PAPU Research plan

Amanda Bruncrona

Preliminary title of the doctoral thesis

Acceleration of MHD stability simulations with machine learning techniques

Field of research (Please use [Academy of Finland's Research field classification](#))

1043 Fluid and plasma physics

Brief summary of the research plan (2000 characters)

Magnetohydrodynamic (MHD) simulations are applied ubiquitously in plasmas physics. In the edge of magnetically confined tokamak plasmas, MHD phenomena have a key role in determining the boundaries of the achievable plasma pressure; hence, providing a boundary condition for the overall performance of the confined plasma in fusion reactor applications [11]. Current tools for predicting the performance of these tokamak edge plasmas, also known as the pedestal region, are EPED, Europed, IPED and IMEP [3 – 6]. All of these include an equilibrium solver, such as HELENA [6], and ideal/resistive MHD stability codes, such as MISHKA [7] and CASTOR [12]. For agile reactor and scenario design as well as real-time scenario optimization, these MHD simulations would have to be conducted significantly faster than achievable with traditional code optimization strategies without unacceptable loss of accuracy. The main aim of this doctoral thesis is to investigate how these goals can be achieved through machine learning (ML) techniques.

Previous studies have implemented ML surrogate models for EPED or Europed [9, 10]. However, the computationally demanding part in EPED-like models is the peeling-ballooning MHD stability calculation, obtained through repeated equilibrium and stability simulations. This work is therefore proposing to focus on ML surrogate models for this part of the simulation. EPED-like models combine the MHD stability simulations with reduced model assumptions about the transport properties of the plasmas. By surrogating the stability simulation only, it becomes possible to improve the physics fidelity of the transport model while using the same surrogate model for the MHD stability part.

Firstly, a database of the full physics model including both HELENA and MISHKA will be generated. This database will be used to train various machine learning models to predict the MHD stability. As an initial proof-of-principle, the database is generated for a subset of Joint European Torus (JET) tokamak plasmas in order to reduce the dimensionality of parameters. Various supervised ML architectures will be explored to generate a regression model on the database. When extending the model beyond the proof-of-principle, physics-informed and active learning approaches will be investigated to achieve the necessary data efficiency. Extension of the surrogate model with resistive MHD features based on a resistive MHD stability code CASTOR will be investigated. The resulting surrogate model will be used to analyse JET plasmas within domains of known resistive MHD features. The fast predictive capabilities will be used for large scale analysis of experimental observation from tokamaks operated under the umbrella of the EUROfusion consortium.

(Reference list in “**The most important literature for the research plan”** section.)

Motivation for conducting doctoral research (2000 characters)

Describe your motivation for conducting doctoral research in general and for this topic in particular. Also explain why you have chosen the doctoral programme you are applying to.

I wish to conduct doctoral research to gain expertise in an interesting field in which I want to continue working as a researcher. It is important for me that I can continue studying and learning through my work and that my work is both fun and purposeful. This doctoral research project would perfectly combine my interests in physics, machine learning and sustainability.

The theoretical background for this research includes plasma physics, magnetohydrodynamics (MHD) and machine learning. MHD is used widely in both fusion physics and space physics, since both concerns plasmas. The Doctoral Programme in Particle Physics and Universe Sciences (PAPU) would therefore be a good fit.

Rationale for the research project (2000 characters)

How is the project linked to previous research? What are the most significant theoretical and methodological premises of the project?

Tokamaks operating in high confinement mode display a characteristic edge plasma pedestal, associated with the edge transport barrier, where the plasma pressure gradient rapidly increases which leads to improved edge confinement. Magnetohydrodynamic (MHD) simulations are used to predict pedestal performance. Fast and accurate pedestal predictions are needed to improve high performance scenarios in tokamaks. Although the MHD stability simulations are an excellent candidate to be accelerated with ML techniques, this type of research has not been conducted before. Therefore, this doctoral thesis project aims to tackle this challenge and, if successful, is expected to lead to a widely adopted set of surrogate models to be used by fusion plasma physics community. Furthermore, the developed ML techniques and approaches are likely to be applicable to other domains of plasma physics, where MHD stability simulations are relevant.

Objectives of the research and scientific impact of research results (2000 characters)

What are the objectives of the project? Shortly present the hypotheses and the research questions. Describe the expected research result and their anticipated novelty value in terms of the research field and the current scientific discussion on the research topic. You may also shortly outline the reach, potential applications and utilization value of the research beyond the scientific community.

The objective is to develop a fast and accurate ML surrogate model for the edge MHD stability calculations for magnetically confined tokamak plasmas. The developed methodology can probably be used to accelerate MHD simulations for other plasma physics applications as well.

Computational methods for simulating the plasma pedestal exist (EPED, Europed, IPED and IMEP [3 – 6]). However, they are computationally demanding, while still relying on reduced model assumptions on the plasma characteristics. Previous research has proposed surrogate models for the entire programs mentioned, including the model reductions. However, these reduced model assumptions are typically fast to evaluate. As a result, little computational throughput gain is obtained by including them in the surrogate model, while the applicability of the established surrogate becomes limited due to the reduced model assumptions being directly encoded in the model. Instead, in this doctoral thesis, a flexible surrogate model will be developed for the computationally demanding part of these simulations, i.e. the MHD stability calculations. Such a model would be much more flexible to be used as an ML surrogate component in EPED, Europed, IPED, IMED, or any other integrated modelling workflow. Within this doctoral thesis, the ML surrogate model will be used both to analyze experimental data and to accelerate the existing plasma pedestal simulations.

Research methods and materials to be used and its significance for the project (3500 characters)

Outline the research methods, described so as to explain how they will contribute to answering the research questions/confirming the hypotheses, or how they will support the chosen approach.

