Physics-Informed Deep Learning for Plasma Turbulence Predictions in Fusion Reactors (DEEPlasma)

DEEPlasma consortium: Simo Särkkä (Aalto) and Aaro Järvinen (VTT)

1 Aim and objectives

1.1 Significance of research project in relation to current knowledge, research based starting points:

Taming of the plasma edge turbulence is of central importance in integrating high energy gain with reactor-relevant duty cycle in future fusion power plants. Due to the extreme computational complexity of plasma turbulence simulations, power plant design has traditionally relied on significant model reductions that mimic the role of turbulence, which substantially increases prediction uncertainties and investment risks. In this project, DEEPlasma, we will leverage the recent advances of high-performance computing (HPC), and in particular, the LUMI supercomputer and physics-informed machine learning, to accelerate high-fidelity edge plasma turbulence predictions to the level required by power plant design and real-time control applications. The project is done in close collaboration with two professors from Texas, USA.

Predicting sustainable fusion power plant operation

Economical and societal development in the 21st century depends on the availability of carbon-free energy sources. Nuclear fusion, the energy source of stars, offers one of the greatest opportunities to solve this existential need. Fusion energy research will soon enter the era of reactor-scale devices as ITER is approaching its first operational campaigns [1]. These reactor-scale facilities are facing a significant core-edge integration challenge. For sustainable energy production, they must confine a high temperature, $\sim 10^8$ K, deuterium-tritium plasma in a strong magnetic field without overheating the reactor components. The unmitigated wall heat loads in large demonstration (DEMO) reactors are predicted to exceed the values on the surface of the Sun, ~ 60 MW/m^2, and hence the thermo-mechanical limits of any available materials [2]. A conventional approach to reduce wall heat loads is to inject radiating impurities into the vessel to convert the plasma heat to radiation, which is distributed more uniformly than conducted heat fluxes [3, 4, 5]. However, excess plasma cooling is detrimental to the fusion performance. The power plant design activities must find an appropriate and controllable plasma scenario that meets these competing demands. However, due to the multiple interacting physical processes and spatio-temporal scales, predicting core-edge integration with high-fidelity physics simulations is one of the most demanding computational and scientific challenges of magnetic confinement fusion energy research.

Plasma edge turbulence simulations

The standard operational scenarios projected to provide reactor-relevant performance require a self-organized transport barrier and formation of a pressure pedestal at the edge. One of the most demanding components of core-edge integration predictions is related to simulations of the residual plasma turbulence that regulates heat and particle transport within this transport barrier [6, 7, 8, 9, 10, 11, and refences therein]. These transport fluxes can be simulated by solving a nonlinear, saturated turbulent state of the gyrokinetic (GK) Vlasov-Maxwell system of equations, which is the most advanced theory for turbulence in strongly magnetized plasmas. GK applies the scale-separation between the fast gyro-motion of charged particles in the magnetic field, $\sim 10^{-9} - 10^{-7}$ s, and the slow turbulence evolution, 10^{-5} s, to reduce the fully kinetic Vlasov-Maxwell system of equations by averaging through the fast dynamics. Despite this reduction, a single nonlinear GK calculations takes of the order of $10^4 - 10^6$ CPU hours. Clearly such computational demands are incompatible with the needs of power plant design or real-time control applications, which require thousands of these flux calculations for macroscopic time-evolution, $\sim$ 1 s, of the plasma scenario. Due to this computational challenge, core-edge integration predictions in reactor design activities typically apply reduced models that bypass the need to solve the GK equations. An example of such approach is the ballooning critical pedestal technique in the EPED model [12]. However, such model reductions propagate the resulting prediction uncertainty as a significant investment risk in the large and expensive reactor-scale facilities. Fortunately, as physics-informed machine learning algorithms and HPC systems, such as LUMI, have advanced, a new paradigm is emerging where it is possible to accelerate the plasma edge turbulence simulations for sufficient throughput for power plant design or real-time control applications.

Accelerating plasma edge turbulence simulations with physics-informed machine learning

Physics-informed machine learning is an emerging area that promises to have a profound and lasting impact in science and engineering, providing many opportunities for high-impact, transformative research over the next decade [13]. This area comprises an entirely novel set of tools and frameworks that are only beginning to be developed and understood by the scientific community [14, 15]. Unlike traditional numerical methods, physics-informed neural networks (PINNs) and physics-informed Gaussian processes (PIGPs) do not require elaborate grids, and unlike traditional neural networks and Gaussian processes, they are able to extrapolate to regions where there is no data. These algorithms take advantage of the enormous progress achieved recently in high-performance algorithms and computational infrastructure for training deep machine learning models.

Typical data-driven machine learning methodologies do not incorporate physical understanding of the problem domain. Furthermore, in many scientific domains, high-fidelity data are expensive or time-consuming to obtain. Physics-informed machine learning addresses both of these problems by introducing regularizing constraints obtained from physical laws, allowing prediction of future performance of complex multiscale, multiphysics systems, using sparse, noisy, and heterogeneous data. Unlike traditional black-box machine learning methods, physics-informed machine learning delivers interpretable models, leading to improved verification and validation. PINNs and PIGPs have already been deployed in a diverse set of problems with remarkable success [16, 17, 18, 19, 20, 21, 22, 23, 24, 25].

This project, DEEPlasma, investigates methods to apply physics-informed machine learning to establish fast, high-fidelity, surrogate representations for local, linear GK simulations of pedestal turbulence conducted with the GENE code [26]. From now on, this surrogate model is called PEGENENN (Pedestal GENE Neural Network). The specific objectives of the project are:

- **O1:** Demonstrate that the modern HPC environments, such as LUMI, provide the necessary computational power to enable machine learning acceleration of the highly demanding pedestal GK simulations.
- **O2:** Develop suitable physics-informed constraints for pedestal GK to limit the need to oversample the training domain with the full computationally costly physics model.
- **O3:** Demonstrate that by applying the physics-informed approach as well as Bayesian methods to populate the training domain with optimal information gain for each full simulation, order of magnitude reduction in the necessary density of training data is achieved.
- **O4:** Building on top of the success in achieving the above objectives, leverage the leading-edge computational capabilities of LUMI to generate a production level PEGENENN.

1.2 Research questions and/or hypotheses:

The hypotheses of the research are the following:

- **H1:** The new LUMI HPC provides sufficient computational throughput for generating local, linear pedestal GK stability databases with GENE for establishing PEGENENN.
- **H2:** Physics-informed machine learning algorithms enable learning accurate surrogate models with orders of magnitude sparser training data than traditional black-box machine learning methods.
- **H3:** By applying Bayesian, acquisition function based, methods for building the training data, the necessary training data density for sufficiently accurate surrogate models can also be significantly reduced.
- **H4:** Building on top of the combined methods and infrastructure established in addressing **H1-3** as well as on the capabilities of LUMI, it is possible to develop a production level PEGENENN.

