

This is a repository copy of *An introduction to Bootstrap for nuclear physics*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/141437/>

Version: Accepted Version

---

**Article:**

Pastore, Alessandro [orcid.org/0000-0003-3354-6432](https://orcid.org/0000-0003-3354-6432) (2019) An introduction to Bootstrap for nuclear physics. *Journal of physics g-Nuclear and particle physics*. 052001. pp. 1-20. ISSN 0954-3899

<https://doi.org/10.1088/1361-6471/ab00ad>

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# An introduction to Bootstrap for nuclear physics

A. Pastore

Department of Physics, University of York, Heslington, York, YO10 5DD, UK

(Dated: January 20, 2019)

This guide aims at providing a general introduction to bootstrap methods. By using simple examples taken from nuclear physics, I discuss how such a method can be used to quantify error bars of an estimator. I also investigate the use of bootstrap to estimate parameters of a simple liquid drop model. In particular, I investigate how the method compares with standard techniques based on likelihood estimator and how to take into account correlations to better evaluate confidence interval of parameters.

## I. INTRODUCTION

In statistics, an *estimator* is a rule that applied to the sample data returns information on the underlying population. Common estimators are the mean and the variance.

Following Ref. [1], a *good* estimator should obey three criteria: consistency, efficiency and being unbiased.

- Consistent: in the limit of a sample size approaching the whole population, the estimator get closer and closer to the true value.
- Efficient: the variance of the estimator, *i.e.* the fluctuations around the true value, should be small.
- Unbiased: for any sample size  $N$ , the estimator should on average always produce an answer close to the true value.

For example, I can define the following estimators to determine the average height of a given population provided a sample of size  $N$ .

- $\hat{a}_1$ : add up all the heights and divide by  $N$ .
- $\hat{a}_2$ : select one random observable. Ignore the rest.
- $\hat{a}_3$ : add up all the heights and divide by  $N-1$ .

One notices that only estimator  $\hat{a}_1$  from the previous list respects all the three criteria at once.  $\hat{a}_2$  is not consistent, while  $\hat{a}_3$  is biased.

A common method used to define an estimator is the Maximum Likelihood Estimator (MLE). The MLE finds the set of parameters one wants to estimate that maximise the likelihood function given that specific data-set. The estimator built from MLE is usually *efficient* and *consistent*, but in most cases *biased* [1]. As a consequence it is important to evaluate the bias and thus use it to correct predictions.

An essential ingredient for the MLE construction is to assume an underlying distribution for the data of the sample. In most cases a Gaussian assumption is reasonable, but not necessary always the most appropriate. For example, in the case of a rare radioactive decay, a Poisson distribution would be more suitable in describing data than a Gaussian. A well-known estimator built from a Gaussian MLE is the least-square (LS) estimator. LS is widely used to determine the parameters of a given model using informations extracted from data. In Ref. [2], the authors have illustrated in much detail all the main features of the MLE and how to use it to derive errors. The latter are obtained by performing a second derivative of the likelihood function with respect to a given parameter. Although the procedure looks quite simple, in reality there are some important numerical issues in how performing numerical derivatives at good level of accuracy and at reasonable computational cost [3].

In this guide, I will discuss bootstrap method as a statistical tool to evaluate the bias and confidence intervals of statistical *estimators* [1] without the explicit use of a likelihood function. Efron in 1979 [4] introduced the basic idea behind bootstrap: given a data-set and a particular estimator; one builds new series of data-sets by resampling the original one. This defines a bootstrap sample. By applying the estimator at each of these bootstrap samples, one obtains the *empirical* distribution. The main hypothesis is that such a distribution is very close to the *true* one, so one can use it to obtain informations about the bias and the errors of the estimator. I refer to the vast available literature on the subject [4–9] to demonstrate the validity of such a statement. The procedure I described is usually referred to as Non-Parametric Bootstrap (NPB), to differentiate it from the case where some *prior* knowledge on the underlying distribution of a given estimator is given. One defines in this case a Parametric-Bootstrap (PB): the new data-sets are randomly extracted from the postulated distribution as done in Ref. [10] and not from the original data-set.

Bootstrap has been rarely used in theoretical nuclear physics [10–14], despite its simplicity and high potential. The aim of the present guide is to introduce the basic formalism and apply it to simple models in nuclear physics.

The guide is organised as follow: in Sec.II, I introduce the basic idea behind bootstrap and I use it to evaluate confidence intervals of the mean and of the Pearson coefficients in two different scenario. In Sec.III, I discuss NPB application to parameter estimates of a simple liquid drop model with and without explicit correlations in the data-set. I present my conclusions in Sec.IV.

## II. BOOTSTRAP: BASIC CONCEPTS

NPB is based on the simple assumption that any experimental data-set contains informations on its parent distribution. As a consequence, if the data-set is sufficiently large, it is possible to replace the original parent distribution via the empirical one obtained using Monte Carlo methods [4].

In a more precise way, let's assume I have a data-set formed by  $n$  independent vectors  $X = (x_1, x_2, \dots, x_n)$  and a real-valued estimator of the parameter  $\hat{\theta}$ . The origin of the data is not specified and thus they may be the result of a real experiment or of a simulation. I now resample them to create a series of new data-sets named  $X^*$  and I apply the estimator  $\hat{\theta}^*$ . For the sake of simplicity I consider  $n = 5$ . A possible outcome may look as

$$\begin{aligned} X_1^* &= (x_1, x_1, x_5, x_2, x_3) , \\ X_2^* &= (x_2, x_1, x_5, x_1, x_4) , \\ X_3^* &= (x_1, x_3, x_4, x_3, x_5) , \\ X_4^* &= (x_2, x_2, x_2, x_2, x_5) . \end{aligned}$$

Notice that the resampling allows for repetitions. By applying the estimator at each one of the four bootstrap samples,  $N_{\text{boot}} = 4$ , I obtain the experimental distribution. From the resulting histogram, I extract confidence intervals of  $\hat{\theta}$  by defining the 68% quantile [79]. The resampling assumes that the observables  $x_i$  are independent. Applying NPB to correlated data destroys correlations and may lead to wrong conclusions. I will discuss in Sec.III B an extension of the technique to take into account the case of strongly correlated data-sets.

Concerning the size of the original sample, there is a strong debate in the literature and I refer the reader to Ref. [9] for more details. The role of the size of the sample can be understood by considering the number of possible combinations (considering a permutation of the elements to be the same combination) one can build is

$$\binom{2n-1}{n} = \frac{(2n-1)!}{n!(n-1)!} . \quad (1)$$

For a small sample of  $n = 10$ , the number of possible bootstrap combinations is 92378. For  $n = 50$  the number of combinations grows to  $\approx 10^{30}$ . This number need to be as big as possible so that one can consider negligible the probability of drawing two times *exactly* the same sample and thus make sure all possible combinations can be extracted with equal probability. As a rule of thumb, one needs at least a sample of size 50, but several authors consider *safe* using NPB applied to sample sizes of  $\approx 10 - 20$  [7, 15]. In the following, I will make use of estimators that are not sensitive to the particular order of the observables in the data-set, as for example the mean. For such a reason I used Eq.1 to estimate the number of possible combinations. In the case of an estimator for which order matters, the number of possible sets is  $n^n$ .

The number of bootstrap samples should also be large to avoid introducing undesired bias. Using central limit theorem (for large samples), I assume that the difference between the empirical and real distribution follows a Gaussian distribution and the variance decreases as  $\approx 1/N_{\text{boot}}$ , where  $N_{\text{boot}}$  is the number of bootstrap samplings. In the present article, due to the low computational cost of Monte-Carlo sampling, I will use typically  $N_{\text{boot}} \approx 10^4$  and thus neglect the bias.

