Optimization of plasma diagnostics using Bayesian probability theory

H. Dreier*, R. Fischer†, A. Dinklage*, M. Hirsch* and P. Kornejew*

*Max-Planck-Institut für Plasmaphysik, EURATOM-Association,
Wendelsteinstr. 1, 17491 Greifswald, Germany,
†Boltzmannstr. 2, 85748 Garching, Germany

Abstract. The diagnostic set-up for Wendelstein 7-X, a magnetic fusion device presently under construction, is currently in the design process to optimize the outcome under given technical constraints. Compared to traditional design approaches, Bayesian Experimental Design (BED) allows to optimize with respect to physical motivated design criterions. It aims to find the optimal design by maximizing an expected utility function that quantifies the goals of the experiment. The expectation marginalizes over the uncertain physical parameters and the possible values of future data. The approach presented here bases on maximization of an information measure (Kullback-Leibler entropy). As an example, the optimization of an infrared multichannel interferometer is shown in detail. Design aspects like the impact of technical restrictions are discussed.

Key Words: Plasma diagnostics, diagnostic design, optimization, Kullback-Leibler entropy.

INTRODUCTION

Nuclear fusion in high temperature plasmas is considered to be one of the most promising candidates for future energy sources. Magnetic fusion devices like the Wendelstein 7-X stellarator, presently under construction, provide data from complementary and redundant measurements. The design of plasma diagnostics is a typical task to be resolved along the preparation of fusion experiments. The design process has to meet with requirements like highest accuracy of measurements, high resolution, robustness and extensibility, as well as with constraints such as accessibility or economic restrictions. The mathematical framework of Bayesian probability theory offers the tools to cover these requirements consistently and provides a quantification of the design result, which allows to compare different diagnostic set-ups directly.

BAYESIAN EXPERIMENTAL DESIGN

The BED approach is based on decision theory and was proposed by Lindley [1]. An appropriate utility function is chosen first reflecting purpose and costs of an experiment. For the quantification of the utility of a set-up the Kullback-Leibler distance U_{KL} is used as a measure for the information gain from the experiment. It compares the knowledge - or better: ignorance - on a quantity of interest ϕ before a measurement with the knowledge after data D are taken:

$$U_{\mathsf{KL}}(D,\eta) = \int P(\phi|D,\eta)\cdot\log\left(\frac{P(\phi|D,\eta)}{P(\phi)}\right)\mathrm{d}\phi \tag{1}$$

The uncertainties are encoded using probability density functions P. The utility function U_{KL} depends both on the data D and the design parameters η, which are to be optimized.

Marginalization over the range of expected data yields the expected utility function EU,

$$EU(\eta) = \int P(D|\eta)U_{\mathsf{KL}}(D,\eta)\mathrm{d}D, \tag{2}$$

now only a function of the design parameters η. In Bayesian Experimental Design the EU is maximized with respect to η.

Using Bayes theorem the EU is given by

$$EU(\eta) = \iint P(\phi)P(D|\phi,\eta)\cdot\log\left(\frac{P(D|\phi,\eta)}{P(D|\eta)}\right)\mathrm{d}\phi\ \mathrm{d}D, \tag{3}$$

where the forward function in the likelihood $P(D|\phi,\eta)$ can be understand as a "virtual" diagnostic. The evidence $P(D|\eta)$ is calculated by

$$P(D|\eta) = \int \mathrm{d}\phi\ P(D|\phi,\eta)\cdot P(\phi). \tag{4}$$

In the conventional approach of BED $P(\phi)$ is a prior function for all unknown parameters. The design of the diagnostic is optimized for the complete range of ϕ. If one is interested in an optimal diagnostic for a subspace of the parameters only, we propose to modify the EU by an appropriate weighting function (see below).

The expected utility function as formulated here is an absolute measure for the information gain, it is expressed in *bit* by using the base-2 logarithm.