1.3 Expected research results and their anticipated scientific impact, potential for scientific breakthroughs and for promoting scientific renewal:

This project is expected to break a new ground in core-edge integration simulations in magnetic confinement fusion facilities. While fast models have been developed for core plasma transport through surrogate model accelerations of reduced turbulence models, such as QLKNN and TGLF-NN [27, 28, 29], machine learning acceleration for pedestal turbulent transport has not been tried before. This is presumably due to the complexity of pedestal plasmas requiring more complete GK treatment, including electromagnetic effects, leading to increased computational requirements in creating the training database for surrogate model development.

DEEPlasma aims to address this challenge by combining the leading-edge computational capabilities of LUMI with physics-informed and Bayesian approaches to reduce the necessary density of training data for accurate surrogate model development. The physics-informed constraints will improve the extrapolation and interpolation performance of the model and the Bayesian methods aim to selectively simulate the training cases that are expected to yield optimal information gain in terms of surrogate model accuracy. PEGENENN would provide fast local, linear stability predictions, which when coupled with appropriate mapping from the linear stability to transport would yield a very fast, GK-based transport model for pedestal plasmas. This mapping from linear GK to transport is the focus of complementary research of our international partners in US. When coupling the two components, (1) PEGENENN from DEEPlasma with (2) the linear to transport mapping from the research in US, the resulting GK-based pedestal transport model can yield a new paradigm in core-edge integrated scenario simulations. Eventually, when coupled with fast models for global pedestal magnetohydrodynamic (MHD) stability and scrape-off layer (SOL) transport, which are being addressed through other research activities within EUROfusion, a fast, high-fidelity, physics-based model for the entire edge from the pedestal top to the reactor wall will become possible. Such a model would represent a scientific breakthrough for power exhaust and core-edge integration predictions in these fusion devices.

The novel physics-informed machine learning algorithms developed in this project more generally facilitate the application of data-driven approaches to solve complex engineering and science problems of societal importance. The developed methods also have a wide range of applications outside of the nuclear fusion field.

1.4 Special objective of call:

The special objectives of the call are addressed as follows:

- **Utilising EuroHPC or the LUMI supercomputer in research**: We aim to use the LUMI supercomputer for both running the GENE simulations and machine learning methods. As this project is academic research conducted within Finnish higher-education and research institutions, the LUMI resources are available to the project free of charge (for more details, please see Section 2.2).

- **International collaboration with research teams in the aforementioned countries/regions (US states of Texas and/or Colorado and/or research teams in Japan):** The project involves close collaboration with Prof. David Hatch from University of Texas and with Prof. Ulisses Braga-Neto from Texas A&M University, along with their research teams. The support letters are included in the application.
- **Scientific research in one or several themes represented by the Finnish Flagships:** The physics-informed machine learning theme leveraged in the project falls within the scope of Finnish Center for Artificial Intelligence (FCAI) flagship. The main PI Simo Särkkä is also member of Steering Group of FCAI and leads the AI Across Fields (AI for X) program of the flagship. In addition to the AI for X program of FCAI, the current project is closely related to the scientific goals of the FCAI research programs "Autonomous AI (R6)", "Next-generation data-efficient deep learning (R3)", and "Agile probabilistic AI (R1)".

2 Implementation

2.1 Work plan and schedule:

The list of work packages (WPs) and tasks (T) is the following:

- **WP1:** Proof-of-principle database generation for training PEGENENN:
  - **T1.1:** Create software framework to execute large batches of GENE simulations on LUMI.
  - **T1.2:** Create software framework to automatically process the GENE results into a database format suitable for the machine learning model training algorithms.
  - **T1.3:** Determine the appropriate input parameter training domain for the proof-of-principle PEGENENN.
  - **T1.4:** Establish a proof-of-principle database for training PEGENENN using the LUMI Regular Access mode resources.
- **WP2:** Physics-informed machine learning for accurate surrogate model development with sparse training data:
  - **T2.1:** Train a selection of traditional black-box machine learning models on the proof-of-principle database established in WP1 to provide a surrogate model performance baseline.
  - **T2.2:** Investigate appropriate physics-informed constraints suitable for PEGENENN.
  - **T2.3:** Train physics-informed machine learning models on the proof-of-principle database.
  - **T2.4:** Compare the surrogate model performance between the traditional black-box and physics-informed approaches.
- **WP3:** Methods to quantify uncertainties and improve surrogate model accuracy through Bayesian approaches:
  - **T3.1:** Investigate probabilistic methods to quantify the uncertainties of the established surrogate models.
  - **T3.2:** Apply these methods to establish uncertainty bounds on the PEGENENN predictions.
  - **T3.3:** Investigate active learning and Bayesian optimization to optimize training data generation.
  - **T3.4:** Develop active learning and Bayesian optimization methods for large scale PEGENENN predictions.

- **WP4:** Scaling PEGENENN from proof-of-principle implementation to a production model with LUMI Extreme Scale Access:
  - **T4.1:** Apply for the LUMI Benchmark Access mode.
  - **T4.2:** Complete the LUMI Benchmark Access project to demonstrate that the approach is ready for LUMI Extreme Scale Access.
  - **T4.3:** Prepare the proposal and apply for the LUMI Extreme Scale Access.
  - **T4.4:** Use the LUMI Extreme Scale Access resources to generate a production level PEGENENN by applying the methodology developed in WP1, WP2, and WP3.

The PERT and Gantt charts for the WPs and Tasks are shown below.

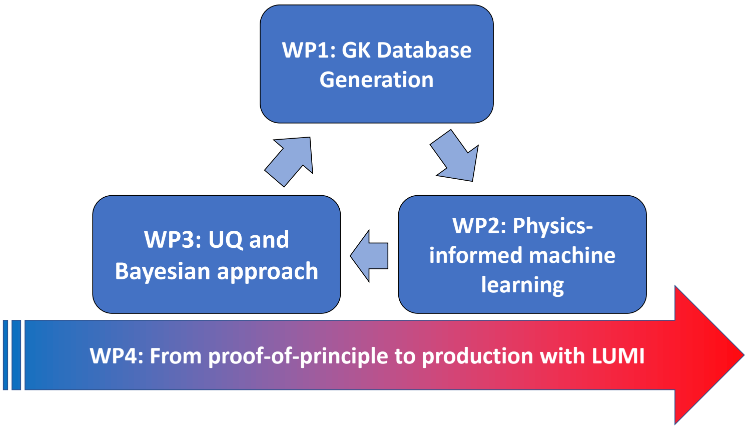

Figure 1: PERT chart showing the WP interactions.

