

# Development and simulation of multi-diagnostic Bayesian analysis for 2D inference of divertor plasma characteristics

C Bowman<sup>1</sup> , J R Harrison<sup>2</sup> , B Lipschultz<sup>1</sup> , S Orchard<sup>1</sup>, K J Gibson<sup>1</sup>, M Carr<sup>2</sup>, K Verhaegh<sup>2</sup>  and O Myatra<sup>1</sup>

<sup>1</sup> York Plasma Institute, Department of Physics, University of York, Heslington, York YO10 5DD, United Kingdom

<sup>2</sup> Culham Centre for Fusion Energy (CCFE), Culham Science Centre, Abingdon, Oxfordshire, OX14 3DB, United Kingdom

Received 31 October 2019, revised 6 February 2020

Accepted for publication 12 February 2020

Published 26 February 2020



## Abstract

We present results of the design, implementation and testing of a Bayesian multi-diagnostic inference system which combines various divertor diagnostics to infer the 2D fields of electron temperature  $T_e$ , density  $n_e$  and deuterium neutral density  $n_0$  in the divertor. The system was tested using synthetic diagnostic measurements derived from SOLPS-ITER fluid code predictions of the MAST-U Super-X divertor which include appropriate added noise. Two SOLPS-ITER simulations in different states of detachment, taken from a scan of the nitrogen seeding rate, were used as test-cases. Taken across both test-cases, the median absolute fractional errors in the inferred electron temperature and density estimates were 10.3% and 10.1% respectively. Differences between the inferred fields and the test-cases were well explained by solution uncertainty estimates derived from posterior sampling. This work represents a step toward a larger goal of obtaining a quantitative, 2D description of the divertor plasma state directly from experimental data, which could be used to gain better understanding of divertor physics phenomena.

Keywords: Bayesian inference, divertor physics, Tokamaks

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

## 1. Introduction

### 1.1. Limitations of conventional divertor diagnostic analysis

The divertor of a magnetic confinement fusion device is a complex system involving transport, atomic, molecular and impurity processes in the plasma as well as at the divertor surfaces, all giving rise to energy, momentum and particle sources and sinks [1]. These processes are influenced by other aspects of the divertor such as the divertor geometry (e.g. what fraction of recycled neutrals escape the divertor) and magnetic topology. All of the above make it difficult to separate out the effects of individual processes to verify whether our physics understanding,

embodied in 2D models of the divertor such as SOLPS-ITER [2], are correct.

Despite the variety of diagnostic systems available in the divertor, they each have limitations such that any single instrument cannot directly determine the 2D fields of plasma characteristics (e.g. electron/ion temperature and density) or the properties of the neutrals (e.g. atomic and molecular densities). For example, Langmuir probes and Thomson scattering systems can directly measure some of these fields, but do so only at a series of isolated points. Filtered camera imaging systems can collect information from a large fraction of the divertor cross-section, but provide line-integrated measurements of spectral line emissivities, which are a complicated function of the underlying plasma fields [3].

Due to these limitations, studying the physics of divertor plasmas has often relied on matching the predictions of codes like SOLPS-ITER to diagnostic measurements, to find a set of



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

plasma fields which are consistent with the available data. This matching process can be extremely time intensive, and typically requires ‘by-hand’ tuning of input parameters (e.g. recycling coefficients, boundary conditions such as upstream density, transport coefficients etc) over many iterations, and a period of weeks to months. Conceptually, in this approach assumptions are made regarding how divertor physics processes and boundary conditions determine the divertor plasma state, which then implies a corresponding set of expected diagnostic measurements.

We propose to instead take the inverse approach, where starting from the diagnostic measurements the plasma state is inferred, and from the inferred plasma state the underlying physics processes can be determined. Here we demonstrate that the first part of this approach, direct inference of the divertor plasma state, is possible using an ‘integrated’ approach to divertor analysis in which data from multiple diagnostic systems are combined.

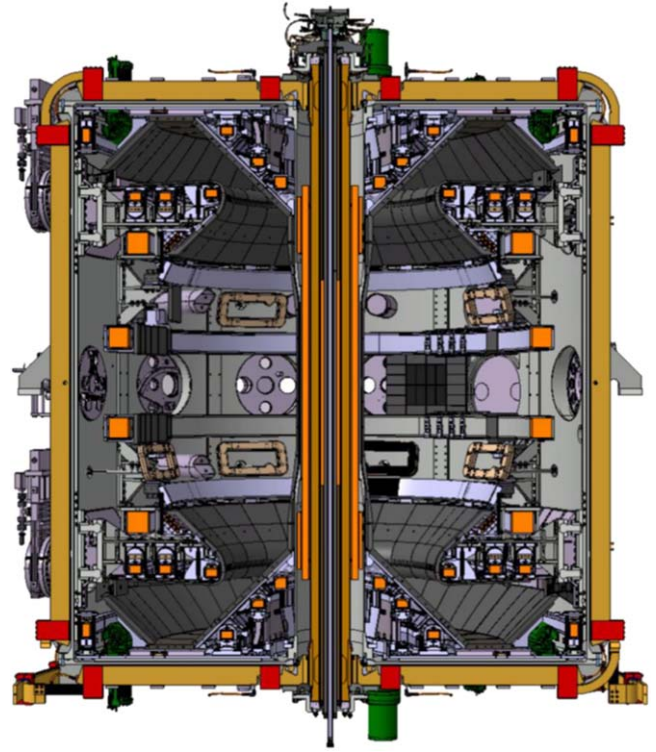
Such an integrated approach, if successful, would not serve as a replacement for 2D divertor modelling codes. Rather it is an alternative path to studying the role of various divertor processes and how they vary during and across discharges. It may provide an independent test of the validity of the physics we believe is responsible for determining the divertor plasma state, and is implemented within 2D divertor modelling codes.

For the purposes of this study we use the geometry and planned diagnostics of the MAST-U spherical tokamak as a test-case to investigate integrated divertor diagnostic analysis [4]. The MAST-U divertor will be well diagnosed, possessing a multi-wavelength imaging (MWI) system based on the MANTIS system at TCV [5], which can simultaneously image the divertor for each of up to 10 atomic lines, spectrometers, bolometers, Langmuir probes and a dedicated divertor Thomson scattering system. A cross-section of MAST-U is shown in figure 1, and the coverage of diagnostics relevant to our analysis is illustrated in figure 2.

### 1.2. An integrated, Bayesian approach to divertor analysis

We will make use of the Bayesian approach to data analysis, in which probability is used as a means of quantifying the information content of experimental data with respect to model parameters. By formalising the information content in this way, we are able to combine data from multiple diagnostics in order to strengthen our knowledge of the plasma fields. This is highly desirable, but comes at the cost of an increase in the complexity and computational expense of the data analysis, as typically all data must be analysed simultaneously. Multi-diagnostic Bayesian analysis has been successfully applied within tokamak plasma studies for profile diagnostic analysis [6, 7] and equilibrium reconstruction [8, 9], but not yet to inference of the 2D divertor plasma state.

Here we discuss the design, implementation and testing of a Bayesian multi-diagnostic inference system for the MAST-U Super-X divertor which aims to infer the fields of plasma electron temperature and density, and hydrogen neutral density, throughout the divertor, including associated uncertainties. In



**Figure 1.** Cross-section of the MAST-Upgrade tokamak. The Super-X outer divertor configuration can be seen at the top and bottom of the device.

section 2 we discuss the parameterisation of the problem and design requirements of the system. In sections 3 and 4 we show how information regarding the plasma fields in both measurement data and prior knowledge may be expressed as probability distributions. In section 5 the construction of synthetic test-cases using SOLPS simulations is discussed. In section 6 we discuss the numerical strategies used to characterise the posterior distribution for the plasma fields. The results of analysing the synthetic data are presented in section 7. A discussion of potential improvements and further work is given in section 8, followed by conclusions in section 9.

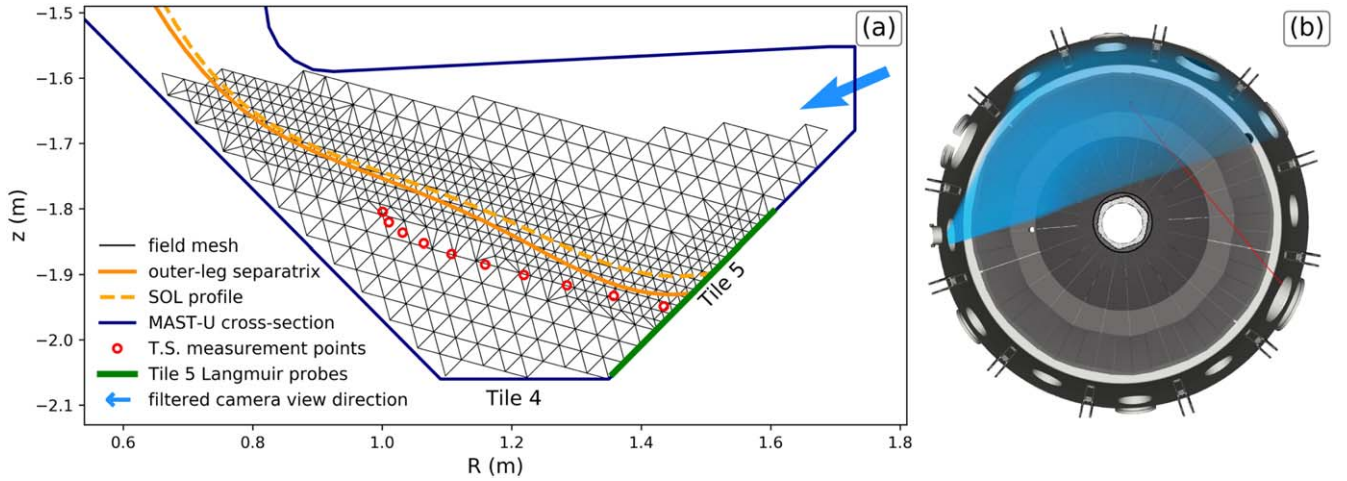
## 2. System design

### 2.1. Parametric representation of plasma fields

We choose to represent each of the 2D plasma fields via linear interpolation on a triangular mesh, shown in figure 2, which covers the relevant areas of the divertor cross-section. Specifically, this means that by defining the value of a field at each vertex of the mesh, that field is defined continuously inside each triangle of the mesh as the plane that connects the three points which define that triangle.

