*30th EPS Conference on Contr. Fusion and Plasma Phys., St. Petersburg, 7-11 July 2003 ECA Vol. **27A**, P-4.80*

From Off-line to Real-time Analysis: Accelerating Bayesian Analysis Codes

A. Dinklage[1], R. Fischer[2], J. Geiger[1], S. Kühner[1], H. Maaßberg[1],
H. Svensson[1] and U. von Toussaint[2]

[1]*Max-Planck-Institut für Plasmaphysik, EURATOM Association, BI Greifswald, Wendelsteinstr. 1, 17491 Greifswald, GERMANY*
[2]*Centre for Interdisciplinary Plasma Science, Boltzmannstr. 2, 85748 Garching, GERMANY*

Abstract

Capabilities of real time performance of a Bayesian analysis codes are explored.

Background

For continuously working fusion devices timely provision of significant physics quantities are desirable for possible interventions in running physics programmes. Typical time scales relevant for control and interventions for the Wendelstein 7-X stellarator are given by MHD fluctuations ($\tau_{MHD} \approx 10^{-3}\,\mathrm{s}$), the energy confinement time ($\tau_E \approx 10^{-2\cdots-1}\,\mathrm{s}$) and the plasma current relaxation time ($\tau_{L/R} \approx 10^{1}\,\mathrm{s}$). Both real time control and possibilities for human interaction may benefit from such a fast analysis. The term 'real-time analysis' refers to the usual timeliness condition for the completion of analysis – here $\tau_A \ll \tau_{L/R}$ – allowing for a synchronous operation of analysis procedures and a sufficiently high repetition rate of analyzes. Possibilities for the acceleration of data analysis codes with focus on such a continuously working data analysis techniques are discussed, such as parallelization. Additional emphasis is led on the real-time provision of significance measures. The required analysis time of codes based on Bayesian models of diagnostics is discussed in figures of approximations entering the full error analysis. Moreover, the linkage of different diagnostic results through mapping procedures is a key issue for integration of different profile diagnostics. Within the preparation of continuously working plasma devices, studies of requirements for real-time data analysis may be a considerable issues for planning of the computational equipment [1].

Analysis by Bayesian modelling in figures of computational costs

In previous studies [2] Bayesian probability theory (BPT) was employed for modelling of diagnostic data and for assessment and improvement of diagnostic capabilities [2, 3]. In Fig. 1 the outcome of the Bayesian analysis of a single spatial position of a YAG Thomson scattering system is shown in order to discuss some specific features of the Bayesian analysis.

As seen from Fig. 1, the resulting probability density function (PDF) may be non-Gaussian, but Gaussians are an essential prerequisite for error propagation laws [3].

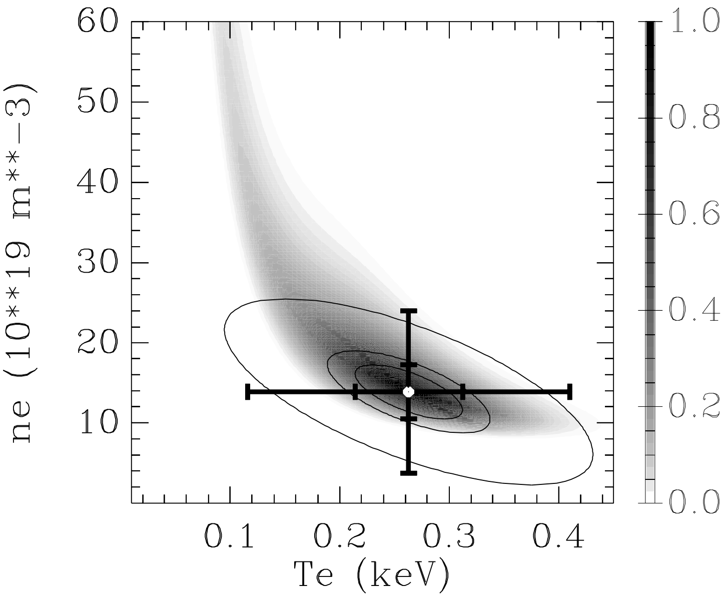

Figure 1: Different approximations on a two dimensional PDF depicting the probability for a n_e–T_e tuple inferred from Thomson scattering data [3]. The grey scale represents the maximum normalized PDF.

Estimates for parameters can be represented by the maximum, mean value or median of the posterior PDF. Uncertainty measures can be derived from the Hessian matrix at the maximum of the posterior PDF or from confidence intervals enclosing a certain amount of probability. But for skew distributions or multi-modal distributions, however, single estimates and uncertainties may be misleading. Superior to single estimates are low-dimensional projections of the full posterior distribution contain the uncertainty structure of the full space which are given by integrating the posterior PDF over a sub-space of the full parameter space. Hence, integration techniques turn out to be the major time consuming part of the inference process.

