Inferring the dense matter equation of state from neutron star observations via artificial neural networks

Ameya Thete\textsuperscript{a} Kinjal Banerjee\textsuperscript{a} Tuhin Malik\textsuperscript{b}

\textsuperscript{a}Department of Physics, Birla Institute of Technology and Science, Pilani – KK Birla Goa Campus, Zuarinagar, Goa 403726, India
\textsuperscript{b}CFisUC, Department of Physics, University of Coimbra, 3004–516, Coimbra, Portugal

E-mail: f20180885@goa.bits-pilani.ac.in, kinjalb@gmail.com, tm@uc.pt

Abstract. The difficulty in describing the equation of state (EoS) for nuclear matter at densities above the saturation density ($\rho_0$) has led to the emergence of a multitude of models based on different assumptions and techniques. These EoSs, when used to describe a neutron star (NS), lead to differing values of observables. An outstanding goal in astrophysics is to constrain the dense matter EoS by exploiting astrophysical and gravitational wave measurements. Nuclear matter parameters appear as Taylor coefficients in the expansion of the EoS around the saturation density of symmetric and asymmetric nuclear matter, and provide a physically-motivated representation of the EoS. In this paper, we introduce a deep learning-based methodology to predict key neutron star observables such as the NS mass, NS radius, and tidal deformability from a set of nuclear matter parameters. Using generated mock data, we confirm that the neural network model is able to accurately capture the underlying physics of finite nuclei and replicate inter-correlations between the symmetry energy slope, its curvature and the tidal deformability arising from a set of physical constraints. We study the validity of our trained model using Bayesian inference and show that the performance of our model is on par with physics based models with the added benefit of much lower computational cost.

Keywords: Neutron stars, Machine learning, Bayesian reasoning
1 Introduction

Neutron Stars (NSs) are some of the most fascinating astrophysical objects in multi-messenger astronomy. They contain matter at extreme conditions far beyond the ones accessible in a terrestrial laboratory. The core of a NS is believed to contain matter at a few times nuclear saturation density, $\rho_0$ [1–4]. While the structure of a NS can be determined using the Tolman-Oppenheimer-Volkoff (TOV) equations [5, 6], this requires the knowledge of the dense matter Equation of State (EoS). Understanding the internal structure of neutron stars in terms of fundamental interactions between its constituents is an open problem in nuclear physics. To understand the physics of dense matter, at low densities of $\rho \sim 1 - 2\rho_0$, we can use \textit{ab initio} approaches derived from chiral effective field theory ($\chi$EFT) [7–11]. At large densities of $\rho \geq 40\rho_0$, perturbative quantum chromodynamics (QCD) calculations converge and provide reliable estimates [11–14]. However, in the intermediate density region around $\rho \sim 2 - 10\rho_0$ which is relevant for most structural descriptions of neutron stars, reliable calculations from first principles are currently unavailable [11, 15]. Quantum field theory calculations based on lattice QCD are challenging at these densities due to the sign problem that arises in Monte Carlo simulations [16]. As a result, structural descriptions of neutron stars rely on relativistic and non-relativistic phenomenological models for the EoS. The nuclear matter parameters (NMPs), which form the basis of the construction of these equations of state for neutron star matter, are not directly accessible. While lower-order NMPs can be empirically extracted through finite nuclei nuclear physics experiments [17–20], in order to constrain higher-order NMPs, we need to rely on astrophysical observations [21–23].

Recent developments in multi-messenger astronomy have provided important information about high-density nuclear matter physics relevant for NSs. Constraining the EoS is a joint task between nuclear physics and astrophysics. Measured astrophysical quantities such as $\rho_0 = 0.16 \text{ fm}^{-3}$...
as NS observables can uncover properties of dense nuclear matter. Narrowing the constraints on NS observables, therefore, has the ability to constrain the behavior of matter under extreme conditions. It is expected that precise and simultaneous measurements of NS properties like mass, radius, moment of inertia and tidal deformability may help constrain the EoS to a narrow range [24–27]. High mass pulsars like PSR J1614–2230 \((M = 1.908 \pm 0.016 M_\odot)\) [28–30], PSR J0348–0432 \((M = 2.01 \pm 0.04 M_\odot)\) [31], PSR J0740+6620 \((M = 2.07 \pm 0.07 M_\odot)\) [32], and very recently PSR J1810+1714 \((M = 2.13 \pm 0.04 M_\odot)\) [33] have already placed tight constraints on the EoS. GW signals emitted from a NS merger event depend on the behaviour of the neutron star matter at high densities [34, 35]. Therefore, the values of tidal deformability obtained from GW events such as GW170817 associated with a binary NS merger, as well as the simultaneous measurement of NS masses and radii from high-precision X-ray space missions, such as NICER (Neutron star Interior Composition ExploreR), may help further constrain the EoS. Some current observational evidences are the simultaneous measurements of NS mass \(1.34^{+0.15}_{-0.16} M_\odot\) and radius \(12.71^{+1.14}_{-1.19} \) km for the pulsar PSR J0030+0451 by NICER [36]. Other independent analyses show that the radius is \(13.02^{+1.24}_{-1.06}\) km and the mass \(1.44^{+0.15}_{-0.14} M_\odot\) [37]. However the recent measurement of the equatorial circumferential radius of the pulsar PSR J0740+6620 with mass \(M = 2.072^{+0.067}_{-0.066} M_\odot\) and \(R = 12.39^{+1.30}_{-0.98}\) km (68 % CI) [38] cannot further constrain the EoS which already predicts a NS with maximum mass more than \(2M_\odot\) [39].

The EoS – to a good approximation – can be expressed in terms of NMPs at saturation density. The NMPs usually considered for constructing the EoS are the incompressibility coefficient, the skewness parameter of the symmetric nuclear matter, the symmetry energy coefficient, its slope, and the curvature parameters characterizing the density dependence of the symmetry energy. Recently, there has been a comprehensive analysis of correlations of tidal deformability and other NS properties with NMPs [21, 40, 41]; however, these correlations are found to be model-dependent [41, 42]. Ref. [43] shows that the correlations are sensitive to finite nuclei properties, which are accessible to laboratories. Therefore, finite nuclei properties are important quantities to consider in a model while determining the correlations between NS properties and NMPs. Of late, the EoSs obtained by several meta-models have gained popularity owing to their cost-effectiveness in big simulations [13, 24, 44]. These models are constrained by \textit{ab initio} theoretical calculations of nucleon-nucleon chiral potentials for low-density neutrons and nuclear matter [45, 46], and perturbative QCD for asymptotically high-density regimes [13]. In the intermediate density region, the EoS is evolved in a thermodynamically consistent manner with either piece-wise polytropic segments [47–49], a speed-of-sound interpolation, or a spectral interpolation [50, 51]. These meta-models are limited to incorporating finite nuclei properties and differ on results in establishing a bridge between NS properties and NMPs. Non-parametric models of the NS EoS have also been proposed based on Gaussian processes (GPs) [52, 53] which use Bayesian methods to infer the EoS from multi-messenger data. These models are usually computationally expensive to implement or are highly sensitive to the choice of the training data sets and therefore, might be limited by the current knowledge of the EoS [54]. Consequently, there is a need to search for alternative approaches to construct a model-independent EoS.

In recent years, deep learning (DL) has been extensively applied to a wide range of technological and scientific tasks. DL algorithms, which are a class of machine learning (ML) algorithms, are highly scalable and distributed computational techniques with the ability to learn intricate relationships from raw data using units called neurons arranged in a stacked fashion. The advent of high-performance computing and the development of parallel devices...
like graphics processing units have rendered DL as the primary choice of algorithms for tasks such as computer vision [55], or natural language processing [56]. DL models have been used as alternatives to conventional statistical framework and have been successfully applied to many problems in physics. Most applications of ML and DL in physics have been in analyzing data obtained from data-intensive experiments like LIGO for the detection and denoising of GW signals [57–59], and the Large Hadron Collider for particle track reconstruction or anomaly detection [60–62]. Significant progress has also been made in using these algorithms in the context of nuclear physics [63] and neutron star physics [54, 64, 65]. While these applications are certainly promising, it remains to be seen up to what extent a DL model can supplement existing physical models. Recent research also aims to address whether a trained DL model can produce correct predictions from experimental data alone and whether such predictions are comparable to mesoscopic phenomenological physics models [66]. Such ML and DL-based models do not have the feature richness possessed by physics-based models but they offer other benefits like cost-effectiveness while dealing with a large amount of experimental or observational data. Recent works [15, 64, 67] have studied the applications of machine learning methods to the neutron star EoS. They employ a feedforward neural network (FFNN) to map neutron star data to EoS parameters. Instead of considering the FFNN as merely an interpolation tool between EoS NMPs and neutron star observables, we adopt an approach similar to [54] by treating the FFNN as a representation of the EoS itself.

