

PAPER • OPEN ACCESS

# Accelerated Bayesian inference of plasma profiles with self-consistent MHD equilibria at W7-X via neural networks

To cite this article: A. Merlo *et al* 2023 *JINST* **18** P11012

View the [article online](#) for updates and enhancements.

# Accelerated Bayesian inference of plasma profiles with self-consistent MHD equilibria at W7-X via neural networks

A. Merlo,<sup>a,\*</sup> A. Pavone,<sup>a</sup> D. Böckenhoff,<sup>a</sup> E. Pasch,<sup>a</sup> G. Fuchert,<sup>a</sup> K.J. Brunner,<sup>a</sup> K. Rahbarnia,<sup>a</sup> J. Schilling,<sup>a</sup> U. Höfel,<sup>a</sup> S. Kwak,<sup>a</sup> J. Svensson,<sup>a</sup> T.S. Pedersen<sup>b</sup> and the W7-X team<sup>1</sup>

<sup>a</sup>Max-Planck-Institute for Plasma Physics,  
17491 Greifswald, Germany

<sup>b</sup>Type One Energy Group,  
Madison, WI, 53703, U.S.A.

E-mail: [andrea.merlo@ipp.mpg.de](mailto:andrea.merlo@ipp.mpg.de)

**ABSTRACT:** High- $\beta$  operations require a fast and robust inference of plasma parameters with a self-consistent magnetohydrodynamic (MHD) equilibrium. Precalculated MHD equilibria are usually employed at Wendelstein 7-X (W7-X) due to the high computational cost. To address this, we couple a physics-regularized artificial neural network (NN) model that approximates the ideal-MHD equilibrium with the Bayesian modeling framework Minerva. We show the fast and robust inference of plasma profiles (electron temperature and density) with a self-consistent MHD equilibrium approximated by the NN model. We investigate the robustness of the inference across diverse synthetic W7-X plasma scenarios. The inferred plasma parameters and their uncertainties are compatible with the parameters inferred using the variational moments equilibrium code (VMEC), and the inference time is reduced by more than two orders of magnitude. This work suggests that MHD self-consistent inferences of plasma parameters can be performed between shots.

**KEYWORDS:** Analysis and statistical methods; Nuclear instruments and methods for hot plasma diagnostics

ARXIV EPRINT: [2305.12889](https://arxiv.org/abs/2305.12889)

<sup>1</sup>See Pedersen et al. 2022 (<https://dx.doi.org/10.1088/1741-4326/ac2cf5>) for the W7-X Team author list.

\*Corresponding author.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Methods</b>	<b>2</b>
2.1	Bayesian analysis	2
2.2	Model graph	3
2.2.1	Prior distributions and forward models	3
2.3	Inference	6
<b>3</b>	<b>Results</b>	<b>6</b>
3.1	2D inference on synthetic data	7
3.2	Inference robustness on synthetic data	8
3.3	Inference of improved confinement scenarios at W7-X	9
<b>4</b>	<b>Conclusions</b>	<b>12</b>
<b>5</b>	<b>Author statement</b>	<b>14</b>
<b>A</b>	<b>Prior distributions of model parameters</b>	<b>15</b>
<b>B</b>	<b>Modeling the uncertainty due to the TS laser misalignment</b>	<b>16</b>

---

## 1 Introduction

Bayesian inference is the statistical identification of model parameters that are consistent with given observed data using Bayes theorem. In a magnetic confinement experiment, the model parameters describe the plasma state, which often includes, but is not limited to the plasma profiles (e.g., temperature and density) and an ideal-magnetohydrodynamic (MHD) equilibrium. Diagnostics raw measurements provide the observed data.

In case of a high  $\beta$  (the ratio between the plasma kinetic to the magnetic pressure) scenario or a large internal plasma current, the equilibrium may differ from the vacuum equilibrium. To correctly infer the plasma state, a self-consistent, finite-beta ideal-MHD equilibrium is required. Wendelstein 7-X (W7-X) is an optimized stellarator that features a stiff equilibrium and a low bootstrap current [1, 2] (the expected bootstrap current at high- $\beta$  in the standard configuration is  $\simeq 80$  kA [3]), therefore, equilibrium changes caused by the plasma internal current are predicted to be small with respects to an equivalent tokamak. Finite-beta effects, however, have been experimentally observed, and they can still modify the vacuum equilibrium [4].

In tokamaks, data analysis workflows regularly include the reconstruction of a self-consistent ideal-MHD equilibrium [5–7]. However, this is not the case with stellarators. The inference of 3D ideal-MHD equilibria have been already demonstrated in several non-axisymmetric experiments,

namely the Compact Toroidal Hybrid (CTH) [8, 9], the National Compact Stellarator Experiment (NCSX) [10], the Helically Symmetric Experiment (HSX) [11–13], Wendelstein 7-AS (W7-AS) [14], and W7-X [4, 15, 16]. Moreover, multiple frameworks exist that allow the reconstruction of 3D ideal-MHD equilibria: V3FIT [17, 18], STELLOPT [19], and Minerva [20, 21]. However, because of the computational cost of calculating 3D ideal-MHD equilibria, equilibrium reconstruction is not routinely used in Bayesian inference procedures and data analysis pipelines. Vacuum fields, or precalculated finite-beta approximations are used instead [16].

Data-driven approaches have been proposed to reduce the computational cost of constructing stellarator equilibria. Function parametrization was used at W7-AS [14, 22] and at W7-X [23–25]. More recently, [26] proposed a physics-regularized artificial neural network (NN) to approximate the ideal-MHD solution operator in W7-X configurations. Adopting machine learning (ML) models to approximate experimental equilibria can drastically accelerate sample-intensive applications (e.g., Bayesian inference) that require the computation of a MHD equilibrium.

However, how does the adoption of MHD equilibria approximated by NN models affect the inferred plasma parameters? In this paper, we show the fast ( $\mathcal{O}(10^2\text{s})$ ) and robust inference of plasma profiles (electron temperature and density) based on the Thomson scattering (TS) [27, 28] and single channel dispersion interferometer (DI) [29] diagnostics with a self-consistent MHD equilibrium. The self-consistent MHD equilibrium is approximated by the physics-regularized NN model proposed in [26]. The NN model allows the efficient exploration of the full posterior distribution of plasma profiles with self-consistent MHD equilibria, yielding an estimation of the profile uncertainties. Specifically, we:

- couple the physics-regularized NN model from [26] in the Bayesian framework Minerva (section 2);
- investigate the robustness of the inferred plasma parameters across diverse W7-X synthetic plasma scenarios (sections 3.1 and 3.2);
- compare the reconstructed experimental plasma parameters and their uncertainties (estimated with Markov chain Monte Carlo (MCMC)) with three different MHD equilibrium models (section 3.3): a fixed, finite-beta variational moments equilibrium code (VMEC) equilibrium; a self-consistent VMEC equilibrium; and a self-consistent NN equilibrium

## 2 Methods

### 2.1 Bayesian analysis

In a model-based simulation, a mechanistic model, the model parameters  $H$ , and observed data  $D$  characterize the phenomenon under study. In Bayesian analysis, the model parameters and observations are fully captured by their joint probability distribution:

$$P(H, D) = P(D|H)P(H), \quad (2.1)$$

where the prior distribution  $P(H)$  describes the prior knowledge on the model parameters before taking any observations into account (e.g., plasma density should be positive), and the conditional

distribution  $P(D|H)$  denotes the probability of the observed data  $D$  given the model parameters  $H$  (when seen as a function of  $H$  for given  $D$ ,  $P(D|H)$  is it also known as the likelihood function).

A parametric distribution function (e.g., Gaussian) usually represents  $P(D|H)$ . The likelihood function is a function that captures the physical model under consideration, and it is often referred to as the forward model. Experimentally derived uncertainties inform the distribution variance.

Bayesian inference updates the model parameters prior distribution to their posterior distribution given the observed data using Bayes formula:

$$P(H|D) = \frac{P(H, D)}{P(D)} = \frac{P(D|H)P(H)}{P(D)}, \quad (2.2)$$

where the *posterior* distribution  $P(H|D)$  describes the probability of the model parameters given the observed data, and  $P(D)$  is the *evidence*, a normalization constant that describes the probability of the observed data given all possible values of the model parameters:

$$P(D) = \int P(D|H)P(H) dH. \quad (2.3)$$

The posterior distribution of the model parameters describes the degree of belief we have in the parameter values, which is based on prior knowledge and updated by the observed data. It encapsulates all the information and uncertainties we have on the model parameters.

