Impact of statistical uncertainties on the composition of the outer crust of a neutron star

A. Pastore,[1, *] D. Neill,[1] H. Powell,[1] K. Medler,[1] and C. Barton[1, †]

[1]*Department of Physics, University of York, Heslington, York, Y010 5DD, UK*
(Dated: January 25, 2020)

By means of Monte Carlo methods, we perform a full error analysis on the Duflo-Zucker mass model. In particular, we study the presence of correlations in the residuals to obtain a more realistic estimate of the error bars on the predicted binding energies. To further reduce the discrepancies between model prediction and experimental data we also apply a Multilayer Perceptron Neural Network. We show that the root mean square of the model further reduces of roughly 40%. We then use the resulting models to predict the composition of the outer crust of a non accreting neutron star. We provide a first estimate of the impact of error propagation on the resulting equation of state of the system.

PACS numbers: 21.30.Fe 21.65.-f 21.65.Mn

I. INTRODUCTION

The study of neutron stars (NS) is important for our understanding of nuclear matter in the most extreme conditions. The recent measurements of gravitational waves [1] followed by optical detection of electromagnetic radiation [2] has opened an entire new way of studying such objects.

Due to the large pressure gradient, matter within NS arranges in layers with different properties. Going from the outside toward the inside (low density to high density), we find the *envelope*: a thin layer of iron atoms. At densities of $\rho_b \approx 10^4 \ g/cm^3$, the outer crust begins where a gas of (ultra) relativistic electrons surrounds ionised nuclei. Throughout the *outer crust* the Fermi energy of the electrons increases with density thus making it energetically favourable for electron capture to occur. As a consequence, we observe an increasing neutronisation of matter leading to the production of very neutron rich nuclei. The process continues until the drip-line is reached and neutrons start dripping off to form a uniform gas [3–9]. This region is called the *inner crust*. By going deeper inside the NS, the density continues increasing and eventually nuclei dissolve into a Fermi liquid, the *core*. The composition of the core is still matter of open debate and several models have been suggested over years [10–14].

In this article, we focus our attention on the sequence of nuclei forming the outer crust and how a full fledged error analysis impacts the prediction of the equation of state (EoS). To determine the sequence of nuclei which constitute the outer crust, we follow the methodology developed by Baym, Pethick and Sutherland (BPS) in 1971 [15]. For each input pressure value, one obtains the number of neutrons (N) and protons (Z) that minimises the Gibbs free energy for a given lattice configuration.

To perform these calculations a key ingredient is the nuclear binding energy. Of course, when experimental measurements are available there is no (or little uncertainty) in doing these calculation. The problem arises when one needs to use binding energies extrapolated from mass models. How reliable are these values? Only recently the scientific community has tried to answer such a question by performing detailed error analysis and error propagation on models [16–19] and use them for astrophysical purposes [20–23].

At present a very limited set of models are published with relevant statistical information to perform a complete error analysis [24–26], and none of these has been fully developed for astrophysical purposes. Of course, we may argue that a model which is very accurate in reproducing known masses may be very bad when used to extrapolate values and vice-versa. In this work, we decided to study one of the widely used mass-models currently used in astrophysical calculations, namely the Duflo-Zucker (DZ) mass model [27] and to perform a full statistical analysis.

The DZ model describes nuclear masses of known nuclei with a small standard deviation of ≈ 500 keV, in its 10 parameter version (DZ10) and ≈ 400 keV in its 33 parameter version (DZ33). The DZ represents an ideal case for our statistical investigations for two main reasons: the parameters have been adjusted only on one type of observable and it is computationally inexpensive. By means of Monte Carlo (MC) methods we have thus performed a full error

---

*Electronic address: alessandro.pastore@york.ac.uk
†Electronic address: charles.barton@york.ac.uk

propagation and studied the impact of these error bars on the resulting sequence of nuclei in the outer crust of a NS. To further improve the quality of the predictions, we equip the DZ with a Neural Network (NN) adjusted on mass residuals. We use such a combined model to predict the composition of the outer crust of NS.

The article is organised as follows: in Sec.II, the values and errors in the parameters of the Duflo-Zucker model are calculated and propagated into the binding energies calculated by the model. In Sec.III, we briefly present the main features of the NN and its application to the DZ model. The BPS method for calculating the composition of the outer crust is detailed in Sec.IV. In this section we present the impact of error bars on nuclear masses on the possible composition of the outer crust. Finally we present our conclusions in Sec.V.

II. DUFLO-ZUCKER MASS MODEL

The Duflo-Zucker [27] is a very successful macroscopic mass model based on a generalised liquid-drop (LD) plus the shell-model monopole Hamiltonian and It is used to obtain the binding energies of nuclei along the whole nuclear chart with high accuracy. Following Refs [28–30], we write the nuclear binding energy (B) for a given nucleus as a sum of ten terms as

$$B_{th} = a_1 V_C + a_2(M + S) - a_3 \frac{M}{\rho} - a_4 V_T + a_5 V_{TS} + a_6 s_3 - a_7 \frac{s_3}{\rho} + a_8 s_4 + a_9 d_4 + a_{10} V_P \ . \tag{1}$$

We defined $A = N + Z$; $2T = |N - Z|$ and $\rho = A^{1/3} \left[1 - \frac{1}{4}\left(\frac{T}{A}\right)^2\right]^2$. The ten different contributions can be grouped in two categories: in the first one we find terms similar to LD model as Coulomb (V_C), symmetry energy (V_T, V_{TS}) and pairing V_P. The other parameters originate from the averaging of shell-model Hamiltonian and they are based on the microscopic single level structure of the nucleus. For a more detailed discussion we refer to Ref. [28], where all these terms are described in great detail. Within the scientific literature, it is possible to find other variants of the DZ model containing 28 and 33 terms [30]. The DZ28 and DZ33 have a different structure compared to Eq.1 and we refer to Ref [27] for more details. In the present article, we will consider only the DZ10 model as given in Eq.1.

