WGFINNs: Weak-form GENERIC formalism informed neural networks

Jun Sur Richard Park (박준서)

Center for AI and Natural Sciences
Korea Institute for Advanced Study (KIAS, 고등과학원)



Joint work with Prof. Yeonjong Shin, Auroni Hashin (NCSU),
Dr. Youngsoo Choi, Dr. Siu Wun Cheng (LLNL)

KIAS Center for AI and Natural Sciences 2024 Fall Workshop

November 6, 2024

Setup: Data driven discovery of dynamics

- Learning a dynamical system (strong form)

$$\dot{\boldsymbol{z}}(t) = F(\boldsymbol{z}(t)),\ t > 0,\ \boldsymbol{z}(0) = \boldsymbol{z}_0$$

where F is unknown.

- Goal: Approximate F by a **neural network** F_{NN} using **trajectory data**

$$\left\{(t_0, \boldsymbol{z}_0^{(k)}), (t_1, \boldsymbol{z}(t_1, \boldsymbol{z}_0^{(k)})), \ldots, (t_N, \boldsymbol{z}(t_N, \boldsymbol{z}_0^{(k)}))\right\}, k = 1, \ldots, K$$

- Given NN parameter θ, generate trajectories from $\dot{\boldsymbol{z}}(t) = F_{\text{NN}}(\boldsymbol{z}(t); \theta)$

$$\left\{(t_0, \boldsymbol{z}_0^{(k)}), (t_1, \boldsymbol{z}_{\text{NN}}^{\theta}(t_1, \boldsymbol{z}_0^{(k)})), \ldots, (t_N, \boldsymbol{z}_{\text{NN}}^{\theta}(t_N, \boldsymbol{z}_0^{(k)}))\right\}, k = 1, \ldots, K$$

- Loss functional

$$\mathcal{L}(\theta) = \frac{1}{KN} \sum_{k=1}^{K} \sum_{i=1}^{N} ||\boldsymbol{z}_{\text{NN}}^{\theta}(t_i, \boldsymbol{z}_0^{(k)}) - \boldsymbol{z}(t_i, \boldsymbol{z}_0^{(k)})||^2$$

Setup: Data driven discovery of dynamics

- Learning a dynamical system (strong form)

$$\dot{\boldsymbol{z}}(t) = F(\boldsymbol{z}(t)),\ t > 0,\ \boldsymbol{z}(0) = \boldsymbol{z}_0$$

where F is unknown.

- Loss functional

$$\mathcal{L}(\theta) = \frac{1}{KN} \sum_{k=1}^{K} \sum_{i=1}^{N} ||\boldsymbol{z}_{\mathrm{NN}}^{\theta}(t_i, \boldsymbol{z}_0^{(k)}) - \boldsymbol{z}(t_i, \boldsymbol{z}_0^{(k)})||^2$$

- Challenge: Inaccuracy with noisy data

$$\boldsymbol{z}_{\mathrm{NN}}^{\theta}(t_i, \boldsymbol{z}_0^{(k)}) = \boldsymbol{z}(t_{i-1}, \boldsymbol{z}_0^{(k)}) + \Delta t F_{\mathrm{NN}}(\boldsymbol{z}(t_{i-1}, \boldsymbol{z}_0^{(k)}), \theta)$$

Setup: Data driven discovery of dynamics

- Learning a dynamical system (strong form)

$$\dot{\boldsymbol{z}}(t) = F(\boldsymbol{z}(t)),\ t > 0,\ \boldsymbol{z}(0) = \boldsymbol{z}_0$$

where F is unknown.

- Loss functional

$$\mathcal{L}(\theta) = \frac{1}{KN}\sum_{k=1}^{K}\sum_{i=1}^{N} \left\| \boldsymbol{z}_{\mathrm{NN}}^{\theta}(t_i, \boldsymbol{z}_0^{(k)}) - \boldsymbol{z}(t_i, \boldsymbol{z}_0^{(k)}) \right\|^2$$

NOISE

- Challenge: Inaccuracy with noisy data

$$\boldsymbol{z}_{\mathrm{NN}}^{\theta}(t_i, \boldsymbol{z}_0^{(k)}) = \boldsymbol{z}(t_{i-1}, \boldsymbol{z}_0^{(k)}) + \Delta t F_{\mathrm{NN}}(\boldsymbol{z}(t_{i-1}, \boldsymbol{z}_0^{(k)}), \theta)$$

NOISE NOISE

Setup: Data driven discovery of dynamics (weak form)

- Learning a dynamical system (weak form)

$$\int_0^T \dot{\boldsymbol{z}}\phi_j(t)dt = \int_0^T F(\boldsymbol{z}(t))\phi_j(t)dt,\ j = 1, \ldots, J$$

where F is unknown, ϕ_j is absolutely continuous test function.

Setup: Data driven discovery of dynamics (weak form)

- Learning a dynamical system (weak form)

$$-\int_0^T \boldsymbol{z}\dot{\phi}_j(t)dt = \int_0^T F(\boldsymbol{z}(t))\phi_j(t)dt,\ j = 1, \ldots, J$$

where F is unknown, ϕ_j is an absolutely continuous test function.

Setup: Data driven discovery of dynamics (weak form)

- Learning a dynamical system (weak form)

$$-\int_0^T \boldsymbol{z}\dot{\phi}_j(t)dt = \int_0^T F(\boldsymbol{z}(t))\phi_j(t)dt,\ j = 1, \ldots, J$$

where F is unknown, ϕ_j is an absolutely continuous test function.

- Goal: Approximate F by a **neural network** F_{NN} using **trajectory data**

$$\left\{(t_0, \boldsymbol{z}_0^{(k)}), (t_1, \boldsymbol{z}(t_1, \boldsymbol{z}_0^{(k)})), \ldots, (t_N, \boldsymbol{z}(t_N, \boldsymbol{z}_0^{(k)}))\right\},\ k = 1, \ldots, K$$

- Loss functional based on the discretization of

$$\left| -\int_0^T \boldsymbol{z}(t)\dot{\phi}_j(t)dt - \int_0^T F_{\mathrm{NN}}(\boldsymbol{z}(t))\phi_j(t)dt \right|,\ j = 1, \ldots, J$$

- Wishlist: Noise in $\boldsymbol{z}$ averaged out with the integrations.

Setup: Data driven discovery of dynamics (weak form)

- Learning a dynamical system (weak form)

$$-\int_0^T \boldsymbol{z}\dot{\phi}_j(t)dt = \int_0^T F(\boldsymbol{z}(t))\phi_j(t)dt,\ j = 1, \ldots, J$$

where F is unknown, ϕ_j is an absolutely continuous test function.

- Goal: Approximate F by a **neural network** F_{NN} using **trajectory data**

$$\left\{(t_0, \boldsymbol{z}_0^{(k)}), (t_1, \boldsymbol{z}(t_1, \boldsymbol{z}_0^{(k)})), \ldots, (t_N, \boldsymbol{z}(t_N, \boldsymbol{z}_0^{(k)}))\right\},\ k = 1, \ldots, K$$

- Loss functional based on the discretization of

$$\left| -\int_0^T \boldsymbol{z}(t)\dot{\phi}_j(t)dt - \int_0^T F_{\mathrm{NN}}(\boldsymbol{z}(t))\phi_j(t)dt \right|,\ j = 1, \ldots, J$$

- Question: How do we model F_{NN}?

Setup: Data driven discovery of dynamics (weak form)

- Learning a dynamical system (weak form)

$$-\int_0^T \boldsymbol{z}\dot{\phi}_j(t)dt = \int_0^T F(\boldsymbol{z}(t))\phi_j(t)dt,\ j = 1, \ldots, J$$

where F is unknown, ϕ_j is an absolutely continuous test function.

- Goal: Approximate F by a **neural network** F_{NN} using **trajectory data**

$$\left\{(t_0, \boldsymbol{z}_0^{(k)}), (t_1, \boldsymbol{z}(t_1, \boldsymbol{z}_0^{(k)})), \ldots, (t_N, \boldsymbol{z}(t_N, \boldsymbol{z}_0^{(k)}))\right\},\ k = 1, \ldots, K$$

- Loss functional based on the discretization of

$$\left| -\int_0^T \boldsymbol{z}(t)\dot{\phi}_j(t)dt - \int_0^T F_{\mathrm{NN}}(\boldsymbol{z}(t))\phi_j(t)dt \right|,\ j = 1, \ldots, J$$

- Question: How do we model F_{NN}? ⟶ Physics! (Thermodynamics)

GENERIC formalism (Strong form)

- The General Equation for Non-Equilibrium Reversible-Irreversible Coupling (GENERIC) formalism (Ottinger HC. 2005)

$$\dot{\boldsymbol{z}}(t) = L(\boldsymbol{z})\nabla E(\boldsymbol{z}) + M(\boldsymbol{z})\nabla S(\boldsymbol{z})$$

subject to $(1)\ L(\boldsymbol{z})\nabla S(\boldsymbol{z}) = 0$ $(2)\ M(\boldsymbol{z})\nabla E(\boldsymbol{z}) = 0$

$(3)\ L = -L^{\top}$ $(4)\ M$ is symmetric positive semidefinite

- $\boldsymbol{E}, \boldsymbol{S}$: $\mathbb{R}^d \to \mathbb{R}$: Total energy and entropy of the system
- $\boldsymbol{L}$: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$: Poisson matrix, skew-symmetric,
- $\boldsymbol{M}$: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$: Friction matrix, symmetric and positive semi-definite

- All functions E, S, L and M are modeled by NNs (GFINNs, Zhang at al. 2022)

GENERIC formalism (Strong form)

- The General Equation for Non-Equilibrium Reversible-Irreversible Coupling (GENERIC) formalism (Ottinger HC. 2005)

$$\dot{\boldsymbol{z}}(t) = L(\boldsymbol{z})\nabla E(\boldsymbol{z}) + M(\boldsymbol{z})\nabla S(\boldsymbol{z})$$

subject to

$$(1)\ L(\boldsymbol{z})\nabla S(\boldsymbol{z}) = 0 \qquad (2)\ M(\boldsymbol{z})\nabla E(\boldsymbol{z}) = 0$$

$$(3)\ L = -L^{\top} \qquad (4)\ M \text{ is symmetric positive semidefinite}$$

- $\boldsymbol{E}, \boldsymbol{S}$: $\mathbb{R}^d \to \mathbb{R}$: Total energy and entropy of the system
- $\boldsymbol{L}$: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$: Poisson matrix, skew-symmetric,
- $\boldsymbol{M}$: $\mathbb{R}^d \to \mathbb{R}^{d\times d}$: Friction matrix, symmetric and positive semi-definite

- Degeneracy conditions guarantee the first & second laws of thermodynamics

GENERIC Formalism Informed Neural Networks (GFINNs)

Zhang, Shin, karniadakis, Philosophical Transaction of the Royal Society A 380 (2229), 20210207

GFINNs comprise of $E_{\mathrm{NN}}(z), S_{\mathrm{NN}}(z), L_{\mathrm{NN}}(z), M_{\mathrm{NN}}(z)$

a) Satisfy degeneracy conditions (1)-(4)

b) Remain expressive enough (universal approximation)