## 2.2 Model graph

The Bayesian modeling framework Minerva [20, 21] is used in this work. In Minerva, a graph describes the relationship between model parameters and observed data, which are both represented as the graph's nodes. The graph's edges represent the probabilistic relationships between the modeled quantities. At W7-X, several diagnostics are already implemented in Minerva, making it the obvious choice to validate the accelerated inference of plasma profiles with the NN model.

Figure 1 shows a simplified picture of the Minerva graph used in this work. Blue denotes the model free parameters, and orange denotes the observed quantities. A MHD equilibrium maps the 1D electron temperature and density profiles to a 3D position of magnetic flux surfaces in space. The TS [27, 28] and the single channel DI [29] diagnostics are used to constrain the plasma state: the TS provides information on the electron density and temperature, while the DI measures the line-integrated electron density along a single line of sight (LOS).

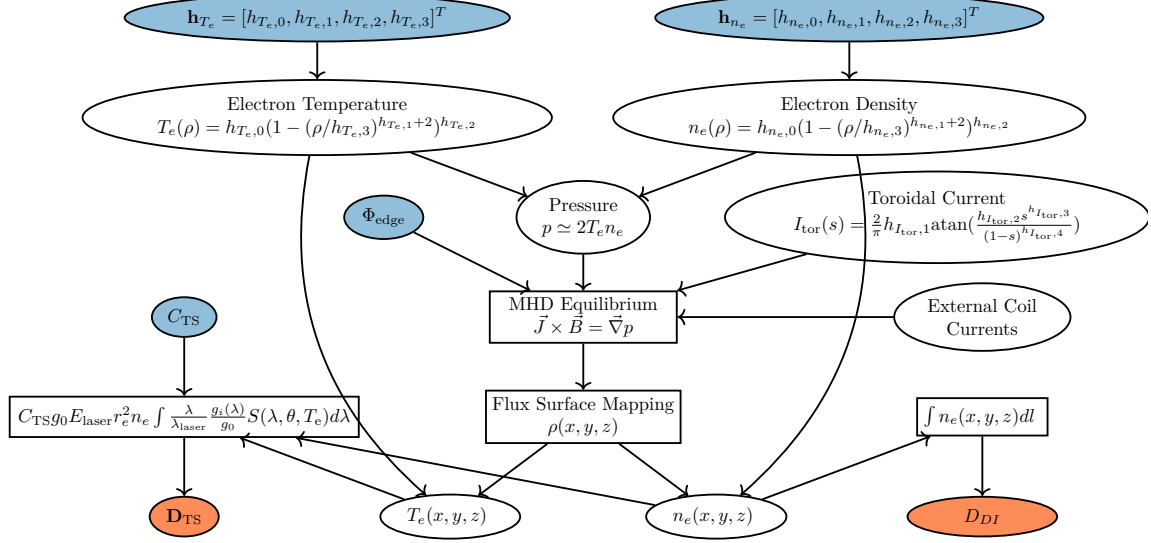
### 2.2.1 Prior distributions and forward models

Modified two power profiles parametrize the electron density and temperature:

$$n_e(\rho) = h_{n_e,0}(1 - (\rho/h_{n_e,3})^{h_{n_e,1+2}})^{h_{n_e,2}}, \quad (2.4)$$

$$T_e(\rho) = h_{T_e,0}(1 - (\rho/h_{T_e,3})^{h_{T_e,1+2}})^{h_{T_e,2}}, \quad (2.5)$$

where  $\rho = \sqrt{\Phi/\Phi_{\text{edge}}}$  is the square root of the normalized toroidal flux. The additional (with respect the standard two power parametrization) parameters  $h_{T_e,3}$  and  $h_{n_e,3}$  allow modeling not-null temperatures and densities beyond the last closed flux surface (LCFS), which are usually observed at W7-X [30]. Uniform distributions represent the prior on the plasma profile parameters (see appendix A).



**Figure 1.** A simplified sketch of the Minerva graph. Blue ellipses represent free parameters, orange ellipses represent observed quantities, white ellipses represent model fixed parameters or model internal quantities, and white boxes represent physical forward models.

Flux surfaces  $\rho(x, y, z)$  computed from a free-boundary ideal-MHD equilibrium map the 1D profiles (e.g.,  $n_e(\rho)$ ) to the 3D real space geometry ( $n_e(x, y, z)$ ). The set of W7-X coil currents fixed to their setpoint values, the net toroidal magnetic flux  $\Phi_{\text{edge}}$  represented by a uniform prior distribution, and the 1D pressure and toroidal current profiles parametrize the free-boundary ideal-MHD equilibrium. In this work, we assume  $T_i \simeq T_e$ ,  $Z_{\text{eff}} \simeq 1$ , and  $n_i \simeq n_e = n$ . As a result, the pressure is simply given by  $p \simeq 2n_e T_e$ . The purpose of this paper is to investigate the use of NN equilibria in a Bayesian inference framework, rather than providing a highly accurate description of W7-X (e.g., by including a X-ray imaging crystal spectroscopy (XICS) to infer the ion temperature, or spectrometers to infer the plasma effective ion charge  $Z_{\text{eff}}$ ). Therefore, the relaxation of these assumptions is not in the scope of this paper.

The NN model sees only the total pressure  $p$  and not  $T_e$  or  $T_i$ . As a result, for the scope of this work, it is not critical how the pressure profile is constructed. The assumptions only limit the set of experimental scenarios that can be used to test the inference procedure. Indeed, the experimental scenarios considered in section 3.3 feature  $T_i \simeq T_e$  within experimental uncertainties. However, in section 3.2, we use synthetic data to investigate the robustness of the inference procedure across diverse, randomly sampled (within the model parameter prior distributions) W7-X scenarios. The toroidal current profile is parametrized as:

$$I_{\text{tor}}(s) = \frac{2}{\pi} h_{I_{\text{tor}},1} \text{atan} \left( \frac{h_{I_{\text{tor}},2} s^{h_{I_{\text{tor}},3}}}{(1-s)^{h_{I_{\text{tor}},4}}} \right), \quad (2.6)$$

where  $s = \rho^2$ . However, given the lack of diagnostics to constrain its profile [25], the profile shape is fixed:  $h_{I_{\text{tor}},2} = h_{I_{\text{tor},3}} = h_{I_{\text{tor},4}} = 1$ . Moreover, given that  $I_{\text{tor}}(s = 1) = h_{I_{\text{tor},1}}$ ,  $h_{I_{\text{tor},1}}$  is held fixed to the net toroidal current as measured by the Rogowski coil [31].

This work considers two MHD equilibrium models: VMEC [32] and a physics-regularized NN [26]. The NN model has been integrated into Minerva with the ONNX Runtime framework [33]. In the proposed Minerva graph, the equilibrium model is employed only to provide the self-consistent flux surface mapping. In this regard, [26] reports that the NN model achieves an average flux surface error of  $\approx 0.6$  mm. The NN model was trained on a large set of W7-X magnetic configurations, including the reference W7-X configurations [34], pressure profiles with  $\langle\beta\rangle$  up to 5%, and toroidal current profiles with  $I_{\text{tor}}$  up to 20 kA. W7-X experimental conditions in previous operational campaigns are well within the training data set [35]. For further information on the model architecture, training, and evaluation, please refer to [26].

The TS observations constrain the electron density and temperature. The TS system's forward model maps  $n_e(x, y, z)$  and  $T_e(x, y, z)$  to the number of scattered photons on a set of volumes along the TS laser beam path. Polychromator filters are used to measure the scattered spectrum by selecting distinct spectral intervals. The time integral of the signal of each channel over the length of the scattered laser pulse corresponds to the number of the scattered photons [28]. Thomson scattering scatters laser photons off of electrons, and the Doppler effect broadens the Thomson scattered spectra due to the thermal motions of the electrons. The electron temperature determines the width of the Thomson scattered spectrum, and electron density determines its amplitude. The predicted TS signal of the  $i$ -th channel is [28]:

$$D_{\text{TS}}^i = C_{\text{TS}} g_0 E_{\text{laser}} r_e^2 n_e \int \frac{\lambda}{\lambda_{\text{laser}}} S(\lambda, \theta, T_e) d\lambda, \quad (2.7)$$

where the integration is over the scattered spectrum wavelength  $\lambda$ .  $g_0$  is the absolute sensitivity of a reference channel,  $E_{\text{laser}}$  is the laser pulse energy,  $r_e$  is the electron radius,  $\lambda_{\text{laser}}$  is the laser wavelength,  $g_i(\lambda)$  is the absolute sensitivity of the  $i$ -th channel,  $\theta$  is the scattering angle, and  $S(\lambda, \theta, T_e)$  is the spectral density function. The  $g_i(\lambda)/g_0$  factor represents the relative calibration of each channel, and the  $C_{\text{TS}}$  is the global calibration factor. For a more detailed description of the TS system, please refer to [27, 28]. The forward model of the DI system maps the 3D electron density to the line-integrated value along the DI LOS. The predicted DI signal is:

$$D_{\text{DI}} = \int n_e(x, y, z) dl, \quad (2.8)$$

where the integration path  $\int dl$  is along the DI LOS. For a more detailed description of the DI system, please refer to [29].

