5D Neural Surrogates for Nonlinear Gyrokinetic Simulations of Plasma Turbulence

Gianluca Galletti[1], Fabian Paischer[1], Paul Setinek[1], William Hornsby[2], Lorenzo Zanisi[2], Naomi Carey[2], Stanislas Pamela[2], Johannes Brandstetter[1,3]

[1] ELLIS Unit, LIT AI Lab, Institute for Machine Learning, JKU Linz, Austria
[2] UKAEA (United Kingdom Atomic Energy Authority), Culham Campus, Abingdon
[3] Emmi AI GmbH, Linz, Austria

TL;DR

Turbulence impairs plasma confinement in nuclear fusion.

We introduce the first **generalized nD swin transformer** and apply it to the **5D distribution function** of nonlinear gyrokinetics.

Our method achieves a **speedup of two orders of magnitude** compared to GKW, while capturing underlying physical quantities.

Nuclear Fusion

- Produce energy by fusing hydrogen isotopes in a plasma
- **Problem:** Turbulence arises due to instabilities in plasma
  ↪ results in particle, heat, and momentum transfer
  ↪ essential for control and power plant design

Plasma Turbulence

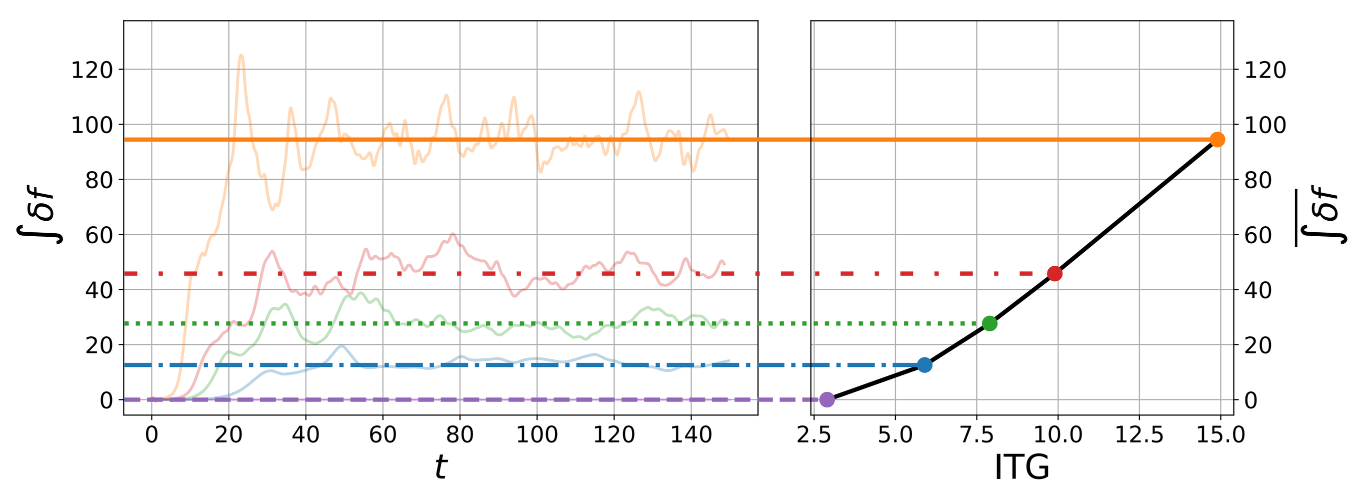
Nonlinear Gyrokinetics is used to model plasma turbulence.

$$\frac{\partial f}{\partial t} + (v_\parallel \mathbf{b} + \mathbf{v}_D) \cdot \nabla f + \mathbf{v}_\chi \cdot \nabla f - \frac{\mu B \mathbf{B} \cdot \nabla B}{m} \frac{\partial f}{B^2} \frac{\partial f}{\partial v_\parallel} = S$$

Turbulent transport is obtained by integration over δf.
↪ But, evolving δf over time is **computationally expensive**.
↪ Can we develop neural surrogates to model δf?

