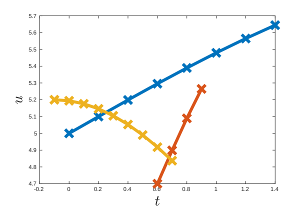
Learning dynamical systems with geometry + UQ

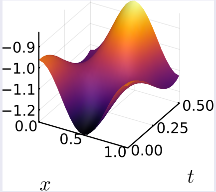
Christian Offen

Department of Mathematics, Paderborn University, Germany

July 2025

Task: Learn dynamical systems from data





time series / (discrete) space-time data

Prior knowledge

Learn with geometric prior knowledge. [incredibly active research community][1]

Motivation: extrapolation, stability, geometric properties, data-requirements

Question

How to include UQ?

[1] Allen-Blanchette, Bertalan, Brunton, Buchfink, Carlberg, Celledoni, Chen, Cranmer, Dehmamy, Ehrhard, Geelen, Glas, Greydanus, Jin, Kramer, Lee, Leok, Mason, Mezić, Mu, Murari, Ortega, Owren, Proctor, Qin, Rath, Schönlieb, Sharma, Tao, Yin, Zhang, Zhu, . . .

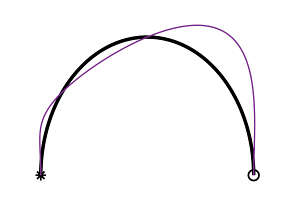
Variational structure

Variational principle

Motions $q\colon [t_0, t_n] \to Q$ are stationary points of

$$S(q) = \int_{t_0}^{t_n} L(t, q(t), \dot{q}(t))\,\mathrm{d}t.$$

subject to $q(t_0) = q_0, \; q(t_n) = q_n$.



Euler–Lagrange equations

Motions solve

$$0 = EL(L) = \frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}} \ddot{q} + \frac{\partial^2 L}{\partial q \partial \dot{q}} \dot{q} - \frac{\partial L}{\partial q}$$

Idea

Learn Lagrangian L. [Cranmer et al 2020], [Qin 2020]

Examples of geometric models

Lagrangian-based

$$(\mathrm{ODE})\ 0 = EL(L) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot q} - \frac{\partial L}{\partial q}$$

$$(\mathrm{PDE})\ 0 = EL(L) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial u_t} + \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial u_x} - \frac{\partial L}{\partial u}$$

Discrete Lagrangian-based

$$(ODE)\ 0 = \nabla_{q^i}\left(L_d(q^{i-1},q^i) + L_d(q^i,q^{i+1})\right)$$

$$(PDE)\ 0 = \nabla_{u^i_j}\Big(L_d(u^i_j, u^{i+1}_j, u^i_{j+1}) + L_d(u^{i-1}_j, u^i_j, u^{i-1}_{j+1}) + L_d(u^i_{j-1}, u^{i+1}_{j-1}, u^i_j)\Big)$$

...

Hamiltonian-based

$$(\mathrm{sympl})\ \dot z = J^{-1}(z)\nabla H(z),\ z=(q,p)^\top$$

$$(\mathrm{multi-sympl})\ Ku_t + Lu_x = \nabla_u H(u)$$

Port-Hamiltonian

$$\dot z = (J(z) - R(z))\nabla H(z) + g(z)f, \quad f \text{ input}$$

$$e = g(z)^\top \nabla H(z) \quad e \text{ output}$$

Discrete Hamiltonian-based

$$(\mathrm{sympl})\ \frac{z^{i+1}-z^i}{h} = J^{-1}\nabla H_d\left(\frac{z^{i+1}+z^i}{2}\right)$$

...

Typical approach

Parametrize L, L_d, H, H_d, ... and fit to data.

Blessings of learning Lagrangians vs. equations of motions

Hope: Less data-requirements, less model uncertainty

Blessings of learning Lagrangians vs. equations of motions

Hope: Less data-requirements, less model uncertainty

Running example: Euler–Lagrange equations

Motions solve

$$0 = EL(L) = \frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}} \ddot{q} + \frac{\partial^2 L}{\partial q \partial \dot{q}} \dot{q} - \frac{\partial L}{\partial q}$$

Problem with learning Lagrangians

Lagrangians are not uniquely determined by motions.

