Predictive power of nuclear mean-field theories for exotic-nuclei problem

Karolina Szczechowska Rybak

Institut Pluridisciplinaire Hubert Curien
Université de Strasbourg

Dissertation

Predictive Power of Nuclear Mean-Field Theories For Exotic-Nuclei Problem

submitted by

Karolina RYBAK

in partial fulfillment of the requirements for the degree of

Docteur en Sciences
at University of Strasbourg, France

Defended on September 21, 2012, in front of the commission:

Mr. Andrzej GÓŹDŹ	*Referee*
Mr. Piet VAN ISACKER	*Referee*
Mrs. Marianne DUFOUR	*Examiner*
Mr. Benoit GALL	*Examiner*
Mrs. Marie-Genevieve PORQUET	*Examiner*
Mr. Jerzy DUDEK	*Thesis Supervisor*

Resumé.

Cette thèse de doctorat vise l'examen critique de certaines théories de champ moyen nucléaire phénoménologiques, en se focalisant sur la description fiable des niveaux de particules individuelles. L'approche suivie ici est nouvelle en ce sens que elle permet non seulement la prédiction des valeurs numériques obtenues avec ce formalisme, mais également une estimation des distributions de probabilités correspondant aux résultats expérimentaux.

Nous introduisons le concept des «erreurs théoriques», visant estimer, dans un cadre mathématique bien établi, les incertitudes relatives aux modélisations théoriques. Il est également introduit une notion subjective de pouvoir prédictif des Hamiltoniens nucléaires, qui est analysé dans le contexte des spectres énergétiques de particules individuelles. Le concept mathématique du «Problème Inverse» est appliqué aux Hamiltoniens de champ moyen réalistes. Cette technique permet la prédiction de propriétés du système partir d'un nombre limité de données.

Afin dapprofondir notre connaissance des Problèmes Inverses, nous focalisons notre attention sur un problème mathématique simple. Une fonction dépendant de quatre paramètres libres est introduite afin de reproduire des données «expérimentales». Nous étudions le comportement des paramètres «fittés», leur corrélation, ainsi que les erreurs associées. Cette étude nous aide comprendre la signification de la formulation correcte du problème en question. Il nous montre également limportance dinclure les erreurs expérimentales et théoriques dans la solution.

Abstract.

This thesis is a critical examination of phenomenological nuclear mean field theories, focusing on reliable description of levels of individual particles. The approach presented here is new in the sense that it not only allows to predict the numerical values obtained with this formalism, but also yields an estimate of the probability distributions corresponding to the experimental results.

We introduce the concept of 'theoretical errors' to estimate uncertainties in theoretical models. We also introduce a subjective notion of 'Predictive Power' of nuclear Hamiltonians, which is analyzed in the context of the energy spectra of individual particles. The mathematical concept of 'Inverse Problem' is applied to a realistic mean-field Hamiltonian. This technique allows to predict the properties of a system from a limited number of data.

To deepen our understanding of Inverse Problems, we focus on a simple mathematical problem. A function dependent on four free parameters is introduced in order to reproduce 'experimental' data. We study the behavior of the 'fitted' parameters, their correlation and the associated errors. This study helps us understand the importance of the correct formulation of the problem. It also shows the importance of including theoretical and experimental errors in the solution.

Contents

List of Figures

List of Tables

Introduction

Nuclear physics is a very specific branch of science. It is probably one of the least discovered fields due to the typical size of an atomic nucleus being of the order of tens of femtometers. On the one hand we have such a small but of complex structure object to experiment upon, on the other, there is a problem of the strong and not sufficiently well known interactions appearing at a distance of a femtometer. An atomic nucleus is a complex object, consisting of tens or hundreds of nucleons squeezed in a very small space. Being fully aware of the fact that nucleons are also complex particles themselves, we treat them as the smallest constituents of the nucleus. However, this simplification proves to be not much of a solution, as the nucleons bound together into a nucleus bring enough of complications in the formal description/modelling.

The theory of nuclear interactions is incomplete, because we are unable to reconstruct the complexity of nucleon-nucleon interactions in the many-body systems. We can only try to approximate their description using approximate, incomplete theories such as nuclear shell-model, mean-field theory or collective model, to mention these with the longest history. Every model created has its own limitations and usually works under certain specific, limiting conditions. Moreover, all these models include free parameters, which need to be fitted to experimental information, that is affected by uncertainties. This very important fact is often ignored, while creating nuclear models. Experimental errors may have a large and/or unpredictable impact on the resulting parameters.

Knowing how inexact and limited the nuclear theories are, we should take a closer look at their ability to predict the properties of the nuclides outside the range the models were fitted to. Without this crucial step, we have no way of knowing, if our predictions are valid. This fact is an important issue when searching for yet unknown exotic nuclei – we have to know where to look for and the expected values should lie close to the real ones.

This work emphasizes the statistical aspect of mathematical models. We introduce the concept of 'Predictive Power' and stress the necessity of incorporating theoretical uncertainties into nuclear theories. Chapter 3 presents the results of a statistical analysis of an analytical mathematical model, showing the discussed properties of the fitting process.

In chapter 4 we present the results concerning the Predictive Power of spherical Woods-Saxon mean-field single-particle Hamiltonian and a way of improving its performance. We also discuss the number of necessary parameters of the Woods-Saxon formalism and present a way of significantly reducing them.

Chapter 1

Predictive Power

This chapter introduces a rather new and often underestimated concept of theoretical errors in microscopic theories of subatomic physics together with a related problem of theory's predictive power. At first we will discuss the notion[1] of predictive power of a given theory in a rather general context, i.e., without any reference to any particular theory focussing on its impact on existing nuclear theories. We will also discover the possible significance of predictive power for future experiments.

General Framework. Our understanding of subatomic world, although still strongly limited at present, is evolving in time. In particular, we do not have the truly microscopic theory of the nucleon-nucleon interactions that would be derived from a totally microscopic starting point for instance describing the structure of nuclear matter in terms of quarks and gluons. Our present-day theories that address processes taking place inside the nucleus are based on rather phenomenological concepts. A uniform theory of atomic nucleus, which would be acceptable for everyone does not exist so far despite the fact of significant progress in e.g. the recent developments of methods treating the nuclear few-body systems – we are still unable to describe all known nuclei and nuclear properties within one general-scope theory.

For many years there have been attempts undertaken to construct the more and more complete theories with increasing degrees of sophistication. Some of them, such as e.g. nuclear mean field theory, which is going to receive a particular attention in what follows, are very successful in reproducing certain physical properties, but are unable or fail in describing a more extended set of observables – the tasks which in turn can be achieved by other models (as e.g. the nuclear collective model). This brings us to an observation that will be of importance when introducing the notion of the 'theoretical errors' – quite generally the structure of the present day Hamiltonians can be written down as

$$\hat{H} = \hat{t} + \hat{v}_1 + \hat{v}_2 + \hat{v}_3 + \ \ldots \tag{1.1}$$

where the symbols $\hat{v}_i$ represent the presently known interactions whereas ellipsis repre-

[1] At present there is a quick growth of interest in the literature of the methods of constructing theories whose predictive power is treated in mathematical terms essentially using the Stochastic Methods and the mathematical theory of the so-called Inverse Problem, the latter will be discussed in more detail in the present work.

sents the interactions that are presently either unknown or intentionally ignored. These are the unknown/ignored terms which introduce:

Theories without or with limited Predictive Power. It sometimes happens that within one theory there are several parameterisations, which can be used depending on the quantity we want to calculate. Theories based on density functionals can serve as an example of such a behaviour. There are over 120 parameterisations in Hartree-Fock formalism with Skyrme interactions but neither one of them can be called universal. Such situation can even be satisfactory locally, but it is far from the ultimate goal: one theory applicable to whole of the nuclear chart, which reproduces all available experimental data and is able to correctly predict properties of yet unknown nuclei. The Hartree-Fock formalism has one more disadvantage, i.e. large number of parameters to be determined, but this will be discussed in more detail later on.

Most of existing theories in nuclear physics have one structural element in common – the coupling constants. These are some unknown parameters which define the relative strength of the terms in the interaction potential or even the functional dependence (for example diffuseness or radius of Woods-Saxon potential). These constants are usually obtained as a result of some fitting procedure, which compares experimental data to its theoretical counterpart. This approach can be quite dangerous. In order to be confident of our results we must chose carefully the data to compare with. Some choices can be more natural than the others, in the sense that it is easier to calculate and measure them – or more problematic. We can try to reproduce the masses of nuclides, rotational bands or single-particle energy levels. Whatever we chose, we have to remember that every experiment we perform is contaminated with errors, which can result, for example, from insufficient precision of used equipment or from complexity of measured quantity. The origin of these errors may be different but they are still unavoidable. However, almost every theoretical calculation in nuclear physics, which relies on experimentally obtained values in recovering parameters of the formalism, neglect the uncertainties of these quantities. Experimental data exploited in various fitting procedures are treated as points instead of probability distributions, thus obtained parameters are idealised and have limited applicability. This may seem as an overstatement, but can we really trust such parameterisations? In many calculations we use experimental data for a small number of nuclides to recover the coupling constants of a Hamiltonian, which is next used to calculate properties of other nuclides – we extrapolate the result. What if the parameterization is not exact, if there is some probability distribution associated with it? The error we make may be small for the data we use; after all we fitted the parameters to this set. However if we extrapolate an erroneous function, the result may be far from correct. The great significance of this observation can be easily recognised when we try to explore the region of heavy and super-heavy nuclei, where we have to rely on theoretical predictions only. This brings us to the most basic question: 'how to determine whether our calculations and models are correct and their parameterisations are stable?'.

Before we try to answer this question, let us first introduce and discuss some basic concepts.

1.1 Experimental Information

In our illustrations of the discussed concepts we focus mainly on single-particle mean-field Hamiltonians with Woods-Saxon or Skyrme-Hartree-Fock interactions restricted to doubly magic spherical nuclei. This phenomenological approach has several important extensions into such more modern realisations as Relativistic Mean Field or density dependent spin-orbit and tensor fields (see below). Furthermore, the single-particle energy levels as a natural choice of experimental data to compare to are also ones of the most difficult to measure. Nonetheless this formalism still has a large educational value due to its intuitive form and rather straightforward interpretation of terms included in the Hamiltonian. The properties of its parameters, i.e. their natural correlation, which will be discussed later, make it also a great tool to present the concepts of this work.

Experimental single-particle energy levels of spherical nuclei are very complex objects. They are extracted from light-projectile nuclear-transfer reactions, where they are populated or depopulated in residual $A \pm 1$ neighbouring nuclei. These reactions provide information about excitation energy, angular momentum and parity of single-particle or single-hole levels. However, these states do not need to be directly comparable to the eigen-values of the single-particle Hamiltonian. It often happens that there is more than one level with the same j^π characteristics in a given energy region. This situation can be explained by mixing of single-nucleonic states with core excitations or coupling to the collective surface vibrations. It is obvious that we have to consider all of the scattered states in determining the position of the eigen-energies in question. The true single-particle levels are thus extracted by calculation of the center of gravity of all levels with the same j^π weighted by the so-called spectroscopic factors, corresponding to the strengths of the excited levels extracted from the transfer reactions. This can be expressed as follows

$$\varepsilon_k = \sum_i S(k,i) e_i^{exc}, \tag{1.2}$$

where e_i^{exc} denotes the energy of the ith measured level and $S(k,i)$ the associated spectroscopic factor, which is defined as the probability that the wave-function of measured state e_i^{exc} contains the single-particle state with the energy ε_k[2]. In principle, we need all levels e_i^{exc} to properly recover the single-particle state j^π. This is ensured if the spectroscopic factors fulfil the sum rule

$$\forall k \quad \sum_i S(k,i) = 1. \tag{1.3}$$

Obviously this is not always the case – sometimes we do not have the full experimental information. Additionally, what is even more important, extraction of the spectroscopic factors can be subject to large errors. To minimise these effects we can modify the expression (1.2) with the help of the the following normalisation

$$\varepsilon_k = \frac{\sum_i S(k,i) e_i^{exc}}{\sum_i S(k,i)}. \tag{1.4}$$

[2]Theoretical justification of this procedure was presented by Baranger in [26]

From the above expression we can clearly see, that there are two types of contribution to the total error possible during extraction of single-particle states from experimental data. On the one hand we have the energy resolution of the used apparatus, on the other there is the determination of spectroscopic factors. They are calculated by dividing the measured cross section by the theoretical one determined for the pure state k:

$$\frac{d\sigma}{d\Omega}(\theta, E_i)\Big|_{\text{exp}} = S(k,i) \cdot \frac{d\sigma}{d\Omega}(\theta, E_k)\Big|_{\text{th}}. \tag{1.5}$$

This means that our 'experimental' objects are in fact theory dependent and require an important contribution from modelling of the reaction process. The situation gets even more complicated – the obtained results strongly vary depending on the chosen model of the reaction and its parameterization. Calculated spectroscopic factors can differ up to 30% between models, therefore in order to be certain of validity of extracted energy levels one must compare a vast number of reaction data. This is a hard and demanding task but also very important, especially for light nuclei. The number of identified levels in ^{40}Ca or ^{48}Ca is still very low which makes it almost impossible to exploit them for obtaining very reliable theoretical extrapolations.

The analysis of the process of extraction of single-particle levels of double-magic nuclei is a distinctive example of how important it is to take into account the experimental uncertainties when trying to build a theory upon them.

1.2 Theoretical Errors

Let us now consider what happens when we create a theoretical model to describe a physical phenomenon. Usually we end up with a set of parameters to be determined through a fitting procedure to some experimental data. If we now take into account experimental uncertainties of these data, we immediately see, that this will result in theoretical error bars for calculated quantities, but also obtained parameterization of the model will not be given as a set of numbers, but rather as a set of probability distributions. Furthermore, calculated errors do not need to correspond to its measured counterparts. This means that even if one experimental point has a very small uncertainty, we cannot expect the same for the predicted value – the width of its probability distribution can be quite the opposite. This is a problem of the modelling process, which can be more sensitive to some points than the others. The conclusion we get is simple: if the experimental data input is given by probability distributions the same applies to theoretical results. This fact should not be neglected when constructing theoretical models.

Another, actually more important issue related to the source of theoretical uncertainties, is the 'accuracy' of the theory, which is a major concern in nuclear physics. We are unable to precisely describe the interactions of many-body systems, therefore our theories are just approximations and as such, they cannot be expected to correctly predict every form of each nuclear phenomenon, actually yet unknown terms introduce extra theory errors.

1.3 Inverse Problem

Every modelling of experimental data, i.e. calculating parameters of the model out of measured values, is by construction an inexact process. This type of approach is often referred to as 'inverse problem', as opposed to 'forward' or 'direct' problem. The latter is simply predicting data based on exact, complete model, which is seldom the case in nuclear physics.

The difference between these two ways of functioning of one and the same model is rather crucial – within the inverse problem we do not have the complete description of the system and we use an incomplete experimental information about it to recover the parameters of modelisation. Another equally important difference is the number of solutions. The forward problem has one, well defined solution, whereas the inverse problem can have infinitely many. As an example let us consider the gravitational field around a massive object. If we know exactly the mass distribution inside the object, we can calculate the field around it – this is an example of a forward problem. On the other hand, if we have a set of measurements of the field, we can create a model of the mass distribution, but there can be more than one solution obtained from the same set of points. The number of solutions or, as in the case of nuclear physics, the number of parameterisations of the model, depend directly on the amount of experimental information we have. The more we already know, the closer we can get to reality in theoretical calculations. It is therefore crucial to extract as much as we can from the experiment.

Unfortunately when we consider the energy levels of protons and neutrons within the mean-field approach we do not have sufficient data to completely recover the form of potential. The main problem is, of course, that we do not know exactly how to parametrize it. Obviously if we chose a wrong type of function to represent some kind of interaction or a part of it, we cannot even hope to recover a meaningful information. This is similar to fitting linear function to data points received from exponential function. We can, of course, do it locally but the discrepancies between data point and results will be large and the model itself will be useless in terms of further modelling, especially extrapolations.

The presence of errors is a natural element when solving the inverse problem. The process of recovering the parameters is inexact and, by definition, theoretical uncertainties have to occur even if we do not include experimental errors. If a complete description of a system is unknown one cannot claim the calculated parameters and extrapolation to be exact. This seems to be very straightforward, but in many calculations the discussion of errors is simply omitted. Moreover, the possibility that obtained result is just one of many solutions is not even mentioned. The problem gets more serious, when the possibly erroneous calculations are extrapolated to describe other, not yet measured systems. In nuclear physics this can be related to predicted, but not yet found super-heavy or exotic nuclei. Numerous experiments are planned and performed based on theoretical predictions with little information on accuracy of these extrapolations.

The methods of solving the inverse problem are well developed and mathematically grounded. They are widely exploited in many scientific domains. Modelling of physical phenomena from limited sets of data is more common than one might think. We encounter inverse problems in astronomy – for example study of interior of stars, medical

imaging – tomography, analysis, prediction of earthquakes or even weather forecasts.

1.3.1 Mathematical Description of the Inverse Problem

The most general formulation of the forward problem, i.e. predicting data $\mathbf{d}$ (in our case: energy levels of a system) based on a given model $\mathbf{m}$[3] (in our case: a set of parameters of a Hamiltonian) can be given as a following expression

$$\mathbf{G}(\mathbf{m}) = \mathbf{d}, \tag{1.6}$$

where $\mathbf{G}$ is called the forward operator such as a Hamiltonian of a quantum system and represents the mathematical model of a system in question. It can be a function, an operator or a system of equations. The above expression can represent a continuos or discrete, both linear or non-linear problems.

As we have already said, experimental data always come with experimental errors, e.g. systematic instrumental errors or simply inaccuracies of the measurement. This means, that observed data $\mathbf{d}$ are not true values of observables with the help of which the true solutions of the model can be found, but instead we have

$$\mathbf{d} = \mathbf{G}(\mathbf{m}_{\text{true}}) + \nu = \mathbf{d}_{\text{true}} + \nu, \tag{1.7}$$

where ν is the uncertainty component. The recovered model $\mathbf{m}$ can be far from the real solution $\mathbf{m}_{\text{true}}$, even if the uncertainties are small. This usually happens, when the mathematical model is just an approximation of the reality or it is formulated in such a way, that some of its parameters are correlated causing the instability of the inverse problem. This aspect will be broadly discussed in this work due to its possible impact on nuclear theories.

In our study, we focus mainly on discrete linear problems[4], therefore equation (1.6) will be interpreted as a system of equations in a matrix representation. This implies that $\mathbf{G}$ is in general an $m \times n$ matrix, $\mathbf{m}$ is an n-element vector of model parameters and $\mathbf{d}$ is the data vector with m entries. The inverse problem could be described formally by

$$\mathbf{m} = \mathbf{G}^{-1}\mathbf{d}, \tag{1.8}$$

but only if $\mathbf{G}$ is square and if the inverse matrix exists. Under such conditions the system of equations gives a unique solution, which on the other hand renders it useless in most of the actual applications, where we deal with a small number of parameters or lack of sufficient experimental information. This formulation can be applied only in a small part of calculations.

If, however, $\mathbf{G}$ is a rectangular matrix and the inverse matrix does not exist, one can multiply from the left by $\mathbf{G}^{\mathrm{T}}$, which results in a set of the, so-called, normal equations

[3] $\mathbf{m}$ could symbolise a mathematical model or a parameter set.

[4] In fact, we are dealing with non-linear least-squares problem, which, however, can be linearised with the help of Taylor expansion. This property justifies the use linear inverse problems in accordance with many publications in this domain of applied mathematics.

$$\mathbf{m} = (\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{d}. \tag{1.9}$$

This time we are also dealing with an inversion of a square matrix, but still one problem remains. The column rank of $\mathbf{G}^{\mathrm{T}}\mathbf{G}$, say, r, i.e. the number of its linearly independent column vectors, must be n for this matrix to be invertible. Once again, this is not always the case – usually we are dealing with *rank-deficient* problems with rank $r < n$, which do not have one unique solution. This stems from the fact, that there exists a nontrivial $(n - r)$-dimensional null space of $\mathbf{G}$, denoted as $N(\mathbf{G})$. It is given by

$$\mathbf{G}\mathbf{m}_0 = \mathbf{0}, \tag{1.10}$$

where $\mathbf{m}_0$ denotes all models, which belong to the null space. Any linear combination of elements from $\mathbf{m}_0$ does not change the data $\mathbf{d}$. Thus when we solve the inverse problem, we may obtain infinitely many mathematically equivalent models from one data set, which have different admixtures of the null space solutions. Fortunately, there are ways of dealing with this problem, for example singular value decomposition, which plays an essential role in this work. Its basic properties will be presented in the next section.

Usually, inverse problems are solved by minimising the expression

$$\delta = \|\mathbf{G}\mathbf{m} - \mathbf{d}\|_2, \tag{1.11}$$

which is basically a general formulation of the least-squares problem, with $\|\cdot\|_2$ denoting the L^2 norm, defined as

$$\|x\|_2 = \sqrt{\sum_i |x_i|^2}. \tag{1.12}$$

This approach allows us to avoid the explicit use of the inverse of matrix $\mathbf{G}$.

Another important feature of inverse problems is the stability of solutions in terms of data perturbation. It often happens, that the solution of the inverse problem are sensitive to 'noise' contamination of $\mathbf{d}$, i.e. a small admixture of ν can cause a large change in obtained model.[5] Inverse problems with this sort of behaviour are called *ill-posed* or *ill-conditioned*. The first term is sometimes reserved only for continuous systems, whereas the latter refers to discrete ones. One can also encounter a definition where 'ill-posedness' means having more than one solution, but we will use this term to describe unstable solutions.[6] Usually, it is possible to stabilize the calculations by implementing one of various regularization methods, which consist of imposing additional constraints. This process will be explained later in this chapter along with an example of Tikhonov regularization. The so-called truncated singular value decompositon, presented in the next section, is also a way to stabilize the inverse problem.

[5] In what follows the word 'model' should be understood as the set of parameters of the Hamiltonian

[6] Actually, the term 'well-posed' was introduced by the French mathematician J.S. Hadamard, who gave the criteria for a problem to be well defined, namely the solution must exist, it must be unique and stable. From this we see, that both unstable and non-unique problems may as well be called ill-posed.

1.4 Singular Value Decomposition

The process of solving the inverse problem in general includes inversion of the matrix $\mathbf{G}$. It often happens, that this matrix is singular and cannot be inverted. As a remedy, we can exploit the singular value decomposition (SVD), which is one of methods for decomposing a matrix, widely exploited in statistics. Its properties are very useful when it comes to determining the predictive power of a model – it allows us to easily identify instabilities related to the $\mathbf{G}$ matrix.

Let us first consider the formal description of this method. Every real matrix $\mathbf{A}$ can be decomposed in the following way (cf. [4], p.55, Eq. (4.1))

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathrm{T}}, \tag{1.13}$$

where, if $\mathbf{A}$ is an $m \times n$ real matrix, then $\mathbf{U}$ is $m \times m$, $\mathbf{D}$ is $m \times n$ and $\mathbf{V}$ is $n \times n$.[7] These three matrices have specific properties: $\mathbf{U}$ and $\mathbf{V}$ are orthogonal (or have orthogonal columns), so that

$$\left.\begin{aligned} \mathbf{U}^{\mathrm{T}}\mathbf{U} &= \mathbb{I} \\ \mathbf{V}^{\mathrm{T}}\mathbf{V} &= \mathbb{I}. \end{aligned}\right\} \tag{1.14}$$

The columns of $\mathbf{U}$ are eigenvectors of matrix $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ and similarly columns of $\mathbf{V}$ are eigenvectors of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$. The eigen-values in both cases are the same, they are the square roots of the diagonal entries of $\mathbf{D}$. In fact, $\mathbf{D}$ is a diagonal matrix, which can be represented as

$$\mathbf{D} = \mathrm{diag}\{d_1, d_2, \ldots, d_{\min(m,n)}\}. \tag{1.15}$$

The entries d_i are nonnegative; they are called singular values of $\mathbf{A}$. They fulfil the conditions

$$\mathbf{A}\mathbf{v} = d\,\mathbf{u} \quad \text{and} \quad \mathbf{A}^{\mathrm{T}}\mathbf{u} = d\,\mathbf{v}, \tag{1.16}$$

where $\mathbf{u}$ and $\mathbf{v}$ are the columns of matrices $\mathbf{U}$ and $\mathbf{V}$ respectively.

If matrix $\mathbf{G}$ is singular, only r first diagonal values of $\mathbf{D}$ are nonzero with r being the rank of $\mathbf{G}$. This allows us to decompose this matrix into

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{1.17}$$

where $\mathbf{D}_r$ is $r \times r$ diagonal matrix. Using this we can express the decomposition of $\mathbf{G}$ as a following expression

[7] For completeness we note here that there exist in the literature another notational convention, where $\mathbf{U}$ is $m \times n$ matrix, $\mathbf{D}$ is $n \times n$ and $\mathbf{V}$ is $n \times n$, cf. [8], but this is often referred to as an example of reduced SVD, which we will discuss later. In both approaches, however, the properties of the matrices are similar.

$$\mathbf{G} = \begin{bmatrix}\mathbf{U}_r, & \mathbf{U}_0\end{bmatrix} \begin{bmatrix}\mathbf{D}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{V}_r, & \mathbf{V}_0\end{bmatrix}^{\mathrm{T}}, \tag{1.18}$$

where $\mathbf{U}_r$ denotes the first r significant columns of $\mathbf{U}$, $\mathbf{U}_0$ stands for last $m-r$ columns. Similarly, $\mathbf{V}_r$ denotes the first r columns and $\mathbf{V}_0$ marks the last $n-r$ columns of $\mathbf{V}$.

The columns in $\mathbf{U}$ form an r-dimensional orthonormal basis of the range of $\mathbf{G}$ denoted as $R(\mathbf{G})$, whereas columns of $\mathbf{V}_0$ are the null space basis of $\mathbf{G}$. On the other hand the columns of $\mathbf{U}_0$ span the null space $N(\mathbf{G}^{\mathrm{T}})$ and $\mathbf{V}_r$ spans the range $R(\mathbf{G}^{\mathrm{T}})$.

1.4.1 Reduced Singular Value Decomposition

The full version of the singular value decomposition is the least useful in practical applications. Usually we have $m \neq n$, so automatically the $\mathbf{D}$ matrix has only $\min(m,n)$ diagonal values, the rest are '0'. This allows us to reduce the problem to the so-called *thin* SVD[8], which can be written as

$$\mathbf{A} = \mathbf{U}_{\min(m,n)}\mathbf{D}_{\min(m,n)}\mathbf{V}^{\mathrm{T}}_{\min(m,n)}, \tag{1.19}$$

which means that $\mathbf{D}_{\min(m,n)}$ is now a square matrix, and the larger one of $\mathbf{U}$ and $\mathbf{V}$ is cut to $m \times n$ size. This version of SVD is especially useful in numerical calculations when there is a large difference between m and n.

Another, more reduced option is *compact* SVD. If rank r of matrix $\mathbf{A}$ is lower than $\min(m,n)$ we can decompose $\mathbf{A}$ in the following way

$$\mathbf{A} = \mathbf{U}_r\mathbf{D}_r\mathbf{V}^{\mathrm{T}}_r, \tag{1.20}$$

where $\mathbf{D}_r$ is now a square matrix of size $r \times r$, $\mathbf{U}_r$ is $m \times r$ and $\mathbf{V}_r$ is $n \times r$.

Truncated SVD. The most common reduction of singular value decomposition is the low-rank approximation called *truncated* SVD. It is similar to compact SVD, but the number of singular values we take into account is smaller than the rank of $\mathbf{A}$. This approach can easily be justified. The singular values of a matrix vary in the order of magnitude. The smallest are often the cause of instability of the solution, but on the other hand they contribute the least to the total SVD. Rows of $\mathbf{U}$ and columns of $\mathbf{V}$ multiplied by a small number have low impact on entries of resulting matrix, thus it is safe to omit the smallest diagonal entries of $\mathbf{D}$. Moreover, in numerical calculations it often happens, that some of the non-zero singular values have strong contribution from the rounding errors.

The TSVD can be written as

$$\mathbf{A}_t = \mathbf{U}_t\mathbf{D}_t\mathbf{V}^{\mathrm{T}}_t, \tag{1.21}$$

[8]It is necessary here to mention that this naming convention, consistent with e.g. Chapter 6 of [6], is just one of many used in the literature. The most common is one definition of reduced or compact SVD, cf. [4], pp. 56.

where $t < r$ represents the number of singular values we subjectively decide to take as significant, $\mathbf{D}_t$ is a $t \times t$ matrix, $\mathbf{U}_t$ is $m \times t$ and $\mathbf{V}_t$ is $n \times t$.

The resulting matrix $\mathbf{A}_t$ is only an approximation of $\mathbf{A}$, but it is still the best approximation among the matrices of the rank t. This statement is known as the Eckart-Young theorem (1936). It can be proven in terms of the so-called Frobenius norm, which is defined as a square root of a sum of absolute values of all the entries of a $m \times n$ matrix, i.e.

$$\|\mathbf{A}\|_F \equiv \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}, \tag{1.22}$$

but it can also be represented in terms of the trace of $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ and singular values of $\mathbf{A}$

$$\|\mathbf{A}\|_F = \sqrt{Tr[\mathbf{A}\mathbf{A}^{\mathrm{T}}]} = \sqrt{\sum_{i=1}^{\min(m,n)} d_i^2}. \tag{1.23}$$

In line with the mentioned theorem, in order to find a matrix $\mathbf{B}_{m\times n}$ with rank t, that is the best approximation of $\mathbf{M}_{m\times n}$ with rank $r \geq t$, we need to minimize the differences between the entries of $\mathbf{M}$ and $\mathbf{B}$, namely

$$\min_{\mathbf{B}} \|\mathbf{M}-\mathbf{B}\|_F = \min_{b_{ij}} \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|m_{ij}-b_{ij}|^2}. \tag{1.24}$$

If we now assume that singular value decomposition of $\mathbf{M}$ is $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathrm{T}}$ and exploit the fact that Frobenius norm is unitarily invariant, i.e. $\|\mathbf{M}\|_F = \|\mathbf{U}^{\mathrm{T}}\mathbf{M}\mathbf{V}\|_F$, we may write

$$\min_{\mathbf{B}} \|\mathbf{M}-\mathbf{B}\|_F = \min_{\mathbf{B}} \|\mathbf{D}-\mathbf{U}^{\mathrm{T}}\mathbf{B}\mathbf{V}\|_F \tag{1.25}$$

To minimize this expression, $\mathbf{U}^{\mathrm{T}}\mathbf{B}\mathbf{V}$ needs to be diagonal since $\mathbf{D}$ is a diagonal matrix. This is possible if the singular value decomposition of $\mathbf{B}$ can be written as $\mathbf{B} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}$, where $\mathbf{S}$ is a diagonal matrix with entries being the singular values of $\mathbf{B}$. Assuming this, we obtain

$$\min_{\mathbf{B}} \|\mathbf{M}-\mathbf{B}\|_F = \min_{\mathbf{B}} \|\mathbf{D}-\mathbf{S}\|_F = \min_{s_i} \sqrt{\sum_{i=1}^{r}(d_i-s_i)^2}. \tag{1.26}$$

If we now recall, that the rank of $\mathbf{B}$ is $t \leq r$ we finally have

$$\min_{\mathbf{B}} \|\mathbf{M}-\mathbf{B}\|_F = \min_{s_i} \sqrt{\sum_{i=1}^{t}(d_i-s_i)^2 + \sum_{i=t+1}^{r} d_i^2}. \tag{1.27}$$

The minimum is obtained when $s_i = d_i; i = 1, \ldots, t$, therefore matrix $\mathbf{S}$ needs to be composed of exactly t largest singular values of $\mathbf{M}$. This means that even if we choose to ignore the smallest entries of $\mathbf{D}$ to obtain better stability of the solution, we still have the best approximation of the original matrix in the range of t-rank matrices. However, one needs to remember that by removing the non-zero singular values we loose precious information, thus the cut-off level has to be chosen very carefully. Unfortunately there is no rule when it comes to this truncation, we have to choose between stability and precision.

1.5 Inverse Problem in Terms of SVD

As it was already said, singular value decomposition can be used to solve the linear inverse problem. If matrix $\mathbf{G}$ is singular, we can exploit the compact SVD defined in eq. (1.20) and perform the so-called *generalised* inverse, or *Moore-Penrose pseudoinverse*. It is defined as

$$\mathbf{G}^{\dagger} = \mathbf{V}_r \mathbf{D}_r^{-1} \mathbf{U}_r^{\mathrm{T}}. \tag{1.28}$$

Now we can write the pseudoinverse solution as

$$\mathbf{m}^{\dagger} = \mathbf{G}^{\dagger}\mathbf{d} = \mathbf{V}_r \mathbf{D}_r^{-1} \mathbf{U}_r^{\mathrm{T}} \mathbf{d}. \tag{1.29}$$

In the index notation we have the following expression

$$m_i^{\dagger} = \sum_{j=1}^{r} \frac{1}{d_j} V_{ij} \sum_{k=1}^{m} U_{jk}^{T} d_k. \tag{1.30}$$

This approach has a very desirable property, i.e. the pseudoinverse always exists and so does the solution. In addition, it can be shown, that the generalised inverse solution is the least-squares solution of minimum length. It means that it minimises both $\|\mathbf{Gm}-\mathbf{d}\|_2$ and $\|\mathbf{m}\|_2$. The last property is a way of choosing the desired solution from the infinite set of solution minimising the first condition. This approach is justified, because any uncertainty admixture increases, in principle, the norm of $\mathbf{m}$. Selecting the solution with the smallest length ensures that the result is as close to the true solution as it can be, given the formulation of the problem. Of course, this is not the only way of solving the problem of infinite number of solutions, but it offers a rather good compromise between the size of the uncertainties and the exactness of the chosen fit.

1.6 Regularization Methods

Truncated singular value decomposition is just one solution for ill-posed inverse problems. Another way of dealing with instability of the calculations are regularisation methods. As we already explained, the instability arises from small singular values, which makes the model very sensitive for small perturbations of data. Thanks to

the specific mathematical form of the problem we can just neglect those values, as in TSVD, which unfortunately changes the precision; or we can introduce an additional parameter to scale the singular values.

1.6.1 Tikhonov Regularization

Tikhonov regularization is probably the most popular method used for linear inverse problems. Its basic assumption is that instead of minimising equation (1.11), we extend it by an additional term

$$\min \left(\|\mathbf{Gm} - \mathbf{d}\|_2^2 + \|\Gamma \mathbf{m}\|_2^2 \right), \tag{1.31}$$

where Γ is the so-called Tikhonov matrix, chosen in such a way as to prefer solutions with given properties. Most common choice of Γ is the identity matrix multiplied by some additional parameter, which means that solutions with the smallest norm of $\mathbf{m}$ are preferable. This version is called the damped least-squares problem and is defined as

$$\min \left(\|\mathbf{Gm} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2 \right), \tag{1.32}$$

where α is called the regularization parameter.

Tikhonov regularization can be easily expressed in terms of singular value decomposition. The definition of Euclidean norm allows to rewrite the equation (1.32) as the following matrix expression

$$\min \left\| \begin{bmatrix} \mathbf{G} \\ \alpha \mathbf{I} \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_2^2, \tag{1.33}$$

where we put $\Gamma = \mathbf{I}$ as the simplest case. For $\alpha \neq 0$ we can solve it as a full-rank least-squares problem, thus we have

$$\begin{bmatrix} \mathbf{G}^{\mathrm{T}} & \alpha \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{G} \\ \alpha \mathbf{I} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{G}^{\mathrm{T}} & \alpha \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}, \tag{1.34}$$

from where one obtains

$$(\mathbf{G}^{\mathrm{T}}\mathbf{G} + \alpha^2 \mathbf{I})\mathbf{m} = \mathbf{G}^{\mathrm{T}}\mathbf{d}. \tag{1.35}$$

The above expression is the *zeroth-order Tikhonov regularization* of our initial problem. It can be further developed by decomposing $\mathbf{G}$ using SVD, which yields

$$(\mathbf{V}\mathbf{S}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} + \alpha^2 \mathbf{I})\mathbf{m} = \mathbf{V}\mathbf{S}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{d}, \tag{1.36}$$

or, equivalently

$$(\mathbf{V}\mathbf{S}^{\mathrm{T}}\mathbf{S}\mathbf{V}^{\mathrm{T}} + \alpha^2 \mathbf{I})\mathbf{m} = \mathbf{V}\mathbf{S}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{d}. \tag{1.37}$$

Solution of this equation is unique and in the index notation it is given as follows

$$m_{\alpha,i} = \sum_{j=1}^{\min(m,n)} \frac{s_j^2}{s_j^2 + \alpha^2} \frac{V_{ij}}{s_j} \sum_{k=1}^{m} U_{jk}^T d_k. \tag{1.38}$$

One introduces a specific notation,

$$f_i = \frac{s_i^2}{s_i^2 + \alpha^2}, \tag{1.39}$$

f_i called *filter factors*. If $s_i \gg \alpha$ then $f_i \approx 1$; if, to the contrary we have $s_i \ll \alpha$ we have $f_i \approx 0$. If we recall eq. (1.30), we notice that both expressions differ only by the factor f_i. This means that the smallest singular values, the most troublesome ones, will have reduced impact on the solution as compared to (1.30), because they are multiplied by a very small number. This method is smoother than truncated SVD, but also, what is more important, it allows to keep the original dimension of $\mathbf{G}$.

1.7 Role of Probability in the Inverse Problems

Before we begin, it would be helpful to introduce the notation by recapitulating some basic concepts of probability theory.

When speaking of probability, one usually thinks of the rate of occurrence of some event in a given sample. By sample we mean a subset of the population, i.e. a part of the whole set of objects or events under consideration. This classical definition is rather intuitive, but we will base our considerations on another, more modern, approach, which exploits measure theory.

1.7.1 Definition and Properties of Probability Function

In a given space Ω (it can be the data or model space) we can define a σ-field Σ of all its subsets, composed of either single points (e.g. points in the data space composed of set of numbers representing various outcomes of series of, say, m measurements) or collections of points. A σ-field is a collection of sets that has specific properties: it has at least one element, the complement $(\Omega \setminus A)$ of a subset A belongs to the field and also all possible unions of subsets belong to the field. These subsets are called 'events' and are denoted with capital letters A, B, C, ...

We can define a measure $P(\cdot)$ over the space Ω, that associates a positive real number to an event A. This measure is the probability measure over Ω, if it satisfies the following Kolmogorov axioms

1. Probability is a nonnegative number, $P(A) \geq 0$ for all $A \in \Sigma$,
2. Probability is normalised, its value over the space is $P(\Omega) = 1$,
3. The probability of a sum of pairwise disjoint events $P(A_i)$ is equal to the sum of individual probabilities, $P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$.

One can observe, that additional properties of the probability function arise from the above statements. First of all, $P(\cdot)$ is a monotonic function, i.e. if one subset is contained in another, its probability is less or equal than the probability of the larger subset. This can be formally written as

$$P(A) \leq P(B) \leq P(C) \ldots \text{ if } A \subseteq B \subseteq C \ldots. \tag{1.40}$$

This statement also implies that the lower bound of P is

$$P(\emptyset) = 0, \tag{1.41}$$

thus for any event $A \in \Omega$ we have the limits of probability

$$0 \leq P(A) \leq 1. \tag{1.42}$$

For event A and its complement $(\Omega \setminus A)$ we get the relation

$$P(A) + P(\Omega \setminus A) = 1. \tag{1.43}$$

Finally, if two events are not disjoint, we may write the sum rule as

$$P(A \cup B) = P(A) + P(B) - P(A \cap B), \tag{1.44}$$

where $P(A \cap B)$ denotes the probability of the intersection of subsets A and B.

1.7.2 Probability Density

Probability distribution over the space Ω can be represented as an integral

$$P(A) = \int_A f(\mathbf{x}) d\mathbf{x}, \tag{1.45}$$

where $\mathbf{x} = \{x_1, x_2, x_3, \ldots\}$ denotes the coordinate system and the integration is over the subset A. The function $f(\mathbf{x})$ is called the *probability density function* (PDF) and its form depends on the coordinate system we choose. In fact, this quantity has a physical dimension, which is inverse to the dimension of the volume element of the chosen space as opposed to the dimensionless value of the probability function itself. This immediately implies, that it depends on the coordinate system. The change of coordinates can be written as

$$\mathbf{x}' = \mathbf{x}'(\mathbf{x}), \tag{1.46}$$

where $\mathbf{x}'$ and $\mathbf{x}$ are the new and old coordinates respectively. If by $f'(\mathbf{x}')$ we denote the probability density in the new coordinates, then we have for the probability distribution

$$P(A) = \int_A f'(\mathbf{x}') d\mathbf{x}'. \tag{1.47}$$

However, from the definition (1.45), for any $A \in \Omega$ we have

$$\int_A f(\mathbf{x})d\mathbf{x} = \int_A f'(\mathbf{x}')d\mathbf{x}'. \tag{1.48}$$

Recalling that the change of coordinates can be written as

$$d\mathbf{x} = |J(\mathbf{x}')|\, d\mathbf{x}', \tag{1.49}$$

where $J(\mathbf{x}')$ is determinant of the Jacobian matrix, we get as the result

$$f'(\mathbf{x}') = |J(\mathbf{x}')|\, f(\mathbf{x}). \tag{1.50}$$

1.7.3 Expectation and Variance

The *expectation* or *mean* of a vector variable $\mathbf{x}$ can be defined using the probability density function. For each component of $\mathbf{x}$, it is given as the weighted average over the whole space Ω, i.e.

$$\mathrm{E}[x_i] = \int x_i f(\mathbf{x})d^N x, \tag{1.51}$$

where $\mathrm{E}[\mathbf{x}]$ denotes the expectation of $\mathbf{x}$ and $f(\mathbf{x})$ is the probability density function defined for the variable $\mathbf{x}$, which plays the role of the weight factors. Actually, one can calculate the expectation of any function, say $g(\mathbf{x})$, as

$$\mathrm{E}[g(\mathbf{x})] = \int g(\mathbf{x})f(\mathbf{x})d^N x. \tag{1.52}$$

For the mean of $\mathbf{x}$ we will reserve a symbol '$\boldsymbol{\mu}$'. This is a multidimensional case, where $\boldsymbol{\mu}$ is a vector.

The *standard deviation* or *variance* of $\mathbf{x}$ is defined as

$$\sigma_i^2 \equiv \mathrm{E}[(x_i - \mu_i)^2] = \int (x_i - \mu_i)^2 f(\mathbf{x})d^N x. \tag{1.53}$$

Exploiting the properties of the expectation, we can expand this expression in the following way

$$\sigma_i^2 \equiv \mathrm{E}[(x_i - \mu_i)^2] = \mathrm{E}[x_i^2] - (\mathrm{E}[x_i])^2. \tag{1.54}$$

All these considerations can be easily transfered into a one-dimesional case with only one variable x or a discrete case, where we are dealing with finite number of possible values of $\mathbf{x}$.

1.7.4 Covariance Matrix

In chapter 1.3.1 we have already briefly mentioned the correlation of parameters of the mathematical model as one of the reasons of ill-posedness of the inverse problem. This undesirable property can unfortunately be hidden and unnoticeable at first glance. The correlations can be established by performing straightforward tests of behaviour of the model under modification of the parameters[9]. There also exist an analytical tool enabling quantitive description of correlations.

Having introduced the expectation in eq.(1.51), one can define the so-called *covariance matrix* element as follows

$$C_{ij} = \mathrm{Cov}(x_i, x_j) = \int (x_i - \mu_i)(x_j - \mu_j) f(\mathbf{x}) d\mathbf{x}, \tag{1.55}$$

where the diagonal entries C_{ij} are the variances σ_i^2. In matrix notation this definition takes the form

$$\mathbf{C} = \int (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T f(\mathbf{x}) d\mathbf{x}. \tag{1.56}$$

The covariance matrix is symmetric, and for any vector $\mathbf{x}$ it has a following property

$$\mathbf{x}^* \mathbf{C}^{-1} \mathbf{x} \geq 0, \tag{1.57}$$

i.e. it is definite nonnegative[10]. In the above expression $\mathbf{x}^*$ denotes the conjugate transpose of $\mathbf{x}$.

This property means that eq.(1.57) can be considered as a norm and thus the expression

$$\|x\| = \mathbf{x}^* \mathbf{C}^{-1} \mathbf{x}, \tag{1.58}$$

is the weighted l_2-norm of a vector $\mathbf{x}$.

We can also define the so-called correlation coefficients ρ_{ij}. They can be expressed in terms of the covariance matrix, and are simply defined as

$$\rho_{ij} = \frac{C_{ij}}{\sigma_i \sigma_j}. \tag{1.59}$$

They measure the relative strength of the correlation between the components of $\mathbf{x}$ (e.g. a set of paramaters of a Hamiltonian or a set of experimental nuclear single-particle levels). The correlation coefficients take the values

$$-1 \geq \rho_{ij} \leq 1. \tag{1.60}$$

Let us now consider the inverse problem defined in eqs. (1.8) and (1.9). In this case, the covariance matrix of the parameter set $\mathbf{m}$ can be calculated as follows[11]

[9] This method will be presented in Chapter 2

[10] c.f. [5], pp. 172.

[11] c.f. [4], pp. 264, Theorem B.5.

$$\mathrm{Cov}(\mathbf{m}) = \mathrm{Cov}(\mathbf{Ad}) = \mathbf{ACA}^{\mathrm{T}}, \tag{1.61}$$

where $\mathbf{A}$ stands for either $\mathbf{G}^{-1}$ (if it is invertible), or for $(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}$ (if it is not) and $\mathbf{C}$ denotes the covariance matrix of $\mathbf{d}$. This expression is valid if $\mathbf{d}$ has a normal distribution.[12] Moreover, this implies that $\mathbf{Ad}$ also has a normal distribution.

Let us focus on the case of normal equations, i.e. the solution of the least-squares problem. We can write the full expression for the covariance matrix as

$$\mathrm{Cov}(\mathbf{m}) = (\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{C}\mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}. \tag{1.62}$$

Now, if we consider the simplest case for the covariance matrix of $\mathbf{d}$, i.e.

$$\mathbf{C} = \sigma^2\mathbf{I}, \tag{1.63}$$

which means that the experimental data are uncorrelated and have the same standard deviation σ, we get

$$\mathrm{Cov}(\mathbf{m}) = \sigma^2(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}. \tag{1.64}$$

The covariance matrix is not necessarilly diagonal. In fact, the model elements (parameters) are linear combinations of data, therefore it is only natural to assume their correlation and occurrence of the non-diagonal entries of the covariance matrix.

In a case of rank-deficient inverse problem, we may use the Singular Value Decomposition. Recalling the generalised inverse solution $\mathbf{m}^{\dagger}$ given by eq. (1.29), we can write

$$\mathrm{Cov}(\mathbf{m}^{\dagger}) = \mathbf{G}^{\dagger}\mathbf{C}(\mathbf{G}^{\dagger})^{\mathrm{T}} = \sigma^2\mathbf{G}^{\dagger}(\mathbf{G}^{\dagger})^{\mathrm{T}}. \tag{1.65}$$

Inserting the decomposition of $\mathbf{G}^{\dagger}$ we obtain

$$\mathrm{Cov}(\mathbf{m}^{\dagger}) = \sigma^2\mathbf{V}_r\mathbf{D}_r^{-2}\mathbf{V}_r^{\mathrm{T}}, \tag{1.66}$$

and in the index notation

$$\mathrm{Cov}(m_i^{\dagger}, m_j^{\dagger}) = \sigma^2\sum_{k=1}^{r} V_{ik}\frac{1}{d_k^2}V_{kj}^{T}. \tag{1.67}$$

It is important to notice, that since the part of $\mathbf{V}$ spanning the null space is excluded from the summation, we obtain in general a nonsingular solution. If, however, it happens, that some of the remaining diagonal values of $\mathbf{D}$ are small enough to cause problems, we can perform one of regularisation methods to avoid instabilities of the calculations.

[12] More precisely, it is a Multivariate Normal Distribution, which is simply a generalisation of one-dimensional normal distribution into higher dimensions.

1.7.5 Confidence Intervals

Another important concept of statistics is the notion of *confidence intervals* (CI). They are calculated from the samples of the population at a givent *confidence level*, which specifies how certain of the true value of a measured variable being inside the interval we wish to be. In other words, calculating the confidence intervals with confidence level, say, $(1-p)$ means that with probability $P = 1-p$ this interval will contain the true mean value. We can write this formally as

$$P(x_1 \leq x \leq x_1) = 1 - p, \tag{1.68}$$

where $(1-p)$ is the confidence level, and x_1, x_2 denote the limits of the confidence interval.

Usually the information about the population is limited, which implies, that the data we possess cannot be treated as having a normal distribution. Instead, we can use a Student's t-distribution, which has the probability density function defined as[13]

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \tag{1.69}$$

where ν denotes the number of degrees of freedom and Γ is the Gamma function. The Student's distribution has a similar shape to normal distribution, but it has 'fatter' tails, which means that more points lie far from the mean value.

To recover the confidence interval of x of a sample of n elements with the mean $\bar{x}$ we define the following variable

$$z \equiv \frac{x - \bar{x}}{s/\sqrt{n}}, \tag{1.70}$$

which has a t-distribution with $(n-1)$ degrees of freedom. In the above, s is the standard deviation of the sample. Using this we can write

$$P(t_{n-1,\frac{p}{2}} \leq \frac{x - \bar{x}}{s/\sqrt{n}} \leq t_{n-1,1-\frac{p}{2}}) = 1 - p, \tag{1.71}$$

where $t_{\nu,\alpha}$ is the critical t value, i.e. an argument of t-distribution with ν degrees of freedom corresponding to the borders of the area of given probability α around the mean.

The $(1-p)$ confidence interval can be thus represented as

$$\mathrm{CI} = \left(\bar{x} + \frac{t_{n-1,\frac{p}{2}}\, s}{\sqrt{n}}, \bar{x} + \frac{t_{n-1,1-\frac{p}{2}}\, s}{\sqrt{n}}\right). \tag{1.72}$$

[13] cf. [4], pp. 256

1.7.6 Likelihood Function

Let us consider the inverse problem given by equation (1.8) from the statistical point of view. We can define the probability density function depending on a given model $\mathbf{m}$ for each data component d_i, e.g. each nuclear single-particle level measured in the experiment. This PDF will be denoted as $f_i(d_i|\mathbf{m})$.

The joint probability density function for the whole data set will be given by

$$f(\mathbf{d}|\mathbf{m}) = f_1(d_1|\mathbf{m}) \cdot f_2(d_2|\mathbf{m}) \dots f_m(d_m|\mathbf{m}). \tag{1.73}$$

It can be considered as a function of parameters $\mathbf{m}$, which gives us the probability of reproducing the data set $\mathbf{d}$ depending on parameters $\mathbf{m}$. Therefore we use a specific notation

$$\mathcal{L}(\mathbf{m}|\mathbf{d}) = f(\mathbf{d}|\mathbf{m}), \tag{1.74}$$

and call this expression the *likelihood function* (or simply *likelihood*). It helps us to determine, whether a given model is likely to fit the data. A simple criterion for the selection of plausible models has been formulated. It is called the *maximum likelihood principle*. The basic assumption for this criterion is that the best parameter set is the one, that maximizes the likelihood function, therefore one simply needs to calculate

$$\frac{\partial \mathcal{L}(\mathbf{m}|\mathbf{d})}{\partial m_i} = 0, \tag{1.75}$$

for each parameter m_i.

1.8 Conditional Probability and Bayes' Theorem

From the poinf of view of this work, one of the most useful notion in probability theory is the Bayes' law originally proposed by Thomas Bayes in the XVIII century. Its contemporary formulation allows us to calculate the so-called *posterior probability*, directly connected with the problem discussed in this section. The term 'bayesian' refers to a whole branch of statistics based on a specific interpretation of probability started by Bayes and Laplace. The probability is generally interpreted as an occurrence of a given event in a large sample, which is a rather intuitive approach. In terms of bayesian statistics, however, we speak of a confidence or a belief that a given 'proposition' is true. The actual meaning of this statement will become clearer once the Bayes' law is shortly recalled.

To introduce all the necessary concepts it will be instructive to consider the following example.

Example. We draw a ball from one of two urns. The probability of drawing from the first one is two times smaller than from the other. In the first urn there are three black balls and one white, whereas in the second one there are two balls of each colour. What is the probability that if the drawn ball was black, it came from the first urn?

The probability of drawing any ball from the first urn U_1 is $P(U_1) = \frac{1}{3}$, and from the second $P(U_2) = \frac{2}{3}$. The probability of drawing a black ball from the first urn is $P(b|U_1) = \frac{3}{4}$, and from the urn U_2 it is $P(b|U_2) = \frac{1}{2}$. This is the *conditional* probability, that the event 'drawing a black ball' occurs at the condition that 'it is drawn from the urn number (1)'. Similarly, we can write $P(w|U_1) = \frac{1}{4}$ and $P(w|U_2) = \frac{1}{2}$ for the white ball. What we are looking for is the conditional probability $P(U_1|b)$, that the ball was drawn from the first urn, provided it was black.

The conditional probability is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{1.76}$$

where $P(A \cap B)$ is the *joint* probability of A and B. In our example $P(b \cap U_1)$ means the probability of drawing a black ball *and* drawing it from the first urn.

Transforming the above definition to calculate the joint probability, we get

$$P(A \cap B) = P(A|B)P(B). \tag{1.77}$$

We can also rewrite this expression in a reversed order as

$$P(B \cap A) = P(B|A)P(A), \tag{1.78}$$

but, evidently, $P(A \cap B) = P(B \cap A)$. This allows us to write the following relation

$$\boldsymbol{P(B|A) = \frac{P(A|B)P(B)}{P(A)}}, \tag{1.79}$$

which is known as the *Bayes' theorem*. $P(B|A)$ is called the *posterior* probability, also reffered to as the 'probability of the causes'. This means that given a result A, we calculate the probability of its cause B. $P(A|B)$ is sometimes reffered to as the *likelihood*. $P(A)$ and $P(B)$ are called the *prior-* or *marginal-* probabilities of A and B respectively.

In our example we wish to calculate

$$P(U_1|b) = \frac{P(b|U_1)P(U_1)}{P(b)}. \tag{1.80}$$

$P(b)$ is the probability of drawing a black ball from the urn, regardless of which one we choose. This quantity can be calculated from the law of total probability

$$P(A) = \sum_n P(A, B_n) = \sum_n P(A|B_n)P(B_n), \tag{1.81}$$

where B_n denotes a set of pairwise disjoint events $B_i \cap B_j = \emptyset$, whose union is the whole space.

According to this, we immediately have

$$P(b) = P(b|U_1)P(U_1) + P(b|U_2)P(U_2) = \frac{3}{4}\cdot\frac{1}{3} + \frac{1}{2}\cdot\frac{2}{3} = \frac{7}{12}, \tag{1.82}$$

and finally the posterior probability that the chosen urn was U_1, if drawn ball was black, becomes

$$P(U_1|b) = \frac{3}{4}\cdot\frac{1}{3} : \frac{7}{12} = \frac{3}{7}. \tag{1.83}$$

If we now combine equations (1.79) and (1.81), we obtain a generalised expression

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_n P(A|B_n)P(B_n)}. \tag{1.84}$$

As it was said in the introduction of this section, the Bayes' theorem was a starting point to a new interpretation of probability. The equation (1.79) can be simply viewed as a relation between the conditional probability $P(A|B)$ and its inverse $P(B|A)$, valid regardles of the approach we choose; but it can also be given a deeper meaning.

Let us suppose that a physical phenomenon can be described by several theoretical models, but based on our knowledge we are unable to determine the correct one. We can estimate our belief or confidence in a chosen model relying on partial evidence provided by acquired data, for example measurement results. This confidence is exactly the posterior probability from the Bayes' theorem and its value changes with every additional data we consider. We can compare this to a learning process, and actually bayesian statistics is exploited in developing artificial intelligence protocols. In fact every process requiring decision making based on incomplete or uncertain information can be programmed and solved in terms of bayesian statistics. Moreover, many cases of inverse problems are treated within this framework.

