

Non-parametric profile gradient estimation

R. Fischer¹, A. Dinklage², Y. Turkin²

¹ *Max-Planck-Institut für Plasmaphysik, EURATOM Association, D-85748 Garching, Germany*

² *Greifswald Branch, D-17493 Greifswald, Germany*

Reliable profile and profile gradient estimates are of utmost importance for many different physical models in fusion science, e.g. transport modeling or mode stabilization. Fitting profiles to a collection of results from different diagnostics defines basic work in plasma physics. The fitting results often crucially depend on the functional representation of the profile. In particular, the estimation uncertainty of the profile and, even worse, the estimation of the profile gradient and its uncertainty is closely coupled with the provided profile flexibility. This is one reason why profile gradient uncertainties are usually not estimated. Profile flexibility to allow for a form-free description of the data often competes with profile reliability. The estimation reliability decreases with increasing number of degree of freedom provided. The problem of the proper choice of the functional representation of the profile is hampered by measurement errors of the usually pointwise measurements of profiles and by lack of information in profile segments. Severe complications arise from systematic deviations due to inconsistent diagnostics.

The aim is to have a robust technique to allow for a reasonable balance between flexibility and reliability. Flexibility is frequently obtained by using non-parametric profile functionals, e.g. linear interpolation between pointwise estimations or cubic or B-splines. The reliability of the profile estimates is determined by the data uncertainties as well as the supported degree of freedoms (DOF). The balance between fitting the significant information content in the data and avoiding noise fitting poses a major problem. In the framework of Bayesian probability theory the competition between flexibility and reliability is tackled in a natural way by marginalizing of (integrating out) all model parameters including the DOF. The result is a marginal posterior probability distribution of the quantity of interest which implicitly considers penalization of DOF. Implicit penalization of DOF together with a comprehensive physical and statistical (including systematic errors!) description of sets of diagnostics in the framework of Integrated Data Analysis (IDA) [1] allows for a proper balance between profile flexibility and reliability.

Frequently, reliability is obtained by either providing a family of tailored parametric functionals or piecewise polynomial functions combined with modified hyperbolic tangent functions (tanh) at the plasma edge. For sufficient flexibility support we propose to use exponential splines. Consider a set of function values f_i given at E support points ξ_i (pivots). The exponen-

tial spline function $S_i(x)$ in the interval $\xi_i \leq x \leq \xi_{i+1}$ is then given by

$$S_i(x) = \alpha_i + \beta_i(x - \xi_i) + \gamma_i\psi_i(x - \xi_i) + \delta_i\phi_i(x - \xi_i). \quad (1)$$

The auxiliary functions ψ_i and ϕ_i contain a stiffness parameter λ_i on the support $[\xi_i, \xi_{i+1}]$ and are given by the hyperbolic functions

$$\psi_i(x - \xi_i) = 2\{\cosh[\lambda_i(x - \xi_i)] - 1\}/\lambda_i^2 \quad (2)$$

$$\phi_i(x - \xi_i) = 6\{\sinh[\lambda_i(x - \xi_i)] - \lambda_i(x - \xi_i)\}/\lambda_i^3 \quad (3)$$

The coefficients $\alpha, \beta, \gamma, \delta$ are determined from the requirement of continuity of function, first and second derivatives at the pivotal points ξ_i . The system can be closed by setting at the end points either the second derivatives to zero or the first derivatives to given values. In the present work the gradients in the plasma center and outside of the SOL are set to zero. In the limiting cases of $\lambda \rightarrow 0$ and $\lambda \rightarrow \infty$ the exponential spline becomes a cubic spline and a polygon, respectively. It allows for less curvature regularization compared to the cubic spline which is known to over-smooth the pedestal region and more flexibility than a linear interpolation scheme. The parameters of the exponential-spline, amplitudes f_i , pivot positions ξ_i and stiffness parameters λ_i as well as the number of spline pivots E are marginalized. The probabilistic marginalization of the nuisance parameters is the main ingredient which allows to be as flexible as necessary to describe the data as well as robust in the sense that the noisy data are not overfitted. Details about the exponential-spline approach within the framework of Bayesian probability theory and a systematic exploration of the method can be found in [2, 3].

