Bayesian Design of Diagnostics: Case Studies for Wendelstein 7-X

H. Dreier[1], A. Dinklage[1], R. Fischer[2], M. Hirsch[1], P. Kornejew[1], E. Pasch[1]

Max-Planck-Institut für Plasmaphysik,
EURATOM Association,
[1]Teilinstitut Greifswald, D-17491 Greifswald ,
[2]D-85748 Garching, Germany

Please Mail to:

Heiko Dreier
Max-Planck-Institut für Plasmaphysik
Wendelsteinstrasse 1
D-17491 Greifswald
Germany

Tel.: +49 3834 88 - 1888
Fax.:   - 2509
eMail: hdreier@ipp.mpg.de

Bayesian Design of Diagnostics: Case Studies for Wendelstein 7–X

H. Dreier[1], A. Dinklage[1], R. Fischer[2], M. Hirsch[1], P. Kornejew[1], E. Pasch[1]

Max–Planck–Institut f¨ur Plasmaphysik, EURATOM Association,

[1]Teilinstitut Greifswald, D–17491 Greifswald / [2]D–85748 Garching, Germany

Abstract: A general method for the design of diagnostics by means of Bayesian probability theory is outlined. Case studies are discussed for two diagnostics planned for Wendelstein 7-X. First, line of sight optimization for interferometry meeting with physical optimization targets is investigated. The filter transmission bandwidth for Thomson scattering in W7-X is discussed as a second example.

I. PRINCIPLE OF DIAGNOSTIC DESIGN

The set of "start-up" diagnostics for Wendelstein 7-X is oriented to resolve physical issues of stellarator optimization [1]. The diagnostics philosophy of W7-X aims at extendable setups which are to be supplemented along the operation of the device. Planning diagnostics directly lead to the requirement to optimize their respective outcome. In many cases, the optimization is much restricted by technical constraints. But even in those cases, figures of robustness of the chosen design settings are beneficial for the diagnostic strategy.

In this paper a framework for diagnostic design is presented employing the Bayesian probability theory. (For an introduction to the field of Bayesian theory see, e.g., [2].) Part of the approach can be regarded as a virtual diagnostic, which describes the future experimental unit including its error statistics and is able to produce "artificial data". The virtual diagnostic is fed with mock-up quantities of interest and hence it enables the designer to compare the outcome of the virtual diagnostic with its input.

For the design of diagnostics in fusion experiments several figures of merit are possible, depending on the optimization goal. Here physical questions like estimation of the density distribution or the density gradient are possible optimization targets as well as technical issues as the robustness of density measurements, e.g, for controlling purposes. Also the estimation of derived quantities - like gradients needed for modeling of the bootstrap current or of the radial electric field - might be considered to be the goal of a diagnostic design.

A. Bayesian Experimental Design

The figure of merit for the optimization is called the utility function $U(D, \eta)$. It depends on the data D and the design parameters η. The data represent the expected data which are to be derived from

1. the expected range of quantities of interest φ (such as a expected density range) and their respective relevance (by means of a 'weight function' allowing to put special interest, e.g., on large density values) – and
2. a diagnostics model which contains the forward function

   $D = f(\varphi, \eta)$

   and the error statistics of all quantities.

Then, an integration over the range of expected data, where the evidence of the data is represented by the probability density function $P(D|\eta)$, yields the expected utility function EU:

$$EU(\eta) = \int dD \ P(D \mid \eta) \ U(D, \eta), \qquad (1)$$

which remains to be a function of the design parameters only.

Diagnostic optimization is the maximization of the expected utility with respect to the design parameters η.

A possible utility function is provided by information theory (see [3]): The so called Kullback–Leibler distance is an absolute information measure which expresses the information gain from the knowledge (or better ignorance) about the parameters of interest φ before a

measurement $P(\varphi)$ to the knowledge on the parameter after a measurement $P(\varphi|D, \eta)$. It is measured in bits, if the base-2 logarithm is used.

$$U_{KL}(D,\eta) = \int d\varphi \; P(\varphi \,|\, D,\eta) \; \log \frac{P(\varphi \,|\, D,\eta)}{P(\varphi)} \qquad (2)$$

The functions $P(\varphi)$ and $P(\varphi|D, \eta)$ are probability density functions encoding the uncertainty of the quantity to be inferred. The term $P(\varphi)$ is the prior probability density, and in data analysis it describes the prior knowledge of the parameter of interest φ (e.g. the density distribution) before the measurement. For the design process the prior probability density function is used to define the range of interest of φ (e.g. densities from $10^{19} - 5 \cdot 10^{20} \, \mathrm{m}^{-3}$).

