Evidence cross-validation and Bayesian inference of MAST plasma equilibria

G. T. von Nessi, M. J. Hole, J. Svensson, and L. Appel

Related Articles

Additional information on Phys. Plasmas

Evidence cross-validation and Bayesian inference of MAST plasma equilibria

G. T. von Nessi,[1,a] M. J. Hole,[1] J. Svensson,[2] and L. Appel[3]

[1]*Research School of Physical Sciences and Engineering, The Australian National University, Canberra ACT 0200, Australia*
[2]*Max-Planck-Institut für Plasmaphysik, D-17491 Greifswald, Germany*
[3]*EURATOM/CCFE Fusion Association, Culham Science Centre, Abingdon, Oxon OX14 3DB, United Kingdom*

In this paper, current profiles for plasma discharges on the mega-ampere spherical tokamak are directly calculated from pickup coil, flux loop, and motional-Stark effect observations via methods based in the statistical theory of Bayesian analysis. By representing toroidal plasma current as a series of axisymmetric current beams with rectangular cross-section and inferring the current for each one of these beams, flux-surface geometry and q-profiles are subsequently calculated by elementary application of Biot-Savart's law. The use of this plasma model in the context of Bayesian analysis was pioneered by Svensson and Werner on the joint-European tokamak [Svensson and Werner, Plasma Phys. Controlled Fusion **50**(8), 085002 (2008)]. In this framework, linear forward models are used to generate diagnostic predictions, and the probability distribution for the currents in the collection of plasma beams was subsequently calculated directly via application of Bayes' formula. In this work, we introduce a new diagnostic technique to identify and remove outlier observations associated with diagnostics falling out of calibration or suffering from an unidentified malfunction. These modifications enable a good agreement between Bayesian inference of the last-closed flux-surface with other corroborating data, such as that from force balance considerations using EFIT++ [Appel *et al.*, "A unified approach to equilibrium reconstruction" *Proceedings of the 33rd EPS Conference on Plasma Physics* (Rome, Italy, 2006)]. In addition, this analysis also yields errors on the plasma current profile and flux-surface geometry as well as directly predicting the Shafranov shift of the plasma core. © *2012 American Institute of Physics*. [doi:10.1063/1.3677362]

I. INTRODUCTION

Bayesian analysis has been used in data analysis in a wide range of contexts, spanning from astronomy to artificial intelligence.[1–4] Over the last decade, the techniques of Bayesian inference in fusion plasmas have been developed.[5–13] The motivation of these developments has largely been to incorporate uncertainty in parameters, model, and data, which is not easily captured in many traditional methods such as least squares fitting. Specifically, least squares fitting is not an intrinsically statistical procedure, in that it will only yield most-probable parameter configurations but not any moments of a probability distribution (e.g., the expectation or covariance matrix). Thus, least squares fits usually have to be augmented with a linearisation of propagated errors through the associated forward model, to obtain uncertainties. In the case of non-linear models, linear error-propagation may yield misleading results; and as the topic of non-linear models is extremely broad, there is no general way to explicitly propagate errors in such systems without employing some level of approximation or Taylor expansion of uncertainties. In the Bayesian probabilistic framework, however, the uncertainty of the plasma parameters, model and data are made explicit: their statistics are related by Bayes' theorem, which relates the probability of inferred parameters given the data (the posterior) to the probability of the data given the parameter (the likelihood) and the probability of the parameter (the prior).

For each initial hypothesis of the plasma parameter, a forward model describes the predicted data signal. Inversion of plasma parameters becomes a probabilistic inference technique, where one must in general compute the posterior probability distribution function (PDF), and marginalising over subspaces by using sampling techniques (e.g., Markov Chain Monte Carlo).

In fusion, Bayesian inversion is more complicated as many plasma diagnostic signals are often functions of multiple plasma parameters, including the magnetic field equilibrium geometry. Moreover, the equilibrium geometry itself has a non-linear dependence on many plasma quantities that are difficult to measure directly (e.g., plasma pressure, flow, and internal poloidal currents). To help address the complexity of this situation, Svensson *et al.*[9] introduced the use of graphical models to fusion, which provide visualisation of the PDF inter-dependance between parameter, data and model. Such graphical models provide an efficient means by which more complex physics and diagnostic forward models can be coded and visualised.

A good example of the latter is Bayesian toroidal current profile tomography, as developed by Svensson and Werner on the joint-European tokamak (JET).[10] In this plasma model, forward models are constructed for magnetic pickup coils, flux loops, and motional Stark effect (MSE) observations, the toroidal plasma current is modelled as a series of axisymmetric current beams, and the current in each beam inferred using Bayesian inversion. Svensson and Werner were able to develop a fast Bayesian inversion by the use of

a)Electronic mail: greg.vonnessi@anu.edu.au.

a conditional auto-regressive (CAR) prior, which makes the posterior a Gaussian PDF. This made exploration of the posterior computationally feasible, as the posterior then had an analytic representation.

In this work, we build on the success of Bayesian inference for current tomography on JET, and implement toroidal current profile tomography on MAST. To this end, we have developed new techniques that identify and remove extreme outlier measurements coming from diagnostics that are not modelled correctly in our analysis (i.e., have undergone some level of failure or have fallen out of calibration); without this, it is not possible to infer a mean configuration of plasma beams that yields any physical flux-surface geometry. Indeed, these techniques are not meant to refine the resolution of existing inferences but are designed to minimise the manipulation of raw data, while still preserving the existence of a physical solution, i.e., where there are no currents outside the plasma volume. Using these techniques, which (in this case) remove both pickup coils and flux loop observations, we show that flux-surfaces calculated from Bayesian current profile tomography show strong agreement with last-closed flux-surfaces (LCFS) generated by EFIT++, a Grad-Shafranov solver.[14–16]

The paper is structured as follows: in Sec. II, a general overview of Bayes' formula is given in the context of diagnostic data analysis. Following this, physical models used for the diagnostic predictions and plasma current reconstruction are presented. New techniques used to identify and remove outlier measurements coming from the magnetic diagnostics are introduced in Sec. IV, and the results are presented in Sec. V for a high-performance MAST discharge. Finally, Sec. VI contains concluding remarks and discusses future directions.

II. OVERVIEW OF BAYESIAN DIAGNOSTIC ANALYSIS

The basis of our work is a probabilistic method that uses diagnostic forward models in the context of Bayes' Formula. First, we introduce the notation $\mathcal{P}(A|B)$, which denotes a probability associated with a set of events A occurring given preconditioned set of events or parameters B. Next, taking H to represent the hypothesis—a collection of parameters to be inferred in the context of a physical model reflecting the quantities that are under observation—and D the diagnostic data, Bayes' formula can be cast as[1,4]

$$\mathcal{P}(H|D) = \frac{\mathcal{P}(D|H) \cdot \mathcal{P}(H)}{\mathcal{P}(D)}. \qquad (1)$$

Implicit in this notation are background assumptions made in the particular context of the inference; the specific assumptions used in the forthcoming analysis will be pointed out explicitly when appropriate. In Eq. (1), $\mathcal{P}(H|D)$ is called the *posterior*, $\mathcal{P}(D|H)$ the *likelihood*, $\mathcal{P}(H)$ the *prior*, and $\mathcal{P}(D)$ the *evidence*. The posterior represents an updated state of knowledge, when a prior state of knowledge is updated with an observation manifested via the likelihood. By subsequently using a posterior as a new prior, one is able to construct an iterative process by which a given prior state of knowledge is updated with any number of observations.

