



Estimation of the Radial Tungsten Concentration Profiles from Soft X-ray Measurements at WEST with Bayesian Integrated Data Analysis

Hao Wu^{1,2} · Axel Jardin³ · Didier Mazon⁴ · Geert Verdoolaege² · The WEST Team

Accepted: 18 March 2024 / Published online: 20 April 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

The accumulation of heavy impurities like tungsten in the plasma core of fusion devices can cause significant radiative power losses or even lead to a disruption. It is therefore crucial to monitor the tungsten impurity concentration. In this paper, we follow the integrated data analysis approach using Bayesian probability theory to jointly estimate tungsten concentration profiles and kinetic profiles from soft X-ray, interferometry and electron cyclotron emission measurements. As the full Bayesian inference using Markov chain Monte Carlo sampling is time-consuming, we also discuss emulation of the inference process using neural networks, with a view to real-time implementation.

Keywords Bayesian inference · Tokamak · Integrated data analysis · Tungsten · Tomography · Gaussian process · Neural network

Introduction

The WEST tokamak is equipped with a full-tungsten actively cooled divertor and serves as a test bed for ITER-like tungsten monoblocks [1]. Tungsten (W) impurities entering the plasma pose a major risk to plasma core performance by radiative cooling. In principle, the tungsten concentration $c_W = n_W/n_e$ must be lower than 10^{-4} for fusion-relevant

operation [2, 3]. Therefore, it is highly desirable to determine reliable, space-resolved estimates of the tungsten density and avoid tungsten accumulation. Diagnostics measuring the plasma emission, e.g., extreme ultra-violet spectroscopy [4], X-ray spectroscopy [5] and bolometry [6] are often used for this purpose. However, inference of the tungsten concentration from the plasma emission also requires the electron density and temperature obtained by other diagnostics. The various sources of uncertainty and the intrinsic interdependencies of multiple diagnostics greatly increase the difficulty of obtaining reliable results that are consistent with all measurements, as well as credible estimates of the uncertainty on individual parameters of interest. Indeed, when the analysis using data from one diagnostic also depends on physical quantities inferred from other diagnostics, the error propagation can be quite difficult to deal with, especially when the errors are not Gaussian [7]. Instead of treating the data from different diagnostics separately and evaluating the uncertainty through complicated conventional error propagation analysis, one can resolve such challenges with a coherent combination of various measurements using the integrated data analysis (IDA) approach based on Bayesian probability theory [8]. Since a number of years, IDA has seen increased adoption in fusion experimental data analysis, where heterogeneous diagnostics often provide complementary, but also partially redundant information. The method can help resolving data inconsistencies and may reduce the uncertainty on estimates

✉ Hao Wu
hao.wu@ugent.be

Axel Jardin
axel.jardin@ifj.edu.pl

Didier Mazon
didier.mazon@cea.fr

Geert Verdoolaege
geert.verdoolaege@ugent.be

The WEST Team
<http://west.cea.fr/WESTteam>

¹ Southwestern Institute of Physics, CNNC, Chengdu 610041, People's Republic of China

² Department of Applied Physics, Ghent University, 9000 Ghent, Belgium

³ Institute of Nuclear Physics, Polish Academy of Sciences, 31-342 Krakow, Poland

⁴ IRFM, CEA, 13108 Saint Paul-lez-Durance, France

of physical quantities. Typical applications of IDA in the fusion community include the inference of temperature and density profiles [7, 9–12], the effective ion charge [13–15], the magnetic equilibrium [16], etc. In order to improve the reliability of physical parameters and allow uncertainty estimates, we apply IDA approach for the joint analysis of measurements of the soft X-ray emissivity at WEST, combined with density and temperature measurements.

This paper is organized as follows. We first introduce some basic concepts of Bayesian inference in “[Bayesian Inference](#)”, followed by the application to tomographic reconstruction of the soft X-ray emissivity at WEST in “[Bayesian Soft X-ray Tomography with Gaussian Process](#)”. “[Integrated Analysis of Tungsten Concentration](#)” concerns the integrated estimation of tungsten concentration profiles from both synthetic and real data. Finally, in “[Preliminary Exploration on Speeding Up the Inference with Surrogate Models](#)” we briefly discuss the possibility of accelerating the inference using neural network surrogate models, with a preliminary application to fast inference of electron density profiles.

Bayesian Inference

Parameter estimation, i.e., to determine values of unobserved parameters of interest $\vec{H} = [H_1, H_2, \dots, H_n]$ from experimental data $\vec{D} = [D_1, D_2, \dots, D_m]$, is a common data analysis problem. In Bayesian inference, this is done by inferring the conditional probability distribution of the parameters \vec{H} given the data \vec{D} , i.e., the *posterior distribution* $p(\vec{H}|\vec{D})$, which follows from Bayes’ theorem:

$$p(\vec{H}|\vec{D}) = \frac{p(\vec{D}|\vec{H})p(\vec{H})}{p(\vec{D})} \propto p(\vec{D}|\vec{H})p(\vec{H}). \tag{1}$$

The *likelihood* $p(\vec{D}|\vec{H})$ quantifies the mismatch between the data obtained from experiments and the predictions by the forward model for given parameter values \vec{H} . The *prior distribution* $p(\vec{H})$ represents the state of knowledge about the parameters before any observations are taken and can be used to implement regularization and physical constraints. The *evidence* or *marginal likelihood* $p(\vec{D}) = \int p(\vec{D}|\vec{H})p(\vec{H})d\vec{H}$ is a normalization constant with no explicit dependence on \vec{H} . It is usually omitted for parameter estimation problems [17].

By maximizing the posterior with respect to \vec{H} , we obtain a single most probable solution, known as the maximum a posteriori (MAP) estimate \vec{H}_{MAP} . However, the MAP estimate is not always optimal, and more generally one needs to calculate the marginal distribution of the individual parameters H_i , by integrating over the others:

$$p(H_i|\vec{D}) = \int p(\vec{H}|\vec{D}) dH_1 \dots dH_{i-1}dH_{i+1} \dots dH_n. \tag{2}$$

In most cases, the posterior distribution cannot be obtained in an analytical form and therefore the marginalization integral also has no closed-form expression. Exceptions can be found only for some special situations, for instance when both the prior and the likelihood are Gaussian and the forward model is linear, as will be explained in the next section. Numerical methods like Markov chain Monte Carlo (MCMC) are usually required in Bayesian inference to sample from the posterior distribution [17]. The values of H_i collected from the joint samples $[H_1, H_2, \dots, H_n]$ of the joint distribution $p(H_1, H_2, \dots, H_n|\vec{D})$ are also samples from the marginal distribution $p(H_i|\vec{D})$ [10].

Bayesian Soft X-ray Tomography with Gaussian Process

Soft X-ray Diagnostic at WEST

A new soft X-ray (SXR) diagnostic with energy discrimination has been developed for WEST [18]. It consists of two gas electron multiplier (GEM) cameras providing vertical and horizontal views, aiming at 2D tomographic reconstruction. However, as currently there is only limited data available from the GEM diagnostics [19], in this work, we use the line-integrated SXR emissivity measured by silicon diodes. This is part of the former SXR diagnostic, DTOMOX, developed for Tore Supra (TS) and still operational on WEST. DTOMOX has 82 lines-of-sight considered, among which 45 are horizontal and 37 are vertical [20]. The viewing geometry is depicted in Fig. 1. However, the vertical camera of DTOMOX has been dismantled for WEST due to the lack

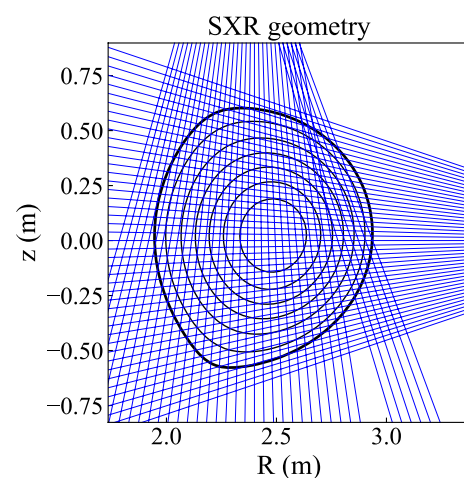


Fig. 1 The geometry of the soft X-ray diagnostic system DTOMOX

of space for plasma observation, as will be explained in the subsection "Test on Real Data".

In order to infer the local SXR emissivity in a poloidal cross-section, we first discretize the cross-section in 3600 pixels using a 60×60 grid, and then only consider the n pixels inside the last closed flux surface (LCFS) for tomography. The n emissivities are represented by a vector $\vec{\epsilon}_{\text{SXR}} = [\epsilon_{\text{SXR}}(\vec{r}_1), \epsilon_{\text{SXR}}(\vec{r}_2), \dots, \epsilon_{\text{SXR}}(\vec{r}_n)]$, where $\epsilon_{\text{SXR}}(\vec{r}_i)$ denotes the emissivity at location \vec{r}_i . Then, based on the line-of-sight approximation [21], the line-integrated emissivity along m viewing chords $\vec{d}_{\text{SXR}} = [d_{\text{SXR},1}, d_{\text{SXR},2}, \dots, d_{\text{SXR},m}]$ can be calculated using the forward model in a matrix form

$$\vec{d}_{\text{SXR}} = \bar{R} \vec{\epsilon}_{\text{SXR}}. \tag{3}$$

Here \bar{R} is the response matrix or transfer matrix describing how much each of the n pixels contributes to the signal recorded in each of the m channels.

