Calibrating global behaviour of equation of state by combining nuclear and astrophysics inputs in a machine learning approach

Sk Md Adil Imam[1,2],* Prafulla Saxena[3], and B. K. Agrawal[1,2]
[1]Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700064, India
[2]Homi Bhabha National Institute, Anushakti Nagar, Mumbai 400094, India and
[3]Malaviya National Institute of Technology, Jaipur, India
(Dated: January 2024)

We implemented symbolic regression techniques to identify suitable analytical functions that map various properties of neutron stars (NSs), obtained by solving the Tolman-Oppenheimer-Volkoff (TOV) equations, to a few key parameters of the equation of state (EoS). These symbolic regression models (SRMs) are then employed to perform Bayesian inference with a comprehensive dataset from nuclear physics experiments and astrophysical observations. The posterior distributions of EoS parameters obtained from Bayesian inference using SRMs closely match those obtained directly from the solutions of TOV equations. Our SRM-based approach is approximately 100 times faster, enabling efficient Bayesian analyses across different combinations of data to explore their sensitivity to various EoS parameters within a reasonably short time.

I. INTRODUCTION

Understanding the behavior of matter at densities beyond those found in atomic nuclei is crucial for advancements in modern astrophysics and nuclear physics. The dead remnants of stellar evolution, neutron stars (NS) are perfect laboratories for studying particles in dense environments as their central densities go up to 5-8 times the nuclear saturation density ($\rho_0 \sim 0.16$ fm^{-3}). So multi-messenger observations related to NS properties such as gravitational mass, radius, and tidal deformability, provide a way to study the matter at extreme conditions.

In this study, we refrain from limiting our parameter space, in contrast to the approach outlined in Ref.[1]. Instead, we employ a symbolic regression technique to construct analytical expressions of neutron star properties in terms of a few key EoS parameters called the nuclear matter parameters (NMPs). In the present investigation, we have developed analytical expressions for NS maximum mass, radii, and tidal deformabilitie. These expressions are then incorporated into a Bayesian framework, utilizing data from nuclear physics experiments and astrophysical observations, to probe the overall behavior of the EoS.

Symbolic regression [2–5], a machine learning approach is utilized to map the NS properties to a few key EoS parameters. The EoS are constructed using $\frac{n}{3}$ expansion as detailed in [6–8]. These EoSs are utilized in computing neutron star properties through the solutions of the TOV equations [9–11]. The chosen EoS satisfy conditions of : (i)thermodynamic stability, (ii)positive symmetry energy, (iii)causality of the speed of sound, and (iv) maximum mass of stable non-rotating neutron stars, $M_{max} \geq 2.15\ M_\odot$. To perform symbolic regression, one has to provide the features (input variables) and their corresponding target values (output variables). A K-fold cross-validation procedure is often used to assess model performance and avoid overfitting. This procedure yields multiple equations in each of the 'K' folds. To choose the best equation, the goodness of fit and model complexity are good metrics to filter out among various searched equations. Several metrics are commonly used for this purpose, two of them are (i) R^2 which is defined as

$$R^2 \ = \ 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \tag{1}$$

Here y_i is the observed or actual value, $\hat{y}_i$ is the predicted value for the i^{th} data point and $\bar{y}$ is the mean of the 'n' observed values. It ranges from 0 to 1, with higher values indicating a better fit.
(ii) RMSE which is defined as,

$$\text{RMSE} = \sqrt{\frac{SS_{residual}}{n}} \tag{2}$$

This also provide information about the accuracy of a regression model, with lower values indicating a better fit.

Once these equations are constructed can be used to compute the NS properties in the Bayesian inference without computing EoS and solving TOV equations. This makes the Bayesian inference much faster than the traditional approach.

