Unveiling Neutron Star Composition and Observables: A Comprehensive Study using Deep Bayesian Neural Networks

Tuhin Malik

CFisUC, University of Coimbra, Portugal

October 10, 2023

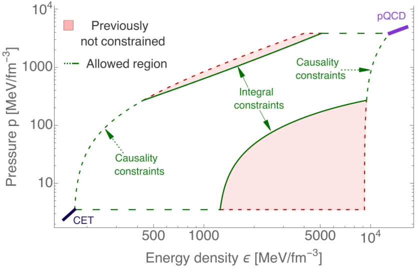
Fundação para a Ciência e a Tecnologia

ECT*
EUROPEAN CENTRE FOR THEORETICAL STUDIES IN NUCLEAR PHYSICS AND RELATED AREAS

Outline

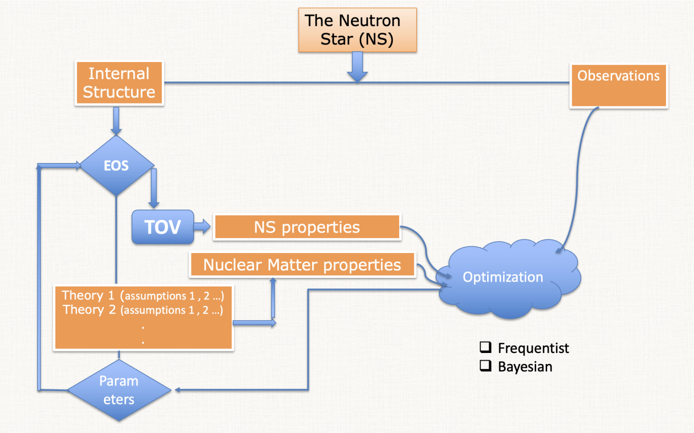
The dense matter equation of state (EOS)

- A neutron star (NS), also known as a pulsar, is one of the densest and most compact objects in the universe.
- A significant probe to reduce uncertainty can be the NS maximum mass, radii, moments of inertia, and tidal Love numbers, which are all accessible to observation.
- The NS core composition remains a mystery

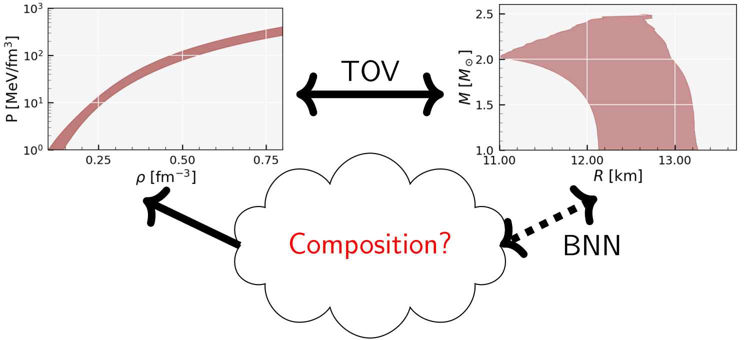


Phys. Rev. Lett. 128, 202701 (2022), 2111.05350

The workflow



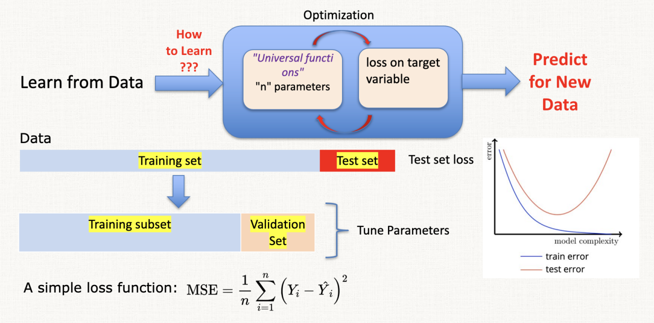
One-to-one correspondence



- General relativity guarantees a unique one-to-one correspondence between static observables of neutron stars(NSs) accessible by multi-messenger astronomy, such as mass-radius or tidal deformability, and the equation of state (EOS) of beta equilibrated matter.