Before moving forward, it is important mentioning a technique very similar to NPB and named *jackknife* [16]. In this case instead of resampling the full data-set, one leaves out  $m$  observables at the time and then one uses the remaining  $N - m$  data to evaluate the estimator. Although it looks very similar to NPB, the mathematical hypothesis behind it are very different [17]. Jackknife is computationally less expensive than NPB, but it is specialised to obtain the variance of the estimator, while NPB gives access to the complete empirical distribution.

### A. Example: confidence interval of the mean

I illustrate the use of bootstrap by applying it to the well-known case of the estimator of the mean  $\hat{\mu}$  of a data-set  $\{X_1, X_2, \dots, X_N\}$  of length  $N$ . In this example, I will consider a sample of size  $N = 10$  extracted from an exponential distribution with mean  $\frac{1}{\lambda} = \frac{1}{2}$ . In Tab.II A, I report the values used for this example.

value	0.068	1.649	0.058	0.165	0.522	0.040	1.078	0.512	0.354	0.449
position	1	2	3	4	5	6	7	8	9	10

TABLE I: Random values extracted from exponential distribution with mean  $\frac{1}{\lambda} = \frac{1}{2}$ .

To calculate the mean of the parent distribution, I use the estimator

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N X_i . \quad (2)$$

In this particular case I obtain  $\hat{\mu} = 0.489$ . The estimator is *efficient*, *consistent* and *unbiased* [1], moreover for this particular case (knowing the distribution of the estimator) it is possible to calculate the resulting error bars using the formula

$$\sigma_M = \frac{\sigma}{\sqrt{N}} , \quad (3)$$

where  $\sigma$  is the root mean square of the data-set. I obtain  $\sigma_M = 0.154$ .

I now repeat the same analysis using NPB. In this case I perform a random sampling ( $N_{boot} = 10^4$ ) and I apply the estimator  $\mu^*$ . I group the results using an histogram as illustrated in Fig.1. On this figure, I also show the confidence interval of the NPB estimates by considering the lower (upper) quantiles defining 68% of the total counts.

The average value of the different means calculated via NPB is  $\bar{\mu}^* = 0.489_{-0.146}^{0.159}$ . As expected the values coincide since, it is known that the estimator in Eq.2 has no bias. The error bars are not symmetric (since I assume no prior Gaussian distribution), but they are extremely close to the estimated value, thus showing that NPB is a reliable method.

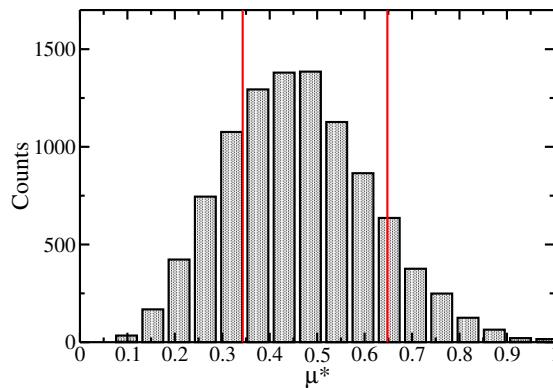


FIG. 1: (Colors online) Confidence interval of the mean: distribution of occurrences of the value  $\mu^*$  for the mean, Eq.2, obtained from various NPB samples. The vertical lines indicate the confidence interval defining 68% of the total counts.

### B. Example: confidence interval on Pearson coefficient

In this section, I consider another estimator, *i.e.* the Pearson coefficient [1]. The standard formulas to estimate errors of the Pearson coefficient are based on the Fisher transformation [18]. The latter is valid under the assumption

of normally distributed data, so in case of outliers it is possible to obtain inconsistent results.

As an example, I consider the possible correlation between the slope of the symmetry energy  $L_0$  and the neutron skin thickness  $\Delta r_{np}$  of heavy nuclei [19, 20].

In Fig.2, I show the neutron skin thickness in  $^{208}\text{Pb}$  as a function of  $L_0$  extracted from several nuclear models. The data-set is composed by 63 nuclear functionals grouped in 3 big families: zero-range Skyrme functionals [21]; finite-range Gogny [22] and Relativistic Lagrangians [23, 24]. For completeness, I report the models used in the calculation in Tab.V with all relevant infinite nuclear matter properties.

The neutron skin thickness is extracted from the full quantal calculation using two point Fermi function (2pF) as detailed in Ref. [25]. The two horizontal lines represent the region compatible with the most recent experimental measurement of neutron skin thickness in  $^{208}\text{Pb}$  [26] with its corresponding experimental error bars. To study the possible correlation between  $L_0$  and  $\Delta r_{np}$ , I calculate the Pearson coefficient as

$$\hat{r} = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sigma_X \sigma_Y}, \quad (4)$$

where  $\bar{X}, \bar{Y}$  represent the mean values and  $\sigma_X, \sigma_Y$  the root mean square of the sample.

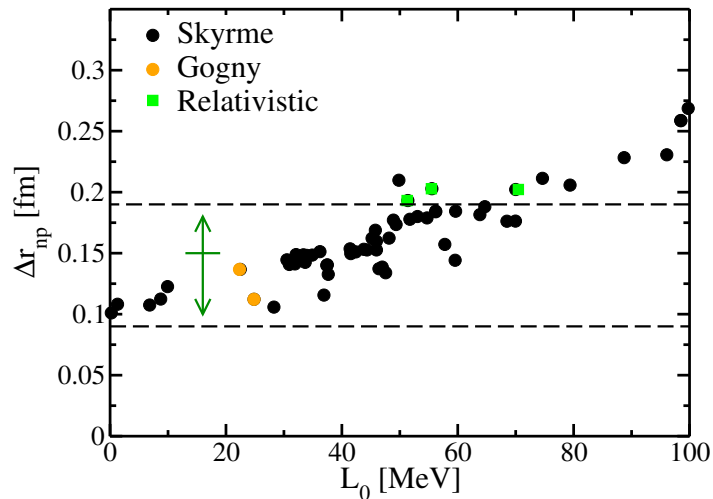


FIG. 2: (Colors online) Neutron skin thickness in  $^{208}\text{Pb}$  as a function of  $L_0$  for various nuclear models. The horizontal dashed line represent the experimental confidence interval as taken from Ref. [26]. See text for details.

For the present case I obtain  $\hat{r} = 0.91_{-0.05}^{0.03}$  where the error bars represent 95% confidence interval obtained using Fisher transformation. Performing NPB on the same data, I obtain a distribution of Pearson coefficient  $r^*$  as shown in Fig.3. The vertical lines represent the region including the 95% of the counts.

The average value of the correlation coefficient extracted from NPB is

$$\bar{r}^* = 0.92_{-0.04}^{0.04}. \quad (5)$$

The results are in excellent agreement between the two methods and with Refs [19, 20]. I conclude that the correlation in this data-set is very robust and very little dependent on the particular choice of the data used. This can be understood by looking at the distribution in Fig.3 and keeping in mind that *not* all data are used in every bootstrap resampling.

I repeat the previous analysis, but considering a different data-set to evaluate the impact of possible outliers when extracting statistical informations.