Gaussian Process Tomography Method

The tomographic reconstruction of the local SXR emissivity at n points from m line integrals is essentially an ill-posed problem, as the number of unknowns n is much larger than m . A unique solution does not exist without additional regularizing assumptions. There are various ways to impose regularization. One of the most commonly used methods is Tikhonov regularization, where a regularization matrix imposes smoothness on the emissivity profiles, for instance by minimizing the Fisher information [22, 23]. In the Bayesian framework, regularization is built into the prior probability distribution, which in our case is a Gaussian process.

Non-stationary Gaussian Process Prior Distribution

Basically, a Gaussian process (GP) is a generalization of multivariate Gaussian distribution to a function space. A Gaussian process $f(\vec{r})$ describing the SXR emissivity over a poloidal cross-section is completely specified by a mean function $\mu(\vec{r})$ and a covariance function $k(\vec{r}, \vec{r}')$, evaluated at positions \vec{r} and \vec{r}' in the cross-section. Any finite number of function values $f(\vec{r}_i)$, i.e. SXR emissivities at pixel positions $\{\vec{r}_i\}$ jointly follow a multivariate Gaussian distribution [24]. The covariance function determines the correlation between any two points and therefore regularizes the problem of estimating the local emissivities by controlling the smoothness of the emissivity profile. A *stationary* covariance function (only depends on $\vec{r} - \vec{r}'$) assumes the same smoothness everywhere, an example of which is the squared exponential covariance function

$$k_{\text{SE}}(\vec{r}_i, \vec{r}_j) = \sigma_f^2 \exp\left(-\frac{\|\vec{r}_i - \vec{r}_j\|^2}{2l^2}\right). \tag{4}$$

Here the *amplitude* σ_f controls the level of variability and the characteristic *length scale* l controls the degree of smoothness, while $\|\vec{r}_i - \vec{r}_j\|$ denotes the Euclidean distance between \vec{r}_i and \vec{r}_j . In general, free parameters like σ_f and l that specify the prior distribution are called *hyperparameters*. In comparison, a nonstationary GP allows varying degrees of smoothness in different locations and is therefore more flexible. Both stationary [25, 26] and nonstationary [27, 28] Gaussian processes have been applied to tomography problems on several fusion devices. The Bayesian tomography method with Gaussian processes is also referred to as *Gaussian process tomography* [29].

To reconstruct emissivity profiles with spatially varying smoothness, we adopt a nonstationary covariance proposed by Gibbs [30]. Under an isotropic assumption of equal lengths scales in the radial (R) and vertical (z) directions ($l_R(\vec{r}) = l_z(\vec{r}) = l(\vec{r})$), the covariance function can be written as

$$k_{\text{NS}}(\vec{r}_i, \vec{r}_j) = \sigma_f^2 \frac{2l(\vec{r}_i)l(\vec{r}_j)}{l^2(\vec{r}_i) + l^2(\vec{r}_j)} \exp\left(-\frac{\|\vec{r}_i - \vec{r}_j\|^2}{l^2(\vec{r}_i) + l^2(\vec{r}_j)}\right), \tag{5}$$

where $l(\vec{r}_i)$ and $l(\vec{r}_j)$ are the length scales employed at spatial positions \vec{r}_i and \vec{r}_j . In order to incorporate information about the magnetic equilibrium, we assume a flux-dependent length scale. This is quite reasonable considering the dominant impurity transport along the flux surface, causing the emissivity field to be strongly correlated with the magnetic equilibrium [26]. In practice, we find that local length scales with linear dependence on the normalized poloidal flux ψ_N are sufficient to model the emissivity profile:

$$l(\psi_N) = (1 - \psi_N)l_c + \psi_N l_e. \tag{6}$$

In this expression, the length scales at the core ($l_c = l(0)$) and the edge ($l_e = l(1)$) are two hyperparameters. Merging all hyperparameters in a vector θ , the prior distribution of the SXR emissivity profile, which is conditional on the hyperparameters, is given by

$$p(\vec{\epsilon}_{\text{SXR}}|\vec{\theta}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{\text{prior}}|^{1/2}} \times \exp\left[-\frac{1}{2} (\vec{\epsilon}_{\text{SXR}} - \vec{\mu}_{\text{prior}})^\top \Sigma_{\text{prior}}^{-1} (\vec{\epsilon}_{\text{SXR}} - \vec{\mu}_{\text{prior}})\right]. \tag{7}$$

The prior mean vector $\vec{\mu}_{\text{prior}}$ is usually set to zero and the prior covariance matrix Σ_{prior} is determined by the covariance function, i.e., $\Sigma_{\text{prior}}[i, j] = k_{\text{NS}}(\vec{r}_i, \vec{r}_j)$.

Likelihood

The observations from different channels are usually assumed to be affected by independent Gaussian noise. Therefore the likelihood is a multivariate Gaussian distribution:

$$p(\vec{d}_{\text{SXR}}|\vec{\epsilon}_{\text{SXR}}, \vec{\theta}) = \frac{1}{(2\pi)^{\frac{m}{2}} |\Sigma_d|^{\frac{1}{2}}} \times \exp \left[-\frac{1}{2} (\vec{d}_{\text{SXR}} - \bar{R}\vec{\epsilon}_{\text{SXR}})^T \Sigma_d^{-1} (\vec{d}_{\text{SXR}} - \bar{R}\vec{\epsilon}_{\text{SXR}}) \right], \tag{8}$$

where the covariance Σ_d is a diagonal matrix given by $\Sigma_d = \text{diag}[\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2]$. Here, the standard deviation of each channel σ_i is assumed to depend on the signal level,

$$\sigma_i = \alpha \cdot d_{\text{SXR},i}, \tag{9}$$

where α here is a hyperparameter controlling the noise level.

Posterior Distribution

With a Gaussian process for the prior and a Gaussian likelihood, and the forward model being linear, the posterior $p(\vec{\epsilon}_{\text{SXR}}|\vec{d}_{\text{SXR}}, \vec{\theta})$ is also a Gaussian process. The mean vector and covariance matrix corresponding to a set of discrete pixels are therefore available in a closed form [29]:

$$\begin{aligned} \vec{\mu}_{\text{post}} &= \vec{\mu}_{\text{prior}} + (\bar{R}^T \Sigma_d^{-1} \bar{R} + \Sigma_{\text{prior}}^{-1})^{-1} \bar{R}^T \Sigma_d^{-1} (\vec{d}_{\text{SXR}} - \bar{R}\vec{\mu}_{\text{prior}}), \\ \Sigma_{\text{post}} &= (\bar{R}^T \Sigma_d^{-1} \bar{R} + \Sigma_{\text{prior}}^{-1})^{-1}. \end{aligned} \tag{10}$$

The posterior mean provides an estimate of the emissivity profile and the diagonal elements of the posterior covariance matrix are used to quantify the uncertainty of the inference.

Hyperparameter Optimization

The quality of the reconstructed emissivity profile strongly relies on the choice of hyperparameters $\vec{\theta} = [\alpha, \sigma_f, l_c, l_e]$. In principle, a fully Bayesian analysis also takes into account the uncertainty on the hyperparameters by marginalizing them out. This approach provides reliable estimate of uncertainty but is computationally expensive. As this fully Bayesian approach is computationally demanding, here we use an *empirical Bayes* approach, determining the hyperparameters from the data [31, 32]. This can be done for instance by maximizing the marginal likelihood

$$p(\vec{d}_{\text{SXR}}|\vec{\theta}) = \int p(\vec{d}_{\text{SXR}}|\vec{\epsilon}_{\text{SXR}}, \vec{\theta}) p(\vec{\epsilon}_{\text{SXR}}|\vec{\theta}) d\vec{\epsilon}_{\text{SXR}}. \tag{11}$$

Again, thanks to the Gaussian prior and likelihood, and the linear forward model, this integration is analytically tractable [24]. The logarithm of the marginal likelihood is

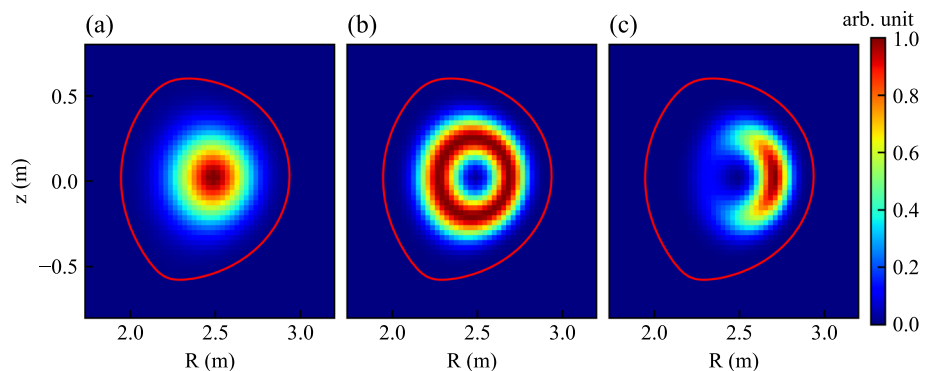
$$\begin{aligned} \ln [p(\vec{d}_{\text{SXR}}|\vec{\theta})] &= -\frac{m}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma_d + \bar{R}\Sigma_{\text{prior}}\bar{R}^T| \\ &\quad - \frac{1}{2} (\vec{d}_{\text{SXR}} - \bar{R}\vec{\mu}_{\text{prior}})^T (\Sigma_d + \bar{R}\Sigma_{\text{prior}}\bar{R}^T)^{-1} (\vec{d}_{\text{SXR}} - \bar{R}\vec{\mu}_{\text{prior}}). \end{aligned} \tag{12}$$

The hyperparameters obtained through maximization of Eq. (12) are then substituted into Eq. (10) to obtain the reconstructed emissivity distribution. This maximization can be done numerically with the help of the optimization algorithms implemented in the python package SciPy [33].

