Constraining the Neutron Star Equation of State: A Deep Learning Approach

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Physik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von

Shriya Soma
aus Hyderabad, Indien

Frankfurt am Main, 2023
(D30)

vom Fachbereich Physik der
Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan: Prof. Dr. Roger Erb

Gutachter: Prof. Dr. Dr. h.c. Horst Stöcker, Prof. Dr. Luciano Rezzolla

Datum der Disputation: 16.04.2024

Diese Arbeit basiert auf folgenden Publikationen:

- **Shriya Soma** and Debades Bandyopadhyay, "Properties of Binary Components and Remnant in GW170817 Using Equations of State in Finite Temperature Field Theory Models", *The Astrophysical Journal,* **890**, 139 (2020) [1]
  DOI: 10.3847/1538-4357/ab6a9e

- **Shriya Soma**, Lingxiao Wang, Shuzhe Shi, Horst Stöcker and Kai Zhou, "Neural network reconstruction of the dense matter equation of state from neutron star observables", *Journal of Cosmology and Astroparticle Physics,* **08**, 071 (2022) [2]
  DOI: 10.1088/1475-7516/2022/08/071

- **Shriya Soma**, Lingxiao Wang, Shuzhe Shi, Horst Stöcker and Kai Zhou, "Reconstructing the neutron star equation of state from observational data via automatic differentiation", *Physical Review D,* **107**, 083028 (2023) [3]
  DOI: 10.1103/PhysRevD.107.083028

- **Shriya Soma**, Lingxiao Wang, Shuzhe Shi, Horst Stöcker and Kai Zhou, "A physics-based neural network reconstruction of the dense matter equation of state from neutron star observables", *EPJ Web of Conferences,* **276**, 06007 (2023) [4]
  Conference: The 20^{th} International Conference on Strangeness in Quark Matter (SQM 2022)
  DOI: 10.1051/epjconf/202327606007

- **Shriya Soma**, Lingxiao Wang, Shuzhe Shi, Horst Stöcker and Kai Zhou, "A neural network reconstruction of the neutron star equation of state via automatic differentiation", *Proceedings of Science,* FAIRness2022, **419**, 055 (2023) [5]
  Workshop: FAIR next generation scientists - 7th Edition Workshop (FAIRness2022)
  DOI: 10.22323/1.419.0055

- **Shriya Soma**, Horst Stöcker and Kai Zhou, "Mass and Tidal parameter extraction from gravitational waves of binary neutron stars mergers using deep learning", *Journal of Cosmology and Astroparticle Physics,* **01**, 009 (2024) [6]
  DOI: 10.1088/1475-7516/2024/01/009

Zusammenfassung

Die Zustandsgleichung extrem dichter Materie unter Einfluss der starken Wechselwirkungen ist zur Zeit nicht vollständig verstanden. Ein Grund hierfür ist die große Herausforderung, die eine theoretische Berechnung direkt aus der Quantenchromodynamik, also der fundamentalen Theorie der starken Wechselwirkung, bei endlichen chemischen Potentialen darstellt. Neutronensterne mit ihren extremen Dichten von 5 bis 6 mal der Kernsättigungsdichte (Englisch: nuclear saturation density, ρ_0) können als kosmische Laboratorien dienen, um die Eigenschaften von Materie unter diesen Bedingungen zu untersuchen. Die Messung von Neutronensternmassen und -radien sowie ihrer Gezeitenverformbarkeit (Englisch: tidal deformability) geben Einsichten in die Eigenschaften von Materie im Inneren dieser Sterne. Die Fortschritte der letzten Jahre bei der Beobachtung von Neutronensternen durch elektromagnetische Signale und Gravitationswellen erlauben nun eine Rekonstruktion der zugrundeliegenden Zustandsgleichung.

Bei niedrigen Dichten ($\sim \rho \lesssim 2\rho_0$) kann die Zustandsgleichung durch Experimente an Atomkernen und theoretische Berechnungen im Rahmen der sogenannten chiralen effektiven Feldtheorie (english: chiral effective field theory, χEFT) eingeschränkt werden. Bei asymptotisch hohen Dichten ($\sim \rho \gtrsim 40\rho_0$) können störungstheoretische Berechnungen der Quantenchromodynamik angewendet werden, um die Zustandsgleichung anzunähern. Im Dichtebereich zwischen diesen beiden Extremen, indem sich Neutronensterne befinden, wird die Zustandsgleichung üblicherweise mit verschiedenen theoretischen Ansätzen modelliert. Dabei werden unterschiedliche physikalische Annahmen für die Zusammensetzung und Eigenschaften stark wechselwirkender Materie getroffen. Modelle können nichtrelativistisch oder relativistisch sein und verschiedene Freiheitsgrade beinhalten. Manche Beschreibungen gehen von rein nukleonischer Materie aus, andere enthalten zusätzlich Hyperonen, Pionen- und Kaon-Kondensate oder reine, freie Quarkmaterie. Außerdem existieren sogenannte hybride Modelle mit einem Phasenübergang von hadronischer Materie zu freier Quarkmaterie.

Mit diesen Modellen werden globale Eigenschaften von Neutronensternen vorhergesagt und mit aktuellen Beobachtungsdaten verglichen. Alternativ kann die Zustandsgleichung aus Beobachtungen durch herkömmliche Methoden wie der Bayes'schen Inferenz rekonstruiert werden. In dieser Arbeit werden hierzu effiziente, alternative Methoden entwickelt. Hierfür untersuchen wir die Anwendbarkeit von Deep Learning Techniken.

Deep Learning ist ein Teilgebiet des maschinellen Lernens, bei dem künstliche, neuronale Netzwerke, sogenannte 'Deep Neural Networks', mit zahlreichen, miteinander verbundenen Schichten verwendet werden. Dieser Ansatz war in den letzten Jahren maßgeblich an Fortschritten der Künstlichen Intelligenz beteiligt und hat für zahlreiche Durchbrüche im Bereich Bilderkennung, Spracherkennung und Sprachverarbeitung gesorgt.

In dieser Arbeit nutzen wir Deep Learning, um Algorithmen für zwei verschiedene Aufgaben zu entwickeln:

1. Die Rekonstruktion der Zustandsgleichung extrem dichter Materie aus Beobachtungsdaten von Neutronensternen. Dies ist eine teilüberwachte Lernmethode, in welcher die Zustandsgleichung mithilfe von automatischem Differenzieren optimiert wird. Ausführliche Erklärungen werden in Kapitel 4 gegeben.

2. Die Analyse von Gravitationswellendaten, um die Parameter der Quelle dieser Wellen abzuschätzen. Hierfür werden überwachte Lernmethoden für die Klassifikation von Quellen und die Regression der Eigenschaften dieser Quellen verwendet. Details dazu werden in Kapitel 5 präsentiert.

0.1 Rekonstruktion der Zustandsgleichung dichter Materie

In Kapitel 4 stellen wir die neu entwickelte Deep Learning Methode für die Rekonstruktion der Zustandsgleichung stark wechselwirkender Materie aus Masse-Radius Beobachtungsdaten von Neutronensternen vor. Dieser einzigartige Algorithmus beinhaltet eine modellunabhängige und flexible Darstellung der Zustandsgleichung mithilfe von neuronalen Netzwerken, welche letztlich den Druck, P, als Funktion der Dichte, ρ, bestimmen. Die Zustandsgleichung wird mithilfe einer bayesianische Perspektive optimiert und die damit verbundenen Unsicherheiten werden durch Importance Sampling bestimmt. Dieser Ansatz ist bedeutend effizienter als eine vollständige Bayes'sche Analyse im multidimensionalen Parameterraum.

Konkret führen wir zwei verschiedene neuronale Netzwerke ein, welche wir mit `EoS Network` und `TOV-Solver Network` bezeichnen. Mit dem `EoS Network` wird die Zustandsgleichung modellunabhängig dargestellt. Wie der Name schon vermuten lässt, wird das `TOV-Solver Network` mit herkömmlichen, überwachten Lernmethoden darauf trainiert, die Tolman-Oppenheimer-Volkoff (TOV) Gleichungen zu lösen. Diese Gleichungen beschreiben die Struktur von statischen und sphärisch symmetrischen Neutronensternen im hydrostatischen Gleichgewicht. Die Grundlage des Algorithmus besteht darin, das `EoS Network` mit dem vorab trainierten `TOV-Solver Network` zu einer Pipeline zu kombinieren. Die einzelnen Wichtungen des `EoS Network` werden mithilfe eines gradientenbasierten Ansatzes im sogenannten Automatic Differentiation Framework durch unüberwachtes Lernen so optimiert, dass die Masse-Radius Kurve der resultierenden Zustandsgleichung am besten zu

den Beobachtungsdaten passt. Die Masse-Radius Kurve wird dabei mit dem `TOV-Solver Network` bestimmt. Mit diesem Algorithmus führen wir mehrere Tests mit simulierten Daten sowohl aus stückweise polytropen (Englisch: piecewise polytropes) als auch aus mikrophysikalischen Zustandsgleichungen durch. Dabei zeigt sich, dass sich die Leistung des Algorithmus verbessert, je geringer das Rauschen in den simulierten Daten ist. Eine verbesserte Präzision bei zukünftigen Messungen globaler Eigenschaften von Neutronensternen mit Teleskopen der nächsten Generation kann daher die Grundlage für eine bessere Rekonstruktion der Zustandsgleichung dichter Materie bieten. Anschließend wenden wir den entwickelten Algorithmus auf aktuell verfügbare, echte Beobachtungsdaten aus verschieden Quellen an. Die rekonstruierte Zustandsgleichung im Dichtebereich 1-7ρ_0 ist in Abbildung. 1 dargestellt. Die Ergebnisse sind mit Rekonstruktionen der Zustandsgleichung aus früheren Arbeiten mit konventionellen Methoden und den Einschränkungen der Gezeitenverformbarkeit durch die Beobachtungen des Gravitationswellenereignisses GW170817 kompatibel. Wir beobachten außerdem ein lokales Maximum in der quadratischen Schallgeschwindigkeit (Englisch: speed of sound squared, c_s^2) bei Dichten von $\sim 5\ \rho_0$. Die Schallgeschwindigkeit scheint für höhere Dichten wieder abzunehmen (für $\rho > 5 \sim 6\rho_0$). Eine solche glatte Änderung von c_s^2 könnte zu einer Kurve führen, welche bei extremen Dichten das sogenannte conformal limit erreicht. Es muss hierbei betont werden, dass die in die-

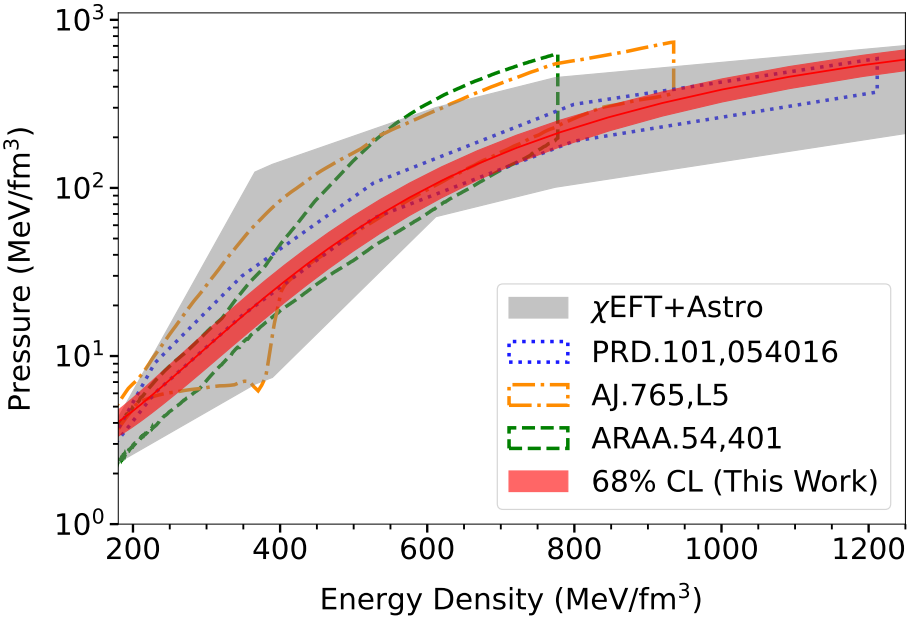

Figure 1: 68% Konfidenzintervall der aus Beobachtungsdaten von 18 Neutronensternen rekonstruierten Zustandsgleichung (rotes Band). Die grauen Bänder markieren die Einschränkungen der Zustandsgleichung aus der χEFT. Gezeigt werden außerdem Resultate anderer Methoden basierend auf Bayes'scher Inferenz und direkter Invertierung mit neuronalen Netzwerken. Details hierzu finden sich in Kapitel 4. Die Abbildung wurde aus Ref. [3] entnommen. (DOI:10.1103/PhysRevD.107.083028).

sem Kapitel verwendeten Daten zur Rekonstruktion der Zustandsgleichung dichter Materie auf Massen- und Radienbeobachtungen von Neutronensternen beschränkt sind. Mit einer Zunahme der Gravitationswellendaten aus Kollisionen von Neutronensternbinärsystemen kann der Algorithmus in Zukunft erweitert werden, um zusätzlich die Beziehung zwischen der Gezeitenverformbarkeit und der Masse eines Neutronensterns zu berücksichtigen.

0.2 Gravitationswellenanalyse mit Deep Learning

In Kapitel 5 demonstrieren wir die Leistungsfähigkeit von neuronalen Netzwerken bei der Analyse simulierter Gravitationswellendaten anhand zwei verschiedener Aufgaben:

(i) der Klassifizierung von Gravitationswellensignalen aus der Verschmelzungen zweier schwarzer Löcher, zweier Neutronensterne und von Signalen aus reinem Rauschen; und,

(ii) der Regression der sogenannten Chirpmasse ($\mathcal{M}$) und der kombinierten Gezeitenverformbarkeit ($\tilde{\Lambda}$) aus simulierten Gravitationswellensignalen von der Verschmelzung zweier Neutronensterne.

Idealerweise sollten zur Analyse von Gravitationswellen aus Neutronensternverschmelzungen Modelle basierend auf dreidimensionalen, magnetohydrodynamischen Simulationen im Rahmen der Allgemeinen Relativitätstheorie verwendet werden. Diese sind zwar einerseits akkurater und modellunabhängiger, andererseits aber auch extrem rechenaufwändig und eignen sich weniger für diese Arbeit. Um die künstlichen Gravitationswellensignale zu generieren, verwenden wir daher das approximative Wellenformmodel der LALSuite Bibliothek. Der verwendete Frequenzbereich wird außerdem auf Frequenzen unterhalb von 2048 Hz beschränkt. Dies schließt eine Analyse der sogenannte Ringdown-Phase des Überrests nach der Verschmelzung aktuell aus. Für eine Erweiterung der entwickelten Algorithmen hin zu höheren Frequenzen ist weitere Arbeit nötig.

Der in diesem Kapitel präsentierte Deep Learning Klassifizierer ist in der Lage, die verschiedenen Signalarten mit einer sogenanten Macro-Averaged Genauigkeit (Englisch: macro-averaged precision) von 0,99 und 0,96 für ein optimales Signal-zu-Rausch Verhältnis (Englisch: optimal signal to noise ratio, oSNR) von jeweils 25 und 20 zu identifizieren.

Zusätzlich zum Klassifizierer entwickeln wir einige Deep Learning Regressoren, um $\mathcal{M}$ und $\tilde{\Lambda}$ aus Signalen von Neutronensternverschmelzungen zu bestimmen. Dabei erreicht das Netzwerk hohe $\mathcal{R}^2$-Werte von jeweils 0,98 und 0,97 für die Regression von $\mathcal{M}$ und $\tilde{\Lambda}$ aus Testdaten. Für diese Testdaten wurde weißes und farbiges Rauschen des Gravitationswellendetektors aLIGO auf simulierte Signale von Neutronensternverschmelzungen addiert. Die Bestimmung des Gezeitenparameters ist dabei besonders wichtig, um die Zustandsgleichung von Neutronensternmaterie zu einzuschränken. Hierbei ist zu beachten, dass das Regressionsnetzwerk weder die Gezeitenverformbarkeiten noch die Massen der einzelnen Sterne aus der Chirpmasse und der kombinierten Gezeitenverformbarkeit

ausgibt. Dies stellt ein aktuell realistisches Szenario dar, bei der Informationen über die Neutronensternzustandsgleichung aus $\mathcal{M}$ und $\tilde{\Lambda}$ bestimmt werden müssen. Dies ist mit einer ausreichenden Anzahl an Beobachtungen möglich. Die in diesen Kapitel vorgestellten Klassifikatoren und Regressoren können in eine umfassende Pipline für die Analyse zukünftiger Gravitationswellensignale integriert werden. Nach der Identifikation eines Ereignisses mit dem Klassifikationsnetzwerk können die relevanten Parameter direkt mit dem Regressionsnetzwerk bestimmt werden. Dieser nahtlose Übergang bildet somit eine ganzheitliche Pipeline für die zukünftige Analyse von Gravitationswellen mit den geplanten Detektoren der nächsten Generation. Von diesen wird erwartet, dass ihre Sensitivität die der heutigen Detektoren um etwa eine Größenordnung übertrifft.

0.3 Thermische Effekte in Kollisionsüberresten

In Kapitel 6 untersuchen wir die Auswirkungen endlicher Temperaturen auf Eigenschaften des Überrests einer Neutronensternverschmelzung. Hierzu verwenden wir verschiedene isentrope Zustandsgleichungenm mit einer konstanten Entropie pro Baryon von $s = 2$, um das Trägheitsmoment und die Kepler-Frequenz des starr rotierenden Überrests des Ereignisses GW170817 zu untersuchen. Dabei ergibt sich eine erhebliche Abnahme der Kepler-Frequenzen verglichen mit den Berechnungen für einen kalten Überrest ($s = 0$).

Wir folgern daraus, dass die Frequenzen in Ref. [7] deutlich überschätzt werden, da diese mit einer kalten Zustandsgleichung berechneten wurden. Zusätzlich geben wir Abschätzungen für die Radien der beiden Neutronensterne an, die im Ereignis GW170817 kollidiert sind. Hierzu verwenden wir die analytischen Beziehungen aus Ref. [8] mit einer größeren Anzahl an kalten Modellen für die Zustandsgleichung. Unter Verwendung der Abschätzungen für die effektive Gezeitenverformbarkeit berechnen wir eine obere und untere Grenze für die Radien der Sterne von etwa 13 km bzw. 8,85 km. Darüber hinaus verwenden wir verbesserte Fitrelationen für die $R - \tilde{\Lambda}$ Beziehung, um die Radien von Neutronensternen mit Massen von 1,2 $M_\odot$ und 1,55 $M_\odot$ zu bestimmen. Es ergeben sich Werte von 13,08 km bzw. 12,99 km. Dabei ist zu beachten, dass diese Abschätzungen unter der Annahme getroffen werden, dass sich die Radien von Sternen im Massenbereich von $1,17 - 1,6\ M_\odot$ kaum unterscheiden. Diese Annahme wäre in Szenarien mit einem großen Dichtesprung der Zustandsgleichung durch einen Phasenübergang erster Ordnung höchstwahrscheinlich nicht mehr zutreffend. Die in dieser Studie verwendeten Relationen für $R - \tilde{\Lambda}$ basieren hauptsächlich auf nukleonischen Zustandsgleichungen.

Zusammengefasst stellt diese Arbeit neuartige Deep-Learning-Methoden vor, die Beobachtungsdaten von Neutronensternen nutzen, um damit die zugrunde liegende Zustandsgleichung einzuschränken.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The behaviour of matter under extreme conditions, such as high densities and temperatures, is currently not completely understood. An accurate description of matter follows from the determination of the equation of state (EoS), which is the relationship between pressure and energy density at variable temperatures and entropies. In other words, the EoS is a mathematical representation of how matter responds to changes in its environment. At densities above saturation density ($\sim$0.16 fm^{-3}), however, the equation of state remains poorly known and is a central topic across various fields of physics. The relevant microscopic degrees of freedom and their fundamental interactions at such high densities are unresolved. While a phase transition from hadronic to deconfined quark matter is expected to occur at high densities or temperatures, it has not been discovered yet. Under such conditions, the matter can described by Quantum Chromodynamics (QCD), the theory of strong interactions.

In terrestrial laboratories, heavy ion collisions (HIC) provide means to create a localized region of hot and dense QCD matter. By varying the collision energies, different temperatures and densities of the QCD matter can be probed. Several experiments like the Large Hadron Collider (LHC) and Super Proton Synchrotron (SPS) at CERN, Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory, and the Schwerionen Synchrotrons (SIS) at GSI and FAIR, are designed to probe different regions of the QCD phase diagram (see Refs. [9,10] for a review on the phase diagram of QCD). Detailed information can be found in the following references [11–30]. The conditions created in heavy ion collisions, however, exist only for timescales typically on the order of 10^{-15}s and involve moderate to high temperatures ($\gtrsim$ 50-300 MeV [31]).

Neutron stars (NSs), on the other hand, provide an opportunity to study cold dense matter. The timescales of long-lived neutron stars can vary from several million to billion years, making it possible to gain an understanding of stable equilibrated high density matter. Despite the huge differences in scales, i.e., 1fm ($=10^{-15}$m) in nuclear collisions, and 10km ($=10^{4}$m) in neutron stars, the microscopic and bulk properties of matter are sim-

ilar: it is a dense composition of strongly-interacting particles. In this thesis, we focus on studying neutron stars and their properties to gain a deeper understanding into their composition and the properties of strongly interacting matter.

1.1 Neutron Stars

Neutron stars (NSs) were first hypothesized in 1931, long before their discovery [32, 33]. In his article which was published in 1932, Lev D. Landau had proposed the existence of dense astrophysical objects which looked like giant nuclei [32]. Interestingly, this occurred prior to Chadwick's discovery of the neutron, which coincidentally took place in the same year [33–35]. A few years later, in 1933, the term 'neutron star' made its first appearance in an article by Baade and Zwicky [36]. This was the first explicit prediction of neutron stars, which resulted from an attempt to explain the immense energy released in supernovae explosions [36–38]. These developments marked the onset of research into the physics of neutron stars. It was Landau's original work that inspired the study of neutron stars within general relativity, by Oppenheimer and Volkoff [39, 40]. An extended explanation of the Tolman-Oppenheimer-Volkoff equations [39, 41] can be found in section 2.2.1 of chapter 2.

It was only after three subsequent decades, in 1967, that the serendipitous discovery of a neutron star was made by Jocelyn Bell-Burnell [42], a graduate research student at that time. It was observed as a pulsar, or a rapid pulsating radio signal. The pulse period was noted to have a precision so high, that initial suspicions included a possible intelligent origin. For this discovery, Jocelyn Bell's supervisor, Anthony Hewish was awarded the Nobel prize in 1974. Several years later, Bell Burnell was awarded the Special Breakthrough Prize in Fundamental Physics in 2018, and the Copley Medal for her work on the major astronomical discovery. Following their finding, research in the field of NSs and their interior composition gained momentum.

Neutron stars are formed in the aftermath of Type II supernovae explosions of massive stars with initial mass of above $8M_\odot$ (see Ref. [43] for a review on core collapse supernovae). When the iron core of the massive star achieves the Chandrasekhar mass, the collapse is initiated [44]. During the compression of the massive star's core and subsequent collapse to a neutron star, most of the angular momentum is retained. This leads to very high rotation speeds in NSs (rotation periods of $\sim$1.4ms-30s). Similarly, the strong magnetic fields in NSs ($\sim 10^{12}$G) are a result of the conservation of magnetic flux [45]. The newly born NS, also known as a proto-neutron star, typically has a mass of $1.4M_\odot$, and a radius of $\sim$10-12km. Therefore, NSs constitute the most dense compact objects in our universe. An extensive review on neutron stars and their properties can be found in Refs. [40, 46–48].

The structure of a neutron star is ultimately governed by the properties of dense matter, or the EoS of strongly interacting matter; and by the influence of gravity. Hence, it is interesting to look at characteristics of neutron stars to constrain the underlying EoS at these densities. In the following sections, we discuss three NS observables, namely the mass,

radius and tidal deformability. This thesis is structured around using these NS properties to constrain the EoS. Consequently, the rest of the chapter contains a brief description of the different techniques that allow for these measurements.

1.1.1 Mass Measurements

The neutron stars' strong magnetic fields coupled with their rapid rotation, yields a prominent observable signature that enables their detection. As the NS rotates around its spin axis, the charged particles experience an acceleration along magnetic field lines within the magnetosphere, resulting in the emission of electromagnetic radiation. Due to the misalignment of the magnetic axis and the spin axis, we observe the so called 'light-house' effect, where the electromagnetic radiation can be detected at every rotation as it crosses the observer's line of sight as a sequence of pulses.

Following their discovery [42], it became evident that pulsars serve as exceptional cosmic clocks. The identification of the millisecond pulsar B1937+21 in 1982 [49], showcased that its period could be gauged with precision down to one part in 10^{13} or more [50]. This exceptional stability of the pulse period gives rise to a multitude of uses, one of which is timekeeping. Therefore, pulsar timing is one of the principal means of measuring neutron star masses precisely. See Refs. [51, 52] for detailed reviews on pulsar timing.

Depending on their evolutionary scenarios, pulsars can exist as isolated objects as well as in binaries with white-dwarf or neutron star companions. The most precise measurements of neutron star masses have been accomplished through radio timing of pulsars situated in binary systems [53, 54]. Most of these known binary pulsar systems can be described by the Keplerian parameters [55], namely, the orbital period (P_b), the projected semi-major orbital axis ($x \equiv a_p \sin i$), the eccentricity (e), the longitude of periastron (ω), and the epoch of periastron (T_0). However, it was observed that for several systems, a fit to one or more relativistic corrections to the Keplerian parameters is required (for example, see Ref. [56]). Therefore, recent timing models incorporate an additional set of "post-Keplerian" parameters [56], which according to Einstein's general relativity [57–59], are given by,

1. the advance of the periastron of the orbit,

$$\dot{\omega} = 3 \left(\frac{P_b}{2\pi} \right)^{-5/3} (T_\odot M)^{2/3} \left(1 - e^2\right)^{-1} , \tag{1.1}$$

2. the combined effect of the gravitational redshift and transverse Doppler shifts due to the pulsar's orbit,

$$\gamma = e \left(\frac{P_b}{2\pi} \right)^{1/3} T_\odot^{2/3} M^{-4/3} m_c \left(m_p + 2m_c\right) , \tag{1.2}$$

3. the orbital period decay due to the emission of gravitational radiation,

$$\dot{P}_b = \frac{-192\pi}{5}\left(\frac{P_b}{2\pi}\right)^{-5/3}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right)(1-e^2)^{-7/2}T_\odot^{5/3}m_p m_c M^{-1/3}, \quad (1.3)$$

4. the range of Shapiro delay,

$$r = T_\odot m_c, \quad (1.4)$$

and,

5. the shape of Shapiro delay,

$$s = x\left(\frac{P_b}{2\pi}\right)^{-2/3} T_\odot^{-1/3} M^{2/3} m_c^{-1}. \quad (1.5)$$

Shapiro delay is the delay caused in the pulsar signal's time of arrival due to its propagation through the gravitational field of its companion [60,61]. In these equations, m_p denotes the pulsar mass; m_c denotes the orbiting companion mass; and $M \equiv m_p + m_c$. In addition, $T_\odot \equiv GM_\odot/c^3 = 4.9255\mu$s and $s \equiv \sin i$, where i is the inclination angle or the angle between the orbital plane and the plane of the sky. Since the masses of the binary components constitute the sole unknowns in the equations, it is sufficient to obtain the measurements for any two post-Keplerian parameters. This provides us with the two mass estimates, m_p and m_c. Measurements of three or more post-Keplerian parameters therefore allow for a consistency check. Although only a small fraction of pulsars in binaries have two or more well-measured relativistic effects, these techniques can facilitate precise measurements of pulsar masses. There have been relatively precise mass measurements of more than a few dozen neutron stars in binary systems [62, 63], most of which are listed in Ref. [64]. Accurate measurements of the most massive pulsars include PSR J0740+6620, with a mass of 2.08 ± 0.07 $M_\odot$ [53] and PSR J0348+0432, with a mass of 2.01 ± 0.04 $M_\odot$ [54]. With the observational support of such massive pulsars, many EoSs were ruled out for being unable to produce such high maximum masses [65–70]. Simultaneously, several microphysical EoS models have been developed to satisfy these constraints, some of which will be introduced in chapter 2.

Note that the masses of isolated neutron stars can also be measured through pulsar glitches, as demonstrated in Ref. [71]. However, this method assumes a prior knowledge of the correct EoS. Additionally, neutron stars with higher masses have been measured more recently through Keck-telescope spectrophotometry [72, 73].

1.1.2 Radii Measurements

Radii measurements of neutron stars are crucial to obtain constraints on the underlying equation of state. However, due to their incredibly small sizes (∼12 km), it is extremely challenging to obtain precise measurements of NS radii. Nevertheless, several astrophysical

observations with the potential to facilitate the determination of neutron star radii were proposed.

The Neutron Star Interior Composition Explorer, or NICER [74], as the name suggests, is an X-ray telescope, developed to probe the interior structure of neutron stars. It was installed on the International Space Station (ISS) in 2017. NICER uses pulse-profile modeling of nearby rotation-powered millisecond pulsars (MSPs) to estimate their radii (Refer to the paragraph on 'Rotation-powered Millisecond Pulsars' in this section). Prior to the era of NICER, however, the leading candidates for inferring the radii of neutron stars were quiescent low-mass X-ray binaries and thermonuclear bursters [62, 63, 75]. In both these cases, measurements of the radii rely on the detection of thermal emission from the neutron stars' surfaces. The study of these emissions is benefited by observations from X-ray instruments like the Rossi X-ray Timing Explorer (RXTE), Chandra X-ray Observatory and XMM-Newton.

Thermonuclear Bursters

Thermally unstable helium (or hydrogen) ignition in the accreted envelope of a neutron star results in X-ray bursts [76]. The ignition triggers a thermonuclear explosion which is observed as an X-ray burst [62,77]. The rise time of the X-ray burst is ∼1s, and is followed by a cooling decay with a period $\sim 10 - 100$s. During the rise period, these sources are believed to have a radiation pressure large enough to temporarily lift the neutron star atmosphere off the surface. Therefore, the photosphere at that time, has a peak luminosity that approaches the Eddington limit,

$$L_{Edd} = 4\pi c G M \kappa^{-1}. \tag{1.6}$$

Note that the Eddington limit is defined as the maximum luminosity that a body can achieve when there is a balance between two opposing forces, namely, the outward radiation pressure and the inward gravitational force. In Eq. (1.6), M is the mass of the star, and κ is the opacity of the atmosphere. Conversely, the luminosity on the tail of the burst is attributed to the thermal radiation emanating from the cooling star and follows,

$$L = 4\pi f_c^{-4} R^2 T_{\rm eff}^4. \tag{1.7}$$

Here, the radius of the neutron star is denoted by R; $T_{\rm eff}$ is the effective blackbody temperature; and the factor f_c accounts for color correction, which approximates the influence of stellar atmospheric conditions on distorting the observed temperature from the effective blackbody temperature. The observed parameter is the flux, so an estimation of the luminosity requires knowledge of the source distance, D. Another observed parameter is the angular area of the photosphere, A. Furthermore, κ and f_c are obtained from atmospheric models. Therefore, assuming a knowledge of D, κ and f_c, the mass and radius of the source can be determined. A few observations were used to extract radii measurements of NSs using this technique, and are catalogued in Ref. [75].

Quiescent Low-Mass X-ray Binaries

Many neutron stars within binary systems are transients, i.e., they might undergo periodic episodes of accretion from a companion star in evolution. These episodes of accretion are separated with extended periods (months to years) of quiescence, i.e., when accretion either ceases or continues at very low levels [62, 63, 78]. Electron capture and pycnonuclear fusion reactions at densities exceeding 10^{12}g cm^{-3} of a transiently accreting neutron star release energy in the deep crust of the NS [79, 80]. It is possible that a fraction of this energy is deposited as heat if it is not lost through neutrino emission [81]. Therefore, as accretion stops and the NS goes into quiescence, the crust reradiates the deposited energy. As opposed to the emission from the accretion disk at the high mass accretion rates, during quiescence, the observed X-ray emission originates primarily from the neutron star surface. The extracted emission spectra of these NSs is then fitted with well-understood hydrogen atmosphere models (due to the very short timescale for heavier nuclei to sink below the photosphere, the atmosphere envelope is expected to be composed of hydrogen, or helium if the companion star is hydrogen-poor). Atmosphere modeling is then used to infer the apparent angular emitting area, which is a function of the mass and size of the neutron star. See Refs. [75, 82] for a list of radius measurements obtained using this method.

Rotation-powered Millisecond Pulsars

Rotation-powered millisecond pulsars (MSPs) originate as an outcome of accreting millisecond pulsars once the accretion process comes to a halt. The magnetic field of an accreting MSP directs infalling material toward the magnetic poles of the star. As a result, a confined region on the surface of the pulsar is heated up, forming a 'hot spot', which is much brighter than its surroundings. NICER targets rotation-powered MSPs to map these surface hotspots through modeling of their X-ray pulses [74,83–85]. Depending on the fraction of the hot spot visible to the observer, an apparent oscillation in the intensity is expected at the rotation period of the pulsar. The pulse waveform has a shape that depends not only on the orientation of the neutron star relative to the observer, but also on physical processes like gravitational light-bending and relativistic Doppler boosting. These relativistic effects are intensified by higher values of NS compactness ($C = M/R$), and therefore enable us to relate to their masses and radii. A major challenge faced in extracting the radii, is due to a lack of understanding in the surface emission patterns. Due to this uncertainty, the data obtained by NICER is independently analyzed by two different groups, leading to the use of different hotspot models. The first mass and radius measurements extracted from NICER were reported in 2019, for the pulsar PSR J0030+0451. The independent analyses by the different groups, however, led to similar radius and mass estimates, i.e., $12.71^{+1.14}_{-1.19}$km and $1.34^{+0.15}_{-0.16}$M$_\odot$, respectively, by Ref. [86]; and $13.02^{+1.24}_{-1.06}$km and $1.44^{+0.15}_{-0.14}$M$_\odot$, respectively, by Ref. [87]. It is worth mentioning that an updated analysis of the NICER data for the pulsar was recently carried out using an upgraded framework. The adopted models were expanded and jointly analyzed with XMM-Newton data to obtain new mass-radius estimates. Details can be found in Ref. [88]. Furthermore, radii measurements for the massive

pulsar PSR J0740+6620 were reported by the two NICER groups in 2021 [89, 90]. Despite being 4 times more distant than PSR J0030+0451, and 20 times fainter, PSR J0740+6620 serves as an interesting target, due its precise high mass measurement [53]. Note that the first results of PSR J0740+6620's radius measurement included joint fits of NICER and XMM-Newton data. Therefore, a reanalysis was carried out using NICER background estimates to avoid cross correlations between the two instruments [91]. If NICER attains the originally targeted precision of 5% relative error on radii measurements, strong constraints can be applied on the underlying EoSs in future.

1.1.3 Tidal Deformability from Gravitational Waves

Neutron stars in binary systems lose substantial amounts of energy via the emission of gravitational waves (GWs). This leads to a continuous decrease in their orbital separation and an eventual merger of the binary components. This effect of orbital shrinkage was first observed in the Hulse-Taylor binary pulsar [92–96], resulting in the first indirect detection of gravitational waves. During the late inspiral stage, both components in the binary neutron star (BNS) system encounter a perturbing tidal gravitational field from their companion star. As a result, the stars experience tidal distortions through an induced quadrupole moment. This subsequently influences the binding energy of the binary system and enhances the gravitational wave emission rate [97, 98]. Therefore, the emitted gravitational waves carry an imprint of the tidal properties of the neutron stars. Consequently, the detection of gravitational waves from BNS inspirals and subsequent mergers is crucial to constrain the tidal parameter.

A direct measurement of gravitational waves is now possible by the existing gravitational wave detectors on Earth. These include the ground-based detectors, namely, advanced LIGO (Livingston and Hanford) and advanced Virgo; as well as the underground Kamioka Gravitational Wave detector, KAGRA. A passing gravitational wave causes space to stretch in one direction and simultaneously squeeze in the perpendicular direction. This effect, though extremely miniscule, can be detected on Earth with the help of an interferometer. The above-mentioned GW detectors are large Michelson interferometers, designed to detect strains of the order of magnitude $\sim 10^{-21} - 10^{-22}$. A simple interferometer comprises a laser source, a photodetector, and two mirrors at the ends of two arms placed perpendicular to each other. A beam-splitter is used to split a laser beam into both these arms. The laser beams from both arms get reflected at the mirrors placed at the end and recombine on a screen to form an interference pattern. Any changes in the length of either of the arms are reflected by changes in the interference pattern. A GW in transit results in an oscillating interference pattern, thereby allowing for a measurement of the waveform signal.

The tidal effects during the inspiral phase leave an imprint on both the amplitude and the phase of the emitted GW signal. The phase, however, is a better measure for harmonic functions such as waves [99]. The gravitational wave phase evolution can be obtained from a post-Newtonian (PN) expansion of the binary orbital frequency [100] (see Ref. [101] for a detailed review on post-Newtonian methods). In this framework, the tidal correction to

the quadrupole appears only at 5PN order [99, 102], and is expressed as a combined tidal parameter, i.e.,

$$\tilde{\Lambda} = \frac{16}{13}\frac{[(M_1 + 12M_2)M_1^4\Lambda_1 + (M_2 + 12M_1)M_2^4\Lambda_2]}{(M_1 + M_2)^5} \ , \tag{1.8}$$

where Λ_1 and Λ_2 are the individual tidal deformabilities of the binary NSs with masses M_1 and M_2, respectively (see chapter 2, section 2.2.2 for a mathematical definition of tidal deformability). Therefore, the 5PN tidal correction term is not sufficient to extract the individual tidal deformabilities. The 6PN tidal correction term, however, is a linear combination of $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ [102, 103]. An estimation of both $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ from the 5PN and 6PN terms can provide us with the individual tidal deformabilities, $\{\Lambda_1, \Lambda_2\}$. The first tidal deformability estimates emerged from the gravitational wave event, GW170817 [104]. This event was the first detection ever of a binary neutron star merger. An analysis of the event sets limits on the estimated combined tidal deformability, i.e., $70 \leq \tilde{\Lambda} \leq 720$ [105] for neutron star components within mass ranges, $1.16 \leq M_2/M_\odot \leq 1.36$ and $1.36 \leq M_1/M_\odot \leq 1.6$. Therefore, GW170817 could rule out a number of extremely stiff EoSs that do not satisfy this constraint [1, 8].

The sensitivity of the existing gravitational wave detectors is however expected to improve in the near future with the implementation of regular upgrades. In the current run, O4, advanced LIGO has the capacity to detect binary neutron star mergers at distances in the range of 130-150 Mpc. The target of the O4 run, however, is to reach 160-190 Mpc [106]. This implies an increased number of BNS merger detections. Therefore, an estimate of the tidal deformability from these events, could help to further constrain the dense matter EoS in NSs. The new gravitational wave detector, IndIGO (LIGO India), is expected to be added to the grid in 2030s. Moreover, third generation telescopes, like the Einstein Telescope (ET) and the Cosmic Explorer (CE) are anticipated to have sensitivities that are an order of magnitude better than the current generation GW detectors. Together, the ET and CE are expected to detect over 100 BNS merger events per year with an oSNR value >30 [107].