Using this approach a field, for example the electron temperature field  $T_e(R, z)$ , is defined as

$$T_e(R, z) = \sum_{k=1}^V T_e^{(k)} \phi_k(R, z) \quad (1)$$



**Figure 2.** Illustration of diagnostic coverage in the MAST-U Super-X divertor region. (a) Poloidal cross-section showing the triangular mesh used to represent the plasma fields. The filtered cameras (MWI) have a view which covers the entire mesh due to the fact their viewing direction has both a poloidal and toroidal component (see (b)). The Thomson scattering measurement points are located in the private-flux region for this particular equilibrium, which is used for MAST-U SOLPS simulations, but can also be located around the separatrix or in the scrape-off layer depending on experimental set-up. Only the Langmuir probes on tile 5 (highlighted in green) are used for this work. However, others exist on most divertor surfaces. The dashed line shows a scrape-off layer flux surface which is later used to produce profiles of plasma properties. (b) Top-down view of the Super-X divertor chamber showing the filtered-camera field-of-view (in blue) and the Thomson-scattering laser path (in red).

where  $T_e^{(k)}$  is the electron temperature at vertex  $k$ ,  $\phi_k(R, z)$  is the linear interpolation basis function for vertex  $k$ , and  $V$  is the number of mesh vertices. This model for the plasma fields has the advantage that the model parameters themselves are the values of each field at each mesh vertex, allowing physics constraints to be easily applied. For example, to ensure that the electron temperature field is greater than zero everywhere, we need only ensure that the parameters which set the temperature at each vertex are greater than zero, i.e. that  $T_e^{(k)} > 0 \forall k$ .

The mesh shown in figure 2 was used to produce all results presented in this paper. It was generated by first creating a mesh of equilateral triangles of side length 35 mm which aligns with a toroidally-symmetric approximation of tile 5, where the outer strike-point will typically be located. In select regions of the mesh covering the expected position of the divertor leg and strike point, triangles were partitioned to produce a higher-resolution area with side-lengths of 17.5 mm, yielding a refined mesh with  $V = 586$  vertices.

## 2.2. Design requirements

To guide the direction of the system design a set of requirements were chosen. Firstly, we want to be able to choose easily which diagnostics are included in the analysis. This means that diagnostic systems should be able to be added or removed from the analysis without making direct alterations to the system code. Instead, there should be a ‘higher-level’ interface for specifying the choice of diagnostics.

A key part of the system are the diagnostic forward models. Also sometimes referred to as ‘synthetic diagnostic’ models, forward models simulate the experimental data we would expect to measure using a particular instrument under a given set of plasma conditions (in this case, the 2D fields of

electron temperature  $T_e$ , electron density  $n_e$  and hydrogen neutral density  $n_0$  defined by the mesh). There may be many possible forward models for a given diagnostic, which vary based on the physics assumptions they make, their level of complexity and their computational cost. As before, we want an interface which allows us to specify which model is used for a given diagnostic system without making changes to the code.

## 2.3. Choice of diagnostics

In order to infer  $T_e$ ,  $n_e$  and  $n_0$  without including additional fields, we must choose diagnostics whose measurements can be predicted using these fields only. The divertor Thomson scattering system is able to make direct measurements of  $T_e$  and  $n_e$  making it an obvious choice [10]. We also include the target Langmuir probes on tile 5 (shown in figure 2) as they are located where the primary heat and particle fluxes are incident on a divertor surface. The probe measurements, under certain assumptions, can also be modelled using only  $T_e$  and  $n_e$  [11].

The emissivity of line radiation has a dependence on both  $T_e$  and  $n_e$ , so any filtered camera data will carry information about the  $T_e$ ,  $n_e$  fields. However, the emissivity of a given line also depends on the density field of the corresponding emitting species, which is charge-state and metastable-state resolved. For the initial development and testing of this technique, we consider only hydrogenic line emission, which can be modelled using only  $T_e$ ,  $n_e$  and  $n_0$ . Consequently we include 4 filtered cameras which view the hydrogen Balmer  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  lines, which correspond to the  $n = 3, 4, 5, 6$  to  $n = 2$  transitions respectively.



This initial version of the system represents a first step toward integrating as many divertor diagnostics as possible; other diagnostics which may be included in future are discussed in section 8.

### 3. Bayesian multi-diagnostic analysis

In Bayesian analysis, our knowledge regarding a set of model parameters is expressed as a probability distribution over those parameters, which can be thought of as the distribution of possible ‘causes’ that could have produced the measured data.

In this case the set of model parameters, which we will call  $\theta$ , are the values of  $T_e$ ,  $n_e$  and  $n_0$  at each vertex of the mesh (which are used by the model in (1) to specify the plasma fields) such that

$$\theta = \{T_e^{(1)}, \dots, T_e^{(V)}, n_e^{(1)}, \dots, n_e^{(V)}, n_0^{(1)}, \dots, n_0^{(V)}\}. \quad (2)$$

Our goal is to learn about the distribution of  $\theta$  constrained by the set of diagnostic data, which is commonly referred to as  $\mathcal{D}$ . This distribution is expressed mathematically as the probability of  $\theta$  given  $\mathcal{D}$ , i.e.  $P(\theta|\mathcal{D})$ . This is called the ‘posterior distribution’, and is given by Bayes’ theorem as

$$P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}. \quad (3)$$

Constructing the posterior distribution and learning about its properties is absolutely central to Bayesian analysis, so it is worthwhile to discuss the terms on the right-hand side of (3) individually.

$P(\theta)$  is the *prior* distribution, and represents any information we have regarding the model parameters before we include information from the diagnostic data. For example, this information may be a physics constraint such as non-negativity of the plasma fields. This information could be encoded into the prior distribution by having the prior probability fall to zero if any field values are negative. Typically the prior distribution must be chosen rather than derived—this choice will be discussed in section 4.

$P(\mathcal{D})$  is usually referred to as the *model evidence*, and is important in model selection problems, however we may ignore it in this analysis as the posterior need only be determined up to a constant of proportionality in order for it to be characterised.

$P(\mathcal{D}|\theta)$  is the *likelihood*, and is the probability that we would observe a dataset  $\mathcal{D}$  assuming the plasma were in a state described by a given  $\theta$ . The use of  $\mathcal{D}$  serves as a useful shorthand to represent distributions over many individual data values. For example, suppose that  $d^{(i)}$  represents a single data value from our full dataset—the likelihood is actually the joint distribution over every individual data value given the model parameters, i.e.  $P(d^{(1)}, d^{(2)}, \dots, d^{(n)}|\theta)$ . By letting  $\mathcal{D} = \{d^{(1)}, d^{(2)}, \dots, d^{(n)}\}$  we may write the likelihood more concisely as  $P(\mathcal{D}|\theta)$ .

If some set of random variables, in this case  $\mathcal{D}$ , are mutually conditionally independent (i.e. the uncertainties of

all data values are independent) then the joint distribution of all the variables can be written as the product over the distributions for each variable such that

$$P(d^{(1)}, d^{(2)}, \dots, d^{(n)}|\theta) = \prod_{i=1}^n P(d^{(i)}|\theta). \quad (4)$$

This assumption of independence may not always be valid and depends on the instruments in question, but it is strongly simplifying so should be made where possible.

#### 3.1. Individual diagnostic likelihoods

In multi-diagnostic inference, it is often practical to separate out the overall likelihood for all data into a product of the likelihoods for each diagnostic system. Let the dataset for the Thomson scattering system and Langmuir probes be labelled  $\mathcal{D}_{ts}$  and  $\mathcal{D}_{lp}$  respectively. We will separate out the data for each filtered camera, such that data for the  $i$ ’th camera is represented by  $\mathcal{D}_{fc,i}$ . Again making the assumption of mutual independence between the datasets, the likelihood for all data can be now written as

$$P(\mathcal{D}|\theta) = P(\mathcal{D}_{ts}|\theta)P(\mathcal{D}_{lp}|\theta) \prod_i P(\mathcal{D}_{fc,i}|\theta). \quad (5)$$

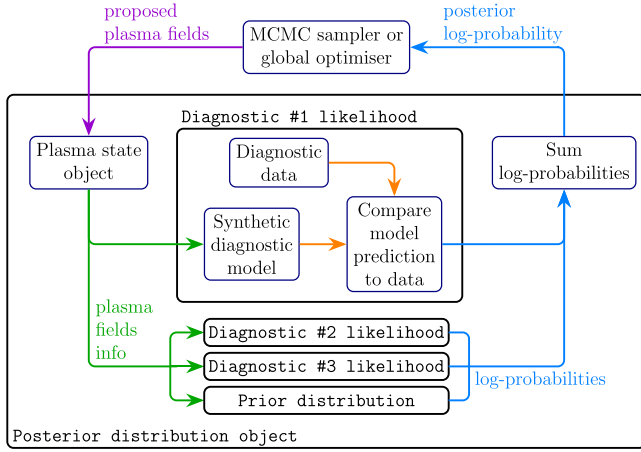
It is common practice to work in log-probabilities, not only for the conceptual simplification that large products of probabilities become sums of log-probabilities, but also for improved numerical stability. Here we use  $\mathcal{L}$  to indicate a log-probability density function, such that  $\mathcal{L}(A|B) = \ln P(A|B)$ . Now combining (3) and (5) we can express the log-posterior distribution  $\mathcal{L}(\theta|\mathcal{D})$  as

$$\begin{aligned} \mathcal{L}(\theta|\mathcal{D}) = & \mathcal{L}(\mathcal{D}_{ts}|\theta) + \mathcal{L}(\mathcal{D}_{lp}|\theta) + \left( \sum_i \mathcal{L}(\mathcal{D}_{fc,i}|\theta) \right) \\ & + \mathcal{L}(\theta) - \mathcal{L}(\mathcal{D}). \end{aligned} \quad (6)$$

All terms in (6) (except  $\mathcal{L}(\mathcal{D})$ , which is in practice discarded) can be evaluated independently. From a programming perspective, this allows each term to be implemented as a separate, self-contained object, encapsulating all experimental data and forward models required to evaluate that term. This approach was used when designing the system, and the resulting structure of the code is illustrated in figure 3. This allows any of the terms to be easily included or excluded from the log-posterior, fulfilling one of the design requirements.

