

Heiko Dreier

**Bayes'sches Design von Experimenten: Anwendungen in der  
Kernfusion**

**Bayesian Experimental Design: Applications in Nuclear Fusion**

**IPP 13/8  
Dezember, 2007**

# Bayesian Experimental Design: Applications in Nuclear Fusion

INAUGURALDISSERTATION

zur

Erlangung des akademischen Grades eines

doctor rerum naturalium (Dr. rer. nat.)

an der Mathematisch-Naturwissenschaftlichen Fakultät

der

Ernst-Moritz-Arndt-Universität Greifswald

vorgelegt von

Heiko Dreier

geboren am 09.06.1977

in Bad Saarow

Greifswald, im Mai 2007

Dekan: Prof. Dr. Klaus Fesser

1. Gutachter: Dr. habil. Andreas Dinklage

2. Gutachter: Prof. Dr. Wolfgang von der Linden

Tag der Promotion: 04. Oktober 2007

## Abstract

This work introduces the idea of Bayesian experimental design to the field of high temperature plasma physics and gives examples for the optimisation of diagnostics as well as for the planning of future experiments.

The method of Bayesian experimental design is based on the idea of maximisation of the information gain expected from the future measurement. This information gain is expressed by the Kullback-Leibler distance, an information measure which compares the prior knowledge about the parameters of interest with the state of knowledge about these parameters given the data. Bayesian theory is applied as a tool for the treatment of different probability distributions.

The key element of this method is the *Expected Utility*, which is an absolute measure for the information gain averaged over all possible data outcome. Bayesian experimental design means the maximisation of the Expected Utility with respect to the design parameters. Furthermore, different Expected Utilities from different designs can be compared quantitatively. Physical questions and problems, encoded as parameters of interest, are implemented directly as design criterion. Boundary conditions like technical limitations can be included as well.

The approach is validated using data sets from the Wendelstein 7-AS experiment: The data was analysed and compared with the respective Expected Utility, which itself was calculated independently. A monotonous relationship between Expected Utility and the deviation  $\chi^2$  of the measurement result was found.

As a first application, a multi-channel interferometer at the Wendelstein 7-X stellarator is optimised using Bayesian experimental design. Three different density effects (measurement of the effects of Core Electron Root Confinement, estimation of high confinement regimes and the analysis of pellet injections) were used as design criteria. The results were analysed and compared, a final proposal for the sight line configuration is given and discussed.

Bayesian experimental design can be used as a tool for experimental planning, too. Here, the optimal experimental setting can be found to gain maximum information about a certain problem. For this, also data from previous experiments are incorporated, their respective usefulness can also be expressed by the Expected Utility. As an example, a data set from the stellarator database was examined with respect to the approach of power scaling laws.

It was found out, that a good signal-to-noise ratio is preferred by Bayesian experimental design. Furthermore, a significant impact of the measurement error and the error statistics could be verified. The final design is determined by a combination of these criteria. For the interferometer, the optimal design is given for a beam line configuration where a change of the parameters of interest has maximum impact on the expected data. In the case of experimental planning, the most informative experimental configurations for future measurements is found for settings which are not yet provided by the given database.



# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
1.1	General motivation: Physical design . . . . .	7
1.2	Specific motivation: Nuclear fusion . . . . .	7
1.2.1	The energy problem . . . . .	7
1.2.2	Nuclear fusion as an energy source . . . . .	8
1.2.3	Physical challenges . . . . .	9
1.2.4	Plasma confinement . . . . .	10
1.2.5	The stellarator principle . . . . .	12
1.2.6	Diagnostic approach for W7-X . . . . .	14
1.3	Physical design . . . . .	15
1.3.1	Overview . . . . .	16
<b>2</b>	<b>Bayesian Experimental Design</b>	<b>19</b>
2.1	Bayesian probability theory . . . . .	19
2.2	A brief introduction to information theory . . . . .	22
2.2.1	Hartley information . . . . .	23
2.2.2	Shannon information . . . . .	24
2.2.3	Kullback–Leibler distance . . . . .	25
2.3	Bayesian experimental design . . . . .	26
<b>3</b>	<b>Interferometry at Wendelstein 7-X</b>	<b>29</b>
3.1	Density measurement with interferometry . . . . .	29
3.1.1	The principle: Phase shift and line integration . . . . .	29
3.1.2	Phase shift measurements . . . . .	30
3.2	The interferometry system at Wendelstein 7-X . . . . .	33
3.2.1	A two-color, multi-channel interferometer . . . . .	33
3.2.2	Technical boundary conditions . . . . .	34
3.2.3	Error sources of the interferometry system . . . . .	36
<b>4</b>	<b>Basic Case Studies and Methodological Validation</b>	<b>41</b>
4.1	Previous diagnostic optimisation results for W7-X . . . . .	41
4.2	Case studies: One-beam interferometer . . . . .	43
4.2.1	The basic setting . . . . .	43
4.2.2	Parametrisation of the density distribution . . . . .	45
4.2.3	One-beam design for density model function . . . . .	48
4.2.4	Impact of the error statistics . . . . .	50
4.3	Data analysis for interferometric measurements . . . . .	52
4.3.1	Abel integral and inversion problem . . . . .	52
4.3.2	Analysis of interferometer data from W7-AS . . . . .	55
4.4	Validation of the <i>EU</i> . . . . .	57

4.4.1	Data set and parametric fit . . . . .	59
4.4.2	Comparison with Expected Utility . . . . .	60
<b>5</b>	<b>Design of a Multi-Channel Interferometer</b>	<b>65</b>
5.1	Physical problems of interest . . . . .	65
5.1.1	Core Electron Root Confinement (CERC) . . . . .	65
5.1.2	High confinement operational scenarios . . . . .	69
5.1.3	Density effects of pellet injection . . . . .	72
5.2	Multi-channel interferometer design . . . . .	74
5.2.1	Four-channel interferometer . . . . .	74
5.2.2	Finding the optimal design by comparison . . . . .	81
5.2.3	Influence of magnetic configurations . . . . .	82
5.2.4	Eight-channel interferometer . . . . .	85
5.3	Discussion . . . . .	86
<b>6</b>	<b>Experimental Planning using Bayesian Design</b>	<b>89</b>
6.1	Motivation . . . . .	89
6.2	The principle of data adaptive planning . . . . .	90
6.2.1	General approach for a linear problem . . . . .	90
6.2.2	A one-dimensional example . . . . .	91
6.3	Scaling Laws for Fusion Devices . . . . .	92
6.3.1	Background . . . . .	92
6.3.2	Energy confinement time . . . . .	93
6.3.3	Confinement energy . . . . .	95
6.4	Data adaptive planning for scaling laws . . . . .	96
6.4.1	<i>DAP</i> for W7-AS measurements . . . . .	96
6.4.2	The W7-AS data set . . . . .	97
6.4.3	Experimental planning: The <i>EU</i> for a new data point . . . . .	99
6.4.4	Influence of the measurement error . . . . .	101
6.5	Discussion . . . . .	103
<b>7</b>	<b>Summary</b>	<b>107</b>
<b>A</b>	<b>Basic Code Structure for <i>EU</i> Calculation</b>	<b>113</b>
A.1	Computing the Expected Utility . . . . .	113
A.1.1	Code structure . . . . .	113
A.1.2	Source code . . . . .	114
A.2	Benchmarking . . . . .	117
<b>B</b>	<b>Magnetic Configurations for W7-X</b>	<b>121</b>
<b>C</b>	<b>Detailed Calculation for Data Adaptive Design</b>	<b>123</b>
<b>D</b>	<b>Post-submit Changes</b>	<b>129</b>
	<b>Bibliography</b>	<b>131</b>
	<b>Acknowledgments</b>	<b>138</b>
	<b>Curriculum Vitae</b>	<b>140</b>
	<b>List of Publications</b>	<b>142</b>



# Chapter 1

## Introduction

### 1.1 General motivation: Physical design

Every experiment is designed with respect to the scientific question of interest. An experiment with an optimal design should give the most informative results for the respective scientific problem. Therefore, an effective design method has to implement the scientific question as a design criterion. It must also cover the boundary conditions which may give limitations to the experiment. This work will introduce a general and consistent framework for experimental design which fulfills these demands.

Experimental design in this context covers two sub-problems:

- The optimisation of a diagnostic unit according to a defined set of design parameters is called *diagnostic optimisation*.
- In *experimental planning*, the best experimental configuration has to be found, with respect to results from previous experiments in this context.

It is the goal of this work to introduce a method based on probability theory, which will meet with these criteria.

The work is embedded into the assembly of Wendelstein 7-X, a stellarator type fusion experiment currently under construction. Examples for both applications, experimental planning and diagnostic design, will be presented.

### 1.2 Specific motivation: Nuclear fusion

This section will introduce the physical background of this work: The main issues of nuclear fusion and, in particular, the experimental challenges for plasma confinement are outlined. Therefore, the stellarator concept is explained shortly.

#### 1.2.1 The energy problem

Since the beginning of the industrial age in the 19th century the global energy consumption has increased rapidly, by nearly a factor of ten since 1850 [1]. One of the major challenges of the future is to find energy sources which can satisfy the continuously increasing demand for energy. In particular, the need for carbon dioxide neutral energy sources becomes evident, since recent findings imply that  $CO_2$  emissions result in significant climate effects [2].

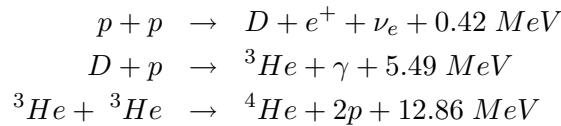
These energy sources have to match three main demands:

1. Sustainability: The necessary resources should be available at long sight.
2. Independency of locally restricted resources to provide global access
3. Avoidance of  $CO_2$  emission and other dangerous waste products

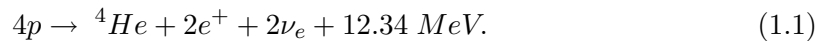
Furthermore, the energy source has to fulfill additional requirements like security of the power plants and the time-independent availability of energy. One proposal for such an energy source is a fusion power plant, which gains energy from the fusion of light nuclei into heavier ones.

### 1.2.2 Nuclear fusion as an energy source

Nuclear fusion was identified as the energy source of the sun by von Weizsäcker in 1937 [3, 4]: . The sun, which mostly consists of hydrogen, gains energy by fusing protons to helium. The main fusion reactions are:



The net reaction remains



The energy release in this reaction is about 6 orders of magnitude larger than in typical chemical reactions (hydrogen ionisation energy:  $13.6 \text{ eV}$ ). It results from the difference of the mass of a complete nucleus and the sum of the masses of its constituents, the nucleons. The mass of the nucleus is slightly smaller, the mass difference corresponds to the nuclear binding energies, which are related to the mass difference via Einsteins energy-mass-relation  $\Delta E = \Delta mc^2$ .

The contributions to the total binding energy of the nucleus by the nucleons on the surface are small because of missing partners. So, smaller nuclei have smaller binding energies per nucleon according to the relatively large surface-to-volume ratio of the nucleus, larger nuclei have larger binding energies per nucleon. This effect reverses at a nucleon number of  $\approx 60$ , then the repulsive Coulomb force starts to play a role.

The highest energy gain per nucleon can therefore be achieved from the fusion of very light atoms like hydrogen (see figure 1.1), the same process takes place in the sun. However, the reaction rates of equation (1.1) are unacceptable small for the use in power plants. This argument does not play a role for the sun, because of its giant volume, but similar conditions cannot be achieved on earth.

Therefore, other fusion reactions have to be taken into account. The most promising one is



Here, the hydrogen isotopes deuterium and tritium are used, the energy gain of  $17.59 \text{ MeV}$  is relatively high. In addition, the reaction rate coefficient  $\langle \sigma v \rangle$  is relatively high compared to other possible reactions below an ion temperature of  $100 \text{ keV}$  (see figure 1.2). Deuterium is found in the oceans, whereas tritium can be bred from lithium by neutron attachment. Lithium itself is found in minerals.

A fusion power plant basing on this reaction turns out to be comparable with other possible energy sources in its costs [6]: The direct costs (constructing, fueling, operating

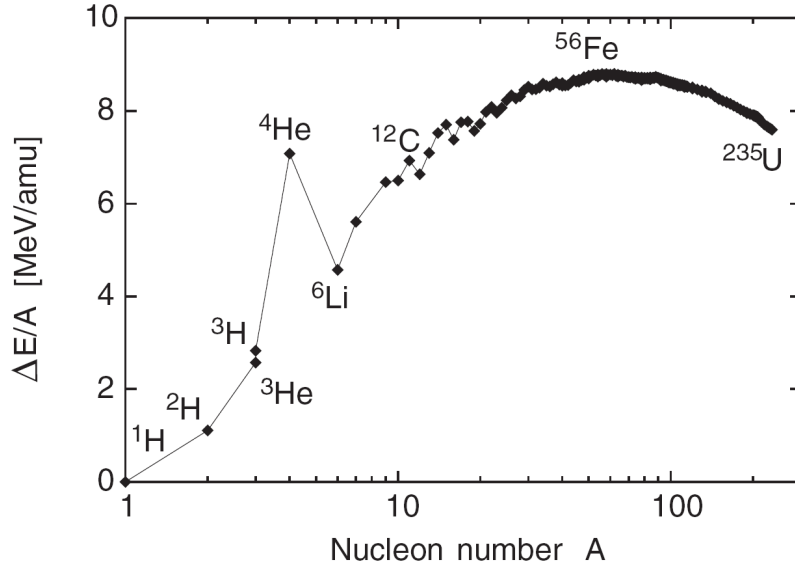


Figure 1.1: Binding energy per nucleon related to the nucleon number (from [5]).

disposing) are comparable with renewable energy sources and "clean coal" (fossil energy source including emission abatement costs). Also, fusion belongs to the class of low external cost power sources. That means that costs from environmental damage or adverse impacts upon health are significantly low (about 2.5% of the direct costs), because of the anticipated high safety standard of a fusion power plant. In result, fusion power turns out to be a technology with a long term potential (approximately 50 years from now).

### 1.2.3 Physical challenges

The main problem in the realisation of a nuclear fusion device is the Coulomb force, which acts repulsive for positively charged nuclei. The typical diameter of an atom lies in the range of  $10^{-10}$  m, whereas the strong force of the nucleus is operant at some  $10^{-15}$  m. The Coulomb barrier ( $Z$  - charge number,  $r$  - distance)

$$E_{Coulomb} = \frac{1}{4\pi\epsilon_0} \frac{Z_1 Z_2 e^2}{r} \quad (1.3)$$

in this range lies about some hundred keV and has to be overcome by the kinetic energy of the fusion reactants. The necessary amount of energy corresponds to temperatures in the order of  $10^9$  K, which is much more than a particle, even in the sun, can be achieved. The reason why nuclear fusion works anyway was found by Gamov [7] in 1928: The Coulomb barrier can be overcome by tunneling.

Nevertheless, for a sufficient number of fusion reactions high temperatures are still necessary: The reaction rate  $R$  for  $D$ - $T$  fusion reactions per volume is given by [5]

$$R = n_D n_T \langle \sigma v \rangle, \quad (1.4)$$

where  $n_D$  and  $n_T$  are the densities for deuterium and tritium, respectively. The reaction rate coefficient  $\langle \sigma v \rangle$  depends on the temperature (figure 1.2), sufficiently high values are given for temperatures of 10 keV (approximately  $10^8$  K). At these conditions the hydrogen atoms are fully ionised, this state is called *plasma*.

The plasma state is often referred to as the fourth state of matter, but one has to remember that whereas phase transitions exist between the other three states (gaseous,

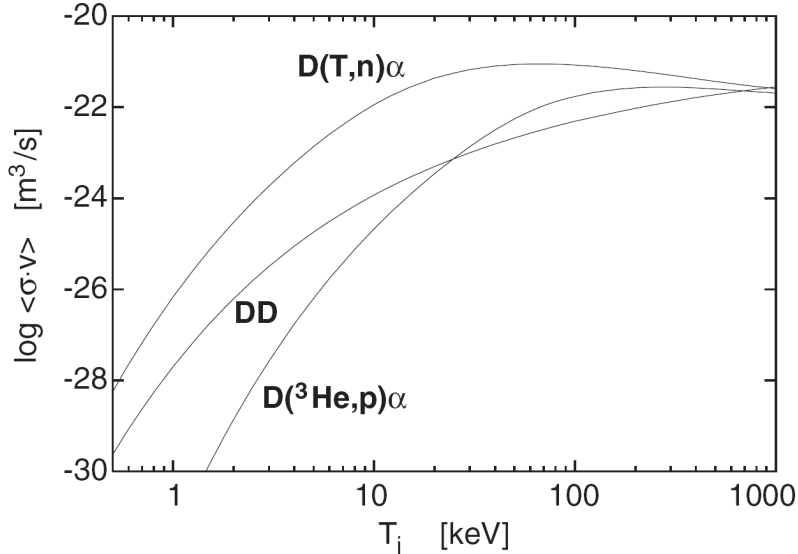


Figure 1.2: Reaction parameter  $\langle \sigma v \rangle$ , related to the ion temperature  $T_i$  for  $D$ - $T$ ,  $D$ - $D$  and  $D$ - ${}^3\text{He}$  fusion reactions (from [5]).

liquid and solid), no exact transition occurs between the gaseous and the plasma state. A plasma offers some unique properties in comparison to the normal gas: Because of the ionisation, the plasma contains of charged particles (electrons, ions). Nevertheless it is assumed as quasi-neutral, i.e. the negative and positive charges neutralise on a macroscopic scale [8].

In an effective fusion reactor a positive energy balance between the energy gain from the fusion reaction and the loss mechanisms (bremsstrahlung, and transport processes), characterised by the *energy confinement time*  $\tau_E$ , is necessary [5, 9]:

$$\underbrace{n_D n_T \langle \sigma v \rangle E_\alpha}_{\text{energy gain}} = \underbrace{3nk_B T / \tau_E}_{\text{transport losses}} + \underbrace{c_1 n^2 Z_{\text{eff}} (k_B T)^{1/2}}_{\text{bremsstrahlung}}. \quad (1.5)$$

Here,  $c_1$  is the bremsstrahlung constant and  $Z_{\text{eff}}$  the effective charge of the plasma, including all species (also impurities).  $E_\alpha$  is the energy of the  $\alpha$  particles (helium nuclei), which are confined due to their charge and contribute to the heating of the plasma.

For  $T \approx 10 \text{ keV}$  one finds  $\langle \sigma v \rangle \propto T^2$ , resulting in a criterion for the *triple product*  $n\tau_E T$

$$n\tau_E T = \frac{12k_B T^2}{\langle \sigma v \rangle E_\alpha - 4c_1 Z_{\text{eff}} (k_B T)^{1/2}}. \quad (1.6)$$

It turns out that the triple product shows a minimum of about  $35 \cdot 10^{20} \text{ keV s/m}^3$  for temperatures around  $10 \text{ keV}$ , this minimum has to be applied by an adequate confinement time  $\tau_E$ .

#### 1.2.4 Plasma confinement

For the confinement of a high temperature plasma, the unique characteristics of a plasma are utilised: Charged particles in electro-magnetic fields are exposed to the Lorentz force

$$F_L = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (1.7)$$

In electric fields  $\mathbf{E}$  they are accelerated along the field lines, in magnetic fields  $\mathbf{B}$  the Lorentz force constraints the particle movement (velocity  $\mathbf{v}$ ) to a helix like trajectory,

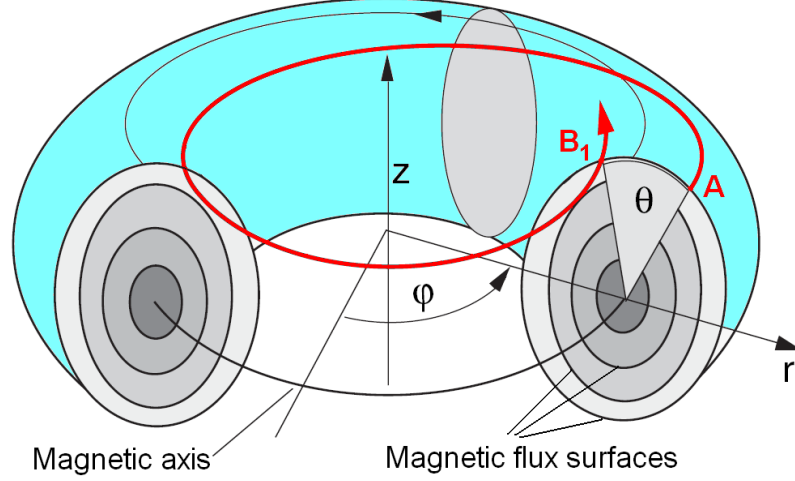


Figure 1.3: Rotational transform (from [11]): After one toroidal circulation a magnetic field line (red), which has started at the point  $A$ , ends up in  $B_1$ . After several circuits the field line forms a magnetic surface.

winded around the magnetic field line. Because the particles cannot move perpendicular to the field line, this seems to be an appropriate method to confine a plasma.

A homogeneous magnetic field can be found inside a long coil. To avoid particle losses at the ends, the coil can be bent to a torus, which leads to closed magnetic field lines inside the coil tube. As an effect of the bending, the distribution of the magnetic field lines becomes inhomogeneous: They are more dense at the inner side of the torus than at the outer side. As a consequence, a particle drift appears perpendicular both to the magnetic field  $\mathbf{B}$  and its spacial derivation  $\nabla\mathbf{B}$ , resulting in a loss of the plasma confinement.

To avoid this, the idea of *rotational transform* was introduced by Spitzer in 1958 [10]. The basic principle is a twisting of the magnetic field lines in a way that one and the same field line can be on the outer and on the inner side of the torus at different toroidal positions.

As a consequence, a magnetic field line does not necessarily match its starting point after one toroidal circulation (see figure 1.3). In fact, only one field line does this for sure, it is called *magnetic axis*. In a poloidal cross section, a field line, which has started in point  $A$ , normally crosses the plane several times in different points  $B_1 \dots B_n$ . The transformation of a set of points  $A$  into a set of points  $B_1$  is called *rotational transformation*  $\iota$ . It is characterised by the relation of poloidal to toroidal circuits:  $\iota = N_p/N_t = \theta/(2\pi)$ , where  $\theta$  is the poloidal angle between  $A$  and  $B_1$ . For closed magnetic field lines this is a rational number. It becomes irrational in case of field lines which are not closed.

The crossing points  $B_1 \dots B_n$  lie on a closed curve in the cross section plane. Taking into account the whole torus, one field line forms a *magnetic flux surface* after many circuits around the torus (see figure 1.3). These surfaces are of strong importance for plasma confinement.

Assuming a simple approximation for plasma equilibrium, the pressure gradient is balanced by the magnetic force:

$$\nabla p = \mathbf{j} \times \mathbf{B}, \quad (1.8)$$

where  $\mathbf{j}$  is the current density, leading to

$$\mathbf{j} \cdot \nabla p = 0 \quad \text{and} \quad \mathbf{B} \cdot \nabla p = 0. \quad (1.9)$$

As a result, on a magnetic surface with a magnetic field  $B_0$  the pressure  $p$  can be assumed as



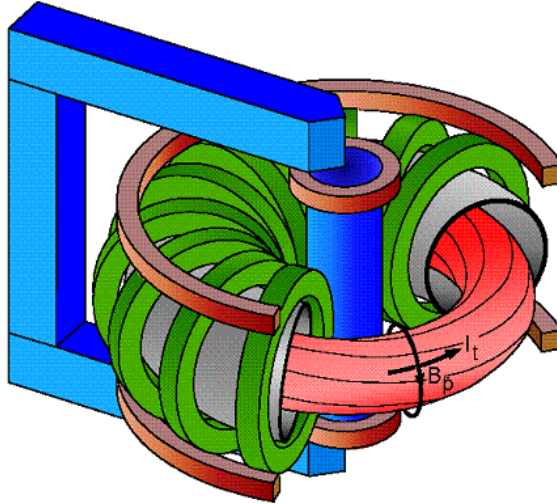


Figure 1.4: Schematic view of a tokamak: The external primary winding (red), toroidal coils (green) and an iron yoke (blue).

constant. In addition, the toroidal magnetic flux is independent on the toroidal position of the cross section plane, therefore the magnetic surfaces are also *flux surfaces*. The enclosed magnetic flux can be used as a label of the surfaces, whereas the innermost one (identical with the magnetic axis) encloses no volume and is often labeled with 0.

To generate the rotational transform, two basic ideas are realised in experiments: the *tokamak* and the *stellarator*.

In the tokamak configuration, the rotational transform is generated by a toroidal plasma current  $I_t$ , which is driven by an additional external coil (*primary winding*) by induction (see figure 1.4). The toroidal plasma current  $I_t$  generates an poloidal magnetic field  $B_p$ , which superposes with the toroidal field. In result the magnetic field is helically twisted. As a disadvantage, the inductive generation of the plasma current prevents steady-state operations for a long time.

The largest fusion experiment in the world is a tokamak: the JET (Joint European Torus) experiment in Culham, England. Here, a significant amount of fusion energy was produced by  $D - T$  experiments in 1997 [12], about 70 % of the heating energy was re-generated.

The next step in the tokamak line will be the International Thermonuclear Experimental Reactor (ITER), currently under construction in Cadarache, France [13].

### 1.2.5 The stellarator principle

The second type of fusion machines which realises the rotational transformation in a torus, is the stellarator. Whereas the early stellarator configurations were not necessarily torus-shaped (eight-shaped versions have also been tested), todays stellarator experiments are always toroidal set-ups. In comparison with other fusion experiments (like pyrotrons, pinches, even tokamaks), the stellarator concept offers at least in principle the advantage of an almost complete confinement and the possibility of steady-state operations. Both properties are advantageous for the use of a fusion device as a power plant.

The main characteristic of the stellarator concept is the absence of an induced toroidal current. In contrast to tokamaks, the rotational transform is produced by external coils only. For this, in addition to the toroidal field coils a system of helical windings is necessary (figure 1.5 (a)). The so-called *advanced stellarators* merge poloidal and helical windings

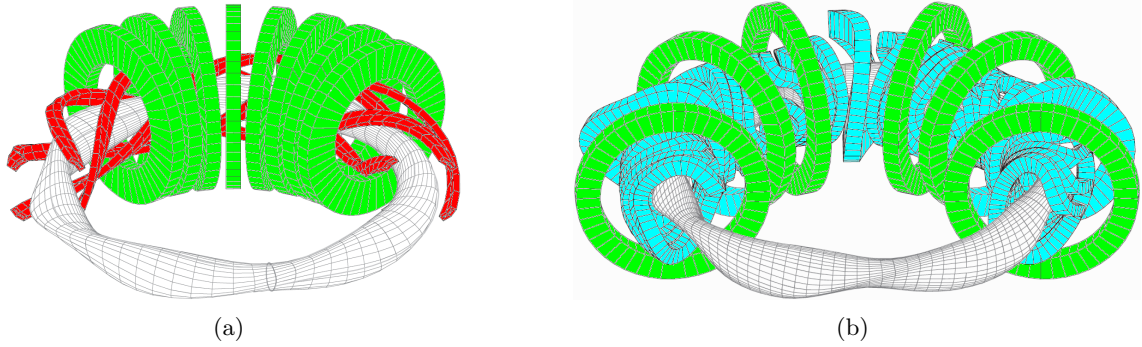


Figure 1.5: Types of stellarators [11]: In W7-A (a) the magnetic confinement was realised by poloidal coils (green) and helical windings (red). The modular stellarator W7-AS (b) used a combination of planar (green) and non-planar coils (light blue). The plasma is indicated by the mesh structure.

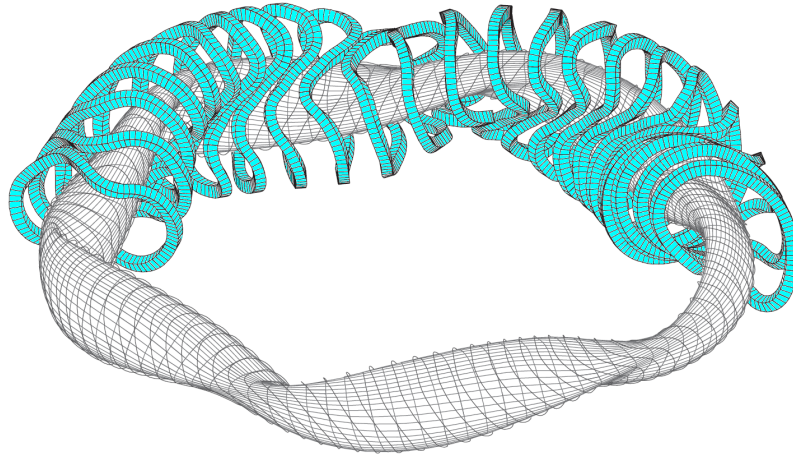


Figure 1.6: The coil system of W7-X (light blue) is a superposition of poloidal and helical windings [11]. The plasma is indicated by the mesh structure.

to a set of modular coils with a geometry of high complexity (figure 1.5 (b)).

The latest approach for an advanced stellarator experiment is Wendelstein 7-X, currently under construction at the Greifswald branch of the Max-Planck Institute for Plasma Physics. Key element is a modular and supra-conducting coil system (see figure 1.6). The design of the experiment was done with respect to several physical criteria [14]:

- Closed and well-shaped magnetic surfaces in the bulk plasma. Magnetic islands and stochastic regions should be avoided here, nevertheless they may be utilized for divertor operations in the plasma edge regions.
- Low Shafranov shift (movement of the plasma centre) and a stiff equilibrium configuration.
- Good magnetohydrodynamic (MHD) stability.
- Reduced influence of the plasma pressure on the rotational transform.
- Vanishing neoclassical particle losses (in the  $1/\nu$  transport regime).

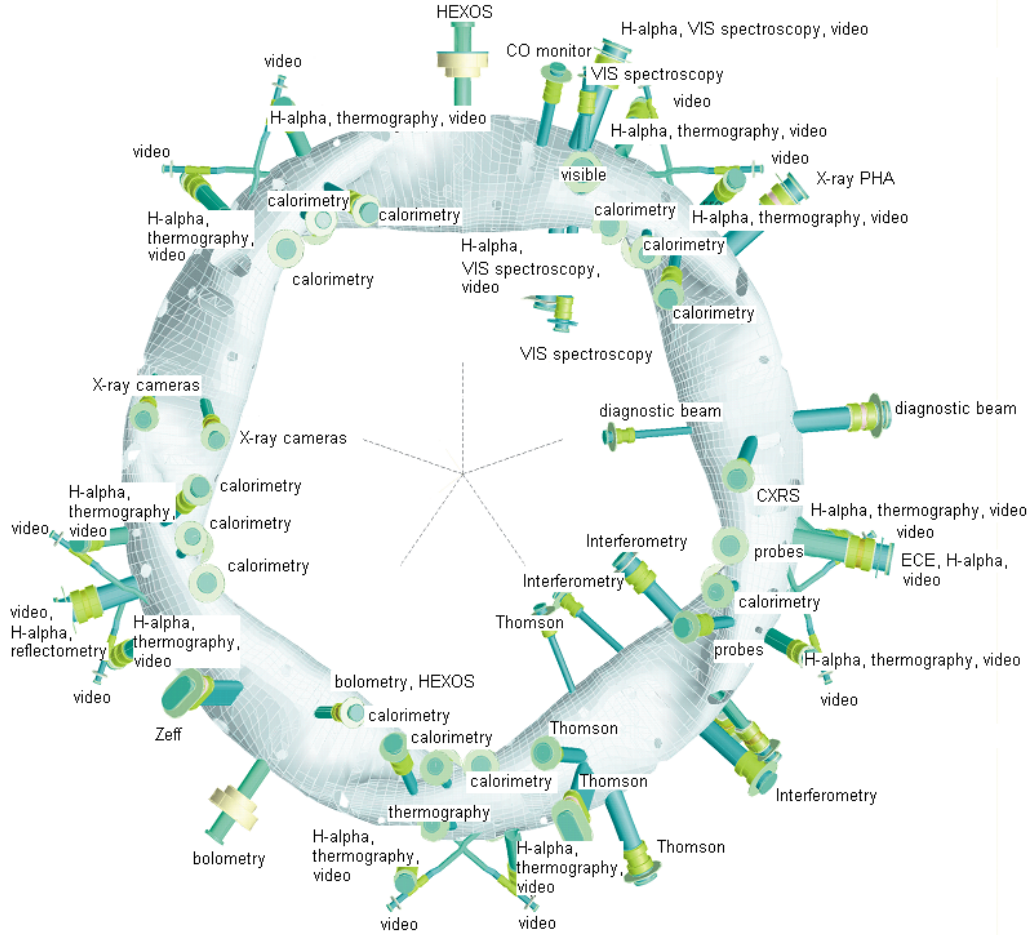


Figure 1.7: Schematic view of the W7-X diagnostic system: The access to the plasma vessel is realised by a complex port system.

- Small bootstrap currents in the low mean free path (lmfp) regime to avoid disturbances of the magnetic field.
- Good collisionless  $\alpha$  particle confinement.
- From the technical point of view, the coil system should provide a good feasibility and sufficient access to the plasma by an adequate large space between the coils and the plasma for diagnostic purposes, shielding and, later, operations with lithium blankets.

### 1.2.6 Diagnostic approach for W7-X

Taking into account these design specifications for W7-X, it is obvious that a well-structured diagnostic concept is needed to test and to analyse whether these criteria are fulfilled. Therefore, the diagnostics have to be designed to assure that the effects of these criteria can be measured in an optimal way. The boundary conditions of the experiment have to be taken into account, too. An overview over the diagnostic strategy of W7-X is given in [15].

Fusion plasmas are characterised by extreme conditions: A direct contact of a diagnostic with the plasma is nearly impossible due to the very high temperatures. The strong

magnetic fields, which are generated to realise the plasma confinement, may disturb the measurements. Furthermore, the complex assembly of the experiment in case of W7-X leads to restrictive technical boundary conditions for the access to the plasma.

For the measurement of different plasma parameters, e.g. the distributions of the plasma density and the temperature, several types of diagnostics have to be applied to gain reliable information. Figure 1.7 shows the diagnostic concept for W7-X in a schematic view, where the different diagnostics are assigned to their respective ports at the plasma vessel.

The complex situation outlined here confirms the need for a consistent treatment of the diagnostic design. A logical choice would be a design approach with respect to the scientific question of interest, in other words, to apply a *physical experimental design* ansatz.

### 1.3 Physical design

The approach for physical experimental design is outlined in figure 1.8 for a diagnostic at a stellarator, e.g. an interferometer. However, the ansatz is also valid for any other kind of experiment.

In the beginning, the scientific background (plasma physics in general, stellarator specific problems in particular) leads to predictions about the expected physical scenarios. These scenarios have to be parametrised in a way that one obtains a mathematical description of the physical problem. A set of *parameters of interest* has to be defined now: The estimation of these parameters will be the goal of the experiment.

Also, the interest in the parameter space has to be specified, a weighting of this interest is possible here. With the resulting limitations of the parameter range, the expected data space can be calculated.

For this, a mathematical model of the measurement (*forward function*) has to be applied, it transforms a physical situation, described by the parameters of interest, into (expected) data points. Here, an exact knowledge about the experiment and the diagnostics is necessary, also a description of the error statistics. The resulting mathematical model can be seen as a *virtual measurement* and can also be used for the analysis of measured data later on.

Finally, an expression of the *utility* of the measurement with respect to the data has to be found. For the expected data used here, it is called *Expected Utility (EU)*. It describes the usefulness of the expected data with respect to the physical problem.

The approach described here would provide some remarkable advantages:

- Physical design: The physical parameters of interest are intrinsically implemented as design criteria.
- Comparability: If the Expected Utility can be quantified, different designs are directly comparable.
- Independency: The framework of this approach is general, so it is possible to apply it to different kinds of diagnostics.

The methodological ansatz presented in this work uses an *information measure* as the expression of the utility. This approach is not only valid for the optimisation of diagnostic units, but also for the planning of experiments and experimental campaigns.

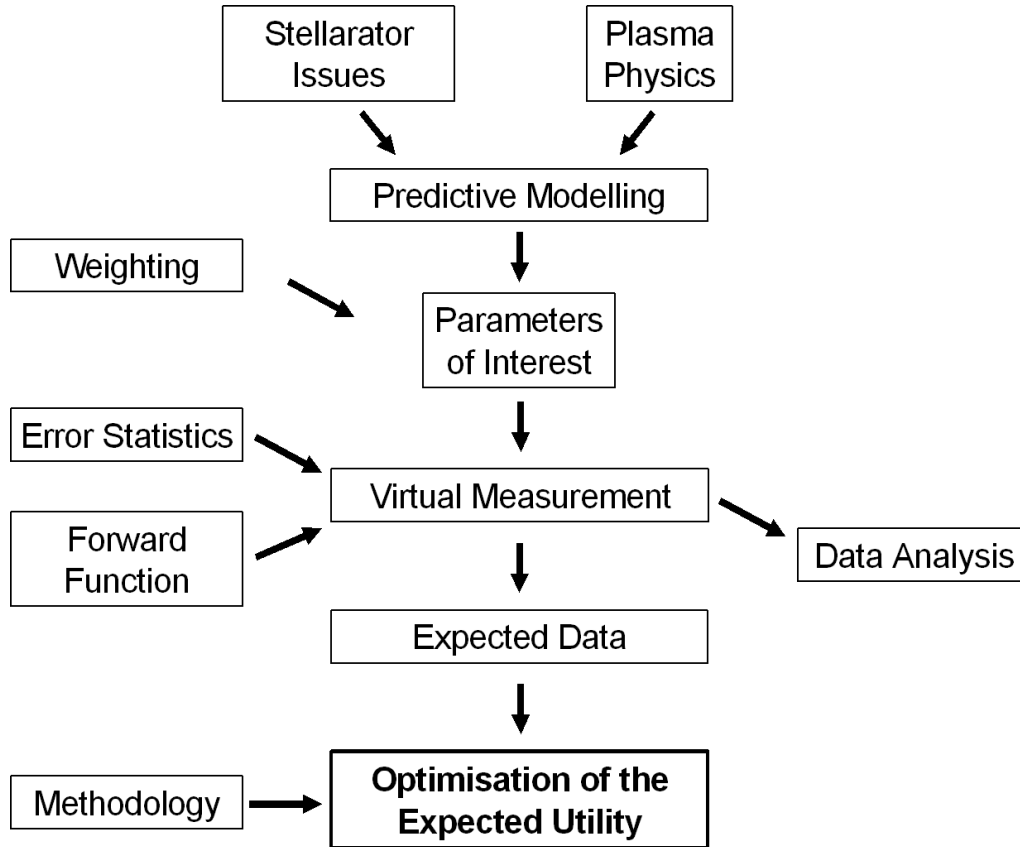


Figure 1.8: Concept for physical experimental design

### 1.3.1 Overview

This work will present a method for physically motivated experimental design by using an information measure as utility function. It will be shown, how stellarator-specific problems can be utilised for the design of a diagnostic unit as well as for the planning of experiments.

In the next chapter, the basics of the design method will be introduced. Therefore, the principles of Bayesian theory, which offer a consistent framework for the handling of probability density distributions, are illustrated. Furthermore, a short introduction into information theory is given, presenting the necessary basics for the quantification of information gain.

The diagnostic to be optimised, the infrared multi-channel interferometer at W7-X, is described in the third chapter. Here, the virtual measurement (see figure 1.8) is discussed by analysing the forward model as well as the error statistics of the diagnostic.

Next, some previous results of diagnostic design for this interferometer will be discussed. The use of an information measure as utility function is validated by case studies. For this, data sets from the W7-AS experiment are analysed, the respective Expected Utility is calculated independently. By comparison, the relationship between measurement result and  $EU$  is displayed and discussed.

The fifth chapter of this work will at first present an overview over some selected topics of physical problems to be solved with W7-X: The problem of the Core Electron Root Confinement (CERC), the estimation of high plasma confinement regimes and the measurement of density effects during the injection of hydrogen pellets. From these physical problems, the parameters of interest are derived. In the second part, the design results for a multi-channel interferometer are presented, according to the different physical prob-

lems. The optimisation results are shown for configurations with and without technical constraints, taking into account several boundary conditions.

The last chapter deals with the problem of experimental planning. The concept of empirical scaling laws will be introduced as a physical question. The use of the design method for the evaluation of given data and for the estimation of the best experimental configuration for a future experiment is shown. The application of this method for experimental planning is outlined.



## Chapter 2

# Bayesian Experimental Design

In this chapter the method for Bayesian experimental design (*BED*) will be introduced. It applies Bayesian probability theory, which itself is derived from first principles. Three advantages of this approach will be pointed out: First, it is independent of the experiment to be optimized, therefore it can be seen as a general method. Secondly, the quality of the design is quantifiable, so different designs can be compared directly. And finally, it allows the inclusion of physical motivated background knowledge (e.g. quantities like temperatures cannot be negative) into the data analysis process.

A brief introduction to Bayesian theory and information measures is given in the first part of this chapter. Both theories are used for the approach of *BED*, which will be explained in the second part.

### 2.1 Bayesian probability theory

Bayesian probability theory combines first principles for the calculation of conditional probabilities. This allows the self-consistent implementation of different sources of information. Examples are, e.g., the inclusion of (physical) background knowledge and the handling of unknown, so-called nuisance parameters. An intelligible and more extensive introduction into Bayesian theory is given, e.g., by Sivia [16] or Gregory [17].

Two different approaches exist to define the term *probability* (see, e.g., Cox [18]). First, the idea of *frequency* in an ensemble and, second, the idea of *uncertainty*. In the frequency approach, probability is defined as the repetition frequency of an event in an ensemble. Here, the ensemble may be a large number of identical experiments or an experiment which is repeated many times in the same way. The term "probability" is therefore defined by the ensemble.

The frequency approach fails in cases where no repetition is possible or, in other words, where no ensemble exists. For instance, the question whether there is life on Mars cannot be treated in this way, because there is only one Mars, not a large number of planets. The same issue arises in the case of physical constants, which are not exactly known: Defining the probability of a constant in such a way would mean that in different experiments different values for the *constant* are possible - which is a contradiction [19].

In the Bayesian approach, probabilities are understood as the uncertainty of the quantities of interest, which have a well defined value, but one is not certain about it. This uncertainty may be diminished by measurements.

Uncertainties can be formulated as probability density functions (*PDF*). The *PDF*  $p(X|Y)$  encodes the probability distribution of the quantity  $X$ , given the quantity  $Y$ , whereas  $p(X, Y|Z)$  expresses the probability distribution of  $X$  and  $Y$  given  $Z$ . As a



common convention, probabilities of a quantity are normalised to 1 if there is no other state of the quantity possible, in other words, there is no uncertainty about this quantity. On the other hand, the probability is zero if the state of a quantity is impossible.

Some basic rules, which can be derived directly from Boolean algebra [18], are used to handle probabilities and probability distributions. The two basic laws are the *product rule* and the *sum rule*.

The sum rule is based on the consideration, that if one knows the probability that  $X$  is *true*, one also knows the probability that  $X$  is not *true*:

$$p(X|I) + p(\bar{X}|I) = 1. \quad (2.1)$$

Here  $\bar{X}$  denotes that  $X$  is not *true*,  $I$  stands for all context information available. If more than one state for  $X$  is possible, one obtains (for disjoint  $X_m$ ,  $M$  is the number of all possible  $X_m$ )

$$\sum_{m=1}^M p(X_m|I) = 1. \quad (2.2)$$

The aforementioned product rule as the second basic law involves conditional probabilities. Here, the probability of a quantity is not independent of other quantities. The product rule covers the following case: If one knows the probability that  $Y$  is true independent of  $X$ , and the probability that  $X$  is true *if  $Y$  is true*, one knows how probable both  $X$  and  $Y$  are true:

$$p(X, Y|I) = p(X|Y, I) \cdot p(Y|I). \quad (2.3)$$

The sum rule in (2.1) can also be written for conditional probabilities: For a *PDF*  $\sum_{m=1}^M p(X, Y_m|I)$  one obtains from (2.3)

$$\sum_{m=1}^M p(X, Y_m|I) = p(X|I) \sum_{m=1}^M p(Y_m|X, I) = p(X|I), \quad (2.4)$$

because of the normalisation

$$\sum_{m=1}^M p(Y_m|X, I) = \sum_{m=1}^M p(Y_m|I) = 1. \quad (2.5)$$

Equations (2.2) and (2.4) finally lead to

$$p(X|I) = \sum_{m=1}^M p(Y_m|I) \cdot p(X|Y_m, I) = \sum_{m=1}^M p(X, Y_m|I) \quad (2.6)$$

as the sum rule for conditional probabilities.

The sum rule (2.6) can be extended to continuous probability distributions:

$$p(X|I) = \int dY p(Y|I) \cdot p(X|Y, I). \quad (2.7)$$

This procedure is called *marginalisation*. It offers a very powerful tool to deal with *nuisance parameters*: Quantities, whose exact value is unknown (here:  $Y$ ), can be integrated out. The resulting probability distribution of the parameters of interest ( $X$ ) becomes broader but it does no more explicitly depend on  $Y$ .

In data analysis, this formalism is used to calculate the influence of the nuisance parameters (e.g. calibration parameters) on the probability distribution of the parameters of interest. Therefore it can be seen as a general description for error propagation.

The marginalisation rule (2.7) is one of the two basic formulas of Bayesian probability theory. The second one, called Bayes theorem, is derived from equation (2.3). Because of  $p(X, Y|I) = p(Y, X|I)$  one obtains

$$p(X|Y, I) \cdot p(Y|I) = p(Y|X, I) \cdot p(X|I), \quad (2.8)$$

which leads to

$$p(X|Y, I) = \frac{p(Y|X, I) \cdot p(Y|I)}{p(X|I)} \quad (\text{Bayesian theorem}). \quad (2.9)$$

This theorem is a rule for inverse conclusions. The probability distributions of the quantities  $X$  and  $Y$ ,  $p(Y|I)$  and  $p(X|I)$  have to be taken into account, if the *PDF*  $p(X|Y, I)$  has to be calculated.

### Bayesian data analysis

The Bayesian approach for data analysis is a framework to merge probability distributions in a consistent way. The result is a probability density distribution for the parameters of interest.

The subsequent part of this section is devoted to Bayesian data analysis. The notation used in the following chapters will be introduced, all quantities may be understood as vectors.

If Bayes theorem is used for data analysis, the quantities and terms are interpreted in the following way: A model  $H$  is used to describe the physical problem. It is characterised by a set of parameters  $\theta$ , the goal of the experiment is to estimate these parameters by a measurement. Therefore,  $\theta$  will be called *parameters of interest*. The knowledge about  $\theta$  without any given data is encoded in a *PDF*  $p(\theta|H, I)$ . This function is called *prior*. Again,  $I$  represents all kind of background information.

The formulation of the prior function contains the available information about the parameters of interest. Typical prior knowledge is, e.g., given by physical constraints (densities and temperatures are positive). Identical knowledge about the parameters of interest will always lead to the same prior, invariance arguments can be used to determine the prior *PDF* uniquely. In addition, the principle of Maximum Entropy allows one to deal with testable constraints [16, 20].

When the experiment is done, *data*  $D$  are available. One knows that the data is uncertain because of systematic and statistical uncertainties in the measurement. Therefore, for a given parameter set  $\theta$  different data sets are possible. The probability distribution of the data is given by the *likelihood*  $p(D|\theta, H, I)$ . It contains the *forward function*, i.e. a mathematical model of the experimental set-up describing the emergence of a data point at a certain state of  $\theta$ . The likelihood also covers the error statistics of the measurement, in other words, it is, in principle, a model of the experiment. A common ansatz for a likelihood function is a Gaussian distribution:

$$p(D|\theta, H, I) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(D - f(\theta))^2}{2\sigma^2}\right], \quad (2.10)$$

where  $f(\theta)$  is the forward function and  $\sigma$  is the standard deviation of the measurement noise.

The result of the data analysis is the *posterior* function  $p(\theta|D, H, I)$ , a *PDF* describing the knowledge about the parameters of interest when data  $D$  are available. With these notations, Bayes theorem (2.9) reads as

$$p(\theta|D, H, I) = \frac{p(D|\theta, H, I) \cdot p(\theta|H, I)}{p(D|H, I)}. \quad (2.11)$$

The term  $p(D|H, I)$  is called *evidence* (sometimes also *prior predictive value (PPV)*), it is the probability for the occurrence of the data  $D$ , independent of the parameters  $\theta$ . The evidence is a normalisation factor which can be ignored if estimates from a posterior are to be derived. However, it becomes important for the comparison of different models where  $p(D|H_1, I) \neq p(D|H_2, I)$ . It also plays a crucial role for experimental design, where it is calculated by marginalisation (equation (2.7)):

$$p(D|H, I) = \int d\theta \ p(D|\theta, H, I) \cdot p(\theta|H, I). \quad (2.12)$$

There are many examples of successful application of the Bayesian formalism in data analysis. In plasma physics one finds, e.g, straight forward analyses from Ertl et al. [21] and Anton et al. [22] for soft X-ray tomography and from Koponen and Dumbrajs [23] for interferometry. Different approaches for the prior function (Maximum Entropy, Minimum Fisher) were applied in these works.

Studies for model comparison were made by Preuss et al. [24, 25] for scaling laws (see chapter 6). And, as a step for future analysis work, the concept of *Integrated Data Analysis (IDA)* was presented by Dinkluge et al. [26]: Several different diagnostic units (Thomson scattering, interferometry, soft X-ray) were analysed together, the capability of handling the interdependencies between the different diagnostics in a consistent way shows the advantages of the Bayesian approach.

## 2.2 A brief introduction to information theory

The goal of every experiment is to learn something about the parameters of interest, in other words, to gain information about them. Information shall be understood here in a way that it constraints our belief about the parameters to be measured (see Caticha and Griffin [27]), i.e., that one gains information if the uncertainty of the parameters after the experiment is not the same as before.

A maximum of information gain in that way seems to be an effective criterion for experimental design. However, it is necessary to define "information" mathematically first. In information theory, several measures are given to quantify the information gain of an experiment.

Information theory basically deals with storage, coding and transmission of information. Electronic communication is a typical example: Words, consisting of letters from an alphabet  $A$ , are encoded using an alphabet  $B$  for transmission. Here,  $B$  may be of advantage from a technical point of view (one may think of the binary code), whereas  $A$  could be the normal alphabet. The decoding of the transmitted signals will give back the original words, if the transmission was not disturbed (e.g. by noise).