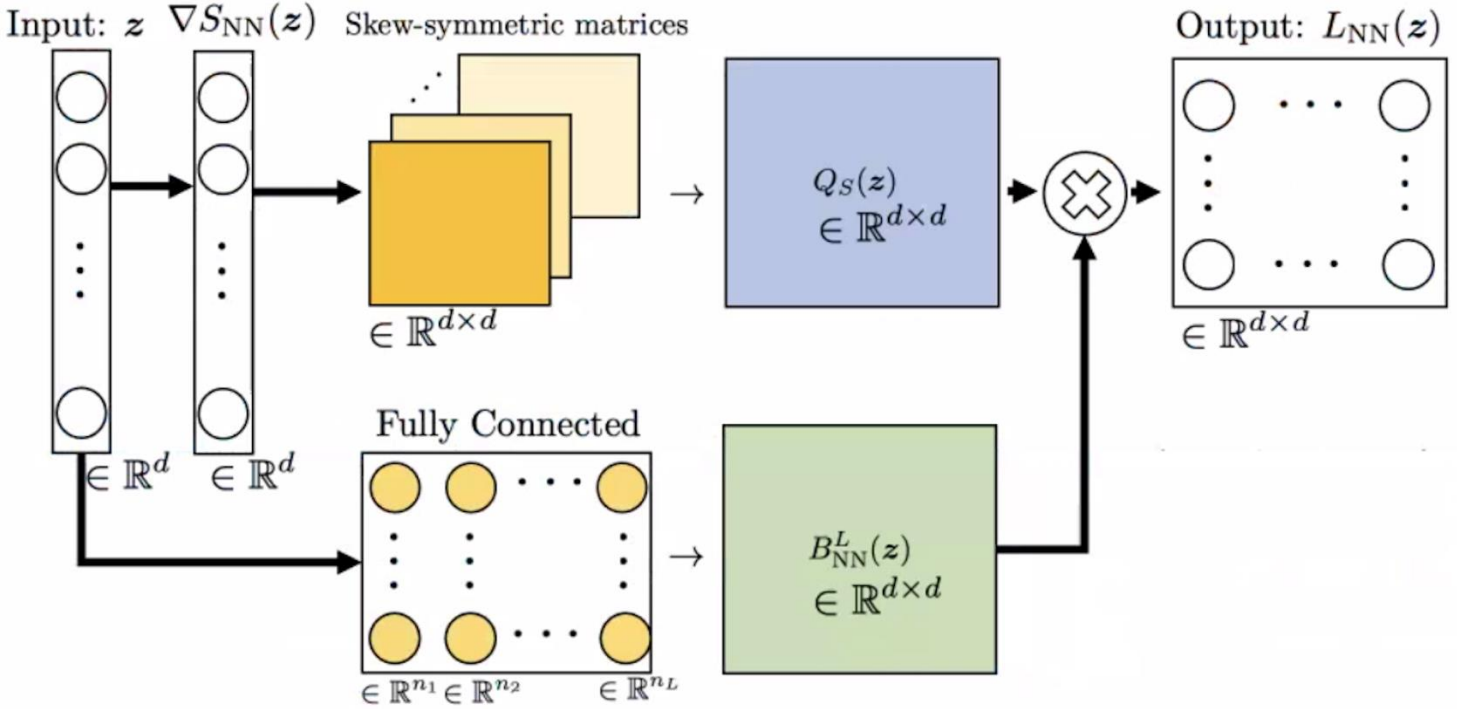
$$\mathrm{GFINN}(z;\theta) = L_{\mathrm{NN}}(z)\nabla E_{\mathrm{NN}}(z) + M_{\mathrm{NN}}(z)\nabla S_{\mathrm{NN}}(z)$$

(1) $L_{NN}\dfrac{\partial S_{NN}}{\partial z}(z) = 0$

(2) $M_{NN}\dfrac{\partial E_{NN}}{\partial z}(z) = 0$

(3) $L_{NN} = -L_{NN}^{\top}$

(4) M_{NN} is S.P.S.D



GENERIC formalism (Weak form)

- For some function ϕ smooth enough, we have

$$\int_a^b \dot{\boldsymbol{z}}(t)\phi(t)\,dt = \int_a^b L(\boldsymbol{z}(t))\nabla E(\boldsymbol{z}(t))\phi(t) + M(\boldsymbol{z}(t))\nabla S(\boldsymbol{z}(t))\phi(t)\,dt \qquad (*)$$

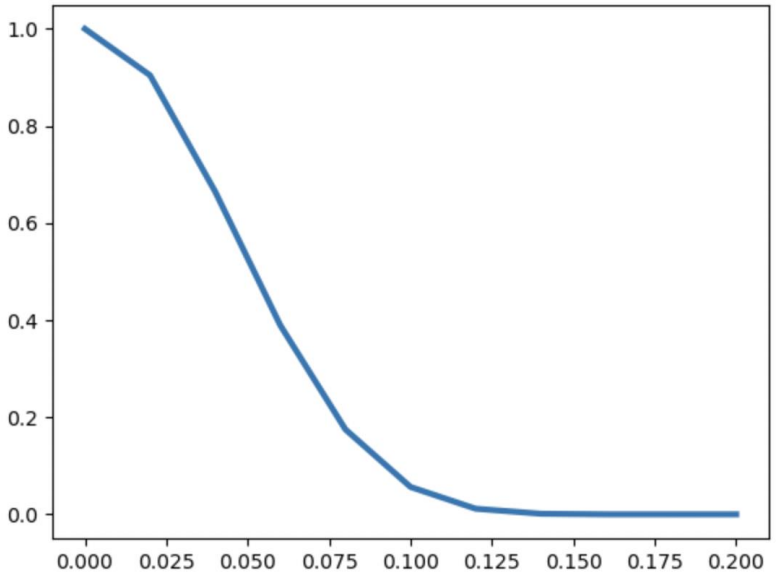
subject to $(1)\ L(\boldsymbol{z})\nabla S(\boldsymbol{z}) = 0$ $\quad (2)\ M(\boldsymbol{z})\nabla E(\boldsymbol{z}) = 0$

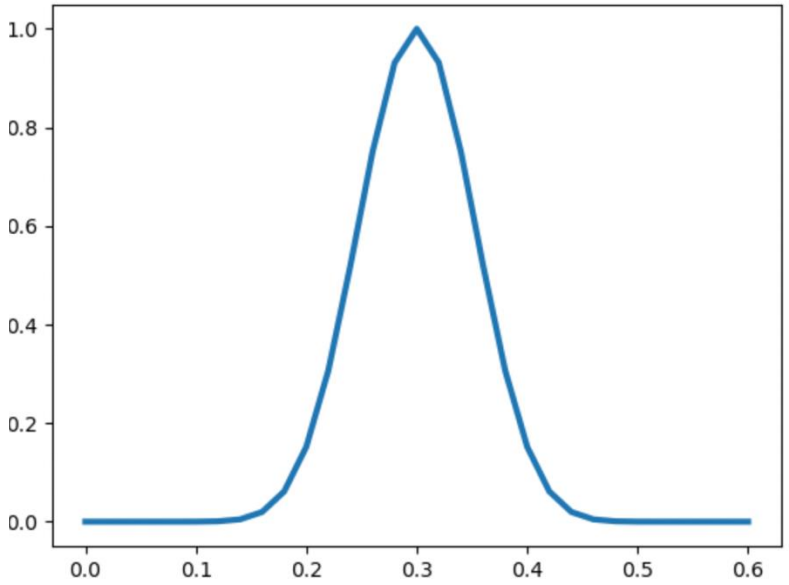
$(3)\ L = -L^\top$ $\quad (4)\ M$ is symmetric positive semidefinite

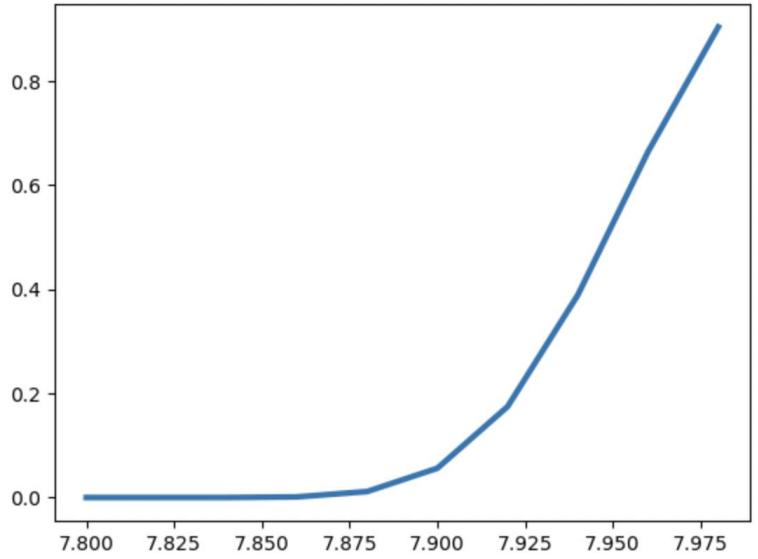
- WGFINNs (Weak GFINNs): All functions E, S, L and M are modeled by NNs like GFINNs, but in the loss, the weak form $(*)$ is used.

- Wishlist: Due to the averaging effect of integral, WGINNs handle noisy data better than GFINNs.

Test functions







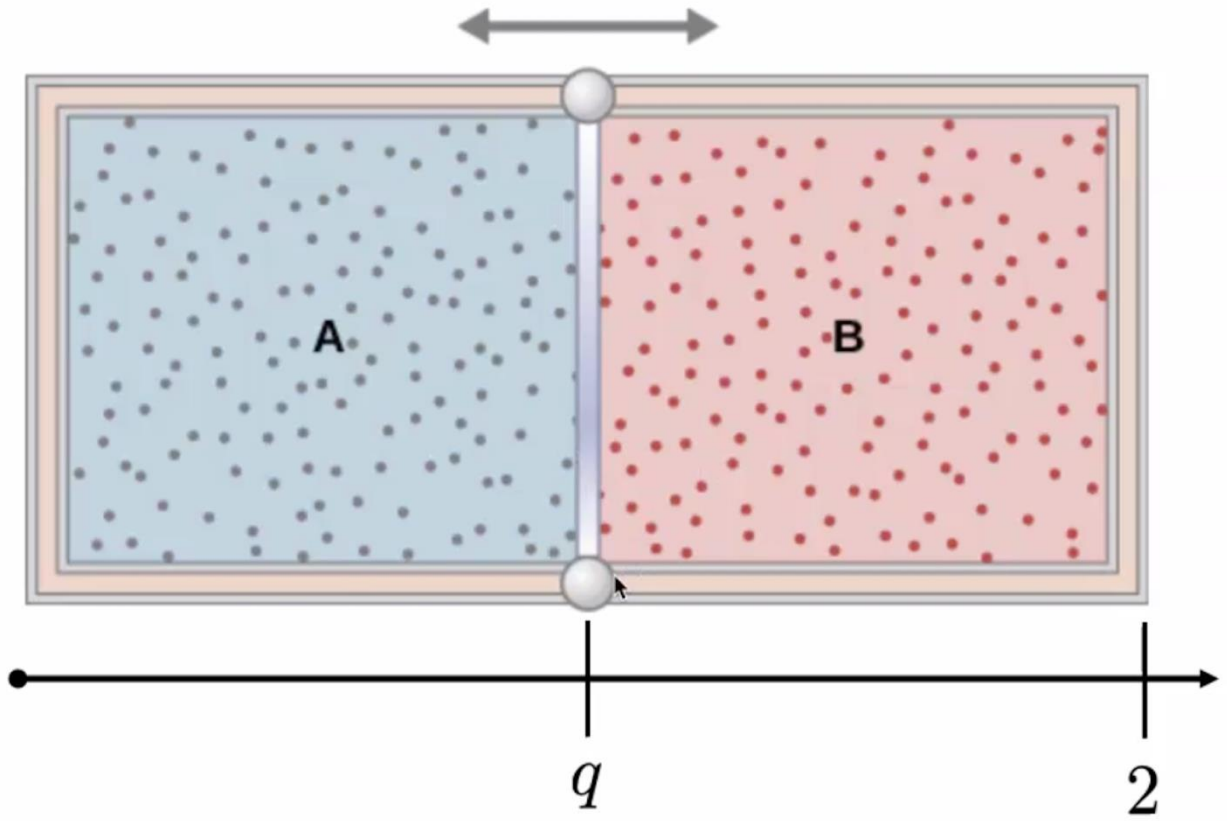
- For each subinterval [a,b], we consider the test function

$$\phi(t) = \begin{cases} C(t-a)^p(b-t)^q & t \in [a,b], \\ 0 & \text{otherwise}, \end{cases}$$

where $(a,b) \subset \mathbf{t} \times \mathbf{t}$ satisfies $a < b$ and $p, q \geq 1$.

