5D NEURAL SURROGATES FOR NONLINEAR GYROKINETIC SIMULATIONS OF PLASMA TURBULENCE

Gianluca Galletti, Fabian Paischer, Paul Setinek, William Hornsby, Naomi Carey, Lorenzo Zanisi, Stanislas Pamela, Johannes Brandstetter

https://arxiv.org/abs/2502.07469

ICLR 2025 MLMP Workshop

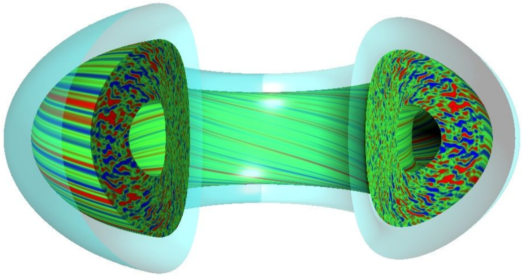
April 27 2025

Nuclear fusion

Produce energy by fusing hydrogen isotopes

We focus on **Tokamaks**, the most established fusion reactor design

- requires magnetic **plasma confinement**
- turbulence arises due to instabilities in plasma [1]
  ↪ leads to particle, energy, and momentum transport [2]
  ↪ **needs to be modeled** for **power plant design** and **control!**



Plasma turbulence visualization [3]

JOHANNES KEPLER UNIVERSITÄT LINZ

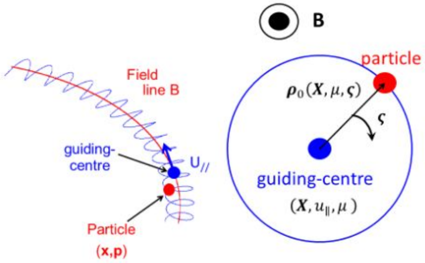
UK Atomic Energy Authority

[1] Introduction to Plasma Physics and Controlled Fusion, Francis F. Chen, 2016
[2] JINTRAC: A system of codes for integrated simulation of tokamak scenarios, Romanelli et al., Plasma Fusion Research, 2014
[3] https://w3.pppl.gov/~hammett/viz/viz.html

Plasma turbulence modelling

<u>Turbulence is a key driver of plasma confinement degradation</u>

[4]



- Described via nonlinear **<u>Gyrokinetics</u> equations** [1]
  - evolve **5D distribution function** over time
  - solving these equations numerically is **very expensive**
    ↪ hours to days for a single simulation on HPC!

- Development of fusion control requires knowledge of transport in “saturated state”
  - saturation is caused by nonlinear phenomena, e.g. **“zonal flows”** [2]
  - **Current reduced-order methods** (quasilinear) **don’t capture zonal flow** [3]

JOHANNES KEPLER UNIVERSITÄT LINZ

UK Atomic Energy Authority

[1] The non-linear gyro-kinetic flux tube code gkw., Peeters et al., Computer Physics Communications, 2009
[2] Physics of zonal flows, Itoh et al., Physics of Plasmas, 2006
[3] Validity of Quasi-Linear Transport Model, Bourdelle et al., Nuclear Fusion, Volume 6, 2021
[4] Garbet, Xavier and Lesur, Maxime. Gyrokinetics. HAL. 2023

Motivation

Quasilinear approximations [1,2]

- rely on saturation rules derived from linear simulations
  ↪ to date there is **no general quasilinear model** [3]
- **neglect nonlinear phenomena**

What about machine learning?

- Current neural surrogates [4] mimic quasilinear models → **no zonal flow**
- Can we learn GK surrogates that **directly evolve the 5D distribution function?**

JOHANNES KEPLER UNIVERSITÄT LINZ

UK Atomic Energy Authority

[1] A new gyrokinetic quasilinear transport model applied to particle transport in tokamak plasmas, Bourdelle et al., Physics of Plasmas, 2007
[2] A theory-based transport model with comprehensive physics, Staebler et al., Physics of Plasmas, 2007
[3] A quasi-linear model of electromagnetic turbulent transport and its application to flux-driven transport predictions for STEP, Giacomin et al. Journal of Plasma Physics, 2024
[4] Fast modeling of turbulent transport in fusion plasmas using neural networks, van de Plassche et al., Data-Driven Plasma Science, 2020

