Incorporating magnetic equilibrium information in Gaussian process tomography for soft X-ray spectroscopy at WEST [a)]

T. Wang[1,2,3,b)], D. Mazon[2], J. Svensson[4], D. Li[1], A. Jardin[2] and G. Verdoolaege[3,5]

[1]*Southwestern Institute for Physics, CNNC, C-610200 Chengdu, China*
[2]*Institute for Magnetic Fusion Research, CEA, F-13115 Saint-Paul-lez-Durance, France*
[3]*Department of Applied Physics, Ghent University, B-9000 Ghent, Belgium*
[4]*Max-Planck-Institut für Plasmaphysik, D-17491 Greifswald, Germany*
[5]*Laboratory for Plasma Physics, Royal Military Academy (LPP-ERM/KMS), B-1000 Brussels, Belgium*

Gaussian process tomography[1,2] (GPT) is a recently developed tomography method applied earlier to soft X-ray (SXR) spectroscopy on WEST. The short execution time of the algorithm makes GPT an important candidate for providing real-time information on impurity transport and for fast MHD control. In earlier work, GPT has shown its flexibility by providing good reconstruction results without background information about the magnetic equilibrium. On the other hand, information about the magnetic flux surface geometry can in general be useful for additional regularization of the solution. In this paper, we develop a way to take into account the equilibrium information, by constructing a covariance matrix of the prior Gaussian process depending on the flux surface geometry. The GPT method is validated using synthetic SXR emissivity profiles relevant to WEST plasmas, and compares favorably with the classical algorithm based on minimization of the Fisher information.

I. INTRODUCTION

The tokamak WEST – for Tungsten (W) Environment in Steady-State Tokamak – has started operating from the end of 2017 as a testbed for the ITER divertor components in long pulse operation. In this context, radiative cooling of heavy impurities like W is a critical issue for the plasma core performance. Thus, reliable tools are required to monitor the local impurity density and avoid W accumulation. Soft X-ray (SXR) spectroscopy is a diagnostic technique that has the potential to deliver valuable information in this respect. This diagnostic can provide very good temporal resolution (up to 1 MHz), which is sufficient for MHD activity and impurity transport studies. Particularly, the plasma is optically thin for SXR radiation in the range from 1 keV to 15 keV, which makes SXR tomography a powerful tool for studying core plasma physics.[3] The tomography problem essentially involves the prediction of high-dimensional physics parameters by inversion of a limited number of measurements. This is an ill-posed problem, as the number of measurements (SXR line integrals) is always lower than the number of unknowns (emissivity value in each cell). There exists a variety of reconstruction algorithms to solve the inversion problem. In the past, various tomographic reconstruction techniques have been applied to SXR, such as the Cormack method,[4] the maximum entropy method,[5] the minimum Fisher

information method,[6] etc. Particularly the minimum Fisher information technique has been widely adopted in the fusion community. This reconstruction method involves χ^2 optimization, regularized by the Fisher information. Intuitively, the goal is to find the least complex solution that is compatible with the data. The method is often implemented, e.g. on Tore Supra and WEST,[7] using additional information concerning the location of the equilibrium magnetic flux surfaces, obtained from magnetic measurements. This paper is focused on a recently developed tomography technique, namely Gaussian process tomography (GPT) applied to the WEST SXR spectroscopy setup. In previous work[8], we have showed the excellent performance of this method even without relying on information regarding the magnetic equilibrium. This is an important asset of GPT in case the equilibrium information is unreliable or in special situations such as in the presence of poloidally asymmetric impurity concentrations. On the other hand, in routine applications of SXR tomography, equilibrium information can be very useful for additional regularization of the tomographic inversion. In this paper, we incorporate for the first time magnetic equilibrium information into the GPT technique. Combined with the computational efficiency of the method, this makes GPT an excellent tool for routine application in real time.

