

Daniel Fabian Böckenhoff

En Route Towards Heat Load Control for Wendelstein 7-X with Machine Learning Approaches

**IPP 2020-01
März 2020**

En Route Towards Heat Load Control for Wendelstein 7-X with Machine Learning Approaches

Inauguraldissertation

zur

Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. Nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Universität Greifswald

vorgelegt von
Daniel Fabian
Böckenhoff
geboren am 20.10.1988
in Detmold

Greifswald, 20.09.2019

Dekan Prof. Dr. Werner Weitschies

1. Gutachter Prof. Dr. Thomas Sunn Pedersen

2. Gutachter Prof. Dr. Tünde Fülöp

Tag der Promotion: 26.02.2020

“And finally [...]”

Monty Python - The meaning of life

Contents

1. Introduction	1
2. Long Term Objective	17
3. Accomplishments	25
3.1. Feature Reconstruction and Heat Load Pre-Processing . . .	25
3.1.1. t reconstruction from limiter heat loads	27
3.1.2. t - ΔR reconstruction from divertor heat loads	29
3.2. Quantification of Critical Heat Loads	31
4. Résumé	33
Bibliography	37
Acronyms	45
List of Symbols	47
List of Figures	49
A. Thesis Articles	51
A.1. Article I	55
A.2. Article II	72
A.3. Article III	82
A.4. Article IV	96
A.5. Article V	114
B. CV	151
C. Acknowledgements	155

Contents

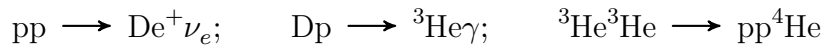
Statutory declaration

157

1

Introduction

Nuclear fusion has the potential to solve the world's energy needs with minimal environmental impact as compared to today's primary energy sources. The challenging physical problem with fusion is overcoming the mutual repulsion of the nuclei due to their positive charge. This is measured by the cross-section σ for proton-proton Coulomb scattering which is substantially higher as compared to the cross-section of all relevant fusion processes (see Figure 1.1). Although the cross-sections for the reactions of the proton-proton-chain



are marginal as compared to the Coulomb cross-section ($\sigma_{pp}/\sigma_{\text{Coulomb}} = \mathcal{O}(10^{-51} \text{ m}^2)/\mathcal{O}(10^{-27} \text{ m}^2)$ [1]^I), the pp-chain is the primary process igniting the stars. Reason is, that the thermonuclear reaction rate

$$r_{\text{nuc}} = n_1 n_2 \langle \sigma \vec{v} \rangle \tag{1.1}$$

scales quadratically with the nuclei's densities per unit volume n and their relative velocities \vec{v} which is proportional to the nucleus (or ion) temperature T_i . Gravitational confinement, self-generated by massive stars in the vacuum of space, lead to naturally high temperatures and especially tremendous densities^{II}. At ultimo, the small value of σ_{pp} provides the conditions

^IThe slowest process of the p-p chain gives the relevant cross-section since it represents a bottleneck for the reaction. Since pp fusion is a process of the weak interaction via the interchange of a W^+ -boson, the cross-section is orders of magnitude below that of Dp and ${}^3\text{He}{}^3\text{He}$ which are processes of the strong interaction.

^{II} $n = 5 \cdot 10^{31} \text{ m}^{-3}$ and $T_i = 1.4 \text{ keV}$ in the sun's core [2]

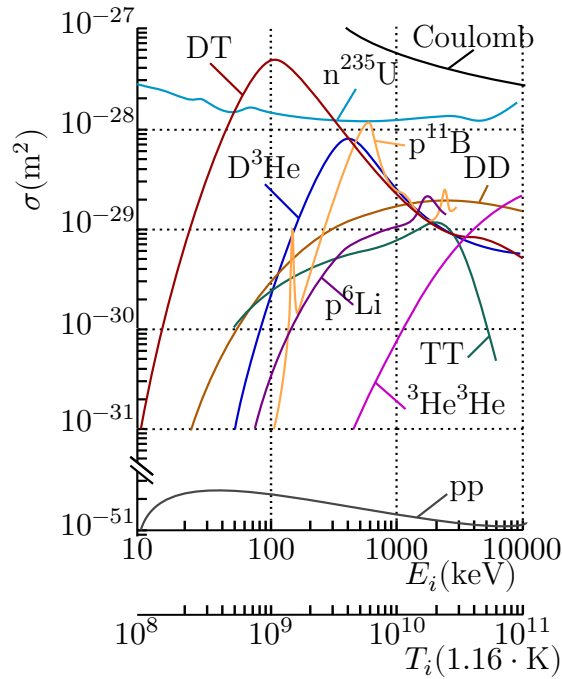
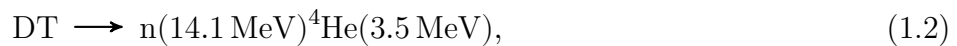


Figure 1.1.: Cross-sections of various nuclear fusion reactions as a function of the center of mass energy of the two particles. For comparison, the cross-section of the much more probable coulomb interaction and the neutron induced fission process $n^{235}\text{U}$ is added. Note the broken axis due to the factor of above 10^{20} lower maximum cross-section for pp fusion as compared to any other fusion reaction relevant for earth bound fusion power plants. The pD cross-section is on the order of 10^{-32} m^2 for the displayed energy range and thus not visible. Data from various sources [1, 3, 4]

necessary for planetary formation and thus the genesis of life since higher reaction rates limit the lifetime of stars drastically as observed in super heavy stars.

However, these reaction rates are not achievable with mankind's technology and fusion processes with much higher cross-sections need to be chosen for a nuclear fusion power plant. Besides the Coulomb and proton-proton chain cross-sections, Figure 1.1 depicts possible choices. Due to resonant tunneling [5], the deuterium-tritium reaction



has the favorable maximum cross-section at a center of mass energy of 100 keV. Attempting the most promising thermonuclear DT fusion, this corresponds to a maximum reaction rate at $T_i \approx 60$ keV because the reactions occur in the high energy tail of the velocity distribution. With a sophisticated particle confinement this process can result in a positive energy balance. Lawson showed that a triple product of n_i , T_i and the energy confinement time τ_E above approximately^{III} 3×10^{21} keV m⁻³ for $T_i \in (10, 40$ keV) is required for net energy output [6, 7].

Pushing up the value of $nT_i\tau_E$ towards this limit is the subject of research since the 1950s and remarkable progress has been made [8]. Reaching the required temperatures of $\mathcal{O}(10$ keV), corresponding to $\mathcal{O}(100 \times 10^6$ K), precludes confinement by reactor walls and alternatives are sought.

Magnetic confinement is a natural approach in the sense that at such temperatures, the material is in the plasma state and electrons and ions are exposed to electric fields \vec{E} and magnetic fields \vec{B} which are coupled with electric charge and current density ρ and \vec{j} by the Maxwell equations

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (\text{Gauss's law}) \quad (1.3)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (\text{Gauss's law for magnetism}) \quad (1.4)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (\text{Maxwell-Faraday equation}) \quad (1.5)$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \left(\vec{j} + \epsilon_0 \frac{\partial \vec{E}}{\partial t} \right) \quad (\text{Ampère's law}). \quad (1.6)$$

A common approach in plasma physics is to treat the whole plasma as a single fluid with pressure p , viscosity $\boldsymbol{\pi}$ and flow velocity \vec{v}_f . In the limit of low plasma resistance^{IV} and low frequency phenomena the so called magnetohydrodynamic (MHD) equations can be formulated as an addition

^{III}The precise value of $n_i T_i \tau_E$ depends on the profiles of n_i and T_i .

^{IV}At fusion relevant temperatures this limit is well fulfilled since the conductivity parallel to the magnetic field lines scales with $T_e^{3/2}$ [9]

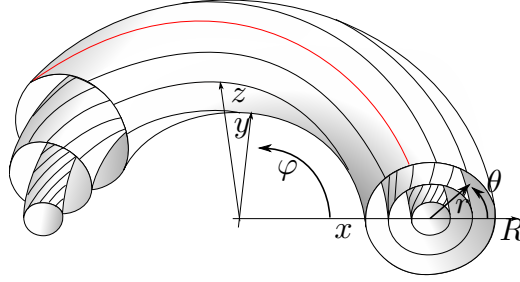


Figure 1.2.: Field lines generating nested magnetic surfaces.

to Equations (1.3) and (1.6)

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \rho \vec{v}_f = 0 \quad (\text{Continuity equation}) \quad (1.7)$$

$$\rho \left(\frac{\partial}{\partial t} + \vec{v}_f \cdot \vec{\nabla} \right) \vec{v}_f = \vec{j} \times \vec{B} - \vec{\nabla} p - \vec{\nabla} \cdot \boldsymbol{\pi} \quad (\text{MHD equation of motion}). \quad (1.8)$$

The MHD equation of motion basically resembles the Navier-Stokes equation [10]. In a steady state without flows, we can constrain

$$\frac{\partial}{\partial t} = 0, \vec{v}_f = 0 \Rightarrow \boldsymbol{\pi} = 0$$

Thus Equation (1.8) yields

$$\vec{j} \times \vec{B} = \vec{\nabla} p. \quad (1.9)$$

Solving this equation has been and still is objective of research and multiple numerical tools such as e.g. variational moments equilibrium code (VMEC) [11] have been created in this effort. Equation (1.9) implies, that both \vec{B} and \vec{j} lie in surfaces of constant pressure which must have toroidal topology (Poincaré-Hopf theorem [12–14]). Magnetic field lines with different pressure labels p form nested surfaces of constant pressure. Such surfaces are called flux surfaces.

With this found, we can define a coordinate system (p, φ, θ) , utilizing the pressure label and the toroidal topology with poloidal angle θ and toroidal angle φ (Figure 1.2). In these coordinates together with Equation (1.4), the magnetic field can be simplified to

$$\vec{B} = \vec{\nabla} \Psi \times \vec{\nabla} \theta + \vec{\nabla} \varphi \times \vec{\nabla} \chi, \quad (1.10)$$

with the toroidal and poloidal magnetic flux

$$\Psi = \int_{S:\varphi=\text{const.}} \vec{B} d\vec{S} \quad (1.11)$$

$$\chi = \int_{S:\theta=\text{const.}} \vec{B} d\vec{S} \quad (1.12)$$

$$= \int \vec{\nabla}_\theta \vec{B} dV \quad (1.13)$$

respectively [15, 16]. In these magnetic coordinates, magnetic field lines are conveniently described as straight lines.

The change of the toroidal flux with respect to the poloidal flux

$${}_\iota(\Psi) = \frac{d\chi}{d\Psi} \quad (1.14)$$

is called the rotational transform. In direct consequence of equations 1.14 and 1.10 we can also write

$${}_\iota(\Psi) = \frac{d\theta}{d\varphi}. \quad (1.15)$$

Magnetic field lines can be labeled by two coordinates $(\Psi, \alpha = \theta - \iota\varphi)$. Tracing a certain field line one toroidal turn, we find a poloidal angle difference of $\Delta\theta = 2\pi\iota$ and thus can understand the nature of the rotational transform. If $\iota \in \mathbb{Q} : \iota = n/m; n, m \in \mathbb{N}$, one field line closes in itself after m toroidal and n poloidal turns and a so called resonant surface establishes. In the presence of radial magnetic perturbations which are periodical in φ ^V magnetic islands are generated at such rational values of ι . In first order, the radial island width can be approximated^{VI} by

$$\Delta r = 4 \sqrt{\frac{rR\delta B_r}{nsB_\varphi}}, \quad (1.16)$$

where

$$s = -r \frac{d\iota}{dr} / \iota \quad (1.17)$$

^VSuch perturbations can be inherent to the ideal magnetic field of a device (in the case of stellarators) or can be caused by magnetic error fields.

^{VI}Assuming circular poloidal cross-section and large aspect ratio

is the magnetic shear and δB_r denotes the magnitude of the radial magnetic perturbation [16].

If otherwise $\iota \in \mathbb{R} \setminus \mathbb{Q}$, one field line defines the whole flux surface. In this case, a direct consequence of »Weyl's lemma« [17] is, that one field line followed an arbitrary number $k \in \mathbb{N}$ of times around the torus covers the whole flux surface to any desired precision, depending on k . This finding allows to display a rational magnetic surface in a plane by tracing one point on this surface k times around the torus and displaying the piercing points in a so called Poincaré representation (see Figure 1.4 as an illustration by the example of the later introduced Wendelstein 7-X magnetic field).

The force balance for charged particles in a magnetized plasma with a constant external force \vec{F} can be written as a differential equation of first order in time

$$\frac{d(m\vec{v})}{dt} - q(\vec{v} \times \vec{B}) = \vec{F}. \quad (1.18)$$

The solution to the homogeneous part of the equation describes the gyration of the particle around a guiding center with cyclotron frequency $\omega_c = qB/m$ and Larmor radius $r_l = v_\perp/\omega_c$. Hence, in an otherwise force free environment, a particle is confined perpendicular to a gradient free magnetic field. In this regard, the freedom of the charged particle parallel to \vec{B} is of no relevance for particle confinement in toroidal magnetic fields. However, \vec{B} forming toroidal flux surfaces gives rise to $\vec{\nabla}|\vec{B}|$ which corresponds to a radial force $F_{\vec{\nabla}|\vec{B}|}$ as well as centrifugal forces F_c . The solution to the inhomogeneous part of 1.18 describes the drift of the guiding center with constant velocity

$$\vec{v}_d = \frac{1}{q} \frac{\vec{F} \times \vec{B}}{|\vec{B}|^2}. \quad (1.19)$$

Thus the radial forces give rise to an up down charge separation accompanied by large magnitudes in \vec{E} . This electric field gives rise to a force $\vec{F} = q\vec{E}$ which, substituted in Equation (1.19), corresponds to a strong drift in radial direction which puts a tight limit on plasma confinement.

To compensate the charge separation due to the guiding center drifts, toroidal magnetic confinement devices need a finite and carefully tuned rotational transform ι . Since \vec{B} is a superposition of the field which is externally

applied by magnetized coils as well as internal toroidal currents, we can also describe ι by separating these two parts:

$$\iota = \frac{\mu_0}{S_{11}} \cdot \frac{I_{\text{tor}}}{\chi} + \iota_{\text{CF}}, \quad (1.20)$$

where

$$I_{\text{tor}} = \int_0^r \vec{j}(r') \cdot \hat{e}_\varphi \frac{V(r')}{2\pi R_0} dr' \quad (1.21)$$

is the toroidal current enclosed by the magnetic surface at r and

$$\iota_{\text{CF}} = -\frac{S_{12}}{S_{11}}. \quad (1.22)$$

The components of the susceptance matrix $S = [S_{ij}]$ only depend on external coil currents and coil geometry and are thus independent of internal plasma currents, hence the index »CF« which abbreviates »current free« [18].

Toroidal devices can be classified into two categories, distinguished by the generation of the poloidal field component, i. e. the main contributing term in Equation (1.20).

In machines of the Tokamak class the magnetic field is a superposition of an externally imposed toroidal field ($S_{11} \neq 0$, $S_{12} = 0 \Rightarrow \iota_{\text{CF}} = 0$) and a poloidal field which is generated by a toroidal current I_{tor} induced by the change of flux in the central solenoid coil acting as primary winding of a transformer. Tokamaks possess relatively good confinement properties but the need to sustain a toroidal current is a major drawback in two ways. Until I_{tor} can be driven without transformer action, the devices are pulsed introducing a huge loss in efficiency. Furthermore, the toroidal current drives instabilities such as disruptions which can potentially harm all components and especially the first wall of future large devices.

Machines of the stellarator class give rise to the poloidal field component purely by external coils ($S_{11} \neq 0$, $S_{12} \neq 0 \Rightarrow \iota_{\text{CF}} \neq 0$) without the necessity of large I_{tor} and thus above mentioned downsides. Nonetheless, plasma confinement in the early stellarators was disappointing. As an example particle orbits with low collisionality, i. e. high energy could not be confined for a long enough time. The fusion-born fast alpha particles fall within that

category and their confinement time is short as compared to the slowing down time, and thus they are lost for plasma self-heating in the early stellarators [16]. However, advances in plasma theory and computation power, in particular in the 1980s and 1990s, introduced the possibility of stellarator optimization, enabling significantly improved confinement.

Wendelstein 7-X (W7-X) is the currently most advanced stellarator, among other properties also optimized for a fast particle confinement, small radial magnetic plasma shift ΔR and a small toroidal current I_{tor} .

I_{tor} is suppressed by reduction of the dominant contribution by means of the bootstrap current I_{bs} . This current is diffusion-driven and result of the interaction between trapped and passing particles in a plasma with finite density and temperature gradient which can be described by neoclassical theory [15, 16, 19, 20]. Due to the very low resistance R of a plasma, scaling with $T_e^{-3/2}$ [21], and the high inductance L , I_{tor} does not build up instantaneous but converges like

$$I_{\text{tor}} = \Theta(t = 0) \cdot I_0 (1 - e^{-t/\tau}), \quad (1.23)$$

as counter currents, shielding the bootstrap current, decay. $\tau = L/R$ has typical values of $\mathcal{O}(10\text{ s})$ for fusion relevant plasmas. For W7-X, I_{bs} is reduced by careful combination of helical components as well as toroidal curvature effects in \vec{B} for every magnetic surface [22]. However, I_{bs} and thus ultimately I_{tor} can still have values of $\mathcal{O}(10\text{ kA})$ which significantly affect the ι profile, potentially causing undesired islands at resonant ι and interfere with island divertor operation (see below).

Some contributions to the aforementioned radial shift ΔR of the magnetic surfaces are constant throughout the discharge and can be corrected for [23]. »A more dangerous displacement is that due to the increase of pressure during the heating of the plasma« called Shafranov shift after its discoverer. »[...] when the pressure varies from zero to $B^2/(2\mu_0)$, this displacement is seen [...] to amount to several centimeters« [23]. As the Shafranov shift is a magnetic configuration effect it can be reduced in stellarators [24].

The magnetic field optimization and the subsequent coil optimization resulted in the coil set depicted in Figure 1.3a with a major radius of $R \approx 5.5\text{ m}$, an aspect ratio of $\epsilon = R/r_a \approx 10$ (plasma radius r_a) and $N_p = 5$ field periods. These five periods can be thought of as constituting five magnetic

mirrors connected such that particles in the loss cone transverse into the mirror of the adjacent module [25]. Each field period (module) possesses an inner flip symmetry (two half-modules), such that the device could be build from ten almost identical half-modules.

Every half-module contains five superconducting non planar coils (1, ..., 5) and two planar coils (A, B) (labeled in Figure 1.3b). This main coil system is capable of generating a magnetic field of 3 T on axis (most experiments run at 2.5 T). All superconducting coils of one type are energized in a joint electrical circuit in series respectively [26]. The non planar coils each consist of $n_{\{1,\dots,5\},w} = 108$ windings of copper stabilized NbTi strands cooled by supercritical He. Planar coils use the same superconductor with $n_{\{A,B\},w} = 36$ windings respectively.

Additionally every module contains a pair of conventional coils, usually referred to as control coils. Although the total of ten control coils are independently powered, they are usually run, following the five-fold rotational symmetry of W7-X (magnetic field of one period describes the complete magnetic field). This reduces the freedom to two independent control coils labeled S1 and S2 for the upper and lower control coil respectively («S» referring to the alternative name »sweep coils«).

This leaves in total nine coil currents which are free to choose. Without loss of generality, the coil currents can be defined as normalized with respect to the non-planar coil 1 and the actually applied coil currents $I_{x,\text{true}}$ are multiplied with their respective number of windings $n_{x,w}$ such that

$$I_x = \frac{I_{x,\text{true}} \cdot n_{x,w}}{I_n}, \forall x \in \{1, \dots, 5, A, B, S1, S2\}. \quad (1.24)$$

In such a way, the magnetic field strength is proportional to $I_n = I_{1,\text{true}} \cdot n_{1,w}$ and the value for I_x is reflecting the amount of magnetic field generated by coil type x .

Large magnetic islands are generally bad for confinement because they allow heat and particles to transit the island region by parallel transport rather than the slow perpendicular transport. In such a way they effectively »short-circuit« the radial confinement and which strongly increases the radial transport. Regarding Equation (1.16), two strategies are feasible in order to enforce small islands. The first is to use high shear to cross resonant ι fast (e. g. in the case of large helical device (LHD) [27]). W7-X belongs to

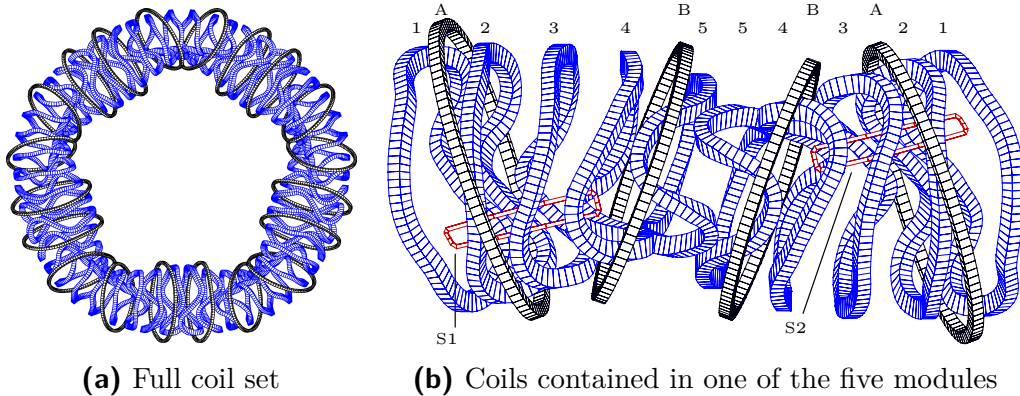


Figure 1.3.: The W7-X coil set. Blue: non planar coils. Black: planar coils.
Red: island control coils.

the second class of devices which avoid low order (small m and n) rationals and use low shear.

The net heating power $P_{\text{heat}}^{\text{VII}}$ is necessarily exhausted either by means of diffusive convection following the magnetic field topology or by radiation i. e.

$$P_{\text{heat}} = P_{\text{conv}} + P_{\text{rad}}.$$

A sophisticated power and particle exhaust concept is required for steady state confinement devices, especially for W7-X which aims for continuous operation with $P_{\text{heat}} = 10 \text{ MW}$ at reactor relevant performance i. e. high n and T_i . The convective power must be directed towards predetermined plasma facing components (PFCs) under controlled conditions to protect weak components and enable density control.

The most direct solution is the limiter concept, where the closed flux surfaces are directly intersected by a dedicated PFC. However, limiters have the disadvantage, that the plasma-wall interaction occurs in direct proximity to the plasma core. As a direct consequence, impurities (e. g. particles sputtered away from the PFC surface) are introduced to the core and neutralized particles can not easily be pumped away. W7-X went into operation at the end of 2015 with five graphite limiters mounted on the inboard side of the vacuum vessel (operation phase (OP) 1.1, see Figure 1.5a). However, this

^{VII}This includes external heating as well as alpha particle heating

minimal PFCs setup was only chosen for the start of operation.

More sophisticated exhaust concepts utilize magnetic edge topologies in order to »divert« core plasma particles which pass the last closed magnetic surface (LCMS) outwards towards the PFCs where the particles can be pumped away after being neutralized and the deposited energy is removed by active cooling. In low shear stellarators this can be done by allowing a large scale resonant island in the plasma edge which intersects a so called island divertor (see Figures 1.4 and 1.5b for the example of W7-X). In this way, the plasma wall interaction is spatially separated from the core plasma which allows efficient impurity screening and refueling under achievable circumstances [28]. Also, reduction of the peak PFC heat load is granted because the plasma carrying field lines in the island of low-shear stellarators are characterized by large connection lengths

$$L_c = \frac{\pi}{2t} \sqrt{\frac{rRB_\varphi}{s\delta B_r n}} \quad (1.25)$$

as described by [28]. Large L_c increase the role of cross-field transport such that the scrape-off layer (SOL) can be populated, leading to a spread in the power deposition profile and thus reduced heat loads.

The heat load pattern on the divertors is very sensitive to changes in the magnetic field topology since the island position, shape and size in the plasma is a highly non-linear function of t , \vec{j} , external coil currents and much more^{VIII}. The W7-X divertor concept foresees ten modular island divertors (one in the upper, one in the lower half of each module) with a geometry optimized for magnetic configurations with 4, 5 and 6 edge islands corresponding to rotational transforms $t = 5/4$, $5/5$ and $5/6$ respectively^{IX}. Each divertor consists of two horizontal target plates intermitted by a low loaded area as well as a vertical target plate. In between the horizontal and vertical target plates two gaps allow neutralized gas to stream towards turbo molecular vacuum pumps [29]. The PFCs surrounding the divertor (baffle, toroidal and poloidal closure - see Figure 1.5b) are designed to guide neutrals towards the pumpings and thus to increasing the pumping efficiency. The first divertors installed after the successfully completed OP 1.1 [30], were non-actively cooled test divertor units (TDUs) for OP 1.2, limiting the

^{VIII}With a change in the magnetic topology divertors can run in a limiter configuration.

^{IX}Also limiter configurations (0 islands at the edge) are foreseen

discharges to a total injected energy of 80 MJ. By operating divertors, a new regime in plasma parameters is achievable and detachment was observed [31]. Detachment is a highly favorable state for divertor operation. It describes the formation of an intermediate gaseous target in front of the divertor as the plasma temperature at the divertor targets is reduced to $\mathcal{O}(1 \text{ eV})$ and volume recombination becomes strong [32]. As a consequence the target heat load and erosion is greatly reduced because radiation cools the edge plasma and uniformly distributes the power to 4π solid angle. Furthermore, since the neutral pressure in front of the target plates is increased, pumping performance increases.

With OP 2 which is planned to start in 2021, W7-X is aiming to be the first machine reaching steady-state operation at fusion relevant plasma parameters [31] with discharge durations of up to 30 min and continuous injected electron cyclotron resonance heating (ECRH) power of up to $P_{\text{heat}} = 10 \text{ MW}$. Despite this sophisticated exhaust concept, heat loads expected for such scenarios are requiring PFC materials at the limit of feasibility. Maximum design heat loads of the most critical PFCs are listed in Table 1 of Article III. The divertor target plates can withstand a heat load of up to 10 MW m^{-2} ^X. However, the edge of the divertor plates close to the pumping gap as well as the low loaded area can not sustain such high heat loads due to reduced cooling [34]. Although the PFCs are designed such that the component design loads are not exceeded for various magnetic configurations and plasma parameters [35, 36], this does not mean, their integrity is ensured. Plasma wall interaction is dynamic since the magnetic topology changes with time due to various mechanisms. In leading order these are the two aforementioned effects:

ι evolution At plasma start-up, ι changes according to Equation (1.20) with the toroidal current which evolves, driven by I_{bs} , changing the radial island position. This change in the magnetic edge topology has a large effect onto the heat load pattern on the divertor and further PFCs for two reasons. First is the small shear in the W7-X ι profiles. Second, the divertor has been designed for shallow field line incident angles to maximize the area hit by the strike lines to reduce the heat load. As a result the heat load pattern is highly susceptible to small changes

^X15 % of the heat load on the surface of the sun [33] or about 1 kW m^{-2} at a distance of 1 au, i. e. the heat load at noon on the earth's equator.

in ι and strike lines (parts of the heat load pattern that are mostly elongated) can appear or disappear with a change in the magnetic edge topology.

Change in ΔR The plasma density and temperature profiles change, mainly due to varying heating or fueling e. g. by pellet injection. A change in $\vec{\nabla}p$ leads to a different I_{bs} (see above), which finally again changes ι . The main effect of plasma pressure change however, is the resulting radial shift of the magnetic surfaces ΔR . Similarly to the topology change due to ι , ΔR leads to heat load changes which are not necessarily described by translation only but can have non-linear effects due to shadowing.

Such non-linear variations of the heat load pattern do not only affect the divertor but potentially all PFCs and the heat load can exceed the design limits and thus result in overload. The implications of excessive heat loads are various, depending on the respective component.

Erosion »Material erosion has to be small enough to guarantee a sufficiently long lifetime of the most loaded components. In addition, the plasma contamination by impurities, which arises from the interplay between wall erosion and plasma transport, has to be compatible with a burning fusion plasma« [37]. As a result of erosion and the following re-deposition of carbon, surface layers are formed [38]. Surface layers are characterized by a very poor thermal connection to the bulk material, leading to a reduced exhaust and being a potential source of systematic uncertainties for heat loads retrieved through the Infra-Red (IR) imaging diagnostic [39]. Furthermore surface layers in future fusion devices will contain high tritium inventories also on less exposed areas of the vessel wall [40, 41]. For safety reasons, the total tritium inventory of a power plant must be limited.

Delamination The divertor is composed of carbon fibre reinforced carbon (CFC) mounted onto a water cooled CuCrZr heat sink with a interlayer of soft Cu to reduce mechanical stress from heat expansion. In case local temperatures exceed 475 °C, the resulting stress can lead to cracks in the interlayer (so called »delaminations«) which propagate with each heat cycle and drastically reduce the heat conductivity [42].

Besides the then reduced heat exhaust capabilities of the respective PFC, higher material temperatures are reached, resulting again in higher sputter yields [43] and erosion intensifies.

Water leakage Thermally induced mechanical stresses between the cooling pipe and the hot PFC-tiles may cause rupture of the cooling pipes to leak after a certain amount of cycles with a particular tile temperature. »The ingress of coolant might cause enough pressurization to break the vacuum vessel and a possible release of tritium and/or activated materials to the outside of the vacuum vessel« [44]. For W7-X a only 2 mm² partial break of a cooling pipe would lead to a pressure increase to ≈ 6 bar [45] and a burst disk is installed to prevent damage to the vacuum vessel. Moreover the introduction of oxygen to the system would open up the window for organic chemistry and surface erosion due to oxidation^{XI}.

^{XI}Which is particularly inconvenient in the case of later fusion reactors with neutron activated PFCs.

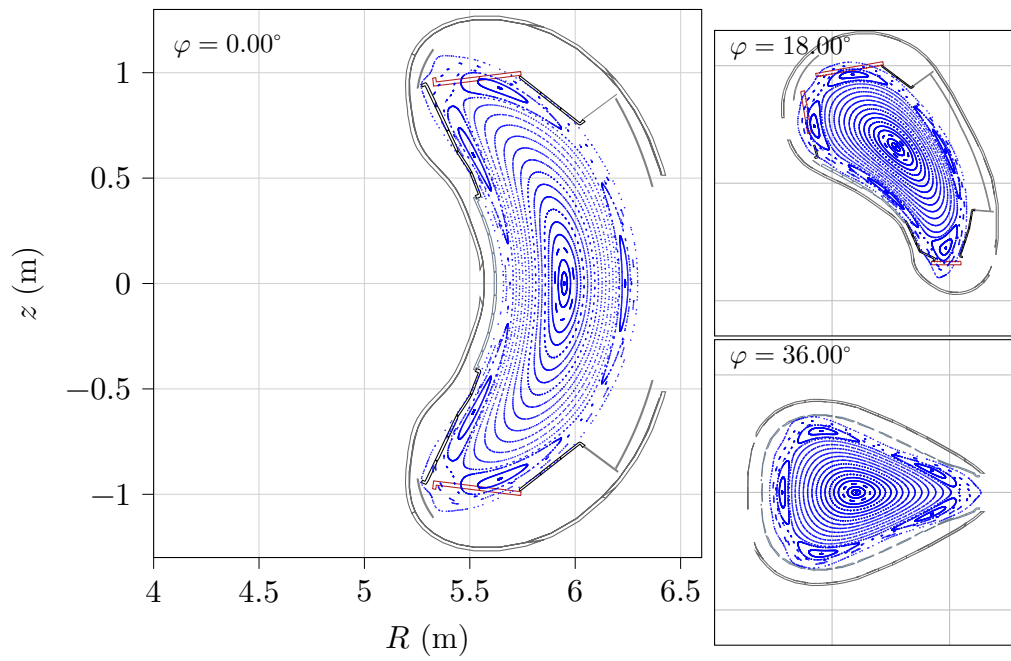


Figure 1.4.: Poincaré representation of the standard reference case magnetic field. Three sections of one half module ($\varphi = 0^\circ$ to 36°) are shown. The $\iota = n/m = 5/5$ island chain is clearly visible in the edge.

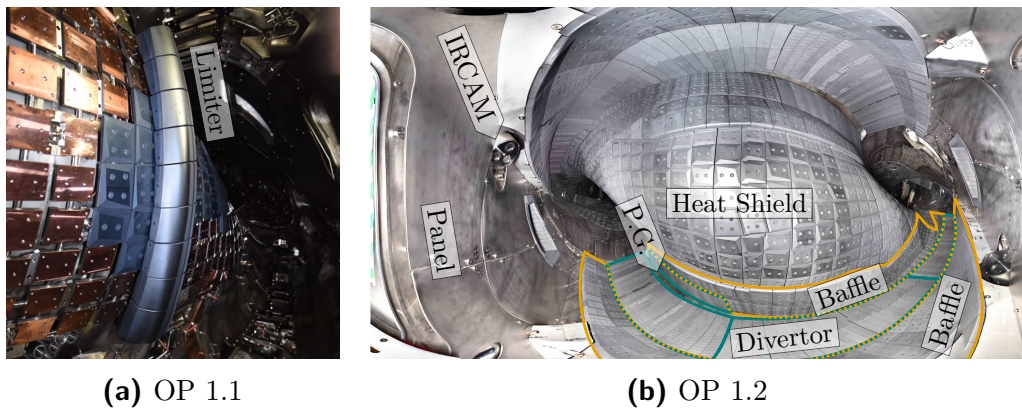


Figure 1.5.: PFCs in the two stages of OP 1. (a): In OP 1.1 a graphite limiter was installed on the inboard side of the vacuum vessel per module. A few baffle graphite tiles were clamped on the CuCrZr cooling structure of the heat shield in the ultimate vicinity of the limiter. (b): In OP 1.2 the same PFC geometries as in the final OP 2 were installed which are mainly the divertor, baffle, heat shield and panel, as labeled above. Additionally, the pumping gap (P.G.) and the end of the IR camera endoscope are indicated. For the upcoming OP 2 the components exposed to the largest heat loads are water cooled.

2

Long Term Objective

Long term objective of this work is to develop a real-time heat load control system which should be applied in the upcoming W7-X steady-state plasma discharges and could be modified for future steady-state machines. Not only should such a system be applied to ensure PFC integrity as motivated before. Plasma property optimization could be done with the same system. Subject to optimization could be every measure that can be retrieved in real time. Examples are:

Pumping performance By moving the strike-line close to the pumping gap the neutral pressure in the vicinity of the turbo pumps would be increased, leading to an increase in pumping efficiency.

Detachment control by means of hydrogen gas puff or impurity seeding with N and Ne. Various experiments in OP 1.2 were performed, demonstrating the possibility with yet pre-programmed settings [31, 46].

Triple product The ultimate goal of fusion research can be described by maximization of the triple product as described introductory. The line integrated density (measured by the interferometer) can be used as a real time signal scaling with the average density.

T_i is only measured by the charge exchange recombination spectroscopy (CXRS) or X-ray imaging crystal spectroscopy (XICS). CXRS requires the neutral beam injection (NBI) beam and hence is only available for 10s which makes it impractical as control diagnostic. XICS can lately deliver T_i in real time ($\mathcal{O}(10\mu\text{s})$) due to the fast evaluation by means of an artificial neural network (NN) [47] but requires

an Ar puff which is not foreseen in every discharge. The measurement of the energy confinement time is based on the energy conservation principle

$$\frac{dW}{dt} = P_{\text{heat}} - \frac{W}{\tau_E}. \quad (2.1)$$

The diamagnetic energy

$$W = \frac{3}{2} \int (n_e T_e + n_i T_i) d^3 r \stackrel{\text{H plasma}}{=} 3V\bar{n}(\bar{T}_e + \bar{T}_i) \quad (2.2)$$

is measured by diamagnetic loops [48] as a function of t . Real time evaluation of P_{heat} is feasible in many cases but more demanding in scenarios with multi-path absorption.

For all such operations it is evidently highly critical to use a coupled system of overload prevention in combination with the respective optimization or all. In fact the combination of multiple optimization goals into one control system would be highly beneficial since often a trade-off between various goals has to be found. For such a combination, the reinforcement learning (RL) approach introduced below is highly suited due to its flexibility and the custom reward description. It should be noted, though, that this is accompanied by the increase in complexity with the amount of sensors and actuators included.

The main sensors for the control system are ten IR cameras, each observing one divertor and the vicinity (especially the baffle tiles) [49]. With these, the heat load can be calculated by solving the heat diffusion equations [50]^I.

Essential actuators are listed in the following.

Island control coils can be used to directly change the island width. For example by applying alternating currents to the coils, the strike-line can be swept (hence the alternative name) and the peak PFC heat

^INote that solving this equation is computationally expensive and possibly not real time capable. At present effort is made to retrieve heat loads in real time. In the case that a numerical evaluation of the heat diffusion equations is not feasible in real time, it is considered to bypass this step by function approximation with supervised learning approaches.

load is reduced. The current in the sweep coils can be changed fast, only delayed by the resistance and inductance of the sweep coils (L/R time of 50 ms [51]).

Trim coils constitute an additional system of five conventional coils and allows the correction of error fields and (thus) balancing of divertor heat loads. This system can ramp with about 20 kA s^{-1} .

Superconducting coils Especially the use of planar coils would be an attractive control opportunity because they give a direct contribution to t and ΔR . However, due to the nature of superconductors, their current may not be changed fast and a limit of 30 A s^{-1} is set which usually corresponds to an impractically small relative current change of 2 ‰ s^{-1} .

Current drive by means of electron cyclotron current drive (ECCD) or neutral beam current drive (NBCD) adds additional freedom to change the rotational transform by manipulation of I_{tor} [52, 53]. It should be noted, that both actuators have limitations. ECCD can not be applied in the advanced ECRH O2-heating scenario because this mode does not allow full single pass power absorption. A multi-beam path-scenario is used in order to reflect shine through back through the plasma allowing for further absorption, which limits the flexibility of the heating and prevents current drive [54]. NBCD is limited by the pulse length of the neutral beams which is at maximum 10 s [55].

Fueling by gas puff and hydrogen ice pellets [56] can increase the plasma density while lowering its temperature. This is connected to a significant impact on especially the PFC heat load intensity but also the shape as motivated before.

Heating with the primary ECRH system or auxiliary ion cyclotron radiation heating (ICRH) and NBI determine P_{heat} and thus have a straightforward influence on the PFC heat load and the triple product. The NBI system also has a fueling effect.

Impurity seeding by fueling with high- Z gases is a possibility to reduce the local convective heat load by increasing P_{rad} .

Depending on the exact control system in question, suiting diagnostics and actuators are to be added (e. g. Bolometers [57] and gas puff for detachment control).

Plasma physics systems are chaotic, highly complex and not predictable in fine detail. For example, presently no theory exists, which allows the deduction of the energy confinement time from first principles and empirical scaling laws for τ_E are used [58]. A further example is perpendicular transport. Modern transport theory is a patchwork of various theories which apply to different scales of collisionality and turbulence. Although a lot of progress has been made, modern transport theory models do not adequately correspond to the measurements [59, 60] and calculations have to be performed with highly computationally expensive codes.

In this domain of highly complicated and computationally expensive plasma physics, machine learning (ML) tools are the adequate choice for real time evaluation and control.

Most ML algorithms today belong to the category of supervised learning. In supervised learning, a machine learning model is trained on a data-set containing pairs of inputs and targets, called training set, such that it provides a mapping which can be applied to new inputs, without knowing the target. A re-training of the model can account for an updated training set.

Highly suitable for the afore motivated case of complex and insufficiently known connection between input and target are model agnostic NNs which can approximate any continuous functions to any degree of accuracy as shown by Cybenko [61]. Artificial NNs are mathematically generalized representations of networks of neurons in nervous systems. Figure 2.1 shows a generic feed-forward fully-connected neural network (FF-FC). The network is built from individual neurons who's individual elements have corresponding counterparts in the biological neurons, shown in brackets in the following. Neuron inputs \vec{x} (potentials at dendrites) dampened or amplified by individual weights \vec{w} (length of dendrites) are summed (signal superposition in the cell body). The key point follows hereafter as this signal is non-linearly transformed by the activation function f (activation potential in the axon) so that the output of neuron j is

$$y_j = f(b_j + \vec{w}_j \vec{x}_j) = f(b_j + \sum_{k=1}^{\nu_x} w_{jk} x_{jk}). \quad (2.3)$$

This non-linearity is essential for the findings of Cybenko.

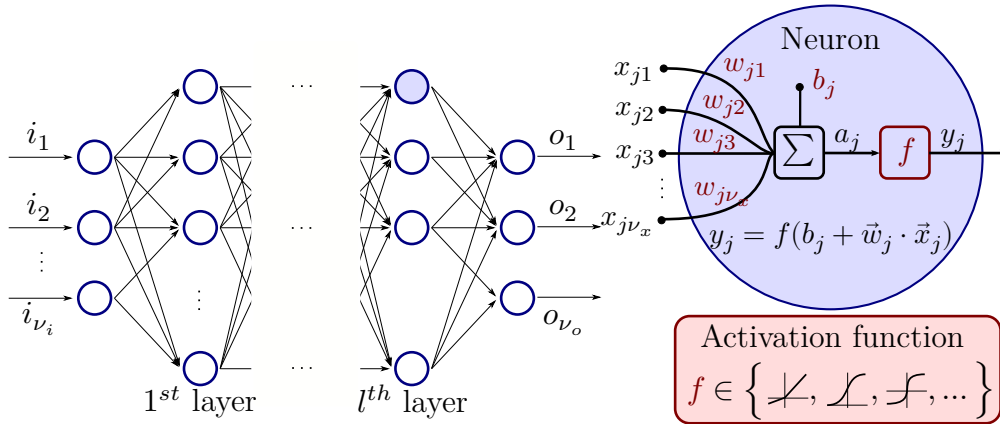


Figure 2.1.: Scheme of a simple NN (FF-FC). A zoom into the interior of one neuron is depicted in blue. Three examples of common activation functions are given in the red inlay. The non-linear nature of the activation functions is crucial.

The process of learning is done by adjusting \vec{w} by the amount

$$\Delta \vec{w} = \lambda \vec{\nabla}_{\vec{w}} \mathcal{L}, \quad (2.4)$$

a process called backpropagation. This requires differentiability of $f(\vec{x})$ because \mathcal{L} generally is not an explicit function of \vec{w} . A common choice for the so called loss function \mathcal{L} is

$$\mathcal{L} = \frac{1}{2} |\vec{o}(\vec{i}) - \vec{t}|^2, \quad (2.5)$$

where $\vec{o}(\vec{i})$ is the output given input \vec{i} and \vec{t} is the target the NN should approximate.

The architecture of NNs has tremendous impact on their abstraction capabilities. It can be shown both empirically as well as in theory that in order to learn complicated models that can represent high-level abstractions, deep (large number l of hidden layers) are superior to shallow architectures with comparable amounts of free parameters [62].

Based on the predictions of a model, further, mostly human-controlled, actions can be triggered.

In order to be able to use machine learning models in environments in which they can learn autonomous action-response events, learning al-

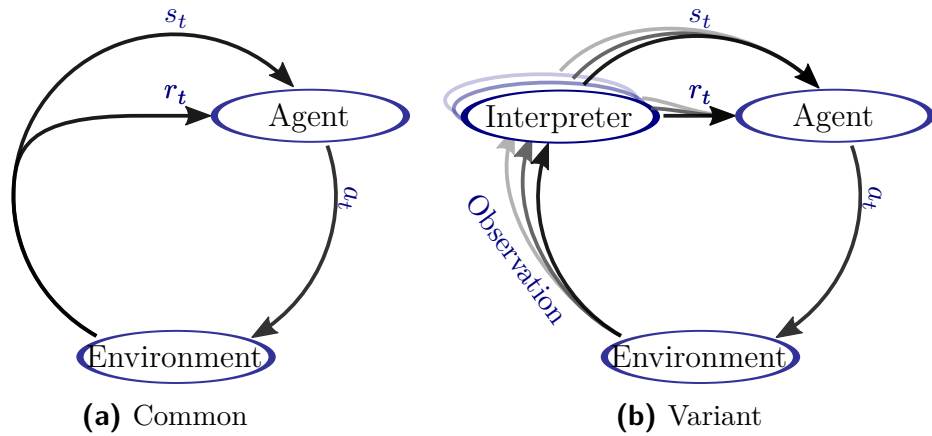


Figure 2.2.: RL schemes. The environment in (a) can be split into environment and (potentially multiple) interpreter as shown in (b).

gorithms are required that take into account the changing dynamics of the environment. This class of algorithms are called RL.

Reinforcement learning describes a learning ansatz which is similar to natural learning behavior of humans (and other vertebrates) and is a label for approaches to Markov decision processes [63, 64] with unknown reward functions. The principal scheme is depicted in Figure 2.2a. At time t , the environment is in a state s_t and a reward r_t can be assigned (small or negative rewards can be interpreted as punishments). An agent chooses an action a_t and thus manipulates the environment. This transfers the environment into a new state s_{t+1} and according reward r_{t+1} . The agent is trained to choose actions such that the cumulative reward is maximized. It is always in a dilemma between using its previously acquired experience on the one hand and exploring new strategies to increase rewards on the other. This is called the exploration-exploitation dilemma.

In the training phase, the agent chooses a random a_t and resulting (r_{t+1}, s_{t+1}) are used in order to update the expected total reward $Q(s_t, a_t)$ for choosing action a_t from state s_t . A major breakthrough in RL was the development of an update rule for the expected total reward

$$Q(s_t, a_t) = (1 - \alpha)Q(s_t, a_t) + \alpha(r_{t+1}(s_t, a_t) + \gamma \max_{a'} Q(s_{t+1}, a')). \quad (2.6)$$

called Q-learning [65, 66]. The parameters α and γ are hyperparameters. $\alpha \in (0, 1]$ is called learning rate and controls to which extent new information is used to update Q . $\gamma \in (0, 1]$ corresponds to the so-called discount factor and controls a trade-off between short-term and long-term reward maximization. It can be shown that the estimate Q converges towards the optimal function $Q \rightarrow Q^*$ as $t \rightarrow \infty$ [67]. After the learning phase, the agent is set to maximize the reward by always choosing the greedy optimal action which can be derived from Q as $a_t = \arg \max_{a'} Q(s_t, a')$.

A major challenge for RL applied to real-world problems is to decide how to represent and store Q . »Unless the states set is finite and small enough to allow exhaustive representation by a lookup table [...] one must use a parameterized function approximation scheme« [67]. At this point deep NNs come in handy and various modifications of standard Q-learning lead to latest impressive results [68, 69]. The concept is called deep Q-network (DQN) and resembles a hybrid between RL and supervised learning.

Curse of dimensionality Given a control problem with a high dimensional state and/or action space, training of a RL agent requires a data set size which scales exponentially with the dimension in order to find reward optimizing actions. This is also called the curse of dimensionality [70] and turns out severe in the control problem stated above.

The curse of dimensionality can be overcome by a reduction of the state space, where care must be taken to not drop information which is particularly important for the control scheme [71]. State space reduction was also done for the latest impressive results of DQNs learning to play super-human levels of computer games [69].

Transferred to the heat load control problem, the state space can be defined in two different levels.

The largest dimensionality occurs when defining the state space as the diagnostic output. Consider the state space to be only the input of heat load images. Since every pixel can take values from $\mathbb{R}_{>0}$, the dimension of the state space becomes the number of pixels times the number of cameras.

State space reduction can be done by replacing such an input by features which are embedded in the heat load image. Such features

should ideally be physics parameters or proxies for such because by principle, the basic plasma physics properties determine the whole system. These features can be extracted by means of supervised learning approaches. However, the important question is, whether the relevant plasma physics properties are chosen and none are missing (otherwise the information content is reduced) and whether labels can be retrieved in real time for supervised learning.

A trade-off between low dimension and minimal reduced information content has to be found and a mixture between both levels is possible.

3

Accomplishments

A RL control system as the one sketched above is highly ambitious and a lot needs to be investigated before such a system can be implemented in the safety-critical control system of W7-X. The preliminary studies carried out in this work are described below. In this presentation, the focus will be on the contributions to the control system as well as the spin-offs thereby created for the W7-X project.

3.1. Feature Reconstruction and Heat Load Pre-Processing

The reconstruction of magnetic edge topology parameters as ι and ΔR or proxies for such from heat load images in real time is desirable for the following reasons.

RL control system

State space reduction: As motivated before, the curse of dimensionality requires solutions such as state space reduction. ι and ΔR , being the two most dynamic parameters which occur during a discharge, are thus two obvious features which should be reconstructed from the highly dimensional heat load image state.

Agent architecture: As long as RL has not been implemented in the way intended, the requirement of state space reduction is a hypothesis, unproven but reasoned. In the hypothetical case that

state space reduction is not required later on, NN architectures which have proven successful in feature reconstruction are a good starting point for the architecture of a RL agent. This is because the agent NN architecture complexity will be at least as complex as successful NN architectures for the reconstruction problem. Such an approach highlights the advantage of ML approaches that NN architectures are very adaptive to changes of their task.

Real time information: Real time information about a proxy for edge magnetic equilibrium properties during experiment can help operators make decisions. In such a way, the information about edge magnetic topology properties can supplement real time information e. g. from magnetic diagnostics as diamagnetic loops (β , Pfirsch-Schlüter currents + diamagnetic currents) and Rogowski coils (toroidal current).

Insight: An often criticized characteristic of ML approaches is the missing insight to their internal mechanics. Especially in basic research and feasibility studies, which W7-X can be assigned to, given a control system that finds routes to optimization it is desirable to know physical equivalents of the state(s) from which a control agent took a certain action.

Mimic scenario development: For discharge planing, often it is of interest to compensate dynamic magnetic field changes. As shown in [72] it is possible to mimic the edge effects of magnetic equilibria by the change of coil currents in vacuum magnetic configurations. Such configurations are called mimic configurations. Once a mimic configuration is found, coil current differences can be inverted in order to compensate the effects which are expected on the basis of equilibrium calculations. So far, such scenarios were developed by a tedious manual search. Neural networks trained on vacuum configuration simulations reconstructing coil current proxies from heat loads can automate this task. A mimic configuration is directly found by feeding the heat load of the magnetic equilibrium calculation of choice to the NN. It should be noted however, that the resulting mimic configuration needs to be checked since an input which is too different from all the heat

load patterns which make up the NN training set could represent an extrapolation which the NN has not learned to describe.

3.1.1. ϵ reconstruction from limiter heat loads

Initial work presented in Articles I and II describes the successful implementation of NNs capable of reconstructing ϵ by means of the proxy I_B from limiter heat load images. Both articles' investigations are performed on the same data set which stems from an ϵ -scan conducted on 09.03.2016 (see Article I).

Article I describes the initial results of ϵ reconstruction by means of the proxy I_B . I_B can be reconstructed from experimental heat load images to a precision quantified by a root-mean-square error (rmse) of 0.010. To put this into perspective, a toroidal current of 44 kA changes the magnetic edge topology such that it can be approximately mimicked by $I_B = -0.1$ [72]. This means that this early NN is capable of resolving strike-line changes due to toroidal currents of approximately 5 kA.

However, the interpolation and especially extrapolation performance of this NN is performing, as expected, worse due to the sparse data set available as shown in Article I. Although more experiments were performed during OP 1.1, others were not exploiting the magnetic edge topology and thus add variation in other parameters which would even further increase the amount of training data required for robust reconstruction. To overcome the shortness of data, field line diffusion (FLD) simulations are performed in the vacuum magnetic field approximation. This is preferred over full equilibrium fields because of calculation time and a good approximation since the discharges in the experimental data set have low β and maximal toroidal current.

Simulation and experiment show systematic differences which is the reason for bad transferability of NNs trained on simulation to experimental data sets. It is shown in Article I, that the extrapolation performance can be increased when supplementing the experimental data set with simulated field line diffusion data during training. The reason is, that a NN confronted with both data set types learns to focus onto features which are present in both data sets and thus overcomes the systematic difference between simulation and experiment. Nonetheless, such a result requires a pre-processing which needs to be carefully adjusted in order to account for unavoidable systematic

differences between simulation and experiment.