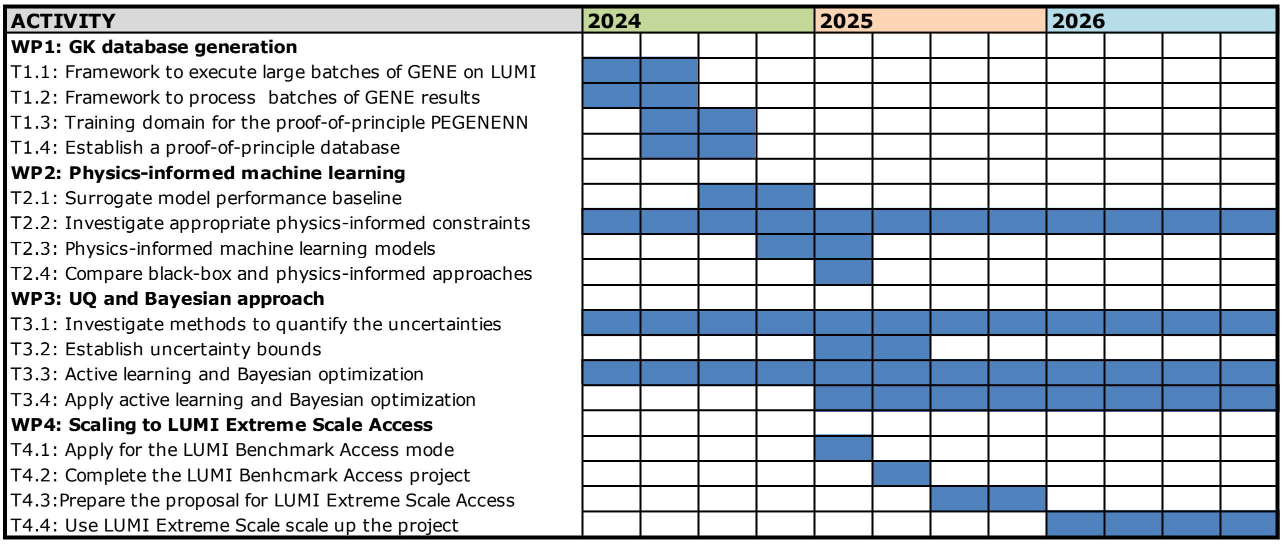

Figure 2: Gantt chart showing the schedule of the project.

2.2 Research data and material, methods, and research environment:

The details of the data and methodology used in the project is given below.

WP1: Proof-of-principle database generation for training PEGENENN

One of the leading challenges in generating surrogate models for GK plasma turbulence simulations is related to the computational cost of generating a sufficiently large training database for the model. A fully nonlinear GK simulation for a single radial location can be as expensive as 10^4 – 10^6 CPU hours. Clearly generating a high-density training database for a multidimensional domain with these computational costs would be an enormous computational task, approaching potentially a billion CPU hours. Furthermore, nonlinear simulations can be susceptible to convergence challenges and numerical difficulties rendering large batch execution of cases very challenging.

Linear GK, on the other hand, provides a stable predictor of stability thresholds, turbulence drives, growth rates, mode frequencies, and eigenfunctions with a reasonable computational cost of the order of 10^3 CPU hours. With the recent progress towards GPU execution, this is expected to be even faster. With these computational costs, it is feasible to generate a proof-of-principle training database of local, linear GK simulations for pedestal parameters with the computational time available through LUMI Regular Access mode (details are given later in this section).

The local GK approach takes the relevant parameters and gradients at a single radial position and does not consider any other radial variation of plasma or equilibrium quantities. This is an efficient way to gain GK insight on the physics, but the local approach has no information on the effective normalized gyro-radius, $\rho*$, and can be misleading if the background quantities vary over short radial scales relative to the radial wave number of the mode [30, 31]. One of the expected challenges in local simulations is the emergence of MHD-like modes, such as the Kinetic Ballooning Mode (KBM), that would potentially be stable in global simulations due to the limited radial extent of pedestal [30]. On the other hand, Electron Temperature Gradient mode (ETG) that drives significant electron heat transport is characterized by sufficiently short radial wavelengths that the local approximation is expected to be quite good [10, 32]. Almost all other modes, such as the Micro-Tearing Mode (MTM), are likely to have sensitivity to finite radial extent of the pedestal [32]. Nevertheless, the local linear approach is still expected to provide sufficiently good approximation of the fundamental turbulence drive that with an appropriate mapping from local to nonlinear, subject to research by Prof. Hatch and his team, a connection to transport with effective $\rho*$ taken into account can be established.

To implement this work, it is expected that dedicated efforts are needed to establish the necessary software infrastructure to robustly execute (T1.1) and analyze (T1.2) large batches of GENE simulations on LUMI. Once this infrastructure is established, the appropriate input parameter space and boundaries are determined based on previous databases of GK simulations for pedestal plasmas, such as described by Hatch et al. in [10] (T1.3). In [10], a reduced model for ETG transport in pedestal was developed based on 61 nonlinear GENE simulations. In DEEPlasma, with computational resources equivalent of a few million CPU hours through the LUMI Regular Access mode, a database of thousands of local, linear GENE simulations will be established (T1.4). Aligning this database with the operational space covered by the nonlinear GENE simulations and spanning multiple devices (DIII-D, JET, C-Mod, AUG) and operational scenarios documented in [10, and references therein] is also expected to expedite the progress in developing the mapping from local, linear to transport by Prof. Hatch and his team. Furthermore, both Dr. Järvinen and Prof. Hatch have strong connections to the experimental fusion research programs within EUROfusion and US, such that the experimental scenario relevant input parameter domains can be further extended whenever the established GK databases have sufficiently covered the previously selected input domains.

The GK plasma turbulence simulations will be conducted with the simulation code GENE, which is a 5D continuum GK code [26] (https://genecode.org/). The international collaborator, Prof. Hatch, has several years of experience in applying GENE for pedestal GK simulations [7, 8, 10, 32, 30]. Dr. Järvinen has also access to the GENE repository maintained by the GENE development team.

WP2: Physics-informed machine learning for accurate surrogate model development with sparse training data

To first establish the regression baseline, traditional machine learning models are trained with the proof-of-principle database established in WP1 (T2.1). These are expected to provide the worst-case performance when only mean square error or similar regression performance metric is applied without any actual physics-informed constraints. These traditional machine learning models will be implemented in TensorFlow and Keras [33, 34] as simple deep learning neural networks [35] with a relatively low number of convolutional and dense layers with ReLU activations, and trained with standard methods such as SGD or Adam.

Once the regression performance baseline is established, physics-informed constraints will be investigated. Finding appropriate physics-laws to constrain machine learning regression models for linear GK stability calculations can be quite challenging. Therefore, this is identified as a separate task in the work (T2.2). Physics-informed constraints can probably be constructed based on dispersion relations, along the lines discussed by Hatch et al. [32] and Parisi et al. [36]. Furthermore, physically plausible combinations of turbulence drives, growth rates, mode frequencies, and eigenfunctions can be deduced from the characteristics of the various modes and these can be provided as constraints to the machine learning model.