1.1.4 Other Properties of Neutron Stars

Apart from the three observable parameters (mass, radius and tidal deformability) used in this thesis, we discuss additional neutron star properties that can help us gain a further understanding of these objects. Some of the astrophysical observations that lead to the extraction of additional NS properties include the following:

1. **Spin-orbit coupling**

   Rapidly rotating NSs can carry a spin-induced quadrupole moment, which causes changes in the orbital dynamics of a binary system [108, 109]. A spin-orbit coupling is realized through the precession of the orbital plane about the direction of the total

angular momentum of the system. This is termed 'geodetic precession'. Due to the conservation of total angular momentum of the system, compensating precessions of the NSs' spins are expected. The resulting spin precessions lead to observable effects like, (i) a change in the spin axis orientation in space through pulsar timing, and (ii) a change in the inclination angle of the orbital plane orientation [110]. See Ref. [111] for how these effects influence pulsar timing. The change in inclination angle provides an estimate for the moment of inertia. Universal relations between moment of inertia and compactness (ratio of mass and radius) of NSs, can in turn be used to derive the radii, given precise mass measurements [8]. Future measurements of moment of inertia are expected by 2030 [112] and can help constrain the underlying NS EoS.

2. **Pulsar glitches**

   In some isolated neutron stars, abrupt changes in the spin, or the so called 'pulsar glitches', have been observed. These glitches are believed to occur due to the transfer of angular momentum between the crust and the liquid interior [113–115]. By studying these glitches, one can gain insights into the properties of neutron star crusts, and the dynamics of the superfluid interiors [113–115].

3. **Neutron star seismology**

   The aim of NS seismology is to probe the dense matter physics through observations of various modes of vibration. Quasi-periodic oscillations observed in the aftermath of giant flares in soft gamma-ray repeaters[1] indicate a correlation between the seismic motion of the NS crust after a major quake and the modes of oscillations in the NS [116]. Such modes can be used to deduce global properties of the NSs, thereby adding constraints to the dense matter EoS [116].

Research on neutron stars has benefited from several observations in the past decades. We utilize some of the observations in this thesis, in order to constrain the cold dense matter EoS, using deep learning techniques.

The thesis is organized as follows:

(i) Chapter 2 provides a brief introduction to the equation of state. We describe the theoretical models utilized in this thesis for the EoS.

(ii) Chapter 3 focuses on the deep learning techniques used for this thesis. We provide an introduction to neural networks and their utility for this work.

(iii) Chapter 4 details a deep learning algorithm based on automatic differentiation to reconstruct the dense matter EoS from mass-radius observations of neutron stars. We further present results based on current NS mass-radius observations. In addition, we show that the resulting EoS is compatible with the tidal deformability constraints obtained from GW170817.

[1]Soft gamma-ray repeaters are NSs or magnetars that emit intense bursts of soft gamma-ray radiation in irregular intervals.

(iv) Chapter 5 introduces deep learning algorithms developed to analyze gravitational waves. We introduce an method to overcome the computational inefficiency in conventional Bayesian analyses to extract tidal deformability from GWs of binary neutron star mergers.

(v) Chapter 6 is centered around the impact thermal effects can have on a binary neutron star merger remnant. We also discuss the properties of the binary components and merger remnant involved with the event GW170817.

(vi) Chapter 7 summarizes the results from this thesis. We offer a perspective on how this work can be further extended and integrated into forthcoming research endeavors.

Chapter 2

Theoretical Models for Neutron Stars

Neutron stars and their properties can be described from the underlying equation of state (EoS). This chapter provides a theoretical description of several EoS models used throughout this thesis. As mentioned in the previous chapter, the equation of state describes matter under different conditions. It is generally a thermodynamic relation between pressure (P) and energy density (ϵ) or baryon density (n_b) at variable temperatures (T) or entropies (S) and composition (Y_e, or electron fraction). Up to nuclear saturation density ($n_0 = 0.16$ fm^{-3}), the equation of state of matter can be constrained by experiments on finite nuclei [117–119] and chiral effective theory [120–125]. However, for densities beyond saturation density, the equation of state is not completely understood due to the limitations on the availability of experimental data. In this regime, several different theoretical approaches to describe the EoS exist. The fundamental physical requirements are,

(i) the causalty of an EoS, i.e., the speed of sound in matter ($c_s = \sqrt{dP/d\epsilon}$) cannot be greater than that of light (c),

(ii) and, a monotonically increasing pressure with increasing energy density, in order to ensure thermodynamic stability.

Further constraints on these theoretical models arise from astrophysical observations. A stringent constraint on the EoS emerges from the pulsar observation with the maximum mass, i.e., currently PSR J0740+6620, with a mass of $2.08 \pm 0.07 M_\odot$ [53]. All theoretical models discussed in this thesis satisfy this constraint[1]. Additional constraints are incorporated from tidal deformability measurements obtained from gravitational waves of binary neutron star mergers. Sections 2.2.1 and 2.2.2 of this chapter are devoted to theoretical calculations of structural properties and tidal deformability of a neutron star, respectively. The rest of the thesis is based on the EoSs described in this chapter, and their compatibility with other observational constraints will be tested.

[1]Note: For the polytropic EoSs in section 2.1.2, we use a less conservative bound of $1.9 M_\odot$. The reasons for this choice are explained in chapter 4.

2.1 Models for the Equation of State

Theoretical models which can generate possible equations of state (EoSs) that attempt to describe the dense matter inside neutron stars (NSs) incorporate different physical assumptions for the strongly interacting dense matter. Non-relativistic and relativistic model EoSs may include only purely nucleonic degrees of freedom [126–135], hybrid (hadrons and quarks) models [31, 136–142], models with hyperons [143–149], models with kaon condensates [150] or pure quark models [151, 152].

Alternatively, EoSs can be parameterized with piecewise polytropes [153–157], speed-of-sound [158–161], or spectral representations [162–164]. Such models benefit from their simplicity. The so-called 'nonparameteric' models for the EoS overcome this by increasing model freedom at the cost of increased complexity. This alternative approach targets higher model flexibility by making use of Gaussian processes (GPs) [165–170]. All these mentioned parameteric and nonparametric models are useful for generation of EoSs in bulk. These EoSs can be used to calculate the global properties of NSs, which can then be confronted with observational data. This way, we can directly constrain the EoS or its parameters through NS properties like maximum mass, radius, and tidal deformability, which are all dependent on the EoS.

2.1.1 Microphysical EoS Models

In this section, we describe some of the theoretical microphysical EoS models used in the thesis. We adopt β-equilibrated and charge neutral EoSs at zero and finite temperatures constructed within the framework of relativistic mean field (RMF) models with and without density dependent (DD) couplings. The baryon-baryon interaction in RMF models is mediated by the exchange of σ-, ω- and ρ-mesons; while the hyperon-hyperon interaction is mediated by the exchange of ϕ-mesons. Furthermore, we consider two classes of EoSs, namely, the unified EoSs and the non-unified EoSs. Unified EoSs refer to those EoSs where the same nucleon-nucleon interaction of RMF models is employed in low and high density matter. For non-unified EoSs, we use the RMF model including non-linear σ-meson self-interaction terms to describe the neutron star matter EoS in the core. This high density EoS is matched with the Baym-Pethick-Sutherland (BPS) outer crust EoS [171] and the Negele-Vautherin (NV) inner crust EoS [172]. On the other hand, for unified EoS, an extended version of the nuclear statistical equilibrium (NSE) is used for matter made of light and heavy nuclei, and unbound nucleons at low temperatures and below the saturation density [173]. The interaction among unbound nucleons is dictated by the RMF models which are also used to describe the matter at high densities.

We exploit different parametrizations of RMF models such as DD2 [131], SFHo [133], SFHx [133], TM1 [128], and TMA [129] EoSs for nuclear matter; BHB$\Lambda\phi$ EoS [144] for hyperonic matter; and a hybrid EoS undergoing a first order phase transition from hadronic to deconfined quark matter [140]. In the following paragraphs, we discuss the unified and non-unified EoSs based on the Boguta model [174]. The Boguta model [174] is also

commonly referred to as the 'non-linear Walecka' model due to the inclusion of non-linear self-interactions which were omitted in the Walecka model [175, 176]. We also discuss the density dependent relativistic hadron (DDRH) field theory model for EoSs at finite temperature, and additionally, the hybrid EoS.

Unified EoSs in the Boguta model

Here we introduce some of the unified EoSs based on the NSE model for matter below the saturation density, and the Boguta (non-linear Walecka) model [174] with additional meson couplings [173,177]. The non-linear Walecka model with cross meson terms is given by [177],

$$\begin{aligned}\mathcal{L} = \bar{\psi}_B(i\gamma_\mu\partial^\mu - m_B + g_{\sigma B}\sigma - g_{\omega B}\gamma_\mu\omega^\mu - \frac{1}{2}g_{\rho B}\gamma_\mu\bar{\tau}_B\cdot\bar{\rho}^\mu)\psi_B \\ +\frac{1}{2}(\partial_\mu\sigma\partial^\mu\sigma - m_\sigma^2\sigma^2) - \frac{1}{4}\omega_{\mu\nu}\omega^{\mu\nu} + \frac{1}{2}m_\omega^2\omega_\mu\omega^\mu - \frac{1}{4}\bar{\rho}_{\mu\nu}\bar{\rho}^{\mu\nu} + \frac{1}{2}m_\rho^2\bar{\rho}_\mu\cdot\bar{\rho}^\mu - U(\sigma) \\ +\frac{\kappa}{24}g_{\omega B}^4(\omega^\mu\omega_\mu)^2 + \frac{\lambda}{24}g_{\rho B}^4(\bar{\rho}^\mu\cdot\bar{\rho}_\mu)^2 + g_{\rho B}^2 f(\sigma,\omega^\mu\omega_\mu)\bar{\rho}^\mu\cdot\bar{\rho}_\mu\ . \end{aligned} \tag{2.1}$$

$\bar{\tau}_B$ is the isospin operator, and $U(\sigma)$ represents the self interaction terms, and can be expanded as

$$U(\sigma) = \frac{\zeta}{6}(g_{\sigma B}\sigma)^3 + \frac{\xi}{24}(g_{\sigma B}\sigma)^4\ , \tag{2.2}$$

and

$$f(\sigma,\omega^\mu\omega_\mu) = \sum_1^6 a_i\sigma^i + \sum_1^3 b_j(\omega^\mu\omega_\mu)^j\ . \tag{2.3}$$

There are 17 parameters in this model, which provide sufficient freedom to fine-tune the low and high density parts of the isospin sector independently [133]. The two EoSs derived from this Lagrangian density are known as SFHo and SFHx where 'o' stands for optimal and 'x' stands for extremal. For the SFHo EoS, the most probable mass-radius curve of Ref. [156] was fitted whereas for the SFHx EoS, the radius of low mass neutron stars was minimised resulting in a low value for the density slope of the symmetry energy at saturation density (23.18 MeV) [133]. By neglecting the last two terms of the Lagrangian density given by Eq. (2.1), it reduces to the TM1 and TMA EoSs [128,129,178]. In this case too, a unified EoS was constructed based on the NSE model for the low density matter. All four EoSs are publicly available on CompOSE [179].

Non-unified EoSs in the Boguta model

We also use non-unified EoSs where the high density part is based on the Boguta model. For these EoSs, we use the model description given by Eq. (2.1), however excluding the last three terms. In other words, we use the Boguta model with self-interaction terms of

only the σ-mesons. As mentioned earlier, the low density part of the EoS is described by the BPS model [171] for the outer crust and NV model [172] for the inner crust. We construct few such non-unified EoSs using different values of incompressibilty and effective mass. In this thesis, we denote these EoSs as Glendenning (G) followed by the value of incompressibility and either 'a' or 'b' for the two values of effective masses, 0.78 and 0.7, respectively. For example, the EoS with incompressibility, K = 240 MeV, and effective mass, $m^*/m = 0.78$, is denoted as G240a.

EoSs in the DDRH field theory model at finite temperature

The Lagrangian density of this model is given by [131, 144],

$$\begin{aligned}\mathcal{L}_B &= \sum_{B=n,p,\Lambda} \bar{\Psi}_B \left(i\gamma_\mu\partial^\mu - m_B + g_{\sigma B}\sigma - g_{\omega B}\gamma_\mu\omega^\mu - g_{\rho B}\gamma_\mu\boldsymbol{\tau}_B\cdot\boldsymbol{\rho}^\mu - g_{\phi B}\gamma_\mu\phi^\mu\right)\Psi_B \\ &+\frac{1}{2}\left(\partial_\mu\sigma\partial^\mu\sigma - m_\sigma^2\sigma^2\right) - \frac{1}{4}\omega_{\mu\nu}\omega^{\mu\nu} + \frac{1}{2}m_\omega^2\omega_\mu\omega^\mu - \frac{1}{4}\boldsymbol{\rho}_{\mu\nu}\cdot\boldsymbol{\rho}^{\mu\nu} + \frac{1}{2}m_\rho^2\boldsymbol{\rho}_\mu\cdot\boldsymbol{\rho}^\mu \\ &-\frac{1}{4}\phi_{\mu\nu}\phi^{\mu\nu} + \frac{1}{2}m_\phi^2\phi_\mu\phi^\mu \ . \end{aligned} \tag{2.4}$$

Here m_B is the bare baryon mass, $\boldsymbol{\tau_B}$ is the isospin operator and Ψ_B denotes the isospin multiplets for baryons. The parameterization of this model involving only nucleons is the DD2 EoS [131]. The Lagrangian can be extended to include Λ hyperons and the resulting model is the BHB$\Lambda\phi$ EoS [144].

The partition function in the mean field approximation can be written as,

$$\begin{aligned}\ln Z_B &= \beta V[-\frac{1}{2}m_\sigma^2\sigma^2 + \frac{1}{2}m_\omega^2\omega_0^2 + \frac{1}{2}m_\rho^2\rho_{03}^2 + \frac{1}{2}m_\phi^2\phi_0^2 + \Sigma^r\sum_{B=n,p,\Lambda} n_B] \\ &+2V\sum_{i=n,p,\Lambda}\int\frac{d^3k}{(2\pi)^3}[\ln(1+e^{-\beta(E^*-\nu_i)}) + \ln(1+e^{-\beta(E^*+\nu_i)})] \ , \end{aligned} \tag{2.5}$$

where the temperature appears as $\beta = 1/T$, $E^* = \sqrt{(k^2 + m_B^{*2})}$ and effective baryon mass $m_B^* = m_B - g_\sigma\sigma$. The chemical potential of i-th baryon is given by

$$\mu_i = \nu_i + g_{\omega B}\omega_0 + g_{\rho B}\tau_{3B}\rho_{03} + g_{\phi B}\phi_0 + \Sigma^r \ , \tag{2.6}$$

and the rearrangement term which takes care of many-body correlations, has the form,

$$\Sigma^r = \sum_{B=n,p,\Lambda}[-\frac{\partial g_{\sigma B}}{\partial n_B}\sigma n_B^s + \frac{\partial g_{\omega B}}{\partial n_B}\omega_0 n_B + \frac{\partial g_{\rho B}}{\partial n_B}\tau_{3B}\rho_{03}n_B + \frac{\partial g_{\phi B}}{\partial n_B}\phi_0 n_B] \ . \tag{2.7}$$

The total partition function of the system is $Z = Z_B Z_L$ where Z_L denotes the partition function for non-interacting leptons.

We obtain the equations of motion for the meson fields by extremising the partition function Z_B. Furthermore, we can compute all the thermodynamic quantities of baryonic matter

using the partition function Z_B. The baryon pressure is written as $P = TV^{-1} \ln Z_B$ and the energy density of baryons is given by,

$$\begin{aligned} \epsilon &= \frac{1}{2}m_\sigma^2\sigma^2 + \frac{1}{2}m_\omega^2\omega_0^2 + \frac{1}{2}m_\rho^2\rho_{03}^2 + \frac{1}{2}m_\phi^2\phi_0^2 \\ &\quad +2\sum_{i=n,p,\Lambda}\int \frac{d^3k}{(2\pi)^3}E^*\left(\frac{1}{e^{\beta(E^*-\nu_i)}+1} + \frac{1}{e^{\beta(E^*+\nu_i)}+1}\right). \end{aligned} \tag{2.8}$$

The number density of $i(= n, p, \Lambda)$-th baryon is,

$$n_i = 2\int \frac{d^3k}{(2\pi)^3}\left(\frac{1}{e^{\beta(E^*-\nu_i)}+1} - \frac{1}{e^{\beta(E^*+\nu_i)}+1}\right). \tag{2.9}$$

The scalar density, n_B^s, for baryon B is

$$n_B^s = 2\int \frac{d^3k}{(2\pi)^3}\frac{m_B^*}{E^*}\left(\frac{1}{e^{\beta(E^*-\nu_B)}+1} + \frac{1}{e^{\beta(E^*+\nu_B)}+1}\right). \tag{2.10}$$

The entropy density of baryons follows from the relation,

$$S = \beta\left(\epsilon + P - \sum_{i=n,p,\Lambda}\mu_i n_i\right), \tag{2.11}$$

and the entropy density per baryon is $s = S/n_b$ where n_b is the total baryon density.

The nucleon-meson couplings in the DDRH model are density dependent. The DD2 parameter set of nucleon-meson couplings is used to describe the nuclear matter properties [131, 180]. The functional forms of density dependent couplings $g_{\sigma N}$ and $g\omega N$ are given by,

$$g_{\alpha N} = g_{\alpha N}(n_0) f_\alpha(x), \quad \text{where} \quad f_\alpha(n_b/n_0) = a_\alpha \frac{1 + b_\alpha(x + d_\alpha)^2}{1 + c_\alpha(x + d_\alpha)^2}. \tag{2.12}$$

Here, n_0 is the saturation density, $\alpha = \sigma, \omega$ and $x = n_b/n_0$. For the ρ-mesons, we have,

$$g_{\rho N} = g_{\rho N}(n_0)\exp\left[-a_\rho(x-1)\right]. \tag{2.13}$$

The coefficients in both Eqs. (2.12) and (2.13), saturation density, nucleon-meson couplings at the saturation density, mass of the σ-mesons, are all obtained by fitting the properties of finite nuclei [131]. The properties of symmetric nuclear matter at the saturation density ($n_0 = 0.149065$ fm^{-3}) are consistent with the experimental values [119]. The symmetry energy (32.73 MeV) and its density slope (57.94 MeV) are in consonance with current experimental constraints and observations of neutron stars [118, 181, 182]. Furthermore, the DD2 EoS is reasonably compatible with that of pure neutron matter obtained in the chiral effective field theory [119, 122].

On the other hand, Λ hyperon-vector meson couplings are determined from the SU(6) symmetry relations [183, 184] and the Λ hyperon-scalar meson coupling is extracted from the hypernuclei data. A potential depth of -30 MeV is considered for the Λ hyperon in normal nuclear matter [185–187]. The hyperon-meson couplings used in this thesis are taken from Ref. [144]. Both DD2 and BHB$\Lambda\phi$ EoSs are publicly available from CompOSE [179].

Hybrid EoS at zero temperature

We also consider an EoS undergoing a first order phase transition from hadronic to quark matter governed by the Gibbs phase rules [46]. In this case, the hadronic matter is described by an extended version of the DD2 Lagrangian density of Eq. (2.4). The Lagrangian density is extended to include all hyperons of the 1/2-spin baryon octet and Δ resonance [140]. The three flavour quark matter is described by the nonlocal extension of the Nambu-Jona-Lasino (NJL) model as introduced in Ref. [140]. This hybrid EoS is calculated at zero temperature.

Chiral Mean Field model at finite temperature

The Chiral Mean Field (CMF) is an effective relativistic field theory model that is chirally invariant [188]. This model is based on the non-linear realization of SU(3) σ-model. Pseudoscalar mesons in this model act as angular parameters for chiral transformation. The CMF Lagrangian density in the mean field approximation is given by,

$$\mathcal{L} = \mathcal{L}_{\rm kin} + \mathcal{L}_{\rm int} + \mathcal{L}_{\rm self} + \mathcal{L}_{\rm SB}\,, \tag{2.14}$$

where $\mathcal{L}_{\rm kin}$ corresponds to the kinetic energy term, $\mathcal{L}_{\rm int}$ to the interaction term for baryons interacting with vector and scalar mesons, $\mathcal{L}_{\rm self}$ to the self-interaction term of vector and scalar mesons, and $\mathcal{L}_{\rm SB}$ to the explicit chiral symmetry breaking term generating masses of pseudoscalar mesons [189]. These terms can be expanded as,

$$\begin{aligned}
\mathcal{L}_{\rm int} &= -\sum_i \bar{\psi}_i[\gamma_0(g_{i\omega}\omega + g_{i\phi}\phi + g_{i\rho}\tau_3\rho) + M_i^*]\psi_i\,, \\
\mathcal{L}_{\rm self} &= \frac{1}{2}\left(m_\omega^2\omega^2 + m_\rho^2\rho^2 + m_\phi^2\phi^2\right) + g_4\left(\omega^4 + 3\omega^2\phi^2 + \frac{\phi^4}{4} + \frac{4\omega^3\phi}{\sqrt{2}} + \frac{2\omega\phi^3}{\sqrt{2}}\right) \\
&- k_0\left(\sigma^2+\zeta^2+\delta^2\right) - k_1\left(\sigma^2+\zeta^2+\delta^2\right)^2 \\
&- k_2\left(\frac{\sigma^4}{2} + \frac{\delta^4}{2} + 3\sigma^2\delta^2 + \zeta^4\right) - k_3\left(\sigma^2-\delta^2\right)\zeta - k_4\ln\frac{(\sigma^2-\delta^2)\zeta}{\sigma_0^2\zeta_0}\,, \\
\mathcal{L}_{\rm SB} &= -m_\pi^2 f_\pi\sigma - \left(\sqrt{2}m_k^2 f_k - \frac{1}{\sqrt{2}}m_\pi^2 f_\pi\right)\zeta\,.
\end{aligned} \tag{2.15}$$

The CMF EoS as a function of density, temperature and positive charge fraction is publicly available in a tabular form on CompOSE [189]. Additionally, see Refs. [31, 190] for adaptations of the described model.

Table 2.1: Saturation properties of nuclear matter such as saturation density (n_0), dimensionless effective nucleon mass (m^*/m), binding energy (BE), incompressibility (K), symmetry energy (S), and density slope of symmetry energy (L) are obtained using the different parameters for the EoSs listed below. The maximum mass of non-rotating neutron stars and the corresponding baryonic mass of the EoSs are also mentioned. The experimental values of nuclear matter properties at saturation density (n_0), as indicated in the final row, are sourced from Refs. [119, 191–193]. For the maximum mass, we quote the mass of the pulsar PSR J0740+6620 [53]. See Ref. [1] for the original publication.

EoS	n_0 (fm^{-3})	m^*/m	BE (MeV)	K (MeV)	S (MeV)	L (MeV)	M_{max} ($\mathrm{M}_\odot$)	M_B ($\mathrm{M}_\odot$)
DD2	0.149	0.56	16.02	243.0	31.67	55.04	2.42	2.89
BHB$\Lambda\phi$	0.149	0.56	16.02	243.0	31.67	55.04	2.10	2.43
SFHo	0.158	0.76	16.19	245.4	31.57	47.10	2.06	2.43
SFHx	0.160	0.72	16.16	238.8	28.67	23.18	2.13	2.53
TM1	0.146	0.63	16.30	281.6	36.95	110.99	2.21	2.30
TMA	0.147	0.64	16.03	318.2	30.66	90.14	2.02	2.30
G230a	0.153	0.78	16.30	230.0	32.50	89.76	2.01	2.31
G230b	0.153	0.70	16.30	230.0	32.50	94.46	2.33	2.75
G240a	0.153	0.78	16.30	240.0	32.50	89.70	2.02	2.75
G240b	0.153	0.70	16.30	240.0	32.50	94.39	2.34	2.75
G300a	0.153	0.78	16.30	300.0	32.50	89.33	2.08	2.40
G300b	0.153	0.70	16.30	300.0	32.50	93.94	2.36	2.78
Hybrid	0.149	0.56	16.02	243	31.67	55.04	2.05	2.39
Exp.	0.151	0.55-0.75	16.00	240±10	29-31.7	45-61.9	2.08±0.07	-

We present the saturation properties of the EoSs described above in Table 2.1. The empirical values of nuclear matter properties are reported in the last row. The reported range of incompressibility (K) values of nuclear matter at the saturation density can be found in Refs. [119, 191]. It was demonstrated that the effective masses of nucleons in the range $0.55 \leq m^*/m \leq 0.75$ lead to the physical solution for pure neutron matter which is compatible with the chiral effective field theory [192]. The bounds on the symmetry energy (S) and the symmetry energy slope (L) are taken from Ref. [181]. From Table 2.1, it is evident that the symmetry energy and slope parameter of several EoS models (SFHx, TM1, TMA, G230a, G230b, G240a, G240b, G300a, G300b) are in tension with the new bounds from Ref. [181] and state-of-the-art calculations in chiral effective field theory [119, 181, 182].

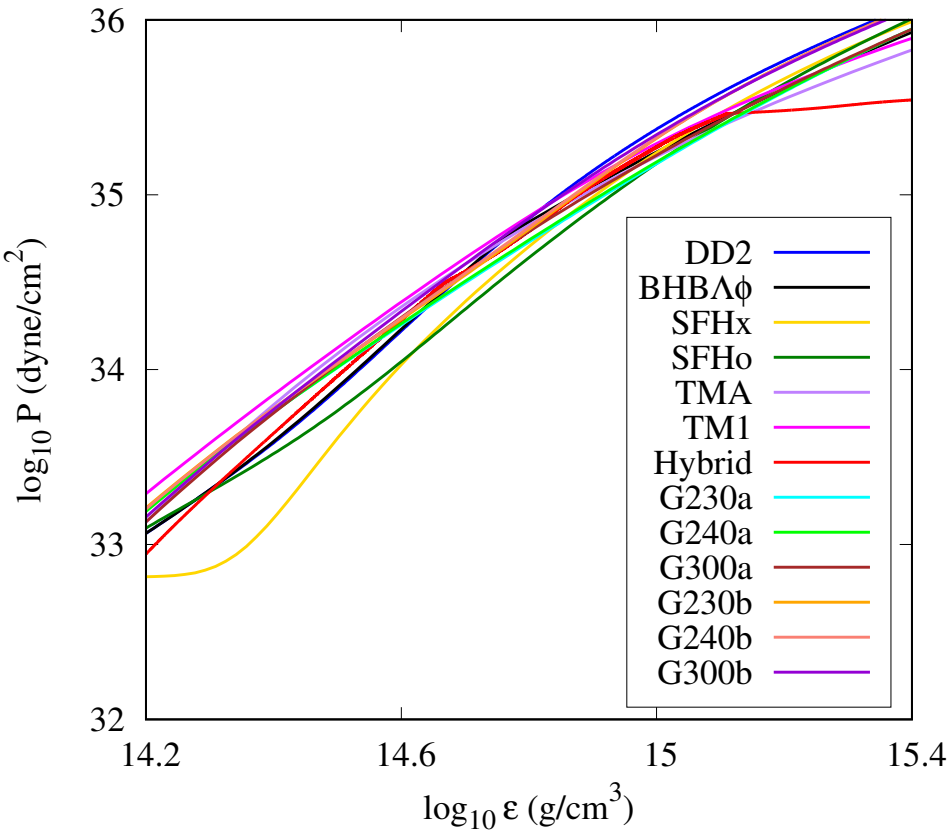

Figure 2.1: Pressure plotted as a function of energy density, for the microphysical EoS models in β-equilibrium, described in this section, at zero temperature. See Ref. [1] for original figure (DOI:10.3847/1538-4357/ab6a9e).

Note that the low value of slope parameter in the SFHx EoS model was obtained in an attempt to minimise the radii of low mass neutron stars [133]. This effect can also been seen in the depiction of the SFHx EoS in figure 2.1. Additionally, figure 2.1 contains other EoS models described in this section, where pressure is plotted as a function of energy density, at zero temperature. The low-density part of the SFHx EoS deviates significantly from other EoSs. Furthermore, it can be seen that the pressure around the saturation density remains constant. This kind of behavior was also noted in the SFHx EoS for pure neutron matter [194]. In addition, the beginning and ending of the mixed phase in the hybrid EoS are observed as two kinks. It can be observed that the hybrid EoS becomes softer once the mixed phase ends. We also plot the results of the Glendenning EoS models.

2.1.2 Piecewise Polytropic EoS Models

We can generate a multitude of equations of state spanning wide ranges in the pressure-density ($P - \rho$) space by parameterizing the EoSs in terms of piecewise polytropes. For

the work carried out in this thesis, the low-density part of each EoS ($\rho < \rho_{\rm sat}$, where $\rho_{\rm sat} \sim 2 \times 10^{14}$ g cm^{-3} is the nuclear saturation density) is assumed to conform to a conventional nuclear EoS. Nevertheless, in order to introduce robustness in the generated dataset, the low-density regime of the EoSs in the present thesis is assumed to comply with one of SLy [130], PS [126], DD2 [131] or TM1 [128] EoSs. For densities greater than the saturation density ($\rho > \rho_{\rm sat}$), one can use n-number of piecewise segments, stitched together to describe the EoS. For the purpose of adding constraints on the NS EoS, several studies with $n > 1$ segments exist in literature [153–157]. Here, we adopt the density segmentation scheme from Ref [157], i.e., any EoS can be reasonably well parameterized with five polytropic segments. The density region of the EoSs for parameterization is chosen within the range [$\rho_{\rm sat}$, $7.4\rho_{\rm sat}$]. This upper limit follows from the evidence that the pressure at this density, $P(\rho = 7.4\rho_{\rm sat})$, affects the maximum mass of the neutron star [154]. The densities are uniformly spaced on a logarithmic scale. The five segments are separated at densities (1.0, 1.4, 2.2, 3.3, 4.9, 7.4) $\rho_{\rm sat}$, as in Ref. [157]. The pressure in the $i^{\rm th}$ segment is given as a function of the density ρ,

$$P = K_i \rho^{\Gamma_i} \quad \text{for } i = [1,5], \tag{2.16}$$

where,

(i) $\rho \in [\rho_{i-1}, \rho_i]$, with ρ_{i-1} and ρ_i being the minimum and maximum densities of segment i.

(ii) $K_i = P_{i-1}/\rho_{i-1}^{\Gamma_i}$.

(iii) Following Ref. [157], the adiabatic index, Γ_i is assigned random values in the range [1, min{5, $\Gamma_{\rm luminal}$}), where $\Gamma_{\rm luminal}$ sets the limit for the causal condition, i.e., the speed of sound, c_s, does not exceed the speed of light, c. Hence, $c_s = \sqrt{dP/d\epsilon} < 1$, or

$$\Gamma \equiv \Gamma_{\rm \ luminal} \quad \text{when} \quad \frac{dP}{d\epsilon} = 1. \tag{2.17}$$

Here, ϵ is the energy density, given by

$$\epsilon = \left(\frac{\epsilon(\rho_{i-1})}{\rho_{i-1}} - \frac{P_{i-1}}{(\Gamma_i - 1)\rho_{i-1}} \right) \rho + \frac{K_i}{\Gamma_i - 1} \rho^{\Gamma_i}\,, \tag{2.18}$$

for $\Gamma \neq 1$, and

$$\epsilon = \frac{\epsilon(\rho_{i-1})}{\rho_{i-1}} \rho + K_i \ln\left(\frac{1}{\rho_{i-1}}\right) \rho - K_i \ln\left(\frac{1}{\rho}\right) \rho\,, \tag{2.19}$$

for $\Gamma = 1$ (see Ref. [157] or [154] for details).

With this prescription, we can generate several hundred thousands of polytropic EoSs. These EoSs may be flexible enough to model multiple stars' masses and radii within known constraints. A few of the generated EoSs are depicted in Figure 2.2.

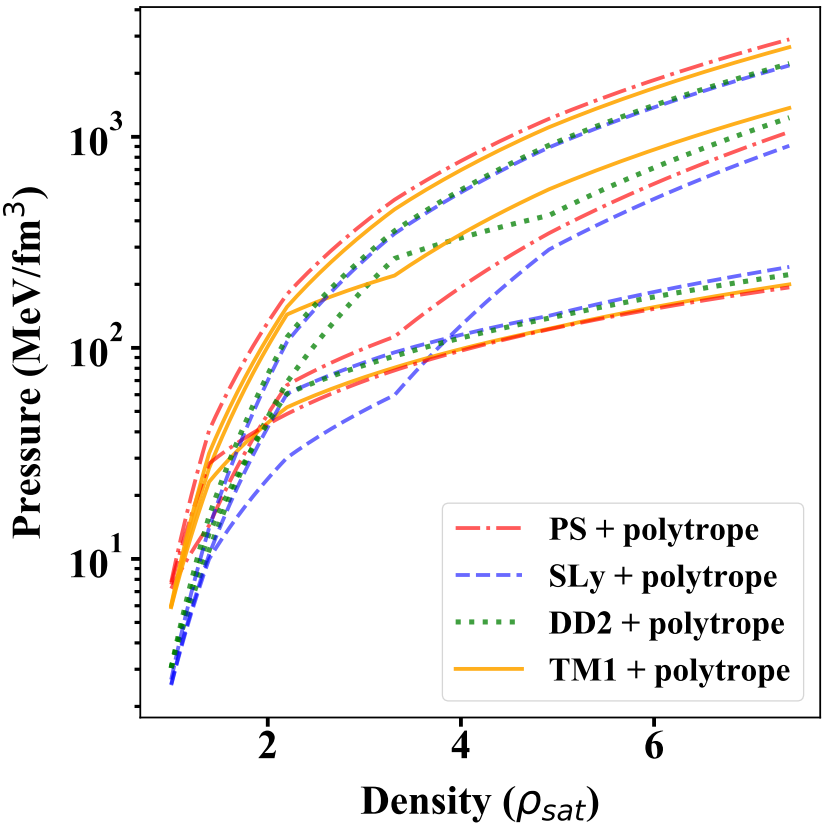

Figure 2.2: A few examples of piece-wise polytropic EoSs that were generated using PS (red), SLy (blue), DD2 (green) or TM1 (gold) EoSs for the low-density region. A major portion of the pressure range is spanned by the several polytropic EoSs. See Ref. [2] for original figure (DOI:10.1088/1475-7516/2022/08/071).

2.2 Neutron Star observables from EoS

The EoS models described in this chapter can be directly translated to observable quantities of neutron stars. Assuming gravity is described by the theory of general relativity [57–59], we provide the theoretical calculations of three observable quantities in the subsections below, namely, mass, radius and tidal deformability.

2.2.1 Stellar Structure from EoS: the TOV equations

The line element of a system containing a static, spherically symmetric star in general relativity is given by,

$$ds^2 = e^{2\nu(r)}dt^2 - e^{2\lambda(r)}dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2. \tag{2.20}$$

This can be compared to the first solutions to Einstein's equations, which were provided by Karl Schwarzschild [195, 196]. Based on the Birkhoff theorem [197, 198], the Schwarzschild solution describes the metric outside of a single non-rotating star. Therefore, for neutron stars, the exterior solution is given by the Schwarzschild metric which is connected to the interior solution, obtained by assuming a static, ideal fluid in hydrostatic equilibrium. Thus, we can derive the so-called the Tolman–Oppenheimer–Volkoff (TOV) equations from this metric, which describe the stellar structure of non-rotating NSs [39, 41]. See Ref. [46] for a detailed explanation. The TOV equations are given as,

$$-\frac{dP}{dr} = \frac{\big[\epsilon(r) + P(r)\big]\big[m(r) + 4\pi r^3 P(r)\big]}{r[r - 2m(r)]}\,, \tag{2.21}$$

and

$$\frac{dm(r)}{dr} = 4\pi r^2 \epsilon(r). \tag{2.22}$$

Here, r is the radial coordinate from the centre of the star, and $m(r)$ is the mass enclosed within the radial coordinate, r. In order to determine the observables, mass (M) and radius (R) of the star, the TOV equations are integrated radially outwards from the centre. The initial conditions taken at the centre of the star are, $r = 0,\ m(r = 0) = 0,\ \epsilon(r = 0) \neq 0,\ P(r = 0) = P_c$, where P_c is the central pressure, obtained from the EoS, usually given as a table. The radius, R, of the star is defined by the vanishing pressure condition at the surface ($P(r = R) = 0$), and the mass enclosed in R is the total mass of the star, i.e. $M = m(R)$. Thus, the TOV equations provide a direct one-to-one mapping between the neutron star EoS and the mass-radius relation. Determining the mass and the radius of a neutron star therefore provides insights on the underlying EoS. Figures 2.3 and 2.3 depict the $M - R$ relations of the microphysical and polytropic EoSs described in the previous section. The gray region on the top left corner of the bottom figure represents the causalty limit. We generate the piecewise polytropic EoSs such that they all obey this limit as mention in section 2.1.2.

2.2.2 Tidal Deformability from EoS

In a neutron star binary system, each star experiences tidal effects from its companion. These tidal effects deform the stars by inducing a quadrupole moment which leaves an imprint on the gravitational wave signal (see chapter 1, section 1.1.3). The parameter describing this effect is the so-called tidal deformability.

Mathematically, the tidal deformability (λ) of a neutron star can be expressed as the ratio of the induced quadrupole (Q_{ij}) to the external tidal field ($\mathcal{E}_{ij}$) exerted by its companion [100, 199, 200], i.e.,

$$\lambda = -\frac{Q_{ij}}{\mathcal{E}_{ij}}. \tag{2.23}$$

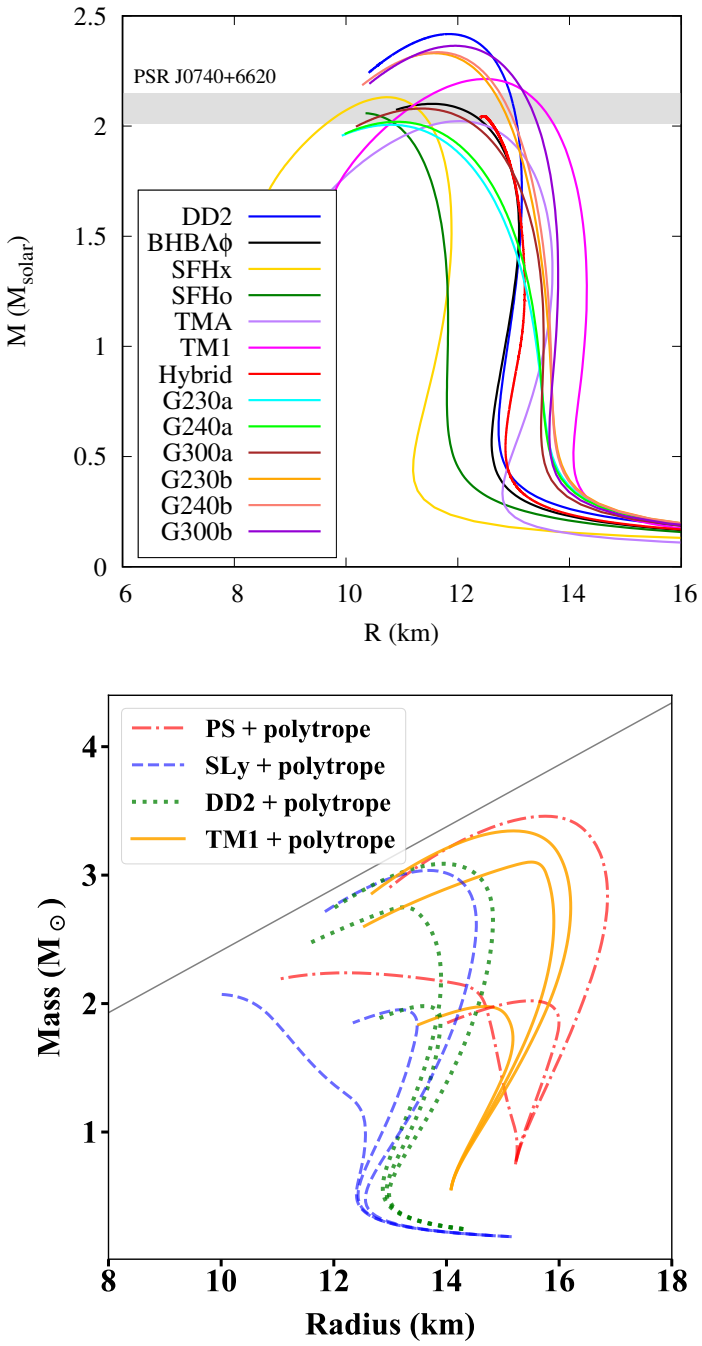

Figure 2.3: Mass-Radius ($M - R$) relations corresponding to the microphysical EoSs (top) and the polytropic EoSs (bottom) shown in figures 2.1 and 2.2 respectively. The gray region in the bottom figure represents the causalty limit. See Refs. [1] and [2] for original figures (DOI:10.3847/1538-4357/ab6a9e, DOI:10.1088/1475-7516/2022/08/071).