1.9 Least-Squares Approach as a Form of Inverse Problem

In our study we use a linearised method of least-squares. The parameters 'p', of the mean-field Hamiltonian are recovered by minimisation of the so-called χ^2-function, which compares various experimental data, for example: single-particle energy levels, radii, level densities, position of the Fermi level and other quantities we may possibly think of, with the respective theoretically calculated quantities. The function in question is defined as

$$\chi^2(p) = \frac{1}{k_\varepsilon - n_p}\cdot\sum_{k=1}^{k_\varepsilon} w_k[\varepsilon_k - e_k(p)]^2. \tag{1.85}$$

In the most elementary realisations, ε_k are the measured energies of single-particle levels and $e_k(p)$ are their theoretical analogues. The weight factors w_k are, by definition, positive numbers chosen arbitrarily by a physicist; they reflect relative importance of particular data points. The symbols k_ε and n_p denote the number of experimental data

and number of parameters, respectively. This version of the minimised function does not include the uncertainties of the experimental data. However, the role of the latter can be verified/tested in a simple and straightforward way by varying the input data to obtain the impact of such variations and repeat the minimisation of the χ^2 (modelling the experimental and theoretical uncertainties).

Let us now examine expression (1.85) aiming at an approximate linearization of the original, generally non-linear problem. We can linearise the problem by observing, that in the vicinity of the real solution the theoretical energy levels calculated in the $(i+1)^{th}$-step can be expanded as follows

$$e_k(p^{(i+1)}) \approx e_k(p^i) + \sum_{j=1}^{n_p} \left(\frac{\partial e_k}{\partial p_j}\right)\Big|_{p=p_j^{(i)}} (p_j^{(i+1)} - p_j^{(i)}). \tag{1.86}$$

To shorten the notation we introduce a scaled Jacobian matrix $J_{kj}^{(i)}$ and an auxiliary vector $b_k^{(i)}$ defined as

$$J_{kj}^{(i)} \equiv \sqrt{w_k}\left(\frac{\partial e_k}{\partial p_j}\right)\Big|_{p=p_j^{(i)}} \quad \text{and} \quad b_k^{(i)} = \sqrt{w_k}[e_k(p^{(i)}) - \varepsilon_k]. \tag{1.87}$$

These new objects allow to rewrite the expression for χ^2 in $(i+1)^{th}$ iteration in an (apparently) more complicated and yet very useful way, as

$$\chi^2(p^{(i+1)}) = \frac{1}{k_\varepsilon - n_p} \cdot \sum_{k=1}^{k_\varepsilon} \left[\sum_{j=1}^{n_p} J_{kj}^{(i)} \cdot (p_j^{(i+1)} - p_j^{(i)}) + b_k^{(i)}\right]^2. \tag{1.88}$$

The necessary condition for the minimum of this function is obtained when partial derivatives with respect to the parameters vanish, namely

$$\frac{\partial \chi^2}{\partial p_i} = 0, \tag{1.89}$$

This necessary condition for the minimum assumes to a good approximation that the minimum has already been reached at the $(i+1)^{st}$ iteration, in which case it can be written in the form of a matrix relation

$$\left[\mathbf{J}^{\mathrm{T}}\mathbf{J}\right]^{(i)} (\mathbf{p}^{(i+1)} - \mathbf{p}^{(i)}) = -\left[\mathbf{J}^{\mathrm{T}}\right]^{(i)} \mathbf{b}^{(i)} \quad \text{where} \quad [J_{kl}^{(i)}] \equiv \mathbf{J}^{(i)}. \tag{1.90}$$

This is our formulation of the forward problem, for the inverse one we instantly obtain

$$\mathbf{p}^{(i+1)} - \mathbf{p}^{(i)} = -\left\{\left[\mathbf{J}^{\mathrm{T}}\mathbf{J}\right]^{(i)}\right\}^{-1} \left[\mathbf{J}^{\mathrm{T}}\right]^{(i)} \mathbf{b}^{(i)}. \tag{1.91}$$

where $\mathbf{J}^{(i)}$ is a $k_\varepsilon \times n_p$ real matrix, $\mathbf{p}^{(i)} \in \mathbb{R}^{n_p}$ and $\mathbf{b}^{(i)} \in \mathbb{R}^{k_\varepsilon}$.

The solution of this equation for $\mathbf{p}^{(i+1)}$ requires inversion of the matrix $[\mathbf{J}^{\mathrm{T}}\mathbf{J}]^{(i)}$. One can easily predict what happens, when this matrix is singular and therefore cannot be inverted – the problem becomes ill-posed and the obtained solution is useless. Of

course if this is known from the beginning, the problem can be fixed by introducing some additional constraints into the χ^2 function. Unfortunately a typical situation is quite different. Numerical codes used in this type of calculations often do not solve equation (1.91), but are rather based on e.g. conjugated-gradient or steepest descent method. As a result there will be some parameterization of the model found, but because of its mathematical instability it will be physically meaningless in terms of further extrapolation. To avoid such situations, one should always carry out a stability check by performing, for example the singular value decomposition.

If we follow the prescription given in Section 1.4 and omit the iteration number we can write

$$\mathbf{J} = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathrm{T}} \tag{1.92}$$

where, for completeness, we recall that $\mathbf{D}$ is a diagonal matrix of the form

$$\mathbf{D} = \mathrm{diag}\{d_1, d_2, \ldots, d_{\min(k_\varepsilon, n_p)}\}. \tag{1.93}$$

This decomposition allows us to write $\mathbf{J}^{\mathrm{T}}\mathbf{J}$ matrix in the following form

$$\mathbf{J}^{\mathrm{T}}\mathbf{J} = \mathbf{V}\mathbf{D}^2\mathbf{V}^{\mathrm{T}}, \tag{1.94}$$

and therefore for the inverse we immediately obtain

$$(\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1} = \mathbf{V}\mathbf{D}^{-2}\mathbf{V}^{\mathrm{T}}, \tag{1.95}$$

where

$$\mathbf{D}^{-2} = \mathrm{diag}\left\{\frac{1}{d_1^2}, \frac{1}{d_2^2}, \ldots, \frac{1}{d_{\min(k_\varepsilon, n_p)}^2}; 0\right\}. \tag{1.96}$$

Performing the singular value decomposition allows us to easily analyse the possible scenarios. If at least one of the diagonal values of matrix $\mathbf{D}$ is zero, the inverse matrix $(\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1}$ does not exist. In numerical calculations this may happen for sufficiently small d_i due to the rounding errors. This implies that the problem is ill-posed and the parameters of the model are correlated, i.e. one ore more parameters are functions of the others.

Let us now insert the the decomposed matrix $\mathbf{J}$ into expression (1.91). We obtain a following relation

$$\mathbf{p}^{(i+1)} - \mathbf{p}^{(i)} = -\left[\mathbf{V}\mathbf{D}^{-1}\mathbf{U}^{\mathrm{T}}\right]^{(i)} \mathbf{b}^{(i)}, \tag{1.97}$$

which in index notation takes the form

$$\left[\mathbf{p}^{(i+1)} - \mathbf{p}^{(i)}\right]_j = -\sum_{l=1}^{n_p} V_{jl}^{(i)} \frac{1}{d_l^{(i)}} \sum_{k=1}^{k_\varepsilon} \left[U_{lk}^{\mathrm{T}}\right]^{(i)} b_k^{(i)}. \tag{1.98}$$

The form of this expression (d_l in denominator) automatically implies, that all vanishing singular values should be excluded from the summation in this expression. However, we can also remove all d_l, which are sufficiently close to zero at the price of introducion errors which need to be analysed on the case-by-case basis. We have a total freedom of choice when it comes to this cut-off level, but we have to remember that this advantage comes with a price. Every entry removed is a part of information lost and we move further away from the real solution.

1.9.1 Covariance Matrix

Simiraly to the linear least-squares problem, we can define the covariance matrix of the model parameters ($\mathbf{p}$ in this case). Unfortunately, we cannot use the equation (1.8), defining the inverse problem, because the relation between data and parameters is not linear. However, we can exploit the result given by eq. (1.61), where role of matrix $\mathbf{G}$ is assumed by the Jacobian matrix. This allows us to write[14]

$$\mathrm{Cov}(p_i, p_j) \sim (\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1}_{ij}, \tag{1.99}$$

where we assumed, that covariance matrix of the data vector is proportional to the identity matrix.

We can immediately conclude, that the possible singularity of the matrix $(\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1}$ directly affects the covariance matrix of the parameters. If the singular values are small, they cause the covariances to become large. This means, that not only parameters become correlated (off-diagonal entries of covariance matrix) but also the confidence intervals of the parameters grow larger, destroying the predictive power of the model.

Another important issue is that the covariance matrix in nonlinear problems strongly depends on the accuracy of the linearisation. The Taylor expansion is valid only in the vicinity of the real solution, but if this is not the case for a particular result, we cannot rely on the covariance matrix. A way of solving this problem is scaling the eq. (1.99) with the value of the $\chi^2(p)$ of this fit. We thus have

$$\mathrm{Cov}(p_i, p_j) = \chi^2(p)(\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1}_{ij}. \tag{1.100}$$

If we now insert the singular-value-decomposed matrix $(\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1}$ we obtain in the index notation

$$\mathrm{Cov}(p_i, p_j) = \chi^2(p) \sum_{l}^{n_p} V_{il} \frac{1}{d_l^2} V_{lj}^T. \tag{1.101}$$

Presenting the covariance matrix in this particular way helps us better visualise what happens when one or more $d_i \to 0$. This unwanted value enters every term of the matrix, thus spoiling our results. Fortunately, the situation can be remedied by simply removing those entries of $\mathbf{D}$, which are too small and can cause the problem to diverge by applying the Truncated Singular Value Decomposition approach, presented

[14] For more detailed derivation, cf. [4], Chapter 9.

in section 1.4.1. Again, we have to remember, that this causes the loss of information, therefore we have to be very careful in our choices.

The Confidence Intervals with given confidence level $(1-p)$ can be calculated by taking

$$\mathrm{CI} = \left(\bar{p}_i + \sqrt{\mathrm{Cov}(p_i,p_i)}\, t_{k_\varepsilon - n_p, \frac{p}{2}},\ \bar{p}_i + \sqrt{\mathrm{Cov}(p_i,p_i)}\, t_{k_\varepsilon - n_p, 1-\frac{p}{2}}\right), \tag{1.102}$$

where $\bar{p}_i$ denotes the parameter values obtained as the best fit. This expression means, that given the confidence level, there is a probability $P = 1-p$ that the interval contains the true value of parameters. The use of t-distribution instead of normal distribution is due to the fact, that the number of degrees of freedom is rather small.

1.10 Monte Carlo Methods

The existence of powerful computers allows to perform large, time-consuming calculations, which would be impossible to be carried out by humans. This possibility is exploited by various methods of error propagation, called Monte Carlo methods. They are useful, when there is no other way of estimating the uncertainties of the theoretical model, e.g. in nonlinear problems but also in high-dimensional cases, where an analytical result can be difficult to interpret. The Monte Carlo methods allow to test the behaviour of the model, its sensitivity to changes of the input.

The basic idea behind these methods is to create a large statistics by repeating the same calculations with given initial probability distribution of data. This is an opposite approach to all regression-type methods, where we have only limited data sets (samples). In Monte Carlo methods we simulate the 'experiment' along with its uncertainties. This is achieved by randomly generating a large number of data vectors, affected by uncertainty component, according to known probability distribution. Each vector is then used, e.g. to recover the parameters of the model, which results in a large number of parameter sets. We can formally write it as

$$\mathbf{m}^{(i)} = \mathbf{G}^{-1}(\mathbf{d}_0 + \nu_i), \tag{1.103}$$

where $\mathbf{d}_0$ is the mean value of the data probability distibution (it can be real experimental data or just a simulation), ν_i denotes the i-th randomly generated uncertainty component and $\mathbf{m}^{(i)}$ is the i-th set of parameters. This way we obtain a probability distribution of the model to be established, giving us the sought information about model uncertainties.

The presented scenario is just one way of applying the Monte Carlo approach. There are many algorithms, which are based on random number generators. They are used in various scientific domains, e.g. particle physics, biology, mathematics etc.

Chapter 2

Spherical Nuclei

A large part of results presented in this work was obtained within the nuclear mean field theory in the special case of spherical symmetry. This particular area of nuclear chart, due to its properties and simplicity, was already widely explored in the past. Our motivation to re-investigate the problem was based on recent suggestion of existence of exotic symmetries, e.g. tetrahedral symmetry in nuclei (for more details, cf. e.g. [12], [13], [14], [15]). This property can manifest itself in some nuclear regions, but, in order to search for experimental proofs, one needs theoretical predictions related, among others, to nuclear transitions probabilities.

Various attempts to obtain such predictions showed how crucial it is to use as good parameterisation of the Hamiltonian, as possible. For the calculations based on deformed Woods-Saxon approach, the so-called universal parameterisation was applied to search for energy minima. In view of the fact, that new, revisited single-particle energy levels were available for spherical nuclei, it was only proper to recalculate this parameterisation. We have chosen seven doubly-magic nuclei for this task: ^{16}O, ^{40}Ca, ^{48}Ca, ^{56}Ni, ^{90}Zr, ^{132}Sn, ^{208}Pb. A new, fast numerical code restricted only to spherically-symmetric nuclei was constructed. Results of its performance will be presented in the next section, along with discussion of its applicability in view of our recent reflections about the predictive power.

2.1 Experimental Information

Single-particle energy levels in case of spherical nuclei are relatively well known. Their positions have long been studied but many physists forget, that still they are very complex quantities. Therefore, since we have chosen energy levels as our observables, we are forced to ensure, that we use them properly.

As we have already mentioned, neutron and proton levels for a given nucleus (Z_0,N_0) are not observed directly, but are established from neighbouring $N_0 \pm 1$ and $Z_0 \pm 1$ nuclei, where single-particle or single-hole states are populated. This is obtained by light-projectile direct nucleon-transfer reactions ('stripping' e.g. (d,p), (t,d) or 'pick-up' e.g. (p,d), (d,t)) of the (Z_0,N_0) nucleus.

2.1.1 Single-Particle Energy Levels of ^{208}Pb

The heaviest stable spherical nucleus is the doubly-magic ^{208}Pb. It is also one of the best studied among the spherical nuclei. Let us therefore analyze its nucleonic structure to understand all the difficulties in extracting the proper information.

The ^{208}Pb nucleus has a lowest excited state at 2.614 MeV, which is a collective-oscillation octupole state. Due to its low energy and strong collectivity it is the main reason for the fragmentation of the single-particle strengths. There are other excited collective states resulting from one-particle one-hole couplings, which may also contribute to the scattering of spectroscopic factors, e.g. 5^- level at 3.198 MeV or 2^+ level at 4.085 MeV.

To recover the single-particle level structure of ^{208}Pb, four neighbouring odd-A nuclei were studied, i.e. $^{209}_{82}\mathrm{Pb}_{127}$, $^{207}_{82}\mathrm{Pb}_{125}$, $^{209}_{83}\mathrm{Bi}_{126}$ and $^{207}_{81}\mathrm{Tl}_{126}$. First reactions carried out to produce these nuclei gave only partial information about nucleonic levels. The range of considered excitation energies was rather small and most of the states were considered pure. Later experiments with different transfer reactions resulted in a more complete picture and the fragmentation of levels became more prominent.

Table 2.1: *Single-particle levels around $Z = 82$ and $N = 126$. Symbol e_I^{exc} represents the energy of the lowest state of each spin value measured in the four presented nuclei. The states written bold are considered as pure single-nucleon states, i.e. their spectroscopic factor is close to unity. The energies denoted with ε_κ are computed according to (1.4), taking into account the N fragments measured in the quoted references. The difference between e_I^{exc} and the averaged value ε_κ is given by the last column. All the energies are given in MeV.*

Proton levels around Z=82					Neutron levels around N=126				
States in $^{209}_{83}\mathrm{Bi}_{126}$					States in $^{209}_{82}\mathrm{Pb}_{127}$				
Level	e_1^{exc}	N[ref]	ε_κ	shift	Level	e_1^{exc}	N[ref]	ε_κ	shift
$\pi h_{9/2}$	**0.0000**	1[27]	0	0	$\nu g_{9/2}$	**0.0000**	1	0	0
$\pi f_{7/2}$	**0.8963**	2[27]	1.31	0.41	$\nu i_{11/2}$	**0.7788**	1	0.779	0
$\pi i_{13/2}$	1.6086	5[27]	1.97	0.36	$\nu j_{15/2}$	1.423(1)	4[30]	1.77	0.35
$\pi f_{5/2}$	**2.8262**	5[27]	3.44	0.61	$\nu d_{5/2}$	**1.5671**	-	-	-
$\pi p_{3/2}$	**3.1195**	-	-	-	$\nu s_{1/2}$	**2.0322**	-	-	-
$\pi p_{1/2}$	3.633(4)	-	-	-	$\nu g_{7/2}$	**2.491(1)**	-	-	-
					$\nu d_{3/2}$	**2.538(2)**	-	-	-
States in $^{207}_{81}\mathrm{Tl}_{126}$					States in $^{207}_{82}\mathrm{Pb}_{125}$				
Level	e_1^{exc}	N[ref]	ε_κ	shift	Level	e_1^{exc}	N[ref]	ε_κ	shift
$\pi s_{1/2}$	**0.0000**	3[31]	0.10	0.10	$\nu p_{1/2}$	**0.0000**	1	0	0
$\pi d_{3/2}$	**0.3510**	1[31]	0.351	0	$\nu f_{5/2}$	**0.5697**	1	0.570	0
$\pi h_{11/2}$	**1.3481**	2[31]	1.44	0.10	$\nu p_{3/2}$	**0.8978**	1	0.898	0
$\pi d_{5/2}$	**1.6827**	6[31]	2.08	0.40	$\nu i_{13/2}$	**1.6334**	9[28]	2.4	0.8
$\pi g_{7/2}$	3.474(6)	5[31]	4.18	0.71	$\nu f_{7/2}$	2.3399	4[29]	3.0	0.7
					$\nu h_{9/2}$	3.414(2)	-	-	-

2.2 Spherical Woods-Saxon Mean-Field

Single-particle energy levels, in view of theoretical considerations, are eigenvalues of the single-particle mean-field Hamiltonian. Proper construction of this Hamiltonian is crucial in order to obtain instructive theoretical information.

The mean-field, as only an approximation to reality, is incomplete and inexact by definition. The true nuclear Hamiltonian, in theory, includes all the interactions between nucleons – the ones we already know of, but also the yet undiscovered ones.

As we do not know the true form of the many-body potential, we can approximate the nucleon-nucleon interaction by an averaged potential, which roughly resembles and reproduces the real interactions. Over the years, many forms of this mean potential, which originated from different assumptions, have been created. Some of them are more successful than others in describing the nuclear structure.

2.2.1 Definition of the Hamiltonian

Starting from the phenomenological point of view, we can write the following expression for the single-particle Hamiltonian

$$\hat{H}(\vec{r}) = \hat{T}(\vec{r}) + \hat{V}_C(\vec{r}) + \hat{V}_{SO}(\vec{r},\hat{\vec{p}},\hat{\vec{s}}) + \hat{V}_T(\vec{r},\hat{\vec{p}},\hat{\vec{s}}) + \left[\hat{V}_{coul}(\vec{r})\right] + \dots \quad . \tag{2.1}$$

Above, $\hat{T}(\vec{r})$ is the kinetic energy term, defined as

$$\begin{aligned} &\hat{T}(\vec{r}) \equiv -\frac{\hbar^2}{2m}\nabla^2, \text{ where} \\ &\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left[\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\right]. \end{aligned} \tag{2.2}$$

$\hat{V}_C(\vec{r})$ is the central term, responsible for the leading radial dependence of the potential. For our calculations, we have chosen the Woods-Saxon potential to play the role of the central term, which in its simplest, spherical form is given by

$$V_C^{WS}(r) = \frac{V_0}{1+e^{(\frac{r-R}{a})}}, \quad R = r_0 A^{1/3}, \tag{2.3}$$

where V_0 is the depth of the potential well, a is the diffuseness parameter, responsible for the thickness of the nuclear surface and R is the radius parameter of the nucleus.

The third term in equation (2.1) is the spin-orbit interaction, introducing the shell effects into the Hamiltonian. Traditionally it is parameterised by

$$\hat{V}_{so}(\vec{r},\hat{\vec{p}},\hat{\vec{s}}) = \frac{1}{r}\frac{dV^{so}(r)}{dr}\hat{\vec{\ell}}\cdot\hat{\vec{s}}, \tag{2.4}$$

where $V^{so}(r)$ is the Woods-Saxon type potential, which means it has three adjustable parameters: the potential depth V_0^{so}, diffuseness a^{so}, and the radius r_0^{so}.

The next term is the tensor interaction, recently intensively re-discussed in nuclear physics within the nuclear shell model [25]. Its formal definition will be presented in a seperate section along with the presentation of its mean-field version.

The last term is the Coulomb interaction, applicable only to protons, as only they have nonzero electric charge [hence the brackets]. In this work we assume the charge of the nucleus to be uniformly distributed inside the sphere of the radius R_{coul}, called the charge radius. This will allow us to present the Coulomb interaction potential as the classical electrostatic potential inside and outside of a uniformly charged sphere:

$$\hat{V}_{coul}(r) = \left\{ \begin{array}{ll} \dfrac{e^2(Z-1)}{4\pi\epsilon_o R_{coul}} \left(\dfrac{3}{2} - \dfrac{1}{2} \dfrac{r^2}{R_{coul}^2} \right) & \text{for} \quad r \leqslant R_{coul} \\ \dfrac{e^2(Z-1)}{4\pi\epsilon_o r} & \text{for} \quad r > R_{coul} \end{array} \right\}, \tag{2.5}$$

where $(Z-1)e$ denotes the source charge, which is contained in the sphere. The charge radius is given by the standard expression, connecting it to the number of nucleons

$$R_{coul} \equiv r_{coul}\, A^{1/3}. \tag{2.6}$$

We allow the charge radius parameter r_{coul} to differ from the nuclear radius and treat is as an additional parameter. This gives us a total of seven parameters for protons, which include three from the central WS potential, another three parametrising the spin-orbit and the Coulomb radius as the seventh one; for neutrons we have six parameters.

The single-particle Hamiltonian can be expanded by adding further terms, if we find them, which will complement the interaction and bring it closer to the true Hamiltonian. As a starting point of our considerations, we take into account only the main, central term and the spin-orbit interaction with its form defined by eq. (2.4). As a next step, we propose a different definition of the spin-orbit term, which includes also the nucleonic densities and allows for a reduction of the number of the parameters. This approach will be discussed in the next section.

2.2.2 The Schrödinger Equation

The theoretical nucleonic single-particle energy levels are obtained by solving the Schrödinger equation with the Hamiltonian defined by equation (2.1). To solve this equation, we use the fact, that for any vector in Hilbert space we can write the following superposition

$$\Psi_n = \sum_k c_{nk} \phi_k \tag{2.7}$$

where $\{\phi_k\}$ is a given set of basis vectors, k denotes all the quantum numbers specifying the basis states and c_{nk} are some unknown coefficients. The Schrödinger equation takes now the form

$$\hat{H} \sum_k c_{nk} \phi_k = E_n \sum_k c_{nk} \phi_k. \tag{2.8}$$

Multiplication of both sides by $\phi^*_{k'}$ and integration over the whole space gives

$$\sum_k c_{nk} \int dV \Phi^*_{k'} \hat{H} \Phi_k = \sum_k c_{nk} E_n \delta_{k'k}. \tag{2.9}$$

where we exploited the orthonormality property for the basis vectors. The resulting integrals are matrix element of the Hamiltonian, for which we introduce a short-hand bra-ket notation

$$\langle k'|H|k\rangle \equiv \int dV \Phi^*_{k'} \hat{H} \Phi_k. \tag{2.10}$$

This way we obtain a system of linear equations to be solved for c_{nk}:

$$\sum_k c_{nk} \left(\langle k'|H|k\rangle - E_n \delta_{k'k}\right) = 0, \tag{2.11}$$

with the non-trivial solutions existing if and only if

$$\mathbf{det}\left(\langle k'|H|k\rangle - E_n \delta_{k'k}\right) = 0; \tag{2.12}$$

the latter expression equivalent to a diagonalisation problem of the matrix $H_{k'k} \equiv \langle k'|H|k\rangle$.

2.2.3 Spherical Basis

As we have limited ourselves only to the spherical case, the natural choice of the basis functions would be the harmonic oscillator basis in spherical coordinates $\{r, \vartheta, \varphi\}$, namely

$$\phi_{n\ell m_\ell}(r, \vartheta, \varphi) = f_{n\ell}(r)\; Y_{\ell m_\ell}(\vartheta, \varphi) = \frac{R_{n\ell}(r)}{r}\, Y_{\ell m_\ell}(\vartheta, \varphi), \tag{2.13}$$

where $Y_{\ell m_\ell}$ are the spherical harmonics and $R_{n\ell}$ are the radial functions defined as

$$R_{n\ell}(r) = N_{n\ell}\, \exp\left[-\tfrac{1}{2}\left(\tfrac{r}{a}\right)^2\right] \left(\tfrac{r}{a}\right)^{\ell+1}\; L_n^{(\ell+\frac{1}{2})}\left[\left(\tfrac{r}{a}\right)^2\right], \tag{2.14}$$

with $L_n^{(\ell+\frac{1}{2})}$ being the generalised Laguerre polynomials, Ref. [11], Eq. No. (22.3.9). The constant $N_{n\ell}$ is the normalisation factor given by

$$N_{n\ell} = \sqrt{\frac{2^{\,n+\ell+2}\, n!}{a(2n+2\ell+1)!!\sqrt{\pi}}}, \tag{2.15}$$

and $a \equiv \sqrt{\hbar/m\omega}$.

However, one needs to remember that we are dealing with fermions and therefore our basis needs to include also the spin basis vectors. We need to construct vectors, that are simultaneus eigenstates of $\hat{j}^{\,2}$, $\hat{\ell}^{\,2}$ and $\hat{s}^{\,2}$ and spherically symmetric harmonic oscillator Hamiltonian.[1] This can be acomplished by introducing the Clebsch-Gordan coefficients

[1] All these operators commute among themselves, which implies the solutions to be their simulaneous eigenstates.

$C\left(\ell m_\ell; s, m_s | j m_j\right)$. Addition of spin and the orbital angular momentum must obey the following rule for the eigenvectors

$$\Phi_{n;jm_j,\ell s}(\vec{r}) = \sum_{m_\ell\, m_s} C\left(\ell m_\ell; s, m_s | j m_j\right) \phi_{n;\ell m_\ell}(\vec{r}) \chi_{s m_s}, \tag{2.16}$$

with the condition

$$m_\ell + m_s = m_j. \tag{2.17}$$

Above, χ_{sm_s} denotes the eigenvectors of the spin operators $\vec{s}^2$ and s_z:

$$\chi_+ \equiv \chi_{\frac{1}{2}\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \chi_- \equiv \chi_{\frac{1}{2}-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.18}$$

In the case of spin 1/2 particles, expression (2.16) will simplify to two terms corresponding to two possible values of $m_s = \pm 1/2$. Each value of ℓ gives us only two possibilities for the total angular momentum: $j = \ell + \frac{1}{2}$ and $j = \ell - \frac{1}{2}$. This implies that four Clebsch-Gordan coefficients will be neccessary. From Ref. [10], Table 8.1, p. 271:

For $j = \ell + \frac{1}{2}$ and fixed m_j

$$\mathcal{C}^{++}_{\ell m_j} \equiv C(\ell,\, m_\ell = m_j - \tfrac{1}{2}; s, m_s = +\tfrac{1}{2} | j = \ell + \tfrac{1}{2},\, m_j) = +\sqrt{\tfrac{\ell + m_j + \frac{1}{2}}{2\ell+1}} \tag{2.19}$$

$$\mathcal{C}^{+-}_{\ell m_j} \equiv C(\ell,\, m'_\ell = m_j + \tfrac{1}{2}; s, m'_s = -\tfrac{1}{2} | j = \ell + \tfrac{1}{2},\, m_j) = +\sqrt{\tfrac{\ell - m_j + \frac{1}{2}}{2\ell+1}} \tag{2.20}$$

For $j = \ell - \frac{1}{2}$ and fixed m_j

$$\mathcal{C}^{-+}_{\ell m_j} \equiv C(\ell,\, m_\ell = m_j - \tfrac{1}{2}; s, m_s = +\tfrac{1}{2} | j = \ell - \tfrac{1}{2},\, m_j) = -\sqrt{\tfrac{\ell + m_j + \frac{1}{2}}{2\ell+1}} \tag{2.21}$$

$$\mathcal{C}^{--}_{\ell m_j} \equiv C(\ell,\, m'_\ell = m_j + \tfrac{1}{2}; s, m'_s = -\tfrac{1}{2} | j = \ell - \tfrac{1}{2},\, m_j) = +\sqrt{\tfrac{\ell - m_j + \frac{1}{2}}{2\ell+1}}. \tag{2.22}$$

If we now shorten the notation

$$\phi^+_{n;\, jm_j,\ell s} \equiv \phi_{n\ell; m_j - \frac{1}{2}\chi_+} \quad \text{and} \quad \phi^-_{n;\, jm_j,\ell s} \equiv \phi_{n\ell; m_j + \frac{1}{2}\chi_-}, \tag{2.23}$$

we can write the following expressions for the wave functions in the form

$$\left.\begin{array}{lrcl} j = \ell + \frac{1}{2}: & \Phi_{n;\, jm_j,\ell s} & = & \mathcal{C}^{++}_{\ell m_j}\phi^+_{n;\, jm_j,\ell s} + \mathcal{C}^{+-}_{\ell m_j}\phi^-_{n;\, jm_j,\ell s} \\ j = \ell - \frac{1}{2}: & \Phi_{n;\, jm_j,\ell s} & = & \mathcal{C}^{-+}_{\ell m_j}\phi^+_{n;\, jm_j,\ell s} + \mathcal{C}^{--}_{\ell m_j}\phi^-_{n;\, jm_j,\ell s} \end{array}\right\}. \tag{2.24}$$

2.2.4 Matrix Elements of the Hamiltonian

In this section the matrix elements of the Hamiltonian defined through Eq.(2.10) will be calculated. First we introduce explicitly the set of quantum numbers denoted before as k. In the following the matrix elements will be labelled with

$$k \equiv \{n;\, jm_j, \ell s\} \quad \text{and} \quad k' = \{n';\, j'm'_j, \ell' s'\}. \tag{2.25}$$

The form of the basis functions implies that the matrix elements will in general consist of four terms. This number reduces to diagonal form in terms of the $\{\chi_+, \chi_-\}$-basis, since the Hamiltonian depends only on $\vec{s}^2$. Therefore the non-zero matrix elements have the following structure

$$\begin{aligned} j = \ell + \tfrac{1}{2}: \quad \langle k'|\hat{H}|k\rangle &= \mathcal{C}^{++}_{\ell' m'_j}\mathcal{C}^{++}_{\ell m_j}\int dV\,(\phi^{+}_{n';\,jm'_j,\ell's'})^*\,\hat{H}\,\phi^{+}_{n;\,jm_j,\ell s} \\ &+ \mathcal{C}^{+-}_{\ell' m'_j}\mathcal{C}^{+-}_{\ell m_j}\int dV\,(\phi^{-}_{n';\,jm'_j,\ell's'})^*\,\hat{H}\,\phi^{-}_{n;\,jm_j,\ell s} \end{aligned} \tag{2.26}$$

and

$$\begin{aligned} j = \ell - \tfrac{1}{2}: \quad \langle k'|\hat{H}|k\rangle &= \mathcal{C}^{-+}_{\ell' m'_j}\mathcal{C}^{-+}_{\ell m_j}\int dV\,(\phi^{+}_{n';\,jm'_j,\ell's'})^*\,\hat{H}\,\phi^{+}_{n;\,jm_j,\ell s} \\ &+ \mathcal{C}^{--}_{\ell' m'_j}\mathcal{C}^{--}_{\ell m_j}\int dV\,(\phi^{-}_{n';\,jm'_j,\ell's'})^*\,\hat{H}\,\phi^{-}_{n;\,jm_j,\ell s}. \end{aligned} \tag{2.27}$$

The basis is composed of the eigenstates of all the operators with which the Hamiltonian commutes i.e. $\vec{j}^2$, $\vec{\ell}^2$ and $\vec{s}^2$. This means that all the above matrix elements vanish unless $\ell' = \ell$, $j' = j$ and $m'_j = m_j$, $s = s'$. Therefore the matrix-element expressions reduce to

$$\begin{aligned} \int dV\,(\phi^{+}_{n';\,jm'_j,\ell's'})^*\,\hat{H}\,\phi^{+}_{n;\,jm_j,\ell s} &= \int dV\,(\phi^{-}_{n';\,jm'_j,\ell's'})^*\,\hat{H}\,\phi^{-}_{n;\,jm_j,\ell s} \\ &= \delta_{\ell\ell'}\,\delta_{ss'}\,\delta_{jj'}\,\delta_{m_j m'_j}\,\langle n'\ell'|\hat{H}|n\ell\rangle, \end{aligned} \tag{2.28}$$

where

$$\langle n'\ell'|\hat{H}|n\ell\rangle \equiv \int_0^\infty r^{\,2}\,dr f^*_{n'\ell'}(r)\hat{H} f_{n\ell}(r). \tag{2.29}$$

In addition the Clebsch-Gordan coefficients satisfy:

$$(\mathcal{C}^{\pm+}_{\ell m_j})^2 + (\mathcal{C}^{\pm-}_{\ell m_j})^2 = 1. \tag{2.30}$$

In the next step we will calculate the matrix elements of the Hamiltonian but treating the kinetic and potential terms separately.

Since the Laguerre polynomials depend on $\left(\frac{r}{a}\right)^2$ it will be useful to introduce a new variable

$$z \equiv \tfrac{r^{\,2}}{a^{\,2}} \quad \rightarrow \quad r = a\sqrt{z} \quad \rightarrow \quad dr = \tfrac{a}{2}\,\tfrac{dz}{\sqrt{z}} \tag{2.31}$$

Using this expression we can write the matrix elements of the operator $\hat{V}$ as follows

$$\boxed{\langle n'\ell|\hat{V}|n\ell\rangle = \frac{a}{2}N_{n'\ell}N_{n\ell}\int_0^\infty dz\,e^{-z}z^{\ell+\frac{1}{2}}\,L^{(\ell+\frac{1}{2})}_{n'}(z)\,\hat{V}(a\sqrt{z})\,L^{(\ell+\frac{1}{2})}_{n}(z)\,.} \tag{2.32}$$

The final expression for the kinetic energy term can be written down as

$$\begin{aligned} \langle n'\ell|\hat{T}|n\ell\rangle &= a\,\hbar\omega\,N_{n'\ell}\,N_{n\ell}\int_0^\infty dz\,e^{-z}z^{\ell-\frac{1}{2}}\left\{\left[\frac{1}{2}(\ell+1-z)L^{\ell+\frac{1}{2}}_{n'}(z) + z\frac{dL^{\ell+\frac{1}{2}}_{n'}}{dz}\right]\right. \\ &\left.\left[\frac{1}{2}(\ell+1-z)L^{\ell+\frac{1}{2}}_{n}(z) + z\frac{dL^{\ell+\frac{1}{2}}_{n}}{dz}\right] + \frac{\ell(\ell+1)}{4}L^{\ell+\frac{1}{2}}_{n'}(z)L^{\ell+\frac{1}{2}}_{n}(z)\right\}. \end{aligned} \tag{2.33}$$

2.3 Density Dependent Spin-Orbit Potential

The spin-orbit interaction defined in the previous section is an artificial construction, which does not include all the effects, which can occur at the surface of the nucleus. The Hartree-Fock formalism provides us with a solution, which can improve our model. Guided by the results of Skyrme-Hartree-Fock calculations we will substitute the Woods-Saxon type form-factor in the definition of the spin-orbit potential by nuclear density defined with the help of nucleonic wave functions. This will ensure, that all fluctuations of distibution of the nucleonic matter will have their contribution to the mean-field potential. As a result, we will obtain a semi-self-consistent procedure, since the wave functions will occur both in the potential and the solution.

Let us first briefly present the Skyrme-Hartree-Fock formalism and then we will derive the expression for the density dependent spin-orbit interaction used in the context of Woods-Saxon form of the central potential.

2.3.1 Skyrme Interaction

One of the interactions used in Hartree-Fock calulations is the Skyrme force. This interaction is local (zero-range), and is expressed by 'delta' functions. The general form of this force

$$\hat{V} = \sum_{i<j} \hat{v}_{ij}^{(2)} + \sum_{i<j<k} \hat{v}_{ijk}^{(3)}, \tag{2.34}$$

includes also the three-body interaction, which simulates the many-body effects. The two-body term has the following general form

$$\hat{v}_{ij}^{(2)} = \delta(\vec{r}_i - \vec{r}_j) v(\vec{k}', \vec{k}), \tag{2.35}$$

where $\vec{k}$ and $\vec{k}'$ are the operators of the relative motion defined as

$$\vec{k} \equiv \frac{1}{2i}(\vec{\nabla}_1 - \vec{\nabla}_2) \qquad \text{acting on the right} \tag{2.36a}$$

$$\vec{k}' \equiv -\frac{1}{2i}(\vec{\nabla}'_1 - \vec{\nabla}'_2) \qquad \text{acting on the left} \tag{2.36b}$$

The three-body term is expressed as

$$\hat{v}_{ijk}^{(3)} = t_3 \delta(\vec{r}_i - \vec{r}_j) \delta(\vec{r}_j - \vec{r}_k). \tag{2.37}$$

The function $v(\vec{k}', \vec{k})$ is a polynomial expansion in powers of $\vec{k}$ and $\vec{k}'$ limited to quadratic terms. It has the following form, as introduced by T.H.R. Skyrme ([16])

$$
\begin{aligned}
v(\vec{k}',\vec{k}) &= t_0(1+x_0\hat{P}_\sigma) \\
&+ \tfrac{1}{2}t_1(1+x_1\hat{P}_\sigma)(\vec{k}^2+\vec{k}'^2) \\
&+ t_2(1+x_2\hat{P}_\sigma)\,\vec{k}'\cdot\vec{k} \\
&+ \tfrac{1}{2}t_e\Big\{[3(\vec{\sigma}_1\cdot\vec{k})(\vec{\sigma}_2\cdot\vec{k})-(\vec{\sigma}_1\cdot\vec{\sigma}_2)\vec{k}^2]+[3(\vec{\sigma}_1\cdot\vec{k}')(\vec{\sigma}_2\cdot\vec{k}')-(\vec{\sigma}_1\cdot\vec{\sigma}_2)\vec{k}'^2]\Big\} \\
&+ t_o\Big\{[3(\vec{\sigma}_1\cdot\vec{k}')(\vec{\sigma}_2\cdot\vec{k})-(\vec{\sigma}_1\cdot\vec{\sigma}_2)\vec{k}'\cdot\vec{k}]\Big\} \\
&+ iW_o(\vec{\sigma}_1+\vec{\sigma}_2)\cdot\vec{k}'\times\vec{k},
\end{aligned}
\tag{2.38}
$$

with $\hat{P}_\sigma$ being the spin-exchange operator. The terms with coupling constants t_0, t_1 and t_2 enter the central part of the potential (along with the three-body $\hat{v}^{(3)}_{ijk}$ term). Constants t_e and t_o characterise the tensor potential, first corresponding to even and second - to odd states with respect to relative orbital motion. The last term with coefficient W_o represents the spin-orbit interaction.

The three-body term can be transformed into a two-body term by averaging over one of the particles. This way we obtain

$$
\hat{v}^{(3)}_{ij} = \tfrac{1}{6}t_3(1+x_3\hat{P}_\sigma)\rho_0\big[\tfrac{1}{2}(\vec{r}_i+\vec{r}_j)\big]\delta(\vec{r}_i-\vec{r}_j), \tag{2.39}
$$

where ρ_0 is the density taken at the center of mass of two nucleons. This term has been modified by introducing nonlinearity in density (most frequently chosen power of $\alpha=\frac{1}{3}$), which improves the compressibility of the nuclear matter.

2.3.2 Single-Particle Spin-Orbit Potential

Solution of the Hartree-Fock equations with the interaction defined in the previous section, gives the possibility of calculating the single-particle spin-orbit potential. It consists of two parts: the 'usual' and the tensor part. The latter originates from central and tensor part of the Skyrme interaction. Following the articles [17] and [18], we start with the definition of the spin-orbit potential for given particle type $q = p, n$ ([17] eq.(20) and (22b))

$$
V^q_{ls} = \overrightarrow{W}_q(\vec{r})\cdot(-i)(\overrightarrow{\nabla}\times\vec{\sigma}) \tag{2.40}
$$

with

$$
\overrightarrow{W}_q(\vec{r}) = \frac{1}{2}W_o(\overrightarrow{\nabla}\rho+\overrightarrow{\nabla}\rho_q) + \underbrace{\frac{1}{8}(t_1-t_2)\vec{J}_q(\vec{r})+\alpha_T\vec{J}_q(\vec{r})+\beta_T\vec{J}_{q'}(\vec{r})}_{\text{tensor spin-orbit}}. \tag{2.41}
$$

The first part, containig the derivatives of particle density ρ_q and total density $\rho = \rho'_q+\rho_q$, is equivalent to standard spin-orbit interaction defined by eq. (2.4). Its form is consistent with the relativistic mean-field approach. The second part of eq. (2.41)

is the tensor part of spin-orbit interaction. It consists of two expressions of different origins. The term with parameter $(t_1 - t_2)$ comes from the central interaction and only this part was considered by Vautherin and Brink in [17], whereas the last two terms are due to the tensor part ([18] eq.(16),(17)). The coefficients α_T and β_T, as given by [18] are expressed as[2]

$$\begin{aligned} \alpha_T &= \frac{5}{12} t_o \\ \beta_T &= \frac{5}{24}(t_o + t_e) \end{aligned} \tag{2.42}$$

The quantity $\vec{J}_q(\vec{r})$ represents the spin density defined as

$$\vec{J}_q(\vec{r}) = (-i) \sum_{i,\sigma\sigma'} \phi_i^*(\vec{r}, \sigma, q)[\overrightarrow{\nabla} \phi_i(\vec{r}, \sigma', q) \times \langle \sigma | \vec{\sigma} | \sigma' \rangle], \tag{2.43}$$

where the sum runs over all occupied states and ϕ_i represents the single-particle states given by

$$\phi_i(\vec{r}, \sigma, q) = \frac{R_\alpha(r)}{r} \sum_{m_l m_s} C(lsm_l m_s | jm)\, Y_{lm_l}(\vartheta, \varphi)\, \chi_{sm_s}(\sigma) \chi_q(\tau). \tag{2.44}$$

Appearance of the Clebsch-Gordan coefficients $C(lsm_l m_s|jm)$ stems from the fact, that this function must be simultaneously an eigenstate of $\hat{j}^2$, $\hat{l}^2$ and $\hat{s}^2$ and thus $i \leftrightarrow \{j, \ell, s\}_i$.

For symmetry reasons (spherical case), the spin density depends only on the radial component (cf. [17]), therefore it can be written as

$$\vec{J}(\vec{r}) = \frac{\vec{r}}{r} J(r) \tag{2.45}$$

with $J(r)$ defined as

$$J(r) = \frac{1}{4\pi r^3} \sum_\alpha (2j_\alpha + 1) \left[j_\alpha(j_\alpha + 1) - l_\alpha(l_\alpha + 1) - \frac{3}{4} \right] R_\alpha^2(r). \tag{2.46}$$

The same applies to the first term of $\overrightarrow{W}_q$, eq. (2.41) and the spin-orbit term can be therefore reduced to

$$V_{ls}^q = \frac{1}{r} W_q(r)\, \vec{l} \cdot \vec{\sigma} \tag{2.47}$$

with

[2] These values are a little bit confusing, because the coefficient in front of the central part of tensor interaction is the same as given by Vautherin and Brink, but there is $\vec{s}$ in the equation instead of $\vec{\sigma}$, so there could de a difference by a factor of $\frac{1}{2}$ in α_T and β_T.

$$W_q(r) = \frac{1}{2} W_o \frac{d}{dr} \left[\rho(r) + \rho_q(r)\right] + \frac{1}{8}(t_1 - t_2) J_q(r) + \alpha_T J_q(r) + \beta_T J_{q'}(r) \tag{2.48}$$

In our approach we have taken a generalised form of this potential, namely[3]

$$V_{ls}^q = \frac{1}{r}\Big[\big(\lambda_{so}^{qq} \frac{d\rho_q}{dr} + \lambda_{so}^{qq'} \frac{d\rho_{q'}}{dr}\big) + \lambda_T^{qq} J_q + \lambda_T^{qq'} J_{q'}\Big] \vec{l} \cdot \vec{\sigma} \tag{2.49}$$

2.3.3 Derivation of Particle Density

The particle density is defined as the sum of all the individual probability distributions over all occupied states, i.e.

$$\rho(r) \equiv \sum_{\alpha} \rho_\alpha(r), \tag{2.50}$$

where α denotes the full set of quantum numbers characterising single particle states. The solutions of the single-particle Hamiltonian are superpositions of basis function, that are simultaneous eigenstates of $\hat{j}^2$, $\hat{l}^2$ and $\hat{s}^2$. This allows us to change the summation index from ℓ to $j\ell \pm 1/2$, thus simplifying the notation. Let us consider only a part of expression (2.50) and start with a density for given N and j numbers. It can be defined as follows

$$\rho_j^N(r) = \sum_{m_j=-j}^{j} \left|\Psi_{N;\, jm_j,\ell}(r)\right|^2 \quad \text{with} \quad \Psi_{N;\, jm_j,\ell}(r) = \sum_n c_{j\ell n}^N \Phi_{n;\, jm_j,\ell s}(r) \tag{2.51}$$

where $\Phi_{n;\, jm_j,\ell s}(r)$ are given by Eq.(2.16). Both projections in the quantum number $j = \ell \pm 1/2$ yield the same result, as the square of Clebsh-Gordan coefficients are the same in both cases, cf. Eq (2.19)-(2.22). Therefore we will introduce a compact notation of these coefficients, namely

$$\mathcal{C}_{jm_j}^{+} \equiv \mathcal{C}_{\ell m_j}^{\pm +} \quad \text{and} \quad \mathcal{C}_{jm_j}^{-} \equiv \mathcal{C}_{\ell m_j}^{\pm -}. \tag{2.52}$$

This allows us to write the following expression for the density function, including the angular and spin components

[3] At this point we have a possibility of considering an alternative set of names of the coupling constants given the fact that

$$W_\nu(r) \sim \frac{d}{dr}(2\rho_\nu + \rho_\pi) \;\text{ and }\; W_\pi(r) \sim \frac{d}{dr}(2\rho_\pi + \rho_\nu)$$

and this leads, strictly speaking, to the following parameterisation

$$\lambda_{so}^{\nu\nu} = 2W_0, \;\; \lambda_{so}^{\pi\pi} = 2W_0, \;\; \lambda_{so}^{\nu\pi} = W_0, \;\text{ and }\; \lambda_{so}^{\pi\nu} = W_0.$$

It will therefore be important to verify on the basis of the fit to the experimental levels what are the *real* proportions between the coupling constants: is really $\lambda_{so}^{\nu\nu} \approx \lambda_{so}^{\pi\pi}$ and $\lambda_{so}^{\nu\pi} \approx \lambda_{so}^{\pi\nu}$?

$$\rho_j^N(r) = \sum_{m_j=-j}^{j} \sum_{n_1} c^N_{j\ell n_1} \left(\mathcal{C}^+_{jm_j} f^*_{n_1\ell} Y^*_{\ell;m_j-\frac{1}{2}} \langle\chi_+| + \mathcal{C}^-_{jm_j} f^*_{n_1\ell} Y^*_{\ell;m_j+\frac{1}{2}} \langle\chi_-| \right) \cdot \sum_{n_2} c^N_{j\ell n_2} \left(\mathcal{C}^+_{jm_j} f_{n_2\ell} Y_{\ell;m_j-\frac{1}{2}} |\chi_+\rangle + \mathcal{C}^-_{jm_j} f_{n_2\ell} Y_{\ell;m_j+\frac{1}{2}} |\chi_-\rangle \right). \quad (2.53)$$

Since the scalar products of antiparallel spin orientations $\langle\chi_-|\chi_+\rangle$ and $\langle\chi_+|\chi_-\rangle$ vanish, the above expression takes the following form

$$\rho_j^N(r) = \sum_{n_1,n_2} c^N_{j\ell n_1} c^N_{j\ell n_2} f_{n_1\ell} f_{n_2\ell} \left(\sum_{m_j=-j}^{j} \left(\mathcal{C}^+_{jm_j} \right)^2 \left| Y_{\ell;m_j-\frac{1}{2}} \right|^2 + \sum_{m_j=-j}^{j} \left(\mathcal{C}^-_{jm_j} \right)^2 \left| Y_{\ell;m_j+\frac{1}{2}} \right|^2 \right). \quad (2.54)$$

Additionally, both terms are exactly the same and the relative alignment of orbital momentum $\vec{\ell}$ and spin $\vec{s}$ is unimportant. This can be easily shown by simply considering all the possible combinations. Let us focus on one of them and set $j = \ell + 1/2$ and $m_j = m_\ell + 1/2$. If we change the summation index, we obtain

$$\sum_{m_j=-j}^{j} \left(\mathcal{C}^+_{jm_j} \right)^2 \left| Y_{\ell;m_j-\frac{1}{2}} \right|^2 = \sum_{m_\ell=-\ell-1}^{\ell} \frac{\ell + m_\ell + 1}{2\ell+1} |Y_{\ell m_\ell}|^2 = \sum_{m_\ell=-\ell}^{\ell} \frac{\ell + m_\ell + 1}{2\ell+1} |Y_{\ell m_\ell}|^2 . \quad (2.55)$$

From [10], p. 150, Eq. (1) and (2) we get

$$\sum_{m_\ell=-l}^{l} |Y_{\ell m_\ell}|^2 = \frac{2\ell+1}{4\pi} \quad \text{and} \quad \sum_{m_\ell=-l}^{l} m_\ell |Y_{\ell m_\ell}|^2 = 0, \quad (2.56)$$

which simplifies the expression to

$$\sum_{m_j=-j}^{j} \left(\mathcal{C}^+_{jm_j} \right)^2 \left| Y_{\ell;m_j-\frac{1}{2}} \right|^2 = \frac{1}{4\pi} \left(j + \frac{1}{2} \right). \quad (2.57)$$

Inserting this result into Eq.(2.51) we get the following expression for the partial density function

$$\rho_j^N(r) = \frac{2j+1}{4\pi} \sum_{n_1,n_2} c^N_{j\ell n_1} c^N_{j\ell n_2} f_{n_1\ell} f_{n_2\ell}. \quad (2.58)$$

Finally, the total density function can be written as

$$\rho(r) = \sum_{N,\ell,n_1,n_2} \frac{1}{2\pi} f_{n_1\ell} f_{n_2\ell} \left[(\ell+1) c^N_{\ell+\frac{1}{2}n_1} c^N_{\ell+\frac{1}{2}n_2} + \ell \cdot c^N_{\ell-\frac{1}{2}n_1} c^N_{\ell-\frac{1}{2}n_2} \right]. \quad (2.59)$$

where we replaced the summation over j with the one over ℓ. The obtained expression does not depend on the angular parts of the wave functions, which agrees with the symmetry of the problem. Furthermore, the degeneracy of nucleonic levels is included in this result.

2.3.4 Derivation of the Vector Spin Density

Begining with the same starting point as in previous derivation, we can rewrite the expression (2.46) into the following

$$J(r) = \frac{1}{4\pi r} \sum_{Nljn_1n_2} (2j+1) \left[j(j+1) - l(l+1) - \frac{3}{4} \right] c^N_{jln_1} c^N_{jln_2} f_{n_1l}(r) f_{n_2l}(r), \quad (2.60)$$

where we used the radial part of the eigenstate of the single-particle Hamiltonian, namely

$$R^N_{lj}(r) = r \sum_n C^N_{jln} f_{nl}(r). \quad (2.61)$$

We notice that for

$$j = l + 1 \quad \rightarrow \quad (2j+1) \left[j(j+1) - l(l+1) - \tfrac{3}{4} \right] = +2l(l+1), \quad (2.62)$$

$$j = l - 1 \quad \rightarrow \quad (2j+1) \left[j(j+1) - l(l+1) - \tfrac{3}{4} \right] = -2l(l+1), \quad (2.63)$$

which gives

$$J(r) = \frac{1}{2\pi r} \sum_{Nln_1n_2} l(l+1) \left[c^N_{l+\frac{1}{2}n_1} c^N_{l+\frac{1}{2}n_2} - c^N_{l-\frac{1}{2}n_1} c^N_{l-\frac{1}{2}n_2} \right] f_{n_1l}(r) f_{n_2l}(r). \quad (2.64)$$

We can now recall that the radial funtion is defined as

$$f_{nl}(r) = \frac{R_{nl}(r)}{r}, \quad (2.65)$$

and in the dimensionless coordinates $z = \left(\frac{r}{a}\right)^2$ we have

$$R_{n\ell}(z) = N_{nl} e^{-\frac{z}{2}} z^{\frac{1}{2}(\ell+1)} L_n^{(\ell+\frac{1}{2})}(z)\,. \quad (2.66)$$

In the above, the normalisation N_{nl} factor contains $\sqrt{\frac{2}{a}}$. This allows us to calculate

$$\frac{1}{r^2} f_{n_1l}(r) f_{n_2l}(r) = \frac{1}{a^4} N_{n_1l} N_{n_2l} e^{-z} z^{l-1} L_{n_1}^{(\ell+\frac{1}{2})}(z) L_{n_2}^{(\ell+\frac{1}{2})}(z) \quad (2.67)$$

and obtain

$$\frac{1}{r}J(r) = \frac{1}{2\pi a^4}\sum_{Nln_1n_2} N_{n_1l}N_{n_2l}e^{-z}z^{l-1}l(l+1)L_{n_1}^{(\ell+\frac{1}{2})}(z)L_{n_2}^{(\ell+\frac{1}{2})}(z) \cdot \left[c^N_{l+\frac{1}{2}n_1}c^N_{l+\frac{1}{2}n_2} - c^N_{l-\frac{1}{2}n_1}c^N_{l-\frac{1}{2}n_2}\right]. \quad (2.68)$$

Let us briefly analyze the obtained result. If the nucleus is spin-saturated, which means both 'spin up' and 'spin down' configurations with the same quantum number ℓ are occupied, the expression in brackets is very close to zero. However, if we are dealing with spin-unsaturated nucleus with only one level of this pair occupied, we obtain a nonzero correction as compared to classical spin-orbit. This explains the reduction of energy gaps in some spin-unsaturated nuclei, as compared to the same gap in other nuclei. An example of this behaviour for two nuclides will be presented in Sec. 4.3.

Comments and Interpretation. We have presented a new formulation of the spin-orbit term of the single-particle Hamiltonian. In this form, the spin-orbit interaction brings important structural differences compared with the original one as in Eq. (2.4). The main difference is the lack of the two 'geometrical' adjustable parameters, namely radius r_o^{so} and diffusivity a^{so}. From one point of view this is an improvement, as we aim for large-scale nuclear structure calculations, especially in exotic nuclear ranges, where physical properties of nuclei are poorly known. The classical spin-orbit parameters are not reliable when extrapolated far from the range, that they were fitted in. Therefore excluding them from the model may increase the stability of the properties of the phenomenological Hamiltonian.

This procedure, however, needs to be validated. We have to be certain, that the improvement is not illusive. The central potential determines the spatial distribution of nucleons inside the nucleus. Its parameters are very reliable in terms of extrapolations far from the range, in which they are fitted. This property is justified if we consider available experimental information. The electron scattering on nuclei gives us an information that the diffusivity of the nuclear surface is nearly constant and does not depend on the mass or size of the nucleus. Thus the central diffuseness parameter may be considered as independent, neither on Z nor on N. Additionally, the hadron scattering experiments allows us to assume that the effective mass distribution radius can be expressed as $R_o = r_o A^{1/3}$, with only one parameter r_o. Morover, the central potential depth parameter is nearly constant throughout te Periodic Table with a modification term proportional to the neutron excess parameter $(N-Z)/(N+Z)$. [4] Knowing all this, we may expect, that the geometrical parameters of the central potential are suitable for determining the nucleonic density distributions, which in turn will be reliable when extrapolated into exotic nucelar regions.

[4] The depth parameter is in this case taken in the form of $V = V_o[1 \pm \kappa(N-Z)/(N+Z)]$, where the plus sign is for protons and the minus sign for neutrons, which means it depends only on two parameters for all nuclei.

Chapter 3

Simulation of Predictive Power

This chapter contains a description of a simple mathematical model illustrating the very basic concepts of predictive power of a mathematical model, in particular difficulties arising from ill-posed models. We performed a througout modelling of performance of this model in terms of uncertainty of the input data. Applying the Monte Carlo techniques allowed us to test the correlation of the model parameters and their uncertainties. The impact of the level of data (im)precision (large or small 'experimental' errors) on the resulting parameters and their performance was also studied, along with the importance of proper data selection (usually referred to as 'sampling'). We tested also the predictive power of the model, its capability to extrapolate the results inside or outside the fitting region (intraneous and extraneous predictive power).

