Predictive Power of Mathematical Modelling for Nuclear Physics

J. DUDEK

Department of Subatomic Research, CNRS/IN_2P_3
and
University of Strasbourg, F-67037 Strasbourg, FRANCE

September 26, 2012

COLLABORATORS:

Bartłomiej SZPAK, IFJ Kraków

Arthur DROMARD, UdS and IPHC, Strasbourg

Andrzej GOZDZ, UMCS Lublin

Karolina RYBAK, UdS and IPHC, Strasbourg

Helena SLIWINSKA, UdS and IPHC, Strasbourg

Hervé MOLIQUE, UdS and IPHC, Strasbourg

Marie-Geneviève PORQUET, CSNSM Orsay

Bogdan FORNAL, IFJ Kraków

A review and a short introduction can be found in:

1) Nuclear Hamiltonians: The Question of their Spectral Predictive Power and the Associated Inverse Problem;

JD, B. Szpak, M-G, Porquet, H. Molique, K. Rybak, B. Fornal
J. Phys. G: Nucl. Part. Phys. **37** (2010) 064031

FOCUS Special Issue: Open Problems in Nuclear Theory

2) Nuclear Mean Field Hamiltonians and Factors Limiting their Predictive Power: Formalism;

JD, K. Rybak, B. Szpak, M-G, Porquet, H. Molique & B. Fornal
Int. J. Mod. Phys. E **19** (2010) 652

3) Statistical significance of theoretical predictions: A new dimension in nuclear structure theories (I);

J. Dudek, B. Szpak, M.-G. Porquet and B. Fornal
Journal of Physics: Conference Series, **267** *(2011) 012062*

4) Statistical significance of theoretical predictions: A new dimension in nuclear structure theories (II);

B. Szpak, J. Dudek, M.-G. Porquet and B. Fornal
Journal of Physics: Conference Series, **267** *(2011) 012063*

5) Nuclear Physics Hamiltonians, Inverse Problem and the Related Issue of Predictive Power;

JD, B. Szpak, A. Dromard, M.-G. Porquet, B. Fornal and A. Góźdź
Int. J. Mod. Phys. **E 21**, *No. 5 (2012) 1250053*

Part I

Nuclear Hamiltonians and Nuclear Theories: Predictive-Power Perspective

Predictive Power... The Issue of the Very Definition

Predictive Power... The Issue of the Very Definition

- **Discussing the problem of Predictive Power, we usually have no doubt in our minds so as to what is meant (... or what we mean...)**

Predictive Power... The Issue of the Very Definition

• **Discussing the problem of Predictive Power, we usually have no doubt in our minds so as to what is meant (... or what we mean...)**

• **Suppose somebody has obtained a modelling result before any experimental verification - Such a result can be called a prediction!**

Predictive Power... The Issue of the Very Definition

• Discussing the problem of Predictive Power, we usually have no doubt in our minds so as to what is meant (... or what we mean...)

• Suppose somebody has obtained a modelling result before any experimental verification - Such a result can be called a prediction!

• After performing the experiment we verify, ex post, whether this prediction was good and claim victory and (good) predictive power!

Predictive Power... The Issue of the Very Definition

• Discussing the problem of Predictive Power, we usually have no doubt in our minds so as to what is meant (... or what we mean...)

• Suppose somebody has obtained a modelling result before any experimental verification - Such a result can be called a prediction!

• After performing the experiment we verify, ex post, whether this prediction was good and claim victory and (good) predictive power!

• At this moment "theory predictions" turn into "modelling result" of the experiment - without anybody doing anything on theory side

Predictive Power... The Issue of the Very Definition

• Discussing the problem of Predictive Power, we usually have no doubt in our minds so as to what is meant (... or what we mean...)

• Suppose somebody has obtained a modelling result before any experimental verification - Such a result can be called a prediction!

• After performing the experiment we verify, ex post, whether this prediction was good and claim victory and (good) predictive power!

• At this moment "theory predictions" turn into "modelling result" of the experiment - without anybody doing anything on theory side

• At this point - what begins - are the issues of lacking precision in very posing of the problem, arbitrariness and semantical confusion, the implied questions, troubles, possibly mathematical non-sense...

What does it mean: Having Predictive Power?

What does it mean: Having Predictive Power?

- **Any result of any modelling of any phenomenon before the right experiments are performed can be called "a theoretical prediction"**

What does it mean: Having Predictive Power?

- **Any result of any modelling of any phenomenon before the right experiments are performed can be called "a theoretical prediction"**
- **... thus performing any model calculation can be called predicting!**

What does it mean: Having Predictive Power?

• Any result of any modelling of any phenomenon before the right experiments are performed can be called “a theoretical prediction”

• ... thus performing any model calculation can be called predicting!

• But if any calculation is a prediction, any theory can predict always and therefore it has always a predictive power ‘no-matter-what’ ...

What does it mean: Having Predictive Power?

• Any result of any modelling of any phenomenon before the right experiments are performed can be called "a theoretical prediction"

• ... thus performing any model calculation can be called predicting!

• But if any calculation is a prediction, any theory can predict always and therefore it has always a predictive power 'no-matter-what' ...

• As a consequence, the very term "predictive power" applies always i.e. means no special property. In our context it will be fair to say: **This term is void of sense - more precisely: does not tell us anything**

What does it mean: Having Predictive Power?

- Any result of any modelling of any phenomenon before the right experiments are performed can be called "a theoretical prediction"

- ... thus performing any model calculation can be called predicting!

- But if any calculation is a prediction, any theory can predict always and therefore it has always a predictive power 'no-matter-what' ...

- As a consequence, the very term "predictive power" applies always i.e. means no special property. In our context it will be fair to say: This term is void of sense - more precisely: does not tell us anything

- ...and one may try using similar, a slightly modified wording: What carries certain interest is, possibly, theory's good predictive power!

What does it mean: Prediction is Good (or Satisfactory)?

What does it mean: Prediction is Good (or Satisfactory)?

- **Being good for someone may not be satisfactory for someone else**

What does it mean: Prediction is Good (or Satisfactory)?

- **Being good for someone may not be satisfactory for someone else**

- **... and it becomes clear that discussions of this type unavoidably involve the elements of arbitrariness and of a subjective judgement**

What does it mean: Prediction is Good (or Satisfactory)?

- Being good for someone may not be satisfactory for someone else
- ... and it becomes clear that discussions of this type unavoidably involve the elements of arbitrariness and of a **subjective judgement**
- Therefore directly related with the notion of “good predictions” are, sine qua non, criteria of distinction between “good” and “poor”

What does it mean: Prediction is Good (or Satisfactory)?

- Being good for someone may not be satisfactory for someone else
- ... and it becomes clear that discussions of this type unavoidably involve the elements of arbitrariness and of a subjective judgement
- Therefore directly related with the notion of "good predictions" are, sine qua non, criteria of distinction between "good" and "poor"
- It is not possible to talk about Predictive Power [whatever it means*)] without specifying the criteria of choice at the same time:

The notion of Predictive Power is relative and/or subjective#)

What does it mean: Prediction is Good (or Satisfactory)?

- Being good for someone may not be satisfactory for someone else
- ... and it becomes clear that discussions of this type unavoidably involve the elements of arbitrariness and of a **subjective judgement**
- Therefore directly related with the notion of "good predictions" are, sine qua non, criteria of distinction between "good" and "poor"
- It is not possible to talk about Predictive Power [whatever it means*)] without specifying the criteria of choice at the same time:

The notion of Predictive Power is relative and/or subjective#)

*) *This notion is still to be defined for you here ...*

#) *So is the very notion of probability (12 'official' definitions and 16 interpretations)*

Theory-Errors Limit Theory's Predictive Power [1]

Theory-Errors Limit Theory's Predictive Power [1]

- What do we usually wish to do is to learn the so-called 'full truth'

$$\boxed{\hat{\mathsf{H}}^{\mathsf{true}}\,\psi_{\mathsf{n}} = \mathsf{e}_{\mathsf{n}}\,\psi_{\mathsf{n}} \quad - \quad \text{where } \underline{\underline{\text{we wish}}} \text{ to know: } \; \hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}}}$$

Theory-Errors Limit Theory's Predictive Power [1]

- What do we usually wish to do is to learn the so-called 'full truth'

$$\hat{\mathsf{H}}^{\mathsf{true}}\,\psi_{\mathsf{n}} = \mathsf{e}_{\mathsf{n}}\,\psi_{\mathsf{n}} \quad - \quad \text{where } \underline{\underline{\text{we wish}}} \text{ to know: } \; \hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}}$$

- The way humans do their research can be summarised like this:

$$\hat{\mathsf{H}} = \underbrace{\hat{\mathsf{H}}_{\mathsf{1}}}_{\text{of 1949}} + \underbrace{\hat{\mathsf{H}}_{\mathsf{2}}}_{\text{of 1964}} + \underbrace{\hat{\mathsf{H}}_{\mathsf{3}}}_{\text{of 2012}} + \cdots + \underbrace{\hat{\mathsf{H}}_{\mathsf{n}}}_{\text{of 2055}} + \cdots \rightarrow \underbrace{\hat{\mathsf{H}}^{\mathsf{true}}}_{\text{say:}\infty}$$

In other words: Human quantum theories are usually incomplete

Theory-Errors Limit Theory's Predictive Power [1]

- What do we usually wish to do is to learn the so-called 'full truth'

$$\boxed{\hat{\mathsf{H}}^{\mathsf{true}}\,\psi_{\mathsf{n}} = \mathsf{e}_{\mathsf{n}}\,\psi_{\mathsf{n}} \quad - \quad \text{where } \underline{\underline{\text{we wish}}} \text{ to know: } \; \hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}}}$$

- The way humans do their research can be summarised like this:

$$\hat{\mathsf{H}} = \underbrace{\hat{\mathsf{H}}_1}_{\text{of } 1949} + \underbrace{\hat{\mathsf{H}}_2}_{\text{of } 1964} + \underbrace{\hat{\mathsf{H}}_3}_{\text{of } 2012} + \cdots + \underbrace{\hat{\mathsf{H}}_{\mathsf{n}}}_{\text{of } 2055} + \cdots \;\rightarrow\; \underbrace{\hat{\mathsf{H}}^{\mathsf{true}}}_{\text{say:}\infty}$$

In other words: Human quantum theories are usually incomplete

- Our Hamiltonians have always a structure: $\underline{\underline{\hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}} + \delta\hat{\mathsf{H}}^{\mathsf{ignor}}}}$

Theory-Errors Limit Theory's Predictive Power [1]

- What do we usually wish to do is to learn the so-called 'full truth'

$$\boxed{\hat{\mathbf{H}}^{\mathbf{true}}\,\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \quad - \quad \text{where } \underline{\underline{\text{we wish}}} \text{ to know: } \quad \hat{\mathbf{H}} = \hat{\mathbf{H}}^{\mathbf{true}}}$$

- The way humans do their research can be summarised like this:

$$\hat{\mathbf{H}} = \underbrace{\hat{\mathbf{H}}_1}_{\text{of 1949}} + \underbrace{\hat{\mathbf{H}}_2}_{\text{of 1964}} + \underbrace{\hat{\mathbf{H}}_3}_{\text{of 2012}} + \cdots + \underbrace{\hat{\mathbf{H}}_n}_{\text{of 2055}} + \cdots \rightarrow \underbrace{\hat{\mathbf{H}}^{\mathbf{true}}}_{\text{say:}\infty}$$

In other words: Human quantum theories are usually incomplete

- Our Hamiltonians have always a structure: $\hat{\mathbf{H}} = \hat{\mathbf{H}}^{\mathbf{true}} + \delta\hat{\mathbf{H}}^{\mathbf{ignor}}$

- **Conclusion:** The desired truth remains unknown to us because of $\delta\hat{H}^{ignor} \rightarrow$ ignorance decreasing with research time

Theory-Errors Limit Theory's Predictive Power [2]

- **There exist effective means of limiting the negative impact of the ignorance originating from incompleteness of information**

Theory-Errors Limit Theory's Predictive Power [2]

- **There exist effective means of limiting the negative impact of the ignorance originating from incompleteness of information**
- **Solution: In cases of doubt we 'parametrize our ignorance'**

Theory-Errors Limit Theory's Predictive Power [2]

• **There exist effective means of limiting the negative impact of the ignorance originating from incompleteness of information**

• **Solution: In cases of doubt we 'parametrize our ignorance'**

• **In other words: Estimate which answer will be more,- and which less-likely 'the right solution'. Expressed alternatively:**

Theory-Errors Limit Theory's Predictive Power [2]

• **There exist effective means of limiting the negative impact of the ignorance originating from incompleteness of information**

• **Solution: In cases of doubt we 'parametrize our ignorance'**

• **In other words: Estimate which answer will be more,- and which less-likely 'the right solution'. Expressed alternatively: Find relative probability of what we think the right answer is!**

• **There exist effective means of limiting the negative impact of the ignorance originating from incompleteness of information**

• **Solution: In cases of doubt we 'parametrize our ignorance'**

• **In other words: Estimate which answer will be more,- and which less-likely 'the right solution'. Expressed alternatively: Find relative probability of what we think the right answer is!**

In Applied Mathematics:

1. Our ignorance is usually represented by a random variable X

• There exist effective means of limiting the negative impact of the ignorance originating from incompleteness of information

• Solution: In cases of doubt we 'parametrize our ignorance'

• In other words: Estimate which answer will be more,- and which less-likely 'the right solution'. Expressed alternatively: Find relative probability of what we think the right answer is!

In Applied Mathematics:

1. Our ignorance is usually represented by a random variable X

2. Mathematically, variable X is represented by a probability distribution $P_X = P_X(x)$, x called 'realisation' of the variable X

Theory-Errors Limit Theory's Predictive Power [2]

• There exist effective means of limiting the negative impact of the ignorance originating from incompleteness of information

• Solution: In cases of doubt we 'parametrize our ignorance'

• In other words: Estimate which answer will be more,- and which less-likely 'the right solution'. Expressed alternatively: Find relative probability of what we think the right answer is!

In Applied Mathematics:

1. Our ignorance is usually represented by a random variable X
2. Mathematically, variable X is represented by a probability distribution $P_X = P_X(x)$, x called 'realisation' of the variable X

Conclusion: Not knowing 'the truth' we may introduce several competing hypotheses & calculate their relative probabilities!

This is a new strategical observation which introduces what we call "Stochastic Theory of Predictive Power"

This is a new strategical observation which introduces what we call "Stochastic Theory of Predictive Power"

Given theory $\mathcal{T}$ of a phenomenon $\mathcal{P}$ generating observables $\hat{\mathcal{F}}_1$, $\hat{\mathcal{F}}_2$, . . . $\hat{\mathcal{F}}_p$.

This is a new strategical observation which introduces what we call
"Stochastic Theory of Predictive Power"

Given theory $\mathcal{T}$ of a phenomenon $\mathcal{P}$ generating observables $\hat{\mathcal{F}}_1$, $\hat{\mathcal{F}}_2$, . . . $\hat{\mathcal{F}}_p$.

These observables are characterized not only by the eigenvalues $\{\hat{\mathcal{F}}_1 : \mathbf{f_1}, \hat{\mathcal{F}}_2 : \mathbf{f_2}, \ldots\ \hat{\mathcal{F}}_p : \mathbf{f_p}\}$

This is a new strategical observation which introduces what we call
"Stochastic Theory of Predictive Power"

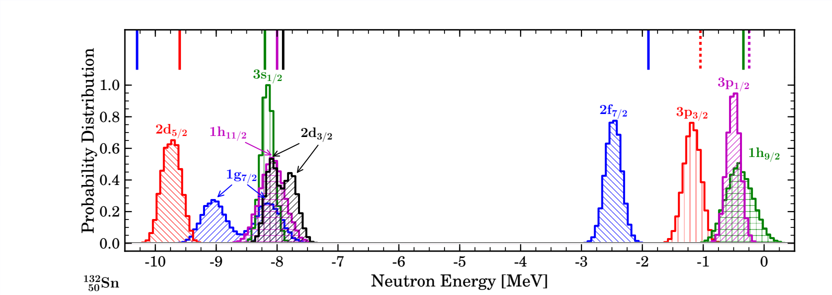
Given theory $\mathcal{T}$ of a phenomenon $\mathcal{P}$ generating observables $\hat{\mathcal{F}}_1$, $\hat{\mathcal{F}}_2$, ... $\hat{\mathcal{F}}_p$.

These observables are characterized not only by the eigenvalues $\{\hat{\mathcal{F}}_1 : f_1, \hat{\mathcal{F}}_2 : f_2, \ldots \hat{\mathcal{F}}_p : f_p\}$

but also by their probability distributions:

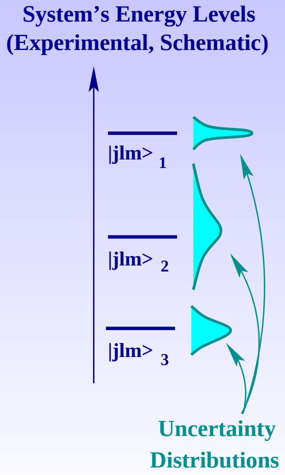
$$P_1 = P_1(f_1), \;\; P_2 = P_2(f_2), \; \ldots \; P_p = P_1(f_p)$$

Inverse Problem and Predictive Power: ^{132}Sn



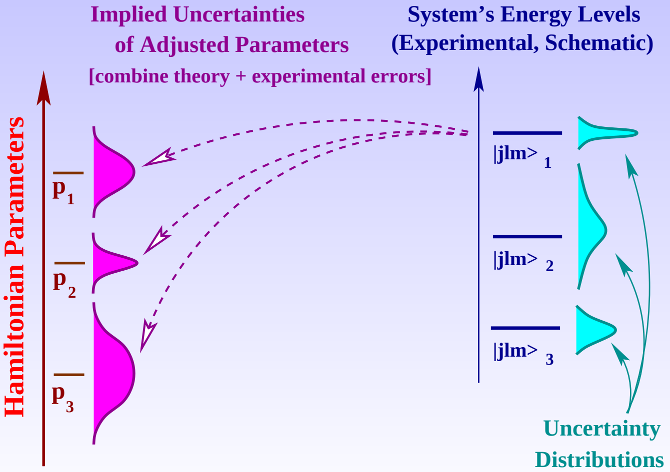
Results of the extrapolation from the ^{208}Pb to the ^{132}Sn nucleus for the neutrons, bars - cf. preceding table. Monte-Carlo simulation with N=20 000 Gaussian-distributed parameter sets, based on ^{208}Pb results; noise width σ=0.1MeV. With each of the so obtained N=20 000 sets of parameters the results for the neutrons in ^{132}Sn nucleus have been obtained. Observe 'pathologies': $1g_{7/2}$ and $1f_{7/2}$ cf. following figures.