Note that λ is a constant, and can be derived as an inherent property of the NS based on the underlying EoS. It can be expressed as a function of the gravitational Love number (k_2) and the NS radius (R),

$$\lambda = \frac{2}{3} k_2 R^5. \tag{2.24}$$

λ is commonly used in its dimensionless form (Λ), defined as $\Lambda \equiv \lambda/M^5$, where M is the mass of the NS. The term tidal deformability in this thesis usually refers to the dimensionless quantity.

The external tidal field, $\mathcal{E}_{ij}$, and the induced quadrupole moment, Q_{ij}, emerge from the time-time component of the metric (g_{tt}) in the local asymptotic rest frame of a spherically symmetric static star of mass M, at large distances to the star r, in mass-centered Cartesian coordinates [201].

It is given by,

$$\frac{1-g_{tt}}{2} = -\frac{M}{r} - \frac{3Q_{ij}}{2r^3}\left(n^i n^j - \frac{1}{3}\delta^{ij}\right) + O\left(\frac{1}{r^4}\right) + \frac{1}{2} n^i n^j \mathcal{E}_{ij} r^2 + O(r^3). \tag{2.25}$$

Here, $n^i = x^i/r$ and δ_{ij} is the Kronecker delta. See Ref. [202] for a review.

The metric of a linear $l = 2$ perturbation on a spherically symmetric star is given by,

$$\begin{aligned} ds^2 &= e^{2\nu(r)}\big[1 + H(r)Y_{20}(\theta,\phi)\big]dt^2 \\ &\quad - e^{2\lambda(r)}\big[1 - H(r)Y_{20}(\theta,\phi)\big]dr^2 \\ &\quad - r^2\big[1 - K(r)Y_{20}(\theta,\phi)\big](d\theta^2 + \sin^2\theta d\phi^2), \end{aligned} \tag{2.26}$$

where,

$$K'(r) = H'(r) + 2H(r)\nu(r), \tag{2.27}$$

and,

$$H'' + H'\Big(\frac{2}{r} + \nu' - \lambda'\Big) + H\Big(-\frac{6e^{2\lambda}}{r^2} - 2\nu'^2 + 2\nu'' + \frac{3}{r}\lambda' + \frac{7}{r}\nu' - 2\nu'\lambda' + \frac{f}{r}(\nu' + \lambda')\Big) = 0. \tag{2.28}$$

Here, $f = d\epsilon/dp$, and the second order differential equation can be written as two first order differential equations. They are given by,

$$\frac{dH}{dr} = \beta, \tag{2.29}$$

and,

$$\begin{aligned} \frac{d\beta}{dr} &= 2\Big(1 - 2\frac{M(r)}{r}\Big)^{-1} H\Big\{ -2\pi[5\epsilon + 9p + f(\epsilon + p)] \\ &\quad + \frac{3}{r^2} + 2\Big(1 - 2\frac{M(r)}{r}\Big)^{-1}\Big(\frac{M(r)}{r^2} + 4\pi r p\Big)^2\Big\} \\ &\quad + \frac{2\beta}{r}\Big(1 - 2\frac{M(r)}{r}\Big)^{-1}\Big\{ -1 + \frac{M(r)}{r} + 2\pi r^2(\epsilon - p)\Big\}. \end{aligned} \tag{2.30}$$

Eqs. (2.29) and (2.30) can be solved simultaneously with the TOV equations (2.21) and (2.22), where the boundary conditions $H(r) = r^2$ and $\beta = 2r$ are applied at $r \to 0$. This way, we can solve for the tidal love number, k_2, which is given by,

$$\begin{aligned} k_2 \;=\; & \frac{8C^5}{5}(1-2C)^2[2+2C(y-1)-y] \;\times \\ & \big\{2C[6-3y+3C(5y-8)]+4C^3[13-11y+C(3y-2)+2C^2(1+y)] \\ & +3(1-2C)^2[2-y+2C(y-1)]\log(1-2C)\big\}^{-1}, \end{aligned} \tag{2.31}$$

where, $y = R\beta(R)/H(R)$ and the compactness, $C = M/R$. Using the estimate for k_2 in Eq. (2.24), we obtain the value for tidal deformability (Λ).

Since the tidal deformability is directly linked to the EoS, a measurement of Λ from gravitational wave observations provides observational constraints on the neutron star EoS. The $M - \Lambda$ relations corresponding to the EoSs in figure 2.1 are depicted in figure 2.4.

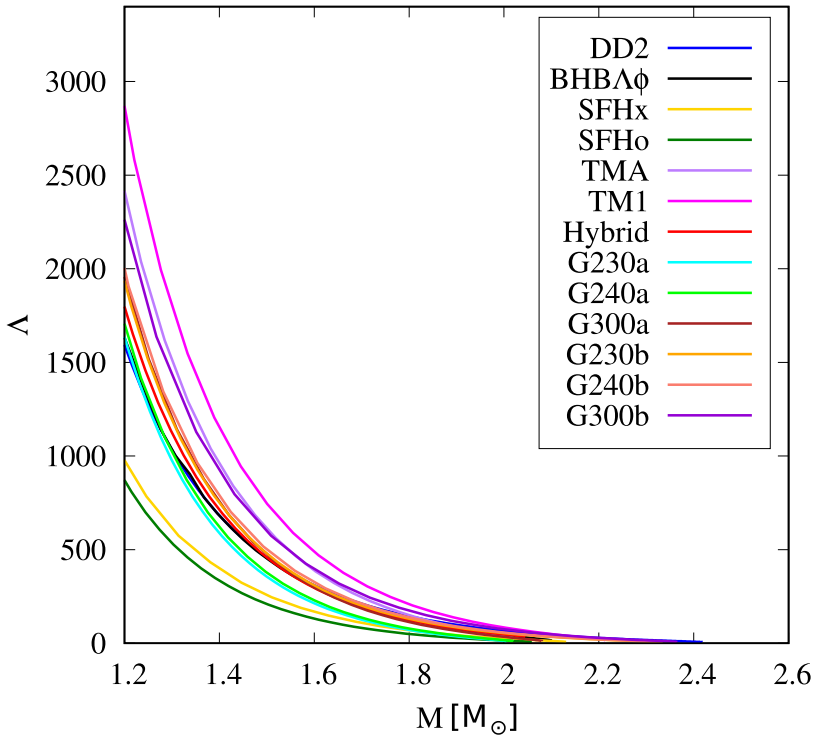

Figure 2.4: Mass-Tidal deformability (M-Λ) relations corresponding to the microphysical EoSs depicted in figure 2.1. See Ref. [6] for original figure.

The next chapter provides a brief introduction to common deep learning methods required for the algorithms used in the thesis.

Chapter 3

Deep Learning

Artificial Intelligence (AI) has been one of the most efficient tools used in the past few years while dealing with the analysis of large amounts of data. The field of AI involves Machine Learning (ML) and Deep Learning (DL), which have gained popularity in the past decade and are well acclaimed in the contemporary world [203].

Machine Learning encompasses a wide range of algorithms that enable computers to learn from data and make predictions or decisions without being explicitly programmed. It is well-suited for a wide range of tasks, including the traditional supervised learning, unsupervised learning, and reinforcement learning. Supervised learning entails algorithms which are trained to learn and make predictions based on labeled training data. The name supervised arises because the process involves a 'supervisor' who presents a set of input-output pairs (input data with corresponding labels or target values) to the algorithm to learn from. In contrast to supervised learning, unsupervised learning involves algorithms that are trained to analyze and find patterns in data without explicit supervision or labeled target outputs. In this thesis, we employ both supervised as well as unsupervised learning algorithms, as we will see in the next chapters (see chapters 4 and 5).

Deep Learning is a subset of Machine Learning that employs artificial neural networks with several interconnected layers or 'deep neural networks' (DNNs). It has been particularly influential in recent AI advancements, enabling breakthroughs in image recognition, natural language processing, and speech recognition. In this thesis, we utilize DL to develop algorithms to perform two different tasks:

1. To reconstruct the dense matter EoS from observational data of neutron stars. This is a semi-supervised learning method that employs automatic differentiation to optimize the EoS. See chapter 4 for detailed explanations.

2. To analyze gravitational wave (GW) data to estimate the source parameters. Here, we employ supervised learning methods to perform classification (of GW signals from different sources) and regression (of source properties). Refer to chapter 5 for details.

This chapter introduces the fundamental concepts and basic essentials required for implementing the tasks mentioned above.

3.1 Artificial Neural Networks

Artificial neural networks are computational models inspired by the structure and function of biological neural networks in the brain. They are powerful machine learning algorithms capable of learning complex patterns from data. A typical neural network comprises of several interconnected nodes called neurons. These neurons are organized into layers, typically consisting of an input layer, one or more hidden layers, and an output layer. Each neuron receives input signals which are then processed using weighted connections to produce an output signal. During the learning process, neural networks adjust the weights assigned to connections between neurons to optimize their performance. This adjustment is achieved through a training phase, where the network learns to map the input to the desired output. The universal function approximation theorem ensures that artificial neural networks can approximate any kind of continuous function with nonlinear activation functions [204–208]. In this section, we introduce a few neural networks that will be used in the upcoming chapters. The sections following briefly describe commonly used activations (Section 3.2), loss functions (Section 3.3) and optimization algorithms (Section 3.4) used in neural networks.

3.1.1 Fully Connected Neural Networks

A fully connected neural network (FCN) as the name suggests consists of layers that are fully connected. In other words, in fully connected layers (also known as dense layers), each neuron is connected to every neuron in the previous layer. A neuron performs a linear transformation on the input vector using a weight matrix. Given an input vector $\mathbf{x}$ of size N, and assuming the layer has M neurons, the output vector $\mathbf{y}$ of the fully connected layer can be computed as,

$$\mathbf{y} = \mathbf{W} \cdot \mathbf{x} + \mathbf{b}. \tag{3.1}$$

Here, $\mathbf{W}$ is the weight matrix of size $M \times N$, where each element W_{ij} represents the weight connecting the ith neuron in the layer to the jth neuron in the previous layer. The vector $\mathbf{b}$ represents the bias term of size M, which provides an additional learnable parameter for each neuron in the layer. Subsequently, a non-linear transformation is applied element-wise to the resulting product, i.e., the output vector $\mathbf{y}$. This transformation is done using a non-linear activation function denoted as f, and can be written as,

$$\mathbf{z} = f(\mathbf{y}). \tag{3.2}$$

The output vector $\mathbf{z}$ serves as the input to the next layer in the neural network architecture. By stacking multiple fully connected layers with appropriate activation functions, neural networks can learn hierarchical representations of the input data, enabling them to model intricate patterns and make accurate predictions.

3.1.2 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are widely used when dealing with structured grid-like data, such as images, audio, and time series data, as they have the potential to extract key features or patterns in the input data. A CNN consists of multiple layers, including convolutional layers, pooling layers, and fully connected layers. Here, we denote the input as $\mathbf{X}$ and the output feature maps as $\mathbf{Y}$. In a convolutional layer, we convolve the input $\mathbf{X}$ with a set of K trainable filters or kernels, denoted as $\mathbf{W}$. Each filter generates a corresponding feature map. Assuming the kernel has a size of $k \times k$, the convolution operation is represented as,

$$\mathbf{Y}_{i,j} = f\left(\sum_{m=1}^{K}\sum_{p=1}^{k}\sum_{q=1}^{k}\mathbf{W}_{p,q,m} \cdot \mathbf{X}_{i+p-1,j+q-1,m}\right), \tag{3.3}$$

where $\mathbf{Y}_{i,j}$ is the element at position (i, j) in the output feature map, $\mathbf{W}_{p,q,m}$ is the weight at position (p, q) of the m^{th} filter, $\mathbf{X}_{i+p-1,j+q-1,m}$ is the input value at position $(i + p - 1, j + q - 1, m)$, and f is the non-linear activation function. Note that this convolution operation can be generalized and carried out in 1D, 2D or 3D. The feature maps output from the convolutional layers are usually fed as input to pooling layers. A pooling layer reduces the spatial dimensionality of the feature maps while retaining important features. Different kinds of pooling layers include max pooling, average pooling, global pooling, sum pooling, adaptive pooling, etc. In this thesis, we make use of max pooling and average pooling layers. The max pooling operation selects the maximum value within a pooling region, and can be expressed as,

$$\mathbf{Y}_{i,j} = \max(\mathbf{X}_{p,q}), \tag{3.4}$$

where $\mathbf{Y}_{i,j}$ is the element at position (i, j) in the pooled feature map, and $\mathbf{X}_{p,q}$ represents the elements within the pooling region. Similarly, the average pooling operation takes the mean of the elements in the pooling region. The output feature maps of the final convolutional/pooling layer are flattened into a vector $\mathbf{v}$. This is followed by the fully connected layer(s). As discussed in section 3.1.1, a fully connected layer applies a linear transformation on vector $\mathbf{v}$, followed by a non-linear activation function. Therefore, by assembling few/several convolutional, pooling, and fully connected layers, CNNs can extract features from raw input data, making them powerful tools for tasks like image classification, object detection, and more.

3.1.3 Recurrent Neural Networks

Recurrent neural networks (RNNs) are typically used for sequential or time series data. They are different from traditional feedforward neural networks in the sense that they have a feedback connection that allows them to maintain an internal state or memory. This allows RNNs to handle sequential information effectively. The key feature of an RNN is the recurrent connection, which allows information to be passed from one step to the

next. This feedback loop enables the network to process input sequences of variable lengths and capture dependencies between elements in the sequence.

At each time step t, the RNN takes an input vector $\mathbf{x}_t$ and the previous hidden state $\mathbf{h}_{t-1}$ as inputs and produces an output $\mathbf{y}_t$ and an updated hidden state $\mathbf{h}_t$. The hidden state serves as the memory of the network, capturing information from past time steps. The update equations for a simple RNN can be written as,

$$\mathbf{h}_t = f(\mathbf{x}_t, \mathbf{h}_{t-1}) \tag{3.5}$$

and

$$\mathbf{y}_t = g(\mathbf{h}_t). \tag{3.6}$$

The functions f and g represent non-linear activation functions, where f is the transition function that updates the hidden state, and g is a function that maps the hidden state to the output. The RNN's hidden state $\mathbf{h}_t$ is a representation of the previous information and is influenced by both the current input $\mathbf{x}_t$ and the previous hidden state $\mathbf{h}_{t-1}$. This permits the network to capture temporal dependencies and learn patterns in the sequential data. The output $\mathbf{y}_t$ can be used for tasks such as sequence prediction or generating new sequences.

However, standard RNNs often suffer from the vanishing or exploding gradient problem, where the gradients become extremely small or large during training, making it difficult to learn long-range dependencies. One solution to these issues is to reduce the number of hidden layers within the neural network. However this approach reduces the complexity in the RNN model. Therefore, various advanced RNN architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) have been developed to incorporate gating mechanisms to improve the flow of information over time. In one of the upcoming chapters in this thesis, we make use of a variant RNN, namely the LSTM. A brief description of the LSTM is given below.

Long Short Term Memory

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture that counteracts the vanishing gradient problem and is capable of capturing long-term dependencies in sequential data. LSTMs address this issue by introducing a more complex cell state $\mathbf{c}_t$ that allows the network to learn when to forget or remember information over long sequences. The LSTM cell has three main components that include an input gate, a forget gate, and an output gate. The cell state $\mathbf{c}_t$ is updated based on these gates, which control the flow of information. The update equations for an LSTM cell at time step t can

be described as follows,

$$\begin{aligned}
\mathbf{i}_t &= f(\mathbf{W_i}\mathbf{x}_t + \mathbf{U_i}\mathbf{h}_{t-1} + \mathbf{b_i}) \quad \text{(input gate)} \\
\mathbf{f}_t &= f(\mathbf{W_f}\mathbf{x}_t + \mathbf{U_f}\mathbf{h}_{t-1} + \mathbf{b_f}) \quad \text{(forget gate)} \\
\mathbf{o}_t &= f(\mathbf{W_o}\mathbf{x}_t + \mathbf{U_o}\mathbf{h}_{t-1} + \mathbf{b_o}) \quad \text{(output gate)} \\
\tilde{\mathbf{c}}_t &= g(\mathbf{W_c}\mathbf{x}_t + \mathbf{U_c}\mathbf{h}_{t-1} + \mathbf{b_c}) \quad \text{(candidate cell state)} \\
\mathbf{c}_t &= \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t \quad \text{(update cell state)} \\
\mathbf{h}_t &= \mathbf{o}_t \odot g(\mathbf{c}_t) \quad \text{(output)}.
\end{aligned}$$

Here, $\mathbf{x}$ is the input vector, and $\mathbf{h}$ is the hidden state. Furthermore, $\mathbf{W}$'s and $\mathbf{U}$'s are the weight matrices; and $\mathbf{b}$'s, represent the bias vectors. We use f and g to represent non-linear activation functions, and $\odot$ denotes element-wise multiplication. The input gate $\mathbf{i}_t$ controls the extent to which the candidate cell state $\tilde{\mathbf{c}}_t$ is incorporated into the updated cell state $\mathbf{c}_t$. The forget gate $\mathbf{f}_t$ determines the extent to which the previous cell state $\mathbf{c}_{t-1}$ is retained. The output gate $\mathbf{o}_t$ regulates the amount of information that is output from the cell. By dynamically updating the cell state and selectively incorporating or forgetting information, LSTMs can capture long-range dependencies and store relevant information over extended sequences.

LSTMs have proven to be effective in various tasks that involve sequences. These include language modeling, machine translation, speech recognition, and time series prediction.

3.2 Activations

Activation functions are significant when it comes to introducing add non-linearities to neural networks. Therefore, they can allow complex functions and let the neural networks learn these powerful operations. In this section, we list the different activation functions used for the different neural networks in this thesis.

1. **Sigmoid Activation:**
   The functional form of the sigmoid activation is,

$$f(X) = \frac{1}{1 + e^{-X}}. \tag{3.7}$$

   Therefore, the output lies within the range (0,1) for all X.

2. **Tanh Activation:**
   The hyperbolic tangent activation function is defined as,

$$f(X) = \frac{e^{X} - e^{-X}}{e^{X} + e^{-X}}. \tag{3.8}$$

   In this case, the output lies within the range (-1,1).

3. **ReLU Activation:**
   The standard Rectified Linear Unit or 'ReLU' activation has the functional form,

$$f(X) = \max(0, X). \tag{3.9}$$

   It is a piecewise linear function which outputs zero for all negative values of X, and the value itself for non-negative X.

4. **ELU Activation:**
   The Exponential Linear Unit or 'ELU' activation is based on the ReLU function. A constant α is introduced to induce smoothness to the function. The ELU activation is defined as,

$$f(X) = \begin{cases} X, & \text{if } X > 0 \\ \alpha(e^X - 1), & \text{if } X <= 0 \end{cases}$$

   where $\alpha > 0$. The default value of α is set to 1.

5. **SELU Activation:**
   The functional form of the Scaled Exponential Unit or 'SELU' is,

$$f(X) = \begin{cases} sX, & \text{if } X > 0 \\ s\alpha(e^X - 1), & \text{if } X <= 0 \end{cases} \tag{3.10}$$

   where s is the scale. Note that when $s = 1$, we get the ELU activation. By default, the Keras Library uses $s \approx 1.05$ and $\alpha \approx 1.67$.

6. **Softmax Activation:**
   The softmax function converts a vector of values to a probability distribution, with values in the range (0,1) and a unit sum. It is defined as,

$$f(\mathbf{X})_i = \frac{e^{X_i}}{\sum_{j=1}^{N} e^{X_j}} \tag{3.11}$$

   where $\mathbf{X} = (X_1, X_2, ..., X_N)$ is the input vector and $i = 1, 2, ..., N$. The output is then a vector of probabilities. Therefore, the softmax function can be used for classification problems (as we will see in chapter 5).

Figure 3.1 summarizes four different activation functions. All four subplots are shown for input values in the range [-10,10]. On the top panels, we present the Sigmoid and Tanh functions. The sigmoid function's output is bounded between (0,1), and the tanh output lies within (-1,1). On the bottom panels, we show the ReLU and ELU activation functions. The output values on the bottom panels are displayed within range [-10,10].

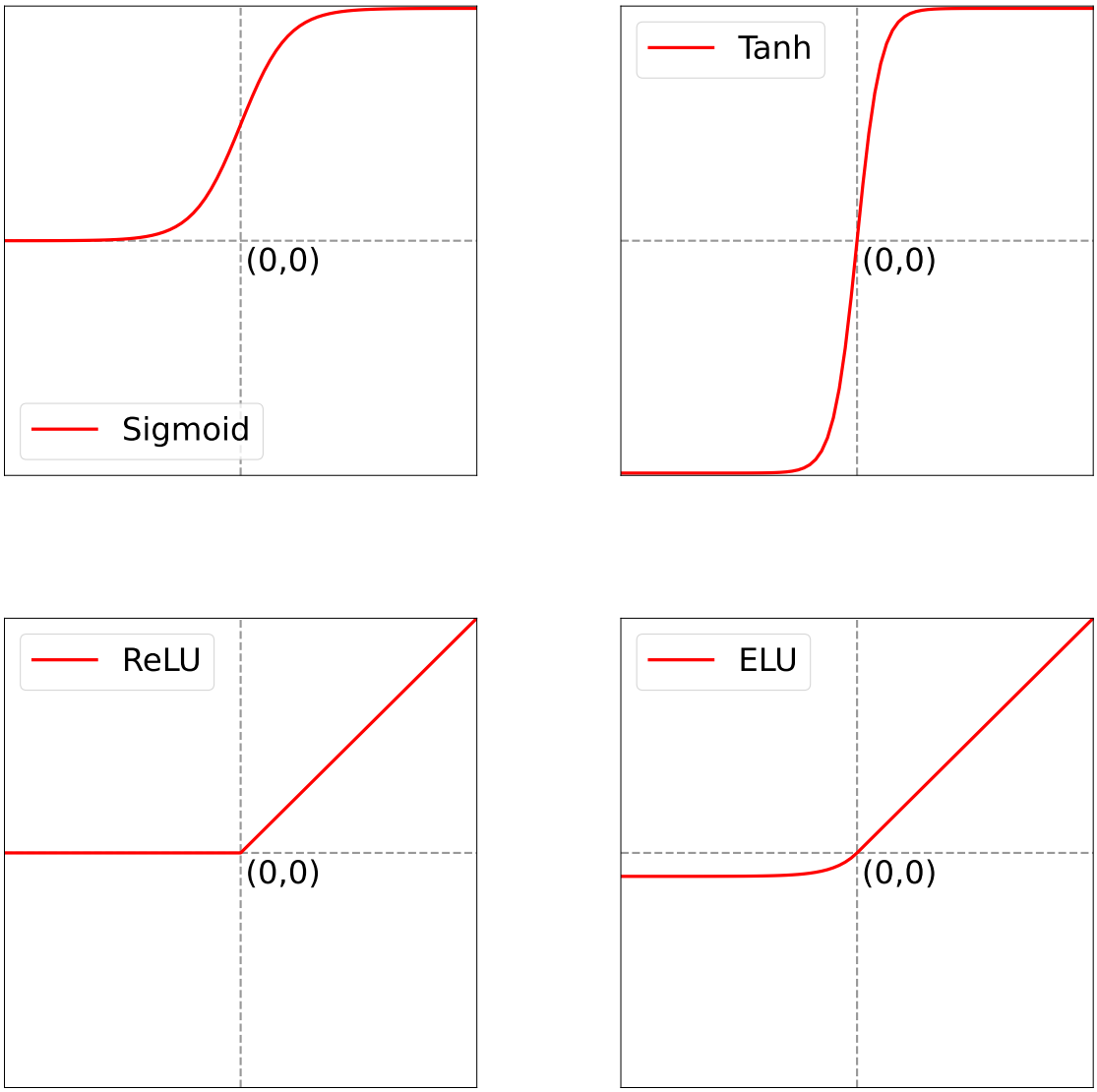

Figure 3.1: Common activation functions used for different layers of neural networks described in this thesis. All activations are displayed for input values in the range [-10,10]. The figures in the top panel depict the sigmoid and tanh activations. The outputs are bounded within (0,1) and (-1,1) for the sigmoid and tanh functions, repsectively. The ReLU and ELU activations are displayed on the bottom panel. The displayed range of the output on the bottom panels is [-10,10].

3.3 Loss Functions

The optimization of the neural networks during the learning process involves the adjustment of the network weights. This is called the training phase, and weights are updated to minimize the error in the network predictions. The quantification of this error is achieved through loss functions. The loss functions used in this thesis are listed below.

1. **MSE Loss:**
   The Mean Squared Error or 'MSE' loss function, as the name suggests, is the mean of the squared difference between the true and predicted values. Mathematically, it is written as,
   $$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{3.12}$$
   where y_i and $\hat{y}_i$ are the true and predicted labels respectively.

2. $\mathcal{R}^2$ **Loss:**
   The $\mathcal{R}^2$ loss function is given by,
   $$\mathcal{R}^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2 + \delta}, \quad \text{where } \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i. \tag{3.13}$$
   Here, y_i is the true label, $\hat{y}_i$ is the corresponding prediction, and $\bar{y}$ is the mean of the true labels. A perfect prediction, i.e., $\hat{y}_i = y_i$ yields $\mathcal{R}^2 = 1$. On the other hand, if the predictions are centered around the mean, $\bar{y}$, we obtain $\mathcal{R}^2 = 0$. .When applying the $\mathcal{R}^2$ loss numerically, we use small values of $\delta = 10^{-7}$ to avoid undefined values when encountered with a division by zero.

3. **Cross Entropy Loss:**
   The cross-entropy loss is used for classification tasks. Assuming a total number of 'C' classes, the CE loss function of a single vector label for each class $i \in C$ is defined as,
   $$\text{CE} = -\sum_{i=1}^{C} y_i \ \ln(\hat{y}_i) \tag{3.14}$$
   where y_i is the true label and $\hat{y}_i$ is the corresponding predicted probability. The loss is averaged over all the vector labels. For a binary classification, where the target labels are either 0 or 1, the CE loss is termed binary cross-entropy. Note that for the multi-class classification implemented in chapter 5 of this thesis, the labels are one-hot encoded. This implies that there is only one element of the true vector label which is non-zero. This is a special case of CE loss, called the 'categorical cross-entropy' loss for multi-class classification.

3.4 Optimization

Before beginning the optimization of the neural network, the first iteration utilizes network weights that are initialized randomly (for example, Xavier initialization [209] or *He* Normal initialization [210]). With each iteration in the training process, the weights are updated using the concept of loss functions and their minimization. There are numerous algorithms that are designed to optimize functions [211]. One of the simplest optimization algorithm is the gradient-descent.

3.4.1 Gradient-Descent

The gradient-descent algorithm iteratively calculates the updated weights using the scaled gradients of the loss function with respect to the current weights. Mathematically, it can be written as,

$$W_{\text{new}} = W_{\text{old}} - \alpha \frac{\partial L}{\partial W_{\text{old}}}. \tag{3.15}$$

The scale, α, is referred to as the learning rate. W_{new} and W_{old} are the new and old weights, respectively. The loss is denoted by L.

However, in order to calculate the derivatives of the loss function, it is necessary to compute it for each example of the training dataset. This is inefficient and a major drawback of the gradient-descent algorithm. Therefore, to overcome this problem, there is an adaptation to the gradient-descent algorithm, known as the mini-batch gradient-descent.

Mini-batch Gradient-descent

The underlying principle of updating the weights in a mini-batch gradient-descent algorithm remains consistent with the gradient-descent algorithm, but the derivative of the loss function is an approximated value, i.e., it is computed only on a small batch (a mini-batch) from the dataset. The choice of batch size for derivative estimation is one of the hyperparameters in building the neural network.

Stochastic Gradient-descent

An extreme alteration of the mini-batch gradient-descent is the stochastic gradient descent (SGD). Here, the batch size is 1, i.e., the weights are updated based on every training sample. Since the derivative in SGD is not exact, the descent is not always in the optimal direction. In order to resolve this issue, a momentum term can be added to the SGD algorithm, giving rise to SGD with momentum[1].

[1]Note that stochastic gradient descent can sometimes be used to refer to the mini-batch gradient descent in references.

Stochastic Gradient-descent with Momentum

Stochastic gradient-descent with momentum is a stochastic optimization method that helps accelerate gradient descent vectors in the optimal directions. Momentum, or the moving average of the gradients, essentially accumulates the gradient of the past steps to determine the right direction. This leads to a quicker convergence. Mathematically, the momentum vector (m_t) at time-step t, can be written as,

$$m_t = \eta m_{t-1} - \alpha \frac{\partial L}{\partial W_{\text{old}}}, \tag{3.16}$$

where $\eta \in [0, 1]$ is a hyperparameter, and α is scaled with (1-η). The weight is updated as,

$$W_{\text{new}} = W_{\text{old}} + m_t. \tag{3.17}$$

Note that the advantage of SGD with momentum is realized only after the first few updates when the knowledge of the previous gradients is utilized to accelerate to the optimal direction.

3.4.2 Root Mean Square Propagation

The fundamental principle of the Root Mean Square Propagation (RMSProp) algorithm, is to use the moving average of the squared gradients for each network weight, and then, as the name suggests, to divide the gradient by the root mean square. RMSProp is an adaptive learning rate optimization method. The mathematical representation of the algorithm is given by the equations,

$$E[g^2]_t = \beta E[g^2]_{t-1} + (1 - \beta) g_t^2 \tag{3.18}$$

and,

$$W_t = W_{t-1} - \frac{\eta}{\sqrt{E[g^2]_t + \epsilon}} g_t. \tag{3.19}$$

Here, $E[g^2]_t$ is the moving average of the squared gradient at time step t, and,

$$g_t = \frac{\partial L}{\partial W_t}, \tag{3.20}$$

is the gradient of the loss function with respect to the weight W, at time-step t. The hyperparameter β is called the decay rate, and η is the learning rate. In order to prevent divisions by zero, ϵ is introduced as a smoothing term. This way, the learning rate is divided by an exponentially decaying value, making the learning process adaptive.

3.4.3 Adaptive Moment Estimation (Adam)

Adam is an adaptive learning rate optimization algorithm that combines RMSProp and SGD with momentum [212]. Adam uses the moving average of past gradients to compute

the direction of parameter update, similar to momentum. Just like RMSProp, Adam also utilizes an exponentially decaying average of past squared gradients, essentially resulting in an adaptive learning rate method. The update equations of the Adam optimizer can be expressed as,

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \tag{3.21}$$

and,

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2. \tag{3.22}$$

m_t and v_t are the first and the second moment estimates of the gradients, respectively. These estimates are not perfect, and are biased towards zero due to the initialization of the m_t and v_t vectors as zeros. Therefore, bias-correction terms to the first and second moment estimates can be written as,

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \tag{3.23}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}. \tag{3.24}$$

Finally, we get the weight update equation as,

$$W_t = W_{t-1} - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t. \tag{3.25}$$

Throughout this thesis, we utilize the Adam optimizer. Additional details on optimizers can be found in Ref. [211]. The computation of gradients in all these methods is carried out via backpropagation [208, 213]. Backpropagation involves the gradient calculation starting with respect to the weights of the penultimate layer, and eventually proceeding backwards, propagating the error through the hidden layers to consequently estimate all the gradients.

Ultimately, the optimization of the weights is an iterative process, where the network learns with the entire training dataset multiple times or 'epochs'. The training process typically includes as many epochs as required until the loss no longer decreases with additional training. The quality of the trained network can then be evaluated on a validation dataset, and higher level hyperparameters can be updated at this stage. Finally, a testing dataset is used to obtain an unbiased evaluation of the trained network. The learning curves, which are a function of the training and testing losses (or accuracies) and epochs, provide us with information on whether the network is overfitting or underfitting the data. A network that overfits the data has a decreasing training loss, but an increasing testing loss as the number of epochs increases. Similarly, if the training loss has not yet reached a plateau, the network is underfit to the data. Therefore, the learning curves are used to monitor the network training process. In order to achieve the best network performance, methods like early stopping can be applied, i.e., the training is discontinued at the epoch when the error on the validation set grows. This way, deep learning applications have the capability to capture complex nonlinear correlations in data, and hence have been employed to solve a number of physical problems. Some of them include: determining

the parton distribution function [214, 215], reconstructing the spectral function [216–218], identifying phase transitions [219–224], assisting lattice field theory calculations [225–228], evaluating centrality distributions for heavy ion collisions [229–231], parameter estimation under detector effects [232, 233], and speeding up hydrodynamic simulations [234].

In this thesis, we apply the deep learning concepts introduced here to develop novel algorithms and achieve high performances of NNs for constraining the dense matter EoS. Chapters 4 and 5 are dedicated to the development of these algorithms.

Chapter 4

A DL Reconstruction of the NS EoS from Observations

In the first two chapters, we have discussed the observable quantities of neutron stars and the theoretical EoS models used to predict the global properties of NSs. Traditionally, the theoretical models are developed and confronted with the available observational data, to either fulfill the current constraints or are ruled out. This method can however be strongly model-dependent. It is also possible that the existing NS observations are not systematically translated to the EoS or the dense matter properties. Other approaches to study dense matter properties include reconstructing the EoS from NS observables. The most conventional method is the widely recognized Bayesian inference method [156, 235, 236]. This statistical method, however, typically also rely on parameterizations of the EoS, like piecewise polytropic EoSs [153, 157] or spectral representations [162]. These parameterizations of the EoS, therefore, make the Bayesian methods model-dependent as well. Alternatively, one could use Gaussian processes [168, 169] or neural network representations of the EoS in order to overcome this model-dependency.

An effective approach to directly exploit the non-linear mapping between a neutron star's structural properties (specifically, the mass-radius or $M - R$ relations) and its underlying EoS, is the machine learning inference in the sense of supervised learning. Recent studies based on this approach include works by Fujimoto et al. [159–161], Morawski et al. [237], Ferreira and Providência [238], and Krastev [239]. However, the method in Refs. [159–161] utilizes the speed of sound representation for the EoS, integrating a model-dependency. In Ref. [237], Morawski et al. employed the encoder-decoder architecture of an Auto-Encoder to reconstruct the EoS without relying on any parametric representation. Ferreira and Providência utilized Support Vector Machines to regress the EoS in terms of nuclear matter parameters in Ref. [238]. Krastev demonstrated in Ref. [239] that, a trained feedforward neural network, when given observational neutron star data, is capable of extracting the density dependence of the nuclear symmetry energy and therefore indirectly the EoS.

In this chapter, we implement for the first time, an unsupervised deep learning algorithm to reconstruct the dense matter EoS. It has been shown in earlier works which incorporated DL methods that deep neural networks (DNNs) can potentially surpass traditional methods in solving inverse design problems [240–242]. Here, we test a DNN's capacity to invert the TOV equations (see section 2.2.1 of chapter 2), in order to reconstruct the dense matter EoS from a limited number of $M-R$ observations of NSs. A physics-based method is introduced in the Automatic Differentiation (AD) framework. We use a neural network representation of the EoS, called the `EoS Network` for model-independency. The EoS reconstruction is achieved through the statistical inference from $M-R$ observations in an unbiased manner. The design of the inverse problem stated necessitates a functional form of the TOV equations that is easily differentiable. For this purpose, we train a DNN to solve the TOV equations, essentially creating an emulator that is called the `TOV-Solver Network`. Details on the implemented method follow in the next section.