One can see the process of a measurement in a similar way: The parameters of interest - described by the alphabet  $A$  - are encoded by the measurement into a new alphabet  $B$ , the data. It is decoded later by the data analysis process, finally the diagnostician gets the original parameters.

In the following section, the principles of information measures are introduced. Here, the argumentation of [28] is followed.

### 2.2.1 Hartley information

An alphabet

$$A : \{a_1, a_2, \dots, a_{|A|}\} \quad (2.13)$$

shall be assumed, as well as a source sending a sequence of letters

$$a_{m_1}, a_{m_2}, \dots, a_{m_T} \quad (2.14)$$

during the time interval  $T$ , where

$$m_t \in \{1, 2, \dots, |A|\} \quad (2.15)$$

is the index at the point in time  $t \in T$ . For the sequence of letters,  $|A|^T$  different possibilities exist, therefore it seems to be a reasonable approach to define the information  $I$  of the sequence as

$$I = T. \quad (2.16)$$

For a sequence twice as long as the original one, one gets double information. If the source did not send any letter, no information is gained:  $I = 0$ .

If one maps  $A$  one-to-one on a new alphabet

$$B : \{b_1, b_2, \dots, b_r\}, \quad (2.17)$$

one also finds a one-to-one transformation of (2.14) on a sequence

$$b_{m_1}, b_{m_2}, \dots, b_{m_T}, \quad (2.18)$$

which contains the same information as (2.14).

If there are less letters in  $B$  than in  $A$ , the transformation cannot be inverted. In this case the Cartesian product

$$\otimes_l B_r = \underbrace{B_r \times B_r \times \dots \times B_r}_{l\text{-times}} \quad (2.19)$$

can be employed, which gives a one-to-one transformation if  $l$  is sufficiently large. As a necessary condition,  $r$  has to be  $r \geq 2$ , because  $|\otimes_l B_r| = r^l$ . The letters of  $\otimes_l B_r$  are called *code words*, in case of  $r = 2$   $B$  is called *binary code* with  $b_1 \equiv 0$  and  $b_2 \equiv 1$ .

When a sequence of letters from alphabet  $A$  of length  $T$  is expressed by the alphabet  $B$ , one gets a new sequence of length  $l \cdot T$ . Defining the information as in (2.16), one obtains  $I = T$  for the old and  $I = l \cdot T$  for the new sequence. Apparently, this is a contradiction, since a one-to-one transformation was used and no additional information was introduced.

To avoid this problem, the smallest alphabet possible has to be chosen to define the information. which is the aforementioned binary code with  $r = 2$  different letters. Now an alphabet  $A$ , containing of  $|A| = 2^l$  letters, can be coded with binary code words  $\otimes_l B_2$  of length  $\text{ld } 2^l = l$ . The overall information is given by

$$I_{all} = lT, \quad (2.20)$$

and (using  $\text{ld}$  for the base-2 logarithm)

$$I_{all}/T = \text{ld}|A| = l \quad (2.21)$$

as the information per letter, respectively. With every letter of  $A$  one gains the information

$$I_H = \text{ld}|A|. \quad (2.22)$$

This *information measure* is called Hartley information. It is the basic measure for all measures of information gain.

For two different alphabets  $X$  and  $Y$ , the Hartley information obeys the following rules:

1. additivity:  $I_H(X \times Y) = I_H(X) + I_H(Y)$
2. monotonicity:  $I_H(X) < I_H(Y)$  for  $|X| < |Y|$
3. unit:  $I_H(X) = 1$  if  $|X| = 1$ . By using the base-2 logarithm, the unit is called *bit*.

The unit *bit* may be interpreted as follows: An easy explanation is given by looking at a cupboard with several drawers. An object is hidden in one of the drawers, and one wants to find it just by asking yes/no questions. For two drawers, one needs exactly one question, two questions for four drawers, three for eight drawers and so on. In other words: The number of questions is equal to the logarithm dualis (ld) of the number of drawers. What happens in the case of five drawers? Sometimes one will need two questions, sometimes three to find the object. In average,  $\text{ld } 5 = 2.3219\dots$  questions are necessary.

This explains how fractional numbers of bits can occur. In the example, a gain of one bit halves the uncertainty about the position of the object. When a probability distribution is given, the correlation between information gain and uncertainty (maybe given in form of a standard deviation) is straight forward, but not necessarily log-linear.

However, the Hartley information measure fails if the letters from the alphabet have different probabilities to occur: Drawing a letter with a small probability should intuitively lead to a higher information gain than a letter with a large probability. Therefore, a new expression for the information gain is needed taking into account that letters with different probabilities must result in different amounts of information.

### 2.2.2 Shannon information

For introducing a probability based weighting to the information measure, the main idea is to divide the original alphabet  $A$  into pairwise disjoint sub - alphabets:

$$A_1, A_2, \dots, A_m, \dots, A_M \quad (2.23)$$

with

$$A = \bigcup_{m=1}^M A_m, \quad A_{m_1} \cap A_{m_2} \quad \forall m_1 \neq m_2. \quad (2.24)$$

Doing so, one obtains

$$p_m = \frac{|A_m|}{|A|} \quad (2.25)$$

as the probability for finding a letter in the sub-alphabet  $A_m$ . The overall information for a letter in sub-alphabet  $m$  can now be divided in two parts:

$$\text{ld}|A| = \underbrace{I_m}_{\text{information that letter in } A_m} + \underbrace{\text{ld}|A_m|}_{\text{information about position in } A_m}, \quad (2.26)$$

leading to

$$I_m = -\text{ld} \frac{|A_m|}{|A|} = -\text{ld } p_m \quad (2.27)$$

as the information, that the letter is in sub-alphabet  $A_m$ . The overall information for all sub-alphabets needs a summation over all  $m$ , weighted with the probability  $p_m$ :

$$I_S = - \sum_{m=1}^M p_m \text{ld } p_m \quad (2.28)$$

This expression is called Shannon information [29]. For the limiting case of uniformly distributed letters ( $p_m = 1/M$ ), the Hartley information (2.22) is resembled.

Some characteristics of the Shannon information should be mentioned:

1.  $I_S$  is always positive, since  $0 \leq p_m \leq 1$ .
2.  $I_S$  is symmetric in all arguments:  

$$I_S(p_1, \dots, p_{m_1}, \dots, p_{m_2}, \dots, M) = I_S(p_1, \dots, p_{m_2}, \dots, p_{m_1}, \dots, M)$$
3. For a discrete probability distribution  $P = \{p_1, \dots, p_m, \dots, p_M\}$  the Shannon information is restricted:  

$$0 \leq I_S(P) \leq \text{ld}M,$$
where the lower limit is reached if all  $p_m$  are zero except one, and the upper limit for a uniform distribution as mentioned before.

For the quantification of the information gain it is necessary to compare different state of knowledge.

### 2.2.3 Kullback–Leibler distance

Again, an alphabet  $A : \{A_1, A_2, \dots, A_M\}$  is used, containing of pairwise disjoint sub-alphabets  $A_m$ . The probability of a letter from  $A$  in  $A_m$  is  $p_m = |A_m|/|A|$ , the distribution of all discrete probabilities is  $P = \{p_m\}_{m=1}^M$ .

Now a second alphabet  $A^*$  with  $A^* \subseteq A$  is considered, containing of sub-alphabets  $A_m^* \equiv A^* \cap A_m$ . Their probabilities are  $q_m = |A_m^*|/|A^*|$ , the probability distribution is  $Q \equiv \{q_m\}_{m=1}^M$ .

The term  $\text{ld}|A|$  is identified as the overall "a priori" information about a drawn letter  $a_i$ , and  $\text{ld}|A^*|$  as the "a posteriori" information, i.e. the information about  $a_i$  if

$$a_i \in A^* \tag{2.29}$$

is known.

The information gain if  $P$  is replaced by  $Q$  is given by

$$\text{ld}|A| - \text{ld}|A^*| = \text{ld} \frac{|A|}{|A^*|}. \tag{2.30}$$

This information can be divided in two parts:

1. The information, from which sub-alphabet  $A_m$  a letter  $a_i$  was drawn, when (2.29) is known. This is the partial information gain one is interested in, it is written as  $I(q_m||p_m)$ .
2. The information, which letter from  $A_m$  was drawn, when (2.29) is known. This can be written as  $\text{ld}|A_m| - \text{ld}|A_m^*| = \text{ld} \frac{|A_m|}{|A_m^*|}$ .

Combining both parts leads to

$$\text{ld} \frac{|A|}{|A^*|} = I(q_m||p_m) + \text{ld} \frac{|A_m|}{|A_m^*|}. \tag{2.31}$$

By combining (2.30) and (2.31) one finds for the  $m$ -th partial information gain

$$I(q_m||p_m) = \text{ld} \frac{q_m}{p_m}. \tag{2.32}$$

In general, one is interested in the average information gain, when  $P$  is replaced by  $Q$ . Therefore, one has to summarise all partial information gains, each weighted by the probability of its occurrence, which is given by  $q_m$  for  $a_i \in A_m$  if (2.29) is fulfilled. One obtains

$$I(Q||P) \equiv \sum_{m=1}^M q_m I(q_m||p_m), \tag{2.33}$$

and with (2.32)

$$I(Q||P) = \sum_{m=1}^M q_m \text{ld} \frac{q_m}{p_m}. \quad (2.34)$$

This expression gives a quantity for the information gain, if the probability distribution  $P$  is replaced by the distribution  $Q$ . In general, this always happens when the uncertainty of a parameter of interest changes by the analysis of a measurement.

Whereas the term  $\text{ld} \frac{q_m}{p_m}$  can be positive and negative, the information gain (2.34) is always positive [28].

For continuous *PDFs* equation (2.34) can be expanded straightforward: Let  $p(x)$  and  $q(x)$  be two continuous probability distributions, the support of  $q(x)$ ,  $\text{supp}(q) \equiv \{x \in R : q(x) > 0\}$ , is contained in  $\text{supp}(p)$ .

Now  $\text{supp}(p)$  is divided into pairwise disjoint intervals  $p_m$  with length  $\Delta x$ , the same is done for  $q(x)$ . This generates the distributions  $P_{\Delta x} = \{p_m\}$  and  $Q_{\Delta x} = \{q_m\}$ . With (2.34) one obtains

$$I(P_{\Delta x}||Q_{\Delta x}) = \sum_{m=1}^M q_m \Delta x \text{ld} \frac{q_m \Delta x}{p_m \Delta x}. \quad (2.35)$$

The  $\Delta x$  - term in the logarithm cancels, so one finally gets:

$$\lim_{\Delta x \rightarrow 0} I(P_{\Delta x}||Q_{\Delta x}) \Rightarrow I_{KL} = \int q(x) \text{ld} \frac{q(x)}{p(x)} dx \quad (2.36)$$

This expression is called Kullback–Leibler information and provides an *absolute measure* for the information gain. It is a measure for the "distance" between two continuous probability distributions  $P$  and  $Q$ , or, in other words, for the information gain if  $P$  is replaced by  $Q$ . It is measured in *bit* in case of using the base-2 logarithm as intended here. Sometimes it is written with a minus sign and called Kullback–Leibler entropy.

## 2.3 Bayesian experimental design

Using the tools of Bayesian probability theory one is able to handle *PDFs* in a consistent way. In addition, the introduced information measures offer an instrument to quantify the information gain of an experiment. For experimental design only the design parameters have to be included now.

As the figure of merit for the experimental design a *utility function*

$$U(D, \eta) \quad (2.37)$$

is defined, which depends on the data  $D$  of an experiment and the design parameters  $\eta$ . Design parameters are understood as quantities of an experiment and diagnostic unit respectively, which are accessible and changeable. Examples are geometric coordinates for a line of sight or the point in time for the next measurement. Experimental design in this approach is a decision-theoretic problem, based on the optimisation (i.e. maximisation) of the utility function with respect to  $\eta$ .

For the data  $D$  different results are expected, because the future data is often uncertain due to systematic and statistical errors and, in addition, the data range changes due to different physical scenarios in the experiment (parameter scans etc.). Nevertheless, the experimental design has to cover all possible data sets. Hence, the utility function has to be marginalised over the expected data space, which results in the so called *Expected Utility (EU)*:

$$EU(\eta) = \int dD p(D|\eta, I) U(D, \eta). \quad (2.38)$$

The *PDF*  $p(D|\eta, I)$  encodes the probability of the occurrence of the data  $D$  from relevant scenarios for the optimal design. The *EU* itself only depends on the design parameters.

As introduced in section 2.2.3 the Kullback-Leibler distance provides an absolute measure for the information gain of an experiment on the parameters of interest  $\theta$ . Therefore, the uncertainty about  $\theta$  before the measurement,  $p(\theta|I)$ , is compared with the knowledge about  $\theta$  when data from the experiment are available,  $p(\theta|D, \eta, I)$ :

$$I_{KL} = \int d\theta p(\theta|D, \eta, I) \text{ld} \frac{p(\theta|D, \eta, I)}{p(\theta|I)} \quad (2.39)$$

In this work, the Kullback-Leibler distance is used as the utility function for experimental design. By inserting (2.39) into (2.38) one obtains

$$EU(\eta) = \iint dD d\theta p(D|I) p(\theta|D, \eta, I) \text{ld} \frac{p(\theta|D, \eta, I)}{p(\theta|I)}. \quad (2.40)$$

During the design process, no data  $D$  are given and  $p(\theta|D, \eta, I)$  is not available. It can be replaced by using Bayes theorem (2.11) and one gets

$$EU(\eta) = \iint d\theta dD p(D|\theta, \eta, I) p(\theta) \text{ld} \frac{p(D|\theta, \eta, I)}{p(D|\eta, I)}. \quad (2.41)$$

This expression is an absolute measure for the information gain from the experiment, now averaged over all possible data sets.

In Bayesian experimental design, the Expected Utility (2.41) is maximised with respect to the design parameters  $\eta$ .

Only the prior *PDF* and the likelihood have to be provided for the design, the evidence is calculated by using these functions according to equation (2.12). The likelihood contains all information about the experiment as described in section 2.1. The meaning of the prior is different in a sense that in *BED* the prior encodes the *area of interest*, in other words, the area of the parameter space one is interested in.

The Expected Utility can be tested with some limiting cases [30]:

1. No data are available: The posterior distribution  $p(\theta|D, \eta, I)$  is equal to the prior  $p(\theta|I)$ , (2.40) vanishes because  $\text{ld} 1 = 0$ , there is no information gain without data.
2. The data  $D$  are totally uninformative of the parameters of interest  $\theta$ : The experiment is completely ignorant about  $\theta$ , leading to  $p(D|\theta, \eta, I) = p(D|I)$ . Equation (2.41) results in  $EU = 0$ . So, if the data tells nothing about the parameters there will be no information gain.
3. The parameters of interest are exactly known before the measurement: In this case,  $p(\theta|I) = p(\theta|D, \eta, I)$ , resulting in  $EU = 0$  – one can not learn anything if everything is known.

Historically, the characterisation of the average information gain from an experiment was first proposed by Lindley [31]. Later, he introduced a decision theoretic approach for experimental design using a utility function [32]. An overview of different approaches for utility functions (e.g. Shannon entropy, Kullback-Leibler entropy, maximisation of the inverse determinant of the covariance matrix) can be found by Chaloner and Verdinelli [33]. In this work, the *BED* approach is also compared with the non-Bayesian design approaches and Bayesian alphabetical optimality.

Applications of Bayesian experimental design in physics are very rare to this point. One example is given by the work of Loredo [34], in which *BED* is used to find the optimal observation points in time for an astrophysical experiment.





## Chapter 3

# Interferometry at Wendelstein 7-X

### 3.1 Density measurement with interferometry

Interferometry is a well established method for the determination of the electron density in a high temperature plasma. It is a line-integrated density measurement, using the effect of a changing optical path length due to the varying refraction index of the plasma. This induces a phase shift on the probing beam, which is detected by interferometric comparison with a reference beam not affected by the plasma. Both beams are usually generated by a laser to establish coherence. Although only the line-integrated density (along the line of sight of the probing beam) can be measured, it is possible to extract a spatially resolved density profile by using a multi-channel interferometer and an appropriate inversion procedure. Interferometry is a robust diagnostic and provides a good temporal resolution of typical changes in the plasma density (e.g. due to injection of hydrogen pellets). These effects can therefore be used as design criteria.

#### 3.1.1 The principle: Phase shift and line integration

The principle of interferometry shall be introduced shortly by following the overview given by Veron [35].

A plasma is transparent for an electromagnetic wave with wavelength  $\lambda$ , if the plasma density  $n$  is smaller than the *cut-off density*  $n_c$ :

$$n_c = \frac{4\pi c^2 \epsilon_0 m_e}{\lambda^2 e^2} \quad (3.1)$$

Here,  $c$  is the speed of light,  $e$  the elementary charge,  $m_e$  the electron mass and  $\epsilon_0$  the dielectric constant for vacuum. In case of an ordinary wave, i.e. the electric field is parallel to the magnetic field, the refraction index of the plasma is

$$N = \sqrt{1 - \frac{n}{n_c}}. \quad (3.2)$$

For  $n \ll n_c$  (3.2) can be approximated by

$$N \approx 1 - \frac{n}{2n_c}. \quad (3.3)$$

Since the speed of light is different in media and the refraction index differs from one, the optical path length in a plasma is not the same as in vacuum. In a plasma the refraction index is typically  $N < 1$ , i.e. a plasma is an optical active medium. An electromagnetic wave, e.g. a laser beam, which is sent through the plasma, gets phase shifted in comparison

to a beam outside the plasma. The phase shift depends on the length of the beam line inside the plasma and the difference of the local refraction index  $N$  from vacuum:

$$\phi = \frac{2\pi}{\lambda} \int_{P_1}^{P_2} [N_{\text{vacuum}} - N(r)] dr. \quad (3.4)$$

Here,  $P_1$  and  $P_2$  are starting and ending points of the probing beam. The refraction index is not necessarily constant along the beam line but depends on the spatial coordinate  $r$ . Using the expression for the refraction index (3.3) one obtains with  $N_{\text{vacuum}} = 1$

$$\begin{aligned} \phi &= \frac{2\pi}{\lambda n_c} \int_{P_1}^{P_2} n(r) dr \\ &= \frac{\lambda e^2}{4\pi\epsilon_0 m_e c^2} \int_{P_1}^{P_2} n(r) dr. \end{aligned} \quad (3.5)$$

The phase shift is proportional to the line-integrated density and the wavelength of the probing beam. Equation (3.5) suggests the use of higher wavelengths for interferometry systems in order to increase  $\phi$ , however, one tends to apply wavelengths in the infrared range to suppress effects of beam diffraction and refraction which will be discussed later.

### 3.1.2 Phase shift measurements

In this section, the basic principles of interferometry and the estimation of the phase shift will be discussed in detail. Technically, different types of interferometers exist. For W7-X a Mach-Zehnder type interferometer is intended to be employed, a schematic view is shown in figure 3.1. The phase shift detection is done by a heterodyne technique as explained in the following, for this the frequency of the reference beam is changed by an acousto-optical modulator (AOM). The interference signal is finally analysed with a phase comparator.

The electric fields of the probing and the reference beam,  $E_p$  and  $E_r$ , are:

$$\begin{aligned} E_p(t) &= E_{0p} \cos(\omega t - \phi) \\ E_r(t) &= E_{0r} \cos(\omega t), \end{aligned} \quad (3.6)$$

where  $\omega$  is the frequency of the beam and  $\phi$  is the phase shift according to equation (3.5).

The power of the interfering beams is proportional to  $(E_p + E_r)^2$ . Using equation (3.6) one obtains

$$P \propto E_{0p}^2 \cos^2(\omega t - \phi) + E_{0r}^2 \cos^2(\omega t) + E_{0p} E_{0r} [\cos(2\omega t - \phi) + \cos(\phi)], \quad (3.7)$$

with  $\cos x \cos y = \frac{1}{2} (\cos(x + y) + \cos(x - y))$  [36].

For infrared or microwave interferometers, the terms varying with the frequency of the probing waves,  $\omega$ , are too fast to detect. The term to be measured in (3.7) is therefore  $S \propto E_{0p} E_{0r} \cos(\phi)$ . Note that this signal is sensitive to the amplitudes  $E_{0p}$  and  $E_{0r}$  of reference and probing beam, which may cause an error for the phase shift measurement if one of the amplitudes changes (e.g. by disturbances of the beam). Furthermore, the sign of  $\phi$  cannot be detected.

To overcome these problems a heterodyne technique is employed. Here, a frequency offset is added to the reference beam by an AOM (see figure 3.1). This offset  $\omega_\psi$  ( $\omega_\psi \ll \omega$ ) is called *intermediate frequency* (IF). The electric field is now given by

$$E_r(t) = E_{0r} \cos(\omega t + \psi); \quad \psi = \omega_\psi t, \quad (3.8)$$

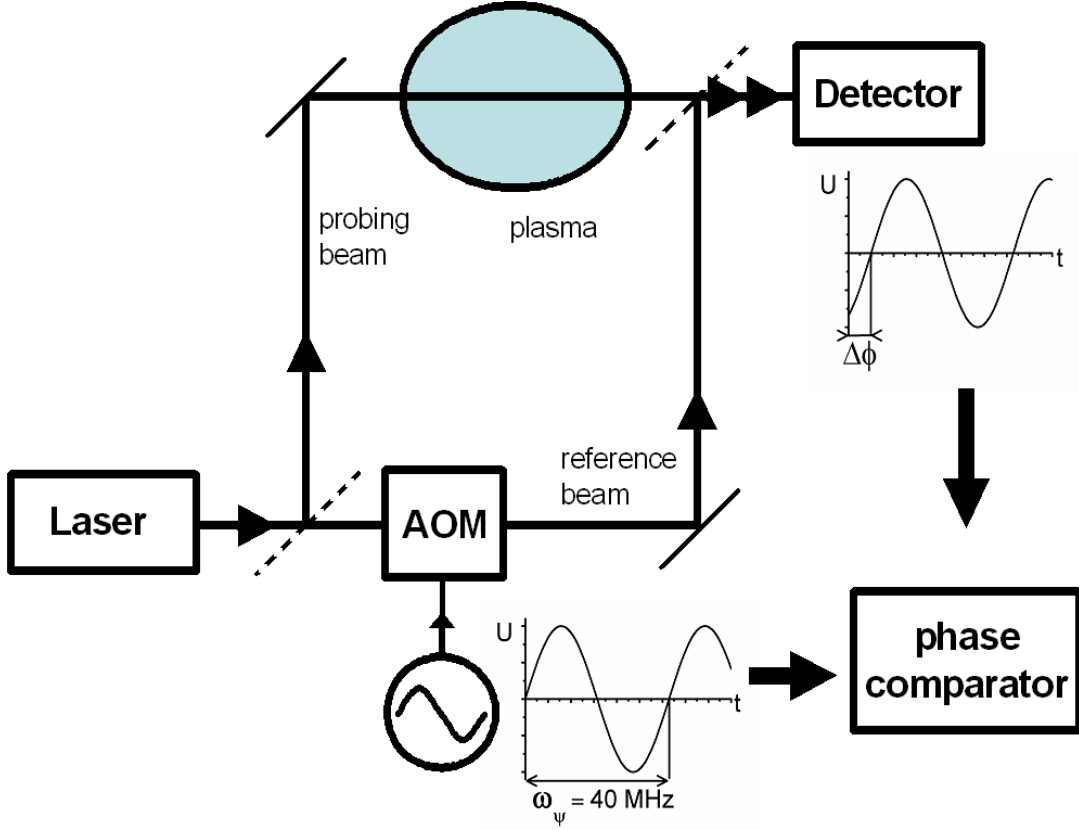


Figure 3.1: Schematic assembly of a Mach-Zehnder heterodyne interferometry system. Reference and probing beam are split, on the reference beam the frequency is changed by an acousto-optical modulator (AOM). After crossing the plasma, where the plasma generated phase shift  $\Delta\phi$  is added to the probing beam, both beams interfere, the interference signal is now modulated with the AOM frequency  $\omega_\psi$ . The modulation signal of the AOM and the detected interference signal are finally analysed by the phase comparator.

and the interference signal (only the part not depending on  $\omega$  in equation (3.7)) becomes

$$S \propto E_{0p}E_{0r} \cos(\phi + \omega_\psi t), \quad (3.9)$$

which is of sinusoidal form. This signal is compared with the modulation signal of the AOM by a phase comparator. To measure the phase shift of the probing beam, the nodes of both signals have to be detected. They are given at  $(k_1, k_2 \in \mathbb{N})$

$$\begin{aligned} \psi_{AOM} &= 2k_1\pi & \text{at the AOM} \\ \psi + \phi &= 2k_2\pi & \text{at the detector} \end{aligned} \quad (3.10)$$

leading to

$$\phi = \psi_{AOM} - \psi + 2\pi(k_2 - k_1). \quad (3.11)$$

Obviously, the estimation of the phase shift is only possible modulo  $m \cdot 2\pi$  ( $m \in \mathbb{N}$ ). For this reason, a uninterruptible measurement has to be realised, only by knowing the "history" of the measurement process phase shifts larger than  $2\pi$  can be detected. In particular, the zero-crossings of the signals have to be counted.

With  $\psi_{AOM} = \omega_\psi t_1$  and  $\psi = \omega_\psi t_2$  and introducing the pseudo-period  $\tau$ :  $\psi(t + \tau) = \psi(t) + 2\pi$  one obtains

$$\phi = \frac{t_1 - t_2}{\tau} \cdot 2\pi. \quad (3.12)$$

Equation (3.12) shows that measuring the instants of the nodes of both the AOM ( $t_1$ ) and the detector signal ( $t_2$ ) the phase shift  $\phi$  can be estimated. Because the phase shift is determined only once per time interval  $\tau$ , this period has to be chosen adequately small for a sufficient phase resolution. The phase shift  $\phi$  is then assumed to be constant during one time period  $\tau$ . In result, the estimation of  $\phi$  does no more depend on the amplitude but only of the nodes of the sinusoidal interference signal  $S$ . In addition, the sign of  $d\phi/dt$  can be derived from the sign of  $\frac{d}{dt} \frac{t_1 - t_2}{\tau}$ .

For interferometry systems operating in the far infrared (FIR) region, the additional frequency shift  $\omega_\psi$  is normally generated by utilising the Doppler shift of a moving mirror or a rotating cylindrical grating. For systems at medium infrared (MIR) or visible (VIS) range, acousto-optic modulators or electro-optic modulators (EOM) are used (see, e.g., [37]): The probing laser beam is sent through a crystal, in which a sound wave is induced, e.g. by a piezo-electric transducer. The sound wave in the crystal forms an optical grid, which leads to a diffraction of the probing beam. In case of a moving sound wave, the diffracted beam is frequency-shifted due to the Doppler effect, whereas the 0-th order diffraction is not influenced. These modulators can therefore be used as beam splitter (separating reference and probing beam) as well as for the shift of the frequency of the reference beam. As an advantage of this method, the number of optical elements in the beam line is reduced. For Wendelstein 7-X, a modulation frequency of  $\omega_\psi = 40 \text{ MHz}$  is planned for the  $CO_2$  laser and  $\omega_\psi = 25 \text{ MHz}$  for the  $CO$  laser (see below).

## Vibration compensation

The leading source of errors in the measurement of the plasma density by interferometry are vibrations. Vibrations are due to, e.g., effect of vacuum pumps or vibrations of the building and cannot be avoided in most cases. If short wavelengths are used for the probing beam (see section 3.2), the phase shift induced by vibrations can be much larger than the shift generated by the plasma density: E.g., for a  $CO_2$  laser ( $\lambda = 10.6 \mu\text{m}$ ) and a line integrated density of  $10^{20} \text{ m}^{-2}$  one obtains a phase shift of  $\phi \approx 3 \text{ rad}$ , whereas a displacement of an optical element (mirror) of  $\delta = 0.1 \text{ mm}$  due to vibrations would cause a phase shift of  $\phi_{vib} = 2\pi \frac{\delta}{\lambda} \approx 59 \text{ rad}$ .

While the phase shift due to the plasma scales with  $\lambda$  (see equation (3.5)), the vibration induced shift is inversely proportional to the wavelength, in result, the ratio  $\phi/\phi_{vib} \propto \lambda^2$ . To reduce the effect of vibrations, the use of higher wavelength is therefore an appropriate choice [38]. But as mentioned before, for an interferometry system working in the sub-millimeter or millimeter range, effects of diffraction and refraction occur as an additional error source.

If wavelengths in the infrared range are used, a common method for vibration compensation is a two-color (or two-wavelength) interferometer, as planned for W7-X. Here, the interferometer set-up of figure 3.1 is doubled: Two probing beams with different wavelengths  $\lambda_1$  and  $\lambda_2$  are guided on an identical path through the optical system, both beams are exposed to the same vibrations and the same plasma density effects.

Vibrations yielding a change  $\delta$  in the optical path generate a phase shift  $2\pi \frac{\delta}{\lambda}$  to each

beam, so the phase shifts of both probing beams are given by

$$\begin{aligned}\phi_1 &= \varphi_1 + 2\pi \frac{\delta}{\lambda_1} \\ \phi_2 &= \varphi_2 + 2\pi \frac{\delta}{\lambda_2};\end{aligned}$$

$\varphi_i$  are the phase shifts generated by the plasma. Because the vibrations are the same for both beams, one obtains

$$\begin{aligned}\phi_1\lambda_1 - \phi_2\lambda_2 &= \varphi_1\lambda_1 - \varphi_2\lambda_2 \\ &= \frac{(\lambda_1^2 - \lambda_2^2)e^2}{4\pi\epsilon_0 m_e c^2} \int n(r)dr\end{aligned}\quad (3.13)$$

and finally

$$\begin{aligned}\int n(r)dr &= \frac{4\pi\epsilon_0 m_e c^2}{e^2} \frac{\phi_1\lambda_1 - \phi_2\lambda_2}{\lambda_1^2 - \lambda_2^2} \\ &\approx \frac{4\pi\epsilon_0 m_e c^2}{e^2} \frac{\phi_1\lambda_1 - \phi_2\lambda_2}{\lambda_1^2} \quad \text{for } \lambda_2 \ll \lambda_1.\end{aligned}\quad (3.14)$$

For  $\lambda_2 \ll \lambda_1$ ,  $\phi_2$  measures the vibration-induced phase shift directly. In practice the diagnostician has to decide between two limits: On one side, a large wavelength difference is beneficial as shown by Kawano et al. [39]: The accuracy of the density measurement (resolution  $\Delta n$ ), assuming a resolution of  $\Delta\phi$  for the phase detection, is given by

$$\Delta n \propto \frac{\Delta\phi}{\lambda_1} \frac{1}{(1 - \lambda_2/\lambda_1)},\quad (3.15)$$

i.e. a large difference in wavelength leads to a better density resolution.

On the other hand, a small difference is favourable for practical applications: Beams with similar wavelengths have similar Gaussian beam characteristics, can use the same optical elements along the beam path and are insensitive to wavelength-dependent disturbances like, e.g., the thermo-optical effect in windows [40].

An effective vibration compensation requires, that as many optical elements as possible are used by both beams simultaneously. Therefore, at W7-X it is planned to use only one detector to measure the interference signal of both wavelengths. This is possible by applying different intermediate frequencies ( $\omega_\psi = 40 \text{ MHz}$  for the  $CO_2$  laser and  $\omega_\psi = 25 \text{ MHz}$  for the  $CO$  laser), the interference signal is later decoupled using a band filter at the phase comparator.

## 3.2 The interferometry system at Wendelstein 7-X

The interferometry system planned for Wendelstein 7-X will be described next. After a short introduction, the technical boundary conditions and error sources will be discussed.

### 3.2.1 A two-color, multi-channel interferometer

For Wendelstein 7-X, a multi-channel interferometer is proposed [15]. The diagnostic will start as a four-channel interferometer, it is planned to be extended to eight or more beams later.

For the compensation of vibrations, the interferometry system will use two wavelengths in the infrared (IR) region, generated by a  $CO_2$  and a  $CO$  laser. The wavelengths and

Laser	Wavelength $\lambda$	Cut-off density $n_c$
<i>CO</i>	5.4 $\mu m$	$9.92 \cdot 10^{24} m^{-3}$
<i>CO<sub>2</sub></i>	10.6 $\mu m$	$3.82 \cdot 10^{25} m^{-3}$

Table 3.1: Laser wavelengths and cut-off densities for the W7-X interferometer

the cut-off densities are given in table 3.1, both densities are far beyond the expected maximum densities of about  $4 \cdot 10^{20} m^{-3}$  for the W7-X plasma.

The probing wavelengths are short compared to previous experiments (see, e.g., [41] for W7-AS), where millimeter waves ( $\lambda \approx 1.87 mm$  in case of W7-AS) have been used. As a result, the phase shift generated by the plasma will be much less for the W7-X system. But this is not necessarily a disadvantage: Because the phase measurement can be performed today with a resolution of  $\Delta\phi < 1/100 \times 2\pi$  and better, counting a large number of full  $2\pi$ -shifts is not necessary for accurate density measurements. Moreover, if only a small phase shift has to be detected, the danger of losing the signal by phase jumps of  $2\pi$  (see section 3.1.2) is reduced.

The laser wavelengths have been chosen according to the arguments given in section 3.1.2: Both wavelengths are close enough to use the same optical components, AOM and detectors. On the other hand, the difference is sufficiently large for a good resolution in density measurement, the degradation factor (equation (3.15)) is given with

$$\frac{1}{(1 - \lambda_2/\lambda_1)} \approx 2. \quad (3.16)$$

Furthermore, both the *CO<sub>2</sub>* and the *CO* laser provide beams with a sufficient high power output, they are commercially available as well as the necessary optical components for guiding and manipulating the beams.

### 3.2.2 Technical boundary conditions

For the interferometry system at W7-X, three ports are available for the access to the plasma vessel and the plasma itself. They are located at a toroidal angle of  $\Phi \approx 195^\circ$  at the plasma vessel (see figure 3.2, left). The three ports allow different poloidal angles, however, the beams will not be able to cross the plasma at the longest possible line (figure 3.2, right).

An important fact regarding the chosen ports is that the beam lines cannot cross the plasma vessel directly, since the ports are localised only at the outer side of the vessel. Therefore, the beams have to be reflected on the wall opposite to the ports and will cross the plasma two times.

This restriction can be seen as an advantage, because crossing the plasma twice leads to a higher phase shift and therefore to a better signal-to-noise ratio. On the other hand, the incident beam has to be separated from the beam leaving the plasma, which makes additional optical elements (semipermeable mirror) necessary, leading to a reduction in the intensity of the laser beam.

Furthermore, the diagnostic is highly exposed to vibrations of the plasma vessel, since the optical system, especially the mirrors inside the plasma vessel (corner cube retro-reflectors), cannot be mechanically decoupled from the machine. In addition, the reflecting components on the wall have to fit to the structure of in-vessel components (see figure 3.3). This leads to restrictions of the mirror positioning, in fact, only 11 fixed reflector positions are available.

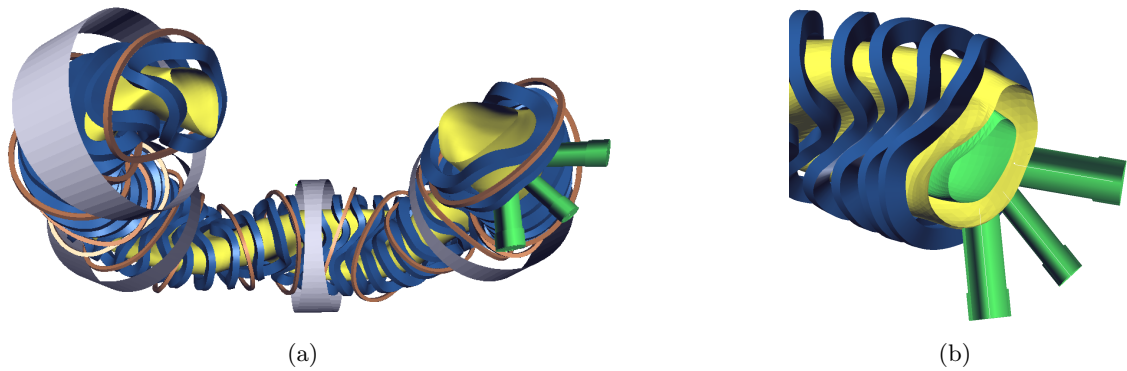


Figure 3.2: A schematic view of the Wendelstein 7-X assembly (a): The vacuum vessel (grey), the coil system (blue and brown) and the plasma vessel (yellow). Three of the 299 ports are dedicated to the interferometry system (green). A cross section of the interferometry plane (b) shows that different poloidal angles are possible for the beam lines (plasma: green area). (The plots were created using the *VirtualWendelstein* software [42].)

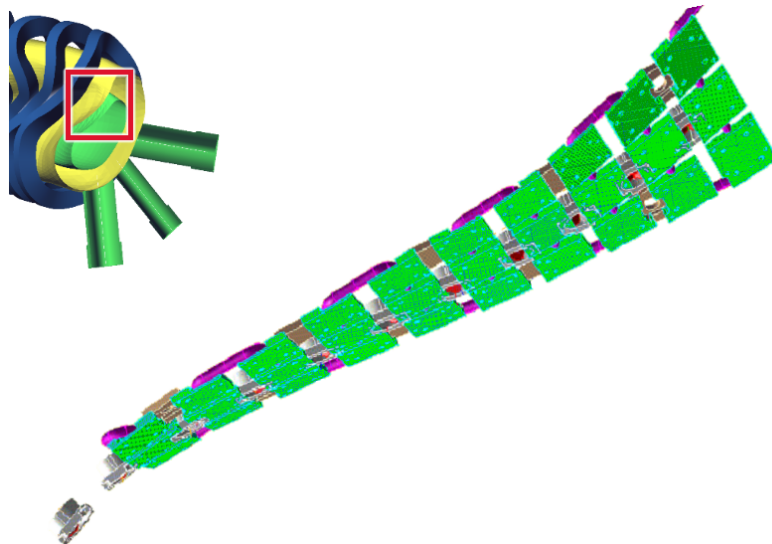


Figure 3.3: Assembly of the wall elements at the inner side of the plasma vessel: The corner cube retro-reflectors (grey) have to fit the in-vessel components, like carbon tiles (green) and the cooling system (violet). The position of the assembly with respect to the ports is indicated by the red square in the small picture.



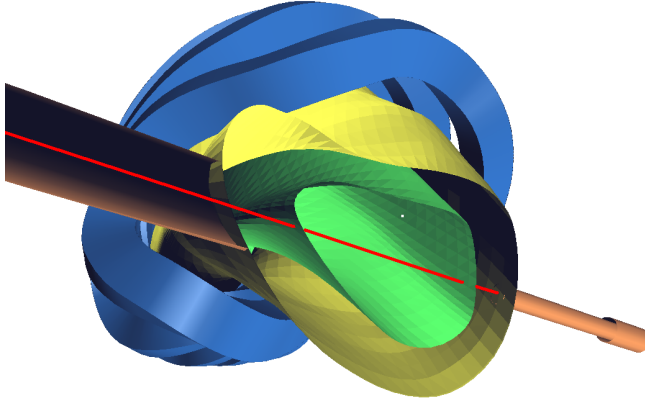


Figure 3.4: Additional interferometry channel: Cross section of the W7-X plasma vessel (yellow) at a toroidal angle of  $\varphi \approx 187^\circ$ , this plane is symmetrical to the diagnostic plane of the Thomson scattering diagnostic. The extra beam (red), the plasma (green), the plasma vessel (yellow) and the coil system (blue) are indicated.

An extra beam line can be positioned at a toroidal angle of approximately  $187^\circ$ . The plasma shape at this plane is nearly triangular (see figure 3.4), the beam is congruent to the probing beam of the Thomson scattering diagnostic due to the symmetric properties of the W7-X stellarator. This offers the possibility to use an interferometry beam as absolute calibration for Thomson scattering, which provides density and temperature profiles with high spacial resolution, but not independent of each other. The extra beam is a proposal resulting from the requirements of an *Integrated Data Analysis (IDA)* [26]. The influence of this beam to the design of the interferometry system will be discussed in chapter 5.

### 3.2.3 Error sources of the interferometry system

For the construction of the likelihood function (section 2.1) the error distribution of the phase shift has to be estimated. The main errors should be therefore discussed here. The leading error is given by the measurement of the phase shift, other errors (table 2.2) are negligible.

The problems arising by refraction and the angular deviation of the probing beam as described by Veron in [35], Chap. VII, can be neglected for the W7-X interferometry system due to the very small wavelengths of the probing beams. The examples used by Veron were made for FIR interferometry systems  $\lambda = 0.337 \text{ mm}$ , therefore the effects are about 1000 times larger than described for W7-X (the error is proportional to  $\lambda^2$ ): One would expect errors in the order of some  $10^{-4}$  percent of the phase shift, which can be ignored.

The effect of Faraday rotation can also be ignored: The presence of a magnetic field leads to a rotation of the polarisation plane of the probing beam. This influences the interference of probing and reference beam, which would lead to intensity effects of the interference signal. This does not play a role for the phase detection, where only the

Error	Description	Comment
Refraction index	$N = \sqrt{1 - \frac{n}{n_c}} \approx 1 - \frac{n}{2n_c}$ for $n \ll n_c$	$\Delta N = 1, 3 \cdot 10^{-11}$ at $n = 1 \cdot 10^{20} m^{-3}$ and $\lambda = 10.6 \mu m$
ordinary and extraordinary wave	$E \parallel B$ not established, refraction index: $N_o = 1 - \frac{\omega_p^2}{2\omega^2}$ ; $N_x = 1 - \frac{\omega_p^2}{2\omega^2} \cdot \frac{1}{1 - \frac{\omega_{ce}^2}{\omega^2 - \omega_p^2}}$	W7-X: $\omega = 1, 777 \cdot 10^{14}$ , $B = 3 T$ , $n = 10^{20} m^{-3} \rightarrow$ $\frac{N_x}{N_o} = 1.0000088$
Veron [35]: Angular deviation of probing and reference beam	$\epsilon_1 \propto \Delta\phi/\phi$ $\epsilon_1 \approx 1.5 \cdot 10^{-4} \%$	additional phase shift for: maximum deflection, parabolic density profile, cylindrical plasma.
Change of optical path length	$\epsilon_2 \propto \Delta\phi/\phi$ $\epsilon_2 \approx -6 \cdot 10^{-4} \%$	parabolic density distribution assumed
Overall error	$\epsilon = \epsilon_1 + \epsilon_2$ $\epsilon \approx -0.0004 \%$	for W7-X
Stability laser	additional phase shift: $\Delta\phi = \frac{2\pi \delta f \Delta l}{c}$	$\delta f$ - frequency shift laser $\Delta l$ - difference in optical path of probing and reference beam $\Delta\phi \approx 0.001 \cdot 2\pi$

Table 3.2: Negligible errors for the W7-X interferometry system. The error resulting from beam deviation is approximated by assuming a circular shaped plasma and a parabolic density distribution, which leads to an estimate for the error.

nodes of the sinusoidal interference signal are detected (see equation (3.12)). Assuming a constant density for one time period  $\tau$ , the Faraday rotation is also constant and does not affect the estimation of the nodes.

An error source for the phase shift may occur, if the laser frequency  $f$  is not stable. The instability of the laser lies in the range of some  $MHz$  in case of the  $CO$  laser: Due to the long resonator of this laser, more than one laser mode may be activated. The  $CO_2$  laser frequency, however, is much more stable. By a shift of  $\Delta f$  an additional phase shift of

$$\Delta\phi = \frac{2\pi \Delta f \Delta l}{c} \quad (3.17)$$

occurs, where  $\Delta l$  is the difference in the optical path between probing and reference beam. For typical values of  $\Delta l \approx 10 \text{ cm}$  the additional phase shift is about  $\Delta\phi \approx 0.001 \cdot 2\pi$ .

Several errors arise during the derivation of equation (3.5). First, the approximation in equation (3.3):

$$N = \sqrt{1 - \frac{n}{n_c}} \approx 1 - \frac{n}{2n_c} \quad (3.18)$$

For  $n_c \approx 10^{25} \text{ m}^{-3}$  and  $n \approx 10^{20} \text{ m}^{-3}$  one gets a difference of  $\Delta N \approx 10^{-11}$ , which will not play any role.

Another error of this kind occurs if  $\mathbf{E} \parallel \mathbf{B}$  is not established as assumed in equation (3.2). In case of the *ordinary wave* where the electric and the magnetic field are perpendicular the refraction index  $N$  is given by

$$N_o = 1 - \frac{\omega_p^2}{2\omega^2} \quad (3.19)$$

with  $\omega_p = \sqrt{\frac{ne^2}{\epsilon_0 m_e}}$  as the plasma frequency. In case of an *extraordinary wave* ( $\mathbf{E} \perp \mathbf{B}$ ) one obtains

$$N_{ext} = 1 - \frac{\omega_p^2}{2\omega^2} \cdot \frac{1}{1 - \frac{\omega_{ce}^2}{\omega^2 - \omega_p^2}}, \quad (3.20)$$

where  $\omega_{ce} = \frac{eB}{m_e}$  is the electron gyro frequency. For  $B = 3 \text{ T}$  one finally gets 1.0000088 as the correction factor for the extraordinary wave, which is also a negligible effect.

The leading error is given by the measurement of the phase shift by the phase comparator, which compares the AOM and the interference signal electronically. Both signals are sinusoidal and are transformed into rectangular shape for better comparison. For phase measurement, the phase comparator counts the edges of both rectangular signals, the phase shift is estimated by counting over time intervals larger than the modulation time  $t_\psi = \frac{2\pi}{\omega_\psi}$  [43]. The phase comparator itself has a very high resolution (up to  $1/500 \cdot 2\pi$  and better, see Kawano et al. [37]), the main error occurs during the transformation of the interference signal into a rectangular shape.

The edges of the rectangular signal are defined by the zero-crossings of the interference signal (see figure 3.5), which varies with the modulation frequency  $\omega_\psi$  of the AOM. This signal is noisy due to several disturbances on the beam line (e.g. vibrations, the vibration compensation is done later by the phase comparator). This leads to an error of the phase shift measurement, because the zero-crossing of the noisy signal may not be measured correctly.

This error is independent of the magnitude of the phase shift and can therefore be seen as a constant background noise. The error distribution will be assumed as Gaussian for further calculations. Because no experimental experience is given so far, the error

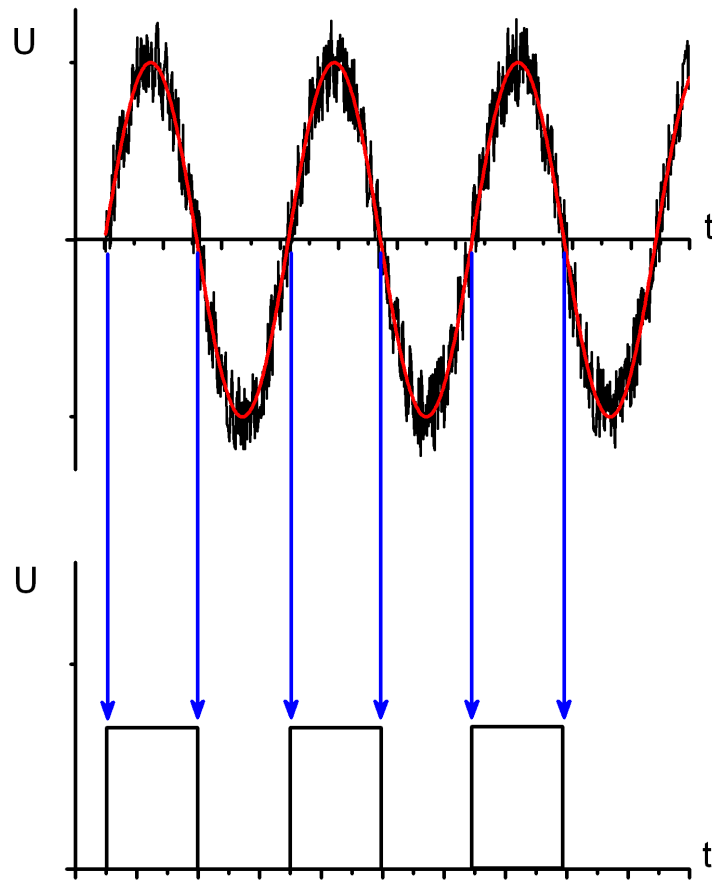


Figure 3.5: Transformation of the sinusoidal interference signal at the detector (above) into rectangular shape (bottom). The edges of the rectangles are given by the zero-crossings of the noisy detector signal.

is conservatively assessed with  $\sigma = 0.1 \text{ rad} \approx 0.016 \cdot 2\pi$ . This corresponds to the expected resolution of a phase meter of about  $0.01 \cdot 2\pi$  [44] and takes into account the error propagation by using two phase meters (one for each wavelength).



## Chapter 4

# Basic Case Studies and Methodological Validation

In this chapter the principle of using the Expected Utility as a measure for the quality of a design will be analysed. As a reference, a previous design approach for the W7-X interferometer is shown first, the results are compared with the outcome of Bayesian experimental design. Case studies for a one-channel interferometer are presented to study the "working principle" of the *EU* on examples which can be explained in figures of the signal-to-noise ratio, modified with respect to the parameters of interest.

Finally, the meaning of the *bit*-unit as an expression of the information gain will be illustrated using old data sets from W7-AS, the predecessor experiment of W7-X. For this, the data sets are analysed, and the *EU* of the respective measurement is calculated independently. A monotonous relationship between the quality of the measurement result, expressed by a global deviation  $\chi^2$ , and the *EU* is found.

### 4.1 Previous diagnostic optimisation results for W7-X

For the interferometer of Wendelstein 7-X an earlier optimisation has been performed by Airila et al. [45]. The results will be discussed here shortly and be compared with the new design later.

The approach presented in the work of Airila et al. is based on the idea of a plasma structured in shells (figure 4.1). They are formed by the magnetic flux surfaces, the density is assumed as constant inside a shell (see also section 4.3.1). Additionally, the shells are of equal width in the effective coordinate system, their number is equal the number of beam lines: four and eight, respectively.

