Nuclear Mean Field Hamiltonians and Their Spectroscopic Predictive Power

Bartłomiej Szpak

Institute of Nuclear Physics, PAN
Kraków, Poland

25th of May, 2010

In collaboration with:
J. Dudek, K. Rybak, M.-G. Porquet, H. Molique, B. Fornal

The aims of a project

Fields of applied mathematics such as the methods of statistical inference and inverse problem theory provide powerful and valuable tools for estimating the predictive power of mathematical models. They usefulness is not fully explored in nuclear structure models.

We wish to use these methods to:

- Examine their possible applications in empirical nuclear structure models, especially insabilities of the modelling
- Obtain reliable estimates and confidence intervals for model parameters and observables
- Improve the predictive power of models rather then getting better *numerical* description of available data

The basics of least-squares methods (1)

Let us model a set of experimental data points y_i using function $f(x_i, \vec{\beta})$. The aim of inference is to determine the n parameters β_k from a set of m measurments y_i taking into account the uncertainties in parameters.

Assuming that the error in each measurment is normally distributed with zero mean and variance σ_i, and that the errors are statistically independent, the likelihood $p(\vec{y}|\vec{\beta})$ is:

$$p(\vec{y}|\vec{\beta}) \propto \prod_i \exp\left[-\frac{(y_i - f(x_i, \vec{\beta}))^2}{\sigma_i^2}\right]$$

and

$$\chi^2 = -2log\left[p(\vec{y}|\vec{\beta})\right] = \sum_i \frac{(y_i - f(x_i, \vec{\beta}))^2}{\sigma_i^2}$$

We define the Jacobian matrix: $\boxed{J_{ij} = \frac{\partial y_i}{\partial \beta_j}}$. The normal equation for β reads:

$$\boxed{\vec{\beta} = (\mathbf{J^T J})^{-1}\mathbf{J^T \vec{y}}}$$

The basics of least-squares methods (2)

The unbiased estimator for rms deviation of residuals $r_i = y_i - f(x_i, \hat{\vec{\beta}})$ is:

$$\boldsymbol{\Delta^2_{rms} = \hat{\sigma}^2 = \frac{\chi^2}{m - n}}$$

The distribution of Δ^2_{rms} follows a χ^2-square distribution with $\nu = m - n$ degrees of freedom.

A number of properties can be estimated with the knowledge of a Jacobian:

- The variance-covariance matrix for parameters $\vec{\beta}$ (assuming equal σ_i):

$$\mathbf{M}^\beta = \boldsymbol{\Delta}^2_{rms}(\mathbf{J}^T\mathbf{J})^{-1}$$

- Correlation matrix:

$$\rho_{ij} = \frac{\mathbf{M}^\beta_{ij}}{\sqrt{\mathbf{M}^\beta_{ii}\mathbf{M}^\beta_{jj}}}$$

- Confidence intervals (using Student t-distribution):

$$\beta_i \in (\hat{\beta}_i - t_{\alpha/2,\nu}\sqrt{\mathbf{M}^\beta_{ii}}, \hat{\beta}_i + t_{\alpha/2,\nu}\sqrt{\mathbf{M}^\beta_{ii}})$$

The illustrative toy model

As an illustrative model we are using a classical Woods-Saxon potential which parameters are fitted to a set of neutron single particle levels in ^{208}Pb. For simplicity of presentation I assume equal weights for all experimental levels. We set $a_0^{so} = 0.6$ fm.

$$
\begin{aligned}
H &= T + V_c + V_{so} \\
V_c &= \frac{V_0}{1 + \exp[(r - R_0)/a_0]} \\
V_{so} &= \frac{V_0^{so}}{r}\frac{d}{dr}\left\{\frac{V_0^{so}}{1 + \exp[(r - R_0^{so})/a_0^{so}]}\right\}
\end{aligned}
$$

We have:

$$
\begin{aligned}
\vec{\beta} &= [V_0, r_0, a_0, V_0^{so}, r_0^{so}] \\
x_i &= \{1h_{9/2}, 2f_{7/2}, 1i_{13/2}, 3p_{3/2}, 2f_{5/2}, 3p_{1/2}, 2g_{9/2}, \\
&\quad 1i_{11/2}, 1j_{15/2}, 3d_{5/2}, 4s_{1/2}, 2g_{7/2}, 3d_{3/2}\} \\
y_i &= \{-10.784, -10.300, -9.800, -8.270, -7.940, -7.370, \\
&\quad -3.940, -3.160, -2.517, -2.370, -1.900, -1.440, -1.400\}
\end{aligned}
$$

How much can we get using classical methods ?

