Spectral predictive power of the nuclear hamiltonians

Bartłomiej Szpak

Institute of Nuclear Physics, PAN
Kraków, Poland

6th of June, 2012

In collaboration with:
J. Dudek, B. Fornal, H. Molique

Motivation behind the project

Following the recent paper of M. Dutra *et. al* PRC 85, 035201 (2012):

This paper presents a detailed assessment of the ability of the **240 Skyrme interaction parameter sets** in the literature to satisfy a series of criteria derived from macroscopic properties of nuclear matter [. . .]

The objective is to identify those parameterizations which best satisfy the current understanding of the physics of nuclear matter over a wide range of applications.

Out of the 240 models, only 16 are shown to satisfy all these constraints. Additional, more microscopic, constraints [. . .] further reduce this number to 5, a very small group of recommended Skyrme parameterizations to be used in future applications of the Skyrme interaction of nuclear matter related observables

Motivation behind the project

Phenomenological fit without proper statistical analysis:

- do not provide error estimates of model parameters
- do not provide error estimates of model predictions
- cause poorly determined parameters to take random values
- can lead to overfit
- gives no method to compare results obtained with different models

Only full regression analysis can answer the question how reliable are our fits and predictions.

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.

Outline of a talk

1. Skyrme functional
2. Regression analysis
3. Origin of ill-posed problems
4. Monte-carlo methods
5. Regularization techniques
6. Conclusions and future work

Form of a Skyrme Functional

We consider Skyrme-Hartree-Fock problem with the spherically-symmetric Hamiltonian density without pairing, which is given by a sum of the kinetic- and potential-energy isoscalar and iso-vector terms:

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

where

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

The C_t^{ρ} "constant" is density-dependent and we have

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

Problem is described by 13 parameters in total:

$$C_t^{\rho}, C_t^{\rho\alpha}, C_t^{\Delta\rho}, C_t^{\tau}, C_t^{J}, C_t^{\nabla J}, \alpha$$

Form of a Skyrme Functional

The functional can be described by other sets of coupling constants. The other choices are:

- (t, x) constants
- Nuclear Matter constants

The parameters of a Skyrme functional are fitted to single particle energies and masses of doubly magic nuclides:

^{16}O, ^{40}Ca, ^{48}Ca, ^{56}Ni, ^{90}Zr, ^{132}Sn, ^{208}Pb

This gives 98 experimental data points in total.

Due to limited set of observables still a toy model.

Nuclear matter properties

The expected values of nuclear matter properties reads:

$$
\begin{aligned}
\rho_c &= 0.16\,\mathrm{fm}^{-3} && \text{saturation point} \\
\frac{E^{NM}}{A} &= -16\,\mathrm{MeV} && \text{total energy per nucleon at equalibrium} \\
K^{NM} &= 220 \pm 10\,\mathrm{MeV} && \text{nuclear matter incompressibility modulus} \\
L_{sym}^{NM} &= 80 \pm 30\,\mathrm{MeV} && \text{density dependence of the symmetry energy} \\
a_{sym}^{NM} &= 28 - 36\,\mathrm{MeV} && \text{symmetry energy at saturation density} \\
M_s^{*-1} &= 1 && \text{isoscalar effective mass} \\
M_v^{*-1} &= 1 && \text{isovector effective mass}
\end{aligned}
$$

The EDF parameters are taken from the typical Skyrme parametrizations $[\mathrm{MeVfm}^5]$:

$$
C_0^{\Delta\rho} = -76.99, \quad C_0^{\nabla J} = -92.25, \quad C_1^{\Delta\rho} = 15.657, \quad C_1^{\nabla J} = -30.75
$$

Forward and inverse problems

We consider:

- Physical model: J
- Model parameters: $\vec{\beta}$
- Data points: $\vec{y}$

Forward problem: having specified J and $\vec{\beta}$ calculate $\vec{y}$.

Inverse problem: Knowing $\vec{y}$ and assuming J obtain the most probable $\vec{\beta}$.

Features:

- non-uniqueness
- ill-posed
- extra assumptions required

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:

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

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:

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

The distribution of $\boldsymbol{\Delta^2_{\mathrm{rms}}}$ follows a $\boldsymbol{\chi^2}$-square distribution with $\boldsymbol{\nu = m - n}$ degrees of freedom.

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

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

Correlation matrix:

$$\boxed{\boldsymbol{\rho_{ij} = \frac{M^{\beta}_{ij}}{\sqrt{M^{\beta}_{ii} M^{\beta}_{jj}}}}}$$

In the least-squares problems a solution is a n-dimensional Gaussian parametrized by a $\vec{\beta}$ and M^{β}. Solution space can be sampled with Monte-Carlo methods to make predictions.

Classical Regression Analysis

The only, recent parametrization of EDF coming with error estimation:
M. Kortelainen, T. Lesinski, W. Nazarewicz et. al. PRC 82 024313 (2010)

Fit to 28 spherical nuclei (binding energies, charge radii) and 44 deformed (b. e.).

Parameter M_v^{*-1} set to 1.249, no tensor.

