Uncertainty Analysis of the nuclear Liquid Drop Model

B. Cauchois, Hongliang Lu, D. Boilley, Guy Royer

Uncertainty Analysis of the nuclear Liquid Drop Model

Bartholomé Cauchois, Hongliang Lü,* and David Boilley†

Grand Accélérateur National d'Ions Lourds (GANIL),
CEA/DRF - CNRS/IN2P3, BP 55027, F-14076 Caen Cedex,
France and Normandie Université, Unicaen, Caen, France

Guy Royer

Laboratoire Subatech, UMR: IN2P3/CNRS-Université-Institut Mines Télécom,
4 rue A. Kastler, 44307 Nantes Cedex 03, France
(Dated: May 14, 2018)

Abstract

The present work investigates the uncertainties in a nuclear liquid-drop model. The model parameters, their uncertainties and correlations, are determined through standard regression analysis that includes a statistical treatment of the errors of the model. The adjustment of the model is based on experimental binding energies corrected by the Thomas-Fermi shell corrections. The uncertainties are then propagated onto observables to reveal the reliability of the predictions and insight is provided to guide further research. Standard regression analysis is further extended to encompass experimental uncertainties as well as correlations. Strong support for such an extension can be found in Bayesian inference. The uncertainties and the correlations regarding microscopic corrections are not currently known, however, when available, these contributions may be included through this extended formalism.

PACS numbers: 21.10.Dr, 21.60.Ev, 02.50.-r, 02.30.Zz

I. INTRODUCTION

The phenomenological Bethe-Weizsäcker formula, proposed in 1935, can describe the binding energy of a broad range of nuclei, using only few elementary concepts [1, 2]. Since then, many refinements have been introduced allowing for a better description of the ever-expanding nuclear chart. The most significant development, was provided by Strutinsky, as semi-classical corrections, i.e., shell corrections, to the original model [3, 4]. An analogous prescription may be applied to procure pairing corrections. Later, deformation was included, both microscopically, i.e., shell and pairing corrections, and macroscopically, i.e., liquid drop model [5, 6]. Although, recent investigations [7, 8] have initiated detailed studies of such family of models, a thorough uncertainty analysis is yet to be provided and the ambition of the present text is to fill the current void.

The liquid drop model has tremendous implications when it comes to nuclear physics. In particular, the values of some parameters are directly related to infinite nuclear matter, e.g., the binding energy per nucleon or the symmetry energy, having profound consequences on our understanding of the equation of state and consequently, on precise stellar features. Furthermore, numerous observables can be deduced from binding energies, e.g., Q-values influencing the energies either required or released by nuclear reactions, thus, affecting nucleosynthesis. The liquid drop parameters are also used to calculate fission barriers, defined as the difference between the ground-state and the saddle-point energies and are crucial quantities for the synthesis of superheavy nuclei. Moreover, separation energies, as well as pairing effects may also be approached directly through masses. Therefore, one can only recognize the consequences of binding energies and their predictions upon our understanding of nuclear physics.

Since the last decade, phenomenological descriptions have being complemented by uncertainty analysis whose aim is to encompass the errors of the model. Methods, such as regression analysis, have been introduced to specifically address this question in precise quantitative terms. Regarding the current inquiry, we adopt the regression formalism presented in Refs. [9, 10] along with an extended version of weighted regression to simultaneously address experimental and theoretical uncertainties as well as their respective correlations.

The parameters of a liquid drop model are determined through standard regression analysis, thus, going beyond least squares by involving the errors of the model, the uncertainties, as well as the mutual influence of the parameters. These uncertainties can then be propagated onto observables to reveal the reliability of the predictions, disprove models or even guide further research.

There are several liquid drop formulas in the scientific literature, however, we shall focus on a single one found in Ref. [11]. Nevertheless, the ideas developed here may be applied to any linear model and can further be extended to account for nonlinearities as exposed in Ref. [8].

---

* Current address: Inria Grenoble Rhône-Alpes, 655 Avenue de l'Europe, 38330 Montbonnot-Saint-Martin, France
† boilley@ganil.fr

II. LIQUID DROP MODEL

The liquid drop model considered in the present work is taken from Ref. [11] and gives the following expression, for the theoretical binding energy,

$$B'_{Th,i} = \left(p_1 + p_2 I_i^2\right) A_i + \left(p_3 + p_4 I_i^2\right) A_i^{\frac{2}{3}} + p_5 \frac{Z_i^2}{A_i^{\frac{1}{3}}}$$
$$+ p_6 \frac{Z_i^2}{A_i} + p_7 \left|N_i - Z_i\right| e^{-\left(\frac{A_i}{50}\right)^2} + p_8 e^{-80 I_i^2}, \tag{1}$$

for the ith nucleus having N_i neutrons, Z_i protons, mass number $A_i = N_i + Z_i$, isospin $I_i = (N_i - Z_i)/A_i$ and where p_j are the parameters entering the model with $j = 1, \ldots, 8$.

The parameters in Eq. (1) are determined through regression using the following expression for the theoretically corrected experimental binding energies,

$$B'_{Exp,i} = B_{Exp,i} + E_{Pair,i} + E_{Shell,i}. \tag{2}$$

It includes the uncorrected experimental binding energy ($B_{Exp,i}$), the average pairing energy ($E_{Pair,i}$) and the shell correction energy ($E_{Shell,i}$). The uncorrected experimental nuclear binding energy is deduced from the atomic mass excess, found in the 2016 mass evaluation, accounting for the masses of the electrons as well as their binding energy [12, 13]. The pairing energy was directly taken from the Thomas-Fermi model [5], and is restated here for the sake of completeness:

$$E_{Pair,i} = \begin{cases} \dfrac{4.8}{N_i^{\frac{1}{3}}} + \dfrac{4.8}{Z_i^{\frac{1}{3}}} - \dfrac{6.6}{A_i^{\frac{2}{3}}} - \dfrac{30}{A_i} & (N_i = Z_i, \text{odd}) \\[3mm] \dfrac{4.8}{N_i^{\frac{1}{3}}} + \dfrac{4.8}{Z_i^{\frac{1}{3}}} - \dfrac{6.6}{A_i^{\frac{2}{3}}} & (N_i \text{ odd}, Z_i \text{ odd}) \\[3mm] \dfrac{4.8}{Z_i^{\frac{1}{3}}} & (N_i \text{ even}, Z_i \text{ odd}) \\[3mm] \dfrac{4.8}{N_i^{\frac{1}{3}}} & (N_i \text{ odd}, Z_i \text{ even}) \\[3mm] 0 & (N_i \text{ even}, Z_i \text{ even}) \end{cases} \tag{3}$$

Notice, the minus sign before the term $30/A_i$ instead of the plus sign appearing in Ref. [5] which seems to be a typographical error. In the following, the Thomas-Fermi shell corrections from Ref. [5] are used.

Note that Eq. (1) is linear and can easily be written in the matrix form,

$$\mathbf{B}'_{Th} = \mathbf{F} \cdot \mathbf{p}, \tag{4}$$

where $\mathbf{B}'_{Th}$ and $\mathbf{p}$ are column vectors, respectively, containing all the values of $B'_{Th,i}$ and p_j while the matrix $\mathbf{F}$ is defined as,

$$\mathbf{F} = \begin{bmatrix} \dfrac{\partial B'_{Th,1}}{\partial p_1} & \dfrac{\partial B'_{Th,1}}{\partial p_2} & \cdots & \dfrac{\partial B'_{Th,1}}{\partial p_{N_p}} \\[3mm] \dfrac{\partial B'_{Th,2}}{\partial p_1} & \dfrac{\partial B'_{Th,2}}{\partial p_2} & \cdots & \dfrac{\partial B'_{Th,2}}{\partial p_{N_p}} \\[3mm] \vdots & \vdots & \ddots & \vdots \\[3mm] \dfrac{\partial B'_{Th,Nn}}{\partial p_1} & \dfrac{\partial B'_{Th,Nn}}{\partial p_2} & \cdots & \dfrac{\partial B'_{Th,Nn}}{\partial p_{N_p}} \end{bmatrix} \tag{5}$$

where the ith line of $\mathbf{F}$ is associated with the ith nucleus while the jth column is related to the parameter p_j. Thus, the matrix $\mathbf{F}$ has as many lines as there are nuclei, N_n and as many columns as there are parameters, N_p and naturally, the vector $\mathbf{p}$ has as many lines as there are parameters.

In order to ease the notations of the following sections, the primes will be dropped while still implicitly referring to the corrected experimental binding energies defined by Eq. (2).

III. REGRESSION ANALYSIS

In this section, the mathematical formalism of standard regression, used throughout this text, is exposed and special attention is directed towards the distinction between the least squares method and regression analysis which is based on additional hypotheses and opens the door to uncertainty analysis. Most of the forthcoming material can be found in Refs. [9, 10].

