CRITEO AI Lab

SORBONNE UNIVERSITÉ

Confluences Mathématiques
Le machine learning au pays des équations
Rencontres chercheur·euse·s et ingénieur·e·s - 6ème édition
Paris, le 21 novembre 2024

Operator learning : solving partial differential equations on general geometries

Patrick Gallinari
Sorbonne University & CRITEO AI Lab – Paris
patrick.gallinari@sorbonne-universite.fr, p.gallinari@criteo.com

Outline

- Neural PDE solvers & neural operators
- Two families of neural operators
  - CORAL: Operator Learning with Neural Fields
    - **Neural fields** for encoding continuous functions
  - AROMA: Attentive Reduced Order model with Attention
    - **Attention/ Transformers** for encoding spatiality in latent space

Neural PDE solvers: Tasks and objectives

Objectives

- Surrogate models for solving PDEs or spatio-temporal forecasting
- Accelerate simulation, complement physical models, design, etc

Approaches

- Pure data-driven approaches
  Learn from observations or simulations
- Hybrid approaches
  Leverage prior physical background + information extracted from data
- Data free approaches
  PDE loss only

Numerical solvers & Neural solvers

- Classical numerical solvers operate on grids or meshes
  - Finite differences, finite elements, finite volumes
- Neural solvers operate on tensors (grids) or graphs (irregular meshes)

Neural PDE solvers: Learning operators

Instead of learning maps between vector space, learn maps between infinite input and output function spaces

Images for example are considered as continuous functions

Key motivations

- Mesh free operators

  Handle general geometries, resolution independence
- Multi-resolution
- High dimensional problems
- Interpolate between function spaces instead of vector spaces

  e.g. solve parametric equations: varying I/B conditions, forcing terms, equation coefficients , ...

Representative neural operators

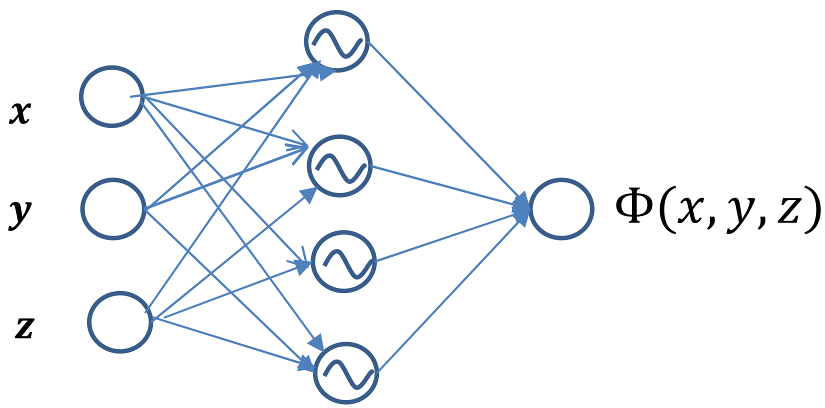
- Fourier Neural operators (Li et al. 2021) - Stanford
- Deep-Onet (Lu et al. 2021) – Brown Univ.

CORAL:COordinate-based model for opeRAtor Learning

Serrano et al. , 2023, Operator Learning with Neural Fields: Tackling PDEs on General Geometries, NeurIPS

Neural Fields (Implicit Neural Representations)

- Coordinate-based approximation of functions
  - Continuous representations of objects as coordinate-dependent functions
  - Appeared initially as a novel way to represent 3D shapes in place of discrete representations
    - Example: signed distance



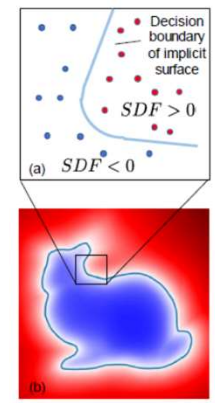
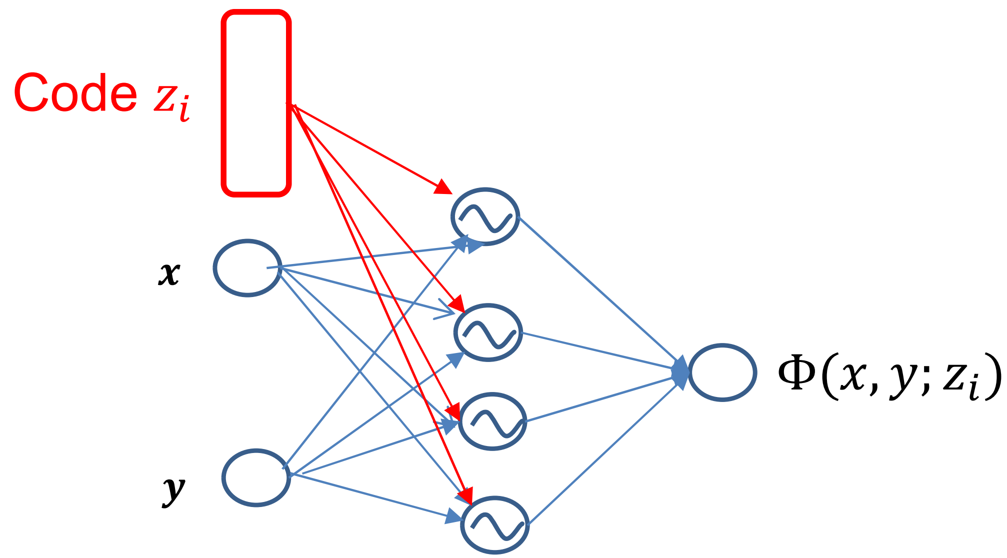


Fig. Park et al. 2019

    - The shape is fully described by the NN parameters
    - Mesh-free approach – independent of the resolution: learn from point sets
    - Lower memory requirements than discrete representations
  - References: Sitzmann et al. 2020, Fathony et al., 2021, Tancik et al. 2020, etc

Neural Fields (Implicit Neural Representations)

- Learning several images
  - A neural field model represents one image
  - How to represent multiple images using a single model?
    - Condition the neural field on a compact code specific of an image



    - This code z_i could be learned e.g. through auto encoding by gradient descent and is specific to an image
    - Conditioning is performed through an hypernetwork
    - Network weights (in blue) are shared across images

CORAL : Operator Learning with Neural Fields (Serrano et al. 2023)

- Tasks

Direct forecasting	$u_0 \rightarrow u_T$
Dynamics modeling	$u_t \rightarrow u_{t+1}$
Geometry aware inference	$a \rightarrow u_{sol}$