As the model is non-linear in its parameters we start the minimization with randomly selected $\vec{\beta}$ and repeat the procedure many times to enhace the chance for obtaining all minima.

The two solutions are found (called non-compact and compact):

V_0	r_0	a_0	V_0^{so}	r_0^{so}	$\mathbf{\Delta}_{\mathbf{rms}}$
-42.025	1.320	0.694	24.781	1.231	**0.349**

V_0	r_0	a_0	V_0^{so}	r_0^{so}	$\mathbf{\Delta}_{\mathbf{rms}}$
-41.653	1.328	0.670	24.007	0.924	**0.425**

This is where many of analysis finish.
How much can we deduce from this pieces of information ?

Lack of information

We can conclude that:

- We fitted the parameters to 13 neutron levels in ^{208}Pb
- We do know nothing about the predictive power of our model.
- The parameters of central potential are "stable" as they are almost the same in two resulting parameterizations.
- The spin-orbit radius parameter can take two distinct values. The model with smaller Δ_{rms} is better so we are tempted to choose the first set of parameters.
- The performance of the model is quite good. The average disagreement between the theory and experiment is around 350keV.

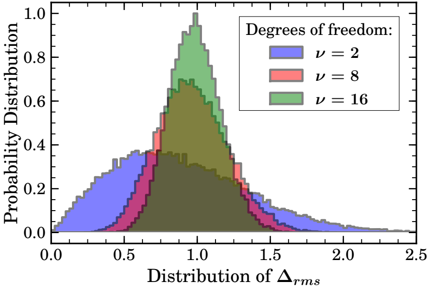
A chance for being too lucky

Suppose we use a set of 12 experimental levels out of 13 to perform the fit. We can get the set of 13 different values for Δ_{rms}:

$$\Delta_{rms} \in \{0.191, 0.289, 0.338, 0.342, 0.351, 0.351, 0.351, 0.359, 0.363, 0.363, 0.365, 0.373, 0.423\}$$

These are the random samples from a $\sqrt{\frac{\chi^2(x;\nu)}{\nu}}$ distribution.

Without the prior knowledge of physical/expected value of Δ_{rms} we shall always remember that Δ^2_{rms} follows a $\chi^2(x;\nu)/\nu$ distribution. This means that Δ^2_{rms} is a poor estimator of $\hat{\sigma}^2$ - it can be biased and have large variance. Moreover there always exist a possibility of overfit.This problem is crucial for the reliable estimation of confidence intervals as $M^\beta = \Delta^2_{rms}(J^T J)^{-1}$.



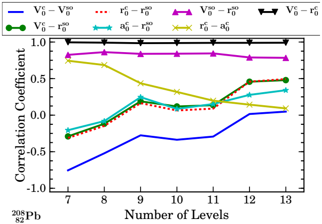
A gain of information with variance-covariance matrix

The variance-covariance matrix provides a basic information about the properties of a fit. Having this information we are able to calculate the standard errors, confidence intervals and correlation matrix for the parameters.

property	V_0	r_0	a_0	V_0^{so}	r_0^{so}
mean	-42.025	1.320	0.694	24.781	1.231
std. error	1.217	0.023	0.031	2.639	0.050
conf. interval	2.264	0.043	0.058	4.907	0.093

Estimate of a correlation matrix:

	V_0	r_0	a_0	V_0^{so}	r_0^{so}
V_0	1.0000	0.9900	0.0096	0.0482	0.4788
r_0		1.0000	0.0923	0.0632	0.4935
a_0			1.0000	0.3393	0.3407
V_0^{so}				1.0000	0.7835
r_0^{so}					1.0000



Making predictions with classical analysis

The information about the correlation among the parameters is essential for estimation of the predictive power of a model. With it, **assuming the output of a model are linearly related to the perturbations in the input**, the variance-covariance matrix for the output is given by:

$$\boxed{\mathbf{M^y = S^T M^\beta S}}$$

where $S_{ij} = \frac{\partial y_i}{\partial \beta_j}$.

Alternatively, one can use Monte-Carlo methods to sample the parameter space. As we will see later, the full information about the correlation matrix is necessary to obtain reliable predictions.

Part 2

A deeper inspection of model's statistical properties using Monte-Carlo analysis

Algorithm of calculations

We can use a Monte-Carlo method to verify how the random noise in the input data influences the obtained parameters of a model.