3.1 Motivations and Basic Assumptions

In Chapter 1 we have briefly presented the theory behind the inverse problems and introduced the concept of theoretical errors, which should be, in our opinion, taken into account when constructing nuclear theories. We believe that every model in subatomic physics, especially such, whose results are extrapolated into different regions of nuclear chart, needs to be tested in terms of its predictive power. We would like to show the significance of this statement by testing an exactly soluble mathematical model, which will be easy to analyze but also instructive enough to illustrate the problem of uncertainties in real calculations.

We choose to fit a function with four parameters to a given set of 'experimental' data. This function is defined as follows

$$y(x) = a + bx + c\sinh(x) + d\cosh(x), \tag{3.1}$$

where a, b, c and d are the parameters to be estimated. The data to be modelled are simply exponent of x scaled by additional factor, i.e.

$$y_{\alpha}^{\exp}(x) = \frac{\exp(x)}{1 + \alpha x^2}, \tag{3.2}$$

where α is an additional parameter allowing us to test the behaviour of the fitting

procedure when approaching the analytical case of $\alpha = 0$. Strictly speaking, eq. 3.2 represents a 'data generating function' which is used as follows. We define a sequence of discrete points x_i and generate the pseudo-experimental input to the fitting procedures by $y_\alpha^{\exp}(x_i)$.

For $\alpha = 0$ we have simply

$$\exp(x) = a + bx + c\sinh(x) + d\cosh(x). \tag{3.3}$$

It can be easily checked, that the only values fulfilling this relation are

$$a = b = 0 \quad \text{and} \quad c = d = 1. \tag{3.4}$$

This is a situation which resembles what is usually referred to as 'exact modelling'; in other words, this way of generating the 'experimental data points' leads to the data set that may be reproduced exactly. Every other case, when $\alpha \neq 0$, will result in only an approximation of the generated data. We are thus interested in examining the predictive power of this problem depending on the value of α.

We perform two different types of error estimation. The first test is a graphical presentation of the correlation of pairs of parameters, where we plot maps of the χ^2 values as a function of chosen two parameters with the remaining two kept fixed. The second one is a test of the impact of experimental errors on the fitted parameters. We introduce these errors into the minimisation process by generating a large number of data sets contaminated with random uncertainty admixtures, i.e. performing a Monte-Carlo-type calculations. This allows us to recover the information about parameter uncertainties.

3.1.1 Solution in Terms of the Inverse Problem

To fit the parameters of the function defined in eq. (3.1) we need a set of experimental data points, generated using the expression eq. (3.2). Let us therefore choose a set of m points, which means that we solve a system of equations

$$\begin{cases} a + bx_1 + c\sinh(x_1) + d\cosh(x_1) = y_\alpha^{\exp}(x_1) \\ a + bx_2 + c\sinh(x_2) + d\cosh(x_2) = y_\alpha^{\exp}(x_2) \\ a + bx_3 + c\sinh(x_3) + d\cosh(x_3) = y_\alpha^{\exp}(x_3) \\ \vdots \\ a + bx_m + c\sinh(x_m) + d\cosh(x_m) = y_\alpha^{\exp}(x_m). \end{cases} \tag{3.5}$$

This system can be represented in a matrix notation as

$$\mathbf{Gm} = \mathbf{d}, \tag{3.6}$$

where the data column vector $\mathbf{d}$ is defined as

$$\mathbf{d} = \begin{bmatrix} y_\alpha^{\exp}(x_1) \\ y_\alpha^{\exp}(x_2) \\ y_\alpha^{\exp}(x_3) \\ \vdots \\ y_\alpha^{\exp}(x_m) \end{bmatrix}, \tag{3.7}$$

the model vector in this case is simply

$$\mathbf{m} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \tag{3.8}$$

and $\mathbf{G}$ is the forward operator defined as

$$\mathbf{G} = \begin{bmatrix} 1 & x_1 & \sinh(x_1) & \cosh(x_1) \\ 1 & x_2 & \sinh(x_2) & \cosh(x_2) \\ 1 & x_3 & \sinh(x_3) & \cosh(x_3) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & \sinh(x_m) & \cosh(x_m) \end{bmatrix}. \tag{3.9}$$

This mathematical model is a linear least-squares problem and in general it should be possible to solve it analytically. However, we intend to introduce a random noise to the data vector, thus the minimisation will be performed using iterative methods.

3.1.2 Monte Carlo Protocol

Our analysis of the problem is based on the Monte Carlo method, presented in Sec. 1.10 . According to its principles, we generate a large number (N_{MC}) of 'experimental' data sets, as defined by Eq. 3.7, but modified by small uncertainty admixtures, namely

$$y_1 = y_\alpha^{\exp}(x_1) + \delta y_1, \quad y_2 = y_\alpha^{\exp}(x_2) + \delta y_2, \quad \ldots \quad y_m = y_\alpha^{\exp}(x_m) + \delta y_m. \tag{3.10}$$

The uncertainties $\{\delta y_i\}$ are generated randomly with probabilities given by a Gaussian distribution with mean value equal to 0 and a given standard deviation σ. For each data set, we perform a minimisation in order to recover the parameters that fit the given data. The best fit is established with the help of minimisation routine, calculating the minimum of a function

$$\chi^2(\mathbf{p}) = \sum_i (y_i - y(x_i))^2, \tag{3.11}$$

where $y(x_i)$ are the theoretical predictions, defined by Eq. 3.1.

As a result of the Monte Carlo calculations, we obtain sets of parameters

$$\{a, b, c, d\} \equiv \{P_1, P_2, P_3, P_4\}_i \quad \textbf{for} \;\; i = 1, 2, \ldots N_{MC}, \tag{3.12}$$

which are naturally affected by the uncertainty admixtures introduced to the data.

Applying the Monte Carlo protocol allows us to recover the influence of the 'experimental' uncertainties on the parameters. We can analyze the resulting 'theoretical' errors and probability distributions of the parameters. We can also test the predictive power of the model by attempting to 'predict' the results which in the actual modelling situation are known before hand.

3.2 Test of Correlation

The correlation of the parameters can be established by calculating the correlation matrix, but it can be also obtained as a result of the Monte Carlo calculations.

Each minimisation yields a set of parameters. By plotting the obtained values for pairs of parameters (e.g. P_1 versus P_2), we can judge, if the parameters are correlated. When this happens, we will observe curves formed by the points corresponding to pairs of parameters that were fitted in the same set. Uncorrelated parameters would result in a more or less uniform distribution on the 'xy' plane. Our mathematical model is parametrised by 4 parameters, therefore we obtain six types of correlation plots for all possible combinations of pairs $P_i - P_j$.

The following set of four figures presents the results of such a correlation test in the case of the 'exact model' (i.e. for $\alpha = 0$) with small 'experimental' uncertainties ($\sigma = 0.001$). Each figure consisting of four plots corresponds to one parameter pair. The plots differ by the sampling option, i.e. the number of points the model is fitted to. The top-left plot is obtained when we fit only to four points (they are calculated for $x_1 = 0,\ x_2 = 0.5,\ x_3 = 1\ , x_4 = 1.5$), every subsequent plot is a result of introducing four more points into the data set (calculated for next equidistant x_i's with $\Delta x = 0.5$). It is important to stress here, that the results are plotted in different scales, depending on the sampling type we choose.

The first figure shows the correlation of P_1 and P_2. What we can observe is a linear correlation between these two parameters, which strongly depends on the data we use. In the case of the smallest sample, the parameters vary in a rather broad range: $P_1 \in [-0.1, +0.1]$ and $P_2 \in [-0.1, +0.1]$. They are poorly constrained, as the points are calculated for small values of x_i, which means, that the exponential is almost negligible. Increasing the sample, we can observe, that P_1 and P_2 become much less affected by the data uncertainties and their correlation is less linear. This happens, because for larger values of x_i, the exponential part has the main contribution to the data points.

The next figure (3.2) presents an analogous set of plots for P_1 and P_3. Again, there is also a linear correlation between these two parameters. Similarly as before, the parameters become more constrained, when the size of the sample grows. This time, however, for the 16-point sample, the correlation is much less visible.

The plots presented in figure 3.3 correspond to the correlation of P_3 and P_4. These are the parameters responsible for the exponential part of the function. It is clearly visible, that they become strongly correlated for the largest-sized sample.

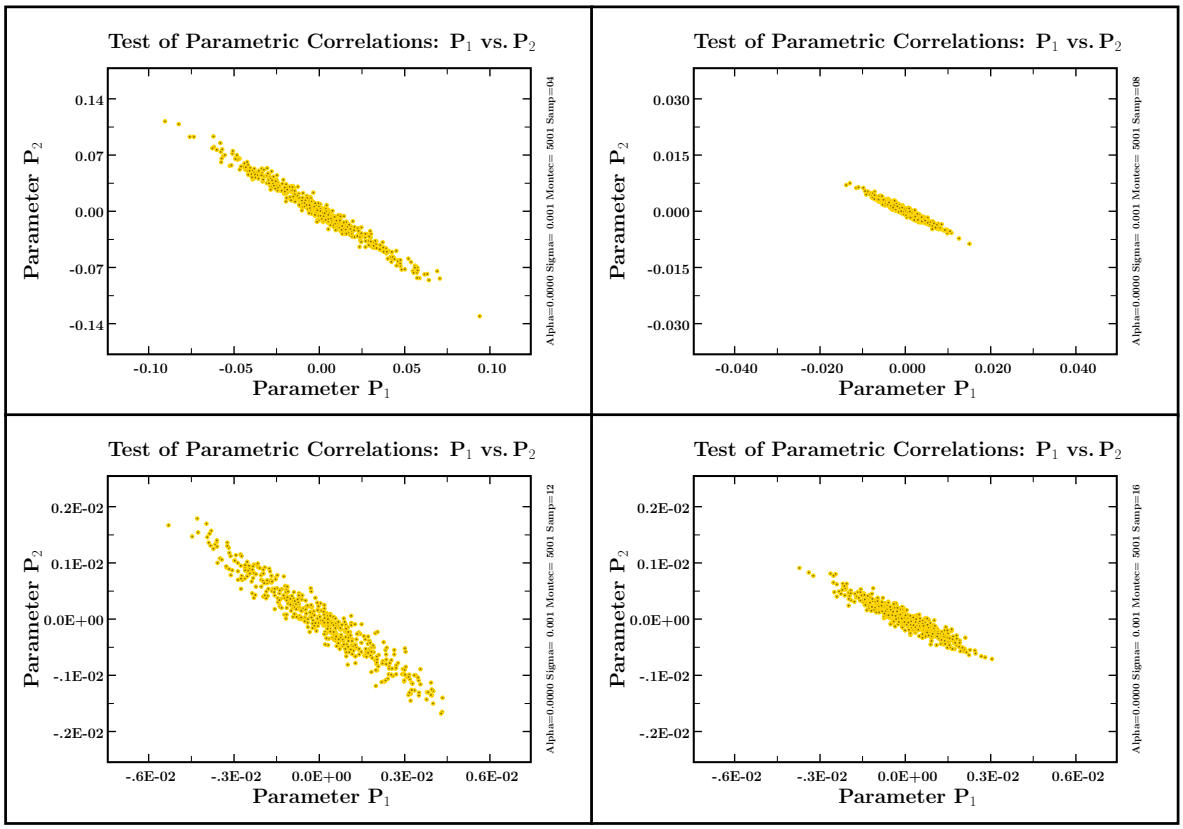

Figure 3.1: *Correlation of parameters P_1 and P_2 for $\alpha = 0$ ('exact model' version) and $\sigma = 0.001$ ('precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

The strong linear correlation for all possible pairs of parameters (similar results are obtained for other pairs) could be explained by the small level of contamination of the data. We may expect, that this will change, if we increase the width of the Gaussian distribution of the data points. However, if we plot similar figures for $\sigma = 0.005$ and $\sigma = 0.025$, we immediately notice, that the only thing that changes is the scale. The parameters are more weakly constrained by the data and thus vary strongly. but still the linear correlation remains.

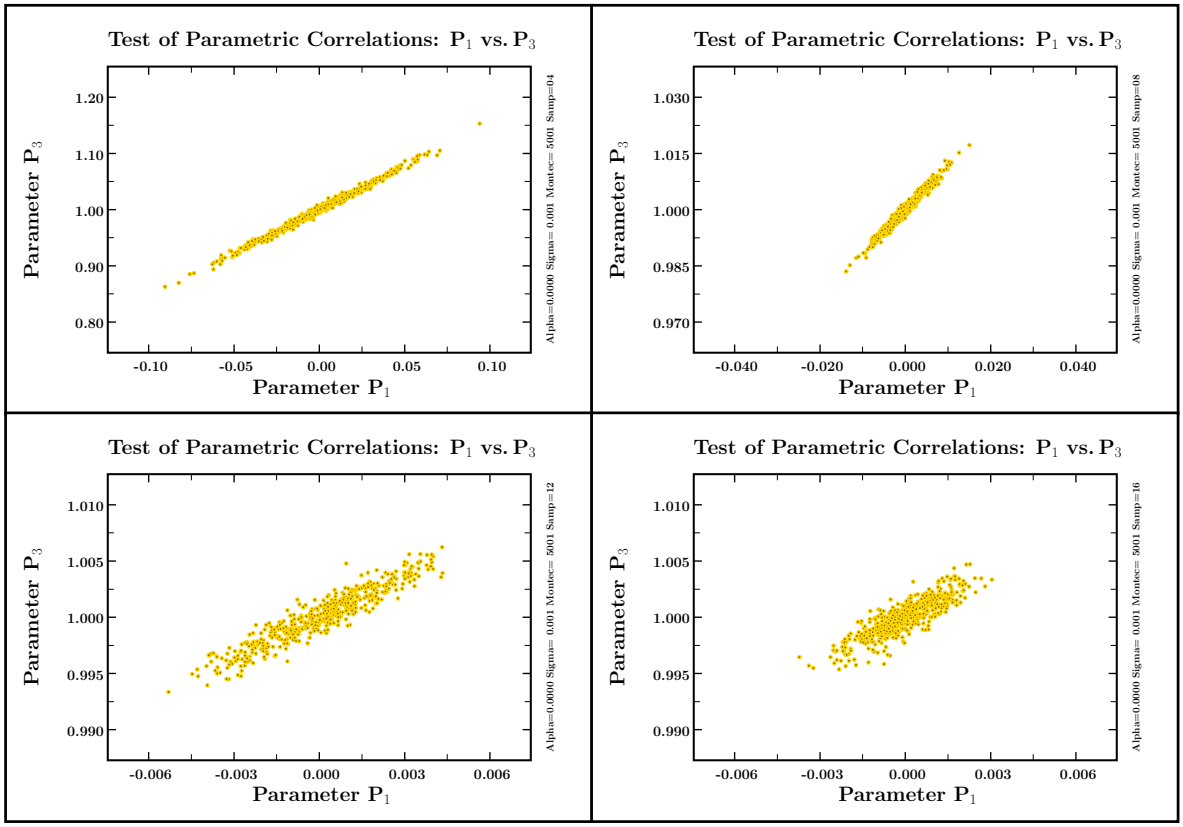

Figure 3.2: *Correlation of parameters P_1 and P_3.*

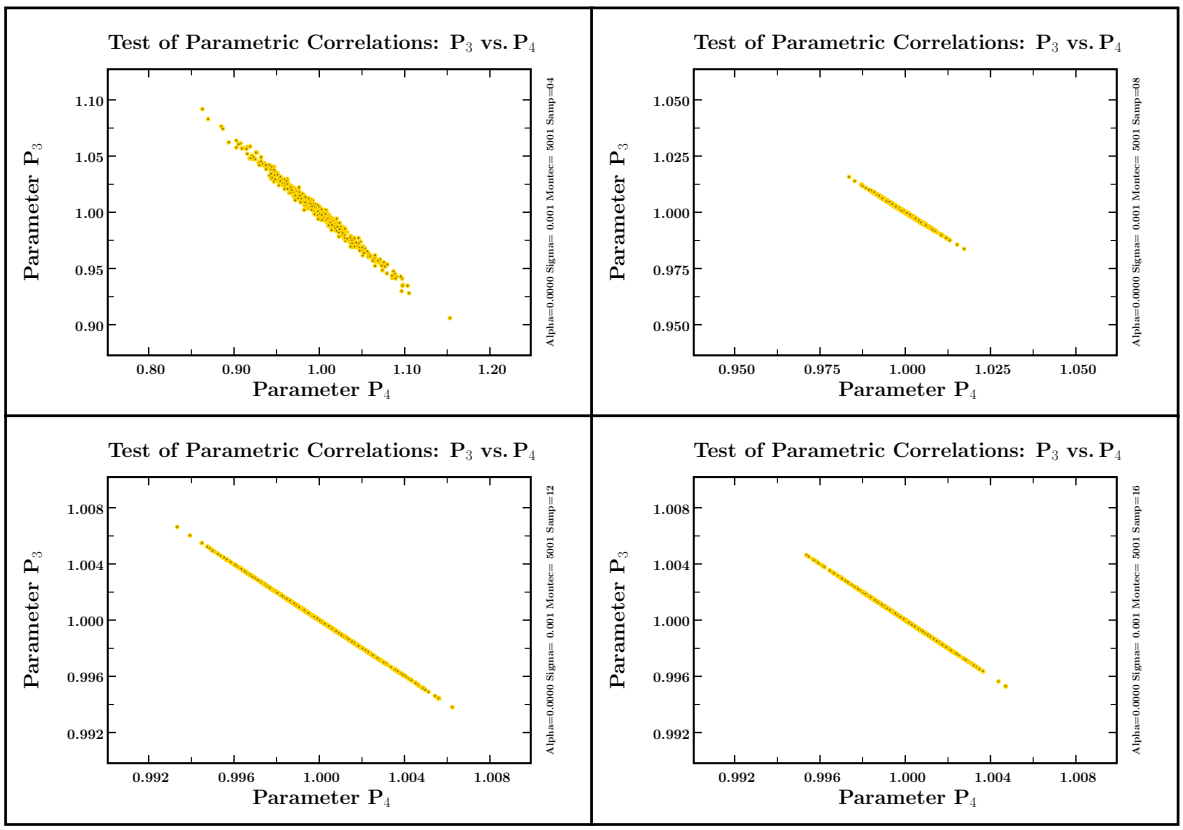

Figure 3.3: *Correlation of parameters P_3 and P_4.*

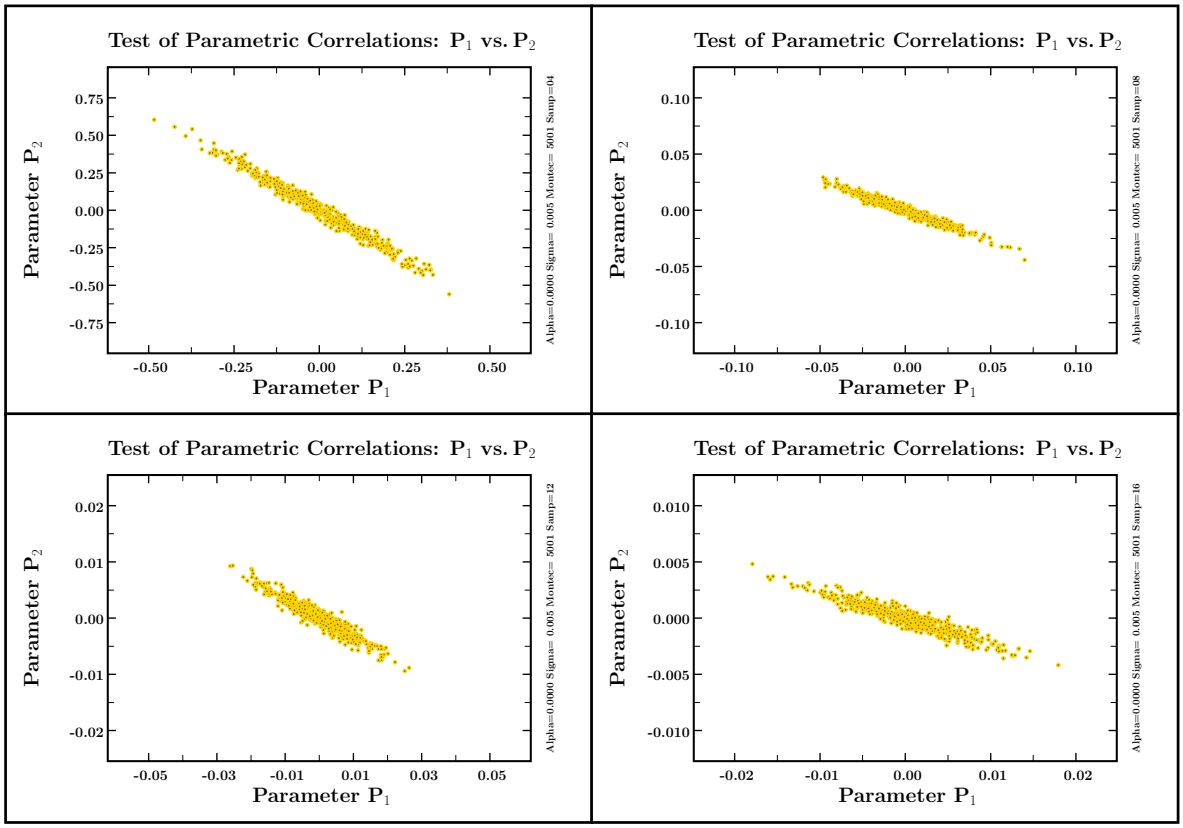

Figure 3.4: *Correlation of parameters P_1 and P_2 for $\alpha = 0$ ('exact model' version) and $\sigma = 0.005$ ('less precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

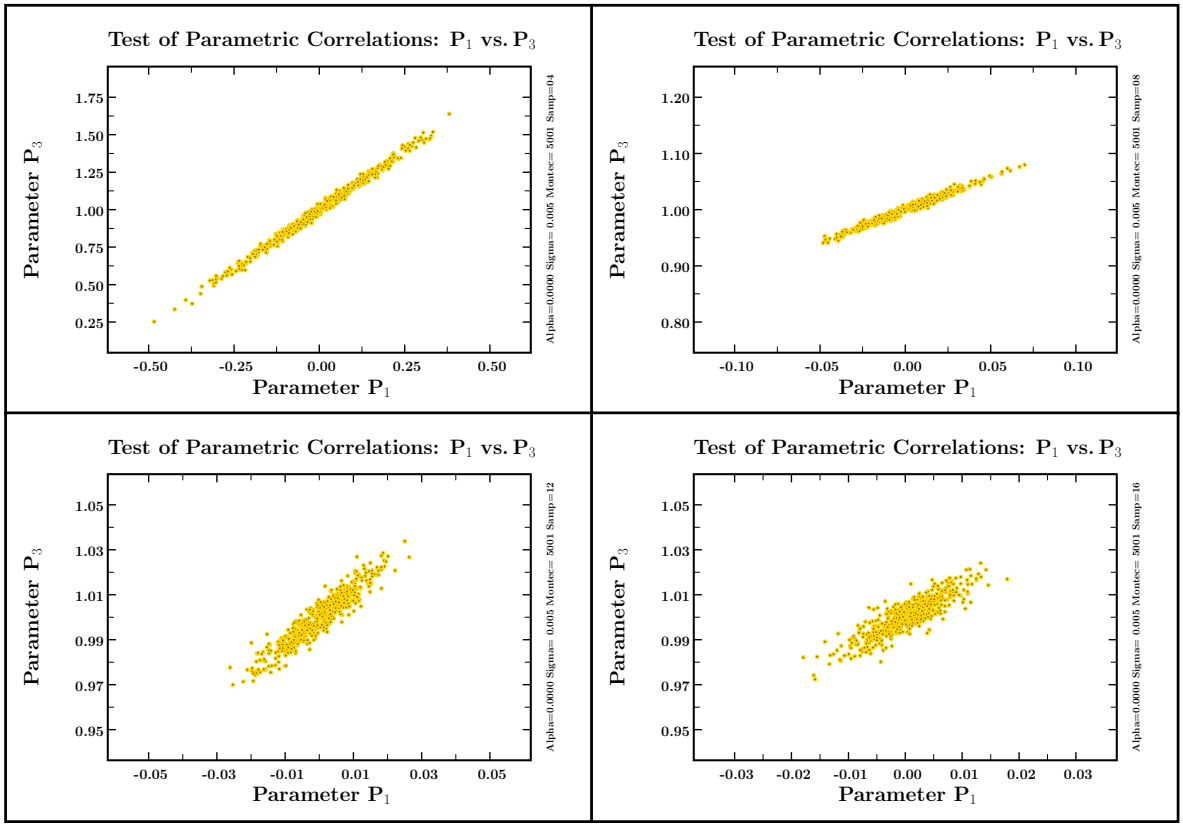

Figure 3.5: *Correlation of parameters P_1 and P_3.*

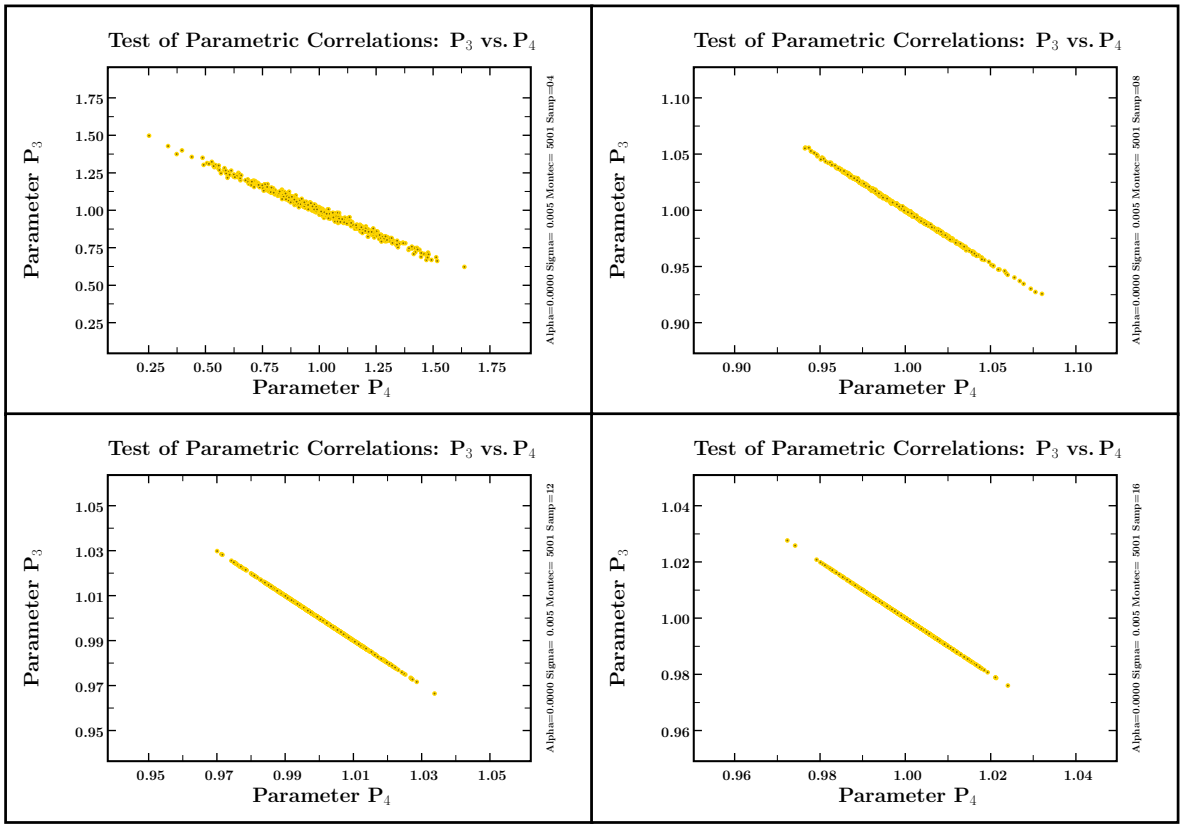

Figure 3.6: *Correlation of parameters P_3 and P_4.*

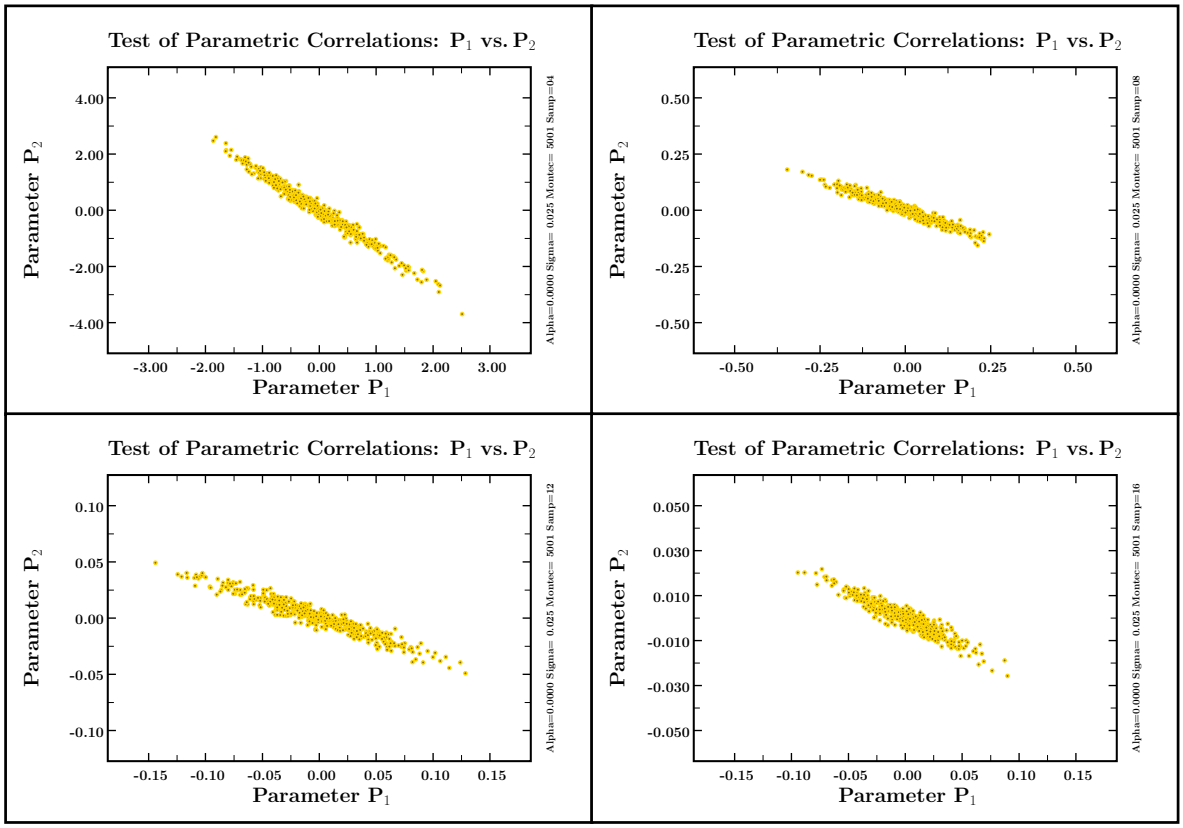

Figure 3.7: *Correlation of parameters P_1 and P_2 for $\alpha = 0$ ('exact model' version) and $\sigma = 0.025$ ('imprecise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

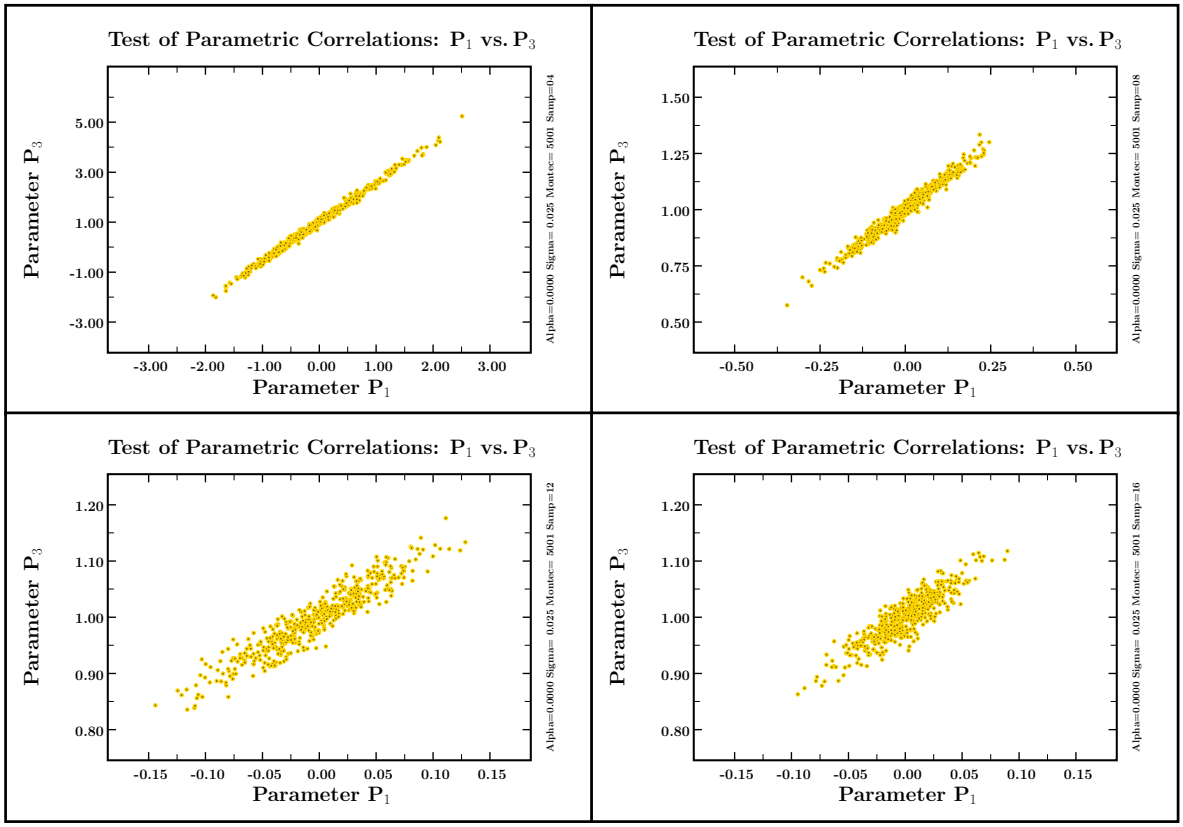

Figure 3.8: *Correlation of parameters P_1 and P_3.*

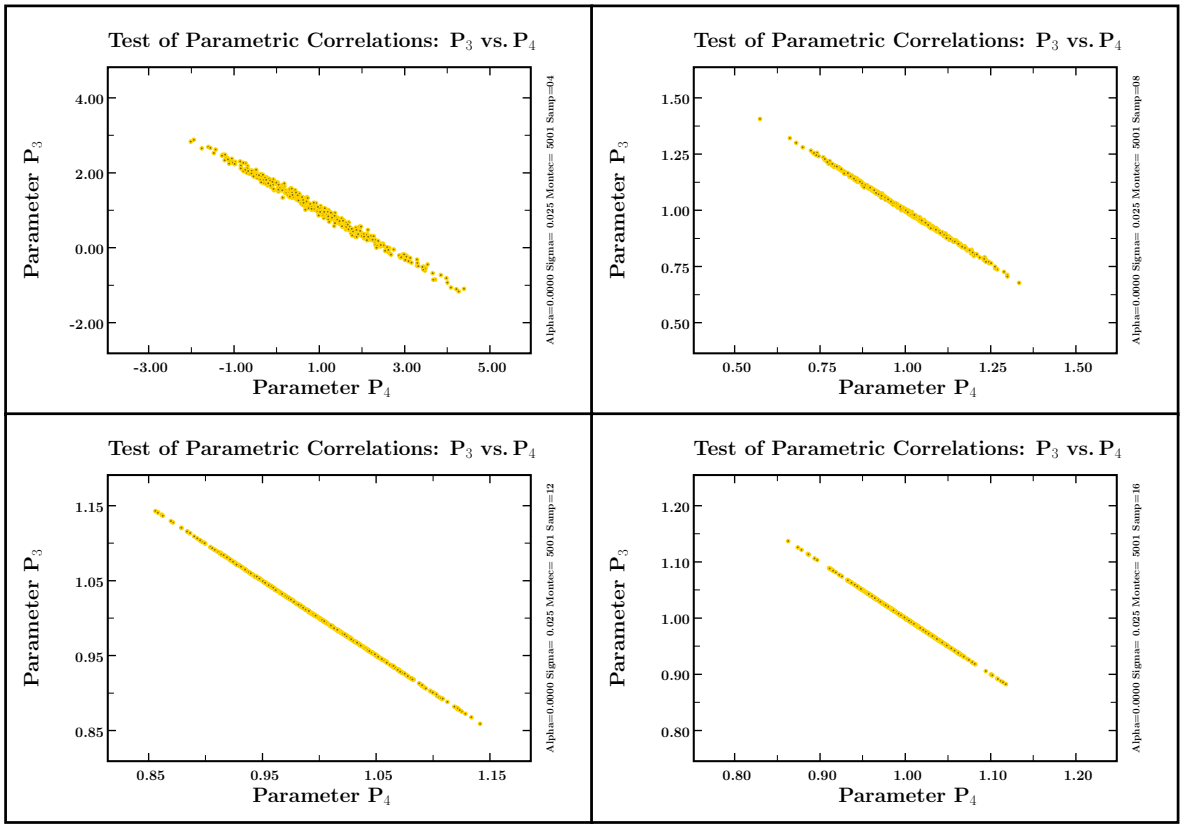

Figure 3.9: *Correlation of parameters P_3 and P_4.*

The following figures 3.10-3.18 and 3.19-3.27 present a similar set of plots, but for the 'inexact' models with $\alpha = 0.001$ and $\alpha = 0.005$, respectively. They are organized in exactly the same way as previously. Observe, that the general trend and the scale of the plots remain roughly unchanged. However, for the largest sample sizes (12 and 16), the increase of the experimental uncertainties has almost no impact on the results.

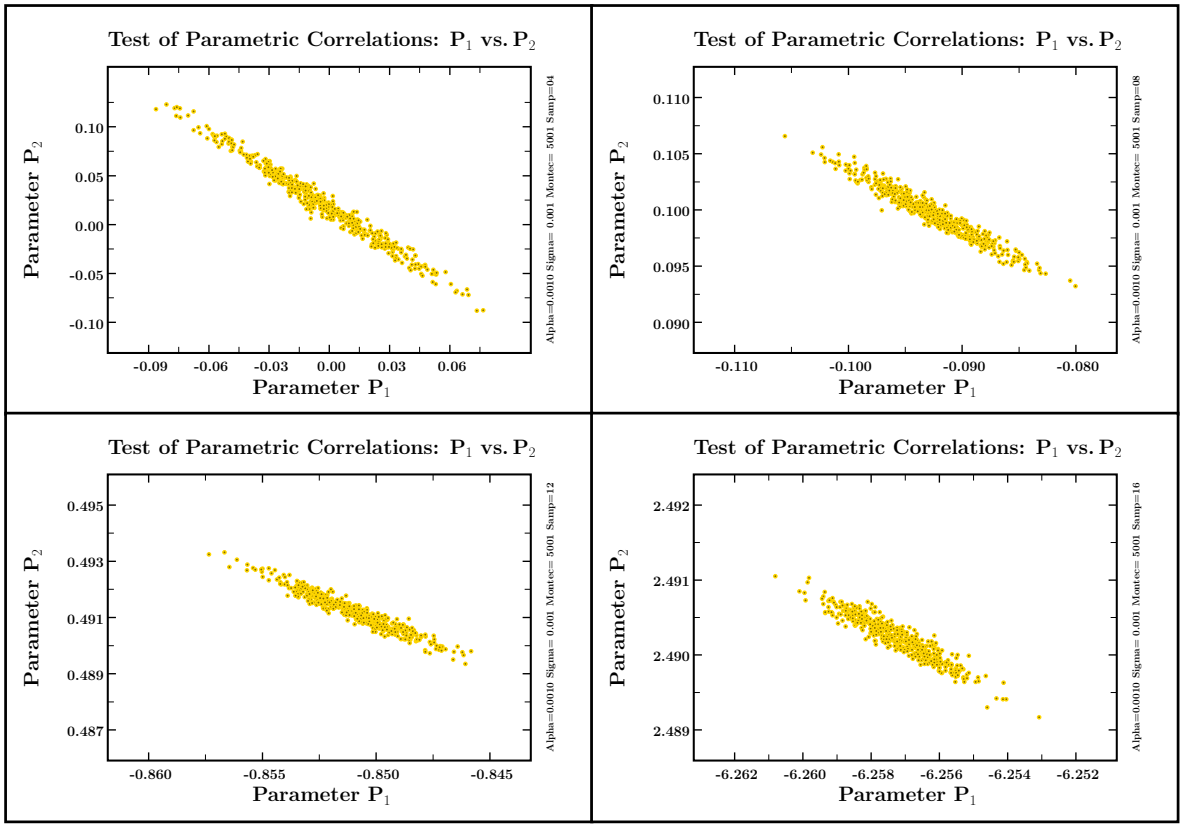

Figure 3.10: *Correlation of parameters P_1 and P_2 for $\alpha = 0.001$ and $\sigma = 0.001$ ('precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

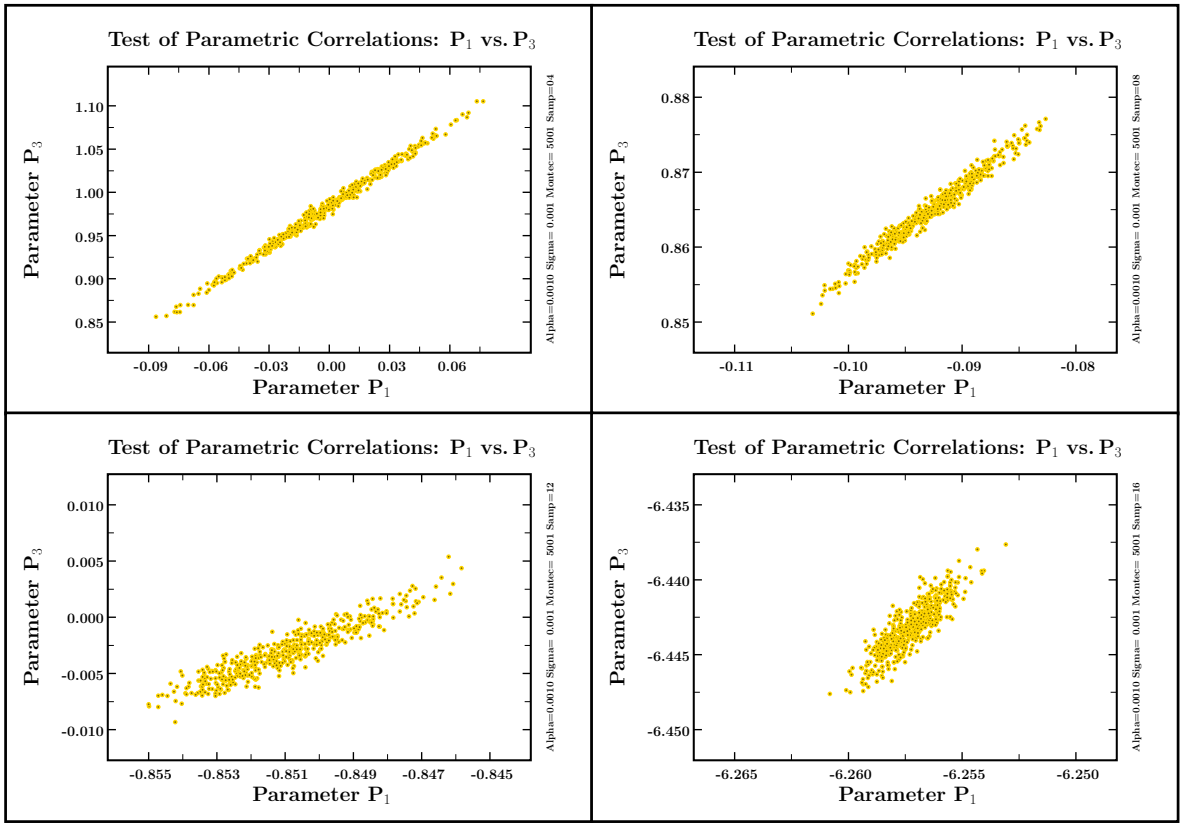

Figure 3.11: *Correlation of parameters P_1 and P_3.*

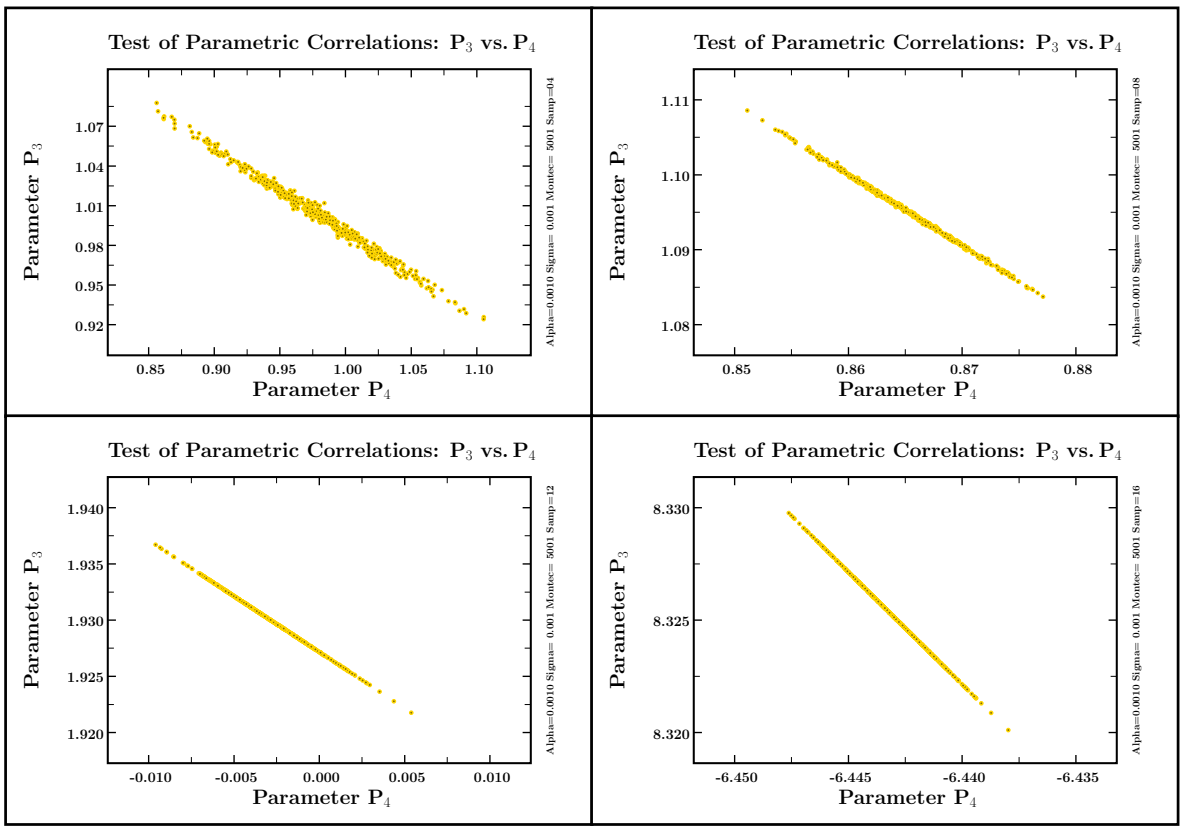

Figure 3.12: *Correlation of parameters P_3 and P_4.*

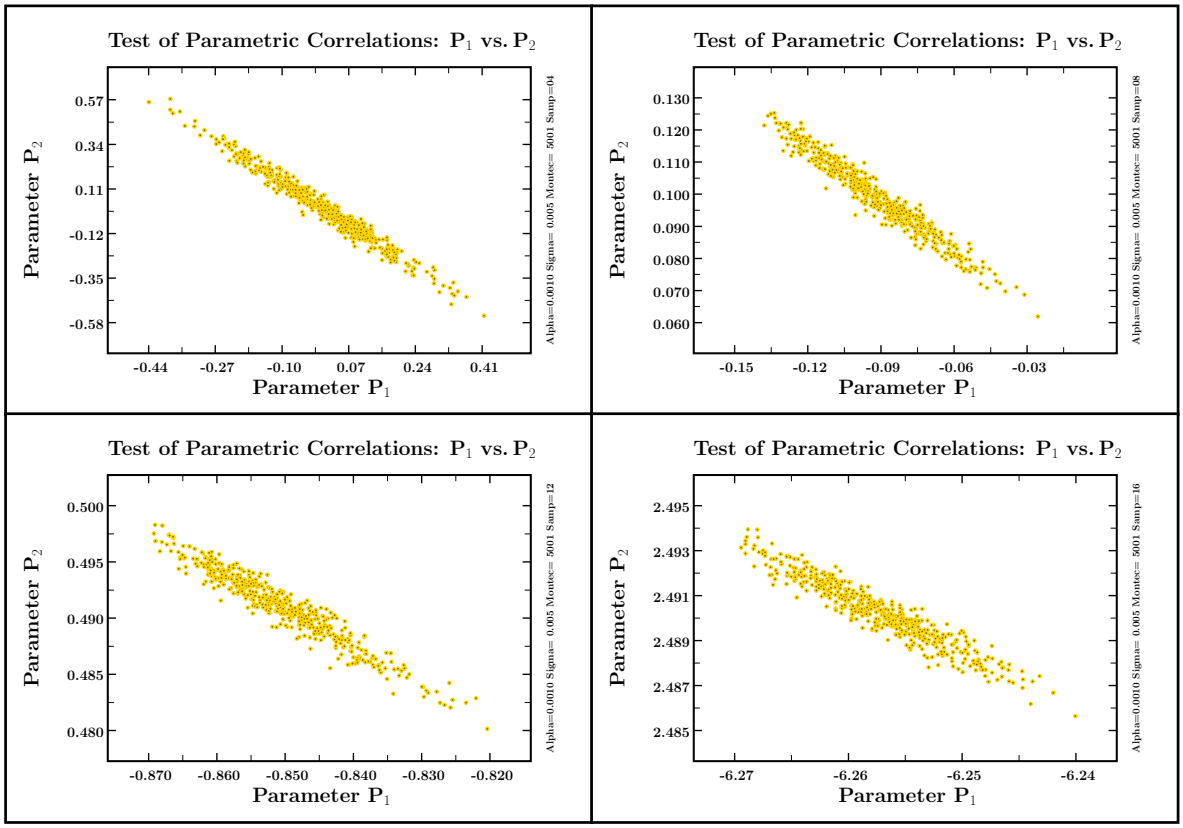

Figure 3.13: *Correlation of parameters P_1 and P_2 for $\alpha = 0.001$ and $\sigma = 0.005$ ('less precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

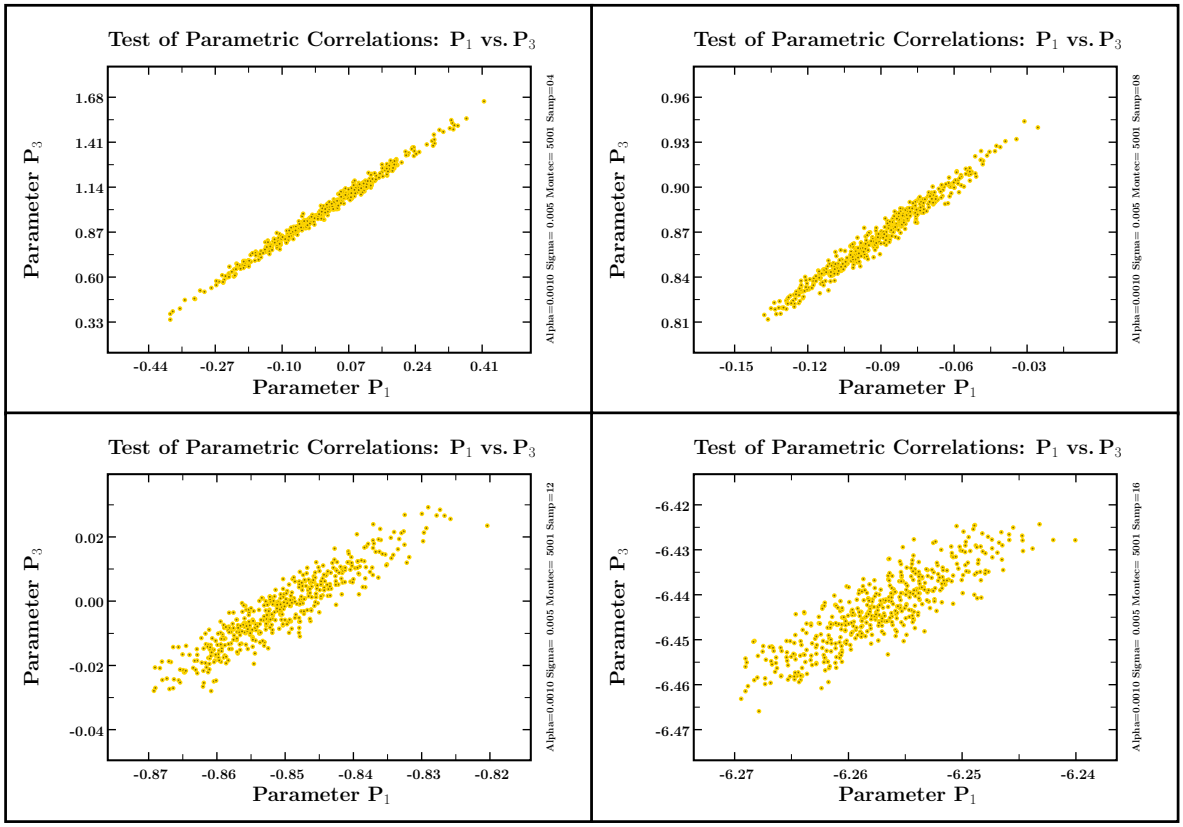

Figure 3.14: *Correlation of parameters P_1 and P_3.*

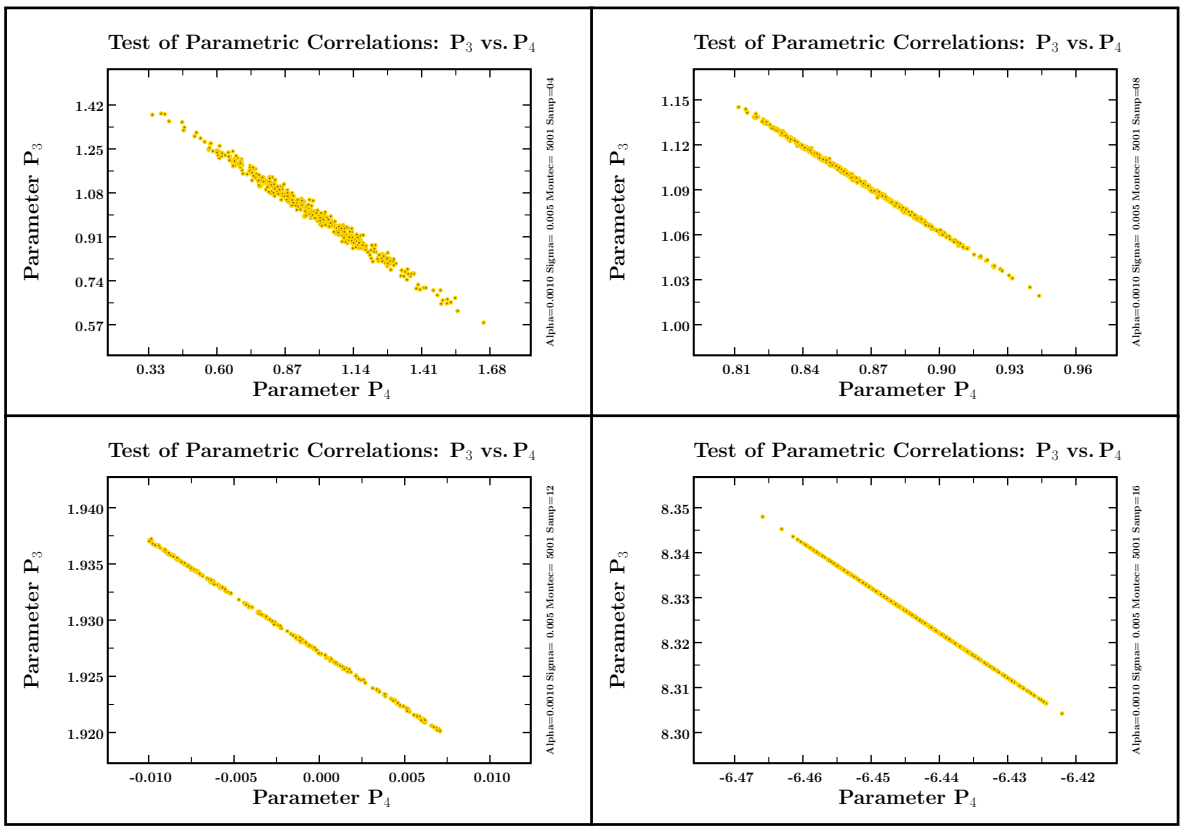

Figure 3.15: *Correlation of parameters P_3 and P_4.*

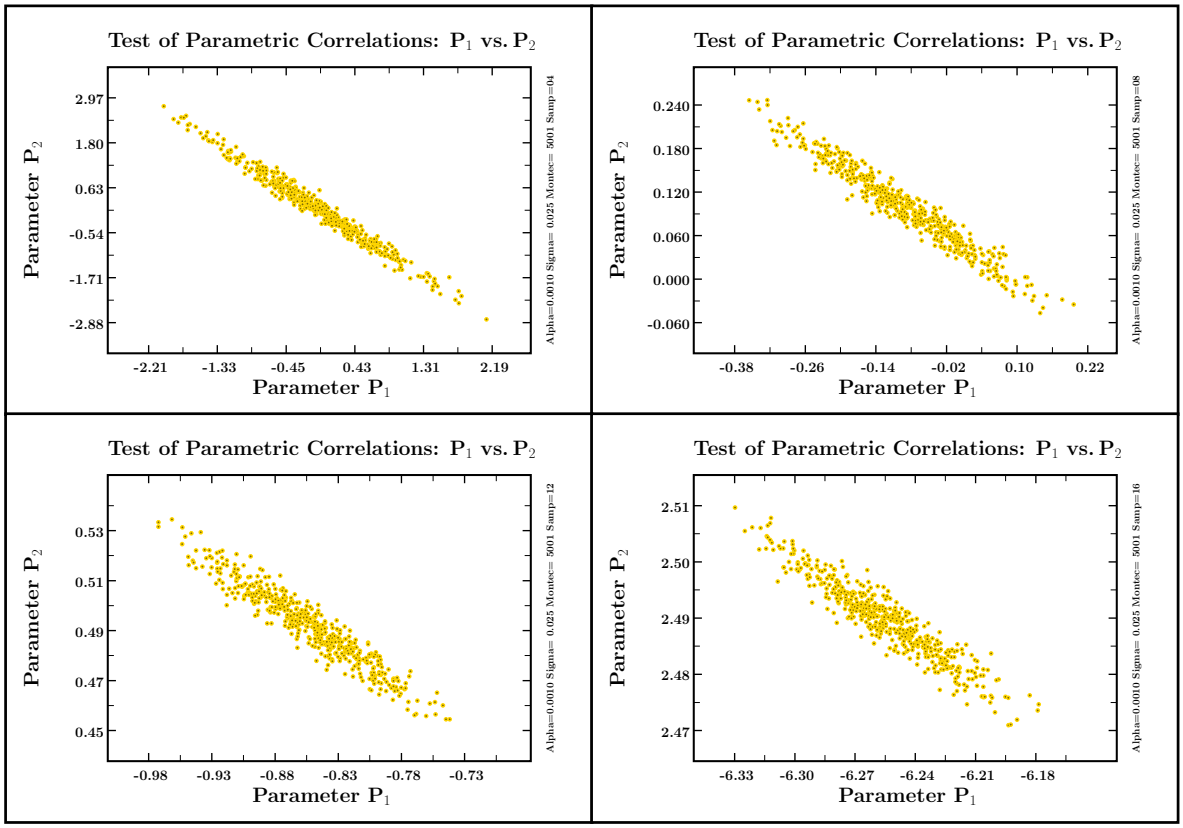

Figure 3.16: *Correlation of parameters P_1 and P_2 for $\alpha = 0.001$ and $\sigma = 0.025$ ('imprecise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