A. The ordinary least squares method

In the literature related to liquid drop models, the least squares method is, more often than not, the favored method to determine the parameters. This method consists in minimizing the sum of the squared errors,

$$S = \sum_{i=1}^{N_n} \left(B_{Exp,i} - B_{Th,i}\right)^2, \tag{6}$$

with respect to the model parameters entering $B_{Th,i}$. The previous expression can be written in matrix form, cf. Eq. (4), as

$$S = \left(\mathbf{B}_{Exp} - \mathbf{F} \cdot \mathbf{p}\right)^T \cdot \left(\mathbf{B}_{Exp} - \mathbf{F} \cdot \mathbf{p}\right) \tag{7}$$

and its minimization with respect to the parameters yields the following solution,

$$\mathbf{p} = \left(\mathbf{F}^T \cdot \mathbf{F}\right)^{-1} \cdot \mathbf{F}^T \cdot \mathbf{B}_{Exp}. \tag{8}$$

The parameters obtained using Eq. (8) are identical to those given in Ref. [11] under the same conditions, thus, insuring the validity of the approach presented hitherto [14].

B. The regression hypotheses

In the present least squares method, the model is assumed exact and experimental uncertainties are ignored, therefore, theoretical uncertainties cannot be reached. However, this may be achieved, using regression analysis and in particular, by involving the errors of the model. It

is assumed, in regression analysis, that the experimental data can be described by

$$\mathbf{B}_{\mathrm{Exp}} = \mathbf{F} \cdot \mathbf{p} + \boldsymbol{\epsilon}, \tag{9}$$

where $\boldsymbol{\epsilon}$ represents the error associated with the model.

Assuming no prior understanding of the experimental data, the binding energies are just a set of distinct scattered values, thus, having a mean and a variance. A model's ambition is, of course, to explain why these values are all so different from each other and to uncover the origin of this variance. Thus, on the one hand, the model $\mathbf{F} \cdot \mathbf{p}$ is responsible for the explained variations, i.e., the corrected binding energy variations that the model is able to account for. On the other hand, the error $\boldsymbol{\epsilon}$ is accountable for the unexplained variations, i.e., the corrected binding energy variations left unexplained by the model and yet present in the experimental data. It should be borne in mind that the model is only meant to describe the corrected experimental binding energies, cf. Eq. (2), while the errors may also inherit flaws from the liquid-drop itself, the average pairing, the shell corrections, as well as other issues having unknown origins.

In order to determine the uncertainties in the parameters, some assumptions, regarding the distribution of these errors, are still required. For the simplest regression analysis, the first hypothesis concerns the mean of the errors which is assumed to vanish, i.e., $\mathrm{E}[\epsilon_i]=0$. The second hypothesis deals with the dispersion and the mutual influence of the errors which are assumed to possess a common variance σ^2 and to be uncorrelated, i.e., $\mathrm{COV}[\epsilon_i,\epsilon_j]=\sigma^2\,\delta_{ij}$. The errors having the same variance, the second hypothesis is also known as the homoscedastic assumption. Thus, according to these hypotheses, the errors follow a gaussian distribution with zero mean and variance σ^2. Highly convenient regression corollaries derive from these two prior assumptions:

$$\mathbf{E}\left[\mathbf{B}_{\mathrm{Exp}}\right] = \mathbf{F} \cdot \mathbf{p} \tag{10}$$

$$\mathbf{COV}\left[\mathbf{B}_{\mathrm{Exp}}\right] = \sigma^2 \mathbf{I}, \tag{11}$$

where $\mathbf{I}$ is the identity matrix.

C.   Unbiased sample estimates

As usual, it is assumed that only a restricted sample of all existing nuclei is known and an even smaller sample is actually considered here, as a result, all of the quantities upon which this study is based, will be sample estimates. It is conventional for these estimates to wear hats in order to distinguish them from quantities inferred from the whole population, i.e., all of the existing nuclei. Whenever, a quantity p, is estimated from a sample, it is said to be a sample estimate and will be noted p̂. Hence, a more appropriate notation for Eq. (8) would be,

$$\hat{\mathbf{p}} = \left(\mathbf{F}^{\mathrm{T}} \cdot \mathbf{F}\right)^{-1} \cdot \mathbf{F}^{\mathrm{T}} \cdot \mathbf{B}_{\mathrm{Exp}}. \tag{12}$$

Furthermore, if the mean of a sample estimate, i.e., E[p̂], is equal the corresponding population quantity, i.e., p, the estimate is said to be unbiased. Using the first regression hypothesis, one can indeed show that E[p̂]=p, thus, making p̂ an unbiased sample estimate of p.

D.   Uncertainty in the parameters

The parameters uncertainties are deduced from the $N_p \times N_p$ covariance matrix,

$$\begin{aligned}\mathbf{COV}\left[\hat{\mathbf{p}}\right] &= \mathbf{COV}\left[\left(\mathbf{F}^{\mathrm{T}} \cdot \mathbf{F}\right)^{-1} \cdot \mathbf{F}^{\mathrm{T}} \cdot \mathbf{B}_{\mathrm{Exp}}\right] \\ &= \sigma^2 \left(\mathbf{F}^{\mathrm{T}} \cdot \mathbf{F}\right)^{-1},\end{aligned} \tag{13}$$

where $\mathbf{F}$ is a constant matrix while $\mathbf{B}_{\mathrm{Exp}}$ is a random vector, implicitly carrying the variability, or said differently the variance, coming from the errors, cf. Eq. (9). The diagonal elements of $\mathbf{COV}[\hat{\mathbf{p}}]$ are the variances of the parameters, i.e., the square of their uncertainties, while the off-diagonal elements are the covariances between the parameters and characterize their mutual influence. This relation concentrates within a unique formula, uncertainties and covariances which are absolutely essential ingredients to any thorough uncertainty analysis.

The corresponding correlation matrix can always be deduced from the covariance matrix through the ensuing expression,

$$\mathbf{COR}\left[\hat{\mathbf{p}}\right] = \mathbf{VAR}\left[\hat{\mathbf{p}}\right]^{-\frac{1}{2}} \cdot \mathbf{COV}\left[\hat{\mathbf{p}}\right] \cdot \mathbf{VAR}\left[\hat{\mathbf{p}}\right]^{-\frac{1}{2}}, \tag{14}$$

where $\mathbf{VAR}[\hat{\mathbf{p}}]$ is formed from the diagonal of $\mathbf{COV}[\hat{\mathbf{p}}]$.

E.   Uncertainties in the observables

Once the covariance of the parameters is known, one can determine the uncertainties in the binding energies expressed in terms of the $N_n \times N_n$ covariance matrix,

$$\begin{aligned}\mathbf{COV}\left[\hat{\mathbf{B}}_{\mathrm{Th}}\right] &= \mathbf{COV}\left[\mathbf{F} \cdot \hat{\mathbf{p}}\right] \\ &= \mathbf{F} \cdot \mathbf{COV}\left[\hat{\mathbf{p}}\right] \cdot \mathbf{F}^{\mathrm{T}}.\end{aligned} \tag{15}$$

This expression only accounts for the uncertainty in the model, describing the corrected binding energies. However, it is also possible to include the uncertainty coming from the error of the model. Thus, further including all the flaws of the model yields the $N_n \times N_n$ covariance matrix,

$$\mathbf{COV}\left[\hat{\mathbf{B}}_{\mathrm{Th}} + \boldsymbol{\epsilon}\right] = \mathbf{COV}\left[\hat{\mathbf{B}}_{\mathrm{Th}}\right] + \sigma^2 \mathbf{I}. \tag{16}$$

Note that this expression differs from second corollary because it involves the estimates of the parameters, i.e., $\hat{\mathbf{B}}_{\mathrm{Th}} = \mathbf{F} \cdot \hat{\mathbf{p}}$ instead of $\mathbf{B}_{\mathrm{Th}} = \mathbf{F} \cdot \mathbf{p}$ in Eqs. (9) and (11).

Accordingly, separation energies and $Q-$values, deduced from differences between binding energies, i.e.,

$\hat{\mathbf{S}}_{\text{Th}} = \hat{\mathbf{B}}_{\text{Th,M}} - \hat{\mathbf{B}}_{\text{Th,D}} = -\hat{\mathbf{Q}}_{\text{Th}}$, have the following $N_n \times N_n$ covariance matrix,

$$
\begin{aligned}
\mathbf{COV}\left[\hat{\mathbf{S}}_{\text{Th}}\right] &= \mathbf{COV}\left[\hat{\mathbf{Q}}_{\text{Th}}\right] \\
&= \mathbf{F}_{\text{MD}} \cdot \mathbf{COV}\left[\hat{\mathbf{p}}\right] \cdot \mathbf{F}_{\text{MD}}^{\text{T}}
\end{aligned}
\tag{17}
$$

where $\mathbf{F}_{\text{MD}} = \mathbf{F}_{\text{M}} - \mathbf{F}_{\text{D}}$ and $\mathbf{F}_{\text{M}}$ together with $\mathbf{F}_{\text{D}}$, refer to matrices similar to $\mathbf{F}$, cf. Eq. (5), and respectively, containing only information regarding the mother and daughter nuclei.

