Nuclear Physics Hamiltonians and the Problem of Their Predictive Power

Jerzy DUDEK

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

May 25, 2010

The question of predictive power of nuclear theories and statistical significance of theoretical predictions is a vast and quickly growing field of nuclear science

The question of predictive power of nuclear theories and statistical significance of theoretical predictions is a vast and quickly growing field of nuclear science

Today we will present merely a selection of keywords and illustrations belonging to the underlying subfield of Applied Mathematics: the Inverse Problem[1]

The question of predictive power of nuclear theories and statistical significance of theoretical predictions is a vast and quickly growing field of nuclear science

Today we will present merely a selection of keywords and illustrations belonging to the underlying subfield of Applied Mathematics: the Inverse Problem[1]

Attention: The term 'Inverse Problem' of Applied Mathematics has NOT much in common with its traditional meaning in quantum mechanics

The question of predictive power of nuclear theories and statistical significance of theoretical predictions is a vast and quickly growing field of nuclear science

Today we will present merely a selection of keywords and illustrations belonging to the underlying subfield of Applied Mathematics: the Inverse Problem[1]

Attention: The term 'Inverse Problem' of Applied Mathematics has NOT much in common with its traditional meaning in quantum mechanics

[1] GOOGLE SEARCH — on Sunday, May 16, 2010, at 10:30 AM

"Nuclear Physics"	3 200 000 results (∼0.18 seconds)
"Inverse Problem"	510 000 results (∼0.36 seconds)

COLLABORATORS:

Bartłomiej SZPAK, IFJ Kraków
Marie-Geneviève PORQUET, CSNSM Orsay

Bogdan FORNAL, IFJ Kraków
Hervé MOLIQUE, DRS Strasbourg
Karolina RYBAK, DRS Strasbourg

In this presentation we use some material from the article:

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 structure theory

... as well as some material from the articles:

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. Nuclear Mean Field Hamiltonians and Factors Limiting their Predictive Power: Illustrations

B. Szpak, JD, K. Rybak, M-G, Porquet, H. Molique & B. Fornal
Int. J. Mod. Phys. E **19** *(2010) 665*

About the Bottom Line

About the Bottom Line

Suppose we let 100 theorists...

About the Bottom Line

Suppose we let 100 theorists...
... to develop 100 models of, say, the neutron drip line:

About the Bottom Line

Suppose we let 100 theorists...
... to develop 100 models of, say, the neutron drip line:

Then: there will be 100 predictions (we are individualists)

About the Bottom Line

Suppose we let 100 theorists...
... to develop 100 models of, say, the neutron drip line:

Then: there will be 100 predictions (we are individualists)
and, eventually, one of them will, perhaps, be ...

About the Bottom Line

Suppose we let 100 theorists...
... to develop 100 models of, say, the neutron drip line:

Then: there will be 100 predictions (we are individualists)
and, eventually, one of them will, perhaps, be ...
in agreement with an a posteriori experimental information.

About the Bottom Line

Suppose we let 100 theorists...
... to develop 100 models of, say, the neutron drip line:

Then: there will be 100 predictions (we are individualists)
and, eventually, one of them will, perhaps, be ...
in agreement with an a posteriori experimental information.

QUESTION: Should we congratulate the happy one?

About the Bottom Line

Suppose we let 100 theorists...
... to develop 100 models of, say, the neutron drip line:

Then: there will be 100 predictions (we are individualists)
and, eventually, one of them will, perhaps, be ...
in agreement with an a posteriori experimental information.

QUESTION: Should we congratulate the happy one?

ANSWER: Yes, we always do when one of our friends

About the Bottom Line

Suppose we let 100 theorists...
... to develop 100 models of, say, the neutron drip line:

Then: there will be 100 predictions (we are individualists)
and, eventually, one of them will, perhaps, be ...
in agreement with an a posteriori experimental information.

QUESTION: Should we congratulate the happy one?

ANSWER: Yes, we always do when one of our friends
winns a bigger amount of money in a poker game...

This and the following presentation are about:

How to help our 100 theorists to arrive at close-lying results

Part I

Nuclear Hamiltonians from the Predictive-Power Perspective

About this Particular Project: What? and Why?

About this Particular Project: What? and Why?

- **What do we usually wish to do is to learn the full truth**

$$\boxed{\hat{\mathbf{H}}\,\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \quad \leftrightarrow \quad \hat{\mathbf{H}} = \hat{\mathbf{H}}^{\mathbf{true}}}$$

About this Particular Project: What? and Why?

- **What do we usually wish to do is to learn the full truth**

$$\boxed{\hat{\mathsf{H}}\,\psi_{\mathsf{n}} = \mathsf{e}_{\mathsf{n}}\,\psi_{\mathsf{n}} \quad \leftrightarrow \quad \hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}}}$$

where by definition $\hat{\mathsf{H}}^{\mathsf{true}}$ contains <u>complete</u> information

About this Particular Project: What? and Why?

- **What do we usually wish to do is to learn the full truth**

$$\boxed{\hat{\mathsf{H}}\,\psi_{\mathsf{n}} = \mathsf{e}_{\mathsf{n}}\,\psi_{\mathsf{n}} \quad \leftrightarrow \quad \hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}}}$$

where by definition $\hat{\mathsf{H}}^{\mathsf{true}}$ contains <u>complete</u> information

- **<u>Complete</u> - both: already known <u>and</u> yet undiscovered**

About this Particular Project: What? and Why?

- **What do we usually wish to do is to learn the full truth**

$$\boxed{\hat{\mathsf{H}}\,\psi_{\mathsf{n}} = \mathsf{e}_{\mathsf{n}}\,\psi_{\mathsf{n}} \quad \leftrightarrow \quad \hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}}}$$

where by definition $\hat{\mathsf{H}}^{\mathsf{true}}$ contains <u>complete</u> information

- **<u>Complete</u> - both: already known <u>and</u> yet undiscovered**

Conclusion 1: The solutions $\{\mathsf{e}_{\mathsf{n}}, \psi_{\mathsf{n}}\}$ remain 'asymptotic' rather fictitious, ideal and to an extent unknown objects

About this Particular Project: What? and Why?

- **What do we usually wish to do is to learn the full truth**

$$\boxed{\hat{\mathsf{H}}\,\psi_{\mathsf{n}} = \mathsf{e}_{\mathsf{n}}\,\psi_{\mathsf{n}} \quad \leftrightarrow \quad \hat{\mathsf{H}} = \hat{\mathsf{H}}^{\mathsf{true}}}$$

where by definition $\hat{\mathsf{H}}^{\mathsf{true}}$ contains <u>complete</u> information

- **<u>Complete</u> - both: already known <u>and</u> yet undiscovered**

Conclusion 1: The solutions $\{\mathsf{e}_{\mathsf{n}}, \psi_{\mathsf{n}}\}$ remain 'asymptotic' rather fictitious, ideal and to an extent unknown objects

Conclusion 2: The desired truth remains unknown to us $\longrightarrow$ lack of knowledge $\longrightarrow$ ignorance imposed[#)] by nature

[#)]**... and thus well excused - because not resulting from our lazyness**

About this Particular Project: What? and Why?

About this Particular Project: What? and Why?

- **Humans do not know any efficient cure against deliberate lazyness, but they did invent a cure against forced ignorance**

About this Particular Project: What? and Why?

- **Humans do not know any efficient cure against deliberate lazyness, but they did invent a cure against forced ignorance**
- **Solution: In cases of doubt we 'parametrize our ignorance'**

About this Particular Project: What? and Why?

- **Humans do not know any efficient cure against deliberate lazyness, but they did invent a cure against forced ignorance**
- **Solution: In cases of doubt we 'parametrize our ignorance'**
- **... more precisely: estimate by all available means which answer will be more- and which less-likely 'the right solution':**

About this Particular Project: What? and Why?

- **Humans do not know any efficient cure against deliberate lazyness, but they did invent a cure against forced ignorance**
- **Solution: In cases of doubt we 'parametrize our ignorance'**
- **... more precisely: estimate by all available means which answer will be more- and which less-likely 'the right solution':**

Find <u>relative</u> probability of what we think the right answer is!

About this Particular Project: What? and Why?

• **Humans do not know any efficient cure against deliberate lazyness, but they did invent a cure against forced ignorance**

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

• **... more precisely: estimate by all available means which answer will be more- and which less-likely 'the right solution':**

Find relative probability of what we think the right answer is!

In Applied Mathematics:

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

About this Particular Project: What? and Why?

• **Humans do not know any efficient cure against deliberate lazyness, but they did invent a cure against forced ignorance**

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

• **... more precisely: estimate by all available means which answer will be more- and which less-likely 'the right solution':**

Find relative probability of what we think the right answer is!

In Applied Mathematics:

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

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

About this Particular Project: What? and Why?

• **Humans do not know any efficient cure against deliberate lazyness, but they did invent a cure against forced ignorance**

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

• **... more precisely: estimate by all available means which answer will be more- and which less-likely 'the right solution':**

Find relative probability of what we think the right answer is!

In Applied Mathematics:

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

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

Conclusion: We need to introduce probabilities of ignorance!

Combining the Theoretical and the Experimental Errors

A Few Comments about the Nature of Errors

- Theories are incomplete whereas experiments plagued with errors:

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

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

A Few Comments about the Nature of Errors

- Theories are incomplete whereas experiments plagued with errors:

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

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

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

A Few Comments about the Nature of Errors

- Theories are incomplete whereas experiments plagued with errors:

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

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

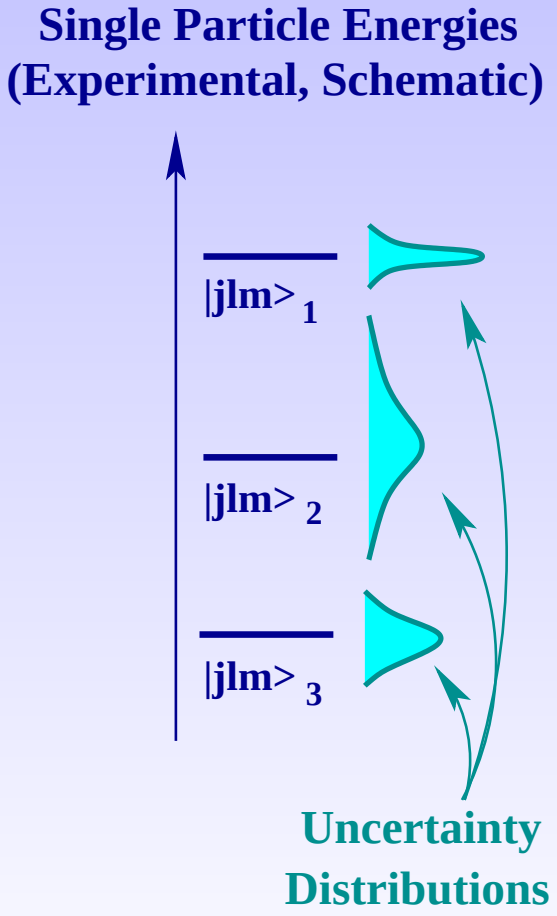
$$\chi^2(p) \sim \sum w_n \Big[\underbrace{\left(\varepsilon_n^{true} + \delta\varepsilon_n^{err}\right)}_{\textbf{Experiment}} - \underbrace{\left(e_n^{true} + \delta e_n^{err}\right)}_{\textbf{Theory}} \Big]^2 \rightarrow \frac{\partial \chi^2}{\partial p} = 0$$

- The final theory predictions will depend on the convolution

$$\mathbf{P^{pred.}(p) = \{P^{th}(e) \; * \; P^{exp}(\varepsilon)\}}$$

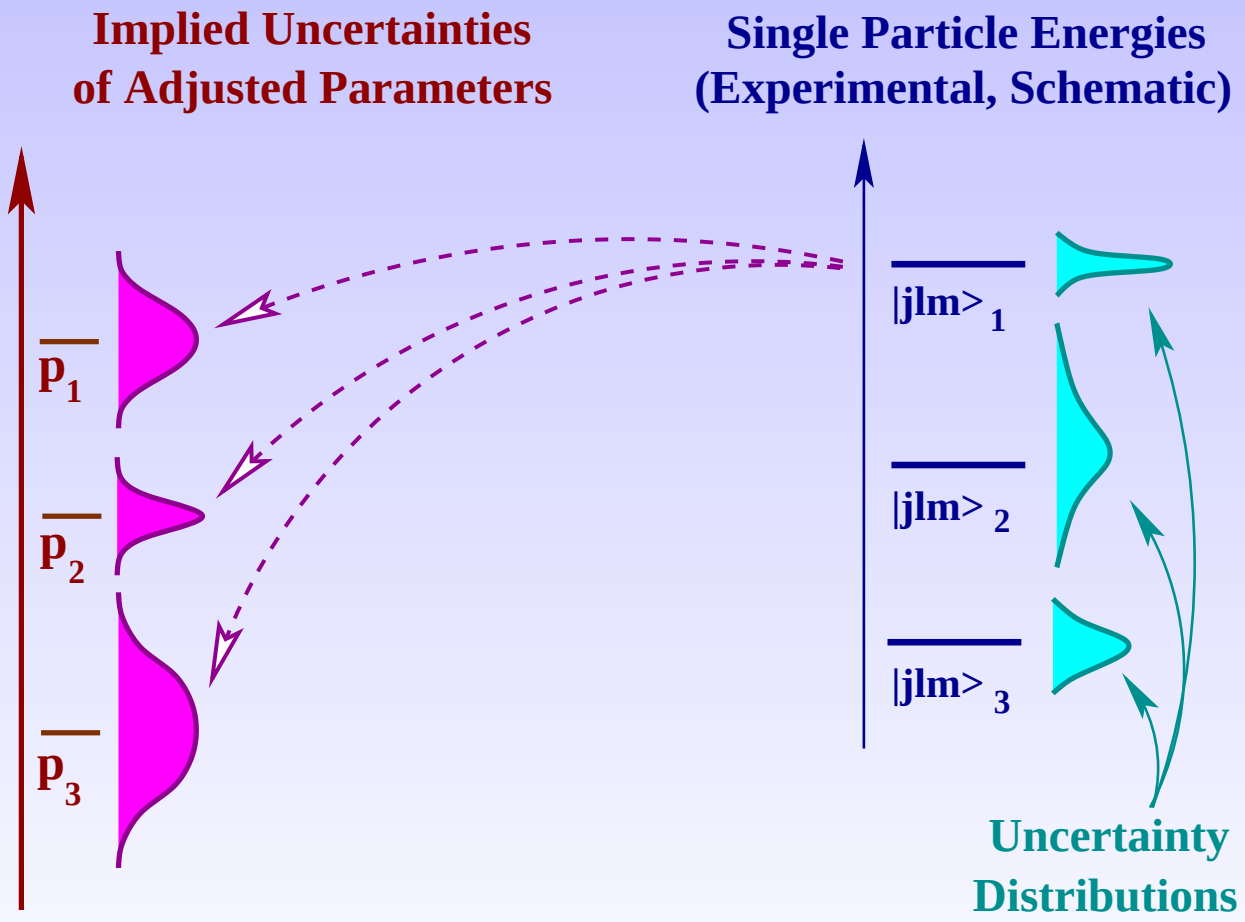
Single-Particle Levels as Probability Distributions

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



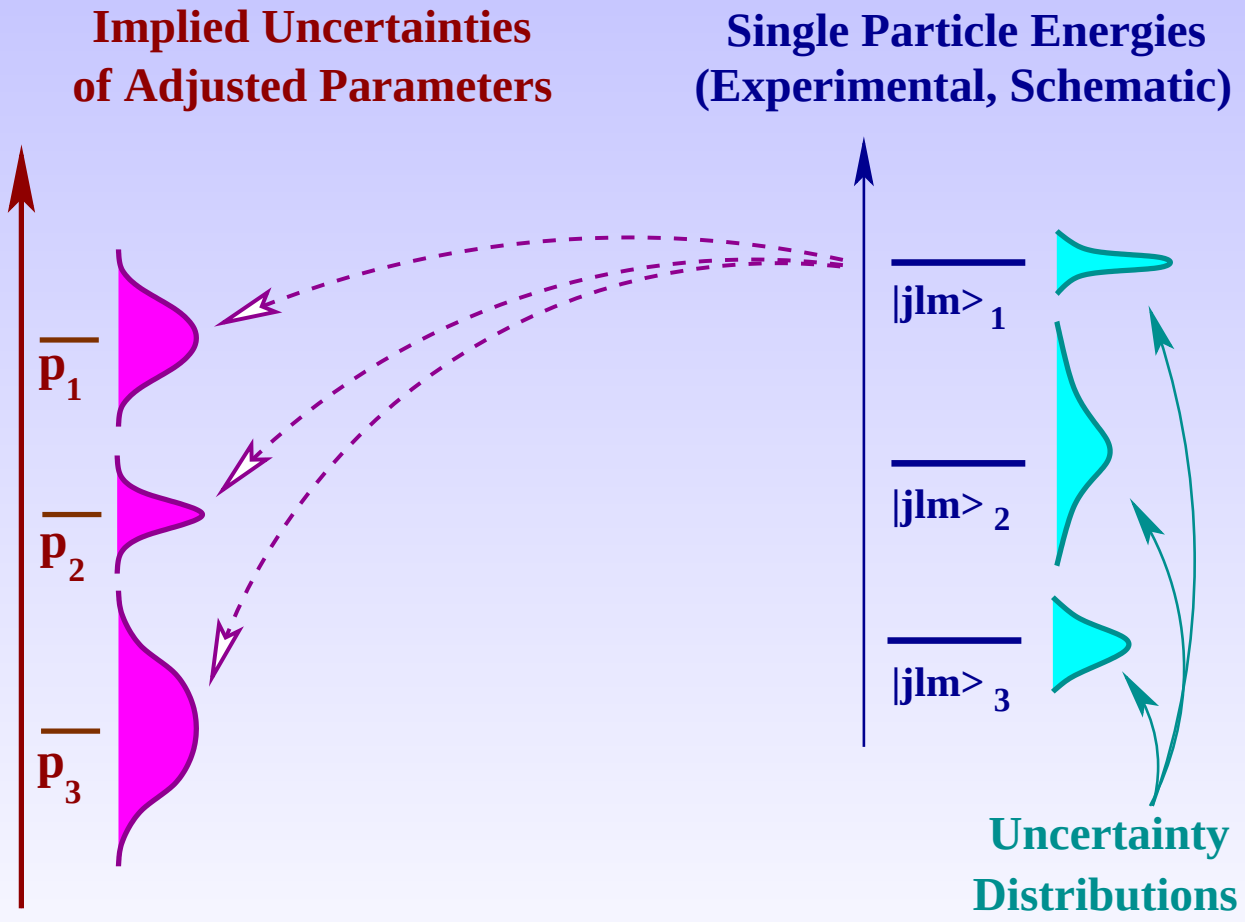
Single-Particle Levels as Probability Distributions