The misalignment of the laser system is a common issue that affects TS systems. In a TS system, observation optics with LOSs that cross the laser beam path capture the scattered light. A misaligned laser system can cause the laser beam path to be outside the nominal scattering volumes; in such a case, the inferred density will be lower than the actual plasma density.

At W7-X, the TS system is believed to be affected by a systematic shift of the laser beam [36, 37]. Laser misalignment is expected to be the dominant source of error in the electron density profiles, affecting both the overall scale and shape of the profile.

In this work, we introduce two additional parameters to account for laser misalignment errors: the global TS calibration factor  $C_{\text{TS}}$ , and a systematic uncertainty  $\sigma_{\text{laser}}$  that is added to the nominal TS uncertainty.  $C_{\text{TS}}$  is a model free parameter, and it is constrained during inference by the

line-integrated density of the DI (i.e., the TS is cross-calibrated during inference). The systematic uncertainty due to laser misalignment is held fixed, and the standard deviation of the  $i$ -th TS channel is modeled as:

$$(\sigma_{\text{TS}}^i)^2 = (\tilde{\sigma}_{\text{TS}}^i)^2 + (\sigma_{\text{laser}}^i)^2, \quad (2.9)$$

$$\alpha_{\text{laser}}^i = \max(\sigma_0, \alpha D_{\text{TS}}^i), \quad (2.10)$$

where  $\tilde{\sigma}_{\text{TS}}^i$  is the nominal TS uncertainty.  $\alpha$  and  $\sigma_0$  are experimentally estimated (see appendix B). Ongoing efforts at W7-X are in place to account for the TS laser misalignment. A comprehensive modeling and exhaustive analysis of this error is beyond the scope of this paper

### 2.3 Inference

The inference procedure is divided in two stages. Firstly, in the so called maximum a posteriori (MAP) optimization, the model free parameters that maximizes the left-hand side of equation (2.2), i.e., the mode of the posterior distribution, are derived:

$$H_{\text{MAP}} = \text{argmax}_H P(H|D). \quad (2.11)$$

The Hooke and Jeeves optimizer [38] drives the optimization. Secondly, the MCMC Metropolis-Hastings algorithm (MHA) [39, 40] numerically approximates the full posterior distribution. The model parameter uncertainties can be derived from the MCMC samples. The inference is carried out on a single core of an Intel(R) Xeon(R) Gold 6136 CPU @ 3 GHz.

We are interested in generating samples from the posterior distribution of the model parameters because they allow propagating their uncertainties into any subsequent quantities. The computation of the samples is not straightforward: either a closed form of the posterior is derived (either analytically, possible only in very seldom cases, or by approximating it with a parametric form as it is done in variational inference), or a sampling algorithm, such as an MCMC method, is used. MCMC methods have desirable mathematical convergence properties, and they do not require assumptions on the parametric form of the posterior distribution. However, they do not scale well to problems where the computation of the likelihood is inefficient because they require several thousands of evaluations of the likelihood function. Replacing the VMEC code with the NN model results in orders of magnitude faster likelihood evaluations, therefore, it brings the Bayesian inference approach to a scale that was not possible before.

## 3 Results

This section investigates the inference of plasma profiles for both synthetic and experimental W7-X data: section 3.1 visualizes the simple reconstruction of two parameters on simulated diagnostic data; section 3.2 investigates the robustness of the inference procedure across diverse W7-X plasma scenarios; section 3.3 compares the reconstruction of plasma profiles and their uncertainties with three different MHD equilibrium models: a fixed, finite-beta VMEC equilibrium; a self-consistent VMEC equilibrium; and a self-consistent NN equilibrium.



### 3.1 2D inference on synthetic data

To introduce the inference procedure, and to visualize the free parameter values during inference, a simplified 2D inference on synthetic data is reported. Apart for the electron density scale factor  $h_{n_e,0}$  and the total toroidal flux enclosed by the plasma  $\Phi_{\text{edge}}$ , all other model parameters are fixed.

In this section, diagnostic data are simulated with this Minerva graph rather than taken from the experiment. Observations are modeled using normal or truncated normal distributions. The predicted data from the forward model determine the distribution means, and the experimentally informed uncertainties determine the distribution standard deviations. The synthetic diagnostic data are then generated by setting the free parameter values to their target one (see paragraph below), running the forward model with VMEC as the equilibrium solver, setting the observation means to the predicted diagnostic data, and finally drawing samples from the observation distributions. We would like to make two important remarks: the generated synthetic data are obtained using an accurate equilibrium provided by VMEC, while the inference process uses equilibria obtained from the NN model; the synthetic data are sampled from distributions whose sigma is informed by experimental uncertainties, ensuring that they are not noise-free.

A W7-X plasma scenario resembling the post-pellet phase with an enhanced confinement [4] is considered (W7X20181016.037 at  $t = 1.7$  s): central density  $n_e(\rho = 0) = 8.0 \times 10^{19} \text{ m}^{-3}$ , density peaking factor  $n_e(\rho = 0)/n_e(\rho = 0.8) \simeq 2.77$ , central temperature  $T_e(\rho = 0) = 3.0 \text{ keV}$ , flat temperature profile till  $\rho \simeq 0.2$ , net toroidal current  $I_{\text{tor}} = 1.3 \text{ kA}$ , and  $\Phi_{\text{edge}} = 1.945 \text{ V s}$  as from the vacuum equilibrium. For simplicity,  $h_{n_e,3} + h_{T_e,3} = 1.0$ . The resulting volume average plasma beta is  $\langle \beta \rangle = 1.31\%$ .

In this simplified 2D inference, the MAP optimization quickly converges to the true values (figure 2): after five iterations, the inferred values (black dots) are already close to the target ones (red star). The inference initial guess (black star) is sampled from the parameter prior distributions. The relative difference between the inferred and true values is less than 1% (table 1).

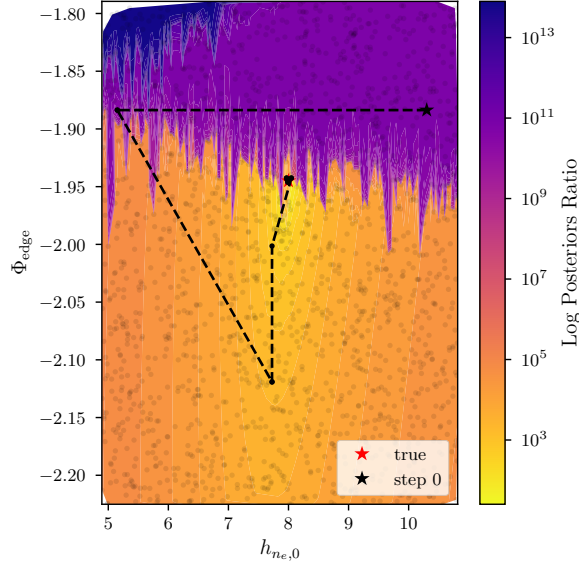
Figure 2 also shows the contours of the full posterior distribution normalized to the posterior distribution of the MAP values in the  $(h_{n_e,0}, \Phi_{\text{edge}})$  space: a dark color indicates that a region of parameter space is less likely than the MAP values, on the other hand, a light color indicates that a region of parameter space is equally likely as the MAP values. 10 000 random samples are used to interpolate the posterior distribution. Figure 2 does also highlight the advantage of a fast MHD equilibrium model: sampling 10 000 parameter values, which requires computing 10 000 self-consistent MHD equilibria, took only about 20 min (in a serial fashion on a single core). Hence, it is now possible to quickly explore the entire posterior distribution.

**Table 1.** Reconstructed free parameter values for the 2D inference on synthetic data. See section 2.2.1 for the description of each parameter.

Parameter	true	initial	inferred	difference [%]
$h_{n_e,0}$	8.00	10.30	7.98	0.19
$\Phi_{\text{edge}}$	-1.95	-1.88	-1.94	0.04

### 3.2 Inference robustness on synthetic data

Is the NN model's approximation of the ideal-MHD equilibrium robust across diverse plasma scenarios? The NN model was trained on a large set of W7-X equilibria for various vacuum magnetic configurations, plasma pressure and toroidal current profiles. In this section, we investigate if the inference procedure is robust across plasma scenarios (e.g., high- and low- $\beta$ , flat or peaked pressure profiles).