F. Error estimates

Notice the presence of σ^2, i.e., the variance of the errors, in all of the previous expressions involving $\mathbf{COV}[\hat{\mathbf{p}}] = \sigma^2(\mathbf{F}^{\text{T}}\mathbf{F})^{-1}$. However, in practice, these relations cannot be applied since we do not dispose of an actual estimate, for either the errors or their variance. Fortunately, intuitive sample estimates for the errors, also known as residuals, can be found to be

$$
\hat{\boldsymbol{\epsilon}} = \mathbf{B}_{\text{Exp}} - \mathbf{F} \cdot \hat{\mathbf{p}}.
\tag{18}
$$

Furthermore, one can show that the following expression,

$$
\hat{\sigma}^2 = \frac{(\mathbf{B}_{\text{Exp}} - \mathbf{F} \cdot \hat{\mathbf{p}})^{\text{T}} \cdot (\mathbf{B}_{\text{Exp}} - \mathbf{F} \cdot \hat{\mathbf{p}})}{N_n - N_p},
\tag{19}
$$

is an unbiased sample estimate of the variance of these errors and therefore, is a precise measure of the variability left unexplained. It also turns out that this relation is the square of a well-known quantity, the root-mean-square deviation (RMS), usually characterizing the goodness-of-fit.

Replacing the variance of the errors by its estimate, i.e., σ^2 becomes $\hat{\sigma}^2$ in the preceding expressions, we gather a new set of useful formulas for the unbiased sample estimates of the covariance matrices, listed above, and at the heart of the present work.

Previously, we have introduced the error of the model that is distinguished from the uncertainty in the model. In the spirit of the international standards presented in Ref. [15] the error was defined as the difference between the "true" experimental binding energy and its theoretical estimate. On the other hand, the theoretical uncertainty characterizes the dispersion of the theoretical binding energy. We have also seen that all the uncertainties are proportional to the standard deviation of the errors. However, note that these two quantities are never equal.

IV. PRIMARY RESULTS

The formalism presented in Sect. III is now exploited in order to examine the uncertainties and the correlations of the parameters entering the liquid drop model,

cf. Eq. (1). A particular emphasis is given to the parameters, their uncertainties and correlations, as well as a diversity of observables.

The following results are based on the nuclear binding energies deduced from the atomic mass excesses found in Refs. [12, 13] for all nuclei satisfying N,Z $\geq$ 8 with uncertainties below 150 keV, thus, a total of 2315 nuclei are considered.

Note that the uncertainties in the experimental binding energies, the shell corrections and the pairing energies are disregarded at this stage.

A. Parameters

The parameters obtained using Eq. (12) are displayed in Table I. They are specific to the model considered in this study and cannot be directly compared to the parameters of other LD formulas. The corresponding uncertainties, also appearing in this table, were inferred from the diagonal elements of the covariance matrix, cf. Eq. (13), given in appendix A. Careful examination of the relative uncertainties establishes that the parameters, p_6, p_8 and to an even greater extent p_7, are loosely constrained with respect to the other parameters.

On the one hand, the parameters p_1 and p_5 are the most constrained as they correspond to the robust volume and Coulomb terms respectively. On the other hand, the parameter p_7 is the less constrained which means that the corresponding term is not well grounded. In Ref. [11], its value changes drastically when other terms are removed or added to the LD formula. Other observables than the mass should be considered to better establish such a term.

TABLE I. The parameters along with their uncertainties and relative uncertainties.

	$\hat{p}$ [MeV]	$\hat{u}(\hat{p})$[MeV]	$|\hat{u}(\hat{p})/\hat{p}|$ [%]
$\hat{p}_1$	15.4829	0.0145	0.1
$\hat{p}_2$	-27.8219	0.0843	0.3
$\hat{p}_3$	-17.5783	0.0505	0.3
$\hat{p}_4$	31.1447	0.3797	1.2
$\hat{p}_5$	-0.7058	0.0008	0.1
$\hat{p}_6$	0.9251	0.0288	3.1
$\hat{p}_7$	-0.2942	0.0293	10.0
$\hat{p}_8$	2.7265	0.1693	6.2

From the covariance matrix, cf. Table XI, one can deduce the correlation matrix, cf. Eq. (14), given in Table II. At first glance, it appears that all the parameters entering the liquid drop model are strongly correlated with one another. Meticulous inspection of this table, shows the existence of two correlation groups, i.e., sets of parameters with particularly strong correlations. The first involving p_1, p_3, p_5 and p_7, the second comprising p_2, p_4 and p_8. Notice, p_6 is less correlated with the other

parameters and the two groups are linked through the partial intrusion of p_1 in the second group and p_8 in the first group. Besides the Wigner term p_7, the first group, contains the historical liquid drop parameters while the second group, consists of isospin dependent corrections. Such an observation may be sufficient to physically explain the emergence of these two groups. Although not referred to as correlation groups, their existence has been studied in Ref. [7] by either removing or adding parameters to the regression and debating the corresponding effects on the remaining parameters. In the current analysis, a similar conclusion is drawn, directly, from a detailed examination of the correlation matrix shown Table II.

TABLE II. (Color online) Correlation matrix of the parameters deduced from the covariance matrix, cf. Table XI. Notations: **group 1**, *group 2* and intruders.

	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
$\hat{p}_1$	**1.00**	-0.71	**-0.93**	0.47	**-0.87**	-0.32	**0.74**	0.82
$\hat{p}_2$	-0.71	*1.00*	0.62	*-0.92*	0.39	0.51	-0.24	*-0.87*
$\hat{p}_3$	**-0.93**	0.62	**1.00**	-0.48	**0.95**	-0.05	**-0.81**	-0.70
$\hat{p}_4$	0.47	*-0.92*	-0.48	*1.00*	-0.22	-0.25	0.07	*0.72*
$\hat{p}_5$	**-0.87**	0.39	**0.95**	-0.22	**1.00**	-0.15	**-0.85**	-0.52
$\hat{p}_6$	-0.32	0.51	-0.05	-0.25	-0.15	1.00	0.15	-0.55
$\hat{p}_7$	**0.74**	-0.24	**-0.81**	0.07	**-0.85**	0.15	**1.00**	0.40
$\hat{p}_8$	0.82	*-0.87*	-0.70	*0.72*	-0.52	-0.55	0.40	*1.00*

B. Observables

1. Theoretical binding energies

Once the parameters are determined, one can easily compute the theoretical binding energy of each nucleus along with its uncertainty.

The root-mean-square deviation associated with the binding energies is $\hat{\sigma}$=602 keV. As already mentioned, it is an estimate of the standard deviation of the errors and quantifies the goodness-of-fit.

The uncertainties in the binding energy predictions, as deduced from Eq. (15), range from 21 keV to 124 keV with a mean of 33 keV. Those uncertainties are shown Fig. 1. Notice, the increase of uncertainties away from the mean proton number of the sample, i.e., 57.9. This is a well-known feature of regression procedures, even extending beyond the regression interval, thus, making extrapolations further and further away from the known data more and more uncertain.

Thus, the uncertainties in the predicted binding energies are rather small compared to the errors of the model ($\hat{\sigma}$). This suggests that once the shell corrections and the average pairing have been specified, the liquid-drop model is extremely constrained. Note that shell corrections were included in the corrected binding energies, however, they were excluded from the uncertainty
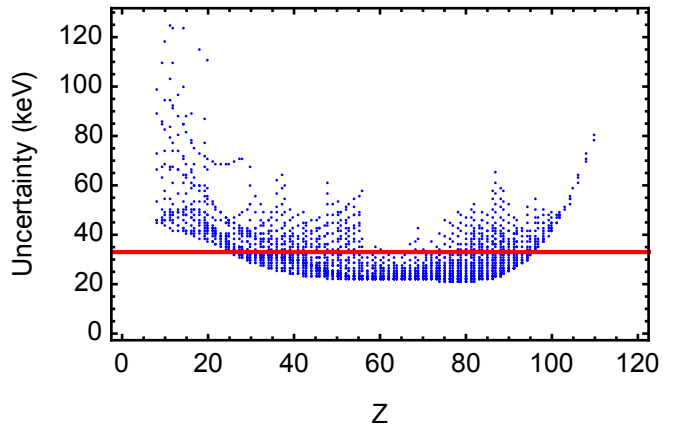

FIG. 1. (Color online) Uncertainties in the predicted binding energies (blue dots) as a function of proton number Z along with the mean of the uncertainties (red line).

analysis. When available, the latter should increase the uncertainties in the model constructed from, both, the liquid-drop and the shell corrections.

Adding the error of the model to the predicted binding energies as in Eq. (16) leads to uncertainties ranging from 602 to 615 keV because they are dominated by the error term.

The correlations between the binding energy of ^{208}Pb and the binding energies of the 2315 nuclei is depicted in Fig. 2. As expected, the correlations with neighbouring nuclei are very strong and fade away from this region.
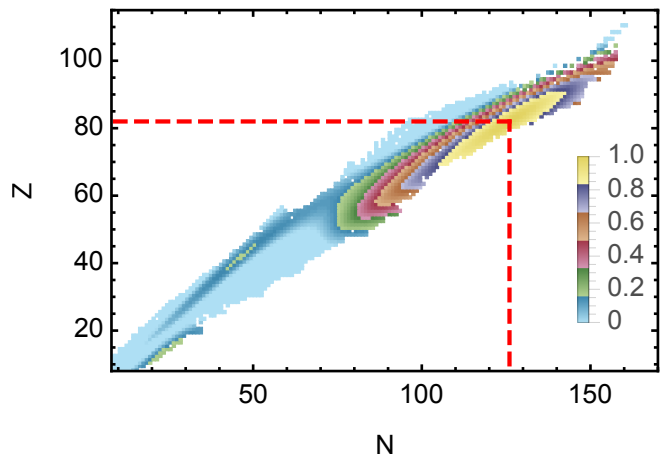

FIG. 2. (Color online) Correlations between the theoretical binding energy of ^{208}Pb and the 2315 binding energies. Notice, the extreme correlations in the region of ^{208}Pb and their downfall away from this region. The position of ^{208}Pb is indicated by the red dashed lines.