The posterior distribution, $P(\varphi|D, \eta)$, is the inference for a set of given data. From eqs. (1) and (2) follows

$$EU(\eta) = \iint dD \; d\varphi \; P(D \,|\, \eta) \; P(\varphi \,|\, D,\eta) \; \log \frac{P(\varphi \,|\, D,\eta)}{P(\varphi)} . \qquad (3)$$

$P(\varphi|D, \eta)$, which is unknown, can be replaced by Bayes theorem

$$P(\varphi \,|\, D,\eta) = \frac{P(D \,|\, \varphi,\eta) \cdot P(\varphi)}{P(D \,|\, \eta)}, \qquad (4)$$

yielding

$$EU(\eta) = \iint dD \; d\varphi \; P(D \,|\, \varphi,\eta) \; P(\varphi) \; \log \frac{P(D \,|\, \varphi,\eta)}{P(D \,|\, \eta)}. \qquad (5)$$

Here the likelihood $P(D|\varphi, \eta)$ is the error statistics of a measurement for a given quantity of interest. The term $P(D|\eta)$ is called prior predictive value or evidence of the data. It is given by

$$P(D \,|\, \eta) = \int d\varphi \; P(\varphi) \; P(D \,|\, \varphi,\eta). \qquad (6)$$

For the design of diagnostics, only the prior $P(\varphi)$ and the likelihood $P(D|\varphi, \eta)$ are necessary.

The expected utility must fulfill some simple properties in the limiting cases (a) no data available, (b) data is totally uninformative about the parameters φ, and (c) exact knowledge about the parameters [4]. Without data, the posterior $P(\varphi|D, \eta)$ is equal the prior $P(\varphi)$, which leads to $EU(\eta) = 0$ according to (3). In other words, there is no information gain without data. In case of (b), the experiment is totally ignorant about the chosen parameters φ, so the likelihood $P(D|\varphi, \eta)$ is equal to the evidence $P(D|\eta)$, the expected utility is zero again. Finally, exact knowledge of the parameters (case c) means $P(\varphi) = P(\varphi|D, \eta)$, which also results in $EU(\eta) = 0$, according to (3), because no information was gained by the measurement.

II. CASE STUDIES

A. Single chord interferometry

First design studies were performed for a single-chord interferometer. Different geometries of the plasma shape, from a simple circular to the complex W7–X profile, were used to analyze the performance of the design algorithm and to study limiting cases.

The virtual interferometer was fed with a parameterized electron density function

$$n_e(\varphi_1, \varphi_2) = \varphi_1 \cdot \frac{1 - 0.2 \cdot r_{eff}^2}{1 + \left(\dfrac{r_{eff}^2}{\varphi_2^2}\right)^{20}}, \qquad (7)$$

representing a typical parametric profile function for W7-AS. One parameter was changed according to Fig. 1, respectively. The resulting phase shift $\Delta\Phi = \lambda e^2 / (4\pi\varepsilon_0 m_e) \cdot \int n_e \, dl$ was deteriorated by normally distributed noise with a constant standard deviation σ. The dependence on the design criterion was investigated for two different criteria: estimation of the density amplitude and of the position of the density edge.

As an example, the effect of the choice of the line of sight was investigated; results are shown in Fig. 2. For parametrization of the line of sight, two angles representing the start θ_1 and end point θ_2 of the chord on a circumventing circle span up a plane on which the

expected utility is determined (cf. Fig 2 lower panels). θ_1 and θ_2 are exchangeable resulting in a symmetric plot of the expected utility. For beams which do not cross the plasma, the expected utility is zero.

