Bayesian modelling of Thomson scattering and multichannel interferometer diagnostics using Gaussian processes

Sehyun Kwak[1,2†]**, J. Svensson**[2]**, S. Bozhenkov**[2]**, J. Flanagan**[3]**,
M. Kempenaars**[3,4]**, A. Boboc**[3]**, Y.-c. Ghim**[1*] **and
JET Contributors**‡

EUROfusion Consortium, JET, Culham Science Centre, Abingdon, OX14 3DB,
United Kingdom of Great Britain and Northern Ireland
[1]Department of Nuclear and Quantum Engineering, KAIST, Daejeon 34141,
Korea, Republic of
[2]Max-Planck-Institut für Plasmaphysik, 17491 Greifswald, Germany
[3]Culham Centre for Fusion Energy, Culham Science Centre, Abingdon, OX14 3DB,
United Kingdom of Great Britain and Northern Ireland
[4]ITER Organization, Route de Vinon-sur-Verdon, CS 90 046, 13067 St Paul Lez
Durance Cedex, France

E-mail: †slayer313@kaist.ac.kr,*ycghim@kaist.ac.kr

Abstract. Electron temperature and density profiles consistent with JET high resolution Thomson scattering (HRTS) and far infrared (FIR) interferometer data are inferred by a Bayesian joint model using Gaussian processes. Forward models predicting diagnostic data including instrument effects such as optics and electronics are developed independently for both diagnostic systems in the Minerva framework, and combined as one joint model. The full joint posterior distribution of the electron temperature and density profiles, the hyperparameters of the Gaussian processes and calibration factor is explored by Markov chain Monte Carlo (MCMC) sampling. The posterior distribution of the electron density(temperature) profile is obtained by marginalising all the possible combinations of the electron temperature(density) profile, the hyperparameters and calibration factor. Therefore, the method removes profile dependency on the hyperparameters completely in addition to eliminating often-used avoidable constraints such as monotonicity and parametrisation on the profiles. The posterior distribution of the calibration factor is also calculated explaining both the HRTS and the FIR interferometer data simultaneously. Thus, absolute electron density can be obtained from the HRTS without additional experiments measuring the calibration factor.

Gaussian process

1. Introduction

Consistent inference on physics parameters is arguably one of the most important and fundamental issues in any scientific field. Since large-scale magnetic confinement fusion experiments such as Joint European Torus (JET) [1] and Wendelstein 7-X (W7-X) [2] have multiple sophisticated diagnostics, it is inevitable to use a framework that is capable of handling and keeping track of parameters, data, assumptions, forward models and analysis codes in order to achieve consistent scientific inference. Resolving discrepancies among various measurements, if they exist, is challenging at least for the following reasons: i) forward models and analysis codes for a complex system, e.g., plasma diagnostics, often contain many hidden assumptions and depend on various uncertain information such as measurement positions, calibration factors and instrument effects as well as our insufficient understanding of physics, and ii) even if an individual forward model, for instance, is well managed, creating a joint model with other diagnostics or physics models can be difficult due to lack of a standardised interface.

The Minerva framework [3, 4] allows us to perform a consistent data analysis for complex experiments as it provides a standardised format (modularisation) for forward models and analysis codes, so called Minerva models, and a standardised interface to connect all of them in a systematic way. For instance, a Minerva (forward) model for a Thomson scattering system is built as an independent module depending on lasers, collecting optics, polychromators, data acquisition systems and physics parameters, i.e., electron temperature and density. Once the Minerva model is created, the input, e.g., laser energy and wavelength, polychromator details, electron temperature and density, and the output, e.g., predicted Thomson scattered signals, can be connected to/from other Minerva models and/or data sources. Minerva thus standardises scientific modelling and represents such joint models as joint probabilities and provides their graphical representation as shown in Figure 1. Some examples of implemented Minerva models are interferometer [5], magnetic sensor [6, 7, 8], Thomson scattering [9], beam emission spectroscopy [10, 11], soft X-ray [12], electron cyclotron emission [13], x-ray imaging crystal spectroscopy [14] and effective ion charge [15] diagnostics. We also note that data analysis based on Minerva models can be accelerated by field-programmable gate arrays (FPGAs) [16] and neural networks [17].

The conventional analysis for the high resolution Thomson scattering (HRTS) and the far infrared (FIR) interferometer diagnostics at JET are carried out individually. The HRTS can provide electron temperature and density profiles with the spatial resolutions of $0.8 - 1.6\,\mathrm{cm}$ [20, 21]. It is well known that electron density measurements with Thomson scattering systems require information on the calibration factors of the system [22]. The FIR interferometer diagnostic system provides line integrated electron densities, i.e., no spatial resolution, without such calibration factors [23, 24, 25, 26]. Bayesian analysis for Thomson scattering [9, 27, 28], interferometer [5, 29] systems and joint
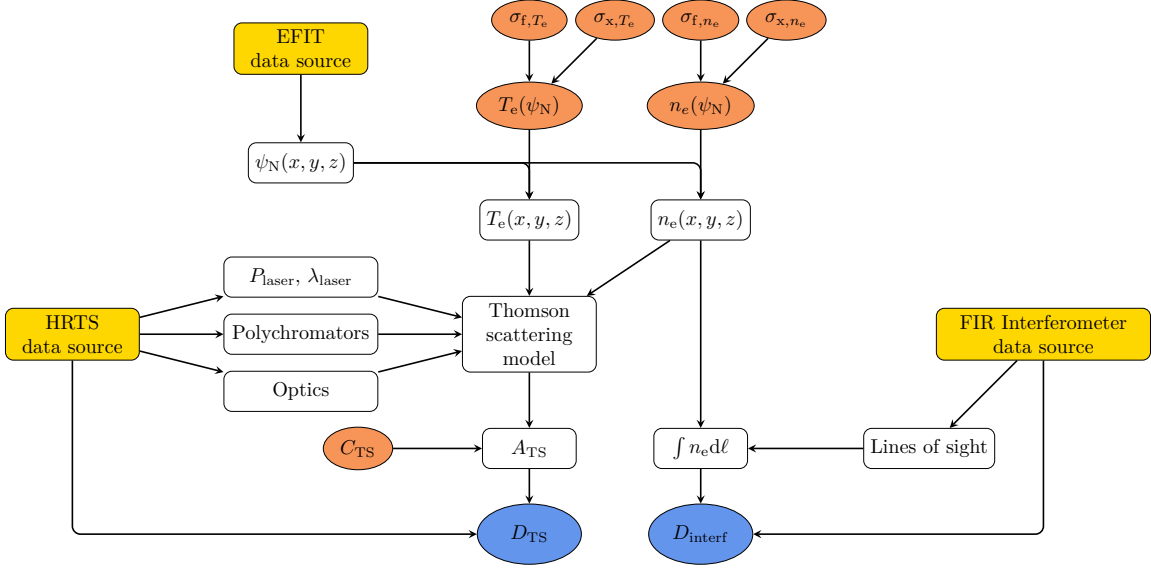
Figure 1. A simplified version of the Minerva graph representing the joint model of the high resolution Thomson scattering (HRTS) and the far infrared (FIR) interferometer systems at JET. The unknown parameters, i.e., the parameters we wish to determine, and observations are shown as red and blue circles, respectively. Note that the observations can be taken from previous inference, e.g., an *observed* Thomson scattered amplitude D_{TS} is taken from the inference results described in Section 3.1. The electron temperature T_e and density n_e are given as a function of the normalised poloidal flux ψ_N and mapped to Cartesian coordinates x, y, z in real space by using the equilibrium fitting (EFIT) code [18, 19]. The predicted Thomson scattered amplitude A_{TS} are calculated by the Minerva model of the Thomson scattering system with the calibration factor C_{TS} treated as an unknown parameter, and they are compared with the observed data. Line integrated electron densities are calculated by the lines of sight integration model mimicking the FIR interferometer system and compared with the observed data. Note that σ_{f,T_e}, σ_{x,T_e}, σ_{f,n_e} and σ_{x,n_e} are the hyperparameters of Gaussian processes which are explained in Section 2.3.

analysis [30] have been previously reported. They made impressive progress, especially evaluating uncertainties of electron temperature and density. This joint analysis [30] made use of spline models for the electron temperature and density profiles with additional monotonicity assumptions, which our method avoids. We use Gaussian processes, with a smoothness determined by a marginalisation over different length scales, thus providing an objective smoothness criterion.

In this work, we present a Bayesian joint model of the HRTS and the FIR interferometer systems at JET to combine the advantages of each diagnostic system while eliminating the disadvantages of them. We have developed Minerva (forward) models for both diagnostics individually and combine them as one joint model with an additional unknown parameter, the HRTS electron density calibration factor C_{TS}, as shown in Figure 1. The conventional analysis for the HRTS system uses a single interferometer chord (channel 3) to calibrate electron density for each discharge, if the

data from the chord 3 is available. It calculates an electron density correction factor (a single value per discharge) for the HRTS system as the ratio of line integrated density from the chord 3 to the predicted line integrated density from a linearly interpolated HRTS density profile with the EFIT mapping. The method in this work infers C_{TS} with uncertainties, self consistently for each time slice. This means that we can automatically and explicitly obtain the posterior distribution of C_{TS} explaining all the HRTS and the FIR interferometer data simultaneously, thereby getting rid of generally required extra calibration procedures, e.g., Raman calibration, to measure C_{TS}.

Electron temperature and density profiles are formulated by Gaussian processes [5, 31], a non-parametric model that can adopt the complexity of profiles within a Bayesian framework. Specifically, we use the hyperparameter model developed by [32] (denoted as $\sigma_{\mathrm{f},T_{\mathrm{e}}}$, $\sigma_{\mathrm{x},T_{\mathrm{e}}}$, $\sigma_{\mathrm{f},n_{\mathrm{e}}}$ and $\sigma_{\mathrm{x},n_{\mathrm{e}}}$ in Figure 1) whose details are explained in Section 2.3. Since we use a non-parametric model, our method does not depend on any predefined parametric model [20, 21, 33, 34, 35, 36] such as a modified hyperbolic tangent function [33] which often limit our *knowledge* on the profiles from measurements. Moreover, Gaussian processes avoid imposing assumptions such as monotonicity [30, 37].

We explore the joint posterior distribution of the electron temperature and density profiles, their hyperparameters and the HRTS electron density calibration factor by using Markov chain Monte Carlo (MCMC) sampling, specifically, we use an adaptive Metropolis-Hastings algorithm [38, 39, 40] implemented in Minerva. Finally, we obtain electron temperature(density) profiles with the associated uncertainties by marginalising out electron density(temperature) profiles, their hyperparameters and the calibration factor. In other words, our final result on the temperature and density profile does not depend on a specific set of values of all the other unknown parameters.

Section 2 describes the forward models of the HRTS and the FIR interferometer systems, and explains Gaussian processes for the electron temperature and density profile modelling. The details on Bayesian inference for electron temperature and density profiles and their results are discussed in Section 3. Our conclusions are provided in Section 4.

