Dimensionally Exact Energy Confinement Scaling in W7-AS

V. Dose, R. Preuss, and W. von der Linden
Max-Planck-Institut für Plasmaphysik, EURATOM Association, D-85740 Garching b. München, Germany
(Received 31 March 1998)

Energy confinement in W7-AS has been analyzed in terms of dimensionally exact form free functions employing Bayesian probability theory. Based upon the international stellarator database, which contains the energy content for a wide variety of variable settings, predictions for single variable scans are made. The scaling functions for density and power scans, respectively, are in quantitative agreement with data collected in W7-AS. Furthermore, the optimal model for the description of the global transport in W7-AS is identified as the collisional low-beta kinetic model. [S0031-9007(98)07405-5]

PACS numbers: 52.55.Pi

Fusion plasma behavior has been described for about 20 years by energy confinement scaling functions [1]. Such confinement relations serve presently primarily two purposes: first, they constitute a convenient summary of machine operation. This allows intermachine comparisons and the characterization of conditions for enhanced confinement regimes. Second, energy confinement scaling provides the basis for the design of future experiments such as ITER representing the tokamak line or W7-X and LHD in the stellarator branch. Confinement scaling relations have been used to predict L-mode tokamak performance with some success [2]. This is notable for several reasons. First, the popular power law functional form of the confinement scaling function was originally assumed for reasons of convenience and simplicity and lacks a physical foundation. This initial choice has subsequently been justified by a surprisingly good characterization of data trends. There have also been, now and then, attempts at improved data representation by more complicated functions as for example the class of offset linear scalings [3]. Since they were not really superior to a single power law term, the latter has accumulated considerable credit just by experience.

Another major shortcoming of the unconstrained power law scaling function is its dimensional incorrectness. Connor and Taylor [4] tried to interpret experimental scaling functions, established by Hugill and Sheffield [1], in terms of constraints derived from the requirement of physical invariance under similarity transformations of the basic equations describing plasma behavior. These attempts were accompanied by considerable frustration since they found that the experimental scaling was incompatible with any of their plasma models. Based on this experience they suggested that the theoretically derived dimensional constraints be incorporated directly in the power law ansatz. This proposal has been followed subsequently on many occasions [5]. It consists of expressing the energy content W of the toroidal magnetic confinement device by [4]

$$\begin{aligned} W^{\text{theo}} &= cna^4RB^2\left(\frac{P}{na^4RB^3}\right)^{x_1}\left(\frac{a^3B^4}{n}\right)^{x_2}\left(\frac{1}{na^2}\right)^{x_3} \\ &= c'g(n,B,P,a;\mathbf{x}), \end{aligned} \tag{1}$$

where n is the average density, a and R minor and major radius of the torus, B the magnetic field, and P the deposited heating power. The particular values of x_1, x_2, x_3 specify the plasma kinetic model as shown in Table I. Since we concentrate in this paper on data from a single machine, parameters which are constant within the examined data set (e.g., R) are absorbed in c'. The number of degrees of freedom in (1) varies—depending upon the model—between one and three, whereas the unconstrained ansatz for a single device with $R =$ const would have four (n, B, P, a). Imposing physical constraints on the power law ansatz reduces the flexibility and leads necessarily to an increased misfit. We will demonstrate in this paper, however, that a dimensionally exact power-law type of ansatz can be formulated which leads to a significantly reduced misfit.

Interestingly enough, already Connor and Taylor [4] proposed to express a general form free energy confinement scaling function as a series of terms of the form (1) for properly chosen $\mathbf{x}_k$ with expansion coefficients c_k. In mathematical terms this is nothing but the expansion in a basis which is dimensionally exact. In this paper we shall exploit this suggestion. Consider a set of measurements of the plasma energy content for N different values of the experimental input variables (n, B, P, a). We represent the theoretical prediction for the energy content by an N-dimensional vector $\boldsymbol{W}^{\text{theo}}$ which may be described for all plasma models by

$$\boldsymbol{W}^{\text{theo}} = \sum_{k=1}^{N} c_k \boldsymbol{f}(\mathbf{x}_k)\,. \tag{2}$$

The ith component of the expansion vector $\boldsymbol{f}(\mathbf{x}_k)$ corresponds to the ith measurement and reads according to (1) $f_i(\mathbf{x}_k) = g(n_i, B_i, P_i, a_i; \mathbf{x}_k)$. In general, N such linearly

TABLE I. Parameter of the Connor-Taylor (CT) kinetic plasma models.