As a first limiting case, the reconstruction of the density amplitude was examined (see Fig. 1 (left)). A practical application of this case is the use of the interferometer for density control. Figure 2 shows the density distribution for a circular, elliptical and W7–X shaped plasma (upper row). The plot of the expected utility function (lower row) shows a maximum which corresponds to the chords traversing the center of the plasma, as shown for the circular, elliptical and W7–X plasma profile. Provided a constant background noise the highest accuracy to estimate the density amplitude is the chord with the highest signal-to-noise ratio. Obviously, this corresponds to the longest possible way through the plasma, crossing the most dense regions, which are located in the plasma center. In addition, the shape of expected utility plot offers estimates for the effect of alignment on the measurement.

Figure 3 shows the design result for the estimation of the position of the largest density gradient that is of relevance for modeling of, e.g., the bootstrap current or the radial electric field. The best estimation of the parameter of interest is given with a beam through the edge of the plasma, i.e. the region of the gradient position as expected from the profile according to Fig. 1 (right). In addition to the radial position, the largest expected utility is attained for chords which maximize the length of the line of sight in the plasma edge region.

Effect of error statistics

Different types of error statistics result in different experimental designs. To study the effect of the error statistics on the design, the density amplitude was chosen as parameter of interest. Three different cases for measurement errors were analyzed: first, a constant background noise was assumed, i.e. the standard deviation σ was independent from the data value, second, a counting experiment was assumed with a standard deviation $\sigma \propto \sqrt{D}$, and third, a constant signal-to-noise ratio (SNR) with standard deviation $\sigma \propto D$.

Figure 4 shows the results for the expected utility for the estimation of the density amplitude for three different uncertainties. For a constant σ a distinct maximum appears,

corresponding to a chord transversing the plasma center (fig. 2 right). For $\sigma \propto D$, on the other hand, the expected utility is constant as long as the line of sight crosses the plasma. For the case $\sigma \propto \sqrt{D}$ a maximum similar to a constant standard deviation appears, but the expected utility distribution is less peaked, which indicates that the choice of the line of sight is less critical than in the case $\sigma = $ const. The most prominent peak in the EU distribution is given for the case $\sigma = $ const. For $\sigma \propto \sqrt{D}$ the SNR is proportional to $\sqrt{D}$, whereas for $\sigma \propto D$ the SNR and, hence, the expected utility is constant.

The calculations show a significant impact of the type of measurement error on the expected utility distribution. This indicates the relevance of error statistics on the design.

B. Thomson scattering

The measurement of electron temperature and density by a polychromator Thomson scattering set-up is investigated as a second example [5]. The scattering function is detected by four spectral filters. The design of the spectral filters is discussed by the choice of the cut–off wavelengths λ_1 and λ_2 of two spectral filters as shown in Figure 5.

Figure 6 shows results for different expected temperature intervals. The maxima of Figs. 6 (upper row) depend on the range of T_e and resemble the wavelength range covered by significant scattering intensity. The results itself can be used to study the robustness of the design. The lower row of Fig. 6 shows that the error statistics has substantial impact on the expected utility as well. It is the different influence of σ on the optimum expected signal-to-noise ratio for a given temperature range which leads not only to different maxima for the same expected T_e range (cf. Fig. 6 lower row left and center) but even to a qualitatively different behavior. For a constant noise (Fig. 6 lower row right) it appears that the optimum corresponds to $\lambda_1 = \lambda_2$ or $\lambda_{1,2} = \lambda_{Laser}$ or, in other words, to remove one filter.

III. SUMMARY AND OUTLOOK

Bayesian probability theory offers general methods for diagnostic design. The approach reveals aspects which could be missed otherwise. For example, the analysis shows that the expected noise behavior of the signal can affect the choice of sightlines for a single chord interferometer. The design of filter band-widths for a Thomson scattering system depends on a correct description of the error statistics as well.

The case studies will be completed towards a virtual multi-chord interferometer for Wendelstein 7-X and compared to traditional design approaches [6]. The results attained indicate the impact of the specific design goal and the error statistics both for the optimum design and the shape of the expected utility function, where the latter can be interpreted as a figure of robustness of the design chosen.