Uncertainties propagate to sought parameters $\{p_1, p_2, \ldots\} \equiv \{p\}$ in $\hat{H}(\hat{r}, \hat{p}; \{p\})\, \psi_n = e_n(p)\, \psi_n$: Parameters $\rightarrow$ Probability Distributions



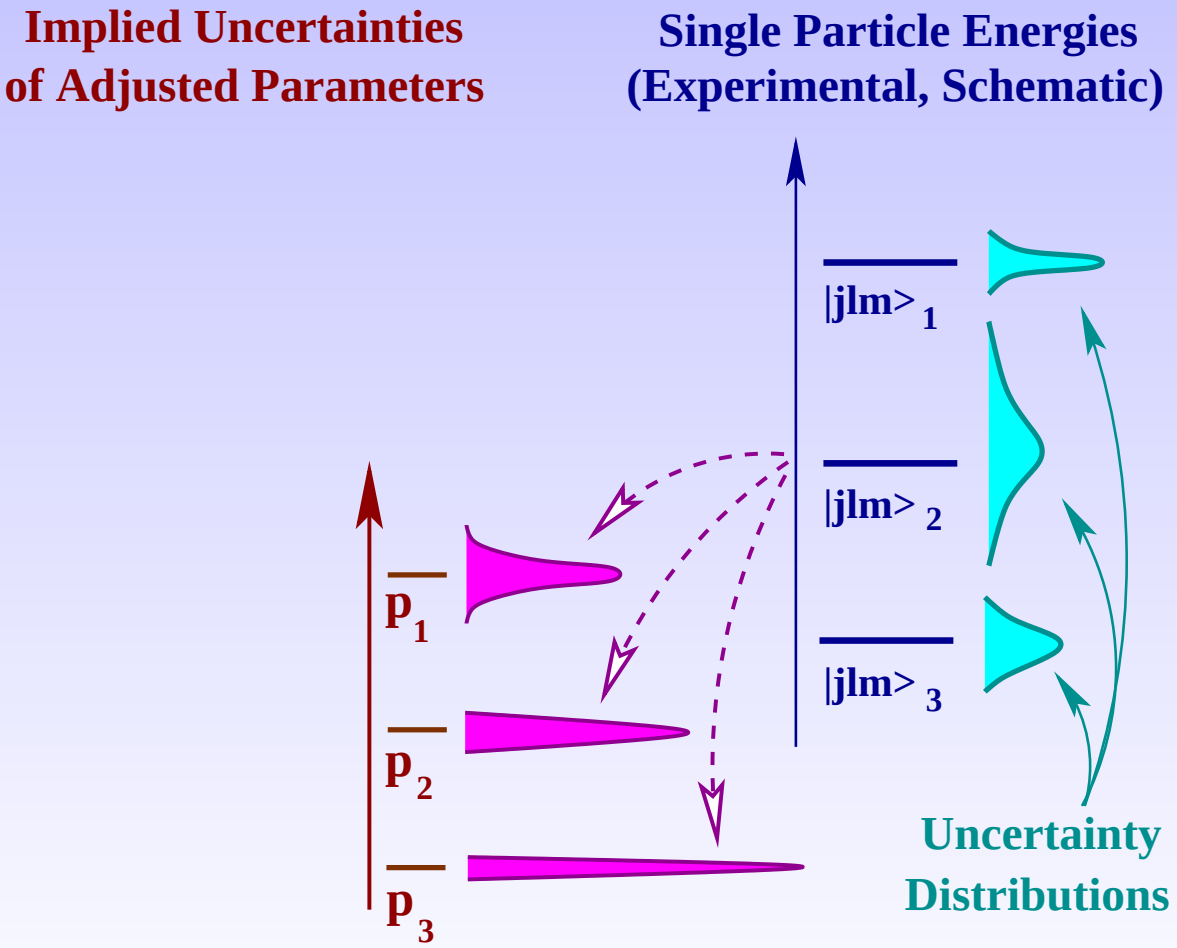
Single-Particle Levels as Probability Distributions

As it turns out, even small uncertainties on some experimental levels may cause very large uncertainties on the adjusted parameters ...



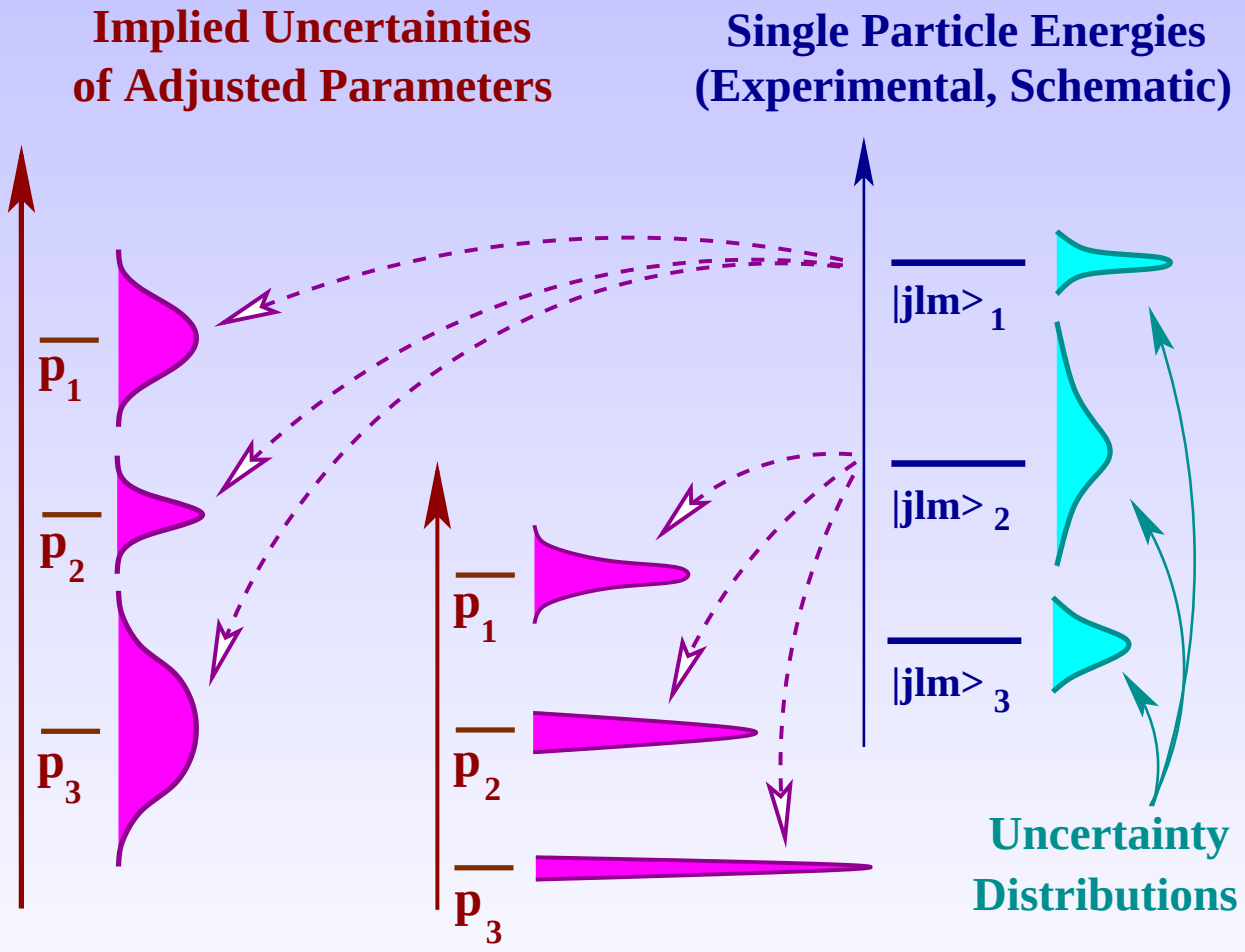
Single-Particle Levels as Probability Distributions

... while at the same time big uncertainties on other levels have a small impact: as in life, all are equal but some more than the others



Single-Particle Levels as Probability Distributions

Conclusion: the quality of adjustement depends strongly on the quality of the data implying the existence of theoretical error bars



There will be no way to present
the full richeness of the problem today

There will be no way to present the full richeness of the problem today

Therefore: Let us illustrate one of the bottom-lines first

There will be no way to present the full richeness of the problem today

Therefore: Let us illustrate one of the bottom-lines first

... and continue with presenting as much as Time

There will be no way to present
the full richeness of the problem today

Therefore: Let us illustrate one of the bottom-lines first

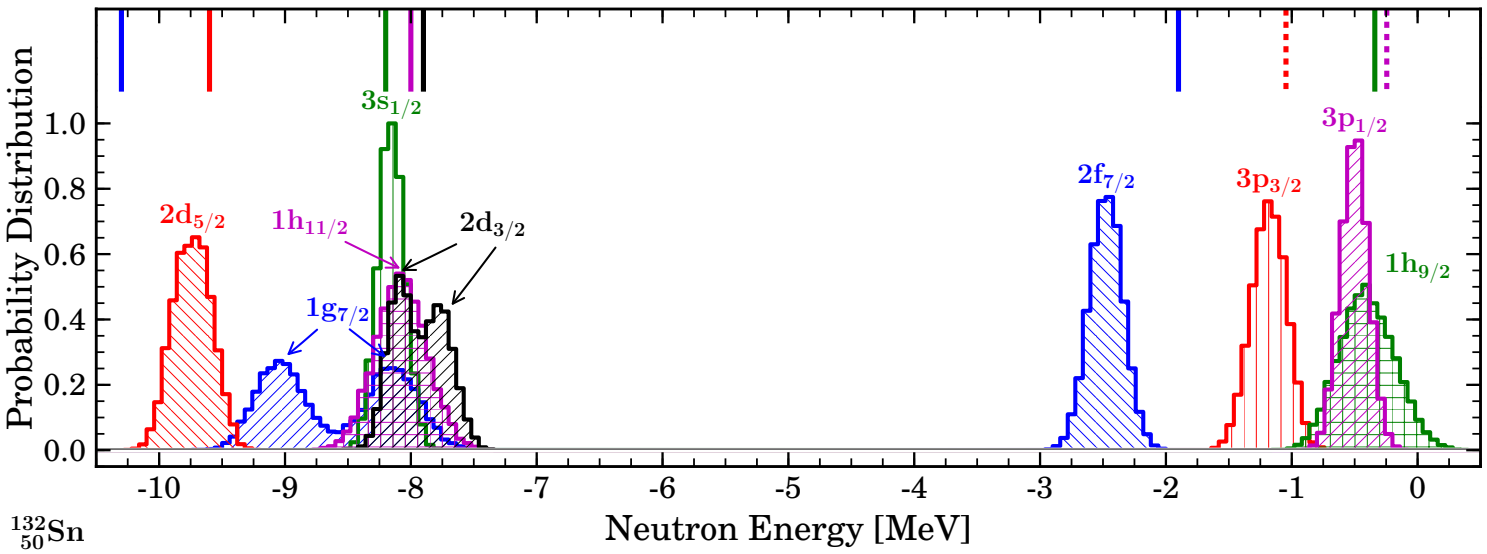
... and continue with presenting as much as Time
(and the Chairman)
permit

Inverse Problem and Predictive Power: ^{132}Sn

Neutron levels around N=82					
Level	States in ^{133}Sn				
	e_1^{exc}	shift 1$^{(a)}$	shift 2$^{(b)}$	$\bar{\varepsilon}$	B. E.
$\nu f_{7/2}$	0.0000	0.2	0.6(4)	0.6(4)	-1.9(4)
$\nu p_{3/2}$	(0.8537)	-	-	-	-
$\nu h_{9/2}$	1.5609	0.1	0.6(5)	2.2(5)	-0.3(5)
$\nu p_{1/2}$	(1.6557)	-	-	-	-
Level	States in ^{131}Sn				
	e_1^{exc}	shift 1$^{(a)}$	shift 2$^{(b)}$	$\bar{\varepsilon}$	B. E.
$\nu d_{3/2}$	0.0000	0.25	0.6(4)	0.6(4)	-7.9(4)
$\nu h_{11/2}$	0.0651	0.3	0.6(3)	0.7(3)	-8.0(3)
$\nu s_{1/2}$	0.3317	0.25	0.6(4)	0.9(4)	-8.2(4)
$\nu d_{5/2}$	1.6545	-	0.6(4)	2.3(4)	-9.6(4)
$\nu g_{7/2}$	2.4341	-	0.6(4)	3.0(4)	-10.3(4)

$^{(a)}$ *Shifts in energy from level fragmentation measured in neighbouring nuclei.*
$^{(b)}$ *The values obtained through analogy by extrapolating from the data on ^{208}Pb.*
The numbers in parentheses give errors in the last digit.

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. Obseve 'pathologies': $1g_{7/2}$ and $1f_{7/2}$ cf. following figures.

Applied Mathematics: About Inference & Inverse Problems

- **Starting from a <u>limited</u> experimental data set $\{e_\mu^{exp}\}$ we wish to obtain the information about <u>all of the single particle levels</u>**

Applied Mathematics: Inference & Inverse Problems

- **Starting from a <u>limited</u> experimental data set $\{e_\mu^{exp}\}$ we wish to obtain the information about <u>all of the single particle levels</u>**

- **In Applied Mathematics this is called the 'Inference Problem'**

Applied Mathematics: Inference & Inverse Problems

- **Starting from a limited experimental data set $\{e_\mu^{exp}\}$ we wish to obtain the information about all of the single particle levels**

- **In Applied Mathematics this is called the 'Inference Problem'**

- **The goal of underlying mathematical theories is to provide statistically sound, meaningful (i.e. stable) predictions: Thus**

THE PREDICTIVE POWER

Applied Mathematics: Inference & Inverse Problems

• Starting from a limited experimental data set $\{e_\mu^{exp}\}$ we wish to obtain the information about all of the single particle levels

• In Applied Mathematics this is called the 'Inference Problem'

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

THE PREDICTIVE POWER

• Statistically sound $\Leftrightarrow$ Instead of saying: $e_{g_{9/2}} = -8.8$ MeV we better provide also the probability function $P = P(e_{g_{9/2}})$

Simple Mathematics of the Nonlinear Inverse Problem

Simple Mathematics of the Nonlinear Inverse Problem

- Consider an example of a spherical mean-field Hamiltonian:

$$\boxed{\mathbf{H_{mf}} = \mathbf{H_{mf}}(\hat{\mathbf{r}}, \hat{\mathbf{p}}, \hat{\mathbf{s}}; \{\mathbf{p}\}); \quad \{\mathbf{p}\} \longrightarrow \text{parameters}}$$

Simple Mathematics of the Nonlinear Inverse Problem

- Consider an example of a spherical mean-field Hamiltonian:

$$\boxed{\mathbf{H_{mf}} = \mathbf{H_{mf}}(\hat{\mathbf{r}}, \hat{\mathbf{p}}, \hat{\mathbf{s}}; \{\mathbf{p}\}); \quad \{\mathbf{p}\} \longrightarrow \text{parameters}}$$

- In looking for 'an as adequate approach as possible in a search for coupling constants' we need some guidelines. Our choice:

The mean-field Hamiltonian should first of all describe optimally the mean field degrees of freedom

$$\boxed{\{\varphi_\nu, \mathbf{e}_\nu\} \ \text{ in } \ \mathbf{H_{mf}}\varphi_\nu(\mathbf{r}, \{\mathbf{p}\}) = \mathbf{e}_\nu(\mathbf{p})\,\varphi_\nu(\mathbf{r}, \{\mathbf{p}\})}$$

Simple Mathematics of the Nonlinear Inverse Problem

• Consider an example of a spherical mean-field Hamiltonian:

$$\boxed{\mathbf{H_{mf}} = \mathbf{H_{mf}}(\hat{\mathbf{r}}, \hat{\mathbf{p}}, \hat{\mathbf{s}}; \{\mathbf{p}\}); \quad \{\mathbf{p}\} \longrightarrow \text{parameters}}$$

• In looking for 'an as adequate approach as possible in a search for coupling constants' we need some guidelines. Our choice:

The mean-field Hamiltonian should first of all describe optimally the mean field degrees of freedom

$$\boxed{\{\varphi_\nu, \mathbf{e}_\nu\} \ \text{ in } \ \mathbf{H_{mf}}\varphi_\nu(\mathbf{r}, \{\mathbf{p}\}) = \mathbf{e}_\nu(\mathbf{p})\, \varphi_\nu(\mathbf{r}, \{\mathbf{p}\})}$$

• In an Appendix we explain at length why not the nuclear masses...

