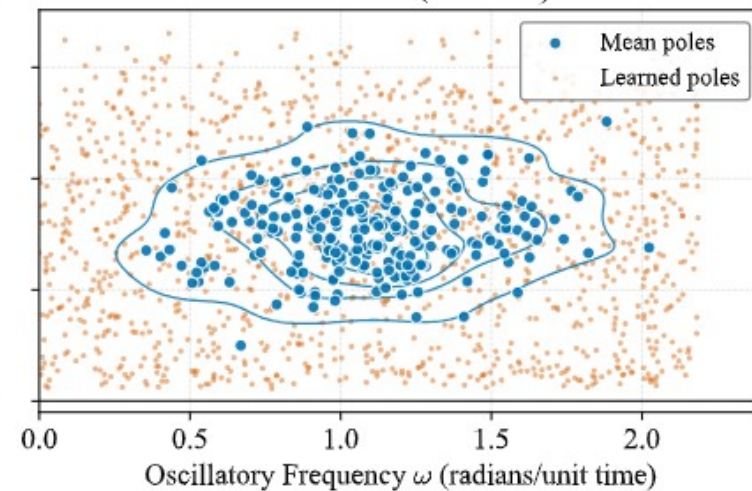
PhysioNet (Uncond)



Crypto (Uncond)

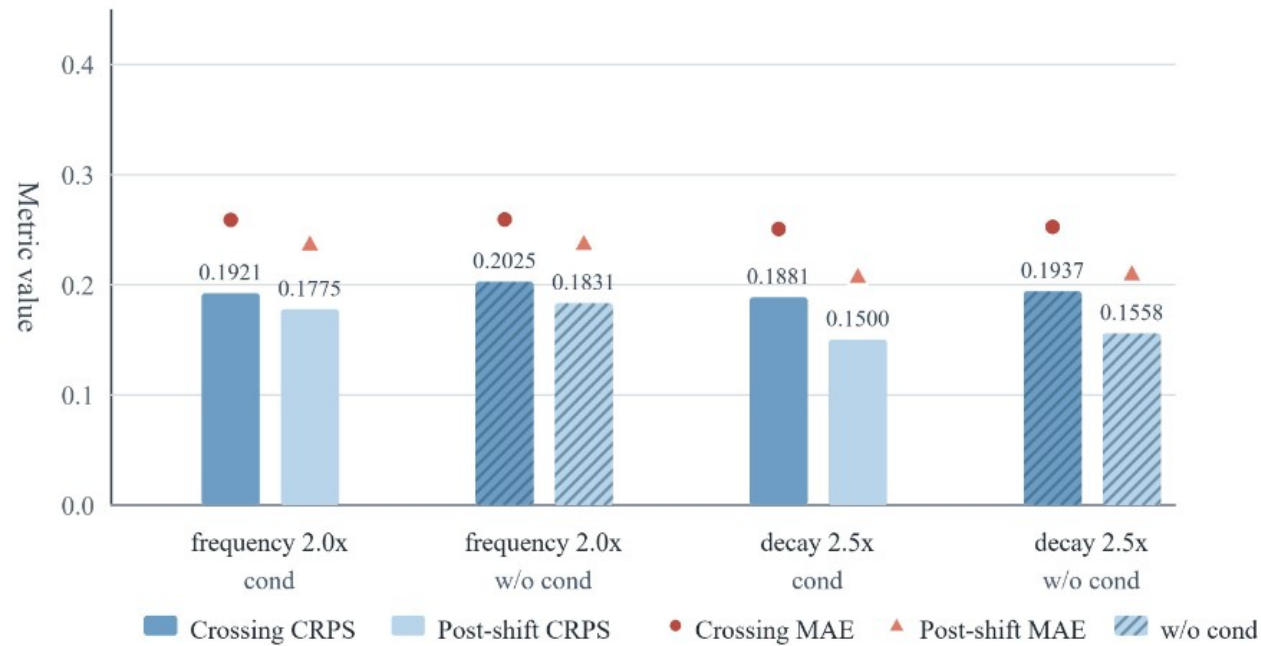


NOAA UK (Uncond)