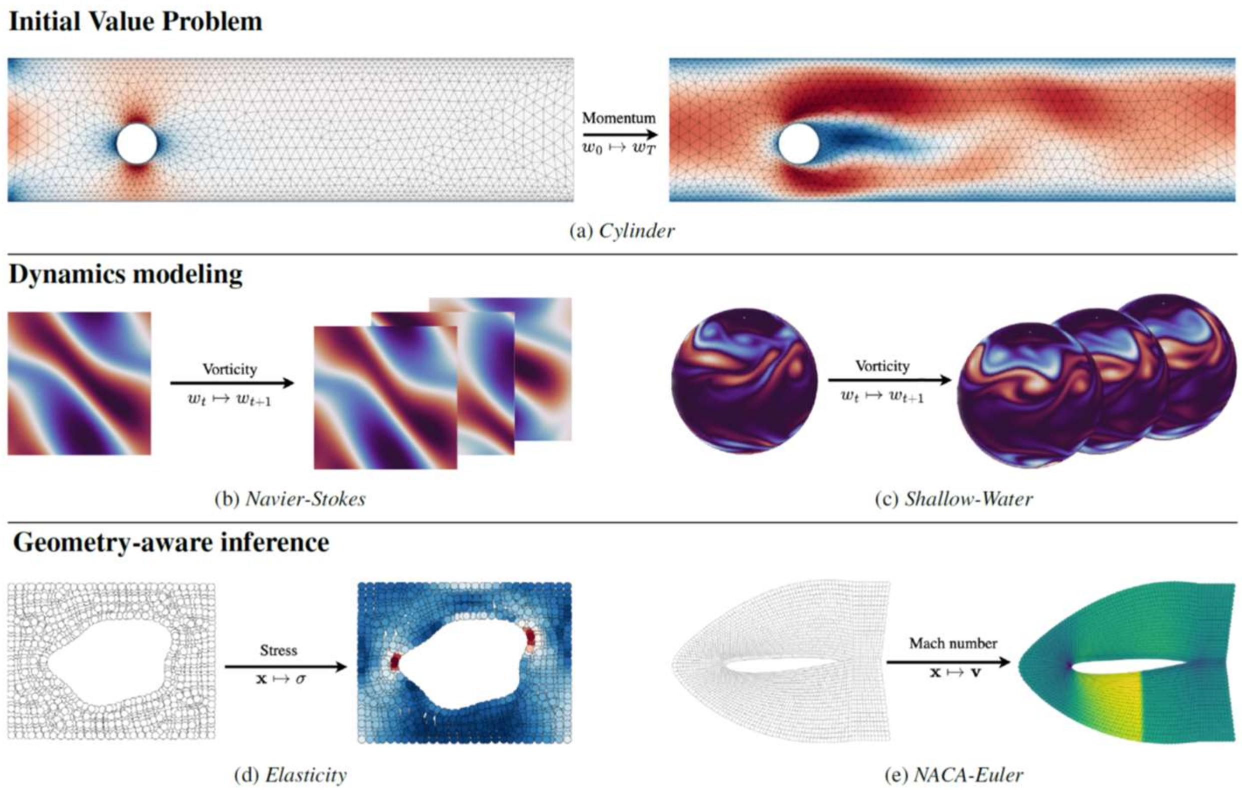
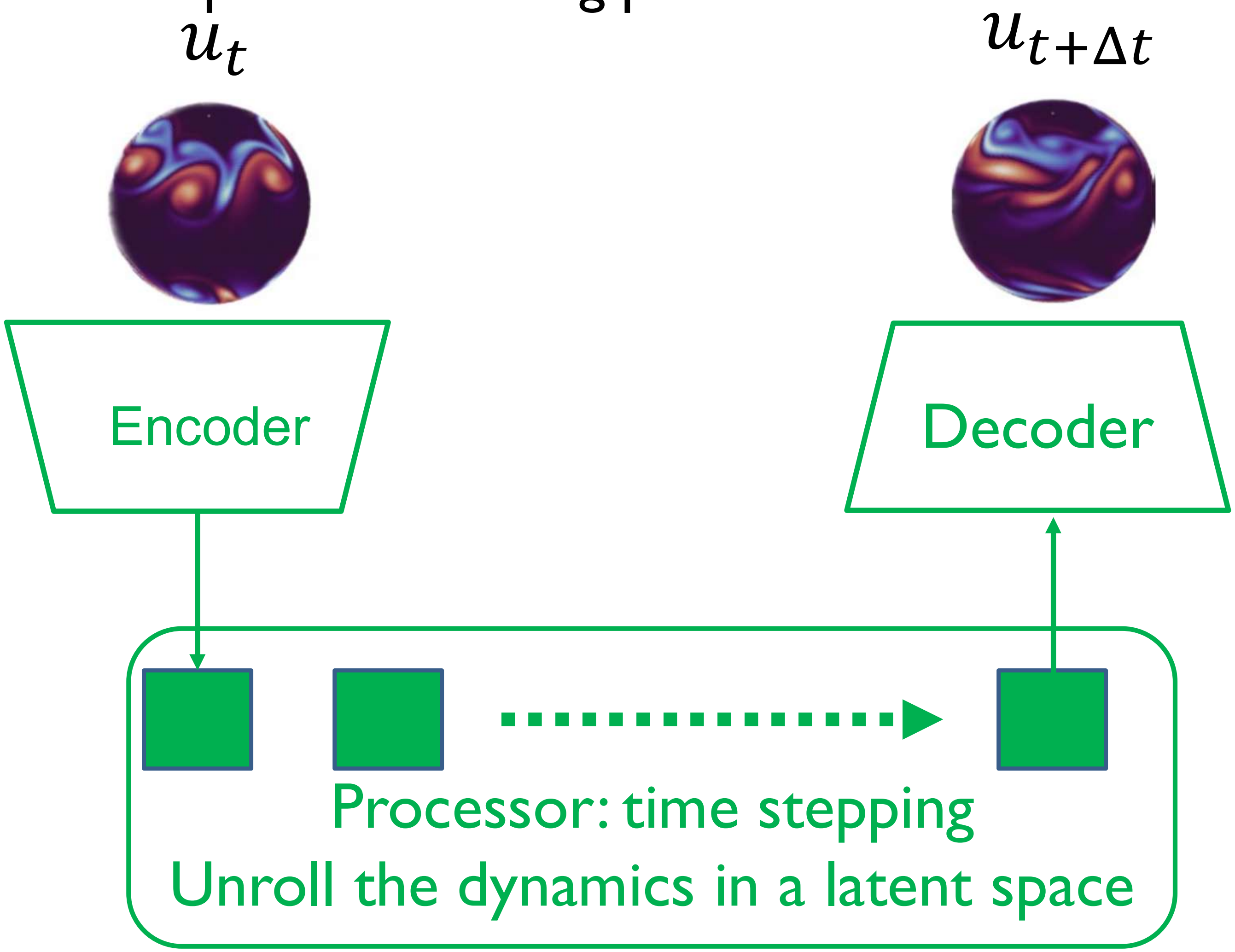