II. GAUSSIAN PROCESS TOMOGRAPHY

A. WEST soft X-ray diagnostic system

The WEST SXR diagnostic system is presently being commissioned with two triple-gas electron multiplier (GEM) cameras, located in the same toroidal cross-section

to allow tomographic reconstruction.[9] The triple-GEM detector is based on photoionization in a flowing gas mixture subjected to an electric field and enclosed in Mylar foil. As depicted in Fig. 1, photoelectrons are produced in the first conversion layer and drift towards a perforated copper-clad Kapton foil. A voltage is applied to the foil, causing electron avalanching, hence amplifying the detector signal. The process is repeated in two successive GEM foils, followed by charge collection on the anode strips (pixels). The GEM detectors work in photon counting mode with energy discrimination.
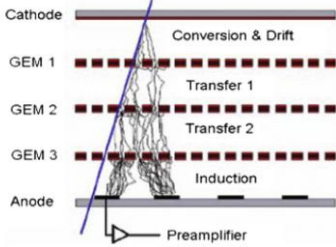

Fig. 1. Schematic of the triple GEM, which implements three copper-clad Kapton foils. After a signal amplification, the electrons are collected on the anode strips.

Compared to photodiode detectors, the GEM concept separates the regions where photoionization, amplification and detection takes place. An additional advantage is that electrons travel fast to the anode in about 50 ns, while the GEM holes are ion-free after ca. 1 μs. Therefore, the system has high-rate capabilities of ca. 10^6-10^7 ph.s^{-1}.mm^{-2}. Further advantages of GEM detectors are their compactness, good spatial and temporal resolution and good neutron-resistance.
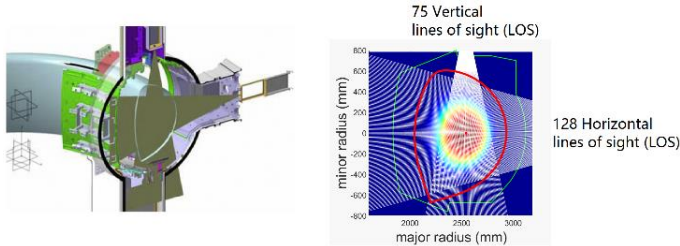

Fig. 2. Tomographic capabilities of the WEST SXR system based on GEM detectors. The horizontal camera views along 128 lines-of-sight. The vertical camera is inside the vertical port and is coupled to 75 sight lines.

As shown in Fig. 2, one of the cameras of the WEST SXR diagnostic views along the horizontal direction through 128 lines-of-sight (LOS) from the low-field-side to the high-field-side. The other camera is located at the top of the device, viewing downwards along 75 lines-of-sight. Hence, the majority of the core plasma region is covered with a good spatial resolution (~1 cm in the equatorial plane). The GEM system can provide a temporal resolution for real-time analysis of 1 kHz (five energy windows within an energy range of 2-15 keV), while the full spectrum will be stored off-line at a rate of up to 10 kHz, with a view to more detailed analysis and study of fast plasma phenomena. This setup provides good capabilities for studying fast MHD activity and impurity transport, in particular for tungsten transport.

A common and simple approach to discretize the emissivity field in a poloidal cross-section uses a square grid. We here impose a 100×100 grid comprised of square cells with a dimension of 16 mm $\times$ 16 mm. The SXR emissivity within each pixel can reasonably be assumed to be constant, so the SXR line-integrated emissivity $\bar{d}_m$ along m viewing chords can be written in the following matrix form:

$$\bar{d}_m = \bar{\bar{R}}_{m \times n} \cdot \bar{E}_n + \bar{\varepsilon}. \tag{1}$$

Here, $\bar{E}_n$ is the unknown vector of local emissivities in $n = 10^4$ cells, while $\bar{\bar{R}}$ is the geometry matrix, whose elements R_{ij} represent the physical length of chord i through cell j. $\bar{\varepsilon}$ denotes an error term to account for measurement uncertainty, which is usually limited to statistical errors only.