In this paper, our focus is on implementing an artificial neural network to emulate a physics model and assess its accuracy through validation and inference tasks. Our primary goal is to thoroughly examine the neural network’s ability to mimic the physics model, and subsequently utilize the trained network for computationally expensive Bayesian inference tasks. To the best of our knowledge, the use of neural networks in Bayesian inference framework has not been employed in nuclear astrophysics literature. We show that we can indeed obtain results comparable to physics models at a fraction of the computational cost. This points out a novel utilization of machine learning and neural network models in research. Let us however clarify that we are not claiming that neural networks can replace physics based models. Our claim is the such models can help us capture some of the features of the EoS at much lower computational costs.

The paper is structured as follows: in Section 2, we discuss the parameterization of the EoS which leads to the emergence of NMPs, followed by a brief review of the theory of artificial neural networks and the DL approach used to map NMPs to NS observables. This is followed by a description of the Bayesian statistical framework. In 3 we test how well does the best fit neural network model capture correlations between EoS parameters. We also perform a Bayesian analysis to compare our models with a physics based model and show that the computational cost is much less for our model. We finally end with our conclusions in Section 4.

2 Framework

In this section, we outline the different facets of the adopted framework for our analysis. In Sec. 2.1, we briefly describe the neutron star EoS and the nuclear matter parameters involved in its construction. Then, in Sec. 2.2, we describe the generation of the data set used to train the neural network model. Sec. 2.3 describes at length the DL approach used to construct a deep neural network model which accepts nuclear matter saturation parameters as inputs and produces neutron star properties obtained from a set realistic nuclear physics EoSs as
Finally, in Sec. 2.4 we present a Bayesian inference framework that is applied on the trained neural network model to verify the performance of our model in comparison with Skryme models.

2.1 Nuclear matter parameters

The structure of neutron stars is obtained by solving the Tolman-Oppenheimer-Volkoff (TOV) equation with a given EoS for the nuclear matter. The EoS is expressed as the variation of pressure \( e \) with density \( \rho \), over a wide range of densities. To a good approximation, any EoS calculated from phenomenological nuclear models can be decomposed into two parts, (i) the EoS for symmetric nuclear matter \( e(\rho, 0) \); and (ii) a term involving the symmetry energy coefficient \( S(\rho) \) and the asymmetry \( \delta \),

\[
e(\rho, \delta) \simeq e(\rho, 0) + S(\rho)\delta^2,
\]

where \( \rho = \rho_n + \rho_p \) is the baryon density, \( \rho_n \) and \( \rho_p \) are the neutron and proton densities respectively, and the asymmetry \( \delta = (\rho_n - \rho_p)/\rho \). We can then characterize the density dependence of the energy density of symmetric matter around the saturation density \( \rho_0 \) in terms of a few bulk parameters by constructing a Taylor expansion around \( \rho_0 \). That is,

\[
e(\rho, 0) = e_0 + \frac{K_0}{2} \left( \frac{\rho - \rho_0}{3\rho_0} \right)^2 + \frac{Q_0}{6} \left( \frac{\rho - \rho_0}{3\rho_0} \right)^3 + O(4) \tag{2.2}
\]

The coefficients \( e_0, K_0, Q_0 \) denote the energy per particle, the incompressibility coefficient, and the third derivative of symmetric matter at saturation density, respectively. Similarly, the behaviour of the symmetry energy around saturation can also be characterized in terms of a few bulk parameters,

\[
S(\rho) = J_{\text{sym},0} + L_{\text{sym},0} \left( \frac{\rho - \rho_0}{3\rho_0} \right) + \frac{K_{\text{sym},0}}{2} \left( \frac{\rho - \rho_0}{3\rho_0} \right)^2 + O(3) \tag{2.3}
\]

where \( J_{\text{sym},0} = S(\rho_0) \) is the symmetry energy at saturation density. The incompressibility \( K_0 \), the skewness coefficient \( Q_0 \), the symmetry energy slope \( L_{\text{sym},0} \), and its curvature \( K_{\text{sym},0} \) evaluated at saturation density, are defined in [68]. These quantities are the key nuclear matter parameters (NMPs) that describe any equation of state (EoS). Hence, an EoS can be represented by a point in the seven-dimensional parameter space of NMPs \( \{e_0, \rho_0, K_0, Q_0, J_{\text{sym},0}, L_{\text{sym},0}, K_{\text{sym},0}\} \) [43]. Symbolically, the \( j \)th EoS in this space is written as

\[
\text{EoS}_j = \{e_0, \rho_0, K_0, Q_0, J_{\text{sym},0}, L_{\text{sym},0}, K_{\text{sym},0}\}_j \approx \mathcal{N}(\boldsymbol{\mu}, \Sigma) \tag{2.4}
\]

where \( \mathcal{N}(\boldsymbol{\mu}, \Sigma) \) is a multivariate Gaussian distribution with \( \boldsymbol{\mu} \) being the mean value of the nuclear matter parameters \( \mathbf{p} \) and a covariance matrix \( \Sigma \). The diagonal elements of \( \Sigma \) represent the variance or the squared error for the parameters \( p_i \). The off-diagonal elements of \( \Sigma \) are the covariances between different parameters \( p_i \) and \( p_j \), and denote the correlation coefficient between them. Hence, given a mean \( \mathbf{\mu} \) and covariance matrix \( \Sigma \), a large number of EoSs can be obtained.

While the Taylor expansion of the symmetric and asymmetric energy is only truly accurate around the saturation density, we can treat these expansions as a parameterization of the EoS — similar to other adopted parameterizations — with the condition that this
representation asymptotically approaches the Taylor expansion in the limit \( \rho \to \rho_0 \) [21, 69]. This lets us ignore any issue arising with the convergence of the approximation. However, the higher-order NMPs obtained via this parameterization might markedly diverge from the actual nuclear matter expansion coefficients. They can be thought of as effective parameters that incorporate the effects of missing higher-order terms. Moreover, Refs. [26, 69] indicate that an EoS obtained from the Skyrme framework can be well-reproduced by considering Taylor coefficients until the third or fourth order.

2.2 Data set generation

In our analysis, the input data used to train an artificial neural network are the seven key nuclear matter parameters that govern the equation of state \{\( e_0 \), \( \rho_0 \), \( K_0 \), \( Q_0 \), \( J_{\text{sym},0} \), \( L_{\text{sym},0} \), \( K_{\text{sym},0} \)\}. Six neutron star properties are the target variables: the maximum NS mass, \( M_{\text{max}} \); the maximum NS radius, \( R_{\text{max}} \); the radius for 1.4\,\( M_\odot \) NS, \( R_{1.4} \); and the tidal deformability \( \Lambda_{\text{M}} \) for NS having mass \( M \in [1.0, 1.4, 1.8] M_\odot \). We generate our data set by sampling points from the multivariate Gaussian distribution \( N(\mu, \Sigma) \), where \( \mu \) is the mean vector with components \( \mu_i = \mathbb{E}[p_i] \) and \( \Sigma \) is the covariance matrix with entries \( \Sigma_{ij} = \mathbb{E}[(p_i - \mu_i)(p_j - \mu_j)] \), for a NMP \( p \). This method closely follows the procedure for generating the Case-II data set in Ref. [43]. We assume an a priori inter-correlation coefficient between \( L_{\text{sym},0} \) and \( K_{\text{sym},0} \) of 0.8, which is reasonable choice for nuclear physics models that satisfy finite nuclear properties. Models which satisfy finite nuclear properties exhibit different correlations between NMPs and NS properties as compared to meta-models or nuclear physics model for infinite nuclear matter that do not respect finite nuclear properties [41, 42]. Therefore, we consider Case-II data of Ref. [43] to mimic the microphysics information of finite nuclear properties. Figure 1 presents the 7 \times 7 matrix for the correlation coefficients between the NMPs in the sampled data. The central values and uncertainties on each NMP in the constructed distribution are listed in Table 1.