In Fig.4, I show the neutron skin thickness of  $^{208}\text{Pb}$  calculated with the same data-set reported in Tab.V, with an addition of extra 5 points (triangles). These points correspond to 5 Skyrme functionals with very large values of  $L_0$ :

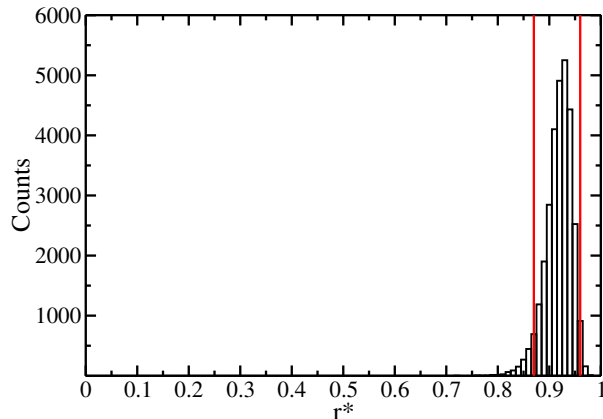


FIG. 3: (Colors online) Confidence interval of the Pearson coefficient: distribution of occurrences of the values of  $r^*$ , Eq.4, obtained using various bootstrap data sets in  $^{208}\text{Pb}$ . The vertical lines represent the region including 95% of the counts.

SEFM68 [27], SGOII [28], SKI1-2-5 [29]. These functionals produce a value of  $L_0$  which is way too large compared to the acceptable values of  $L_0$  [27]. In this guide, I assume for simplicity that each observables can be extracted with the same probability. Given the *prior* knowledge on the acceptable values of  $L_0$  [27], one may use instead the Bayesian version of bootstrap. I refer the reader to Refs. [30, 31] for more details on the method.

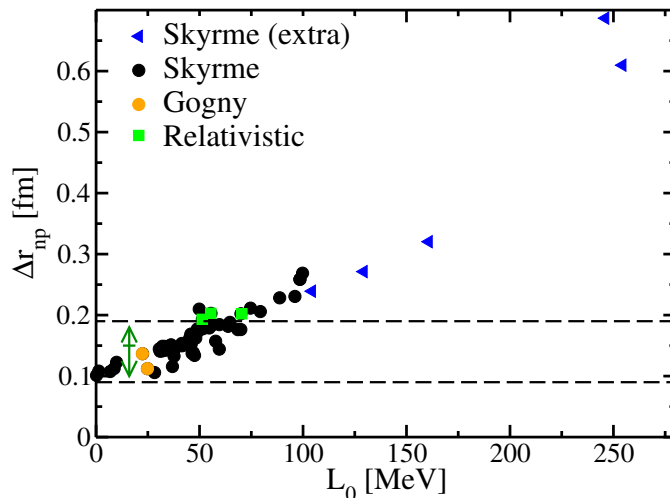


FIG. 4: (Colors online) Same as Fig.2, but with 5 additional Skyrme functional (triangles).

I performed the same NPB analysis as before in  $^{208}\text{Pb}$ , but for this extended data-set. I obtain from NPB the value  $\bar{r}^* = 0.96_{-0.03}^{+0.03}$ . The distribution of  $r^*$  resembles very much the one reported in Fig.3, although much more narrow. The extra points do not change the conclusions obtained using the data-set reported in Fig.2.

### C. Example: confidence interval on Pearson coefficient with outliers

I now consider a different close shell nucleus:  $^{100}\text{Sn}$ . In a previous study [13], the authors have shown that there is no real correlation within  $\Delta r_{np}$  and  $L_0$  in the case of a close shell nucleus with equal number of neutrons and protons, since the possible presence of a proton skin is mainly driven by Coulomb interaction more than by symmetry energy properties.

This case is quite suitable to illustrate how the presence of outliers may lead to wrong conclusions when using the Pearson coefficient and the Fisher transformation to estimate error bars. In Fig.5, I show the neutron skin thickness of  $^{100}\text{Sn}$  calculated with the same data-set reported in Tab.V, with an addition of extra 5 points (triangles). This is exactly the same pool of models used to produce Fig.4.

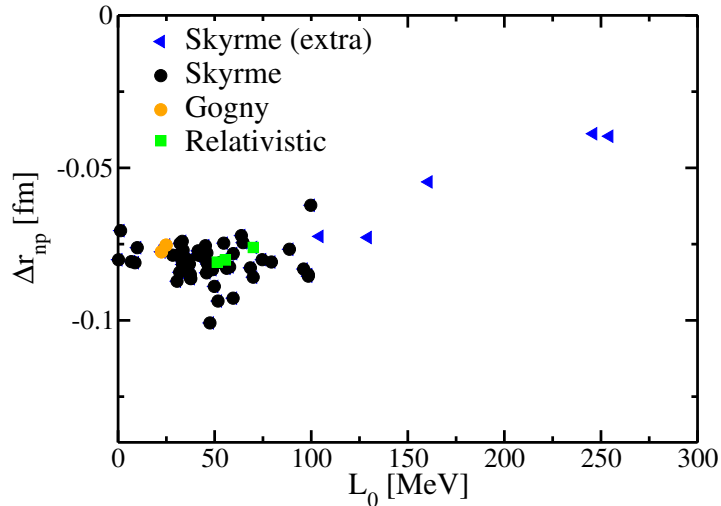


FIG. 5: (Colors online) Neutron skin thickness in  $^{100}\text{Sn}$  as a function of  $L_0$  for various models. See text for details.

The Pearson coefficient calculated with these extra 5 Skyrme functionals leads to  $\hat{r} = 0.68_{-0.15}^{0.11}$ . The correlation coefficient is smaller than the one obtained in  $^{208}\text{Pb}$ , but I can exclude with statistical significance the null-hypothesis of a non-correlated data-set.

In this case NPB may be very useful due to its random resampling nature: if the data are contaminated by few outliers, during random sampling most likely one will produce new data-sets with no outliers and thus getting strong variation on the estimator. Similar results would also be obtained using a much simpler *jackknife* method as illustrated in Ref. [4, 32]. Actually this is a quite common procedure used to validate how the conclusions of a statistical analysis depend on the particular choice of the data sample.

In Fig.6, I report the empirical distribution of the Pearson estimators  $r^*$  calculated using  $N_{boot} = 3 \times 10^4$  Monte-Carlo samplings. Contrary to Fig.3, the distribution appears to be much more broad. The calculated mean value with 95% confidence error bar is  $\bar{r}^* = 0.58_{-0.59}^{0.25}$ . Although the mean value of the estimator obtained with bootstrap is in good agreement with the direct calculation of the Pearson coefficient on the original data-set, I observe that the error bars are much larger. As a consequence, it is not possible to exclude the null-hypothesis of non-correlated data with a good statistical significance.

In this example, NPB is not used to identify which point is an outlier, but inspecting Fig.6, I can clearly observe a very large variance in the resulting estimator depending on the very specific choice of the observables of the data-set. This outcome should be used as a signal for a further investigation of the observables contained in the data-set.

In this specific case, I decided to exclude the extra 5 points as done in Ref. [13], by advocating the criteria on  $L_0$  given in Ref [27]. I obtain a result with a much more narrow distribution and centred on  $r^* = 0 \pm 0.3$ . It is important to stress the crucial role of error bars in this example. In both cases (with and without the 5 extra points) although the mean Pearson coefficient is very different, the error bars obtained via NPB lead always to the same conclusion: one can not exclude the null-hypothesis of non-correlated data.