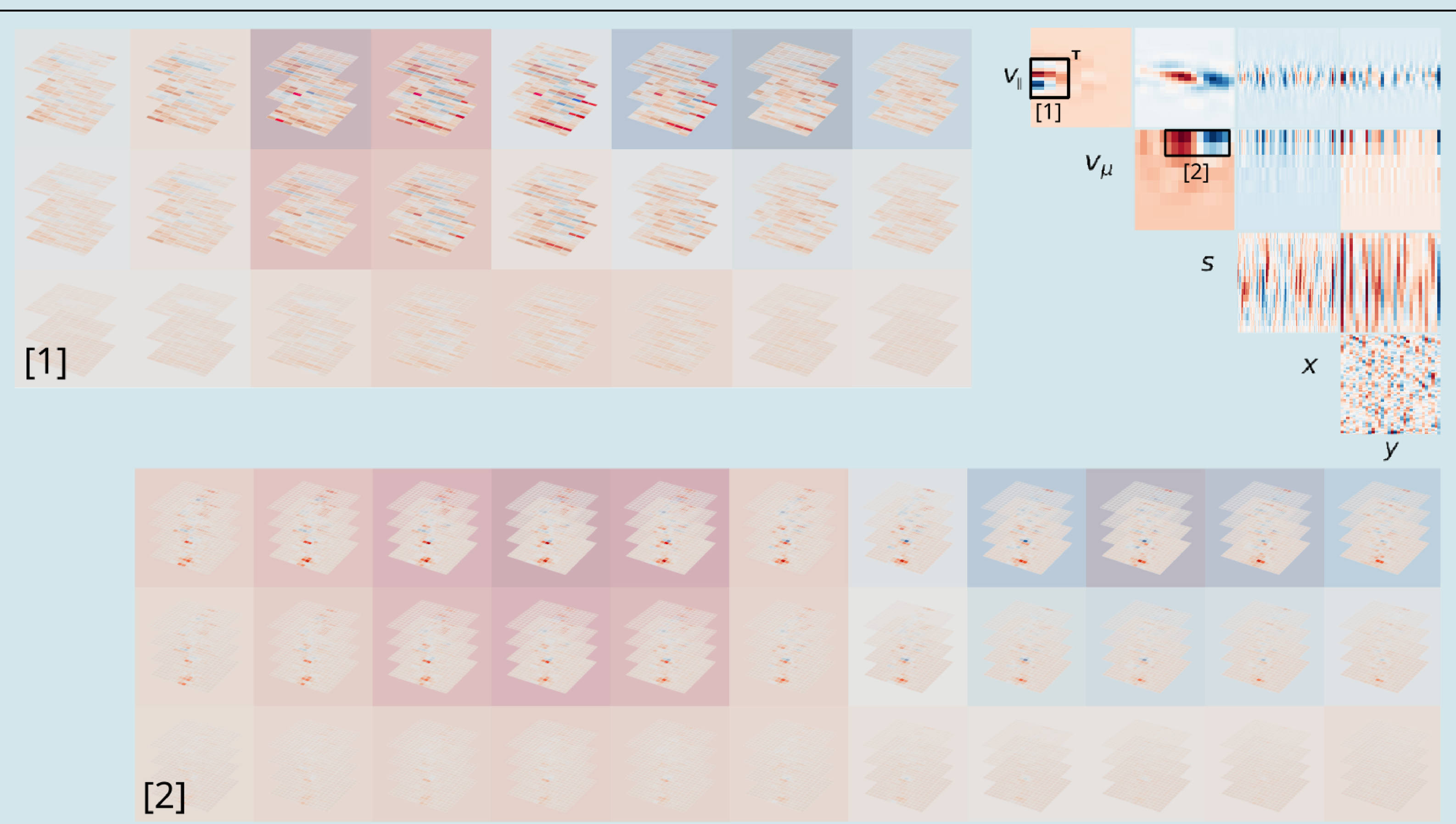
Data generation

We leverage GKW and vary Ion Temperature Gradient (ITG) to collect data for five simulations for the adiabatic electron approximation.



Flux trace over time (left) and averaged (right) for selected ITG values.