Validation on Synthetic Data

To validate our method, we created three different synthetic SXR emissivity profiles from a known magnetic equilibrium, with characteristic patterns of the Gaussian shape, the hollow shape and the poloidally asymmetric banana shape, as shown in Fig. 2. The phantom models used to generate the emissivity profiles were based on an a priori assumption of strong correlation between emissivity and magnetic flux. Synthetic line-integrated measurements were generated from these artificial emissivity profiles using all 82 lines-of-sight of DTOMOX, with 5% additive Gaussian noise. Then, Gaussian process tomography with the nonstationary covariance function in Eq. (5) was applied to reconstruct the original emissivity profiles from the artificial noisy measurements. Since the noise level of the synthetic data was already known, the covariance matrix of the measurements was simply $\Sigma_d = \text{diag}[(0.05d_{\text{SXR},1})^2, (0.05d_{\text{SXR},2})^2, \dots, (0.05d_{\text{SXR},m})^2]$. After optimizing the hyperparameters, the posterior mean and posterior covariance matrix of the emissivity profiles were immediately available according to Eq. (10). The reconstructions are shown in Fig. 3. Here the

Fig. 2 Synthetic SXR emissivity profiles with **a** Gaussian shape, **b** hollow shape and **c** banana shape. The red line represents the last closed flux surface



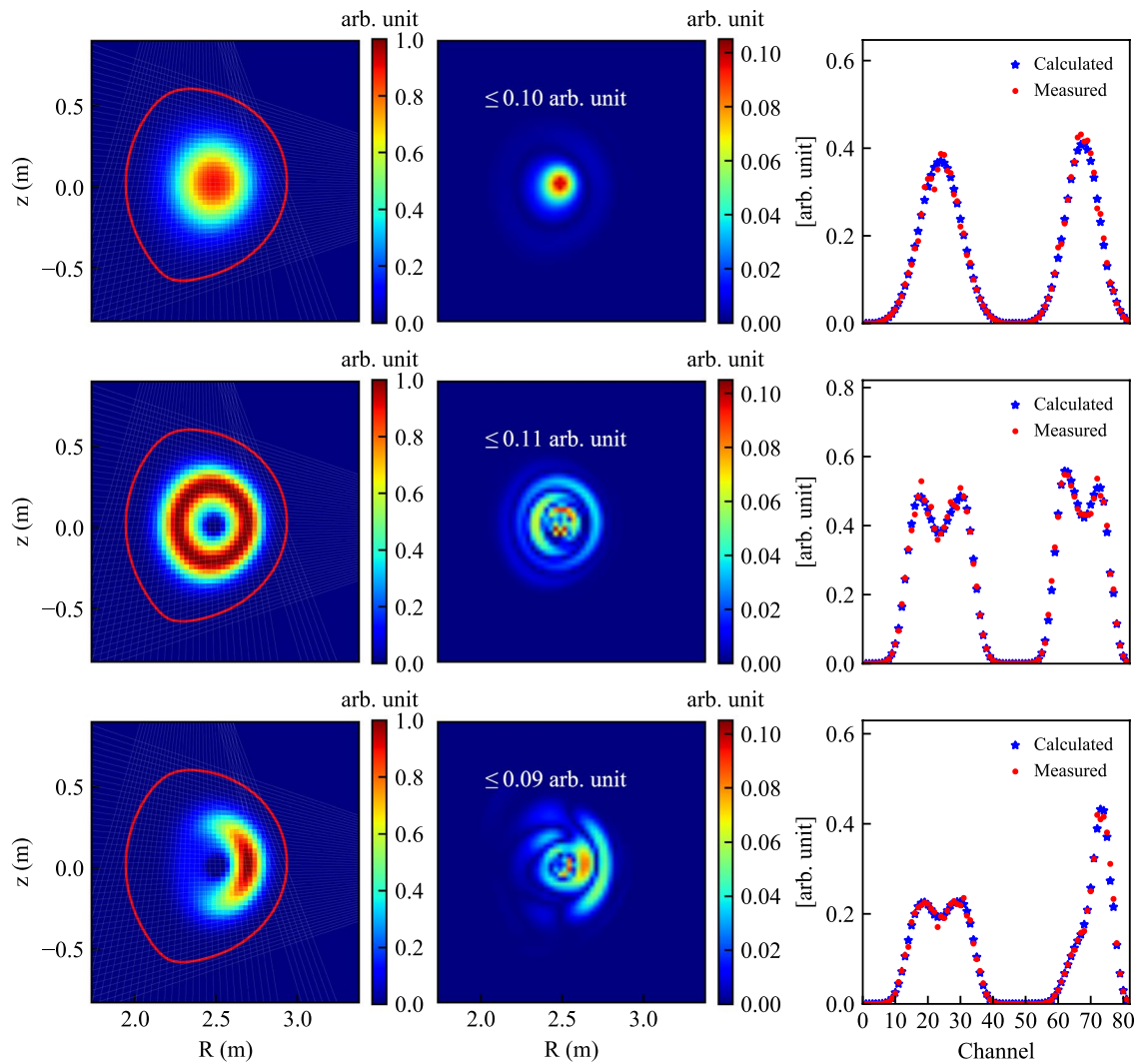


Fig. 3 Reconstructions of SXR emissivity profiles from synthetic data with 5% noise level. Left column: reconstructed emissivity $\epsilon_{\text{SXR}}^{\text{rec}}$. Middle column: error map defined as the absolute difference between

the true and the reconstructed emissivity $|\epsilon_{\text{SXR}}^{\text{rec}} - \epsilon_{\text{SXR}}^{\text{true}}|$. Right column: measured and calculated line-integrated emissivity. The white lines correspond to the lines-of-sight of the DTOMOX system

reconstruction error is defined as the absolute difference between the true and reconstructed emissivity. The characteristic shapes of all three emissivity profiles were well recovered by our Gaussian process tomography method, including the asymmetric banana shape, which is often challenging for conventional tomographic techniques. For all three reconstructions, the local errors are less than 0.05 arb. unit on most area of the poloidal section and the largest errors are below 0.12 arb. unit. The line integrals calculated from the reconstructed emissivity profiles also show good agreement with the synthetic measurements.

Further test has also been performed on synthetic data with a higher noise level of 10%. The results are illustrated in Fig. 4. Although the reconstruction error for the hollow and banana shapes gets larger with the increased

noises, the overall performance is still quite good in all three cases.

Test on Real Data

The application on real WEST data turned out to be challenging for several reasons. Due to the installation of an upper divertor, many vertically viewing diodes from DTOMOX are masked on WEST. Therefore, in practice only 45 horizontally viewing lines-of-sight are available [18], which poses difficulties for resolving certain spatial emissivity distributions with horizontal asymmetry (see appendix ‘Phantom Test with Only Horizontal Lines-of-Sight’). Moreover, the SXR data can also contain outliers caused by saturation or insufficient signal-to-noise ratio. Furthermore, we found