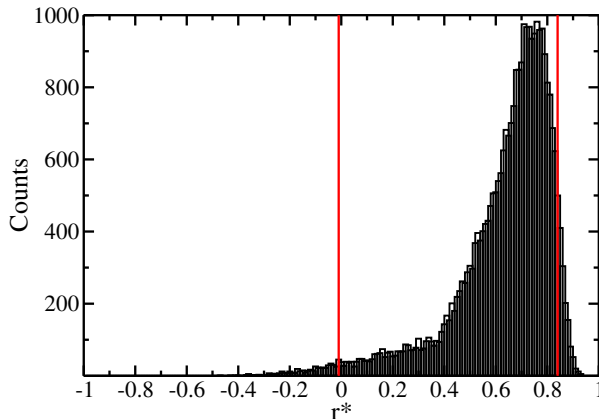


FIG. 6: (Colors online) Confidence interval of the Pearson coefficient: distribution of occurrences of the values of  $r^*$ , Eq.4, obtained using various bootstrap data sets in  $^{100}\text{Sn}$ . The vertical lines represent the region including 95% of the counts.

### III. REGRESSION ANALYSIS

Most of the models currently applied in theoretical nuclear physics rely on some adjustable parameters, usually obtained via LS minimisation [2]. In more general terms, the model function  $M(\mathbf{x}, \mathbf{p})$  depends on some input variables  $\mathbf{x}$  and parameters  $\mathbf{p}$ . This function is derived under some assumptions and it contains our knowledge of a particular (physical) phenomenon. The least-square estimator is used to determine the optimal set of parameters  $\mathbf{p}_0$  to maximise (minimise) the likelihood function of the parameters assuming an underlying Gaussian distribution.

Given the observable  $Y(\mathbf{x})$ , I have the following equation

$$Y(\mathbf{x}) = M(\mathbf{x}, \mathbf{p}_0) + \mathcal{E}(\mathbf{x}), \quad (6)$$

where  $\mathcal{E}(\mathbf{x})$  represents a residual error. The standard assumption is that the residuals are uncorrelated and they follow a normal distribution with zero mean.

In this article, I discuss how to use NPB to estimate the parameters  $\mathbf{p}_0$  and their error bars in both cases of non-correlated and strongly correlated residuals. The main advantage of NPB compared to the method presented in Ref. [2] is that having access to the empirical distribution, I can extract error bars on parameters directly from the histogram without performing numerical derivatives in parameter space [3, 33].

Since the present article aims at illustrating the methodology, I will consider the simple liquid drop (LD) model to calculate binding energies,  $B$ , for various nuclei as a function of neutron ( $N$ ) and proton ( $Z$ ) number. The binding energy is calculated as

$$B_{th}(N, Z) = a_v A - a_s A^{2/3} - a_c \frac{Z(Z-1)}{A^{1/3}} - a_a \frac{(N-Z)^2}{A} - \delta \frac{\text{mod}(Z, 2) + \text{mod}(N, 2) - 1}{A^{1/2}}, \quad (7)$$

where  $A = N + Z$ . The set of parameters I want to determine is  $\mathbf{p} = \{a_v, a_s, a_c, a_a, \delta\}$ . These parameters are named volume ( $a_v$ ), surface ( $a_s$ ), Coulomb ( $a_c$ ), asymmetry ( $a_a$ ) and pairing term ( $\delta$ ) and they refer to specific physical properties of the underlying nuclear system. I refer to Ref. [34], for a more detailed discussion on the physical meaning of these parameters. This model has a linear dependence on the parameters, so in principle, it is possible to perform least-square minimisation analytically by simple matrix inversion [1]. I will not use such a specific feature of the model, since I want to keep the discussion as general as possible. The experimental nuclear binding energies,  $B_{exp}(N, Z)$ , are extracted from Ref. [35]. In my calculation, I exclude all nuclei with  $A < 14$ , since Eq.7 is not suitable to describe very light systems. I define the penalty function as

$$\chi^2 = \sum_{N, Z \in \text{data-set}} \frac{[B_{exp}(N, Z) - B_{th}(N, Z)]^2}{\sigma^2(N, Z)}. \quad (8)$$



Parameter	[MeV]	Error [MeV]
$a_v$	15.69	$\pm 0.025$
$a_s$	17.75	$\pm 0.08$
$a_c$	0.713	$\pm 0.002$
$a_a$	23.16	$\pm 0.06$
$\delta$	11.8	$\pm 0.9$

TABLE II: Parameters of LD model obtained using least square fitting. On the last column I report the errors on the fitted parameters.

	$a_v$	$a_s$	$a_c$	$a_a$	$\delta$
$a_v$	1				
$a_s$	0.993	1			
$a_c$	0.984	0.960	1		
$a_a$	0.913	0.897	0.879	1	
$\delta$	0.036	0.035	0.038	0.036	1

TABLE III: Correlation matrix for LD parameters obtained from LS fitting.

Moreover, I will consider in my calculation only measured binding energies with an experimental error lower than 100 keV. So in first approximation I set  $\sigma^2(N, Z) = 1$ , thus assuming equal weight to all data.

The LS fit provides the values reported in Tab.II. Following Refs. [2, 36], the penalty function, Eq.8, normalised by the number of degrees of freedom should be equal to one at the minimum. This is not the case here, I have thus introduced a global scaling constant, usually named Birge factor [37], corresponding to the root mean square deviation of the residuals presented in Fig.7. The errors are obtained by taking the square root of the diagonal part of the variance matrix.

As already stated in Ref. [12], the errors obtained from LS under the assumption of non-correlated observables are quite unrealistic compared to the structure of the residuals  $\mathcal{E}(N, Z)$  shown in Fig.7.

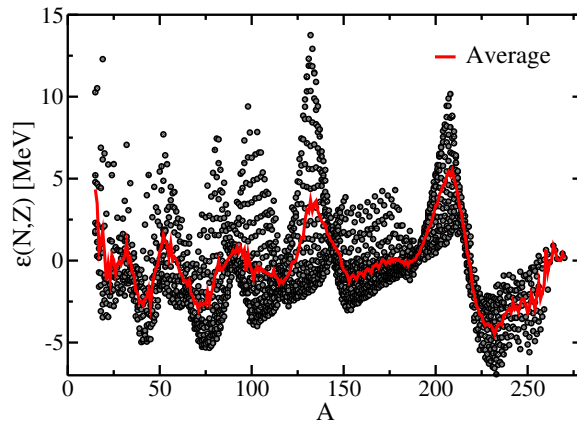


FIG. 7: (Colors online) Residuals  $\mathcal{E}(N, Z)$  for LD model obtained using least square minimisation. The solid line represents the average residual for fixed values of  $A$  as defined in Eq.10.

The variance of the residuals is quite large ( $\approx$  few MeV). I tested that the null hypothesis that the residuals are distributed according to a normal distribution is rejected at the 5% level based on the Kolmogorov-Smirnov test [1]. This means that there is a strong signal left in the residuals and Eq.6 is not valid. In this case, I conclude that the

model given in Eq.7 is not satisfactory and extra terms should be included to improve the fit [38].

### A. Uncorrelated Bootstrap

I now illustrate how to apply NPB to estimate the parameters of the model in Eq.7. The crucial aspect of NPB is to generate a new series of residuals  $\mathcal{E}^*(A)$  from random sampling the original one  $\mathcal{E}(A)$  and use them to generate a new set of observables (in this case the binding energy) as

$$B_{exp}^*(N, Z) = B_{exp}(N, Z) + \mathcal{E}^*(N, Z) , \quad (9)$$

and then use  $B_{exp}^*(N, Z)$  as a new input for a LS minimisation. This is only one possible variant of NPB, for more details I refer to Ref. [5]. By performing 5000 bootstrap samplings, I derived the empirical distribution of all parameters. I show in Fig.8 the two and one dimensional distribution of the parameters obtained with NPB. This figure is in excellent agreement with the marginalised likelihood obtained with Markov-Chain Monte Carlo in Ref.[39]. From the one dimensional histogram I extract the 68% quantile. For the first term of Eq.7, I obtain  $a_v = 15.69_{-0.023}^{0.024}$ , these values are in good agreement with the error bars provided in Tab.II. Similarly for the other parameters, the error bars obtained with LS and NPB are in agreement. This means NPB provides the same level of accuracy of a standard LS fitting as already spotted in Ref. [12].