- Ill-defined inverse problem
- degenerate solutions

Additional complication

- Ill-posed inverse problem $\Longrightarrow$ huge uncertainty from model ambiguity

Way out

Regularisation that

- does not restrict generality of the ansatz
- remove (some) model ambiguity

Additional complication

- Ill-posed inverse problem $\Longrightarrow$ huge uncertainty from model ambiguity

Way out

Regularisation that

- does not restrict generality of the ansatz
- remove (some) model ambiguity

Lagrangian ode model

$$0 = EL(L) = \frac{\partial^2 L}{\partial \dot q \partial \dot q} \ddot q + \frac{\partial^2 L}{\partial q \partial \dot q} \dot q - \frac{\partial L}{\partial q}$$

Example: Linear conditions for L

$(q^{(j)}, \dot q^{(j)}, \ddot q^{(j)})$ data. For $\tau_c, p_b, c_b \neq 0$, $(q^*, \dot q^*, \ddot q^*)$ not a motion

$$\mathrm{EL}(L)(q^{(j)}, \dot q^{(j)}, \ddot q^{(j)}) = 0 \quad \forall j$$
$$\mathrm{EL}(L)(q^*, \dot q^*, \ddot q^*)_k = \tau_c \neq 0$$
$$\frac{\partial L}{\partial \dot q}(q_b, \dot q_b) = p_b \neq 0, \; L(q_b, \dot q_b) = c_b \neq 0.$$

Theorem

Non-deg. assumptions. The conditions

- do not restrict the generality
- do **not** determine L uniquely
- (infinite data limit) uniquely determine $\mathrm{EL}(L)$, sympl. struct. + Hamiltonian, acceleration ...
- safeguard from degenerate solutions

Gaussian Process - Learning of a Lagrangian

New perspective: pde solving

$$0 = EL(L) = \frac{\partial^2 L}{\partial \dot{q} \partial \dot{q}} \ddot{q} + \frac{\partial^2 L}{\partial q \partial \dot{q}} \dot{q} - \frac{\partial L}{\partial q}$$

$+$ regularisation $\leftrightsquigarrow$ boundary conditions.

Fits to some theory, data-driven methods for pdes [Owhadi,Scovel,Henning,Pförtner,...].

Idea

Solve this ill-posed, linear pde for L with a data-driven method.

1. Model Lagrangian L as GP (prior)
2. Condition on
   - differential equation $\mathrm{EL}_{\hat{q}_j}(L) = 0$ at data points
   - (linear) regularisation
3. The a posteriori process is again a GP.

Covariance matrix

$$\Phi = \big(\mathrm{EL}, \quad P, \quad \mathrm{ev}\big)$$

$$\theta = \kappa(\Phi,\Phi) = \begin{pmatrix} (\mathrm{EL}^x_{\hat q_i}\mathrm{EL}^y_{\hat q_j}(k))_{ij} & (\mathrm{EL}^x_{\hat q_i} P^y_{\overline{q}_b}(k))_i & (\mathrm{EL}^x_{\hat q_i}\mathrm{ev}^y_{\overline{q}_b}(k))_i \\ (P^x_{\overline{q}_b}\mathrm{EL}^y_{\hat q_j}(k))_j & P^x_{\overline{q}_b} P^y_{\overline{q}_b}(k) & P^x_{\overline{q}_b}\mathrm{ev}^y_{\overline{q}_b}(k) \\ (\mathrm{ev}^x_{\hat q_b}\mathrm{EL}^y_{\hat q_j}(k))_j & \mathrm{ev}^x_{\overline{q}_b} P^y_{\overline{q}_b}(k) & \mathrm{ev}^x_{\overline{q}_b}\mathrm{ev}^y_{\overline{q}_b}(k) \end{pmatrix}, \quad k(x,y) \text{ kernel}$$

Conditional mean L_Φ

$$L_\Phi(\overline{q}) = y^\top \theta^{-1}\kappa_\Phi = \big(0, \quad \ldots, \quad 0, \quad p_b, \quad c_b\big)\,\theta^{-1} \begin{pmatrix} \mathrm{EL}_{\hat q_1}(k(x,\overline{q})) \\ \vdots \\ \mathrm{EL}_{\hat q_N}(k(x,\overline{q})) \\ P_{\bar q_b}(k(x,\overline{q})) \\ k(\overline{q}_b,\overline{q}) \end{pmatrix}, \quad \overline{q} \in \Omega = TQ$$