References:

[1] H.J. Hartfuss, R. Brakel, M. Endler, P. Grigull, J.V. Hofmann, J. Junker, M. Kick, G. Kühner, H. Niedermeyer, H. Ringler, A. Teubel, F. Wagner, and A. Weller, *Diagnostic strategy of the W7--X stellarator*, Rev. Sci. Instruments **68**, No. 2, pp. 1244-1249 (1997)

[2] D.S. Sivia, Data Analysis, A Bayesian Tutorial, Clarendon, Oxford (1996)

[3] K. Chaloner, I. Verdinelli, *Bayesian Experimental Design: A Review*, Stat. Sci. **10**, No. 3, pp. 273-304 (1995)

[4] R. Fischer, *Bayesian Experimental Design – Studies for Fusion Diagnostics,* in Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Proceedings MaxEnt
July 25-30 2004, Garching Germany, published: AIP, Melville, NY, 2004, vol. 735, pp. 76-83

[5] 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, Proceedings MaxEnt
Aug 7-12 2005, San José CA USA, published: AIP, Melville, NY, 2005, vol. 803, pp. 440-447

[6] M.I. Airila, O. Dumbrajs, H.-J. Hartfuss, M. Hirsch, J. Geiger and U. Neuner, S*ightline optimization of the multichannel laser interferometer for W7--X*, Rev. Sci. Instruments **76**, 023501 (2005)
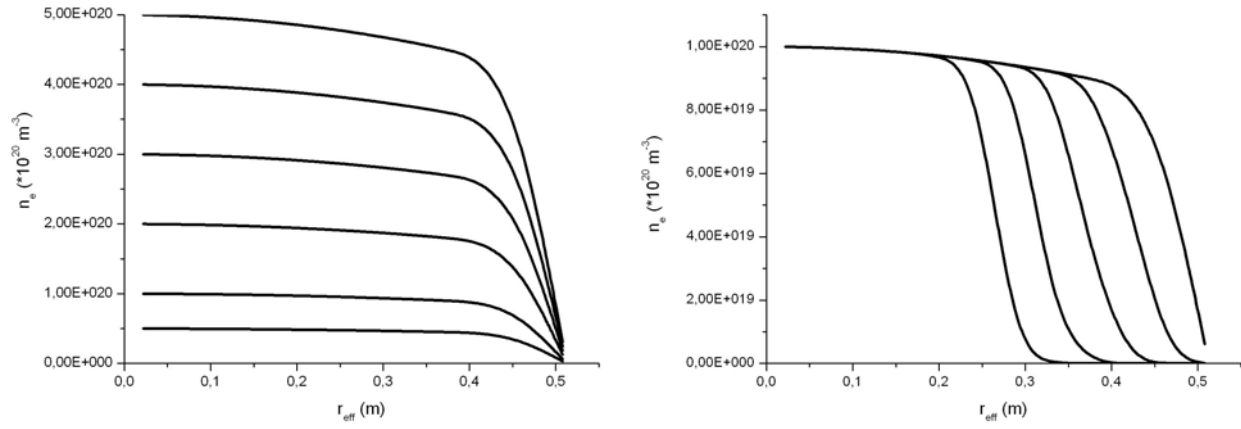
FIG. 1: Density distribution: (left) amplitude φ_1 of density distribution changed; (right) position of density edge φ_2 changed
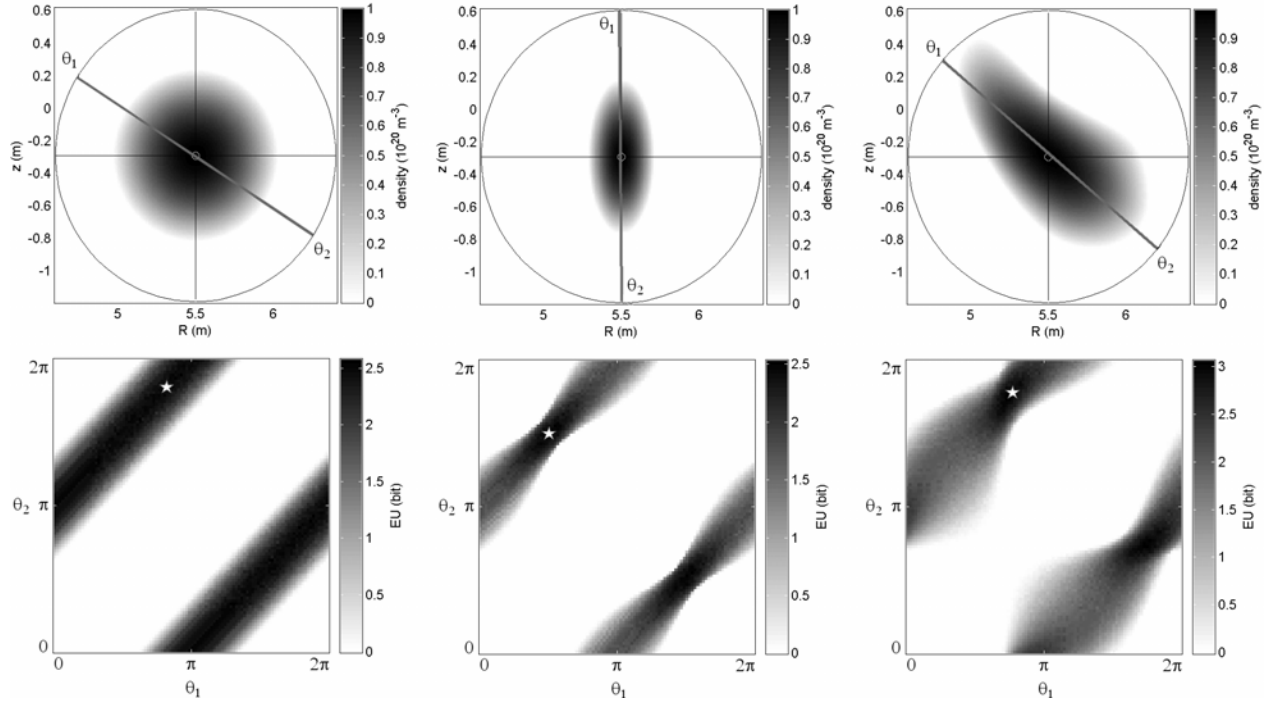
FIG. 2: Density distribution and expected utility for the estimation of the density amplitude, calculated for circular-shaped (left), elliptical-shaped (middle) and W7–X shaped (right, toroidal angle 194°) plasma. The interferometer beam chords in the upper panels correspond to the symbols (star) in the lower panels, respectively. The angles θ_1 and θ_2 parametrize the lines of sight. In regions of EU = 0 the beam does not enter the plasma.
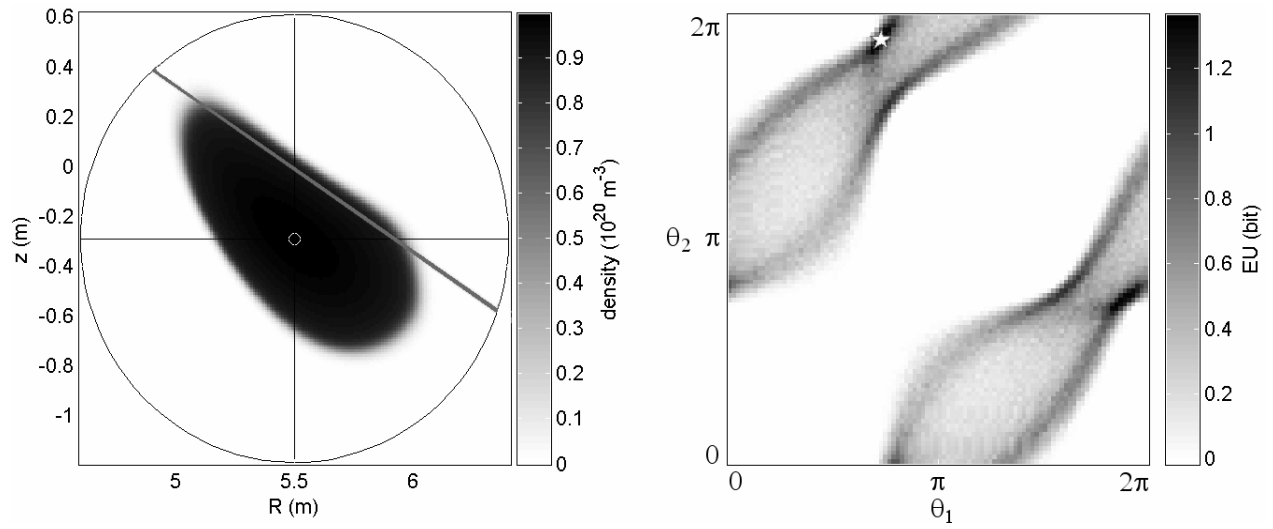
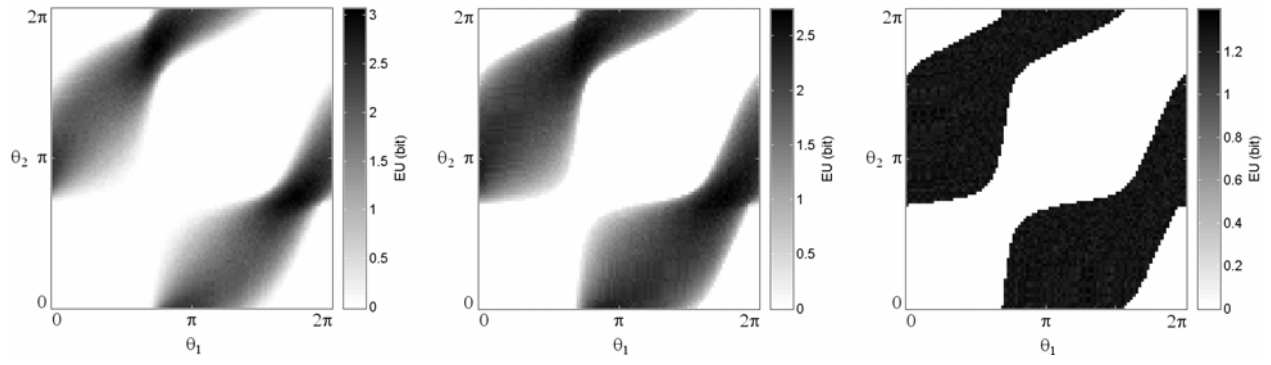FIG. 3: Estimation of the gradient position with one chord: density plot (left), expected utility (right)