Finally, the evidence is the renormalisation constant that ensures that the right-hand side of Eq. (1) is indeed a probability distribution. In this research, likelihood distributions are fixed, modulo translations determined by diagnostic observations. This is quite general, as diagnostic observations are often associated with Gaussian distributions, whose standard deviation corresponds to the given error of the diagnostic. Assuming likelihoods to be Gaussian serves to minimise the information contained in these distributions, when only diagnostic observations and uncertainties are known.[4] In this scenario, one may think of the evidence as a measure of the consistency all observations with each other and the prior knowledge. This can be understood by thinking of two arbitrary, freely translating, Gaussian probability distributions with fixed variances multiplied together. The integral of the result will decrease as the expectation of these distributions move away from each other and is maximised when the expectations are identical.

In order to clarify the relationship of Eq. (1) in the context of diagnostic observations, we associate a vector of model parameters, $\bar{\lambda}$, with H and a vector of given observations $\bar{\mu}$ with a corresponding error vector $\bar{\sigma}$ with the data D. With this, one can rewrite Eq. (1) as

$$\mathcal{P}(H|D) \propto \mathcal{P}(\mu_i|\bar{\lambda}, \sigma_i) \cdot \mathcal{P}(\bar{\lambda}), \qquad (2)$$

for the specific observation corresponding to μ_i and σ_i. As $\bar{\mu}$ and $\bar{\sigma}$ are given and thus assumed to be constant, $\mathcal{P}(D) = \mathcal{P}(\bar{\mu}|\bar{\sigma})$ is also just a constant, which justifies the proportionality in Eq. (2). Implicit to $\mathcal{P}(\mu_i|\bar{\lambda}, \sigma_i)$, is the *forward model*, $\overline{\mathcal{F}}(\bar{\lambda})$, that relates an arbitrary configuration of model parameters to a given set of observations, in a sense that $\overline{\mathcal{F}}(\bar{\lambda})$ is constructed to take in a set of model parameters and generate a vector of *predictions* meant to directly correspond to the given set of observations.

Using Eq. (2) in Eq. (1) and iterating over all diagnostic observations, one finds that

$$\mathcal{P}(H|D) \propto \left(\prod_i \mathcal{P}(\mu_i|\bar{\lambda}, \sigma_i) \right) \cdot \mathcal{P}(\bar{\lambda}); \qquad (3)$$

and thus, the posterior reflects a probability distribution over the space of model parameter configurations. Equation (3) is the general representation of the posterior by which sampling statistics can be used to construct moments of the posterior distribution, which directly correspond to expectation values and associated errors for the model parameters.

The software used to apply the above techniques to MAST diagnostic data was written in the MINERVA framework. MINERVA is a JAVA-based framework, developed by Svensson and Werner, used to develop and analyse Bayesian graphical models in the context of diagnostic data analysis.[17]

III. PLASMA CURRENT AND DIAGNOSTIC MODELS FOR THE MAST EXPERIMENT

MAST is equipped with an array of precision diagnostics, including a MSE system[18] and over 100 routine equilibrium magnetic diagnostics. We solve for the toroidal current

profile utilising equilibrium magnetics (pickup coils and flux loops) and MSE diagnostics.

A. Plasma beam current model

Our development of plasma current profile tomography via Bayesian inference closely follows the work of Svensson and Werner on JET. In that work, the plasma is represented as a finite grid of toroidally extended, axisymmetric current beams, each with rectangular cross-section and carrying a uniform current density. Each plasma beam current in this model corresponds to a model parameter of the hypothesis in Bayes' Formula depicted in Eq. (1). In particular, this means that the posterior will reflect a PDF, whose domain dimension is equal to the dimension of current beam elements in our plasma model, and that it is a *distribution* of possible current values that will be inferred for each plasma current beam.

In MAST current tomography these beams are placed to fill-out a large region inside the first-wall, as depicted in Fig. 1. The ratio of first-wall to plasma volume is much larger on MAST than most other tokamaks. Indeed, MAST was not designed to have it is first wall tightly enclose the plasma volume, which makes the first-wall boundary a poor choice in selecting the extent of the plasma beam model. In particular, filling out the first-wall with beams would yield inference with greater uncertainties and requires stronger priors to exclude unphysical solutions corresponding to artificial current screening. This contrasts the situation in
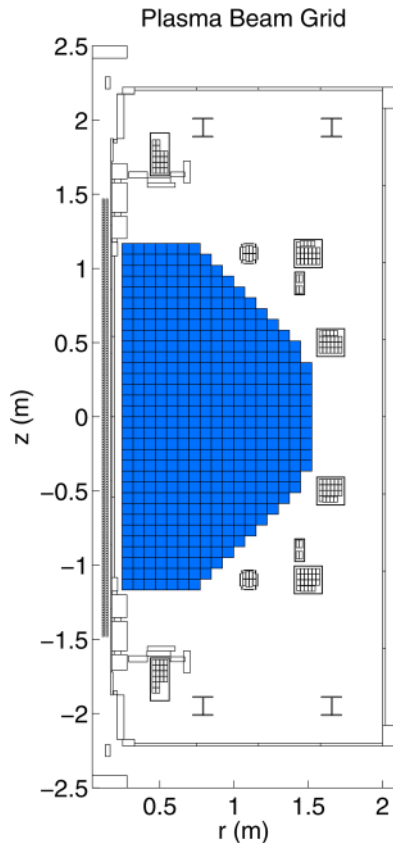

FIG. 1. (Color online) Plasma beam cross-sections fill out a large region inside the limiter that encompasses the plasma volume for MAST discharges.

Ref. 10, where JET—having a first-wall that tightly fits the plasma volume—had its first wall as a canonical choice for the cross-section beam extent. Thus, the beam extent for our MAST inferences was determined by randomly selecting shots and times between discharge #21000 and #24500 and picking a beam configuration that encompassed the LCFSs of all the tested shots. This initial survey was carried out by having the beams fill out the entire first-wall region and subsequently determining the corresponding LCFSs. It was then verified that the LCFSs calculated using the reduced set of beam currents were in a very good agreement with those coming from the initial survey (i.e., no visually discernible differences in the contours).

The key advantage to using a series of current beams with finite cross-section to model the toroidal plasma current is that the semi-analytic expressions for the corresponding magnetic field and vector potential have no singularities, even at points within the current beam itself.[19] Indeed, if one were to use filaments to model the toroidal plasma current, there would be many singular points in the calculated magnetic field within the plasma that would make subsequent flux-surface calculations difficult and prone to (possibly large) numerical errors.[10] A limitation of this current model is that poloidal currents are not described.

