Thesis for The Degree of Licentiate of Engineering

Machine learning applications for predicting the pedestal in tokamak plasmas

Andreas Gillgren

Department of Space, Earth and Environment
Chalmers University of Technology
Gothenburg, Sweden, 2023

Machine learning applications for predicting the pedestal in tokamak plasmas

Andreas Gillgren

Department of Space, Earth and Environment
Division of Astronomy and Plasma Physics
Chalmers University of Technology
SE-412 96 Göteborg,
Sweden
Phone: +46(0)31 772 1000

Printed by Chalmers Digitaltryck,
Gothenburg, Sweden 2023.

Machine learning applications for predicting the pedestal in tokamak plasmas

Andreas Gillgren

Department of Space, Earth and Environment
Chalmers University of Technology

Abstract

Magnetic confinement fusion is a field of research that strives to develop an environmental friendly energy source to assist in powering our society. By confining a plasma with magnetic fields, conditions that enable nuclear fusion can be achieved. However, gaining a high efficiency has proven to be a challenging task. In the 1980s, it was discovered that steep temperature and density gradients are formed near the plasma edge when the external heating passes a certain threshold leading to an increased energy and particle confinement. The region with steep gradients at the edge is referred to as the pedestal. As of today the formation and behaviour of the pedestal is still not fully understood from a theoretical standpoint. However, the enhanced performance of plasmas with a developed pedestal is routinely exploited in current fusion experiments, and is a key element in extrapolating to future devices.

The purpose of this thesis is to explore machine learning methodologies to help improve the understanding and predictive capabilities of the pedestal. Specifically, a neural network for predicting pedestal characteristics has been developed and integrated with core transport models. Additionally, another neural network has been developed to enhance the temporal resolution of the main diagnostics used to analyse the pedestal. The thesis incorporates additional machine learning applications for plasma physics that extend beyond a specific focus on the pedestal.

Keywords

Magnetic confinement fusion, Tokamak, Pedestal, Machine learning, Neural Networks

List of Publications

Appended publications

This thesis is based on the following publications:

[**Paper I**] **A. Gillgren**, E. Fransson, D. Yadykin, L. Frassinetti, P. Strand and JET contributors, *Enabling adaptive pedestals in predictive transport simulations using neural networks*
Nuclear Fusion 62 (2022).

[**Paper II**] E. Fransson, **A. Gillgren**, A. Ho, J. Borsander, O. Lindberg, W. Rieck, M. Åqvist and P. Strand, *A fast neural network surrogate model for the eigenvalues of QuaLiKiz*
Accepted in Physics of Plasmas.

[**Paper III**] D. R. Ferreira, **A. Gillgren**, A. Ludvig-Osipov, P. Strand and JET contributors, *High temporal resolution of pedestal dynamics via machine learning on density diagnostics*
Accepted in Plasma Physics and Controlled Fusion.

Other publications

The following publications were published during my PhD studies, or are currently in submission/under revision. However, they are not appended to this thesis, due to contents overlapping that of appended publications or contents not related to the thesis.

[**a**] C. J. Ham, A. Bokshi, D. Brunetti, G. Bustos-Ramirez, B. Chapman, J. W. Connor, D. Dickinson, A. R. Field, L. Frassinetti, **A. Gillgren**, *Towards understanding reactor relevant tokamak pedestals* *Nuclear fusion 61 (2021).*

Acknowledgment

First, I would like to thank my supervisors Pär Strand and Dmytro Yadykin for having supported me and for having created a research environment that stimulates creativity. I would also like to express my gratitude to my colleagues Emil Fransson and Andrei Osipov, who have served as important mentors.

More than anything, I would like to thank my wife Katarina for being the best possible companion and for making life fun. I would also like to thank my parents Gunilla and Mårten, and the rest of my family for unconditionally having supported me in whatever life choices I have made.

Finally, I would like to acknowledge that part of the work presented in this thesis has been supported under the Vetenskapsrådet (VR) project grant 2020-05465 and the EUROfusion project ENR-MOD.01.FZJ "Development of machine learning methods and integration of surrogate model predictor schemes for plasma-exhaust and PWI in fusion".

Contents

Abstract i

List of Publications iii

Acknowledgement v

I Summary 1

1 Introduction 3
- 1.1 Fusion . . . 3
  - 1.1.1 Fusion on Earth . . . 4
  - 1.1.2 Potential benefits with fusion . . . 6
- 1.2 Machine learning . . . 7
- 1.3 Objectives of thesis . . . 8

2 An overview of plasma physics 9
- 2.1 Definition of a plasma . . . 9
- 2.2 Theoretical descriptions of plasmas . . . 10
  - 2.2.1 Particle motion in electric and magnetic fields . . . 10
  - 2.2.2 Kinetic description of plasmas . . . 13
  - 2.2.3 Two-fluid theory . . . 15
  - 2.2.4 Magnetohydrodynamics . . . 16
- 2.3 Perturbations, linear models, nonlinear models, and plasma instabilities . . . 17
- 2.4 Gyro-average models . . . 18
- 2.5 Collisions . . . 18

3 The tokamak 21
- 3.1 Power plant concept . . . 21
- 3.2 Plasma geometry . . . 21
- 3.3 Reactor chamber and gas fueling . . . 24
- 3.4 Plasma heating . . . 25
  - 3.4.1 Ohmic heating . . . 26
  - 3.4.2 Neutral beam injection (NBI) . . . 26
  - 3.4.3 Radio frequency (RF) heating . . . 26

3.5 Plasma diagnostics . . . 27
3.5.1 High resolution thomson scattering (HRTS) . . . 27
3.5.2 Reflectometry . . . 27
3.6 Plasma profiles . . . 28
3.6.1 Two-dimensional profiles . . . 28
3.6.2 One-dimensional profiles . . . 29
3.7 Heat and particle transport in a tokamak . . . 30
3.7.1 Classical transport . . . 30
3.7.2 Neoclassical transport . . . 31
3.7.3 Turbulent transport . . . 31
3.8 Integrated modelling . . . 32
3.9 The pedestal . . . 32
3.9.1 Edge localized mode types . . . 34
3.9.2 Pedestal modelling . . . 34
3.9.3 Empirical pedestal scalings . . . 35

4 Machine learning fundamentals 37
4.1 The neural network node . . . 37
4.2 Example: Linear single-node model . . . 38
4.2.1 The loss and cost function . . . 38
4.2.2 Optimization . . . 39
4.2.3 Training procedure and result . . . 40
4.3 Dense neural networks . . . 40
4.4 Nonlinear activation functions . . . 41
4.5 Hyperparameters . . . 42
4.6 Training, validation, and testing . . . 43
4.7 Classification models . . . 44
4.8 Ensemble learning . . . 44
4.9 Surrogate models . . . 44
4.10 Interpretability and the black-box problem . . . 45

5 Summary of Appended Papers 47
5.1 Enabling adaptive pedestals in predictive transport simulations using neural networks . . . 47
5.2 A fast neural network surrogate model for the eigenvalues of QuaLiKiz . . . 48
5.3 High temporal resolution of pedestal dynamics via machine learning 48

6 Future Work 51

Bibliography 53

II Appended Papers 59

Paper I - Enabling adaptive pedestals in predictive transport simulations using neural networks

Paper II - A fast neural network surrogate model for the eigenvalues of QuaLiKiz

Paper III - High temporal resolution of pedestal dynamics via machine learning on density diagnostics

Part I

Summary

Chapter 1

Introduction

In the ever-evolving saga of human progress, both the narratives of energy and artificial intelligence (AI) currently stand at the forefront. The dream of emulating the energy from the stars, coupled with the ascent of AI, has lead to the rapid development of an interdisciplinary field: the application of machine learning in fusion technology.

1.1 Fusion

Fusion energy is the energy released in the process when two atomic nuclei are merged into a heavier element. This is the counterpart of fission, where heavy nuclei are split into lighter elements. For instance, the hydrogen isotopes deuterium (D) and tritium (T) can fuse into a helium-4 (He^4) nucleus, producing an additional neutron (n) in the process

$$D + T \longrightarrow He^4 + n + \text{energy}. \tag{1.1}$$

Although the helium-4 nucleus is heavier than the deuterium and tritium nuclei respectively, the sum of the mass of the deuterium and tritium nuclei is larger than the sum of the mass of the helium-4 nucleus and the neutron [1]. This is due to the difference in binding energies of the different nuclei [2], and the loss of mass m is converted to kinetic energy E according to Einstein's equation

$$E = mc^2, \tag{1.2}$$

where c is the speed of light in vacuum (299 792 458 m/s) [1]. Since c is a large number, even a small change of mass leads to a large amount of energy. The energy release in chemical reactions is also due to this principle. However, chemical reactions are associated with the rearrangement of bonds between outer shell electrons and nuclei. These bonds are carried by the electromagnetic force, which is significantly weaker than the strong nuclear force binding nuclei [3]. Therefore, a smaller change in binding energy in chemical reactions leads to a smaller change in mass and energy in (1.2). As an illustrative example, a

kilogram of fusion fuel can yield nearly four million times more energy compared to the combustion of a kilogram of coal or oil [4].

Nuclear reactions are based on probability rooted in quantum theory [2], which means that we can never be certain if two individual nuclei will fuse. However, when considering a macroscopic scale involving numerous nuclei, the cumulative probability manifests itself in an observable reaction rate [3]. Fundamentally, the probability of a fusion reaction occurring is increased when two nuclei are in a close proximity for an extended duration. Nevertheless, two positively charged nuclei repel each other according to the Coulomb repulsion [5], and to overcome this barrier, the nuclei must have sufficient kinetic energy. However, too high kinetic energy will also lower the probability of a fusion reaction occurring [3], since the nuclei will pass each other more rapidly, which reduces the time that the nuclei are in a close proximity.

On a macroscopic scale, temperature acts as a measurable quantity for the average kinetic energy of the particles in a system [6]. Therefore, there is an optimal temperature that maximizes the rate of fusion reactions. Due to properties derived from quantum theory, the optimal temperature varies between different pairs of nuclei [3]. This is illustrated in Figure 1.1, where the cross-section σ [3], which is proportional to the reaction probability, is plotted as a function of temperature for different fusion reactions. It is also shown how much energy is converted to kinetic energy in each reaction in mega electron volts (MeV).

It is however not only the temperature that affects the rate of fusion reactions in a system. The density plays a vital part since a more densely populated volume means that a nucleus will cross paths with more nuclei along its trajectory, which increases the probability of a reaction.

In the sun, and other stars, the tremendous gravitational pull leads to sufficiently high temperatures and densities in the core to enable fusion, which powers our solar system and illuminates the universe. At these high energy conditions, atoms are in a plasma state, where electrons have been stripped from their nuclei, resulting in a mixture of positively charged ions and free negatively charged electrons. On Earth, it is unrealistic to use gravity to confine particles and achieve fusion. Therefore, we need to find other ways if we wish to exploit this feature of nature in a controlled manner.

1.1.1 Fusion on Earth

The two main branches of fusion power on Earth are magnetic confinement fusion (MCF) [4] and inertial confinement fusion (ICF) [8]. In MCF, the confinement is achieved by utilizing magnetic fields that pass through a chamber, leveraging the principles of the Lorentz force [5], which dictates that charged particles are confined in a gyrating motion along the magnetic field lines. As in the sun, the particles in a MCF device are heated to a plasma state to enable the conditions for fusion. This technology has been developed since the 1950s, including configurations such as the mirror machine, the tokamak, and the stellarator. As of today, the two latter are the leading candidates. For example, ITER [9] is currently under construction as the largest tokamak to

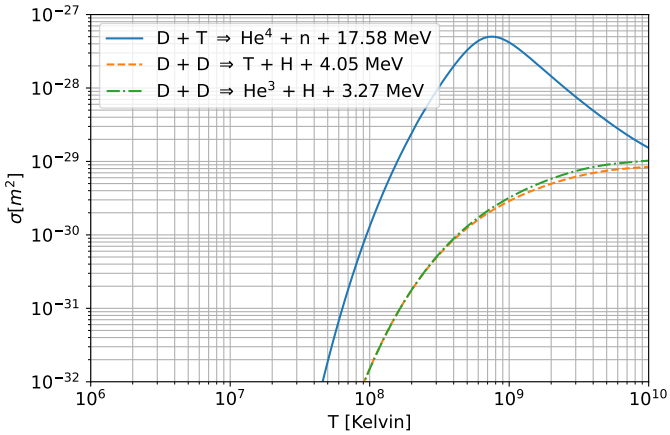

Figure 1.1: The cross section σ of different reactions, which is related to fusion rate, or reaction probability, versus temperature. To achieve a significant rate, temperatures exceeding tens of millions of degrees Kelvin are necessary. The DT reaction provides two advantages compared to the DD reaction as its cross-section peaks at a lower temperature, which would be easier to achieve on Earth, and more energy is converted per reaction. Additionally, out of the two DD reactions, the graph indicates that if a DD reaction occurs, there is approximately a 50% chance to produce a T nucleus, and approximately a 50% chance to produce a He^3 nucleus. Here, H refers to a hydrogen nucleus without neutrons, which is simply a proton. It is noteworthy to mention that there are other possible reactions for fusion. The reactions shown here are the ones that require the least amount of energy. The data in this graph was gathered from the International Atomic Energy Agency (IAEA) Nuclear Data Section [7].

date in Cadarache, France. This project is a collaborative effort involving the European Union, the United States, China, the Russian Federation, Japan, South Korea, India. As all of the work presented in this thesis is related to the tokamak, no further description of mirror machines and stellarators is presented in the context of MCF in this thesis. More information about these concepts can be found in [4].

In ICF, a pellet containing the fusion fuel is placed in a chamber. The pellet is compressed and heated by external laser beams to provide sufficient conditions for fusion. This technology has been developed since the 1970s, and the largest operational ICF experiment to date is the National Ignition Facility (NIF) [10] in the United States.

Both of these technologies are still at the experimental stage, and several challenges must be solved before fusion power can provide more electricity to the grid compared to the electricity needed to operate the devices. An important aspect of this research is performing simulations, such as those for the plasma in a MCF device. Simulations play a crucial role in minimizing the

overall cost and time of fusion research, and there is a need to explore innovative approaches that enhance the speed and accuracy of these simulations.

Both MCF and ICF are expected to use deuterium and tritium as the main fuel in a commercial power plant [4], [10]. This is due to accessibility and reaction properties compared to other possible fusion reactions. Deuterium [11] is not radioactive, and easily accessible and abundant since it can be extracted from water. For instance, relying solely on deuterium-deuterium (DD) fusion reactions to meet the current annual energy demand could potentially provide a supply lasting for billions of years [11]. Tritium however, is radioactive with a half-life of 12.32 years. For this reason, it is not abundant in nature, although it can be produced by irradiation of lithium (Li) with fast neutrons [4]. Lithium is a relatively abundant element in the Earth's crust, but as the demand for lithium for use in batteries for electric vehicles and portable electronics continues to rise, efforts to explore and develop new lithium deposits are also ongoing [12]. In principle, it is possible to only use deuterium as the fuel for fusion, which is a long-term goal. However as illustrated in Figure 1.1, this puts a higher demand on the technology to achieve appropriate conditions for fusion. Nevertheless, Figure 1.1 also illustrates a 50% probability of generating a tritium nucleus in a DD reaction. This tritium nucleus can later fuse with deuterium nuclei which increases the efficiency compered to if only DD reactions occur. The enhanced efficiency is however dependent on the burn-up fraction, which is the percentage of fuel that undergoes fusion. In tokamaks, this is usually a few percent [13], and a key challenge for future machines is to increase this rate.