FIG. 4: Expected utility for the estimation of the density amplitude for three different uncertainties: $\sigma = $ const (left), $\sigma \propto \sqrt{D}$ (middle), and $\sigma \propto D$ (right).
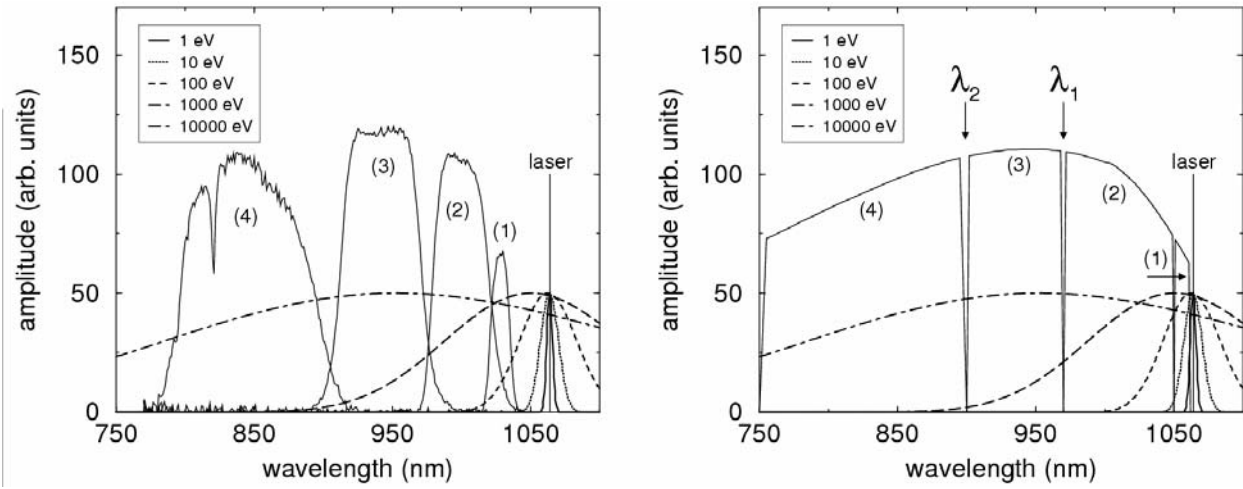
FIG. 5: Transmission curves of spectral filters of a polychromator detector and expected Thomson scattering functions $T_e = 1eV - 10keV$. Left: Set of filters as used for W7–AS Thomson scattering, right: model filters. λ_1 and λ_2 are the cut-off wavelengths of filters allowed to vary for design studies.
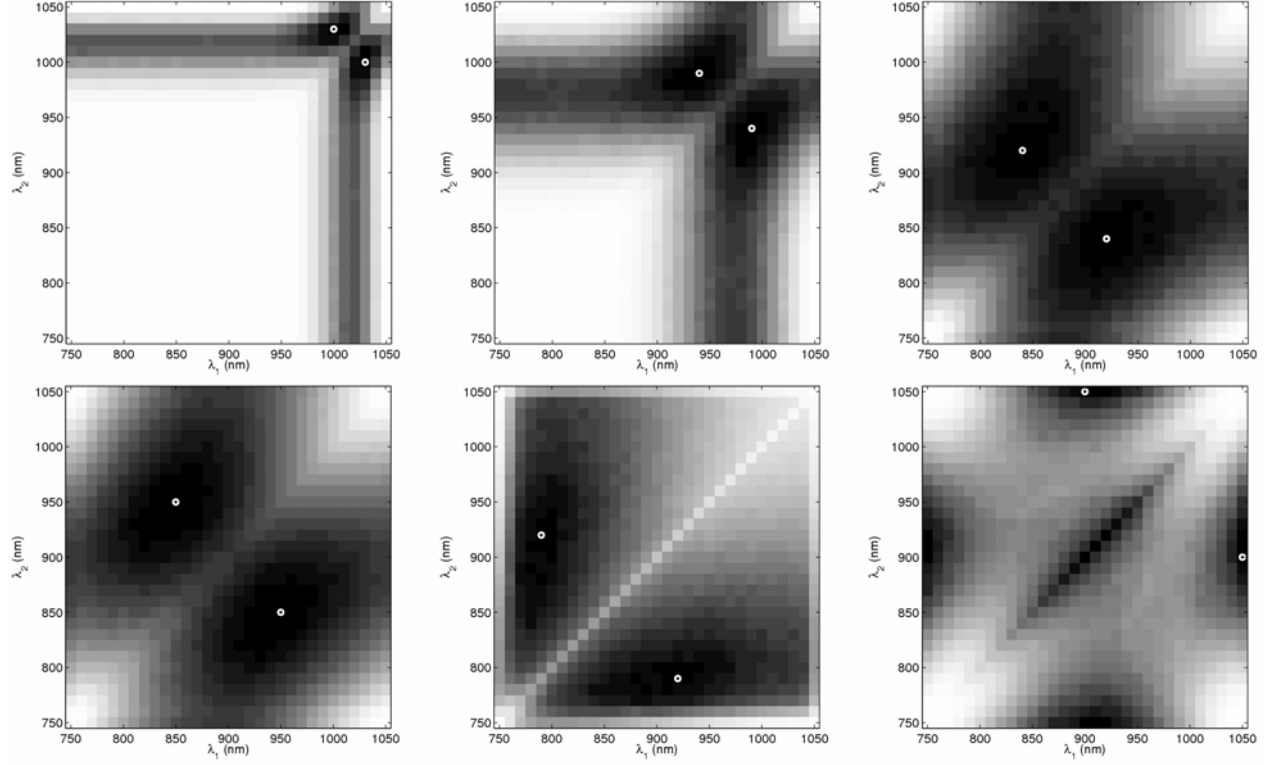
FIG. 6: Expected utility as a function of the cut–off wavelengths λ_1 and λ_2. The upper row shows the dependence on the choice of the parameter range of temperature with the same error statistics ($\sigma \propto \sqrt{D}$): $T_e = 0.01 - 0.25\text{keV}$ (left); $T_e = 0.25 - 1\text{keV}$ (middle); $T_e = 1 - 10\text{keV}$ (right). The lower row shows the dependence on error statistics with the same expected range of temperatures ($T_e = 1 - 10\text{keV}$): $\sigma \propto \sqrt{D}$ (left); $\sigma \propto D$ (center); $\sigma = \text{const}$ (right). The right plot in the upper row is repeated as left plot in the lower row for completeness.