#### 3.2. Thomson scattering and langmuir probe likelihoods

A single Langmuir probe or spatial channel of the divertor Thomson scattering system accumulate their signal over a volume which can be thought of as a spatial instrument function. However, if the extent of this instrument function is small compared to the scale lengths over which the relevant plasma fields vary, we may approximate them to be point measurements. Making this approximation de-couples the analysis of the raw Thomson and Langmuir data from the problem of inferring the fields. For example, the posterior distribution for electron temperature and density for a single



**Figure 3.** Flow chart illustrating the code structure of the system. The posterior distribution is encapsulated as a single object which takes model parameters (which define the plasma fields) as inputs and returns the posterior log-probability. The prior distribution and the likelihood for each diagnostic system are separated into self-contained objects, which can independently request the specific information they require about the plasma fields from the plasma state object. The diagnostic and prior objects each return a log-probability value, which are summed to produce the posterior log-probability. The ‘MCMC sampler or global optimiser’ is responsible for choosing the next set of parameters to pass to the posterior object, with the objective of either sampling from the posterior or finding its maximum.

Thomson channel can be computed in advance and stored, and then referred to when assessing the likelihood of that spatial channel with respect to a set of proposed fields.

This approximation while convenient is not strictly necessary, and in future when the system is applied to real experimental data we may forgo this assumption and forward-model from the proposed fields directly to the raw Thomson scattering and Langmuir probe data. Presently however, we seek only to demonstrate that the multi-diagnostic inference approach has value, so we are free to prescribe a sensible likelihood for the data of a point measurement given  $T_e$  and  $n_e$ . For this purpose, we use an uncorrelated bivariate normal distribution such that

$$\mathcal{L}(\mathcal{D}_{ts}|\theta) = -\frac{1}{2} \sum_i \left[ \left( \frac{\sum_j W_{ij} T_e^{(j)} - \mu_{T_e}^{(i)}}{\sigma_{T_e}^{(i)}} \right)^2 + \left( \frac{\sum_j W_{ij} n_e^{(j)} - \mu_{n_e}^{(i)}}{\sigma_{n_e}^{(i)}} \right)^2 \right]. \quad (7)$$

As in (1),  $T_e^{(j)}$ ,  $n_e^{(j)}$  refer to the model parameters which specify the temperature and density at the  $j$ ’th vertex of the mesh. For the electron temperature and density respectively,  $\mu_{T_e}^{(i)}$ ,  $\mu_{n_e}^{(i)}$  are the measured values and  $\sigma_{T_e}^{(i)}$ ,  $\sigma_{n_e}^{(i)}$  are the uncertainties for the  $i$ ’th spatial measurement point.  $W_{ij}$  is a matrix of pre-calculated linear interpolation weights which give the prediction of the fields at the spatial measurement points. As the Langmuir probes are also treated as being point measurements of  $T_e$  and  $n_e$ , the

total log-likelihood for the Langmuir probes  $\mathcal{L}(\mathcal{D}_{lp}|\theta)$  is also of the form given in (7).

### 3.3. Filtered camera system likelihood

The emissivity at the  $j$ ’th mesh vertex for a given hydrogen spectral line  $\mathcal{E}_j$  is approximated as a sum of excitation and recombination emission such that

$$\mathcal{E}_j = n_e^{(j)} n_0^{(j)} \text{PEC}_{\text{ex}}(T_e^{(j)}, n_e^{(j)}) + (n_e^{(j)})^2 \text{PEC}_{\text{rec}}(T_e^{(j)}, n_e^{(j)}) \quad (8)$$

where  $\text{PEC}_{\text{ex}}$ ,  $\text{PEC}_{\text{rec}}$  are the photon emissivity coefficients for excitation and recombination respectively, whose values are taken from the ADAS database [3]. This model assumes that only atomic emission channels contribute meaningfully to the hydrogenic spectral emission and that  $Z_{\text{eff}} = 1$ . The experimental data are camera images, each of which are analysed as vector of pixel-brightness values  $\mathbf{b}$ . The brightness at the  $i$ th pixel  $b_i$  is modelled as the integral along that pixel’s line-of-sight through the emissivity field defined by the values in the emissivity vector  $\mathcal{E}$ . As the fields are defined through Barycentric interpolation, which is linear, this line-integral can be represented exactly by a weighted sum of the emissivities at each mesh vertex. Given a particular mesh, and a set of lines-of-sight for the pixels, these weights can be pre-calculated and stored as a ‘geometry matrix’  $\mathbf{G}$  such that the product of this matrix with the emissivities  $\mathbf{G}\mathcal{E}$  yields a prediction of the pixel brightness values. We represent the filtered camera likelihood as multivariate normal such that

$$\mathcal{L}(\mathcal{D}_{\text{fc}}|\theta) = -\frac{1}{2} (\mathbf{G}\mathcal{E} - \mathbf{b})^\top \Sigma^{-1} (\mathbf{G}\mathcal{E} - \mathbf{b}). \quad (9)$$

Experimental calibration of filtered camera systems typically finds the variance of the pixel brightnesses (assuming the pixel is not near saturation) to be linear [12] such that

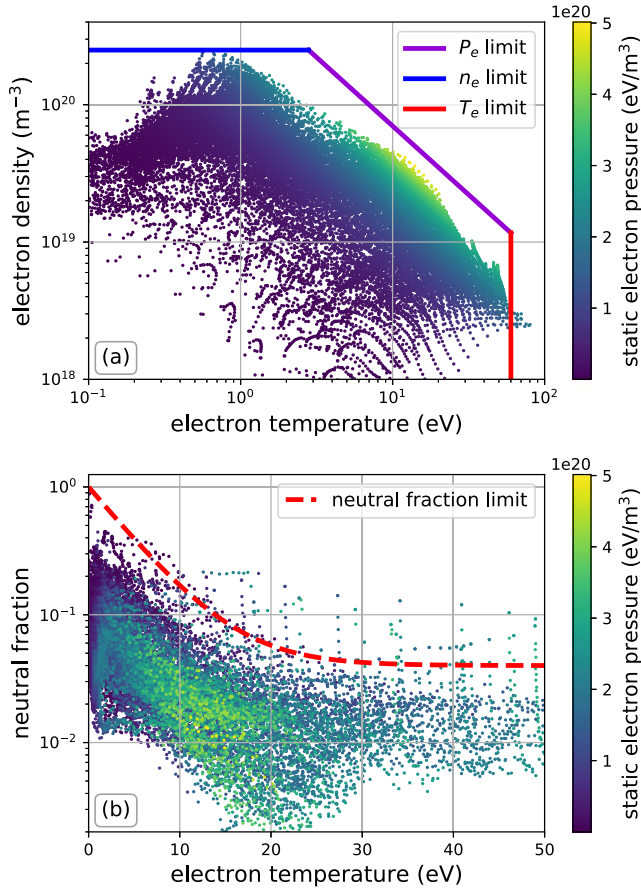
$$\Sigma_{ii} = \alpha b_i + \beta \quad (10)$$

where  $\alpha$ ,  $\beta$  are constants determined as part of the calibration. For our synthetic camera model, we re-parametrise (10) so that the coefficients are more easily interpreted. First suppose that the error at zero brightness can be expressed as some fraction  $f_0$  of the maximum brightness  $b_{\text{max}}$ . Second, we fix the fractional error at the maximum brightness to be a constant  $f_{\text{max}}$ . Under these assumptions the variance may be expressed as

$$\Sigma_{ii} = b_i (f_{\text{max}}^2 - f_0^2) b_{\text{max}} + f_0^2 b_{\text{max}}^2. \quad (11)$$

## 4. Prior constraints

We are always forced to choose a prior distribution—even omitting the prior is equivalent to using a uniform prior (i.e. one which deems all possible sets of  $\theta$  to be equally likely), which is itself a choice. Our goal here is to construct a prior which excludes unrealistic plasma conditions. In order to do this, we require information about the space of realistic plasma



**Figure 4.** The space of realistic plasma conditions in the MAST-U divertor as predicted by a group of SOLPS simulation results spanning various experimental configurations. (a) The space of realistic  $(T_e, n_e)$  can be well approximated by placing upper limits on  $T_e$ ,  $n_e$  and  $P_e$ . (b) The space of realistic  $n_0/(n_e + n_0)$  can be bounded by placing an upper limit on  $n_0/(n_e + n_0)$  as a function of  $T_e$ .

conditions that exist within the divertor. To gain insight into this, we examined a collection of 25 MAST-U SOLPS simulations which were carried out in support of a study on enhancements to the plasma exhaust operational space of MAST-U [13]. These simulations cover a range of plasma densities at the core grid boundary ( $3.6 \times 10^{18} \rightarrow 1.5 \times 10^{20} \text{ m}^{-3}$ ), and heating powers ( $1.7 \rightarrow 2.5 \text{ MW}$ ). In each case, the fields of  $T_e$ ,  $n_e$  and  $n_0$  in the lower divertor were extracted, and the values of each field in the lower divertor were gathered into a single dataset. By plotting the gathered field values against one another we are able to derive simple but useful constraints on plasma conditions which dictate whether a given triple of  $(T_e, n_e, n_0)$  is considered realistic—these results are summarised in figure 4.