1.1.2 Potential benefits with fusion

The process of fusion as an energy source is intrinsically environmental friendly and sustainable, since no greenhouse gasses are produced in the reaction, and the fuel is accessible as discussed in the previous subsection. This of course assumes environmental friendly transport of fuel and construction of facilities. Nevertheless, since each fusion reaction converts significantly more energy than chemical reactions in coal, oil, and biogas facilities, less fuel needs to be transported for the same amount of converted energy.

A fusion reactor will produce some radioactive byproducts through neutron activation of reactor materials, although the hazard differs significantly compared to fission facilities. The most common fission fuels, uranium-235 and plutonium-239, produce highly radioactive products that remain hazardous for thousands of years, which is not the case for fusion [1], [11]. Future fusion reactor designs will include a lithium based blanket for tritium breeding in the reactor chamber [4], which eliminates the need to transport or store any radioactive material. The risk of a nuclear meltdown, for instance due to a natural disaster, is also not a concern in fusion since fusion reactions are highly dependent on specific conditions that are challenging to maintain consistently. In practice, any major disturbance will immediately terminate the process in relevant designs [4]. At any given time, a reactor contains only a minimal quantity of active materials, such as tritium, ensuring that any

potential environmental impact remains localized to the reactor.

Fusion, much like fission and combustion, holds the advantage of not being contingent on specific weather conditions or location for its operation—unlike solar, wind, and hydropower. It is however crucial to emphasize that the goal of fusion is not necessarily to replace other renewables like solar and wind power. Instead, fusion aims to complement existing sustainable energy efforts to contribute to a more diverse and resilient energy landscape.

1.2 Machine learning

On another scientific front, machine learning and AI are advancing with a tremendous momentum. From pioneering handwritten digit recognition in the 1990s [14], [15] to surpassing the world champion in Go during the 2010s [16] and creating imaginative images and texts with models like ChatGPT in the 2020s [17], [18], AI researchers have achieved remarkable milestones. Although AI is often mentioned in the context of these popular applications, it has also proven to assist in other research disciplines, such as detection of diseases [19], [20], environmental modelling [21], [22], but also fusion [23], [24].

AI and machine learning are terms that often are used to describe the same thing. However, machine learning is a subset of AI that involves the development of algorithms that allow systems to learn from data, identify patterns, and make decisions or predictions without being explicitly programmed for each scenario. A neural network is an example of a machine learning model which we will explore in detail in Chapter 4.

However, there are not only benefits associated with the development of machine learning algorithms. As with many inventions and technologies, machine learning can be utilized in harmful and unethical ways, and can potentially pose a threat to human civilisation as we know it. Most machine learning models also pose challenges in terms of interpretability, lacking transparency in explaining the rationale behind a decision or prediction, which is a topic that is becoming more relevant as AI technology progresses [25], [26]. Nevertheless, there is a big difference in the ethical risk between, for instance, making AI-based judgements in court, and assisting in the development of a clean energy source. There is also a big difference between the current models and Artificial General Intelligence (AGI), which is a hypothetical system with general cognitive abilities comparable to those of humans, which is one of the main concerns related to the future of AI [27].

In essence, the application of current machine learning algorithms in fusion present minimal risks in relation to the primary concerns associated with artificial intelligence [28]. While interpretability remains a challenge, it is noteworthy that models lacking complete transparency can still offer valuable applications.

1.3 Objectives of thesis

This thesis delves into the applications of machine learning in the field of fusion. Specifically, the goal is to examine various applications related to a specific region in the plasma in a tokamak referred to as the pedestal. The appended papers encompasses ways in which machine learning applications related to the pedestal can accelerate simulations and contribute to the analysis of reactor diagnostics.

In the next chapter, an overview of the theoretical descriptions of plasmas is presented. Chapter 3 delves into the key aspects of the tokamak that hold the most relevance to the context of the scientific papers appended in this thesis. In the end of Chapter 3, an introduction of the pedestal is outlined. Moving to Chapter 4, the fundamentals of machine learning are presented. Chapter 5 summarizes the appended papers, and finally in Chapter 6, avenues for future research are discussed.

Chapter 2

An overview of plasma physics

As mentioned in the introduction, to create conditions suitable for fusion, it is necessary to heat the particles enough, resulting in the formation of a plasma. Therefore, almost all aspects of fusion research, in particular for magnetic confinement fusion, are related to how plasma behaves for the relevant conditions. In this section, an overview of plasma physics based on the contents in [4] is presented.

2.1 Definition of a plasma

A simple description of a plasma is that, due to high energy in a system of particles, the electrons are no longer bounded to the nuclei, which results in a collection of free electrons and ions. A more thorough definition is that plasma is a quasi-neutral gas consisting of charged and neutral particles that show collective behaviour. Quasi-neutrality means that the plasma is neutral in charge on a macroscopic scale but occasionally deviate from neutrality on the microscopic level. However, while these microscopic deviations produce electrical fields that operate on a relatively small spatial scale, they remain long-range compared to the collisional forces between two neutral particles in a conventional gas. Therefore, in contrast to pair-wise collisions in a conventional gas, a particle in a plasma interacts with many particles simultaneously, hence the term 'collective behaviour'.

One of the main properties of a plasma is the ability to quickly screen out changes in electric potential, which is why deviations in quasi-neutrality occur on a microscopic scale. For instance, if a positive charge q was to be placed in an otherwise neutral plasma, electrons will rapidly move towards the charge, since they are much lighter than ions and can accelerate faster. They will then arrange such that outside a sphere with the charge q in the center, which is called the Debye sphere, the influence from the charge q becomes negligible. For fusion plasmas, the radius of this sphere, which is called the Debye length

λ_D, is on the order of 7×10^{-6} m. The calculation of the Debye length is based on the assumption that the electron density follows a Boltzmann distribution with respect to the electric potential. Since this is a statistical assumption, the number of particles in the Debye sphere must be large for the assumption to hold. In a bigger picture, since the screening ability is a key property of a plasma, we can now formulate conditions for the plasma definition to hold

- The macroscopic spatial dimension of the plasma must be much larger than the Debye-sphere.
- The particle density must be sufficiently high such that many particles populate the Debye sphere.
- Additionally, the frequency of collisions between plasma particles and neutral particles cannot be too high, such that it disrupts the plasma dynamics governed by the electromagnetic force.

2.2 Theoretical descriptions of plasmas

Although the fusion process itself is a quantum mechanical phenomenon, plasma physics is governed by classical mechanics. In principle, a plasma can be accurately described with the equation of motion for each individual particle in the plasma, together with a self-consistent set of Maxwell's equations for the electromagnetic fields. In practice however, finding a solution to this is not possible, even with supercomputers, due to the large number of particles and the inherent complexity. Additionally, the inability to determine the initial conditions for every individual particle further complicates the challenge. This has led to the development of different statistical approaches to describe a plasma, for which we will summarize in the following sections. However, we will start by considering the motion of charged particles in electric and magnetic fields, that are assumed to be known, since this can provide valuable information in addition to the approaches we will discuss later.

2.2.1 Particle motion in electric and magnetic fields

Consider a particle with the charge q in an electric field $\vec{E}$ and a magnetic field $\vec{B}$. The acceleration $\vec{a}$ of the particle is governed by the Lorentz force

$$m\vec{a} = \vec{F} = q(\vec{E} + \vec{v} \times \vec{B}), \tag{2.1}$$

where m is the mass of the particle, and where $\vec{v}$ is the velocity of the particle. Now assume a scenario where $\vec{E} = 0$ such that

$$\vec{a} = \frac{q}{m}(\vec{v} \times \vec{B}). \tag{2.2}$$

If $\vec{B}$ is uniform and constant, then the solution to (2.2) is a combination of a circular particle motion, which is referred to as the gyro-motion, perpendicular to $\vec{B}$, and a constant velocity parallel to $\vec{B}$. This is illustrated in Figure 2.1,

where the center of the gyro-motion is referred to as the guiding center, or the gyro-center.

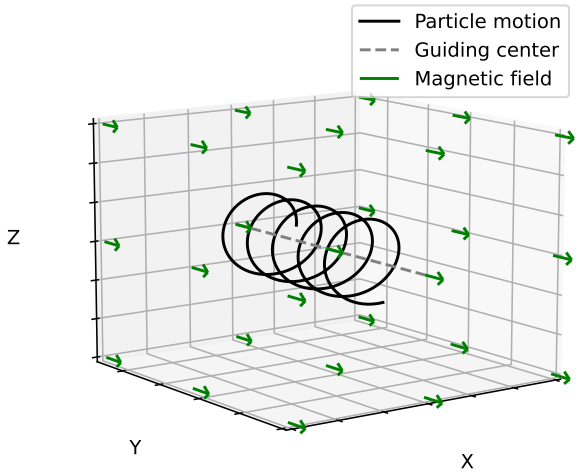

Figure 2.1: The motion of a charged particle in a uniform magnetic field (arrows). X, Y and Z represent spatial coordinates. The particle follows a circular motion around and along the guiding center, which is in the same direction as the magnetic field (Y-direction). Therefore, the particle is confined to gyrate around the field lines.

The solution of (2.2) also provides the gyro-radius, which is called the Larmor radius

$$r_L = \frac{mv_\perp}{|q||\vec{B}|}, \tag{2.3}$$

where $v_\perp$ is the perpendicular velocity of the particle with respect to $\vec{B}$. Equation (2.3) indicates that ions have a larger Larmor radius compared to electrons since their mass is larger. The gyro-frequency ω, in units radians/seconds, can also be obtained through

$$\omega = \frac{v_\perp}{r_L} = \frac{|q||\vec{B}|}{m}. \tag{2.4}$$

Now consider a scenario where we add an arbitrary force $\vec{F}$ to the equation of motion (2.2), such that

$$\vec{a} = \frac{q}{m}(\vec{v} \times \vec{B}) + \frac{\vec{F}}{m}. \tag{2.5}$$

If $\vec{F}$ has a parallel component with respect to $\vec{B}$, this will lead to acceleration of the particle along the magnetic field lines. The orthogonal component of the force $\vec{F_\perp}$ will lead to a drift away from the magnetic field lines, where this particle drift is orthogonal to both $\vec{F_\perp}$ and $\vec{B}$. This can be derived by introducing a constant drift velocity $\vec{v_D}$ such that

$$\vec{v_\perp} = \vec{v_D} + \vec{u}. \tag{2.6}$$

By inserting (2.6) into the orthogonal part of (2.5), the solution will show that $\vec{u}$ represents the gyro-motion, or Larmor motion, and that

$$\vec{v_D} = \frac{1}{q}\frac{\vec{F_\perp} \times \vec{B}}{B^2}. \tag{2.7}$$

For instance, an electrical field $\vec{E}$ orthogonal to $\vec{B}$ will generate a force $\vec{F} = q\vec{E}$, which will lead to the drift

$$\vec{v_D} = \frac{\vec{E} \times \vec{B}}{B^2} = \vec{v}_{E\times B}. \tag{2.8}$$

Note that this $\vec{E} \times \vec{B}$ drift is independent of mass and charge, which means that ions and electrons will drift in the same direction with the same velocity. The $\vec{E} \times \vec{B}$ drift is illustrated in Figure 2.2.

The $\vec{E} \times \vec{B}$ drift is not the only possible drift. By replacing the arbitrary force $\vec{F}$ in (2.5) with, for instance, the centripetal force in a curved magnetic field, we obtain the curvature drift

$$\vec{v_c} = \frac{mv_\parallel^2}{qB^2}\frac{\vec{R}_c \times \vec{B}}{R_c^2}, \tag{2.9}$$

where $v_\parallel$ is the particle velocity parallel to the magnetic field, and where $\vec{R}_c$ is the curvature radius vector. Additionally, a curved magnetic field will not satisfy Maxwell's equations if $\vec{B}$ is homogeneous, which implies that there is a nonzero gradient ∇B. Therefore, over the course of one gyration, a simplified description is that the particle will experience a stronger magnetic field for one half of the gyration compared to the other half of the gyration. Since the strength of the magnetic field affects the Larmor radius (2.3), the radius of one half of the gyration will be smaller compared to the other half of the gyration, which unfolds as a drift. This drift, which is called the ∇B drift, or grad B drift, is illustrated in Figure 2.3.

A derivation of the ∇B drift, which is more complicated compared to the other mentioned drifts, can be found in [4], which yields the result

$$\vec{v}_{\nabla B} = \frac{\mu}{q}\frac{\vec{B} \times \nabla B}{B^2}, \tag{2.10}$$

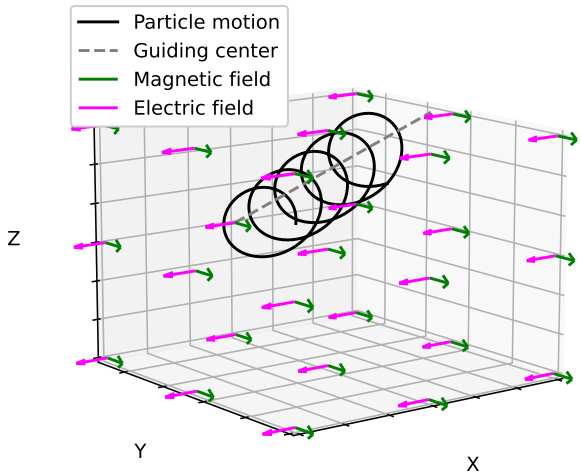

Figure 2.2: The motion of a charged particle in an uniform magnetic field (green arrows in Y-direction) and an uniform electric field (magenta arrows in X-direction) orthogonal to the magnetic field. The particle follows a circular motion, but in contrast to the case illustrated in Figure 2.1, the guiding center, which travels in the Y-direction, now also drifts in an orthogonal direction (Z-direction) compared to both the electric field and the magnetic field as a consequence of the $\vec{E} \times \vec{B}$ drift.

where μ is the magnetic moment of the charged particle

$$\mu = \frac{m v_\perp^2}{2|\vec{B}|}, \tag{2.11}$$

which is a conserved quantity. In summary, understanding these drifts is essential for magnetic confinement fusion, which we will further explore in the chapter about tokamaks.