Contributions

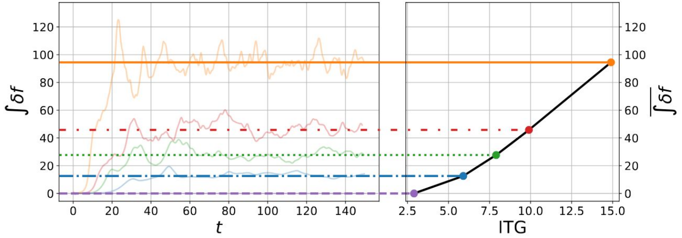
- **Data generation and inspection**
- **5D neural surrogate model**
  - Based on hierarchical vision transformers
- **Evaluation of physical quantities**
  - Electrostatic potentials
  - Heat flux

JOHANNES KEPLER UNIVERSITÄT LINZ
UK Atomic Energy Authority

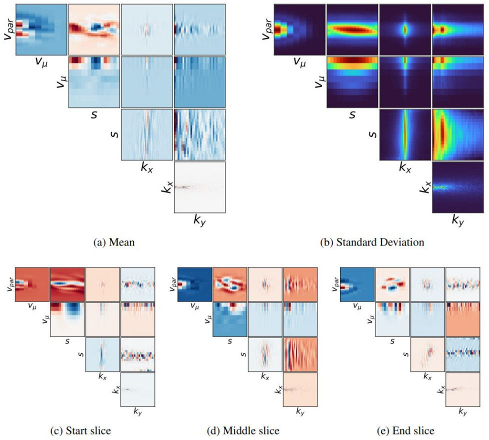
Contributions

- **Data generation and inspection**

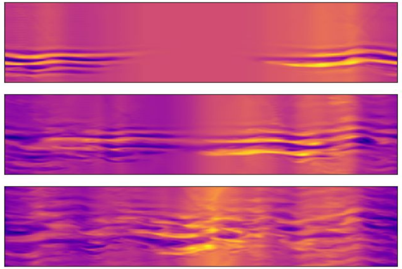
Numerical code **GKW** [1]
Sampled ion temperature gradient (ITG)



Heat fluxes for our dataset



5D density function



3D electrostatic potential

JOHANNES KEPLER UNIVERSITÄT LINZ

UK Atomic Energy Authority

[1] Peeters, Arthur et al. The nonlinear gyro-kinetic flux tube code GKW, Computer Physics Communications, 180, 2009

Contributions

- **Data generation and inspection**

- **5D neural surrogate model**
  - Based on hierarchical vision transformers

Problem: How can we build neural surrogates that can process 5D data?

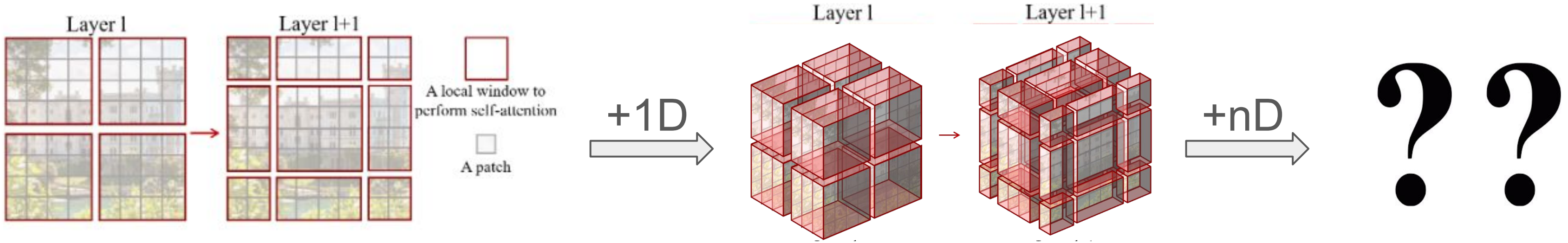
JOHANNES KEPLER
UNIVERSITÄT LINZ

UK Atomic
Energy
Authority

Contributions - neural surrogate

How can we build neural surrogates that can process 5D data?