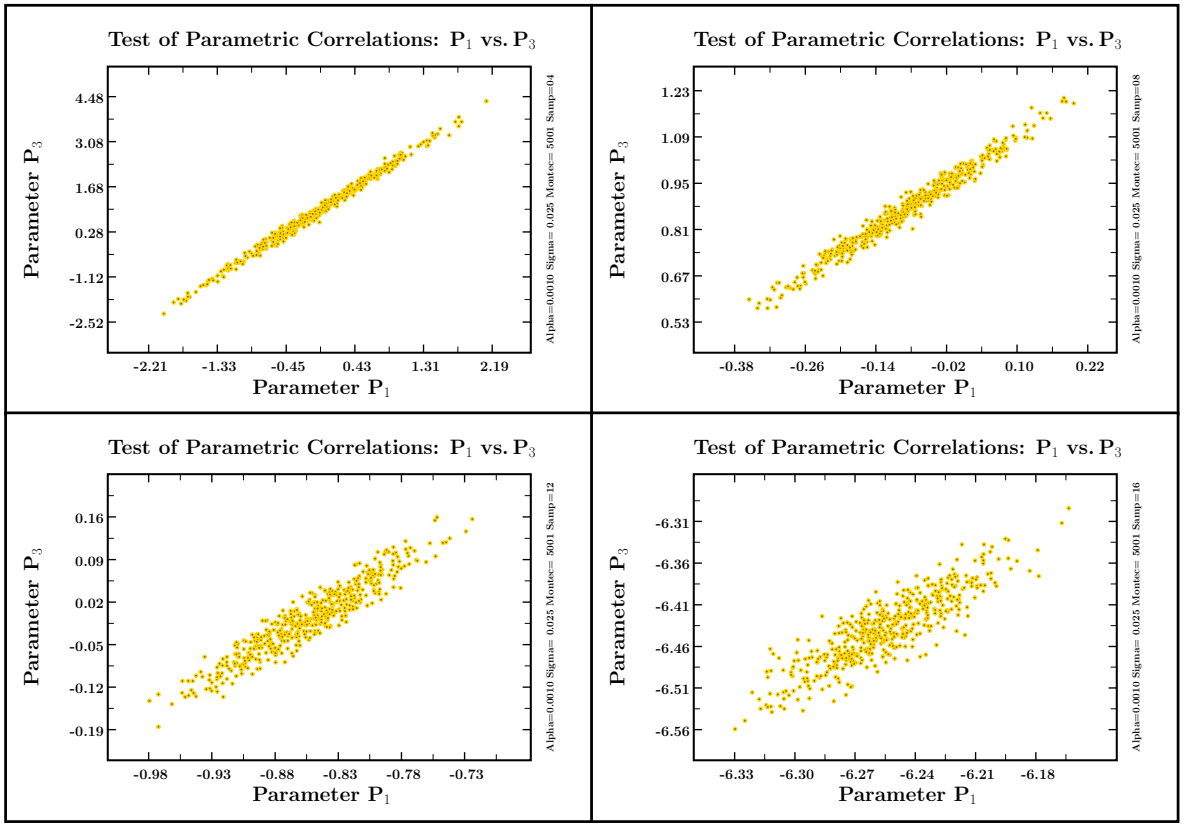

Figure 3.17: *Correlation of parameters P_1 and P_3.*

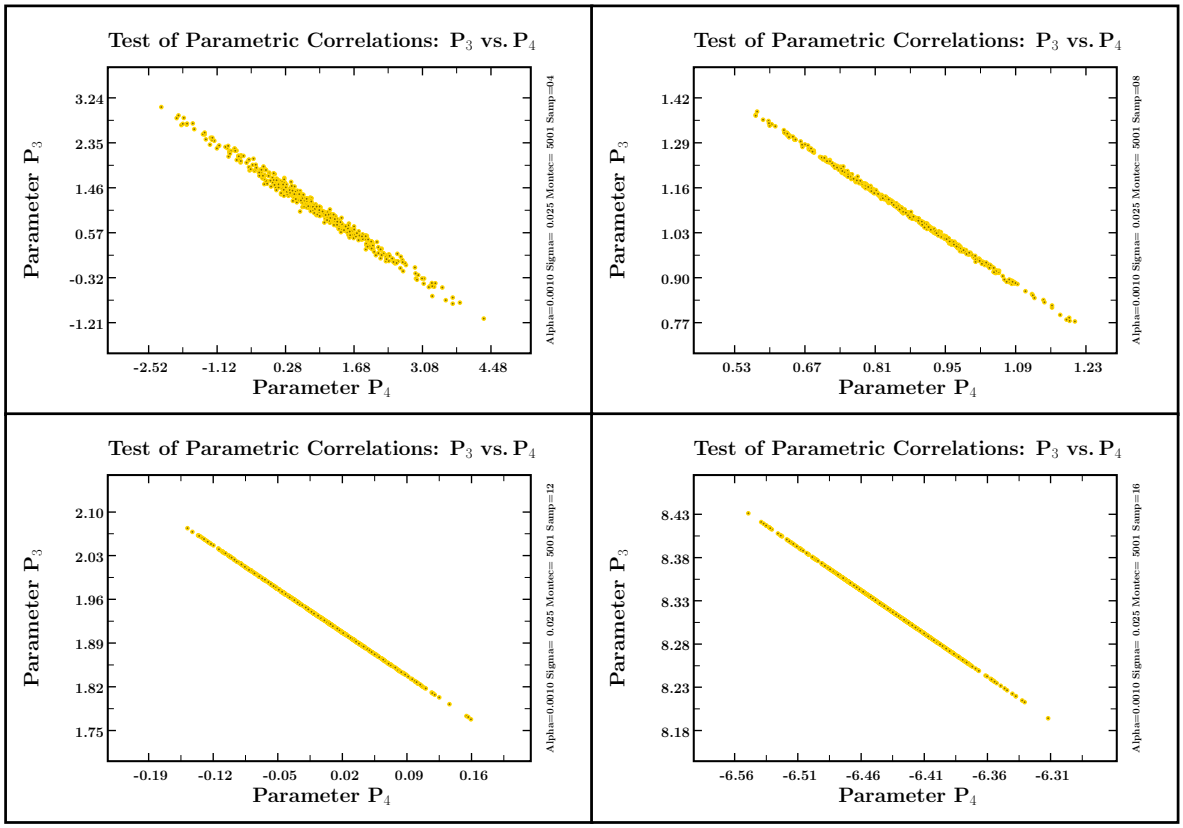

Figure 3.18: *Correlation of parameters P_3 and P_4.*

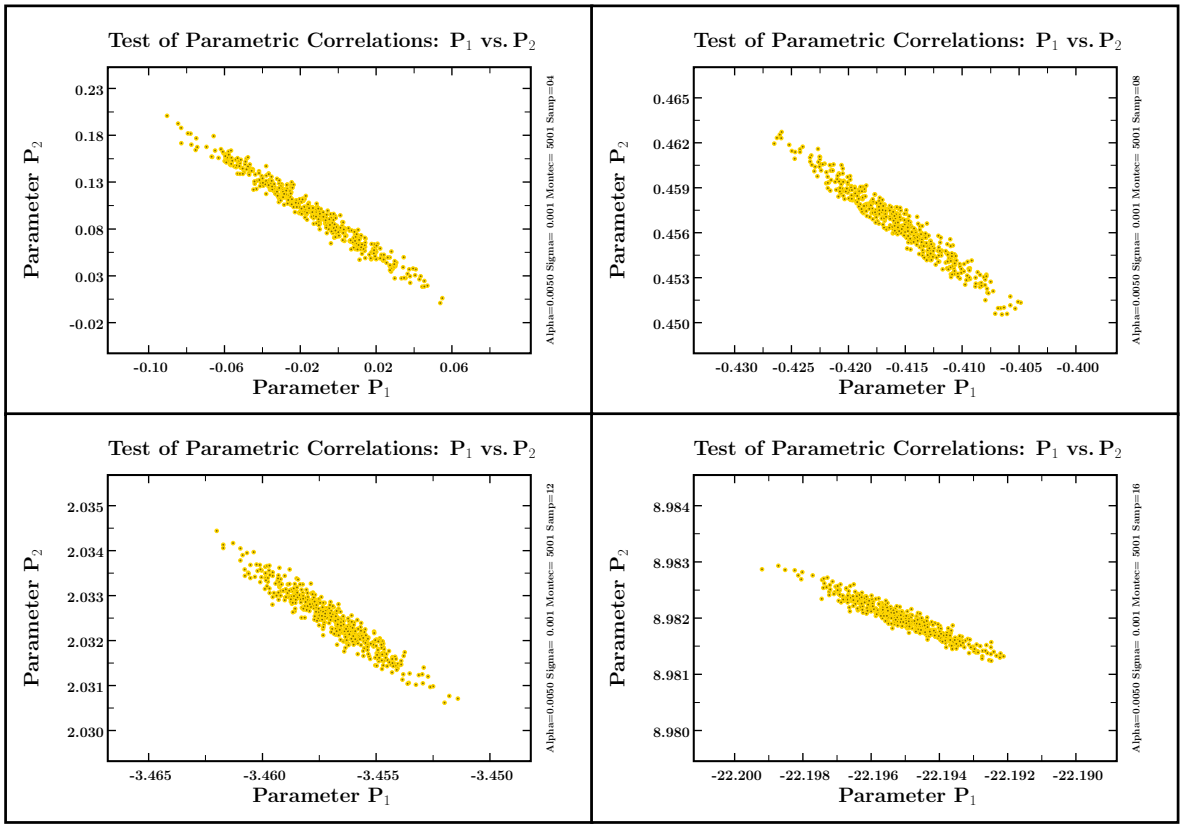

Figure 3.19: *Correlation of parameters P_1 and P_2 for $\alpha = 0.005$ and $\sigma = 0.001$ ('precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

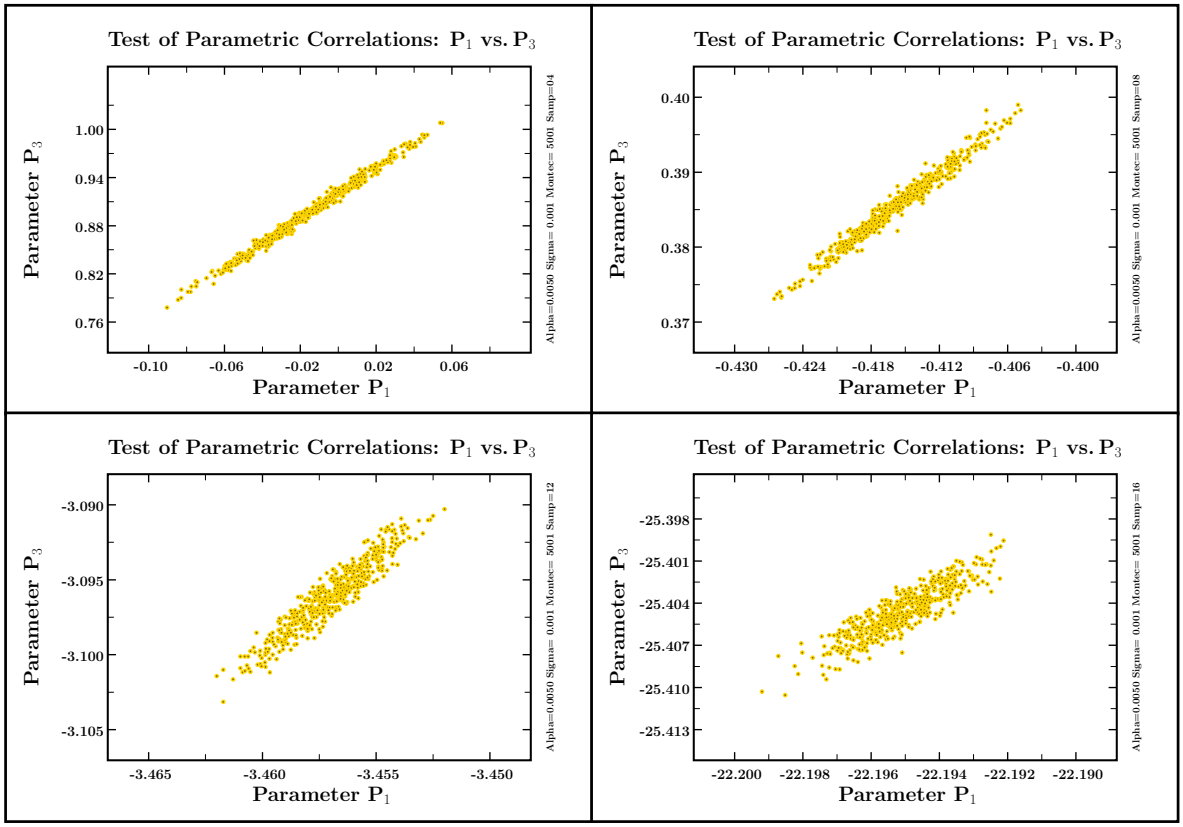

Figure 3.20: *Correlation of parameters P_1 and P_3.*

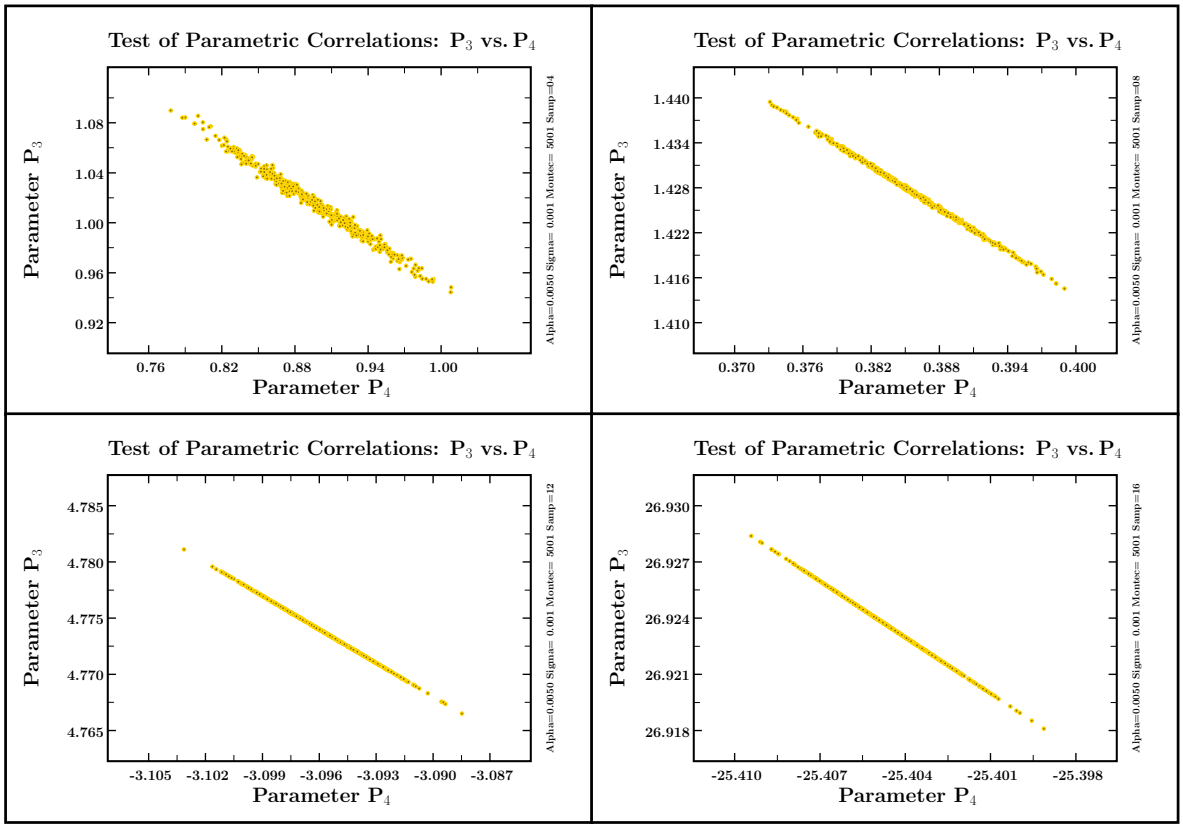

Figure 3.21: *Correlation of parameters P_3 and P_4.*

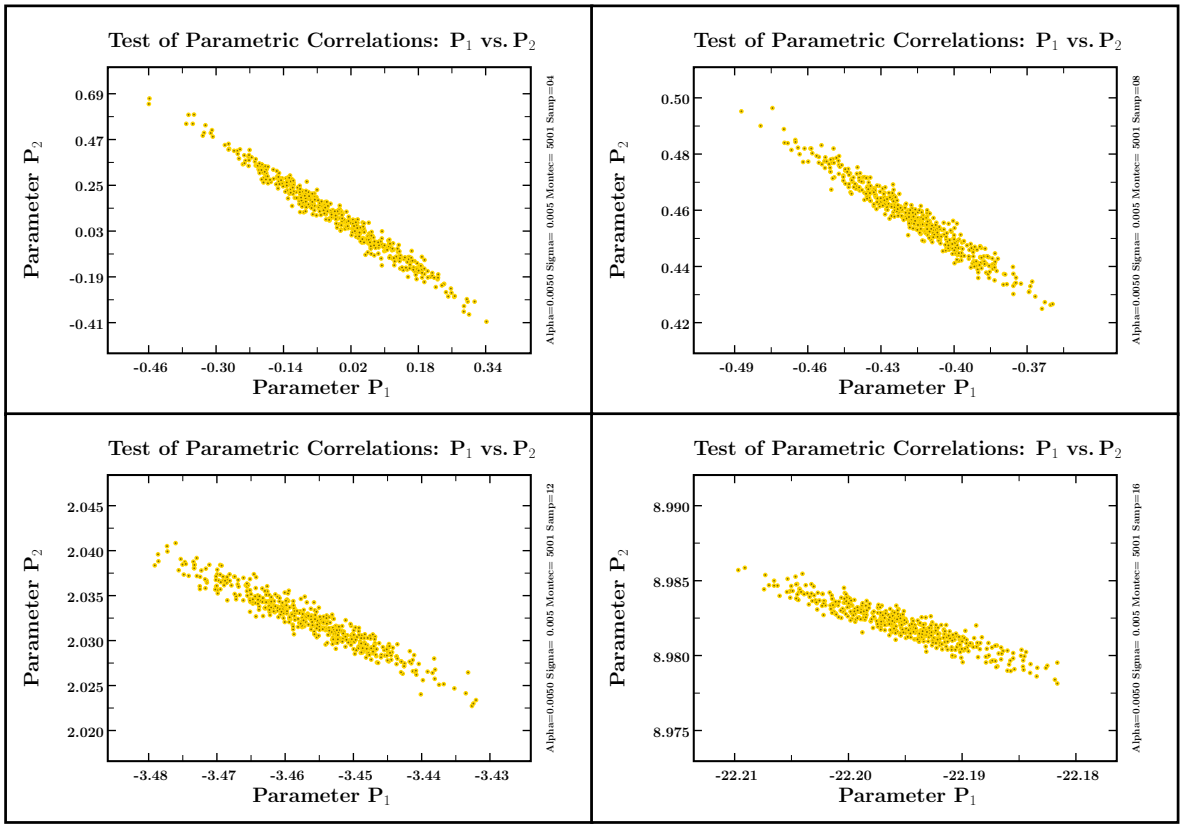

Figure 3.22: *Correlation of parameters P_1 and P_2 for $\alpha = 0.005$ and $\sigma = 0.005$ ('less precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

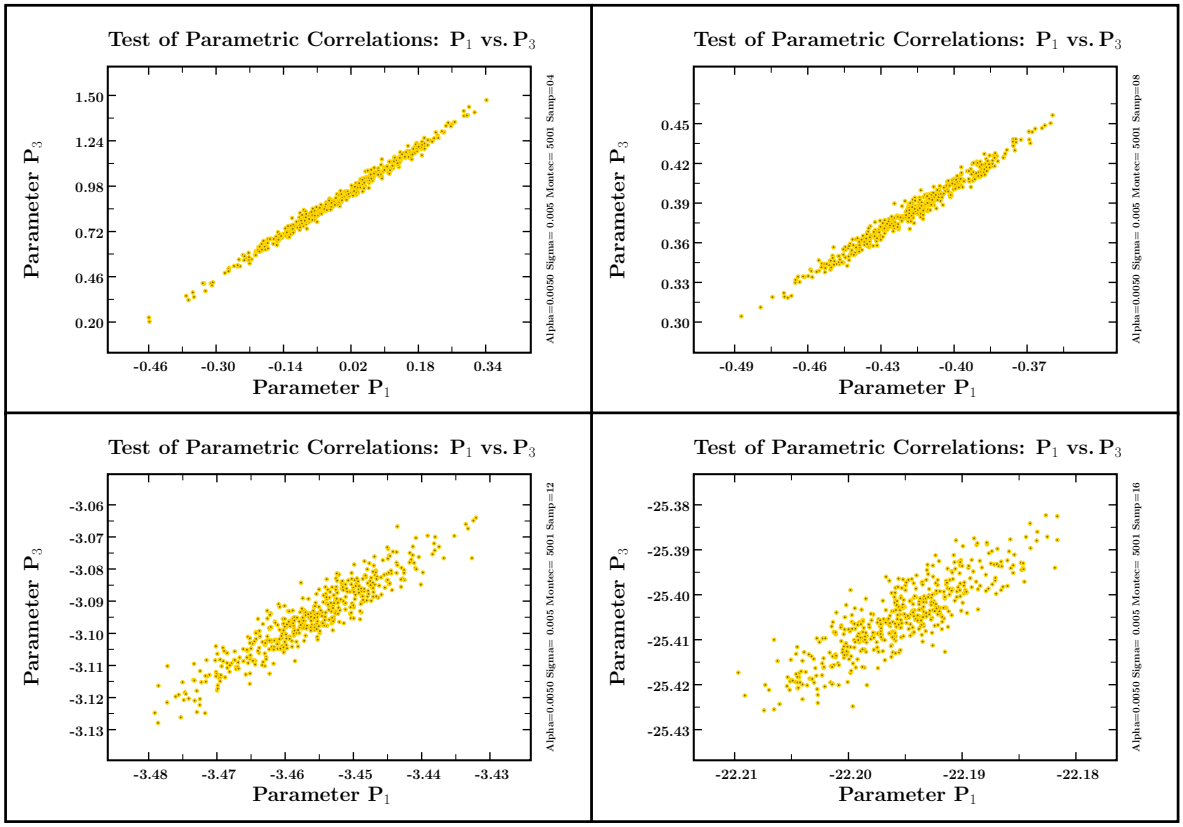

Figure 3.23: *Correlation of parameters P_1 and P_3.*

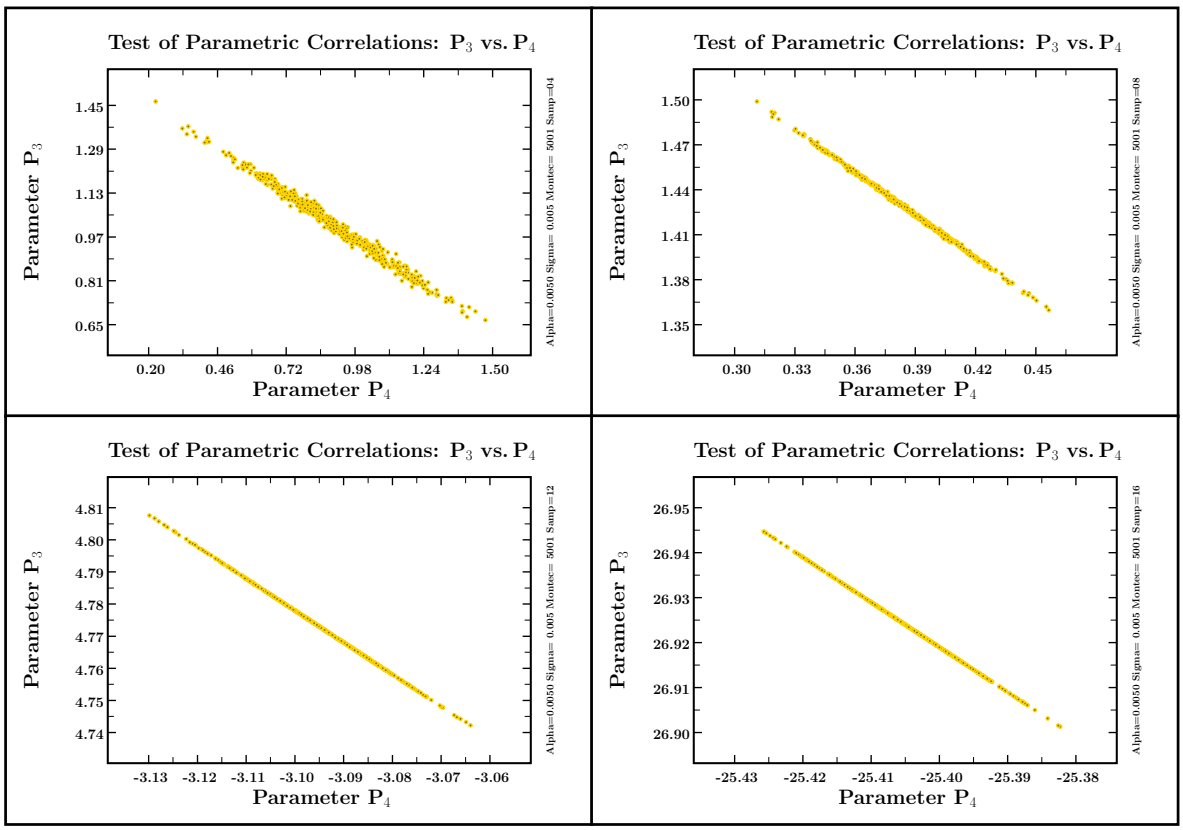

Figure 3.24: *Correlation of parameters P_3 and P_4.*

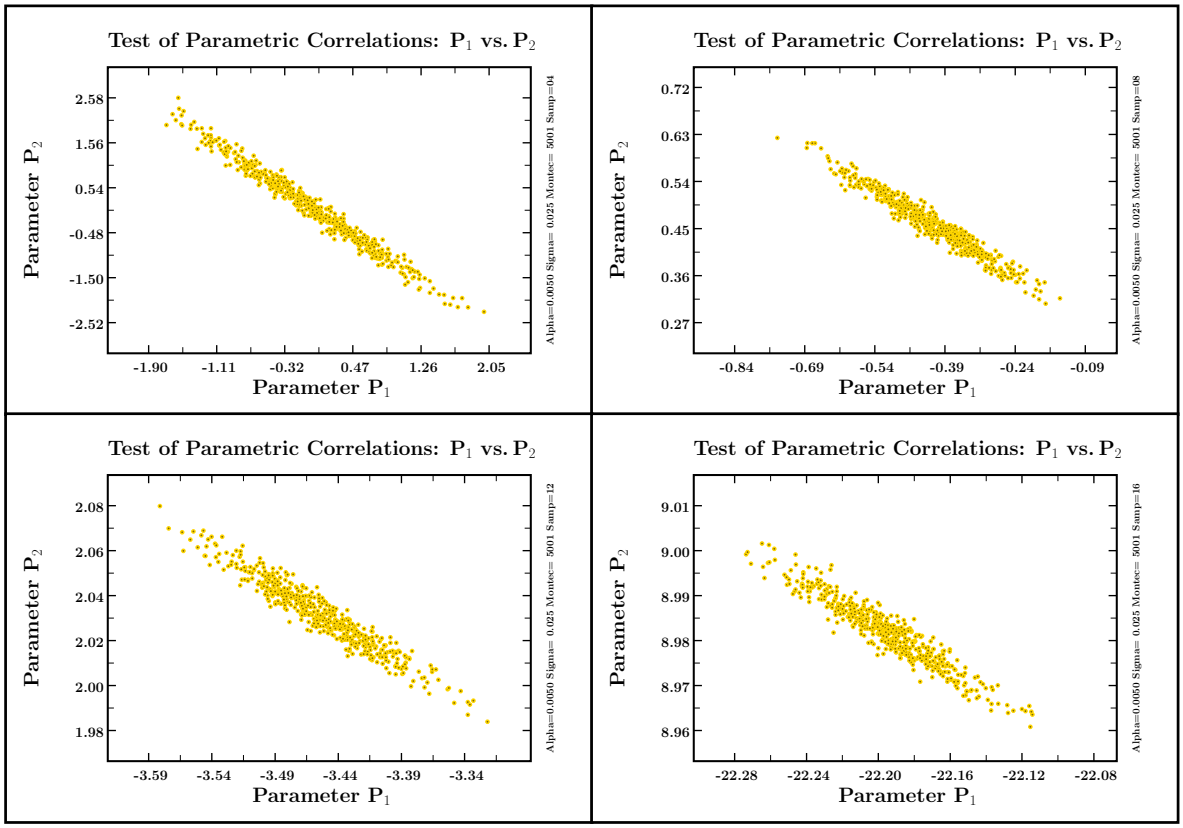

Figure 3.25: *Correlation of parameters P_1 and P_2 for $\alpha = 0.005$ and $\sigma = 0.025$ ('imprecise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

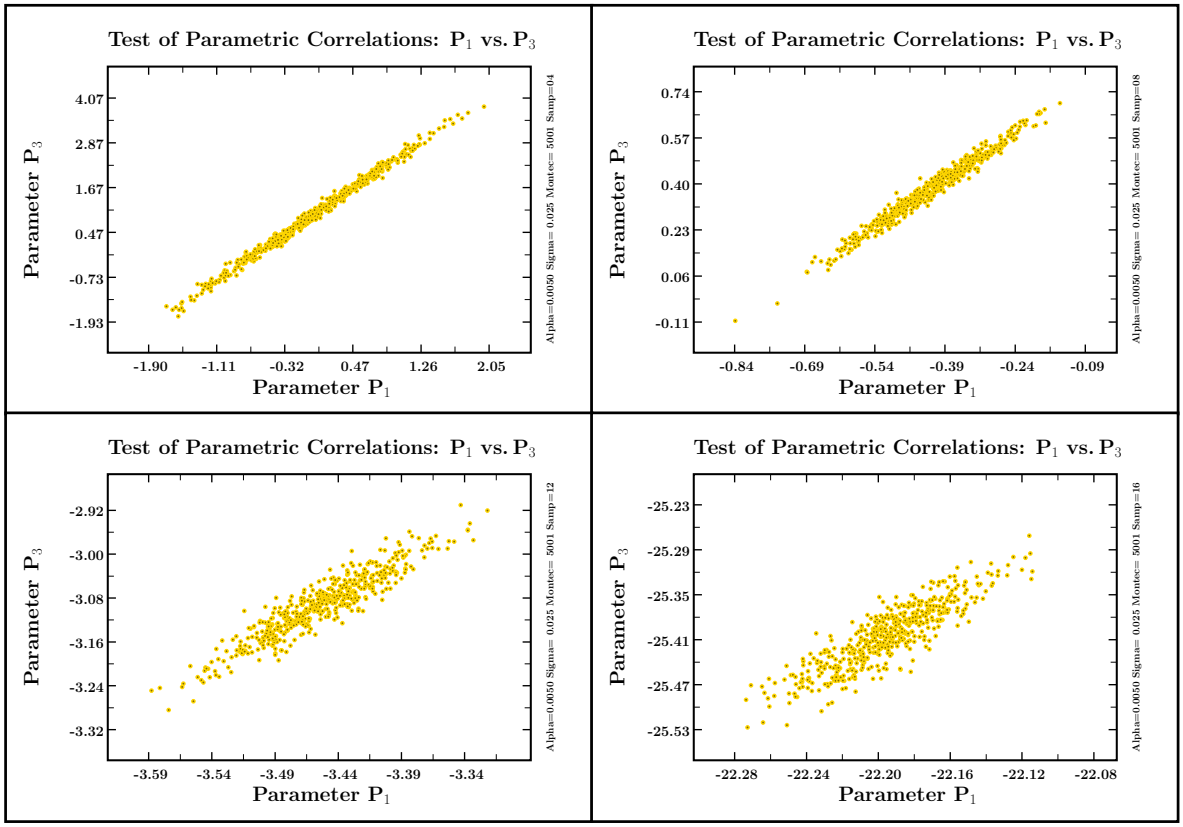

Figure 3.26: *Correlation of parameters P_1 and P_3.*

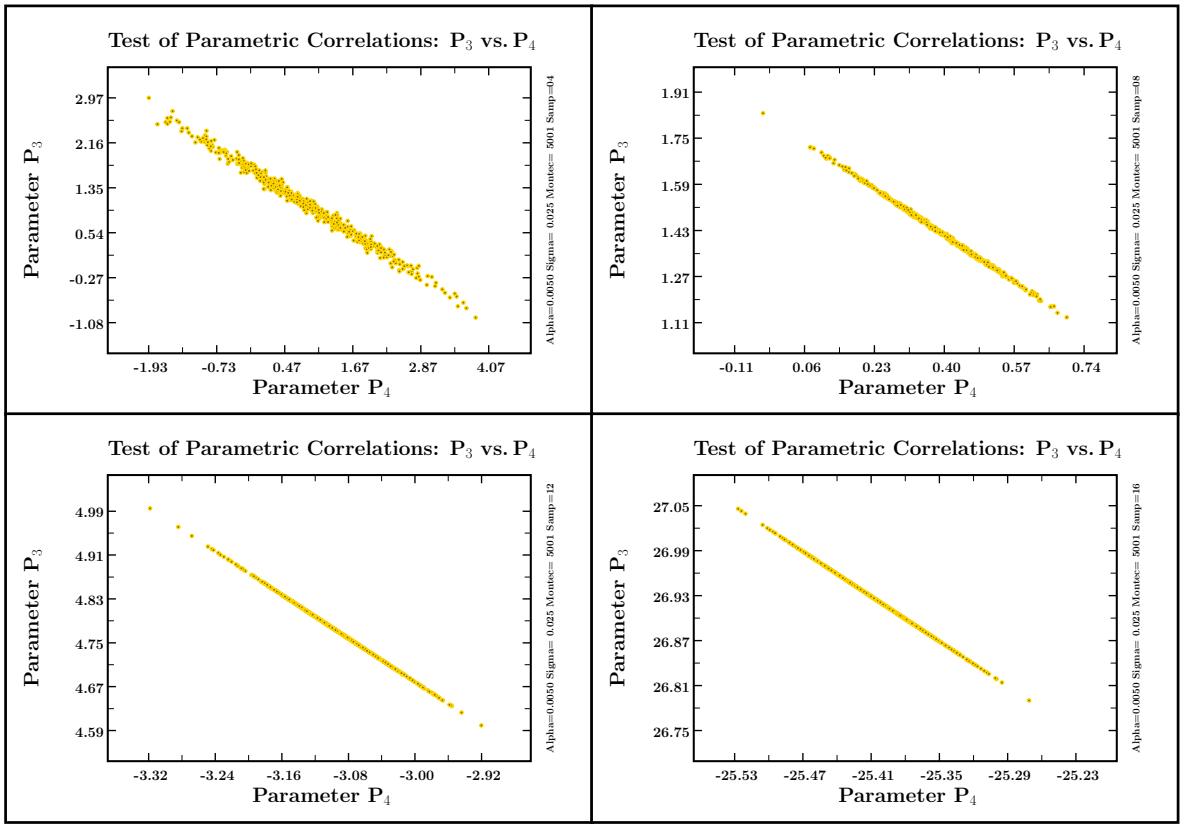

Figure 3.27: *Correlation of parameters P_3 and P_4.*

3.3 Over-fitting

The procedure of fitting the model to a given data set requires that the χ^2 function defined in Eq. 3.11 is as close to zero as possible. However, knowing that the data are affected by the uncertainty component, we should expect, that this level of adjustment corresponding to $\chi^2 \approx 0$ will not be obtained. Unfortunately, this is not as simple as it would seem. If the model is ill-posed and there a parametric correlation between its parameters, there is no possibility to recover the optimal parameter set. The parameters will vary depending on the data we fit, every time resulting in a very small values of the χ^2 function. This is known as the 'over-fitting' mechanism and may cause large problems if one does not consider the importance of the proper formulation of the model. It can be remedied by increasing the size of the sample, which may remove or diminish the impact of parametric correlations of the model on the data predictions.

To present the over-fitting phenomenon, we plot the histograms for the χ^2 function, which means that we count the number of occurrences of a given χ^2 value. The results are presented in the following figures 3.28-3.33. Each figure consists of six plots organised in such a way, that one column corresponds to one sampling option, i.e. the model can be fitted to 6,8,12 or 16 data points. The plots in a given column present the results for different type of input data: the top figure is plotted for 'precise data' variant with $\sigma = 0.001$, the middle one correspond to 'less precise data' ($\sigma = 0.005$) and the bottom one present the results for the 'imprecise data' variant ($\sigma = 0.025$).

The first two plots correspond to the 'exact model' version with $\alpha = 0$. The χ^2 values remain rather small, although they differ by orders of magnitude. What we can observe is that the smaller the size of the sample, the closer the maximum of the distribution is to zero and actually, for sample of size 4 (which is not plotted here), it becomes a single line at zero. Increasing the size of the sample results in a slightly shifted distribution (sample of 6 points – cf. Fig. 3.28, left column) but still the maximum remains at zero. Futher adding of data points shifts the maximum towards right. We can also compare the influence of the data precision on the χ^2 function distribution. As it should be expected, the distributions become broader if we increase the σ values, but additionaly, they also shift toward right.

The figures 3.30 and 3.31 present the result for non-analytical case of $\alpha = 0.001$, whereas the last pair 3.32 and 3.33 show the results for larger value of $\alpha = 0.005$. What we can observe here is a rather similar pattern as before in the analytical case of exponential function. Small sample results in over-fitting phenomenon, even if the experimental uncertainties are large. However, increasing the sample shifts the position of the maxium much further away from zero than before. This effect is the strongest in the case of the largest sample of 16 points for $\alpha = 0.005$ – the maximum is around $\chi^2 = 325$. The width of the χ^2 distribution also grows for this case.

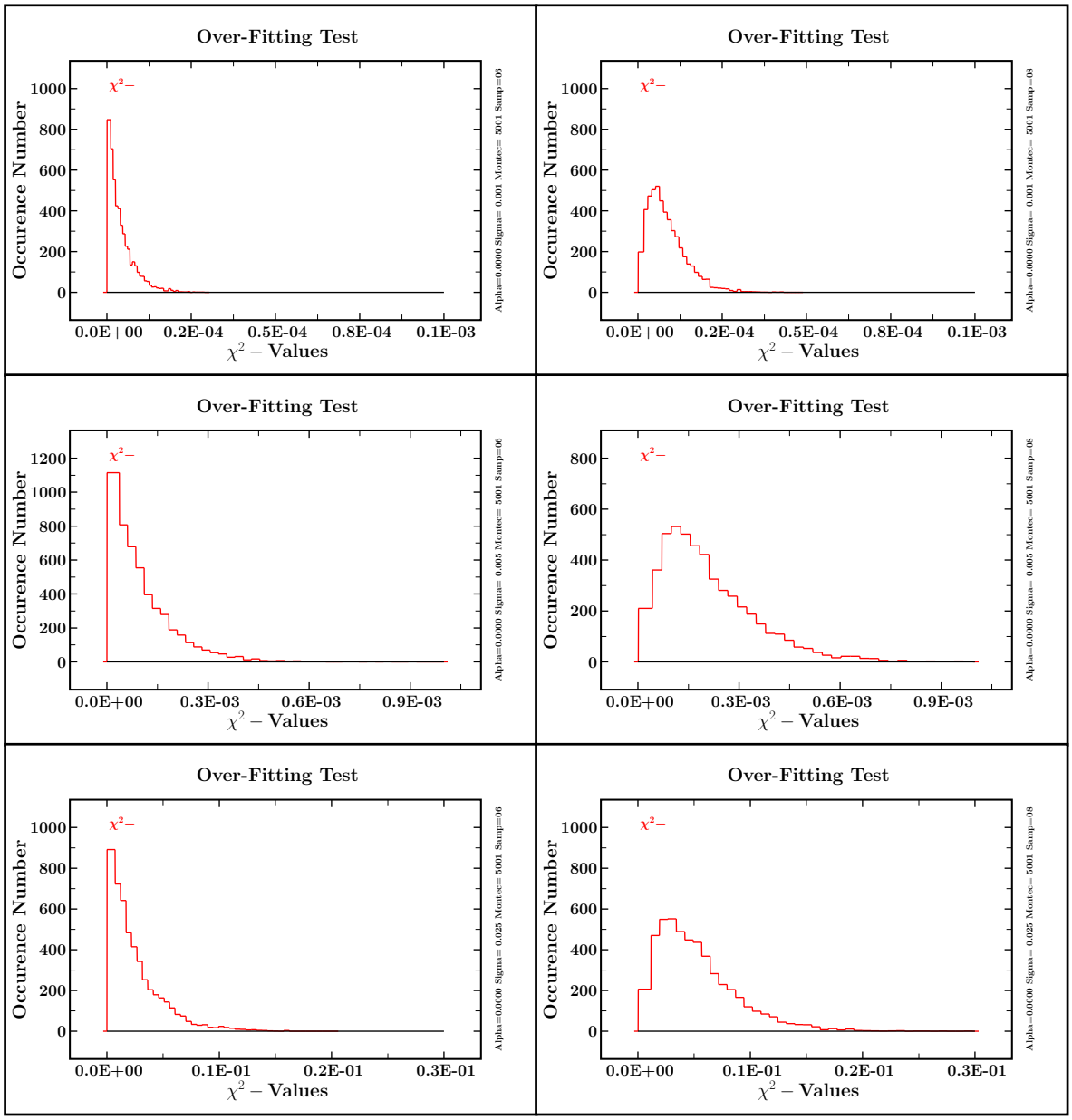

Figure 3.28: *Over-fitting test for $\alpha = 0$ ('exact model' version) with different types of input data for two sampling options: 6 points – left column, 8 points – right column; precise data (small σ) - top, moderately precise data ('medium' σ) - middle; poorly determined data (relatively large σ) - bottom.*

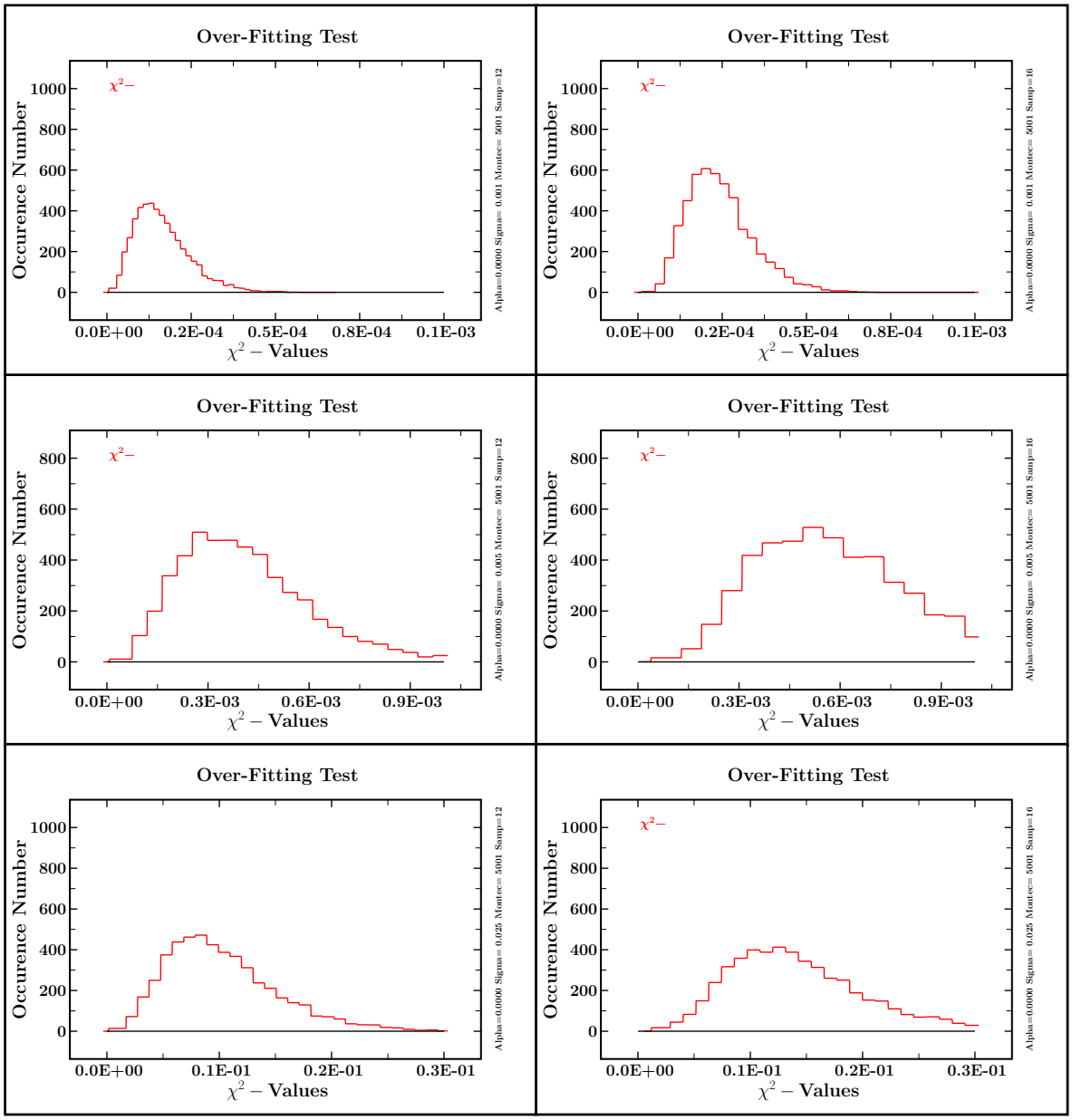

Figure 3.29: *Over-fitting test for $\alpha = 0$ ('exact model' version) with different types of input data for two sampling options: 12 points – left column, 16 points – right column; precise data (small σ) - top, moderately precise data ('medium' σ) - middle; poorly determined data (relatively large σ) - bottom.*

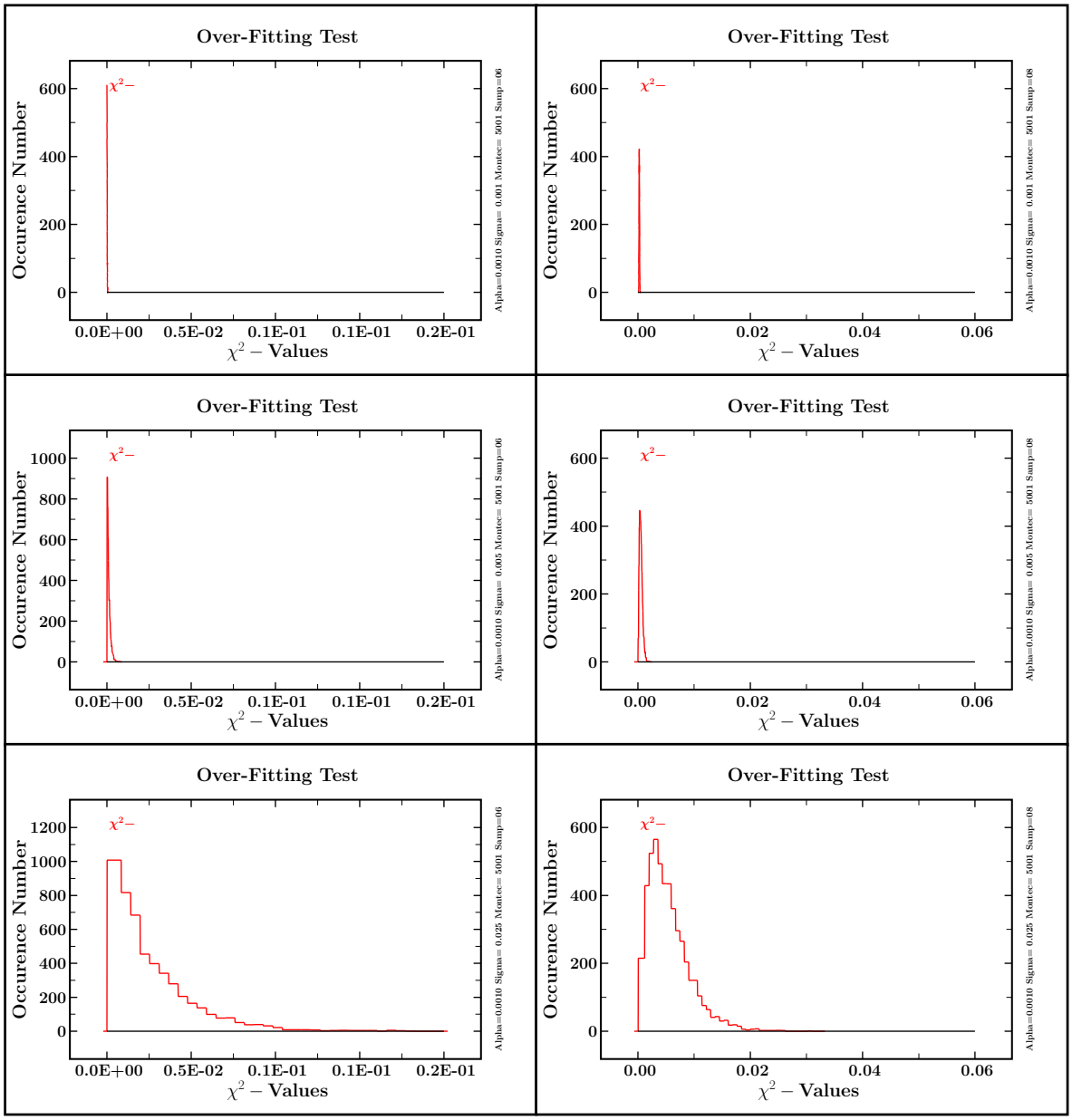

Figure 3.30: *Over-fitting test for $\alpha = 0.001$ with different types of input data for two sampling options: 6 points – left column, 8 points – right column; precise data (small σ) - top, moderately precise data ('medium' σ) - middle; poorly determined data (relatively large σ) - bottom.*

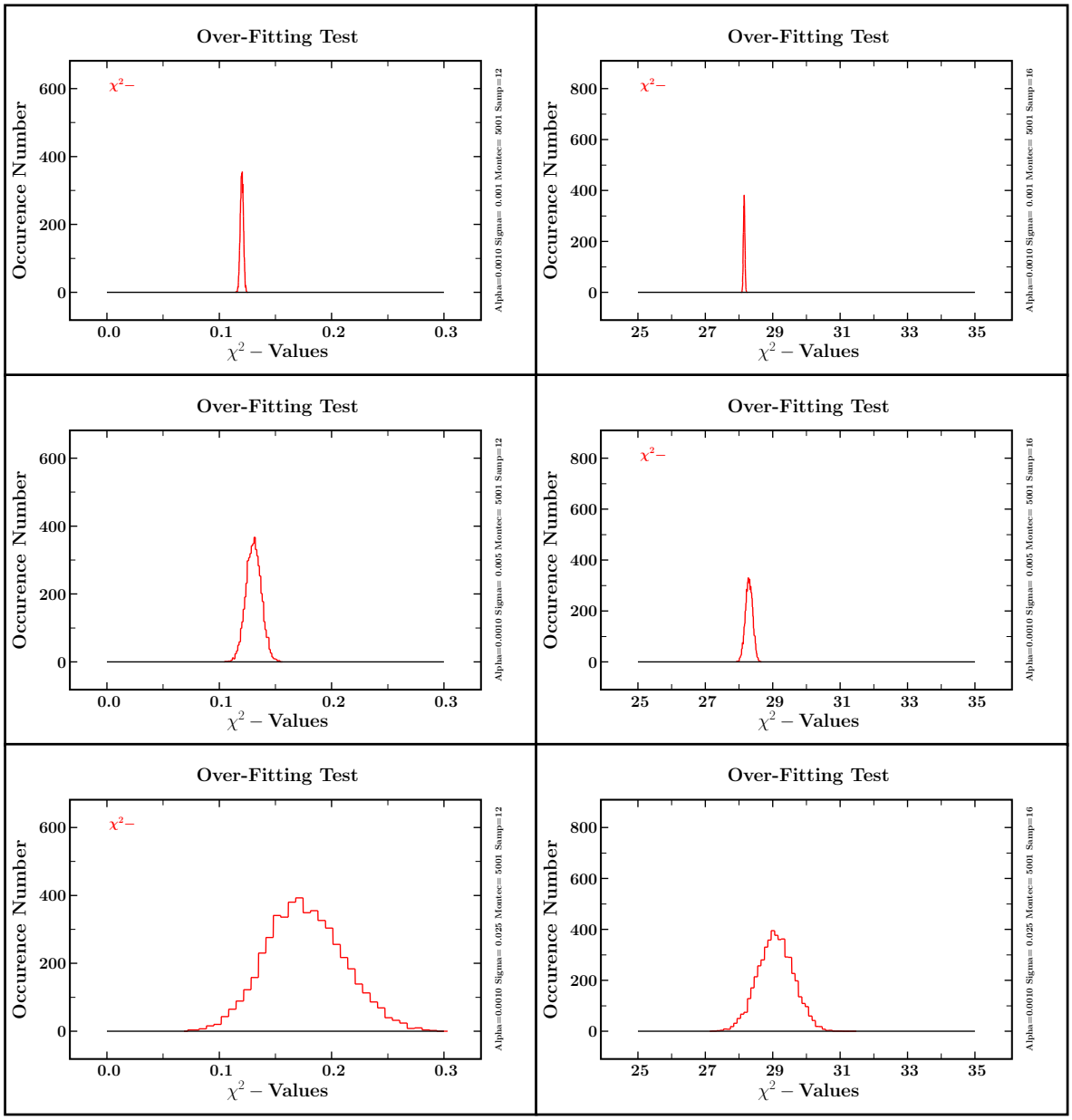

Figure 3.31: *Over-fitting test for $\alpha = 0.001$ with different types of input data for two sampling options: 12 points – left column, 16 points – right column; precise data (small σ) - top, moderately precise data ('medium' σ) - middle; poorly determined data (relatively large σ) - bottom.*