B. Concepts of Bayesian probability theory

In fusion, as in many scientific activities, we do not have direct access to the physical properties of the system under study, in this case the plasma. The experimentalist must instead devise a measurement technique for a diagnostic, which returns numbers (usually voltages) that are related to the quantity of interest. If the properties of the physical system were known precisely, together with the full details of the measurement process, the corresponding measurement values could be computed straightforwardly. This would require the measurement process to be encoded in a mathematical model: the forward model. In practice, a forward model is a simplified mathematical representation of the measurement process, possibly focusing on one important aspect, such as the model in Eq. (1). In this paper, we use Bayesian probability theory for reconstructing a probability distribution for the quantities of interest (here $\bar{E}_n$), starting from a distribution of the measurements (here the line integrals $\bar{d}_m$), through the forward model (1). Using Bayes' formulae, the SXR tomography problem can be expressed in the following form:

$$p(\bar{E}_n|\bar{d}_m) = \frac{p(\bar{d}_m|\bar{E}_n)\,p(\bar{E}_n)}{p(\bar{d}_m)} \sim p(\bar{d}_m|\bar{E}_n)\,p(\bar{E}_n), \tag{2}$$

$$p(\bar{d}_m) = \int p(\bar{d}_m, \bar{E}_n)d\bar{E}_n = \int p(\bar{d}_m|\bar{E}_n)\,p(\bar{E}_n)d\bar{E}_n. \tag{3}$$

$\bar{E}_n$ *Vector of emissivity values in all n pixels*

$\bar{d}_m$ *Vector of m line-integrated GEM array measurements*

In Eq. (2), the likelihood term $p(\bar{d}_m|\bar{E}_n)$ measures the mismatch between the measured line integrals $\bar{d}_m$ and their predictions by the forward model, under the assumption of some emissivity field $\bar{E}_n$. The evidence (marginal likelihood) $p(\bar{d}_m)$ depends on the particular forward measurement model, which we will assume to be fixed. Therefore, it can be considered as a normalization factor, independent of the emissivity. The posterior probability distribution $p(\bar{E}_n|\bar{d}_m)$ quantifies our uncertainty on the estimated emissivity field, given the model, prior knowledge and the measured data. Thus, Bayesian inference

yields probabilities for all possible results consistent with the model.

C. Gaussian process framework

Gaussian process tomography (GPT) is a new technique whereby the prior distribution regularizes the tomographic reconstruction process, by imposing a smoothness level dictated by the correlation between pixels. Briefly, a Gaussian process is a generalization of the multivariate normal (Gaussian) distribution to a function space. It is described by a mean function $\bar{\mu}$ and a covariance function $\bar{\bar{\Sigma}}$, where $GP \sim \mathcal{N}(\bar{\mu}, \bar{\bar{\Sigma}})$. Being nonparametric, Gaussian process tomography does not assume any functional form for the emissivity field, hence leaving a lot of flexibility. Instead, the emissivity field is regularized through the covariance matrix of the Gaussian process. Specifically, GPT assumes that the prior joint distribution of the emissivity in the n cells with coordinates $\vec{r}_i$ is multivariate Gaussian with covariance matrix $\bar{\bar{\Sigma}}_E$ given by:

$$\bar{\bar{\Sigma}}_E = \begin{pmatrix} k(\vec{r}_1, \vec{r}_1) & \cdots & k(\vec{r}_1, \vec{r}_n) \\ \vdots & \ddots & \vdots \\ k(\vec{r}_n, \vec{r}_1) & \cdots & k(\vec{r}_n, \vec{r}_n) \end{pmatrix}. \qquad (4)$$

Here, $k(\vec{r}_i, \vec{r}_j) = cov[E(\vec{r}_i), E(\vec{r}_j)]$, with $E(\vec{r}_i) = E_i$ the emissivity in pixel i, is the *covariance kernel function*, for which we choose the common squared-exponential form:

$$k_{SE} = \sigma_f{}^2 exp\left(-\left(\frac{d_\perp{}^2}{2\sigma_{l\perp}{}^2} + \frac{d_\parallel{}^2}{2\sigma_{l\parallel}{}^2}\right)\right). \qquad (5)$$

