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

September 28, 2011

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 ($\sim$0.18 seconds)
"Inverse Problem"	510 000 results ($\sim$0.36 seconds)

COLLABORATORS:

Arthur DROMARD, UdS/IPHC-CNRS Strasbourg
Bartłomiej SZPAK, IFJ Kraków
Marie-Geneviève PORQUET, CSNSM Orsay

Bogdan FORNAL, IFJ Kraków
Hervé MOLIQUE, UdS/IPHC-CNRS Strasbourg
Karolina RYBAK, UdS/IPHC-CNRS 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: 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**

$$\boxed{\hat{H}\,\psi_n = e_n\,\psi_n \quad \leftrightarrow \quad \hat{H} = \hat{H}^{true}}$$

- **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 complete information

About this Particular Project: What? and Why?

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

$$\boxed{\hat{H}\,\psi_n = e_n\,\psi_n \quad \leftrightarrow \quad \hat{H} = \hat{H}^{true}}$$

where by definition $\hat{H}^{true}$ contains complete information

- **Complete - both: already known and yet undiscovered**

About this Particular Project: What? and Why?

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

$$\boxed{\hat{H}\,\psi_n = e_n\,\psi_n \quad \leftrightarrow \quad \hat{H} = \hat{H}^{true}}$$

where by definition $\hat{H}^{true}$ contains complete information

- **Complete - both: already known and yet undiscovered**

Conclusion 1: The solutions $\{e_n, \psi_n\}$ 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 complete information

- **Complete - both: already known and 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 relative 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 <u>relative</u> 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 Theoretical and Experimental Errors

A Few Comments about the Nature of Errors

- Theories are incomplete whereas experiments plagued with errors:

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

e_n and ε_n are random variables $\longrightarrow$ 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:

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

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

- Errors propagate to the theory predictions through parameter fits

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

so that the parameters $p \equiv \{p_1, p_2, \ldots \; p_f\}$ are random variables and as such are characterised by the probability distributions

A Few Comments about the Nature of Errors

- Theories are incomplete whereas experiments plagued with errors:

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

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

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

so that the parameters $p \equiv \{p_1, p_2, \ldots \; p_f\}$ are random variables and as such are characterised by the probability distributions

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

A Few Comments about the Nature of Errors

- Theories are incomplete whereas experiments plagued with errors:

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

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

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

so that the parameters $p \equiv \{p_1, p_2, \dots \; p_f\}$ are random variables and as such are characterised by the probability distributions

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

- Conclusion: All the predictions have the probability distributions!

This is a new strategical observation which introduces what we call a "Theory about Constructing Theories"

This is a new strategical observation which introduces what we call a
"Theory about Constructing Theories"

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 a "Theory about Constructing Theories"

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 $\{\hat{\mathcal{F}}_1 : f_1, \hat{\mathcal{F}}_2 : f_2, \dots \; \hat{\mathcal{F}}_p : f_p\}$

This is a new strategical observation which introduces what we call a
"Theory about Constructing Theories"

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:

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

Mean-Field Theory and Its General Importance

• **We assume that the Nuclear Mean-Field Theory will remain one of the fundamental nuclear-physics tools for a good while**

Mean-Field Theory and Its General Importance

• **We assume that the Nuclear Mean-Field Theory will remain one of the fundamental nuclear-physics tools for a good while**

• **Existing developments, especially the Relativistic Mean-Field (explicit link with particle physics) and Density Functionals (Kohn-Sham theorem, its potentialities in nuclear physics) do represent the best we can and should do in the coming future**

• We assume that the Nuclear Mean-Field Theory will remain one of the fundamental nuclear-physics tools for a good while

• Existing developments, especially the Relativistic Mean-Field (explicit link with particle physics) and Density Functionals (Kohn-Sham theorem, its potentialities in nuclear physics) do represent the best we can and should do in the coming future

• Possibilities of extrapolations to unknown, extreme isospin areas remain - using the mildest formulation - unsatisfactory!

Mean-Field Theory and Its General Importance

• We assume that the Nuclear Mean-Field Theory will remain one of the fundamental nuclear-physics tools for a good while

• Existing developments, especially the Relativistic Mean-Field (explicit link with particle physics) and Density Functionals (Kohn-Sham theorem, its potentialities in nuclear physics) do represent the best we can and should do in the coming future

• Possibilities of extrapolations to unknown, extreme isospin areas remain - using the mildest formulation - unsatisfactory!

• Questions of statistical significance of the theory modeling the most basic questions asked in other domains of research are seldom - if ever - posed $\longrightarrow$ This will be our subject today

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!*

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!*

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!*

• 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']

The So-Called 'Experimental' Single-Particle Levels

The So-Called 'Experimental' Single-Particle Levels

- From Audi et al., *Nucl. Phys. A729 (2003) 345*, we find that:

Neutron separation from ^{91}Zr: $S_n = 7194.5$ keV

The So-Called 'Experimental' Single-Particle Levels

- From Audi et al., *Nucl. Phys. A729 (2003) 345*, we find that:

Neutron separation from ^{91}Zr: $S_n = 7194.5$ keV

- An example: the data for $\nu d_{3/2}$ states $(3/2^+)$:

$$\begin{array}{lll} E_1 = 2042\ keV & \text{with} & S_1 = 0.78 \\ E_2 = 2871\ keV & \text{with} & S_1 = 0.08 \\ E_3 = 3083\ keV & \text{with} & S_1 = 0.16 \\ E_4 = 3290\ keV & \text{with} & S_1 = 0.22 \\ E_2 = 3681\ keV & \text{with} & \underline{S_1 = 0.16} \end{array}$$

$$\boxed{\langle \mathbf{E} \rangle = \left(\textstyle\sum_\mathbf{i} \mathbf{E_i} \times \mathbf{S_i}\right)/\left(\textstyle\sum_\mathbf{i} \mathbf{S_i}\right)} \quad \longrightarrow\longrightarrow\longrightarrow\longrightarrow \quad \langle E \rangle = 2592\ keV$$

$$\boxed{E_{d_{3/2}} = -S_n + \langle E \rangle = (-7194.5 + 2592)\ keV = -4602.5\ keV}$$

The So-Called 'Experimental' Single-Particle Levels

1. Many energy levels need to be measured in order to extract one, average energy, interpreted as an experimental single particle energy

The So-Called 'Experimental' Single-Particle Levels

1. Many energy levels need to be measured in order to extract one, average energy, interpreted as an experimental single particle energy

2. To obtain reliable results, attention should be payed to measuring all the j^π states that contribute significantly to the weighted average

The So-Called 'Experimental' Single-Particle Levels

1. Many energy levels need to be measured in order to extract one, average energy, interpreted as an experimental single particle energy

2. To obtain reliable results, attention should be payed to measuring all the j^π states that contribute significantly to the weighted average

3. Vibration-coupling contributions need to be measured and one should aim at subtracting them (or alternatively: include in theory)!

The So-Called 'Experimental' Single-Particle Levels

1. Many energy levels need to be measured in order to extract one, average energy, interpreted as an experimental single particle energy

2. To obtain reliable results, attention should be payed to measuring all the j^π states that contribute significantly to the weighted average

3. Vibration-coupling contributions need to be measured and one should aim at subtracting them (or alternatively: include in theory)!

4. Levels depend, among others, on the spectroscopic factors, defined in the presence of simplifying assumptions in reaction theory; the latter may facilitate the self-control through the sum rule tests

The So-Called 'Experimental' Single-Particle Levels

1. Many energy levels need to be measured in order to extract one, average energy, interpreted as an experimental single particle energy

2. To obtain reliable results, attention should be payed to measuring all the j^π states that contribute significantly to the weighted average

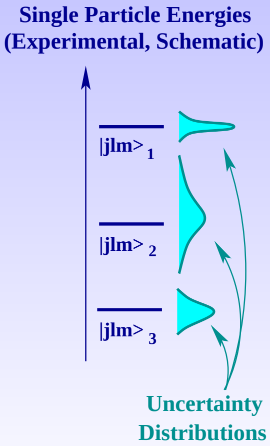
3. Vibration-coupling contributions need to be measured and one should aim at subtracting them (or alternatively: include in theory)!

4. Levels depend, among others, on the spectroscopic factors, defined in the presence of simplifying assumptions in reaction theory; the latter may facilitate the self-control through the sum rule tests

5. Paradoxally, the so-called experimental single-particle levels are highly complicated, model-dependent objects - this leads to errors!