**Figure 2.** Trajectory of the free parameters during inference. In this case, the free parameters are the electron density scale factor  $h_{ne,0}$  and the net toroidal flux enclosed by the plasma  $\Phi_{edge}$ . The reconstruction process starts from the black star, which represents the initial guess, and converges to the red star, which represents the true state. The dashed black line indicates the order of parameter values encountered during inference. The synthetic observations are generated using the Minerva graph and VMEC as equilibrium model, while the inference is performed using the NN as equilibrium model. The logarithm of the ratio of the sample posterior probability to the MAP posterior probability is also shown. The light dark dots represent 10 000 random samples used to interpolate the posterior distribution.

Using synthetic data provides the advantage of being able to conveniently scan the parameter space in different trials: for each inference trial, the true state and the initial guess are sampled from the model parameter prior distributions. The plasma state is represented by the vector  $\vec{h} = (h_{Te,0}, h_{Te,1}, h_{Te,2}, h_{ne,0}, h_{ne,1}, h_{ne,2})^T$ , while  $h_{Te,3} = h_{ne,3} = 1$  are fixed. Similarly to section 3.1, synthetic data are generated by injecting the true plasma state into the graph, and by setting the observed data to samples drawn from the observation distributions, using VMEC as equilibrium model. The free parameters are then initialized to their initial guesses, and the posterior distribution is maximized for 100 iterations. The NN model approximates the MHD equilibrium during inference. The inference procedure is robust across plasma state and initial guess (figures 3a to 3f). Figures 3a to 3f show the trajectory of the free parameters (normalized to their true values) as a function of the MAP iterations. Because the free parameters are normalized to their true values, a final value of 1 (red dashed line) represents a successfully converged inference. The trajectory means (solid line) and 5–95 % quantiles (dashed lines) are estimated with 25 independent inference

trials. The profile scale factors ( $h_{n_e,0}$  and  $h_{T_e,0}$ ) are accurately inferred, whereas the distributions of the profile shape parameters show a non-finite spread in their estimated values. Yet, the inferred values are on average close to the target values. As a representative case, table 2 shows the true, initial, and inferred parameter values in case of the median trial (i.e., the trial that has the median average percentage inference error across all trials).

**Table 2.** Target, initial guess, and inferred parameter values for the median trial. See section 2.2.1 for the description of each parameter.

Parameter	true	initial	inferred	difference [%]
$h_{T_e,0}$	3.87	4.37	3.87	0.20
$h_{T_e,1}$	2.70	2.17	2.87	6.44
$h_{T_e,2}$	3.19	0.20	3.38	5.93
$h_{n_e,0}$	13.92	12.20	13.92	0.03
$h_{n_e,1}$	4.93	2.95	4.96	0.62
$h_{n_e,2}$	6.73	2.90	6.72	0.22

### 3.3 Inference of improved confinement scenarios at W7-X

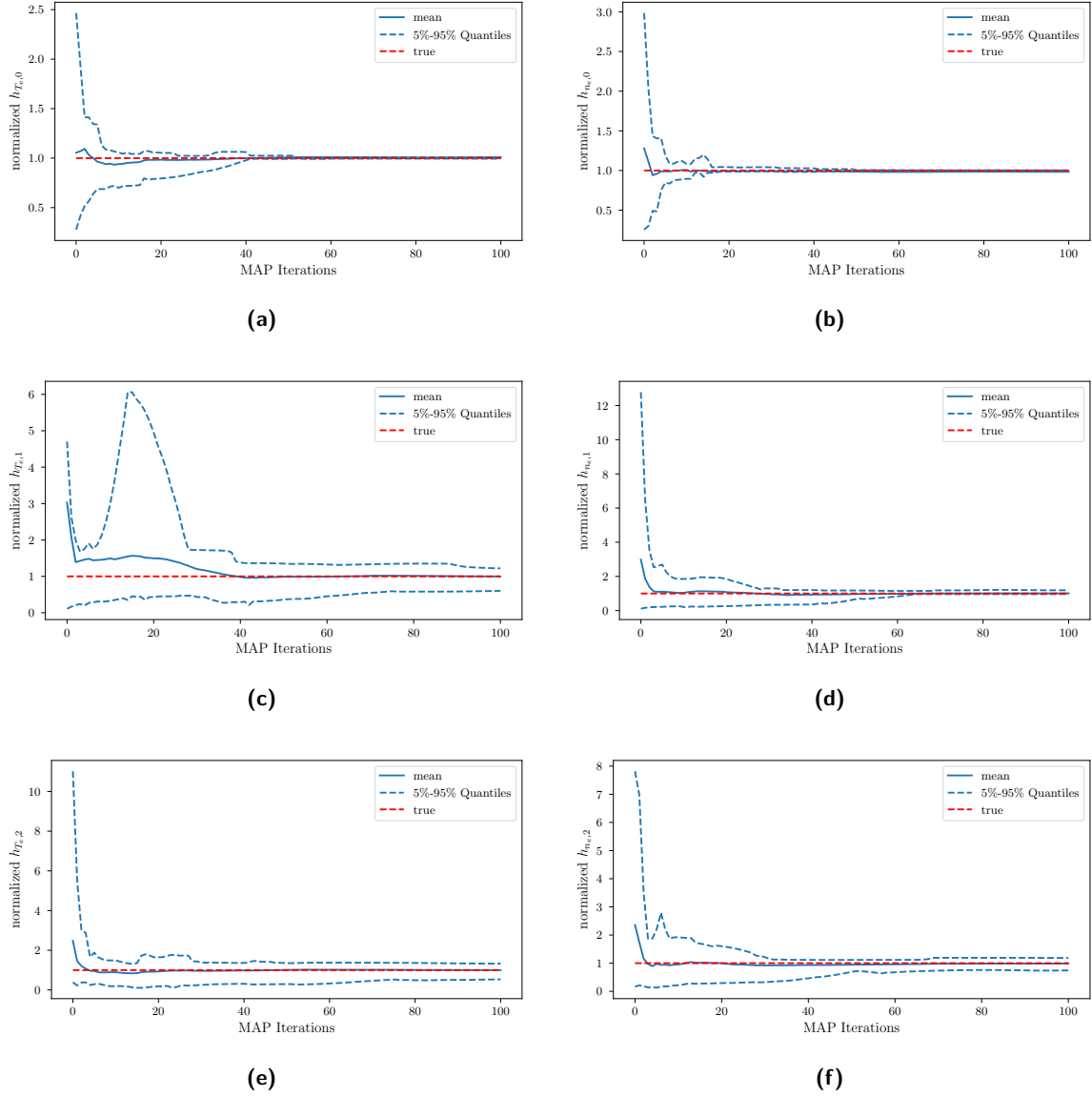
In this section, we replace the synthetic diagnostic data with experimental data. We address two questions: how does the approximation of the equilibria by the NN model affect the inferred plasma parameters? What are the implications of a self-consistent MHD equilibrium for plasma parameter estimation?

To address both questions, two plasma scenarios with  $T_i \simeq T_e$  at different  $\langle\beta\rangle$  and magnetic configurations are considered:

- W7X20181016.37,  $t = 1.7$  s: a  $\langle\beta\rangle > 1\%$ , improved confinement scenario following pellet injections in the standard configuration [4].
- W7X20180808.5,  $t = 20.0$  s: a low-power, improved confinement scenario in the high-iota configuration [41].

In this section, we compare three different ideal-MHD equilibrium models: a reference, fixed, finite-beta VMEC equilibrium (see description below); a self-consistent VMEC equilibrium; and a self-consistent equilibrium. Comparing the two self-consistent NN equilibrium models allows the investigation of the approximations provided by the NN model, while comparing the fixed and self-consistent equilibrium models enables the analysis of the implications of a self-consistent equilibrium on the inferred plasma profiles.

The reference VMEC equilibrium is taken from the set of precalculated finite-beta equilibria commonly used at W7-X [16]. The pressure profile is assumed to be  $p_0(1-s)^{a_2}$ , where  $p_0$  is the pressure scale, and  $a_2$  determines the pressure peaking factor. Estimating  $a_2$  from the experimental profiles, and assuming  $p_0$  so that  $W_{\text{kin}}$  matches the measured  $W_{\text{dia}}$  value (with a plasma volume of  $V_p = 30 \text{ m}^3$ ), the closest precalculated equilibrium in terms of Euclidean distance in the  $(p_0, a_2)$  space is selected. The net toroidal current is less than 2 kA in these experiments and can be neglected.  $\Phi_{\text{edge}}$  is set to the toroidal flux enclosed by the LCFS of the vacuum field obtained from field line tracing.



**Figure 3.** Convergence of the model parameters during inference for 25 independent trials. Each trial yields a randomly sampled set of true and initial guess values. Each figure shows the mean (solid line) and 5%–95% quantiles (dashed lines) of the normalized (to their true value) free parameters as a function of the MAP optimization iterations. A red dashed line guides the eye to indicate the normalized target value (i.e., 1). See section 2.2.1 for the description of each parameter.