With these findings a more thorough investigation of the effect of data set mixing and pre-processing is presented in Article II. In a systematic way limiter heat load resolution as well as parametrization is investigated with respect to the influence onto the reconstruction performance of three different NN architectures. It is shown that the performance can be increased by finding an optimal pre-processing, where strike-line segmentation and parametrization is investigated. Regarding the segmentation resolution a trade-off has to be made between considering limited simulation statistics (coarser resolution necessary) and sufficient detail (finer resolution). The shape of Figure 11 in Article II shows the expected optimized performance given moderate resolutions, falling off to extremes. Among different parametrizations a simple area averaged heat flux is found to be superior to moments of the heat flux. The crucial test of the NNs, trained on both, synthetic as well as experimental data, with a complementary experimental data set shows increased performance as compared to NNs trained solely on an experimental data set which must be the benchmark. As expected, convolutional neural network (CNN) outperform FF-FC architectures.

Article I describes the NN performance by means of the rmse. Since multiple NNs are trained until the best performance is found with the pre-processing and architecture described in Article I, this performance corresponds to the minimal rmse of all previously trained NNs. The performance metric in Article II is the mean rmse, labeled \overline{rmse} , of all NN trained within one category of NN architecture and pre-processing. Therefore, the findings in both articles can not be compared directly and we make up for that here. Figure 3.1a shows the best performing NN and can be compared with Figure 11 from Article I. By training with a combined training set of simulated and experimental limiter heat loads, the rmse can be decreased by about a factor of 2 from the aforementioned 0.010 to 0.006. Especially promising is the interpolation performance as depicted by Figure 3.1b with a rmse of 0.006 equal to the performance of the full test set^I as compared to the former Figure 12(a) from Article I with $rmse = 0.023$.

Extrapolation performance is improved from $rmse = 0.047$ to $rmse = 0.022$ (cf. Figure 3.1d with Figure 12(b) from Article I). The degree of extrapolation is of course critical. Figure 3.1c shows the extrapolation, when five

^Ionly in the next digits, the full test set appears to be better performing

instead of four of the six magnetic configurations are included in the training/validation set and only the upper value at $I_B = 0.13$ is extrapolating. The performance of extrapolation is very close to the performance of the full data set with an rmse of 0.008^{II}.

The method of mixing simulation and experiment despite systematic differences which leads to beneficial performance as measured by the relevant real-world data set is one which is not only relevant for this work but in general for the ML community. Success is not guaranteed and will depend on the respective case and the chosen pre-processing but an approach to overcome the limitations of small data sets especially for the data hungry NNs was long sought.

Preliminary results with generative adversarial neural networks (GAN) architectures have indicated even further performance increase and will be described in later publications.

3.1.2. ι - ΔR reconstruction from divertor heat loads

Motivated by these findings and in preparation of divertor heat loads for OP 1.2 a large heat load simulation data set with approximately 30 000 distinct magnetic configurations in ι - ΔR space is generated and introduced in Article III. The generation of such a large data set is computationally costly. A speed up in computation time can be achieved by defining an optimal resolution based on physics and engineering constraints. Such a speed up of above a factor of 20 can be found without reducing the statistical significance of the simulated heat loads. This finding is highly beneficial not only with respect to approaching the long-term objective but for heat load simulation by Monte Carlo (MC) methods in general.

Careful partitioning additionally leads to a NN input with well-defined neighborhood relations such that CNN architectures can be applied in a way which was found beneficial in Article II.

With Article IV we investigate optimal NN architectures for the reconstruction of ι and ΔR from the divertor heat loads contained in this data set. Six different architecture types, varying from shallow and simple FF-FC NNs to deep and specialized Inception-ResNet (IRNN), are investigated.

^{II}Note, that comparing Figures 3.1c and 3.1d the amount of training data varies which influences reconstruction performance.

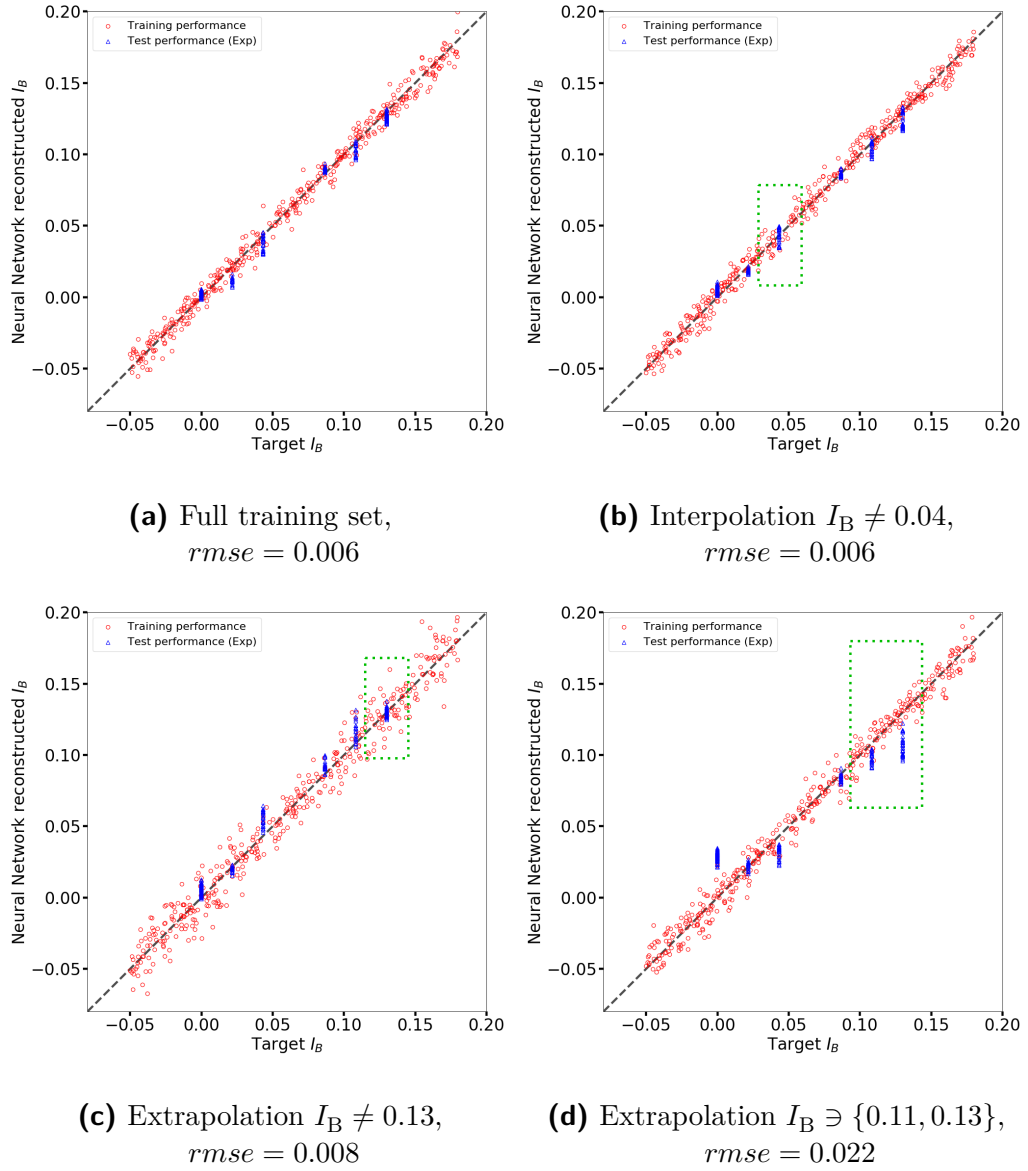


Figure 3.1.: Best performance of I_B reconstruction from limiter heat load images, based on optimal pre-processing and architecture found in Article II for comparison with results from Article I

The shallow architectures behave surprisingly well but their performance can not be entirely trusted as can be seen due to performance variation with changed cross-validation set. CNNs based architectures show stability across all cross-validation sets and furthermore increase performance. As expected from excellent results on image classification tasks [73, 74], the IRNN architecture adaption to our regression problem performs best, however with the longest training time.

3.2. Quantification of Critical Heat Loads

With the large-size, high-resolution data set introduced in Article III, generated in the t - ΔR plane, a further possibility arises. The detailed scan allows to generate a map of critical heat loads under the assumption of a chosen convective power P_{conv} . Care must be taken to account for statistical uncertainties due to under- or over-fluctuation of the MC method. In Article III a method is derived to determine a statistical measure for overload description, based on viewing the FLD simulation as independent Bernoulli processes. With the aid of this tool set, the overload of all primary PFCs is evaluated and severe critical regions are found. By the example of a mimicked toroidal current developing from the standard magnetic reference configuration the problem outlined above is sketched as can be seen in Figure 14(a) from Article III. Overload occurs as I_{tor} evolves. Sampling the higher dimension of island size by means of an island control coil scan we demonstrate how overload can be prevented (see Figure 14(b) from Article III). In this way we further motivate the use of sweep coils as actuators in the envisioned control system.

This statistical method, as well as the divertor partitioning enabling fast simulation at high statistical significance, has already been applied in two further investigations. In Article V it is described how divertor overload as predicted with VMEC equilibrium calculations^{III} in the transient phase of a

^{III}Since VMEC assumes only closed flux surfaces, VMEC equilibria can not describe magnetic islands. However, magnetic islands in the plasma edge are crucial to adequately model PFC heat loads. For this reason the VMEC solution is only used for the core magnetic field and the edge field is added by the virtual casing principle code (EXTENDER) [75]. A more holistic approach can be done with the Princeton iterative equilibrium solver (PIES) code [76] which however needs orders of magnitudes more

high performance discharge due to evolving toroidal currents can be avoided by tuning n and P_{heat} . Furthermore the methodology is applied for a study of baffle and heat shield overload. The overload measure is applied to FLD simulations for P_{conv} and a radiation solver code “loads on Wendelstein” (LOWENDEL) [77]^{IV} for P_{rad} on various equilibrium fields with the aim of identifying critical baffle tiles for safety relevant thermo-couple installation. Thereby the generic nature of the developed tool set is demonstrated. A manuscript is in preparation.

iterations to converge.

^{IV}The overload measure is not only applicable to FLD simulations but all Bernoulli like simulations.

4

Résumé

With this thesis, studies which form the bedrock for the long term goal of first wall heat load control and optimization for the advanced stellarator W7-X are developed, described and put into context. Because W7-X possesses a low shear rotational transform profile and the angle between PFCs and edge magnetic field is designed to be shallow for heat load reduction, the heat load pattern is highly susceptible to changes in the magnetic edge topology. This makes heat load control a highly important task since small effects of plasma dynamics as e. g. toroidal current evolution have a strong impact onto the heat load distribution and potential damage to the PFCs can occur. Real time capable heat load control is envisioned by means of a RL technique and it is laid out how reconstruction of features of the edge magnetic field from PFC heat loads by NNs is an important first step. Since this work is done in parallel to the first operation of W7-X, properly processed experimental data is still sparse. On the basis of a series of experiments following a scan in the rotational transform t we find that NN are capable of reconstructing a proxy for the edge rotational transform from the limiter heat loads. As expected, the reconstruction performance is limited by the amount of available training data. Interpolation and especially extrapolation performance is not satisfying. FLD simulation is employed to simulate additional limiter heat load images to be supplemented in the training process. With a thorough investigation of pre-processing and NN architecture the reconstruction performance is found to be strongly increased and much more stable in inter- and extrapolation when adding the simulation data set in the training process.

A detailed study of PFC heat load distribution and potential overloads (di-

vector set-up) is made in first order approximation of the impact of the two main plasma dynamic effects, the toroidal current evolution and the radial plasma shift ΔR . Since this involves heavy usage of FLD simulation. A tool-set for overload evaluation at an above 20-fold decrease of required simulation time at remaining significance is developed. This tool-set is applied already to other use cases as shown on the example of high performance discharge planning and machine safety evaluation.

On the basis of this massive data set together with the findings of optimized reconstruction performance found with limiter discharges we demonstrate that NNs are capable of reconstructing ι and ΔR from the heat loads. Deep IRNNs are found to most reliably and accurately reconstruct both parameters.

Outlook There are many open questions and challenges which need to be addressed. A few are outlined below:

- At present the massive amount of OP 1.2 data are being processed and a thorough analysis of available data and the according parameter spaces is being done. With an OP 1.2 ι -scan, the results from Article II shall be repeated on divertor geometry such that the gap between Articles I to IV is bridged.
- So far the existence of real-time available heat loads was presupposed. However the translation of IR temperature data to heat loads is costly. Efforts are being made to short-cut the heat load calculation by ML techniques. This is an important missing link for heat load control. At the same time it provides heat load data for various other applications. To name one example: Integrated data analysis with the Minerva framework [78] has not yet included the highly restrictive edge heat loads for plasma physics evaluation and could be expanded such.
- On the route towards RL control various pre-studies are mandatory. With the overload evaluation from Article III which can be defined as reward map and the feature reconstruction or NN architectures found beneficial in Article IV for a RL agent, presently experiments with RL control on simulation basis are made.

-
- So far the focus in the dimensions of plasma simulation was on the effects expected from the main plasma dynamics. However, especially direct control in the dimension of t is not easily possible. As already elaborated in Chapter 2, current changes in the superconducting coils are very slow and need to be calculated in advance whereas current drive is limited to the X2-heating scenario only (ECCD) or only available for 10 s (NBI). By adding a third dimension of the island width by means of island control coils to the t - ΔR scan we aim to expand an important dimension since the island control coils are highly advantageous actuators (fast actuation, direct influence to edge magnetic field, no need for pre-discharge force calculations). This was already touched briefly in Section 4.3 of Article III and revealed an attractive route to avoid PFC overload.
 - This work pursues an approach of steadily increasing the problem complexity. In order to include radiation cooling and detachment, a next step would be to utilize Edge Monte Carlo 3D code coupled to the neutral transport code EIRENE (EMC3-Eirene) calculations which opens up the door to radiation cooling and pumping optimization with flexible actuators (e. g. gas puff).

Bibliography

- [1] »ENDF database at the IAEA«. URL: <http://www-nds.iaea.org/exfor/>.
- [2] C. ALLEN. *Astrophysical Quantities*. 3rd. London: The Athlone Press, 1973.
- [3] W. M. NEVINS and R. SWAIN. »The thermonuclear fusion rate coefficient for p-11B reactions«. In: *Nuclear Fusion*, Vol. 40.4 (2000), pages 865–872. DOI: 10.1088/0029-5515/40/4/310.
- [4] OECD/NEA. *Plutonium fuels*. Technical report. Paris, 1989.
- [5] X. Z. LI et al. »Fusion cross-sections for inertial fusion energy«. In: *Laser and Particle Beams*, Vol. 22.4 (2004), pages 469–477. DOI: 10.1017/S026303460404011X.
- [6] J. D. LAWSON. »Some Criteria for a Power Producing Thermonuclear Reactor«. In: *Proceedings of the Physical Society. Section B*, Vol. 70.1 (1957), pages 6–10. DOI: 10.1088/0370-1301/70/1/303.
- [7] J. WESSON. *Tokamaks*. 2nd ed. Oxford University Press, 1997.
- [8] H. YAMADA. »Handbook of climate change mitigation«. In: *Handbook of Climate Change Mitigation*. Edited by W.-Y. CHEN et al. Vol. 1-4. Springer, 2012, pages 1183–1215. DOI: 10.1007/978-1-4419-7991-9.
- [9] M. KAUFMANN. *Plasmaphysik und Fusionsforschung*. 1st ed. B. G. Teubner, 2003, page 288.
- [10] G. G. STOKES. »On the theories of the internal friction of fluids in motion«. In: *Transactions of the Cambridge Philosophical Society*, Vol. 8 (1845), pages 287–305. DOI: 10.1017/CB09780511702242.

- [11] S. HIRSHMAN, W. VAN RIJ and P. MERKEL. »Three-dimensional free boundary calculations using a spectral Green's function method«. In: *Computer Physics Communications*, Vol. 43.1 (1986), pages 143–155. DOI: 10.1016/0010-4655(86)90058-5.
- [12] H. POINCARÉ. »Mémoire sur les courbes définies par une équation différentielle (I)«. In: *Journal de mathématiques pures et appliquées*, Vol. 7 (1881), pages 375–422.
- [13] H. POINCARÉ. »Mémoire sur les courbes définies par une équation différentielle (II)«. In: *Journal de mathématiques pures et appliquées*, Vol. 7 (1882), pages 251–296.
- [14] H. HOPF. *Abbildungen geschlossener Mannigfaltigkeiten auf Kugeln in n-Dimensionen*. Technical report. Jahresbericht Deutsche Math.-Vereinig., 1926, pages 130–133.
- [15] M. WAKATANI. *Stellarator and Heliotron Devices (International Series of Monographs on Physics)*. Oxford University Press, 1998. URL: <https://www.amazon.com/Stellarator-Heliotron-Devices-International-Monographs/dp/0195078314?SubscriptionId=0JYN1NVW651KCA56C102&tag=techkie-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=0195078314>.
- [16] P. HELANDER. »Theory of plasma confinement in non-axisymmetric magnetic fields«. In: *Reports on Progress in Physics*, Vol. 77.8 (2014), page 087001. DOI: 10.1088/0034-4885/77/8/087001.
- [17] H. WEYL. »Über die Gleichverteilung von Zahlen mod. Eins«. In: *Mathematische Annalen*, Vol. 77.3 (1916), pages 313–352. DOI: 10.1007/BF01475864.
- [18] P. I. STRAND and W. A. HOULBERG. »Magnetic flux evolution in highly shaped plasmas«. In: *Physics of Plasmas*, Vol. 8.6 (2001), pages 2782–2792. DOI: 10.1063/1.1366618.
- [19] M. KIKUCHI, K. LACKNER and M. Q. TRAN, editors. *Fusion Physics*. 1st ed. Vienna: International Atomic Energy Agency, 2012.
- [20] P. HELANDER, J. GEIGER and H. MAASSBERG. »On the bootstrap current in stellarators and tokamaks«. In: *Physics of Plasmas*, Vol. 18.9 (2011), page 092505. DOI: 10.1063/1.3633940.

-
- [21] F. F. CHEN. *Introduction to Plasma Physics and Controlled Fusion*. 2nd ed. Boston, MA: Springer US, 1984. DOI: 10.1007/978-1-4757-5595-4.
- [22] C. BEIDLER et al. »Physics and Engineering Design for Wendelstein VII-X«. In: *Fusion Technology*, Vol. 17.1 (1990), pages 148–168. DOI: 10.13182/FST90-A29178.
- [23] V. D. SHAFRANOV. »Equilibrium of a toroidal pinch in a magnetic field«. In: *Soviet Atomic Energy*, Vol. 13.6 (1963), pages 1149–1158. DOI: 10.1007/BF01312317.
- [24] T. KOBUCHI et al. »Magnetic configuration dependence of the shafranov shift in the Large Helical Device«. In: *Plasma Physics and Controlled Fusion*, Vol. 48.6 (2006), pages 789–797. DOI: 10.1088/0741-3335/48/6/006.
- [25] J. GEIGER et al. »Physics in the magnetic configuration space of W7-X«. en. In: *Plasma Physics and Controlled Fusion*, Vol. 57.1 (2015), page 014004. DOI: 10.1088/0741-3335/57/1/014004.
- [26] T. RUMMEL et al. »The Superconducting Magnet System of the Stellarator Wendelstein 7-X«. In: *IEEE Transactions on Plasma Science*, Vol. 40.3 (2012), pages 769–776. DOI: 10.1109/TPS.2012.2184774.
- [27] A. IYOSHI et al. »Overview of the Large Helical Device project«. In: *Nuclear Fusion*, Vol. 39.9Y (1999), pages 1245–1256. DOI: 10.1088/0029-5515/39/9y/313.
- [28] Y. FENG et al. »Physics of island divertors as highlighted by the example of W7-AS«. In: *Nuclear Fusion*, Vol. 46.8 (2006), pages 807–819. DOI: 10.1088/0029-5515/46/8/006.
- [29] H. RENNER et al. »The capabilities of steady state operation at the stellarator W7-X with emphasis on divertor design«. In: *Nuclear Fusion*, Vol. 40.6 (2000), pages 1083–1093. DOI: 10.1088/0029-5515/40/6/306.
- [30] T. S. PEDERSEN et al. »Key results from the first plasma operation phase and outlook for future performance in Wendelstein 7-X«. In: *Physics of Plasmas*, Vol. 24.5 (2017), page 055503. DOI: 10.1063/1.4983629.

- [31] T. S. PEDERSEN et al. »First results from divertor operation in Wendelstein 7-X«. In: *Plasma Physics and Controlled Fusion*, Vol. 61.1 (2019), page 014035. DOI: 10.1088/1361-6587/aaec25.
- [32] P. STANGEBY. *The Plasma Boundary of Magnetic Fusion Devices*. Edited by S. LAURENSEN. 1st ed. London: IOP Publishing, 2000, page 691.
- [33] R. C. WILLSON and H. S. HUDSON. »The Sun’s luminosity over a complete solar cycle«. In: *Nature*, Vol. 351.6321 (1991), pages 42–44. DOI: 10.1038/351042a0.
- [34] A. PEACOCK et al. »Status of High Heat Flux Components at W7-X«. In: *IEEE Transactions on Plasma Science*, Vol. 42.3 (2014), pages 524–532. DOI: 10.1109/TPS.2014.2302542.
- [35] J. KISSLINGER. »Divertor Concept for the WENDELSTEIN 7-X Stellarator: Theoretical Studies of the Boundary and Engineering«. In: *20th European Physical Society Conference on Plasma Physics*. Lisboa: Plasma Physics and Controlled Fusion, 1993, page 587.
- [36] H. GREUNER et al. »DIVERTOR ENGINEERING FOR THE STELLARATOR WENDELSTEIN W 7-X«. In: *Fusion Technology 1996*. Elsevier, 1997, pages 463–466. DOI: 10.1016/B978-0-444-82762-3.50086-0.
- [37] J. ONGENA et al. »Magnetic-confinement fusion«. In: *Nature Physics*, Vol. 12.5 (May 2016), pages 398–410. DOI: 10.1038/nphys3745.
- [38] J. LIKONEN et al. »Structural studies of deposited layers on JET MkII-SRP inner divertor tiles«. In: *Journal of Nuclear Materials*, Vol. 363-365.1-3 (2007), pages 190–195. DOI: 10.1016/j.jnucmat.2007.01.007.
- [39] D. HILDEBRANDT et al. »Thermographic observation of the divertor target plates in the stellarators W7-AS and W7-X«. In: *Journal of Nuclear Materials*, Vol. 313-316.SUPPL. (2003), pages 738–742. DOI: 10.1016/S0022-3115(02)01490-3.
- [40] J. ROTH. »Chemical erosion of carbon based materials in fusion devices«. In: *Journal of Nuclear Materials*, Vol. 266-269 (1999), pages 51–57. DOI: 10.1016/S0022-3115(98)00658-8.

-
- [41] V. PHILIPPS et al. »Erosion and redeposition of wall material in controlled fusion devices«. In: *Vacuum*, Vol. 67.3-4 (2002), pages 399–408. DOI: 10.1016/S0042-207X(02)00238-5.
- [42] H. GREUNER et al. »Review of the high heat flux testing as an integrated part of W7-X divertor development«. In: *Fusion Engineering and Design*, Vol. 84.2-6 (2009), pages 848–852. DOI: 10.1016/j.fusengdes.2008.12.083.
- [43] J. ROTH, J. BOHDANSKY and K. WILSON. »Erosion of carbon due to bombardment with energetic ions at temperatures up to 2000 K«. In: *Journal of Nuclear Materials*, Vol. 111-112.C (1982), pages 775–780. DOI: 10.1016/0022-3115(82)90304-X.
- [44] M. OGAWA and T. KUNUGI. »Thermohydraulic experiments on a water jet into vacuum during ingress of coolant event in a fusion experimental reactor«. In: *Fusion Engineering and Design*, Vol. 29.C (1995), pages 233–237. DOI: 10.1016/0920-3796(95)80029-W.
- [45] T. KALIATKA, E. USPURAS and A. KALIATKA. »Analysis of Ingress of Coolant Accident to the Vacuum Vessel of the W7-X Fusion Experimental Facility«. In: *Fusion Science and Technology*, Vol. 72.2 (2017), pages 176–187. DOI: 10.1080/15361055.2017.1320496.
- [46] F. EFFENBERG et al. »First Demonstration of Radiative Power Exhaust with Impurity Seeding in the Island Divertor at Wendelstein 7-X«. In: *To be published*, (2019).
- [47] A. PAVONE et al. »Bayesian uncertainty calculation in neural network inference of ion and electron temperature profiles at W7-X«. In: *Review of Scientific Instruments*, Vol. 89.10 (2018), 10K102. DOI: 10.1063/1.5039286.
- [48] K. RAHBARNIA et al. »Diamagnetic energy measurement during the first operational phase at the Wendelstein 7-X stellarator«. In: *Nuclear Fusion*, Vol. 58.9 (2018), page 096010. DOI: 10.1088/1741-4326/aacab0.
- [49] M. JAKUBOWSKI et al. »Infrared imaging systems for wall protection in the W7-X stellarator (invited)«. In: *Review of Scientific Instruments*, Vol. 89.10 (2018), 10E116. DOI: 10.1063/1.5038634.

- [50] A. HERRMANN et al. »Energy flux to the ASDEX-Upgrade diverter plates determined by thermography and calorimetry«. In: *Plasma Physics and Controlled Fusion*, Vol. 37.1 (1999), pages 17–29. DOI: 10.1088/0741-3335/37/1/002.
- [51] F. FÜLLENBACH et al. »Final test of the W7-X control coils power supply and its integration into the overall control environment«. In: *Fusion Engineering and Design*, Vol. 82.5-14 (2007), pages 1391–1395. DOI: 10.1016/j.fusengdes.2007.03.052.
- [52] Y. TURKIN. »Current Control by ECCD for W7-X«. In: *Fusion Science and Technology*, Vol. 50.3 (2006), pages 387–394. DOI: 10.13182/FST06-5.
- [53] Y. GAO et al. »Effects of toroidal plasma current on divertor power depositions on Wendelstein 7-X«. In: *Nuclear Fusion*, (2019). URL: <http://iopscience.iop.org/10.1088/1741-4326/ab32c2>.
- [54] T. STANGE et al. »Advanced electron cyclotron heating and current drive experiments on the stellarator Wendelstein 7-X«. In: *EPJ Web of Conferences*, Vol. 157 (2017). Edited by J. HILLAIRET, page 02008. DOI: 10.1051/epjconf/201715702008.
- [55] P. MCNEELY. »Ion Sources for Neutral Beam Heating of Fusion Plasmas«. In: *32nd International Electric Propulsion Conference*, (2011), pages 1–7.
- [56] J. BALDZUHN et al. »Pellet fueling experiments in Wendelstein 7-X«. In: *To be published*, (2019), pages 1–30.
- [57] D. ZHANG et al. »Design criteria of the bolometer diagnostic for steady-state operation of the W7-X stellarator«. In: *Review of Scientific Instruments*, Vol. 81.10 (2010), 10E134. DOI: 10.1063/1.3483194.
- [58] A. DINKLAGE. *Plasma Physics*. Edited by A. DINKLAGE et al. Volume 670. Lecture Notes in Physics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pages 1–20—. DOI: 10.1007/b103882.
- [59] A. DINKLAGE et al. »Magnetic configuration effects on the Wendelstein 7-X stellarator«. In: *Nature Physics*, Vol. 14.8 (2018), pages 855–860. DOI: 10.1038/s41567-018-0141-9.

-
- [60] A. J. WOOTTON et al. »Fluctuations and anomalous transport in tokamaks«. In: *Physics of Fluids B: Plasma Physics*, Vol. 2.12 (1990), pages 2879–2903. DOI: 10.1063/1.859358.
- [61] G. CYBENKO. »Approximation by Superpositions of a Sigmoidal Function«. In: *Mathematics of control, signals and systems*, Vol. 2.4 (1989), pages 303–414. DOI: 10.1016/j.aml.2003.11.006.
- [62] Y. BENGIO. »Learning Deep Architectures for AI«. In: *Foundations and Trends® in Machine Learning*, Vol. 2.1 (2009), pages 1–127. DOI: 10.1561/22000000006. arXiv: 0500581 [submit].
- [63] R. BELLMAN. »A Markovian Decision Process«. In: *Indiana University Mathematics Journal*, Vol. 6.4 (1957), pages 679–684. DOI: 10.1512/iumj.1957.6.56038.
- [64] J. RUST. »Using Randomization to Break the Curse of Dimensionality«. In: *Econometrica*, Vol. 65.3 (May 1997), page 487. DOI: 10.2307/2171751.
- [65] C. WATKINS. *Learning from Delayed Rewards. (Doctoral Dissertation)*. Ph.D. King’s College, 1989.
- [66] J. N. TSITSIKLIS. »Asynchronous Stochastic Approximation and Q-Learning«. In: *Machine Learning*, Vol. 16.3 (1994), pages 185–202. DOI: 10.1023/A:1022689125041.
- [67] P. MONTAGUE. *Reinforcement Learning: An Introduction*, by Sutton, R.S. and Barto, A.G. 2nd ed. Volume 1. MIT Press, 2017, page 360. DOI: 10.1016/S1364-6613(99)01331-5. arXiv: 1603.02199.
- [68] V. MNIH et al. »Human-level control through deep reinforcement learning«. In: *Nature*, Vol. 518.7540 (2015), pages 529–533. DOI: 10.1038/nature14236.
- [69] O. VINYALS et al. »StarCraft II: A New Challenge for Reinforcement Learning«. In: (2017). arXiv: 1708.04782. URL: <http://arxiv.org/abs/1708.04782>.
- [70] S. BENGIO and Y. BENGIO. »Taking on the curse of dimensionality in joint distributions using neural networks«. In: *IEEE Transactions on Neural Networks*, Vol. 11.3 (May 2000), pages 550–557. DOI: 10.1109/72.846725.

- [71] D. SHARMA. *Combining Reinforcement Learning and Feature Extraction*. PhD thesis. TU Darmstadt, 2012, page 23. URL: http://www.ias.informatik.tu-darmstadt.de/uploads/Theses/Sharma_BScThesis_2012.pdf.
- [72] H. HÖLBE. *Control of the magnetic topology and plasma exhaust in the edge region of Wendelstein 7-X : a numerical study*. PhD thesis. Universität Greifswald, 2016, page 121. URL: https://pure.mpg.de/rest/items/item_2250904/component/file_2250907/content.
- [73] K. HE et al. »Deep residual learning for image recognition«. In: *IEEE conference on computer vision and pattern recognition*. Las Vegas USA, 2016, pages 770–778. DOI: 10.1109/CVPR.2016.90. arXiv: 1512.03385.
- [74] C. SZEGEDY et al. »Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning«. In: *AAAI*. San Francisco, USA, 2017, page 12. DOI: 10.1016/j.patrec.2014.01.008. arXiv: 1602.07261.
- [75] M. DREVLAK, D. MONTICELLO and A. REIMAN. »PIES free boundary stellarator equilibria with improved initial conditions«. In: *Nuclear Fusion*, Vol. 45.7 (2005), pages 731–740. DOI: 10.1088/0029-5515/45/7/022.
- [76] A. REIMAN and H. GREENSIDE. »Calculation of three-dimensional MHD equilibria with islands and stochastic regions«. In: *Computer Physics Communications*, Vol. 43.1 (1986), pages 157–167. DOI: 10.1016/0010-4655(86)90059-7.
- [77] T. EICH and A. WERNER. »Numerical Studies on Radiative Heat Loads to Plasma-Facing Components for the W7-X Stellarator«. In: *Fusion Science and Technology*, Vol. 53.3 (2008), pages 761–779. DOI: 10.13182/FST08-A1733.
- [78] J. SVENSSON and A. WERNER. »Large Scale Bayesian Data Analysis for Nuclear Fusion Experiments«. In: *2007 IEEE International Symposium on Intelligent Signal Processing*. IEEE, 2007, pages 1–6. DOI: 10.1109/WISP.2007.4447579.

Acronyms

CFC carbon fibre reinforced carbon	LOWENDEL “loads on Wendelstein”
CNN convolutional neural network	MC Monte Carlo
CXRS charge exchange recombination spectroscopy	MHD magnetohydrodynamic
DQN deep Q-network	ML machine learning
ECCD electron cyclotron current drive	NBCD neutral beam current drive
ECRH electron cyclotron resonance heating	NBI neutral beam injection
EMC3-Eirene Edge Monte Carlo 3D code coupled to the neutral transport code EIRENE	NN artificial neural network
EXTENDER virtual casing principle code	OP operation phase
FF-FC feed-forward fully-connected neural network	PFC plasma facing component
FLD field line diffusion	PIES Princeton iterative equilibrium solver
GAN generative adversarial neural networks	RL reinforcement learning
ICRH ion cyclotron radiation heating	rmse root-mean-square error
IR Infra-Red	SOL scrape-off layer
IRNN Inception-ResNet	TDU test divertor unit
LHD large helical device	VMEC variational moments equilibrium code
LCMS last closed magnetic surface	W7-X Wendelstein 7-X
	XICS X-ray imaging crystal spectroscopy

List of Symbols

Symbol	Description	Unit
I_{bs}	Bootstrap current.	A
I_{tor}	Toroidal plasma current.	A
t_{CF}	Current free rotational transform.	
t	Rotational transform.	
β	Plasma beta.	
ΔR	Radial shift of the magnetic axis.	m
σ	Reaction cross-section.	m
n_e	Electron density.	m^{-3}
n_i	Ion density.	m^{-3}
n	Particle density, $n = N/V$.	m^{-3}
\vec{B}	Magnetic field.	T = $kg A^{-1} s^{-2}$
\vec{E}	Electric field.	$kg m A^{-1} s^{-3}$
ρ	Electric charge density.	$A s m^{-3}$
\vec{j}	Electric current density.	$A m^{-2}$
W	Plasma kinetic energy.	$kg m^2 s^{-2}$
R	Electric resistance.	Ω

List of Symbols

Symbol	Description	Unit
L	Inductance.	H = kg m ² A ⁻² s ⁻²
L_c	Connection length.	m
\mathcal{L}	Loss function.	
N_p	Number of field periods.	
P_{conv}	Convective power.	W
P_{heat}	Total heating power.	W
P_{rad}	Radiated power.	W
p	Plasma pressure.	kg m ⁻¹ s ⁻²
R	Major radius.	m
ϵ	Aspect ratio.	
r_a	minor radius of the last closed flux surface.	m
r	minor radius.	m
T_e	Electron temperature.	eV
T_i	Ion temperature.	eV
τ_E	Energy confinement time.	s
\vec{v}_f	Plasma flow velocity.	m s ⁻¹
\vec{v}	Particle velocity.	m s ⁻¹
π	Plasma viscosity.	kg m ⁻¹ s ⁻¹
V	Plasma volume.	m ³

List of Figures

1.1.	Cross-sections of various nuclear fusion reactions as a function of the center of mass energy of the two particles.	2
1.2.	Field lines generating nested magnetic surfaces.	4
1.3.	The W7-X coil set.	10
1.4.	Poincare representation of the standard reference case magnetic field.	15
1.5.	PFCs in the two stages of OP 1.	16
2.1.	Scheme of a simple NN (FF-FC)	21
2.2.	RL schemes	22
3.1.	Best performance of I_B reconstruction from limiter heat load images, based on optimal pre-processing and architecture found in Article II for comparison with results from Article I	30



Thesis Articles

Authors contributions:

The W7-X Team conducted the experiments and performed the analysis of data relevant to the characterization of the experiments.

Article I

»Reconstruction of Magnetic Configurations in W7-X using Artificial Neural Networks«

D. BÖCKENHOFF, M. BLATZHEIM, H. HÖLBE, T. SUNN PEDERSEN and R. LABAN

Nuclear Fusion, Vol. 58.5 (2018), DOI: 10.1088/1741-4326/aab22d

The work was done in close collaboration of D. Böckenhoff and M. Blatzheim. D. Böckenhoff and M. Blatzheim wrote the manuscript. D. Böckenhoff generated the simulation training data. F. Pisano and H. Niemann provided the raw heat load data. The pre-processing was developed by D. Böckenhoff and M. Blatzheim. M. Blatzheim did network training. D. Böckenhoff developed the method of mixing simulation and experimental data for network performance improvement given limited experimental data. D. Böckenhoff and M. Blatzheim did the analysis of the neural network results. The work was edited by H. Hölbe, R. Labahn and T. Sunn Pedersen.

Article II

»Neural network performance enhancement for limited nuclear fusion experiment observations supported by simulations«

M. BLATZHEIM, D. BÖCKENHOFF, H. HÖLBE, T. S. PEDERSEN
and R. LABAHN

Nuclear Fusion, Vol. 59.1 (2019), DOI: 10.1088/1741-4326/aaefaf

The work was done in close collaboration of D. Böckenhoff and M. Blatzheim. D. Böckenhoff and M. Blatzheim wrote the manuscript. The pre-processing was developed by D. Böckenhoff and M. Blatzheim. M. Blatzheim did the network training. M. Blatzheim developed the affine transformation in the pre-processing. D. Böckenhoff introduced the statistical analysis. D. Böckenhoff and M. Blatzheim did the analysis of the neural network results. The work was edited by H. Hölbe, R. Labahn and T. Sunn Pedersen.

Article III

»Application of improved analysis of convective heat loads on plasma facing components to Wendelstein 7-X«

D. BÖCKENHOFF and M. BLATZHEIM

Nuclear Fusion, Vol. 59.8 (2019), DOI: 10.1088/1741-4326/ab201e

The manuscript was written by D. Böckenhoff and edited by M. Blatzheim. D. Böckenhoff did the major work for this publication. M. Blatzheim was involved in creating a decision tree for the numerical implementation of the divertor mapping and a dialectic partner for the development of the overload definition.

Article IV

»Neural network regression approaches to reconstruct properties of magnetic configuration from Wendelstein 7-X modeled heat load patterns«

M. BLATZHEIM and D. BÖCKENHOFF

Nuclear Fusion, Vol. 59.12 (2019), DOI: 10.1088/1741-4326/ab4123

The manuscript was written by M. Blatzheim and edited by D. Böck-

enhoff. D. Böckenhoff was responsible for simulation data generation, pre-processing and data analysis. M. Blatzheim was responsible for network training. D. Böckenhoff and M. Blatzheim did the analysis.

Article V

»Scenario with Combined Density and Heating Control to Reduce the Impact of Bootstrap Current in Wendelstein 7-X«

P. SINHA, D. BÖCKENHOFF, M. ENDLER, J. GEIGER, H. HÖLBE, H. M. SMITH, T. SUNN PEDERSEN and Y. TURKIN

Nuclear Fusion, accepted for publication, (2019), DOI: 10.1088/1741-4326/ab37bb

P. Sinha did the VMEC simulations and wrote the manuscript. H. M. Smith and Y. Turkin provided the DKES and NTSS simulations. D. Böckenhoff provided the overload analysis. M. Endler and J. Geiger did major editing. P. Sinha, M. Endler and J. Geiger did the major discussion of the results. The final editing involved all the co-authors (D. Böckenhoff, M. Endler, J. Geiger, H. Hölbe, H. M. Smith, T. S. Pedersen, Y. Turkin).

Confirmed by Supervisor,
Greifswald, 20.09.2019

PROF. DR. THOMAS SUNN PEDERSEN

Confirmed by doctoral candidate,

Greifswald, 20.09.2019

DANIEL BÖCKENHOFF

Appendix A. Thesis Articles

A.1. Article I

»Reconstruction of Magnetic Configurations in W7-X using Artificial Neural Networks«

D. BÖCKENHOFF, M. BLATZHEIM, H. HÖLBE, T. SUNN PEDERSEN
and R. LABAN

Nuclear Fusion, Vol. 58.5 (2018), DOI: 10.1088/1741-4326/aab22d

Reconstruction of Magnetic Configurations in W7-X using Artificial Neural Networks

Daniel Böckenhoff¹, Marko Blatzheim^{1, 2}, Hauke Hölbe¹, Holger Niemann¹,
Fabio Pisano³, Roger Labahn², Thomas Sunn Pedersen¹ and the W7-X Team¹

¹Max Planck Institute for Plasma Physics, Wendelsteinstraße 1, 17491
Greifswald, Germany

²Institute for Mathematics, University of Rostock, Ulmenstraße 69, 18057
Rostock, Germany

³Department of Electrical and Electronic Engineering, University of Cagliari,
Via Marengo 2, 09123 Cagliari, Italy

daniel.boeckenhoff@ipp.mpg.de, marko.blatzheim@ipp.mpg.de

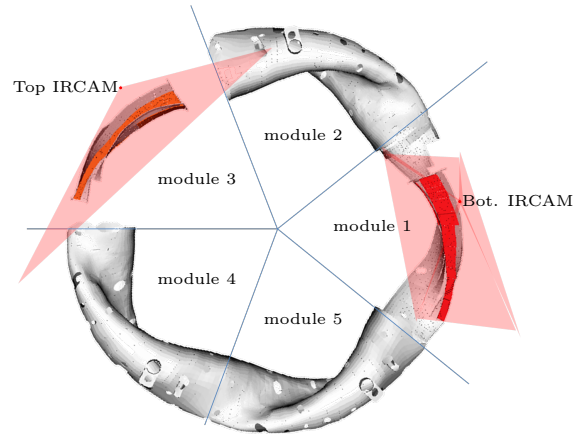
Abstract

It is demonstrated that artificial neural networks can be used to accurately and efficiently predict details of the magnetic topology at the plasma edge of the Wendelstein 7-X stellarator, based on simulated as well as measured heat load patterns onto plasma-facing components observed with infrared cameras. The connection between heat load patterns and the magnetic topology is a challenging regression problem, but one that suits artificial neural networks well. The use of a neural network makes it feasible to analyze and control the plasma exhaust in real-time, an important goal for Wendelstein 7-X, and for magnetic confinement fusion research in general.

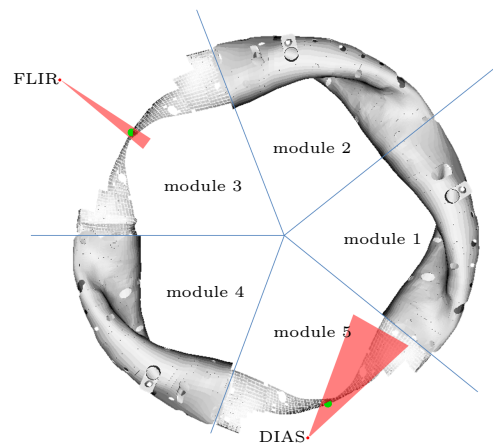
1 Introduction

Wendelstein 7-X (W7-X) is a pioneering experiment of the latest generation of optimized stellarators [1, 2]. It aims to demonstrate steady state capability of the confinement concept with plasma parameters near those required in an energy-producing fusion reactor [3, 4], and to demonstrate the steady-state capabilities of stellarators also at these parameters. The experiment started operation (Operation Phase 1.1, OP1.1) in 2015 [5], has started its next operation phase (OP1.2) in 2017, and will, when the water-cooled plasma-facing components have been fully installed in operation phase 2 (OP2), have discharge times of 30 min with continuous injected power of up to 10 MW. This discharge duration is about an order of magnitude longer than the longest characteristic time scale of the plasma, which is the one on which the net toroidal current, I_{tor} , evolves. The dominant contributor to I_{tor} is the bootstrap current I_{bs} , and the time evolution of I_{tor} can be approximated as: $I_{\text{tor}} = I_{\text{bs}} \cdot e^{-\frac{t}{\tau}}$ with L/R -time τ being the L/R time scale, of order 1 min for high-performance plasmas. On these time scales, power exhaust is an important issue. Ten specially designed discrete island divertor modules [6] take the major part of heat flux that passes the Last Closed Magnetic Surface (LCMS), which defines the edge of the confined plasma. The actively cooled divertors, necessary for OP2, are designed to sustain local power loads up to 10 MW/m^2 . Other first wall components, however, can only be loaded with a fraction of this heat flux, e. g. the less cooled divertor edge withstands roughly 3 MW/m^2 [7, 8]. Overloading the divertors implies a risk of delamination, water leaks (cf. [9, 10]) and impurity buildup, partially as a consequence of the first two, and must therefore be avoided.

To ensure the safety of the first wall, in particular the divertor, and to protect the plasma from impurities, real-time control is highly desirable, in particular for steady-state oper-



(a) Divertor setup with cutaways in modules one and three as used in OP1.2 and afterwards. Depicted are only two of the 10 divertors and the associated IRCAMS. Module one only reveals the top divertor while module three shows the bottom divertor.



(b) Limiter setup with cutaways in modules three and five as used in OP1.1. On this scale and view, the limiters are small (green).

Figure 1: Top-down CAD views of the W7-X inner vessel, showing sight lines of the IR camera systems in the two differing setups.

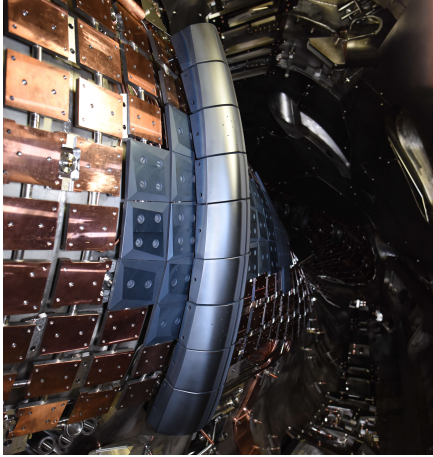


Figure 2: *Side view of the W7-X limiter in module 5 after OP1.1*

ation. An important part of such a real-time control system would be one that optimizes elements of the divertor operation, such as pumping performance and detachment. We aim to develop and ultimately implement such a feedback controller in the device operation system.

The most important sensor system for this control loop will be ten endoscopes that will allow IR and visible observation of the ten divertor units. Prototypes of these will be in operation in the now running intermediate operation phase OP1.2, each observing one divertor [11] (cf. Figure 1a). In OP1.1 five graphite limiters were placed vertically in the bean shaped cross section of the plasma vessel (cf. Figure 1b). In the confinement region, the magnetic field lines form nested toroidal surfaces that are not intersected by material objects. The aforementioned LCMS marks the transition to the exhaust region, called the Scrape-Off Layer (SOL) which is characterized by field lines that do intersect material objects, the Plasma-Facing Components (PFCs). The plasma enters the SOL through diffusive and convective transport processes perpendicular to the magnetic field, and then enters and populates the SOL field lines. These guide the outflowing plasma heat and particles to the PFCs. A common term for the shape- and intensity distribution of the resulting heat load pattern on the PFCs is “strike line” since it is

usually highly elongated. A complex interplay of component shape, magnetic field structure, cross- and parallel field transport results in the strike lines observed.

The IR camera views of the strike-line patterns have a complicated relationship with the magnetic topology and the transport processes in the plasma and the main task for such a feedback system is the abstraction and real time processing of these camera images which are being recorded at frame rates of 50 Hz. Artificial Neural Networks (NN, see Section 2.1) have significantly improved in the last ten years and should be well suited for this task.

A machine learning approach for real time control is an active object of research in the nuclear fusion community for several different tasks [12–14].

The first step towards the ambitious goal of controlling the plasma divertor operation for W7-X in real-time is presented here: the training of NNs to reconstruct the edge magnetic topology from IR camera observations of the strike-lines.

So far, IR picture series are available from OP1.1 discharges only (cf. Section 2.4.1), with OP1.2 barely started at the time of writing this paper. With a very limited set of different magnetic configurations, these experimental data available is no solid basis for a NN training set. To alleviate this problem, simulated strike lines function as NN training set data, and the performance of the NN is verified with experimental data. Since in OP1.1 no divertor was installed only the aforementioned limiters, this study is done entirely with the OP1.1 limiter geometry and associated magnetic configurations.

The limiters are centered on planes with up-down symmetry, which reduces the geometric complexity of the strike line. Complexity of the intersection between magnetic field and limiter is moreover minimized by limiter design. These two factors restrain the limiter strike line variability to a fraction of the divertor strike line. Because of this, the mag-

netic configuration reconstruction from limiter strike lines is expected to be much less accurate than that of more variable divertor strike lines.

The relevant information regarding the used NN architecture, W7-X, training data creation and NN input pre-processing are given in Section 2. In Section 3 we present the results of the challenging magnetic field reconstruction from limiter strike lines, which give reason for optimism with respect to future performance in a divertor geometry.

2 Methods

2.1 Artificial Neural Network

Machine learning tackles the question "How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?" as stated by Tom Mitchel in [15]. Artificial Neural Networks are an approach to solving machine learning problems. This capability allows them to reproduce the behavior of unknown, complicated functions from many examples instead of explicitly programming the function traditionally. The NNs used to predict the magnetic configuration are all based on ideas from 1975, when the method of backpropagation for NNs was introduced [16]. Such a NN depends on numerous parameters that are optimized in a so-called training process. In the previous decades NNs were not as popular as they have now become, especially because the parameter convergence during the training turned out to be slow and the amount of training data was small. Recently NNs are used more frequently because of newly developed network architectures [17, 18], the improvement of NN training algorithms [19, 20], increased computer performance [21] and open source NN libraries [22, 23]. A simple NN is based on fully connected lay-

ers (see Figure 4a), where an input vector \vec{x} is linearly transformed, followed by an element wise non-linearity ϕ , usually referred to as activation function (see Figure 3) such that

$$f(\vec{x}) = \phi(\mathbf{W}\vec{x} + \vec{b}) \quad (1)$$

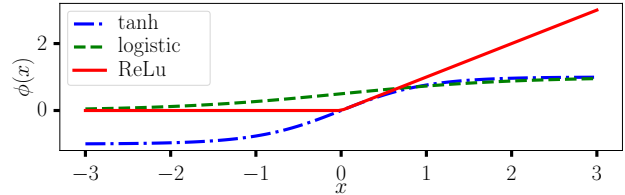
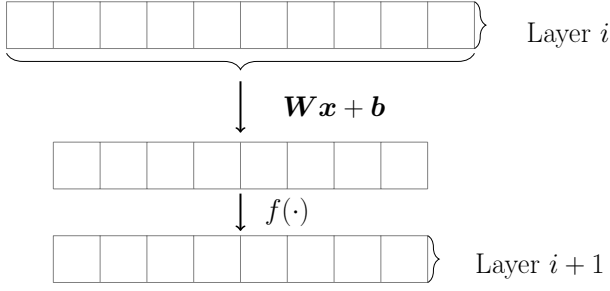


Figure 3: *Activation functions*

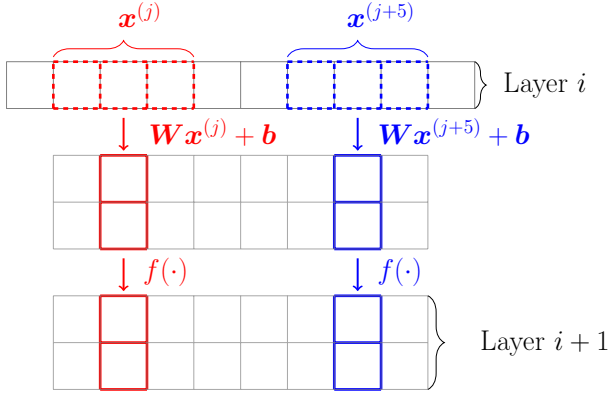
represents one layer. The free parameters \mathbf{W} and \vec{b} define the linear transformation. The activation function mimics the transmission of action potential in physiological neurons, which only create a signal if a certain threshold potential is reached. It is crucial to choose ϕ so that it is smooth and differentiable almost everywhere to allow an efficient gradient-based optimization. A NN consists of multiple of such or more complex layers where the output of one layer is treated as the input of the following one.

A more complex layer is the so-called convolutional layer [24], which has been developed to recognize patterns independent of the location by applying the same linear transformation on multiple smaller subsets of the input. Another positive effect is the reduction of free parameters, which decreases the risk of overfitting [25]. Convolutional layers (see Figure 4b) are used particularly frequently in image pattern recognition. For example the AlexNet architecture [26], which successfully won the 2012 ImageNet challenge [27], consists of several such layers. Most more complex layer structures, that are not recurrent, are based on the concept of convolutional layers.

