Proceedings of PLASMA 2003 'Research and Applications of Plasmas',
Warsaw, Poland, Sept. 9-12 (2003)

BAYESIAN METHODS FOR INTEGRATED DATA ANALYSIS

A. Dinklage[1], R. Fischer[2] and J. Svensson[1]
[1] *Max-Planck-Institut für Plasmaphysik, EURATOM Association, Wendelsteinstr. 1, 17493 Greifswald, GERMANY*
[2] *Centre for Interdisciplinary Plasma Science, Boltzmannstr. 2, 85748 Garching, GERMANY*

A brief introduction to Bayesian probability theory is given with emphasis on data analysis. The analysis of Thomson scattering data is discussed. Capabilities for assessing diagnostic set-ups are outlined. Extension to more complex analyses aiming at an *Integrated Data Analysis* of fusion diagnostics are presented. The efforts are motivated both by physics and technical requirements of steady-state fusion devices.

1. Introduction

In magnetic confinement research, data analysis benefits from the linkage of different, heterogeneous sources of information. The inclusion of interdependencies increases the reliability of the quantities of interest. Moreover, the underlying physics of the confined plasma has to enter the analysis process if profiles of physics quantities are to be derived. A typical example in fusion research is the mapping of profiles on flux surface coordinates to be derived from equilibrium calculations depending itself on the profiles of plasma quantities. E.g. in high-beta stellarator plasmas, the Shafranov shift needs to be included for proper spatial assignment of density and temperature measurements. Since these enter equilibrium calculations in form of plasma pressure, the derivation of the abscissas and ordinates of profiles becomes an interdependent problem, which is usually solved iteratively. In addition, more sophisticated physics studies, such as transport modeling, can be used to check the consistency of sets of data from different measurements. Therefore, a framework allowing for reasoning on the basis of different uncertain, interdependent information is desirable.

On an earlier stage of comparison of different diagnostics results, the issue of comparability and reliability of uncertainty measures arises. This issue may turn out to be a complicated matter if measurements are differently afflicted with systematic and statistical errors. Also physics models may lead to systematic deviations. An example is the interpretation of electron cycloctron emission from the low field side of magnetically confined plasmas, where the assumption of black body radiation needs to be replaced by a proper treatment of the radiation transport.

A different motivation for new approaches for data analysis is due to the requirements for the data analysis of steady-state fusion devices, such as Wendelstein 7-X presently under construction in Greifswald. In addition to the vast amount of diagnostic data to be analyzed, in a continuously operating device, also a fast analysis of data in conjunction with fast consistency checks of physics quantities must be provided for interventions in running physics programs. For a comprehensive (off-line) analysis of data, the integration of different sources of information leads to most reliable analysis results. This issue becomes crucial for resolving physics motivated questions. For Wendelstein 7-X, this includes the verification of optimization criteria, divertor operation, diminished neoclassical transport, impurity transport, and more [1].

This paper aims at a discussion of Bayesian probability theory for the analysis of fusion data, and physics reasoning based on this data. In the following section, some generic results of Bayesian data analysis will be discussed and the notions resulting from Bayesian methods will be introduced. In the next section, capabilities of applications will be discussed explicitly using the example of Thomson scattering. Possibilities for assessing diagnostic set-ups will also be covered. This issue enters considerations for diagnostic design. The next section is devoted to the integration of different sources of information. Here, basic concepts as well as practical combinations of different diagnostics are discussed. Moreover, some results regarding uncertainties which may enter fitting procedures will be covered briefly.

2. Principles of Bayesian probability theory

Instead of giving a thorough introduction to the field of data analysis employing Bayesian probability theory - which can be found in the literature, such as the monograph by Sivia [2] – this section is intended to give a brief overview of the basic ideas and to introduce typical notions of probability theory.

Probability theory deals with conditional probabilities encoded as probability density functions (PDF). E.g., $p(T_e|data)$ is to be read as the probability density for finding a value of T_e given some *data*. This is the formulation of inference within a probabilistic framework.

