UKAEA-CCFE-PR(20)09

C. Bowman, J. R. Harrison, B. Lipschultz, S. Orchard,
K. J. Gibson, M. Carr, K. Verhaegh, O. Myatra

Inference of divertor plasma characteristics using Bayesian multi-diagnostic analysis

Inference of divertor plasma characteristics using Bayesian multi-diagnostic analysis

C. Bowman, J. R. Harrison, B. Lipschultz, S. Orchard, K. J. Gibson, M. Carr, K. Verhaegh, O. Myatra

Inference of divertor plasma characteristics using Bayesian multi-diagnostic analysis

C. Bowman,[1] J.R. Harrison,[2] B. Lipschultz,[1] S. Orchard,[1] K.J. Gibson,[1] M. Carr,[2] K. Verhaegh,[2] and O. Myatra[1]

[1]*York Plasma Institute, Department of Physics, University of York, Heslington, York YO10 5DD,*
United Kingdom
[2]*Culham Centre for Fusion Energy (CCFE), Culham Science Centre, Abingdon, Oxfordshire, OX14 3DB,*
UK

Advancing our understanding of divertor plasma physics is limited by an inability to directly determine the plasma characteristics (density, temperature, etc) over the entire divertor cross-section. At best, diagnostics are able to measure n_e and T_e at isolated points. More commonly however, diagnostics only measure higher-level quantities (e.g. emissivities) which are functions of n_e and T_e. Consequently, a single diagnostic cannot usefully constrain the fields of interest without perturbing the plasma state. We address this problem by using a Bayesian approach to combine information from multiple diagnostic systems to infer the 2D fields of n_e and T_e. We present results of the successful design, implementation and testing of a simple, proof-of-principle system. The synthetic diagnostic measurements used in this testing are derived from SOLPS-ITER fluid code predictions of the MAST-U Super-X divertor, and include appropriate added noise. In these synthetic tests we are able to infer the plasma fields at resolutions better than 2 cm over the outer divertor leg. The mean absolute error between the inferred fields and the SOLPS-ITER test-case was $T_e \sim 1.1$ eV and $n_e \sim 1.5 \times 10^{18}$ m^{-3}, where the test-case has a mean temperature of 6.4 eV and mean density of 1.2×10^{19} m^{-3}.

I. INTRODUCTION

A. 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 usefully constrain the 2D fields of plasma electron temperature and density or the neutral density (at least without perturbing the plasma, for example by strike-point sweeping), which are key in improving our understanding of divertor physics. 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 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 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 over multiple iterations, and a period of weeks to months. Conceptually, in this approach assumptions are made about divertor physics processes, which imply a particular plasma state, and 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. The coverage of diagnostics relevant to our analysis is illustrated in figure 1.

B. 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 diagnos-

tics 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 fusion to 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 section II we discuss the parametrisation of the problem and design requirements of the system. In sections III and IV we show how information regarding the plasma fields in both measurement data and prior knowledge may be expressed as probability distributions. In section V the construction of synthetic test-cases using SOLPS simulations is discussed. In section VI 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 VII. A discussion of potential improvements and further work is given in section VIII, followed by conclusions in section IX.

II. SYSTEM DESIGN

A. Parametric representation of plasma fields

We wish to learn about the continuous fields of plasma properties, such as densities and temperatures, but we cannot infer them directly - a model must be chosen which can adequately describe the types of fields we might reasonably expect to exist. The parameters of this model for the fields are the unknowns which we will seek to infer.

Due to toroidal field ripple and error-field effects, the magnetic field across the divertor is not toroidally symmetric, and this will also be true of the plasma fields. In fact, the layout of the plasma-facing components in the MAST-U divertor is itself not toroidally symmetric, having a periodicity which matches the toroidal field ripple in order to more evenly distribute power across the material surfaces. To know perfectly the plasma state in the divertor would therefore require a 3D description of the fields. Inferring these 3D fields from the available measurements, even if we were to make assumptions regarding the toroidal periodicity of the fields, would be extremely challenging. Instead, as a starting point we will assume toroidal symmetry of the plasma fields in our analysis to make the problem more tractable.

We choose to represent each of the 2D plasma fields via linear interpolation on a triangular mesh, shown in figure 1, 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

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

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 $T_e^{(k)} > 0 \, \forall k$.

A further advantage is the flexibility offered by triangular meshes, such as the ability to selectively partition triangles in order to increase the density of the mesh in areas where greater resolution is required, for example at the strike point and around the divertor leg. The field model in (1) is also linear, which has important implications for efficient calculation of the expected diagnostic signals for a given set of plasma fields, and the derivatives of those signals with respect to the model parameters. These are important considerations for finding the set of plasma fields which are most consistent with the diagnostic data, and are discussed further in later sections.

The mesh shown in figure 1 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, resulting in a total of 586 vertices.

The total number of model parameters has a strong impact on the computational complexity of the inference problem, and because this is proportional to the number of vertices in the field mesh, the number of vertices should be made as small as possible. The mesh described here was chosen to strike a reasonable balance between spatial resolution and computational complexity.

B. 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
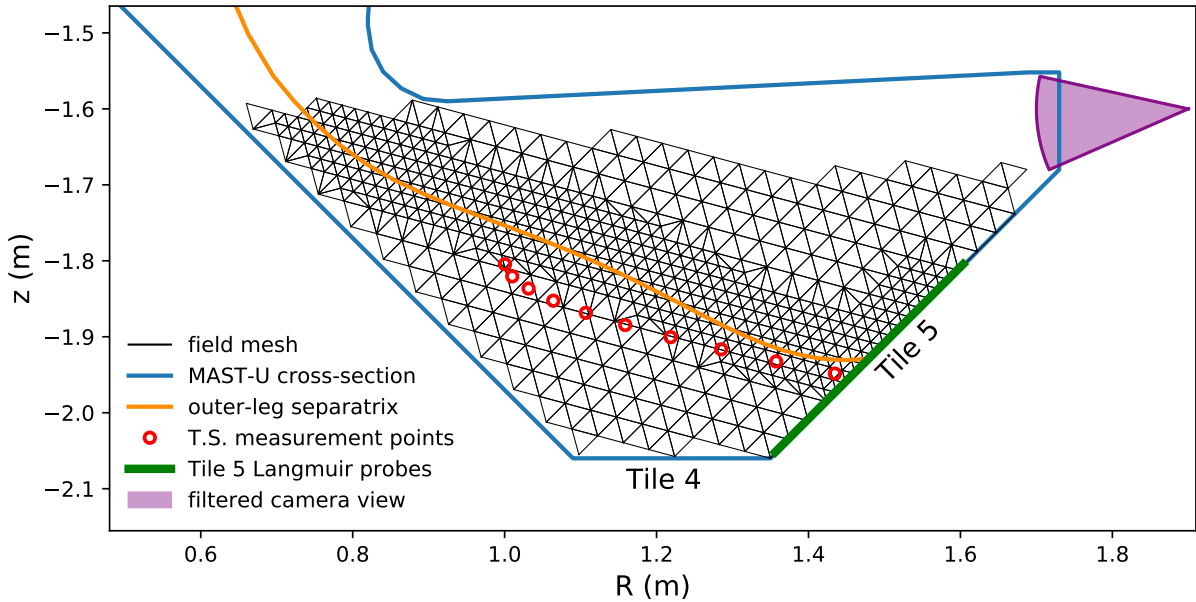
FIG. 1. Illustration of diagnostic coverage in the MAST-U Super-X divertor region, and the triangular mesh used to represent the plasma fields. Only the field-of-view of camera pixels viewing the plasma tangentially is shown here, but the camera is also able to observe the strike point on tile 5 via pixels with non-tangential views. 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.