A overview on classical and Bayesian design is given in [2]. A more detailed discussion of the BED approach can be found in [3]. Previous results for fusion diagnostics are shown in [4] and [5].

Limited range of interest

In some cases one is interested in an optimal design for a subset of the data only. Therefore, we propose to modify the expected utility by an appropriate weighting function. In experimental design, the utility function $U(D,\eta)$ is the mathematical representation of preferences over an alternative set of designs η with uncertain output D. We choose the Kullback-Leibler divergence to quantify the information gain for the data D measured by design η. If we are not interested in all possible outcomes of the experiment, we will set the utility for those outcomes to zero. Since the outcome D of an experiment is usually subject to noise, we might want to specify our interest in the parameter space ϕ. Let us assume $P_{\mathsf{W}}(\phi)$ represents a subset and weighting of the parameter space for which we want to have an optimal design. Then,

$$P_{\mathrm{W}}(D|\eta) = \int \mathrm{d}\phi \;\; P(D|\phi,\eta)\cdot P_{\mathrm{W}}(\phi) \tag{5}$$

is the modified weighting function representing all data generated by the parameters η which we are interested in. Please note, that, due to the measurement uncertainty, the data might be experimental outcome of parameter values not of interest.

This changes the expected utility to

$$EU(\eta) = \int \mathrm{d}D \, P_{\mathrm{W}}(D|\eta) \int \mathrm{d}\phi \frac{P(\phi)P(D|\phi,\eta)}{P(D|\eta)} \cdot \log\left(\frac{P(D|\phi,\eta)}{P(D|\eta)}\right), \tag{6}$$

where one now has two expressions for the parameters of interest ϕ: On one hand, $P(\phi)$ is the prior distribution for ϕ, describing the uncertainty about the parameters. $P_{\mathrm{W}}(\phi)$ on the other hand defines the interest on a subset of ϕ by generating the weighting function $P_{\mathrm{W}}(D|\eta)$ (eqn. 5).

The effect can be demonstrated on a simple academic example: We assume an experiment with two different states, ϕ_1 and ϕ_2. The goal of the experiment is to identify the the state. The probability to measure the state is given by the relation:

$$\begin{aligned} P(D=\phi_1|\phi_1) &= \eta; \qquad P(D=\phi_2|\phi_1) = 1-\eta \\ P(D=\phi_2|\phi_2) &= \eta; \qquad P(D=\phi_1|\phi_2) = 1-\eta, \end{aligned} \tag{7}$$

where $\eta: \; (0 \leq \eta \leq 1)$.

If one is interested in identifying ϕ_1 only, the range of interest is given by $P_{\mathrm{W}}(\phi_1) = 1$, $P_{\mathrm{W}}(\phi_2) = 0$. Then,

$$P_{\mathrm{W}}(D|\eta) = P(D|\phi_1,\eta). \tag{8}$$

For the prior knowledge about ϕ we assume to be totally ignorant: $P(\phi_1) = P(\phi_2) = \frac{1}{2}$. This leads to

$$\begin{aligned} P(D|\eta) &= 1/2\cdot[P(D|\phi_1,\eta)+P(D|\phi_2,\eta)] \\ &= 1/2, \end{aligned} \tag{9}$$

the evidence is independent of D and η.

The utility U_{KL} (eqn. 1) is then given by

$$\begin{aligned} U_{\mathrm{KL}}(D,\eta) &= \sum_{\mathrm{i}=1}^{2} \frac{P(D|\phi_{\mathrm{i}},\eta)\cdot 1/2}{1/2} \log\left[\frac{P(D|\phi_{\mathrm{i}},\eta)}{1/2}\right] \\ &= \sum_{\mathrm{i}=1}^{2} P(D|\phi_{\mathrm{i}},\eta)\log\left[2P(D|\phi_{\mathrm{i}},\eta)\right], \end{aligned} \tag{10}$$

therefore the expected utility becomes

$$\begin{aligned} EU(\eta) &= \sum_{\mathrm{j}=1}^{2} P(D_{\mathrm{j}}|\phi_1,\eta) \;\sum_{\mathrm{i}=1}^{2} P(D_{\mathrm{j}}|\phi_{\mathrm{i}},\eta)\log\left[2P(D_{\mathrm{j}}|\phi_{\mathrm{i}},\eta)\right] \\ &= (1-\eta)\log\left[2(1-\eta)\right] + \eta\log(2\eta). \end{aligned} \tag{11}$$