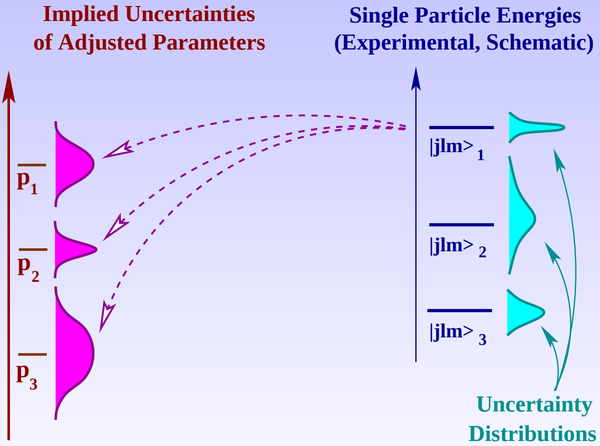
Single-Particle Levels as Probability Distributions

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

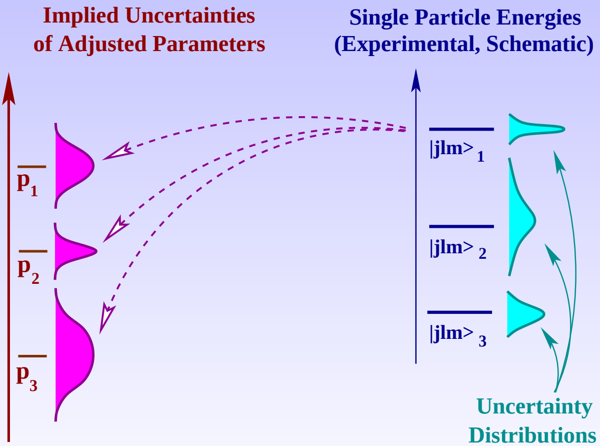


Single-Particle Levels as Probability Distributions

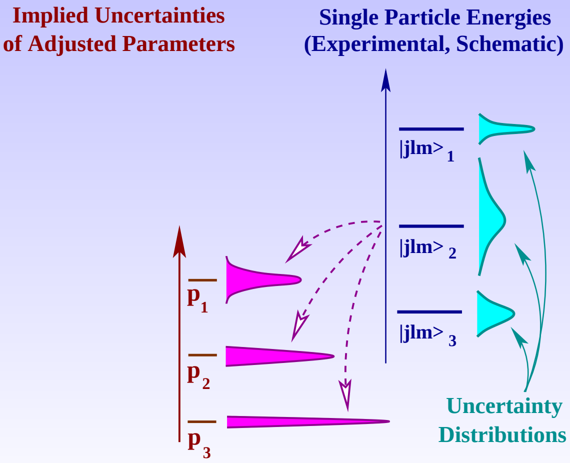
Uncertainties propagate to sought parameters $\{p_1, p_2, \ldots\} \equiv \{p\}$ in $\hat{H}(\hat{r}, \hat{p}; \{p\})\,\psi_n = e_n(p)\,\psi_n$: Parameters $\rightarrow$ Probability Distributions



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

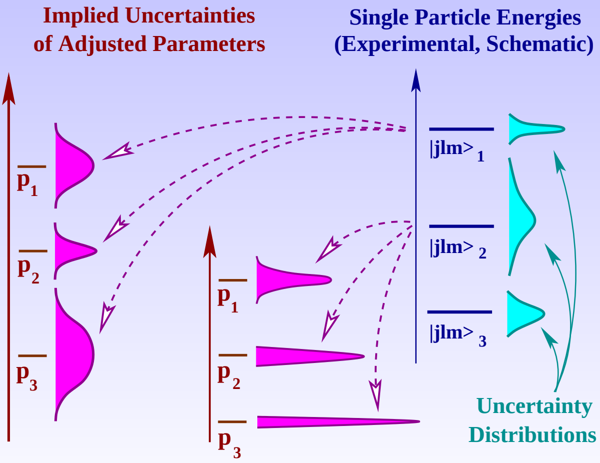


... while at the same time big uncertainties on other levels have a small impact: as in life, all are equal but some more than the others



Single-Particle Levels as Probability Distributions

Conclusion: the quality of adjustement depends strongly on the quality of the data implying the existence of <u>theoretical error bars</u>



Let Us Come Back to the Predictive Power

Let Us Come Back to the Predictive Power

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

Let Us Come Back to the Predictive Power

• 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.

Let Us Come Back to the Predictive Power

• 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!

Let Us Come Back to the Predictive Power

• 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 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 <u>before</u> the right experiments are performed can be called "a theoretical prediction"**

- **... thus performing any model calculation can be called <u>predicting</u>!**

- **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 criteria of choice at the same time:

$\boxed{\text{The notion of Predictive Power is relative and/or subjective}^{\#)}}$

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 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)*

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

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

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

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

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

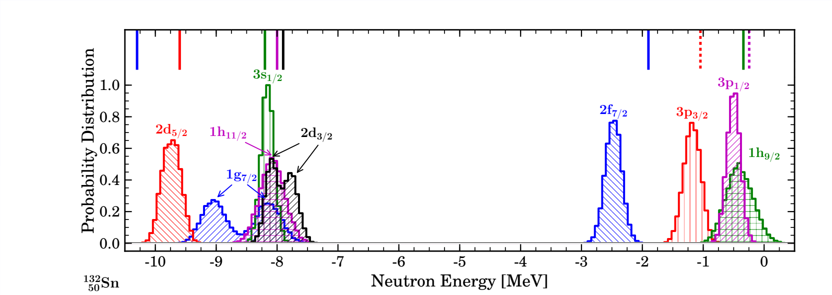
... and continue with presenting as much as Time
(and the Chairman)
permit

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

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;

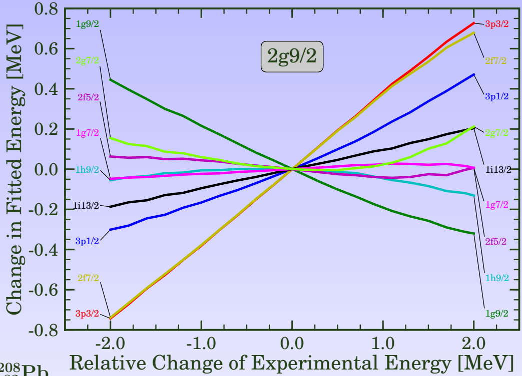
• Chose one level, say $e^o_\kappa \in \{e^o_\nu\}$, and arbitrarily modify its position:

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

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

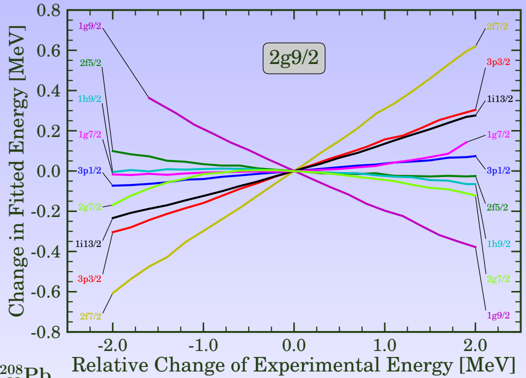
Example: Error Response Functions to $2g_{9/2}$-Orbital



$^{208}_{82}$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



$^{208}_{82}$Pb

<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 the 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 II

Inverse Problem of Applied Mathematics

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>**

- **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'**

- **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'**

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

THE PREDICTIVE POWER

- 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:

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

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

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

- 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

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

- 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

- 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 $e_j^{th}(p) \rightarrow f_j$]

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

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

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

χ^2-Problem Using Applied-Mathematics' Jargon

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

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

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

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

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

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

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

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:

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

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

- 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 $D = \mathrm{diag}\{\underbrace{d_1, d_2, \ \dots \ d_{min(m,n)}}_{\text{decreasing order}}\}$

• 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

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

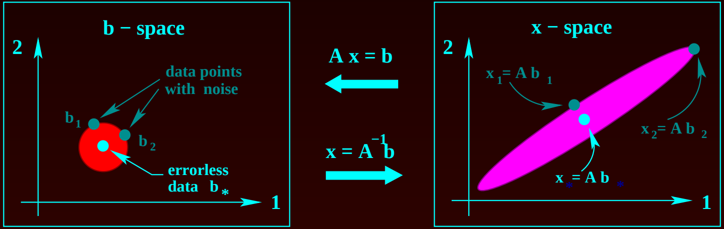
Ill-Conditioned Problems: Qualitative Illustration