Parameters of interest are now the densities in the respective shells. For the estimation of the phase shift, equation (3.5) changes to

$$\phi_i(\boldsymbol{\theta}) = \frac{\lambda e^2}{4\pi\epsilon_0 m_e c^2} \sum_j A_{ij} \cdot \theta_j, \quad (4.1)$$

where  $\theta_j$  is the density in shell  $j$ . The element  $A_{ij}$  of the projection matrix  $\mathbf{A}$  describes the length of beam  $i$  inside shell  $j$ . Here, the mapping from magnetic to real space coordinates is implemented in the projection matrix, too. The design ansatz of Airila et al. is to find the optimal measurement of the density in every shell.<sup>1</sup> For this, every single beam line is assigned to a single shell.

---

<sup>1</sup>This is the reverse of the inversion procedure which will be introduced later in section 4.3.1.

This leads directly to the requirement that the projection matrix has to be as diagonal as possible: The contribution of shell  $i$  to beam  $i$  (described by  $A_{ii}$ ) should be large, so that the phase shift of beam  $i$  is mainly determined by matrix element  $A_{ii}$ :

$$\phi_i \approx A_{ii} \cdot \theta_i. \quad (4.2)$$

The optimisation goal of Airila et al. was to find a sight line configuration for which the diagonality of the projection matrix is maximal. In result (see fig 4.1), the beam lines cross "their" shell at the longest possible way (restricted by the technical limitations due to the port system). In particular, the beam lies as close as possible to the next inner shell.

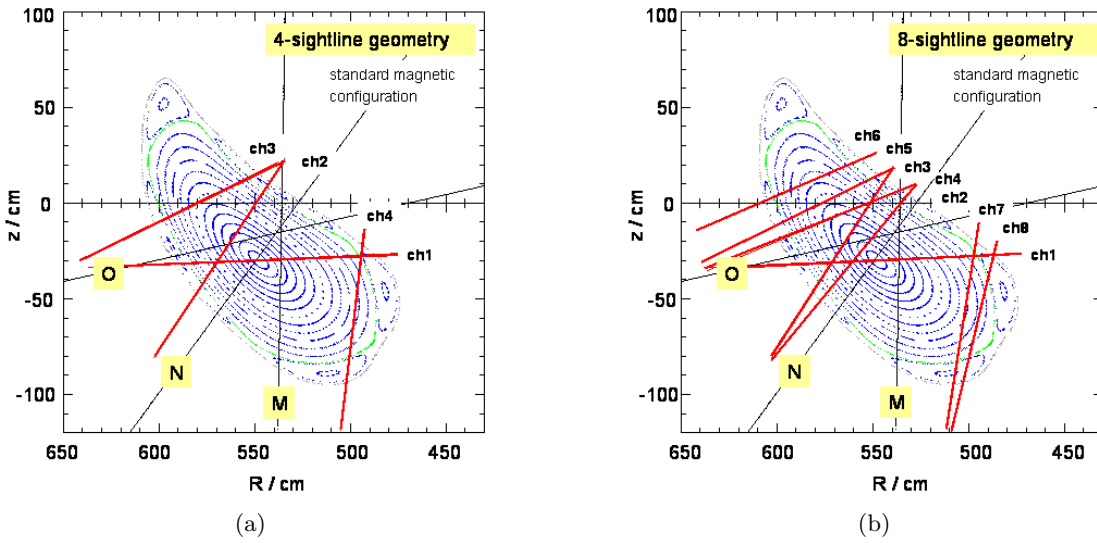


Figure 4.1: Optimisation result for a four (a) and eight (b) channel interferometer at W7-X from Airila et al. [45]. The beams are tangent to the next-inner shell, respectively.

## Comments on the results

The design results were tested with five different density distributions: Using the optimised interferometer for the generation of artificial data, the distributions could be reconstructed quite well. Different to Bayesian experimental design, this test was done after the design of the interferometer was finished. The reconstruction of the density distributions was not implemented in the optimisation procedure, the directions of the beam lines are completely determined by the parametrisation, i.e. mapping expressed by the shell structure. In contrast, in the *BED* approach classes of density distribution functions are implemented as parameters of interest.

The design of Airila et al. is fixed to the approach of equally sized shell widths (in the space of effective coordinates), a change of the widths is not intended. As will be shown later (section 4.2.2), this leads to a design where the optimised beam lines are strongly related to the assumed shell structure.

## 4.2 Case studies: One-beam interferometer

### 4.2.1 The basic setting

It seems to be useful to start the presentation of the results from Bayesian experimental design with some simple cases to introduce, on the one hand side, the parametrisation chosen for the problem. On the other hand, the optimal design of such examples is often known, which offers the possibility to benchmark the results from *BED*.

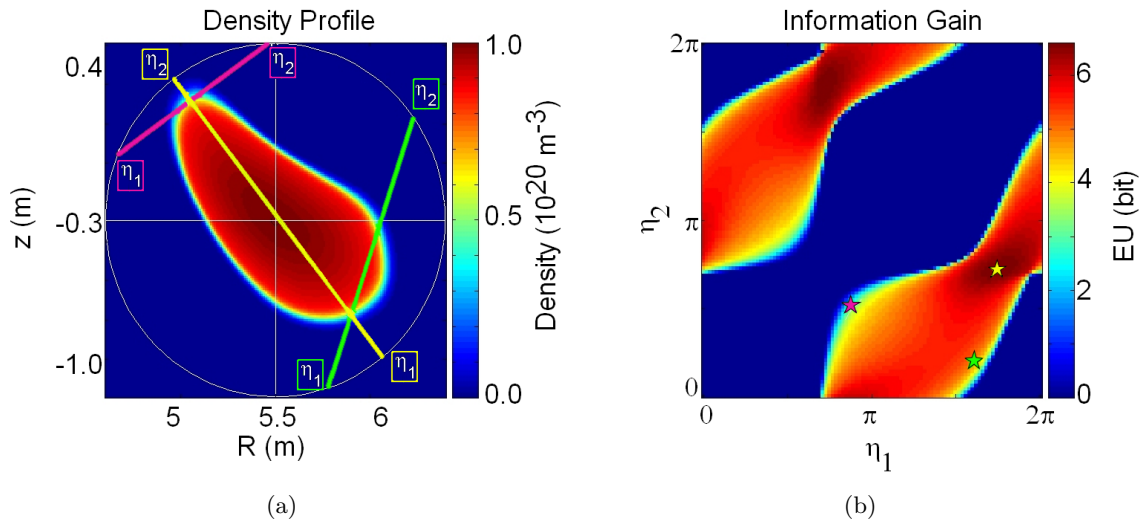


Figure 4.2: Density distribution for a cross section of the W7-X plasma at the interferometry plane (a) and the expected information gain for one beam (b). The colour-coded lines of sight in the left picture correspond to the star symbols on the right side, respectively.

To parametrise the line of sight of a probing beam, the starting and ending points of the beam may lie on a circumventing circle outside the plasma whose centre is identical with the plasma centre. Both starting and ending point can therefore be described by two angles  $\eta_1$  and  $\eta_2$ . These angles are the design parameters which have to be varied in order to optimise the diagnostic design. The concept is shown in figure 4.2: For a cross section of the W7-X plasma at the diagnostic plane of the interferometer (toroidal angle  $\Phi \approx 195^\circ$ ), several possible beam lines are indicated. The beam lines correspond to the star symbol in the plot of the information gain (figure 4.2 (b)).

A simple optimisation problem occurs, e.g., for a cylindrical plasma. In the first example, the maximum density of the plasma,  $\theta_{max}$ , shall be measured with only one line of sight, assuming a quadratic density profile:

$$n(r_{eff}) = \theta_{max} \cdot (1 - r_{eff}^2) \quad (4.3)$$

This question may arise, e.g., if the density control in a plasma is realised by an interferometer beam. For simplification, the probing beam has to be perpendicular to the axis of the plasma cylinder. For the error statistics a constant noise level independent of the data is assumed.

Figure 4.3 shows the varied density profile and a cross section of the plasma. The beam line is described by only two angles, the Expected Utility (*EU*) is plotted with respect to both design parameters. The plot is symmetric in  $\eta_1$  and  $\eta_2$ , because the start and end point are exchangeable.



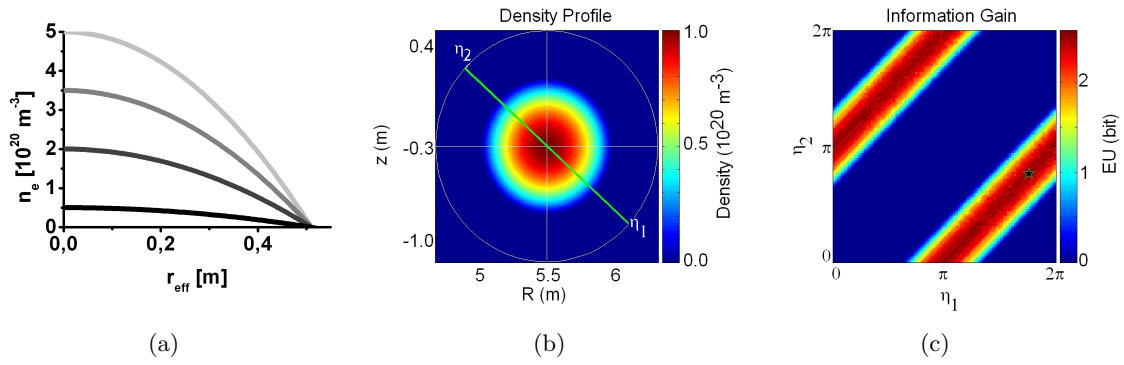


Figure 4.3: Estimation of the maximum density with one beam: The amplitude of a quadratic density profile varies from  $0.5 - 5 \cdot 10^{20} \text{ m}^{-3}$  (a) in a cylindrical plasma. The plasma cross section is shown in (b). The probing laser beam is parametrised by the angles  $\eta_1$  and  $\eta_2$ . The *EU* is displayed in (c), the star symbol corresponds to the beam line in (b).

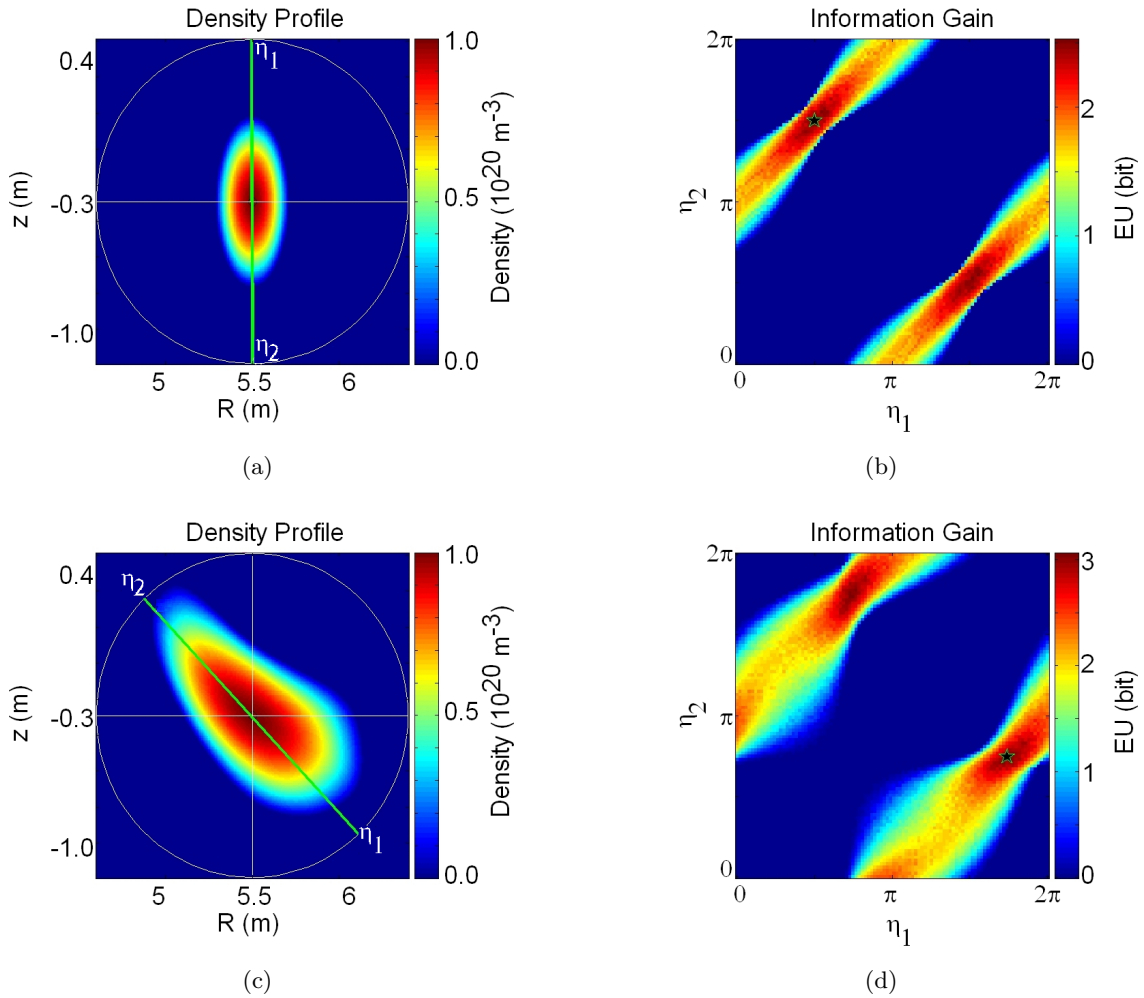


Figure 4.4: Design for a one-channel interferometer for an elliptical plasma (upper row) and for a W7-X shaped plasma cross section (lower row). The green beam lines in the left pictures correspond to the star symbols in the respective right picture.

As one can see in figure 4.3 (c), the  $EU$  shows a maximum for beam lines crossing the plasma centre. This was already expected: For a parabolic density profile, these lines of sight yield a maximum phase shift and therefore an optimal signal-to-noise ratio (SNR) in case of a constant background noise. Also, the  $EU$  is zero for beams outside the plasma: No information can be gained from these measurements.

For the same physical problem (estimation of the maximum density for a quadratic distribution), a slightly more complex geometry is analysed in figure 4.4, upper row: In an elliptical plasma (similar to elongated tokamak configurations), the  $EU$  shows a clear maximum. It corresponds to a beam crossing the plasma through its centre at the longest possible way: The SNR is maximised for this beam. In the lower row of the figure, the design result is shown for a plasma cross section of Wendelstein 7-X, here for the interferometry plane at a toroidal angle of  $\Phi = 195^\circ$ . Again, the beam line traversing the plasma on the longest way is preferred by the design.

Nevertheless, while this first example shows the principle of designing the interferometry system, it has to be extended to more realistic conditions for the further design process. As a first point, the parametrisation of the density function should not be restricted to a parabolic approach, since this would limit the educible space of distributions. The parametrisation of the density profile functions is therefore discussed next.

## 4.2.2 Parametrisation of the density distribution

The goal of interferometry measurement is to determine the density distribution of the plasma. The parameters of interest are therefore the variables in the mathematical description of the distribution. Several approaches for such descriptions are possible, starting from the already mentioned stepwise constant distribution to other classes of model functions. In this section, these approaches are discussed, the model function which are later used for optimisation are introduced.

### Stepwise constant distributions

For the analysis of interferometric data it is possible to use stepwise constant density functions, as it is shown later in section 4.3.1. Here, the parameters of interest are the densities in the different shells. The question arises whether this parametrisation is also useful for diagnostic optimisation.

Figure 4.5 shows a density distribution consisting of four shells and the expected utility for the optimisation of a single beam line. The design parameters are the starting and the ending point of the line of sight, parametrised by two angles,  $\eta_1$  and  $\eta_2$ . The parameters of interest are the densities in the individual shells, they were assumed as independent of each other and varied between  $0.0 - 2.0 \cdot 10^{20} m^{-3}$ .

The expected utility (figure 4.5 (b)) shows maxima in ribbon-like structures, corresponding to beam lines tangential to the next-inner shell, respectively. These lines of sight cross a density shell on the longest possible way, this result is equivalent to the diagonality criterion of Airila et al. (section 4.1 and [45]). The  $EU$  drops rapidly if the next-inner shell is intersected.

However, for the optimisation of the diagnostic the approach of shells with constant densities does not seem to be appropriate: The maxima of the  $EU$  are clearly influenced by the shell structure. But a design depending on the shell configuration seems to be disadvantageous, since the number of shells and their width are not fixed quantities.

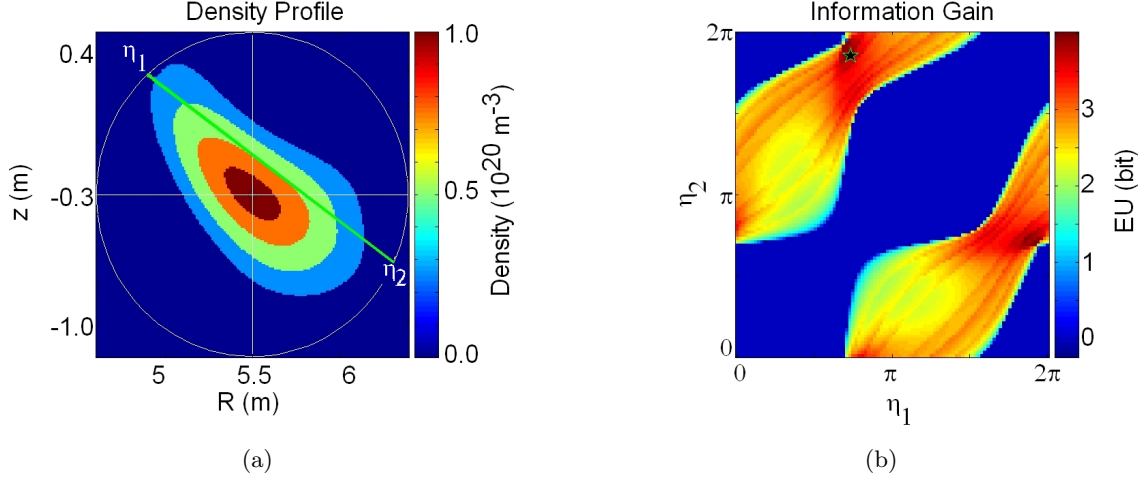


Figure 4.5: Optimisation for stepwise constant, shell-structured density distribution, a constant density is assumed for each shell (a). The star symbol in the  $EU$  distribution (b) corresponds to the beam line in the density plot.

### Form-free parametrisation

To overcome the problem of stepwise constant density distributions, the use of other fitting functions is also applied in data analysis, like linear, quadratic or cubic spline solutions (see, e.g., [46]) as well as exponential splines [47]. These methods provide flexibility and, in most cases, access to the first and higher derivatives of the profile function.

Parameters of interest as criterion for diagnostic design would be the number and position of the spline knots as well as the spline coefficients. However, for the optimisation process the range of these parameters of interest has to be restricted according to the physical problem. This is different to data analysis, where the result of the measurement is not known and therefore in principle all combinations of the parameters of interest are allowed.

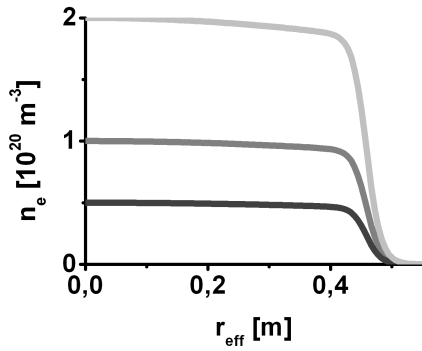
For diagnostic optimisation, physical questions as described later in section 5.1 would lead to very restrictive limitations for the parameters of form-free distribution functions. But these limitations may be hard to define due to the interdependencies of the spline parameters.

### Parametrised model function

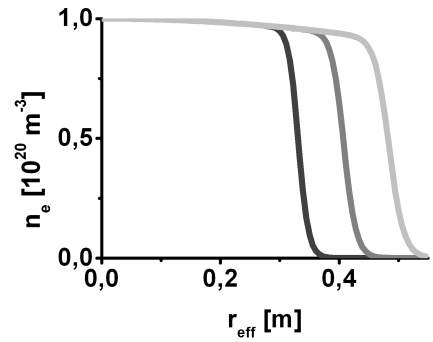
For the illustration of density profiles at W7-AS some parametrised model functions were applied [48]. An example is

$$n_e(r_{eff}) = \theta_1 \cdot 10^{20} m^{-3} \cdot \left[ \frac{1 - \theta_4 \cdot (r_{eff}^2/a^2)}{1 + (r_{eff}^2/(\theta_2 \cdot a)^2)^{\theta_3}} \right]. \quad (4.4)$$

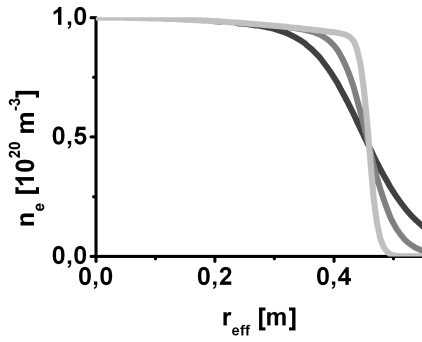
Here,  $a$  is the minor radius of the plasma, i.e. the effective radius of the plasma at the last closed magnetic flux surface. These function turns out to be very flexible, although it is determined by only four parameters,  $\theta_1 - \theta_4$ . These parameters describe unique characteristics of the function: maximum density ( $\theta_1$ ), position of the steepest gradient ( $\theta_2$ ), steepness of the density decay at the edge ( $\theta_3$ ) and the pronouncement of the bulge ( $\theta_4$ ). In figure 4.6 the effect of the parameters is shown.



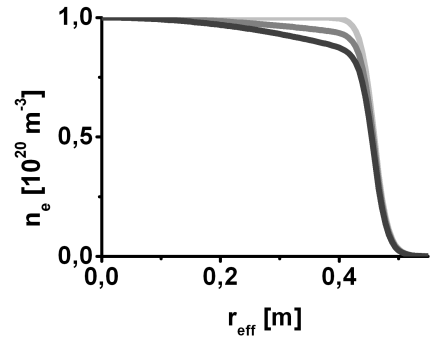
(a)



(b)



(c)



(d)

Figure 4.6: Effect of the parameters  $\theta_1 - \theta_4$  from the density model function (4.4): Maximum density  $\theta_1$  (a):  $0.5 \leq \theta_1 \leq 2.0$ ; position of the steepest gradient  $\theta_2$  (b):  $0.6 \leq \theta_2 \leq 0.95$ ; steepness at the plasma edge  $\theta_3$  (c):  $5 \leq \theta_3 \leq 30$ ; and the bulge of the profile  $\theta_4$  (d):  $-0.2 \leq \theta_4 \leq 0.0$ .

The advantage of this model function lies in the independency of the single parameters: Changing one has no significant effect on the distribution properties described by another parameter. Therefore, different effects can be analysed separately by varying the respective parameter.

One may argue that using a model function restricts oneself to the class of density distributions described by this model. This is correct, and in case of data analysis, where the outcome of the measurement should not be biased by the model function, the use of such parametrisation seems questionable. For diagnostic optimisation, on the other hand, the diagnostic has to be designed according to some well-defined physical effects. If these effects can be adequately described by a model function like (4.4), parametrised functions can be applied. Using a form-free parametrisation like splines, on the other hand, may lead to complex limitations of its parameter space, if the same physical effect shall be described. This could also lead to an increase in computation time.

The function presented here can be easily extended to describe hollow profiles (equation (5.8)) or peaked profiles (equation (5.10)). Therefore, it will be used in section 5.1 to parametrise the density effects of interest.

### 4.2.3 One-beam design for density model function

The effect of the parameters in the density model function (4.4) on the design of a one-beam interferometer shall be pointed out next. Therefore, the parameters were chosen one by one as parameters of interest for the design of one line of sight. The computational aspects for the calculation of the  $EU$  are described in appendix A.

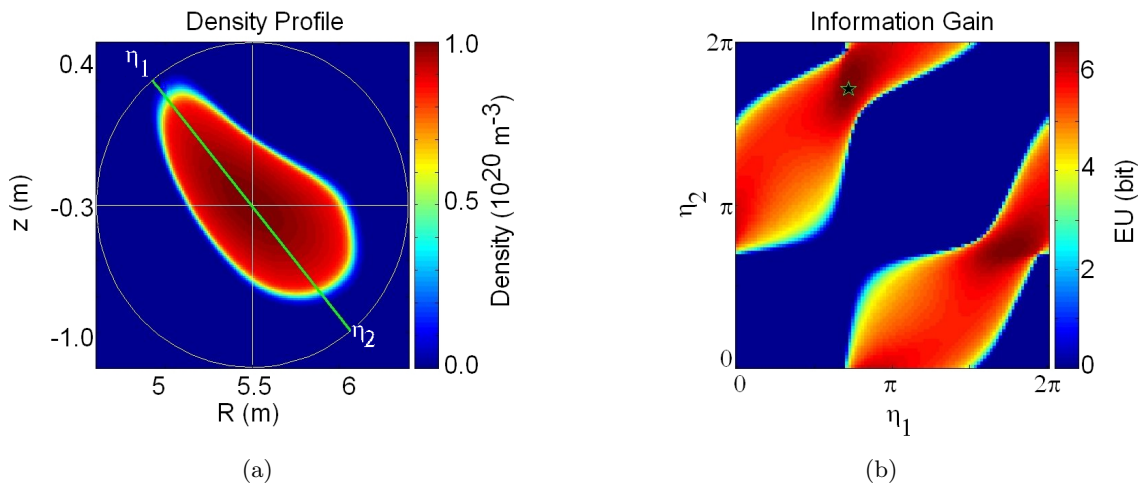


Figure 4.7: Optimal beam line (a) and distribution of the  $EU$  (b) for the estimation of the maximum density according to the model function (4.4).

For the estimation of the maximum density, e.g. for density control in a fusion plasma device, the expected utility shows a behaviour similar to the case of the quadratic density distribution (see section 4.2.1): The maximum of the  $EU$  is given for a beam line crossing the plasma through its centre on the longest path possible (figure 4.7). However, the maximum of the  $EU$  is not as distinct as in the case of the quadratic distribution. This can be explained by the plateau in the density distribution described by the new model function (figure 4.6). Moving the beam slightly out of the plasma centre leads to only small changes in the density and therefore in the signal-to-noise ratio. A broad maximum of the  $EU$  distribution characterises a robust design: Small changes of the optimal beam

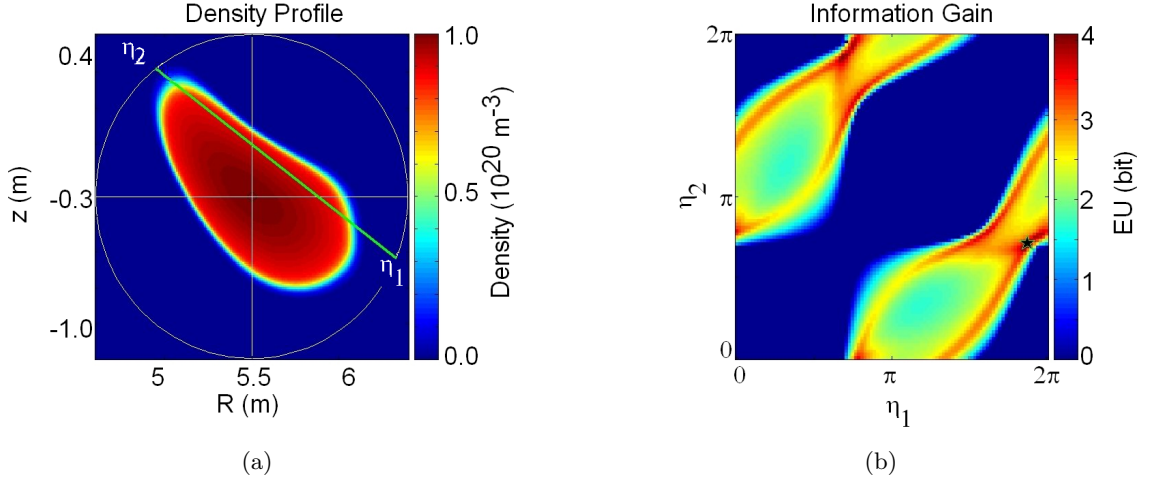


Figure 4.8: Optimal design for measuring the position of the steepest gradient: Density distribution with the optimal beam line (a), the star symbol at the  $EU$  distribution (b) corresponds to this line of sight.

line have only a small effect on the expected information gain of the measurement.

The expected utility for the best design to estimate the position of the steepest gradient has a different shape: The maximum is given for a beam line at the edge of the plasma, however, again a line of sight crossing the plasma on a long path is preferred (figure 4.8). This result is reasonable: An effect occurring at the plasma edge is best measured with a beam line at the edge. It should be mentioned that the variance of the parameter of interest,  $\theta_2$ , ranges over a quite large interval as can be seen in figure 4.6 (b). For a smaller range (e.g. as described in section 5.1.2) the design is slightly different: The optimal beam moves outwards to the edge region of the plasma where the effect is located.

For the steepness of the density decay at the plasma edge described by  $\theta_3$ , the  $EU$  shows a strongly peaked maximum (figure 4.9). This indicates a very sensitive design: Changes of the beam line have a strong effect on the expected information gain. Again,

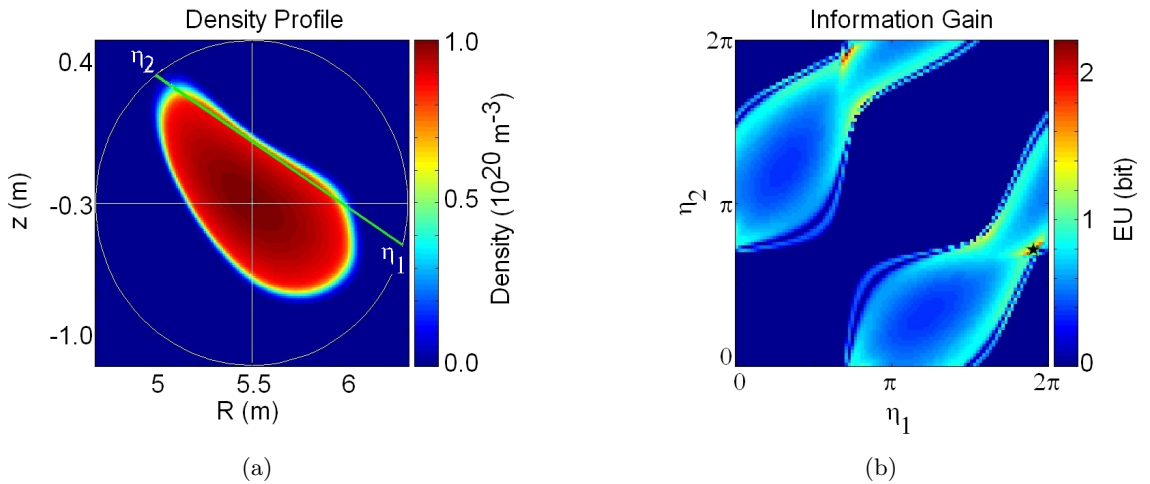


Figure 4.9: Plasma cross section with the optimal line of sight for the estimation of the steepness of the density distribution (a), the corresponding maximum of the  $EU$  distribution is marked with a star symbol (b).

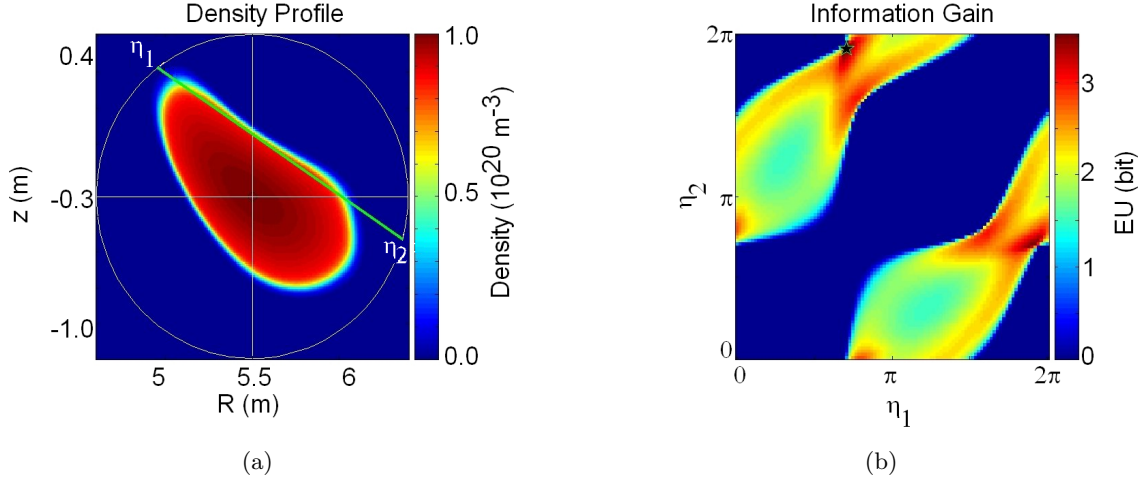


Figure 4.10: Optimal beam line configuration for the measurement of the bulge of the density distribution (a), the corresponding maximum of the  $EU$  is marked with a star symbol (b).

one finds for the optimal line of sight a beam at the plasma edge. The  $EU$  distribution is also characterised by a ribbon-like structure for beam lines at the very edge of the plasma, indicating that the steepness can also be measured there. This result can be explained by comparison with the density plot in figure 4.6 (c): Changing the steepness of the density distribution leads to an effect in the region of the plasma edge and at the brink of the profile at higher densities. This results in a ribbon-like structure at the plasma edge and a maximum of the Expected Utility near the density plateau.

The last parameter to be used as an optimisation goal is the bulge of the density distribution ( $\theta_4$ , see 4.6 (d)). As in the examples before, the most distinct effect in the density distribution can be found at the brink of the plasma profile, therefore the maximum of the  $EU$  corresponds again with a beam line at the edge (figure 4.10). In comparison to the third case, the maximum here is much broader, the design is therefore not as sensitive against small changes of the beam line.

For the design with respect to  $\theta_3$  and  $\theta_4$  one has to be aware that there is a dependency of the optimisation outcome on the value of  $\theta_2$ : The position of the steepest gradient determines the localisation of the density effects of  $\theta_3$  and  $\theta_4$  and therefore the corresponding design. From this point of view, the design examples presented here can be seen as case studies, for a complete design all parameters need to be varied simultaneously.

However, it turns out that the positioning of beam lines at the edge region of the plasma is very beneficial. Unfortunately, the preferred beam line positions are not accessible at W7-X due to the port system (see section 3.2.2, figure 3.2). In figure 4.11 the allowed regions are shown in the  $EU$  plot. The effect of the boundary conditions on the diagnostic design will be pointed out in the next chapter.

#### 4.2.4 Impact of the error statistics

A description of the measurement error at the interferometry diagnostic was given in section 3.2.3, however, the influence of the error statistic on the diagnostic design shall be demonstrated in more detail. As an academic example, the density distribution from equation (4.4) is used, parameter of interest is the maximum density,  $\theta_1$  [49].

Figure 4.12 shows the expected utility for three different error statistics: In the first



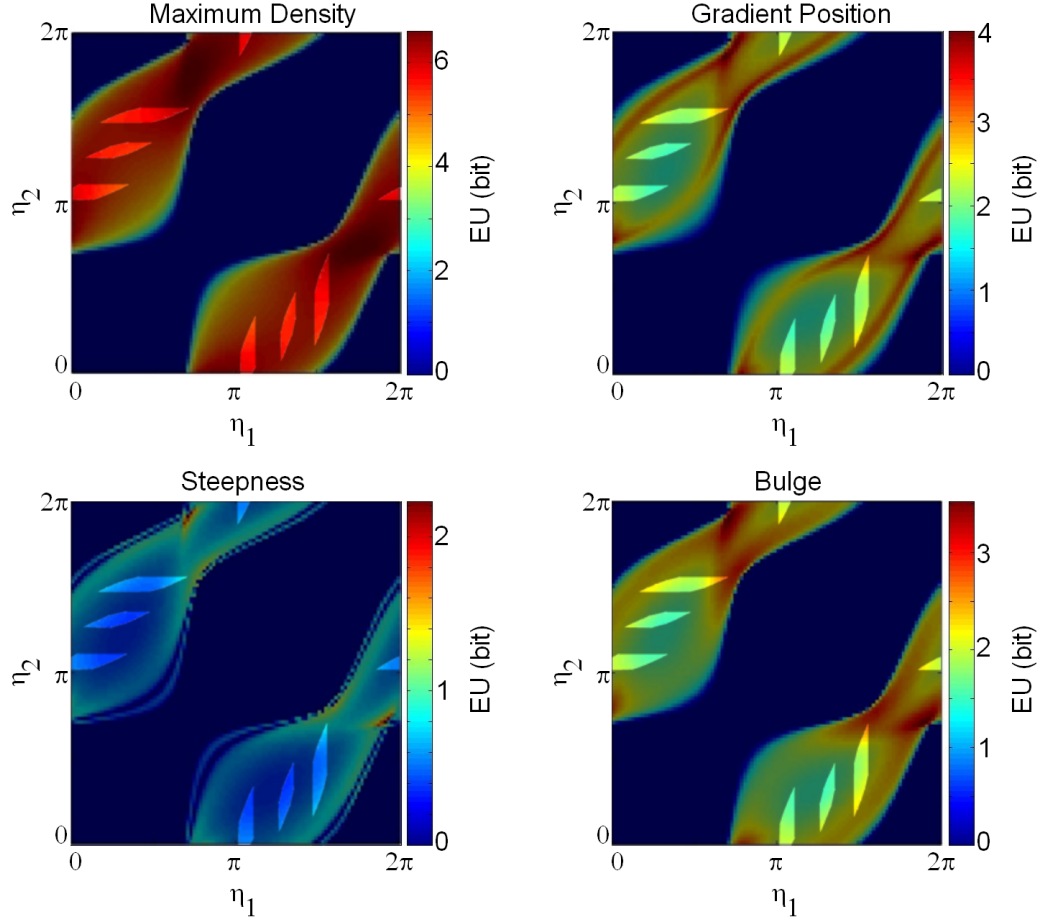


Figure 4.11: The expected utility for the parameters  $\theta_1 - \theta_4$ : The shaded areas are not accessible due to the port system of Wendelstein 7-X.

case, the standard deviation  $\sigma$  was chosen to be independent from the data value, describing a constant background noise. Second, the standard deviation was proportional to the square root of the data value, which is typically the case in counting experiments. As a last example, an error proportional to the data value, i.e. a "relative" error, was assumed.

The  $EU$  for a constant background noise shows a distinct maximum for a beam line crossing the plasma through its centre on the longest possible way. The same maximum appears for the case where the standard deviation is proportional to the square root of the data, but the maximum is not as distinct as in the first case. Finally, for the error proportional to the data, the  $EU$  shows no maximum at all, but is equally distributed.

This result can be explained with the dependence of the  $EU$  on the signal-to-noise ratio: The SNR is proportional to the data in case of the constant background noise - the beam line with the largest data value - a line of sight crossing the plasma centre on the longest way possible - is preferred. For the second case, the SNR is only proportional to the square root of the data, resulting again in a maximum for the  $EU$  at large data values, but a more smooth distribution - changes in the data value have a smaller impact on the SNR and therefore on the  $EU$  itself. In case of a measurement error proportional to the data the SNR remains constant for every line of sight: For a known density distribution, the maximum density  $\theta_{max}$  can be measured everywhere. Here, the  $EU$  shows a constant distribution.

Although the examples discussed here are limiting cases, they point out the tremendous



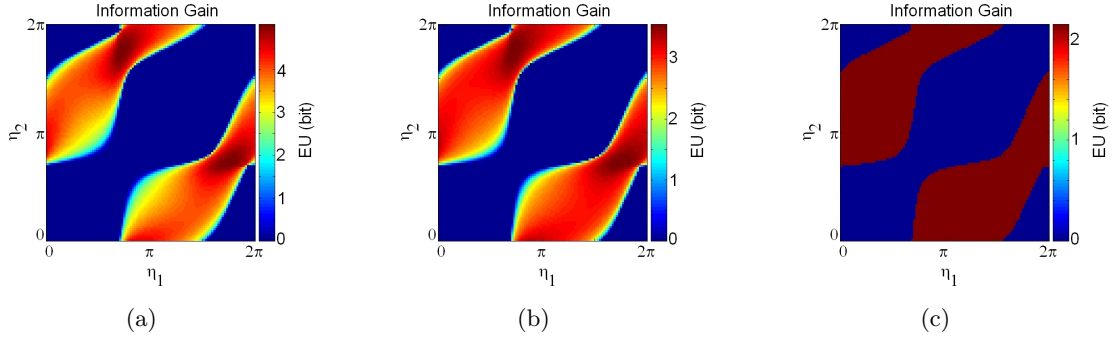


Figure 4.12: Expected Utility for the estimation of the maximum density assuming a measurement error independent from the data (a), proportional to the square root of the data (b) and proportional to the data (c) [49].

influence of the error statistic. A detailed error analysis for the diagnostic is therefore of particular importance even in the design phase.

### 4.3 Data analysis for interferometric measurements

The Expected Utility as introduced in chapter 2 is an absolute measure for the average information gain of a future measurement. The averaging is done over the expected data space, every possible datum corresponding to the parameter range of interest is taken into account. This leads to an important characteristic of the *EU*: It does not give a hint about the outcome of a single future experiment.

Even for an optimal experimental design extreme measurement errors may occur due to the error statistics. In average, the optimised experiment would produce better measurements than a non-optimal experiment, but no statement can be made about the quality of a single data point from a future measurement. This characteristic makes it finally difficult to prove, by a measurement, that the optimal design was found by the maximisation of the *EU*.

However, in a few cases the influence of the *EU* can be illustrated. In this section, interferometry measurements from W7-AS are used to illustrate the consequences of high and low *EU* values. In the first part an introduction into the inversion procedure necessary for the analysis of line-integrated measurements is given, in the second part the effect of the *EU* on a measurement is examined.

#### 4.3.1 Abel integral and inversion problem

To extract a spatially resolved density distribution from the line integrated measurement of the interferometer, an inversion procedure has to be applied. This is only possible by using a multi-channel interferometer.

If the probing beam of an interferometer crosses a circular plasma with radius  $R$  at the distance  $r_0$  (closest point to the plasma centre, see figure 4.13), the phase shift is given with [35]

$$\phi(r_0) = \frac{2\pi}{\lambda n_C} \int_r^{r_0} \frac{n(r)}{(r^2 - r_0^2)^{1/2}} r \, dr. \quad (4.5)$$

This integral can be inverted, and one obtains an expression for the electron density

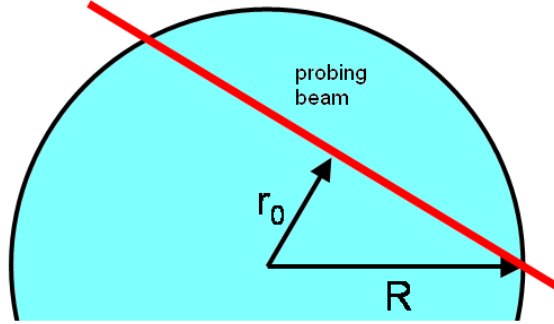


Figure 4.13: Abel inversion: The inversion formula (4.6) is given for a circular shaped plasma (light blue) with radius  $R$ ,  $r_0$  is the smallest distance of the probing beam to the plasma centre.

distribution:

$$n(r) = -\frac{\lambda n_c}{\pi^2} \int_r^{r_0} \frac{d\phi(r_0)}{dr_0} \frac{dr_0}{(r_0^2 - r^2)^{1/2}}. \quad (4.6)$$

Integrals of the kind  $\int f(x, y(x)) dx$  are called *Abel integrals*, the inversion procedure is often referred to as Abel inversion. It is solvable analytically for parabolic density distributions.

For many practical applications this is not an appropriate assumption: First, non-parabolic density distributions  $n(r)$  occur, and second, like in W7-X, the plasma is not of circular shape. A general approach for inversion is therefore introduced next.

Plasma density distributions are often expressed in an appropriate coordinate system to overcome the problem of a non-circular cross section of the plasma. One approach is to transform the volume inside a magnetic flux surfaces into a cylinder with the same volume. The radius of this theoretical cylinder is called the *effective radius* of the flux surface. This conversion is called *mapping*, and it has to be provided for an interferometry measurement. The mapping of the magnetic flux surfaces can be calculated or measured with adequate diagnostics (e.g. soft X-ray tomography [21, 22]). The final outcome is then a density distribution in effective coordinates:  $n(r_{eff})$ .

For analysis purposes, the density distribution is discretised along  $r_{eff}$ , and the density is assumed to be constant in every segment. The number of segments is in the order of the number of beam lines, but not necessarily identical: If additional information is provided by the regularisation functional (see below), the number of segments can be larger than the number of beam lines. In real space, the segments correspond to onion-like shells given by the magnetic flux surfaces. They are often chosen as equidistant, although this is not required.

Two possibilities exist to extract the (discretised) density distribution from the measurement provided by a multi-channel interferometer: If a *backward calculation* is applied, the density of the outermost shell is estimated first by the measurement using the outermost interferometer channel. The density of the second shell is then given by the density of the first shell, combined with the measurement of the interferometer channel next to the outermost one. The densities of the inner shells are estimated iteratively in the same way. The disadvantage of this method is its error propagation: The error of the density in the innermost shell is large, since it accumulates the errors of all other shells. This result is due to the analysis procedure (backward calculation) and is questionable, because the measurement error of the interferometer channel at the plasma centre is the same as for the other channels. Furthermore, the error does not decrease below the accumulated error

of all other shells, no matter how many additional beams are applied through the plasma centre.

The second method establishes *forward calculation* and can be regarded as a virtual diagnostic: The data  $d$  is measured by the diagnostic and simultaneously the forward function  $f(\boldsymbol{\theta})$  is calculated in relation to the parameters of interest  $\boldsymbol{\theta}$  (here: the densities in the different shells,  $\theta_i = n_e(r_{eff}^i)$ ). Then, both data and forward function are compared:

$$\chi^2 = \sum_i^{N_d} \left( \frac{d_i - f_i(\boldsymbol{\theta})}{\sigma_i} \right)^2 \quad (4.7)$$

Here,  $N_d$  beam lines are applied,  $\sigma_i$  is the error of the respective data point. To estimate the density distribution,  $\chi^2$  has to be minimised with respect to  $\boldsymbol{\theta}$ .

For piecewise constant density distributions, equation (3.5) becomes the sum

$$\phi_i(\boldsymbol{\theta}) = \frac{\lambda e^2}{4\pi\epsilon_0 m_e c^2} \sum_j A_{ij} \cdot \theta_j ; \quad (4.8)$$

where  $A_{ij}$  is a matrix element of the *projection matrix*  $\mathbf{A}$  containing the length of beam  $i$  in shell  $j$ , and taking into account the mapping of the flux surfaces. Setting  $f_i(\boldsymbol{\theta}) = \phi_i(\boldsymbol{\theta})$ , equation (4.7) becomes a system of linear equation to be minimised.

In case of noisy data the minimisation of (4.7) is an ill-posed problem, which requires an additional boundary condition to find an unique solution. Some possible methods regarding an interferometry system are presented by Koponen and Dumbrajs [23] for the multichannel interferometer at W7-AS.

The boundary condition is implemented by the *regularisation functional*  $R$ . In detail, the term

$$F = \frac{1}{2} \chi^2 - \alpha R \quad (4.9)$$

has to be minimised instead of  $\chi^2$  only.  $\alpha$  is a Langrage parameter, which weights the regularisation functional so that is is neither over- nor underestimated.

Different regularisation functionals are possible and are strongly related to the inversion problem to be solved. The approach to be chosen has to reflect the physical background knowledge about the parameters of interest. For instance, information about the functional behaviour of  $\boldsymbol{\theta}$  like smoothness or curvature can be encoded in  $R$ . Two approaches are used in the fusion research (see, e.g., [23, 21, 22]) and shall be introduced briefly: the method of Maximum Entropy (MaxEnt) and the Minimum Fisher approach.

The *maximum entropy method* is based on the Shannon entropy measure and prefers the most uninformative solution. This avoids "virtual" information, i.e. effects of the noise are not interpreted as signal by mistake. The regularisation functional  $R$  is given by the information entropy

$$R = \sum_i \left[ \theta_i - m_i - \theta_i \ln \left( \frac{\theta_i}{m_i} \right) \right]. \quad (4.10)$$

Here,  $\mathbf{m}$  describes the model for  $\boldsymbol{\theta}$  in absence of any data, in case of no further information it is a flat distribution  $\mathbf{m} = const$  [21].

As a disadvantage, physical knowledge is completely ignored by the MaxEnt method. For instance, in case of the plasma density one would always assume a smooth distribution without large steps between the different shells. But using MaxEnt, the shells are assumed as independent of each other, so that, in principle, large steps could occur.

A second method is the principle of *Minimum Fisher regularisation* [22, 23]. Here, the regularisation functional is given by the Fisher information

$$R = I_F = \int \frac{[\boldsymbol{\theta}'(r_{eff})]^2}{\boldsymbol{\theta}(r_{eff})} dr_{eff} \quad (4.11)$$

where  $'$  denotes the derivation with respect to  $r_{eff}$ . It has been shown that a minimum of (4.9) using this regularisation functional provides the smoothest result possible [50].

The principle of Minimum Fisher regularisation is to cancel out noise effects of the background, where the values of  $\boldsymbol{\theta}$  are small: Rapid changes of the distribution function lead to large values of its derivative  $\boldsymbol{\theta}'$ . The term  $\frac{[\boldsymbol{\theta}'(r_{eff})]^2}{\boldsymbol{\theta}(r_{eff})}$  becomes large for small values of  $\boldsymbol{\theta}$  and vice versa, in other words, the division by  $\boldsymbol{\theta}$  assures that regions of low  $\boldsymbol{\theta}$  and high derivatives  $\boldsymbol{\theta}'$  considerably contribute to the regularisation functional  $R$ . Minimisation of equation (4.9) leads therefore to a smoothing of these regions of  $\boldsymbol{\theta}$ . This effect was illustrated for data analysis of interferometric measurements by Koponen and Dumbrajs [23] and for X-ray tomography by Anton et al. [22].