Following the methodology illustrated in Ref. [31], we fit DZ10 using the nuclear mass table of 2016 [32]. We define a penalty function of the form

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

In the fit, we have excluded all nuclei with $A < 14$; moreover we have considered only measured binding energies with an experimental error lower than 100 keV. Ignoring the small experimental uncertainties, we set $\sigma^2(N,Z) = 1$, thus assuming equal weights to all data. In total we considered 2293 nuclei. The optimal parameter set $\mathbf{a}^0 = \{a_1, a_2, a_3, \ldots, a_{10}\}$ is found using standard Nelder Mead method [33] and it is given in Tab.II. The DZ10 values are compatible with previous analysis illustrated in Ref. [28] and based on 1995 mass table [34] and more recent Ref. [35] based on 2012 mass table [36]. In the second and third column of the table, we report the errors on the parameters named type-1 and type-2 obtained by calculating the Hessian matrix of the χ^2 at its minimum and using the Block-Bootstrap method illustrated in Sec. II B.

The basic assumption used in fitting a model is that it satisfies the following equation

$$B_{exp}(N,Z) = B_{th}(N,Z|\mathbf{a}^0) + \varepsilon(N,Z) \ , \tag{3}$$

where $\varepsilon(N,Z)$ represents the residuals. In an ideal case, $\varepsilon(N,Z)$ follow a normal distribution with zero mean and variance σ^2.

In Fig.1, we report $\varepsilon(N,Z)$ as a function of the mass number A. The root mean square deviation is $\sigma = 0.572$ keV. Using the data shown in Fig.1, we have performed a simple Kolmogorov-Smirnov (KS) test to check whether or not the residuals follow a normal distribution. We obtained a very low p-value thus rejecting the null-hypothesis of normally distributed data [37]. Following Ref. [38], we define an average residual for each value of A as

$$\sigma_A^2 = \frac{1}{N_A} \sum_{Z+N=A} (\mathcal{E}(N,Z) - \mathcal{E}_A(A))^2 \ . \tag{4}$$

	Value	Error type-1	Error type-2
a_1	0.70454	0.00037	0.0010
a_2	17.740	0.0068	0.020
a_3	16.244	0.023	0.069
a_4	37.500	0.042	0.119
a_5	53.56	0.20	0.53
a_6	0.4573	0.0069	0.018
a_7	2.071	0.035	0.093
a_8	0.0210	0.0002	0.0004
a_9	41.43	0.21	0.66
a_{10}	6.162	0.088	0.103

TABLE I: Parameters of the DZ10 mass formula given in Eq.1. The quantities are expressed in MeV. The last two columns represent the error on the parameters neglecting/considering correlations among residuals. See text for details.
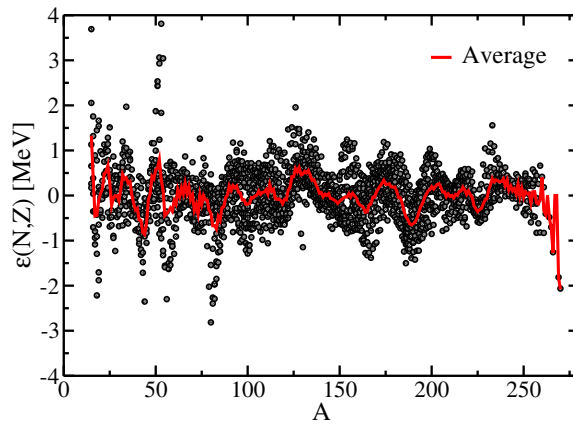

FIG. 1: Residuals of the DZ10 model. The solid line represents an average as given in Eq.4.
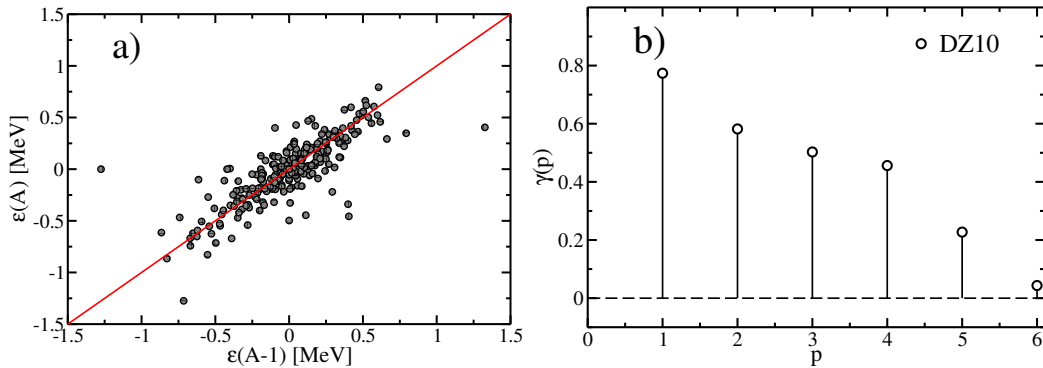
FIG. 2: (Colour online). Left panel: lag-plot of order 1 for DZ10 model. Right panel: evolution of the self-correlation coefficient defined in Eq.5 for different values of the lag. See text for details.

where $\mathcal{E}_A(A)$ is the average value of the residuals for the nuclei with mass A. The averaging is reported in Fig.1 as a solid line. Although we loose some information in the $(N - Z)$ direction, σ_A^2 still carries some useful information concerning the trend along the A direction of the nuclear chart.