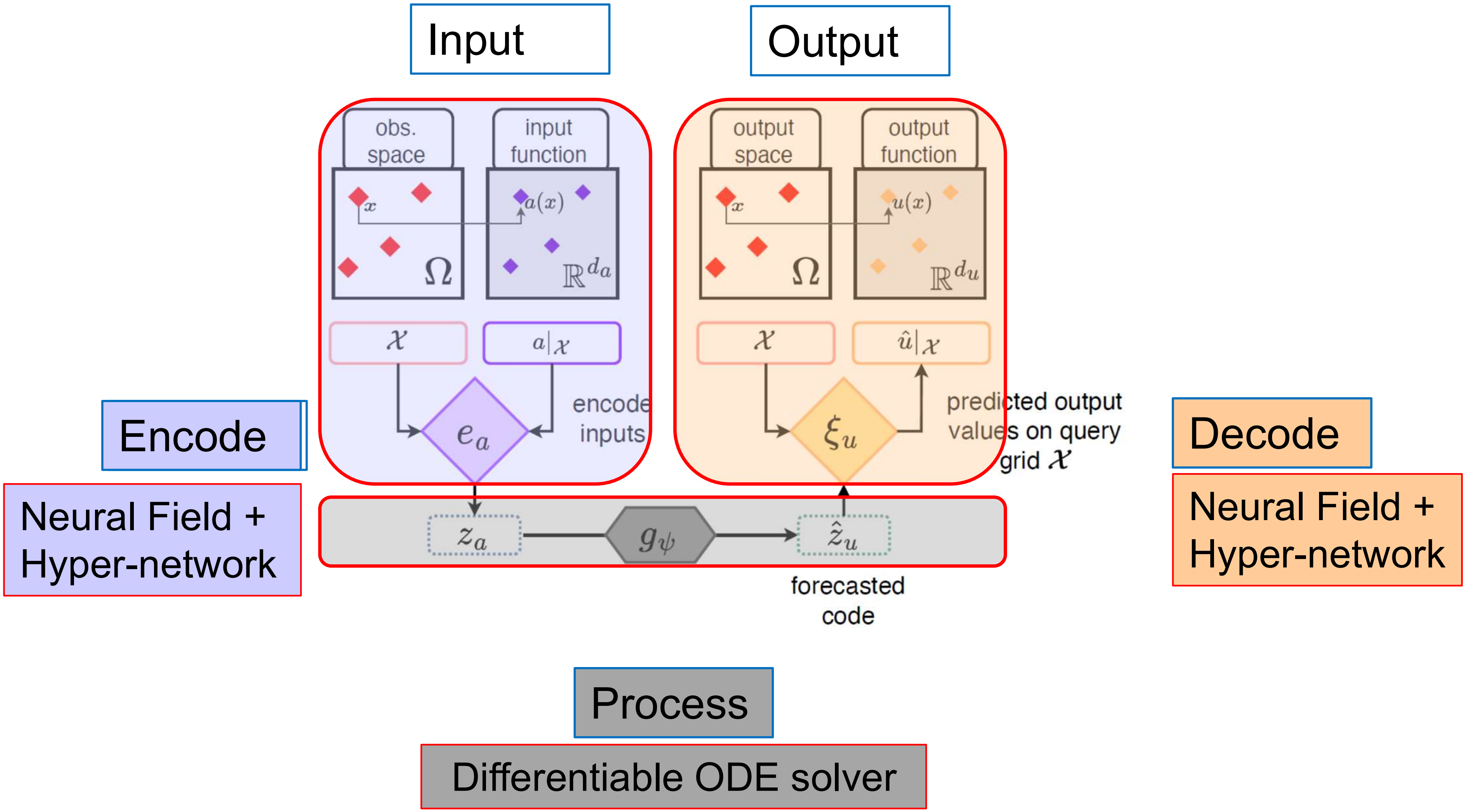
Figure 1: Illustration of the problem classes addressed in this work: Initial Value Problem (IVP) (a), dynamic forecasting (b and c) and geometry-aware inference (d and e).

C3: Neural operator
Encode – Process – Decode framework

Encode-Process-Decode has become the standard framework for many spatio-temporal forecasting problems



CORAL : Operator Learning with Neural Fields
Inference



CORAL : Operator Learning with Neural Fields

- Example: IVP on Airfoil (predict pressure, density, velocity)

CORAL can forecast physical fields from different initial conditions with different boundary conditions

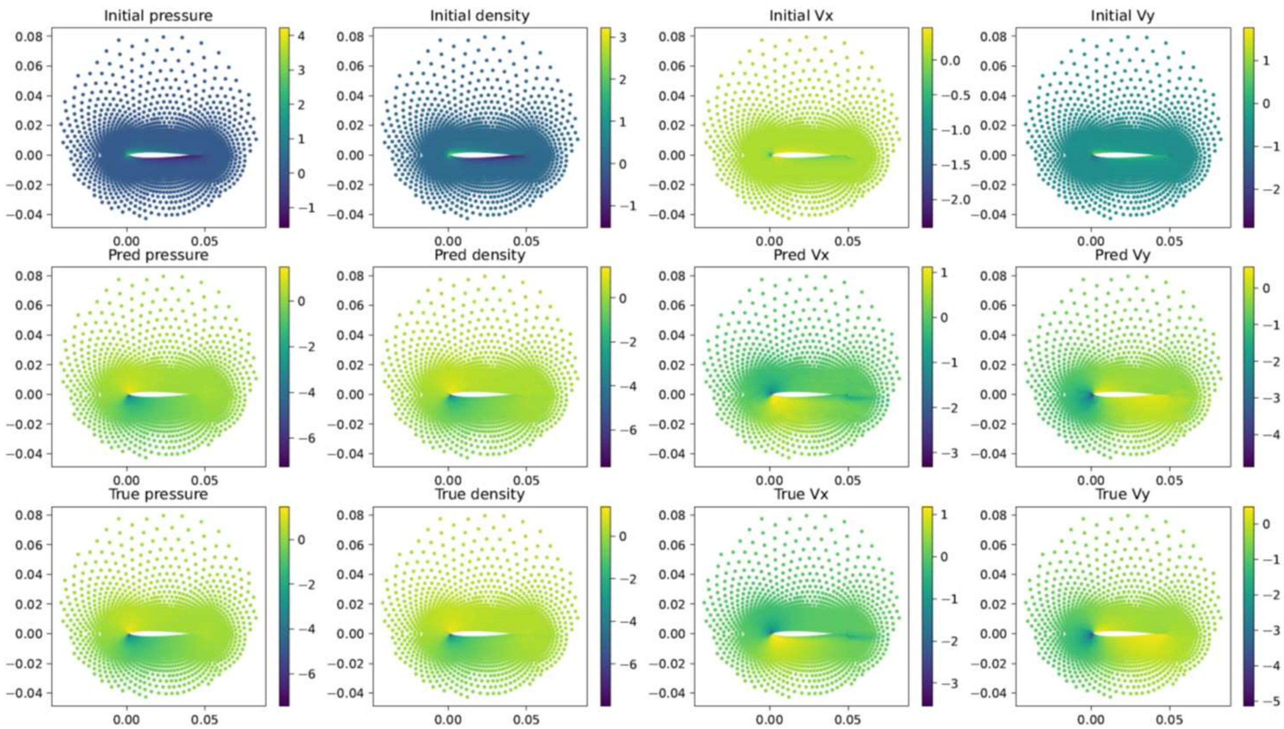


Figure 11: CORAL prediction on *Airfoil*

CORAL : Operator Learning with Neural Fields

- Example: forecasting on Shallow-Water (vorticity)