Beside the methods of maximum entropy and minimum Fisher, other regularisation functionals are also possible and used for inversion: minimisation of the norm ( $R = |\boldsymbol{\theta}|^2$ ), minimisation of the gradient (first order regularisation,  $R = |\boldsymbol{\theta}'|^2$ ) or curvature (second order regularisation,  $R = |\boldsymbol{\theta}''|^2$ ) of the density distribution [22]. All these methods have different meanings as a boundary condition for (4.9), therefore their application strongly depends on the physical problem. However, these boundary conditions seem to be inappropriate for the reconstruction of a density distribution.

### 4.3.2 Analysis of interferometer data from W7-AS

As examples for the analysis of interferometric data two shots of the W7-AS experiment shall be analysed (shot numbers # 35502 and # 35504). The effect of the regularisation method and the influence of the measurement error will be illustrated.

The multichannel interferometer at W7-AS was a microwave system with wavelengths around  $\lambda = 1.87 \mu m$ . The channels were arranged nearly horizontal at a toroidal angle of approximately  $30^\circ$  (see figure 4.14). In total, nine channels were attached to the experiment, for the shots analysed here only 7 channels were available. A detailed description of the diagnostic is given by Geist et al. [41].

As a first example, discharge # 35502 is analysed by using the aforementioned regularisation methods, Maximum Entropy and Minimum Fisher. The time traces of the interferometer signals are displayed in figure 4.15, the analysis was done for  $t = 0.5 s$ .

For the analysis, the data traces were corrected with the offset from the time interval after the discharge ( $t > 1 s$ ). Because the line integrated density  $\int n dl$  was originally estimated

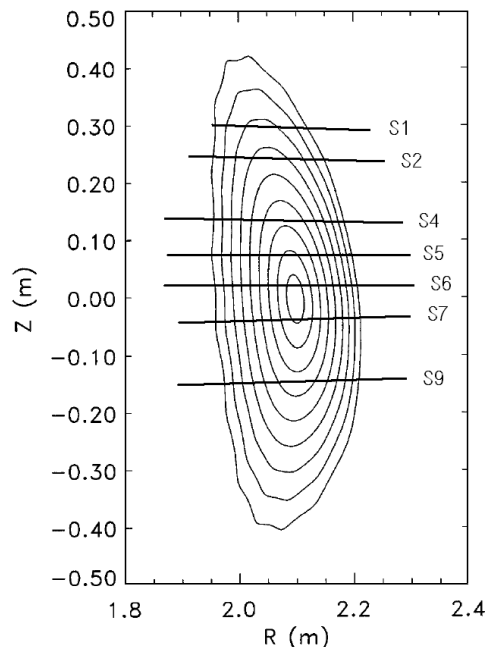


Figure 4.14: Seven chords of the multichannel interferometer at W7-AS with a contour plot of the magnetic flux surfaces (from [23]).

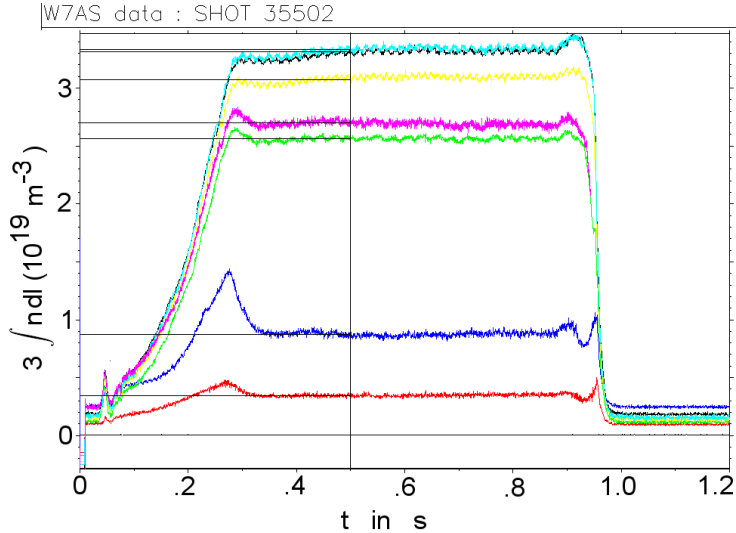


Figure 4.15: Line integrated density measurement from shot # 35502 of the W7-AS experiment (a), the time traces are colour coded according to the respective channel: S1 - red, S2 - dark blue, S4 - green, S5 - yellow, S6 - black, S7 - light blue and S9 - magenta.

for an average wavelength  $\lambda = 1.87 \text{ mm}$ , a correction factor  $c = \lambda_i/1.87$  was used for the data of every channel. Here,  $\lambda_i$  is the wavelength of the  $i$ -th channel in  $\text{mm}$  as listed in table 4.1.

The errors bars in the density plot were computed according to [21]: For the error estimation the covariance matrix

$$H_{ij} = \frac{\alpha}{\theta_i} \delta_{ij} + \sum_l A_{li} \frac{1}{\sigma_l^2} A_{lj} \quad (4.12)$$

was calculated. Here,  $\sigma_l$  is the error of channel  $l$ ,  $\theta_i$  the density in shell  $i$  and  $\mathbf{A}$  the projection matrix from equation (4.8). For this study, only the statistical error given by the noise of the data traces was taken into account. The error of the density is now given with

$$(\Delta\theta_i)^2 = (H^{-1})_{ii}. \quad (4.13)$$

It should be mentioned here, that this error is a function of the Lagrange parameter  $\alpha$  (equation (4.12)), too.

Figure 4.16 (a) shows the result of the data analysis using the Maximum Entropy regularisation, comparing the outcome for the density distribution with and without chord S9. This differentiation was done because of the larger variation in the S9 data time trace: The noise for the other beams was determined with  $\Delta(\int n dl) = 0.01 \cdot 10^{19} \text{ m}^{-2}$ , whereas for S9  $\Delta_{S9}(\int n dl) = 0.03 \cdot 10^{19} \text{ m}^{-2}$  was assumed. In result, the solution including S9 shows a poor performance, leading to the assumption that the S9 signal can be regarded as an outlier. Therefore, the S9 chord was excluded from the data set for the following calculations.

beam:	S1	S2	S3	S4	S5	S6	S7	S9
$\lambda \text{ (mm):}$	1.873	1.872	1.87	1.869	1.867	1.865	1.863	1.858

Table 4.1: Wavelengths of the interferometer chords at W7-AS (from [41]).

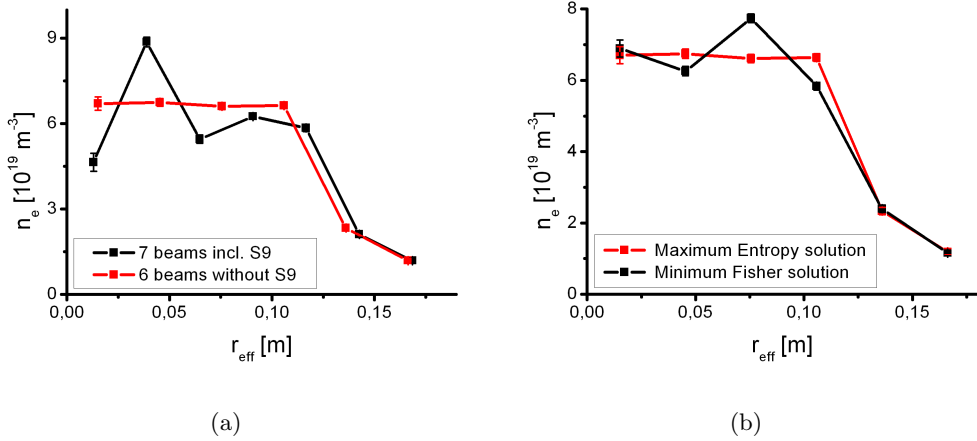


Figure 4.16: Reconstruction of the density profile of W7-AS shot # 35502: Comparison of the solution with and without beam S9 (a), and of different regularisation methods (b): Maximum Entropy (black) and Minimum Fisher inversion (red).

Considering six beams (S1, S2, S4 - S7), the comparison between MaxEnt and Minimum Fisher inversion is done in figure Figure 4.16 (b). The density distribution reconstructed with Minimum Fisher is characterised by smaller slope compared to the Maximum Entropy result: The distribution calculated by MaxEnt shows a density plateau in the plasma centre and a steep gradient at the plasma edge, whereas the edge gradient from Minimum Fisher regularisation is somewhat more flat. This may imply that introducing this smoothness would be misleading in case of density distributions with steep gradients at the plasma edge.

It is important to mention that the inversion result strongly depends on the measurement error. To illustrate this, another discharge from W7-AS (# 35504) was analysed. During this discharge, 7 channels of the interferometer were available (see figure 4.17, (a)), the data from channel S7 was slightly more noisy than the others. The error of the other beams was given with  $\Delta(\int n \, dl) = 0.01 \cdot 10^{19} \text{ m}^{-3}$ , for S7  $0.03 \cdot 10^{19} \text{ m}^{-3}$ . The error of S7 was varied, figure 4.17 (b) displays the results for MaxEnt inversion.

Even the small difference in the error of chord S7 results in a significant impact to the density distribution: Assuming equal errors for all beams leads to scattered distribution in the density plateau region, whereas the distribution becomes flat if a higher error for S7 is taken into account, which seems to be more reasonable. The strong impact of the error variation in this example illustrates the importance of an accurate estimation of the measurement errors.

## 4.4 Validation of the $EU$

In this section, the impact of different beam line configurations shall be compared with their respective Expected Utility, which is calculated separately. With this method, the correlation between measurement outcome and the value of the  $EU$  can be demonstrated.

Basically, the meaning of the Expected Utility can hardly be interpreted by experiments quantitatively. The reason for that is mainly the integration over the expected data space which is done for the  $EU$  calculation (see equation (2.38)). This would mean that for a reconstruction of the  $EU$  in principle every possible datum is needed to be measured, resulting in the necessity for an infinite, or at least very large number of measurements.

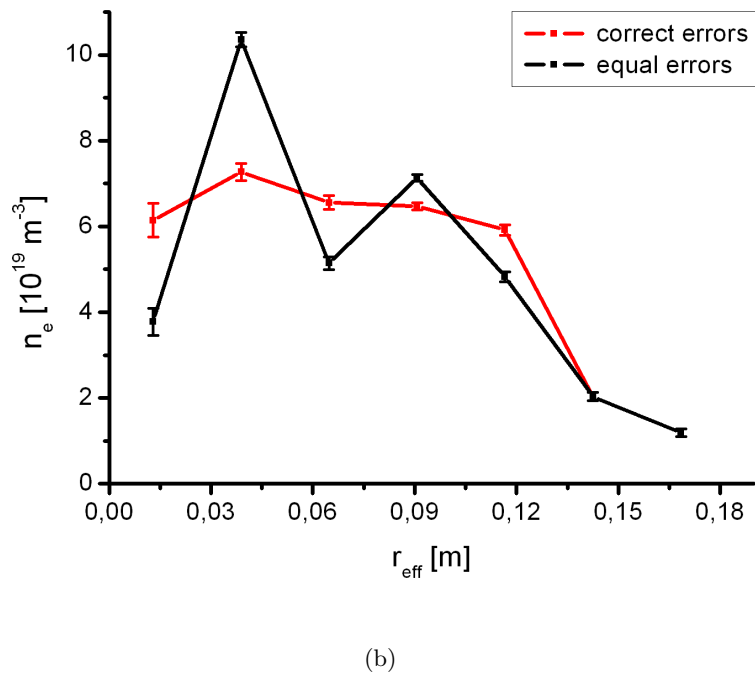
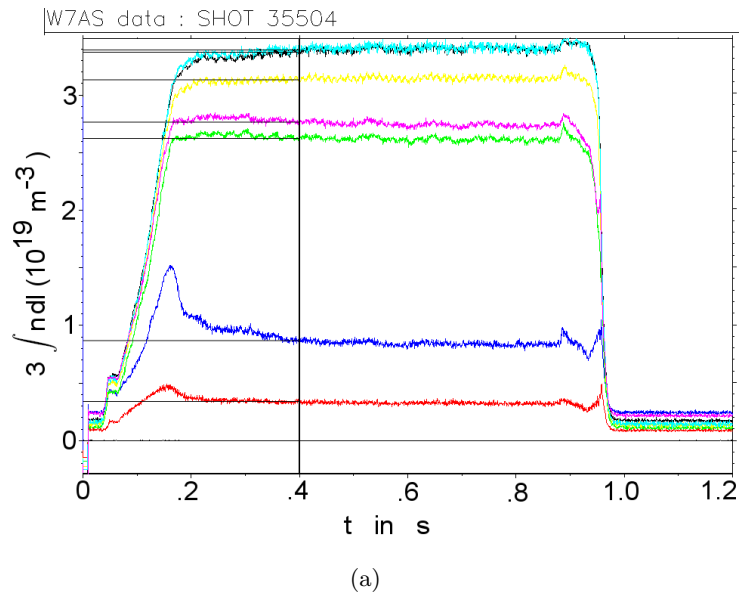


Figure 4.17: Influence of the error in MaxEnt inversion: Data traces of shot # 35504 of the W7-AS experiment ((a), see figure 4.15 for the colour codes). The error of channel S7 (light blue) is slightly more noisy than the others. In (b) the inversion result is given for assuming equal errors (black) and a slightly larger error for S7 (red).

The integration over the parameter range of interest would lead to similar difficulties.

Nevertheless, a higher *EU* should correspond to a better estimation of the physical parameters of interest. This can be illustrated also with single data sets in a qualitative way. In the example chosen here the reconstruction of a density distribution using different numbers and different combinations of interferometer beam lines is compared with the density distribution calculated from the complete interferometer data set combined with data from Thomson scattering.

#### 4.4.1 Data set and parametric fit

Using a measurement for the validation of the Expected Utility has the disadvantage that the "real" physical situation, in this case the distribution of the plasma density, is unknown. But by combining different data sets one may get a good estimate of the distribution which can be used as a "best estimate" for comparison in the following.

For this, shot # 35502 of the W7-AS experiment was analysed in the following way: Six beam lines of the interferometer (S1, S2, S4 - S7) were used to calculate the density distribution by MaxEnt regularisation. As discussed in the previous section, the S9 chord was removed from the data set. In addition, the results from the Thomson scattering diagnostic as shown in [23] were taken into account. Figure 4.18 shows both data sets and a parametric fit. The fit function applied here is given in equation (4.4) and is commonly used for the description of density distributions in W7-AS [48].

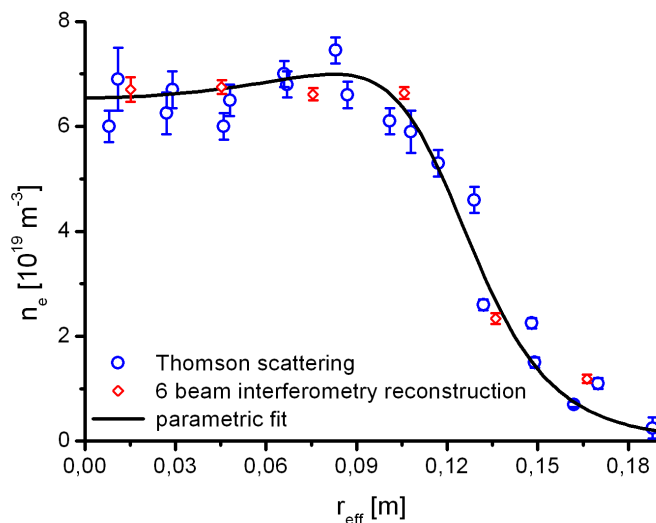


Figure 4.18: W7-AS shot # 35502: Reconstructed density distribution: Interferometry (red) and Thomson scattering data (blue) [23], combined with a parametric fit (black).

The density profile is now reconstructed by using only a subset of beams from the interferometer: A four-beam interferometer, consisting of the innermost chords S4 - S7 (see figure 4.14), and two five-beam interferometers, consisting of the four-beam configuration extended by the chords S1 and S2, respectively.

Figure 4.19 shows the results in comparison to the parametric fit from the complete data set of interferometer and Thomson scattering. The solution using only four beams gives the poorest result, which is not surprising: Without a beam through the plasma edge, a good reconstruction of the plasma profile becomes difficult.



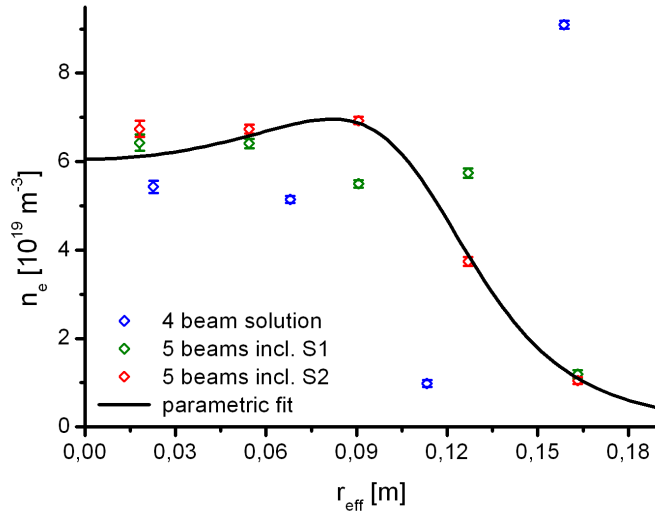


Figure 4.19: Reconstruction of the density distribution with different beam line configurations: Four beams (S4 - S7, blue), five chords including S1 (green) and including S2 (red), respectively. The parametric fit from Figure 4.18 (black) is given for comparison.

Adding chord S1, which is the outermost beam, to the set-up leads to a significant improvement of the result. At the supporting points in the plasma centre and at the very edge the reconstructed profile shows a good agreement with the parametric fit. Differences are found in the regions of the strongest density decay, again a region not covered by one of the beam lines.

Finally, the four-beam set-up is extended by chord S2, which is located slightly closer to the plasma centre than S1. The density distribution calculated from this data set shows the best agreement with the parametric fit function, especially in the region of the density decay.

#### 4.4.2 Comparison with Expected Utility

The result of the different reconstructions shall now be compared with the corresponding Expected Utility, which is calculated independently from the data analysis. For this, a parameter range of interest has to be defined (see section 2.3). Here, the physical interest is expressed by the parameters of the fit function (4.4). The variations of the different parameters are displayed in figure 4.20.

The  $EU$  is now calculated for a simple design problem: Given the beams S4 - S7, what would be the best position for a fifth interferometer chord to measure these variations in the density distribution? As a boundary condition, the new beam line was assumed to be horizontal, only the  $z$  coordinate was varied.

In Figure 4.21 the distribution of the  $EU$  with respect to the  $z$  coordinate is displayed. The plot shows a nearly symmetric shape due to the fact that the plasma is approximately symmetric at the interferometer plane (figure 4.14), and the given chords S4 - S7 are also nearly symmetric to the plasma centre.

The flat parts of the distribution for large positive and negative values of  $z$  correspond to beam lines which do not cross the plasma, in other words, these regions show the  $EU$  for only four chords. Here, the  $EU$  is given with  $EU = (14.20 \pm 0.07) \text{ bit}$ .

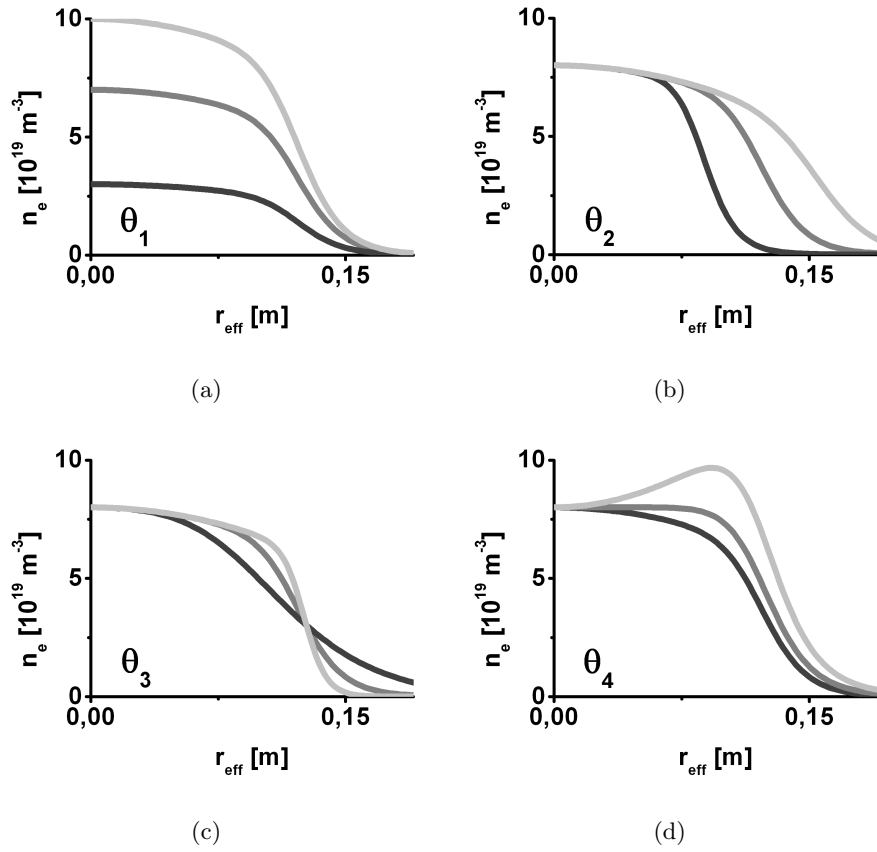


Figure 4.20: Variation of the parameter range of interest according to equation 4.4: Maximum density ( $\theta_1 : 0.3 \leq \theta_1 \leq 2.0$ , (a)), position of the edge decay ( $\theta_2 : 0.5 \leq \theta_2 \leq 0.9$ , (b)), steepness ( $\theta_3 : 0.01 \leq \theta_3 \leq 0.2$ , (c)) and density behaviour at the plasma centre ( $\theta_4 : -1.0 \leq \theta_4 \leq 2.0$ , (d)).

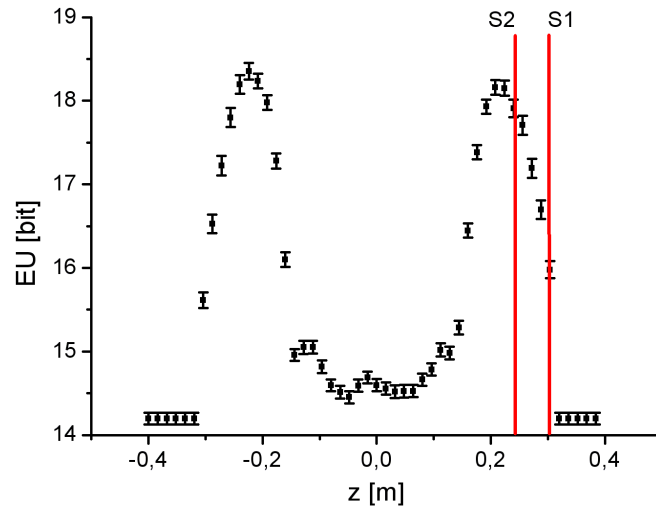


Figure 4.21: Expected Utility for a fifth beam line shifted on the  $z$  axis, given the chords S4 - S7. The positions of the beams S1 and S2 are indicated.

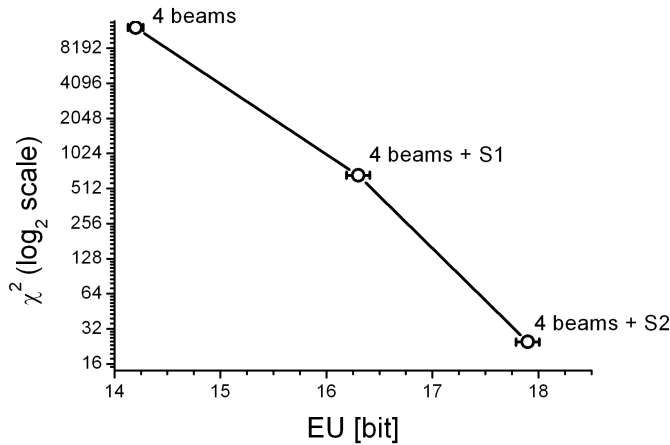


Figure 4.22:  $\chi^2$  of the density reconstructions and the parametric fit, plotted with respect to the Expected Utility  $EU$ . The  $\chi^2$  is displayed on a  $\log_2$  scale, resulting in an approximately linear relation.

A small maximum of the  $EU$  occurs for beam lines through the plasma centre ( $z = 0$ ). This corresponds to beams with a high signal-to-noise ratio, similar to the findings of section 4.2.3 and comparable to the results for W7-X as will be explained in detail in chapter 5. However, this extremum is only marginal compared to the maxima at  $z \approx \pm 0.23$ . In these regions the density distribution as parametrised in figure 4.20 shows its strongest changes, mainly the effects of the density decay are localised here. In addition, no given beam covers this range of  $z$  directly, the given beams S4-S7 are located near the plasma centre (figure 4.14). Therefore, a fifth chord traversing the plasma at these  $z$  coordinates would be very "valuable".

The positions of the two available beam lines S1 and S2 are also indicated in figure 4.21. Their position is only an approximation, because the chords at W7-AS were not exactly horizontal. As can be seen in the figure, the  $EU$  for S2 is much larger as for S1. Calculated directly one obtains

$$\begin{aligned} EU(S1) &= (16.3 \pm 0.1) \text{ bit} \\ EU(S2) &= (17.9 \pm 0.1) \text{ bit}. \end{aligned}$$

The  $EU$  of S1 is smaller, because the effect of the density distribution at the coordinates of S1 is less distinct as in the region where S2 is located.

Finally, the outcome of the density reconstruction and the calculation of the  $EU$  can be compared. For this, the quadratic deviation  $\chi^2$  between the inversion results in figure 4.19 and the parametric fit was calculated for all three designs:

$$\begin{aligned} \chi^2(4 \text{ beams}) &= 12352.1 \\ \chi^2(\text{incl. S1}) &= 665.1 \\ \chi^2(\text{incl. S2}) &= 24.6 \end{aligned}$$

The  $\chi^2$  values are plotted with respect to the respective  $EU$  in figure 4.22.

The plot shows a monotonous relation between Expected Utility and the  $\chi^2$ . Moreover, the relation is linear on a  $\log_2$  scale: In this example, an increase of the information gain by approximately 3 *bit* leads to a reduction of  $\chi^2$  by a factor two.

## Discussion

As pointed out at the beginning of this section, this result can only be interpreted qualitatively. It depends on the specific problem and is therefore not valid in general. The result can be seen as a proof of a monotonous relationship between the Expected Utility and the accuracy of the reconstruction of a density distribution, since both quantities,  $EU$  and  $\chi^2$ , have been calculated independently. However, the specific relation between  $EU$  and  $\chi^2$  found here can not be translated directly to other examples. Especially the meaning of a gain of 1 *bit* may be different, if the physical question or the parameter range of interest changes.

Nevertheless, the working principle of the  $EU$  is illustrated in this example: A larger  $EU$  leads to a better reconstruction of the parameters of interest.



## Chapter 5

# Design of a Multi-Channel Interferometer

This chapter presents the design results for a multi-channel interferometer at Wendelstein 7-X. The diagnostic is optimised according to three different physical problems, which are introduced and parametrised first. The discussion of the design results follows in the second part of the chapter.

### 5.1 Physical problems of interest

The physical and technological challenges and issues for the development of a stellarator reactor are manifold and widespread. It is nearly impossible to cover the whole area, therefore the focus of this work lies on three selected topics: analysis of the Core Electron Root Confinement (CERC), measurement of density effects of different high plasma confinement scenarios and the observation of the density changes during the injection of hydrogen pellets. This section will motivate the interest in these problems and introduce the physical background.

The diagnostic to be designed is the infrared multi-channel interferometer at W7-X (see chapter 3), which is a diagnostic for the estimation of density profiles from line-integrated measurements. Therefore, the effects discussed here are analysed with respect to changes in the density distribution.

#### 5.1.1 Core Electron Root Confinement (CERC)

The Core Electron Root Confinement (CERC, see [51] and [52] for an overview) is a plasma confinement scenario characterised by a strong positive radial electric field  $E_r$ . It is a stellarator-specific effect which does not appear in tokamaks. CERC occurs in the regime of long mean free paths and confines the electrons, which would be lost due to  $\nabla\mathbf{B}$ -drifts otherwise. Another characteristic is the highly peaked electron temperature profile (see figure 5.2 right).

The electron root follows from a possible solution of the ambipolarity condition

$$\sum_{\alpha} Z_{\alpha} e \Gamma_{\alpha} = 0. \quad (5.1)$$

Here,  $Z_{\alpha}$  denotes the species of the particles, it is  $-1$  for electrons and  $1$  for hydrogen ions.  $\Gamma_{\alpha}$  are the fluxes of species  $\alpha$ . So, equation 5.1 describes the stationary state of particle balance:

$$\Gamma_e = \Gamma_i = \frac{1}{v'} \int v' \text{Sp}(r) dr, \quad (5.2)$$

where  $\Gamma_e$  and  $\Gamma_i$  are the electron and the ion flux,  $\text{Sp}(r)$  is the particle source term and  $v'$  is the derivation of the particle velocity with respect to  $r$ .

The particle flux is given by [51]

$$\Gamma = -nD \left[ \frac{n'}{n} - \frac{ZE_r}{T} + \delta_{12} \frac{T'}{T} \right], \quad (5.3)$$

where  $n$  is the density and  $D$  the diffusion coefficient, which depends on the collision frequency  $\nu$ . The coefficient  $\delta_{12}$  is of order one [48], for strongly peaked temperature profiles the last term is large. To reduce the particle flux, high values for the radial electric field  $E_r$  are therefore favourable.

The three-dimensional transport regime in stellarators at low collisionalities is the so-called  $1/\nu$  - regime, where  $D \propto 1/\nu$ . It arises at small  $E_r$  and is suppressed for large positive radial electric fields [52]. This transport regime is unfavourable due to its dependency on the electron temperature [53]:

$$D_{1/\nu} \propto T_e^{7/2}. \quad (5.4)$$

This proportionality finally leads to a significant increase in diffusion even for a small increase in the temperature. In consequence, it may limit a fusion experiment to lower  $T_e$  regions. Suppressing the  $1/\nu$  - regime is therefore one optimisation goal for stellarators [14].

In the case where the electrons are in the  $1/\nu$  - regime with  $T_e \approx T_i$ , only one solution for ambipolarity (5.1) occurs: In the long mean free path regime the electrons and ions are helically trapped. Nevertheless, the ions are going to leave the plasma more quickly than the electrons (the viscosity for ions is by a factor  $(m_i/m_e)^{1/2}$  larger than for electrons [53], the same dependency is found for the ratio  $D_i/D_e$  at  $E_r = 0$ ), which leads to a negative radial electric field generated by the still confined electrons. This field finally holds back the ions, the solution is called *ion root*.

Other transport regimes occur at higher values of the radial electric field, namely the  $\sqrt{\nu}$  - and the  $\nu$  - regime, where  $D \propto \sqrt{\nu}$  and  $D \propto \nu$ , respectively. These regimes show more favourable dependencies on  $T_e$ :

$$D_{\sqrt{\nu}} \propto T_e^{5/4} \quad (5.5)$$

$$D_{\nu} \propto T_e^{1/2}. \quad (5.6)$$

For these transport regimes, more than one solution for the ambipolarity equation may occur. Normally, electrons and ions appear in different regimes: If  $T_e \gg T_i$  the electron flux  $\Gamma_e$  rapidly grows, because the transport in the  $1/\nu$ - regime scales with  $T^{7/2}$ . The electrons escape faster than the ions, resulting in a strong and positive radial electric field generated by the ions. Now, the  $1/\nu$ - transport is suppressed, because the  $\mathbf{E} \times \mathbf{B}$  drift influences the particle transport [53], the  $\sqrt{\nu}$  - or  $\nu$  - regime, respectively, is established instead. This leads to a better confinement due to a smaller dependency on  $T_e$  of the diffusion coefficients. This state is called *electron root*.

In general, for diverse species different collisionalities are found, resulting in different solutions of ambipolarity. Typically, the electron root is established in the plasma centre, whereas the ion root is found at the plasma edge. In between, multiple solutions are possible [54]. Therefore, for the electron root scenario the term *Core Electron Root Confinement* was established [52]. In figure 5.1 both the ion and electron flux are plotted for the ion root and the CERC case.

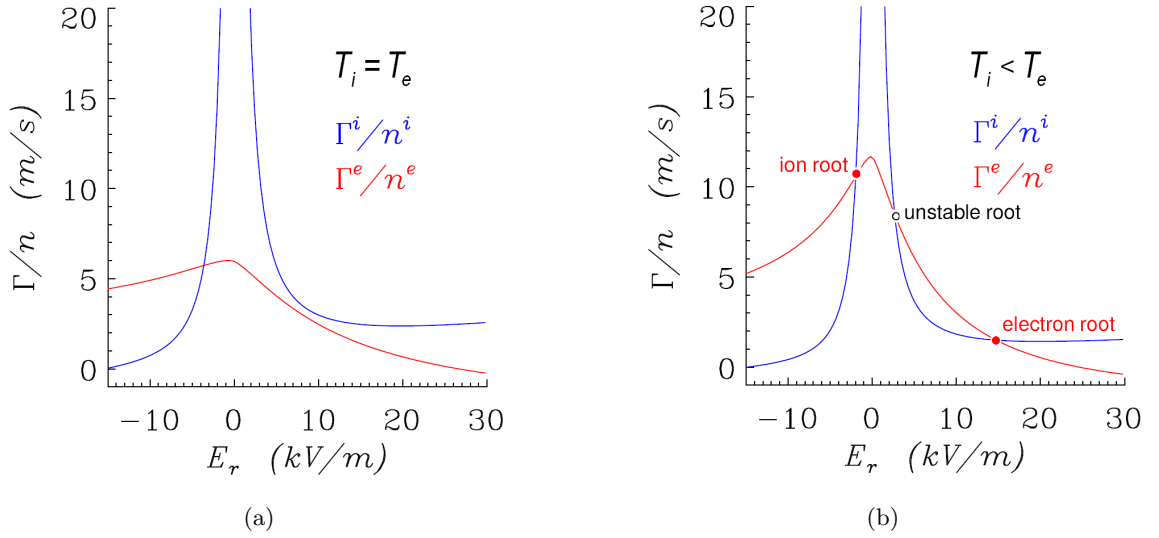


Figure 5.1: Electron (red) and ion flux (blue) with respect to the radial electric field  $E_r$ : Only one solution for ambipolarity condition  $\Gamma_i = \Gamma_e$  leads to pure ion root (left), whereas in case of multiple solutions the electron root occurs (right).

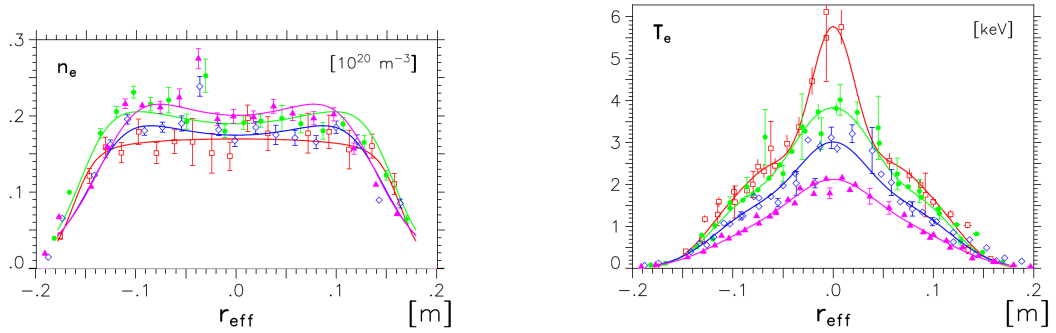


Figure 5.2: Electron density profiles for Core Electron Root Confinement for the standard configuration at the W7-AS experiment (from [51]): Electron density (left) and temperature (right) ECRH power scan ( $P_{\text{ECRH}} = 0.23 \text{ MW}$  (magenta),  $0.46 \text{ MW}$  (blue),  $0.77 \text{ MW}$  (green) and  $1.23 \text{ MW}$  (red)), CERC conditions are characterised by peaked  $T_e$  profiles.



In the case of three solutions of ambipolarity, the third one (in figure 5.1 (b)) is thermodynamically unstable, because

$$\sum_{\alpha} Z_{\alpha} e \frac{\partial \Gamma_{\alpha}}{\partial E_r} < 0 \quad (5.7)$$

has to be fulfilled [55]. This is equivalent to the situation of first order phase transitions in non-linear thermodynamics [56], where the middle solution for the characteristic "S" curve behaviour vanishes.

To achieve the Core Electron Root Confinement, the electron collisionality has to be decreased. Two possibilities are at hand: heating the electrons or decreasing the electron density. A possible way is using electron cyclotron resonance heating (ECRH) with highly peaked power deposition for electron heating.

Measurements have indicated the occurrence of CERC at different experiments like LHD, TJ-II, CHS and W7-AS [52]. A common feature is a strongly peaked profile of the electron temperature  $T_e$  emerging from the electron heating via ECRH, whereas the electron density profiles differ from each other. Also, the ECRH power and the magnetic field strength were different in the experiments.

CERC was found at density ranges from  $0.15 \cdot 10^{19} m^{-3}$  at LHD up to  $5.3 \cdot 10^{19} m^{-3}$  at W7-AS. Figure 5.2 shows density profiles at CERC conditions for a power scan with ECRH and a density scan at W7-AS [51]. The profiles show the tendency to become "hollow", meaning that the highest values for electron density are not found in the plasma centre.

This feature was found at several CERC experiments, one of the most outstanding examples was given by Maaßberg et al. [57] (see figure 5.3). A theoretical explanation may be that the outward directed electron flux driven by  $\partial_r T_e$  (outward pinch) is stronger than the inward pinch from the radial electric field  $E_r$  [52, 57]. This would theoretically lead to an inward diffusion driven by  $\partial_r n_e$ , however, a detailed analysis of the particle balance and the estimation of the particle sources are still needed at this point.

With respect to the capabilities of a interferometry system, the estimation of the hollowness of the density profile can be chosen as an optimisation goal for the best configuration of the lines of sight. Therefore, the hollowness effect has to be parametrised next.

### Parametrisation of the CERC effects on the electron density

Calculations and measurements of the indentation in the density profile implicate different depths and widths for the hollow part. In addition, the electron root was discovered for a certain range of maximal densities. Therefore, the density effect of CERC can be described by these three parameters: depth, width and maximum density.

The functional description of the density distribution is discussed in section 4.2.2. For the parametrisation of CERC it is chosen with ( $\rho = r_{eff}/a$ , where  $a$  is the effective radius

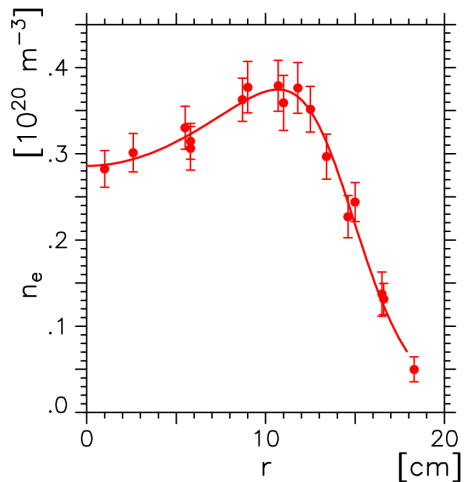


Figure 5.3: Hollow density profile at W7-AS, shot # 15155 [57].

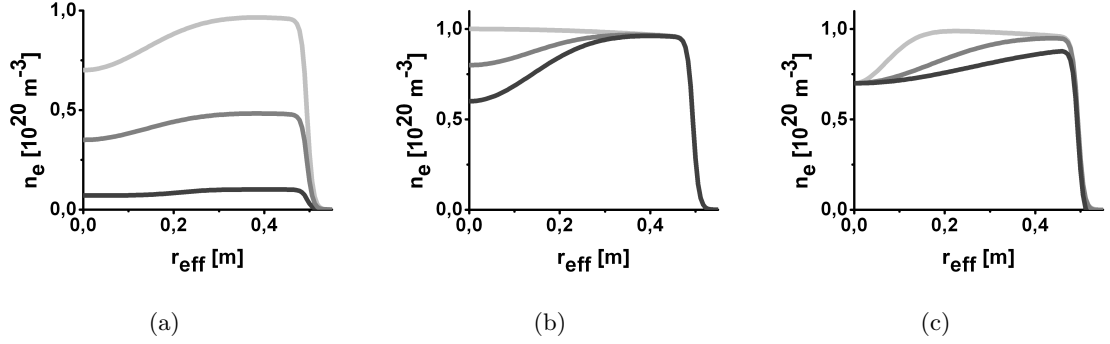


Figure 5.4: Variation of the hollow part of the electron density profile: (a) Maximum density ( $\theta_1 = 0.1$  dark,  $\theta_1 = 0.5$  medium,  $\theta_1 = 1.0$  light grey), (b) depth of the hollow part ( $\theta_2 = 0.4$  dark,  $\theta_2 = 0.2$  medium,  $\theta_2 = 0.0$  light grey) and (c) width of the indentation ( $\theta_3 = 0.4$  dark,  $\theta_3 = 0.25$  medium,  $\theta_3 = 0.1$  light grey).

at the last closed magnetic flux surface)

$$n_e(r_{eff}) = \theta_1 \cdot 10^{20} m^{-3} \cdot \left[ \frac{1 - 0.05 \cdot \rho^2}{1 + (\rho^2/0.97^2)^{20}} - \theta_2 \cdot \exp\left(-\frac{r_{eff}^2}{\theta_3^2}\right) \right]. \quad (5.8)$$

The parameters of interest can be identified as follows:  $\theta_1$  describes the maximum density,  $\theta_2$  depth of the indentation and  $\theta_3$  its width. The variation range is shown in figure 5.4.

The maximum density can be very small ( $\approx 0.1 \cdot 10^{20} m^{-3}$  as seen in LHD), but can also reach regions up to  $1 \cdot 10^{20} m^{-3}$ . Therefore,  $\theta_1$  is varied from 0.1 to 1.0 to cover all possibilities. The depth  $\theta_2$  ranges from 0.0 (absence of hollowness) to 0.4 (decrease of central density to 60% of maximum density). At last, the width of the indentation,  $\theta_3$ , reaches from 0.1, which means 10% of the plasma radius, to 0.4, where the edge of the indentation reaches the edge of the plasma, describing a concave density plateau. To measure the effect of these variations, one expects lines of sight in the central region of the plasma.

### 5.1.2 High confinement operational scenarios

At Wendelstein 7-AS, three regimes with improved confinement characteristics relative to their preceding state have been found: the Optimum Confinement regime, the H-mode and the so-called High Density H-mode. The effects on the density distribution generated by these regimes are predestinated as an optimisation goal.

All three regimes are ion-root solutions of the ambipolarity condition (5.1). They show high energy confinement times, making them favourable for possible operation modes for fusion devices.

#### Optimum Confinement regime

The maximum values of the ion temperature  $T_i$ , the energy confinement time  $\tau_E$  and the triple product  $n \cdot \tau_E \cdot T_i$  were achieved in the Optimum Confinement regime (OC) at low to medium densities with relatively steep density gradients at the edge of the confinement region ([58], see figure 5.5). The density decay is clearly located inside the last closed magnetic surface (LCMS), so the density at the plasma edge is low.

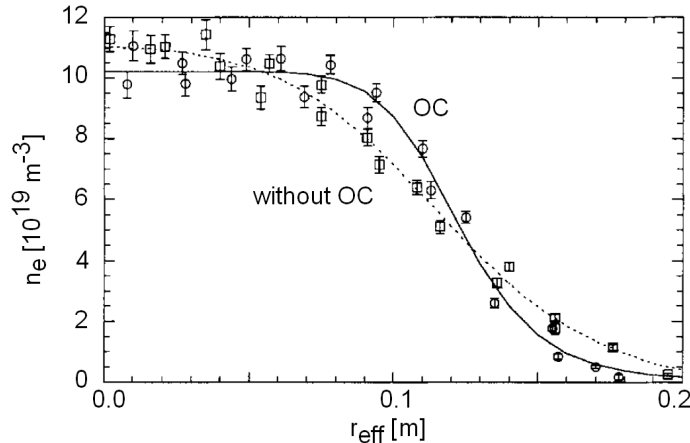


Figure 5.5: OC regime (W7-AS discharge # 39142) compared with a reference discharge (# 39161) without OC [58].

Compared to the ISS95 scaling (see chapter 6) the energy confinement time in OC regime is enhanced by a factor  $\tau_E^{OC}/\tau_E^{ISS95} \approx 2$  [59]. With OC, the highest confinement time of W7-AS ( $\tau_E \approx 60$  ms) and a maximum triple product of  $n \cdot \tau_e \cdot T_i = 5 \cdot 10^{21}$  eV s/m was established [60]. The maximum medium density under these conditions was stable at  $n = 1.1 \cdot 10^{20} \text{ m}^{-3}$  (figure 5.5). An upper density limit is given due to the increasing core radiation, for instance at a medium density of  $1.2 \cdot 10^{20} \text{ m}^{-3}$  for a heating power of  $P = 0.35$  kW by neutral beam injection (NBI).

Establishing the OC regime requires sufficient ion heating, the neutral beam injection (NBI) is favoured as a direct ion heating method. For the maximum confinement time as described before, a minimum heating power of  $P \approx 0.25$  MW was needed at W7-AS. Additional heating with electron cyclotron resonance heating (ECRH) is possible, for sufficient high densities ( $n_e \approx 10^{20} \text{ m}^{-3}$ ), where the energy exchange between electrons and ions is high, OC can be established with ECRH only ( $P_{ECRH} > 500$  kW) [58, 61].

The typical density profiles in the OC regime at W7-AS decline a few centimeters before the last closed magnetic surface. The decay length is about 5 cm. For W7-X, similar decay lengths in total are expected: The penetration depth of neutral atoms, injected by, e.g., gas puffing to refuel the plasma, is defined by atomic processes (ionisation by collisions). These processes do not depend on the plasma size, therefore the typical length scale of the density decay at the edge will approximately remain the same in total for W7-X as it was at W7-AS [48].

## H-mode in stellarators

The H-mode regime was originally discovered at tokamaks [62]. Later, it was also observed in stellarators, first at W7-AS [63]. The stellarator H-mode shows similar properties as its pendant on tokamaks: It is mainly characterised by an edge transport barrier which strongly reduces turbulent transport, the barrier is clearly located inside the LCMS. The transition to H-mode occurs spontaneously and is characterised by a sudden drop and a reduction in fluctuation of  $H_\alpha$  radiation. In contrast to the tokamak regime the stellarator H-mode does not show a dependence on the hydrogen isotopes in the plasma [64].

Two different states of the H-mode are known: In the *quiescent H-mode*, also H\*-mode, a layer of nearly completely suppressed turbulence appears at the plasma edge. This layer is localised in the outer 3 – 4 cm of the confined plasma, in a region with the strongest gradients in density  $n$ , electron and ion temperature,  $T_e$  and  $T_i$ . In this mode,

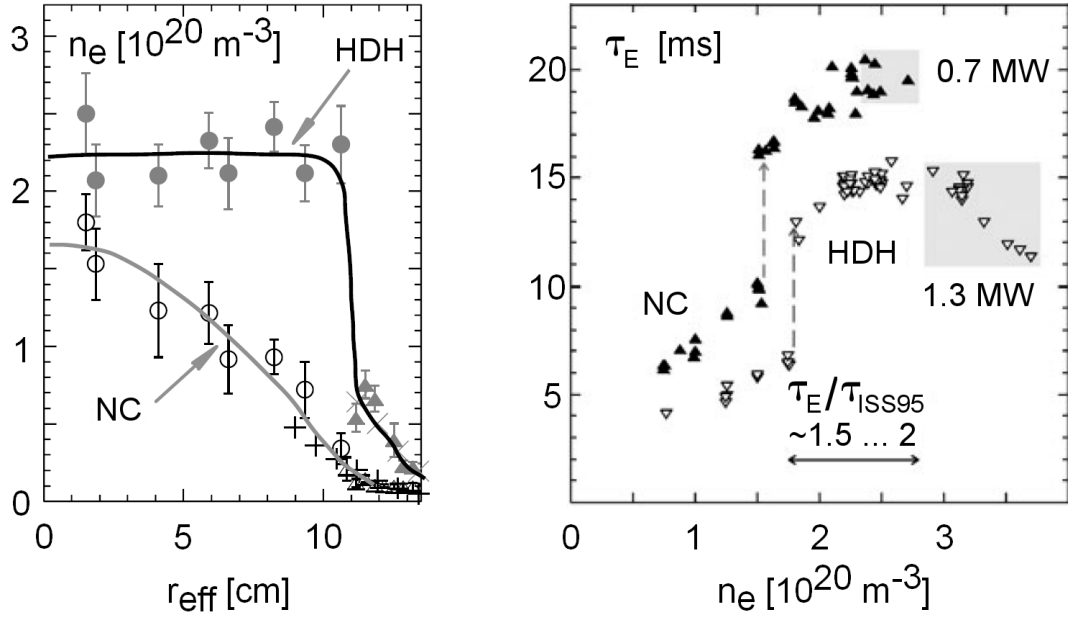


Figure 5.6: HDH regime at W7-AS: The density profile shows a very steep gradient at the edge of the confinement region in comparison to normal confinement (NC) regime (left, from [67]). The energy confinement time  $\tau_E$  is also increased, by a factor of 1.5 – 2 to the ISS95 scaling (right, [68]).

line integrated densities of about  $2 \cdot 10^{20} m^{-3}$  have been observed [65].

The *ELMy H-mode* is characterised by occurrence of edge-localised modes (ELM), which are transport events for particles and heat. A single ELM causes an energy loss of  $< 4\%$  of the stored energy, with a repetition frequency of  $1 - 2 kHz$  (at energy confinement times of about  $25 ms$ ) ELMs cause a major part of non-radiative energy loss [66].

The H-mode can be achieved only within three narrow regions of the rotational transform  $\iota$ :  $\iota \approx 0.474$ ,  $\iota = 0.525$  and  $\iota = 0.556$  [64]. At these small windows ( $\Delta\iota/\iota \approx 1\%$ ), H-mode occurs independently of heating scenario, gas program and toroidal magnetic field as well as with and without divertor. The regime is highly reproducible.

Like the OC regime, H-mode is able to achieve high energy confinement times compared to the ISS95 scaling:  $\tau_E^{H mode} / \tau_E^{ISS95} \approx 2$  [65].