2. Other observables

Additional observables can be deduced from the binding energies. We have selected here a few mass filters for which experimental values are available over the whole nuclear chart.

For each observable the uncertainties in the predictions of the model were calculated. We present in Table III, the minimum and maximum uncertainties over the nuclear chart, as well as, a statistical indicator, the mean of the uncertainties, defined as

$$ \text{MU} = \frac{1}{N_n} \sum_{i=1}^{N_n} \text{VAR}[O_{\text{Th},i}]^{\frac{1}{2}}, \qquad (20) $$

for the observable O. Note that the indicator MU, has no profound statistical significance, and was only introduced, as a substitute for tables of unreasonable length, containing 2315 uncertainties for each observable. It appears that observables based on the difference between binding energies are less uncertain than the binding energies themselves. This is a consequence of the severe correlations existing between binding energies of neighbouring nuclei, cf. Fig. 2.

We have also compared the predicted values to the experimental ones. As a statistical indicator, the root-mean-square deviation (RMS) of an observable O, defined as

$$ \text{RMS} = \left[\frac{1}{N_n - N_p} \sum_{i=1}^{N_n} [O_{\text{Th},i} - O_{\text{Exp},i}]^2 \right]^{\frac{1}{2}}, \qquad (21) $$

is presented in Table III. As the model was adjusted to experimental binding energies, it is natural that this quantity possesses the smallest root-mean-square deviation.

TABLE III. The minimum, mean and maximum of the uncertainties (MinU, MU and MaxU) are presented in the uppermost portion of the table and the root-mean-square deviation (RMS) appears below, for the 2315 nuclei.

	$\hat{B}_{\text{Th}}$	$\hat{S}_n$	$\hat{S}_{2n}$	$\hat{S}_p$	$\hat{S}_{2p}$	$\hat{Q}_\alpha$	$\hat{Q}_\beta$
MinU [keV]	21	2	3	3	5	2	4
MU [keV]	33	6	11	8	15	5	13
MaxU [keV]	124	72	120	70	120	36	120
RMS [keV]	602	1273	1050	1367	985	1346	1847

The correlations between the theoretical Q_α-value of ^{208}Pb and the 2315 Q_α-values is exposed in Fig. 3. As formerly argued, the correlations with neighbouring nuclei are very strong and vanish away from this region. However, one can also distinguish more sophisticated patterns which were absent from the correlations between binding energies, cf. Fig. 2. Notice, these observations can further be extended to the remaining observables considered here. These complex patterns are owed to the involvement of twice as many nuclei. Indeed, binding energy correlations, only, involve two distinct nuclei while those between separation energies, or Q-values, require four, thus, making the structure much more intricate.
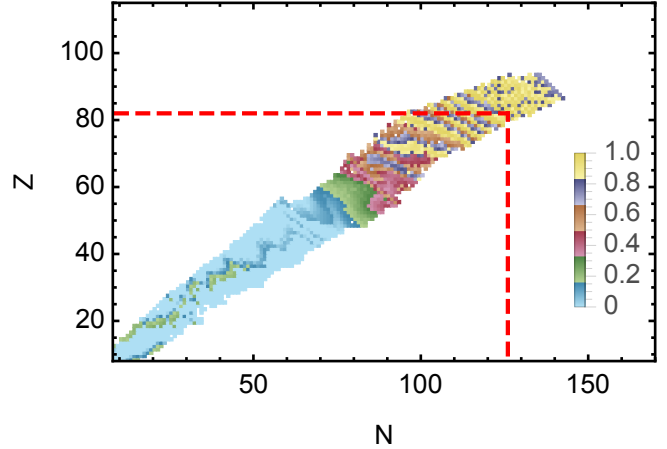

FIG. 3. (Color online) Correlations between the theoretical Q_α-value of ^{208}Pb and the 2315 Q_α-values. Notice, the extreme correlations in the region of ^{208}Pb and their downfall away from this region, as well as the complex patterns due to the involvement of a greater number of nuclei. The position of ^{208}Pb is indicated by the red dashed lines.

C. Assessment of the regression hypotheses

The results presented beforehand were procured through regression analysis which is based on hypotheses, regarding the distribution of the errors, as described in Sect. III B. This distribution of errors, associated with the binding energies, is shown Fig. 4. The corresponding mean and standard deviation are, respectively, 2 keV and $\hat{\sigma} = 602$ keV. These values were used to plot the normal distribution that exhibits a fair agreement with the histogram representing the distribution of the errors.

Other, more elaborate, methods are available to support the normality of the errors. The quantile-quantile plot is a graphical tool used to confirm that a set of data follows a specific probability distribution, e.g., a normal distribution. This is usually done by plotting the quantiles of the observed distribution of errors against the quantiles of the normal distribution having the same mean and variance, i.e., 2 keV and 602 keV. Indeed, if the observed errors do follow a normal distribution, then, the quantiles should be equal and the plot should form a straight line. The corresponding quantile-quantile plot is given Fig. 5 and confirms that most of the observed errors follow a normal distribution. However, the plot also makes it clear that the tails of distribution of the errors do not correspond to those of a normal distribution. Nevertheless, the central part, $\hat{\epsilon} \in$ [-1,1] MeV containing
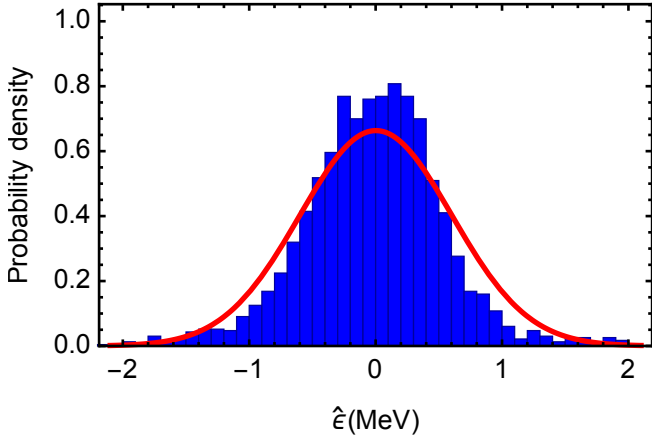
6

FIG. 4. (Color online) Histogram (blue bars) formed of 50 bins restricted to $\hat{\epsilon} \in$ [-2.1;2.1] MeV and representing the distribution of the errors ($\hat{\epsilon} = \mathbf{B}_{\mathrm{Exp}} - \mathbf{F} \cdot \hat{\mathbf{p}}$). The normal distribution constructed from the mean (2 keV) and the standard deviation ($\hat{\sigma} = 602$ keV) of the errors, is also provided (red line).

around 95% of the errors, does follow a normal distribution. Thus, the normality assumption, although not perfectly satisfied, should be enough to insure that the outcome of the current endeavour offers serious results.
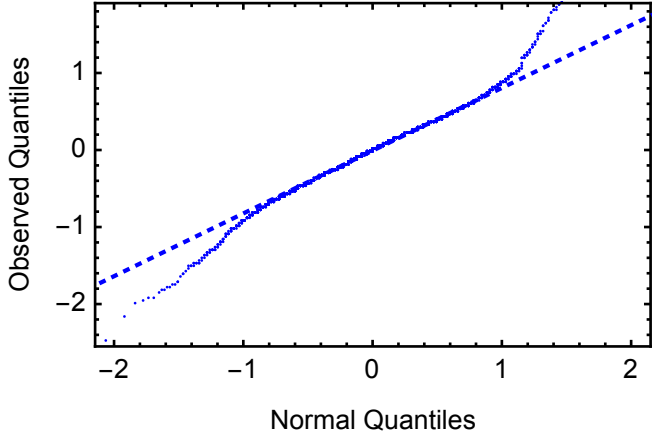

FIG. 5. (Color online) Quantile-quantile plot (blue dots) for the distribution of the errors restricted to $\hat{\epsilon} \in$ [-2.1;2.1] MeV and the expected theoretical straight line (dashed blue) in the case where normality is perfectly satisfied.

Notice, the mean of these errors does not exactly vanish and one can show that this hypothesis is only strictly satisfied by models including an intercept which macroscopic models do not, in general, contain. We now investigate the impact of adding a constant parameter p_0 to the liquid drop model, Eq. (1).

The parameters and their uncertainties were recalculated and the results are reported in Table IV. It appears that the additional intercept does not significantly alter

the previous outcome, i.e., p_6, p_7 and p_8 are still loosely constrained as well as the intercept p_0.

TABLE IV. The parameters along with their uncertainties and relative uncertainties obtained for the model with intercept.