Resulting from the product rule for the conditional probability for the joint occurrence of a result given the data and some additional information I, Bayes' theorem relates the probability of finding the quantities (e.g. the electron temperature and the electron density $p(T_e,n_e|d,I)$) to the likelihood function and the prior probability density function (PDF).

Eq. 1 $$p(T_e, n_e \mid d, I) \propto p(d \mid T_e, n_e, I) \times p(T_e, n_e \mid I)$$

The likelihood PDF (Eq. 1: $p(d|T_e,n_e,I)$) contains the forward calculation and may include nuisance parameters. Nuisance parameters are those quantities in a data descriptive equation, which are required for the evaluation of data rather than being the quantities to be inferred. The likelihood is afflicted with uncertainties of the data. Additional PDFs (not shown in Eq. 1) quantify the uncertainties in the nuisance parameters. The posterior PDF (left hand side of Eq. 1) contains all uncertainties entering the problem.

The prior PDF (Eq. 1: $p(T_e,n_e|I)$) describes the *a priori* knowledge on the quantities to be inferred. The principle of maximum entropy can be used for the formulation of most uninformative priors subjected to constraints such as the exponential distribution if the mean value is known. Prior PDFs can also be used to incorporate physical information. For example, the density and the temperature cannot become negative and therefore a prior PDF using this information will vanish for negative values of these quantities. The additional information I abbreviates the physical model of the inference process. For the example discussed later on, this could be considered as the assumption that the Thomson scattering signal is due to classical dipole radiation of an ensemble of free electrons at temperature T_e. Since it plays the role of a normalization factor, the so-called evidence PDF ($p(d| I)$) is omitted in Eq. 1 and needs not to be discussed for our purposes.

Although the formulation of the inference problem in Eq. 1 is straightforward, practical application of Bayes theorem implies computational capabilities. Eq. 1 represents generally a high-dimensional function, the dimensionality of which is given by the number of degrees of freedom in Eq. 1. In order to arrive at the quantities of interest, the sum rule of probability theory can be employed. This procedure is called marginalization:

Eq. 2 $$p(x) = \int p(x,y)dy = \int p(x \mid y)p(y)dy$$

A low dimensional example is shown in Fig. 1 [3]. Fig 1(a) shows the two-dimensional projection of the marginal posterior PDF representing a result of the analysis described in the

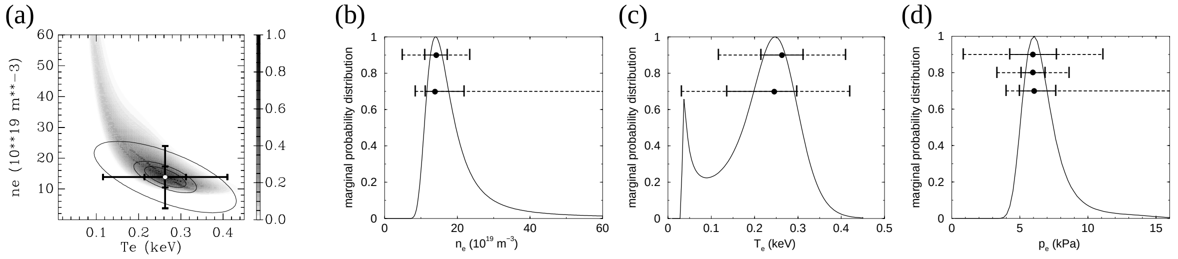

Fig. 1. Marginal posterior PDF (the probability is encoded by the gray scale) representing the result of a Thomson scattering system (a), for details see section 3. Figs. (b) and (c) show the projection of (a) onto the corresponding axes. (d) shows the marginal posterior PDF for a derived quantity, namely the electron pressure. For details of the different error bars see Ref. [3].

next section. The subsequent figures (Fig. 1(b)-(c)) depict the aforementioned projections of the two-dimensional PDF (Fig. 1(a)). Finally, Fig. 1(d) shows a derived quantity, which is the electron pressure $p_e = n_e \times kT_e$. This PDF for p_e can be determined by:

Eq. 3 $$p(p_e \mid d) = \int\int p(p_e, n_e, T_e \mid d)dn_e dT_e$$

Following the rules of probability theory, the PDF on the right hand side of Eq.3 can be resolved as $p(p_e,n_e,T_e|d) = p(p_e|n_e,T_e,d) \times p(n_e,T_e|d)$, where the latter factor is given (see. Fig. 1(a)) and the probability to find the electron pressure, given the density and the temperature is the delta function $p(p_e|n_e,T_e)=\delta(p_e-n_e \times kT_e)$.