This expression provides the results as expected:

- For $\eta = 0.5$ (maximum measurement uncertainty) one gets $EU = 0$, because the states cannot be resolved.
- The maximum information gain is obtained by $\eta = 0$ (no uncertainty) and $\eta = 1$ (exact the inverse result), the expected utility is $EU_{\mathrm{max}} = \log 2$.
- In case of the base-2 logarithm the maximum information gain is $\ln 2 = 1$ *bit*, which is the correct information gain from a two state experiment.

This academic example shows that it might be useful to choose an appropriate expected utility function by a weighting function if one is interested only in a subset of the data space.

DESIGN OF A MULTICHANNEL INTERFEROMETER

An interferometer provides a precise and robust measurement of the electron density of a plasma. The detected phase shift of a probing laser beam is proportional to $\lambda \int n(r)\mathrm{d}l$, where λ is the wavelength of the probing beam and $n(r)$ the electron density distribution, integrated along the line of sight $r(l)$. An interferometer is used for plasma control and, using several beams, for reconstruction of density profiles.

Design parameters are given by the geometrical position of the probing laser beams. Although not all positions are feasible due to the technical restrictions of the plasma vessel, it is useful to optimize the diagnostic without these boundary conditions first.

For the creation of virtual data a parametrized density function is used:

$$n(r) = \phi_1 \cdot \frac{1 + \phi_4 \cdot (r/r_{\mathrm{max}})^2}{1 + \left(\frac{(r/r_{\mathrm{max}})^2}{\phi_2^2}\right)^{\phi_3}} \tag{12}$$

The physical parameters $\phi_1 \ldots \phi_4$ represent the maximum density, position of the edge gradient, steepness and bulge of the density distribution.

It has been shown that the error statistics of the measurement has crucial impact on the expected utility [5]. For the examples presented here, a constant error level σ is chosen which is in the order of a few percent of typical interferometer data. For the error statistics a Gaussian distribution is assumed, the likelihood is given by:

$$P(D|\phi,\eta) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{(\lambda \int n(r)\mathrm{d}l - D)^2}{2\sigma^2}\right] \tag{13}$$

The prior distributions are assumed to be uniformly distributed for every parameter ϕ_i. Because we are interested in the whole parameter range, the conventional ansatz for BED can be used (eqn. 3), so that $P_W(D|\eta) = P(D|\eta)$.

For parametrization of the line of sight, which was the target of the optimization process, two angles represent the start (η_1) and end point (η_2) of the chord on a circumventing circle around the plasma. The expected utility is determined with respect to these angles (cf. Fig 1 lower panels). η_1 and η_2 are exchangeable, resulting in a symmetric expected utility. For beams which do not cross the plasma, the expected utility is zero.