4.1 Automatic Differentiation to reconstruct the EoS

Figure 4.1 summarizes our designed scheme in a flow-chart. We introduce a primary neural network called the `TOV-Solver Network`, in Part(a) of figure 4.1. The `TOV-Solver Network`, as the name suggests, is trained using supervised learning methods, to solve the TOV equations, i.e. to map an EoS to its corresponding $M-R$ curve. The trained network model is then saved as a mapping from the EoS (x) to its mass-radius curve ($\mathbf{z}$), where $\mathbf{z} = f(x)$. Here, the EoS is input to the `TOV-Solver Network`, represented as $x = P_i(\rho_i)$, where P_i is the pressure at density ρ_i (Note that ρ is a 1D array of densities arranged in increasing order). The output of the network, $\mathbf{z} = (M_i, R_i)$, represents the corresponding mass-radius pairs. Furthermore, the `TOV-Solver Network` is designed with several layers, a series of differentiable modules, which incorporate both linear transformations and nonlinear activation functions. The pre-trained network is then loaded into the next step of our scheme, as shown in Part(b) of figure 4.1. This means that the well-trained parameters of the `TOV-Solver Network` are frozen in subsequent procedures, and are not subject to further training. Figure 4.1(b) additionally depicts a secondary neural network, the `EoS Network`, which is tailored to permit unbiased and flexible representations of the EoS. The EoS representation can be written as $P_\theta(\rho)$, where $\{\theta\}$ is the parameter set of the `EoS Network`. The output of the `EoS Network` is linked to the well-trained `TOV-Solver Network`, as shown in figure 4.1(b), and is optimized in an unsupervised manner to fit the output of the `TOV-Solver Network` to $M-R$ observations. The optimization procedure of the `EoS Network` involves reducing the loss function, which is the standard χ^2 between the observations and the predictions of our designed pipeline "`EoS Network` $\rightarrow$ `TOV-Solver Network`". Assuming we have N_{obs} number of $M-R$ observations, the loss function is defined by

$$\chi^2 = \sum_{i=1}^{N_{obs}} \frac{(M_i - M_{obs,i})^2}{\Delta M_{obs,i}^2} + \frac{(R_i - R_{obs,i})^2}{\Delta R_{obs,i}^2}. \tag{4.1}$$

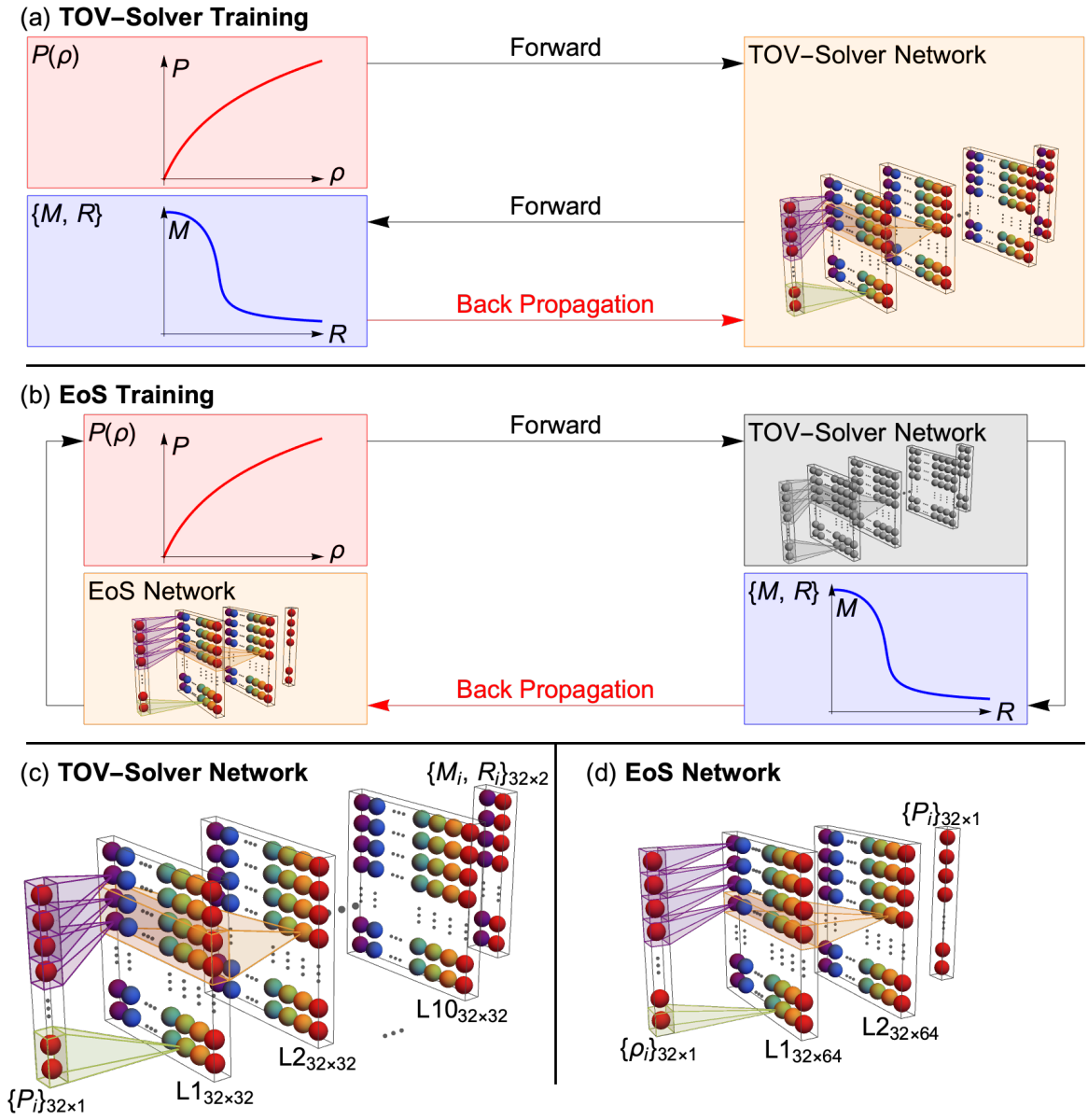

Figure 4.1: Flow chart outlining the methodology implemented toreconstruct the EoS. It consists of two main stages: training a `TOV-Solver Network` (a) and training an `EoS Network` (b). Note that in (b) the `TOV-Solver Network` is well-trained and the weights are frozen. The architectures of the `TOV-Solver Network` and `EoS Network` are illustrated in (c) and (d) respectively. The network nodes are represented by colored spheres, and the colors indicate the indices across the width of each layer. The connections between nodes across layers are depicted as shadow lines, representing a 1D convolutional kernel operation with trainable parameters. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

In Eq. (4.1), if the $M-R$ observations are denoted as (M_{obs}, R_{obs}), and their corresponding uncertainties as $(\Delta M_{obs}, \Delta R_{obs})$, then (M_i, R_i) denotes the predicted output of the i^{th} observation, i.e. $(M_{obs,i}, R_{obs,i})$. Since the parameters of the `TOV-Solver Network` are fixed, we can derive their gradients as

$$\frac{\delta\chi^2}{\delta\theta} = \frac{\delta\chi^2}{\delta\mathbf{z}}\frac{\delta\mathbf{z}}{\delta P_\theta}\frac{\delta P_\theta}{\delta\theta}. \tag{4.2}$$

The last two terms of Eq. (4.2) by definition, are computed in the back-propagation algorithm [208] within the AD framework for the designed scheme in figure 4.1(b).

In other words, we are essentially fine-tuning the parameters of the `EoS Network` to obtain the desired $M-R$ curve, after going through the frozen `TOV-Solver Network`, with guidance from a limited set of observational data. The architectures of the `TOV-Solver Network` and the `EoS Network` are shown in Figure 4.1 as part (c) and (d), respectively. We utilized the Python Library Keras [243], which is built on the Tensorflow platform [244], to setup the network models and to perform the AD calculations for optimizing the NS EoS.

4.2 TOV-Solver Network

In this section, we provide details on the `TOV-Solver Network` like data generation and network architecture. The data preparation includes techniques described in chapter 2.

4.2.1 TOV-Solver Network: Data Generation

In order to train the `TOV-Solver Network` on mapping any given EoS to its corresponding $M-R$ curve, we need sufficient training data. Since the number of available microphysical EoS models are limited, we use parameteric representatinons to generate EoSs in bulk. In particular, we employ the piecewise polytropic EoS model. The description of these parameterized EoS models is detailed in section 2.1.2 of chapter 2. Note that we initially use only 3 different EoSs (SLy [130], PS [126], or DD2 [131] EoS) to describe matter at low densities ($\rho < \rho_{sat}$). For training the `TOV-Solver Network`, we generate 100,000 polytropic EoSs for each category of low-density EoSs (SLy, PS, and DD2), leading to a total of 300,000 EoSs. These EoSs serve as training and testing input to the `TOV-Solver Network`. The $M-R$ sequences corresponding to these EoSs are calculated using the Tolman–Oppenheimer–Volkoff (TOV) equations [39, 41], which are described in section 2.2.1 of chapter 2.

This way, we prepare the data for the `TOV-Solver Network`. From this set of generated data, we exclude all the EoSs (and their corresponding $M-R$ sequences) which fail to accommodate a neutron star of mass $1.9M_\odot$. This preference of a conservative limit follows from the observations [53, 54, 245]. Moreover, the data is merely used to create an emulator for solving the TOV-equations, irrespective of the validity of the EoS. A lower limit however, can be used as the lower bound output by the network. On applying the cut-off, we are

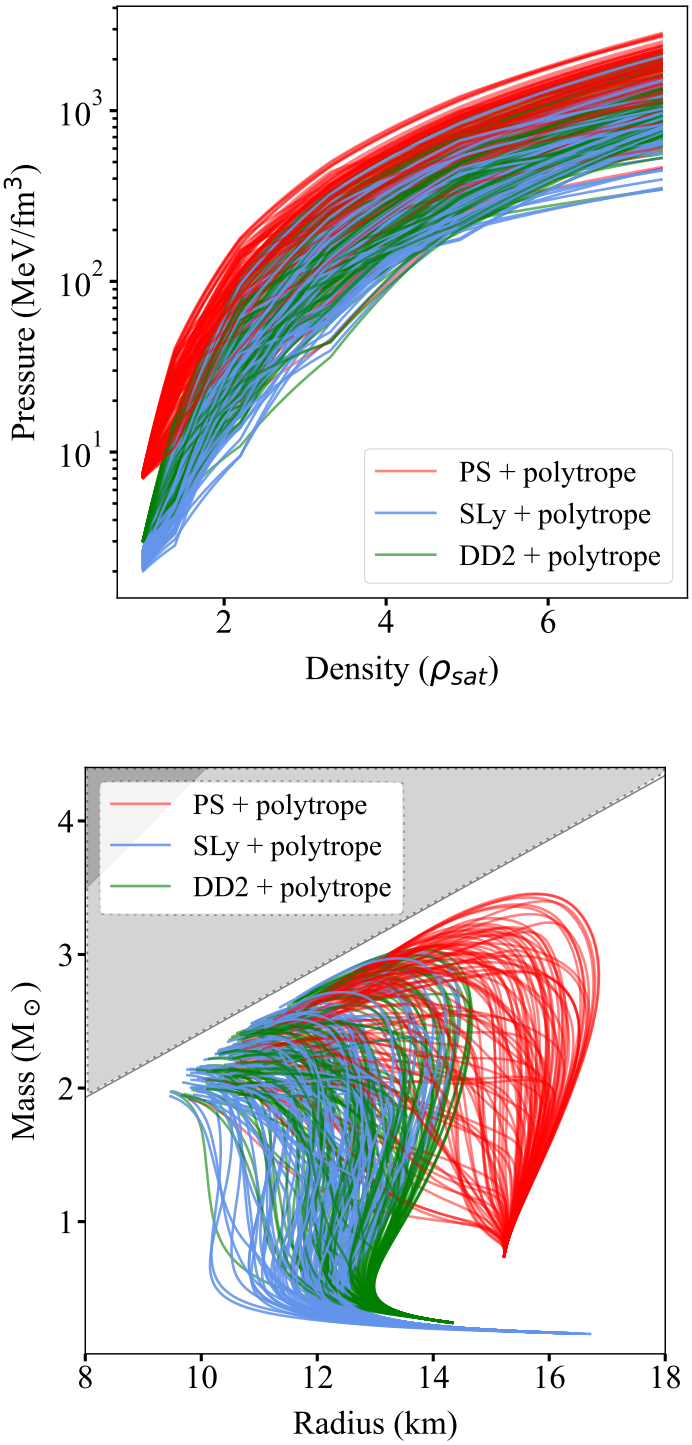

Figure 4.2: A few hundred piecewise polytropic EoSs (top) and their corresponding $M-R$ curves (bottom) for training/testing the `TOV-Solver Network`. At sub-saturation densities, the EoSs are fixed as either PS (red), SLy (blue) or DD2 (green). The boundary to the grey region on the $M-R$ plot depicts the causalty limit. See Ref. [2] for original figure.

left with the remaining EoSs. This includes 94,462 EoSs corresponding to PS; 61,273 EoSs corresponding to SLy; and 72,834 EoSs corresponding to DD2; implying a total of 228,569 EoSs for training and validating the `TOV-Solver Network`. A few hundred of the EoSs and their corresponding $M-R$ curves are depicted in figure 4.2 in the top and bottom panels, respectively. The boundary to the grey region on the $M-R$ plot depicts the causalty limit.

4.2.2 TOV-Solver Network: Architecture

The input to the `TOV-Solver Network`, i.e., the EoS, is represented in a discretized format, denoted as $P_i \equiv P(\rho_i)$. The contribution to the EoSs from the low density region ($\rho < \rho_{sat}$) are omitted in the network as these values have been established in the initial `TOV-Solver Network` to follow one of the three conventional nuclear EoSs: PS, SLy or DD2. Thus, we designate the input layer as an array of pressure values, $P_i \equiv P(\rho_{sat} \leq \rho_i \leq 7.4\rho_{sat})$, i.e., a single channel with shape $(N_\rho, 1)$. Here, N_ρ is the number of discrete density values. The output however has a shape $(N_\rho, 2)$, i.e., it has two channels for the mass M, and radius, R, respectively. Two different resolutions for the EoS representation are used here, namely $N_\rho = 128$ or 32. It was realized in subsequent assessments that a coarse resolution suffices to accomplish the desired results (see figure 4.4). The model uses logarithmic values of pressure as the input. Both the input and the output arrays are normalized as,

$$y_{\text{norm}} = (y - y_{\text{min}})/(y_{\text{max}} - y_{\text{min}}), \tag{4.3}$$

where $y = \{\log(P), (M, R)\}$. Therefore, the normalized arrays constitute elements that lie within the range (0, 1).

We construct several neural network architectures for the `TOV-Solver Network`. More specifically, we employ the Fully Connected neural network (FCN), Convolutional neural network (CNN), Long Short Term Memory (LSTM) and WaveNet. A brief description of the different neural network structures is given below. See section 3.1 of chapter 3 for a description of the CNN, FCN and LSTM models. WaveNet is a generative neural network model that is a modification of the CNN, with autoregressive properties, which mimics the concept of autoregression utilized in solving the TOV equations (refer to section 2.2.1). This model is therefore chosen here to resemble the conventional numerical methods.

WaveNet for TOV-Solver Emulator

WaveNet was first introduced in the year 2016, for generating raw audio waveforms [246]. It is based on a specific design of the convolutional neural network (see section 3.1.2 of chapter 3 for an introduction to CNNs). It is a deep neural network with autoregressive properties, meaning that the predictive distribution for each audio sample is conditioned on all previous ones.
The model operates directly on the raw audio waveform. The joint probability of a wave-

form $\mathbf{x} = \{x_1, ..., x_T\}$ can be factorised as a product of conditional probabilities:

$$p(\mathbf{x}) = \prod_{t=1}^{T} p(x_t \mid x_1, ..., x_{t-1}). \tag{4.4}$$

Therefore, each audio sample x_t is conditioned on the samples at all previous timesteps.

The autoregressive property of WaveNet comes in handy for the `TOV-Solver Network`. This is because the same concept of autoregression is exercised in solving the TOV equations. Therefore, we model and test a neural network, based on WaveNet, to map an EoS to its corresponding $M - R$ curve. The concept of autoregression in WaveNet is reproduced by using causal, dilated convolutional layers [246, 247] in the network. A dilated convolution effectively allows the network to operate on a coarser scale than with a normal convolution. This is similar to pooling or strided convolutions, but here the output has the same size as the input. The causal convolutions ensure that the model does not violate the ordering of the input data, which in this case is the EoS, or $x = P(\rho)$. In other words, the prediction output by the model at timestep t, i.e. $p(x_{t+1} \mid x_1, ..., x_t)$, depends only on the previous ones and cannot depend on any of the future timesteps $x_{t+1}, x_{t+2}, ..., x_T$. The same goes for the $M - R$ pair, (M_i, R_i), predicted at central density ρ_i : the prediction of the $M - R$ pair with central density ρ_i, depends on the EoS until ρ_i and not at higher densities ($\rho > \rho_i$). Additionally, in order to ensure information from the lowest densities is efficiently captured to make an $M - R$ prediction, we use dilated convolutions. A dilated convolution, also known as atrous convolution, capacitates the `TOV-Solver Network` to have a larger receptive field without increasing the number of parameters at the same rate [248]. It is a convolution which makes use of a kernel that is inflated by inserting holes between its consecutive elements. Therefore, the kernel is applied over a larger length (area/volume if the convolutions are 2D/3D respectively) by skipping input values with a step (this step corresponds to the dilation factor, described below). This is an efficient way of replacing a convolution that uses a larger kernel (derived from the original one by dilating it with zeros) with a convolution that uses inflated kernels. The dilation factor tells us how much a kernel is inflated, or how many step to skip in the input values, while applying the convolution operator. If $F(\mathbf{s})$ is the input, and $k(\mathbf{t})$ is the kernel, we can define the convolution operator as,

$$(F * k)(\mathbf{p}) = \sum_{\mathbf{s}+\mathbf{t}=\mathbf{p}} F(\mathbf{s})k(\mathbf{t}). \tag{4.5}$$

This formula can be generalized as,

$$(F *_l k)(\mathbf{p}) = \sum_{\mathbf{s}+l\mathbf{t}=\mathbf{p}} F(\mathbf{s})k(\mathbf{t}). \tag{4.6}$$

On keeping the value of $l = 2$, we skip 1 pixel ($l - 1$ pixels) while mapping the filter onto the input, thus covering more information in each step.

Therefore, a 1-dilated convolution is the standard convolution we described in chapter 3. By applying a series of dilated convolutional layers (with different dilation factors as shown in table 4.1), along with a *'causal'* padding, we ensure the autoregressive behaviour of the network, without forfeiting any information. The network architecture is presented in Table 4.1, for both $N_\rho = 128$ and 32.

The kernel parameters are initialized from the Xavier or Glorot uniform distribution [209], and the L_2 regularization ($\lambda = 10^{-7}$) is applied. The ELU activation function is applied on the all the layers but the last (Sigmoid activation).

Table 4.1: The WaveNet model architecture used for the `TOV-Solver Network`. Table taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

Layer Index	Layer	Dilation	Dimension	
			$N_\rho = 128$	$N_\rho = 32$
Input	-	-	(128,1)	(32,1)
1	Convolution 1D	-	(128,128)	(32,32)
2	Convolution 1D	1	(128,128)	(32,32)
3	Convolution 1D	2	(128,128)	(32,32)
4	Convolution 1D	4	(128,128)	(32,32)
5	Convolution 1D	8	(128,128)	(32,32)
6	Convolution 1D	16	(128,128)	(32,32)
7	Convolution 1D	32	(128,128)	(32,32)
8	Convolution 1D	16	(128,128)	(32,32)
9	Convolution 1D	32	(128,128)	(32,32)
10 (Output)	Convolution 1D	64	(128,2)	(32,2)

Alternatives for the TOV-Solver Emulator

In this section, we describe three other neural networks trained to map the EoS to its corresponding $M-R$ curve. We describe their network architectures in the following paragraphs. Note that for each DL model, we design architectures with both $N_\rho = 128$ and 32.

Fully connected Neural Network for the TOV-Solver Emulator

The fully connected neural network (FCN) was introduced in section 3.1.1 of chapter 3. Here, we test the use of a FCN or the basic dense neural network for the `TOV-Solver Network`. The network models pertaining to the task were designed to have the architectures as detailed in Table 4.2 for both $N_\rho = 128$ and 32. In both cases, the ELU activation

function was applied on the hidden layers, and Sigmoid on the output layer. The model weights were initialized with the *He* normal distribution [210]. An L_2 regularization penalty ($\lambda = 10^{-7}$) was applied on the layer weights. The FCN models were trained for 15,000 epochs, given their significantly superior training times when compared to the rest of the DL models. Further details of performance are given in Table 4.5.

Table 4.2: The FCN model architecture used for the `TOV-Solver Network`. Table taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

Layer Index	Layer	Dimension	
		$N_\rho = 128$	$N_\rho = 32$
Input	-	128	32
1	Dense	128	32
2	Dense	64	64
3	Dense	32	128
4	Dense	64	64
-	Add (2,4)	64	64
5	Dense	128	32
-	Add (1,5)	128	32
6	Dense	256	64
Output	Reshape	(128,2)	(32,2)

Convolutional Neural Network for the TOV-Solver Emulator

Convolutional Neural Networks (CNNs) are widely exploited for their potential to extract key features or patterns in the input data (see section 3.1.2 of chapter 3 for an introduction to the functionality of CNNs). This type of a DL model is explored for the `TOV-Solver Network` to ensure that the network is not rendered obsolete when encountered with stark changes in the pressure gradient. The number of layers in the two CNN models that we built and their dimensions are listed in Table 4.3. A kernel size of 3×1 is used for both the models, with the weights initialized from the *He* normal distribution [210], and penalized with the L_2 regularizer ($\lambda = 10^{-6}$). Stride values of 1 and 2 are applied alternatively on the convolutional layers. We employ the ELU activation function on the hidden layers and Sigmoid on the output layer. The padding *'same'* is applied to the convolutional layers to disallow loss of information at the boundaries or a change in dimension of the following layer. The performance of CNNs for the `TOV-Solver Network` in comparison to other networks can be found in table 4.5.

Table 4.3: The CNN model architecture used for the `TOV-Solver Network`. Table taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

Layer Index	Layer	Dimension	
		$N_\rho = 128$	$N_\rho = 32$
Input	-	(128,1)	(32,1)
1	Convolution 1D	(128,128)	(32,32)
2	Convolution 1D	(64,64)	(16,64)
3	Convolution 1D	(64,64)	(16,64)
-	Add (2,3)	(64,64)	(16,64)
4	Convolution 1D	(32,64)	(8,64)
5	Convolution 1D	(32,64)	(8,64)
-	Add (4,5)	(32,64)	(8,64)
6	Convolution 1D	(16,32)	(4,32)
7	Convolution 1D	(16,32)	(4,32)
-	Add (6,7)	(16,32)	(4,32)
-	Reshape	512	128
8	Dense	128	32
9	Dense	256	64
Output	Reshape	(128,2)	(32,2)

Long Short Term Memory Network for the TOV-Solver Emulator

The Long Short Term Memory (LSTM) network is variant of the recurrent neural network (RNN) as introduced in section 3.1.3 of chapter 3. Due to the fact that the output layer is an $M - R$ sequence, the prospects of LSTM for the `TOV-Solver Network` are examined. The recurring units or memory cells in LSTM carry the dependency across time sequences, leading to longer training times (see Table 4.5 for a comparison to other models). The model descriptions are given in Table 4.4 for $N_\rho = 128$, 32. The kernel parameters of all the layers were initialized from the Xavier or Glorot uniform distribution [209], and the L_2 regularization ($\lambda = 10^{-7}$) was applied. In this case, we used the Tanh activation function for the LSTM layers. We provide the details on the network performance in table 4.5. Note that the use of recurrent neural networks for sequence based data outputs the minimum error.

Table 4.4: The LSTM model architecture used for the `TOV-Solver Network`. Table taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

Layer Index	Layer	Dimension	
		$N_\rho = 128$	$N_\rho = 32$
Input	-	(128,1)	(32,1)
1	LSTM	(128,128)	(32,32)
2	LSTM	(128,64)	(32,64)
3	LSTM	(128,128)	(32,64)
4	LSTM	(128,128)	(32,32)
-	Add (1,4)	(128,128)	(32,32)
5 (Output)	LSTM	(128,2)	(32,2)

4.2.3 TOV-Solver Network: Performance

The generated EoSs and corresponding $M - R$ sequences from the previous section 4.2.1 are used to train and test the `TOV-Solver Network`. We use 52,000 samples from each low-density EoS, for training, i.e., 156,000 training samples in total. The remaining 72,569 EoSs are used for testing. This implies a 68%-32% segregation of the training-testing data. All four neural network models are tested on two different resolutions, and their performances are tabulated in table 4.5.

Table 4.5: Comparison of the performance of different Neural Networks for solving the TOV equations. Table taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

N_ρ	NN	$\mathcal{R}^2$	MSE ($\times 10^{-5}$)	Parameters (#)	Epochs ($\times 10^3$)	Time ($\times 10^3 sec$)
128	CNN	0.9999	1.743	170,176	3.5	7.35
	FCN	0.9999	1.052	70,304	15	4.91
	LSTM	0.9998	0.741	347,416	3	32.5
	WaveNet	0.9998	3.003	296,706	3	64.7
32	CNN	0.9999	3.019	58,912	3.5	2.15
	FCN	0.9999	1.179	23,936	15	2.79
	LSTM	0.9999	0.814	74,904	3	4.03
	WaveNet	0.9999	3.047	18,882	3	10.7

Table 4.5 shows that all the DNNs are capable of finding the TOV solutions of the testing

input EoS samples at precisions that reach $\mathcal{R}^2$=99.9%. The LSTM and WaveNet models are typically used for training series or sequences. Therefore, we observe a boost in their performances when combined with the `EoS-Network` (described in section 4.3 below). The network predictions for $N_\rho = 32$ and the true TOV solutions to a few testing EoS samples are presented in figure 4.3.

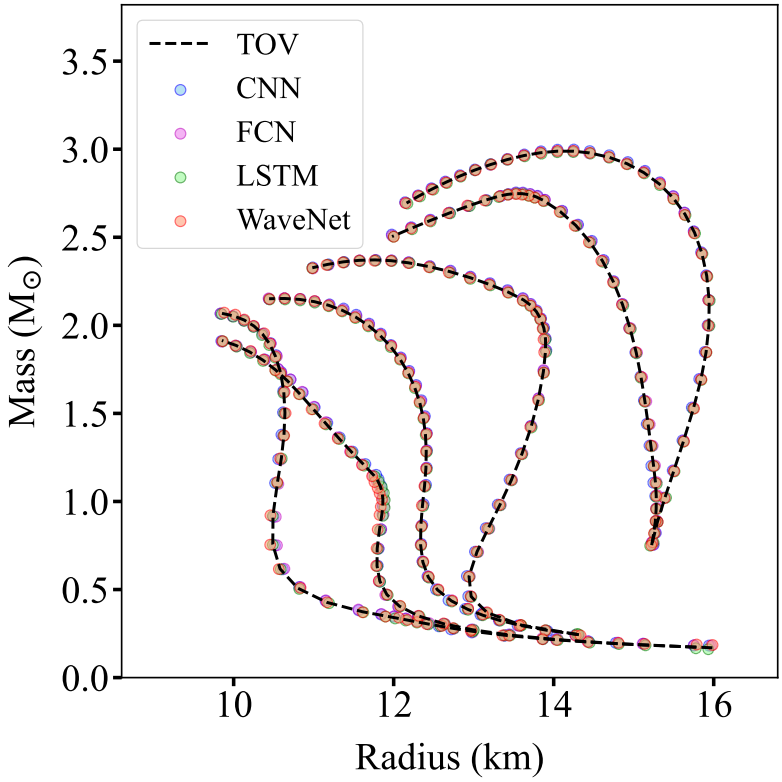

Figure 4.3: Predictions of the four different neural network models for the `TOV-Solver Network` with resolution $N_\rho = 32$, represented by different colored markers. The true TOV solutions to the EoSs are shown as dashed black lines and labelled 'TOV'.

Figure 4.3 depicts the performance of all the four neural networks employed for mapping an EoS to its corresponding $M-R$ curve. The results from the networks overlap. For EoSs that incorporate stark gradients in pressure, we observe that the predicted $M-R$ curves from the different DL models cannot all capture the kinks resulting from the pressure gradient. For these kinds of EoSs, we also observe a better performance from WaveNet compared to LSTM.

Furthermore, we depict the performance of one of the four NN models, in particular, the LSTM model for $N_\rho = 128$ and 32 in figure 4.4. We observe that the performance of the `TOV-Solver Network` on $M-R$ curves with kinks reduces very slightly as N_ρ is reduced

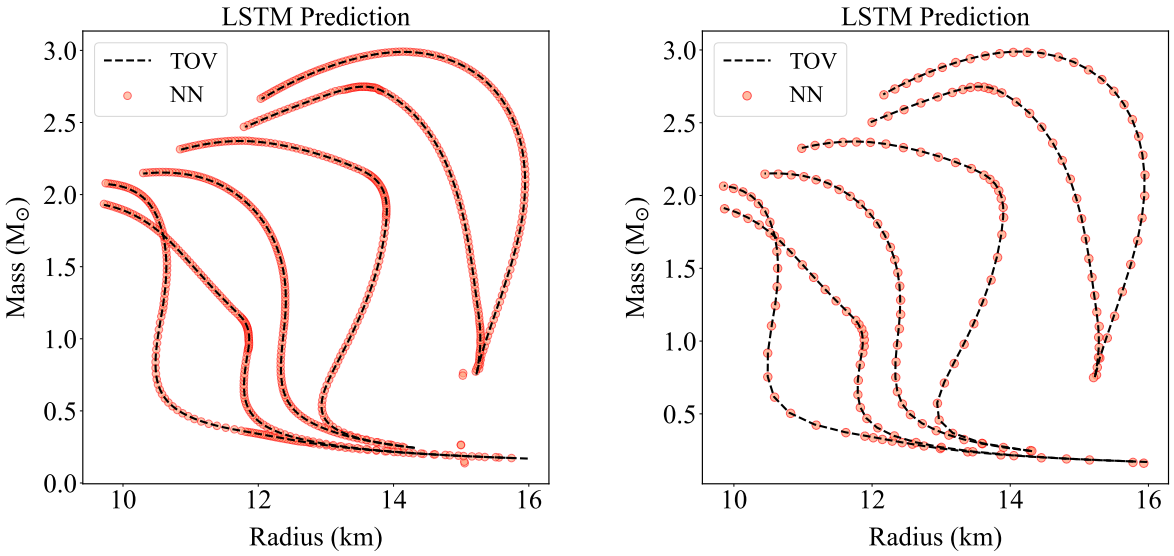

Figure 4.4: Comparison of the LSTMs' predictions for different input-output resolutions, $N_\rho = 128$ and 32, in the left and right panels, respectively.

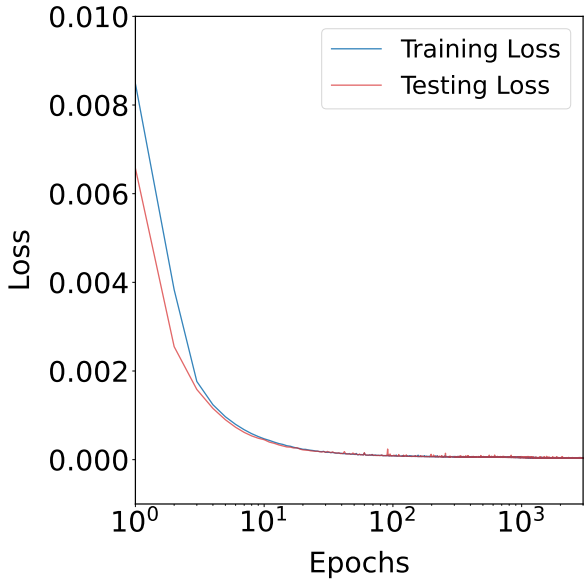

Figure 4.5: Learning curves of the WaveNet model show no signs of overfitting.

from 128 to 32. For this reason, we continue to work with the lower resolution, i.e., $N_\rho = 32$.

Based on the results of the NN models for the `TOV-Solver Network`, we employ the

WaveNet deep learning model as an emulator for this task. The WaveNet model is trained for 3000 epochs (the learning curves are shown in figure 4.5, and prove that the model does not overfit to the training data). We present the results from the trained network on a few testing samples in figure 4.6.

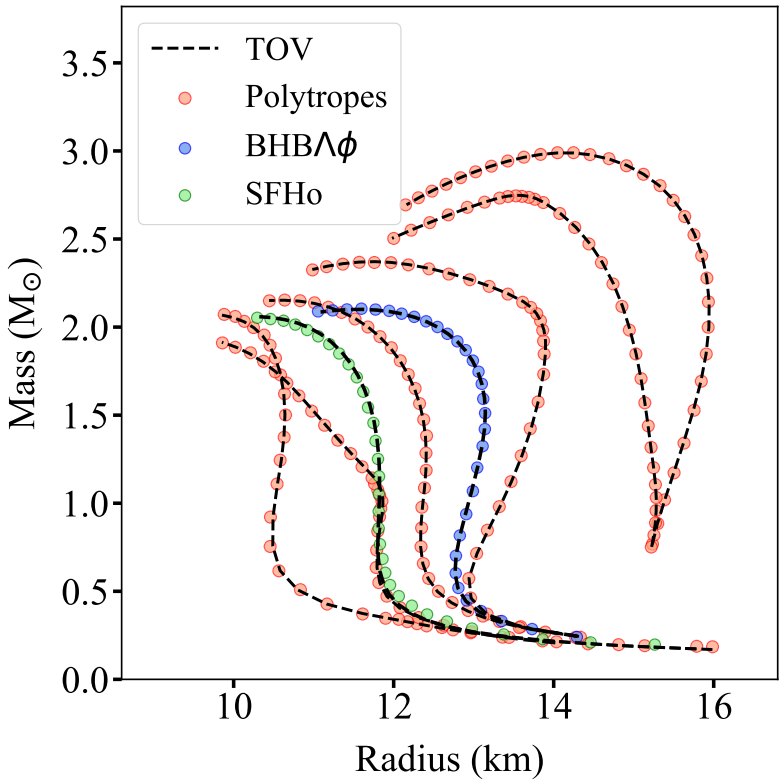

Figure 4.6: The `TOV-Solver Network` predictions using the WaveNet model for a few unseen test EoSs are shown against the ground truth values of the $M-R$ curves. The dashed black line is output from TOV equations and the markers denote the network predictions. The blue and green dotted $M-R$ curves represent the network predictions for the BHB$\Lambda\phi$ and SFHo EoSs respectively. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

We observe from figure 4.6 that the trained `TOV-Solver Network` is successful in capturing the mapping of randomly given EoSs to the corresponding $M-R$ curves. It can also be seen that the trained WaveNet also predicts the $M-R$ curves of the SFHo and BHB$\Lambda\phi$ EoSs up to a high precision as shown in Figure 4.6. The $M-R$ curves of all the test EoSs are successfully reproduced by the DL models with the coefficient of determination $\mathcal{R}^2 \sim 99.9\%$. This illustrates that DNNs have the capacity to replace conventional numerical methods used for solving the TOV equations. In comparison to numerical calculations like the Euler or Runge-Kutta methods, the network emulator used here, is (i) superior in

computational efficiency, ($\sim 10^6$ *sec* faster), and (ii) easily differentiable. This is critical for applying back-propagation in the AD framework, deployed in the following section for statistical inference of the EoS reconstruction. The trained `TOV-Solver Network` model and weights are frozen and saved for the forthcoming procedures.

4.3 EoS Network

Once the `TOV-Solver Network` is trained, the `EoS Network` is introduced into the pipeline (see Figure 4.1(b)). The `EoS Network` takes density, ρ, as input and produces the corresponding pressure, $P_\theta(\rho)$, as output, effectively modeling the equation of state. The EoS output ($P_\theta(\rho)$) from the `EoS Network` is further input to the well-trained `TOV-Solver Network`, thus connecting the two networks. This integration allows the `TOV-Solver Network` to incorporate the learned EoS at every iteration step and employ the saved weights to output the mass-radius sequence. The `TOV-Solver Network` parameters are set as non-trainable weights prior to the EoS optimization process. The trainable weights of the `EoS Network` are then optimized in this pipeline to fit the predicted $M - R$ output to mock observational data, thereby inverting the TOV equations. Compared to other studies which captured the inverse mapping from $M-R$ observations to the EoS [159–161, 237–239, 249–251], the proposed method belongs to the unsupervised learning paradigm. The unsupervised learning algorithm developed in this thesis can be cast as a generalized Bayesian inference, with augmentations in the following aspects: (i) the EoS is represented in an unbiased manner as a DNN, thus the parameters to optimize are the network weights and biases; (ii) the traditional numerical methods to solve the TOV equations are replaced by a well-trained `TOV-Solver network`, thus simplifying and speeding up the following AD process; and (iii) the optimization uses a gradient-descent based approach within the AD framework as depicted in Figure 4.1(b). The network architecture and the optimization procedure are described in the following section.

4.3.1 EoS Network: Architecture and Optimization

The input density to our pipeline consists of a 1D array of length N_ρ= 32, evenly spaced on the linear scale and normalized to range between $(0, 0.1)$. As mentioned earlier, the trained WaveNet model of the same resolution is deployed as the `TOV-Solver Network`. The architecture of the `EOS Network` includes three 1D convolutional layers: two hidden layers and one output layer, as illustrated in Figure 4.1(d). Each of the hidden layers consists of 64 feature maps. We employ a 1×1 kernel size, with the kernel weights initialized from a *He* normal distribution [210]. Regularization with L_2 penalty ($\lambda = 10^{-8}$) is exercised on the weights, and the *same* padding is applied to all layers. The output layer employs the Sigmoid activation function, while the hidden layers utilize the ELU activation function. By using kernel size 1×1, we ensure that each input element shares the same parameters in the above architecture. Hence, the induced relationship between the i^{th} input density neuron, ρ_i, and corresponding output pressure, P_i, follows a structure similar to a fully connected

neural network with two hidden layers, each comprising 64 neurons. Consequently, the `EoS Network` is characterized by a total of 4353 parameters, defining $P_\theta(\rho)$. We further specify the weights of each layer in the `EoS Network` to be non-negative. This preserves the order of the input layer sequence up to the output layer, $P(\rho)$. As a consequence, we establish monotonicity in the represented function $P(\rho)$, a condition that is required by any physical EoS. Additionally, this approach guarantees a well-correlated function for the reconstructed EoS. Contrary to the common practice of utilizing large datasets for training neural network, the approach designed here uses an optimization process that necessitates only one mass-radius ($M-R$) sequence. Hence, each training epoch operates with a batch size of 1, i.e, in the ideal situation, we are required to only optimize the unique dense matter EoS to fit $M-R$ observations. However, we take into account the realistic scenario where all observational data comes with sizeable uncertainties and adapt the developed scheme to meet the associated requirements (see section 4.3.2 for further details).

If we denote $N_{\rm obs}$ as the number of reliable $M-R$ observations, the loss function for the training the `EoS Network` is given by Eq. (4.1). It is defined as the distance between the observations and the $M-R$ curve as predicted from the network pipeline above, i.e. the likelihood of observations given an EoS and its corresponding $M-R$ curve from the `TOV-Solver Network`. Nevertheless, the observational data is unevenly distributed in the $M-R$ space, potentially causing discontinuities in the $M-R$ curve due to measurement uncertainties. Furthermore, the data is limited and the central density, (ρ_{ci}), corresponding to an uncertain observation is unknown. In other words, the $i^{\rm th}$ $M-R$ observation, (M_i, R_i), does not necessarily correspond to the $i^{\rm th}$ central density, ρ_{ci}, in the input layer [1]. In order to optimize the loss function effectively, we employ the 'closest approach' method as implemented in Ref. [235]. Thus, during each iteration of the training process, we evaluate the loss by,

$$\chi^2 = \sum_{i=1}^{N_{\rm obs}} \frac{(M(\rho_{ci}) - M_{{\rm obs},i})^2}{\Delta M_i^2} + \frac{(R(\rho_{ci}) - R_{{\rm obs},i})^2}{\Delta R_i^2}\,. \tag{4.7}$$

Here, ρ_{ci} for each $i^{\rm th}$ observation is updated according to the formula,

$$\rho_{ci} = \arg\min_{\rho_c} \frac{(M(\rho_c) - M_{{\rm obs},i})^2}{\Delta M_i^2} + \frac{(R(\rho_c) - R_{{\rm obs},i})^2}{\Delta R_i^2}\,. \tag{4.8}$$

Hence, Eq. (4.8) is employed to calculate the central densities of $(M_{\rm obs}, R_{\rm obs})$ that result in the minimal distance between the $M-R$ curve generated by the `TOV-Solver Network` and the actual $M-R$ observations. The Adam optimizer [212] was utilized here, with varying learning rates (α's) or different stages of the training process. Usually, smaller learning rates are used at later stages to stabilize the training. The network's learning process is regulated by reducing the loss function χ^2. The learning rate for one such reconstruction in our study, was scheduled as follows: 1000 epochs ($\alpha = 0.001$), 1000 epochs ($\alpha = 0.005$), 1500

[1] In such circumstances, a finer resolution ($N_\rho = 128$) might prove useful.

epochs ($\alpha = 0.003$), 1600 epochs ($\alpha = 0.0001$) and, finally, 2000 epochs ($\alpha = 3 \times 10^{-5}$), in the order specified. The choice of applying small changes in the learning rates is attributed to the minimal impact they have on the final results. Through this approach, the EoS is reconstructed with an uncertainty which is proportional to the statistical uncertainty of the observations. In certain situations where the causal condition is not fulfilled, the DNN fails to produce a realistic EoS. This lack of convergence is evident from the non-decreasing loss during the optimization process. To address this problem, we enforce the causal condition to reject EoS solutions that do not fulfill the physical constraints. The next section describes the reconstruction of the EoS using mock $M - R$ data.