Introduction: χ^2-Problem and Its Linearisation

Introduction: χ^2-Problem and Its Linearisation

- Parameter adjustement usually corresponds to a minimisation of:

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

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

Introduction: χ^2-Problem and Its Linearisation

- Parameter adjustement usually corresponds to a minimisation of:

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

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

- Introduce linearisation procedure [after simplification $e_j^{th}(p) \rightarrow f_j$]

$$f_j(p) \approx f_j(p_0) + \sum_{i=1}^{n}\left(\frac{\partial f_j}{\partial p_i}\right)\Big|_{p=p_0}(p_i - p_{0,i})$$

$$J_{jk} \equiv \sqrt{W_j}\left(\frac{\partial f_j}{\partial p_k}\right)\Big|_{p=p_0} \quad \text{and} \quad b_j = \sqrt{W_j}\,[e_j^{exp} - f_j(p_0)]$$

$$\boxed{\chi^2(\mathbf{p}) = \frac{1}{\mathbf{m}-\mathbf{n}}\sum_{\mathbf{j}=1}^{\mathbf{m}}\left[\sum_{\mathbf{i}=1}^{\mathbf{n}} \mathbf{J_{ji}}\cdot(\mathbf{p_i} - \mathbf{p_{0,i}}) - \mathbf{b_j}\right]^2}$$

χ^2-Problem Using Applied-Mathematics' Jargon

χ^2-Problem Using Applied-Mathematics' Jargon

- **One may easily show that within the linearized representation**

$$\boxed{\frac{\partial \chi^2}{\partial \mathbf{p_i}} = 0 \quad \longrightarrow \quad (\mathbf{J^T J}) \cdot (\mathbf{p} - \mathbf{p}_0) = \mathbf{J^T\, b}}$$

χ^2-Problem Using Applied-Mathematics' Jargon

- **One may easily show that within the linearized representation**

$$\boxed{\frac{\partial \chi^2}{\partial \mathbf{p_i}} = 0 \quad \longrightarrow \quad (\mathbf{J^T J}) \cdot (\mathbf{p} - \mathbf{p_0}) = \mathbf{J^T\, b}}$$

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

$$\boxed{\{\mathbf{p}\} \leftrightarrow \mathbf{x} \leftrightarrow \text{'causes'} \text{ and } \{\mathbf{e}\} \leftrightarrow \mathbf{b} \leftrightarrow \text{'effects'} \leftrightarrow \mathbf{A} \cdot \mathbf{x} = \mathbf{b}}$$

From the measured effects represented by 'b' we extract information about the causes 'x' by inverting the matrix: $\mathbf{A} \longrightarrow \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$

χ^2-Problem Using Applied-Mathematics' Jargon

- **One may easily show that within the linearized representation**

$$\boxed{\frac{\partial \chi^2}{\partial \mathbf{p_i}} = 0 \quad \longrightarrow \quad (\mathbf{J^T J}) \cdot (\mathbf{p} - \mathbf{p_0}) = \mathbf{J^T\, b}}$$

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

$$\boxed{\{\mathbf{p}\} \leftrightarrow \mathbf{x} \leftrightarrow \text{'causes'} \text{ and } \{\mathbf{e}\} \leftrightarrow \mathbf{b} \leftrightarrow \text{'effects'} \leftrightarrow \mathbf{A} \cdot \mathbf{x} = \mathbf{b}}$$

From the measured effects represented by 'b' we extract information about the causes 'x' by inverting the matrix: $\mathbf{A} \longrightarrow \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$

- Some mathematical details: $J \equiv A \in \mathbb{R}^{m \times n}$; $x \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$

A Powerful Tool: Singular-Value Decomposition

A Powerful Tool: Singular-Value Decomposition

• The problems with instabilities (i.e. ill-conditioning) can be easily illustrated using the so-called Singular-Value Decomposition of A:

$$\boxed{\mathbf{A} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^{\mathbf{T}} \text{ with } \mathbf{U} \in \mathbb{R}^{\mathbf{m}\times\mathbf{m}}, \quad \mathbf{V} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}, \quad \mathbf{D} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}}$$

where diagonal matrix has a form $D = \mathrm{diag}\{\underbrace{d_1, d_2, \ \ldots \ d_{min(m,n)}}_{\text{decreasing order}}\}$

A Powerful Tool: Singular-Value Decomposition

• The problems with instabilities (i.e. ill-conditioning) can be easily illustrated using the so-called Singular-Value Decomposition of A:

$$\boxed{\mathbf{A} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^{\mathbf{T}} \text{ with } \mathbf{U} \in \mathbb{R}^{\mathbf{m}\times\mathbf{m}}, \quad \mathbf{V} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}, \quad \mathbf{D} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}}$$

where diagonal matrix has a form $D = \mathrm{diag}\{\underbrace{d_1, d_2, \ \ldots \ d_{min(m,n)}}_{\text{decreasing order}}\}$

• Formally (but also in practice), the solution 'x' is expressed as

$$\boxed{\mathbf{x} = \mathbf{A}^{\mathbf{T}}\mathbf{b}; \quad \mathbf{A}^{\mathbf{T}} = \mathbf{V} \cdot \mathbf{D}^{\mathbf{T}} \cdot \mathbf{U}^{\mathbf{T}}}$$

where

$$\mathbf{D}^{\mathbf{T}} = \mathbf{diag}\left\{\frac{1}{d_1}, \frac{1}{d_2}, \ \ldots \ \frac{1}{d_p}; 0, 0, \ \ldots \ 0\right\}$$

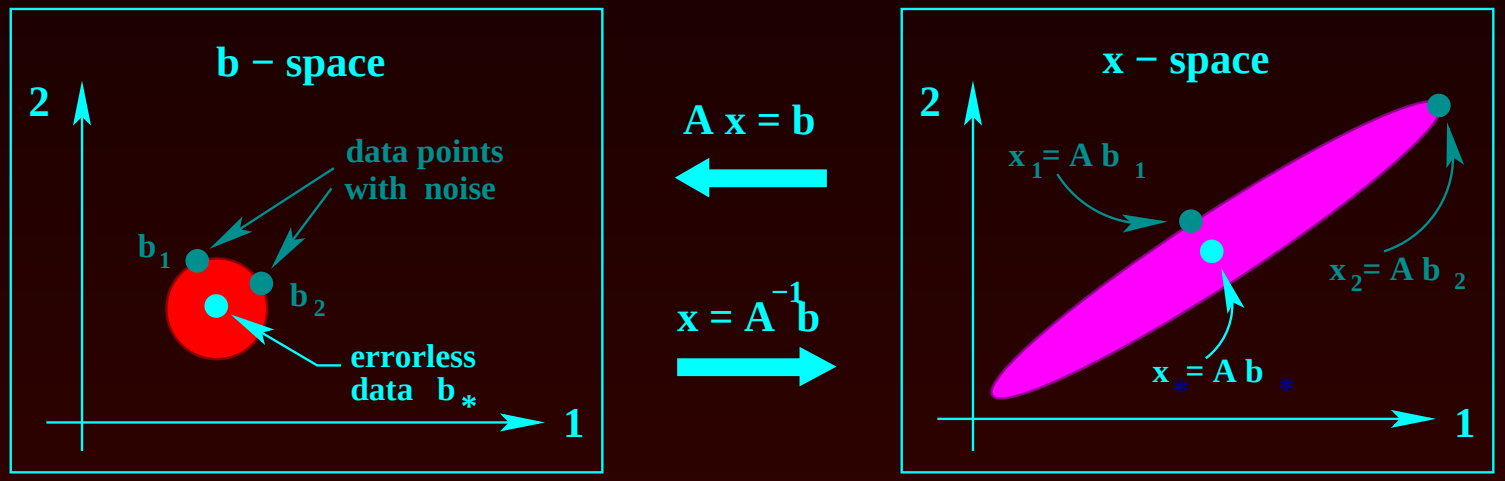
Ill-Conditioned Problems: Qualitative Illustration

Ill-Conditioned Problems: Qualitative Illustration

- An academic 2×2 problem: A has eigenvalues: d_1 and $d_2 \sim \frac{1}{10} d_1$

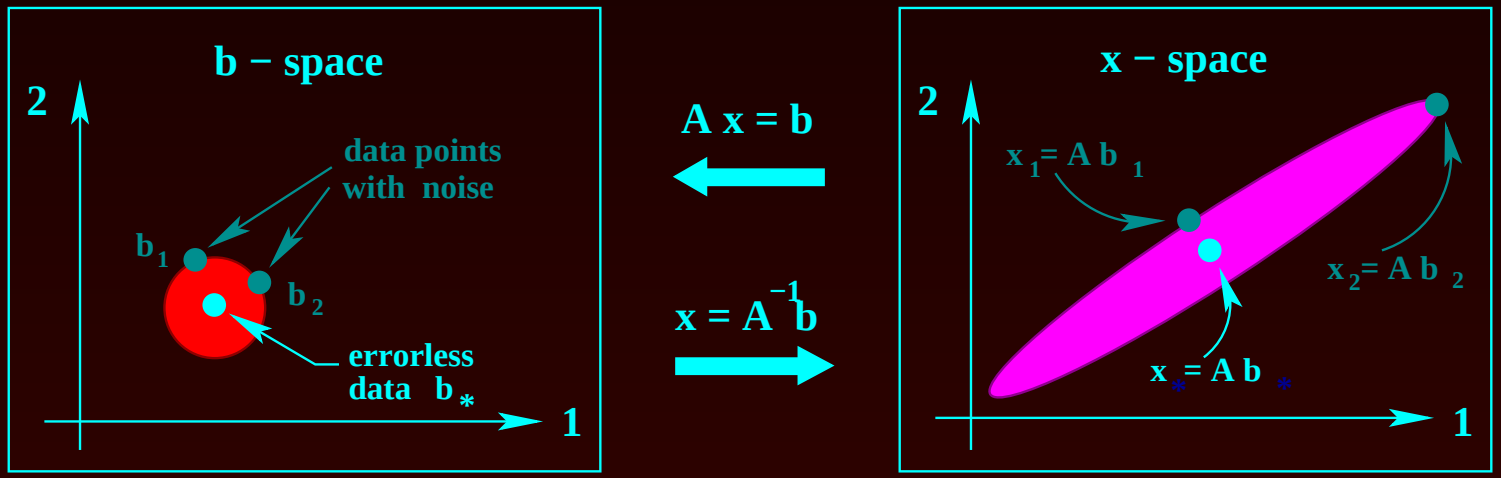
Ill-Conditioned Problems: Qualitative Illustration

- An academic 2×2 problem: A has eigenvalues: d_1 and $d_2 \sim \frac{1}{10} d_1$
- Introduce errorless data (b_*) and uncertain (noisy) data: b_1 & b_2



Ill-Conditioned Problems: Qualitative Illustration

- An academic 2×2 problem: A has eigenvalues: d_1 and $d_2 \sim \frac{1}{10} d_1$
- Introduce errorless data (b_*) and uncertain (noisy) data: b_1 & b_2



Left: Red circle represents points equally distant from 'noisless' data b_*. Right: Purple oval represents the image of the circle through $Ax=b$. One may show that instability is directly dependent on the condition number

$$\boxed{cond(A) \equiv \frac{d_1}{d_r}}$$

- the bigger the condition number the more 'ill-conditioned' the problem

Fitting, Implied Errors and Confidence Intervals

Fitting, Implied Errors and Confidence Intervals

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

$$\frac{\partial \chi^2}{\partial \mathbf{p_j}} \quad \longrightarrow \quad (\mathbf{J^T J})(\mathbf{p} - \mathbf{p_0}) = \mathbf{J^T b}$$

Fitting, Implied Errors and Confidence Intervals

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

$$\frac{\partial \chi^2}{\partial p_j} \quad \rightarrow \quad (J^T J)(p - p_0) = J^T b$$

- Using the Singular-Value Decomposition written down explicitly

$J_{ik} = \sum_{\ell=1}^{r} U_{i\ell}\, d_\ell\, V^T_{\ell k}$ one shows that $(J^T J)^{-1} = V \frac{1}{d^2} V^T$

Fitting, Implied Errors and Confidence Intervals

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

$$\frac{\partial \chi^2}{\partial p_j} \quad \rightarrow \quad (J^T J)(p - p_0) = J^T b$$

- Using the Singular-Value Decomposition written down explicitly

$J_{ik} = \sum_{\ell=1}^{r} U_{i\ell}\, d_\ell\, V^T_{\ell k}$ one shows that $(J^T J)^{-1} = V \frac{1}{d^{\,2}} V^T$

- Independently one derives the expression for the correlation matrix

$$\boxed{\langle (p_i - \langle p_i \rangle) \cdot (p_j - \langle p_j \rangle) \rangle = \chi^2(p)\, t^2_{\alpha/2, m-n}\, (J^T J)^{-1}_{ij}}$$

Fitting, Implied Errors and Confidence Intervals

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

$$\frac{\partial \chi^2}{\partial p_j} \quad \longrightarrow \quad (J^T J)(p - p_0) = J^T b$$

- Using the Singular-Value Decomposition written down explicitly

$J_{ik} = \sum_{\ell=1}^{r} U_{i\ell}\, d_\ell\, V^T_{\ell k}$ one shows that $(J^T J)^{-1} = V \frac{1}{d^{\,2}} V^T$

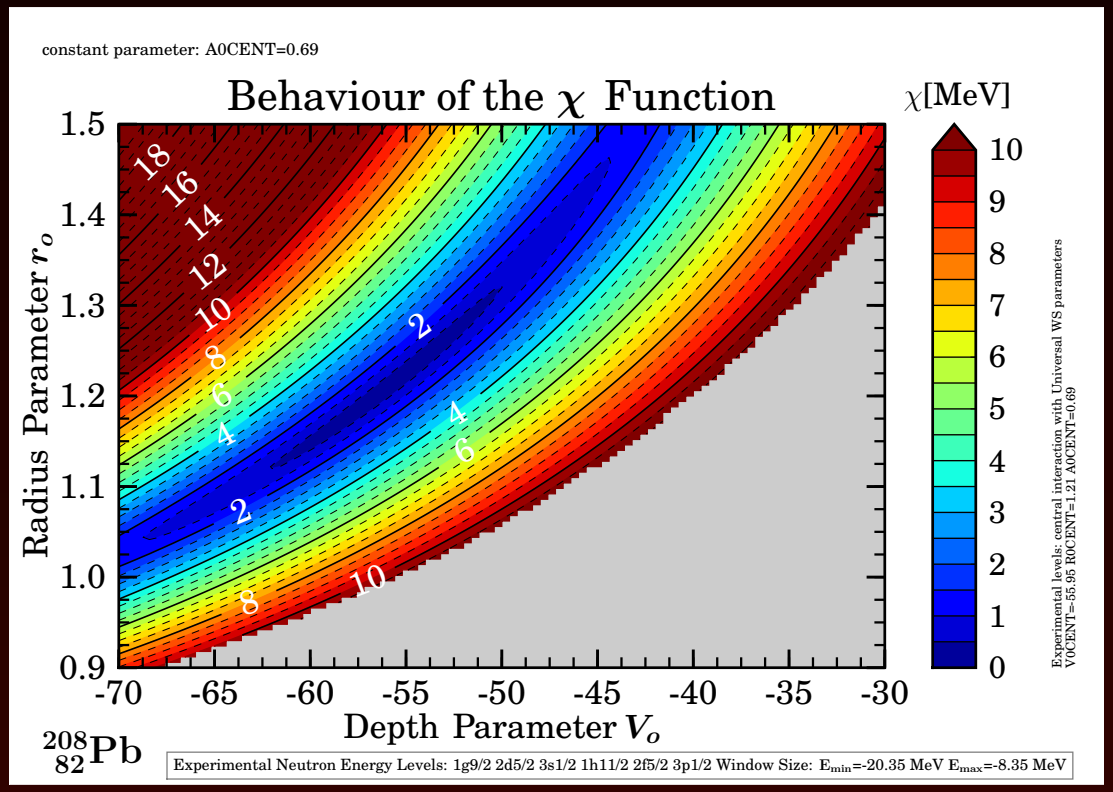
- Independently one derives the expression for the correlation matrix

$$\boxed{\langle (p_i - \langle p_i \rangle) \cdot (p_j - \langle p_j \rangle) \rangle = \chi^2(p)\, t^2_{\alpha/2, m-n}\, (J^T J)^{-1}_{ij}}$$