CT model	x_1	x_2	x_3	$p(M_j \vert \boldsymbol{W}^{\text{exp}}, \boldsymbol{\sigma}, I)$
(1) Collisionless low-β	x	0	0	4×10^{-12}%
(2) Collisional low-β	x	y	0	99.7%
(3) Collisionless high-β	x	0	z	0.25%
(4) Collisional high-β	x	y	z	0.025%

independent vectors form a complete basis in the N-dimensional data space and would therefore allow a pointwise reconstruction of the data. This is neither desirable, nor with respect to physics correct, since the corresponding vector of measured energy contents $\boldsymbol{W}^{\mathrm{exp}}$ is corrupted by noise. What we really want is an expansion, truncated at some appropriate upper limit E and describing the physics, while the residual $N - E$ terms in the expansion (2) fit only noise. Further, we would like to identify the plasma physics model which describes the data best. An important topic are single variable scans (e.g., the variation of the energy content as function of the density with all other variables fixed) which constitute a very stringent test on any energy confinement function. Such scans are not directly accessible from published databases. On the other hand, single variable scans are experimentally cumbersome, and expensive experiments have to be performed for each and every input variable of interest. It is therefore highly desirable to extract single variable scans from existing databases by employing improved data analysis techniques. A comprehensive answer to these and other questions is provided by Bayesian probability theory [6,7].

Bayesian probability theory rests on the strict application of two basic rules. The product rule

$$p(A,B \mid I) = p(A \mid I)p(B \mid A,I) = p(B \mid I)p(A \mid B,I) \tag{3}$$

allows one to expand the joint probability $p(A,B \mid I)$ for propositions A and B on the basis of some background information I into conditional probabilities for only one variable. Comparison of the two alternative forms in (3) allows one to infer $p(A \mid B,I)$ from the knowledge of $p(B \mid A,I)$. This is called Bayes theorem. A second rule of probability theory allows one to get rid of variables (marginalize over) in probability densities which may be necessary at intermediate stages of the calculation but do not enter into the distribution of interest. It reads

$$p(A \mid I) = \int p(B \mid I)p(A \mid B,I)\,dB\,. \tag{4}$$

The starting point for the analysis of experimental data is the likelihood function

$$p(\boldsymbol{W}^{\mathrm{exp}} \mid \omega, \boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\sigma}, E, M_j, I) = \left(\frac{\omega}{2\pi}\right)^{N/2} \prod_i^N \sigma_i^{-1} \exp\left\{-\omega \sum_i \left[W_i^{\mathrm{exp}} - \sum_k^E c_k f_i(\mathbf{x}_k)\right]^2 \Big/ 2\sigma_i^2 \right\}. \tag{5}$$

M_j denotes the plasma kinetic model used to generate the expansion vectors $\boldsymbol{f}(\mathbf{x}_k)$. $\boldsymbol{\sigma}$ is the vector of experimental uncertainties associated with the measurement $\boldsymbol{W}^{\mathrm{exp}}$. The uncertainty in the energy content $\boldsymbol{W}^{\mathrm{exp}}$ contains the direct distributions from the diamagnetic measurement as well as indirect contributions from the finite precision in the input variables (n, B, P, a). Both contributions have been estimated to the best of our knowledge. To allow for possible deviations from the true errors we introduce an overall correction factor ω. The Bayesian analysis yields *a posteriori* $\omega^{-0.5} \approx 0.8$ for the most probable model and the optimum expansion order which indicates an overestimation of $\boldsymbol{\sigma}$ of 20%. $\boldsymbol{c}$ are the expansion coefficients and E is as before the expansion order. We illustrate the type of calculation performed for the case of model comparison, e.g., a probabilistic answer to the question which one of the kinetic plasma models discussed by Connor and Taylor is most likely on the basis of the data. To this end we need $p(M_j \mid \boldsymbol{W}^{\mathrm{exp}}, \boldsymbol{\sigma}, I)$. By application of Bayes theorem we obtain