4.3.2 Tests on Mock Mass-Radius data

Ideal Scenario

The reconstruction method presented in this work is validated by closure tests that are first performed in an ideal case of mock data, i.e, synthetically generated $M - R$ mock data without systematic and statistical uncertainties. We use two example sets of such mock data from $M - R$ curves which correspond to two randomly chosen EoSs (labelled Mock Test I and II) as depicted in Figures 4.7 and 4.8. The black solid lines in both the figures are the true curves used to test the reconstruction method. We are compelled to use mock NS data with masses above $1M_{\odot}$ due to the absence of naturally occurring low-mass ($< 1M_{\odot}$) neutron stars. Figure 4.7 depicts two example sets of mock observations each containing 11 $M - R$ points, which are chosen from the region $M > 1M_{\odot}$ along the $M - R$ curve. These 11 highlighted black points lie on the $M - R$ curve (ground truth). Note that the mock observations are spread quite uniformly on the mass scale in the left panel of figure 4.7. However, this is a very unlikely situation in real observational data. Therefore, we ensure that the mock data is randomly spread on the mass scale in the right panel of figure 4.7. With the proposed method, we reconstruct the EoSs from these two example sets of mock $M - R$ data. The EoSs reconstructed by the DNN in this work are plotted as dashed red lines in the left and right panels of Figure 4.8. These reconstructed curves are compared to the ground truth EoSs, which are displayed as solid black lines. Evidently, for ideal scenarios, the reconstruction of the EoSs is fairly successful in the range $M > 1M_{\odot}$. The EoSs predicted by the DNN are remarkably close to the ground truth EoSs in the high-density region (i.e., corresponding to the high mass region in $M - R$ curve). In the left panel of figure 4.8, the deviation of the reconstructed EoS from the ground truth in the low-density regime is attributed to the lack of mock data below $1M_{\odot}$. This is also demonstrated in figures 4.8 and 4.7, where the 11 black solid points on the EoSs correspond to the 11 mock data points on the $M - R$ curves. The red dashed curves in figure 4.7 are the $M - R$ curves corresponding to the reconstructed EoSs as predicted by the `TOV-Solver Network`. The two $M - R$ curves agree with each other reasonably well in the region where the mock data is sufficiently available. For further validation, one can produce the $M - R$ curve of the reconstructed EoS by directly solving the TOV equations and then compare the obtained results.

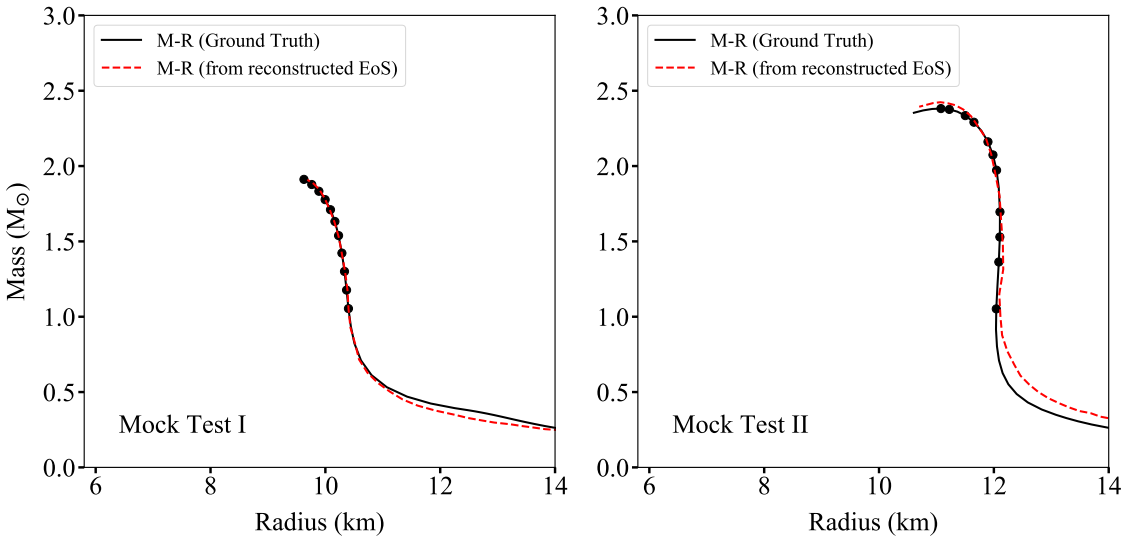

Figure 4.7: Examples of ideal mock $M - R$ data used to test the reconstruction of the corresponding EoSs with the method proposed in this work. The 11 $M-R$ mock data points (black markers) are located in the region $M > 1M_\odot$ along the ground truth $M - R$ curve (black solid line). A reasonable agreement of the $M-R$ curves from the reconstructed EoSs (red dashed line) with the ground truth curves is observed in the mass region $M > 1M_\odot$. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

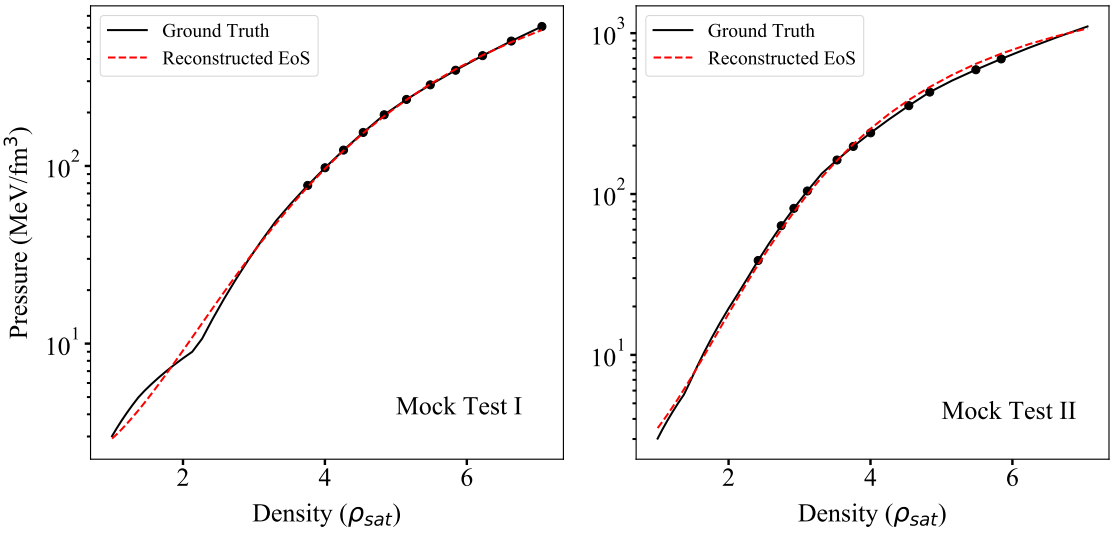

Figure 4.8: Comparison of the EoSs reconstructed by the described method (dashed red lines) and the ground truth EoSs (solid black lines) using ideal mock $M - R$ data. The black points shown are central densities of the NSs corresponding to the mock $M-R$ data in figure 4.7. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

Realistic Scenario

We now proceed to the practical situation where observations inevitably involve uncertainties. In the next step, the proposed method is evaluated on the same mock $M-R$ data as in figure 4.7, albeit with uncertainties included. We incorporate statistical noise, for both the mock tests, by sampling multiple $M-R$ curves from a normal distribution centered around the true $M-R$ values, $\mathcal{N}(M_i; R_i; \sigma_{M_i}; \sigma_{R_i})$. A relative standard deviation of 10% is applied to both mass and radius, represented as $\Delta M_i = 0.1 M_i$ and $\Delta R_i = 0.1 R_i$. Consequently, 500 sets of mock $M-R$ data samples are drawn from a normal distribution with $\sigma_M = 0.1M$, $\sigma_R = 0.1R$. The true $M-R$ curve is represented by the black solid line in figure 4.9, and the highlighted black circles mark the 11 $M-R$ mock data, which are the mean values of the respective normal distributions. The blue point-cloud in the figure depicts all 500 sampled sets of the mock data points resulting from the true $M-R$ curve, where each sampled set contains just 11 $M-R$ pairs, all above $\sim 1 M_\odot$, to be consistent with the current real observations. The optimization procedure is carried out for all 500 samples. The results of the optimization are shown in figure 4.10. For every sampled set of mock data, the proposed algorithm can reconstruct exactly one EoS based on the maximum a posteriori probability (MAP). During this closure test, the uncertainty of the reconstruction method is evaluated by fitting the reconstructed EoSs to a uni-variate normal distribution, denoted as $\mathcal{N}(\mu_i, \sigma_i)$. This process is straightforward and the reconstructed, model-independent EoS is shown in figure 4.10. The 2σ, 95% confidence level of the EoS is depicted as the orange shaded region. The dashed red curve in the shaded region of the figure is the mean of the reconstructed EoS. The `TOV-Solver Network` predicts the $M-R$ curves corresponding to each of the 500 reconstructed EoSs. The $M-R$ predictions of the `TOV-Solver Network` from the reconstructed EoSs are first filtered to satisfy the causal limit and then subject to a normal distribution fit. In this case, all $M-R$ points with the same central density $\rho_{c,i}$ (i.e. the ensemble of each element from the output sequence) are fitted with individual bi-variate Gaussian distributions, given by $p((M_i, R_i)|P_j) = \mathcal{N}(\mu_\mathbf{i}, \sigma_\mathbf{i})$. The 95% confidence level is determined from the $2.44\sigma_i$ interval for 2D distributions. This uncertainty band of the $M-R$ curve is illustrated by the orange band in figure 4.9. The mean of the resulting $M-R$ curve band from the bi-variate distribution is represented by the dashed red line in the figure. It is evident that the width of the $M-R$ curves obtained from the reconstructed EoSs is considerably smaller, when compared to the width spanned by the uncertainty of the mock $M-R$ data samples, represented as blue point clouds. This illustrates the potential of the novel algorithm for the EoS reconstruction from NS observables. We demonstrate its effectiveness in the presence of large uncertainties associated with the mock $M-R$ data. In a few instances of the sampled $M-R$ curves, however, the NN fails to produce a realistic EoS. The failure can be attributed to the large uncertainties of individual $M-R$ pairs, causing outliers in the $M-R$ curve. Moreover, the points on the $M-R$ curve are assumed to be ordered in the ascending order of their mean mass. However, accounting for large uncertainties while sampling from a Gaussian distribution, may result in a disordering of the $M-R$ pairs. The causal condition is employed to filter out these EoSs which fail to converge.

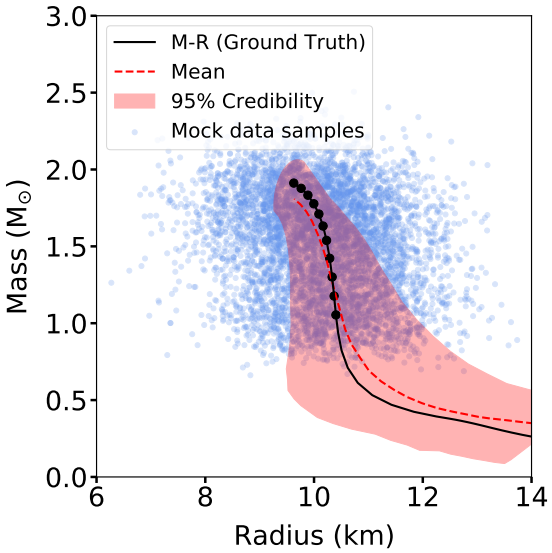

Figure 4.9: The blue point cloud represents the $M-R$ mock data ensembles, sampled with 10% uncertainties from the ground truth $M-R$ curve (black solid line). The 95% confidence level of the $M-R$ band from the reconstructed EoSs is depicted as the orange shaded region, and the mean as a red dashed line. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

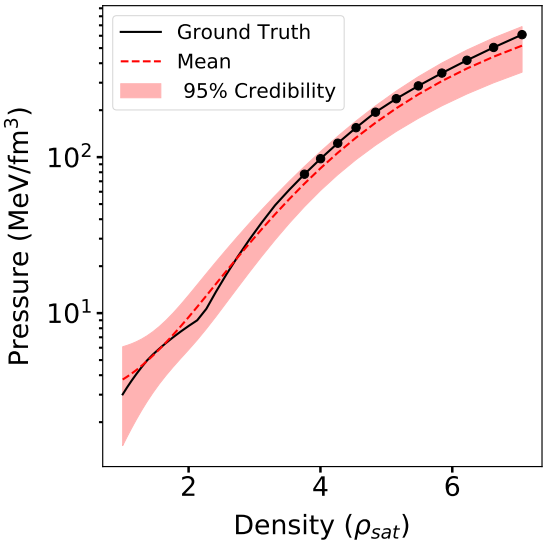

Figure 4.10: Comparison of the ground-truth (black solid line) and the reconstructed EoS (red dashed line) with uncertainty (orange band), using $M-R$ mock data from figure 4.9. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

Additional results are presented in figures 4.11 and 4.12. In these figures, we demonstrate the performance of the EoS reconstruction for both mock test I and II (the same EoSs and $M-R$ curves used in figures 4.8 and 4.7). However, different observational uncertainties are assumed here: the relative noise level of mock data are set to 10% and 5% of the mean $M-R$ mock data. For this method, the systematic error can be quantified based on both the uncertainty of the `TOV-Solver Network`, and of the `EoS Network`. The results shown in figure 4.12 demonstrate that the mean of the reconstructed EoSs is closer to the ground truth in the case with 5% error as compared to the case with 10% error. The uncertainty band narrows down with decreasing error. Hence, the availability of more precise observations in future offers significant potential for the reconstruction of a better constrained EoS using the described method. All the closure tests described above confirm that the Automatic Differentiation method developed in this work, utilizing deep neural networks, effectively reconstructs the underlying EoS for dense neutron star matter. This is achieved even with limited mock $M-R$ observations of neutron stars that contain substantial measurement errors.

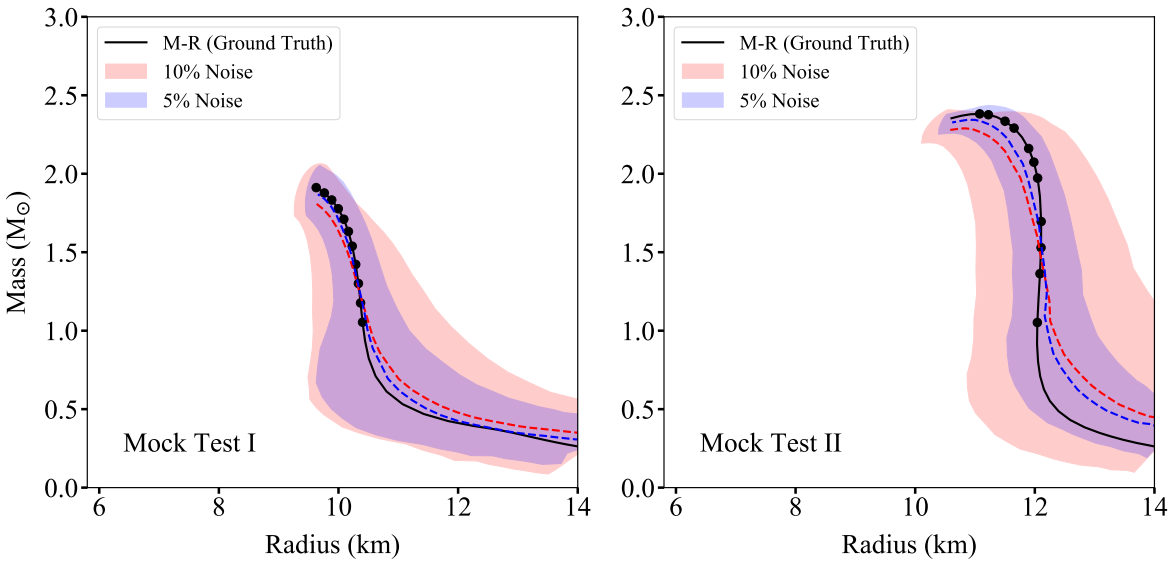

Figure 4.11: The ground truth $M-R$ curve is depicted as a black solid line and the black dotted markers represent mean of the individual $M-R$ mock data points used for reconstructing the EoS. The $M-R$ relationships from the reconstructed EoSs with means (red dashed line for 10% noise level, blue for 5%) and uncertainties (orange band for 10% noise level and blue for 5%) are also depicted. Note that the $M-R$ bands are an output from the `TOV-Solver Network`. The left and right panels represent the Mock Test I and Mock Test II, respectively. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

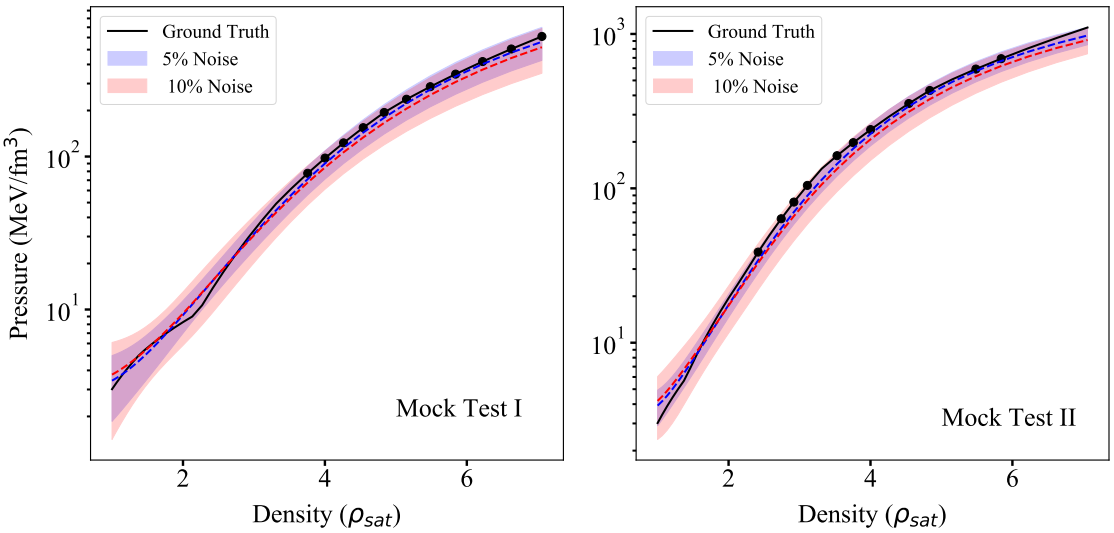

Figure 4.12: The EoSs reconstructed using different levels of uncertainties on the $M-R$ mock data, with the blue band denoting a 5% noise level and the orange band denoting a 10% noise level. The corresponding dashed lines represent the respective mean of the reconstructed EoSs. The left and right panels represent the results for the Mock Tests I and II, respectively. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

Mock Tests on SFHo and BHB$\Lambda\phi$ EoSs.

So far, we have presented results of the reconstructed EoSs that used a few $M-R$ curve examples from the testing data set (piecewise-polytropes), selected at random. In this section, we test the performance of the neural networks in the developed method on two microphysical EoSs, namely the SFHo [133] and BHB$\Lambda\phi$ [144] EoSs (see section 2.1.1 of chapter 2 for a description of these EoS models). We use random $M-R$ points on the $M-R$ curves, such that $M > 1M_{\odot}$, corresponding to the SFHo and BHB$\Lambda\phi$ EoSs. We repeat the optimization procedure to reconstruct the two EoSs without assuming any uncertainties. The reconstructed SFHo and BHB$\Lambda\phi$ EoSs obtained from the `EoS Network` are depicted as red dashed curves in the left panels of figure 4.13 and figure 4.14 respectively. The results are compared with the true EoSs which are also depicted in the figures as solid black curves. Furthermore, we depict the $M-R$ curves obtained from the reconstructed EoSs using the `TOV-Solver Network` (red dashed curves) against the true $M-R$ curves (solid black curves) in the right panels of figure 4.13 and figure 4.14, for the SFHo and BHB$\Lambda\phi$ EoSs, respectively. It can be observed from figure 4.14 that the reconstructed EoS deviates from the true EoS only at densities larger than the central density corresponding to the maximum mass neutron star produced by the EoS. The `EoS Network` outputs an EoS that reproduces the true $M-R$ curve until the maximum mass, up to a very high precision. Therefore, these figures showcase a drawback of the proposed method. The lack of information regarding whether or not the maximum mass is reached for a given $M-R$ curve affects the EoS

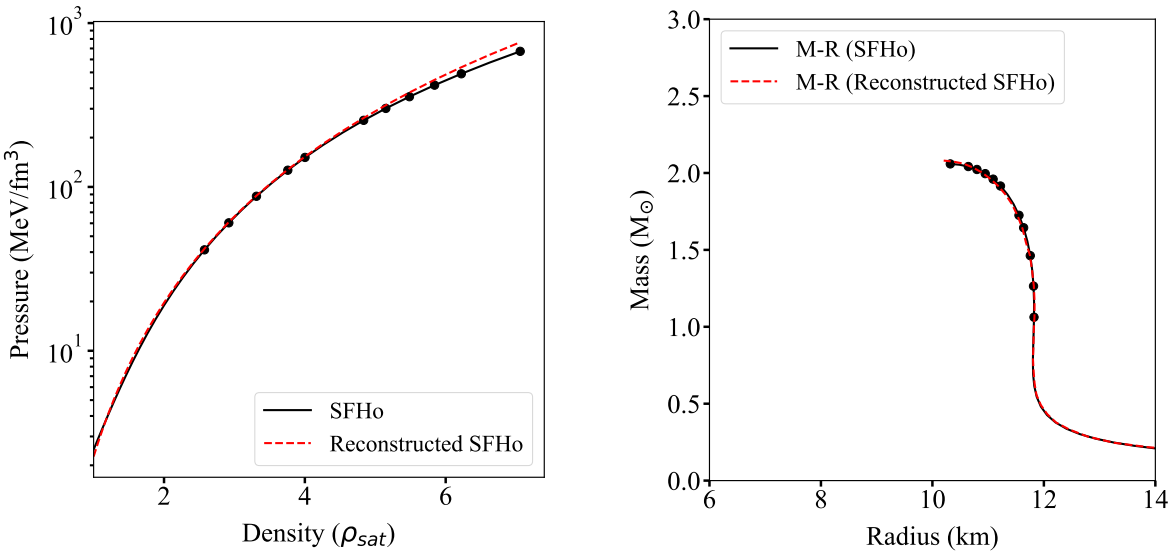

Figure 4.13: EoS reconstruction test on the SFHo EoS [133]. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

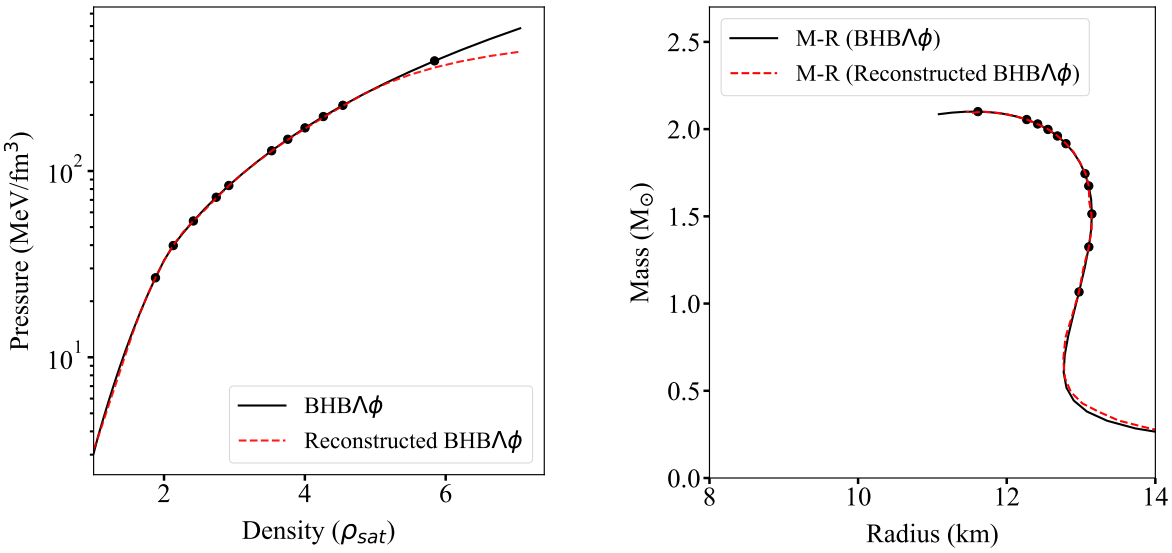

Figure 4.14: EoS reconstruction test on the BHB$\Lambda\phi$ EoS [144]. Figure taken from Ref. [2] (DOI:10.1088/1475-7516/2022/08/071).

reconstruction. This remains unknown for real observations. Nevertheless, several attempts have been made to predict the maximum mass of NSs [252–255].

4.3.3 Application on Real Mass-Radius Observations

In this section, we apply the physics-based DL algorithm to reconstruct the dense matter EoS using real $M-R$ observations of NSs. For this, we re-train the `TOV-Solver Network` with additional data. It can be seen from the lower panel of figure 4.2 that for NSs with $M \approxeq 1M_{\odot}$, there is insufficient data in the region where radius $R \approxeq 13-15$km. Therefore, we prepare additional training data that is based on the TM1 EoS [128] for sub-saturation densities (more specifically, we prepare 100,000 piece-wise polytropic EoSs and apply the $2M_{\odot}$ constraint to obtain 90,234 additional EoSs). See figure 2.2 and lower panel of figure 2.3 for a few examples. Hence, we now have four different low density ($\rho < \rho_0$) EoSs included in the data generation for re-training the `TOV-Solver Network`: PS, SLy, DD2 and TM1. The `TOV-Solver Network` is updated in the pipeline (see figure 4.1), with this modification.

We now present the list of current $M-R$ observations [75, 82, 86, 87, 89, 90] used in this thesis in table 4.6. The first six observations in the table come from NSs in low mass X-ray binaries in quiescence [75] (see section 1.1.2 for an introduction). The next six observations in the list are observations from thermonuclear bursters [75] (see section 1.1.2). X5 and X7 are quiescent neutron star low-mass X-ray binaries in the globular cluster 47 Tuc [82]. By fitting the atmosphere models of X-ray bursting NSs directly to the observed spectra, mass and radius measurements of the pulsar 4U 1702-429 were obtained [256]. A radius measurement of PSR J0437–4715 is obtained from its ultraviolet and soft X–ray thermal emissions [257]. The last two pulsars in the table (PSR J0030+0451 and PSR J0740+6620) are observations from NICER [86, 87, 89, 90], which targets rotation-powered millisecond pulsars for radii estimations (see section 1.1.2 for more information on the measurement techniques). It is important to note that the systematic uncertainties vary for the different measurements. Therefore, the joint analysis of the observations listed in the table would provide results in this work that are not completely accurate for interpretation. Furthermore, there have been revisions on the radius measurement of pulsars, PSR J0030+0451 and PSR J0740+6620, so as to have estimates from combined analyses with XMM-Newton data, as well as without [88, 91]. We use the first results on both pulsars. Due to limited number of observations by NICER, whose systematic uncertainties are believed to be better accounted for, we utilize all the observations listed in table 4.6 for reconstructing the dense matter EoS.

The masses and radii measurements of some of these pulsars are provided as 2D confidence contours [75, 82]. In such cases, we integrate the 2D distribution onto 1D for both M and R, by integrating over the probability distributions of R and M, respectively (we follow a similar procedure from Refs. [160,161]). Normal distributions are then used to fit the masses and radii individually. We repeat this for the first 14 observations in the list. Regarding the NICER observations, we take an average of the results from the two independent groups. The Gaussian fitted masses and radii observations of neutron stars based on their marginalized distributions are presented in the second and third column of table 4.6, respectively. We use this data to reconstruct the EoS and adopt Importance Sampling (a

Bayesian perspective) to the estimate the uncertainty of the NN reconstruction.

Table 4.6: The individual 1D Gaussian fitted $M - R$ observations of neutron stars based on their marginalized distributions [75, 82, 86, 87, 89, 90, 256, 257].

Pulsar	Mass ($M_\odot$)	Radius (km)
M13	1.42±0.49	11.71±2.48
M28	1.08±0.30	8.89±1.16
M30	1.44±0.48	12.04±2.30
NGC 6304	1.41±0.54	11.75±3.47
NGC 6397	1.25±0.39	11.48±1.73
ωCen	1.23±0.38	9.80±1.76
4U 1608-52	1.60±0.31	10.36±1.98
4U 1724-207	1.79±0.26	11.47±1.53
4U 1820-30	1.76±0.26	11.31±1.75
EXO 1745-248	1.59±0.24	10.40±1.56
KS 1731-260	1.59±0.37	10.44±2.17
SAX J1748.9-2021	1.70±0.30	11.25±1.78
X5	1.18±0.37	10.05±1.16
X7	1.37±0.37	10.87±1.24
4U 1702-429	1.90±0.30	12.40±0.40
PSR J0437–4715	1.44±0.07	13.60±0.85
PSR J0030+0451	1.44±0.15	13.02±1.15
PSR J0740+6620	2.08±0.07	13.70±2.05

Uncertainty Estimation: Importance Sampling

The optimization of the `EoS Network` to reconstruct the EoS corresponding to the observational data follows from section 4.3.1. In order to evaluate the uncertainty of the reconstruction, we employ the Bayesian approach, focusing on the posterior distribution of the EoSs given the astrophysical observations under consideration, denoted as Posterior($\boldsymbol{\theta}_{\text{EoS}}$|data). We first draw an ensemble of $M - R$ samples from the fitted Gaussian distributions for the real observations listed in table 4.6. From these Gaussian distributions, we deterministically infer the corresponding EoS with maximum likelihood estimation. This can be termed the 'proposal'. Given the ensemble of reconstructed EoSs, we apply importance

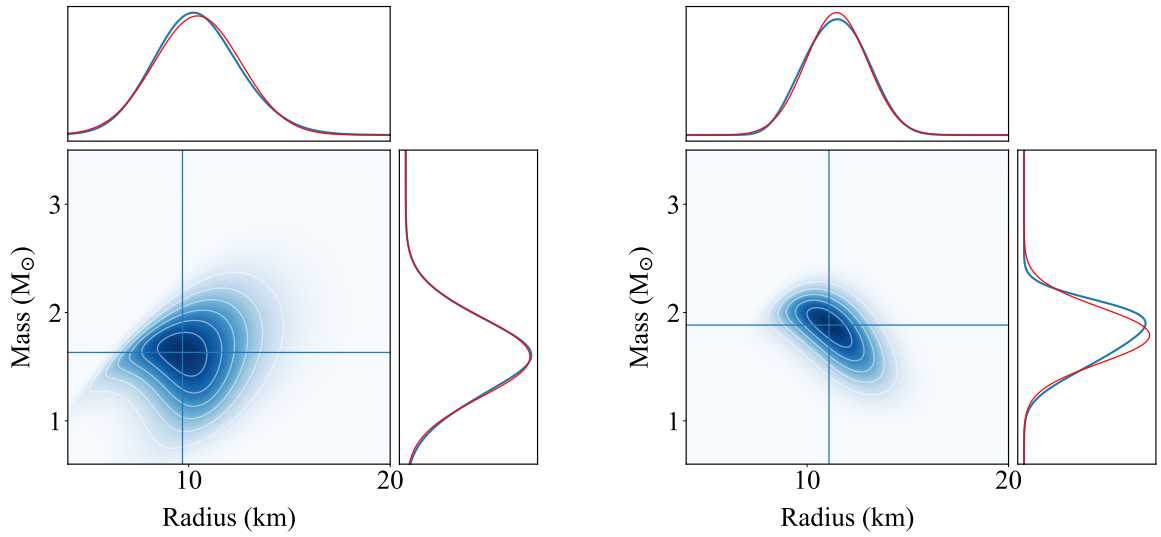

Figure 4.15: 2D contour distributions of mass and radius for pulsars KS 1731-260 (left) and 4U 1724-207 (right) depicted as blue shaded regions. The marginalized distributions of M and R are obtained by integrating over the probability distributions of R and M, respectively (see Refs. [160, 161] for similar procedure) and plotted as blue curves outside the center plots. Normal distributions (red curves) are then used to fit the marginalized distributions of masses and radii individually.

sampling (see Ref. [258] for a review on importance sampling) to estimate the uncertainty related to the desired posterior distribution, where a proper weight (w) is evaluated to each EoS.

In general, a physical variable $\hat{O}$ can be estimated as,

$$\bar{O} = \langle \hat{O} \rangle = \sum_{j}^{N_{\text{samples}}} w^{(j)} O^{(j)}, \tag{4.9}$$

and the standard deviation can also be estimated as $(\Delta O)^2 = \langle \hat{O}^2 \rangle - \bar{O}^2$. The weights are given by,

$$\begin{aligned} w^{(j)} &= \frac{\text{Posterior}(\boldsymbol{\theta}^{(j)}_{\text{EoS}}|\text{data})}{\text{Proposal}(\boldsymbol{\theta}^{(j)}_{\text{EoS}})} \\ &\propto \frac{p(\text{data}|\boldsymbol{\theta}^{(j)}_{\text{EoS}})\text{Prior}(\boldsymbol{\theta}^{(j)}_{\text{EoS}})}{p(\boldsymbol{\theta}^{(j)}_{\text{EoS}}|\text{samples}^{(j)})p(\text{samples}^{(j)}|\text{data})}, \end{aligned} \tag{4.10}$$

where, j indicates the index of reconstructed EoS samples; $\boldsymbol{\theta}_{\rm EoS}$ is the parameters set for representing the EoS; and $p(\text{samples}|\text{data}) = \mathcal{N}(M_{\rm obs}, \Delta M^2)\mathcal{N}(R_{\rm obs}, \Delta R^2)$ dictates the probability for samples we draw from the fitted Gaussian distributions for the observations in table 4.6. Furthermore,

$$p(\boldsymbol{\theta}^{(j)}_{\rm EoS}|\text{samples}^{(j)}) = 1,$$

since the reconstruction technique is capable of locating the deterministic corresponding EoS given the sampled M-R points; and

$$p(\text{data}|\boldsymbol{\theta}^{(j)}_{\rm EoS}) \propto \exp\left(-\chi^2(M_{\boldsymbol{\theta}^{(j)}_{\rm EoS}}, R_{\boldsymbol{\theta}^{(j)}_{\rm EoS}})\right),$$

is the likelihood function of EoS parameters invoked here (see Eq. (4.7) for the loss function). In practical calculations, weights should be normalized as $\tilde{w}^{(j)} = w^{(j)}/\sum_j w^{(j)}$. In addition, a cut-off is applied to avoid outliers in samples. Consequently, note that during the training process, target observables of χ^2-fitting in Eq. (4.7) are changed from $(M_{{\rm obs},i}, R_{{\rm obs},i})$ to $(\tilde{M}^{(j)}_i, \tilde{R}^{(j)}_i)$. Based on this approach, we sample 10,000 $M-R$ curves and reconstruct the EoS with uncertainties based on importance sampling. The results are presented in figure 4.16 and are labelled as 'This Work'.

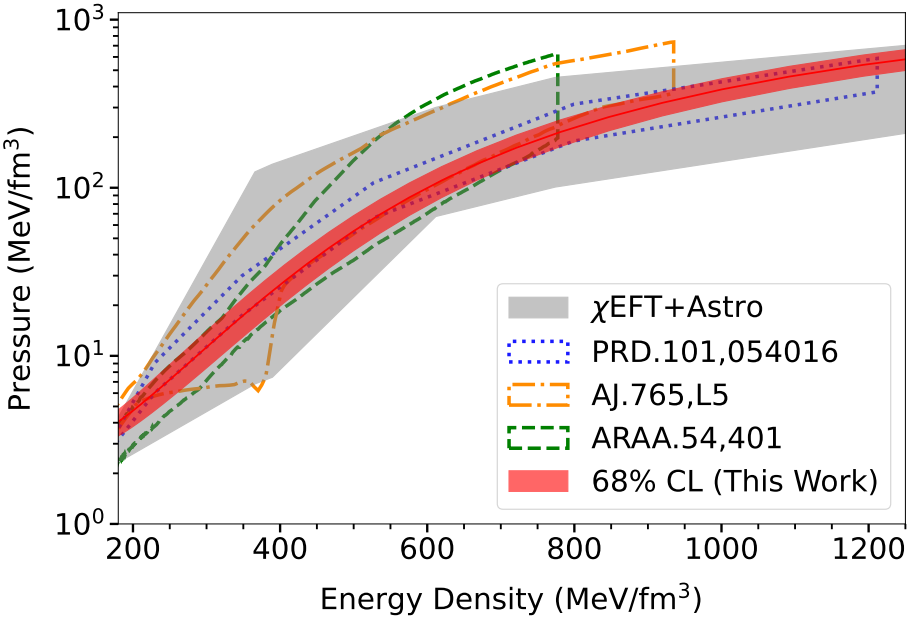

Figure 4.16: 68% confidence level (CL) of the EoS reconstructed from observational data of 18 neutron stars, depicted as a red shaded band. The gray band depicts the constraints on the EoS from χEFT [122]. Results from Bayesian methods (AJ.765,L5 [259] and ARAA.54,401 [260]) and the direct inverse mapping using NNs (PRD.101,054016 [160]) are also included here. Figure taken from Ref. [3] (DOI:10.1103/PhysRevD.107.083028).

The red band in the figure depicts the 68% confidence level (CL) of the reconstructed EoS. Constraints on the EoS from χEFT [122] are presented as a gray band. In addition, we compare our results with previous works based on Bayesian methods labelled 'AJ.765,L5' [259]

and 'ARAA.54,401' [260]. The results from the direct inverse mapping using NNs are also shown as 'PRD.101,054016' [160].

The M-R band and contours corresponding to the reconstructed EoS results from figure 4.16 are demonstrated in figure 4.17. The $M-R$ observations with their uncertainties from table 4.6 are illustrated as gray points. It can be observed that the reconstructed EoS strongly supports the existence of massive neutron stars ($>2M_{\odot}$). Moreover, the EoS yields $R_{1.4} = 11.6 \pm 0.43$,km (at 68% CL) for a canonical $1.4M_{\odot}$ neutron star. This result is in agreement with recent constraints derived from multi-messenger observations [261].

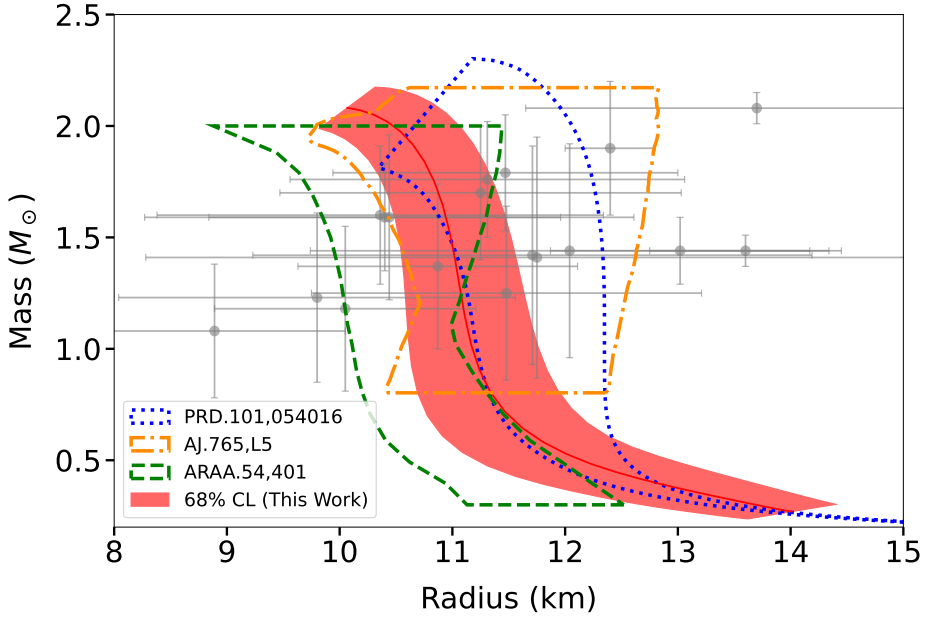

Figure 4.17: 68% CL of the M-R band corresponding to the reconstructed EoS in figure 4.16, shown as a red band. The M-R contours corresponding to the other EoS results from figure 4.16 are also depicted. The gray points with uncertainties are the $M-R$ observations with uncertainties listed in table 4.6. Figure taken from Ref. [3] (DOI:10.1103/PhysRevD.107.083028).