Methods for marginalization are integration techniques. For high dimensional integrals, Markov chain Monte Carlo methods can be applied. For estimators, optimization techniques can be used and uncertainties can be estimated employing corresponding approximations (e.g. Laplace approximation) using covariance measures.

The interpretation of the role of prior PDFs in Bayes' theorem is a continuous source of debates. In Bayesian probability theory the PDFs are to be interpreted as probability densities rather than distributions of samples. Therefore, reasoning even on the basis of a few samples becomes possible without knowledge of the entire sampling distribution of quantities.

A recipe for application of Bayesian probability theory comprises in the formulation of the likelihood PDF, the formulation of the prior PDF (which encodes the a priori knowledge), and the calculation of (high dimensional) integrals in order to arrive at the quantities of interest. This conceptual simplicity makes Bayesian probability theory a universal tool for the analysis of data. The next section gives an example for an application in fusion data analysis, different examples can be found in the literature, e.g. [4-6].

3. Bayesian Modeling of Thomson Scattering Data

As an application of Bayesian probability theory, Thomson scattering data from the stellarator Wendelstein 7-AS were analyzed. The Thomson scattering system under consideration (Nd:YAG Laser light source, polychromator system for scattered light detection) was designed to measure central densities of about $n_e = 5 \times 10^{19} \ldots 4 \times 10^{20}$ m^{-3}. During the final campaigns of Wendelstein 7-AS, the central electron temperatures of the scenarios under investigation were in the range from a few hundred eV up to some keV. This section summarizes many more details described in Refs. [3,7].

Following the recipe introduced in the preceding section, as a first step the likelihood function needs to be formulated. The likelihood relates the data d to the physics results (n_e, T_e) to be resolved. For the discussion of the principles of the Bayesian approach, we restrict ourselves to a simplified data descriptive equation:

Eq. 4 $$d \propto n_e \times \int \tau(\lambda)\, S(T_e, \lambda, \ldots)\, d\lambda$$

The data descriptive equation (Eq. 4) also contains parameters (not shown here) which are necessary for the description of the measurement, but which are of no further interest. In Bayesian probability theory, these nuisance parameters are to be marginalized out. The marginalization technique provides the generalization of error propagation laws.

As a practical remark, the formulation of the likelihood turned out to become the most time-consuming effort, since the determination of uncertainties of the different nuisance parameters required detailed laboratory efforts.

For the analysis to be discussed here, the prior PDF for n_e and T_e was chosen to be constant for positive values of the density and the temperature. Negative values of the density and the temperature were excluded.

A result for a single spatial channel of the Thomson scattering data is shown in Fig. 1(a). Here, the two-dimensional marginal posterior PDF for n_e and T_e is shown revealing strong correlations between the quantities of interest. These correlations are close to isobars in the density - temperature plot. The probability density resembles a 'banana shape'. If the values of the electron temperature and density were independent, a two-dimensional PDF figure with principal axes parallel to the abscissa and the ordinate would be expected. For comparison, Fig. 1(a) also shows first order (linear) n_e-T_e correlations which appear in the plots as tilted ellipses.

This finding can be explained by Eq.4 , which relates the data d to the product of the density n_e and the overlap of the scattering function S and the transmission characteristics τ. In first order expansion, the integral on the right hand side is proportional to the temperature T_e. Hence, single data are determined by electron pressure rather than electron density or electron temperature independently. However, by combination of different spectral channels, density and temperature can be disentangled. The less the information from different spectral channels is, the more pronounced the banana shaped PDF appears. In this sense, the example shown in Fig. 1(a) is an example of poor resolution but depicting the basic structures of the underlying diagnostics model most significantly. This example demonstrates, that Bayesian probability theory is capable to recover the complete informational content of data including interdependencies. As a consequence, for the case presented here the error bars for the electron pressure is not well described by the error propagation of the separate error bars of n_e

and T_e. Due to the correlation the relative error of p_e is smaller than the relative errors both of the density and the temperature [3].