| NMP       | MVGD      | Tranning Set |
|-----------|-----------|--------------|
|           | \( \mu_{p_i} \) | \( \sqrt{\Sigma_{p_ip_i}} \) | \( \mu_{p_i} \) | \( \sqrt{\Sigma_{p_ip_i}} \) |
| \( e_0 \) | -16.0     | 0.25         | -15.99        | 0.25 |
| \( \rho_0 \) | 0.16      | 0.005        | 0.158         | 0.005 |
| \( K_0 \) | 230       | 20           | 242           | 15   |
| \( Q_0 \) | -300      | 100          | -307          | 70   |
| \( J_{\text{sym},0} \) | 32        | 3            | 32.65         | 2.8  |
| \( L_{\text{sym},0} \) | 60        | 20           | 73            | 12   |
| \( K_{\text{sym},0} \) | -100      | 100          | -20           | 50   |

Table 1. The mean value \( \mu_{p_i} \) and error \( \sqrt{\Sigma_{p_ip_i}} \) for the nuclear matter parameters \( p_i \) employed for the multivariate Gaussian distribution. All quantities are in units of MeV except for \( \rho_0 \) which is in units of \( \text{fm}^{-3} \).

We then construct the Skyrme EoS from a drawn NMP sample and check whether the EoS (a) predicts a NS maximum mass above 2\,\( M_\odot \); (b) predicts a tidal deformability for a NS with \( M = 1.4 M_\odot \) below 800; and (c) satisfies the causality condition, i.e., the speed of sound \( c_s = \sqrt{\frac{\rho_0}{\rho c^2}} \leq c \) at the center of maximum mass NS, where \( c \) is the speed of

\[ \mathbb{E}[X] \] is the expectation value of a random variable \( X \) and is defined as \( \mathbb{E}[X] = \int_{-\infty}^{\infty} xp(x) \), where \( p(x) \) is the probability density function.
Figure 1. Correlations among the various NMPs for the sampled data set. Correlations among the off-diagonal pairs corresponding to $L_{\text{sym},0} - K_{\text{sym},0}$ are the only nontrivial values. Negligible correlations between other NMPs are visible as the figure plots sample statistics.

light in vacuum. Any samples that generate EoSs which do not satisfy these conditions are discarded. For each EoS, the TOV and deformability equations [70] are solved to obtain the six aforementioned NS observables. Naturally, realistic NS observations accrue experimental and instrumental errors, which result in corresponding uncertainties while reconstructing the NMPs. For simplicity, we choose to disregard NS observational errors and uncertainties while training the model. Following this procedure, we obtain 2106 filtered NMPs and corresponding NS observables which form the data set.

2.3 Artificial neural networks

Artificial Neural Networks (ANN) are a class of machine learning algorithms that have found widespread use over the past decade. The popularity of ANNs arises from their ability to model complex nonlinear relationships within data and their potential to generalize to a wide class of functions. An ANN with sufficiently many layers or neurons is theoretically capable of representing any continuous function [71, 72]. Therefore, as a general rule, a more complex network with a larger number of parameters is able to learn more abstract features from the data.

Feedforward Neural Networks (FFNN) are the simplest type of neural network architecture. For an input $x$, the objective of a feedforward neural network is to approximate some true mapping $f^*(x)$ by a mapping $f(x; \theta)$, parameterized by a set of weights $\theta$. A typical feedforward neural network consists of a number of processing units called neurons, arranged into one or many layers composed in a sequential fashion. A neuron performs a linear operation by aggregating weighted inputs received from neurons in the previous layer. A FFNN generally consists of an input layer, followed by one or more hidden layers, and a
final output layer consisting of one or more neurons. Computation in an FFNN flows in a linear fashion, starting at the input layer and moving successively through the hidden layers until it reaches the output layer. Figure 2 provides an illustration of a simple feedforward neural network with one input layer, two hidden layers and an output layer.

The parameters \( \theta \) typically represent the weights assigned to a connection between neurons in adjacent layers. The training data provides noisy, approximate examples of \( f^*(x) \) and each \( x \) is paired with a corresponding label \( y \). The goal of a learning algorithm is to learn a particular value of \( \theta \) that results in the best function approximation. During the training procedure, the learning algorithm does not say what each layer does but instead decides how to use these layers to produce an optimal approximation of \( f^* \). As the training data does not show the desired output for intermediate layers, they are called hidden layers. The number of hidden layers decides the depth of the network, and the dimensionality or the number of neurons in each hidden layer determines the width of the network. To introduce nonlinearity in the computation between two successive layers within the model, a nonlinear function, called the activation function, acts element-wise on the output of one hidden layer, and the output of the function is passed to the next layer in the computation. For most modern neural networks, the default recommendation is to use the rectified linear unit, or ReLU [73, 74] activation, defined as \( g(z) = \max\{0, z\} \). Other commonly used choices for the activation function include the tanh function or the logistic function. For a textbook review of neural networks and training algorithms, see Ref. [75].

The ability of neural networks to model highly nonlinear relationships between input and output variables makes them ideal for estimating neutron star properties from the equation of state parameters. This is important because the relationships between these two quantities are expected to be nontrivial and involve multiple intermediate steps composed of nonlinear operations. Moreover, a neural network offers two major advantages over a conventional approach with traditional physics models:

1. An ANN can efficiently map the sample NMPs to NS properties without calculating the EoS from nuclear physics models. Similar works [76, 77] demonstrate that ANNs offer up to a two-fold speedup over conventional astrophysical models

2. ANNs can also accurately capture finite nuclear information, which can computationally expensive to verify with a traditional physics model in a Bayesian setting.

At this step, we wish to note that other machine learning models can also be used for an identical purpose, albeit with varying degrees of success. We performed a preliminary comparison of FFNNs with other ML models, specifically linear regression, support vector regression and eXtreme Gradient Boosting (XGBoost) [78]. We observed that FFNNs outperformed all the other models in the study, and therefore, we chose to use FFNNs for this work.

In supervised learning instances, the data set is generally partitioned into non-overlapping training, testing, and validation sets. The training data set is used by the neural network to learn, the validation set is used to see whether the network is learning properly during the training procedure and to select the optimal values for the hyperparameters, and the testing data set is used to assess the performance of the trained model. It should be noted that the model does not see any instances from the testing data set until after the training is completed. Accordingly, we randomly partition the original generated data set following a 70%-20%-10% split into a training set, a validation set, and a testing set consisting of 1515, 422, and 169 instances, respectively. Before training, the features in the split datasets are
Figure 2. An example of a fully-connected feedforward neural network with two hidden layers. Each hidden layer consists of \( m \) neurons, denoted by \( a^{(l)}_j \), where \( l \) denotes an index over the hidden layers and the index \( j \) runs over the dimension of the hidden layer. This particular network accepts an \( n \)-dimensional input through the input layer (in green), passes the input through the hidden layers (in lavender), and produces a \( k \)-dimensional output at the output layer (in red). Solid lines between individual nodes denote weighted connections that are learned during the training procedure. Activation functions are not shown in this figure.

Standardization of data is a common requirement for many ML algorithms as they tend to perform poorly if the individual features are not standard normally distributed, i.e. Gaussian with unit variance and zero means. Standardization is performed using Scikit-Learn’s [79] `StandardScaler` method. The output of the network is scaled back to the original distribution before computing the loss.

The choice of the ANN architecture, such as the number of layers and the number of neurons in each layer, represents a trade-off between the quality of fit and the overfitting problem. The model should contain sufficiently many trainable parameters to learn the ground truth function accurately. At the same time, the ANN must avoid overfitting the training data by losing its ability to generalize well. As there is no panacea for choosing an optimal model configuration, the final architecture for the ANN was chosen based on empirical tests performed on the data and selecting the set of configurations that performed best on the validation set. Our network consists of two hidden layers with 15 neurons each and uses the rectified linear unit (ReLU) activation function, which is a standard choice for most deep learning applications and is known to mitigate the vanishing gradients problem [80]. The ReLU activation function behaves as an identity function for positive inputs and saturates at 0 for negative inputs. No activation function is applied to the input or the output layer. The design of our feedforward architecture is summarized in Table 2.