In case of the self-consistent MHD equilibria, the equilibrium is not fixed, but evaluated consistently with the proposed density and temperature profiles. Nine free parameters represent the plasma state: four profile parameters each for  $n_e$  and  $T_e$ , and a global calibration factor for the TS (see section 2.2.1).  $\Phi_{\text{edge}}$  is held fixed. As in section 3.2, the toroidal current profile shape is fixed, but the net toroidal current is set to the value measured by the Rogowski coil.

To improve the inference’s convergence, diagnostics raw data inform the initial guess:  $h_{n_e,0}$  is set to be consistent with the line average integrated density from the DI (assuming a flat density

profile and a LOS length of 1.1 m), and  $h_{Te,0}$  is chosen so that the plasma's kinetic energy matches the measured diamagnetic energy from the diamagnetic loop (assuming a plasma volume of  $V_p = 30 \text{ m}^3$ , and a linear pressure profile in normalized toroidal flux).

For both considered experiments, the inferred profile parameters with the three MHD equilibrium models are compatible (tables 3 and 4): after 100 MAP iterations, the parameter values are within one standard deviation away from each other. The standard deviations are estimated using 30 000 MCMC samples, of which the first 10 000 are discarded, and only 1 in every 50 are kept to reduce the sample autocorrelation.

It is worth highlighting the time needed to perform each inference (100 MAP iterations on a single core, given here in case of the *W7X20181016.37* shot): using the self-consistent VMEC equilibrium took 1108 min and 45 s, using the fixed VMEC equilibrium took 9 min and 35 s, and using the NN model took only 3 min and 31 s. The inference time is reduced by more than two orders of magnitude.

**Table 3.** Comparison of the inferred free parameters in case of the *W7X20181016.37* shot at  $t = 1.7 \text{ s}$  with the three ideal-MHD models: a fixed, finite-beta precalculated VMEC equilibrium; a self-consistent VMEC equilibrium; and the NN model described in section 2.2.1. The free parameter values are obtained via a MAP optimization, and the standard deviation (in brackets) is estimated with MCMC. Because of the computational time required to run VMEC for each MCMC sample, the MCMC procedure has not been performed in case of the self-consistent VMEC equilibrium.

Parameter	Fixed VMEC	Self-consistent VMEC	Self-consistent NN
$h_{Te,0}$	3.39 (0.16)	3.403	3.40 (0.17)
$h_{Te,1}$	1.04 (0.36)	1.038	1.03 (0.38)
$h_{Te,2}$	4.7 (1.9)	4.834	4.7 (1.9)
$h_{Te,3}$	1.129 (0.081)	1.138	1.130 (0.080)
$h_{ne,0}$	8.27 (0.51)	8.341	8.35 (0.50)
$h_{ne,1}$	0.16 (0.32)	0.140	0.14 (0.33)
$h_{ne,2}$	2.77 (0.74)	2.726	2.71 (0.89)
$h_{ne,3}$	1.330 (0.072)	1.330	1.328 (0.079)

As depicted in tables 3 and 4, the three ideal-MHD models yield compatible plasma profiles. This is true not only in terms of MAP profiles, but also for their posterior distributions estimated with MCMC (figures 4a to 4d).

The inferred profiles are also consistent with a independently inferred electron temperatures and densities per TS volume (figures 4a to 4d). For the reference Minerva inference, the density and temperature values at each TS volume are independently reconstructed using the fixed, finite-beta VMEC equilibrium as MHD model, and the density is further scaled to match the line-integrated value from the DI. See [28, section 5] for a description of the inference procedure.

Tables 5 and 6 show selected quantities of interest of the resulting equilibria. In case of the inference performed with the NN model, the mean and standard deviation (in brackets) of the MCMC samples are reported.

The inferred MHD equilibria obtained using the self-consistent VMEC and NN models feature compatible properties (tables 5 and 6): the inferred values with the self-consistent NN model are within one standard deviation with respect to the values inferred with the self-consistent VMEC

**Table 4.** Comparison of the inferred free parameters in case of the *W7X20180808.5* shot at  $t = 20.0$  s with the three ideal-MHD models. See table 3 for a description of each row.

Parameter	Fixed VMEC	Self-consistent VMEC	Self-consistent NN
$h_{T_e,0}$	1.930 (0.089)	1.928	1.915 (0.088)
$h_{T_e,1}$	0.001 (0.323)	0.001	0.002 (0.344)
$h_{T_e,2}$	4.1 (1.4)	4.126	4.4 (1.4)
$h_{T_e,3}$	1.330 (0.087)	1.330	1.330 (0.086)
$h_{n_e,0}$	4.32 (0.49)	4.258	4.21 (0.50)
$h_{n_e,1}$	2.52 (0.65)	2.524	2.35 (0.68)
$h_{n_e,2}$	8.2 (1.9)	8.535	6.4 (1.9)
$h_{n_e,3}$	1.079 (0.075)	1.090	1.051 (0.072)

equilibrium. It is worth noting that, for the two experimental scenarios investigated, the assumed pressure profile in the precalculated fixed equilibrium (first column in the tables) differs from the inferred pressure profiles (second and third columns in the tables).

**Table 5.** Comparison of selected quantities of the inferred equilibria of experiment *W7X20181016.37* at  $t = 1.7$  s with three MHD models: a fixed, finite-beta precalculated VMEC equilibrium; a self-consistent VMEC equilibrium; and the self-consistent NN model as described in section 2.2.1. In case of the inference performed with the self-consistent NN model, the mean and standard deviation (in brackets) of the MCMC samples are reported. In these experimental conditions, the measured diamagnetic energy is  $W_{\text{dia}} = 1137.661$  kJ.

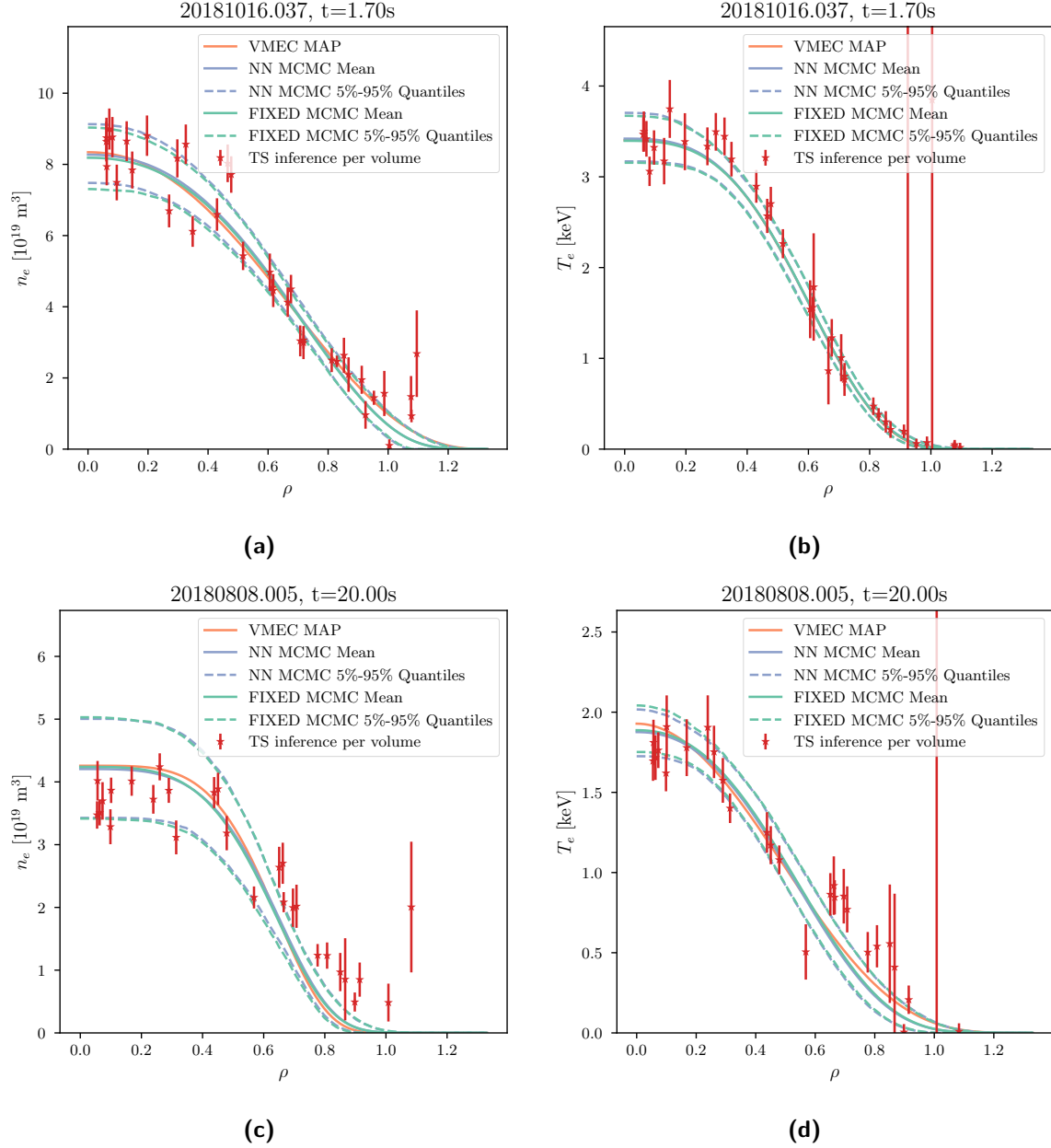
Quantity	Fixed VMEC	Self-consistent VMEC	Self-consistent NN
$\langle\beta\rangle$ [%]	1.176	1.067	1.083 (0.064)
$t_{\text{axis}}$	0.862	0.857	0.8597 (0.0012)
$t_{\text{LCFS}}$	0.971	0.974	0.97414 (0.00014)
$W_{\text{kin}}$ [kJ]	1198	1089	1090 (64)
$R_{\text{axis}}(\varphi = 0)$ [m]	5.983	5.987	5.9876 (0.0028)
$p_0$ [kPa]	80.0	91.1	90.6 (6.7)
pressure peaking factor	2.98	3.74	3.65 (0.19)