Our chosen prior is made up of three components: a constraint on the static electron pressure, a constraint on the neutral fraction and a constraint on the spatial ‘smoothness’ of the plasma fields. This can be expressed mathematically by writing the log-prior  $\mathcal{L}(\theta)$  as a sum of three terms, one for each constraint such that

$$\mathcal{L}(\theta) = \mathcal{L}(\theta)_{\text{pressure}} + \mathcal{L}(\theta)_{\text{fraction}} + \mathcal{L}(\theta)_{\text{smoothness}}. \quad (12)$$

We will now discuss each of these terms individually.

#### 4.1. Static electron pressure prior

The prior on the static electron pressure for each vertex is uniform if the pressure is less than the chosen limit  $P_e^{\text{max}}$ , and Gaussian for values above the limit. The resulting static electron pressure log-prior is

$$\mathcal{L}(\theta)_{\text{pressure}} = -\frac{1}{2\sigma_{\text{prs}}^2} \sum_i \max\left(\frac{n_e^{(i)} T_e^{(i)}}{P_e^{\text{max}}} - 1, 0\right)^2. \quad (13)$$

The value of  $\sigma_{\text{prs}}$  can be thought of as a ‘fractional tolerance’ of the limit  $P_e^{\text{max}}$ , i.e. by what fraction the limit may be violated before the prior probability drops significantly. Based on the SOLPS data we set  $P_e^{\text{max}} = 2 \times 10^{20} \text{ eV m}^{-3}$  and  $\sigma_{\text{prs}} = 0.1$ .

#### 4.2. Neutral fraction prior

The upper limit on the neutral fraction at each vertex  $f_i^{\text{max}}$  is set as a function of the temperature at each vertex such that

$$f_i^{\text{max}} = (1 - c) \exp(-T_e^{(i)}/l) + c, \quad (14)$$

where  $c = 0.04$  and  $l = 5 \text{ eV}$ . In figure 4 this limit is shown to be greater than 99.5% of neutral fractions in the SOLPS dataset. The neutral fraction prior has the same form as that used for the static electron pressure in (13) such that

$$\mathcal{L}(\theta)_{\text{fraction}} = -\frac{1}{2\sigma_{\text{frc}}^2} \sum_i \max\left(\frac{1}{f_i^{\text{max}}} \frac{n_0^{(i)}}{n_e^{(i)} + n_0^{(i)}} - 1, 0\right)^2, \quad (15)$$

where  $\sigma_{\text{frc}} = 0.1$ .

#### 4.3. Spatial smoothness prior

We want our prior to favour solutions where the plasma fields are spatially smooth. This requires us to choose a metric for the overall ‘roughness’ of the fields, so that solutions which are too rough can be penalised by assigning them a lower prior probability. Suppose  $\mathbf{v}$  is a vector of field values at each mesh vertex, and define the ‘umbrella’ matrix operator  $\mathbf{U}$  such that

$$U_{ij} = \begin{cases} -1 & \text{if } i = j \\ 1/n & \text{if vertex } j \text{ is one of } n \text{ vertices connected to } i. \\ 0 & \text{if vertex } j \text{ is not connected to } i \end{cases} \quad (16)$$

The product  $\mathbf{U}\mathbf{v}$  is then a vector of differences between the field value at each vertex and the average field value of all vertices to which it is connected. For a purely equilateral mesh, if the value of a vertex and all its neighbours lie in a plane, then this difference will be exactly zero. In this sense the umbrella operator measures how much the field deviates from a plane in the local region of each vertex. We therefore take the sum of the squares of the umbrella differences,  $|\mathbf{U}\mathbf{v}|^2$ , as our metric for the total ‘roughness’ of a field.

It is helpful to consider whether the fields can be transformed such that the expected solutions for the transformed

**Table 1.** Upper and lower bounds placed on the values of the plasma fields at all mesh vertices.

	$T_e$	$n_e$	$n_0$
Lower bound	0.2 eV	$1 \times 10^{16} \text{ m}^{-3}$	$1 \times 10^{15} \text{ m}^{-3}$
Upper bound	60 eV	$2.5 \times 10^{20} \text{ m}^{-3}$	$2 \times 10^{20} \text{ m}^{-3}$

fields better satisfy the assumption of smoothness. Enforcing smoothness on these transformed fields means that real features of the fields, which we want to preserve, are less likely to be penalised by the smoothing prior.

For this reason we enforce spatial smoothness on the natural log of the plasma fields, rather than the fields themselves. Let  $\tilde{T}_e$ ,  $\tilde{n}_e$ ,  $\tilde{n}_0$  represent the vectors of log-temperature, log-density and log-neutral density at each vertex of the mesh, and define  $\mathbf{S} = \mathbf{U}^\top \mathbf{U}$ . The roughness of one of the log-fields, for example the log-temperature, can now be written as  $|\mathbf{U}\tilde{T}_e|^2 = \tilde{T}_e^\top \mathbf{S} \tilde{T}_e$ . By introducing a constant  $\sigma_{\text{smth}}$ , which determines how strongly overly-rough fields are penalised, we can define the smoothing log-prior as

$$\mathcal{L}(\theta)_{\text{smoothness}} = -\frac{1}{2\sigma_{\text{smth}}^2}(\tilde{T}_e^\top \mathbf{S} \tilde{T}_e + \tilde{n}_e^\top \mathbf{S} \tilde{n}_e + \tilde{n}_0^\top \mathbf{S} \tilde{n}_0). \quad (17)$$

Unlike the priors on the static electron pressure and neutral fraction, which effectively set upper limits on those quantities, the smoothness prior has a strong impact on the entire posterior distribution. Consequently, additional work is required to select an appropriate value for  $\sigma_{\text{smth}}$ —this is discussed further in section 7.

#### 4.4. Bounds on field values

Upper and lower bounds are placed on the electron temperature, density and neutral density at every vertex. These bounds, chosen based on the SOLPS data, are given in table 1. The bounds could be imposed by including an additional term in the definition of the log-prior in (12), but in practice it is easier to allow the bounds to be enforced by the optimisation or sampling algorithm which is being used to characterise the posterior distribution.

## 5. Production of synthetic test-case data

### 5.1. SOLPS test-cases

In order to test the system we require synthetic data for each instrument, and that this data is as representative as possible of the real experimental data which will be measured during MAST-U operation. For this purpose we use results from SOLPS simulations of the MAST-U edge and divertor to prescribe the fields of electron temperature, density and neutral density from which the synthetic data will be derived.

Here we consider two SOLPS cases taken from a scan of the nitrogen seeding rate to detachment. Both cases have the

same magnetic equilibrium, 2.5 MW of heating power and a deuterium fuelling rate of  $2 \times 10^{21} \text{ s}^{-1}$ . The two cases, which we will from now refer to as the low- and high-seeding cases, have nitrogen seeding rates into the divertor of  $2 \times 10^{20} \text{ s}^{-1}$  and  $5 \times 10^{20} \text{ s}^{-1}$  respectively. These two cases are not part of the set used to inform the prior constraints discussed in section 4, however their field values lie well inside the limits set by the chosen prior.

Note that although in the SOLPS data itself the electron density and hydrogen ion density fields maybe be different due to the presence of the seeded nitrogen, we set them to be equal when producing synthetic data, as this equality is assumed in the emission model in (8).

The field values on the SOLPS grid are interpolated on to the triangular mesh prior to producing the synthetic data, such that the resulting mesh representation of the fields becomes a test-case which we will attempt to reconstruct. The mesh-representations of the plasma fields for each of the two test-cases are shown in figure 5.

### 5.2. Addition of simulated noise to synthetic data

After synthetic measurements for each instrument are generated using their respective forward-models, simulated noise is added to the data. For the filtered camera images, the variance of the noise added to each pixel is set according to (11), where  $f_{\text{max}} = 0.025$  and  $f_0 = 1/256$ .

The point measurements of electron temperature taken by the Thomson scattering and Langmuir probes systems have an assigned uncertainty of  $\sigma_{T_e} = T_e/10 + 0.1 \text{ eV}$ , and the corresponding electron density measurements have an assigned uncertainty of  $\sigma_{n_e} = n_e/10 + 10^{18} \text{ m}^{-3}$ .

## 6. Characterising the posterior distribution

Now that all terms in (6) which have a dependence on  $\theta$  have been defined, the posterior log-probability can be evaluated for any chosen set of plasma fields. The posterior must now be characterised in a way which allows us to extract useful information about the plasma fields.