B. Gaussian process (GP) prior

To make the inference on plasma beam currents, a prior is used that effectively imposes a smoothness criterion across the 2D cross-section of plasma beams. The degree of smoothening is associated with a length-scale hyper-parameter, τ, whose selection is detailed in Sec. IV B. We use the theory of Gaussian processes to construct such a prior. The departure from Svensson and Werner's selection of the CAR prior for their JET analysis[10] is due to the fact that the results from using a GP prior have a much clearer interpretation as a smoothing assumption than that of a CAR prior. Indeed, the CAR prior enforces smoothening only locally[20] and is thus prone to noisy sample data. On the other hand, the GP priors enforce smoothening globally and generate samples that are not nearly as noisy as those coming from the CAR prior. The application of Gaussian processes to Bayesian analysis has been explored in several different contexts, including machine learning.[21] Indeed, in more recent works, Svensson has also made use of GP priors[22] in the context of current tomography. In this analysis, the theory is used to construct a covariance matrix for the prior distribution via a covariance function, $k(R_p, Z_p, R_q, Z_q)$, where (R_p, Z_p) and (R_q, Z_q) are the positions of two, arbitrary plasma currents I_p and I_q (respectively), taken from the collection of beams in our plasma model. The explicit relation between the covariance and k in our analysis is given by

$$\text{Cov}(I_p, I_q) = k(R_p, Z_p, R_q, Z_q) :$$
$$= \exp\left(-\frac{1}{2\tau^2}\left[(R_p - R_q)^2 + (Z_p - Z_q)^2\right]\right), \quad (4)$$

where the units on the covariance function values are A^2. As the positions of plasma beams are fixed, the beams in our model can be linearly indexed. Subsequently, Eq. (4) can be

used to construct a covariance matrix, $\bar{\bar{K}}$, associated with the vector of indexed plasma beams, $\bar{I}$. The specific prior distribution used in our analysis is taken to be a zero-mean Gaussian with $\bar{\bar{K}}$ as its covariance matrix, i.e.

$$\mathcal{P}(\bar{I}) \propto \exp\left(-\frac{1}{2}\bar{I}^T\bar{\bar{K}}^{-1}\bar{I}\right), \quad (5)$$

where the superscript T denotes the transpose. A Gaussian is a canonical choice for our prior, as normal distributions minimise the Shannon information over all probability distributions, when only the mean and covariance matrix is given (This is an elementary result, whose derivation can be seen in any number of textbooks including.[4]) From Eqs. (4) and (5), one can see that as $\tau \to \infty$, the smoothening constraint increases in strength, in a sense that beams have to have a great spatial separation to have a relatively low correlation. Specifically, as $\tau \to \infty$, the beam model reduces to a single current beam and $\mathcal{P}(I)$ becomes a δ-distribution centred at 0 current. In contrast, if $\tau \to 0$, the current beams become independent, with the variance of each current element becoming infinite. The choice of τ will have an affect on the errors propagated through calculations based on the inferred current tomography, with larger values of τ generally corresponding to larger error bars on results. This point will be discussed further in Sec. IV B.

One will notice that Eq. (4) indicates that our covariance matrix will have ones on its diagonal. While one can generalise Eq. (4) to have more hyper-parameters that control the diagonal values of the covariance matrix, this makes hyper-parameter selection a much more difficult problem, as the techniques in Sec. IV B would have to incorporate multidimensional scans that subsequently would require orders of magnitude more inversions being made to carry out the hyper-parameter determination. Thus, we use a simpler prior in this research and will explore more complex variations thereof in future works.

Finally, while Eq. (4) effectively enforces a smoothening effect across current beams, the strength of this effect is homogenous across the beam extent. This prior is appropriate in the current situation, as there are not enough internal diagnostic observations within the plasma to resolve spatial inhomogenaities of the τ hyper-parameter. Indeed, the MAST MSE system currently only provides internal magnetic field measurements in a highly localised region across the plasma mid-plane. However, this prior can easily be generalised to include parameterisations of τ as a function of cross-section coordinate. This type of prior would be appropriate to use when there are multiple internal magnetic field measurements for several different flux surfaces within the plasma volume.

C. Diagnostic forward models

In toroidally axisymmetric magnetic-confinement devices, there exists a direct relation between the poloidal flux, ψ, and the toroidal component of the plasma current vector. For convenience, we use the standard cylindrical coordinate system (R, Z, ϕ) with Z pointing upwards along the major

axis of the tokamak. Taking $D(R, Z)$ to denote a vertically oriented disk centred on the major axis of the tokamak at a height Z and radius R, the poloidal flux function, ψ, is commonly defined as

$$\psi(R, Z, \bar{I}) \equiv \int_{D(R,Z)} \bar{B}(R, Z, \bar{I}) \cdot \overline{dS}$$

$$= \oint_{\partial D(R,Z)} \bar{A}(R, Z, \bar{I}) \cdot \overline{d\ell} = 2\pi r A_\phi(R, Z, \bar{I}), \quad (6)$$

where $\bar{B}$ is the magnetic field and $\bar{A}$ is the vector potential. A_ϕ is related to the toroidal current I_ϕ via Biot-Savart's law

$$A_\phi(R, Z, \bar{I})$$

$$= \frac{\mu_0}{4\pi} \int \frac{I_\phi(R', Z')R'}{\sqrt{(Z - Z')^2 + R^2 + R'^2 - 2RR'\cos\phi'}} \, dR' dZ' d\phi'. \quad (7)$$

It is clear that when only considering I_ϕ, $\bar{A}$ will have zero R and Z components. Thus, taking the curl of Eq. (6) and using the definition of $\nabla \times \bar{A} = \bar{B}$, one has

$$B_R(R, Z, \bar{I}) = -\frac{1}{R}\frac{\partial \psi}{\partial Z}, \quad (8)$$

$$B_Z(R, Z, \bar{I}) = \frac{1}{R}\frac{\partial \psi}{\partial R}, \quad (9)$$

where B_R, B_Z, and B_ϕ indicate the components of the magnetic field in cylindrical coordinates. From Eqs. (6)–(9), it is clear that if I_ϕ is represented by an axisymmetric beam of rectangular cross section and uniform current distribution, then ψ, B_R, and B_Z will vary linearly with respect to the beam currents. The importance of this will become clear in the next two paragraphs.