Ill-Conditioned Problems: Qualitative Illustration

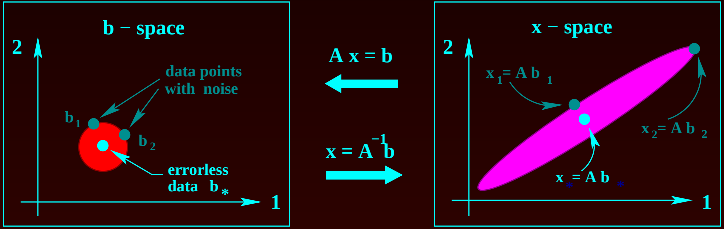
- An academic 2×2 problem: A has eigenvalues: d_1 and $d_2 \sim \frac{1}{10} d_1$

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



- 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

- 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:

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

- Using the Singular-Value Decomposition written down explicitly

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

Fitting, Implied Errors and Confidence Intervals

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

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

- Using the Singular-Value Decomposition written down explicitly

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

- Independently one derives the expression for the correlation matrix

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

Fitting, Implied Errors and Confidence Intervals

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

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

- 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

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

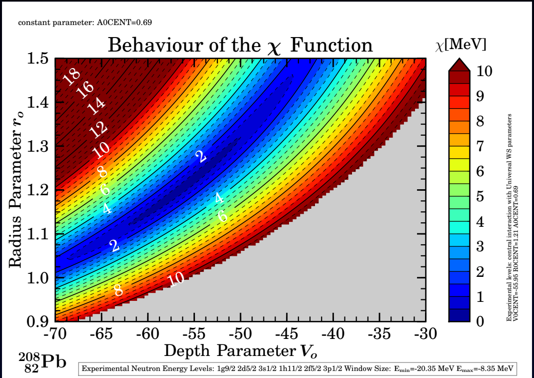
- **If one or more $d_k \rightarrow 0$ then $(J^T J)^{-1}$ tends to infinity and generally, the confidence intervals of all parameters diverge**

Part III

Totally Undesired Parameteric Correlations

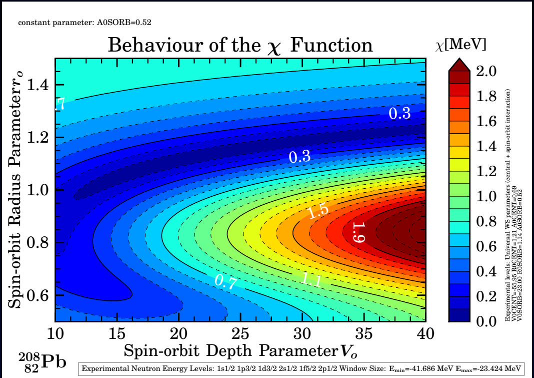
Undesired Parametric Correlations: Illustrative Examples

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.

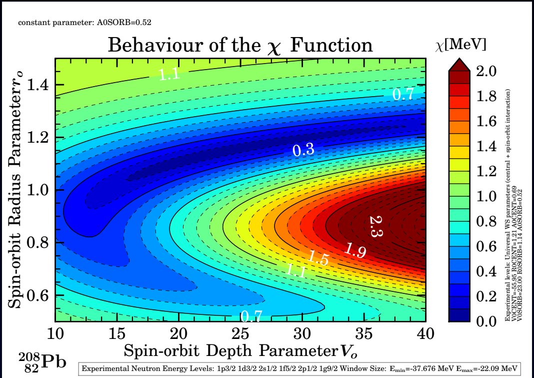
Parametric Correlations and Their Consequences



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

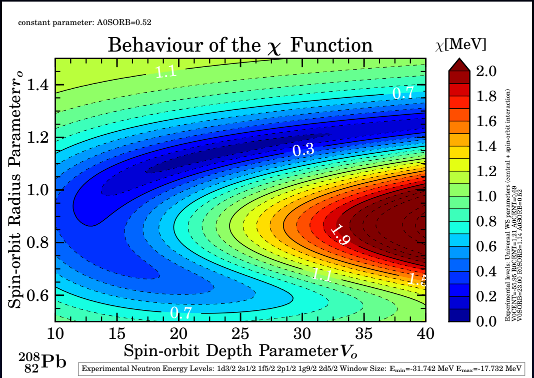
We will gradually increase the energy of the six-level window to approach the nucleon binding region and thus simulate the present-day experimental situation

Limited Experimental Input: How Little is Sufficient?



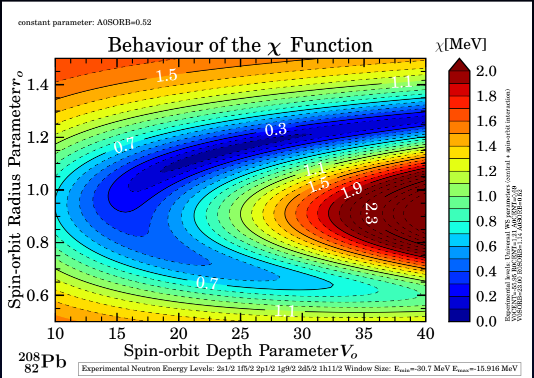
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly!

Limited Experimental Input: How Little is Sufficient?



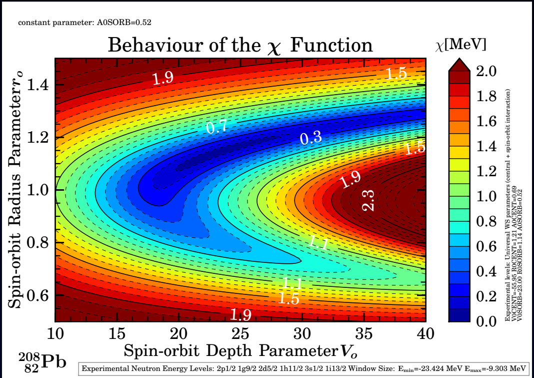
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



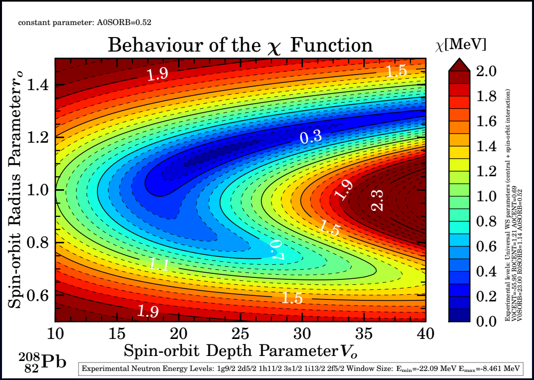
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



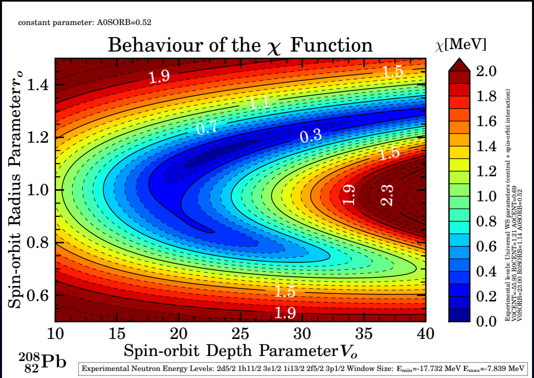
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



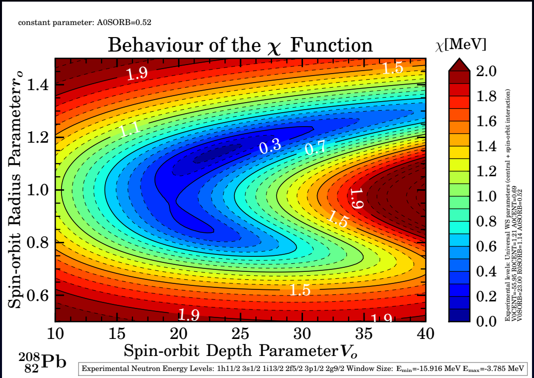
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly...

Limited Experimental Input: How Little is Sufficient?



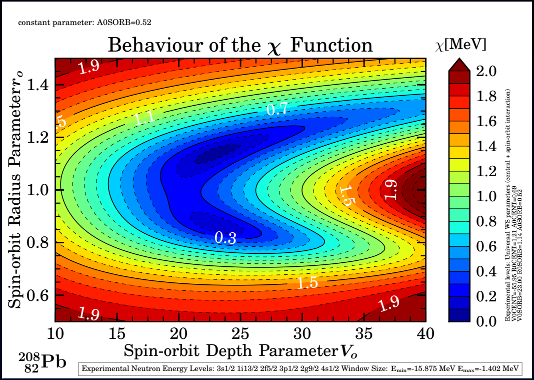
Increasing the energy of the six levels helps localising the spin-orbit strength only very slowly... whereas gradually another solution ...