Experimental levels represent, from both quantum-mechanical and experimental points of view an ensemble of probability distributions



Energy-Levels as Probability Distributions

The biggest uncertainties of Hamiltonian Parameters originate not so much from the experimental but rather from the theory uncertainties



Combining Theoretical and Experimental Errors

Stochastic Nature of Theoretical Predictions

- Theories are incomplete whereas experiments plagued with errors:

$$\text{Theo.} \rightarrow \boxed{e_n = e_n^{true}(p) + \delta e_n^{error}} \quad \& \quad \boxed{\varepsilon_n = \varepsilon_n^{true} + \delta\varepsilon_n^{err}} \leftarrow \text{Exp.}$$

e_n and ε_n are random variables $\rightarrow$ distributions $P_n^{th.}(e_n)$ and $P_n^{exp}(\varepsilon_n)$

Stochastic Nature of Theoretical Predictions

- Theories are incomplete whereas experiments plagued with errors:

$$\text{Theo.} \rightarrow \boxed{e_n = e_n^{true}(p) + \delta e_n^{error}} \quad \& \quad \boxed{\varepsilon_n = \varepsilon_n^{true} + \delta\varepsilon_n^{err}} \leftarrow \text{Exp.}$$

e_n and ε_n are random variables $\rightarrow$ distributions $P_n^{th.}(e_n)$ and $P_n^{exp}(\varepsilon_n)$

- Errors propagate to the theory predictions through parameter fits

$$\chi^2(p) \sim \sum w_n \Big[\underbrace{\left(\varepsilon_n^{true} + \delta\varepsilon_n^{err}\right)}_{\textbf{Experiment}} - \underbrace{\left(e_n^{true} + \delta e_n^{err}\right)}_{\textbf{Theory}} \Big]^2 \rightarrow \frac{\partial \chi^2}{\partial p} = 0$$

thus the optimal parameter values $p \equiv \{p_1, p_2, \ldots\ p_f\}$ are random variables and consequently characterised by probability distributions

Stochastic Nature of Theoretical Predictions

- Theories are incomplete whereas experiments plagued with errors:

$$\text{Theo.} \rightarrow \boxed{e_n = e_n^{true}(p) + \delta e_n^{error}} \quad \& \quad \boxed{\varepsilon_n = \varepsilon_n^{true} + \delta\varepsilon_n^{err}} \leftarrow \text{Exp.}$$

e_n and ε_n are <u>random variables</u> $\rightarrow$ distributions $P_n^{th.}(e_n)$ and $P_n^{exp}(\varepsilon_n)$

- Errors propagate to the theory predictions through parameter fits

$$\chi^2(p) \sim \sum w_n \Big[\underbrace{\big(\varepsilon_n^{true} + \delta\varepsilon_n^{err}\big)}_{\textbf{Experiment}} - \underbrace{\big(e_n^{true} + \delta e_n^{err}\big)}_{\textbf{Theory}} \Big]^2 \rightarrow \frac{\partial \chi^2}{\partial p} = 0$$

thus the optimal parameter values $p \equiv \{p_1, p_2, \ldots\ p_f\}$ are random variables and consequently characterised by probability distributions

$$\boxed{P(p) = \{P^{th}(e) \ast\ P^{exp}(\varepsilon)\}}$$

Stochastic Nature of Theoretical Predictions

- Theories are incomplete whereas experiments plagued with errors:

$$\text{Theo.} \rightarrow \boxed{e_n = e_n^{true}(p) + \delta e_n^{error}} \quad \& \quad \boxed{\varepsilon_n = \varepsilon_n^{true} + \delta\varepsilon_n^{err}} \leftarrow \text{Exp.}$$

e_n and ε_n are random variables $\rightarrow$ distributions $P_n^{th.}(e_n)$ and $P_n^{exp}(\varepsilon_n)$

- Errors propagate to the theory predictions through parameter fits

$$\chi^2(p) \sim \sum w_n \Big[\underbrace{\left(\varepsilon_n^{true} + \delta\varepsilon_n^{err}\right)}_{\textbf{Experiment}} - \underbrace{\left(e_n^{true} + \delta e_n^{err}\right)}_{\textbf{Theory}} \Big]^2 \rightarrow \frac{\partial \chi^2}{\partial p} = 0$$

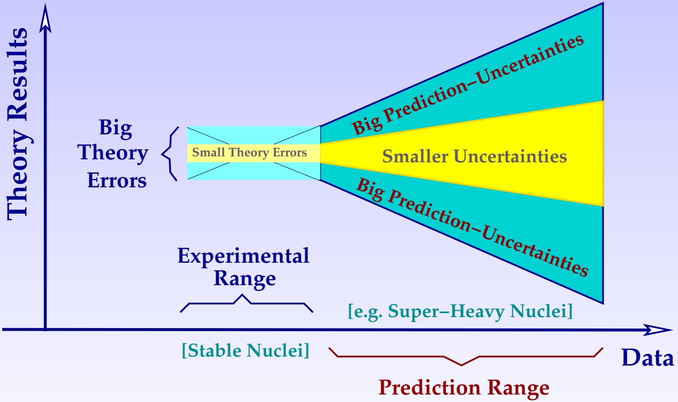
thus the optimal parameter values $p \equiv \{p_1, p_2, \dots\ p_f\}$ are random variables and consequently characterised by probability distributions

$$\boxed{P(p) = \{P^{th}(e) \ast P^{exp}(\varepsilon)\}}$$

- Conclusion: All predictions have their probability distributions!

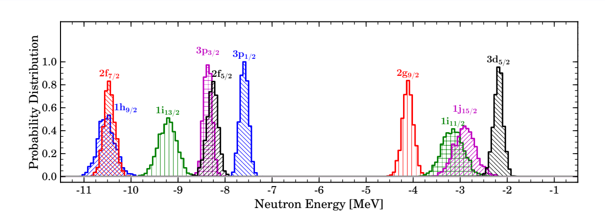
Smaller Theory Errors vs. Bigger Predictive-Power

- Constraining theory errors may help stabilising theory predictions: The necessary although not sufficient condition of model's stability

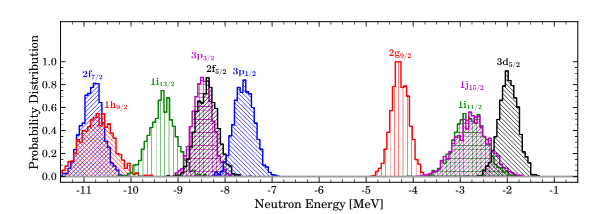


Theoretical Predictions & Probability Distributions

- Neutron levels for ^{208}Pb. Top: WS, bottom: HF Hamiltonians



Realistic phenomenological Woods-Saxon Hamiltonian



Realistic Skyrme-Hartree-Fock Hamiltonian

Part II

Nuclear Theories: Inference & Inverse Problem

• Starting from a limited experimental data set, e.g. energies $\{e_\mu^{exp}\}$,
- we wish to obtain the information about all energies of the system

Nuclear Theories: Inference & Inverse Problems

• Starting from a limited experimental data set, e.g. energies $\{e_\mu^{exp}\}$,
- we wish to obtain the information about all energies of the system

• In Applied Mathematics this approach is called *Inference Problem*

• Starting from a limited experimental data set, e.g. energies $\{e_\mu^{exp}\}$,
- we wish to obtain the information about all energies of the system

• In Applied Mathematics this approach is called *Inference Problem*

• The goal of the underlying mathematical theories is to provide statistically sound, meaningful (i.e. stable) predictions and therefore

THE PREDICTIVE POWER

• Starting from a limited experimental data set, e.g. energies $\{e_\mu^{exp}\}$,
- we wish to obtain the information about all energies of the system

• In Applied Mathematics this approach is called *Inference Problem*

• The goal of the underlying mathematical theories is to provide statistically sound, meaningful (i.e. stable) predictions and therefore

THE PREDICTIVE POWER

• All the theory predictions depend on the Hamiltonian parameters

Nuclear Theories: Inference & Inverse Problems

• Starting from a limited experimental data set, e.g. energies $\{e_\mu^{exp}\}$,
- we wish to obtain the information about all energies of the system

• In Applied Mathematics this approach is called *Inference Problem*

• The goal of the underlying mathematical theories is to provide statistically sound, meaningful (i.e. stable) predictions and therefore

THE PREDICTIVE POWER

• All the theory predictions depend on the Hamiltonian parameters

• Hamiltonian parameters fitted by physicists reflect at the same time both the form of the interactions and the data sampling (choice)

PARAMETERS INVOLVE ARBITRARY JUDGEMENT

- Consider an arbitrary, e.g. many-body, theory with its Hamiltonian:

$$\boxed{\hat{\mathbf{H}} = \hat{\mathbf{T}} + \hat{\mathbf{V}}_{\mathbf{int}}(\ldots\{\mathbf{p}\}); \quad \{\mathbf{p}\} \longrightarrow \text{Optimal parameters}}$$

- Consider an arbitrary, e.g. many-body, theory with its Hamiltonian:

$$\boxed{\hat{H} = \hat{T} + \hat{V}_{int}(\ldots\{p\}); \quad \{p\} \longrightarrow \text{Optimal parameters}}$$

- If we know the parameters, we are able to solve the Direct Problem:

$$\boxed{\hat{H}\,\varphi_j(\ldots,\{p\}) = e_j^{th}(\ldots,\{p\})\,\varphi_j(\ldots,\{p\})}$$

• Consider an arbitrary, e.g. many-body, theory with its Hamiltonian:

$$\hat{H} = \hat{T} + \hat{V}_{int}(\ldots\{p\}); \quad \{p\} \longrightarrow \text{Optimal parameters}$$

• If we know the parameters, we are able to solve the Direct Problem:

$$\hat{H}\,\varphi_j(\ldots,\{p\}) = e_j^{th}(\ldots,\{p\})\,\varphi_j(\ldots,\{p\})$$

• However, before any comparison theory-experiment, and even more generally: Before any calculation we must solve the Inverse Problem:

Determine the optimal parameters of the Hamiltonian

Inverse Problem in Quantum Theories

- Parameter adjustment usually corresponds to the χ^2-minimisation

$$\chi^2(p) = \sum_{j=1}^{d}[e_j^{exp} - e_j^{th}(p)]^2 \;\rightarrow\; \frac{\partial \chi^2}{\partial p_k} = 0,\; k = 1 \ldots m$$

where: d - number of data points; m - number of model parameters

- Parameter adjustment usually corresponds to the χ^2-minimisation

$$\chi^2(p) = \sum_{j=1}^{d} [e_j^{exp} - e_j^{th}(p)]^2 \;\rightarrow\; \frac{\partial \chi^2}{\partial p_k} = 0,\; k = 1 \ldots m$$

where: d - number of data points; m - number of model parameters

- Usually we iterate this non-linear problem using Taylor linearization

$$e_j^{th}(p^{[it+1]}) \approx e_j^{th}(p^{[it]}) + \sum_{k=1}^{m} \left(\frac{\partial e_j^{th}}{\partial p_k} \right) \Big|_{p=p^{[it]}} \left(p_k^{[it+1]} - p_k^{[it]} \right)$$

Short-hand notation: $J_{jk}^{[it]} \overset{df}{=} \left(\frac{\partial e_j^{th}}{\partial p_k} \right) \Big|_{p=p^{[it]}}$ and $b_j^{[it]} = \left[e_j^{exp} - e_j^{th}(p^{[it]}) \right]$

Inverse Problem in Quantum Theories

- Parameter adjustment usually corresponds to the χ^2-minimisation

$$\chi^2(p) = \sum_{j=1}^{d} [e_j^{exp} - e_j^{th}(p)]^2 \;\rightarrow\; \frac{\partial \chi^2}{\partial p_k} = 0,\; k = 1 \ldots m$$

where: d - number of data points; m - number of model parameters

- Usually we iterate this non-linear problem using Taylor linearization

$$e_j^{th}(p^{[it+1]}) \approx e_j^{th}(p^{[it]}) + \sum_{k=1}^{m} \left(\frac{\partial e_j^{th}}{\partial p_k} \right) \Big|_{p=p^{[it]}} \left(p_k^{[it+1]} - p_k^{[it]} \right)$$

Short-hand notation: $J_{jk}^{[it]} \overset{df}{=} \left(\frac{\partial e_j^{th}}{\partial p_k} \right) \Big|_{p=p^{[it]}}$ and $b_j^{[it]} = \left[e_j^{exp} - e_j^{th}(p^{[it]}) \right]$

- Inserting the above into $\chi^2(p)$ gives the Linearized Representation

$$\boldsymbol{\chi^2(p^{[it+1]}) = \sum_{j=1}^{d} \left[\sum_{k=1}^{m} J_{jk}^{[it]} \cdot \left(p_k^{[it+1]} - p_k^{[it]} \right) - b_j^{[it]} \right]^2}$$

Inverse Problem in Linearized Representation

- One may easily show that within the new, linearized representation

$$\frac{\partial \chi^2}{\partial p_i} = 0 \quad \rightarrow \quad (J^T J) \cdot p = J^T b \quad \leftrightarrow \quad J^T J \stackrel{df}{=} \mathcal{A}$$

- One may easily show that within the new, linearized representation

$$\frac{\partial \chi^2}{\partial \mathbf{p}_i} = 0 \quad \longrightarrow \quad (\mathbf{J}^T \mathbf{J}) \cdot \mathbf{p} = \mathbf{J}^T \mathbf{b} \quad \leftrightarrow \quad \mathbf{J}^T \mathbf{J} \stackrel{\text{df}}{=} \mathcal{A}$$

- In *Applied Mathematics* we slightly change wording and notation:

$$\{\mathbf{p}\} \rightarrow \mathcal{P} : \textbf{`Causes'} \text{ and } \{\mathbf{J}^T \mathbf{b}\} \rightarrow \mathcal{D} : \textbf{`Effects'} \Rightarrow \mathcal{A} \cdot \mathcal{P} = \mathcal{D}$$

Inverse Problem in Linearized Representation

- One may easily show that within the new, linearized representation

$$\boxed{\frac{\partial \chi^2}{\partial \mathbf{p_i}} = \mathbf{0} \quad \longrightarrow \quad (\mathbf{J^T J}) \cdot \mathbf{p} = \mathbf{J^T\, b} \quad \leftrightarrow \quad \mathbf{J^T J} \stackrel{\mathbf{df}}{=} \mathcal{A}}$$

- In *Applied Mathematics* we slightly change wording and notation:

$$\boxed{\{\mathbf{p}\} \longrightarrow \mathcal{P} : \textbf{`Causes'} \text{ and } \{\mathbf{J^T b}\} \longrightarrow \mathcal{D} : \textbf{`Effects'} \Rightarrow \mathcal{A} \cdot \mathcal{P} = \mathcal{D}}$$

- From the measured 'Effects', called Data, represented by $\mathcal{D}$, we extract information about the optimal parameters, $\mathcal{P}$, by inverting the matrix $\mathcal{A}$:

$$\boxed{\underbrace{\mathcal{A} \cdot \mathcal{P} = \mathcal{D}}_{\text{Direct Problem}} \quad \longrightarrow \quad \underbrace{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D}}_{\text{Inverse Problem}}}$$

Stability of Solutions of Nuclear Inverse Problem

- We consider linear equations: $\boxed{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \leftrightarrow \mathcal{P} = \mathcal{C} \cdot \mathcal{D}}$

$$\begin{bmatrix} \mathcal{P}_1 \\ \mathcal{P}_2 \\ \dots \\ \mathcal{P}_m \end{bmatrix} = \underbrace{\begin{bmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} & \cdots & \mathcal{C}_{1d} \\ \mathcal{C}_{21} & \mathcal{C}_{22} & \cdots & \mathcal{C}_{2d} \\ \dots & \dots & \dots & \dots \\ \mathcal{C}_{m1} & \mathcal{C}_{m2} & \cdots & \mathcal{C}_{md} \end{bmatrix}}_{\text{m} \times \text{d rectangular matrix}} \begin{bmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \\ \dots \\ \mathcal{D}_d \end{bmatrix}$$

Stability of Solutions of Nuclear Inverse Problem

- We consider linear equations: $\boxed{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \leftrightarrow \mathcal{P} = \mathcal{C} \cdot \mathcal{D}}$

$$\begin{bmatrix} \mathcal{P}_1 \\ \mathcal{P}_2 \\ \dots \\ \mathcal{P}_m \end{bmatrix} = \underbrace{\begin{bmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} & \cdots & \mathcal{C}_{1d} \\ \mathcal{C}_{21} & \mathcal{C}_{22} & \cdots & \mathcal{C}_{2d} \\ \dots & \dots & \dots & \dots \\ \mathcal{C}_{m1} & \mathcal{C}_{m2} & \cdots & \mathcal{C}_{md} \end{bmatrix}}_{\text{m} \times \text{d rectangular matrix}} \begin{bmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \\ \dots \\ \mathcal{D}_d \end{bmatrix}$$

- $[\mathcal{C}_{ik}]$ depend on: 1) Hamiltonian, and 2) Selection of data points

Stability of Solutions of Nuclear Inverse Problem

- We consider linear equations: $\boxed{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \leftrightarrow \mathcal{P} = \mathcal{C} \cdot \mathcal{D}}$

$$\begin{bmatrix} \mathcal{P}_1 \\ \mathcal{P}_2 \\ \cdots \\ \mathcal{P}_m \end{bmatrix} = \underbrace{\begin{bmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} & \cdots & \mathcal{C}_{1d} \\ \mathcal{C}_{21} & \mathcal{C}_{22} & \cdots & \mathcal{C}_{2d} \\ \cdots & \cdots & \cdots & \cdots \\ \mathcal{C}_{m1} & \mathcal{C}_{m2} & \cdots & \mathcal{C}_{md} \end{bmatrix}}_{\text{m} \times \text{d rectangular matrix}} \begin{bmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \\ \cdots \\ \mathcal{D}_d \end{bmatrix}$$

- $[\mathcal{C}_{ik}]$ depend on: 1) Hamiltonian, and 2) Selection of data points

- If one of the parameters is a function of another, say, $p_k = f(p_{k'})$ then one may show, that two columns of $\mathcal{A}$ are linearly dependent

Stability of Solutions of Nuclear Inverse Problem

- We consider linear equations: $\boxed{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \leftrightarrow \mathcal{P} = \mathcal{C} \cdot \mathcal{D}}$

$$\begin{bmatrix} \mathcal{P}_1 \\ \mathcal{P}_2 \\ \dots \\ \mathcal{P}_m \end{bmatrix} = \underbrace{\begin{bmatrix} \mathcal{C}_{11} & \mathcal{C}_{12} & \cdots & \mathcal{C}_{1d} \\ \mathcal{C}_{21} & \mathcal{C}_{22} & \cdots & \mathcal{C}_{2d} \\ \dots & \dots & \dots & \dots \\ \mathcal{C}_{m1} & \mathcal{C}_{m2} & \cdots & \mathcal{C}_{md} \end{bmatrix}}_{\text{m}\times\text{d rectangular matrix}} \begin{bmatrix} \mathcal{D}_1 \\ \mathcal{D}_2 \\ \dots \\ \mathcal{D}_d \end{bmatrix}$$

- $[\mathcal{C}_{ik}]$ depend on: 1) Hamiltonian, and 2) Selection of data points

- If one of the parameters is a function of another, say, $p_k = f(p_{k'})$ then one may show, that two columns of $\mathcal{A}$ are linearly dependent

- If this happens $\rightarrow$ $\mathcal{C}$-matrix becomes singular [Ill-Posed Problem]

Ill-Posed: Correlation between parameters and the data is lost!

A Mathematical Model
of
Predictive Power

A Mathematical Model for Predicting Data

A Mathematical Model for Predicting Data

- We will generate a set of pseudo-experimental data using function

$$f(x) \equiv \frac{\exp(\beta x)}{1+\alpha(\beta x)^2}; \quad \rightarrow \quad \boxed{\{f_i^{exp} \equiv f(x_i);\ i = 1, 2, \ \ldots n_S\}}$$

A Mathematical Model for Predicting Data

- We will generate a set of pseudo-experimental data using function

$$f(x) \equiv \frac{\exp(\beta x)}{1+\alpha(\beta x)^2}; \quad \rightarrow \quad \boxed{\{f_i^{exp} \equiv f(x_i);\ i = 1, 2,\ \ldots n_S\}}$$

- We wish to be able to describe three mechanisms important here:
  - Sampling: Controlling the number- and type of data points
  - Precision (imprecision, errors) of the experimental input data
  - Exact vs. in-exact theories - more generally: Inexact modelling

A Mathematical Model for Predicting Data

- We will generate a set of pseudo-experimental data using function

$$f(x) \equiv \frac{\exp(\beta x)}{1+\alpha(\beta x)^2}; \quad \longrightarrow \quad \boxed{\{f_i^{exp} \equiv f(x_i);\; i = 1, 2, \; \dots n_S\}}$$

- We wish to be able to describe three mechanisms important here:
  - Sampling: Controlling the number- and type of data points
  - Precision (imprecision, errors) of the experimental input data
  - Exact vs. in-exact theories - more generally: Inexact modelling

- Concerning the Sampling: We define sampling by fixing "$\mathbf{n_S}$"

A Mathematical Model for Predicting Data

- We will generate a set of pseudo-experimental data using function

$$f(x) \equiv \frac{\exp(\beta x)}{1+\alpha(\beta x)^2}; \quad \rightarrow \quad \boxed{\{f_i^{exp} \equiv f(x_i);\ i = 1, 2, \ \ldots n_S\}}$$

- We wish to be able to describe three mechanisms important here:
  - Sampling: Controlling the number- and type of data points
  - Precision (imprecision, errors) of the experimental input data
  - Exact vs. in-exact theories - more generally: Inexact modelling

- Concerning the Sampling: We define sampling by fixing "$\mathbf{n_S}$"

- We introduce the pseudo-experimental errors δf_i by setting

$$f_i^{exp} \rightarrow f_i^{exp} + \delta f_i$$

where δf_i are random numbers, here: Gaussian $\mathbf{N(0, \sigma)}$-distribution

How to Parametrize Exact- vs. Inexact-Theory?

How to Parametrize Exact- vs. Inexact-Theory?