- **If one or more $d_k \longrightarrow 0$ then $(J^T J)^{-1}$ tends to infinity and generally, the <u>confidence intervals of all parameters diverge</u>**

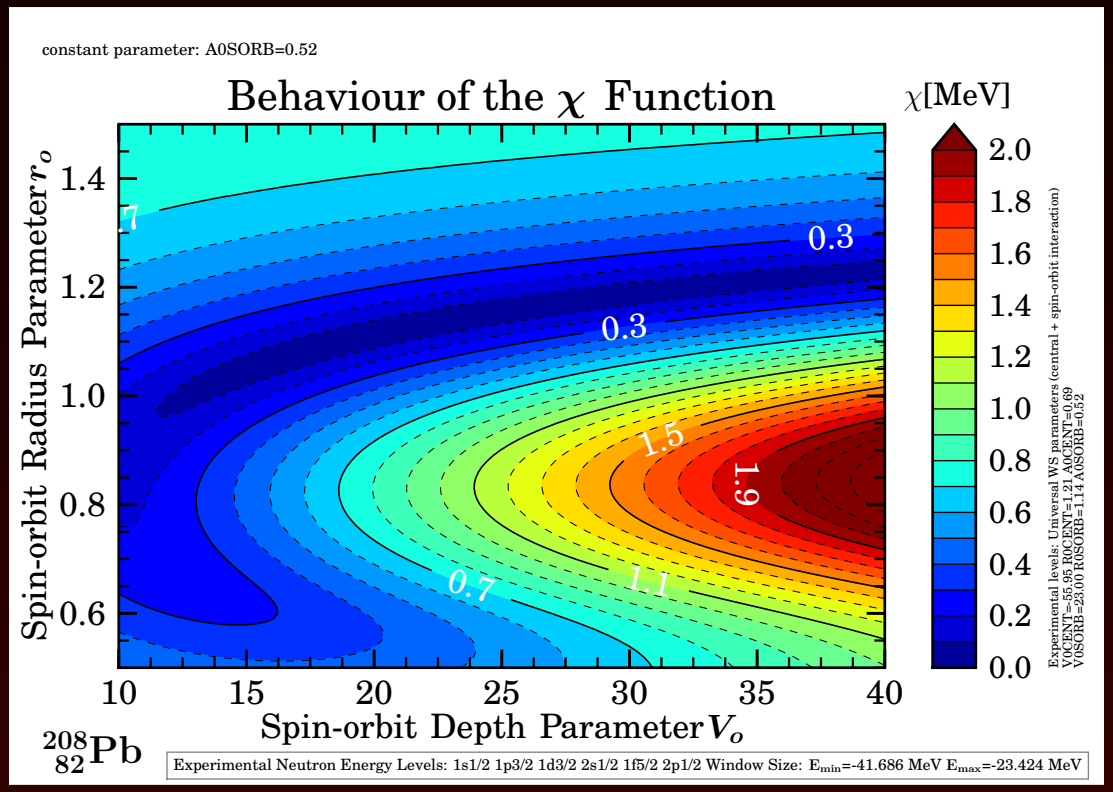
Undesired Parametric Correlations: Illustrative Examples

Begin with a Well Known: V_o vs. r_o Are Correlated



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

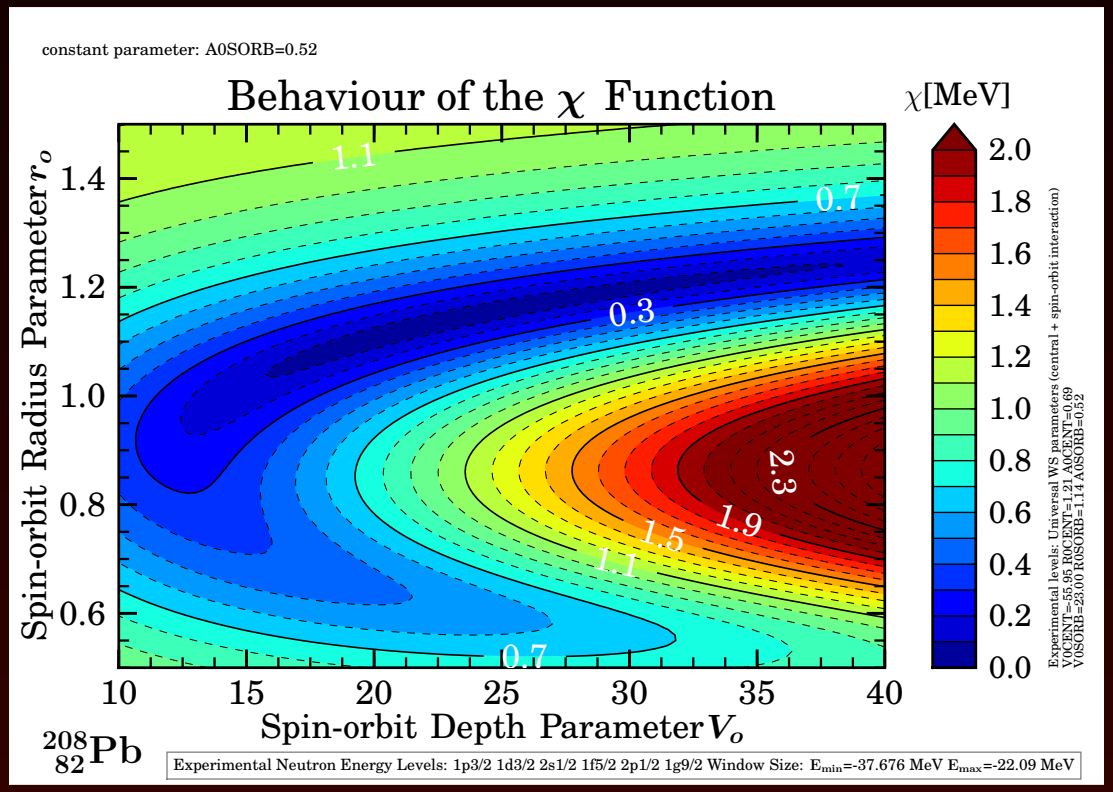
Parametric Correlations and Their Consequences



We plot the χ^2 in function of the S-O strength (horizontal) and the S-O radius (vertical) axis. We start with six very lowest levels. Note: no way to fix reliably the spin-orbit strength in the interval from 15 to 40 units

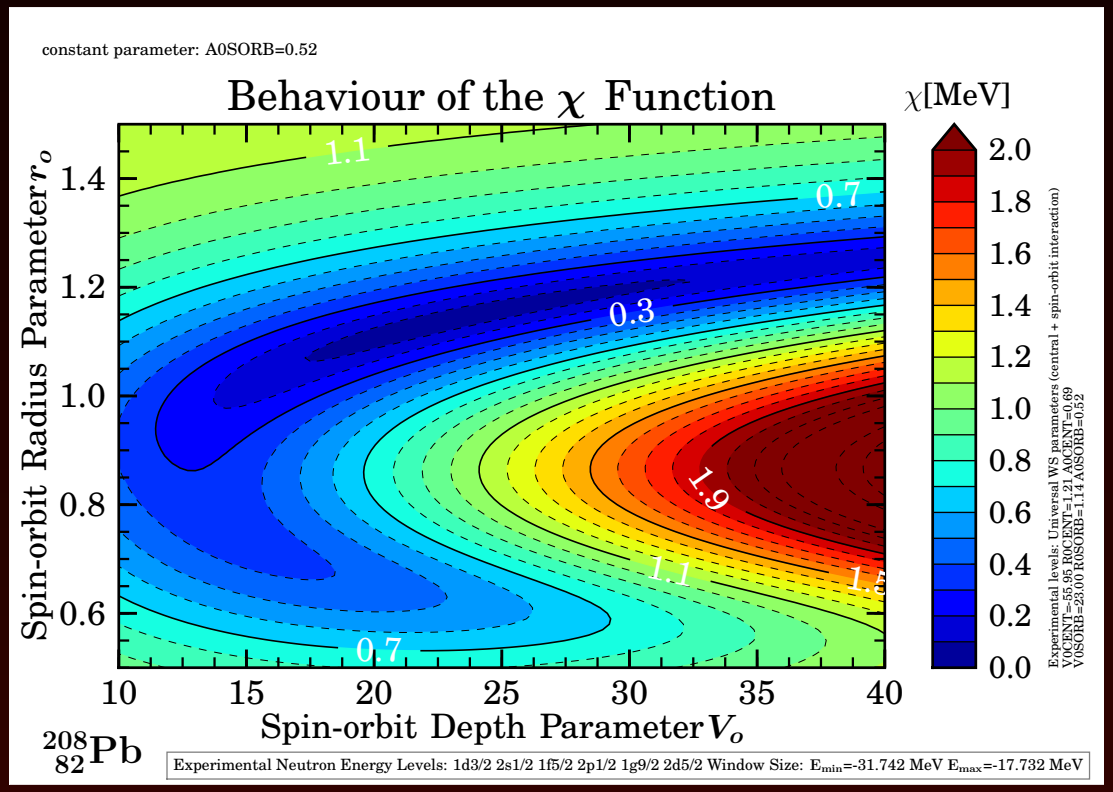
We will gradually increase the energy of the six-level window to approach the nucleon binding region and thus simulate the present-day experimental situation

Limited Experimental Input: How Little is Sufficient?



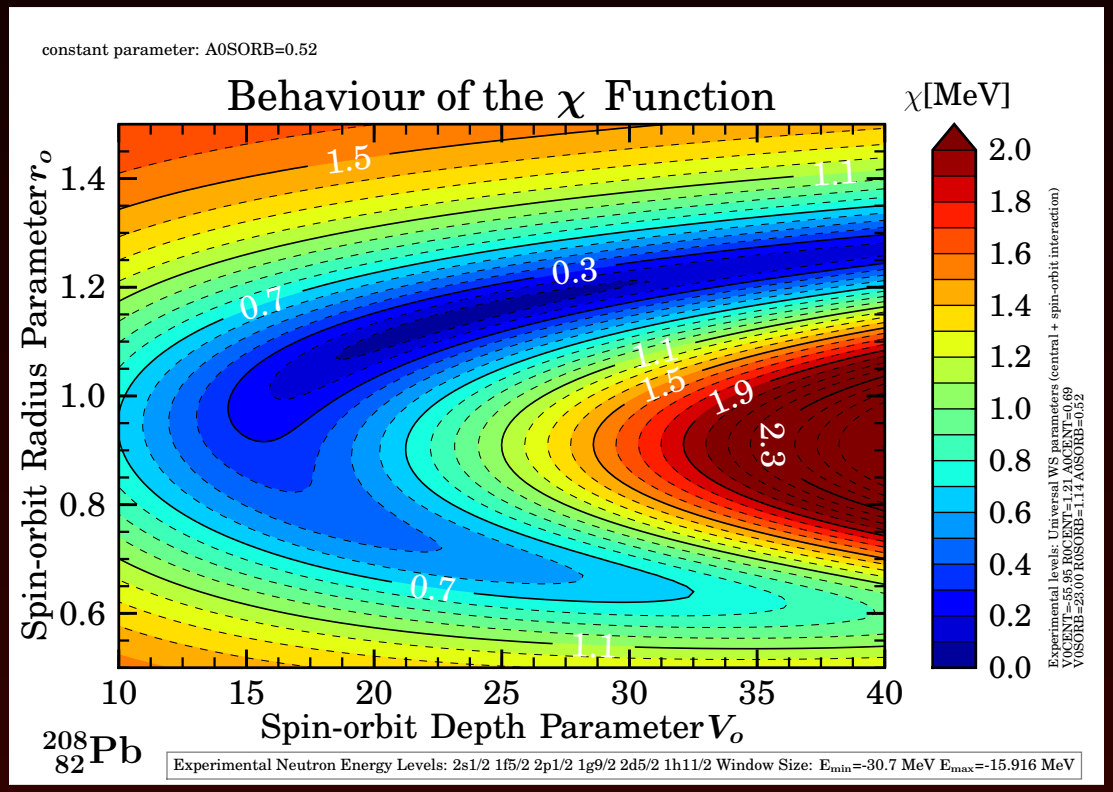
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly!

Limited Experimental Input: How Little is Sufficient?



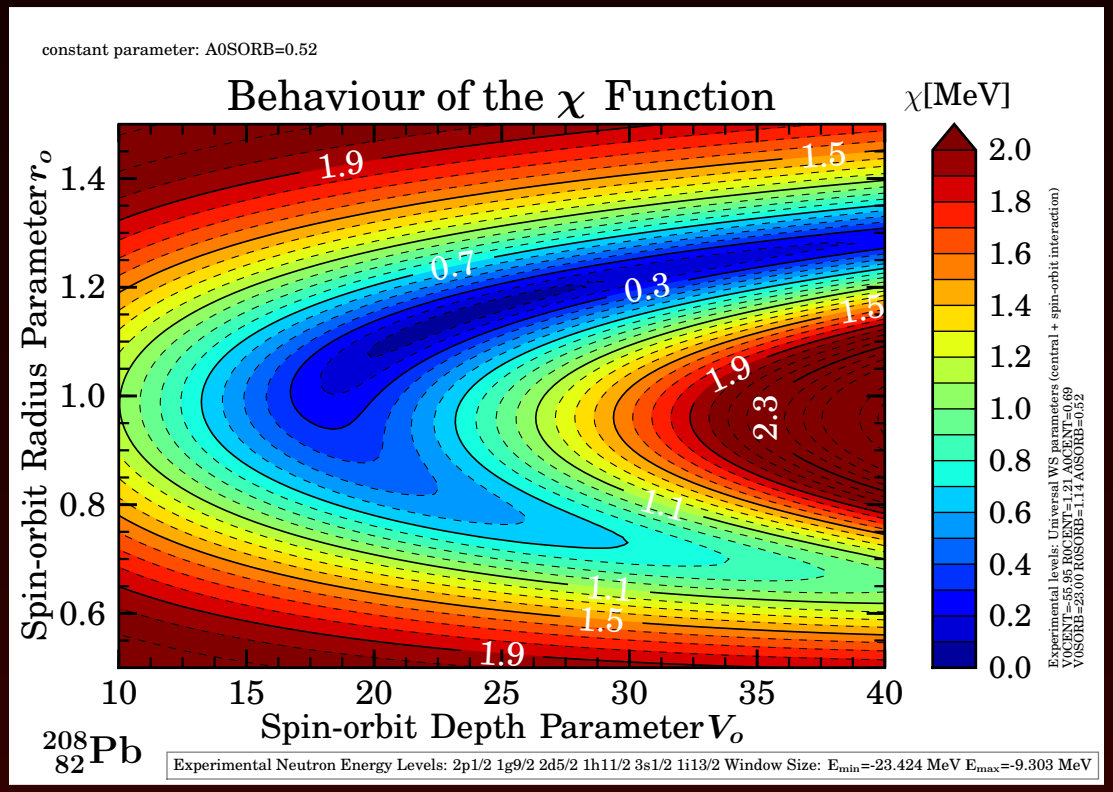
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



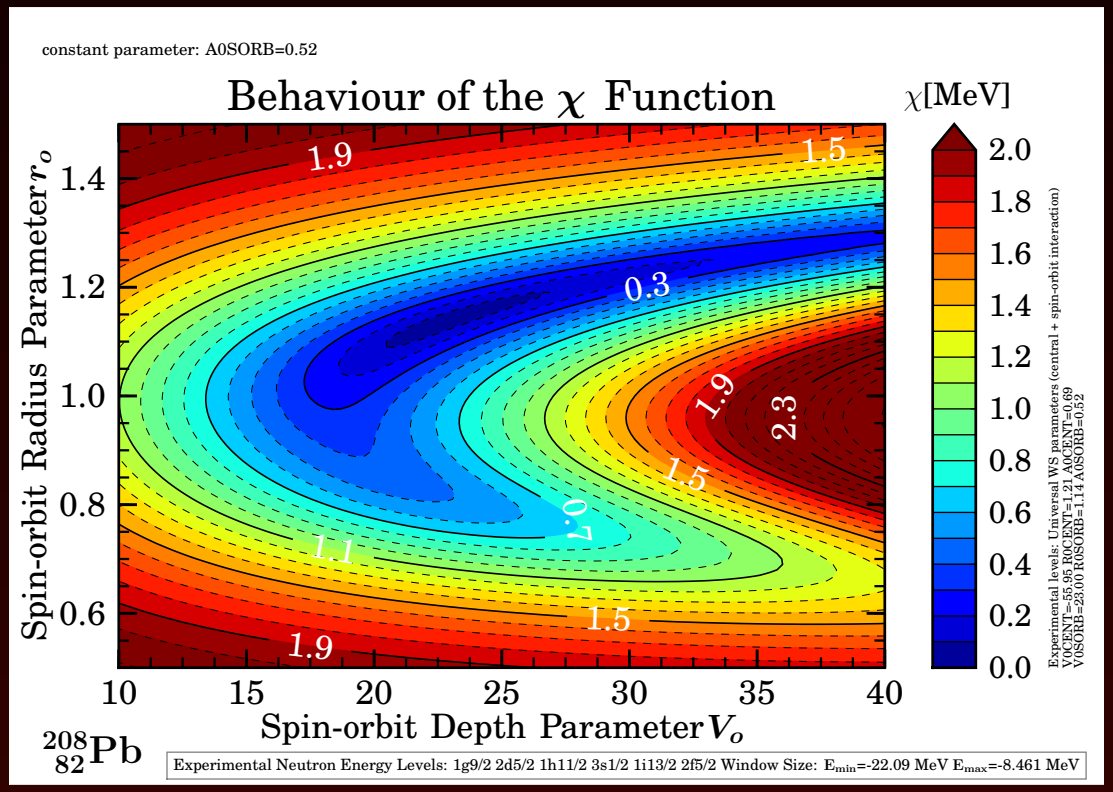
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



