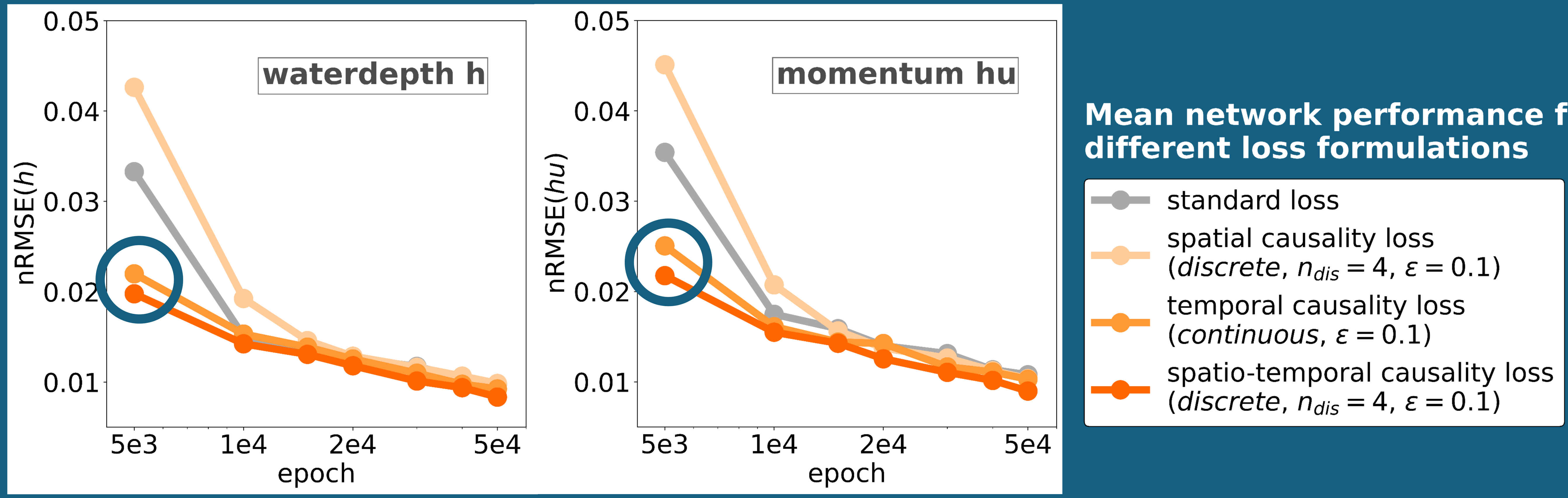
Respecting Temporal or Spatio-Temporal Evolution of the Underlying Partial Differential Equation in the Loss Term Speeds Up the Convergence of Physics-Informed Neural Networks (PINNs)



Mean network performance for different loss formulations

Impact of Causality-Based Losses on Physics-Informed Neural Networks for 1D Shallow Water Equations

Henrik Schmieder*, Andreas Malcherek

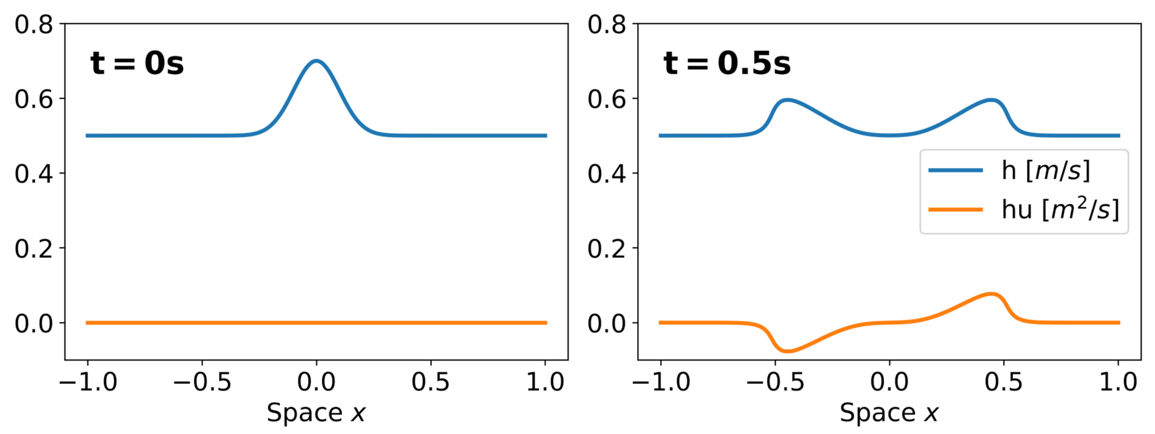
Introduction

- Physics Informed Neural Networks (PINNs) include an additional residual loss term $\mathcal{L}_{res}$ to evaluate performance against the system underlying partial differential equation.
- Previous studies showed that if this loss is reformulated to respect temporal causalities of the PDE, PINNs converge for complex, high fidelity PDEs where standard PINNs fail.