2. The forward model

2.1. JET high resolution Thomson scattering system

The HRTS diagnostic [20] measures the electron temperature and density from 63 spatial locations along the laser beam across the low-field side of JET from the major radius of $R = 3.0\,\mathrm{m}$ to $3.9\,\mathrm{m}$ near the middle plane ($Z \sim 0.1\,\mathrm{m}$) with spatial resolutions of $0.8-1.6\,\mathrm{cm}$. The laser wavelength of $1064\,\mathrm{nm}$ and typical energy level of $5.0\,\mathrm{J}$ with $20\,\mathrm{ns}$ pulse duration and $20\,\mathrm{Hz}$ repetition rate are used. The laser photons are scattered off of electrons via Thomson scattering, and the Thomson scattered spectra are broadened by the Doppler effect due to the thermal motions of the electrons. Electron temperature T_{e} and density n_{e} determine the width and the area, respectively, of a Thomson scattered
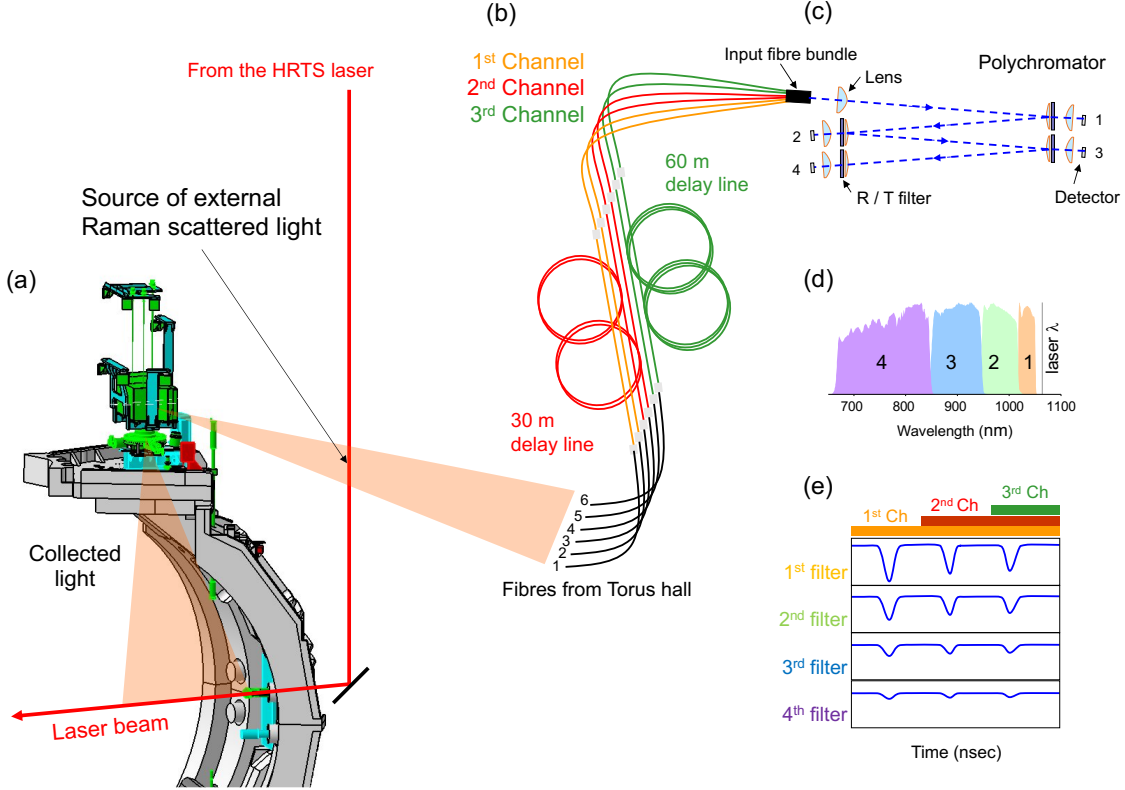
Figure 2. A schematic diagram of the HRTS system showing (a) the geometry of the laser beam and collecting optics, (b) three spatial locations covered by a polychromator with three sets of fibres (in this case, two optical fibres per location), (c) a polychromator with four reflection/transmission interference filters and detectors, and (d) an example of spectral response functions $\phi(\lambda)$. (e) Four filtered signals (cartoon drawings) detected by a polychromator from three different spatial locations are distinguished by time delays. External Raman scattered signals are originated from where the laser beam crosses the optical path between the collecting optics and the fibres.

spectrum. The predicted Thomson scattered signal amplitude A_{TS} is given as a function of T_{e}, n_{e} and the HRTS electron density calibration factor C_{TS} including the electronics, optics and geometric effects:

$$A_{\mathrm{TS}} = C_{\mathrm{TS}}\, n_{\mathrm{e}} E_{\mathrm{laser}} \int \frac{\phi(\lambda)}{\phi^{1,1}(\lambda_{\mathrm{N}})} \frac{\lambda}{hc} r_{\mathrm{e}}^2 \frac{S(\lambda, \theta, T_{\mathrm{e}})}{\lambda_{\mathrm{laser}}}\, \mathrm{d}\lambda, \tag{1}$$

where E_{laser} is the laser energy, $\phi(\lambda)$ are spectral response functions of the HRTS system, $\phi^{1,1}(\lambda_{\mathrm{N}})$ is a normalisation factor for the spectral response functions (the value of the spectral response function of the first spectral channel of the first spatial position at the wavelength $\lambda_{\mathrm{N}} = 1020\,\mathrm{nm}$), λ a scattered wavelength, h the Planck constant, c the speed of light, r_{e} the classical electron radius, $S(\lambda, \theta, T_{\mathrm{e}})$ the spectral density function [41], θ a scattering angle and λ_{laser} the laser wavelength. Further details of the Thomson scattering model are described in Appendix A.

Each polychromator detects Thomson scattered signals from three spatial locations

collected by three or six fibres, i.e., one or two fibres per location. The fibres for the second and third locations are set to be 30 m and 60 m longer than the first location, respectively, to separate these signals by time delays (Figure 2). This allows 63 spatial locations to be addressed by 21 polychromators.

Thomson scattered signals from each spatial location are resolved over time by the detectors as shown in Figure 2(e). The shape of the Thomson scattered signals depends on the shape of the laser pulses and detector electronics. By assuming that the HRTS system has Gaussian-shaped laser pulses and low-pass filter electronics [42], the Thomson scattered signals as a function of time $V_{\text{TS}}(t)$ are formulated as

$$V_{\text{TS}}(t) = A_{\text{TS}} \int_0^t \frac{1}{\sqrt{2\pi}\tau_{\text{laser}}} \frac{1}{\tau_{\text{electronics}}} \exp\left[-\frac{(t'-t_{\text{TS}})^2}{\tau_{\text{laser}}^2}\right] \exp\left(-\frac{t-t'}{\tau_{\text{electronics}}}\right) dt', \quad (2)$$

where A_{TS} is the amplitude (see Equation (1)), τ_{laser} is the width of Gaussian-shaped laser pulses, $\tau_{\text{electronics}}$ is the characteristic time of the electronics, and t_{TS} is the time when the laser energy is the maximum for Thomson scattered signals.

Raman scattered photons also get into the polychromators as the laser beam crosses the optical path between the collecting optics and fibres outside of the JET vacuum vessel (Figure 2(a)) in addition to stray light signals. These unintended *external* Raman scattered and stray light signals, which we call *parasitic* signals, are occasionally overlapping the Thomson scattered signals, therefore they must be separated out. The parasitic signals can be measured separately by firing laser pulses into the vessel before every JET discharge. While the temporal evolution of the parasitic signals is not changing, the amplitudes of the signals may slightly change over plasma discharges due to the changes of gas pressure and the laser energy. As the parasitic signals are collected by the same polychromators and fibres we can use Equation (2) to express these signals as

$$V_{\text{PS}}(t) = A_{\text{PS,R}} \int_0^t \frac{1}{\sqrt{2\pi}\tau_{\text{laser}}} \frac{1}{\tau_{\text{electronics}}} \exp\left[-\frac{(t'-t_{\text{PS,R}})^2}{\tau_{\text{laser}}^2}\right] \exp\left(-\frac{t-t'}{\tau_{\text{electronics}}}\right) dt'$$

$$+ A_{\text{PS,S}} \int_0^t \frac{1}{\sqrt{2\pi}\tau_{\text{laser}}} \frac{1}{\tau_{\text{electronics}}} \exp\left[-\frac{(t'-t_{\text{PS,S}})^2}{\tau_{\text{laser}}^2}\right] \exp\left(-\frac{t-t'}{\tau_{\text{electronics}}}\right) dt', \quad (3)$$

where $A_{\text{PS}} = [A_{\text{PS,R}}, A_{\text{PS,S}}]$ is the amplitude, and $t_{\text{PS}} = [t_{\text{PS,R}}, t_{\text{PS,S}}]$ is the time when the laser energy is the maximum for these parasitic signals (external Raman scattered and stray light signals with the subscript R or S, respectively).

By taking account of the Thomson scattered signals $V_{\text{TS}}(t)$, the parasitic signals $V_{\text{PS}}(t)$ and the constant background offset V_{B}, we can express the predicted HRTS signals $V_{\text{HRTS}}(t)$ as

$$V_{\text{HRTS}}(t) = V_{\text{TS}}(t) + V_{\text{PS}}(t) + V_{\text{B}}. \quad (4)$$

Our method infers the Thomson scattered and parasitic signals simultaneously, and results of the method are discussed in Section 3.1.
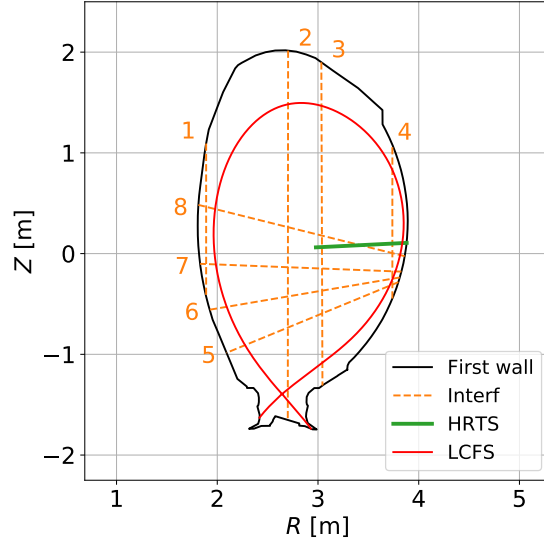
Figure 3. The lines (dashed orange) of sight of the FIR interferometer (Interf) system and the spatial positions of the HRTS system (green). The HRTS system measures the electron temperature and density along the green line with spatial resolutions of $0.8-1.6$ cm, and the FIR interferometer system measures line integrated densities along the dashed orange lines. Red line indicates the plasma boundary of a typical JET plasma discharge.

2.2. JET far infrared (FIR) interferometer system

The JET far infrared (FIR) interferometer [23, 24, 25, 26] diagnostic measures line integrated electron densities along the eight lines of sight (four vertical and four lateral directions as shown in Figure 3) with an approximately $10-200\,\mu$s temporal resolution (depending on the channels). The 1st and 4th channels measure the line integrated densities near the first wall, constraining the electron density at the edge, and the other lines of sight cover the plasma core. Our forward model calculates the line integrated electron density of the ith channel V_{interf}^i as:

$$V_{\text{interf}}^i = \int n_{\text{e}}\left(x, y, z\right) \mathrm{d}\ell, \tag{5}$$

where the integration path $\int \mathrm{d}\ell$ is all the way along the lines of sight inside the vacuum vessel.

2.3. Gaussian process prior

A Gaussian process [31] is a probabilistic function that associates each element of a domain to a single element of a random vector following a multivariate Gaussian distribution whose mean at every point and covariance between any two points within the domain are defined. Thus, the function that is modelled by a Gaussian process is

not constrained by any specific parametric representation, but properties given by their mean and covariance functions.

Zero mean, which does not usually limit the inference solutions [31], and squared exponential covariance functions are one of the most common specifications of a Gaussian process, and they are given as

$$\mu\left(x\right) = 0 \tag{6}$$

$$\Sigma\left(x^i, x^j\right) = \sigma_{\mathrm{f}}^2 \exp\left(-\frac{\left(x^i - x^j\right)^2}{2\sigma_{\mathrm{x}}^2}\right) + \sigma_{\mathrm{y}}^2 \delta^{ij}, \tag{7}$$

where μ is the mean function, and x is a scalar or vector input, e.g., time or space. x^i and x^j denote i^{th} and j^{th} elements of the input domain, respectively. Σ is the covariance function with an overall scale σ_{f}, a length scale σ_{x} and a noise scale σ_{y}. These scales are hyperparameters, and σ_{f} and σ_{x} determine the smoothness of the function. σ_{y} is chosen to be a relatively small number with respect to the overall scale, i.e., $\sigma_{\mathrm{y}}/\sigma_{\mathrm{f}} = 10^{-3}$ in this work, to avoid any numerical instability [10]. δ^{ij} is the Kronecker delta function.

We note that Gaussian processes have been introduced to the fusion community with a non-parametric tomography for the JET FIR interferometer data [5] and used for current tomography [6, 7, 43, 44], beam emission spectroscopy [10, 11], soft X-ray spectroscopy [12], X-ray imaging crystal spectroscopy [14] and profile regressions [32]. In this work, Gaussian processes are used to model electron temperature and density profiles to constrain their smoothness (gradient scale length) without imposing any specific profile shapes or assumptions such as monotonicity.