- Observe that for $\alpha = 0$ we can express our 'sampling function' as

$$\begin{aligned} \mathbf{f(x)}\Big|_{\alpha=0} &= \mathbf{exp(\beta x)} \quad \longleftarrow \text{ “Exact” A,B,C,D-Model } \longrightarrow \\ &= \mathbf{A} + \mathbf{B}\cdot\beta\mathbf{x} + \mathbf{C}\cdot\mathbf{sinh}(\beta\mathbf{x}) + \mathbf{D}\cdot\mathbf{cosh}(\beta\mathbf{x}) \end{aligned}$$

How to Parametrize Exact- vs. Inexact-Theory?

- Observe that for $\alpha = 0$ we can express our 'sampling function' as

$$\begin{aligned} \mathbf{f(x)}\Big|_{\alpha=0} &= \mathbf{exp}(\beta \mathbf{x}) \quad \longleftarrow \text{ ``Exact'' A,B,C,D-Model } \longrightarrow \\ &= \mathbf{A} + \mathbf{B} \cdot \beta \mathbf{x} + \mathbf{C} \cdot \mathbf{sinh}(\beta \mathbf{x}) + \mathbf{D} \cdot \mathbf{cosh}(\beta \mathbf{x}) \end{aligned}$$

- For small $\beta\mathbf{x}$ we have an approximate linear dependence

$$\mathbf{exp}(\beta \mathbf{x}) \approx \mathbf{A} + \mathbf{B} \cdot \mathbf{x}$$

How to Parametrize Exact- vs. Inexact-Theory?

- Observe that for $\alpha = 0$ we can express our 'sampling function' as

$$\mathbf{f(x)}\Big|_{\alpha=0} = \mathbf{exp(\beta x)} \quad \longleftarrow \text{ "Exact" A,B,C,D-Model } \longrightarrow$$
$$= \mathbf{A} + \mathbf{B} \cdot \beta\mathbf{x} + \mathbf{C} \cdot \mathbf{sinh}(\beta\mathbf{x}) + \mathbf{D} \cdot \mathbf{cosh}(\beta\mathbf{x})$$

- For small $\beta\mathbf{x}$ we have an approximate linear dependence

$$\mathbf{exp(\beta x)} \approx \mathbf{A} + \mathbf{B} \cdot \mathbf{x}$$

- We call the 'A,B,C,D'-model exact since generally we have

$$\mathbf{exp(\beta x)} \equiv \Big[\mathbf{A}+\mathbf{B}\cdot\mathbf{x}+\mathbf{C}\cdot\mathbf{sinh}(\beta\mathbf{x})+\mathbf{D}\cdot\mathbf{cosh}(\beta\mathbf{x})\Big]\Big|_{\mathbf{A=B=0,\, C=D=1}}$$

How to Parametrize Exact- vs. Inexact-Theory?

How to Parametrize Exact- vs. Inexact-Theory?

- By convention we generate the pseudo-experimental errors using
$$\delta \mathbf{f}(\mathbf{x};\sigma) = 1/\Big(\sqrt{2\pi}\sigma\Big)\,\mathbf{exp}\,\big[-\mathbf{x}^2/(2\sigma^2)\big]$$
- We say that
  - The value of $\sigma = 0.0001$ represents 'precise' measurements
  - The value of $\sigma = 0.0005$ represents 'average' measurements
  - The value of $\sigma = 0.0050$ corresponds to 'poor' measurements

How to Parametrize Exact- vs. Inexact-Theory?

- By convention we generate the pseudo-experimental errors using

$$\boldsymbol{\delta f(x;\sigma) = 1/\left(\sqrt{2\pi}\sigma\right) \exp\left[-x^2/(2\sigma^2)\right]}$$

- We say that
  - The value of $\sigma = 0.0001$ represents 'precise' measurements
  - The value of $\sigma = 0.0005$ represents 'average' measurements
  - The value of $\sigma = 0.0050$ corresponds to 'poor' measurements

- Should x be interpreted as energy in [MeV], $\sigma = 0.0001$ implies the precision of the order of a few of hundreds of eV

How to Parametrize Exact- vs. Inexact-Theory?

- By convention we generate the pseudo-experimental errors using

$$\delta\mathbf{f}(\mathbf{x};\sigma) = 1/\left(\sqrt{2\pi}\sigma\right)\mathbf{exp}\left[-\mathbf{x}^2/(2\sigma^2)\right]$$

- We say that
  - The value of $\sigma = 0.0001$ represents ‘precise’ measurements
  - The value of $\sigma = 0.0005$ represents ‘average’ measurements
  - The value of $\sigma = 0.0050$ corresponds to ‘poor’ measurements

- Should x be interpreted as energy in [MeV], $\sigma = 0.0001$ implies the precision of the order of a few of hundreds of eV

- We consider two cases:
  - The so-called ‘Exact Theory’ (with $\alpha = 0$), and:
  - The so-called ‘In-exact Theory’ (with $\alpha = 0.001$)

How to Parametrize Exact- vs. Inexact-Theory?

- By convention we generate the pseudo-experimental errors using

$$\boldsymbol{\delta f(x;\sigma) = 1/\left(\sqrt{2\pi}\sigma\right)\exp\left[-x^2/(2\sigma^2)\right]}$$

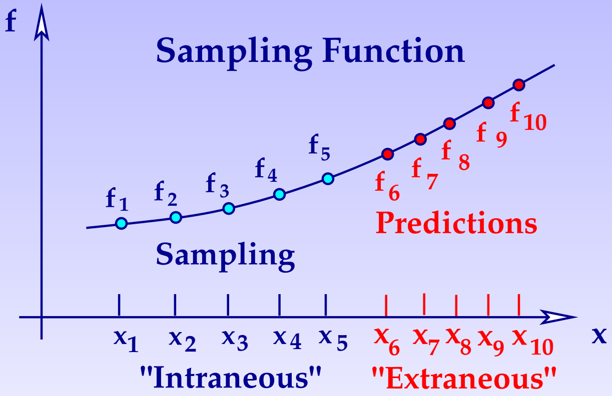
- We say that
  - The value of $\sigma = 0.0001$ represents 'precise' measurements
  - The value of $\sigma = 0.0005$ represents 'average' measurements
  - The value of $\sigma = 0.0050$ corresponds to 'poor' measurements

- Should x be interpreted as energy in [MeV], $\sigma = 0.0001$ implies the precision of the order of a few of hundreds of eV

- We consider two cases:
  - The so-called 'Exact Theory' (with $\alpha = 0$), and:
  - The so-called 'In-exact Theory' (with $\alpha = 0.001$)

- When $\alpha \neq 0 \rightarrow$ The 'a,b,c,d' formula can, in the best case, only approximate the above exponential, but it becomes exact at $\alpha \rightarrow 0$

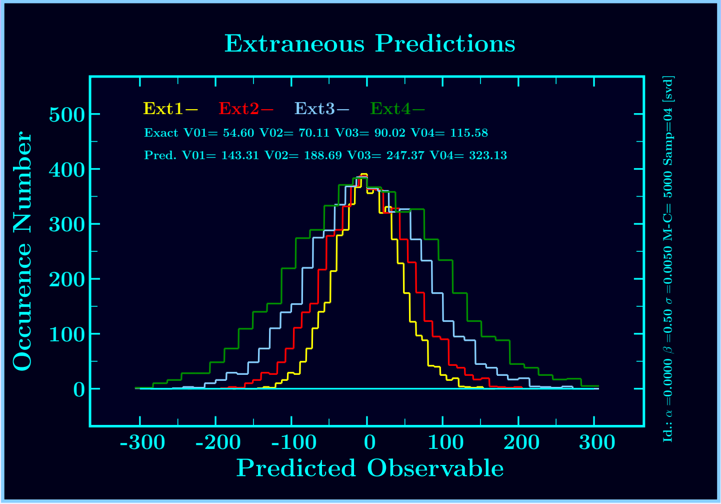
Intraneous vs. Extraneous Predictions: Summary



- Observe: From now on we 'forget about the $\{x_j\}$' $\longrightarrow$ focus on $\{f_j\}$
- Pseudo-experiment: $\{f_j\}$ $\longrightarrow$ We add random error (distributions)

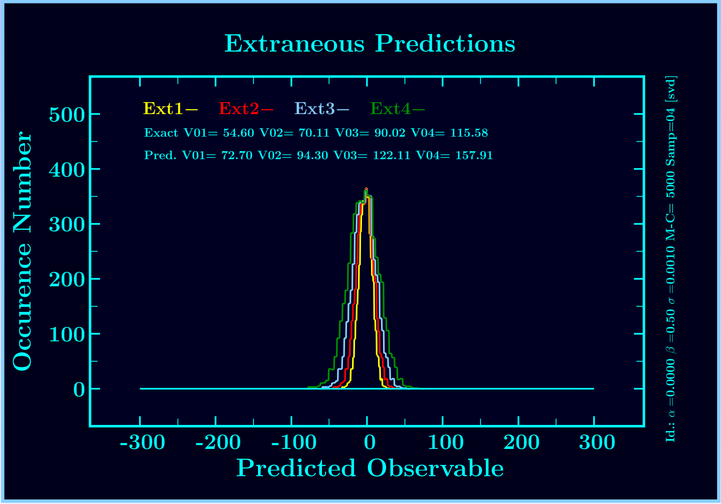
Extraneous Regime:

The Impact of Decreasing Experimental Error in the Case of an Exact Theory



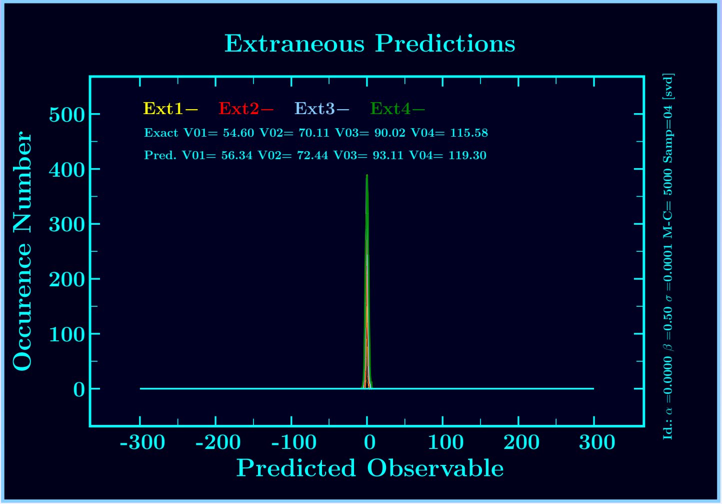
- Conditions: Big errors and weak sampling → No Predictive Power
  [Sampling: 4 points; Big Error $\sigma = 0.005$; Model: $\alpha = 0$]

Extraneous Predictions for an Exact Theory

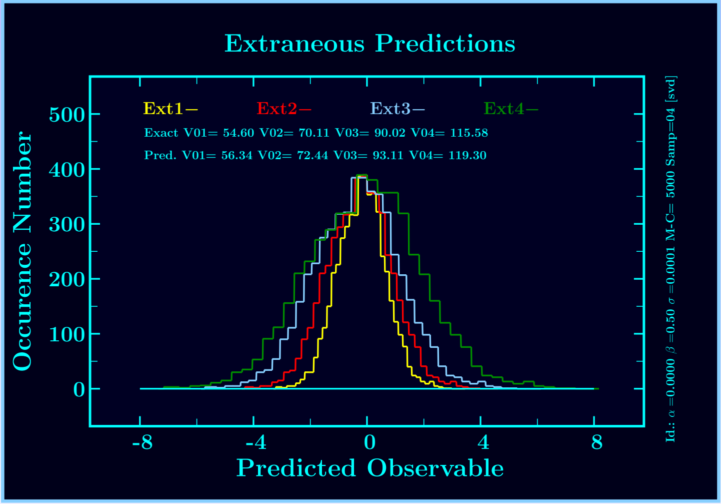


- Smaller errors (a factor of 5) → But: No 'Good' Predictive Power

[Sampling: 4 points; Moderate Error $\sigma = 0.001$; Model: $\alpha = 0$]



- Smaller errors (a factor of 10) → Here: Some Predictive Power
  [Sampling: 4 points; Small Error $\sigma = 0.0001$; Model: $\alpha = 0$]



- Error Impact → The same as before but using an enlarged scale

[Sampling: 4 points; Small Error : $\sigma = 0.0001$; Model: $\alpha = 0$]

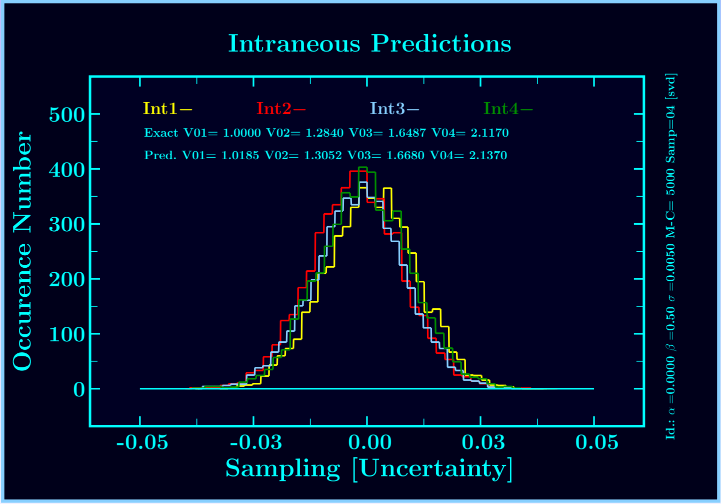
Conclusion:

Experimental errors may totally ruin the Extraneous Predictive Power even in the case of an Exact Theory

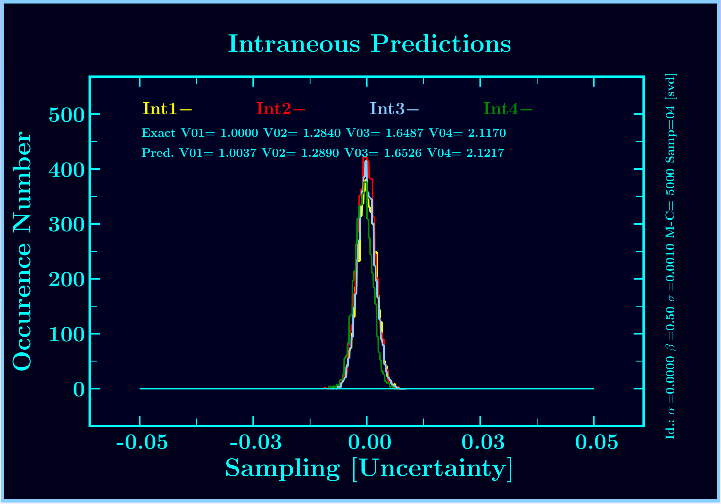
Intraneous Regime:

The Impact of Decreasing Experimental Error in the Case of an Exact Theory

Intraneous Predictions for an Exact Theory

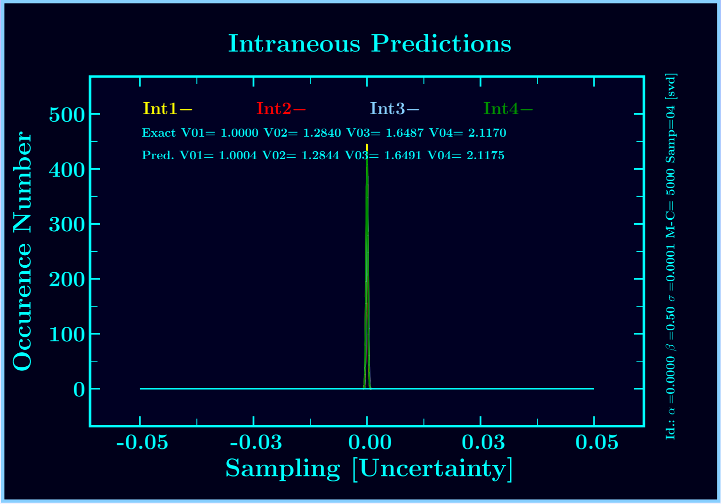


- Big errors → Small sampling → Very good fit → χ-by-the-eye
  [Sampling: 4 points; Big Error : $\sigma = 0.005$; Model: $\alpha = 0$]



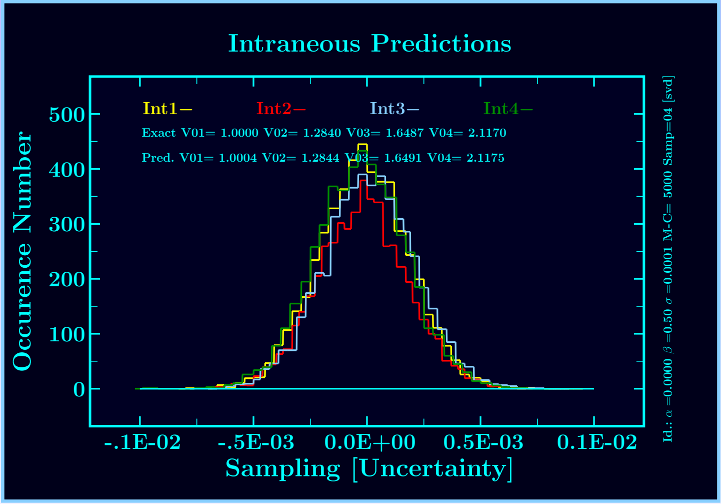
• Smaller errors → Small sampling → Very good fit → χ-by-the-eye
[Sampling: 4 points; Moderate Error : $\sigma = 0.001$; Model: $\alpha = 0$]

Intraneous Predictions for an Exact Theory



- Smaller errors → Small sampling → Excellent Fit → χ-by-the-eye

[Sampling: 4 points; Small Error : $\sigma = 0.0001$; Model: $\alpha = 0$]



- Same information, x-axis scaled → Excellent Fit → χ-by-the-eye

[Sampling: 4 points; Small Error : $\sigma = 0.0001$ Model: $\alpha = 0$]

Conclusions:

Even very large experimental errors may have a rather small impact on the Intraneous Predictive Power*)

----

*) **This is what is usually called the chi-by-the-eye "method"**

About Chi-by-the-Eye "Method"

- After laborious theoretical constructions, we get terribly exhausted and forget that: *Parameter determination is a noble, mathematically sophisticated procedure based on the statistical theories often more involved than the physical problems under study!*

About Chi-by-the-Eye "Method"

• After laborious theoretical constructions, we get terribly exhausted and forget that: *Parameter determination is a noble, mathematically sophisticated procedure based on the statistical theories often more involved than the physical problems under study!*

• In their introduction to the chapter *'Modelling of Data'*, the authors of *'Numerical Recipes"* (p. 651), observe with sarcasm:

About Chi-by-the-Eye "Method"

• After laborious theoretical constructions, we get terribly exhausted and forget that: *Parameter determination is a noble, mathematically sophisticated procedure based on the statistical theories often more involved than the physical problems under study!*

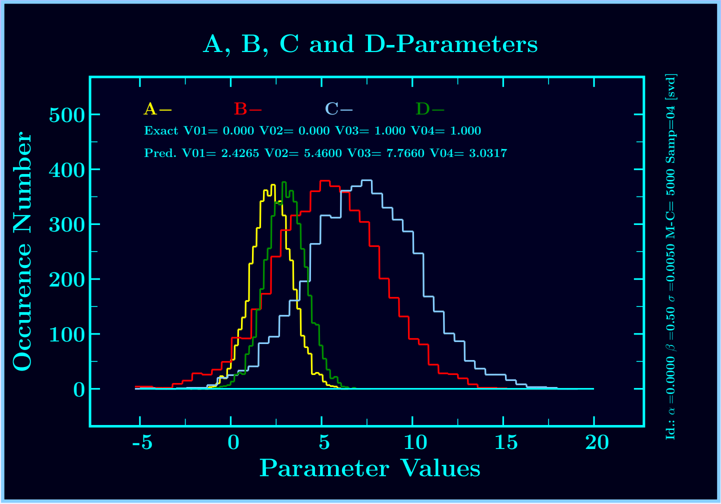
• In their introduction to the chapter *'Modelling of Data'*, the authors of *'Numerical Recipes"* (p. 651), observe with sarcasm:

"Unfortunately, many practitioners of parameter estimation never proceed beyond determining the numerical values of the parameter fit. They deem a fit acceptable if a graph of data and model 'l o o k s   g o o d'. This approach is known as chi-by-the-eye. Luckily, its practitioners get what they deserve" [i.e. - what is meant is: "they" get a 'statistical nonsense']

The Mechanism: Why?

Why are the Intraneous and Extraneous components of Predictive Power so strongly decorrelated?