The idea is that the correlation of the emissivity among any two pixels depends on the distance between those pixels, adapted here to the flux surface geometry obtained from an equilibrium reconstruction. In particular, $\boldsymbol{d}_\parallel$ represents the distance between pixel i and pixel j along a magnetic flux surface, and $\boldsymbol{d}_\perp$ is the perpendicular distance between the surfaces on which pixel i and j are located. This is where the magnetic equilibrium information is implemented in our algorithm. An example of a distance map is given in Fig. 3.
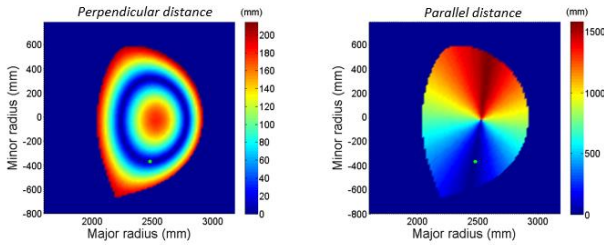

Fig. 3. Example of a distance map w.r.t. to a reference pixel indicated by the green point. The color maps represent the perpendicular distance (left) and parallel distance (right) between the reference pixel and the other pixels.

In turn, the kernel function depends on three parameters (summarized by the vector $\bar{\theta}$ in the remainder): the signal standard deviation σ_f, and the perpendicular and parallel characteristic length scales $\sigma_{l\perp}$ and $\sigma_{l\parallel}$. In Bayesian terminology, these parameters of the prior distribution are called *hyperparameters* and in this case they determine the smoothness of the emissivity field. In principle, the hyperparameters can be marginalized from the problem (i.e.

integrated out), but this would greatly increase the computational complexity of the method, thereby defeating the goal of real-time application. Instead, we employ a common approximation wherein a fixed set of hyperparameters is determined by maximizing the evidence $p(\bar{d}_m|\bar{\theta})$, and plugging those estimates into the Gaussian process prior:

$$p(\bar{E}_n|\bar{\theta}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\bar{\bar{\Sigma}}_E|^{\frac{1}{2}}} exp\left[-\frac{1}{2}(\bar{E}_n - \bar{\mu}_E)^T \bar{\bar{\Sigma}}_E^{-1} (\bar{E}_n - \bar{\mu}_E)\right]. \quad (6)$$

Here, $\bar{\mu}_E$ is the prior mean, which will be fixed at 0, or it may be chosen on the basis of earlier experiments or expert knowledge. Under the reasonable assumption of a normal distribution of the measurement uncertainty on the emissivity line integrals, described by the variable $\bar{\epsilon}$ in Eq. (2), the likelihood can be written as:

$$p(\bar{d}_m|\bar{E}_n, \bar{\theta}) =$$
$$\frac{1}{(2\pi)^{\frac{m}{2}}|\bar{\bar{\Sigma}}_d|^{\frac{1}{2}}} exp\left[-\frac{1}{2}(\bar{\bar{R}} \cdot \bar{E}_n - \bar{d}_m)^T \bar{\bar{\Sigma}}_d^{-1} (\bar{\bar{R}} \cdot \bar{E}_n - \bar{d}_m)\right]. \quad (7)$$

Here, $\bar{\bar{\Sigma}}_d$ is the covariance of the emissivity, describing measurement uncertainty and correlation on the vector $\bar{d}_m$ of measured line-integrals. We will assume that the various line-integrated measurements are uncorrelated and choose a 5% noise level, based on previous experience at Tore Supra. Therefore,

$$\bar{\bar{\Sigma}}_d = \begin{pmatrix} (0.05 \cdot d_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (0.05 \cdot d_m)^2 \end{pmatrix}. \qquad (8)$$