Electron temperature and density profiles can have substantially different gradients in the core and edge regions [45]. In order to represent such spatially varying gradients, we choose non-stationary covariance functions [46], given as

$$\Sigma\left(x^i, x^j\right) = \sigma_{\mathrm{f}}^2 \left(\frac{2\sigma_{\mathrm{x}}\left(x^i\right)\sigma_{\mathrm{x}}\left(x^j\right)}{\sigma_{\mathrm{x}}\left(x^i\right)^2 + \sigma_{\mathrm{x}}\left(x^j\right)^2}\right)^{\frac{1}{2}} \exp\left(-\frac{\left(x^i - x^j\right)^2}{\sigma_{\mathrm{x}}\left(x^i\right)^2 + \sigma_{\mathrm{x}}\left(x^j\right)^2}\right) + \sigma_{\mathrm{y}}^2 \delta^{ij}, \tag{8}$$

where the length scale σ_{x} is an arbitrary function of the input x. The length scale function σ_{x} specifies the smoothness at any point on the domain. We need a function with different length scales in the core and edge regions and a smooth transition between the two. For this reason, we choose a hyperbolic tangent function [32],

$$\sigma_{\mathrm{x}}\left(x^i\right) = \frac{\sigma_{\mathrm{x}}^{\mathrm{core}} + \sigma_{\mathrm{x}}^{\mathrm{edge}}}{2} - \frac{\sigma_{\mathrm{x}}^{\mathrm{core}} - \sigma_{\mathrm{x}}^{\mathrm{edge}}}{2} \tanh\left(\frac{x^i - x_0}{\sigma_{\mathrm{x}}^{\mathrm{width}}}\right), \tag{9}$$

where $\sigma_{\mathrm{x}}^{\mathrm{core}}$ and $\sigma_{\mathrm{x}}^{\mathrm{edge}}$ are the core and edge length scales, respectively. x_0 is the transition position of the length scale, and $\sigma_{\mathrm{x}}^{\mathrm{width}}$ determines how fast the transition occurs. With this length scale function, the smoothness (gradient scale length) of the profiles is changing gradually from the core to the edge.

The Gaussian processes of electron temperature and density profiles are set as a function of the normalised poloidal flux, i.e., $x = \psi_{\mathrm{N}}$, as shown in the Figure 1. By using

the equilibrium fitting (EFIT) code [18, 19] with only the magnetic diagnostics, electron temperature and density profiles are mapped onto Cartesian coordinates x, y, z in real space. The forward models of the HRTS and the FIR interferometer systems access values of T_e and n_e at their corresponding spatial positions and calculate the predictions, A_{TS} and V_{interf}^i, given by Equation (1) and (5), respectively. These predictions are directly compared to the observations for inference.

3. Bayesian inference

In Bayesian probability theory, we express a probability of generating (observing) experimental data D given parameter values H as a conditional probability $P(D|H)$, a likelihood. Our state of knowledge on the parameters H before any observations are taken into account is formulated as a prior probability $P(H)$. For instance, $P(T_e)$ may be a uniform distribution from zero to some maximum, e.g., $100\,\mathrm{keV}$. Note that the parameters H are what we wish to infer by conducting experiments. Bayes' formula states that our state of knowledge on the parameters H given the observed data D is the posterior probability $P(H|D)$, expressed as

$$P(H|D) = \frac{P(D|H)\,P(H)}{P(D)},\tag{10}$$

where the denominator is a normalisation factor $P(D)$ (also called the model evidence). A more detailed description of Bayesian inference can be found in [47, 48].

The posterior distribution is high dimensional if there exist many parameters of interest. Thus, to calculate lower dimensional distributions of parameters of interest, a marginalisation has to be carried out:

$$P(H_1) = \int P(H_1, H_2)\,\mathrm{d}H_2.\tag{11}$$

The marginal distributions provide full information of the parameters of interest by taking into account all possible values of the other parameters. We can perform this integration by collecting values of the parameters of interest H_1 from joint samples $[H_1, H_2]$ of the joint distribution $P(H_1, H_2)$. These collected values of H_1 are equivalent to samples drawn from the marginal distribution $P(H_1)$.

In this work, the parameters are electron temperature and density profiles at 50 normalised poloidal flux surfaces, $\mathbf{T}_e = [T_e(\psi_N^{i=1}), T_e(\psi_N^{i=2}), \cdots, T_e(\psi_N^{i=s})]$ and $\mathbf{n}_e = [n_e(\psi_N^{i=1}), n_e(\psi_N^{i=2}), \cdots, n_e(\psi_N^{i=s})]$ with $s = 50$, the calibration factor C_{TS} and the hyperparameters $\mathbf{M}_{T_e} = [\sigma_{f,T_e}, \sigma_{x,T_e}]$ and $\mathbf{M}_{n_e} = [\sigma_{f,n_e}, \sigma_{x,n_e}]$ (the red circles in Figure 1). σ_{x,T_e} and σ_{x,n_e} are given by Equation (9) with two sets of four parameters σ_x^{core}, σ_x^{edge}, σ_x^{width}, and x_0. The data are obtained by the HRTS and the FIR interferometer diagnostics. With $\mathbf{D}_{TS}$ denoting the amplitudes of Thomson scattered signals from all the spatial locations and $\mathbf{D}_{interf}$ standing for the line integrated electron densities from all the lines

of sight, the posterior is expressed as

$$
\begin{aligned}
&P\left(\mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}, C_{\mathrm{TS}}, \mathbf{M}_{T_{\mathrm{e}}}, \mathbf{M}_{n_{\mathrm{e}}} \mid \mathbf{D}_{\mathrm{TS}}, \mathbf{D}_{\mathrm{interf}}\right) \\
&= \frac{P\left(\mathbf{D}_{\mathrm{TS}}, \mathbf{D}_{\mathrm{interf}} \mid \mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}, C_{\mathrm{TS}}, \mathbf{M}_{T_{\mathrm{e}}}, \mathbf{M}_{n_{\mathrm{e}}}\right) P\left(\mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}, C_{\mathrm{TS}}, \mathbf{M}_{T_{\mathrm{e}}}, \mathbf{M}_{n_{\mathrm{e}}}\right)}{P\left(\mathbf{D}_{\mathrm{TS}}, \mathbf{D}_{\mathrm{interf}}\right)} \\
&= \frac{P\left(\mathbf{D}_{\mathrm{TS}} \mid \mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}, C_{\mathrm{TS}}\right) P\left(\mathbf{D}_{\mathrm{interf}} \mid \mathbf{n}_{\mathrm{e}}\right) P\left(\mathbf{T}_{\mathrm{e}} \mid \mathbf{M}_{T_{\mathrm{e}}}\right) P\left(\mathbf{n}_{\mathrm{e}} \mid \mathbf{M}_{n_{\mathrm{e}}}\right) P\left(\mathbf{M}_{T_{\mathrm{e}}}\right) P\left(\mathbf{M}_{n_{\mathrm{e}}}\right) P\left(C_{\mathrm{TS}}\right)}{P\left(\mathbf{D}_{\mathrm{TS}}\right) P\left(\mathbf{D}_{\mathrm{interf}}\right)}.
\end{aligned}
\tag{12}
$$

Notice that Figure 1 exactly expresses Equation (12) which is automatically generated by the Minerva framework such that conditional dependences among the parameters and observations can be easily verified.

Raw data from the HRTS system contain the parasitic signals, i.e., the external Raman scattered and stray light signals as explained in Section 2.1. In order to proceed the profile inference, we have to first extract the Thomson scattered signals from the raw HRTS data which is discussed in Section 3.1. Then, we present how electron temperature and density profiles are inferred in Section 3.2.

3.1. Inference on the amplitudes of the Thomson scattered signals

We assume that the time series of raw HRTS data from the ith spatial position (total of 63 spatial positions with 21 polychromators, $i = 1, 2, \ldots, 63$) and the jth spectral channel (four spectral channels for each polychromator $j = 1, \ldots, 4$) denoted as $\mathbf{D}_{\mathrm{HRTS}}^{i,j}$§ is following a multivariate Gaussian distribution whose mean is the HRTS predicted signals $\mathbf{V}_{\mathrm{HRTS}}^{i,j}$ given by Equation (4) with the covariance $\boldsymbol{\Sigma}_{\mathrm{HRTS}}^{i,j}$ as a combination of the electronics and photon noises. Thus, the probability $P\left(\mathbf{D}_{\mathrm{HRTS}}^{i,j} \mid A_{\mathrm{TS}}^{i,j}, A_{\mathrm{PS}}^{i,j}, V_{\mathrm{B}}^{i,j}, t_{\mathrm{TS}}^{i,j}, t_{\mathrm{PS}}^{i,j}, \tau_{\mathrm{laser}}^{i,j}, \tau_{\mathrm{electronics}}^{i,j}\right)$ is written as

$$
\begin{aligned}
&P\left(\mathbf{D}_{\mathrm{HRTS}}^{i,j} \mid A_{\mathrm{TS}}^{i,j}, A_{\mathrm{PS}}^{i,j}, V_{\mathrm{B}}^{i,j}, t_{\mathrm{TS}}^{i,j}, t_{\mathrm{PS}}^{i,j}, \tau_{\mathrm{laser}}^{i,j}, \tau_{\mathrm{electronics}}^{i,j}\right) = \\
&\frac{1}{\sqrt{(2\pi)^m \left|\boldsymbol{\Sigma}_{\mathrm{HRTS}}^{i,j}\right|}} \exp\left(-\frac{1}{2}\left(\mathbf{D}_{\mathrm{HRTS}}^{i,j} - \mathbf{V}_{\mathrm{HRTS}}^{i,j}\right)^{\mathrm{T}} {\boldsymbol{\Sigma}_{\mathrm{HRTS}}^{i,j}}^{-1} \left(\mathbf{D}_{\mathrm{HRTS}}^{i,j} - \mathbf{V}_{\mathrm{HRTS}}^{i,j}\right)\right).
\end{aligned}
\tag{13}
$$

$\mathbf{D}_{\mathrm{HRTS}}^{i,j}$ and $\mathbf{V}_{\mathrm{HRTS}}^{i,j}$ are time series vectors, e.g., $\mathbf{D}_{\mathrm{HRTS}}^{i,j} = \left[D_{\mathrm{HRTS}}^{i,j}\left(t = t_1\right), \cdots, D_{\mathrm{HRTS}}^{i,j}\left(t = t_{\mathrm{m}}\right)\right]$ with $m = 500$ covering $500\,\mathrm{ns}$. The covariance $\boldsymbol{\Sigma}_{\mathrm{HRTS}}^{i,j}$ is given as a diagonal matrix assuming no correlation with other time points, that is

$$
\boldsymbol{\Sigma}_{\mathrm{HRTS}}^{i,j} = \begin{pmatrix}
\left(\sigma_{\mathrm{HRTS}}^{i,j}\left(t = t_1\right)\right)^2 & 0 & \cdots & 0 \\
0 & \left(\sigma_{\mathrm{HRTS}}^{i,j}\left(t = t_2\right)\right)^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \left(\sigma_{\mathrm{HRTS}}^{i,j}\left(t = t_m\right)\right)^2
\end{pmatrix},
\tag{14}
$$

§ Note that $\mathbf{D}_{\mathrm{TS}}$ in Equation (12) are the amplitudes of the Thomson scattered signals *extracted* from the raw HRTS data $\mathbf{D}_{\mathrm{HRTS}}$.
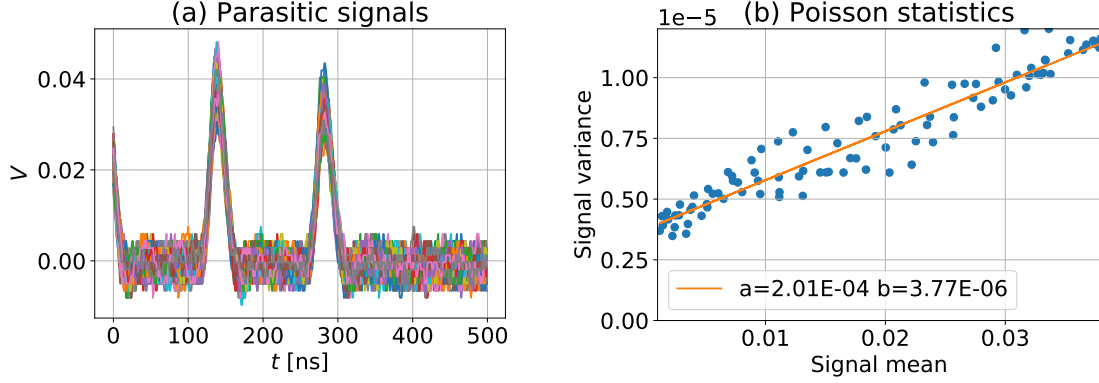
Figure 4. (a) An example of the parasitic signals obtained by firing the laser pulses before a plasma discharge. Different signal levels from different laser pulses (different colours) at a fixed time are dominated by the photon noise. (b) Poisson statistics are calculated from the linear relation between the signal mean and variance from the different parasitic signals, and we find that $\sigma_{\text{photon}}^2 = aV_{\text{HRTS}}$ with $a = 2.01 \times 10^{-4}$ in this example.

where $\left(\sigma_{\text{HRTS}}^{i,j}(t)\right)^2 = \left(\sigma_{\text{electronics}}^{i,j}\right)^2 + \left(\sigma_{\text{photon}}^{i,j}(t)\right)^2$ is the uncertainty of the raw HRTS data.