$$p(M_j \mid \boldsymbol{W}^{\mathrm{exp}}, \boldsymbol{\sigma}, I) = \frac{p(M_j \mid \boldsymbol{\sigma}, I)}{p(\boldsymbol{W}^{\mathrm{exp}} \mid \boldsymbol{\sigma}, I)} \times p(\boldsymbol{W}^{\mathrm{exp}} \mid M_j, \boldsymbol{\sigma}, I)\,. \tag{6}$$

Comparison of model M_k to model M_j leads to the odds ratio

$$\frac{p(M_j \mid \boldsymbol{W}^{\mathrm{exp}}, \boldsymbol{\sigma}, I)}{p(M_k \mid \boldsymbol{W}^{\mathrm{exp}}, \boldsymbol{\sigma}, I)} = \frac{p(M_j \mid \boldsymbol{\sigma}, I)}{p(M_k \mid \boldsymbol{\sigma}, I)} \frac{p(\boldsymbol{W}^{\mathrm{exp}} \mid M_j, \boldsymbol{\sigma}, I)}{p(\boldsymbol{W}^{\mathrm{exp}} \mid M_k, \boldsymbol{\sigma}, I)}\,. \tag{7}$$

The first term on the right-hand side of (7) is called the prior odds. It expresses the preference for one model over another prior to the analysis of the data. Clearly, we shall put this factor equal to unity. The second term is called the Bayes factor and is the ratio of marginal likelihoods conditional on models M_j and M_k, respectively. $p(\boldsymbol{W}^{\mathrm{exp}} \mid M_j, \boldsymbol{\sigma}, I)$ is obtained upon application of (4)

$$p(\boldsymbol{W}^{\mathrm{exp}} \mid M_j, \boldsymbol{\sigma}, I) = \sum_E p(E \mid M_j, I) \int d\omega\, d\boldsymbol{c}\, d\boldsymbol{x} \times p(\boldsymbol{W}^{\mathrm{exp}}, \omega, \boldsymbol{c}, \boldsymbol{x} \mid E, M_j, \boldsymbol{\sigma}, I)\,. \tag{8}$$

Expansion of the integrand in (8) by application of (3) and omission of all irrelevant logical conditionings [e.g., $p(\boldsymbol{c} \mid \boldsymbol{x}, E, M_j, \boldsymbol{\sigma}, I)$ does not depend on $\boldsymbol{\sigma}$] leads to

$$p(\boldsymbol{W}^{\mathrm{exp}} \mid M_j, \boldsymbol{\sigma}, I) = \sum_E p(E \mid M_j, I) \int d\omega\, d\boldsymbol{c}\, d\boldsymbol{x}\, p(\boldsymbol{W}^{\mathrm{exp}} \mid \omega, \boldsymbol{c}, \boldsymbol{x}, \boldsymbol{\sigma}, E, M_j, I) p(\omega \mid I) \times p(\boldsymbol{x} \mid E, M_j, I) p(\boldsymbol{c} \mid \boldsymbol{x}, E, M_j, I)\,. \tag{9}$$

The first term in the integrand of (9) is the already known likelihood function (5). The other three terms are prior probabilities for ω, $\boldsymbol{x}$, and $\boldsymbol{c}$, respectively. For the scale variable ω we use Jeffreys' prior $p(\omega \mid I) = 1/\omega$ [8]. For $\boldsymbol{c}$ we take an uninformative prior, $p(\boldsymbol{c} \mid \boldsymbol{x}, E, M_j, I)$, which is constant within an allowed region defined by

$\sum_i[\sum_k c_k f_i(\mathbf{x}_k)]^2/\sigma_i^2 \leq r^2$ of volume twice as large as given by Bessel's inequality for the exponent of the likelihood function (5). This describes our minimal prior knowledge. The prior has no influence on parameter estimation. It merely introduces a constant factor relevant for model comparison. Finally, for $\boldsymbol{x}$ we choose an uninformative prior derived from transformation invariance arguments. The $\boldsymbol{c}$ and ω integrations in (9) can be performed analytically [9] while the remaining x integration is carried out with Markov-chain Monte Carlo techniques.