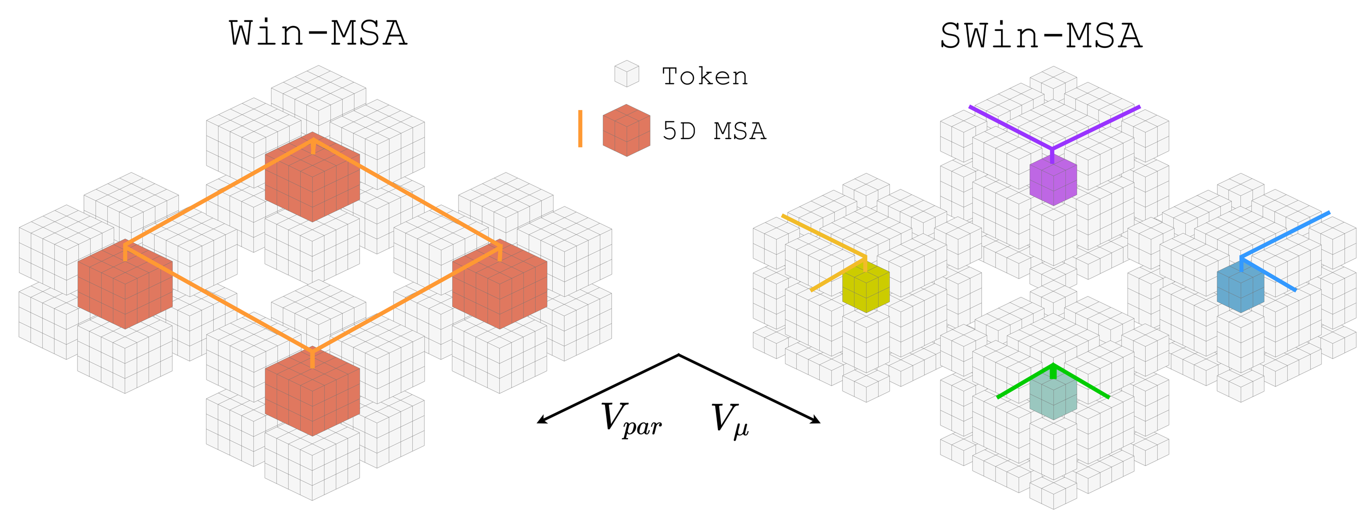
Visualizing 5D information



Real-space 5D δf displayed as 2D planes ([1]: $(v_\mu, v_\parallel)$, [2]: (mu, s)), where each tile contains a 3D cube of the remaining dimensions ([1]: (s, x, y), [2]: $(v_\parallel, x, y)$). Upper triangular plot is a reduced visualization with $\binom{5}{2}$ combinations of axes.

5D Vision Transformers

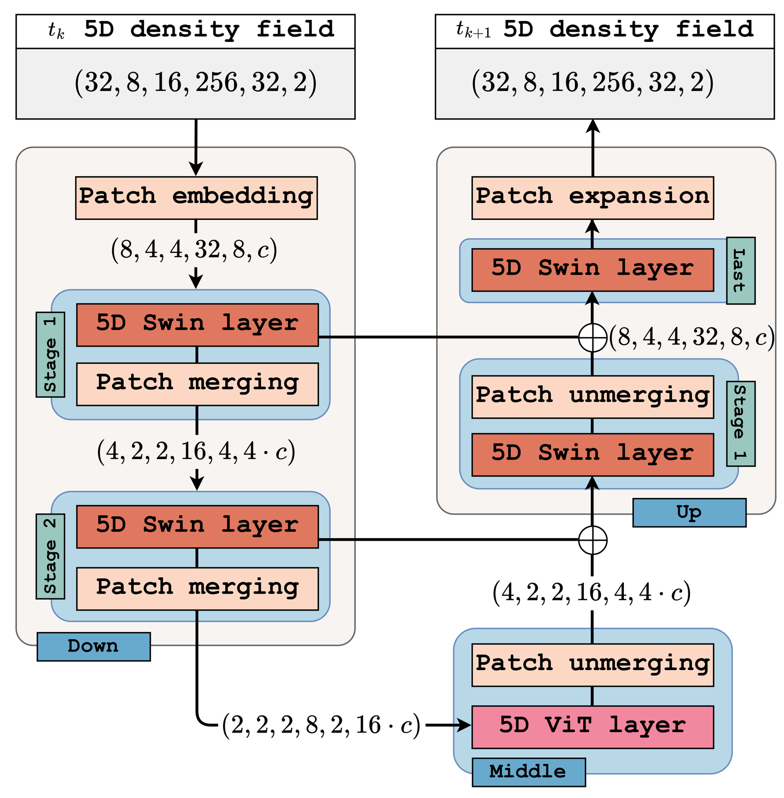
nD shifted-window attention



5-dimensional swin attention ilustrated as a 2D plane of 3D windowed tokens.