The electronics noise $\sigma_{\text{electronics}}^{i,j}$ is set as a zero mean Gaussian noise whose standard deviation is estimated from the electronics noise signals measured before the laser pulses and plasma discharges, i.e., without the laser pulses and plasmas, and assumed not to vary over time.

The photon noise is estimated from the mean and variance of the parasitic signals. Figure 4(a) shows an example of the parasitic signals obtained by firing many laser pulses (different colours) before a single plasma discharge, and (b) shows the linear relation between the signal mean and the variance over many laser pulses with the estimated slope of $a = 2.01 \times 10^{-4}$. Here, the photon noise is assumed to be the dominant contribution to the signal fluctuation [36]. Once we have the value of a, then the photon noise can be estimated as

$$\left(\sigma_{\text{photon}}^{i,j}(t)\right)^2 = aV_{\text{HRTS}}^{i,j}(t). \tag{15}$$

More details on the method of the photon noise estimation can be found in [10].

Assuming that a prior probability of each parameter ($A_{\text{TS}}^{i,j}$, $A_{\text{PS}}^{i,j}$, $V_{\text{B}}^{i,j}$, $t_{\text{TS}}^{i,j}$, $t_{\text{PS}}^{i,j}$, $\tau_{\text{laser}}^{i,j}$ and $\tau_{\text{electronics}}^{i,j}$) is given as a uniform distribution, the posterior probability is written as

$$P\left(A_{\text{TS}}^{i,j}, A_{\text{PS}}^{i,j}, V_{\text{B}}^{i,j}, t_{\text{TS}}^{i,j}, t_{\text{PS}}^{i,j}, \tau_{\text{laser}}^{i,j}, \tau_{\text{electronics}}^{i,j}|\mathbf{D}_{\text{HRTS}}^{i,j}\right) =$$
$$P\left(\mathbf{D}_{\text{HRTS}}^{i,j}|A_{\text{TS}}^{i,j}, A_{\text{PS}}^{i,j}, V_{\text{B}}^{i,j}, t_{\text{TS}}^{i,j}, t_{\text{PS}}^{i,j}, \tau_{\text{laser}}^{i,j}, \tau_{\text{electronics}}^{i,j}\right) \times$$
$$\frac{P\left(A_{\text{TS}}^{i,j}\right) P\left(A_{\text{PS}}^{i,j}\right) P\left(V_{\text{B}}^{i,j}\right) P\left(t_{\text{TS}}^{i,j}\right) P\left(t_{\text{PS}}^{i,j}\right) P\left(\tau_{\text{laser}}^{i,j}\right) P\left(\tau_{\text{electronics}}^{i,j}\right)}{P\left(\mathbf{D}_{\text{HRTS}}^{i,j}\right)}. \tag{16}$$

The posterior distribution of $A_{\text{TS}}^{i,j}$, $A_{\text{PS}}^{i,j}$, $V_{\text{B}}^{i,j}$, $t_{\text{TS}}^{i,j}$, $t_{\text{PS}}^{i,j}$, $\tau_{\text{laser}}^{i,j}$ and $\tau_{\text{electronics}}^{i,j}$ is explored by Markov chain Monte Carlo (MCMC) sampling, specifically, we use the adaptive
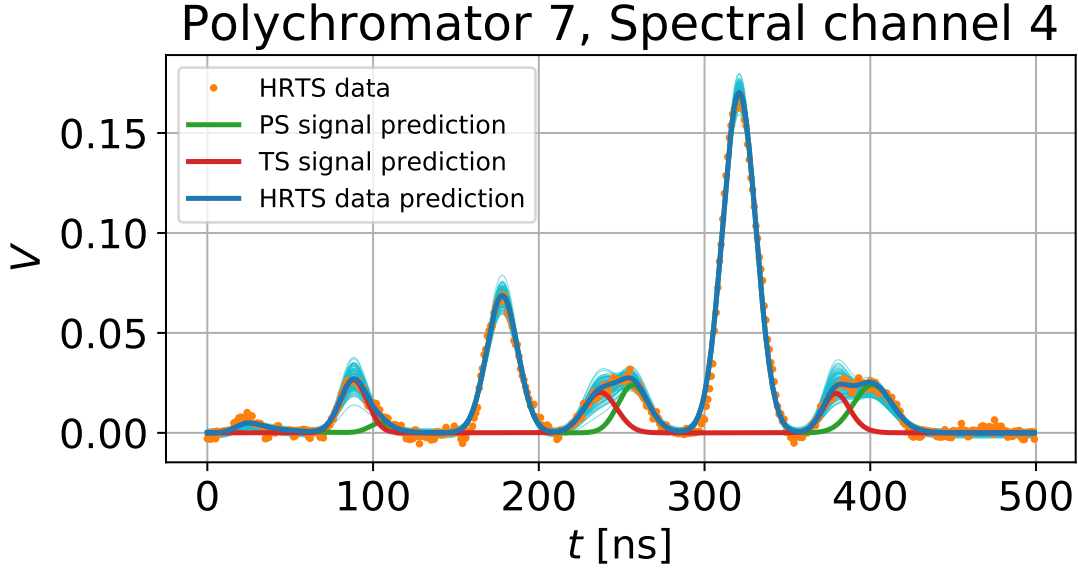
Figure 5. An example of the measured raw HRTS data $\mathbf{D}_{\mathrm{HRTS}}^{i,j}$ (orange dots) and the mean of predicted HRTS signals $\mathbf{V}_{\mathrm{HRTS}}^{i,j}$ (dark blue line) from the fourth spectral channel ($j = 4$) of the seventh polychromator (corresponding to $i = 19, 20$ and 21) during a plasma discharge over one laser pulse. The light blue lines show samples of the posterior (Equation (16)) explored by the MCMC method. The red and green lines are the predicted Thomson scattered signals $\mathbf{V}_{\mathrm{TS}}^{i,j}$ and the predicted parasitic signals $\mathbf{V}_{\mathrm{PS}}^{i,j}$ calculated through Equations (2) and (3), respectively, given the posterior mean. There are three peaks of Thomson scattered signals as one polychromator receives the signals from three different spatial positions as described in Section 2.1

Metropolis-Hastings algorithm [38, 39, 40]. The mechanism of this algorithm is: i) to propose a random sample x' in parameter space from a proposal distribution $Q\left(x'|x_t\right)$ given a previous sample x_t and ii) to accept the candidate if $P\left(x'\right)/P\left(x_t\right) \geq u$, otherwise reject it ($P\left(x\right)$ is the posterior probability of x, and u is a random number from a uniform distribution on $[0, 1]$). After some iterations, the algorithms will collect a set of samples, which are drawn from the posterior distribution [38, 39]. We use an adaptive rule [40] to modify the proposal distribution Q in every iteration to sample the posterior distribution effectively.

The samples drawn from the posterior distribution given by Equation (16) are shown in Figure 5. The orange dots are the measured raw HRTS data $\mathbf{D}_{\mathrm{HRTS}}^{i,j}$ (the fourth spectral channel of the seventh polychromator in this example) during a plasma discharge over one laser pulse. The dark blue line is the mean of predicted HRTS signals $\mathbf{V}_{\mathrm{HRTS}}^{i,j}$, whereas the light blue lines are samples of the posterior. The red and green lines are the predicted Thomson scattered signals $\mathbf{V}_{\mathrm{TS}}^{i,j}$ and the predicted parasitic signals $\mathbf{V}_{\mathrm{PS}}^{i,j}$ calculated through Equations (2) and (3), respectively, given the posterior mean.

As electron temperature and density profiles are inferred based on the *observed*‖

‖ As a matter of fact, $\mathbf{D}_{\mathrm{TS}}$ is an inferred quantity (rather than being an observed quantity) from the
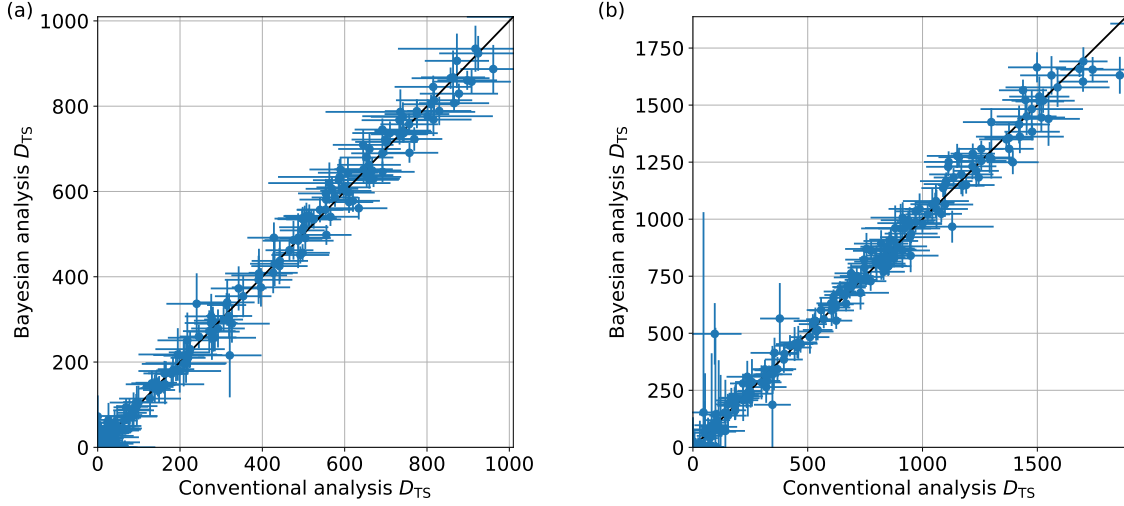
Figure 6. Comparisons between the mean values of the Thomson scattered amplitudes with one standard deviation error bars from the Bayesian approach and the conventional method for (a) JET discharge #88630 at 8.024 s and (b) JET discharge #89380 at 11.776 s. The black line is: $y = x$.

amplitudes of the Thomson scattered signals $\mathbf{D}_{\mathrm{TS}}$ from all the spatial locations, we need to obtain $\mathbf{A}_{\mathrm{TS}}$ given the measured raw HRTS data (see Equation (2)). Thus, we marginalise out all the other parameters except $A_{\mathrm{TS}}^{i,j}$ from the posterior as

$$
P\left(A_{\mathrm{TS}}^{i,j}|\mathbf{D}_{\mathrm{HRTS}}^{i,j}\right) =
$$
$$
\int P\left(A_{\mathrm{TS}}^{i,j}, A_{\mathrm{PS}}^{i,j}, V_{\mathrm{B}}^{i,j}, t_{\mathrm{TS}}^{i,j}, t_{\mathrm{PS}}^{i,j}, \tau_{\mathrm{laser}}^{i,j}, \tau_{\mathrm{electronics}}^{i,j}|\mathbf{D}_{\mathrm{HRTS}}^{i,j}\right) \mathrm{d}A_{\mathrm{PS}}^{i,j}\mathrm{d}V_{\mathrm{B}}^{i,j}\mathrm{d}t_{\mathrm{TS}}^{i,j}\mathrm{d}t_{\mathrm{PS}}^{i,j}\mathrm{d}\tau_{\mathrm{laser}}^{i,j}\mathrm{d}\tau_{\mathrm{electronics}}^{i,j}.
$$
$$
(17)
$$

We perform the marginalisation by collecting the values of $A_{\mathrm{TS}}^{i,j}$ from sample vectors $\left[A_{\mathrm{TS}}^{i,j}, A_{\mathrm{PS}}^{i,j}, V_{\mathrm{B}}^{i,j}, t_{\mathrm{TS}}^{i,j}, t_{\mathrm{PS}}^{i,j}, \tau_{\mathrm{laser}}^{i,j}, \tau_{\mathrm{electronics}}^{i,j}\right]$ of the posterior given by Equation (16). The collected values of $A_{\mathrm{TS}}^{i,j}$ are equivalent to samples drawn from the marginal posterior distribution $P\left(A_{\mathrm{TS}}^{i,j}|\mathbf{D}_{\mathrm{HRTS}}^{i,j}\right)$ given by Equation (17). Finally, we define the mean and variance of the amplitudes of the Thomson scattered signal from the jth spectral channel of the ith spatial position as $D_{\mathrm{TS}}^{i,j}$ and $\left(\sigma_{\mathrm{TS}}^{i,j}\right)^2$ for profile inference in the following section, respectively, which are calculated as

$$
D_{\mathrm{TS}}^{i,j} = \int A_{\mathrm{TS}}^{i,j} \, P\left(A_{\mathrm{TS}}^{i,j}|\mathbf{D}_{\mathrm{HRTS}}^{i,j}\right) \mathrm{d}A_{\mathrm{TS}}^{i,j}, \tag{18}
$$

$$
\left(\sigma_{\mathrm{TS}}^{i,j}\right)^2 = \int \left(A_{\mathrm{TS}}^{i,j} - D_{\mathrm{TS}}^{i,j}\right)^2 P\left(A_{\mathrm{TS}}^{i,j}|\mathbf{D}_{\mathrm{HRTS}}^{i,j}\right) \mathrm{d}A_{\mathrm{TS}}^{i,j}. \tag{19}
$$