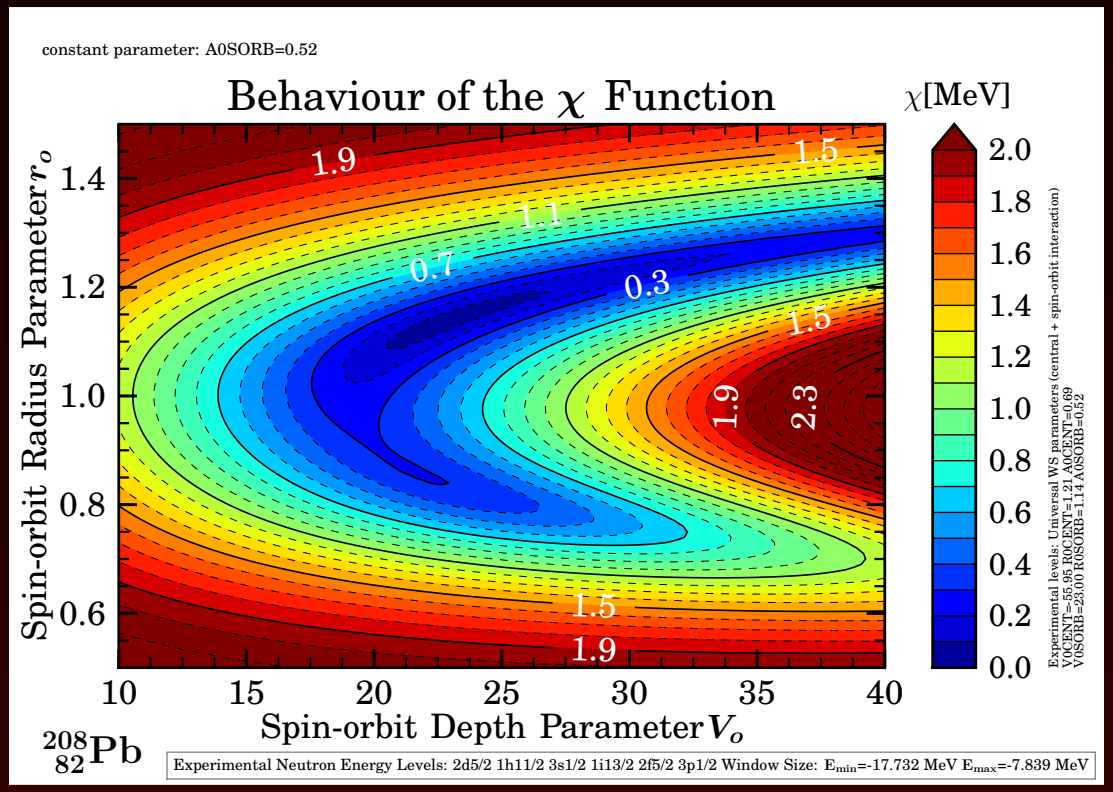
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



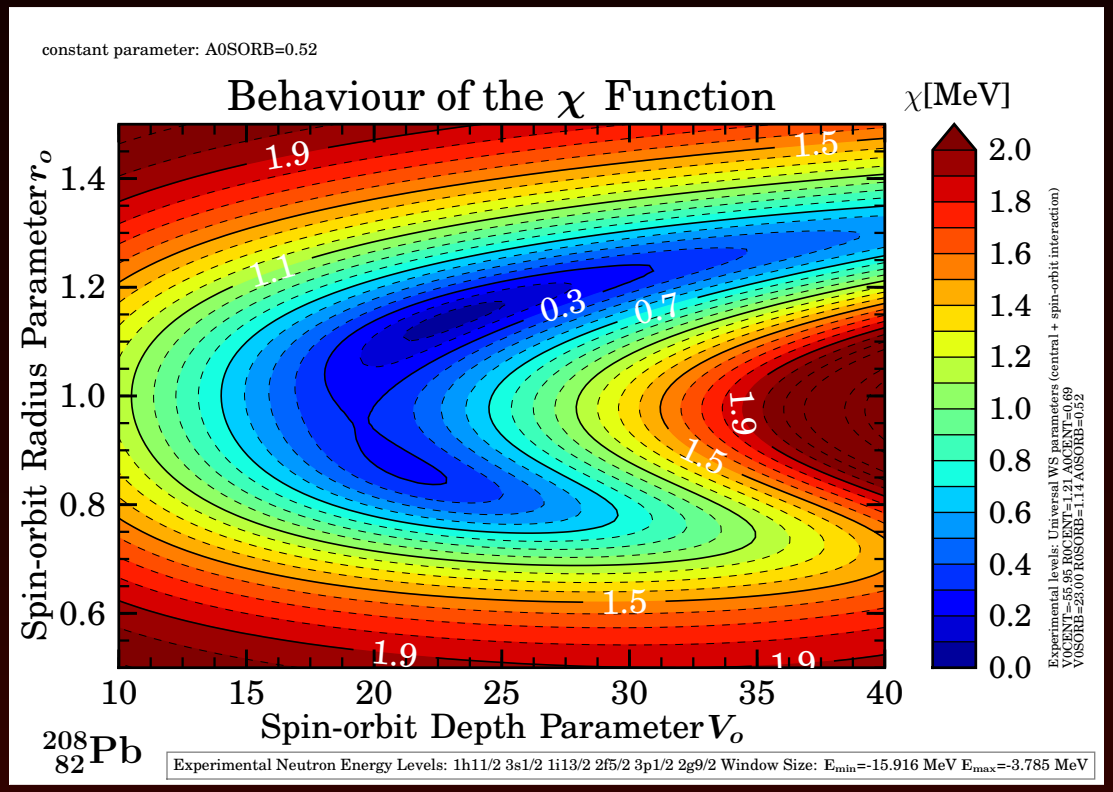
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



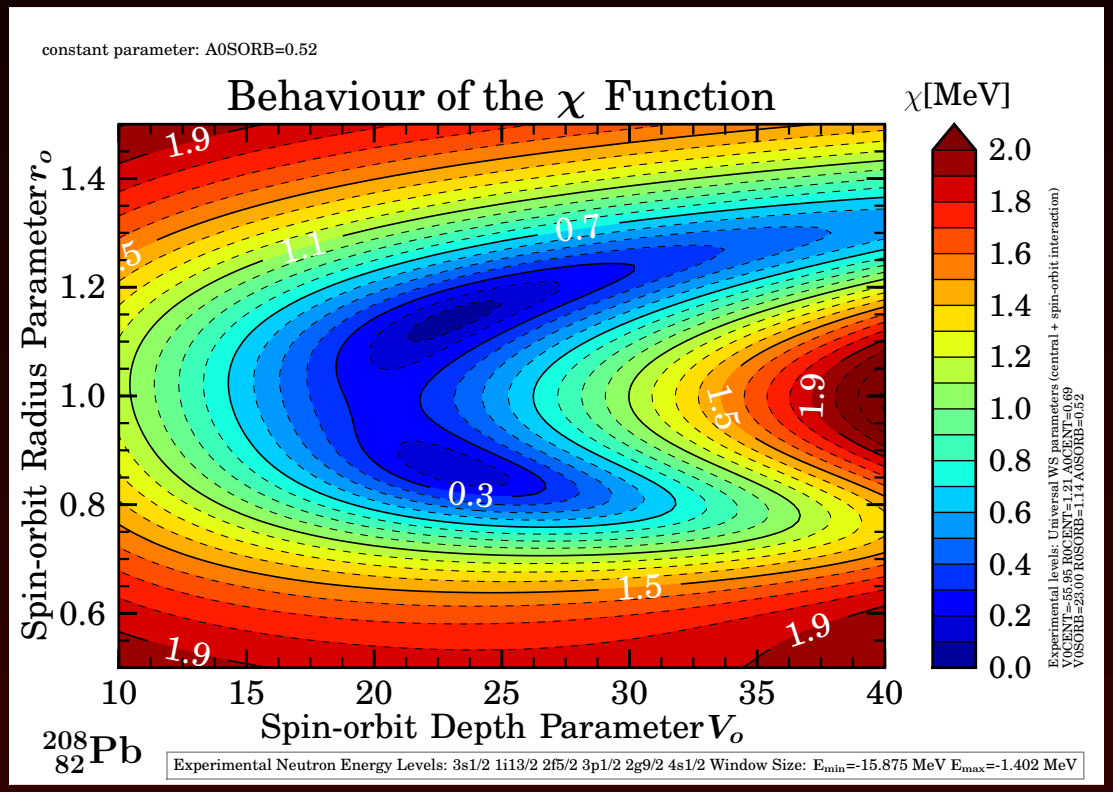
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly... whereas gradually another solution ...

Limited Experimental Input: How Little is Sufficient?



Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly... *Attention: Second solution is coming !*

Limited Experimental Input: How Little is Sufficient?



... and here we discover the existence of <u>two solutions</u> - we call them compact and non-compact.

What Can We Conclude from the This Set of Tests?

- **First of all, the <u>fitted</u> spin-orbit strength may <u>vary widely</u> from one doubly-magic nucleus to another - there exists a considerable softness in χ^2 dependence on λ_{so}**

What Can We Conclude from the This Set of Tests?

- **First of all, the fitted spin-orbit strength may vary widely from one doubly-magic nucleus to another - there exists a considerable softness in χ^2 dependence on λ_{so}**

- **Secondly - there are correlations between the parameters: two relatively distant sets of parameters may be, what we call spectroscopically equivalent ('iso-spectroscopic')**

What Can We Conclude from the This Set of Tests?

- **First of all, the fitted spin-orbit strength may vary widely from one doubly-magic nucleus to another - there exists a considerable softness in χ^2 dependence on λ_{so}**

- **Secondly - there are correlations between the parameters: two relatively distant sets of parameters may be, what we call spectroscopically equivalent ('iso-spectroscopic')**

- **Clearly: The 'spectroscopic predictive power' is in danger !!! - especially since the results depends on the sampling thus on experimentalist's $\rightarrow$ with all the implied subjectivity**

What Can We Conclude from the This Set of Tests?

- **First of all, the fitted spin-orbit strength may vary widely from one doubly-magic nucleus to another - there exists a considerable softness in χ^2 dependence on λ_{so}**

- **Secondly - there are correlations between the parameters: two relatively distant sets of parameters may be, what we call spectroscopically equivalent ('iso-spectroscopic')**

- **Clearly: The 'spectroscopic predictive power' is in danger !!! - especially since the results depends on the sampling thus on experimentalist's $\rightarrow$ with all the implied subjectivity**

- **We discover a possibility of double-valued solutions giving rise to compact and non-compact spin-orbit parametrisations**

What Can We Conclude from the This Set of Tests?

- **First of all, the fitted spin-orbit strength may vary widely from one doubly-magic nucleus to another - there exists a considerable softness in χ^2 dependence on λ_{so}**

- **Secondly - there are correlations between the parameters: two relatively distant sets of parameters may be, what we call spectroscopically equivalent ('iso-spectroscopic')**

- **Clearly: The 'spectroscopic predictive power' is in danger !!! - especially since the results depends on the sampling thus on experimentalist's $\rightarrow$ with all the implied subjectivity**

- **We discover a possibility of double-valued solutions giving rise to compact and non-compact spin-orbit parametrisations**

- **We confirm the presence of iso-spectral lines also in the space of the spin-orbit potential parameters**

A Kind of Summary and a Historical Analogy

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers,

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers, but on the other hand we are confident of our errors

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers, but on the other hand we are confident of our errors

A similar problem has been encountered,
according to Umberto Eco, about 1327 ("Il nome della rosa")

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers, but on the other hand we are confident of our errors

A similar problem has been encountered,
according to Umberto Eco, about 1327 ("Il nome della rosa")

"So you don't have unique answers to your questions?"

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers, but on the other hand we are confident of our errors

A similar problem has been encountered,
according to Umberto Eco, about 1327 ("Il nome della rosa")

"So you don't have unique answers to your questions?"

"Adson, if I had, I would teach theology in Paris"

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers, but on the other hand we are confident of our errors

A similar problem has been encountered,
according to Umberto Eco, about 1327 ("Il nome della rosa")

"So you don't have unique answers to your questions?"

"Adson, if I had, I would teach theology in Paris"

"Do they always have a right answer in Paris"

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers, but on the other hand we are confident of our errors

A similar problem has been encountered,
according to Umberto Eco, about 1327 ("Il nome della rosa")

"So you don't have unique answers to your questions?"

"Adson, if I had, I would teach theology in Paris"

"Do they always have a right answer in Paris"

"Never", said William,

A Kind of Summary and a Historical Analogy

If the confidence intervals diverge we loose unique answers, but on the other hand we are confident of our errors

A similar problem has been encountered,
according to Umberto Eco, about 1327 ("Il nome della rosa")

"So you don't have unique answers to your questions?"

"Adson, if I had, I would teach theology in Paris"

"Do they always have a right answer in Paris"

"Never", said William,
"but there they are quite confident of their errors".

On the Way to the Land-of-Meaningful-Probabilities

our brave saylors will see a lot of sharks and rocks

ILLUSTRATIONS

The Following Messages

The Following Messages are intended

The Following Messages
are intended
for Mature Audiences

About Contemporary Skyrme-Type Functionals

About Contemporary Skyrme-Type Functionals

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

About Contemporary Skyrme-Type Functionals

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

- **Their total energy density reads**

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

where $C^{m'l',n'L'v'J'}_{ml,nLvJ,Q}$ are coupling constants

About Contemporary Skyrme-Type Functionals

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

- **Their total energy density reads**

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

where $C^{m'l',n'L'v'J'}_{ml,nLvJ,Q}$ are coupling constants

- **It is instructive to think about the extentions of the EDF based approaches in terms of increasing number of coupling constants and the preceding illustrations...**

About Contemporary Skyrme-Type Functionals

Numbers of terms of different orders in the EDF up to N^3LO. 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 invariance is assumed, respectively. To take into account both isospin channels, the numbers of terms should be multiplied by a factor of two.

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

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

Theory Predictions and Their Statistical Significance

Consider a mean-field Hamiltonian: RMF, HF, Phenomenological ...

$$\boxed{\mathbf{H_{mf} = H_{mf}(\hat{r}, \hat{p}, \hat{s}; \{p\})}; \quad \{\mathbf{p}\} \longrightarrow \text{parameters}}$$

After laborious constructions of H_{mf}, we often 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!*

Theory Predictions and Their Statistical Significance

Consider a mean-field Hamiltonian: RMF, HF, Phenomenological ...

$$\boxed{\mathbf{H_{mf} = H_{mf}(\hat{r}, \hat{p}, \hat{s}; \{p\})}; \quad \{\mathbf{p}\} \longrightarrow \text{parameters}}$$

After laborious constructions of H_{mf}, we often 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!*

Theory Predictions and Their Statistical Significance

Consider a mean-field Hamiltonian: RMF, HF, Phenomenological ...

$$\boxed{\mathbf{H_{mf} = H_{mf}(\hat{r}, \hat{p}, \hat{s}; \{p\})}; \quad \{p\} \longrightarrow \text{parameters}}$$

After laborious constructions of H_{mf}, we often 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 *'Modeling of Data'*, the authors of *'Numerical Recipes"* (p. 651), observe with sarcasme:

"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 ment is: "they" get a 'statistical nonsense']

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

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!

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.

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

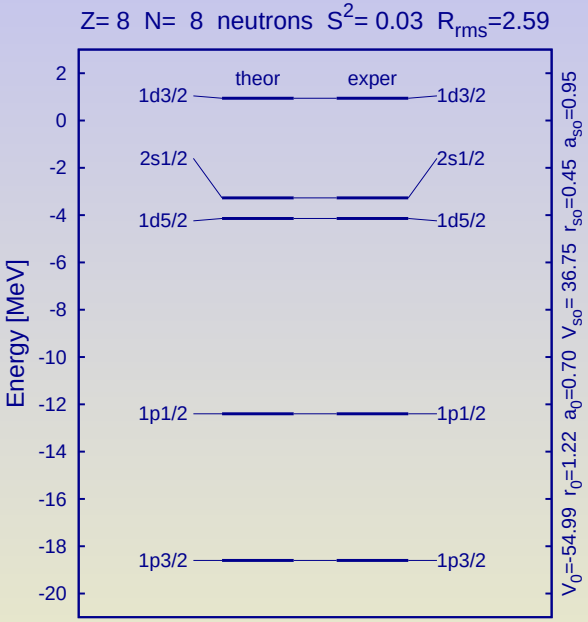
• We can treat $\{e_\mu\}$ 'as experimental'; by trying to reproduce them through fitting we know an exact solution!