The main diagnostics used to infer the current distribution are pickup coils, denoted $\overline{\mathcal{F}}_P$, full flux loops, $\overline{\mathcal{F}}_F$ and the polarisation angle γ of the emitted light from neutral excited species during neutral beam injection due to the MSE. The responses of these diagnostics to the collection of beam currents are given by

$$\overline{\mathcal{F}}_P(R, Z, \bar{I}; \theta) = B_R\cos(\theta) + B_Z\sin(\theta), \quad (10)$$

$$\overline{\mathcal{F}}_F(R, Z, \bar{I}) = 2\pi\psi, \quad (11)$$

$$\overline{\mathcal{F}}_M(R, Z, \bar{I}; \bar{A}) \equiv \tan\gamma(R, Z, \bar{I}; \bar{A}) = \frac{A_0 B_Z + A_1 B_R + A_2 B_\phi}{A_3 B_Z + A_4 B_R + A_5 B_\phi}, \quad (12)$$

where θ is the angle between a pickup coil's normal and the mid-plane and A_0, A_1, A_2, A_3, A_4, A_5 are constants for the particular MSE viewing geometry. Both $\overline{\mathcal{F}}_P$ and $\overline{\mathcal{F}}_F$ have a linear dependence on ψ, B_R, and B_Z and hence they are linearly dependent on the total current going through each toroidal current beam. The function $\tan\gamma$ has a non-linear dependence on B_R and B_Z; however, given that both neutral beams and MSE viewing optics all are on the mid-plane for MAST, A_1, A_2, A_3, A_4 are all approximately zero. Using

these properties, $\tan \gamma$ can be written as a linear combination of B_R and B_Z

$$\tan \gamma(r, z, \bar{I}; A) = \frac{A_0 B_Z + A_1 B_R}{A_5 B_\phi}. \tag{13}$$

Even though the A_1 term is small, it is still included in the MSE forward model, as the model is still linear and the added computational complexity is negligible. Since the current beam plasma model does not take into account poloidal currents, vacuum field values for B_ϕ are used in Eq. (13) to perform the current tomography calculations. Also, Eqs. (12) and (13) both ignore the effect of the radial electric field on $\tan \gamma$. Thus, given that $\overline{\mathcal{F}}_P(R, Z, \bar{I}; \theta)$, $\overline{\mathcal{F}}_F(R, Z, \bar{I})$, and $\overline{\mathcal{F}}_M(R, Z, \bar{I}; A)$ all have a linear dependence on the current flowing through $\bar{I}$, one may write out a generalised prediction vector $\bar{p}$ as

$$\bar{p} = \overline{\overline{M}} \bar{I} + \bar{C}, \tag{14}$$

where $\overline{\overline{M}}$ is the response matrix of the current vector $\bar{I}$ corresponding to all the plasma beams modelling the plasma and $\bar{C}$ represents the contribution to surface currents present in the vessel walls.

IV. POSTERIOR INFORMATION AND EXTREME OUTLIER OBSERVATIONS

As MAST represents and state-of-art device with many complex, inter-dependent engineering subsystems put under strenuous duty-cycles, diagnostic malfunctions do happen. The equilibrium magnetics on MAST are particularly hard to maintain as these diagnostics lie within the vacuum vessel itself. As a result, routine equilibrium reconstruction on MAST requires that at least 20 of these diagnostics be disabled or have their signal rescaled from their calibration parameters. There is no quantifiable algorithm that guides the implementation of this process outside of reworking signals to the point where the equilibrium reconstruction is consistent with observations of other, independent diagnostics and MAST standard operational scenarios. In this section, we present an algorithm which is able to robustly and consistently pinpoint and remove malfunctioning diagnostics, using methods from information theory.

We build on the model in Sec. III, developed by Svensson and Werner in Ref. 10, to include an algorithm to identify and remove extreme outlier diagnostic observations, within particular diagnostic subsystems, from the current tomography inference. Extreme outlier observations are those observations coming from diagnostics that our analysis fails to model correctly. This situation may be due to an unidentified diagnostic malfunction/failure or the diagnostic falling out of calibration. Although attempts have been made to accurately model the many systematic errors associated with the magnetics diagnostics, it is difficult (in practice) to ascertain if a particular pickup coil or flux loop or pickup coil is malfunctioning or has somehow fallen out of calibration, as the information they provide is non-localised relative to the plasma. The MSE system on MAST is fortunately easier to maintain and diagnose, as its primary components lie outside

of the MAST vacuum vessel; this is not the case for magnetics. Moreover, MSE provides information that is most sensitive to a highly localised region inside the plasma, and thus, it is easy to identify malfunctioning channels. Given this situation, the below procedure for identifying malfunctioning diagnostics is done without the use of MSE data.

Removal of extreme outliers is necessary on MAST to ensure that the vast majority (over 95%, to allow for sheath currents) of current flows only in the plasma volume, thereby ensuring an accurate current profile inference that is consistent with all diagnostic observations. The model is flexible enough to infer toroidal edge current just outside the LCFS. The motivation is that inference of plasma currents using the model in Sect. III is not correct, with large currents far outside the LCFS (beyond uncertainties), if diagnostics producing extremely outlying measurements are included. We introduce a new method to effectively identify and remove these measurements and thereby remove solutions with more than 5% of the current flowing outside the last closed flux surface.

A. Evidence cross-validation

The Bayesian notion of evidence was introduced in Sec. II as a measure of consistency between observations with each other and the prior knowledge, when likelihood distributions are fixed, modulo translation. Diagnostics in fusion experiments often have their calibrations cross-validated against other diagnostics measuring related quantities in the experiment. By using evidence as a tool in a generalised notion of observation cross-checking, we have developed a new technique, which we label "evidence cross-validation" (ECV). The philosophy behind ECV is to rank diagnostic observation according to how consistent they are with all other measurements against a given, fixed prior and set of given observational errors. The diagnostic observations are then removed from the analysis, with the most inconsistent diagnostics being removed first. As diagnostics are removed, the resulting posterior evidence is recorded. Diagnostics are only removed up to the point where the evidence is maximised during the procedure.

The technique proceeds as follows:

1. A baseline posterior, labeled $\mathcal{P}_0$ is calculated with all diagnostics using Eq. (3).
2. One diagnostic observation, label it o_i, is removed from the inference, a new posterior is calculated, call it $\mathcal{P}_i$, and the associated log-evidence $\mathcal{E}_i := \ln(\mathcal{P}(D))$ is computed and associated with o_i. We repeat this process for all diagnostics.
3. The diagnostic with the lowest associated log-evidence is then removed from future inversions, and a new baseline posterior calculated without the most conflicting diagnostic observation being present. The evidence of this new posterior is then recorded and associated with the removed diagnostic.
4. Steps (1)–(3) are repeated to generate a curve of posterior evidence versus the number of diagnostics removed.
5. Diagnostics are removed such that the posterior evidence recorded in Step (3) is maximised.

Computationally, calculating the evidence of high-dimensional posteriors is an extremely difficult problem.[4] To address this, we consider likelihood distributions to have the following form:

$$\mathcal{P}(\bar{\mu}|\bar{\lambda}|, \bar{\bar{\Sigma}})$$
$$= \frac{1}{(2\pi)^{N_o/2}|\bar{\bar{\Sigma}}|^{1/2}} \exp\left[-\frac{1}{2}(\bar{\mathcal{F}}(\bar{\lambda}) - \bar{\mu})^T \bar{\bar{\Sigma}}^{-1}(\bar{\mathcal{F}}(\bar{\lambda}) - \bar{\mu}) \right],$$

(15)

where $\bar{\bar{\Sigma}}$ is the covariance matrix associated with the given diagnostic errors; N_o is the number of observations; and $\bar{\mathcal{F}}$ is the forward model associated with the diagnostic observation. Since the forward models being employed in this research are all linear, Eq. (15) shows that the associated likelihood distributions will all be multivariate Gaussian distributions. This subsequently implies that the posterior distribution itself will also be a multivariate Gaussian, as the GP prior is a multivariate Gaussian distribution. Thus, the evidence associated with the posterior distribution will have an analytic expression based on the covariance matrix of the posterior. It is this computational simplification that makes this procedure computationally tractable, able to be completed on the order of minutes on an average desktop computer. For the results presented here, the ECV procedure calculates the evidence for up to 3850 diagnostic configurations. The entire ECV procedure takes 50-60 min to complete on a 2.9 GHz desktop CPU running in a serial implementation. A second motivation for selecting Gaussian PDFs for our likelihoods comes from the fact that Gaussian distributions also minimises the amount of information contained in a distribution when only the mean and variance of a distribution are known.[4] Thus, picking a multivariate Gaussian distribution is a canonical choice for the likelihood, in the current scenario.