## 4 Conclusions

In this work, we found that employing the NN model from [26] as ideal-MHD equilibrium model in the Minerva Bayesian framework yields inferred plasma parameters compatible with the parameters obtained using VMEC. We have also investigated the robustness of the NN accelerated inference using synthetic data, and we observed that it is robust across diverse W7-X plasma scenarios.

The inference time can be reduced by more than two orders of magnitude when using the NN model. Moreover, given the fast computational time, plasma parameter posterior distributions can be approximated using MCMC samples with self-consistent ideal-MHD equilibria.

This work has two major limitations: no magnetic diagnostics (except for the Rogowski coil) are included in the Minerva graph, owing to the lack of accuracy of the NN model in faithfully computing



**Figure 4.** Comparison of the inferred electron density (figures 4a and 4c) and temperature (figures 4b and 4d) profiles with the following MHD models: a fixed, finite-beta precalculated VMEC equilibrium; a self-consistent VMEC equilibrium; and the self-consistent NN model described in section 2.2.1. The solid lines represent the mean profile across all MCMC samples (in case of the self-consistent VMEC equilibrium, the solid line is the MAP profile), and dashed lines represent the 5%–95% quantiles. In addition, the values (as well as their standard deviation) of the electron density and temperature at each TS volume location, from a reference Minerva inference, are also shown (red stars). The shot id and time of each reconstructed slice are shown in each figure title.

**Table 6.** Comparison of selected quantities of the inferred equilibria of experiment *W7X20180808.5* at  $t = 20.0$  s with the three MHD models, analogous to table 5.  $W_{\text{dia}} = 232.188$  kJ.

Quantity	Fixed VMEC	Self-consistent VMEC	Self-consistent NN
$\langle\beta\rangle$ [%]	0.253	0.273	0.268 (0.034)
$t_{\text{axis}}$	1.012	1.003	1.00253 (0.00026)
$t_{\text{LCFS}}$	1.193	1.172	1.17385 (0.00012)
$W_{\text{kin}}$ [kJ]	218	236	228 (29)
$R_{\text{axis}}(\varphi = 0)$ [m]	5.976	5.977	5.97641 (0.00092)
$p_0$ [kPa]	20.0	26.3	25.3 (3.2)
pressure peaking factor	3.47	4.22	4.13 (0.31)

the plasma current density,<sup>1</sup> which is required to predict the magnetic diagnostics observations (the plasma contribution to the magnetic field at a location outside the plasma volume depends on a volume integral of the plasma current density [42]);  $T_i \sim T_e$  and  $Z_{\text{eff}} \simeq 1$  have been assumed, as a result, the approach can be tested only on a limited set of experimental conditions where these assumptions are close to being satisfied. To infer additional plasma parameters, more diagnostics can be included in the Minerva graph. A XICS can be used to constrain the ion temperature [43], and a spectrometer can be used to infer the effective charge  $Z_{\text{eff}}$  [44].

Using NN ideal-MHD models, it is possible to quickly infer plasma parameters while taking into account finite-beta effects in the equilibrium. If diagnostic data are available, the MHD self-consistent inference of plasma parameters can theoretically be performed between shots, providing valuable insights for the conduction of operational campaigns.

## 5 Author statement

The contributions to this paper are described using the CRediT taxonomy [45]:

- *Andrea Merlo*: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing — original draft, Writing — review & editing.
- *Andrea Pavone*: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Writing — review & editing.
- *Daniel Böckenhoff*: Methodology, Software, Supervision, Validation, Writing — review & editing.
- *Ekkehard Pasch*: Data Curation.
- *Kai Jakob Brunner*: Data Curation.
- *Kian Rahbarnia*: Data Curation.
- *Jonathan Schilling*: Software.

<sup>1</sup>The plasma current density depends up to the second-order derivatives of the equilibrium solution, and the equilibrium solution provided by the NN model does not have smooth second-order radial derivatives [26].



- *Udo Höfel*: Software.
- *Sehyun Kwak*: Software.
- *Jakob Svensson*: Software.
- *Thomas Sunn Pedersen*: Funding acquisition, Supervision.

## Acknowledgments

We are indebted to the communities behind the multiple open-source software packages on which this work depends: hydra [46], matplotlib [47], numpy [48], pymc3 [49], pytorch [50], pytorch lightning [51], scipy [52].

Financial support by the European Social Fund (ID: ESF/14-BM-A55-0007/19) and the Ministry of Education, Science and Culture of Mecklenburg-Vorpommern, Germany via the project “NEISS” is gratefully acknowledged. This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

## A Prior distributions of model parameters

Table 7 lists the prior distribution of the model free parameters.

**Table 7.** The prior distributions of the model parameters.  $U(a, b)$  means a uniform distribution defined on  $(a, b)$ . The bounds of uniform distribution of  $\Phi_{\text{edge}}$  are set by the distribution of the training data on which the NN model has been trained.

Parameter	Distribution	Unit
Electron temperature		
$h_{T_e,0}$	$U(1.0 \times 10^0, 6.0 \times 10^0)$	keV
$h_{T_e,1}$	$U(10^{-3}, 8.0 \times 10^0)$	-
$h_{T_e,2}$	$U(10^{-1}, 10^1)$	-
$h_{T_e,3}$	$U(1.0 \times 10^0, 1.33 \times 10^0)$	-
Electron density		
$h_{n_e,0}$	$U(1.0 \times 10^0, 15.0 \times 10^0)$	$10^{19} \text{ m}^{-3}$
$h_{n_e,1}$	$U(10^{-3}, 8.0 \times 10^0)$	-
$h_{n_e,2}$	$U(10^{-1}, 10^1)$	-
$h_{n_e,3}$	$U(1.0 \times 10^0, 1.33 \times 10^0)$	-
Net toroidal flux		
$\Phi_{\text{edge}}$	$U(-2.31 \times 10^0, -1.35 \times 10^0)$	Wb

## B Modeling the uncertainty due to the TS laser misalignment

This section briefly introduces how the uncertainty due to the TS laser misalignment is modeled in the Minerva graph.

To estimate the uncertainties due to the TS laser misalignment, we make use of the fact that the inferred electron density linearly depends on the TS observations (see equation (2.7)). In this work, we assume that only two factors affect electron density: a change in the macroscopic plasma state (e.g., pellet injection, neutral beam injection (NBI) heating) or a shift in the TS laser. To rule out the former, we consider time windows from multiple shots where all the other plasma parameters apart from the inferred electron density and temperature from the TS are constant. Previous Minerva reconstruction are used for this estimation. Table 8 lists the considered shots and time windows.

**Table 8.** The W7-X shots and time windows used to inform the modeling of the uncertainty due to the TS laser misalignment. The magnetic configuration and line integrated density (as measured by the DI) are also listed.