	$\hat{p}$ [MeV]	$\hat{u}(\hat{p})$[MeV]	$|\hat{u}(\hat{p})/\hat{p}|$ [%]
$\hat{p}_0$	1.5896	0.3542	22.3
$\hat{p}_1$	15.5399	0.0193	0.1
$\hat{p}_2$	-27.9838	0.0914	0.3
$\hat{p}_3$	-17.9250	0.0922	0.5
$\hat{p}_4$	31.8989	0.4138	1.3
$\hat{p}_5$	-0.7088	0.0010	0.1
$\hat{p}_6$	1.0017	0.0334	3.3
$\hat{p}_7$	-0.3448	0.0313	9.1
$\hat{p}_8$	2.5587	0.1727	6.7

Table V reveals noticeable changes in the correlations, however, the group structure remains quite similar. There are still two correlation groups, a first consisting of p_1, p_3 and p_5, the second involving p_2 and p_4. Parameters, outside these groups, display moderate correlations and there are no intruders in this particular case.

TABLE V. (Color online) Correlation matrix obtained for the model with intercept deduced from the covariance matrix, cf. Table XII.
Notations: **group 1** and *group 2*.

	$\hat{p}_0$	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
$\hat{p}_0$	1.00	0.66	-0.39	-0.84	0.41	-0.64	0.51	-0.36	-0.22
$\hat{p}_1$	0.66	**1.00**	-0.75	**-0.93**	0.59	**-0.93**	0.13	0.28	0.46
$\hat{p}_2$	-0.39	-0.75	*1.00*	0.64	*-0.93*	0.53	0.20	-0.06	-0.70
$\hat{p}_3$	-0.84	**-0.93**	0.64	**1.00**	-0.58	**0.94**	-0.45	-0.11	-0.19
$\hat{p}_4$	0.41	0.59	*-0.93*	-0.58	*1.00*	-0.42	0.01	-0.09	0.55
$\hat{p}_5$	-0.64	**-0.93**	0.53	**0.94**	-0.42	**1.00**	-0.43	-0.38	-0.25
$\hat{p}_6$	0.51	0.13	0.20	-0.45	0.01	-0.43	1.00	-0.07	-0.57
$\hat{p}_7$	-0.36	0.28	-0.06	-0.11	-0.09	-0.38	-0.07	1.00	0.44
$\hat{p}_8$	-0.22	0.46	-0.70	-0.19	0.55	-0.25	-0.57	0.44	1.00

Regarding the binding energy, the mean uncertainty and root-mean-square deviation, respectively, become MU = 34 keV and $\hat{\sigma}$=600 keV. Thus, the differences between the uncertainties and the predictions obtained with and without an intercept are negligible. Note that, apart from the binding energy, the other observables are not affected by the intercept.

This assessment confirms that, although a more rigorous analysis should include an intercept in order to satisfy the first regression hypothesis, not doing so, does not significantly affect the outcomes of the study.

D. The impact of shell effects

In this study, an attempt was made to examine the impact of shell effects by using both, the Thomas-Fermi [5] and the Nix-Möller [6] microscopic corrections. The results show that the parameters marginally vary with the change of shell corrections, similarly, the absolute and relative uncertainties are barely altered and the correlation matrices are identical. The meagre influence of shell corrections is not surprising since they are identical for nuclei having N, Z > 30 in Refs. [5] and [6]. As only light nuclei contribute to this change, the effects are rather limited. Therefrom, no conclusions could be drawn about the exact impact of shell effects on either the estimation of the parameters, their uncertainties or even their correlations. Nevertheless, the outcome of the study suggested that the parameters could be significantly affected by the microscopic corrections while uncertainties and correlations suffered more minor changes.

E. Primary comments

It should be mentioned that, for all observables considered in this work, the uncertainties in the model predictions are insufficiently broad to account for the errors, thus, pointing towards some missing variability. Indeed, Fig. 6 reveals errors spreading way beyond the bounds of uncertainties which is coherent with prior results and can be understood from the root-mean-square deviation being virtually 20 times larger than the mean uncertainty.

Certainly, the scope of our uncertainty analysis, i.e., disregarding pairing and shell corrections, may be too crude and the resulting uncertainties too modest, to properly describe the uncertainties of the entire "macroscopic-microscopic" model. Nevertheless, it serves to distinguish the precision of the predictions from the uncertainties originating from only a limited portion of the model. This suggests that, once the microscopic corrections have been fixed, the experimental data leaves hardly any room for macroscopic variability, i.e., the macroscopic model is extremely constrained. Consequently, any further improvement of such model must be focused on constraining and evaluating uncertainties in the microscopic corrections.

Moreover, there is no model in nuclear physics that is able to reproduce the mass of the nuclei of the complete chart with a RMS lower than few hundreds of keV. Although it is a great achievement, this shows shortcomings in our understanding of the binding energy.

V. WEIGHTED REGRESSION

At this stage, experimental uncertainties or correlations were not taken into account. In the present section, we extend the regression analysis exposed in Sect. III to encompass this empirical evidence. Considering both the
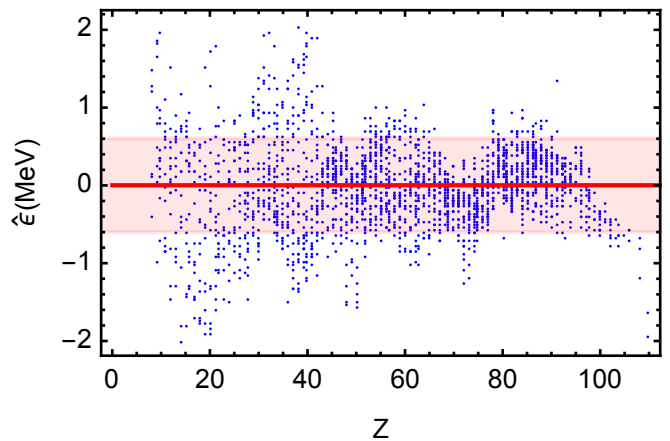

FIG. 6. (Color online) The errors ($\hat{\boldsymbol{\epsilon}} = \mathbf{B}_{\mathrm{Exp}} - \mathbf{F} \cdot \hat{\mathbf{p}}$) restricted to $\hat{\boldsymbol{\epsilon}} \in$ [-2.1;2.1] MeV are shown (blue dots) as a function of proton number Z. The mean of the errors (red line) and the standard deviation of the errors (red band), i.e., the root-mean-square deviation, are also displayed.

uncertainties and correlations or only the uncertainties is investigated in the following. Strong support for such an extension can be found in Bayesian inference [16].

Unfortunately, the uncertainties or the correlations regarding microscopic corrections are not currently known, however, when available, these contributions may be included through this extended formalism. For now, we shall focus on experimental uncertainties and their correlations.

A. Generalized regression

The uncertainties in the corrected binding energies were disregarded in the previous analysis and can be accounted for through a weighting matrix. In the current study, only experimental uncertainties are included. In this section, we present such a method where N_n experimental binding energies along with their uncertainties and correlations are used to determine N_p parameters. The formalism developed hereafter can be found in Ref. [10]. In this weighted formulation, the sum of squared errors becomes,

$$\mathrm{S} = (\mathbf{B}_{\mathrm{Exp}} - \mathbf{F} \cdot \mathbf{p})^{\mathrm{T}} \cdot \mathbf{W} \cdot (\mathbf{B}_{\mathrm{Exp}} - \mathbf{F} \cdot \mathbf{p}), \qquad (22)$$

where $\mathbf{W}$ is referred to as a weight matrix and assigns a weight to each nucleus or pair of nuclei if their experimental binding energies are correlated. Minimizing this expression with respect to the parameters leads to,

$$\hat{\mathbf{p}} = (\mathbf{F}^{\mathrm{T}} \cdot \mathbf{W} \cdot \mathbf{F})^{-1} \cdot (\mathbf{F}^{\mathrm{T}} \cdot \mathbf{W} \cdot \mathbf{B}_{\mathrm{Exp}}). \qquad (23)$$

Note that the regression assumptions are revised in this extension and the corollaries now become:

$$\mathbf{E} \left[\mathbf{B}_{\mathrm{Exp}} \right] = \mathbf{B}_{\mathrm{Th}} \qquad (24)$$

$$\mathbf{COV} \left[\mathbf{B}_{\mathrm{Exp}} \right] = \sigma^2 \mathbf{W}^{-1}. \qquad (25)$$

As before, the parameters are correlated and the co-variance matrix reads,

$$\mathbf{COV}\,[\hat{\mathbf{p}}] = \sigma^2\,(\mathbf{F}^{\mathrm{T}} \cdot \mathbf{W} \cdot \mathbf{F})^{-1}. \qquad (26)$$

The covariance matrix of the binding energies can be obtained from Eq. (15) and accordingly, observables deduced from differences between binding energies have the covariance matrix given by Eq. (17) where in both cases $\mathbf{COV}[\hat{\mathbf{p}}]$ is now defined by Eq. (26) instead of Eq. (13).