CORAL shows strong robustness to changes of grid and can extrapolate in time

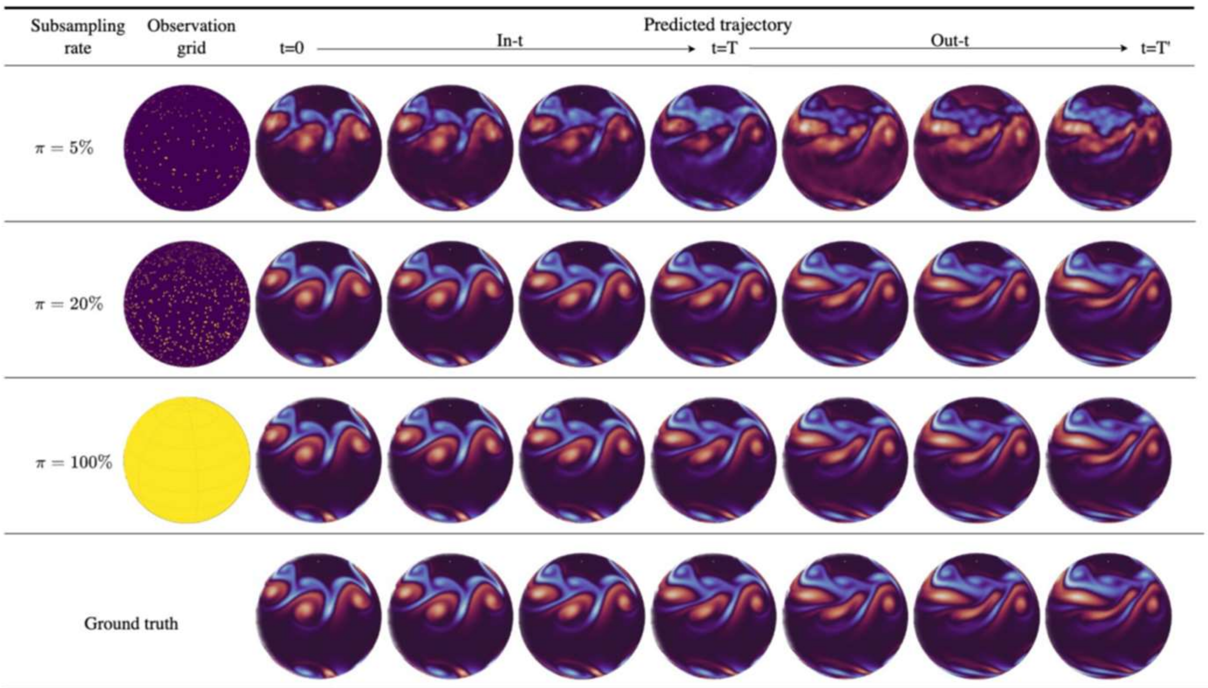


Figure 13: Prediction MSE per frame for CORAL on *Shallow-Water* with its corresponding training grid $\mathcal{X}$. Each row corresponds to a different sampling rate and the last row is the ground truth. The predicted trajectory is predicted from $t = 0$ to $t = T'$. In our setting, $T = 19$ and $T' = 39$.

CORAL : Operator Learning with Neural Fields

- Geometry aware inference: NACA-Euler (Mach number)

CORAL is on par with baseline for geometry-aware inference

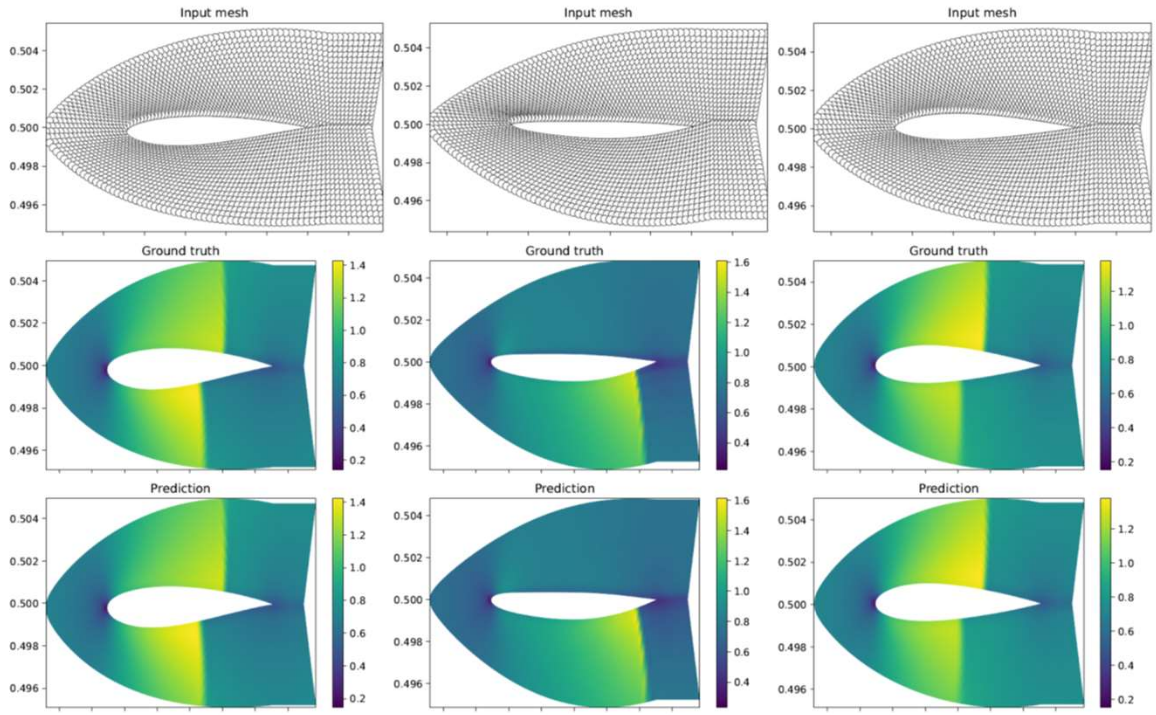


Figure 14: CORAL predictions on *NACA-Euler*

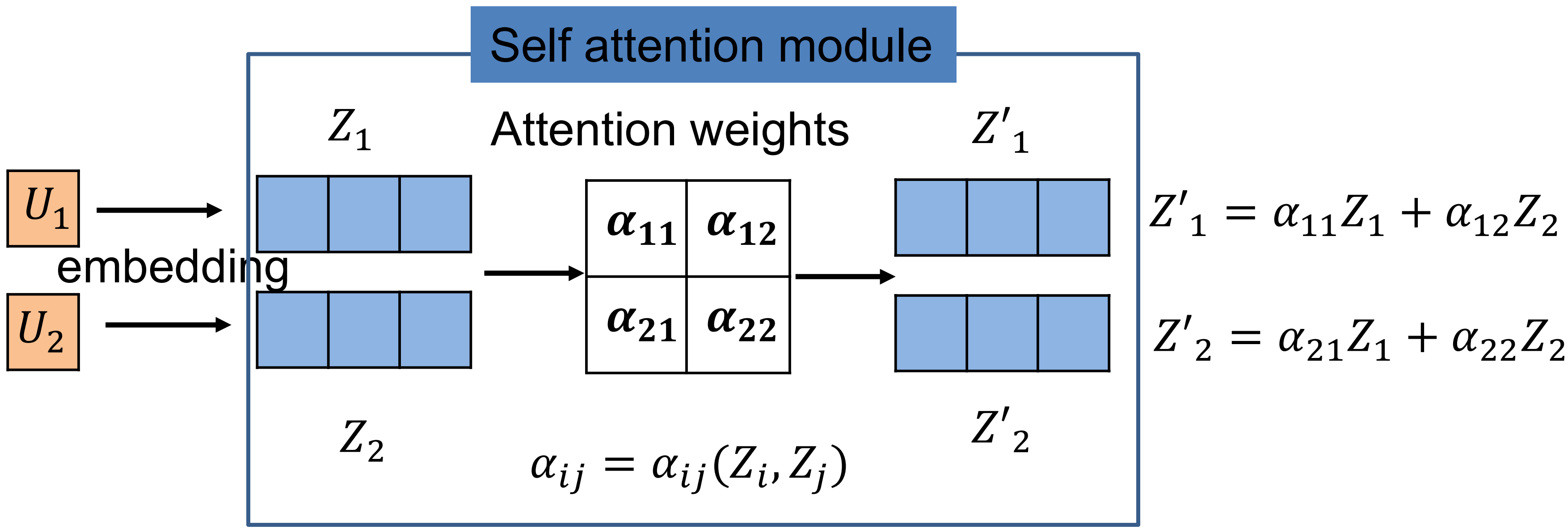
AROMA: Attentive Reduced Order Model with Attention

Serrano et al. , 2024, AROMA: Preserving Spatial Structure for Latent PDE Modeling with Local Neural Fields, NeurIPS