The two-dimensional histograms inform us on possible correlations among parameters. Getting a cigar-shape along the diagonal of the figure means strong correlations, while a circle means low or no correlation. I observe that all LD parameters  $a_v, a_s, a_c, a_a$  are strongly correlated among each other, while the pairing term  $\delta$  is not. In this model,  $\delta$  is less constrained compared to other parameters as one can also see from the typical error bars obtained in Tab.II. These results can be compared with the one given in Tab.III. Using the data illustrated in Fig.8, I have been able to extract the full covariance matrix which is in perfect agreement with the one calculated using the Hessian matrix of the  $\chi^2$  [1]. This means I did not make use of any explicit numerical derivative in parameter space.

The covariance matrix it is not only useful to provide indication concerning error bars and correlations, but it can also be used to extract useful informations about the geometrical properties of the  $\chi^2$  surface around the minimum. NPB by construction explores the vicinity of the minimum of Eq.8, thus producing the corner plot shown in Fig.8. By analysing this figure, it is possible to assess qualitatively which directions in parameter space are more *stiff*. I refer to Refs. [40, 41] for a more detailed discussion on the topic.

### B. Block-Bootstrap

As seen in Fig.7, ignoring the autocorrelation between residuals is quite a major approximation. In this section, I illustrate how to correct the NPB to take into account correlations in a simple way.

As a first step, I observe that the variation in the direction N+Z is much more important than in the direction N-Z, so I define an *average* residual as

$$\mathcal{E}_A(A) = \frac{1}{N_A} \sum_{Z+N=A} \mathcal{E}(N, Z) . \quad (10)$$

The result is reported in Fig.7 as a solid line. The averaging is not washing out the signal in the data and it is used to simplify the bootstrap treatment. The information along the variation in the N-Z is kept as an isotopic-dependent variance function defined as

$$\sigma_A^2 = \frac{1}{N_A} \sum_{Z+N=A} (\mathcal{E}(N, Z) - \mathcal{E}_A(A))^2 . \quad (11)$$

I start determining the degree of correlation between data, by drawing a lag-plot. This is a well know technique to analyse correlation in time series [42]. The lag-plot of lag- $p$  consists in plotting the series against itself, but shifted by  $p$  units. In case of non-correlated residuals, the lag-plot should not show any pattern, *i.e* one should observe a cloud of points. In Fig. 9 (left panel), I show the lag-1 plot of the averaged residuals. They cluster around the diagonal thus showing a strong autocorrelation.

To quantify the correlation, I define a self-correlation coefficient as [42]

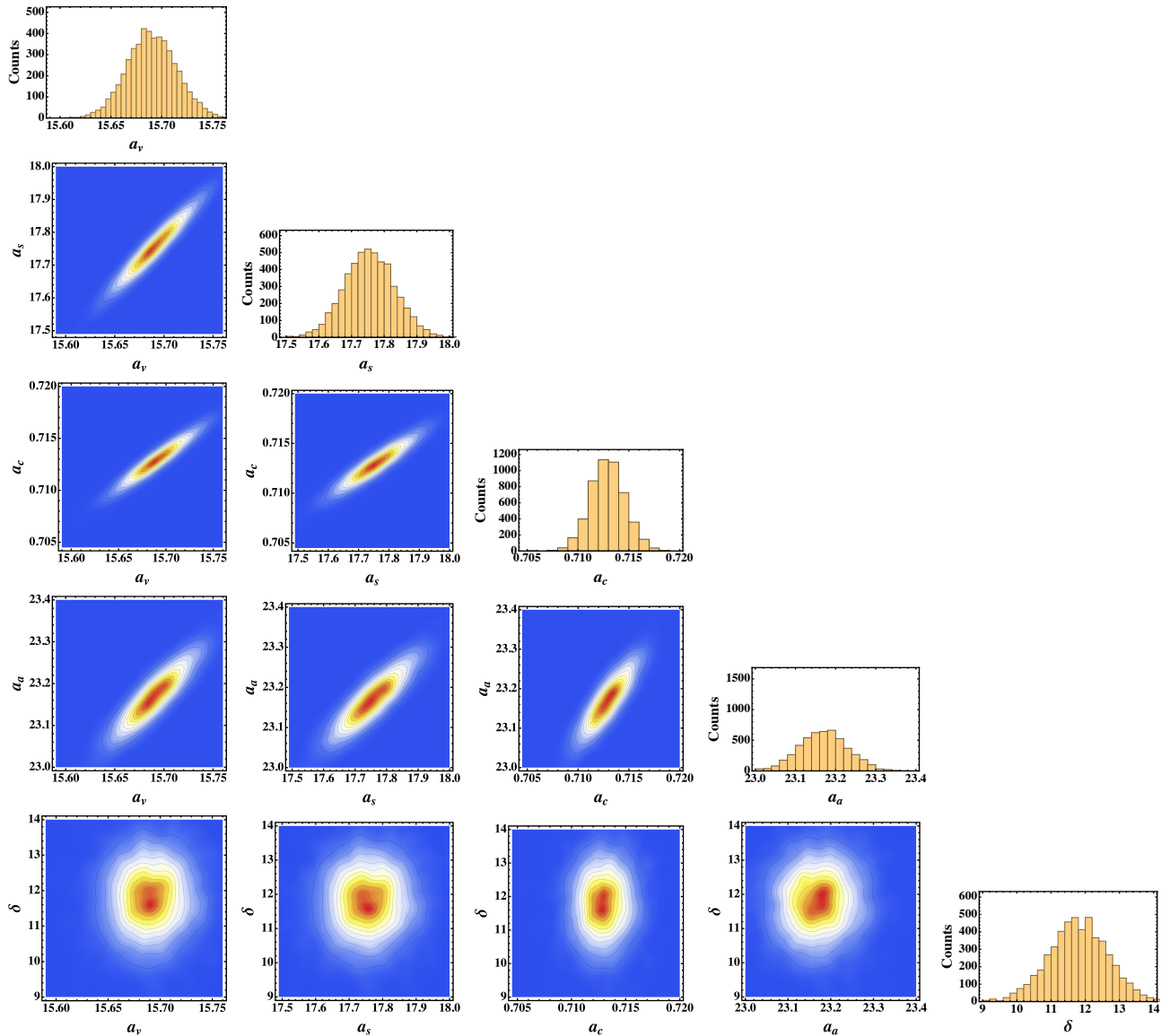


FIG. 8: One and two dimensional histograms of the LD model parameters obtained with NPB assuming uncorrelated residuals. The colour code of the 2-dimensional histogram is an heat-map. The total number of counts is fixed for all the plots  $N_{Boot} = 5000$ . See text for details.

$$\gamma(p) = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sigma_X \sigma_Y}, \quad (12)$$

where  $Y_i = X_{i-p}$  is the delayed series. In Fig.9 (right panel), I present the value of such a coefficients for different lags. I observe that there is a non-negligible correlation (i.e.  $\gamma(p) \geq 0.5$ ) up to  $p \approx 8$ ; this means that the residual of a nucleus  $A$  is strongly correlated with all other residuals within the interval  $[A - 8, A + 8]$ . This is in contradiction with the assumption made in Eq.8, where the  $\sigma^2$  matrix is assumed to be diagonal.