### The HDH regime

As a last example for operational regimes with high energy confinement properties, the *High Density H-mode* (HDH) shall be introduced. It was one of the first discoveries at W7-AS after installing an island divertor.

The density profile in HDH mode shows a flat profile with a very steep gradient at the plasma edge (figure 5.6, left). The decay of the profile is located at or even beyond the LCMS so the density at the edge of the confinement region is significant.

In HDH mode, very high densities up to  $4 \cdot 10^{20} m^{-3}$  have been achieved [67], therefore it is sometimes called *improved confinement regime* (IC). Furthermore, HDH is characterised by high energy confinement times ( $\tau_e$  lies about a factor of 2 above ISS95 scaling, see figure 5.6, right), low impurity confinement times and edge localised radiation. The accumulation of impurities is avoided [67, 68].

For establishing HDH regime the NBI heating was the only choice, because the ECRH

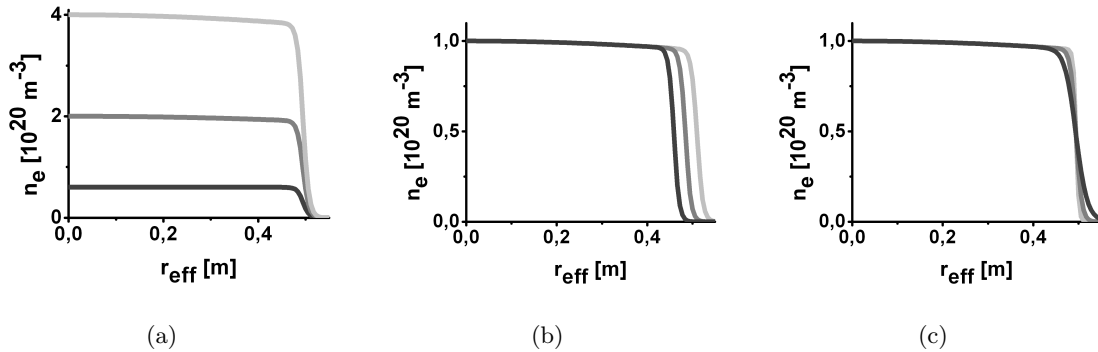


Figure 5.7: Variation of the electron density profile according to the estimated profiles for high confinement regimes: (a) Maximum density ( $\theta_1 = 0.6$  dark,  $\theta_1 = 2.0$  medium,  $\theta_1 = 4.0$  light grey), (b) position of the steepest gradient ( $\theta_2 = 0.9$  dark,  $\theta_2 = 0.95$  medium,  $\theta_2 = 1.0$  light grey) and (c) steepness ( $\theta_3 = 0.2$  dark,  $\theta_3 = 0.4$  medium,  $\theta_3 = 0.8$  light grey).

system at W7-AS operating at  $140\text{ GHz}$  was restricted by its cut-off density at approximately  $n_c = 1.2 \cdot 10^{20}\text{ m}^{-3}$ . HDH was established at NBI heating power of  $P = 0.4\text{ MW}$ .

Whereas the build-up of an HDH mode is accompanied with typical H-mode phenomena like the edge-localised transport barrier and the occurrence of ELMs, the HDH regime itself is completely ELM-free. For establishing HDH, a balance is needed between central fueling (via NBI) and edge density buildup by gas puffing [67].

Due to its very beneficial properties, the HDH mode is one of the most interesting scenarios for future experiments at W7-X, which will be equipped with an island divertor like W7-AS. The unique characteristics are therefore a logical choice as design criteria for diagnostic optimisation.

### Parametrisation of the density effects

Summarising the density effects of the three high confinement regimes, one finds three parameters as unique characteristics: the maximum density, the position of the steepest gradient at the edge of the confinement region and the steepness of this decay.

For a mathematical description, a similar parametrisation as used in the last section is applied:

$$n_e(\rho) = \theta_1 \cdot 10^{20}\text{ m}^{-3} \cdot \frac{1 - 0.05 \cdot \rho^2}{1 + (\rho^2/\theta_2^2)^{\theta_3}} \quad (5.9)$$

Again,  $\rho$  is the normalised effective radius:  $\rho = r_{eff}/a$ .

The parameter  $\theta_1$  describes the maximum density. According to the measurement results at W7-AS, it is varied from 0.6 to 4.0. Parameter  $\theta_2$  gives the position of the steepest gradient at the edge and is varied from 0.9 to 1.0, corresponding to a shift of the density decay of about  $5\text{ cm}$  in effective coordinates. The steepness parameter  $\theta_3$  lies between 0.2 and 0.8, resulting in decay lengths of about  $1 - 5\text{ cm}$ . The effect of the variations is shown in figure 5.7.

#### 5.1.3 Density effects of pellet injection

As a third example for a stellarator specific physical problem, a new plasma state found at the LHD experiment shall be a matter of discussion. The Large Helical Device (LHD) is a

superconducting helical confinement fusion device of heliotron/torsatron-type, in operation at the National Institute for Fusion Science (NIFS) in Toki, Japan [69, 70].

### *Super Dense Core at the LHD experiment*

Recent findings at the Japanese experiment give indications for a new high dense plasma regime, called *Super Dense Core* (SDC), created by injection of hydrogen pellets [71]. The measured density profiles show maxima of about  $4.6 \cdot 10^{20} \text{ m}^{-3}$ , which are the highest densities ever diagnosed in a stellarator. In addition, the peak seems to be stable for a few milliseconds, until the next pellet is injected. This procedure may be a possible method to reach fusion conditions in a stellarator.

Because pellet injection for refueling the plasma is also planned, the scenario discovered at LHD may be also of interest for experiments at W7-X. Here, the infrared interferometry system is an appropriate diagnostic to measure the expected high and peaked densities with the necessary temporal resolution.

Fortunately, the results from LHD are not needed to be scaled for W7-X due to the fact, that the plasma radii are of similar size (about  $0.5 \text{ m}$ ). The effects in size are therefore taken as absolute values.

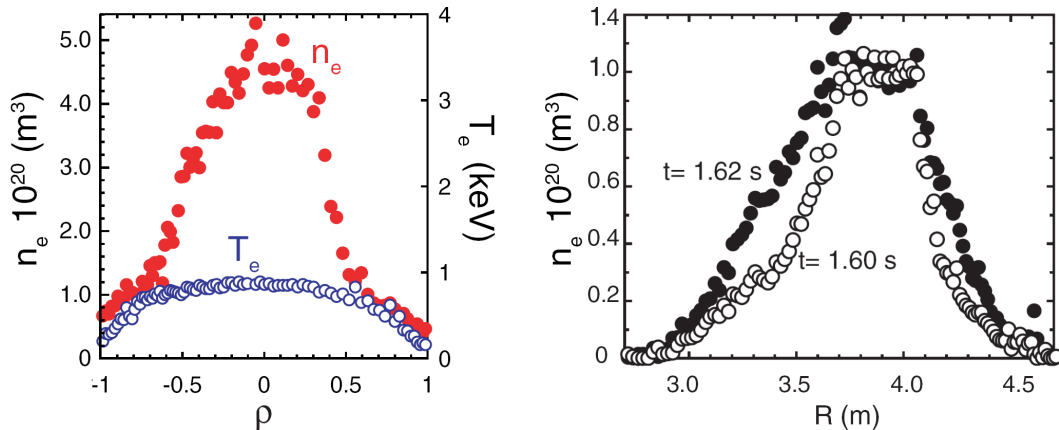


Figure 5.8: The super dense core plasma found at LHD [71]: Peaked density profile with very high densities in the core region (left,  $\rho = r/a$ ), width of the centre density peak before and after pellet injection (right).

The plasma density profile shows a strong peak at the plasma centre (see figure 5.8). At the plasma core, a small plateau region with a diameter of  $0.3 - 0.4 \text{ m}$  is formed. The decay from plasma centre to the plasma edge is not very steep, near the plasma edge indications for a small step structure can be found.

As pointed out by Ohyabu et al. [71], the core region is characterised by an internal fusion barrier at a normalised radius of  $0.3 \geq \rho \geq 0.5$ . The position of the barrier can be influenced with an appropriate choice of the magnetic configuration.

After the pellet injection, a shift to a broader peak was observed (figure 5.8, right). Then, the peak width decreases, until a new pellet is injected.

### **Parametrisation of the SDC effect**

To parametrise the density effects shown above, three parameters of interest should be estimated by the interferometer: the maximum density, the width of the density peak and its decay length. For this, a combination of two of the density distributions from section

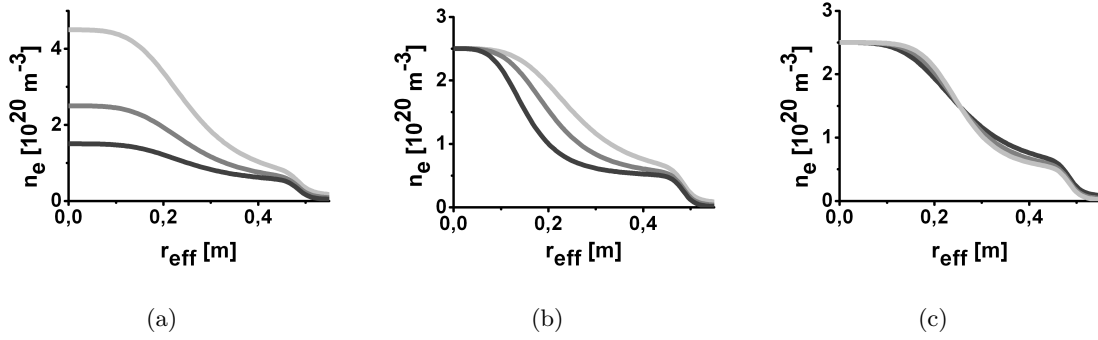


Figure 5.9: The parametrization of the density peak as expected for pellet injection: (a) Maximum density ( $\theta_1 = 1.0$  dark,  $\theta_1 = 2.0$  medium,  $\theta_1 = 4.0$  light grey), (b) peak width ( $\theta_2 = 0.3$  dark,  $\theta_2 = 0.4$  medium,  $\theta_2 = 0.5$  light grey) and (c) steepness of the peak decay ( $\theta_3 = 2.0$  dark,  $\theta_3 = 2.5$  medium,  $\theta_3 = 3.0$  light grey).

4.2.2 is chosen: The first one generates a basic plateau at  $0.5 \cdot 10^{20} m^{-3}$  (to form the step effect at the plasma edge) and a second one generating the central peak:

$$n_e(\rho) = 10^{20} m^{-3} \cdot \left[ \underbrace{0.5 \cdot \frac{1 - 0.05 \cdot \rho^2}{1 + (\rho^2/0.97^2)^{20}}}_{\text{basic plateau}} + \underbrace{\theta_1 \frac{1 - 0.05 \cdot \rho^2}{1 + (\rho^2/\theta_2^2)^{\theta_3}}}_{\text{central peak}} \right] \quad (5.10)$$

The parameter  $\theta_1$  describes the height of the density peak as shown in figure 5.9 (a). Very high densities should be possible, therefore  $\theta_1$  varies from 1.0 to 4.0 (corresponding to central densities of  $1.5 - 4.5 \cdot 10^{20} m^{-3}$ ).  $\theta_2$  parametrises the width of the density peak, lying in the interval of  $0.3 \geq \theta_2 \geq 0.5$  like estimated for the position of the internal transport barrier. The steepness of the peak decay ( $\theta_3$ ) changes from 2.0 to 3.0, the effect is illustrated in figure 5.9 (c).

It has to be mentioned that the parameter ranges, according to the actuality of the SDC finding, are only an estimate from today's point of view. But by the measurement results achieved so far, the parametrization shown here seems to be appropriate.

## 5.2 Multi-channel interferometer design

In this section, the results for the multichannel interferometer will be presented and compared. First, the designs for a four-channel interferometer are presented, based on the different physical problems introduced in the last section. For a comparison, also "ideal" designs without technical restrictions of the port system were computed, so that the loss of information due to the port system can be illustrated.

The influence of the magnetic configuration and the effect of the plasma outside the last closed flux surface is analysed, too. Finally, the design of an eight-channel interferometer, using the four-channel configuration as starting point, is presented.

### 5.2.1 Four-channel interferometer

The technical boundary conditions, namely the port system at W7-X which restricts the directions of the lines of sight, have been described in section 3.2.2. Given the 11 allowed

positions for retro-reflectors inside the plasma vessel and the three port tubes (see figure 3.2, (b)), 101 beams in total can be realised in the interferometry plane: Four vacuum flanges per port can be applied, and the beam lines have to fit the port system. Beams which intersect with the plasma vessel and the in-vessel components were eliminated by ray tracing calculations. As a remark it should be mentioned that every "allowed" beam crosses the plasma, no beam line lies outside the separatrix.

In addition to these beams, the extra beam at the plane symmetric to the Thomson scattering diagnostic (figure 3.4) is also available. The designs presented in the following are calculated with and without this beam.

Because of the discrete nature of the problem, no optimisation routine could be used to find the optimal design. Therefore, the following procedure was applied: First, one beam was optimised by scanning all possible beam configurations with respect to one parameter of interest (normally the maximum density). Secondly, the best position for the second chord was computed for two parameters of interest, while the first beam was kept fixed. The calculation for the third and fourth beam was done for all three parameters of interest. As a next step, all beams were varied again individually one after another, until the highest *EU* value of the scan did not change anymore. In case of a configuration including the (fixed) extra beam, only three beam lines were optimised, respectively.

In the scanning process, identical beam lines were not forbidden by default. Whereas in reality identical beam lines cannot be realised (because it is impossible to separate them after crossing the plasma), they were not excluded from the calculations to study their possible influence on the design.

For the results without technical constraints, a scanning is not necessary, an optimisation routine can be applied instead. Here, the *Amoeba* routine from [46] was used. Design parameters were starting and ending point, lying on a circle around the plasma cross section (see section 4.2.3).

A description of the algorithms and the source code used for the computation of the *EU* is given in appendix A.

## Core Electron Root Confinement

The variation of density distributions for Core Electron Root Confinement (CERC) was introduced in section 5.1.1: The density profile is characterised by a hollow part at the plasma centre. Using the parametrisation of the parameter range of interest given there, the Expected Utility was calculated to find the optimal beam configuration.

The density effect of CERC is mainly located at the plasma centre, for the design of the four-channel interferometer a configuration with beam lines through the inner part of the plasma is expected. The result of the calculations is shown in figure 5.10.

As expected, the lines of sight for the optimal configuration are located at the inner part of the plasma. The chords fan out to cover the part of the plasma where the effect of the hollowness occurs. In the configuration without the extra beam, two beam lines cross the plasma centre directly, whereas only the extra beam does so in the second design. However, calculating the Expected Utility, the latter is preferred:

$$\begin{aligned} EU &= (6.43 \pm 0.01) \textit{ bit} \quad \text{without and} \\ EU &= (6.86 \pm 0.01) \textit{ bit} \quad \text{including the extra beam.} \end{aligned}$$

The explanation for this result can be given by taking into account the very long distance the extra beam has to cross through the plasma: If, e.g., the maximum density is changed from  $\theta_1 = 0.1$  to  $\theta_1 = 1.0$ , the difference in the phase shift for the extra beam is  $\Delta\phi =$



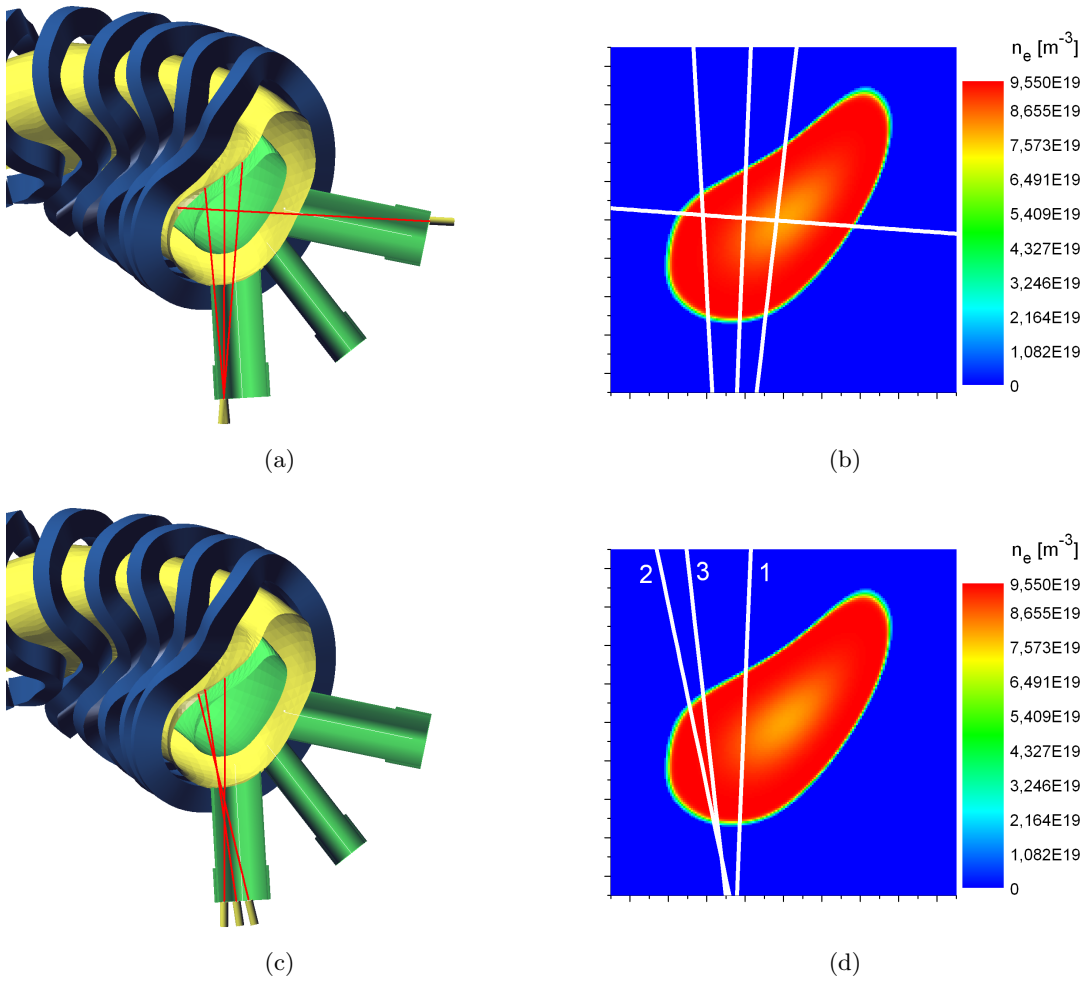


Figure 5.10: Design results for a four-channel interferometer with respect to the measurement of CERC: Without the extra beam (upper row) and including (lower row). The beams are shown with respect to the port system (left) and in relation to the density distribution (right).

6.172 rad (for  $\theta_2 = 0.2$  and  $\theta_3 = 0.25$ , see figure 5.4). The beam lines crossing the plasma centre in figure 5.10 (b), will only show a phase shift difference of  $\Delta\phi = 4.0944 \text{ rad}$  (horizontal beam) and  $\Delta\phi = 4.1626 \text{ rad}$ , respectively. Assuming a constant background noise as leading error (as has been done here), the extra beam will show a better signal-to-noise ratio for the change of  $\theta_1$ , which seems to be preferred by the *EU*.

This hypothesis can be confirmed by looking at the design result without technical restrictions (figure 5.11). Here, all beam lines take the longest way possible through the plasma, while they stay fanned out. Therefore, a change of the parameters of interest results in a significant change of the measured phase shift: For the situation described above, the maximum difference in the phase shift is given for beam 2 (figure 5.11) with  $\Delta\phi = 7.209 \text{ rad}$ , which is even larger than for the extra beam in the constrained configuration.

The situation is similar for the other parameters of interest: If  $\theta_2$  is changed from 0.0 to 0.4 with  $\theta_1 = 0.5$  and  $\theta_3 = 0.25$ , the maximum change in phase shift for the unconstrained configuration is given for beam 1 with  $\Delta\phi = 0.782 \text{ rad}$ . For the constrained configuration including the extra beam the maximum difference in the phase shift is smaller:  $\Delta\phi = 0.708 \text{ rad}$  for the extra beam,  $\Delta\phi = 0.266 \text{ rad}$  for beam 1 (figure 5.10 (d)). A similar result is given for the variation of  $\theta_3$ .

The Expected Utility for the unconstrained configuration is calculated with  $EU = (8.71 \pm 0.01) \text{ bit}$ . Approximately 2 bit of the possible information gain are therefore lost due to the technical boundary conditions. Taking into account the impact of the parameter changes on the phase shift, one can conclude here that the *EU* prefers designs where the impact of the parameter of interest is strong.

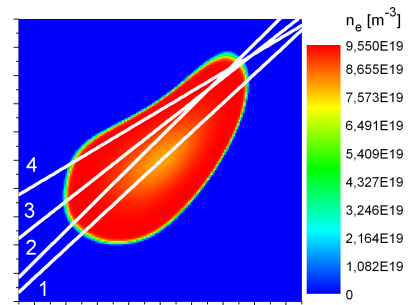


Figure 5.11: Design result for a four-channel interferometer with respect to the best estimation of CERC density effects, without technical restrictions.

## High confinement regimes

As pointed out in section 5.1.2 the density effects of the different high confinement regimes (optimal confinement, H mode and High Density H mode) are mainly localised at the plasma edge. As a consequence for the interferometer design one would expect a configuration with beams in this region of the plasma. This may become complicated, because only a few beam lines matching this criterion are accessible.

Figure 5.12 shows the result of the optimisation, again for a configuration with and without the extra beam. In both configurations, a chord nearly tangential to the plasma is established, also a beam line crossing the plasma centre.

Whereas the design including the extra beam shows two chords near the plasma edge, only one such beam is preferred for the other configuration. This is surprising, but can be explained by a closer look at the calculation: The two nearly horizontal chords are not the optimal design solution, the highest *EU* is given for a beam line which is identical to the beam at the plasma edge. Because this is not realisable for technical reasons, the second and third best result has to be taken, leading to the configuration as shown in the figure. Nevertheless, this result again points out the meaning of the signal-to-noise ratio, which would be theoretically increased by two beams propagating along the same line of sight.

As in the example before, the configuration including the extra beam shows a higher

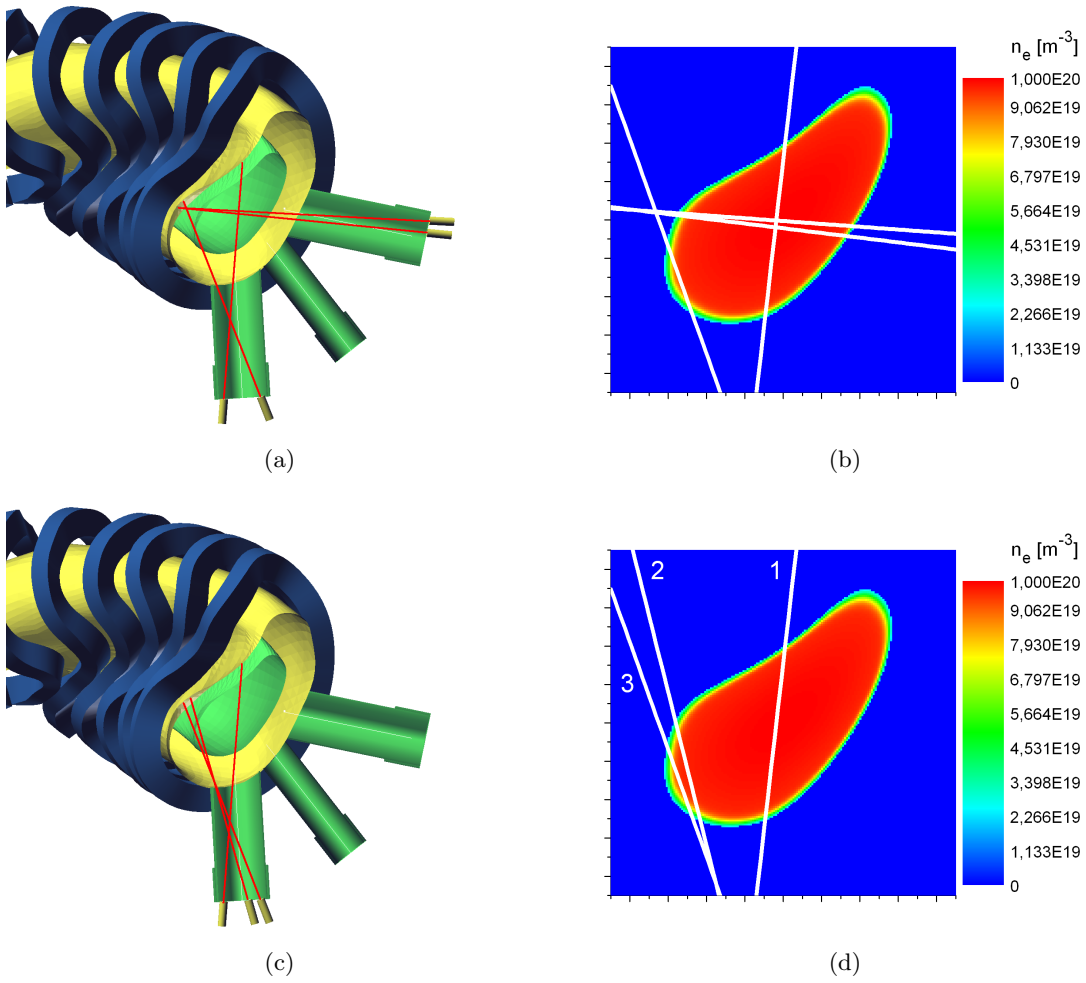


Figure 5.12: Design results for a four-channel interferometer for high confinement regimes: Without the extra beam (upper row) and including (lower row); with respect to the port system (left) and in relation to the density distribution (right).

$EU$ :

$$\begin{aligned} EU &= (8.53 \pm 0.01) \text{ bit} \quad \text{without and} \\ EU &= (8.85 \pm 0.01) \text{ bit} \quad \text{including the extra beam.} \end{aligned}$$

The difference in this case is slightly smaller, which could be explained by the importance of the measurements of the effects at the plasma edge, which cannot be provided by the extra beam.

For comparison, the beam design without technical boundary conditions was computed, too. The result is shown in figure 5.13. Here, one finds two beams tangential to the plasma edge, but traversing the plasma on a very long path. The other beams cross the inner part of the plasma. The  $EU$  is much larger than in the restricted case:  $EU = (28.3 \pm 0.2) \text{ bit}$ , i.e. the information gain is reduced to less than one third by the boundary conditions. This tremendous loss can be explained, if the special situation of the chords at the plasma edge is taken into account: On the one hand side, the beams in the restricted configurations cross the plasma only on a very small distance, whereas the path of the unrestricted beams is about three to four times longer. On the other hand, the density in this part of the plasma is low, again reducing the signal of the probing beams.

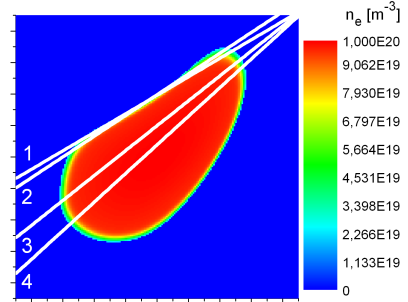


Figure 5.13: Design result for a four-channel interferometer with respect to the problem of high confinement regimes, without technical restrictions.

Quantitatively, this effect can be demonstrated by calculating the  $EU$  with only one chord at the plasma edge for the unconstrained configuration. This leads to  $EU = 20.5 \text{ bit}$  by removing beam 1 in figure 5.13, and  $EU = 17.8 \text{ bit}$  by removing beam 2, respectively. Without both beams, the  $EU$  is given with only  $7.48 \text{ bit}$ .

These differences can be explained by taking into account the effect of the parameters of interest: As shown in the hollowness example before, the  $EU$  increases if the variation of the parameters of interest leads to significant changes in the measured phase shift. In table 5.1 this effect is shown: The parameters  $\theta_2$  and  $\theta_3$ , which describe the edge behaviour of the density distribution, are varied, the effect on the measurement is given for the edge beams of the unconstrained configuration (figure 5.13) and for the beams 2 and 3 of the constrained design (figure 5.12 (d)).

The impact of the parameter change on the unconstrained configuration is three to six times larger for the variation of  $\theta_2$ , and five to nine times larger for  $\theta_3$ . This remarkable

configuration and beam	$\Delta\phi$ for $0.9 \leq \theta_2 \leq 1.0$	$\Delta\phi$ for $0.2 \leq \theta_3 \leq 0.8$
unconstrained, beam 1	4.53 rad	0.46 rad
unconstrained, beam 2	5.93 rad	0.03 rad
constrained, beam 2	0.96 rad	0.05 rad
constrained, beam 3	1.53 rad	0.10 rad

Table 5.1: Change of the measured phase shift  $\Delta\phi$  if the parameters of interest for high confinement regimes are varied. If  $\theta_2$  was changed,  $\theta_1 = 1.0$  and  $\theta_3 = 0.4$  was set. For the variation of  $\theta_3$ ,  $\theta_1 = 1.0$  and  $\theta_2 = 0.95$  was chosen.

difference explains the increase in the  $EU$  for the unconstrained design – the beam line design in the restricted configuration obviously suffers from short path lengths through the edge part of the plasma, where the variations of the parameters of interest lead to only small effects on the data. For the variation of  $\theta_3$ , the difference in the measured phase shift is in the order of the assumed measurement error:  $\Delta\phi = 0.05 \text{ rad}$  and  $\Delta\phi = 0.1 \text{ rad}$ , whereas the error is given with  $\sigma = 0.1 \text{ rad}$  (see section 3.2.3). Therefore, it may be difficult to estimate  $\theta_3$  with the constrained beam line configuration.

### Pellet injection

As well as the problem of CERC, the injection of hydrogen pellets causes density effects mainly localised in the centre regions of the plasma, so one expects beam lines crossing the inner part of the plasma. The parametrisation and the variation of the parameters of interest are given in section 5.1.3. Again, an optimal beam line configuration with and without the extra beam was calculated, the results are displayed in figure 5.14.

The configuration of the chords confirms the expectations: The beam lines cross the plasma through the central regions, in this case only marginally fanned out, but more parallel than in the CERC case. In case of the extra beam configuration, the three optimised

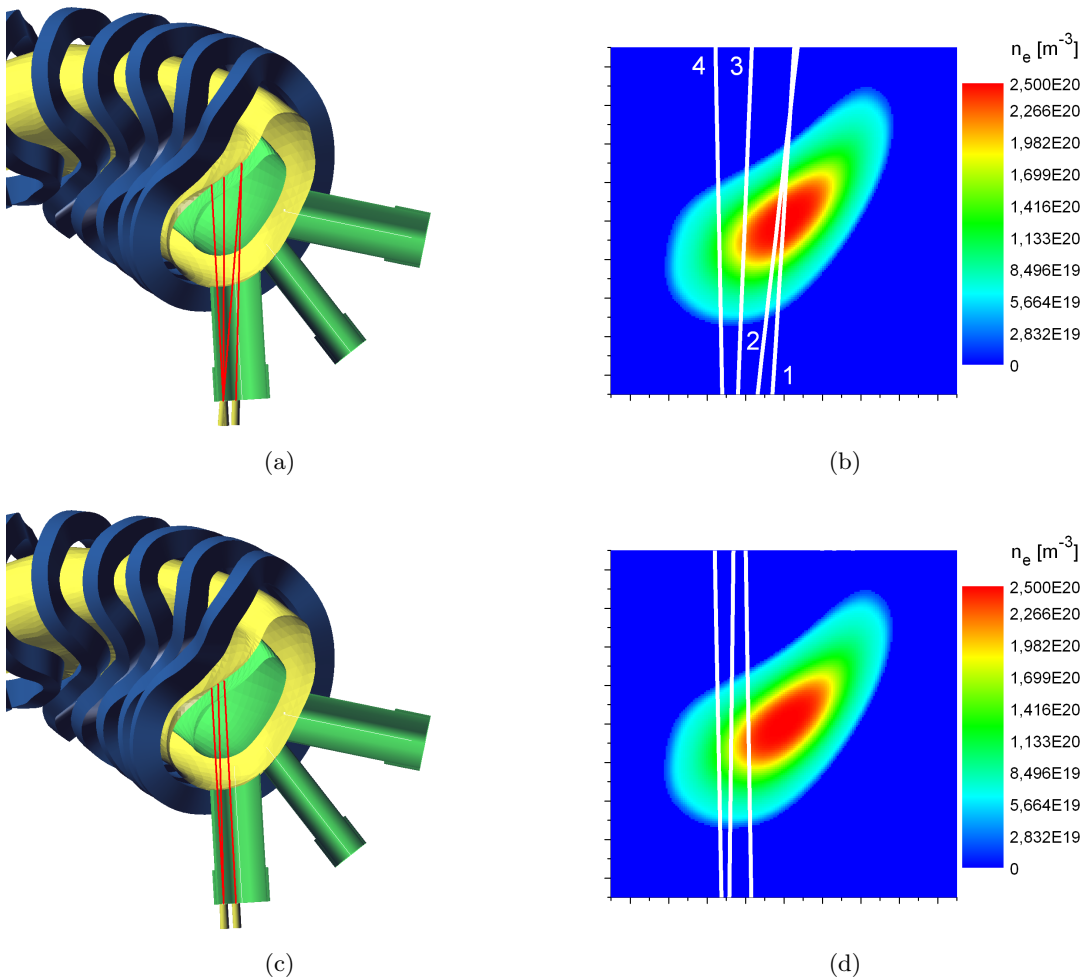


Figure 5.14: Four-channel interferometer design for the problem of pellet injection: Without the extra beam (upper row) and including (lower row); with respect to the port system (left) and in relation to the density distribution (right).

beam lines are localised near the expected edge of the pellet driven cone structure of the density distribution in the plasma centre. Without the extra beam, two beams cross the plasma approximately through its centre. As for the other two physical problems, the Expected Utility for the configuration including the extra beam is the highest:

$$\begin{aligned} EU &= (8.24 \pm 0.01) \text{ bit} \quad \text{without and} \\ EU &= (8.86 \pm 0.01) \text{ bit} \quad \text{including the extra beam.} \end{aligned}$$

Again, this can be explained with the effect of the parameters of interest on the data: Varying the maximum density  $\theta_1$  from 1.0 to 4.0 (with  $\theta_2 = 0.4$  and  $\theta_3 = 2.5$ , see figure 5.9), one finds a difference in the measured phase shift of  $\Delta\phi = 10.45 \text{ rad}$  for the extra beam and  $\Delta\phi = 6.52 \text{ rad}$  for beam 1 in figure 5.14 (b). This finding confirms the results from the previous problems: The  $EU$  increases if the effects of the parameters of interest on the measurement become larger.

At last, again the optimisation result with technical restrictions shall be compared with the unconstrained design (figure 5.15). The result is similar to the case of CERC: The  $EU$  for the unconstrained design is higher:  $EU = (10.97 \pm 0.01) \text{ bit}$ , also the difference to the restricted case ( $\approx 2 \text{ bit}$ ) is comparable to the CERC problem. As expected, the beam lines crossing the plasma on very long distances, covering the inner part of the plasma, were the density effects of pellet injection are localised.

One can also conclude here that the extra beam is the most valuable one for all three physical problems, therefore it makes sense to discuss only configurations including this chord in the following.

At this point, three different designs for three different physical problems are calculated. Only one configuration can be realised, the selection is made by quantitative comparison of the different designs next.

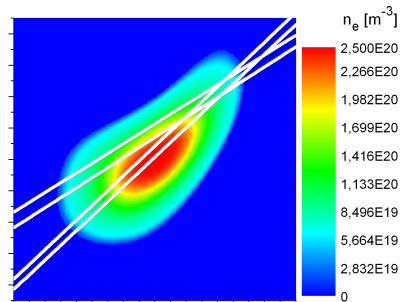


Figure 5.15: Design result for a four-channel interferometer with respect to the problem of pellet injection, without technical restrictions.

## 5.2.2 Finding the optimal design by comparison

The quantitative description of a design using the Expected Utility can be applied to compare different designs by calculating the  $EU$  for every configuration and for every physical problem, respectively. In table 5.2 the results for the three four-channel designs and the three physical questions are listed. Only the designs including the extra beam have been considered, because the designs without the extra beam always showed a smaller  $EU$ . In addition, the design from Airila et al. [45] is listed for comparison.

The design for the problem of CERC shows a quite high  $EU$  for the pellet injection problem, and vice versa. The difference is only  $0.26 \text{ bit}$  and  $0.1 \text{ bit}$  to the  $EU$  from the optimal result, respectively, which is about  $1.5 - 3\%$ . This is not surprising, because both effects are localised in the centre region of the plasma – similar effects obviously result in similar designs. On the other hand, both configurations result in a lower  $EU$  for the problem of high confinement regimes (with a difference of  $1 \text{ bit}$  and higher compared to the optimal design), which causes density effects at the plasma edge. The other way around, the diagnostic configuration for the high confinement problem gives less information about the centre region of the plasma. This results in a low  $EU$  for CERC and pellet injection, with a difference of  $0.4$  up to  $> 2 \text{ bit}$  to the respective optimal design, which is  $> 20\%$ .

Configuration	$EU$ [bit] for: CERC	$EU$ [bit] for: High Confinement	$EU$ [bit] for: Pellet injection	Sum $EU$ [bit]
CERC	$6.86 \pm 0.01$	$7.86 \pm 0.01$	$8.60 \pm 0.01$	23,32
High Confinement	$6.47 \pm 0.01$	$8.85 \pm 0.01$	$6.59 \pm 0.01$	21,90
Pellet injection	$6.76 \pm 0.01$	$7.52 \pm 0.01$	$8.86 \pm 0.01$	23,14
Airila et al. [45]	$5.845 \pm 0.005$	$8.24 \pm 0.01$	$6.78 \pm 0.01$	20,87

Table 5.2: Comparison of interferometer designs: The Expected Utility for different design configurations (rows) with respect to the three different physical problems (columns).

To find the design which is the best for all three analysed problems, the different  $EU$ s are added (see table 5.2). One has to be aware, that such a simple summation is only possible if the physical problems are exclusive: Expected Utilities, calculated according to equation (2.41), can only be added if the ranges of the parameters of interest  $\theta_i$  do not overlap. Only in this case the integrals for the calculation of the different  $EU$ s can be separated. So, for the problems analysed here, the summation is only an approximation, because the parameter spaces are not completely separated (e.g., the maximum density is a parameter of interest for all three problems), and the overlap can hardly be quantified.

The configuration for the estimation of CERC shows the highest sum. This was expected because the design gives a high  $EU$  value for the problems of CERC and pellet injection, and only a difference of 1 *bit* to the optimal design for the high confinement problem.

Also in table 5.2 the  $EU$ s for the design from [45] are listed. They show small values especially for the two problems located at the plasma centre. The result for the confinement regime problem is slightly better, but if all results are summarised this diagnostic configuration shows the worst performance for the problems analysed here.

Summarising that, the beam line configuration for the CERC problem including the extra beam is the best design according to the problems analysed here. One has to keep in mind that this result is valid only for the three problems of CERC, high confinement regimes and pellet injections; other physical questions may result in different designs.

### 5.2.3 Influence of magnetic configurations

In the previous sections the results presented have been calculated for the so-called *standard case* at W7-X. This magnetic configuration is characterised by a rotational transform of  $\iota = 1.0$  at the last closed flux surface and a minor radius of  $a = 0.52$  m [72]. The remarkable flexibility of the W7-X design also allows different magnetic configurations, two cases with significant impact on the plasma shape - the so-called high- $\iota$  case and high- $\beta$  case - are discussed in this section (see appendix B for details). In addition, the influence of the plasma outside the separatrix will be discussed.

The influence of these effects is analysed for the four-channel interferometer including the extra beam. For the physical problem the measurement of CERC was chosen, because the design resulting from this problem provided high  $EU$  values for the other physical questions, too.

#### High- $\iota$ case

For this plasma configuration with  $\iota = 1.2$  at the last closed flux surface the minor radius is given with  $a = 0.48$  m [72]. In the interferometry plane, the plasma is slightly elongated,

as it can be seen in figure 5.16 (b). In the same figure one can see the optimal sight line configuration for the measurement of CERC, which is exact the same configuration as for the standard case. So, in this case, the magnetic configuration has no effect on the design.

### High- $\beta$ case

The shape of the magnetic flux surfaces changes if the effect of the plasma is taken into account. Whereas the aforementioned cases are vacuum configurations, i.e. the effect of the plasma is neglected, the next example will consider a non-vanishing plasma pressure for the standard case. As the describing quantity,  $\beta$  is the ratio of the volume-averaged kinetic pressure of the plasma and the magnetic field pressure from the toroidal magnetic field:  $\beta = 2\mu_0 \langle p/B^2 \rangle$ . It was chosen as  $\beta = 0.04$  [73] for this analysis. This has significant impact on the plasma shape (see figure 5.16 (c)). The fan-like structure of the optimal configuration of the beam line changes, but the beams still cover regions near the plasma centre.

### Plasma outside the separatrix

In the former analyses it was assumed that no plasma is located outside the last closed magnetic surface (LCMS). This is not a realistic assumption, so the effect of this plasma is analysed in the last example. For the standard case, a "ribbon" of plasma with a width of 5 cm in effective coordinates was added outside the LCMS. The density distribution inside this ribbon was chosen as an exponential decay, starting with the plasma density at the separatrix, the decay length was also 5 cm.

In principle, the mapping routine used for the calculation of the magnetic coordinates works only inside the separatrix, but the flux surfaces can be extended by numerical simulation. The 5 cm chosen for the plasma outside the LCMS are given by the numerical resolution of this method [74].

The exponential decay is used here for a rough estimate of the influence of the plasma outside the separatrix on the density measurement. The density effects in this region are of enormous complexity (see, e.g., [75, 76]). However, the density is of two or three orders of magnitude smaller than in the confined plasma region, and the path lengths of the probing beams in the edge regions are small, so the effect of this approximation on the data should be negligible.

In figure 5.16 (d) the plasma with an exponential decay outside the LCMS is shown, also the configuration of the optimal beam lines for the measurement of CERC. Only one beam has changed in comparison to the original design without an additional plasma edge. This is not surprising, because the influence of the plasma edge on the measurement signal should be small for beam lines crossing the centre part of the plasma. Therefore, its influence on the design is not very large, too.

Comparing the *EUs* for the different designs (see table 5.3) one can state that the original design leads to a sufficient good configuration even for different plasma conditions. For the case of the high- $t$  configuration the beam line configuration is identical, in case of plasma outside the separatrix the changes are minimal. Here, the error intervals of both *EUs* overlap. The difference between the standard case design and the optimal design for the  $\beta = 0.04$  case differ a little bit more, however, the difference is only about one percent of the *EU*.

In general, the configurations shown here are very similar from a technical point of view: They all use the same port to access the plasma vessel (see figure 5.12). It may be possible to realise all configurations, e.g. by a movable mirror at the outer side of the port,



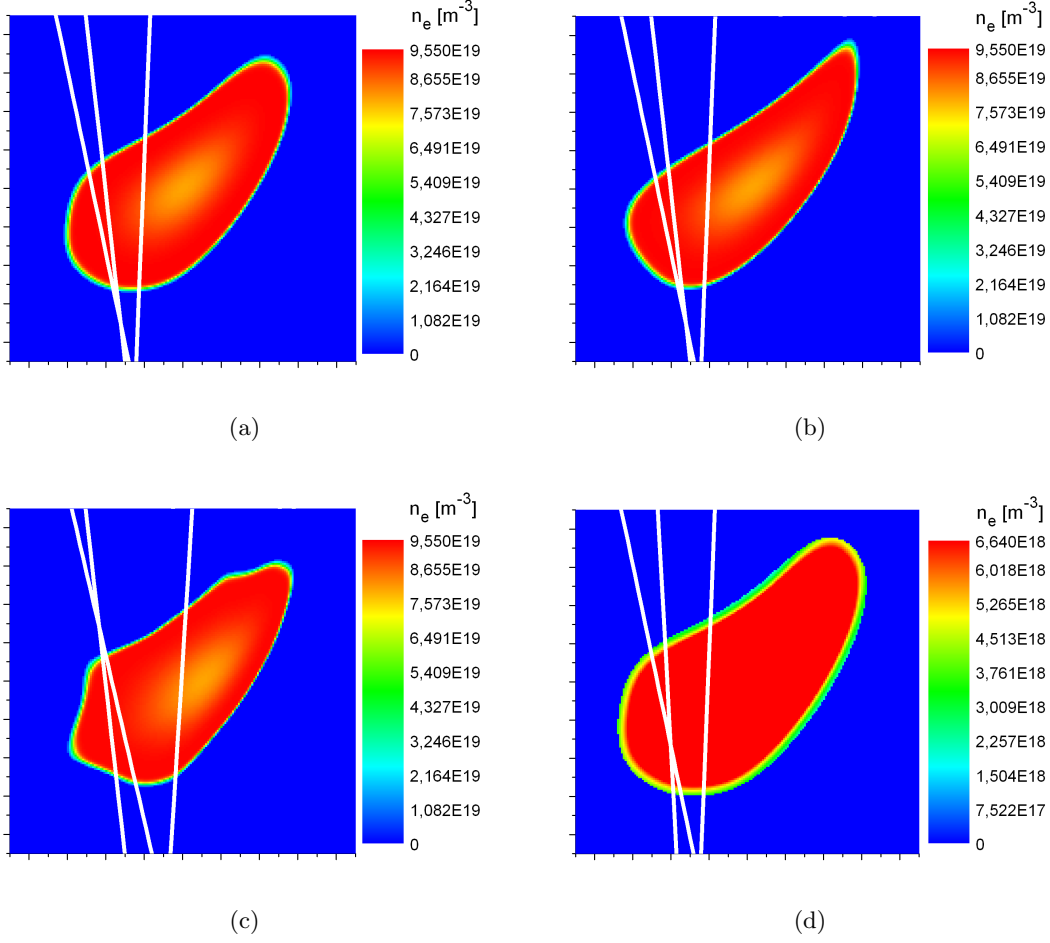


Figure 5.16: Different plasma configurations: Standard case (a), high- $t$  case (b),  $\beta = 0.04$  case (c) and plasma outside the separatrix (d). In the latter case, the plasma inside the LCMS is indicated as red for a better visualisation of the edge plasma.

Configuration	$EU$ in bit for: High- $t$ case	$EU$ in bit for: $\beta = 0.04$ case	$EU$ in bit for: Density outside LCMS
CERC design	$6.741 \pm 0.005$	$6.748 \pm 0.005$	$6.839 \pm 0.005$
Configuration specific design	$6.741 \pm 0.005$	$6.822 \pm 0.005$	$6.847 \pm 0.005$

Table 5.3: Different magnetic configurations: The performance of the original design from standard case (upper row) and the best design for the respective configuration (lower row).

so that the different magnetic configurations can be treated with the respective optimal beam line configuration.

#### 5.2.4 Eight-channel interferometer

The interferometer at Wendelstein 7-X is proposed to be a four-beam configuration in the start-up configuration. However, it is also planned to extend the set-up later to an eight-channel diagnostic. A possible configuration is outlined next.

First, it is assumed that the eight-channel configuration will be an extension of the four-beam interferometer, so four chords (here: the design result for the measurement of CERC) are taken as given. The configuration of the new four channels has to be designed.

The problem of the physical question of interest has also be discussed again in this context. With eight lines of sight, it is possible to fulfill more than one design goal. Two approaches are possible: A new parametrisation for the density profile with more than three parameters of interest, or using the sum of several *EUs* calculated for problems similar to the four-channel diagnostic.

For the interferometer, the first approach turns out to be unpracticable due to the strongly increasing computation times (see table A.1 in appendix A): Even if calculated on eleven double-core CPUs at the same time by using a parallelisation algorithm, the computation of one *EU* for five parameters of interest takes about 10 hours. At least, one and a half month would be necessary for the complete scanning of one beam over all 101 beam positions.

The second approach, finding the maximum of the sum of several *EUs*, does not suffer from this problem, in addition, it also offers the advantage to implement a weighting to the process: If necessary, the effect of one of the sub-problems can be highlighted by an appropriate weighting factor. On the other hand, for two problems, where the spaces of the parameters of interest overlap, this summation is only an approximation, as stated out before in section 5.2.2.

For the example presented here, the combination of two problems was chosen: measurement of CERC and the estimation of high confinement regimes. The effect of the first problem is localised at the plasma centre, the latter at the plasma edge. Both problems have been taken without weighting, so figure of merit for the optimisation was simply the sum of both *EUs*.

For the optimisation, the chords from the four-channel interferometer (the design result for the measurement of CERC) were taken as constant. Start configuration for the four new channels was the design result for the estimation of high confinement regimes without the extra beam (see figure 5.12, (a) and (b)).

As a result, it turns out that only one beam line of the new chords slightly changes in comparison to the start configuration. The new arrangement is shown in figure 5.17.

The sum of the *EUs* for the two sub-problems, which was to be optimised, turns out to be 18.01 *bit*: 7.37 *bit* for the measurement of CERC, and 10.64 *bit* for the high confinement problem. The *EU* of the sub-problems is larger than in the four-channel case, especially the information gain for the high confinement regime increases. This effect is explained by the fact, that now eight channels contribute to the estimation of the respective parameters of interest. It has to be mentioned that the differences in the summarised *EU* for different designs are small, the available beam line configurations do not provide significant differences in the expected information gain for the physical problems analysed here.