Machine Learning

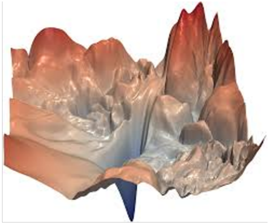
Machine Learning aims to build a mathematical function that solves a human task.

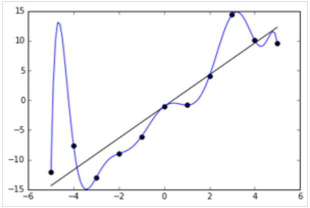


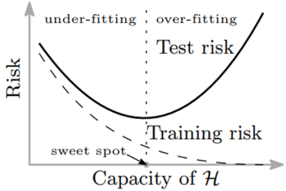
Challenges of ML

- Not enough training data.
- Poor Quality of data.
- Irrelevant features.
- Overfitting and Underfitting.

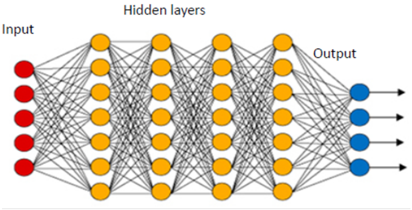
Global Minima in the loss function may look like this:







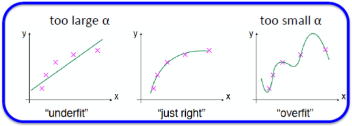
Deep Neural Network



Loss function: $L(\mathbf{W}) = \frac{1}{N_{\text{train}}} \sum_{t=1}^{N_{\text{train}}} \left(F(\mathbf{x}_i) - y_t\right)^2 + \alpha \|\mathbf{W}\|_2^2$

mean squared error

regularization (penalizes large weights)



Relativistic description of the neutron star equation of state

(a Bayesian approach)

EOS: relativistic mean field description

RMF Lagrangian for stellar matter

- Lagrangian density
  - Lorentz-covariant Lagrangian with baryon densities and meson fields
  - causal by construction

$$\mathcal{L} = \mathcal{L}_N + \mathcal{L}_M + \mathcal{L}_{NL},$$

- Baryonic contribution:

$$\mathcal{L}_N = \bar{\Psi}\Big[\gamma^{\mu}\left(i\partial_{\mu} - \Gamma_{\omega} A_{\mu}^{(\omega)} - \Gamma_{\varrho} \boldsymbol{t}\cdot\boldsymbol{A}_{\mu}^{(\varrho)}\right) - (m - \Gamma_{\sigma}\phi)\Big]\Psi,$$

- Meson contribution

$$\mathcal{L}_M = \frac{1}{2}\left[\partial_{\mu}\phi\partial^{\mu}\phi - m_{\sigma}^2\phi^2\right] - \frac{1}{4}F_{\mu\nu}^{(\omega)}F^{(\omega)\mu\nu} + \frac{1}{2}m_{\omega}^2\omega_{\mu}\omega^{\mu} - \frac{1}{4}\boldsymbol{F}_{\mu\nu}^{(\varrho)}\cdot\boldsymbol{F}^{(\varrho)\mu\nu} + \frac{1}{2}m_{\varrho}^2\boldsymbol{\varrho}_{\mu}\cdot\boldsymbol{\varrho}^{\mu}.$$

- Non-linear meson terms

$$\mathcal{L}_{NL} = -\frac{1}{3}bg_{\sigma}^3(\sigma)^3 - \frac{1}{4}cg_{\sigma}^4(\sigma)^4 + \frac{\xi}{4!}(g_{\omega}\omega_{\mu}\omega^{\mu})^4 + \Lambda_{\omega}g_{\varrho}^2\boldsymbol{\varrho}_{\mu}\cdot\boldsymbol{\varrho}^{\mu}g_{\omega}^2\omega_{\mu}\omega^{\mu}$$