We use the lag-plot of lag-p [39] to determine the degree of correlation between data. The lag-plot consists in

plotting the data against themselves, but shifted by p units. In case of non-correlated residuals, the lag-plot does not have any pattern, $i.e$ one should observe a cloud of points. In Fig. 2 (panel a), we show the lag-1 plot of the averaged residuals. They cluster around the diagonal, meaning a strong auto-correlation. To be more quantitative, we also define a self-correlation coefficient as

$$\gamma(p) = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sigma_X \sigma_Y} , \qquad (5)$$

where $Y_i = X_{i-p}$ is the delayed series. In Fig.2b, we present the values of $\gamma(p)$ for different lags. We observe that there is a non-negligible correlation (i.e. $\gamma(p) \geq 0.5$) for DZ10 up to $p \approx 4$; this means that the residual of a nucleus A is strongly correlated with all other residuals within the interval $[A - 4, A + 4]$.

These correlations should be taken into account to provide a more realistic estimate of the error of the parameters. See detailed discussion in Refs [31, 38]. A simple approach to perform a more consistent error analysis is based on Block Bootstrap (BB) [31] and it discussed in the following sections.

A.   Error propagation

Leaving aside for a moment the problem of correlations in the residuals as discussed in Fig.2, we now focus on the propagation of errors in the DZ model on the binding energies. For the present analysis, we will use the standard method based on a truncated Taylor expansion of χ^2 function around its minimum

$$\sigma_{th}^2 = \mathcal{G}\mathcal{C}\mathcal{G}^T , \qquad (6)$$

where $\mathcal{G}$ is the vector containing the derivative of the model in parameter space

$$\mathcal{G}_j = \frac{\partial B_{th}}{\partial a_j} . \qquad (7)$$

The latter are done numerically using the prescription given in Ref. [25]. The covariance matrix $\mathcal{C}$ is obtained by evaluating the Hessian matrix of Eq.2 at its minimum. We name this method of evaluating errors type-1 to distinguish from the Block Bootstrap one, see Tab.II. Having access to errors, we compare how the DZ10 performs against experimental data. We expect that 68% of known masses differ from the model prediction no more than $\sigma = \sigma_{th} + \sigma_{exp}$, where σ_{exp} is the experimental error reported in Ref. [32]. By increasing the error bar by a factor of 2 and 3 we should obtain 95% and 99.7% of experimental binding energies falling into the interval. In Tab.II A, we have reported the actual percentage of nuclei using type-1 errors falling within the error band for the full nuclear chart and some different mass regions.

type-1 errors			
	1σ	2σ	3σ
Full chart	13.6%	27.2%	39.5%
$50 \leq A < 150$	14.7%	26.8%	37.2%
$20 \leq Z \leq 50$	11.5 %	22.2%	31.4%
$A \geq 150$	14.8%	30.8 %	45.8%

TABLE II: Percentage of nuclei included in the total error bars in different sectors of the nuclear chart. See text for details.

Focusing on the nuclei relevant for the NS crust composition, $i.e.$ $20 \leq Z \leq 50$, we observe that only 11.5% of measured nuclei fall within the error bars of DZ model. This means that either Eq.3 is not valid and one should improve the model or the estimate of the error bars is not done properly. In the next section we discuss how to improve error bars taking into account residual correlations, while in Sec.III, we discuss how to improve the model using a neural network.

B. Block Bootstrap

Bootstrap is a simple Monte Carlo procedure commonly used to estimate error bars of various statistical estimators [40–43]. Firstly introduced by Efron in the 70s [44], it has become a very common statistical tool in several scientific domains [45–50]

The crucial aspect of non-parametric bootstrap (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

$$B^*(N, Z) = B_{exp}(N, Z) + \mathcal{E}^*(N, Z) . \tag{8}$$

$B^*(N, Z)$ is used as input for a least square minimisation. As discussed in Ref. [31], NPB gives access to the full likelihood of the model from which one can easily extract confidence intervals on the estimated parameters.

Another important advantage of Bootstrap is the possibility of removing the usual approximation of Eq.6 and to use a full MC error propagation instead. The latter is a side product of any NPB procedure since effectively, we produce several DZ fits and thus we can infer the error on each predicted mass by simply defining the 68% quantile on each mass distribution. In the current case, we did not find any noticeable difference between the full MC error propagation and simpler approximation provided in Eq. 6.

A crucial step to improve error bar estimates is to take into account correlations in the residuals. A possible way would be to model a non diagonal variance matrix entering the minimisation procedure of the χ^2 function given in Eq.2. In the present article, we stick to Bootstrap methods, but using the Block-Bootstrap (BB) variant [51]. The data are separated in overlapping blocks to which we apply the Monte Carlo procedure discussed before. In such a way, we preserve the correlations among residuals of neighbouring nuclei.

Our data set presents correlations along the A and $(N - Z)$ direction. In Refs [31, 38], the length of correlations in the A directions was more important that the $(N - Z)$ one; as a consequence it has been possible to simplify the problem by taking the average value and correct it by a Gaussian error extracted from the isotopic dependent variance defined in Eq.4. From Fig.2, we observe that the length of correlations along the A direction is quite short, few units, thus the approximation done before in Ref. [31, 38] is not *a priori* justified. We have thus decided to develop a more consistent framework for error analysis using a 2-dimensional BB. The main issue here is that the residuals represent an irregular grid. We have created square blocks of size l around each nucleus using the following procedure: when less than half terms (l^2) are empty, since there are no data, we discard such a block. Otherwise we keep it. Finally we inspect in all retained blocks for the presence of missing entries, these are replaced by a random number extracted from a normal distribution with the variance equal to the one of the DZ10 model.