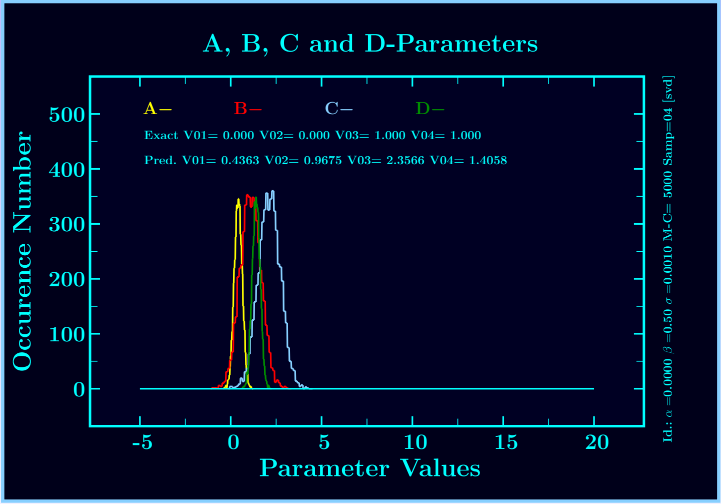
Fitted Parameters for an Exact Theory



- Parameters totally wrong, but: → Excellent Fit → χ-by-the-eye

[Sampling: 4 points; Big Error: $\sigma = 0.005$; Model: $\alpha = 0$]

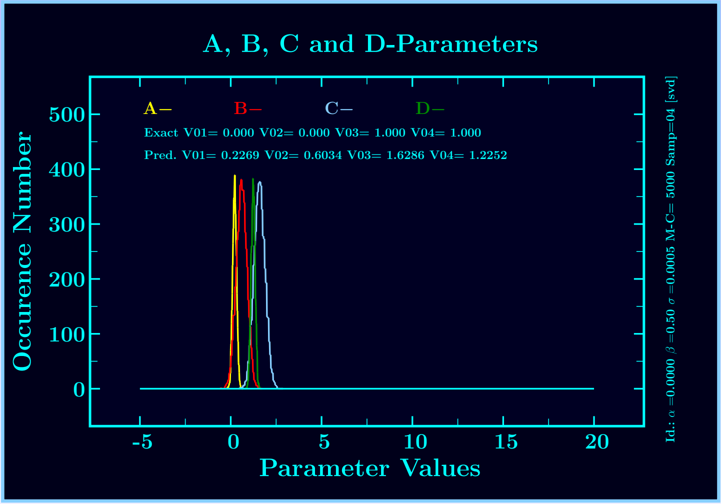
Fitted Parameters for an Exact Theory



- Parameters still quite wrong: → Excellent Fit → χ-by-the-eye

[Sampling: 4 points; Moderate Error: $\sigma = 0.001$; Model: $\alpha = 0$]

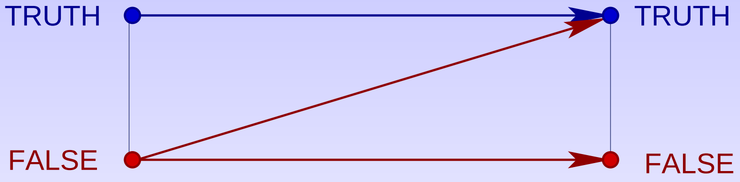
Fitted Parameters for an Exact Theory



- Parameters not really good, but: $\rightarrow$ Excellent Fit $\rightarrow$ χ-by-the-eye

[Sampling: 4 points; Small Error: $\sigma = 0.0005$; Model: $\alpha = 0$]

- As it is well known in logic: **An error may imply the truth!**



- **Parameters were totally wrong, and yet: $\longrightarrow$ Excellent Fit**
- **Exact theories/models are rare but extremely instructive**

Conclusions:

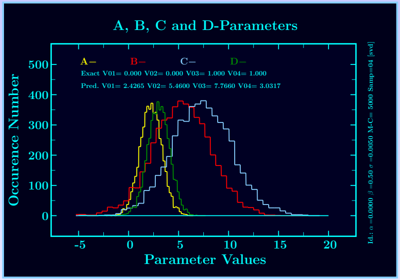
1. **We may easily obtain an excellent fit with totally wrong parameters**
2. **This mechanism is a known sign of an ill-posed Inverse Problem**

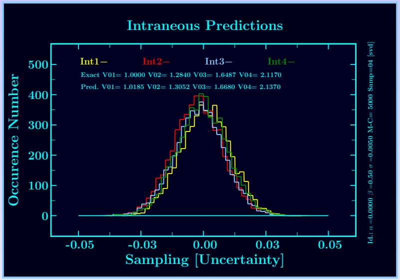
Illustrations:

A Comparative Study
of Various Quantities of the Model

Fit vs. Intraneous Predictive Power

- There is a risk of fooling oneself with the chi-by-the-eye technique
- ... and yet: The reproduction of the input may seem excellent ...

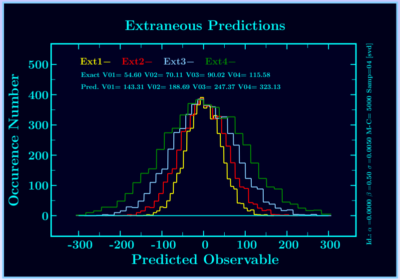


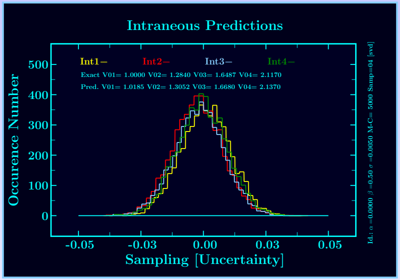


- Parameters totally wrong, but: $\longrightarrow$ Excellent Fit $\longrightarrow$ χ-by-the-eye

[Sampling: 4 points; Big Error: $\sigma = 0.005$; Model: $\alpha = 0$]

- There is a risk of fooling oneself with the chi-by-the-eye technique
- Although: The reproduction of the input may seem excellent...





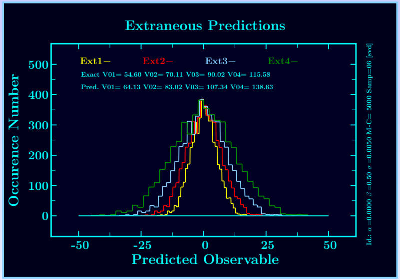
- There is no extraneous predictive power whatsoever = 'Good' Fit

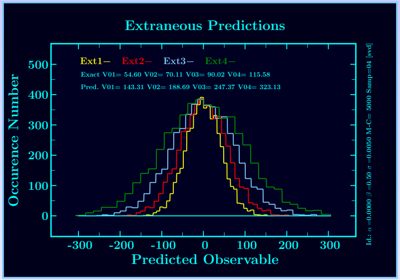
[Sampling: 4 points; Big Error: $\sigma = 0.005$; Model: $\alpha = 0$]

Increasing the Sampling vs. Predictive Power

- Big errors but increasing sampling → Improving Predictive Power?

  [Sampling: 6 points [left]; 4 points [right]; Error $\sigma = 0.005$]

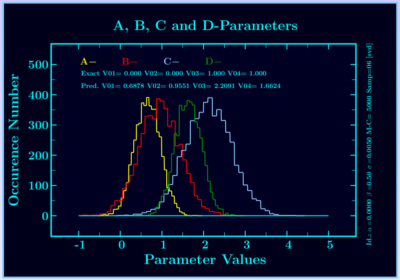


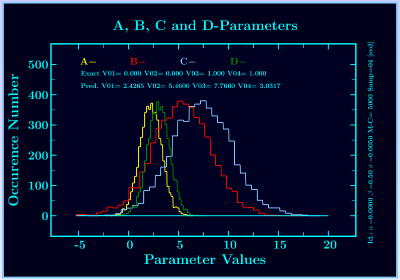


- Increasing sampling at a constant experimental error modelling <u>decreased the relative percentage errors</u> by ∼an order of magnitude

Increasing the Sampling: Intraneous vs. Extraneous

- Big errors but increasing sampling → Improving Predictive Power?
  [Sampling: 6 points [left]; 4 points [right]; Error $\sigma = 0.005$]

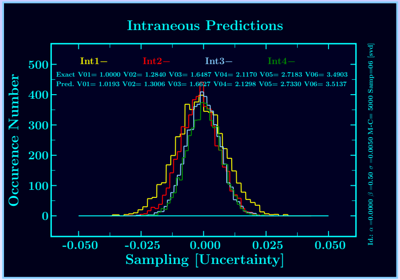


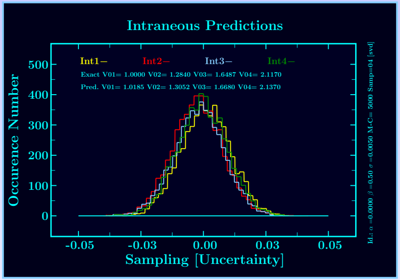


- Increasing sampling at a constant experimental error modelling we restore the order of solutions and their approximate magnitude

Increasing the Sampling: Intraneous Predictions

- Big errors but increasing sampling → Improving Predictive Power
  [Sampling: 6 points [left]; 4 points [right]; Error $\sigma = 0.005$]



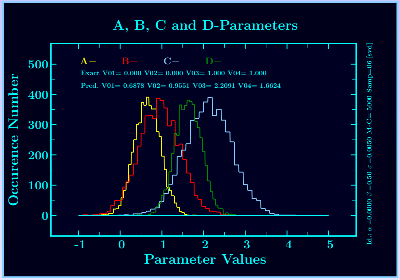


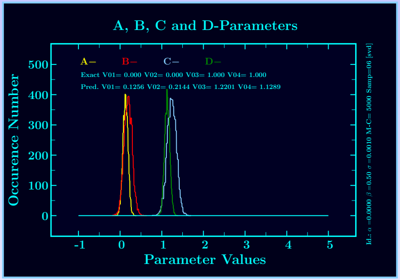
- Increasing sampling at a constant experimental error modelling <u>has no impact</u> on the intraneous performance of predictive power

Possible Improvements:

The Focus
on the Experimental Errors
& Their Impact on Parameters

- In how much decreasing experimental errors improves modelling?
  [Sampling: 6 points; Error $\sigma = 0.005$ (left) $\sigma = 0.001$ (right)]



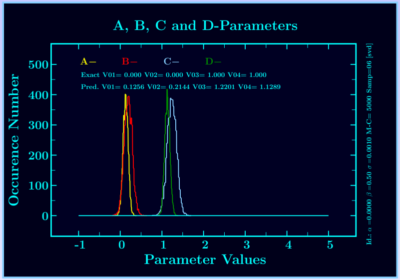


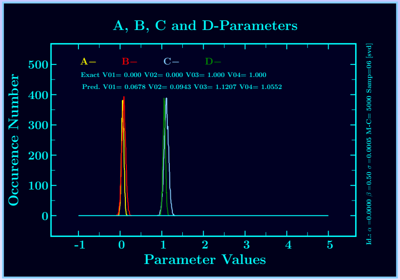
- Decreasing the experimental error by a factor of 5 at constant sampling implies a significant improvement in fitting parameters

Decreasing Experimental Errors: Fitted Parameters

- In how much decreasing experimental errors improves modelling?

[Sampling: 6 points; Error $\sigma = 0.001$ (left) $\sigma = 0.0005$ (right)]



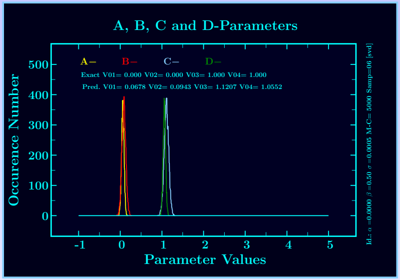


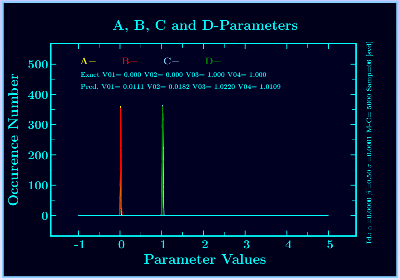
- Decreasing the experimental error by a factor of 5 at constant sampling implies more significant improvement in fitting parameters

Decreasing Experimental Errors: Fitted Parameters

- In how much decreasing experimental errors improves modelling?

[Sampling: 6 points; Error $\sigma = 0.0005$ (left) $\sigma = 0.0001$ (right)]





- Decreasing the experimental error by a factor of 5 at constant sampling implies a definite improvement in fitting parameters

Conclusions & Questions

1. **By increasing the experimental precision we definitely approach the right parameters of the Exact Theory**

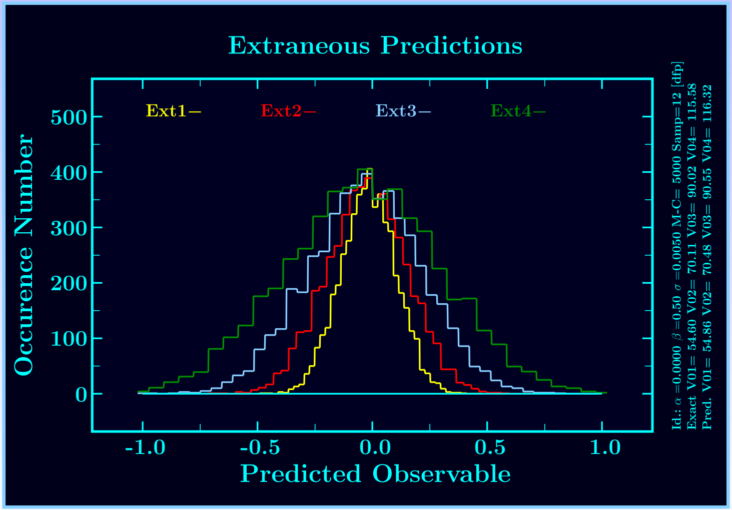
2. **Are we definitely solving the issue of the ill-posed Inverse Problem?**

Possible Improvements:

The Focus
on the Improved Sampling:
Impact on Extraneous Predictions

- We fix sampling at 12 points and see how far we can go improving?

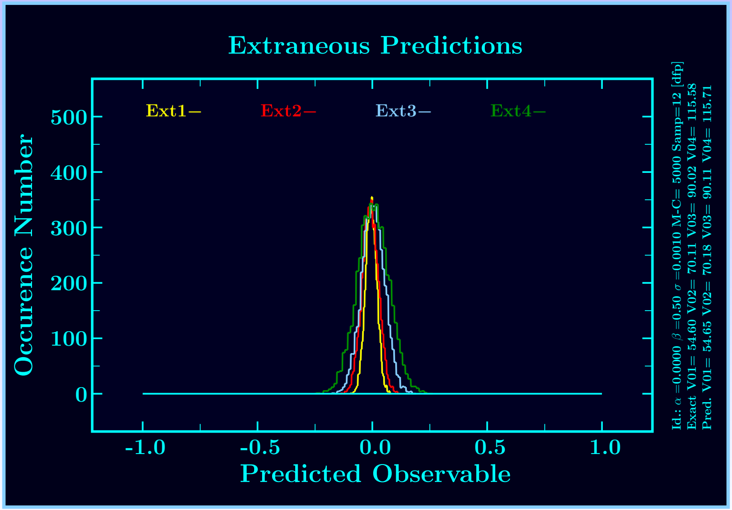
[Sampling: 12 points; Decreasing Error, here: $\sigma = 0.005$]



Extraneous Predictions at Sufficient Sampling

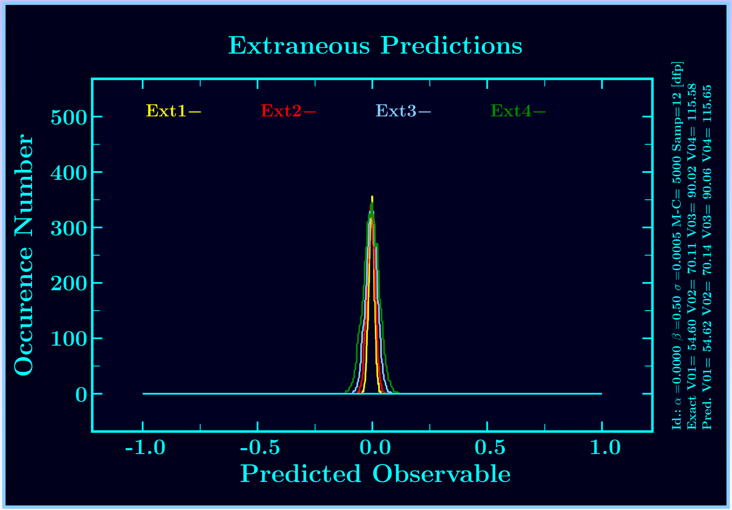
- We fix sampling at 12 points and see how far we can go improving?

  [Sampling: 12 points; Decreasing Error, here: $\sigma = 0.001$]

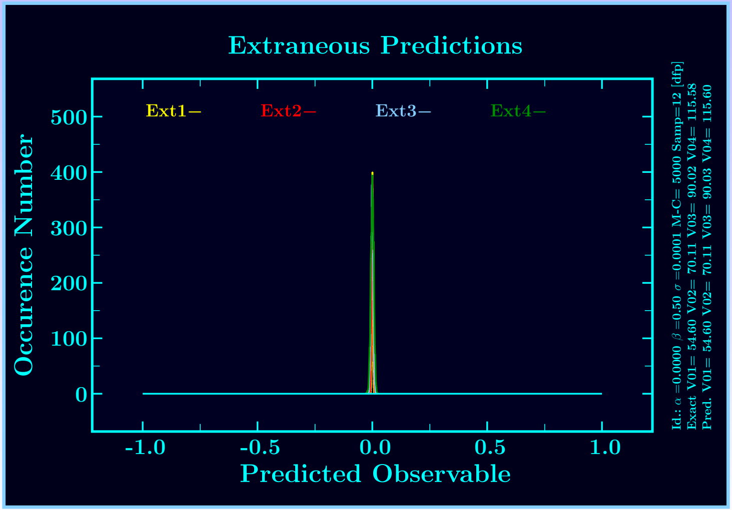


Extraneous Predictions at Sufficient Sampling

- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Decreasing Error, here: $\sigma = 0.0005$]



- We fix sampling at 12 points and see how far we can go improving?
[Sampling: 12 points; Decreasing Error, here: $\sigma = 0.0001$]



Conclusions & Questions

1. **By increasing the number of fit data-points we definitely arrive at "predicting" of our extraneous data-points**

2. **Again: Are we definitely solving the issue of the ill-posed Inverse Problem?**

A So-Far Ignored Mechanism:

Parametric Correlations in Mathematical Modelling

Principles of a Simple Monte-Carlo Technique

• We generate pseudo-experimental errors: Here we will use random numbers following the Gaussian distribution $N(0, \sigma)$ for $n_S = 50\,000$

• We repeat the parameter fit 50 000 times thus obtaining 50 000 "optimal parameter sets" - they are denoted: P_1, P_2, P_3 and P_4

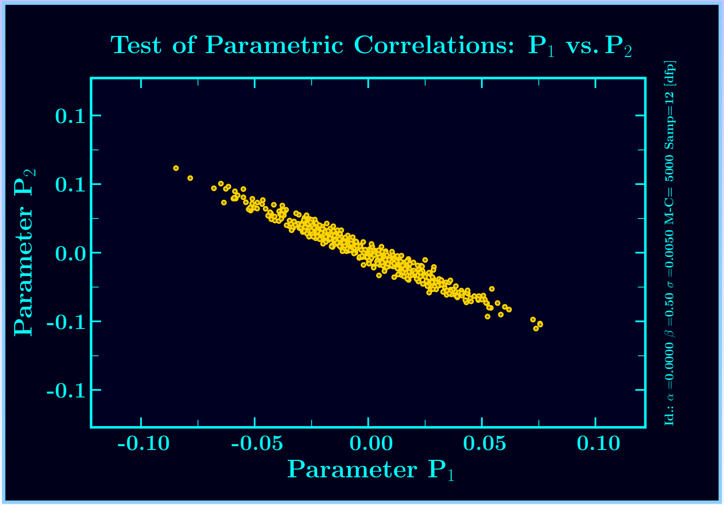
• We plot two-dimensional projections in the form of points with the coordinates P_i vs. P_j on the x-y plane (in principle: 50 000 points)

• If there are no parametric correlations - the parameters fill in a certain sub-set on the x-y plane: a circle, an ellipsoid, etc.