TABLE I: Prior distributions of the NMPs (θ). Here ρ_0 = 0.16 fm^{-3}, e_0 = -16.0 MeV. All other parameters are uniformly distributed between a minimum (θ_{min}) and a maximum (θ_{max}). For the symbolic regression all the NMPs are reduced to the standard variable, $\hat{\theta} = \frac{(\theta - \bar{\theta})}{\sigma_\theta}$ with mean $\bar{\theta}$ and standard deviation σ_θ. The unit of all the NMPs are in MeV.

	K_0	Q_0	Z_0	J_0	L_0	$K_{sym,0}$	$Q_{sym,0}$	$Z_{sym,0}$
θ_{min}	180	-1500	-3000	25	20	-300	300	-3000
θ_{max}	330	1500	3000	40	120	300	1500	3000
$\bar{\theta}$	256.54	-0.52	-1049.42	31.99	76.04	-41.12	987.38	200.38
σ_θ	48.99	266.00	1558.57	4.28	27.57	154.15	331.14	1694.79

*Electronic address: mdadil.imam@saha.ac.in

II. RESULTS AND DISCUSSIONS

We construct the Symbolic Regression Models(SRMs) to express the NS maximum mass, radii and tidal deformabilities in terms of various nuclear matter parameters. Once SRMs are obtained using the solutions of the TOV equations, they can facilitate constraining the EoS through Bayesian inference. We have modelled NS radii and tidal deformabilities with masses between 1.1 - 2.15 $M_\odot$ in the interval of 0.05 $M_\odot$ and maximum mass ($M_{\rm max}$) of NS in terms of the nuclear matter parameters directly through the symbolic regression. The equations are selected based on a balance between achieving a high R^2 (Eq 1), low RMSE (Eq 2), and less model complexity. These equations are provided in the Appendix of ref[12]. Few such equations are -

$$M_{\rm max} = 0.14\hat{K}_0 - 0.02\hat{L}_0 + 0.27\hat{Q}_0 + 0.16\hat{Z}_0 + 2.33 \quad (3)$$

$$R_{1.4} = -0.31(\hat{J}_0 - \hat{Q}_0 - \hat{Q}_{\rm sym0}) - 0.62\hat{K}_{\rm sym0} + 1.75\hat{L}_0 + 14.22 \quad (4)$$

$$\Lambda_{1.4} = -54.02(\hat{J}_0 - \hat{K}_0 - \hat{Q}_0 - \hat{Q}_{\rm sym0}) + 76.09\hat{K}_{\rm sym0} + 178.12\hat{L}_0 + 737.93 \quad (5)$$

The values of R^2 are $\sim$ 0.99, 0.93 and 0.92 for NS radii, tidal deformabilities and maximum mass, respectively. These equation are then employed in the Bayesian inference. In the Bayesian inference we employed nested sampling with priors as given in TableI and likelihood functions are constructed according to the data. The data contains : (i) 12 experimental data on symmetry energy, pressure due to symmetry energy and symmetric nuclear matter as in [12].
(ii) 2 empirical data on pure neutron matter and slope of incompressibility at the crossing density (ρ_c = 0.1 fm^{-3}) [13, 14]. It is worth mentioning that these two data along with experimental data are able constrain the NS EoS further [12]
(iii) Astrophysical data contains mass-radius posterior distribution for PSR J0030+0451 [15, 16] and PSR J0740+6620 [17, 18] and posterior distribution for dimensionless tidal deformability for binary neutron star components from the GW170817 event [19].

Bayesian inference is employed with both SRMs and solutions from TOV equations to calculate various properties of neutron stars. The figure 1 displays L_0 is strongly correlated with J_0 and $K_{\rm sym0}$ with Pearson's correlation coefficient r $\sim$ 0.9. There is also a stronger correlation between J_0 and $K_{\rm sym0}$ (r $\sim$ 0.7). Any other pairs of NMPs do not show significant correlations. It is also clear that the results of Bayesian inference using SRMs and solutions of TOV equations are very similar.