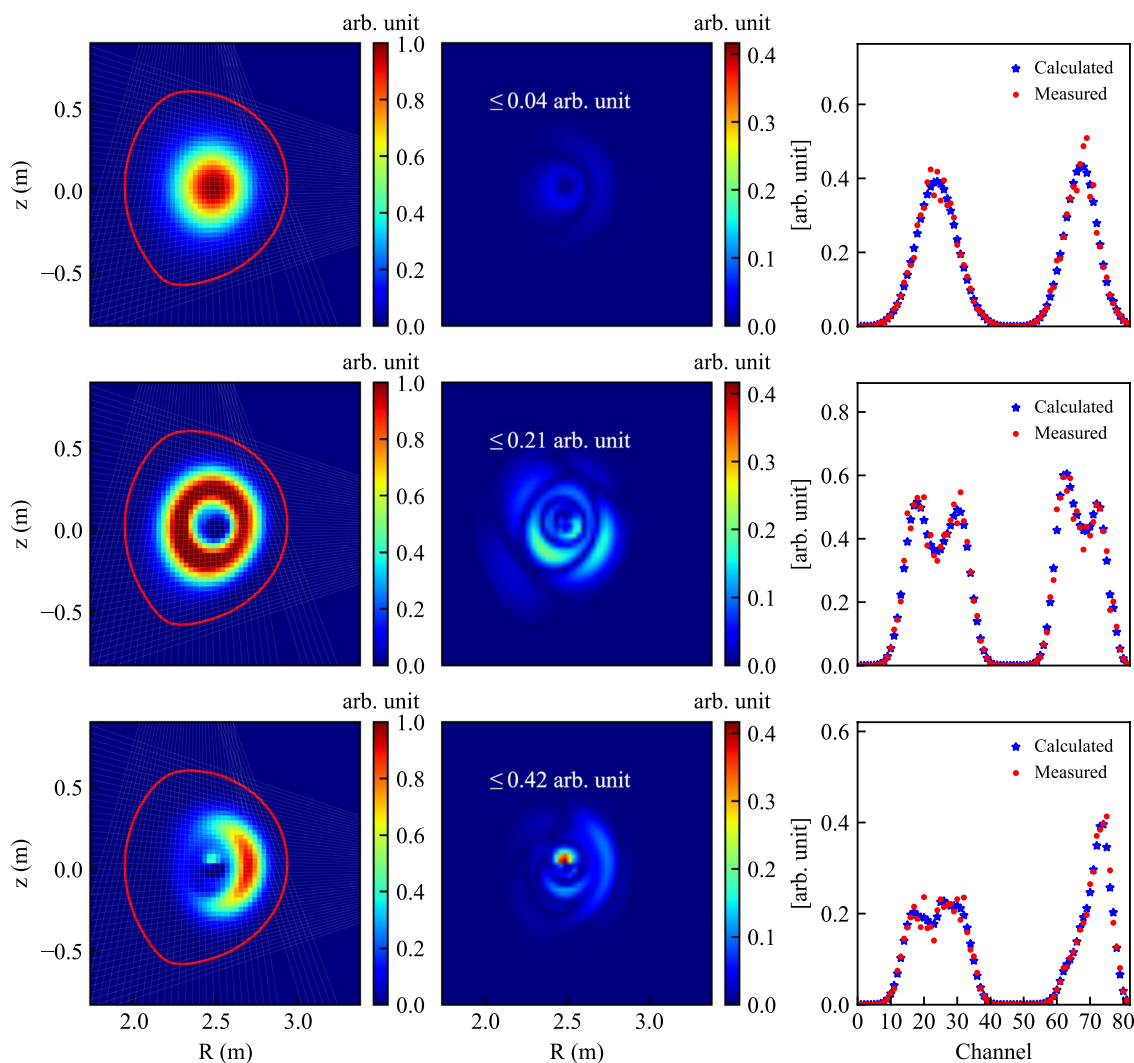


Fig. 4 Similar to Fig. 3, except the noise level of the synthetic line integrals is now 10%

that the error model described in Eq. (9) tends to underestimate the measurement uncertainty, especially for low-signal channels. Therefore, a new error model was used, consisting of background noise with variance $\sigma_{bg,i}^2$ (estimated without plasma), in addition to a proportional measurement error:

$$\sigma_i^2 = \sigma_{bg,i}^2 + (\alpha \cdot d_{SXR,i})^2. \tag{13}$$

This method has been applied to the SXR data from a real WEST pulse (#55191) with high n_e and T_e at the flat-top phase, for example $t = 1.40$ s. The result is shown in Fig. 5. Compared with the previous error model, the new model fits the measured line integrals better and significantly improves the posterior uncertainty by reducing the largest posterior standard deviation from $66.6 \text{ W} \cdot \text{m}^{-3}$ to $25.0 \text{ W} \cdot \text{m}^{-3}$. Without vertical lines-of-sight, the reconstructed emissivity profile hardly exhibits any horizontal asymmetry. The

double-lobed structure in the uncertainty profile Fig. 5 (f) also suggests the limits of having only the horizontal camera. Considering this lack of information, in the remainder we will consider 1D radial profiles $c_W(\rho)$, $n_e(\rho)$ and $T_e(\rho)$ instead of their 2D distributions.

Integrated Analysis of Tungsten Concentration

SXR Emissivity and Cooling Factors

The measured SXR emissivity from an ion species S in the plasma, ϵ_S (in $\text{W} \cdot \text{m}^{-3}$), is usually modelled using the cooling factor (filtered by the spectral response of the detector) L_S (in $\text{W} \cdot \text{m}^3$): $\epsilon_S = n_e \cdot n_S \cdot L_S$, with n_e the electron density and n_S the density of S. The cooling factor, defined as the energy

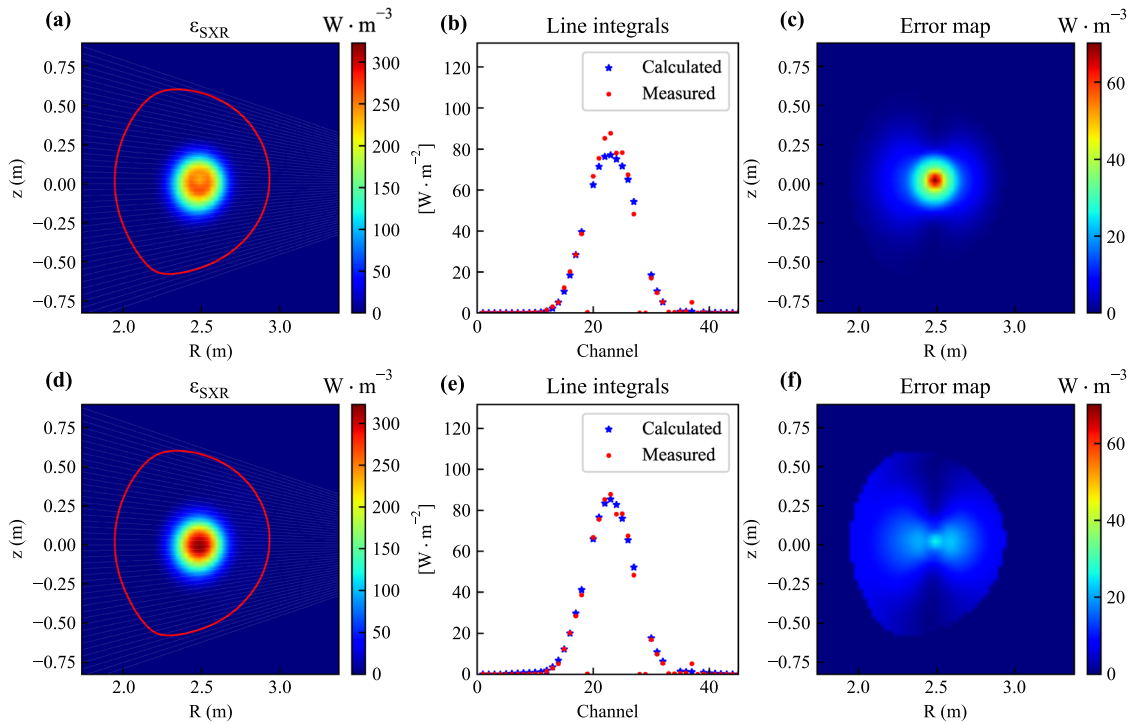


Fig. 5 Reconstructed SXR emissivity, line-integrated emissivity and error map (given by the posterior standard deviation) for WEST pulse #55191, $t = 1.40$ s, without **a–c** and with **d–f** background noise considered

loss rate per free electron per ion due to radiative cooling, is also called the radiative cooling rate [34] or the radiation loss parameter [35]. It can be calculated theoretically using atomic physics [2, 36]. In a hydrogenic plasma with dominant tungsten impurities at trace concentration ($c_W = n_W/n_e \ll 1$), the detected SXR emissivity is approximately the sum of contributions from the main ion species (denoted here as H) and tungsten:

$$\epsilon_{SXR} = n_e^2 [L_H(T_e) + c_W \cdot L_W(T_e)]. \tag{14}$$

The cooling factors of deuterium and tungsten are calculated using atomic data from the OPEN-ADAS database [37]. They depend primarily on the electron temperature T_e , as shown in Fig. 6. Under the above assumptions, the tungsten concentration can in principle be estimated on the basis of measurements of the SXR emissivity, n_e and T_e .