The free parameters are arranged in the vector



(a) Fully connected layer



(b) 1D convolutional layer with kernel size 3 and 2 output feature maps

Figure 4: Schematic representation of neural network layer architectures

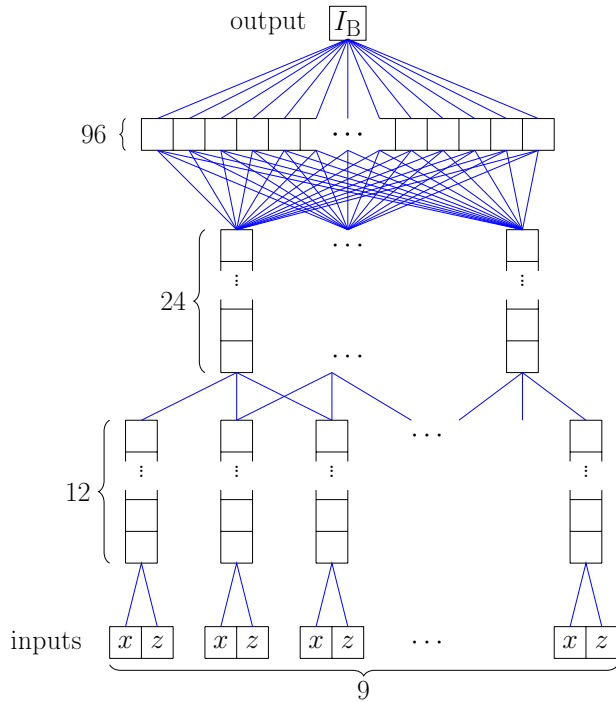


Figure 5: The used neural network architecture with two convolutional layers and two fully connected layers

$\vec{\theta}$. To train the NN, the error

$$E(\vec{\theta}) = \sum_{i=1}^N \left(y(\vec{x}_i, \vec{\theta}) - y_i^{\text{target}} \right)^2, \quad (2)$$

defined as the sum of squared differences between all NN result values $y(\vec{x}_i, \vec{\theta})$ with inputs \vec{x}_i and target values y_i^{target} of the corresponding set with set size N , is minimized with respect to all free parameters. By calculating $\vec{\nabla}_{\theta} E$, the gradient of the error with respect to $\vec{\theta}$, efficient gradient based minimization algorithms can be used. The basic approach is the gradient descent

$$\vec{\theta}_{i+1} = \vec{\theta}_i - \lambda \vec{\nabla}_{\theta} E \quad (3)$$

with step size λ . This algorithm starts with an initial $\vec{\theta}_0$ and calculates new $\vec{\theta}_{i+1}$ with every iteration.

The data set is separated in training, validation and test sets. Weights are adjusted with the training set only. Training terminates if the error on the validation set reaches a minimum. Once the NN training stops, the generalization quality is checked on the independent test set. To increase the rate of convergence, this method has been improved by the stochastic gradient descent where the error gradient is not calculated for all training sets but only for a stochastic selection, the mini-batch. After each mini-batch θ is updated. Other improvements focus on the choice of $\vec{\theta}_0$ [28] or the step size λ , where, instead of a constant value, an adaptive method depending on the previous λ and the derivative $\vec{\nabla}_{\theta} E$, such as AdaGrad [19] and Adam [20], determines the update rule. Though in the W7-X experiment IR cameras will monitor the vessel wall temperature, the input for the used NN is not an image, but characteristics extracted from it, as further explained in Section 2.4.3. Experimental data is scarce so the main data set is based on simulations (see Section 2.4.2). The network structure (see Figure 5) consists of two 1D convolutional layers followed by two fully connected layers. This choice is based on the better performance as compared to fully connected layer based NNs. The convolutional

layers consist of 12 and 24 feature maps respectively, the first fully connected layer consists of 96 neurons. The last layer in this NN leads to the target value.

The network is trained with the Adam optimizer and the recommended parameters. A dropout [29] of 0.5 is applied during the training. The weights are initialized by a normal distribution with standard deviation 0.5 and 0.1 for the first three layers and the last layer respectively. The activation functions as introduced in Figure (3) are Rectified Linear Units (ReLU) in the convolutional layers. In the first fully connected layer the logistic function has been chosen as activation function. Before the first layer, the NN input is linearly transformed in such a way, that the training set input has mean 0 and standard deviation 1 to improve convergence. The simulated data set has a size of 3993 and the experimental data set has a size of 319 cf. Table 2. The training set is nine times larger than the validation set while the test set size depends on the respective experiment.

The NN quality is evaluated by the root-mean-square error (*rmse*), defined as

$$rmse = \sqrt{\frac{E(\vec{\theta})}{N}}. \quad (4)$$

2.2 W7-X Coil System

Toroidal magnetic confinement devices need a rotational transform \bar{l} to compensate for magnetic drifts which would otherwise cause charge separation. \bar{l} is defined as

$$\bar{l} = c \cdot \frac{I_{\text{tor}}}{\Theta} + \bar{l}_{CF}, \quad (5)$$

where

$$I_{\text{tor}} = \int_0^r j_{\text{tor}}(r') \frac{V(r')}{2\pi R_0} dr' \quad (6)$$

is the toroidal current enclosed by the magnetic surface at r . $\Theta = \int \vec{\nabla}_\theta \cdot \vec{B} dV$ represents the toroidal magnetic flux. c denotes a constant that depends on the machine coil currents. \bar{l}_{CF} is the current free part of \bar{l} , only

occurring in stellarators and as well depending on the coil geometry and currents [30].

The W7-X coil system is shown in Figure 6a. The magnetic coil system consists of five identical modules each of which is point symmetric towards the module center, creating a five-fold symmetry in the magnetic field (cf. Figure 6b). Each half module consists of five different non-planar and two planar coils, depicted in red and blue. The modular coils are used to confine the plasma. The planar coils provide further variability of the magnetic field.

The modular and planar coils consist of 108 and 36 windings respectively. For convenience and clearness, we indicate the corresponding coil currents I_1, \dots, I_5 and I_A, I_B in terms of coil currents that would yield the same magnetic field with one winding only:

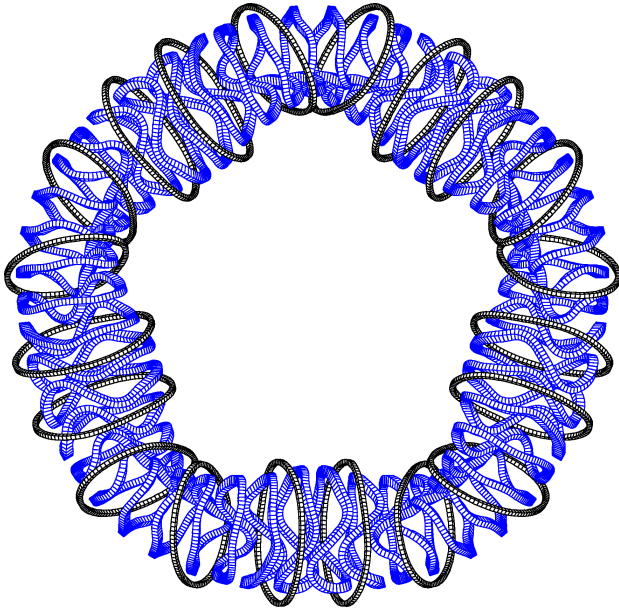
$$I_x = \frac{n_w \cdot I_{x,true}}{I_n}, \quad (7)$$

$\forall x \in \{1, 2, \dots, 5, A, B\}$, where the nominal current I_n is defined such, that all I_x are dimensionless. $I_{x,true}$ is the current that is actually applied to the coils and n_w denotes the number of windings per coil. This follows the convention used in [32].

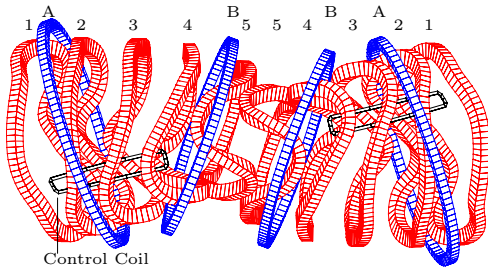
The parameter space spanned by the coil currents is six-dimensional since the effects of the magnetic field strength are negligible.

2.3 Parameter Space Choice

The strike line is mainly influenced by the topology of the magnetic field \vec{B} at the edge. The topology of \vec{B} is to a large degree determined by the rotational transform \bar{l} (cf. Section 2.2), which is a measure of how much a field line moves poloidally for one full toroidal turn. In stellarators, \bar{l} is primarily determined by the currents in the magnetic coils rather than those running inside the plasma. We refer to this plasma-current free contribution as \bar{l}_{CF} . \bar{l} is heavily dependent on the values of I_A and I_B . If $I_A + I_B < 0$, \bar{l} is increased and vice versa. Plasma currents can similarly change



(a) All 50 modular (blue) and 20 planar coils (black) [31]



(b) Coils contained in one module with modular coils 1 – 5 (red) and planar coils A, B (blue). Adapted after [32]

Figure 6: Overview of the W7-X main coil system

\bar{i} but these plasma effects can be effectively mimicked by adjusting coil currents [31]. In particular, the \bar{i} changes due to plasma currents are roughly equivalent to changes caused by the currents I_A and I_B in the planar coil systems. While it is complicated and computationally intensive to calculate the plasma currents that change \bar{i} , it is straight-forward to calculate the essentially equivalent changes in \bar{i} caused by I_A and I_B . For this reason, the NNs are trained to reconstruct the sum of the modular currents I_A and I_B of an \bar{i} -scan rather than reconstructing the toroidal plasma current.

Table 1: Planar currents I_A and I_B of OP1.1 \bar{i} -scan experiment and simulation. The number of distinct configurations in the scan is labeled “# configs”.

	I_A	I_B	# configs
experiment	0.13	[0.00, 0.13]	6
simulation	0.13	[-0.05, 0.18]	1001

On 2016/03/09 several experiments were conducted, where a modest \bar{i} -scan was made (see e.g. [5]). I_B was varied from its standard OP1.1 normalized value of 0.13 (corresponding to 5 kA per turn), down to 0 while I_A was kept constant at 0.13. Strike lines from this experiment were also simulated (cf. 2.4.2) with a much closer spacing than that obtained with actual OP1.1 plasmas. The simulated data also extends this range by 0.05 in both directions. Table 1 lists the normalized currents of this experiment as well as those of the simulations.

2.4 Data Set Creation

In the following we discuss all steps necessary for the generation of NN training data. First the experimental IR data processing is discussed. Subsequently the process of strike line simulation is described. For several reasons, a post processing applied to both experimental and synthetic data is necessary before the data

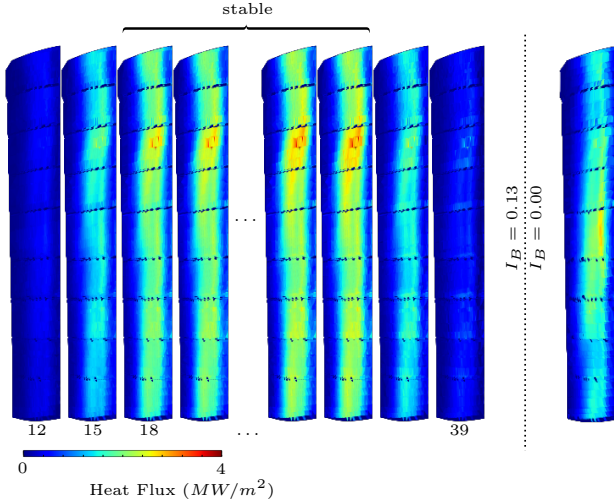
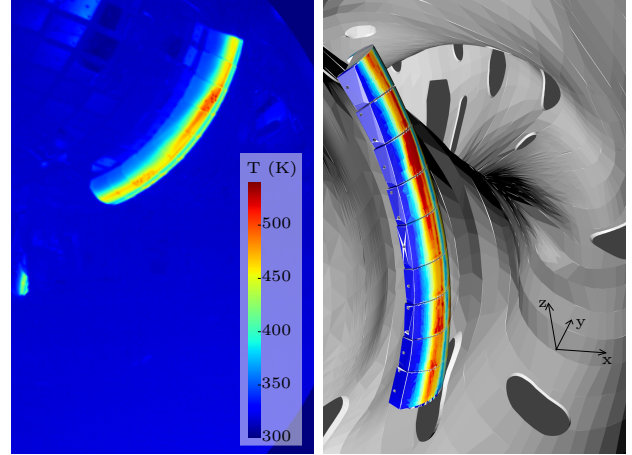


Figure 7: Selection of eight frames from the DIAS diagnostic, recorded during discharge “20160309.007” at $I_B = 0.13$ in heat flux representation. The chosen frames (respective frame numbers written below) display the plasma build up and break down. For this experiment, a stable plasma-wall interaction occurs from frame 18 till 34. The frames depicted only show that half of the limiter which was observed by the camera. For a comparison of strike lines at different I_B , one frame from discharge “20160309.035” ($I_B = 0.00$) is shown on the right.

can be used as a NN input. This is elaborated in Section 2.4.3.

2.4.1 IR Data Processing

The OP1.1 \bar{i} -scan covers I_B specified in Table 2. As mentioned above, two IR cameras were installed in OP1.1 (cf. Figure 1b). One long-wavelength (8 - 14 μm) μ -Bolometer camera from the company DIAS, able to operate in a 3 T magnetic field and a mid-wavelength (3 - 5 μm) CCD chip camera from the company FLIR. Both observe only parts of the limiter while other parts are shadowed (obstructed views). The FLIR camera observes only the center of the limiter. The DIAS camera faces a full side of module 5 limiter (see Figure 8a).



(a) DIAS IR camera temperature data **(b)** Limiter heat flux density corresponding to Figure 8a in three dimensions

Figure 8: Pseudo-color display of raw and processed data of discharge “20160309.007”

Table 2: Summary of available DIAS data for OP1.1 \bar{i} -scan from 2016/03/09. All modular coil currents were equal, that is $I_1, \dots, I_5 = 1$. $I_A = 0.1300$ for all experiments of that day. I_B is given in the first column. The following column gives the number of discharges where reliable IR data were available. The column labeled “# frames” denotes the total number of frames with stable plasma-wall interaction, i. e. durable strike line.

I_B	# discharges	# frames
0.1300	3	52
0.1083	2	35
0.0867	2	42
0.0433	2	40
0.0217	2	29
0.0000	5	121

Since the limiter strikeline is expected to be symmetric about the center of the limiter, observing one side of the limiter is sufficient to infer the full strike line. Thus we focus on the DIAS camera as an input.

Each IR video is processed in several steps to obtain the heat flux q (in W/m^2) [33].

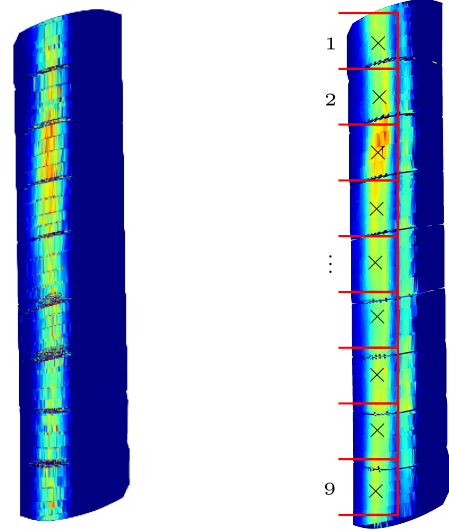
Mapping 2D to 3D In order to account for geometry effects, the sequence of such IR pictures is mapped onto the 3D CAD model of the limiter [34]. Result of this procedure is a 3D temperature map, assuming the surface to behave like a perfect black body with an emissivity of 1.

Emissivity correction In reality, the emissivity of graphite is smaller than 1. Furthermore, over the three months of OP1.1 the emissivity of the limiter tiles became nonuniform. This effect was explained in terms of surface layer depositions created during plasma operation [35], was quantified, and the data were corrected accordingly, yielding absolutely calibrated surface temperatures of the limiter as a function of time.

Heat flux This surface temperature is the result of the heat capacity and heat conductivity of the graphite combined with the time-history of the incident heat flux on the carefully shaped limiter surface. The THEODOR (THERmal Energy Onto DivertOR) algorithm [36] utilizes the temperature-dependent material properties and the measured time-dependent temperatures to deduce the heat fluxes. This typically reveals a pair of strikelines on either side of the limiter, one of which is the subject of this analysis.

Interpolation As a last step, we interpolate the resulting discrete heat flux map to retrieve a continuous heat flux distribution on the limiter surface as shown in figures 7 and 8b.

Dark frames are removed. Furthermore, the plasma build up and break down is discarded



(a) Heat load, represented by strike point density from a field line diffusion process (b) DIAS camera data with nine regions defined for the NN input (enclosed by red lines) and $\mu(\vec{r})$ (black crosses)

Figure 9: Front views of heat load representations on limiter in module 5 at discharge “20160309.007”.

(see Figure 7). The remaining frames i. e. time spans of stable plasma-limiter contact, serve as the basis for further processing.

2.4.2 Strike Line Simulation

The size of the training data set strongly determines the performance of the NN after it has been trained. Since the amount of data for limiter operation is low, especially for the OP1.1 subset of the \bar{t} -scan (cf. Table 1), the data set is complemented by simulated strike lines. For the proof of principle, we make use of several approximations, described below. The input creation can be broken down into sequential parts.

\vec{B} creation A fast algorithm based on the Biot–Savart law generates \vec{B} from the W7-X coil currents. β -effects like Pfirsch–Schlüter currents can be neglected since

the discharge β are small yet, where $\beta = \frac{p}{B^2/2\mu_0}$ is the ratio of thermal to magnetic pressure. The coil geometry used was created by modifying the ideal coil set, taking into account preload and electromagnetic forces as well as dead weight cool down at modular coil currents of 6 kA and planar coil currents of 2.3 kA each [37].

Field line diffusion The approach to simulate particles and heat transport from the plasma core to the edge is as follows: Pseudo-randomly distributed particles are generated on the LCMS, assuming that the energy flux across the LCMS is uniform on the surface. Then particles follow their field lines, taking small random steps perpendicular to the field to simulate cross-field diffusion. If a particle trace intersects a material component the trajectory is stopped and the particle is registered as having hit the surface of the object through which the particle entered the object. The algorithm is described in more detail in [38], and it neglects the effects of magnetic drifts. The relevant parameters for the diffusion process are perpendicular diffusion coefficient D_{\perp} , collisional mean free path λ and velocity v . Values assumed for these constants are $D_{\perp} = 1 \text{ m}^2/\text{s}$, $\lambda = 0.1 \text{ m}$ and four different values for v between 1.4×10^5 and $2.8 \times 10^5 \text{ m/s}$. They are empirically justified in [31, Chapter 3.3]. A number of 8000 diffusion traces proved to be fast as well as robust in terms of strike line characteristics.

DIAS emulation The strike zone is typically two strike lines, one on either side of the limiter. As only one side is fully covered by the DIAS view, the field diffusion points on the far side are mirrored (with stellarator symmetry: $x \rightarrow -x$, $z \rightarrow -z$) to that side. The simulation procedure described above provides the strike line heat load, represented by the strike point density (e.g. Figure 9a).

2.4.3 Strike Line Partitioning

To make the NN robust towards noise, averaging quantities are extracted from the strike line and serve as input.

The raw simulation and experiment data are significantly different because simplifying assumptions were made in the simulation (cf. Section 2.4.2). NNs trained on synthetic data but applied to reality may perform poorly. This problem is known in the machine learning community [39, Chapter 2.3.3].

In the following, processing steps are described to transform the simulated and experimental data sets in such a way that the NN can hardly distinguish them. The 80 % quantile is applied to the IR data as the threshold limit q_{thr} for the local heat load. For each triangle i on the CAD triangulated limiter, this threshold is implemented by setting $q_i = 0$ for $q_i < q_{\text{thr}}$. Each data set is split into nine regions, roughly following the natural geometry of the nine limiter tiles as shown in Figure 9b. For each of those nine regions, we compute the center of mass of the heat flux

$$\vec{\mu} = \frac{1}{\sum_{i=1}^n A_i q_i} \sum_{i=1}^n A_i q_i \vec{\kappa}_i, \quad (8)$$

with n triangles per section, where each triangle i has the properties centroid $\vec{\kappa}_i$, area A_i and heat flux q_i .

The NN inputs are

$$\Delta\vec{\mu} = \vec{\mu} - \vec{\mu}_{\text{ref}}, \quad (9)$$

with $\vec{\mu}_{\text{ref}} = \vec{\mu}(I_B = 0)$, separately for experiment and simulation. This processing is introduced in order to be independent of the origin of the coordinate system in each region and diminish the offset between the two data sets.

As an example, Figure 10 depicts all three spatial components of $\vec{\mu} = (\mu_x, \mu_y, \mu_z)^T$ for the section corresponding to tile 4. $\Delta\mu_x$ and $\Delta\mu_y$ show a similar behavior, except for a constant

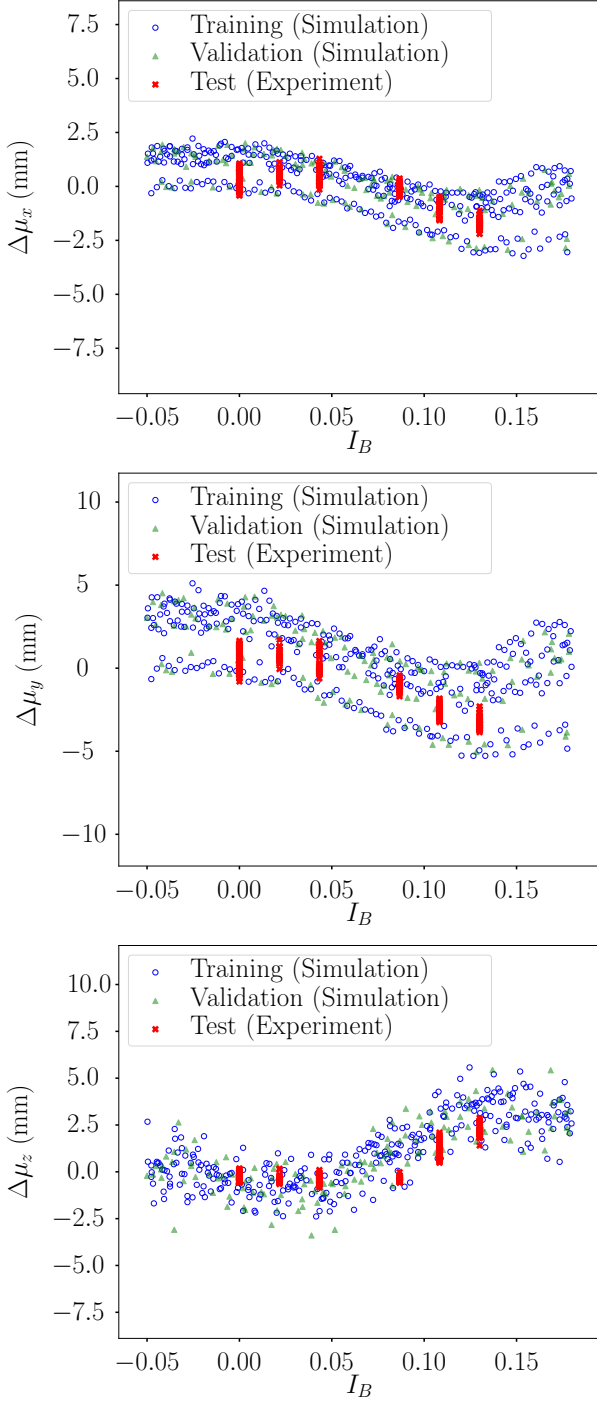


Figure 10: *The images show three different NN inputs depending on the planar coil current. The data presented in each graph are training set (blue), validation set (green) which are both based on simulations, and the experimental test set (red). The training and validation set are based on four different field line diffusion velocities to illustrate different collisionalities (cf. Section 2.4.2) which leads to distinguishable curves.*

scaling factor. Since the NN will rescale the input to a standard deviation of 1 (cf. 2.1), this factor vanishes. Thus only $\Delta\mu_x$ and $\Delta\mu_z$ are given to the NN. As expected, we observe systematic differences between simulation and experiment. For example, $\Delta\mu_z$ is slightly above the mean simulation result for $I_B < 0.05$ while it is below the simulation trend for the other half.

The variability of $\Delta\vec{\mu}$ stems from the random processes and different velocities used for field line diffusion simulation described in Section 2.4.2.

3 Results

In the following, we present the results of the I_B prediction. They were created with the same input parametrization (cf. Section 2.4.3) and NN architecture (see Section 2.1).

3.1 IR Data

A NN is trained and tested on all available IR data from the OP1.1 \bar{i} -scan. This means in particular, only six different I_B values are available for training, validation and testing. Figure 11 shows I_B , reconstructed from the strike line, in dependence of the actual target value of I_B . Here, as well as in all following figures, the NN validation is displayed by means of red circles. The independent test sets are depicted in terms of blue triangles. In case of an optimal reconstruction, target and NN output turn out equal, indicated by the dashed line. It can be seen that the NN is able to reconstruct I_B sufficiently well, since the points for the test set scatter closely and symmetrically around the target I_B . This is quantified by a *rmse* of only 0.010 for the validation as well as the test set. Because only six different magnetic configurations are available here, the NN is not expected to perform well for the interpolation between the

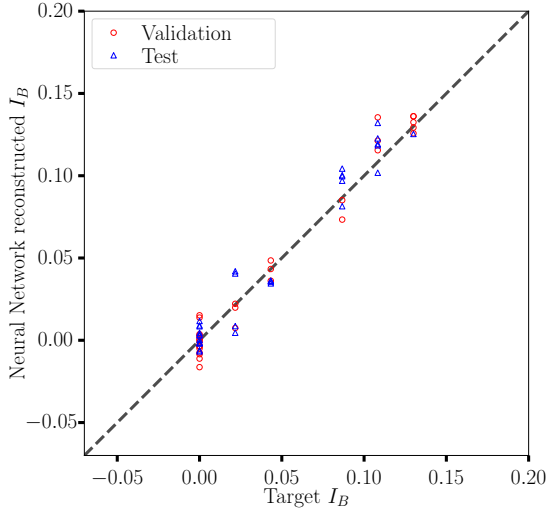


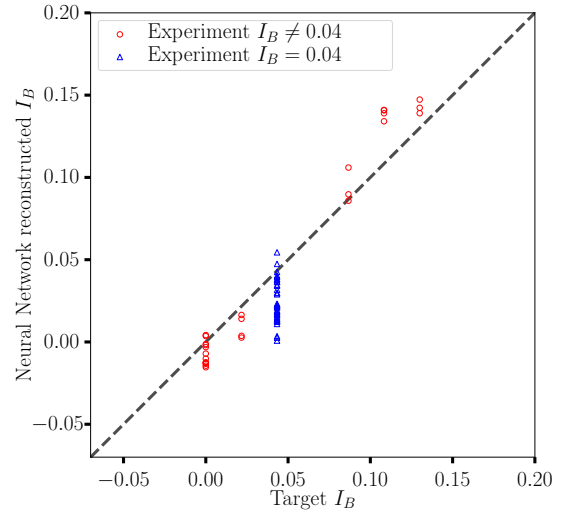
Figure 11: *NN performance, on IR data only. Red circles represent the validation set while blue triangles identify members of the test set.*

given currents. This can be seen in Figure 12a, where *rmse* for validation is 0.016 and 0.023 for the test set which is 40% larger. As one might expect, the *rmse* for extrapolation differs by more almost one order of magnitude, namely 0.007 for validation and 0.047 for the test set. Figure 12b shows this behavior.

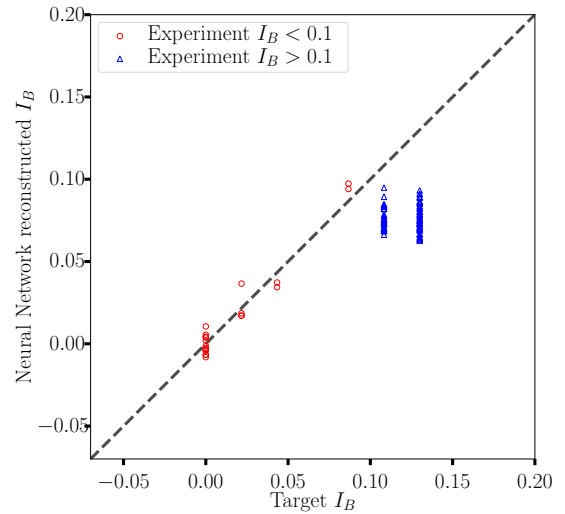
3.2 Combination of IR and Synthetic Data

To obtain a sufficient amount of training data, we simulate the limiter \bar{i} -scan described in Section 2.3. As with the previous result on experimental data, I_B is well reconstructed by a NN, trained with simulated data leading to an *rmse* of 0.014 in both validation and test set. Additionally, the NN trained with synthetic data can now be tested with IR data, i.e. in the process of training, the NN was never exposed to any subset of this test set, neither implicitly (validation set) nor explicitly (training set).

The test is performed with all DIAS frames



(a) *Interpolation performance, demonstrated by training and validating with five out of the six I_B , and testing on the remaining $I_B = 0.04$*



(b) *Extrapolation performance, demonstrated by training and validating with the lowest four out of the six I_B , and testing on the remaining two I_B values*

Figure 12: *NN performance on IR data corresponding to Figure 11. Red circles represent the validation set while blue triangles identify members of the test set.*

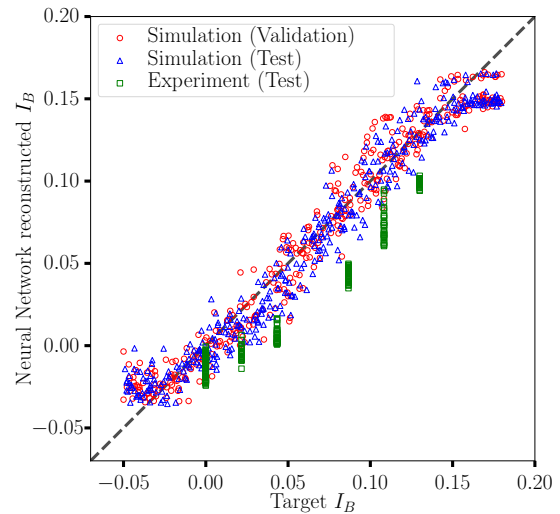
given in Table 1. The green squares in Figure 13a show the performance for this data set. The experimental data follow the trend of the dashed line with an $rmse$ of 0.029. We observe a small but clear and systematic underestimation of I_B of -0.026 ± 0.001 on average; this causes the larger $rmse$ as compared to the validation set. The reason for this offset is not yet clear to us but is under investigation. Though the test set does not perform as well as the validation set, the performance can already be considered a success, considering the relatively small discrepancies and the complexity of the problem (cf. Section 2.4.3).

We examine now whether training with a subset of experimental data improves the NN performance for such experimental data with no corresponding target I_B in neither training nor validation set. The only change as compared to Figure 13a is introduced by adding all IR data with target I_B below 0.1 to the training and validation set. This way, the NN should learn to neglect systematic differences between simulation and experiment. The discrepancy is significantly reduced as compared to the only simulation trained NN, quantified by an $rmse$ of 0.023. Still, the reconstruction underestimates I_B .

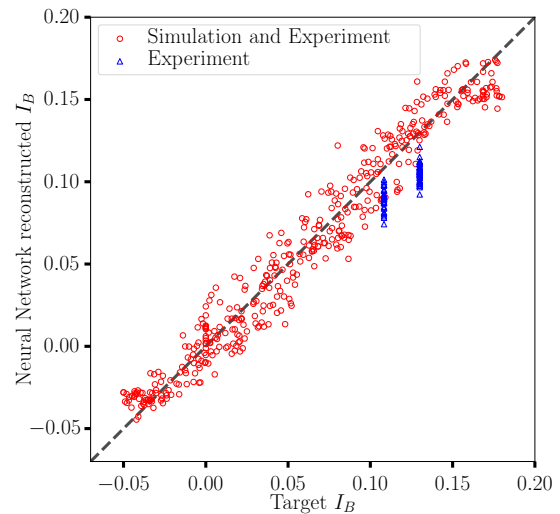
Note that the validation $rmse$ of 0.013 for this NN is similar to that of the NN presented in Figure 13a, although it is trained with a mixture of experimental and simulated data. In Section 3.1, we demonstrated the poor interpolation as well as extrapolation ability of the NN with few target I_B . The NN presented here clearly outperforms the extrapolation result discussed in Section 3.1.

4 Conclusion

It was shown here that neural networks are capable of reconstructing the magnetic configuration in terms of its \bar{l} value, with experimental observations of limiter heat loads performed



(a) *NN trained on simulation and tested with simulated and experimental data*



(b) *Training and validation set are supported by the full experimental data set with lower 2/3 of I_B to improve the reconstruction performance of the NN extrapolating towards upper 1/3 of I_B .*

Figure 13: *Performance of NN based on simulation data. Red circles represent the validation set while blue triangles and green squares identify members of the test set.*

during an \bar{l} -scan during the first experimental phase of Wendelstein 7-X. The NNs operated on real IR as well as simulated data. The NNs trained on simulated data can reconstruct configurations from IR data with sufficient accuracy. An even better performance is reached when the NN training set is based on a combination of simulated and parts of the IR data. It is expected that this approach will lead to a capability of extracting key physics parameters (of which \bar{l} is perhaps the most important) in real-time in future divertor operation, based on IR camera data and possibly other diagnostic signals as well. Such a system would play a major role in a real-time divertor operation control system on Wendelstein 7-X.

Further improvements in the ensemble of NN architecture and the data parametrization are subjects of ongoing work. In the upcoming OP1.2 with installed divertors we plan experiments with iota scans and non-negligible I_{tor} .

5 Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. We wish to acknowledge the helpful discussions with H.Hölbe, L. Rudischhauser, G. Leifert and T. Grüning. Our gratitude goes to H. Niemann and F. Pisano for providing the IR data geometry and heat flux mapping. We are also thankful for the support from the CITlab (University of Rostock) and Planet-AI teams.

Bibliography

- [1] Jr. Lyman Spitzer. “The Stellarator Concept”. In: *Physics of Fluids* 1.1958 (1958), pp. 253–264. ISSN: 00319171.
- [2] G. Grieger and I. Milch. “Das Fusionsexperiment WENDELSTEIN 7-X”. In: *Physik Journal* 49.11 (1993), pp. 1001–1005. ISSN: 00319279.
- [3] Craig Beidler et al. “Physics and Engineering Design for Wendelstein VII-X”. In: *Fusion Science and Technology* 17.1 (1990), pp. 148–168.
- [4] J. Nührenberg and R. Zille. “Stable stellarators with medium β and aspect ratio”. In: *Physics Letters A* 114.3 (1986), pp. 129–132. ISSN: 03759601.
- [5] Thomas Sunn Pedersen et al. “Key results from the first plasma operation phase and outlook for future performance in Wendelstein 7-X”. In: *Physics of Plasmas* 24.5 (2017), p. 055503. ISSN: 1070-664X.
- [6] Hermann Renner et al. “Physical Aspects And Design of the Wendelstein 7-X Divertor”. In: *Fusion Science and Technology* 46.2 (2004), pp. 318–326.
- [7] J. Boscary et al. “Design improvement of the target elements of Wendelstein 7-X divertor”. In: *Fusion Engineering and Design* 87.7-8 (2012), pp. 1453–1456. ISSN: 09203796.
- [8] R Brakel et al. *Specification of design loads for in-vessel components of w7-x*. Tech. rep. PLM: 1-AC-S0005.1. Max-Planck-Institut für Plasmaphysik, 2017.
- [9] Henri Greuner et al. “Cyclic heat load testing of improved CFC/Cu bonding for the W 7-X divertor targets”. In: *Journal of Nuclear Materials* 386-388 (2009), pp. 772–775. ISSN: 00223115.
- [10] Alan Peacock et al. “Status of High Heat Flux Components at W7-X”. In: *IEEE Transactions on Plasma Science* 42.3 (2014), pp. 524–532. ISSN: 0093-3813.

- [11] M. W. Jakubowski et al. “Development of infrared and visible endoscope as the safety diagnostic for steady- state operation of Wendelstein 7-X”. In: *QIRT 100* (2014).
- [12] Christopher M Bishop et al. “Real-time control of a Tokamak plasma using neural networks”. In: *Neural Computation*. Vol. 7. 1. 1995, pp. 1007–1013.
- [13] Barbara Cannas et al. “Dynamic Neural Networks for Prediction of Disruptions in Tokamaks”. In: *10th International Conference on Engineering Applications of Neural Networks (EANN 2007)* 284 (2007).
- [14] Jakob Svensson. “Artificial Neural Network Modelling in Nuclear Fusion Research and Engineering at the HET Tokamak”. Doctoral thesis. Royal Institute of Technology Stockholm, 2000.
- [15] Tom Michael Mitchell. *The discipline of machine learning*. Vol. 9. Carnegie Mellon University, School of Computer Science, Machine Learning Department, 2006.
- [16] P. J. Werbos. “Beyond regression: New tools for prediction and analysis in the behavioral sciences”. PhD thesis. Harvard University, 1974.
- [17] Christian Szegedy et al. “Going deeper with convolutions”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 1–9.
- [18] Kaiming He et al. “Deep residual learning for image recognition”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 770–778.
- [19] John Duchi, Elad Hazan, and Yoram Singer. “Adaptive subgradient methods for online learning and stochastic optimization”. In: *Journal of Machine Learning Research* 12.Jul (2011), pp. 2121–2159.
- [20] Diederik P. Kingma and Jimmy Lei Ba. “Adam: a Method for Stochastic Optimization”. In: *International Conference on Learning Representations 2015* (2015), pp. 1–15. ISSN: 09252312. arXiv: 1412.6980.
- [21] Sharan Chetlur et al. “cudnn: Efficient primitives for deep learning”. In: *arXiv preprint arXiv:1410.0759* (2014).
- [22] Martín Abadi et al. “Tensorflow: Large-scale machine learning on heterogeneous distributed systems”. In: *arXiv preprint arXiv:1603.04467* (2016).
- [23] James Bergstra et al. “Theano: A CPU and GPU math compiler in Python”. In: *Proc. 9th Python in Science Conf.* 2010, pp. 1–7.
- [24] Yann LeCun, Yoshua Bengio, et al. “Convolutional networks for images, speech, and time series”. In: *The handbook of brain theory and neural networks* 3361.10 (1995), p. 1995.
- [25] Henrique Steinherz Hippert, Carlos Eduardo Pedreira, and Reinaldo Castro Souza. “Neural networks for short-term load forecasting: A review and evaluation”. In: *IEEE Transactions on power systems* 16.1 (2001), pp. 44–55.
- [26] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “Imagenet classification with deep convolutional neural networks”. In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.
- [27] Olga Russakovsky et al. “Imagenet large scale visual recognition challenge”. In: *International Journal of Computer Vision* 115.3 (2015), pp. 211–252.
- [28] Xavier Glorot and Yoshua Bengio. “Understanding the difficulty of training deep feedforward neural networks.” In: *Aistats*. Vol. 9. 2010, pp. 249–256.
- [29] Ian H Witten et al. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

- [30] P. I. Strand and W. A. Houlberg. “Magnetic flux evolution in highly shaped plasmas”. In: *Physics of Plasmas* 8.6 (2001), pp. 2782–2792. ISSN: 1070-664X.
- [31] H. Hölbe. “Control of the magnetic topology and plasma exhaust in the edge region of Wendelstein 7-X : a numerical study”. PhD thesis. Universität Greifswald, 2016.
- [32] T. Andreeva. *Vacuum magnetic configurations of Wendelstein 7-X*. Tech. rep. Greifswald: Max-Planck-Institut für Plasmaphysik, 2002.
- [33] Holger Niemann et al. “First Power Decay Length Studies in the Limiter Phase of Wendelstein 7-X”. Unpublished Manuscript. 2017.
- [34] B Cannas et al. “Towards a new VIS/IR image processing system at W7-X: from spatial calibration to characterization of thermal events”. Unpublished Manuscript. 2017.
- [35] G. A. Wurden et al. “Limiter observations during W7-X first plasmas”. In: *Nuclear Fusion* 57.5 (2017), p. 056036. ISSN: 0029-5515.
- [36] A. and The ASDEX Upgrade Team Herrmann. “Limitations for divertor heat flux calculations of fast events in tokamaks”. In: *Proc. 28th EPS Conf. on Controlled Fusion and Plasma Phys.* Vol. 25A. 2001, p. 2109.
- [37] K Egorov. *ANSYS GM 504: FE Model Description*. Tech. rep. Greifswald: Max-Planck-Institut für Plasmaphysik, 2011.
- [38] S.A. Bozhnikov et al. “Service oriented architecture for scientific analysis at W7-X. An example of a field line tracer”. In: *Fusion Engineering and Design* 88.11 (2013), pp. 2997–3006. ISSN: 09203796.
- [39] Jens Kober and Jan Peters. “Learning Motor Skills: From Algorithms to Robot Experiments”. In: *Springer Tracts in Advanced Robotics* 97 (2014). ISSN: 1610742X.

A.2. Article II

»Neural network performance enhancement for limited nuclear fusion experiment observations supported by simulations«

M. BLATZHEIM, D. BÖCKENHOFF, H. HÖLBE, T. S. PEDERSEN and R. LABAHN

Nuclear Fusion, Vol. 59.1 (2019), DOI: 10.1088/1741-4326/aaefaf

Neural network performance enhancement for limited nuclear fusion experiment observations supported by simulations

Marko Blatzheim^{1,2}, Daniel Böckenhoff¹, Hauke Hölbe¹, Thomas Sunn Pedersen¹, Roger Labahn² and The W7-X Team¹

¹ Max Planck Institute for Plasma Physics, Wendelsteinstraße 1, Greifswald 17491, Germany

² Institute for Mathematics, University of Rostock, Ulmenstraße 69, Rostock 18057, Germany

E-mail: marko.blatzheim@ipp.mpg.de and daniel.boeckenhoff@ipp.mpg.de

Received 5 September 2018, revised 19 October 2018

Accepted for publication 9 November 2018

Published 5 December 2018



Abstract

It has recently been shown that artificial neural networks (NNs) are able to establish nontrivial connections between the heat fluxes and the magnetic topology at the edge of Wendelstein 7-X (W7-X) (Böckenhoff *et al* 2018 *Nucl. Fusion* **58** 056009), a first step in the direction of real-time control of heat fluxes in this device. We report here on progress on improving the performance of these NNs. A particular challenge here is that of generating a suitable training set for the NN. At present, experimental data are sparse, and simulated data, which are much more abundant, do not match the experimental data closely. It is found that the NNs show significantly improved performance on experimental data when experimental and simulated data are combined into a common training set, relative to training performed on only one of the two data sets. It is also found that appropriate pre-processing of the data improves performance. The architecture of the NN is also discussed. Overall a significant improvement in NN performance was seen—the normalized error reduced by more than a factor of three over the previous results. These results are important since heat flux control in a W7-X, as well as in a future fusion power plant, is likely a key issue, and must start with a very limited set of experimental training data, complemented by a larger, but not necessarily fully realistic, set of simulated data.

Keywords: neural networks, Wendelstein 7-X, plasma edge, magnetic topology

(Some figures may appear in colour only in the online journal)

1. Introduction

In machine learning, the size of the data set on which an artificial neural network (NN) is trained is crucial for a good performance. However, there are many cases where the cardinality of real (i.e. measured) data is limited. This issue may be addressed by complementing the available data with simulations. The underlying model for a simulation is generally

not complete. Simplifications are made, in particular to make the model computationally solvable within a reasonable time, given available computing power. Therefore, simulations and experiments often show systematic differences. On the other hand, these reduced models are useful since they reproduce many, and often the most important, measured features of the experimental data.

Other important factors for the performance of a NN are the architecture of the NN [2, 3] and appropriate parametrization [4] of the input data.

These rather general issues are present, and will be addressed in this article, for the specific case of heat loads onto the plasma-facing components of the fusion experiment



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

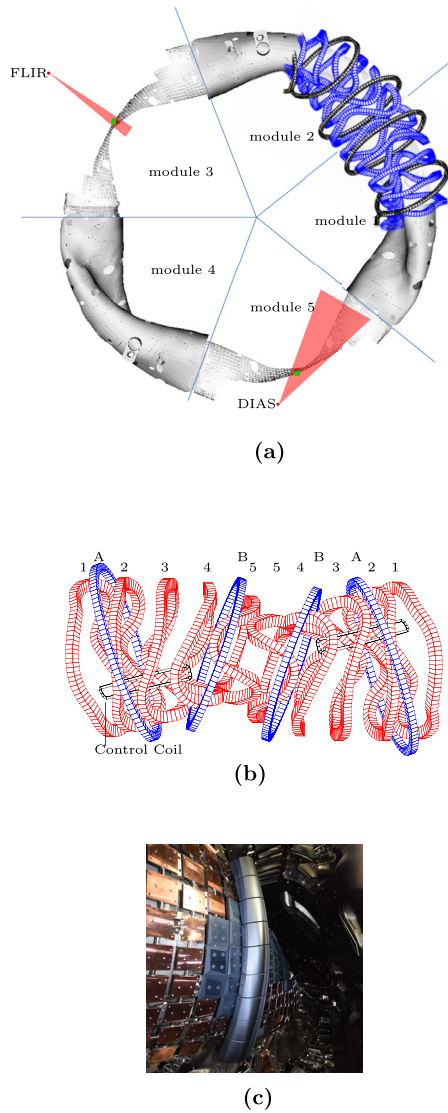


Figure 1. Overview of essential parts of W7-X in its first experimental campaign. (a) Top-down CAD view of the W7-X inner vessel, showing sight lines of the IR camera system for the limiter setup with cutaways in modules three and five as used in the first experimental campaign. On this scale and view, the limiters are small (green). One segment of the total 50 modular (blue) and 20 planar coils (black) is overlaid in modules 1 and 2. (b) Coils contained in one module with modular coils 1–5 (red) and planar coils A, B (blue). One module is point symmetrical towards its center. Adapted with permission from [8]. (c) Side view of the W7-X limiter in module 5.

Wendelstein 7-X (W7-X) [5], which recently went into operation. Experimental data are still scarce, and the interactions between the edge plasma and the plasma-facing components are multifaceted and complicated, involving at least plasma physics, atomic physics, chemistry, and solid state physics, so that state-of-the-art codes such as EMC3-Eirene [6, 7] do not yet capture all the important dynamics. The focus is on developing NNs that, trained on simulated data, supplemented with a very limited number of experimental data, generalize well onto experimental data. Our specific example is the training of a NN using simulated and experimental observations of W7-X

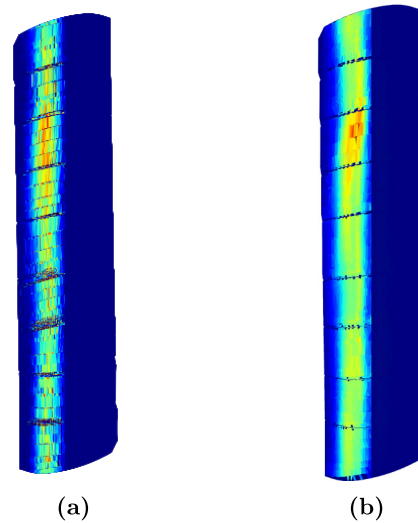


Figure 2. Front views of heat load representations on limiter in module 5 at discharge ‘20160309.007’. The right side of the limiter is ignored since it is largely obstructed from the camera view. (a) Heat load, represented by strike point density from a field line diffusion simulation. (b) Heat load, calculated from DIAS camera data.

heat load patterns to reconstruct ι , a property of the magnetic field at the plasma edge that determines the heat load patterns. The future practical application of this NN will be a part of a real-time control system, ensuring the control and safety of all W7-X plasma facing components.

In the next section we briefly explain the essentials regarding W7-X, heat load pattern retrieval from infra-red cameras and a proper formulation of ι , followed by a definition of the data set composition. The two used NN architectures are described and the parametrizations introduced before the NN performance is presented and analyzed.

2. Methods

2.1. Context for the neural network application

W7-X is a world-leading, relatively large experiment of the stellarator type [5] (16 m outer diameter, 30 m³ confinement volume). It has a carefully tailored magnetic field configuration designed to confine hot plasma and aims to explore if this concept can be scaled up to yield net energy production from thermonuclear fusion processes. This paper focuses on the optimization of a neural network that is to become an important piece of a real-time control system for the heat and particles exhausted by the plasma onto the components specially designed to absorb these heat and particle loads. For the present studies, these plasma-intercepting components were so-called limiters (figure 1(c)), but later divertor modules are used. Perhaps the most important parameter determining the spatial heat load distribution onto these components is the magnetic winding number ι . It describes how many full poloidal turns a magnetic field line performs when performing one full toroidal turn, and is around 0.9 for these studies.

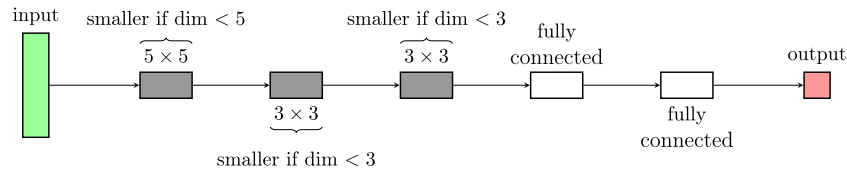


Figure 3. Structure of NN with three convolutional layers: the grey elements represent convolutional layers with their associated kernel sizes indicated.

Therefore, we started investigating if a NN could determine t given, as input, measurements by infrared (IR) cameras of heat load patterns. Figure 1(a) shows an overview of W7-X and IR camera views. Our first results were recently published showing that indeed this is possible [1]. The present paper is a systematic attempt to further improve the performance of NNs for this application, by investigating how the pre-processing of the input data, the character and quantity of training sets, and the NN architecture affect the performance.

Following the previous paper, we train the NNs to estimate the current in coil set B (figure 1(b)), I_B , appropriately normalized, instead of t itself, but this is a technical detail. For these studies, numerical and experimental, there is a one-to-one correspondence between the two.

For details regarding the underlying physics and data origin (experimental as well as synthetic) we refer to [1]. Figure 2 shows an example of a heat load pattern on the limiter from W7-X module 5 for vacuum field line diffusion simulation (figure 2(a)) as well as IR data (figure 2(b)) at the same magnetic configuration. The right half of the limiter is shadowed from the IR camera view. Some basic characteristics are similar, e.g. the maximum heat load is located at the third limiter tile. However, in detail, the structure of the IR observation is not reproduced.

2.2. Data sets

The investigated data sets result from experimental and simulated t scans of W7-X. The scans were performed by varying the current in one coil set, planar coil set B (see [1, section 2.3]). The simulation set \mathbb{S} was created by the field line diffusion approach described in [1, section 2.4.2] with $|\mathbb{S}| = 3993$.

The experimental set of processed IR data \mathbb{I} comes from six different experimental magnetic configurations that were investigated. For each magnetic configuration, two to five sets of data were taken, giving a total of 16 IR videos (see [1, section 2.4.1]).

Each IR video corresponding to one experiment contributes on average 20 frames for the same value of I_B , leading to a total cardinality of $|\mathbb{I}| = 319$.

A subset $\mathbb{I}_C \subset \mathbb{I}$ with $|\mathbb{I}_C| = 190$ is defined such that \mathbb{I}_C as well as $\mathbb{I} \setminus \mathbb{I}_C$ cover all six experimental settings. Each of these two sets contain either all or none of the frames of each single experiment.

A mixed set is defined as $\mathbb{M} = (\mathbb{S} \cup \mathbb{I}) \setminus \mathbb{I}_C$.

To determine NN quality, three disjoint subsets, namely training set, validation set and test set have to be defined.

2.3. NN architectures

Three NN architectures were compared. Since the data set is very limited, the number of free parameters must be kept low to avoid overfitting. The first one is a fully-connected feed-forward neural network (FFNN) with 3 layers of 64 nodes each with tanh activation function and L2-regularization with a factor of 0.001. The second one is mainly based on convolutional layers [9] as shown in figure 3. Three consecutive convolutional layers are followed by two fully-connected layers with 64 and 1 nodes, respectively. The number of filters in the convolutions are (in ascending order) 4, 8 and 16 with a kernel slide of 1. The third NN architecture starts with an inception module as defined in [10] but instead of a very deep structure as suggested for the large ImageNet data set [11] it is followed by pooling [12], a 1×1 convolutional layer and two fully-connected layers, as shown in figure 4. The number of filters in all convolutions within the inception is 4, it is 8 for the following convolution and there are 16 and 1 nodes within the last two fully-connected layers, respectively. The pool kernel within the inception module has the dimension 3×3 and the pool kernel dimension is 2×2 for the one following the inception module. All the kernel slides are 1 except for the pool following the inception, which has a slide of 2. In the last two architectures all activation functions except for the last layer are rectified linear units (ReLU). Because the NNs are designed to solve a regression problem, the activation function of the last layer is the identity. The analysis of the parameters presented in the following section 2.4 is mainly performed using the convolution neural network. For the other two networks only examples are shown.