Shot	Time Window [s]	Configuration	$\int n_e dl$ [ $10^{19} \text{m}^{-2}$ ]
<i>W7X20180808.005</i>	$45 \times 10^0 - 54 \times 10^0$	FTM+252 $\times 10^0$	$3 \times 10^0$
<i>W7X20180809.039</i>	$3 \times 10^0 - 4 \times 10^0$	KKM+252 $\times 10^0$	$5 \times 10^0$
<i>W7X20181016.013</i>	$4 \times 10^0 - 15 \times 10^0$	EJM+262 $\times 10^0$	$12 \times 10^0$
<i>W7X20181016.014</i>	$4 \times 10^0 - 19 \times 10^0$	EJM+262 $\times 10^0$	$12 \times 10^0$
<i>W7X20181016.026</i>	$16 \times 10^0 - 36 \times 10^0$	EJM+262 $\times 10^0$	$6 \times 10^0$
<i>W7X20181017.016</i>	$12 \times 10^0 - 32 \times 10^0$	FTM+251 $\times 10^0$	$5 \times 10^0$
<i>W7X20181017.025</i>	$8 \times 10^0 - 28 \times 10^0$	EJM+262 $\times 10^0$	$12 \times 10^0$

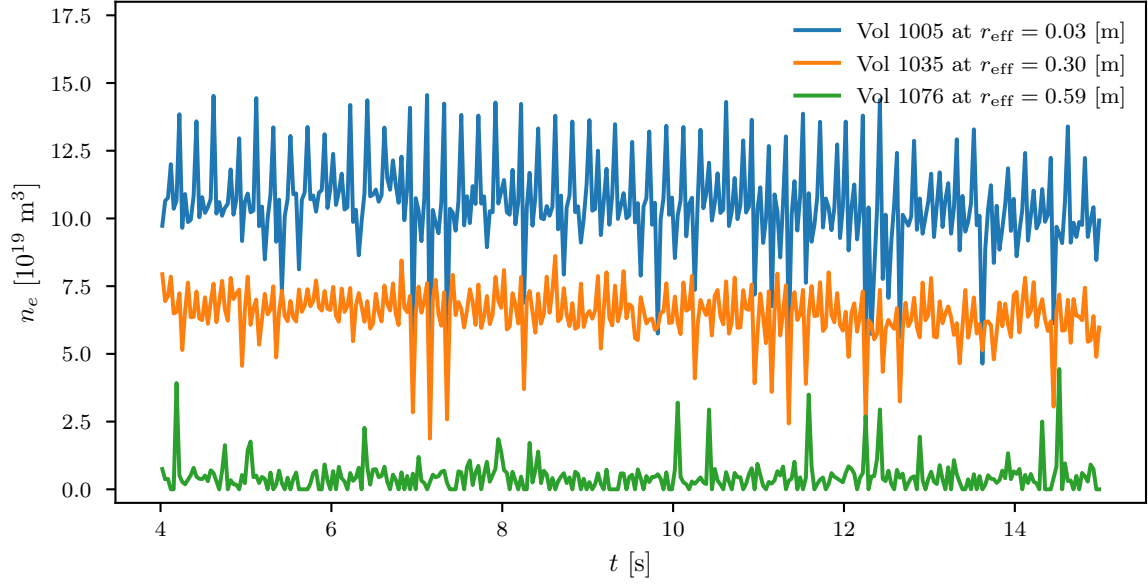
Figure 5 shows an example of these shot, in which the inferred electron density of three TS volumes is plotted over time. Given that all the other plasma parameters are constant in this time window, we assume that even the actual electron density is constant as well, and that the variations in the inferred values are due to the TS laser shifts. Therefore, we assume the time average (over the whole time window) of the electron density to be the actual density, and we consider the standard variation (in time) of the inferred electron density as a proxy to model the uncertainty due to the TS misalignment.

How does the standard deviation depend on the inferred value? Figure 6 shows the relationship between the standard deviation and average electron density for all the evaluated shots. When considering only the volumes inside the LCFS (stars), a simple linear model reasonably describes the data:  $\sigma_{n_e} \sim TN(\alpha \mu_{n_e}, \sigma_{\sigma_{n_e}})$ , where  $TN$  stands for a truncated normal distribution. However, the volumes outside the LCFS (dots) do not seem to follow the same trend. The posterior distributions of  $\alpha$  and  $\sigma_{\sigma_{n_e}}$  are numerically estimated with 50 000 MCMC samples. The mean of the posterior distribution yields  $\alpha = 0.18$ .

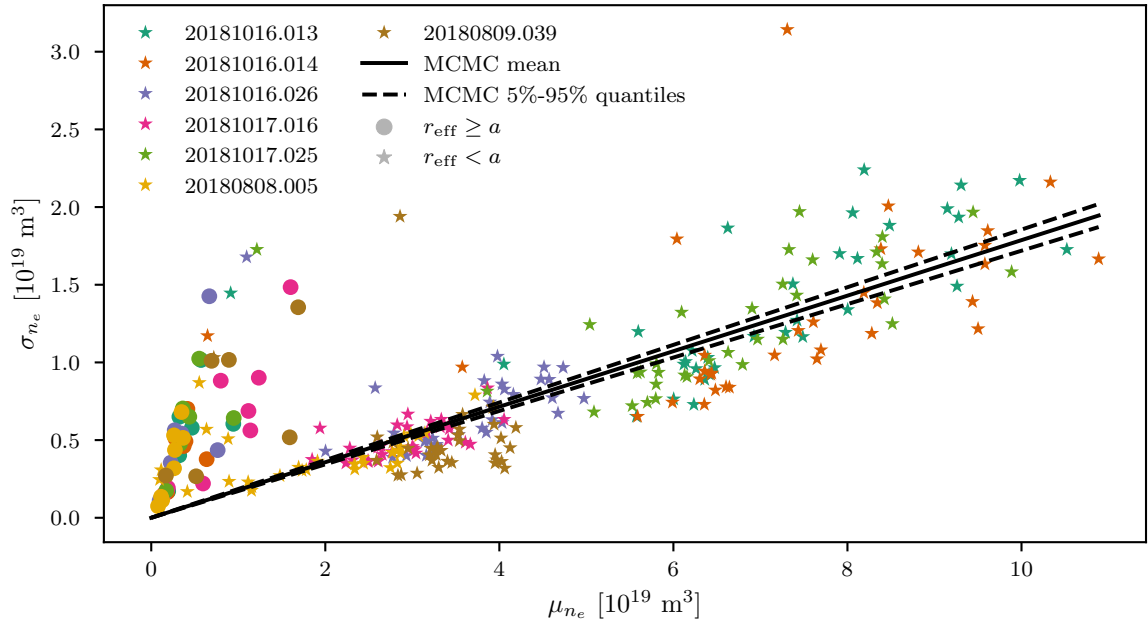
Finally, we model the uncertainty of a generic TS volume due to TS laser misalignment using the linear relationship between the electron density and the TS observed data as follows:

$$\sigma_{\text{laser}}^i = \max(\sigma_0, \alpha D_{\text{TS}}^i), \quad (\text{B.1})$$

where  $\sigma_0$  accounts for the standard deviation of the TS volumes outside the LCFS not well modeled by the linear relationship.  $\sigma_0$  is usually set to a small value compared to the observed TS data:  $\sigma_0 \simeq \frac{2.5}{100} \max_i(D_{\text{TS}}^i)$ .



**Figure 5.** The inferred electron density for three TS volumes in stationary plasma conditions. In the selected time window, all other macroscopic plasma parameters are nearly constant. Volumes in the core (1005), mid-radius (1035), and plasma edge (1076) are depicted. Data from previous Minerva reconstructions are used.



**Figure 6.** The inferred electron density standard deviation as a function of the average value in each time window. Colors indicate different shots, stars indicate TS volumes inside the LCFS, and dots indicate TS volumes outside the LCFS. Given a simple linear model of the form  $\sigma_{n_e} \sim TN(\alpha\mu_{n_e}, \sigma_{\sigma_{n_e}})$ , where  $TN$  stands for a truncated normal distribution, the solid and dashed lines represent the mean and 5%–95% quantiles of  $\alpha$ , respectively.