The neural network is implemented and trained using the Python Keras library [81] with a TensorFlow [82] backend. Neural network parameters are initialized with the Glorot uniform distribution [83]. We use the Adam optimizer [84] to update the weights of the ANN. For training, we use an initial learning rate of \( 1 \times 10^{-3} \) and the training data is batched into batches of size 16. Training is performed for a total of 40 epochs, or until the validation loss stops decreasing. The model is trained on a 2-core Intel Xeon CPU @ 2.20GHz. The network seeks to optimize a root mean squared error (RMSE) loss between the predictions and actual
| Layer  | Number of neurons | Activation function |
|--------|-------------------|---------------------|
| 0 (Input) | 7               | None                |
| 1, 2   | 15              | ReLU                |
| 3 (Output) | 6               | None                |

Table 2. Neural network architecture used in this work. The input layer consists of seven neurons corresponding to nuclear matter parameters. In the output layer, the six neurons correspond to the six neutron star properties.

2.4 Bayesian statistical framework

The Bayesian statistical framework allows us to carry out a detailed analysis of the parameters of a model for a given set of fit data [85–87]. The hypothesis or the prior knowledge of the model parameters and various constraints on them is encoded through the prior distributions. The technique yields a joint posterior distribution of the model parameters by updating the probability for the hypothesis using the available observational data according to Bayes’ theorem. The posterior distribution of the model parameters $\theta$ can be written as

$$P(\theta|D) = \frac{\mathcal{L}(D|\theta)P(\theta)}{Z}$$

where $\theta$ and $D$ denote the set of model parameters and the fit data, respectively. $P(\theta|D)$ is the joint posterior probability of the parameters, $\mathcal{L}(D|\theta)$ is the likelihood function, $P(\theta)$ is the prior for the model parameters and $Z$ is the evidence. The posterior distribution of a given parameter $\theta_i$ can be obtained by marginalizing $P(\theta|D)$ over remaining parameters. The marginalized posterior distribution for a parameter $\theta_i$ is obtained as

$$P(\theta_i|D) = \int P(\theta|D) \prod_{j \neq i} d\theta_j$$

A major contribution of this study is to demonstrate how a neural network can be effectively integrated within a Bayesian inference framework for astrophysical use cases. In particular, we demonstrate the effect of intercorrelations between symmetry energy parameters on NS properties. We sample NMPs within a Bayesian framework in two different ways: (i) sampling all the NMPs randomly without any intercorrelations among them, (ii) sample all the NMPs randomly, except for the pair $L_{\text{sym,0}}-K_{\text{sym,0}}$ which is set to a correlation coefficient of 0.9. Then, we compare the posterior distribution of the NMPs for both cases under certain constraints on NS properties.

For the sampling process, we use the Gaussian likelihood function defined as,

$$\mathcal{L}(D|\theta) = \prod_j \left( \frac{1}{\sqrt{2\pi} \sigma_j} \right) \exp \left[ -\frac{1}{2} \sum_{i=1}^{N_d} \left( \frac{d_i - m_j(\theta)}{\sigma_j} \right)^2 \right]$$

Here the index $j$ runs over all the data, $d_j$ and $m_j$ are the data and the corresponding model values, respectively. The model $m$ is parameterized by a set of parameters, $\theta$. $\sigma_j$ are the adopted uncertainties. The evidence is used to compare the compatibility of different models with the available data. In our present work, the evidence $Z$ is not relevant and thus can be
Table 3. The mean value $\mu_{p_i}$ and error $\sqrt{\Sigma_{p_i,p_i}}$ for the nuclear matter parameters $p_i$ in the prior multivariate Gaussian distribution. All quantities are in units of MeV except for $\rho_0$ which is in units of fm$^{-3}$.

| NMP     | $p_i$ | $\mu_{p_i}$ | $\sqrt{\Sigma_{p_i,p_i}}$ | $\mu_{p_i}$ | $\sqrt{\Sigma_{p_i,p_i}}$ |
|---------|-------|-------------|----------------------------|-------------|----------------------------|
| $e_0$   | -15.99| 0.25        | -15.99                     | 0.50        |
| $\rho_0$| 0.158 | 0.005       | 0.158                      | 0.01        |
| $K_0$   | 242   | 15          | 242                        | 30          |
| $Q_0$   | -307  | 70          | -307                       | 140         |
| $J_{\text{sym,0}}$ | 32.65 | 2.8        | 32.65                      | 5.6         |
| $L_{\text{sym,0}}$ | 73    | 12         | 73                         | 24          |
| $K_{\text{sym,0}}$ | -20   | 50         | -20                        | 100         |

2.4.1 Log Likelihood

The log-likelihood for mass, radius at maximum mass, and radius at $1.4M_\odot$ are defined as follows:

$$\log \chi(M_{\text{max}}) = \log \left( \frac{1}{\epsilon \exp \left[ \frac{M_{\text{cal}} - M_{\text{obs}}}{-0.01} \right]} + 1 \right)$$  \hspace{1cm} (2.8)

$$\log \chi(R_{2.07}) = -0.5 \left[ \frac{R_{2.07,\text{cal}} - R_{2.07,\text{obs}}}{\Delta R_{2.07}} \right]^2 + \log \left[ 2\pi \Delta R_{2.07}^2 \right]$$  \hspace{1cm} (2.9)

$$\log \chi(R_{1.4}) = -0.5 \left[ \frac{R_{1.4,\text{cal}} - R_{1.4,\text{obs}}}{\Delta R_{1.4}} \right]^2 + \log \left[ 2\pi \Delta R_{1.4}^2 \right]$$  \hspace{1cm} (2.10)

The calculations are performed for a set of NS observational properties, such as maximum mass ($M_{\text{max}}$), the radius at NS mass ($R_{2.07}$), and at $1.4M_\odot$ ($R_{1.4}$) listed in Table 4.

Bayesian parameter estimation is commonly carried out using the Markov Chain Monte Carlo (MCMC) algorithm. This algorithm updates the existing parameters with a new set of parameters with a probability proportional to the ratio of the two points. However,
| Constraints | Quantity | Value/Band |
|-------------|----------|------------|
| $N_{max}$   | > 2.0 $M_\odot$ | [91] |
| $R_{2.07}$  | 12.4 ± 1.0 km | [91] |
| $R_{1.4}$   | 13.02 ± 1.24 km | [37] |

Table 4. The constraints imposed in the Bayesian inference: Observed maximum mass of NS, Radius of 2.07 $M_\odot$ NS, Radius of 1.4 $M_\odot$ NS.

MCMC approaches can lead to issues with converging to a stable posterior. To overcome this problem, we use the dynamic nested sampling algorithm [88–90]. In dynamic nested sampling, the posterior is broken into many nested “slices” with an initial n-live number of points that vary dynamically as the sampling progresses; samples are generated from each of them and then recombined to construct the posterior distribution. We use the Dynasty dynamic nested sampling algorithm interfaced in BILBY [87] with 5000 n-live points to sample from the posterior distributions of the nuclear matter parameters. Dynasty enables flexible Bayesian inference over complex, multi-modal distributions without needing to converge to the posterior before generating valid samples [90].