Two gas containers exchanging heat and volume

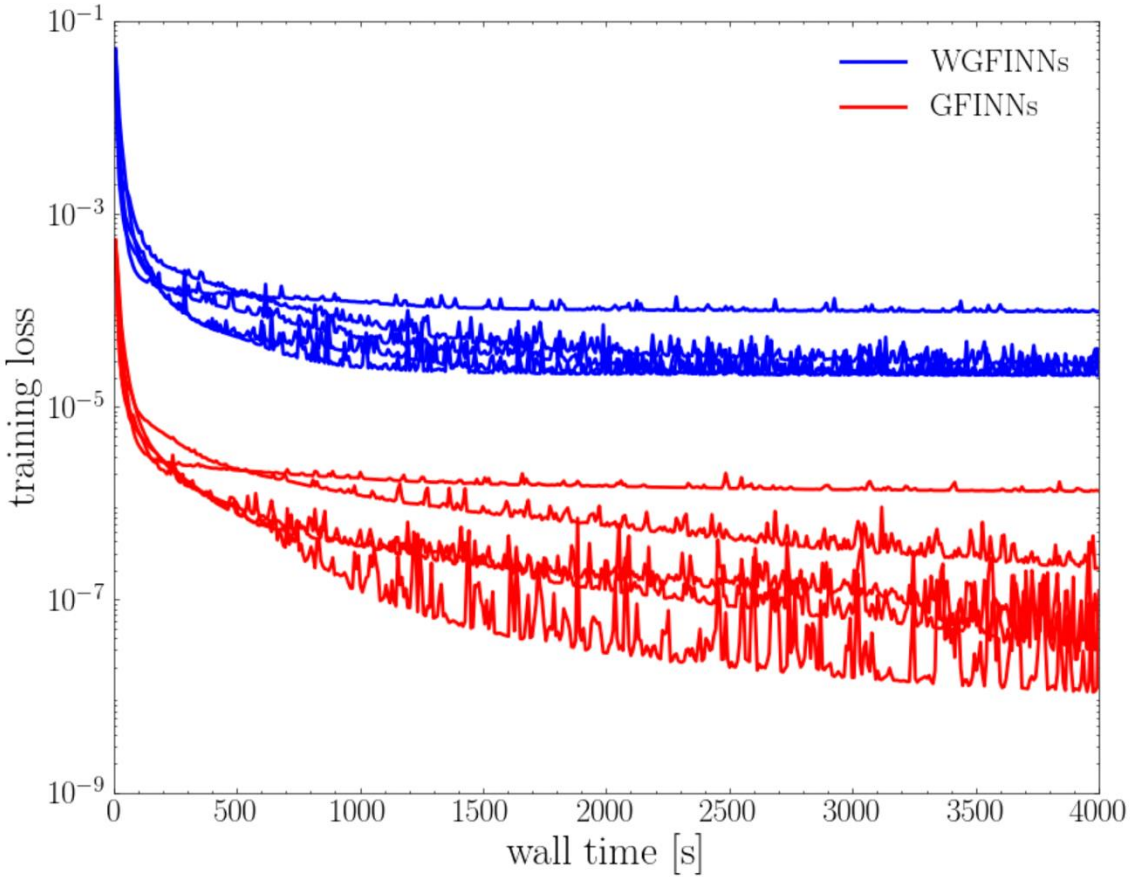
$$\begin{pmatrix} \dot{q} \\ \dot{p} \\ \dot{S}_1 \\ \dot{S}_2 \end{pmatrix} = \begin{pmatrix} \frac{p}{m} \\ \frac{2}{3}\left(\frac{E_1}{q} - \frac{E_2}{2-q}\right) \\ \frac{\alpha}{T_1}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \\ -\frac{\alpha}{T_2}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \end{pmatrix}$$



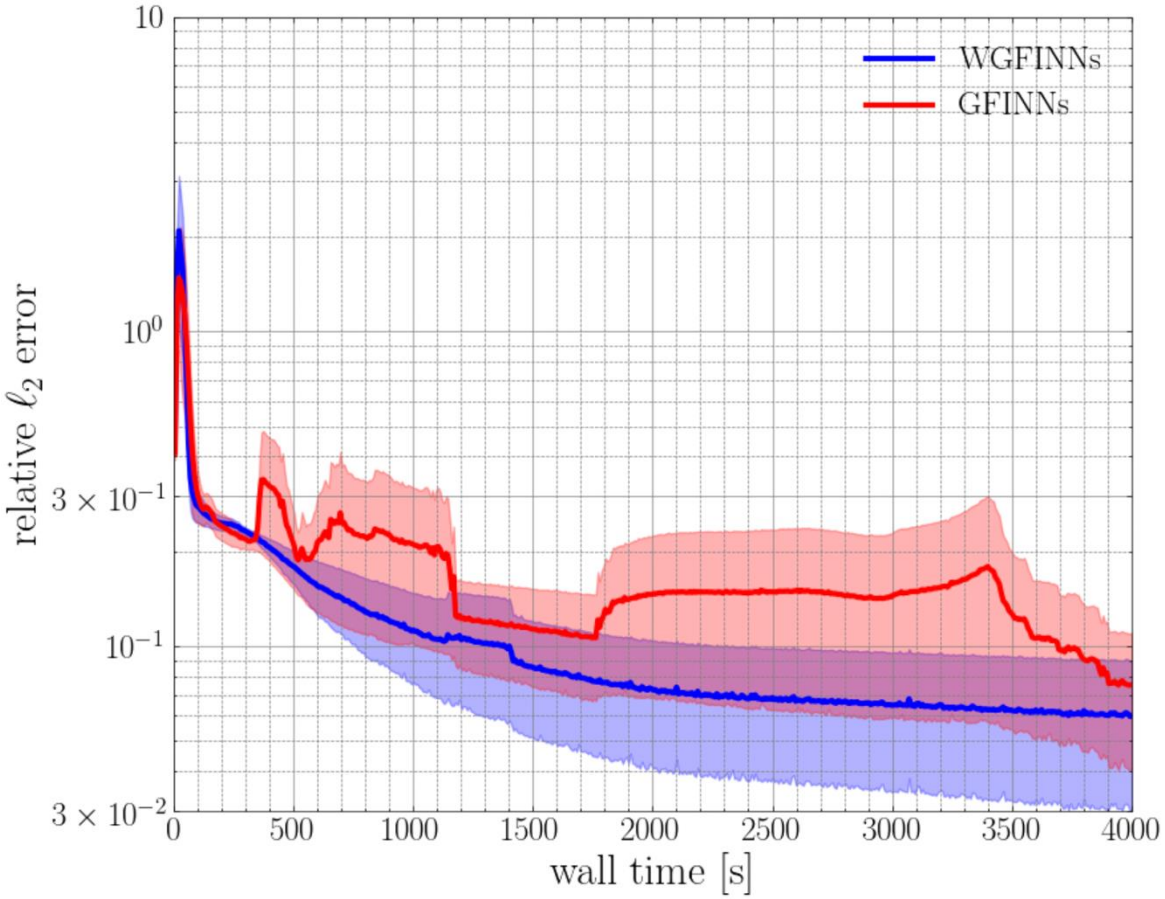
- Goal: Given 80 training trajectories, predict trajectories with new initial conditions