We further validate the proposed method by confronting the tidal deformability, Λ, predicted from the reconstructed EoS, with the constraint obtained from the gravitation wave event, GW170817 [262, 263] (see section 1.1.3 of chapter 1 for an introduction to tidal deformability). Since the EoS output by the `EoS Network` corresponds to densities greater than the saturation density, we adopt the DD2 EoS for sub-saturation densities. Consequently, we compute the corresponding Λ from the reconstructed EoSs. The resulting curve as a function of NS mass is shown in figure 4.18. Note that the constraint from GW170817, on a canonical $1.4\mathrm{M}_{\odot}$ NS, $\Lambda_{1.4} = 190^{+390}_{-120}$ [262] at 90% credible level, is also shown. For this reason, we also depict the 90% CL band in a fainter shade of red. The tidal deformability of a $1.4\mathrm{M}_{\odot}$ NS obtained from the reconstructed EoSs in this work, $\Lambda_{1.4} = 209.12^{+110.8}_{-110.8}$ (at 90% credible level) is in agreement with the bounds estimated from GW170817.

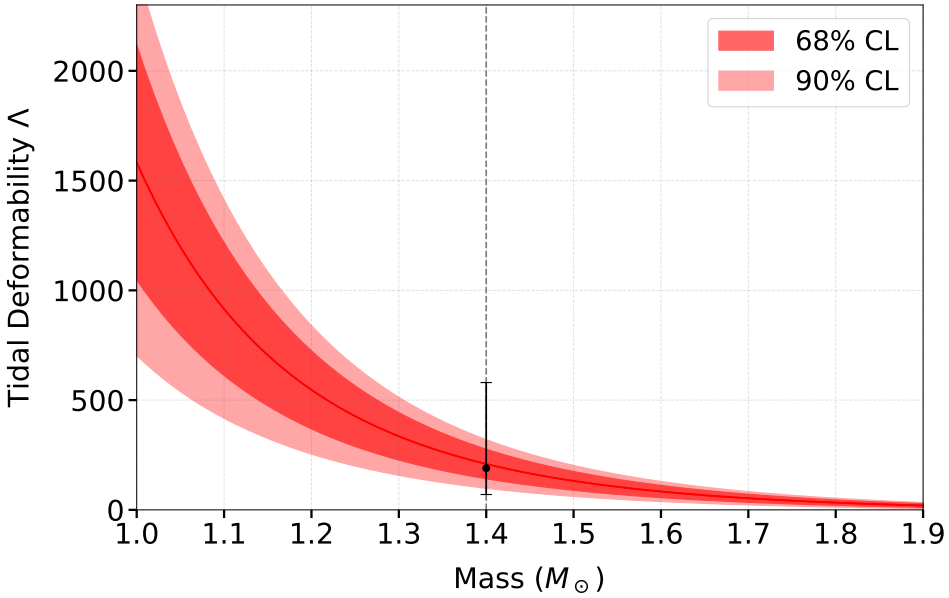

Figure 4.18: 68% and 90% CLs of tidal deformability, Λ, derived from the reconstructed EoS as a function of mass, depicted as darker and fainter shades of red, respectively. The 90% CL constraint from GW170817 is also shown here. Figure taken from Ref. [3] (DOI:10.1103/PhysRevD.107.083028).

The speed of sound in matter is an important characteristic of the EoS. Figure 4.19 displays the c_s^2 corresponding to the reconstructed EoS (with natural unit $c = 1$). While in the low and medium density regions ($\rho < 3\rho_0$), c_s^2 shows a slow increase and has a relatively narrow band, it exceeds the conformal limit (an upper bound for massless ultra-relativistic matter) in the high-density region ($\rho > 3\rho_0$), implying the existence of strongly interacting dense matter. In addition, it appears from the figure that c_s^2 likely decreases as density increases (at $\rho > 5 \sim 6\rho_0$). Such a smooth change for c_s^2 might result in the curve that eventually reaches the conformal limit. These findings were also published in other recent works [160, 264]. The explanation follows as a consequence of the perturbative-QCD (pQCD) constraint, where the asymptotically free quarks and gluons exist in the picture of hadron-quark continuity [265, 266]. Note that in this work, the conformal limit/pQCD constraint was not used as a prior for reconstructing the EoS, but is rather a possible outcome of the reconstructed EoS. Furthermore, from the results obtained in this work, there is no significant trace of a phase transition from the hadronic to deconfined quark matter at densities below $\approx 7\rho_0$. However, due to limited number of observations and the huge uncertainties that accompany them, the possibility of a phase transition remains ambiguous. If the number of observations and their accuracy are both improved significantly in the near future, the existence and order of possible phase transitions can be recognized in the developed framework. We demonstrate this in figures 4.20 and 4.21. In figure. 4.20, we illustrate that the `EoS Network`, with the same network structure, is capable of producing an EoS which contains a phase transition. We use one of the

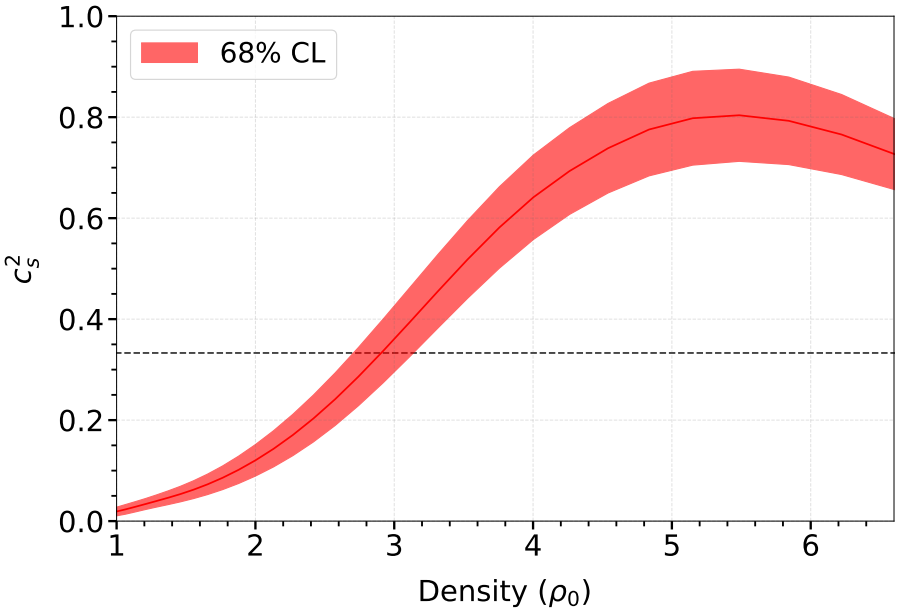

Figure 4.19: Speed of sound (68% CL) corresponding to the reconstructed EoS in figure 4.16. The dotted horizontal line represents the conformal limit, $c_s^2 = 1/3$.

DD2F-SF EoS models [141] here. In this case, we merely fit the network to the DD2F-SF EoS and observe that the kinks in the curve can be represented by the `EoS Network` output.

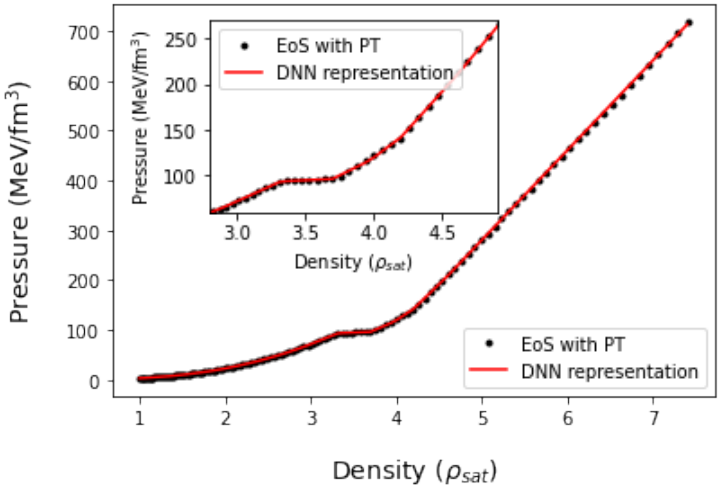

Figure 4.20: Demonstration that the `EoS Network` is capable of represent an EoS with phase transition. We use the DD2F-SF model here as an example.

In figure. 4.21, we employ the developed framework to test its ability in reconstructing a test EoS with a phase transition. We use the same DD2F-SF EoS [141], as in figure 4.20. It

is evident from the left panel of the figure that the `EoS Network` is capable of reconstructing the EoS (dashed red line) when compared with the true EoS (solid black line). On the right panel, we present the $M-R$ curve corresponding to the reconstructed EoS as a dashed red line, compared to the true $M-R$ curve (solid black line). However, it must be noted that in this case, we did not apply uncertainties to the $M-R$ data. We also used a significantly larger number of mock $M-R$ points.

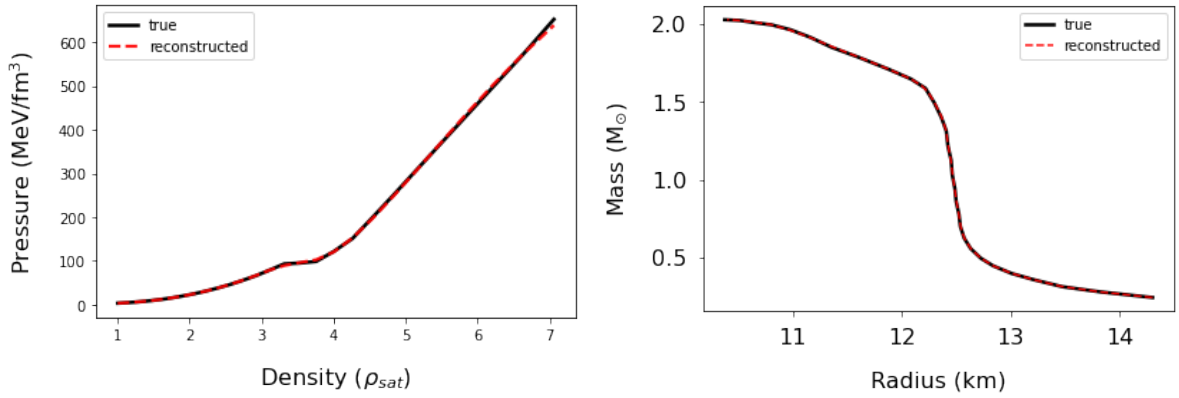

Figure 4.21: Testing the devised framework on an example EoS with phase transitions (DD2F-SF [141]). *Left:* The reconstructed EoS (dashed red line) is compared to the true test EoS that includes a phase transition (solid black line). *Right:* the M-R relationships corresponding to the EoSs in the left panel.

4.4 Discussions

The method developed here introduces a model-independent and flexible neural network representation of the dense matter EoS, which is obtained from observations of structural properties of neutron stars. While Bayesian approaches are capable of reconstructing a model-independent or a parameterized EoS, the optimization in a multi-dimensional parameter space is computationally inefficient. Moreover, as the Automatic Differentiation framework uses a pre-trained DNN to emulate the solving of the TOV equations, the method proposed here is fast and efficient. When compared to supervised learning approaches on the same research problem, the proposed method possesses a natural Bayesian picture for interpretation. Besides providing a novel alternative for the EoS reconstruction, this physics-based unsupervised learning algorithm uses importance sampling for estimating the uncertainties related to the observational noise. This work can be further extended to incorporate a Bayesian Neural Network in future. In such a set-up, we could include additional priors from heavy ion collisions (HIC) (see Ref. [267] for a reconstruction of the QCD EoS from Bayesian analysis of HIC data). A combined study with the results from Ref. [267] is in foresight. An interesting aspect which can be explored is the inclusion of microphysical EoSs which undergo a first-order phase transition (FOPT) or a crossover

from hadronic matter to deconfined quark matter. There is a possibility of such an occurrence of a FOPT in the cores of NSs, which may have dramatic consequences also for binary neutron star mergers, BNSMs [268], and for core collapse supernovae [269]. In the current study we do not explicitly include such EoSs for training the `TOV-Solver Network` (e.g., constant pressure with increasing energy density).

A similar study can be carried out by incorporating the $M-\Lambda$ relationships of NSs in addition to the $M-R$ relations. The tidal deformability estimates obtained from GWs however, have even larger uncertainties. Future measurements from next generation GW detectors, like LIGO India, KAGRA, Einstein Telescope, Cosmic Explorer and LISA, with greater sensitivities are expected to yield better constrained estimates for the tidal deformability. The present study is restricted to $M-R$ observations of long-lived cold neutron stars. However, it can be further extended to observations from proto-neutron stars and to remnants of BNSMs, where the high temperatures cannot be ignored. The thermal contributions in the pressure can be incorporated separately in a piecewise polytropic approach, such as, $P = P_{\rm cold} + P_{\rm thermal}$. The adiabatic index of the thermal pressure has been examined in earlier works and can be set to $\Gamma_{\rm th} \sim 1.5$-2.0 [270]. With an increase in the number of GW event detections, there is scope for the study of long-lived remnants from BNSMs using a non-zero temperature EoS in future.

Chapter 5

Gravitational Wave Analysis using DL

Gravitational waves (GWs) are ripples caused by the disruption of space-time, for example, due to massive accelerating objects like neutron stars or black holes orbiting around each other. These ripples of distorted space-time, travel through the universe carrying information about their cataclysmic origins (see section 1.1.3 of chapter 1 for a brief introduction to GWs and GW detectors). The first detection of GWs from the collision of a binary black hole (BBH) system by Advanced LIGO on 14^{th} September, 2015 [271], marked the beginning of the gravitational wave astronomy era. Several GW events have been detected by the LIGO-Virgo Scientific collaboration subsequently. More specifically, 11 events were observed in the first and second observing runs (O1 and O2), and 79 in the third observing run (O3) [272–275]. The fourth observing run, O4, was initiated on 24^{th} May 2023, and is expected to include 18 months of observation time [276]. The GW events observed so far include binary black hole mergers (BBHMs), binary neutron star mergers (BNSMs), neutron star-black hole mergers (NSBHMs) [277] and also mergers of component objects from the 'mass gap' [278]. The first detection of a BNSM event, GW170817, by Advanced LIGO and Virgo has opened venues for significant advancements in the ongoing study on neutron stars (NSs) and the underlying equation of state (EoS) [104]. It was also observed across the electromagnetic spectrum [279], making it a multi-messenger observation. This unique event therefore provided insights not only on the maximum mass of NSs [252,253], but also on the formation of r-process elements [280], short gamma-ray bursts (sGRBs) [281], and the Hubble constant [282]. A significant aspect of GW170817, however, was the extraction of tidal deformability (Λ) (see section 2.2.2 of chapter 2), for the very first time [105].

The dependence of tidal deformability on radius enables predictions for the radius of a canonical 1.4 $M_\odot$ neutron star [1, 283, 284]. These estimates provide information complementary to the independent radii measurements by the NICER (Neutron Star Interior Composition Explorer) collaboration (see section 1.1.2 of chapter 1 for more on NICER and its radii measurements). The combined results from NICER and GW170817 allow for constraints on the EoS, and provide estimates for the radius and moment of inertia of a 1.4 $M_\odot$ NS [285]. These serve as valuable probes for dense matter in the relatively

low density range (n_s - $3n_s$, where n_s here denotes the nuclear saturation density) [286]. Additionally, by analyzing the electromagnetic observations post GW170817, certain studies have provided estimates on the upper limit of the maximum non-rotating neutron star mass [252, 253]. Shortly afterwards, a new lower bound on the maximum mass was evaluated from the analysis of event GW190814, by assuming the secondary mass was a rapidly spinning neutron star [287]. These results have the potential to constrain the EoS in the high density regime ($3n_s$ - $5n_s$). All these findings demonstrate that gravitational waves are instrumental in providing direct or indirect constraints on the dense matter EoS.

The large amount of data associated with GW observations calls for efficient methods of data analysis. This makes DL techniques particularly promising in the field of GW astronomy [288–293]. More specifically, it was demonstrated that DL is very efficient for the real-time GW detection of BBH mergers and reliable parameter estimation [288, 294–296], forecasting BNS mergers [297], rapid identification of transient GW signals from BNS mergers [289], identifying glitches in GW data [298], denoising GWs [290], and for the detection and classification of NSBH mergers [299]. So far, DL methods for GW analysis have focused mainly on binary black hole mergers. Although deep learning has also been used on binary neutron star mergers signals [289, 297, 300], the inference of tidal deformability has received little attention as of now. In this chapter, we focus on analyzing GWs of BNSMs for constraining the dense matter EoS from the tidal parameter, using deep learning (DL) methods. We start by introducing a classification scheme to distinguish these events from binary black hole mergers or noise. We then present a regression method to infer chirp mass and combined tidal deformability from the GW signal of the BNS system. We further illustrate that the extraction of both $\mathcal{M}$ and $\tilde{\Lambda}$ could provide direct constraints on the NS EoS.

5.1 Data Generation

The standard conventional method to analyze gravitational wave data for signal-detection and for parameter-estimation is the template matched filtering [301]. This is a technique that employs template banks containing waveforms from a vast parameter space. However, this method is computationally intensive. In order to circumvent this challenge, we use deep neural networks (DNNs) to analyze simulated gravitational waveforms. The major computational costs while using DNNs occur during the training process, which requires large amounts of data. This GW data is generated using waveform-approximant models from the LALSuite library [302] due to the computational expenses associated with Numerical Relativity (NR) simulations. Model waveforms like the PhenomDNRT, PhenomPNRT, SEOBNRT, etc, which use the aligned-spin point-particle model (with and without precession effects), and the aligned-spin point-particle effective-one-body (EOB) model [303–309], perform up to a high level of accuracy when tested with NR simulations. These waveforms are generated by adding a tidal amplitude correction to their corresponding BBH baseline waveforms. Here, we use the inspiral-merger-ringdown (IMR) precession models "IMRPhe-

nomPv" and "IMRPhenomPv2_NRTidalv2" [310] for simulating gravitational waveforms of BBHMs and BNSMs, respectively. Note that for the BBHM waveforms, we use the same baseline waveform, however, without tidal effects. We generate the signals in the frequency domain, as opposed to Ref. [289], where the waveforms were analyzed in the time domain. All gravitational waveforms used in this analysis are generated without spin.

For binary black hole mergers, the component masses are chosen within the range of [5, 50]$M_\odot$, with the mass selection following from a uniform distribution in the specified range. The binary neutron star systems in our study have component masses ranging from 1.2 $M_\odot$ to 2.2 $M_\odot$, with the higher mass designated as the primary mass (m_1) and the lower mass as the secondary mass (m_2). In order to generate BNS signals, we also input the tidal deformabilities (Λ_1, Λ_2) of the individual stars. For this, we adopt generalized $\mathrm{M}-\Lambda$ relations, which do not rely on specific equations of state. Note that we present two different cases for the classifier, where one of them uses BNSMs based only on the DD2 EoS. Further information on the second case is provided at the end of this section. To define a broad region in the $\mathrm{M}-\Lambda$ parameter space that can include several possibilities of the $\mathrm{M}-\Lambda$ curve, we employ two fitting functions: $g_1(x) = a_1 \exp(b_1 x)$ and $g_2(x) = a_2 \exp(b_2 x)$. These fit functions serve as upper and lower envelopes encompassing several microphysical EoS models, as depicted in figure 5.1. The tidal deformability for each of the binary stars is randomly assigned a value within the range bounded by the envelopes, the dotted ($g_1(x)$) and dashed ($g_2(x)$) black lines in figure 5.1. For a chosen primary mass, m_1, the corresponding tidal deformability, Λ_1, is randomly assigned a value between the two fit functions $g_1(m_1)$ and $g_2(m_1)$. Here, we use a_1 = 6.45e+05, b_1 = -4.386, a_2 = 2.45e+05, and b_2 = -6.16. The tidal deformability, Λ_2, of the secondary mass (m_2) is then selected from a uniform distribution within the range (Λ_1, $g_1(m_2)$]. This approach ensures that Λ is a monotonically decreasing function of M within the specified region of interest. Note that in this approach, we ignore the possibility of stable twin stars, where the tidal deformability depicts two distinct branches for the same mass NS [311]. Furthermore, it is important to mention that the EoSs in figure 5.1 are not all necessarily within the Λ-constraints set by GW170817 as shown in [1]. However, these choices of upper and lower envelopes ensure that the relevant range of $\mathrm{M}-\Lambda$ is enclosed by the parameters we use for training the regressor network.

All the waveforms are generated within a specified frequency range, $f \in [f_{\mathrm{low}}, f_{\mathrm{high}}]$ Hz. The exact frequency values for f_{low} and f_{high} are mentioned while describing the various network architectures. To account for the encountered noise characteristics, we utilize conventional methods for noise generation in gravitational wave astronomy. Firstly, we generate a colored noise background based on the power spectral density (PSD) of the "Advanced LIGO Zero-Detuned High Power" configuration. This background noise is representative of the colored detector noise present in GW observations. Subsequently, the simulated BBHM and BNSM signals are injected into this noise background. The injection is performed to achieve optimal signal-to-noise ratio (oSNR) values falling within the range

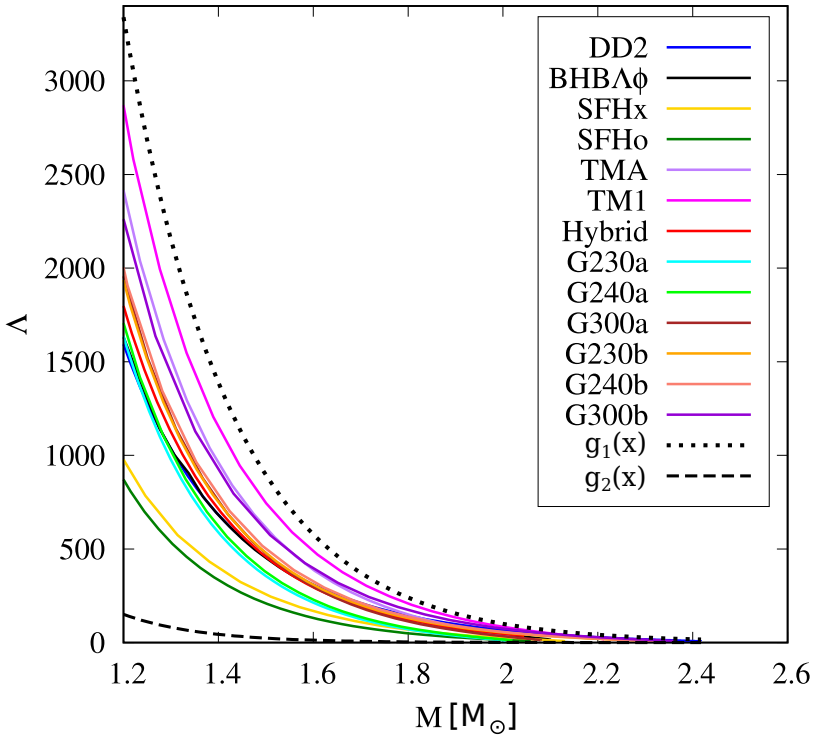

Figure 5.1: The M − Λ curves of several EoSs taken from [1] are displayed as solid colored curves. The black dotted and dashed curves labelled as $g_1(x)$ and $g_2(x)$, respectively, are chosen functions to span a wide range in the M − Λ space. The M − Λ data points enveloped in this region are used to simulate GW signals of BNSMs in the regression task. See Ref. [6] for original figure.

of 20 to 30. The optimal SNR (ρ_{opt}) is defined as

$$\rho_{opt} = 2\left[\int_0^\infty df \frac{|\, h(f)\,|^2}{S_n(f)}\right]^{1/2} \tag{5.1}$$

where $h(f)$ is the gravitational wave signal in the frequency domain, and $S_n(f)$ is the detector PSD [312]. These simulated signals in the colored noise background are then whitened for further analysis in the classification and regression networks. We present an example each of the BBH and BNS signals generated with this quantification of noise, in figure 5.2. We depict the clean signals in red on both panels. On the left panel, we show a BBHM signal with $m_1 = m_2 = 15\ M_\odot$. The BNSM signal on the right is simulated with $m_1 = 1.75\ M_\odot, m_2 = 1.5\ M_\odot$ and $\Lambda_1 = 176.94, \Lambda_2 = 450.05$ (note that these values for the plot are chosen based on the DD2 EoS). The grey waveforms represent the signals injected into aLIGO's colored noise. We whiten these signals before using them in the classifier/regressor networks. A total of 75,000 such waveforms are generated for each class.

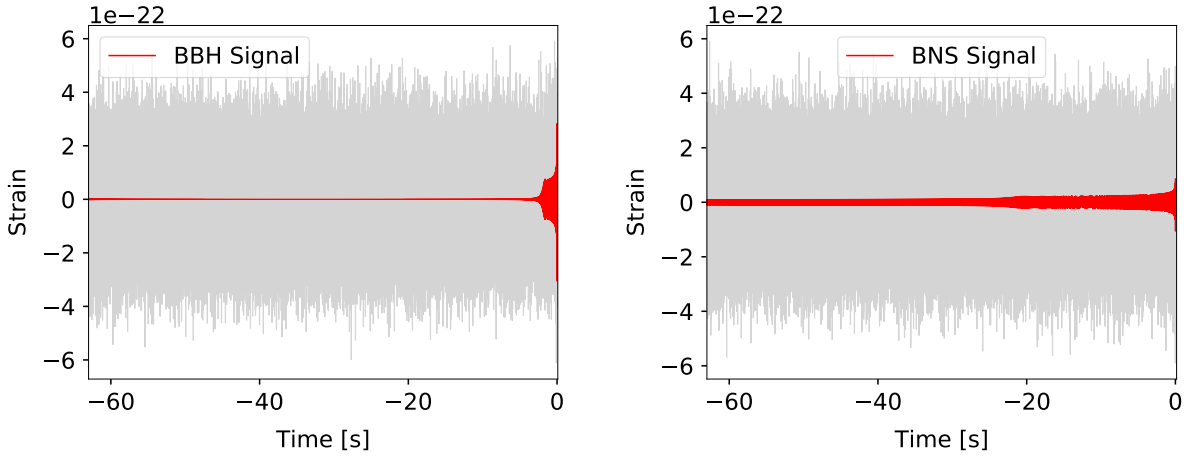

Figure 5.2: Examples of time-domain inspiral-merger signals of a BBH and a BNS injected in aLIGO's noise. Here, the optimal signal to noise ratio (oSNR) is 20. The grey waveforms represent the BBHM (left) and BNSM (right) signals injected into colored noise. The clean signals are depicted in red. For the BBHM signal, we chose the component masses, $m_1 = m_2 = 15\ M_\odot$. As for the BNSM signal, the component masses are $m_1 = 1.75\ M_\odot$ and $m_2 = 1.32\ M_\odot$. The corresponding tidal deformabilities for this waveform are chosen based on the DD2 EoS, i.e., $\Lambda_1 = 176.94$ and $\Lambda_2 = 450.05$. See Ref. [6] for original figure.

Note that for the classifier network, we use two different cases of noise quantification. Apart from using the oSNR, we also use the peak signal to noise ratio (pSNR) to measure the strength of the GW signals. We follow the methods from Ref. [289] for this case (dubbed as Case II in subsection 5.2.1). For BBH mergers, we use the same mass range [5, 50] $M_\odot$ in a uniform distribution. As for binary neutron star (BNS) mergers, the component masses are based on the DD2 equation of state (EoS) [131], and are therefore selected to lie within the range of [1, 2.4] $M_\odot$. The corresponding tidal deformation values for these masses are computed from the DD2 EoS and fixed accordingly. For this case of the classifier network, a simple noise model is applied, based on the peak signal-to-noise ratio, or pSNR. The pSNR is defined as the ratio of the peak amplitude of the simulated signal to the standard deviation of the noise [288, 289]. Several realizations of Gaussian noise with zero mean and unit variance, $\mathcal{N}(0, 1)$, are generated, and the simulated gravitational wave (GW) signals are injected with varying amplitudes. The simulated waveforms are scaled to achieve peak signal-to-noise ratio (pSNR) values of 0.5, 0.75, and 0.9. Similar to figure 5.2, two examples of simulated waveforms in the time-domain are presented in figure 5.3 using pSNR values of 0.5. On the left, we see a BBHM signal injected into Gaussian noise. The clean signal of an equal mass BBHM ($m_1 = m_2 = 50\ M_\odot$) is shown in red, while the grey waveform represents the BBHM signal injected into noise. The right panel depicts a waveform where a BNS inspiral-merger signal is injected into Gaussian noise. The grey waveform represents the injected BNSM signal embedded in the simulated noise. The clean inspiral-merger signal of an equal-mass BNS system ($m_1 = m_2 = 1.32\ M_\odot$

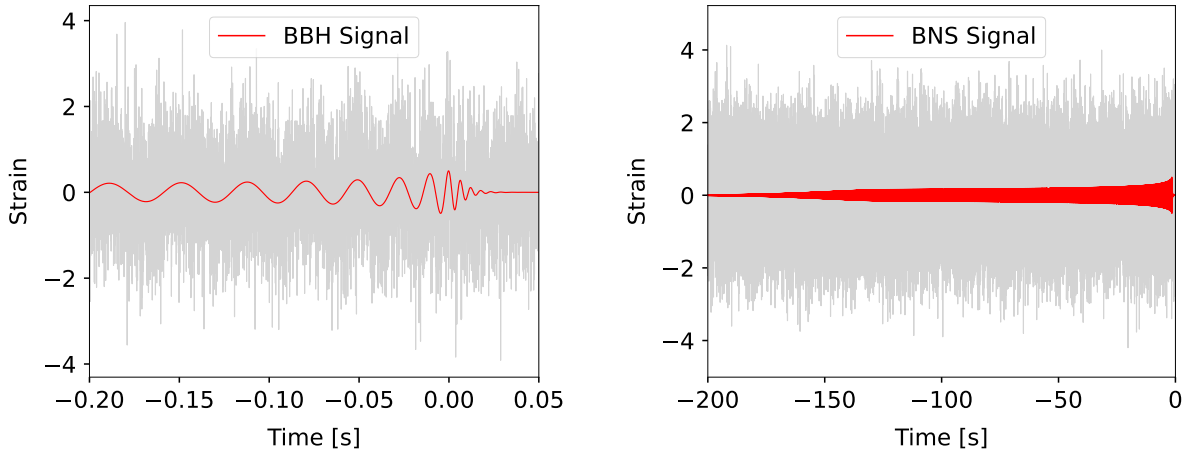

Figure 5.3: Examples of inspiral-merger signals of a BBH and a BNS in the time-domain. We choose peak signal to noise ratio (pSNR), or ratio of the peak amplitude of signal to the standard deviation of the noise, to be 0.5 in both subfigures. *Left:* The grey waveform represents the signal injected into simulated Gaussian noise. The clean BBH signal is depicted in red, and was simulated for component masses, $m_1 = m_2 = 50\ M_\odot$. *Right:* Same as left panel for BNS signal injected into simulated white Gaussian noise. The signal in red depicts a clean BNS signal simulated for equal component masses, $m_1 = m_2 = 1.32\ M_\odot$ and $\Lambda_1 = \Lambda_2 = 964$, obtained from the DD2 EoS. See Ref. [6] for original figure.

and $\Lambda_1 = \Lambda_2 = 964$, obtained from the DD2 EoS) is shown in red.

5.2 GW Signal detection from BBHMs and BNSMs

In this section, we explore the detection of signals from binary black hole and binary neutron star mergers in gravitational wave data. The following subsections provide details on the architecture of the classification network (subsection 5.2.1) as well as the associated results (subsection 5.2.2).

5.2.1 Classifier: Deep Learning Models

We employ the methods described in section 5.1 to generate sufficient training and testing data. The network architectures are based on the convolutional neural network (CNN), first introduced in chapter 3 (section 3.1.2). The classifier networks each consist of an input layer, multiple 1D convolutional layers, and two dense layers. The output layer of each classifier consists of three nodes, corresponding to the classification of the input waveform into one of three categories: binary black hole mergers, binary neutron star mergers, or waveforms with just noise (no signals). We perform our analysis on GW simulations that represent clean signals, i.e. they do not incorporate any noise, as well as those which

incorporate noise that is increased in consecutive steps. Two cases based on the noise quantification, i.e., oSNR and pSNR, are discussed below.

Case I: For this case, we generate the GW signals of binary black hole and binary neutron star mergers for classification using the noise quantification method based on oSNR, as described in section 5.1. The network designed for classification consists of the input layer, several 1D convolutional layers and two dense layers. The output layer of the network contains three nodes for classifying the input waveform into one of three categories: BBH mergers, BNS mergers or no signals. The structure of the classification network is detailed in table 5.2. We employ $f_{\rm low} = 23$ Hz, $f_{\rm high} = 1024$ Hz and $\Delta f = 1/64$. Therefore, we obtain the length of the input waveform as 64063. The convolution layers have 32 or 64 filters as mentioned in the table. In addition, these convolutional layers use kernels of size 16 and strides of 4, 4, 2, 2 and 2 in that respective order. The leaky rectified linear unit or 'LeakyReLU' non-linear activation ($\alpha = 0.05$) is applied to all the hidden layers, and the 'softmax' function is applied to the last layer. The target labels are one-hot encoded to facilitate the classification task. A waveform in the frequency domain, $h(f)$, is written as $A(f)\exp{(i\phi(f))}$ where A is the Fourier amplitude or absolute value and ϕ is the phase or argument. Both the absolute and phase values of the input waveforms are input to the classification network. The utilization of two input channels to the network results in the shape of the network (64063, 2). The first channel, or the Fourier amplitude is first normalized to have a unit integral value. The waveforms are then scaled (with the maximum values) to lie within the range [0, 1]. The second channel, i.e., the phase, ϕ, is scaled with π, such that it also lies within [0, 1]. Furthermore, we use the cross entropy loss function (see Eq. (3.14) of chapter 3) and an Adam optimizer [212] (see section 3.4.3 of chapter 3) with a learning rate of 10^{-4} for this case of the classifier. 66,000 waveforms from each category, i.e. a total of 198,000 waveforms from the three categories, are shuffled and used to train the classification network. A batch size of 16 is used to fit the network model to the training data. The validation data uses 9,000 waveforms from each category, adding up to a total of 27,000 waveforms. These validation or testing samples are used to assess the performance of the network during training. The training is initiated with the clean signals and then continued on signals with lower oSNR values (30, 25, and 20). The network is trained for 15 epochs in the case of clean signals. However, we increase the number of epochs as we include higher levels of noise. For oSNR values of 30, 25 and 20, we train the network for 30, 60 and 100 epochs, respectively. The results of the classifier for different oSNR values are discussed in next subsection 5.2.2. Below, we present Case II, which uses pSNR to quantify the signal to noise ratio in the gravitational waveforms.

Case II: In this case, we use pSNR to describe the strength of the GW signals. The detailed structure of the classification network in this case is provided in table 5.2. For the input waveforms here, we set $f_{\rm low} = 23$ Hz, $f_{\rm high} = 1024$ Hz, and $\Delta f = 1/128$. Consequently, the length of the input waveform amounts to 128,257. Each convolutional layer has 64 filters, and a convolutional kernel of size 3 is used here. The strides for the convolutional layers are set to 4, 4, 2, 2, and 2 in that respective order. The rectified linear unit (ReLU) activation function is applied to all hidden layers, while the final layer

Table 5.1: Classification model architecture used for classifying the simulated input GW signals into 3 classes, i.e., BBH mergers, BNS mergers or noise. The network consists of 2,094,787 trainable parameters. See Ref. [6] for the original table.

Layer Index	Layer	Dimension
1	Input	64063 x 2
2	Convolution 1D	16012 x 32
3	Convolution 1D	4000 x 32
4	Convolution 1D	1993 x 64
5	Convolution 1D	989 x 32
6	Convolution 1D	487 x 32
-	Flatten	15584
7	Dense Layer	128
-	Dropout	128
8	Dense Layer	3
-	Output	3

utilizes the softmax function. The target labels are one-hot encoded. For the classifier in this case, we only input the absolute values to the network. The input waveforms are normalized to have an integral value of unity. However, the amplitude is then amplified by rescaling the waveforms such that they lie within the range [0, 1]. The classifier utilizes the cross-entropy loss function, along with the Adam optimizer [212], just as in the Case I. The optimizer is configured with a learning rate of 10^{-5} and helps in effectively training the network for accurate classification. To train the neural network, 5,000 waveforms from each category are utilized, resulting in a total of 15,000 training samples. These waveforms from the three categories are shuffled before being used as training data for the classifier. A batch size of 32 is employed during the training process. For the validation set, 2,473 waveforms from each category are selected, totaling 7,419 waveforms. The training process is initiated using clean signals and subsequently continued with gravitational waveforms having lower pSNR values of 0.90, 0.75, and 0.50. The network is trained for 110 epochs when using clean signals and a pSNR of 0.9. However, as the pSNR value decreases, we increase the number of epochs. For pSNR values of 0.75 and 0.5, the network is trained for 150 and 200 epochs, respectively. Note that the training takes $\sim$1 hour for about 100 epochs. The results from the trained classifier for different pSNR values are discussed in the subsequent subsection.

Table 5.2: Model architecture of the classifier used for segregating the simulated input GW signals into 3 classes, i.e., BBH mergers, BNS mergers or noise. The network consists of 8,250,371 trainable parameters. See Ref. [6] for the original table.