3 Results

In this section, we present the main findings of our analysis. After selecting the best model, which we call the NS-ANN, by performing a grid search over hyperparameters such as the model depth, model width, learning rate, etc., we determine its performance on the test set. We wish to reemphasize that the test set is never used during training or the validation phase. Evaluating the model’s performance on this set quantifies the generalization capacity of the model, i.e., its predictive power on unseen data. The RMSE values obtained for each NS observable on the testing data set are summarized in Table 5. We also include the root mean squared relative error as it provides a scale-independent measure of the generalization capacity and eases comparison between multiple dependent variables. In Figure 3, we plot the loss as a function of training time, which is called the learning curve. The learning curve plots the root mean squared error on the y-axis against the number of elapsed epochs — or, the number of full passes over the training data set — on the x-axis. We plot two different learning curves for a single training instance: one for the training loss (in blue) and the other for the validation loss (in red). Figure 3 also plots a 1σ (68% confidence interval) region centered around the training curves computed for 10 independent runs to indicate the degree of variability observed by training the model on different subsets of the original data set. The training and validation losses are the loss functions computed for the training set and validation sets, respectively. The former is a metric of how well the ANN learns the training data, while the latter indicates how well the model is able to predict by generalizing the learned information. For a NS mass of 1.4$M_\odot$, we have a prediction error below 2% for the radius $R_{1.4}$ and 5% for the tidal deformability $\Lambda_{1.4}$. This implies that using the ANN, we can infer $\Lambda_{1.4}$ and $R_{1.4}$ with an average error of 19.239 and 0.194 km, respectively. The NS maximum mass, $M_{max}$, and the maximum radius $R_{max}$, are also predicted with an error below 2%. Moreover, the saturation of the loss curves for the training and validation sets indicates that the training has converged to a minimum and that the model can generalize sufficiently well.
Table 5. Root Mean Squared Error (RMSE) on the test set, defined as \(\sqrt{(1/N) \sum_{i=1}^{N} (\hat{y}_i - y_i)^2}\), where \(N\) is the total number of samples in the test set. We also show the root mean squared relative error, \((\text{RMSE}/\bar{y}) \times 100\), where \(\bar{y} = (1/N) \sum_{i=1}^{N} y_i\) is the mean of the true value. All quantities are dimensionless except \(R_{\text{max}}\) and \(R_{1.4}\), which are in units of km, and \(M_{\text{max}}\), which is in units of \(M_\odot\).

| \(\hat{y}\) | RMSE | (RMSE/\(\bar{y}\)) \times 100 |
|-------------|------|-------------------------------|
| \(M_{\text{max}}\) | 0.024 | 1.1                          |
| \(R_{\text{max}}\) | 0.088 | 0.8                          |
| \(R_{1.4}\) | 0.194 | 1.5                          |
| \(\Lambda_{1.0}\) | 123.800 | 3.7                        |
| \(\Lambda_{1.4}\) | 19.239 | 4.3                         |
| \(\Lambda_{1.8}\) | 5.344 | 8.2                         |

Once we have successfully and efficiently learned the non-linear mapping between the empirical nuclear matter parameters and the NS observables, we must first prove the correctness of the NS-ANN model before using it in the context of a Bayesian inference framework to constrain the NMPs. In Figure 4 we plot the 1\(\sigma\) confidence ellipses for \(\Lambda_M\) versus \(L_{\text{sym},0}\) and \(K_{\text{sym},0}\) for NS mass \(M = 1.0, 1.4,\) and \(1.8M_\odot\) predicted by the NS-ANN for three different pseudo-sets of NMPs. These pseudo-NMP distributions were constructed as a multivariate Gaussian distribution by setting the inter-correlation between \(L_{\text{sym},0}\) and \(K_{\text{sym},0}\) to \(r = 0.3, 0.6\) and 0.9, and all other parameters of the distribution to values listed in Table 1. Once again, all NMPs that generate EoSs which do not satisfy the conditions from Sec. …
2.3 are discarded. We use the NS-ANN model for each NMP in the dataset to generate a corresponding set of NS observables. Through this analysis, we wish to establish that a neural network is capable of replicating the underlying microphysical information conveyed by a conventional EoS model. The values of the correlation coefficients for the results illustrated in the figure are summarized in Table 6. For the first set, we observe that the correlations of $\Lambda^1_{0.1,0.8}$ with $K_{\text{sym},0}$ are $\chi \sim 0.5 - 0.8$, and those with $L_{\text{sym},0}$ are $\chi \sim 0.2 - 0.8$. For the second set, we see a non-trivial narrowing of the confidence ellipses, indicating stronger correlations of $\Lambda^1_{0.1}$ and $L_{\text{sym},0}$ and $K_{\text{sym},0}$, $\chi \sim 0.7 - 0.8$, while these correlations become moderate for $\Lambda^1_{1.8}$. We also observe that the $\Lambda_M - L_{\text{sym},0}$ correlations decrease with increasing mass of NS, while an opposite trend is observed for $\Lambda_M - K_{\text{sym},0}$ correlations. Moreover, for the same NS mass, $\Lambda_M - L_{\text{sym},0}$ correlations are much more sensitive to $L_{\text{sym},0} - K_{\text{sym},0}$ correlations over $\Lambda_M - K_{\text{sym},0}$ correlations. These results help emphasize that the correlations of the tidal deformability with $L_{\text{sym},0}$ and $K_{\text{sym},0}$ are sensitive to the physical correlations among the empirical NMPs arising from a set of physical constraints [26, 27]. Additionally, the observations made in this analysis are in excellent agreement with results from previous studies performed on observing the effect of correlations of the tidal deformability with the slope of the symmetry energy, and its curvature [40, 43, 92]. At this point we also wish to note that there exists a strong model-independent intercorrelation between $L_{\text{sym},0}$ and $J_{\text{sym},0}$, but this correlation does not disrupt the relationships between some NS properties, such as tidal deformability, radius, and NMPs [43], and as a result is not covered in this study.

| $r$ | $L_{\text{sym},0}$ | $\Lambda^1_0$ | $\Lambda^1_4$ | $\Lambda^1_{1.8}$ |
|-----|-------------------|--------------|--------------|------------------|
| 0.3 | $K_{\text{sym},0}$ | 0.81         | 0.56         | 0.23             |
|     | $K_{\text{sym},0}$ | 0.57         | 0.78         | 0.81             |
| 0.6 | $L_{\text{sym},0}$ | 0.86         | 0.69         | 0.43             |
|     | $K_{\text{sym},0}$ | 0.69         | 0.80         | 0.79             |
| 0.9 | $L_{\text{sym},0}$ | 0.91         | 0.84         | 0.71             |
|     | $K_{\text{sym},0}$ | 0.86         | 0.86         | 0.79             |

Table 6. The values of the correlation coefficients for $\Lambda^1_{0.1,0.4,1.8}$ with $L_{\text{sym},0}$ and $K_{\text{sym},0}$ for three different $L_{\text{sym},0} - K_{\text{sym},0}$ correlation coefficients, $r$

As a next step, we apply our trained NS-ANN model to a dataset acquired from a different class of physics models in order to assess the model-dependent uncertainty related to neutron star properties.

The nuclear matter parameters are known to be model-dependent and hence also the properties of stars or EoSs [93–95]. In Table 7, we compare the NS properties obtained using a dataset generated by a density-dependent meson coupling framework called the DDB model. This model is constructed using a Bayesian inference approach in conjunction with minimal constraints on the nuclear saturation properties, the maximum mass of a neutron star exceeding $2M_\odot$, and a low-density equation of state (EoS) calculated using chiral effective field theory. In the first horizontal panel, we show that NS properties predicted by the NS-ANN ML model match well with the theoretical predictions of the Skyrme model. In the second horizontal panel, then we apply our trained NS-ANN model to a data set generated by DDB. It can be seen that the residual errors on the NS properties are much larger in this case. We posit that this wider uncertainty might arise due to input parameters (NMPs) in
Figure 4. The 1σ confidence ellipses in the planes of $\Lambda_M - L_{\text{sym},0}$ (top) and $\Lambda_M - K_{\text{sym},0}$ (bottom) with $M = 1.0, 1.4$ and $1.8M_\odot$ generated by the NS-ANN model for the correlation coefficient, $r$, between $L_{\text{sym},0}$ and $K_{\text{sym},0}$ set to 0.3, 0.6 and 0.9.

the DDB dataset lying outside the parameter space covered by the training set. To avoid any extrapolation error, we then filter the DDB according to the parameter space of the training sets. We call this set DDB-FL and apply the NS-ANN on this set. It can be seen that the dispersion is still large. The NS-ANN model is trained on data obtained from the non-relativistic Skryme model, so it is unsurprising that it fails to completely capture the features of the relativistic DDB model.

3.1 Validity of NS-ANN model in Bayesian Inference

In this section, we apply our NS-ANN model in a Bayesian inference framework. The primary objective is to validate the trained model's effectiveness in computationally expensive calculations of this type. The results will be analyzed in two parts: i) The posterior of neutron star properties will be compared between trained NS-ANNs and Skyrme models in Bayesian Inference, ii) the CPU time will be compared between the two models.