Integrated Data Analysis for c_W

With the SXR emissivity profile reconstructed using Gaussian process tomography, the electron density profile measured by interferometry (INT) and the electron temperature profile measured by electron cyclotron emission (ECE), the tungsten concentration can directly be estimated by reformulating Eq. (14):

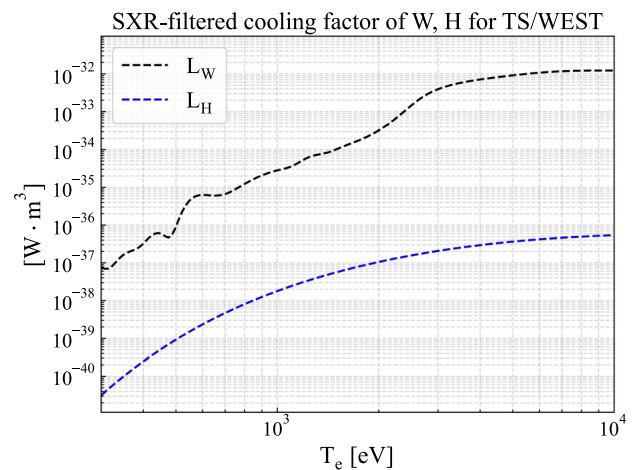


Fig. 6 T_e dependence of W and H cooling factors filtered by the SXR spectral response

$$c_W = \frac{\epsilon_{SXR}}{n_e^2 L_W(T_e)} - \frac{L_H(T_e)}{L_W(T_e)}. \tag{15}$$

Such direct estimation of W concentration based on the tomographic reconstruction of plasma emission and the T_e dependent cooling factors has been applied to several devices, for example JET [38] and HL-2A [6]. However, this calculation does not take into account the sources of

uncertainty from individual diagnostics or their interdependencies. In order to treat the data in an integrated way, we follow an integrated data analysis (IDA) approach using Bayesian probability theory [7]. IDA considers the joint posterior distribution of \vec{c}_W , \vec{n}_e and \vec{T}_e :

$$p(\vec{c}_W, \vec{n}_e, \vec{T}_e | \vec{d}_{SXR}, \vec{d}_{INT}, \vec{d}_{ECE}) \propto p(\vec{d}_{SXR} | \vec{c}_W, \vec{n}_e, \vec{T}_e) p(\vec{d}_{INT} | \vec{n}_e) p(\vec{d}_{ECE} | \vec{T}_e) p(\vec{c}_W) p(\vec{n}_e) p(\vec{T}_e). \tag{16}$$

The prior distributions of c_W , n_e and T_e are modeled by three stationary Gaussian processes (so far with fixed hyperparameters). An additional constraint of zero gradient at the core $c'_W(0) = 0$ is imposed on the W concentration profile by adding an artificial ‘observation’ [31]. The likelihoods are also assumed to be Gaussian in analogy with the distribution

expressed in Eq. (8), with the following corresponding forward models:

$$\begin{aligned} \vec{d}_{SXR} &= \bar{\bar{R}}_{SXR} \vec{\epsilon}_{SXR}(\vec{n}_e, \vec{T}_e, \vec{c}_W), \\ \vec{d}_{INT} &= \bar{\bar{R}}_{INT} \vec{n}_e, \\ \vec{d}_{ECE} &= \vec{T}_e. \end{aligned} \tag{17}$$

The interferometry system measures the line-integrated electron density and the geometry of the 10-channel WEST interferometry is illustrated in Fig. 7. Its forward model is linear and can be expressed in a matrix form with the help of a response matrix $\bar{\bar{R}}_{INT}$. Under the blackbody assumption and considering the plasma being optically thick, the Rayleigh-Jeans approximation tells us that the intensity of the electron cyclotron radiation is proportional to the electron temperature and the forward model of the ECE system should also be linear. However, in the WEST ECE system, empirical corrections are required during data processing in order to yield reliable temperature estimates. This makes it relatively difficult to specify a detailed forward model, which is why we use the (validated) T_e measurements from ECE as the raw data \vec{d}_{ECE} . Nevertheless, due to the nonlinear forward model used in $p(\vec{d}_{SXR} | \vec{c}_W, \vec{n}_e, \vec{T}_e)$, the posterior distribution does not have an analytical expression and needs to be approximated numerically with MCMC sampling methods.

Validation on Synthetic Data

For validation purposes, we generated synthetic data for ECE, interferometry and SXR diagnostics, starting from an artificial Gaussian-shaped radial tungsten concentration profile with a peak value of $c_W = 10^{-4}$ at the core, and radial profiles of density and temperature taken from real measurements at WEST. These synthetic diagnostic data were then used as the input for IDA, and samples of c_W , n_e and T_e were drawn from the joint posterior distribution in Eq. (16). An example result is shown in Fig. 8. Here, the reconstructed

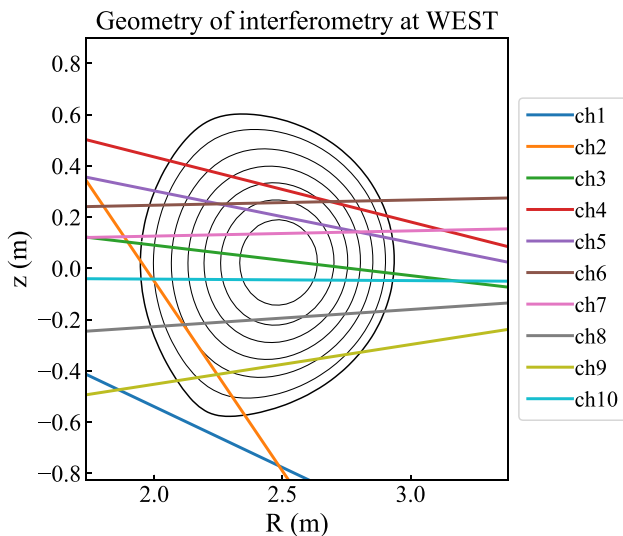
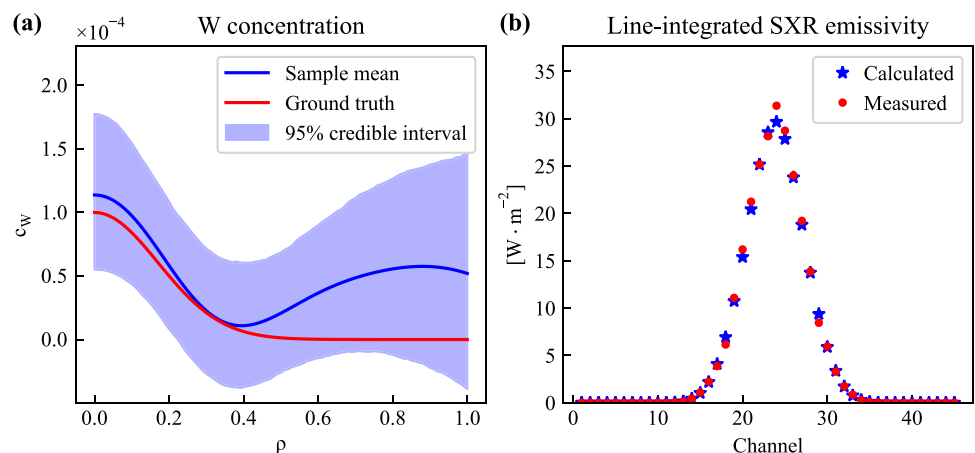


Fig. 7 The beam chord trajectories of the WEST interferometry system

Fig. 8 a W concentration estimated from synthetic data and b re-integrated SXR emissivity



W concentration profile (blue) is estimated by taking the average of posterior samples. It matches the original profile (red) well in the center of the plasma and the re-integrated emissivity is in good agreement with the synthetic measurements. However, further outside the central plasma the uncertainty of the reconstructed c_W quickly grows, because the measured SXR emissivity is less sensitive to the tungsten concentration at low n_e and T_e . In other words, the SXR measurements provide little information on the edge tungsten concentration.

Test on Real Data

Despite the difficulties with reconstruction of the W concentration profiles outside the very core plasma, a preliminary test was carried out with the same method using real data from WEST. This yielded a reasonable estimate of the core tungsten concentration (Fig. 9), although further outside the decreasing trend of c_W is not realistic. As the tungsten impurity at the edge has little impact on the measured SXR emissivity, the reconstructed tungsten concentration profile is quite sensitive to the choices of hyperparameters. In fact, the inferred tungsten concentration profile can have quite a

different trend towards the edge under another set of hyperparameters, even if the measured data remain unchanged. An example tungsten concentration profile inferred under a different set of hyperparameters is shown in Fig. 10. In addition, the slight mismatch between the calculated and measured line-integrated SXR emissivity may arise from uncertainty on the magnetic equilibrium, combined with the rather strong assumption of constant W concentration on each flux surface.