Conditional covariance κ_Φ

$$\kappa_\Phi(\varphi) = \varphi(\kappa\varphi) - \kappa(\varphi,\Phi)\theta^{-1}\kappa(\Phi,\varphi), \qquad \varphi \in \mathcal{U}^*$$

Example $\varphi = H_{\overline{q}}$, $\overline{q} \in \Omega$, $\quad H_{\bar q}(L) = P_{\bar q}(L) - L(\bar q)$

$$\varphi(\kappa\varphi) = H^x_{\overline{q}} H^y_{\overline{q}}(k), \quad \kappa(\varphi,\Phi) = \kappa(\Phi,\varphi)^\top = \Big((H^x_{\overline{q}}\mathrm{EL}^y_{\hat q_j}(k))_j \quad H^x_{\overline{q}} P^y_{\overline{q}_b}(k)) \quad H^x_{\overline{q}}\mathrm{ev}^y_{\overline{q}_b}(k)) \Big)$$

Examples of geometric models

Lagrangian-based

$$(\mathrm{ODE})\ 0 = EL(L) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q}$$

$$(\mathrm{PDE})\ 0 = EL(L) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial u_t} + \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial u_x} - \frac{\partial L}{\partial u}$$

Discrete Lagrangian-based

$$(ODE)\ 0 = \nabla_{q^i}\left(L_d(q^{i-1}, q^i) + L_d(q^i, q^{i+1})\right)$$

$$(PDE)\ 0 = \nabla_{u^i_j}\Big(L_d(u^i_j, u^{i+1}_j, u^i_{j+1})$$

$$+ L_d(u^{i-1}_j, u^i_j, u^{i-1}_{j+1}) + L_d(u^i_{j-1}, u^{i+1}_{j-1}, u^i_j)\Big)$$

...

Hamiltonian-based

$$(\mathrm{sympl})\ \dot{z} = J^{-1}(z)\nabla H(z),\ z = (q,p)^\top$$

$$(\mathrm{multi-sympl})\ K u_t + L u_x = \nabla_u H(u)$$

Port-Hamiltonian

$$\dot{z} = (J(z) - R(z))\nabla H(z) + g(z) f, \quad f \text{ input}$$

$$e = g(z)^\top \nabla H(z) \qquad e \text{ output}$$

Discrete Hamiltonian-based

$$(\mathrm{sympl})\ \frac{z^{i+1} - z^i}{h} = J^{-1}\nabla H_d\left(\frac{z^{i+1} + z^i}{2}\right)$$

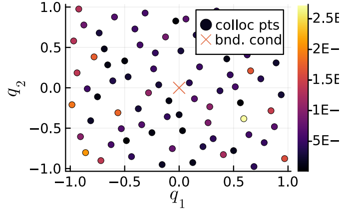
...

Idea

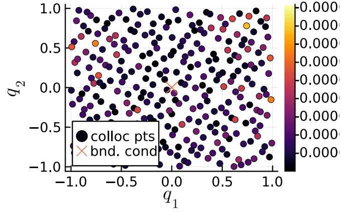
- Training geom. model $\iff$ solving pde for L, L_d, H, H_d, ...
- The pde is linear ☺

Experiment: Coupled harmonic oscillator on $[-1,1]^4 \subset TQ^2$

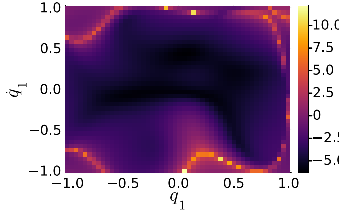
80



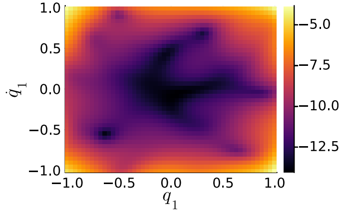
300



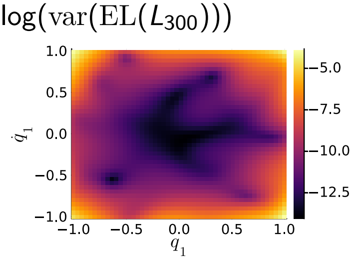
$\log(\mathrm{var}(\mathrm{EL}(L_{80})))$