3.2 Comparison between NS-ANN and Skyrme

As discussed above, Bayesian inference requires three ingredients: a) the model, b) the fit data or likelihood, and c) prior. We will perform identical Bayesian inference calculations with both Skyrme and NS-ANN models and compare their posteriors. The input priors are the nuclear matter parameters, namely $e_0, \rho_0, K_0, Q_0, J_{\text{sym},0}, L_{\text{sym},0}$, and $K_{\text{sym},0}$. In Table 3, we have defined two multivariate Gaussian priors, P1 and P2. P1 represents the prior validity in which the neural network is originally trained, while in P2, we have doubled the uncertainty to evaluate the performance of the NS-ANN model. The likelihood for our calculation is defined in section 2.4.1 for the fit data described in 4. In the likelihood, we have only considered observed NS properties because we are motivated to compare the performance of the NS-ANN trained model with the Skyrme nuclear physics model within observational constraints.

We used the Pymultinest sampler to sample our parameter. In Figure 5, we compared the Bayesian posterior results for the neutron star properties obtained from the prior set P1 with the NS-ANN and Skyrme model. This included the NS maximum mass ($M_{\text{max}}$), radius
for a 2.07 M⊙ NS, radius (R_{1.4}), tidal deformability (\Lambda_{1.4}) for a 1.4 M⊙ NS, and the square of the speed of sound (C_{sym}^2) at the center of the maximum mass NS. We also compared the same for the prior set P2 in Figure 6. In both figures, the joint distribution of NS properties is represented by corner plots, which are obtained from Bayesian Inference posteriors. Each variable in a corner plot is plotted on one of the axes, and its distribution is shown by its diagonal. Off-diagonal plots show the joint distributions of each pair of variables. Using it to understand how variables are related is a useful tool.

In Table 8, we compared the relative % residuals of the NS-ANN model with the Skyrme model for both P1 and P2. It is worth noting that the results for P1 are very comparable to the original physics model. This is expected because P1 is the prior range in which the NS-ANN was originally trained. The relative % residuals for the NS maximum mass, radius for 2.07 and 1.4 M⊙ NS, and tidal deformability for 1.4 M⊙ NS are ~ 1%, while the square

| Input | Output |
|-------|--------|
| NMP | Skyrme | ML | \sigma_{rr} |
| | median | 90% CI | min | max | | median | 90% CI | min | max | | median | 90% CI | min | max |
| \(e_0\) | -16.01 | -16.39 | -15.59 | | | 2.09 | 1.87 | 2.24 | 2.08 | 1.89 | 2.24 | 2.2 |
| \(\rho_0\) | 0.160 | 0.152 | 0.168 | | | 11.74 | 9.77 | 11.85 | 10.75 | 9.80 | 11.84 | 1.5 |
| \(K_0\) | 231 | 200 | 261 | | | 12.82 | 11.56 | 14.10 | 12.81 | 11.60 | 14.02 | 2.1 |
| Skyrme \(Q_0\) | -292 | -457 | -140 | | | 3279 | 1994 | 5323 | 3279 | 1973 | 5280 | 6.8 |
| \(J_{sym,0}\) | 32 | 27 | 37 | | | 32 | 27 | 37 | 32 | 27 | 37 | 7.0 |
| \(L_{sym,0}\) | 67 | 42 | 94 | | | 435 | 234 | 729 | 430 | 232 | 722 | 7.0 |
| \(K_{sym,0}\) | -63 | -169 | 73 | | | 63 | 21 | 118 | 63 | 22 | 116 | 15.5 |
| \(\rho_0\) | 0.153 | 0.147 | 0.158 | | | 2.14 | 2.02 | 2.35 | 2.21 | 1.95 | 2.39 | 9.4 |
| \(K_0\) | 232 | 201 | 275 | | | 12.63 | 12.07 | 13.22 | 12.26 | 11.35 | 13.46 | 7.7 |
| DDB \(Q_0\) | -108 | -256 | 129 | | | 2802 | 2171 | 3702 | 2829 | 1869 | 3911 | 16.5 |
| \(J_{sym,0}\) | 32 | 29 | 35 | | | 455 | 339 | 626 | 385 | 247 | 582 | 34.7 |
| \(L_{sym,0}\) | 41 | 24 | 65 | | | 79 | 52 | 125 | 70 | 28 | 110 | 65.4 |
| \(K_{sym,0}\) | -116 | -150 | 73 | | | 246 | 216 | 380 | 257 | 164 | 391 | 16.5 |

Table 7. The median and the minimum (min) and maximum (max) values of the associated 90% confidence intervals for the NMPs and the NS properties obtained from Skyrme, DDB, and DDB-FL datasets. We also indicate these quantities for maximum relative residual errors \(\sigma_{rr}\) on the NS properties.
Table 8. The maximum relative residual for the maximum mass of a neutron star (NS), denoted as \( M_{\text{max}} \), is attained. Additionally, we determine the radius \( R_{2.07} \) for a 2.07 solar mass \((M_{\odot})\) NS, the radius \( R_{1.4} \) and the dimensionless tidal deformability \( \Lambda_{1.4} \) for a 1.4 \( M_{\odot} \) NS, as well as the squared speed of sound at the core of the NS with maximum mass. These quantities are obtained using the prior sets P1 and P2.

| Units       | Model | P1   | P2   |
|-------------|-------|------|------|
| M_{\odot}   | ANN   | 2.23 | 3.08 |
| km          | Skyrme| 16h  | 19h  |
| km          |       | 27 min | 55 min |
| ...         |       |       |      |
| c^2         |       |       |      |

Table 9. CPU inference time estimates for the ANN model and a Skyrme model to infer NS observations from a set of NMPs for prior sets P1 and P2. The timing tests were performed on a 12-core Intel i7-8700K CPU @ 3.70 GHz. The inference is performed with a batch size of one.

of the speed of sound is \( \sim 3\% \) at the center of the maximum mass NS. However, in the case of P2, all the residuals are noticeably increased, with the tidal deformability of 1.4 \( M_{\odot} \) NS and the square of the speed of sound at the center of the maximum mass NS being around \( \sim 5\% \). It should be emphasized that the neural network model can be trusted for interpolation but used with caution for extrapolation. The main advantage of this type of ANN model is that it can produce very fast draft results for computationally expensive calculations.

3.3 CPU Time
In Table 9, we compared the CPU time for NS-ANN and Skyrme model for both prior sets. The timing tests were performed on a 12-core Intel i7-8700K CPU @ 3.70 GHz. It is noteworthy that the NS-ANN model is approximately 500 times faster than the Skyrme model. It is exciting that the neural network model can be employed in computationally expensive calculations to obtain fast draft results. This will help in reducing carbon emissions and help reduce the environmental cost of using high computational power for the preliminary testing of any model.

4 Conclusions
We have demonstrated the application of ANNs to analyze NS observables like the radius, mass, and tidal deformability from a set of seven parameters, or nuclear saturation properties that characterize the equation of state of cold dense matter. We have shown that a neural network that models such a mapping is able to learn the microphysics information of finite nuclei, which are the intercorrelations arising between the NMPs. Moreover, we delved into the utilization of Artificial Neural Networks (ANN) as a complement to theoretical models, enabling them to undertake computationally demanding tasks, such as Bayesian Inference studies to constrain nuclear models. This study demonstrates that in situations where speed and computational efficiency are desired, a trained neural network can function as a surrogate for traditional physics-based EoS models, and possibly also provide insights that might be
Figure 5. Comparison of neutron star properties obtained from Bayesian posterior for prior set P1 with ANN and Skyrme model, including NS maximum mass ($M_{\text{max}}$), radius for a $2.07 \, M_\odot$ NS, radius ($R_{1.4}$) and tidal deformability ($\Lambda_{1.4}$) for a $1.4 \, M_\odot$ NS, and square of the speed of sound ($C_s^2$) at the center of the maximum mass NS. The vertical lines represent the 68% CIs, and the light and dark intensities represent the $1\sigma$, $2\sigma$, and $3\sigma$ CIs, respectively.

difficult to obtain otherwise. However, it is essential to clarify that this methodology only complements other approaches, and does not seek to replace them in any form.