	ρ_c	$\frac{E^{NM}}{A}$	K^{NM}	a_{sym}^{NM}	L_{sym}^{NM}	M_s^{*-1}	$C_0^{\Delta\rho}$	$C_1^{\Delta\rho}$	$C_0^{\nabla J}$	$C_1^{\nabla J}$
mean	0.151	-16.06	337.9	32.45	70.2	0.96	-49.1	33.5	-78.4	63.9
error	0.001	0.04	26.8	2.75	45.0	0.1	4.78	27.2	5.1	30.9

Parametrization obtained in the current work:

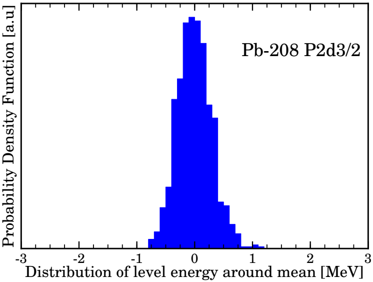
	ρ_c	$\frac{E^{NM}}{A}$	K^{NM}	a_{sym}^{NM}	L_{sym}^{NM}	M_s^{*-1}	$C_0^{\Delta\rho}$	$C_1^{\Delta\rho}$	$C_0^{\nabla J}$	$C_1^{\nabla J}$
mean	0.144	-15.80	344.6	25.6	-85.8	0.92	-44.4	25.6	-77.2	-29.9
error	0.01	0.26	59.8	0.93	33.9	0.07	9.48	11.6	11.1	28.4

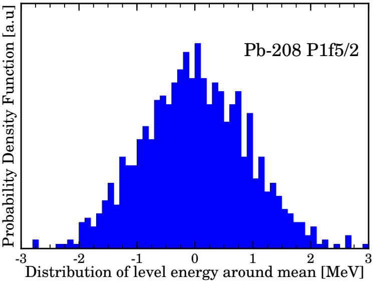
Correlation matrix obtained in present work:

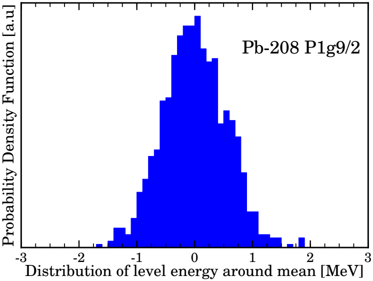
	ρ_c	$\frac{E^{NM}}{A}$	K^{NM}	a_{sym}^{NM}	L_{sym}^{NM}	M_s^{*-1}	$C_0^{\Delta\rho}$	$C_1^{\Delta\rho}$	$C_0^{\nabla J}$	$C_1^{\nabla J}$
ρ_c	1.00	-0.93	0.41	0.74	-0.11	-0.74	0.79	0.10	0.88	0.49
$\frac{E^{NM}}{A}$		1.00	-0.66	-0.84	0.31	0.79	-0.87	-0.02	-0.92	-0.38
K^{NM}			1.00	0.60	-0.54	-0.65	0.81	-0.10	0.68	0.12
a_{sym}^{NM}				1.00	-0.05	-0.62	0.70	0.36	0.71	0.39
L_{sym}^{NM}					1.00	0.28	-0.36	0.72	-0.30	0.32
M_s^{*-1}						1.00	-0.93	-0.13	-0.88	-0.41
$C_0^{\Delta\rho}$							1.00	0.05	0.94	0.42
$C_1^{\Delta\rho}$								1.00	0.05	0.53
$C_0^{\nabla J}$									1.00	0.42
$C_1^{\nabla J}$										1.00

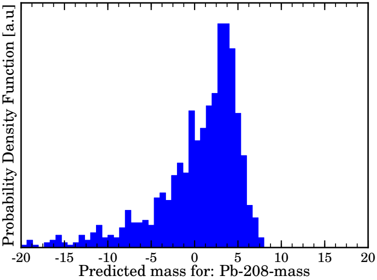
Statistically significant correlations between large number of parameters.

Classical Regression Analysis - Predictions

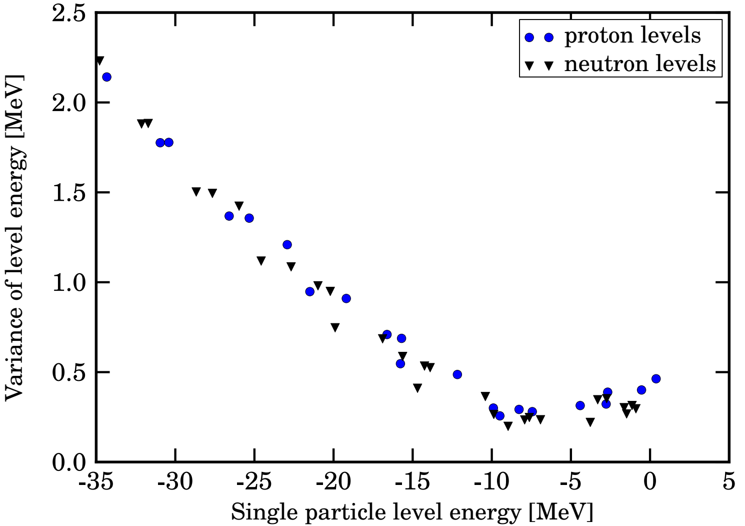








Classical Regression Analysis - Predictions



Monte-carlo methods

Utilize a concept of synthetic data sets:

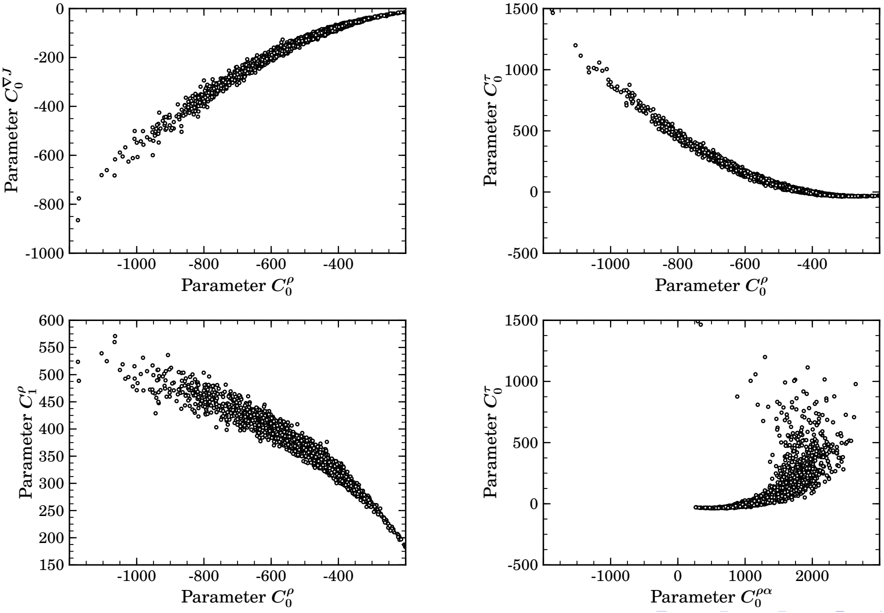
1. generate data using known $\vec{\beta}$
2. add random noise with known σ
3. use the data in simulation
4. repeat the process as many times as possible

Number of important advantages:

- known unbiased solution
- test minimization procedures
- test influence of weights, role of different observables
- test over overfit

Disadvantage: time consuming.

Monte-carlo methods



The normal equation for β reads:

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

Let us consider two Hilbert matrices which matrix elements are given by:

$$H_{ij} = \frac{1}{i + j - 1} \quad H'_{ij} = H_{ij} \times X$$

where X is a uniformly distributed random number from range $(0.999, 1.001)$. The H^{-1}/H'^{-1} can look like:

$$H^{-1}/H'^{-1} = \begin{pmatrix} 6 & 20 & -1174 & -38 & -25 \\ 16 & 48 & -535 & -62 & -38 \\ 64 & 289 & -221 & -96 & -67 \\ -115 & -123 & -137 & -155 & -173 \\ -44 & -60 & -102 & -280 & 561 \end{pmatrix}$$

The Conditional Number of H matrix is of order 10^5.

Regularized solutions of ill-posed problems (1)

Any matrix can be factorized as:

$$J = USV^T$$

where U, V are orthogonal matrixes spanning the space of data points and model parameters.
Matrix S is diagonal with entries $\lambda_i = S_{ii}$ called singular values.

An idea is to replace:

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

by:

$$\vec{\beta} = VS^{-1}U^T \vec{y}$$

New model parameters:

$$\vec{z} = SV^T \vec{\beta}$$

To each element of $\vec{z}$ we assign singular value.

Regularized solutions of ill-posed problems (2)

Concept of regularization:
Fit only parameters that are well determined by available experimental data. Keep other constant.

SVD: Importance of parameters characterized by its singular value.

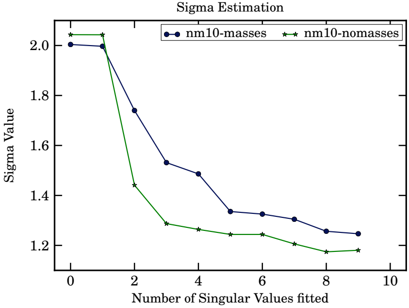
We fit only first p parameters:

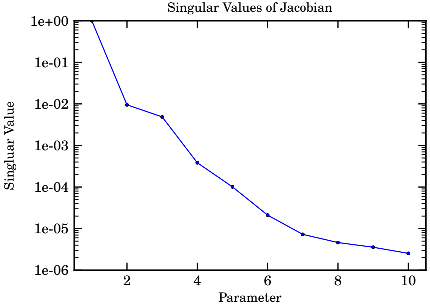
$$\vec{\beta} = V_p S_p^{-1} U_p^T \vec{y} = \sum_{i=1}^{p} \frac{U_{\cdot,i}^T \vec{y}}{\lambda_i} V_{\cdot,i}$$

$$M^\beta = \Delta^2 \sum_{i=1}^{p} \frac{V_{\cdot,i} V_{\cdot,i}^T}{\lambda_i^2}$$

Disadvantage: the result of estimation may depend on a prior choice of parameters. Ambiguity in choosing p.