AROMA: Attentive Reduced Order Model with Attention (Serrano et al. 2024)

- Principled Framework:
  - Properties
    - Handle diverse geometries: inputs and outputs may consist in point sets, grids, meshes
    - Can be queried at any spatial position
  - Demonstrates how modern NN components allow building versatile PDE solvers
    - Encode/ Process/ Decode framework
      - Encoding: cross-attention maps variable-size inputs to a fixed-size compact latent token space encoding local spatial information
      - Processing: a diffusion transformer architecture to model dynamics and exploit spatial relations locally and globally via self-attention + model uncertainty
      - Decoding: uses a conditional neural field + cross-attention to query forecast values at any spatial point within the equation's domain

AROMA: Attentive Reduced Order Model with Attention
Attention – Self attention



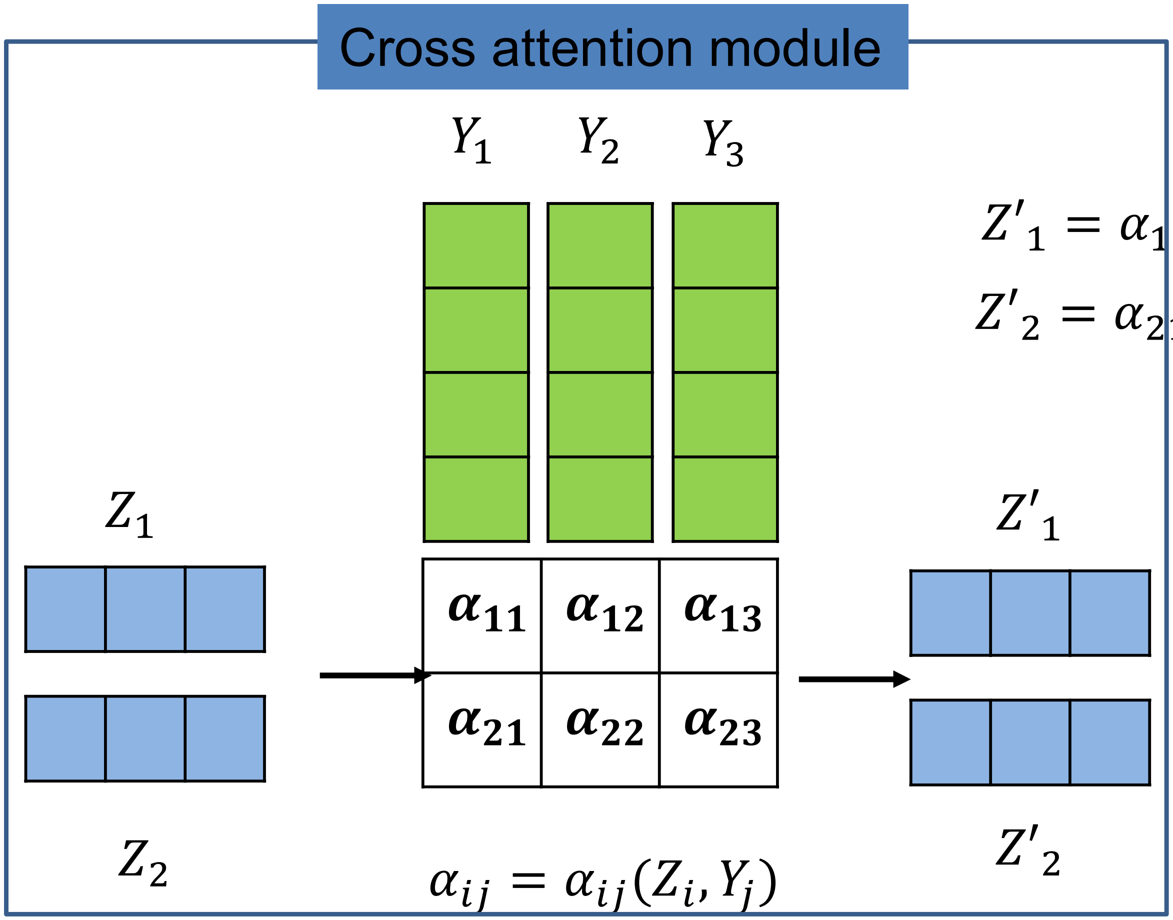
Captures contextual representation of the inputs

Complexity $O(N^2)$ with N the size of the input sequence

AROMA: Attentive Reduced Order Model with Attention

Attention – Cross attention

Cross attention maps a sequence of vector (Y) of variable size N into a sequence of vector (Z') of fixed size M



$$Z'_1 = \alpha_{11} W Y_1 + \alpha_{12} W Y_2 + \alpha_{13} W Y_3$$

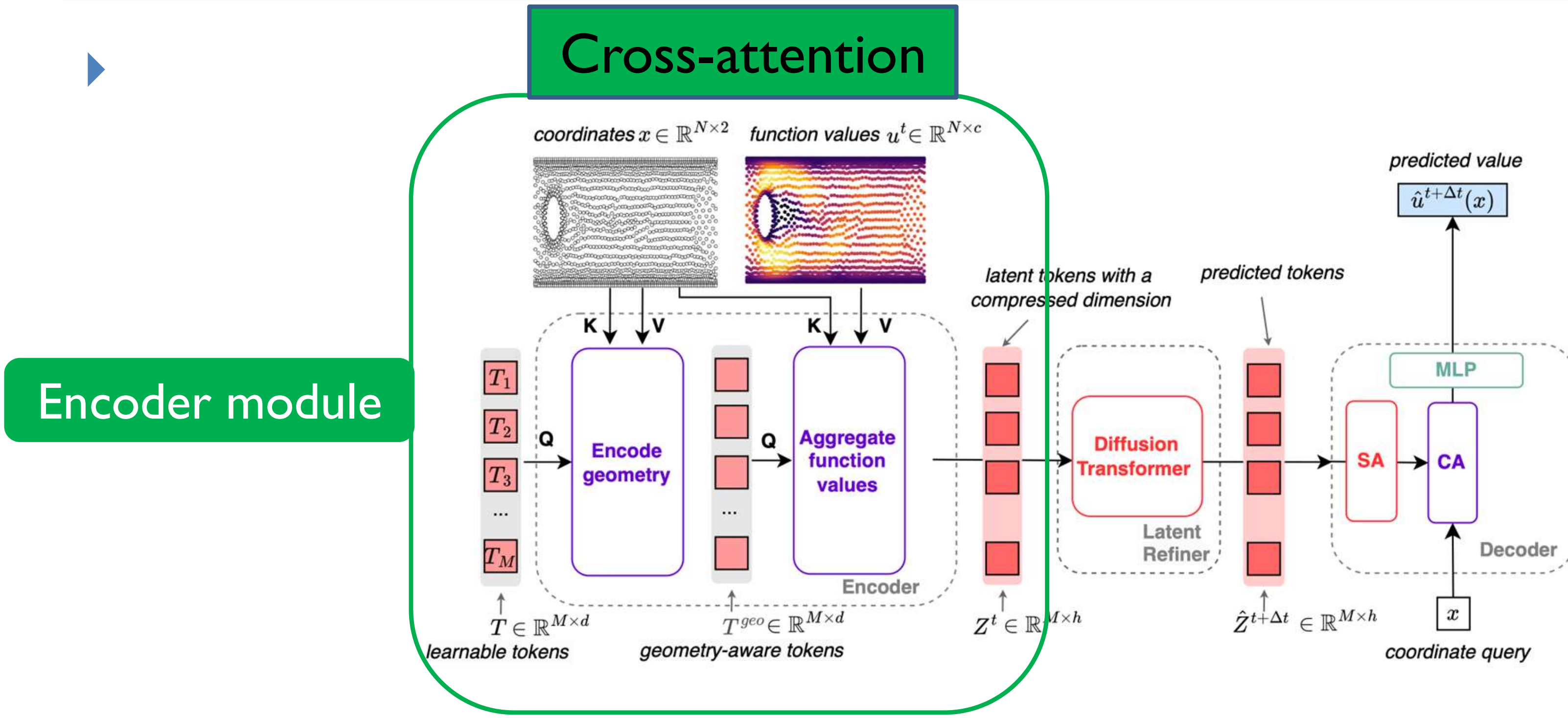
$$Z'_2 = \alpha_{21} W Y_1 + \alpha_{22} W Y_2 + \alpha_{13} W Y_3$$