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.

Both of these requirements are similar in that they demand a modular design, where the code responsible for each diagnostic system is self-contained. An object-oriented programming approach is ideally suited to meeting these requirements. For this reason, and its growing importance as a data analysis language both in fusion and science in general, we chose to develop the system in Python.

Finding the set of plasma fields which are most consistent with the diagnostic data is a challenging optimisation problem in which we must maximise a function called the 'posterior log-probability' (this function will be discussed in detail later). It was our opinion that solving this optimisation problem would require the ability to efficiently calculate the derivative of the posterior log-probability with respect to the model parameters (i.e. the field values at each vertex). This was chosen as the final requirement for the system design.

C. Choice of diagnostics

To include a diagnostic system in our analysis, we must be able to make predictions about what data we expect that diagnostic to collect for a given set of plasma fields (which are specified by the model parameters). In this way, it is the choice of diagnostics which dictates what plasma fields must be modelled. For example, the Thomson scattering (T.S.) system will estimate the electron temperature and density, so to include the T.S. system in the analysis we must model the fields of T_e and n_e.

The number of modelled fields has a large impact on the complexity of the problem, and as such we would like to model a small number fields constrained by as many diagnostics as possible. On account of its ability to make direct measurements of T_e, n_e the divertor T.S. system was chosen as a starting point[10]. The Langmuir probe data (under certain assumptions) can also be modelled using only the T_e and n_e fields, so the probes can be included without the incurring the cost of an additional field[11].

The emissivity of atomic 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. Including multiple filtered cameras in the system while keeping the number of modelled fields to a minimum therefore requires that we choose carefully which lines are being measured.

This could be achieved by viewing a group of lines from the same emitting species and charge-state, but this limits the possible choices of lines. Another option would be to choose lines from a single species but different charge states, and then specify a physics model for the metastable and transport-resolved ionisation balance of that species[12,13]. Rather than modelling a density field for each charge-state, we could model the concentration and transport residence time of the species, and then derive the density of each charge-state from the ionisation balance model.

However, specifying a model for the ionisation balance is a strong physics assumption. If the real ionisation balance present in the divertor plasma cannot be adequately reproduced by the chosen model, the inferred plasma fields may be erroneous. For this reason, we will initially attempt to infer the fields without making assumptions regarding the ionisation balance.

Divertor bolometry will also be available in MAST-U, however we expect that the total radiated power measured by the bolometers would contain non-negligible contributions from impurities (e.g. Carbon). Accurate modelling of the bolometer signals may therefore require the addition of multiple impurity charge-state density fields. As such we chose not to include bolometry in the system at this stage.

The initial version of the system we present here represents a first step toward integrating as many divertor diagnostics as possible in order to better infer the divertor plasma state. This initial development includes 4 filtered cameras which view the Hydrogen Balmer α, β, γ and δ lines, which correspond to the $n = 3, 4, 5, 6$ to $n = 2$ transitions respectively. Under certain assumptions discussed in the following section, data for all system diagnostics may be modelled from three fields; the electron temperature, electron density and hydrogen neutral density.

III. BAYESIAN MULTI-DIAGNOSTIC ANALYSIS

When using experimental data to estimate some parameters of interest, rather than conceiving of the parameters having some fixed 'true' value that could give rise to a variety of possible experimental measurements, the Bayesian view conceptualises the model parameters as fundamentally uncertain by expressing them as a probability distribution, which we can think of as the distribution of possible causes that could have given rise to the measured data.

The Bayesian approach to data analysis therefore uses probability as a means of formally quantifying the information that a given piece of experimental data provides with respect to the model parameters. This allows information from different sources, for example completely separate diagnostic systems, to be combined together to strengthen our knowledge regarding the physical quantities in which we are interested.

In this case the set of model parameters, which we will call θ, are the values of T_e, n_e and n_0 at each vertex of the mesh such that

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

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

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

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 IV.

$P(\mathscr{D})$ 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.

$P(\mathscr{D}|\theta)$ is the *likelihood*, and is the probability that we would observe a dataset $\mathscr{D}$ assuming the plasma were in a state described by a given θ. The use of $\mathscr{D}$ 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. $P(d^{(1)}, d^{(2)}, \ldots, d^{(n)}|\theta)$. By letting $\mathscr{D} = \{d^{(1)}, d^{(2)}, \ldots, d^{(n)}\}$ we may write the likelihood more concisely as $P(\mathscr{D}|\theta)$.

If some set of random variables, in this case $\mathscr{D}$, 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

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

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.

A. Individual diagnostic likelihoods

When analysing data from a single diagnostic, constructing the likelihood for the whole data set as a single product of the likelihoods for all individual data values makes sense, as the $P(d^{(i)}|\theta)$ will be similar in structure as they all relate to the same instrument. For example, we could treat each pixel of a filtered camera system as an individual diagnostic with its

own measurement, but the camera has hundreds of thousands of pixels, and the likelihood will have the same form for each. It therefore makes more sense to treat the entire camera as one instrument and express its likelihood as the product of the likelihoods of each individual pixel.

In the case of multi-diagnostic inference, it is more practical to separate out 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 $\mathscr{D}_{\text{ts}}$ and $\mathscr{D}_{\text{lp}}$ respectively. We will separate out the data for each filtered camera, such that data for the i'th camera is represented by $\mathscr{D}_{\text{fc},i}$. Again making the assumption of mutual independence between the datasets, the likelihood for all data can be now written as

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

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 $\mathscr{L}$ to indicate a log-probability density function, such that $\mathscr{L}(A|B) = \ln P(A|B)$. Now combining (3) and (5) we can express the log-posterior distribution $\mathscr{L}(\theta|\mathscr{D})$ as

$$\mathscr{L}(\theta|\mathscr{D}) = \mathscr{L}(\mathscr{D}_{\text{ts}}|\theta) + \mathscr{L}(\mathscr{D}_{\text{lp}}|\theta) + \left(\sum_i \mathscr{L}(\mathscr{D}_{\text{fc},i}|\theta)\right)$$
$$+ \mathscr{L}(\theta) - \mathscr{L}(\mathscr{D}). \qquad (6)$$

At this point it is useful to consider how the structure of the log-posterior supports the object-oriented design approach discussed in section II B. All terms in (6) (except $\mathscr{L}(\mathscr{D})$, which is in practice discarded) can be evaluated independently - this allows each term to be represented as a single object, which contains all experimental data and forward models required to evaluate that term. The resulting structure of the system is illustrated in figure 2. This approach means that any of the terms can be easily included or excluded from the log-posterior, fulfilling one of the design requirements.