The method to design the eight-channel interferometer presented here turns out to be practicable, however, the two physical problems applied should be seen as an example. For

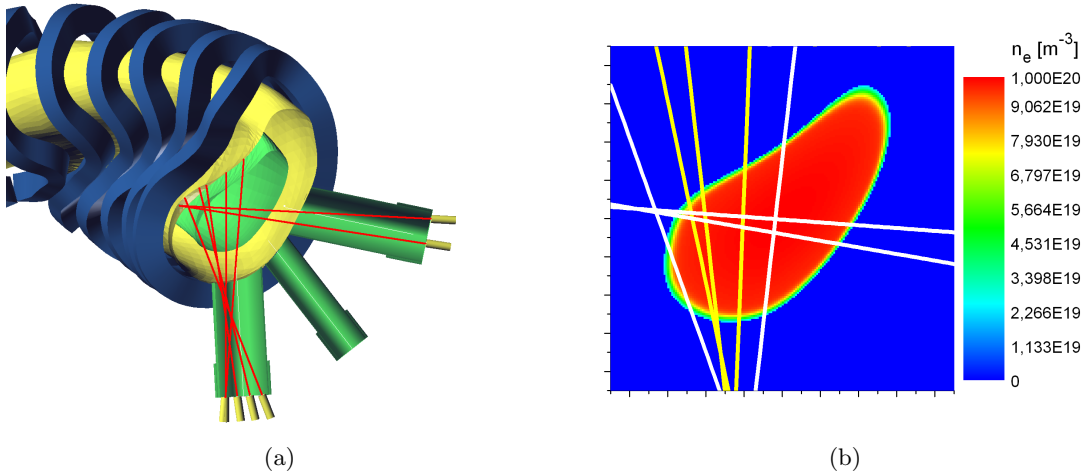


Figure 5.17: Design results for an eight-channel interferometer, optimised according to the measurement of CERC and high confinement regimes: With respect to the port system (a) and the density distribution (b). The yellow beam lines were taken as given (four-channel interferometer) and not changed for the design.

the final design of the eight-channel interferometer, first experimental results from W7-X may be available, so that the design problem can be adapted to the respective physical questions in the future.

### 5.3 Discussion

In this section three different designs considering three different physical problems for a four-channel interferometer at Wendelstein 7-X were presented. The results have been compared, and the solution for the measurement of CERC is proposed as the best beam line configuration for the analysed problems. It shows the highest sum of the  $EU$ s of the considered physical problems. In average, quite high  $EU$  values for every problem are achieved, the design is widely insensitive against the discussed changes of the considered magnetic configuration.

The highest loss of information gain for this design is found for the problem of high confinement regimes, which is mainly an effect of the plasma edge, whereas the CERC configuration proposed here consists of beam lines through the plasma centre. One may argue here that interferometry is not a good diagnostic for density effects at the plasma edge anyway - the nature of a line integrated measurement gives limitations for a spatial resolved analysis. Other diagnostics (lithium beam, reflectometry) may be more sufficient here, their results, however, can be combined with interferometric measurements using the approach of integrated data analysis [26]. For the future it also may be also possible to fan out single interferometer beams to achieve a better spacial resolution.

For all four-beam configurations it was beneficial to implement an extra beam, which was localised in a plane symmetric to the Thomson scattering diagnostic. The beam was crossing the plasma on a very long path. For all analysed physical problems, the designs including this beam showed the highest  $EU$  values. A possible explanation is its large signal-to-noise ratio: E.g., given the density distribution from equation (5.9) with  $\theta_1 = 1.0$ ,  $\theta_2 = 0.95$  and  $\theta_3 = 0.2$ , one obtains a SNR of 77 for the extra beam (for a measurement error  $\sigma = 0.1 \text{ rad}$ ). For the seven channels in figure 5.17, the best SNR is only 52. According to this, the extra beam may be also favourable for the purpose of

density control.

The proposal for a four-channel interferometer (the optimal configuration to measure the CERC effect) was chosen according to the analysed physical problems and conditions: the physical problems, magnetic configurations and the influence of the plasma edge outside the LCMS. It has to be kept in mind that the proposal results from these boundary conditions and is therefore only valid in this context.

The different physical problems were treated equally, but it would also be possible to introduce a weighting: If the optimal design is estimated by summarising the single *EUs*, a weighting factor on every summand can be introduced. The same principle can be applied for the eight-channel interferometer: If the design is optimised according to a sum of different *EUs*, here, also a weighting can be applied.

One analysed boundary condition is the influence of the magnetic configuration and the plasma outside the LCMS. It was found out that, for the analysed configurations, the optimal beam line designs were very similar. As a proposal resulting from this finding, it may be practicable to adjust the beam line configuration for every experiment by an appropriate optical arrangement, according to the respective magnetic configuration of the experiment.

In addition, some general findings about the method of Bayesian experimental design can be formulated: First, one can state that the signal-to-noise ratio is of strong importance for the design. A good SNR is always part of the optimum design, in one-beam case (section 4.2.3) as well as for the multi-channel interferometer, where the unconstrained design with long pathways through the plasma always yields a higher *EU*. A further argument for this finding is the fact that in the design process identical beam lines are sometimes preferred (but technically not realisable)- measuring the same effect on the same position would also reduce the SNR. The final argument is that a constant SNR leads to a constant *EU* distribution (see section 4.2.4)

The second finding is the fact that the optimal beam lines are always located in the regions where the density effect one wants to estimate is maximal. This was shown by comparing the unconstrained design results with the restricted configurations: The highest *EU* was given for beam lines for which changes of the parameters of interest led to a maximum change of the measured phase shift. In result, in every design the beam lines are arranged to cover the area of the plasma where the effects to be measured are localised.

As a remarkable fact, both findings - the dependency of the *EU* on the SNR and the measurement of the parameters of interest where they show a maximum effect on the data - are the *outcome* of *BED*, not a design criterion. This is different compared to former design methods, e.g., as used in [45].



## Chapter 6

# Experimental Planning using Bayesian Design

This chapter presents an additional application for the principle of Bayesian experimental design: the planning of experiments. In particular, it is discussed how already measured data sets from previous experiments can be implemented in the planning process. After introducing the methodology in the first part, the approach is applied to the problem of scaling laws for fusion devices in the second.

### 6.1 Motivation

Besides the optimisation of diagnostic devices, Bayesian experimental design can also be used for the planning of experiments: The approach of *data adaptive planning* (*DAP*) additionally offers the possibility to implement already measured data sets in the design process. With this, on the one hand side, one can quantify the importance of already measured data with respect to the whole data set. On the other hand, the best experimental conditions for a maximum information gain from the next measurement, given the existing experiments can be identified. Some examples may illustrate the usefulness:

- A typical application for this method is given for experiments with limited access to the measuring instrument, e.g. telescopes in astronomical research. Loredó [34] used the *DAP* approach for the estimation of the optimal observation time to measure the orbital parameters of an extra solar planet.
- One may think of experiments where the change of the experimental settings is technically impossible or would result in high costs, for instance the change of the geometrical parameters of fusion experiments (e.g. the major radius  $R$ ). Here, also an estimation of the best configuration can be done by *DAP* taking into account the findings from previous experiments.
- The importance of already measured data points with respect to the complete data set can be expressed by their utility. So, by comparing the utility of different data points, the influence of physical parameters and diagnostic variables, like the measurement error, on the significance of the data points can be analysed.

Requirements for the application of *DAP* are a physical model connecting the model parameters  $\alpha$  with the data point  $D$ :  $D = f(\alpha)$ , as well as the corresponding error statistics. Both are combined in the likelihood. Also, a database with already measured data  $\mathbf{d}$  is needed.

## 6.2 The principle of data adaptive planning

### 6.2.1 General approach for a linear problem

The planning of experiments is based on the same design approach as described in chapter 2: Maximisation of the expected information gain. This information gain is again expressed by the Kullback–Leibler distance (equation 2.39).

Suppose a database containing existing data  $\mathbf{d}$ , where  $d_i$  was measured at experimental configurations  $\mathbf{X}_i$ , with  $X_{i,1} \dots X_{i,p}$  as the  $p$  configuration variables for the data point  $i$ . The measurement error of  $d_i$  is given with  $s_i$ ,  $q$  different data points may exist. The parameters of interest are named  $\boldsymbol{\alpha} : \alpha_1 \dots \alpha_p$ .

For this work, a linear model function shall be analysed:

$$\mathbf{d} = \mathbf{X} \cdot \boldsymbol{\alpha} \quad (6.1)$$

and

$$D = \boldsymbol{\xi}^T \cdot \boldsymbol{\alpha}, \quad (6.2)$$

respectively. Now, the Kullback-Leibler distance for the new datum  $D$  is given with

$$U_{KL}(D, \mathbf{d}, \boldsymbol{\xi}) = \int d\boldsymbol{\alpha} p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi}) \ln \left[ \frac{p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi})}{p(\boldsymbol{\alpha} | \mathbf{d})} \right]. \quad (6.3)$$

Marginalisation over the expected data space for  $D$  leads to the Expected Utility:

$$EU(\mathbf{d}, \boldsymbol{\xi}) = \int dD p(D | \mathbf{d}, \boldsymbol{\xi}) \cdot U(D, \mathbf{d}, \boldsymbol{\xi}), \quad (6.4)$$

with the evidence *PDF* given by

$$\begin{aligned} p(D | \mathbf{d}, \boldsymbol{\xi}) &= \int d\tilde{\boldsymbol{\alpha}} p(D, \tilde{\boldsymbol{\alpha}} | \mathbf{d}, \boldsymbol{\xi}) \\ &= \int d\tilde{\boldsymbol{\alpha}} p(\tilde{\boldsymbol{\alpha}} | \mathbf{d}) p(D | \tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi}). \end{aligned}$$

Here, the term  $p(D | \tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi})$  does not explicitly depend on  $\mathbf{d}$  because of (6.1). The parameters of interest are written as  $\tilde{\boldsymbol{\alpha}}$  to distinguish between both integrations over the parameter space when (6.3) and (6.4) are combined:

$$\begin{aligned} EU(\mathbf{d}, \boldsymbol{\xi}) &= \int d\tilde{\boldsymbol{\alpha}} p(\tilde{\boldsymbol{\alpha}} | \mathbf{d}) \cdot \int dD p(D | \tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi}) \\ &\quad \cdot \int d\boldsymbol{\alpha} p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi}) \ln \left[ \frac{p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi})}{p(\boldsymbol{\alpha} | \mathbf{d})} \right]. \end{aligned} \quad (6.5)$$

The dependencies on  $\mathbf{s}$ ,  $\sigma$  and  $\mathbf{X}$  are not displayed for purposes of clarity. The individual terms are described in detail next.

Using a Gaussian error distribution for the existing data  $\mathbf{d}$  and assuming a flat prior in  $\boldsymbol{\alpha}$ , one obtains for the old posterior

$$\begin{aligned} p(\boldsymbol{\alpha} | \mathbf{d}) &\propto p(\mathbf{d} | \boldsymbol{\alpha}) \cdot p(\boldsymbol{\alpha}) \quad (\text{Bayes' theorem}) \\ &\propto \exp \left[ -\frac{1}{2} (\mathbf{d} - \mathbf{X}\boldsymbol{\alpha})^T \mathbf{C} (\mathbf{d} - \mathbf{X}\boldsymbol{\alpha}) \right], \end{aligned} \quad (6.6)$$

where the matrix  $\mathbf{C}$  contains the measurement errors:  $C_{ii} = 1/s_i^2$ . The non-diagonal elements are zero, the data sets are assumed as not correlated.

Taking into account the new datum  $D$  with the same error statistic as for (6.6), an extended likelihood can be formed:

$$\begin{aligned} p(D, \mathbf{d} | \boldsymbol{\alpha}, \boldsymbol{\xi}) &= p(\mathbf{d} | \boldsymbol{\alpha}) \cdot p(D | \boldsymbol{\alpha}, \boldsymbol{\xi}) \\ &= \frac{1}{(2\pi)^{q/2} \prod_{i=1}^q s_i} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \\ &\quad \cdot \exp \left[ -\frac{1}{2} (\mathbf{d} - \mathbf{X}\boldsymbol{\alpha})^T \mathbf{C} (\mathbf{d} - \mathbf{X}\boldsymbol{\alpha}) - \frac{1}{2\sigma^2} (D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2 \right] \end{aligned} \quad (6.7)$$

This is directly proportional to the posterior of  $\boldsymbol{\alpha}$ :

$$p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi}) \propto p(D, \mathbf{d} | \boldsymbol{\alpha}, \boldsymbol{\xi}) \cdot p(\boldsymbol{\alpha}). \quad (6.8)$$

Finally, the new datum shall be described by the model:

$$p(D | \tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi}) = \delta(D - \boldsymbol{\xi}^T \tilde{\boldsymbol{\alpha}}). \quad (6.9)$$

The  $EU$  in equation (6.5) can now be computed analytically, the calculation is somewhat extensive and is given in appendix C. The final Expected Utility is obtained with

$$EU(\mathbf{d}, \boldsymbol{\xi}) = \frac{1}{2} \left[ \log(1 + G) - \frac{G}{(1 + G)^2} \right]; \quad G = \frac{\boldsymbol{\xi}^T (\mathbf{X}^T \mathbf{C} \mathbf{X})^{-1} \boldsymbol{\xi}}{\sigma^2}. \quad (6.10)$$

The result does not explicitly depend on the data values of the existing data  $\mathbf{d}$ , since the model function (6.1) is linear in  $\boldsymbol{\alpha}$ , it is only related to the error of the former measurements given by  $\mathbf{C}$ , the experimental configurations of these measurements  $\mathbf{X}$ , the error  $\sigma$  of the future datum and its experimental configuration  $\boldsymbol{\xi}$ . This is reasonable because the  $EU$  does not give any prediction for  $\boldsymbol{\alpha}$  (For this,  $p(\boldsymbol{\alpha} | \mathbf{d}, D, \boldsymbol{\xi})$  would be needed to be calculated, which explicitly depends on  $\mathbf{d}$ , see eq. (6.7)). The  $EU$  only gives the expected information gain with respect to the experimental configuration for the next measurement,  $\boldsymbol{\xi}$ . Therefore, only two information are needed: How dense the configuration space has been scanned in the former experiments (this is given by  $\mathbf{X}$ ), and how exact these scans have been made (expressed by  $\mathbf{C}$ ).

Furthermore, the term  $G$  is proportional to  $1/\sigma^2$ , which means that the  $EU$  decreases for an increasing measurement error of the new datum. In other words, the utility of a new datum is lower if its error is large, which can be intuitively understood.

### 6.2.2 A one-dimensional example

To illustrate the working principle, a simple example describing the estimation of the slope  $\alpha$  of a straight line passing the origin shall be analysed. For a one-dimensional problem with

$$\mathbf{d} = \mathbf{x} \cdot \alpha \quad \text{and} \quad D = \xi \cdot \alpha \quad (6.11)$$

one obtains from equation (6.10)

$$EU(\mathbf{d}, \xi) = \frac{1}{2} \left[ \log \left( 1 + \frac{\xi^2 s^2}{\mathbf{x}^2 \sigma^2} \right) - \frac{\frac{\xi^2 s^2}{\mathbf{x}^2 \sigma^2}}{\left( 1 + \frac{\xi^2 s^2}{\mathbf{x}^2 \sigma^2} \right)^2} \right]. \quad (6.12)$$

The  $EU$  is large for high values of  $\xi$ , which means that the best way to estimate the slope is to measure as far away as possible from the origin.



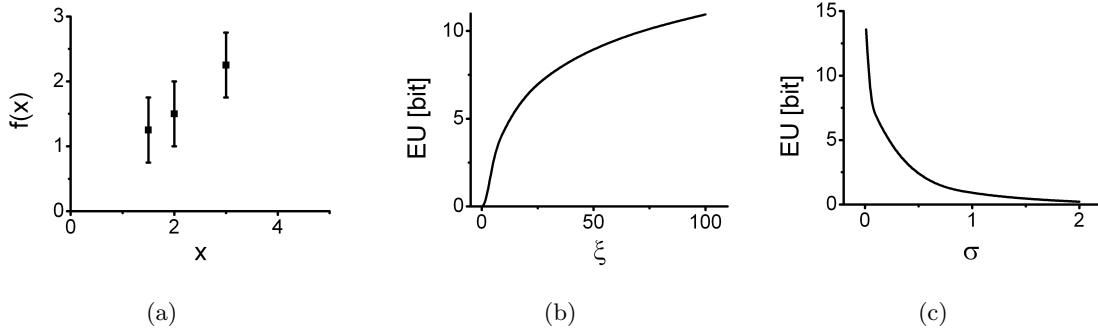


Figure 6.1: Estimation of the slope of a straight line: Given three data points (a), the distribution of the  $EU$  is calculated for the next measurement with respect to  $\xi$  for  $\sigma = 0.5$  (b) and with respect to the measurement error  $\sigma$  at  $\xi = 5.0$  (c).

As an example, one may consider a function  $y = f(x) = \alpha \cdot x$  and three given (already measured) data points at  $\mathbf{x} = \{1.5, 2.0, 3.0\}$ , their errors are  $s_1 = s_2 = s_3 = 0.5$  (figure 6.1 (a)). Now, the  $EU$  is calculated for a new datum, where  $\sigma = 0.5$  is assumed, too. For different  $\xi$  one obtains, e.g.,

$$\begin{aligned} EU(\xi = 0.5) &= 4.34 \text{ bit}, \\ EU(\xi = 1.0) &= 6.35 \text{ bit}, \\ EU(\xi = 5.0) &= 10.99 \text{ bit}. \end{aligned}$$

The complete  $EU$  distribution is given in figure 6.1 (b). The expected information gain shows a strong rise for increasing values of  $\xi$  near the origin. However, the increase of the  $EU$  becomes smaller at larger values of  $\xi$ . A shift from  $\xi = 0.5$  to  $\xi = 1.0$  has therefore a stronger impact on the information gain than a shift from, e.g., 90 to 90.5, in other words, the  $EU$  is sensitive to the relative change  $\Delta\xi/\xi$ .

Furthermore, equation (6.12) also allows to study the influence of the error statistics. As already mentioned, the  $EU$  decreases for an increasing measurement error  $\sigma$ , and vice versa. In figure 6.1 (c)  $\sigma$  is varied for a new measurement at  $\xi = 5.0$ . The  $EU$  is very sensitive for small values of  $\sigma$ , a small increase of the error leads to a large descent of the expected information gain. For large  $\sigma$  this effect is less distinct, but the  $EU$  is always decreasing if the error is increased.

### 6.3 Scaling Laws for Fusion Devices

Scaling laws for fusion devices are a predestined field of application for *DAP*: First, databases containing several hundred data sets exist [77], which have to be assessed with respect to the importance of the respective data points. Secondly, it is possible to plan experimental campaigns at running experiments by identifying the most useful operational conditions. And finally, the scalings are also used to plan future experiments, so *DAP* can be applied, e.g., to find the optimal geometrical configuration.

In the special case of a logarithmised scaling law the *DAP* approach for a multi-dimensional linear problem can be used.

#### 6.3.1 Background

Scaling laws connect plasma and machine parameters with quantities relevant for plasma and energy confinement. They are widely used for

- inter-machine comparison of different discharges and
- prediction of the confinement performance of future experiments.

The first point is motivated by the limitations in the parameter range for a single experiment (e.g. restrictions in geometrical variations), this can be overcome by a comparison of different machines. The second application arises from the fact that the interdependencies between the experimental parameters (like density  $n$ , temperature  $T$ , the magnetic field  $B$ , effective minor radius  $a$ , major radius  $R$  and total heating power  $P_{tot}$ ) and the quantities describing the plasma confinement (the energy confinement time  $\tau_E$  or the confinement energy  $W$ ) are not fully understood, which makes theoretical predictions for future experiments difficult. Scaling laws give a semi-empirical approach to connect the different quantities, involving data sets from different experiments.

A typical scaling is given as a power law of the form

$$\tau_E = \prod_j x_j^{\alpha_j}, \quad (6.13)$$

where the  $x_j$  are the plasma and machine parameters (configuration variables), and the  $\alpha_j$  are the scaling exponents describing the dependencies of the quantity of interest (here:  $\tau_E$ ). In most cases these scaling parameters have been found by regression procedures of available data sets.

In addition to pure empirical studies, it is always the goal of scaling laws to connect the experimental findings with justification based on physical principles. A remarkable attempt was made by Connor and Taylor [78], who connected the plasma regime with the scaling law resulting in different numbers of scaling terms (i.e. different numbers of scaling parameters). Therefore, they differentiated between collisional and collisionless cases, high- and low- $\beta$  plasmas and ideal and resistive MHD models. The theoretical background of the approach was simply the consideration of charge neutrality, the Vlasov equation with and without collision term, respectively, and the Maxwell equations. The findings of Connor and Taylor were employed in high- $\beta$  studies at W7-AS [79].

### 6.3.2 Energy confinement time

#### Tokamaks

The first approaches on scaling laws have been made for Tokamaks and were mostly of empirical or semi-empirical nature. An overview above the early findings is given, e.g., by Goldston [80]. In this work, previous scaling attempts, mainly consisting of parameter scans on different machines, are reviewed and connected. In result, two scaling laws were found for ohmic and auxiliary heated plasmas by combining the data from several experiments. A comparison with the constraints of Connor and Taylor showed good agreement within the error bars. Furthermore, Goldston pointed out the necessity for multi-dimensional parameter scans.

It was found to be valuable to collect the data sets of different experiments in a confinement database to put the scaling laws on a much broader basis [81]. In addition, parameter scans on minor and major radius are not applicable on single machines. One approach was done by Yushmanov et al. combining data sets from eight major tokamak experiments [82]. Here, the plasma current  $I_p$ , the elongation of the plasma cross section  $k$  and the isotopic mass  $M$  were also taken into account as plasma parameters. To overcome geometrical differences of the several machine configurations, the shape index  $f_s = f(R, a)$  was introduced.

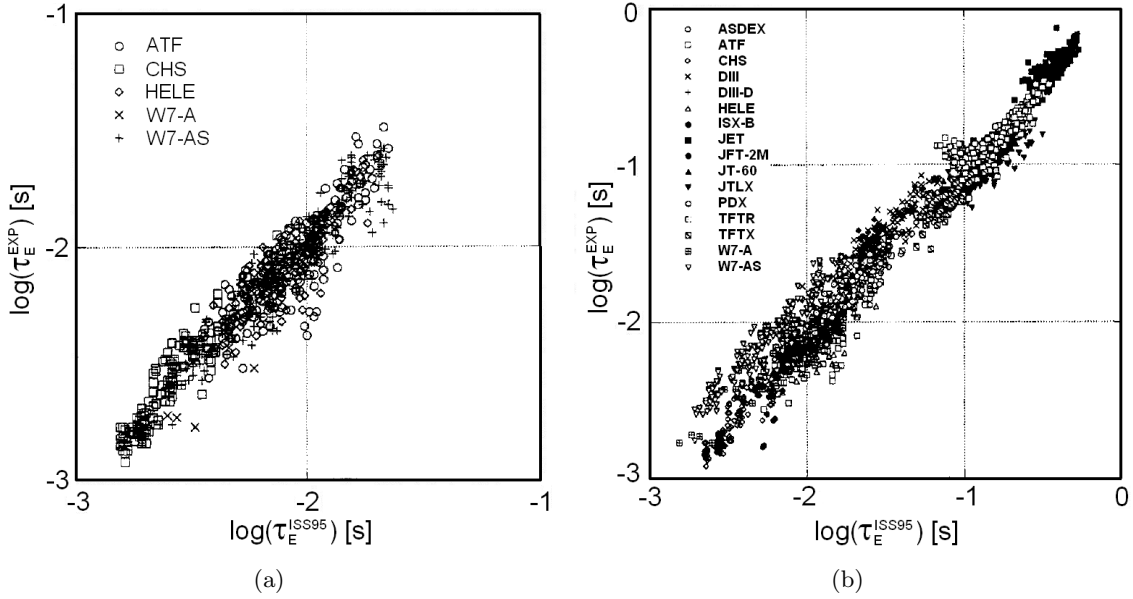


Figure 6.2: Comparison of the experimental confinement time  $\tau_E^{EXP}$  with the result from ISS95 scaling  $\tau_E^{ISS95}$  (from [86]): Only for stellarator experiments (a) and including tokamaks (b).

## Stellarators

Similar to the confinement databases for tokamaks (see, e.g., [83, 84]) the need for such a database and confinement relations turned out to be important for stellarators, too. Whereas earlier scaling approaches used Tokamak scalings like, e.g. the semi-empirical expression from Lackner and Gottardi [85], Stroth et al. introduced an international database for stellarators and a energy confinement time scaling law [86]. Five stellarators contribute to the database: ATF, CHS, Heliotron-E, W7-A and W7-AS. The  $\tau_E$  scaling was related to the minor radius  $a$ , major radius  $R$ , total heating power  $P_{tot}$ , line averaged density  $\bar{n}$ , the magnetic field strength  $B$  and the rotational transform  $t$ , typically given by the rotational transform at  $r_{eff} = \frac{2}{3}a$ ,  $t_{2/3}$ .

The scaling law was then in the form of

$$\tau_E = 10^{\alpha_x} 10^{s\alpha_s} a^{\alpha_r} R^{\alpha_R} P_{tot}^{\alpha_P} \bar{n}^{\alpha_n} B^{\alpha_B} t_{2/3}^{\alpha_t}. \quad (6.14)$$

Here,  $s$  is an additional parameter introduced to differentiate between heliotrons and torsatrons like ATF, CHS and Heliotron-E ( $s = 1$ ), and the shearless stellarators W7-A and W7-AS ( $s = 0$ ).

In result, the International Stellarator Scaling ISS95 was proposed:

$$\tau_E^{ISS95} = 0.079 a^{2.21} R^{0.65} P_{tot}^{-0.59} \bar{n}^{-0.51} B^{0.83} t_{2/3}^{0.4} \quad (6.15)$$

Additionally, this scaling was compared with the tokamak L-mode scaling by Lackner and Gottardi [85]. Here, the plasma current  $I_p$ , which does not occur in stellarators, was replaced by an expression containing the rotational transform  $t$  and an elliptical elongation  $\kappa$  as a geometrical factor. It turned out that the Tokamak data were also well described by the ISS95 scaling (see figure 6.2).

In 2004 the stellarator database was extended by new experiments, which led to a new scaling expression ISS04 ([77], see figure 6.3). The experiments Heliotron-J, HSX, LHD

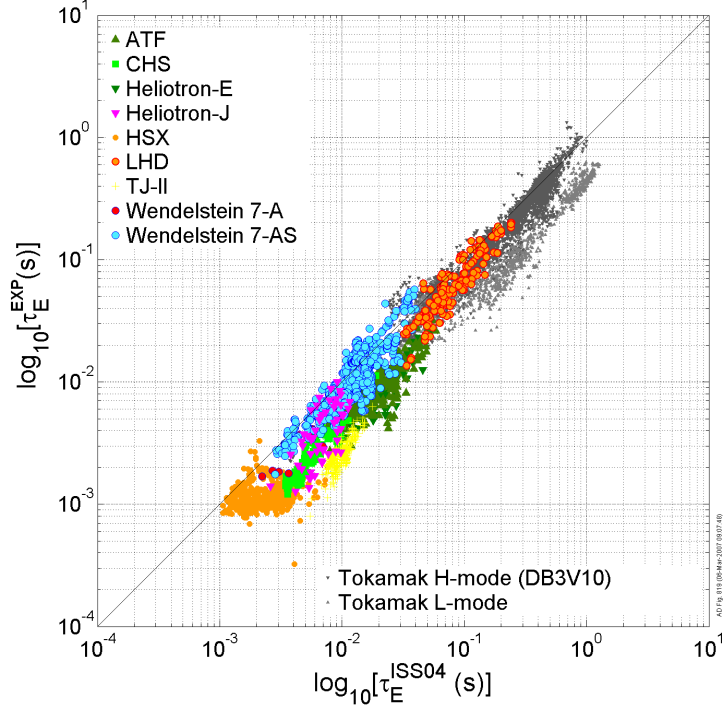


Figure 6.3: Results from ISS04 stellarator scaling (from [87]): Experimental result  $\tau_E^{EXP}$  vs. scaling expression  $\tau_E^{ISS04}$ . Data sets from tokamak experiments are shown for comparison.

and TJ-II were additionally implemented in the database, new data sets were also added from the other machines.

Instead of the shear parameter  $s$  the ISS04 scaling uses a single renormalisation factor  $f_{\tau_{en}}$  for each configuration. It does not make a distinction between experiments, but between experimental configurations. For instance, for W7-AS three different factors were applied ( $t_{2/3} < 0.48$ ,  $t_{2/3} \geq 0.48$  and high- $\beta$  case). The final scaling turned out to be

$$\tau_E^{ISS04} = 0.134 a^{2.28} R^{0.64} P_{tot}^{-0.61} \bar{n}^{0.54} B^{0.84} t_{2/3}^{0.41}. \quad (6.16)$$

Here, the dependency of the confinement time on the density and the total heating power as found in the ISS95 scaling was confirmed.

### 6.3.3 Confinement energy

Since the energy confinement time  $\tau_E$  cannot be measured directly, the plasma energy content  $W$  should be used in data analysis instead. For given electron and ion density and temperature profiles  $W$  is obtained by [53]

$$W = \frac{3}{2} \int (n_e T_e + n_i T_i) dV. \quad (6.17)$$

However, this method is mostly used as a cross check because  $T$  and  $n$  profiles are often not available or uncertain. In practice the energy content of the plasma is measured, e.g., by diamagnetic measurements [38].

The relation between the energy content and the energy confinement time are given for stationary conditions via the net heating power  $P_{net}$  with

$$\tau_e = \frac{W}{P_{net}}, \quad (6.18)$$

in case of negligible loss terms  $P_{net}$  is replaced by the total absorbed heating power  $P_{tot}$  [53]. Therefore, the scaling laws for  $\tau_E$  have only be reduced by a factor  $P_{tot}$  to achieve the energy content scaling of the form

$$W = e^{\alpha_c} \bar{n}^{\alpha_n} B^{\alpha_B} P_{tot}^{\alpha_P} a^{\alpha_r}. \quad (6.19)$$

An early attempt for the use of  $W$  instead of  $\tau_E$  can be found by Rebut et al. [88]. However, major work was done by Preuss et al. [24, 25, 79, 89] evaluating data from W7-AS with respect to the aforementioned constraints of Connor and Taylor [78]. It was shown that the data sets from W7-AS are best explained by their respective physical model, expressed in the parameter constraints of Connor and Taylor. This finally led to the conclusion that different physical situations, such as variations in the rotational transform  $\iota$ , have to be described with different scaling laws.

## 6.4 Data adaptive planning for scaling laws

The principle of Data adaptive planning shall now be implemented for a data set of the W7-AS stellarator. For this, 153 data sets with a rotational transform of  $\iota = 1/3$  from the ISS95 database were extracted and analysed. The impact of additional data is also pointed out.

### 6.4.1 DAP for W7-AS measurements

As mentioned before, the parameters of interest are now the scaling parameters  $\alpha$  of equation (6.19). The design parameters, however, are the quantities the energy content depends on: minor radius  $a$ , magnetic field  $B$  and the heating power  $P_{tot}$  (simply  $P$  in the following). Note, that - in contrast to the previous chapters - the line averaged density  $\bar{n}$  is also a design parameter and therefore treated differently.

Goal of *DAP* is to find the design parameter set  $\{a, B, \bar{n}, P\}$  for which a new measurement will provide a maximum of information about the scaling parameters  $\alpha$  (with  $\alpha = \{\alpha_c, \alpha_r, \alpha_B, \alpha_n, \alpha_P\}$ ). A set of existing data  $\mathbf{d}$  is thereby considered.

The power law approach enables one to linearise the problem by taking the logarithm of (6.19):

$$\ln W = \alpha_c + \alpha_r \ln a + \alpha_P \ln P + \alpha_n \ln \bar{n} + \alpha_B \ln B. \quad (6.20)$$

The design parameters are now given by  $\xi = \{1, \ln a, \ln P, \ln \bar{n}, \ln B\}$ . The existing data  $\mathbf{d}$  is a set of logarithmised values of the energy content  $\{\ln W_i\}$ . For further variables, the following notation is used in this section:

- $q$  number of existing data points
- $p$  number of scaling parameters
- $\mathbf{d}$  existing data
- $\mathbf{s}$  errors of the existing data
- $\mathbf{X}$  experimental configuration of the existing data set:  
 $\mathbf{X}_i = \{\ln a_i, \ln B_i, \ln \bar{n}_i, \ln P_i\}$
- $D$  new datum (to be measured)
- $\sigma$  error of the new datum
- $\xi$  experimental configuration of the new datum  $\rightarrow$  design parameters
- $\alpha$  scaling parameters  $\rightarrow$  parameters of interest

The error  $s_i$  is calculated from simple error propagation law according to equation (6.20) with

$$s_i^2 = s_{i,\ln W}^2 + \alpha_r^2 s_{i,\ln r}^2 + \alpha_P^2 s_{i,\ln P}^2 + \alpha_n^2 s_{i,\ln n}^2 + \alpha_B^2 s_{i,\ln B}^2. \quad (6.21)$$

The single errors  $s_{i,\ln W}$ ,  $s_{i,\ln r}$ ,  $s_{i,\ln P}$ ,  $s_{i,\ln n}$  and  $s_{i,\ln B}$  are the errors of the logarithmised coordinates in the parameter space and can be extracted from the database. The coefficients  $\alpha_i$  are the scaling exponents. For the ISS95 database, the results from Lackner and Gottardi [85] were used as a prior guess.

Given these definitions, the expression for the Expected Utility (6.10) can now be applied.

#### 6.4.2 The W7-AS data set

The ISS04 database provides values and errors for every parameter  $a$ ,  $P$ ,  $\bar{n}$ ,  $B$  and  $W$ . In the following, a distinction is made between the measured value of  $W$ ,  $W_{exp}$ , and the theoretical value  $W_{theo}$ , calculated by the power scaling law. The error of  $W_{exp}$  is given by the measurement error of the diamagnetic energy  $W_{diag}$  and can be found in the database. For  $W_{theo}$  the error is computed by error propagation according to equation (6.21).

The scaling law for the 153 data sets from W7-AS is calculated by regression with

$$\ln W_{theo} = -0.831 + 2.264 \ln a + 0.452 \ln P + 0.440 \ln \bar{n} + 0.618 \ln B \quad (6.22)$$

with the following uncertainties:

$$\begin{aligned} \alpha_c &= -0.831 \pm 0.082 \\ \alpha_r &= 2.264 \pm 0.093 \\ \alpha_P &= 0.452 \pm 0.035 \\ \alpha_n &= 0.440 \pm 0.022 \\ \alpha_B &= 0.618 \pm 0.049. \end{aligned}$$

These values are different from the findings of ISS04 (equation (6.16)), because only a subset of the data (with  $\iota = 1/3$ ) was taken into account, mainly consisting of measurements at low densities. Using these parameters,  $W_{theo}$  can be calculated and compared with the experimental value  $W_{exp}$ .

*DAP* allows one to determine the utility of individual measurements from a given data set. For this purpose, the respective measurement is extracted from the data set, and its utility is calculated with respect to the remaining  $(q - 1)$  data points by equation (6.10): Here,  $\{\xi, \sigma\}$  is the experimental configuration of the extracted datum, the other data sets are described by  $\{\mathbf{X}, \mathbf{s}\}$ . Note, that the *EU* calculated for different individual measurements is not comparable quantitatively, because the reference data set  $\{\mathbf{X}, \mathbf{s}\}$  is different in every case.

Figure 6.4 shows the  $W_{theo}$  plot for all 153 data points. The utility of every single datum is expressed in the colour code. While most data points show a small utility, mainly for the data at higher confinement energies higher *EU* values are calculated. However, exceptions for lower  $W_{theo}$  can be found, too.

Since equation (6.10) incorporates all reference data, the Expected Utility is a context sensitive quantity, it depends on what has been measured beforehand. Therefore, it may therefore be instructive to analyse the reason for the different *EU* values of the data points in figure 6.4.

W7-AS was operated at two values for the magnetic field, at  $B \approx 2.54 T$  and at  $B \approx 1.24 T$ . In figure 6.5 (a) the *EU* is plotted over the magnetic field. Whereas the maximum *EU* value is given for the smaller field, in general high values of the *EU* are found for both fields, so that no general conclusion can be made.

A similar picture is given for the minor radius  $a$ : Most data sets are given for radii around  $a \approx 0.175 m$  and  $a \approx 0.12 m$  (figure 6.5 (b)). Again, no clear assumption according to the relation between  $a$  and the Expected Utility can be made.

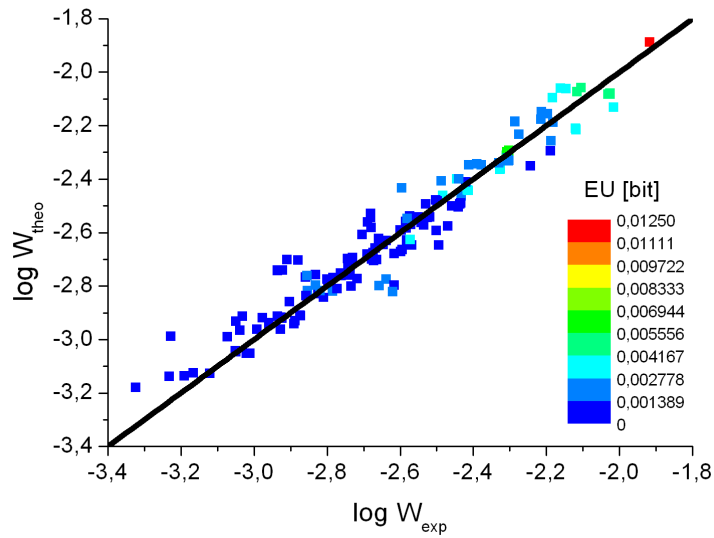


Figure 6.4: Comparison of  $W_{exp}$  and  $W_{theo}$  for 153 data sets from W7-AS. The colour code expresses the utility of the respective data point. For comparison, the function  $W_{theo} = W_{exp}$  is also displayed (black line).

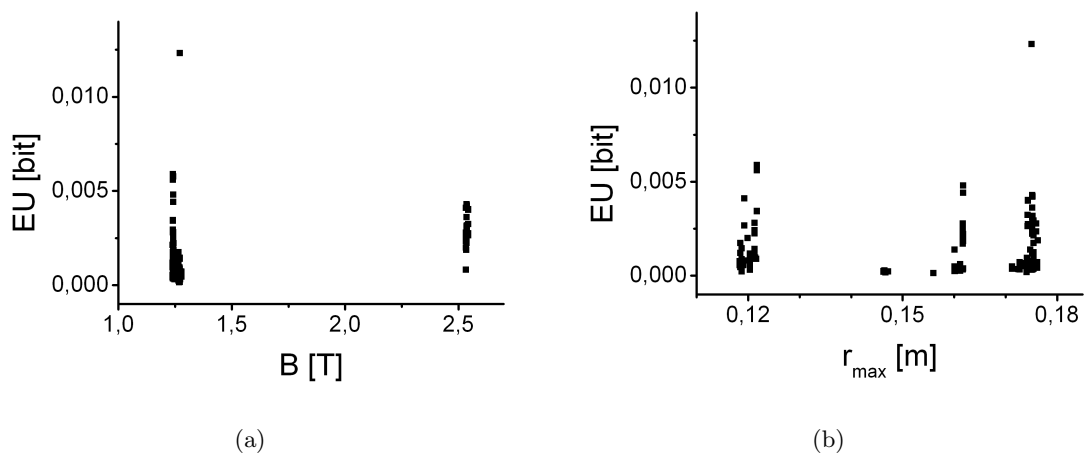


Figure 6.5: Dependence of the  $EU$  on the magnetic field (a) and the minor radius  $a$  (b) for the 153 W7-AS data sets.

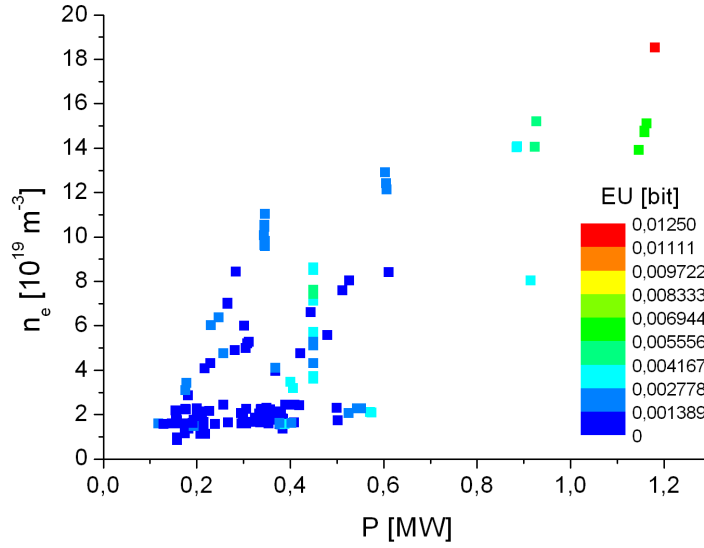


Figure 6.6: Power - density dependence of the utility, the  $EU$  is decoded in the colour scheme.

Most variation in the parameter space can be seen for the line-averaged density  $\bar{n}$  and the heating power  $P$ . To show possible interdependencies, the  $EU$  is displayed in relation to both parameters in figure 6.6. High utilities are found in regions of high heating power and high densities. In these regions of the  $P - \bar{n}$  plane only a few data points are available. The most separated point shows the highest utility.

At this point one can conclude that data sets in regions of the configuration parameters space with a low sample rate are most valuable. A new measurement should therefore lie in a parameter range not yet covered by previous experiments. This assumption can be proven by calculating the  $EU$  of a possible new data point.

### 6.4.3 Experimental planning: The $EU$ for a new data point

For the planning of a future experiment the expected utility of the new data point has to be calculated. In case of the W7-AS data sets it makes sense to calculate the  $EU$  of all possible data points in the  $P - \bar{n}$  plane. The principle is illustrated in figure 6.7: Here, the datum with the highest utility from the database was scanned over the complete plane, the  $EU$  is illustrated as background colour. For comparison, the other data are also displayed.

The  $EU$  for this "new" data point shows the lowest value in regions which are already covered by the other measurements, whereas the uncovered regions show a high utility.

Next, a completely new data point shall be assumed: Given all 153 data points, what would be the best parameter combination for a 154th measurement? Again, the  $P - \bar{n}$  plane was sampled, the result is shown in figure 6.8. For the fictional new data point, the parameter settings for the error, minor radius and magnetic field have been taken from the data point with the highest  $EU$  from the given data set.

The result confirms the conclusions from the previous example: The regions with a high  $EU$  are restricted to areas where no previous measurement exists. By comparison with figure 6.7 one also finds, that the area of smaller  $EU$  increases slightly around the data point with the highest  $EU$ . This shows the impact of this measurement on the  $EU$



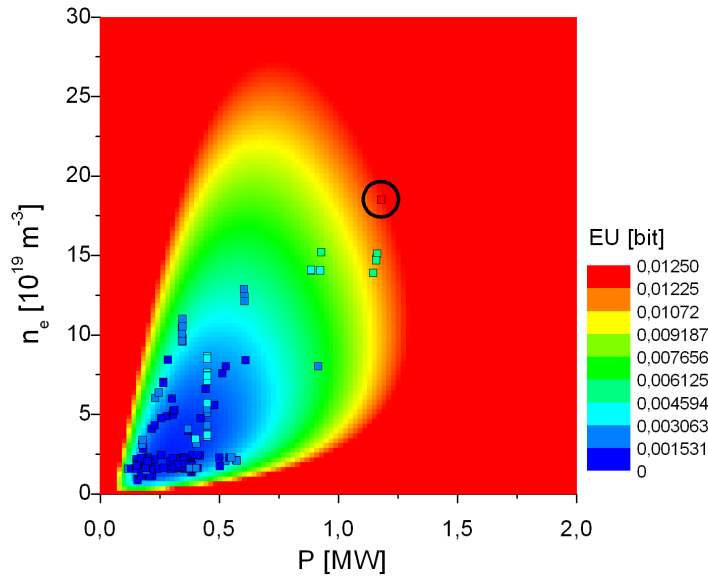


Figure 6.7:  $EU$  distribution, sampled over the  $P - \bar{n}$  plane, the point with the highest utility (encircled) of the W7-AS data set was used for this sampling. Therefore, its colour coincides with the background. All other data sets are shown for comparison, too.

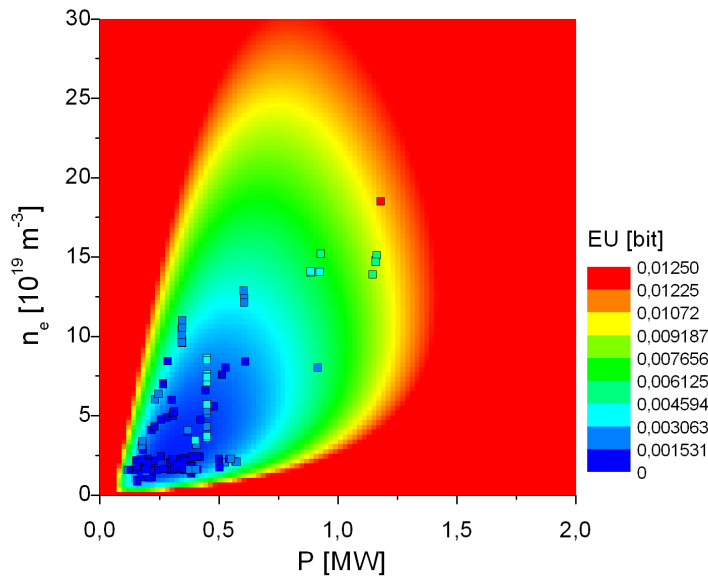


Figure 6.8: The same plot as figure 6.7, but now for a fictive 154th data point, given the 153 data sets from W7-AS. These sets are also shown for comparison.

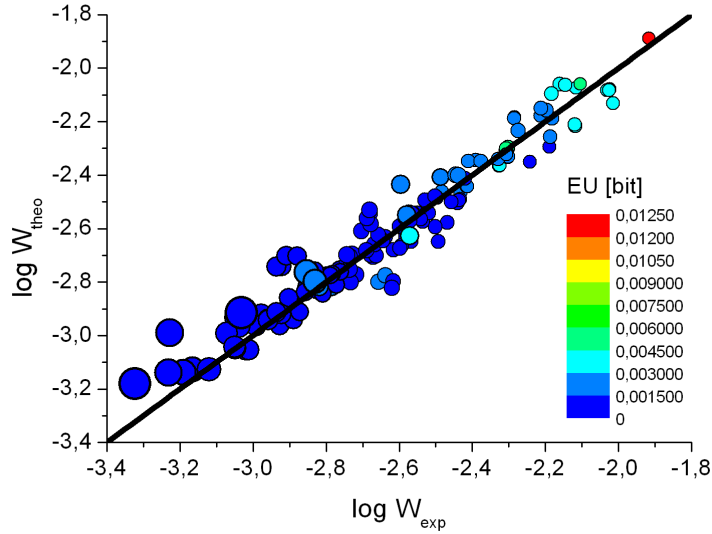


Figure 6.9: Comparison of  $W_{exp}$  and  $W_{theo}$  for 153 data sets from W7-AS, the  $EU$  is colour coded. The symbol size is given by the error of  $\log W_{theo}$ , scaled by a factor of 200. The datum with the largest  $EU$  (red) shows a very small error.

distribution: Because this data point is now taken into account, measurements with similar  $P - \bar{n}$  settings become less informative.

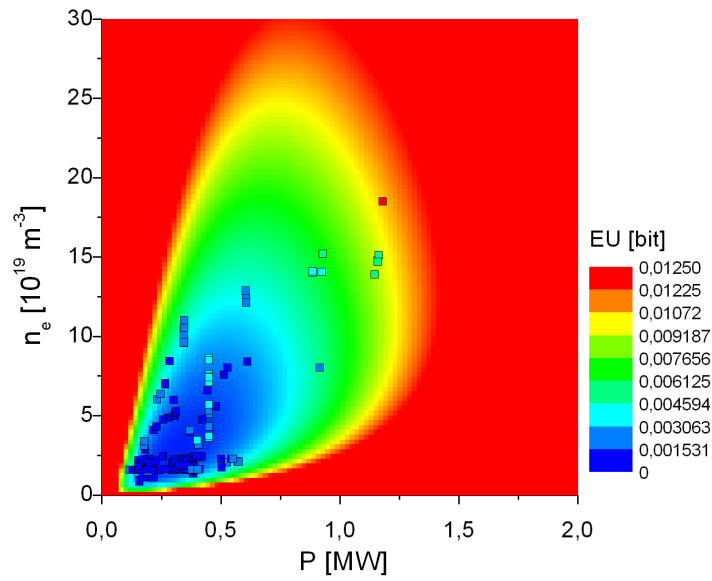
#### 6.4.4 Influence of the measurement error

As already shown in section 4.2.4, the design of an experiment may strongly depend on the measurement error (figure 6.9), its influence on experimental planning shall therefore be discussed, too. Two cases will be analysed: First, the change of the  $EU$  if the same error is considered for all data points, and secondly, the influence of an increasing or decreasing error for a possible new measurement.

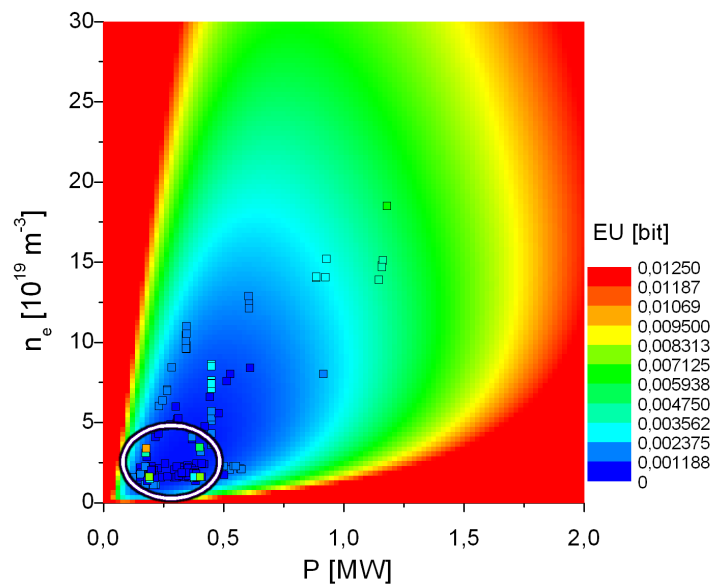
For the first case, the diagonal elements of the error matrix  $\mathbf{C}$  are calculated from the error  $s_{eumax} = 0.057$  of the most informative point (the encircled one in figure 6.7):  $C_{ii} = 1/s_{eumax}^2 \forall i$ . The same value is applied for the error of the new datum,  $\sigma$ . The new  $EU$  distribution can be seen in figure 6.10, the former result, considering the true measurement error, is given for comparison. In the plots, all 153 data points are displayed, also the  $EU$  distribution in the  $P - \bar{n}$  plane for a possible additional measurement.