The data which we have used in our calculations are the 153 $\iota \approx 1/3$ W7-AS data from the international stellarator energy confinement data base [10] (ι: rotational transform). We have selected the $\iota \approx 1/3$ data only since the single variable scans which we present further down, have been performed at this value of the rotational transform. The limitation has the further advantage that we get rid of an additional, dimensionless, and ill-defined variable which would otherwise have to be considered in the ansatz (2). Odds ratios obtained from (7) are converted back to probabilities using $\sum_j p(\boldsymbol{W}^{\text{exp}} \mid M_j, \boldsymbol{\sigma}, I) = 1$. The resulting model probabilities are depicted in the last column of Table I. We see that the $\iota \approx 1/3$ W7-AS data are best described by the collisional low beta Connor Taylor model. The high beta models follow in second and third place with much lower probability. The collisionless low beta model is clearly inappropriate to describe the transport physics in W7-AS.

The terms $p(\boldsymbol{W}^{\text{exp}} \mid M_j, \boldsymbol{\sigma}, E, I)$ arising in (9) can be inverted using Bayes theorem to obtain the probability for the expansion order $p(E \mid \boldsymbol{W}^{\text{exp}}, \boldsymbol{\sigma}, M_j, I)$ in the light of the data and a particular model. This quantity is displayed for the most probable plasma model, the collisional low beta case, in Fig. 1 as full circles. The error bars indicate Monte Carlo integration uncertainties. The open squares, with associated error bars, depict the misfit between data and model prediction (2) as a function of the expansion order. We observe quite a typical behavior. Up to three terms in the expansion (2) lead to a rapid decrease of the misfit. Though it does decrease further monotonically with increasing expansion order the probability for a given E decreases rapidly, so that contributions of higher expansion orders become very small. This is a demonstration of Occam's razor automatically included in Bayesian theory. This principle dictates that a simpler model should be preferred unless a more complicated one leads to a substantially better fit to the data. Note that the present optimum three term expansion reduces the misfit to about 65% of its initial value. The possibility for such an effect has previously been pointed out by Kaye *et al.* [2].

There is, however, a much more demanding test of the present semiempirical theory of global transport. Such a test is provided by a comparison of measured single variable scans to the predictions from the present theory. The latter is the expectation value $\langle W \mid \boldsymbol{W}^{\text{exp}}, M_j, \boldsymbol{\sigma}, \boldsymbol{v}, I\rangle$ of energy content obtained from averaging W over

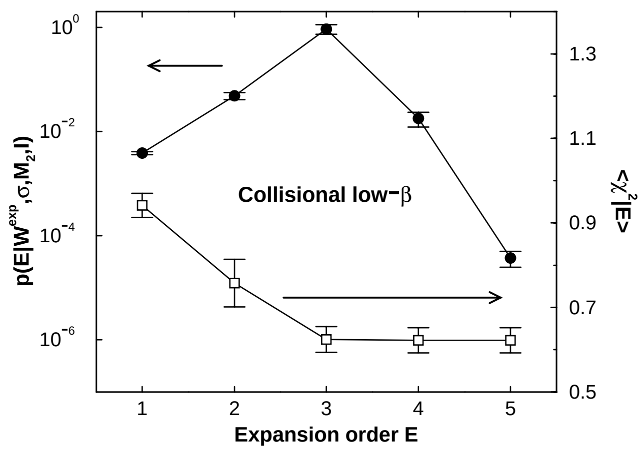

FIG. 1. The maximum of the probability of expansion order E in Eq. (5) is obtained for $E = 3$ (full circles, left scale). Higher order terms do not further reduce the misfit with E (open squares, right scale). Both graphs were obtained for the collisional low beta case.