2.2.2 Kinetic description of plasmas

We will now look at statistical approaches to describe a plasma instead of describing the exact position and velocity of every single particle. In the kinetic description, this is done by defining a distribution function $f(\vec{r}, \vec{v}, t)$ that describes the probability of finding a particle at the position $(\vec{r})$ with the velocity $(\vec{v})$ at some time t. For instance, particle density in real space

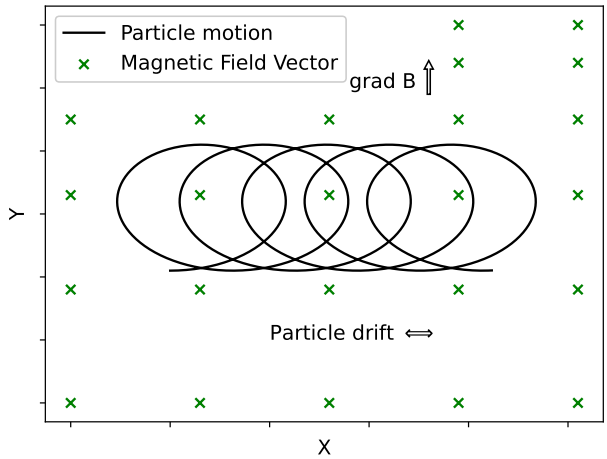

Figure 2.3: Illustration of the grad B drift. The magnetic field lines that are orthogonal to the XY-plane, illustrated by crosses, are closer to each other in the upper part of the figure to illustrate that the magnetic field is stronger for higher Y. This leads to a larger Larmor radius for lower Y according to (2.3), resulting in a drift in the X-direction. The direction of the particle drift depends on the sign of the charge of the particle, as indicated by (2.10), which means that ions and electrons drift in opposite directions for this drift.

$n(\vec{r}, t)$, which is a measurable quantity, can be obtained by integrating f over the velocity space

$$n(\vec{r}, t) = \int d^3v f(\vec{r}, \vec{v}, t). \tag{2.12}$$

Similarly, the mean velocity of the particles $\vec{u}(\vec{r}, t)$ can be defined as

$$\vec{u}(\vec{r}, t) = \frac{1}{n} \int d^3v \vec{v} f(\vec{r}, \vec{v}, t). \tag{2.13}$$

Both n and $\vec{u}$ are called velocity moments of f, where n is the zeroth-order moment (f is multiplied with 1, and where $\vec{u}$ is a first-order moment (f is multiplied with $\vec{v}^1$). Similarly to the equation of motion for individual particles, the kinetic equation determines the distribution function f. A derivation if the kinetic equation can be seen in [4]. If collisions between particles are neglected, the kinetic equation is described by the Vlasov equation

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} + \frac{q}{m}(\vec{E} + \vec{v} \times \vec{B}) \cdot \frac{\partial f}{\partial \vec{v}} = 0, \tag{2.14}$$

which applies separately for different species of particles. Here, the acceleration $\vec{a}$ of the particles have been assumed to only be dependent on the Lorentz force.

To include collisional effects, a collision operator is added to the right-hand-side of (2.14), which yields the Boltzmann equation

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{r}} + \vec{a} \cdot \frac{\partial f}{\partial \vec{v}} = (\frac{\partial f}{\partial t})_c. \tag{2.15}$$

The collision operator describes how f changes due to collisions, and there is no exact version of it. In summary, the research related to kinetic theory in a plasma is related to modelling the Vlasov and Boltzmann equation with different collision operators, and solving them with numerical methods to obtain f, which then can be used to quantify measurable quantities such as the particle density. Kinetic models are the highest fidelity programming codes that can be implemented in practice to make predictions of the plasma in a reactor. However, due to the high dimensionality of the theory, they are generally computationally heavy, since both the velocity space and real space need to be solved for.

2.2.3 Two-fluid theory

One approach to reduce the dimensionality of kinetic theory is to treat the plasma as a composite of fluids. Instead of solving equations for the distribution function f as in the previous case, fluid theory strives to solve for the macroscopic quantities directly, such as the particle density n (2.12) and mean velocity $\vec{u}$ (2.13). Therefore, it is not necessary to solve for the velocity space, which reduces the dimensionality by 3.

As mentioned previously, the macroscopic quantities are related to moments of the distribution function, and a moment $\langle\psi\rangle$ is defined as the velocity average of the function $\psi(\vec{v})$

$$\langle\psi\rangle = \frac{1}{n}\int d^3v\psi(\vec{v})f, \tag{2.16}$$

where the zeroth-order moment is obtained by setting $\psi = 1$, and where the first-order and second-order moments correspond to $\psi = \vec{v}$ and $\psi = \vec{v}\vec{v}$ respectively. The general moment equation is obtained by integrating the Boltzmann equation from kinetic theory with respect to velocity, which yields

$$\frac{\partial}{\partial t}(n\langle\psi\rangle) + \nabla \cdot (n\langle\vec{v}\psi\rangle) - \frac{nq}{m}\left\langle (\vec{E} + \vec{v} \times \vec{B}) \cdot \frac{\partial \psi}{\partial \vec{v}} \right\rangle = \frac{\partial}{\partial t}(n\langle\psi\rangle)_c. \tag{2.17}$$

For instance, the zeroth order ($\psi = 1$) moment equation is the continuity equation [4]

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\vec{u}) = 0, \tag{2.18}$$

which includes the particle density n and the mean velocity $\vec{u}$. Other macroscopic quantities, such as kinetic pressure (temperature) and heat flux, can be obtained through higher order moment equations. However, a challenge with solving moment equations is that each equation includes the next higher-order

moment, which can be seen in the second term in both (2.17) and (2.18). For instance, the continuity equation, which is the zeroth-order moment equation, includes the mean velocity $\vec{u}$, which is a first-order moment. This creates an infinite chain of dependent equations, which must be broken through assumptions and approximations, such that at some point a moment equation is not dependent on the next-order moment. Such approximations are referred to as moment closure since they lead to a finite set of equations to solve. It is also important to note that each particle species has their own set of moment equations since they are treated as separate fluids.

Similarly to the kinetic theory case, researchers model plasmas on computers with fluid theory and experiment with different closures and numerical methods to solve the moment equations. Fluid theory is not of equally high fidelity compared to kinetic theory, since it introduces more assumptions. However, it is less computationally expensive due to the dimensionality reduction.

2.2.4 Magnetohydrodynamics

Magnetohydrodynamics (MHD) is a theoretical description that modifies fluid theory such that the plasma can be treated as one fluid. For instance, the plasma can be characterized by parameters like the mass density ρ_m and the centre-of-mass velocity $\vec{V}$, which are defined as

$$\rho_m := \sum_\alpha n_\alpha m_\alpha, \tag{2.19}$$

$$\vec{V} := \frac{1}{\rho_m} \sum_\alpha n_\alpha m_\alpha \vec{u}_\alpha, \tag{2.20}$$

where the summation over α refers to the sum over the different particle species, and where $\vec{u}_\alpha$ again is the mean velocity, in this case for species α. The MHD equations can be obtained by adjusting the moment equations from two-fluid theory according to these new MHD parameters. For instance, the first MHD equation can be derived by multiplying the continuity equation with the mass, and by adding the equations for the different species

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \vec{V}) = 0, \tag{2.21}$$

which reflects the conservation of mass. Other MHD equations can be obtained by modifying higher order moment equations with MHD parameters and combining them with Maxwell's laws. Inherently, the equations in a specific MHD model depend on the closure that is used in the two-fluid model that the MHD model is based on. Moreover, two distinct formulations of MHD exist: ideal MHD and resistive MHD. The former treats the plasma as a perfect conductor, while the latter accounts for resistive effects, which affects the MHD equations.

As with the previous cases, MHD equations can be solved with numerical methods for different scenarios. These models are generally computationally cheaper compared to their two-fluid model counterparts, since fewer parameters

are needed to be solved for. Additionally, resistive MHD models are generally more computationally expensive compared to their ideal MHD model counterparts, due to added complexity in the equations. Of course, the computational requirements of models vary based on several factors. For example, a 'linear' fluid model may entail lower computational costs compared to a 'nonlinear' MHD model. In other words, the general statements about computational demand are not universally applicable to all scenarios. In the next section, we will explore what is meant by a linear model and a nonlinear model.

2.3 Perturbations, linear models, nonlinear models, and plasma instabilities

The theoretical descriptions of plasmas discussed so far can be used, for instance, to calculate the steady state of a plasma. In practice, this means setting $\partial/\partial t = 0$ in the equations employed to characterize the plasma. However, small perturbations from the steady state can lead to waves and instabilities. These can be simulated by introducing perturbations in the plasma parameters when solving a chosen set of plasma equations numerically. Another approach is to analyse the effect of perturbations analytically. For instance, a plasma parameter $\vec{X}$ can be expanded by assuming that it is a sum of the steady state/background solution $\vec{X}_0$ and a series of perturbations such that

$$\vec{X} = \vec{X}_0 + \epsilon\vec{X}_1 + \epsilon^2\vec{X}_2 + \cdots, \tag{2.22}$$

where ϵ is a small parameter. Let us now consider the continuity equation (2.18) and insert (2.22) with first order perturbations (only $\vec{X}_0$ and $\vec{X}_1$), such that the continuity equation becomes

$$\frac{\partial n_0}{\partial t} + \epsilon\frac{\partial n_1}{\partial t} + \nabla\cdot(n_0\vec{u}_0 + \epsilon n_0\vec{u}_1 + \epsilon n_1\vec{u}_0 + \epsilon^2 n_1\vec{u}_1) = 0. \tag{2.23}$$

Here, the terms that are not multiplied with ϵ represent the steady state solution. We now wish to solve the ϵ^1 equation, where we discard higher order ϵ-terms. We also use that $\vec{u}_0 = 0$ since the average velocity vector of the particles is zero in steady state. This gives the result

$$\frac{\partial n_1}{\partial t} + \nabla\cdot(n_0\vec{u}_1) = 0, \tag{2.24}$$

where we see that the perturbations of the first order, n_1 and $\vec{u}_1$, have been linearized due to the exclusion of higher order ϵ-terms. Let us now, for example, look at monochromatic wave solutions of n_1 and $\vec{u}_1$ such that $\vec{X}_1 \sim \exp[i(\vec{k}\cdot\vec{r} - \omega t)]$, where ω is the angular frequency of the wave, and where $\vec{k}$ is the wave vector. We may then use that $\nabla \sim i\vec{k}$ and $\partial/\partial t \sim -i\omega$, such that (2.24) becomes

$$-i\omega n_1 + in_0\vec{k}\cdot\vec{u}_1 = 0, \tag{2.25}$$

which is called a dispersion relation. The same principle is applied to all equations in a coupled system, which leads to a set of linear equations that can be solved analytically or numerically. Specifically, we can solve for relations between ω and $\vec{k}$ that provide non-trivial solutions to $\vec{X}_1$. In other words, we seek to find the eigenvalues

$$\omega(\vec{k}) = \omega_r(\vec{k}) + i\gamma(\vec{k}). \tag{2.26}$$

The real part of ω is called the real frequency ω_r and the imaginary part of ω is called the growth rate γ. If γ is negative, the amplitude of the perturbation we assumed to be a wave attenuate, and if γ is positive, the amplitude of the wave grows over time. The latter case can be problematic as it may lead to growing instabilities in the plasma that are detrimental for the confinement. Therefore, researchers perform simulations of equations with perturbations to find the growth rate of instabilities as a function of the spatial scale dictated by $\vec{k}$ for different cases.

The difference between linear models and nonlinear models is that nonlinear terms, such as the ϵ^2-term in (2.24), are neglected in linear models. This works well when the associated perturbations are small. However, if there is a positive growth rate, the perturbation will eventually become large enough such that nonlinear terms can no longer be neglected. In reality this will lead to a saturation of the the amplitude of the perturbations. Nonlinear models are thus more accurate for such cases, but also more costly to solve numerically since the equations can no longer be solved using linear algebra.

2.4 Gyro-average models

As discussed in Section 2.2.1, a charged particle gyrates along a guiding center in the presence of a magnetic field. The gyro motion, dictated by the gyro frequency (2.4) and Larmor radius (2.3), occur on a much shorter timescale and spatial scale compared to other dynamics in the plasma for reactor relevant conditions. By gyro-averaging, a model can capture the average behavior of particles, which is of higher interest compared to the detailed gyro motion, while significantly reducing the computational complexity. This is commonly utilized in kinetic theory, and it can be used in the derivation of fluid equations.

2.5 Collisions

The concluding topic in this brief plasma overview is a description of collisions. As mentioned in the beginning of this chapter, charged particles in a plasma do not collide in a pair-wise manner as in a gas. They are rather deflected by the combined interaction with many particles due to the electromagnetic force, which typically occur on a length scale similar to the Debye sphere, since electric fields are screened out outside this sphere. The collision time is defined as the average time it takes for a particle to be deflected 90°, and it is proportional to $T^{3/2}$, where T is temperature, which means that collisions in a plasma occur less frequently at high temperatures.

In a plasma where particles are confined to applied magnetic field lines, collisions play a part in how energy and particles diffuse and are transported perpendicularly to the field lines, which we will discuss more in the next chapter.

Chapter 3

The tokamak

The tokamak [29] is the most developed reactor design for magnetic confinement fusion research, and given that the work presented in this thesis is linked to this reactor type, we will now delve into its key components and inherent features. Additionally, since the work in this thesis is specifically linked to the Joint European Torus (JET) tokamak in Culham, UK, most examples will be based on this device.

3.1 Power plant concept

The goal of a future power plan based on the tokamak is to create plasma conditions such that fusion reactions occur at a high rate, and to confine those conditions for sufficiently long times. This is often expressed in terms of obtaining a high engineering Q-value, which is the ratio between the total electrical power output and the power required to operate the reactor. In the fusion reactions, neutrons with high kinetic energy are produced. These escape the plasma due to their charge neutrality, and pass through the plasma facing wall. In a future power plant, the neutrons will be captured in the breeding blanket and will simultaneously heat a coolant that will drive a turbine, which in turn will generate electricity [29].

3.2 Plasma geometry

As discussed previously, the main concept of magnetic confinement is that charged particles approximately follow the magnetic field lines if we do not consider drifts, waves, and collisions. The idea of the tokamak is to create a closed magnetic field geometry, such that charged particles in a plasma are confined in a torus-like shape, which is also more informally referred to as a donut shape. The geometry of a torus is illustrated in Figure 3.1.

Note that in Figure 3.1, the cross section of the torus is circular to simplify the explanation of the coordinates. In real tokamaks, the shape of the cross

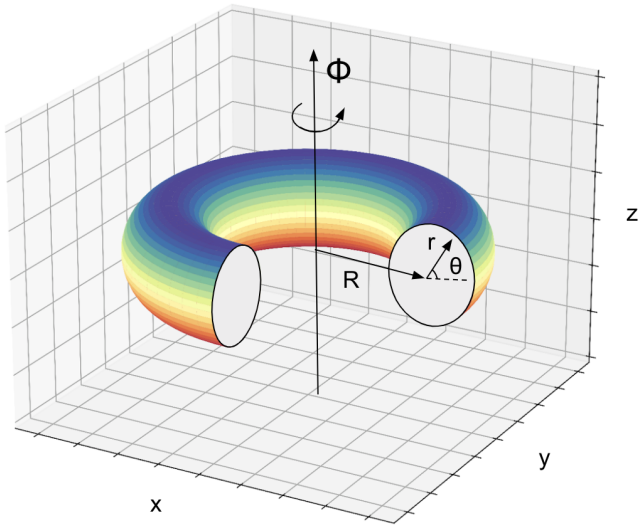

Figure 3.1: The geometry of a torus, which is described by toroidal angle Φ, poloidal angle θ, major radius R, and minor radius r. Here, x, y, and z represent Cartesian spatial coordinates.

section of the plasma can be varied. Therefore, in this thesis the plasma configuration is referred to as 'torus-like'.