Let us briefly summarize what we have done in this work. First, we generate a set of pseudo-NMP data using the Skyrme model. A suitable set of NMPs are chosen so that the resulting neutron star EoSs are consistent with the currently observed maximum mass of $\sim 2M_\odot$ and satisfy causality constraints. Next, we train an artificial neural network on this data set to learn a non-linear mapping between the set of nuclear matter parameters and NS observables. We demonstrate that the ANN is capable of inferring, with reasonable accuracy, NS observables from empirical parameters, as compared to conventional physics models which require the computation of an EoS followed by tedious equation-solving. This NS-ANN model is then validated to ensure that it satisfies the microphysics information of
Figure 6. Comparison of neutron star properties obtained from Bayesian posterior for prior set P2 with ANN and Skyrme model, including NS maximum mass ($M_{\text{max}}$), radius for a 2.07 M⊙ NS, radius ($R_{1.4}$) and tidal deformability ($\Lambda_{1.4}$) for a 1.4 M⊙ NS, and square of the speed of sound ($C_s^2$) at the center of the maximum mass NS. The vertical lines represent the 68% CIs, and the light and dark intensities represent the 1σ, 2σ, and 3σ CIs, respectively.

finite nuclear matter. More specifically, we study whether a trained ANN model is able to capture correlations between the tidal deformability for a NS with mass 1.4M⊙, $\Lambda_{1.4}$ and the symmetry energy slope $L_{\text{sym,0}}$ as well as its curvature $K_{\text{sym,0}}$. We find that the NS-ANN model learns a mapping that is sensitive to $L_{\text{sym,0}} - K_{\text{sym,0}}$ correlations in agreement with previous studies. Using the Bayesian inference framework, we determine the extent to which the neural network can replicate the physics model, and then utilize the trained neural network for inference tasks, which typically involve computationally expensive calculations. By doing so, we are able to achieve results that are comparable to the physics model while significantly reducing the computational time required.

Presently, our framework is a proof-of-concept that demonstrates the applicability of ANNs to NS physics. For a more realistic application of our framework, empirical uncertain-
ties ought to be considered. This can be achieved, for example, by using a class of ANNs called Bayesian Neural Networks to perform the inference task. Bayesian networks have the ability to cast the problem into a probabilistic domain by inferring probability distributions over a prior of NMP inputs. A model which predicts uncertainties will also be able to potentially further reveal the effects of NS observational constraints on the NMPs. Presently, the entire domain of validity of the ANN cannot be automatically ensured. Due to a lack of training data for some parts in the modeled output space, the ANN might be unable to accurately predict regions that have not been learned.

This work is limited only to hadronic NS compositions, i.e., the set of $\beta$-equilibrated EoSs employed in this work are composed of neutrons, protons, electrons, and muons. Present observational constraints on NS properties cannot rule out the possibility of exotic degrees of freedom or deconfined quark phases inside the NS core. In future work, a detailed and systematic analysis with an ANN trained on a set of EoSs with different compositions of particles is required to investigate the uncertainties on higher-order NMPs with available observational constraints. The ANN map can also be further extended to determine the required number of NS observations and their precision to conclude the existence of quark phases inside the NS core.

Data availability

The trained network and corresponding data sets are made publicly available through a GitHub repository. 3

Acknowledgements

T.M. would like to thank national funds from FCT (Fundação para a Ciência e a Tecnologia, I.P, Portugal) under Projects No. UID/FIS/04564/2019, No. UIDP/04564/2020, No. UIDB/04564/2020, and No. POCI-01-0145-FEDER-029912 with financial support from Science, Technology, and Innovation, in its FEDER component, and by the FCT/MCTES budget through national funds (OE). The authors acknowledge the Laboratory for Advanced Computing at the University of Coimbra for providing HPC resources that have contributed to the research results reported within this paper.

References

[1] N.K. Glendenning, *Compact stars: Nuclear physics, particle physics and general relativity*, Springer Science & Business Media (2012).

[2] P. Haensel, A.Y. Potekhin and D.G. Yakovlev, *Neutron stars 1: Equation of state and structure*, vol. 326, Springer, New York, USA (2007), 10.1007/978-0-387-47301-7.

[3] E.R. Most, L.R. Weih, L. Rezzolla and J. Schaffner-Bielich, *New constraints on radii and tidal deformabilities of neutron stars from GW170817*, *Phys. Rev. Lett.* 120 (2018) 261103 [1803.00549].

[4] S.B. Ruester, M. Hempel and J. Schaffner-Bielich, *The outer crust of non-accreting cold neutron stars*, *Phys. Rev. C* 73 (2006) 035804 [astro-ph/0509325].

[5] R.C. Tolman, *Static solutions of Einstein’s field equations for spheres of fluid*, *Phys. Rev.* 55 (1939) 364.

---

3 [https://github.com/ameya1101/NS-ANN](https://github.com/ameya1101/NS-ANN)
[25] S. Altiparmak, C. Ecker and L. Rezzolla, On the Sound Speed in Neutron Stars, 2203.14974.

[26] J. Margueron, R. Hoffmann Casali and F. Gulminelli, Equation of state for dense nucleonic matter from metamodeling. I. Foundational aspects, Phys. Rev. C 97 (2018) 025805 [1708.06894].

[27] J. Margueron, R. Hoffmann Casali and F. Gulminelli, Equation of state for dense nucleonic matter from metamodeling. II. Predictions for neutron star properties, Phys. Rev. C 97 (2018) 025806 [1708.06895].

[28] P. Demorest, T. Pennucci, S. Ransom, M. Roberts and J. Hessels, Shapiro Delay Measurement of A Two Solar Mass Neutron Star, Nature 467 (2010) 1081 [1010.5788].

[29] E. Fonseca et al., The NANOGrav Nine-year Data Set: Mass and Geometric Measurements of Binary Millisecond Pulsars, Astrophys. J. 832 (2018) 167 [1708.06895].

[30] NANOGrav collaboration, The NANOGrav 11-year Data Set: High-precision timing of 45 Millisecond Pulsars, Astrophys. J. Suppl. 235 (2018) 37 [1801.01837].

[31] J. Antoniadis et al., A Massive Pulsar in a Compact Relativistic Binary, Science 340 (2013) 6131 [1304.6875].

[32] E. Fonseca et al., Refined Mass and Geometric Measurements of the High-mass PSR J0740+6620, Astrophys. J. Lett. 915 (2021) L12 [2104.00880].

[33] R.W. Romani, D. Kandel, A.V. Filippenko, T.G. Brink and W. Zheng, PSR J1810+1744: Companion Darkening and a Precise High Neutron Star Mass, Astrophys. J. Lett. 908 (2021) L46 [2101.09822].

[34] J. Faber, Status of neutron star-black hole and binary neutron star simulations, Class. Quant. Grav. 26 (2009) 114004.

[35] M.D. Duez, Numerical relativity confronts compact neutron star binaries: a review and status report, Class. Quant. Grav. 27 (2010) 114002 [0912.3529].

[36] T.E. Riley et al., A NICER View of PSR J0030+0451: Millisecond Pulsar Parameter Estimation, Astrophys. J. Lett. 887 (2019) L21 [1912.05702].

[37] M.C. Miller et al., PSR J0030+0451 Mass and Radius from NICER Data and Implications for the Properties of Neutron Star Matter, Astrophys. J. Lett. 887 (2019) L24 [1912.05705].

[38] T.E. Riley et al., A NICER View of the Massive Pulsar PSR J0740+6620 Informed by Radio Timing and XMM-Newton Spectroscopy, Astrophys. J. Lett. 918 (2021) L27 [2105.06980].

[39] T. Malik, M. Ferreira, B.K. Agrawal and C. Providência, Relativistic description of dense matter equation of state and compatibility with neutron star observables: a Bayesian approach, 2201.12552.

[40] E. Annala, T. Gorda, A. Kurkela, J. Nättäli and A. Vuorinen, Evidence for quark-matter cores in massive neutron stars, Nature Phys. 16 (2020) 907 [1903.09121].
K. Hebeler, J.M. Lattimer, C.J. Pethick and A. Schwenk, *Equation of state and neutron star properties constrained by nuclear physics and observation*, *Astrophys. J.* **773** (2013) 11 [1303.4662].

C. Drischler, K. Hebeler and A. Schwenk, *Asymmetric nuclear matter based on chiral two- and three-nucleon interactions*, *Phys. Rev. C* **93** (2016) 054314 [1510.06728].

J.S. Read, B.D. Lackey, B.J. Owen and J.L. Friedman, *Constraints on a phenomenologically parameterized neutron-star equation of state*, *Phys. Rev. D* **79** (2009) 124032 [0812.2163].