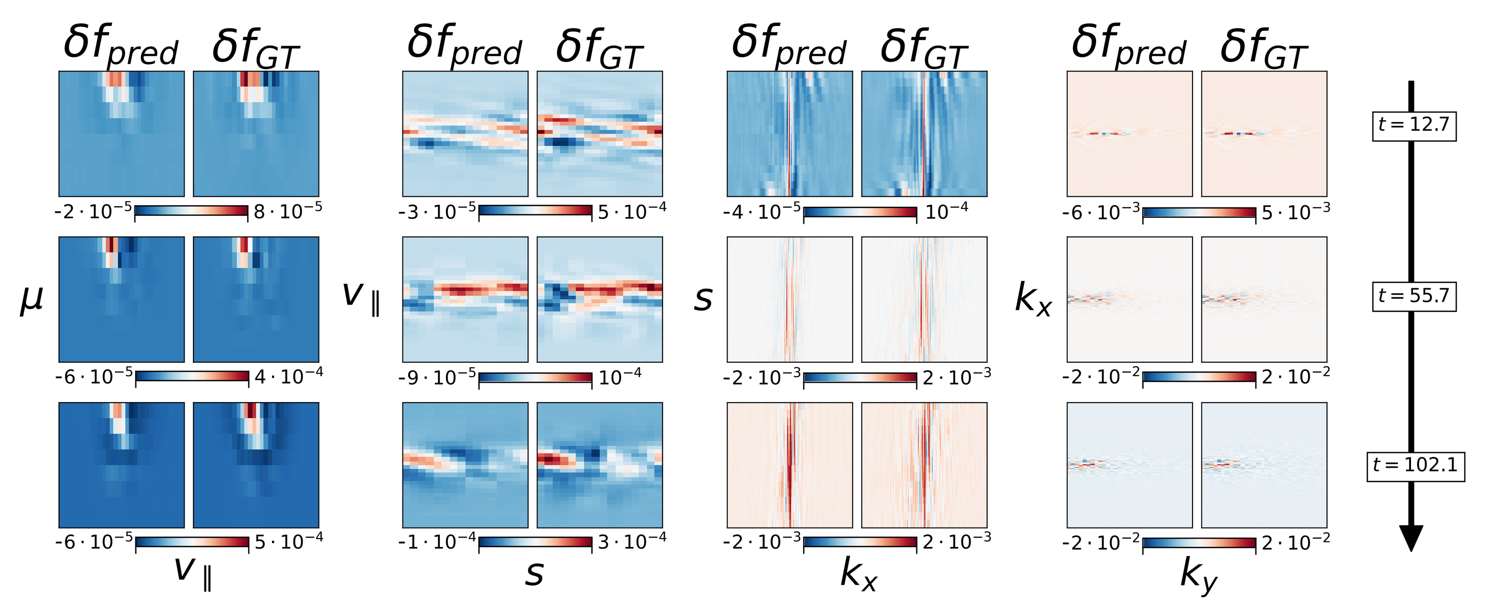
5D Swin-UNet

- **5D CNNs** → **unfeasible!**
- **5D** = **very long sequences**:
  ↪ global ViT is expensive!
  ↪ **local attention** (swin)
- **Linear** in sequence length!
- General **nD window attention**
- **Hierarchical** UNet structure



Results

5D distribution function is well reproduced!



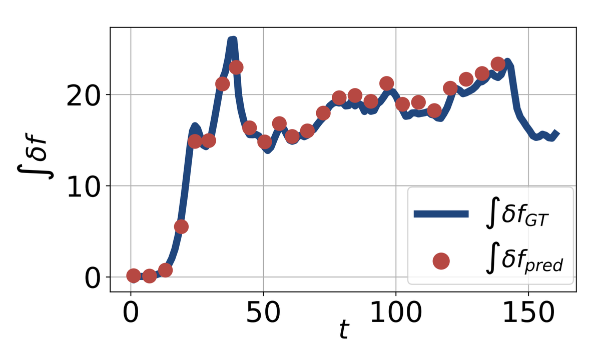
One-step model predictions δf_{pred} versus ground truth δf_{GT} over time.

Physics Verification

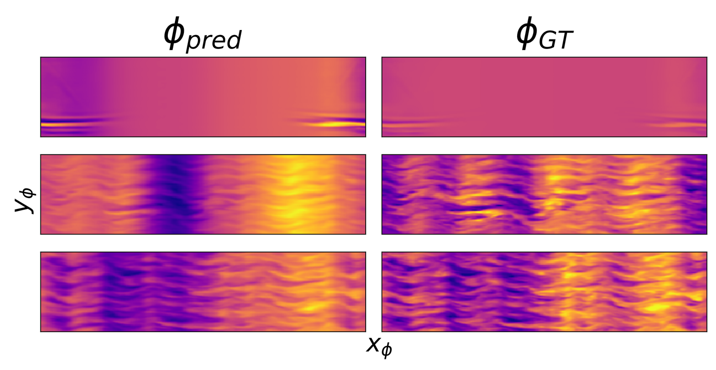
Apply GKW to the predicted δf to obtain

- **Heat flux:** Key quantity for reactor design
- **Electrostatic potential** ϕ: shows nonlinear phenomena

Heat Flux Time Trace



Electrostatic Potential ϕ



Takeaways

- First 5D neural surrogate for nonlinear gyrokinetics
- Two orders of magnitude faster than numerical solver (GKW)
- Reproduces physical quantities (heat flux and potentials)

galletti@ml.jku.at