Nevertheless, to obtain a torus-like shaped plasma, the first requirement is a magnetic field in the toroidal direction B_Φ. This can be achieved by placing several toroidal field coils in a circle, as illustrated in Figure 3.2. In the JET tokamak, a toroidal field on the order of $B_\Phi = 3.5$ T can be achieved, which is approximately 100 000 times stronger than the magnetic field of Earth at the equator [1].

As discussed in the previous chapter, curved magnetic field lines lead to a nonzero magnetic field gradient, in this case in the R-direction in Figure 3.1. This will lead to a ∇B drift of the particles in the z-direction. Specifically, the toroidal field in a tokamak varies as $B_\Phi \propto 1/R$. Since the ∇B drift moves electrons and ions in opposite directions, this will lead to an electric field in the z-direction, which in turn leads to an $\vec{E} \times \vec{B}$ drift in the R-direction. There are also other drifts, such as the curvature drift, that negatively affects confinement in this design. In other words, a tokamak with only a toroidal field cannot confine particles well since they drift out from the center of the plasma.

To solve this issue, a poloidal magnetic field component B_θ can be introduced. In this setup, particles do not only travel around the toroidal axis, but also around the poloidal axis such that the resulting motion parallel to the field lines has a helical pattern. For instance, if there is a drift in the z-direction upwards in Figure 3.1, then for the upper half of the poloidal orbit, the particle will drift away from the plasma (r increasing), and for the lower part of the

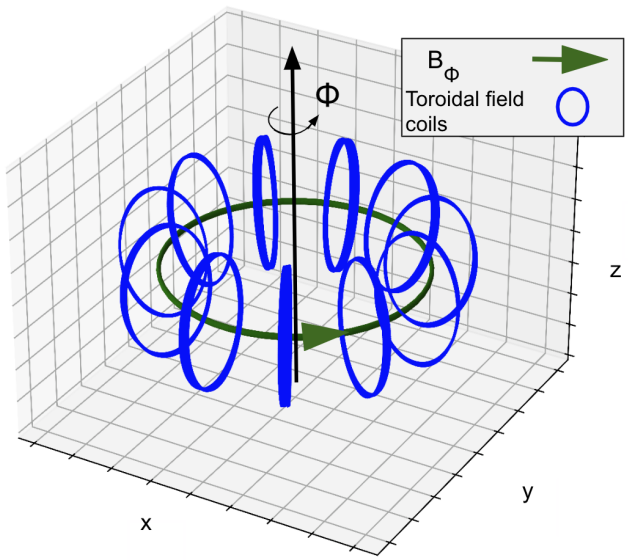

Figure 3.2: An illustration of the toroidal field coils in a tokamak. By running current through the coils, a toroidal magnetic field B_Φ can be generated.

poloidal orbit, the particle will drift towards the plasma (r decreasing). In total, the drift approximately cancels out over the course of one full poloidal turn, which leads to improved confinement.

There are different alternatives for creating a poloidal magnetic field component. In tokamaks, this is done by generating a plasma current I_P in the toroidal direction. For instance, one of the methods is to induce a toroidal electric field, which drives a toroidal current, by gradually increasing a current in a solenoid placed in the middle of the tokamak. Since the solenoid current must keep increasing in order to uphold induction, there is an intrinsic limit for how long a tokamak with induced plasma current can contain a plasma before the solenoid current becomes too high. Therefore, tokamak operations that rely only on an induced plasma current are run in pulses. There are however other non-inductive mechanisms that can contribute to the plasma current, such as the plasma self generating 'bootstrap current' [4] and currents driven by the heating systems which we explore later in this chapter. In Figure 3.3, an illustration of the helical field lines and particle paths are shown, in this case due to a plasma current.

Although a poloidal field is necessary for good confinement, it is important to emphasise that the strength of the poloidal field should not be too high in relation to the toroidal field, as this can lead to the growth of instabilities [4], [29]. The safety factor q is used in tokamaks and other MCF devices to monitor the relation between the poloidal field and the toroidal field

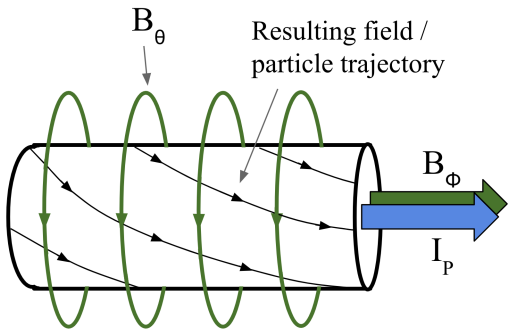

Figure 3.3: A segment of a tokamak plasma (shown as a straight cylinder although it is curved in a tokamak). A toroidal plasma current I_P leads to a poloidal magnetic field B_θ, which together with the toroidal field B_Φ results in helical fields lines. In this illustration, the Larmor motion of the particles are not drawn.

$$q = \frac{r \cdot B_\Phi}{R \cdot B_\theta}, \tag{3.1}$$

where different conditions of q are important for different regions and instabilities in the plasma.

3.3 Reactor chamber and gas fueling

A tokamak contains a plasma in a chamber. Much like the torus-shaped plasma within a tokamak, the chamber itself adopts a toroidal structure. This configuration is necessary to allocate space in the central region of the torus, beyond the plasma-facing reactor wall. This designated space accommodates the solenoid and segments of the toroidal field coils since these are situated outside the chamber.

Prior to initiating a pulse, vacuum pumps are employed to evacuate the chamber of gasses. Thereafter, gasses of the desired particle species are injected to create the initial plasma. Additional gas can be injected periodically or continuously to replenish the fuel lost during fusion reactions and due to particle transport, which will be discussed later in this chapter. Injection of gasses can also be used to intentionally cool the plasma. For instance, in the case of disruptions [30], wherein the plasma loses its confinement and stability, it is desirable to cool the plasma to mitigate potential heat damage to the reactor components. Impurities, which are elements with higher atomic numbers than the main fuel ions, can also be injected to the plasma for cooling and stabilization purposes. A useful parameter for quantifying the ion species content in the plasma is the effective atomic number

$$Z_{eff} = \sum_{\alpha} \frac{n_\alpha}{n_e} Z_\alpha^2, \tag{3.2}$$

where n_α and Z_α are the number density and atomic number of the ion species α respectively, and n_e is the electron number density.

The choice of material for the plasma facing wall is also an important consideration due to the extreme conditions inside the chamber, but also since particles from the wall can contaminate the plasma, where the impact depends on the particle species. In the JET tokamak, the wall consists of tungsten and beryllium due to properties such as high melting points, heat conductivity, and resistance to erosion. As mentioned in the introduction, future power plants will also include a lithium blanket on the wall in the reactor chamber to enable continuous breeding of tritium during a pulse.

In Figure 3.4, the reactor chamber of the JET tokamak is shown.

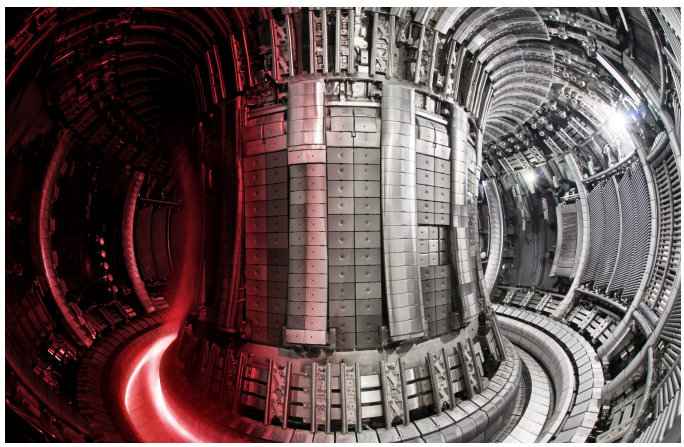

Figure 3.4: The inside of the reactor chamber of the Joint European Torus (JET) tokamak at Culham, UK, with a superimposed image of the hot plasma to the left; credit UKAEA; courtesy of EUROfusion.

3.4 Plasma heating

Optimally, in future power plants the high kinetic energy of the charged fusion products will be sufficient to heat the plasma and trigger new fusion reactions without external heating, resulting in 'ignition' of a burning plasma. The conditions needed to achieve ignition is dictated by the Lawson criterion [4], which can be expressed as

$$nT\tau_E \geq \frac{12}{E_{ch}} \frac{T^2}{\langle \sigma v \rangle}, \tag{3.3}$$

where n is the particle density, and where T is the temperature. τ_E is the confinement time, which measures the rate at which the plasma loses energy.

E_{ch} is the average kinetic energy of the charged fusion products in a given reaction, for instance, the He^4-ions (α-particles) in D-T reactions. $\langle \sigma v \rangle$ is the velocity average of the product between the reaction cross section and the particle velocity. The left-hand side of (3.3) is called the 'triple product', which fusion researchers strive to increase to achieve an efficient reactor.

Nevertheless, a burning plasma has not been achieved yet in magnetic confinement fusion, and a reactor that operate below the Lawson criterion with external heating can still be sufficiently efficient if it provides a high Q-value. Additionally, external heating will still be necessary in any device for control and for heating at the startup stage of a pulse.

3.4.1 Ohmic heating

When currents run through the plasma, heat is generated due to resistivity, which is called ohmic heating. However, the plasma resistivity η is dependent on the temperature as $\eta \propto T^{-3/2}$ [29], such that when the temperature increases, the resistivity drops, reducing the effect of ohmic heating. This puts an upper bound on the temperature that can be achieved with ohmic heating alone, and this temperature is not high enough for a sufficiently high rate of fusion reactions. Nevertheless, ohmic heating is a useful heating mechanism, especially, in the initial stages of a pulse.

3.4.2 Neutral beam injection (NBI)

One of the external heating systems in a tokamak is the Neutral Beam Injection (NBI) system. The concept of NBI is that fast neutral particles are shot into the plasma and ionized through collisions with plasma particles. The injected particles need to be neutral to not be shielded by the magnetic field lines in the tokamak before reaching the plasma. An important aspect of NBI heating, which needs careful consideration, is where the beam deposits its energy in the plasma. For instance, too low energy will lead to collisions and energy deposits only at the plasma edge, and too high energetic beams will shine through the plasma and hit, and potentially damage, the reactor walls [29].

If a NBI source is not aimed towards the center of the tokamak, but rather along the toroidal axis, it can generate a toroidal plasma rotation. Additionally, when the fast neutral particles are ionized, the ions from the neutral beam have higher momentum in the toroidal direction compared to the electrons from the neutral beam due to their larger mass. This contributes to the plasma current since the electrons will lose their toroidal velocity more rapidly.

In JET, the NBI system can shoot fast beams into the plasma up to a power of approximately 34 MW [31].

3.4.3 Radio frequency (RF) heating

The other main external heating method used in tokamaks is the application of Radio Frequency (RF) waves. It is based on launching electromagnetic waves into the plasma, which are tuned to be effectively absorbed by a resonance

mechanism associated with the motion of the plasma particles. Collisional absorption also occurs to some extent, but tends to be weak in a hot fusion plasma since it scales like $T^{-3/2}$). The two principal resonance mechanisms are cyclotron absorption and Landau damping [4]. In the former the wave frequency is tuned to resonate with the cyclotron motion of one or several particle species in the plasma, while Landau damping occurs for particles travelling with a parallel velocity matching the parallel phase velocity of the waves. In fact, the two most important heating methods employed in tokamaks today are ion cyclotron resonance heating (ICRH) and electron cyclotron resonance heating (ECRH) [29]. Because the magnetic field in a tokamak varies roughly as $1/R$, and the cyclotron frequency is proportional to the magnetic field, resonant interaction occurs around a major radius where the wave frequency matches the fundamental cyclotron frequency of a species or a harmonic of it. This allows one, at least to some extent, to tailor where the power is absorbed in the plasma. Schemes based Landau damping are used for current drive applications, e.g. Lower Hybrid Current Drive (LHCD) and Fast Wave Current Drive (FWCD) [32].

3.5 Plasma diagnostics

Diagnostic tools play a crucial role in tokamak research by facilitating measurements and analysis of the plasma. In this thesis, we focus on the two diagnostics that are the most relevant in the appended papers.

3.5.1 High resolution thomson scattering (HRTS)

In this technique, laser beams are directed into the plasma, where electrons scatter the light, causing a frequency shift. This shift is directly linked to the temperature of the electrons. The intensity of the scattered light yields information about the electron density. Consequently, through detailed analysis of the scattered light's spectrum, researchers can pinpoint the electron temperature and density at the specific location where scattering occurs. 'High resolution' refers to the spatial accuracy of the diagnostics, reaching an order of 1 cm at JET [33], [34], which is relatively accurate considering that JET has a minor radius of 1.25 m. However, one drawback of HRTS is the relatively low sampling rate of approximately 20 Hz at JET, which makes it difficult to study phenomena that occur on faster timescales.

3.5.2 Reflectometry

This method similarly relies on the interaction of light with the electrons in a plasma. However, in contrast to HRTS, where the frequency is in the visible light / near-infrared range, reflectometry systems [35] send out microwaves with lower frequencies. The frequency of a wave dictates at which electron density it will be reflected, and the total travel time of a microwave can reveal the position of a specific density. In other words, reflectometry is therefore another method to obtain the spatial distribution of the electron density in the

plasma. The reflectometry system at JET [35], which is referred to as KG10, can provide the spatial distribution of the electron density at a rate up to 10 kHz, which greatly exceeds the sampling rate of HRTS. However, reflectometry is also associated with uncertainties that occasionally propagate to large errors during the data processing to obtain the density distribution.

3.6 Plasma profiles

3.6.1 Two-dimensional profiles

To study tokamak diagnostics, such as the spatial distribution of the plasma temperature and density, obtaining a detailed interpretation from the complete three-dimensional representation of the toroidal geometry may prove overly comprehensive. Fortunately, due to the toroidal symmetry in a tokamak, we can instead study the two-dimensional (2D) cross section of the plasma.

One can for instance study the 2D projection of 'flux surfaces' [4], [29], which are surfaces where the magnetic flux remains constant. An illustration of flux surfaces is shown in Figure 3.5. Here, the field lines are arranged such that the flux surfaces are closed in the core, but open in the outer edge of the plasma. It may seem counter-intuitive to generate open flux surfaces such that particles can escape the plasma and hit 'divertor plates'. However this allows for controlled exhaust of particles and energy, which is important both for maintaining the desired conditions in the plasma, as well as for reducing the damage of other plasma facing components. The divertor plates are specifically designed to handle larger heat loads compared to other components. The contour where the flux surfaces go from closed to open is referred to as the Last Closed Flux Surface (LCFS), or the separatrix, and the region outside the LCFS with open flux surfaces is referred to as the Scrape-Off layer (SOL).