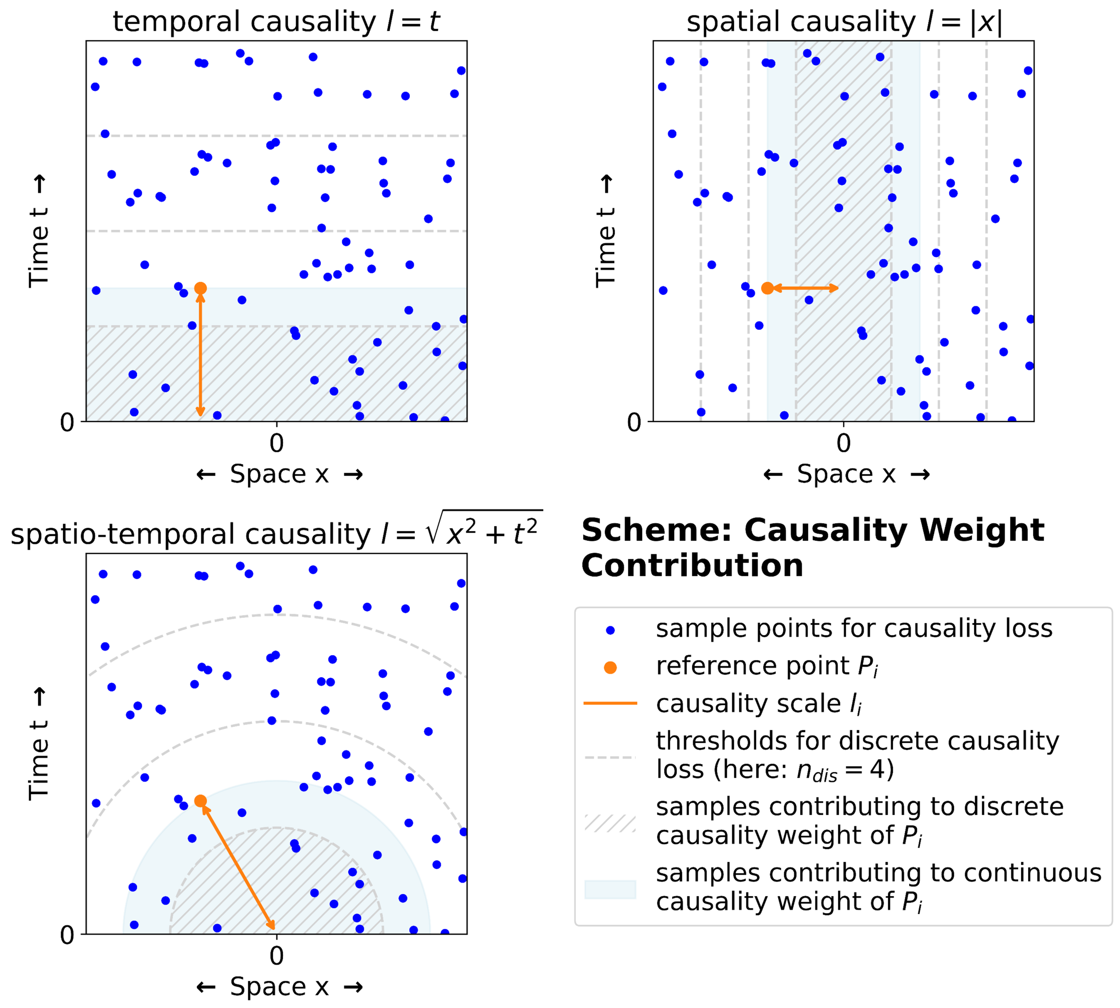
Task

- We examine whether the causality loss improves performance if applied to physical systems where PINNs with standard loss also converge.
- We extend the loss principle to spatial and spatio-temporal dimension.
- We investigate this tasks on the example of 1D Shallow Water equations.

1D Shallow Water Equations (SWE)



- SWE describes shallow fluid flow influenced by gravity, accounting for variations in water depth (h) and momentum (hu).
- The reference case is an initial Gaussian water level with zero momentum in the domain $x \in [-1,1]$, $t \in [0,1]$.