Preliminary Exploration on Speeding Up the Inference with Surrogate Models

A fully Bayesian inference with MCMC sampling is computationally expensive and does not lend itself to a potential real-time implementation. A possible approach for speeding up the inference is using neural network surrogate models trained on synthetic data [39]. As a first step, we investigate the application on a single diagnostic, the interferometry, for the reconstruction of electron density profiles. Preliminary results are obtained with a single-hidden-layer neural network called extreme learning machine [40], based

Fig. 9 **a** Estimated W concentration for WEST pulse #55191, $t = 1.40$ s and **b** re-integrated SXR emissivity, when $\sigma_{f,c_W} = 0.56$ and $l_{c_W} = 0.4$

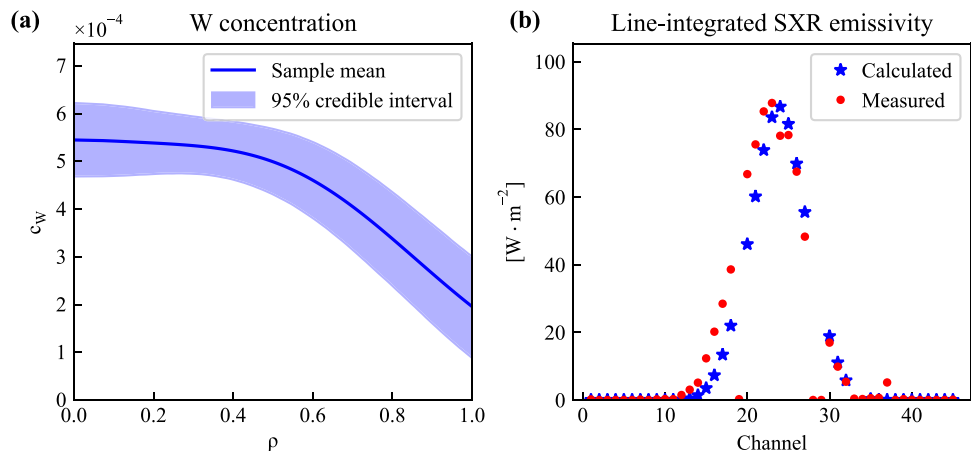
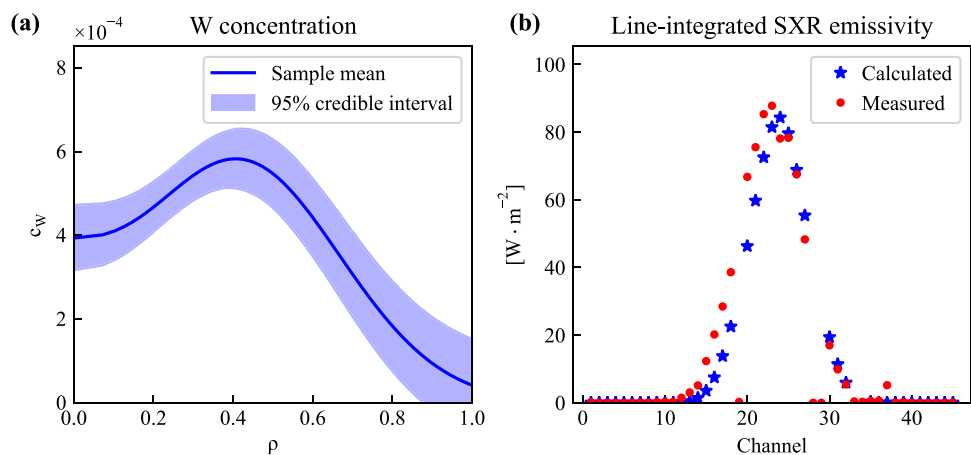


Fig. 10 **a** Estimated W concentration for WEST pulse #55191, $t = 1.40$ s and **b** re-integrated SXR emissivity, when $\sigma_{f,c_W} = 0.56$ and $l_{c_W} = 0.25$



on the implementation by a python package PyRCN [41]. The input-to-hidden layer weights of an extreme learning machine are initialized randomly and left untrained, while the hidden-to-output weights are determined using linear regression, therefore the training is very fast. The model was trained on 165,000 synthetic line-integrated density measurements corresponding to realistic n_e profiles sampled from a prior distribution (Fig. 11) and evaluated on real data from interferometry on WEST (Fig. 12). The training took about 2.5 s on a normal laptop with Core i7 CPU and the evaluation time for a single input is shorter than 1 ms. The good agreement of the density profile inferred by the neural network and that calculated by the equilibrium code NICE [42] demonstrates the potential of the surrogate model. However, further validations on a larger scale are still required. Ultimately, this approach will be extended to the joint estimation of impurity concentration, density and temperature profiles.

Summary

In this paper, tomographic reconstruction of SXR emissivity profiles in tokamak plasmas has been demonstrated using nonstationary Gaussian processes in a Bayesian probabilistic approach. The validation on synthetic data

generated from emissivity profiles of several different characteristic shapes verifies the robustness of this method for different noise levels. By incorporating information about the magnetic equilibrium, the reconstruction from synthetic data of only horizontal lines-of-sight also gives reliable estimates for emissivity profiles with limited asymmetry. This was supported by an application to real WEST SXR data: using a realistic error model, the reconstruction of SXR emissivity profiles on WEST achieved relatively good performance, with the advantage of reliable uncertainty estimates. Moreover, the IDA approach was followed for joint estimation of tungsten concentrations and kinetic profiles from heterogeneous diagnostics on WEST. This allowed a realistic estimate of the tungsten concentration in the core plasma, but revealed difficulties for reconstructing the full profile based on these measurements alone. Due to the lack of information from low ionization states of tungsten, which primarily radiate outside the observed SXR range, the edge tungsten concentration inferred from SXR is not reliable. As a result, the estimated tungsten concentration profile is quite sensitive to the choice of hyperparameters. Therefore, in a next step we intend to incorporate additional diagnostics, such as bolometry. In addition, we aim to extend neural network

Fig. 11 Training set consisting of (left) density profiles sampled from the prior distribution and (right) synthetic noisy line-integrated measurements

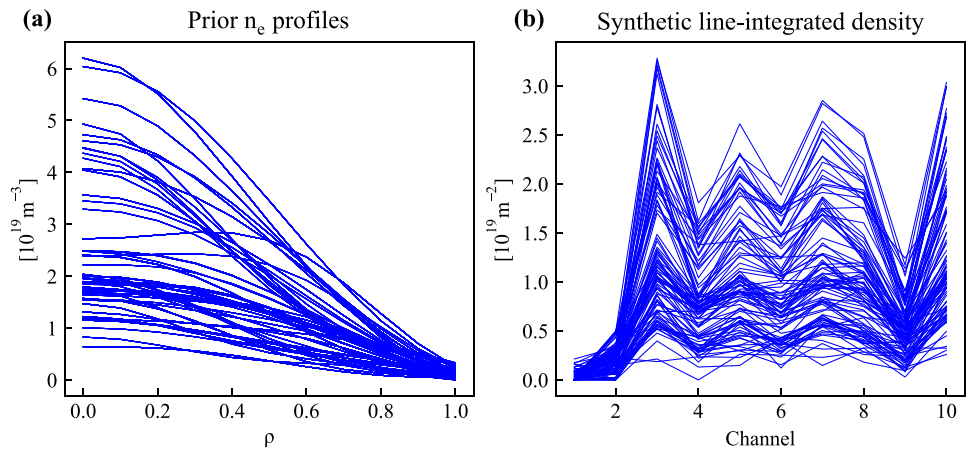
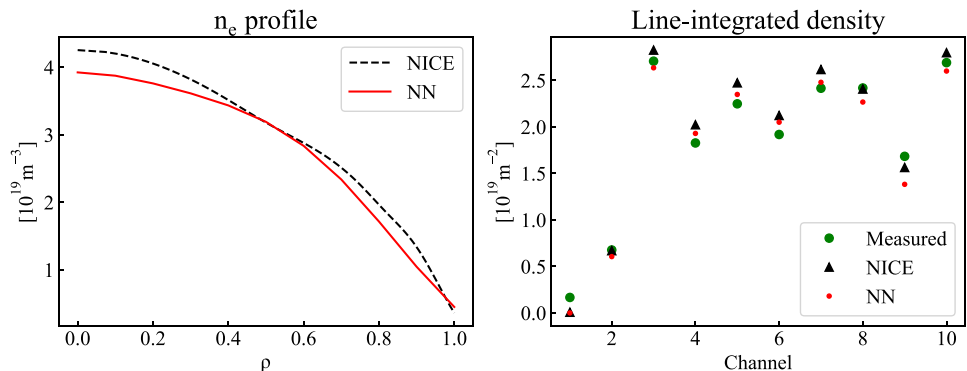


Fig. 12 Left: density profiles reconstructed by the neural network (red) and by the equilibrium code NICE (black) for WEST pulse #53259 at $t = 4$ s. Right: comparison of calculated and measured line-integrated densities