As illustrated in Figure 3.5, the plasma does not have to be perfectly circular. In this case, the plasma height b is larger than the minor radius a, which leads to an elongated plasma, where plasma elongation is defined as

$$\kappa = \frac{b}{a}. \tag{3.4}$$

The plasma shape is also characterized by triangularity. For instance, the upper triangularity δ_{up} is defined as

$$\delta_{up} = \frac{R_{geo} - R_{upper}}{a}, \tag{3.5}$$

where R_{upper} is the major radius at the highest vertical point of the LCFS. R_{geo} is the major radius at the geometric axis, which is defined as

$$R_{geo} = \frac{R_{max} + R_{min}}{2}, \tag{3.6}$$

where R_{max} and R_{min} correspond to the maximum and minimum major radius of the LCFS. In the example illustrated in Figure 3.5, the upper triangularity is 0 since $R_{geo} \approx R_{upper}$. A triangulated plasma at JET is more D-shaped compared to the elongated cylinder in Figure 3.5.

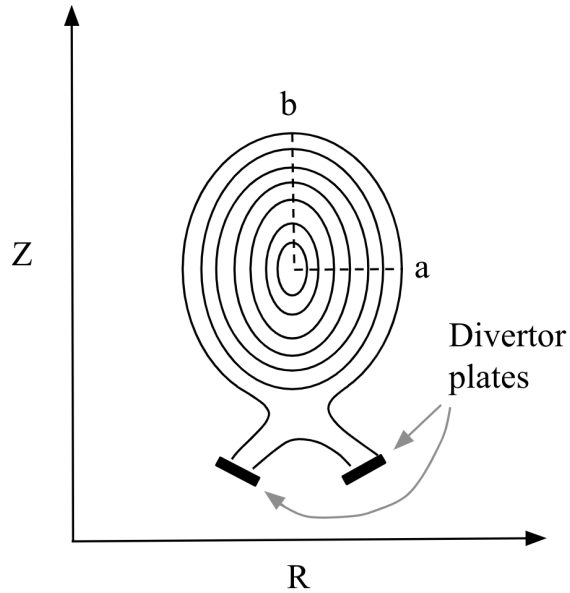

Figure 3.5: Illustration of the flux surfaces in a 2D cross section of a tokamak plasma. In the core of the plasma, the surfaces are closed. At the edge of the plasma, the field lines are configured such that the flux surfaces are open. This allows particles to be exhausted to divertor plates. The parameters a and b represent the minor radius and height of the plasma respectively. Z and R are the same spatial coordinates as in Figure 3.1.

3.6.2 One-dimensional profiles

Due to the fast motion of particles along the field lines in a tokamak, and due to derivations in MHD, many properties in a plasma are approximately constant on the flux surfaces described in the previous section [4], [29]. Therefore, it is often sufficient to further simplify the analysis to one spatial dimension (1D) by looking at flux surface averages. For instance, Figure 3.6 illustrates the 1D profile of the flux surface average electron temperature T_e. In such profiles, it is common to have a flux label ψ on the x-axis, and it is normalized to be 1 at the LCFS / separatrix. A simplified description of ψ is that it acts as a proxy for how far we are away from the center of the plasma.

It is common to study different quantities in such 1D profiles, such as the safety factor q, ion temperature T_i, electron and ion density n_e and n_i, current density j, electron and ion pressure p_e and p_i, and also the ratio of the plasma pressure to the magnetic pressure, which is defined as

$$\beta = \frac{p}{B^2/2\mu_0}, \tag{3.7}$$

where μ_0 is the magnetic permeability in vacuum. β is an important parameter for both stability and confinement, and it is also often expressed in terms of its

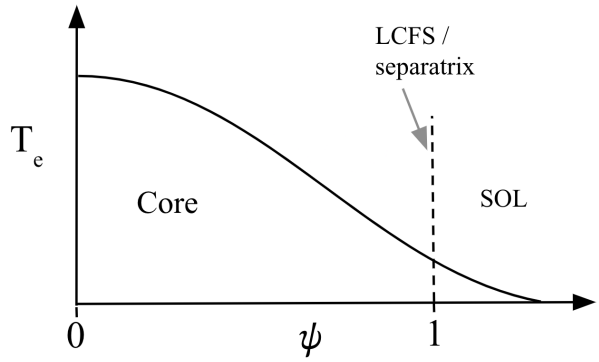

Figure 3.6: An illustration of an 1D flux surface average electron temperature profile. Here, the temperature T_e is drawn to be highest in the core of the plasma, and it decreases towards the SOL region. The flux label ψ indicates how far out we are in the plasma.

normalized version

$$\beta_N = \beta \frac{aB_T}{I_p}. \tag{3.8}$$

In simulations, it is often of interest to find steady state solutions, or the evolution of 1D profiles since their characteristics reveal dynamics of the plasma.

3.7 Heat and particle transport in a tokamak

For most of this thesis, it has been discussed how particles are confined to magnetic field lines if we neglect drifts. In this section, other mechanisms that lead to heat and particle transport perpendicular to the magnetic field lines, and how they pose challenges for confinement, are summarized.

3.7.1 Classical transport

This type of transport is based on collisions between charged particles in the plasma. By calculating average collision times and treating the particle movement as a random walk process [6], the total travel distance of heat and particles after a certain time can be calculated if the spatial step size is known. Since particles gyrate in tokamak plasmas, the spatial step size is on the same order as the Larmor radius. Classical transport typically contributes a minor fraction to the overall transport in a tokamak plasma, which corresponds to a long confinement time τ_E [29].

3.7.2 Neoclassical transport

Classical transport can be extended to neoclassical transport by including effects from the toroidal geometry, such as $B_\Phi \propto 1/R$. When particles travel from a lower magnetic field to a larger field, their perpendicular velocity must increase as the magnetic moment (2.11) of a particle is a conserved quantity. For energy to also be conserved, the parallel velocity must decrease. Some particles with too low initial parallel velocity on the low field side will at some point at higher field strength reach $v_\parallel = 0$, and then bounce back in the opposite direction. This leads to a fraction of particles that are trapped and travel back and fourth on the low field side in orbits referred to as banana orbits due to how they look in a 2D projection at fixed Φ. The width of these banana orbits, due to drifts, leads to a larger spatial step size due to collisions compared to classical transport, which makes confinement more difficult since particles escape more rapidly.

3.7.3 Turbulent transport

Turbulent transport [36] is conceptually different from the previous transport processes as it is not rooted in collisions, but rather arises from the complex and chaotic behavior of plasma fluctuations. There are different mechanisms that can contribute to turbulent transport. For instance, micro instabilities can lead to a large heat and particle flux that almost always exceeds the transport contribution from classical and neoclassical transport in current machines. Understanding and controlling turbulence in tokamaks is therefore a prioritized research topic in MCF to improve confinement.

The turbulent transport from micro instabilities can be calculated by solving differential equations describing perturbations as discussed in Section 2.3. In simulations, this is done numerically, either through solving the nonlinear equations or by discarding the nonlinear terms, although the previous is far more computationally costly. Certain models, such as EDWM [37], TGLF [38], and QuaLiKiz [39] are quasi-linear models. These calculate the growth rate of instabilities with the linear approach, and then applies a saturation rule to account for the nonlinear interactions resulting in a transport estimation without having to explicitly solve nonlinear equations. The most common turbulent transport contributing instabilities, which are also referred to as drift waves, are

- Ion Temperature Gradient (ITG) mode - As the name suggests, this instability is driven by gradients in the ion temperature.
- Electron Temperature Gradient (ETG) mode - Similar to the previous but due to gradients in the electron temperature.
- Trapped Electron Mode (TEM) - Arises from the fact that some electrons are trapped in banana orbits and therefore cannot cancel out fluctuations in the electric field to the same extent as if there were no trapped electrons.

A typical feature of drift waves is that they result in stiff profiles [36]. This means that the shape of the temperature profile is roughly the same above a certain threshold of the normalized gradient $\nabla T/T$, regardless of the applied heating and fueling profiles. Additionally, the heat flux of drift waves is sensitively dependent on the normalized gradient $\nabla T/T$ above this threshold [36]. These two features implies than an overall increase in temperature greatly impacts the heat flux of drift waves. As we will see later in this chapter, a pedestal can form near the edge in the temperature profile given the right circumstances. This leads to an overall elevation in the entire temperature profile due to the stiffness, which implies that the pedestal at the edge impacts the turbulent transport, even in the core.

3.8 Integrated modelling

Due to the many phenomena and engineering aspects that are involved in a tokamak plasma, several models that explain different processes are often used together in simulations. This is called integrated modelling, and is by itself a complicated and essential research topic in magnetic confinement fusion. In integrated modelling, modules are configured in a numerical flowchart, where the outgoing information of a model can be forwarded to other models. This enables flexibility as different models that aim to calculate the same thing can be exchanged based on the preference of the user. Individual models can be specialized to, for instance, account for the NBI heating, or to estimate the turbulent transport from a chosen theoretical description.

However, integrated modelling frameworks can be limited by slow, computationally expensive models. The models used are often employed as a compromise between accuracy and computational demand, and ongoing research, in particular in the machine learning domain as we will discuss in the next chapter, strives to reduce the computational requirement of expensive models [40].

An example of an integrated modelling framework is the European Transport Simulator (ETS) [41], which integrates a catalog of modules to simulate, for instance, the different transport mechanisms in a tokamak plasma.

3.9 The pedestal

In the 1980s, it was discovered that confinement suddenly increased at high applied heating power in the ASDEX tokamak, at Garching, Germany [42]. It turns out that transport had been suppressed near the LCFS, which resulted in steep temperature and density gradients in the profiles, as illustrated in Figure 3.7. This operational regime was named the High-confinement mode (H-mode), and the elevated temperature and density near the LCFS was named the pedestal due to its visual shape.

The characteristics of the pedestal are to this day still not fully understood. Essentially, the forming of the pedestal consists of two parts; 1) local transport suppression which leads to its build-up; 2) rapid drops in the pedestal top

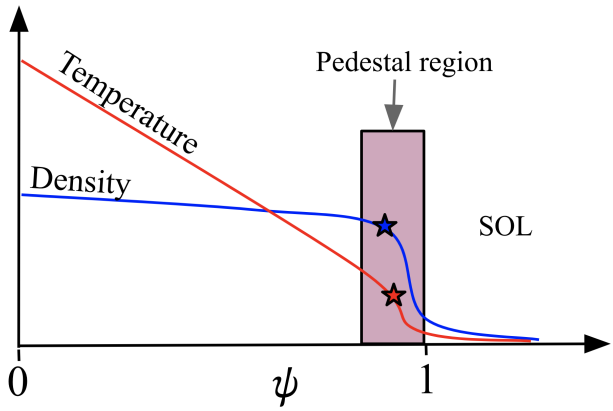

Figure 3.7: An illustration of the pedestal in tokamaks. In the pedestal region, which is sometimes referred to as the plasma edge, transport is suppressed such that steep gradients are formed. The values at the top of the pedestal are highlighted as stars.

temperature and density due to instabilities called edge localized modes (ELMs) that can be triggered by different mechanisms [43]. This results in a cyclic pattern of the pedestal called the ELM-cycle [44], which is illustrated in Figure 3.8.

The ELMs impose an upper limit on the temperature and density at the top of the plasma pedestal, a limit influenced by machine parameters such as total plasma current and plasma shape [45]. For integrated modelling applications, these pedestal top values are important as they act as boundary conditions when simulating phenomena in the core. For instance, turbulent transport in the core is affected by the profile shift induced by the pedestal as discussed previously. Consequently, there is a demand for a model capable of predicting these pedestal top values from machine parameters. In a more general sense, the enhanced confinement due to the pedestal is a key element in the operation of current machines as well as for extrapolating to future machines, further motivating the improvement of predictive capabilities.

Considering the characteristics of the pedestal is additionally important for heat load management. Certain types of ELMs, when triggered, can release a considerable amount of energy in a brief duration [45]. This poses a threat to plasma-facing components, including divertor plates. In the context of larger future machines, such as ITER, more energy is going to be stored in the plasma. Therefore, mitigating ELM types associated with excessive energy release becomes paramount. Optimally, further understanding of the pedestal may help design operational scenarios where a balance in transport is achieved before ELMs are triggered or where less disruptive types of ELMs are intentionally induced [46]. For this purpose, improved understanding of both the build-up of the pedestal as well as the ELMs will be important.

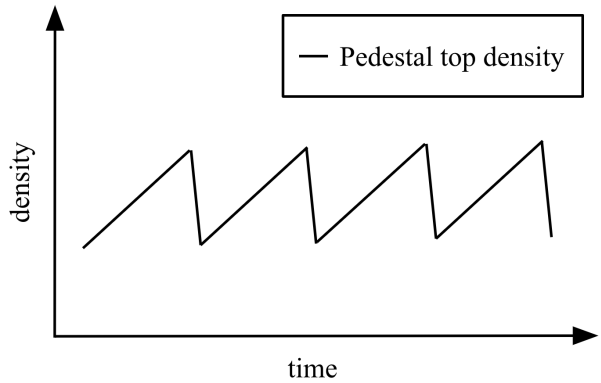

Figure 3.8: Illustration of the cyclic dynamics of the pedestal, in this case the pedestal top density. Due to local transport suppression, the pedestal density increases before an instability referred to as an ELM is triggered, which rapidly lowers the pedestal density. The increase of the pedestal is drawn as linear in this illustration, however this is not necessarily always the case.

3.9.1 Edge localized mode types

The physics that trigger ELMs are not fully understood and they are generally categorized by characteristics from experiments [47]. Two of the ELM types that have been observed at JET are

- Type I ELMs - These are characterised by a positive correlation between the ELM-cycle frequency f_{ELM} and heating power. They are usually found when the heating power is well above the H-mode threshold. Type I ELMs can lead to large energy deposits which makes them dangerous for future machines.
- Type III ELMs - These are characterised by a decreasing f_{ELM} when increasing the heating power, and are found near the H-mode power threshold. Compared to Type I ELMs, Type III ELMs are associated with lower energy deposits, posing less threat to plasma facing components.

3.9.2 Pedestal modelling

One approach to model the pedestal is through ideal MHD. Essentially, ideal MHD describes that large currents at the edge can destabilize 'peeling' modes, and that large pressure gradients can destabilize 'ballooning' modes. The Peeling-Ballooning (PB) [48] model incorporates these effects and calculates a stability boundary in $j - \alpha$ phase space where j is the current density at the edge, and α is the normalized pressure gradient. This phase space has a stable region and an unstable region, separated by the boundary which shifts depending on machine parameters and plasma conditions. According to this model, an ELM is triggered when the pedestal has grown such that the stability

boundary is reached. The pressure gradient in the steep region of the pedestal depends on the pedestal top pressure and the pedestal width in the pressure profile. Hence, the PB model calculates a relation between the pedestal height and width. However, since the width is unknown prior to an experiment, the PB model alone cannot make a prediction for the pedestal top.