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 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}**

What Can We Conclude from 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 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 $\longrightarrow$ with all the implied subjectivity

What Can We Conclude from 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 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 $\longrightarrow$ 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

If the confidence intervals diverge we loose unique answers,

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

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".

Part IV

Predicitive Power in Terms of Soluble Models

Predictive Power in Terms of Soluble Models

- **A central issue about predictive power of any given theory is the determination of its Hamiltonian (read: its parameters)**

Predictive Power in Terms of Soluble Models

- **A central issue about predictive power of any given theory is the determination of its Hamiltonian (read: its parameters)**

- **This implies: Solving the Inverse Problem (not χ-by-the eye)**

Predictive Power in Terms of Soluble Models

- **A central issue about predictive power of any given theory is the determination of its Hamiltonian (read: its parameters)**

- **This implies: Solving the Inverse Problem (not χ-by-the eye)**

- **The problems (seemingly not known in sub-atomic physics):**
  - **Hamiltonians contain inter-dependent parameters**
  - **This implies that the Inverse Problem is ill-posed**
  - **This implies that the theory has no predictive power**

Predictive Power in Terms of Soluble Models

- **A central issue about predictive power of any given theory is the determination of its Hamiltonian (read: its parameters)**
- **This implies: Solving the Inverse Problem (not χ-by-the eye)**
- **The problems (seemingly not known in sub-atomic physics):**
  - **Hamiltonians contain inter-dependent parameters**
  - **This implies that the Inverse Problem is ill-posed**
  - **This implies that the theory has no predictive power**
- **The problem is serious as illustrated below using an exactly soluble modelling**

Predictive Power in Terms of Soluble Models

- **A central issue about predictive power of any given theory is the determination of its Hamiltonian (read: its parameters)**

- **This implies: Solving the Inverse Problem (not χ-by-the eye)**

- **The problems (seemingly not known in sub-atomic physics):**
  - **Hamiltonians contain inter-dependent parameters**
  - **This implies that the Inverse Problem is ill-posed**
  - **This implies that the theory has no predictive power**

- **The problem is serious as illustrated below using an exactly soluble modelling but the presence on the market of over 130 non-equivalent parametrizations of the Skyrme-HF Hamiltonian is a strong signal!**

Predictive Power in Terms of Soluble Models

- **A central issue about predictive power of any given theory is the determination of its Hamiltonian (read: its parameters)**

- **This implies: Solving the Inverse Problem (not χ-by-the eye)**

- **The problems (seemingly not known in sub-atomic physics):**
  - **Hamiltonians contain inter-dependent parameters**
  - **This implies that the Inverse Problem is ill-posed**
  - **This implies that the theory has no predictive power**

- **The problem is serious as illustrated below using an exactly soluble modelling but the presence on the market of over 130 non-equivalent parametrizations of the Skyrme-HF Hamiltonian is a strong signal! - And one has to stop the non-sense!!**

Predictive Power in Terms of Soluble Models

- **A mathematical model with a Hamiltonian will be replaced by an equivalent one [same mechanisms]: Fitting a function**

Predictive Power in Terms of Soluble Models

- **A mathematical model with a Hamiltonian will be replaced by an equivalent one [same mechanisms]: Fitting a function**
- **The issue: Find parameters a, b, c and d in the formula**

$$\boxed{\exp(x) = a \cdot x + b + c \cdot \sinh(x) + d \cdot \cosh(x)}$$

Predictive Power in Terms of Soluble Models

- **A mathematical model with a Hamiltonian will be replaced by an equivalent one [same mechanisms]: Fitting a function**
- **The issue: Find parameters a, b, c and d in the formula**

$$\boxed{\exp(x) = a \cdot x + b + c \cdot \sinh(x) + d \cdot \cosh(x)}$$

- **Introduce 'experiment' by selecting the points $x_1, x_2, \ldots x_n$:**

$$\boxed{\text{Experiment:} \quad y_1 = \exp(x_1), \ldots y_n = \exp(x_n)}$$

- **A mathematical model with a Hamiltonian will be replaced by an equivalent one [same mechanisms]: Fitting a function**
- **The issue: Find parameters a, b, c and d in the formula**

$$\boxed{\exp(x) = a \cdot x + b + c \cdot \sinh(x) + d \cdot \cosh(x)}$$

- **Introduce 'experiment' by selecting the points $x_1, x_2, \ldots x_n$:**

$$\boxed{\text{Experiment:} \quad y_1 = \exp(x_1), \ldots y_n = \exp(x_n)}$$

- **One may think: Naive people! The exact solution is known:**

$$\boxed{a = 0,\ b = 0,\ c = 1,\ d = 1}$$

Predictive Power in Terms of Soluble Models

- **A mathematical model with a Hamiltonian will be replaced by an equivalent one [same mechanisms]: Fitting a function**
- **The issue: Find parameters a, b, c and d in the formula**

$$\boxed{\exp(x) = a \cdot x + b + c \cdot \sinh(x) + d \cdot \cosh(x)}$$

- **Introduce 'experiment' by selecting the points x_1, x_2, ... x_n:**

$$\boxed{\text{Experiment:} \quad y_1 = \exp(x_1), \ldots y_n = \exp(x_n)}$$

- **One may think: Naive people! The exact solution is known:**

$$\boxed{a = 0,\ b = 0,\ c = 1,\ d = 1}$$

- **What useful could we learn from such a 'naive' formulation?**

Predictive Power in Terms of Soluble Models

- **First of all we will need to model the experimental conditions**

- **First of all we will need to model the experimental conditions**
- **Therefore we introduce Gaussian experim. error distributions**

$$\boxed{y_i \longrightarrow \bar{y}_i + \delta y_i; \quad \delta y_i \sim \frac{1}{2\pi\sigma^2} \exp[-(y_i - \bar{y}_i)^2/2\sigma^2]}$$

Predictive Power in Terms of Soluble Models

- **First of all we will need to model the experimental conditions**
- **Therefore we introduce Gaussian experim. error distributions**

$$\boxed{y_i \to \bar{y}_i + \delta y_i; \quad \delta y_i \sim \frac{1}{2\pi\sigma^2}\exp[-(y_i - \bar{y}_i)^2/2\sigma^2]}$$

- **We generate a big number N (say $N \sim 40\,000$) of n-tuplets**

$$\boxed{\{y_1 \ldots y_n\}_1; \;\; \{y_1 \ldots y_n\}_2; \;\; \ldots \;\; \{y_1 \ldots y_n\}_N;}$$

Predictive Power in Terms of Soluble Models

- **First of all we will need to model the experimental conditions**
- **Therefore we introduce Gaussian experim. error distributions**

$$\boxed{y_i \to \bar{y}_i + \delta y_i; \quad \delta y_i \sim \frac{1}{2\pi\sigma^2}\exp[-(y_i - \bar{y}_i)^2/2\sigma^2]}$$

- **We generate a big number N (say N $\sim$ 40 000) of n-tuplets**

$$\boxed{\{y_1 \ldots y_n\}_1; \;\; \{y_1 \ldots y_n\}_2; \;\; \ldots \;\; \{y_1 \ldots y_n\}_N;}$$

- **We launch the minimisation $\to$ N quadruplets of $\{a, b, c, d\}$**

Predictive Power in Terms of Soluble Models

- **First of all we will need to model the experimental conditions**
- **Therefore we introduce Gaussian experim. error distributions**

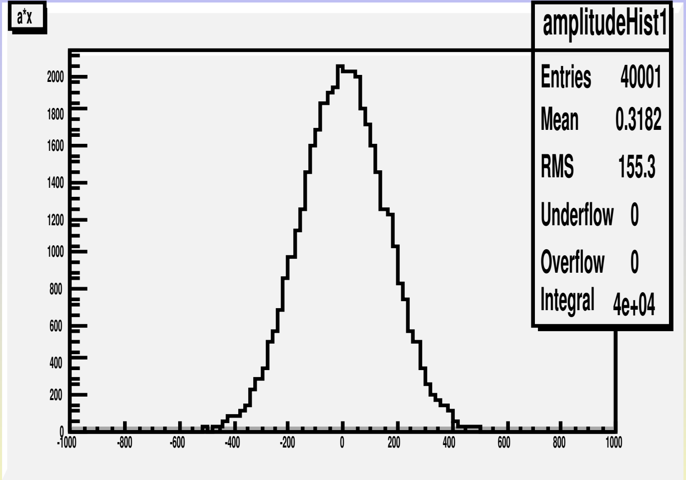
$$\boxed{y_i \to \bar{y}_i + \delta y_i; \quad \delta y_i \sim \frac{1}{2\pi\sigma^2}\exp[-(y_i - \bar{y}_i)^2/2\sigma^2]}$$