• Any pattern that resembles a line will be interpreted here as the corresponding parametric correlation $P_i vs. P_j(remaining\ parameters)$

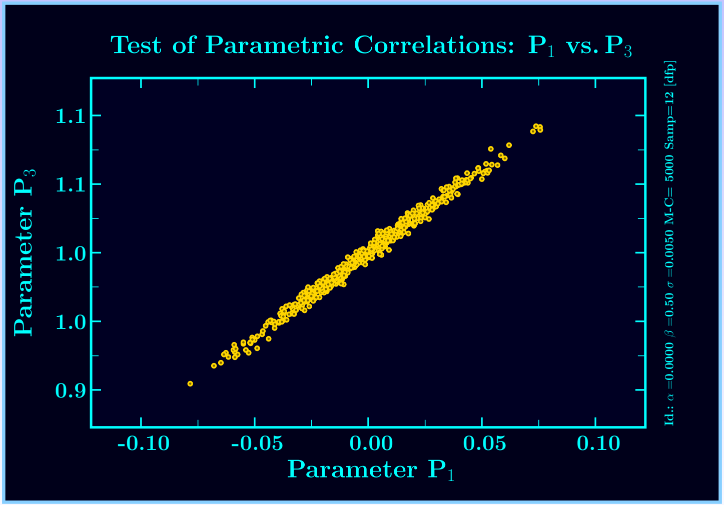
- Not done at all! Discover a disaster whose name is: Correlations!!

[Parametric Correlations between parameters A=P_1 and B=P_2]



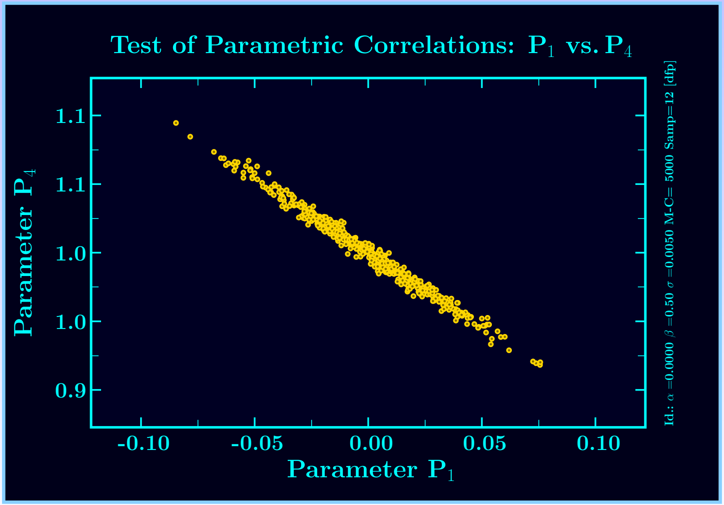
- Not done at all! Discover a disaster whose name is: Correlations!!

[Disaster continues: Correlations between A=P_1 and C=P_3]



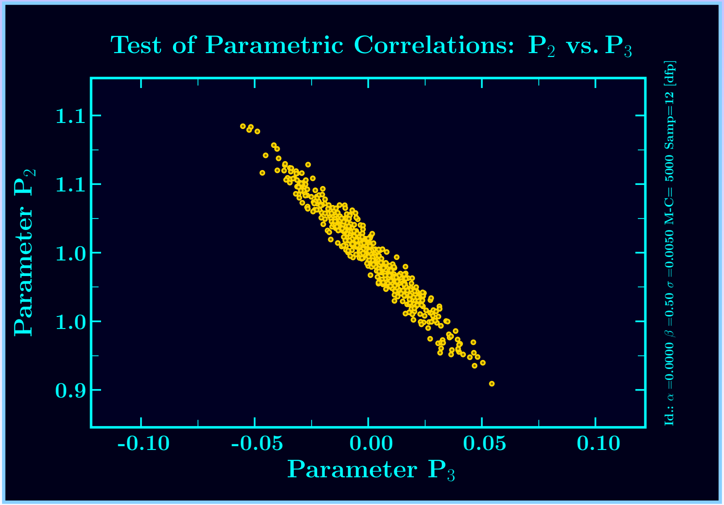
- Not done at all! Discover a disaster whose name is: Correlations!!

[Disaster continues: Correlations between A=P_1 and D=P_4]



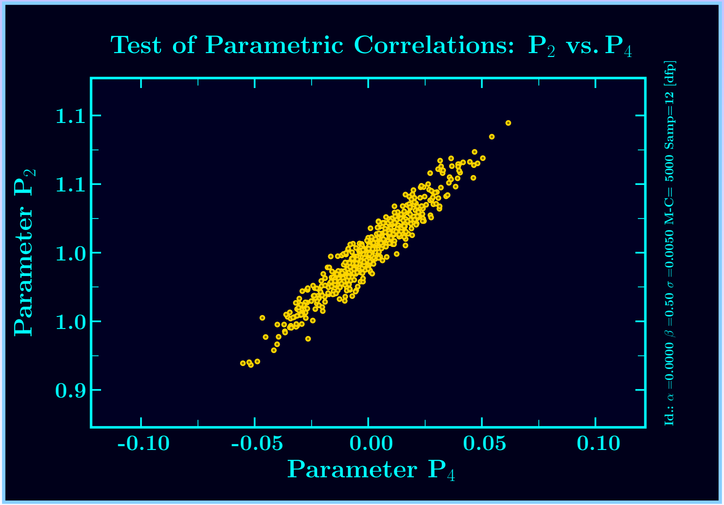
- Not done at all! Discover a disaster whose name is: Correlations!!

[From bad to worse: Correlations between B=P_2 and C=P_3]



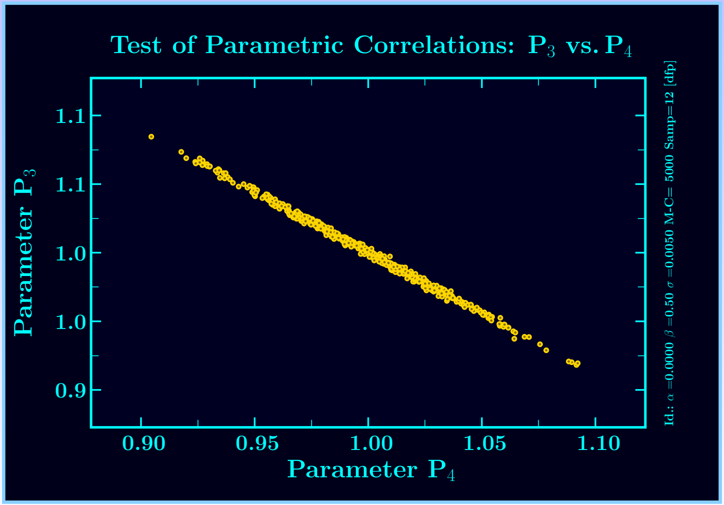
- Not done at all! Discover a disaster whose name is: Correlations!!

[Bad luck continues: Correlations between B=P_2 and D=P_4]



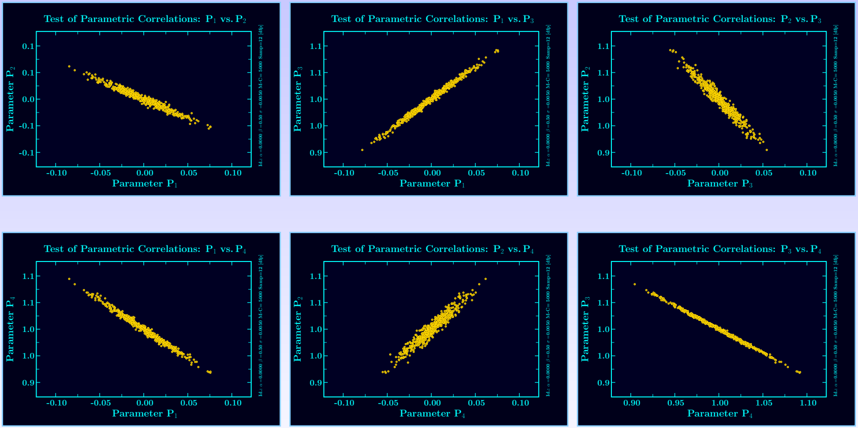
- Not done at all! Discover a disaster whose name is: Correlations!!

[If that was not enough: Correlations between C=P_3 and D=P_4]



All Model Parameters Are Perfectly Correlated!

- This is the worst what may happen: All parameters correlated imply the ill-posedeness of the inverse problem: No predictive power



What Did We Learn?

1. **An exact theory may contain parametric correlations**

What Did We Learn?

1. An exact theory may contain parametric correlations

2. Correlations can be studied and illustrated with the help of the Monte-Carlo 2-D projections as shown above

What Did We Learn?

1. An exact theory may contain parametric correlations

2. Correlations can be studied and illustrated with the help of the Monte-Carlo 2-D projections as shown above

3. For exact theories & null-errors they can be ignored...

What Did We Learn?

1. **An exact theory may contain parametric correlations**

2. **Correlations can be studied and illustrated with the help of the Monte-Carlo 2-D projections as shown above**

3. **For exact theories & null-errors they can be ignored...**

4. **... but when shall we have the null errors?**

What Did We Learn?

1. An exact theory may contain parametric correlations

2. Correlations can be studied and illustrated with the help of the Monte-Carlo 2-D projections as shown above

3. For exact theories & null-errors they can be ignored...

4. ... but when shall we have the null errors?

5. Importantly: In the general case they imply Ill-Posed Inverse Problem: No stability in theory Predictive Power

Parameter-Correlations in Skyrme-HF

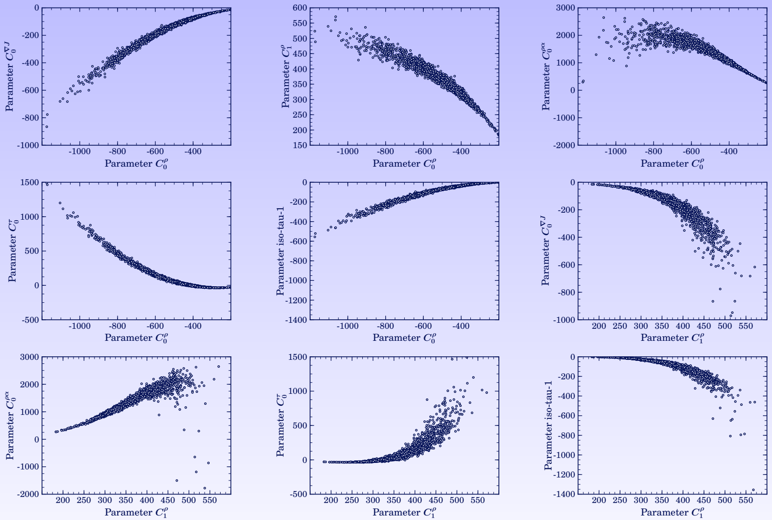
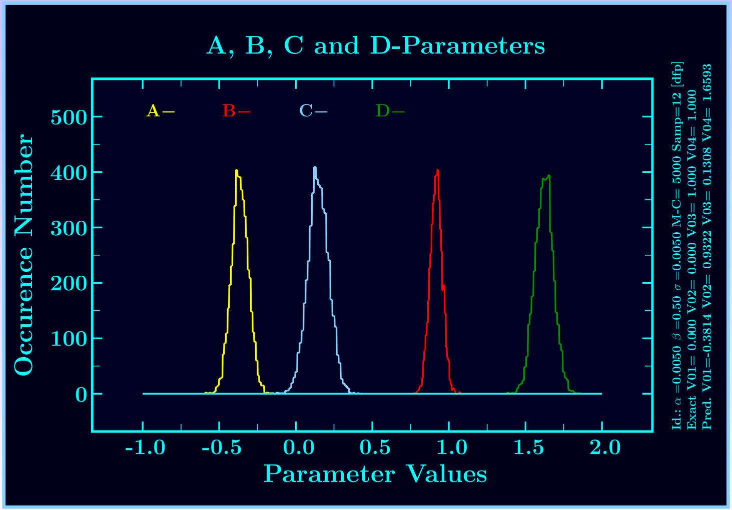


Illustration suggesting that there are rather very few independent parameters

The Case of an Inexact Theory:

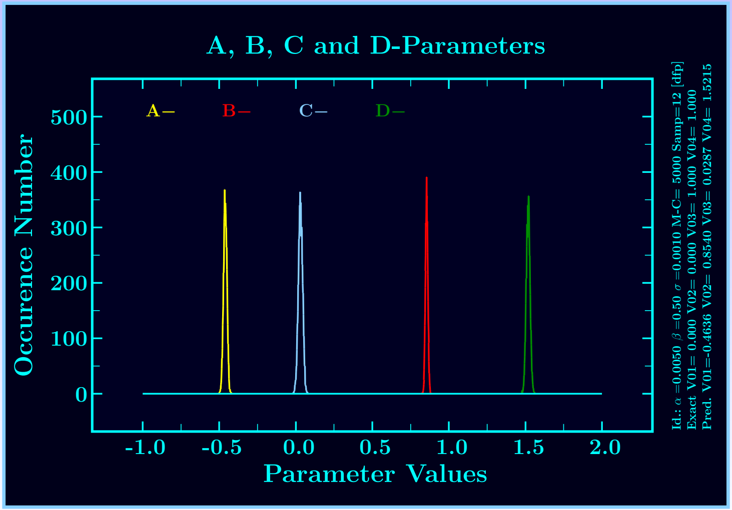
The Number of Factors to Consider and of Mechanisms to Analyse - Increases: Things Get More Complicated [but perfectly doable]

- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Error: $\sigma = 0.005$; Model: $\alpha = 0.005$]



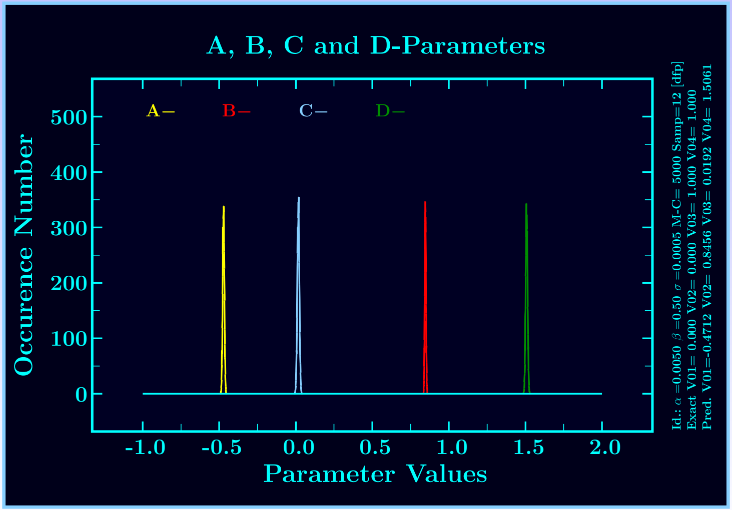
Decreasing Experimental Errors: Inexact Theory

- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Error: $\sigma = 0.001$; Model: $\alpha = 0.005$]



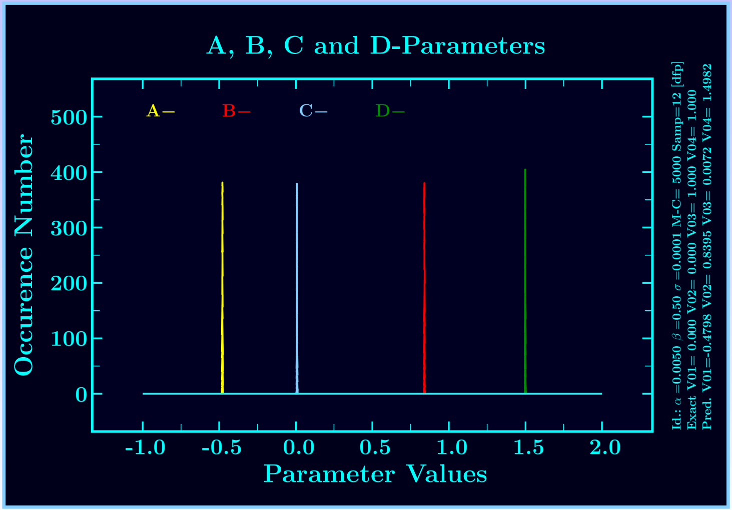
Decreasing Experimental Errors: Inexact Theory

- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Error: $\sigma = 0.0005$; Model: $\alpha = 0.005$]

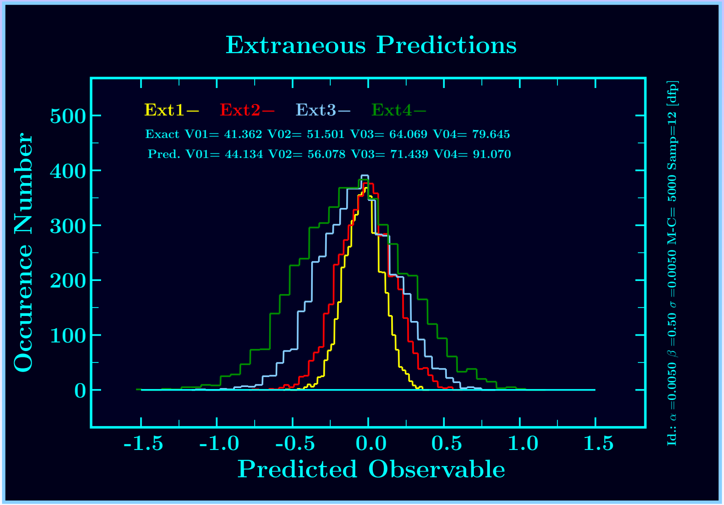


- We fix sampling at 12 points and see how far we can go improving?

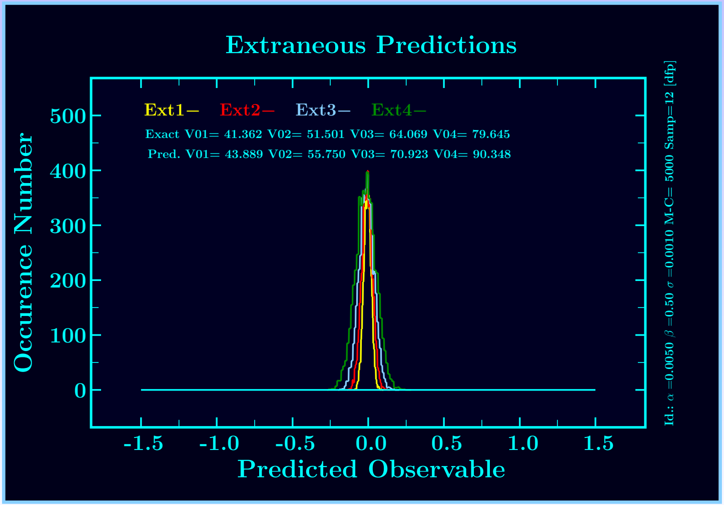
[Sampling: 12 points; Error: $\sigma = 0.0001$; Model: $\alpha = 0.005$]



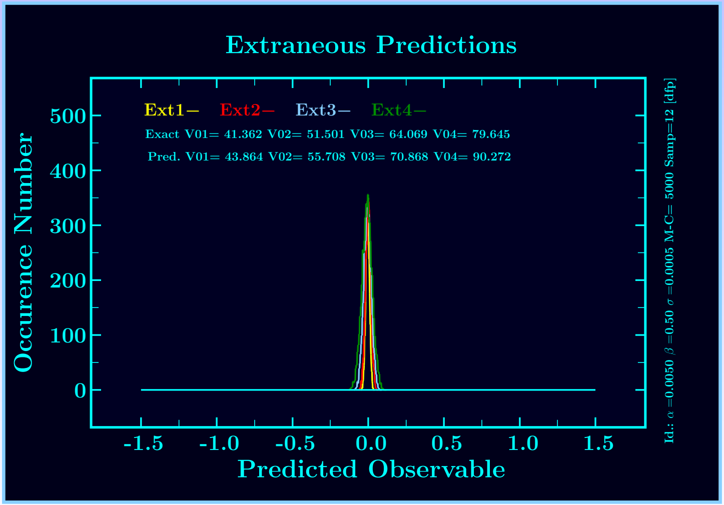
- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Error: $\sigma = 0.005$; Model: $\alpha = 0.005$]



- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Error: $\sigma = 0.001$; Model: $\alpha = 0.005$]

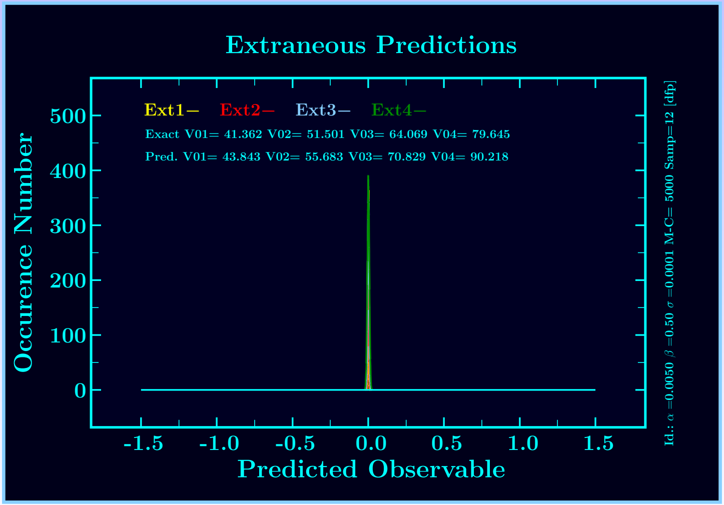


- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Error: $\sigma = 0.0005$; Model: $\alpha = 0.005$]