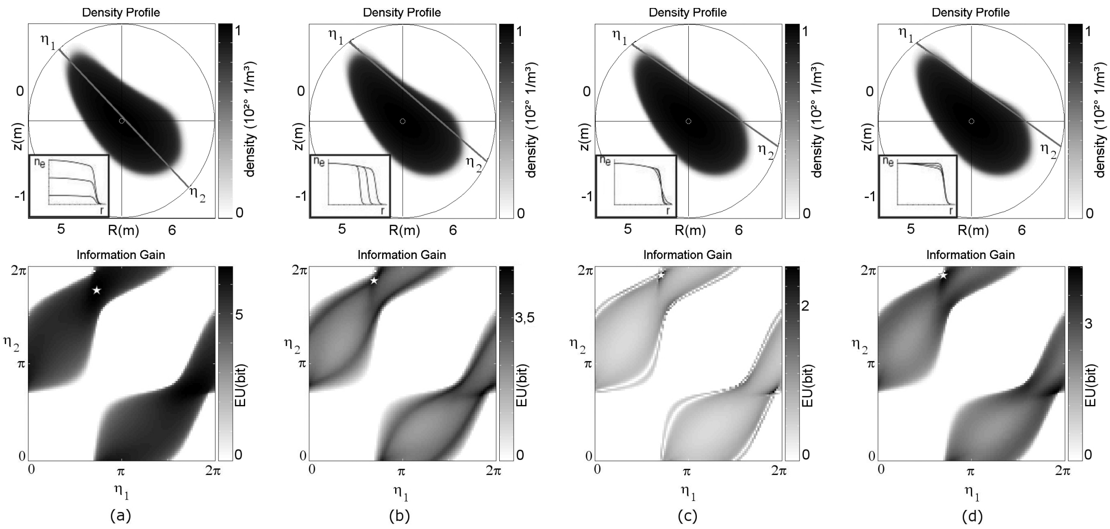

FIGURE 1. Optimized chords for interferometer at W7-X (toroidal angle 195°, upper row) and expected utility (lower row) for estimation of maximum density (a), gradient position (b), steepness (c) and bulge (d). The star symbol in the *EU* plot corresponds to the beam line in the density plot. The data space is generated by a variation of (a) $\phi_1 = 0 \ldots 5 \times 10^{20}\mathrm{m}^{-3}$, (b) $\phi_2 = 0.6 \ldots 0.95$, (c) $\phi_3 = 1 \ldots 30$ and (d) $\phi_4 = -1 \ldots 0$. The insets in the upper row show the corresponding density profile variation where the maximum ordinate is $n_e = 1 \times 10^{20}\mathrm{m}^{-3}$ and the abscissas are effective radii (r_{eff}/a).

Figure 1 shows the design results for four single beams, optimized to estimate the parameters of interest ϕ_1 - ϕ_4. Therefore, the *EU* was calculated for a single parameter, the others were kept constant, respectively . The beam line in the density plots (upper row) corresponds to the maximum of the expected utility. The *EU* is displayed in figures of the two angles η_1 and η_2 (lower row).

As one can see from the results, the calculated *EU* is determined by the shape of the plasma. E.g. for the estimation of maximum density (Fig. 1(a)) a beam traversing the plasma center on the longest possible way yields the maximum of the *EU*. The optimal beam line in this case represents the maximum signal-to-noise ratio (SNR) chord. Since the effects of the other parameters are most distinct at the plasma edge, the resulting reconstruction has maximum information gain for sightlines traversing the edge region.

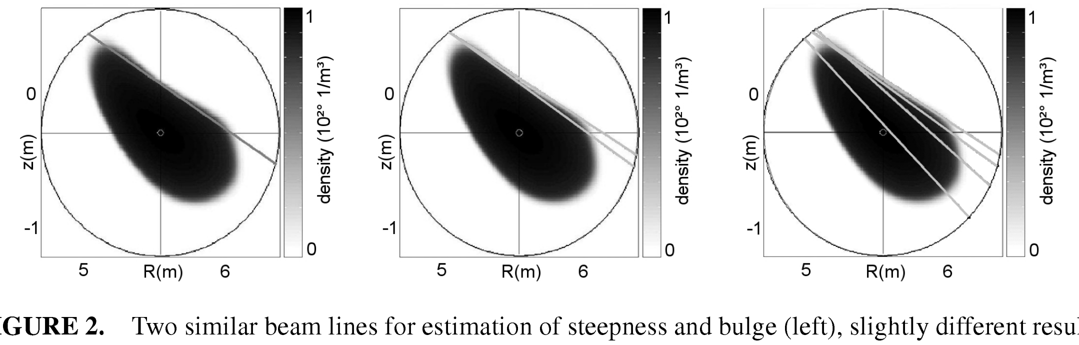

FIGURE 2. Two similar beam lines for estimation of steepness and bulge (left), slightly different result after combined optimization (middle), complete 4-channel interferometer (right).