The different techniques for Bayesian inference can be subdivided with respect to performance in figures of approximation and speed. Maximization of distributions is straightforward with commercial numerical packages, but calculating uncertainties from the Hessian at the maximum can be more cumbersome. Integrals can be performed with (adaptive) quadrature techniques in low-dimensional cases (dimension m<10) for which numerical packages exist. Integrals with large dimensions (m>5) must be performed with more elaborated tools. The Laplace ansatz approximates the PDF by a multi-dimensional Gaussian distribution which can be integrated analytically over an appropriate subspace. For PDFs where the Laplace approximation fails one has to employ Markov Chain Monte Carlo (MCMC) techniques for integration.

Bayesian neural networks provide an increasingly popular technique for reasoning under

Method	Marg.	Est.	Unc.	CPU time	Par.	Approx.
Full MCMC	Yes	Moments	full PDF	$> 10^3$ s $N \times M$	Yes	Convergence
Maximization	No	Maxima	No	$> 10^1$ s $N^2 \times M$	Partially	Single point
Hessian	Yes	No	Yes	$> 10^1$ s $N^3 \times M$	Yes	Laplace
GME	(Yes)	Yes	(Yes)	$10^{0...2}$ s M^3	Partially	Entropic
Neural networks	Yes	Yes	Yes	$10^{-3...-2}$ s $N \times k$	Yes	training set knots, layers nets k

Table 1: Approximations vs. computational costs. CPU time refers to 1 GHz Pentium 3 processors. (Marg.: Marginals; Est.: Estimates, Unc.: Uncertainties; Par.: Parallelization; Approx.: Approximation).

uncertainty. The main benefit of neural nets is the flexibility to model any unknown physical situation only from a training set. A second advantage of trained neural nets is given by the performance in time-critical situations. The training data sets are based on full Bayesian calculations including all relevant information and uncertainties whereas for an on-line analysis of measured data only the trained net is used. In addition to profile information neural nets may provide fast magnetic configurations [5]. The Generalized Maximum Entropy (GME) method is based on the joint entropy maximization of noise and parameter distributions. GME is fast for linear problems (e.g. tomography). GME is a robust technique since the entropy norm is much less sensitive to outliers than the usual L2-norm. Uncertainties can be obtained via bootstrap methods which can be fully parallelized. Costs depend on the number of data rather than on the number of parameters which can be an enormous benefit.

As a technique for code acceleration, parallelization can be performed by dividing the inference problem into subtasks, all of which can then be solved simultaneously. E.g., for MCMC calculations the performance of codes is directly proportional to the number of processors. Generally, the performance of the different techniques depend on the sensitivity to the number of data M and the number of parameters N. Table 1 summarizes different techniques with respect to computational performance and outcome.

Towards Integrated Data Analysis: The impact of equilibrium calculations

In figures of computational effort, equilibrium calculations for W7-AS by means of the NEMEC code require about 3000 s for a standard case (CPU 1 GHz Pentium 3, 55 $\times$ 24 spatial grid, 8 Fourier harmonics). For W7-X the computational effort is about a factor of

10 larger (CPU 1 GHz Pentium 3, 99×16 spatial grid, 12 Fourier harmonics), however, the number of iterations for attaining convergence is a matter of experience. One order of magnitude in computational speed can be gained using larger facilities (e.g. NEC SX-7) but at much higher costs.

So far, the approach to resolve the interdependency of configurations and profile information is an iterative analysis of profiles and equilibrium calculations. A much faster approach is the function parameterization (FP) of Fourier components representing equilibria by means of principal component analyzes of pre-calculated equilibria [5, 6]. A single FP equilibrium calculation including mapping of one sightline takes about $0.3\,\mathrm{s}$ on a 1 GHz Pentium 3 PC. This time includes a 3D calculation of the distance between the computed last closed flux surface and the nearest found in-vessel component, necessary for optimization of plasma volume. The speed of this approach makes it possible to map different diagnostics on magnetic simultaneously with the estimation of pressure profiles inferred jointly from different diagnostic signals [6].

Outlook

Present techniques of Bayesian inference allow for the fast provision of analysis results. Possibilities for parallelization exist such that real-time analysis based on appropriate approximations can be expected to become realized if required. Moreover, the computer capabilities to expected double each eighteen months according to *Moore's law*. A proof of principle application within a running data storage environment has to be demonstrated as a next step.

References

[1] J.A. How, J.W. Farthing, V. Schmidt, *Trends in Computing Systems for Large Fusion Experiments*, 22nd Symposium on Fusion Technology, Helsinki, Finnland, to be published in Fusion Engineering and Design.

[2] R. Fischer, C. Wendland, A. Dinklage, S.Gori, V.Dose, Plasma Phys. Contr. Fusion **44**, 1501 (2002).

[3] R. Fischer, A. Dinklage, E. Pasch, Plasma Phys. Contr. Fusion **45**, 1095 (2003).

[4] H. Callaghan, P. Mc Carthy, J. Geiger, Plasma Phys. Contr. Fusion **42**, 1013 (2000).

[5] A. Sengupta *et al.*, Contribution P1-13, this conference.

[6] J. Svensson *et al.*, Contribution P1-65, this conference.