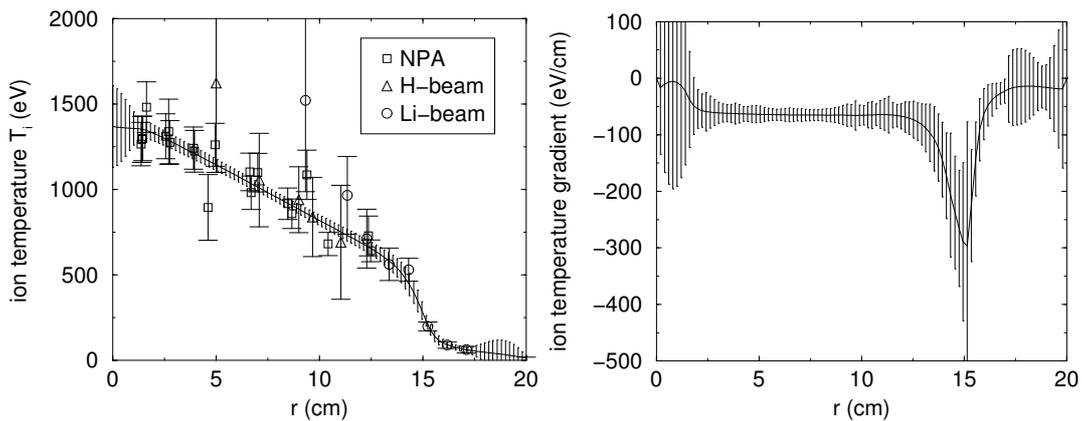


Figure 1: Ion temperature profile and profile gradient of #54285 of W7-AS.

To demonstrate the approach, data of a study of the neoclassical predictions for the radial electric field E_r in the stellarator Wendelstein 7-AS were reanalyzed [4]. Data from a set of

diagnostics consisting of Thomson scattering (TS) (T_e, n_e), neutral particle analyzer (NPA), H-beam (T_i), and Li-beam (T_i, n_e) were used. The left panel of figure 1 shows an estimate of the T_i profile and its uncertainty. The solid line and the error bars represent the mean value and the ± 1 standard deviation of the profile posterior probability distribution calculated using a Markov Chain Monte Carlo (MCMC) sampling technique. The shape of the exponential spline is close to linearity for $r < 13$ cm and $r > 17$ cm, and close to a cubic spline in between. The error bars depict a large variability close to the plasma center and at the plasma edge where the data are sparse. The right panel of figure 1 shows an estimate of the T_i profile gradient calculated from the MCMC samples of the analytic derivative of the exponential spline. The gradient profile is well determined at the linear region where the number of data is large. The gradient error bars are larger close to the plasma center and in the plasma edge region where the data are sparse. Due to the methodologically inherent competition between flexibility and reliability the approach provides a reliable tool for gradient estimation.

Although the gradient of the exponential spline at $r = 0$ and at $r = 20$ cm was constrained to be zero, there is enough flexibility to have non-zero gradients in the closest neighborhoods. Additional physical constraints have to be applied to consider power balance in the plasma. Figure

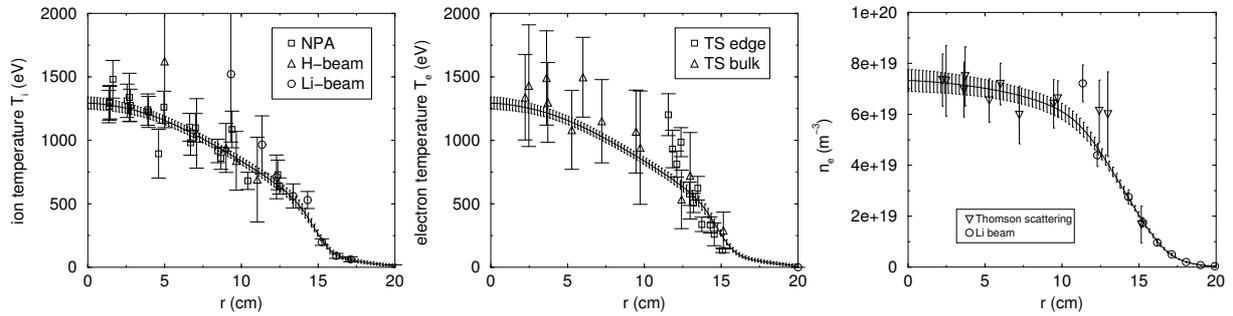


Figure 2: Profiles of ion and electron temperature, and electron density simultaneous sampled from a joined posterior probability distribution.