Layer Index	Layer	Dimension
1	Input	128257
2	Reshape	128257 x 1
3	Convolution 1D	32064 x 64
4	Convolution 1D	8016 x 64
5	Convolution 1D	4007 x 64
6	Convolution 1D	2003 x 64
7	Convolution 1D	1001 x 64
-	Reshape	64064
8	Dense Layer	128
-	Dropout	128
9	Dense Layer	3
-	Output	3

5.2.2 Classifier: Results

Here, we discuss the results from the trained classifier networks obtained on testing data. For each test sample, the classifiers output a corresponding class that the waveform belongs to, i.e., a binary black holes merger, a binary neutron star merger or a waveform that contains neither, but only noise. We can compare the predicted classes with the true class labels in a confusion matrix. A confusion matrix, also known as an error matrix, of a binary classifier summarizes the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) in a table of two rows and two columns. These values help quantify the performance of the classifier. Two common metrics used in a binary classification are 'Precision', P, and 'Recall', R, where P = TP/(TP+FP) and R = TP/(TP+FN). The same concept can be extended to multi-classifiers, where P and R are computed for each class and the obtained values are then averaged over the number of classes. This method is termed as macro-averaging and the resulting value is the macro-average precision. We present the results obtained from the Case I and Case II classifiers as confusion matrices in figures 5.4 and 5.5, respectively. In figure 5.4, starting from the top row, the figure depicts the confusion matrices for clean signals and oSNR values of 30 on the left and right corners, respectively. In the bottom row, we show the classifier results for oSNR=25 and 20 in the left and right panels, respectively. We observe a high accuracy obtained by the classification networks in the top row. A macro-average precision

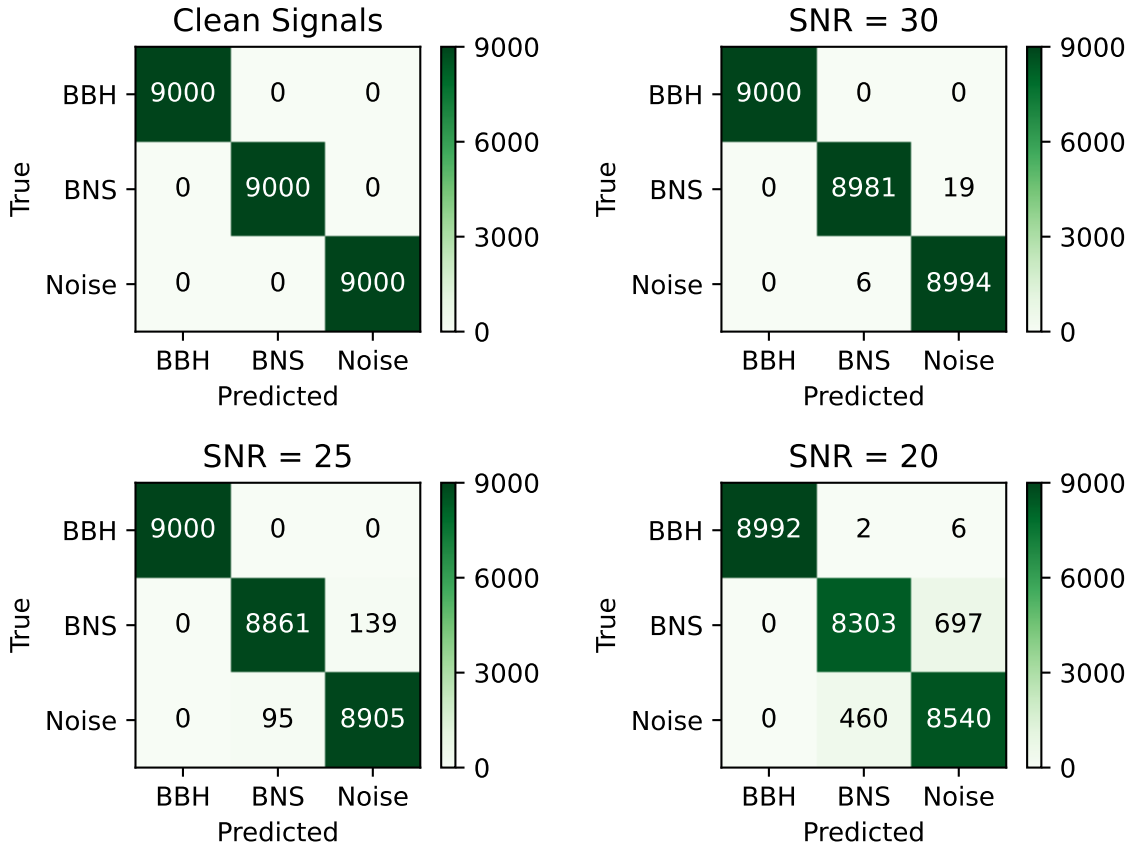

Figure 5.4: Performance of the classification network represented as confusion matrices for clean signals and different values of oSNR (30, 25 and 20). Moving from clean signals to increasing levels of noise, the classifiers are trained for 15, 30, 60 and 100 epochs, respectively. This is intuitive as a larger number of epochs are required for higher levels of noise in data. With decreasing oSNR, the number of misclassifications increases. Macro-average precision values of 1.0, ∼1.0, 0.99 and 0.96 are obtained for clean signals, oSNR values of 30, 25 and 20, respectively. See Ref. [6] for the original figure.

of 1.0 is obtained for the clean signals and ∼1.0 for oSNR=30. As the signal amplitude is further reduced with respect to the noise in the bottom row, we note that network requires longer training times to reach higher accuracies. High precision values are nevertheless maintained, i.e., macro-average precisions of 0.99 and 0.96 are obtained for oSNR=25 and 20, respectively. Similarly for Case II, we present the results in the four confusion matrices, where noise levels are decreased as one moves from top to bottom row (starting from the top left corner). The top panels include clean signals and pSNR=0.9, while the lower panels depict the results for pSNR=0.75 and 0.5. Macro-average precisions of 0.99, 0.99, 0.99 and 0.89 are obtained for clean signals, pSNR=0.9, 0.75 and 0.5, respectively. We remark that even in this case longer training times are required as the noise levels are increased with

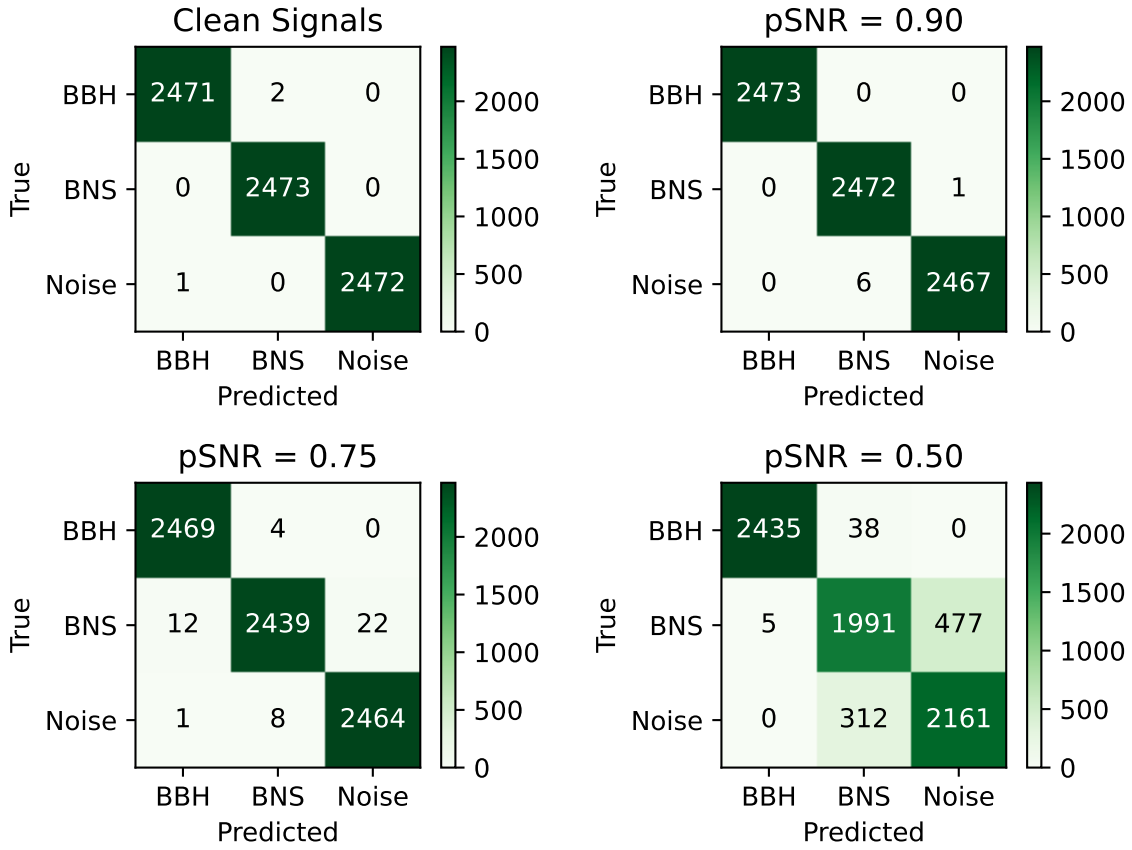

Figure 5.5: Performance of the classification network represented as confusion matrices for four different values of pSNR (ratio of the peak amplitude of signal to the standard deviation of noise). In the case of clean signals and pSNR=0.9, we train the classification network for 110 epochs. With pSNR=0.75 and 0.5, the network is trained for 150 and 200 epochs respectively. We observe that with decreasing pSNR, the number of misclassifications increases. Macro-average precision values of 0.99, 0.99, 0.99 and 0.89 are obtained for clean signals, pSNR values of 0.9, 0.75 and 0.5, respectively. See Ref. [6] for the original figure.

respect to the signal. These longer training times are reflected in figures 5.6 and 5.7, where we present the learning curves for the four different training processes in Case I and Case II, respectively. Here, we observe that as we decrease the oSNR from 25 to 20, or from pSNR=0.9 to 0.75, the performance of the classifier does not decrease but the classifiers require a larger number of training epochs to reach a high accuracy. Moreover, the learning curves for both cases also show that the classifiers does not overfit to the training data. Note that due to the simplicity of the model in Case II, a larger number of training epochs are required for reaching high accuracies. The number of samples used for training also plays a role in the performance of the classifiers. Due to the large dataset used in Case I,

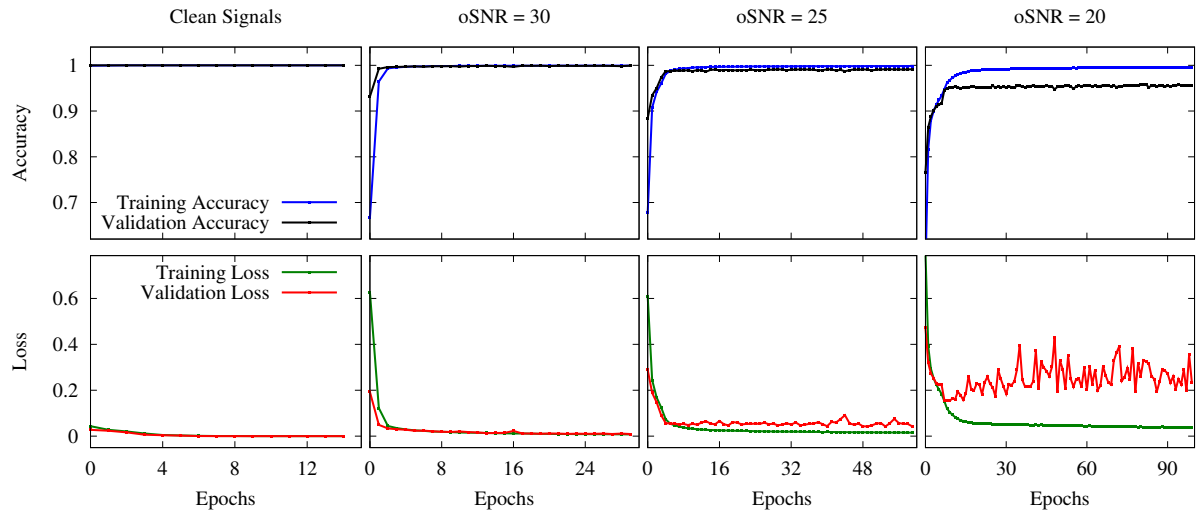

Figure 5.6: Learning curves for Case I of the classification network for different oSNR values (clean signals, oSNR=30, 25 and 20). Top and bottom panels depict the training and validation accuracies and losses for the four different networks as a function of the number of epochs, respectively. Note that for lower oSNR values, we train the network for a larger number of epochs in order to reach higher accuracies.

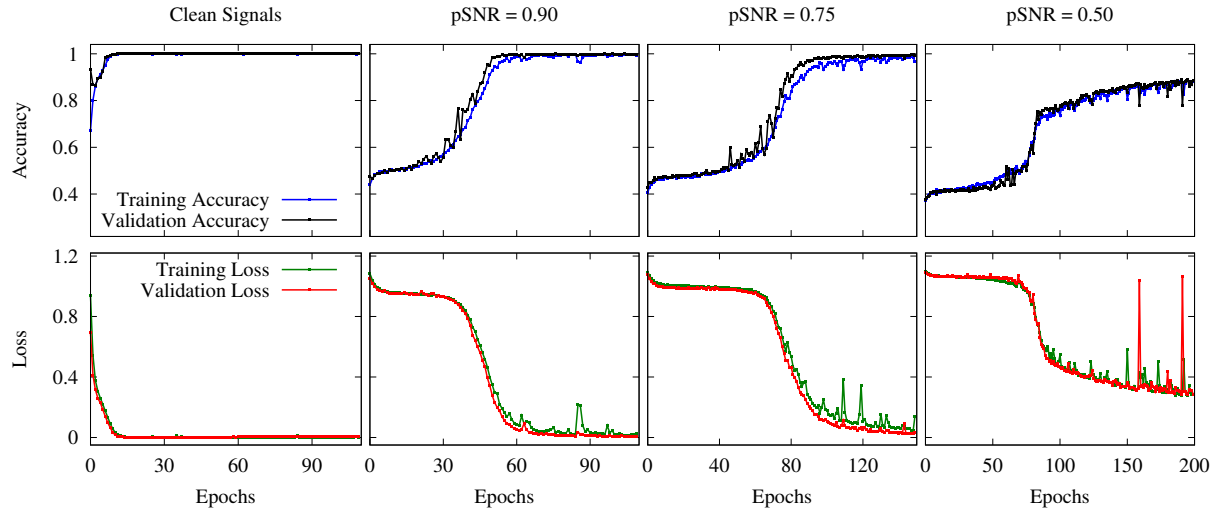

Figure 5.7: Similar to figure 5.6, we depict the learning curves for Case II of the classification network for different pSNR values (clean signals, pSNR=0.9, 0.75 and 0.5). Note that we use larger number of epochs for training compared to Case I, due to the simplicity of the network structure used here. The smaller number of training samples also plays a role in the training times.

we see that the network reaches high accuracies in the first few epochs compared to Case II.

5.3 Estimation of $\mathcal{M}$ and $\tilde{\Lambda}$ from BNSMs

In this section, we discuss the application of deep neural networks for parameter estimation from gravitational wave signals of binary neutron star mergers. The estimation of the individual binary masses in BNSMs has been studied using deep learning techniques [300]. In this study, however, instead of directly estimating the component masses, $\{m_1, m_2\}$ and individual tidal deformabilities, $\{\Lambda_1, \Lambda_2\}$, we propose the use of DL methods to estimate $\mathcal{M}$ and $\tilde{\Lambda}$ from simulated GW signals of BNSMs. The reasons for this approach are the obstacles faced in estimating the mass ratio, q. A precise estimation of the mass ratio is crucial to separate the individual masses from the chirp mass. However, obtaining an accurate estimate of q is challenging due to degeneracy with the aligned spin values in higher order post-Newtonian (PN) expansion terms [313]. Similarly, in order to untangle the individual tidal deformabilities, an estimate of the corrections to the tidal parameter $\delta\tilde{\Lambda}$ is necessary in addition to $\tilde{\Lambda}$. However, this correction term only appears at 6PN order, making $\tilde{\Lambda}$ a more feasible choice for regression [99, 100, 102, 200, 314]. In the following subsection 5.3.1, we demonstrate that the estimation of the chirp mass and combined tidal deformability can yield valuable constraints on the underlying equation of state (EoS). Details on data generation methods for training the network are already provided in section 5.1. Specifics of various DL models and the corresponding results follow in subsection 5.3.2.

5.3.1 Constraining the EoS from $\mathcal{M}$ and $\tilde{\Lambda}$ observations

The appearance of the chirp mass at the leading order in a PN expansion [99] of a gravitational wave makes it an easy target. The combined tidal deformability is chosen as the second target for the regressor. Despite appearing only at 5PN order, it helps that $\tilde{\Lambda}$ is of the order of a few 100 or more (when compared to the mass ratio which is < 1), making it more pronounced in a GW signal [99]. When provided with estimates of $\mathcal{M}$ and $\tilde{\Lambda}$ alone, we are faced with systematic uncertainties and difficulties in disentangling the component masses and the individual tidal deformabilities from the GW signal. Therefore, it is not feasible to determine a unique mass-tidal deformability ($\mathrm{M}-\Lambda$) relation. As a consequence, the underlying EoS remains ambiguous too. However, with sufficient GW events of BNSMs, the estimated $\mathcal{M} - \tilde{\Lambda}$ relations can still provide valuable constraints on the properties of dense matter. This is demonstrated in figure 5.8. The left and right sub-figures display the $\mathcal{M} - \tilde{\Lambda}$ relations as solid points and the $\mathrm{M} - \Lambda$ relations as dashed lines. Note that this figure can also be compared to Fig. 3 of [99]. The $\mathcal{M} - \tilde{\Lambda}$ relations are derived from all possible combinations of (m_1, Λ_1) and (m_2, Λ_2) of a particular equation of state (EoS), ensuring $m_1 \gtrsim m_2$. On the left panel, we present the relations for the SFHo and SFHx EoSs [133] (see chapter 2 for a brief description of these EoSs), indicated in green and red, respectively. A clear separation between the $\mathcal{M} - \tilde{\Lambda}$ points from the two EoSs is evident. We choose a small region from the $\mathcal{M}-\tilde{\Lambda}$ relations of SFHo EoS in the inset image. Within the chosen range, when the chirp mass is altered by $\mathcal{M}\pm$ 0.01 $M_\odot$, only small fluctuations of $\tilde{\Lambda}\pm5$, can be observed. This indicates that the analysis of gravitational wave (GW) signals from binary neutron star (BNS) mergers can provide additional constraints on the

EoS through the $\mathcal{M} - \tilde{\Lambda}$ relations. The relations for the DD2 [131] and BHB$\Lambda\phi$ [144] EoSs are shown in the right panel, in blue and black, respectively. In this case, the $\mathrm{M} - \Lambda$ curves of the two EoSs overlap until they reach a certain mass. The DD2 EoS is a nucleonic model, while the BHB$\Lambda\phi$ is an extension of the DD2 EoS to further include the Λ-hyperons (see chapter 2 for more on these EoSs). Consequently, there is a corresponding overlap in the $\mathcal{M} - \tilde{\Lambda}$ relations as well. This results from the softening of the EoS due to the hyperons that populate at higher densities. Nevertheless, the same argument can be made for distinct EoSs, as also shown on the right panel. It is also well-established that the parameter $\tilde{\Lambda}$ demonstrates minimal sensitivity to the mass ratio, q [8, 99]. Hence, by considering a given chirp mass, $\mathcal{M}$, and estimates with narrow ranges for $\tilde{\Lambda}$, it becomes possible to exclude numerous EoS models. Nonetheless, obtaining an adequate number of observations is crucial to get sufficient $\mathcal{M} - \tilde{\Lambda}$ estimates that cover a wide range in the parameter space. Despite the relatively large uncertainties in the combined tidal deformability derived from GW170817 and AT2017gfo ($197 \leq \tilde{\Lambda} \leq 720$) [272, 315], these measurements are already capable of ruling out extremely stiff EoSs [1]. Therefore, as the number of binary neutron star merger detections and their confidence increases, the extraction of combined tidal deformability will play a significant role in further constraining the properties of dense matter and the underlying EoS.

5.3.2 Regressor: Deep Learning Models and Results

The BNS waveforms generated from the inspiral-merger-ringdown approximant as described in section 5.1, with oSNR in the range [20,30], are segregated for training and testing the regression network. The regressor is trained and tested on data with clean signals as well as those with noise. We use convolutional neural networks (CNNs) for all DL regressor models (see section 3.1.2 of chapter 3 for an introduction to CNNs). We begin by training two independent networks, each to output only one of the two parameters, either chirp mass ($\mathcal{M}$) or combined tidal deformability ($\tilde{\Lambda}$). The analysis is carried forward by training one network to regress both parameters simultaneously. Different DL models were explored and the details are described below.

Independent Parameter Regressor

In this section, we discuss the training and testing of two independent regressors, one to estimate the chirp mass ($\mathcal{M}$), and the other to estimate combined tidal deformability ($\tilde{\Lambda}$) from the signals of BNS mergers. This means that for training the $\mathcal{M}$-regressor, for each event, the network only learns the chirp mass. No explicit information of the $\tilde{\Lambda}$ value is given to the $\mathcal{M}$-regressor network. The same condition holds for the $\tilde{\Lambda}$-regressor. Any information on $\mathcal{M}$ to the $\tilde{\Lambda}$-regressor is only implicit through the gravitational waveform itself. Both the $\mathcal{M}$-regressor and the $\tilde{\Lambda}$-regressor take as input the amplitude (A) and phase (ϕ) of the waveform, $h(f)(= A(f)\exp(i\phi(f)))$. For this simple case, the input waveform is sparsely spaced on the frequency scale. We choose $f_{\mathrm{low}} = 128$ Hz, $f_{\mathrm{high}} = 1024$ Hz and $\Delta f = 1/16$. Therefore, this yields an input shape of (14336, 2) to the regressors. Note

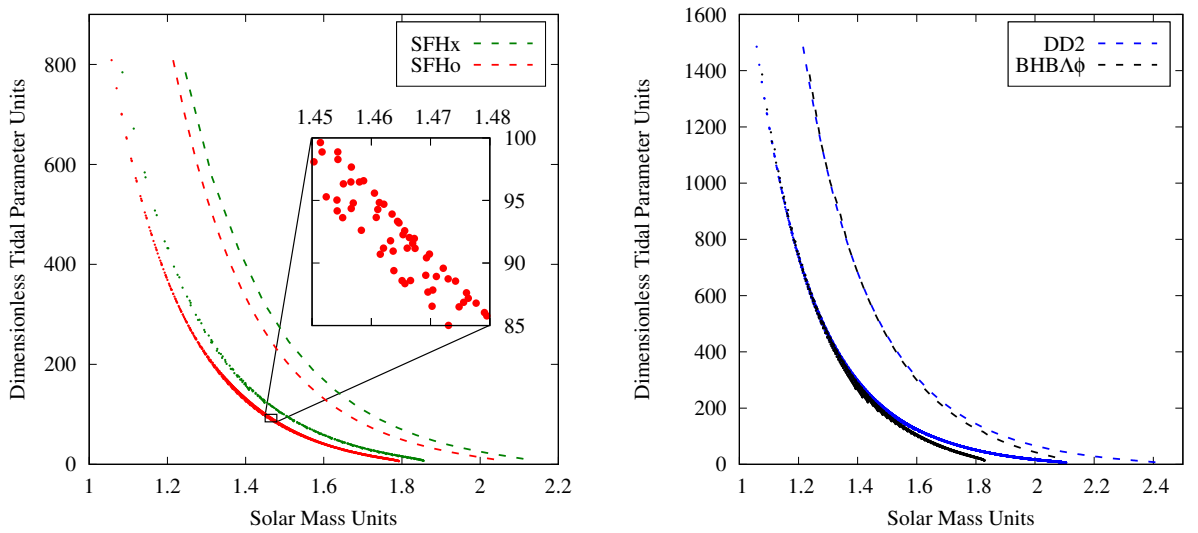

Figure 5.8: *Left:* M − Λ and $\mathcal{M} - \tilde{\Lambda}$ curves of SFHx (green) and SFHo (red) EoSs. The dashed lines depict the M − Λ curve. The solid points are $\mathcal{M} - \tilde{\Lambda}$ values, obtained using several plausible combinations of component masses and tidal deformabilities of a particular EoS. The inset image displays a finer resolution of the $\mathcal{M} - \tilde{\Lambda}$ space within a specific range. The area spanned by the $\mathcal{M} - \tilde{\Lambda}$ points of the SFHo EoS in the zoomed-in region is not broader than ± 5 dimensionless units for $\pm 0.01 M_{\odot}$. *Right:* Same as the left panel, but for the DD2 (blue) and BHBΛϕ (black) EoSs. In this case, we observe the distinction of the two M − Λ curves (dashed lines) around $1.6 M_{\odot}$. This is due to the appearance of hyperons at the corresponding density. This effect is consequently also observed in the $\mathcal{M} - \tilde{\Lambda}$ relations (solid points). See Ref. [6] for the original figure.

that the dimension 2 appears due to the two channels (A and ϕ) of the network input. Each regressor is architectured with five 1D convolutional layers. We place max-pooling layers for down-sampling (pool size = 2) following the second and the fourth convolutional layer. Two dense layers follow the last convolutional layer, marking the ultimate and penultimate layers of the regressor networks. Starting from the first convolutional layer and moving deeper into the network in order, we use 16, 16, 32, 32 and 16 filters, and kernels of sizes 32, 32, 24, 24 and 16 respectively. The weights of the convolutional kernels are initialized from a *He* normal distribution [210]. We use a stride of size 3 when applying the convolutional operator for the first four layers, and a stride of 2 for the last convolutional layer. The 'ReLU' activation function is employed to all the convolutional layers and the penultimate dense layer. The ultimate dense layer uses the 'sigmoid' activation. Here, the loss function is defined as the mean squared error (MSE). An Adam optimizer [212] with a learning rate of 0.0001 is applied. Accuracy of the regressors is quantified using the $\mathcal{R}^2$ metric[1], or the coefficient of determination (refer to Eq. (3.13), section 3.3 of chapter 3).

[1]Note that the $\mathcal{R}^2$ metric was also used in both the `TOV-Solver Network` and the `EoS-Network` in Chapter 4 for EoS-reconstruction.

In order to normalize the labels (here, $\mathcal{M}$ or $\tilde{\Lambda}$) of the regressor, we use the normalization,

$$y_{\rm norm} = (y - y_{\rm min})/(y_{\rm max} - y_{\rm min}), \tag{5.2}$$

where, $y = \{\mathcal{M}, \tilde{\Lambda}\}$. The normalized labels then lie in the range [0,1]. The input to the regressors is normalized too. The two channels, Fourier amplitude and the phase, use the normalization method described for the Case I classifier, under subsection 5.2.1.

As described in section 5.1, the complete waveform data set is generated using the component masses (m_1, m_2) and individual tidal deformabilities (Λ_1, Λ_2) of a BNS system. When m_1, m_2, Λ_1, and Λ_2 are represented in terms of $\mathcal{M}$ and $\tilde{\Lambda}$, respectively, they span the region depicted in figure 5.9.

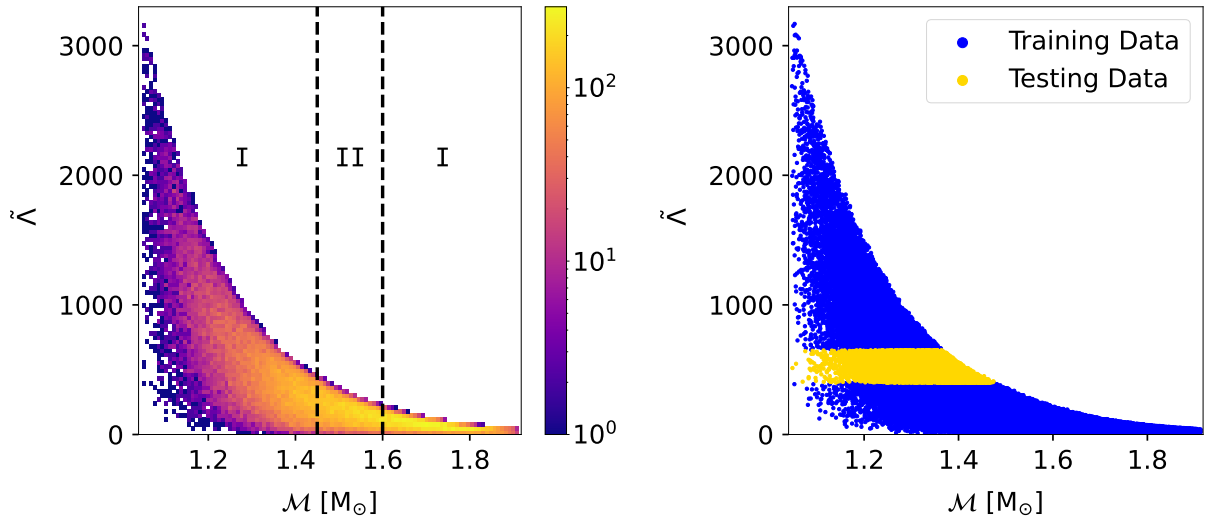

Figure 5.9: Training and testing data choice for independent parameter regression. *Left*: Density plot displaying the distribution of chirp mass and tidal deformability ($\mathcal{M} - \tilde{\Lambda}$) values for all simulated gravitational wave (GW) events. The training samples for the $\mathcal{M}$-regression are located within region `I`, while the testing samples are represented by region `II`. The total number of simulated events sum up to 75,000. *Right*: Individual GW events used for training and testing the $\tilde{\Lambda}$-regressor depicted by blue and yellow points, respectively. Note that the distribution of the events in the $\mathcal{M}-\tilde{\Lambda}$ space for the $\tilde{\Lambda}$-regressor, is consistent with that in the left panel. See Ref. [6] for original figure.

On the left panel of the figure 5.9, we show the data as a density plot. The plot covers a total of 75,000 simulated BNSM events. The regions marked by `I` and `II` correspond to training and testing data used for the $\mathcal{M}$-regressor, respectively. On the right panel, we illustrate the training and testing data for the $\tilde{\Lambda}$-regressor as blue and yellow points, respectively. Each of the points represents the data labels ($\mathcal{M}$ or $\tilde{\Lambda}$) of one simulated GW event. The distribution of the events across the parameters, $\mathcal{M}$ and $\tilde{\Lambda}$, is the same as on the left panel. The data segregation implemented in both panels shows a complete range of parameters

entirely stripped from the network during training process. These parameter values are only used for testing. This helps asses the robustness of the regressor. A good performance ensures that the network has the ability to extrapolate to regions that are obscured in the training procedure. The $\mathcal{M}$-regressor uses BNSM signals with $\mathcal{M} \in [1.45, 1.6]\ M_\odot$ as test data, amounting to 20,224 samples. The remaining data is used for training, making up a total of 54,776 samples. Therefore, of all the 75,000 generated events, we use 73% for training, and 27% for testing the $\mathcal{M}$-regressor. The $\tilde{\Lambda}$-regressor on the other hand is tested on data using BNSM signals generated with $\tilde{\Lambda} \in [400, 650]$. This amounts to 64,532 and 10,468 samples for training and testing, respectively. Both $\mathcal{M}$- and $\tilde{\Lambda}$-regressors are trained in batches of size 16, for 50 epochs each.

The trained $\mathcal{M}$- and $\tilde{\Lambda}$-regressors produce the results on the testing samples as depicted in the left and right panels of figure 5.10, respectively. We show the networks' predicted parameters plotted against the true labels. A 2D-histogram representation is used for both subfigures. The density of the points on the plots is given by the color map. A perfect prediction of the parameters would imply that all the points lie on the $y = x$ curve, illustrated by the diagonal black solid line. We observe from both the left and right panels that a high density of parameter values fall around the solid black line. The distribution of the data in the plots shows that the number of points decreases as we move away from the $y = x$ curve. The $\mathcal{M}$- and $\tilde{\Lambda}$-regressors produce $\mathcal{R}^2$ values of 0.95 and 0.94, respectively.

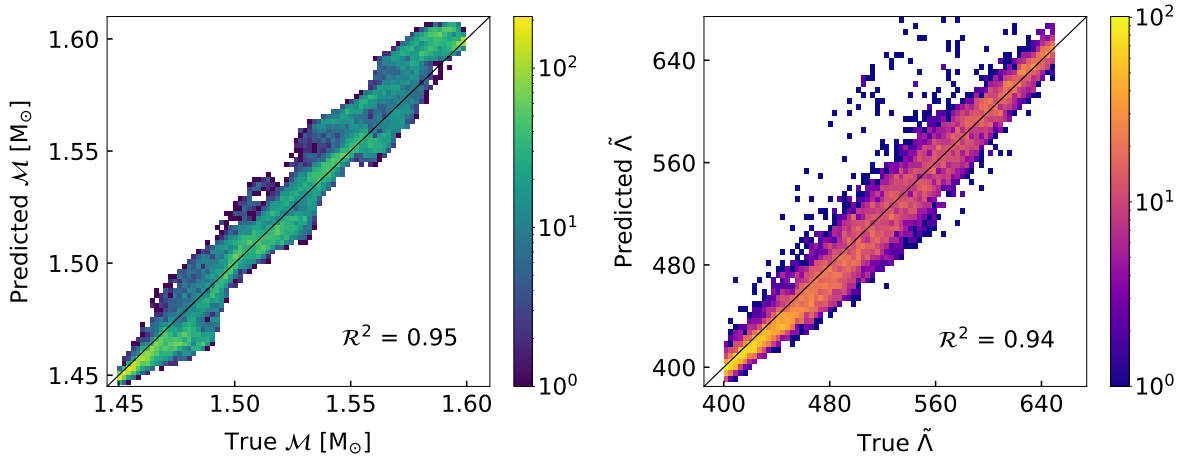

Figure 5.10: Predicted parameters plotted against true parameters for the case of independent parameter regression (without noise) for $\mathcal{M}$ (left panel) and $\tilde{\Lambda}$ (right panel). Both subfigures are shown as density plots. As seen from the distribution, the number of points becomes sparse as one moves away from the diagonal line. The total number of testing samples used for the independent $\mathcal{M}$- and $\tilde{\Lambda}$-regressors amount to 20,224 and 10,468, respectively. See Ref. [6] for original figure.

The corresponding learning curves are presented in figure 5.11. It can be observed that

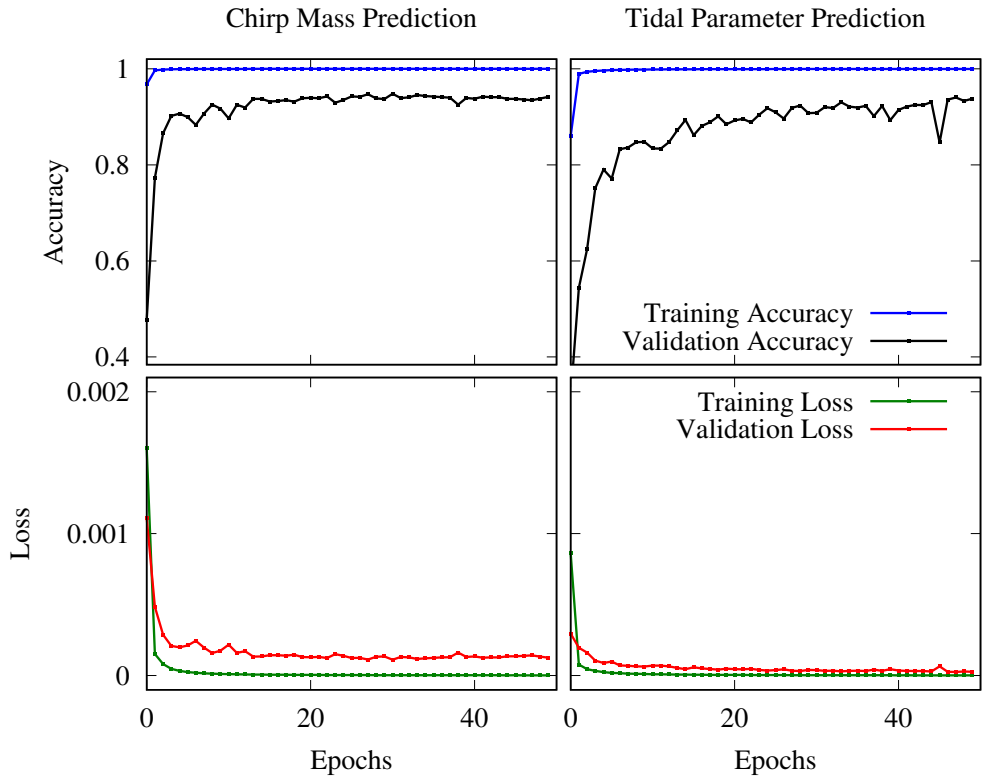

Figure 5.11: Learning curves of the $\mathcal{M}$- and $\tilde{\Lambda}$-regressors shown on the left and right panels, respectively. These curves depict that the regressors do not overfit to the training data.

the network does not overfit to the training data.

Simultaneous Parameter Regressors

Without Noise

Three distinct cases are explored for simultaneous regression without noise, Case I, II and III. Below, we discuss each case in detail and the modifications implemented from one case to the other. Note that while the training-testing data-segregation might differ from case to case, the distribution of all simulated GW events in the $\mathcal{M} - \tilde{\Lambda}$ space remains consistent with figure 5.9.

Case I: We train the weights of the regression network to output both $\mathcal{M}$ and $\tilde{\Lambda}$ simultaneously, using the same network structure as described for independent parameter regression. However, we make a slight modification to the last dense layer. Table 5.3 provides the network structure and layer dimensions. In this case, the network is no longer a sequential model. The last dense layer (layer index 10) branches into two dense layers (10a and 10b), with one layer outputting the chirp mass and the other layer the combined tidal deforma-

bility. Here, we apply the same values of $f_{\rm low} = 128$ Hz, $f_{\rm high} = 1024$ Hz, and $\Delta f = 1/16$ as used for independent parameter regression. The loss functions, activations, and normalization techniques for the input and labels remain the same as well. To evaluate the network's performance, we conduct a simple variation of data segregation. The training and testing data are illustrated in figure 5.12, as the blue and yellow points, respectively. The event distribution corresponds to the density plot shown in left panel of figure 5.9. For this particular test set, we consider $\mathcal{M} \in [1.3, 1.57]M_{\odot}$ and $\tilde{\Lambda} \in [250, 895]$. This testing data then consists of 19,582 samples, while the remaining 55,418 samples are utilized for training the network. We train the network for 100 epochs, using a batch size of 16.

Table 5.3: Example of a model architecture used for the simultaneous regression of chirp mass $\mathcal{M}$ and combined tidal deformability $\tilde{\Lambda}$ (Case I and Case II). The network comprises of 66,866 trainable parameters. See Ref. [6] for the original table.

Layer Index	Layer	Dimension
1	Input	14336 x 2
2	Convolution 1D	4769 x 16
3	Convolution 1D	1580 x 16
4	Max Pooling	790 x 16
5	Convolution 1D	256 x 32
6	Convolution 1D	78 x 32
7	Max Pooling	39 x 32
8	Convolution 1D	12 x 16
-	Reshape	192
9	Dense Layer	64
10a	Dense Layer	1
10b	Dense Layer	1

The simultaneous regressor, for this case, outputs the results which are presented in Figure 5.13. The left and right panels of the figure depicts the network's capability in retrieving $\mathcal{M}$, and $\tilde{\Lambda}$, respectively. It is noteworthy to mention that in this case and the subsequent ones, both parameters are estimated simultaneously using a single network. The trained network demonstrates high accuracy values for both parameters, with $\mathcal{R}^2 = 0.99$ for the $\mathcal{M}$-regression and $\mathcal{R}^2 = 0.95$ for the $\tilde{\Lambda}$-regression. This improved accuracy can be attributed to the fact that the network is exposed to a wide range of $\mathcal{M}$ and $\tilde{\Lambda}$ values during the training process, avoiding the complete absence of any specific parameter range. The learning curves for this network are shown in figure 5.14.