W projection matrix

Attention weights

Complexity: $O(N(Y)N(Z))$

AROMA: Attentive Reduced Order Model with Attention (Serrano et al. 2024)
General framework

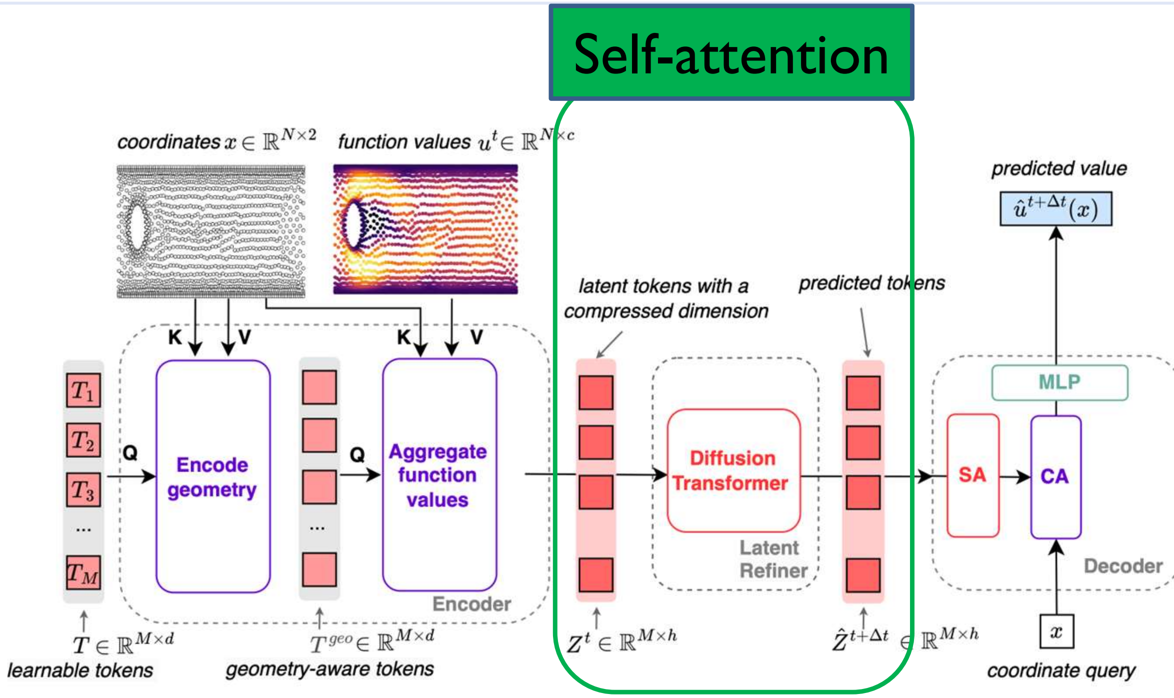


Cross-attention encoder: $u^t \rightarrow Z^t$

- Encodes variable size discretized input $u()$ into a fixed size & small dimensional sequence of latent embedding tokens Z
- Z encodes local spatial information on problem geometry + variable values

AROMA: Attentive Reduced Order Model with Attention (Serrano et al. 2024)
General framework

Processor



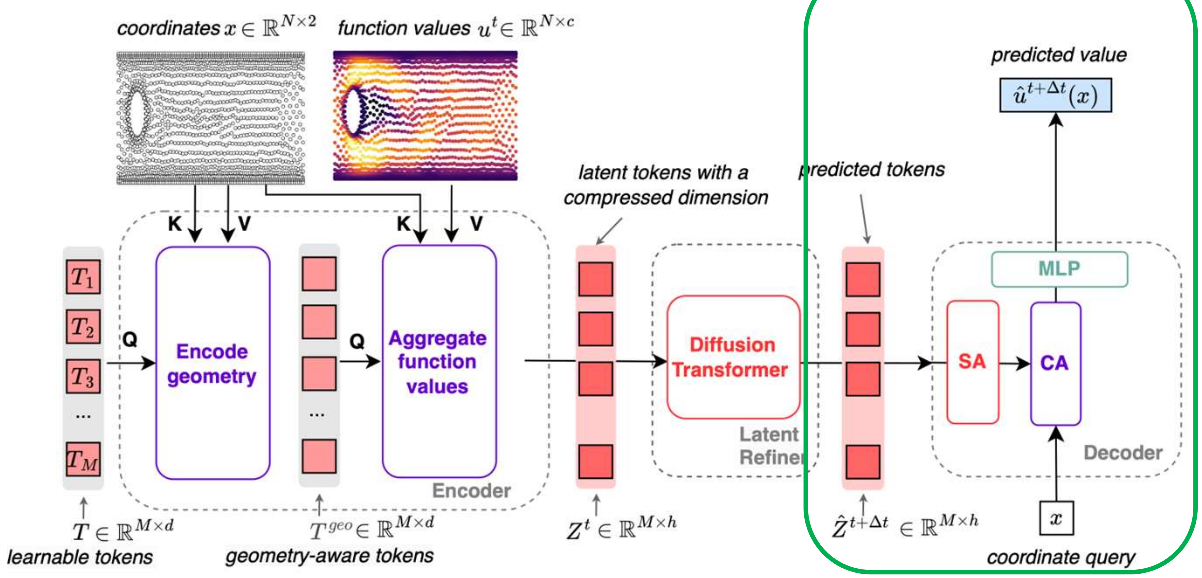
Time stepping transformer: $Z^t \rightarrow Z^{t+\Delta t}$

- Learns the dynamics in the small dimensional latent space
- Self attention models relations between spatial latent tokens
- Inference: dynamics is enrolled in the latent space starting from an initial condition– low complexity
- Diffusion: introduces a stochastic component

AROMA: Attentive Reduced Order Model with Attention (Serrano et al. 2024)
General framework

Cross-attention

Decoder module



Cross-attention neural fields decoder: $Z^{t+\Delta t} \rightarrow u^{t+\Delta t}$

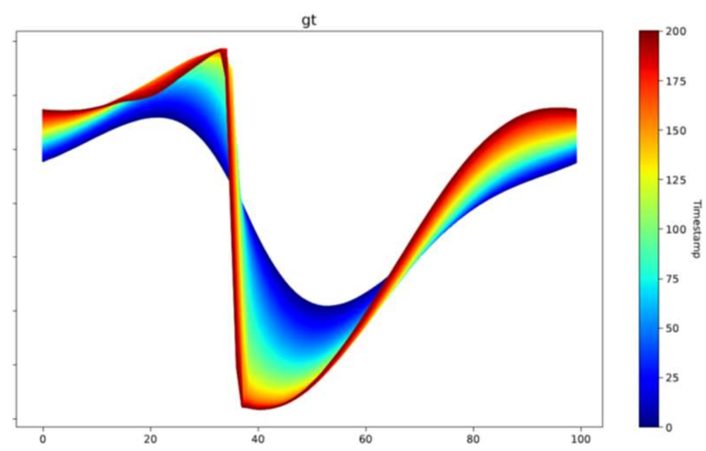
- Maps the latent representation $Z^{t+\Delta t}$ to the original physical space
- Can be queried at any position x of the physical space