Rather than passing the full set of parameters to each diagnostic, we introduce a 'plasma state' object which acts a mediator between the values in θ and the diagnostic objects. Each diagnostic object has access to the plasma state object, and can request from it the value of a given field at a given point in space to allow the forward model to be evaluated, which predicts the expected experimental measurements. In this way, the field parametrisation can be changed (for example by switching to a different mesh) without making any alteration to the likelihood or prior objects.

B. 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 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 a uncorrelated bivariate normal distribution such that

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

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 $\mathscr{L}(\mathscr{D}_{\text{lp}}|\theta)$ is also of the form given in (7).

C. filtered camera system likelihood

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

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

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 $\mathscr{E}$. As the fields are defined through Barycentric interpolation, which is linear, this line-integral can be represented exactly by a weighted sum of the
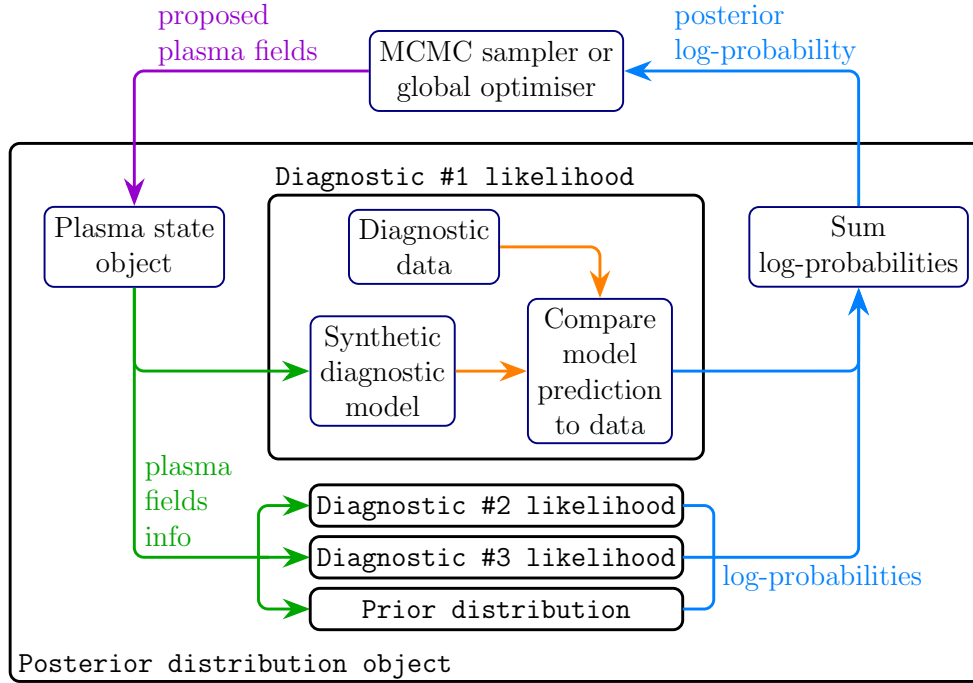
FIG. 2. Flow chart illustrating the code structure of the system. The posterior distribution is encapsulated as a single object which takes model parameters (which define the plasma fields) as inputs and returns the posterior log-probability. The prior distribution and the likelihood for each diagnostic 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.

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 $\mathbf{G}\mathscr{E}$ yields a prediction of the pixel brightness values. Although the pixel brightnesses are strictly non-negative, for simplicity we presently represent the filtered camera likelihood as multivariate normal such that

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

Experimental calibration of filtered cameras systems typically find the variance of the pixel brightnesses (assuming the pixel is not near saturation) to be linear such that

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

where α, β 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 $b_{\max}$. Second, we fix the fractional error at the maximum brightness to be a constant $f_{\max}$. Under these assumptions the variance may be expressed as

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

IV. 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 $\boldsymbol{\theta}$ to be equally likely), which is itself a choice. As we don't have prior information from another source of experimental data, our goal should be to choose a prior which excludes unrealistic plasma conditions. In order to do this, we require information about the space of realistic plasma 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[14]. 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 $(1.7 \rightarrow 2.5$ 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 across all simulations were concatenated. 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 3.
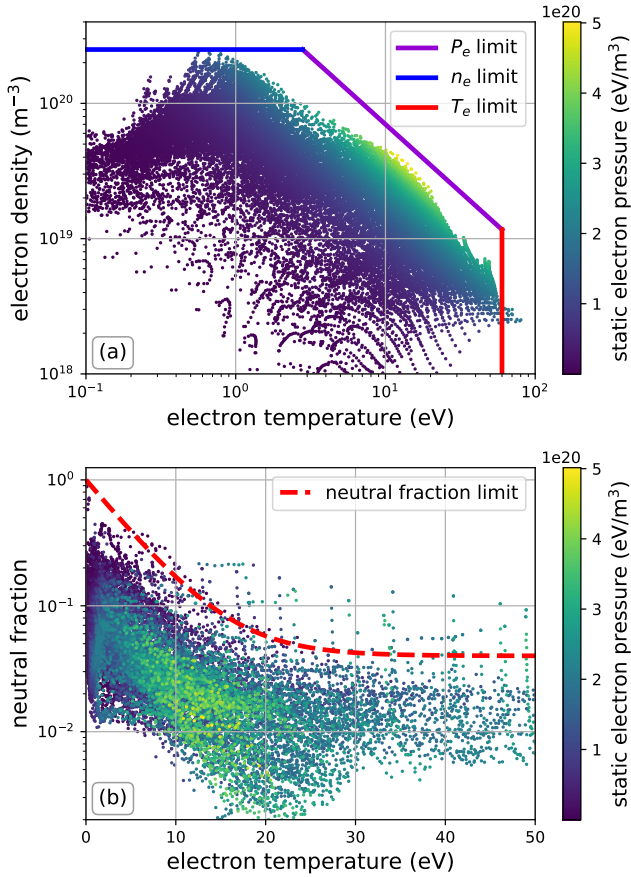
FIG. 3. The space of realistic plasma conditions in the MAST-U divertor as predicted by a group of SOLPS simulation results spanning various experimental configurations. (a) The space of realistic (T_e, n_e) can be well approximated by placing upper limits on T_e, n_e and P_e. (b) The space of realistic $n_0/(n_e + n_0)$ can be bounded by placing an upper limit on $n_0/(n_e + n_0)$ as a function of T_e.

TABLE I. 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	$1 \times 10^{16}\,\mathrm{m}^{-3}$	$1 \times 10^{15}\,\mathrm{m}^{-3}$
upper bound	60 eV	$2.5 \times 10^{20}\,\mathrm{m}^{-3}$	$2 \times 10^{20}\,\mathrm{m}^{-3}$

A. 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 I. Imposing these bounds explicitly by using a uniform prior introduces discontinuities to the posterior which can cause problems for gradient-based sampling and optimisation algorithms. Instead, the bounds are imposed by the sampling/optimisation algorithms themselves.