To take into account the presence of a signal in the residuals, one needs to introduce a minor modification to NPB. Following Ref. [43], instead of resampling individual residuals, I do re-sample blocks of fixed size to preserve the correlations between data. This procedure is named Block Bootstrap (BB).

Given a data-set composed by  $n$  elements  $\{X_1, X_2, \dots, X_n\}$ , I consider an integer  $l$  satisfying  $1 \leq l \leq n$ . I define  $\mathcal{B}_N$  overlapping blocks of length  $l$  as

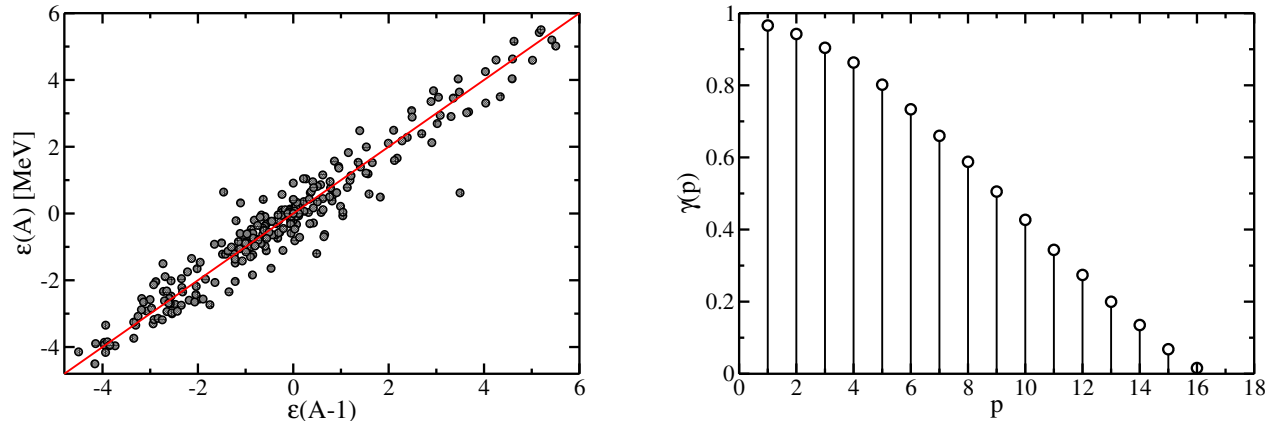


FIG. 9: (Colors online) Left panel: lag-plot of order 1 . Right panel: evolution of the self-correlation coefficient defined in Eq.12 for different values of the lag.

Parameter	[MeV]	Error [MeV]
$a_v$	15.70	$\pm 0.14$
$a_s$	17.82	$\pm 0.44$
$a_c$	0.713	$\pm 0.009$
$a_a$	23.20	$\pm 0.35$
$\delta$	11.78	$\pm 0.80$

TABLE IV: Parameters of LD models obtained using BB and  $l = 16$ .

$$\begin{aligned}
 \mathcal{B}_1 &= (X_1, X_2, \dots, X_l) \\
 \mathcal{B}_2 &= (X_2, X_3, \dots, X_{l+1}) \\
 \dots &= \dots \\
 \mathcal{B}_N &= (X_{n-l+1}, \dots, X_n)
 \end{aligned}$$

where  $N = n - l + 1$ . The uncorrelated version of bootstrap discussed in previous section is a particular case of BB when  $l = 1$  [80]. I now treat the different blocks  $\mathcal{B}_N$  as independent and I thus apply the standard bootstrap algorithm to them. Once the new residuals are produced, I restore the explicit  $(N, Z)$  dependence by adding an additional error extracted from a Gaussian distribution with variance  $\sigma^2(A)$  and defined in Eq.11. This is the same procedure followed in Ref. [12] in the case of Frequency Domain Bootstrap (FDB).

The choice of  $l$  is somehow arbitrary, but in the present case case, I use the lag-plot as indication. I observe that the data are correlated over an extension of at least 8 units. I have performed a test on the LD model using BB and various values of the length  $l$  of the block.

In Fig.10, I report the experimental distribution obtained using BB for the  $a_v$  parameter as a function of the size of the block.

By increasing the size of the blocks, I observe that the distributions gets wider and wider thus showing that I'm taking into account the effect of correlations. I see that  $l = 8$  captures quite well the correlations, and beyond  $l = 16$  the distribution does not change anymore. This is due to the fact that the size of the blocks is getting bigger than the size of the correlations. Such a result confirms the outcomes of the lag-plot shown in Fig.9.

From the distributions, I now extract the new error bars of the parameters using the 68% quantile definition. The results are reported in Tab.IV. For simplicity I defined the errors to be symmetric around the mean value. This is a minor approximation that has no effect on my conclusions.

I observe that the errors are more important using this method and they are most likely more realistic than the

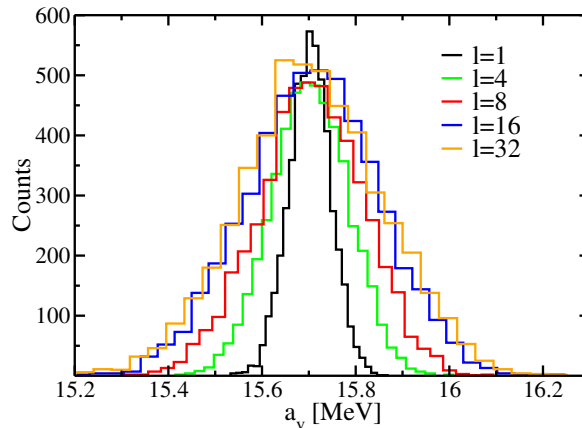


FIG. 10: (Colors online) Distribution of the  $a_v$  parameter obtained using BB and different length of the block. See text for details.

one obtained in Tab.II. The only exception is the error on  $\delta$  parameter which remains more or less constant. This is related to the fact that it is weakly correlated with the other terms. See Tab.III. The BB does not change the structure of correlations, the results of a corner plot for BB case are quite similar to the ones shown in Fig.8. These results are in good agreement with the one reported in Ref. [12] using FDB. Since the covariance matrix changes due to the effect of correlations, it is interesting to compare the effects on the error propagation for the LD model. For example the error on the binding energy of  $^{208}\text{Pb}$  when I use standard bootstrap is  $\pm 110$  keV, while when I use BB I obtain  $\pm 670$  keV, thus a factor of 6 larger. It would be now interesting to apply BB to more realistic models as the ones presented in Refs. [38, 44] to study how error propagation [45, 46] of model parameters is affected.

#### IV. CONCLUSIONS

In the present article, I have discussed in detail bootstrap methods and how they can be used to quantify uncertainties for statistical estimators. In particular, I have shown how the minimal hypothesis behind NPB provides a better understanding of the data-set under investigation. The NPB is consistent with findings based on other methods, but it has the remarkable advantage of being able to spot possible outliers in the data-set in a very simple way, without visual investigation of the data. This feature may have interesting applications in the case of automatised statistical analysis.

In Sec.III, I have shown how NPB can be used to fit parameters of a model, giving at the same time a reasonable estimation of errors and correlations without performing explicitly derivative in parameter space. Contrary to the standard parabolic approximation used to derive errors in a MLE method, the NPB explores the surroundings of the minimum in hyper-parameter space, thus having the possibility of finding possible defects in the hyper-surface.