For all practical purposes, the estimate for the variance of the errors is,

$$\hat{\sigma}^2 = \frac{(\mathbf{B}_{\mathrm{Exp}} - \mathbf{F} \cdot \hat{\mathbf{p}})^{\mathrm{T}} \cdot \mathbf{W} \cdot (\mathbf{B}_{\mathrm{Exp}} - \mathbf{F} \cdot \hat{\mathbf{p}})}{\mathrm{N_n} - \mathrm{N_p}}. \qquad (27)$$

B. Empirical weights

The experimental binding energies are correlated through the specific method employed by the mass evaluation, these correlations are contained in the empirical covariance matrix $\mathbf{ECM}[\mathbf{B}_{\mathrm{Exp}}]$ also included in Refs. [12, 13] and having the structure below,

$$\mathbf{ECM}\,[\mathbf{B}_{\mathrm{Exp}}] = \begin{bmatrix} u_1^2 & u_{1,2} & \dots & u_{1,n} \\ u_{1,2} & u_2^2 & \dots & u_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1,n} & u_{2,n} & \dots & u_n^2 \end{bmatrix} \qquad (28)$$

where u_i^2 and $u_{i,j}$ are, respectively, the variance of the ith binding energy and the covariance between the ith and the jth binding energies. The uncertainties, deduced from the diagonal of the empirical covariance matrix, range from 160 meV to 149 keV with a mean of 7 keV, thus, spanning 6 orders of magnitude and making some nuclei significantly more important than other in the regression. It should also be mentioned that most of the binding energies are weakly correlated, e.g., 94% of the entries in the corresponding correlation matrix are in the range [-0.01;0.01].

We hereby describe three possible choices for the weight matrix all of which are constructed from the empirical covariance matrix.

(i) When all binding energies are presumed equivalent, i.e., experimental uncertainties and correlations are ignored, the weight matrix is simply the identity, $\mathbf{W}_1 = \mathbf{I}$. This is the favored literature standard investigated beforehand in Sects. III and IV, also corresponding to the homoscedastic hypothesis.

(ii) While neglecting correlations and only considering experimental uncertainties, the weight matrix preserves its diagonal form, the ith element being the inverse uncertainty squared associated with the ith binding energy, i.e., u_i^{-2}. Accordingly, the weight matrix is in fact the inverse of the empirical variance matrix, i.e., $\mathbf{W}_2 = \mathbf{EVM}[\mathbf{B}_{\mathrm{Exp}}]^{-1}$. Thus, in this particular weighing scheme, binding energies having larger uncertainties lose their influence over the regression.

(iii) Finally, the inclusion of correlations along with uncertainties leads to the weight matrix, $\mathbf{W}_3 = \mathbf{ECM}[\mathbf{B}_{\mathrm{Exp}}]^{-1}$. This weighing scheme incorporates every single empirical evidence at our disposal. Binding energies having larger uncertainties or being correlated, both lose their leverage over the regression. The latter is important such that no over-counting of information occurs.

As argued in Ref. [8], macroscopic models, such as the one considered in the present work, together with microscopic corrections deliver fairly inaccurate predictions, i.e., away from the measured values and far beyond experimental uncertainties, usually making the regression (ii) inadequate. Nevertheless, the approach (iii) has never been studied and could turn out to be more attractive while going further than approach (ii) through the inclusion of empirical correlations.

VI. SECONDARY RESULTS

The extended formalism exposed in Sect. V is applied in order to determine the uncertainties and the correlations in the parameters entering the model. A special attention is directed towards the choices of empirical weights and their effects on the outcomes.

The subsequent results are based on the nuclear binding energies deduced from the atomic mass excesses found in Refs. [12, 13] for all nuclei satisfying N,Z $\geq$ 8 with uncertainties below 150 keV and present in the empirical covariance matrix. Thus, a total of 1088 nuclei were considered.

A. The parameters

The parameters obtained using Eq. (23) are presented in Table VI with the corresponding uncertainties, deduced from the diagonal elements of the covariance matrix, cf. Eq. (26), given in appendix B.

Since the sets of nuclei involved are not identical, the parameters obtained with the regression (i) differ from those obtained in Sect. IV A. The inclusion of experimental uncertainties, i.e., regressions (ii) and (iii), changes the weight of each nucleus, thus, influencing the values taken by the parameters. In particular, p_2, p_4, p_7 and p_8 suffer dramatic changes and correspond to isospin dependent terms. On the one hand, exotic nuclei, characterized by larger isospins, have larger uncertainties and therefore, loose their influence over the regression. On the other hand, p_1 and p_5 are quite stable as the volume and Coulomb terms are constrained by larger number of nuclei, i.e., all nuclei independently of their isospin. This is consistent with the fact that they have the smallest

relative uncertainty. The prior argument is supported by the Fig. 7 proving that exotic nuclei have larger uncertainties. Furthermore, the distribution of the uncertainties presented in Fig. 8 is very peaked around the origin where nearly half of the population is grouped. Combining the conclusions drawn from Figs. 7 and 8 demonstrates that exotic nuclei should significantly loose their influence over the regressions (ii) and (iii).
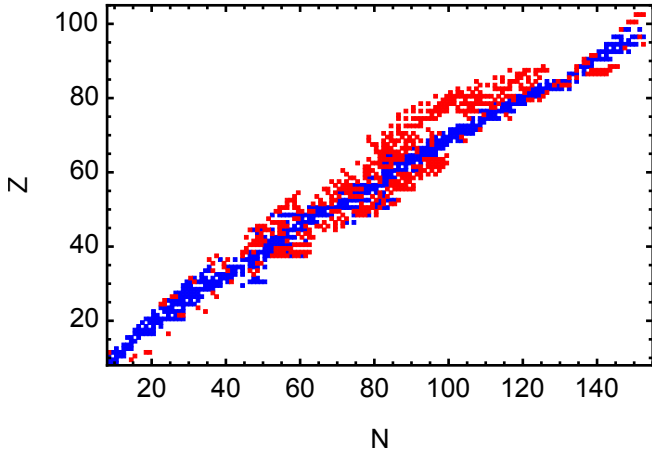

FIG. 7. (Color online) The 1088 nuclei involved in the study were divided into two equal groups each containing 544 nuclei. The first group (blue) includes those having a binding energy uncertainty below 2.6 keV while the second (red) contains those having a binding energy uncertainty above 2.6 keV. The second group is composed of more exotic nuclei than the first group.
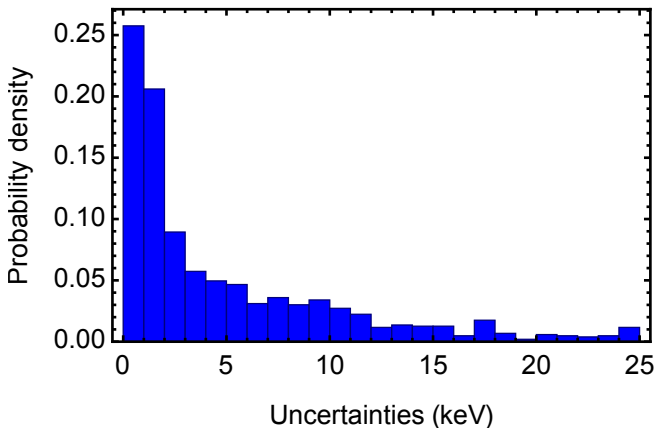

FIG. 8. (Color online) Histogram formed of 25 bins, representing distribution of the binding energies uncertainties, restricted to the range [0;25] keV and including 1029 nuclei. A fourth of the 1088 nuclei have uncertainties below 1 keV and half below 2.6 keV.

The inclusion of experimental covariances seems to reduce the effects of the experimental uncertainties, consequently, the departure away from the results obtained with the homoscedastic regression (i) is diminished.

TABLE VI. The parameters along with their uncertainties and relative uncertainties as determined using the regressions (i), (ii) and (iii).

| | $\hat{p}$ [MeV] | | | $\hat{u}(\hat{p})$ [MeV] | | | $|\hat{u}(\hat{p})/\hat{p}|$ [%] | | |
	$\mathbf{W}_1$	$\mathbf{W}_2$	$\mathbf{W}_3$	$\mathbf{W}_1$	$\mathbf{W}_2$	$\mathbf{W}_3$	$\mathbf{W}_1$	$\mathbf{W}_2$	$\mathbf{W}_3$
$\hat{p}_1$	15.343	15.651	15.599	0.025	0.020	0.016	0.2	0.1	0.1
$\hat{p}_2$	-27.034	-44.040	-38.012	0.141	0.635	0.531	0.5	1.4	1.4
$\hat{p}_3$	-16.980	-19.340	-18.540	0.087	0.162	0.159	0.5	0.8	0.9
$\hat{p}_4$	27.423	118.319	87.799	0.633	3.820	3.488	2.3	3.2	4.0
$\hat{p}_5$	-0.698	-0.683	-0.690	0.001	0.002	0.002	0.2	0.3	0.2
$\hat{p}_6$	0.840	1.188	0.739	0.045	0.130	0.151	5.3	11.0	20.4
$\hat{p}_7$	-0.669	-0.070	-0.445	0.060	0.110	0.117	8.9	157.6	26.3
$\hat{p}_8$	1.365	12.130	9.858	0.255	0.401	0.385	18.7	3.3	3.9

A cautious analysis of the correlation matrices, given in Tables VII, VIII and IX, shows, in the case of homoscedastic regression (i), the same two correlation groups as in Sect. IV A. The first involving p_1, p_3, p_5 and p_7 and the second implicating p_2, p_4 and p_8 while p_6 is almost uncorrelated with the other parameters. These two groups communicate mostly through the incursions of p_1 in the second group and p_8 in the first group. For both regressions (ii) and (iii), two groups can also be distinguished, a weakly correlated first group consisting of p_1, p_5 and a second with stronger correlations containing all of the six remaining parameters. In both cases, the two groups are well separated and are not coupled through decisive correlations, i.e., there are no intruders. Notice, regression (i) leads to overall weaker correlations than regression (iii) which, itself, leads to weaker correlations than regression (ii). Thus, the empirical evidence, contained in the weights matrices, seems to globally strengthen the correlations and to change the group structure.