- **We generate a big number N (say $N \sim 40\,000$) of n-tuplets**

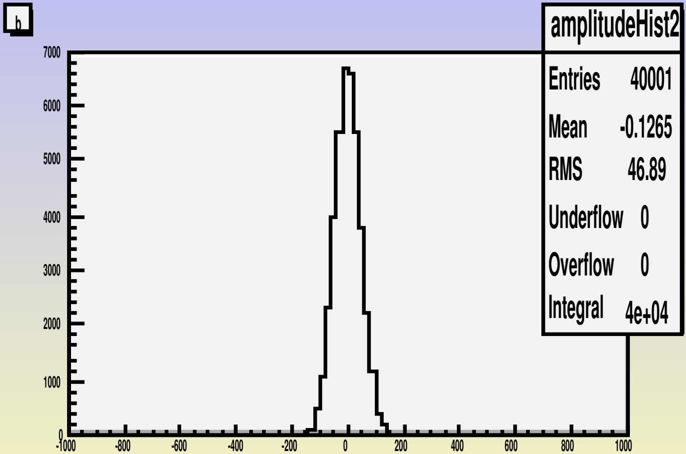
$$\boxed{\{y_1 \dots y_n\}_1; \;\; \{y_1 \dots y_n\}_2; \;\; \dots \;\; \{y_1 \dots y_n\}_N;}$$

- **We launch the minimisation $\longrightarrow$ N quadruplets of $\{a, b, c, d\}$**
- **We construct histograms of occurrence of each parameter; After normalisation $\longrightarrow$ probability distributions of a, b, c & d**

Predictive Power in Terms of Soluble Models

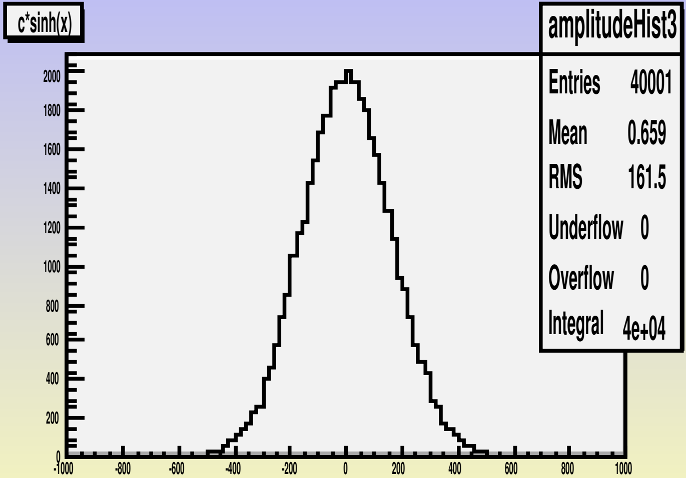


Probability Distribution of the *a*-parameter of the ‘theory’

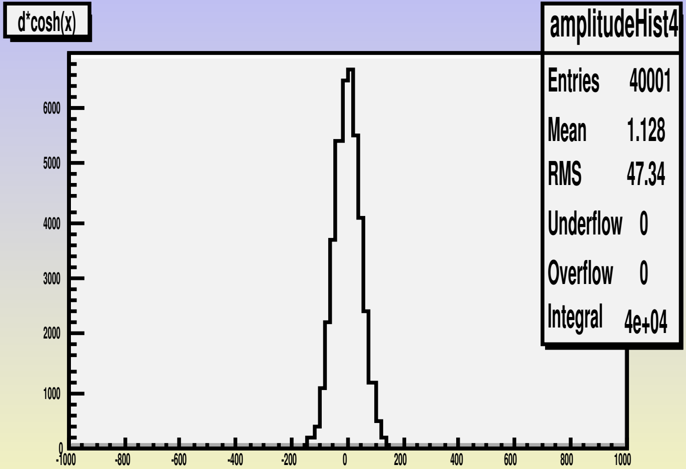


Probability Distribution of the *b*-parameter of the ‘theory’

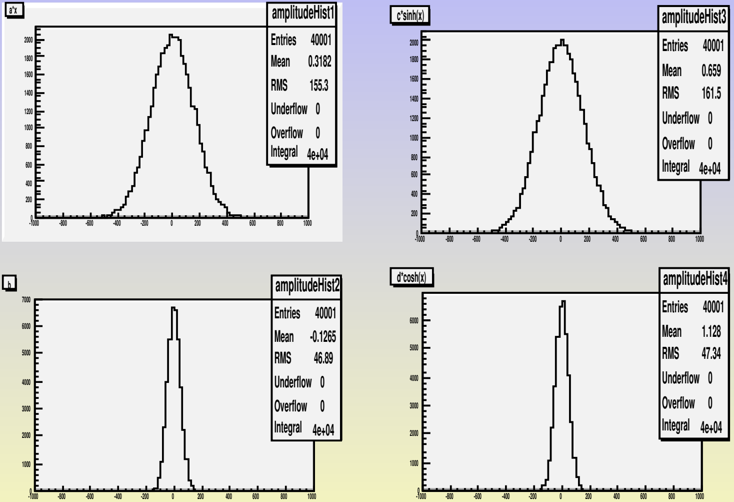
Predictive Power in Terms of Soluble Models



Probability Distribution of the *c*-parameter of the 'theory'



Probability Distribution of the *d*-parameter of the 'theory'

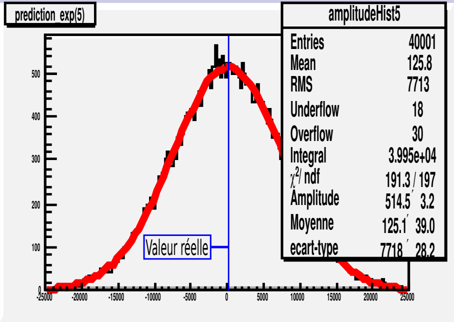


Observe different behaviour of the positive and negative parities

Predictive Power in Terms of Soluble Models

- **Extraneous Predictive Power $\leftrightarrow$ Extrapolations by theory:**

$$\boxed{\mathbf{exp(5) = ?}}$$



Probability Distribution of the 'exact theory' prediction for $\exp(5) = 148.4$

What Are the First Conclusions ?

- **We are learning that a theory can give not only the result but also its probability distribution - but in this particular case:**

What Are the First Conclusions ?

• We are learning that a theory can give not only the result but also its probability distribution - but in this particular case:

- The 'theory' is great! (down to an absurd!) a failure!!!

What Are the First Conclusions ?

• We are learning that a theory can give not only the result but also its probability distribution - but in this particular case:

- The 'theory' is great! (down to an absurd!) a failure!!!

- It will obviously have nearly no predictive power [why?]

What Are the First Conclusions ?

- **We are learning that a theory can give not only the result but also its probability distribution - but in this particular case:**

  - **- The 'theory' is great! (down to an absurd!) a failure!!!**
  - **- It will obviously have nearly no predictive power [why?]**

- **Observe that uncertainty distributions are extremely broad!**

What Are the First Conclusions ?

• **We are learning that a theory can give not only the result but also its probability distribution - but in this particular case:**

- **The 'theory' is great! (down to an absurd!) a failure!!!**
- **It will obviously have nearly no predictive power [why?]**

• **Observe that uncertainty distributions are extremely broad!**

• **More precisely: With the parameter values of the order of 10^0 the distribution widths are of the order of 10^3 (3 orders)**

What Are the First Conclusions ?

• We are learning that a theory can give not only the result but also its probability distribution - but in this particular case:

- The 'theory' is great! (down to an absurd!) a failure!!!
- It will obviously have nearly no predictive power [why?]

• Observe that uncertainty distributions are extremely <u>broad</u>!