Finally, the posterior distribution, conditioned on the optimized hyperparameters, reads up to a constant factor,

$$p(\bar{E}_n|\bar{d}_m, \bar{\theta}) \sim p(\bar{d}_m|\bar{E}_n, \bar{\theta}) \cdot p(\bar{E}_n|\bar{\theta})$$
$$\sim exp\left[-\frac{1}{2}(\bar{\bar{R}} \cdot \bar{E}_n - \bar{d}_m)^T \bar{\bar{\Sigma}}_d^{-1} (\bar{\bar{R}} \cdot \bar{E}_n - \bar{d}_m)\right] \cdot$$
$$exp\left[-\frac{1}{2}(\bar{E}_n - \bar{\mu}_E)^T \bar{\bar{\Sigma}}_E^{-1} (\bar{E}_n - \bar{\mu}_E)\right]. \qquad (9)$$

The major advantage of normal distributions and a linear forward model now becomes clear. Indeed, it follows from standard probability calculus that the product of two normal distributions is also Gaussian, with mean vector and covariance matrix given by:

$$\bar{\mu}_E^{post} = \bar{\mu}_E + \left(\bar{\bar{R}}^T \bar{\bar{\Sigma}}_d \bar{\bar{R}} + \bar{\bar{\Sigma}}_E^{-1}\right)^{-1} \bar{\bar{R}}^T \bar{\bar{\Sigma}}_d^{-1} (\bar{d}_m - \bar{\bar{R}} \cdot \bar{\mu}_E), \quad (10)$$

$$\bar{\bar{\Sigma}}_E^{post} = \left(\bar{\bar{R}}^T \bar{\bar{\Sigma}}_d \bar{\bar{R}} + \bar{\bar{\Sigma}}_E^{-1}\right)^{-1}. \qquad (11)$$

The posterior mean is thus available in a closed form and can be used as an estimate of the emissivity field, which can be calculated in real time. In addition, the diagonal elements of the posterior covariance matrix $\bar{\bar{\Sigma}}_E^{post}$ quantify the uncertainty on the inference result.

III. Phantom test and comparison with MFI

To assess the performance of the GPT and compare with existing methods, we first use synthetic data (phantom test). We perform a benchmark with the standard minimum Fisher information method (MFI) which has been thoroughly tested with several fusion diagnostics, e.g. the SXR systems at JET, Tore Supra and TCV. Four different shapes were used for the phantom tests, corresponding to

various situations that are expected to be relevant for WEST SXR emission: Gaussian shape, hollow shape, left-right kidney shape and up-down kidney shape. The white dashed contours in Fig. 4 represent equilibrium magnetic flux surface positions provided by the EQUINOX[10] equilibrium code.
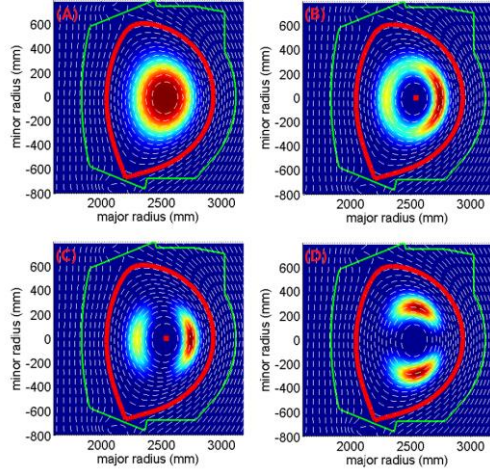

Fig. 4. Four phantom emissivity fields are used in our test: (a) Gaussian shape, (b) hollow shape, (c) left-right kidney shape, (d) up-down kidney shape. The green curve in the figure represents the vacuum vessel and the red curve is the last-closed flux surface. The white dashed curves provide the flux surface geometry. Note that the phantom emissivity has been normalized for the benefit of numerical computation.