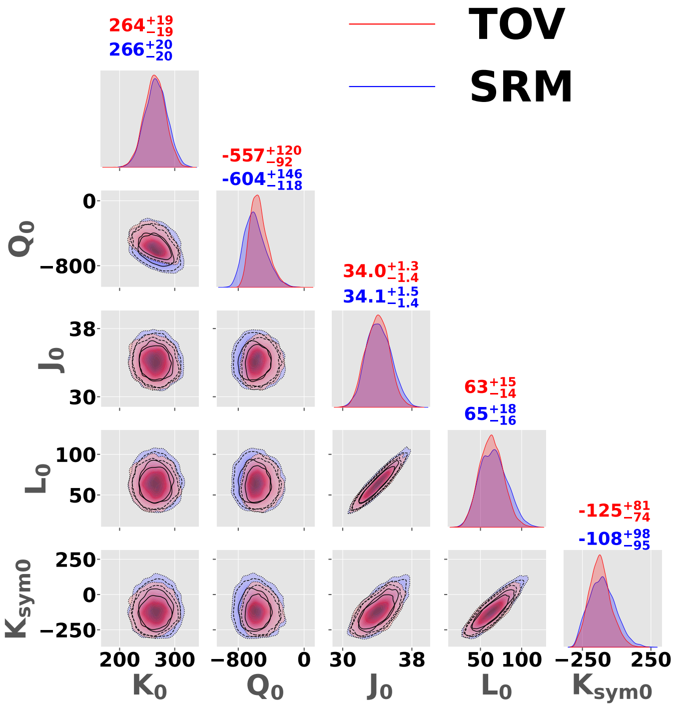

FIG. 1: The marginalized posterior distributions of the NMPs obtained in the Bayesian inference with the realistic data. We present the results obtained using SRMs (blue) and solutions of TOV equations (red) to compute NS properties in the inference.

[1] S. M. A. Imam, A. Mukherjee, B. K. Agrawal, and G. Banerjee, Phys. Rev. C **109**, 025804 (2024).
[2] J. R. Koza, *Genetic programming: on the programming of computers by means of natural selection* (MIT Press, Cambridge, MA, USA, 1992).
[3] M. Schmidt and H. Lipson, Science (New York, N.Y.) **324**, 81—85 (2009).
[4] S. Sun, R. Ouyang, B. Zhang, and T.-Y. Zhang, MRS Bulletin **44**, 559 (2019).
[5] N. Makke and S. Chawla, Interpretable scientific discovery with symbolic regression: A review (2023), arXiv:2211.10873 [cs.LG] .
[6] J. M. Lattimer and M. Prakash, Phys. Rept. **621**, 127 (2016).
[7] H. Gil, P. Papakonstantinou, C. H. Hyun, T.-S. Park, and Y. Oh, Acta Phys. Polon. B **48**, 305 (2017).
[8] S. M. A. Imam, N. K. Patra, C. Mondal, T. Malik, and B. K. Agrawal, Phys. Rev. C **105**, 015806 (2022).
[9] J. R. Oppenheimer and G. M. Volkoff, Phys. Rev. **55**, 374 (1939).
[10] R. C. Tolman, Phys. Rev. **55**, 364 (1939).
[11] T. Hinderer, Astrophys. J. **677**, 1216 (2008).
[12] S. M. A. Imam, P. Saxena, T. Malik, N. K. Patra, and B. K. Agrawal, (2024), arXiv:2407.08553 [nucl-th] .
[13] B. A. Brown, Phys. Rev. Lett. **111**, 232502 (2013).
[14] E. Khan and J. Margueron, Phys. Rev. C **88**, 034319 (2013).
[15] T. Riley *et al.*, Astrophys. J. Lett. **887**, L21 (2019).
[16] M. C. Miller *et al.*, Astrophys. J. **887**, L24 (2019).
[17] M. C. Miller *et al.*, Astrophys. J. Lett. **918**, L28 (2021).
[18] T. E. Riley *et al.*, Astrophys. J. Lett. **918**, L27 (2021).
[19] B. P. Abbott *et al.*, Phys. Rev. X **9**, 011001 (2019).