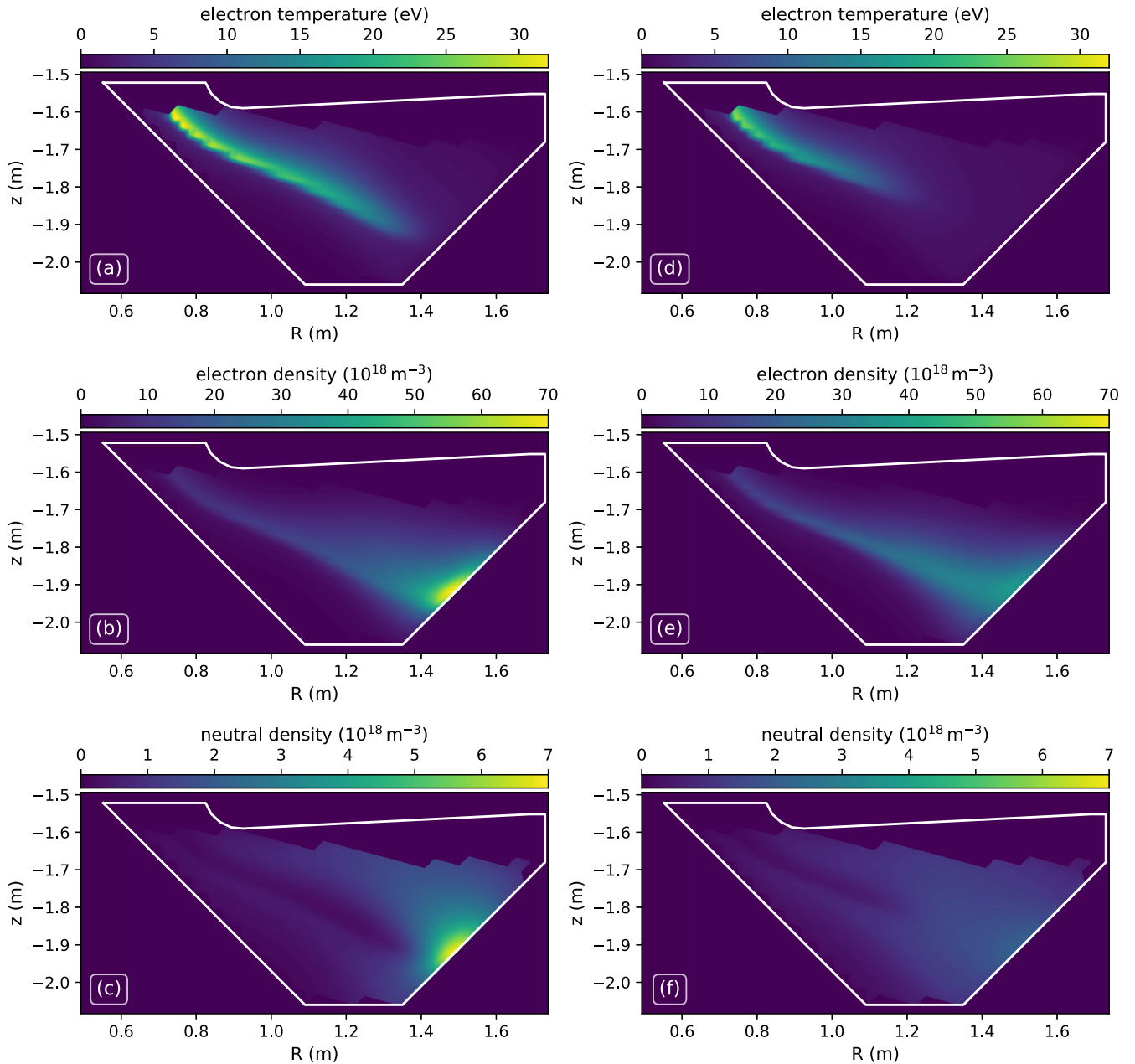
### 6.1. Maximum a posteriori estimation

The first stage of characterising the posterior is to find the set of model parameters which maximises its value, referred to as ‘*maximum a posteriori*’ (MAP) estimation [14]. To locate this maximum, we employ a ‘hybrid’ approach which combines both stochastic and gradient-based optimisation—a more detailed description of this approach is given in the [appendix](#).

### 6.2. Hamiltonian Monte-Carlo sampling

Although MAP estimation yields a useful single-value estimate of the model parameters, it does not provide any information regarding the uncertainties associated with that estimate. To characterise these uncertainties we employ ‘Hamiltonian Monte-Carlo’ (HMC), a gradient-based sampling algorithm which is





**Figure 5.** Plots showing the mesh-representations of the electron temperature, electron density and neutral density taken from the nitrogen-seeded SOLPS predictions. Plots (a)–(c) and (d)–(f) show the low- and high-seeding cases respectively.

particularly effective (in comparison to other approaches) in cases where the number of model parameters is large, and strong correlations are present [15].

### 6.3. Marginal expectation

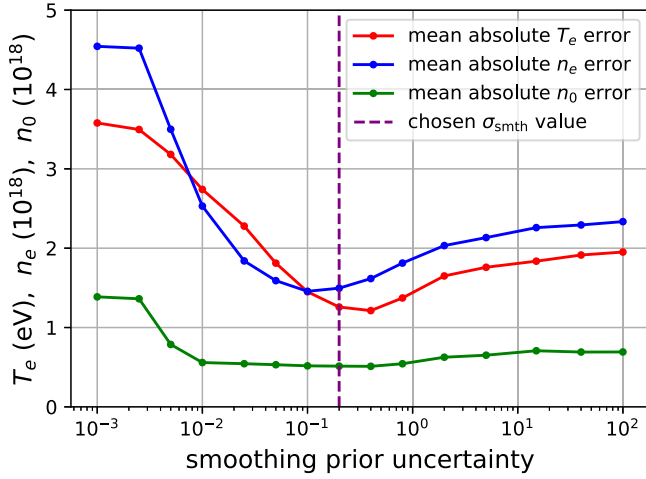
An alternative to the MAP estimate is the marginal expectation (MEX), where the value of each parameter is taken to be the mean of its corresponding marginal distribution. Whereas the MAP is the single most probable solution, the marginal expectation can be thought of as the ‘average’ of the possible solutions. The marginal expectation estimate can be calculated easily by taking the mean of the sample generated using Hamiltonian Monte-Carlo.

## 7. Results of synthetic testing

### 7.1. Selecting the smoothing prior uncertainty

The value of the smoothing prior uncertainty  $\sigma_{\text{smth}}$ , which appears in (17), can have a strong impact on the posterior distribution and therefore the MAP estimate. To assess this impact we evaluated the mean absolute difference between the MAP estimate and the low-seeding case at all vertices for a range of values of  $\sigma_{\text{smth}}$ —the results of this scan are shown in figure 6. The minima in the error for each field are fairly broad, but do not all occur at the same value of  $\sigma_{\text{smth}}$ . The results presented here used a value of  $\sigma_{\text{smth}} = 0.2$  which provides a good balance between low error in electron temperature and electron density.





**Figure 6.** Mean absolute difference between the field values of the low-seeding case and those of the corresponding MAP estimate for each of the plasma fields as a function of the smoothing prior uncertainty  $\sigma_{\text{smth}}$ . The chosen value of  $\sigma_{\text{smth}} = 0.2$  is indicated by the dashed vertical line.

In an applied case using experimental data we cannot select the smoothing uncertainty in this fashion as the true values of the fields are unknown. As such, testing selection criteria for the smoothing uncertainty which are applicable to experimental data will be the subject of further work.

## 7.2. Comparison of inferred fields and test-cases

The maximum a posteriori (MAP) and marginal expectation (MEX) estimates were evaluated as described in section 6 for both test-cases. The inferred field values from the MEX estimate are compared with those from the corresponding test-case in figure 7. The mean-absolute-difference between the inferred and test-case fields is used to quantify the accuracy of the estimates. The mean-absolute-difference for the electron temperature is given by

$$\langle |\Delta T_e| \rangle = \frac{1}{V} \sum_{k=1}^V |T_{e,\text{inf}}^{(k)} - T_{e,\text{test}}^{(k)}| \quad (18)$$

where  $T_{e,\text{inf}}^{(k)}$  and  $T_{e,\text{test}}^{(k)}$  are the inferred and test-case electron temperature at vertex  $k$  respectively. The mean-absolute-differences for each test-case are given in table 2 for both the MAP and MEX estimates. The marginal expectation appears to outperform the MAP estimate for these cases, but with the exception of the high-seeding neutral density estimation, the differences in the mean absolute error values are less than 10%.

We note that the estimate of the electron temperature becomes less reliable above  $\sim 10$  eV—this may occur because at these higher temperatures the emission is almost purely due to excitation, which is very insensitive to electron temperature above 10 eV for the Balmer series.

Conversely, in regions where the temperature is very low, the emission becomes dominated by recombination, which has no dependence on the neutral density. We suspect this is the cause of the large errors in the neutral density estimation for the high-seeding test-case, which is more

strongly detached than the low-seeding case, and therefore has a large region of recombination-dominated emission.

The absolute fractional error  $|T_{e,\text{inf}}^{(k)} / T_{e,\text{test}}^{(k)} - 1|$  is another useful metric for gauging the overall accuracy of the inferred fields. Averaged over both the low and high-seeding test-cases, the median absolute fractional errors in the electron temperature and density estimates were 10.3% and 10.1% respectively. We use the median rather than the mean in this case as it is more robust against large outliers that can occur when field values get very small.

The inferred fields for the low-seeding case, along with the differences between the inferred fields and the test-case are shown in figure 8. These difference plots highlight spatial structure in the estimation errors, such as the under-estimation of the temperature along the separatrix. The peak in the electron temperature at the separatrix is a very sharp feature which will be penalised by the spatial smoothing prior. This, combined with the relatively weak temperature dependence of the emission in that region, is likely the reason for the under-estimate of the separatrix temperature. This highlights a common difficulty of regularising solutions which possess a wide range of spatial scale-lengths—any level of smoothing which suppresses non-physical fluctuations in regions with a long scale-length will also over-smooth in regions with short-scale lengths.