TABLE VII. (Color online) Correlation matrix deduced from the covariance matrix, cf. Table XIII, with regression (i). Notations: **group 1**, *group 2* and intruders.

	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
$\hat{p}_1$	**1.00**	-0.75	**-0.94**	0.52	**-0.89**	-0.27	**0.76**	0.88
$\hat{p}_2$	-0.75	*1.00*	0.67	*-0.92*	0.47	0.44	-0.43	*-0.90*
$\hat{p}_3$	**-0.94**	0.67	**1.00**	-0.55	**0.96**	-0.07	**-0.84**	-0.79
$\hat{p}_4$	0.52	*-0.92*	-0.55	*1.00*	-0.32	-0.18	0.33	*0.74*
$\hat{p}_5$	**-0.89**	0.47	**0.96**	-0.32	**1.00**	-0.16	**-0.83**	-0.65
$\hat{p}_6$	-0.27	0.44	-0.07	-0.18	-0.16	1.00	0.17	-0.44
$\hat{p}_7$	**0.76**	-0.43	**-0.84**	0.33	**-0.83**	0.17	**1.00**	0.55
$\hat{p}_8$	0.88	*-0.90*	-0.79	*0.74*	-0.65	-0.44	0.55	*1.00*

TABLE VIII. (Color online) Correlation matrix deduced from the covariance matrix, cf. Table XIV, with regression (ii). Notations: **group 1** and *group 2*.

	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
$\hat{p}_1$	**1.00**	-0.42	-0.58	0.37	**-0.76**	0.20	0.44	0.54
$\hat{p}_2$	-0.42	*1.00*	*0.91*	*-0.99*	0.17	*-0.85*	*-0.88*	*-0.94*
$\hat{p}_3$	-0.58	*0.91*	*1.00*	*-0.94*	0.47	*-0.91*	*-0.97*	*-0.99*
$\hat{p}_4$	0.37	*-0.99*	*-0.94*	*1.00*	-0.22	*0.92*	*0.92*	*0.96*
$\hat{p}_5$	**-0.76**	0.17	0.47	-0.22	**1.00**	-0.26	-0.33	-0.38
$\hat{p}_6$	0.20	*-0.85*	*-0.91*	*0.92*	-0.26	*1.00*	*0.94*	*0.90*
$\hat{p}_7$	0.44	*-0.88*	*-0.97*	*0.92*	-0.33	*0.94*	*1.00*	*0.98*
$\hat{p}_8$	0.54	*-0.94*	*-0.99*	*0.96*	-0.38	*0.90*	*0.98*	*1.00*

TABLE IX. (Color online) Correlation matrix deduced from the covariance matrix, cf. Table XV, with regression (iii). Notations: **group 1** and *group 2*.

	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
$\hat{p}_1$	**1.00**	-0.37	-0.32	0.25	**-0.71**	-0.00	0.04	0.26
$\hat{p}_2$	-0.37	*1.00*	*0.90*	*-0.98*	0.39	*-0.79*	*-0.80*	*-0.91*
$\hat{p}_3$	-0.32	*0.90*	*1.00*	*-0.95*	0.56	*-0.94*	*-0.95*	*-0.99*
$\hat{p}_4$	0.25	*-0.98*	*-0.95*	*1.00*	-0.41	*0.89*	*0.90*	*0.96*
$\hat{p}_5$	**-0.71**	0.39	0.56	-0.41	**1.00**	-0.41	-0.39	-0.48
$\hat{p}_6$	-0.00	*-0.79*	*-0.94*	*0.89*	-0.41	*1.00*	*0.99*	*0.94*
$\hat{p}_7$	0.04	*-0.80*	*-0.95*	*0.90*	-0.39	*0.99*	*1.00*	*0.96*
$\hat{p}_8$	0.26	*-0.91*	*-0.99*	*0.96*	-0.48	*0.94*	*0.96*	*1.00*

TABLE X. The root-mean-square deviation (RMS) as well as the mean of the uncertainties (MU) are calculated for various observables and for the regressions (i), (ii) and (iii).

		$\hat{B}_{Th}$	$\hat{S}_n$	$\hat{S}_{2n}$	$\hat{S}_p$	$\hat{S}_{2p}$	$\hat{Q}_\alpha$	$\hat{Q}_\beta$
	$\mathbf{W}_1$	42	8	14	11	19	7	17
MU [keV]	$\mathbf{W}_2$	258	35	68	53	106	42	86
	$\mathbf{W}_3$	296	29	55	42	83	33	68
	$\mathbf{W}_1$	524	1333	1158	1394	986	1371	1758
RMS [keV]	$\mathbf{W}_2$	4841	1480	1658	1670	1918	1500	2261
	$\mathbf{W}_3$	2365	1371	1267	1465	1165	1400	1891

C. Secondary comments

Although, the weights considered did not improve the adjustment of the model, incoming theoretical developments, leading to enhanced model precision and the future, large scale, mass measurements campaigns, will further increase the importance of empirical uncertainties and correlations, consequently, weighted regression should replace standard regression. Moreover, uncertainties in microscopic corrections are expected to be highly correlated, therefore, when available, their inclusion will require the use of regression (iii). Finally, as the parameters seem to be better constrained and the corresponding predictions are more accurate, the present work seems to suggest that employing regression (iii) may be preferable to using regression (ii).

VII. CONCLUSION

Mass models are widely used in nuclear physics especially when it comes to make predictions. However, we have seen that the errors of the model lead to uncertainties in the predictions which should be taken into account in order to make comparisons with experimental data [17]. In our case, we are specially interested in improving the predictive power of the models used to estimate the fusion-evaporation cross-sections to synthesize super-heavy elements by the mean of uncertainty analysis [18].

Regression analysis is a standard method based on a large mathematical corpus that accounts for the errors of the model in the uncertainty estimate. Thus, we applied it to the liquid drop model which lead us to the covariance of the parameters and more precisely, to their uncertainties and correlations which were then propagated onto observables, thus, providing all of the elements for a thorough uncertainty analysis.

The first part of the present work confirms the presence of strong correlations between the parameters and furthermore, the emergence of correlation groups, i.e., involving many parameters and coupled together through the intrusion of members of other groups. It was also shown that although some parameters are better con-

B. The observables

The set of observables considered in Sect. IV B is reexamined with special care regarding the impact of empirical weights on predictions as well as uncertainties. The same statistical indicators are summarized in Table X.

In the case of the homoscedastic regression (i), the regression encompasses a reduced number of nuclei and leads to a smaller root-mean-square deviation than in the primary analysis. Since it is no longer the the quantity being minimized, the RMS in the heteroscedastics regressions (ii) and (iii) are far larger than in the homoscedastic (i) one. On the other hand, the uncertainties in the prediction of the model are larger for the homoscedastic regression (i). A closer look at Table X shows a different outcome regarding regressions (ii) and (iii) where the binding energy exhibits the worst predictions as well as the most uncertain ones, i.e., associated with the largest RMS and mean uncertainty.

Analyzing the values obtained for the RMS, shows that regression (iii) yields systematically superior predictions than regression (ii).

strained than others, once the microscopic corrections are fixed, the experimental evidence leaves hardly any room for macroscopic variability, consequently, further research must be focused on constraining microscopic corrections in order to improve the model as a whole.

The second part, focuses on the influence of empirical uncertainties and correlations on the results obtained in the first part of the study. The errors being much greater than the empirical variances or covariances, the weights did not improve the adjustment of the model. Nevertheless, it was shown that the inclusion of empirical correlations yields more suitable results than those obtained only accounting for experimental uncertainties. With future experimental and theoretical achievements, weighted regression, involving empirical correlations, will prove to be more attractive than only considering experimental uncertainties. This method opens the doors to the inclusion of uncertainties associated with the microscopic corrections which are expected to be highly correlated.

Independently of the method, the errors of the model are far larger than the uncertainties in the liquid drop model. The extent of our uncertainty analysis, i.e., neglecting pairing and shell corrections, may be too coarse to properly describe the uncertainties of the entire "macroscopic-microscopic" model. Moreover, the coherence, between the macroscopic and microscopic contributions, can only be achieved if the same deformation is assumed for both of these components. In this development, this coherence was lost, as the macroscopic model considered above, bears no deformation while microscopic corrections do, thus making the two contributions inconsistent with one another. Furthermore, the present work disregards any uncertainty or correlation for either the shell or the pairing corrections and are hereby, presumed, independent of the macroscopic component. As such, all of the conclusions drawn from this analysis were reached consciously neglecting these inconsistencies and hoping that further investigations will fill those holes.

Finally, it is worth recalling that no model can reduce the RMS on the whole nuclear chart to values lower than few hundreds keV. Therefore, extended uncertainty evaluation to the microscopic part of the model will still give results that remain lower than the errors of the model.

Appendix A: Covariance of primary results

In this appendix, the covariance between the parameters, obtained from regression analysis using Eq. (13), are provided in their matrix form. These results correspond to those contained in the first part of the present study.

In the case of the model without an intercept, the correlations between the parameters, presented in Table II, can be deduced from the covariance matrix given Table XI, using Eq. (14).