In the EPED model [49], the PB model is combined with another dependence between the pedestal top and width that together enable predictions based on where the dependencies intersect. The other dependence is based on kinetic theory and called the kinetic ballooning mode (KBM) criterion [49]. This framework has shown to provide accurate predictions for the pedestal pressure for many Type I ELM plasmas at JET. However, there is also a non-negligible subset of Type I ELM plasmas at JET that do not agree with the PB model [45]. Additionally, to obtain the pedestal temperature, which is the output of EPED, one has to assume the pedestal density since $p \propto nT$ where p is the pressure. Recent work [50] has shown promising results in alleviating this obstacle by predicting the pedestal density through an extended version of the neutral ionisation model [51].

Machine learning based surrogate models, which are described in the next chapter, are currently being developed to emulate EPED. These will provide a useful application in integrated modelling frameworks as surrogate models can provide fast predictions of the pedestal compared to the original EPED model. However, as these surrogate models are trained to emulate EPED, they will also fail to represent pedestals that do not agree with the PB model.

3.9.3 Empirical pedestal scalings

Extensive empirical studies have previously been performed to improve the understanding of the characteristics of the pedestal [45], [52]–[55]. It is of interest to study how pedestal properties, such as top values and width, correlate with machine parameters, such as the plasma current and heating power. It is additionally interesting to examine for which operational scenarios the critical pressure gradient deviates from the one predicted by the PB model. More recently, it has been investigated for how the pedestal changes when going from a deuterium and/or hydrogen plasma to a deuterium-tritium plasma [56].

Empirical analysis can be done in multiple ways. It is common to select a small set of pulses where most machine and plasma parameters are approximately constant. By only varying one machine parameter, its dependency with different pedestal properties can be investigated. However, a caveat with this approach is that the dependency is found for a specific operational scenario. It is not guaranteed that such dependencies remain constant when changing the parameters that were constant in this sub set of pulses. An optional approach is to investigate the dependency between a machine parameter and the pedestal across a large data set, as in [45]. This method is useful for visualizing general trends that hold for large operational domains. However, as other parameters are not constant across these large data sets, it becomes challenging to isolate dependencies. To counter this issue, multi-variate analysis can be performed by curve fitting. By assuming a functional mapping between several machine

parameters and pedestal properties, it is possible to separate the contribution from different parameters to the pedestal. For instance, in [45], the pedestal is empirically modelled at JET by assuming a power scaling law, which has yielded the results

$$T_{e,ped} = (0.05 \pm 0.03) I^{0.00\pm0.2} P^{0.74\pm0.12} \delta^{-0.23\pm0.15} \Gamma^{-0.16\pm0.05} M^{0.3\pm0.4}, \quad (3.9)$$

$$n_{e,ped} = (9.9 \pm 0.3) I^{1.24\pm0.19} P^{-0.34\pm0.11} \delta^{0.62\pm0.14} \Gamma^{0.08\pm0.04} M^{0.2\pm0.2}, \quad (3.10)$$

where $T_{e,ped}$ is the pedestal top temperature in keV, $n_{e,ped}$ the pedestal top density in $\mathrm{m}^{-3}(10^{19})$, I the plasma current in MA, P the total heating power in MW, δ the triangularity which is unitless, Γ the fueling rate in charge per second (10^{22}), and M the effective mass which is unitless.

The scaling law for $n_{e,ped}$ is rather accurate, where it has achieved a R^2-value of 0.80. The R^2-value is the coefficient of determination [57], which is 1 when a model perfectly describes all data points, and 0 when the model is equally insufficient as a model that always outputs the mean value of the output variable as its prediction. The R^2-value for $T_{e,ped}$ is lower (0.70), which implies that there are either input parameters missing, or that the functional mapping between machine parameters and $T_{e,ped}$ is more complicated than what can be described with a power scaling law.

In summary, general dependencies related to the pedestal are relatively mapped. However, there is still a discrepancy between these relatively simple models and experimental data which motivates further investigation and development of more expressive empirical models.

Chapter 4

Machine learning fundamentals

The purpose of this chapter is to provide an additional background for the machine learning methodology used for the plasma physics applications in the appended papers. Given that the models developed in these papers are specifically grounded in neural networks, we will focus on this approach, and gradually introduce their components. The concepts in this chapter are based on [57], which provides more thorough explanations.

4.1 The neural network node

The node is the fundamental building block in neural networks. As illustrated in Figure 4.1, it takes a set of input parameters, in this example $[x_1, x_2, x_3]$, and predicts an output $\hat{y}$.

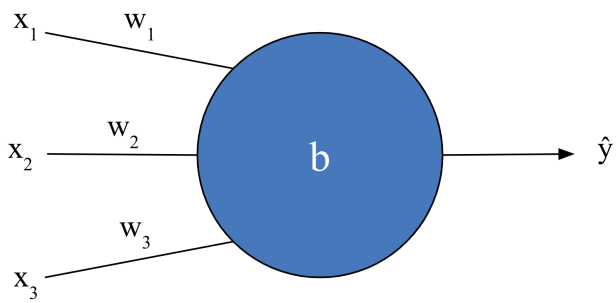

Figure 4.1: An illustration of a node in a neural network.

The calculation of $\hat{y}$ is done in two steps. First, input parameters are multiplied with the weights, in this example $[w_1, w_2, w_3]$, and a bias b is added

such that

$$z = \sum_i w_i x_i + b, \tag{4.1}$$

where z is an intermediate parameter called the pre-activation value. The output of the node is then obtained by applying a suitable activation function g where

$$\hat{y} = g(z). \tag{4.2}$$

4.2 Example: Linear single-node model

To simplify the explanation for how a machine learning model 'learns from experience', let us now consider an example where we employ one node as a model. Assume that we want our model to accurately estimate a quantity y by predicting it from one input parameter x, where the true underlying relationship that describe the data is

$$y = 0.8x - 0.5. \tag{4.3}$$

In a real scenario, we are often not aware of the true relationship between the input and output parameters. We might however have access to a data set with tabular values of y and x that can be used to train our model.

If we assume the simplest possible activation function, which is the identity mapping $g(z) = z$, the functional map of our model becomes

$$\hat{y} = wx + b. \tag{4.4}$$

This type of activation is occasionally referred to as a linear activation, as the output now is a straightforward linear function of the input parameter.

4.2.1 The loss and cost function

In machine learning, an initial guess is made for the weight and bias parameters. For instance, let the initial guess be $w = 0.6$ and $b = 0.6$ in our example. If the first row in a tabular data set is $[x = 1, y = 0.3]$, then our model will predict $\hat{y} = 1.2$, which clearly is wrong since the correct answer is $y = 0.3$ according to the underlying function that describes the data (4.3).

A loss function L provides a method to quantify the error of individual predictions. For instance, a common loss function is the squared error

$$L = (y - \hat{y})^2, \tag{4.5}$$

which also can be expressed as a function of the weight and bias parameters by inserting (4.4) into (4.5)

$$L = (y - wx - b)^2. \tag{4.6}$$

The derivative of the loss L with respect to w and b can now be obtained, which we will make use of in the next subsection

$$\frac{\partial L}{\partial w} = 2(y - wx - b) \cdot (-x), \tag{4.7}$$

$$\frac{\partial L}{\partial b} = -2(y - wx - b). \tag{4.8}$$

In total, we have defined a method to estimate the error of our model using a loss function, and we have analytically derived how this loss function depends on the weight and bias parameters of the model.

The difference between a loss function and a cost function is that the latter is the average loss across several data entries, although the two terms are often used interchangeably. The cost function in our example is the mean squared error (MSE)

$$\text{MSE} = \frac{1}{N}\sum_{k=0}^{N}(y_k - \hat{y}_k)^2, \tag{4.9}$$

where N is the total number of data entries being evaluated. The derivative of the MSE cost function with respect to, for instance w, can be obtained for our example by expanding (4.7)

$$\frac{\partial}{\partial w}\text{MSE} = \frac{1}{N}\sum_{k=0}^{N} 2(y_k - x_k w - b) \cdot (-x_k). \tag{4.10}$$

4.2.2 Optimization

The next step in achieving an accurate model is to adjust the weight and bias parameters iteratively, which are referred to as trainable parameters. The specific strategy is to iteratively adjust the trainable parameters to gradually minimize the loss and cost functions. Fortunately, the analytical expression for the loss as a function of the trainable parameters is known. For instance, a data entry that gives $\partial L/\partial w > 0$ indicates that if w is increased, the loss is also increases. Since the goal is to minimize the loss and cost function, w should instead be decreased. In general, optimization algorithms in machine learning are based on the concept that the trainable parameters should be shifted in the opposite direction as the sign of the partial derivative of the loss/cost. Stochastic Gradient Descent (SGD) is an optimization algorithm that follows this concept, and it defines the update of an arbitrary trainable parameter, that is, a weight or bias parameter θ

$$\theta^{l+1} = \theta^l - \eta\frac{\partial L}{\partial \theta}, \tag{4.11}$$

where η is called the learning rate, which controls the step size of the parameter update. The superscript l refers to the iteration number.

4.2.3 Training procedure and result

To train our model, we use a data set with 100 data entries generated with the underlying function (4.3) in this example. We use the same parameter initialization as was mentioned before ($w = 0.6$ and $b = 0.6$), and we use SGD with $\eta = 0.05$ as the optimization algorithm. Instead of the weight and bias update for each data entry individually, we calculate the full MSE cost function on the entire data. This is called full-batch training, where the batch size, which is defined as the number of data entries being evaluated in a parameter update, is equal to the number of rows in the data set.

The training result is shown in Figure 4.2, which shows the evolution of w, b, and the cost function in the iterative training process. After approximately 200 parameter updates, which also is referred to as training iterations, the model has successfully been able to find $w = 0.8$ and $b = -0.5$, which has yielded a low cost function value.

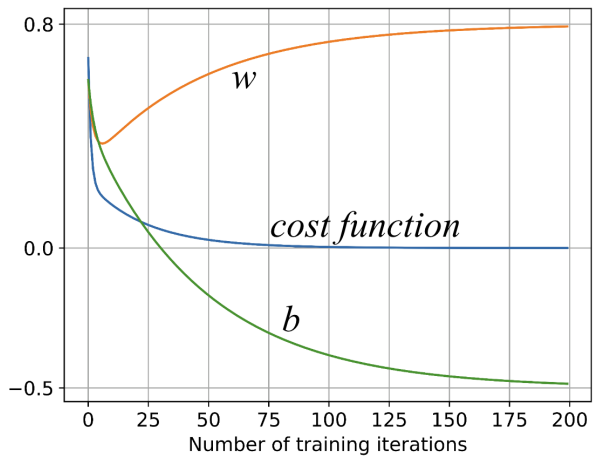

Figure 4.2: The result of the iterative training process. Even though the parameter w decreases at the initial iterations, the algorithm later finds that w needs to increase and saturate at 0.8 to minimize the cost function.

4.3 Dense neural networks

The one-node model in the previous example was able to find the true underlying relationship between the input and output parameter due to the simplicity of the problem. However, there are many problems, in particular in plasma physics, that are much more complicated.

To create a models that can approximate a larger family of functions, multiple nodes can be arranged in several layers to form a dense neural network, as illustrated in Figure 4.3.

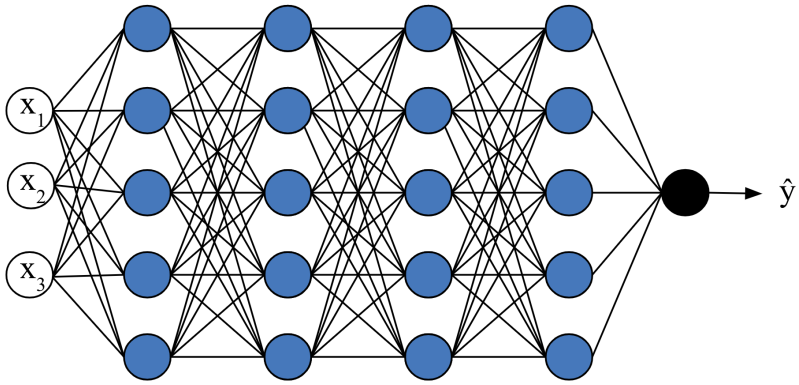

Figure 4.3: The architecture of a dense neural network. Here, the output of a node is forwarded as an input to each node in the next layer (left to right). This particular network consists of 3 input parameters in the input layer (white), 4 hidden layers (blue) with the layer size 5 (5 nodes each), and an output layer with one output node (black). This is a relatively small network, and many real application use networks with a significantly larger number of nodes. The input and output dimensions of a network does not either need to be small. For instance, networks that either process or predict image data have many input or output nodes to represent pixel values. The inputs of a model are also usually referred to as features, and the outputs are also referred to as labels.

In neural networks, the final output prediction $\hat{y}$ is based on the cumulative calculation of all individual nodes, and its functional mapping can be represented as

$$\hat{y} = f(\vec{x}, \vec{\theta}), \tag{4.12}$$

where $\vec{x}$ represents all of the input parameters, and where $\vec{\theta}$ represents all of the trainable parameters in all of the nodes. The training procedure of a neural network follows the same principle as for the case with the single node; a cost function is differentiated with respect to all of the trainable parameters $\vec{\theta}$, which can be done analytically through automatic differentiation in a programming script. The trainable parameters are then iteratively updated through an optimization algorithm to minimize the cost function.

4.4 Nonlinear activation functions

When the activation function for all nodes in a neural network is set as the identity function, the overall functional mapping is effectively reduced to a linear function. Therefore, to allow for learning of more complicated relationships, nonlinear activation functions are needed. Common nonlinear

activation functions include the Rectified linear unit (ReLU), and the sigmoid function σ, which are defined as

$$g_{\text{ReLU}}(z) = \max[0, z], \tag{4.13}$$

$$\sigma(z) = \frac{1}{1 + e^{-z}}, \tag{4.14}$$

where z again is the pre-activation value in the nodes (4.1). Both of these activation functions are visualized in Figure 4.4.

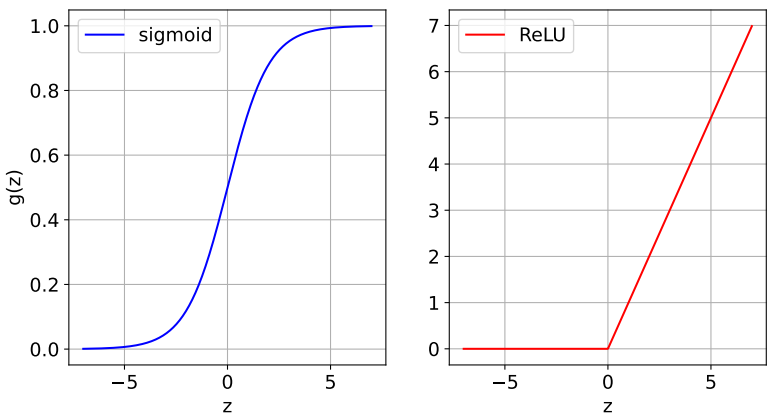

Figure 4.4: The sigmoid activation function (left) is characterized by being bounded to 0 and 1. The ReLU function (right) is characterized as a piece-wise linear function where $g(z) = 0$ when $z < 0$, and where $g(z) = z$ when $z \geq 0$.

Although nonlinear activation functions are essential for enabling complicated functional mappings, the output layer of a model that predicts a non-discrete value is usually set to be linear to avoid a bounded output.