The choice of l is somehow arbitrary, but in the present case, by using the lag-plot shown in Fig.2, we can assess that a block size of $l = 8$ will take into account all correlations. In Fig.3, we illustrate the histogram of the a_5 parameter as a function of the block size for $N_{boot} = 5000$ Monte Carlo samplings. We see that beyond $l = 8$ the increase in the size of the blocks does not lead to any major change in the distribution of the values of a_5. As discussed in Refs [31, 38], correlations do not change the mean value of the distribution, but they modify the width. As a consequence the resulting error bars become larger.

From the marginalised likelihood, we extract the full correlation matrix $\mathcal{R}$, defined as

$$\mathcal{R}_{ij} = \frac{\mathcal{C}_{ij}}{\sqrt{\mathcal{C}_{ii}\mathcal{C}_{jj}}} . \tag{9}$$

By construction, the matrix elements of $\mathcal{R}$ are normalised between -1 and 1. The resulting matrix for the DZ10 model is reported in Fig.4. We observe a strong correlation between the two *master* terms a_2, a_3 and Coulomb coupling a_1. We also notice a degree of correlation between these three terms and the two related to symmetry energy a_4, a_5. The coupling constants a_6, a_7 show a very strong correlation between them. Most likely it could be possible to express a_7 as a function of a_6 and remove one of these parameters [52]. Finally, the a_8, a_9 terms are mildly correlated to each other and completely uncorrelated to the other surface terms. The last term a_{10} takes into account the pairing effects. It is completely uncorrelated to all terms, thus showing that pairing correlations within DZ10 are absorbed only by a_{10}.

We repeat the comparison between theoretical and experimental values using the new error bars. The results are reported in Tab.II B. We see that the current estimate perform much better than the one based on type-1 errors. The percentage of nuclei falling into the error bars is not still the expected value (*i.e.* 68% of counts for $1\,\sigma$ error), although the error bars are probably more realistic than in the previous case. Given the discrepancy of the results observed in Tab.II B, we can conclude, also observing Fig.1, that to improve the reproduction of the data, we need to act on the model itself. To this purpose we employ a Neural Network.
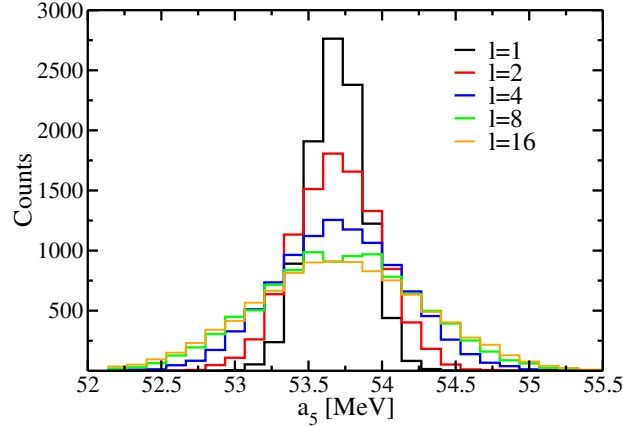
FIG. 3: (Colour online). Histogram of the a_5 parameter, Eq.1, as a function of the block size. See text for details.
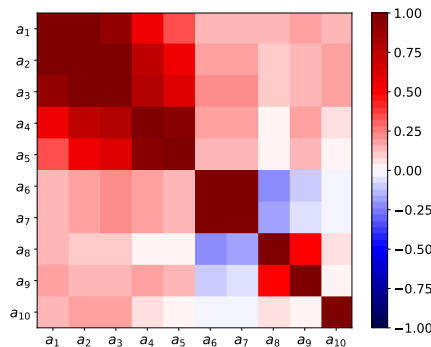

FIG. 4: (Colour online).Correlation matrix obtained using BB. See text for details.

	type-2 errors		
	1σ	2σ	3σ
Full chart	34.5%	60.4%	77.9%
$50 \leq A \leq 150$	31.8%	55.5%	74.2%
$20 \leq Z \leq 50$	27.9 %	52.8%	71.9%
$A > 150$	39.9%	69.4 %	85.6%

TABLE III: Percentage of nuclei included in the total error bars using BB error bars for different sectors of the nuclear chart.

III.   MULTILAYER PERCEPTRON

To capture the fluctuating function needed to accurately predict the binding energy of nuclei, we used a class of feedforward artificial neural network (NN) called Multilayer Perceptron (MLP) [53, 54]. A NN is a system of connected algorithms, called nodes, that are designed to mimic the working of a biological brain. These nodes are arranged into layers and we allow for connections only between nodes of two sequential layers, with the input nodes having no node connected behind them and the output nodes not connecting to any further nodes.

Following the work of Ref [20, 35], we built a NN composed by two layers densely connected having 45 and 84 nodes each. We use a swish function [55] as activation and stochastic gradient method to determine the values of

the weights of each neural connection. Several architecture of the NN are possible, but from our analysis we did not obtain any significant improvement.

To train the NN, we use part of the residuals shown in Fig.1. To avoid over-fitting, we split the data set in two: a *training* set, composed by 2/3 of the data and a *validation* set, composed by the remaining 1/3.The splitting is done randomly. From the training set we still remove randomly 10% of the data and we use them as *control* to validate at each epoch the actual improvement of the loss function [53]. Notice that the *validation* set is never used to train the network.