Nuclear matter properties at saturation

- Taylor expansion, parabolic approximation

$$
\begin{aligned}
\frac{E_{\text{nuc}}}{A}(n,\delta) &= \frac{E_{\text{SNM}}}{A}(n) + S(n)\,\delta^2, \\
S(n) &= \frac{1}{2}\,\frac{\partial^2 E_{\text{nuc}}/A}{\partial\delta^2}\Big|_{\delta=0}, \\
\frac{E_{\text{SNM}}}{A}(n) &= E_0 + \frac{K_0}{2}\eta^2 + \frac{J_0}{3!}\eta^3 + \frac{Z_0}{4!}\eta^4, \\
S(n) &= E_{\text{sym}} + L_{\text{sym}}\eta + \frac{K_{\text{sym}}}{2}\eta^2 + \frac{J_{\text{sym}}}{3!}\eta^3 + \frac{Z_{\text{sym}}}{4!}\eta^4,
\end{aligned}
$$

$$\delta = (n_p - n_n)/n, \qquad \eta = (n - n_0)/(3n_0)$$

Bayesian estimation of model parameters

Bayesian Inference:

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

- The $\boldsymbol{\theta}$ is the model parameter vector and D is the set of fit data.
- $P(\boldsymbol{\theta} \mid D)$ is the joint posterior distribution of the parameters.
- $\mathcal{L}(D \mid \boldsymbol{\theta})$ is the likelihood function.
- $P(\boldsymbol{\theta})$ is the prior distribution for the model parameters.
- $\mathcal{Z}$ is the evidence. It can be obtained by complete marginalization of the likelihood function.

The marginalized posterior distribution for a parameter θ_i:

$$P(\theta_i \mid D) = \int P(\boldsymbol{\theta} \mid D) \prod_{k \neq i} d\theta_k$$

Gaussian likelihood function

$$\mathcal{L}(D \mid \boldsymbol{\theta}) = \prod_j \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{1}{2}\left(\frac{d_j - m_j(\theta)}{\sigma_j}\right)^2}$$

- The index j runs over all the data points.
- The d_j and m_j are the data and corresponding model values, respectively.
- The σ_j are the uncertainties for every data point.

Sampling EOS

Markov Chain Monte Carlo sampling (we do):

- Cost-function guided random walk
- Sample the posterior

we use the nested sampling algorithm, first proposed in J Skilling, American Institute of Physics Conference Series, Vol. 735, edited by R. Fischer, R. Preuss, and U. V. Toussaint (2004) pp. 395–405.

- suitable for low-dimensional problems
- approximately 25K samples we have obtained in the posterior

Public available data: 10.5281zenodo.7854112

Constraints			
Quantity		Value/Band	Ref
NMP (MeV)	ρ_0	0.153 ± 0.005	Typel & Wolter (1999)
	ϵ_0	-16.1 ± 0.2	Dutra et al. (2014)
	K_0	230 ± 40	Todd-Rutel & Piekarewicz (2005); Shlomo et al. (2006)
	$J_{sym,0}$	32.5 ± 1.8	Essick et al. (2021a)
PNM (MeV fm^{-3})	$P(\rho)$	$2 \times N^3LO$	Hebeler et al. (2013)
NS mass ($M_\odot$)	M_{max}	>2.0	Fonseca et al. (2021)

Monte Carlo sampling (we dont):

- Generate random uniform samples in the parameter hyperspace.
- Apply filter
- Analyze filtered samples' properties

Results

Structuring of Data

- We generate two types of datasets that share the output **Yi** structure but with different input **Xi** structures.
- The **Y** as proton fraction y_p or square of speed of sound v_s^2 at 15 fixed baryonic densities n_k, e.g., $y_p(n) = [y_p(n_1), y_p(n_2), ..., y_p(n_{15})]$.
- $X1 = [M_1, ..., M_5, R_1, ..., R_5]$ corresponding to five $M_i(R_i)$ simulated observations
- $X2 = [M_1, .., M_5, R_1, .., R_5, M_1^{\prime}, .., M_5^{\prime}, \Lambda_1, .., \Lambda_5]$ corresponding to five $M_i(R_i)$ and five $\Lambda_j(M_j^{\prime})$ simulated observations.