First it has to be noticed, that the utility for the given data points changes if the measurement error is set as equal. For some data the utility decreases, especially in the case of the data point with maximum  $EU$  in figure 6.10 (a). The effect can be explained as follows: The measurement error of this datum is relatively small (see figure 6.9), therefore it is more informative compared with the other data points. This advantage is cancelled if the error is equal for all measurements.

The largest  $EU$  values are now found for four data points with relative low values of  $P$  and  $\bar{n}$  (see the encircled region in figure 6.10 (b)). It turns out that these points are the only measurements from the database at  $a = 0.12 m$  and  $B = 2.54 T$ . Compared with the equalised error  $s_{eumax}$  assumed for this analysis, the true errors of these points are about two times larger: 0.110, 0.138, 0.081 and 0.105, respectively. This leads to the assumption, that the value of the  $EU$  is influenced by the measurement error as well as



(a)



(b)

Figure 6.10: The influence of the measurement error on the  $EU$  for scaling laws:  $EU$  distribution for a fictive 154th measurement, the error was taken (a) as measured, (b) equal for every data point. The highest  $EU$  values in the case of equal errors are given for relatively low values of  $P$  and  $\bar{n}$  (encircled area).

by the configuration variables: If all errors are set as equal, the data points with rare experimental configurations become the most informative.

The distribution of the  $EU$  for the new measurement (background colour in figure 6.10 (b)) also depends on the error: For equal errors, the less informative part in the  $P - \bar{n}$  plane becomes larger. Again, the reason is given by the equal error: In figure 6.10 (a) for the new measurement the error from the most informative point was taken:  $\sigma = s_{eumax}$ . Therefore, in comparison to the other data, the new measurement becomes very informative, which can be seen in the large areas with a high  $EU$ . Now, in figure 6.10 (b), this advantage is removed, because all data points have the same measurement error  $s_{eumax}$ . So, only the experimental configuration of the new datum is important, leading to high  $EU$  values only at unusual experimental conditions (e.g. large values for  $P$  and  $\bar{n}$ ).

If only the error for the new measurement,  $\sigma$ , is varied, the  $EU$  distribution in the  $P - \bar{n}$  plane changes dramatically: In figure 6.11  $\sigma$  was halved (a) and doubled (b). Here,  $\sigma$  was taken as the average error of all measurements from the database. When  $\sigma$  is halved, the  $EU$  becomes large for nearly every combination of  $P$  and  $\bar{n}$ . A new data set with an error just two times smaller than the average error becomes very informative. On the other hand, if  $\sigma$  is just doubled, for only small regions at the edge of the  $P - \bar{n}$  plane the  $EU$  remains high (figure 6.11 (b)), because these experimental configurations would be very unusual compared with the other data.

As a numerical example, the  $EU$  for  $P = 1 MW$  and  $\bar{n} = 1 \cdot 10^{19} m^{-3}$  is given with  $5.67 \cdot 10^{-3} bit$  for a normal error (figure 6.10 (a)). If the measurement error is doubled, the expected information gain becomes  $3.91 \cdot 10^{-4} bit$ , whereas it is  $6.428 \cdot 10^{-2} bit$  if the error is halved.

One can conclude that the influence of the measurement error is significant, data points with large errors are not beneficial. Therefore, the measurement error has to be taken into account for experimental planning, and its value has to be estimated carefully. This result confirms the outcome from section 4.2.4, where a strong influence of the measurement error on the design result was found, too.

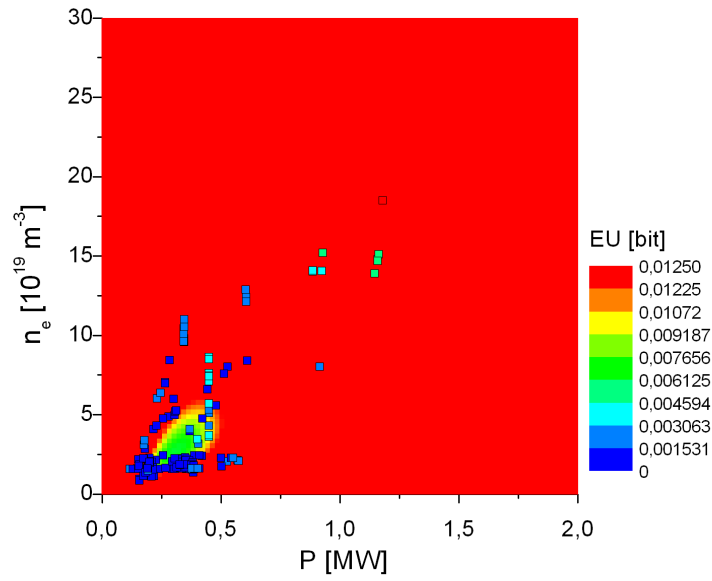
## 6.5 Discussion

In this chapter it was shown that the  $BED$  approach can be used for the planning of a future measurement, taking into account a database with measured data. Given a multi-dimensional linear physical problem, the Expected Utility of a new datum depends on the experimental configuration of the existing data sets and of the new datum. Also, the  $EU$  is significantly influenced by the measurement errors. The data values of the existing measurements, however, do not play a role (equation (6.10)), if the model function of the data is linearly depending on the parameters of interest.

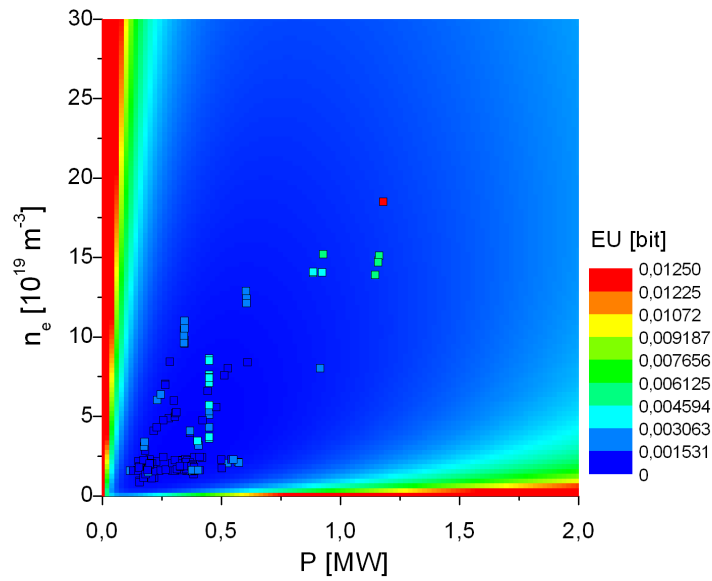
It can easily be seen from equation (6.10) that the  $EU$  decreases with an increasing measurement error  $\sigma$ . Consequently, for a future experiment the measurement error should be kept as small as possible.

Concerning the dependency on the experimental configuration one can conclude that measurements in regions of the configuration space with a sparse number of given data points show the highest Expected Utility, in other words, one should choose experimental configurations which have not been applied before to gain maximum information. For the analysed data set from W7-AS the regions with the highest  $EU$  were given for high  $\bar{n}$  and low  $P$ , and vice versa.

For the planning of the next measurement, the experimental configuration with the highest  $EU$  should be applied. Not every configuration may be realisable, technical and/or physical constraints may limit the configuration space.



(a)



(b)

Figure 6.11: Influence of the modified error for an additional measurement on the  $EU$  distribution, given 153 data sets: The error was given with (a)  $0.5 \cdot \sigma$  for the new measurement and (b)  $2 \cdot \sigma$  for the new measurement in comparison to figure 6.8.

Furthermore, one can also implement a *cost function* which expresses the necessary effort for the realisation of a certain parameter combination. The optimal experimental configuration is then given by the best relation of cost and expected information gain.

For the problem presented here, the *EU* is independent of the data values from the database. This allows one to design experimental campaigns with more than one new measurement: If the experimental conditions of the next measurement were estimated by *DAP*, this configuration can be added to the existing data set. Then, the next but one experiment can be planned using the *DAP* approach, and so on. The only assumption to be made is about the measurement error of future experiments, which may be derived from previous measurements. With this method, more than one future experiment can be planned. Therefore, the approach of *DAP* can also be used for campaign design.



# Chapter 7

## Summary

In this work the approach of Bayesian experimental design was applied to optimisation problems related to plasma physics: the design of a multi-channel interferometer and the planning of experiments for the determination and validation of energy confinement time scalings. The probabilistic ansatz of *BED* allows to use the tools of Bayesian probability theory. The work is related to the assembly of the W7-X stellarator, a nuclear fusion experiment.

### Design principle

The basic idea of Bayesian experimental design is the maximisation of the expected information gain (Expected Utility) from future measurements with respect to the design parameters. To express the information gain of an experiment, an information measure (Kullback-Leibler distance) is applied as utility function. It quantifies the information gain by comparison of the state of knowledge before and after the measurement. The expected information gain is given by integration of the utility function over the expected data space. This data space is calculated according to the mathematical model of the experiment (forward function) and the error statistics.

In result, the Expected Utility (*EU*) is an absolute measure of the expected information gain and therefore provides a quantification of the design: The *EU* is commonly expressed in *bit*. This quantification can be used to compare different designs.

The *BED* formalism is independent of the experiment to be designed. It allows one to implement the parameters of interest, which describe the physical problem and have to be estimated by the measurement, directly as a design criterion: As a part of the forward function, the parameters of interest have a direct impact to the expected information gain. By applying an appropriate distribution for these parameters it is also possible to implement a weighting to the range of interest.

In this work, the *BED* approach was adapted to the field of nuclear fusion research, in particular to physical challenges of the Wendelstein 7-X experiment. Using data sets from a previous experiment, the meaning of the *EU* was validated by comparing the *EU* values with the respective measurement results. The flexibility of *BED* was then demonstrated by applying the ansatz to two different fields of experimental design: the optimisation of a diagnostic and the planning and validation of experiments.

### Diagnostic optimisation: multi-channel interferometer

The diagnostic to be optimised was the multi-channel interferometer at the Wendelstein 7-X experiment. The interferometer provides line-integrated measurements of the plasma density, and is intended to be a start-up diagnostic at W7-X.



For this work, case studies for a one-beam interferometer have been made to evaluate the working principle of *BED* first. An interferometer provides line-integrated measurements of the plasma density, therefore the effects of different parameters of a density distribution on the beam configuration were studied. The influence of different error statistics was analysed, too. As a result, one can state that the optimal beam line crosses the plasma on a long path to achieve a good signal-to-noise ratio. In addition, the design is strongly affected by the error statistics.

The *BED* approach was then compared to a former design result for the W7-X interferometer. It was found out, that the parametrisation of the density distribution, which was used for this design, led to a significant and misleading impact on the design outcome: The assumption of a stepwise constant density distribution resulted in a beam line design which was strongly influenced by the mapping of the density distribution. This could be demonstrated by using the *BED* approach for the same parametrisation.

To validate the *EU*, data sets from the interferometer at the W7-AS stellarator have been examined according to the relationship between Expected Utility and the outcome of the data analyses. The data analysis was done independently from the calculation of the *EU*. It was shown that the *EU* criterion can be monotonically mapped on a global deviation ( $\chi^2$ ). This proves that the *EU* is an expression for the quality of the design: A design with a higher *EU* would on average lead to a better measurement outcome.

As a next step for the design of a multi-channel interferometer at W7-X, three actual problems from stellarator physics were used to formulate the optimisation criteria: the measurement of Core Electron Root Confinement, as predicted from neoclassical theory, measurement of the density effects of high plasma confinement regimes and the variation of the density by the injection of hydrogen pellets. The parameters of interest were given by the functional description of the respective problem.

According to these problems, three four-channel interferometers were designed. The results were compared by calculating the *EU* for every design with respect to every physical question. For the three analysed cases, the design for the measurement of CERC turned out to be the best solution in average. In addition, the implementation of an extra beam, which is congruent to the beam line of the Thomson scattering diagnostic and crosses the plasma on a very long path through the plasma centre, was found to be beneficial in all three cases. The installation of this beam can therefore be recommended.

All designs have been calculated with and without technical boundary conditions, namely the restriction of the possible beam lines by the port system at W7-X and the possible positions of the retro-reflectors. It turned out, that the influence of these restrictions is remarkable for the physical problems analysed here.

The four-beam design was also tested for different magnetic configurations of W7-X and for the influence of the plasma outside the last closed magnetic flux surface. These variations result only in minor changes of the beam line configuration. The *EU* of the original proposed design was similar to the respective *EUs* from the different magnetic configurations. According to this, the original four-channel design can also be applied for the other configurations analysed here without a significant loss of expected information. Alternatively, it may be possible to realise all beam line designs by a variable optical system, because the designs differ only marginally.

Finally, a possible proposal for an eight-channel interferometer was calculated. Here, two of the previous problems were used as design criteria: the measurement of CERC and the high confinement problem. The proposal for the four-channel interferometer was taken as "given". These four beam lines were not changed in the design process.

The result for eight lines of sight was, in principle, a combination of two four-beam interferometer, only slightly varied. It has to be mentioned that this proposal may change

due to the fact, that the eight-beam interferometer will be applied subsequently: Different physical questions of interest may arise from the first experimental results of W7-X, which may lead to a new design. Therefore, it is proposed here to revise the design and the optimisation goals according to the physical issues resulting in the first operation campaigns of W7-X.

Already indicated in the case studies with the one-beam interferometer, the importance of a good signal-to-noise ratio is reflected in the optimisation results: Beams crossing the plasma on a long distance show a high  $EU$  and are therefore preferred by the design process. This result was confirmed by the optimisation results for the four-channel interferometer. Here, the technical boundary conditions of the port system had to be taken into account, the restriction-free design was used for comparison. Again, beams with long path lengths inside the plasma show high  $EU$  values, in all cases the optimal design without restrictions differs significantly from the lines of sight allowed by the port system.

An important finding was, that the maximum of the expected information gain corresponds to lines of sight crossing the regions of the plasma, where the change in the parameters of interest shows the largest effect on the expected data.

## Experimental planning

As a second example for experimental design,  $BED$  was applied to the planning of a future experiment. In particular, it was asked for the best operational condition of a stellarator experiment to gain maximum information about the exponents of a scaling law for the energy confinement time. Data sets from previous experiments were available and have been implemented into the design, therefore this application is called data adaptive planning ( $DAP$ ).

The  $DAP$  approach was applied to a data set from the International Stellarator Scaling database ISS04. In particular, 153 data points from W7-AS with a rotational transform  $\iota = 1/3$  were selected.

$DAP$  was also applied to describe the value of a single data set: The datum was excluded from the database and its utility was calculated using the other data sets as "given". With this method, the most informative data points in a database could be identified.

Next, the utility for a possible future measurement, depending on the operational conditions, was calculated and the most informative regions in the space of the operational variables were identified. As a result it was found out, that the utility of a data point significantly depends on the error of the measurement as well as on the error of the measurements from the database. Also, the constellation of operational variables was important, too.

In general, two findings can be stated: First, a datum becomes informative, if it is measured in a region of the space of the operational variables not covered by other data points. Secondly, data points with a small measurement error show a high  $EU$ , which corresponds to the SNR criterion for the interferometer design.

Beside these results, the  $DAP$  approach is found to be a very effective method to validate given data and to find optimal operational conditions for future experiments. This may be of interest, especially for experiments with high operating costs or with low availability.

## General remarks

Bayesian experimental design was successfully applied to different tasks of plasma physics: the design of a multi-channel interferometer, and the planning of future experiments with

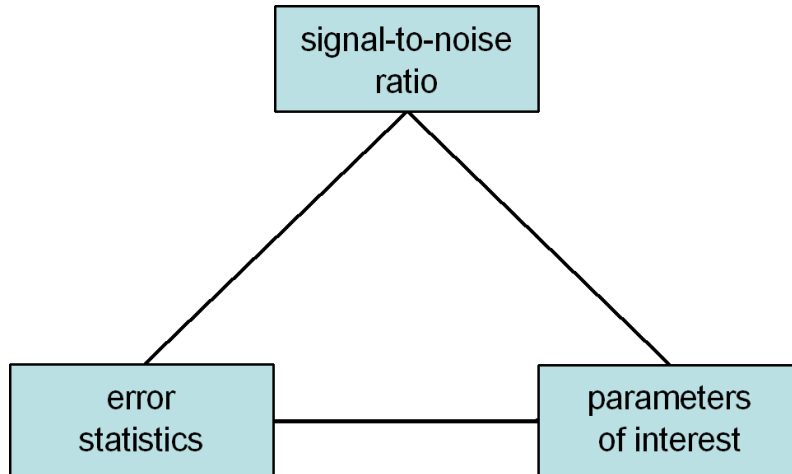


Figure 7.1: Estimation of the Expected Utility: Interaction between error statistics, signal-to-noise ratio and the parameters of interest.

respect to a given database. In general, the following statements can be made:

- For the diagnostic optimisation, high values of the expected information gain are given for a configuration where a change of the parameters of interest has maximum impact on the expected data.
- In experimental planning, experimental configurations which are not provided by the given database are preferred by the design for future measurements.
- The design strongly depends on the measurement error of the future data. In particular, the error statistics has a significant impact on the design result. This points out the necessity of a detailed error analysis of an experiment even in the design phase.
- The Expected Utility is affected by the signal-to-noise ratio of the measurement, in the cases analysed here a higher SNR leads to a higher  $EU$ .

However, it turns out that the dependency of the  $EU$  on these criteria is complex, the  $EU$  value is determined by an interaction of error statistics, SNR and parameters of interest (figure 7.1). The effect of the different aspects is thereby specific for the concrete design problem.

The influence of a particular criterion can be analysed in detail as it has been done in this work, e.g., for the error statistics. For the final design, however, the role of a single aspect is difficult to interpret in the most cases.

It is notable that these findings illustrate one of the main differences between Bayesian experimental design and other design approaches like, e.g., in [45]: Whereas in the latter case a good SNR is the design *criterion*, for  $BED$  it is the *outcome* of the design procedure. The design criteria, on the other hand, are the parameters of interest itself.

As a drawback it has to be mentioned that the computational costs of  $BED$  are high. This is a consequence resulting from the multi-dimensional integrals, which have to be solved for the calculation of the  $EU$ . Their computation strongly depends on the design problem itself and may take weeks, even if calculated on many computers using parallelisation routines (as in case of the interferometer), or just a few seconds on a single machine (as in case of experimental planning for scaling laws). For experiments like

Wendelstein 7-X, which is destined to be a long-run enterprise, it may be at least effective to spend the necessary effort in the optimisation of the diagnostics. Depending on the specific problem, for other experiments the cost-benefit ratio of the application of *BED* needs to be analysed.

## Outlook

As a result of this work one can state that Bayesian experimental design can be used for the optimisation of fusion diagnostics. It therefore seems to be the logical choice to apply the method to other diagnostics like Thomson scattering, X-ray tomography and so on. Given the forward calculations, it should also be possible to design sets of different diagnostics together, according to the approach of integrated data analysis [26]. However, the computation of the *EU* for such sets may become extensive due to the number of parameters involved.

The proposal for the four-channel interferometer was calculated for three different scenarios, other physical problems are also possible as a design criterion. In particular, the questions given for the initial phase of the W7-X operation may be of special interest due to the fact that the four-channel interferometer is a start-up diagnostic.

Some possible future developments have already been mentioned in this work: Due to the similar design of the interferometer for different magnetic configurations, it may be an option to realise all of these designs by a variable optical system. For the analysis of density effects at the plasma edge, a beam which is fanned out may be a possible arrangement.

A new development on the field of interferometer techniques is the *dispersion interferometer*. Here, only the dispersive elements of the optical arrangement contribute to the phase shift of the probing beam, vibrations are intrinsically cancelled out. The prototype of a multi-channel dispersion interferometer is currently under construction at the TEXTOR experiment [90], an interferometer of this type may become an option for W7-X. However, its unique properties would possibly lead to new error sources, which could influence the design.

With respect to the results from Data Adaptive Planning it can be concluded, that this method turns out to be very promising for the planning of future experiments and experimental campaigns. If a physical problem different from the scaling law approach is applied as design criterion, an analytical calculation of the *EU* may become impossible. In this case, a numerical integration similar to the interferometer example has to be used.



# Appendix A

## Basic Code Structure for *EU* Calculation

### A.1 Computing the Expected Utility

#### A.1.1 Code structure

The calculation of a multi-dimensional integral (equation (2.41)) is necessary for the computation of the Expected Utility. Because of the structure of the involved *PDFs*, the integral cannot be solved analytically in most cases, so a numerical integration routine has to be applied.

For this work, the Monte-Carlo method of *simple sampling* was used: Given a multiple integral of the form

$$F = \int_{a_x}^{b_x} dx \int_{a_y}^{b_y} dy f(x, y), \quad (\text{A.1})$$

random samples of the parameters  $x_r \in [a_x, b_x]$  and  $y_r \in [a_y, b_y]$  are drawn first. Then, the function values  $f_r(x_r, y_r)$  are calculated, summarised and divided by the number of samples  $R$ :

$$F = \frac{1}{R} \sum_{r=1}^R f_r. \quad (\text{A.2})$$

As an advantage, this method can, in principle, be applied for every kind of function  $f$ . As a drawback, a large number of samples is necessary to achieve a sufficient accuracy of the result, the accuracy is thereby proportional to  $\sqrt{R}$ .

In addition to the integration over the parameter and the data space, for the computation of the evidence (equation (2.12)) a further integration over the parameter space is needed. For this, the `vegas` routine from Numerical Recipes [46] was used, slightly modified for a reasonable first guess of the initial grid [30].

A commented outline of the source code can be found below. It covers the central part of the *EU* computation and does therefore not depend on a certain diagnostic.

For the optimisation of the multi-channel interferometer an additional integration was needed to calculate the different data values: The integration of the phase shift along the line of sight, see equation (3.5). For this, the phase shift was computed for beam segments along the line of sight with a length of  $0.5 \text{ mm}$ , for which the plasma density was assumed as constant. Finally, the particular phase shifts were summarised.

The computation of the *EU* finally consisted of four different integrations: Along the beam line for the calculation of the phase shift, over the parameter space for the calculation

of the evidence and finally over parameter and data space. To reduce the computation time, the calculation was done by parallelisation: The different Monte-Carlo (MC) samples were computed on different CPUs and finally combined.

The calculation was done on a blade cluster computer system, consisting of several dual core CPUs with hyper-threading capability. One hundred sub-programs were started, every sub-program itself calculated a certain number of MC samples.

In most cases, a number of 10000 samples in total (100 sub-programs with 100 samples, respectively) was sufficient to achieve a accuracy of  $> 1\%$  of the  $EU$  value. If a higher accuracy was needed to differentiate between two designs, the number of samples in the sub-program was increased.

The computation time depends mainly on the number of parameters of interest, but also on the number of data points (e.g. number of beam lines for an interferometer). Table A.1 shows the computation time for the calculation of the  $EU$  for a multi-channel interferometer (2, 3, 4 and 8 beams) and different physical problems (3, 4 or 5 parameters of interest). Here, the density parametrisation as used in chapter 4 was applied. For the

Configuration	Computation time
$N_{par} = 3, N_{dat} = 2$	6 min
$N_{par} = 3, N_{dat} = 3$	10 min
$N_{par} = 3, N_{dat} = 4$	16 min
$N_{par} = 4, N_{dat} = 4$	130 min
$N_{par} = 5, N_{dat} = 4$	285 min
$N_{par} = 5, N_{dat} = 8$	660 min

Table A.1: Computation time for  $EU$  calculation for different number of parameter of interest ( $N_{par}$ ) and different number of data points ( $N_{dat}$ ).

calculation shown here, 100 samples on 100 sub-programs were computed, respectively. The sub-programs were run on 11 CPUs, 4 sub-programs were started at the same time on one CPU (i.e. maximum 44 sub-programs were running at the same time). This was the typical configuration used for the  $EU$  calculation.

For four and five parameters of interest the computation time is about several hours, which is quite high if some 100  $EUs$  have to be calculated for the optimisation of the diagnostic. Therefore, a parametrisation of the physical problem, using three parameters of interest, is preferred and was applied in this way in chapter 5.

### A.1.2 Source code

In the following, the main parts of the source code used for the calculation of the Expected Utility will be shown. The programming language is Fortran90. The code will be commented in parts.

```
subroutine EU_quadrature_KL_MC (N_loop, exp_util, exp_util_unc)

integer(ikind), intent(in)  :: N_loop
real(rkind),   intent(out) :: exp_util, exp_util_unc

real(rkind), dimension(:), allocatable :: par
real(rkind), dimension(:), allocatable :: ln_ppv
real(rkind), dimension(:), allocatable :: dat
```

This subroutine computes the *EU* `exp_util` and its uncertainty `exp_util_unc` by evaluating `N_loop` Monte-Carlo runs. The variables `par` and `dat` are allocated with the number of parameters of interest and the number of data points, respectively. The logarithmised evidence `ln_ppv` is allocated with the number of MC loops, `N_loop`.

```
do eu_loop = 1, N_loop
  call make_prior(par)
  call make_likelihood(par, dat, ln_likelihood=ln_likelihood)
```

Here, a parameter sample `par` is drawn from the prior distribution by the function `make_prior`. This sample is then used to compute a data sample from the likelihood *PDF* by the function `make_likelihood`. For the interferometer, the data sample was calculated with

$$d = f + \epsilon. \quad (\text{A.3})$$

Here,  $\epsilon$  is the error of the measurement, randomly sampled from a Gaussian distribution with mean zero and standard deviation  $\sigma$ .  $f$  is the forward function, given by an integration of the electron density profile along the probing beam according to equation (3.5). For the mapping from real space to magnetic coordinates the `mconf` routine from Yu. Turkin was used [91].

The variable `ln_likelihood` gives the logarithmic value of the likelihood for parameters `par` and data `dat`.

```
ppv = prior_predictive_value(dat)
if (ABS(the_ppv)>1.d-99) then
  loop = loop + 1
  ln_ppv(loop) = ln_likelihood - log(ppv)
endif
ln_ppv(loop) = ln_ppv(loop) / log(2.)
enddo
```

The evidence `ppv` is now computed by the function `prior_predictive_value` (see below). The value is accepted if it is positive. Then, the quotient of likelihood and evidence is calculated and logarithmised (see equation (2.41)). To achieve the result in *bit*, the term is finally divided by the natural logarithm of 2.

```
exp_util      = SUM(ln_ppv(1:loop)) / loop
exp_util2     = SUM(ln_ppv(1:loop)**2) / loop
exp_util_var  = (exp_util2 - exp_util**2) / loop
exp_util_unc  = SIGN(1.d0,exp_util_var) * SQRT(ABS(exp_util_var))
```

```
end subroutine EU_quadrature_KL_MC
```

Finally, the Expected Utility is calculated from the result of the single loops. The uncertainty of the *EU* is given by the standard deviation of `exp_util`.

### Calculation of the evidence

The evidence - or prior predictive value - is the integral over the parameter space of the product from likelihood and prior (see equation (2.12)). It is computed by the `vegas` routine [46]:



```

function prior_predictive_value(dat)

do i = 1, SIZE(par)
  xi_start(1,i) = par(i)/upli(i) - (upli(i)-loli(i))/1.d4
  if (xi_start(1,i)<0.5d0*par(i)/upli(i))
    xi_start(1,i) = 0.5d0*par(i)/upli(i)
  xi_start(2,i) = par(i)/upli(i) + (upli(i)-loli(i))/1.d4
  if (xi_start(2,i)>0.5d0*(1.d0+par(i)/upli(i)))
    xi_start(2,i) = 0.5d0*(1.d0+par(i)/upli(i))
  xi_start(3,i) = 1.0
enddo
where (xi_start(:, :) < 1.d-99) xi_start(:, :) = 1.d-99

```

For the integration over the parameter space the `vegas` routine was slightly modified by choosing beneficial starting conditions `xi_start` [30]. The variables `loli` and `upli` are the lower and upper limits for the parameters of interest and are given by  $p(\theta)$  (see chapter 2).

```

init = 0
ncall = 300
call nr_vegas(joint_distr_dat_par_vegas, loli, upli,
              init, integral, sd, xi_start=xi_start)

init = 1
ncall = 200
vegas: do i = 1, loops
  call nr_vegas(joint_distr_dat_par, loli, upli,
                init, integral, sd)
  if (ABS(sd/integral) < acc_vegas) exit vegas
  ncall = 2*ncall
enddo vegas

prior_predictive_value = integral

end function prior_predictive_value

```

The `vegas` routine `nr_vegas` is called up for the first time using the initial guess `xi_start` and then repeated several times, until the necessary accuracy `acc_vegas` is achieved. The result is accepted, if the ratio of the standard deviation `sd` of the integral and the result `integral` is below 0.05.

The function to be integrated by the `vegas` routine, `joint_distr_dat_par`, is simply the product from likelihood and prior, according to equation (2.12):

```

function joint_distr_dat_par

call make_likelihood(par, likelihood=likelihood)
call make_prior(par, prior=prior)
joint_distr_dat_par = likelihood * prior

end function joint_distr_dat_par

```

So far, the routines shown here are in principle not restricted to a certain kind of diagnostics. The influence of the diagnostic is expressed by the forward function, which is part of the `make_likelihood` routine.

## A.2 Benchmarking

To benchmark the code, a simple-but-not-too-simple example was computed and compared with a quasi-analytic result. A uniform distribution for the prior in the limits  $\{\theta_{min}, \theta_{max}\}$ ,

$$p(\theta) = \frac{1}{\theta_{max} - \theta_{min}}, \quad (\text{A.4})$$

and a Gaussian likelihood

$$p(D | \theta, \eta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(D-f)^2}{2\sigma^2}\right) \quad (\text{A.5})$$

is applied. Here,  $\sigma$  is the standard deviation of the data  $D$ , which is taken as constant, and  $f$  is the forward function. For simplification, the dependence on the hypothesis  $H$  and the background information  $I$  will not be explicitly written down as far as it is not necessary. However, the dependence is always suggested implicitly.

To construct the forward function, a cylindrical plasma with a simple parabolic density profile is assumed:

$$n_e(r) = \theta \cdot (1 - r^2). \quad (\text{A.6})$$

For an interferometer, the data describing function is

$$f = C_1 \lambda \int n(r) dr. \quad (\text{A.7})$$

Here,  $\lambda$  is the wavelength of the probing laser beam and was chosen for an infrared  $CO_2$  laser ( $\lambda = 10.6 \mu m$ ).  $C_1$  is an aggregation of several physical constants (see chapter 3). Using the density profile in cartesian coordinates, one obtains

$$f = C_1 \lambda \cdot \theta \int_0^{\sqrt{1-x^2}} dy \left[ 1 - \left( \sqrt{x^2 + y^2} \right)^2 \right]. \quad (\text{A.8})$$

The goal is to find the best position of a probing beam, which is moved on the x-axis to estimate  $\theta$ . So, the x-coordinate becomes the design parameter  $\eta = x$ :

$$\begin{aligned} f &= C_1 \lambda \cdot \theta \int_0^{\sqrt{1-\eta^2}} dy [1 - \eta^2 - y^2] \\ &= \frac{2}{3} C_1 \lambda \theta (1 - \eta^2)^{3/2} \end{aligned} \quad (\text{A.9})$$

This leads to a likelihood function

$$p(D | \theta, \eta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(D - \frac{2}{3} C_1 \lambda \theta (1 - \eta^2)^{3/2})^2}{2\sigma^2}\right) \quad (\text{A.10})$$

using (A.5) and (A.9). Substituting  $C_2 = C_1 \lambda \cdot 2/3 \cdot (1 - \eta^2)^{3/2}$  one obtains

$$p(D | \theta, \eta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(D - C_2 \theta)^2}{2\sigma^2}\right). \quad (\text{A.11})$$

To compute the evidence  $p(D|\eta)$ , equation (2.12) is used:

$$\begin{aligned} p(D|\eta) &= d\theta p(\theta) p(D|\theta, \eta) \\ &= \frac{1}{\theta_{max} - \theta_{min}} \cdot \frac{1}{2C_2} \left[ \text{Erf} \left( \frac{D - C_2\theta_{min}}{\sqrt{2}\sigma} \right) - \text{Erf} \left( \frac{D - C_2\theta_{max}}{\sqrt{2}\sigma} \right) \right] \end{aligned} \quad (\text{A.12})$$

Now the Expected Utility function can be constructed by combining the prior, the likelihood and the evidence as shown in equation (2.41):

$$\begin{aligned} EU(\eta) &= \int_{-\infty}^{\infty} dD \int_{\theta_{min}}^{\theta_{max}} d\theta \frac{1}{\theta_{max} - \theta_{min}} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(D - C_2\theta)^2}{2\sigma^2}\right) \\ &\quad \cdot \text{ld} \frac{\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(D - C_2\theta)^2}{2\sigma^2}\right)}{\frac{1}{\theta_{max} - \theta_{min}} \cdot \frac{1}{2C_2} \cdot \left[ \text{Erf} \left( \frac{D - C_2\theta_{min}}{\sqrt{2}\sigma} \right) - \text{Erf} \left( \frac{D - C_2\theta_{max}}{\sqrt{2}\sigma} \right) \right]} \\ &= \text{ld} \left( \frac{4C_1\lambda(1 - \eta^2)^{3/2}(\theta_{max} - \theta_{min})}{3\sigma\sqrt{2\pi}e} \right) \\ &\quad - \int_{-\infty}^{\infty} dD \frac{1}{\theta_{max} - \theta_{min}} \cdot \frac{3}{4C_1\lambda(1 - \eta^2)^{3/2}} \\ &\quad \cdot \left[ \text{Erf} \left( \frac{D - 2/3 \cdot C_1\lambda(1 - \eta^2)^{3/2}\theta_{min}}{\sqrt{2}\sigma} \right) \right. \\ &\quad \quad \left. - \text{Erf} \left( \frac{D - 2/3 \cdot C_1\lambda(1 - \eta^2)^{3/2}\theta_{max}}{\sqrt{2}\sigma} \right) \right] \\ &\quad \cdot \text{ld} \left[ \text{Erf} \left( \frac{D - 2/3 \cdot C_1\lambda(1 - \eta^2)^{3/2}\theta_{min}}{\sqrt{2}\sigma} \right) \right. \\ &\quad \quad \left. - \text{Erf} \left( \frac{D - 2/3 \cdot C_1\lambda(1 - \eta^2)^{3/2}\theta_{max}}{\sqrt{2}\sigma} \right) \right] \end{aligned} \quad (\text{A.13})$$

Here, the last integration over  $D$  cannot be solved analytically, however, in this state of calculation the  $EU$  can be computed using the **Mathematica** software package [92].

The  $EU$  was also calculated using the Monte Carlo algorithms as described above. Figure A.1 shows the quasi-analytical solution (red), calculated with **Mathematica**, and the results from the MC sampling code (black dots) for two different standard deviations. The MC sampling code was tested successfully, both results show a good agreement.

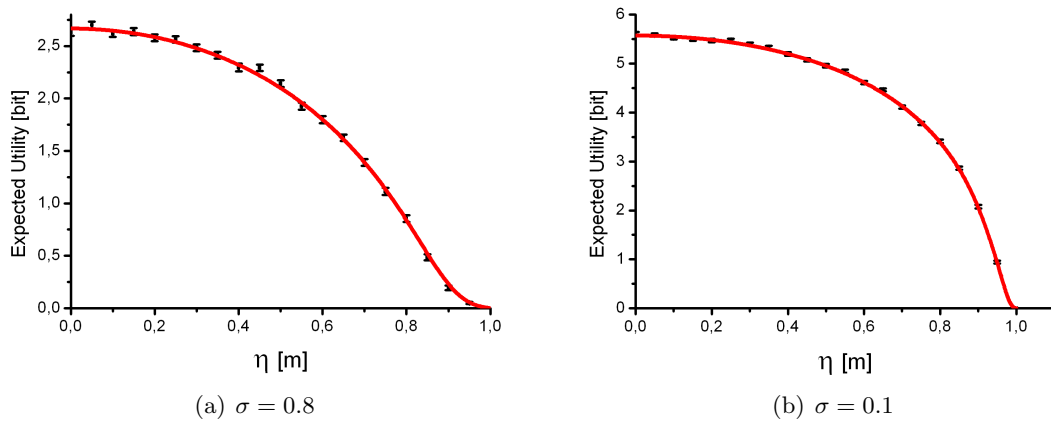


Figure A.1: Expected Utility solved quasi-analytically according to eq. (A.13) (red) and with Monte Carlo sampling (black dots) for two different standard deviations.



## Appendix B

# Magnetic Configurations for W7-X

The magnetic configurations used in chapter 5 shall be illustrated here shortly. For the standard case and the high- $\iota$  case, detailed descriptions can be found in [72]. The `mconf` program, with `mcviewer` for visualisation [91], was used for the calculation of the flux surface plots.

The plots show the configuration of the magnetic flux surfaces at W7-X at a toroidal angle of  $\Phi = 195^\circ$ , which is the position of the interferometer diagnostic. The relative distribution of the magnetic field strength  $|B|$  in toroidal and poloidal direction is shown, too. The standard configuration is displayed in figure B.1, figure B.2 shows the high- $\iota$  configuration and figure B.3 the high- $\beta$  case. Also, in figure B.4 the differences of the radial  $\iota$  profile for the standard and the high- $\iota$  configuration are illustrated.

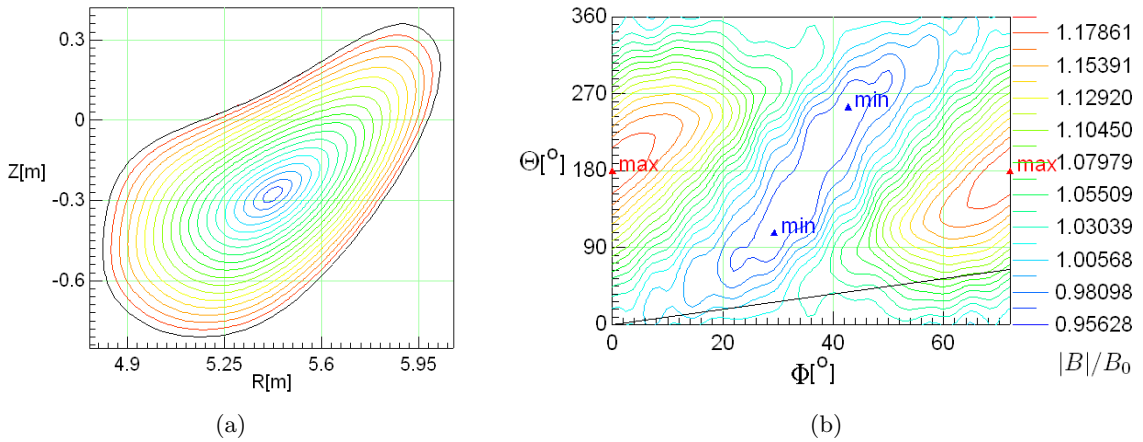


Figure B.1: Standard configuration at W7-X: Contour plot of the magnetic flux surfaces at a toroidal angle of  $\Phi = 195^\circ$ , where the interferometer is located (a). Figure (b) shows the magnetic field strength  $|B|/B_0$  with respect to the toroidal and poloidal angle,  $\Phi$  and  $\Theta$ .  $B_0$  is the minimum of  $|B|$  on the axis.

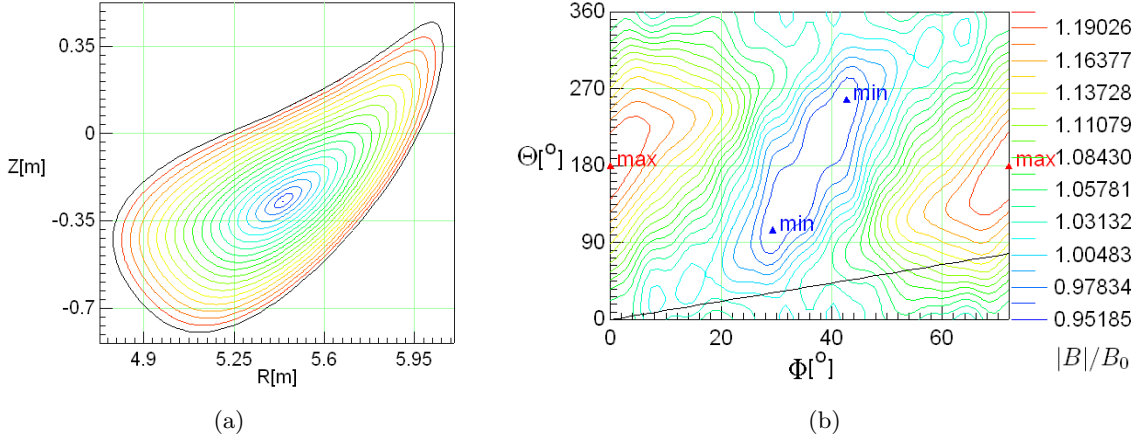


Figure B.2: High- $t$  configuration at W7-X: Contour plot of the magnetic flux surfaces at a toroidal angle of  $\Phi = 194^\circ$ , where the interferometer is located (a). Figure (b) shows the magnetic field strength  $|B|/B_0$  with respect to the toroidal and poloidal angle,  $\Phi$  and  $\Theta$ .  $B_0$  is the minimum of  $|B|$  on the axis.

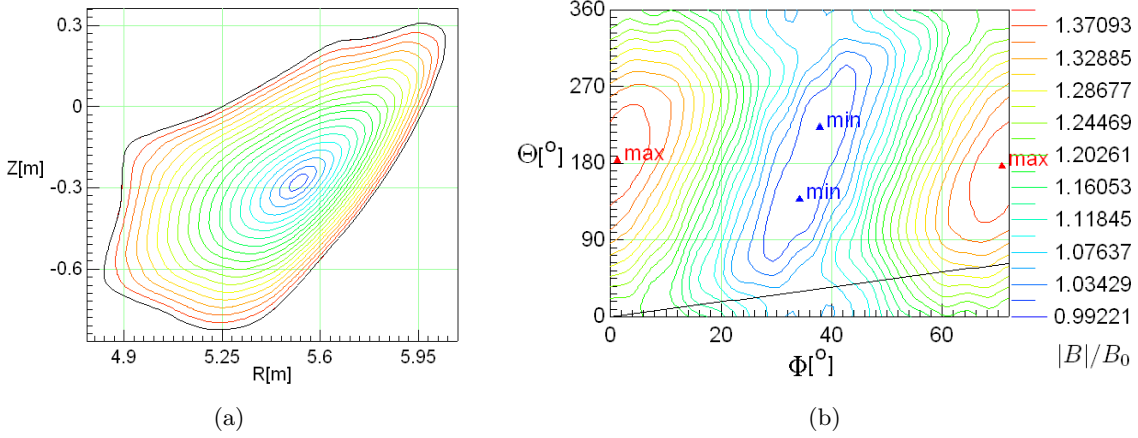


Figure B.3: High- $\beta$  configuration at W7-X[73]: Contour plot of the magnetic flux surfaces at a toroidal angle of  $\Phi = 195^\circ$ , where the interferometer is located (a). Figure (b) shows the magnetic field strength  $|B|/B_0$  with respect to the toroidal and poloidal angle,  $\Phi$  and  $\Theta$ .  $B_0$  is the minimum of  $|B|$  on the axis.

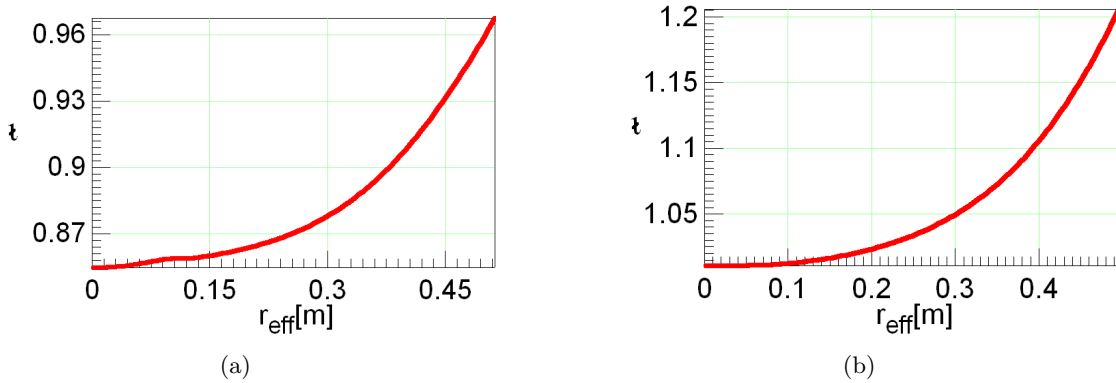


Figure B.4: Radial  $t$  profile for the standard case (a) and the high- $t$  case (b).

## Appendix C

# Detailed Calculation for Data Adaptive Design

In chapter 6 the Expected Utility for a future measurement was given with respect to previous experiments. The calculations are somewhat extensive and shall be shown here in detail [93, 94].