While the ECV ordering of diagnostics is relatively robust against various choices of τ in the prior, some ranking variations can occur. This is discussed at the end of Sec. IV B.

B. Plasma current smoothening and τ selection

The parameter τ is a pre-set value that determines the degree of smoothening enforced by the prior across the collection of current beams. As the GP prior represents a zero-mean Gaussian distribution, making τ too large (i.e., enforcing more smoothening) will reduce the total plasma current inferred in our analysis. On the other hand, making τ too small allows for unphysical "current screening" to occur that infers currents only in plasma beams that are closest to the positions where diagnostic observations are made. While it is possible to treat τ as a nuisance parameter to be marginalised (i.e., integrated out), this would make the GP prior no longer a multivariate Gaussian distribution, relative to the free-parameters of the problem. The subsequent sampling and evidence calculation on the posterior would require a completely different set of algorithms, which would require substantially more computational effort than the present analysis employs; and thus, this calculation is beyond the scope of the research presented here.

In light of the above statements, our analysis is carried out using a single τ, which is chosen to maximise the evidence associated with the posterior, when using data from both magnetics and MSE. Indeed, MSE data will significantly impact which τ maximises the posterior evidence. This method of prior parameter selection in Bayesian inference has been very successful in the context of machine learning.[21] While it is possible that the evidence make be maximised only in the limit of $\tau \rightarrow 0$, a distinct, non-zero global maximum has always been observed in the dozens of shots analysed using this procedure.

As mentioned previously, the ECV procedure can have its result affected by different selections of τ. However, we have not seen the results of the ECV change (i.e., ranking or number of diagnostics removed) for variations of τ under 1 m on MAST discharge data. For τ shifts of more than 1 m, the ranking of a particular diagnostic has been seen to shift by up to one place with the overall number of removed diagnostics varying by up to three. Given the robustness of ECV results for τ variances under 1 m, the ECV and τ-selection procedures are wholly compatible in a sense that by iterating the ECV and τ-selection procedures, one will converge to a particular τ and collection of magnetic diagnostics that have been removed. Practically, τ is usually about 0.25 m for MAST discharges with MSE data. If using this as an initial guess, the iteration of these processes will find the optimal configuration of diagnostics and τ after one iteration.

V. RESULTS

To demonstrate the Bayesian inference of MAST current profiles, we analysed MAST discharge #22254 at 350 ms and #24600 at 330 ms. Discharge #22254 is a strongly shaped double-null diverter (DND) deuterium plasma that has been heated with 3.1 MW of neutral beam power. The time at 350 ms corresponds to highest poloidal beta of the discharge where the plasma is in H-mode, just before neutral beam turn-off. Discharge #24600 at 330 ms is also a DND plasma heated with 3 MW of neutral beam power but is in L-mode. Bayesian inference of current profiles was calculated using approximately 76 pickup coil, 24 flux loop, and 31 MSE observations.

First, we tested the ECV procedure's ability to pick out failing diagnostics by randomly selecting 20% of the diagnostics and rescaling their signals by a randomly selected factor of either 0.5 or 1.5. These scalings were chosen, as pickup coils on MAST are actually installed in pairs (i.e., having the same cross-sectional coordinate), with their signals being averaged to produce the signal recorded on the scheduler. Thus, if one of these paired pickup coils fail, the signals are scaled down by 0.5. If MAST personal recognises this failure, they will rescale to compensate. However, if the failed diagnostic comes back online later, then the average signal will be scaled by 1.5 times its true value. Applying this artificial rescaling procedure both artificial (based on the configuration depicted in #22254 at 350 ms) and real (#22254 at 350 ms and #24600 at 330 ms), the ECV procedure was able to consistently identify all diagnostics that had underwent this random rescaling. Moreover, the ECV procedure was able to pinpoint these diagnostics, even when the
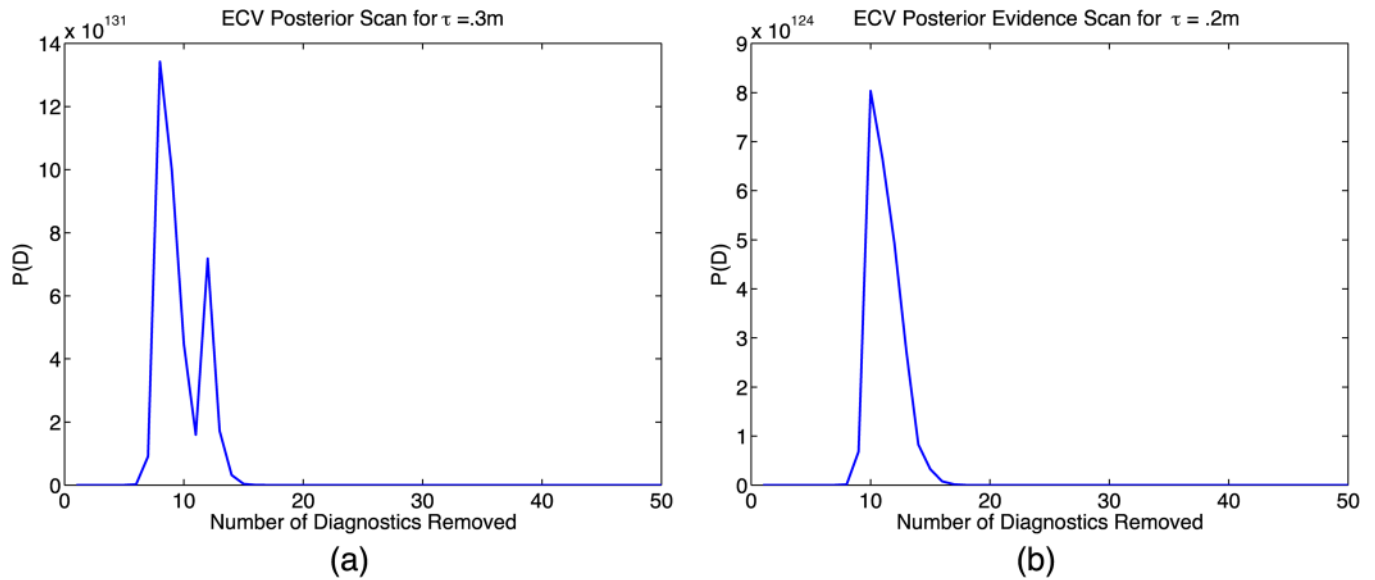
FIG. 2. (Color online) Evidence values for diagnostics removed via the ECV procedure. (a) Corresponds to shot #22254 at 350 ms, with $\tau = 0.3$m; and (b) corresponds to shot #24600 at 330 ms, with $\tau = 0.2$m.

artificial signal corruption leads to inferences that significantly altered the shape of the plasma (due to the random selection corrupting many diagnostics around the X point of the plasma).