Noise-free case: Losses and Prediction errors from 5 simulations

Training losses for GFINNs and WGFINNs

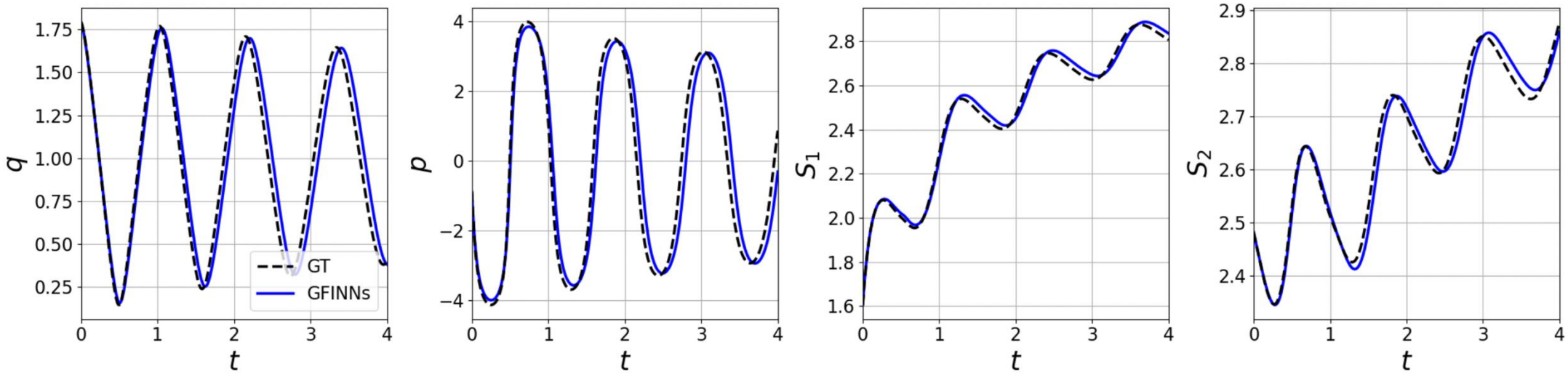


Prediction errors for GFINNs and WGFINNs

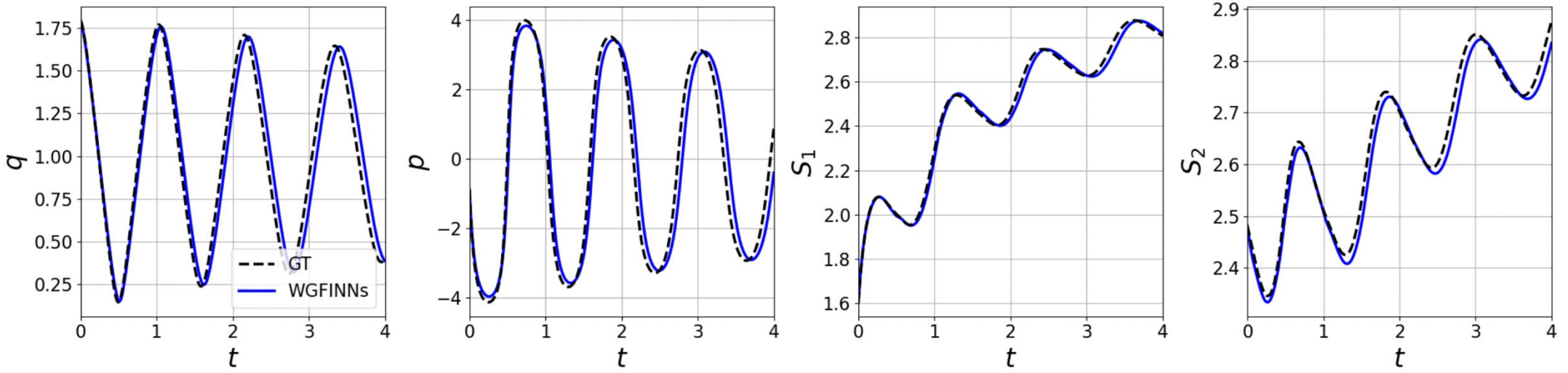


Noise-free case: Predictions with smallest losses

GFINNs

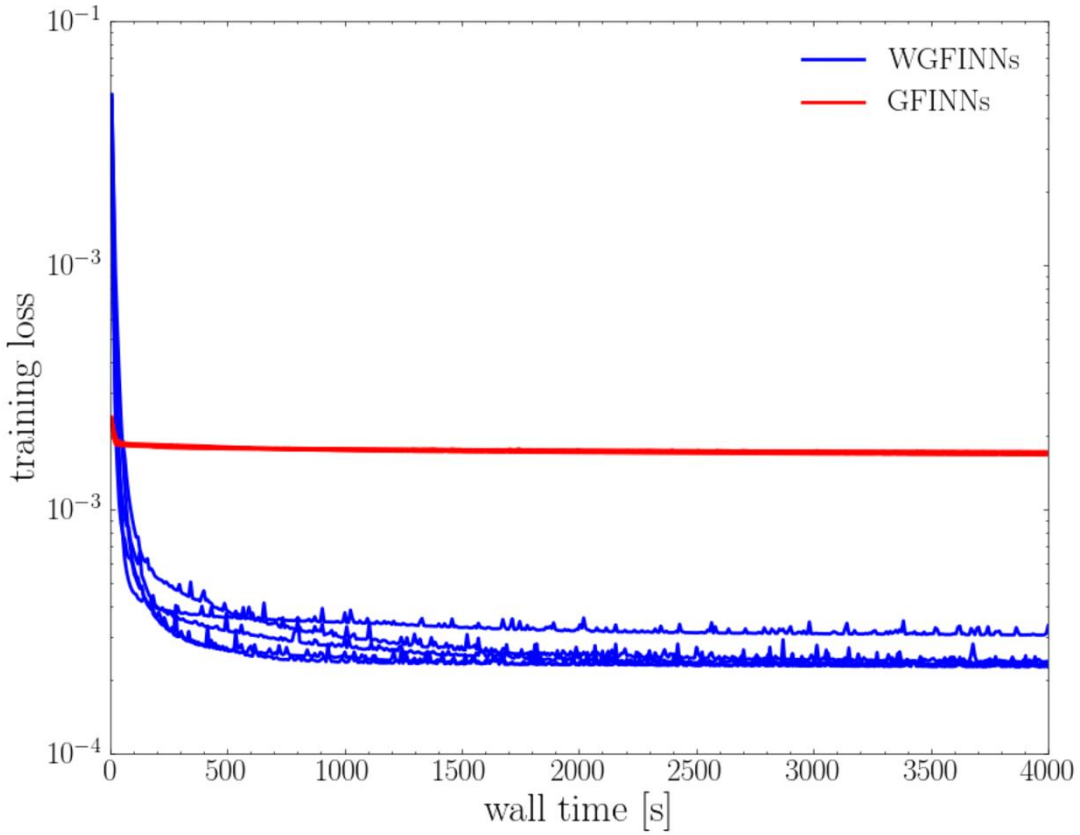


WGFINNs

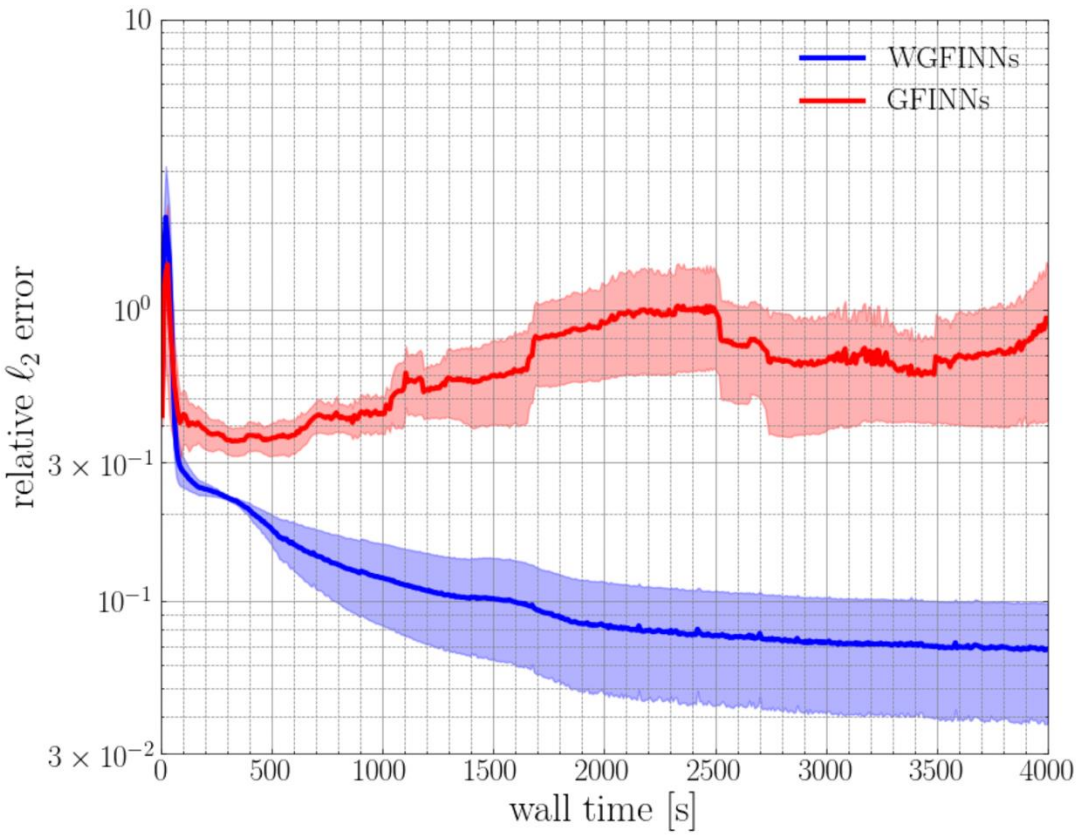


5%-noise case: Losses and Prediction errors from 5 simulations

Training losses for GFINNs and WGFINNs

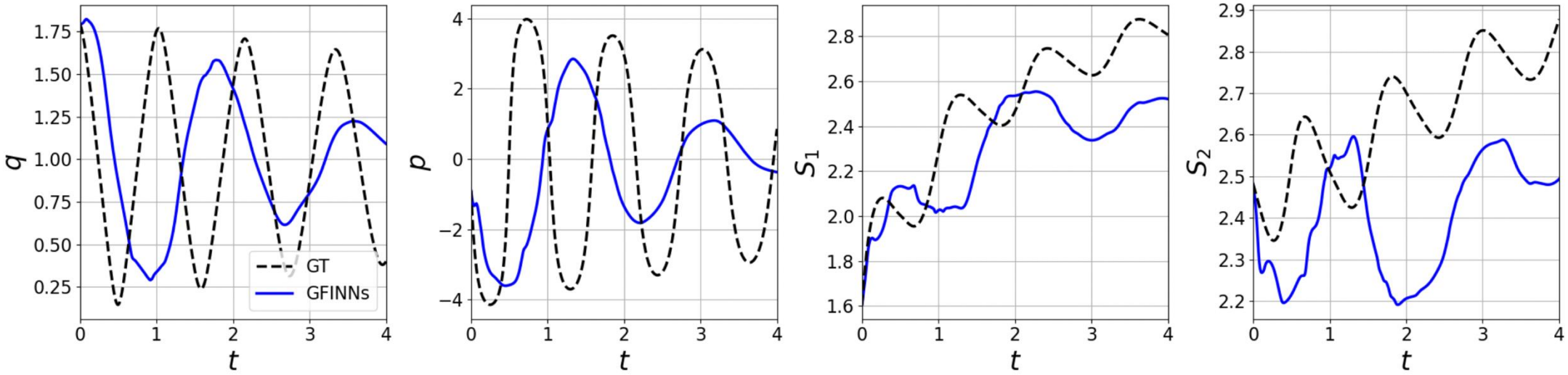


Prediction errors for GFINNs and WGFINNs

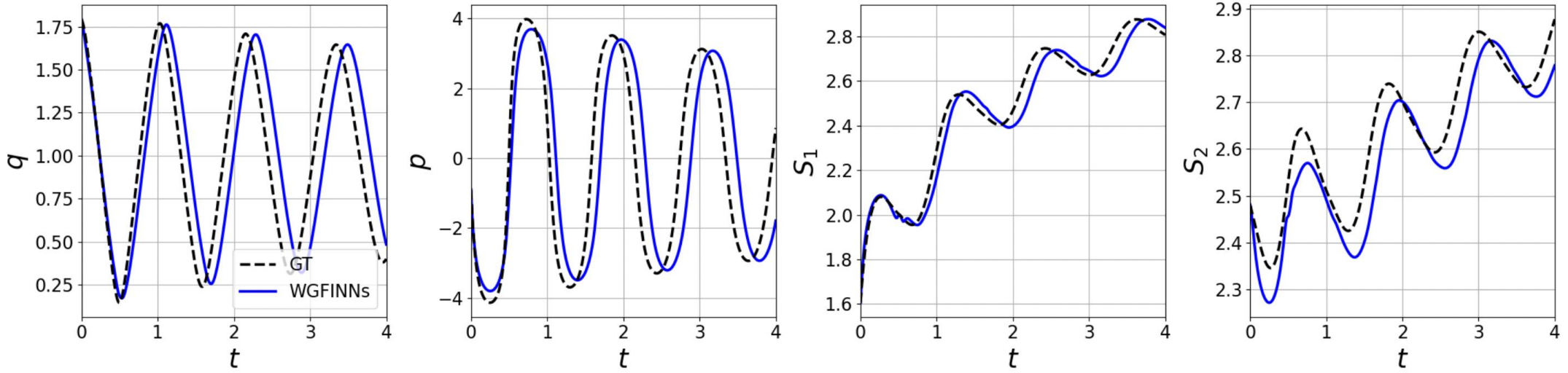