Having trained the NN, we then replace in Eq.3, the new theoretical model composed by the DZ10 plus the NN. We can assess the performance of such a model by comparing with the nuclear masses. We obtain a root mean square deviation of $\sigma_{DZ+NN} = 0.344$ keV. This means that the NN has been able to grasp a signal in the residual and thus further improve the agreement with experimental data. Such a result is in agreement with previous findings [21, 56]. To check the predictive power of our NN, we also calculate the root mean square deviation of the DZ+NN on the validation and training set only. In this case we obtain $\sigma_{DZ+NN}^{validation} = 0.358$ keV and $\sigma_{DZ+NN}^{training} = 0.324$ keV. Such a test is important to prove that the NN has clearly grasped some residual signal in the whole data and not only fitted the training set.
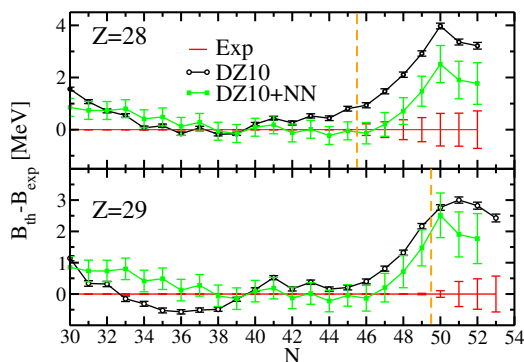

FIG. 5: (Colour online) In the two panel we compare the evolution of the difference between the theoretical and experimental binding energies obtained using only a DZ10 model (open dots) and a DZ10 plus NN (full squares). On the horizontal line we also report the experimental error on measured masses. The vertical dashed line separates masses that have been directly measured from estimated ones.

In Fig.5, we illustrate the evolution of the difference between the theoretical binding energies and the experimental one [32, 57] obtained using the DZ10 (open dots) and DZ10+NN(full squares). On the figure, we also provide the theoretical and experimental error bar. The typical error bars on masses are not visible on the energy scale, apart from the estimated values of masses. These values are separated from the measured one by a vertical dashed line. The theoretical bar associated with the DZ10 has been obtained using BB method, while for the DZ10+NN model, we have defined the error bar by adding in quadrature the error of the original DZ10 model (type-2) and the variance of the NN over the nuclear chart. This is not the optimal way to estimate the error and a more sophisticated analysis is necessary [58]. In general, we notice that the inclusion of NN reduces the discrepancy between the theory and the nuclear masses, thus showing that the DZ10 is missing some physical effect. Further studies are still necessary to draw solid conclusions, but we have notice that even using more powerful NN it is hard to further reduce such a discrepancy [59]. This may be explained in terms of chaotic motion inside the nucleus as discussed in Ref.[60].

IV. OUTER CRUST

To determine the chemical composition of the outer crust, we follow the methodology illustrated in Ref. [61], *i.e.* we determine the optimal combination of neutrons and protons that minimises the baryonic chemical potential μ_b at a given pressure P

$$\mu_b = \frac{W_N + \frac{4}{3}W_L + Z\mu_e}{A} \ . \tag{10}$$

The three terms appearing in this formula are the lattice energy W_L, the nuclear energy W_N and the electron chemical potential μ_e. The minimisation is done assuming charge neutrality and β-stable matter. At zero temperature the

Gibbs free energy is related to the total energy of the system as $\mu_b = G/A$.

Following Ref.[15], the outer crust is formed by a body-centre-cubic (BCC) lattice of nuclei surrounded by free electrons [62]. The energy contribution of the lattice reads

$$W_L = -1.819620 \frac{Z^2 e^2}{4\pi\varepsilon_0 a}, \tag{11}$$

where a is the lattice constant which is given by $n_N a^3 = 2$. The density of the nuclei is related to the electronic density by the relation

$$n_N = \frac{n_e}{Z}. \tag{12}$$

The energy contribution coming from finite nuclei reads

$$W_N = m_n(A - Z) + m_p Z - B(N, Z). \tag{13}$$

Finite nuclei do not contribute to the total pressure, which is just the sum of the lattice P_L and electron contribution P_e

$$P = P_e + P_L \tag{14}$$
$$= n_e \mu_e - E_e + \frac{1}{3} W_L \frac{n_e}{Z}, \tag{15}$$

where $\mu_e = \frac{\partial E_e}{\partial n_e}$ and E_e is the total energy of the electrons. In our case, this is just the kinetic term

$$E_e = \frac{m_e^4}{8\pi^2 n_e} \left[x_F y_F (x_F^2 + y_F^2) - \ln(x_F + y_F) \right], \tag{16}$$

where $x_F = p_{Fe}/m_e$ and $y_F = \sqrt{1 + x_F^2}$.

The full minimisation using the DZ mass-model leads to the result shown in Tab.IV. When the value of the binding energy is known experimentally, we read such a value from AME2016 mass table [32], while if it is estimated or absent we use the DZ10 model to produce the binding energy. As already observed by other authors [12, 15, 63, 64] we see that the outer crust of a non-accreting neutron star is composed (almost independently of the model used) by semi-magic nuclei. From Tab.IV, we observe that at low pressure region, nuclei with N=50 are favoured and at $P \approx 10^{-4}$ MeVfm^{-3} the N=82 nuclei appear.