Weights are initialized randomly as recommended in [13] and biases are initialized as zero. The weights and biases are iteratively improved by the *adam* optimizer [14] during the training process. Early-stopping is applied as soon as the validation loss-function increases. The implementation is done in TensorFlow [15]. The calculations are computed on a workstation with two Intel Xeon CPU E5 – 2650 and four NVIDIA GeForce GTX 1080 Ti.

2.4. Parametrization

The heat load is given on an unstructured triangular grid that represents the CAD structure of one limiter. This data is transformed affine from W7-X coordinates to an orthogonal ξ, η, ζ -coordinate system, where ζ points in the direction of \vec{n}_m , the normalized mean of the normals of all triangles forming the limiter. The rotation is achieved by the matrix \mathbf{R} with

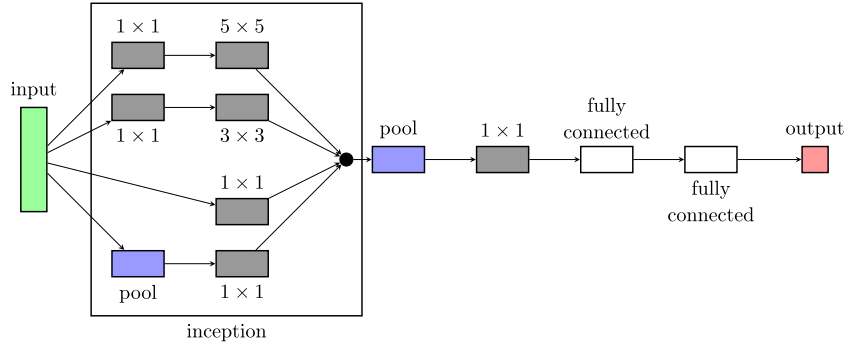


Figure 4. Structure of NN with inception module: the grey elements represent convolutional layers with associated kernel sizes indicated. The blue elements are max-pooling layers and the white elements are fully-connected layer.

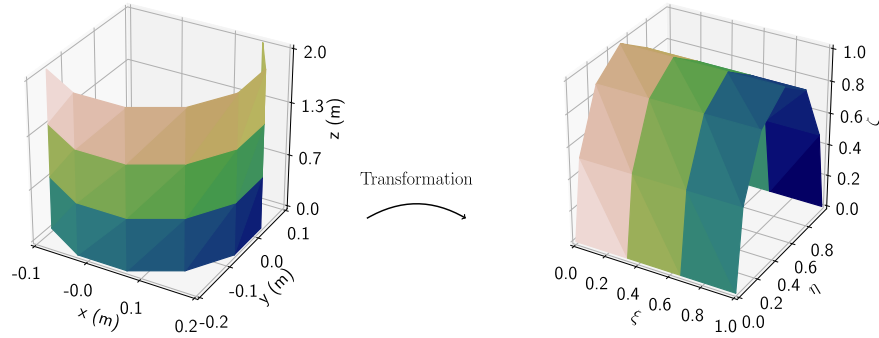


Figure 5. Sketch of the affine transformation from Cartesian coordinates x, y, z to ξ, η, ζ .

$$\mathbf{R} = \begin{pmatrix} 1 - \frac{n_{m,1}^2}{1+n_{m,3}} & -\frac{n_{m,1}n_{m,2}}{1+n_{m,3}} & -n_{m,1} \\ -\frac{n_{m,1}n_{m,2}}{1+n_{m,3}} & 1 - \frac{n_{m,2}^2}{1+n_{m,3}} & -n_{m,2} \\ n_{m,1} & n_{m,2} & n_{m,3} \end{pmatrix}. \quad (1)$$

Each axis is scaled such that the minimal value in this coordinate is 0 and the maximum value is 1. An example for such an affine transformation is shown in figure 5 for a half cylinder with remote resemblance to the W7-X limiter. Only triangles within a tight bounding box around the limiter are considered for the parametrization.

2.4.1. Partitioning. The ξ, η -space is then divided by various 2D grids with sizes described in table 1 and shown for simulations and experiments in figures 6(a) and (b) respectively. There are no divisions in the ζ direction.

The first three columns in table 1 show one dimensional partitionings. The inception NN is applied only for inputs of dimensionality of at least 5×5 . However, for the one-dimensional inputs the convolutional kernel dimensionality is reduced from that presented in figure 3 such that it does not exceed the input dimensionality.

2.4.2. Extracted characteristics. For each element p of the partitionings, characteristic values can be extracted. One is the heat-load-weighted spatial mean

$$\vec{\mu}_p = \Xi_p \mathbf{W}_p, \quad (2)$$

with triangle centroids

$$\Xi_p = \left((\xi_{pi}, \eta_{pi}, \zeta_{pi})_{i=1, \dots, n_p} \right)^T \in \mathbb{R}^{3 \times n_p} \quad (3)$$

and weight vector \mathbf{W}_p defined as

$$\mathbf{W}_p = \left(\frac{A_{pi} q_{pi}}{\sum_{j=1}^{n_p} A_{pj} q_{pj}} \right)_{i=1, \dots, n_p} \in \mathbb{R}^{n_p}, \quad (4)$$

with number of triangles per partitioning n_p , triangle areas A_{pi} , and triangle heat load q_{pi} (in Wm^{-2}). The weighted covariance matrix is defined as

$$\begin{aligned} \mathbf{Cov}(\Xi_p) \\ = \left(\Xi_p - \vec{\mu}_p \mathbf{J}_{n_p}^T \right) \left(\left(\Xi_p - \vec{\mu}_p \mathbf{J}_{n_p}^T \right) \odot \mathbf{W}_p \right)^T \in \mathbb{R}^{3 \times 3}, \end{aligned} \quad (5)$$

with vector of ones \mathbf{J}_{n_p} of dimension n_p and element-wise multiplication \odot . Another statistical characteristic is the spatial standard deviation $\vec{\sigma}_p$ calculated as

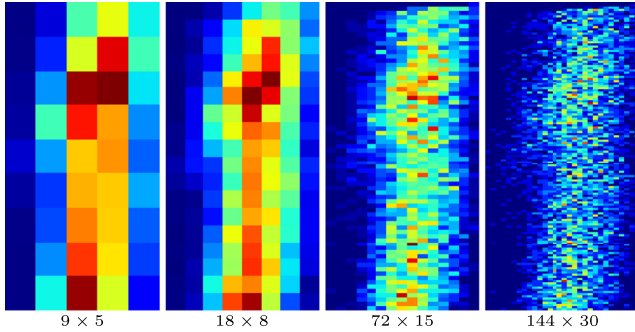
$$\vec{\sigma}_p = \sqrt{\text{diag}(\mathbf{Cov}(\Xi_p))}, \quad (6)$$

with element-wise square root and the operator $\text{diag}()$ which extracts the diagonal elements of a matrix. The direction vector $\vec{\delta}_p$, calculated as the eigenvector corresponding to the largest eigenvalue λ_{\max} of $\mathbf{Cov}(\Xi_p)$ can be characteristic as well.

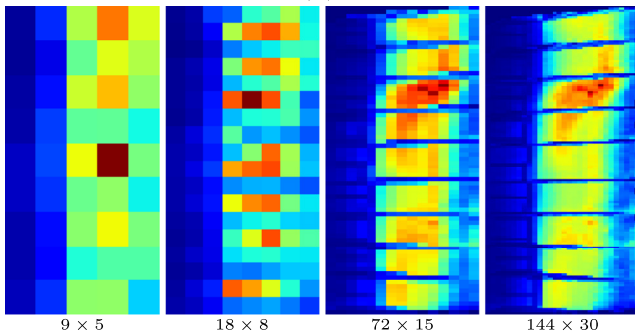
This parameter is inspired by divertor heat load patterns which show a more complicated shape [16]. The last examined parameter is the relative heat load

Table 1. Dimensions of the used partitionings and total number of elements.

n_ξ	2	4	9	9	15	18	27	36	72	144
n_η	1	1	1	5	6	8	10	12	15	30
$n_\xi \cdot n_\eta$	2	4	9	45	90	144	270	432	1080	4320



(a)



(b)

Figure 6. Comparison of relative heat load between simulation and experiment for four exemplary partitionings. The color scale ranges from 0 to the maximal value of each plot respectively.

(a) Simulation. (b) Experiment.

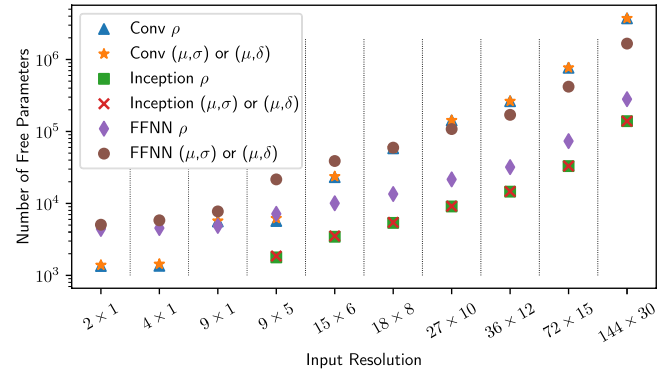
$$\rho_p = \frac{\hat{q}_p}{\sum_{p'=1}^m \hat{q}_{p'}} \quad (7)$$

with number of partitions m and

$$\hat{q}_p = \frac{\sum_{i=1}^{n_p} A_{pi} q_{pi}}{\sum_{i=1}^{n_p} A_{pi}} \quad (8)$$

defined as the absolute heat load.

Three combinations of those parameters are studied as NN input: (μ, σ) , (μ, δ) and ρ exclusively. In figure 6 ρ is shown for simulation and experiment of the same physical condition. Note that the input dimension of the NNs is batch size $\times n_\xi \times n_\eta \times$ input channels, with input channels being 6 for the cases (μ, σ) as well as for (μ, δ) but 1 for ρ . In the case of ρ as parameter, the 1×1 convolutions within the inception module (see section 2.3) perform rather trivial operations. Whether this affects the performance was not investigated. Pooling layers are used for all partitionings to achieve comparable architectures, although they may not be important for smaller partitionings. The number of free parameters resulting from the chosen NN type, partitioning and characteristic is shown in figure 7.

**Figure 7.** Free parameters of the NNs depending on the partitioning, the parametrization and architecture type logarithmic scaled in the y-axis. Note, that for the convolutional NNs and the inception NNs the number of free parameter is almost independent from the parametrization.

We define the following notation to describe the NN settings:

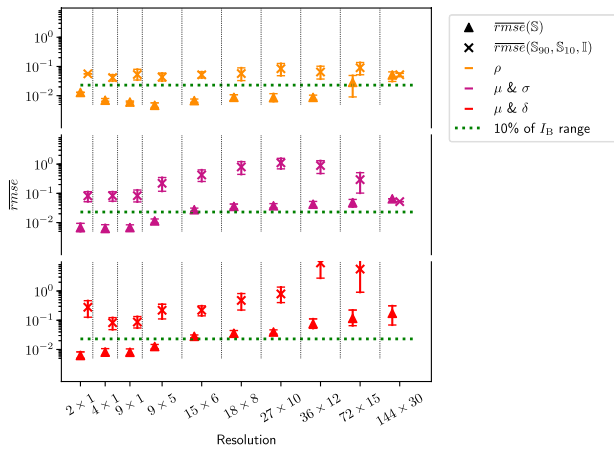
$$f(\text{train, validate, test})_{\text{partition, architecture}}^{\text{parametrization}}, \quad (9)$$

where f can be any function.

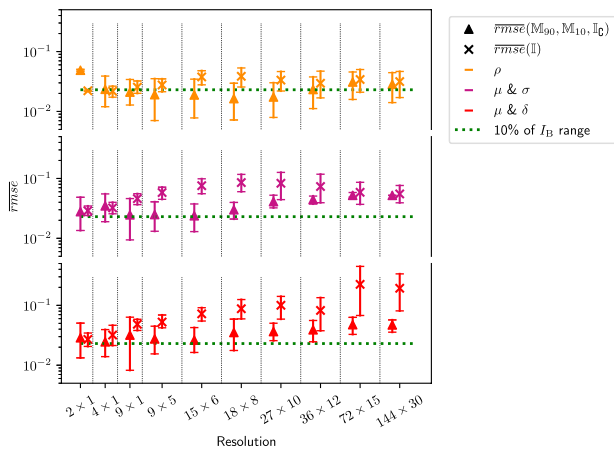
For example $rmse(\mathbb{S}_{90}, \mathbb{S}_{10}, \mathbb{I})_{9 \times 5, \text{inception}}^\rho$ refers to the root mean square error of an inception NN trained as well as validated on samples from set \mathbb{S} and tested with experimentally observed IR data \mathbb{I} , requiring a 9×5 input of relative intensities ρ . $\mathbb{S}_{90} \subset \mathbb{S}$ and $\mathbb{S}_{10} \subset \mathbb{S}$ with $\mathbb{S}_{90} \cap \mathbb{S}_{10} = \emptyset$ refer to two disjunct sets for training and validation consisting of 90% and 10% of the samples within \mathbb{S} respectively. Given a single term in the brackets implies that training, validation and test data sets are disjunct and subsets of the same super set. If one describing parameter is omitted, the entirety of all possible parameters of that kind is referred to. So $(\mathbb{S})_{9 \times 5}^\rho$ describes all NNs parametrized by ρ and partitioned into 9 times 5 parts with both of the two considered architectures. They are trained, validated and tested on subsets of \mathbb{S} .

3. Results

The convolutional NN performance for all tested settings is shown in figure 8. These two graphics depict the $rmse$ dependence on the partitionings defined in section 2.4.1. The NN performances measured in terms of the $rmse$ are divided into 10 groups representing the partitioning. Within each group the parameter choices are shown. To avoid false conclusions by statistical outliers, an ensemble of 27 NNs trained with the same settings has been calculated. Variations are the randomness of the weight initialization and the mini batch sampling as well as different learning rates and batch sizes. Markers and bars indicate the mean $rmse$ (\overline{rmse}) and the associated 95% confidence interval for the mean respectively. The confidence interval has been calculated by bootstrapping [17]. The true possible values of I_B range between -0.05 and 0.18 . In order to facilitate the evaluation of the NN reconstruction quality, a reference value of 10% of the total I_B range is marked by the dotted, green line. The figures 9–11 depict subsets of the



(a)



(b)

Figure 8. Average $rmse$ with 95% confidence level for various parametrizations and limiter partitionings. In case of smaller $rmse$ than 10% of the I_B range a sufficient reconstruction of I_B has been achieved. (a) Logarithmic representation of $\overline{rmse}(\mathbb{S})$ and $\overline{rmse}(\mathbb{S}_{90}, \mathbb{S}_{10}, \mathbb{I})$. (b) Logarithmic representation of $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})$ and $\overline{rmse}(\mathbb{I})$.

outcomes shown in figure 8 to clarify the observations, while figure 13 compares the different NN architectures as an additional hyperparameter.

3.1. Simulation trained NNs

The results of the NNs trained on \mathbb{S} are shown in figure 8(a). It can be seen that $(\mathbb{S})^\rho$ performs well with both NN architectures for partitionings up to 36×12 . The \overline{rmse} is minimal at the partitioning 9×5 and increases with finer as well as coarser resolution. The NNs $(\mathbb{S})^{\mu \& \sigma}$ and the $(\mathbb{S})^{\mu \& \delta}$ perform especially well for the coarse partitionings between 2×1 and 9×5 . For finer partitionings the performance decreases gradually. This can be understood contemplating the decreasing information content per section with shrinking section size.

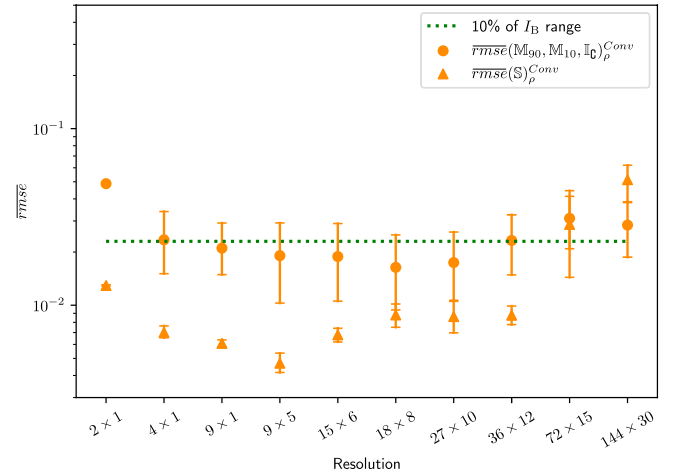


Figure 9. The development of $\overline{rmse}(\mathbb{S})_\rho^{\text{Conv}}$ and $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_\rho^{\text{Conv}}$ are compared for the increasingly refined partitionings. Especially for coarse and extremely fine partitionings the behavior is similar.

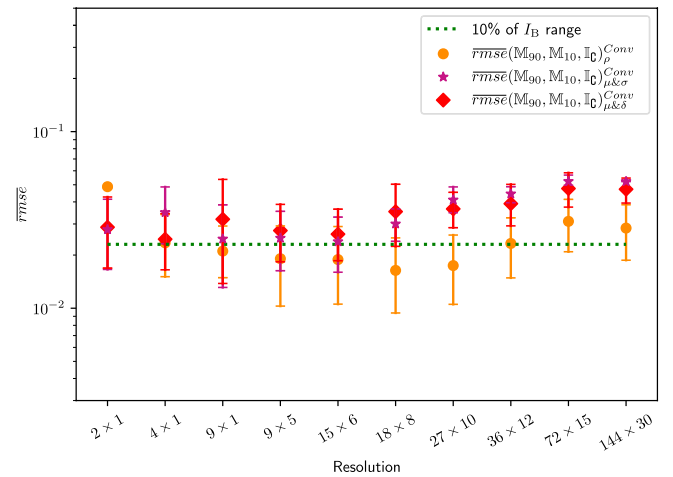


Figure 10. The development of $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_\rho^{\text{Conv}}$, $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\mu \& \sigma}^{\text{Conv}}$ and $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\mu \& \delta}^{\text{Conv}}$ are compared for the increasingly refined partitionings. Over all partitionings $\mu \& \sigma$ and $\mu \& \delta$ behave similarly while ρ performs clearly better for fine partitionings.

On the basis of figure 8(a) we investigate the performance of NNs characterized by (\mathbb{S}) versus $(\mathbb{S}_{90}, \mathbb{S}_{10}, \mathbb{I})$. A good performance of the NN $(\mathbb{S}_{90}, \mathbb{S}_{10}, \mathbb{I})$ would be advantageous since it would indicate applicability to new, never conducted experiments. Although it would be preferable, the evaluation indicates that this is not the case.

We observe not only that $\overline{rmse}(\mathbb{S}) \ll \overline{rmse}(\mathbb{S}, \mathbb{S}, \mathbb{I})$, but also that $\overline{rmse}(\mathbb{S}, \mathbb{S}, \mathbb{I})$ significantly exceeds the 10% I_B range. The NNs $(\mathbb{S}_{90}, \mathbb{S}_{10}, \mathbb{I})$ are specialized onto patterns of \mathbb{S} . Those patterns are not suitable to determine I_B from experimental data. The fundamentally different magnitude of the width of the \overline{rmse} confidence intervals $CI(\mathbb{S}, \mathbb{S}, \mathbb{I})$ as compared to $CI(\mathbb{S})$ points towards the same reason.

Both parametrizations including μ show a decreasing performance with growing resolution. Only a marginal difference between $rmse(\mathbb{S})^{\mu \& \delta}$ and $rmse(\mathbb{S})^{\mu \& \sigma}$ can be observed. The $rmse(\mathbb{S})$ range seems independent of the architecture.

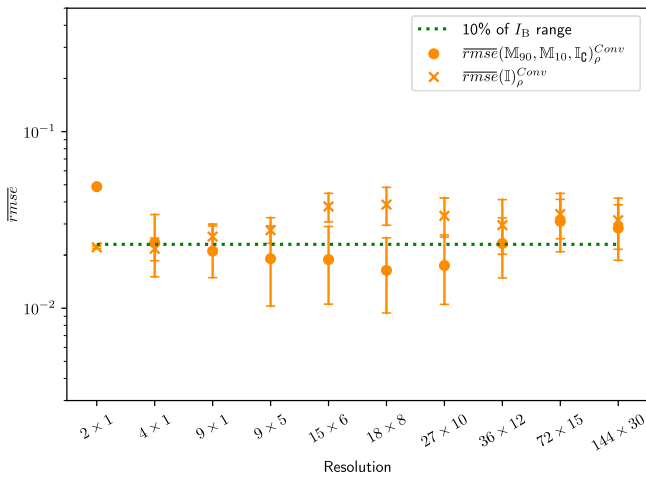


Figure 11. The development of $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\rho}^{\text{Conv}}$ and $\overline{rmse}(\mathbb{I})_{\rho}^{\text{Conv}}$ are compared for the increasingly refined partitionings. Especially for partitionings between 9×5 and 27×10 the mixed training leads to a clearly better NN. The performance for 18×8 is shown in more detail in figure 12.

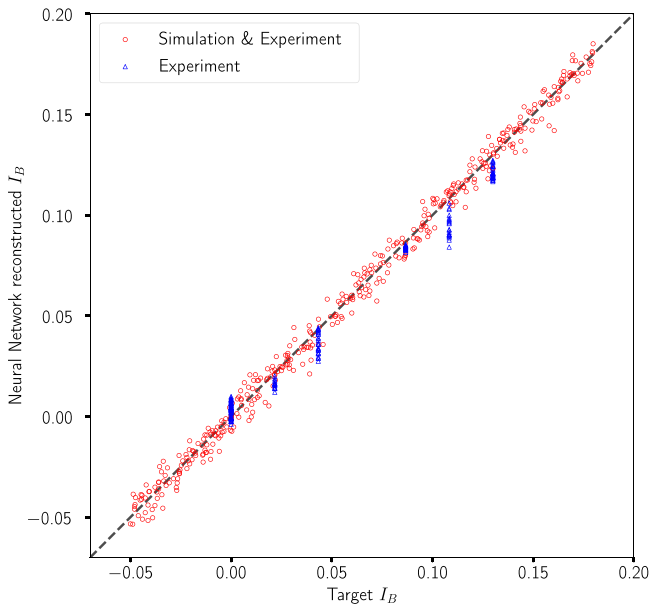


Figure 12. For the median of $rmse(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\rho}^{18 \times 8, \text{Conv}}$ the fitting performance on the validation set (red circles) and test set $\mathbb{I}_{\mathbb{C}}$ (blue triangles) is shown. On the x -axis are the target values and on the y -axis are the NN reconstructions. An ideal result would be the identity, shown by the dashed line.

3.2. NN trained with simulation and experiment

Since \mathbb{S} based training did not lead to sufficient NN performance for application to \mathbb{I} , some samples from \mathbb{I} are provided during training and validation, i.e. \mathbb{M} and $\mathbb{I}_{\mathbb{C}}$ as defined in section 2.2 are used. With this procedure, we intend to force the NNs to consider patterns present in both \mathbb{S} and \mathbb{I} during training. The performance of the NNs $(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})$ is compared to NNs trained, validated and tested with the small amount of available experimental data only, i.e. $(\mathbb{I}_{37}, \mathbb{I}_4, \mathbb{I}_{\mathbb{C}})$. Note, that the performance is tested with the same set $\mathbb{I}_{\mathbb{C}}$. Figure 8(b) depicts this comparison. The upper end of the occurring \overline{rmse} range is reduced by two orders of magnitude in this figure as compared

Table 2. Mean calculation times in ms for one training cycle until the early-stopping criteria is reached, and calculation time for a single evaluation of the validation set for the best performing partitioning with 18×8 elements and the largest one with 144×30 elements.

Partitioning	Conv	Inception	FFNN
18×8 Training	$0.9 \cdot 10^5$	$4.2 \cdot 10^5$	$1.0 \cdot 10^5$
144×30 Training	$1.1 \cdot 10^5$	$1.9 \cdot 10^5$	$0.6 \cdot 10^5$
18×8 Validation	8	10	5
144×30 Validation	17	19	9

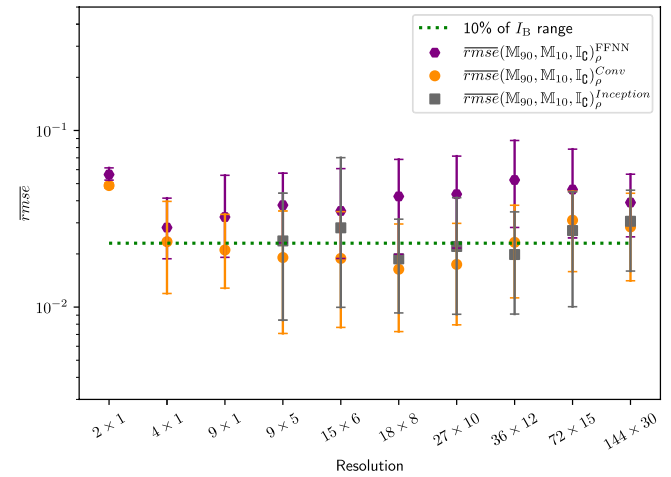


Figure 13. The performance between the fully-connected feed-forward NN, the convolutional NN and the inception NN is compared by examining the corresponding $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\rho}$. While the feed-forward NN is clearly outperformed regarding the precision, it can not be observed that the convolutional architecture is significantly better than the inception NN.

to 8(a). As in the case of NNs trained with \mathbb{S} , the parametrizations by ρ yield the best \overline{rmse} for partitionings between 9×5 and 36×12 while the μ based parametrizations are best for one dimensional partitionings decreasing in performance with growing resolution as visualized in figure 10.

For two-dimensional partitionings, we observe in figure 11 $\overline{rmse}(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\rho} < \overline{rmse}(\mathbb{I})$, as intended by training on \mathbb{M} . In the case of coarse partitioning, training with \mathbb{M} has no advantages above direct training with \mathbb{I} .

The most impressive results were found for resolutions between 9×5 and 36×12 , for $(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\rho}^{\text{Conv}}$. In order to display the good reconstruction quality, we choose a representative $rmse$, namely median $rmse$, of the NNs $(\mathbb{M}_{90}, \mathbb{M}_{10}, \mathbb{I}_{\mathbb{C}})_{\rho}^{18 \times 8, \text{Conv}}$, as seen in figure 12. The $rmse$ is 0.008 which is 3.5% of the I_B range, so the median $rmse$ is clearly below the 10% I_B range threshold. The results clearly outperform those of the NN in [1], where the best $rmse$ was 0.029.

In table 2 the mean training and validation times of NN with 18×8 and 144×30 data input are listed, which represent the best performing NNs and the largest NNs respectively. The training time is measured from the beginning to the end of the optimization process, while the validation time represents the time it takes to calculate the loss for the validation set once. Training time is in the order of minutes and does not only depend on the NN size but also on the early-stopping

criteria. The complete validation set is evaluated within milliseconds, which guarantees real-time applicability.

Changing the NN architecture to an inception model seems to slightly but not significantly improve the performance for finer partitioning as seen in figure 13. This behavior is not completely unexpected, as the inception model does not take advantage of the ‘parallelization’ unless very large problem sizes occur. Such larger sizes will be relevant in the future when dealing with the more complex divertor of W7-X instead of the limiter. The FFNN is clearly outperformed by the convolutional NN as well as by the inception NN.

4. Conclusion and future work

It was shown here that it is possible to reconstruct an important property of the W7-X edge magnetic field structure from limiter heat load patterns with even better accuracy than earlier reported in [1]. The main challenge was to deal with sparse experimental data given. A naive approach to apply NNs trained and validated with synthetic data to experimental data showed good performance only in a minority of cases. After such a training process most NNs focus on patterns not present in the experimental observations.

For a more targeted training and validation, a mixture of experimental and synthetic data is formed for the training process. This approach resulted in convincing NN performance for certain NN input processing. Partitioning the limiter with resolutions between 9×5 and 36×12 and defining the NN input as the heat load in each part divided by the maximum heat load of all parts results in better performance compared to NNs trained, validated and tested with experimental data only. The low number of experimental results probably leads to overfitting in these nets but the added simulation data diminished these effects. We created NNs that extract relevant patterns from experimental as well as from synthetic data sets to reconstruct an important parameter of the magnetic field at the edge. With this systematic approach NNs were found to outperform the results found in [1].

The upgraded W7-X with installed divertors will be the next object of interest. We will start the investigation with a parametrization based on a two-dimensional partitioning of the heat load. Favoring one of the two examined NN architectures *a priori* and excluding the other is not possible at this stage because neither consistently outperforms the other. The reached results are satisfactory, however it remains future work to investigate other methods such as generative adversarial nets [18] to further enhance the reconstruction performance when dealing with simulated and experimental data. These could learn to generate ‘experimental data’ and provide an additional supplement for real experimental data to train NNs which reconstruct plasma properties.

Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from

the Euratom research and training programme 2014–2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

We wish to acknowledge the helpful discussions with G. Leifert, T. Grüning, T. Strauss, M. Weidemann and J. Michael. Our gratitude goes to H. Niemann and F. Pisano for providing the IR data geometry and heat flux mapping. We are also thankful for the support from the CITlab (University of Rostock) and Planet-AI teams.

ORCID iDs

Thomas Sunn Pedersen  <https://orcid.org/0000-0002-9720-1276>

References

- [1] Böckenhoff D. et al 2018 Reconstruction of magnetic configurations in W7-X using artificial neural networks *Nucl. Fusion* **58** 056009
- [2] Hinton G.E. et al 2012 Improving neural networks by preventing co-adaptation of feature detectors 1–18 (arXiv:1207.0580v1)
- [3] He K. et al 2016 Deep residual learning for image recognition *IEEE Conf. on Computer Vision and Pattern Recognition (Las Vegas USA, 27–30 June 2016)* pp 770–8
- [4] Cireşan D. et al 2012 Multi-column deep neural network for traffic sign classification *Neural Netw.* **32** 333–8
- [5] Grieger G. and Milch I. 1993 Das fusionsexperiment Wendelstein 7-X *Phys. J.* **49** 1001–5
- [6] Feng Y. et al 2008 EMC3/EIRENE transport modelling of the island divertor in W7-X *35th European Physical Society Conf. on Plasma Physics. 10th Int. Workshop on Fast Ignition of Fusion Targets (Hersonissos, Crete, 9–13 June 2008)* (http://epsppd.epfl.ch/Hersonissos/pdf/P2_085.pdf)
- [7] Effenberg F. et al 2017 Numerical investigation of plasma edge transport and limiter heat fluxes in Wendelstein 7-X startup plasmas with EMC3-EIRENE *Nucl. Fusion* **57** 36021
- [8] Andreeva T. 2002 Vacuum magnetic configurations of Wendelstein 7-X *Technical Report* Max-Planck-Institut für Plasmaphysik, Greifswald
- [9] LeCun Y. and Bengio Y. 1995 Convolutional networks for images, speech, and time series *The Handbook of Brain Theory and Neural Networks* p 3361 (<http://yann.lecun.com/exdb/publis/pdf/lecun-bengio-95a.pdf>)
- [10] Szegedy C. et al 2015 Going deeper with convolutions *Proc. IEEE conference on computer vision and pattern recognition (Boston, USA, 7–12 June 2015)* pp 1–9
- [11] Deng J. et al 2009 ImageNet: a large-scale hierarchical image database *IEEE Conf. on Computer Vision and Pattern Recognition (Miami USA, 20–25 June 2009)* pp 248–55
- [12] Krizhevsky A., Sutskever I. and Hinton G.E. 2012 ImageNet classification with deep convolutional neural networks *Adv. Neural Inf. Process. Syst.* **25** 1097–105
- [13] Glorot X. and Bengio Y. 2010 Understanding the difficulty of training deep feedforward neural networks *Proc. 13th Int. Conf. on Artificial Intelligence and Statistics (Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010)* pp 249–56 (<http://proceedings.mlr.press/v9/glorot10a.html>)
- [14] Kingma D.P. and Ba J.L. 2015 Adam: a method for stochastic optimization *Int. Conf. on Learning Representations 2015*

- (*San Diego USA, 7–9 May, 2015*) pp 1–15 (<https://iclr.cc/archive/www/2015.html> arXiv:1412.6980)
- [15] Abadi M. *et al* 2016 TensorFlow: a system for large-scale machine learning *OSDI* pp 265–83 (<https://ai.google/research/pubs/pub45381>)
- [16] Renner H. *et al* 2004 Physical aspects and design of the Wendelstein 7-X divertor *Fusion Sci. Technol.* **46** 318–26
- [17] Diccio T.J. and Efron B. 1996 Bootstrap confidence intervals *Stat. Sci.* **11** 189–228
- [18] Goodfellow I.J. *et al* 2014 Generative adversarial nets, *Advances in neural information processing systems (Montreal, Canada, 8–13 December 2014)* pp 2672–80 (<http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>)

A.3. Article III

»Application of improved analysis of convective heat loads on plasma facing components to Wendelstein 7-X«

D. BÖCKENHOFF and M. BLATZHEIM

Nuclear Fusion, Vol. 59.8 (2019), DOI: 10.1088/1741-4326/ab201e

Application of Improved Analysis of Convective Heat Loads on Plasma Facing Components to Wendelstein 7-X

Daniel Böckenhoff¹, Marko Blatzheim^{1, 2} and the W7-X Team¹

¹Max Planck Institute for Plasma Physics, Wendelsteinstraße 1, 17491 Greifswald, Germany

²Institute for Mathematics, University of Rostock, Ulmenstraße 69, 18057 Rostock, Germany

daniel.boeckenhoff@ipp.mpg.de, marko.blatzheim@ipp.mpg.de

Abstract

Within this research we present a method to speed up the simulation of convective heat loads onto the plasma-facing components (PFCs) of the Wendelstein 7-X (W7-X) stellarator by a factor of approximately 20 with the same statistical precision as compared to the previous standard. The geometric models developed for this are also designed to unravel the complex 3D PFCs onto a 2D picture-like input format which gives access to the full set of image analysis tools like for example wavelet analysis or the applicability of convolutional neural network (CNN) architectures.

The significant speedup of heat load calculation allows to simulate a massive data set of heat loads for approximately 3×10^4 magnetic configurations in edge rotational transform-radial axis shift space. To first order, plasma dynamic effects like toroidal current development as well as beta effects mainly influence this space which motivates the simulation in this scope.

A criterion to evaluate the safety of a magnetic configuration with respect to the convective heat load onto the plasma facing components has been developed taking statistical fluctuations of the simulation into account. This criterion, applied to the introduced data set, provides a map relevant for discharge planning and machine safety.

The methods and concepts introduced herein could contribute to the safety evaluation of magnetic confinement devices in general and are not specific for W7-X.

1 Introduction

Wendelstein 7-X (W7-X) is a stellarator type nuclear fusion experiment [1, 2]. Three main goals were defined for the project. The first, demonstrating the feasibility of building a superconducting modular stellarator with the required precision, has been reached, as shown by the many successful experiments conducted already [3–5]. A second objective is to demonstrate the accessibility of plasma parameters close to those of a future fusion power plant, which is ongoing work [6–9]. The last is to prove the possibility of continuous operation [6]. To accomplish this, it is essential to protect the plasma facing components (PFCs) from heat loads exceeding the design specifications. We refer to this case as ‘overload’.

Monte Carlo field line diffusion simulation is the standard tool to calculate the heat load onto the PFCs. It is a heavily used method for discharge planning and evaluation. In order to obtain PFC heat loads with statistical significance, it has so far been necessary to simulate with $\mathcal{O}(10^6)$ single field line diffusion traces, resulting in run times of roughly 10 hours. Due to evolving plasma density, temperature and currents,

the PFC heat load is dynamic [5, 10–12] and multiple simulations have to be performed. This can result in days to weeks of simulation time.

With this work, we introduce new geometrical models of the most important W7-X PFCs, i. e. the divertor, baffle and heat shield. The new models have the same geometric fidelity as the original models but are partitioned in a way that introduces the following beneficial properties:

- a) They reduce the necessary field line diffusion simulation time by a factor of 20 with the same or higher statistical significance. Due to the significant decrease in simulation time, high resolution multi-dimensional magnetic configuration scans become feasible.
- b) The new partitions take into account physics and engineering constraints of the respective PFCs.
- c) They can be mapped to a two dimensional data format well suited for heat load display and analysis. This mapping is not restricted to simulation results but also applicable to experimental heat loads calculated from Infra-Red (IR) temperature data [13]. Notably, the format allows the applica-

tion of convolutional neural network (CNN) architectures to the problem of PFC heat load feature extraction since it is equivalent to a picture-like array, preserving neighborhood relations between the “pixels”. To be very clear: In the scope of the present article no artificial neural networks (NNs) or other machine learning (ML) techniques are applied although they are envisioned. So far heat load visualizations have either been realized by a camera-matrix projection, where the complex curved geometry is distorted [14, 15], or by one-dimensional heat load quantification along a predefined trajectory [15].

Furthermore, a new method is introduced that allows a probabilistic assessment of the overload evaluation for a given magnetic field and a power convected along this field. This method considers each particle trace of a field line diffusion simulation as a Bernoulli process[16].

A high resolution two-dimensional magnetic configuration scan in rotational transform ι and radial shift ΔR with in total 27 181 magnetic configurations and respective simulated heat load patterns was performed. These two properties have a major impact on the plasma edge and thus the PFC heat load pattern [17] and are closely related to the effect the toroidal plasma current and plasma beta have onto the plasma edge topology [5, 11, 18, 19]. The overload evaluation method is applied to this main data set and reveals safe or critical operation regimes for each magnetic configuration in dependence of the convective power onto the PFCs. Such overload maps are highly relevant for the W7-X operation, as they allow to quantify the risk of overload and hint at possible critical states that could be reached by external actuation or occur due to the development of plasma parameters.

In accordance with the line of argumentation presented above, the remainder of this paper is organized as follows. In Section 2 we introduce W7-X and briefly discuss its magnetic field. Section 3 holds the detailed motivation and implementation of the new geometrical models and presents the simulation time speedup introduced thereby. In Section 4 the overload estimation function is introduced and evaluated on the basis of the ι - ΔR scan which is initially described. The paper concludes with a summary and a short discussion of future work which can be pursued on the basis of this work.

2 The Wendelstein 7-X Magnetic Field

W7-X consists of ten half modules. Within each half module, five superconducting non planar coils labeled 1 to 5 are providing the toroidal magnetic flux Ψ as well as the poloidal magnetic flux χ for a low shear

rotational transform profile

$$\iota = \frac{d\chi}{d\Psi}. \quad (1)$$

A finite ι is necessary for the compensation of the effects of the $\vec{\nabla}B$ drift. Two further superconducting planar coils (labeled A and B) per half module allow the superposition of a toroidal magnetic flux component. Behind each divertor unit [20] (see Figure 1(a) and Section 3.1), a so-called (island) control coil allows the local alteration of the magnetic field providing the possibility of changing phase and size of the outer magnetic islands. All control coils can be energized independently. Subsequently, only stellarator symmetric fields are considered, which means that the control coils behind the upper and lower divertor respectively are controlled jointly with their counterparts in all modules. The upper and lower control coils are labeled S1 and S2 (“S” referring to the alternative name “sweep coils”). All coils described above are depicted in Figure 1(b).

We normalize the relative winding currents to the relative winding current of the non planar coil 1, i. e.

$$I_x = \frac{n_{x,w} \cdot I_{x,true}}{I_n}, \quad (2)$$

$\forall x \in \{1, 2, \dots, 5, A, B, S1, S2\}$, where $I_{x,true}$ corresponds to the current applied to the coils and $n_{x,w}$ denotes the number of windings per coil. In the following, the normalization current I_n is equal to $n_{1,w} \cdot I_{1,true}$. With this representation, the contribution to the overall magnetic field component created by coil type x linearly scales with I_x .

Nine so-called reference magnetic configurations are defined to represent particular points in the W7-X configuration space [17]. They can be arranged in a three-dimensional space of rotational transform (ι), radial shift (ΔR) and mirror ratio (m).

Among these nine reference scenarios, five share similar values of m . Changes in the magnetic configuration due to internal plasma currents predominantly take place in two of those dimensions, namely ι and ΔR . The mirror ratio can also be changed by internal currents but the magnitude of the change is much smaller than that of ι and ΔR .

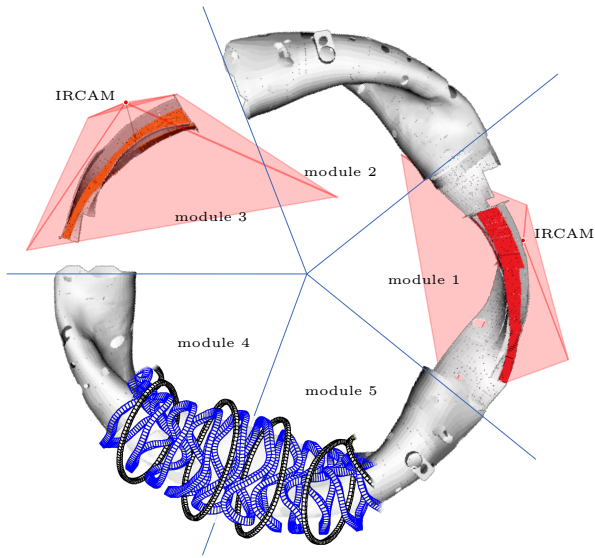
In leading order, the rotational transform as well as the radial shift are linear combinations of the planar coil currents I_A and I_B [11], so for the rest of this paper we shall use normalized versions of ι and ΔR , which are identified by $\tilde{\iota}$ and $\tilde{\Delta R}$ and defined such that

$$\begin{pmatrix} \tilde{\Delta R} \\ \tilde{\iota} \end{pmatrix} = \mathbf{R} \begin{pmatrix} I_A \\ I_B \end{pmatrix}, \quad (3)$$

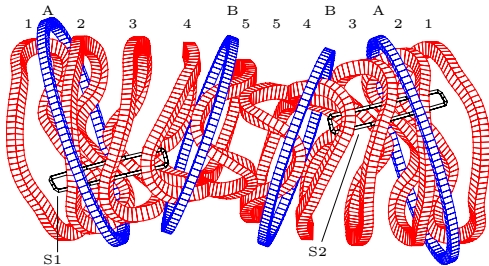
where

$$\mathbf{R} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \quad (4)$$

is the matrix describing a rotation of $\frac{3}{4}\pi$. A high $\tilde{\iota}$ value corresponds to high rotational transform and



(a) Top-down CAD view of the W7-X inner vessel, showing the field of view of the IR camera system for the divertor setup. In the cutaway of module 1, the field of view of the IR camera observing an upper divertor is depicted. In module 3 the lower divertor is depicted together with the field of view of the IR camera observing it. One fifth of the total 50 modular (blue) and 20 planar coils (black) is overlaid in modules 4 and 5.



(b) Coils contained in two adjacent half modules with modular coils 1 – 5 (red), planar coils A, B (blue) and control coils (black). Each module is symmetric with respect to a radial axis through its center. Adapted from [17].

Figure 1: Overview of essential parts of W7-X

positive $\Delta\tilde{R}$ corresponds to an outward radial shift. Second-order effects are slight deviations especially in the mirror ratio. In the $I_A = -I_B$ axis the value of m is altered slightly due to the opposite sign of I_A and I_B , so the inward and outward shifted reference configuration have a small adjustment of the mirror term by means of non-planar coil current adjustments.

3 Fast and Statistically Significant Convective Heat Load Simulation

Field line diffusion (FLD) is frequently used to estimate heat loads onto PFCs of magnetic plasma confinement experiments. It is a Monte-Carlo approach where particles are traced along the magnetic field lines. A diffusive behavior is modeled by random displacements in $[0, \sqrt{12D_\perp\lambda/v}]$ perpendicular to the field lines after a random length x with the distribution $p(x) = \frac{1}{\lambda}e^{-x/\lambda}$, where λ models the mean free path length between two collisions, D_\perp is the orthogonal diffusion coefficient and v is the mean particle velocity [21]. The FLD simulation for a particle terminates, when it collides with any PFC. D_\perp/v is a physical parameter describing the ratio of perpendicular to parallel transport. Since D_\perp actually describes turbulent transport, λ must not be identified with a physical mean free path of particles in the plasma edge. Rather it acts as a numerical parameter which must be small compared to the parallel connection length in the scrape-off layer.

Compared with edge transport codes as e.g. EMC3-Eirene [22], the FLD model has the advantage of much lower calculation time, whereas the transport process, in particular the dominant anormaleous transport perpendicular to the magnetic field, is simulated in much the same way. Effects occurring at high edge densities like plasma detachment [9] are of course not reflected by FLD simulations.

Assuming a convective power P_{conv} reaching the PFCs, the convective heat load q_i onto the PFC segment i , characterized by position and area A_i can be calculated as

$$q_i = \frac{n_i P_{conv}}{\zeta n_{tot} A_i}, \quad (5)$$

with n_{tot} the overall number of hits and the number of hits n_i per PFC segment with area A_i . Symmetries in magnetic fields and PFCs can be utilized to increase the statistics by mapping the hit points onto the smallest symmetric unit. This operation has to be accounted for by introducing ζ in Equation (5).

In the case of the ideally symmetric W7-X magnetic field and for the PFCs used in the simulation, the assumption of a five-fold toroidal periodicity ($q_i(\phi, z) = q_i(\phi + 2\pi/5, z)$) describing the W7-X modules, together with a symmetry inherent to each module ($q_i(\phi_c + \phi, z) = q_i(\phi_c - \phi, -z)$, where ϕ_c is the toroidal angle at the center of each module) is

justified. This allows for an increase in statistics or, equivalently, the decrease in calculation time by the factor of $\zeta = 10$.

On the one hand it is important that the computer-aided design (CAD) components resemble the actually installed PFCs appropriately. This means that the CAD surface has to consist of triangles of different sizes which are usually smaller than the demanded spatial heat load resolution on the other hand. In order to retrieve statistically significant heat load patterns onto these triangles, it is usually necessary to start $\mathcal{O}(10^6)$ separate traces which will require a computation time of above ten hours^I per magnetic configuration.

As long as only few heat load simulations are required, this poses no huge problem. However, FLD is a heavily employed tool. FLD applications can mostly be assigned to one or more of the following categories:

- Exploration problem: Magnetic configuration space exploration or FLD parameter scans in the search for magnetic configurations or parameters with particular properties [23, 24]
- Inverse problem: or discharge reconstruction [15, 25]
- Forward problem: or discharge planning [26]

As motivated before, evolving plasma parameters induce changes in the magnetic configuration and thus dynamic PFC heat loads, which results in the need of multiple simulations for forward problems and even more so for inverse problems. Thus all of the categories above require multiple simulations.

The exploration problem that is going to be introduced in this report requires $\mathcal{O}(10^4)$ different heat load simulations and a rough estimate yields above three years of computation time for one core given the 10^6 separate traces mentioned before^{II}. Thus it would be beneficial to reduce the number of necessary traces per configuration without losing statistical significance.

3.1 PFC Partitioning

The problem of varying and too small A_i , accompanied by the need of large simulation times (see last chapter), can be solved together with the question of proper heat load display and analysis. Unfolding the complex 3D geometry onto a 2D surface would have

^IOptimal performance of a virtual machine cluster with cumulative clock speed of 10 GHz and 16 GB allocated RAM

^{II}Each field line trace is an explicit function of the start position and the magnetic field. Thus both parameters can be parallelized. The parallelization of field lines for a given magnetic field is implemented in the 'FIELDLINES' [27] code. In our simulation space the number of available cores is much smaller than both the number of field line traces per configuration as well as the number of configurations, so that no speedup is gained by exploiting the parallelization of both parameters.

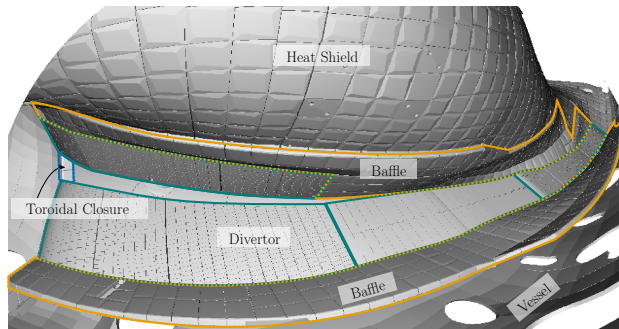


Figure 2: IR camera view of the main loaded W7-X PFCs

benefits for a variety of applications. For example in ML research groups 2D rectangular array data as given in images or spectrograms of audio are the preferred input for so-called deep CNNs [28, 29]. To achieve such an architecture, pixel-like input with proper local neighborhood relations is required. Following the current work we aim to apply such CNN architectures to the PFC heat loads. Besides ML applications, the full image processing and analysis tool box can be applied[30].

By splitting and combining triangles of the CAD PFC geometry, both problems are solved. A python library [31] was created with the aim to handle such mesh operations with a focus on operation speed and the abstractness to handle both, simulated FLD hit points as well as experimental heat load patterns.

Divertor Partitioning For W7-X the energy and particle exhaust is realized by an island divertor concept which proved successful in the predecessor experiment Wendelstein 7-AS (W7-AS) [32, 33]. Ten divertor target plates are arranged corresponding to the five-fold toroidal symmetry and the up-down flip symmetry inherent to the W7-X device. Every divertor unit is composed of four different parts, labeled and depicted in Figure 3(a). The newly meshed model for the divertor provides a two-dimensional map with well-defined neighborhood relations (see bottom part of Figure 3(a)) and significantly larger areas A_i as compared to the varying sizes within the original triangulations. Each part of the new segmentation and a corresponding scalar value are represented by one pixel in the newly defined grid and the pixel value, respectively. Pixels that do not have a corresponding counterpart on the divertor are marked white.

The new mesh is inherited from the original mesh and cut into smaller parts, following the natural tile geometry of the target modules (TM) TM1h-4h, TM7h-9h and TM1v-3v (see grey tiles in right half of the zoom in Figure 3(a)). An artificial tile-like substructure is applied to the TM5h-6h target elements (see the left half of the zoom in Figure 3(a)) which ensures similar sized areas. The partitioning of the divertor following mainly the tile sizes is also physically motivated

since the poloidal heat load pattern scale length is on the same order of magnitude as the typical divertor tile width [34]. It may be that for certain plasma scenarios the poloidal heat load pattern scale length will be below the tile size. If so, our method can be extended to smaller partitions. The computational effort would of course go up accordingly. So far, experimental evidence points more to a confirmation of our assumptions than the need for smaller partitions [35]. Since we want to access the heat load q with respect to the maximal design heat load $q_{d,i}$, we furthermore discriminate on the basis of $q_{d,i}$ [36] which is color coded in Figure 3(a). This means in particular that the divertor edges facing the pumping gap between TM1h-4h and TM1v-3v are subdivided along the so-called end top tile and gap tile structures which correspond to lower $q_{d,i}$ of 5 MW and 2 MW respectively (see the right half of the zoom in Figure 3(a)). The gap tiles are subdivided into two parts parallel to the main target surface (see subdivision of the yellow colored parts in the right half of the zoom in Figure 3(a)). These cuts are implemented in order to account for the large heat load gradient along this side.