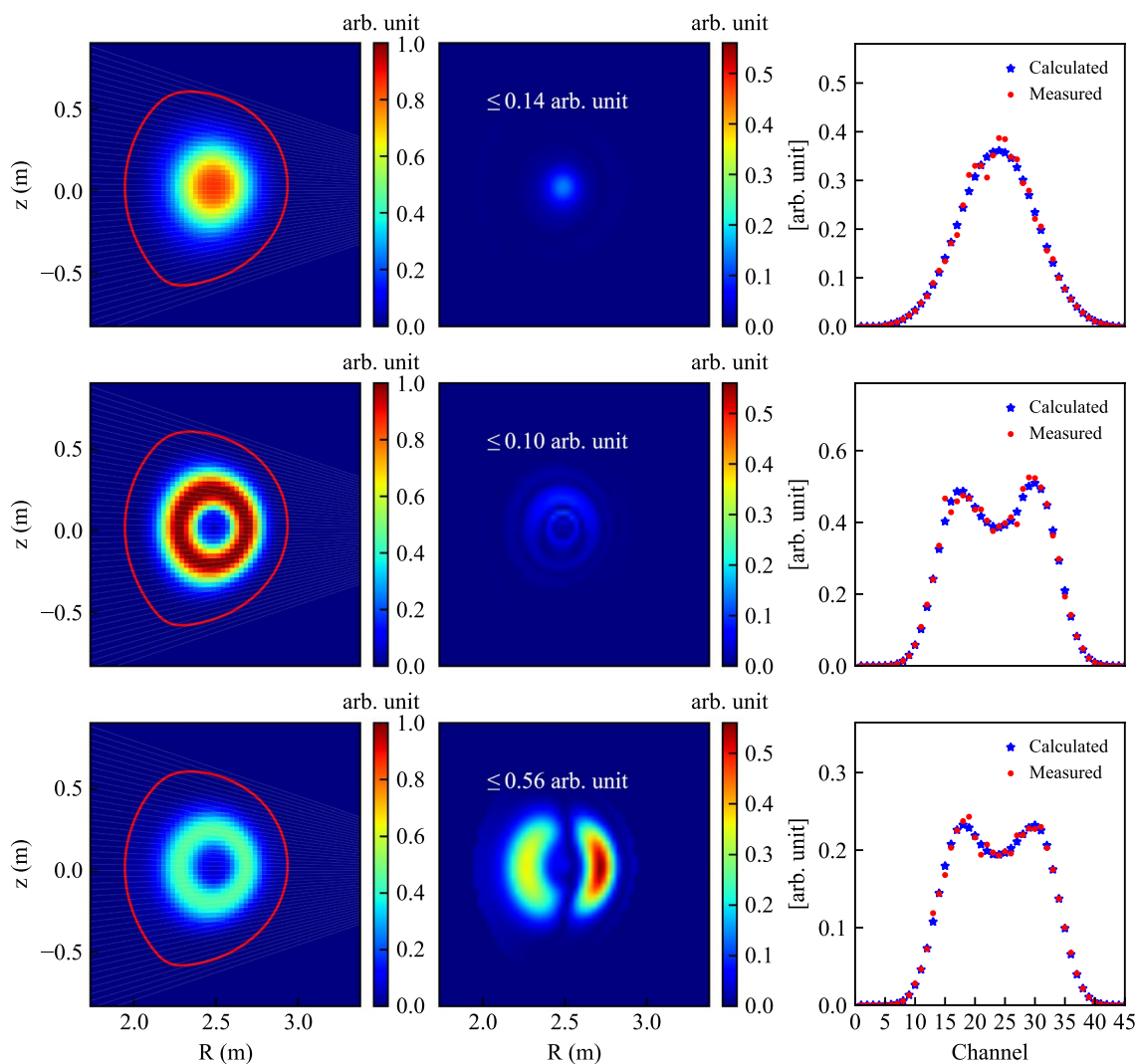


Fig. 13 Reconstructions of SXR emissivity profiles from synthetic SXR data with only horizontal lines-of-sight, for three different phantoms (from top to bottom: Gaussian, hollow and banana-like shapes). From left to right: reconstructed emissivity, error map and line integrated data

surrogate modeling to the full inference process of the tungsten concentration profile and the kinetic profiles.

Appendix: phantom test with only horizontal lines-of-sight

In order to illustrate the difficulty of reconstructing horizontally asymmetric emissivity profiles with only horizontal lines-of-sight at WEST, we also perform a test using synthetic data generated from the three different phantom shapes mentioned previously. As shown in Fig. 13, with the information from magnetic equilibrium, the reconstruction for symmetric Gaussian and hollow shapes are still quite satisfactory. However, comparing the synthetic data corresponding to the asymmetric banana-shaped and

the hollow-shaped profiles, the horizontal camera registers lower signals from the banana-shaped emissivity distribution, but it does not capture the differences arising from the asymmetry. Therefore the Gaussian process tomography method is not able to restore the asymmetry, to distinguish between the hollow and the banana shapes with these horizontal lines-of-sight alone.

Acknowledgements This work is supported by China Scholarship Council affiliated with the Ministry of Education of the P.R. China (File No. 202008510111) and Natural Science Foundation of Sichuan Province, China under Grant No. 24NSFSC7404.

Author contributions HW wrote the main manuscript text. GV made important revision to the manuscript. DM provided the data source. AJ calculated the cooling factors of tungsten and hydrogen. All authors reviewed the manuscript.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest Not applicable