• More precisely: With the parameter values of the order of 10^0 the distribution widths are of the order of <u>10^3 (3 orders)</u>

$$\boxed{\boxed{\textbf{What is the fundamental origin of the `theory' failure?}}}$$

The Inverse Problem Related to the Theory

- **Our 'theory' is dangerously near the parametric correlations**

$$\boxed{\mathbf{a = 0,\ b = 0,\ c = 1,\ d = 1 \quad \longrightarrow \quad a = f(b)\ \ and\ \ c = g(d)}}$$

The Inverse Problem Related to the Theory

- **Our 'theory' is dangerously near the parametric correlations**

$$\boxed{\mathbf{a = 0,\; b = 0,\; c = 1,\; d = 1 \quad \longrightarrow \quad a = f(b) \;\; and \;\; c = g(d)}}$$

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

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

The Inverse Problem Related to the Theory

- **Our 'theory' is dangerously near the parametric correlations**

$$\boxed{a = 0,\ b = 0,\ c = 1,\ d = 1 \quad \longrightarrow \quad a = f(b) \ \text{ and } \ c = g(d)}$$

- **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 Singular-Value Decomposition written down explicitly**

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

The Inverse Problem Related to the Theory

- Our 'theory' is dangerously near the parametric correlations

$$\boxed{a = 0,\ b = 0,\ c = 1,\ d = 1 \quad \longrightarrow \quad a = f(b) \ \text{ and } \ c = g(d)}$$

- 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 Singular-Value Decomposition written down explicitly

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

- When there are correlations $[p_i = f(p_j)]$ at least one $d \to 0$

$$(p - p_0) = \underbrace{\left[(J\, J^T)^{-1}\, J^T\right]}_{d \to 0 \ \text{singularity}} b; \quad \text{p-parameters;} \quad \text{b-data}$$

Observations and Partial Conclusions

- **Since d $\longrightarrow$ 0: $\Rightarrow$ the correlations between d and p disappear**

Observations and Partial Conclusions

- **Since d $\longrightarrow$ 0: $\Rightarrow$ the correlations between d and p disappear**
- **Consequently - big variations of p may cause no effect on d**

Observations and Partial Conclusions

- **Since d $\longrightarrow$ 0: $\Rightarrow$ the correlations between d and p disappear**
- **Consequently - big variations of p may cause no effect on d**
- **Big modifications of parameters get mutually compensated**

Observations and Partial Conclusions

- **Since d $\longrightarrow$ 0: $\Rightarrow$ the correlations between d and p disappear**
- **Consequently - big variations of p may cause no effect on d**
- **Big modifications of parameters get mutually compensated**
- **As the result we obtain very broad parametric distributions**

Observations and Partial Conclusions

- **Since d $\longrightarrow$ 0: $\Rightarrow$ the correlations between d and p disappear**
- **Consequently - big variations of p may cause no effect on d**
- **Big modifications of parameters get mutually compensated**
- **As the result we obtain very broad parametric distributions**

In other words : 'Nothing Depends on Nothing'

Observations and Partial Conclusions

- **Since d $\longrightarrow$ 0: $\Rightarrow$ the correlations between d and p disappear**
- **Consequently - big variations of p may cause no effect on d**
- **Big modifications of parameters get mutually compensated**
- **As the result we obtain very broad parametric distributions**

$$\boxed{\boxed{\textbf{In other words : `Nothing Depends on Nothing'}}}$$

$$\boxed{\boxed{\textbf{At the end : Predictive Power Disappears}}}$$

Truncated Singular Value Decomposition (TSVD)

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

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

Truncated Singular Value Decomposition (TSVD)

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

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

- **Using Singular-Value Decomposition written down explicitly**

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

Truncated Singular Value Decomposition (TSVD)

- 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 Singular-Value Decomposition written down explicitly

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

- But when singular values $d \longrightarrow 0 \Rightarrow$ they do not contribute on the left-hand side, and the <u>effective size</u> of J gets smaller!

Truncated Singular Value Decomposition (TSVD)

- **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 Singular-Value Decomposition written down explicitly**

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

- **But when singular values $d \longrightarrow 0 \Rightarrow$ they do not contribute on the left-hand side, and the <u>effective size</u> of J gets smaller!**

- **Paradoxically, the divergence through the zero (singular!) values is a false conclusion since <u>from the beginning</u> they do not contribute! Ignoring them is called 'truncation' (TSVD).**

Truncated Singular Value Decomposition (TSVD)

- **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 Singular-Value Decomposition written down explicitly**

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

- **But when singular values d $\longrightarrow 0 \Rightarrow$ they do not contribute on the left-hand side, and the effective size of J gets smaller!**

- **Paradoxically, the divergence through the zero (singular!) values is a false conclusion since from the beginning they do not contribute! Ignoring them is called 'truncation' (TSVD).**

- **The real problem: How about small but $\neq 0$ singular values?**

- **The central point of considerations: Singular Values $\mathbf{d_i \longrightarrow \frac{1}{d_i}}$**

- **The central point of considerations: Singular Values $d_i \longrightarrow \frac{1}{d_i}$**

$$(J^T J)^{-1} = V\left[D^{-1}\right]^2 V^T; \quad D^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1},\ \frac{1}{d_2},\ \ldots\ \frac{1}{d_N}\right\}$$

Hidden Charms of the TSVD Method

- **The central point of considerations: Singular Values $\mathsf{d_i} \longrightarrow \frac{1}{\mathsf{d_i}}$**

$$(J^T J)^{-1} = V\big[D^{-1}\big]^2 V^T; \quad D^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1},\ \frac{1}{d_2},\ \dots\ \frac{1}{d_N}\right\}$$

- **We order the singular values (decreasing sequence): $\mathsf{d_N} \approx 0$**

Hidden Charms of the TSVD Method

- **The central point of considerations: Singular Values** $\mathbf{d_i \longrightarrow \frac{1}{d_i}}$

$$(J^T J)^{-1} = V\left[D^{-1}\right]^2 V^T; \quad D^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1},\ \frac{1}{d_2},\ \dots\ \frac{1}{d_N}\right\}$$

- **We order the singular values (decreasing sequence):** $\mathbf{d_N \approx 0}$

- **If** $\mathbf{d_N = 0}$ **strictly, it does not contribute to the sum in** $\mathbf{J_{ik}}$

- **The central point of considerations: Singular Values $d_i \longrightarrow \frac{1}{d_i}$**

$$(J^T J)^{-1} = V \left[D^{-1}\right]^2 V^T; \quad D^{-1} \equiv \text{diag.} \left\{\frac{1}{d_1}, \ \frac{1}{d_2}, \ \ldots \ \frac{1}{d_N}\right\}$$

- **We order the singular values (decreasing sequence): $d_N \approx 0$**

- **If $d_N = 0$ strictly, it does not contribute to the sum in J_{ik}**

$$J_{ik} = \sum_{\ell=1}^{r} U_{i\ell}\, d_\ell\, V^T_{\ell k} \quad \text{no term with } d_N = 0$$

Hidden Charms of the TSVD Method

- The central point of considerations: Singular Values $d_i \longrightarrow \frac{1}{d_i}$

$$(J^T J)^{-1} = V\left[D^{-1}\right]^2 V^T; \quad D^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1}, \frac{1}{d_2}, \ldots \frac{1}{d_N}\right\}$$

- We order the singular values (decreasing sequence): $d_N \approx 0$

- If $d_N = 0$ strictly, it does not contribute to the sum in J_{ik}

$$J_{ik} = \sum_{\ell=1}^{r} U_{i\ell}\, d_\ell\, V^T_{\ell k} \quad \text{no term with } d_N = 0$$

- There is no divergence in the solution for the parameters

$$(p - p_0) = \underbrace{\left[(J\, J^T)^{-1}\, J^T\right]}_{\text{NO singularity}} b; \quad \underline{\underline{\text{Problem is well posed}}}$$

Hidden Charms... But What if $d \neq 0$, Yet Small?

- **This is the heart of the problem:** $\mathbf{d_1 > d_2 > \rightarrow \ldots d_N \rightarrow 0}$

$$\boxed{(\mathbf{J}^T\mathbf{J})^{-1} = \mathbf{V}\left[\mathbf{D}^{-1}\right]^2\mathbf{V}^T; \quad \mathbf{D}^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1},\ \frac{1}{d_2},\ \ldots\ \frac{1}{d_N}\right\}}$$

- **This is the heart of the problem:** $\mathbf{d_1 > d_2 > \rightarrow \ldots d_N \rightarrow 0}$

$$\boxed{(\mathbf{J}^T\mathbf{J})^{-1} = \mathbf{V}\big[\mathbf{D}^{-1}\big]^2\mathbf{V}^T; \quad \mathbf{D}^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1},\ \frac{1}{d_2},\ \ldots\ \frac{1}{d_N}\right\}}$$

- **If we remove small but non-zero values, <u>we introduce errors</u>!**

- **This is the heart of the problem:** $\mathbf{d_1 > d_2 > \rightarrow \ldots d_N \rightarrow 0}$

$$\boxed{(\mathbf{J}^T\mathbf{J})^{-1} = \mathbf{V}\big[\mathbf{D}^{-1}\big]^2\mathbf{V}^T; \quad \mathbf{D}^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1},\ \frac{1}{d_2},\ \ldots\ \frac{1}{d_N}\right\}}$$

- **If we remove small but non-zero values, <u>we introduce errors</u>!**
- **But if we do not remove - <u>we have 'an almost' divergence</u>!**

Hidden Charms... But What if $d \neq 0$, Yet Small?