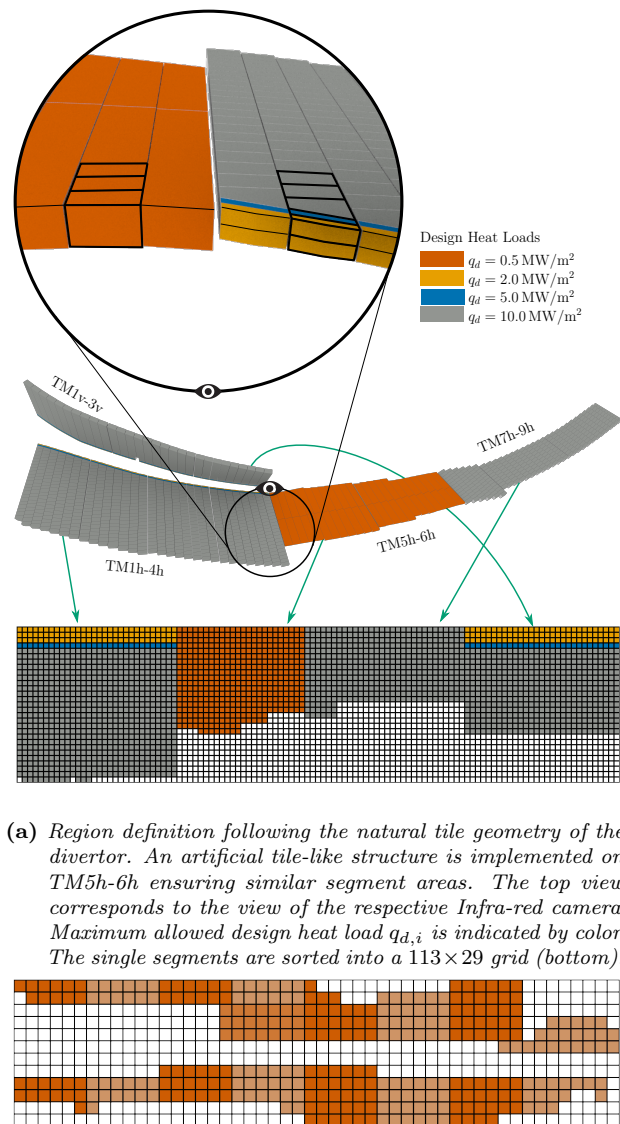
Baffle Partitioning Baffle tiles covering the wall in the vicinity of the divertor are designed to protect the wall from the loads radiated by the divertor plasma. The mapping of the baffle geometry onto a 50×12 grid, splitting the meshed model along the single baffle tiles, is displayed in Figure 3(b) (cf. Figure 2). In order to guide the eye, the single baffle modules are colored alternately. The heat load limit of 0.5 MW m^{-2} originates from the requirement to keep the temperature of the CuCrZr-structure at the bottom of each tile connecting the fine grained graphite with the cooling pipe below a critical temperature [37]. Because of that, the partitioning according to the baffle tile geometry is not only geometrically evident but also motivated from an engineering point of view.

Heat Shield Partitioning The partitioning of the heat shield is motivated and carried out analogous to the baffle partitioning, since the heat shield is constituted from tiles of similar dimension and quality as the baffle tiles (see Figure 2). A corresponding 2D visualization with defined neighborhood relations is not realized yet but easily expandable.

The overload analysis described below is not affected by this complexity reduction. The convective heat load onto the further PFCs is negligible for P_{conv} up to 10 MW.

3.2 Field Line Diffusion Parameters

In the following the parameters used for the field line diffusion simulations are motivated. The perpendicular diffusion coefficient D_{\perp} is adjusted to $1 \text{ m}^2 \text{ s}^{-1}$



(a) Region definition following the natural tile geometry of the divertor. An artificial tile-like structure is implemented on TM5h-6h ensuring similar segment areas. The top view corresponds to the view of the respective Infra-red camera. Maximum allowed design heat load $q_{d,i}$ is indicated by color. The single segments are sorted into a 113×29 grid (bottom).

(b) Region definition due to separation of the single baffle tiles into a 50×12 grid. As indicated by the orange color, the maximum q for the whole baffle is 0.5 MW m^{-2} . The variations in the orange color are highlighting the different baffle modules.

Figure 3: Partitioning the divertor and baffle tiles to a two-dimensional grid with proper neighborhood relations.

which was encountered in the predecessor experiment W7-AS [38]. Evaluation of D_{\perp} for W7-X is currently done using transport calculations in comparison with IR camera measurements and hints to smaller values but of the same magnitude [39].

The average particle velocity at the plasma edge can be approximated as

$$\langle v \rangle = \sqrt{\frac{k_B(T_e + T_i)}{m_i + m_e}} \stackrel{m_i \gg m_e}{\approx} \sqrt{\frac{k_B(T_e + T_i)}{m_i}} \quad (6)$$

$$\stackrel{T_e \approx T_i}{\approx} \sqrt{\frac{2k_B T_e}{m_i}},$$

where the last assumption $T_e \approx T_i$ is only valid at the edge [5]. With values of $T_e = T_i = 100$ eV [5] this results in $v = 1.4 \times 10^5$ m s⁻¹. A diffusive step perpendicular to a particles field line is evaluated after $\lambda = 0.1$ m, which is sufficiently small as compared to typical connection lengths (cf. Section 3).

description	$q_{d,i}$ [MW m ⁻²]
divertor	
high loaded area	10.0000
gap tiles	5.0000
end top tiles	2.0000
TM5h-6h	0.5000
baffle	0.5000
toroidal closure	0.5000
heat shield	0.5000
wall panels	0.2000
vessel	0.0019

Table 1: PFCs used for field line tracing. $q_{d,i}$ from [37]

Besides the divertor, baffle and heat shield the PFCs included in the FLD process are those installed in the final operational phase (OP2) which are depicted in Figure 2. Together with the local design heat load $q_{d,i}$, they are listed in Table 1. All in-vessel components (IVCs) are designed to withstand an electron cyclotron resonance heating (ECRH) stray radiation power density of 0.05 MW m⁻² or are shielded against it [40]. $q_{d,i}$ as well as $A - i$ were chosen in discussion with engineers on the basis of [37] and takes into account failure mechanisms for the component segments represented by our mesh.

3.3 Heat Load Simulation Speedup

The remaining field line diffusion parameter to be chosen is n_{tot} . Obviously, n_{tot} should be sufficiently high in order to achieve a smooth heat load pattern without statistical noise. In other words the calculated heat loads on the segments should no longer show a significant dependence on the choice of n_{tot} . For exactly this reason in the newly partitioned mesh, small segments of the raw mesh are combined with

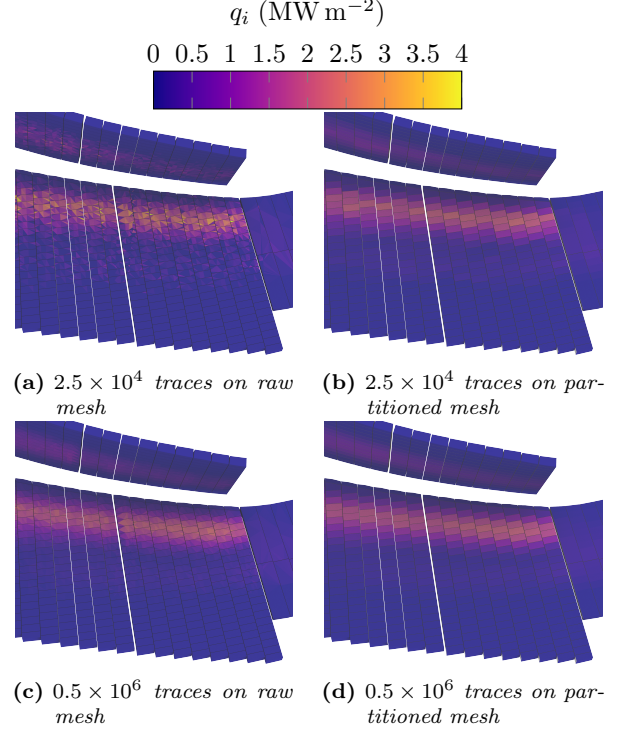


Figure 4: Zoom into horizontal part of divertor heat load for simulations of the standard reference case magnetic configuration. The geometrical model (raw vs. partitioned) as well as the total number of field line diffusion traces n_{tot} changes in between the sub figures.

adjacent segments into the relevant scale size (see Section 3.1).

We perform field line diffusion simulations with ten different n_{tot} and the parameters introduced in the previous section on the original (“raw”) and partitioned mesh. The value of $n_{tot} = 0.5 \times 10^6$ shown in Figure 4(c) is the commonly used value[41] and results in a smooth pattern. In contrast, the raw heat load pattern for $n_{tot} = 2.5 \times 10^4$ has insufficient statistics which leads to a high fluctuation and extreme maximal heat loads (see Figure 4(a)). This is not evident on the partitioned mesh for either n_{tot} (see Figures 4(b) and 4(d)).

To examine the smoothness quantitatively, we examine the fluctuation of the strike-line mean as a function of n_{tot} normalized to $n_{tot} = 10^6$

$$\frac{\langle q_i(n_{tot}) \rangle_{sl}}{\langle q_i(10^6) \rangle_{sl}}. \quad (7)$$

The “strike-line” (sl) is defined by those two percent of the total area with the highest heatload. Figure 5

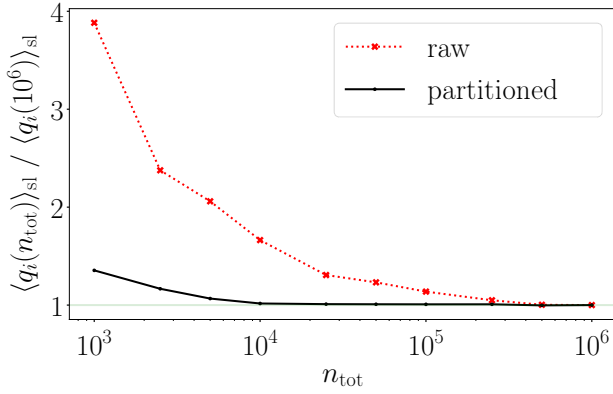


Figure 5: Sensitivity of the average strike-line (sl) heat load with respect to the total number of field line traces n_{tot} for the original (“raw”) or partitioned CAD model (indicated red and black respectively). In order to show convergence the value is normalized to the reference strike-line at $n_{tot} = 10^6$. The data are based on ten independent simulations (varying n_{tot}) with the W7-X standard reference magnetic configuration.

demonstrates the convergence of this metric towards unity. As expected, the raw mesh requires a higher n_{tot} to converge as compared to the newly partitioned mesh. Sufficient convergence can be approximated as $n_{tot} \geq 0.5 \times 10^6$ and $n_{tot} \geq 2.5 \times 10^4$ for the raw and partitioned mesh respectively.

This is also evident in comparing Figures 4(b) and 4(d) where no further improvement is seen with increasing n_{tot} for the partitioned mesh. The quality of Figure 4(c) is achieved in Figure 4(b) with the partitioned mesh requiring a factor of 20 fewer traces. In conclusion a simulation speed up of a factor of 20 can be reported and hereafter all simulations are performed with $n_{tot} = 2.5 \times 10^4$.

3.4 Heat Load Visualization

The segmentation technique described in Section 3.1 is demonstrated for various reference configurations in figures 6 and 7 for the divertor and baffle respectively with assumed 5 MW convective power. Well-known patterns as e. g. the standard reference case with the main heat load on TM1h-4h and TM1v-3v [20] can be identified (see left and right quarters of Figure 6(a) respectively).

In the same configuration, the convective heat load onto the baffle as depicted in Figure 7 appears mostly on a baffle tile which is adjacent to the vertical divertor target (TM1v-3v).

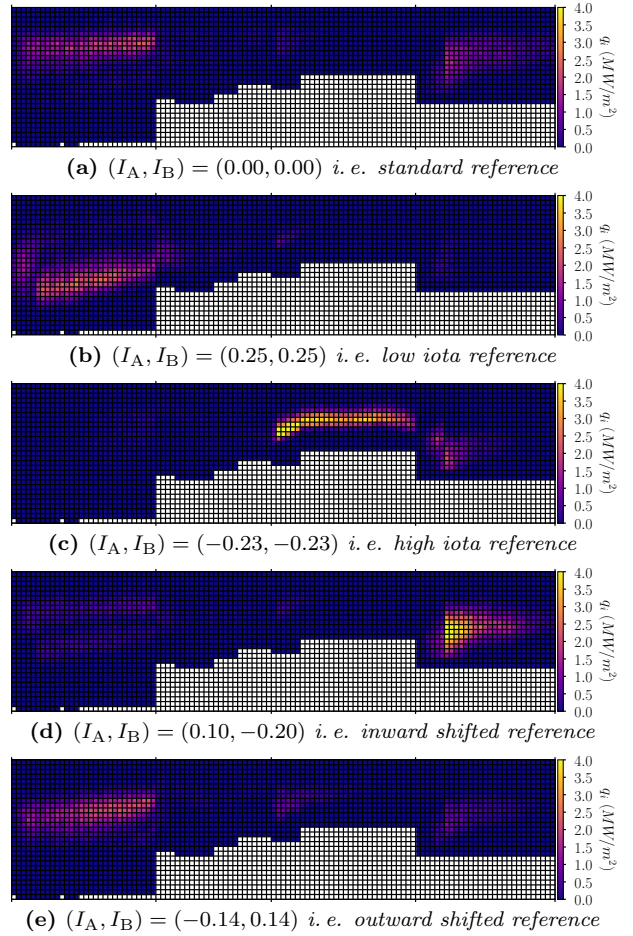


Figure 6: Divertor heat loads visualized in 2D as described in Section 3.1 at $P_{conv} = 5$ MW

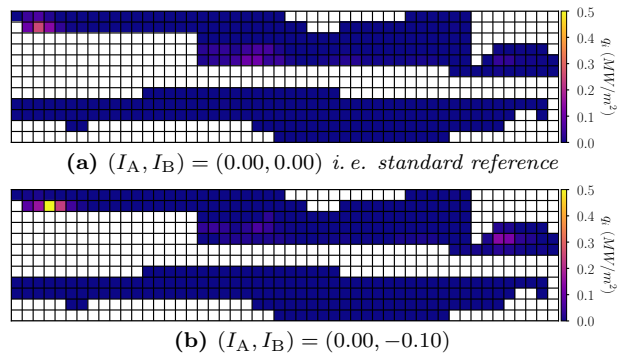


Figure 7: Baffle heat loads visualized in 2D as described in Section 3.1 at $P_{conv} = 5$ MW

4 Overload Evaluation

With the speed up in simulation time (cf. Section 3.3) it becomes feasible to perform heat load simulations for magnetic configurations in a narrow spaced two dimensional parameter space with those two plasma parameters affecting the magnetic edge topology the most. Therefore, as motivated before, it is sensible to simulate on the basis of a grid built in the $\tilde{\tau}$ - $\Delta\tilde{R}$ space (cf. Figure 8) and transformed into the I_A - I_B space according to

$$\begin{pmatrix} I_A \\ I_B \end{pmatrix} = \mathbf{R}^{-1} \begin{pmatrix} \tilde{\tau} \\ \Delta\tilde{R} \end{pmatrix}. \quad (8)$$

The non planar coil currents I_1 to I_5 are all set to 1. The resulting grid is reduced such that $\max(|I_A|, |I_B|) \leq 0.3$, corresponding to the technical limitations of the coil system. Figure 9 depicts the resulting 27 181 grid points in the I_A - I_B plane.

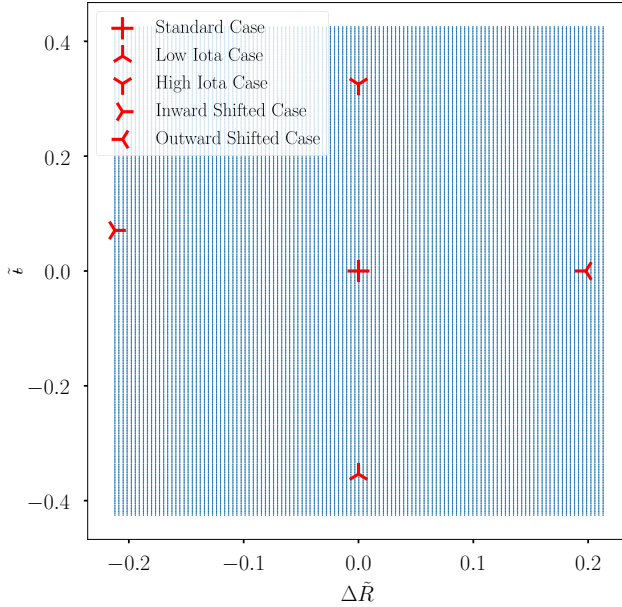


Figure 8: Grid creation with $\Delta\tilde{R}$: 121 evenly spaced points from $\Delta\tilde{R} = -\frac{0.3}{\sqrt{2}}$ to $\Delta\tilde{R} = \frac{0.3}{\sqrt{2}}$ and $\tilde{\tau}$: 301 evenly spaced points from $\tilde{\tau} = -\frac{0.6}{\sqrt{2}}$ to $\tilde{\tau} = \frac{0.6}{\sqrt{2}}$, i. e. 36421 grid points. W7-X reference magnetic configurations are indicated by symbols.

4.1 Overload Definition

Overload occurs, if any q_i exceeds $q_{d,i}$. It is important to consider the noisy character of the field line diffusion simulation and its impact onto the overload, i. e. $q_i - q_{d,i}$. With Equation (5) the critical number of hits per area can be calculated as

$$n_{crit,i} = \lfloor q_{d,i} \frac{\zeta n_{tot} A_i}{P_{conv}} \rfloor, \quad (9)$$

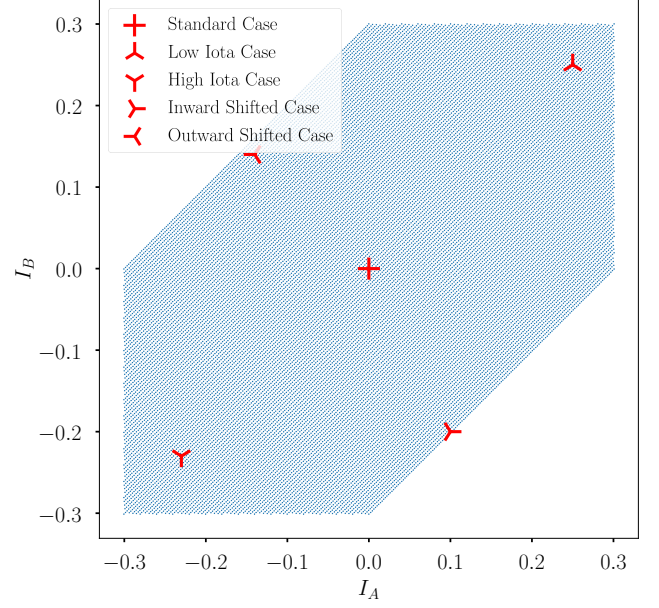


Figure 9: Rotated grid with 27181 points after eliminating points with $\max(|I_A|, |I_B|) > 0.3$

where the $\lfloor \cdot \rfloor$ operator indicates conservative rounding to the next lower integer.

For independent single Monte-Carlo events, where $n_i(I_A, I_B) \ll n_{tot}$, $n_i(I_A, I_B)$ is a good estimate for the rate parameter of a Poisson distribution,

$$X \sim \text{Poisson}_{n_i}(k), \quad (10)$$

where the random variable X describes the number of hits on segment i in the case of a repeated experiment. The cumulative distribution function is given as

$$\Pr_{n_i}(X \leq k) = e^{-n_i} \sum_{j=0}^k \frac{n_i^j}{j!}. \quad (11)$$

We can utilize Equation (11) to calculate the probability of safe operation (i. e. no overload) for each target element as

$$\Pr_{n_i}(X \leq n_{crit,i}) = e^{-n_i} \sum_{j=0}^{n_{crit,i}} \frac{n_i^j}{j!}. \quad (12)$$

The varying parameters in Equation (9) are A_i and $q_{d,i}$.

With Equation (12), we can combine probabilities for arbitrary PFCs assuming independency:

$$\Pr(X \leq n_{crit}; I_A, I_B) = \prod_i \Pr_{n_i}(X \leq n_{crit,i}). \quad (13)$$

This is what we define as the safety estimation function

$$\Theta(I_A, I_B) := \Pr(X \leq n_{crit}; I_A, I_B). \quad (14)$$

4.2 Overload Analysis

In the following, an evaluation of the probability of safe operation given a certain convective power P_{conv} , i. e. the safety estimation function introduced above (Equation (14)) is presented. The choice of P_{conv} has to be seen in the context of the planned heating scenarios for OP2 as well as the actual fraction of radiative losses per discharge. For stage I and II total maximum heating powers from 10 MW^{III} to 24 MW^{IV} are planned depending on the heating scenario [37]. Recent experiments have demonstrated, that a detached plasma can significantly reduce the convective load onto the divertor by a factor of approximately 10 [9]. This study provides four exemplary values of $P_{conv} = 2, 5, 8$ and 10 MW but could be easily extended to other values.

In order to work out the main influences and most critical parts, we shall discuss the different PFCs separately.

4.2.1 Divertor

Figure 10 shows the safety evaluation function as a function of I_A and I_B for four different values of P_{conv} on the basis of the introduced data set. At $P_{conv} = 2$ MW a safe operation is apparent. With growing P_{conv} regions of significant overload i. e. $q > q_{d,i}$ appear and grow.

In the configuration space shown here, the most frequently overloaded part of the divertor is TM5h-6h as shown in Figure 11. It depicts the safety estimation function for each pixel i of the partitioned divertor at $P_{conv} = 5$ MW and three different tuples (I_A, I_B) . Four of the five reference configurations contained in the I_A - I_B space simulated (see Figure 9) can safely be run even at $P_{conv} = 10$ MW (see Figure 10(d)). The low iota configuration however shows first signs of potential overload at $P_{conv} = 5$ MW (see fig. 11(a)) and is significantly overloaded at higher P_{conv} .

4.2.2 Baffle

Figure 12 reveals a similar picture for the safety of the baffle. Four of five reference configurations are safe up to 8 MW inclusive. However, the low iota reference configuration is safe for all P_{conv} whereas the inward shifted case becomes critical in between 5 and 8 MW.

4.2.3 Heat Shield

The evaluation of the heat shield overload (see Figure 13) reveals some critical load in magnetic fields with high ι for a range of outward Shafranov shifts. Furthermore, there is an unsafe region at medium negative ι and inward Shafranov shift. However, for the

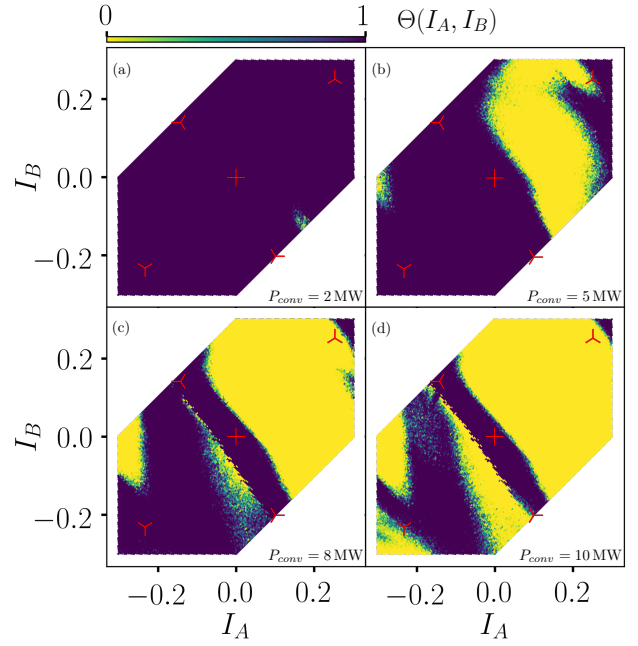


Figure 10: Divertor probability of safe operation for $P_{conv} = 2, 5, 8$ and 10 MW shown by (a) - (d) respectively. The markers indicate five W7-X reference magnetic configurations (see Figure 9).

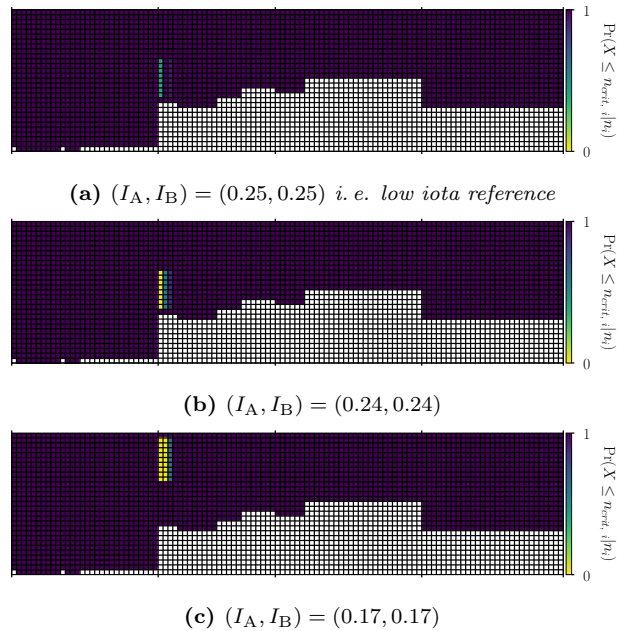


Figure 11: Probability for safe operation for three different magnetic configurations. Depicted is the evaluation of $\Pr_{n_i}(X \leq n_{crit,i})$ according to Equation (12) for each part of the segmented divertor as described in Section 3.1 at $P_{conv} = 5$ MW

^{III}Maximum pure ECRH heating power

^{IV}Maximum neutral beam injection (NBI) + ion cyclotron radiation heating (ICRH) input power for up to 10s in stage II

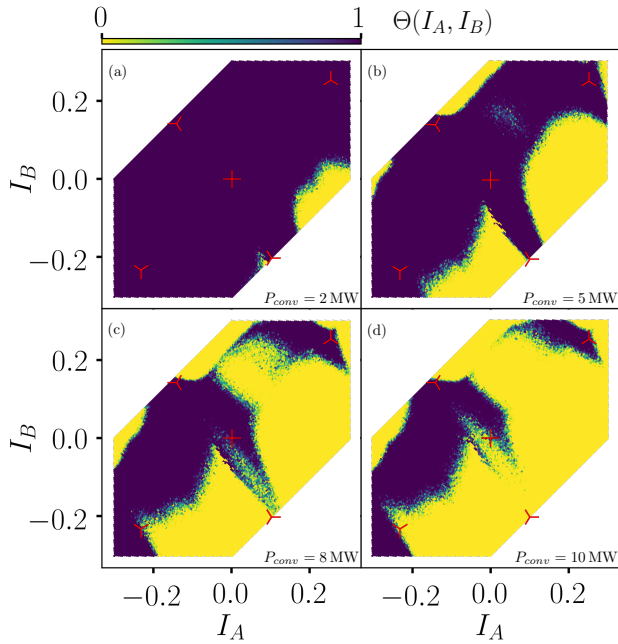


Figure 12: Baffle probability of safe operation for $P_{conv} = 2, 5, 8$ and 10 MW shown by (a) - (d) respectively. The markers indicate five W7-X reference magnetic configurations (see Figure 9).

convective powers and the parameter space considered, this second region is always a subset of the insecure regions identified for the divertor.

The further components that were included in the FLD simulation, i.e. toroidal closure, wall panels and vessel do not show any critical loads even at $P_{conv} = 10$ MW. Multiplying the probabilities in the I_A - I_B maps for all PFCs (see Equation (14)), we get a complete picture of the remaining configurations which do not show significant overload risk of any PFC. As an example, Figure 14(a) shows the result for $P_{conv} = 8$ MW. The high iota, the outward shifted and the standard reference configurations can be run with no overload on the PFCs at this convective power if plasma effects are negligible.

4.3 Sweep Coils as Additional Degree of Freedom

Although W7-X is designed to minimize toroidal currents, their magnitude is rarely zero. Depending on the magnetic configuration and the heating scenario, a toroidal current I_{tor} converging with the bootstrap current of $\mathcal{O}(10$ kA) is changing ι and thus influences the heat load pattern on the PFCs. In a first approximation, the change of the edge magnetic configuration due to I_{tor} corresponds to a simultaneous reduction of I_A and I_B toward negative values [14]. This poses a challenge for the operation in the standard configuration at high convective powers as $P_{conv} = 8$ MW, as can be seen from Figure 14(a), where the orange arrow

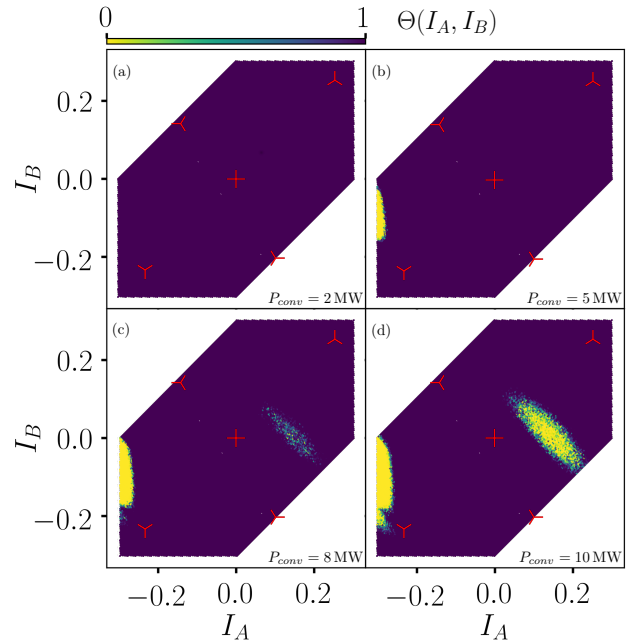


Figure 13: Heat shield probability of safe operation for $P_{conv} = 2, 5, 8$ and 10 MW shown by (a) - (d) respectively. The markers indicate five W7-X reference magnetic configurations (see Figure 9).

indicates the path corresponding to a change of I_{tor} from 0 to 40 kA. It is apparent, that such a discharge would potentially damage the divertor as well as the baffle.

Changing the control coil currents I_{S1} and I_{S2} introduces a third dimension in the configuration space and allows to avoid overload on those PFCs as indicated by the green arrow in Figure 14(b), circumnavigating the critical configuration above the I_A - I_B plane. Such actuators like the control coils can provide additional freedom as demonstrated here.

5 Conclusion and Future Work

A redefinition of the CAD models representing the most important W7-X PFCs has been implemented successfully by merging and splitting faces without loss in geometrical precision. As a result of the model redefinition the required field line diffusion simulation time is reduced by a factor of 20, engineering constraints to the geometry are taken into account and a mapping to unravel the complex 3D PFCs onto 2D rectangular array data can be provided.

The speed-up of the heat load calculation enables simulations based on a fine grid in the rotational transform-radial shift parameter space.

A new criterion to evaluate the overload risk has been developed taking statistical fluctuations of the MC based FLD simulation into account. This criterion,

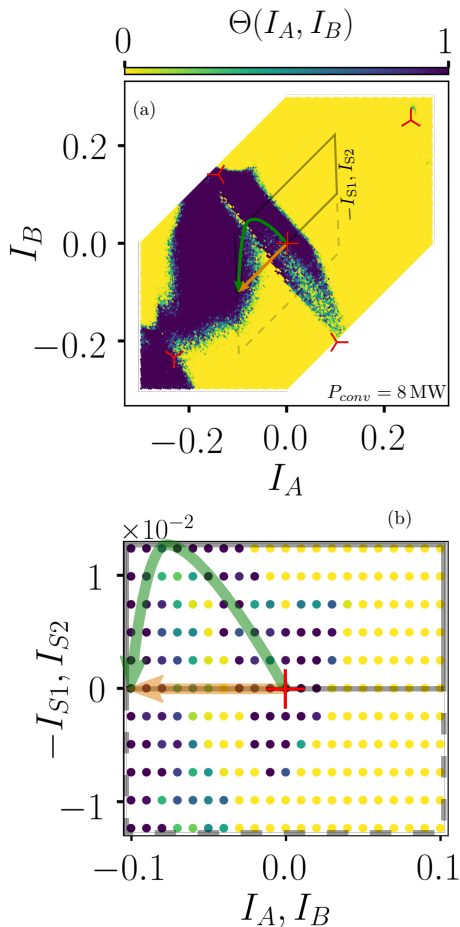


Figure 14: Probability of safe operation for all PFCs and $P_{conv} = 8\text{MW}$. (a): Set of magnetic configurations introduced in the introduction of Section 4. (b): Additional data set with control coil variations, which intersects the main data set in the way indicated by grey lines in Figure 14(a). The red markers indicate W7-X reference magnetic configurations (cf. Figure 9). The orange arrow indicates I_A - I_B mimicked configurations for $I_{tor} \in [0, 40\text{kA}]$ (see text) starting from the standard case at $(I_A, I_B, I_{tor}) = (0, 0, 0)$. This corresponds to a path through the configuration space as it would occur due to the time evolution of the toroidal plasma current during the first several 10s of a discharge [42]. The green arrow demonstrates a possible route to avoid overload by using the control coils S1 and S2 such that $I_{S1} = -I_{S2}$.

applied to the simulated data, provides overload maps that will be relevant for machine operation.

These tools developed to improve the heat load analysis constitute a huge improvement for various tasks beyond the applications presented herein. The underlying concepts are applicable to further Monte Carlo based ray tracing simulations, calculating heat loads onto other PFCs without restrictions to the specific W7-X magnetic field and geometry.

In [25, 43] it was shown, that NNs are capable of reconstructing plasma properties from W7-X limiter heat load data. It was proven, that the performance of those networks applied to experimental data was improved by supplementing the training data set with simulations. The present work provides the methods necessary to investigate these findings with the divertor and further PFCs which is ongoing work. The new visualization mapping is applicable to simulation as well as experimental data and thus provides an interface between the two. The rectangular nature of the new heat load representation allows the implementation of state of the art NN architectures as for example CNNs, and beyond this gives access to all kinds of image manipulation and analysis tools.

As a long term objective we pursue heat load pattern control in W7-X for which this work constitutes a highly relevant building block. We envision reinforcement learning approaches which require reward functions in order to have a quantity to optimize for [44, 45]. The overload maps found in Section 4 could be used as such.

6 Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

We wish to acknowledge the helpful discussions with M. Endler, L. Rudischhauser, J. Fellingner, J. Geiger, R. Labahn and T. Sunn Pedersen.

Bibliography

- [1] H.-S. Bosch et al. “Technical challenges in the construction of the steady-state stellarator Wendelstein 7-X”. In: *Nuclear Fusion* 53.12 (2013), p. 126001. ISSN: 0029-5515.
- [2] J. Nührenberg et al. “Overview on Wendelstein 7-X Theory”. In: *Fusion Technology* 27.Suppl. S (1995), pp. 71–78.

- [3] T. Sunn Pedersen et al. “Confirmation of the topology of the Wendelstein 7-X magnetic field to better than 1:100000”. In: *Nature Communications* 7 (2016), p. 13493. ISSN: 2041-1723.
- [4] T. Klinger et al. “Performance and properties of the first plasmas of Wendelstein 7-X”. In: *Plasma Physics and Controlled Fusion* 59.1 (2017), p. 014018. ISSN: 0741-3335.
- [5] A. Dinklage et al. “Magnetic configuration effects on the Wendelstein 7-X stellarator”. In: *Nature Physics* (2018), pp. 1–6. ISSN: 17452481.
- [6] C. Beidler et al. “Physics and Engineering Design for Wendelstein VII-X”. In: *Fusion Science and Technology* 17.1 (1990), pp. 148–168.
- [7] J. Nührenberg and R. Zille. “Stable stellarators with medium β and aspect ratio”. In: *Physics Letters A* 114.3 (1986), pp. 129–132. ISSN: 03759601.
- [8] Thomas Sunn Pedersen et al. “Key results from the first plasma operation phase and outlook for future performance in Wendelstein 7-X”. In: *Physics of Plasmas* 24.5 (2017), p. 055503. ISSN: 10897674.
- [9] Thomas Sunn Pedersen et al. “First results from divertor operation in Wendelstein 7-X”. In: *Plasma Physics and Controlled Fusion* 61.1 (2019), p. 014035. ISSN: 0741-3335.
- [10] V. Erckmann, U. Gasparino, and H. Maaßberg. “Current drive and bootstrap current in stellarators”. In: *Plasma Physics and Controlled Fusion* 34.13 (1992), pp. 1917–1923. ISSN: 0741-3335.
- [11] P. I. Strand and W. A. Houlberg. “Magnetic flux evolution in highly shaped plasmas”. In: *Physics of Plasmas* 8.6 (2001), pp. 2782–2792. ISSN: 1070664X.
- [12] Y. Feng et al. “Impact of Island Geometry on Island Divertor Performance”. In: *Koch, R.; Lebedev, S.: 30th EPS Conference on Plasma Physics and Controlled Fusion, European Physical Society (2003)* (2003).
- [13] M. Jakubowski et al. “Development of infrared and visible endoscope as the safety diagnostic for steady-state operation of Wendelstein 7-X”. In: *QIRT* 100 (2014).
- [14] H. Hölbe et al. “Access to edge scenarios for testing a scraper element in early operation phases of Wendelstein 7-X”. In: *Nuclear Fusion* 56.2 (2016), p. 026015. ISSN: 0029-5515.
- [15] Y. Gao et al. “Approaches for quantitative study of divertor heat loads on W7-X”. In: *14th Quantitative InfraRed Thermography Conference*. 2018, pp. 6–7.
- [16] Achim Klenke. *Probability theory : a comprehensive course*. 2014. ISBN: 1848000472.
- [17] T. Andreeva. *Vacuum magnetic configurations of Wendelstein 7-X*. Tech. rep. IPP-III/270. Greifswald: Max-Planck-Institut für Plasmaphysik, 2002.
- [18] J. Geiger et al. “Physics in the magnetic configuration space of W7-X”. In: *Plasma Physics and Controlled Fusion* 57.1 (2015), p. 014004. ISSN: 0741-3335.
- [19] Per Helander. “Theory of plasma confinement in non-axisymmetric magnetic fields”. In: *Reports on Progress in Physics* 77.8 (2014), p. 087001. ISSN: 0034-4885.
- [20] Hermann Renner et al. “Physical Aspects And Design of the Wendelstein 7-X Divertor”. In: *Fusion Science and Technology* 46.2 (2004), pp. 318–326.
- [21] S.A. Bozhenkov et al. “Service oriented architecture for scientific analysis at W7-X. An example of a field line tracer”. In: *Fusion Engineering and Design* 88.11 (2013), pp. 2997–3006. ISSN: 09203796.
- [22] Y. Feng et al. “Recent Improvements in the EMC3-Eirene Code”. In: *Contributions to Plasma Physics* 54.4-6 (2014), pp. 426–431. ISSN: 08631042.
- [23] P. Sinha et al. “Numerical studies of scrape-off layer connection length in Wendelstein7-X”. In: *Nuclear Fusion* 58.1 (2018), p. 016027. ISSN: 0029-5515.
- [24] F. Effenberg et al. “Numerical investigation of plasma edge transport and limiter heat fluxes in Wendelstein 7-X startup plasmas with EMC3-EIRENE”. In: *Nuclear Fusion* 57.3 (2017), p. 036021.
- [25] D. Böckenhoff et al. “Reconstruction of Magnetic Configurations in W7-X using Artificial Neural Networks”. In: *Nuclear Fusion* 58.5 (2018), p. 056009.
- [26] J. D. Lore et al. “Modeling and preparation for experimental testing of heat fluxes on W7-X divertor scraper elements”. In: *IEEE Transactions on Plasma Science* 46.5 (2018), pp. 1387–1392. ISSN: 00933813.
- [27] Samuel A. Lazerson et al. “First measurements of error fields on W7-X using flux surface mapping”. In: *Nuclear Fusion* 56.10 (2016), p. 106005. ISSN: 0029-5515.
- [28] Yann Lecun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *Nature* 521.7553 (2015), pp. 436–444. ISSN: 14764687. arXiv: arXiv:1312.6184v5.

- [29] Jürgen Schmidhuber. “Deep learning in neural networks: an overview.” In: *Neural networks : the official journal of the International Neural Network Society* 61 (2015), pp. 85–117. ISSN: 1879-2782. arXiv: 1404.7828.
- [30] J. R. Parker. *Algorithms for Image Processing and Computer Vision*. 2nd ed. Wiley, 2011, p. 481. ISBN: 978-0-470-64385-3.
- [31] Daniel Böckenhoff and Marko Blatzheim. *tfIELDS: numpy + sympy implementation of tensor fields with attached coordinate systems*. <https://pypi.org/project/tfields/>. 2018.
- [32] P. Grigull et al. “First island divertor experiments on the W7-AS stellarator”. In: *Plasma Physics and Controlled Fusion* 43.12A (2001), A175–A193. ISSN: 0741-3335.
- [33] K. McCormick et al. “Island divertor experiments on the Wendelstein 7-AS stellarator”. In: *Journal of Nuclear Materials*. Vol. 313-316. SUPPL. 2003, pp. 1131–1140. ISBN: 0022-3115.
- [34] M Endler et al. “Managing leading edges during assembly of the Wendelstein 7-X divertor”. In: *Plasma Physics and Controlled Fusion* 61.2 (2019), p. 025004. ISSN: 0741-3335.
- [35] H. Niemann et al. “Large wetted areas of divertor power loads at Wendelstein 7-X”. In: *Physical Review Letters* (2019), To be published.
- [36] J. Boscary et al. “Design improvement of the target elements of Wendelstein 7-X divertor”. In: *Fusion Engineering and Design* 87.7-8 (2012), pp. 1453–1456. ISSN: 09203796.
- [37] R. Brakel et al. *Specification of design loads for in-vessel components of w7-x*. Tech. rep. 1-AC-S0005.1. Max-Planck-Institut für Plasmaphysik, 2017, p. 57.
- [38] F. Wagner et al. “Overview on W7-AS Results with Relevance for Wendelstein 7-X and the Low-Shear Stellarator Line”. In: *Fusion Energy Proc. 17th Int. Conf. Yokohama*: IAEA Vienna, CD-Rom file and https://inis.iaea.org/collection/NCLCollectionStore/_Public/31/011/31011276.pdf, 1998, pp. 115–128.
- [39] F. Effenberg et al. “Investigation of 3D effects on heat fluxes in performance-optimized island divertor configurations at Wendelstein 7-X”. In: *Nuclear Materials and Energy* 18 (2019), pp. 262–267. ISSN: 23521791.
- [40] R. Brakel et al. *Technische Richtlinie Einsatz von Materialien und Verwendung von magnetisierbaren Werkstoffen an W7-X*. Tech. rep. 1-NBT-T0010.12. 2017.
- [41] M. Endler et al. *Private Communication*. 2019.
- [42] Arnold Lumsdaine et al. “Design and analysis of the W7-X divertor scraper element”. In: *Fusion Engineering and Design* 88.9–10 (2013), pp. 1773–1777. ISSN: 09203796.
- [43] Marko Blatzheim et al. “Neural network performance enhancement for limited nuclear fusion experiment observations supported by simulations”. In: *Nuclear Fusion* 59.1 (2019), p. 016012. ISSN: 0029-5515.
- [44] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. Vol. 3. 9. 1998, p. 360. ISBN: 0262193981. arXiv: 1603.02199.
- [45] Timothy P. Lillicrap et al. “Continuous Control with Deep Reinforcement Learning”. In: *International Conference on Learning Representations 2016* (2016). arXiv: 1509.02971.

A.4. Article IV

»Neural network regression approaches to reconstruct properties of magnetic configuration from Wendelstein 7-X modeled heat load patterns«

M. BLATZHEIM and D. BÖCKENHOFF

Nuclear Fusion, Vol. 59.12 (2019), DOI: 10.1088/1741-4326/ab4123

Neural Network Regression Approaches to Reconstruct Properties of Magnetic Configuration from Wendelstein 7-X Modeled Heat Load Patterns

Marko Blatzheim^{1, 2}, Daniel Böckenhoff¹ and the Wendelstein 7-X Team¹

¹Max Planck Institute for Plasma Physics, Wendelsteinstraße 1, 17491 Greifswald, Germany

²Institute for Mathematics, University of Rostock, Ulmenstraße 69, 18057 Rostock, Germany

marko.blatzheim@ipp.mpg.de, daniel.boeckenhoff@ipp.mpg.de

Abstract

Convective heat loads onto the plasma facing components of magnetic confinement devices contain information about edge magnetic field properties which are not yet fully exploited. Machine learning approaches are a promising technique to automatically extract information about such properties from heat load images. In this study, we present the successful reconstruction of proxies for two independent, important edge magnetic field properties given simulated heat load images on the Wendelstein 7-X divertor target plates.

Six different artificial neural network architectures from shallow and simple feed-forward fully-connected neural network to deep Inception ResNets with 24 223 to 804 804 free parameters are investigated. The relative reconstruction error is between 1 and 2% with calculation times on the order of milliseconds. A competing benchmark method without machine learning reaches slightly smaller errors but exceeds the calculation time by three orders of magnitude. The experiments demonstrate that machine learning is also a powerful tool in this particular field of nuclear fusion research and deep convolutional neural networks are identified as favorable algorithms for the stated problem. The findings of this paper build a basis for future real time discharge optimization and control by means of machine learning methods.

1 Introduction

In recent time artificial neural networks (NNs) have become a popular tool in data-driven scientific practice. Most NN tasks are classification problems, from the classic machine learning set of the MNIST database [1], to handwritten text recognition [2] and machine translation [3]. The NNs described in this study, on the other hand, solve regression problems. The general form [4] uses feed-forward fully-connected neural networks (FF-FCs), but more recent research on regression also considers convolutional neural networks (CNNs) when dealing with images [5]. Pioneering approaches with fusion relevance investigated the reconstruction of plasma equilibrium parameters [6] and plasma shape parameters [7], both considering tokamaks. More recent studies dealt

with neoclassical transport [8], real-time disruption predictions in tokamaks [9] or modelling turbulent transport [10]. While these NNs are all FF-FCs, the positive impact of CNNs on magnetic topology reconstruction performance is analyzed in [11]. With a new approach to unravel the complex 3D geometry of the plasma facing components (PFCs) monitored by the Infra-Red (IR) cameras onto a 2D picture-like input format [12], it is possible to investigate the performance of various NN architectures that have been optimized for 2D-input and proven good.

Wendelstein 7-X (W7-X) is the most advanced stellarator type nuclear fusion experiment optimized for nested magnetic surfaces with small islands, a small shafranov shift ($\beta \approx 5\%$), small neoclassical transport, magneto hydro dynamics (MHD) stability, fast α particle confinement and

small bootstrap current I_{bs} with feasible modular coils [13–16]. W7-X has already accomplished several major goals [17, 18], demonstrating the fusion relevance and high performance of the W7-X magnetic confinement concept.

One major advantage over the Tokamak confinement concept [19] is the steady state capability inherent to stellarators. With respect to technical restraints though, it is yet to be demonstrated that W7-X is capable of arbitrarily long operations at fusion-relevant plasma parameters (temperatures of the order of 100×10^6 K at densities of $1 \times 10^{20} / \text{m}^3$). Approaching such parameters, heat loads to the PFCs occur on the order of the critical heat loads of the components. Usually, experiments are planned such that actuator behaviour is determined in advance. However, going towards a minute and beyond in the discharge duration, a real time control system becomes crucial. Although, as mentioned before, W7-X is optimized for small I_{bs} as well as small Shafranov shift the impact of both quantities on machine performance and safety must not be neglected. I_{bs} is the origin of a toroidal current I_{tor} , establishing itself on time scales of minutes. The Shafranov shift, caused by the Pfirsch-Schlüter current, follows the time evolution of plasma pressure [20]. The rotational transform ι and the radial plasma shift ΔR are linearly dependent on I_{tor} and the plasma pressure respectively. Since the PFC heat load pattern is heavily influenced by ι and ΔR , significant changes of the pattern can occur during a discharge which are potentially critical [12, 21]. Thus, it is crucial to monitor and control the heat load onto the PFCs.

The heat load can be calculated from temperature data recorded by infra-red cameras [22]. Information about equilibrium properties such as ι and shift contained within the heat load pattern by means of proxy measures can be extracted by NNs in real-time. The relations between the magnetic edge properties, and the heat load pattern on the PFCs are highly complex. The heat load patterns move in a non-trivial, non-linear way with the underlying parameters and may appear or disappear on certain divertor parts depending on the respective magnetic configuration (see the later introduced Figure 5). This is another reason to apply NNs.

¹Either way, further direct diagnostics would be additional inputs for such systems. For example I_{tor} is directly measured by continuous Rogowski coils [25–27].

Feature reconstruction with NNs is relevant for the following reasons: The main purpose of the application of NNs to evaluate the infra-red diagnostic data is to identify heat load patterns in real-time for safety reasons. The applied feature reconstruction can, however, also connect heat load distributions to plasma parameters also known from other diagnostics, e.g. strike-line movements to toroidal current measurements with Rogowski-coils. Thus the applied feature reconstruction NN may be able to combine safety and control issues and help to steer divertor and discharge performance by safely controlling the heat load pattern on the divertor. The approach would be to relate the plasma parameters, which change the heat deposition patterns, i.e. the edge magnetic field topology (mainly the toroidal current and the plasma beta), to the heat deposition patterns in terms of proxies. These proxies could then be provided to classical control systems or even to a trained reinforcement learning [23, 24] control system ¹. In order to determine how successful a NN-based reinforcement agent can handle W7-X divertor heat load images, a first step would be to show the capability of NNs to reconstruct the aforementioned proxies for edge magnetic field properties. At this point we aim at testing and identifying which NN-setup is most suitable to our task in terms of its capabilities and its performance. We are therefore not using MHD-equilibrium fields but use the simpler approach of using certain variations of the vacuum fields allowed by the coil system of W7-X which can be somehow related to the previously mentioned plasma parameters. As said, the details of the relation is of no concern in the context of the present work.

In preceding work [11, 28], it has been demonstrated that the one-dimensional reconstruction of a proxy $\tilde{\iota}$ for the rotational transform from heat load patterns is feasible. These experimental data and according simulations are heat load patterns on the graphite limiters which were the main PFC during the first operational phase. Since 2017 ten island divertors (see Figure 1) are installed, spatially separating the core plasma from the plasma edge and thus making W7-X suitable for high performance discharges [29–31]. For geometri-

cal reasons², the limiter heat load patterns are much less diverse as compared to the heat load distribution on the divertor.

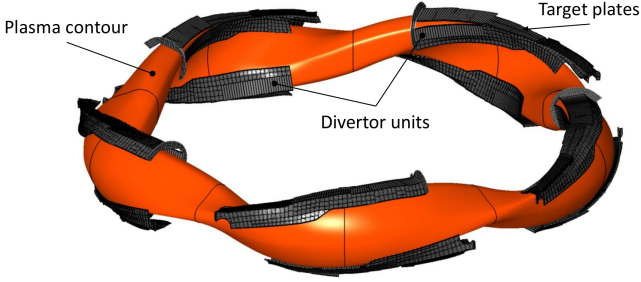


Figure 1: Ten divertor units surrounding the last closed flux surface (orange)

A synthetic data set based on a scan in the proxies $\Delta\tilde{R}$ and $\tilde{\iota}$ of the “plasma shift” and “rotational transform” respectively was introduced in [12]. In Section 2, we will give a brief overview of the training data, followed by an introduction of various NN architectures for the reconstruction of $\tilde{\iota}$ and $\Delta\tilde{R}$. The results in Section 3 focus on the NN regression capability, differences for $\Delta\tilde{R}$ and $\tilde{\iota}$ reconstruction and dependencies on the region within the $\tilde{\iota}$ - $\Delta\tilde{R}$ field for regression.

2 Methods

The recent W7-X operational phases provided experimental time for multiple proposals. Only a few of them focused on exploring the configuration space in detail, so for initial NN analysis with the divertor geometry, simulations are necessary to provide sufficient training data. Simulations allow to focus on the two most dynamic and thus most important parameters, the radial shift $\Delta\tilde{R}$ and the edge rotational transform $\tilde{\iota}$. As mentioned before, a Shafranov shift ΔR_S , induced by the Pfirsch-Schlüter current, corresponds to a radial shift and an establishing toroidal current I_{tor} influences the rotational transform. The two parameters $\Delta\tilde{R}$ and $\tilde{\iota}$ are proxies for the real radial shift and edge rotational transform assuming a superposition principle, i. e. ΔR and ι are linear

combinations

$$\begin{aligned}\Delta R &= c_1 \cdot \Delta\tilde{R}(I_A, I_B) + c_2 \cdot \Delta R_S \\ \iota &= c_3 \cdot \tilde{\iota}(I_A, I_B) + c_4 \cdot \iota_{I_{\text{tor}}}(I_{\text{tor}}) + \iota_0\end{aligned}$$

with constants c_1 , c_2 , c_3 , c_4 and ι_0 . In the case of ι this superposition principle is supported by theory [32]. The dimensionless proxies are based on a linear transformation with

$$\begin{pmatrix} \Delta\tilde{R} \\ \tilde{\iota} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \begin{pmatrix} I_A \\ I_B \end{pmatrix},$$

resulting in a rotation of $\frac{3}{4}\pi$ with the W7-X planar coil currents I_A and I_B which are defined relative to the coil current of non-planar coil 1, $I_{1,\text{true}}$ [12]. The simulation assumes a simplified vacuum configuration and therefore plasma pressure and I_{tor} are both zero, which reduces the simulation time by orders of magnitude. The planar coil currents are adjustable parameters of the field line diffusion (FLD) simulation [33] which is employed to compute the heat loads onto the W7-X PFCs by random diffusion processes which start sufficiently close to the plasma core. The non-planar coil currents $I_{1.5}$ are all set to $I_{1,\text{true}}$. With this normalization the actual value of $I_{1,\text{true}}$ is only a parameter for the magnetic field strength and does not influence the outcome of the simulation. The simulation is performed with a diffusion coefficient $D_{\perp} = 1 \text{ m}^2 \text{ s}^{-1}$ based on the predecessor experiment Wendelstein 7-AS (W7-AS) [34], an estimated particle velocity of $v = 1.4 \times 10^5 \text{ m s}^{-1}$ and a mean free path length of $\lambda = 0.1 \text{ m}$. A detailed discussion of the choice of these parameters in the FLD simulation can be found in [12]. The local heat loads are formed at the intersection of magnetic fluxes and PFCs as shown in Figure 3(a).