The following notations were used:

- $q$  number of existing data points
- $p$  number of scaling parameters
- $\mathbf{d}$  existing data
- $\mathbf{s}$  errors of the existing data
- $\mathbf{X}$  experimental configuration of the existing data set:  
 $\mathbf{X}_i = \{a_i, B_i, \bar{n}_i, P_i\}$
- $D$  new datum (to be measured)
- $\sigma$  error of the new datum
- $\boldsymbol{\xi}$  experimental configuration of the new datum  $\rightarrow$  design parameters
- $\boldsymbol{\alpha}$  scaling parameters  $\rightarrow$  parameters of interest

The  $EU$  is given with

$$\begin{aligned}
 EU(\mathbf{d}, \boldsymbol{\xi}) &= \int dD p(D|\mathbf{d}) \cdot U(D) \\
 &= \int d\tilde{\boldsymbol{\alpha}} p(\tilde{\boldsymbol{\alpha}}|\mathbf{d}) \cdot \int dD p(D|\tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi}) \cdot \\
 &\quad \int d\boldsymbol{\alpha} p(\boldsymbol{\alpha}|D, \mathbf{d}, \boldsymbol{\xi}) \ln \left[ \frac{p(\boldsymbol{\alpha}|D, \mathbf{d}, \boldsymbol{\xi})}{p(\boldsymbol{\alpha}|\mathbf{d})} \right].
 \end{aligned} \tag{C.1}$$

With a constant prior for  $\boldsymbol{\alpha}$ ,  $p(\boldsymbol{\alpha}) = \text{const}$ , and the likelihood for the old data

$$p(\mathbf{d}|\boldsymbol{\alpha}) = \frac{\sqrt{\det \mathbf{C}}}{(2\pi)^{q/2}} \exp \left[ -\frac{1}{2}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha})^T \mathbf{C}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha}) \right], \tag{C.2}$$

one obtains the posterior, using Bayes theorem (equation (2.9)), with

$$p(\boldsymbol{\alpha}|\mathbf{d}) = \frac{1}{Z} \exp \left[ -\frac{1}{2}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha})^T \mathbf{C}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha}) \right]. \tag{C.3}$$

Using the normalisation  $\int p(\boldsymbol{\alpha}|\mathbf{d}) d\boldsymbol{\alpha} = 1$  the normalisation factor  $Z$  is given by

$$Z = \int d\boldsymbol{\alpha} \exp \left[ -\frac{1}{2}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha})^T \mathbf{C}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha}) \right]. \tag{C.4}$$



This can be calculated by completing the square:

$$\begin{aligned} \mathbf{d}^T \mathbf{C} \mathbf{d} - \mathbf{d}^T \mathbf{C} \mathbf{X} \boldsymbol{\alpha} - \boldsymbol{\alpha}^T \mathbf{X}^T \mathbf{C} \mathbf{d} + \boldsymbol{\alpha}^T \mathbf{X}^T \mathbf{C} \mathbf{X} \boldsymbol{\alpha} &= (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_0)^T \hat{\mathbf{A}} (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_0) + \hat{B} \\ &= \boldsymbol{\alpha}^T \hat{\mathbf{A}} \boldsymbol{\alpha} - \boldsymbol{\alpha}^T \hat{\mathbf{A}} \hat{\boldsymbol{\alpha}}_0 - \hat{\boldsymbol{\alpha}}_0^T \hat{\mathbf{A}} \boldsymbol{\alpha} + \hat{\boldsymbol{\alpha}}_0^T \hat{\mathbf{A}} \hat{\boldsymbol{\alpha}}_0 + \hat{B}. \end{aligned} \quad (\text{C.5})$$

By comparison of the terms one gets

$$\begin{aligned} \hat{\mathbf{A}} &= \mathbf{X}^T \mathbf{C} \mathbf{X} \\ \hat{\boldsymbol{\alpha}}_0 &= \hat{\mathbf{A}}^{-1} \mathbf{X}^T \mathbf{C} \mathbf{d} \\ \hat{\boldsymbol{\alpha}}_0^T &= \mathbf{d}^T \mathbf{C} \mathbf{X} \hat{\mathbf{A}}^{-1} \\ \hat{B} &= \mathbf{d}^T \mathbf{C} \mathbf{d} - \mathbf{d}^T \mathbf{C} \mathbf{X} (\mathbf{X}^T \mathbf{C} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{C} \mathbf{d}, \end{aligned}$$

and with

$$\begin{aligned} Z &= \int d\boldsymbol{\alpha} \exp \left[ -\frac{1}{2} (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_0)^T \hat{\mathbf{A}} (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_0) \right] \exp \left[ -\frac{1}{2} \hat{B} \right] \\ &= \exp \left[ -\frac{1}{2} \hat{B} \right] \frac{(2\pi)^{p/2}}{\sqrt{\det \hat{\mathbf{A}}}} \end{aligned} \quad (\text{C.6})$$

finally

$$\begin{aligned} p(\boldsymbol{\alpha} | \mathbf{d}) &= \frac{\sqrt{\det \hat{\mathbf{A}}}}{(2\pi)^{p/2}} \exp \left[ \frac{1}{2} \hat{B} \right] \exp \left[ -\frac{1}{2} (\mathbf{d} - \mathbf{X} \boldsymbol{\alpha})^T \mathbf{C} (\mathbf{d} - \mathbf{X} \boldsymbol{\alpha}) \right] \\ &= \frac{\sqrt{\det \hat{\mathbf{A}}}}{(2\pi)^{p/2}} \exp \left[ -\frac{1}{2} (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_0)^T \hat{\mathbf{A}} (\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}_0) \right]. \end{aligned} \quad (\text{C.7})$$

For a new datum  $D$  one obtains an extended likelihood (see equation (6.7)):

$$\begin{aligned} p(D, \mathbf{d} | \boldsymbol{\alpha}, \boldsymbol{\xi}) &= \frac{1}{Z} \exp \left[ -\frac{1}{2} (\mathbf{d} - \mathbf{X} \boldsymbol{\alpha})^T \mathbf{C} (\mathbf{d} - \mathbf{X} \boldsymbol{\alpha}) - \frac{1}{2\sigma^2} (D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2 \right] \\ &= p(\mathbf{d} | \boldsymbol{\alpha}) \cdot p(D | \boldsymbol{\alpha}, \boldsymbol{\xi}). \end{aligned} \quad (\text{C.8})$$

The factoring in the last line is possible because  $D$  and  $\mathbf{d}$  are independent of each other. Normalisation and completing the square similar to the equations (C.5) and (C.6) lead to

$$\begin{aligned} \mathbf{A} &= \mathbf{X}^T \mathbf{C} \mathbf{X} + \frac{1}{\sigma^2} \boldsymbol{\xi} \boldsymbol{\xi}^T \\ \boldsymbol{\alpha}_0 &= \mathbf{A}^{-1} \left( \mathbf{X}^T \mathbf{C} \mathbf{d} + \frac{1}{\sigma^2} \boldsymbol{\xi} D \right) \\ \boldsymbol{\alpha}_0^T &= \left( \mathbf{d}^T \mathbf{C} \mathbf{X} + \frac{1}{\sigma^2} D \boldsymbol{\xi}^T \right) \mathbf{A}^{-1} \\ B &= \mathbf{d}^T \mathbf{C} \mathbf{d} + \frac{1}{\sigma^2} D^2 \\ &\quad - \left( \mathbf{d}^T \mathbf{C} \mathbf{X} + \frac{1}{\sigma^2} D \boldsymbol{\xi}^T \right) \left[ \mathbf{X}^T \mathbf{C} \mathbf{X} + \frac{1}{\sigma^2} \boldsymbol{\xi} \boldsymbol{\xi}^T \right]^{-1} \left( \mathbf{X}^T \mathbf{C} \mathbf{d} + \frac{1}{\sigma^2} \boldsymbol{\xi} D \right) \end{aligned}$$

and to the new posterior

$$\begin{aligned} p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi}) &= \frac{\sqrt{\det \mathbf{A}}}{(2\pi)^{p/2}} \exp \left[ \frac{1}{2} B \right] \\ &\quad \cdot \exp \left[ -\frac{1}{2} (\mathbf{d} - \mathbf{X} \boldsymbol{\alpha})^T \mathbf{C} (\mathbf{d} - \mathbf{X} \boldsymbol{\alpha}) - \frac{1}{2\sigma^2} (D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2 \right]. \end{aligned} \quad (\text{C.9})$$

To solve the integral in equation (C.1), one starts with the last integration: With

$$\ln \left[ \frac{p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi})}{p(\boldsymbol{\alpha} | \mathbf{d})} \right] = \ln \sqrt{\frac{\det \mathbf{A}}{\det \widehat{\mathbf{A}}}} + \frac{1}{2}(B - \widehat{B}) - \frac{1}{2\sigma^2}(D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2 \quad (\text{C.10})$$

one obtains for the integration over  $\boldsymbol{\alpha}$ :

$$\begin{aligned} I_1 &= \int d\boldsymbol{\alpha} p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi}) \ln \left[ \frac{p(\boldsymbol{\alpha} | D, \mathbf{d}, \boldsymbol{\xi})}{p(\boldsymbol{\alpha} | \mathbf{d})} \right] \\ &= \ln \sqrt{\frac{\det \mathbf{A}}{\det \widehat{\mathbf{A}}}} + \frac{1}{2}(B - \widehat{B}) \\ &\quad - \frac{1}{2\sigma^2} \frac{\sqrt{\det \mathbf{A}}}{(2\pi)^{p/2}} \int d\boldsymbol{\alpha} \exp \left[ \frac{1}{2}B \right] \cdot (D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2 \\ &\quad \cdot \exp \left[ -\frac{1}{2}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha})^T \mathbf{C}(\mathbf{d} - \mathbf{X}\boldsymbol{\alpha}) - \frac{1}{2\sigma^2}(D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2 \right] \\ &= \ln \sqrt{\frac{\det \mathbf{A}}{\det \widehat{\mathbf{A}}}} + \frac{1}{2}(B - \widehat{B}) \\ &\quad - \frac{1}{2\sigma^2} \frac{\sqrt{\det \mathbf{A}}}{(2\pi)^{p/2}} \int d\boldsymbol{\alpha} \exp \left[ \frac{1}{2}B \right] \cdot (D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2 \\ &\quad \cdot \exp \left[ -\frac{1}{2}((\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^T \mathbf{A}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_0) + B) \right]. \end{aligned} \quad (\text{C.11})$$

Using the substitution  $\mathbf{u} = \boldsymbol{\alpha} - \boldsymbol{\alpha}_0$ ,  $\boldsymbol{\alpha} = \mathbf{u} + \boldsymbol{\alpha}_0$  it becomes

$$\begin{aligned} I_1 &= \ln \sqrt{\frac{\det \mathbf{A}}{\det \widehat{\mathbf{A}}}} + \frac{1}{2}(B - \widehat{B}) \\ &\quad - \frac{1}{2\sigma^2} \frac{\sqrt{\det \mathbf{A}}}{(2\pi)^{p/2}} \int d\mathbf{u} (D - \boldsymbol{\xi}^T \mathbf{u} - \boldsymbol{\xi}^T \boldsymbol{\alpha}_0)^2 \\ &\quad \cdot \exp \left[ -\frac{1}{2}\mathbf{u}^T \mathbf{A} \mathbf{u} \right] \\ &= \ln \sqrt{\frac{\det \mathbf{A}}{\det \widehat{\mathbf{A}}}} + \frac{1}{2}(B - \widehat{B}) \\ &\quad - \frac{1}{2\sigma^2} \frac{\sqrt{\det \mathbf{A}}}{(2\pi)^{p/2}} \int d\mathbf{u} ((D - \boldsymbol{\xi}^T \boldsymbol{\alpha}_0)^2 + \mathbf{u}^T \boldsymbol{\xi} \boldsymbol{\xi}^T \mathbf{u}) \\ &\quad \cdot \exp \left[ -\frac{1}{2}\mathbf{u}^T \mathbf{A} \mathbf{u} \right]. \end{aligned} \quad (\text{C.12})$$

With

$$\int d\boldsymbol{\alpha} \boldsymbol{\alpha}^T \mathbf{H} \boldsymbol{\alpha} \cdot \exp \left[ -\frac{1}{2}\boldsymbol{\alpha}^T \mathbf{G} \boldsymbol{\alpha} \right] = \frac{(2\pi)^{p/2}}{\sqrt{\det \mathbf{G}}} \cdot \text{Tr}(\mathbf{G}^{-1} \mathbf{H}) \quad (\text{C.13})$$

one obtains

$$\begin{aligned} I_1 &= \ln \sqrt{\frac{\det \mathbf{A}}{\det \widehat{\mathbf{A}}}} + \frac{1}{2}(B - \widehat{B}) - \frac{1}{2\sigma^2} \frac{\sqrt{\det \mathbf{A}}}{(2\pi)^{p/2}} \\ &\quad \cdot \left[ (D - \boldsymbol{\xi}^T \boldsymbol{\alpha}_0)^2 \frac{(2\pi)^{p/2}}{\sqrt{\det \mathbf{A}}} + \frac{(2\pi)^{p/2}}{\sqrt{\det \mathbf{A}}} \text{Tr}(\mathbf{A}^{-1} \boldsymbol{\xi} \boldsymbol{\xi}^T) \right] \\ &= \ln \sqrt{\frac{\det \mathbf{A}}{\det \widehat{\mathbf{A}}}} + \frac{1}{2}(B - \widehat{B}) - \frac{1}{2\sigma^2} [(D - \boldsymbol{\xi}^T \boldsymbol{\alpha}_0)^2 + \text{Tr}(\mathbf{A}^{-1} \boldsymbol{\xi} \boldsymbol{\xi}^T)]. \end{aligned} \quad (\text{C.14})$$

Now, the expected utility can be written as

$$EU = \ln \sqrt{\frac{\det \mathbf{A}}{\det \hat{\mathbf{A}}}} - \frac{1}{2\sigma^2} \text{Tr}(\mathbf{A}^{-1} \boldsymbol{\xi} \boldsymbol{\xi}^T) + \int d\tilde{\boldsymbol{\alpha}} p(\tilde{\boldsymbol{\alpha}} | \mathbf{d}) \cdot \int dD p(D | \tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi}) \cdot \left[ \frac{1}{2} (B - \hat{B}) - \frac{1}{2\sigma^2} (D - \boldsymbol{\xi}^T \boldsymbol{\alpha}_0)^2 \right]. \quad (\text{C.15})$$

The integration over  $D$  is performed by using  $p(D | \tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi}) = \delta(D - \boldsymbol{\xi}^T \tilde{\boldsymbol{\alpha}})$  (equation (6.9)) and the expressions for  $\hat{\mathbf{A}}$ ,  $\mathbf{A}$ ,  $\hat{B}$  and  $B$  as well as  $\mathbf{Z} = \frac{\boldsymbol{\xi} \boldsymbol{\xi}^T}{\sigma^2}$ :

$$\begin{aligned} \frac{1}{2} \int dD p(D | \tilde{\boldsymbol{\alpha}}, \boldsymbol{\xi}) \cdot \left[ (B - \hat{B}) - \frac{1}{\sigma^2} (D - \boldsymbol{\xi}^T \boldsymbol{\alpha}_0)^2 \right] = \\ \frac{1}{2} \left[ \mathbf{d}^T \mathbf{C} \mathbf{X} \left( \hat{\mathbf{A}}^{-1} - \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \right) \mathbf{X}^T \mathbf{C} \mathbf{d} \right. \\ \left. - \tilde{\boldsymbol{\alpha}}^T \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{X}^T \mathbf{C} \mathbf{d} - \mathbf{d}^T \mathbf{C} \mathbf{X} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \tilde{\boldsymbol{\alpha}} \right. \\ \left. + \tilde{\boldsymbol{\alpha}}^T \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}^T \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \tilde{\boldsymbol{\alpha}} \right]. \end{aligned} \quad (\text{C.16})$$

For the final integration over  $\tilde{\boldsymbol{\alpha}}$  this expression falls into two parts: The first one is independent of  $\tilde{\boldsymbol{\alpha}}$  (see equation (C.7)):

$$\begin{aligned} \frac{1}{2} \int d\tilde{\boldsymbol{\alpha}} \frac{\sqrt{\det \hat{\mathbf{A}}}}{(2\pi)^{p/2}} \exp \left[ -\frac{1}{2} (\tilde{\boldsymbol{\alpha}} - \hat{\boldsymbol{\alpha}}_0)^T \hat{\mathbf{A}} (\tilde{\boldsymbol{\alpha}} - \hat{\boldsymbol{\alpha}}_0) \right] \cdot \\ \mathbf{d}^T \mathbf{C} \mathbf{X} \left( \hat{\mathbf{A}}^{-1} - \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \right) \mathbf{X}^T \mathbf{C} \mathbf{d} = \\ \frac{1}{2} \mathbf{d}^T \mathbf{C} \mathbf{X} \left( \hat{\mathbf{A}}^{-1} - \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \right) \mathbf{X}^T \mathbf{C} \mathbf{d}. \end{aligned} \quad (\text{C.17})$$

In the second part the substitution  $\mathbf{u} = \tilde{\boldsymbol{\alpha}} - \hat{\boldsymbol{\alpha}}_0$  is used:

$$\begin{aligned} \frac{1}{2} \frac{\sqrt{\det \hat{\mathbf{A}}}}{(2\pi)^{p/2}} \int d\mathbf{u} \exp \left( -\frac{1}{2} \mathbf{u}^T \hat{\mathbf{A}} \mathbf{u} \right) \cdot \\ \left[ -(\mathbf{u} - \hat{\boldsymbol{\alpha}}_0)^T \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{X}^T \mathbf{C} \mathbf{d} - \mathbf{d}^T \mathbf{C} \mathbf{X} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} (\mathbf{u} - \hat{\boldsymbol{\alpha}}_0) \right. \\ \left. + (\mathbf{u} - \hat{\boldsymbol{\alpha}}_0)^T (\mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} - \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z}) (\mathbf{u} - \hat{\boldsymbol{\alpha}}_0) \right]. \end{aligned} \quad (\text{C.18})$$

This integral can be calculated using (C.13) and finally one obtains

$$\begin{aligned} EU(\mathbf{d}, \boldsymbol{\xi}) = \ln \sqrt{\frac{\det \mathbf{A}}{\det \hat{\mathbf{A}}}} - \frac{1}{2\sigma^2} \text{Tr}(\mathbf{A}^{-1} \boldsymbol{\xi} \boldsymbol{\xi}^T) \\ + \frac{1}{2} \mathbf{d}^T \mathbf{C} \mathbf{X} \left[ \hat{\mathbf{A}}^{-1} - \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \right. \\ \left. - \hat{\mathbf{A}}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \hat{\mathbf{A}}^{-1} \right. \\ \left. + \hat{\mathbf{A}}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \hat{\mathbf{A}}^{-1} - \hat{\mathbf{A}}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \hat{\mathbf{A}}^{-1} \right] \mathbf{X}^T \mathbf{C} \mathbf{d} \\ + \frac{1}{2} \text{Tr} \left( \hat{\mathbf{A}}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \right) - \frac{1}{2} \text{Tr} \left( \hat{\mathbf{A}}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \mathbf{A}^{-1} \mathbf{Z} \right). \end{aligned} \quad (\text{C.19})$$

Using the identities [46]

$$\left( \hat{\mathbf{A}} + \beta \mathbf{g} \mathbf{g}^T \right)^{-1} = \hat{\mathbf{A}}^{-1} - \frac{\beta \hat{\mathbf{A}}^{-1} \mathbf{g} \mathbf{g}^T \hat{\mathbf{A}}^{-1}}{1 + \beta \mathbf{g}^T \hat{\mathbf{A}}^{-1} \mathbf{g}} \quad (\text{Sherman-Morrison formula}) \quad (\text{C.20})$$

and

$$\det(\widehat{\mathbf{A}} + \beta \mathbf{g}\mathbf{g}^T) = \det \widehat{\mathbf{A}} (1 + \beta \mathbf{g}^T \widehat{\mathbf{A}}^{-1} \mathbf{g}) \quad (\text{C.21})$$

as well as the expressions for  $\widehat{\mathbf{A}}$ ,  $\mathbf{A}$  and  $\mathbf{Z}$  one gets the final result for the Expected Utility:

$$\underline{\underline{EU(\mathbf{d}, \boldsymbol{\xi}) = \frac{1}{2} \left[ \log(1 + G) - \frac{G}{(1 + G)^2} \right] ; \quad G = \frac{\boldsymbol{\xi}^T (\mathbf{X}^T \mathbf{C} \mathbf{X})^{-1} \boldsymbol{\xi}}{\sigma^2}.}} \quad (\text{C.22})$$



# Appendix D

## Post-submit Changes

This chapter documents the modifications of the thesis after submission to the faculty. Please note, that all changes are only corrections of typing or compiling errors, not of the content. All calculations and their results as shown in the figures are correct, if no further statements are made.

### July 16th, 2007

- Reference [79] was changed from "to be published" to "submitted".
- In equations (C.2) and (C.3), a minus sign was missing:  $\exp[-\frac{1}{2} \dots]$

### July 18th, 2007

- The term  $(D - \boldsymbol{\xi}^T \boldsymbol{\alpha})^2$  was missing in equation (C.11) (L<sup>A</sup>T<sub>E</sub>X- typo).
- In equation (C.13), **G** and **H** were interchanged (L<sup>A</sup>T<sub>E</sub>X- typo).
- A typo in the final *EU* for *DAP* (eq. (6.10) and (C.22)) was corrected, in the submitted version, the  $(\dots)^{-1}$  was missing:  $G = \frac{\boldsymbol{\xi}^T (\mathbf{X}^T \mathbf{C} \mathbf{X})^{-1} \boldsymbol{\xi}}{\sigma^2}$ .

### July 19th, 2007

- In equation (6.9),  $\boldsymbol{\alpha}$  was changed to  $\tilde{\boldsymbol{\alpha}}$  (according to equation (6.5)).

### July 30th, 2007

- According to the correction of equation (6.10), equation (6.12) was also corrected. Here, figure 6.1 (b) and (c) had to be changed, too. The changes made were only of quantitative nature (scale of the *EU*), the conclusion given in the text are correct.

### August 28th, 2007

- In figure 4.21, the markings *S1* and *S2* were exchanged.

### December 12th, 2007

- In equation (6.21), squares are added to the  $\alpha_i$  terms according to the error propagation law (L<sup>A</sup>T<sub>E</sub>X- typo).

I would like to thank all colleagues who pointed out these errors to me.



# Bibliography

- [1] T. Hamacher, The Possible Role of Nuclear Fusion in the 21st Century, in *Plasma Physics*, edited by A. Dinklage et al., volume 670 of *Lecture Notes in Physics*, pages 461–482, Springer-Verlag Berlin Heidelberg, 2005.
- [2] Climate Change 2007: The Physical Science Basis, Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), 2007, [www.ipcc.ch](http://www.ipcc.ch).
- [3] C. F. von Weizsäcker, Über Elementumwandlungen im Innern der Sterne, Tl. 1, *Physikalische Zeitschrift* **38**, 176 (1937).
- [4] C. F. von Weizsäcker, Über Elementumwandlungen im Innern der Sterne, Tl. 2, *Physikalische Zeitschrift* **39**, 633 (1938).
- [5] H.-S. Bosch, Nuclear Fusion, in *Plasma Physics*, edited by A. Dinklage et al., volume 670 of *Lecture Notes in Physics*, pages 445–460, Springer-Verlag Berlin Heidelberg, 2005.
- [6] G. Borrelli, I. Cook, T. Hamacher, K. Lackner, P. Lako, R. Saez, G. C. Tosato, and D. J. Ward, Socio-Economic Research on Fusion, Summary of EU research 1997 - 2000, European Fusion Development Agreement (EFDA), 2001.
- [7] G. Gamow, Zur Quantentheorie des Atomkerns, *Zeitschrift für Physik* **51**(3), 204–212 (1928).
- [8] U. Schuhmacher, Basics of Plasma Physics, in *Plasma Physics*, edited by A. Dinklage et al., volume 670 of *Lecture Notes in Physics*, pages 3–20, Springer-Verlag Berlin Heidelberg, 2005.
- [9] J. D. Lawson, Some Criteria for a Useful Thermonuclear Reactor, Report GP/R 1807, United Kingdom Atomic Energy Research Establishment, 1955.
- [10] L. Spitzer, The Stellarator Concept, *Physics of Fluids* **1**(4), 253–264 (1958).
- [11] F. Wagner and H. Wobig, Magnetic Confinement, in *Plasma Physics*, edited by A. Dinklage et al., volume 670 of *Lecture Notes in Physics*, pages 137–172, Springer-Verlag Berlin Heidelberg, 2005.
- [12] M. Keilhacker, A. Gibson, C. Gormezano, P. Lomas, P. Thomas, M. Watkins, P. Andrew, B. Balet, D. Borba, C. Challis, I. Coffey, G. Cottrell, H. D. Esch, N. Deliyanakis, A. Fasoli, C. Gowers, H. Guo, G. Huysmans, T. Jones, W. Kerner, R. König, M. Loughlin, A. Maas, F. Marcus, M. Nave, F. Rimini, G. Sadler, S. Sharapov, G. Sips, P. Smeulders, F. Söldner, A. Taroni, B. Tubbing, M. von Hellermann, D. Ward, and the JET Team, High fusion performance from deuterium-tritium plasmas in JET, *Nuclear Fusion* **39**(2), 209–234 (1999).



- [13] ITER Physics Basis Editors, the ITER Physics Expert Group Chairs and Co-Chairs, and the ITER Joint Central Team and Physics Integration Unit, ITER Physics Basis Chapter 1: Overview and summary, *Nuclear Fusion* **39**(12), 2137–2174 (1999).
- [14] G. Grieger, W. Lotz, P. Merkel, J. Nührenberg, J. Sapper, E. Strumberger, H. Wobig, R. Burhenn, V. Erckmann, U. Gasparino, L. Giannone, H. J. Hartfuß, R. Jaenicke, G. Kühner, H. Ringler, A. Weller, F. Wagner, the W7-X Team, and the W7-AS Team, Physics optimization of stellarators, *Physics of Fluids B: Plasma Physics* **4**(7), 2081–2091 (1992).
- [15] H. J. Hartfuß, R. Brakel, M. Endler, T. Geist, P. Grigull, J. 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, *Review of Scientific Instruments* **68**(2), 1244–1249 (1997).
- [16] D. Sivia, *Data Analysis – A Bayesian Tutorial*, Oxford Science Publications, Clarendon Press, Oxford, 1996.
- [17] P. C. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica Support*, Cambridge University Press, 2005.
- [18] R. T. Cox, Probability, Frequency and Reasonable Expectation, *American Journal of Physics* **14**(1), 1–13 (1946).
- [19] V. Dose, Bayesian estimate of the Newtonian constant of gravitation, *Measurement Science and Technology* **18**, 176–182 (2007).
- [20] V. Dose, Bayes in five days, Lecture Notes, 2002, Greifswald.
- [21] K. Ertl, W. von der Linden, V. Dose, and A. Weller, Maximum Entropy based Reconstruction of Soft X-Ray Profiles in W7-AS, *Nuclear Fusion* **36**(11), 1477–1488 (1996).
- [22] M. Anton, H. Weisen, M. Dutch, W. von der Linden, F. Buhlmann, R. Chavan, B. Marletaz, P. Marmillod, and P. Paris, X-Ray Tomography on the TCV Tokamak, *Plasma Physics and Controlled Fusion* **38**(11), 1849–1878 (1996).
- [23] J. Koponen and O. Dumbrajs, Electron density profile reconstruction from multi-channel microwave interferometer data at W7-AS, *Review of Scientific Instruments* **68**(11), 4038–4042 (1997).
- [24] R. Preuss, V. Dose, and W. von der Linden, Model Comparison in Plasma Energy Confinement Scaling, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by V. Dose et al., pages 145–152, Dordrecht, 1999, Kluwer Academic.
- [25] R. Preuss and A. Dinklage, Model Comparison in Plasma Energy Confinement Scaling revisited, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by A. Mohammad-Djafari, pages 280–287, Melville, New York, AIP Conference Proceedings.
- [26] R. Fischer, A. Dinklage, and E. Pasch, Bayesian modelling of fusion diagnostics, *Plasma Physics and Controlled Fusion* **45**(7), 1095–1111 (2003).

- [27] A. Caticha and A. Griffin, Updating Probabilities, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by A. Mohammad-Djafari, pages 31–42, Melville, New York, 2006, AIP Conference Proceedings.
- [28] B. Pompe, Einführung in die Informationstheorie, Lecture Notes, 2004, Greifswald.
- [29] C. E. Shannon, A Mathematical Theory of Communication, Bell System Technical Journal **27**, 379–423, 623–656 (1948).
- [30] R. Fischer, Bayesian Experimental Design – Studies for Fusion Diagnostics, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by R. Fischer et al., pages 76–83, Melville, New York, 2004, AIP Conference Proceedings.
- [31] D. V. Lindley, On a Measure of the Information Provided by an Experiment, The Annals of Mathematical Statistics **27**(4), 986–1005 (1956).
- [32] D. V. Lindley, *Bayesian Statistics – A Review*, SIAM, Philadelphia, 1972.
- [33] K. Chaloner and I. Verdinelli, Bayesian Experimental Design: A Review, Statistical Science **10**(3), 273–304 (1995).
- [34] T. Loredò, Bayesian Adaptive Exploration, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by G. Erickson and Y. Zhai, pages 330–346, Melville, New York, 2003, AIP Conference Proceedings.
- [35] D. Veron, Submillimeter Interferometry of High Density Plasma, in *Infrared and Millimeter Waves*, edited by K. Button, volume 2, pages 69–135, Academic, New York, 1979.
- [36] I. N. Bronstein, K. A. Semendjajew, G. Grosche, V. Ziegler, and D. Ziegler, *Teubner-Taschenbuch der Mathematik*, B.G. Teubner Stuttgart Leipzig, 1996.
- [37] Y. Kawano, A. Nakashima, K. Tsuchiya, S. Gnuji, S. Chiba, and T. Hatae, Tangential  $CO_2$  Laser Interferometer for Large Tokamaks, Journal of Plasma and Fusion Research **73**(8), 870–891 (1997).
- [38] I. Hutchinson, *Principles of Plasma Diagnostics*, Cambridge University Press, 2nd edition, 2002.
- [39] Y. Kawano, A. Nakashima, S. Ishida, T. Fukuda, and T. Matoba,  $CO_2$  Laser Interferometer for Electron Density Measurement in JT-60U Tokamak, Review of Scientific Instruments **63**(10), 4971–4973 (1992).
- [40] M. Sanchez and J. Sanchez, Thermo-optical effect in zinc selenide windows for two-color interferometer for fusion plasma diagnostics, Review of Scientific Instruments **76**(4), 046104–1 – 4 (2005).
- [41] T. Geist, E. Würsching, and H. J. Hartfuß, Multichannel millimeter wave interferometer for W7-AS, Review of Scientific Instruments **68**(2), 1162–1167 (1997).
- [42] H. Wobig, private communication.
- [43] H. Horbach, HCN-Laser – Elektronik: Analoger Phasen Komparator, IPP Report III/79, Max-Planck-Institut für Plasmaphysik, 1982.

- [44] P. Kornejew, M. Hirsch, T. Bindemann, A. Dinklage, H. Dreier, and H. J. Hartfuß, Design of multichannel laser interferometry for W7-X, *Review of Scientific Instruments* **77**(10), 10F128 (2006).
- [45] M. Airila, O. Dumbrajs, M. Hirsch, J. Geiger, U. Neuner, and H. J. Hartfuß, Sightline optimization of the multichannel laser interferometer for W7-X, *Review of Scientific Instruments* **76**(2), 023501 (2005).
- [46] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 77*, Oxford Science Publications, Cambridge University Press, 2nd edition, 1986, 1992.
- [47] R. Fischer and V. Dose, Flexible and reliable profile estimation using exponential splines, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by A. Mohammad-Djafari, pages 296–303, Melville, New York, 2006, AIP Conference Proceedings.
- [48] H. Maaßberg, private communication.
- [49] H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew, and E. Pasch, Bayesian design of diagnostics: Case studies for Wendelstein 7-X, *Fusion Science and Technology* **50**(2), 262–267 (2006).
- [50] B. Frieden, Applications to Optics and Wave Mechanics of the Criterion of Maximum Cramer-Rao Bound, *Journal of Modern Optics* **35**(8), 1297–1316 (1988).
- [51] H. Maaßberg, C. D. Beidler, U. Gasparino, M. Rome, the W7-AS Team, K. S. Dya-bilin, N. B. Marushchenko, and S. Murakami, The neoclassical “Electron Root” feature in the Wendelstein 7-AS stellarator, *Physics of Plasmas* **7**(1), 295–311 (2000).
- [52] M. Yokoyama, H. Maaßberg, C. Beidler, V. Tribaldos, K. Ida, F. Castejón, T. Estrada, A. Fujisawa, T. Minami, T. Shimozuma, Y. Takeiri, J. Herranz, S. Murakami, and H. Yamada, Common Features of the Core Electron-Root Confinement in Helical Devices, *Fusion Science and Technology* **50**(3), 327–342 (2006).
- [53] U. Stroth, Transport in toroidal plasmas, in *Plasma Physics*, edited by A. Dinklage et al., volume 670 of *Lecture Notes in Physics*, pages 213–267, Springer-Verlag Berlin Heidelberg, 2005.
- [54] H. Mynick and W. Hitchon, Effects of the Ambipolar Potential on Stellarator Confinement, *Nuclear Fusion* **23**(8), 1053–1059 (1983).
- [55] D. Hastings, W. Houlberg, and K. Shang, The Ambipolar Electric Field in Stellarators, *Nuclear Fusion* **25**(4), 445–454 (1985).
- [56] S.-K. Chan, Steady-state kinetics of diffusionless first order phase transformations, *The Journal of Chemical Physics* **67**(12), 5755–5762 (1977).
- [57] H. Maaßberg, R. Brakel, R. Burhenn, U. Gasparino, P. Grigull, M. Kick, G. Kühner, H. Ringler, F. Sardei, U. Stroth, and A. Weller, Transport in stellarators, *Plasma Physics and Controlled Fusion* **35**(SB), B319–B332 (1993).
- [58] J. Baldzuhn, L. Giannone, M. Kick, K. McCormick, and W7-AS Team, Optimized confinement discharges in the stellarator W7-AS, *Plasma Physics and Controlled Fusion* **42**(4), 463–487 (2000).

- [59] U. Stroth, J. Baldzuhn, J. Geiger, T. Geist, L. Giannone, H.-J. Hartfuß, M. Hirsch, R. Jaenicke, M. Kick, J. P. Koponen, G. Kühner, F.-P. Penningfeld, F. Wagner, and the W7-AS Team, High-confinement NBI discharges in the W7-AS stellarator, *Plasma Physics and Controlled Fusion* **40**(8), 1551–1565 (1998).
- [60] F. Wagner, S. Bäuml, J. Baldzuhn, N. Basse, R. Brakel, R. Burhenn, A. Dinklage, D. Dorst, H. Ehmler, M. Endler, V. Erckmann, Y. Feng, F. Gadelmeier, J. Geiger, L. Giannone, P. Grigull, H.-J. Hartfuß, D. Hartmann, D. Hildebrandt, M. Hirsch, E. Holzhauer, Y. Igitkhanov, R. Jaenicke, M. Kick, A. Kislyakov, J. Kisslinger, T. Klinger, S. Klose, J. P. Knauer, R. König, G. Kühner, H. P. Laqua, H. Maaßberg, K. McCormick, H. Niedermeyer, C. Nührenberg, E. Pasch, N. Ramasubramanian, N. Ruhs, N. Rust, E. Sallander, F. Sardei, M. Schubert, E. Speth, H. Thomsen, F. Volpe, A. Weller, A. Werner, H. Wobig, E. Würsching, M. Zarnstorff, and S. Zolotnik, W7-AS: One step of the Wendelstein stellarator line, volume 12, page 072509, 2005.
- [61] A. Kreter, J. Baldzuhn, A. Dinklage, H. Ehmler, M. Kick, H. Maaßberg, and the W7-AS Team, Plasma profiles in optimized confinement discharges with high ion temperatures after installation of the island divertor in W7-AS, in *Proceedings of the 14th International Stellarator Workshop, Greifswald, 2003*.
- [62] F. Wagner, G. Becker, K. Behringer, D. Campbell, A. Eberhagen, W. Engelhardt, G. Fußmann, O. Gehre, J. Gernhardt, G. v. Gierke, G. Haas, M. Huang, F. Karger, M. Keilhacker, O. Klüber, M. Kornherr, K. Lackner, G. Lisitano, G. G. Lister, H. M. Mayer, D. Meisel, E. R. Müller, H. Murmann, H. Niedermeyer, W. Poschenrieder, H. Rapp, H. Röhr, F. Schneider, G. Siller, E. Speth, A. Stäbler, K. H. Steuer, G. Venus, O. Vollmer, and Z. Yü, Regime of Improved Confinement and High Beta in Neutral-Beam - Heated Divertor Discharges of the ASDEX Tokamak, *Physical Review Letters* **49**(19), 1408–1412 (1982).
- [63] V. Erckmann, F. Wagner, J. Baldzuhn, R. Brakel, R. Burhenn, U. Gasparino, P. Grigull, H. J. Hartfuß, J. V. Hofmann, R. Jaenicke, H. Niedermeyer, W. Ohlendorf, A. Rudyj, A. Weller, S. D. Bogdanov, B. Bomba, A. A. Borschevsky, G. Cattanei, A. Dodhy, D. Dorst, A. Elsner, M. Endler, T. Geist, L. Giannone, H. Hacker, O. Heinrich, and G. Herre, H mode of the W7-AS stellarator, *Physical Review Letters* **70**(14), 2086–2089 (1993).
- [64] F. Wagner, J. Baldzuhn, R. Brakel, R. Burhenn, V. Erckmann, T. Estrada, P. Grigull, H.-J. Hartfuß, G. Herre, M. Hirsch, J. V. Hofmann, R. Jaenicke, A. Rudyj, U. Stroth, A. Weller, and the W7-AS Teams, H-mode of W7-AS Stellarator, *Plasma Physics and Controlled Fusion* **36**(7A), A61–A74 (1994).
- [65] K. McCormick, P. Grigull, J. Baldzuhn, Y. Feng, S. Fiedler, L. Giannone, H. Hartfuß, A. Herrmann, D. Hildebrandt, M. Hirsch, J. Kißlinger, J. P. Knauer, G. Kühner, R. König, D. Naujoks, F. Penningfeld, J. Sallander, F. Sardei, H. Wobig, and the W7-AS Team, Core-edge studies with boundary island configurations on the W7-AS stellarator, *Plasma Physics and Controlled Fusion* **41**(12B), B285–B304 (1999).
- [66] E. Holzhauer, M. Hirsch, T. Grossmann, B. Brañas, and F. Serra, Theoretical and experimental investigation of the phase-runaway in microwave reflectometry, *Plasma Physics and Controlled Fusion* **40**(11), 1869–1886 (1998).

- [67] K. McCormick, P. Grigull, R. Burhenn, R. Brakel, H. Ehmler, Y. Feng, F. Gadelmeier, L. Giannone, D. Hildebrandt, M. Hirsch, R. Jaenicke, J. Kisslinger, T. Klinger, S. Klose, J. Knauer, R. König, G. Kühner, H. Laqua, D. Naujoks, H. Niedermeyer, E. Pasch, N. Ramasubramanian, N. Rust, F. Sardei, F. Wagner, A. Weller, U. Wenzel, and A. Werner, *New Advanced Operational Regime on the W7-AS Stellarator*, *Physical Review Letters* **89**(1), 015001 (2002).
- [68] R. Jaenicke, S. Bäuml, J. Baldzuhn, R. Brakel, R. Burhenn, H. Ehmler, M. Endler, V. Erckmann, Y. Feng, F. Gadelmeier, J. Geiger, L. Giannone, P. Grigull, H. J. Hartfuß, D. Hartmann, D. Hildebrandt, M. Hirsch, E. Holzhauser, M. Kick, J. Kisslinger, T. Klinger, S. Klose, J. Knauer, R. König, G. Kuehner, H. Laqua, H. Maaßberg, K. McCormick, R. Narayanan, H. Niedermeyer, E. Pasch, N. Ruhs, N. Rust, J. Saffert, F. Sardei, F. Schneider, M. Schubert, E. Speth, F. Wagner, A. Weller, U. Wenzel, A. Werner, and E. Würsching, *A new quasi-stationary, very high density plasma regime on the W7-AS stellarator*, *Plasma Physics and Controlled Fusion* **44**(12B), B193–B205 (2002).
- [69] M. Fujiwara, O. Motojima, Y. Hamada, T. Watari, M. Okamoto, J. Yamamoto, T. Kuroda, the LHD Group, and A. Iiyoshi, *Progress in the LHD (large helical device) program*, *Plasma Physics and Controlled Fusion* **39**(5A), A261–A274 (1997).
- [70] A. Iiyoshi, A. Komori, A. Ejiri, M. Emoto, H. Funaba, M. Goto, K. Ida, H. Idei, S. Inagaki, S. Kado, O. Kaneko, K. Kawahata, T. Kobuchi, S. Kubo, R. Kuramazawa, S. Masuzaki, T. Minami, J. Miyazawa, T. Morisaki, S. Morita, S. Murakami, S. Muto, T. Mutoh, Y. Nagayama, Y. Nakamura, H. Nakanishi, K. Narihara, K. Nishimura, N. Noda, S. Ohdachi, N. Ohyabu, Y. Oka, M. Osakabe, T. Ozaki, B. Peterson, A. Sagara, S. Sakakibara, R. Sakamoto, H. Sasao, M. Sasao, K. Sato, M. Sato, T. Seki, T. Shimozuma, M. Shoji, H. Suzuki, Y. Takeiri, K. Tanaka, K. Toi, T. Tokuzawa, K. Tsumori, K. Tsuzuki, K. Watanabe, T. Watari, H. Yamada, I. Yamada, S. Yamaguchi, M. Yokoyama, R. Akiyama, H. Chikaraishi, K. Haba, S. Hamaguchi, M. Iima, S. Imagawa, N. Inoue, K. Iwamoto, S. Kitagawa, J. Kodaira, Y. Kubota, R. Maekawa, T. Mito, T. Nagasaka, A. Nishimura, C. Takahashi, K. Takahata, Y. Takita, H. Tamura, T. Tsuzuki, S. Yamada, K. Yamauchi, N. Yanagi, H. Yonezu, Y. Hamada, K. Matsuoka, K. Murai, K. Ohkubo, I. Ohtake, M. Okamoto, S. Satoh, T. Satow, S. Sudo, S. Tanahashi, K. Yamazaki, M. Fujiwara, and O. Motojima, *Overview of the Large Helical Device project*, *Nuclear Fusion* **39**(9Y), 1245–1256 (1999).
- [71] N. Ohyabu, T. Morisaki, S. Masuzaki, R. Sakamoto, M. Kobayashi, J. Miyazawa, M. Shoji, H. Funaba, J. Harris, Y. Hirooka, S. Inagaki, K. Itoh, K. Narihara, N. Nakajima, Y. Narushima, S. Ohdachi, B. Peterson, S. Sakakibara, R. Sanchez, K. Tanaka, K. Watanabe, M. Yokoyama, K. Ida, T. Shimozuma, H. Yamada, Y. Nagayama, O. Kaneko, T. Mutoh, K. Kawahata, A. Komori, S. Sudo, and O. Motojima, *Properties of the LHD plasmas with a large island – super dense core plasma and island healing*, *Plasma Physics and Controlled Fusion* **48**(12B), B383–B390 (2006).
- [72] T. Andreeva, *Vacuum Magnetic Configurations of Wendelstein 7-X*, IPP Report III/270, Max–Planck–Institut für Plasmaphysik, 2002.
- [73] J. Geiger, private communication.
- [74] Y. Turkin, private communication.

- [75] D. Sharma, Y. Feng, F. Sardei, and D. Reiter, Three-dimensional Monte Carlo simulations of W7-X plasma transport: density control and particle balance in steady-state operations, *Nuclear Fusion* **45**(8), 825–836 (2005).
- [76] D. Sharma, Y. Feng, and F. Sardei, A 3D Monte-Carlo study of the W7-X island divertor transport for different magnetic configurations, *Nuclear Fusion* **46**(4), S127–S138 (2006).
- [77] H. Yamada, J. Harris, A. Dinklage, E. Ascasibar, F. Sano, S. Okamura, J. Talmadge, U. Stroth, A. Kus, S. Murakami, M. Yokoyama, C. Beidler, V. Tribaldos, K. Watanabe, and Y. Suzuki, Characterization of energy confinement in net-current free plasmas using the extended International Stellarator Database, *Nuclear Fusion* **45**(12), 1684–1693 (2005).
- [78] J. Connor and J. Taylor, Scaling Laws for Plasma Confinement, *Nuclear Fusion* **17**(5), 1047–1055 (1977).
- [79] R. Preuss, A. Dinklage, A. Weller, and the W7-AS team, Energy confinement scaling for W7-AS high- $\beta$  plasmas, *Physical Review Letters*, 2007, submitted.
- [80] R. Goldston, Energy Confinement Scaling in Tokamaks: Some Implications of Recent Experiments with Ohmic and Strong Auxiliary Heating, *Plasma Physics and Controlled Fusion* **26**(1A), 87–103 (1984).
- [81] S. M. Kaye, C. W. Barnes, M. G. Bell, J. C. DeBoo, M. Greenwald, K. Riedel, D. Sigmar, N. Uckan, and R. Waltz, Status of global energy confinement studies, *Physics of Fluids B: Plasma Physics* **2**(12), 2926–2940 (1990).
- [82] P. Yushmanov, T. Takizuka, K. Riedel, O. Kardaun, J. Cordey, S. Kaye, and D. Post, Scalings for Tokamak Energy Confinement, *Nuclear Fusion* **30**(10), 1999–2006 (1990).
- [83] J. Christiansen, J. Cordey, K. Thomsen, A. Tanga, the JET team, J. DeBoo, D. Schissel, T. Taylor, the DIII-D team, O. Kardaun, F. Wagner, F. Ryter, the ASDEX team, S. Kaye, the PBX-M, the PDX teams, Y. Miura, N. Suzuki, M. Mori, T. Matsuda, H. Tamai, T. Takizuka, S.-I. Itoh, K. Itoh, and the JFT-2M group, Global energy confinement H-mode database for ITER, *Nuclear Fusion* **32**(2), 291–338 (1992).
- [84] K. Thomsen, D. Campbell, J. Cordey, the JET team, O. Kardaun, F. Ryter, U. Stroth, A. Kus, the ASDEX team, J. DeBoo, D. Schissel, the DIII-D team, Y. Miura, N. Suzuki, M. Mori, T. Matsuda, H. Tamai, T. Takizuka, S.-I. Itoh, K. Itoh, the JFT-2M team, S. Kaye, the PBX-M, and the PDX teams, ITER H mode confinement database update, *Nuclear Fusion* **34**(1), 131–167 (1994).
- [85] K. Lackner and N. Gottardi, Tokamak Confinement in Relation to Plateau Scaling, *Nuclear Fusion* **30**(4), 767–770 (1990).
- [86] U. Stroth, M. Murakami, R. Dory, H. Yamada, S. Okamura, F. Sano, and T. Obiki, Energy Confinement Scaling from the International Stellarator Database, *Nuclear Fusion* **36**(8), 1063–1077 (1996).
- [87] A. Dinklage, H. Maaßberg, R. Preuss, Y. Turkin, E. Ascasibar, C. D. Beidler, R. Brakel, J. Geiger, J. Harris, A. Kus, S. Murakami, S. Okamura, F. Sano, U. Stroth, Y. Suzuki, J. Talmadge, V. Tribaldos, K. Y. Watanabe, A. Weller, H. Yamada, and M. Yokoyama,

Assessment of Global Stellarator Confinement: Status of the International Stellarator Confinement Data Base, 21st IAEA Fusion Energy Conference, EX/P7-1 and Nuclear Fusion, submitted 2007.

- [88] P. Rebut, P. Lallia, and M. Watkins, The critical temperature gradient model of plasma transport: Applications to JET and future tokamaks, in *Plasma Physics and Controlled Nuclear Fusion Research 1988, Vol. 2*, pages 191–200, Vienna, 1989, International Atomic Energy Agency (IAEA).
- [89] R. Preuss, V. Dose, and W. von der Linden, Dimensionally exact form-free energy confinement scaling in W7-AS, *Nuclear Fusion* **39**(7), 849–862 (1999).
- [90] P. A. Bagryansky, A. D. Khilchenko, A. N. Kvashnin, A. A. Lizunov, R. V. Voskoboynikov, A. L. Solomakhin, H. R. Koslowski, and the TEXTOR team, Dispersion interferometer based on a  $CO_2$  laser for TEXTOR and burning plasma experiments, *Review of Scientific Instruments* **77**(5), 053501 (2006).
- [91] Max-Planck-Institut für Plasmaphysik, Annual Report 2005, [www.ipp.mpg.de/de/for/publikationen/AR2005\\_nav.pdf](http://www.ipp.mpg.de/de/for/publikationen/AR2005_nav.pdf), 2005.
- [92] [www.wolframresearch.com](http://www.wolframresearch.com).
- [93] V. Dose, private communication.
- [94] R. Preuss, private communication.

## Acknowledgments

Three years of research are recapitulated and summarised in this thesis. Nevertheless, no one is able to handle such a work without any help. Fortunately I had many colleagues who shared their time, knowledge and experience with me whenever it was needed. Therefore, I want to take the opportunity to thank them here.

First of all, I wish to express my gratitude to my advisor, Dr. Andreas Dinklage, for his support, his patience, his ideas and suggestions and his friendly but clear critics whenever necessary. The door to his office was always open to me, and yes, I used that opportunity very often. Thank you very much!

It was Prof. Volker Dose, head of the former data analysis group at the IPP in Garching, who gave me the opportunity to learn all about Bayesian theory from the experts in his group. He also came up with the idea and first calculations for experimental planning, and I am very grateful for this.

In Garching, my first steps on Bayesian territory were made under the guidance of Dr. Rainer Fischer and Dr. Roland Preuss, but also Dr. Udo von Toissant and Silvio Gori. Thank you a lot for your patience, for your help, support and code sharing to make me become a FORTRAN user, for your kind hospitality and for your continuing interest in my DVD collection.

Special thanks go to Dr. Roland Preuss for his help on the calculations for experimental planning.

Some colleagues made my work easier by sharing and formating their code packages for my special purposes. My acknowledgments go to Dr. Yuriy Turkin for the `mconf` mapping routines and to Dr. Andreas Werner for programming the `intersector` interface and web service.

The construction of the interferometry system at Wendelstein 7-X lies in the hands of Dr. Matthias Hirsch and Dr. Petra Kornejew. Both of them helped me many times with details about the diagnostic. Their ideas, comments and suggestions have always been appreciated. In addition, I want to thank Dr. Hirsch for his help on finding and extracting the W7-AS data.

I am very thankful to a lot of colleagues for the discussions we had, especially Dr. Henning Maaßberg, Dr. Joachim Geiger and Prof. Friedrich Wagner.

In many cases the first person one asks a question is the person at the next table. And in many cases, the result of such questions is a longer discussion with a group of people around a whiteboard or a sheet of paper. For bringing together science and fun the credits go to my roommates Dirk Dodt, Maciej Krychowiak, Stefan Schmuck and Dr. Stefan Zegenhagen. Guys, whenever I had the desire for my own office, it was not your fault!

Many thanks go to the members of the Nabla – Arbeitsgruppe für Film, Fun und Füsik for the proof reading and all the other things what friends are for: Marcel Fuhrmann, Matthias Günther and Randy Mathes.

And finally, I wish to thank Anja Ritter for her help on correcting my English – just for a coffee.





# Curriculum Vitae

## Heiko Dreier

geboren am 9. Juni 1977 in Bad Saarow

## Eltern

Dietmar Dreier, geboren am 23. Juli 1955,  
selbständiger Fuhrunternehmer

Monika Dreier, geboren am 15. Juni 1953,  
Angestellte

## Geschwister

Anke Dreier, geboren am 28. April 1982,  
Studentin



## Lebenslauf

1984 - 1997	Schul Ausbildung an der Polytechnischen Oberschule bzw. Gesamtschule Storkow
Juli 1997 - Juni 1998	Wehrdienst
Oktober 1998 - Januar 2004	Studium der Physik (Diplom) an der Humboldt - Universität zu Berlin
Januar 2004	Abschluß der Diplomarbeit: "Spektroskopie hochgeladener Argon-Ionen an der EBIT"
März 2004 - April 2007	Doktorarbeit am Max-Planck-Institut für Plasmaphysik, Teilinstitut Greifswald: "Bayesian Experimental Design"

Fellow of the

"International Max-Planck Research School on Bounded Plasmas"



# List of publications

## Contributions to paper journals

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew and E. Pasch: *Bayesian Design of Diagnostics: Case Studies for Wendelstein 7-X*, Fusion Science and Technology, Volume 50, No. 2, pages 262-267 (2006)

P. Kornejew, M. Hirsch, T. Bindemann, A. Dinklage, H. Dreier, and H.-J. Hartfuß: *Design of multichannel laser interferometry for W7-X*, Review of Scientific Instruments, Volume 77, 10F128 (2006)

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch and P. Kornejew: *Bayesian Design of Plasma Diagnostics*, Review of Scientific Instruments, Volume 77, 10F323 (2006)

H. Dreier, R. Preuss, A. Dinklage and V. Dose: *Data Adaptive Exploration for Scaling Laws* (in preparation)

## Conference proceedings

A. Dinklage, R. Fischer, H. Dreier, J. Svensson and Yu. Turkin: *Integrated approaches in fusion data analysis*, in: Bayesian Inference and Maximum Entropy Methods in Science and Engineering 2004, AIP Conference Proceedings 735, pages 43-51, Melville, NY, 2004

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 2005, AIP Conference Proceedings 803, pages 440-447, Melville, NY, 2005

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch and P. Kornejew: *Optimization of plasma diagnostics using Bayesian probability theory*, in: Bayesian Inference and Maximum Entropy Methods in Science and Engineering 2006, AIP Conference Proceedings 872, pages 304-310, Melville, NY, 2006

## Poster contributions on international conferences

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, J. Svensson: *Concepts for diagnostics performance optimization*, 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Garching, Germany, 2004

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew, E. Pasch and Yu. Turkin: *Design of Diagnostics: Case Studies for Wendelstein 7-X*, 15th International Stellarator Workshop, Madrid, Spain, 2005

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew: *Physical Diagnostics Design*, 16th Topical Conference on High-Temperature Plasma Diagnostics, Williamsburg, USA, 2006

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew: *Bayesian design of a multi-channel interferometer*, 33rd EPS Conference on Plasma Physics, Rome, Italy, 2006

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew: *Optimization of plasma diagnostics using Bayesian probability theory*, 26th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Paris, France, 2006

### **Contributions to the Spring Meeting of the German Physical Society – Beiträge zur Frühjahrstagung der Deutschen Physikalischen Gesellschaft (DPG)**

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, and H.-J. Hartfuß: *Physical Diagnostic Design and Optimization*, Verhandlungen der DPG, Reihe VI, Band 40, S. 116, (2005), *Poster*

H. Dreier, A. Dinklage, R. Fischer, M. Hirsch, P. Kornejew: *Physikalisches Design eines Interferometers für Wendelstein 7-X*, Verhandlungen der DPG, Reihe VI, Band 41, S. 53, (2006), *Talk - Vortrag*

H. Dreier, R. Preuss, A. Dinklage, V. Dose: *Data adaptive control parameter estimation for scaling laws*, Verhandlungen der DPG, Reihe VI, Band 42, S. 120, (2007), *Poster*