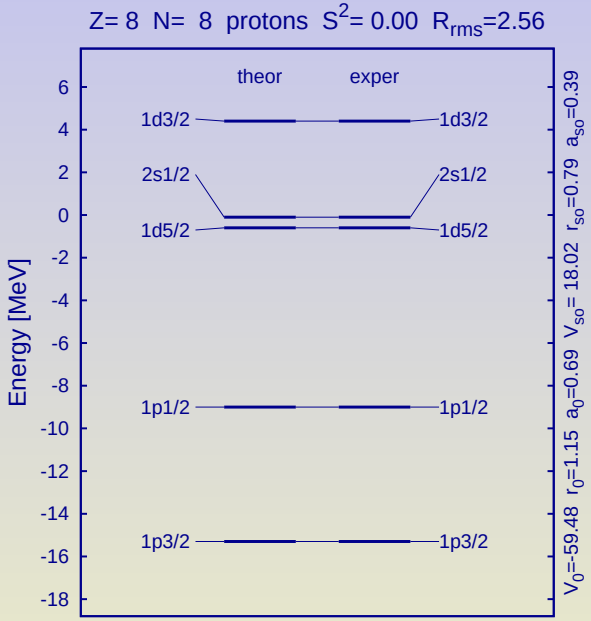
• Extra advantage: we may introduce the notion of 'noise', usually a random variable distributed according to a certain probability fct.

• 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 $\{p\}$

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

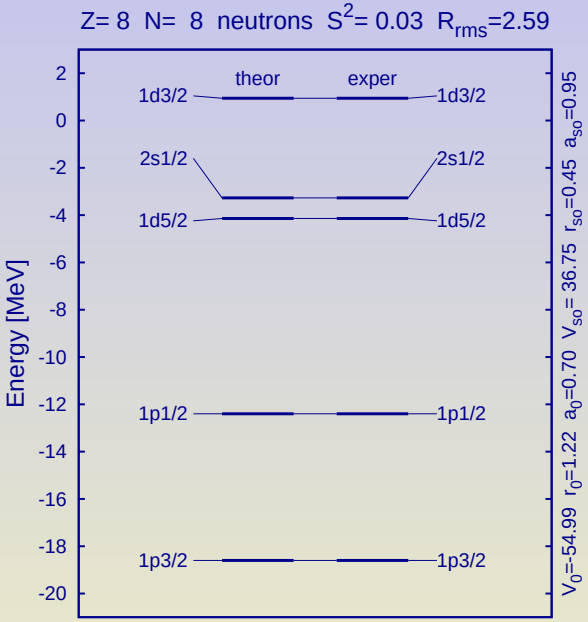
- We fit the single-particle experimental levels in ^{16}O using Woods-Saxon potential (six parameters for protons and neutrons each)

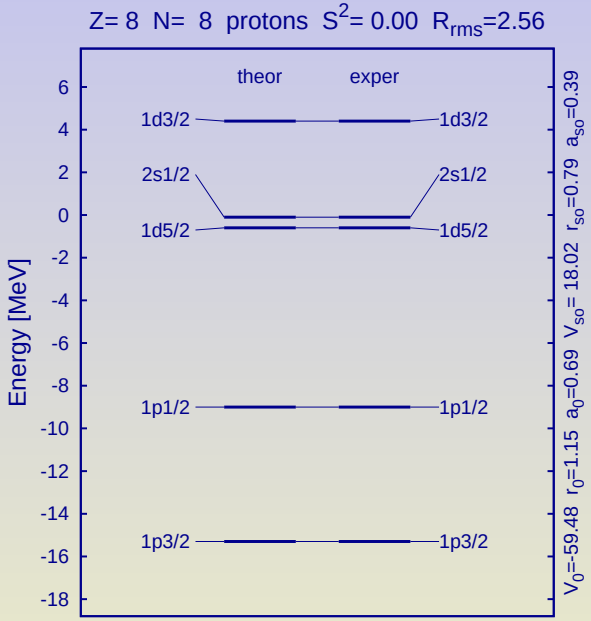




‘Chi-by-the-eye’ Results May Look Attractive...

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

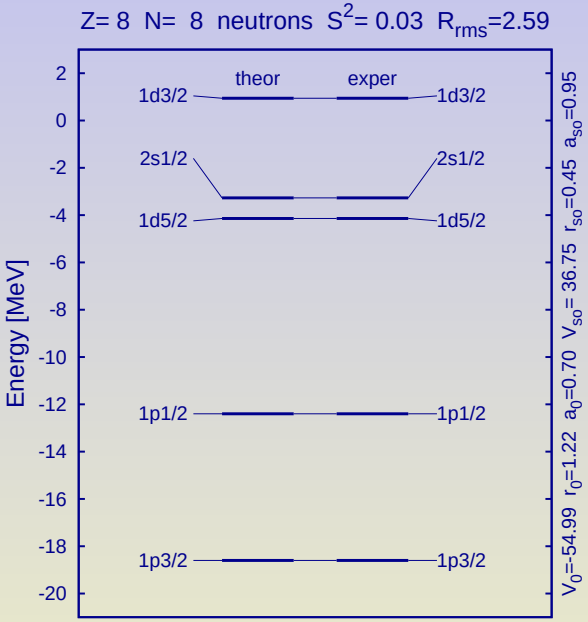


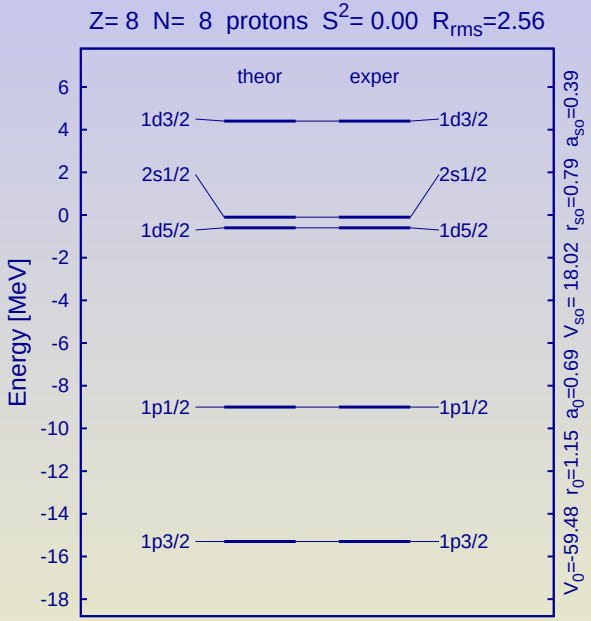


• This result may look surprising: the quality of the fit is such that *graphical illustrations are <u>insufficient to show it</u>* !!!

‘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 <u>insufficient to show it</u>* !!!

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

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!

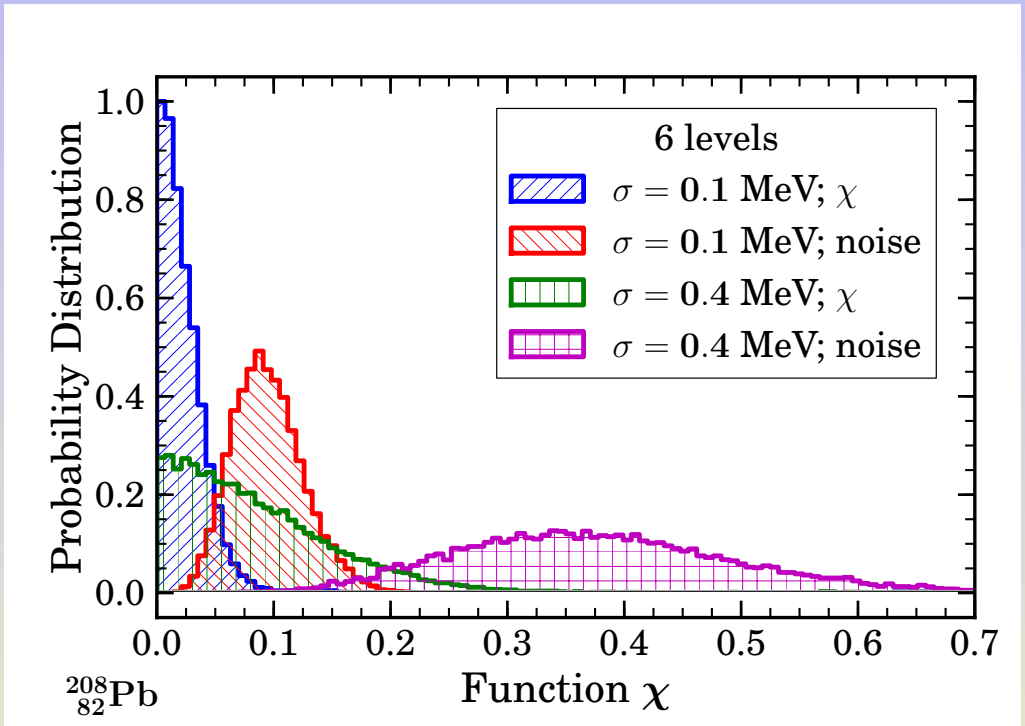
→ The standard Woods-Saxon Hamiltonian has been used:

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

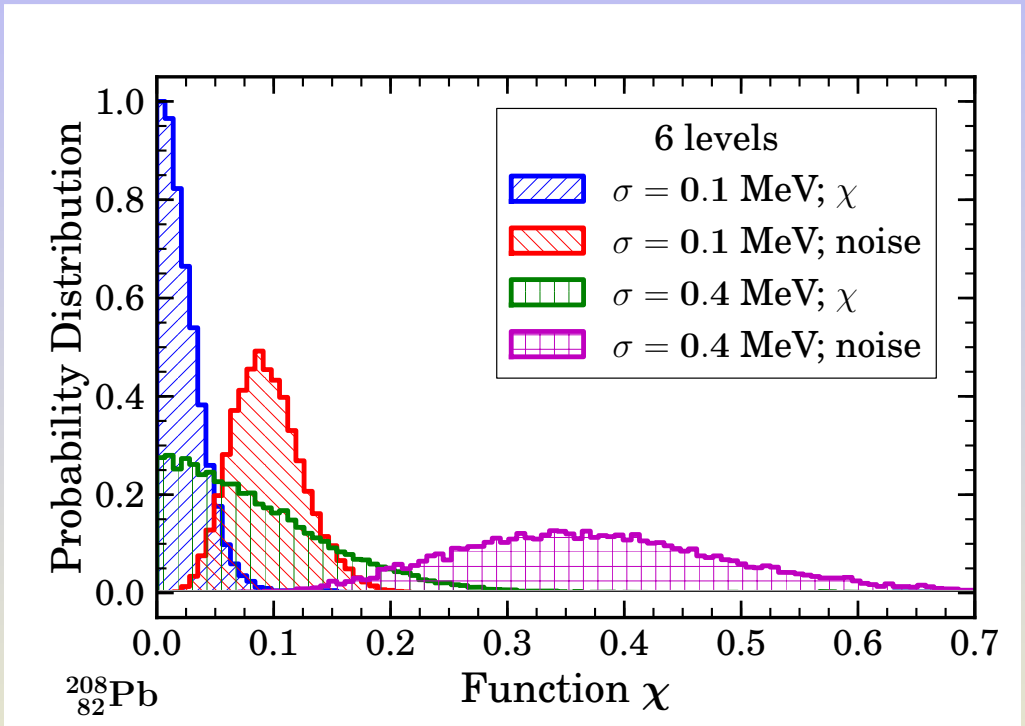
→ Couple of questions may come to one's mind...:

- *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?*

Unprecedented Precision ... Known as 'Overfitting'

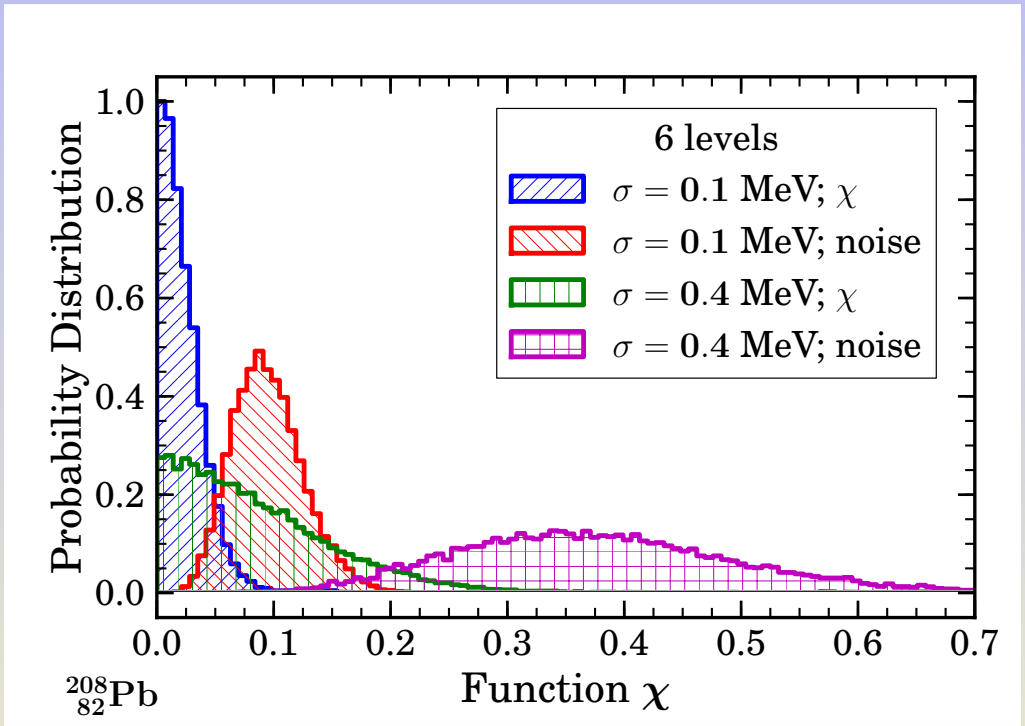


Unprecedented Precision ... Known as 'Overfitting'



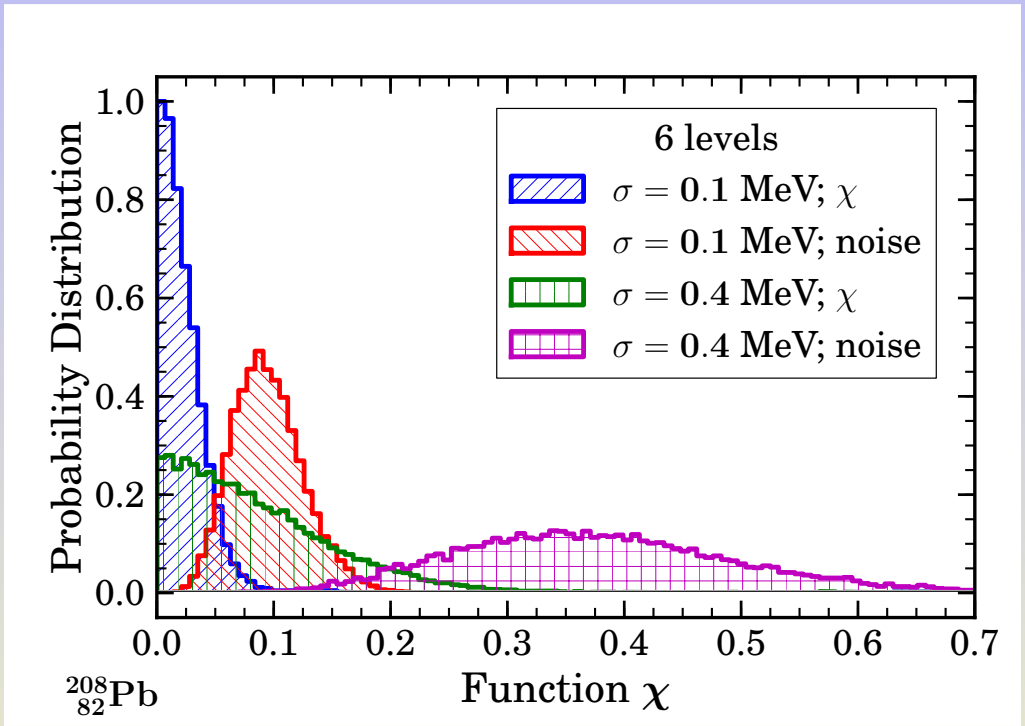
- We introduce the Gaussian noise into the experimental-level input,

Unprecedented Precision ... Known as 'Overfitting'



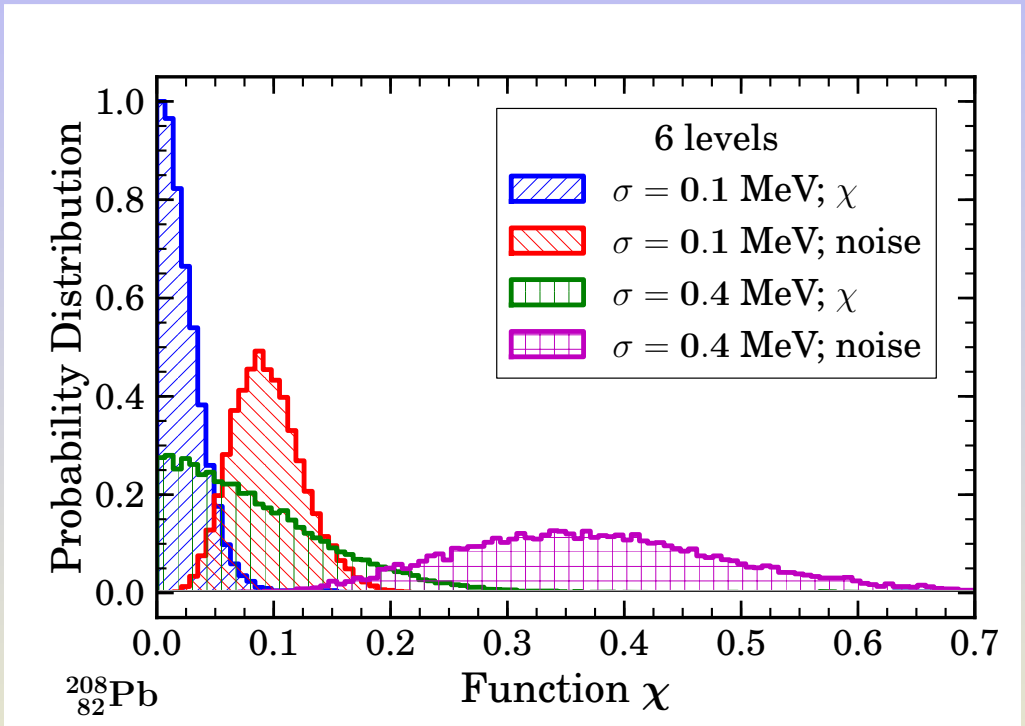
• We introduce the Gaussian noise into the experimental-level input, repeat the χ^2-fit

Unprecedented Precision ... Known as 'Overfitting'



• 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 'Overfitting'



• 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 $N = \infty^6$ exact fits possible. Is it totally trivial?