The reconstructed emissivity fields based on line integrals with 5% statistical uncertainty are shown in Fig. 5. The quality of the reconstructions can be quantified through a relative error map, showing the difference between the phantom and reconstructed field, normalized by the maximum phantom emissivity:

$$\xi_i = \frac{\left|E_{n,i}^{(rec)} - E_{n,i}\right|}{max\{\bar{E}_n\}}. \tag{12}$$

In addition, in order to quantitatively compare the quality of the reconstructions, the root-mean-square deviation (RMSD) was calculated for each result, given by

$$RMSD = \sqrt{\frac{\sum_{t=1}^{n}(E_{t,i}^{(rec)} - E_{t,i})^2}{n}}. \tag{13}$$

In case of the Gaussian shape, the maximum relative error is around 2%, 8% for the hollow shape, 7% for the left-right kidney shape and 8% for the up-down kidney shape. In general, more complex emissivity field structures are more difficult to reconstruct, the error level depending greatly on the coverage and field of view of the optical system. Nevertheless, in all cases the characteristic shape of the phantom is recovered relatively well by GPT. In addition, one can compare the line integrals obtained from the original phantom, with those calculated from the reconstructed emissivity field. As shown in Fig. 5, good agreement is achieved in all cases. On a typical PC environment with Matlab, each time slice takes about 100 ms calculation time.
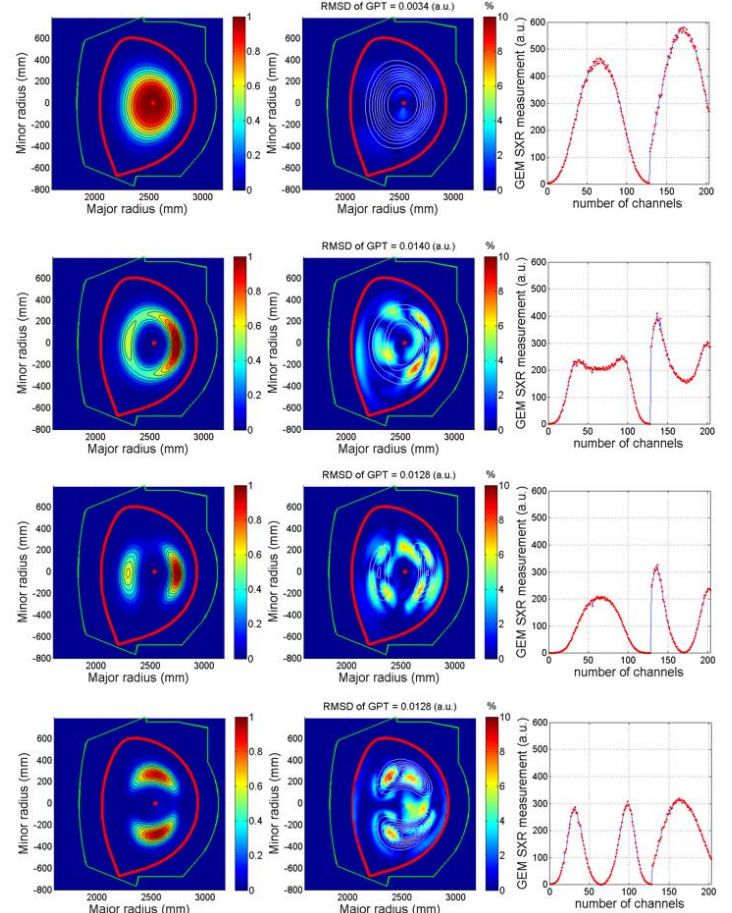