The procedure of calculations is following:

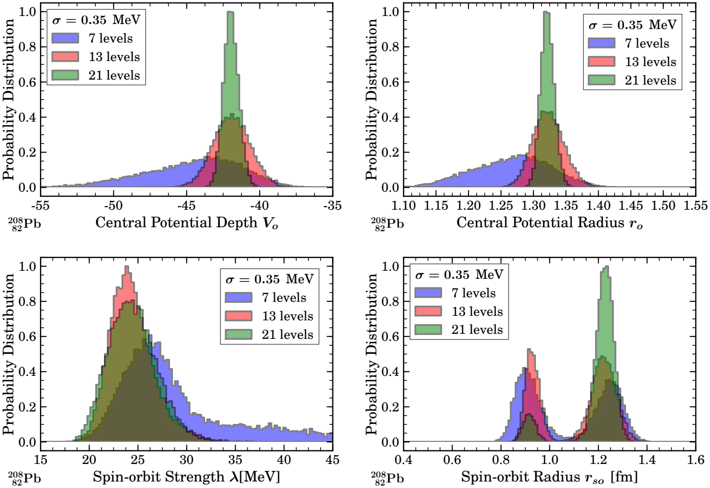
1. Fit the experimental data within the examined model.
2. Use the resampling techniques for inspection.
3. Add a Gaussian random noise with $\sigma = \Delta_{rms}$ to the "pseudoexperimental data" obtained from the fit. Generate many sets of pseudoexperimental data sets.
4. Fit the parameters to each item in the set.

Advantages over previous approach

While using Monte-Carlo analysis:

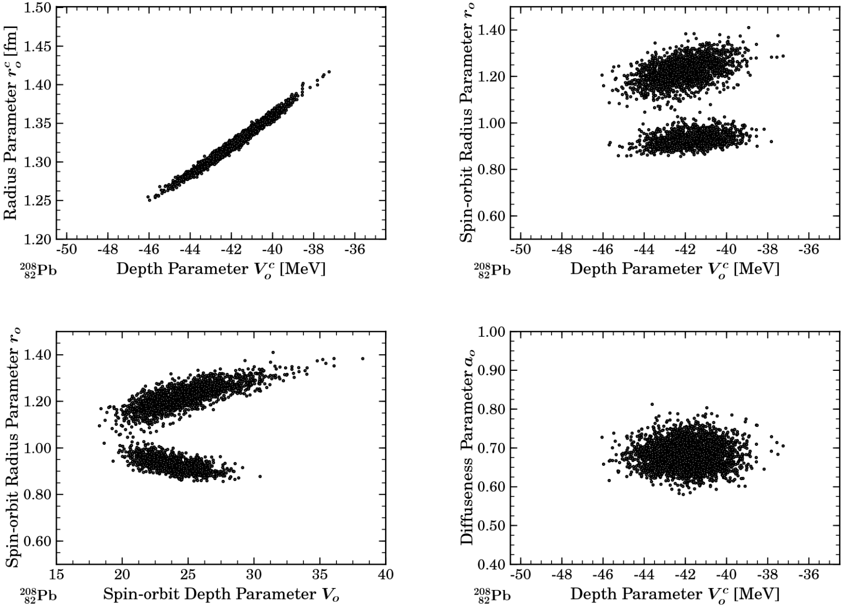
- We introduce no other assumptions about the properties of a model.
- We obtain the Probability Density Function (PDF) for parameters and observables.
- We can easily inspect the correlations among parameters and residuals.
- We can generate different kinds of pseudoexperimental data to determine which data can more- and which less-effectively constrain e.g. the confidence intervals for parameters.

Confidence intervals for parameters.



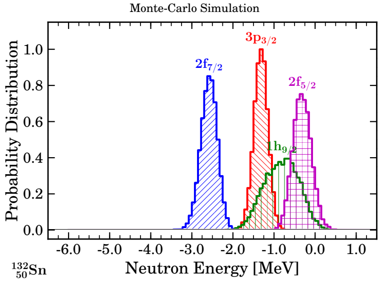
In case of well-posed problems the confidence intervals obtained in "classical calculations" are well reproduced by monte-carlo simulations. This is not true in the case of ill-posed problems.

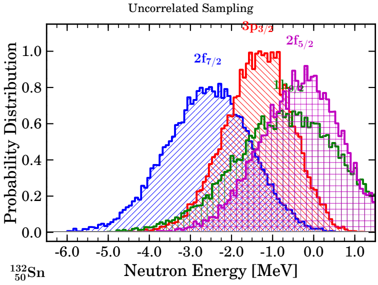
Correlations between parameters.