5%-noise case: Predictions with smallest losses

GFINNs

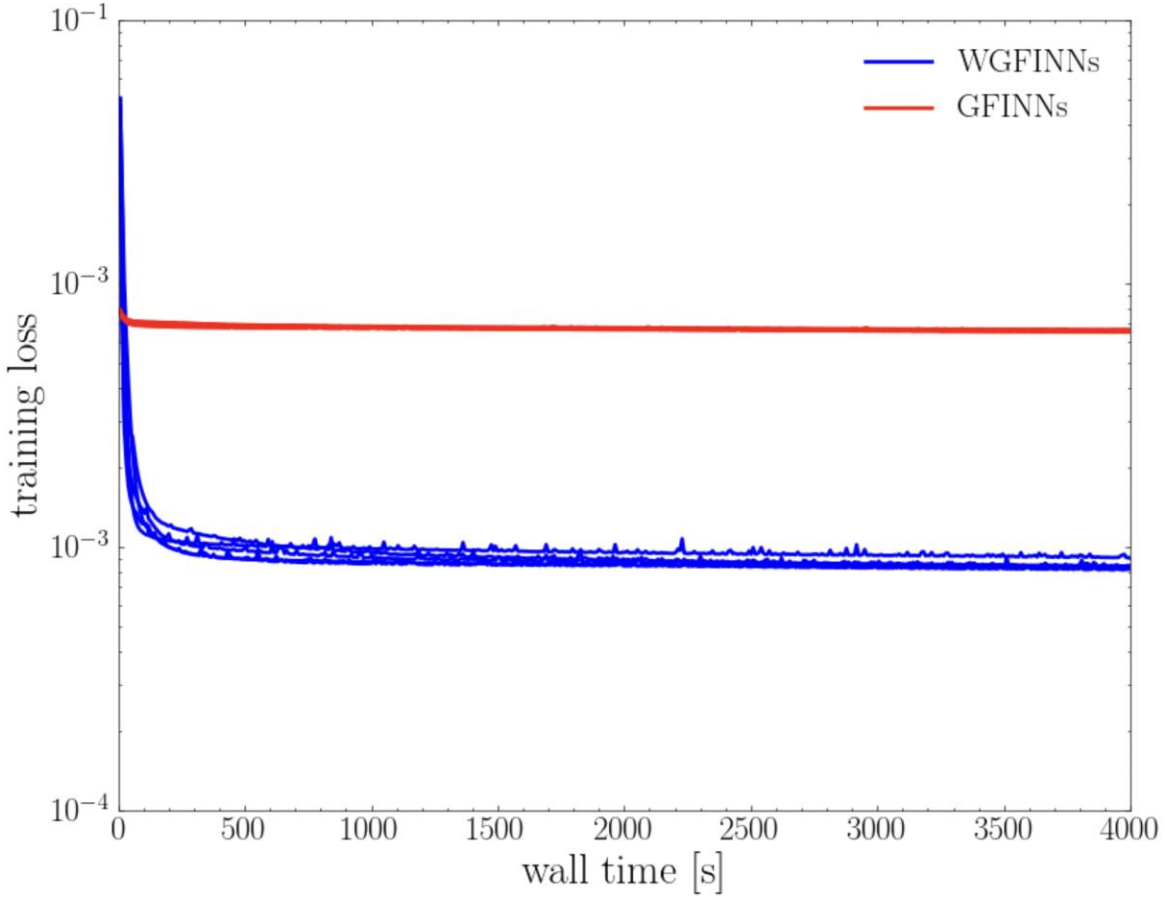


WGFINNs

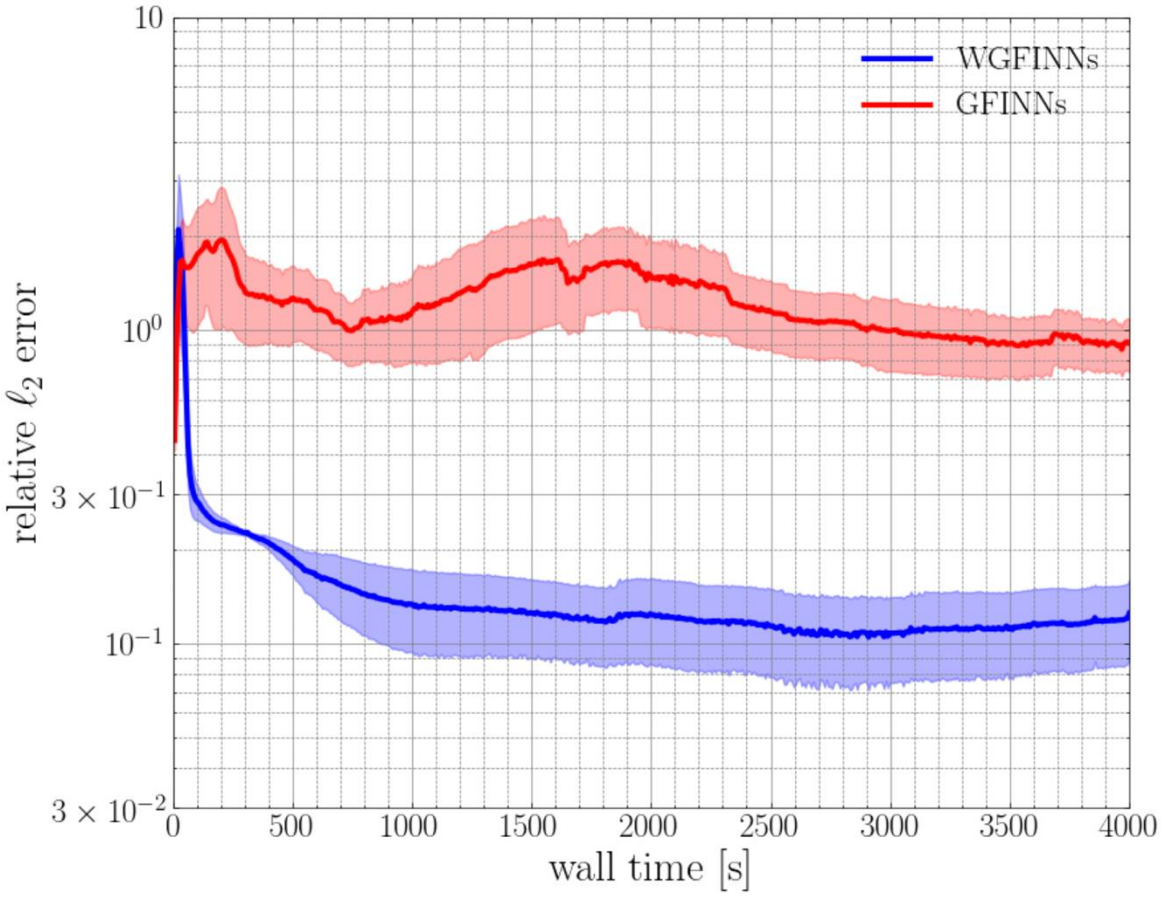


10%-noise case: Losses and Prediction errors from 5 simulations

Training losses for GFINNs and WGFINNs

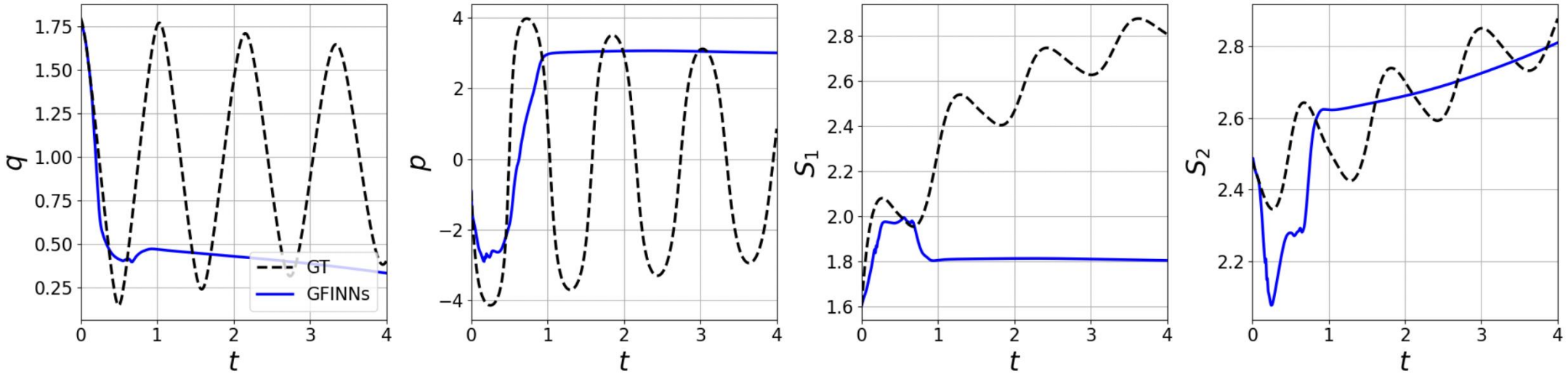


Prediction errors for GFINNs and WGFINNs

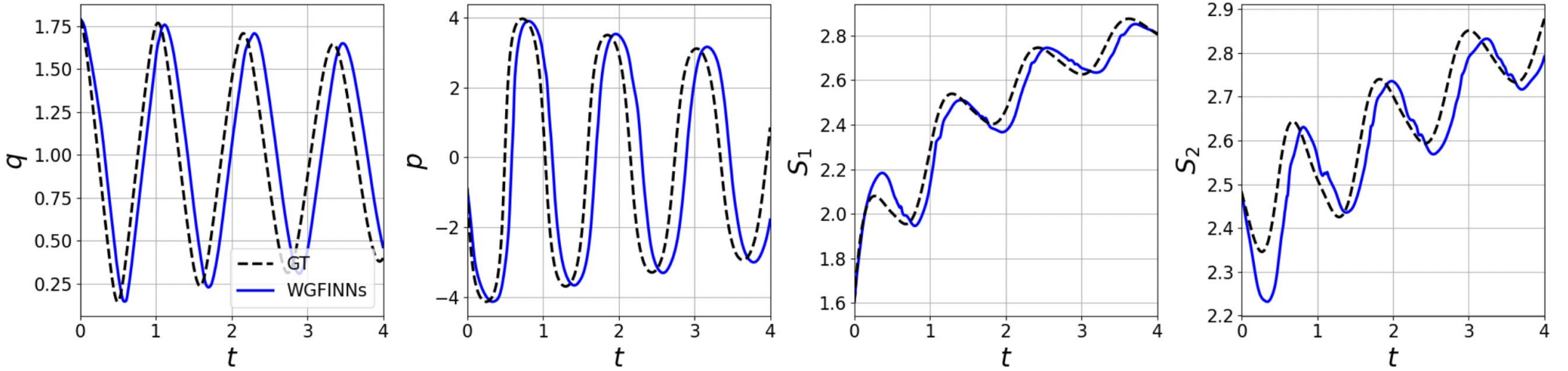


10%-noise case: Predictions with smallest losses

GFINNs



WGFINNs



Thermoelastic double pendulum

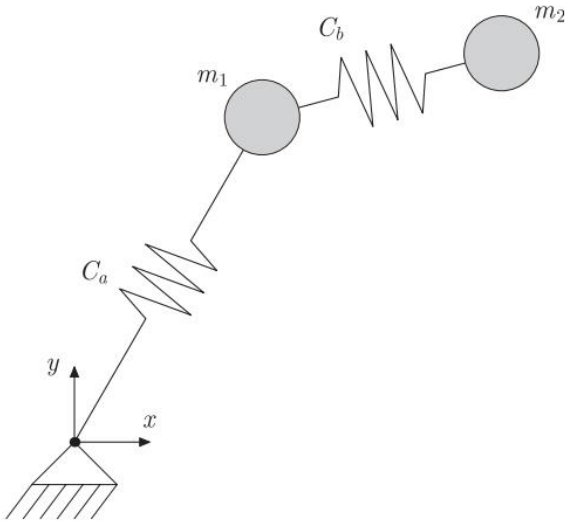


Figure 1. Model problem: a double thermoelastic pendulum.