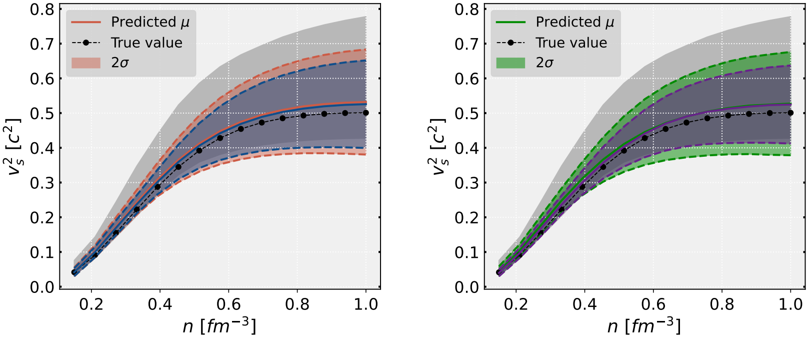
For each EoS, we randomly select

$$M_i^{(0)} \sim \mathcal{U}(1, M_{\text{max}}) \quad (\text{in units of } \mathrm{M}_\odot)$$
$$R_i \sim \mathcal{N}\left(R\left(M_i^{(0)}\right), \sigma_R^2\right)$$
$$M_i \sim \mathcal{N}\left(M_i^{(0)}, \sigma_M^2\right), \quad i = 1, .., 5$$

Generation parameters for each dataset. $\hat{\sigma}(M_j)$ denotes the standard deviation of $\Lambda(M)$ calculated on the train set.

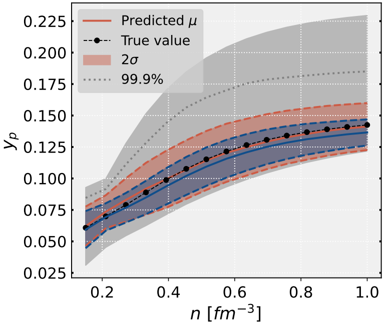
Dataset	σ_M [$M_\odot$]	σ_R [km]	$\sigma_\Lambda(M_j)$
1	0.05	0.15	—
2	0.1	0.3	—
3	0.1	0.3	$0.5\hat{\sigma}(M_j)$
4	0.1	0.3	$2\hat{\sigma}(M_j)$

The BNNs predictions for square of speed of sound v_s^2



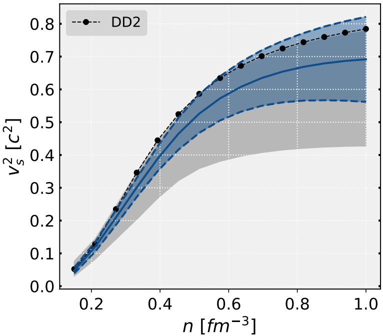
The models trained on datasets 1 (blue) and 2 (orange) are in the left figure while datasets 3 (purple) and 4 (green) models are in the right figure. The prediction mean values (solid lines) and 2σ confidence intervals are shown. The true values are shown in black dots and the range of $v_s^2(n)$ from the train set is indicated by the grey region.

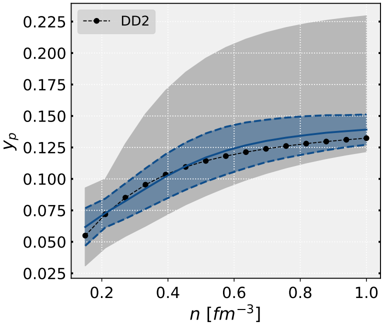
The BNNs predictions for proton fraction y_p



The models trained on datasets 1 (blue) and 2 (orange) are in the left figure while datasets 3 (purple) and 4 (green) models are in the right figure. The prediction mean values (solid lines) and 2σ confidence intervals are shown. The true values are shown in black dots and the range of y_p from the train set is indicated by the grey region.