References

1. M. Missirlian, J. Bucalossi, Y. Corre, F. Ferlay, M. Firdaouss, P. Garin, A. Grosman, D. Guilhem, J. Gunn, P. Languille et al., The WEST project: current status of the ITER-like tungsten divertor. *Fusion Eng. Des.* **89**(7–8), 1048–1053 (2014). <https://doi.org/10.1016/j.fusengdes.2014.01.050>
2. T. Pütterich, R. Neu, R. Dux, A. Whiteford, M. O'Mullane, H. Summers, A.U. Team et al., Calculation and experimental test of the cooling factor of tungsten. *Nucl. Fusion* **50**(2), 025012 (2010). <https://doi.org/10.1088/0029-5515/50/2/025012>
3. J. Bucalossi, 9 - Tore Supra WEST, in *Magnetic Fusion Energy*. ed. by G.H. Neilson (Woodhead Publishing, Cambridge, 2016), pp.261–293. <https://doi.org/10.1016/B978-0-08-100315-2.00009-X>
4. R. Guirlet, C. Desgranges, J. Schwob, P. Mandelbaum, M. Boumendjel, W. Team et al., Extreme UV spectroscopy measurements and analysis for tungsten density studies in the WEST tokamak. *Plasma Phys. Controlled Fusion* **64**(10), 105024 (2022). <https://doi.org/10.1088/1361-6587/ac8d2c>
5. T. Nakano, A. Shumack, C. Maggi, M. Reinke, K. Lawson, I. Coffey, T. Pütterich, S. Brezinsek, B. Lipschultz, G. Matthews et al., Determination of tungsten and molybdenum concentrations from an X-ray range spectrum in JET with the ITER-like wall configuration. *J. Phys. B: At. Mol. Opt. Phys.* **48**(14), 144023 (2015). <https://doi.org/10.1088/0953-4075/48/14/144023>
6. T. Wang, B. Li, J. Gao, W. Zhong, H. Li, Z. Yang, J. Min, K. Fang, G. Xiao, Y. Zhu et al., Monitoring of two-dimensional tungsten concentration profiles on the HL-2A tokamak. *Plasma Phys. Controlled Fusion* **64**(8), 084003 (2022). <https://doi.org/10.1088/1361-6587/ac77b9>
7. R. Fischer, C. Fuchs, B. Kurzan, W. Suttrop, E. Wolftrum, A..U. Team, Integrated data analysis of profile diagnostics at ASDEX Upgrade. *Fusion Sci. Technol.* **58**(2), 675–684 (2010). <https://doi.org/10.13182/FST10-110>
8. R. Fischer, A. Dinklage, E. Pasch, Bayesian modelling of fusion diagnostics. *Plasma Phys. Controlled Fusion* **45**(7), 1095 (2003). <https://doi.org/10.1088/0741-3335/45/7/304>
9. B.P. Van Milligen, T. Estrada, E. Ascasibar, D. Tafalla, D. López-Bruna, A.L. Fraguas, J. Jiménez, I. García-Cortés, A. Dinklage, R. Fischer, Integrated data analysis at TJ-II: The density profile. *Rev. Sci. Instrum.* (2011). <https://doi.org/10.1063/1.3608551>
10. S. Kwak, J. Svensson, S. Bozhnikov, J. Flanagan, M. Kempnaars, A. Boboc, Y.-C. Ghim, J. Contributors, Bayesian modelling of Thomson scattering and multichannel interferometer diagnostics using Gaussian processes. *Nucl. Fusion* **60**(4), 046009 (2020). <https://doi.org/10.1088/1741-4326/ab686e>
11. P. Wenan, W. Tianbo, W. Zhibin, Y. Yonghao, W. Hao, G. Verdoolaege, Y. Zengchen, L. Chunhua, G. Wenping, L. Bingli et al., Integrated data analysis on the electron temperature profile of HL-2A with the Bayesian probability inference method. *Plasma Sci. Technol.* **24**(5), 055601 (2022). <https://doi.org/10.1088/2058-6272/ac5c25>
12. J. Chen, Z. Wang, T. Wang, Y. Yang, H. Wu, Y. Li, G. Xiao, G. Verdoolaege, D. Mazon, Z. Yang et al., Integrated data analysis on the electron density profile of HL-2A with the Bayesian probability inference method. *Plasma Phys. Controlled Fusion* **65**(5), 055027 (2023). <https://doi.org/10.1088/1361-6587/acc60e>
13. G. Verdoolaege, R. Fischer, G. Van Oost, J.-E. Contributors et al., Potential of a Bayesian integrated determination of the ion effective charge via bremsstrahlung and charge exchange spectroscopy in tokamak plasmas. *IEEE Trans. Plasma Sci.* **38**(11), 3168–3196 (2010). <https://doi.org/10.1109/TPS.2010.2071884>
14. S. Rathgeber, R. Fischer, S. Fietz, J. Hobirk, A. Kallenbach, H. Meister, T. Pütterich, F. Ryter, G. Tardini, E. Wolftrum et al., Estimation of profiles of the effective ion charge at ASDEX Upgrade with integrated data analysis. *Plasma Phys. Controlled Fusion* **52**(9), 095008 (2010). <https://doi.org/10.1088/0741-3335/52/9/095008>
15. S. Kwak, U. Hergenahn, U. Höfel, M. Krychowiak, A. Pavone, J. Svensson, O. Ford, R. König, S. Bozhnikov, G. Fuchert et al., Bayesian inference of spatially resolved Zeff profiles from line integrated bremsstrahlung spectra. *Rev. Sci. Instrum.* (2021). <https://doi.org/10.1063/5.0043777>
16. S. Kwak, J. Svensson, O. Ford, L. Appel, Y.-C. Ghim, J. Contributors, Bayesian inference of axisymmetric plasma equilibrium. *Nucl. Fusion* **62**(12), 126069 (2022). <https://doi.org/10.1088/1741-4326/ac9c19>
17. D. Sivia, J. Skilling, *Data Analysis: A Bayesian Tutorial* (Oxford University Press, Oxford, 2006)
18. D. Mazon, M. Chernyshova, G. Jiolat, T. Czarski, P. Malard, E. Kowalska-Strzeciwiłk, S. Jablonski, W. Figacz, R. Zagorski, M. Kubkowska et al., Design of soft-x-ray tomographic system in WEST using GEM detectors. *Fusion Eng. Des.* **96**, 856–860 (2015). <https://doi.org/10.1016/j.fusengdes.2015.03.052>
19. M. Chernyshova, D. Mazon, K. Malinowski, T. Czarski, I. Ivanova-Stanik, S. Jabłoński, A. Wojeński, E. Kowalska-Strzeciwiłk, K.T. Poźniak, P. Malard et al., First exploitation results of recently developed SXR GEM-based diagnostics at the WEST project. *Nuclear Mater. Energy* **25**, 100850 (2020). <https://doi.org/10.1016/j.nme.2020.100850>
20. D. Mazon, D. Vezinet, D. Pacella, D. Moreau, L. Gabelieri, A. Romano, P. Malard, J. Mlynar, R. Masset, P. Lotte, Soft x-ray tomography for real-time applications: present status at Tore Supra and possible future developments. *Rev. Sci. Instrum.* **83**(6), 063505 (2012). <https://doi.org/10.1063/1.4730044>
21. L. Ingesson, C. Maggi, R. Reichle, Characterization of geometrical detection-system properties for two-dimensional tomography. *Rev. Sci. Instrum.* **71**(3), 1370–1378 (2000). <https://doi.org/10.1063/1.1150466>
22. M. Odstrcil, J. Mlynar, T. Odstrcil, B. Alper, A. Murari, J.E. Contributors, Modern numerical methods for plasma tomography optimisation. *Nucl. Instrum. Methods Phys. Res., Sect. A* **686**, 156–161 (2012). <https://doi.org/10.1016/j.nima.2012.05.063>
23. J. Mlynar, M. Tomes, M. Imrisek, B. Alper, M. O'Mullane, T. Odstrcil, T. Puetterich et al., Soft x-ray tomographic reconstruction of JET ILW plasmas with tungsten impurity and different spectral response of detectors. *Fusion Eng. Des.* **96**, 869–872 (2015). <https://doi.org/10.1016/j.fusengdes.2015.04.055>
24. C.E. Rasmussen, C.K.I. Williams et al., *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, 2006)
25. T. Wang, D. Mazon, J. Svensson, D. Li, A. Jardin, G. Verdoolaege, Gaussian process tomography for soft x-ray spectroscopy at WEST without equilibrium information. *Rev. Sci. Instrum.* **89**(6), 063505 (2018). <https://doi.org/10.1063/1.5023162>
26. T. Wang, D. Mazon, J. Svensson, D. Li, A. Jardin, G. Verdoolaege, Incorporating magnetic equilibrium information in Gaussian process tomography for soft x-ray spectroscopy at WEST. *Rev. Sci. Instrum.* **89**(10), 10–103 (2018). <https://doi.org/10.1063/1.5039152>

27. D. Li, J. Svensson, H. Thomsen, F. Medina, A. Werner, R. Wolf, Bayesian soft x-ray tomography using non-stationary Gaussian processes. *Rev. Sci. Instrum.* **84**(8), 083506 (2013). <https://doi.org/10.1063/1.4817591>
28. K. Moser, A. Bock, P. David, M. Bernert, R. Fischer, A.U. Team et al., Gaussian process tomography at ASDEX Upgrade with magnetic equilibrium information and nonstationary kernels. *Fusion Sci. Technol.* **78**(8), 607–616 (2022). <https://doi.org/10.1080/15361055.2022.2072659>
29. J. Svensson, A. Werner, J.-E. Contributors et al., Current tomography for axisymmetric plasmas. *Plasma Phys. Controlled Fusion* **50**(8), 085002 (2008). <https://doi.org/10.1088/0741-3335/50/8/085002>
30. M.N Gibbs, Bayesian Gaussian processes for regression and classification. PhD thesis, University of Cambridge (1998)
31. M. Chilenski, M. Greenwald, Y. Marzouk, N. Howard, A. White, J. Rice, J. Walk, Improved profile fitting and quantification of uncertainty in experimental measurements of impurity transport coefficients using gaussian process regression. *Nucl. Fusion* **55**(2), 023012 (2015). <https://doi.org/10.1088/0029-5515/55/2/023012>
32. J. Leddy, S. Madireddy, E. Howell, S. Kruger, Single Gaussian process method for arbitrary tokamak regimes with a statistical analysis. *Plasma Phys. Controlled Fusion* **64**(10), 104005 (2022). <https://doi.org/10.1088/1361-6587/ac89ab>
33. ...P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, Í Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0 contributors: SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020). <https://doi.org/10.1038/s41592-019-0686-2>
34. D.E. Post, R. Jensen, C. Tarter, W. Grasberger, W. Lokke, Steady-state radiative cooling rates for low-density, high-temperature plasmas. *At. Data Nucl. Data Tables* **20**(5), 397–439 (1977). [https://doi.org/10.1016/0092-640X\(77\)90026-2](https://doi.org/10.1016/0092-640X(77)90026-2)
35. R. Neu, R. Dux, A. Kallenbach, T. Pütterich, M. Balden, J. Fuchs, A. Herrmann, C. Maggi, M. O'Mullane, R. Pugno et al., Tungsten: an option for divertor and main chamber plasma facing components in future fusion devices. *Nucl. Fusion* **45**(3), 209 (2005). <https://doi.org/10.1088/0029-5515/45/3/007>
36. T. Pütterich, R. Neu, R. Dux, A. Whiteford, M. O'Mullane, t ASDEX Upgrade Team et al., Modelling of measured tungsten spectra from ASDEX Upgrade and predictions for ITER. *Plasma Phys. Controlled Fusion* **50**(8), 085016 (2008). <https://doi.org/10.1088/0741-3335/50/8/085016>
37. The Open ADAS Project (Atomic Data and Analysis Structure). <https://open.adas.ac.uk>
38. C. Angioni, P. Mantica, T. Pütterich, M. Valisa, M. Baruzzo, E. Belli, P. Belo, F. Casson, C. Challis, P. Drewelow et al., Tungsten transport in JET H-mode plasmas in hybrid scenario, experimental observations and modelling. *Nucl. Fusion* **54**(8), 083028 (2014). <https://doi.org/10.1088/0029-5515/54/8/083028>
39. A. Pavone, J. Svensson, A. Langenberg, U. Höfel, S. Kwak, N. Pablant, R. Wolf et al., Neural network approximation of Bayesian models for the inference of ion and electron temperature profiles at W7-X. *Plasma Phys. Controlled Fusion* **61**(7), 075012 (2019). <https://doi.org/10.1088/1361-6587/ab1d26>
40. G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications. *Neurocomputing* **70**(1–3), 489–501 (2006). <https://doi.org/10.1016/j.neucom.2005.12.126>
41. P. Steiner, A. Jalalvand, S. Stone, P. Birkholz, PyRCN: a toolbox for exploration and application of reservoir computing networks. *Eng. Appl. Artif. Intell.* **113**, 104964 (2022). <https://doi.org/10.1016/j.engappai.2022.104964>
42. B. Faugeras, An overview of the numerical methods for tokamak plasma equilibrium computation implemented in the NICE code. *Fusion Eng. Des.* **160**, 112020 (2020). <https://doi.org/10.1016/j.fusengdes.2020.112020>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.