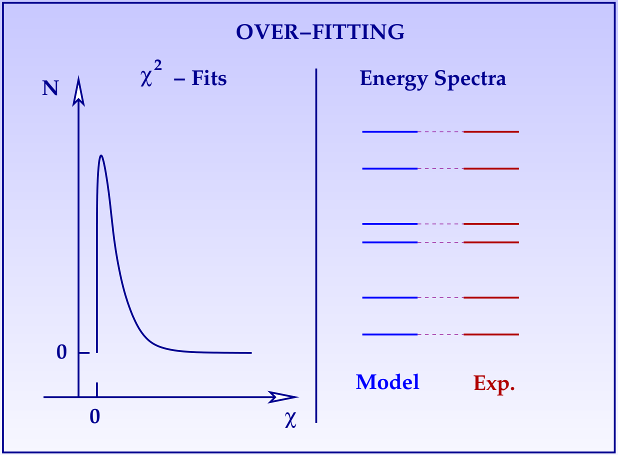
Decreasing Experimental Errors: Inexact Theory

- We fix sampling at 12 points and see how far we can go improving?
  [Sampling: 12 points; Error: $\sigma = 0.0001$; Model: $\alpha = 0.005$]

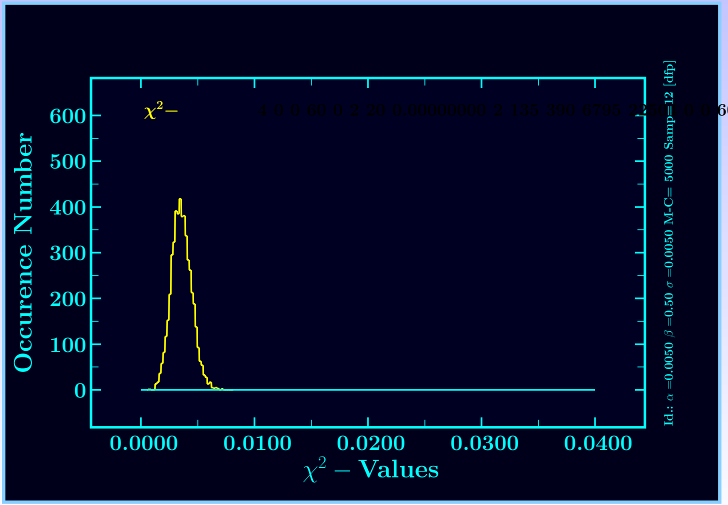


The Mechanism of Over-fitting

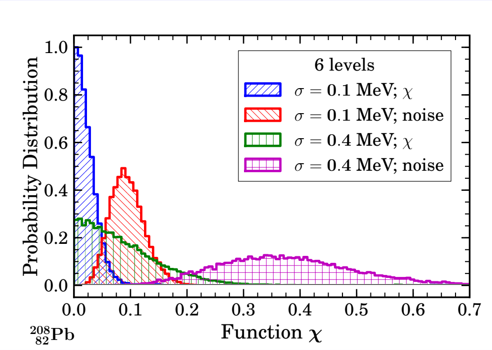
- Over-fitting: A mechanism according to which the model adjusts itself to 'any' data set with $\chi^2 \approx 0$ (data do not constrain the model)



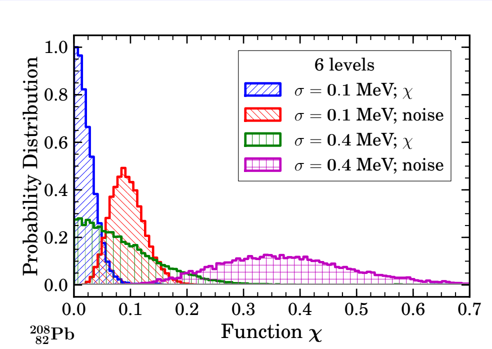
- Over-fitting: A mechanism according to which the model adjusts itself to 'any' data set with $\chi^2 \approx 0$ (data do not constrain the model)



Unprecedented Precision ... Known as 'Over-Fitting'

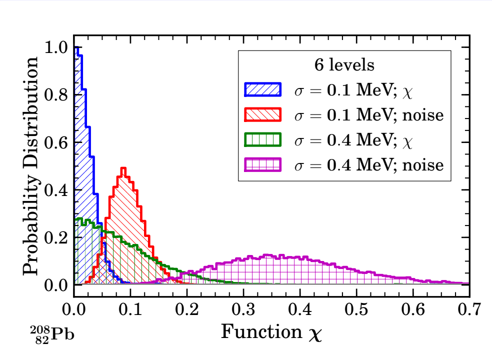


Unprecedented Precision ... Known as 'Over-Fitting'

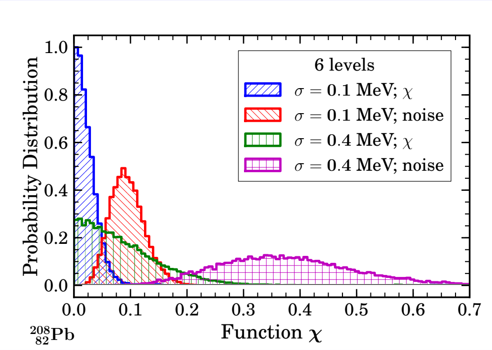


- We introduce the Gaussian noise into the experimental-level input,

Unprecedented Precision ... Known as ‘Over-Fitting’

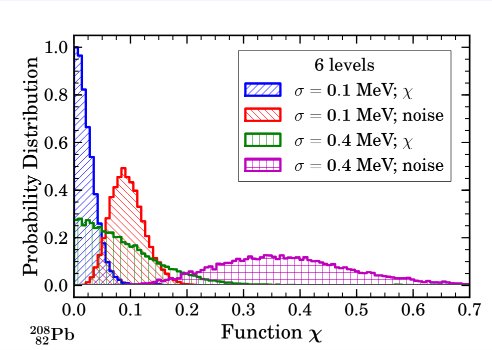


- We introduce the Gaussian noise into the experimental-level input, repeat the χ^2-fit



• We introduce the Gaussian noise into the experimental-level input, repeat the χ^2-fit - and plot the histograms in function of χ^2.

Unprecedented Precision ... Known as 'Over-Fitting'

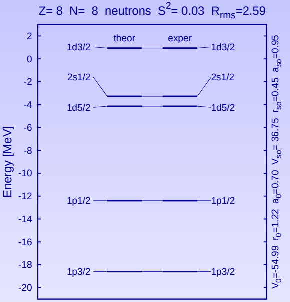


• We introduce the Gaussian noise into the experimental-level input, repeat the χ^2-fit - and plot the histograms in function of χ^2.

• Under the mathematical conditions discussed there are a large number of exact fits possible. Over-Fitting - is a form of ill-posedenenss

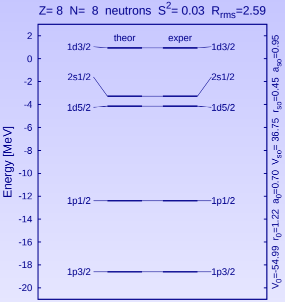
'Chi-by-the-Eye' Results May Look Attractive...

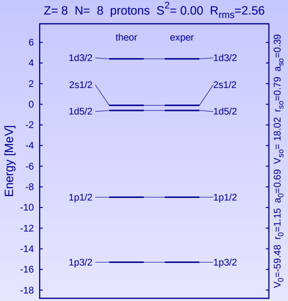
- We fit the single-particle experimental levels in ^{16}O using Woods-Saxon potential (six parameters for protons and neutrons each)



‘Chi-by-the-Eye’ Results May Look Attractive...

• We fit the single-particle experimental levels in ^{16}O using Woods-Saxon potential (six parameters for protons and neutrons each)

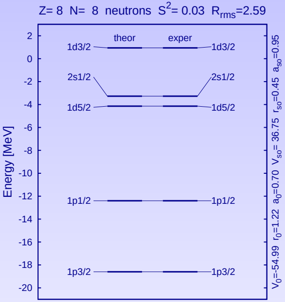


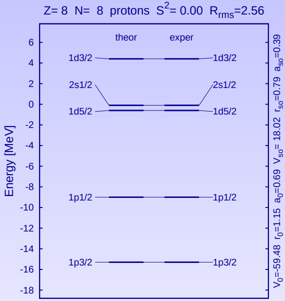


• This result may look surprising: the quality of the fit is such that *graphical illustrations are insufficient to show it* !!!

'Chi-by-the-Eye' Results May Look Attractive...

• We fit the single-particle experimental levels in ^{16}O using Woods-Saxon potential (six parameters for protons and neutrons each)





• This result may look surprising: the quality of the fit is such that *graphical illustrations are* *insufficient to show it* !!!

• On the other hand: If we trust the model - we may hope that also the remaining levels are close to the experimental results to come

Summary and Conclusions (I)

Summary and Conclusions (I)

- The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem

b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem
b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

• Inexact theories involve always theory uncertainties (which must be estimated) and related probability distributions can be modelled

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem
b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

• Inexact theories involve always theory uncertainties (which must be estimated) and related probability distributions can be modelled

• In the future theoretical approaches: Theory provides not only the numerical predictions but also probability distributions of the associated uncertainties

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem
b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

• Inexact theories involve always theory uncertainties (which must be estimated) and related probability distributions can be modelled

• In the future theoretical approaches: Theory provides not only the numerical predictions but also probability distributions of the associated uncertainties

• We believe that quite often it is easier to estimate the uncertainties of the present theory rather than to document a new interaction term

Summary and Conclusions (II)

Summary and Conclusions (II)

- Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters
C. We increase the number of data points (if we can pay for it...)

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters
C. We increase the number of data points (if we can pay for it...)

• Suppose we have already used out all the existing experimental data: as theorists we can modify models / analyse uncertainties...

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters
C. We increase the number of data points (if we can pay for it...)

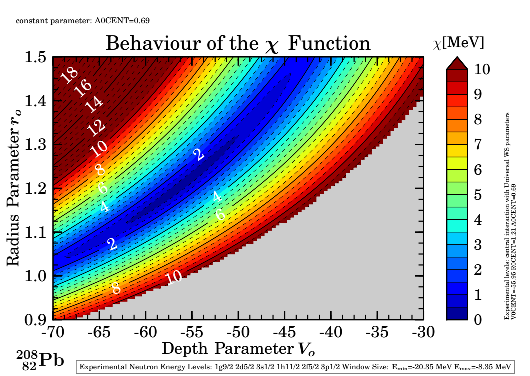
• Suppose we have already used out all the existing experimental data: as theorists we can modify models / analyse uncertainties...

• In other words: We improve predictive power of our theory by reducing the number of parameters, by regularising the associated Inverse Problem, but first of all through including all interactions

Part III

Ill-Posed Problems with Parametric Correlations: Illustrative Examples with Realistic Hamiltonians

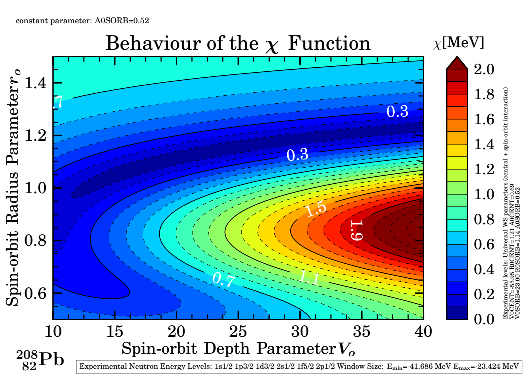
- The valley on the χ^2-plot showing correlation: $r_0 = f(V_0)$



A map of χ^2 from the fit based on six exp. levels close to the Fermi level

Spherical Woods-Saxon and Correlations V_o^{so} vs. r_o^{so}

- Valley on the χ^2-plot showing parametric correlations for $V_{WS}^{so}(r)$



We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels: $\boxed{r_0^{so} = F(V_0^{so})}$

Parameter Correlations and Correlation Matrix [WS]

- Given random variables X and Y. Correlation matrix in this case:

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

Parameter Correlations and Correlation Matrix [WS]

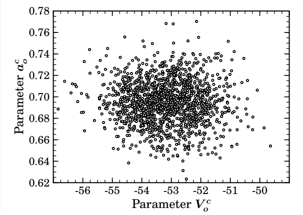
- Given random variables X and Y. Correlation matrix in this case:

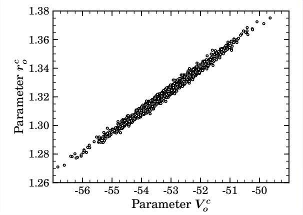
$$\mathrm{corr}(X,Y) = \frac{\sum_i[(X_i-\bar{X})(Y_i-\bar{Y})]}{\sqrt{\sum_i(X_i-\bar{X})^2}\sqrt{\sum_i(Y_i-\bar{Y})^2}}; \quad \bar{X} \equiv \frac{1}{n}\sum_{i=1}^{n} X_i, \ \bar{Y} \equiv \frac{1}{n}\sum_{i=1}^{n} Y_i$$

- Generally: $\{X, Y\} \to \{X_k\} = \{V_0^c, r_0^c, a_0^c, V_0^{so}, r_0^{so}\}$ we obtain:

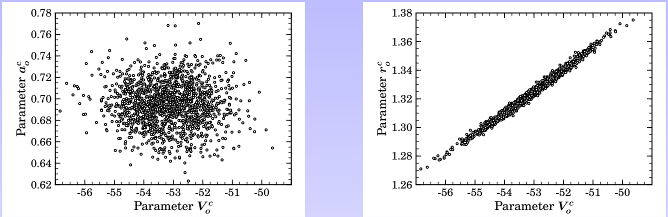
Correlation matrix for the Woods-Saxon Hamiltonian parameters
as obtained from the Monte-Carlo simulation

	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





Monte-Carlo fitting results for ^{208}Pb with the Woods-Saxon potential
Left:(a_0^c vs. V_0^c)-plane and Right: (r_0^c vs. V_0^c)-plane



Monte-Carlo fitting results for ^{208}Pb with the Woods-Saxon potential
Left:(a_0^c vs. V_0^c)-plane and Right: (r_0^c vs. V_0^c)-plane

Correlation matrix for the Woods-Saxon Hamiltonian parameters

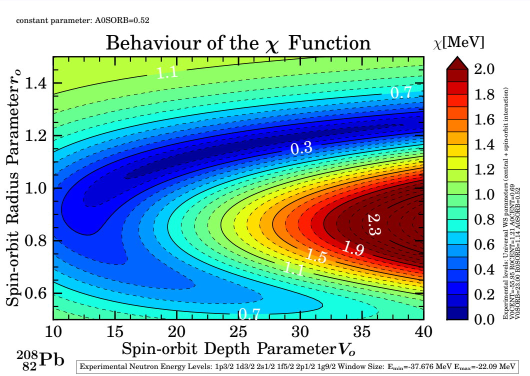
	V_0^c	r_0^c	a_0^c	V_0^{so}	r_0^{so}
V_0^c	1.000	0.994	-0.028	0.000	0.265
r_0^c	0.994	1.000	0.016	0.005	0.270
a_0^c	0.028	0.016	1.000	0.259	0.288
V_0^{so}	0.000	0.005	0.259	1.000	0.506
r_0^{so}	0.265	0.270	0.288	0.506	1.000

Sampling and Parametric Correlations

We will gradually increase the energy of the six-level window to approach the nucleon binding region and thus gradually approach the present-day experimental situation

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

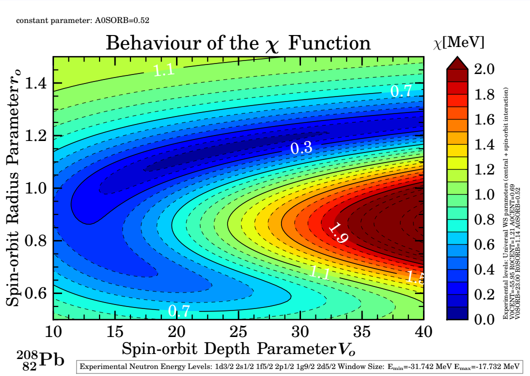
- Impact of sampling (choice of data) on Parametric Correlations



We plot the χ^2 *in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels:* $\boxed{r_0^{so} = F(V_0^{so})}$

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

- Impact of sampling (choice of data) on Parametric Correlations

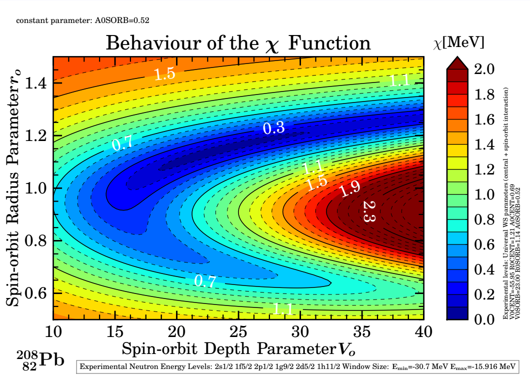


We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels:

$$\boxed{r_0^{so} = F(V_0^{so})}$$

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

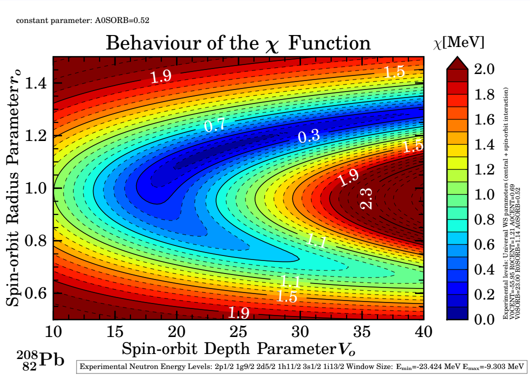
- Impact of sampling (choice of data) on Parametric Correlations



We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels: $r_0^{so} = F(V_0^{so})$

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

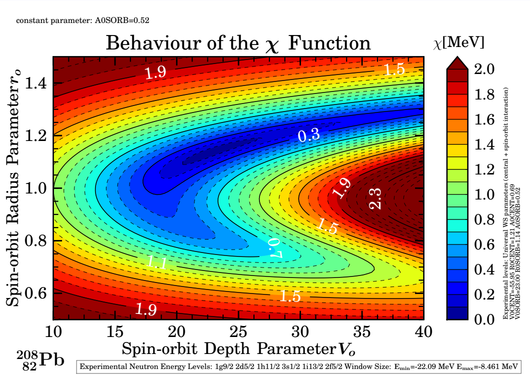
- Impact of sampling (choice of data) on Parametric Correlations



We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels: $\boxed{r_0^{so} = F(V_0^{so})}$

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

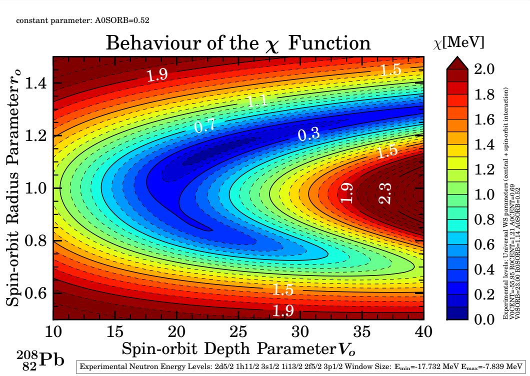
- Impact of sampling (choice of data) on Parametric Correlations



We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels: $r_0^{so} = F(V_0^{so})$

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

- Impact of sampling (choice of data) on Parametric Correlations

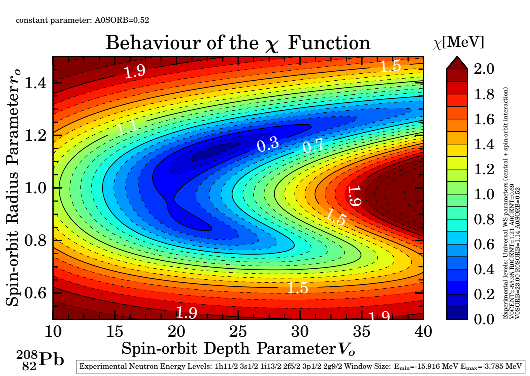


We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels:

$$r_0^{so} = F(V_0^{so})$$

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

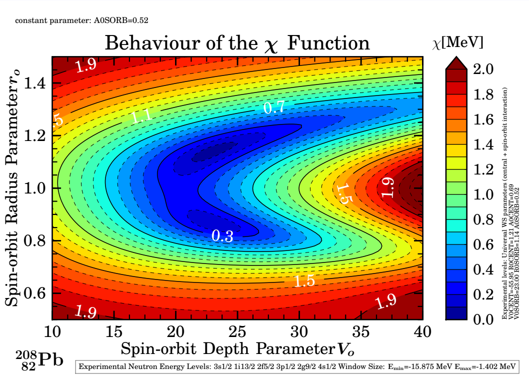
- Impact of sampling (choice of data) on Parametric Correlations