$p(W \mid \boldsymbol{W}^{\text{exp}}, M_j, \boldsymbol{\sigma}, \boldsymbol{v}, I)$. The additional condition $\boldsymbol{v}$ specifies the "input data vector" $\boldsymbol{v}^T = (n, B, P, a)$ at which the energy content measurement was performed. $\boldsymbol{W}^{\text{exp}}, M_j$, and $\boldsymbol{\sigma}$ have still their previous meaning. The confidence ranges were calculated from $\sigma_{\langle W\rangle} = \sqrt{\langle W^2\rangle - \langle W\rangle^2}$, with $\langle W^2\rangle = \langle W^2 \mid \boldsymbol{W}^{\text{exp}}, M_j, \boldsymbol{\sigma}, \boldsymbol{v}, I\rangle$ and $\langle W\rangle = \langle W \mid \boldsymbol{W}^{\text{exp}}, M_j, \boldsymbol{\sigma}, \boldsymbol{v}, I\rangle$. In the following we will show the result for density and power scans obtained on the one hand from the present theory and on the other hand from experiments at W7-AS. Because these data are not included in the stellarator confinement data base $\boldsymbol{W}^{\text{exp}}$ the analysis is based on, this test shows the predictive power of our approach.

The full circles in Fig. 2 represent experimental results for the density scan. Representative error bars signify the precision level of these data. The continuous curve depicts the result of the present semiempirical theory along with the confidence range indicated by the gray shaded area. The stellarator energy confinement data base is represented by the open circles which are spread all over since they were obtained for various settings of the variables (B, P, a). The histogram at the base line indicates the number of shots at the respective density and gives an impression about the range our result for the single variable scan is best supported by the data base. Last but not least the dashed curve represents the density dependence as inferred from an unrestricted single power law conventional least-squares fit resting on the same data $\boldsymbol{W}^{\text{exp}}$ as our Bayesian result. It does only hit the progression of the data at two points, while staying out of the data scatter (of the full circles) most of the time. Within the density range of the single variable scan the prediction of the semiempirical theory runs straight through the data and exhibits clearly the previously supposed density saturation [10,11], which can never be obtained by a single power law term at all. Outside this range the data set $\boldsymbol{W}^{\text{exp}}$ is too sparse, which is reflected in the

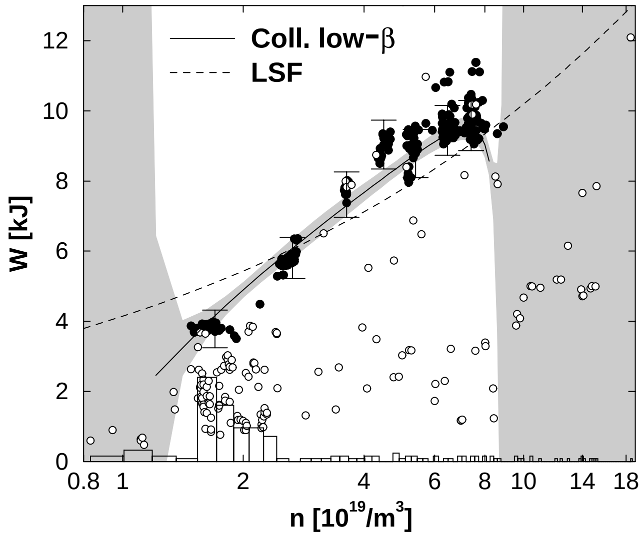

FIG. 2. Experimental results of a single variable density scan (full circles) compared to the predictions of the present semiempirical theory (continuous line, shaded area represents the error) for $B = 2.5$ T, $P = 0.45$ MW, $a = 0.176$ m. The input data (open circles) are shown regardless of additional variations in B, P, a, and are therefore spread all over. The histogram accounts for their distribution over the density axis. A least-squares fit (LSF) of the input data would yield $W \sim n^{0.39}$ (dashed line).