P_{max} [MeVfm^{-3}]	ρ_b^{max} [fm^{-3}]	N	Z
$3.30 \cdot 10^{-10}$	$4.67 \cdot 10^{-6}$	30	26
$4.36 \cdot 10^{-8}$	$1.52 \cdot 10^{-4}$	34	28
$3.56 \cdot 10^{-7}$	$7.46 \cdot 10^{-4}$	36	28
$4.02 \cdot 10^{-7}$	$8.57 \cdot 10^{-4}$	38	28
$1.03 \cdot 10^{-6}$	$1.73 \cdot 10^{-3}$	50	36
$5.59 \cdot 10^{-6}$	$6.56 \cdot 10^{-3}$	50	34
$1.76 \cdot 10^{-5}$	$1.54 \cdot 10^{-2}$	50	32
$1.58 \cdot 10^{-4}$	$8.94 \cdot 10^{-2}$	50	28
$1.82 \cdot 10^{-4}$	$1.03 \cdot 10^{-1}$	82	42
$3.31 \cdot 10^{-4}$	$1.73 \cdot 10^{-1}$	82	40
$4.83 \cdot 10^{-4}$	$2.29 \cdot 10^{-1}$	82	38
$4.86 \cdot 10^{-4}$	$2.39 \cdot 10^{-1}$	82	36

TABLE IV: Composition of the outer crust of a NS using the DZ mass model. In the first two columns we report the maximum value of pressure and baryonic density at which the nucleus is found using the minimisation procedure. The horizontal line separates the measured mass reported in AME2016 [32].

From Tab.IV, we see that thanks to major experimental effort several nuclei in the N=50 region relevant for the NS have been measured with extreme accuracy [57], but still the masses of nuclei in the region N=82 need to be

P_{max} [MeVfm^{-3}]	ρ_b^{max} [fm^{-3}]	N	Z
$3.30 \cdot 10^{-10}$	$4.67 \cdot 10^{-6}$	30	26
$4.36 \cdot 10^{-8}$	$1.52 \cdot 10^{-4}$	34	28
$3.56 \cdot 10^{-7}$	$7.46 \cdot 10^{-4}$	36	28
$4.02 \cdot 10^{-7}$	$8.57 \cdot 10^{-4}$	38	28
$1.03 \cdot 10^{-6}$	$1.73 \cdot 10^{-3}$	50	36
$5.59 \cdot 10^{-6}$	$6.56 \cdot 10^{-3}$	50	34
$1.76 \cdot 10^{-5}$	$1.54 \cdot 10^{-2}$	50	32
$2.67 \cdot 10^{-5}$	$2.19 \cdot 10^{-2}$	50	30
$1.45 \cdot 10^{-4}$	$8.15 \cdot 10^{-2}$	50	28
$1.84 \cdot 10^{-4}$	$1.04 \cdot 10^{-1}$	82	42
$3.33 \cdot 10^{-4}$	$1.68 \cdot 10^{-1}$	82	40
$4.83 \cdot 10^{-4}$	$2.29 \cdot 10^{-1}$	82	38

TABLE V: Composition of the outer crust of a NS using the DZ10+NN mass model. In the first two columns we report the maximum value of pressure and baryonic density at which the nucleus is found using the minimisation procedure. The horizontal line separates the measured mass reported in AME2016 [32].

predicted using some particular mass model [65]. Following the discussion in Sec.III, we have also performed a full minimisation of the outer crust using the DZ10+NN. The result is reported in Tab.IV

The inclusion of NN in the mass models has an impact over the general composition. It affects both the threshold pressures and the observed sequence of nuclei in the crust. The major difference between the results reported in Tab.IV and IV is the absence (presence) of ^{80}Zr and the presence (absence) of ^{118}Kr at the transition between outer and inner crust.

A.   Error propagation in the outer crust

Having studied in detail the error propagation during the fitting procedure, we now assess the impact of such errors on the composition of the outer crust. The case of type-1 errors have been already presented in Ref. [66], here we apply the same methodology to the case of type-2 errors only. To this purpose we build 10^4 new mass tables [67] defined as

$$B_{th}^*(N, Z) = B_{th}(N, Z) + \mathcal{N}(0, \sigma(N, Z)) , \tag{17}$$

where $B_{th}(N, Z)$ is the value predicted by DZ10 model and $\mathcal{N}(0, \sigma(N, Z))$ represents a random normal error using as a variance the error on masses discussed in Sec.II B. When a mass is known experimentally, we use the experimental value. When the mass is not know or estimated, we use the DZ10 model. Having access to such a large number of samples, we define an *existence* probability for each nucleus as a function of pressure within the DZ model. The probability is defined in the usual *frequentist* approach [37] as the number of successful event over the total number of events. In Fig.6 (a), we present the probability of appearance of a given nucleus as a function of the pressure of the star. The explicit inclusion of error bars on nuclear masses does not change the main sequence reported in Tab.IV, but we observe the transitional region between N=50 and N=82 are no more sharp. Another interesting aspect is that ^{80}Zr is not included in the main sequence obtained with DZ10, but when considering error bars, we have a probability of 60% of observing it.

Using the same data set, we can define in a similar way a statistical uncertainty for the EoS: by counting the 10^4 EoS built before, we define the 68%, 95% and 99% quantile of the counts, *i.e*, one, two or three σ deviation. The result is reported in Fig.6 (b). We interpret the bands as statical uncertainties propagated from our model. As expected, we observe quite large uncertainties corresponding to the densities at which the chemical composition changes. By using these error bars, it is now possible to compare the EoS obtained with various models and check if the results are compatible or not within the statistical errors.