B. 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^{max}, and Gaussian for values above the limit. The resulting static electron pressure log-prior is

$$\mathcal{L}(\theta)_{\mathrm{pressure}} = -\frac{1}{2\sigma_{\mathrm{prs}}^2} \sum_i \max\left(\frac{n_e^{(i)} T_e^{(i)}}{P_e^{\mathrm{max}}} - 1, 0\right)^2. \quad (12)$$

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

C. Neutral fraction prior

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

$$f_i^{\mathrm{max}} = (1-c)\exp\left(-T_e^{(i)}/l\right) + c, \quad (13)$$

where $c = 0.04$ and $l = 5\,\mathrm{eV}$. In figure 3 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 (12) such that

$$\mathcal{L}(\theta)_{\mathrm{fraction}} = -\frac{1}{2\sigma_{\mathrm{frc}}^2} \sum_i \max\left(\frac{1}{f_i^{\mathrm{max}}} \frac{n_0^{(i)}}{n_e^{(i)} + n_0^{(i)}} - 1, 0\right)^2, \quad (14)$$

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

D. Spatial smoothness prior

Suppose $\mathbf{v}$ is a vector of field values at each mesh vertex, and define the 'umbrella' matrix operator $\mathbf{U}$ such that

$$\mathrm{U}_{ij} = \begin{cases} -1 & \text{if } i = j \\ 1/n & \text{if vertex } j \text{ is one of } n \text{ vertices connected to } i \\ 0 & \text{if vertex } j \text{ is not connected to } i \end{cases} \quad (15)$$

The product $\mathbf{Uv}$ 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. Consequently, the sum of the squares of the umbrella differences, $|\mathbf{Uv}|^2$, is a useful starting point from which to construct a prior that favours spatial smoothness.

It is helpful to consider whether the fields can be transformed such that the expected solutions for the transformed 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 $\mathbf{S} = \mathbf{U}^\top \mathbf{U}$. The un-smoothness (quantified by the sum of the squares of the umbrella differences) 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 σ_{smth}, which determines how strongly un-smooth fields are penalised, we can define the smoothing log-prior as

$$\mathscr{L}(\boldsymbol{\theta})_{\mathrm{smoothness}} = -\frac{1}{2\sigma_{\mathrm{smth}}^2} \left(\tilde{T}_e^\top \mathbf{S} \tilde{T}_e + \tilde{n}_e^\top \mathbf{S} \tilde{n}_e + \tilde{n}_0^\top \mathbf{S} \tilde{n}_0 \right). \tag{16}$$

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 σ_{smth} - this is discussed further in section VII.

The log-prior distribution can be defined as the sum of the various components described in this section such that

$$\mathscr{L}(\boldsymbol{\theta}) = \mathscr{L}(\boldsymbol{\theta})_{\mathrm{pressure}} + \mathscr{L}(\boldsymbol{\theta})_{\mathrm{fraction}} + \mathscr{L}(\boldsymbol{\theta})_{\mathrm{smoothness}}. \tag{17}$$

V. PRODUCTION OF SYNTHETIC TEST-CASE DATA

A. 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 $2 \times 10^{21}\,\mathrm{s}^{-1}$. 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 $2 \times 10^{20}\,\mathrm{s}^{-1}$ and $5 \times 10^{20}\,\mathrm{s}^{-1}$ respectively. Both cases are accessible via the MDSplus database[15], stored as runs number 121844 for the low-seeding case and 121847 for the high-seeding case. These two cases are not part of the set used to inform the prior constraints discussed in section IV, and 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 Nitrogren, 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 4.

B. 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 $f_{\mathrm{max}} = 0.025$ and $f_0 = 1/256$. The Deuterium-α emissivity field derived from the low-seeding case, and the corresponding synthetic data image are shown in figure 5.

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

VI. CHARACTERISING THE POSTERIOR DISTRIBUTION

Now that all terms in (6) which have a dependence on θ 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 that allows us to extract useful information about the plasma fields.

A. Gradient-based approach

The filtered camera data, which constitutes the majority of the information in the problem, introduces strong correlations between the model parameters due to both the line-integrated nature of the measurement and the dependence of the emissivities on the plasma fields. This results in a high-dimensional, highly-correlated posterior distribution, which strongly favours 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. It was decided therefore that efficient calculation of the log-posterior derivative should be a core part of the system design, and this was added to the list of design requirements discussed in section II B. As a result, evaluating the derivative of the log-posterior takes approximately a factor of 3 longer than evaluating the log-posterior itself - this is approximately 600 times faster than evaluating the derivative using finite-difference for the current number of model parameters.
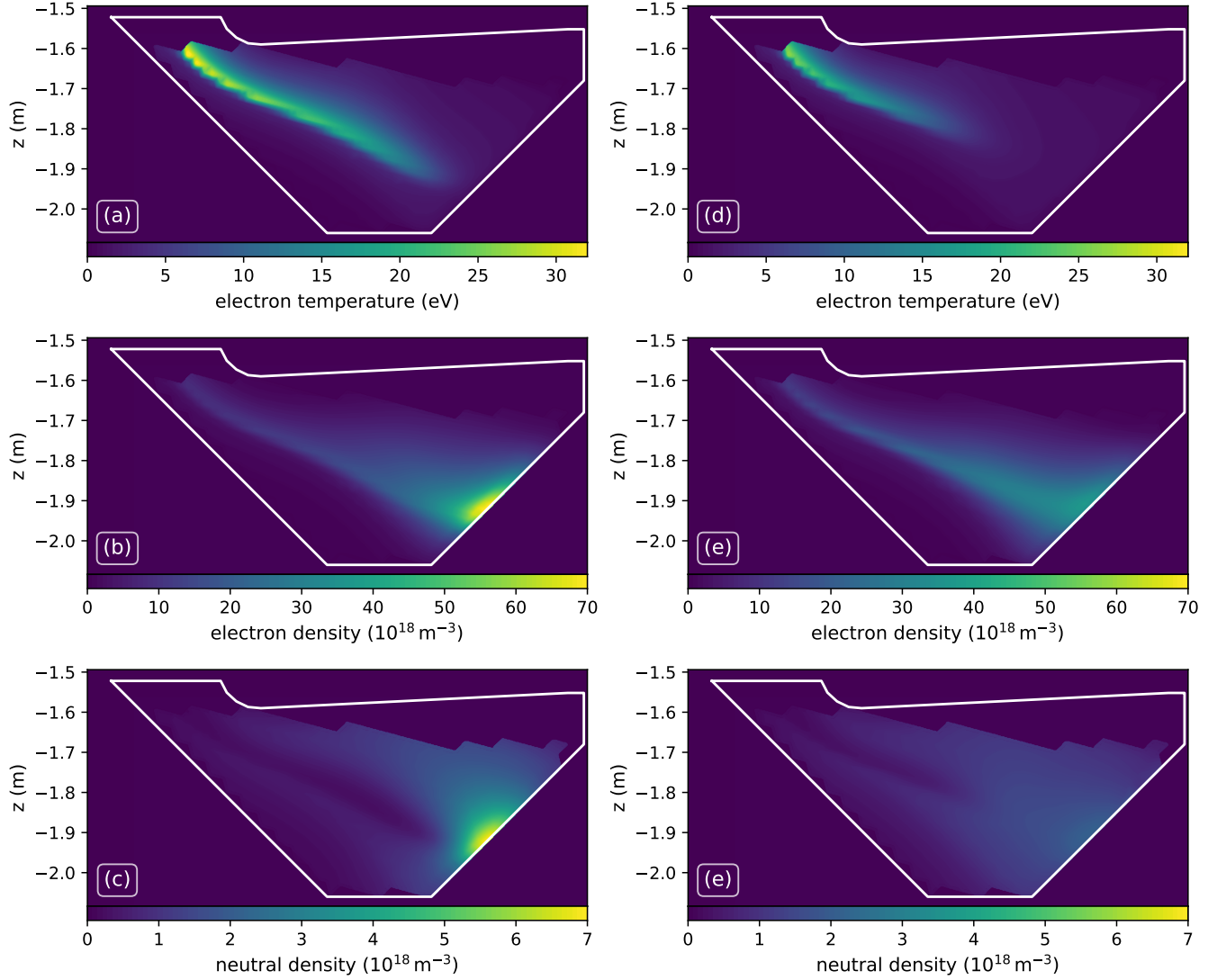
FIG. 4. Plots showing the mesh-representations of the electron temperature, electron density and neutral density taken from the Nitrogen-seeded SOLPS predictions. Plots (a,b,c) and (d,e,f) show the low- and high-seeding cases respectively