Figure 2 shows the evidence corresponding to the removed diagnostic during successive iterations of the ECV procedure for two different MAST discharges. Specifically, these plots correspond to the final ECV run, in the ECV/τ-selection iteration described in Sec. IV. It is clear from Figure 2, that the evidence obtains a sharp global maximum when nine to ten diagnostics are removed. In general, it is not necessary to remove all the diagnostics leading up to the peak of the evidence to get the existence of a physical solution (i.e., little plasma current flowing outside of the LCFS beyond uncertainties). However, removing diagnostics

to maximise the evidence, according to the methods outlined above, has been seen empirically as a sufficient condition to guarantee the inference of a physical plasma across dozens of MAST discharge snapshots, without exception.

Corresponding to Fig. 2 in the ECV/τ-selection iteration is the final τ-scan plot in Fig. 3. This figure clearly indicates that $\tau = 0.3$m maximises the evidence of the posterior inference, when nine diagnostics are removed via the ECV procedure for discharge 22254 at 350 ms. Likewise, $\tau = .2$m maximises the evidence for discharge 24600 at 330 ms. The well-defined maxima shown in Figs. 2 and 3 have been consistently seen qualitatively, when analysing dozens of different MAST discharge snapshots. Across all discharges analysed with MSE data, the results of the ECV/τ-selection iteration have produced a range of τ between 0.21 m and
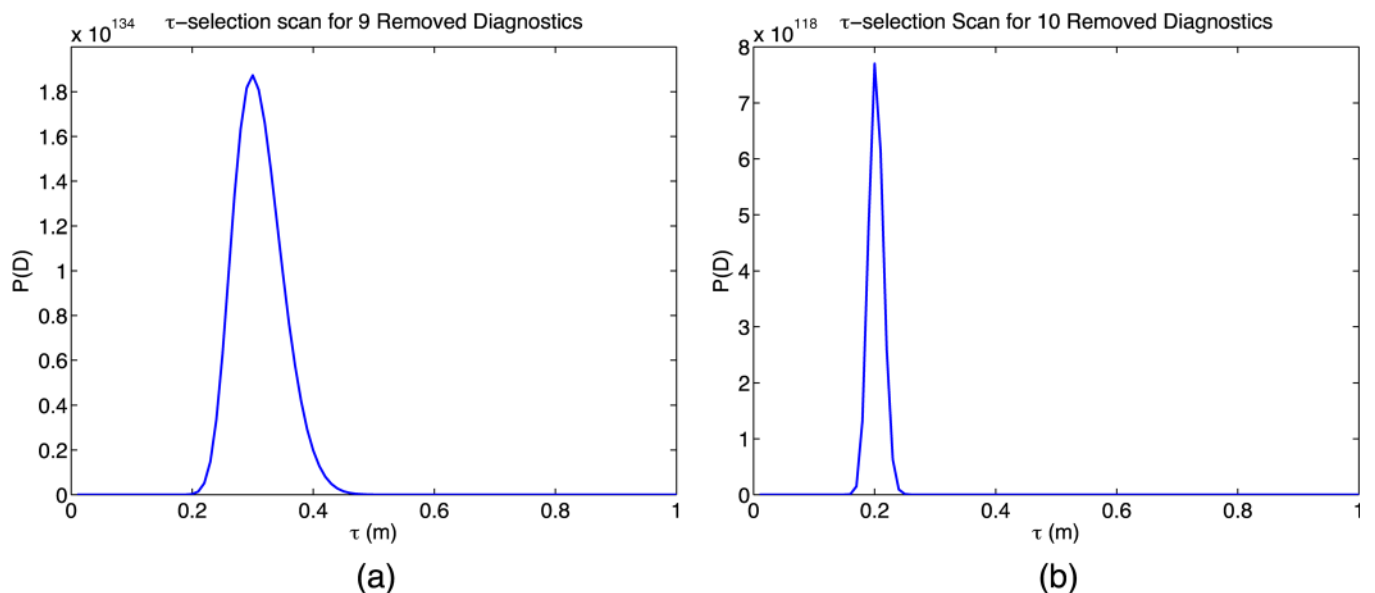

FIG. 3. (Color online) Calculated evidence for scanned values of τ. (a) Corresponds to shot #22254 at 350 ms with nine magnetic diagnostics removed via ECV and (b) to shot #24600 at 330 ms with 10 magnetic diagnostics removed.
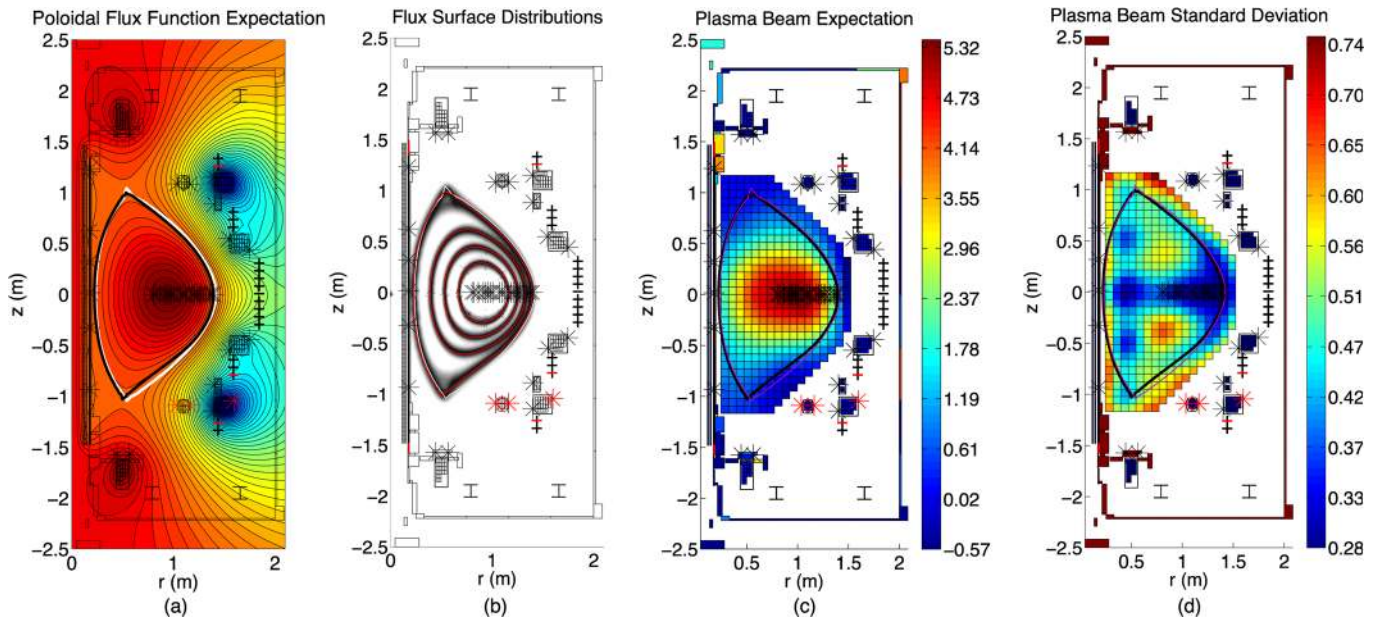
FIG. 4. (Color online) Current profile data inferred for MAST shot #22254 at 350 ms using pickup coils, flux loops, and MSE. The inferred LCFS is drawn in black on (a), (c), and (d); the EFIT++ LCFS is under-layered on all figures in white/thin solid lines for comparison only. Flux loop locations are indicated by stars outside the plasma region; position and orientation of pickup coils are indicated via heavy bars on the out-board edge of the first wall and as a vertically oriented column line along the solenoid; and MSE observation positions are indicated by the stars across the mid-plane inside the plasma region. Non-black diagnostics are those which have been removed by the ECV procedure. (a) Poloidal flux-surfaces corresponding to the expectation configuration of plasma beams. (b) Line-density plot showing the uncertainties in the 0.25, 0.50, 0.75, and 0.99 normalised flux-surfaces as calculated from 200 posterior samples; the solid lines indicate the expectation of these surfaces. (c) Shows the expected plasma beam configuration, the color bar indicating individual beam currents in kA. (d) Indicates the standard deviation of plasma beam currents in kA.