In the case of the model with an intercept, the correlations between the parameters, presented in Table V, can be deduced from the covariance matrix given Table XII, using Eq. (14).

Appendix B: Covariance of secondary results

In this appendix, the covariance between the parameters, obtained from weighted regression analysis using Eq. (26), are provided in their matrix form. These results correspond to those contained in the second part of the present study, using the regressions i, ii and iii which were defined in Sect. V B.

The correlations between the parameters, presented in Table VII, can be deduced from the covariance matrix given Table XIII, using Eq. (14).

The correlations between the parameters, presented in Table VIII, can be deduced from the covariance matrix given Table XIV, using Eq. (14).

The correlations between the parameters, presented in Table IX, can be deduced from the covariance matrix given Table XV, using Eq. (14).

[1] C. von Weizsäcker, Z. Phys. **96**, 431 (1935).
[2] H. Bethe and R. Backer, Rev. Mod. Phys. **8**, 82 (1936).
[3] V. Strutinsky, Nucl. Phys. A **3**, 420 (1967).
[4] V. Strutinsky, Nucl. Phys. A **122**, 1 (1968).
[5] W. Myers and W. Swiatecki, LBL report 36803 (Dec. 1994) uC-413.
[6] W. Myers, W. Swiatecki, P. Möller, and J. Nix, At. Data Nucl. Data Tables **59**, 185 (1995).
[7] M. Kirson, Nucl. Phys. A **798**, 29 (2008).
[8] J. Toivanen, J. Dobaczewski, M. Kortelainen, and K. Mizuyama, Phys. Rev. C **78**, 426 (2008).
[9] S. Chatterjee and A. S. Hadi, *Regression analysis by example* (Wiley, 2006).
[10] Y. Dodge and V. Rousson, *Analyse de régression appliquée* (Dunod, 1999).
[11] G. Royer and A. Subercaze, Nucl. Phys. A **917**, 1 (2013).
[12] W. Huang and e. a. G. Audi, Chin. Phys. C **41** (2017).
[13] M. Wang and e. a. G. Audi, Chin. Phys. C **41** (2017).
[14] The tables containing rounded experimental data and the shell corrections were provided by the main author of Ref. [11].
[15] BIPM, IEC, IFCC, ILAC, ISO, and IUPAC, JCGM **100** (2008).
[16] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial* (Oxford University Press, 2006).
[17] T. Editors, Phys. Rev. A **83**, 040001 (2011).

TABLE XI. Covariance matrix between the parameters for the model without intercept.

$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
0.0002112255	-0.0008757963	-0.0006794524	0.0025757144	-0.0000102194	-0.0001320633	0.0003152269	0.0020147598
-0.0008757963	0.0071133452	0.0026224132	-0.0293433212	0.0000265519	0.0012374354	-0.0005926571	-0.0124508863
-0.0006794524	0.0026224132	0.0025454952	-0.0092533611	0.0000386468	-0.0000681243	-0.0012004781	-0.0059446316
0.0025757144	-0.0293433212	-0.0092533611	0.1441782758	-0.0000678873	-0.0027366304	0.0007282969	0.0460073012
-0.0000102194	0.0000265519	0.0000386468	-0.0000678873	0.0000006481	-0.0000035655	-0.0000200434	-0.0000701814
-0.0001320633	0.0012374354	-0.0000681243	-0.0027366304	-0.0000035655	0.0008284236	0.0001229168	-0.0026805039
0.0003152269	-0.0005926571	-0.0012004781	0.0007282969	-0.0000200434	0.0001229168	0.0008585116	0.0019737292
0.0020147598	-0.0124508863	-0.0059446316	0.0460073012	-0.0000701814	-0.0026805039	0.0019737292	0.0286534371

TABLE XII. Covariance matrix between the parameters for the model with an intercept.

$\hat{p}_0$	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
0.1254667086	0.0044991307	-0.0127792513	-0.0273692844	0.0595335998	-0.0002367796	0.0060480260	-0.0039936493	-0.0132486054
0.0044991307	0.0003708226	-0.0013268437	-0.0016553016	0.0046893470	-0.0000186260	0.0000859004	0.0001694244	0.0015230991
-0.0127792513	-0.0013268437	0.0083564328	0.0053885001	-0.0351656109	0.0000504503	0.0006112401	-0.0001810128	-0.0109990240
-0.0273692844	-0.0016553016	0.0053885001	0.0084948821	-0.0221638752	0.0000899799	-0.0013868790	-0.0003194269	-0.0030056724
0.0595335998	0.0046893470	-0.0351656109	-0.0221638752	0.1712405408	-0.0001796800	0.0001556571	-0.0011726707	0.0393423395
-0.0002367796	-0.0000186260	0.0000504503	0.0000899799	-0.0001796800	0.0000010896	-0.0000149500	-0.0000123417	-0.0000446014
0.0060480259	0.0000859004	0.0006112401	-0.0013868790	0.0001556571	-0.0000149500	0.0011131480	-0.0000706053	-0.0032970883
-0.0039936493	0.0001694244	-0.0001810128	-0.0003194269	-0.0011726707	-0.0000123417	-0.0000706053	0.0009785672	0.0023791976
-0.0132486054	0.0015230991	-0.0109990240	-0.0030056724	0.0393423395	-0.0000446014	-0.0032970883	0.0023791976	0.0298166653

[18] H. Lü, D. Boilley, Y. Abe, and C. Shen, Phys. Rev. C **94**, 034616 (2016).

TABLE XIII. Covariance matrix between the parameters using regression (i).

$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
0.0006082522	-0.0025984253	-0.0020193653	0.0081677050	-0.0000304887	-0.0002948485	0.0011131294	0.0055410398
-0.0025984253	0.0199312841	0.0082768795	-0.0824713403	0.0000920849	0.0027833281	-0.0036521942	-0.0322768133
-0.0020193653	0.0082768795	0.0076229685	-0.0306027904	0.0001158646	-0.0002666849	-0.0043696777	-0.0176287206
0.0081677050	-0.0824713403	-0.0306027904	0.4000871096	-0.0002827051	-0.0051879688	0.0123767162	0.1200788307
-0.0000304887	0.0000920849	0.0001158646	-0.0002827051	0.0000019082	-0.0000097643	-0.0000681658	-0.0002287967
-0.0002948485	0.0027833281	-0.0002666849	-0.0051879688	-0.0000097643	0.0019854801	0.0004409849	-0.0050006166
0.0011131294	-0.0036521942	-0.0043696777	0.0123767162	-0.0000681658	0.0004409849	0.0035569094	0.0082877888
0.0055410398	-0.0322768133	-0.0176287206	0.1200788307	-0.0002287967	-0.0050006166	0.0082877888	0.0649826247

TABLE XIV. Covariance matrix between the parameters using regression (ii).

$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
0.0003819971	-0.0052531624	-0.0018352767	0.0277998983	-0.0000260665	0.0004993266	0.0009365340	0.0042188415
-0.0052531624	0.4034694831	0.0940782882	-2.3921066309	0.0001885700	-0.0699538700	-0.0617906364	-0.2389822932
-0.0018352767	0.0940782882	0.0262915121	-0.5825888711	0.0001329309	-0.0191423309	-0.0173332023	-0.0645221741
0.0277998983	-2.3921066309	-0.5825888711	14.5912280071	-0.0014604903	0.4572853381	0.3881599979	1.4663478275
-0.0000260665	0.0001885700	0.0001329309	-0.0014604903	0.0000030415	-0.0000600708	-0.0000636914	-0.0002641005
0.0004993266	-0.0699538700	-0.0191423309	0.4572853381	-0.0000600708	0.0169847741	0.0134145574	0.0472166671
0.0009365340	-0.0617906364	-0.0173332023	0.3881599979	-0.0000636914	0.0134145574	0.0121098120	0.0431063890
0.0042188415	-0.2389822932	-0.0645221741	1.4663478275	-0.0002641005	0.0472166671	0.0431063890	0.1604514629

TABLE XV. Covariance matrix between the parameters using regression (iii).

$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	$\hat{p}_7$	$\hat{p}_8$
0.0002444162	-0.0030966569	-0.0007898135	0.0138980286	-0.0000171969	-0.0000074066	0.0000744017	0.0015884903
-0.0030966569	0.2820927098	0.0756541571	-1.8124261615	0.0003224603	-0.0631928071	-0.0495966682	-0.1865929987
-0.0007898135	0.0756541571	0.0252066878	-0.5265266243	0.0001383686	-0.0225791476	-0.0176880512	-0.0607295316
0.0138980286	-1.8124261615	-0.5265266243	12.1680573525	-0.0021906101	0.4706768938	0.3653808853	1.2958749960
-0.0000171969	0.0003224603	0.0001383686	-0.0021906101	0.0000023828	-0.0000955509	-0.0000702254	-0.0002826310
-0.0000074066	-0.0631928071	-0.0225791476	0.4706768938	-0.0000955509	0.0227856426	0.0174302753	0.0548222841
0.0000744017	-0.0495966682	-0.0176880512	0.3653808853	-0.0000702254	0.0174302753	0.0136646136	0.0432448396
0.0015884903	-0.1865929987	-0.0607295316	1.2958749960	-0.0002826310	0.0548222841	0.0432448396	0.1484077708