Having identified suitable physics-informed constraints to the models, the physics-informed models will be trained with the established database (T2.3) and compared to the regression performance obtained without physical constraints (T2.4). For the training we aim to leverage physics-informed neural network methods already available in TensorDiffEq package [37], along with Gaussian process based PDE methods [38, 39, 40, 41, 42, 17, 43], as well as their combinations. To implement Gaussian process models in TensorFlow, we initially use GPFlow [44], but the more complicated physics-informed models and hybrid models will need to be implemented as custom layers and cost-functions on top of Keras.

An example of physics-informed neural network solution of (viscous) Burgers' equation computed by using the TensorDiffEq package [37] is shown in Figure 3. While this only a toy example of a physics-informed neural network, it shows the utility of these algorithms in solving convection-diffusion type partial differential equations (PDE). In WP2, a significant part of the work is to investigate approaches to apply these methods for significantly more complicated simulations of continuum GK of fusion plasma. This will also help to develop the physics-informed machine learning methodology for other computationally complicated PDE solvers.

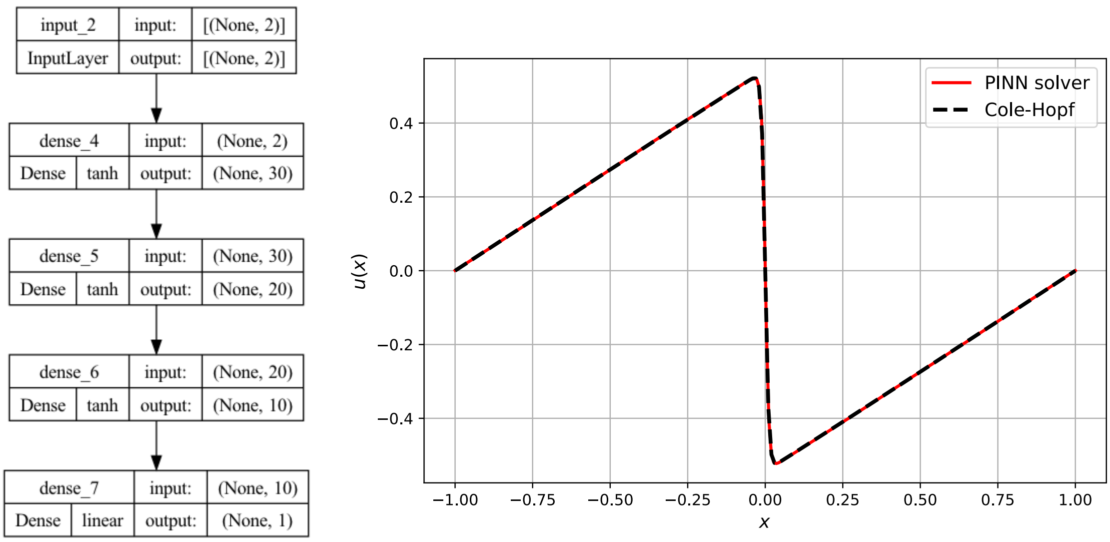

Figure 3: Neural network architecture used in the PINN solver (left) and the solver result compared to a Cole-Hopf reference solution (right). It can be seen that the results agree very well.

WP3: Methods to quantify uncertainties and improve surrogate model accuracy through Bayesian approaches

The established surrogate model is an approximation of the ground truth full model and uncertainty quantification is, therefore, essential to gauge the quality of the surrogate model predictions. We aim to use Bayesian machine learning methods for neural networks and Gaussian processes [45, 46] to model the uncertainties caused by the model uncertainty (T3.1). When applied to surrogate model training on the dataset, they can provide uncertainty bounds on the PEGENENN predictions (T3.2). For computations, we aim to initially use the variational Bayesian methods and Markov chain Monte Carlo methods already available in TensorFlow probability [47], but we will significantly need to extend them for the physics-informed models. Figure 4 shows an example of a simple Gaussian process-based probabilistic solution to the (viscous) Burger's equation using a yet-unpublished PDE-extension of the method from [43].

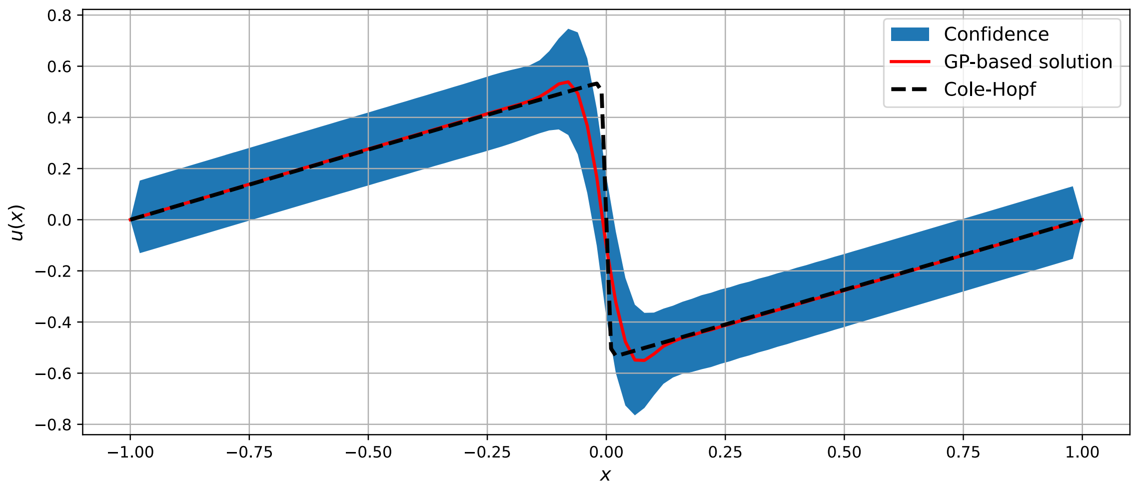

Figure 4: Gaussian process (GP)-based probabilistic solution to Burger's equation. The Gaussian process-based solver also provides the uncertainty (confidence interval) for the solution.

As the model becomes aware about the inference uncertainty, such an uncertainty can be applied in an acquisition function to recommend new evaluations of the computationally costly full model that would lead to optimal reduction of the prediction uncertainty (T3.3). In this project this active learning task is done using Bayesian optimization methods [48] which are methods that can do active learning while properly taking the uncertainties into account. Once appropriate probabilistic methods and acquisition functions are established, these will be used to enhance the training dataset to optimize the accuracy of the surrogate model with the given computational resources (T3.4).

WP4: Scaling PEGENENN from proof-of-principle implementation to a production model with LUMI Extreme Scale Access

To establish a production level PEGENENN trained on a multidimensional input space, extreme computational resources are expected to be needed. The philosophy of this research plan is to build the necessary algorithmic and software infrastructure and proof-of-principle implementation in WP1, WP2, and WP3 to enable leveraging the extreme scale LUMI resources in WP4. To do this, as soon as sufficient level of baseline software infrastructure is developed in the preceding WPs, an application for the LUMI Benchmark Access mode is prepared (T4.1). Assuming a successful demonstration of the approach in the LUMI Benchmark mode (T4.2), a proposal for the LUMI Extreme Scale access will be prepared (T4.3). Once granted, the approach prepared in WP1, WP2, and WP3 will be harnessed to generate production level PEGENENN with the LUMI Extreme Scale resources.