We emphasise that our method infers the amplitudes of the Thomson scattered signals $A_{\mathrm{TS}}^{i,j}$ by taking into account all possible values of all the other parameters $A_{\mathrm{PS}}^{i,j}$, $V_{\mathrm{B}}^{i,j}$, $t_{\mathrm{TS}}^{i,j}$, $t_{\mathrm{PS}}^{i,j}$, $\tau_{\mathrm{laser}}^{i,j}$ and $\tau_{\mathrm{electronics}}^{i,j}$. As discussed in Section 2.1, the parasitic signals measured raw data $\mathbf{D}_{\mathrm{HRTS}}$. We treat the inferred $\mathbf{D}_{\mathrm{TS}}$ as an observed one for the profile inference.

might change over laser pulses during plasma discharges, and our method takes that into account by marginalisation, to arrive at a proper distribution of $A_{\mathrm{TS}}^{i,j}$ from the raw HRTS measurements. In addition, our method uses a more realistic signal model for the Thomson scattered and parasitic signals (Equations (2) and (3)) and uncertainty model which takes into account electronics and Photon noises. In contrast, the conventional method uses a Gaussian signal model and does not perform the marginalisation to calculate $A_{\mathrm{TS}}^{i,j}$. The mean values of the Thomson scattered amplitudes with one standard deviation error bars from the Bayesian approach and the conventional method are compared as shown in Figure 6. In general, the two results are comparable to each other, but there are notable differences on the size of error bars, especially when the signal level is low.

3.2. Inference on the electron temperature and density profiles

3.2.1. Likelihood. To obtain the posterior probability of electron temperature and density profiles, $\mathbf{T}_{\mathrm{e}}$ and $\mathbf{n}_{\mathrm{e}}$, the calibration factor C_{TS}, and the hyperparameters, $\mathbf{M}_{T_{\mathrm{e}}}$ and $\mathbf{M}_{n_{\mathrm{e}}}$, given $\mathbf{D}_{\mathrm{TS}}$ and $\mathbf{D}_{\mathrm{interf}}$, we need to model two likelihoods, i.e., one for the Thomson scattered data $P\left(\mathbf{D}_{\mathrm{TS}}|\mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}, C_{\mathrm{TS}}\right)$ and the other for the FIR interferometer data $P\left(\mathbf{D}_{\mathrm{interf}}|\mathbf{n}_{\mathrm{e}}\right)$ as described in Equation (12).

We assume that the conditional probability of the Thomson scattered data $P\left(\mathbf{D}_{\mathrm{TS}}|\mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}, C_{\mathrm{TS}}\right)$ is a multivariate Gaussian distribution whose mean is $\mathbf{D}_{\mathrm{TS}} = \left\{D_{\mathrm{TS}}^{i,j}|i = 1, 2, \cdots, 63, j = 1, \cdots, 4\right\} = \left[D_{\mathrm{TS}}^{1,1}, D_{\mathrm{TS}}^{1,2}, \cdots, D_{\mathrm{TS}}^{63,3}, D_{\mathrm{TS}}^{63,4}\right]$ (63 spatial positions and four spectral channels for each polychromator resulting in total of $63 \times 4 = 252$ amplitudes) with the covariance $\boldsymbol{\Sigma}_{\mathrm{TS}}$ as

$$\boldsymbol{\Sigma}_{\mathrm{TS}} = \begin{pmatrix} \left(\sigma_{\mathrm{TS}}^{1,1}\right)^2 & 0 & \cdots & 0 & 0 \\ 0 & \left(\sigma_{\mathrm{TS}}^{1,2}\right)^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \left(\sigma_{\mathrm{TS}}^{63,3}\right)^2 & 0 \\ 0 & 0 & \cdots & 0 & \left(\sigma_{\mathrm{TS}}^{63,4}\right)^2 \end{pmatrix}. \tag{20}$$

We have calculated $D_{\mathrm{TS}}^{i,j}$ and $\left(\sigma_{\mathrm{TS}}^{i,j}\right)^2$ in Section 3.1. Thus, we have

$$P\left(\mathbf{D}_{\mathrm{TS}}|\mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}, C_{\mathrm{TS}}\right) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}_{\mathrm{TS}}|}} \exp\left(-\frac{1}{2}\left(\mathbf{D}_{\mathrm{TS}} - \mathbf{A}_{\mathrm{TS}\star}\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\mathrm{TS}}^{-1}\left(\mathbf{D}_{\mathrm{TS}} - \mathbf{A}_{\mathrm{TS}\star}\right)\right), \tag{21}$$

where $\mathbf{A}_{\mathrm{TS}\star} = \left\{A_{\mathrm{TS}\star}^{i,j}|i = 1, 2, \cdots, 63, j = 1, \cdots, 4\right\} = \left[A_{\mathrm{TS}\star}^{1,1}, A_{\mathrm{TS}\star}^{1,2}, \cdots, A_{\mathrm{TS}\star}^{63,3}, A_{\mathrm{TS}\star}^{63,4}\right]$ is a *predicted* quantity (denoted with an additional subscript $\star$) calculated by Equation (1) given the parameters of $\mathbf{T}_{\mathrm{e}}, \mathbf{n}_{\mathrm{e}}$ and C_{TS}. Notice the difference between $\mathbf{A}_{\mathrm{TS}\star}$ and $\mathbf{A}_{\mathrm{TS}}$, where the latter is an *observed* quantity from the raw HRTS data. Here, k is the total number of Thomson scattered amplitudes, i.e. $k = 63 \times 4 = 252$. We formulate the likelihood for the Thomson scattered data as a multivariate Gaussian distribution for computational efficiency.

With the same argument, the conditional probability of the FIR interferometer data $P\left(\mathbf{D}_{\text{interf}}|\mathbf{n}_{\text{e}}\right)$ is

$$P\left(\mathbf{D}_{\text{interf}}|\mathbf{n}_{\text{e}}\right) = \frac{1}{\sqrt{(2\pi)^l\,|\mathbf{\Sigma}_{\text{interf}}|}}\exp\left(-\frac{1}{2}\left(\mathbf{D}_{\text{interf}}-\mathbf{V}_{\text{interf}}\right)^{\text{T}}\mathbf{\Sigma}_{\text{interf}}^{-1}\left(\mathbf{D}_{\text{interf}}-\mathbf{V}_{\text{interf}}\right)\right),$$

(22)

where $\mathbf{V}_{\text{interf}} = \{V_{\text{interf}}^i|i=1,2,\cdots,8\} = [V_{\text{interf}}^{i=1},V_{\text{interf}}^{i=2},\cdots,V_{\text{interf}}^{i=8}]$ is a set of the predicted line integrated densities calculated by Equation (5) given the parameter of $\mathbf{n}_{\text{e}}$. Here, $l = 8$ representing the eight lines of sight. The covariance matrix $\mathbf{\Sigma}_{\text{interf}}$ is an $l \times l$ diagonal matrix, where the diagonal elements describe the uncertainties, i.e., variance, of the eight channels of the FIR interferometer system. Relevant measured quantities $\mathbf{D}_{\text{interf}}$ and $\mathbf{\Sigma}_{\text{interf}}$ are retrieved from the data source [5, 29], i.e., JET database.

3.2.2. Prior. We now turn our attention to model the prior probabilities, which are $P\left(\mathbf{T}_{\text{e}}|\mathbf{M}_{T_{\text{e}}}\right)$, $P\left(\mathbf{n}_{\text{e}}|\mathbf{M}_{n_{\text{e}}}\right)$, $P\left(\mathbf{M}_{T_{\text{e}}}\right)$, $P\left(\mathbf{M}_{n_{\text{e}}}\right)$ and $P\left(C_{\text{TS}}\right)$, in Equation (12).

The prior probabilities of $\mathbf{T}_{\text{e}}$ and $\mathbf{n}_{\text{e}}$ are modelled by using the Gaussian processes with the zero mean function (Equation (6)) and the covariance function $\Sigma\left(\psi_{\text{N}}^i,\psi_{\text{N}}^j\right)$ (Equation (8)). We form the Gaussian process priors as

$$P\left(\mathbf{T}_{\text{e}}|\mathbf{M}_{T_{\text{e}}}\right) = \frac{1}{\sqrt{(2\pi)^s\,|\mathbf{\Sigma}_{\mathbf{T}_{\text{e}}}|}}\exp\left(-\frac{1}{2}\left(\mathbf{T}_{\text{e}}-\mathbf{0}\right)^{\text{T}}\mathbf{\Sigma}_{\mathbf{T}_{\text{e}}}^{-1}\left(\mathbf{T}_{\text{e}}-\mathbf{0}\right)\right),\qquad(23)$$

$$P\left(\mathbf{n}_{\text{e}}|\mathbf{M}_{n_{\text{e}}}\right) = \frac{1}{\sqrt{(2\pi)^s\,|\mathbf{\Sigma}_{\mathbf{n}_{\text{e}}}|}}\exp\left(-\frac{1}{2}\left(\mathbf{n}_{\text{e}}-\mathbf{0}\right)^{\text{T}}\mathbf{\Sigma}_{\mathbf{n}_{\text{e}}}^{-1}\left(\mathbf{n}_{\text{e}}-\mathbf{0}\right)\right),\qquad(24)$$

where $s = 50$ denotes the number of flux surfaces we use to infer the profiles in this work. $\mathbf{0}$ is the zero vector, and the covariance matrix $\mathbf{\Sigma}_{\oplus}$ (where the subscript $\oplus$ representing either $\mathbf{T}_{\text{e}}$ or $\mathbf{n}_{\text{e}}$) is defined as

$$\mathbf{\Sigma}_{\oplus} = \begin{pmatrix} \Sigma_{\oplus}(\psi_{\text{N}}^{i=1},\psi_{\text{N}}^{i=1}) & \Sigma_{\oplus}(\psi_{\text{N}}^{i=1},\psi_{\text{N}}^{i=2}) & \cdots & \Sigma_{\oplus}(\psi_{\text{N}}^{i=1},\psi_{\text{N}}^{i=s}) \\ \Sigma_{\oplus}(\psi_{\text{N}}^{i=2},\psi_{\text{N}}^{i=1}) & \Sigma_{\oplus}(\psi_{\text{N}}^{i=2},\psi_{\text{N}}^{i=2}) & \cdots & \Sigma_{\oplus}(\psi_{\text{N}}^{i=2},\psi_{\text{N}}^{i=s}) \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{\oplus}(\psi_{\text{N}}^{i=s},\psi_{\text{N}}^{i=1}) & \Sigma_{\oplus}(\psi_{\text{N}}^{i=s},\psi_{\text{N}}^{i=2}) & \cdots & \Sigma_{\oplus}(\psi_{\text{N}}^{i=s},\psi_{\text{N}}^{i=s}) \end{pmatrix}.\qquad(25)$$

We set the prior probabilities of the hyperparameters $P\left(\mathbf{M}_{T_{\text{e}}}\right)$ and $P\left(\mathbf{M}_{n_{\text{e}}}\right)$ to be uniform distributions. Likewise, the prior of the calibration factor $P\left(C_{\text{TS}}\right)$ is set to be a uniform distribution.