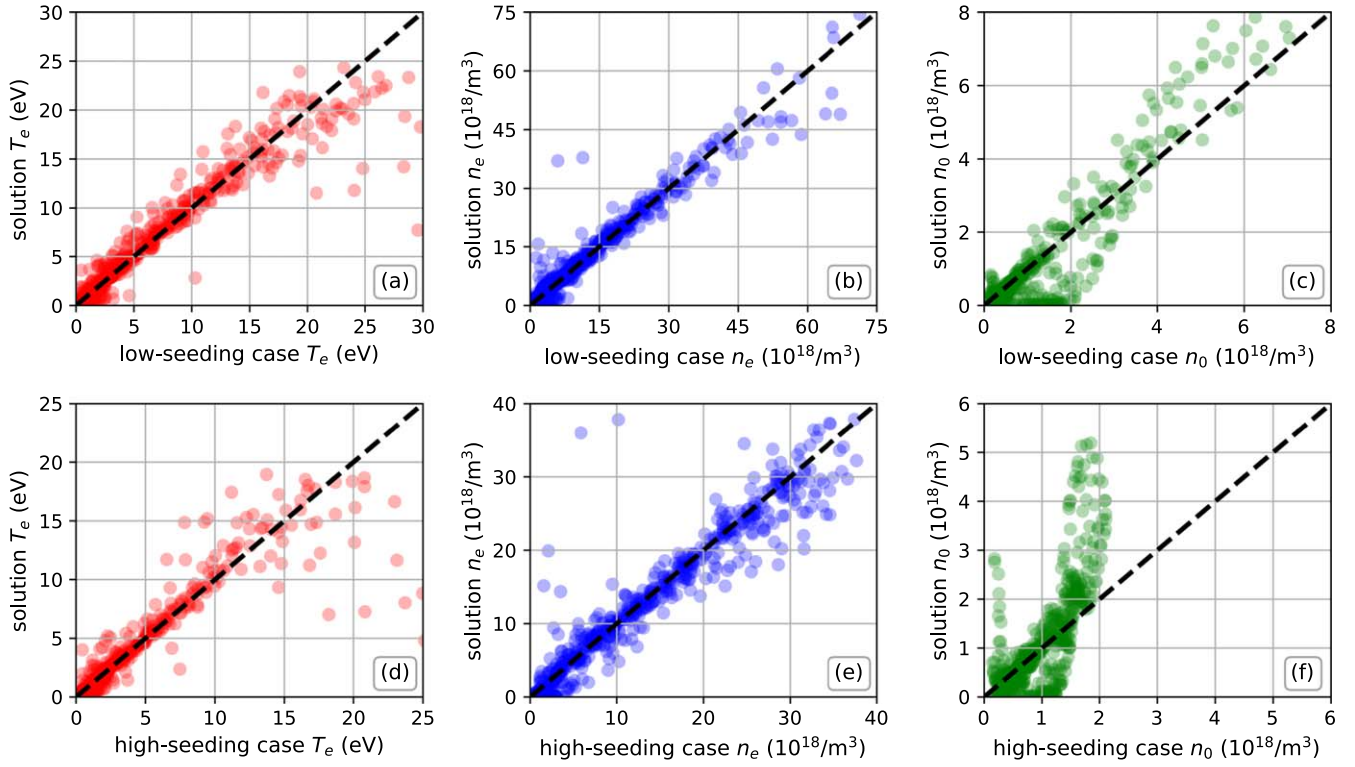
Tests were also carried out wherein only measurements from a single filtered camera were used to constrain the plasma fields in order to verify the effects of a multi-diagnostic approach. The MAP estimates of the fields obtained in these tests were completely erroneous, and posterior sampling predicted very large uncertainties in the solution. Although imaging data for a single emission line does provide some information about  $T_e$ ,  $n_e$  and  $n_0$ , it is not a strong enough constraint to estimate them with useful accuracy. This is because the set of plasma fields which reproduces the measured image to within experimental error is too large and too varied. Inclusion of additional filtered images of different Balmer lines improves the estimate significantly because the size of this set of potential solutions is greatly reduced, as now a valid solution must reproduce *all* of the images simultaneously rather than just one.

## 7.3. Uncertainty estimation

Uncertainties in the inferred fields for both test-cases were estimated by sampling from the posterior distribution using Hamiltonian Monte-Carlo.

Figure 9 shows a comparison of the test-cases and inferred fields along a scrape-off layer flux-surface (shown in figure 2), and shows the 95% highest-density interval derived from the sample. We see that the differences between the test-case values and the inferred fields are well explained by the estimated uncertainties almost everywhere. One notable exception is that the uncertainty in the electron temperature and density appears to be under-estimated close to the target.

For an inverse problem of this type the posterior is typically highly multi-modal. It is possible that the Markov-chains used to generate the sample were trapped near the



**Figure 7.** Scatter plots of the field values of the test-cases versus those of the MEX estimate at each vertex of the mesh. Plots (a)–(c) and (d)–(f) show results from the low- and high-seeding cases respectively.

**Table 2.** The mean-absolute-difference as defined in (18) for  $T_e$ ,  $n_e$  and  $n_0$  in each test-case, and for both the MAP and MEX estimates.

	$\langle  \Delta T_e  \rangle$ (eV)	$\langle  \Delta n_e  \rangle$ ( $\text{m}^{-3}$ )	$\langle  \Delta n_0  \rangle$ ( $\text{m}^{-3}$ )
Low-seeding MAP	1.12	$1.57 \times 10^{18}$	$3.50 \times 10^{17}$
Low-seeding MEX	1.09	$1.46 \times 10^{18}$	$3.38 \times 10^{17}$
High-seeding MAP	0.81	$1.60 \times 10^{18}$	$7.37 \times 10^{17}$
High-seeding MEX	0.74	$1.65 \times 10^{18}$	$4.05 \times 10^{17}$

maxima corresponding to the MAP estimate, and were unable to explore other maxima which may feature more varied configurations of the fields near the target. To investigate this we plan to test extensions to standard Markov-chain Monte-Carlo which are designed specifically to allow exploration of multi-modal distributions such as parallel tempering [16].

#### 7.4. Inference of physical processes

The long-term goal of developing this analysis is to help advance our understanding of divertor physics by providing direct information about the 2D divertor plasma state. We are therefore interested not only in plasma fields like  $T_e$  and  $n_e$ , but also the behaviour of physical processes like ionisation and recombination which are partially responsible for determining the plasma state.

If these processes can be modelled using only the inferred plasma fields, then their values and uncertainties can also be inferred from the posterior sample. An example of this is shown in figure 10, where the deuterium ionisation rate along

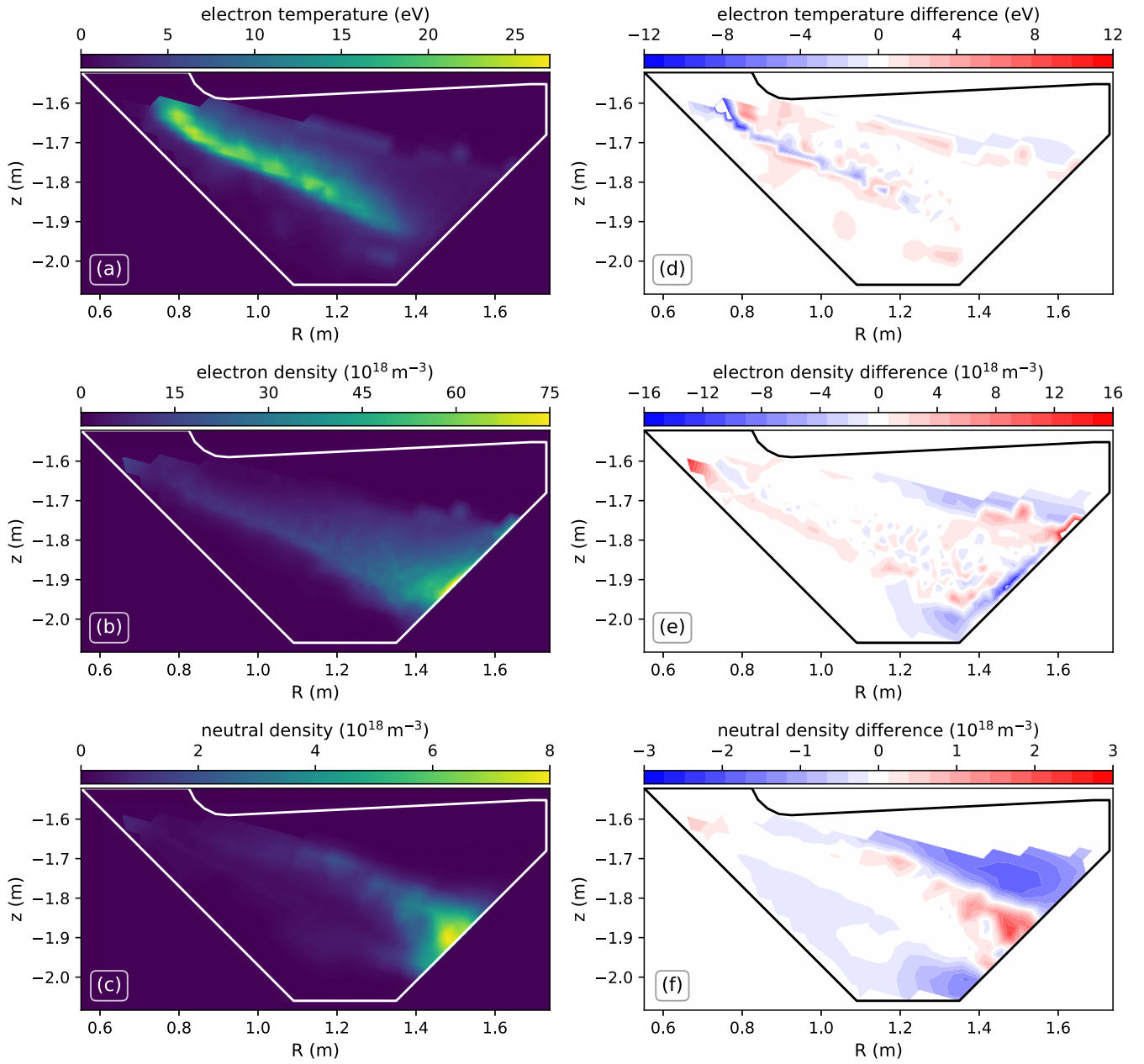
a scrape-off layer flux-surface was calculated from the posterior sample, and is compared with the corresponding ionisation calculated from the test-cases.