Research infrastructures and environments

LUMI is the primary computing resource used in this project. LUMI is located in the data center of CSC in Kajaani and is hosted by the LUMI consortium, consisting of Finland, Belgium, the Czech Republic, Denmark, Estonia, Iceland, Norway, Poland, Sweden, and Switzerland. LUMI is the fastest supercomputer in Europe and the third fastest globally with a sustained computing power of 375 petaflops in its final configuration. As this project consists of academic research conducted within Finnish higher-education and research institutions, the LUMI resources are available to the project free of charge. For the Regular Access mode, users can apply anytime and decisions on allocation are made every three weeks. The Regular Access mode is intended for middle size projects with maximum requirements of 8 million CPU core-h, 0.5 million GPU hours, and 570 000 TiB hours. These are foreseen to be sufficient for the WP1, WP2, and WP3 of the project that aim to establish the proof-of-principle local, linear pedestal GENE surrogate as well as the necessary infrastructure for training database population. Within these work packages, the research team plans to complete the benchmark test within the Benchmark Access mode, such that Extreme Scale Access can be applied for WP4 to generate the production ready local, linear pedestal GENE surrogate. Within the Extreme Scale Access, the available resources are limited between 0.5 – 3 million GPU hours, and maximum of 16 million CPU core-hours. The Extreme Scale Access application proposals will go through a technical review and scientific evaluation in an international panel. Therefore, the Extreme Scale Access application is planned for the second year of the project to have access to the resources during the third year of the project.

Finnish Center for Artificial Intelligence (FCAI) is a nation-wide competence center for Artificial Intelligence in Finland, initiated by Aalto University, University of Helsinki, and VTT Technical Research Centre of Finland. In 2019 FCAI gained the Academy of Finland flagship status, a standing-only granted to a few centers of excellence of high quality and high societal impact. Prof. Särkkä leads the AI Across Fields (AI for X) program in FCAI.

Aalto University. Prof. Simo Särkkä is with the Department of Electrical Engineering and Automation (EEA) at Aalto University which is a multidisciplinary department which combines expertise in many traditional fields of electrical engineering with automation and systems science and applies it in business, industry, and health care. Aalto University is very well equipped with high-end cluster-computer and storage facilities for high-performance computing and modeling, and large scale data-handling and data-mining purposes.

VTT Technical Research Centre of Finland (VTT). Dr. Aaro Järvinen is with VTT which is one of Europe's leading research institutions and it is owned by the Finnish state. VTT advances the utilization and commercialization of research and technology in commerce and society. Through scientific and technological means, it turns large global challenges into sustainable growth for businesses and society. It brings together people, business, science and technology to solve the biggest challenges of our time.

2.3 Risk assessment and alternative implementation strategies:

The primary risk in this project is related to access to sufficient computational time. Since usage of LUMI supercomputer was one of the special objectives of this call, it is expected that LUMI resources will be available as well. However, in the unlikely event that LUMI resources will not be available, it is possible to conduct parts of this work with the other HPC systems of CSC, such as MAHTI and PUHTI. In the very unlikely event of no access to CSC resources at all, work can also be conducted at smaller scale on the computational clusters of Aalto University (Triton) and VTT (The Doctor).

As this is ambitious scientific research, there is always a risk of not fully achieving all the objectives. However, the project is anyway expected to significantly advance the understanding on how to build fast and accurate surrogate models for computationally very expensive forward models, such as applied for pedestal GK in this work.

As this project will be conducted by PhD students and postdocs, a human resources risk is always associated with this type of work. The mitigation strategy is to carefully select the most suitable candidates for the project and carefully tutor them to get on the right track with the project.

2.4 Added value of consortium (if the application is a consortium application):

This consortium together with the international collaborators in Texas will bring together several decades' worth of experience on both magnetic confinement fusion plasma physics (Dr. Järvinen at VTT and Prof. Hatch at University of Texas in Austin) as well as machine learning, including the modern physics-informed approaches (Prof. Särkkä at Aalto University and Prof. Braga-Neto at Texas A&M University). Taming edge plasma turbulence with modern machine learning algorithms is a very ambitious and demanding goal that is likely to truly challenge both the scientific machine learning experts and plasma physicists to the very edge of their capabilities. Forming such a consortium brings together the critical mass of expertise that when provided the LUMI supercomputer resources has a great potential to reach this ambitious goal.

3 Research team and collaborators

3.1 Project personnel and their project-relevant key merits:

Särkkä group (sensor informatics and medical technology), Aalto University (Aalto): Simo Särkkä, D.Sc., Associate Professor, partner PI at Aalto, Department of Electrical Engineering and Automation. Prof. Särkkä's and his group's (~15 members) research interests are in multi-sensor data processing systems, artificial intelligence, and machine learning with applications including health and medical technology. He has authored or coauthored 186 peer-reviewed scientific articles (h-index 42, 10278 total citations, Google Scholar) and his books "Bayesian Filtering and Smoothing" (1st and 2nd edition) and "Applied Stochastic Differential Equations" along with the Chinese translation of the former have been published via the Cambridge University Press. He has acquired external funding as PI amounting to approx. 4M€. He is a Fellow of European Laboratory for Learning and Intelligent Systems (ELLIS), and the leader of AI Across Fields (AIX) program at Finnish Center for Artificial Intelligence (FCAI). Prof. Särkkä is responsible of the machine learning and AI expertise in the project, supervises post-docs and Ph.D. students, and acts as a contact to related international experts in the field. He is a Senior Member of IEEE and serving as a Senior Area Editor of IEEE Signal Processing Letters.

M.Sc. Sahel Mohammad Iqbal will work as a PhD student level researcher in the project. He received his Integrated Master's degree in Physics with thesis title "Dataset debiasing and network pruning" in 2022. Currently, he is a PhD student in Prof. Simo Särkkä's research group at the Department of Electrical Engineering and Automation (EEA) at Aalto University. He has experience working on computational physics and machine learning projects throughout his Bachelor's and Master's studies. As part of his PhD, he currently works on physics-informed machine learning methods.

Dr. Hany Abdulsamad (PhD 2021) will work as a post-doc in the project and guide Sahel during the first year of the project to get started efficiently. Hany's expertise lies at an intersection of control theory and machine learning. His prior work focused on structured representation learning for identification and control of physical systems. He is currently working under the Finnish Center for Artificial Intelligence (FCAI) flagship funding under Prof. Simo Särkkä's and FCAI director Samuel Kaski's guidance.