Finally, it would be interesting to observe how such uncertainties impact other observables on NS as for example radii [68]. We leave such an aspect for future investigations.
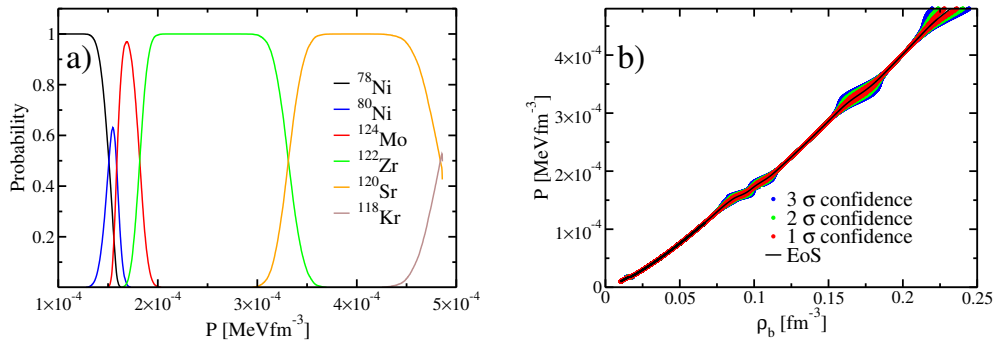
FIG. 6: (Colour online) Panel a: existence probability of a given nucleus within the outer crust as a function of the pressure. On panel b, the equation of state of the system including statistical uncertainties. See text for details.

V.   CONCLUSIONS

Astrophysical models strongly rely on nuclear data and even more on estimates based on theoretical models. In the current article, we have performed a full statistical analysis of a widely used mass model DZ10. We have shown that the standard approximation of adopting independent residuals is not correct, moreover to assess the quality of extrapolation based on the global variance of the model it is only a crude approximation. We have illustrated how using Block Bootstrap, it is possible to identify correlations among the parameters of the model and to provide a more *realistic* estimate of its error bars. We have then used such error bars to propagate the statistical uncertainties on the equation of state of the outer crust of a neutron star. We have observed that the most favourable nuclear configuration are the semi-magic one with N=50 and N=82. The presence of error impacts strongly the transition between these two magic number. We have provided a prediction in terms of probabilities of the most favourable nuclei to happen. By applying such a systematic method to various mass model, we aim at providing a robust guidance for future experimental measurements to identify key nuclei to improve our knowledge of the outer crust of a neutron star.

Finally, we have also studied the use of neural networks to further reduce the discrepancy between the DZ10 model and the experimental masses. Although a more systematic approach is required [19], we have obtained very encouraging results, in good agreement with previous findings [21]. The use of NN allows for a further reduction of the global root mean square deviation of the model to $\sigma_{DZ+NN} = 0.344$ keV, a value which is quite close to the most accurate mass model available in the literature [69].

[1] B. P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, et al., Physical Review Letters **119**, 161101 (2017).

[2] B. P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, et al., Astrophys. J. Lett **848**, L12 (2017).

[3] A. Pastore, S. Baroni, and C. Losa, Physical Review C **84**, 065807 (2011).

[4] A. Pastore, Physical Review C **86**, 065802 (2012).

[5] N. Chamel, Physical Review C **85**, 035801 (2012).

[6] J. Pearson, N. Chamel, S. Goriely, and C. Ducoin, Physical Review C **85**, 065803 (2012).

[7] A. Pastore, J. Margueron, P. Schuck, and X. Viñas, Phys. Rev. C **88**, 034314 (2013), URL https://link.aps.org/doi/10.1103/PhysRevC.88.034314.

[8] A. Gezerlis, C. Pethick, and A. Schwenk, arXiv preprint arXiv:1406.6109 (2014).

[9] A. Pastore, M. Shelley, S. Baroni, and C. Diget, Journal of Physics G: Nuclear and Particle Physics **44**, 094003 (2017).