## 8. Discussion

### 8.1. Potential improvements to instrument modelling

All synthetic diagnostic models are ‘idealised’ to some extent, as they cannot reasonably capture every subtlety of the experimental set-up perfectly. Our goal however should be to make these models more realistic where possible, and this will be the focus of further work on the system before it is applied to real experimental data.

For example, uncertainty in the absolute brightness calibration of filtered cameras is a potentially important effect for which we do not currently account. This can be achieved by including the calibrations as so-called ‘nuisance parameters’. This process involves allowing the calibration values



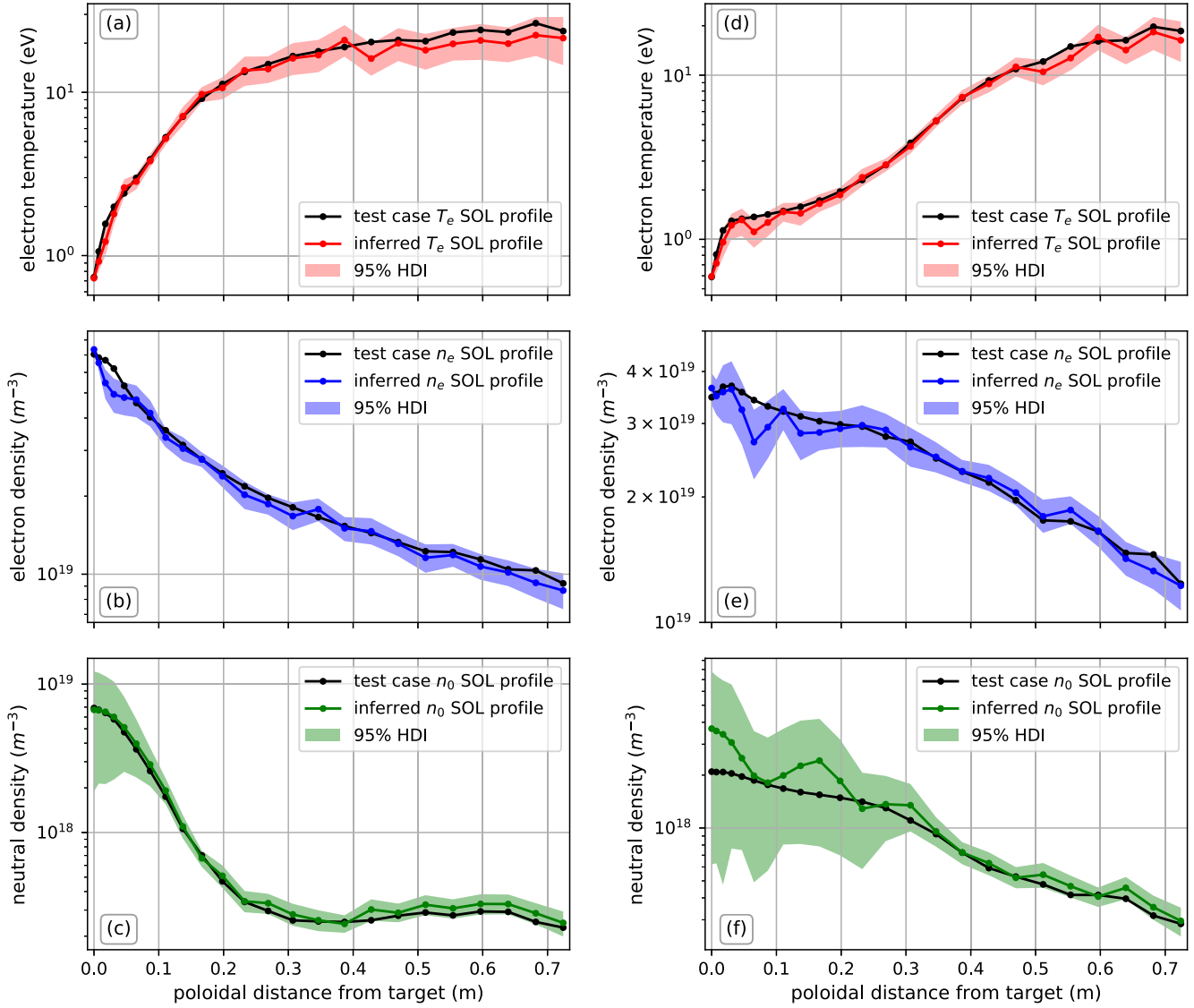
**Figure 8.** The marginal expectation estimate of the electron temperature, electron density and neutral density fields for the low-seeding case are shown in (a)–(c) respectively. The corresponding differences between the marginal expectation estimate of a field and the test-case values are shown in (d)–(f).

themselves to be free parameters in the system, with a prior distribution determined by the measurement of those calibration values. By allowing the calibration values to vary in this way, effects of the uncertainty in their value are reflected in the inferred plasma fields.

Some of the light collected by the MWI system will have reached the camera after being reflected by a material surface in the divertor. The algorithm used here to calculate the Geometry matrix  $\mathbf{G}$ , which appears in (9), only accounts for light which has travelled directly from the plasma to the camera. It is however possible to account for reflections from material surfaces by using a more sophisticated approach to calculating the geometry matrix [17], with an associated increase in the computational cost of the filtered camera forward-model.

It may be the case that in practice, the Langmuir probes are unable to measure the electron temperature with an uncertainty comparable to that which we assume when generating synthetic data when the temperature drops below 5 eV. In such cases, the probes may only provide an ‘upper limit’ measurement on the temperature. Accounting for this will require forward-modelling to produce synthetic probe data, which can be analysed to calculate joint-distributions of  $T_e$ ,  $n_e$  to be used in place of the assumed Gaussian errors.

In this work the SOLPS test-cases were treated as if no impurities were present. However, in real experiments with strong impurity seeding, the presence of impurities can affect measured diagnostic signals. Consequently, modifications to the diagnostic forward-models may be required to properly



**Figure 9.** Comparison of profiles along a scrape-off layer flux-surface (shown in figure 2) derived from the inference results and the test-cases. Plots (a)–(c) and (d)–(f) show results from the low- and high-seeding cases respectively. The coloured areas indicate the 95% highest-density intervals (HDI) derived from the sampling results.

account for these effects. For example, the emission model in (8) assumes that the hydrogen ion and electron densities are equal, which may no longer be reasonable with strong seeding. Although the effect is expected to be small, impurities can also impact Langmuir probe measurements [18].

### 8.2. Inclusion of additional diagnostic systems

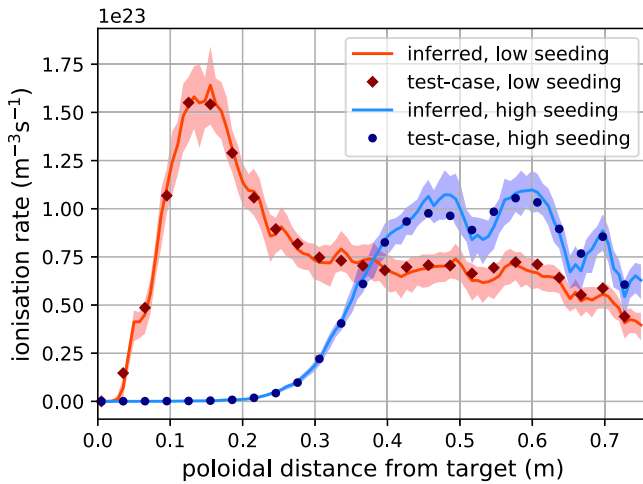
The MAST-U divertor spectroscopy system could be a useful additional source of information for inferring the plasma fields. The system will observe a large number of spectral lines, including many from various impurities, so modelling all data produced by the spectrometers is not feasible. However, if we restrict the analysis to spectral lines which are already being viewed by the MWI system, then this data can be modelled without greatly increasing the number of model parameters. The brightness of these lines as measured by the spectrometers would

provide a cross-check on the brightnesses measured by the MWI system, and it may also be possible to constrain the electron density along the spectrometer line-of-sight using information encoded in the spectral line-shape due to Stark-broadening [19].

### 8.3. Choice of imaged spectral lines

The emissivity model in (8) does not account for emission resulting from the production of excited deuterium atoms due to plasma-molecule interactions. This emission may be a non-negligible component of low- $n$  Balmer series emissivities, particularly for deuterium Balmer- $\alpha$ , in strongly detached conditions [19, 20]. Deuterium- $\alpha$  through  $\delta$  were chosen as a starting point from which to develop and test the system, but there are many possible choices of atomic lines, including higher- $n$  Balmer lines and impurity emission lines.





**Figure 10.** The deuterium ionisation rate, calculated using ADAS, along a scrape-off layer flux surface (shown in figure 2). The points show ionisation calculated from the test-cases, and the lines show the inferred ionisation (calculated from the inferred plasma fields). Uncertainties on the inferred ionisation are given by the coloured areas, which show the 95% highest-density interval.