3.2.3. Posterior. Equation (12) provides us the joint posterior probability of $\mathbf{T}_{\text{e}}$, $\mathbf{n}_{\text{e}}$, C_{TS}, $\mathbf{M}_{T_{\text{e}}}$ and $\mathbf{M}_{n_{\text{e}}}$ with the likelihoods, i.e., Equations (21) and (22), and the prior probabilities, i.e., Equations (23) and (24) together with the uniform distributions for C_{TS}, $\mathbf{M}_{T_{\text{e}}}$ and $\mathbf{M}_{n_{\text{e}}}$. The joint posterior distribution of $\mathbf{T}_{\text{e}}$, $\mathbf{n}_{\text{e}}$, C_{TS}, $\mathbf{M}_{T_{\text{e}}}$ and $\mathbf{M}_{n_{\text{e}}}$ is explored by MCMC sampling with the adaptive Metropolis-Hastings algorithm [38, 39, 40]. The inference results for JET discharge #88630 at 10.526 s (L-mode) are shown in Figure 7:
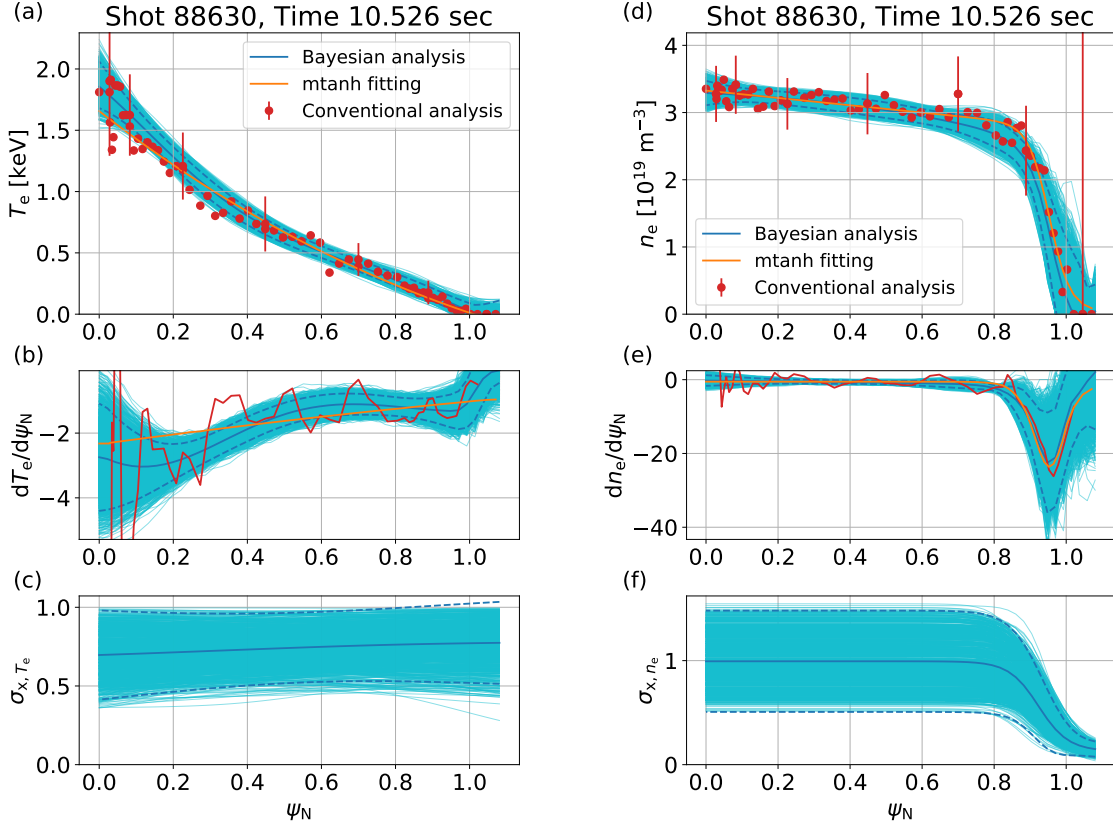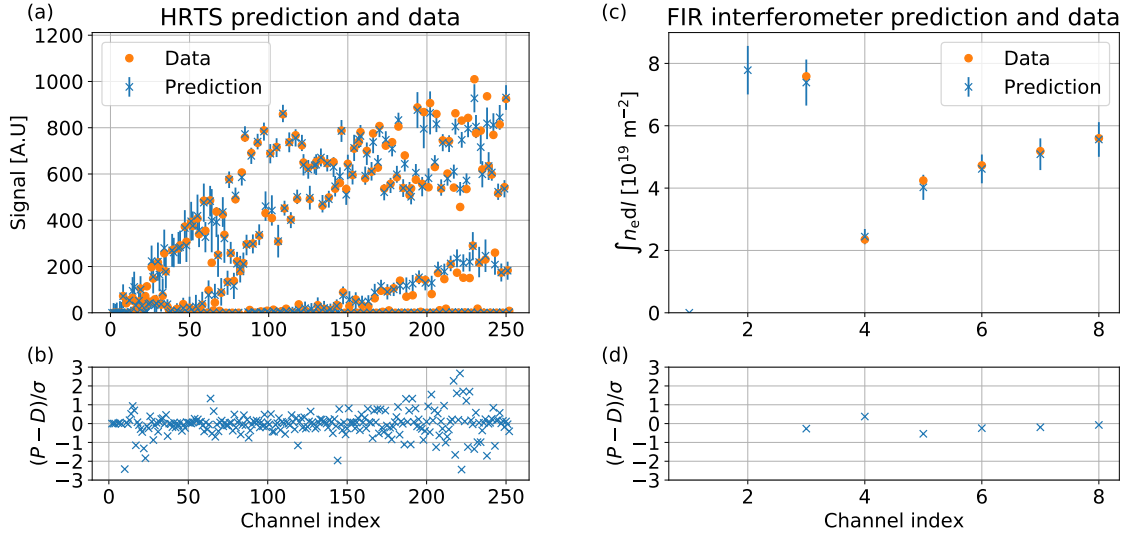
Figure 7. Inference results for JET discharge #88630 at $10.526\,\mathrm{s}$ (L-mode): (a) T_e profiles, (b) T_e gradient, (c) length scale of T_e profiles, (d) n_e profiles, (e) n_e gradient and (f) length scale of n_e profiles. The thick and light blue lines are the mean and samples, respectively, of the marginal joint posterior distributions. The blue dashed lines are the lower and upper boundaries of two standard deviation ($\pm 2\sigma$) marginal posterior uncertainty bands. For comparison, the electron temperature and density profiles (red dots with $\pm 2\sigma$ error bars on a few points) and gradients (red lines) from the conventional analysis of the HRTS system [20] are shown. The results (orange lines) of a fitted modified hyperbolic tangent function [33] to the conventional analysis are also presented here. The red vertical line in (d) indicates a large uncertainty of the conventional analysis outside of the last closed flux surface due to a small signal-to-noise ratio.

(a) T_e profiles, (b) T_e gradient, (c) length scale of T_e profiles, (d) n_e profiles, (e) n_e gradient and (f) length scale of n_e profiles. The electron temperature(density) profiles are obtained by marginalising over the electron density(temperature) profiles, the calibration factor C_TS and hyperparameters $\mathbf{M}_{T_\mathrm{e}}$ and $\mathbf{M}_{n_\mathrm{e}}$. Similarly, the length scale profiles of T_e and n_e are obtained by marginalising over all the other parameters. The gradient profiles are calculated numerically from these marginalised T_e and n_e profiles. The thick and light blue lines are the mean and samples, respectively, of the marginal joint posterior distributions. The blue dashed lines are the lower and upper boundaries of two standard deviation ($\pm 2\sigma$) marginal posterior uncertainty bands.

For comparison, we show, in Figure 7, the electron temperature and density profiles

Figure 8. The data (orange dots) and predictions (blue crosses with $\pm1\sigma$ error bars) of (a) the HRTS and (c) the FIR interferometer systems for JET discharge #88630 at $10.526\,\mathrm{s}$. The differences between the predictions P and data D divided by the uncertainties σ are shown in (b) and (d). Note that the channels #1 and #2 of the FIR interferometer system were not available for this discharge, hence no measured line integrated densities for these channels.

(red dots with $\pm2\sigma$ error bars on a few points) and T_e and n_e gradient profiles (red lines) from the conventional analysis of the HRTS system [20]. The gradient profiles from the conventional analysis are calculated from two neighbouring points then smoothened by a simple moving mean with five points. We also present the results (orange lines) of a fitted modified hyperbolic tangent function [33] to the conventional analysis. The Bayesian method reproduces the gradient profiles which agree with those from the conventional method within the marginal posterior uncertainties.

The data (orange dots) and predictions (blue crosses with $\pm1\sigma$ error bar) of the HRTS and the FIR interferometer systems are shown in Figure 8 (a) and (c). The predictions are calculated through Equations (1) and (5) given the mean of the joint posterior distribution. The differences between the predictions P and data D divided by the uncertainties σ are also shown in Figure 8 (b) and (d). The data and predictions agree within the uncertainties.

An example of the electron temperature and density profiles of a JET H-mode discharge (#89380 at $11.776\,\mathrm{s}$) is shown in Figure 9. The Gaussian processes infer the pedestal gradients of the electron temperature and density profiles by taking account of all possible hyperparameter values and marginalising them out. Note that we have not included the spatial instrument response of the HRTS system because all the HRTS data used in this work are obtained from the present HRTS configuration that affects the pedestal gradient minimally [21].

We emphasise that our method inferring profiles of electron temperature and density
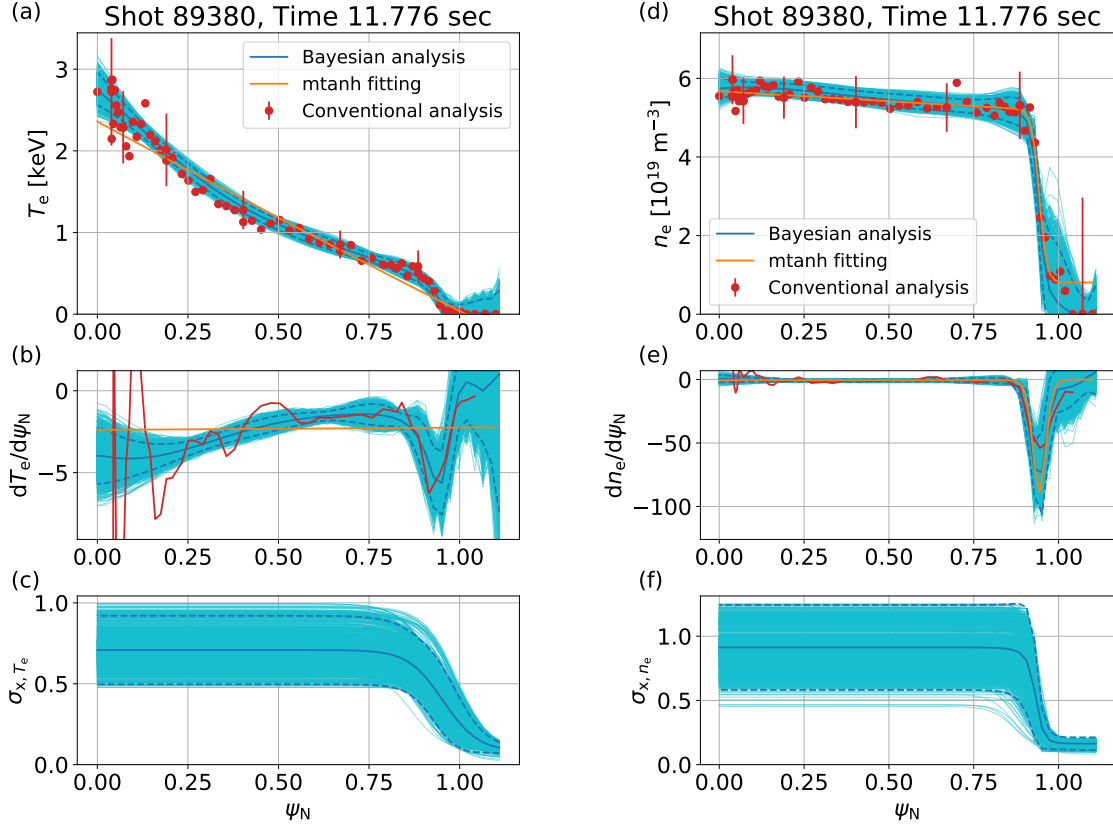
Figure 9. Same as Figure 7 for a JET H-mode discharge (#89380 at 11.776 s).

do not depend on any parametrisation due to the usage of non-parametric Gaussian processes with the hyperparameters. The choice of Gaussian process class will result in different prior probabilities for different profile shapes. At the same time, the Gaussian process is still a universal approximator, so given enough data the actual underlying profile would be recovered. It is very hard to estimate, even define, the bias introduced by the choice of Gaussian process family, but within a family the marginalisation over hyperparameters should objectively choose proper smoothnesses. The choice of Gaussian process family (in this case the hyperbolic tangent function) will in the end correspond to a physics assumption, though a much weaker one than a strict parameterised function. In no case though will the shape be related (not even weakly) to a hyperbolic tangent curve. Thus, non-hyperbolic tangent profile shapes, e.g., hollow profiles, can be inferred as shown in Figure 10 where the profile shapes are fully determined by the data, taking account of all possible combinations of C_{TS}, $\mathbf{M}_{T_e}$ and $\mathbf{M}_{n_e}$.