We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels: $\boxed{r_0^{so} = F(V_0^{so})}$

Spherical Woods-Saxon - Correlations V_o^{so} vs. r_o^{so}

- Impact of sampling (choice of data) on Parametric Correlations



We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with the six lowest levels: $r_0^{so} = F(V_0^{so})$

Sampling and Parametric Correlations

[Illustrations for Skyrme SIII Hartree-Fock Hamiltonian]

Parameter-Correlations and Correlation Matrix [HF]

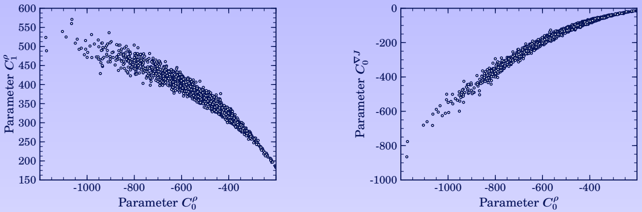
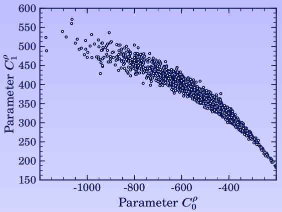


Illustration analogous to the preceding one; here Skyrme Hartree-Fock

Parameter-Correlations and Correlation Matrix [HF]



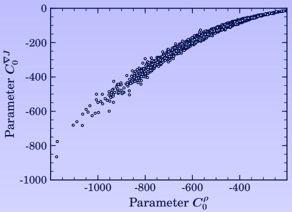
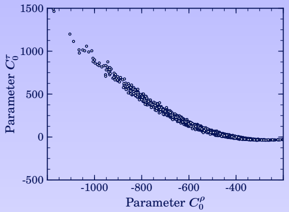


Illustration analogous to the preceding one; here Skyrme Hartree-Fock

Correlation matrix for the Skyrme-Hartree-Fock Hamiltonian parameters

	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

Parameter-Correlations and Correlation Matrix [HF]



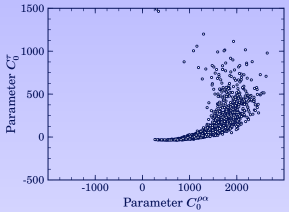


Illustration analogous to the preceding one; here Skyrme Hartree-Fock

Correlation matrix for the Skyrme-Hartree-Fock Hamiltonian parameters

	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

Parameter-Correlations in Skyrme-HF

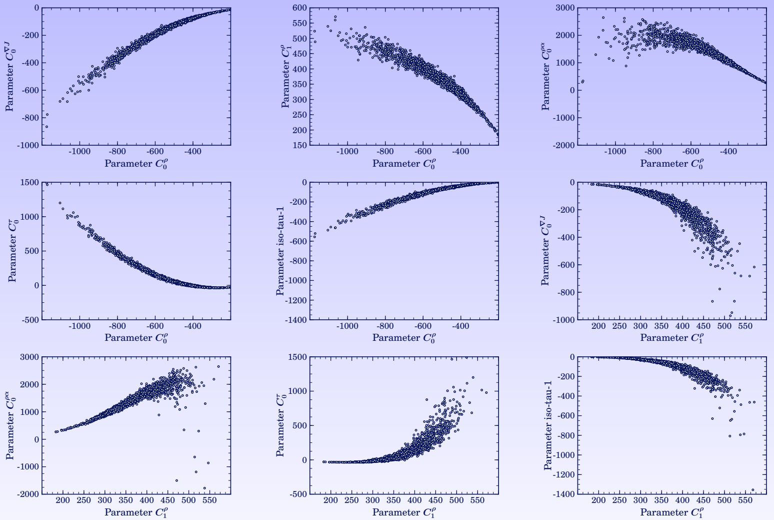


Illustration suggesting that there are rather very few independent parameters

The Following Messages

The Following Messages are intended

The Following Messages are intended for Mature Audiences

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

Skyrme-Hartree-Fock $\hat{V}_{int}(\{p\}) = \hat{v}_{Skyrme}(\vec{r}_i, \vec{r}_j)$

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

$$
\begin{aligned}
v^{tensor}(\vec{r}_i,\vec{r}_j) = {} & \tfrac{1}{2}\,t_e\,\big\{\,[\,3(\hat{\sigma}_i\cdot\hat{k}')\,(\sigma_j\cdot\hat{k}') - (\hat{\sigma}_i\cdot\hat{\sigma}_j)(\hat{k}')^2\,]\,\delta(\vec{r}_{ij}) \\
& + \;\delta(\vec{r}_{ij})\;\;[\,3(\hat{\sigma}_i\cdot\hat{k}\,)\;(\hat{\sigma}_j\cdot\hat{k}\;) - (\hat{\sigma}_i\cdot\hat{\sigma}_j)(\hat{k}\,)^2\,]\big\} \\
& + \; t_o\;\big\{3(\sigma_i\cdot\hat{k}')\,\delta(\vec{r}_{ij})(\hat{\sigma}_j\cdot\hat{k}) - (\hat{\sigma}_i\cdot\hat{\sigma}_j)\big[\hat{k}'\big]\cdot\big[\delta(\vec{r}_{ij})\hat{k}\big]\big\}
\end{aligned}
$$

Skyrme-Hartree-Fock $\hat{V}_{int}(\{p\}) = \hat{v}_{Skyrme}(\vec{r}_i, \vec{r}_j)$

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

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

12 Params.: $\{p\} \overset{df}{=} \left\{\{t_0, t_1, t_2, t_3\};\;\{x_0, x_1, x_2, x_3\};\;\{W_0\};\;\{t_e, t_o\};\;\{\alpha\}\right\}$

• In a comprehensive study Carlsson, Dobaczewski and Kortelainen introduce Skyrme nuclear density functionals up to the sixth order (the standard Skyrme is of second order)

Skyrme-HF in the EDF Formulation up to N^3LO

- In a comprehensive study Carlsson, Dobaczewski and Kortelainen introduce Skyrme nuclear density functionals up to the sixth order (the standard Skyrme is of second order)

- Their total energy density contains all these rather than ∼15 terms

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

where $C^{m'I',n'L'v'J'}_{mI,nLvJ,Q}$ are corresponding necessary coupling constants

Skyrme-HF in the EDF Formulation up to N^3LO

• In a comprehensive study Carlsson, Dobaczewski and Kortelainen introduce Skyrme nuclear density functionals up to the sixth order (the standard Skyrme is of second order)

• Their total energy density contains all these rather than ∼15 terms

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

where $C^{m'I',n'L'v'J'}_{mI,nLvJ,Q}$ are corresponding necessary coupling constants

• It is instructive to think about the extentions of the EDF based approaches in terms of the increasing number of coupling constants

• In a comprehensive study Carlsson, Dobaczewski and Kortelainen introduce Skyrme nuclear density functionals up to the sixth order (the standard Skyrme is of second order)

• Their total energy density contains all these rather than ∼15 terms

$$\mathcal{H}(\vec{r}) = \sum_{\substack{m'l',n'L'v'J' \\ ml,nLvJ,Q}} C_{ml,nLvJ,Q}^{m'l',n'L'v'J'} \times T_{ml,nLvJ,Q}^{m'l',n'L'v'J'}(\vec{r}),$$

where $C_{ml,nLvJ,Q}^{m'l',n'L'v'J'}$ are corresponding necessary coupling constants

• It is instructive to think about the extentions of the EDF based approaches in terms of the increasing number of coupling constants

• ... in view of all the couplings present already at the leading order formulations which suggest a totally ill-posed inverse problem ⟶⟶

Skyrme-HF in the EDF Formulation up to N^{3}LO

• Numbers of terms depending on the time-even and time-odd densities are given separately. The last two columns give numbers of terms when the Galilean or gauge[1] invariance is assumed, respectively.

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^{3}LO	2x312	2x346	2x658	2x188	2x97
	624	692	1316	376	194

• Let us observe a very fast-growing number of terms. To take into account both isospin channels, the number of terms is multiplied by a factor of two

[1] *For comments about Skyrme HF gauge invariance cf. e.g. J. Dobaczewski and J. Dudek, PRC 52 (1995) 1827*

Parametric Correlations - Partial Conclusions

- Parametric correlations are overwhelmingly present and - as it is very well known - they imply an ill-posedeness of the inverse problem

Parametric Correlations - Partial Conclusions

- Parametric correlations are overwhelmingly present and - as it is very well known - they imply an ill-posedeness of the inverse problem

- The solutions of an ill-posed inverse problem are generally unstable

Parametric Correlations - Partial Conclusions

• Parametric correlations are overwhelmingly present and - as it is very well known - they imply an ill-posedeness of the inverse problem

• The solutions of an ill-posed inverse problem are generally unstable

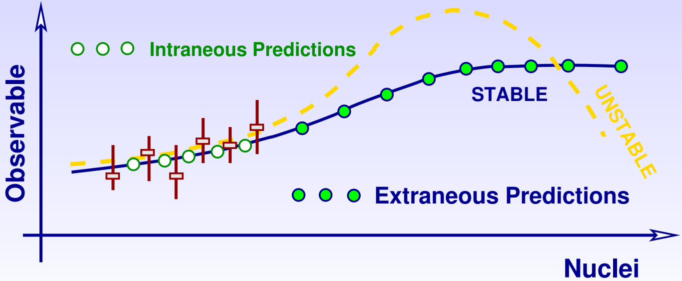
• Lack of stability means, by definition, that a small modification in the input causes a big (perhaps exponential) modification on output

Parametric Correlations - Partial Conclusions

• Parametric correlations are overwhelmingly present and - as it is very well known - they imply an ill-posedeness of the inverse problem

• The solutions of an ill-posed inverse problem are generally unstable

• Lack of stability means, by definition, that a small modification in the input causes a big (perhaps exponential) modification on output



Part IV

Ill-Posed Inverse Problem in Nuclear Theories [Regularisation, Singular Value Decomposition]

A Powerful Tool: Singular-Value Decomposition

A Powerful Tool: Singular-Value Decomposition

• We have demonstrated that the finding the parameters of the Hamiltonian is equivalent to solving the algebraic *Inverse Problem:*

$$\boxed{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \;\text{ with }\; \mathcal{A} = \mathbf{J} \cdot \mathbf{J}^{\mathbf{T}} \;\text{ where }\; \mathbf{J} \equiv \text{Jacobian}}$$

A Powerful Tool: Singular-Value Decomposition

• We have demonstrated that the finding the parameters of the Hamiltonian is equivalent to solving the algebraic *Inverse Problem:*

$$\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \;\text{ with }\; \mathcal{A} = \mathbf{J} \cdot \mathbf{J}^{\mathbf{T}} \;\text{ where }\; \mathbf{J} \equiv \text{Jacobian}$$

• One can demonstrate that an arbitrary rectangular $m \times d$ matrix J can be decomposed as a product of three matrices (D-diagonal)

$$\mathbf{J} = \mathbf{U}\,\mathbf{D}\,\mathbf{V}^{\mathbf{T}} \text{ with } \mathbf{U} \in \mathbb{R}^{\mathbf{m}\times\mathbf{m}},\; \mathbf{V} \in \mathbb{R}^{\mathbf{d}\times\mathbf{d}},\; \mathbf{D} \in \mathbb{R}^{\mathbf{m}\times\mathbf{d}}$$

A Powerful Tool: Singular-Value Decomposition

- We have demonstrated that the finding the parameters of the Hamiltonian is equivalent to solving the algebraic *Inverse Problem:*

$$\boxed{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \;\text{ with }\; \mathcal{A} = \mathbf{J} \cdot \mathbf{J}^{\mathbf{T}} \;\text{ where }\; \mathbf{J} \equiv \text{Jacobian}}$$

- One can demonstrate that an arbitrary rectangular $m \times d$ matrix J can be decomposed as a product of three matrices (D-diagonal)

$$\boxed{\mathbf{J} = \mathbf{U}\,\mathbf{D}\,\mathbf{V}^{\mathbf{T}} \text{ with } \mathbf{U} \in \mathbb{R}^{\mathbf{m}\times\mathbf{m}},\; \mathbf{V} \in \mathbb{R}^{\mathbf{d}\times\mathbf{d}},\; \mathbf{D} \in \mathbb{R}^{\mathbf{m}\times\mathbf{d}}}$$

- Diagonal elements, δ_i, are called "singular values" and we have

$$D = \text{diag}\{\underbrace{\delta_1, \delta_2, \;\ldots\; \delta_{min(m,d)}}_{\text{decreasing order}}\}$$

A Powerful Tool: Singular-Value Decomposition

• We have demonstrated that the finding the parameters of the Hamiltonian is equivalent to solving the algebraic *Inverse Problem:*

$$\boxed{\mathcal{P} = \mathcal{A}^{-1} \cdot \mathcal{D} \;\text{ with }\; \mathcal{A} = \mathbf{J} \cdot \mathbf{J}^{\mathbf{T}} \;\text{ where }\; \mathbf{J} \equiv \text{Jacobian}}$$

• One can demonstrate that an arbitrary rectangular $m \times d$ matrix J can be decomposed as a product of three matrices (D-diagonal)

$$\boxed{\mathbf{J} = \mathbf{U}\,\mathbf{D}\,\mathbf{V}^{\mathbf{T}} \text{ with } \mathbf{U} \in \mathbb{R}^{\mathbf{m}\times\mathbf{m}},\; \mathbf{V} \in \mathbb{R}^{\mathbf{d}\times\mathbf{d}},\; \mathbf{D} \in \mathbb{R}^{\mathbf{m}\times\mathbf{d}}}$$

• Diagonal elements, δ_i, are called "singular values" and we have

$$D = \mathrm{diag}\{\underbrace{\delta_1, \delta_2, \;\ldots\; \delta_{min(m,d)}}_{\text{decreasing order}}\}$$

• We find easily that

$$\mathbf{J}^{\mathbf{T}} = \mathbf{V} \cdot \mathbf{D}^{\mathbf{T}} \cdot \mathbf{U}^{\mathbf{T}} \;\text{ where }\; \mathbf{D}^{\mathbf{T}} = \mathbf{diag}\{\tfrac{1}{\delta_1}, \tfrac{1}{\delta_2}, \;\ldots\; \tfrac{1}{\delta_d}; 0, 0, \;\ldots\; 0\}$$

- Let us come back to the shown earlier χ^2-minimum condition:

$$\frac{\partial \chi^2}{\partial \mathbf{p_j}} \;\; \rightarrow \;\; (\mathbf{J}^{\mathbf{T}}\mathbf{J}) \cdot \mathcal{P} = \mathcal{D}$$

- Let us come back to the shown earlier χ^2-minimum condition:

$$\frac{\partial\chi^2}{\partial \mathbf{p_j}} \;\rightarrow\; (\mathbf{J}^\mathbf{T}\mathbf{J})\cdot\mathcal{P} = \mathcal{D}$$

- From the Singular-Value Decomposition for matrix J it follows

$$\boxed{(\mathbf{J}^\mathbf{T}\mathbf{J})^{-1} = \mathbf{V}\cdot(1/\delta^2)\cdot\mathbf{V}^\mathbf{T}}$$

Fitting, Inverse Problem and Confidence Intervals

- Let us come back to the shown earlier χ^2-minimum condition:

$$\frac{\partial \chi^2}{\partial \mathbf{p_j}} \;\rightarrow\; (\mathbf{J^T J}) \cdot \mathcal{P} = \mathcal{D}$$

- From the Singular-Value Decomposition for matrix J it follows

$$\boxed{(\mathbf{J^T J})^{-1} = \mathbf{V} \cdot (1/\delta^2) \cdot \mathbf{V^T}}$$

- Independently one derives the expression for the correlation matrix

$$\boxed{\mathrm{Corr}(\mathcal{P}_i, \mathcal{P}_j) \rightarrow \langle (\mathcal{P}_i - \langle \mathcal{P}_i \rangle) \cdot (\mathcal{P}_j - \langle \mathcal{P}_j \rangle) \rangle \sim \chi^2(\mathbf{p})\, (\mathbf{J^T J})^{-1}_{ij}}$$

Fitting, Inverse Problem and Confidence Intervals

- Let us come back to the shown earlier χ^2-minimum condition:

$$\frac{\partial \chi^2}{\partial \mathbf{p_j}} \;\; \rightarrow \;\; (\mathbf{J^T J}) \cdot \mathcal{P} = \mathcal{D}$$

- From the Singular-Value Decomposition for matrix J it follows

$$\boxed{(\mathbf{J^T J})^{-1} = \mathbf{V} \cdot (1/\delta^2) \cdot \mathbf{V^T}}$$

- Independently one derives the expression for the correlation matrix

$$\boxed{\mathrm{Corr}(\mathcal{P}_\mathrm{i}, \mathcal{P}_\mathrm{j}) \rightarrow \langle(\mathcal{P}_\mathrm{i} - \langle\mathcal{P}_\mathrm{i}\rangle) \cdot (\mathcal{P}_\mathrm{j} - \langle\mathcal{P}_\mathrm{j}\rangle)\rangle \sim \chi^2(\mathbf{p})\, (\mathbf{J^T J})^{-1}_{\mathbf{ij}}}$$

- If one or more $\delta_k \rightarrow 0$ then $(J^T J)^{-1} \rightarrow \infty$ and generally, the confidence intervals of all parameters diverge [null predictive power]

- One may show that the parametric instability of the solutions of the inverse problem is directly proportional to the condition number

$$\boxed{\mathrm{Cond}(A) \equiv \delta_{\text{biggest}}/\delta_{\text{smallest}}}$$

Sing.-Value Decomposition & Conditional Number

• One may show that the parametric instability of the solutions of the inverse problem is directly proportional to the condition number

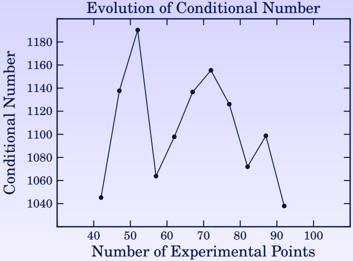
$$\boxed{\mathrm{Cond}(A) \equiv \delta_{\mathsf{biggest}}/\delta_{\mathsf{smallest}}}$$

• For the well posed Inverse Problem, Cond(A) is of the order of 1

- One may show that the parametric instability of the solutions of the inverse problem is directly proportional to the condition number

$$\boxed{\mathrm{Cond}(A) \equiv \delta_{\mathsf{biggest}}/\delta_{\mathsf{smallest}}}$$

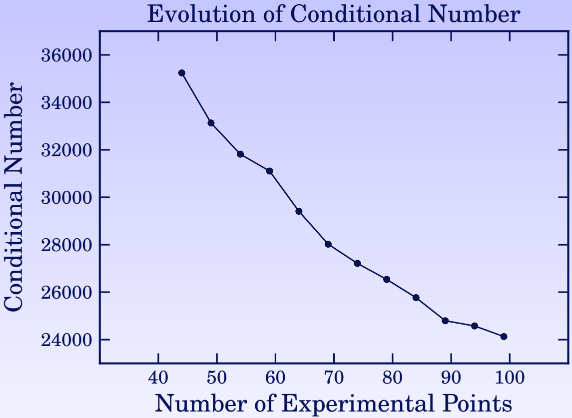
- For the well posed Inverse Problem, Cond(A) is of the order of 1



Conditional number of the SLY4-type Hamiltonian, parameters fitted to the single-particle energies only, observe HUGE values of Cond(A)

The Catastrophe of Fitting to the Masses

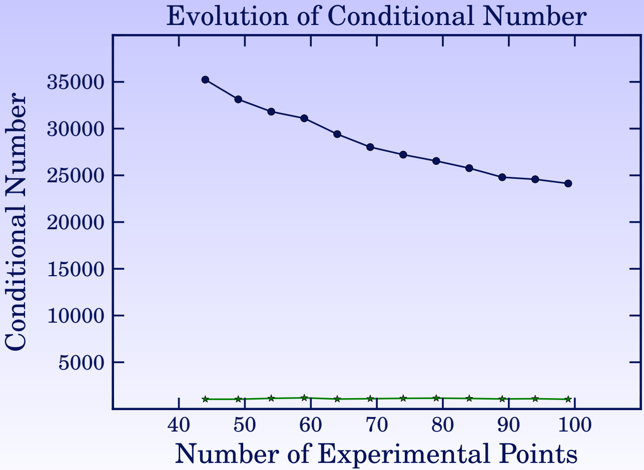
- When fitting the Skyrme Hartree-Fock parameters to the single particle energies and to the masses we obtain Cond(A)$\sim 10^5$