0.32 m with 8 to 12 diagnostics being removed. The reason for the discrepancy between the peak evidence values between Figs. 2 and 3 is due to the fact that the ECV procedure does not use MSE data, as detailed in Sec. IV A.

Figure 4 contains plots of the expected current beam profile and of the flux-surface distribution, calculated by applying a Ampere's law to current beam posterior samples for discharge #22254. As the poloidal flux surface geometry is very similar for both shots, Figure 5 only shows the current beam plots for #24600. The EFIT++ LCFS has been included only for comparison. The key feature of Figure 4 is that errors on flux-surface can be ascertained from the Bayesian inference of current-beams; it is not possible to extract such errors when using conventional Grad-Shafranov reconstruction techniques, in a mathematically rigorous way (stemming from the fact that equilibrium reconstruction requires solving the GS PDE with only a finite number of diagnostic measurements and not a fully specified boundary condition). This is of particular importance, as many diagnostics require data from the flux-surface geometry in order for their observations to be correctly interpreted. The results of Figure 4 are inferred from the combined observations coming from magnetics and MSE.

In addition to giving errors on flux-surfaces, Bayesian current tomography can also yield errors on any quantities that can be calculated from a toroidal current profile. Figure 6 shows statistics on the *q*-profile and toroidal current density on the mid-plane, calculated from samples of the current beam posterior for discharge #22254 at 350 ms. Very similar results come from the analysis of #24600 at 330 ms. The *q*-profile is important for understanding and predicting

plasma stability. The plot of the mid-plane current is included to illustrate the fact that the Bayesian current tomography yields a probability distribution for the 2D toroidal current profile; and thus it is possible to extract errors on any collection of beam currents. The main reason for the *q*-profile being so well resolved relative to the mid-plane current is due to the highly localised MSE observations almost
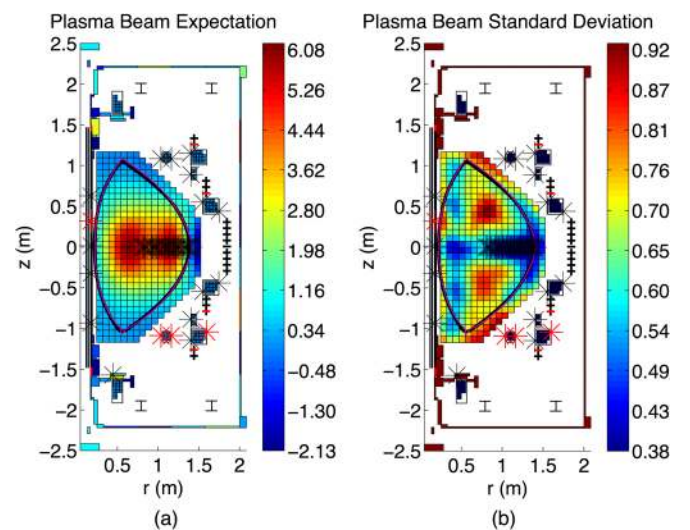

FIG. 5. (Color online) Current profile data inferred for MAST shot #24600 at 330 ms using magnetic and MSE with LCFS curves and diagnostic markers depicted as they were in Figure 4. (a) Shows the expected plasma beam configuration, the color bar indicating individual beam currents in kA. (b) Indicates the standard deviation of plasma beam currents in kA. Poloidal flux statistics are difficult to distinguish from Figures 4(a) and 4(b) and thus have not been included.
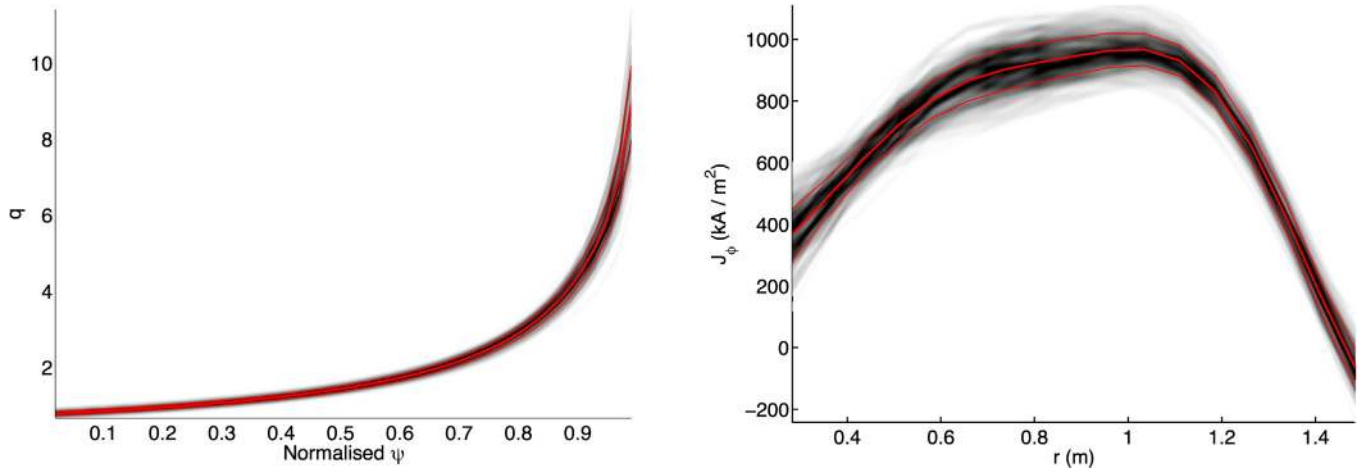
FIG. 6. (Color online) Data inferred for MAST shot #22254 at 350 ms using pickup coils, flux loops, and MSE. In both plots, the heavy lines indicates the expected value, with thin lines indicate the first standard deviation away from the mean; and the gray shading shows the profile distribution, as calculated from 200 samples of the current beam posterior. (a) q-profile as a function of normalised flux. (b) Plasma current density across the mid-plane as a function of radial coordinate.

directly corresponding to the pitch-angle of the magnetic field. Such plots help to illustrate how uncertainties in the current tomography propagate to errors in the flux-surfaces and q-profiles.