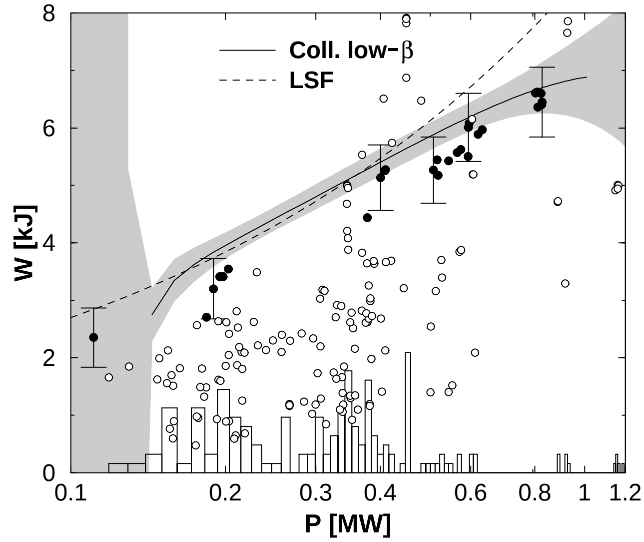

FIG. 3. Same as Fig. 2 for a single variable power scan with $n = 2.4 \times 10^{19}$ m^{-3}, $B = 2.5$ T, $a = 0.176$ m. The histogram is again with respect to the input data (open circles and bars). Our result is represented by the solid line with the shaded area as the error. A least-squares fit of the input data would yield $W \sim P^{0.5}$ (dashed line).

rapidly widening error band. In contrast to the robust but erroneous power law scaling the present theory indicates where the extrapolation becomes unreliable. It might be an unfortunate but honest conclusion that an extrapolation beyond the parameter regime supported by the data base is not possible. However, one has to consider that the seeming predictability of the commonly used power law scaling performs even worse, producing an ever increasing function which misses the saturation entirely. Note that the comparison between the single variable scan and the prediction of our analysis holds on absolute scales. Neither in Fig. 2 nor in the subsequent Fig. 3 adjustable scale parameters are necessary. This means that experiments in W7-AS have an impressive reproducibility.

Figure 3 finally displays a similar comparison for a power scan in W7-AS. Again the semiempirical theory shown as the continuous line predicts the measured energy content—on an absolute scale—within experimental error and corroborates the experimentally observed power degradation. The dashed curve is from a power law fit and is, as for the density scan, convex while the present model shows concave dependence in both cases.

In summary, form free dimensionally exact energy confinement functions derived from data of the international stellarator data base and optimized according to the rules of Bayesian probability theory have identified out of the set of four Conner-Taylor models the collisional low beta one to be the most probable plasma physics model for W7-AS. Moreover, single variable scans were reproduced in quantitative agreement with experiments. The result of a single variable scan is therefore already hidden in the data obtained for arbitrary variable choices and can be extracted from the latter by a proper data analysis. The presented approach has two major advantages over the traditional best-fit approach: Besides its predictive power even on a quantitative level it indicates where the prediction can be trusted. Finally, the importance of this method is far beyond plasma confinement. The possibility to establish a semiempirical theory with controlled prediction properties may well be interesting in other fields of physics.

The authors acknowledge discussions of errors associated with the energy content measurements with R. Jänicke, V. Erckmann, F. P. Penningsfeld, U. Stroth, and F. Wagner. One of us (V. Dose) is indebted to D. Pfirsch for discussions on the Connor-Taylor paper.

[1] J. Hugill and J. Sheffield, Nucl. Fusion **18**, 15 (1978).
[2] S. M. Kaye *et al.,* Phys. Fluids B **2**, 2926 (1990).
[3] J. G. Cordey *et al.,* Plasma Phys. Controlled Nucl. Fusion Res. **3**, 443 (1991).
[4] J. W. Connor and J. B. Taylor, Nucl. Fusion **17**, 1047 (1977).
[5] J. P. Christiansen, J. G. Cordey, and K. Thomsen, Nucl. Fusion **30**, 1183 (1990).
[6] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation,* (Springer Press, Berlin, 1988).
[7] *Kendall's Advanced Theory of Statistics*, *Bayesian Inference,* edited by A. O'Hagan (John Wiley Sons, New York, 1994), 1st ed., p. 293ff.
[8] H. Jeffreys, *Theory of Probability*, (Oxford University Press, Oxford, 1961), 3rd rev. ed.
[9] Details will be published elsewhere.
[10] U. Stroth *et al.,* Nucl. Fusion **36**, 1063 (1996).
[11] U. Stroth, Plasma Phys. Controlled Fusion **40**, 9 (1998).