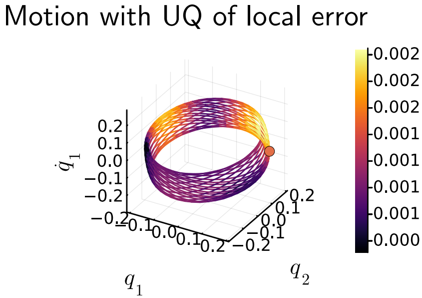


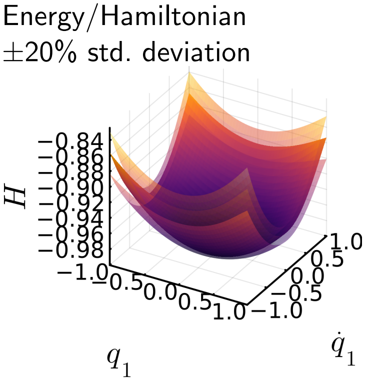
$\log(\mathrm{var}(\mathrm{EL}(L_{300})))$



Coupled harmonic oscillator





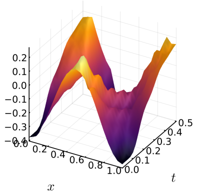


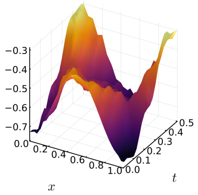
UQ

- Efficient UQ without sampling.
- No artificial uncertainty in equations of motions, Hamiltonian, sympl. structure

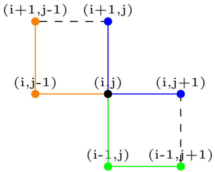
Wave equation

data (2 samples)

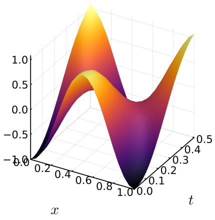




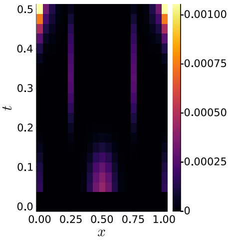
stencil $L_d(u_j^i, u_j^{i+1}, u_{j+1}^i)$



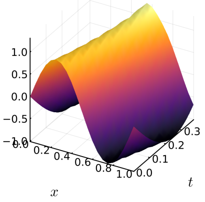
predicted



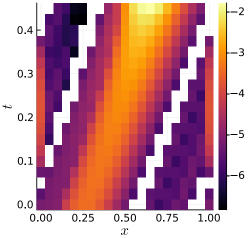
var



predicted



log(var)



Theorem

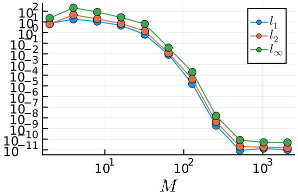
*The method converges in $\|\cdot\|_{\mathcal{C}^2}$ to **a** true L (RKHS compatibility assumptions).*

Theorem

The speed of convergence in L^2-norm of $\mathrm{EL}(L)$, the acceleration field, their variance is

$$h^r .$$

h: max. distance between data points
r: smoothness of underlying dynamics.
... provided that the kernel is smooth enough, r big enough.



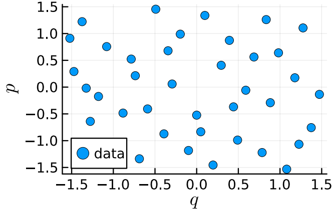
Proof.

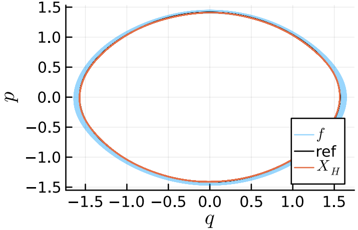
Interpolation & Smoothening theory. $\square$

C. Offen (June 2025). "Machine learning of continuous and discrete variational ODEs with convergence guarantee and uncertainty quantification". In: *Mathematics of Computation*. ISSN: 0025-5718. DOI: 10.1090/mcom/4120.