The $\tilde{\iota}$ value range is twice as high as the $\Delta\tilde{R}$ value range. Because of the different scaling, normalized radial shift and rotational transform are introduced as

$$\Delta\hat{R} = \frac{\Delta\tilde{R}}{\sigma_{\Delta\tilde{R}}}, \quad \hat{\iota} = \frac{\tilde{\iota}}{\sigma_{\tilde{\iota}}}$$

² The heat load pattern shape and relative intensity on the PFCs are in first order the result of the interaction between magnetic field geometry and PFC shape. The poloidally elongated limiter was installed on the inboard side of the center of each module and adjusted to limit the plasma radially. Each of the ten divertors on the other hand spans a toroidal angle of 47° with an approximately 20 times larger surface area as compared to the limiter. The divertors are designed to intersect the large magnetic resonances called magnetic islands which are highly sensitive to changes in the edge rotational transform. Thus, the limiter heat load pattern is a much weaker function of varying plasma properties and magnetic configurations as compared to the divertor heat load pattern (compare limiter heat load pattern variations [28] with Figure 5).

where $\sigma_{\Delta\hat{R}}$ and $\sigma_{\hat{t}}$ represent the standard deviation of $\Delta\hat{R}$ and \hat{t} respectively. The data points grid is hexagon-shaped as shown in Figure 2 with spacing in the $\Delta\hat{R}$ -axis between 1 and 121 points and in \hat{t} -axis between 151 and 301 points, so the data resolution in \hat{t} is higher.

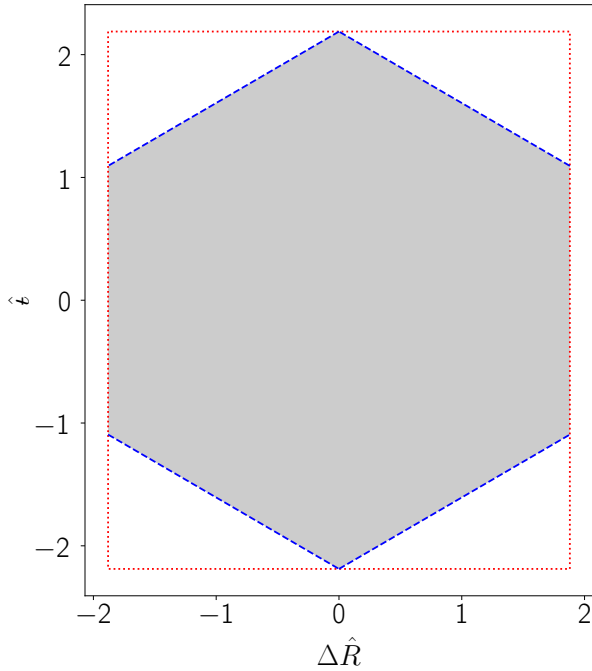


Figure 2: The area enclosed by the red dotted lines represents 36421 grid points, with 121 evenly spaced in $\Delta\hat{R}$ and 301 evenly spaced in \hat{t} . The simulated data set is reduced to 27181 points within the grey hexagon. This reflects technical limitations of the coil system, indicated by the blue dashed lines.

A distinction is made between two different pre-processings defined as extracted characteristics input (ECI) and pixel-based input (PBI). The result of a single FLD simulation is a set of points in 3D-space representing the impact location of particles onto terminating PFCs.

ECI The simulation generates a 3D hit point cloud of size 25 000 describing the heat load. In order to reduce the input dimensionality for ECI these points are clustered using DBSCAN [35] as shown in Figure 3³. Three characteristics are extracted from each of the n_c clusters, namely the number of points, the center of mass and eigenvectors of the covariance matrix. This leads to the number of extracted characteristics within each

cluster $m_c = 13$. In this case the input vector $u_{\text{ECI}} \in \mathbb{R}^{n'}$ with $n' = 195$ is initialized with zeros and filled with $n_c \cdot m_c$ entries, where n_c is the variable number of clusters found by DBSCAN and realized values are between 2 and 15.

PBI An approach to partition the complicated 3D shaped divertor heat flux onto a 2D array like structure, addressing the physics as well as the engineering constraints of the divertor presented in [12], can be utilized to generate the PBI with $u_{\text{PBI}} \in \mathbb{R}^{m \times n}$, where $m = 113$ and $n = 29$. The geometric mapping as well as an exemplary PBI for a high iota reference magnetic configuration is depicted in Figure 4 with the heat load given in arbitrary units. In comparison with Figure 3 differences in the heat load patterns seem to appear. However, these discrepancies are artifacts of the misleading impact points representation in Figure 3 as it highly overrates very low point densities. Four exemplary PBI images at different plasma configurations are shown in Figure 5.

An advantage of ECI is, that features are already separated by the clustering and characteristic extraction. However, there are two important drawbacks. Each array entry within one ECI is of a different quantity (scalar intensity, vector direction), limiting the choice of NNs to FF-FCs. More importantly, an equivalent representation based on experimental data is not possible, so the performance of a trained NN on real data is not measurable and training sets cannot be mixed as shown in [11]. Although this study only considers simulated heat loads, it is of interest, whether the actual experimentally observed heat loads are roughly similar such that the results of this paper (concerning the optimized input and architecture) can be applied to experimental data. A comparison of the simulation with the experimentally found heat loads in the PBI representation is shown in Figure 6. Most importantly, the main heat load patterns are similar and the significant heat load is present on target modules 1h-4h. Deviations from the idealized simulated heat load pattern can be seen for example in the occurrence of increased heat load on target module 7h in the experiment as compared to the simulation. Furthermore the simulation assumes stellarator symmetry and error fields are not considered. However, experimental observations

³The two cluster parameters are chosen as $\text{eps} = 0.065 \text{ m}$ and $\text{min samples} = 40$

show, that the divertor heat loads are asymmetric (the deviation in the total power reaching upper and lower divertor in module 2 is approximately a factor 2). It is not in the scope of this work, to go into further detail about the differences. Such considerations are covered in much detail in other works [36–38].

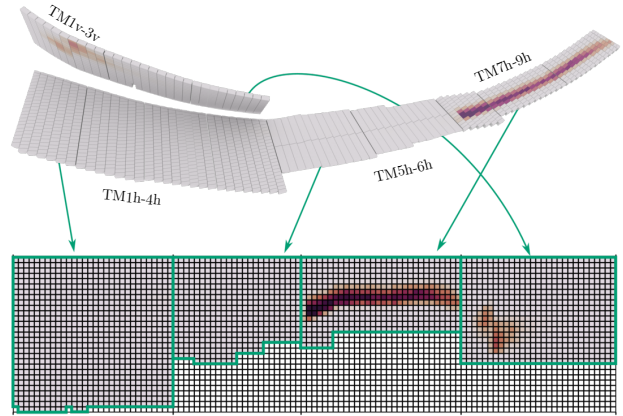
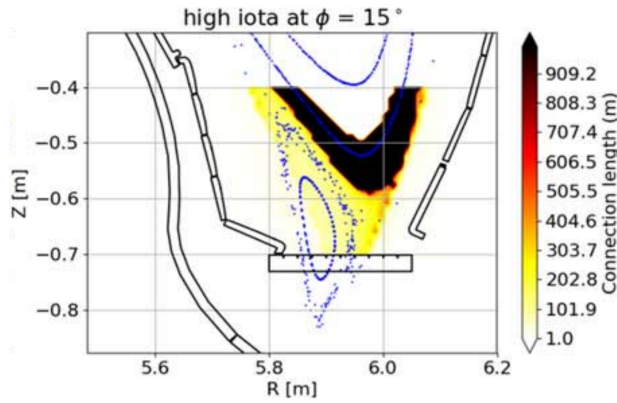
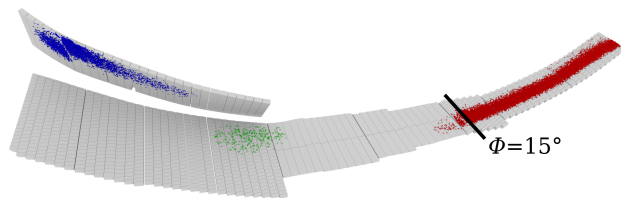


Figure 4: Heat load pattern in arbitrary units retrieved from stellarator symmetric field line diffusion simulation in the so called high iota reference magnetic configuration on the divertor meshed model (top) and the mapping to the 2D array (bottom) which represents the PBI.



(a) Poincare plot at $\Phi = 15^\circ$ (indicated in (b)). Adapted from [18]



(b) Impact points of a stellarator-symmetric field line diffusion simulation

Figure 3: Convective heat exhaust in the high iota reference configuration. The plasma populating the edge island, carrying heat towards the intersecting divertor is shown exemplarily in (a) for a cross section at $\Phi = 15^\circ$ which is a cut through divertor target TM7h-9h (compare Figure 4). In (b) the divertor with the impact points of a stellarator-symmetric field line diffusion simulation in the same reference configuration is shown. In blue, red and green, three clusters resulting from a DBSCAN clustering are color-coded. Based on these clusters statistical properties are taken as an input for the ECI.

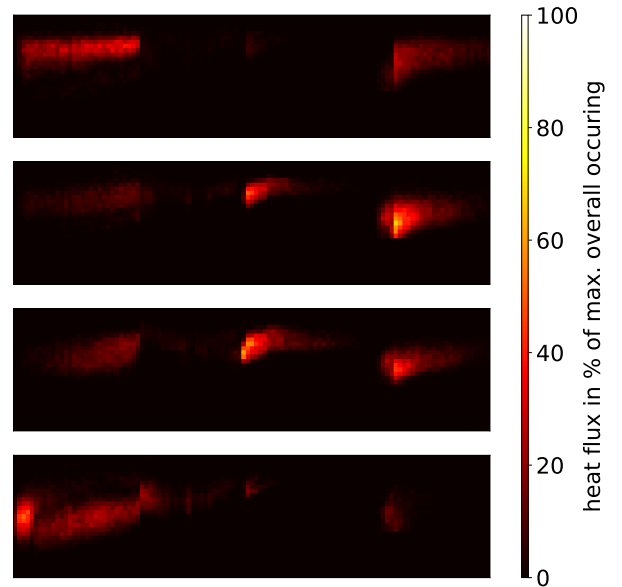


Figure 5: An example of four PBI images representing a pair of $(\Delta\hat{R}, \hat{\epsilon})$ with $(0.0, 0.0)$, $(-1.0, 1.0)$, $(1.0, 1.0)$ and $(0.0, -1.0)$ from top to bottom respectively. To compare the heat loads, a percentage with respect to the highest heat load within all images is given.

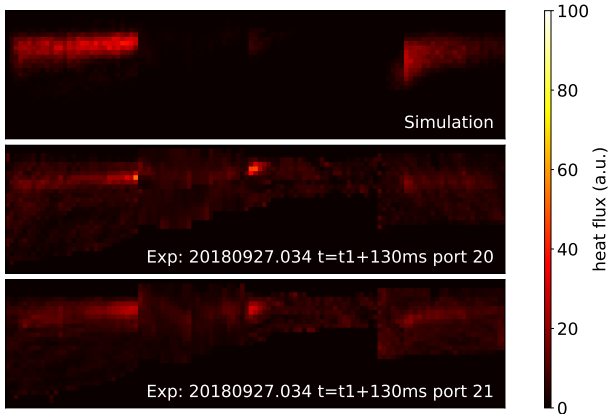


Figure 6: Comparison of three PBI images for standard reference case heat loads. From top to bottom: simulated standard case ($\Delta\hat{R} = 0$ and $\hat{\epsilon} = 0$), experimental measurements of the lower and upper divertor of module 2 for the same frame ($t = t_1 + 130$ ms after trigger 1) of the same experiment (Program 20180927.034). It should be noted that the measured total power reaching the lower divertor is approximately twice the power as compared to the upper divertor. The experimental images show heat load on the high iota tail (center of the picture) whereas the simulation does not.

A brief introduction to the used NN architectures is given in the following, and a detailed layer-for-layer description for the respective architectures can be found in Appendix A. The FF-FC for both ECI and PBI consists of 3 layers following the input layer with 80, 100 and 2 nodes respectively. The performance with the PBI is analyzed not only on FF-FCs but on CNNs, deep convolutional neural networks (DCNNs), deep inception neural networks (DINNs) and Inception-ResNets (IRNNs). An increased depth of 8 additional layers and batch normalization [39] are the key differences between CNNs and DCNNs. The DINN starts with three convolutional layers followed by five inception blocks [40] to calculate the regression. Based on [41] and [42], the IRNN architecture is expanding the DINN by additional twelve residual-inception blocks. These two architectures are adaptations of *Inception-ResNet-v2* [42], with changed reductions in the image width and height induced by the rectangular input shape of 113×29 , reduced repetitions of the residual blocks and a regression loss function. Batch normalization is applied to all convolutional layers within

DINN and IRNN architectures. To solve the regression task all NNs end with a fully-connected layer without an activation function. Unless otherwise stated, the activation functions are Rectified Linear Units (ReLU).

All input data is pre-processed such that the mean is 0 and the standard deviation is 1, which in our experience is necessary for convergence. In case of PBI the sections not corresponding to divertor parts are set to 0 afterwards (i. e. grey pixels in Figure 4). The desired NN outputs (targets) are pairs of $\Delta\hat{R}$ and $\hat{\epsilon}$. The data set is shuffled and separated into five parts of almost equal size. By taking four parts as the training set and splitting the remaining part to validation and test set, it is possible to create five cross-validation sets, resulting in a training set size of 21745, validation set size of 2718 and test set size of 2718. The NNs are updated by the adam optimizer [43] with learning rates and batch sizes given in Table 1. Those two hyperparameters are chosen based on few trainings with learning rates of 0.001, 0.0005 and 0.0001 and batch sizes of 25 and 100 on cross-validation set 1. Batches are sampled randomly from the training set and every second training epoch the PBI training data is augmented by a Gaussian filter with a random Gaussian kernel standard deviation between 0 and 1 to increase the generalization of the NNs. The loss function is the mean squared error (mse) defined as

$$\text{mse} = \frac{1}{N} \sum_k^N (y_k - t_k)^2,$$

with reconstructions y_k , targets t_k and set size N . The error estimation is visualized by the root-mean-square error (rmse) defined as

$$\text{rmse} = \sqrt{\text{mse}}.$$

A scalar reconstruction performance measure $\text{rmse}_{\text{total}}$ is defined based on the reconstructed values for $\hat{\epsilon}$ and $\Delta\hat{R}$.

The total rmse is defined as

$$\text{rmse}_{\text{total}}(\Delta\hat{R}, \hat{\epsilon}) = 0.5 \left(\text{rmse}(\Delta\hat{R}) + \text{rmse}(\hat{\epsilon}) \right).$$

The rmse on the validation set suits as an early-stopping criteria if for 1000 consecutive NN weight updates the rmse is not improving. For each cross-validation distribution the NN training process is performed 32 times with different weight initializations. The software core is based on the TensorFlow library [44]. It runs on a workstation

with two Intel Xeon CPU E5 – 2650 and four NVIDIA GeForce GTX 1080 Ti. Each training is performed on a single GPU.

NN	Trainable Parameters	Learning Rate	Batch Size
FF – FC _{ECI}	23 983	0.0005	25
FF – FC _{PBI}	270 543	0.0010	25
CNN	63 671	0.0005	25
DCNN	89 073	0.0001	100
DINN	313 071	0.0001	100
IRNN	804 804	0.0001	100

Table 1: Comparing the number of trainable parameters as well as the chosen learning rate and batch size for different NNs

It is not possible to determine if human-level reconstruction ability is exceeded because an experimental setting to estimate the error of such a reconstruction can hardly be realized. The random noise of the Monte-Carlo simulation alone makes an error-free reconstruction impossible. To evaluate the NNs, the rmse for two additional benchmarks is defined that serve as examples of non-NN performance. The first one calculates the structural similarity (SSIM) [45] between each PBI of the test set with each of the training set and is referred to as SSIM rmse. The estimated $\Delta\hat{R}$ and $\hat{\epsilon}$ for the test set are those of the training set with the highest SSIM. The neighbor-rmse is the second benchmark which is the spacing distance in $\Delta\hat{R}$ and $\hat{\epsilon}$ direction. Note, that the SSIM method requires a high calculation time that depends on the size of the training set and the grid spacing is not considering any input data knowledge. However, the information that is actually asked for is required, so this method is not applicable in practice and only introduced for comparison purposes.

3 Results

The stated reconstruction problem is far more complex compared to the reconstruction of a single value in the W7-X limiter configuration. Due to non-linear effects between $\Delta\hat{R}$ and $\hat{\epsilon}$ the number of points in an equidistant grid of configurations increases exponentially with the number of considered parameters. There are NNs that converge to functions which solve the regression problem appropriately among all presented architectures.

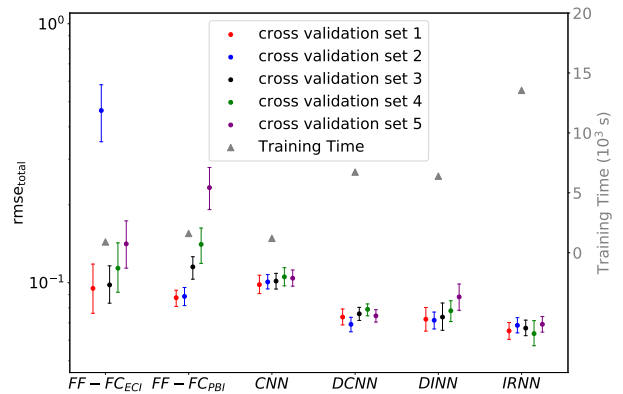


Figure 7: Mean rmse with error bars and training times for each cross-validation set and NN type

An overview of all trained NN results is given in Figure 7. It shows the mean $rmse_{total}$ with respect to all trainings and the 95% confidence interval of the mean calculated with bootstrapping [46] separately for the cross validation sets and architectures.

It can be seen that some FF-FCs with ECI perform well. However, the performance is not consistent across the five cross validation sets and especially cross validation set 3 shows poor performance. Although in [11] a similar concept was outperformed by CNNs on PBI, some of the trained FF-FCs with ECI behave surprisingly well. The FF-FC architectures with PBI show a similar inconsistent behavior. Again, the mean NN performance strongly varies when applied to the different cross validation sets. The FF-FC architectures are thus not recommendable, neither for ECI nor for PBI.

The mean performances of those NN architectures with convolutional layers on PBI are stable with respect to the cross-validation. Note that this stability hints to a sufficient generalization property, but because of the complex deep architectures used, it cannot be guaranteed. The three-layer CNN shows decent performance but it is clearly outperformed by deeper architectures. The first deeper structure we trained after the CNN is the IRNN because of its good performance on other image analysis problems [42]. The gain from CNN to IRNN is impressive but the training time as well as the calculation time increase approximately by a factor of 10, which is still real-time capable. By investigating architectures with

NN	rmse($\Delta\hat{R}$) (10^{-2})	rmse($\hat{\epsilon}$) (10^{-2})	t_{train} (10^3 s)	$t_{\text{calculation}}$ (ms)
FF – FC _{ECI}	7.3	4.6	0.90	0.32
FF – FC _{PBI}	9.5	15.8	1.61	1.06
CNN	10.2	7.1	1.20	1.74
DCNN	7.5	4.9	6.73	7.86
DINN	7.9	5.4	6.39	9.46
IRNN	6.7	5.3	13.55	18.39
SSIM	7.2	2.4	–	8489.27
Neighbor	3.1	1.5	–	–

Table 2: Comparing the different NNs best rmse

depths in between the previously mentioned ones, such as DCNN and DINN, we pursue a trade-off between reconstruction performance and run-time. As can be seen in Figure 7, the NN performance is almost similar to the much deeper IRNN whereas the training time is halved. In Table 2 we compare the NNs, SSIM rmse and neighbor rmse($\hat{\epsilon}$) and rmse($\Delta\hat{R}$) based on the worst performing cross-validation set and within those for the best converged NN. Additionally, the table lists the training times for each NN architecture and the necessary time to calculate the estimate for one input image. Note that these NNs are not optimized with respect to run time. When comparing the NN results with both benchmarks on the one hand it can be noticed that $\hat{\epsilon}$ can be reconstructed a little better (rmse difference of $2.9 \cdot 10^{-2}$) by the SSIM approach and clearly better by taking the next neighbor (rmse difference of $3.8 \cdot 10^{-2}$). On the other hand for $\Delta\hat{R}$ the NNs perform equally good or slightly better compared to the SSIM benchmark (rmse difference of $0.5 \cdot 10^{-2}$) but the neighbor rmse is still smaller (rmse difference of $3.6 \cdot 10^{-2}$). Since the neighbor benchmark does not depend on the input data but only on the grid spacing, it is not a competitive method but it shows that the errors of NNs are of the same order of magnitude. Note that the SSIM value between neighbors is not necessary the largest. These results show that there is still potential for better NN performance. The computation time for SSIM however is not sufficient for real-time applications while up to 18ms for NNs are acceptable.

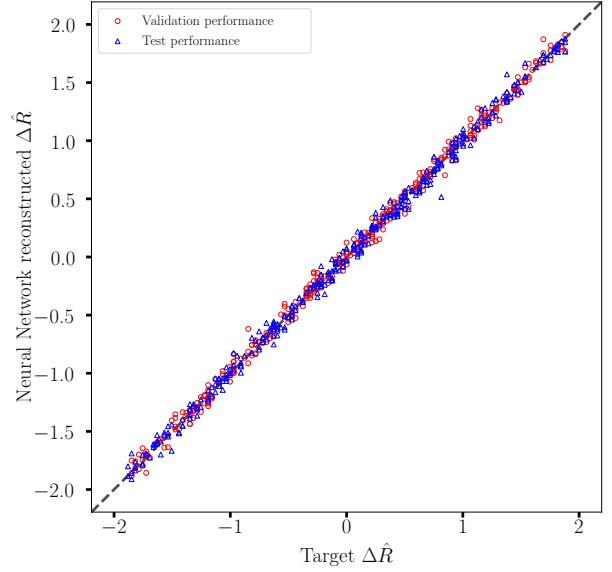


Figure 8: $\Delta\hat{R}$ reconstruction with best IRNN and a rmse of 0.056

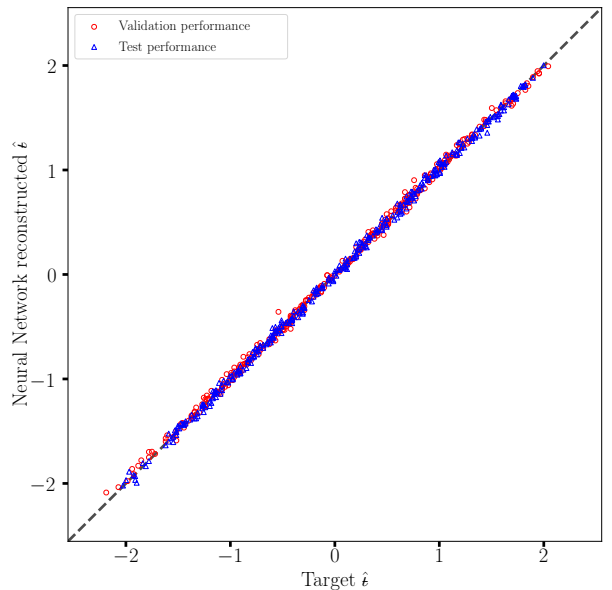


Figure 9: $\hat{\epsilon}$ reconstruction with best IRNN and a rmse of 0.034

The relative error in the $\Delta\hat{R}$ reconstruction is larger compared to that of the $\hat{\epsilon}$ reconstruction. This can be seen in Table 2 but also when visualizing a single NN result in more detail. In Figure 8 and Figure 9 the grey dashed lines mark the identity and thus perfect reconstruction. The points represent the estimated value for each target value and the distance to the grey dashed line resembles the error. In this case, as an example the performance of the best IRNN with a total rmse of 0.045 is shown. When comparing the $\Delta\hat{R}$ and $\hat{\epsilon}$ reconstruction (Figures 8 and 9 respectively), it can be seen that $\hat{\epsilon}$ is estimated better than $\Delta\hat{R}$. However, the NN calculation is appropriately following the identity for both values, indicating that a connection between heat flux image and the respective property has been found. The relative error with respect to the target range is 1.4% for the radial shift and 0.9% for the rotational transform. The reconstruction appears to be symmetric, so the targets are not systematically over- or underestimated.

In the following, the uniformity of the reconstruction error in the $\Delta\hat{R}$ - $\hat{\epsilon}$ space is investigated. When comparing the best-performing NN of each architecture, there are two different types of error distributions, which are represented by the CNN and IRNN results in Figure 10. The four sub-figures depict the difference in the NN performance in a detailed 2D scatter, where the absolute residuals of $\hat{\epsilon}$ and $\Delta\hat{R}$ separately are color-coded. The first type, seen in Figures 10(a) and 10(b), shows clear clusters of higher residuals, especially along the boundary of available data. This phenomenon occurs for $\Delta\hat{R}$ and $\hat{\epsilon}$ reconstruction, however, the cluster of higher residuals are not located at the same region for those two values. This is also the case for PBI FF-FCs. The deeper structures DCNN, DINN and IRNN show a homogeneous error distribution except for a few outliers but no specific structure can be identified. In Figures 10(c) and 10(d) this is shown for the example of IRNN. Notably these residuals do not increase when moving towards the boundary. The ECI FF-FCs show error clusters for $\Delta\hat{R}$ and a more homogeneous error distribution for $\hat{\epsilon}$.

4 Conclusion

The work presented in this article demonstrates the capability of NNs to find a connection be-

tween simulated W7-X divertor heat flux images and proxies for edge magnetic field properties. Former experiments showed similar promising results for less complex limiter heat flux images and a single plasma parameter. The choice of the suiting NN architecture is a trade-off between performance and calculation times. Deeper and more branched NN architectures are associated with decreased regression errors. When stable convergence in combination with small regression errors is the main focus, a customized IRNN can be recommended. FF-FCs are not advisable due to their unpredictable performance on various cross validation sets. A benchmark approach without using machine learning achieves slightly smaller errors but fails to be real-time capable.

There are several possible future investigations arising from the results of this first step. Since the SSIM benchmark example is not outperformed by a machine learning application, further architecture and hyperparameter adjustments may reduce the reconstruction errors. With NNs found feasible to reconstruct proxies for edge magnetic field properties from simulated heat load patterns, it would be the most interesting and consequent pursue to apply the NN to experimental data, which is ongoing work. The PBI allows effortless interchange of simulation and experimental data [12]. This opens the possibility to support experimental data with simulations with the aim to improve the performance of the proxy reconstruction as compared to the proxy reconstruction on NNs trained on experimental data only. Such an approach is inspired by the results in [11] and can be improved by techniques such as transfer learning and generative adversarial networks to reduce necessary experimental data in the training set as much as possible.

More advanced open problems are the consideration of time-dependent effects and using the safety criterion developed in [12] combined with NN architectures from this study for a deep Q-value estimator [24] in reinforcement learning approaches to further approach the long-term objective of real-time heat load control and performance optimization. Besides reinforcement learning, common control theory approaches are candidates for a control loop which would directly use inputs such as the features which have been herein proven to be reconstructable from synthetic heat load images in addition to more direct diag-

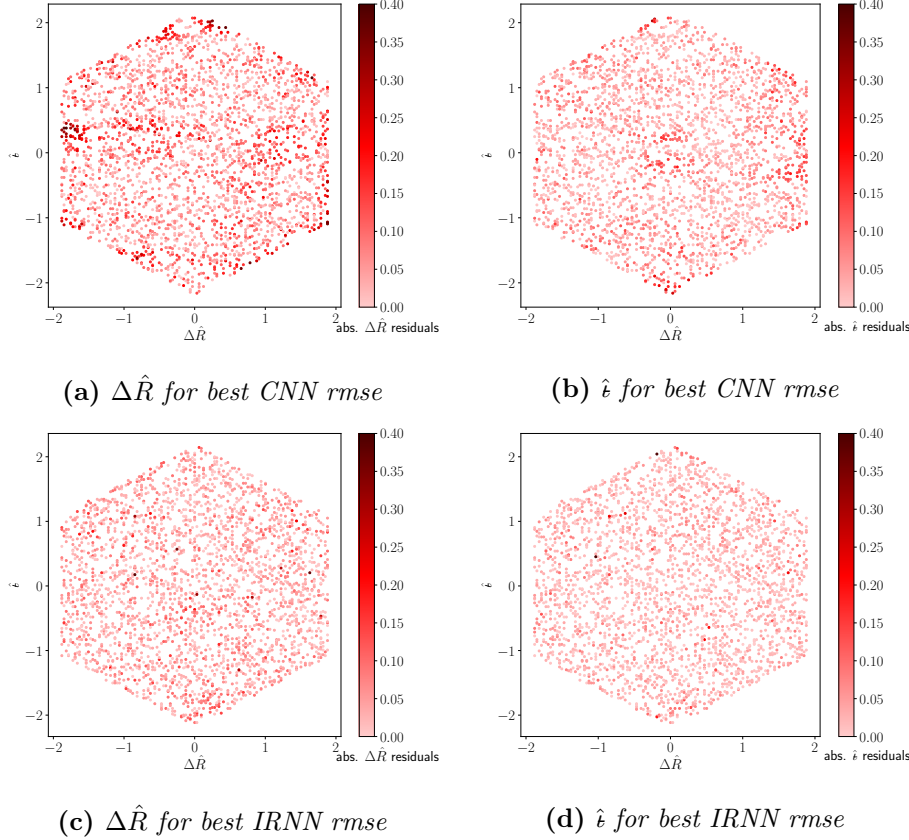


Figure 10: Absolute residuals color-coded depending on iota and shift position

nostic inputs such as Rogowski coils and further. Beyond the scope of heat load control, the possibility to reconstruct proxies for the rotational transform, radial axis shift or further edge magnetic field properties is a desirable goal in itself since it can provide important status information to the operators of W7-X in real-time.

5 Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training program 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

We wish to acknowledge the helpful discussions with J. Geiger, T. Sunn Pedersen, R. Labahn, G. Leifert, T. Grüning, T. Strauss, M. Weidemann and J. Michael. We are also thankful for the support from the CITlab (University of Rostock) and Planet-AI teams.

Bibliography

- [1] Yann LeCun. *The MNIST database of handwritten digits*. 1998.
- [2] Gundram Leifert et al. “Cells in Multidimensional Recurrent Neural Networks”. In: *The Journal of Machine Learning Research* 17.1 (2016), pp. 3313–3349. arXiv: arXiv:1412.2620v2.
- [3] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. “Neural Machine Translation by Jointly Learning to Align and Translate”. In: *Proceedings of 3rd International Conference on Learning Representation*. San Diego, USA, 2015, pp. 1–15. arXiv: arXiv:1409.0473v7.
- [4] Donald F Specht. “A general regression neural network”. In: *IEEE transactions on neural networks* 2:February 1991 (1991), pp. 568–576. DOI: 10.1109/72.97934.
- [5] Philipp Fischer, Alexey Dosovitskiy, and Thomas Brox. “Image Orientation Estimation with Convolutional Networks”. In: *German Conference on Pattern Recognition*.

- Aachen, Germany: Springer, 2015, pp. 368–378.
- [6] J B Lister and H Schnurrenberger. “Fast non-linear extraction of plasma equilibrium parameters using a neural network mapping”. In: *Nuclear Fusion* 31.7 (1991), pp. 1291–1300. DOI: 10.1088/0029-5515/31/7/005.
- [7] E Coccoresse, C Morabito, and R Martone. “Identification of noncircular plasma equilibria using a neural network approach”. In: *Nuclear Fusion* 34.10 (1994), pp. 1349–1363. DOI: 10.1088/0029-5515/34/10/i05.
- [8] Arimitsu Wakasa, Sadayoshi Murakami, and Shun-ichi Oikawa. “Study of Neoclassical Transport in LHD Plasmas by Applying the DCOM/NNW Neoclassical Transport Database”. In: *Plasma and Fusion Research* 3 (2008), S1030. DOI: 10.1585/pfr.3.S1030.
- [9] B. Cannas et al. “An adaptive real-time disruption predictor for ASDEX upgrade”. In: *Nuclear Fusion* 50.7 (2010), p. 075004. DOI: 10.1088/0029-5515/50/7/075004.
- [10] J. Citrin et al. “Real-time capable first principle based modelling of tokamak turbulent transport”. In: *Nuclear Fusion* 55.9 (2015), p. 092001. DOI: 10.1088/0029-5515/55/9/092001. arXiv: 1502.07402.
- [11] Marko Blatzheim et al. “Neural network performance enhancement for limited nuclear fusion experiment observations supported by simulations”. In: *Nuclear Fusion* 59.1 (2019), p. 016012. DOI: 10.1088/1741-4326/aaefaf.
- [12] Daniel Böckenhoff and Marko Blatzheim. “Application of improved analysis of convective heat loads on plasma facing components to Wendelstein 7-X”. In: *Nuclear Fusion* 59.8 (2019), p. 086031. DOI: 10.1088/1741-4326/ab201e.
- [13] J. Nührenberg and R. Zille. “Stable stellarators with medium β and aspect ratio”. In: *Physics Letters A* 114.3 (1986), pp. 129–132. DOI: 10.1016/0375-9601(86)90539-6.
- [14] Craig Beidler et al. “Physics and Engineering Design for Wendelstein VII-X”. In: *Fusion Technology* 17.1 (1990), pp. 148–168. DOI: 10.13182/FST90-A29178.
- [15] J. Nührenberg et al. “Overview on Wendelstein 7-X Theory”. In: *Fusion Technology* 27.3T (1995), pp. 71–78. DOI: 10.13182/FST95-A11947048.
- [16] H.-S. Bosch et al. “Technical challenges in the construction of the steady-state stellarator Wendelstein 7-X”. In: *Nuclear Fusion* 53.12 (2013), p. 126001. DOI: 10.1088/0029-5515/53/12/126001.
- [17] Thomas Sunn Pedersen et al. “Key results from the first plasma operation phase and outlook for future performance in Wendelstein 7-X”. In: *Physics of Plasmas* 24.5 (2017), p. 055503. DOI: 10.1063/1.4983629.
- [18] Thomas Sunn Pedersen et al. “First results from divertor operation in Wendelstein 7-X”. In: *Plasma Physics and Controlled Fusion* 61.1 (2019), p. 014035. DOI: 10.1088/1361-6587/aaec25.
- [19] LA. Artsimovich. “Tokamak Devices”. In: *Nuclear Fusion* 12.2 (1972), p. 215. DOI: 10.1088/0029-5515/12/2/012.
- [20] Vitaly Dmitrijevitsch Shafranov. “Equilibrium of a toroidal plasma in a magnetic field”. In: *Journal of Nuclear Energy. Part C, Plasma Physics, Accelerators, Thermonuclear Research* 5.4 (1963), pp. 251–258. DOI: 10.1088/0368-3281/5/4/307.
- [21] Arnold Lumsdaine et al. “Design and analysis of the W7-X divertor scraper element”. In: *Fusion Engineering and Design* 88.9–10 (2013), pp. 1773–1777. DOI: 10.1016/j.fusengdes.2013.05.075.
- [22] M.W. Jakubowski et al. “Development of infrared and visible endoscope as the safety diagnostic for steady-state operation of Wendelstein 7-X”. In: *Proceedings of the 2014 International Conference on Quantitative InfraRed Thermography*. Vol. 100. Bordeaux, France: QIRT Council, 2014. DOI: 10.21611/qirt.2014.100.
- [23] P. Read Montague. *Reinforcement Learning: An Introduction, by Sutton, R.S. and Barto, A.G.* 2nd ed. Vol. 1. MIT Press, 2017, p. 360. DOI: 10.1016/S1364-6613(99)01331-5. arXiv: 1603.02199.

- [24] Volodymyr Mnih et al. “Human-level control through deep reinforcement learning”. In: *Nature* 518.7540 (2015), pp. 529–533. DOI: 10.1038/nature14236.
- [25] K. Rahbarnia et al. “Diamagnetic energy measurement during the first operational phase at the Wendelstein 7-X stellarator”. In: *Nuclear Fusion* 58.9 (2018), p. 096010. DOI: 10.1088/1741-4326/aacab0.
- [26] T. Andreeva et al. “Equilibrium evaluation for Wendelstein 7-X experiment programs in the first divertor phase”. In: *Fusion Engineering and Design* 146.October 2018 (2019), pp. 299–302. DOI: 10.1016/j.fusengdes.2018.12.050.
- [27] Andreas Werner et al. “W7-X magnetic diagnostics: Rogowski coil performance for very long pulses”. In: *Review of Scientific Instruments* 79.10 (2008), 10F122. DOI: 10.1063/1.2957933.
- [28] Daniel Böckenhoff et al. “Reconstruction of Magnetic Configurations in W7-X using Artificial Neural Networks”. In: *Nuclear Fusion* 58.5 (2018), p. 056009. DOI: 10.1088/1741-4326/aab22d.
- [29] Y Feng et al. “Impact of Island Geometry on Island Divertor Performance”. In: *30th EPS Conference on Plasma Physics and Controlled Fusion*. Ed. by M. Pick and P. Helfenstein. St. Petersburg: European Physical Society, 2003.
- [30] Hermann Renner et al. “Physical Aspects And Design of the Wendelstein 7-X Divertor”. In: *Fusion Science and Technology* 46.2 (2004), pp. 318–326. DOI: 10.13182/FST04-A570.
- [31] Y. Gao et al. “Approaches for quantitative study of divertor heat loads on W7-X”. In: *Proceedings of the 2018 International Conference on Quantitative InfraRed Thermography*. Berlin, Germany: QIRT Council, 2018, pp. 6–7. DOI: 10.21611/qirt.2018.p23.
- [32] Yu Turkin. “Current Control by ECCD for W7-X”. In: *Fusion Science and Technology* 50.3 (2006), pp. 387–394. DOI: 10.13182/FST06-5.
- [33] S.A. Bozhenkov et al. “Service oriented architecture for scientific analysis at W7-X. An example of a field line tracer”. In: *Fusion Engineering and Design* 88.11 (2013), pp. 2997–3006. DOI: 10.1016/j.fusengdes.2013.07.003.
- [34] F. Wagner et al. *Overview on W7-AS Results with Relevance for Wendelstein 7-X and the Low-Shear Stellarator Line*. Tech. rep. 1998, p. 33029492.
- [35] Martin Ester et al. “Density-Based Clustering Algorithms for Discovering Clusters”. In: *Comprehensive Chemometrics 2* (2010), pp. 635–654. DOI: 10.1016/B978-044452701-1.00067-3. arXiv: 10.1.1.71.1980.
- [36] Samuel A. Lazerson et al. “First measurements of error fields on W7-X using flux surface mapping”. In: *Nuclear Fusion* 56.10 (2016), p. 106005. DOI: 10.1088/0029-5515/56/10/106005.
- [37] Y. Gao et al. “Methods for quantitative study of divertor heat loads on W7-X”. In: *Nuclear Fusion* 59.6 (2019), p. 066007. DOI: 10.1088/1741-4326/ab0f49.
- [38] Samuel A Lazerson et al. “Error fields in the Wendelstein 7-X stellarator”. In: *Plasma Physics and Controlled Fusion* 60.12 (2018), p. 124002. DOI: 10.1088/1361-6587/aae96b.
- [39] Sergey Ioffe and Christian Szegedy. “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift”. In: *Proceedings of the 32nd International Conference on Machine Learning*. Lille, France, 2015. DOI: 10.1007/s13398-014-0173-7.2. arXiv: 1502.03167.
- [40] Christian Szegedy et al. “Rethinking the Inception Architecture for Computer Vision”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Las Vegas USA, 2016, pp. 2818–2826. DOI: 10.1109/CVPR.2016.308. arXiv: 1512.00567.
- [41] K He et al. “Deep residual learning for image recognition”. In: *IEEE conference on computer vision and pattern recognition*. Las Vegas USA, 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90. arXiv: 1512.03385.

- [42] Christian Szegedy et al. "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning". In: *AAAI*. San Francisco, USA, 2017, p. 12. DOI: 10.1016/j.patrec.2014.01.008. arXiv: 1602.07261.
- [43] Diederik P. Kingma and Jimmy Lei Ba. "Adam: A Method for Stochastic Optimization". In: *International Conference on Learning Representations 2015*. San Diego, USA, 2015, pp. 1–15. DOI: 10.1145/1830483.1830503. arXiv: 1412.6980.
- [44] Martín Abadi et al. "TensorFlow: A System for Large-Scale Machine Learning". In: *USENIX Symposium on OSDI*. Savannah, USA, 2016, pp. 265–283.
- [45] Zhou Wang et al. "Image Quality Assessment: From Error Visibility to Structural Similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. DOI: 10.1109/TIP.2003.819861.
- [46] Thomas J Diccio and Bradley Efron. "Bootstrap Confidence Intervals". In: *Statistical Science* 11.3 (1996), pp. 189–228. DOI: 10.1214/ss/1032280214.

A Neural Network Architectures

There are many types of NNs and by combining them and modifying parameters the number is infinite. The analysis done in "Neural Network Regression Approaches to Reconstruct Plasma Properties from Wendelstein 7-X Heat Load Images" is limited to five different types introduced in detail in the following. The hyperparameter of a convolutional layer are shown

A.1 Feed-Forward Fully-Connected Network

The first hidden layer consists of 80 neurons, the second layer consists of 100 neurons leading to the output of 2 values. The hidden layer activation functions are ReLU, the last layer activations are the identity.

A.2 Convolutional Network

The convolutional layer hyperparameter are given as shown in Figure 11. A visual representation of the three layers CNN is given in Figure 12. The architecture starts with a convolutional layer with 5×5 kernel followed by one with a 3×3 kernel. The dimension is reduced by a 2×2 Max-Pooling with stride 2×2 . The last convolutional layer has a 3×3 kernel and is again followed by 2×2 Max-Pooling with stride 2×2 . The last two layers are fully-connected, the first one has 24 neurons and the last one leads to the two output values. All activation functions are ReLU, except for the last layer, where activations are the identity.

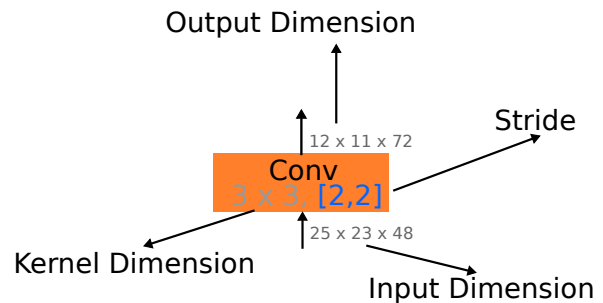


Figure 11: The hyperparameter for convolutional layer are given as shown here. The padding is given indirectly by the input and output dimension. If the width and height stay the same, padding is applied, otherwise, it is not applied.

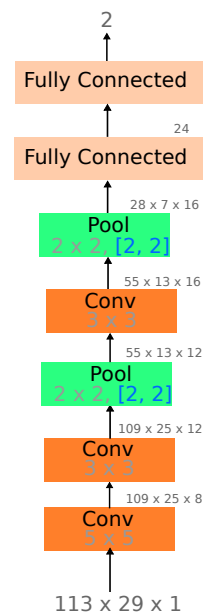


Figure 12: Overview of the CNN

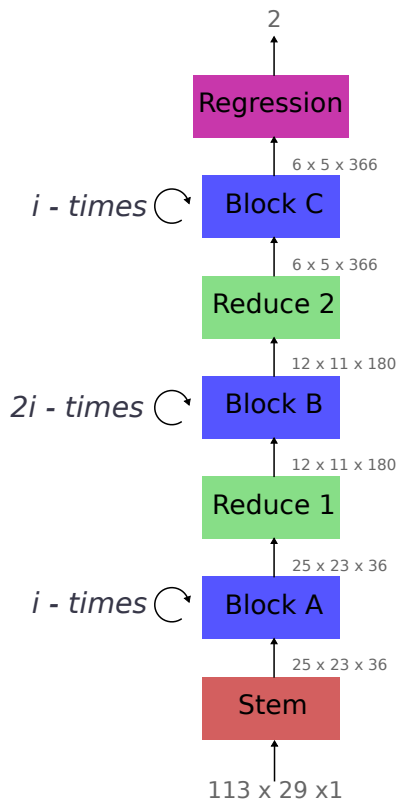


Figure 14: Overview of the IRNN. The parameter i determines the NN depth by giving the numbers of repetitions of each block. The individual blocks are described in Figures 15 to 21.

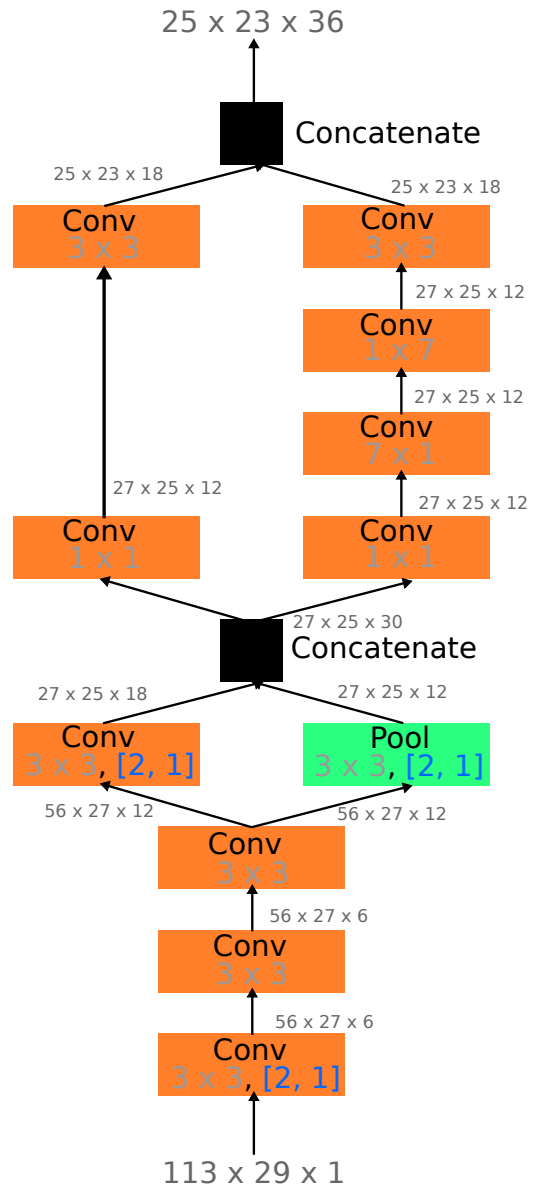


Figure 15: Stem of IRNN which structure overview is given in Figure 14

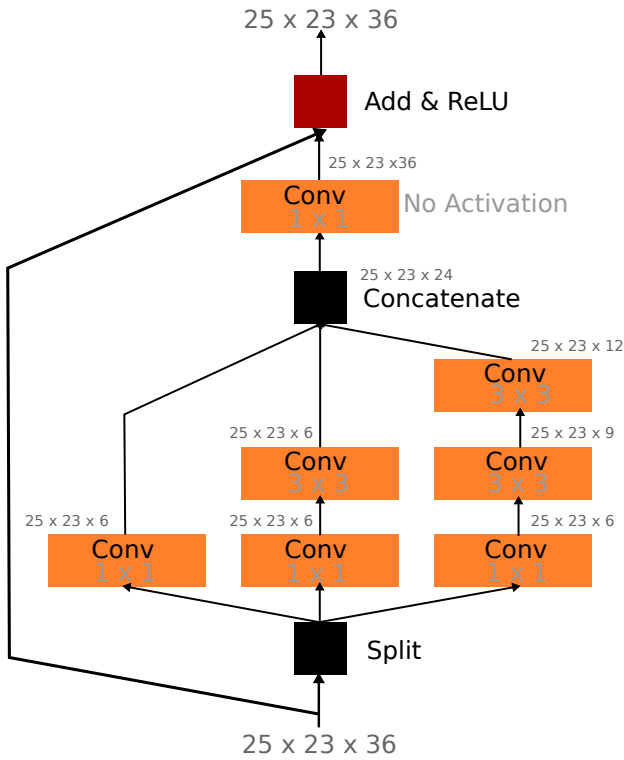


Figure 16: Residual block A of residual-inception NN which structure overview is given in Figure 14

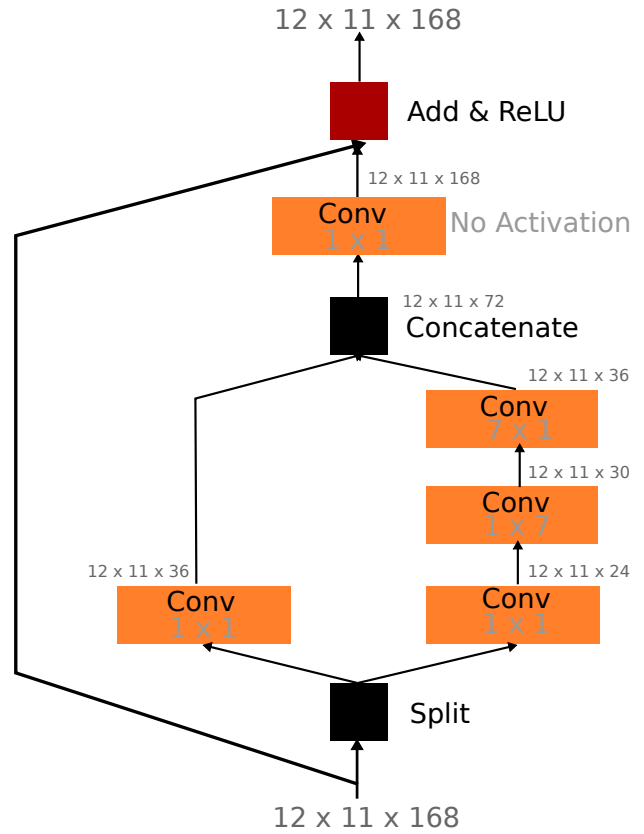


Figure 18: Residual block B of IRNN which structure overview is given in Figure 14

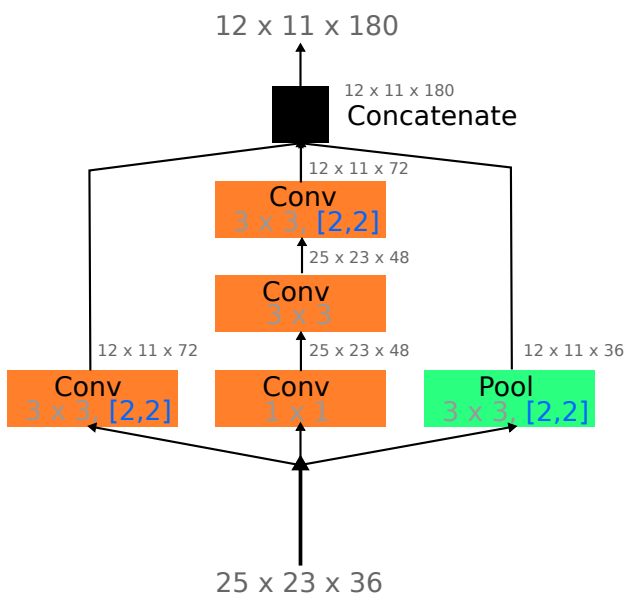


Figure 17: Reduce 1 module of IRNN which structure overview is given in Figure 14

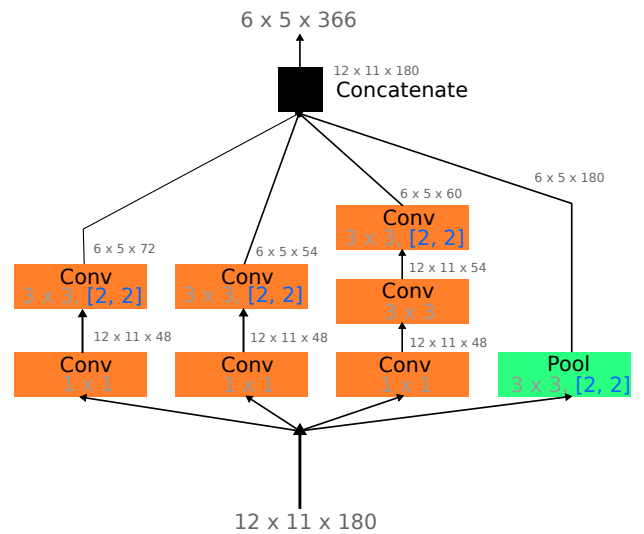


Figure 19: Reduce 2 module of IRNN which structure overview is given in Figure 14

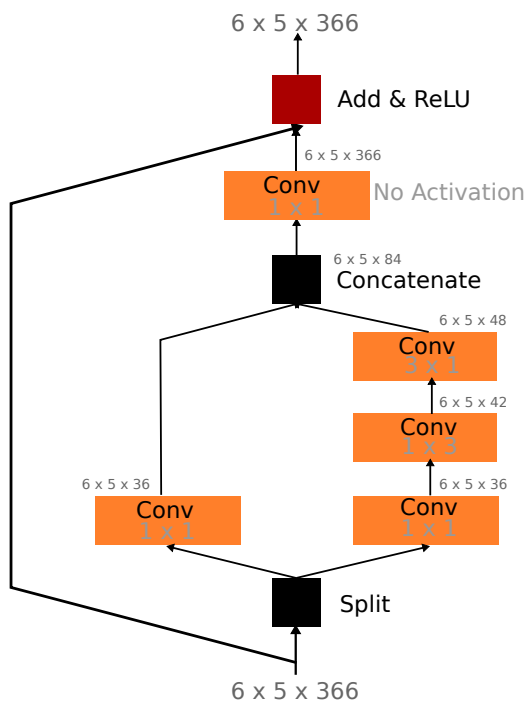


Figure 20: Residual block C of IRNN which structure overview is given in Figure 14

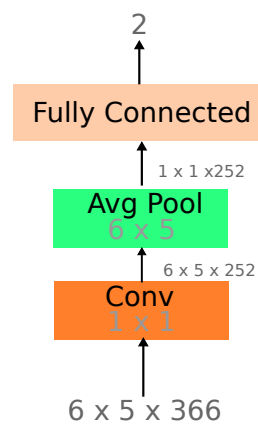


Figure 21: Reduce 3 module and regression of IRNN which structure overview is given in Figure 14

A.5 Deep Inception Network

The DINN is a special case of the IRNN, where the residual parameter i is equal to 0 which can be seen as skipping block A – C in Figure 14. Even without the residual blocks this NN has a deep and complex structure with multiple inception modules.