B. 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[16]. To locate this maximum, we employ a 'hybrid' approach which combines a genetic algorithm with the L-BFGS algorithm[17]. 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 convergences 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, and these evaluations constitute the vast majority of the computational cost of the approach. These computations can be efficiently distributed across multiple CPUs, up to a limit where each candidate solution is assigned to a separate CPU, thus reducing the total computation time significantly.

The convergence of the solution log-probability using this approach is shown in figure 6. In this case a population of 20 candidate solutions was distributed over 20 threads of a Intel Xeon E5-2695 v3. The maximum log-probability had converged sufficiently after 80 generations, which in total took approximately 80 minutes to evaluate. In the first few generations, we see that many of the L-BFGS runs converge to local maxima with probability densities significantly below that of the global maximum. This highlights the necessity of employ-
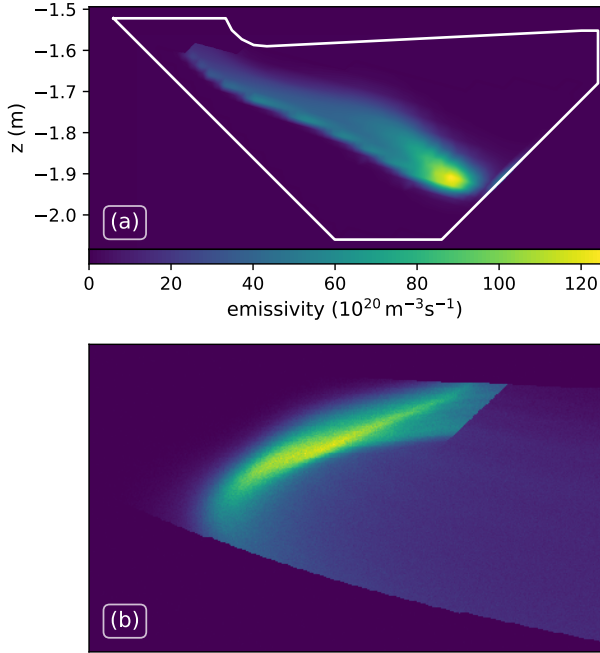
FIG. 5. (a) The Deuterium-α emissivity field as predicted by (8) for the low-seeding case shown in figure 4. (b) The corresponding synthetic Hydrogen-α filtered-camera image with added simulated noise.

ing a 'global' optimisation strategy such as that described here when determining the MAP estimate.

C. 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 particularly effective (in comparison to other approaches) in cases where the number of model parameters is large, and strong correlations are present[18].

D. 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.
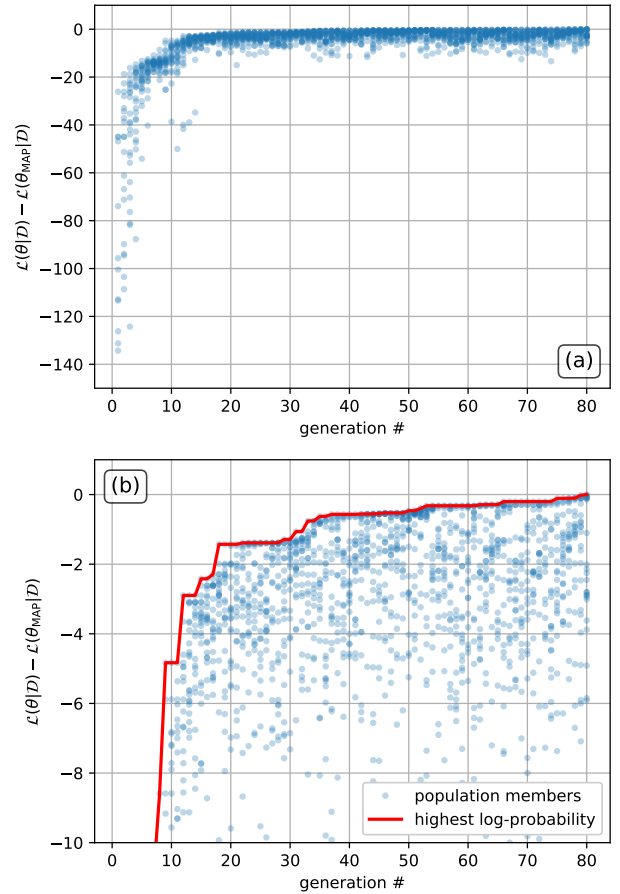

FIG. 6. Plots of the posterior log-probability $\mathscr{L}(\theta|\mathscr{D})$ of each member of the population of candidate solutions at each generation of the optimisation, minus the highest observed log-probability $\mathscr{L}(\theta_{\mathrm{MAP}}|\mathscr{D})$. (a) all solutions; (b) only solutions with probabilities near the maximum, where the red line shows the highest log-probability observed so far at each generation.

VII. RESULTS OF SYNTHETIC TESTING

A. Selecting the smoothing prior uncertainty

The value of the smoothing prior uncertainty σ_{smth}, which appears in (16), 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 σ_{smth} - the results of this scan are shown in figure 7. The minima in the error for each field are fairly broad, but do not all occur at the same value of σ_{smth}. The results presented here used a value of $\sigma_{\mathrm{smth}} = 0.2$ which provides a good balance between low error in electron temperature and electron density.