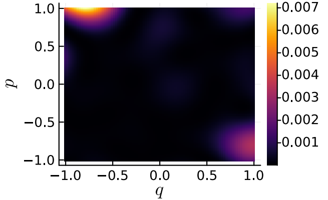
Geometry reduces model uncertainty

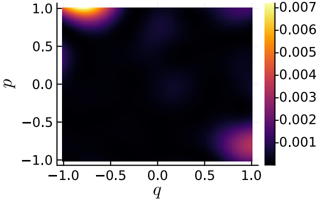
Pendulum
data $\dot z^{(j)} = (\dot q^{(j)}, \dot p^{(j)})$



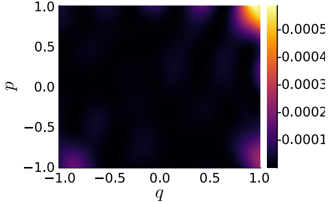


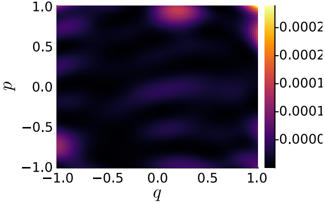
model $\dot z = f_\theta(z)$. Plot $\mathrm{var}(f_\theta)$





model $\dot z = X_{H_\Theta} = J^{-1}\nabla H_\theta(z)$. Plot $\mathrm{var}(X_{H_\Theta})$





10x less variance in equation of motions

Summary

Viewpoint

Training geometric model $\iff$ solving linear pde

Benefits

- Efficient UQ without sampling (but you need to think about regularisation)
- Convergence analysis
- Quantitative tool to describe benefit of geometry

C. Offen (June 2025). "Machine learning of continuous and discrete variational ODEs with convergence guarantee and uncertainty quantification". In: *Mathematics of Computation*. ISSN: 0025-5718. DOI: 10.1090/mcom/4120.

C. Offen (July 2024). *Machine learning of discrete field theories with guaranteed convergence and uncertainty quantification*. arXiv: 2407.07642.

References

- Offen, C. (July 2024). *Machine learning of discrete field theories with guaranteed convergence and uncertainty quantification*. arXiv: 2407.07642.
- — (June 2025). "Machine learning of continuous and discrete variational ODEs with convergence guarantee and uncertainty quantification". In: *Mathematics of Computation*. ISSN: 0025-5718. DOI: 10.1090/mcom/4120.

GitHub Repositories

Neural Networks	`https://github.com/Christian-Offen/DLNN_pde`
Gaussian Processes	`https://github.com/Christian-Offen/Lagrangian_GP`
	`https://github.com/Christian-Offen/Lagrangian_GP_PDE`

Ministry of Culture and Science
of the State of
North Rhine-Westphalia

PhoQS

Geometric versus non-geometric stencil

Geometric model

$$\begin{aligned} 0 &= \nabla_1 L_d(u_j^i, u_j^{i+1}, u_{j+1}^i) \\ &+ \nabla_2 L_d(u_j^{i-1}, u_j^i, u_{j+1}^{i-1}) \\ &+ \nabla_3 L_d(u_{j-1}^i, u_{j-1}^{i+1}, u_j^i) \end{aligned}$$

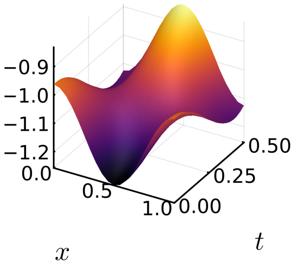
$\implies$ Learn $L_d \colon (\mathbb{R}^n)^3 \to \mathbb{R}$ / linear pde

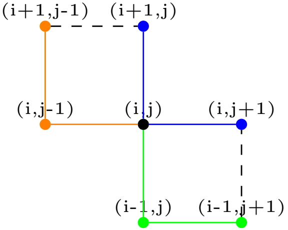
Non-geometric model

$$0 = F(u_j^i, u_j^{i+1}, u_{j+1}^i, u_j^{i-1}, u_{j+1}^{i-1}, u_{j-1}^i, u_{j-1}^{i+1}).$$

$\implies$ Learn $F \colon (\mathbb{R}^n)^7 \to (\mathbb{R}^n)^7$ (stencil)

$u_{j-1}^{i+1} = f(u_j^i, u_{j+1}^i, u_j^{i-1}, u_{j+1}^{i-1}, u_{j-1}^i, u_j^{i+1}) \implies$ stability issues





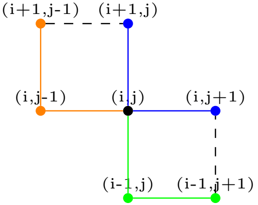
PDE Example

Task

Learn a discrete Lagrangian density $L_d(u_j^i, u_j^{i+1}, u_{j+1}^i)$ to model (unknown) PDE.