For the estimation of steepness (ϕ_3) and bulge (ϕ_4) of the density distribution (Fig. 1 (c) and (d)) a similar beam line results from the optimization. So, as a next step, two beam lines are optimized in one step by integrating over both parameters of interest (ϕ_3 and ϕ_4). Figure 2 shows the result of the combined optimization of the two similar beam lines. The combined optimization led to a slightly different result compared with the single beam optimization.

As mentioned before, no technical boundary conditions were taken into account so far. Figure 3 indicates accessible chords in figures of the design parametrization chosen. For the accessible regions optimal beam line position resulting from the maximum of the EU is identical for all parameters ϕ_i. The effect of the technical restrictions can be quantified and compared to unrestricted access to the plasma.

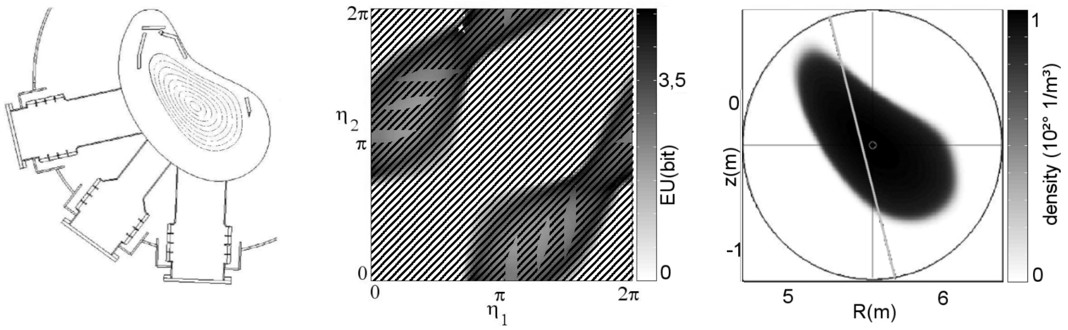

FIGURE 3. Technical boundary conditions: port system on W7-X at a toroidal angle of 195° (left); EU with allowed regions (lightened areas, middle); optimal beam line for all parameters ϕ_i (right).

CONCLUSION AND OUTLOOK

As an extension to the conventional approach of Bayesian experimental design approach an ansatz for a limited parameter range of interest is presented. Future work will adapt and evaluate this for the design of plasma diagnostics, so that the design can be focused on physical interesting and relevant scenarios.

Furthermore, results for the design of a multichannel interferometer are shown. An extension to eight channels and the consideration of more boundary conditions is planned as a next step.

ACKNOWLEDGMENTS

We want to thank D. Dodt for many helpful and inspiring discussions.

REFERENCES

1. Lindley, D. V., 1972. *Bayesian Statistics - A Review*, SIAM (Philadelphia).
2. K. Chaloner, I. Verdinelli, *Bayesian Experimental Design: A Review*, Stat. Sci. **10**, No. 3, pp. 273-304 (1995)
3. R. Fischer, *Bayesian Experimental Design - Studies for Fusion Diagnostics*, in Bayesian inference and Maximum Entropy Methods in Science and Engineering, AIP Conference Proc. **735** (2004), pp. 76-83

4. R. Fischer, H. Dreier, A. Dinklage, B. Kurzan and E. Pasch, *Integrated Bayesian Experimental Design*, in Bayesian inference and Maximum Entropy Methods in Science and Engineering, AIP Conference Proc. **803** (2005), pp. 440-447
5. H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew, E. Pasch, *Bayesian Design of Diagnostics: Case Studies for Wendelstein 7-X*, Fusion Sci. Technol. (Aug. 2006) at press.