Determining the optimal group of lines for inferring the plasma fields is complex—one needs to consider not only the information content of the lines with respect to the plasma fields, but also how well those lines can be measured (considering their brightness, wavelength and contamination from neighbouring spectral lines), how accurately their emissivity can be modelled, and the total number of plasma fields required in order to model them. Testing alternative groups of atomic lines which best meet these criteria will be an important part of the ongoing development of the system.

#### 8.4. Imposing physics constraints using equilibrium information

The spatial structure of the  $T_e$  and  $n_e$  fields in the divertor is closely tied to that of the magnetic field and the resulting flux-surfaces. If the mesh used to parametrise the plasma fields were constructed such that every vertex of the mesh lay on one of a chosen set of flux-surfaces, this would allow additional physics constraints to be imposed.

For example, we could include an additional term in the prior distribution which requires that the electron temperature decrease monotonically along each flux surface when approaching the strike-point. It would also allow for a more powerful constraint on spatial smoothness of the fields, as we could require that the fields vary much more smoothly along flux-surfaces than perpendicular to them.

There are some disadvantages to this approach however. The equilibrium reconstruction required to build such a grid is itself an inverse problem whose solution is uncertain. Given equilibrium reconstructions which include uncertainty estimates (which are typically unavailable), propagating this uncertainty through the system so that it is reflected in the uncertainties on the plasma fields would be difficult, as the grid remains fixed throughout the analysis.

Imposing physics constraints in this way, and the effect this has on the accuracy of the inferred 2D plasma fields will be explored as part of further work.

#### 8.5. Toroidal asymmetry effects

The major radius of tile 5 in the MAST-U divertor (visible in figure 2) varies as a function of toroidal angle, with a periodicity that matches the toroidal-field ripple of the device in order to distribute power more evenly across the tile surface.

Cases in which there is strong emission close to the surface of tile 5 may introduce 3D effects into filtered camera measurements which cannot be reproduced by the 2D model we use here for the plasma fields. The significance of this effect could be explored theoretically using 3D simulations of the MAST-U divertor which account for the toroidal asymmetry of tile 5.

## 9. Summary and conclusions

We have presented details of the first design, implementation and testing of a Bayesian multi-diagnostic inference system, which can infer the 2D fields of electron temperature, density and neutral density over the divertor cross-section. The system has been designed to be modular and flexible, so that the diagnostics utilised by the system, and the underlying fields that are to be obtained by the analysis, can be changed easily.

For this initial test of the system we restricted the inferred plasma fields to only electron temperature, electron density and deuterium neutral density. These fields were inferred using simulated experimental data, which included appropriate added noise, derived from two SOLPS-ITER simulations taken from a scan of the nitrogen seeding rate. The synthetic diagnostic models used to generate the simulated data included four filtered cameras as part of the MWI system viewing the first four Balmer lines, as well as divertor Thomson-scattering system and target Langmuir probes.

These system tests have demonstrated that for the given synthetic data, the 2D plasma fields can be inferred with enough accuracy to give powerful insight into the physics of plasma behaviour in the divertor. It was also demonstrated that uncertainties in the inferred plasma fields can be reliably estimated using Hamiltonian Monte-Carlo sampling, which would allow conclusions to be drawn from the results with greater confidence.

This first effort at Integrated data analysis for the divertor has thus been successful in demonstrating that the use of a Bayesian, multi-diagnostic approach to infer the plasma solution merits further investigation. Future work will focus on the inclusion of additional diagnostic systems, and the application of this analysis to real diagnostic data from tokamak experiments.

## Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014–2018 and

2019–20 under grant agreement no 633 053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. This work has also received funding from the EPSRC under the grant EP/N023846/1 and the research by B Lipschultz was funded in part by the Wolfson Foundation and UK Royal Society through a Royal Society Wolfson Research Merit Award as well as by the RCUK Energy Programme (EPSRC grant number EP/I501045).

## Appendix. Optimisation methodology

The line-integrated nature of filtered camera measurements, and the nonlinear relationship between line emissivities and the plasma fields introduces strong nonlinear correlations between the model parameters. The presence of correlations and a large number of free parameters (around 1800 in this case) usually necessitates the use of gradient-based optimisation and sampling techniques.

Such techniques require the derivative of the log-posterior with respect to the model parameters. However approximating this derivative via finite-difference is prohibitively expensive as the number of model parameters is large. The system was therefore designed such that the gradient of the log-posterior can be calculated analytically. As a result, evaluating the gradient takes around 3 times longer than evaluating the log-posterior itself—this is approximately 600 times faster than evaluating the gradient using finite-difference for the current number of model parameters.

However, we also found the posterior distribution to be highly multi-modal, which causes issues for gradient-based optimisation algorithms which tend to converge to local rather than global maxima.

To address this we employ a ‘hybrid’ approach which combines a genetic algorithm with the L-BFGS algorithm [21]. In this approach a set of candidate solutions is created (initially by random sampling), and then each candidate is used as a starting-guess for the L-BFGS algorithm, which converges to a (typically local) maximum in the posterior log-probability density. Based on the resulting set of local maxima, a new set of candidate solutions is generated using the genetic algorithm, and this process is repeated until the highest observed log-probability converges.

Evaluating the L-BFGS algorithm for each candidate solution is an independent computation, allowing them to be efficiently distributed across multiple CPUs. The results presented here used a population of 20 candidate solutions distributed over 20 threads of a Intel Xeon E5-2695 v3. The maximum log-probability had converged sufficiently after 80 generations, taking around 80 min in total.

## ORCID iDs

C Bowman  <https://orcid.org/0000-0002-0813-2206>  
 J R Harrison  <https://orcid.org/0000-0003-2906-5097>  
 B Lipschultz  <https://orcid.org/0000-0001-5968-3684>  
 K Verhaegh  <https://orcid.org/0000-0002-0500-2764>

## References

- [1] Stangeby P 2000 The Plasma Boundary of Magnetic Fusion Devices
- [2] Wiesen S *et al* 2015 The new solps-iter code package *J. Nucl. Mater.* **463** 480–4
- [3] Summers H P, Dickson W J, O’Mullane M G, Badnell N R, Whiteford A D, Brooks D H, Lang J, Loch S D and Griffin D C 2006 Ionization state, excited populations and emission of impurities in dynamic finite density plasmas: I. The generalized collisional-radiative model for light elements *Plasma Phys. Control. Fusion* **48** 263–93
- [4] Morris W, Harrison J R, Kirk A, Lipschultz B, Militello F, Moulton D and Walkden N R 2018 MAST upgrade divertor facility: a test bed for novel divertor solutions *IEEE Trans. Plasma Sci.* **46** 1217–26
- [5] Vijvers W, Mumgaard R, Andrebe Y, Classen I, Duval B and Lipschultz B 2017 *J. Instrum.* **12** C07006
- [6] Fischer R, Fuchs C J, Kurzan B, Suttrop W and Wolfrum E 2010 Integrated data analysis of profile diagnostics at ASDEX upgrade *Fusion Sci. Technol.* **58** 675–84
- [7] Ford O P, Svensson J, Beurskens M, Boboc A, Flanagan J, Kempenaars M and McDonald D C 2009 Bayesian combined analysis of jet lidar, edge lidar and interferometry diagnostics *36th EPS Conf. on Plasma Physics* vol 33, pp 1–10
- [8] Svensson J and Werner A 2008 Current tomography for axisymmetric plasmas *Plasma Phys. Control. Fusion* **50** 085002
- [9] Von Nessi G T, Hole M J, Svensson J and Appel L 2012 Evidence cross-validation and Bayesian inference of MAST plasma equilibria *Phys. Plasmas* **19** 012506
- [10] Hawke J, Scannell R, Harrison J, Huxford R and Bohm P 2013 Outline of optical design and viewing geometry for divertor Thomson scattering on MAST upgrade *J. Instrum.* **8** C11010
- [11] F  vrier O, Theiler C, De Oliveira H, Labit B, Fedorczak N and Baillod A 2018 Analysis of wall-embedded langmuir probe signals in different conditions on the tokamak configuration variable *Rev. Sci. Instrum.* **89** 053502
- [12] Silburn S A 2014 A doppler coherence imaging diagnostic for the mega-amp spherical tokamak *PhD Thesis* Durham University
- [13] Harrison J, Moulton D, Carr M, Chman I, Keeling D, Kirk A and Walkden N 2016 Enhancements to mast upgrade to address the eurofusion plasma exhaust strategy *43rd EPS Conf. on Plasma Physics*
- [14] Von Toussaint U 2011 Bayesian inference in physics *Rev. Mod. Phys.* **83** 943–99
- [15] Neal R M 2012 MCMC using hamiltonian dynamics *arXiv:1206.1901*
- [16] Sambridge M 2013 A parallel tempering algorithm for probabilistic sampling and multimodal optimization *Geophys. J. Int.* **196** 357–74
- [17] Carr M *et al* 2019 Physically principled reflection models applied to filtered camera imaging inversions in metal walled fusion machines *Rev. Sci. Instrum.* **90** 043504
- [18] Stangeby P C 1987 Interpretation of langmuir, heat-flux, deposition, trapping and gridded energy analyser probe data for impure plasmas *J. Phys. D: Appl. Phys.* **20** 1472–8
- [19] Verhaegh K *et al* 2019 An improved understanding of the roles of atomic processes and power balance in divertor target ion current loss during detachment *Nucl. Fusion* **59** 126038
- [20] Wunderlich D and Fantz U 2016 Evaluation of state-resolved reaction probabilities and their application in population models for he, h, and h2 *Atoms* **4** 26
- [21] Nocedal J and Liu D C 1989 On the limited memory BFGS method for large scale optimization *Math. Program.* **45** 503–28