Regularized solutions of ill-posed problems

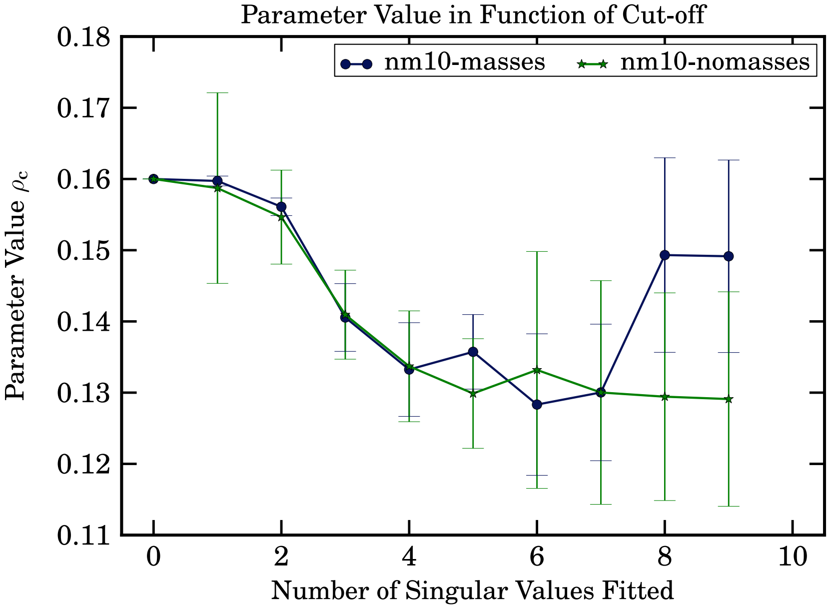




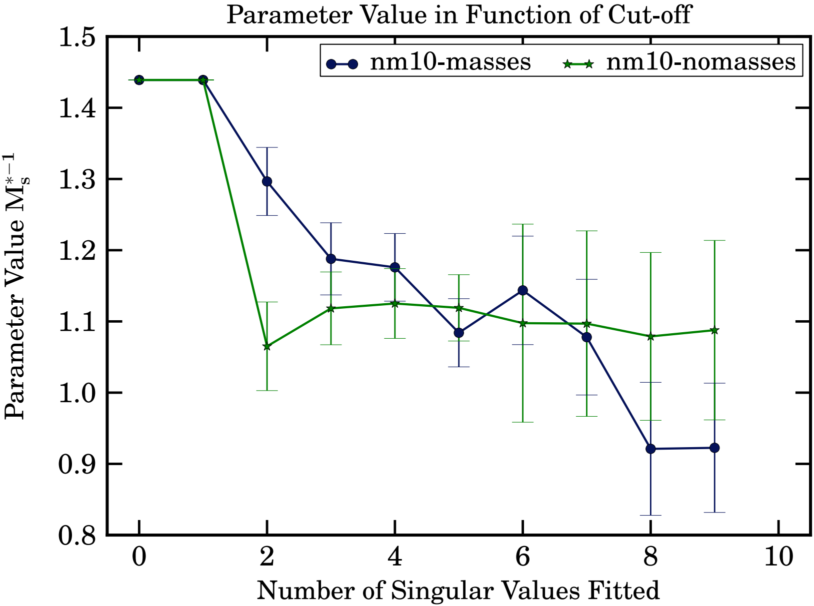
SVD provides a natural scheme to classify importance of parameters.

	ρ_c	$\frac{E^{NM}}{A}$	K^{NM}	a^{NM}_{sym}	L^{NM}_{sym}	M_s^{*-1}	$C_0^{\Delta\rho}$	$C_1^{\Delta\rho}$	$C_0^{\nabla J}$	$C_1^{\nabla J}$
z0	**1.000**	0.043	0.000	0.001	-0.000	-0.042	-0.001	-0.000	0.000	0.000
z1	0.031	0.250	-0.001	0.028	-0.001	**1.000**	0.013	-0.000	-0.000	-0.000
z2	-0.054	**1.000**	0.001	0.137	-0.010	-0.252	-0.001	-0.003	0.009	0.000
z3	-0.005	0.135	-0.006	**-1.000**	0.082	-0.005	-0.043	0.020	0.084	0.001
z4	0.000	-0.017	0.078	0.095	-0.060	0.014	**-1.000**	-0.020	**0.718**	0.011
z5	0.001	-0.017	**-0.414**	0.069	0.156	-0.006	**0.681**	0.067	**1.000**	-0.132
z6	-0.000	-0.002	**1.000**	0.031	0.222	-0.001	0.151	**0.705**	0.145	**-0.617**
z7	-0.000	0.004	**-0.798**	0.047	**0.389**	0.002	**-0.301**	**1.000**	-0.276	-0.148
z8	0.000	-0.003	**0.332**	0.081	**0.932**	-0.000	0.046	0.083	0.081	**1.000**
z9	-0.000	0.001	-0.155	0.062	**1.000**	0.001	-0.176	**-0.731**	-0.162	**-0.804**