## References

- [1] G. Grieger and I. Milch, *Das Fusionsexperiment WENDELSTEIN 7-X*, *Physikalische Blätter* **49** (1993) 1001.
- [2] R.C. Wolf, *A Stellarator Reactor Based on the Optimization Criteria of Wendelstein 7-X*, *Fusion Eng. Des.* **83** (2008) 990.
- [3] J. Geiger et al., *Physics in the magnetic configuration space of W7-X*, *Plasma Phys. Control. Fusion* **57** (2014) 014004.
- [4] S.A. Bozhnikov et al., *High-performance plasmas after pellet injections in Wendelstein 7-X*, *Nucl. Fusion* **60** (2020) 066011.
- [5] L.L. Lao et al., *Reconstruction of current profile parameters and plasma shapes in tokamaks*, *Nucl. Fusion* **25** (1985) 1611.
- [6] L.L. Lao et al., *Equilibrium analysis of current profiles in tokamaks*, *Nucl. Fusion* **30** (1990) 1035.
- [7] B.J. Braams, *The interpretation of tokamak magnetic diagnostics*, *Plasma Phys. Control. Fusion* **33** (1991) 715.
- [8] X. Ma et al., *Non-axisymmetric equilibrium reconstruction of a current-carrying stellarator using external magnetic and soft x-ray inversion radius measurements*, *Phys. Plasmas* **22** (2015) 122509.
- [9] M.D. Pandya et al., *The magnetic structure of long-wavelength magnetohydrodynamic modes in current-carrying stellarator plasmas*, *Phys. Plasmas* **28** (2021) 112503.
- [10] N. Pomphrey et al., *Plasma Control for NCSX and Development of Equilibrium Reconstruction for Stellarators*, Technical report (2005).
- [11] E. Chlechowicz et al., *Magnetic diagnostic optimization for plasma equilibrium reconstruction in the HSX stellarator*, *Nucl. Fusion* **55** (2015) 113012.
- [12] J.C. Schmitt, J.N. Talmadge and D.T. Anderson, *Measurement of a Helical Pfirsch-Schlüter Current with Reduced Magnitude in HSX*, *Nucl. Fusion* **53** (2013) 082001.
- [13] J.C. Schmitt, J.N. Talmadge, D.T. Anderson and J.D. Hanson, *Modeling, measurement, and 3-D equilibrium reconstruction of the bootstrap current in the Helically Symmetric Experiment*, *Phys. Plasmas* **21** (2014) 092518.
- [14] H.P. Callaghan, P.J. McCarthy and J. Geiger, *Fast equilibrium interpretation on the W7-AS stellarator using function parameterization*, *Nucl. Fusion* **39** (1999) 509.
- [15] J. Schilling, *Experimental MHD Equilibrium Analysis of Magnetic Configurations in the Wendelstein 7-X Stellarator*, MSc Thesis IPP 2018-20, Max-Planck-Institute for Plasma Physics, Germany (2018).
- [16] T. Andreeva et al., *Equilibrium evaluation for Wendelstein 7-X experiment programs in the first divertor phase*, *Fusion Eng. Des.* **146** (2019) 299.
- [17] J.D. Hanson et al., *V3FIT: a code for three-dimensional equilibrium reconstruction*, *Nucl. Fusion* **49** (2009) 075031.
- [18] E.C. Howell and J.D. Hanson, *Development of a non-parametric Gaussian process model in the three-dimensional equilibrium reconstruction code V3FIT*, *J. Plasma Phys.* **86** (2020) 905860102.
- [19] S.A. Lazerson, *Three-dimensional equilibrium reconstruction on the DIII-D device*, *Nucl. Fusion* **55** (2015) 023009.

- [20] J. Svensson and A. Werner, *Large Scale Bayesian Data Analysis for Nuclear Fusion Experiments*, in the proceedings of the 2007 IEEE International Symposium on Intelligent Signal Processing, Alcala de Henares, Spain 2007, [DOI:10.1109/wisp.2007.4447579].
- [21] J. Svensson et al, *Connecting Physics Models and Diagnostic Data Using Bayesian Graphical Models*, in 37th EPS Conference on Plasma Physics, European Physical Society, Dublin, Ireland (2010).
- [22] H.P. Callaghan, P.J. McCarthy and J. Geiger, *Pressure profile recovery on W7-AS with idealized external magnetic measurements using function parametrization*, *Plasma Phys. Control. Fusion* **42** (2000) 1013.
- [23] A. Sengupta, P.J. McCarthy, J. Geiger and A. Werner, *Fast recovery of vacuum magnetic configuration of the W7-X stellarator using function parametrization and artificial neural networks*, *Nucl. Fusion* **44** (2004) 1176.
- [24] A. Sengupta, J. Geiger and P.J.M. Carthy, *Statistical analysis of the equilibrium configurations of the W7-X stellarator*, *Plasma Phys. Control. Fusion* **49** (2007) 649.
- [25] J. Geiger, M. Endler and A. Werner, *Magnetic Diagnostics for Equilibrium Reconstruction at W7-X*, *Contrib. Plasma Phys.* **50** (2010) 736.
- [26] A. Merlo et al., *Physics-regularized neural network of the ideal-MHD solution operator in Wendelstein 7-X configurations*, *Nucl. Fusion* **63** (2023) 066020.
- [27] E. Pasch et al., *The Thomson scattering system at Wendelstein 7-X*, *Rev. Sci. Instrum.* **87** (2016) 11E729.
- [28] S.A. Bozhnikov et al., *The Thomson scattering diagnostic at Wendelstein 7-X and its performance in the first operation phase*, *2017 JINST* **12** P10004.
- [29] J. Knauer et al, *A New Dispersion Interferometer for the Stellarator Wendelstein 7-X*, in 43rd EPS Conference on Plasma Physics, European Physical Society (2016).
- [30] C. Killer et al., *Characterization of the W7-X scrape-off layer using reciprocating probes*, *Nucl. Fusion* **59** (2019) 086013.
- [31] M. Endler et al., *Engineering design for the magnetic diagnostics of Wendelstein 7-X*, *Fusion Eng. Des.* **100** (2015) 468.
- [32] S.P. Hirshman and J.C. Whitson, *Steepest-descent moment method for three-dimensional magnetohydrodynamic equilibria*, *Phys. Fluids* **26** (1983) 3553.
- [33] ONNX Runtime developers, *ONNX Runtime*, <https://onnxruntime.ai/>.
- [34] T. Andreeva, *Vacuum Magnetic Configurations of Wendelstein 7-X*, (2002), [https://pure.mpg.de/rest/items/item\\_2138236/component/file\\_2138235/content](https://pure.mpg.de/rest/items/item_2138236/component/file_2138235/content).
- [35] T. Klinger et al., *Performance and properties of the first plasmas of Wendelstein 7-X*, *Plasma Phys. Control. Fusion* **59** (2016) 014018.
- [36] W7-X TEAM collaboration, *Quantification of systematic errors in the electron density and temperature measured with Thomson scattering at W7-X*, *2023 JINST* **18** P09019 [arXiv:2111.03562].
- [37] G. Fuchert et al., *A novel technique for an alignment-insensitive density calibration of Thomson scattering diagnostics developed at W7-X*, *2022 JINST* **17** C03012.
- [38] R. Hooke and T.A. Jeeves, “Direct Search” Solution of Numerical and Statistical Problems, *J. ACM* **8** (1961) 212.
- [39] N. Metropolis et al., *Equation of state calculations by fast computing machines*, *J. Chem. Phys.* **21** (1953) 1087.

- [40] W.K. Hastings, *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, *Biometrika* **57** (1970) 97.
- [41] G.A. Wurden et al, *A Special Case of Long-Pulse High Performance Operation in W7-X*, in *48th EPS Conference on Plasma Physics*, European Physical Society (2022).
- [42] S.P. Hirshman et al., *Magnetic diagnostic responses for compact stellarators*, *Phys. Plasmas* **11** (2004) 595.
- [43] A. Langenberg et al., *Inference of temperature and density profiles via forward modeling of an x-ray imaging crystal spectrometer within the Minerva Bayesian analysis framework*, *Rev. Sci. Instrum.* **90** (2019) 063505.
- [44] A. Pavone et al., *Measurements of Visible Bremsstrahlung and Automatic Bayesian Inference of the Effective Plasma Charge Z<sub>eff</sub> at W7-X*, *2019 JINST* **14** C10003.
- [45] A. Brand et al., *Beyond authorship: attribution, contribution, collaboration, and credit*, *Learned Publishing* **28** (2015) 151.
- [46] O. Yadan, *Hydra*, Meta Research (2022).
- [47] J.D. Hunter, *Matplotlib: A 2D Graphics Environment*, *Comput. Sci. Eng.* **9** (2007) 90.
- [48] C.R. Harris et al., *Array programming with NumPy*, *Nature* **585** (2020) 357 [[arXiv:2006.10256](#)].
- [49] J. Salvatier, T. Wiecki and C. Fonnesbeck, *Probabilistic Programming in Python using PyMC*, *PeerJ Comput. Sci.* **2** (2016) e55 [[arXiv:1507.08050](#)].
- [50] A. Paszke et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, [arXiv:1912.01703](#) [[DOI:10.48550/arXiv.1912.01703](#)].
- [51] W. Falcon et al., *Pytorch Lightning*, [DOI:10.5281/ZENODO.3828935](#).
- [52] P. Virtanen et al., *SciPy 1.0 — Fundamental Algorithms for Scientific Computing in Python*, *Nature Meth.* **17** (2020) 261 [[arXiv:1907.10121](#)].