$$\begin{pmatrix} \dot{\boldsymbol{q}}_1 \\ \dot{\boldsymbol{q}}_2 \\ \dot{\boldsymbol{p}}_1 \\ \dot{\boldsymbol{p}}_2 \\ \dot{S}_1 \\ \dot{S}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ -\frac{\partial}{\partial \boldsymbol{q}_1}(E_1 + E_2) \\ -\frac{\partial}{\partial \boldsymbol{q}_2}(E_1 + E_2) \\ T_1^{-1} T_2 - 1 \\ T_1 T_2^{-1} - 1 \end{pmatrix} = L(\boldsymbol{x}) \frac{\partial E(\boldsymbol{x})}{\partial \boldsymbol{x}} + M(\boldsymbol{x}) \frac{\partial S(\boldsymbol{x})}{\partial \boldsymbol{x}},$$

$$\lambda_1 = \|\boldsymbol{q}_1\|, \; \lambda_1 = \|\boldsymbol{q}_2 - \boldsymbol{q}_1\|$$

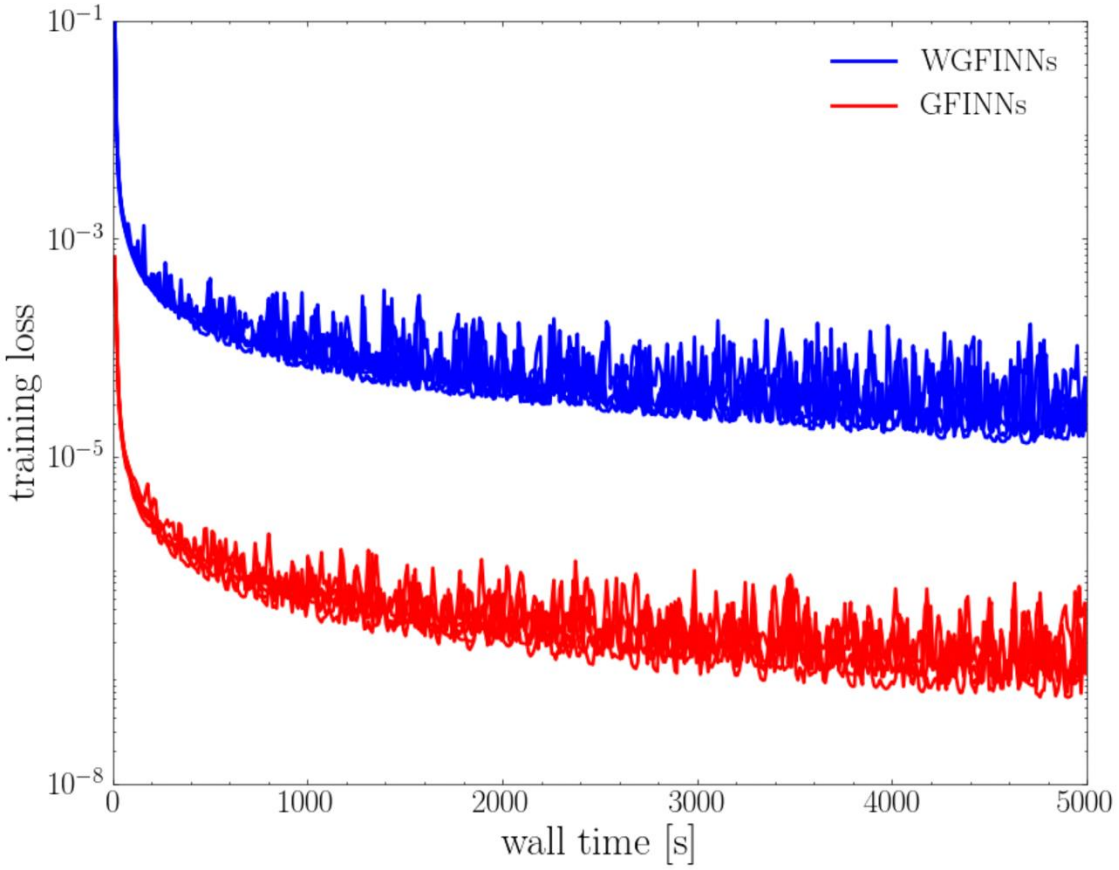
$$\boldsymbol{x} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{p}_1, \boldsymbol{p}_2, S_1, S_2) \in \mathbb{R}^{10}$$

Position, momentum of ith mass, entropy of ith spring

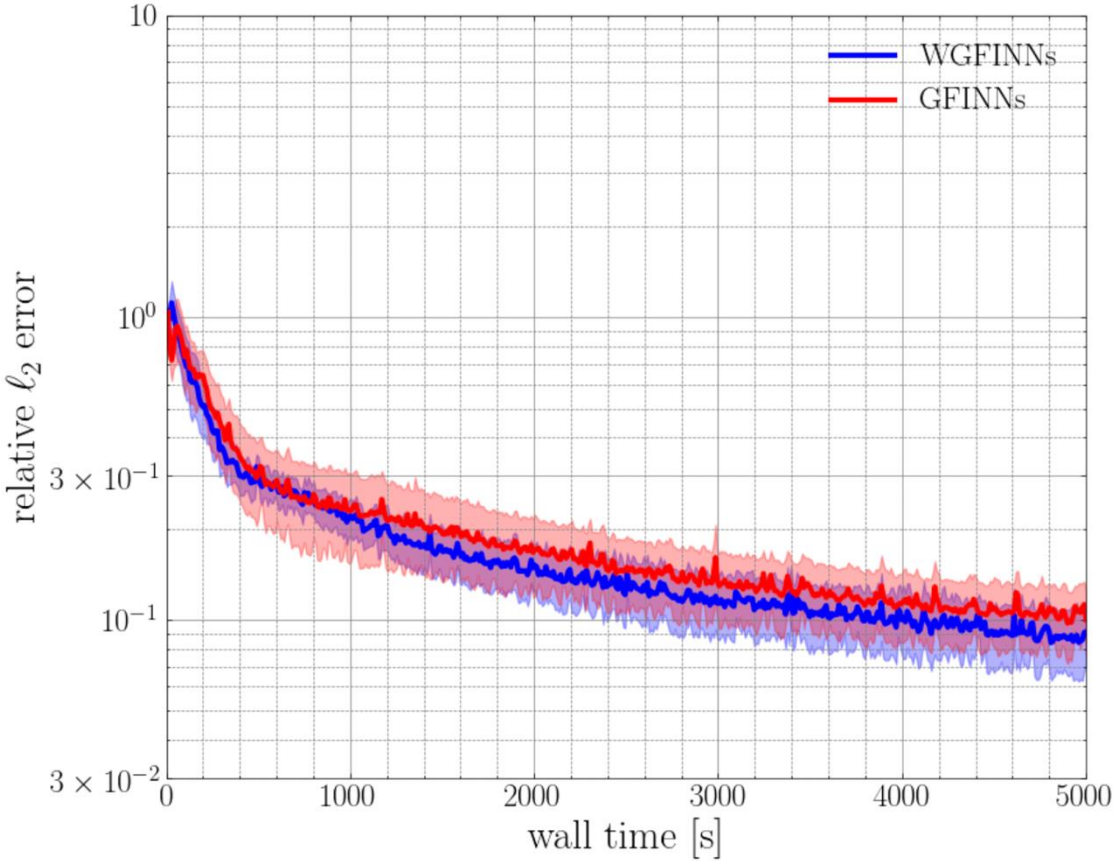
- Goal: Given 80 training trajectories, predict trajectories with new initial conditions

Noise-free case: Losses and Prediction errors from 5 simulations

Training losses for GFINNs and WGFINNs

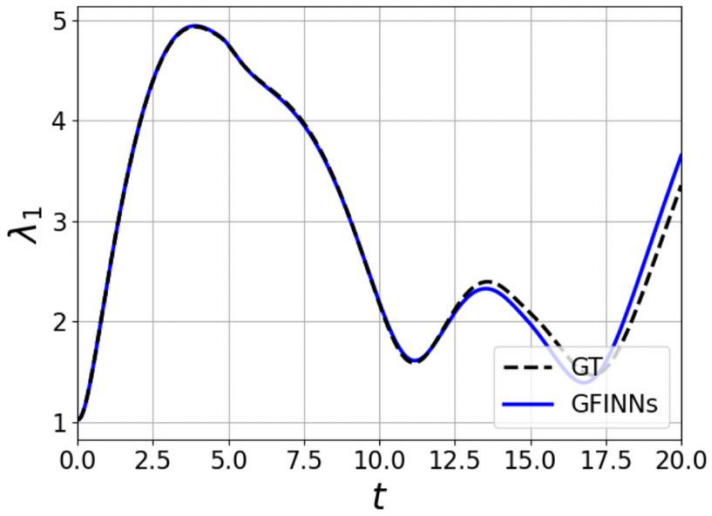


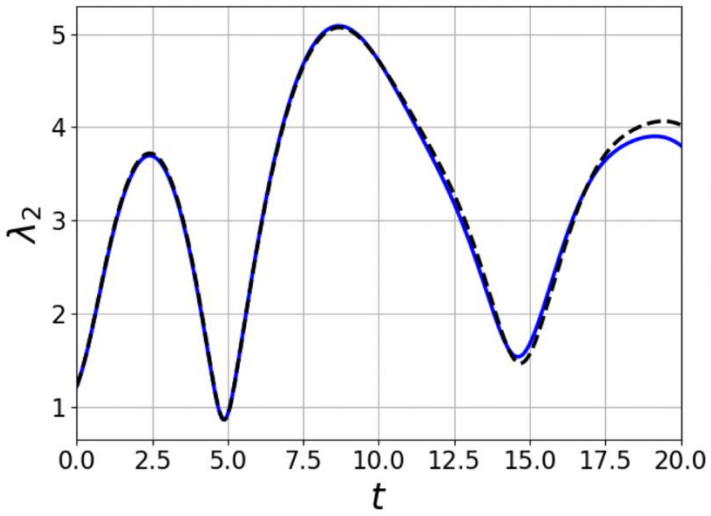
Prediction errors for GFINNs and WGFINNs