A database of tokamak edge plasma MHD stability simulations will be established using the CSC high-performance computing resources. The equilibrium solver HELENA [6] will be used to generate equilibrium data. The ideal MHD stability code MISHKA [7] and the resistive MHD stability code, CASTOR [12], will be used to simulate edge MHD instability growth. Once an initial database is established, a first proof-of-principle surrogate model will be generated with ML techniques.
Various ML approaches will be explored and compared for surrogate accuracy and computational throughput. Possibilities to inform the developed surrogate model with physics-informed constraints will be investigated, as such methods would reduce the amount of training data to reach the desired surrogate accuracy. Furthermore, a typical feature of these plasma stability simulations is a decision boundary in the operational space, where there is a stability threshold at which the instabilities start to grow. It is expected that building a neural network surrogate for the database directly will lead to blurring of this decision boundary and inaccurate predictions of the stability boundary. It is quite probably that the model will need to be split into categorizing component that predicts whether any instability will grow and a regression component that predicts the growth rates only in the case where the categorizing model predicts and instability to occur. However, there might also be physics-informed approaches for this challenge that circumvent the need for model splitting. Once a satisfactory accurate is achieved with the surrogate model, a demonstration of its capability to reproduce the MHD stability simulations in an EPED-like workflow will be conducted. This is followed by comparison to experimental observations from JET. The preference is to use JET experimental observation that are available within the EUROfusion pedestal database [Frassinetti_NF_2021].
To extend the coverage of the surrogate model training database, acquisition function based active learning approaches will be considered to reduce full model oversampling in regions where a good coverage is available already. Such approaches become extremely valuable when adding resistive MHD features with CASTOR to the surrogate model. These simulations are computationally significantly more demanding than ideal MHD simulations, while the role of resistive MHD is only important for a subset of the training database. The acquisition function in active learning workflows can data efficiently guide the database generation to prefer full CASTOR simulations to be conducted in regions where the resistive MHD features impact the predicted MHD stability.

These approaches are expected to lead to fast and accurate surrogate model for the edge MHD stability in tokamaks, which is expected to be widely adopted by the fusion research community.

Preliminary plan on the collection, usage and storage of the research material (1500 characters)

Briefly describe how you plan to collect the research material and use it. Are there any data protection or copyright issues related to data storage that need to be taken into account? Is it possible to make the data available for the use of other researchers?

The database will be generated using the equilibrium solver HELENA and MHD stability solvers MISHKA and CASTOR. Experimental data from JET will be used to test the surrogate model's predictive capabilities. The development of the model and generation of the database will be done on the CSC supercomputers. The generated simulation databases will be made publicly available as the associated studies are published. The experimental data from the JET tokamak is governed by EUROfusion. VTT is the representative entity of Finland in the EUROfusion consortium and the employees of VTT and linked third parties, such as University of Helsinki, can gain access to EUROfusion data through signing a user agreement. The applicant already has the necessary user account to access the data at JET. The review and release guidelines of the data are determined by EUROfusion, and the release of large-scale databases, such as the EUROfusion pedestal database, is discussed at a higher level outside this doctoral thesis work.

Ethical issues (1000 characters)

Are there ethical issues (e.g. ethical governance procedures, informed consent, and anonymity of subjects) that need to be taken into account when conducting the research? Does conducting the research require a research permit or a permit from the ethical board and/or the Animal Experiment Board?

There are no ethical issues.

The most important literature for the research plan (2500 characters)

Reference list

[1] L. Frassinetti, et al. Nucl. Fusion 61, 016001 (2021).

[2] E. Stefanikova, PhD Thesis, KTH, 2020.

[3] P. B. Snyder et al., Phys. Plasmas 16, 056118 (2009).

[4] S. Saarelma et al., Plasma Phys. Control. Fusion 60, 014042 (2018).

[5] M. G. Dunne et al., Plasma Phys. Control. Fusion 59, 025010 (2017).

[6] T. Luda et al., Nucl. Fusion 60, 036023 (2020).

[7] G. Husymans et al., Conf. Comp. Plasma Physics 371 (1991).

[8] A. B. Mikhailovskii et al., Plasma Phys. Rep. 23 844 (1997).

[9] O. Meneghini et al., Nuclear Fusion, 57, 086034 (2017).

[10] A. Panera Alvarez, et al. Manuscript to be submitted to Phys. Plasmas 2024.

[11]

[12] Kerner, W., Goedbloed, J. P., Huysmans, G. T. A., Poedts, S., & Schwarz, E. (1998). CASTOR: Normal-Mode Analysis of Resistive MHD Plasmas. In *JOURNAL OF COMPUTATIONAL PHYSICS* (Vol. 142).

Publication plan and timetable for the articles or Preliminary outline of the monograph (1500 characters)

The following articles are planned:

1. Proof of principal for surrogate model for the ideal MHD instability code MISHKA focused on the parameter ranges of JET tokamak, Plasma physics and controlled physics / Physics of Plasmas, target submission 2024.
2. Using the surrogate model for large scale analysis of JET tokamak plasmas, Nuclear fusion, target submission 2025.
3. Exploring extension to resistive MHD with CASTOR, Plasma physics and controlled fusion, 2026.
4. Testing of the model for JET plasma in domains with known resistive MHD features, Nuclear fusion, 2027.

Preliminary timetable for your research (2000 characters)

November 2023: Started working at VTT.

Beginning of 2024: Work with HELENA & MISHKA, generate database and preliminary machine learning models

2024: Write first article about the proof of principle.

2025: Second article.

2026: Third article.

2027: Fourth article and dissertation.

Planned funding for the research project including received and applied notable funding thus far (1000 characters)

Employment at VTT in the Fusion energy and decommissioning team.