Conditional number of the SLY4-type Hamiltonian, parameters fitted to the single-particle energies and masses

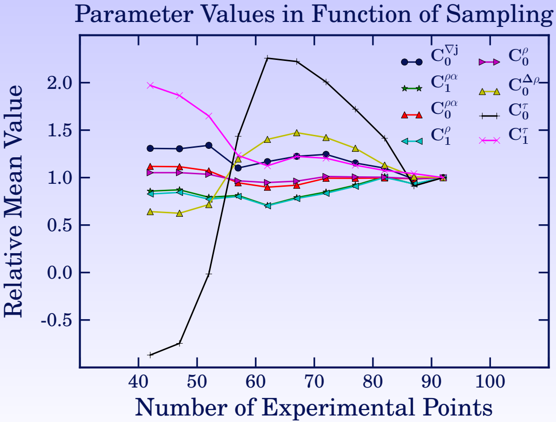
Smaller Theory Errors vs. Bigger Predictive-Power

• Constraining theory errors may help stabilising theory predictions: The necessary although not sufficient condition of model's stability



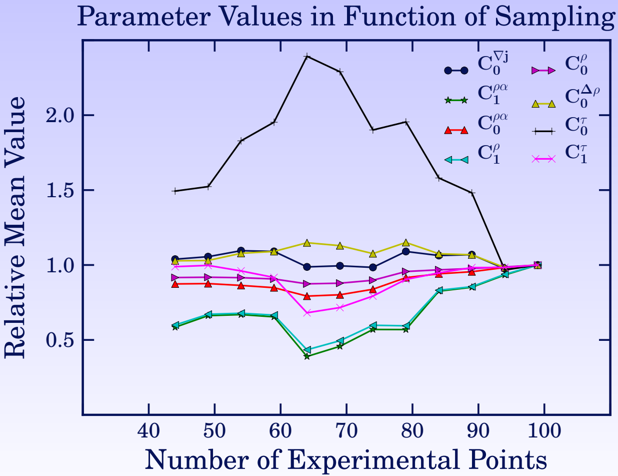
- Parameters expressed using Density-Functional representation

$$\{p\} \leftrightarrow \{C_t^{\rho 0},\ C_t^{\rho\alpha},\ C_t^{\Delta\rho},\ C_t^{\tau},\ C_t^{J},\ C_t^{\nabla J},\ t_e,\ t_o \text{ and } \alpha\ \}$$



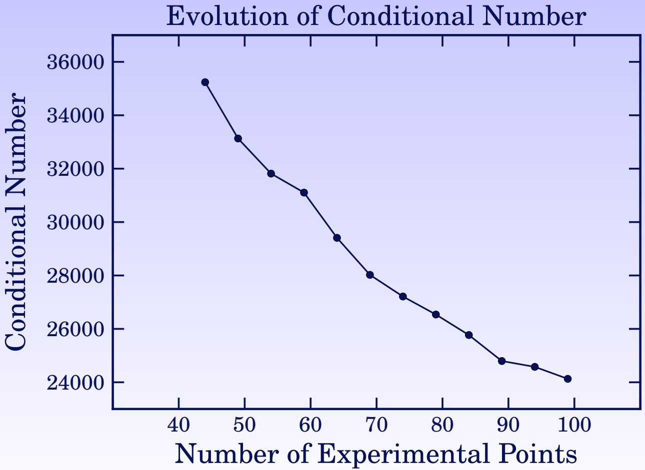
Smaller Theory Errors vs. Bigger Predictive-Power

- Constraining theory errors may help stabilising theory predictions: The necessary although not sufficient condition of model's stability



Smaller Theory Errors vs. Bigger Predictive-Power

- Constraining theory errors may help stabilising theory predictions: The necessary although not sufficient condition of model's stability



Part V

Controlling Experiment with the Help of Noise Simulations

Single-Particle Levels - Noise-Simulation Example

• Consider a single particle spectrum $\{e^o_\nu\} \leftrightarrow H\varphi^o_\nu = e^o_\nu\,\varphi^o_\nu$ obtained with the 'optimal' set of parameters $\{p\}_o$ as in the preceding Table;

• Define the "pseudo-experimental" levels $\{e^{exp}_\nu\} \equiv \{e^o_\nu\}$. Applying the minimisation procedure will now reproduce those $\{e^o_\nu\}$ exactly;

• Chose one level, say $e^o_\kappa \in \{e^o_\nu\}$, and arbitrarily modify its position:

$$e^o_\kappa \to e_\kappa \equiv (e^o_\kappa - e) \;\; \text{with, say} \;\; e \in [-2, +2] \text{ MeV};$$

then refit the χ^2-test $\to$ all other levels will move to new positions

• Collect these new positions: they are functions $e_\nu = e_\nu(e_\kappa)$, below referred to as 'error response functions' $\to$ see illustrations

Single-Particle Levels - Noise-Simulation Example

• Consider a single particle spectrum $\{e^o_\nu\} \leftrightarrow H\varphi^o_\nu = e^o_\nu\,\varphi^o_\nu$ obtained with the 'optimal' set of parameters $\{p\}_o$ as in the preceding Table;

• Define the "pseudo-experimental" levels $\{e^{exp}_\nu\} \equiv \{e^o_\nu\}$. Applying the minimisation procedure will now reproduce those $\{e^o_\nu\}$ exactly;

• Chose one level, say $e^o_\kappa \in \{e^o_\nu\}$, and arbitrarily modify its position:

$$e^o_\kappa \to e_\kappa \equiv (e^o_\kappa - e) \;\text{ with, say }\; e \in [-2, +2] \text{ MeV};$$

then refit the χ^2-test $\to$ all other levels will move to new positions

• Collect these new positions: they are functions $e_\nu = e_\nu(e_\kappa)$, below referred to as 'error response functions' $\to$ see illustrations

Single-Particle Levels - Noise-Simulation Example

• Consider a single particle spectrum $\{e^o_\nu\} \leftrightarrow H\varphi^o_\nu = e^o_\nu\,\varphi^o_\nu$ obtained with the 'optimal' set of parameters $\{p\}_o$ as in the preceding Table;

• Define the "pseudo-experimental" levels $\{e^{exp}_\nu\} \equiv \{e^o_\nu\}$. Applying the minimisation procedure will now reproduce those $\{e^o_\nu\}$ exactly;

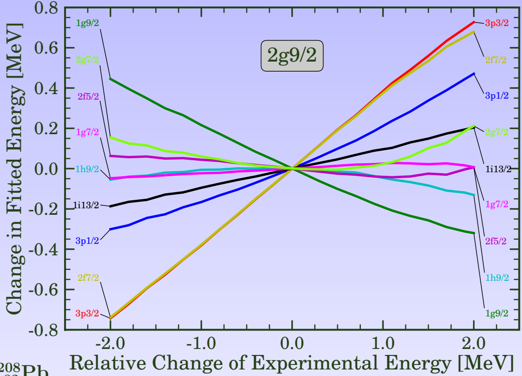
• Chose one level, say $e^o_\kappa \in \{e^o_\nu\}$, and arbitrarily modify its position:

$$e^o_\kappa \to e_\kappa \equiv (e^o_\kappa - e) \;\; \text{with, say} \;\; e \in [-2, +2] \text{ MeV};$$

then refit the χ^2-test $\to$ all other levels will move to new positions

• Collect these new positions: they are functions $e_\nu = e_\nu(e_\kappa)$, below referred to as 'error response functions' $\to$ see illustrations

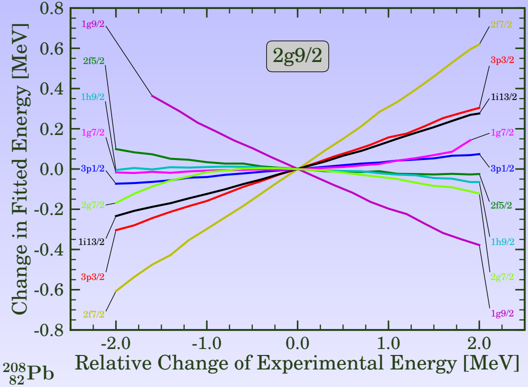
Example: Error Response Functions to $2g_{9/2}$-Orbital



$^{208}_{82}$Pb

To determine precisely the parameters through fitting the energies of $3p_{3/2}$, $2f_{7/2}$ etc. the right position of $2g_{9/2}$ must be analyzed particularly carefully (associated spectroscopic factors precise, particle vibration subtracted, pairing effect subtracted)

Example: Alternative Representation for $2g_{9/2}$-Orbital



<u>Attention:</u> The figure may look similar but it contains a totally opposite information: All the curves represent the $2g_{9/2}$-level - this is how the fitting will modify $2g_{9/2}$ if we vary the indicated levels

Conclusions from Error Response-Function Tests

- Observe rather precise indications as to *'which levels influence which'* what allows to discuss the experimental strategies precisely

- The low-ℓ orbitals (such as $3p_{1/2}$, $3p_{3/2}$) have relatively small impact on the error-response functions ...

- ... while some pairs of orbitals couple very strongly

- The highest-ℓ orbitals do not couple in the strongest way

- ... all that in a particular case presented; analysis of this type may require a case-by-case mode of operating...

Part VI

Predictive Power and Over-Fitting Mechanism

- Let us calculate $\{e_\mu\}$-levels for a given W-S parameter set, here:

Woods-Saxon parameters for the neutrons in ^{208}Pb reproduce the experimental levels with the r.m.s. deviation of 0.164 MeV and maximum error of 0.353 MeV.

V_o^c	r_o^c	a_o^c	λ	r_o^{so}	a^{so}
-39.520	1.371	0.694	26.133	1.255	0.500

A Realistic Toy Model - Noise-Simulation Example

- Let us calculate $\{e_\mu\}$-levels for a given W-S parameter set, here:

Woods-Saxon parameters for the neutrons in ^{208}Pb reproduce the experimental levels with the r.m.s. deviation of 0.164 MeV and maximum error of 0.353 MeV.

V_o^c	r_o^c	a_o^c	λ	r_o^{so}	a^{so}
-39.520	1.371	0.694	26.133	1.255	0.500

- We can treat $\{e_\mu\}$ 'as experimental'; by trying to reproduce them through fitting we know an exact solution!

- Let us calculate $\{e_\mu\}$-levels for a given W-S parameter set, here:

Woods-Saxon parameters for the neutrons in ^{208}Pb reproduce the experimental levels with the r.m.s. deviation of 0.164 MeV and maximum error of 0.353 MeV.

V_o^c	r_o^c	a_o^c	λ	r_o^{so}	a^{so}
-39.520	1.371	0.694	26.133	1.255	0.500

- We can treat $\{e_\mu\}$ 'as experimental'; by trying to reproduce them through fitting we know an exact solution!

- Extra advantage: we may introduce the notion of 'noise', usually a random variable distributed according to a certain probability fct.

A Realistic Toy Model - Noise-Simulation Example

- Let us calculate $\{e_\mu\}$-levels for a given W-S parameter set, here:

Woods-Saxon parameters for the neutrons in ^{208}Pb reproduce the experimental levels with the r.m.s. deviation of 0.164 MeV and maximum error of 0.353 MeV.

V_o^c	r_o^c	a_o^c	λ	r_o^{so}	a^{so}
-39.520	1.371	0.694	26.133	1.255	0.500

- We can treat $\{e_\mu\}$ ‘as experimental’; by trying to reproduce them through fitting we know an exact solution!

- Extra advantage: we may introduce the notion of ‘noise’, usually a random variable distributed according to a certain probability fct.

- We will obtain the response of all the levels to a ‘linear noise’ - vary a level position within a window and refit the H-parameters $\{p\}$

Unprecedented Precision of the Fits: 10^{-1} keV!

→ The standard Woods-Saxon Hamiltonian has been used:

No.	E_{calc}	E_{exp}	Level	Err.(th-exp)
1.	-15.300	-15.300	$1p_{3/2}$	-0.0001
2.	-9.000	-9.000	$1p_{1/2}$	-0.0001
3.	-0.600	-0.600	$1d_{5/2}$	0.0000
4.	-0.100	-0.100	$2s_{1/2}$	0.0000
5.	4.400	4.400	$1d_{3/2}$	0.0001

Unprecedented Precision of the Fits: 10^{-1} keV!

→ The standard Woods-Saxon Hamiltonian has been used:

No.	E_{calc}	E_{exp}	Level	Err.(th-exp)
1.	-15.300	-15.300	$1p_{3/2}$	-0.0001
2.	-9.000	-9.000	$1p_{1/2}$	-0.0001
3.	-0.600	-0.600	$1d_{5/2}$	0.0000
4.	-0.100	-0.100	$2s_{1/2}$	0.0000
5.	4.400	4.400	$1d_{3/2}$	0.0001

→ Couple of questions may come to one's mind...:

Unprecedented Precision of the Fits: 10^{-1} keV!

→ The standard Woods-Saxon Hamiltonian has been used:

No.	E_{calc}	E_{exp}	Level	Err.(th-exp)
1.	-15.300	-15.300	$1p_{3/2}$	-0.0001
2.	-9.000	-9.000	$1p_{1/2}$	-0.0001
3.	-0.600	-0.600	$1d_{5/2}$	0.0000
4.	-0.100	-0.100	$2s_{1/2}$	0.0000
5.	4.400	4.400	$1d_{3/2}$	0.0001

→ Couple of questions may come to one's mind...:

- *Is this property limited to one single nucleus?*

Unprecedented Precision of the Fits: 10^{-1} keV!

→ The standard Woods-Saxon Hamiltonian has been used:

No.	E_{calc}	E_{exp}	Level	Err.(th-exp)
1.	-15.300	-15.300	$1p_{3/2}$	-0.0001
2.	-9.000	-9.000	$1p_{1/2}$	-0.0001
3.	-0.600	-0.600	$1d_{5/2}$	0.0000
4.	-0.100	-0.100	$2s_{1/2}$	0.0000
5.	4.400	4.400	$1d_{3/2}$	0.0001

→ Couple of questions may come to one's mind...:

- *Is this property limited to one single nucleus?* *Not at all!*

Unprecedented Precision of the Fits: 10^{-1} keV!

→ The standard Woods-Saxon Hamiltonian has been used:

No.	E_{calc}	E_{exp}	Level	Err.(th-exp)
1.	-15.300	-15.300	$1p_{3/2}$	-0.0001
2.	-9.000	-9.000	$1p_{1/2}$	-0.0001
3.	-0.600	-0.600	$1d_{5/2}$	0.0000
4.	-0.100	-0.100	$2s_{1/2}$	0.0000
5.	4.400	4.400	$1d_{3/2}$	0.0001

→ Couple of questions may come to one's mind...:

- *Is this property limited to one single nucleus?* *Not at all!*
- *Can a simple phenomenology achieve the precision of hundreds of electronvolts in nearly all doubly-magic nuclei?*

Unprecedented Precision of the Fits: 10^{-1} keV!

→ The standard Woods-Saxon Hamiltonian has been used:

No.	E_{calc}	E_{exp}	Level	Err.(th-exp)
1.	-15.300	-15.300	$1p_{3/2}$	-0.0001
2.	-9.000	-9.000	$1p_{1/2}$	-0.0001
3.	-0.600	-0.600	$1d_{5/2}$	0.0000
4.	-0.100	-0.100	$2s_{1/2}$	0.0000
5.	4.400	4.400	$1d_{3/2}$	0.0001

→ Couple of questions may come to one's mind...:

- *Is this property limited to one single nucleus?* *Not at all!*
- *Can a simple phenomenology achieve the precision of hundreds of electronvolts in nearly all doubly-magic nuclei?* *Is it trivial?*

Unprecedented Precision of the Fits: 10^{-1} keV!

→ The standard Woods-Saxon Hamiltonian has been used:

No.	E_{calc}	E_{exp}	Level	Err.(th-exp)
1.	-15.300	-15.300	$1p_{3/2}$	-0.0001
2.	-9.000	-9.000	$1p_{1/2}$	-0.0001
3.	-0.600	-0.600	$1d_{5/2}$	0.0000
4.	-0.100	-0.100	$2s_{1/2}$	0.0000
5.	4.400	4.400	$1d_{3/2}$	0.0001

→ Couple of questions may come to one's mind...:

- *Is this property limited to one single nucleus? Not at all!*
- *Can a simple phenomenology achieve the precision of hundreds of electronvolts in nearly all doubly-magic nuclei? Is it trivial?*
- *What is the mathematical/physical significance of the result?*

'Over-Fitting' - What Does It Imply?

'Over-Fitting' - What Does It Imply?

- Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always: 2 conclusions

'Over-Fitting' - What Does It Imply?

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always: 2 conclusions

a. The physics modeling is not totally wrong, but also

'Over-Fitting' - What Does It Imply?

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always: 2 conclusions

a. The physics modeling is not totally wrong, but also
b. The number of data points is too weak to constrain the model.

'Over-Fitting' - What Does It Imply?

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always: 2 conclusions

a. The physics modeling is not totally wrong, but also
b. The number of data points is too weak to constrain the model.

• We have two ways out of the overfitting problem:

A. We modify the model <u>de</u>creasing the number of parameters, or:

'Over-Fitting' - What Does It Imply?

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always: 2 conclusions

a. The physics modeling is not totally wrong, but also
b. The number of data points is too weak to constrain the model.

• We have two ways out of the overfitting problem:

A. We modify the model <u>de</u>creasing the number of parameters, or:
B. We increase the number of data points (if we can...)

'Over-Fitting' - What Does It Imply?

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always: 2 conclusions

a. The physics modeling is not totally wrong, but also
b. The number of data points is too weak to constrain the model.

• We have two ways out of the overfitting problem:

A. We modify the model <u>de</u>creasing the number of parameters, or:
B. We increase the number of data points (if we can...)

• We have already used out all the experimental data: as theorists we can only modify the model...

'Over-Fitting' - What Does It Imply?

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always: 2 conclusions

a. The physics modeling is not totally wrong, but also
b. The number of data points is too weak to constrain the model.

• We have two ways out of the overfitting problem:

A. We modify the model decreasing the number of parameters, or:
B. We increase the number of data points (if we can...)

• We have already used out all the experimental data: as theorists we can only modify the model...

• We improve the model by reducing the number of parameters

Summary and Conclusions (I)

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem
b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem
b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

• Inexact theories involve always theory uncertainties (which must be estimated) and related probability distributions can be modelled

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem
b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

• Inexact theories involve always theory uncertainties (which must be estimated) and related probability distributions can be modelled

• In the future theoretical approaches: Theory provides not only the numerical predictions but also probability distributions of the associated uncertainties

Summary and Conclusions (I)

• The concept of the 'exact theories' gives an extremely useful insight and guidance into the functioning of Mathematical Predicting:

a. Intraneous Predictions: Works usually very well no matter the well-posed or ill-posed Inverse Problem
b. Extraneous Predictions: An ill-posed Inverse Problem generally eliminates the possibility of any stable predictive power

• Inexact theories involve always theory uncertainties (which must be estimated) and related probability distributions can be modelled

• In the future theoretical approaches: Theory provides not only the numerical predictions but also probability distributions of the associated uncertainties

• We believe that quite often it is easier to estimate the uncertainties of the present theory rather than to document a new interaction term

Summary and Conclusions (II)

Summary and Conclusions (II)

- Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters
C. We increase the number of data points (if we can pay for it...)

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters
C. We increase the number of data points (if we can pay for it...)

• Suppose we have already used out all the existing experimental data: as theorists we can modify models / analyse uncertainties...

Summary and Conclusions (II)

• Even if there are too few data points and too many parameters - one generally cannot obtain $\chi^2 = 0$ solutions always:

a. The physics modeling is possibly not totally wrong, however
b. The number of data points is too weak to constrain the model

• We have at least three ways out of the Ill-Posedeness:

A. We apply one of Regularisation Methods [not discussed today]
B. We modify the model by <u>de</u>creasing the number of parameters
C. We increase the number of data points (if we can pay for it...)

• Suppose we have already used out all the existing experimental data: as theorists we can modify models / analyse uncertainties...

• In other words: We improve predictive power of our theory by reducing the number of parameters, by regularising the associated Inverse Problem, but first of all through including all interactions