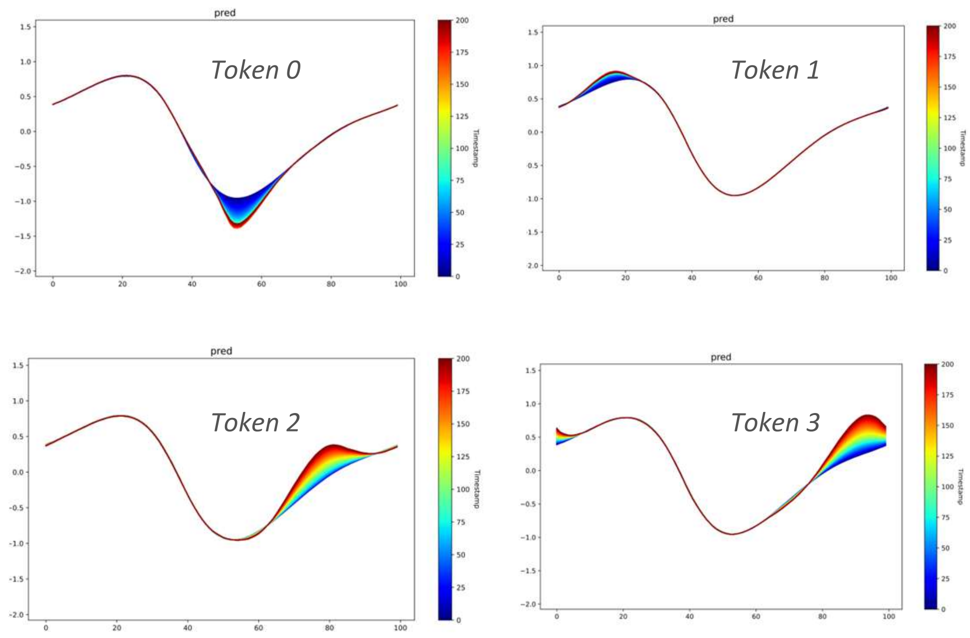
AROMA: Attentive Reduced Order Model with Attention
Cross-attention encoder captures spatial attention

Example: Burgers equation – perturbation analysis on the tokens

Burgers equation ground truth

Tokens encode local spatial information

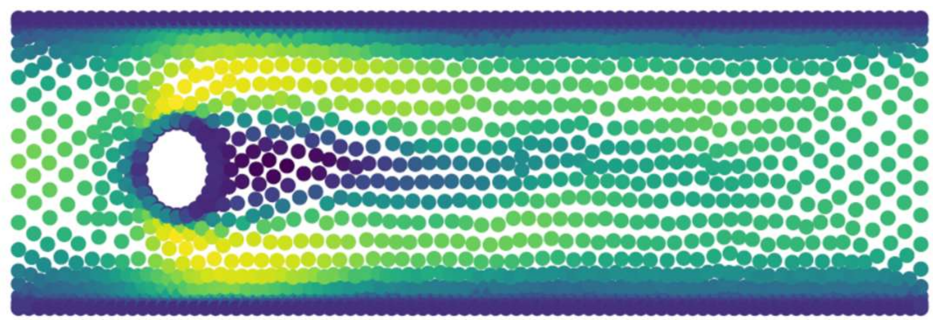




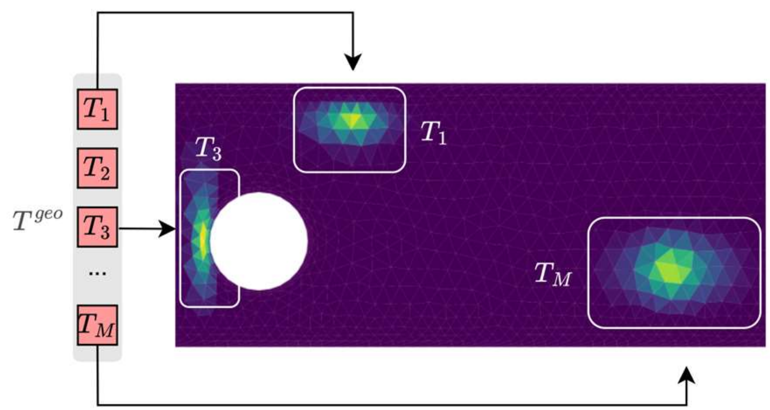
AROMA: Attentive Reduced Order Model with Attention
Cross-attention encoder captures spatial attention

Example: Cross attention on cylinder flow

Cylinder flow ground truth

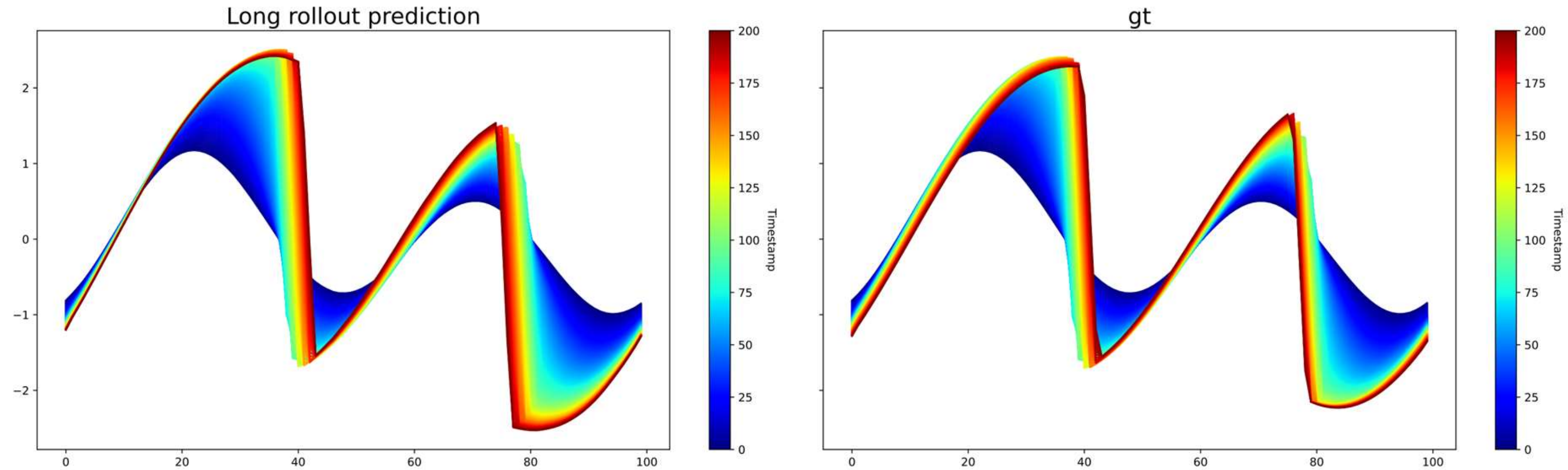


Tokens encode local spatial information – cross attention between T^{geo} tokens and "x"