- CNNs [1] ?
  ↪ Impractical due to runtime and memory constraints
- (Vision) transformers [2]?
  ↪ 5D results in extremely long sequences, self-attention scales quadratically...
  ↪ **Swin** [3,4]: **performs attention in parallel in local windows**

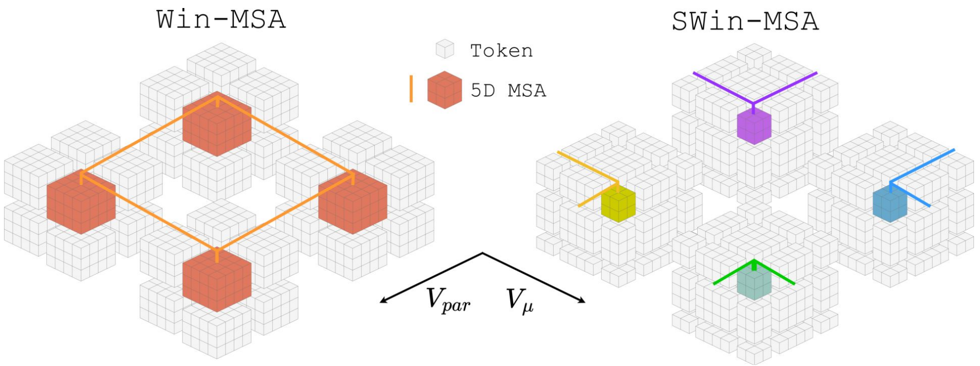


JOHANNES KEPLER UNIVERSITÄT LINZ
UK Atomic Energy Authority

[1] Fukushima, Kunihiko. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. Biological Cybernetics 1980
[2] Dosovitskiy, Alexey, et al. "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". ICLR 2020
[3] Liu, Ze, et al. "Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows". ICCV 2021
[4] Liu, Ze, et al. "Video Swin Transformer". CVPR 2022

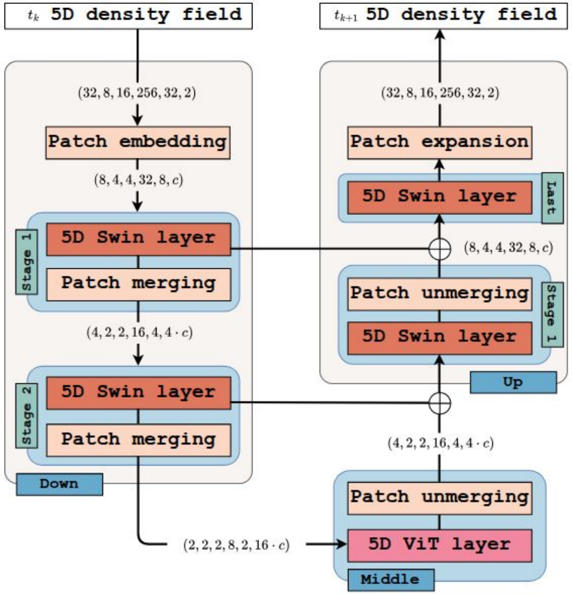
Contributions - 5D Hierarchical Swin Transformer

5D swin attention → generalized nD swin



Hierarchical **UNet** structure [1]

- Downsampling (Patch merging)
- Upsampling (Patch expansion)



JOHANNES KEPLER UNIVERSITÄT LINZ

UK Atomic Energy Authority

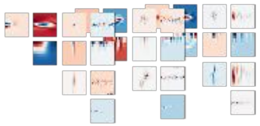
[1] U-Net: Convolutional Networks for Biomedical Image Segmentation, Ronneberger et al., MICCAI 2015

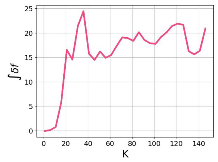
Contributions

- **Data generation and inspection**
- **5D neural surrogate model**
    - Based on hierarchical vision transformers
- **Evaluation of physical quantities**
    - Electrostatic potentials
    - **Heat flux**