To take properly into account correlations, I have also discussed a variant of NPB named Block Bootstrap. I have proved that such a method has the same quality of results of Frequency Domain Bootstrap [12], but having the advantage of not using an explicit Fourier transformation of the data and thus potentially being computationally faster. I have also highlighted that the main feature of BB and FDB is that they do not require an explicit modelling of correlations by introducing an *ad-hoc* covariance matrix into the fit.

#### Acknowledgments

This work has been supported by STFC Grant No. ST/P003885/1. I acknowledge the useful discussions with P. Becker who inspired this work. I would also thank the two anonymous referees for their precious comments.

	$\rho_0$ [fm <sup>-3</sup> ]	$J_0$ [MeV]	$L_0$ [MeV]		$\rho_0$ [fm <sup>-3</sup> ]	$J_0$ [MeV]	$L_0$ [MeV]
BSK16 [50]	0.159	30.00	34.88	SKA25S20 [27]	0.161	33.78	63.82
BSK18 [51]	0.159	30.00	36.22	SKB [52]	0.155	23.89	47.56
BSK19 [53]	0.160	30.00	31.9	SKM [54]	0.160	30.75	49.36
BSK20 [53]	0.160	30.00	37.39	SKM* [55]	0.160	30.03	45.78
BSK21 [53]	0.158	30.13	46.97	SKRA [56]	0.159	31.32	53.05
BSK22 [57]	0.158	32.00	68.49	SKS1 [58]	0.161	28.75	30.52
BSK23 [57]	0.158	31.00	57.77	SKS3 [58]	0.161	28.84	51.74
BSK24 [57]	0.158	30.00	46.40	SLYMR1 [59]	0.155	33.34	49.84
BSK25 [57]	0.159	29.00	36.90	SKSC10 [60]	0.161	28.11	0.21
BSK26 [57]	0.159	30.00	37.49	SKT [61]	0.148	24.90	28.27
F+ [62]	0.162	32.00	41.54	SKT1 [63]	0.161	32.02	56.19
F- [62]	0.162	32.00	43.79	SKT1A [27]	0.161	32.02	56.20
F0 [62]	0.162	32.00	42.41	SKT5 [63]	0.164	37.00	98.50
KDE [64]	0.164	31.97	41.42	SKT5A [27]	0.164	37.00	98.50
KDE0V [64]	0.161	32.98	45.21	SKT9 [63]	0.160	29.76	33.77
KDE0V1 [64]	0.165	34.58	54.69	SKT9A [27]	0.160	29.76	33.77
LNS1 [65]	0.162	29.91	30.93	SKX [66]	0.155	31.11	33.23
MSK1 [67]	0.157	30.00	33.92	SKXCE [66]	0.155	30.21	33.67
MSK2 [67]	0.157	30.00	33.35	SKXM [66]	0.159	31.20	32.09
MSK3 [67]	0.158	27.99	6.83	SLY230A [68]	0.160	31.99	44.33
NRAPR [69]	0.161	32.78	59.63	SLY230B [68]	0.160	32.01	45.97
RATP [70]	0.160	29.26	32.43	SLY4 [68]	0.160	32.00	45.96
SEFM074 [27]	0.160	33.40	88.71	SLY5 [68]	0.161	32.01	48.15
SEFM081 [27]	0.161	30.76	79.38	SQMC750	0.171	33.75	64.67
SEFM09 [27]	0.161	27.78	69.95	SV [71]	0.155	32.82	96.07
SEFM1 [27]	0.161	24.81	59.55	V100 [72]	0.157	28.00	8.73
SGII [73]	0.158	26.83	37.66	D1M [74]	0.164	28.50	24.85
SGOI [28]	0.168	45.20	99.76	D1S [22]	0.163	31.12	22.46
SI [75]	0.155	29.25	1.29	DD-PC1 [24]	0.152	33.00	70.00
SIH [76]	0.145	28.16	9.91	DD-ME1 [77]	0.152	33.06	55.53
SKA [52]	0.155	32.91	74.62	DD-ME2 [77]	0.152	32.31	51.39
SN2LO1 [78]	0.162	31.96	48.89				

TABLE V: Infinite matter properties of the data-set used in the current analysis.

### Appendix A: Infinite matter properties

I report here the explicit nuclear models used in my analysis together with their basic infinite nuclear matter properties. I refer the reader to Refs. [27, 47–49] for more details on the definition of infinite matter properties and how they relate to the parameters of the functionals.

- 
- [1] R.J.Barlow, *A Guide to the Use of Statistical Methods in the Physical Sciences* (John Wiley, 1989).
  - [2] J. Dobaczewski, W. Nazarewicz, and P. Reinhard, *Journal of Physics G: Nuclear and Particle Physics* **41**, 074001 (2014).
  - [3] X. Roca-Maza, N. Paar, and G. Colò, *Journal of Physics G: Nuclear and Particle Physics* **42**, 034033 (2015).
  - [4] B. Efron, *Annals of Statistics* **7**, 1 (1979).
  - [5] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap* (CRC press, 1994).
  - [6] A. C. Davison and D. V. Hinkley, *Bootstrap methods and their application*, vol. 1 (Cambridge university press, 1997).
  - [7] M. R. Chernick, W. González-Manteiga, R. M. Crujeiras, and E. B. Barrios, in *International encyclopedia of statistical science* (Springer, 2011), pp. 169–174.