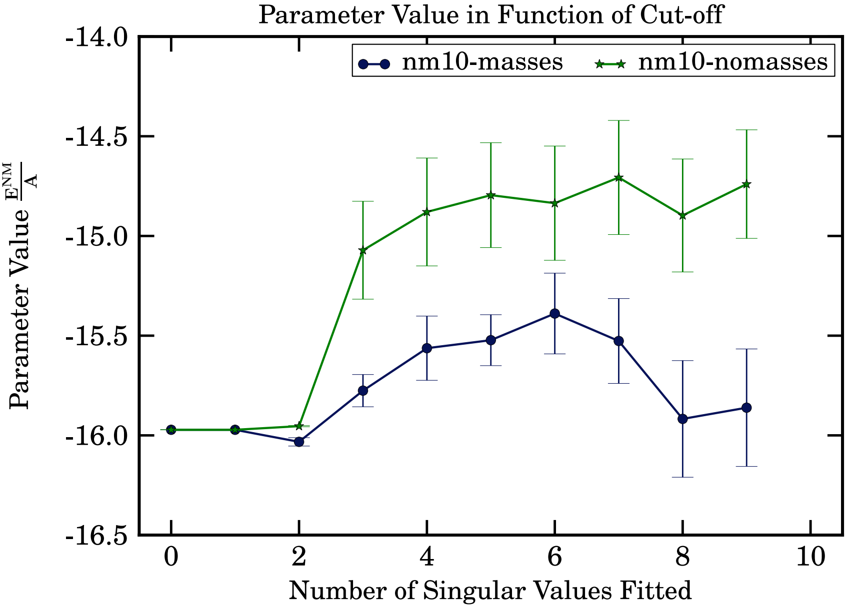
Regularized solutions



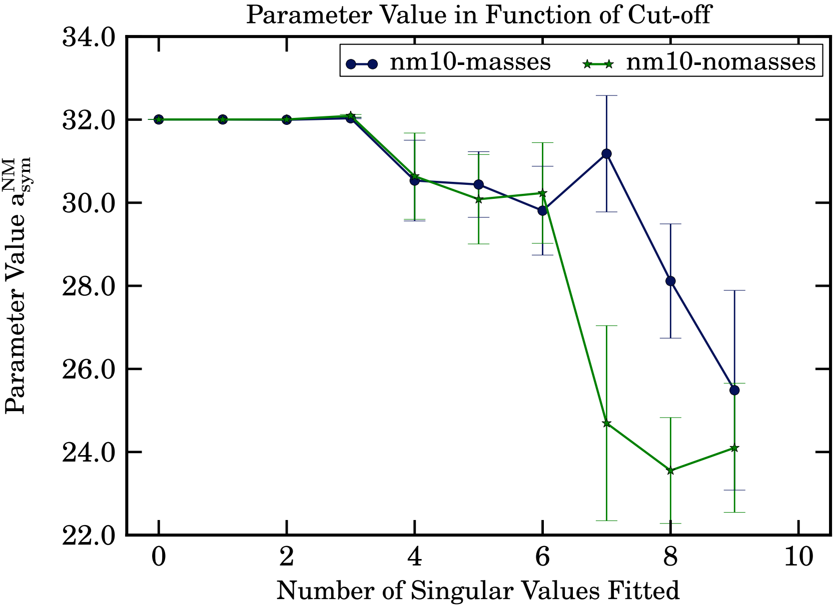
Regularized solutions



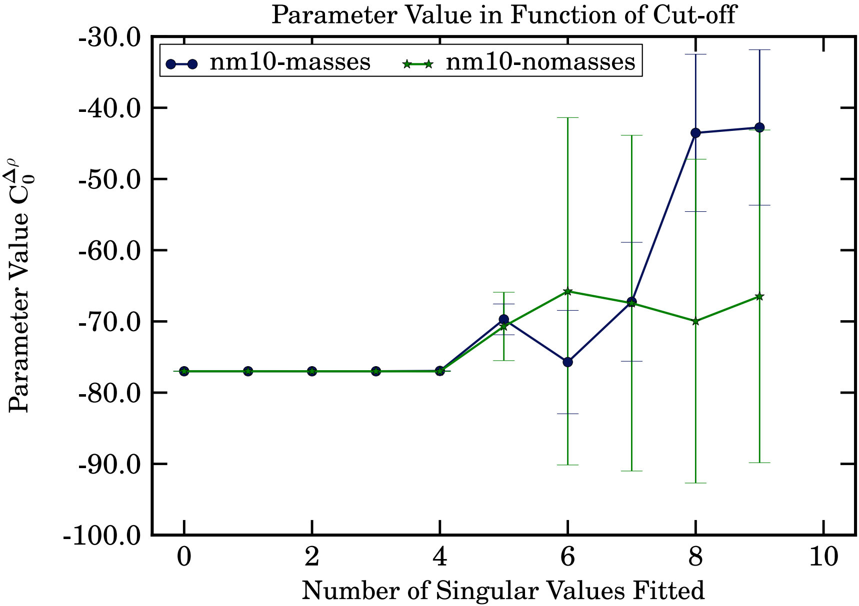
Regularized solutions



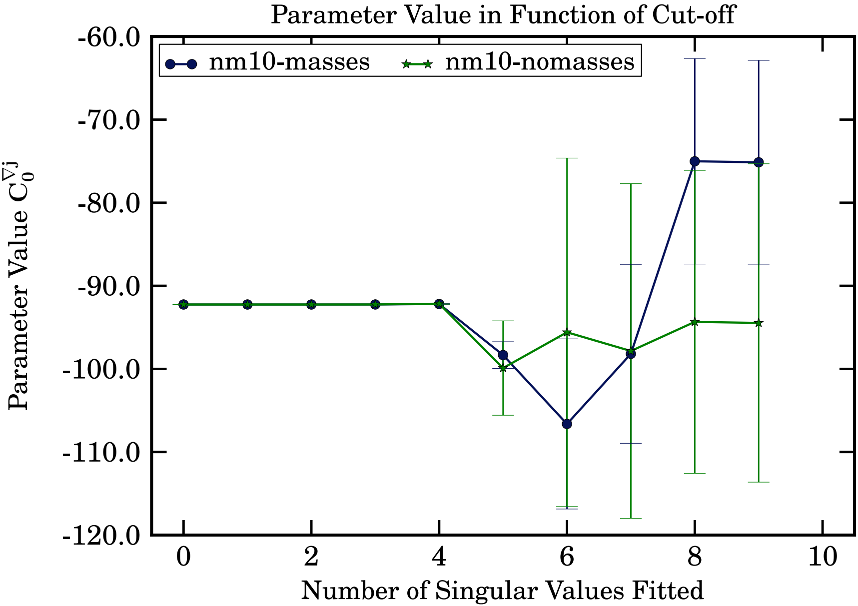
Regularized solutions



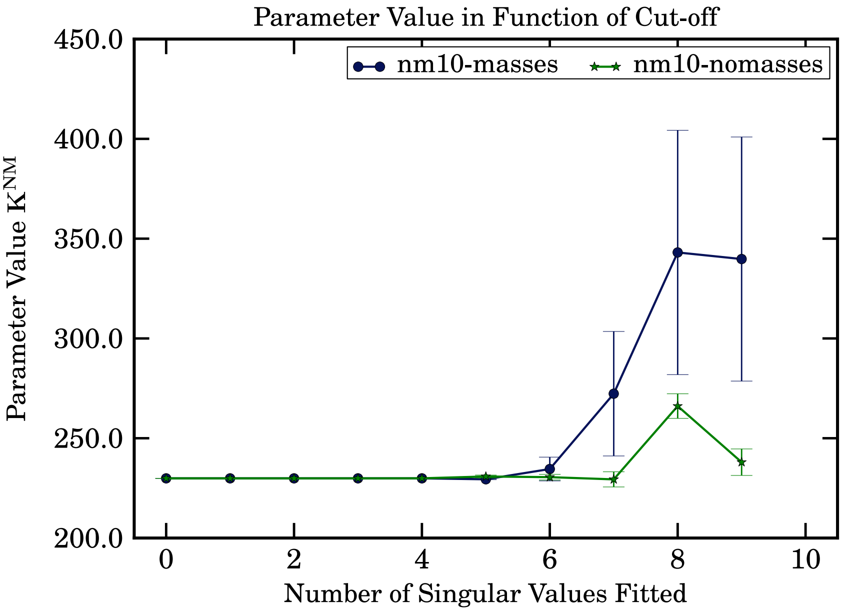
Regularized solutions



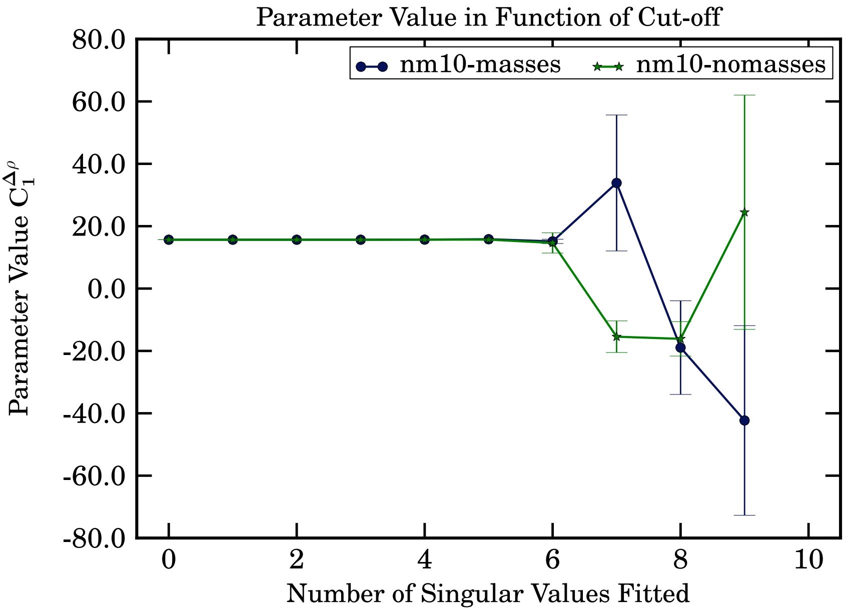
Regularized solutions



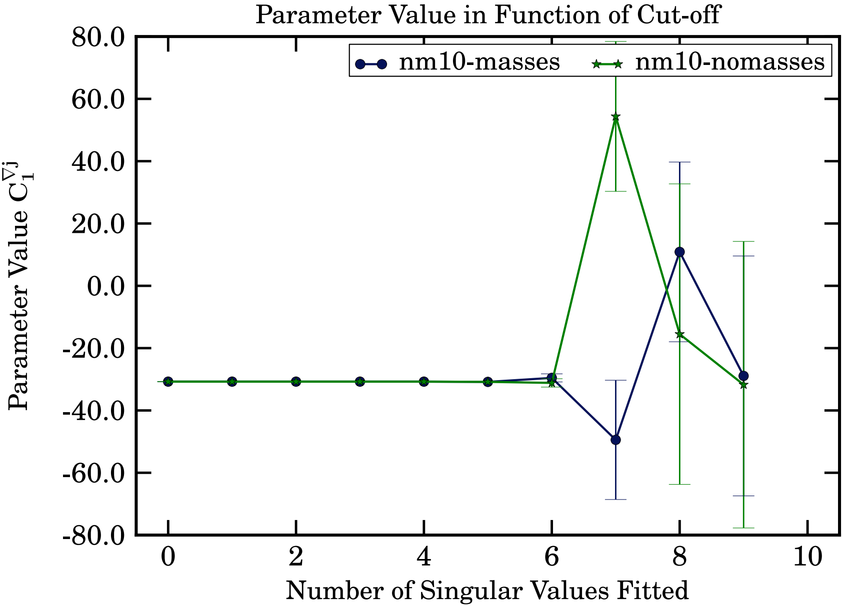
Regularized solutions