This Gaussian process family well reproduces all profile shapes we have examined including highly non-monotonic profiles. Nevertheless, there might be cases which are not well represented by this Gaussian process family. Extensive usage of our method will show how versatile this Gaussian process parameterisation is.
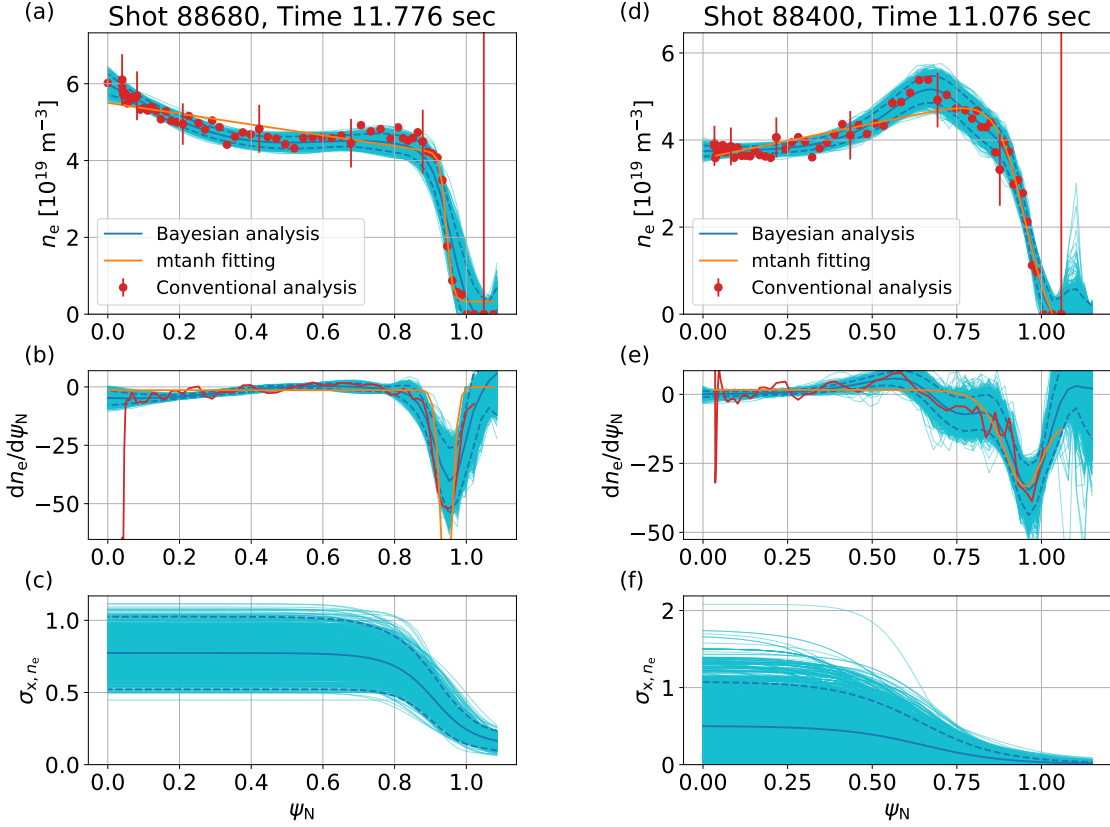
Figure 10. Same as Figure 7 for n_e inference results for (a), (b) and (c) JET discharge #88680 at 11.776 s (H-mode plasma with 6 MW neutral beam injection) and (d), (e) and (f) #88400 at 11.076 s (H-mode plasma with pellet injection).

3.3. Discussion on the calibration factor C_{TS}

The electron density calibration factor C_{TS} for the HRTS system is set as a single unknown parameter for all the spectral channels of all the polychromators and inferred by sampling the joint posterior distribution, i.e., Equation (12), that can explain both the measured HRTS and FIR interferometer data. Figure 11 shows the mean (thick segmented blue lines) and the samples (light blue lines) from the marginalised C_{TS} posterior distribution over all the other parameters which are electron temperature, density and their associated hyperparameters for multiple time points of several JET discharges.

The inferred calibration factor may provide us useful information on the HRTS system. For instance, relatively large variation of the calibration factor within a single discharge, e.g., JET discharge #88380 in Figure 11, may be caused by: i) a fluctuation of the laser energy, ii) a laser misalignment or iii) over/under-estimation of the plasma volume via the EFIT code. The overall trend of the calibration factor over many discharges may suggest us unforeseen slow aberration of the instruments such as contamination on the collecting optics (by dust for example). Our method can handle the temporal
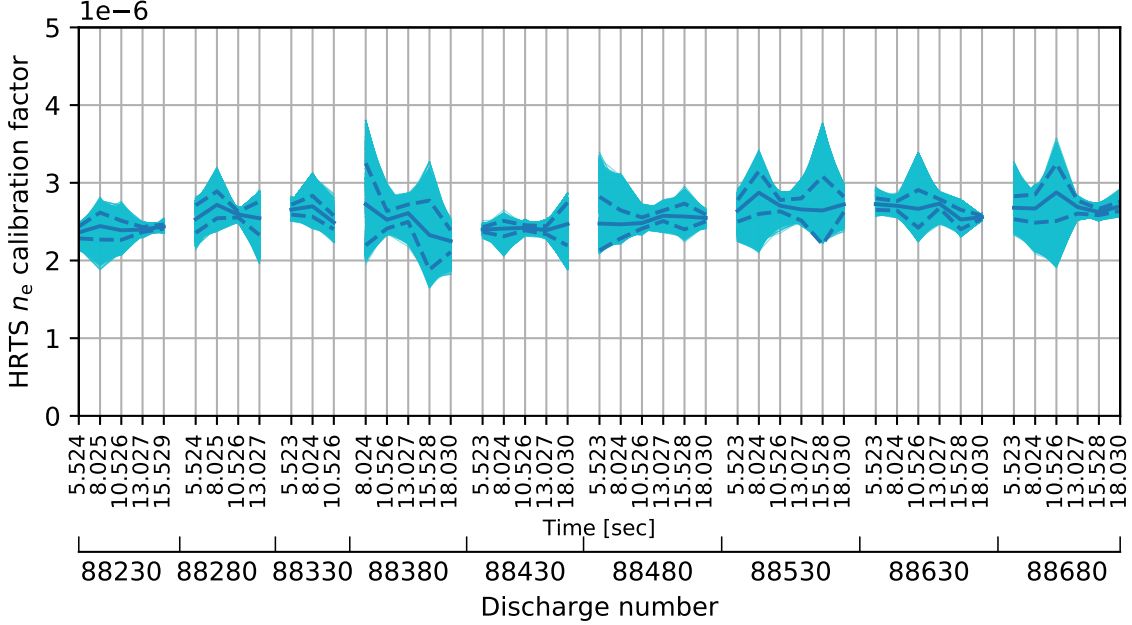
Figure 11. Evolution of the electron density calibration factor C_{TS} over multiple JET discharges. The upper and lower labels on the abscissa are the time points and the discharge numbers, respectively. The thick segmented blue lines show the mean of the samples (light blue lines) from the marginalised C_{TS} posterior distribution. Evolution of C_{TS} may provide us how the HRTS system vary over time. The blue dashed lines are the lower and upper boundaries of one standard deviation marginal posterior uncertainty bands.

variation of C_{TS} as well as the uncertainties of C_{TS} and its influence on the profile quantities.

4. Conclusions

We have developed and presented a Bayesian inference scheme for electron temperature and density profiles using non-parametric Gaussian processes consistent with the high resolution Thomson scattering (HRTS) and the far infrared (FIR) interferometer data from JET. The forward models of both systems are constructed within the Minerva framework, individually and combined together. Our method consists of two steps: i) extracting the amplitudes of the Thomson scattered signals from the raw HRTS measurements, and ii) inferring the electron temperature and density profiles.

The raw HRTS measurements contain not only the Thomson scattered but also the parasitic signals which are external Raman scattered and stray light signals. These signals are carefully modelled by taking account of the instrument effects with the associated uncertainties which are electronics and photon noises. The joint posterior distribution of the Thomson scattered and the parasitic signals are explored by Markov chain Monte Carlo (MCMC) sampling, and the amplitudes of the Thomson scattered signals are

obtained by marginalising out all the other parameters. The mean and variance of the amplitudes define the observed quantities for the profile inference.

The electron temperature and density profiles are modelled by non-parametric Gaussian processes. The profiles are inferred from the observed amplitudes of the Thomson scattered signals and the FIR interferometer data by exploring (with MCMC sampling) the joint posterior distribution of the electron temperature and density profiles, the electron density calibration factor of the HRTS system and the hyperparameters. The electron temperature and density profiles are obtained by marginalising out all the other parameters. Therefore, these profiles do not depend on a specific value of hyperparameters nor any parametric regressions which may restrain the shape of the profiles significantly. In addition, combining the HRTS and the FIR interferometer data allows us to infer the calibration factor and its uncertainty, and it may provide us knowledge on unforeseen aberration of the diagnostic systems over time.

These inference results, including the profile samples, can be used for further advanced investigation such as transport analysis with TRANSP [49]. Samples of gradient profiles can be fed to transport codes to extract uncertainty information on calculated physics parameters.

5. Acknowledgement

Appendix A. Details on the Thomson scattering model, i.e., Equation (1)

Thomson scattered energy E per unit solid angle Ω per unit wavelength λ that depends on the electron temperature T_{e} and density n_{e} is given by Naito formula [41]

$$\frac{\partial^2 E}{\partial \Omega \partial \lambda} = r_{\mathrm{e}}^2 n_{\mathrm{e}} E_{\mathrm{laser}} L \frac{S\left(\lambda, \theta, T_{\mathrm{e}}\right)}{\lambda_{\mathrm{laser}}}, \tag{A.1}$$

where r_{e} is the classical electron radius, E_{laser} the energy of incident laser, L a scattering length, S the spectral density function which depends on scattering wavelength λ and angle θ in addition to the electron temperature T_{e}. Since $N = \frac{\lambda}{hc}E$ where N is the number of photons at the wavelength λ, Equation (A.1) can be rewritten as

$$\frac{\partial^2 N}{\partial \Omega \partial \lambda} = \frac{\lambda}{hc} \frac{\partial^2 E}{\partial \Omega \partial \lambda}, \tag{A.2}$$

where h is the Planck constant and c the speed of light.

Taking account of spectral response functions of Thomson scattering diagnostic systems (in this case, the HRTS system) $\phi(\lambda)$, which include the transmittance of optics and polychromator filter functions, the number of collected photons N_{TS} is

$$
\begin{aligned}
N_{\mathrm{TS}} &= \iint \phi(\lambda)\, \frac{\partial^2 N}{\partial\Omega\partial\lambda}\, \mathrm{d}\Omega\mathrm{d}\lambda \\
&= \iint \phi(\lambda)\, \frac{\lambda}{hc} r_{\mathrm{e}}^2 n_{\mathrm{e}} E_{\mathrm{laser}} L \frac{S(\lambda,\theta,T_{\mathrm{e}})}{\lambda_{\mathrm{laser}}}\, \mathrm{d}\Omega\mathrm{d}\lambda \\
&\approx L\Delta\Omega\, n_{\mathrm{e}} E_{\mathrm{laser}} \int \phi(\lambda)\, \frac{\lambda}{hc} r_{\mathrm{e}}^2 \frac{S(\lambda,\theta,T_{\mathrm{e}})}{\lambda_{\mathrm{laser}}}\, \mathrm{d}\lambda,
\end{aligned}
\tag{A.3}
$$

and the last line is obtained by approximating $\int \mathrm{d}\Omega$ to $\Delta\Omega$.

Detectors convert the collected photons to an electronics signal A_{TS} with their gain factor G, thus we have

$$
A_{\mathrm{TS}} = G\, L\Delta\Omega\, \phi^{1,1}(\lambda_{\mathrm{N}})\, n_{\mathrm{e}} E_{\mathrm{laser}} \int \frac{\phi(\lambda)}{\phi^{1,1}(\lambda_{\mathrm{N}})} \frac{\lambda}{hc} r_{\mathrm{e}}^2 \frac{S(\lambda,\theta,T_{\mathrm{e}})}{\lambda_{\mathrm{laser}}}\, \mathrm{d}\lambda,
\tag{A.4}
$$

where $\phi^{1,1}(\lambda_{\mathrm{N}})$ is a normalisation factor for the spectral response functions (the value of the spectral response function of the first spectral channel of the first spatial position at the wavelength $\lambda_{\mathrm{N}} = 1020\,\mathrm{nm}$ in our case).