F. Ozel, D. Psaltis, T. Guver, G. Baym, C. Heinke and S. Guillot, *The Dense Matter Equation of State from Neutron Star Radius and Mass Measurements*, *Astrophys. J.* **820** (2016) 28 [1505.05155].

C.A. Raithel, F. Özel and D. Psaltis, *From Neutron Star Observables to the Equation of State. II. Bayesian Inference of Equation of State Pressures*, *Astrophys. J.* **844** (2017) 156 [1704.00737].

L. Lindblom, *Spectral Representations of Neutron-Star Equations of State*, *Phys. Rev. D* **82** (2010) 103011 [1009.0738].

L. Lindblom and N.M. Indik, A Spectral Approach to the Relativistic Inverse Stellar Structure Problem, *Phys. Rev. D* **86** (2012) 084003 [1207.3744].

R. Essick, P. Landry and D.E. Holz, *Nonparametric Inference of Neutron Star Composition, Equation of State, and Maximum Mass with GW170817*, *Phys. Rev. D* **101** (2020) 063007 [1910.09740].

P. Landry, R. Essick and K. Chatziioannou, *Nonparametric constraints on neutron star matter with existing and upcoming gravitational wave and pulsar observations*, *Phys. Rev. D* **101** (2020) 123007 [2003.04880].

M.-Z. Han, J.-L. Jiang, S.-P. Tang and Y.-Z. Fan, *Bayesian Nonparametric Inference of the Neutron Star Equation of State via a Neural Network*, *Astrophys. J.* **919** (2021) 11 [2103.05408].

K. He, X. Zhang, S. Ren and J. Sun, *Deep residual learning for image recognition*, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016, DOI.

T. Young, D. Hazarika, S. Poria and E. Cambria, Recent trends in deep learning based natural language processing [review article], *IEEE Computational Intelligence Magazine* **13** (2018) 55.

D. George and E.A. Huerta, *Deep Neural Networks to Enable Real-time Multimessenger Astrophysics*, *Phys. Rev. D* **97** (2018) 044039 [1701.00008].

D. George and E.A. Huerta, *Deep Learning for Real-time Gravitational Wave Detection and Parameter Estimation: Results with Advanced LIGO Data*, *Phys. Lett. B* **778** (2018) 64 [1711.03121].

M. Carrillo, M. Gracia-Linares, J.A. González and F.S. Guzmán, *Parameter estimates in binary black hole collisions using neural networks*, *Gen. Rel. Grav.* **48** (2016) 141 [1608.02491].

A.J. Larkoski, I. Moult and B. Nachman, *Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Machine Learning*, *Phys. Rept.* **841** (2020) 1 [1709.04464].

D. Guest, K. Cranmer and D. Whiteson, *Deep Learning and its Application to LHC Physics*, *Ann. Rev. Nucl. Part. Sci.* **68** (2018) 161 [1806.11484].

D. Bourilkov, *Machine and Deep Learning Applications in Particle Physics*, *Int. J. Mod. Phys. A* **34** (2020) 1930019 [1912.08245].

P. Bedaque, A. Boehnlein, M. Cromaz, M. Diefenthaler, L. Elouadrhiri, T. Horn et al., *Report from the A.I. for nuclear physics workshop*, *2006.05422*. 

– 22 –
[64] Y. Fujimoto, K. Fukushima and K. Murase, Mapping neutron star data to the equation of state using the deep neural network, *Phys. Rev. D* **101** (2020) 054016 [1903.03400].

[65] F. Morawski and M. Bejger, Neural network reconstruction of the dense matter equation of state derived from the parameters of neutron stars, *Astron. Astrophys.* **642** (2020) A78 [2006.07194].

[66] M.U. Anil, K. Banerjee, T. Malik and C. Providência, The neutron star outer crust equation of state: a machine learning approach, *JCAP* **01** (2022) 045 [2004.14196].

[67] Y. Fujimoto, K. Fukushima and K. Murase, Methodology study of machine learning for the neutron star equation of state, *Phys. Rev. D* **98** (2018) 023019 [1711.06748].

[68] I. Vidana, C. Providência, A. Polls and A. Rios, Density dependence of the nuclear symmetry energy: A Microscopic perspective, *Phys. Rev. C* **80** (2009) 045806 [0907.1165].

[69] M.a. Ferreira and C. Providência, Unveiling the nuclear matter EoS from neutron star properties: a supervised machine learning approach, *JCAP* **07** (2021) 011 [1910.05554].

[70] T. Hinderer, Tidal Love numbers of neutron stars, *Astrophys. J.* **677** (2008) 1216 [0711.2420].

[71] K. Hornik, M. Stinchcombe and H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* **2** (1989) 359.

[72] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals and Systems* **2** (1989) 303.

[73] K. Jarrett, K. Kavukcuoglu, M. Ranzato and Y. LeCun, What is the best multi-stage architecture for object recognition?, in 2009 IEEE 12th International Conference on Computer Vision, pp. 2146–2153, 2009, DOI.

[74] V. Nair and G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML’10, (Madison, WI, USA), p. 807–814, Omnipress, 2010.

[75] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press (2016).

[76] W. Wei, E.A. Huerta, M. Yun, N. Loutrel, M.A. Shaikh, P. Kumar et al., Deep Learning with Quantized Neural Networks for Gravitational-wave Forecasting of Eccentric Compact Binary Coalescence, *Astrophys. J.* **919** (2021) 82 [2012.03963].

[77] A.J.K. Chua and M. Vallisneri, Learning Bayesian posteriors with neural networks for gravitational-wave inference, *Phys. Rev. Lett.* **124** (2020) 041102 [1909.05966].

[78] T. Chen and C. Guestrin, XGBoost: A scalable tree boosting system, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16, (New York, NY, USA), pp. 785–794, ACM, 2016, DOI.

[79] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al., Scikit-learn: Machine learning in python, *Journal of Machine Learning Research* **12** (2011) 2825.

[80] Y. LeCun, Y. Bengio and G. Hinton, Deep learning, *Nature Cell Biology* **521** (2015) 436.

[81] F. Chollet et al., “Keras.” [https://keras.io](https://keras.io), 2015.
[85] S. Wesolowski, N. Klco, R.J. Furnstahl, D.R. Phillips and A. Thapaliya, Bayesian parameter estimation for effective field theories, *J. Phys. G* **43** (2016) 074001 [1511.03618].

[86] R.J. Furnstahl, N. Klco, D.R. Phillips and S. Wesolowski, Quantifying truncation errors in effective field theory, *Phys. Rev. C* **92** (2015) 024005 [1506.01343].

[87] G. Ashton et al., BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy, *Astrophys. J. Suppl.* **241** (2019) 27 [1811.02042].

[88] J. Skilling, Nested Sampling, in Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, R. Fischer, R. Preuss and U.V. Toussaint, eds., vol. 735 of American Institute of Physics Conference Series, pp. 395–405, Nov., 2004, DOI.

[89] E. Higson, W. Handley, M. Hobson and A. Lasenby, Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation, *Statistics and Computing* **29** (2019) 891.

[90] J.S. Speagle, dynesty: a dynamic nested sampling package for estimating bayesian posteriors and evidences, *Monthly Notices of the Royal Astronomical Society* **493** (2020) 3132.

[91] M.C. Miller et al., The Radius of PSR J0740+6620 from NICER and XMM-Newton Data, *Astrophys. J. Lett.* **918** (2021) L28 [2105.06979].

[92] M. Ferreira, M. Fortin, T. Malik, B.K. Agrawal and C. Providência, Empirical constraints on the high-density equation of state from multimessenger observables, *Phys. Rev. D* **101** (2020) 043021 [1912.11131].

[93] B.K. Pradhan, D. Chatterjee, R. Gandhi and J. Schaffner-Bielich, Role of vector self-interaction in neutron star properties, *Nucl. Phys. A* **1030** (2023) 122578 [2209.12657].

[94] T. Malik, M. Ferreira, M.B. Albino and C. Providência, Spanning the full range of neutron star properties within a microscopic description, *Phys. Rev. D* **107** (2023) 103018 [2301.08169].

[95] J. Zhou, J. Xu and P. Papakonstantinou, Bayesian inference of neutron-star observables based on effective nuclear interactions, *Phys. Rev. C* **107** (2023) 055803 [2301.07904].