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 experimen-
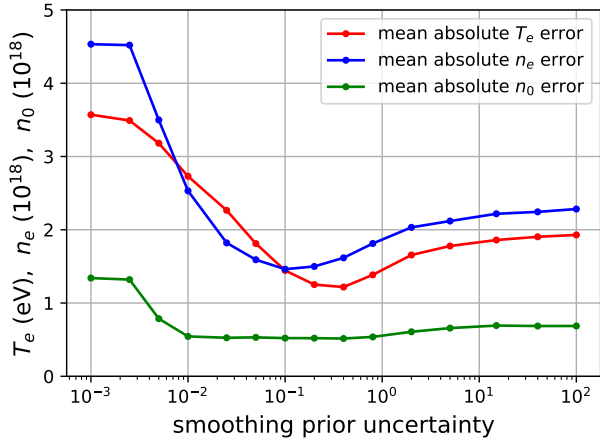
FIG. 7. Mean absolute difference between the field values of the low-seeding case and those of the corresponding MAP estimate for each of the plasma fields as a function of the smoothing prior uncertainty σ_{smth}.

tal data will be the subject of further work.

B. Comparison of inferred fields and test-cases

The maximum a postiori (MAP) and marginal expectation (MEX) estimates were evaluated as described in section VI for both test-cases. The inferred field values from the MEX estimate are compared with those from the corresponding test case in figure 8, and the mean absolute differences for each field are given in table II for both the MAP and MEX. 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 ∼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 inferred fields for the low-seeding case, along with the differences between the inferred fields and the test-case are shown in figure 9. 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. In all such tests with a single diagnostic, it was not possible to make any sensible estimates of the 2D plasma fields.

C. Uncertainty estimation

Uncertainties in the inferred fields for both test-cases were estimated by sampling from the posterior distribution using Hamiltonian Monte-Carlo. Twenty separate Markov-chains were initiated in parallel from different positions close to the MAP estimate, and were allowed to evolve for 24 hours. Any samples in each chain prior to the point where the simulation step-size had converged were discarded as burn-in, and the remaining samples from all chains were then aggregated.

Figure 10 compares the test-cases and inferred fields along a flux surface which lies in the centre of the scrape-off layer, 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 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[19].

VIII. DISCUSSION

A. 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'.

TABLE II. The mean absolute difference between the field values at each vertex for each test case and its corresponding inferred value, for both the MAP and MEX estimates.

	mean $|\Delta T_e|$ (eV)	mean $|\Delta n_e|$ (m^{-3})	mean $|\Delta n_0|$ (m^{-3})
low-seeding MAP	1.12	1.57×10^{18}	3.50×10^{17}
low-seeding MEX	1.09	1.46×10^{18}	3.38×10^{17}
high-seeding MAP	0.81	1.60×10^{18}	7.37×10^{17}
high-seeding MEX	0.74	1.65×10^{18}	4.05×10^{17}
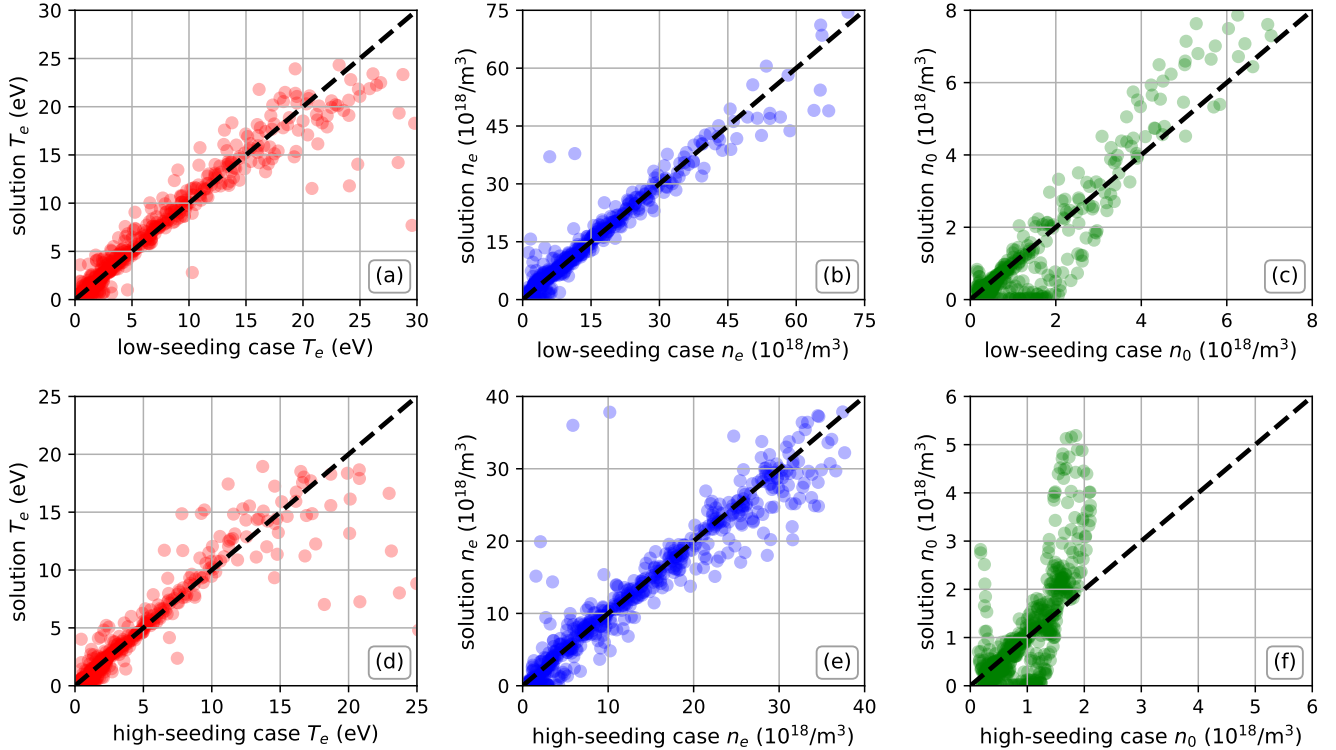
FIG. 8. Scatter plots of the field values of the test-cases versus those of the MEX estimate at each vertex of the mesh. Plots (a,b,c) and (d,e,f) show results from the low- and high-seeding cases respectively.

This process involves allowing the calibration values 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 **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 detailed approach to calculating the geometry matrix[20], 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.