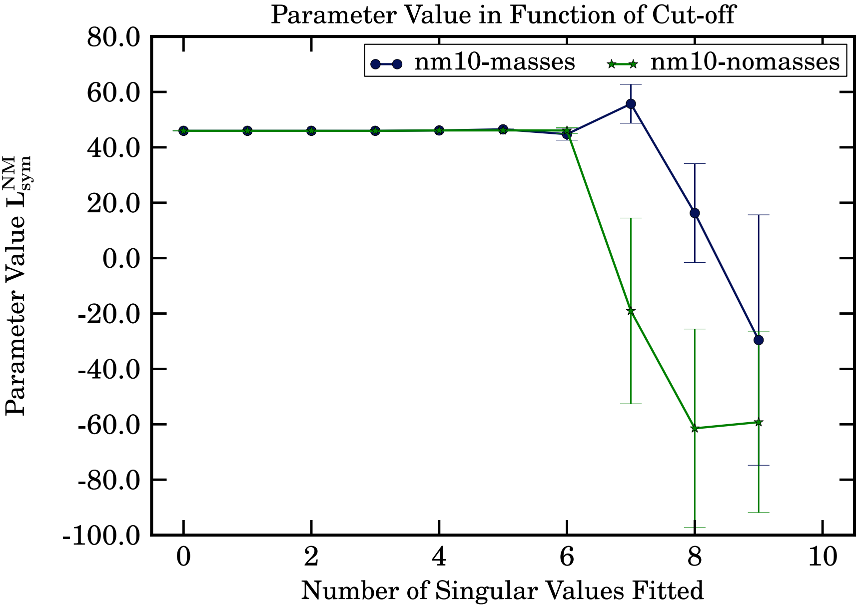
Regularized solutions



Regularized solutions

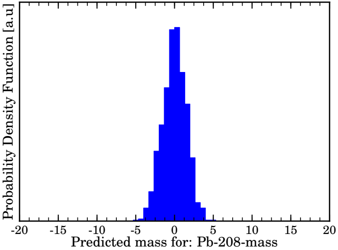


Regularized solutions

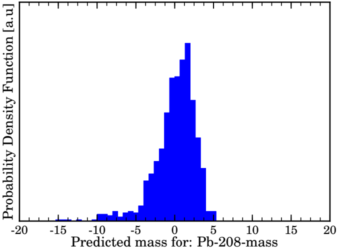


Predictions with regularized solutions

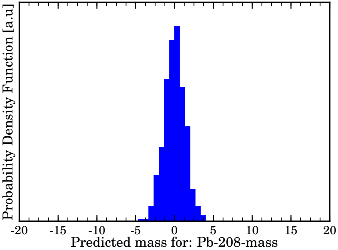
Params fitted: 1



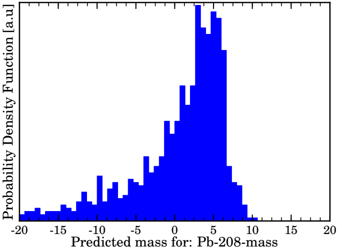
Params fitted: 7



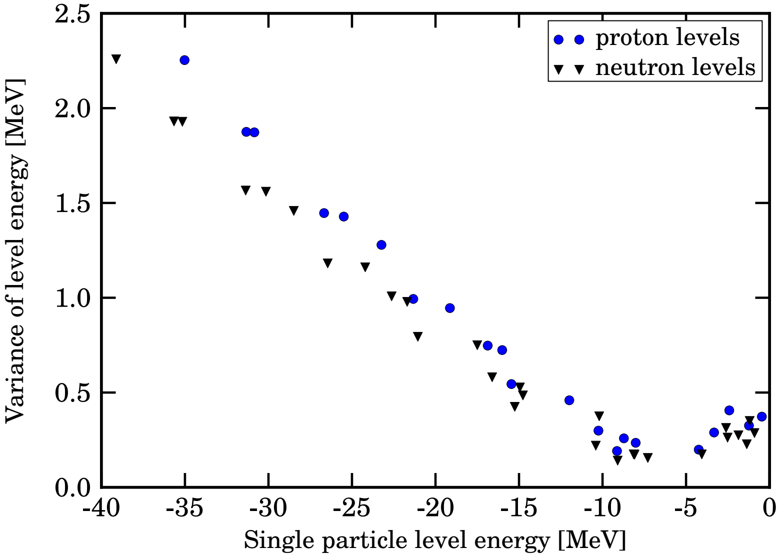
Params fitted: 5



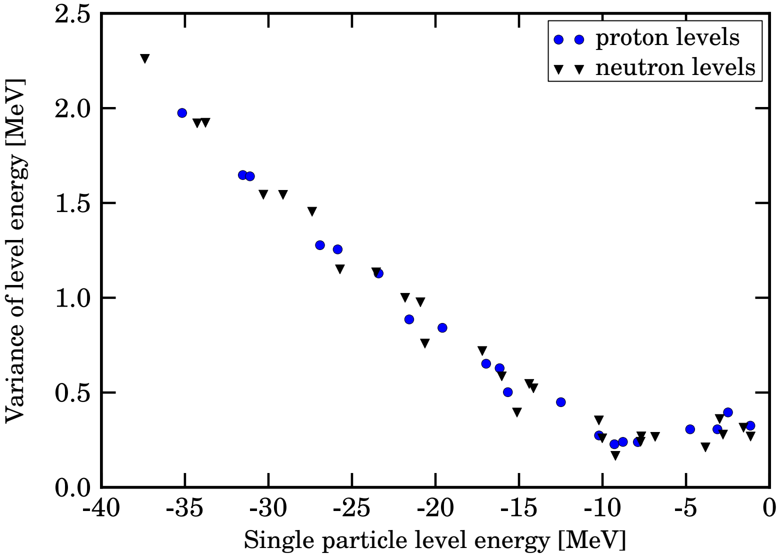
Params fitted: 9