The BNNs predictions for unknown data





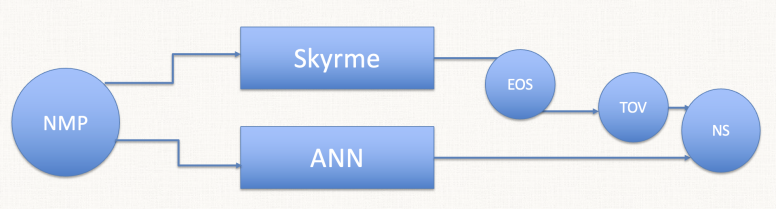
The BNN model predictions, v_s^2 (left) and y_p (right), for one mock observation of the DD2 EoS, the blue area represents the 95.4% confidence interval, and the solid line the mean.

Discussion

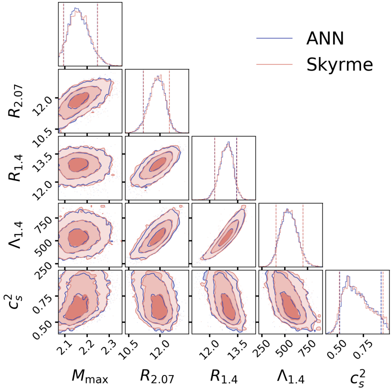
- We have explored Bayesian Neural Networks (BNNs), a probabilistic machine learning model, to predict the proton fraction and speed of sound of neutron star matter from a set of NS mock observations. This method is based upon the usual neural networks but with the crucial advantage of attributing an uncertainty measurement to its predictions.
- The tidal deformability data with a smaller uncertainty improved the speed of sound prediction, but not the proton fraction. This is because the proton fraction has a correlation with the symmetry energy slope, which is weaker with the increase of the NS mass.

@ Phys.Rev.D 108 (2023) 4, 043031

Study 2



The Neural Network fun for very expensive simulation



@ Phys.Rev.D 108 (2023) 6, 063028

Table: CPU inference time estimates for the ANN model and a Skyrme model to infer NS observations from a set of NMPs. 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.

Model	Time
ANN	2.23 min
Skyrme	16h 27 min

Conclusion

- The role of theoretical models go way beyond producing numbers. A theoretical model also indicates the actual physical mechanisms behind the properties being predicted. Since each term in the model is physically motivated, a theoretical model which comes close to experimental predictions also identifies what are the actual physical processes which are important in that energy scale. To have a theoretical understanding of any system, a physics based model is necessary. ML algorithms cannot replace physics modeling in that respect.
- However an interesting area of future work might be in combining the theoretical model and the machine learning methods to arrive at a better physical models. Our theoretical knowledge may help determine which features are physically relevant in a given data set while ML algorithms will help us find patterns and make predictions.

References

- Ameya Thete, Kinjal Banerjee, Tuhin Malik, Phys.Rev.D 108 (2023) 6, 063028
- Valéria Carvalho, Márcio Ferreira, Tuhin Malik, Constança Providência, Phys.Rev.D 108 (2023) 4, 043031
- Shriya Soma, Lingxiao Wang, Shuzhe Shi, Horst Stöcker, Kai Zhou, JCAP 08 (2022) 071
- de Tovar et al. (2021) Rev. D, 104, 123036; Imam et al. (2021) Phys. Rev. C 105, 015806 (2022); Mondal & Gulminelli (2021) Phys. Rev. D 105, 083016 (2022)

Acknowledgements

Collaborators

- Constança Providência, CFisUC, Coimbra
- B. K. Agrawal, SINP, India
- Márcio Ferreira, CFisUC, Coimbra
- Kinjal Banerjee, BITS-Pilani, Goa, India
- Valéria Carvalho, CFisUC, Coimbra
- Ameya Thete, Wisconsin U., Madison, USA

I would like to express my gratitude to all my collaborators for their invaluable contributions.

Thank You!

Email: tm@uc.pt, tuhin.malik@gmail.com