Here

$$\begin{aligned}\mathrm{DEL}(L_d) = \frac{\partial}{\partial u_j^i}\Big(&L_d(u_j^i, u_j^{i+1}, u_{j+1}^i)\\ &+L_d(u_j^{i-1}, u_j^i, u_{j+1}^{i-1})\\ &+L_d(u_{j-1}^i, u_{j-1}^{i+1}, u_j^i)\Big)\end{aligned}$$

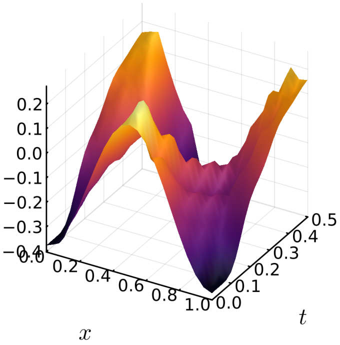


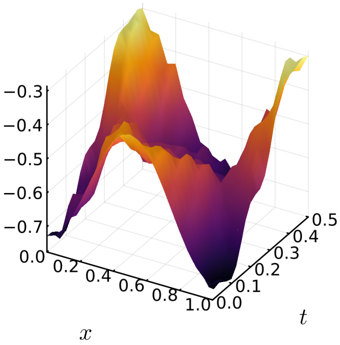
Normalisation

Normalise abs. value and local conjugate momentum of L_d at base point.

Wave equation (Lagrangian/discrete)

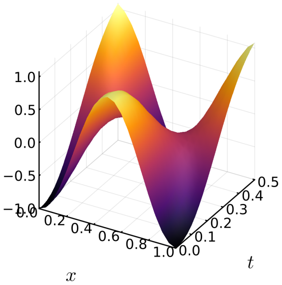
Training data (2 Samples)



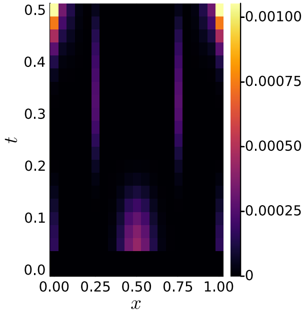


Wave equation (Lagrangian/discrete)

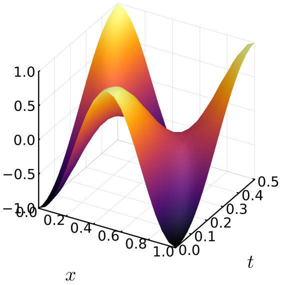
predicted



var(DEL)

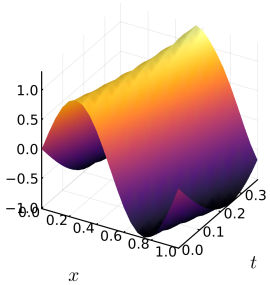


reference

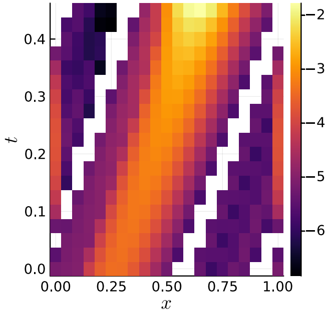


Wave equation (Lagrangian/discrete)

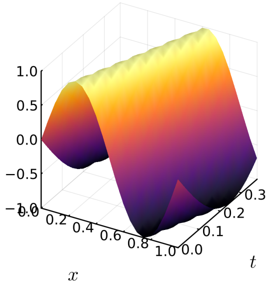
predicted



$\log(\mathrm{var}(\mathrm{DEL}(L_d)))$



reference



C. Offen (July 2024). *Machine learning of discrete field theories with guaranteed convergence and uncertainty quantification*. arXiv: 2407.07642.