A.5. Article V

»Scenario with Combined Density and Heating Control to Reduce the Impact of Bootstrap Current in Wendelstein 7-X«

P. SINHA, D. BÖCKENHOFF, M. ENDLER, J. GEIGER, H. HÖLBE,
H. M. SMITH, T. SUNN PEDERSEN and Y. TURKIN

Nuclear Fusion, accepted for publication, (2019), DOI: 10.1088/1741-4326/ab37bb

Scenario with Combined Density and Heating Control to Reduce the Impact of Bootstrap Current in Wendelstein 7-X

P. Sinha,* D. Böckenhoff, M. Endler, J. Geiger, H. Hölbe,
H. M. Smith, T. S. Pedersen, Y. Turkin, and W7-X Team

*Max-Planck-Institut für Plasmaphysik,
Wendelsteinstraße 1, 17491 Greifswald, Germany*

(Dated: September 10, 2019)

Abstract

Wendelstein 7-X is a low-shear stellarator with an island divertor, formed by natural magnetic islands at the plasma edge and ten modular divertor units for particle and energy exhaust. For the island divertor concept to work properly, the device is optimized for small internal currents. In particular, the bootstrap current is minimized. Previous studies predicted a thermal overload of the targets at a particular location, due to the slow evolution of the toroidal net current in the initial phase of certain otherwise desirable high-power discharges. The present numerical study explores the neoclassical predictions for the bootstrap current in more detail and demonstrates, as a proof of principle, that a path from low density and low heating power to high density and full heating power exists, on which the bootstrap current remains constant. This offers the possibility to reach the predetermined toroidal net current at low heating power, where no overload will occur in the transient phase.

I. INTRODUCTION

Wendelstein 7-X (W7-X) is a large stellarator with five field periods, operated at the Greifswald site of the Max-Planck-Institut für Plasmaphysik (IPP) [1, 2]. Its magnetic field configuration was optimized for good MHD equilibrium and stability properties, low neoclassical transport and small bootstrap current in the long-mean-free-path (lmfp) regime, and good fast particle confinement at high β values. It aims to show that stellarator research can provide a magnetic configuration suitable for a reactor. The superconducting coil system consists of 50 non-planar coils of 5 different coil types and 20 planar coils of 2 different coil types. The 5 non-planar coil types and 2 planar coil types form one half module of the magnet system, two such half modules in stellarator symmetry form one module, which is repeated along the torus, following the fivefold symmetry of the device. Independent control of the currents in each of the coil types allows for considerable flexibility in changing the magnetic configuration. The aim is to operate W7-X in quasi-steady-state discharges at fusion relevant plasma parameters with a duration of up to 30 min.

The main heating system for this operation mode is Electron Cyclotron Resonance Heating (ECRH) [3], with ten continuous wave (cw) gyrotrons providing up to 1 MW each at 140 GHz for operating at a resonance field of 2.5 T. At medium plasma densities, the ECRH is operated in X2-mode (cut-off density at $1.2 \times 10^{20} \text{ m}^{-3}$). For densities beyond the X2 cut-off, O2-mode heating is used up to densities of about $2 \times 10^{20} \text{ m}^{-3}$ (cut-off at $2.4 \times 10^{20} \text{ m}^{-3}$), which has already been successfully demonstrated in the first divertor operation phase of W7-X [4, section 4]. At even higher densities O-X-B mode conversion is envisaged, which was tested in the predecessor experiment Wendelstein 7-AS (W7-AS) [5, 6]. In addition to heating, electron cyclotron current drive (ECCD) can be achieved by an appropriate setting of the ECRH launching mirrors. Further heating systems are neutral beam injection (NBI) with about 5 MW in 10 s pulsed operation [7], which has gone into operation in 2018, and ion cyclotron resonance heating (ICRH) with 4 MW [8] foreseen for the next operational phase, to start in 2021.

In order that the HELIAS concept of W7-X [9] be viable for a fusion reactor [10], a solution for particle and energy exhaust has to be provided. For the low-shear configuration of W7-X, an island divertor concept was chosen [11], which was successfully tested in the predecessor experiment W7-AS [12]. For the island divertor, the edge magnetic field topology

has a specific resonant structure in which magnetic islands formed by a naturally occurring low-order-rational value of the rotational transform ι at the boundary of the plasma generate a separatrix, which guides the particle and power fluxes from the main plasma to the divertor targets designed for high power loads [13]. The target plates are arranged in 10 divertor units of equal shape, corresponding to the fivefold toroidal symmetry and to the up-down flip symmetry (stellarator symmetry) of the device (see Fig. 1).

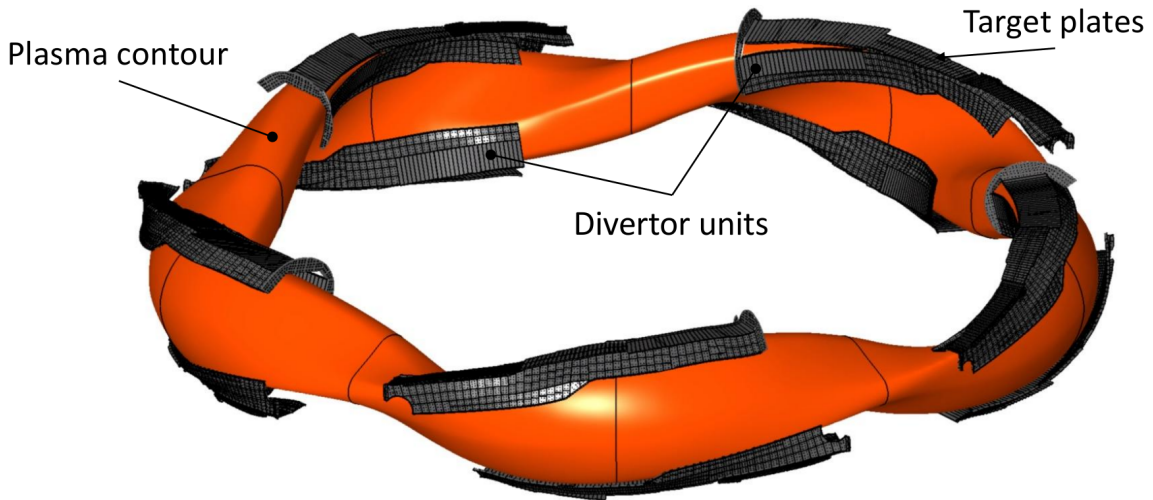


FIG. 1: Plasma column with the ten divertor units, positioned according to periodicity and stellarator symmetry [14].

The High Heat Flux (HHF) divertor units to be installed for future long-pulse operation are designed for stationary heat loads of 10 MW/m^2 in the highest loaded areas [15]. However, due to technical restrictions, there are regions of the target plates where the allowable power density to the surface is significantly lower. A design view of one divertor unit is shown in Fig. 2.

The edge rotational transform is of extreme importance for the island divertor operation, as it determines the radial position of the magnetic islands. If there is a net toroidal current I_{tor} , it changes the rotational transform and thus displaces the island chain radially relative to the vacuum configuration. Because of the low shear, the variation of the toroidal current must therefore be limited to rather small values on the order of 10 kA, or the impact on the magnetic configuration must be balanced by an appropriate adaptation of the field coil currents (or by ECCD [16]) (see section II).

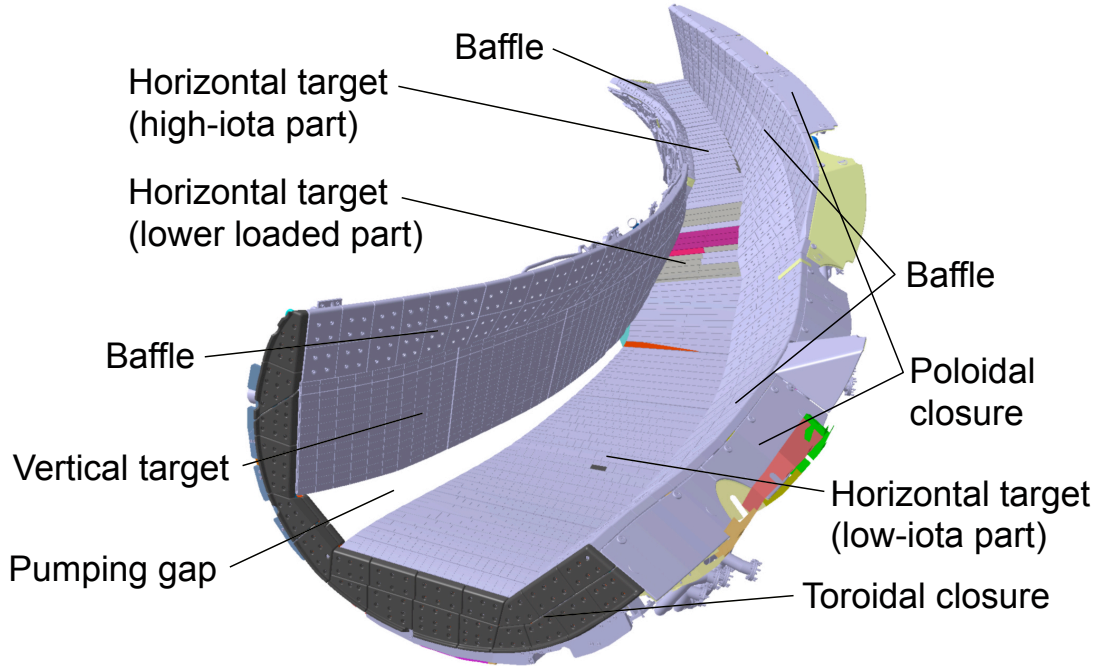


FIG. 2: Lower divertor unit showing the different components.

Wendelstein 7-X is optimized for small bootstrap current I_{bs} [17]. However, in some scenarios, I_{bs} , as calculated from neoclassical theory, can still assume values on the order of 50 kA. The net toroidal current (consisting of the I_{bs} and an exponentially decaying shielding current which is induced by the self-inductance of the plasma) reaches steady state after several L/R times which, for W7-X plasma parameters, may require several tens of seconds:

$$I_{tor} = I_{bs} (1 - \exp(-t/\tau)), \quad (1)$$

where $\tau = L/R$ is the decay time constant of a current in a circular conductor with inductance L and resistance R (see [18] for a detailed discussion of the evolution of currents in a stellarator, including the role of bootstrap current and current drive). An example of the impact of the evolution of I_{tor} on the edge configuration is shown in Fig. 3.

The impact of the variation of the net toroidal current on the heat load pattern on the target plates can be significant and may have to be minimized. Specifically, the area of maximum heat load typically forms one or several narrow stripes with a width of a few centimeters in the poloidal direction and extending along the target in the toroidal direction. These stripes are called strike line(s), and their poloidal location on the targets is sensitive to I_{tor} . In particular, an attractive discharge scenario was identified [16] where, however,

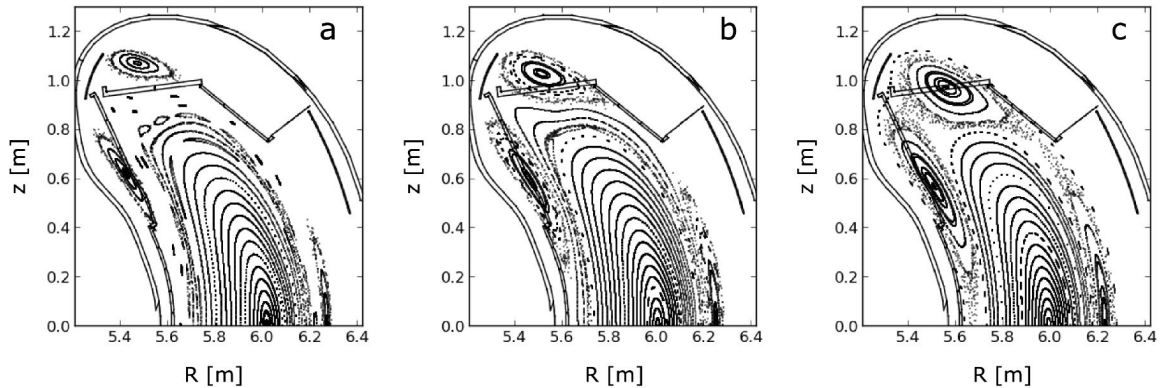


FIG. 3: Poincaré plot of different stages of evolution of the toroidal current in the “scraper element reference scenario” [16] (see section II): (a) 0 kA (b) 22 kA (transitional state) (c) 43 kA (final state) (from [19]).

the strike line would approach the pumping gap during the toroidal current equilibration, thus overloading the target edges bounding the pumping gap and possibly wall components behind the pumping gap (see Figs. 6 and 3) [20]. The analysis in [16], however, used neoclassical transport coefficients in the calculation of radial profiles, which had been calculated for a single density case in the magnetic “standard reference” configuration (which has similar vacuum field properties). The MHD equilibrium, neoclassical transport coefficients and radial profiles of plasma pressure and toroidal current density were therefore not consistent. Part of the present work is to calculate, in a proof-of-principle study, stationary plasma states with different values of I_{bs} and consistency between MHD equilibrium, neoclassical transport and radial profiles, which will be called “self-consistent solutions” in the following. The aim of this work is to demonstrate the existence of a sequence of such stationary states, where, first, a certain value of I_{bs} is achieved at low power and appropriate plasma density. During the evolution of this first plasma state from startup to the desired final stationary state, the configuration with significant power loads on sensitive locations of the targets will be passed at a low power level. The further plasma states of the sequence will then be chosen in such a way as to balance the scalings of I_{bs} with pressure gradient and with collisionality, such that I_{bs} is kept constant. The solution to this task is not entirely self-evident, since I_{bs} itself depends on MHD equilibrium (not only on the structure of $|B|$, but also on t , i. e., the current density profile must be considered), plasma profiles (collisionality and radial electric field) and on the thermal transport calculated from the mono-energetic neoclassical

transport coefficients in a subtle, non-linear way in a magnetic configuration like that of W7-X, which was optimized, among other criteria, for vanishing bootstrap current.

We note that we do not aim at describing a dynamic time evolution by solving time-dependent equations, but we investigate stationary states. For low power, where the overload is avoided, a slow evolution of the toroidal current can be tolerated, and for the phase with the final I_{bs} , the toroidal current should be constant in this approach with no change of the location of the boundary islands. The detailed time evolution is therefore not considered here.

The paper is organized as follows. After revisiting the different solutions suggested to solve the overload problem and introducing our new approach in section II, we shall discuss the numerical method in section III. We shall then show a potential path from a low-density, low-power plasma to high-performance conditions without overloading the target edges at the pumping gap in section IV. After discussing these results and the underlying assumptions in section V, we summarize our conclusions in section VI.

II. SCHEMES TO MITIGATE THE IMPACT OF THE VARIATION OF THE TOROIDAL CURRENT

Different ways have been proposed to avoid the overloading of components during the evolution of the toroidal current density toward its steady-state profile. They can be grouped into two categories:

1. Transient modification of the magnetic configuration by external current drive or by changing the currents in the field coils,
2. modification of the plasma-facing components,

In this section, we shall first revisit the basic idea behind each of the schemes proposed so far and shortly discuss their respective disadvantages. We shall then introduce a third category, making use of the dependance of the bootstrap current on plasma density and temperature.

A. Current Drive

To keep the net value of I_{tor} in an acceptable range, Electron Cyclotron Current Drive (ECCD) can be applied [18]. The two major drawbacks of using this technique are, firstly, that it works efficiently only for intermediate or low plasma densities. Secondly, since the current is driven in a highly localized radial position, the rotational transform profile is strongly modified, giving rise to many low-order rational magnetic surfaces. While compensating the net current, ECCD therefore sacrifices the low-shear profile of the rotational transform, which may destabilize the plasma [16, 21, 22]. Experimentally, strong MHD activity has already been observed in discharges with ECCD [23].

B. Transient Change of Field Coil Currents

The most pronounced change in the plasma edge due to a net toroidal plasma current is the radial shift of the island chain in configurations with edge islands. This could also be compensated by adjusting the ratio of the planar and non-planar field coil currents. The edge rotational transform of the vacuum magnetic field would then be dynamically adjusted to compensate for the effect of the toroidal plasma current. First, a feed-forward scheme could be applied if the current evolution is known, as it was demonstrated on other stellarators (see, e. g., [24, 25]). There are, however, limits for the rate at which the currents in the superconducting field coils of W7-X can be changed. A first experiment with limited scope has been performed to test this option, the analysis of which is still ongoing. Ultimately, an ι control using the field coils would require a feedback system controlling the power supplies of the magnet system. This has not been planned originally. Technically, it would require considerable adjustments in the control and safety system of W7-X. We shall not discuss this method further here.

C. Addition of a Further Target Component

Another possibility to protect the edges of the pumping gap is the installation of a set of new actively cooled passive protection elements called Scraper Elements (SE) [20]. They prevent the heat flux from reaching the critical regions of the divertor, where it might overload the components during the transition to a stationary current state. However,

numerical simulations indicate that the SEs could deteriorate the plasma performance since they significantly reduce the pumping efficiency [26]. This scheme has been tested in the first divertor operation phase with a set of only two SEs for 2 out of the 10 target units. Detailed analysis of the results is ongoing [27].

D. Controlling the Variation of Bootstrap Current by Control of Density and Heating Power

In the previous analysis of the overload problem, a final high-performance plasma state with a certain density and heating power was taken as the goal, and the resulting bootstrap current was calculated by a transport code (see section III). It was then assumed that density and heating power would quickly assume their target values, whereas the toroidal current would slowly evolve on the L/R time scale, overloading the target edges at the pumping gap during several tens of seconds.

However, the bootstrap current naturally depends not only on the magnetic configuration but also non-linearly on density and heating power. This offers the option to have for some desired magnetic configuration the same bootstrap current at low density and low heating power as at the high density, high heating power target state. The critical range of I_{tor} could then be reached and exceeded at reduced heating power levels, avoiding an overload. After that, a path with constant I_{bs} to high power and high density would be followed by a controlled increase of power and density. While ECRH power and its deposition profile is easily controlled, this is not as easy a task for plasma density and the density profile. Here, we shall nevertheless assume that these quantities can be sufficiently well controlled during a plasma discharge.

For the present proof-of-principle study, we choose a magnetic configuration which is as desirable as the one for which the scraper elements were conceived (SE reference configuration) [20, 28] but slightly different. The reason is that in this configuration, which has the same low neoclassical losses (see section IV C) and good MHD stability at a heating power of 10 MW as the standard reference configuration, I_{bs} is predicted to reach ~ 30 kA. Therefore the vacuum configuration is slightly adjusted to achieve divertor heat load patterns compatible with the technical limits of the water-cooled HHF divertor in the steady-state phase. As announced in section I, we aim to develop a path consisting of stationary plasma

scenarios, each of which has a consistent set of MHD equilibrium, neoclassical transport coefficients and plasma profiles from transport simulations, ranging from low power and low density to the desired high-density steady-state without overloading the divertor. The field coil currents for this new configuration are listed together with those of the SE reference configuration in Table I. In the following sections, we shall describe the computer codes and

TABLE I: Coil currents for the vacuum magnetic fields with 2.52 T on the magnetic axis at the ECR heating position for the configuration discussed in this paper (‘new’) and for the scraper element reference configuration [28] (‘SE ref’). Here, I_1 to I_5 are the currents in the non-planar coils and I_A , I_B the ones in the planar coils. The offset in edge rotational transform of the ‘new’ vacuum field relative to the target configuration has been chosen to

account for $I_{\text{tor}} = 30$ kA.

[kA]	I_1	I_2	I_3	I_4	I_5	I_A	I_B
new	12.10	11.96	12.22	13.48	13.60	7.68	-4.06
SE ref	11.95	11.82	12.07	13.31	13.43	8.86	-2.86

their assumptions and the resulting discharge path toward a high-performance plasma state.

III. MODELLING APPROACH

In this chapter, we discuss the building blocks of our analysis. In section III A, a general outline of their interrelation will be given. After that, we shall list the assumptions made in these numerical tools and indicate where more detailed descriptions can be found. We shall start with the three codes used to calculate self-consistent magnetic equilibrium configurations (in the sense defined in section I), the Variational Moments Equilibrium Code (VMEC) [29] and the EXTENDER code [30] (section III B), the Drift-Kinetic Equation Solver (DKES) [31] (section III C), and the Neoclassical Transport Simulation for Stellarators (NTSS) code [32] (section III D). We shall continue in section III E with a description how these codes have been used iteratively to obtain self-consistent radial profiles. We shall conclude this chapter by describing the calculation of thermal loads on the targets in a field line diffusion code [33] in section III F.

A. Interrelation of Numerical Tools

The radial profiles of toroidal current density and pressure for some W7-X configuration can be calculated by the NTSS transport code, which takes as input a radial density profile, a radial heating power profile, and coefficients describing neoclassical and anomalous transport. Neoclassical transport is important in stellarators in particular in the Imfp regime.

A table of (mono-energetic) neoclassical transport coefficients is obtained from the DKES code, which requires as input the magnetic equilibrium configuration (in terms of a Fourier representation of the magnetic flux surfaces and a radial profile of the rotational transform).

This description of the magnetic equilibrium configuration can be calculated by the VMEC code, which takes as input the vacuum magnetic field and the radial profiles of pressure and of toroidal current density.

Obviously, there exists a circular dependance between the output of these three codes and (at least part of) their input, as shown in Fig. 4. Until now, this iteration has not been automated, since the output of the codes (in particular of DKES) requires some assessment with regard to the accuracy of the results. In the past, the sequence of codes was iterated to a self-consistent solution (within the assumptions) for one scenario with very low bootstrap current (on the order of a few kA) and for a second scenario with ECCD to keep the toroidal net current equally small [22]. In this paper, we present for the first time self-consistent iteration results for configurations with larger bootstrap current. This has an impact on the iteration since the bootstrap current changes the t profile which in itself influences the bootstrap current.

The iteration is started with assumed shapes for the radial profiles of pressure and toroidal current density (see section III E), together with an average β and a value of the toroidal net current.

Once the changes between iterations are negligible, the magnetic equilibrium configuration calculated by VMEC for the confinement region is used by the EXTENDER code to calculate the full magnetic field in the entire vacuum chamber, produced by the plasma in magnetohydrodynamic (MHD) equilibrium and by the external coils. This full field is then used to calculate the thermal load to the targets and other wall components in a field line diffusion code.

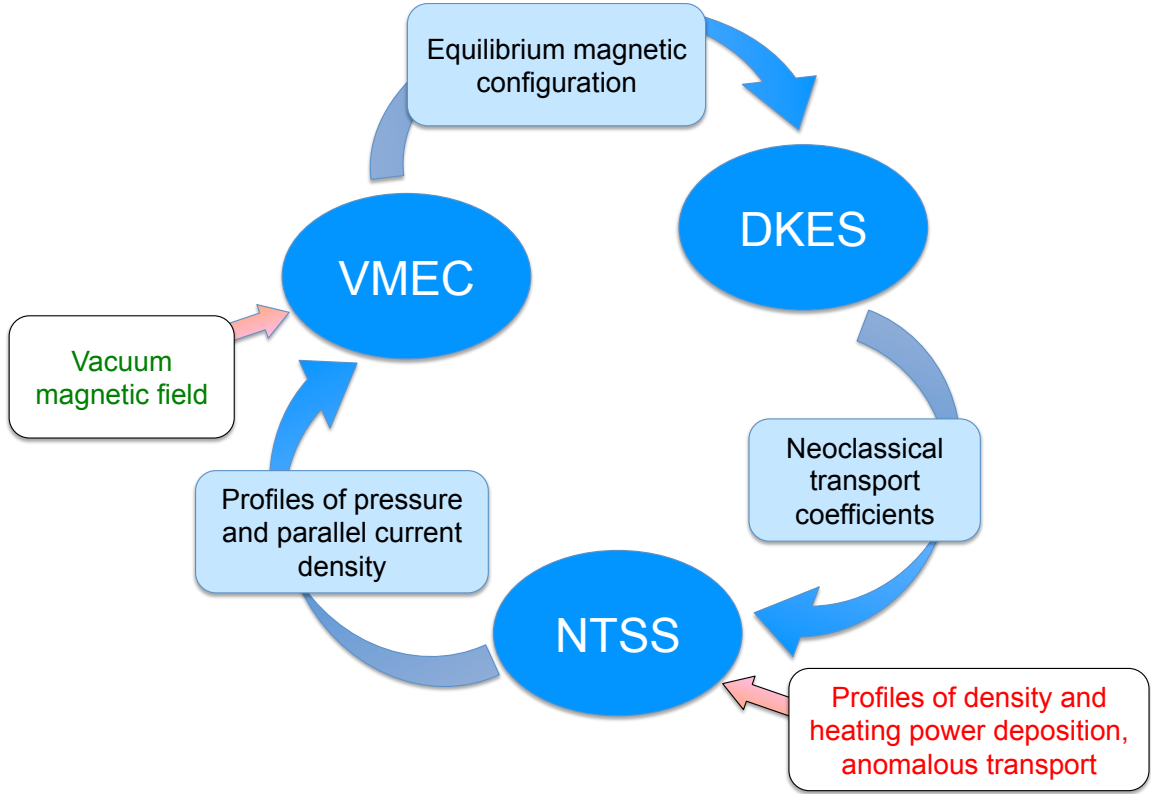


FIG. 4: Iterative modelling using the codes VMEC, DKES and NTSS to calculate a magnetic equilibrium configuration with consistent radial profiles and neoclassical transport coefficients.

B. VMEC and EXTENDER

VMEC is used to calculate the magnetic field in the presence of a plasma by solving the ideal MHD equations for a stationary equilibrium

$$\nabla \cdot \mathbf{B} = 0, \quad (2)$$

$$\mathbf{J} \times \mathbf{B} = \nabla p, \quad (3)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \quad (4)$$

VMEC assumes that the magnetic field forms nested flux surfaces, which are labeled by the value of the normalized toroidal flux s (VMEC's radial coordinate, running from 0 in the plasma center to 1 on the VMEC domain boundary). This surface limiting the computational domain is also assumed to be a flux surface. In a free-boundary calculation this defines the volume of the plasma. As basic input, VMEC requires the toroidal flux contained within this

limiting surface. VMEC by itself has no way to determine whether these assumptions are valid or not and cannot judge on the size of its computational boundary. For our iterations aiming at a certain value of I_{bs} and hence at a fixed radial location of the bounding island chain, we used a fixed value of the toroidal flux. Due to the assumption of nested flux surfaces, VMEC cannot treat islands or stochastic regions. However, the full field can be used to investigate the location of the boundary islands and of the separatrix to a leading order, and this information can be used to adjust the input parameters of VMEC accordingly to exclude the island region from the calculation domain, as described in [16]. In the sequence of calculations we kept the volume of the VMEC domain almost constant at $\sim 28 \text{ m}^3$. With this value, the 5/5 boundary islands are outside the VMEC domain for $I_{\text{tor}} = 30 \text{ kA}$. For simplicity, this volume was also used for the low-power, lower-density phase, where, with smaller toroidal currents, the boundary islands are further out, but where an overload of the critical divertor parts does not occur (see section IV C).

The model of nested flux surfaces is necessary for the iterative use with DKES and NTSS, which assume that transport is local and calculate the integral fluxes through (radial) and within (parallel) the flux surfaces.

Further basic input data to VMEC are the radial profiles of plasma pressure $p(s)$ and toroidal current density $j_{\text{tor}}(s)$.

To investigate the effects of the plasma currents on the edge islands and to study their impact on the interaction of the plasma with the divertor targets and the wall components, magnetic fields in the entire plasma vessel are required. For this purpose the EXTENDER code is employed which uses the virtual casing principle [34, 35] to calculate the fields outside of the VMEC domain generated by the plasma and combines them with the vacuum field to obtain the full field. For the field in the VMEC domain we do not use the VMEC field itself, but rather a combination of the VMEC field, the EXTENDER solution inside the VMEC domain and the vacuum field, which is calculated by Biot-Savart's law from the currents in the field coils, as described in [16]. The full field built in this way is well suited to investigate the edge region of the magnetic configuration and its interaction with the divertor structures, as used in [27, 28].

C. DKES

The neoclassical part of the particle and energy transport can be calculated with transport coefficients derived from a drift-kinetic description of the plasma. The so-called Drift-Kinetic Equation (DKE) [36] allows a steady-state solution for the distribution function f , which is assumed to differ from the local Maxwellian distribution function f_M only by a “small” perturbation $f_1 = f - f_M$. The DKE for f_1 can be linearized on the assumption that the transport processes are radially (on flux surfaces) local. For an expression of the linearized DKE allowing it to assume a so-called mono-energetic form see [37]. Thus the initially 5-dimensional DKE can be reduced by 2 dimensions, since the corresponding variables, flux surface label and particle energy, appear only as parameters, thus simplifying the numerical treatment. To allow the mono-energetic ansatz, energy diffusion is neglected in the collision operator and only pitch-angle scattering is retained. This has no consequences for the radial transport in stellarators but causes errors for the parallel transport (along field lines) because it violates parallel momentum conservation. However, techniques for parallel momentum correction are available [38] to cure this deficiency in the parallel flows at a later stage and thereby recover correct results.

The solution of the DKE, which is numerically calculated with the DKES code, is used to obtain the transport coefficients for fluxes or flows through (radial) or within (parallel) a flux surface by integrating the properly weighted perturbation of the distribution function over the remaining 3 dimensions, i. e. the pitch angle and the two angular coordinates (flux surface average). The resulting so-called mono-energetic transport coefficients $D_{ij}^{(\alpha)}$ of each particle species α

$$D_{ij}^{(\alpha)} = D_{ij}^{(\alpha)} \left(r, \frac{\nu_\alpha}{v_\alpha}, \frac{E_r}{v_\alpha B_0} \right) \quad (5)$$

capture the transport properties of particles with a specific energy, thus the designation. Here, ν_α , v_α , E_r and B_0 are the collision frequency, velocity (corresponding to the chosen energy), radial electric field and magnetic field, respectively [32]. The flux-surface-averaged flows $I_i^{(\alpha)}$ and the thermodynamic forces $A_j^{(\alpha)}$ driving these flows can then be written as

$$I_i^{(\alpha)} = -n^{(\alpha)} \sum_{j=1}^3 L_{ij}^{(\alpha)} A_j^{(\alpha)}, \quad (6)$$

where $n^{(\alpha)}$ is the particle density, $I_1^{(\alpha)}$ is related to the radial particle flux density, $I_2^{(\alpha)}$ to the radial energy flux density, and $I_3^{(\alpha)}$ to the toroidal current density (for stellarators without

an externally imposed electric field this is essentially the bootstrap current — for exact definitions see [37]).

The so-called thermal transport coefficients $L_{ij}^{(\alpha)}$ are calculated from the mono-energetic ones by an energy convolution with a local Maxwellian with appropriate energy weightings (see also [37]).

For each magnetic configuration calculated with VMEC, a table of mono-energetic transport coefficients $D_{ij}^{(\alpha)}$ is generated, reflecting the dependences shown in Eq. (5) to allow a fast convolution for arbitrary plasma parameters in a 1D transport simulation code like NTSS. The generation of these DKES tables (for several minor radii, collisionalities and radial electric field values) is the most computationally expensive step of the entire iteration cycle, in particular the calculations for low collisionalities. Nevertheless, since the tables are later used in interpolating functions, the results need to be checked manually for their quality and suitability, especially the low-collisionality values. This is currently a necessary part of the cycle.

D. NTSS

The NTSS code has been designed to solve the time-dependent set of transport equations encompassing the particle and energy balance of electrons and ions together with a diffusion equation for the radial electric field E_r and an equation for the evolution of the poloidal flux, the latter being equivalent to the evolution of the toroidal current density [18, 32]. For our purpose, we use only the ability to provide the stationary solutions of these equations. In particular, we do not solve the time evolution of the toroidal current density or poloidal flux but only use the calculated bootstrap current distribution resulting from the transport coefficients and plasma profiles.

Particle and energy balance equations contain on the one hand the neoclassical transport, based on first principles (see [32]), using the transport coefficients calculated with DKES. On the other hand NTSS uses heuristic models to account for the turbulence-driven, so-called anomalous, transport. This is simulated just by a simple diffusion model, where the heat diffusion coefficient is, in our case, $\chi_{\text{base}} = 1 \text{ m}^2\text{s}^{-1}$ at the boundary of the model region and is inversely proportional to the density. This model serves two purposes: first, allowing sufficient transport at the plasma boundary, where neoclassical transport is far too

low to produce reasonable temperatures, and second, ensuring that neoclassical transport will display its performance limiting role in the core plasma. The model is based on certain high-performance experiments in the Wendelstein 7-AS stellarator, which showed a core region with a dominantly neoclassical transport [39, sections 3.3 and 5.3.2].

In the NTSS simulations used here, the particle balance equation is not solved, but the density profile is kept fixed (see Fig. 5), since the modelling of particle sources would require

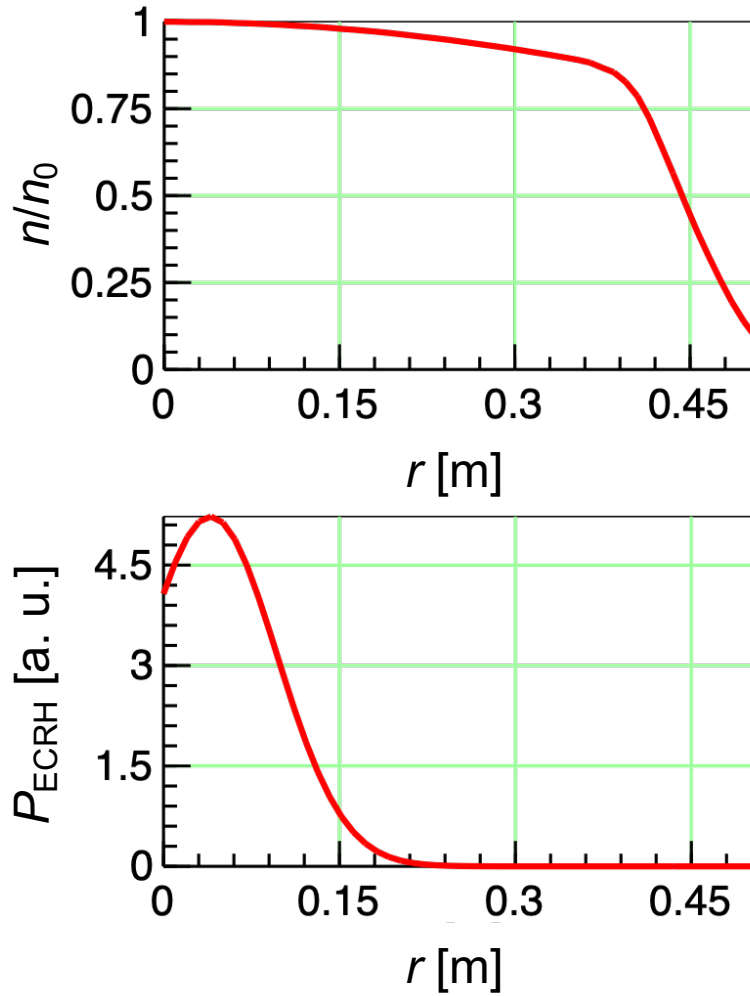


FIG. 5: Fixed density (top) and heating power (bottom) profiles used in the NTSS simulations.

the coupling with further codes. Because the density profile has been chosen rather broad, this results in a fast decay of the anomalous contribution in the energy balance equation to about a tenth of the boundary value in most of the plasma volume.

We aim at a high-performance plasma scenario at high density, requiring ECR O2-mode

heating, which in turn asks for central deposition for high electron temperatures and relies on a three-path absorption, fixing the launching angles of the beams [40]. For the power deposition profile we therefore use a fixed standard profile approximating the real O2 deposition profile with sufficient accuracy (see Fig. 5). Other heating methods with different power deposition profiles are not considered here.

The calculation of E_r is performed with a diffusion equation to avoid the discontinuities which occur for the non-linear ambipolarity condition $Z_i\Gamma_i = \Gamma_e$ (balancing ion and electron particle fluxes $\Gamma_{i/e}$) at a bifurcation point where electron-root and ion-root solutions are interchanged.

The bootstrap current is calculated from the parallel neoclassical flows and can be used in the time-dependent version in the evolution of the poloidal flux (equivalent to the toroidal current density — see [18] for details). As stated before, using the DKES transport coefficients for the parallel transport equations requires the use of momentum correction techniques [38]. These are implemented in NTSS.

Thus, in principle, given sources and sinks for particles, energy and current (e.g. by ECCD), NTSS can calculate the the time evolution of the radial profiles of electron and ion density and temperature, of the radial electric field and of the bootstrap current. As stated before, we omit the solution of the particle transport equation in this work and instead keep the density profile fixed, and we only consider stationary solutions. One should note, that these calculated profiles depend on

1. the MHD equilibrium, through the configuration dependency in the DKES calculation,
2. the assumed experimental scenario with density and heating and
3. the anomalous transport model.

E. Iterative Modelling for consistency between magnetic configuration, neoclassical transport and radial profiles

As indicated in section III A, part of the input and output of the three codes VMEC, DKES and NTSS cyclically depend on each other. We start the cycle by assuming initial profiles for pressure and toroidal current density $p_{\text{in}(1)}(s) = p_{\text{ini}}(s)$ and $j_{\text{tor,in}(1)}(s) = j_{\text{tor,ini}}(s)$, where $j_{\text{tor}}(s)$ represents the toroidal net current density, averaged on the flux surface with

label s . In terms of VMEC's numerical procedure,

$$\frac{dI_{\text{tor}}(s)}{ds} = j_{\text{tor}}(s) \frac{dA}{ds} \quad (7)$$

is used rather than $j_{\text{tor}}(s)$, where $I_{\text{tor}}(s)$ is the total toroidal current inside the flux surface with label s , and $A(s)$ is the toroidally averaged cross-sectional area of this flux surface. VMEC uses this input only for the functional shape of the profile and scales it to obtain the target value of the bootstrap current $I_{\text{bs,target}}$, which is a separate input. For the sake of simplicity, we shall here refer to $j_{\text{tor}}(s)$ as input profile of the toroidal current density, which satisfies

$$\int_0^1 j_{\text{tor,ini}}(s) \frac{dA}{ds} ds = I_{\text{bs,target}}. \quad (8)$$

A VMEC run under these assumptions yields a magnetic equilibrium configuration, which is used to obtain from DKES a table of mono-energetic neoclassical transport coefficients, as described in section III C. The magnetic configuration from VMEC and the transport coefficients from DKES are then used for a number of NTSS simulation runs with the same heating power P and power deposition profiles, the same predefined density profile shape, but different values of the central density n_0 . In addition, anomalous transport is simulated in the NTSS runs by a predefined heat diffusivity (see section III D).

The result after iteration cycle no. k are new pressure and toroidal current density profiles $p_{(k)}(s)[P, n_0]$ and $j_{\text{tor,(k)}}(s)[P, n_0]$, with different values of the bootstrap current

$$\int_0^1 j_{\text{tor,(k)}}(s)[P, n_0] \frac{dA}{ds} ds = I_{\text{bs,(k)}}[P, n_0]. \quad (9)$$

For the next iteration $k + 1$, we choose the profile shapes of that density value n_0 , for which $I_{\text{bs,(k)}}$ is closest to the target value of the bootstrap current $I_{\text{bs,target}}$. We denote the selected value of n_0 as $n_{0,\text{base}(k)}$. The iterations are performed separately for each value of P , and also $I_{\text{bs,target}}$ may differ for different values of P . For the next VMEC run, we use, instead of $I_{\text{bs,(k)}}[P, n_{0,\text{base}(k)}]$, again $I_{\text{bs,target}}$ as input for the total toroidal current since this is the target value of the entire iteration: The configuration calculated by VMEC will thus change between iterations only due to the different profile shape, and the rotational transform at the boundary of the VMEC domain will only change very little between iterations. Hence,

for the next iteration cycle, the input profiles for VMEC are

$$p_{\text{in}(k+1)}(s)[P] = p_{(k)}(s)[P, n_{0,\text{base}(k)}] \quad (10)$$

$$j_{\text{tor},\text{in}(k+1)}(s)[P] = \frac{I_{\text{bs,target}}}{I_{\text{bs},(k)}[P, n_{0,\text{base}(k)}]} \times j_{\text{tor},(k)}(s)[P, n_{0,\text{base}(k)}]. \quad (11)$$

From the new VMEC equilibrium, a new table of neoclassical transport coefficients is generated by DKES. NTSS is run next for a range of n_0 values around $n_{0,\text{base}(k)}[P]$, possibly including an interpolated value of n_0 to more closely hit $I_{\text{bs,target}}$ in the result.

In section IV A, we shall present the resulting bootstrap current $I_{\text{bs},(1)}[P, n_0]$ of the first iteration cycle for a broad range of (P, n_0) values to show the general dependence on these parameters. We shall then proceed up to the 4th iteration cycle for selected power steps to demonstrate that the convergence of the profiles and of I_{bs} to the desired target value is achieved in our procedure.

F. Field Line Diffusion and Assessment of Critical Regions

To calculate the heat load on different regions of the divertor and other plasma facing components (PFCs), the technique of field line diffusion is used, as described, e. g., in [33, 41]. To this end, a number N_{tot} of starting points is chosen, which are distributed uniformly on a flux surface a few millimeters inside the last closed magnetic surface. From each starting point, the magnetic field lines are followed in both directions with small random perpendicular steps taken to imitate perpendicular diffusion. The parameters to describe this diffusion process and their values chosen for this work are the diffusion coefficient $D_{\perp} = 1 \text{ m}^2\text{s}^{-1}$, the scale of the path length between two perpendicular steps $\lambda = 0.1 \text{ m}$ and the parallel “velocity” (which, together with D_{\perp} , determines the ratio between parallel and perpendicular transport) $v = 5 \times 10^5 \text{ ms}^{-1}$. Each trace is followed until it intersects a PFC. The location of each such hit point is recorded.

To calculate the power density, an appropriate segmentation of the PFCs is chosen, and for each segment i the number of hits on this segment N_i is divided by the segment area A_i . The power density on this segment is therefore calculated as

$$q_i = \frac{N_i P_{\text{conv}}}{N_{\text{tot}} A_i}, \quad (12)$$

where P_{conv} is the total available convective power [42] (e. g., for a worst-case estimate, the total heating power). Here, the segmentation size is chosen small enough to resolve the regions of the PFCs with different technical heat flux limits and the scale size of the heat load patterns. For the target plates, mostly segments corresponding to the individual tiles on the high heat flux target elements (with a typical size of $50 \text{ mm} \times 25 \text{ mm}$) are chosen. The calculated power loads q_i are compared to the maximum design heat load $q_{d,i}$ (see Table II and Fig. 6). Note the sides facing the pumping gap, where the reduced $q_{d,i}$ values are critical

TABLE II: Technical limits $q_{d,i}$ for target components in W7-X [43].

Component	$q_{d,i}$ [MW/m ²]
High loaded area	10.0
- End Top Tile	5.0
- Edge Tile	2.0
Lower loaded area	0.5

from the overload analysis point of view.

The ratio $q_i/q_{d,i}$ together with the number of hits N_i out of N_{tot} , on which q_i is based, can be used for a statistical analysis of the probability of overloading a particular segment [42].

IV. RESULTS

A. Bootstrap Current for Variation of Density and Heating Power

To show the general trends, we started one iteration cycle by creating MHD equilibria with VMEC, based on a vacuum magnetic field generated by the coil currents of Table I, assuming a linear pressure profile

$$p_{\text{ini}}(s) = p_c(1 - s) \quad (13)$$

and a parabolic current profile

$$j_{\text{tor,ini,A}}(s) = j_{c,A} s(1 - s) \quad (14)$$

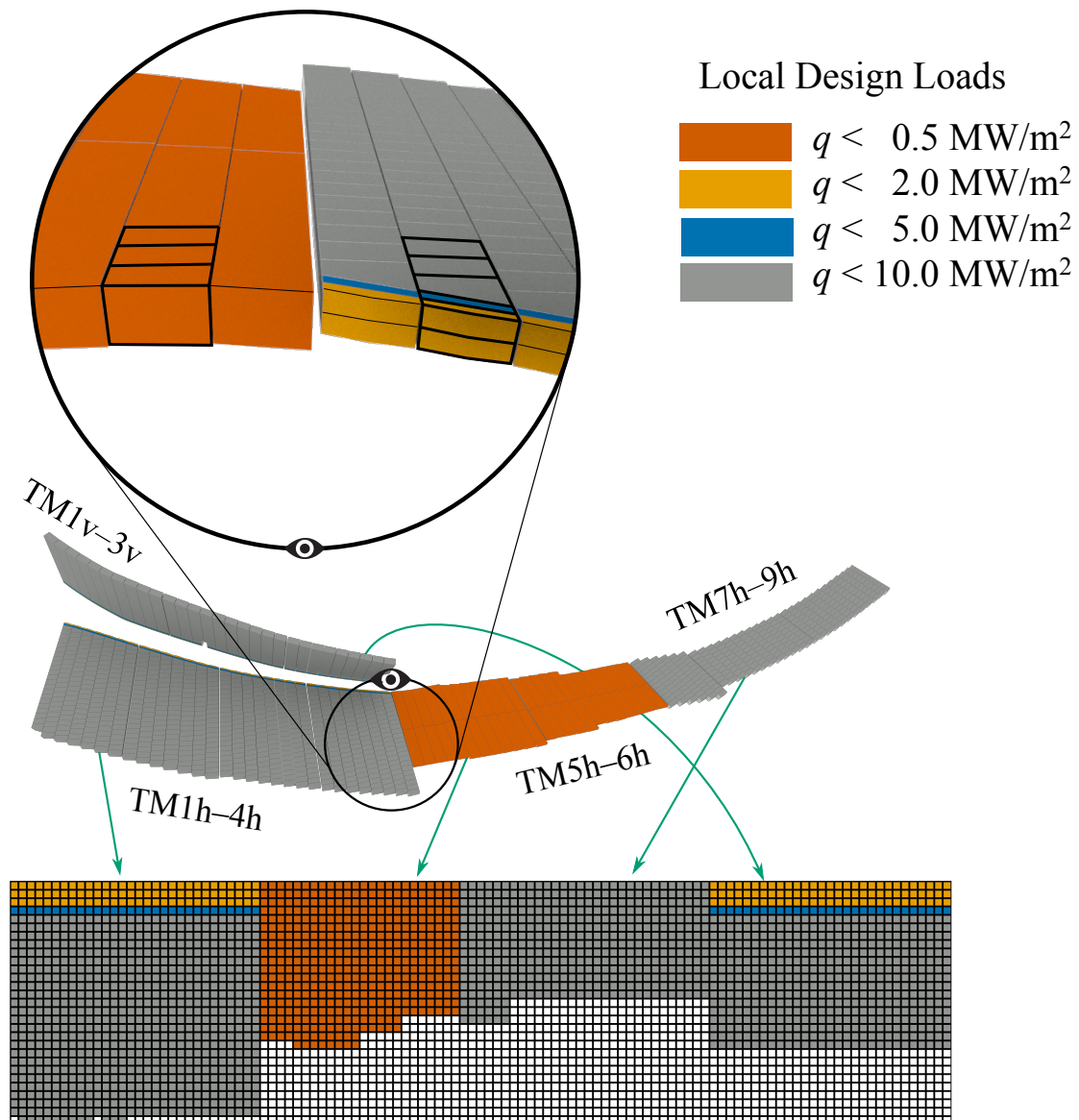


FIG. 6: Mapping of a divertor unit to a 2D grid, as used for representing the power density in section IV. The colors indicate the maximum design heat load (Table II) for each surface element [42]. The magnified view from the inboard shows in particular the critical region adjacent to the pumping gap in a golden-orange color.

with p_c and $j_{c,A}$ chosen such that the $\langle\beta\rangle$ value is 2% or 3% and the total toroidal current is $I_{bs,target} = 30$ kA. Mono-energetic neoclassical transport coefficients were calculated for these equilibria and used in NTSS for a range of central densities n_0 and heating powers P . The resulting bootstrap currents $I_{bs,(1),A}[P, n_0]$ are shown in Fig. 7. Following the 30 kA

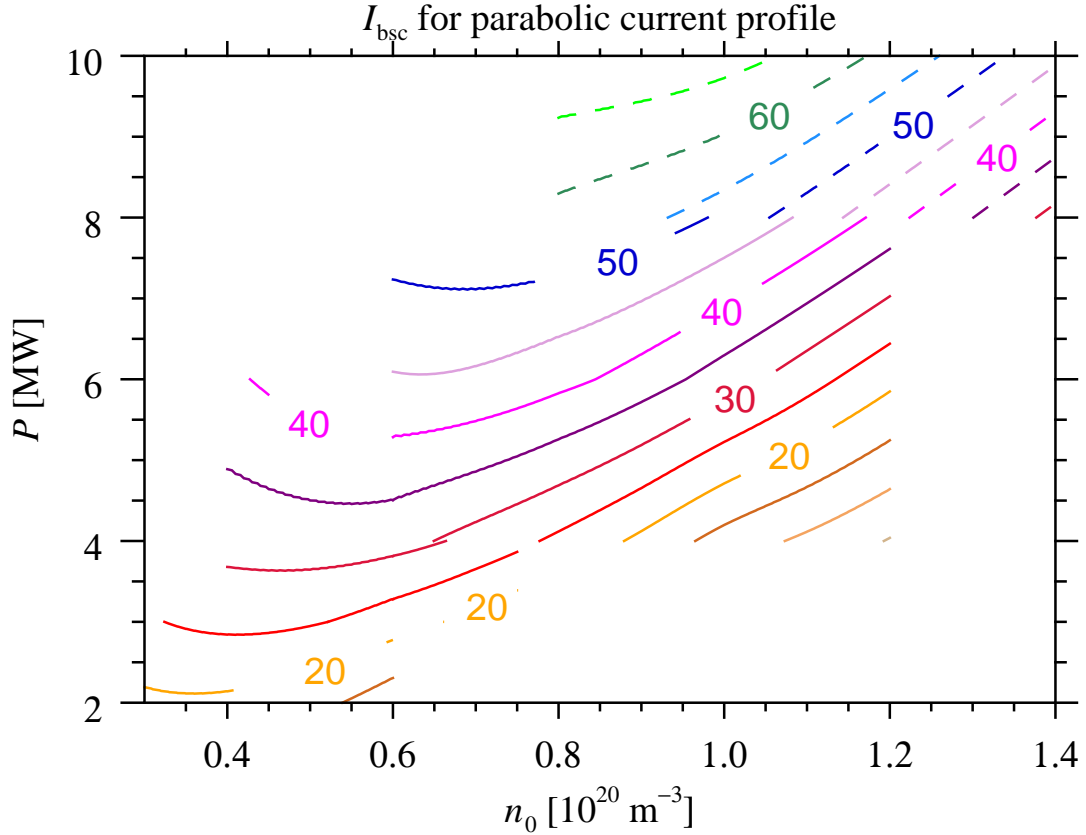


FIG. 7: Bootstrap current after first iteration cycle, starting from a linear pressure profile and a parabolic profile of the toroidal current density. The resulting values ($I_{\text{bs,(1),A}}$) are indicated in kA at the contour levels. The calculations represented by solid (dashed) lines are performed for MHD equilibria with $\langle\beta\rangle$ of 2 (3) %. For the further assumptions, see text.

contour in that figure would outline a path from low power (ca. 4 MW) and low density (0.4 to $0.6 \times 10^{20} \text{ m}^{-3}$) to high power (8 MW) and high density ($1.4 \times 10^{20} \text{ m}^{-3}$) without significant changes in the edge rotational transform.