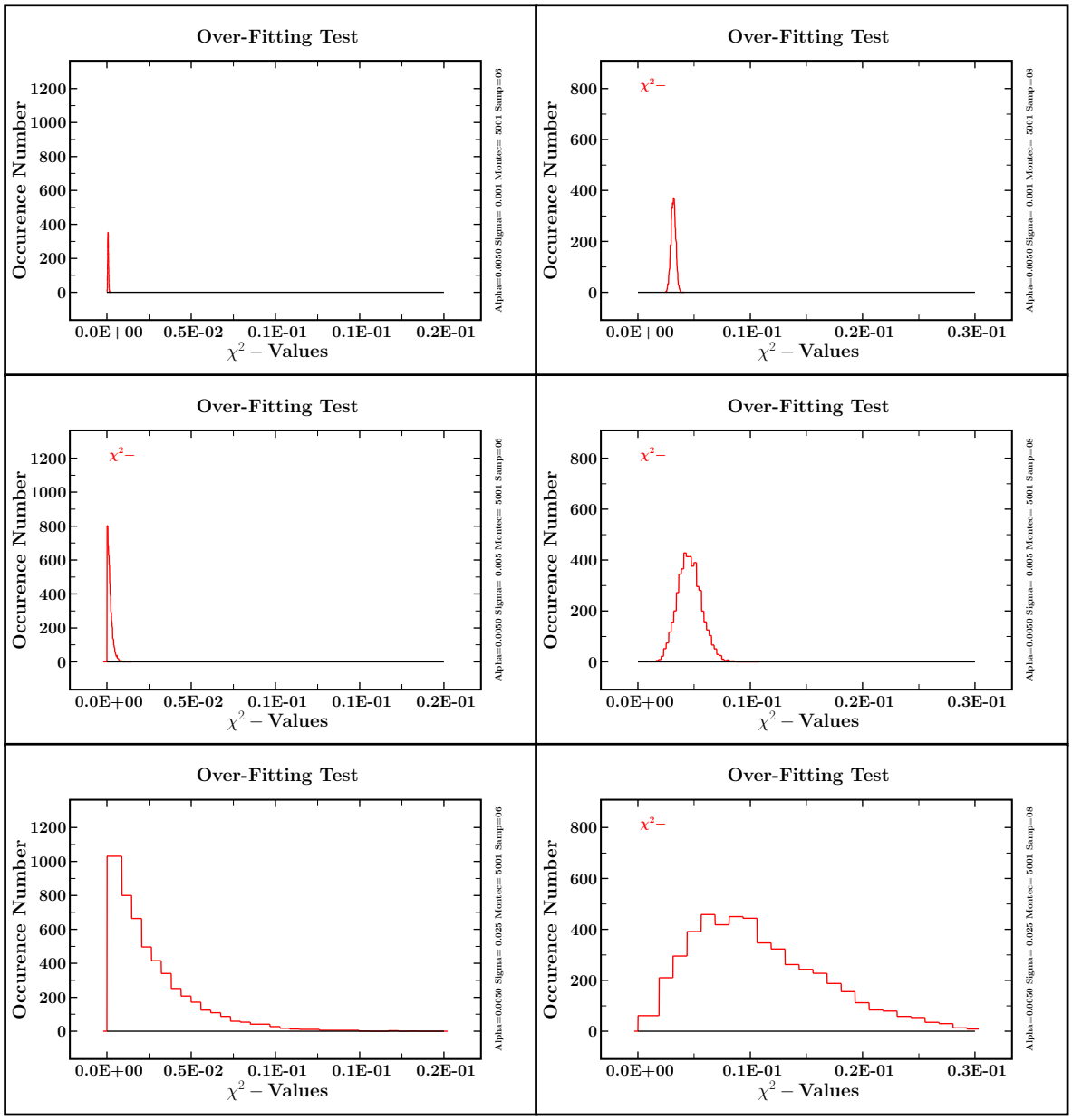

Figure 3.32: *Over-fitting test for $\alpha = 0.005$ with different types of input data for two sampling options: 6 points – left column, 8 points – right column; precise data (small σ) - top, moderately precise data ('medium' σ) - middle; poorly determined data (relatively large σ) - bottom.*

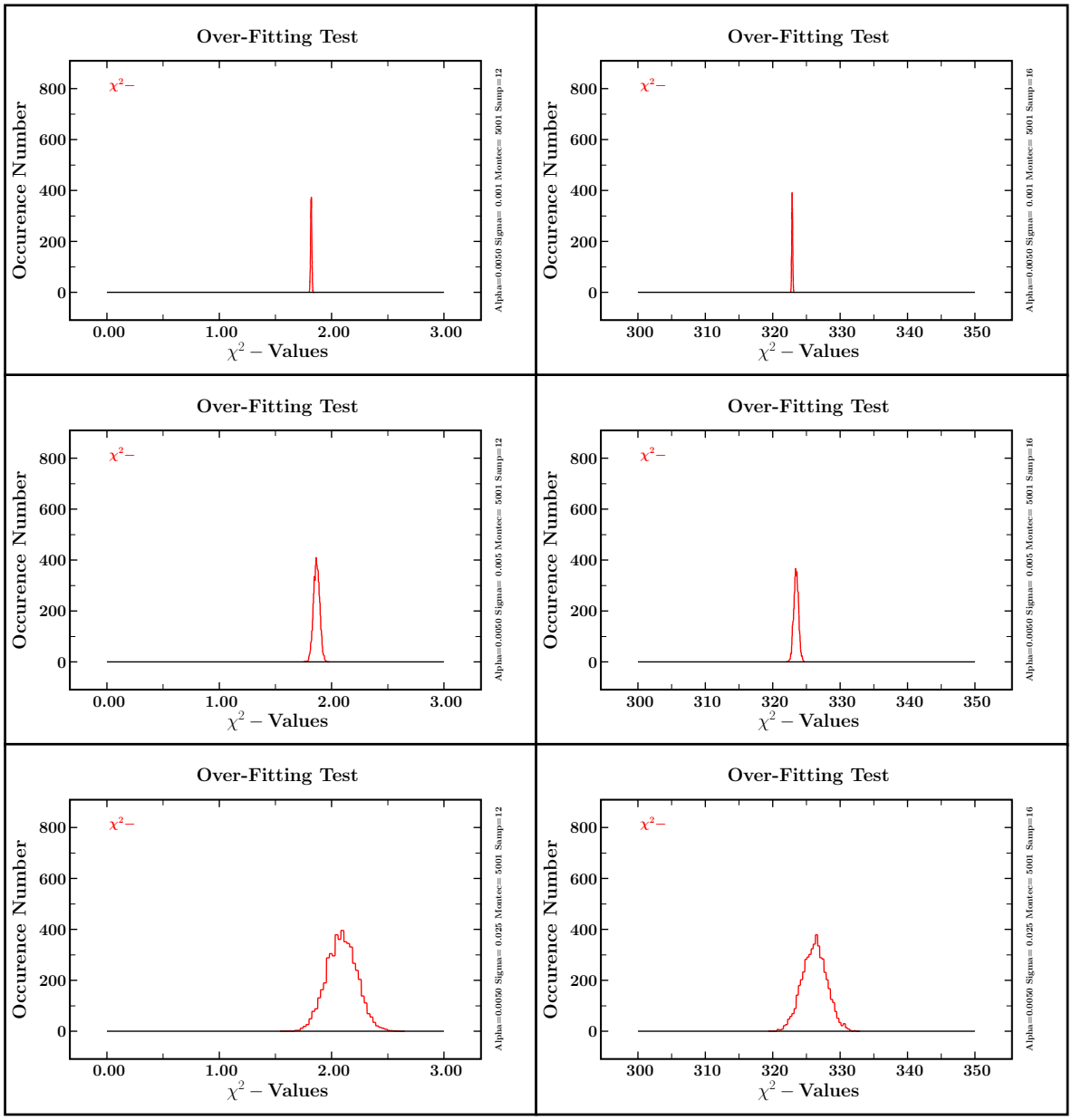

Figure 3.33: *Over-fitting test for $\alpha = 0.005$ with different types of input data for two sampling options: 12 points – left column, 16 points – right column; precise data (small σ) - top, moderately precise data ('medium' σ) - middle; poorly determined data (relatively large σ) - bottom.*

3.4 Monte Carlo Error Estimation

The concept of 'theoretical' errors, i.e. the uncertainties of the parameters of the model can be easily presented using the Monte Carlo approach. This time we simply plot the histograms for parameter values, which gives us the information about the frequency of occurrence of a given value.

The plots are presented in figures 3.34-3.42. Each figure contains the results for one data variant and a given version of the model. We can observe how the increasing imprecision of the data affects the widths of the distributions for every model, regardless of the α parameter.

The figures illustrate also how the 'optimal fit parameters' deviate from the exact solution ($P_1 = 0,\ P_2 = 0,\ P_3 = 1,\ P_4 = 1$) when the three quantities that characterise the model (sampling, experimental input imprecision, inexactitude level of the modelling) are modified.

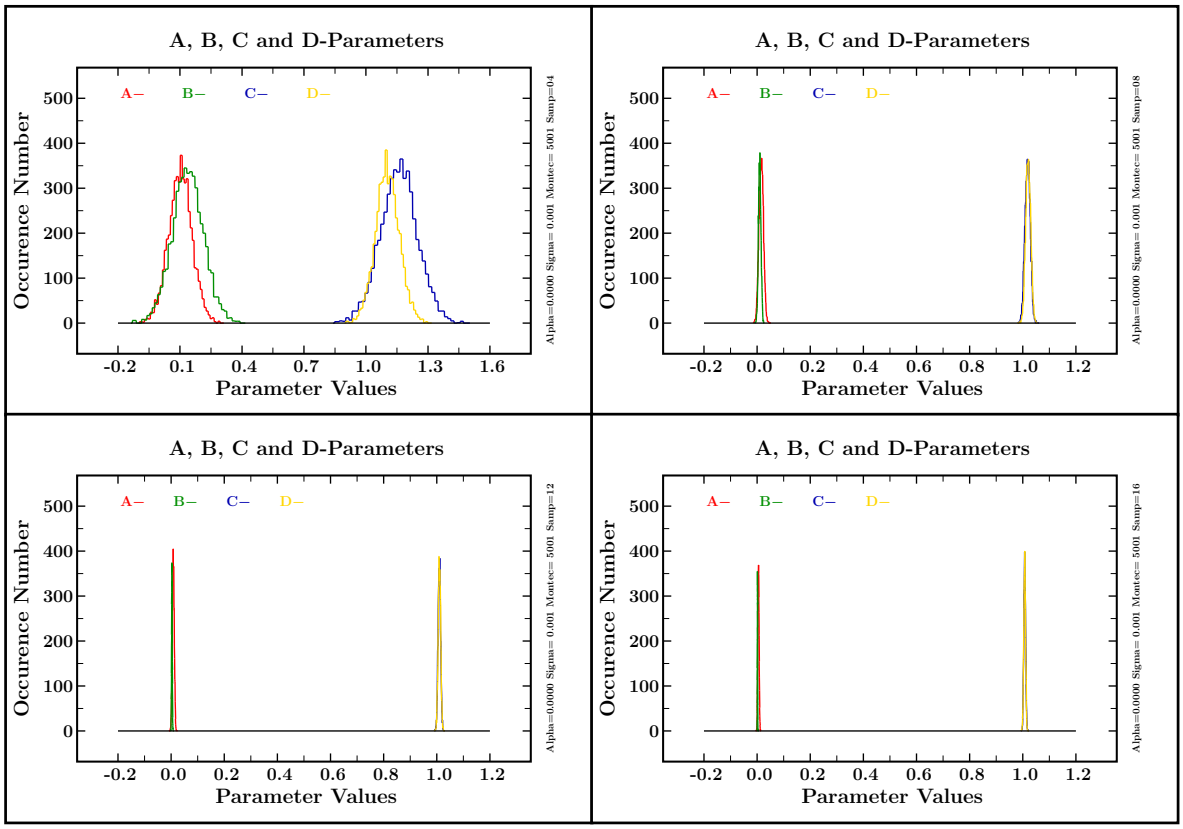

Figure 3.34: *Probability distributions of parameters for $\alpha = 0$ ('exact model' version) and $\sigma = 0.001$ ('precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

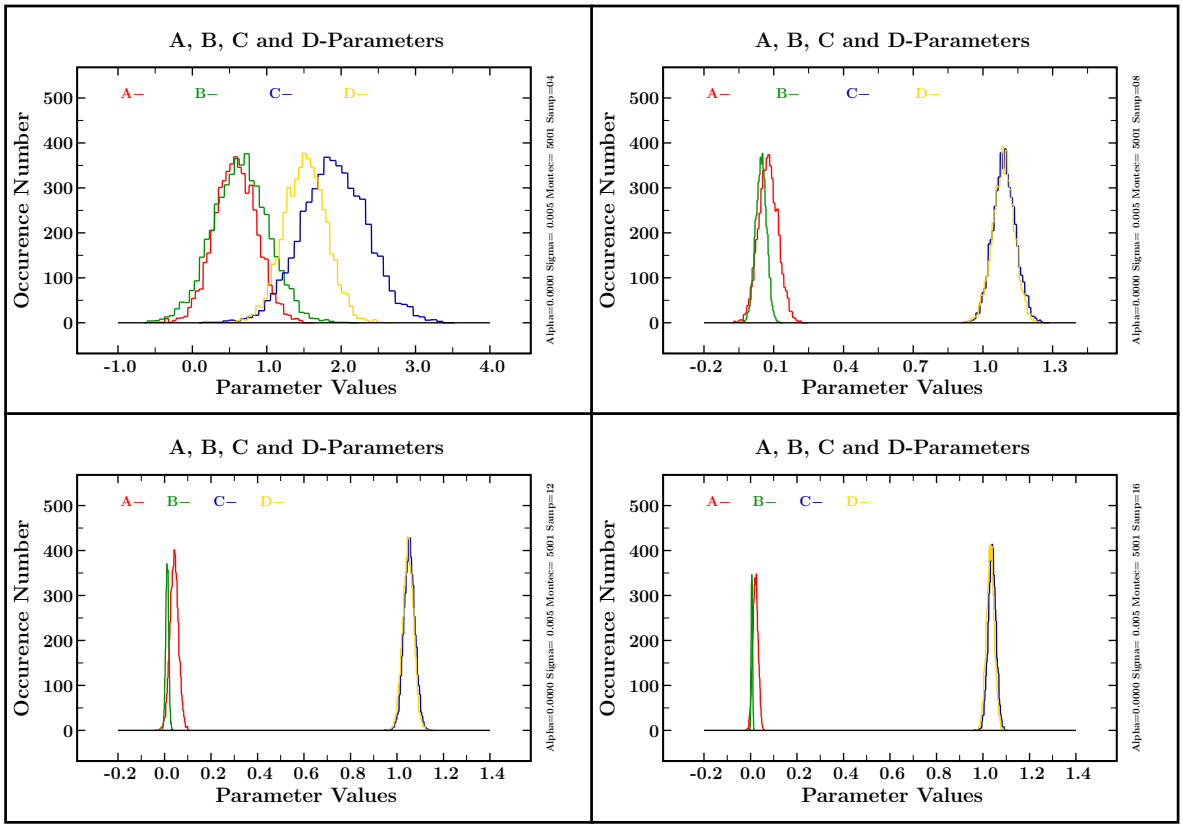

Figure 3.35: *Probability distributions of parameters for $\sigma = 0.005$ ('less precise data').*

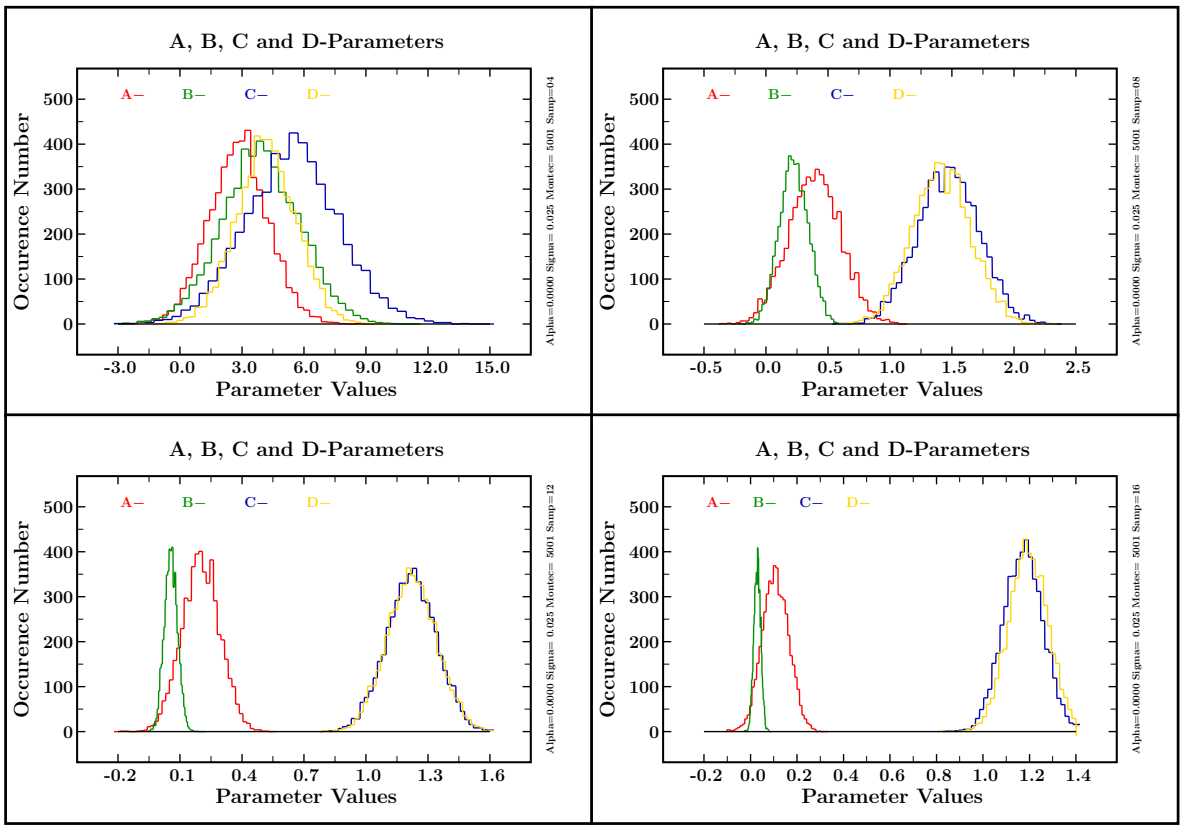

Figure 3.36: *Probability distributions of parameters for $\sigma = 0.025$ ('imprecise data').*

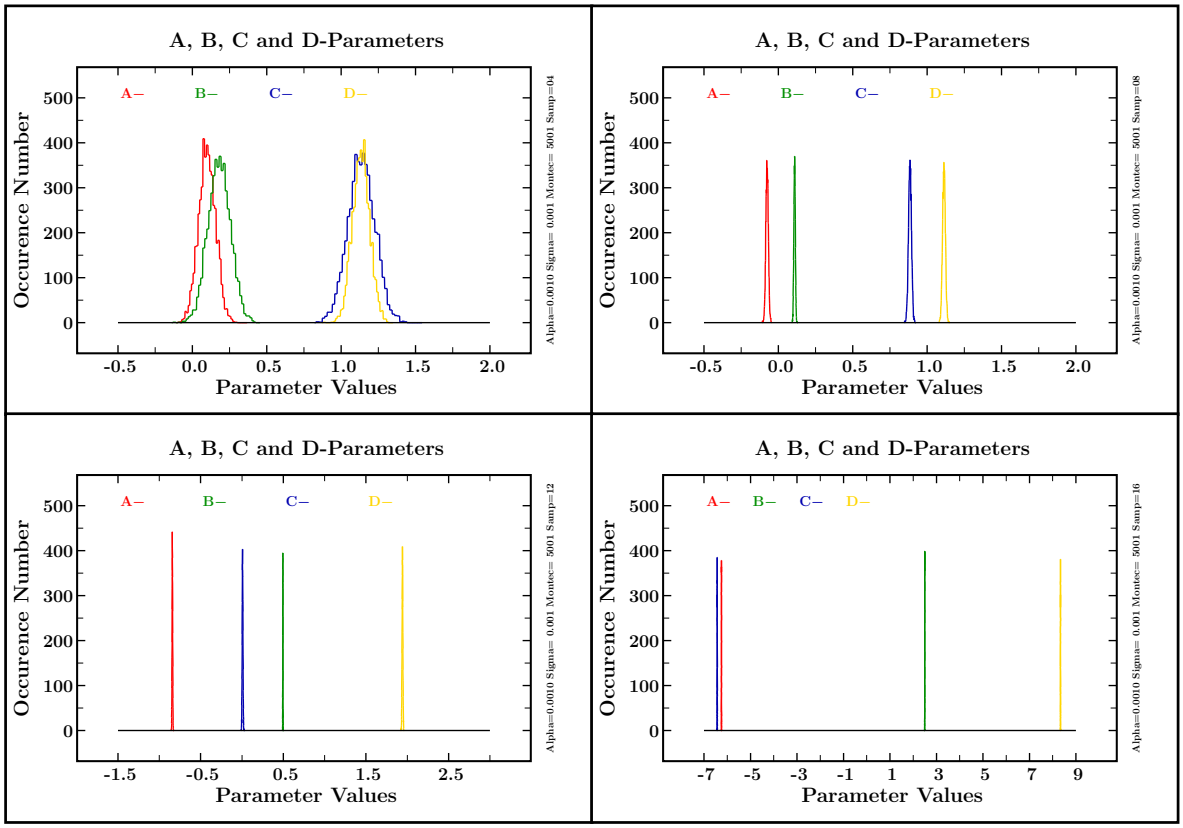

Figure 3.37: *Probability distributions of parameters for $\alpha = 0.001$ and $\sigma = 0.001$ ('precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

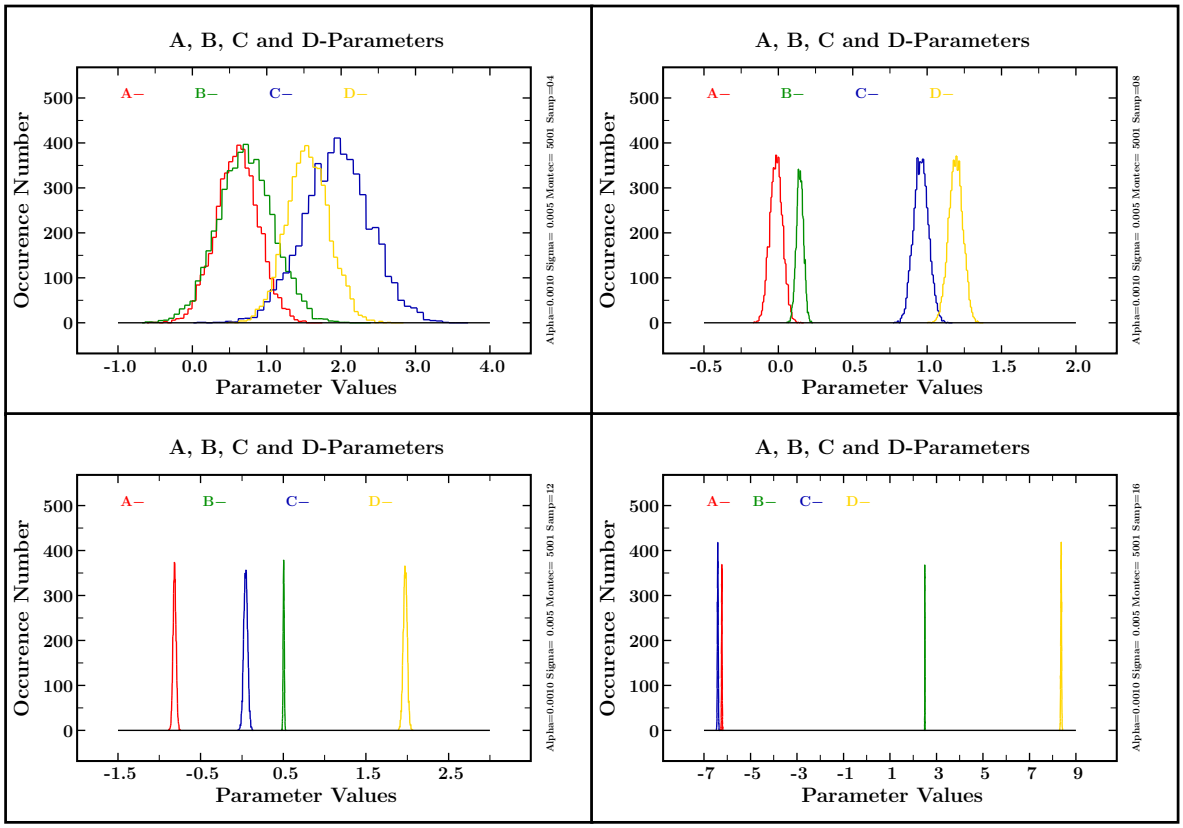

Figure 3.38: *Probability distributions of parameters for $\sigma = 0.005$ ('less precise data').*

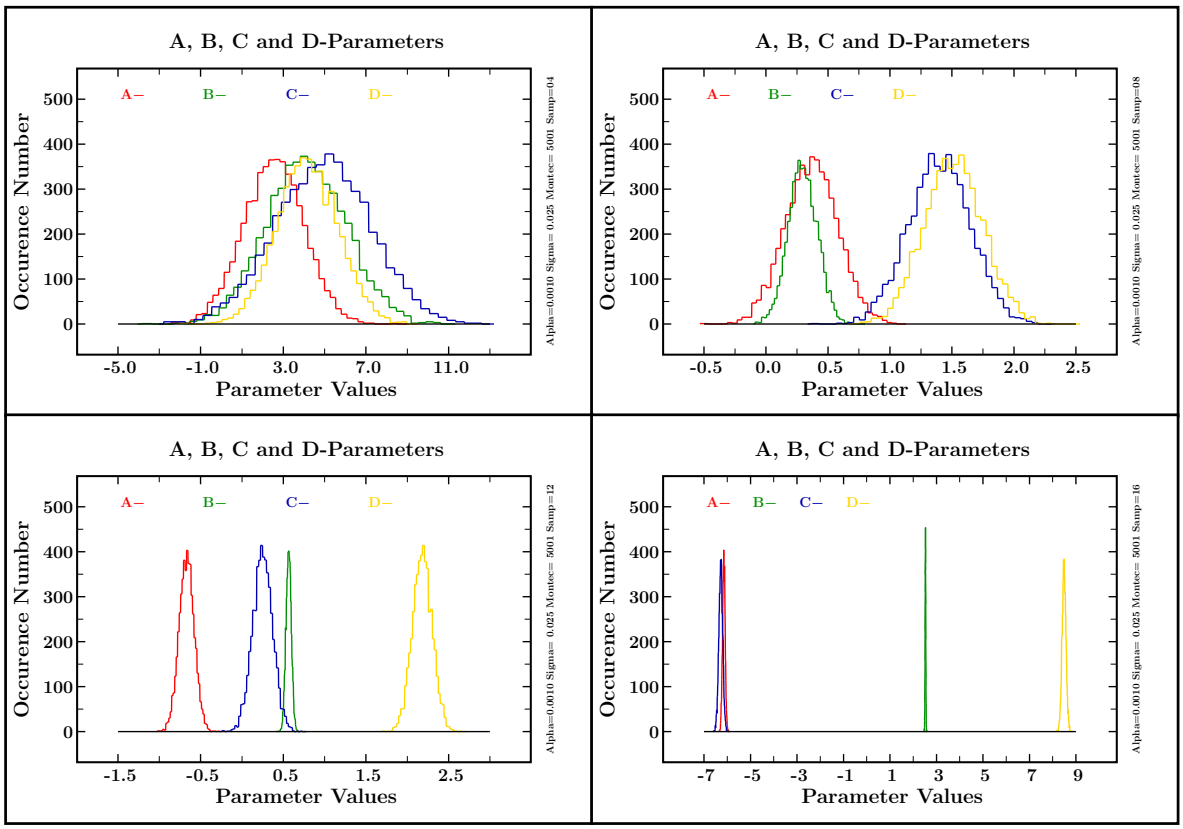

Figure 3.39: *Probability distributions of parameters for $\sigma = 0.025$ ('imprecise data').*

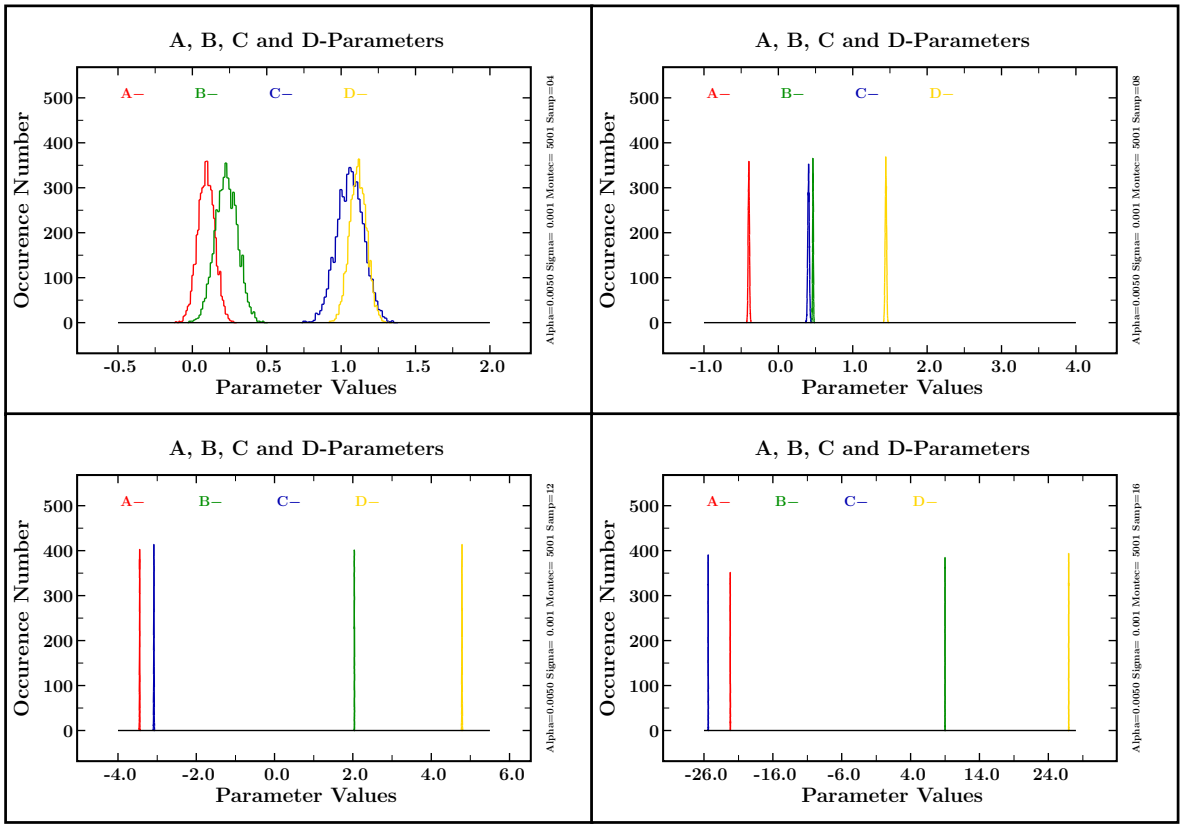

Figure 3.40: *Probability distributions of parameters for $\alpha = 0.005$ and $\sigma = 0.001$ ('precise data' variant) for four different sampling options: top-left – 4 points, top-right – 8 points, bottom-left – 12 points and bottom-right – 16 points.*

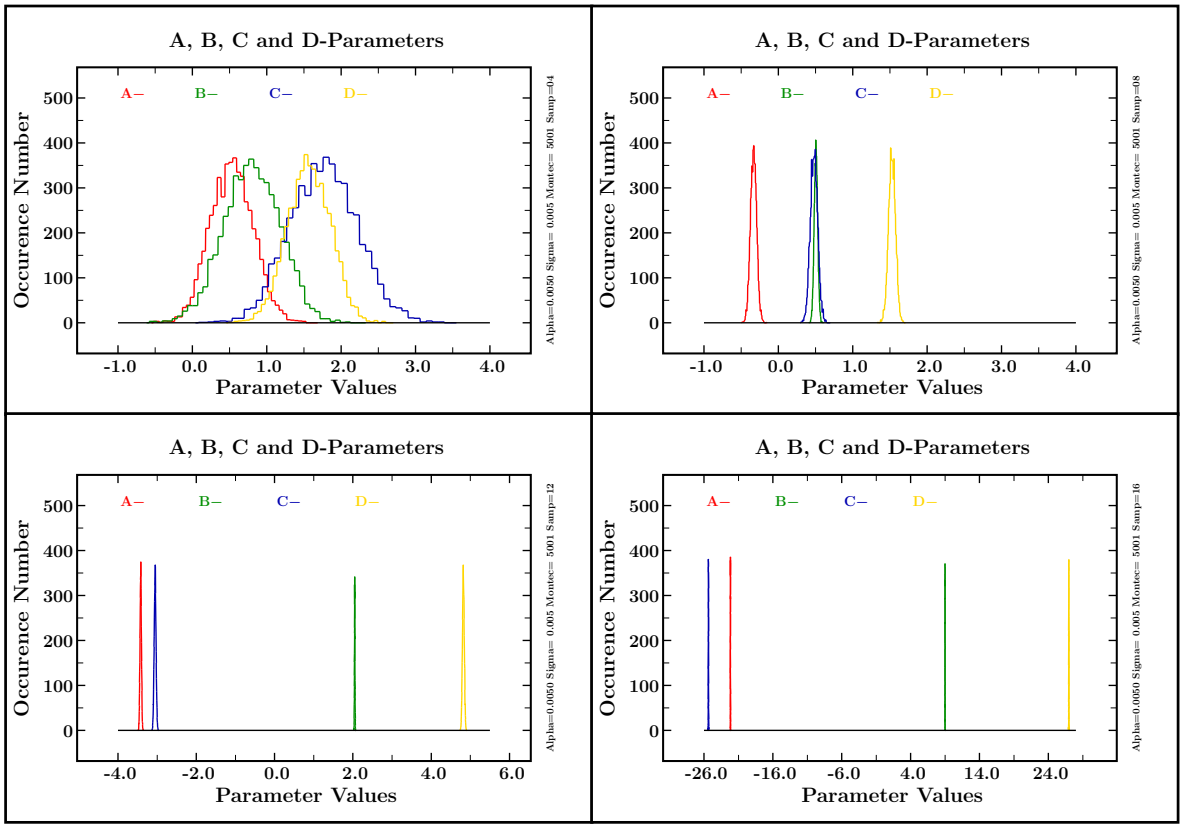

Figure 3.41: *Probability distributions of parameters for $\sigma = 0.005$ ('less precise data').*

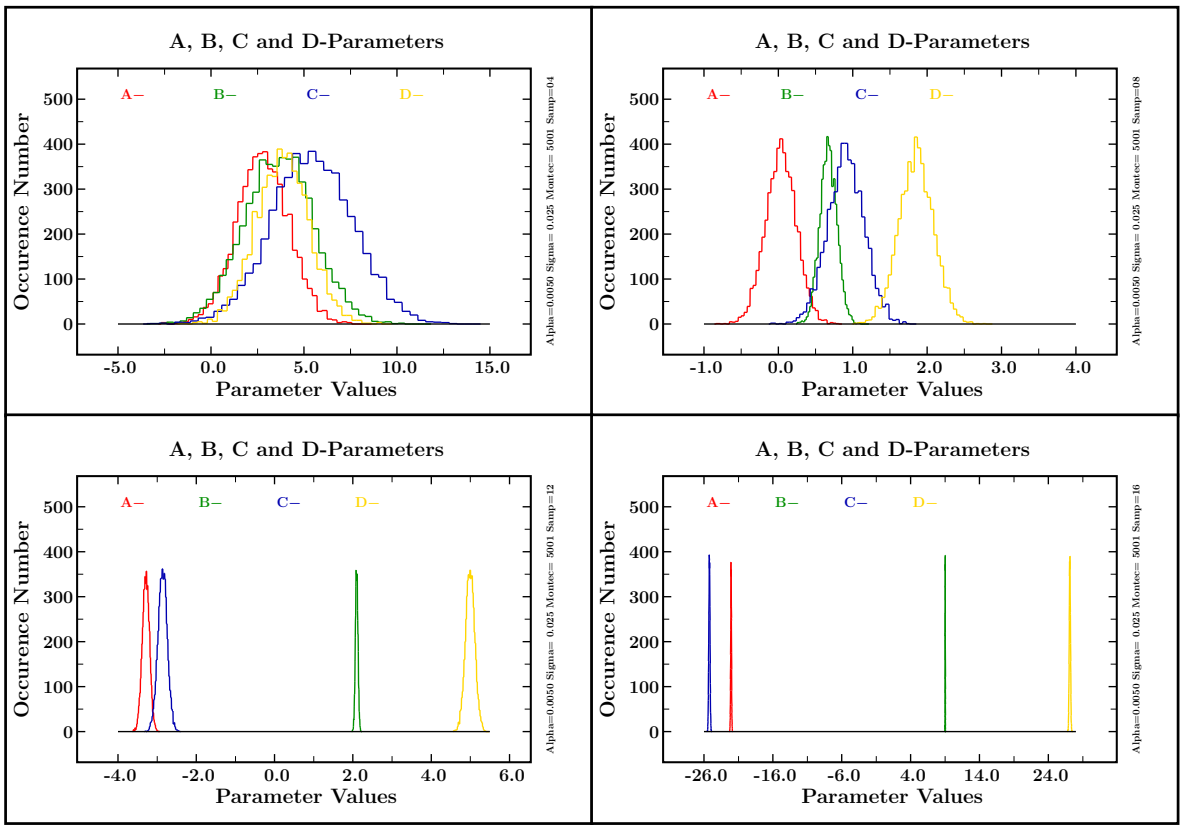

Figure 3.42: *Probability distributions of parameters for $\sigma = 0.025$ ('imprecise data').*

3.5 Test of Predictive Power

The most important aspect of a mathematical model or a theory is its applicability and usefulness. If we create a model and fit it to given data that are available to us, we hope it can be applied in other cases as well. Usually, the obtained set of parameters can be used to calculate the data points (i.e. the predictions of the model) within the same range, that the parameters were obtained from. However, an ultimate test of a model is its ability to reliably extrapolate to regions lying outside the fitting domain. This is not an easy task, especially in nuclear physics, where available experimental data are rather scarce but also the underlying physics is quite complex.

As a next test, we calculate two types of predictions. The first one, for points lying inside the fitting range, will be reffered to as 'intraneous predictions', whereas the second one, where we try to extrapolate beyond the fitting domain, will be called 'extraneous'.

3.5.1 Intraneous Predictions

The following six figures 3.43-3.48 present the histograms for intraneous predictions. They are organised in such a way, that, similarly as in previous sections, the columns correspond to given sampling option and the plots in these columns are arranged by decreasing precision of the data. For simplicity and to increase the clarity, the points are shifted, so the maximum is at zero. The number of intraneous points calculated for each sampling option is the same, as the number of points in the sample. The histograms are labeled by the number of the point with a prefix 'Int'.

The first two figures contain plots for the exact case with $\alpha = 0$, the next two are for $\alpha = 0.001$, whereas the last two present predictions for $\alpha = 0.005$. All three types of calculations show similar behaviour. Large experimental uncertainties result in broad distributions. However, increasing the size of the sample influences the resulting histograms only to a small extent when it comes to their size.

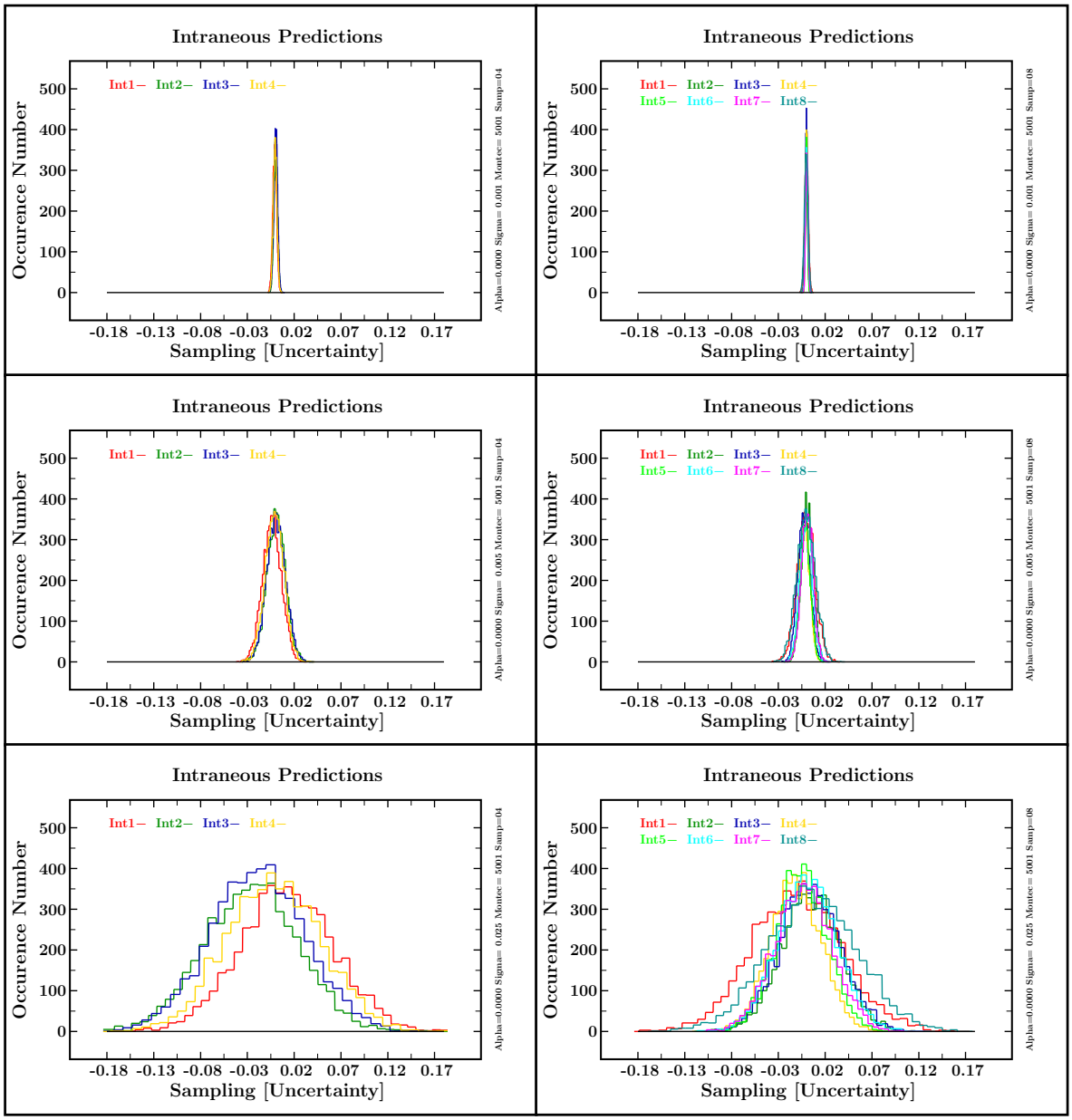

Figure 3.43: *Probability distributions of intraneous predictions for $\alpha = 0$ ('exact model' version) with different types of input data for two sampling options: 4 points – left column, 8 points – right column.*

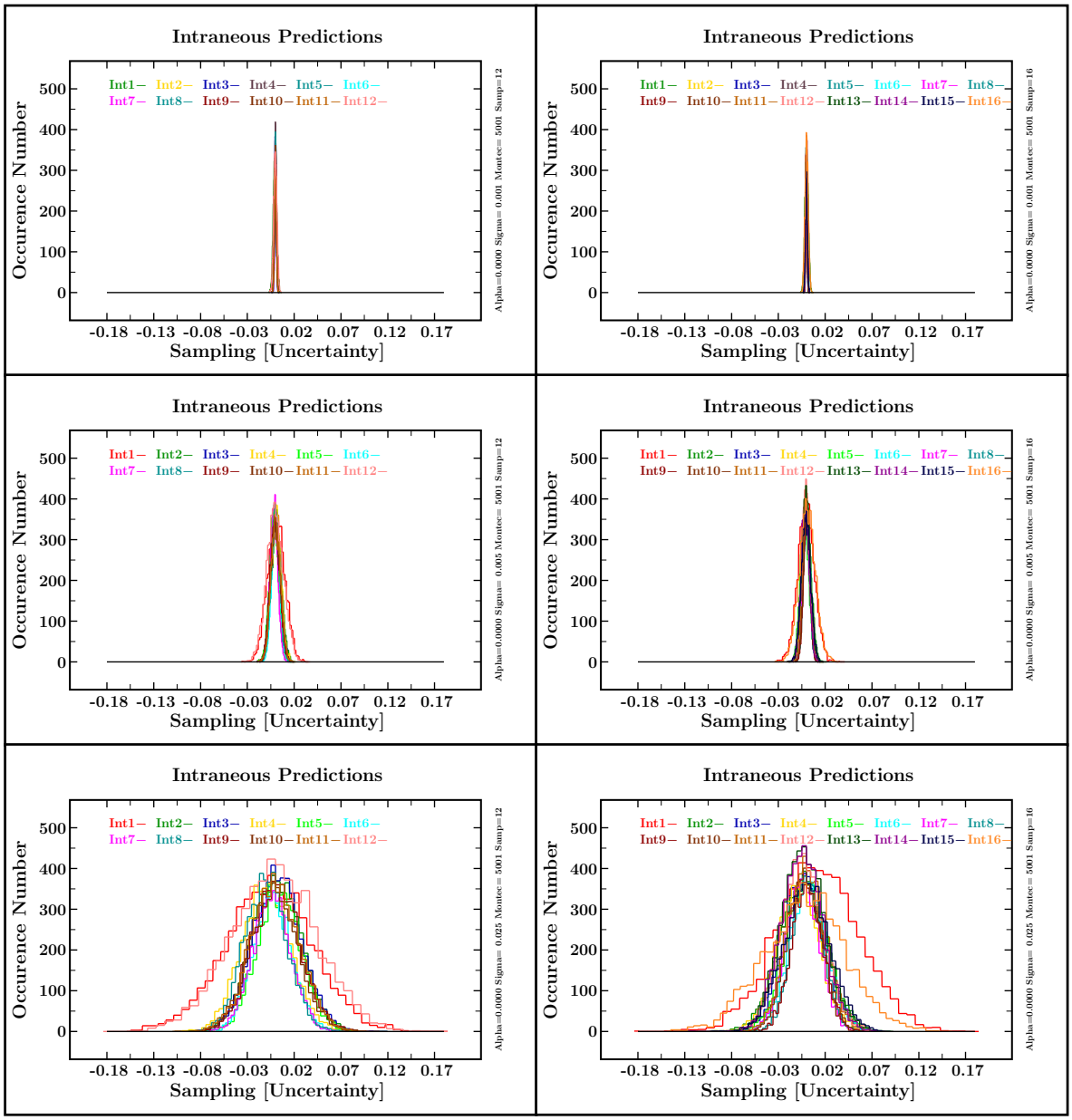

Figure 3.44: *Probability distributions of intraneous predictions for $\alpha = 0$ ('exact model' version) with different types of input data for two sampling options: 12 points – left column, 16 points – right column.*

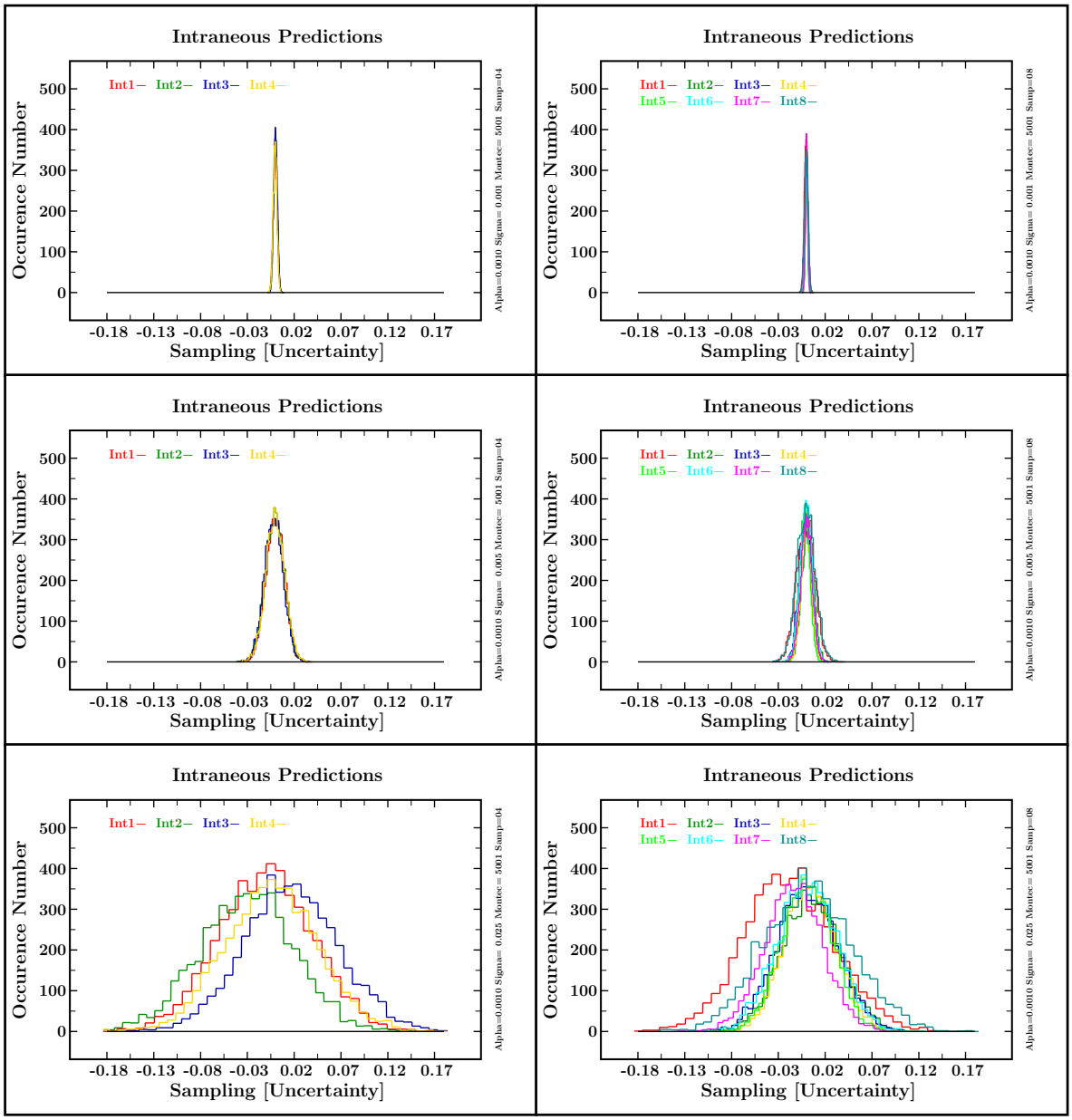

Figure 3.45: *Probability distributions of intraneous predictions for $\alpha = 0.001$ with different types of input data for two sampling options: 4 points – left column, 8 points – right column.*

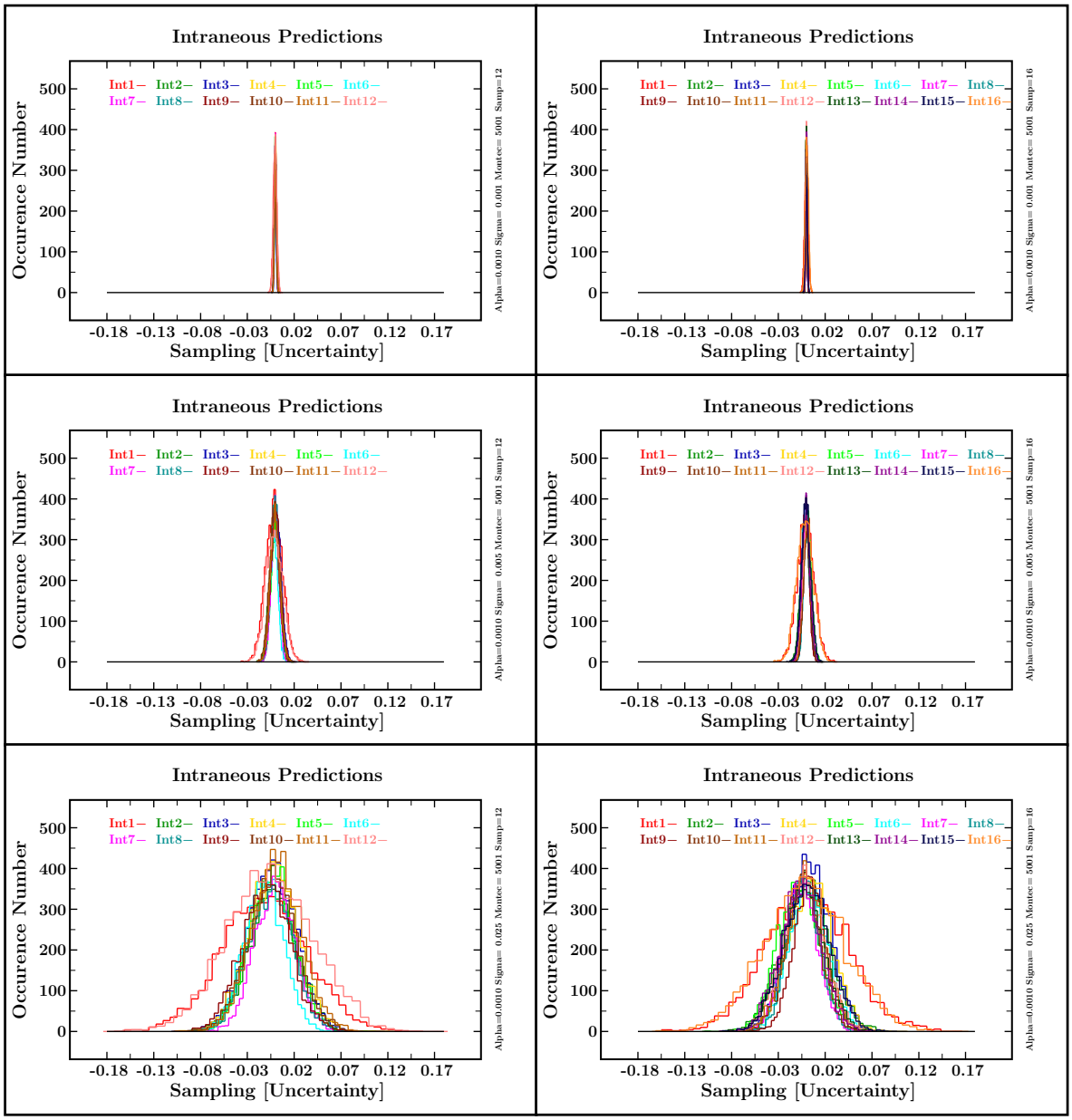

Figure 3.46: *Probability distributions of intraneous predictions for $\alpha = 0.001$ with different types of input data for two sampling options: 12 points – left column, 16 points – right column.*