Fig. 5. GPT phantom test with 5% noise level for Gaussian shape, hollow shape, left-right kidney shape and up-down kidney shape phantoms. From left to right, the first column contains the reconstructions, the second column shows the relative error maps according to Eq. (12) (the white contours represent the original phantom), and the third column gives the comparison between the line integrals obtained from the phantom (red dots) and from the reconstructed emissivity fields (blue curves).
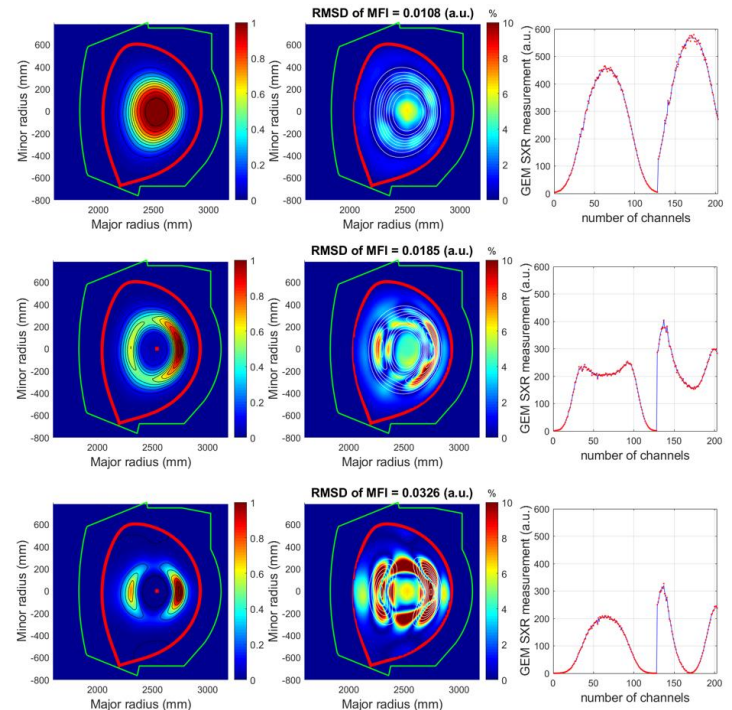

Fig. 6. MFI phantom test with 5% noise level for Gaussian shape, hollow shape and

left-right kidney shape phantoms. From left to right, the first column contains the reconstructions, the second column shows the relative error maps according to Eq. (12) (the white contours represent the original phantom), and the third column gives the comparison between the line integrals obtained from the phantom (red dots) and from the reconstructed emissivity fields (blue curves).

The MFI tomography reconstruction results are shown in Fig. 6. It can be noticed that MFI performs very well in the Gaussian shape case, the maximum relative error being around 6%. However, for the more complex structures like the hollow shape and the left-right kidney shape, errors are much higher: 14% maximum relative error for the hollow shape and over 20% for the left-right kidney shape. Since in GPT, the smoothness along the magnetic flux surface is constrained softly with the characteristic length scale, but in MFI, the smoothness preference is given by a direction weight factor; this gives MFI a lower degree of freedom comparing to GPT. Furthermore, the computational load of MFI is considerably higher than for GPT, in our case amounting to several seconds, compared to 100 ms for GPT.

TABLE I. Error estimation and Root-mean-square deviation.

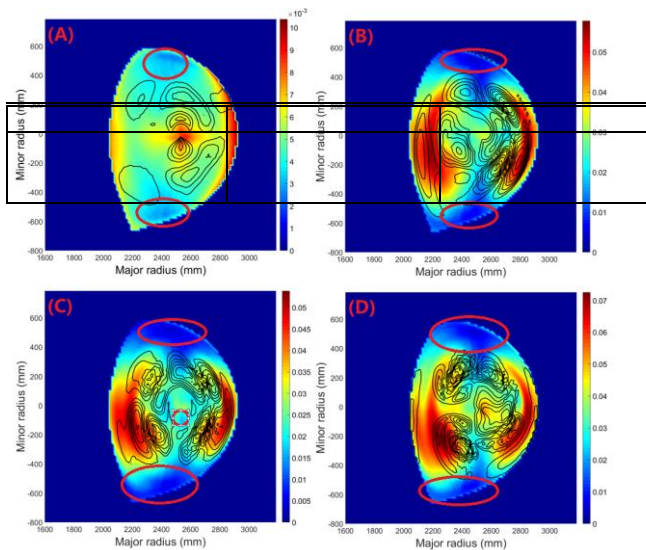| Phantoms	GPT	MFI
Gaussian	2 % max error 0.0034 RMSD (a.u.)	6 % max error 0.0108 RMSD (a.u.)
Hollow	8 % max error 0.0140 RMSD (a.u.)	14 % max error 0.0185 RMSD (a.u.)
Left-right kidney	7 % max error 0.0128 RMSD (a.u.)	Over 20 % max error 0.0326 RMSD (a.u.)
Up-down kidney	8 % max error 0.0128 RMSD (a.u.)	