Fusion Energy and Decommissioning group, VTT Technical Research Centre of Finland (VTT): Aaro Järvinen, D.Sc., Senior Scientist, partner PI at VTT, Fusion Energy and Decommissioning. Dr. Järvinen has a broad previous experience of investigating power exhaust and core-edge integration in magnetic confinement fusion devices, including experience with multiple codes, diagnostic systems, machine learning approaches to fusion, as well as experiment coordination at several fusion facilities, such as JET, DIII-D, AUG, TCV, and MAST-U. Dr. Järvinen obtained his doctoral degree from Aalto University in 2015 on radiative divertor studies in JET. The years following completing the doctoral degree, he was employed by the Lawrence Livermore National Laboratory (LLNL) and was appointed as an area leader for scrape-off layer (SOL) flow studies, a co-leader of the core-edge integration task force at DIII-D, as well as co-chair of the boundary physics expert group in the USA community planning process for magnetic fusion energy research. In 2021, Dr. Järvinen joined the Fusion Energy and Decommissioning research team at VTT as a senior scientist. Despite being an early-career scientist, he is an author or co-author of more than 60 peer-reviewed publications and has given several invited and contributed oral presentations in international conferences. At the moment, he is leading a sub-project on developing machine learning methods for data-driven pedestal models within a broader EUROfusion project focused on machine learning applications for plasma-exhaust and plasma-wall interaction in fusion. For the past two years, he has been instructing a student at the University of Helsinki, investigating plasma state representation learning applications for fusion research. In 2023, he has also been instructing another student at the University of Helsinki investigating machine learning feature storage solutions for fusion research. Since 2022, Dr. Järvinen has been leading a project at VTT investigating the role of cross-field drifts in divertor power exhaust in the UKAEA Spherical Tokamak for Energy Production project. In 2023, he was nominated as an ITER Science Fellow for the Divertor and SOL simulations including drifts, currents and 3D fields area of research.

A PhD student level researcher will be hired at VTT for this project. As soon as positive funding decision is obtained, recruitment process is started. Based on previous experience, many competent candidates have indicated interest for machine learning applications in fusion and it is foreseen that such an open position will be filled by the time the project is projected to start.

Links to previous research

The project will heavily benefit from on-going NSF project of the Texas collaborator Prof. Ulisses Braga-Neto on "A Bayesian Paradigm for Physics-Informed Machine Learning" in which Prof. Särkkä has Academy of Finland counterpart within the NSF-AKA joint funding scheme. Furthermore, the previous projects "Probabilistic Deep Learning via Hierarchical Stochastic Partial Differential Equations" and "Parallel and distributed computing for Bayesian graphical models" from Academy of Finland ICT2023 program, in which Prof. Särkkä was the PI provide essential core methodology for the current project.

Dr. Järvinen has been working in the area of core-edge integration and power exhaust in magnetic confinement fusion devices more than a decade. Within the past few years, he has also applied machine learning and data science approaches to fusion studies, including deep learning for experimental database and Bayesian optimization for calibration of computationally costly forward models. At the moment, he is a sub-project leader in a EUROfusion project investigating machine learning applications for power exhaust in fusion.

Researcher training

Both of the PIs are dedicated to hands-on guidance of the doctoral researchers and post-docs working in the project. The doctoral researchers (PhD students) working in the project will write scientific articles in the project which will steer them towards finishing the doctoral dissertation. The grant will also support post-doc's early career development by helping them to mature as scientists and by connecting them to international networks of researchers. The learnings from the project will also be used to create teaching material for courses given to next generation of researchers.

3.2 Collaborators and their project-relevant key merits:

The main collaborators in the project are Profs. Hatch and Braga-Neto from Texas, USA. Their backgrounds and contributions to the project are in turbulence modeling and physics-informed machine learning, respectively. In addition to regular Zoom/Teams online meetings, the collaboration will involve longer research visits to USA.

Prof. David R. Hatch is research professor at the Institute for Fusion Studies in the Physics Department at the University of Texas at Austin. His research interests are oriented around both fundamental and applied aspects of nonlinear plasma dynamics. A major part of his work exploits the power of high-performance computing to understand and optimize turbulent transport in fusion devices. In particular, in this project will be related to the pioneering the use of gyrokinetic simulations in the critical edge transport barriers that will largely determine the appeal and practicality of fusion as an energy source.

Prof. Ulisses M. Braga-Neto provides physics-informed machine learning expertise in the project. He is Professor at the Department of Electrical and Computer Engineering, Texas A&M University. Prof. Braga-Neto received his Ph.D. degree in electrical and computer engineering from The Johns Hopkins University, Baltimore, Maryland. He is currently a professor in the Department of Electrical and Computer Engineering at Texas A&M University, College Station. His research interests include pattern recognition, machine learning, and statistical signal processing. He is the author of the recent textbook, Fundamentals of Pattern Recognition and Machine Learning.

4 Responsible science

4.1 Research ethics, equality and non-discrimination, open science, and sustainable development:

Main ethical and legal guidelines for this project are set by the EU Regulation of Ethical AI and the EU General Data Protection Regulation (GDPR). These regulations cover AI, robotics, related techniques, software, and data. In addition, we will follow ethical guidelines of the Academy of Finland, and Aalto University and VTT. The project does not involve collection of personal data which would require anonymization.

PIs at Aalto and VTT are fully committed to the promotion of equal opportunities, fairness, and gender equality. We will recruit on the basis of merit, ability, and potential without regard to race, color, national origin, sex, sexual orientation, genetic information, gender identity, gender expression, religion, disability, or similar. However, mathematics, physics and engineering have traditionally been strongly male-dominated fields. Hence, we encourage individuals belonging to groups known to have been discriminated against or underrepresented previously.

In publishing, we will follow open access policies and the Academy of Finland rules. For this purpose, we can use the open-access agreements made by FinELib, who coordinates national-level open publication fee payments, and specific arrangements made by Aalto University. These arrangements cover open access fees for prominent publishers, including Elsevier, IEEE, Springer Nature, and Wiley. We will choose the publication forums in such a way that we can maximally utilize these arrangements. Moreover, following community standards, all the manuscripts will be made public in arXiv and through personal web pages.

We will also follow open data and open-source software publishing policies. Open data: the large GK simulation databases generated within this work will be made available as much as possible through the already established community mainted databases, such as the multiscale GK database (MGKDB) (https://github.com/alblackmon/mgkdb or the GK database (GKDB) https://gitlab.com/gkdb/gkdb/-/wikis/home or a similar community shared storage solution for GK simulations. When locally stored, the datasets will be stored on password protected servers owned by Aalto University or VTT, and the access will only be given to project personnel. Open-source codes: All the developed software (when possible) will be published through GitHub under free software licenses. We also use GitHub for version control throughout the project.

We target the United Nations' sustainable development goals (SDG) 9 and 13 on resiliency and climate action, respectively. With respect to SDG9, the developed technology for nuclear fusion will allow for cheaper energy and hence availability of energy and wealth more widely and equally in the world. As for SDG13, the developed energy production technology is also aimed to be one of the solutions to the climate crisis.