These examples show, that Bayesian probability theory allows for a thorough analysis of data on the basis of diagnostic models containing information of all uncertainties. As a consequence of this approach, there is no distinction between so-called statistical and systematic errors. Possible systematic errors can be introduced for instance as unknown, but existing scaling factors. Then, the additional nuisance parameter is to be marginalized in order to incorporate the effect of this systematic deviance.

The benefits for physical inference needs to be discussed in the context of subsequent analysis steps. If – for example – the derived profiles enter transport analyses, the aforementioned correlation has to enter the analysis of density and temperature gradient driven fluxes. An increase of density results from the Thomson scattering system consequently gives a diminished temperature and vice versa, also affecting the interdependencies of gradients.

The full benefit of Bayesian probability theory, however, appears if different diagnostic models are to be combined. Before discussing this issue in more detail, the capabilities for diagnostics optimization within the framework discussed will be presented.

4. Assessment of diagnostics set-ups

Once a diagnostic model is derived, this model can also be regarded as a tool for the assessment of diagnostics capabilities in order to optimize existing set-ups [8]. Moreover, if artificial data were provided, optimization can be done already in the stage of diagnostics design. A final goal of procedures for diagnostic design within the Bayesian framework is the optimization of utility functions formulated in the context of physics goals.

Fig. 2 shows some results of the assessment of the Thomson scattering system under consideration. Here, the impact of different nuisance parameters entering the full diagnostics model are discussed.

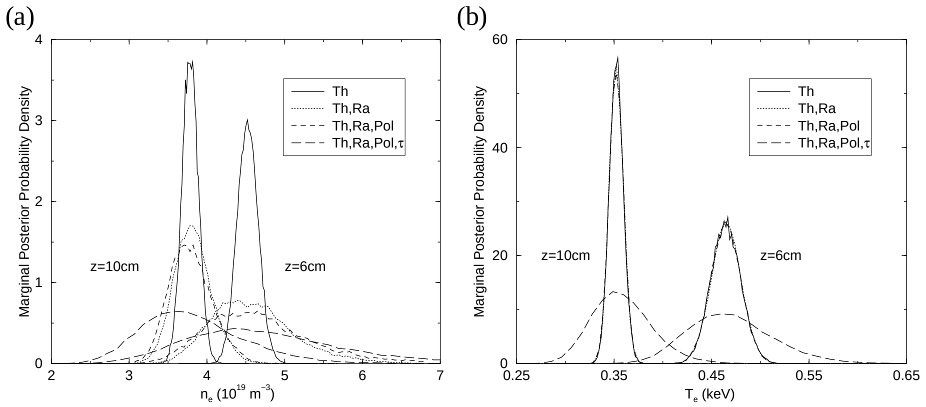

Fig. 2. Marginal posterior PDF for the density measurement (a) and the temperature measurement (b) of two spatial channels (z = 6 cm and z = 10 cm) of the Thomson scattering system under consideration. Here, different nuisance parameters are switched on subsequently.

The most relevant nuisance parameters in the Thomson scattering model are calibration factors (Ra for Raman scattering), alignment of optical components (Pol for polarizing elements in the optical path) and the transmission characteristics (τ) of spectral filters. The uncertainties of these nuisance parameters are switched on subsequently resulting in a broadening of the marginal posterior PDFs and even in a small shift of the maximum (Fig. 2(a)). Fig. 2(a) shows the results for the density, whereas Fig. 2(b) displays the impact of nuisance parameters on the determination of temperature. From the latter panel it can be concluded, that it is the accuracy of the transmission characteristics, which determines the error of the temperature measurement in addition to the scattering of the data (plot Th in Figs. 2(a)-(b)). The density measurement is strongly affected by τ but also by the calibration measurement as to be expected from Eq. 4. From this example it can be concluded that the optimization of diagnostic set-ups requires different nuisance parameters to be adapted for various targets of the utility function.