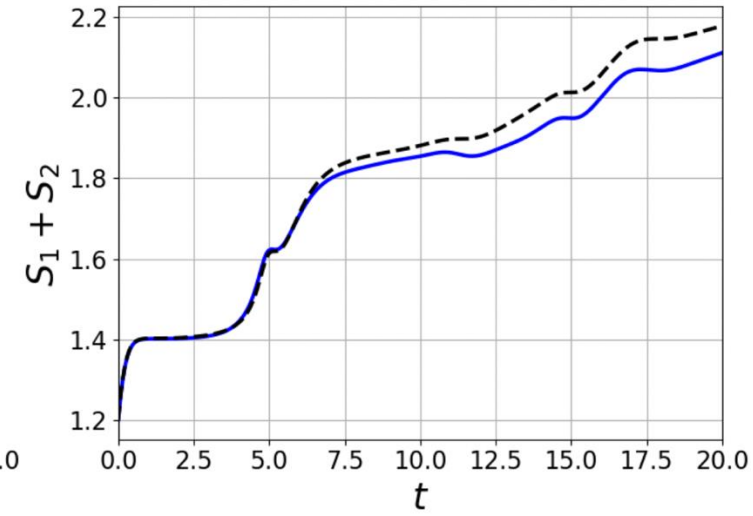


Noise-free case: Predictions with smallest losses

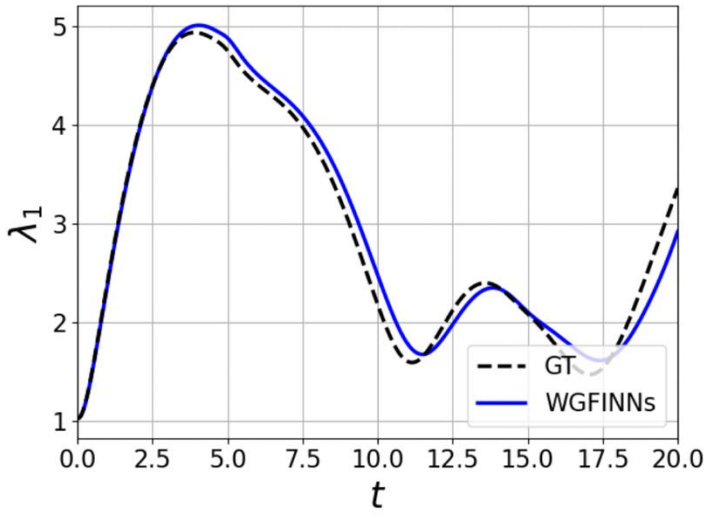
GFINNs

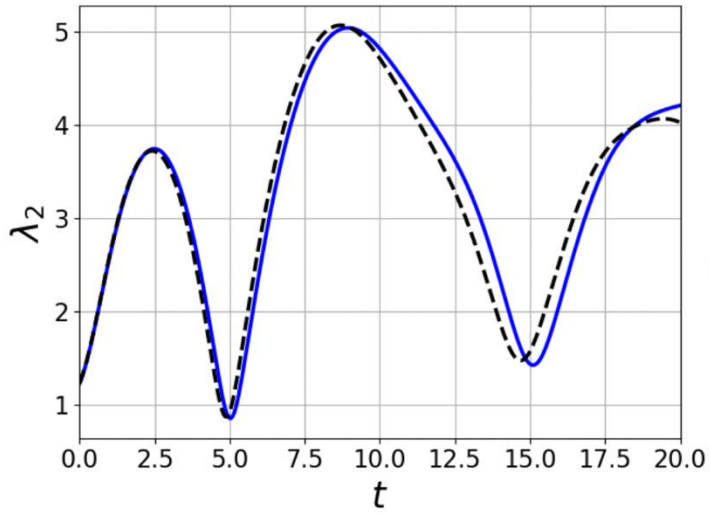


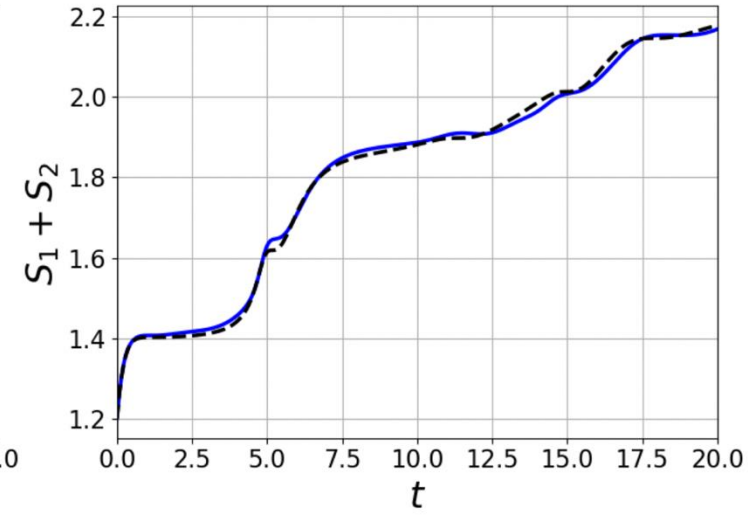




WGFINNs

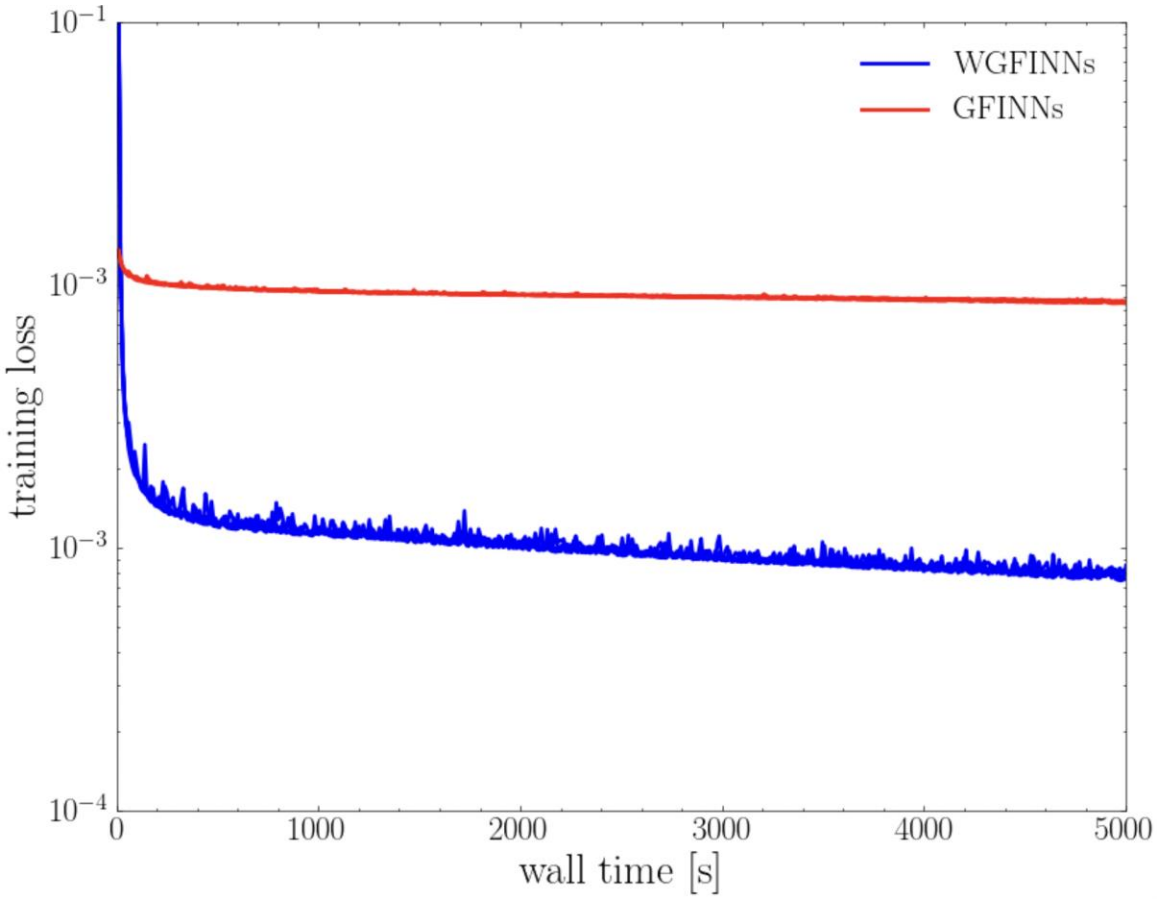




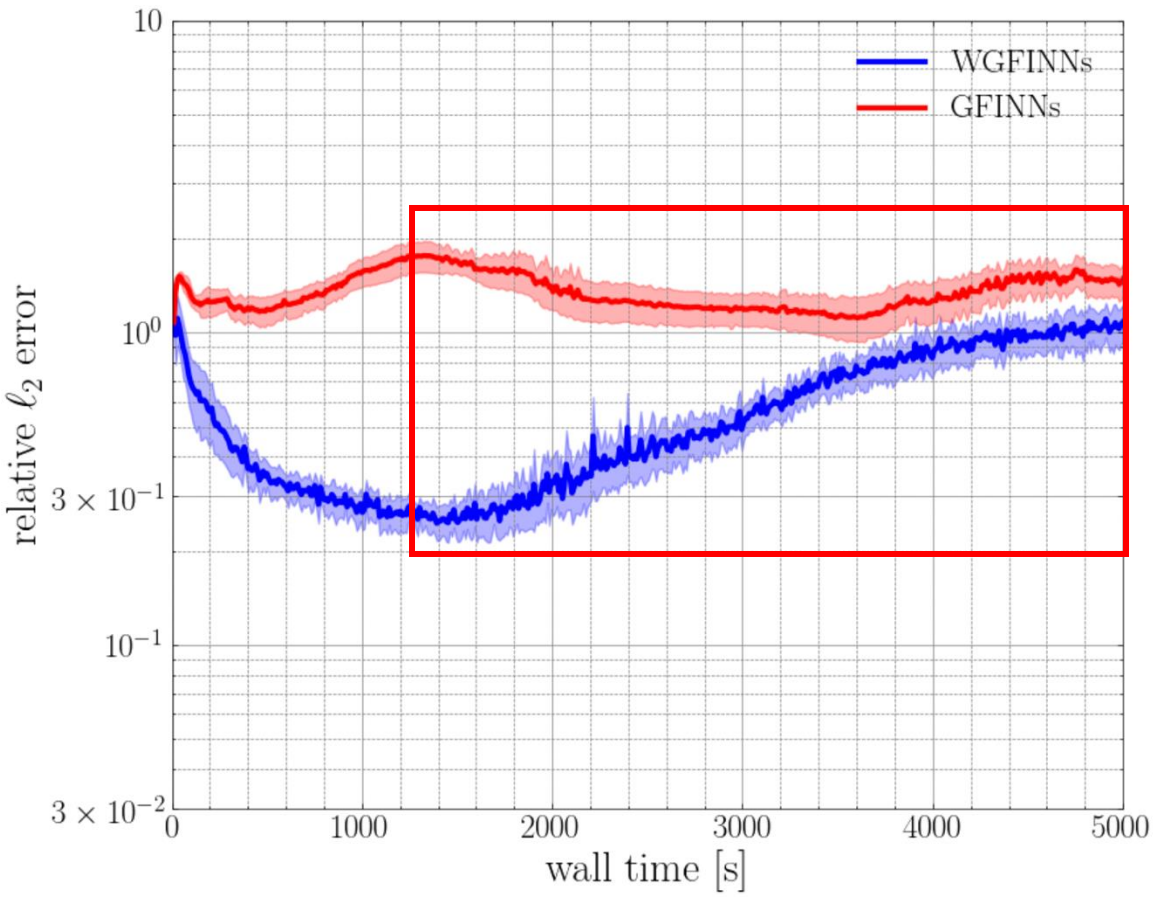


5%-noise case: Losses and Prediction errors from 5 simulations

Training losses for GFINNs and WGFINNs



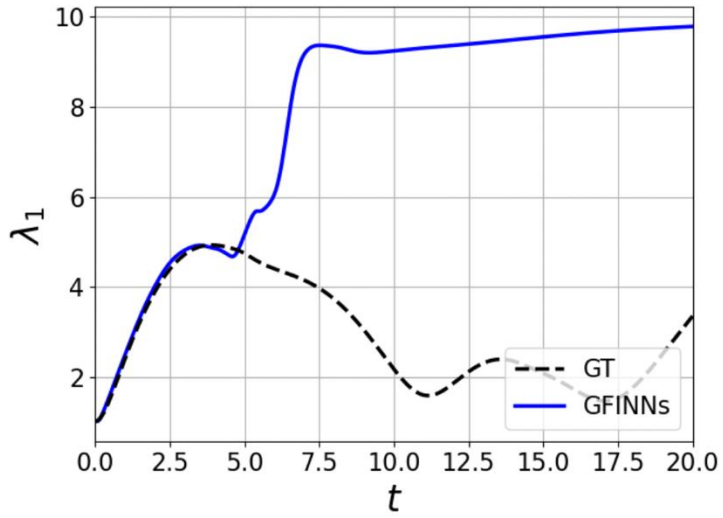
Prediction errors for GFINNs and WGFINNs