5 Societal effects and impact

5.1 Effects and impact beyond academia:

Nuclear fusion offers a virtually unlimited, carbon free energy source that is evenly distributed around the globe. Since climate change is considered as one of the primary threats of modern society, it is fair to say that fusion energy, if realized, will be one of the greatest scientific achievements of the 21st century. Due to this reason, sufficient resources should be allocated for research to harness fusion energy globally as an energy source to replace carbon intensive alternatives as soon as possible. The AI algorithms developed and evaluated throughout this project are expected to be applicable to a very broad range of academic and commercial applications.

References

The project PIs and collaborators are marked as bold.

[1] ITER Organization, "ITER Research Plan within the Staged Approach (Level III – Provisional Version)," Tech. Rep. ITR-18-003, 2018.

[2] M. Wischmeier, "High density operation for reactor-relevant power exhaust," Journal of Nuclear Materials, vol. 463, pp. 22–29, 2015. PLASMA-SURFACE INTERACTIONS 21.

[3] A. Kallenbach, M. Bernert, R. Dux, L. Casali, T. Eich, L. Giannone, A. Herrmann, R. McDermott, A. Mlynek, H. W. Müller, F. Reimold, J. Schweinzer, M. Sertoli, G. Tardini, W. Treutterer, E. Viezzer, R. Wenninger, M. Wischmeier, and the ASDEX Upgrade Team, "Impurity seeding for tokamak power exhaust: from present devices via ITER to DEMO," Plasma Physics and Controlled Fusion, vol. 55, p. 124041, nov 2013.

[4] **A. E. Jaervinen**, S. L. Allen, D. Eldon, M. E. Fenstermacher, M. Groth, D. N. Hill, A. W. Leonard, A. G. McLean, G. D. Porter, T. D. Rognlien, C. M. Samuell, and H. Q. Wang, "E × B Flux Driven Detachment Bifurcation in the DIII-D Tokamak," Phys. Rev. Lett., vol. 121, p. 075001, Aug 2018.

[5] **A. Jaervinen**, S. Allen, D. Eldon, M. Fenstermacher, M. Groth, D. Hill, C. Lasnier, A. Leonard, A. McLean, G. Porter, T. Rognlien, C. Samuell, H. Wang, and J. Watkins, "Impact of drifts on divertor power exhaust in DIII-D," Nuclear Materials and Energy, vol. 19, pp. 230–238, 2019.

[6] C. Chang, S. Ku, A. Loarte, V. Parail, F. Köchl, M. Romanelli, R. Maingi, J.-W. Ahn, T. Gray, J. Hughes, B. LaBombard, T. Leonard, M. Makowski, and J. Terry, "Gyrokinetic projection of the divertor heat-flux width from present tokamaks to ITER," Nuclear Fusion, vol. 57, p. 116023, aug 2017.

[7] **D. Hatch**, M. Kotschenreuther, S. Mahajan, P. Valanju, F. Jenko, D. Told, T. Görler, and S. Saarelma, "Microtearing turbulence limiting the JET-ILW pedestal," Nuclear Fusion, vol. 56, p. 104003, aug 2016.

[8] **D. Hatch**, M. Kotschenreuther, S. Mahajan, P. Valanju, and X. Liu, "A gyrokinetic perspective on the JET-ILW pedestal," Nuclear Fusion, vol. 57, p. 036020, jan 2017.

[9] M. T. Curie, J. L. Larakers, **D. R. Hatch**, A. O. Nelson, A. Diallo, E. Hassan, W. Guttenfelder, M. Halfmoon, M. Kotschenreuther, R. D. Hazeltine, S. M. Mahajan, R. J. Groebner, J. Chen, D.-D. Team, C. Perez von Thun, L. Frassinetti, S. Saarelma, C. Giroud, J. Contributors, and M. M. Tennery, "A survey of pedestal magnetic fluctuations using gyrokinetics and a global reduced model for microtearing stability," Physics of Plasmas, vol. 29, 04 2022. 042503.

[10] **D. R. Hatch**, C. Michoski, D. Kuang, B. Chapman-Oplopoiou, M. Curie, M. Halfmoon, E. Hassan, M. Kotschenreuther, S. M. Mahajan, G. Merlo, M. J. Pueschel, J. Walker, and C. D. Stephens, "Reduced models for ETG transport in the tokamak pedestal," Physics of Plasmas, vol. 29, 06 2022. 062501.

[11] B. Chapman-Oplopoiou, **D. Hatch**, A. Field, L. Frassinetti, J. Hillesheim, L. Horvath, C. Maggi, J. Parisi, C. Roach, S. Saarelma, J. Walker, and J. Contributors, "The role of ETG modes in JET-ILW pedestals with varying levels of power and fuelling," Nuclear Fusion, vol. 62, p. 086028, jun 2022.

[12] P. Snyder, R. Groebner, J. Hughes, T. Osborne, M. Beurskens, A. Leonard, H. Wilson, and X. Xu, "A first-principles predictive model of the pedestal height and width: development, testing and ITER optimization with the EPED model," Nuclear Fusion, vol. 51, p. 103016, aug 2011.

[13] N. Baker, F. Alexander, T. Bremer, A. Hagberg, Y. Kevrekidis, H. Najm, M. Parashar, A. Patra, J. Sethian, S. Wild, K. Willcox, and S. Lee, "Workshop report on basic research needs for scientific machine learning: Core technologies for artificial intelligence," 2 2019. 18(19)

[14] S. Wang, Y. Teng, and P. Perdikaris, "Understanding and mitigating gradient pathologies in physics-informed neural networks," arXiv preprint arXiv:2001.04536, 2020.

[15] S. Wang, X. Yu, and P. Perdikaris, "When and why PINNs fail to train: A neural tangent kernel perspective," arXiv preprint arXiv:2007.14527, 2020.

[16] X. Yang, D. Barajas-Solano, G. Tartakovsky, and A. M. Tartakovsky, "Physics-informed cokriging: A Gaussian-process-regression-based multifidelity method for data-model convergence," Journal of Computational Physics, vol. 395, pp. 410–431, 2019.

[17] Y. Chen, B. Hosseini, H. Owhadi, and A. M. Stuart, "Solving and learning nonlinear PDEs with Gaussian processes," Journal of Computational Physics, vol. 447, p. 110668, 2021.

[18] M. Raissi and G. E. Karniadakis, "Hidden physics models: Machine learning of nonlinear partial differential equations," Journal of Computational Physics, vol. 357, pp. 125–141, 2018.

[19] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," Journal of Computational Physics, vol. 378, pp. 686–707, 2019.

[20] D. Liu and Y. Wang, "Multi-fidelity physics-constrained neural network and its application in materials modeling," Journal of Mechanical Design, vol. 141, no. 12, 2019.

[21] D. Liu and Y. Wang, "A dual-dimer method for training physics-constrained neural networks with minimax architecture," Neural Networks, vol. 136, pp. 112–125, 2021.