B. 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 bright-
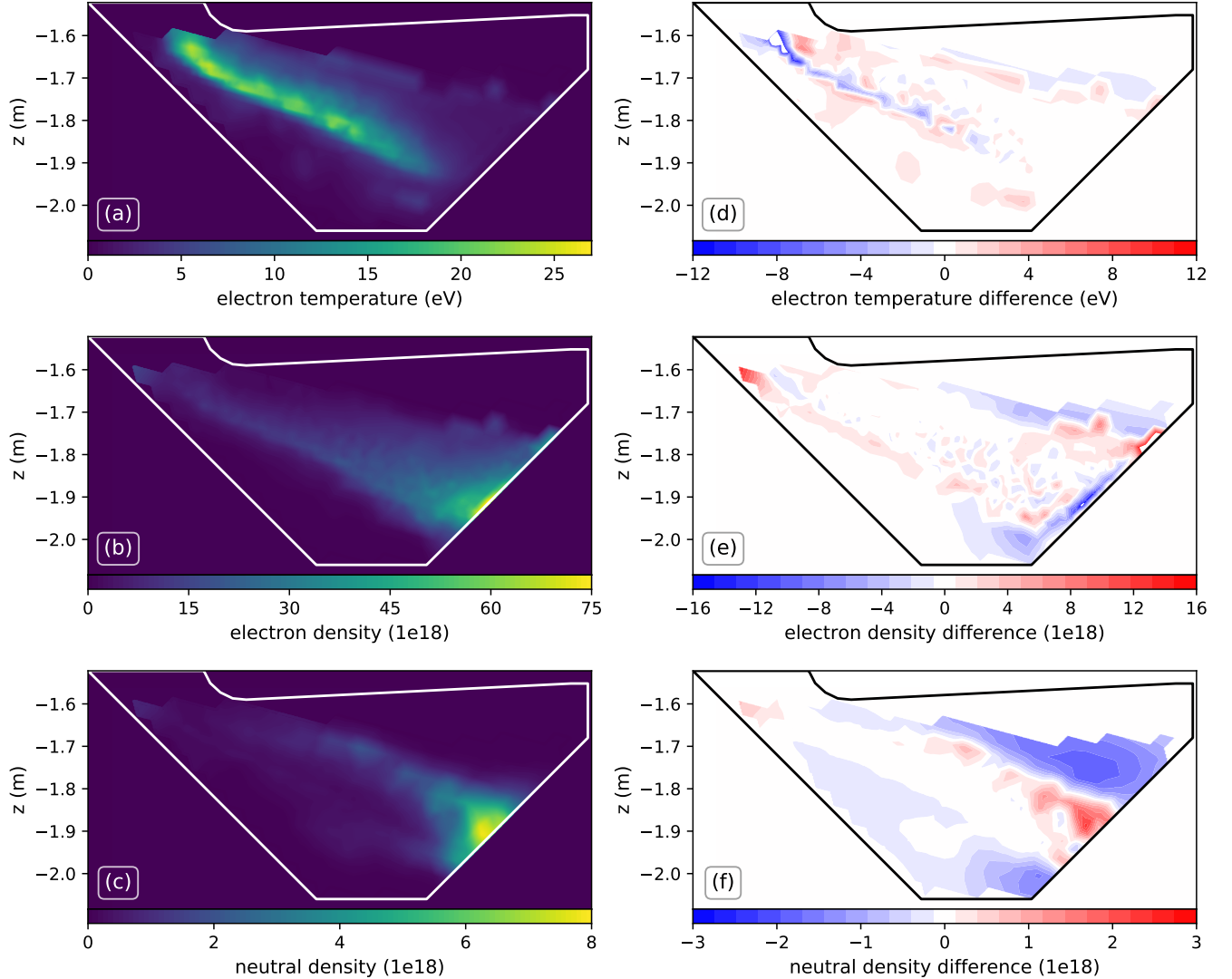
FIG. 9. The marginal expectation estimate of the electron temperature, electron density and neutral density fields for the low-seeding case are shown in (a), (b) and (c) respectively. The corresponding differences between the marginal expectation estimate of a field and the test-case values are shown in (d), (e) and (f).

nesses 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[21].

A coherence imaging system is planned for MAST-U, and will share the same view as the MWI system. Forward-modelling to the raw coherence imaging data would be challenging, as this requires information regarding plasma flows. However, it would be possible to model quantities derived from post-processing of the coherence imaging data, such as the emission-weighted contrast. Inclusion of this data in the analysis may provide powerful additional information, due to the difference in electron density dependence between the emission-weighted contrast and other line-brightness measurements.

C. Choice of imaged spectral lines

The emissivity model in (8) does not include molecular emission, which may be a non-negligible component of low-n Balmer series emissivities, particularly for Deuterium Balmer-α, in strongly detached conditions. Deuterium-α through δ were chosen as a sensible 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.

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 (con-
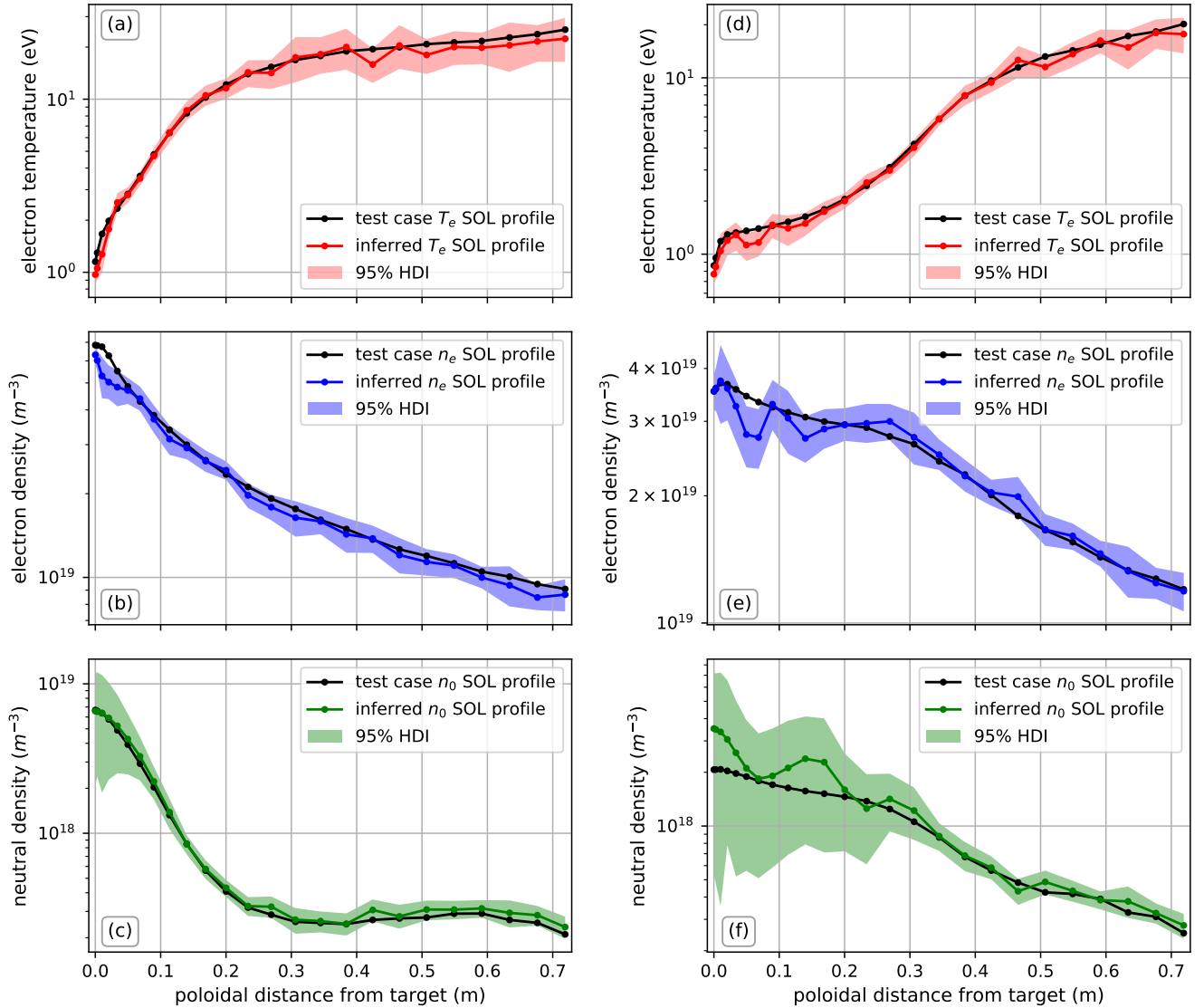
FIG. 10. Comparison of profiles from the centre of the scrape-off layer derived from the inference results and the test-cases. Plots (a,b,c) and (d,e,f) show results from the low- and high-seeding cases respectively. The coloured areas indicate the 95% highest-density intervals derived from the sampling results.

sidering 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.