- [8] B. F. Manly, vol. 70 (CRC press, 2006).
- [9] M. R. Chernick, *Bootstrap methods: a guide for practitioners and researchers*, vol. 1 (Wiley-Interscience, 2008).
- [10] R. N. Pérez, J. Amaro, and E. R. Arriola, *Physics Letters B* **738**, 155 (2014).
- [11] E. Nieves, J. and Ruiz Arriola, *The European Physical Journal A* **8**, 377 (2000), ISSN 1434-601X.
- [12] G. Bertsch and D. Bingham, *Physical review letters* **119**, 252501 (2017).
- [13] D. Muir, A. Pastore, J. Dobaczewski, and C. Barton, *Acta Physica Polonica B* **49** (2018).
- [14] B. Pasquini, P. Pedroni, and S. Sconfiatti, *Physical Review C* **98**, 015204 (2018).
- [15] N. I. Fisher and P. Hall, *Geophysical Journal International* **101**, 305 (1990).
- [16] R. G. Miller, *Biometrika* **61**, 1 (1974).
- [17] J. Tukey, *Ann. Math. Statist.* **29**, 614 (1958).
- [18] D. M. Corey, W. P. Dunlap, and M. J. Burke, *The Journal of general psychology* **125**, 245 (1998).
- [19] M. Centelles, X. Roca-Maza, X. Viñas, and M. Warda, *Phys. Rev. Lett.* **102**, 122502 (2009), URL <http://link.aps.org/doi/10.1103/PhysRevLett.102.122502>.
- [20] X. Roca-Maza, M. Centelles, X. Viñas, and M. Warda, *Phys. Rev. Lett.* **106**, 252501 (2011), URL <http://link.aps.org/doi/10.1103/PhysRevLett.106.252501>.
- [21] E. Perlińska, S. Rohoziński, J. Dobaczewski, and W. Nazarewicz, *Physical Review C* **69**, 014316 (2004).
- [22] J. Dechargé and D. Gogny, *Physical Review C* **21**, 1568 (1980).
- [23] M. Bender, P.-H. Heenen, and P.-G. Reinhard, *Reviews of Modern Physics* **75**, 121 (2003).
- [24] T. Nikšić, D. Vretenar, and P. Ring, *Physical Review C* **78**, 034318 (2008).
- [25] M. Warda, X. Viñas, X. Roca-Maza, and M. Centelles, *Phys. Rev. C* **81**, 054309 (2010), URL <http://link.aps.org/doi/10.1103/PhysRevC.81.054309>.
- [26] C. Tarbert, D. Watts, D. Glazier, P. Aguar, J. Ahrens, J. Annand, H. Arends, R. Beck, V. Bekrenev, B. Boillat, et al., *Physical review letters* **112**, 242502 (2014).
- [27] M. Dutra, O. Lourenço, J. S. Sá Martins, A. Delfino, J. R. Stone, and P. D. Stevenson, *Phys. Rev. C* **85**, 035201 (2012), URL <http://link.aps.org/doi/10.1103/PhysRevC.85.035201>.
- [28] Q.-b. Shen, Y.-l. Han, H.-r. Guo, et al., *Physical Review C* **80**, 024604 (2009).
- [29] P.-G. Reinhard and H. Flocard, *Nuclear Physics A* **584**, 467 (1995).
- [30] D. B. Rubin, *The annals of statistics* pp. 130–134 (1981).
- [31] D. D. Boos and J. F. Monahan, *Biometrika* **73**, 77 (1986).
- [32] B. Efron and C. Stein, *The Annals of Statistics* pp. 586–596 (1981).
- [33] R. Jodon, Ph.D. thesis, Université de Lyon (2014).
- [34] K. S. Krane and D. Halliday, *Introductory nuclear physics*, vol. 465 (Wiley New York, 1988).
- [35] M. Wang, G. Audi, A. Wapstra, F. Kondev, M. MacCormick, X. Xu, and B. Pfeiffer, *Chinese Physics C* **36**, 1603 (2012).
- [36] R. N. Pérez, J. Amaro, and E. R. Arriola, *Journal of Physics G: Nuclear and Particle Physics* **42**, 034013 (2015).
- [37] P. Klüpfel, P.-G. Reinhard, T. Bürvenich, and J. Maruhn, *Physical Review C* **79**, 034310 (2009).
- [38] K. Pomorski and J. Dudek, *Physical Review C* **67**, 044316 (2003).
- [39] M. Shelley, P. Becker, A. Gratton, and A. Pastore, arXiv preprint arXiv:1811.09130 (2018).
- [40] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna, *PLoS computational biology* **3**, e189 (2007).
- [41] T. Nikšić and D. Vretenar, *Physical Review C* **94**, 024333 (2016).
- [42] P. J. Brockwell and R. A. Davis, *Time series: theory and methods* (Springer Science & Business Media, 2013).
- [43] J.-P. Kreiss and S. N. Lahiri, in *Handbook of statistics* (Elsevier, 2012), vol. 30, pp. 3–26.
- [44] J. Dufflo and A. Zuker, *Physical Review C* **52**, R23 (1995).
- [45] J. Toivanen, J. Dobaczewski, M. Kortelainen, and K. Mizuyama, *Physical Review C* **78**, 034306 (2008).
- [46] J. McDonnell, N. Schunck, D. Higdon, J. Sarich, S. Wild, and W. Nazarewicz, *Physical review letters* **114**, 122501 (2015).
- [47] R. Sellahewa and A. Rios, *Physical Review C* **90**, 054327 (2014).
- [48] M. Dutra, O. Lourenço, S. Avancini, B. Carlson, A. Delfino, D. Menezes, C. Providência, S. Typel, and J. Stone, *Physical Review C* **90**, 055203 (2014).
- [49] D. Davesne, A. Pastore, and J. Navarro, *Astronomy & Astrophysics* **585**, A83 (2016).
- [50] N. Chamel, S. Goriely, and J. Pearson, *Nuclear Physics A* **812**, 72 (2008).
- [51] N. Chamel and S. Goriely, *Physical Review C* **82**, 045804 (2010), ISSN 0556-2813.
- [52] H. Köhler, *Nuclear Physics A* **258**, 301 (1976).
- [53] S. Goriely, N. Chamel, and J. M. Pearson, *Phys. Rev. C* **82**, 035804 (2010), URL <https://link.aps.org/doi/10.1103/PhysRevC.82.035804>.
- [54] H. Krivine, J. Treiner, and O. Bohigas, *Nuclear Physics A* **336**, 155 (1980).
- [55] J. Bartel, P. Quentin, M. Brack, C. Guet, and H.-B. Håkansson, *Nuclear Physics A* **386**, 79 (1982).
- [56] M. Rashdan, *Modern Physics Letters A* **15**, 1287 (2000).
- [57] S. Goriely, N. Chamel, and J. Pearson, *Physical Review C* **88**, 024308 (2013).
- [58] J. Gómez, C. Prieto, and J. Navarro, *Nuclear Physics A* **549**, 125 (1992).
- [59] J. Sadoudi, M. Bender, K. Bennaceur, D. Davesne, R. Jodon, and T. Duguet, *Physica Scripta* **2013**, 014013 (2013).
- [60] M. Onsi, H. Przywiecniak, and J. Pearson, *Physical Review C* **50**, 460 (1994).
- [61] C. Ko, H. Pauli, M. Brack, and G. Brown, *Nuclear Physics A* **236**, 269 (1974).
- [62] T. Lesinski, K. Bennaceur, T. Duguet, and J. Meyer, *Physical Review C* **74**, 044315 (2006).
- [63] F. Tondeur, M. Brack, M. Farine, and J. Pearson, *Nuclear Physics A* **420**, 297 (1984).

- [64] B. Agrawal, S. Shlomo, and V. K. Au, *Physical Review C* **72**, 014310 (2005).
- [65] D. Gambacurta, L. Li, G. Colo, U. Lombardo, N. Van Giai, and W. Zuo, *Physical Review C* **84**, 024301 (2011).
- [66] B. A. Brown, *Physical Review C* **58**, 220 (1998).
- [67] F. Tondeur, S. Goriely, J. Pearson, and M. Onsi, *Physical Review C* **62**, 024308 (2000).
- [68] E. Chabanat, P. Bonche, P. Haensel, J. Meyer, and R. Schaeffer, *Nuclear Physics A* **627**, 710 (1997).
- [69] A. W. Steiner, M. Prakash, J. M. Lattimer, and P. J. Ellis, *Physics reports* **411**, 325 (2005).
- [70] M. Rayet, M. Arnould, G. Paulus, and F. Tondeur, *Astronomy and Astrophysics* **116**, 183 (1982).
- [71] M. Beiner, H. Flocard, N. Van Giai, and P. Quentin, *Nuclear Physics A* **238**, 29 (1975).
- [72] J. Pearson and S. Goriely, *Physical Review C* **64**, 027301 (2001).
- [73] N. Van Giai and H. Sagawa, *Physics Letters B* **106**, 379 (1981).
- [74] S. Goriely, S. Hilaire, M. Girod, and S. Péru, *Physical review letters* **102**, 242501 (2009).
- [75] D. Vautherin, *Phys. Rev. C* **5**, 626 (1972).
- [76] M. Giannoni and P. Quentin, *Physical Review C* **21**, 2076 (1980).
- [77] G. Lalazisis, T. Nikšić, D. Vretenar, and P. Ring, *Physical Review C* **71**, 024312 (2005).
- [78] P. Becker, D. Davesne, J. Meyer, J. Navarro, and A. Pastore, *Physical Review C* **96**, 044330 (2017).
- [79] The value of the quantile is arbitrary, other choices are possible as 90%, 95%...
- [80] In this case there is no overlap between blocks and one should not use BB terminology for such a case.