[10] M. Alford, D. Blaschke, A. Drago, T. Klähn, G. Pagliara, and J. Schaffner-Bielich, Nature **445**, E7 (2007).
[11] N. Chamel and P. Haensel, Living Reviews in relativity **11**, 10 (2008).
[12] B. Sharma, M. Centelles, X. Viñas, M. Baldo, and G. Burgio, Astronomy & Astrophysics **584**, A103 (2015).
[13] D. Chatterjee and I. Vidana, The European Physical Journal A **52**, 29 (2016).
[14] I. Vidana, M. Bashkanov, D. Watts, and A. Pastore, Physics Letters B **781**, 112 (2018).
[15] G. Baym, C. Pethick, and P. Sutherland, The Astrophysical Journal **170**, 299 (1971).
[16] J. Dobaczewski, W. Nazarewicz, and P. Reinhard, Journal of Physics G: Nuclear and Particle Physics **41**, 074001 (2014).
[17] N. Schunck, J. D. McDonnell, J. Sarich, S. M. Wild, and D. Higdon, Journal of Physics G: Nuclear and Particle Physics **42**, 034024 (2015).
[18] J. McDonnell, N. Schunck, D. Higdon, J. Sarich, S. Wild, and W. Nazarewicz, Physical review letters **114**, 122501 (2015).
[19] L. Neufcourt, Y. Cao, W. Nazarewicz, F. Viens, et al., Physical Review C **98**, 034318 (2018).
[20] R. Utama, J. Piekarewicz, and H. Prosper, Physical Review C **93**, 014311 (2016).
[21] R. Utama and J. Piekarewicz, Physical Review C **96**, 044308 (2017).
[22] R. Utama and J. Piekarewicz, Physical Review C **97**, 014306 (2018).
[23] T. Sprouse, R. N. Perez, R. Surman, M. Mumpower, G. McLaughlin, and N. Schunck, arXiv preprint arXiv:1901.10337 (2019).
[24] M. Kortelainen, Journal of Physics G: Nuclear and Particle Physics **42**, 034021 (2015).
[25] X. Roca-Maza, N. Paar, and G. Colo, Journal of Physics G: Nuclear and Particle Physics **42**, 034033 (2015).
[26] P. Becker, A. Pastore, D. Davesne, and J. Navarro, arXiv preprint arXiv:1811.07866 (2018).
[27] J. Duflo and A. Zuker, Physical Review C **52**, R23 (1995).
[28] J. Mendoza-Temis, J. G. Hirsch, and A. P. Zuker, Nuclear Physics A **843**, 14 (2010).
[29] A. Zuker, in *11th Symposium on Nuclei in the Cosmos* (SISSA Medialab, 2011), vol. 100, p. 083.
[30] C. Qi, Journal of Physics G: Nuclear and Particle Physics **42**, 045104 (2015).
[31] A. Pastore, Journal of Physics G: Nuclear and Particle Physics **46**, 052001 (2019).
[32] M. Wang, G. Audi, F. Kondev, W. Huang, S. Naimi, and X. Xu, Chinese Physics C **41**, 030003 (2017).
[33] W. T. Vetterling, S. A. Teukolsky, W. H. Press, and B. P. Flannery, *Numerical recipes*, vol. 3 (Cambridge University Press Cambridge:, 1994).
[34] G. Audi and A. Wapstra, Nuclear Physics A **595**, 409 (1995).
[35] R. Utama and J. Piekarewicz, Physical Review C **96**, 044308 (2017).
[36] M. Wang, G. Audi, A. Wapstra, F. Kondev, M. MacCormick, X. Xu, and B. Pfeiffer, Chinese Physics C **36**, 1603 (2012).
[37] R.J.Barlow, *A Guide to the Use of Statistical Methods in the Physical Sciences* (John Wiley, 1989).
[38] G. Bertsch and D. Bingham, Physical review letters **119**, 252501 (2017).
[39] P. J. Brockwell and R. A. Davis, *Time series: theory and methods* (Springer Science & Business Media, 2013).
[40] B. Efron and R. J. Tibshirani, *An introduction to the bootstrap* (CRC press, 1994).
[41] J.-P. Kreiss and E. Paparoditis, Journal of the Korean Statistical Society **40**, 357 (2011).
[42] M. R. Chernick, W. González-Manteiga, R. M. Crujeiras, and E. B. Barrios, in *International encyclopedia of statistical science* (Springer, 2011), pp. 169–174.
[43] D. Muir, A. Pastore, J. Dobaczewski, and C. Barton, Acta Physica Polonica B **49** (2018).
[44] B. Efron, Annals of Statistics **7**, 1 (1979).
[45] J. L. Rasmussen, Psychological Bulletin **101**, 136 (1987).
[46] N. I. Fisher and P. Hall, Geophysical Journal International **101**, 305 (1990).
[47] W. Sauerbrei and M. Schumacher, Statistics in medicine **11**, 2093 (1992).
[48] B. F. Manly, vol. 70 (CRC press, 2006).
[49] R. N. Pérez, J. Amaro, and E. R. Arriola, Physics Letters B **738**, 155 (2014).
[50] B. Pasquini, P. Pedroni, and S. Sconfietti, Physical Review C **98**, 015204 (2018).
[51] J.-P. Kreiss and S. N. Lahiri, in *Handbook of statistics* (Elsevier, 2012), vol. 30, pp. 3–26.
[52] T. Nikšić and D. Vretenar, Physical Review C **94**, 024333 (2016).
[53] K. Gurney, *An introduction to neural networks* (CRC press, 2014).
[54] D. Svozil, V. Kvasnicka, and J. Pospichal, Chemometrics and intelligent laboratory systems **39**, 43 (1997).
[55] P. Ramachandran, B. Zoph, and Q. V. Le, arXiv preprint arXiv:1710.05941 (2017).
[56] Z. Niu and H. Liang, Physics Letters B **778**, 48 (2018).
[57] A. Welker, N. Althubiti, D. Atanasov, K. Blaum, T. E. Cocolios, F. Herfurth, S. Kreim, D. Lunney, V. Manea, M. Mougeot, et al., Physical review letters **119**, 192502 (2017).
[58] Y. Gal and Z. Ghahramani, arXiv preprint arXiv:1506.02142 (2015).
[59] A. Idini, arXiv preprint arXiv:1904.00057 (2019).
[60] J. Barea, A. Frank, J. G. Hirsch, and P. Van Isacker, Physical review letters **94**, 102501 (2005).
[61] X. Roca-Maza and J. Piekarewicz, Physical Review C **78**, 025807 (2008).
[62] R. A. Coldwell-Horsfall and A. A. Maradudin, Journal of Mathematical Physics **1**, 395 (1960).
[63] S. B. Rüster, M. Hempel, and J. Schaffner-Bielich, Physical Review C **73**, 035804 (2006).
[64] J. Pearson, S. Goriely, and N. Chamel, Physical Review C **83**, 065810 (2011).
[65] A. Sobiczewski and Y. A. Litvinov, Physical Review C **90**, 017302 (2014).
[66] D. Neill, K. Medler, A. Pastore, and C. Barton, arXiv preprint arXiv:1909.11496 (2019).
[67] M. W. Kirson, Nuclear Physics A **893**, 27 (2012).
[68] M. Fortin, C. Providência, A. R. Raduta, F. Gulminelli, J. Zdunik, P. Haensel, and M. Bejger, Physical Review C **94**,

035804 (2016).

[69] M. Liu, N. Wang, Y. Deng, and X. Wu, Phys. Rev. C **84**, 014333 (2011), URL `https://link.aps.org/doi/10.1103/PhysRevC.84.014333`.