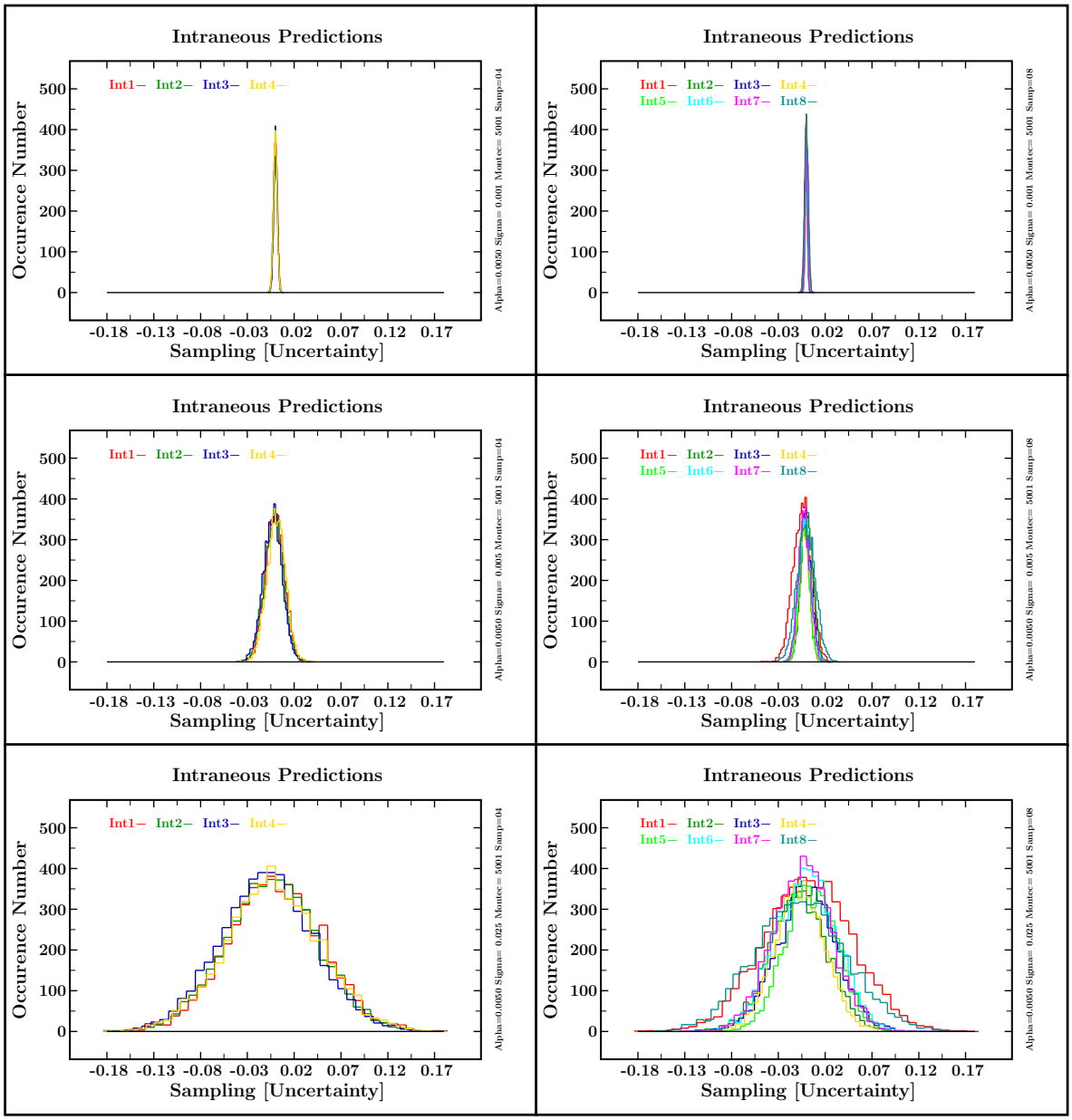

Figure 3.47: *Probability distributions of intraneous predictions for $\alpha = 0.005$ with different types of input data for two sampling options: 4 points – left column, 8 points – right column.*

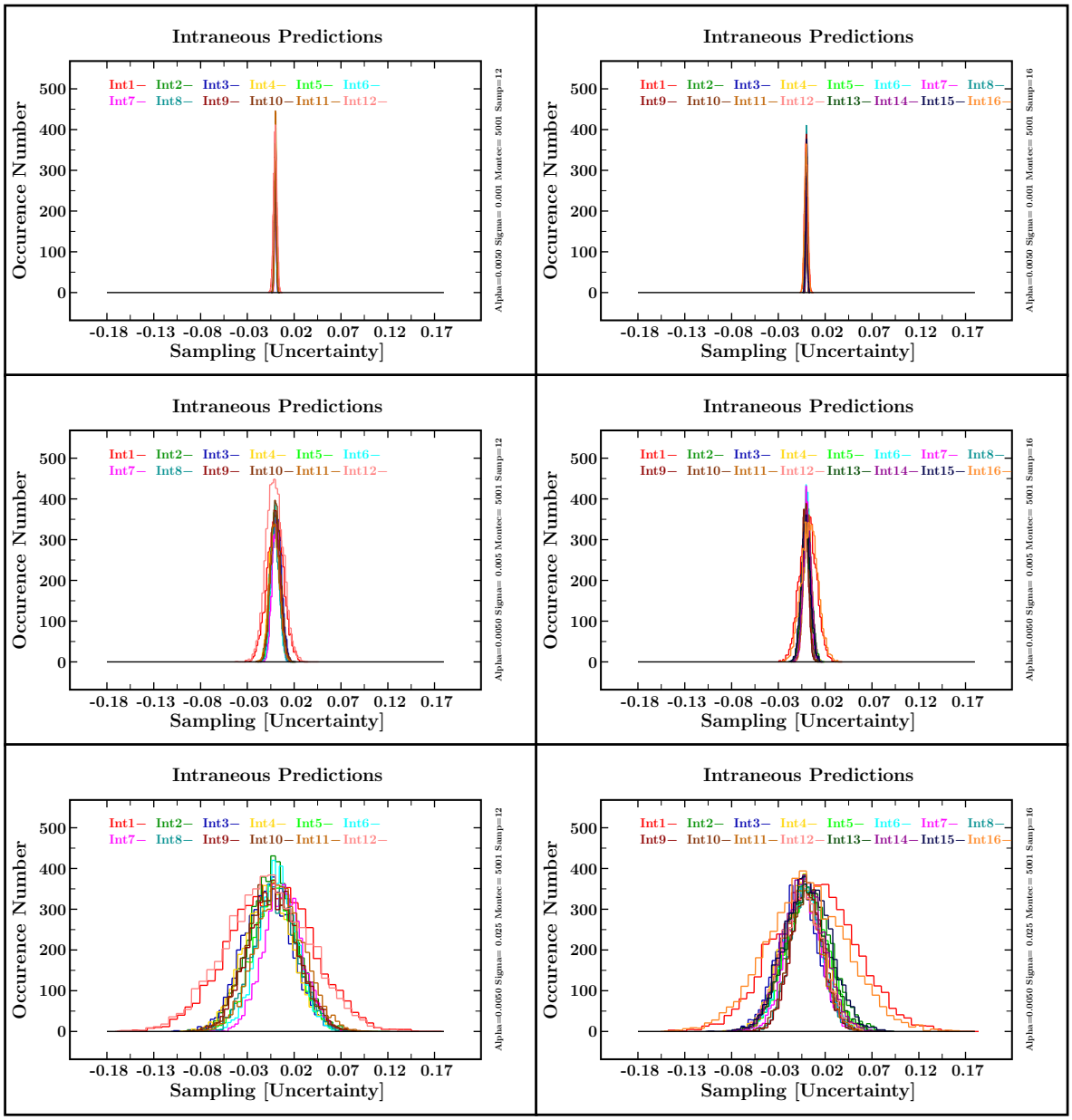

Figure 3.48: *Probability distributions of intraneous predictions for $\alpha = 0.005$ with different types of input data for two sampling options: 12 points – left column, 16 points – right column.*

As we explained at the beginning, the plots contain only a part of the information. The second part is presented in the following set of tables 3.1-3.12, which show the comparison between the 'real' (unperturbed) data and the maxima of the histograms. Each table represents one sampling option for three different data variants (as usual: 'precise' with $\sigma = 0.001$, 'moderately precise' with $\sigma = 0.050$ and 'imprecise' data with $\sigma = 0.025$). Examining the tables 3.1-3.4, we immediately see that, as it was expected, the precision of experimental input has a significant influence on the predictions of the 'exact' model. Increasing the size of the sample from 4 to 8 improves the result, however further addition of point has almost no impact on the performance of the model.

Table 3.1: *Intraneous predictions for 4-point sampling option and $\alpha = 0$ for different data precision variants.*

$\alpha = 0$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	1.00	0.00	1.02	0.02	1.08	0.08
1.65	1.65	0.00	1.67	0.02	1.77	0.12
2.72	2.72	0.00	2.74	0.02	2.86	0.14
4.48	4.49	0.01	4.50	0.02	4.57	0.09

Table 3.2: *Intraneous predictions for and 8-point sampling option $\alpha = 0$ for different data precision variants.*

$\alpha = 0$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	1.00	0.00	1.02	0.02	1.10	0.10
1.65	1.65	0.00	1.66	0.01	1.70	0.05
2.72	2.72	0.00	2.73	0.01	2.77	0.05
4.48	4.48	0.00	4.49	0.01	4.55	0.07
7.39	7.39	0.00	7.40	0.01	7.45	0.06
12.18	12.18	0.00	12.20	0.02	12.24	0.06
20.09	20.09	0.00	20.10	0.01	20.14	0.05
33.12	33.12	0.00	33.13	0.01	33.20	0.08

The situation is slightly different when we consider the 'inexact models' (i.e. with nonzero α parameter), cf. tables 3.5-3.12. The results for 4 and 8 sampling points are rather similar. However, increasing the size of the sample, we immediately see the worsening of the fit. Additionally, the input uncertainties seem to have a very low impact on the result – the errors are similar for all data variants. The more we deviate from the exact model, the more the predictions differ from the expected values.

Table 3.3: *Intraneous predictions for 12-point sampling option $\alpha = 0$ for different data precision variants.*

$\alpha = 0$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	1.00	0.00	1.02	0.02	1.08	0.08
1.65	1.65	0.00	1.66	0.01	1.70	0.05
2.72	2.72	0.00	2.73	0.01	2.76	0.04
4.48	4.48	0.00	4.49	0.01	4.54	0.06
7.39	7.39	0.00	7.40	0.01	7.43	0.04
12.18	12.18	0.00	12.19	0.01	12.22	0.04
20.09	20.09	0.00	20.10	0.01	20.12	0.03
33.12	33.12	0.00	33.13	0.01	33.17	0.05
54.60	54.60	0.00	54.61	0.01	54.64	0.04
90.02	90.02	0.00	90.03	0.01	90.06	0.04
148.41	148.42	0.01	148.42	0.01	148.47	0.06
244.69	244.70	0.01	244.71	0.02	244.79	0.10

Moreover, in major part of the cases, the real solution lies far beyond the obtained distributions. Only the smallest samples (4 and 8 points) with rather high-precision data are accurate enough to reproduce the data.

Table 3.4: *Intraneous predictions for 16-point sampling option $\alpha = 0$ for different data precision variants.*

$\alpha = 0$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	1.00	0.00	1.02	0.02	1.09	0.09
1.65	1.65	0.00	1.66	0.01	1.70	0.05
2.72	2.72	0.00	2.73	0.01	2.77	0.05
4.48	4.48	0.00	4.49	0.01	4.54	0.06
7.39	7.39	0.00	7.40	0.01	7.43	0.04
12.18	12.18	0.00	12.19	0.01	12.22	0.04
20.09	20.09	0.00	20.09	0.00	20.12	0.03
33.12	33.12	0.00	33.12	0.00	33.15	0.03
54.60	54.60	0.00	54.60	0.00	54.63	0.03
90.02	90.02	0.00	90.02	0.00	90.06	0.04
148.41	148.41	0.00	148.42	0.01	148.46	0.05
244.69	244.69	0.00	244.70	0.01	244.75	0.06
403.43	403.43	0.00	403.44	0.01	403.50	0.07
665.14	665.14	0.00	665.15	0.01	665.21	0.07
1096.63	1096.63	0.00	1096.64	0.01	1096.68	0.05
1808.04	1808.05	0.01	1808.06	0.02	1808.12	0.08

Table 3.5: *Intraneous predictions for 4-point sampling option and $\alpha = 0.001$ for different data precision variants.*

$\alpha = 0.001$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	1.00	0.00	1.02	0.02	1.12	0.12
1.65	1.65	0.00	1.67	0.02	1.77	0.12
2.72	2.72	0.00	2.74	0.02	2.79	0.07
4.47	4.48	0.01	4.49	0.02	4.57	0.10

Table 3.6: *Intraneous predictions for and 8-point sampling option $\alpha = 0.001$ for different data precision variants.*

$\alpha = 0.001$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	1.01	0.01	1.02	0.02	1.12	0.12
1.65	1.65	0.00	1.65	0.00	1.69	0.04
2.72	2.72	0.00	2.73	0.01	2.77	0.05
4.47	4.48	0.01	4.49	0.02	4.52	0.05
7.36	7.37	0.01	7.38	0.02	7.42	0.06
12.11	12.11	0.00	12.12	0.01	12.17	0.06
19.91	19.90	-0.01	19.91	0.00	19.98	0.07
32.71	32.72	0.01	32.74	0.03	32.80	0.09

Table 3.7: *Intraneous predictions for 12-point sampling option $\alpha = 0.001$ for different data precision variants.*

$\alpha = 0.001$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	1.08	0.08	1.10	0.10	1.17	0.17
1.65	1.57	-0.08	1.58	-0.07	1.62	0.03
2.72	2.62	-0.10	2.62	0.10	2.67	0.05
4.47	4.42	-0.05	4.43	-0.04	4.47	0.00
7.36	7.38	0.02	7.39	0.03	7.43	0.07
12.11	12.20	0.09	12.20	0.09	12.25	0.14
19.91	20.03	0.12	20.03	0.12	20.07	0.16
32.71	32.81	0.10	32.82	0.11	32.85	0.14
53.74	53.74	0.00	53.75	0.01	53.80	0.06
88.23	88.11	-0.12	88.12	-0.11	88.16	0.07
144.79	144.62	-0.17	144.63	-0.16	144.67	0.12
237.51	237.64	0.13	237.65	0.14	237.74	0.23

Table 3.8: *Intraneous predictions for 16-point sampling option $\alpha = 0.001$ for different data precision variants.*

$\alpha = 0.001$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
1.00	2.07	1.07	2.09	1.09	2.15	1.15
1.65	1.02	-0.63	1.03	-0.62	1.07	-0.58
2.72	1.51	-1.21	1.52	-1.20	1.55	-1.18
4.47	3.34	-1.13	3.35	-1.12	3.38	-1.09
7.36	6.68	-0.68	6.68	-0.68	6.73	-0.63
12.11	12.04	-0.07	12.05	-0.06	12.07	-0.04
19.91	20.48	0.57	20.49	0.58	20.52	0.61
32.71	33.85	1.14	33.85	1.14	33.88	1.17
53.74	55.22	1.48	55.23	1.49	55.25	1.51
88.23	89.74	1.51	89.75	1.52	89.78	1.55
144.79	145.91	1.12	145.92	1.13	145.95	1.16
237.51	237.74	0.23	237.75	0.24	237.79	0.28
389.41	388.36	-0.95	388.37	-0.94	388.40	-0.91
638.18	635.89	-2.29	635.89	-2.29	635.94	-2.24
1045.41	1043.19	-2.22	1043.20	-2.21	1043.24	-2.17
1711.76	1713.92	2.16	1713.94	2.18	1714.00	2.24

Table 3.9: *Intraneous predictions for 4-point sampling option and $\alpha = 0.005$ for different data precision variants.*

$\alpha = 0.005$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	Shift	$y(x)$	Shift	$y(x)$	Shift
1.00	1.00	0.00	1.02	0.02	1.10	0.10
1.65	1.65	0.00	1.67	0.02	1.75	0.10
2.70	2.71	0.01	2.73	0.03	2.81	0.11
4.43	4.44	0.01	4.45	0.02	4.53	0.10

Table 3.10: *Intraneous predictions for and 8-point sampling option $\alpha = 0.005$ for different data precision variants.*

$\alpha = 0.005$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	Shift	$y(x)$	Shift	$y(x)$	Shift
1.00	1.02	0.02	1.04	0.04	1.10	0.10
1.65	1.63	-0.02	1.63	-0.02	1.69	0.04
2.70	2.70	0.00	2.71	0.01	2.76	0.06
4.43	4.45	0.02	4.46	0.03	4.51	0.08
7.24	7.27	0.03	7.28	0.04	7.31	0.07
11.81	11.82	0.01	11.82	0.01	11.88	0.07
19.22	19.19	-0.03	19.20	-0.02	19.26	0.04
31.20	31.22	0.02	31.24	0.04	31.30	0.10

Table 3.11: *Intraneous predictions for 12-point sampling option $\alpha = 0.005$ for different data precision variants.*

$\alpha = 0.005$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	Shift	$y(x)$	Shift	$y(x)$	Shift
1.00	1.32	0.32	1.33	0.33	1.40	0.40
1.65	1.33	-0.32	1.34	-0.31	1.38	-0.27
2.70	2.31	-0.39	2.31	-0.39	2.36	-0.34
4.43	4.23	-0.20	4.24	-0.19	4.29	-0.14
7.24	7.34	0.10	7.35	0.11	7.38	0.14
11.81	12.17	0.36	12.18	0.37	12.21	0.40
19.22	19.69	0.47	19.70	0.48	19.72	0.50
31.20	31.56	0.36	31.57	0.37	31.60	0.40
50.55	50.55	0.00	50.56	0.01	50.60	0.05
81.74	81.26	-0.48	81.27	-0.47	81.31	-0.43
131.92	131.25	-0.67	131.26	-0.66	131.29	-0.63
212.54	213.04	0.50	213.06	0.52	213.12	0.58

Table 3.12: *Intraneous predictions for 16-point sampling option $\alpha = 0.005$ for different data precision variants.*

$\alpha = 0.005$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
$y_\alpha^{\exp}(x)$	$y(x)$	Shift	$y(x)$	Shift	$y(x)$	Shift
1.00	4.73	3.73	4.74	3.74	4.80	3.80
1.65	-0.58	-2.23	-0.57	-2.22	-0.54	-2.19
2.70	-1.52	-4.22	-1.52	-4.22	-1.47	-4.17
4.43	0.52	-3.91	0.53	-3.90	0.57	-3.86
7.24	4.92	-2.32	4.93	-2.31	4.96	-2.29
11.81	11.66	-0.15	11.67	-0.14	11.70	-0.11
19.22	21.31	2.09	21.32	2.10	21.34	2.12
31.20	35.18	3.98	35.19	3.99	35.22	4.02
50.55	55.68	5.07	55.69	5.08	55.71	5.10
81.74	86.88	5.14	86.89	5.15	86.91	5.17
131.92	135.61	3.69	135.62	3.70	135.65	3.73
212.54	213.15	0.61	213.16	0.62	213.19	0.66
341.89	338.16	-3.73	338.17	-3.72	338.20	-3.69
549.14	541.39	-7.75	541.40	-7.74	541.44	-7.70
880.83	873.58	-7.25	873.59	-7.24	873.63	-7.20
1411.16	1418.37	7.21	1418.38	7.22	1418.45	7.29

3.5.2 Extraneous Predictions

The following figures 3.49-3.54 present the second type of predictions, i.e. an attempt to estimate the positions of points lying beyond the fitting range. For each type of calculations, we present four predictions for $x_{17} = 8.0,\ x_{18} = 8.5,\ x_{19} = 9.0; x_{20} = 9.5$. The plots are organised in such a way, that each figure presents the histograms for two out of four sampling options for a given version of the model. Each column presents all three data variants: the 'precise' data at the top, the 'moderately precise' in the middle and the 'lowest precision' at the bottom. All the histograms are shifted and centered around zero and are labeled by the number of the point with a prefix 'Ext'.

As we expected, the precision of the data has a large influence on the width of the histograms. The number of points used for the fit also plays a significant role and actually its impact is larger than the differences resulting from the levels of data precision (the widths of the histograms change by orders of magnitude from one sampling option to another). Comparing the ensembles of plots for different versions of the model, we can observe no significant difference between them. All the corresponding plots are of the same scale.

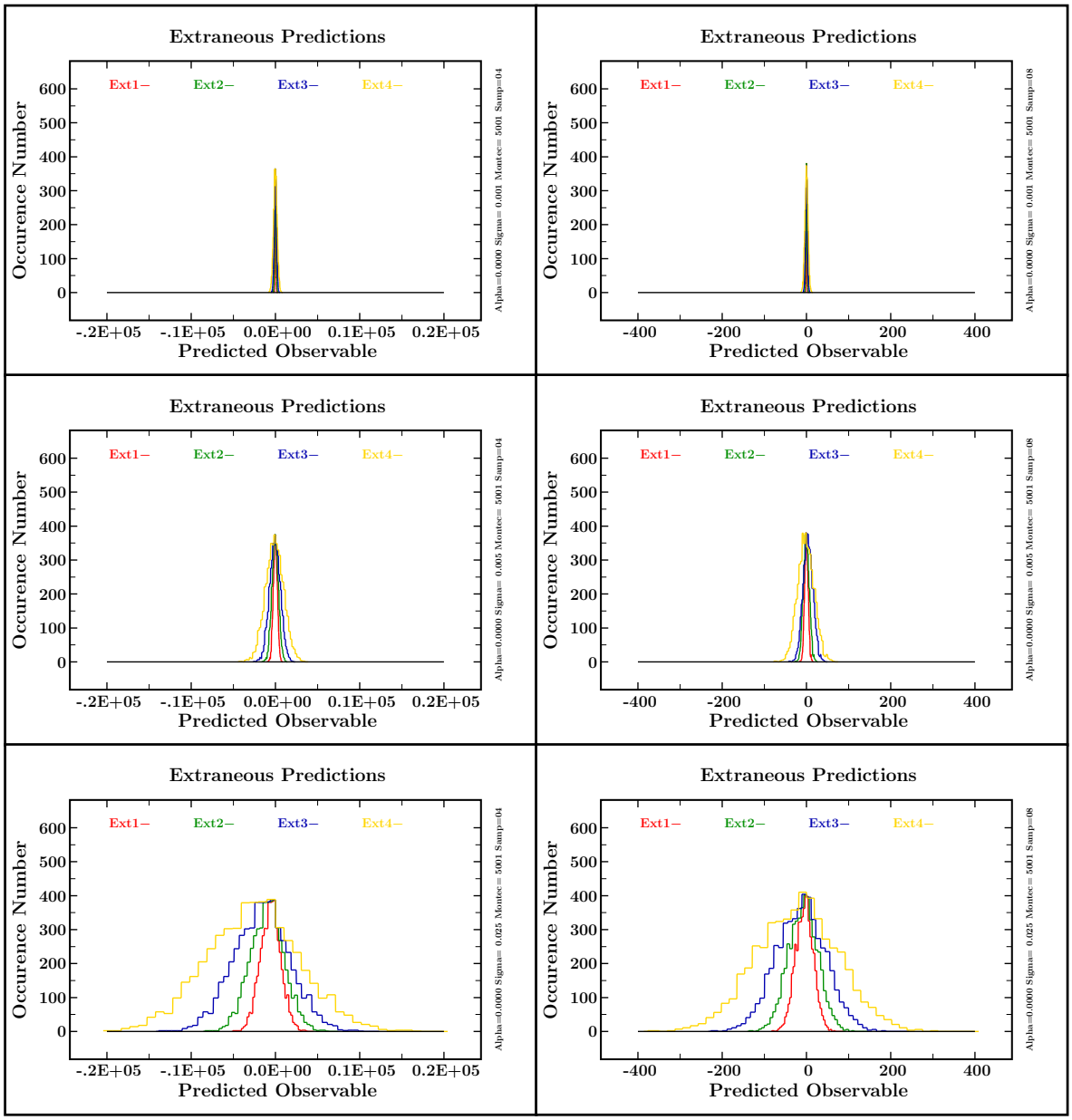

Figure 3.49: *Probability distributions of extraneous predictions for $\alpha = 0$ ('exact model' version) with different types of input data for two sampling options: 4 points – left column, 8 points – right column.*

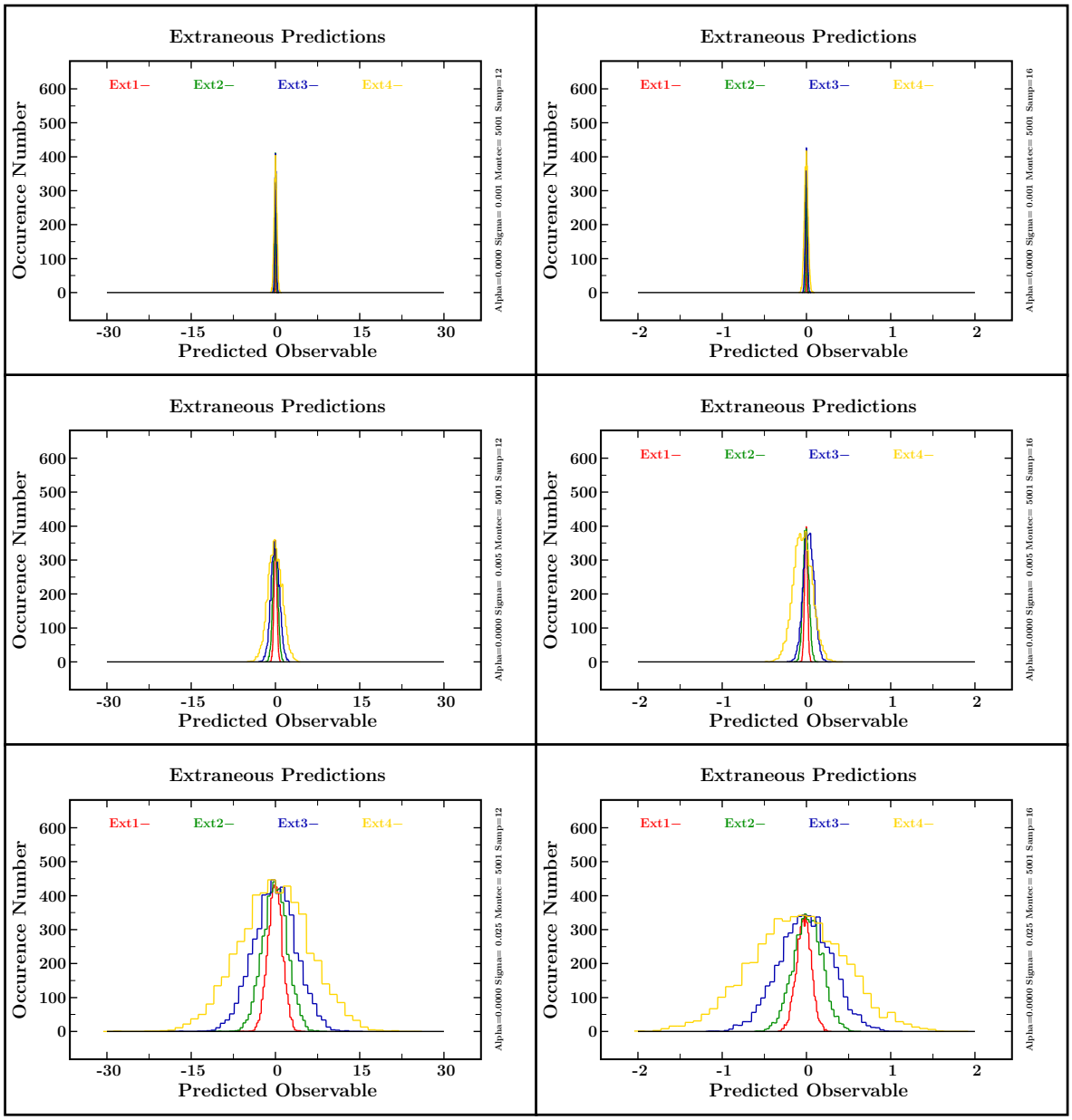

Figure 3.50: *Probability distributions of extraneous predictions for $\alpha = 0$ ('exact model' version) with different types of input data for two sampling options: 12 points – left column, 16 points – right column.*

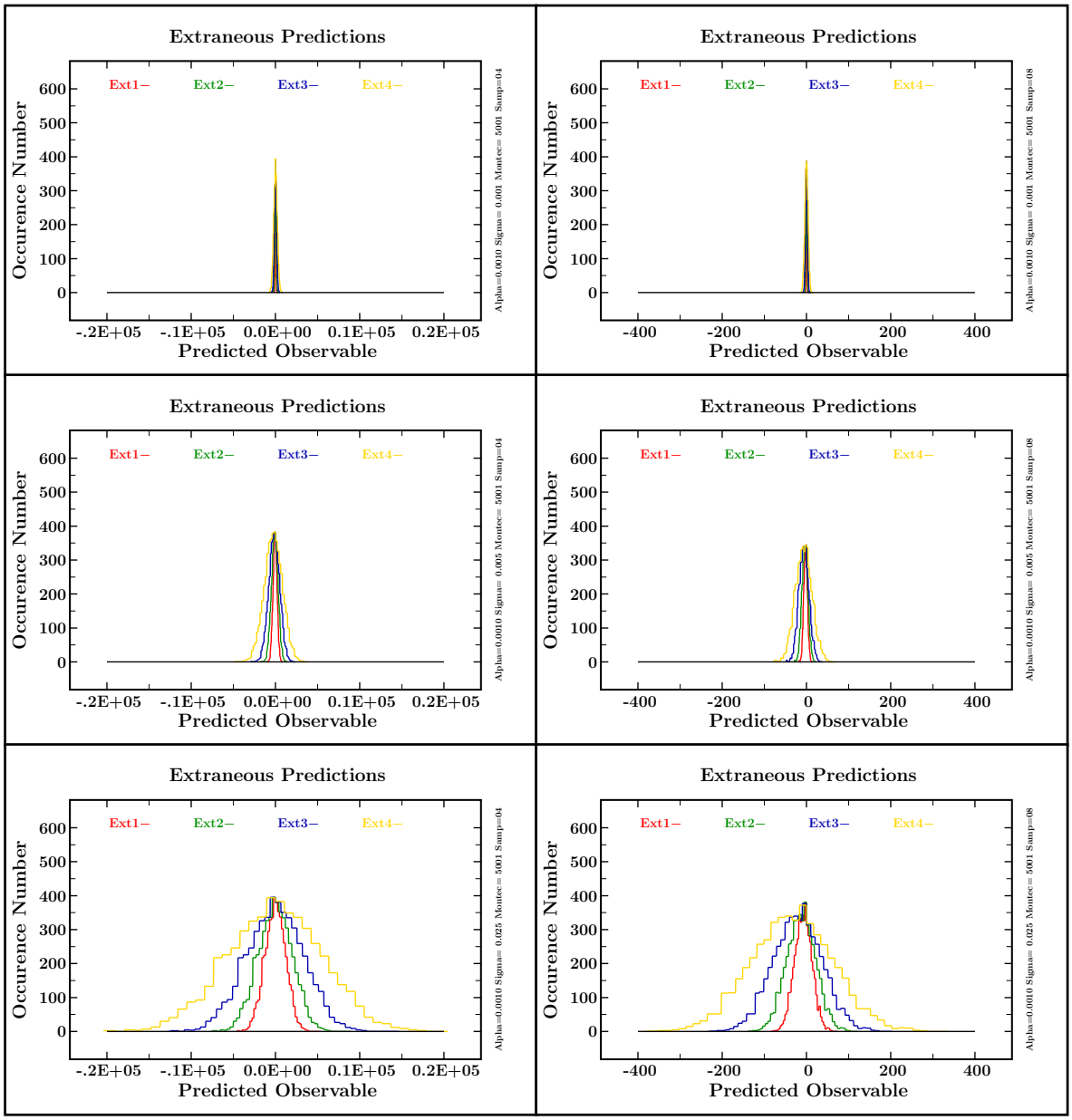

Figure 3.51: *Probability distributions of extraneous predictions for $\alpha = 0.001$ with different types of input data for two sampling options: 4 points – left column, 8 points – right column.*

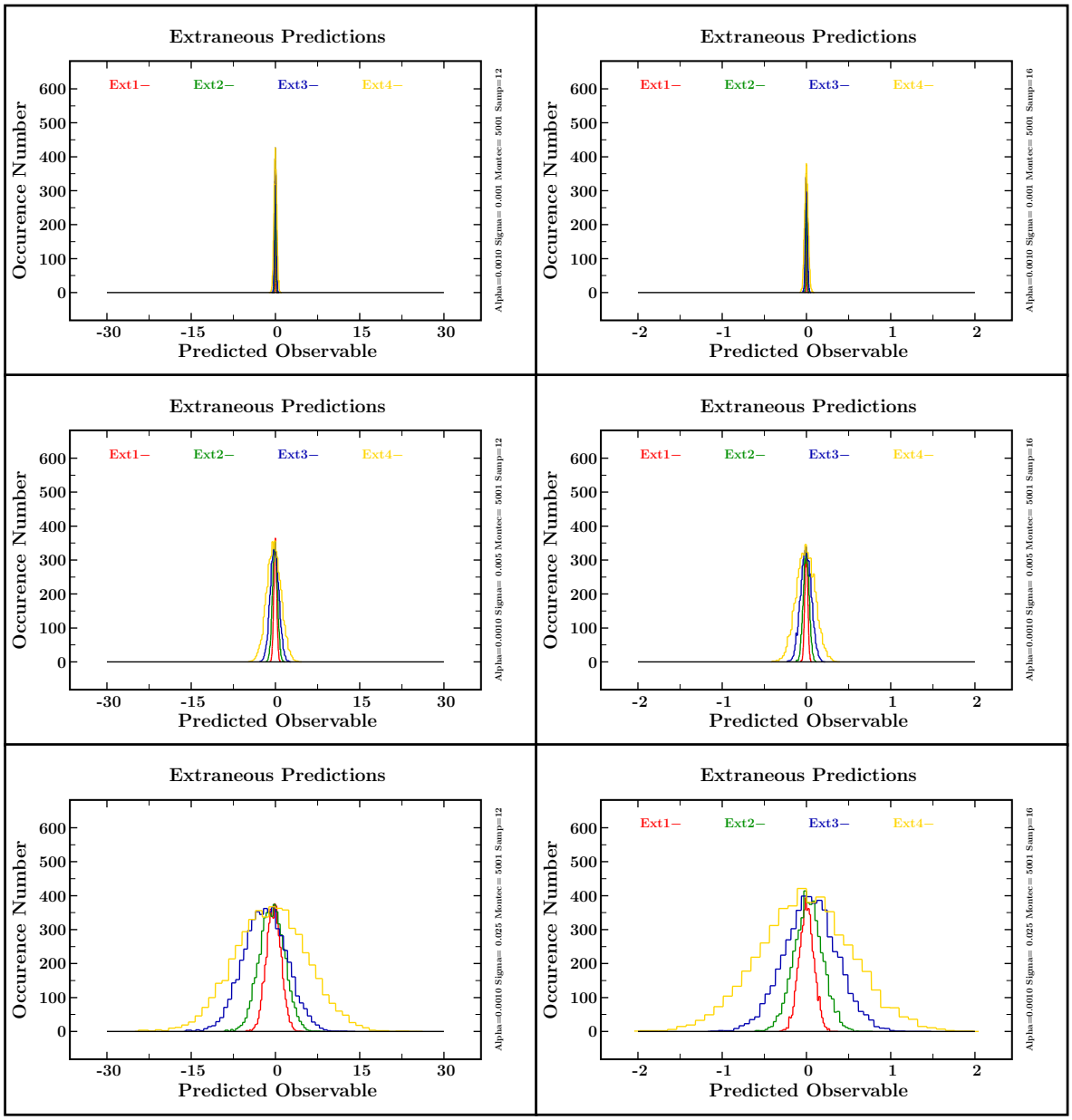

Figure 3.52: *Probability distributions of extraneous predictions for $\alpha = 0.001$ with different types of input data for two sampling options: 12 points – left column, 16 points – right column.*

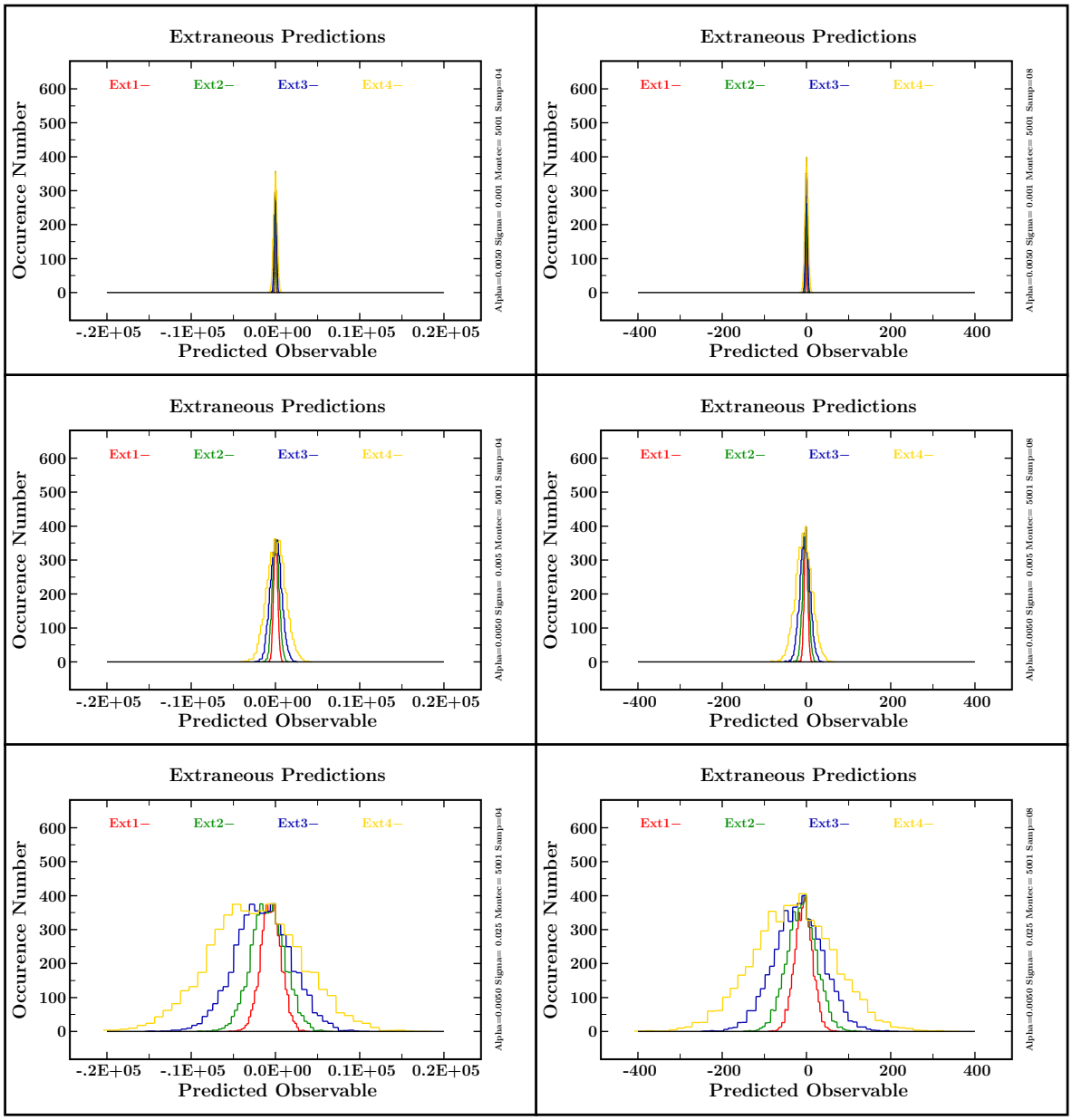

Figure 3.53: *Probability distributions of extraneous predictions for $\alpha = 0.005$ with different types of input data for two sampling options: 4 points – left column, 8 points – right column.*

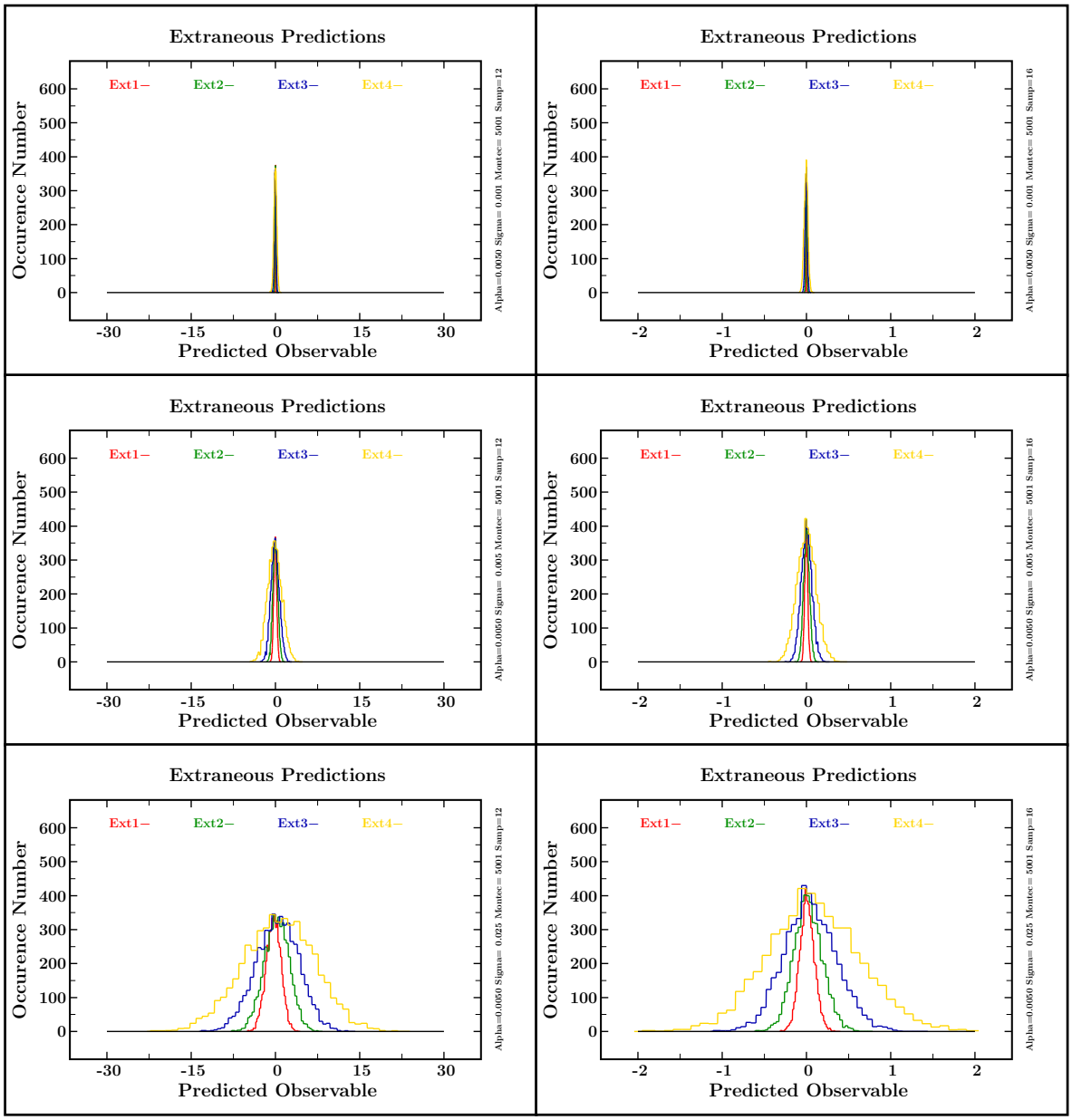

Figure 3.54: *Probability distributions of extraneous predictions for $\alpha = 0.005$ with different types of input data for two sampling options: 12 points – left column, 16 points – right column.*

The following tables 3.13-3.15 present the actual positions of the maxima of the histograms as compared to 'real' values. Each table contains all the results for a given version of the model. The second column contains the 'experimental' values. The next three pairs of columns give the maximum of the histogram and the difference between this value and the exact position of a given point for different data precision.

Let us briefly examine every table. Increasing the inaccuracy of the data input in the case of exact model (table 3.13) causes a drastic deterioration of the accuracy of predictions (notice the 100% error in the case of 4-point sampling fit to 'imprecise' data). Increasing the size of the sample improves the predictions; doubling the size of the sample gives two orders of magnitude improvement of the predictions.

For the inexact model options (tables 3.14 and 3.15) we obtain much worse performance. However, there is one significant difference. For sufficiently large sample size the model seems to be insensitive to data precision.

Table 3.13: *Extraneous predictions for four sampling options and $\alpha = 0$ for different data precision variants.*

Sample	$\alpha = 0$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
	$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
m=4	2980.96	3073.14	92.18	3488.33	507.37	5857.38	2876.42
	4914.77	5067.31	152.54	5754.57	839.80	9675.65	4760.79
	8103.08	8355.18	251.10	9491.23	1388.15	15972.35	7869.27
	13359.73	13776.01	416.28	15652.18	2292.45	26355.26	12995.53
m=8	2980.96	2982.52	1.56	2989.63	8.67	3025.77	44.81
	4914.77	4917.37	2.60	4929.23	14.46	4989.49	74.72
	8103.08	8107.40	4.32	8122.79	19.71	8227.18	124.10
	13359.73	13366.88	7.15	13399.52	40.21	13565.31	205.58
m=12	2980.96	2981.06	0.10	2981.48	0.52	2984.36	2.40
	4914.77	4914.94	0.17	4915.68	0.91	4920.60	5.83
	8103.08	8103.37	0.29	8104.63	1.55	8112.95	9.87
	13359.73	13360.21	0.48	13362.32	3.41	13376.27	16.54
m=16	2980.96	2980.97	0.01	2981.00	0.04	2981.14	1.16
	4914.77	4914.78	0.01	4914.85	0.08	4915.07	0.30
	8103.08	8103.11	0.03	8103.19	0.11	8103.69	0.61
	13359.73	13359.78	0.05	13360.00	0.27	13360.88	1.15

Table 3.14: *Extraneous predictions for four sampling options and $\alpha = 0.001$ for different data precision variants.*

Sample	$\alpha = 0.001$	$\sigma = 0.001$		$\sigma = 0.005$		$\sigma = 0.025$	
	$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
m=4	2801.65	3041.86	240.21	3526.91	725.26	5288.75	2487.10
	4583.60	5015.68	432.08	5818.47	839.80	8735.14	4760.79
	7495.91	8270.02	774.11	9596.89	2100.98	14418.47	6922.56
	12253.82	13635.54	1381.72	15826.73	3572.91	23789.91	11536.09
m=8	2801.65	2924.34	122.69	2932.50	130.85	2968.72	167.07
	4583.60	4821.04	237.44	4834.67	251.07	4894.98	311.38
	7495.91	7948.15	452.24	7970.79	474.88	8070.87	574.96
	12253.82	13103.85	850.03	13137.97	884.15	13307.07	1053.25
m=12	2801.65	2875.57	73.92	2875.97	74.32	2878.30	76.65
	4583.60	4739.27	155.67	4739.97	156.37	4743.99	160.39
	7495.91	7811.83	315.92	7813.02	317.11	7820.57	324.66
	12253.82	12877.47	623.65	12879.46	625.64	12889.78	635.96
m=16	2801.65	2818.96	17.31	2818.99	17.34	2819.12	17.47
	4583.60	4640.06	56.46	4640.12	56.52	4640.35	56.75
	7495.91	7641.75	145.84	7641.84	145.93	7642.29	146.38
	12253.82	12589.88	336.06	12590.05	336.23	12590.95	337.13

Table 3.15: *Extraneous predictions for four sampling options and $\alpha = 0.005$ for different data precision variants.*

Sample	$\alpha = 0.001$	$\sigma = 0.005$		$\sigma = 0.005$		$\sigma = 0.025$	
	$y_\alpha^{\exp}(x)$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$	$y(x)$	$y - y_\alpha^{\exp}$
m=4	2258.30	2942.02	683.72	3306.87	1048.57	5962.17	3703.87
	3610.48	4850.72	1240.24	5454.60	1844.12	9850.30	6239.82
	5767.32	7997.64	2230.32	8995.81	3228.49	16262.33	10495.01
	9205.67	13186.06	3980.39	14834.46	5628.79	26835.56	17629.89
m=8	2258.30	2707.47	449.17	2715.69	457.39	2754.59	496.29
	3610.48	4462.02	851.54	4475.75	865.27	4540.56	930.08
	5767.32	7354.65	1587.33	7377.48	1610.16	7485.07	1717.75
	9205.67	12123.64	2917.97	12161.49	2955.82	12339.66	3133.99
m=12	2258.30	2513.93	255.63	2514.35	256.06	2516.54	258.24
	3610.48	4137.49	527.01	4138.13	527.65	4141.20	530.72
	5767.32	6813.65	1046.33	6814.71	1047.39	6819.92	1052.60
	9205.67	11225.12	2019.45	11226.99	2021.32	11235.76	2030.09
m=16	2258.30	2313.67	55.37	2313.71	55.41	2313.82	55.52
	3610.48	3786.87	176.39	3786.93	176.45	3787.16	176.68
	5767.32	6212.85	445.53	6212.96	445.64	6213.37	446.05
	9205.67	10209.71	1004.04	10209.89	1004.22	10210.62	1004.95

3.6 Conclusions

In this chapter we have presented the application of the concepts of Predictive Power to a mathematical model. For this purpose an exactly soluble model was created, which was then modified with the help of a small parameter α, allowing us to test the performance and limitations of inexact modelling. Using the Monte-Carlo techniques, we have examined various properties of the data modelling.

It was shown, that correlations of parameters of a model are a crucial aspect of the fitting process. The underlying parametric dependencies strongly affect the outcome.

Our calculations showed the strong impact of the selection of data on the performance of the model. Almost every aspect of this study was affected by the sampling option. This shows how important it is to carefully choose a proper set of data suitable for a given model. The intraneous and extraneous prediction test clearly proved the significance of this statement. The parameters fitted to a limited range of data points have almost none predictive power when it comes to extrapolating the results far from the fitting domain.

The proper data selection refers also to the precision of the input. If the 'experimental' points are affected by large uncertainties, the obtained results may strongly deviate from the 'real' solution and the obtained model will have a very limited ability to predict any further results.

Chapter 4

Numerical Results

In this chapter we will present the results of our numerical calculations, concerning spherical nuclei. The results were obtained using a Fortran 77 code written especially for this task. The code solves the spherical single-particle Schrödinger equation introduced in the previous chapter. This limitation to spherical symmetry has a strong advantage: the code is fast and simple in its structure and allows for rapid alterations if necessary, e.g. a modification of the interaction potential.

The main purpose of the calculations was to recover the parameters of the single-particle potential from experimental information including first of all proton and neutron energy levels, but also the nuclear radii. Sperical nuclei are the best and most natural choice for this purpose. They are well-studied experimentally and their properties can easiliy be reproduced by simple theoretical models, like the one we used in our work. Furthermore, the parameters of the single-particle spherical Woods-Saxon potential can be incorporated into calculations of deformed nuclei. The calculations we performed, revealed the mutual dependence of the parameters. This property was throughly tested and studied and led to the concept of Predictive Power of a theory.

Our second task was to test the properties and behaviour of the modification of spin-orbit interaction by implementation of density dependence discussed in section 2.3. We will show, that this approch has several advantages and provides a more natural description of spin-orbit interaction.

4.1 Traditional Form of the Spin-Orbit Potential

For the first part of our study we have chosen to recalculate the so-called ‘universal’ parameterisation of Woods-Saxon single-particle potential that enter equations 2.1-2.6 with a new set of experimental data. This parameterisation is used in many applications, e.g. to predict the properties of other, non-sperical nuclei, especially the ones, which are not yet sufficiently explored. We strongly believe, that the revision of the experimental input will result in an improved set of parameters and in turn lead to better and more precise predictions in other regions of nuclear chart.

4.1.1 Experimental Input

The single-particle levels were obtained from the recently revised analysis of experimental data, cf. [22], [24]. Among all the available levels for a given nucleus, these with large experimental uncertainties have not been taken into consideration. Below in tables 4.1-4.7 we compare the experimental data used in our work with the previously used data [23]. We can see, that the change in the positions of some levels is rather significant and as such, it cannot be neglected. Therefore this aspect should be given more consideration.

Table 4.1: *Experimental single particle levels for ^{16}O*

	State	Ref.[24]	Ref.[23]
1	$\pi 1p_{3/2}$	-15.30	-
2	$\pi 1p_{1/2}$	-9.00	-
3	$\pi 1d_{5/2}$	-0.60	-
4	$\pi 2s_{1/2}$	-0.10	-
5	$\pi 1d_{3/2}$	4.40	-
1	$\nu 1p_{3/2}$	-18.60	-
2	$\nu 1p_{1/2}$	-12.40	-
3	$\nu 1d_{5/2}$	-4.14	-
4	$\nu 2s_{1/2}$	-3.27	-
5	$\nu 1d_{3/2}$	0.94	-

Table 4.2: *Experimental single particle levels for ^{40}Ca*

	State	Ref.[24]	Ref.[23]
1	$\pi 1d_{5/2}$	-	-14.27
2	$\pi 2s_{1/2}$	-8.94	-11.02
3	$\pi 1d_{3/2}$	-6.42	-8.33
4	$\pi 1f_{7/2}$	-1.09	-1.71
5	$\pi 2p_{3/2}$	0.69	0.07
6	$\pi 2p_{1/2}$	2.41	1.79
7	$\pi 1f_{5/2}$	-	3.69
1	$\nu 1d_{5/2}$	-	-22.39
2	$\nu 2s_{1/2}$	-16.70	-18.19
3	$\nu 1d_{3/2}$	-14.10	-15.64
4	$\nu 1f_{7/2}$	-8.36	-8.62
5	$\nu 2p_{3/2}$	-5.86	-6.76
6	$\nu 2p_{1/2}$	-4.40	-4.76
7	$\nu 1f_{5/2}$	-1.38	-3.38

The minimisation procedure implemented in the numerical code compares experimental levels with the eigenvalues of the Schrödinger equation calculated for a given set of parameters. In each step, the parameters are adjusted in order to obtain better agreement with the data. However, in case of the lightest nucleus of ^{16}O, the number of experimental levels we can use is rather small and causes difficulties. For both protons and neutrons, we have only 5 levels to compare to, which means that we solve five equations for six or seven parameters. In view of our discussion from chapter 1, this problem has infinitely many solutions and in this form it is not accepted. To avoid this problem, we add additional constraints to our test function, which increases the number of equations.