- **This is the heart of the problem:** $\mathbf{d_1 > d_2 > \rightarrow \; \ldots d_N \rightarrow 0}$

$$\boxed{(J^T J)^{-1} = V\left[D^{-1}\right]^2 V^T; \quad D^{-1} \equiv \text{diag.}\left\{\frac{1}{d_1},\ \frac{1}{d_2},\ \ldots\ \frac{1}{d_N}\right\}}$$

- **If we remove small but non-zero values, we introduce errors!**

- **But if we do not remove - we have 'an almost' divergence!**

- **We arrive at conflicting factors and thus a problem to solve**

Hidden Charms... But What if $d \neq 0$, Yet Small?

- **This is the heart of the problem:** $\mathbf{d_1 > d_2 > \to \ldots d_N \to 0}$

$$\boxed{(\mathbf{J}^T\mathbf{J})^{-1} = \mathbf{V}\big[\mathbf{D}^{-1}\big]^2\mathbf{V}^T; \quad \mathbf{D}^{-1} \equiv \text{diag.}\Big\{\frac{1}{d_1},\ \frac{1}{d_2},\ \ldots\ \frac{1}{d_N}\Big\}}$$

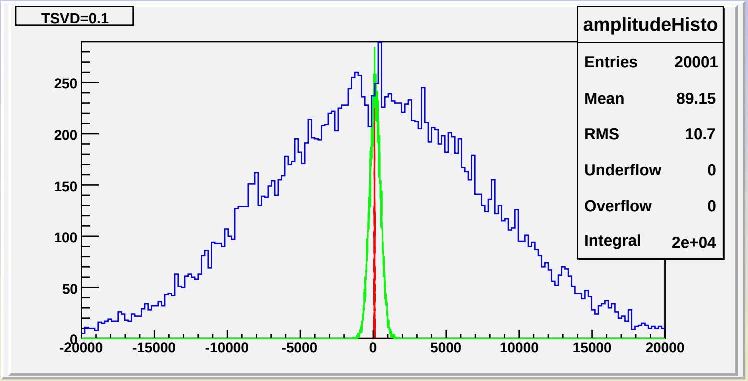
- **If we remove small but non-zero values, we introduce errors!**

- **But if we do not remove - we have 'an almost' divergence!**

- **We arrive at conflicting factors and thus a problem to solve**

- **Before we start looking for a compromise $\to$ an illustration**

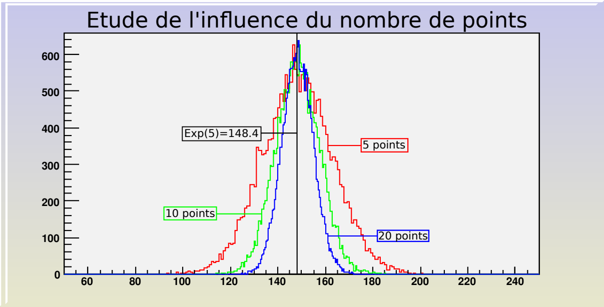
- **TSVD: $\Rightarrow$ efficiently counteracts the problem of instability**



Blue: no TSVD truncation, Green = cut off = 0.01, Red = cut off 0.1

Sampling: Increasing the Number of Experimental Points

- **Divergence depend on the 'Hamiltonian' <u>and</u> on data points**



Increasing the no. of data points increases the constraint on the model and as a consequence - stabilises the final solution

The Following Messages

The Following Messages are intended

The Following Messages
are intended
for Mature Audiences

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**

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

where $C^{m'I',n'L'v'J'}_{mI,nLvJ,Q}$ 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**

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

where $C^{m'I',n'L'v'J'}_{mI,nLvJ,Q}$ 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

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.

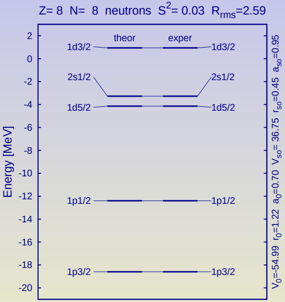
V_o^c	r_o^c	a_o^c	λ	r_o^{so}	a^{so}
-39.520	1.371	0.694	26.133	1.255	0.500

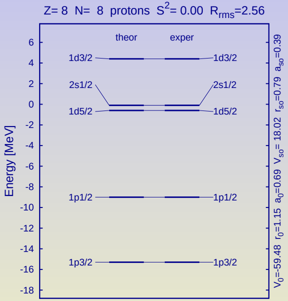
• We can treat $\{e_\mu\}$ 'as experimental'; by trying to reproduce them through fitting we know an exact solution!

• 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 $\{p\}$

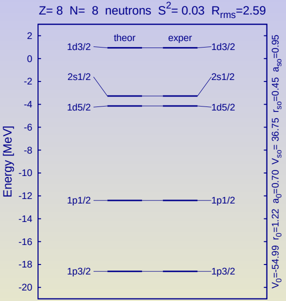
- We fit the single-particle experimental levels in ^{16}O using Woods-Saxon potential (six parameters for protons and neutrons each)

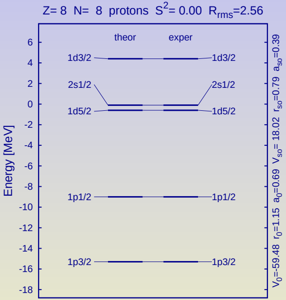




'Chi-by-the-eye' Results May Look Attractive...

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

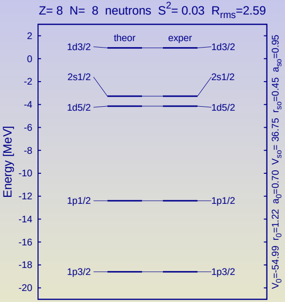


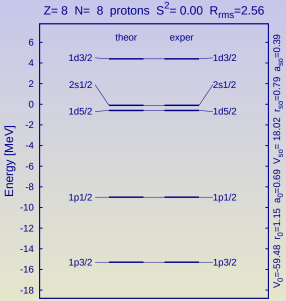


- This result may look surprising: the quality of the fit is such that *graphical illustrations are insufficient to show it* !!!

'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

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!

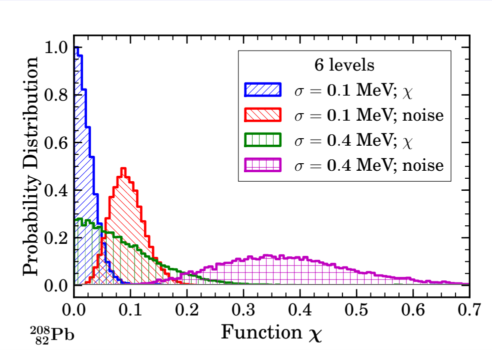
→ The standard Woods-Saxon Hamiltonian has been used:

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

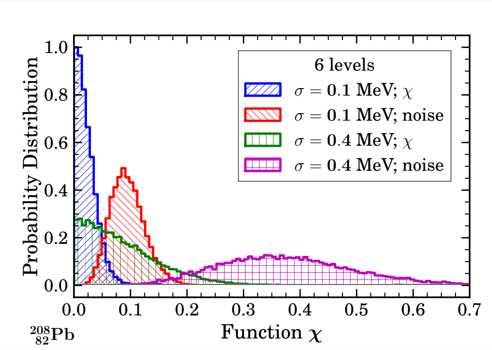
→ Couple of questions may come to one's mind...:

- *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'

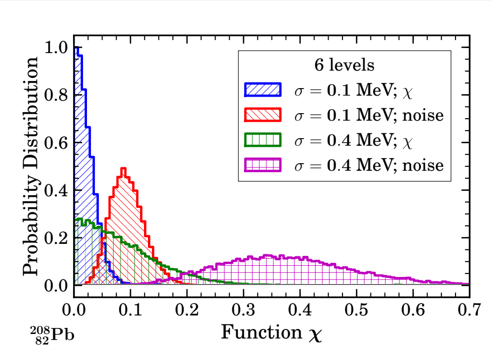


Unprecedented Precision ... Known as ‘Overfitting’