AROMA: Attentive Reduced Order Model with Attention
Stability on long rollouts

- Trained to predict next step on 50 time steps trajectories
- Unrolled for 200 steps



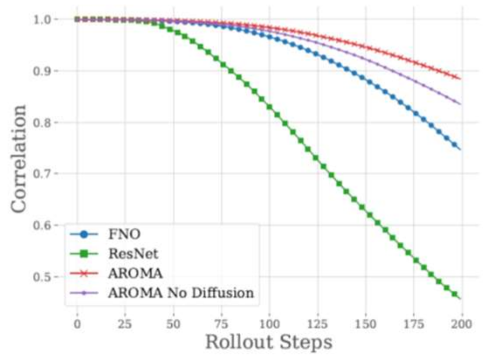
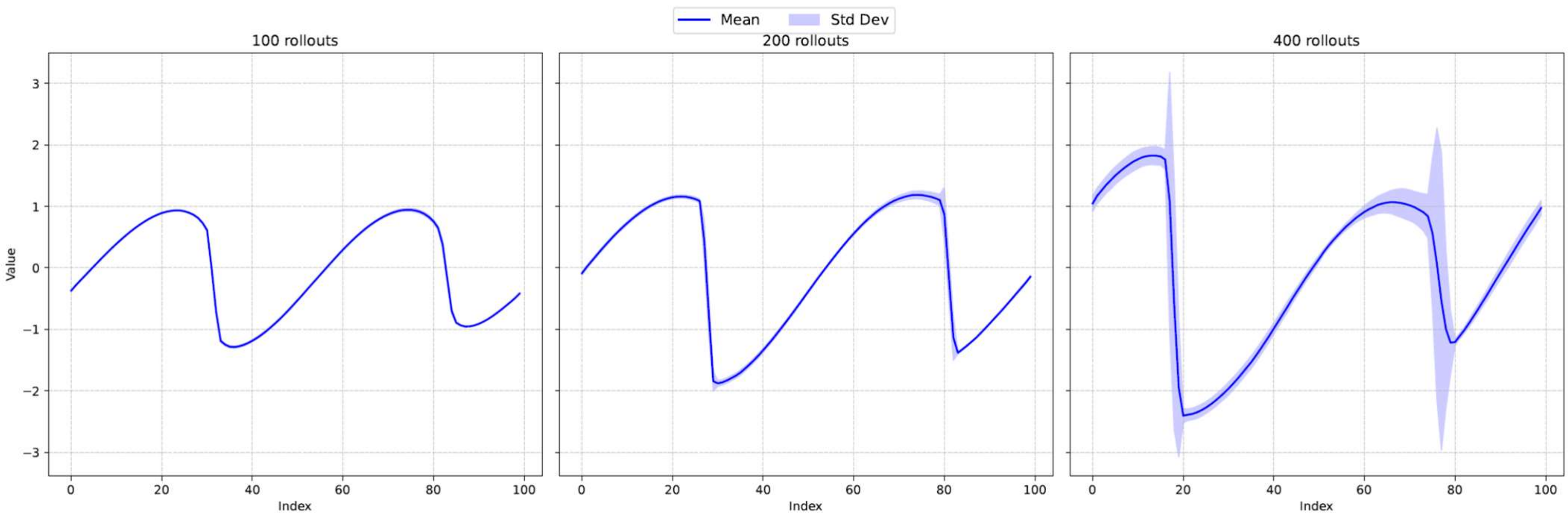


Figure 3: Correlation over time for long roll-outs with different methods on *Burgers*

AROMA: Attentive Reduced Order Model with Attention Uncertainty

- Uncertainty indicator
  - The processor is a diffusion transformer
    - Incorporates stochastic components
  - Example: Burgers equation
  - Mean and variance after 100, 200, 400 rollouts computed on 5 runs



- Thanks for your attention!!

References used in the presentation

- Belbute-Peres, F. de A., Economon, T. D., & Kolter, J. Z. (2020). Combining Differentiable PDE Solvers and Graph Neural Networks for Fluid Flow Prediction. ICML.
- Brandstetter, J., Worrall, D. E., & Welling, M. (2022). Message Passing Neural PDE Solvers. ICLR.
- Dona, J., Dechelle, M., Gallinari, P., & Levy, M. (2022). Constrained Physical-Statistical Models for Dynamical System Identification and Prediction. ICLR.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). an Image Is Worth 16X16 Words: Transformers for Image Recognition At Scale. *ICLR 2021 - 9th International Conference on Learning Representations*.
- Dupont, E., Kim, H., Eslami, S. M. A., Rezende, D., & Rosenbaum, D. (2022). From data to functa: Your data point is a function and you can treat it like one. ICML. http://arxiv.org/abs/2201.12204
- Fathony, R., Sahu, A. K., Willmott, D., & Kolter, J. Z. (2021). Multiplicative Filter Networks. ICLR, 1–10.
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422–440.
- Kashtanova, V., Ayed, I., Arrieula, A., Potse, M., Gallinari, P., & Sermesant, M. (2022). Deep Learning for Model Correction in Cardiac Electrophysiological Imaging. MIDL,
- Kochkov, D., Smith, J. A., Alieva, A., Wang, Q., Brenner, M. P., & Hoyer, S. (2021). Machine learning accelerated computational fluid dynamics. Proceedings of the National Academy of Sciences, 118(21).
- Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2021). Fourier Neural Operator for Parametric Partial Differential Equations. ICLR, 1–16.
- Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, *3*(3), 218–229. https://doi.org/10.1038/s42256-021-00302-5
- Park, J.J, Florence, P, , Straub, J, Newcombe, R, and Lovegrove,, S, "DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1

References used in the presentation

- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
- Serrano, L., Boudec, L. Le, Koupaï, A. K., Wang, T. X., Yin, Y., Vittaut, J.-N., & Gallinari, P. (2023). Operator Learning with Neural Fields: Tackling PDEs on General Geometries. NeurIPS, http://arxiv.org/abs/2306.07266
- Serrano, L., Wang, T., Le Naour, E., Vittaut, J.-N., & Gallinari, P. (2024). *AROMA : Preserving Spatial Structure for Latent PDE Modeling with Local Neural Fields*.
- Sitzmann, V., Martel, J. N. P., Bergman, A. W., Lindell, D. B., Wetzstein, G., & University, S. (2020). Implicit Neural Representations with Periodic Activation Functions. Neurips.
- Stevens R., Taylor V., Nichols J., Maccabe A. B., Yelick K., Brown D., 2020, AI for Science: Report on the Department of Energy (DOE)
- Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., & Ng, R. (2020). Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains. Neurips.
- Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems. ACM Computing Surveys, 1(1), 1–35. https://doi.org/10.1145/3514228
- Yagoubi, M., Danan, D., Leyli-abadi, M., Brunet, J., Gmati, M., Systemx, I. R. T., Schoenauer, M., & Gallinari, P. (2024). ML4CFD Competition : Harnessing Machine Learning for Computational Fluid Dynamics in Airfoil Design. *NeurIPS Competition Track*.
- Yin, Y., Kirchmeyer, M., Franceschi, J.-Y., Rakotomamonjy, A., & Gallinari, P. (2023). Continuous PDE Dynamics Forecasting with Implicit Neural Representations. ICLR, 1–19. http://arxiv.org/abs/2209.14855
- Yin, Y., Le Guen, V., Dona, J., de Bezenac, E., Ayed, I., Thome, N., & Gallinari, P. (2021). Augmenting Physical Models with Deep Networks for Complex Dynamics Forecasting. ICLR.