Physical quantities:
electrostatic potential, **heat flux**

Visual inspection





JOHANNES KEPLER UNIVERSITÄT LINZ

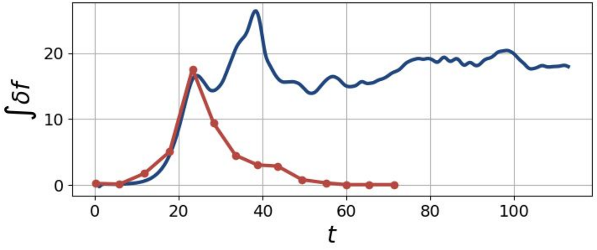
UK Atomic Energy Authority

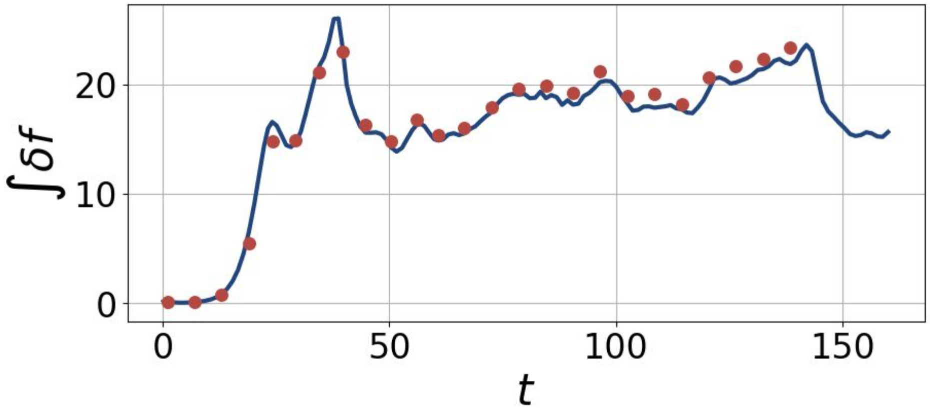
Evaluation - heat flux time trace

- Heat flux time trace for single-step prediction is decent ✅
- (At least) **two orders of magnitude speedup vs GKW** ✅
- Generalization to unseen ion temperature gradient ✅