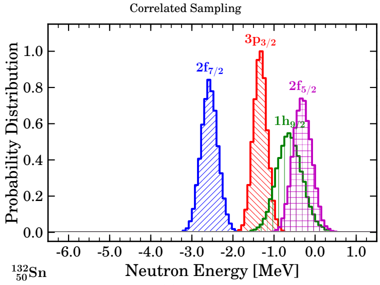


Estimating the predictive power (1)

The PDF distributions obtained from Monte-Carlo analysis and correlated sampling shows the great similarities. *Uncorrelated Sampling* shows the results obtained from classical analysis when only standard errors for the parameters are taken into account.



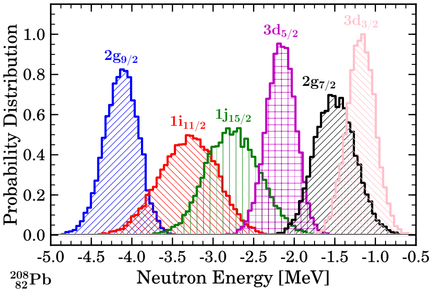




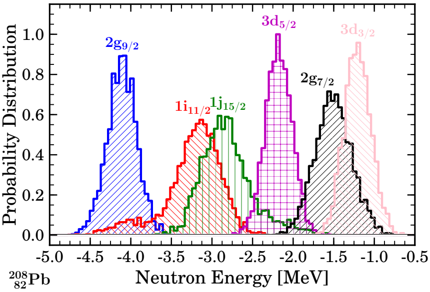
PDF for observables

We wish to learn about the performance of the model when increasing the data sample: example - increase the input from 13 to 21 levels ...

13 levels



21 levels

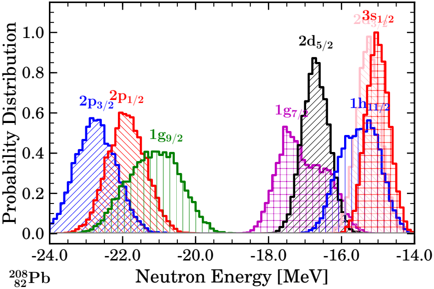


Comparison for levels lying above the Fermi levels.

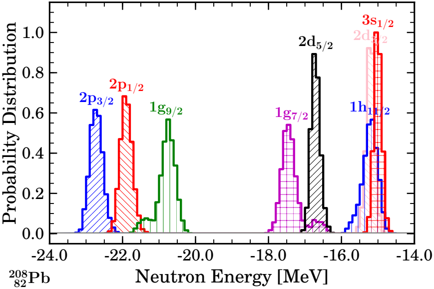
All shown levels are included in fit in both calculations.

PDF for observables

13 levels



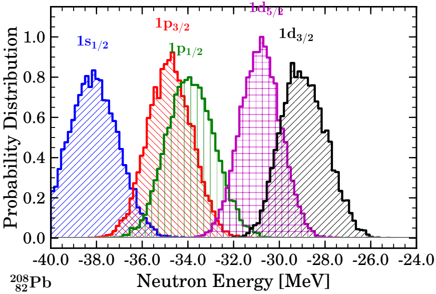
21 levels



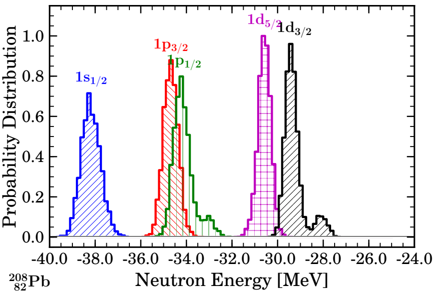
Left: Deep-lying states **not included** in the fit.
Right: Deep-lying states **included** in the fit.

Estimating the predictive power (2)

13 levels



21 levels



The constraining effect on the levels not included in the fit.

Conclusion

Conclusions:

- Without analysis of statistical significance we risk making **more speculations than predictions**.
- Possibility to optimize the models from predictive power point of view.
- The discrepiency between the **predicted** value and new experimental results can indicate interesting physics.
- The correlations cannot be neglected.

The things that were not discussed:

- The techniques to deal with the singularity - SVD, regularization
- Inclusion of prior information - bayesian approach
- The alternative techniques of resampling/bootstrap/bootstraping residuals
- Physical implications and further applications.