5. From one to many: INTEGRATED DATA ANALYSIS of different diagnostics

The analysis of diagnostic data within Bayesian probability theory allows for the linkage of measurements from different diagnostics. Fig.3 shows PDFs for different information to be linked. In addition to the Thomson scattering results, two more diagnostics were included, chosen for exemplification of the effect of different stages of information: As an example of 'weak' information, Fig. 3(b) shows the PDF for the operation of a microwave interferometer. Here, the PDF encodes that an operating microwave traversing the plasma indicates that all densities have to be below the cut-off density. The smooth fall-off towards this limiting density encodes the diminishing reliability due to diffractive effects. Even though being quite coarse information, the result shown in Fig. 3(b) cuts off the physically irrelevant long tails for high densities in Fig. 3(a), which are, however, still consistent with the data of the Thomson scattering system alone. Fig. 3(c) is an example of an additional measurement, where the PDF representing the soft X-ray temperature is determined by a forward calculation rather than representing a complete model of that diagnostics. For the purpose of clarity, density effects were not included in Fig. 3(c) which results in an n_e invariant PDF.

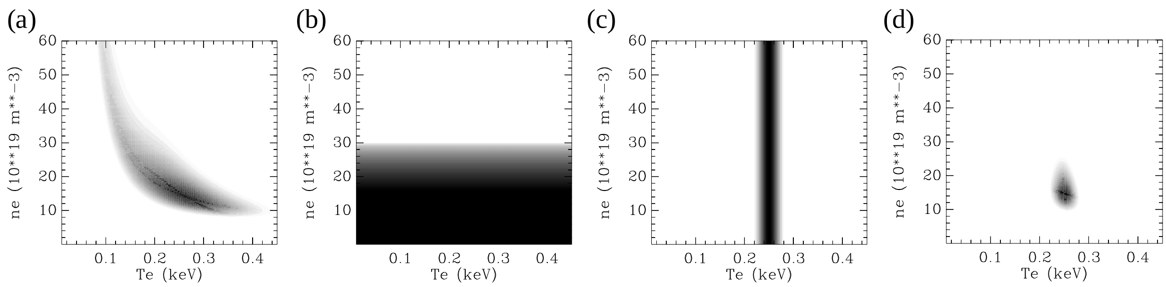

Fig. 3.Integrated Data Analysis: The combination of PDFs from different diagnostics for a single spatial position. (a) is the Thomson scattering result (b) indicates operation of a microwave interferometer, (c) is a temperature measurement from soft X ray measurements. Panel (d) is the combination of the (a-c).

As suggested by the result shown in Fig. 3(d), which shows the product of the PDFs in Figs. 3(a-c), the final outcome is mostly determined by the high accuracy of the soft X-ray temperature measurement. But as it becomes evident from Fig. 3(d) also the accuracy of the density measurement is enhanced by the combination of the PDFs. Again, this result is due to the strong correlation of density and temperature from the Thomson scattering measurement. It is the inclusion of interdependencies which allows for an increase of significance of the outcome of a set of diagnostics [9,10].

6. Bayesian Methods for complex diagnostics models

When attempting to solve the general problem of integration of a large set of fusion diagnostics, the number of interdependencies between parameters of the type mentioned above increases rapidly. Fig. 4 shows an implemented Bayesian graphical model for Wendelstein 7-AS, where dependencies (functional or probabilistic) between different parameters (nodes) are explicitly shown as directed edges between the nodes. The model currently includes the above mentioned Thomson scattering diagnostic, a microwave interferometer, diamagnetic energy measurement, ion temperature profile from a neutral