‘Overfitting’ - What Does It Imply?

‘Overfitting’ - 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

‘Overfitting’ - 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

‘Overfitting’ - 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.

‘Overfitting’ - 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:

‘Overfitting’ - 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...)

‘Overfitting’ - 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...

'Overfitting' - 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

Part II

Overparametrized Nuclear Hamiltonians: Why? How to Deal with the Problem?

Decrease Parametric Freedom: 'Be More Microscopic'

• We suggest to replace a too neat W-S spin-orbit parametrisation by using the density gradient

$$\mathbf{V_{ws}^{so}} \sim \frac{\mathbf{1}}{\mathbf{r}}\frac{\mathbf{d}}{\mathbf{dr}}\left\{\frac{\boldsymbol{\lambda}}{\mathbf{1+\exp[(r-R_o)/a_o]}}\right\} \quad \leftrightarrow \quad \mathbf{V_{ws}^{so}} \sim \frac{\boldsymbol{\lambda}'}{\mathbf{r}}\frac{\mathbf{d}\boldsymbol{\rho}}{\mathbf{dr}}$$

• The nucleonic density can be seen as describing the interaction source: in systems with short range interactions, on the average, the higher the density (gradient) - the more chance to S-O interact.

• Similarly, in the relativistic approach

$$\mathbf{V_{rel}^{so}} \sim \frac{\mathbf{1}}{\mathbf{r}}\frac{\mathbf{d}}{\mathbf{dr}}[\mathbf{S(r)-V(r)}]$$

with S and V expressed by the densities of (source) mesons.

Decrease Parametric Freedom: 'Be More Microscopic'

• We suggest to replace a too neat W-S spin-orbit parametrisation by using the density gradient

$$\mathbf{V_{ws}^{so}} \sim \frac{\mathbf{1}}{\mathbf{r}}\frac{\mathbf{d}}{\mathbf{dr}}\left\{\frac{\boldsymbol{\lambda}}{\mathbf{1+\exp[(r-R_o)/a_o]}}\right\} \quad \leftrightarrow \quad \mathbf{V_{ws}^{so}} \sim \frac{\boldsymbol{\lambda'}}{\mathbf{r}}\frac{\mathbf{d}\boldsymbol{\rho}}{\mathbf{dr}}$$

• The nucleonic density can be seen as describing the interaction source: in systems with short range interactions, on the average, the higher the density (gradient) - the more chance to S-O interact.

• Similarly, in the relativistic approach

$$\mathbf{V_{rel}^{so}} \sim \frac{\mathbf{1}}{\mathbf{r}}\frac{\mathbf{d}}{\mathbf{dr}}[\mathbf{S(r)-V(r)}]$$

with S and V expressed by the densities of (source) mesons.

Decrease Parametric Freedom: 'Be More Microscopic'

• We suggest to replace a too neat W-S spin-orbit parametrisation by using the density gradient

$$\mathbf{V_{ws}^{so}} \sim \frac{\mathbf{1}}{\mathbf{r}}\frac{\mathbf{d}}{\mathbf{dr}}\left\{\frac{\boldsymbol{\lambda}}{\mathbf{1+\exp[(r-R_o)/a_o]}}\right\} \quad \leftrightarrow \quad \mathbf{V_{ws}^{so}} \sim \frac{\boldsymbol{\lambda}'}{\mathbf{r}}\frac{\mathbf{d}\boldsymbol{\rho}}{\mathbf{dr}}$$

• The nucleonic density can be seen as describing the interaction source: in systems with short range interactions, on the average, the higher the density (gradient) - the more chance to S-O interact.

• Similarly, in the relativistic approach

$$\mathbf{V_{rel}^{so}} \sim \frac{\mathbf{1}}{\mathbf{r}}\frac{\mathbf{d}}{\mathbf{dr}}[\mathbf{S(r)-V(r)}]$$

with S and V expressed by the densities of (source) mesons.

A New Toy-Model, Half-Microscopic Hamiltonian

• Consider a given nucleus with the density ρ and a series of neighbouring nuclei with extra occupied orbitals $j_1 \leftrightarrow \rho_{j_1}$, $j_2 \leftrightarrow \rho_{j_2}$, etc. We expect that the density-dependent spin-orbit potentials

$$\mathbf{V^{so}} \sim \frac{\mathbf{d}\rho}{\mathbf{dr}}, \quad \frac{\mathbf{d}\rho_{\mathbf{j_1}}}{\mathbf{dr}}, \quad \frac{\mathbf{d}\rho_{\mathbf{j_2}}}{\mathbf{dr}} \quad \ldots$$

account much better for these extra orbitals than just a flat WS potential introduced long ago for numerical simplicity

• Therefore we will test the following Hartree-Fock like hypothesis

$$\boxed{V_\pi^{so} = \frac{\lambda_{\pi\pi}}{r}\frac{d\rho_\pi}{dr} + \frac{\lambda_{\pi\nu}}{r}\frac{d\rho_\nu}{dr}} \quad \text{and} \quad \boxed{V_\nu^{so} = \frac{\lambda_{\nu\pi}}{r}\frac{d\rho_\pi}{dr} + \frac{\lambda_{\nu\nu}}{r}\frac{d\rho_\nu}{dr}}$$

with the central WS potentials

$$\boxed{V_\pi^{c} = \frac{V_o^\pi}{1+\exp[(r-R_o^\pi)/a_o^\pi]}} \quad \text{and} \quad \boxed{V_\pi^{c} = \frac{V_o^\nu}{1+\exp[(r-R_o^\nu)/a_o^\nu]}}$$

A New Toy-Model, Half-Microscopic Hamiltonian

• Consider a given nucleus with the density ρ and a series of neighbouring nuclei with extra occupied orbitals $j_1 \leftrightarrow \rho_{j_1}$, $j_2 \leftrightarrow \rho_{j_2}$, etc. We expect that the density-dependent spin-orbit potentials

$$\mathbf{V^{so}} \sim \frac{\mathbf{d}\rho}{\mathbf{dr}}, \quad \frac{\mathbf{d}\rho_{\mathbf{j_1}}}{\mathbf{dr}}, \quad \frac{\mathbf{d}\rho_{\mathbf{j_2}}}{\mathbf{dr}} \quad \ldots$$

account much better for these extra orbitals than just a flat WS potential introduced long ago for numerical simplicity

• Therefore we will test the following Hartree-Fock like hypothesis

$$\boxed{V_\pi^{so} = \frac{\lambda_{\pi\pi}}{r}\frac{d\rho_\pi}{dr} + \frac{\lambda_{\pi\nu}}{r}\frac{d\rho_\nu}{dr}} \quad \text{and} \quad \boxed{V_\nu^{so} = \frac{\lambda_{\nu\pi}}{r}\frac{d\rho_\pi}{dr} + \frac{\lambda_{\nu\nu}}{r}\frac{d\rho_\nu}{dr}}$$

with the central WS potentials

$$\boxed{V_\pi^{c} = \frac{V_o^\pi}{1+\exp[(r-R_o^\pi)/a_o^\pi]}} \quad \text{and} \quad \boxed{V_\pi^{c} = \frac{V_o^\nu}{1+\exp[(r-R_o^\nu)/a_o^\nu]}}$$

A New Toy-Model, Half-Microscopic Hamiltonian

• We will try to test two strategies: First of all, the central Woods-Saxon potentials seem to be very robust when 'moving' from one corner of the Periodic Table to another - as independent tests show

• On the other hand we will modify the HF idea of self-consistency from the usual variational context to the spectroscopic context:

$$\mathbf{H}(\rho)\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \;\rightarrow\; \rho = \sum \psi^{*}\psi \;\rightarrow\; \mathbf{H}(\rho)\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \ldots$$

We will iterate to obtain the self-consistency that in this context we call 'auto-reproduction' - it is not a result of energy minimisation!

• **In other words: If at $\mathbf{i}^{\mathbf{th}}$ iteration the spectrum is $\{\mathbf{e}^{\mathbf{i}}_{\mathbf{n}}\}$ and at $\mathbf{i}+\mathbf{1}^{\mathbf{st}} \rightarrow \{\mathbf{e}^{\mathbf{i}+\mathbf{1}}_{\mathbf{n}}\}$, we stop iterating when**

$$|\mathbf{e}^{\mathbf{i}+\mathbf{1}}_{\mathbf{n}} - \mathbf{e}^{\mathbf{i}}_{\mathbf{n}}| < \epsilon, \quad \forall\,\mathbf{n}$$

A New Toy-Model, Half-Microscopic Hamiltonian

• We will try to test two strategies: First of all, the central Woods-Saxon potentials seem to be very robust when 'moving' from one corner of the Periodic Table to another - as independent tests show

• On the other hand we will modify the HF idea of self-consistency from the usual variational context to the spectroscopic context:

$$\mathbf{H}(\rho)\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \;\rightarrow\; \rho = \sum \psi^{*}\psi \;\rightarrow\; \mathbf{H}(\rho)\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \ldots$$

We will iterate to obtain the self-consistency that in this context we call 'auto-reproduction' - it is not a result of energy minimisation!

• **In other words: If at $\mathbf{i}^{\mathbf{th}}$ iteration the spectrum is $\{\mathbf{e}^{\mathbf{i}}_{\mathbf{n}}\}$ and at $\mathbf{i}+\mathbf{1}^{\mathbf{st}} \rightarrow \{\mathbf{e}^{\mathbf{i}+\mathbf{1}}_{\mathbf{n}}\}$, we stop iterating when**

$$|\mathbf{e}^{\mathbf{i}+\mathbf{1}}_{\mathbf{n}} - \mathbf{e}^{\mathbf{i}}_{\mathbf{n}}| < \epsilon, \quad \forall\, \mathbf{n}$$

A New Toy-Model, Half-Microscopic Hamiltonian

• We will try to test two strategies: First of all, the central Woods-Saxon potentials seem to be very robust when 'moving' from one corner of the Periodic Table to another - as independent tests show

• On the other hand we will modify the HF idea of self-consistency from the usual variational context to the spectroscopic context:

$$\mathbf{H}(\rho)\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \;\rightarrow\; \rho = \sum \psi^{*}\psi \;\rightarrow\; \mathbf{H}(\rho)\psi_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}\,\psi_{\mathbf{n}} \ldots$$

We will iterate to obtain the self-consistency that in this context we call 'auto-reproduction' - it is not a result of energy minimisation!

• In other words: If at $\mathbf{i}^{\mathbf{th}}$ iteration the spectrum is $\{\mathbf{e}^{\mathbf{i}}_{\mathbf{n}}\}$ and at $\mathbf{i}+\mathbf{1}^{\mathbf{st}} \rightarrow \{\mathbf{e}^{\mathbf{i}+\mathbf{1}}_{\mathbf{n}}\}$, we stop iterating when

$$|\mathbf{e}^{\mathbf{i}+\mathbf{1}}_{\mathbf{n}} - \mathbf{e}^{\mathbf{i}}_{\mathbf{n}}| < \epsilon, \quad \forall\,\mathbf{n}$$

Self-consistent Formulation: Minimum Coupling

• For the first tests we apply what we call a minimum coupling hypothesis

$$\boxed{\lambda_{\pi\pi} = \lambda_{\pi\nu} = \lambda_{\nu\pi} = \lambda_{\nu\nu} \stackrel{\mathrm{df}}{=} \lambda}$$

• We adjust parameters of the central potential together with $|\lambda|$

• **In other words: We will reduce the number of spin-orbit parameters from 6 (3 per particle kind: λ, r_o^{so} and a^{so}) to one!**

Self-consistent Formulation: Minimum Coupling

• For the first tests we apply what we call a minimum coupling hypothesis

$$\boxed{\lambda_{\pi\pi} = \lambda_{\pi\nu} = \lambda_{\nu\pi} = \lambda_{\nu\nu} \stackrel{\mathrm{df}}{=} \lambda}$$

• We adjust parameters of the central potential together with $|\lambda|$

• **In other words: We will reduce the number of spin-orbit parameters from 6 (3 per particle kind: λ, r_o^{so} and a^{so}) to one!**

Self-consistent Formulation: Minimum Coupling

• For the first tests we apply what we call a minimum coupling hypothesis

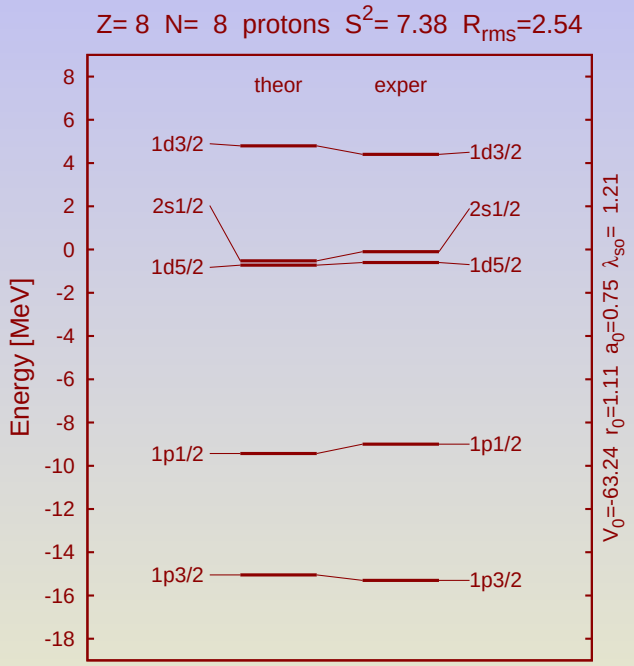
$$\boxed{\lambda_{\pi\pi} = \lambda_{\pi\nu} = \lambda_{\nu\pi} = \lambda_{\nu\nu} \stackrel{\mathrm{df}}{=} \lambda}$$

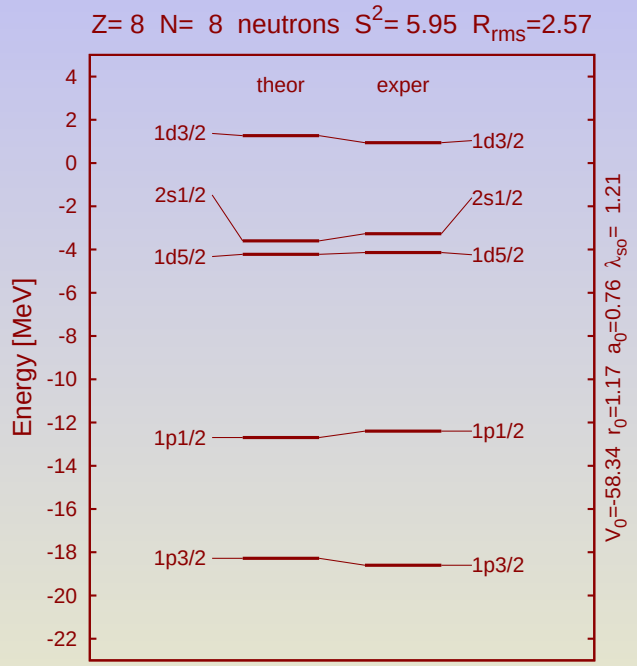
• We adjust parameters of the central potential together with $|\lambda|$

• **In other words: We will reduce the number of spin-orbit parameters from 6 (3 per particle kind: λ, r_o^{so} and a^{so}) to one!**

Comparing with Experimental Results: Example ^{16}O

Single particle proton and neutron states in ^{16}O

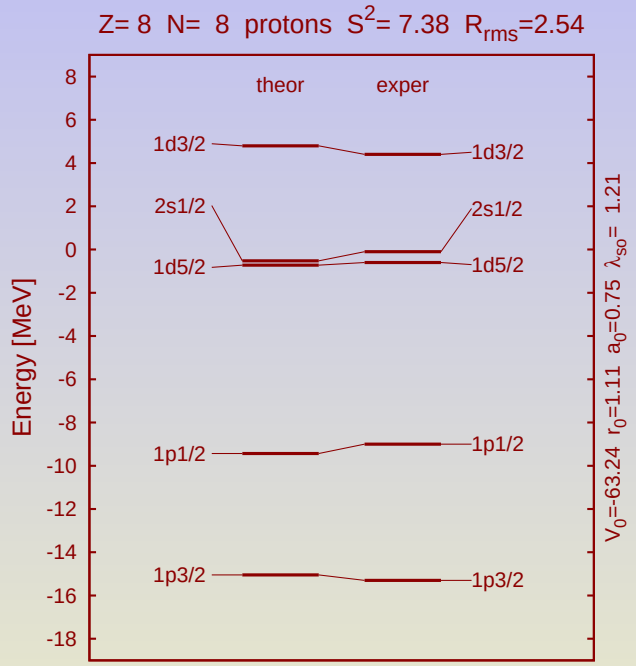


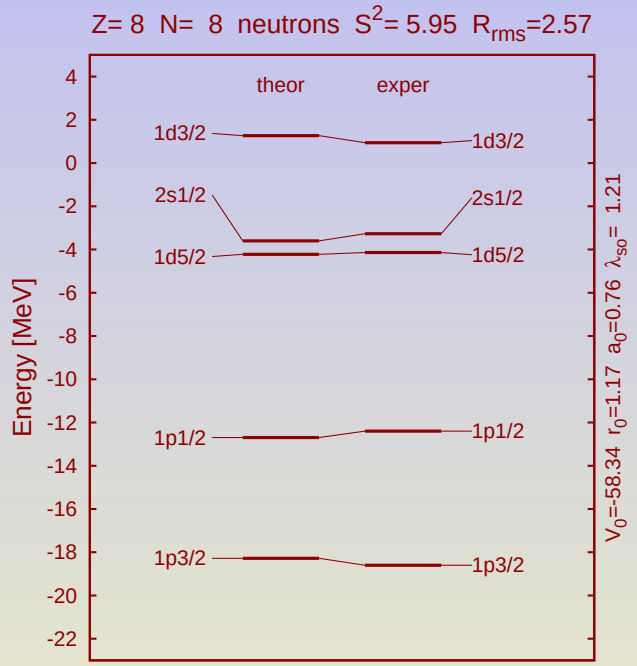


These results corresponds to just one λ-parameter fit instead of 6.