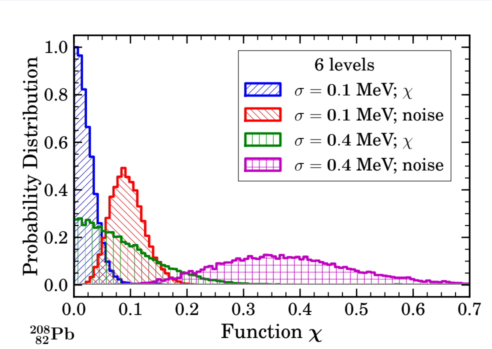
- We introduce the Gaussian noise into the experimental-level input,

Unprecedented Precision ... Known as ‘Overfitting’

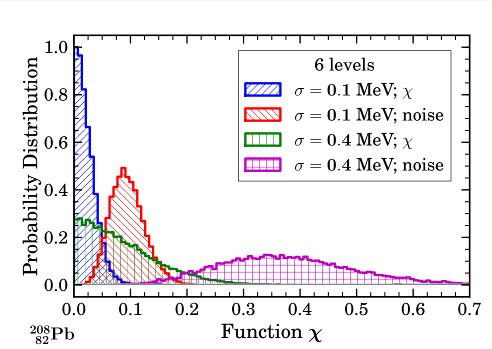


- We introduce the Gaussian noise into the experimental-level input, repeat the χ^2-fit

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.



- 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 <u>de</u>creasing 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 <u>de</u>creasing 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 <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...

'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 <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...

• We improve the model by reducing the number of parameters

Part V

Nuclear Hamiltonians: Sampling vs. Microscopic Structure

Decrease Parametric Freedom: 'Be More Microscopic'

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

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

• 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

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

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

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

• 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

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

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

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

• 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

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

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

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

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

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

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:

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

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 $\mathbf{i^{th}}$ iteration the spectrum is $\{\mathbf{e_n^i}\}$ and at $\mathbf{i+1^{st}} \rightarrow \{\mathbf{e_n^{i+1}}\}$, we stop iterating when**

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

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:

$$\mathbf{H}(\rho)\psi_n = e_n\,\psi_n \;\rightarrow\; \rho = \sum \psi^*\psi \;\rightarrow\; \mathbf{H}(\rho)\psi_n = e_n\,\psi_n \ldots$$

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 i^{th} iteration the spectrum is $\{e_n^i\}$ and at $i+1^{st} \rightarrow \{e_n^{i+1}\}$, we stop iterating when**

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

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:

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

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 $\{\mathbf{e}_{\mathbf{n}}^{\mathbf{i}}\}$ and at $\mathbf{i+1}^{\mathbf{st}} \to \{\mathbf{e}_{\mathbf{n}}^{\mathbf{i+1}}\}$, we stop iterating when**

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

Self-consistent Formulation: Minimum Coupling

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

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

• 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

$$\boxed{\lambda_{\pi\pi} = \lambda_{\pi\nu} = \lambda_{\nu\pi} = \lambda_{\nu\nu} \overset{\text{df}}{=} \lambda}$$

• 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

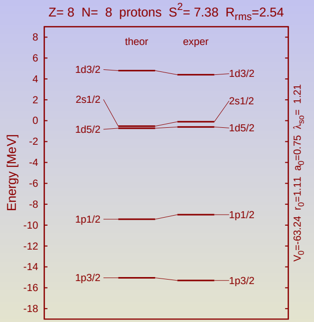
- For the first tests we apply what we call a minimum coupling hypothesis

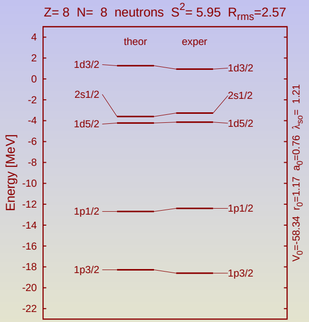
$$\boxed{\boldsymbol{\lambda_{\pi\pi} = \lambda_{\pi\nu} = \lambda_{\nu\pi} = \lambda_{\nu\nu} \overset{\mathrm{df}}{=} \lambda}}$$

- 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, $\mathbf{r_o^{so}}$ and $\mathbf{a^{so}}$) to one!**

Comparing with Experimental Results: Example ^{16}O

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

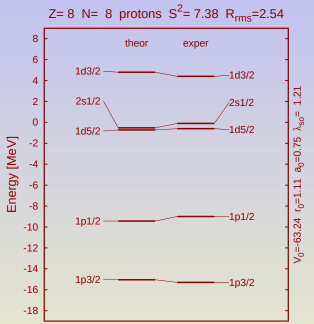


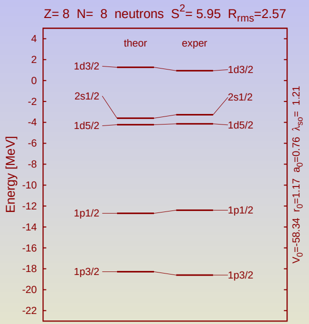


These results corresponds to just one λ-parameter fit instead of 6.

Comparing with Experimental Results: Example ^{16}O

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



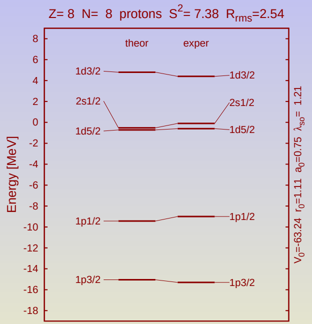


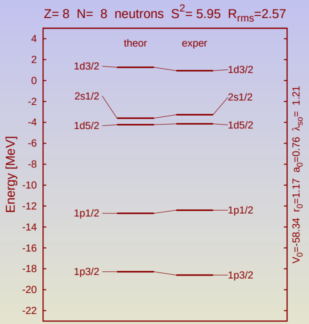
These results corresponds to just <u>*one λ-parameter fit instead of 6.*</u>

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



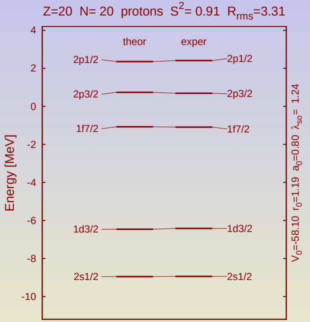


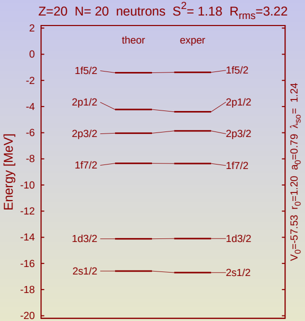
These results corresponds to just one λ-parameter fit instead of 6.

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

The maximum error $\sim$ 200 keV

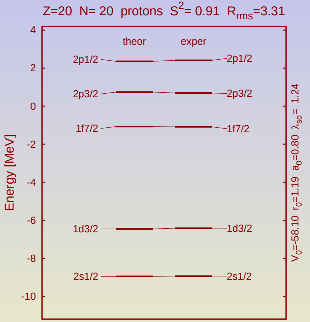
Similarly the single particle proton and neutron states in ^{40}Ca

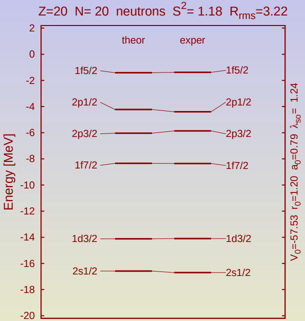




Comparing with Experimental Results: Example ^{40}Ca

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

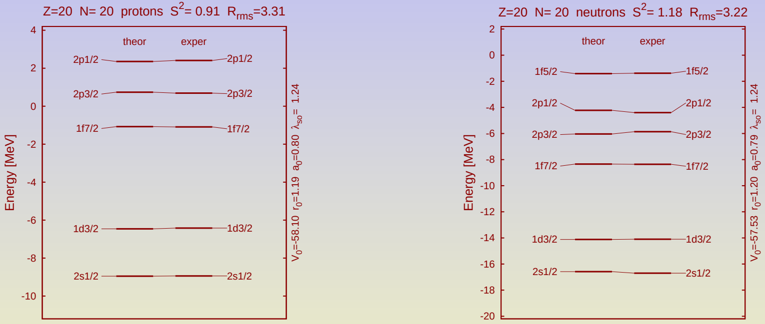




This result corresponds to just one (λ) parameter fit instead of 6; similar results hold for nuclei up to ^{208}Pb.

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:**

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 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

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:**

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

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
   - 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 we have presented some simple semi-quantitative illustrations