2 shows profiles for T_i , T_e and n_e where the ion and electron pressure profiles were constrained to be monotonically decreasing. Such hard constraints as well as any other physical consideration can be easily incorporated into the MCMC approach. In addition to the monotonicity constraint, an empirical soft constraint was applied considering power balance in the plasma. A detailed comparison of various physically motivated constraints on profiles and combinations of profiles will be given in [3]. The combined analysis of the profiles T_i , T_e and n_e within the IDA approach allows straightforwardly to estimate derived physical quantities, e.g. the radial electric field E_r and the electron-ion energy exchange term Q_{ei} . The left panel of figures 3 depicts the E_r profile

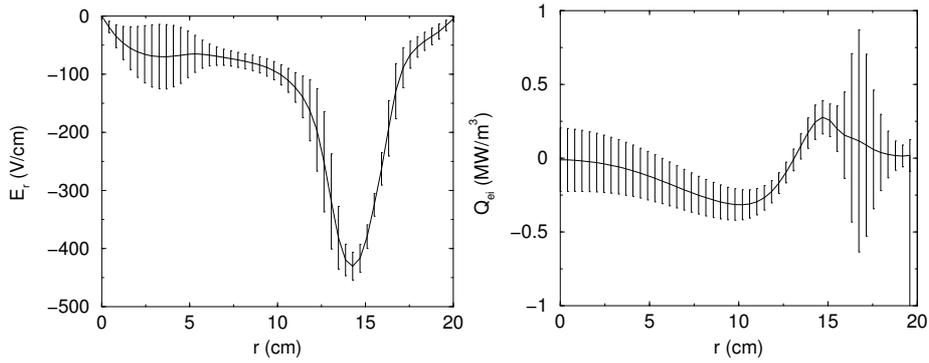


Figure 3: Radial electric field and electron-ion energy exchange term.

calculated from the diffusion equation for E_r (equation (1) in [4]) and the right panel shows the Q_{ei} profile. The marginal posterior probability distributions for both derived profiles are summarized by their mean values (solid line) and standard deviations (error bar). Numerically, they are calculated from the mean and variance of the samples of $E_r(r)$ and $Q_{ei}(r)$ calculated using the MCMC samples of the profiles T_i , T_e and n_e summarized in figure 2. The difference of the T_i and T_e profiles result in a $Q_{ei}(r)$ profile which deviates marginally significant from zero only at about $r = 10$ cm and $r = 14$ cm.

In conclusion, the non-parametric exponential-spline approach for profile and profile gradient estimation provides a robust method for a useful balance between flexibility and reliability. Uncertainties of profiles and gradients is readily derived from data uncertainties. The DOF is determined by the significant information content of the data. Profiles of derived physical quantities including their uncertainties are readily estimated within the probabilistic approach.

References

- [1] R. Fischer and A. Dinklage. Integrated data analysis of fusion diagnostics by means of the Bayesian probability theory. *Rev. Sci. Instrum.*, 75:4237–4239, 2004.
- [2] V. Dose and R. Fischer. Function estimation employing exponential splines. In K.H. Knuth, A.E. Abbas, R.D. Morris, and J.P. Castle, editors, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, volume Conf. Proc. 803, pages 67–71, Melville, NY, 2005. AIP.
- [3] R. Fischer, A. Dinklage, Y. Turkin, and V. Dose, 2006. to be published.
- [4] H. Ehmler, Y. Turkin, C.D. Beidler, H. Maassberg, A. Dinklage, T. Klinger, and W7-AS Team. Experimental check of neoclassical predictions for the radial electric field in a stellarator. *Nucl. Fusion*, 43:L11–L13, 2003.