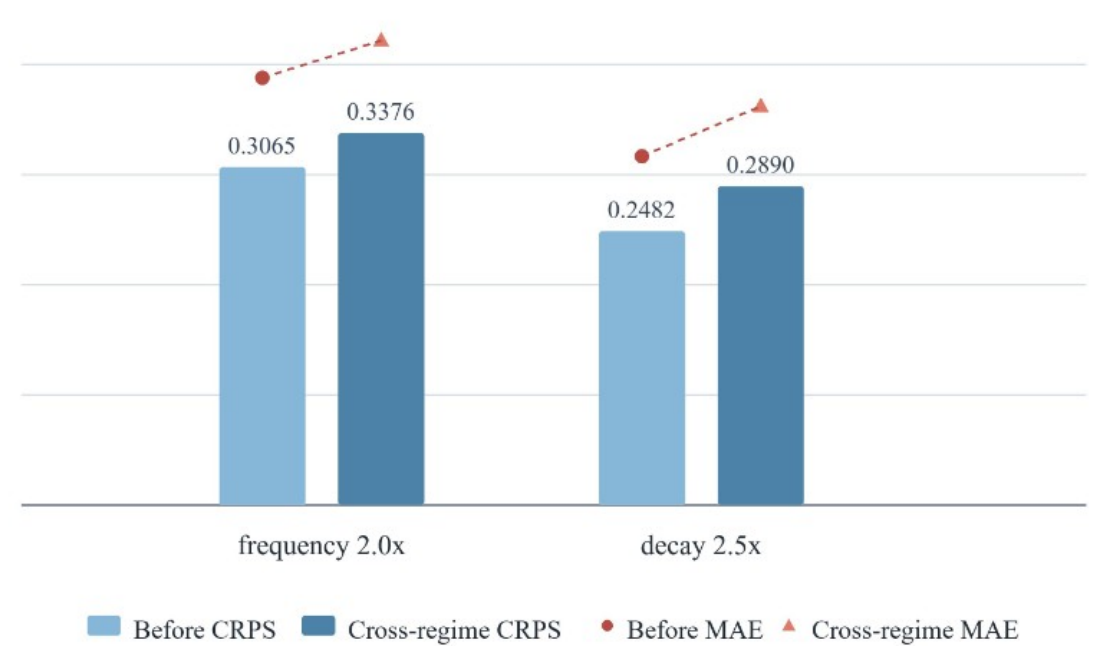


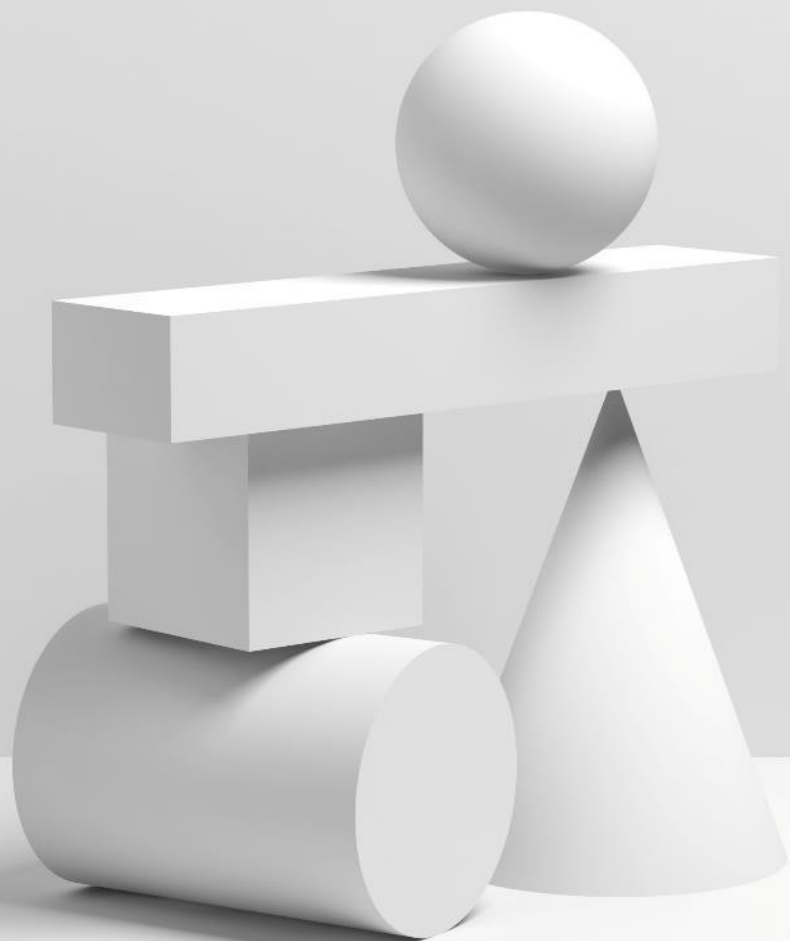
ROBUSTNESS UNDER MISSINGNESS AND REGIME SHIFTS

A. Strict unseen-regime extrapolation



B. Boundary-crossing robustness





THANK YOU!