The total plasma current for #22254 at 350 ms was computed to be 808.70 ± 30.44 kA, as compared to the 822.35 kA measured from the Rogowski coil. The log-evidence, $\ln(\mathcal{P}(D))$, was computed to be 323.88, which corresponds to the peak in Fig. 3. The inferred total plasma current for #24600 at 330 ms was found to be 819.28 ± 32.81 kA and is within uncertainty of the measured total plasma current of 841.73 kA.

In addition to the above analysis, several other shots and times have been analysed using the techniques outlined in this paper, all of which yield very similar results, figures and agreement with the EFIT++ LCFS, as the ones presented in this section.

VI. CONCLUSIONS

We have demonstrated the successful implementation and extension of the work of Svensson *et al.* to the MAST experiment. In particular, a plasma beam current model has been used in conjunction with Bayesian analytic techniques to infer the equilibrium 2D toroidal current profile for a high-performance MAST discharge. The ECV technique was developed to identify and remove extreme outlier observations associated with diagnostics not correctly modelled in the analysis. Using this, the most likely flux-surface geometry was computed, along with uncertainties in various specific flux-surfaces. Moreover, the corresponding q-profile for the discharge was also calculated, along with uncertainty. A major advantage of the techniques discussed is that the inference involves analysing and sampling from a Gaussian posterior distribution, which are computations that can be efficiently done using standard algorithms, even for distributions over high-dimensional domains. The use, however, of non-analytic distribution functions makes Bayesian inference computationally demanding, in both sampling and finding

the maximum of corresponding posterior distributions. The treatment of such statistics is outside the scope of the work and will be treated elsewhere.

While only nine diagnostics were removed via the ECV procedure, physical current profiles are consistently inferred removing anywhere from ~5% to ~20% of all magnetics diagnostics. In some cases, physical inferences can be still made, even if removing over 20% of the diagnostics. The primary advantage of the ECV procedure is that the inference of toroidal current beams pre-supposes no detailed knowledge of the plasma, and thus is implemented in a completely automated way. By caching results from these procedures, one might design real time systems that are actively able to correct for a diagnostic failing or falling out of calibration.

The present analysis does not account for poloidal currents nor the radial electric field, and these are present research foci. In MAST, modelling has suggested the correction to the q-profile is of order 10% for these effects. Preliminary modelling of poloidal currents through correction to the toroidal flux functions has been recently presented.[23] Moreover, the discrepancies between the computed and EFIT++ LCFS, along with the total plasma current discrepancy, may be resolved by making a more detailed analysis of outlier errors via modification of the likelihood distributions. This analysis creates a non-Gaussian posterior, which requires new computational techniques to handle. While such methods may enable better handling of outlier data and provide support for non-Gaussian likelihoods, they entail significantly more computation to perform and are only suited for post-analysis studies. This is another focus of current research.

[1] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, UK, 2003).

[2] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 2nd ed. (Chapman & Hall/CRC, New York, 2004).

[3] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica Support* (Cambridge University Press, Cambridge, 2005).

[4] D. S. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, 2nd ed. (Oxford University Press, Oxford, 2006).

[5] G. A. Cottrell, "Maximum entropy and plasma physics," in *Maximum Entropy in Action: A Collection of Expository Essays*, edited by B. Buck and V. A. Macaulay (Oxford University Press, Oxford, UK, 1991).

[6] A. P. Millar, D. C. McDonald, and D. A. Diver, Plasma Phys. Controlled Fusion **42**, 337 (2000).

[7] R. Fischer, A. Dinklage, and E. Pasch, Plasma Phys. Controlled Fusion **45**, 1095 (2003).

[8] A. Dinklage, R. Fischer, and J. Svensson, Fusion Sci. Technol. **46**, 355 (2004).

[9] J. Svensson, A. Dinklage, J. Geiger, A. Werner, and R. Fischer, Rev. Sci. Instrum. **75**, 4219 (2004).

[10] J. Svensson and A. Werner, Plasma Phys. Controlled Fusion **50**, 085002 (2008).

[11] O. Ford, J. Svensson, A. Boboc, and D. C. McDonald, Rev. Sci. Instrum. **79**, 10F324 (2008).

[12] M. Reginatto and A. Zimbal, Rev. Sci. Instrum. **79**, 023505 (2008).

[13] M. J. Hole, G. T. von Nessi, J. Bertram, J. Svensson, L. C. Appel, B. D. Blackwell, R. L. Dewar, and J. Howard, "Model data fusion: Developing Bayesian inversion to constrain equilibrium and mode structure," in *Proceedings of the 7th General Scientific Assembly of the Asia Plasma and Fusion Association (APFA2009)* (Aomori, Japan, 2009).

[14] L. C. Appel, G. T. A. Huysmans, L. L. Lao, P. J. McCarthy, D. Muir, E. R. Solano, D. Taylor, and W. Zwingmann, "A unified approach to equilibrium reconstruction," in *33rd EPS Conference on Plasma Physics* (Rome, Italy, 2006), P-2.184.

[15] J. Wesson, *Tokamaks*, 3rd ed. (Oxford University Press, Oxford, UK, 2004).

[16] I. H. Hutchinson, *Principles of Plasma Diagnostics*, 2nd ed. (Cambridge University Press, Cambridge, 2002).

[17] J. Svensson and A. Werner, "Large scale Bayesian data analysis for nuclear fusion experiments," in *2007 IEEE International Symposium on Intelligent Signal Processing* (IEEE, Alcala de Henares, Spain, 2007), pp. 1–6.

[18] M. F. M. De Bock, N. J. Conway, M. J. Walsh, P. G. Carolan, and N. C. Hawkes, Rev. Sci. Instrum. **79**, 10F524 (2008).

[19] L. K. Urankar, IEEE Trans. Magn. **18**, 1860 (1982).

[20] J. Besag, J. R. Stat. Soc. Ser. B (Methodol.) **36**, 192 (1974).

[21] C. E. Rasmussen and K. I. Williams, *Gaussian Processes in Machine Learning*, 2nd ed. (MIT, Cambridge, MA, 2006).

[22] J. Svensson, "Non-parametric tomography using Gaussian processes," IEEE Trans. Image Process. (submitted).

[23] L. C. Appel, G. T. von Nessi, M. J. Hole, and J. Svensson, "Bayesian inference applied to magnetic equilibrium on MAST," in *Europhysics Conference Abstracts* (Dublin, Ireland, 2010).