9th Asia-Pacific Conference on Plasma Physics, 21-26 Sep, 2025 at Fukuoka

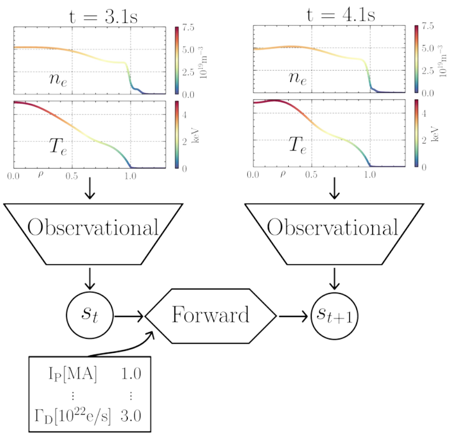
On Physics-Data Generative Modeling for Core-Edge Integration in Tokamaks

A. Kit[1], A.E Järvinen[1], V. Menkovski[3], A. Bruncona[1], S. Wiesen[2]
[1] VTT Technical Research Centre of Finland, [2] DIFFER Dutch Institute for Fundamental Energy Research, [3] Eindhoven University of Technology
e-mail (speaker): adam.kit@vtt.fi

Deep learning based models have become power tools for modeling the time evolution of tokamak plasmas, both in terms of surrogating/emulating physics models[1] as well as learning from experimental data [2]. Some deep learning models, while being fast and potentially very accurate, have the disadvantage of being black-boxes that have low interpretability and struggle to generalise beyond the training data [3]. Generative models, a subset of deep learning models, based on probabilistic variational autoencoders (VAEs) [4], seek to learn a representation of the underlying system, here the dynamic evolution of a tokamak plasma. The benefit of learning a representation via a probablistic generative model is that the model yields an interpretable representation that, ideally, encodes physically meaningful information (Figure 1), and provides uncertainty intervals over the predictions.

To this end, VAEs have been trained on experimental data to predict kinetic profiles of the JET and AUG tokamaks. By introducing semi-supervision to the latent variables models, an the learned latent representation is

Figure 1: Graphical representation of the generative model. Electron profiles are encoded to a representation s via a convolution neural network. The representation s is pushed forward in time via a multi-layered perceptron.



The latent representation can be decoded into profiles and other observables via another convolution neural network.

established to additionally map machine parameters to profiles. The models have been investigated for uses in the prediction of steady-state edge conditions [5], machine size scalings [6], and time evolution of electron kinetic profiles [7] (Figure 2).

Ongoing work is focused on exploring including physics simulations into the representation, alongside the experimental data. The physics simulations are focused around core transport (TGLF) and ideal-MHD stability (MISHKA) and reduced pedestal transport models.

References

[1] S. Wiesen et al., *Nucl. Fusion* **64** 086046 (2024)
[2] Y. Poels et al., (under review), preprint: arxiv:2502.17397 (2025)
[3] J. Abbate et al., *Nucl. Fusion* **61** 046027 (2021)
[4] D. Kingma et al., ICLR 8-9 arxiv:1312.6114 (2014)
[5] A. Kit et al., NME **34** 101347 (2023)
[6] A.E. Järvinen et al., PoP **31** 032508 (2023)
[7] A. Kit et al., PoP **31** 032504 (2024)

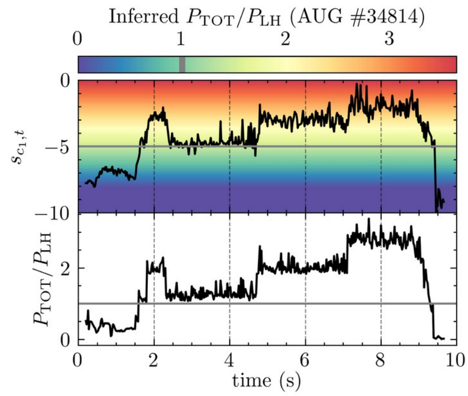


Figure 2: The time traces of (botom) the ratio of total power injected to the Martin L-H scaling for AUG #34814 and (top) the learned representation encoded by the latent variable model. The horizontal line (vertical on colorbar) marks unity for PTOT/PLH, thus an inferred onset of H-mode (Martin scaling). Although the network has been trained for profile prediction, the advantage of the additional representation learning objectives leads to an interpretable model.