Equation (1) is obtained by letting the electron density calibration factor $C_{\mathrm{TS}} = G\, L\Delta\Omega\, \phi^{1,1}(\lambda_{\mathrm{N}})$, that is

$$
A_{\mathrm{TS}} = C_{\mathrm{TS}}\, n_{\mathrm{e}} E_{\mathrm{laser}} \int \frac{\phi(\lambda)}{\phi^{1,1}(\lambda_{\mathrm{N}})}\,(\lambda)\, \frac{\lambda}{hc} r_{\mathrm{e}}^2 \frac{S(\lambda,\theta,T_{\mathrm{e}})}{\lambda_{\mathrm{laser}}}\, \mathrm{d}\lambda.
\tag{A.5}
$$

References

[1] Litaudon X *et al.* 2017 *Nuclear Fusion* **57** 102001 ISSN 0029-5515 URL http://stacks.iop.org/0029-5515/57/i=10/a=102001?key=crossref.17a0f33a13c4a2bbaaafb54b4eb7df30

[2] Wolf R *et al.* 2017 *Nuclear Fusion* **57** 102020 URL https://doi.org/10.1088%2F1741-4326%2Faa770d

[3] Seed eScience Research *The Minerva framework* URL https://seed-escience.org/

[4] Svensson J and Werner A 2007 Large Scale Bayesian Data Analysis for Nuclear Fusion Experiments *2007 IEEE International Symposium on Intelligent Signal Processing* (IEEE) pp 1–6 ISBN 978-1-4244-0829-0 URL http://ieeexplore.ieee.org/document/4447579/

[5] Svensson J 2011 *JET report, EFDA–JET–PR(11)24* URL http://www.euro-fusionscipub.org/wp-content/uploads/eurofusion/EFDP11024.pdf

[6] Svensson J and Werner A 2008 *Plasma Physics and Controlled Fusion* **50** 085002 ISSN 0741-3335 URL http://stacks.iop.org/0741-3335/50/i=8/a=085002?key=crossref.8b9e7d2d10e66d8a740fcec9ab77227d

[7] von Nessi G T, Hole M J, Svensson J and Appel L 2012 *Physics of Plasmas* **19** 012506 ISSN 1070-664X URL http://aip.scitation.org/doi/10.1063/1.3677362

[8] von Nessi G T and Hole M J 2014 *Plasma Physics and Controlled Fusion* **56** 114011 ISSN 0741-3335 URL http://stacks.iop.org/0741-3335/56/i=11/a=114011?key=crossref.d58e08b3c9223c89bab07f760e99229e

[9] Bozhenkov S, Beurskens M, Molin A D, Fuchert G, Pasch E, Stoneking M, Hirsch M, Höfel U, Knauer J, Svensson J, Mora H T and Wolf R 2017 *Journal of Instrumentation* **12** P10004–P10004 ISSN 1748-0221 URL http://stacks.iop.org/1748-0221/12/i=10/a=P10004?key=crossref.21c7d556bea6caf012906a098b133ef0

[10] Kwak S, Svensson J, Brix M and Ghim Y c 2016 *Review of Scientific Instruments* **87** 023501 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.4940925

[11] Kwak S, Svensson J, Brix M and Ghim Y C 2017 *Nuclear Fusion* **57** 036017 ISSN 0029-5515 URL http://stacks.iop.org/0029-5515/57/i=3/a=036017?key=crossref.78317e8d5c69c0d93e4bfd4af120a500

[12] Li D, Svensson J, Thomsen H, Medina F, Werner A and Wolf R 2013 *Review of Scientific Instruments* **84** 083506 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.4817591

[13] Schmuck S, Svensson J, De La Luna E, Figini L, Johnson T, Alper B, Beurskens M, Fessey J, Gerbaud T and Sirinelli A 2011 Bayesian derivation of electron temperature profile using jet ece diagnostics *38th EPS Conference on Plasma Physics 2011, EPS 2011 : Europhysics Conference Abstracts* (*Europhysics Conference Abstracts* no 35:2) pp 1512–1515 qC 20140828

[14] Langenberg A, Svensson J, Thomsen H, Marchuk O, Pablant N A, Burhenn R and Wolf R C 2016 *Fusion Science and Technology* **69** 560–567 ISSN 15361055 URL https://www.tandfonline.com/doi/full/10.13182/FST15-181

[15] Krychowiak M 2016 *Review of Scientific Instruments* **87** 11D304 ISSN 10897623 URL http://scitation.aip.org/content/aip/journal/rsi/87/11/10.1063/1.4964376

[16] Mora H T, Bozhenkov S, Knauer J, Kornejew P, Kwak S, Ford O, Fuchert G, Pasch E, Svensson J, Werner A, Wolf R and Timmermann D 2017 FPGA acceleration of Bayesian model based analysis for time-independent problems *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (IEEE) pp 774–778 ISBN 978-1-5090-5990-4 URL http://ieeexplore.ieee.org/document/8309065/

[17] Pavone A, Svensson J, Langenberg A, Pablant N, Hoefel U, Kwak S, Wolf R C and Team W X 2018 *Review of Scientific Instruments* **89** 10K102 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.5039286

[18] Lao L, St John H, Stambaugh R, Kellman A and Pfeiffer W 1985 *Nuclear Fusion* **25** 1611–1622 ISSN 0029-5515 URL http://stacks.iop.org/0029-5515/25/i=11/a=007?key=crossref.382b4e7e430c8741af0f7248e9a56c09

[19] O'Brien D, Lao L, Solano E, Garribba M, Taylor T, Cordey J and Ellis J 1992 *Nuclear Fusion* **32** 1351–1360 URL https://doi.org/10.1088%2F0029-5515%2F32%2F8%2Fi05

[20] Pasqualotto R, Nielsen P, Gowers C, Beurskens M, Kempenaars M, Carlstrom T and Johnson D 2004 *Review of Scientific Instruments* **75** 3891–3893 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.1787922

[21] Frassinetti L, Beurskens M N A, Scannell R, Osborne T H, Flanagan J, Kempenaars M, Maslov M, Pasqualotto R and Walsh M 2012 *Review of Scientific Instruments* **83** 013506 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.3673467

[22] Scannell R 2007 *Investigation of H-mode edge profile behaviour on MAST using Thomson scattering* Ph.D. thesis University College Cork

[23] Braithwaite G, Gottardi N, Magyar G, O'Rourke J, Ryan J and Véron D 1989 *Review of Scientific Instruments* **60** 2825–2834 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.1140666

[24] Boboc A, Gelfusa M, Murari A and Gaudio P 2010 *Review of Scientific Instruments* **81** 10D538 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.3478146

[25] Boboc A, Gil C, Pastor P, Spuig P, Edlington T and Dorling S 2012 *Review of Scientific Instruments* **83** 10E341 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.4737420

[26] Boboc A, Bieg B, Felton R, Dalley S and Kravtsov Y 2015 *Review of Scientific Instruments* **86** 091301 ISSN 0034-6748 URL http://aip.scitation.org/doi/10.1063/1.4929443

[27] Fischer R, Wendland C, Dinklage A, Gori S, Dose V and Teams W A 2002 *Plasma Physics and*

Controlled Fusion **44** 1501–1519 ISSN 07413335 URL `http://stacks.iop.org/0741-3335/44/i=8/a=306?key=crossref.90d7ee0715d647125b222b2bd64bea16`

[28] Park K R, Kim K h, Kwak S, Svensson J, Lee J and Ghim Y c 2017 *Journal of Instrumentation* **12** C11022–C11022 ISSN 1748-0221 URL `http://stacks.iop.org/1748-0221/12/i=11/a=C11022?key=crossref.3a526848b7d41b2c69eed8b828346290`

[29] Ford O P 2010 *Tokamak plasma analysis through Bayesian diagnostic modelling* Ph.D. thesis Imperial College London

[30] Fischer R, Fuchs C J, Kurzan B, Suttrop W and Wolfrum E 2010 *Fusion Science and Technology* **58** 675–684 ISSN 1536-1055 URL `https://www.tandfonline.com/doi/full/10.13182/FST10-110`

[31] Rasmussen C E and Williams C K I 2006 *Gaussian Processes for Machine Learning* (MIT Press)

[32] Chilenski M, Greenwald M, Marzouk Y, Howard N, White A, Rice J and Walk J 2015 *Nuclear Fusion* **55** 023012 ISSN 0029-5515 URL `http://stacks.iop.org/0029-5515/55/i=2/a=023012?key=crossref.b22d32b1ac570adef0ea0869bc2c1789`

[33] Groebner R J, Mahdavi M A, Leonard A W, Osborne T H and Porter G D 2002 *Plasma Physics and Controlled Fusion* **44** 326 ISSN 07413335 URL `http://stacks.iop.org/0741-3335/44/i=5A/a=326?key=crossref.ab9f87434e30708c202f3d17568e5cc5`

[34] Carlstrom T, Burrell K, Groebner R, Leonard A, Osborne T and Thomas D 1999 *Nuclear Fusion* **39** 1941–1947 URL `https://doi.org/10.1088%2F0029-5515%2F39%2F11y%2F338`

[35] Connor J W and Wilson H R 1999 *Plasma Physics and Controlled Fusion* **42** R1–R74 URL `https://doi.org/10.1088%2F0741-3335%2F42%2F1%2F201`

[36] Leyland M J, Beurskens M N A, Flanagan J C, Frassinetti L, Gibson K J, Kempenaars M, Maslov M and Scannell R 2016 *Review of Scientific Instruments* **87** 013507 ISSN 0034-6748 URL `http://aip.scitation.org/doi/10.1063/1.4939855`

[37] Fischer R, Wolfrum E and Schweinzer J 2008 *Plasma Physics and Controlled Fusion* **50** 085009 ISSN 0741-3335 URL `http://stacks.iop.org/0741-3335/50/i=8/a=085009?key=crossref.66dc3a9cef01c87d8417ae7a2c0d3d6a`

[38] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 *The Journal of Chemical Physics* **21** 1087–1092 ISSN 0021-9606 URL `http://aip.scitation.org/doi/10.1063/1.1699114`

[39] Hastings W K 1970 *Biometrika* **57** 97–109 ISSN 1464-3510 URL `https://academic.oup.com/biomet/article/57/1/97/284580`

[40] Haario H, Saksman E and Tamminen J 2001 *Bernoulli* **7** 223 ISSN 13507265 URL `https://www.jstor.org/stable/3318737?origin=crossref`

[41] Naito O, Yoshida H and Matoba T 1993 *Physics of Fluids B: Plasma Physics* **5** 4256–4258 ISSN 0899-8221 URL `http://aip.scitation.org/doi/10.1063/1.860593`

[42] Kurzan B, Jakobi M, Murmann H and Team A U 2004 *Plasma Physics and Controlled Fusion* **46** 299–317 ISSN 0741-3335 URL `http://stacks.iop.org/0741-3335/46/i=1/a=019?key=crossref.90ae7dc904113006bcfc986aa1ddac7a`

[43] Romero J and Svensson J 2013 *Nuclear Fusion* **53** 033009 ISSN 0029-5515 URL `http://stacks.iop.org/0029-5515/53/i=3/a=033009?key=crossref.e63bc2641d1ea5a6766bf4b0dc1db9ee`

[44] Romero J A, Dettrick S A, Granstedt E, Roche T and Mok Y 2018 *Nature Communications* **9** 691 ISSN 2041-1723 URL `http://www.nature.com/articles/s41467-018-03110-5`

[45] ASDEX Team 1989 *Nuclear Fusion* **29** 1959–2040 ISSN 17414326 URL `http://stacks.iop.org/0029-5515/29/i=11/a=010?key=crossref.f8eff2d027380946fd88742ac88d929c`

[46] Higdon D, Swall J and Kern J 1999 Non-stationary spatial modeling *Bayesian Statistics 6 Proceedings of the Sixth Valencia International Meeting* pp 761–768

[47] Devinderjit Sivia J S 2006 *Data Analysis: A Bayesian Tutorial* (Oxford University Press) ISBN 0-198-56831-2

[48] Jaynes E T 2003 *Probability Theory: The Logic of Science* (Cambridge University Press) ISBN 0-521-59271-2

[49] Hawryluk R 1981 AN EMPIRICAL APPROACH TO TOKAMAK TRANSPORT *Physics of Plasmas Close to Thermonuclear Conditions* (Elsevier) pp 19–46 URL `https://linkinghub.elsevier.com/retrieve/pii/B9781483283852500091`