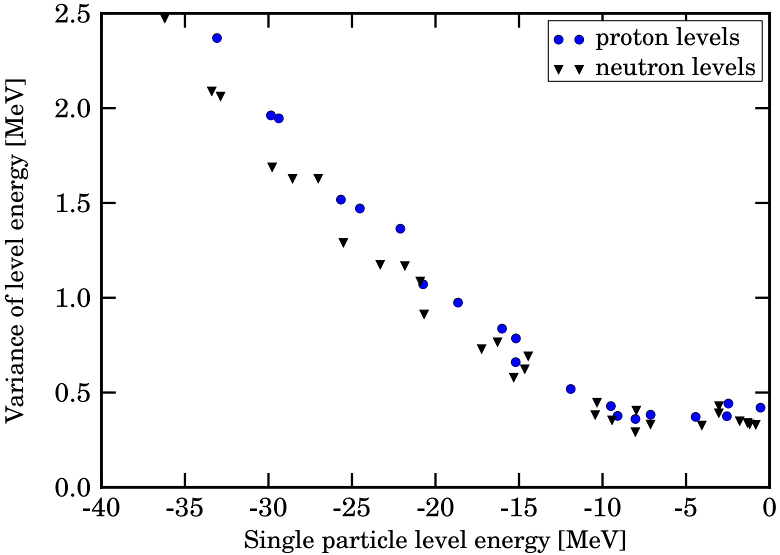
Predictions with regularized solutions: 5 params



Predictions with regularized solutions: 7 params



Predictions with regularized solutions: 9 params



Summary

Work done within COPIGAL collaboration:

- Development of flexible statistical analysis framework
- Analysis of SHF and Woods-Saxon models
- Promising results with classical analysis and SVD regularization
- Other methods tested: Monte-Carlo, bootstrap, Thikonov regularization
- Idea of synthetic data sets and MC works, questionable applications in true models

Conclusions

- Determination of EDF parameters - clearly a ill-posed problem
- Statistical analysis necessary when publishing any next parametrization
- "Technical aspects" of inverse problems not less important then "physical aspects"
- Regularization methods seem to be a useful method in avoiding overfit and obtaining unphysical values of some parameters.