Scheme: Causality Weight Contribution

Causality Loss

- Standard Approach: Residual loss of each sample contributes equally to total loss per epoch.
- Idea: Inspired by PDE evolution in time and space, causality loss $\mathcal{L}_{caus}$ prioritizes earlier sampling points, addressing later points only after reducing earlier losses.
- Formulation based on Wang et. al [2022] †:

$$\mathcal{L}_{caus} = \frac{1}{n}\sum_{I=1}^{n} w_i \mathcal{L}_{res,i}\ ;\ w_i = e^{-\epsilon \sum_{j=1}^{n} \mathcal{L}_{res,j}(l_j < l_i)}$$

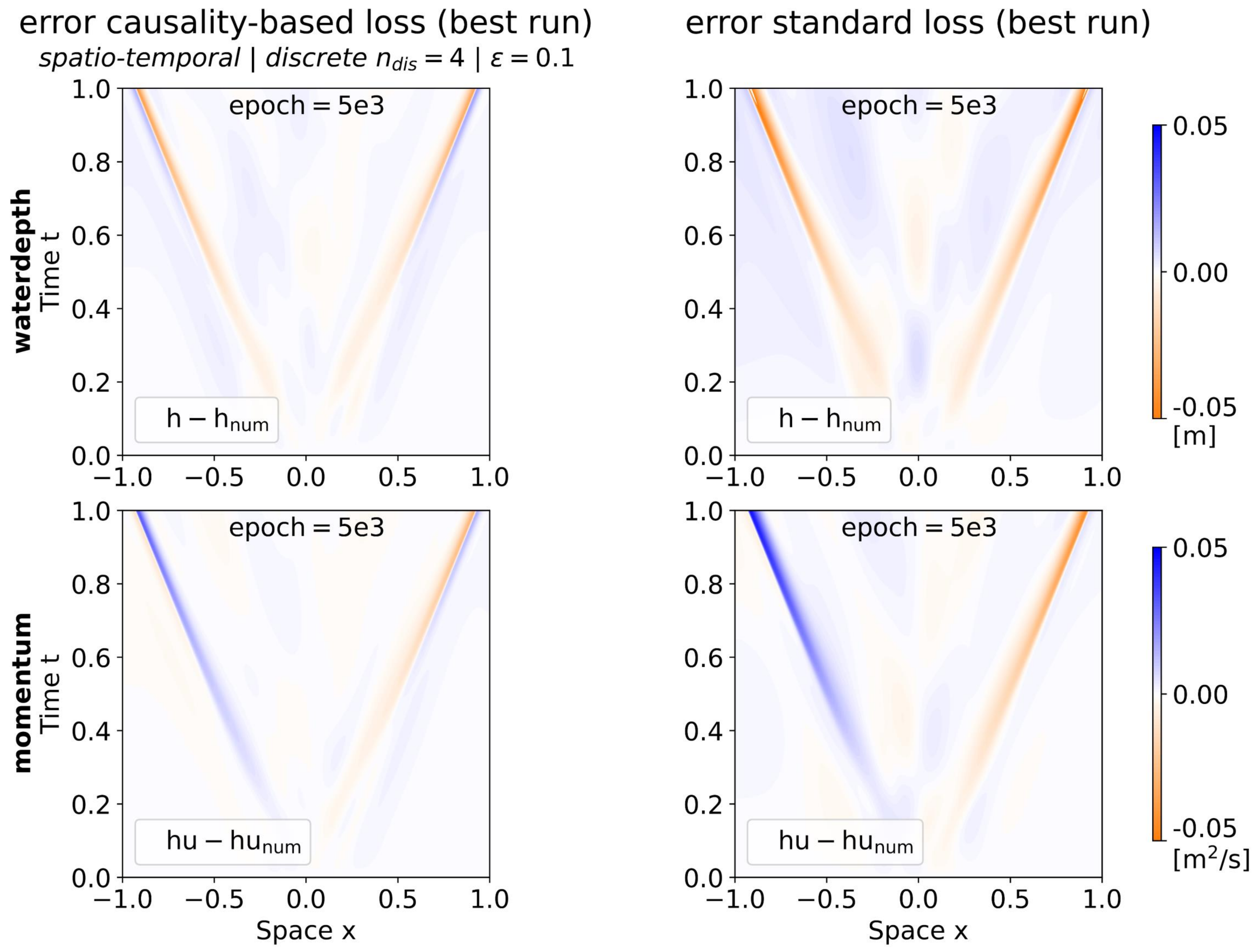
- Causality scale l defines respected system evolution dimension: spatio-temporal: $l = \sqrt{x^2 + t^2}$
  temporal: $l = t$; spatial: $l = |x|$;
- Causality weight w_i of a sample point P_i is small if previous ($l < l_i$) losses are large and vice versa.
- ϵ: hyperparameter controlling steepness of w
- Additionally, causality loss is formulated in discrete form (here: $n_{dis} = 4{,}16$).
- Continuous and discrete settings are tested for $\epsilon = 0.01, 0.1, 0.5, 1, 10$.