[22] X. Jin, S. Cai, H. Li, and G. E. Karniadakis, "NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible Navier-Stokes equations," Journal of Computational Physics, vol. 426, p. 109951, 2021.

[23] S. Cai, Z.Wang, L. Lu, T. A. Zaki, and G. E. Karniadakis, "DeepM&Mnet: Inferring the electroconvection multiphysics fields based on operator approximation by neural networks," Journal of Computational Physics, vol. 436, p. 110296, 2021.

[24] M. T. Rad, A. Viardin, G. Schmitz, and M. Apel, "Theory-training deep neural networks for an alloy solidification benchmark problem," Computational Materials Science, vol. 180, p. 109687, 2020.

[25] Q. Zhu, Z. Liu, and J. Yan, "Machine learning for metal additive manufacturing: predicting temperature and melt pool fluid dynamics using physics-informed neural networks," Computational Mechanics, vol. 67, no. 2, pp. 619–635, 2021.

[25] F. Jenko, W. Dorland, M. Kotschenreuther, and B. N. Rogers, "Electron temperature gradient driven turbulence," Physics of Plasmas, vol. 7, pp. 1904–1910, 05 2000.

[27] J. Citrin, S. Breton, F. Felici, F. Imbeaux, T. Aniel, J. Artaud, B. Baiocchi, C. Bourdelle, Y. Camenen, and J. Garcia, "Real-time capable first principle based modelling of tokamak turbulent transport," Nuclear Fusion, vol. 55, p. 092001, jul 2015.

[28] K. L. van de Plassche, J. Citrin, C. Bourdelle, Y. Camenen, F. J. Casson, V. I. Dagnelie, F. Felici, A. Ho, S. Van Mulders, and J. Contributors, “Fast modeling of turbulent transport in fusion plasmas using neural networks,” Physics of Plasmas, vol. 27, 02 2020. 022310.

[29] O. Meneghini, S. Smith, P. Snyder, G. Staebler, J. Candy, E. Belli, L. Lao, M. Kostuk, T. Luce, T. Luda, J. Park, and F. Poli, “Self-consistent core-pedestal transport simulations with neural network accelerated models,” Nuclear Fusion, vol. 57, p. 086034, jul 2017.

[30] B. F. McMillan, X. Lapillonne, S. Brunner, L. Villard, S. Jolliet, A. Bottino, T. Görler, and F. Jenko, “System size effects on gyrokinetic turbulence,” Phys. Rev. Lett., vol. 105, p. 155001, Oct 2010.

[31] **D. Hatch**, M. Kotschenreuther, S. Mahajan, M. Pueschel, C. Michoski, G. Merlo, E. Hassan, A. Field, L. Frassinetti, C. Giroud, J. Hillesheim, C. Maggi, C. P. von Thun, C. Roach, S. Saarelma, D. Jarema, F. Jenko, and J. Contributors, “Microtearing modes as the source of magnetic fluctuations in the JET pedestal,” Nuclear Fusion, vol. 61, p. 036015, feb 2021.

[32] **D. Hatch**, M. Kotschenreuther, S. Mahajan, G. Merlo, A. Field, C. Giroud, J. Hillesheim, C. Maggi, C. P. von Thun, C. Roach, S. Saarelma, and J. Contributors, “Direct gyrokinetic comparison of pedestal transport in JET with carbon and ITER-like walls,” Nuclear Fusion, vol. 59, p. 086056, jul 2019.

[33] M. Landreman and D. R. Ernst, “Local and global Fokker–Planck neoclassical calculations showing flow and bootstrap current modification in a pedestal,” Plasma Physics and Controlled Fusion, vol. 54, p. 115006, oct 2012. 19(19)

[34] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., “Tensor flow: A system for large-scale machine learning,” in 12th fUSENIXg Symposium on Operating Systems Design and Implementation (fOSDIg 16), pp. 265–283, 2016.

[35] F. Chollet, Deep learning with Python. Simon and Schuster, 2021.

[36] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016.

[37] J. F. Parisi, F. I. Parra, C. M. Roach, C. Giroud, W. Dorland, **D. R. Hatch**, M. Barnes, J. C. Hillesheim, N. Aiba, J. Ball, P. G. Ivanov, and J. contributors, “Toroidal and slab ETG instability dominance in the linear spectrum of JET-ILW pedestals,” Nuclear Fusion, vol. 60, p. 126045, oct 2020.

[38] L. D. McClenny, M. A. Haile, and **U. M. Braga-Neto**, “TensorDiffEq: Scalable multi-GPU forward and inverse solvers for physics informed neural networks,” arXiv preprint arXiv:2103.16034, 2021.

[39] T. Graepel, “Solving noisy linear operator equations by Gaussian processes: Application to ordinary and partial differential equations,” in ICML, vol. 3, pp. 234–241, 2003.

[40] A. Solin, M. Kok, N. Wahlström, T. B. Schön, and **S. Särkkä**, “Modeling and interpolation of the ambient magnetic field by Gaussian processes,” IEEE Transactions on Robotics, vol. 34, no. 4, pp. 1112–1127, 2018.

[41] F. Tronarp, H. Kersting, **S. Särkkä**, and P. Hennig, “Probabilistic solutions to ordinary differential equations as nonlinear Bayesian filtering: a new perspective,” Statistics and Computing, vol. 29, no. 6, pp. 1297–1315, 2019.

[42] **S. Särkkä**, M. A. Álvarez, and N. D. Lawrence, “Gaussian process latent force models for learning and stochastic control of physical systems,” IEEE Transactions on Automatic Control, vol. 64, no. 7, pp. 2953–2960, 2019.

[43] A. Solin and **S. Särkkä**, “Hilbert space methods for reduced-rank Gaussian process regression,” Statistics and Computing, vol. 30, pp. 419–446, 2020.

[44] F. Tronarp, **S. Särkkä**, and P. Hennig, “Bayesian ODE solvers: the maximum a posteriori estimate,” Statistics and Computing, vol. 31, no. 23, 2021.

[45] A. G. d. G. Matthews, M. van der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. León-Villagrá, Z. Ghahramani, and J. Hensman, “GPflow: A Gaussian process library using TensorFlow,” Journal of Machine Learning Research, vol. 18, pp. 1–6, apr 2017.

[46] R. M. Neal, Bayesian learning for neural networks, vol. 118. Springer Science & Business Media, 2012.

[47] C. E. Rasmussen and C. K. Williams, Gaussian processes for machine learning. MIT press, 2006.

[48] J. V. Dillon, I. Langmore, D. Tran, E. Brevdo, S. Vasudevan, D. Moore, B. Patton, A. Alemi, M. Ho man, and R. A. Saurous, “Tensorflow distributions,” 2017. arXiv preprint arXiv:1711.10604.

[49] R. Garnett, Bayesian Optimization. Cambridge University Press, 2023.