Previously, we discussed how the loss function can be expressed as a function of the trainable parameters. This is also true when nonlinear activation functions are implemented, however, the differentiation of the loss function with respect to the trainable parameters must be adjusted based on which activation function is used.

4.5 Hyperparameters

In machine learning, hyperparameters refer to configurations that are set before the training procedure starts. For neural networks, hyperparameters include: the number of hidden layers, the number of nodes in each layer, choice of optimization algorithm, learning rate, batch size, choice of loss/cost function, choice of activation function. The number of training iterations is also a hyperparameter, and the number of 'epochs' refers how many times the full data set is parsed through the model during the training.

4.6 Training, validation, and testing

Large neural networks with many trainable parameters are desired for certain problems in the sense that it allows for highly complicated functional mappings to be learned. However, there is usually a degree of noise in the data in real problems. As the sole goal of a model is to minimize the cost function during training, a model with many trainable parameters pose the risk of learning noise, or to memorize specific data entries if it is trained for too many iterations. This phenomenon is referred to as overfitting, which is important to mitigate to improve generalization capabilities.

Due to the risk of overfitting, it is often meaningless to only evaluate the cost function on the data that has been used in the training. Therefore, a data set is usually split into three parts:

- Training set - As the name suggests, this data set is used for the training of the model.
- Validation set - This set is held out during training and allows for a more unbiased evaluation when comparing different combinations of hyperparameters to find the most suitable configuration. This set is also used to monitor when a model shows signs of overfitting, as illustrated in Figure 4.5.
- Test set - To enable a fully unbiased evaluation of the model, a test set is held out both during the training and the search for the optimal hyperparameters.

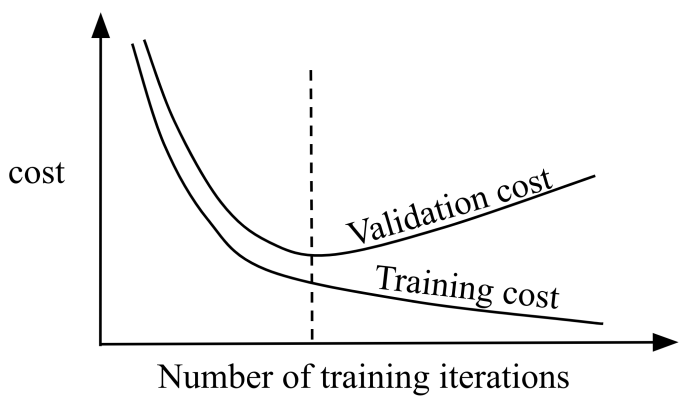

Figure 4.5: An illustration of a cost curve. For the early training iterations, both the validation and training cost decrease. However, at a certain point, which is marked by the dashed line, the validation cost begins to increase while the training cost keeps decreasing. This is an indication that past this point, the model is overfitting to a high degree.

A conventional data split often involves allocating approximately 70-80%

for training, 10-15% for validation, and another 10-15% for testing, although it can vary depending on the application.

4.7 Classification models

The introductory example in this chapter illustrated a regression task since the model was trained to predict a non-discrete output. However, neural networks and other machine learning models can also be used for classification tasks by using a sigmoid activation function in the output layer. As the sigmoid function is bounded between 0 and 1, data sets can be curated to represent classes with ones or zeroes. Although traditional loss functions, such as squared error, can be used in classification tasks, it is more common to use loss functions that consider probability distribution of predictions and labels, such as the binary cross-entropy loss function

$$L = -(y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})). \tag{4.15}$$

4.8 Ensemble learning

Once a neural network has been trained, it can be used to make predictions in a desired field of application. However, when tasked with making predictions far beyond the scope of its training, the model is prone to generating inaccurate or unreliable results. This is often referred to as the extrapolation problem, or the domain adaptation problem. Additionally, it might not be obvious for the user that predictions are made far beyond the training domain.

Ensemble learning is a strategy, not necessarily to enable domain adaptation, but to identify when the model is making predictions outside its training domain. For this purpose, several models are trained separately to perform the same task. The models have to vary to some extent, either through different hyperparameters, different sets of initial weights, or that the models are based on different machine learning models. A small subset of the training data can also be excluded for each model such that no pair of models are trained on the exact same data. Post training, the idea is that the models should agree to a higher degree when they make prediction in the training domain compared to when they make predictions far from the training domain. As the models have not been encouraged to follow the same trends in the data outside the training domain, predictions are assumed to be more spread out. This can, for instance, be quantified with the standard deviation from the mean of the predictions, which serves as a proxy for estimating the confidence of the ensemble of models in a given prediction.

4.9 Surrogate models

As discussed, machine learning models can be trained on empirical data to represent highly complicated functional mappings, which is particularly useful

when the underlying dynamics of a system are unknown. However, neural networks can also be trained on data generated by other models where the underlying theory is known. This is called surrogate modelling. In plasma physics, it is common to solve differential equations numerically, which sometimes requires a lot of computational power. This can be particularly limiting in integrated modelling frameworks, where the entire simulation might be significantly slowed down by one model in the pipeline. For instance, if we use integrated modelling as a tool to study model A, but the computational demand of another model B inhibits fast simulations, it may be beneficial to replace model B with a surrogate model that has been trained to emulate model B. The rationale behind this replacement lies in the efficiency of a single forward pass through a classical neural network, making it a swift process. This of course requires that there exists a data set that has been generated by standalone simulations of model B, which should cover the relevant domains for the integrated modelling cases. Since the entire premise of using surrogate models is that the underlying model is computationally costly, ongoing research on a topic referred to as 'active learning' strives to optimize the data generation process, such that as few data entries as possible are required to cover a wide representative domain [58].

4.10 Interpretability and the black-box problem

Although neural networks, and machine learning models in general, have been shown to be useful in solving many complicated tasks, one of the fundamental challenges is interpretability. Interpretability, or explainability, refers to how transparent a model is in its decision making process when predicting an output. In theory, a neural network is transparent, as the calculation process for the output of each node is fully defined. However, in practice, the interpretation of a model becomes challenging for humans when dealing with thousands, and at times millions, of parameters. For this reason, neural networks are often referred to as black boxes. As machine learning becomes more prominent in different applications, interpretability will incrementally become a greater concern, which ongoing research strives to address [25], [26].

Chapter 5

Summary of Appended Papers

5.1 Enabling adaptive pedestals in predictive transport simulations using neural networks

In this paper we present a neural network model that predicts the electron temperature and electron density at the top of the pedestal in tokamaks, which was named PEdestal Neural Network (PENN). The purpose of this work was to create a model that can be used in integrated modelling frameworks. Specifically, transport simulations of the core of the plasma are dependent on boundary conditions near the plasma edge that are determined by the top values of the pedestal. By enabling an adaptive pedestal through predictions, simulations are not limited to static boundary conditions. Although there exists other, first-principle based, models (EPED [49]) that predicts the top of the pedestal, they are computationally costly, and not accurate for all scenarios [45]. Recent work has shown that some of these inaccuracies can be mitigated by including resistive effects [59], although it is not certain when a surrogate based on such a model will be available, due to the computational demand of generating data when considering resistive effects.

The model proposed in this work was trained as an ensemble of networks using approximately 1500 data entries from the JET tokamak for a wide range of the machine parameters in the H-mode regime. The 12 input parameters were: β_N, plasma current, toroidal field, minor radius, elongation, NBI power, total power, upper triangularity, lower triangularity, plasma volume, the safety factor, and the effective atomic number of the plasma. Evaluation on a test set showed that an R^2-value of 0.93 could be achieved for the pedestal temperature, and that a R^2-value of 0.91 could be achieved for the pedestal density.

As a demonstration, PENN was applied in the integrated modelling framework ETS [41]. Two pulses with different NBI power were simulated, where results showed that PENN was able to replicate the difference in the pedestal region due to the variation in NBI power.

The main conclusion of the paper is that indeed, it is possible to get accurate predictions of the pedestal at JET from a set of machine parameters using neural networks.

5.2 A fast neural network surrogate model for the eigenvalues of QuaLiKiz

QuaLiKiz [39] is a quasi-linear model that calculates turbulent transport in a plasma in two steps; 1) calculation of eigenvalues of micro instabilities using linear theory; 2) a saturation rule is applied for the eigenvalues to obtain the transport fluxes. QuaLiKiz can be used in integrated modelling frameworks, although it can be time consuming during long simulations or extensive analysis. Therefore, a surrogate model for QuaLiKiz (QLKNN [60]) was previously developed. However, a caveat with QuaLiKiz is that the saturation rule is calibrated using experimental data, which makes it challenging to predict for future machines since there is no guarantee that the saturation rule translates well.

In this paper we present a neural network based surrogate model for the part of QuaLiKiz that is robust and translate between machines, which is the calculation of the linear theory to obtain the eigenvalues of the instabilities. This is also the most computationally costly part of QuaLiKiz. Specifically, our objective is to explore the key considerations that must be taken into account from a machine learning perspective when solving this problem. A data set with over 8 million entries was available for the training and evaluation of the model. The output in this problem consists of the growth rate and real frequency of instabilities at 18 different spatial scales, where the shortest scales generally correspond to the ETG-mode, and where the longer scales generally correspond to the ITG-mode and the TEM-mode. In the data set, the growth rate (one of the outputs) at each spatial scale is either positive (growing instability / unstable mode), or zero (stable or damped mode). Results showed that splitting the model into a stable/unstable classifier and a regression model for the unstable entries, yielded an accurate surrogate model as a whole. The accuracy could be further improved by using a weighted loss function for the classifier due to class imbalance between the stable/unstable classes.

The main conclusion of this paper is that task splitting helped improve the accuracy of the surrogate model, and that this may also apply for future surrogate model applications related to eigenvalues of quasi-linear models.

5.3 High temporal resolution of pedestal dynamics via machine learning

For most H-mode plasmas, the pedestal follows a cyclic pattern [44], where it builds up due to transport suppression and then suddenly drops as instabilities (ELMs) are triggered. This dynamical process is of interest to study to better understand the pedestal. However, the cyclic pattern usually occurs on much

faster timescales than what can be captured with the HRTS diagnostics at JET (sampling rate: 20 Hz) [33], [34]. Reflectometry [35] provides an alternative diagnostics for high temporally resolved density measurements (10 kHz), although it is not as consistently accurate as HRTS.

In this paper we present a neural network model that predicts the HRTS 1D electron density profile from the reflectometry 1D electron density profile. A data set was created by pairing HRTS profiles with the corresponding reflectometry profiles that were closest in time, which yielded 62140 entries from approximately 200 pulses. As the spatial grid varies from example to example in reflectometry, the model takes as input both 100 density values and the 100 corresponding position values. The model predicts 63 values: the density at 63 fixed spatial grid points corresponding to the HRTS density measurements. Post training, the model can predict what the HRTS profile would have looked like at all time points where reflectometry data exists between the actual HRTS measurements.

Results showed that the model was able to produce accurate predictions on a test set (pulses that were not included in the training). To demonstrate the applicability of the model, the dynamics of the pedestal was visualized for the record-breaking pulse 99869 at JET in terms of energy production. The predicted HRTS profiles showed reasonable and consistent pedestal dynamics signals, that additionally were less noisy compared to the pedestal dynamics signals obtained via reflectometry alone.

The main conclusion of this paper is that with machine learning, we can obtain a predicted diagnostics that combines the high temporal resolution of reflectometry with the high spatial accuracy of HRTS.

Chapter 6

Future Work

The appended papers in this thesis encompass strategies for predicting the pedestal in the JET tokamak with neural networks. Although results have shown that accurate predictions can be made, the models lack transparency for why a certain prediction is made. There are two aspects to this problem in fusion research; 1) The main goal of magnetic confinement fusion is to create an environmental friendly and sustainable energy source. If machine learning models can assist in reaching this goal, they should not be neglected solely because they are not fully transparent; 2) On the other hand, reaching this goal requires an improved understanding of how plasmas behave for reactor relevant conditions. If machine learning models can be made more interpretable, they may assist in increasing the knowledge.

A specific example of future work would be to try other, more interpretable models to predict the pedestal from machine parameters. The goal would be to fully understand the functional mapping from input parameters to output parameters. This is important, not only for understanding the pedestal, but also for enabling traceability when using pedestal prediction models in integrated modelling. In addition, if comprehensive pedestal data sets were to be available from other tokamaks, the functional mapping of different machines could be compared. Improved understanding of how the behaviour of the pedestal differs between machines will be useful for the speculation of how the pedestal will behave in future machines.

Another example of future work is to use machine learning to find low-dimensional representations of 1D plasma profiles. Specifically, it would be of interest to find low-dimensional representations of the edge of the plasma profiles where the pedestal is located. The concept of compressing high dimensional tokamak data to lower dimensional representations using, for instance, autoencoders [57], is not new [61]. However, it remains a challenge for how to interpret this low-dimensional representation. Autoencoders combined with interpretable machine learning could potentially assist in finding interesting features related to the pedestal. Discoveries that improve the understanding of the plasma core-pedestal interaction may also be valuable for optimizing how pedestal models are implemented in integrated modelling frameworks.

As these topics are oriented towards finding relationships in data, future work is inherently associated with a more extensive experimental analysis compared to the research done so far in this PhD project.

Bibliography

[1] C. Nordling and J. Österman, *Physics Handbook*, 9th ed. Studentlitteratur, 1980 (cit. on pp. 3, 6, 22).

[2] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 2nd ed. Cambridge University Press, 1985 (cit. on pp. 3, 4).

[3] C. G. Tully, *Elementary Particle Physics in a Nutshell*, 1st ed. Princeton University Press, 2011 (cit. on pp. 3, 4).

[4] F. F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, 3rd ed. Springer, 2016 (cit. on pp. 4–6, 9, 12, 14, 15, 23, 25, 27–29).

[5] D. G. Griffiths, *Introduction to Electrodynamics*, 4th ed. Cambridge University Press, 1999 (cit. on p. 4).

[6] F. Mandl, *Statistical Physics*, 2nd ed. Wiley, 2023 (cit. on pp. 4, 30).

[7] IAEA, *Iaea nuclear data services*, `https://www-nds.iaea.org/`, Accessed: 2023-11-20, Year Published: 2007/ Last Updated: 2023 (cit. on p. 5).

[8] K. A. Brueckner, *Inertial Confinement Fusion*, 1st ed. Springer, 1998 (cit. on p. 4).

[9] S. Geng, "An overview of the iter project," *Journal of Physics: Conference Series*, vol. 2386, no. 1, p. 012 012, 2022. DOI: `10.1088/1742-6596/2386/1/012012`. [Online]. Available: `https://dx.doi.org/10.1088/1742-6596/2386/1/012012` (cit. on p. 4).

[10] E. Moses and the NIC Collaborators, "The national ignition campaign: Status and progress," *Nuclear Fusion*, vol. 53, no. 10, p. 104 020, 2013. DOI: `10.1088/0029-5515/53/10/104020`. [Online]. Available: `https://dx.doi.org/10.1088/0029-5515/53/10/104020` (cit. on pp. 5, 6).