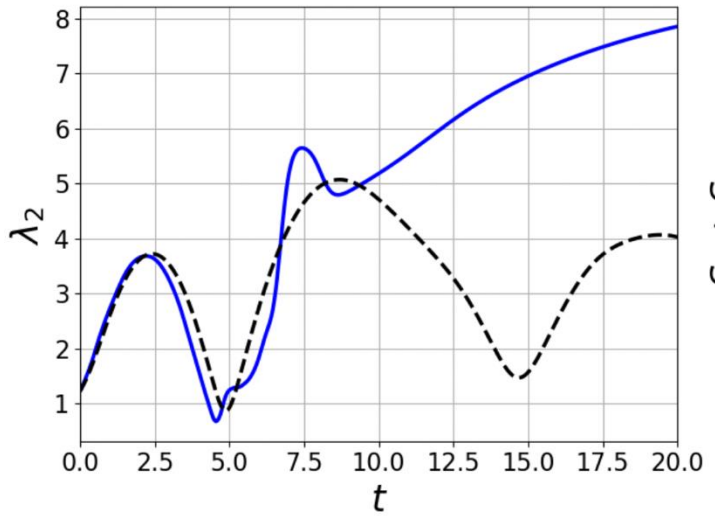


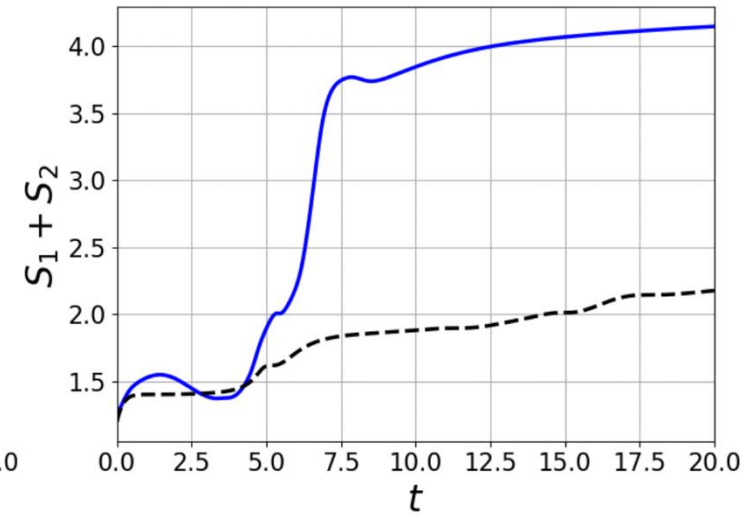
- Overfitting issue should be resolved

5%-noise case: Predictions with smallest losses

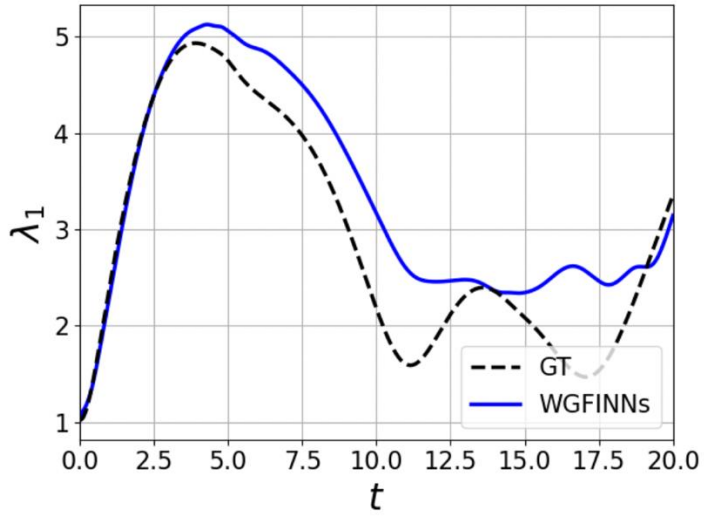
GFINNs

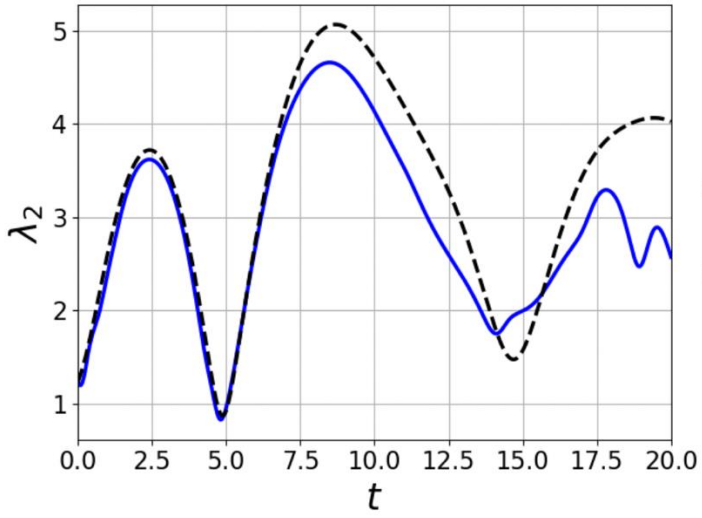


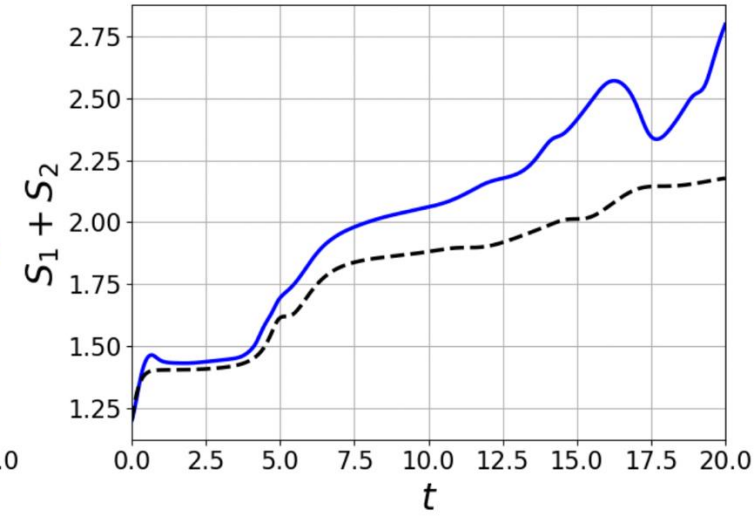




WGFINNs







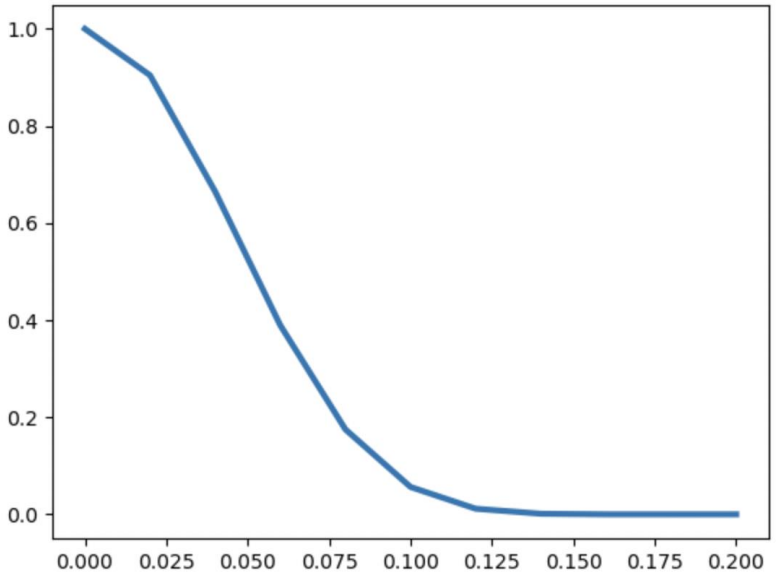
Conclusions

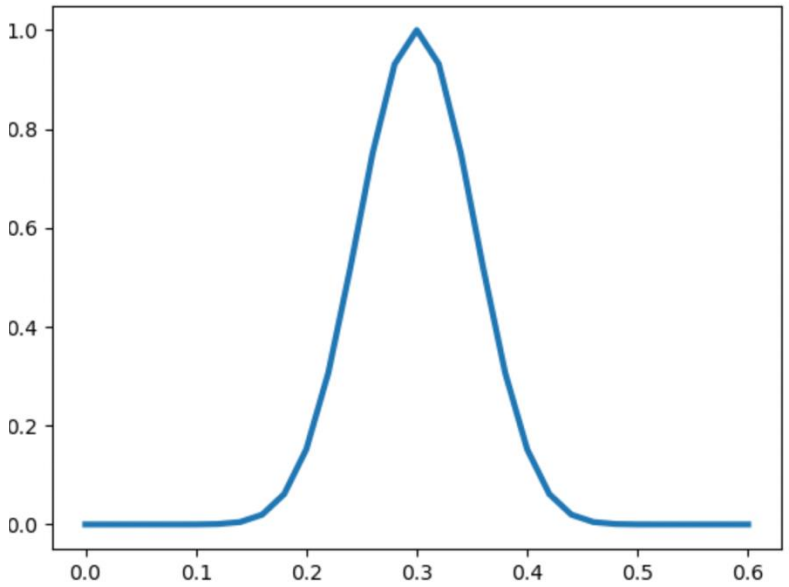
- We considered discovering underlying thermodynamics via GENERIC formalism using trajectory data and NN model.
- Weak-from based loss functional is utilized to handle noisy data (WGFINNs)
- Weak-from GFINNs outperforms GFINNs due to the effect of averaging out the noise
- Future works: Further enhancing the accuracy of the approximations

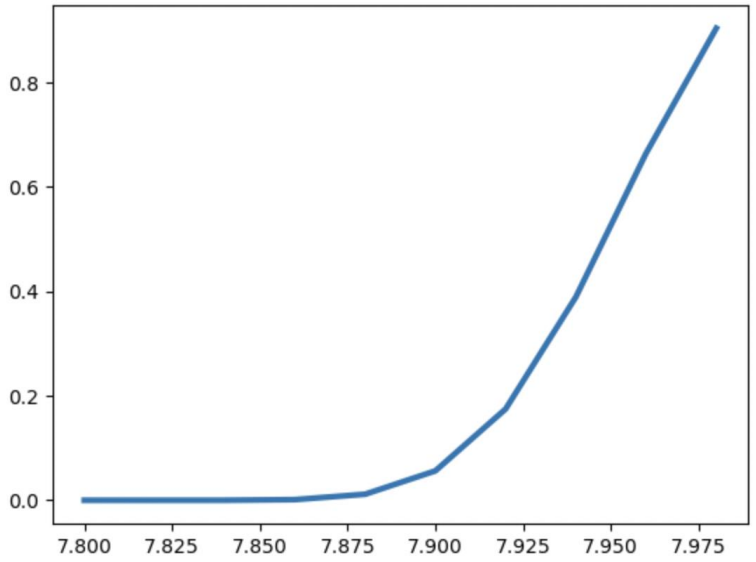
Thank you

Jun Sur Richard Park: junsurpark@kias.re.kr

Test functions







- For each subinterval [a,b], we consider the test function

$$\phi(t) = \begin{cases} C(t-a)^p(b-t)^q & t \in [a,b], \\ 0 & \text{otherwise}, \end{cases}$$

where $(a,b) \subset \mathbf{t} \times \mathbf{t}$ satisfies $a < b$ and $p, q \geq 1$.