We choose to control the size of the nucleus by comparing the nuclear radius to its experimental value (or 'extrapolated' values calculated according to Pomorski formula). This allows us to obtain a proper geometry of the nucleus. Analogous result can be obtained by adding a constraint for the density distribution. It can be implemented for several nuclei, where the experimental denisty profiles are known. Both methods, however, yield similar results. In case of protons for ^{16}O, we still need one more constraint. The problem is solved by adding a maximum error control in values of

Table 4.3: *Experimental single particle levels for ^{48}Ca*

	State	Ref.[24]	Ref.[23]
1	$\pi 1d_{5/2}$	-21.50	-20.81
1	$\pi 1d_{3/2}$	-16.68	-15.80
2	$\pi 2s_{1/2}$	-16.39	-15.51
3	$\pi 1f_{7/2}$	-9.63	-9.99
5	$\pi 2p_{3/2}$	-6.55	-7.30
6	$\pi 2p_{1/2}$	-	-5.16
7	$\pi 1f_{5/2}$	-	-5.07
1	$\nu 1d_{5/2}$	-	-17.31
2	$\nu 2s_{1/2}$	-12.54	-13.16
3	$\nu 1d_{3/2}$	-12.52	-12.01
4	$\nu 1f_{7/2}$	-9.94	-9.68
5	$\nu 2p_{3/2}$	-4.60	-5.25
6	$\nu 2p_{1/2}$	-2.86	-3.58
7	$\nu 1f_{5/2}$	-1.20	-1.66

Table 4.4: *Experimental single particle levels for ^{56}Ni*

	State	Ref.[24]	Ref.[23]
1	$\pi 1d_{3/2}$	-8.60	-
2	$\pi 2s_{1/2}$	-7.90	-
3	$\pi 1f_{7/2}$	-5.00	-6.88
4	$\pi 2p_{3/2}$	-0.69	-0.99
5	$\pi 2p_{1/2}$	0.34	0.84
6	$\pi 1f_{5/2}$	0.42	0.13
7	$\pi 1g_{9/2}$	2.82	-
1	$\nu 1d_{3/2}$	-18.40	-
2	$\nu 2s_{1/2}$	-17.80	-
3	$\nu 1f_{7/2}$	-14.60	-16.35
4	$\nu 2p_{3/2}$	-10.25	-10.36
5	$\nu 1f_{5/2}$	-9.48	-9.53
6	$\nu 2p_{1/2}$	-9.14	-8.48
7	$\nu 1g_{9/2}$	-7.24	-

Table 4.5: *Experimental single particle levels for ^{90}Zr*

	State	Ref.[24]	Ref.[23]
1	$\pi 1f_{7/2}$	-	-14.66
2	$\pi 1f_{5/2}$	-10.09	-10.10
3	$\pi 2p_{3/2}$	-9.86	-9.87
4	$\pi 2p_{1/2}$	-8.35	-8.36
5	$\pi 1g_{9/2}$	-5.15	-5.68
6	$\pi 2d_{5/2}$	-1.30	-1.92
7	$\pi 1g_{7/2}$	0.40	-0.13
8	$\pi 3s_{1/2}$	-	0.10
9	$\pi 2d_{3/2}$	-	0.11
1	$\nu 1f_{7/2}$	-	-14.76
2	$\nu 1f_{5/2}$	-13.50	-13.05
3	$\nu 2p_{3/2}$	-13.00	-12.74
4	$\nu 2p_{1/2}$	-12.60	-12.37
5	$\nu 1g_{9/2}$	-11.97	-11.69
6	$\nu 2d_{5/2}$	-7.15	-7.20
7	$\nu 3s_{1/2}$	-5.63	-5.64
8	$\nu 2d_{3/2}$	-4.60	-4.78
9	$\nu 1h_{11/2}$	-	-4.70
10	$\nu 1g_{7/2}$	-4.40	-4.63

Table 4.6: *Experimental single particle levels for ^{132}Sn*

	State	Ref.[24]	Ref.[23]
1	$\pi 2p_{1/2}$	16.01	-15.77
2	$\pi 1g_{9/2}$	-15.71	-15.48
3	$\pi 1g_{7/2}$	-9.68	-10.15
4	$\pi 2d_{5/2}$	-8.72	-9.06
5	$\pi 2d_{3/2}$	-7.24	-7.23
6	$\pi 1h_{11/2}$	-	-7.20
7	$\pi 3s_{1/2}$	-	-6.65
1	$\nu 1g_{7/2}$	-9.75	-9.63
2	$\nu 2d_{5/2}$	-8.97	-9.10
3	$\nu 3s_{1/2}$	-7.64	-7.55
4	$\nu 1h_{11/2}$	-7.38	-7.42
5	$\nu 2d_{3/2}$	-7.31	-7.17
6	$\nu 2f_{7/2}$	-2.47	-2.29
7	$\nu 3p_{3/2}$	-1.62	-1.31
8	$\nu 1h_{9/2}$	-0.91	-0.91
9	$\nu 3p_{1/2}$	-	-0.72
10	$\nu 1i_{13/2}$	-	-0.39
11	$\nu 2f_{5/2}$	-0.47	-0.35

Table 4.7: *Experimental single particle levels for* ^{206}Pb

	State	Ref.[24]	Ref.[23]
1	$\pi 1g_{7/2}$	-	-11.55
2	$\pi 2d_{5/2}$	-	-9.80
3	$\pi 1h_{11/2}$	-9.35	-9.22
4	$\pi 2d_{3/2}$	-8.35	-8.18
5	$\pi 3s_{1/2}$	-8.10	-7.85
6	$\pi 1h_{9/2}$	-3.80	-4.19
7	$\pi 2f_{7/2}$	-2.49	-3.07
8	$\pi 1i_{13/2}$	-1.83	-2.35
9	$\pi 2f_{5/2}$	-0.40	-1.05
10	$\pi 3p_{3/2}$	-	-0.92
11	$\pi 3p_{1/2}$	-	-0.47
1	$\nu 1h_{9/2}$	-	-10.66
2	$\nu 2f_{7/2}$	-10.30	-9.93
3	$\nu 1i_{13/2}$	-9.80	-8.85
4	$\nu 3p_{3/2}$	-8.27	-8.12
5	$\nu 2f_{5/2}$	-7.94	-7.78
6	$\nu 3p_{1/2}$	-7.37	-7.22
7	$\nu 2g_{9/2}$	-3.94	-3.73
8	$\nu 1i_{11/2}$	-3.16	-3.11
9	$\nu 3d_{5/2}$	-2.37	-2.22
10	$\nu 1j_{15/2}$	-	-2.05
11	$\nu 4s_{1/2}$	-1.90	-1.81
12	$\nu 2g_{7/2}$	-1.44	-1.35
13	$\nu 3d_{3/2}$	-1.40	-1.33

energy levels. This is realised by introducing the absolute value of the maximum difference between experimental and theoretical energy levels in a given spectrum into the χ^2-function. This simply means that we accept only these solutions, where all the levels are rather well reproduced.

Both constraints are also useful in other cases, where we have enough experimental data to perform mathematically correct minimisation. Controlling the size of the nucleus allows to restrict the number of solutions to only these, which give physical results. The same applies for the second constraint. We are only interested in obtaining theoretical level schemes that strongly resemble reality. Therefore, controlling the maximum error forces the code to search only for these parameter sets, that satisfy this condition.

4.1.2 Results of Calculations

Pictures 4.1-4.7 present the best fits for all nuclei that we were able to obtain through the minimisation procedure. However, these results are an example of the so-called overfitting mechanism, which occurs mainly when the model has more parameters

than the size of the data sample.

The graphs show the comparison between experimental and theoretical single-particle levels. The experimental data are on the right-hand side of the plot, whereas the calculated levels are on the left. Each level is labeled and its value is indicated in MeV. At the top of the plot there is an information about the number of protons and neutrons, the type of particles that were considered (left picture of each pair shows protons, right – neutrons). The top line contains also the χ^2 value (denoted S^2) and the value of the nuclear radius (R_{rms}) given in femtometers. On the right-hand side of the plot there is a set of parameters that were obtained as a result of the minimisation.

What we see at the first glance is that the agreement of the theoretical levels with experimental data is extremely good. With the exception of neutrons in case of ^{132}Sn and ^{208}Pb, all the results are almost exact. Moreover, in some cases the accuracy of the fit exceeds the experimental uncertainties. This may seem as an extremely attractive result, but if we take a closer look at parameters obtain through the minimisation, we immediately realize the problem – some of the spin-orbit parameters have strange and undesirable values, eg. $a_{so} = 0.07$ in case of neutrons for ^{40}Ca nucleus.[1]

Performing calculations with minimisation procedures such as the one we present in this work, we force the numerical code to choose parameters from a specific range, which is more or less acceptable from the point of view of nuclear theory. This means that we limit the outcome and search for 'physical' parameters. The results we present were obtained because we allowed the code to go outside the normal range. While this may seem pointless, as we are not really interested in non-physical outcomes, it revealed the limitations of the model. The parameters of the central part of the chosen potential fall in the expected range, the most problematic part is the spin-orbit interaction. The three parameters seem to be random and thus non-transferable to other nuclei. We have no way of knowing which set to choose or how can we find averaged values of the parameters that will work for other nuclei.

Other aspect, that can be raised at this point, is the reliability of the model. It seems extremely unlikely for a model as simple as Woods-Saxon based mean-field to be able to reproduce with such a high level of accuracy the complexity of the subatomic structures. This suggest that the model is ill-posed in the so-called 'overfit' regime and unsuitable for this task. A possible solution, i.e. a replacement of the classical spin-orbit with density-dependent potential, which amount to diminishing the number of parameters was proposed in chapter 2.3. The next section wil present results of calculation with such a potential.

Conclusions. These results are a crucial step in the search for a simple theory applicable to a wide range of nuclear chart. They undeniably revealed the limitations of classical spin-orbit formulations and proved the model to be incomplete and unreliable. They also showed how important it is to consider and test the predictive power of a given model. One can be easily misguided by seemingly good performance of the calculations without proper control of the mathematical and statistical side of the model.

By increasing the number of constraints we eliminate overfitting and the model may

[1] The diffuseness parameter is expected to be around 0.60, which is one order of magnitude larger than the value obtained. Our result gives almost square shape for the WS formfactor in the spin-orbit interaction.

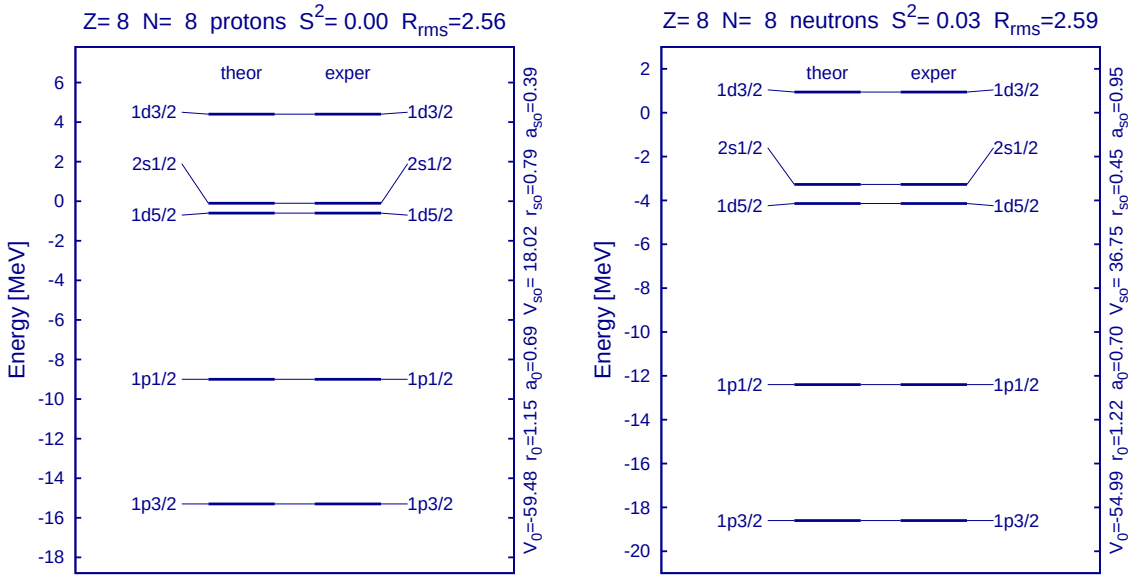

Figure 4.1: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{16}O

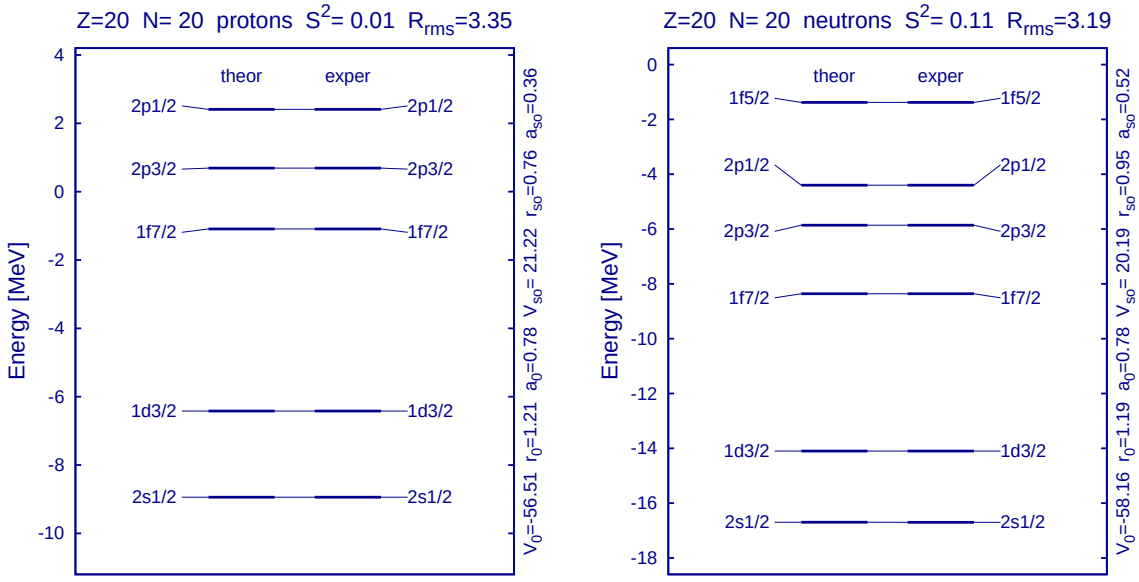

Figure 4.2: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{40}Ca

acquire predictive power.

This is also a strong indication that a new way of calculating the spin-orbit splitting is necessary.

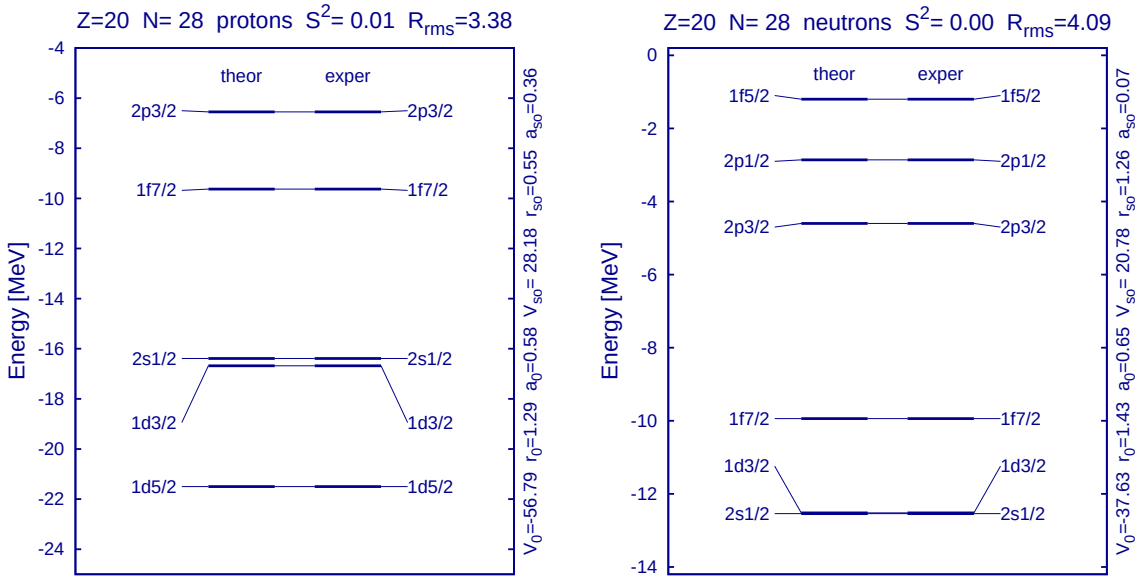

Figure 4.3: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{48}Ca

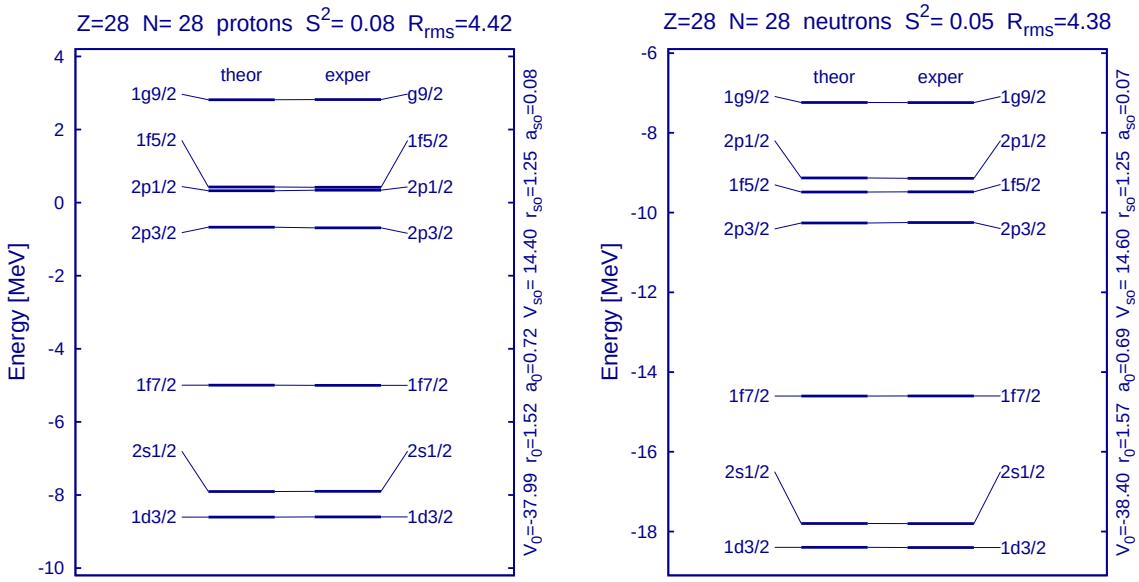

Figure 4.4: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{56}Ni

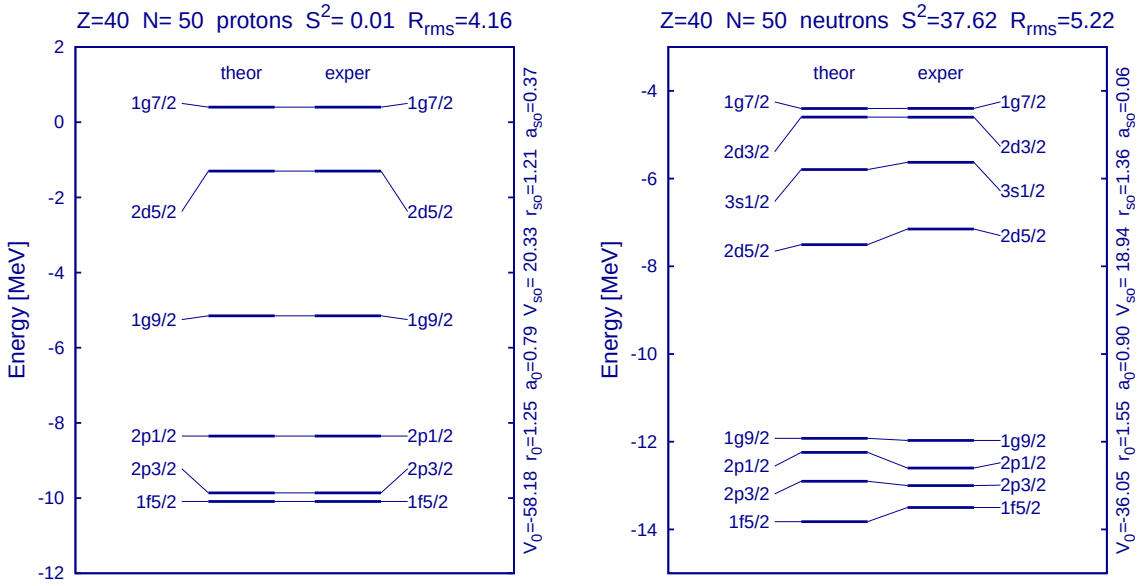

Figure 4.5: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{90}Zr

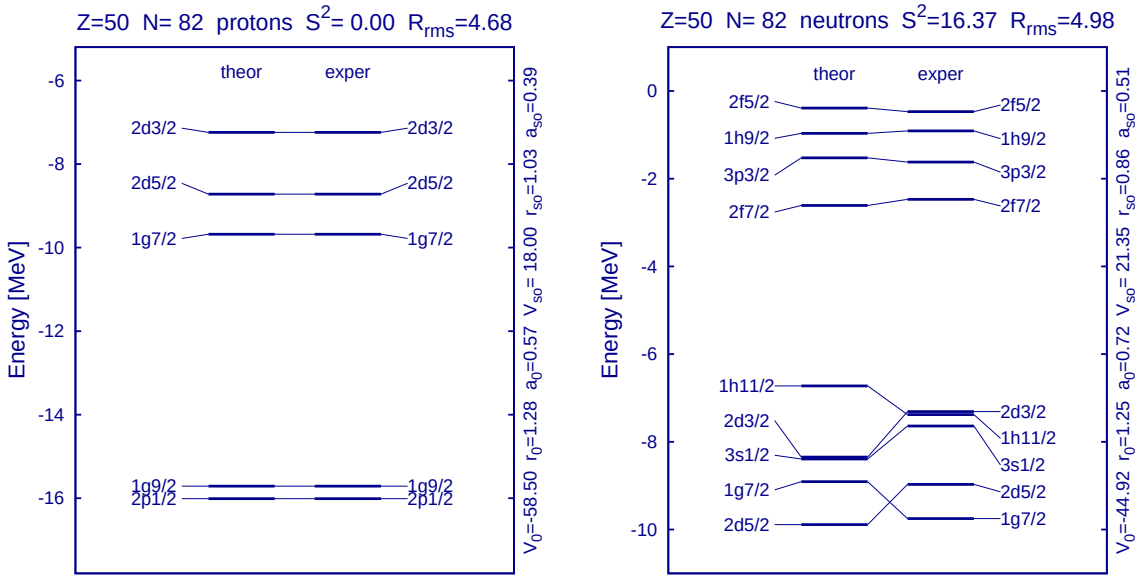

Figure 4.6: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{132}Sn

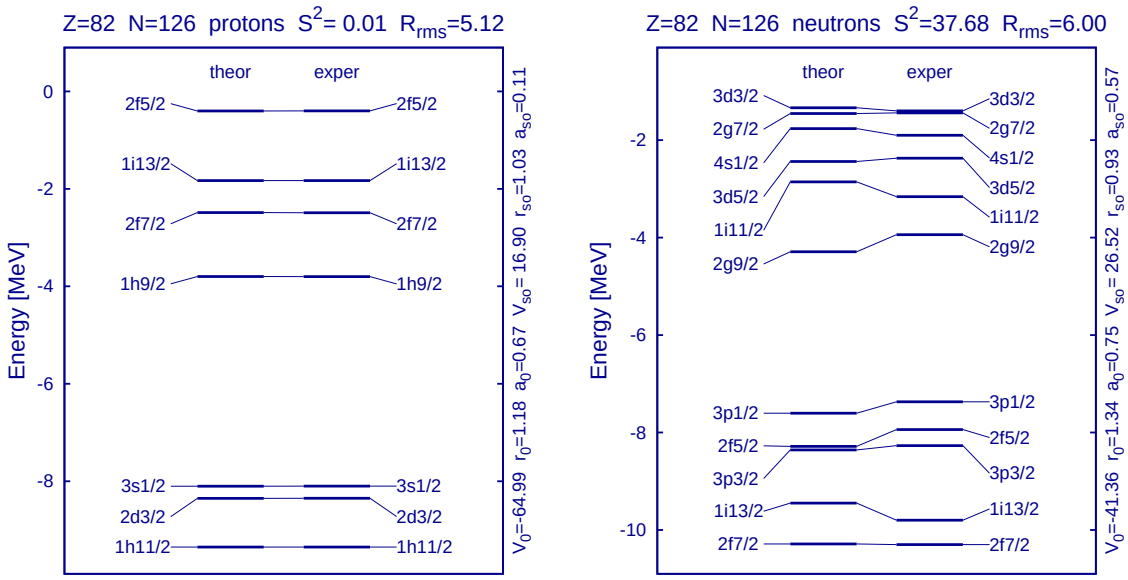

Figure 4.7: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{208}Pb.

4.2 Density Dependent Spin-Orbit Potential

On the basis of the discussion of the previous sections, we may say, that the Woods-Saxon standard approach, which includes the spin-orbit interaction parameterised using a WS form-factor (eq. (2.4)) has many disadvantages. It depends on a set of 6 parameters for neutrons and 7 for protons (additional Coulomb radius), which means 13 different (and in theory uncorrelated) parameters, which need to be adjusted to a limited number of data points. This is, unfortunately, a troublesome feature of this model as it may lead to overfitting, especially for lighter nuclei. Moreover, this way of calculating the spin-orbit interaction is artificial and does not represent all the fluctuations of the nuclear matter. Using the Woods-Saxon-type functions for the central part of the mean-field potential is justified by the bulk properties of the nucleus (cf. section 2.3.4). The spin-orbit interaction is more sensitive to changes in the mutual positions of nucleons.

As an alternative solution, we have presented a density-dependent form of the potential, cf. eq. (2.49). This formulation, guided by Hartree-Fock formalism, allows to incorporate all the advantages of self-consistent approach into a simple model based on Woods-Saxon central interaction. The first step of our work included only the first part of equation (2.49), containing the particle densities, i.e. the equivalent of the classical spin-orbit. As a result of replacing the WS-type spin-orbit potential with this new formulation, we obtain an appropriately reduced set of parameters. In the worst case scenario the model has 11 adjustable parameters, but we will present results suggesting that this number can be further reduced to 8 (or 7, if we consider the Coulomb radius parameter as not related to nuclear properties studied here and keep it fixed) parameters, with only one responsible for the spin-orbit interaction.

The method of calculating the density-dependent spin-orbit interaction potential is a semi self-consistent process. The particle densities, which enter the potential, are obtained iteratively. This is continued until the obtained single-particle energy levels are the same in two consecutive iterations up to a certain level of accuracy. This method requires both types of particles to be calculated simultaneously as both proton and neutron desities enter the spin-orbit potential for each of them. This way we attempt to treat the nucleus as a whole system, which is another advantage of the method bringing us closer to microscopic methods. Yet still it preserves the simplicity of Woods-Saxon approach with the great advantage of the parametric 'robustness' of the WS central potential.

The potential used in our calculations is defined as follows

$$V = V_C^{WS}(r) + \frac{1}{r}\left(\lambda^{qq}\frac{d\rho_q}{dr} + \lambda^{qq'}\frac{d\rho_{q'}}{dr}\right)\vec{\ell}\cdot\vec{\sigma} + [V_{\text{coul}}(r)]\,. \tag{4.1}$$

Generally, for each type of particles it depends on 5 parameters: 3 for the central potential and 2 for the spin-orbit interaction. The Coulomb radius parameter is chosen to have a fixed value. Based on preliminary results suggesting a linear correlation of all four spin-orbit parameters, we have set

$$\lambda^{qq} = \lambda^{qq'} = \lambda^{q'q} = \lambda^{q'q'} \stackrel{def}{=} \lambda. \tag{4.2}$$

This additional constraint means we are left with only one parameter for both protons and neutrons, which results in a total of 7 parameters.

Below, we present a set of plots comparing, similarly as in previous section, theoretical levels with experimental data. On the right-hand side of the frame of each plot, there is a set of parameters obtained as a result of the minimisation procedure. The last one of four is the spin-orbit strength parameter. As we can see, the levels are rather well reproduced. The best agreement was, of course, obtained for the two lightest nuclei of ^{16}O and ^{40}Ca, but also other results, like for example ^{56}Ni should be noted. For other nuclei the agreement is still very good, however some levels are on the incorrect relative positions as compared to the experimental data.

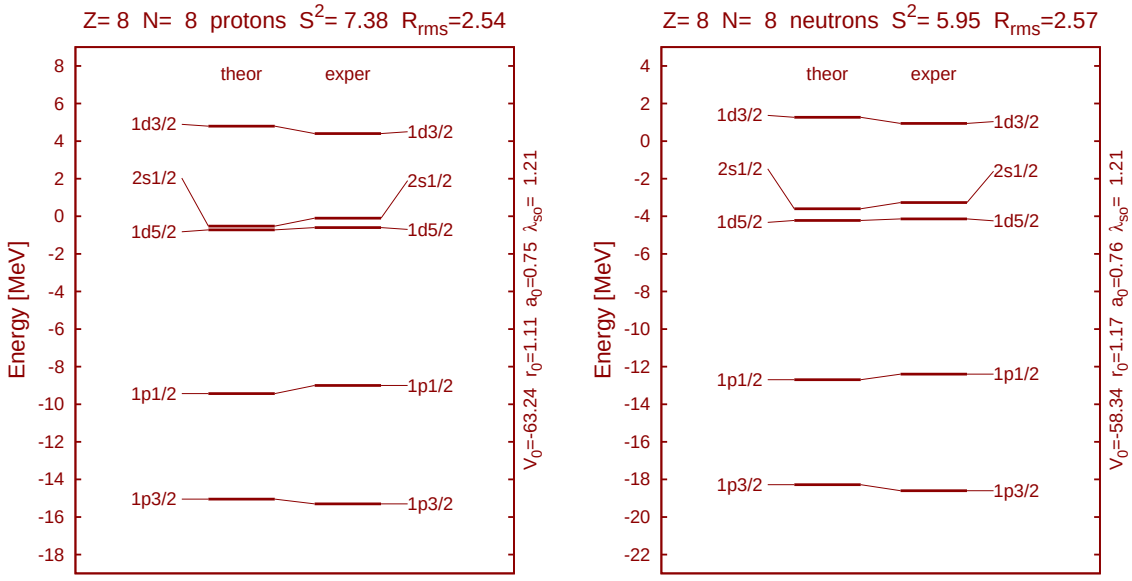

Figure 4.8: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for ^{16}O*

Conclusions. We have presented a method combining the classsical phenomenological Woods-Saxon approach with more advanced and modern self-consistent Hartree-Fock methods. Our calculations show great possibilities of this model. It strongly reduces the number of parameters required to completely describe the mean-field potential, which is a great improvement as compared to standard Woods-Saxon parameterisation. In particular we may get rid of the artificial WS form-factor in the spin-orbit interaction replacing it with nucleonic densities which are a natural part of the Schrödinger equation and certainly the most 'natural' theoretical description of a particle inside a nucleus.

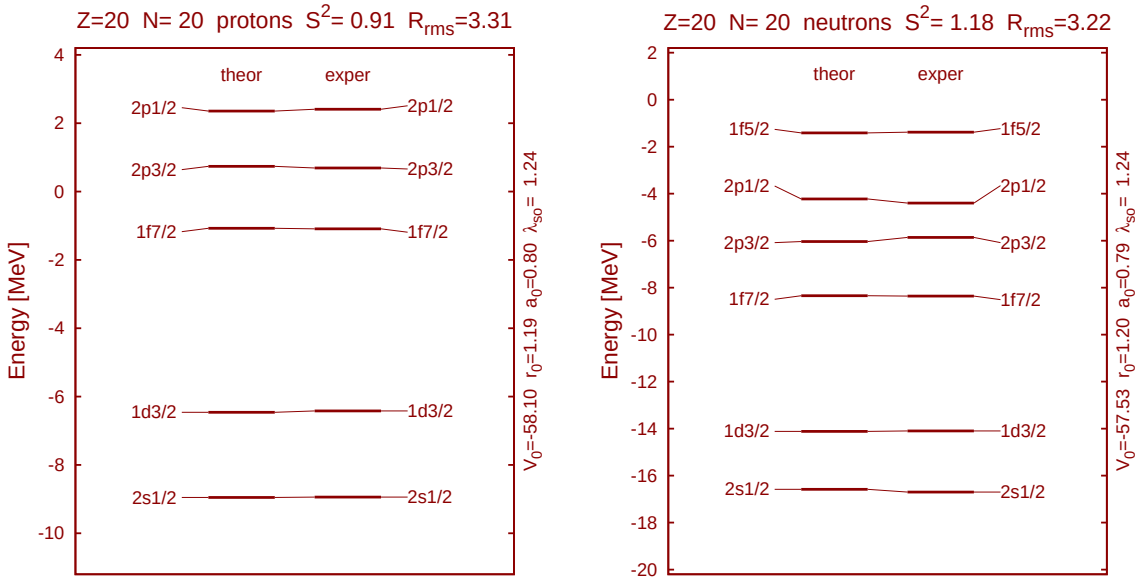

Figure 4.9: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{40}Ca

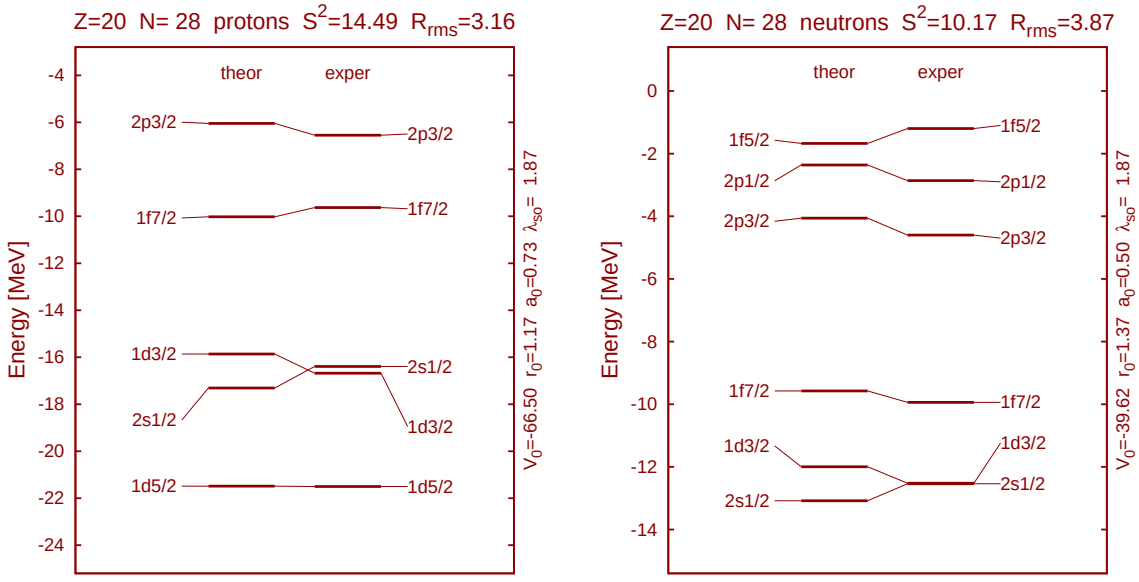

Figure 4.10: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{48}Ca

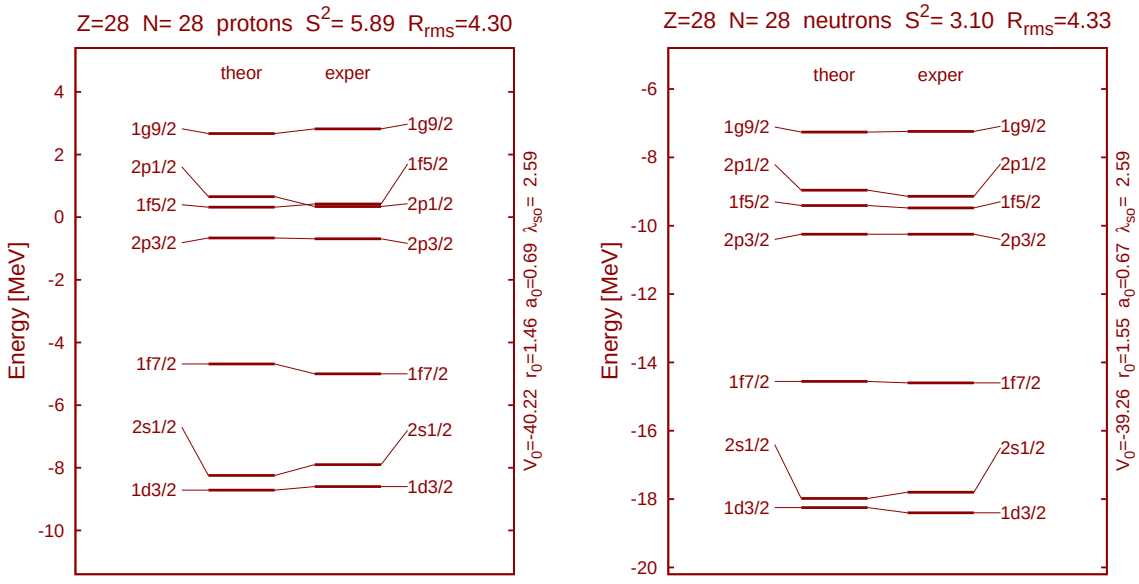

Figure 4.11: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{56}Ni

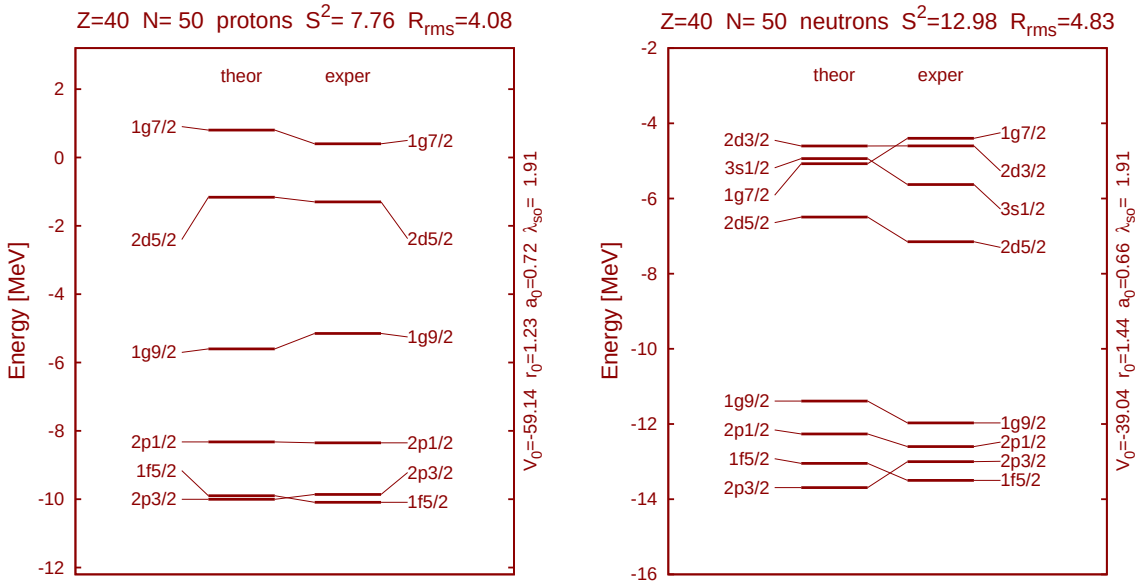

Figure 4.12: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{90}Zr

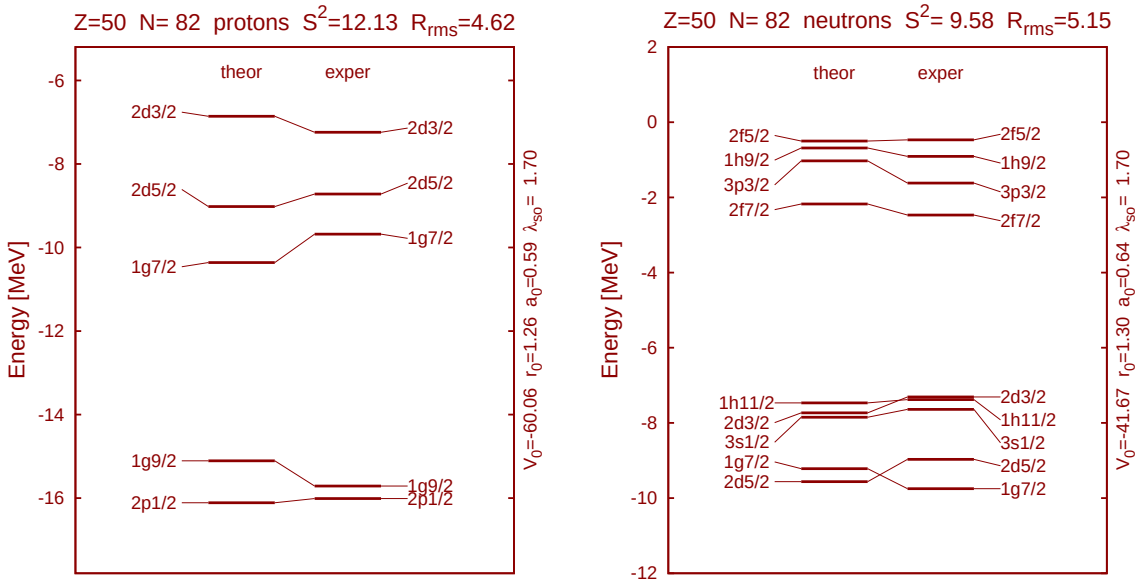

Figure 4.13: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{132}Sn

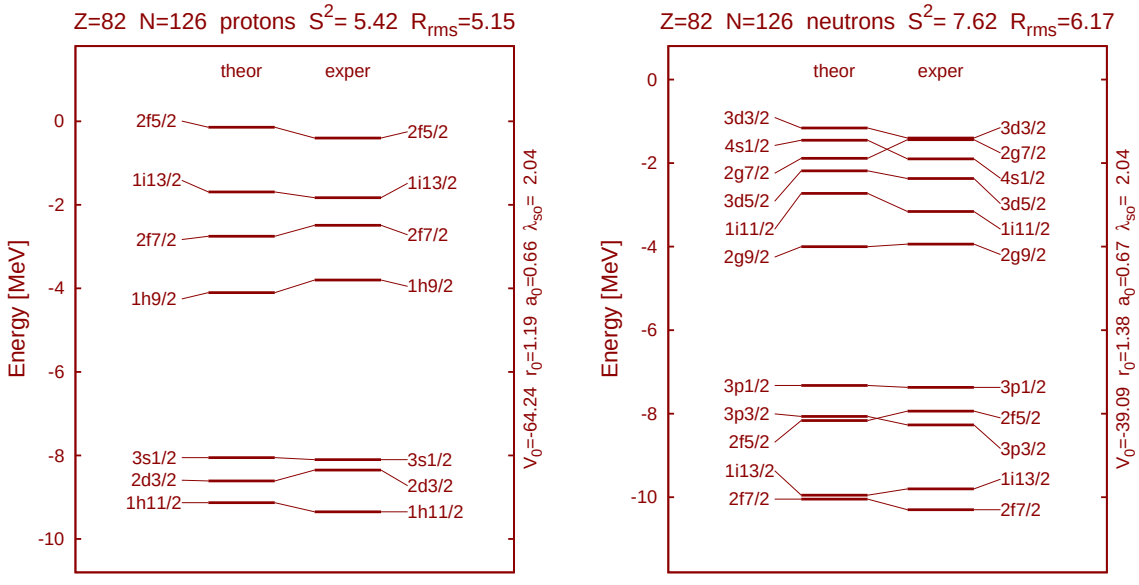

Figure 4.14: *Comparison of experimental and theoretical single-particle energy spectra for protons (left) and neutrons (right) for* ^{208}Pb.

4.3 Tensor Interaction

The choice of nuclei used in our calculations allowed us to test the specific behaviour of energies of nucleonic states, where tensor part of spin-orbit splitting may possibly influence their positions. As it was already mentioned, in nuclei in which full j orbitals are occupied the tensor interaction affects only the spin-unsaturated nuclei; the strongest impact is in the nuclei where only one level from a spin-orbit split-pair is occupied.

Guided by this information, we have chosen to test the behaviour of the spin-orbit splitting in the case of spin-saturated ^{40}Ca and spin-unsaturated ^{48}Ca. In the latter nucleus the last occupied neutron level is $\nu 1f_{7/2}$, while its spin-orbit partner $\nu 1f_{5/2}$ is empty. This means that if the tensor interaction is included, we should observe a reduction of the spin-orbit splitting of protons in ^{48}Ca as compared to the behaviour of the same gap in ^{40}Ca.

We have implemented the tensor part of the spin-orbit splitting according to Eq. (2.49). The full potential can be thus written as

$$V = V_C^{WS}(r) + \frac{1}{r}\left(\lambda^{qq}\frac{d\rho_q}{dr} + \lambda^{qq'}\frac{d\rho_{q'}}{dr}\right)\vec{\ell}\cdot\vec{\sigma} + \frac{1}{r}\left(\alpha J_q + \beta J_{q'}\right)\vec{\ell}\cdot\vec{\sigma}. \tag{4.3}$$

The calculations were performed assuming the following constraints.

$$\lambda^{qq} = \lambda^{qq'} = \lambda > 0 \quad \text{and} \quad \alpha + \beta \simeq 0; \quad \alpha < 0;\ \beta > 0. \tag{4.4}$$

The first constraint concerning the ordinary spin-orbit interaction originates from the results of our calculations, which was presented in the previous section and it differs from the parameterisation given in [17]. However, in this context it is an unimportant difference as we are comparing the influence of the tensor part of the spin-orbit splitting. As for the second constraint, we follow [18].

The sizes of the spin-orbit splitting of proton $1d$ partners are presented in the figure 4.15. The choice of levels for this plot was dictated by the relative strength of the spin-orbit splitting. In case of calcium, it will be the largest for these levels. Both results are calculated using the same set of parameters, which is given at the top of the graph. The plot on the left side shows the levels of ^{40}Ca, whereas the one on the right is for ^{48}Ca. We can observe, that even if α is nonzero, nothing happens for spin-saturated nucleus. The spin-orbit splitting is exactly the same as for normal spin-orbit term. The situation is completely different for spin-unsaturated ^{48}Ca. In this case, the gap is slightly reduced if the tensor part of the spin-orbit interaction is included.

This fact can be easily explained, if we plot the radial components of the spin density in both cases, cf. figure 4.16. We can see, that in the case of ^{40}Ca both proton and neutron parts of the vector density are small. This changes when we go to spin-unsaturated case. The neutron part calculated for ^{48}Ca is rather large and positive. If we now recall the exact expression for vector spin-orbit term in spherical symmetry given by eq. (2.64), we can immediately attribute this difference to the missing spin partner of level $\nu 1f_{7/2}$.

We can also compare the two terms of spin-orbit potential in both cases, i.e. for protons

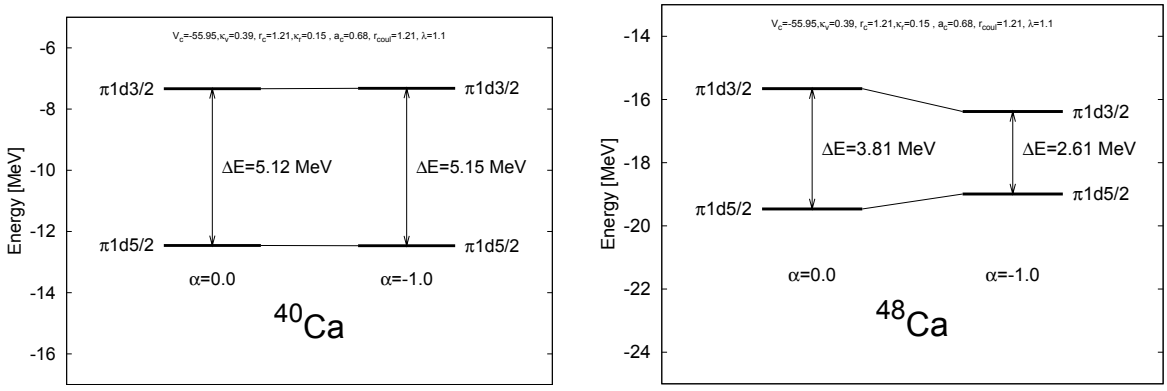

Figure 4.15: *Comparison of energies of proton spin-orbit partners $\pi 1d_{3/2}$ and $\pi 1d_{5/2}$ with (α=-1.0) and without (α=0.0) tensor interaction for ^{40}Ca (left) and ^{48}Ca (right).*

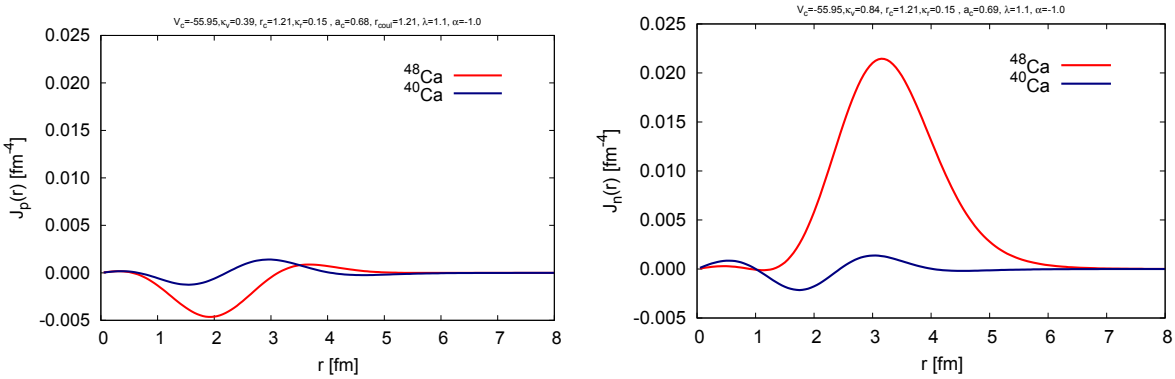

Figure 4.16: *Comparison of radial component of vector spin-orbit densities for protons (left) and neutrons (right).*

in ^{40}Ca and ^{48}Ca. They are given by the following expressions

$$V_\rho^p = r^2(1/r)\lambda\left[\frac{d\rho_p}{dr}+\frac{d\rho_n}{dr}\right], \tag{4.5}$$

$$V_T^p = r^2(1/r)\left[\alpha J_p+\beta J_n\right] = r^2(1/r)\alpha\left[J_p-J_n\right]. \tag{4.6}$$

These two potentials are presented in fig. 4.17 for both nuclei. The left panel shows ^{40}Ca and the right one presents ^{48}Ca. We can observe, that for the spin-saturated nucleus, the tensor contribution to the total spin-orbit potential is practically zero. In case of spin-unsaturated ^{48}Ca there is a small positive potential. It causes the 'abnormal' spin-orbit splitting (due to its opposite sign as compared to standard spin-orbit potential) and thus the reduction of the energy gap.

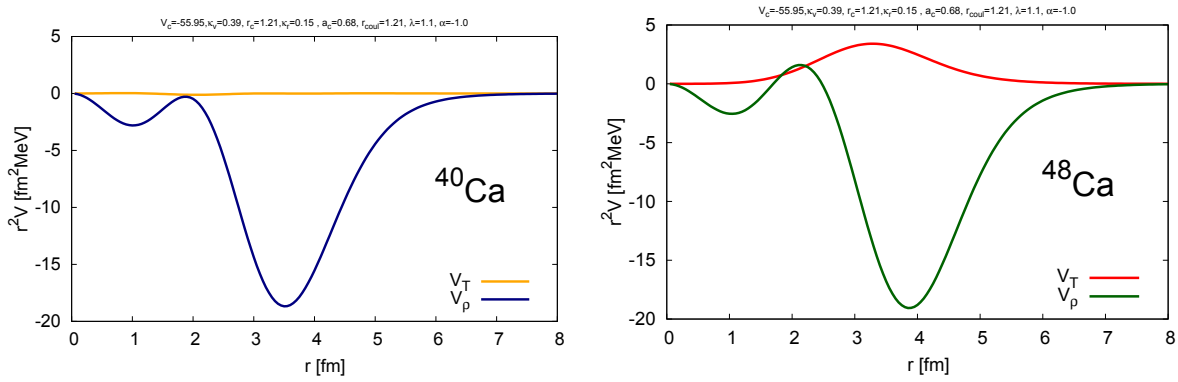

Figure 4.17: *Comparison of terms of spin-orbit interaction for ^{40}Ca (left) and ^{48}Ca (right).*

4.4 Correlation of Woods-Saxon Parameters

The Predictive Power of a mathematical model is a crucial aspect of development in each theory and understanding the underlying physics. The correlation of parameters may cause instability of the calculations (small change of parameters results in large changes in the output), but also, what is important from the statistical point of view, the confidence intervals of parameters tend to infinity. Guided by this, we performed a very straightforward test to reveal the correlation between the parameters of the Woods-Saxon Hamiltonian. From previous research, we already know, that for the central term there is a correlation between the depth of the potential well and its radius, namely $V_o r_o^2 \approx const$. Our calculations showed other, more intresting results.

The main idea behind this test was to try to recover an exact solution of our Hamiltonian defined in chapter 2.2.1 through a fitting procedure. A set of single-particle levels was first generated using the so-called Universal WS parameterisation. It was then treated as a 'new' experimental input for the code. The six parameters form two groups, one set is for the central part, the other for spin-orbit potential. Thus, we choose to test them separately. In search for the correlation we fix one parameter of the set and map the remaining two. This means that their values are varied within a given range and at each point the χ^2 function is calculated cf. the following figures in this section. This allows us to deduce the parametric correlations by examining the behaviour of the χ^2-test.

In real calculations we are strongly limited in the number of available single-particle levels, therefore we decided to mimic this by reducing the number of levels we compare to only six. Additionally, the experimental information is usually limited to the levels lying close to the Fermi level, which originates from the nature of experiment used to obtain this information. However, as it turns out, it is not the best choice for the purpose of recovering the stable sets of paramaters of the WS potential. The results we present were performed for neutrons in ^{208}Pb nucleus as an academic test case. The set of six 'experimental' levels was chosen based on its position, starting form the bottom of the spectrum up to the Fermi level.

The plots in figures 4.18 and 4.19 show the dependence of the χ^2 function on two

central parameters V_o and r_o with a_o remaining fixed. In the first case the lowest states were used, namely $1s_{1/2}$, $1p_{3/2}$, $1d_{3/2}$, $2s_{1/2}$, $1f_{5/2}$ and $2p_{1/2}$. The second figure presents the results of fitting to $1h_{11/2}$, $2f_{5/2}$, $3p_{1/2}$, $1i_{13/2}$, $2g_{9/2}$ and $3d_{3/2}$. In both cases we can see the valleys marked in blue instead of single points, as we would wished to obtain, corresponding to the best fit. However, for the low-lying states it seems that the minimum is more distinct, whereas for the second case it would be rather impossible to give any reasonable solution for the parameters. The next four plots in figures 4.20, 4.21, 4.22, 4.23 present the results for fixed r_o (first two) and V_o (last two) in the same manner[2]. Once again the parametric behaviour is visible. This time, the levels around the Fermi surface behave slightly better, although there would still be a problem of the confidence intervals. This result means that we are unable to fit the parameters to six experimental levels only. We can also reverse this conclusion and notice, that with this type of parametric dependence fitting an arbitrary set of levels may turn out to be easy and give an 'exact' result.

The next set of figures 4.24-4.29 shows the result of the same procedure for the spin-orbit interaction. This time the best behaviour is found for the levels near the Fermi surface. This is not surprising as the shell effects are strongest for those levels. What is new is the exsistence of double solution visible in figure 4.25. We call it the 'compact' and 'non-compact' solution. The first one refers to the small value of r_{so}, whereas the latter is for the large value.

Conclusions. From these tests we can conclude, that a typical experimental information of six single-particle levels lying close to Fermi level may not be sufficient to recover the parameters of the mean-field potential. Deeply bound levels seem to constrain central potential relatively stronger, therefore they should be used as experimental input in this type of calculations.

The spin-orbit interaction defined in a classical way manifests two-valued parametric dependencies. In fact there are two different solutions referred to as 'compact' and 'non-compact' with r_{so}/a_{so} 'small' or 'large'. For recovering this part of parameterisation, one needs the levels around the Fermi surface.

The presence of the valleys indicates also a possible way of eliminating the undesired parametric correlations. Indeed, writing down an equation for the curve connecting all the points at the bottom of the valleys we can express e.g. $V_o = V_o(r_o)$ or, alternatively $a = a(r_o)$ etc. Such a manipulation can limit the number of really independent parameters and stabilise, according to the stochastic theory of the inverse problem, the final performance of the fit.

[2]It should be noticed that the set of levels used in each plot may differ, the size and position of the energy window is unchanged.