PINN Settings

Setting	Details
Input, Output	[x, t], [h, hu]
Constraint Strategy	hard constrained initial conditions
Layout	5 layer, 30 neurons each
Optimizer	ADAM ($lr = 0.001$)
Activation	Tanh
Batch approach	Full batch
Sample data	Boundary Condition Loss.: 1e3 Causality loss: 1e4
Trials	10 for each setting
Evaluation	normalized RMSE w.r.t numerical solution

Results

- Improved initial convergence for temporal and spatio-temporal causal losses (reduced error up to 40% after 5e3 epochs).
- Final performance is similar for all loss formulations.
- Possible reason: Simplified initial loss surface may guide the network effectively to optimal regions.



Primer on PINNs

- general first-order temporal PDE:
  $\boldsymbol{u}_t + \mathrm{N}(\boldsymbol{u}) + c(t, \boldsymbol{x}) = 0, t \in [0, T], x \in \Omega$
- with I.C. and B.C.: $\boldsymbol{u}(t = 0, \boldsymbol{x}) = f_{I.C.}(\boldsymbol{x})$
  $\boldsymbol{u}(t, \boldsymbol{x} \in \partial\Omega) = f_{B.C.}(\mathrm{t}, \mathbf{x})$
- approximation of $\boldsymbol{u}$ by a MLP with tunable parameters $\boldsymbol{\theta}$:
  $\boldsymbol{u} \approx \tilde{\boldsymbol{u}} = PINN(t, \boldsymbol{x}, \boldsymbol{\theta})$
- residual loss formulation:

$$\mathcal{L}_{res} = \frac{1}{n}\sum_{i=1}^{n} \mathcal{L}_{res,i} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial \tilde{\boldsymbol{u}}_i}{\partial t} + \mathrm{N}(\tilde{\boldsymbol{u}}_i) + c(t_i, \boldsymbol{x}_i)\right)^2$$

- no knowledge of true $\boldsymbol{u}$ is required for $\mathcal{L}_{res}$
- partial derivatives of $\boldsymbol{u}$ are derivatives w.r.t. network inputs and can be calculated with auto-differentiation

1D Shallow Water Equations

mass: $\frac{\partial h}{\partial t} + \frac{\partial hu}{\partial x} = 0$; momentum: $\frac{\partial hu}{\partial t} + \frac{\partial}{\partial x}\left(hu^2 + \frac{1}{2}gh^2\right) = 0$

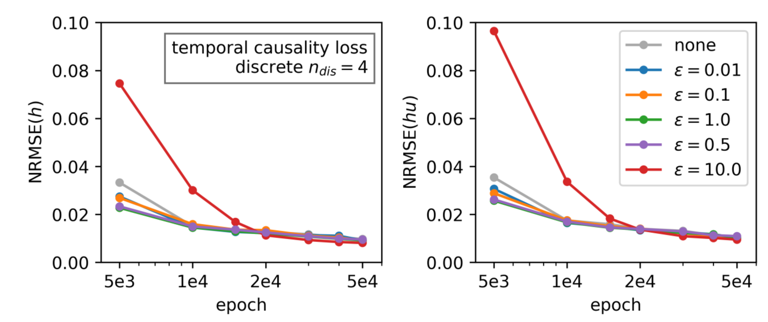
PINN Settings

- hard constrained I.C.: $\begin{bmatrix}\tilde{h}\\ \widetilde{hu}\end{bmatrix} = \mathrm{f_{IC}} + \mathrm{t} * \mathrm{PINN}(\mathrm{x}, \mathrm{t}, \boldsymbol{\theta})$
- total loss: $\mathcal{L} = \lambda_1\mathcal{L}_{caus,h} + \lambda_2\mathcal{L}_{caus,h} + \lambda_3\mathcal{L}_{B.C.,h} + \lambda_4\mathcal{L}_{B.C.,h}$ with $\lambda_{1-4} = 1$
- add. comp. costs for $\mathcal{L}_{caus}$ are moderate ($\mathcal{L}_{res}$ is computed anyway, autograd can be turned off for causality weights w)

Discrete Causality Weight

- $w_i = w_{dis,k} = e^{-\epsilon\sum_{j=1}^{n}\mathcal{L}_{res,j}(l_j < l_{dis,k})}$; if $l_{dis,k} < l_i < l_{dis,k+1}$
- $l_{dis,k} = \frac{k-1}{n_{dis}} L$; L: max. value of l in domain

Effect of ϵ



†: Wang S, Sankaran S, Perdikaris P. *Respecting causality for training physics-informed neural networks*. Computer Methods in Applied Mechanics and Engineering. 2024

*: henrik.schmieder@unibw.de
Institute for Hydroscience, Bundeswehr University Munich
building 41/100 4128, phone +49 89 6004 3376

Example code and poster:

SCAN ME

Institute of Hydroscience

Universität der Bundeswehr München