IX. SUMMARY AND CONCLUSIONS

We have presented details of the first design, implementation and testing of a Bayesian multi-diagnostic inference system for the MAST-U divertor. The system has been designed to be modular and flexible, so that the diagnostics modelled by the system can be changed easily.

In order to test the system, synthetic experimental measurements for filtered cameras viewing the first four Balmer lines, the divertor Thomson scattering system and the Langmuir probes were derived from SOLPS simulations, and included appropriate added noise. These 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 'so-

lution' merits further investigation.

Ongoing developement of the system will focus on the inclusion of uncertainties resulting from instrument calibrations, determining the optimal set of spectral lines to image in order to best constrain the plasma fields, and consideration of other diagnostic systems for inclusion in the system, such as chordal spectroscopy.

ACKNOWLEDGMENTS

[1] P. Stangeby, *The Plasma Boundary of Magnetic Fusion Devices* (2000).

[2] S. Wiesen, D. Reiter, V. Kotov, M. Baelmans, W. Dekeyser, A. Kukushkin, S. Lisgo, R. Pitts, V. Rozhansky, G. Saibene, I. Veselova, and S. Voskoboynikov, "The new solps-iter code package," Journal of Nuclear Materials **463**, 480 – 484 (2015), plasma-Surface Interactions 21.

[3] H. P. Summers, W. J. Dickson, M. G. O'Mullane, N. R. Badnell, A. D. Whiteford, D. H. Brooks, J. Lang, S. D. Loch, and D. C. Griffin, "Ionization state, excited populations and emission of impurities in dynamic finite density plasmas: I. The generalized collisional-radiative model for light elements," Plasma Physics and Controlled Fusion **48**, 263–293 (2006).

[4] W. Morris, J. R. Harrison, A. Kirk, B. Lipschultz, F. Militello, D. Moulton, and N. R. Walkden, "MAST Upgrade Divertor Facility: A Test Bed for Novel Divertor Solutions," IEEE Transactions on Plasma Science **46**, 1217–1226 (2018).

[5] W. Vijvers, R. Mumgaard, Y. Andrebe, I. Classen, B. Duval, and B. Lipschultz, "2017 Jinst 12 C07006," (2017).

[6] R. Fischer, C. J. Fuchs, B. Kurzan, W. Suttrop, and E. Wolfrum, "Integrated data analysis of profile diagnostics at ASDEX upgrade," Fusion Science and Technology **58**, 675–684 (2010).

[7] O. P. Ford, J. Svensson, M. Beurskens, A. Boboc, J. Flanagan, M. Kempenaars, and D. C. Mcdonald, "Bayesian Combined Analysis of JET LIDAR

[8] , Edge LIDAR and Interferometry Diagnostics," 36th EPS Conference on Plasma Physics **33**, 1–10 (2009).

[9] J. Svensson and A. Werner, "Current tomography for axisymmetric plasmas," Plasma Physics and Controlled Fusion **50** (2008), 10.1088/0741-3335/50/8/085002.

[10] G. T. Von Nessi, M. J. Hole, J. Svensson, and L. Appel, "Evidence cross-validation and Bayesian inference of MAST plasma equilibria," Physics of Plasmas **19** (2012), 10.1063/1.3677362.

[11] J. Hawke, R. Scannell, J. Harrison, R. Huxford, and P. Bohm, "Outline of optical design and viewing geometry for divertor Thomson scattering on MAST upgrade," Journal of Instrumentation **8** (2013), 10.1088/1748-0221/8/11/C11010.

[12] O. Février, C. Theiler, H. De Oliveira, B. Labit, N. Fedorczak, and A. Baillod, "Analysis of wall-embedded Langmuir probe signals in different conditions on the Tokamak à Configuration Variable," Review of Scientific Instruments **89** (2018), 10.1063/1.5022459.

[13] S. Potzel, R. Dux, H. W. Müller, A. Scarabosio, and M. Wischmeier, "Electron density determination in the divertor volume of ASDEX Upgrade via Stark broadening of the Balmer lines," Plasma Physics and Controlled Fusion **56** (2014), 10.1088/0741-3335/56/2/025010.

[14] S. Henderson, M. Bernert, S. Brezinsek, M. Carr, M. Cavedon, R. Dux, B. Lipschultz, M. O'Mullane, F. Reimold, and M. Reinke, "Determination of volumetric plasma parameters from spectroscopic N II and N III line ratio measurements in the ASDEX Upgrade divertor," Nuclear Fusion **58**, 016047 (2018).

[15] J. Harrison, D. Moulton, M. Carr, I. Chapman, D. Keeling, A. Kirk, and N. Walkden, "Enhancements to mast upgrade to address the eurofusion plasma exhaust strategy," 43rd EPS Conference on Plasma Physics '2016).

[16] Specifically, the ASDEX-Upgrade MDSplus server 'solps-mdsplus.aug.ipp.mpg.de:8001::solps'.

[17] U. Von Toussaint, "Bayesian inference in physics," Reviews of Modern Physics **83**, 943–999 (2011).

[18] J. Nocedal and D. C. Liu, "On the limited memory BFGS method for large scale optimization," Mathematical Programming **45**, 503–528 (1989).

[19] R. M. Neal, "MCMC using Hamiltonian dynamics," (2012), arXiv:1206.1901.

[20] M. Sambridge, "A parallel tempering algorithm for probabilistic sampling and multimodal optimization," Geophysical Journal International **196**, 357–374 (2013).

[21] M. Carr, A. Meakins, S. A. Silburn, J. Karhunen, M. Bernert, C. Bowman, A. Callarelli, P. Carvalho, C. Giroud, J. R. Harrison, S. S. Henderson, A. Huber, B. Lipschultz, T. Lunt, D. Moulton, and F. Reimold, "Physically principled reflection models applied to filtered camera imaging inversions in metal walled fusion machines," Review of Scientific Instruments **90**, 043504 (2019).

[22] K. Verhaegh, B. Lipschultz, B. Duval, O. Février, A. Fil, C. Theiler, M. Wensing, C. Bowman, D. Gahle, J. Harrison, B. Labit, C. Marini, R. Maurizio, H. de Oliveira, H. Reimerdes, U. Sheikh, C. Tsui, N. Vianello, W. Vijvers, and and, "An improved understanding of the roles of atomic processes and power balance in divertor target ion current loss during detachment," Nuclear Fusion **59**, 126038 (2019).