Comparing with Experimental Results: Example ^{16}O

Single particle proton and neutron states in ^{16}O



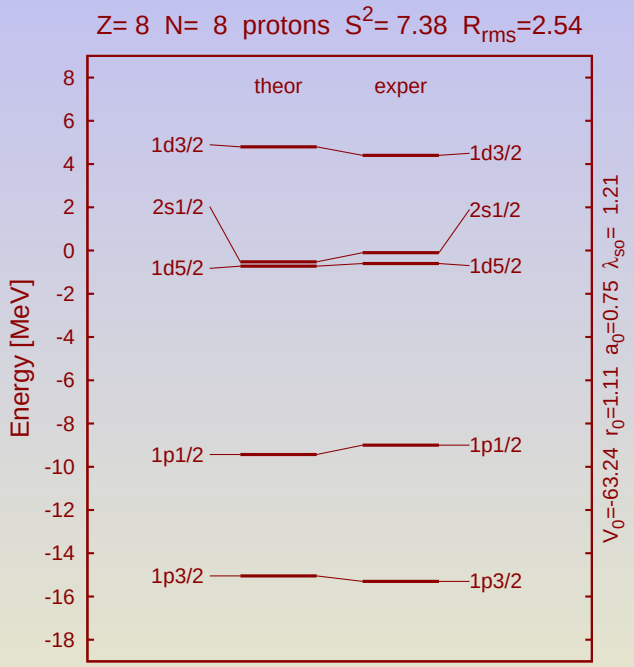


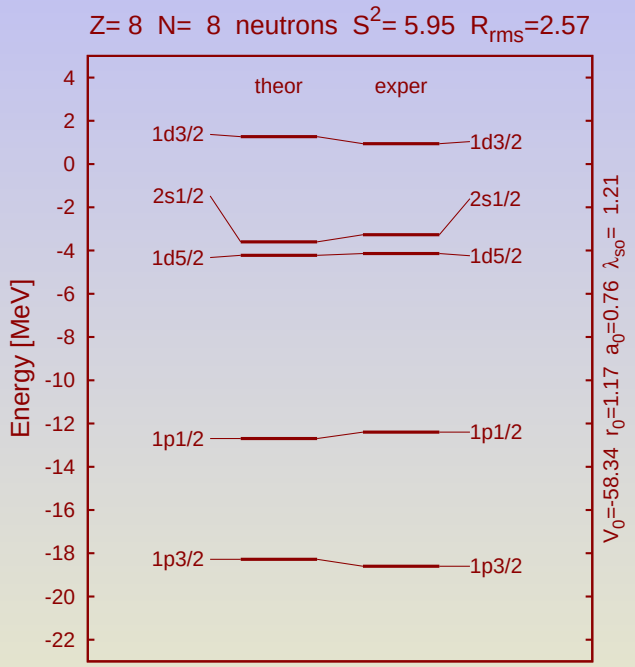
These results corresponds to just one λ-parameter fit instead of 6.

Recall: spin-orbit strength parameter $\lambda = \lambda_{\nu\nu} = \lambda_{\nu\pi} = \lambda_{\pi\nu} = \lambda_{\pi\pi}$.

Comparing with Experimental Results: Example ^{16}O

Single particle proton and neutron states in ^{16}O





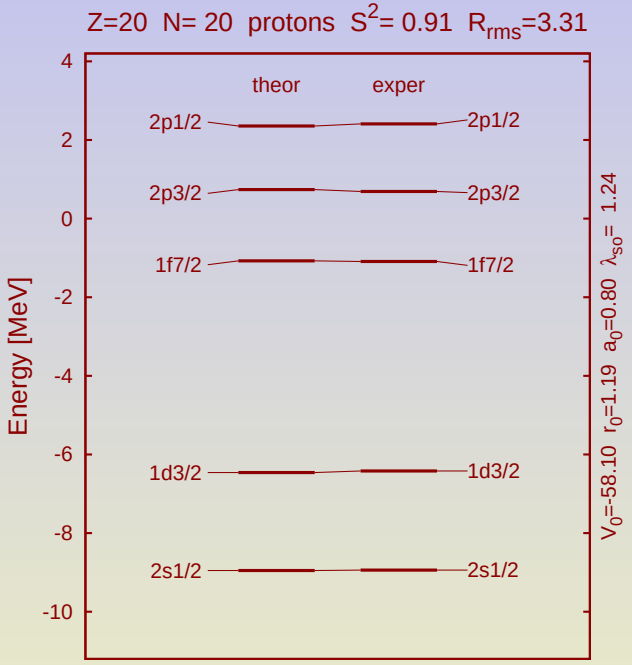
These results corresponds to just one λ-parameter fit instead of 6.

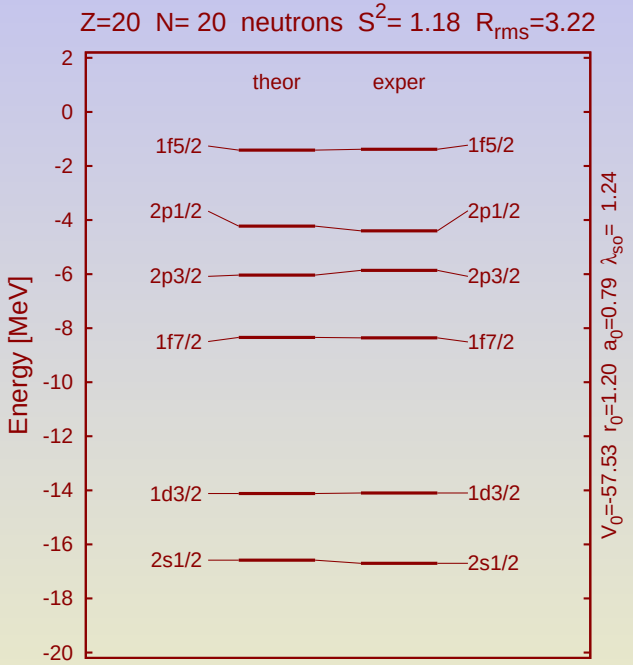
Recall: spin-orbit strength parameter $\lambda = \lambda_{\nu\nu} = \lambda_{\nu\pi} = \lambda_{\pi\nu} = \lambda_{\pi\pi}$.

The maximum error $\sim$ *200 keV*

Comparing with Experimental Results: Example ^{40}Ca

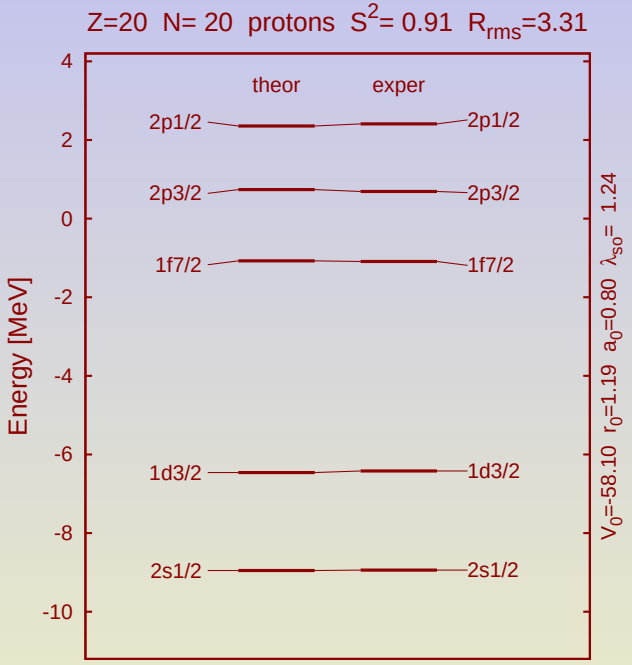
Similarly the single particle proton and neutron states in ^{40}Ca

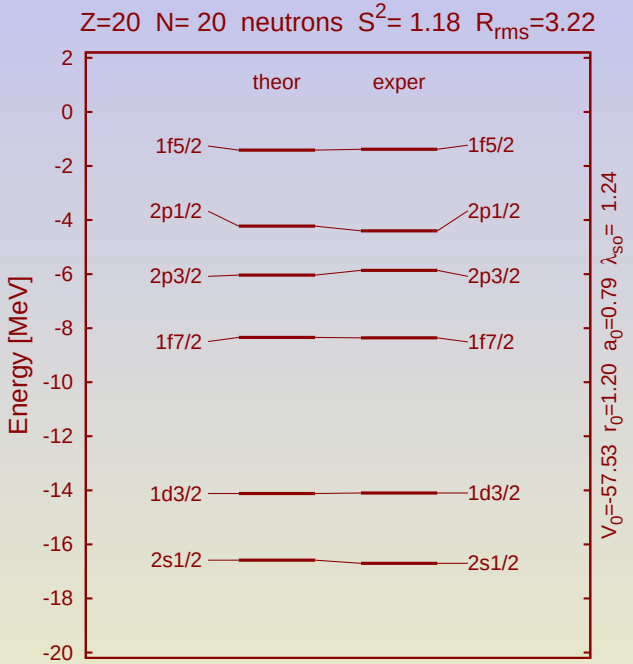




Comparing with Experimental Results: Example ^{40}Ca

Similarly the single particle proton and neutron states in ^{40}Ca

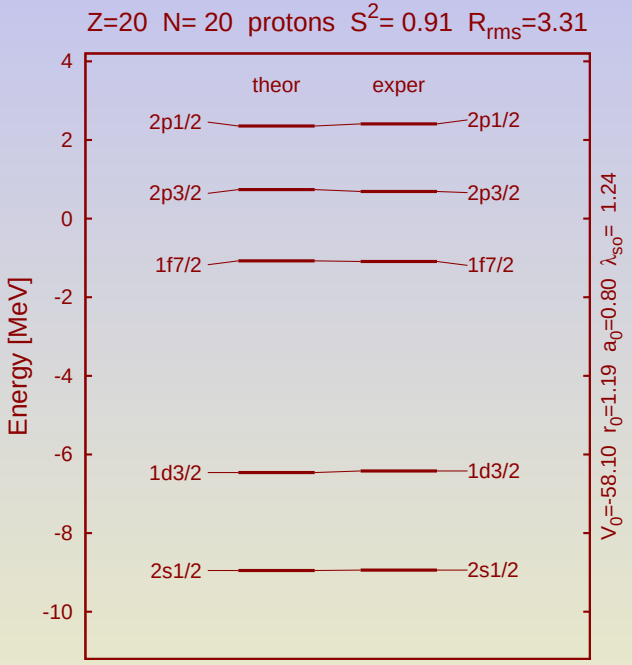


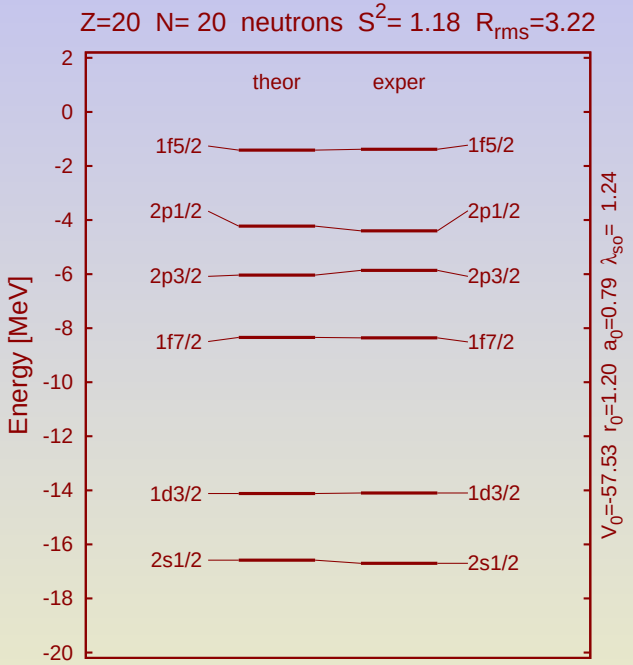


This result corresponds to just one (λ) parameter fit instead of 6; similar results hold for nuclei up to ^{208}Pb.

Comparing with Experimental Results: Example ^{40}Ca

Similarly the single particle proton and neutron states in ^{40}Ca





This result corresponds to just one (λ) parameter fit instead of 6; similar results hold for nuclei up to ^{208}Pb.

We have V_o^π, V_o^ν, r_o^π, r_o^ν and $a_o^\pi = a_o^\nu$ parameters. In the case of ^{208}Pb we have $13_\nu + 11_\pi$ data points.

Removing the Overfit: Increase Physical Significance

Removing the Overfit: Increase Physical Significance

1. **One can fully appreciate the physics success realising that:**

Removing the Overfit: Increase Physical Significance

1. **One can fully appreciate the physics success realising that:**
   - We have replaced <u>six</u> adjustable parameters by a single one

Removing the Overfit: Increase Physical Significance

1. **One can fully appreciate the physics success realising that:**
   - We have replaced <u>six</u> adjustable parameters by a single one
   - Flat-bottom W-S functions are replaced by nucleon densities that carry more physical information through various j-orbitals

Removing the Overfit: Increase Physical Significance

1. **One can fully appreciate the physics success realising that:**
   - We have replaced <u>six</u> adjustable parameters by a single one
   - Flat-bottom W-S functions are replaced by nucleon densities that carry more physical information through various j-orbitals

2. **This result implies that:**

Removing the Overfit: Increase Physical Significance

1. **One can fully appreciate the physics success realising that:**
   - We have replaced <u>six</u> adjustable parameters by a single one
   - Flat-bottom W-S functions are replaced by nucleon densities that carry more physical information through various j-orbitals

2. **This result implies that:**
   - The new, simple and in a way natural notion of self-consistency works in a powerful manner

Removing the Overfit: Increase Physical Significance

1. **One can fully appreciate the physics success realising that:**
   - We have replaced six adjustable parameters by a single one
   - Flat-bottom W-S functions are replaced by nucleon densities that carry more physical information through various j-orbitals

2. **This result implies that:**
   - The new, simple and in a way natural notion of self-consistency works in a powerful manner
   - Most importantly, Fits show that the density fluctuations are needed for the gradients in the realistic spin-orbit terms!

Conclusions:

Conclusions:

• **We presented a few guide-lines followed in our project that aims at quantifying the Question of Predictive Power of Nuclear Hamiltonians**

Conclusions:

• We presented a few guide-lines followed in our project that aims at quantifying the Question of Predictive Power of Nuclear Hamiltonians

• The audience is warned not to be mislead by the simplicity of the illustrations based on the toy model (here: spherical Woods-Saxon) vs. generality and importance of the Inverse-Problem Theory which applies to all realistic Hamiltonians

Conclusions:

• We presented a few guide-lines followed in our project that aims at quantifying the Question of Predictive Power of Nuclear Hamiltonians

• The audience is warned not to be mislead by the simplicity of the illustrations based on the toy model (here: spherical Woods-Saxon) vs. generality and importance of the Inverse-Problem Theory which applies to all realistic Hamiltonians

• Needless to say - we aim at the microscopic level (theories), in particular HF - but today there will still be some simple semi-quantitative illustrations $\rightarrow$ presented by another speaker