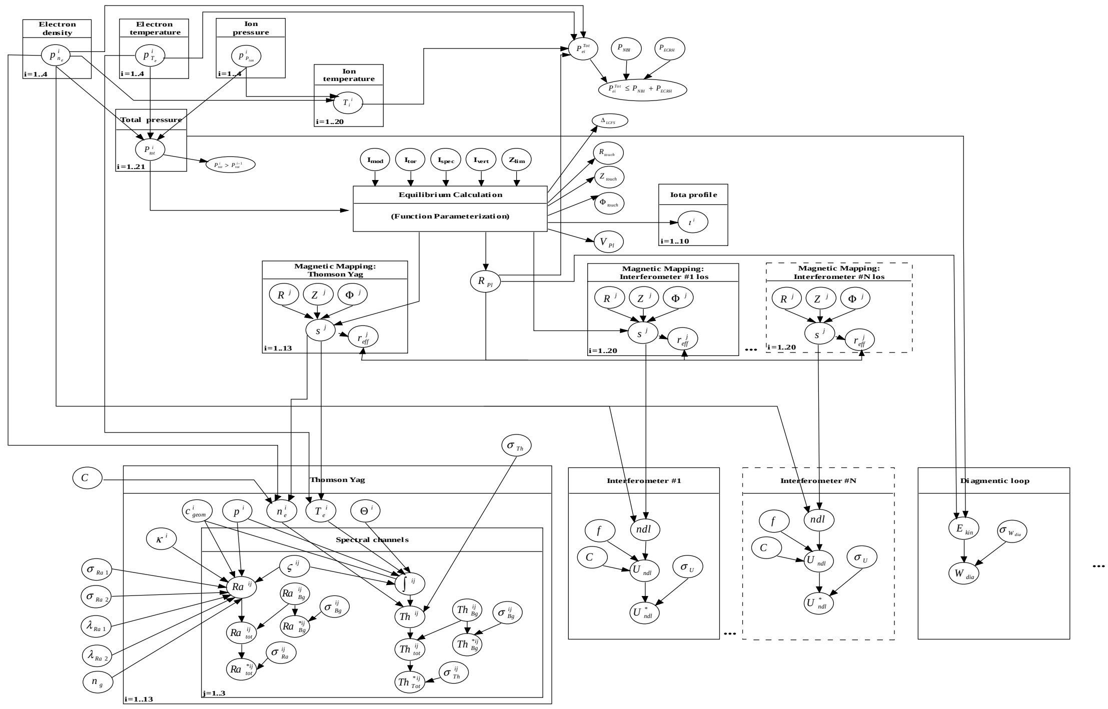

Fig. 4. Bayesian Graphical Model for Wendelstein 7-AS diagnostics.

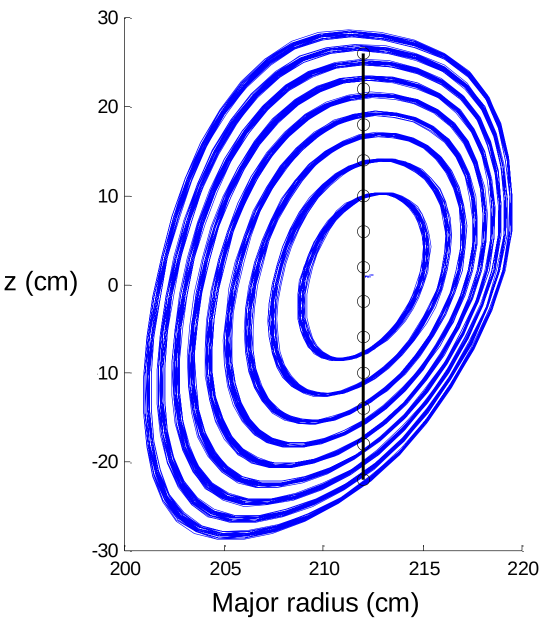

Fig. 5 Uncertainty of magnetic mapping. Thomson scattering line of sight superimposed.