[11] J. Lilley, *Nuclear Physics: Principles and Applications*, 1st ed. Wiley, 2001 (cit. on p. 6).

[12] A. Chagnes and J. Swiatowska, *Lithium Process Chemestry: Resources, Extraction, Batteries and Recycling*, 1st ed. Elsevier, 2015 (cit. on p. 6).

[13] H. Xie, V. S. Chan, R. Ding *et al.*, "Evaluation of tritium burnup fraction for cfetr scenarios with core-edge coupling simulations," *Nuclear Fusion*, vol. 60, no. 4, p. 046 022, 2020. DOI: `10.1088/1741-4326/ab742b`. [Online]. Available: `https://dx.doi.org/10.1088/1741-4326/ab742b` (cit. on p. 6).

[14] Y. LeCun, B. Boser, J. Denker *et al.*, "Handwritten digit recognition with a back-propagation network," in *Advances in Neural Information Processing Systems*, D. Touretzky, Ed., vol. 2, Morgan-Kaufmann, 1989 (cit. on p. 7).

[15] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278 –2324, Dec. 1998. DOI: 10.1109/5.726791 (cit. on p. 7).

[16] D. Silver, A. Huang, C. Maddison *et al.*, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, pp. 484–489, Jan. 2016. DOI: 10.1038/nature16961 (cit. on p. 7).

[17] T. Brown, B. Mann, N. Ryder *et al.*, "Language models are few-shot learners," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan and H. Lin, Eds., vol. 33, Curran Associates, Inc., 2020, pp. 1877–1901 (cit. on p. 7).

[18] A. Ramesh, M. Pavlov, G. Goh *et al.*, "Zero-shot text-to-image generation," in *Proceedings of the 38th International Conference on Machine Learning*, M. Meila and T. Zhang, Eds., ser. Proceedings of Machine Learning Research, vol. 139, PMLR, 2021, pp. 8821–8831 (cit. on p. 7).

[19] A. Esteva, B. Kuprel, R. A. Novoa *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–118, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:3767412 (cit. on p. 7).

[20] V. Gulshan, L. Peng, M. Coram *et al.*, "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," *JAMA*, vol. 316, no. 22, pp. 2402–2410, Dec. 2016, ISSN: 0098-7484. DOI: 10.1001/jama.2016.17216. eprint: https://jamanetwork.com/journals/jama/articlepdf/2588763/joi160132.pdf. [Online]. Available: https://doi.org/10.1001/jama.2016.17216 (cit. on p. 7).

[21] O. Dehzangi and M. Farooq, "Ssvep recognition using discriminative score fusion and transformation for portable brain computer interface," in *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, 2018, pp. 1–4. DOI: 10.1109/BHI.2018.8333355 (cit. on p. 7).

[22] J. F. Gemmeke, D. P. W. Ellis, D. Freedman *et al.*, "Audio set: An ontology and human-labeled dataset for audio events," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 776–780. DOI: 10.1109/ICASSP.2017.7952261 (cit. on p. 7).

[23] J. Kates-Harbeck, A. Svyatkovskiy and W. Tang, "Predicting disruptive instabilities in controlled fusion plasmas through deep learning," *Nature*, vol. 568, Apr. 2019. DOI: 10.1038/s41586-019-1116-4 (cit. on p. 7).

[24] S. Dasbach and S. Wiesen, "Towards fast surrogate models for interpolation of tokamak edge plasmas," *Nuclear Materials and Energy*, vol. 34, p. 101 396, 2023, ISSN: 2352-1791. DOI: https://doi.org/10.1016/j.nme.2023.101396. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352179123000352 (cit. on p. 7).

[25] R. Agarwal, L. Melnick, N. Frosst *et al.*, "Neural additive models: Interpretable machine learning with neural nets," in *Advances in Neural Information Processing Systems*, A. Beygelzimer, Y. Dauphin, P. Liang and J. W. Vaughan, Eds., 2021. [Online]. Available: https://openreview.net/forum?id=wHkKTW2wrmm (cit. on pp. 7, 45).

[26] S.-M. Udrescu and M. Tegmark, "Ai feynman: A physics-inspired method for symbolic regression," *Science Advances*, vol. 6, no. 16, eaay2631, 2020. DOI: 10.1126/sciadv.aay2631. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.aay2631. [Online]. Available: https://www.science.org/doi/abs/10.1126/sciadv.aay2631 (cit. on pp. 7, 45).

[27] J. Babcock, J. Kramar and R. Yampolskiy, "The AGI Containment Problem," *arXiv e-prints*, arXiv:1604.00545, arXiv:1604.00545, Apr. 2016. DOI: 10.48550/arXiv.1604.00545. arXiv: 1604.00545 [cs.AI] (cit. on p. 7).

[28] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman and D. Mané, *Concrete problems in ai safety*, 2016. arXiv: 1606.06565 [cs.AI] (cit. on p. 7).

[29] J. P. Freidberg, *Plasma Physics and Fusion Energy*, 1st ed. Cambridge University Press, 2007 (cit. on pp. 21, 23, 26–30).

[30] D. S. D. George V. Pereverzev and T. P. Group, *Plasma Disruptions in Tokamaks*, 1st ed. Springer, 2004 (cit. on p. 24).

[31] H. Xie, V. S. Chan, R. Ding *et al.*, "Evaluation of tritium burnup fraction for cfetr scenarios with core-edge coupling simulations," *Nuclear Fusion*, vol. 60, no. 4, p. 046 022, 2020. DOI: 10.1088/1741-4326/ab742b. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/ab742b (cit. on p. 26).

[32] G. A. Houlberg, *Radio Frequency Plasma Heating*, 1st ed. CRC Press, 1993 (cit. on p. 27).

[33] R. Pasqualotto, P. Nielsen, C. Gowers *et al.*, "High resolution Thomson scattering for Joint European Torus (JET)," *Review of Scientific Instruments*, vol. 75, no. 10, pp. 3891–3893, Oct. 2004 (cit. on pp. 27, 49).

[34] L. Frassinetti, M. N. A. Beurskens, R. Scannell *et al.*, "Spatial resolution of the JET Thomson scattering system," *Review of Scientific Instruments*, vol. 83, no. 1, p. 013 506, Jan. 2012 (cit. on pp. 27, 49).

[35] A. Sirinelli, B. Alper, C. Bottereau *et al.*, "Multiband reflectometry system for density profile measurement with high temporal resolution on JET tokamak," *Review of Scientific Instruments*, vol. 81, no. 10, p. 10D939, Oct. 2010 (cit. on pp. 27, 28, 49).

[36] E. Fransson, "Doctoral thesis," Ph.D. dissertation, Chalmers University of Technology, 2023 (cit. on pp. 31, 32).

[37] J. Weiland, *Stability and Transport in Magnetic Confinement Systems*, 1st ed. Springer, 2012 (cit. on p. 31).

[38] G. M. Staebler, J. E. Kinsey and R. E. Waltz, "Gyro-Landau fluid equations for trapped and passing particles," *Physics of Plasmas*, vol. 12, no. 10, p. 102 508, 2005 (cit. on p. 31).

[39] C. Bourdelle, X. Garbet, F. Imbeaux *et al.*, "A new gyrokinetic quasilinear transport model applied to particle transport in tokamak plasmas," *Physics of Plasmas*, vol. 14, no. 11, p. 112 501, Nov. 2007 (cit. on pp. 31, 48).

[40] G. Dong, X. Wei, J. Bao, G. Brochard, Z. Lin and W. Tang, "Deep learning based surrogate models for first-principles global simulations of fusion plasmas," *Nuclear Fusion*, vol. 61, no. 12, p. 126 061, 2021. DOI: 10.1088/1741-4326/ac32f1. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/ac32f1 (cit. on p. 32).

[41] D. P. Coster, V. Basiuk, G. Pereverzev *et al.*, "The european transport solver," *IEEE Transactions on Plasma Science*, vol. 38, no. 9, pp. 2085–2092, 2010. DOI: 10.1109/TPS.2010.2056707 (cit. on pp. 32, 47).

[42] F. Wagner, G. Becker, K. Behringer *et al.*, "Regime of improved confinement and high beta in neutral-beam-heated divertor discharges of the asdex tokamak," *Phys. Rev. Lett.*, vol. 49, pp. 1408–1412, 19 1982. DOI: 10.1103/PhysRevLett.49.1408. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.49.1408 (cit. on p. 32).

[43] A. W. Leonard, "Edge-localized-modes in tokamaksa)," *Physics of Plasmas*, vol. 21, no. 9, p. 090 501, Sep. 2014 (cit. on p. 33).

[44] C. Maggi, L. Frassinetti, L. Horvath *et al.*, "Studies of the pedestal structure and inter-elm pedestal evolution in jet with the iter-like wall," *Nuclear Fusion*, vol. 57, no. 11, p. 116 012, 2017 (cit. on pp. 33, 48).

[45] L. Frassinetti, S. Saarelma, G. Verdoolaege *et al.*, "Pedestal structure, stability and scalings in jet-ilw : The eurofusion jet-ilw pedestal database," *Nuclear Fusion*, vol. 61, no. 1, 016001, 2021, QC 20210113. DOI: 10.1088/1741-4326/abb79e (cit. on pp. 33, 35, 36, 47).

[46] L. Gil, C. Silva, T. Happel *et al.*, "Stationary elm-free h-mode in asdex upgrade," *Nuclear Fusion*, vol. 60, no. 5, p. 054 003, 2020. DOI: 10.1088/1741-4326/ab7d1b. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/ab7d1b (cit. on p. 33).

[47] E. J. Doyle, R. J. Groebner, K. H. Burrell *et al.*, "Modifications in turbulence and edge electric fields at the L–H transition in the DIII-D tokamak," *Physics of Fluids B: Plasma Physics*, vol. 3, no. 8, pp. 2300–2307, Aug. 1991 (cit. on p. 34).

[48] J. W. Connor, R. J. Hastie, H. R. Wilson and R. L. Miller, "Magnetohydrodynamic stability of tokamak edge plasmas," *Physics of Plasmas*, vol. 5, no. 7, pp. 2687–2700, Jul. 1998, ISSN: 1070-664X. DOI: 10.1063/1.872956. eprint: https://pubs.aip.org/aip/pop/article-pdf/5/7/2687/12753690/2687_1_online.pdf. [Online]. Available: https://doi.org/10.1063/1.872956 (cit. on p. 34).

[49] P. B. Snyder, R. J. Groebner, A. W. Leonard, T. H. Osborne and H. R. Wilson, "Development and validation of a predictive model for the pedestal heighta)," *Physics of Plasmas*, vol. 16, no. 5, p. 056 118, May 2009, ISSN: 1070-664X. DOI: 10.1063/1.3122146. eprint: https://pubs.aip.org/aip/pop/article-pdf/doi/10.1063/1.3122146/14032267/056118_1_online.pdf. [Online]. Available: https://doi.org/10.1063/1.3122146 (cit. on pp. 35, 47).

[50] S. Saarelma, J. Connor, P. Bilkova *et al.*, "Testing a prediction model for the h-mode density pedestal against jet-ilw pedestals," *Nuclear Fusion*, vol. 63, no. 5, p. 052 002, 2023. DOI: 10.1088/1741-4326/acc084. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/acc084 (cit. on p. 35).

[51] R. Groebner, M. Mahdavi, A. Leonard *et al.*, "Comparison of h-mode barrier width with a model of neutral penetration length," *Nuclear Fusion*, vol. 44, no. 1, p. 204, 2003. DOI: 10.1088/0029-5515/44/1/022. [Online]. Available: https://dx.doi.org/10.1088/0029-5515/44/1/022 (cit. on p. 35).

[52] C. Challis, J. Garcia, M. Beurskens *et al.*, "Improved confinement in jet high plasmas with an iter-like wall," *Nuclear Fusion*, vol. 55, no. 5, p. 053 031, 2015. DOI: 10.1088/0029-5515/55/5/053031. [Online]. Available: https://dx.doi.org/10.1088/0029-5515/55/5/053031 (cit. on p. 35).

[53] E. Stefanikova, L. Frassinetti, S. Saarelma *et al.*, "Effect of the relative shift between the electron density and temperature pedestal position on the pedestal stability in jet-ilw and comparison with jet-c," *Nuclear Fusion*, vol. 58, no. 5, p. 056 010, 2018. DOI: 10.1088/1741-4326/aab216. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/aab216 (cit. on p. 35).

[54] C. Bowman, D. Dickinson, L. Horvath *et al.*, "Pedestal evolution physics in low triangularity jet tokamak discharges with iter-like wall," *Nuclear Fusion*, vol. 58, no. 1, p. 016 021, 2017. DOI: 10.1088/1741-4326/aa90bc. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/aa90bc (cit. on p. 35).

[55] J. Cordey, for the ITPA H-Mode Database Working Group and the ITPA Pedestal Database Working Group, "A two-term model of the confinement in elmy h-modes using the global confinement and pedestal databases," *Nuclear Fusion*, vol. 43, no. 8, p. 670, 2003. DOI: 10.1088/0029-5515/43/8/305. [Online]. Available: https://dx.doi.org/10.1088/0029-5515/43/8/305 (cit. on p. 35).

[56] L. Frassinetti, C. P. von Thun, B. Chapman-Oplopoiou *et al.*, "Effect of the isotope mass on pedestal structure, transport and stability in d, d/t and t plasmas at similar n and gas rate in jet-ilw type i elmy h-modes," *Nuclear Fusion*, vol. 63, no. 11, p. 112 009, 2023. DOI: 10.1088/1741-4326/acf057. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/acf057 (cit. on p. 35).

[57] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org (cit. on pp. 36, 37, 51).

[58] R. C. M. C. M.-T. M. D. Shields K. Gurley and F. J. Masters, "Active learning applied to automated physical systems increases the rate of discovery," *Nature, Sci Rep*, no. 13, 2023 (cit. on p. 45).

[59] H. Nyström, L. Frassinetti, S. Saarelma *et al.*, "Effect of resistivity on the pedestal mhd stability in jet," *Nuclear Fusion*, vol. 62, no. 12, p. 126 045, 2022. DOI: 10.1088/1741-4326/ac9701. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/ac9701 (cit. on p. 47).

[60] A. Ho, J. Citrin, C. Bourdelle *et al.*, "Neural network surrogate of QuaLiKiz using JET experimental data to populate training space," *Physics of Plasmas*, vol. 28, no. 3, p. 032 305, Mar. 2021, ISSN: 1070-664X. DOI: 10.1063/5.0038290. eprint: https://pubs.aip.org/aip/pop/article-pdf/doi/10.1063/5.0038290/12361366/032305_1_online.pdf. [Online]. Available: https://doi.org/10.1063/5.0038290 (cit. on p. 48).

[61] Y. Wei, J. Levesque, C. Hansen, M. Mauel and G. Navratil, "A dimensionality reduction algorithm for mapping tokamak operational regimes using a variational autoencoder (vae) neural network," *Nuclear Fusion*, vol. 61, no. 12, p. 126 063, 2021. DOI: 10.1088/1741-4326/ac3296. [Online]. Available: https://dx.doi.org/10.1088/1741-4326/ac3296 (cit. on p. 51).