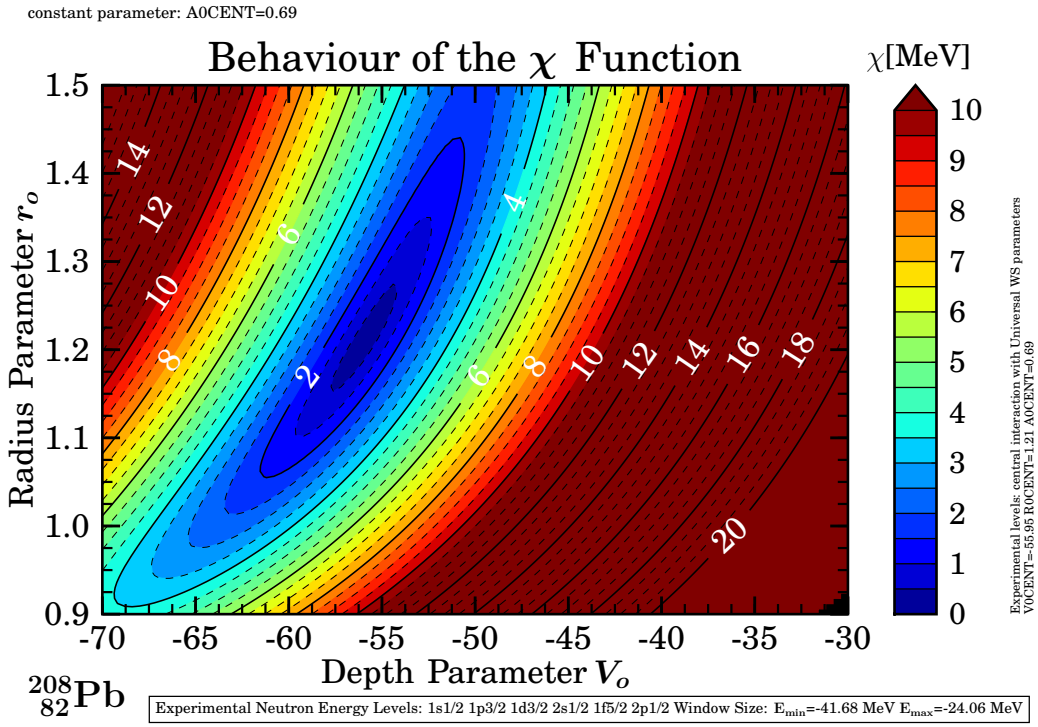

Figure 4.18: *Correlation of parameters of the central potential V_o vs. r_o with energy window far from Fermi level.*

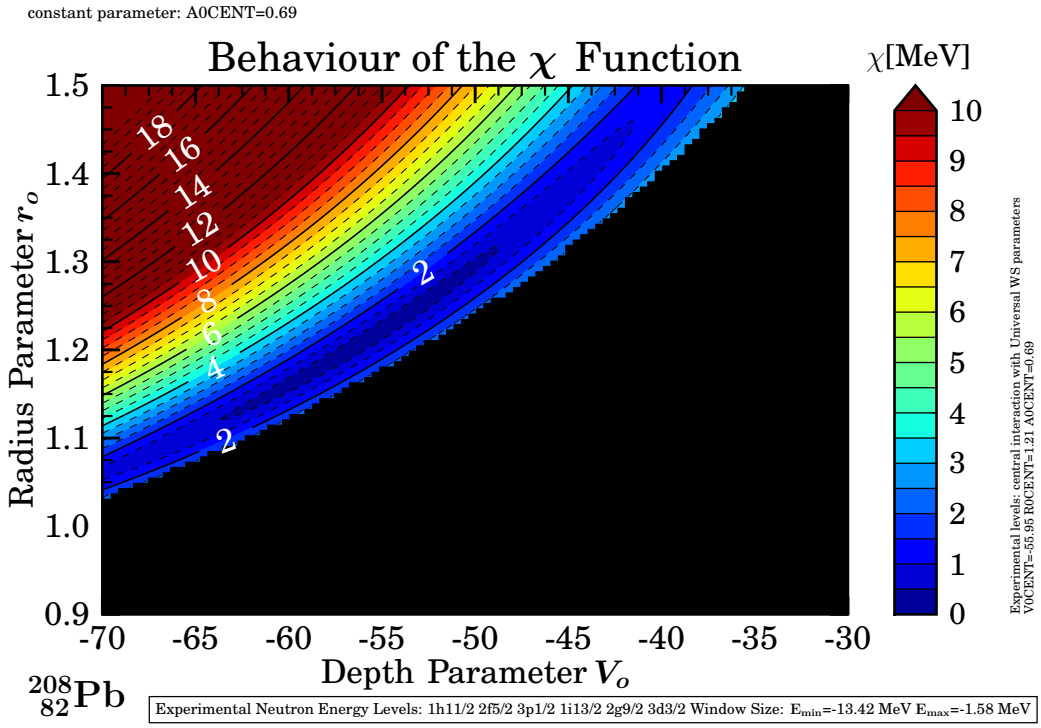

Figure 4.19: *Correlation of parameters of the central potential V_o vs. r_o with energy window close to Fermi level.*

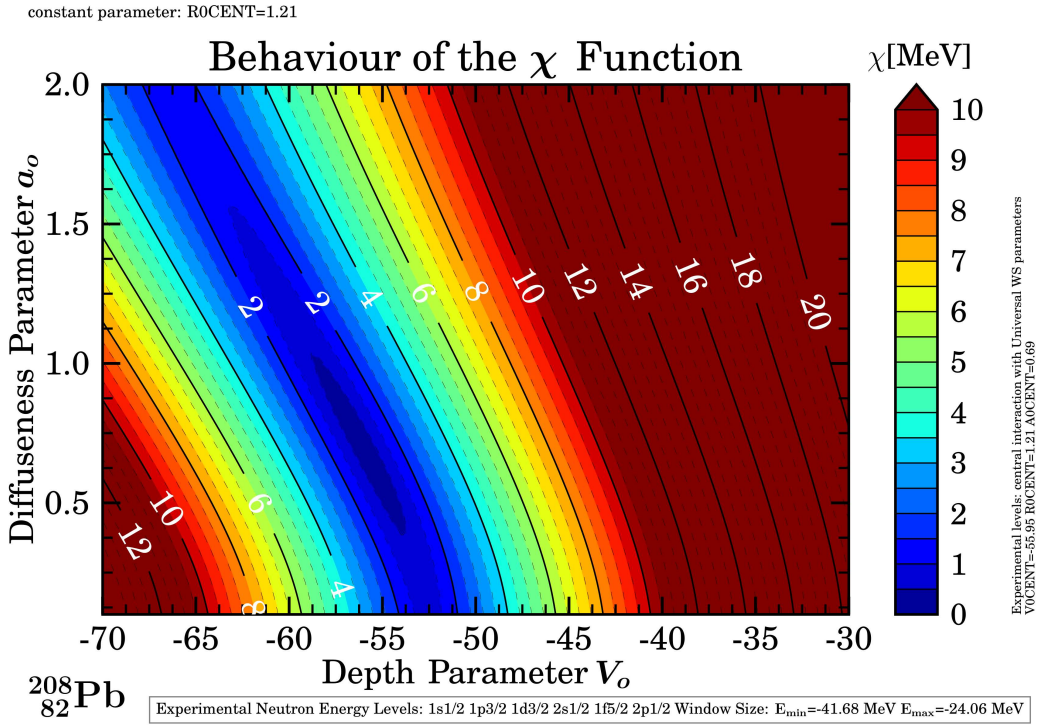

Figure 4.20: *Correlation of parameters of the central potential V_o vs. a_o with energy window far from Fermi level.*

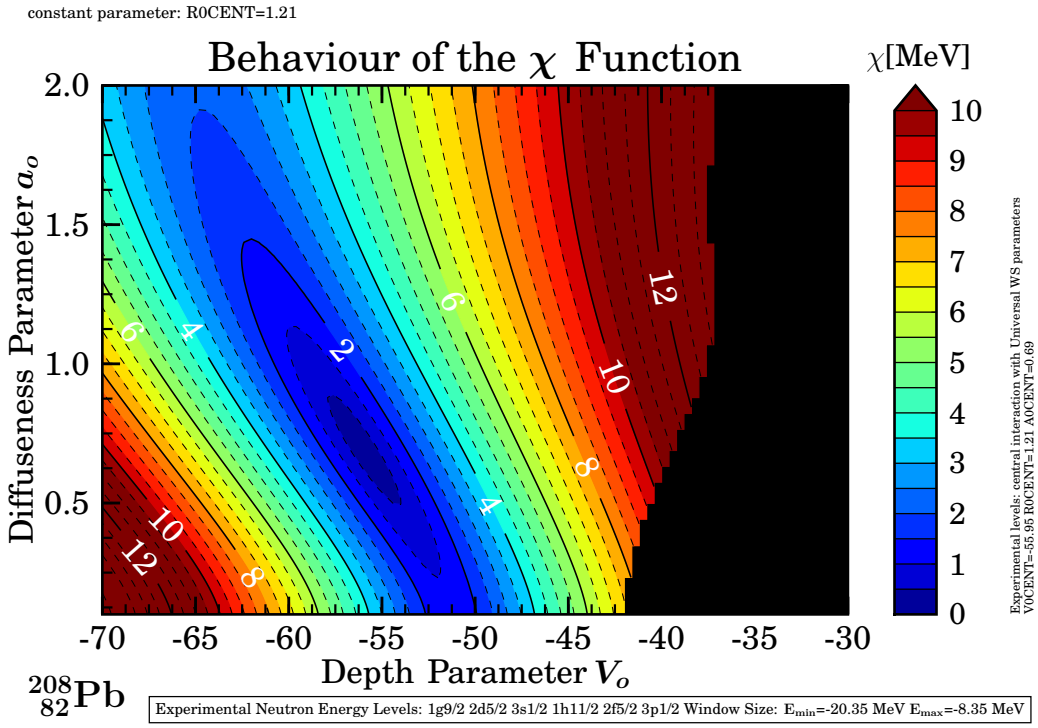

Figure 4.21: *Correlation of parameters of the central potential V_o vs. a_o with energy window close to Fermi level.*

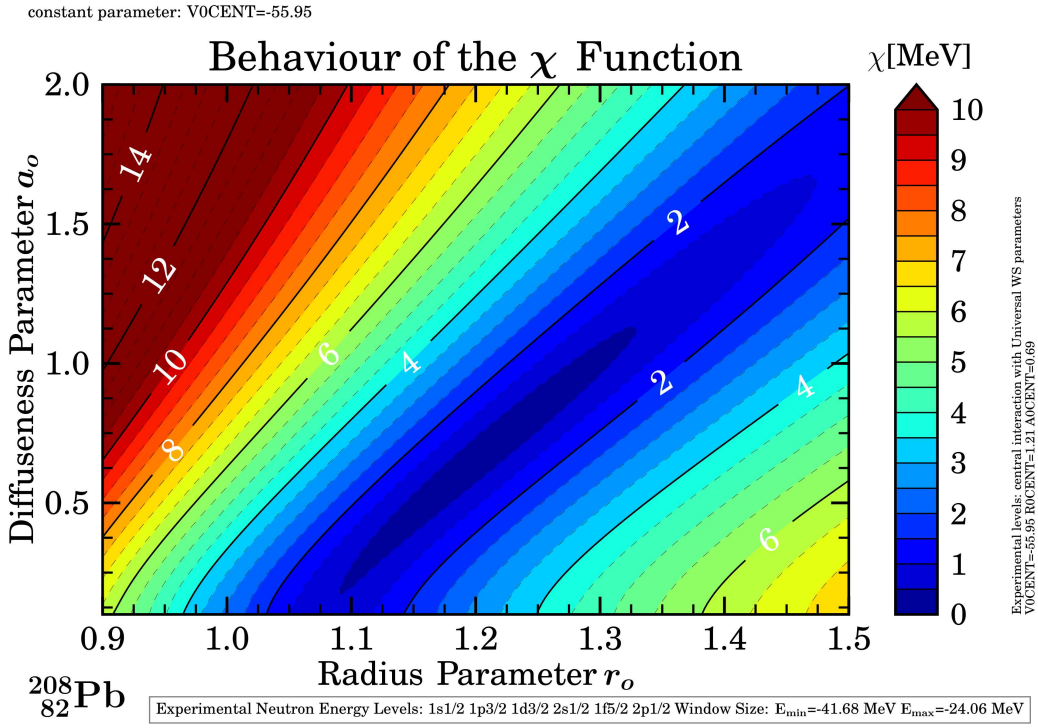

Figure 4.22: *Correlation of parameters of the central potential a_o vs. r_o with energy window far from Fermi level.*

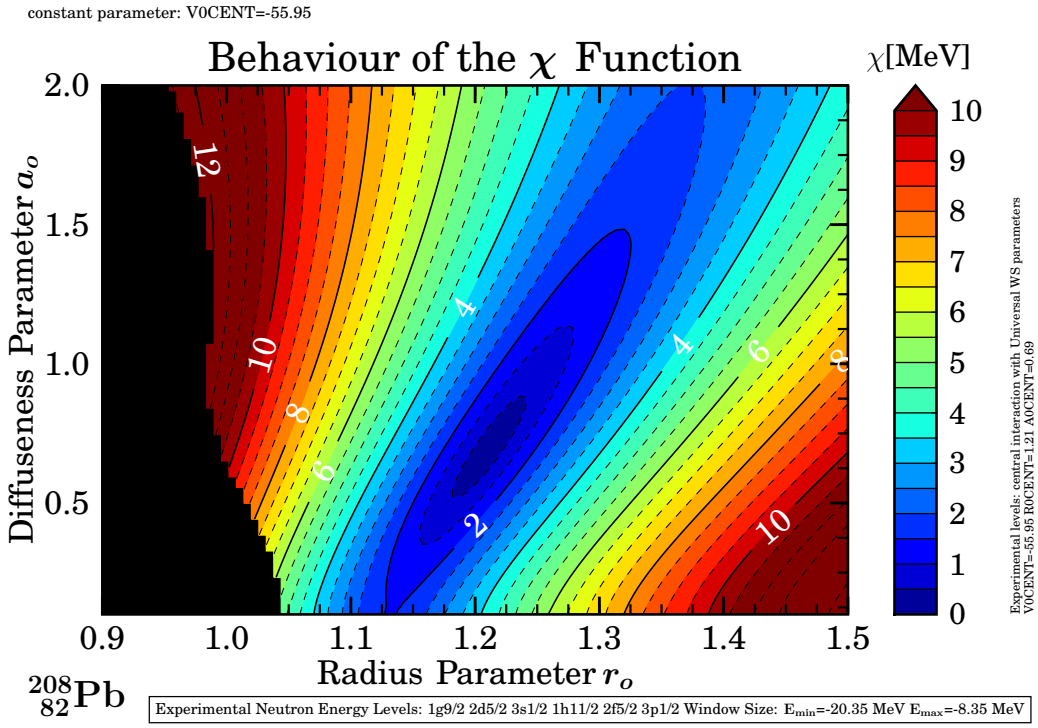

Figure 4.23: *Correlation of parameters of the central potential a_o vs. r_o with energy window close to Fermi level.*

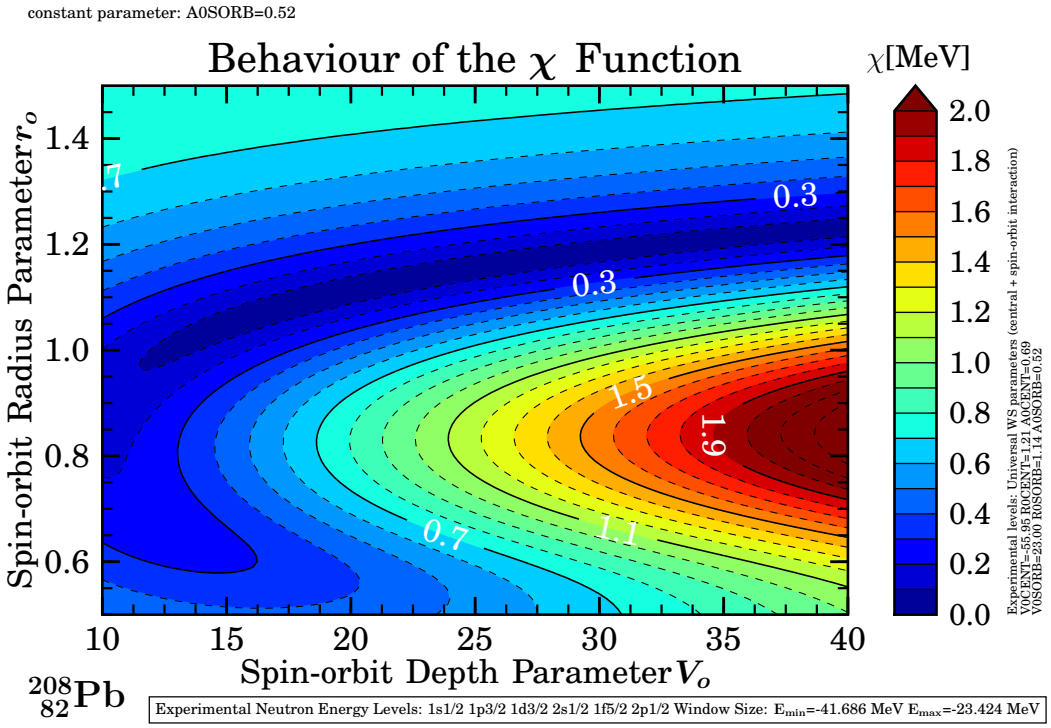

Figure 4.24: *Correlation of parameters of the spin-orbit potential V_{so} vs. r_{so} with energy window far from Fermi level.*

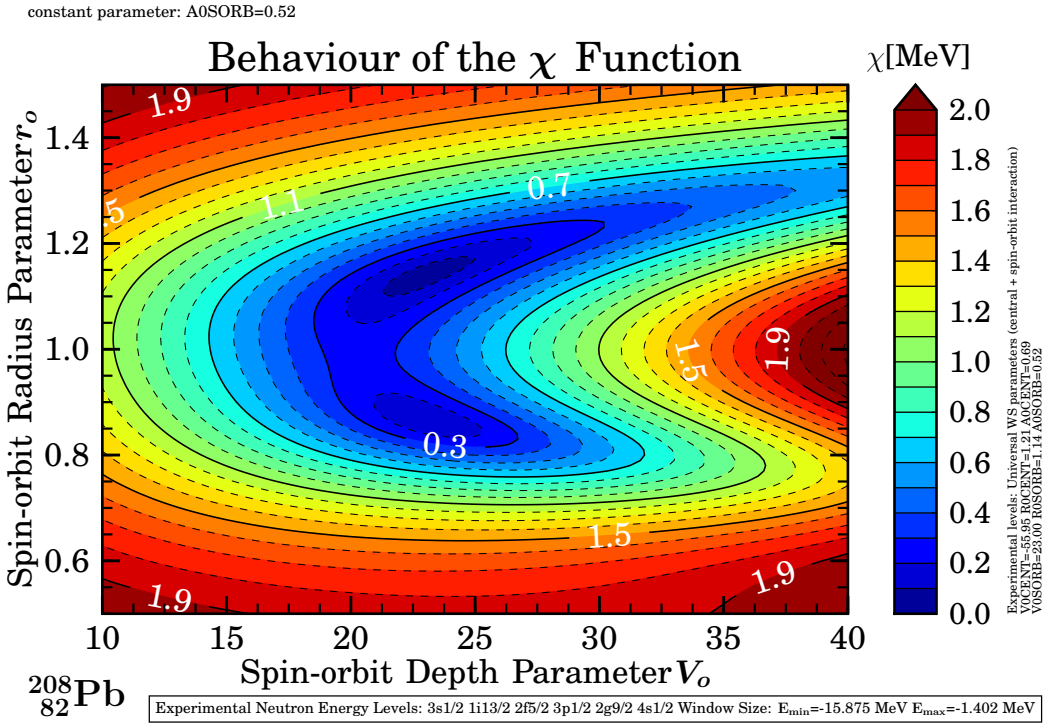

Figure 4.25: *Correlation of parameters of the spin-orbit potential V_{so} vs. r_{so} with energy window close to Fermi level.*

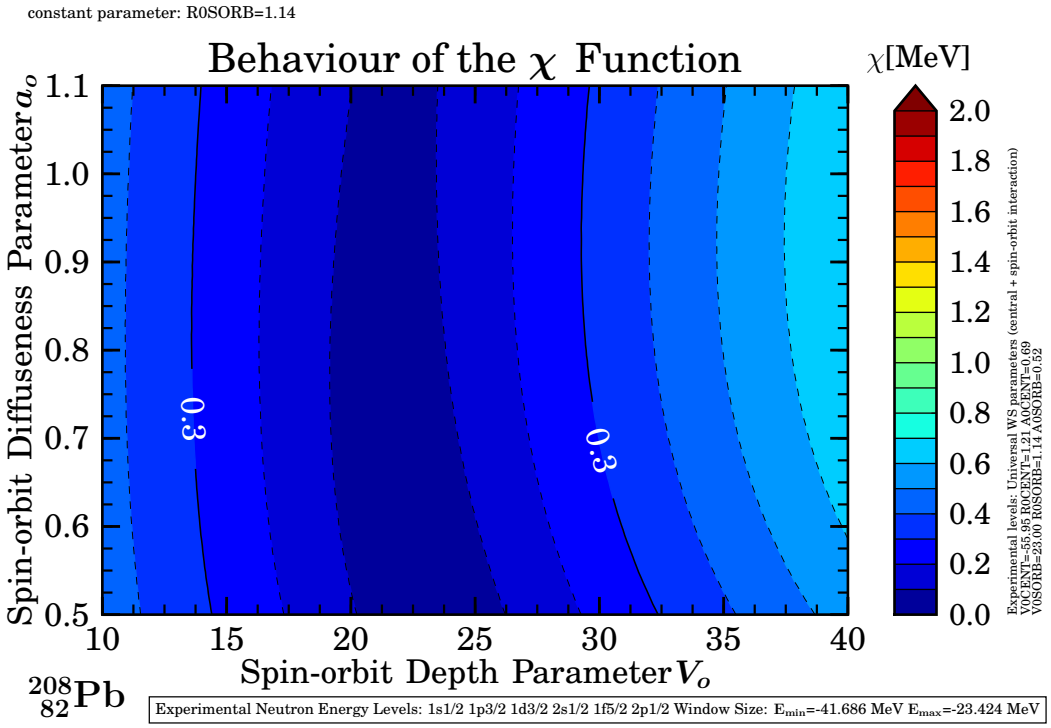

Figure 4.26: *Correlation of parameters of the spin-orbit potential V_{so} vs. a_{so} with energy window far from Fermi level.*

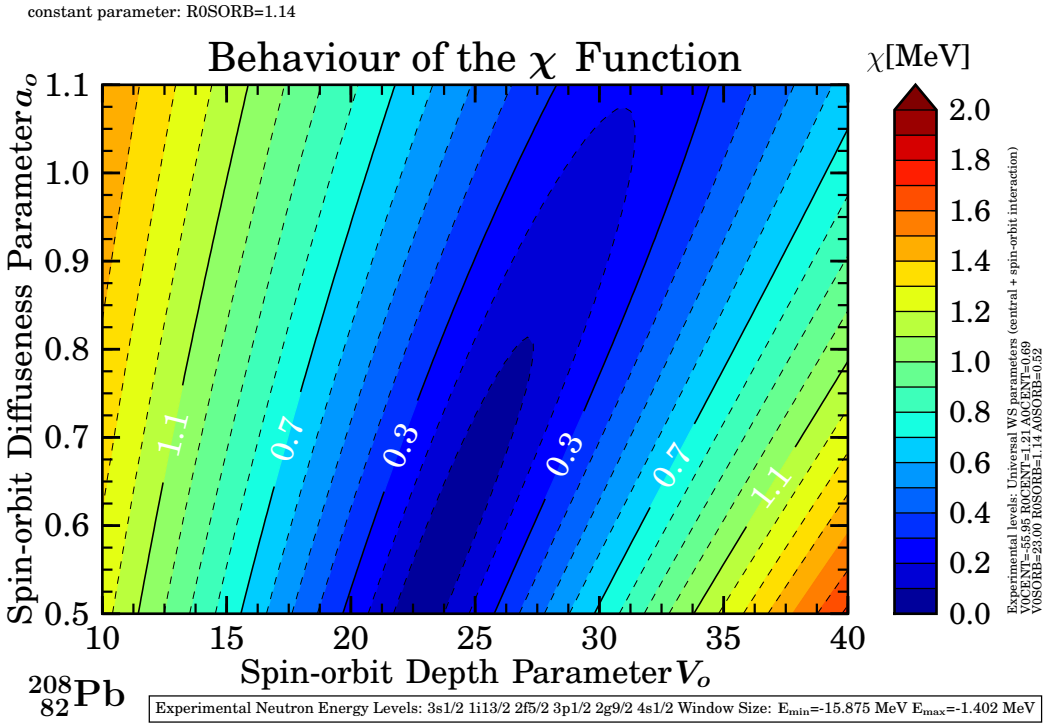

Figure 4.27: *Correlation of parameters of the spin-orbit potential V_{so} vs. a_{so} with energy window close to Fermi level.*

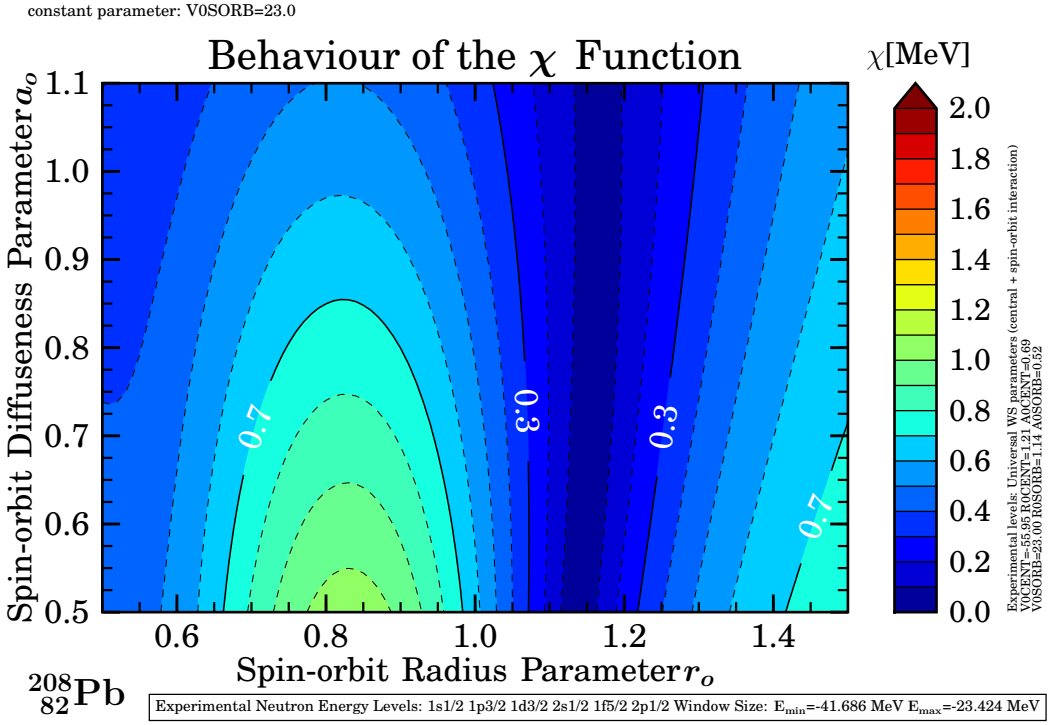

Figure 4.28: *Correlation of parameters of the spin-orbit potential a_{so} vs. r_{so} with energy window far from Fermi level.*

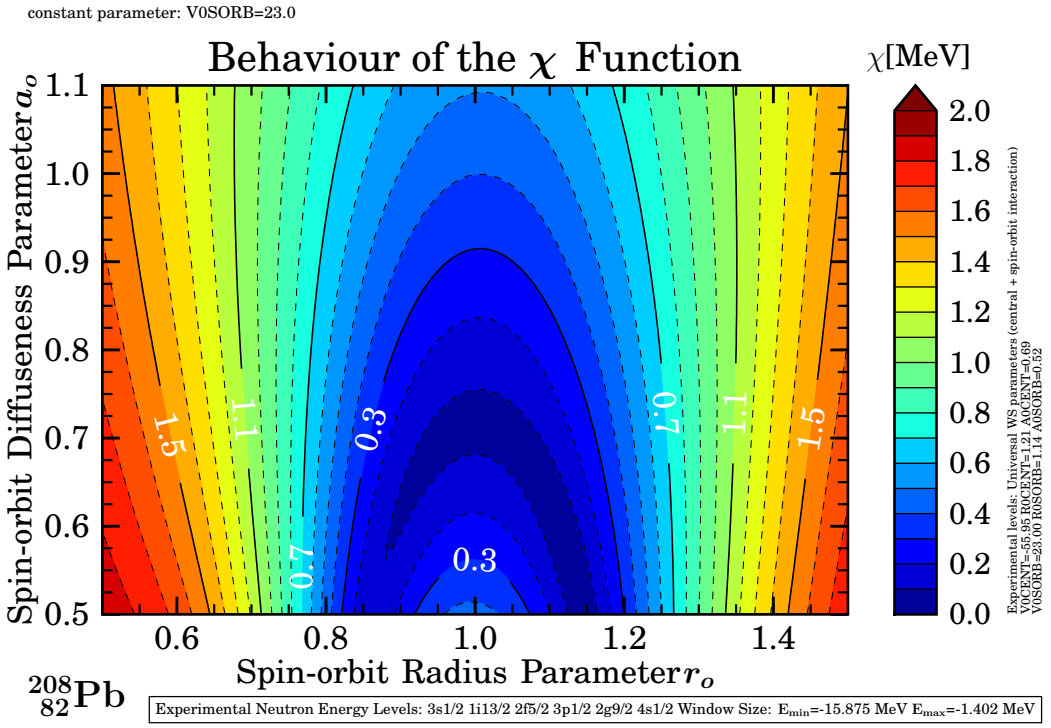

Figure 4.29: *Correlation of parameters of the spin-orbit potential a_{so} vs. r_{so} with energy window close to Fermi level.*

4.5 MonteCarlo Calculations for Spherical Woods-Saxon Hamiltonian

A similar test to the one presented in the previous section was also performed using Monte Carlo techniques [32], [33]. A set of parameters was first obtained through a fitting procedure (with spin-orbit diffuseness fixed at $a_{so} = 0.6$fm) and then used to produce the 'experimental' levels. The positions of the levels were then randomly modified in a set of 10,000 runs with probabilities given by Gaussian distributions. The results of these calculations are presented in figure 4.30. The plot on the left-hand side shows the positions of pairs of points (a_0^c vs. V_0^c), whereas the one on the right-hand side presents the positions of (r_0^c vs. V_0^c).

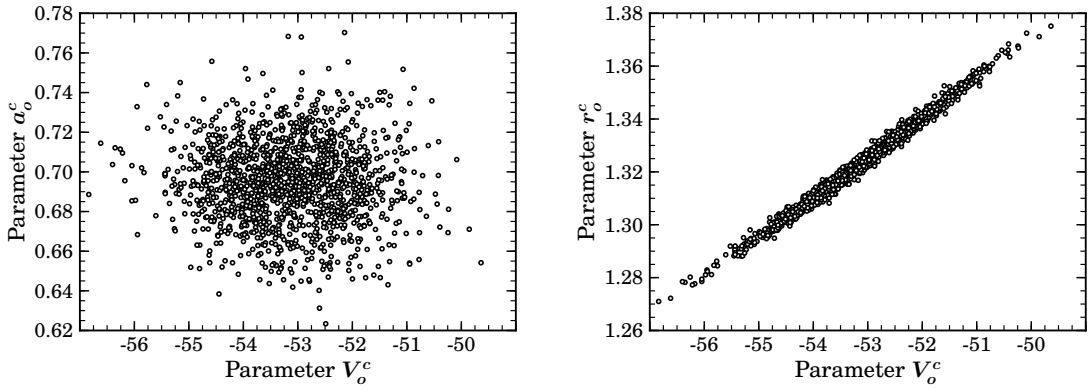

Figure 4.30: *The results of Monte Carlo fitting for ^{208}Pb projected on the (a_0^c vs. V_0^c) plane (left) and (r_0^c vs. V_0^c) plane (right).*

What we can conclude from this figure is the presence of a strong correlation between central parameters r_0^c and V_0^c, and almost no correlation between a_0^c and V_0^c. We can compare this result with a straightforward calculations of the correlation matrix defined as

$$\mathrm{corr}(X,Y) = \frac{\sum_i[(X_i-\bar{X})(Y_i-\bar{Y})]}{\sqrt{\sum_i(X_i-\bar{X})^2}\sqrt{\sum_i(Y_i-\bar{Y})^2}};\quad \bar{X}\equiv\frac{1}{n}\sum_{i=1}^{n}X_i,\ \bar{Y}\equiv\frac{1}{n}\sum_{i=1}^{n}Y_i, \tag{4.7}$$

where X_i and Y_i represnt the parameters of the Hamiltonian. Table 4.8 shows the correlation matrix calculated from the results of Monte Carlo simulation.

Comparing the value corresponding to the pair (V_0^c,r_0^c) with the one for (V_0^c,a_0^c), we clearly see the strong correlation of the first pair, as it was shown in the plot on the right-hand side in Fig. 4.30.

Table 4.8: *The correlation matrix for the parameters of the Woods-Saxon Hamiltonian as obtained from the Monte Carlo simulation.*

	V_0^c	r_0^c	a_0^c	V_0^{so}	r_0^{so}
V_0^c	1.000	0.994	-0.028	0.000	0.265
r_0^c	0.994	1.000	0.016	0.005	0.270
a_0^c	0.028	0.016	1.000	0.259	0.288
V_0^{so}	0.000	0.005	0.259	1.000	0.506
r_0^{so}	0.265	0.270	0.288	0.506	1.000

4.6 Monte Carlo Calculations for Spherical Hartree-Fock Hamiltonian

The same procedure as presented in the previous section was performed using sperically-symmetric Skyrme Hartree-Fock Hamiltonian [32], [33], which is defined as a sum of kinetic-energy term and isoscalar and iso-vector potentials

$$\mathcal{H}(r) = \frac{\hbar^2}{2m}\tau_0 + \mathcal{H}_0(r) + \mathcal{H}_1(r). \tag{4.8}$$

The potential terms are given by

$$\mathcal{H}_t(r) = C_t^{\rho'}\rho_t^2 + C_t^{\Delta\rho}\rho_t\Delta\rho_t + C_t^{\tau}\rho_t\tau_+\frac{1}{2}C_t^J J_t^2 + C_t^{\nabla J}\rho_t\nabla\cdot J_t\,, \tag{4.9}$$

where C_t^{ρ} is defined as

$$C_t^{\rho'} = C_t^{\rho} + C_t^{\rho\alpha}\rho_0^{\alpha}\,. \tag{4.10}$$

Among the twelve parameters, only six was allowed to vary in the fitting procedure, namely $C_0^{\rho}, C_1^{\rho}, C_0^{\rho\alpha}, C_0^{\tau}, C_1^{\tau}, C_0^{\nabla J}$. The remaining six were kept fixed or constrained by relations with other parameters.[3]

The following figures show the results of Monte Carlo simulation for the Skyrme Hartree-Fock Hamiltonian. In this case, we can observe stronger correlations as compared to the Woods-Saxon Hamiltonian. On the left-hand side of Fig. 4.31 we can see the correlation between C_0^{ρ} and C_1^{ρ}, the right-hand side shows $C_0^{\nabla J}$ vs. C_0^{ρ}. Figure 4.32 presents the correlation of C_0^{ρ} and C_0^{τ} (left) and also C_0^{τ} and $C_0^{\rho\alpha}$. For the latter we can also observe a tendency for a double-valued correlation.

It is also instructive to calculate the correlation matrix in this case. As presented in Table 4.9, a major part of the parameters is strongly correlated, not only the ones presented in Fig. 4.31 and 4.32. It can be compared with plots presented in Fig. 4.32[4], where we observe other examples of correlations of parameters. These results suggest that there are three rather than twelve independent parameters within the Syrme Hartree-Fock formalism.

[3] For detailed information, cf. [32], sect. 4.

[4] Plots presented in this figure come from [33].

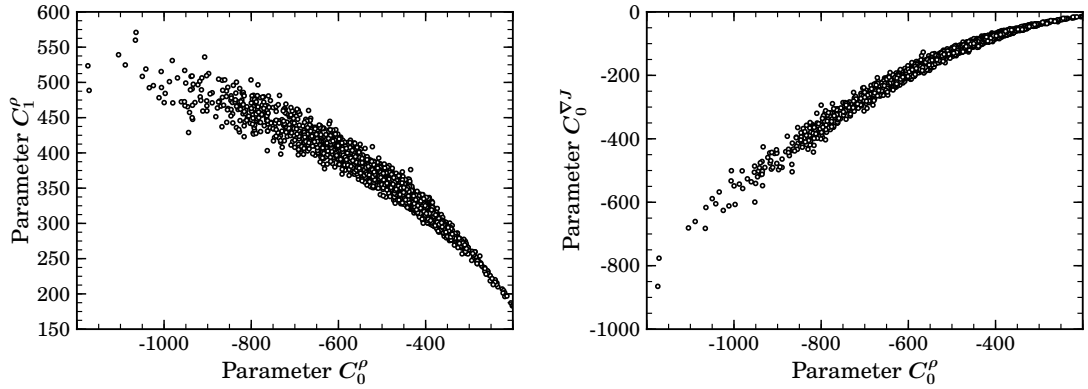

Figure 4.31: *The results of Monte Carlo fitting for ^{208}Pb projected on the (C_1^ρ vs. C_0^ρ) plane (left) and ($C_0^{\nabla J}$ vs. C_0^ρ) plane (right).*

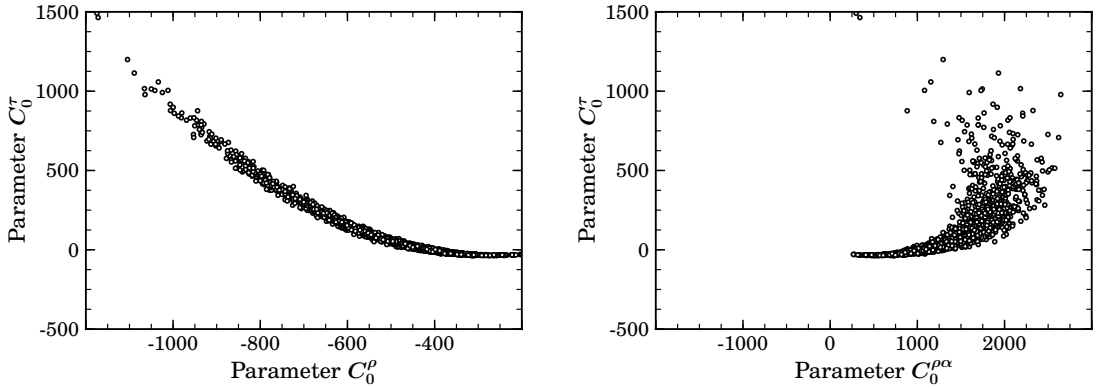

Figure 4.32: *The results of Monte Carlo fitting for ^{208}Pb projected on the (C_0^τ vs. C_0^ρ) plane (left) and (C_0^τ vs. $C_0^{\rho\alpha}$) plane (right).*

Table 4.9: *The correlation matrix for the parameters of the Skyrme-Hartree-Fock Hamiltonian as obtained from the Monte Carlo simulation.*

	C_0^ρ	C_1^ρ	$C_0^{\rho\alpha}$	C_0^τ	C_1^τ	$C_0^{\nabla J}$
C_0^ρ	1.000	-0.948	-0.506	-0.902	0.952	0.965
C_1^ρ	-0.948	1.000	0.682	0.745	-0.838	-0.854
$C_0^{\rho\alpha}$	-0.506	0.682	1.000	0.102	-0.243	-0.290
C_0^τ	-0.902	0.745	0.102	1.000	-0.985	-0.977
C_1^τ	0.952	-0.838	-0.243	-0.985	1.000	0.993
$C_0^{\nabla J}$	0.965	-0.854	-0.290	-0.977	0.993	1.000

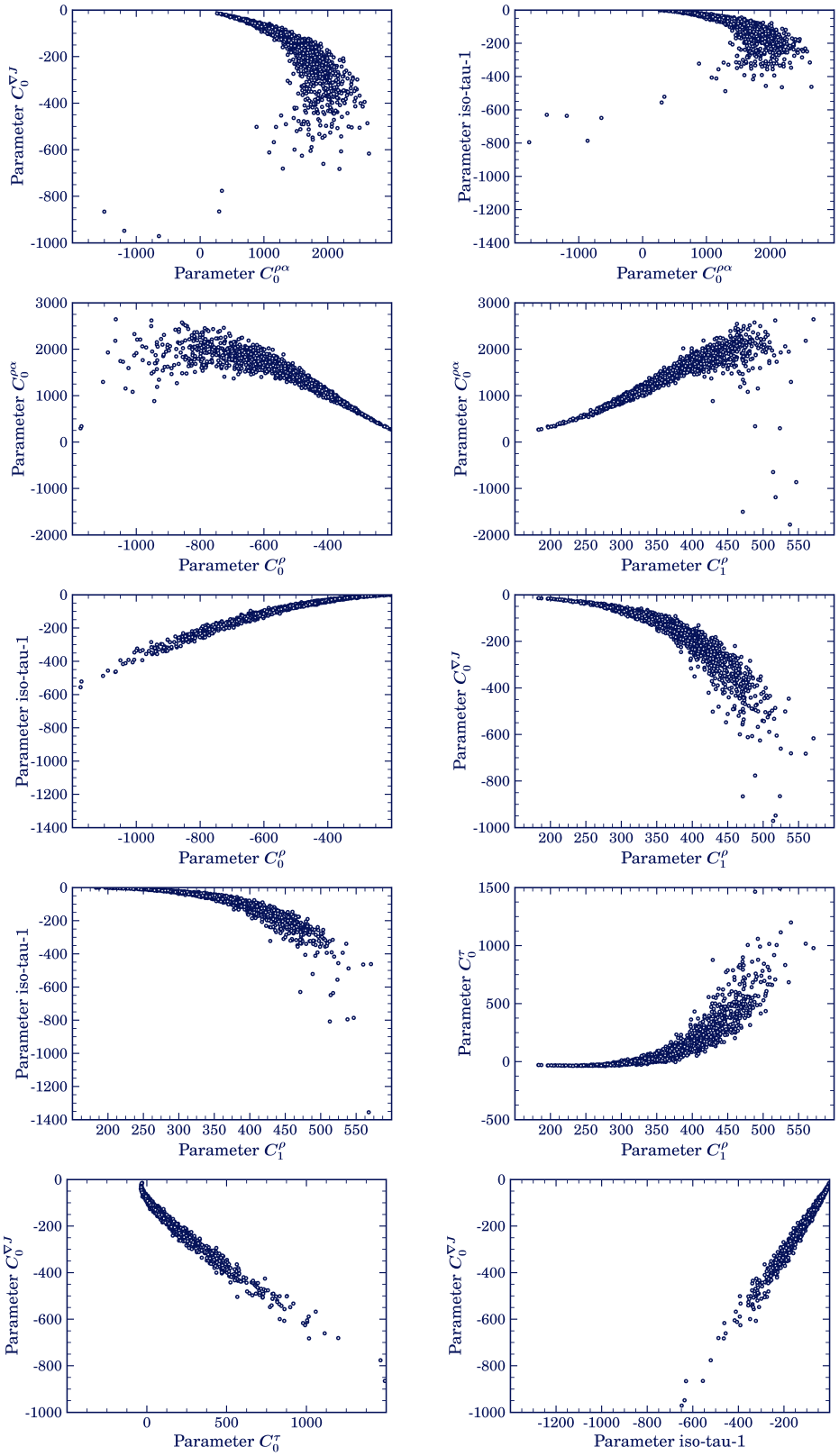

Figure 4.33: *The results of Monte Carlo fitting for ^{208}Pb projected on different planes.*

4.7 Comments on the Skyrme-Hartree-Fock Formalism

Let us recall the formulation of the Skyrme Hartree-Fock approach presented briefly in Sect. 2.3.1, where the Skyrme interaction is given by the following expression

$$\begin{aligned}\hat{v}_{Skyrme}(\vec{r}_i,\vec{r}_j) &= t_0(1+x_0\hat{P}_\sigma)\,\delta(\vec{r}_{ij})\\ &+ \tfrac{1}{2}t_1(1+x_1\hat{P}_\sigma)\left[\vec{k}'^2\delta(\vec{r}_{ij})+\delta(\vec{r}_{ij})\vec{k}^2\right]\\ &+ t_2\,(1+x_2\hat{P}_\sigma)\left[\vec{k}'\right]\cdot\left[\delta(\vec{r}_{ij})\,\vec{k}\right]\\ &+ \tfrac{1}{6}t_3(1+x_3\hat{P}_\sigma)\,\rho^\alpha(\vec{R})\left[\delta(\vec{r}_{ij})\right]\\ &+ iW_0\,(\vec{\sigma}_i+\vec{\sigma}_j)\,\cdot\left[\vec{k}'\times\delta(\vec{r}_{ij})\,\vec{k}\right]\\ &+ v^{tensor}(\vec{r}_i,\vec{r}_j),\end{aligned} \tag{4.11}$$

with the tensor interaction defined as

$$\begin{aligned}v^{tensor}(\vec{r}_i,\vec{r}_j) = \;& \tfrac{1}{2}\,t_e\,\Big\{\,[\,3(\vec{\sigma}_i\cdot\vec{k}')\,(\sigma_j\cdot\vec{k}')-(\vec{\sigma}_i\cdot\vec{\sigma}_j)(\vec{k}')^2\,]\,\delta(\vec{r}_{ij})\\ &+\,\delta(\vec{r}_{12})\,[\,3(\vec{\sigma}_1\cdot\vec{k}\,)\,(\vec{\sigma}_2\cdot\vec{k}\,)-(\vec{\sigma}_1\cdot\vec{\sigma}_2)(\vec{k}\,)^2\,]\Big\}\\ &+ t_o\Big\{3(\sigma_1\cdot\vec{k}')\delta(\vec{r}_{12})(\vec{\sigma}_2\cdot\vec{k})-(\vec{\sigma}_1\cdot\vec{\sigma}_2)\left[\vec{k}''\right]\cdot\left[\delta(\vec{r}_{12})\vec{k}\right]\Big\}.\end{aligned} \tag{4.12}$$

As we have already mentioned, this formulation requires, by definition, a total of twelve parameter, namely

$$\{p\}\stackrel{df}{=}\big\{\{t_0,t_1,t_2,t_3\};\;\{x_0,x_1,x_2,x_3\};\;\{W_0\};\;\{t_e,t_o\};\;\{\alpha\}\,\big\}, \tag{4.13}$$

this only being an expansion of the second order. However, higher orders have also been studied, e.g. cf. [34], where an expansion of the sixth order was presented. Let us briefly disscus their results.

In the presented study, the total energy density is given by

$$\mathcal{H}(\vec{r}\,)=\sum_{\substack{m'I',n'L'v'J'\\ mI,nLvJ,Q}} C^{m'I',n'L'v'J'}_{mI,nLvJ,Q}\times T^{m'I',n'L'v'J'}_{mI,nLvJ,Q}(\vec{r}\,), \tag{4.14}$$

where $C^{m'I',n'L'v'J'}_{mI,nLvJ,Q}$ denotes the coupling constants. This means, that the total number of terms (and also coupling constants) grows to a very large number. An evaluation of this number is presented in Table 4.10. The number of terms for time-even and time-odd densities are given in column 1 and 2 respectively, whereas the last two columns contain the numbers of terms with Galilean and gauge[5] invariance.

[5]For more detailed information about Skyrme HF gauge invariance cf. e.g. [35]

Table 4.10: *The number of terms in Skyrme Hartree-Fock depending on the order of expansion. The multiplication by a factor of two take into account both isospin channels.*

Order	T-even	T-odd	Total	Galilean	Gauge
0	1	1	2	2	2
2	8	10	18	12	12
4	53	61	114	45	29
6	250	274	524	129	54
N^3LO	2x312	2x346	2x658	2x188	2x97
	624	692	1316	376	194

The number of terms as compared to the 'ordinary', second order Skyrme Hartree-Fock formulation is enormous. If we consider this as an Inverse Problem, as it should be treated, we can clearly see the difficulties arising from such a large number of parameters. The stability of this type of fitting procedure is questionable, not to mention its Predictive Power.

From the point of view of our work, expanding the Skyrme-Hartree-Fock formalism into higher orders like 'next-to-next-to-to-next' order disscused in this section based on [34] is rather pointless. It produces hundreds, or even thousands of coupling constants to be fitted, while for the ordinary Hartree-Fock formulation there are twelve coupling constants and at least nine parametric correlations among them, which leaves us with only *three* independent parameters.

Conclusions

In the present work we have examined the functioning of the modelling of selected nuclear observables from the point of view of the predictive power and the inverse problem – the mathematical theory underlying the modern stochastic theories of the predictive power. The importance of the subject can be summarized as follows.

- In the coming years the uncertainties of the theoretical predictions will be a required condition for publishing the articles in various journals.
- Our work shows that the number of parameters often exceeds the mathematical needs of the modelling, e.g. the standard spherical Woods-Saxon single-particle potential.
- This manifests itself in the ill-posedeness of the inverse problem what causes in turn the instability in terms of the extraneous predictions.
- One way out is to reanalyze the problem and decrease the number of parameters, as we have shown in the case of spherical Woods-Saxon potential – by introducing a density dependent spin-orbit interaction.
- Another, more 'automatic' option is to apply the regularisation methods such as e.g. the Singular Value Decomposition or Tikhonov regularisation.
- The inclusion of experimental uncertainties in fitting procedures should become a crucial point in all theoretical calculations.

Our work presents and illustrates certain first steps and concepts whose more systematic use is already becoming a reality in the 'routine work' of various groups of physicists.

Publications

1. H. Molique, J. Dudek, K. Rybak, M.-G. Porquet, *A New Mean Field Approach for Exotic Nuclei*, Acta Physica Polonica B, Vol. 40, No. 3, March 2009 (597)

2. J. Dudek, K. Rybak, B. Szpak, M.-G. Porquet, H. Molique and B. Fornal, *Nuclear Mean-Field Hamiltonians and Factors Limiting Their Predictive Power*, Int. J. Mod. Phys. E, vol. 19, No. 4 (2010), 652-664

3. B. Szpak, J. Dudek, M.-G. Porquet, K. Rybak, H. Molique and B. Fornal, *Nuclear Mean-Field Hamiltonians and Factors Limiting Their Spectroscopic Predictive Power: Illustrations*, Int. J. Mod. Phys. E, vol. 19, No. 4 (2010), 665-671

4. J. Dudek, B. Szpak, M.-G. Porquet, H. Molique, K. Rybak and B. Fornal, *Nuclear Hamiltonians: the question of their spectral predictive power and the associated inverse problem*, J. Phys. G: Nucl. Part. Phys., 37 (2010), 064031

Bibliography

[1] P. Ring, P. Schuck, *The Nuclear Many-Body Problem*, Springer-Verlag, New York 1980

[2] A. Bohr, B. R. Mottelson, *Nuclear Structure*, vol. I, Benjamin, New York 1969

[3] W. Greiner, J.A. Maruhn, *Nuclear Models*, Springer-Verlag, Berlin Heidelberg, 1996

[4] R.C. Aster, B. Borchers and C.H. Thurber, *Parameter Estimation and Inverse Problems*, Elsevier Academic Press, 2005

[5] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, USA, 2005

[6] *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, Editors: Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst, SIAM, 2000

[7] G.H. Golub, C.F van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1996

[8] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, 'Singular Value Decomposition.' 2.6 in *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Cambridge University Press, pp. 51-63, 1992.

[9] C. R. Vogel, *Computational Methods for Inverse Problems*, Society for Industrial and Applied Mathematics, Philadelphia, USA, 2002

[10] D. A. Varshalovich, A. N. Moskalev, V. K. Khersonskii, *Quantum Theory of Angular Momentum*, World Scientific, 1988

[11] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Dover Publications, Inc., New York, 1965

[12] N. Schunck, J. Dudek, A. Góźdź, P.H. Regan, *Tetrahedral symmetry in ground and low-lying states of exotic A∼110 nuclei*, Phys. Rev. C 69, 061305-1, (2004)

[13] J. Dudek, J. Dobaczewski, N. Dubray, A. Góźdź, V. Pangon, N. Schunck, *Nuclei with Tetrahedral Symmetry*, Int. J. of Modern Physics (2006)

[14] J. Dudek, D. Curien, N. Dubray, J. Dobaczewski, V. Pangon, P. Olbratowski, N. Schunck, *Island of Rare Earth Nuclei with Tetrahedral and Octahedral Symmetries: Possible Experimental Evidence*, Phys. Rev. Lett. 97, 072501-1 (2006)

[15] J. Dudek, A. Góźdź, D. Curien, V. Pangon and N. Schunck, *Nuclear Tetrahedral Symmetry and Collective Rotation*, Acta Physica Polonica B, Vol. 38, No. 4, April 2007 (1389)

[16] T.H.R. Skyrme, *The effective Nuclear Potential*, Nuc. Phys. 9, 615-634 (1959)

[17] D. Vautherin, D.M. Brink, *Hartree-Fock Calculations with Skyrme's Interaction. I. Spherical Nuclei*, Phys. Rev. **C5** 626-647 (1972)

[18] D.M. Brink, Fl. Stancu, *Evolution of nuclear shells with the Skyrme density dependent interaction*, Phys. Rev. C, vol. 75, 064311 (2007)

[19] H. Molique, J. Dudek, K. Rybak, M.-G. Porquet, *A New Mean Field Approach for Exotic Nuclei*, Acta Physica Polonica B, Vol. 40, No. 3, March 2009 (597)

[20] J. Dudek, K. Rybak, B. Szpak, M.-G. Porquet, H. Molique and B. Fornal, *Nuclear Mean-Field Hamiltonians and Factors Limiting Their Predictive Power*, Int. J. Mod. Phys. E, vol. 19, No. 4 (2010), 652-664

[21] B. Szpak, J. Dudek, M.-G. Porquet, K. Rybak, H. Molique and B. Fornal, *Nuclear Mean-Field Hamiltonians and Factors Limiting Their Spectroscopic Predictive Power: Illustrations*, Int. J. Mod. Phys. E, vol. 19, No. 4 (2010), 665-671

[22] J. Dudek, B. Szpak, M.-G. Porquet, H. Molique, K. Rybak and B. Fornal, *Nuclear Hamiltonians: the question of their spectral predictive power and the associated inverse problem*, J. Phys. G: Nucl. Part. Phys., 37 (2010), 064031

[23] A.M. Oros, Ph.D. thesis, University of Köln, 1996

[24] M.-G. Porquet, private communication.

[25] T. Otsuka, T. Suzuki, R. Fujimoto, H. Grawe, Y. Akaishi, *Evolution of Nuclear Shells due to the Tensor Force*, Phys. Rev. Lett. 95, (2005), 232502

[26] M. Baranger,*A definition of the single-nucleon potential*, Nucl. Phys. A 149, 225-240 (1970)

[27] S. Galès, C.P. Massolo, S. Fortier, J.P. Schapira, P. Martin, and V. Comparat, *Proton stripping to outer subshells and the damping of single-particle states:* ^{116}Sn, ^{144}Sm, *and* $^{208}(\alpha,t)$ *reactions at 80 MeV*, Phys. Rev. C 31, 94-110 (1985)

[28] H. Langevin-Joliot, J. van de Wiele, J. Guillot, E. Gerlic, L.H. Rosier, A. Willis, M. Morlet, G. Duhamel-Chretien, E. Tomasi-Gustafsson, N. Blasi, S. Micheletti, and S.Y. van der Werf, *Spin determination of valence and inner hole states via the* $^{208}Pb(d\rightarrow,t)^{207}Pb$ *reaction at* E_d*=200 MeV*, Phys. Rev. C 47, 15711585 (1993)

[29] S. Galès, G.M. Crawley, D. Weber, and B. Zwieglinki, *Inner hole states in* ^{207}Pb *via the* $^{208}Pb(^3He, \alpha)\ ^{207}Pb$ *reaction at 70 MeV*, Phys. Rev. C 18, 24752485 (1978)

[30] R. Tickle and W.S. Gray, *The* $^{208}Pb(\alpha,\ ^3He)^{209}Pb$ *reaction at 58 MeV* , Nucl. Phys. A, vol. 247, 187-194 (1975)

[31] P. Grabmayr, A. Mondry, G.J. Wagner, P. Woldt, G.P.A. Berg, J. Lisantti, D.W. Miller, H. Nann, P.P. Singh and E.J. Stephenson, *Search for $3s_{1/2}$ hole-strength fragments via the $^{208}Pb(d, ^{3}He)^{207}Tl$ reaction*, J. Phys. G: Nucl. Part. Phys. 18, 1753 (1992)

[32] J. Dudek, B. Szpak, A. Dromard, M.-G. Porquet, B. Fornal, A. Góźdź, *Nuclear Physics Hamiltonians, Inverse Problem and the Related Issue of Predictive Power*, Int. J. Mod. Phys. E, vol. 21, No. 5 (2012), 1250053

[33] B. Szpak et al., private communication.

[34] B. G. Carlsson, J. Dobaczewski, and M. Kortelainen, *Local nuclear energy density functional at next-to-next-to-next-to-leading order*, Phys. Rev. C 78, 044326 (2008)

[35] J. Dobaczewski and J. Dudek, *Time-odd components in the mean field of rotating superdeformed nuclei*, Phys. Rev. C 52, 1827 (1995)