... BUT

↪ **autoregressive** ❌



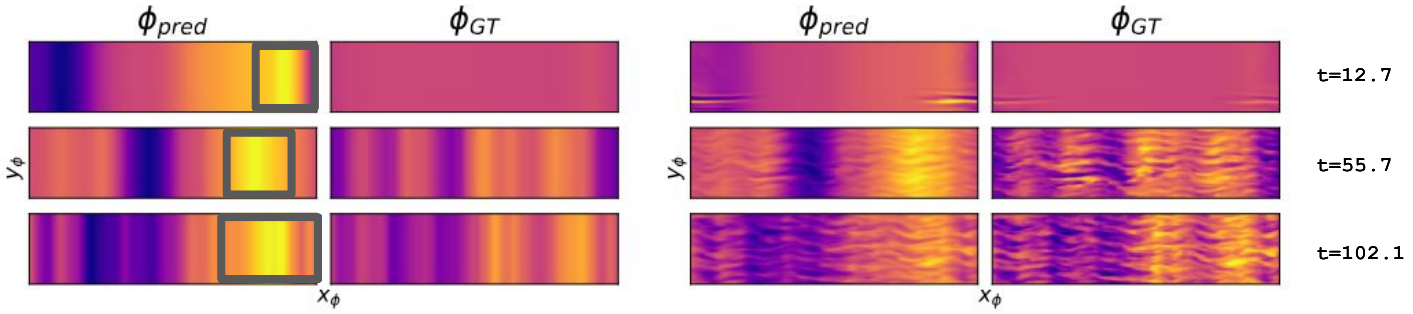


JOHANNES KEPLER UNIVERSITÄT LINZ

UK Atomic Energy Authority

Evaluation - electrostatic potentials

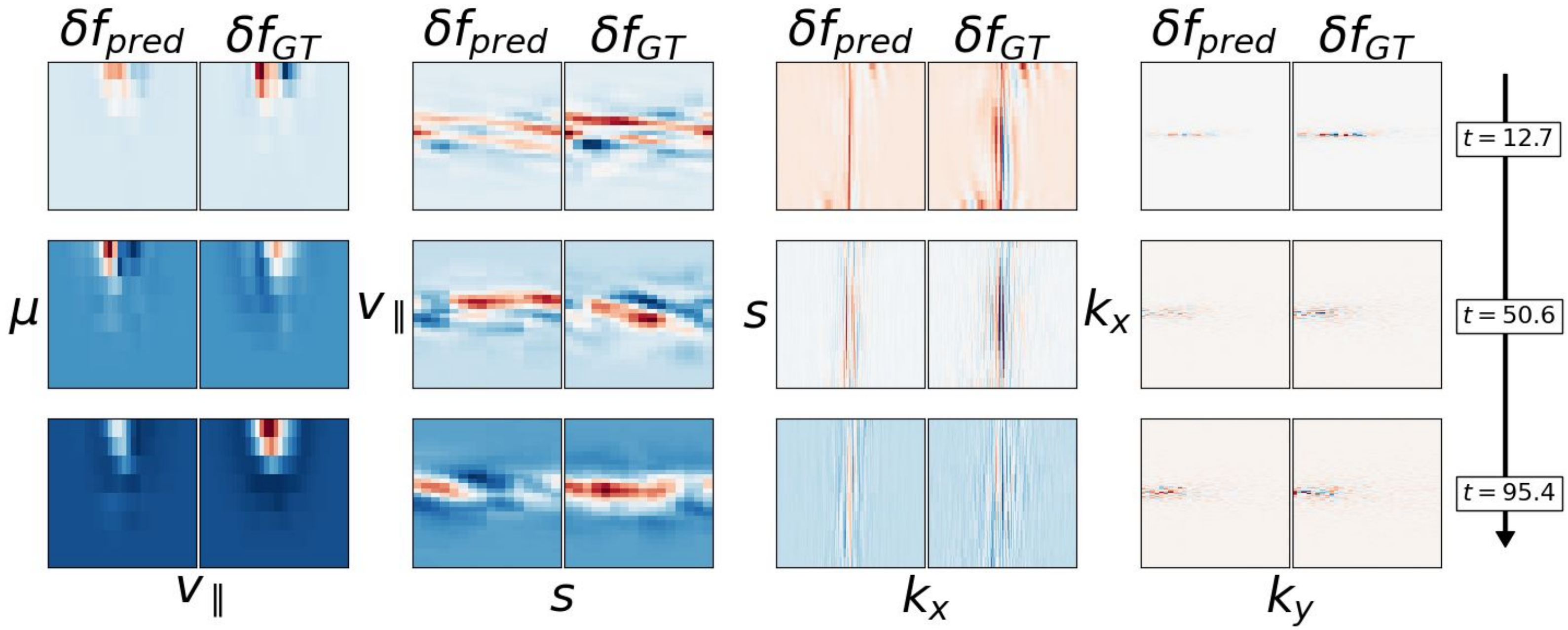
- Y wave vector is well reproduced ✅
  ↪ explains good alignment of predicted heat flux
- **Zonal flow is overestimated** ⚠️
  ↪ dampens turbulence during autoregressive rollout



(a) Zonal mode of the electrostatic potential.
(b) Electrostatic potential with zonal mode.

JOHANNES KEPLER UNIVERSITÄT LINZ
UK Atomic Energy Authority

Evaluation - visual inspection



UK Atomic Energy Authority

JOHANNES KEPLER UNIVERSITÄT LINZ

Conclusions and future work

First 5D neural surrogate for nonlinear gyrokinetics

- 5D density function rollouts, one-step heat flux
- ~two orders of magnitude faster than GKW
  ↪ increases for higher fidelity simulations

Open challenges

- **data storage** ⚠️
- physical grounding
- verification

Future directions

- Improve zonal flow modeling → remedy heat flux decay (autoregressive)
- Move from low fidelity to high fidelity

JOHANNES KEPLER UNIVERSITÄT LINZ

UK Atomic Energy Authority

Thank you for listening!

JOHANNES KEPLER UNIVERSITÄT LINZ

UK Atomic Energy Authority

Gianluca Fabian Paul Johannes

William Naomi Lorenzo Stan

2502.07469