Fig. 7. Examples of a comparison between the posterior variance map (color map) and relative error map (black contours) on a 5% noise level: (a) Gaussian shape, (b) hollow shape, (c) left-right kidney shape, (d) up-down kidney shape.

Another valuable advantage of GPT is that it provides uncertainty estimates on the reconstructed emissivity field through the the posterior covariance matrix; Eq. (11). Indeed, whereas the relative error field is not available when performing tomography on real WEST data, the posterior variance can still be calculated. The good correspondence of the posterior variance with the relative error field is confirmed in Fig. 7. One could notice that in the lower uncertainty area, which has been marked in red circles, the error value also keeps in a lower level. The posterior variance plots can be used to optimize the viewing geometry of the diagnostic, which will be part of future work.

IV. Conclusion and perspectives

In this paper, a new non-parametric SXR tomography algorithm for WEST based on Gaussian processes has been supplemented with magnetic equilibrium information. Compared to the traditional tomography techniques, GPT has several advantages. First, GPT is not based on an optimization process, therefore it is a computationally light-weight solution, rendering GPT a potential candidate for real-time tomography applications. Second, tests on four typical WEST phantom emissivity fields have pointed out that the quality of the GPT reconstructions compares favorably to results provided by the classic minimum Fisher technique. Third, GPT intrinsically provides uncertainty estimates on the reconstructed emissivity fields, obtained from the posterior Gaussian process. This can be exploited for online self-checking of the algorithm's performance, and can contribute to hardware design optimization. Fourth, it is important to note that GPT also provides good SXR reconstructions without equilibrium information[6]. Finally, the Bayesian probabilistic framework naturally enables an integrated approach, combining the SXR tomography with other forward models in a joint probability model, such as the magnetic equilibrium reconstruction.

V. Acknowledgement

References
[1] J. Svensson and JET-EFDA Contributors, JET Intern. Rep. (2010).
[2] D. Li, J. Svensson, H. Thomsen, F. Medina, A. Werner, and R. Wolf, Rev. Sci. Instrum. **84**, 83506 (2013).
[3] T. Putterich, R. Neu, R. Dux, A.D. Whiteford, M.G. O'Mullane, and H. Summers, Nucl. Fusion **50**, 25012 (2010).
[4] L. Wang and R.S. Granetz, Rev. Sci. Instrum. **62**, 842 (1991).
[5] K. Ertl, W. von der Linden, V. Dose, and A. Weller, Nucl. Fusion **36**, 1477 (1996).
[6] M. Anton, H. Weisen, M. Dutch, P. Marmillod, and P. Paris, Plasma Phys. Control. Fusion **38**, 1849 (1996).
[7] A. Jardin, D. Mazon, M. O'Mullane, J. Mlynar, V. Loffelmann, M. Imrisek and C.B. and P.M. M. Chernyshova, T. Czarski, G. Kasprowicz, A.Wojenski, J. Instrum. **11**, C07006 (2016).
[8] T. Wang, D. Mazon, J. Svensson, D. Li, A. Jardin and G. Verdoolaege, Gaussian process tomography for soft-X ray spectroscopy at WEST without equilibrium information, submitted to Rev. Sci. Instrum., 2018.
[9] D. Mazon et al, Fusion Eng. Des. **96–97**, 856 (2015).
[10] D. Mazon et al, Validation of the New Real-time Equilibrium Code EQUINOX on JET and Tore Supra, EFDA-JET-CP(12)04/07, (2012).