However, the resulting profiles $p_{(1)}(s)$ and $j_{\text{tor,(1)}}(s)$ are different from the initial profiles and different for each set (P, n_0) . In particular, the current density profiles in the NTSS results show a more centrally peaked form. Since the bootstrap current also depends on the profile of t and different current density profiles result in different t profiles, the hope was to achieve faster convergence with a more educated guess for the profile shapes. Before continuing the iteration (section IV B), we therefore investigate the effect of starting with a

more peaked choice for the profile of the toroidal current density,

$$j_{\text{tor,ini,B}}(s) = j_{c,B} s \{0.1(1-s) + (1-s)^2 + 10(1-s)^8\}, \quad (15)$$

with $j_{c,B}$ again chosen such that the total toroidal current is 30 kA. This function is displayed in Fig. 9 (top panel) as ‘modified profile’. However, a rather similar behaviour of the resulting bootstrap current $I_{\text{bs,(1),B}}[P, n_0]$ is found as that of $I_{\text{bs,(1),A}}$ depicted in Fig. 7. The 30 kA contour is only slightly shifted in the (P, n_0) plane. In Fig. 8, we show the differences $I_{\text{bs,(1),A}} - I_{\text{bs,(1),B}}$ in the (P, n_0) plane.

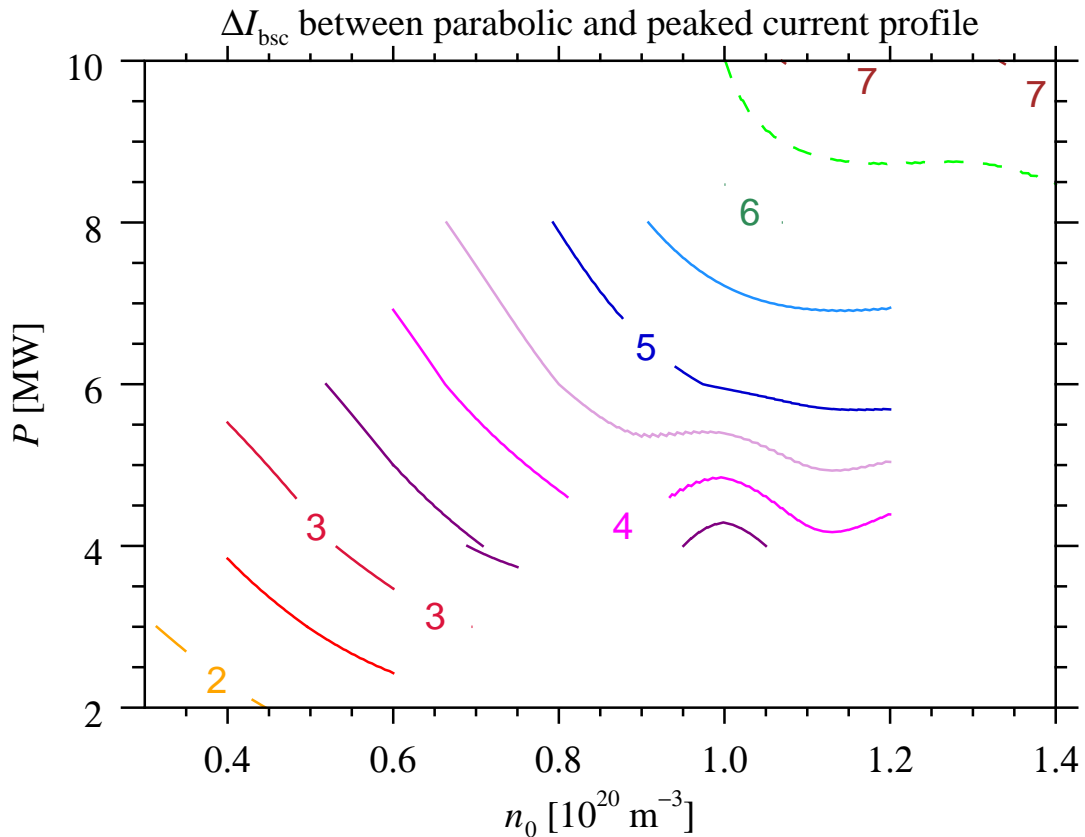


FIG. 8: Differences of the resulting bootstrap current after the first iteration cycle for start with a parabolic and a peaked profile of the toroidal current density, $I_{\text{bs,(1),A}} - I_{\text{bs,(1),B}}$.

B. Convergence of pressure and current density profiles

As described in section III E, for the initial VMEC run, guessed profiles for pressure and toroidal current density, $p_{\text{ini}}(s)$ and $j_{\text{tor,ini}}(s)$, are chosen. After cycling through DKES and

TABLE III: Convergence of successive iterations of the radial profiles $f_{(k)}(s)$, quantified by a measure D_k .

$$D_k \equiv \frac{\sqrt{\sum_{i=1}^N (f_k(s_i) - f_{k-1}(s_i))^2}}{\sqrt{\sum_{i=1}^N (f_{k-1}(s_i))^2}}$$

$f(s)$	$j_{\text{tor}}(s)$	$p(s)$
D_1	0.0311	0.0608
D_2	0.019	0.0026
D_3	0.008	0.0017
D_4	0.0017	0.00015

NTSS, we arrive at a set of different profiles, depending on the P and n_0 values chosen in NTSS. In iteration no. k , each profile has a different total current $I_{\text{bs},(k)}[P, n_0]$. The input profiles for the next iteration for a given value of P is decided based on the value of $I_{\text{bs},(k)}$ closest to 30 kA. As an example, we demonstrate in Fig. 9 the convergence of the profiles for a heating power of $P = 8$ MW. After 4 iterations, the profiles do no longer change visibly (see Table III), and a solution with consistent input and output of the three codes is reached.

C. Path to High Performance

We have computed such self-consistent solutions for heating powers of 2, 4, 6, 8 and 10 MW. From these, we assemble a final path starting at $P = 2$ MW, $n_0 = 0.4 \times 10^{20} \text{ m}^{-3}$ with $I_{\text{bs}} = 17$ kA. Heating power and density are then increased to (4 MW, $0.6 \times 10^{20} \text{ m}^{-3}$) with $I_{\text{bs}} = 28$ kA and to (6 MW, $1.0 \times 10^{20} \text{ m}^{-3}$) with $I_{\text{bs}} = 30$ kA. After this point, I_{bs} can be kept constant by a coordinated increase of heating power and density through (8 MW, $1.31 \times 10^{20} \text{ m}^{-3}$) to (10 MW, $1.57 \times 10^{20} \text{ m}^{-3}$). The whole path is represented in Fig. 10, together with further operation points for each power step from our parameter scan.

The L/R times for the 2 MW and 4 MW cases are 20–40 s. This is the time scale required to achieve the I_{bs} target value. The time scale for changes in the total toroidal current remains on the same order of magnitude for the subsequent plasma states, but the transition

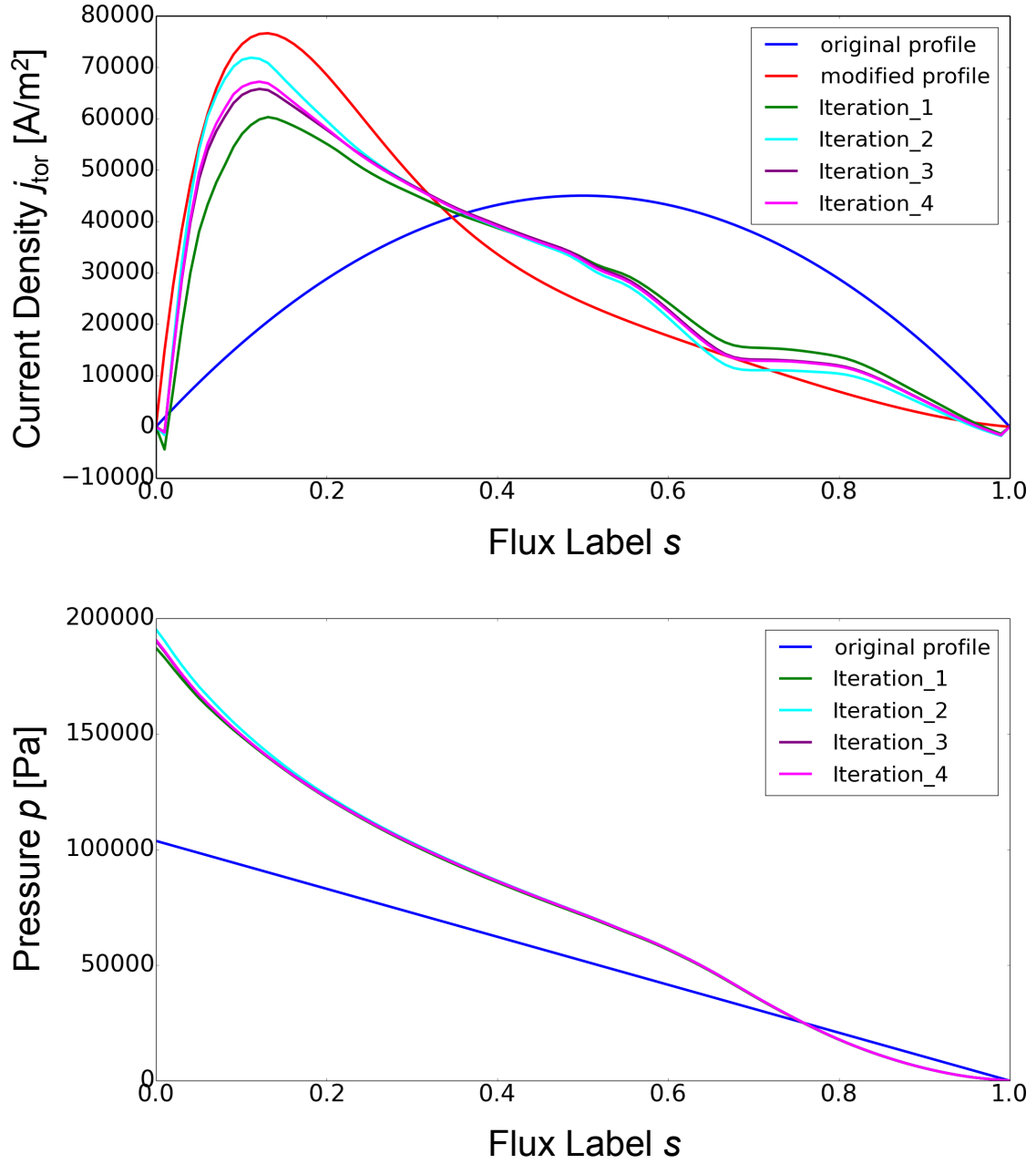


FIG. 9: Evolution of the radial profiles of toroidal current density $j_{\text{tor}}(s)$ (top) and pressure $p(s)$ (bottom) for a heating power of $P = 8$ MW and 4 iteration cycles. The iteration procedure is described in section III E. For $j_{\text{tor}}(s)$ the two initial profiles $j_{\text{tor,ini,A}}$ (“original profile”) and $j_{\text{tor,ini,B}}$ (“modified profile”) are shown.

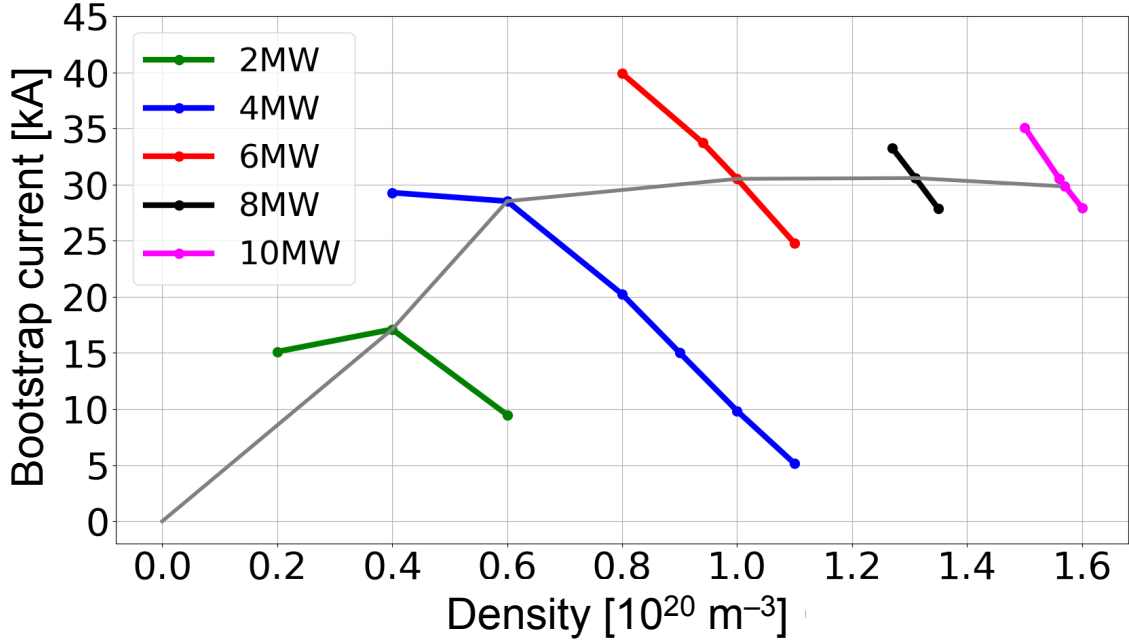


FIG. 10: Stationary values of the bootstrap current on the final path (in grey) from low power and low density toward high power and high density. For each power step, the bootstrap currents for the neighbouring values of the density n_0 are also displayed (in colours).

to higher density and heating power can be faster, since I_{bs} does no longer change. Should the evolution of the discharge deviate from the planned path, sufficient time (L/R) is available to switch off the heating before the island geometry is modified in a way to endanger the sensitive areas of the targets.

For a fixed heating power, the plasma density cannot be chosen arbitrarily high. Due to edge radiation, there exists a density limit n_c in stellarators, which can be described by the empirical scaling

$$n_c = c_{\text{conf}} \frac{P^{0.6}}{f_{\text{imp}}^{0.4}} \quad (16)$$

for W7-X [44]. Here, n_c is the line-averaged electron density in 10^{19} m^{-3} , P is the heating power in MW, f_{imp} is the impurity fraction and c_{conf} is a numerical factor, which depends weakly on the magnetic configuration and is ~ 0.56 in our case. In Table IV we compare the line-averaged densities on our discharge path with n_c for an impurity fraction of 1%, which was found to be a realistic value in the most recent operation phase of W7-X after boronization [44]. We note that the (P, n_0) values on our path do not violate the limit of

TABLE IV: Comparison of line-averaged densities \bar{n} with the empirical density limit n_c of Eq. (16) [44], assuming an impurity fraction of 1%.

P [MW]	2	4	6	8	10
n_0 [10^{19} m^{-3}]	4	6	10	13.1	15.7
\bar{n}/n_c	0.63	0.62	0.81	0.90	0.94

Eq. (16).

The magnetic configurations of the chosen path were provided as full fields (i. e., including the effects of the plasma currents) with VMEC-EXTENDER (see section III B) and were used to calculate the power load to the divertor within the field line diffusion model (see Fig. 11). The predicted heat load for the divertor at the edges near the pumping gap is well within the design limits given in section III F for all five magnetic configurations, since the configuration with a bootstrap current of 17 kA, which places the strike line to an area with reduced design load at the edge of the pumping gap, is passed at low heating power (2 MW, 1/5 the final power).

We also compared the neoclassical transport properties in the lmf regime between the 10 MW configuration of our path and the SE reference configuration proposed in [28] (see Table I for the coil currents). Fig. 12 shows the radial profiles of the effective helical ripple ε_{eff} [37], the corresponding figure of merit, which demonstrates that both configurations have very similar neoclassical transport properties, in the vacuum configuration as well as in the version with finite β and bootstrap current.

V. DISCUSSION

The discharge scenario and the path from plasma startup to a high-performance plasma presented in this work are of a “proof-of-principle” type: The simulations are based on a number of assumptions, like the neglect of radiation effects, the fixed shape of the density profile, or the diffusive ansatz for anomalous transport and the value of the corresponding heat diffusion coefficient.

Whereas neoclassical transport is based on a theory derived from first principles, the anomalous transport in W7-X has still to be explored and described experimentally as

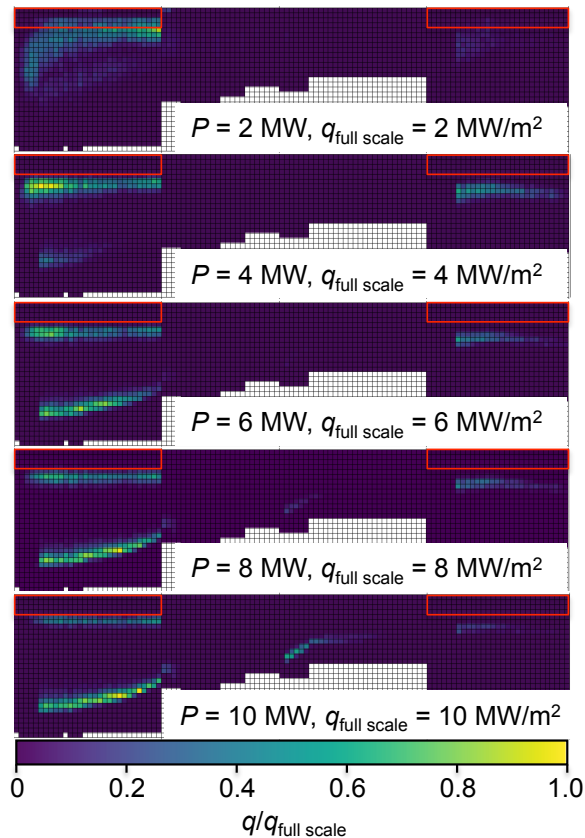


FIG. 11: Heat load patterns on the targets for different heating power P . For each value of P , the magnetic configuration from the last iteration of VMEC+EXTENDER was used for field line diffusion with 20000 starting points. The targets are represented as introduced in Fig. 6, and we have marked the areas with reduced design load by red rectangles. The color scale is adapted to the heating power for a better visibility of the changes in the heat load pattern.

well as theoretically. Although the physical mechanism of anomalous transport is generally believed to be turbulence rather than diffusion, diffusion with an empirical coefficient is widely used in transport codes to simulate it. Our choice, as introduced in section III D, is motivated by observations in various devices, as summarized, e. g., in [45, p. 548] as anomalous electron heat diffusion coefficient $\chi_e = C/n$ with $C = (1 \dots 5) \times 10^{-19} \text{ m}^{-1} \text{ s}^{-1}$. Likewise, on Wendelstein 7-AS, the comparison between measured temperature profiles and predictions of neoclassical transport indicated dominating anomalous heat transport in the edge, and the values of the inferred anomalous heat diffusion coefficients were of the same order of magnitude and they displayed a similar radial dependence [46, section III] as used

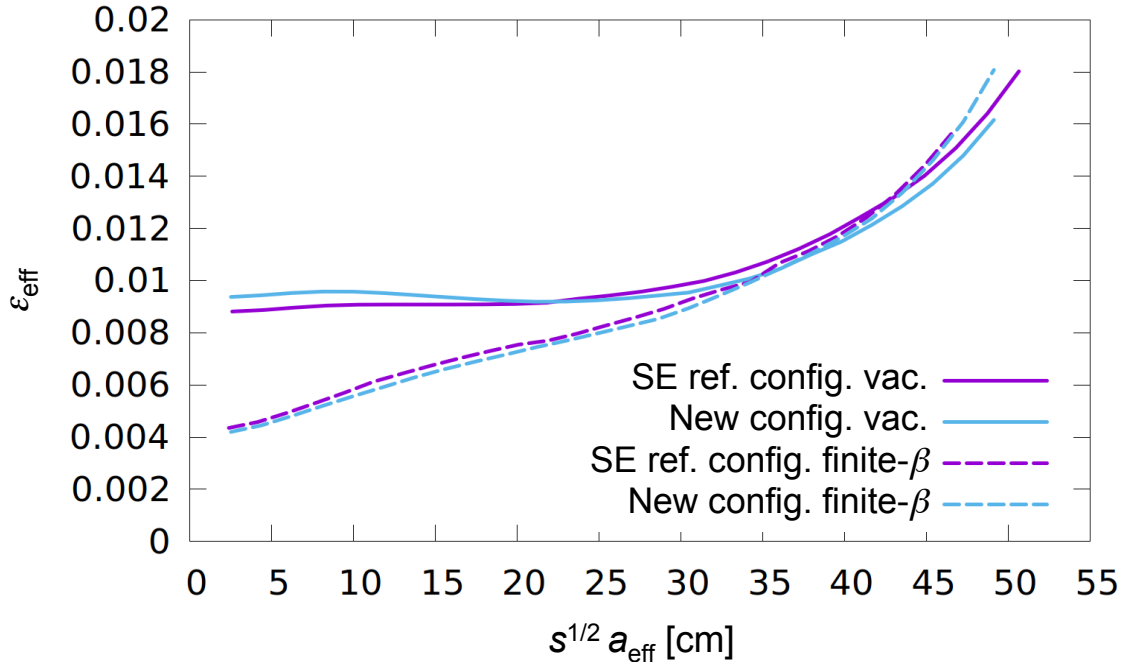


FIG. 12: Comparison of ε_{eff} between the vacuum versions and the finite- β versions of the scraper element (SE) reference configuration with 43 kA bootstrap current and our newly developed configuration with 30 kA bootstrap current and 10 MW heating power. Neither the vacuum versions nor the finite- β versions show significant differences.

in NTSS for the present study.

To investigate the impact of different values of the anomalous transport coefficient, we increased the base value from $1 \text{ m}^2/\text{s}$ to 2 and $4 \text{ m}^2/\text{s}$ in NTSS, without performing the iteration cycle to obtain consistent magnetic configurations with VMEC and new mono-energetic transport coefficients with DKES. As expected, the central temperatures were reduced, and we obtained lower values for the bootstrap current (see Table V). On the other hand side, we find that a bootstrap current of 30 kA can be achieved if the heating power is increased from 8 MW to 10.6 MW (16.8 MW) for $\chi_{\text{base}} = 2 \text{ m}^2/\text{s}$ ($4 \text{ m}^2/\text{s}$) in the case with $n_0 = 1.31 \times 10^{20} \text{ m}^{-3}$, or from 10 MW to 13.9 MW (20.9 MW) for $\chi_{\text{base}} = 2 \text{ m}^2/\text{s}$ ($4 \text{ m}^2/\text{s}$) in the case with $n_0 = 1.57 \times 10^{20} \text{ m}^{-3}$ (all cases of this sensitivity study were calculated in NTSS with the density, magnetic configuration and mono-energetic transport coefficients of the corresponding $1 \text{ m}^2/\text{s}$ base cases, without iterating VMEC, DKES and NTSS for consistency).

Changing the anomalous transport will affect the simulation results in various ways, i. e.,

TABLE V: Variation of the heat diffusion coefficient in NTSS for the two last steps on the path of section IV C (without iterating with VMEC and DKES for consistent results). The

$1/n$ dependance is retained, but the base value at the boundary of the computation domain is chosen as listed. The indirect effect of an increased radial heat diffusivity on the bootstrap current is similar to the effect of an increase in density.

χ_{base} [m ² /s]	$P = 8 \text{ MW}, n_0 =$		$P = 10 \text{ MW}, n_0 =$	
	$1.31 \times 10^{20} \text{ m}^{-3}$		$1.57 \times 10^{20} \text{ m}^{-3}$	
	$\langle\beta\rangle$ [%]	I_{bs} [kA]	$\langle\beta\rangle$ [%]	I_{bs} [kA]
1	3.1	30.6	3.6	30.5
2	2.4	9.3	2.8	6.3
4	1.6	-5.7	1.9	-9.3

via the equilibrium, through the different collisionalities due to changed profiles and the radial electric field, and via the convolution of the mono-energetic transport coefficients with the Maxwellian. It should be noted that the bootstrap current itself is a purely neoclassical effect, without anomalous contribution. The lower value of I_{bs} with increased anomalous transport in our model is therefore entirely due to the indirect effects of reduced temperatures and changed profiles. For the magnetic configuration under consideration, an increase of the anomalous transport at fixed heating power has a similar effect as an increase in plasma density.

We note that the plasma-pressure-induced currents can also modify the magnetic boundary structure by changing the island size. This MHD-equilibrium effect is included in our simulation through the use of the magnetic configurations calculated by VMEC-EXTENDER. For our case with a boundary t of 5/5, the island size increases with increasing β (see, e. g., [47, Fig. 1]). Non-MHD-equilibrium effects influencing the island size (as observed in LHD, leading to island healing [48]) are not considered here, since they apply to internal islands and not to islands intersected by divertor structures.

VI. CONCLUSION

Although W7-X is optimized for low bootstrap current operation, the remaining toroidal net current of up to several 10 kA has long been identified as a challenge for the operation of the island divertor [18]. In particular, the slow evolution of the toroidal net current on the time scale of several tens of seconds may, in a high-power discharge, thermally overload regions of the target plates close to the pumping gap, which are rated for reduced power load. Different solutions to this problem were already proposed, as listed in section II.

In this paper, we demonstrated in a numerical proof-of-principle study, in the framework of neoclassical theory, and assuming a certain amount of anomalous transport, that the dependence of the bootstrap current on the profile shapes can be utilized to keep the bootstrap current constant over a wide range of density and heating power, as long as both quantities are varied in a coordinated way. Even in a magnetic configuration displaying a bootstrap current which needs compensation to allow proper divertor operation, it is thus in principle possible to find a non-critical path toward high plasma density at high heating power without overloading the edges of the targets at the pumping gap, and without the need to feedback-control the superconducting field coils, to apply ECCD, or to install costly additional target components. In contrast to most previous studies of this topic, we verified that our results remain valid if mutually consistent magnetic configurations, neoclassical transport coefficients and radial profiles of plasma pressure and toroidal current density are used. We note that the density values of our path are also below the empirical density limit of W7-X [44] for each power step.

A number of simplifying assumptions were made in this study, such that, in reality, the bootstrap current may assume substantially different values from our simulation results. In particular, a different amount of anomalous transport or different power deposition profiles will modify the radial temperature profiles. Likewise, the density profile may deviate from the assumed fixed shape. Nevertheless, as soon as the model assumptions are adapted appropriately by comparison with observations or by advances in theory, the scheme in itself allows to predict whether such a safe path to a desired operation point exists. A coordinated control of plasma density and heating power, guided by such calculations as ours as well as by experimental experience, could well prove to be an effective and attractive approach to control the bootstrap current evolution, and thereby control the dynamics of

the divertor heat load patterns. A comparison with experimental observations will be the logical continuation of this work.

ACKNOWLEDGEMENTS

We would like to thank S. Bozhenkov for his help with the IPP Webservices and C. D. Beidler for instructive discussions on neoclassical theory.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014–2018 and 2019–2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

* email: priyanjana.sinha@ipp.mpg.de

¹ G. Grieger, C. Beidler, E. Harmeyer, W. Lotz, J. Kießlinger, P. Merkel, J. Nührenberg, F. Rau, E. Strumberger, and H. Wobig, *Fusion Technology* **21**, 1767 (1992).

² H.-S. Bosch *et al.*, *Nucl. Fusion* **53**, 126001 (2013).

³ V. Erckmann, P. Brand, H. Braune, G. Dammertz, G. Gantenbein, W. Kasperek, H. P. Laqua, H. Maassberg, N. B. Marushchenko, G. Michel, M. Thumm, Y. Turkin, M. Weissgerber, A. Weller, W7-X ECRH Team at IPP Greifswald, W7-X ECRH Team at FZK Karlsruhe, and W7-X ECRH Team at IPF Stuttgart, **52**, 291 (2007).

⁴ R. C. Wolf, S. Bozhenkov, A. Dinklage, G. Fuchert, Y. O. Kazakov, H. P. Laqua, S. Marsen, N. B. Marushchenko, T. Stange, M. Zanini, I. Abramovic, A. Alonso, J. Baldzuhn, M. Beurskens, C. D. Beidler, H. Braune, K. J. Brunner, N. Chaudhary, H. Damm, P. Drewelow, G. Gantenbein, Y. Gao, J. Geiger, M. Hirsch, U. Höfel, M. Jakubowski, J. Jelonnek, T. Jensen, W. Kasperek, J. Knauer, S. B. Korsholm, A. Langenberg, C. Lechte, F. Leipold, H. Trimino Mora, U. Neuner, S. K. Nielsen, D. Moseev, H. Oosterbeek, N. Pablant, E. Pasch, B. Plaum, T. Sunn Pedersen, A. Puig Sitjes, K. Rahbarnia, J. Rasmussen, M. Salewski, J. Schilling, E. Scott, M. Stejner, H. Thomsen, M. Thumm, Y. Turkin, F. Wilde, and Wendelstein 7-X Team, *Plasma Phys. & Controlled Fusion* **61**, 014037 (2019).

- ⁵ H. P. Laqua, H. Maassberg, N. B. Marushchenko, F. Volpe, A. Weller, W7-AS Team, W. Kasparek, and ECRH-Group, *Phys. Rev. Lett.* **90**, 075003 (2003).
- ⁶ H. P. Laqua, H. Maassberg, F. Volpe, W7-AS Team, and ECRH-Group, *Nucl. Fusion* **43**, 1324 (2003), <http://stacks.iop.org/NF/43/1324>.
- ⁷ P. McNeely, M. Barlak, J. Baldzuhn, S. Bozhenkov, M. Drevlak, G. Gawlik, B. Heinemann, D. Holtum, J. Jagielski, R. Kairys, R. Nocentini, R. Riedl, P. Rong, N. Rust, R. Schroeder, E. Speth, A. Stähler, A. Turos, and R. Wolf, *Fusion Eng. & Design* **88**, 1034 (2013), Proceedings of the 27th Symposium on Fusion Technology (SOFT-27), Liège, Belgium, September 24–28, 2012.
- ⁸ J. Ongena, A. Messiaen, D. van Eester, B. Schweer, P. Dumortier, F. Durodie, Y. O. Kazakov, F. Louche, M. Vervier, R. Koch, A. Krivska, A. Lyssovian, M. van Schoor, T. Wauters, V. Borsuk, O. Neubauer, O. Schmitz, G. Offermans, Y. Altenburg, C. Baylard, D. Birus, S. Bozhenkov, D. A. Hartmann, J. P. Kallmeyer, S. Renard, R. C. Wolf, and T. Fülöp, **21**, 061514 (2014).
- ⁹ C. Beidler, G. Grieger, F. Herrnegger, E. Harmeyer, J. Kießlinger, W. Lotz, H. Maaßberg, P. Merkel, J. Nührenberg, F. Rau, J. Sapper, F. Sardei, R. Scardovelli, A. Schlüter, and H. Wobig, *Fusion Technology* **17**, 148 (1990).
- ¹⁰ F. Warmer, C. D. Beidler, A. Dinklage, R. Wolf, and The W7-X Team, *Plasma Phys. & Controlled Fusion* **58**, 074006 (2016), Special Issue on the 20th International Stellarator-Heliotron Workshop, Greifswald, Germany, 5–9 October 2015.
- ¹¹ P. Grigull, K. McCormick, J. Baldzuhn, R. Burhenn, R. Brakel, H. Ehmler, Y. Feng, F. Gadelmeier, L. Giannone, D. Hartmann, D. Hildebrandt, M. Hirsch, R. Jaenicke, J. Kisslinger, J. Knauer, R. König, G. Kühner, H. Laqua, D. Naujoks, H. Niedermeyer, N. Ramasubramanian, N. Rust, F. Sardei, F. Wagner, A. Weller, U. Wenzel, and W7-AS Team, *Plasma Phys. & Controlled Fusion* **43**, 175 (2001).
- ¹² K. McCormick, P. Grigull, R. Burhenn, R. Brakel, H. Ehmler, Y. Feng, R. Fischer, F. Gadelmeier, L. Giannone, D. Hildebrandt, M. Hirsch, E. Holzhauser, R. Jaenicke, J. Kisslinger, T. Klinger, S. Klose, J. P. Knauer, R. König, G. Kühner, H. P. Laqua, D. Naujoks, H. Niedermeyer, E. Pasch, N. Ramasubramanian, N. Rust, F. Sardei, F. Wagner, A. Weller, U. Wenzel, and A. Werner, *J. Nucl. Mat.* **313-316**, 1131 (2003), Plasma-Surface Interactions in Controlled Fusion Devices 15, Gifu, Japan, 26–31 May 2002.
- ¹³ Y. Feng, F. Sardei, P. Grigull, K. McCormick, J. Kießlinger, and D. Reiter, *Nucl. Fusion* **46**,

- 807 (2006), <http://stacks.iop.org/NF/46/807>.
- ¹⁴ A. Lumsdaine, T. Bjorholm, J. Harris, D. McGinnis, J. Lore, J. Boscary, J. Tretter, E. Clark, K. Ekici, J. Fellingner, *et al.*, in *2015 IEEE 26th Symposium on Fusion Engineering (SOFE)* (IEEE, 2015) pp. 1–8.
 - ¹⁵ J. Boscary, A. Peacock, R. Stadler, B. Mendelevitch, H. Tittes, J. Tretter, M. Smirnow, and C. Li, **64**, 263 (2013), Proceedings of the Twentieth Topical Meeting on the Technology of Fusion Energy (TOFE-2012) (Part 1), Nashville, Tennessee, August 27–31, 2012.
 - ¹⁶ J. Geiger, C. D. Beidler, M. Drevlak, H. Maaßberg, C. Nührenberg, Y. Suzuki, and Y. Turkin, **50**, 770 (2010), Special Issue: Proceedings of 17th International Stellarator/Heliotron Workshop, October 12–16, 2009, Princeton Plasma Physics Laboratory (PPPL), Princeton, New Jersey, USA (Part II).
 - ¹⁷ A. Dinklage, C. D. Beidler, P. Helander, G. Fuchert, H. Maaßberg, K. Rahbarnia, T. Sunn Pedersen, Y. Turkin, R. C. Wolf, A. Alonso, T. Andreeva, B. Blackwell, S. Bozhnikov, B. Buttenschön, A. Czarnecka, F. Effenberg, Y. Feng, J. Geiger, M. Hirsch, U. Höfel, M. Jakubowski, T. Klinger, J. Knauer, G. Kocsis, A. Krämer-Flecken, M. Kubkowska, A. Langenberg, H. P. Laqua, N. Marushchenko, A. Mollén, U. Neuner, H. Niemann, E. Pasch, N. Pablant, L. Rudischhauser, H. M. Smith, O. Schmitz, T. Stange, T. Szepesi, G. Weir, T. Windisch, G. A. Wurden, D. Zhang, and W7-X Team, *Nature Physics* (2018), <https://doi.org/10.1038/s41567-018-0141-9>.
 - ¹⁸ Y. Turkin, H. Maassberg, C. D. Beidler, J. Geiger, and N. B. Marushchenko, **50**, 387 (2006), Selected Papers from Fifteenth International Stellarator Workshop (Part 2), Madrid, Spain, October 3–7, 2005.
 - ¹⁹ H. Hölbe, T. Sunn Pedersen, J. Geiger, S. Bozhnikov, R. König, Y. Feng, J. Lore, A. Lumsdaine, and Wendelstein 7-X Team, *Nucl. Fusion* **56**, 026015 (2016).
 - ²⁰ A. Lumsdaine, J. Tipton, J. Lore, D. McGinnis, J. Canik, J. Harris, A. Peacock, J. Boscary, J. Tretter, and T. Andreeva, *Fusion Eng. & Design* **88**, 1773 (2013), Proceedings of the 27th Symposium on Fusion Technology (SOFT-27), Liège, Belgium, September 24–28, 2012.
 - ²¹ **51**, 99 (2011), erratum no. 2.
 - ²² J. Geiger, C. D. Beidler, Y. Feng, H. Maaßberg, N. B. Marushchenko, and Y. Turkin, *Plasma Phys. & Controlled Fusion* **57**, 014004 (2015).
 - ²³ M. Zanini, H. P. Laqua, T. Stange, C. Brandt, M. Hirsch, U. Höfel, N. Marushchenko, U. Neuner,

- K. Rahbarnia, J. Schilling, H. Thomsen, R. C. Wolf, and W7-X Team, in *EPJ Web Conf.* (2019) p. 02013, Proceedings of the 20th Joint Workshop on Electron Cyclotron Emission (ECE) and Electron Cyclotron Resonance Heating (ECRH) (EC-20), Greifswald, Germany, 14–17 May 2018, <https://doi.org/10.1051/epjconf/201920302013>.
- ²⁴ M. Murakami, B. A. Carreras, L. R. Baylor, G. L. Bell, T. S. Bigelow, A. C. England, J. C. Glowienka, H. C. Howe, T. C. Jernigan, D. K. Lee, V. E. Lynch, C. H. Ma, D. A. Rasmussen, J. S. Tolliver, M. R. Wade, J. B. Wilgen, and W. R. Wing, *Phys. Rev. Lett.* **66**, 707 (1991).
- ²⁵ D. López-Bruna, J. A. Romero, R. Jiménez-Gómez, M. A. Pedrosa, M. Ochando, T. Estrada, A. López-Fraguas, F. Medina, J. Herranz, T. Kalhoff, E. Ascasíbar, A. de la Peña, and F. Lapayese, *Nucl. Fusion* **49**, 085016 (2009).
- ²⁶ H. Hölbe, *Control of the magnetic topology and plasma exhaust in the edge region of Wendelstein 7-X: A numerical study*, Ph.D. thesis, Ernst-Moritz-Arndt-Universität Greifswald (2015), <https://d-nb.info/1080570160/34>.
- ²⁷ J. D. Lore, M. Cianciosa, H. Frerichs, J. Geiger, H. Hoelbe, J. Boscary, and W7-X Team, *IEEE Trans. Plasma Sci.* **46**, 1387 (2018), Special Issue on Selected Papers from 27th Symposium on Fusion Engineering (SOFE), Princeton, NJ, 2017.
- ²⁸ J. D. Lore, T. Andreeva, J. Boscary, S. Bozhenkov, J. Geiger, J. H. Harris, H. Hoelbe, A. Lumsdaine, D. McGinnis, A. Peacock, and J. Tipton, *IEEE Trans. Plasma Sci.* **42**, 539 (2014).
- ²⁹ S. P. Hirshman and W. I. van Rij, *Comp. Phys. Comm.* **43**, 143 (1986).
- ³⁰ M. Drevlak, D. Monticello, and A. Reiman, *Nucl. Fusion* **45**, 731 (2005).
- ³¹ W. I. van Rij and S. P. Hirshman, *Phys. Fluids B: Plasma Phys.* **1**, 563 (1989).
- ³² Y. Turkin, C. D. Beidler, H. Maaßberg, S. Murakami, V. Tribaldos, and A. Wakasa, **18**, 022505 (2011).
- ³³ S. A. Bozhenkov, J. Geiger, M. Grahl, J. Kießlinger, A. Werner, and R. C. Wolf, *Fusion Eng. & Design* **88**, 2997 (2013).
- ³⁴ V. D. Shafranov and L. E. Zakharov, *Nucl. Fusion* **12**, 599 (1972).
- ³⁵ J. D. Hanson, *Plasma Phys. & Controlled Fusion* **57**, 115006 (2015).
- ³⁶ S. P. Hirshman, K. C. Shaing, W. I. van Rij, C. O. Beasley Jr., and E. C. Crume Jr., *Phys. Fluids* **29**, 2951 (1986).
- ³⁷ C. D. Beidler, K. Allmaier, M. Y. Isaev, S. V. Kasilov, W. Kernbichler, G. O. Leitold, H. Maaßberg, D. R. Mikkelsen, S. Murakami, M. Schmidt, D. A. Spong, V. Tribaldos, and

- A. Wakasa, Nucl. Fusion **51**, 076001 (2011).
- ³⁸ H. Maaßberg, C. D. Beidler, and Y. Turkin, **16**, 072504 (2009).
- ³⁹ M. Hirsch, J. Baldzuhn, C. Beidler, R. Brakel, R. Burhenn, A. Dinklage, H. Ehmler, M. Endler, V. Erckmann, Y. Feng, J. Geiger, L. Giannone, G. Grieger, P. Grigull, H.-J. Hartfuß, D. Hartmann, R. Jaenicke, R. König, H. P. Laqua, H. Maaßberg, K. McCormick, F. Sardei, E. Speth, U. Stroth, F. Wagner, A. Weller, A. Werner, H. Wobig, S. Zoletnik, and W7-AS Team, Plasma Phys. & Controlled Fusion **50**, 053001 (2008).
- ⁴⁰ N. B. Marushchenko, Y. Turkin, and H. Maassberg, Comp. Phys. Comm. **185**, 165 (2014).
- ⁴¹ P. Sinha, H. Hölbe, T. S. Pedersen, S. Bozhnikov, and W7-X Team, Nucl. Fusion **58**, 016027 (2018).
- ⁴² D. Böckenhoff, M. Blatzheim, and W7-X Team, “Application of improved analysis of convective heat loads on plasma facing components to Wendelstein 7-X,” (2019), Nucl. Fusion, in press, <https://doi.org/10.1088/1741-4326/ab201e>.
- ⁴³ R. Brakel, M. Köppen, A. Peacock, A. Werner, and M. Jakubowski, *Specification of Design Loads for In-vessel Components of W7-X*, Tech. Rep. (Max-Planck-Institut für Plasmaphysik, Greifswald, 2011) W7-X report no. 1-AC-S0005.0.
- ⁴⁴ G. Fuchert, K. J. Brunner, K. Rahbarnia, T. Stange, D. Zhang, J. Baldzuhn, S. A. Bozhnikov, C. D. Beidler, M. N. A. Beurskens, S. Brezinsek, R. Burhenn, H. Damm, A. Dinklage, Y. Feng, M. Hirsch, Y. Kazakov, J. Knauer, A. Langenberg, H. P. Laqua, S. Lazerson, N. A. Pablant, E. Pasch, T. Sunn Pedersen, E. R. Scott, F. Warmer, V. R. Winters, R. C. Wolf, and W7-X Team, in *Proceedings of the 27th IAEA Fusion Energy Conference, 22–27 October 2018, Gandhinagar, India* (IAEA, Vienna, 2019) pp. 286–291, paper IAEA-EX/3-5, https://conferences.iaea.org/indico/event/151/papers/5833/files/4764-iaea_template_fuchert.pdf, submitted to Nucl. Fusion.
- ⁴⁵ P. C. Liewer, Nucl. Fusion **25**, 543 (1985).
- ⁴⁶ H. Maaßberg, C. D. Beidler, U. Gasparino, M. Romé, and W7-AS Team, **7**, 295 (2000).
- ⁴⁷ Y. Suzuki, J. Geiger, M. Drevlak, and J. Nührenberg, in *33rd EPS Conference on Plasma Physics (Rome), 19 - 23 June 2006*, Europhysics Conference Abstracts, Vol. 30I, edited by F. De Marco and G. Vlad (European Physical Society, 2006) paper P-2.119, <http://epsppd.epfl.ch/Roma/pdf/P2.119.pdf>.
- ⁴⁸ Y. Narushima, S. Sakakibara, Y. Suzuki, K. Y. Watanabe, S. Ohdachi, Y. Takemura, M. Yosh-

inuma, K. Ida, F. Castejón, D. López-Bruna, C. C. Henga, M. Kobayashi, H. Tanaka, T. Akiyama, N. Ohno, and LHD Experimental Group, Nucl. Fusion **57**, 076024 (2017), <https://doi.org/10.1088/1741-4326/aa6dce>.

Daniel Böckenhoff

Personal Data

Name Daniel Böckenhoff
Date of Birth 20.10.1988
Place of Birth Detmold, Germany
Nationality German

Education

Studies



PhD, *Max Planck Institute for Plasma Physics (Greifswald branch) and University of Greifswald*

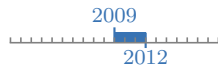
- Member of the International Helmholtz Graduate School for Plasma Physics
- Invited Talk at 3rd IAEA Technical Meeting on Fusion Data Processing, Validation and Analysis



Master of science, *RWTH Aachen University*, Particle Physics and Astroparticle Physics (1.1):

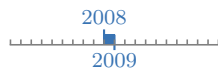
Master thesis: »Untersuchung der prompten Gamma-Emission bei der Protonentherapie«

- Named to the Dean's List for outstanding academic achievement



Bachelor of science, *RWTH Aachen University*, Physics:

Bachelor thesis: »Bestimmung von Triggereffizienzen von Di-Lepton Ereignissen am CMS-Experiment bei $\sqrt{s} = 8 \text{ TeV}$ «



Remote studies Physics, *TU Kaiserslautern*, »Früheinstieg in das Physikstudium« (FiPS) program for persons in civil service



General qualification for university entrance, *Christian-Dietrich-Grabbe-Gymnasium Detmold*, Main courses::

Math, Physics

Extracurricular



Seminar, 643. *WE-Heraeus-Seminar* on »Impact of 3D magnetic fields on hot plasmas«



Summer School, *IPP summer university for plasma physics*



Summer School, *BND summer school on high energy physics*



Internship, *HIMA Paul Hildebrandt GmbH*, All departments



Internship, *Phoenix Contact*, High voltage protection

Experience

Teaching

2014

2015

2017

2018

Lab courses, *elementary particle physics*

Lab courses, *statistical physics, thermodynamics, mechanics and electrodynamics*

2017

2018

Internship supervision, *Max Planck Institute for Plasma Physics (Greifswald branch)*, Neural network application for pattern recognition

2016

2016

Public talks and visitor Guide, *Max Planck Institute for Plasma Physics (Greifswald branch)*, The nuclear fusion experiment Wendelstein 7-X

Work

2015

2016

Research scientist, *RWTH Aachen University*, III Physical Institute B, Double Chooz experiment:

Responsibilities: Planing and installation of a clean room and muon background analysis and suppression by installation of a veto scintillation detector

2008

2009

Civil service, *St. Elisabeth Stiftung*, Caretaking of children with alleged behavioral problems

Miscellaneous

2019

2018

Member of the core team of Scientists for Future Greifswald

Peer reviewer for Nuclear Fusion (IOP)

Publications

- [1] M. BLATZHEIM and D. BÖCKENHOFF. »Neural Network Regression Approaches to Reconstruct Plasma Properties from Wendelstein 7-X Heat Load Images«. In: *Accepted for publication*, (2019).
- [2] P. SINHA et al. »Scenario with Combined Density and Heating Control to Reduce the Impact of Bootstrap Current in Wendelstein 7-X«. In: *Nuclear Fusion*, (2019). DOI: 10.1088/1741-4326/ab37bb.
- [3] D. BÖCKENHOFF and M. BLATZHEIM. »Application of improved analysis of convective heat loads on plasma facing components to Wendelstein 7-X«. In: *Nuclear Fusion*, Vol. 59.8 (2019), page 086031. DOI: 10.1088/1741-4326/ab201e.
- [4] M. BLATZHEIM et al. »Neural network performance enhancement for limited nuclear fusion experiment observations supported by simulations«. In: *Nuclear Fusion*, Vol. 59.1 (2019), page 016012. DOI: 10.1088/1741-4326/aaefaf.
- [5] D. BÖCKENHOFF et al. »Reconstruction of Magnetic Configurations in W7-X using Artificial Neural Networks«. In: *Nuclear Fusion*, Vol. 58.5 (2018), page 056009. DOI: 10.1088/1741-4326/aab22d.
- [6] A. CHROBAK et al. »Comparison of Various Models of Monte Carlo Geant 4 Code in Simulations of Prompt Gamma Production«. In: *Acta Physica Polonica B*, Vol. 48.3 (2017), page 675. DOI: 10.5506/APhysPolB.48.675.
- [7] K. RUSIECKA et al. »Determination of Gamma Angular Distribution from the Shape of Spectral Line for the First Excited State of Carbon Nucleus«. In: *World Journal of Nuclear Science and Technology*, Vol. 06.01 (2016), pages 63–70. DOI: 10.4236/wjnst.2016.61006.
- [8] A. WRONSKA et al. »Gamma emission in hadron therapy — Towards new tools of quality assurance«. In: *2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE, 2015, pages 1–3. DOI: 10.1109/NSSMIC.2015.7582186.
- [9] A. WRONSKA et al. »Gamma Emission in Hadron Therapy — Experimental Approach«. In: *Acta Physica Polonica B*, Vol. 46.3 (2015), page 753. DOI: 10.5506/APhysPolB.46.753.



Acknowledgements

This work was not possible without the help of many people. First and foremost I would like to express my gratitude to Prof. Thomas Sunn-Pedersen for giving me the wonderful opportunity to work in such an important field with a great amount of freedom in my research. Very special thanks go to Michael Endler, Hauke Hölbe and Joris Fellingner for intensive discussions, constructive criticism and advice. I greatly appreciate the fruitful collaboration with my colleague Marko Blatzheim. An important thing about constructive work are friends who lend an ear, advice, and help where needed. Such I would like to express my special thanks to Lukas Rudischhauser and Udo Höfel with whom I could also discuss very technical questions. I highly appreciate the very open and positive work environment at IPP and could write a personalised appreciation to all my friends and most colleagues.

Besides work, I would like to mention my wife Anna who is the best person to have as a partner and to my parents. Without their support I couldn't have come to where I am now.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Statutory declaration

Hiermit erkläre ich, dass diese Arbeit bisher von mir weder an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Greifswald noch einer anderen wissenschaftlichen Einrichtung zum Zwecke der Promotion eingereicht wurde. Ferner erkläre ich, dass ich diese Arbeit selbstständig verfasst und keine anderen als die darin angegebenen Hilfsmittel und Hilfen benutzt und keine Textabschnitte eines Dritten ohne Kennzeichnung übernommen habe.

Greifswald, 20.09.2019

DANIEL BÖCKENHOFF