particle analyzer diagnostic, and geometry information for calculation of the position of the last closed flux surface. A fundamental cause of interdependencies in this model is the mapping of different diagnostic signals to a common magnetic coordinate system. This mapping is itself dependent on measurements (e.g. pressure profiles) through an equilibrium calculation, and the integrated model must therefore include a self-consistent model for calculating the magnetic coordinate system from estimated profiles which themselves rely on the estimated mapping [11]. The mapping will therefore itself have errors originating from statistical and systematic errors in the whole system. The self-consistent calculation, which includes a *Function Parameterization* equilibrium calculation [12], further has to be applied to the mapping of the full set of diagnostics. From the integrated model, marginal posterior probability densities for all nodes in Fig. 4 can be derived. As an example, Fig. 5 shows the Thomson scattering laser line superimposed on a graph showing inferred flux surfaces with an uncertainty indicated by the thickness of each surface.

7. Conclusion

Capabilities of Bayesian probability theory for the analysis of data from fusion diagnostics were outlined. Bayesian probability theory is a universal tool for a concise treatment of uncertainties of any kind, which are to be encoded as probability density functions. If applied consequently, comparable data evaluation from different diagnostics can be assured. Different approximations allow for fast, but coarse analyses, or for detailed, but slow evaluations. The use of fast approximations is a promising option for on-line data analysis and validation in steady-state fusion devices. The underlying diagnostics models are valuable for the assessment of diagnostic and the design of future set-ups.

The combination of different PDFs in order to arrive at an INTEGRATED DATA ANALYSIS is then a straightforward matter. Bayesian graphical models allow for a clearly arranged formulation of complex diagnostics models. As a concrete benefit, the proper incorporation of interdependencies enhance the level of reliability strongly. This allows for the inclusion of subsequent theoretical analysis steps in a concise manner.

Acknowledgements
Helpful discussions with J. Geiger, H.-J. Hartfuß, M. Hirsch, H. Maaßberg, E. Pasch and A. Weller are appreciated.

References

[1] G. Grieger, C.D. Beidler, E. Harmeyer, *et al.*, *Plasma Physics and Controlled Nuclear Fusion Research*, Proceedings of the 12th International Conference, Nice, 1988 (IAEA, Vienna, 1989), Vol. 2, p. 369; G. Grieger, W. Lotz, P. Merkel, *et al*., Phys. Fluids B **4**, (1992), 2081.
[2] D. S. Sivia, *Data Analysis, a Bayesian Tutorial*, Clarendon (Oxford), 1996.
[3] R. Fischer, A. Dinklage and E. Pasch, Plasma Phys. Control. Fusion **45**, (2003) 1095.
[4] A.P. Millar, D.A. Mc Donald, D.A. Diver, Plasma Phys. Control. Fusion **42**, (2000) 337;
[5] V. Dose, R. Preuss, R. Fischer, Phys. Rev. Lett. **81** (1998) 3407.
[6] R. Giannella, J. Lasalle, M. Fois, *28th EPS Conference on Contr. Fusion and Plasma Phys. Funchal, 18-22 June 2001* ECA Vol. **25A,** 381 (2001).
[7] R. Fischer, C. Wendland, A. Dinklage, S. Gori, V. Dose, Plasma Phys. Control. Fusion **44** (2002) 1501.
[8] H. Bindslev, Rev. Sci. Instrum. **70**, (1999) 1093.
[9] A. Dinklage, R. Fischer, M. Hirsch, E. Pasch, A. Weller, *30th EPS Conference on Contr. Fusion and Plasma Phys., St. Petersburg, 7-11 July 2003* ECA Vol. **27A**, P-1.52 (2003).
[10] A. Dinklage, R. Fischer, G. Kühner, H. Maaßberg, E. Pasch, J. Svensson, *29th EPS Conference on Plasma Phys. and Contr. Fusion Montreux, 17-21 June 2002* ECA Vol. **26B**, P-5.103 (2002).
[11] J. Svensson, A. Dinklage, J. Geiger, R. Fischer, *30th EPS Conference on Contr. Fusion and Plasma Phys., St. Petersburg, 7-11 July 2003,* ECA Vol. **27A**, P-1.65 (2003).
[12] H. Callaghan, J. Geiger, P. Mc Carthy, Plasma Phys. Contr. Fusion **42**, 1013 (2000).