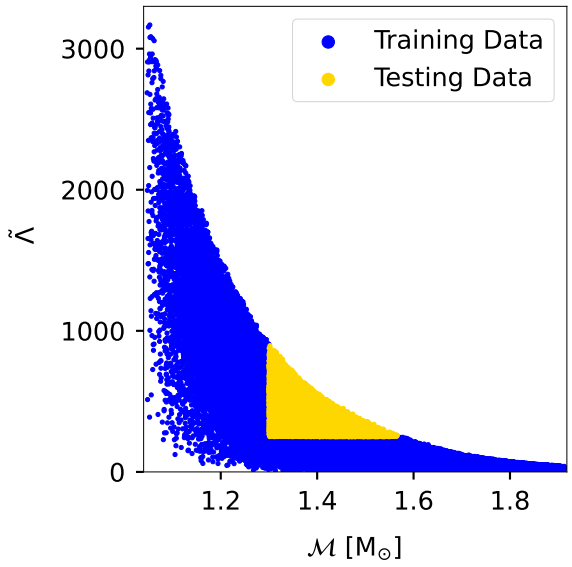

Figure 5.12: Individual training (blue) and testing (yellow) samples chosen for Case I of the simultaneous regression of $\mathcal{M}$ and $\tilde{\Lambda}$. The distribution of the points is depicted as a density plot in the left panel of figure 5.9. See Ref. [6] for original figure.

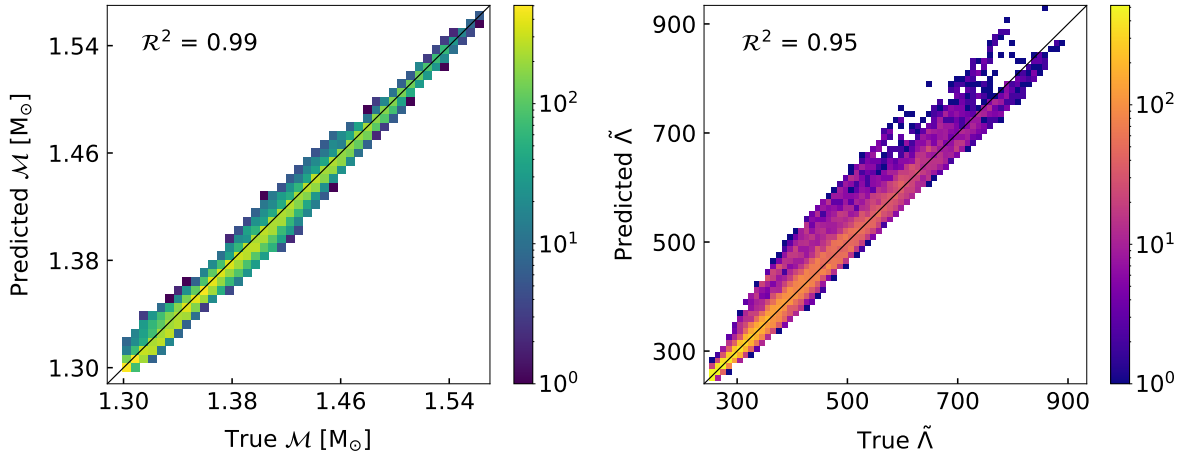

Figure 5.13: Density plots of the predicted parameters plotted against the true parameters in the case of simultaneous parameter regression (Case I, without noise) for $\mathcal{M}$ (left panel) and $\tilde{\Lambda}$ (right panel). Testing samples amount to 19,582 in this case. An improvement in accuracy from figure 5.10 is observed here due to the choice of test-train data segregation. In this case, the network is not completely obscure to an entire range of parameters during the training process, see figure 5.12. See Ref. [6] for original figure.

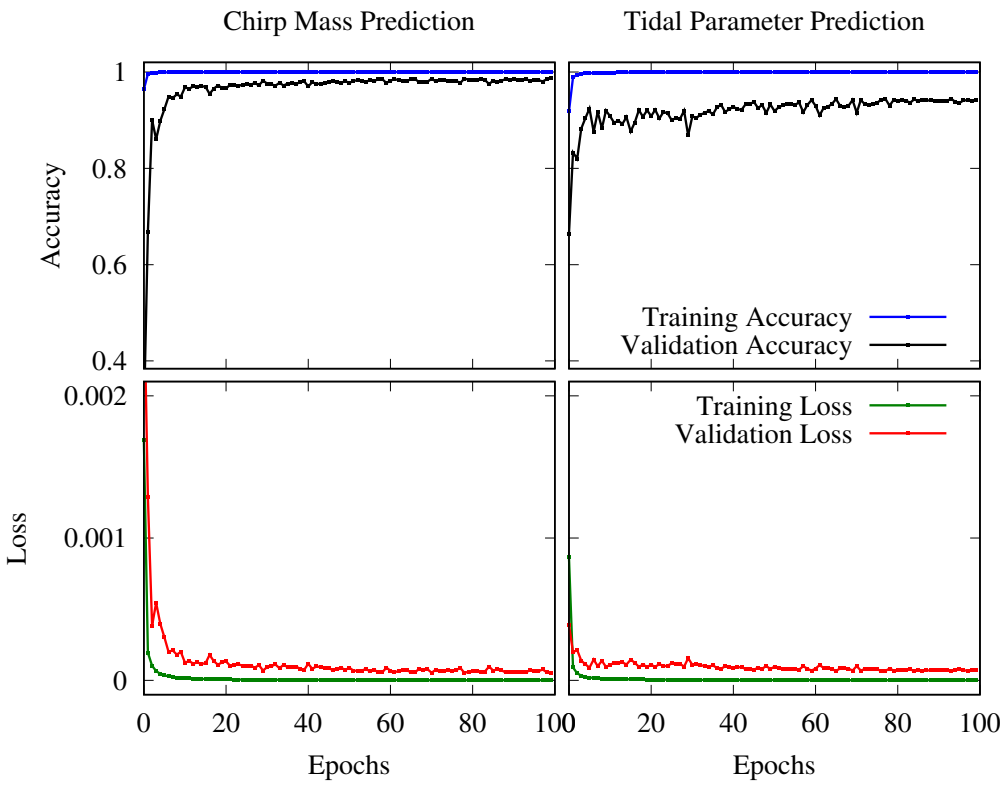

Figure 5.14: The learning curves of the Case I simultaneous parameter regressor for $\mathcal{M}$ (left) and $\tilde{\Lambda}$ (right) prove that no overfitting was encountered while training the network.

Case II: To enhance the robustness of the simultaneous parameter regressor, we introduce a generalized approach to segregate the training and testing data, building upon the previous case of independent parameter regression. For this, we provide the network with training and testing data as depicted in figure 5.15. The blue points represent the chosen training set, accounting for 33,686 waveform samples. In order to mask an entire range of both parameters, we exclude the points depicted in red from the training set. We choose the testing data such that the samples are completely exclusive of the parameter ranges used for training. Consequently, we obtain 18,718 testing samples marked in yellow. The red points are deemed redundant for our analysis. More specifically, gravitational wave (GW) signals generated from the parameter range $\mathcal{M} \in [1.4, 1.6]M_{\odot}$ and $\tilde{\Lambda} \in [150, 500]$ are used as test data. This results in a 64%-36% of training-testing samples. Considering the unusually large ratio of testing to training (~ 0.56) in this case, we train the network for 200 epochs. The longer training duration is necessary to accommodate the aforementioned ratio and ensure optimal performance.

The regressor is faced with new challenges when we eliminate an entire range of $\mathcal{M}$ and $\tilde{\Lambda}$ in the training process. The results in figure 5.16 display the true and predicted param-

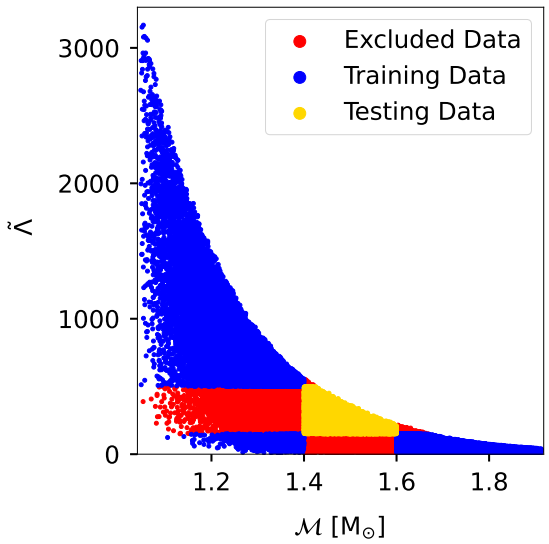

Figure 5.15: Individual training (blue) and testing (yellow) samples chosen for Case II of the simultaneous regression of $\mathcal{M}$ and $\tilde{\Lambda}$. The red data points in this figure are excluded in the training process in order to test the robustness of the regressor. See Ref. [6] for original figure.

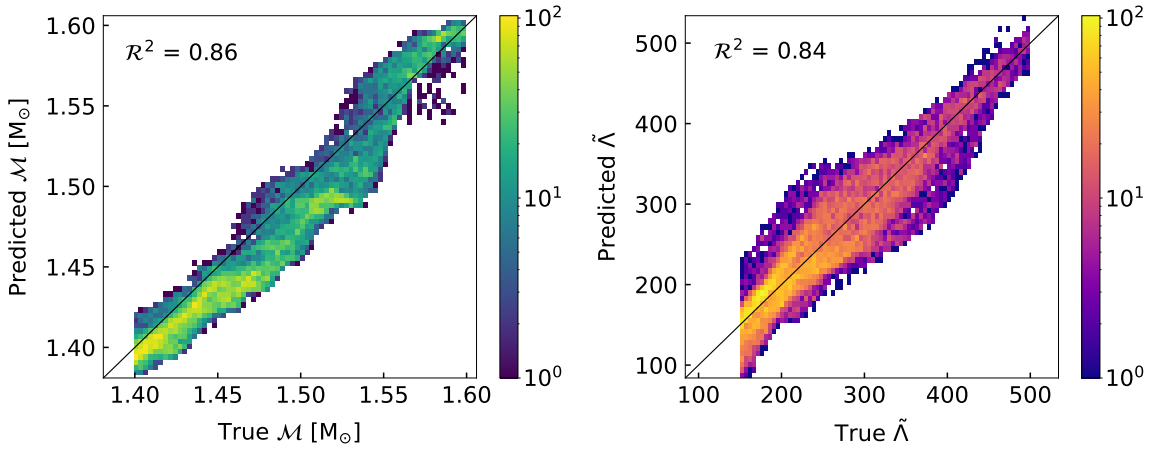

Figure 5.16: Density plots of the predicted parameters plotted against the true parameters for Case II of the simultaneous parameter regression. A total of 18,718 samples were used for testing. The regressor is faced with challenges when an entire parameter range is masked during the training process. Additionally, the ∼65%-35% splitting between the training and the testing data here, leads to considerably lower accuracies in this case. See Ref. [6] for original figure.

eters, presented as 2D-histograms. The network outputs $\mathcal{R}^2$= 0.86 for the $\mathcal{M}$-regression and $\mathcal{R}^2$= 0.84 for the $\tilde{\Lambda}$-regression. The decrease in $\mathcal{R}^2$ values for both parameters can be attributed to the challenges the network faces when it is blind to an entire range of parameters. The small number of training samples when compared to the testing samples (or the large value of test-to-train samples) also plays a role in the reduction of $\mathcal{R}^2$ when compared to Case I. The regressor requires longer training times in order to reach an optimum value of accuracy. This can be seen in figure 5.17, where the learning curves of the network for $\mathcal{M}$ and $\tilde{\Lambda}$ parameters are shown in the left and right panels, respectively.

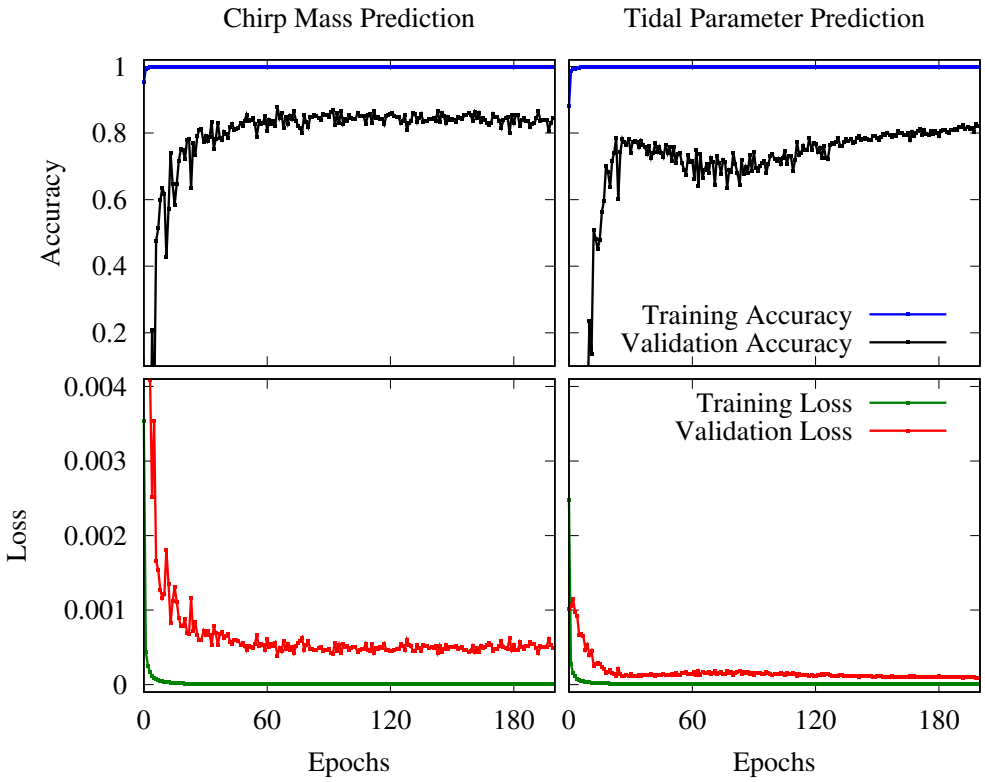

Figure 5.17: The learning curves of the simultaneous parameter regressor (Case II) for $\mathcal{M}$ (left) and $\tilde{\Lambda}$ (right), depict that the network does not overfit during the training.

Case III: In the third scenario, we further tune the hyperparameters of the network. Due to the highly non-linear dependence of $\tilde{\Lambda}$ on $\mathcal{M}$, we assess the simultaneous regression of the parameters by substituting the scale of tidal deformability labels from linear to log, i.e., we use $\log(\tilde{\Lambda})$ as the target label. In addition, we modify the input data format for training the network in this particular case. In contrast to the previous representation of the frequency-domain waveforms as amplitude and phase, here, we represent the input waveforms as real and imaginary parts. A finer resolution of the waveforms is implemented here, i.e., we choose a frequency range [23, 2048] Hz, with $\Delta f = 1/128$. Therefore, the input dimension is (259454, 2). For the normalization, we ensure that the integral sum of

the real and imaginary parts, are both of unit value. A rescaling is then applied to both channels, so that the values lie within the range [0,1]. Normalization of the tidal parameter is similar to Eq. (5.2), but in this case, $y = \log(\tilde{\Lambda}+1)$. Furthermore, we use average-pooling layers as opposed to max-pooling layers in the previous cases. The pooling layer is used for down-sampling, but an average-pooling layer is chosen to avoid loss of information on the finer representation used here. We further introduce an L2 regularizer to the convolutional kernels. The number of nodes in the dense layer (index 9 in table 5.3) is increased from 64 to 128. We modify the activation function of this particular dense layer to 'Scaled Exponential Linear Unit (SELU)'. Additionally, the activation functions acting on the rest of the convolutional layers are updated from 'ReLU' to 'Exponential Linear Unit (ELU)'. Due to the unusually large ratio of testing to training samples in Case II, we reduce the range of the testing samples in this case. We choose the ranges $\mathcal{M} \in [1.4, 1.55]\ M_\odot$ and $\tilde{\Lambda} \in [150, 400]$, as testing data. This results in 14,808 testing samples. The training and testing data are depicted in figure 5.18 [2]. This choice of testing-training data-segregation produces 41,588 training samples. This way, we improve the ratio of the testing-training samples to ~ 0.36.

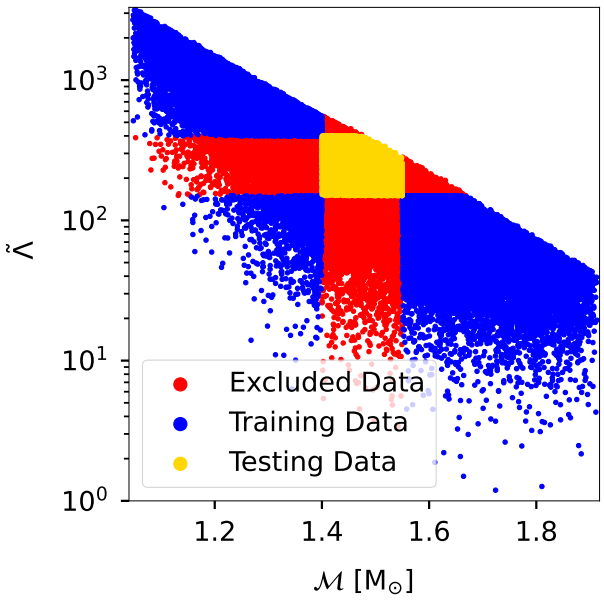

Figure 5.18: Training (blue) and testing data (yellow) for Case III of the simultaneous parameter regression. Note that we depict the tidal parameter as a log-scale here as the data labels for the tidal parameter are substituted with $\log(\tilde{\Lambda})$ in this case. See Ref. [6] for original figure.

The results illustrating the true and predicted parameters in this case are presented in figure 5.19. The range of both parameters that the network is uninformed about, is reduced

[2]Since the network is trained on $\log(\tilde{\Lambda})$ in Case III, we use a log-scale for $\tilde{\Lambda}$ in the figure

to $\mathcal{M} \in [1.4, 1.55]$ and $\tilde{\Lambda} \in [150, 400]$. Alongside several modifications to the regressor's hyperparameters, the update from $\tilde{\Lambda}$ labels to $\log(\tilde{\Lambda})$ leads to an improvement in the $\mathcal{R}^2$ values for both parameters. We achieve $\mathcal{R}^2 = 0.98$ for the $\mathcal{M}$-regression and $\mathcal{R}^2 = 0.88$ for the $\tilde{\Lambda}$-regression. The ratio of the test-train samples also helps producing more reliable results. The corresponding learning curves are depicted in figure 5.20. Once again, the learning curves prove that the model does not overfit to the data during the training process.

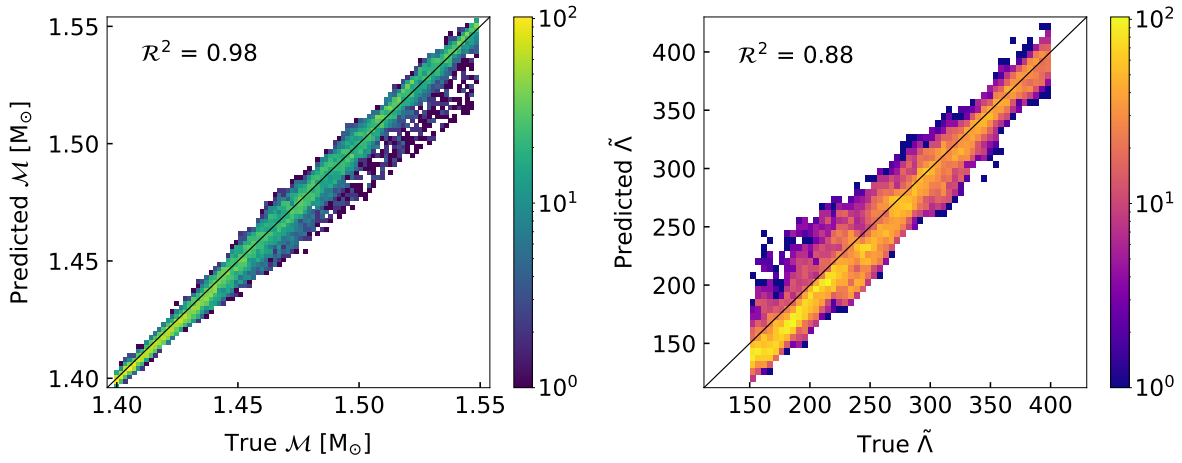

Figure 5.19: Density plots of the predicted versus true parameters for Case III of the simultaneous parameter regression. We use a total of 14,808 testing samples. The use of different scaling for combined tidal deformability, apart from the updates to certain hyperparameters in the training process, leads to an improvement in resulting accuracies (when compared to Case II, figure 5.16). See Ref. [6] for original figure.

With Noise

We further extend our work to incorporate noise. To achieve this, we first generate gravitational waveforms using aLIGO's detector noise, as described previously. These waveforms are then preprocessed (whitened) to ensure that the noise contributes equally to the overall signal at different frequency intervals. The data is fed to the network in the form of the real and imaginary parts of the frequency domain waveform. The waveforms are simulated with $f_{low} = 23$ Hz, $f_{high} = 2048$ Hz, and $\Delta f = 1/128$, leading to an input dimension of (259454, 2). The same activation functions as in Case III are used here, i.e., ELU is applied on the convolutional layers, and SELU on the first dense layer. Additionally, the activation function of the Dense layer, which outputs $\tilde{\Lambda}$, is altered from 'sigmoid' to 'tangent hyperbolic (tanh)'. The steep gradients are expected to increase the learning step size in the tidal deformability regression. The first step in normalizing the tidal parameter is the same as in Eq. (5.2), but in this case $y = \log(\tilde{\Lambda} + 1)$ is used. The second step involves scaling y_{norm} from [0, 1] to [-1, 1], which is a requirement for the 'tanh' activation function. Therefore,

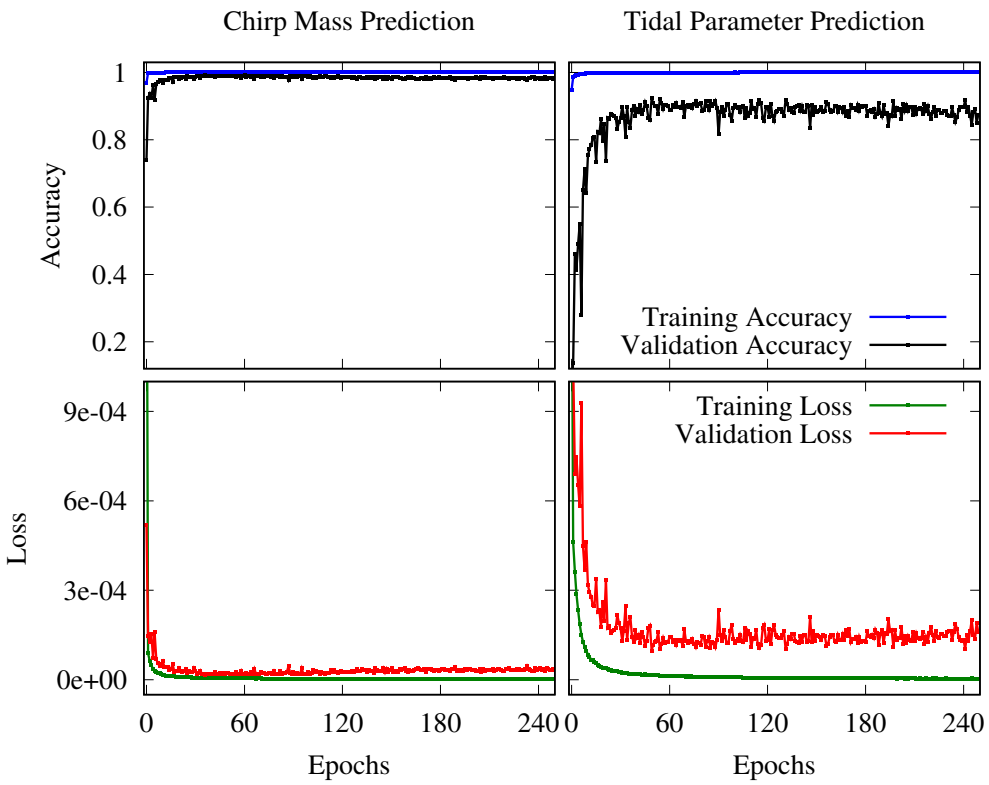

Figure 5.20: The learning curves for Case III of the simultaneous parameter regressor for $\mathcal{M}$ (left) and $\tilde{\Lambda}$ (right). No overfitting is encountered during the training process.

we further scale the output of Eq. (5.2) as $y'_{\rm norm} = 2(y_{\rm norm}) - 1$. In this case, we do not use a specific range of parameters for the testing set. We choose the training and testing data randomly, spread across the entire range of the simulated waveforms. We generate a total of 48,000 waveforms with noise, where 36,000 samples are randomly selected for training, and the remaining 12,000 samples for testing the network. We train the network in batches of 50 samples for 80 epochs.

The results of the trained regressor on waveforms that incorporate noise are shown in figure 5.21. The choice of random $\mathcal{M}$ and $\log(\tilde{\Lambda})$ values permits an easier prediction of the parameters. The inclusion of the complete range of parameters without withholding data from a specific plausible range of $\mathcal{M}$ and $\tilde{\Lambda}$, produces more accurate results. This is because the network has the capacity to interpolate between the points in the training data. The regressor achieves $\mathcal{R}^2$ values of 0.98 and 0.97 for the prediction of $\mathcal{M}$ and $\tilde{\Lambda}$ on the testing set, respectively. The learning curves corresponding to this regressor (with noise) are depicted in figure 5.22. The model does not overfit to the data during the training process, as seen from the figure.

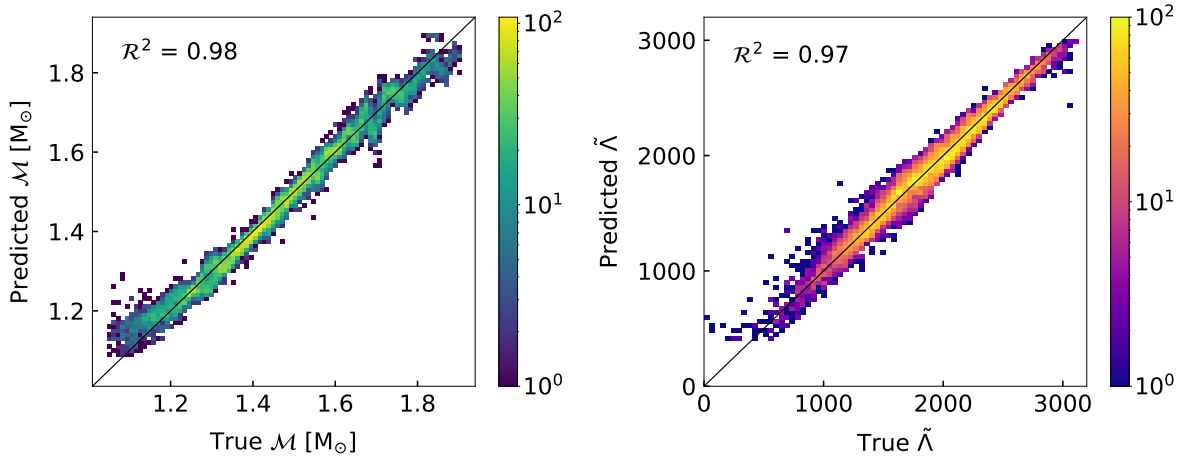

Figure 5.21: Density plots of the predicted parameters plotted against true parameters in the case of simultaneous parameter regression (with noise) for $\mathcal{M}$ (left panel) and $\tilde{\Lambda}$ (right panel). 12,000 samples were used for testing. In this case, we use a random segregation of testing-training data, spanning the entire range. The network therefore does not face challenges when dealing with testing data (as observed in figures 5.16 and 5.19, i.e., Case II and III of simultaneous parameter regression without noise). Hence, we obtain high accuracies for both parameters simultaneously in this case, despite the inclusion of noise. See Ref. [6] for original figure.

5.4 Discussions

To summarize, we have analyzed simulated GW data using deep neural networks. In particular, we develop DL algorithms for (a) signal-detection of BBHMs and BNSMs, and (b) estimation of mass and tidal deformability from BNSMs. The efficient inference of Λ is of specific interest here, as it carries valuable information on the NS EoS. This observable appears at late times during the inspiral where orbital frequencies are high. At such frequencies, the detector noise is comparatively large and hence, the inference of tidal deformability is considerably affected [102]. With the algorithm designed here, we show that NNs are capable of extracting the combined tidal parameter, $\tilde{\Lambda}$, of a BNS system from its inspiral-merger GW signal.

Several universal relations have been established with the tidal parameter and other quantities of interest. The famous $I - Love - Q$ relations [316], for example, allow for the inference of a NS's moment of inertia (I) as well as the quadrupole moment (Q), if $\tilde{\Lambda}$ is known. A universal relation between tidal deformability and the GW frequency at peak amplitude was established for equal mass binaries [317]. A similar correlation was found between tidal deformability and the dominant post-merger GW frequency for hadronic EoSs [318–321]. However, as demonstrated in Ref. [320], this universality can be broken when a phase transition from hadronic to deconfined quark matter takes place in the

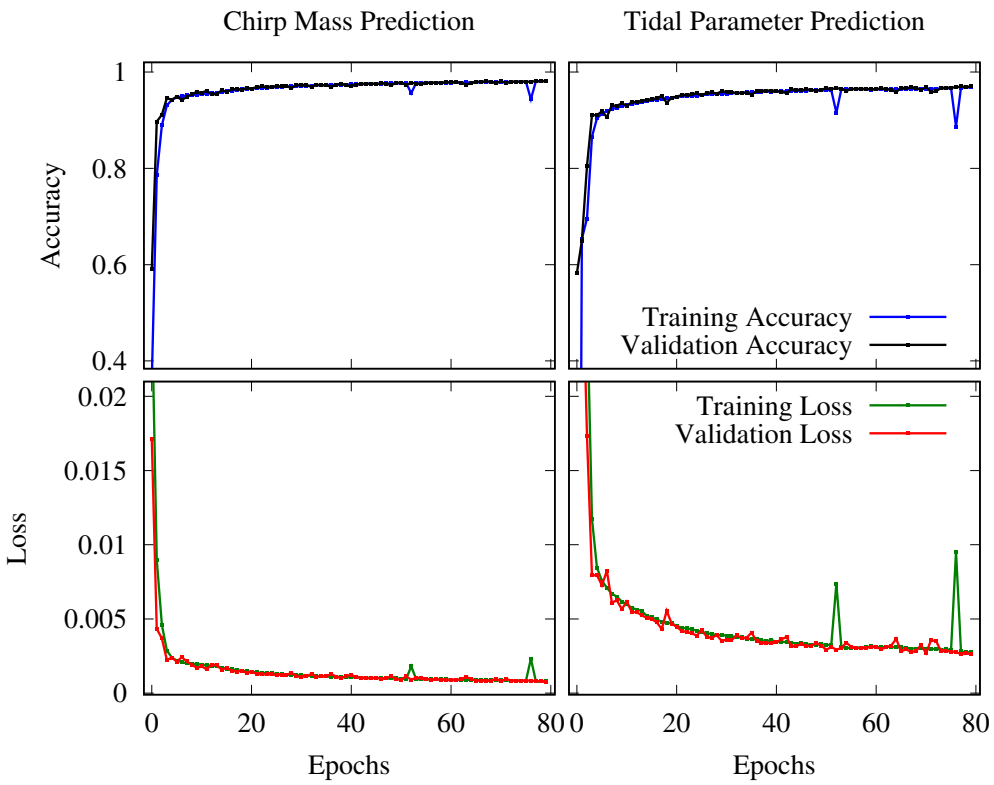

Figure 5.22: The learning curves for the simultaneous parameter regressor (with white noise) for $\mathcal{M}$ (left) and $\tilde{\Lambda}$ (right). The network does not encounter overfitting during the training process.

merger remnant. Therefore, evaluations of the tidal parameter from GW observations are necessary to enhance our knowledge of the dense matter EoS. With the ongoing development of current and next-generation GW detectors and the anticipated increase in their sensitivities, significantly more observations are expected in the near future [106, 107]. The analysis of these detections can benefit from the regressors developed in this work.

Chapter 6

BNSM Remnant: Studies using isentropic EoSs

In the previous chapter, we discussed the use of inspiral-merger gravitational waveforms of binary neutron star collisions for constraining the cold dense matter equation of state (EoS). In this chapter, we perform studies on the post-merger gravitational wave (GW) signal to extract information on the EoS at finite temperatures. The post-merger GW signal from the collision of binary neutron stars contains valuable information about the underlying EoS [318–321]. During the post-merger phase, the frequency of gravitational waves can be high, reaching a few kHz. However, the sensitivity of the current generation detectors such as Advanced LIGO and VIRGO is limited to lower frequencies. This explains the absence of any post-merger signal detection from the remnant of GW170817 [322]. Consequently, the electromagnetic observations that followed the GW event were crucial in providing the limited information we have about the merger remnant in GW170817. An intriguing finding from the electromagnetic counterpart of GW170817 was the observed kilonova ejecta [279, 323, 324]. From the amount of blue kilonova ejecta, the incompatibility with a prompt collapse to a black hole was deduced. Instead, three possible fates of the binary neutron star merger (BNSM) remnant were proposed: (i) a delayed collapse to a black hole, i.e., the remnant could be a hypermassive neutron star (HMNS) supported by differential rotation, which eventually collapses to a black hole after a short duration (approximately $1s$), (ii) a supramassive neutron star (SMNS), i.e., the remnant might be supported by rigid body rotation, surviving for a longer period before finally collapsing to form a black hole, or, (iii) alternatively, a permanently stable neutron star, i.e., the remnant could remain stable as a neutron star without collapsing. The actual fate of the remnant depends on both the maximum mass of a non-rotating neutron star and the threshold mass for prompt black hole formation. These factors play a role in determining which of the three possibilities mentioned above would occur. Predicting some properties of the remnant would be possible if it eventually settles into a state of uniform rotation. In previous works, the LIGO-VIRGO collaboration (LVC) estimated the moment

of inertia and maximum rotation rate of the uniformly rotating remnant based on a large set of zero-temperature EoSs [7]. However, it has been shown that mergers of binary neutron stars result in the formation of remnants characterized by extreme conditions of high temperature and density [325, 326]. Numerical relativity simulations performed by various research groups [268, 323, 327–330] have shed light on this phenomenon. According to these simulations, the maximum temperature within the remnant could reach ~ 70 MeV or even higher, while the maximum density could reach about 5 times that of normal nuclear matter density. Additionally, the simulations indicate that the entropy per baryon (s_B) within the central region of the remnant, shortly after the merger is $\lesssim 2k_B$ [331]. The bulk of the remnant, excluding the unshocked core, however, exhibits an entropy per baryon few times k_B [331]. Although there may be considerable variation in temperature and entropy initially, the conditions become homogeneous at a later time. The early evolution of the remnant is primarily driven by gravitational wave radiation, which dominates over a timescale of 10-20 ms. Subsequently, the evolution is influenced by viscosity over a period of about 100 ms and neutrino cooling over 2-3 s [325, 328, 332–334]. Furthermore, the magnetic field exerts a significant influence on the remnant's evolution. An effective viscosity is expected to be generated in the remnant through the magnetorotational instability (MRI). The other effect competing with the MRI is the magnetic winding during differential rotation of the remnant [332]. Both these effects are responsible for transporting angular momentum and eradicating the differential rotation, making the remnant a rigidly rotating body [330]. Given these observations, it would be intriguing to explore the thermal effects on the structures of uniformly rotating neutron stars and the associated Keplerian frequencies. Understanding the behavior of the remnant under these extreme conditions could offer valuable insights into the physics of neutron stars and help constrain the underlying EoS. The thermal effects on the remnant in BNSMs were earlier studied using EoSs at fixed temperatures [335]. Following the findings from the LVC on the uniformly rotating remnant [7], our motivation is to investigate the properties of the remnant in the numerical library LORENE [336] using EoSs at fixed entropy per baryon, as opposed to zero-temperature EoSs.

6.1 LORENE: Formulation and Implementation

Here, we discuss the formalism implemented in the numerical library LORENE [336]. We assume that the remnant in GW170817 becomes a rigidly rotating body once the differential rotation of the remnant is shed over the effective viscous and magnetic winding timescale ~ 100 ms. A stationary, axisymmetric spacetime is assumed for the study of this rigidly rotating remnant. Such rapidly rotating star models can be studied within general relativity in the 3+1 dimensional space plus time framework [337]. Here, the spacetime manifold is foliated into a family of nonintersecting space-like hypersurfaces Σ_t parameterized by coordinate time, t. By defining three spatial coordinates (x^i) on each hypersurface, one can write the line element in terms of lapse function N and shift vector (β^i) as,

$$ds^2 = -N^2 dt^2 + \gamma_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt) \ , \tag{6.1}$$

where γ_{ij} is the 3-metric on each Σ_t [337]. For this scenario, the coordinates are chosen based on considerations of spacetime symmetries and the foliation within the 3+1 framework. The underlying assumption here is the stationarity, axisymmetry, and asymptotic flatness of the spacetime. This implies that there are two mutually commutative Killing vector fields, $\mathbf{e}_0 = \partial/\partial t$ and $\mathbf{e}_3 = \partial/\partial\phi$ in the coordinates $(x^0 = t, x^1, x^2, x^3 = \phi)$. The x^1, x^2 coordinates are cosen as spherical coordinates, i.e., $x^1 = r$ and $x^2 = \theta$. In addition, $\beta^r = \beta^\theta = 0$ and $\gamma_{r\phi} = \gamma_{\theta\phi} = 0$. In a quasi-isotropic gauge, $\gamma_{r\theta}$ vanishes and the line element reduces to the form [147],

$$ds^2 = -N^2dt^2 + A^2(dr^2 + r^2d\theta^2) + B^2r^2\sin^2(d\phi - N^\phi dt)^2 \ , \tag{6.2}$$

where the metric functions, N, β^ϕ, A and B, depend on coordinates r and θ. Then, four gravitational field equations are obtained as a set of four coupled elliptic partial differential equations involving the energy-momentum tensor in source terms [337].

The matter is described by the energy-momentum tensor of a perfect fluid,

$$T^{\mu\nu} \ = \ (\varepsilon + P)u^\mu u^\nu + Pg^{\mu\nu} \ . \tag{6.3}$$

The fluid log-enthalpy is

$$H = \ln\left(\frac{\varepsilon + P}{nm_B}\right) \ , \tag{6.4}$$

where n and m_B are baryon density and rest mass, respectively.

The equation of the fluid equilibrium follows from the conservation of energy-momentum tensor

$$H(r,\theta) + \ln N - \ln\Gamma(r,\theta) = \frac{Te^{-H}}{m_B}\partial_i s - u_\phi u^t\partial_i\Omega, \tag{6.5}$$

where Γ is the Lorentz factor of the fluid with respect to the Eulerian observer, and s is the entropy per baryon in Boltzmann units. The last term of Eq. (6.5) is zero, as we consider only rigid rotation, i.e., $\Omega =$ constant. Therefore, it is shown that the equilibrium Eq. (6.5) finally reduces to the zero temperature expression [147],

$$H(r,\theta) + \ln N - \ln\Gamma(r,\theta) = \text{constant}. \tag{6.6}$$

The formulation described above is implemented in LORENE. We use different equations of state at constant entropy per baryon to calculate the properties of the rigidly rotating remnant in GW170817.

6.2 Thermal Effects on BNSM remnant

Here, we discuss finite temperature EoSs and their impact on the properties of the remnant. We present the plots depicting the EoSs, as well as the relationships between mass and radius for sequences of non-rotating neutron stars in figures 6.1 and 6.2, respectively. These

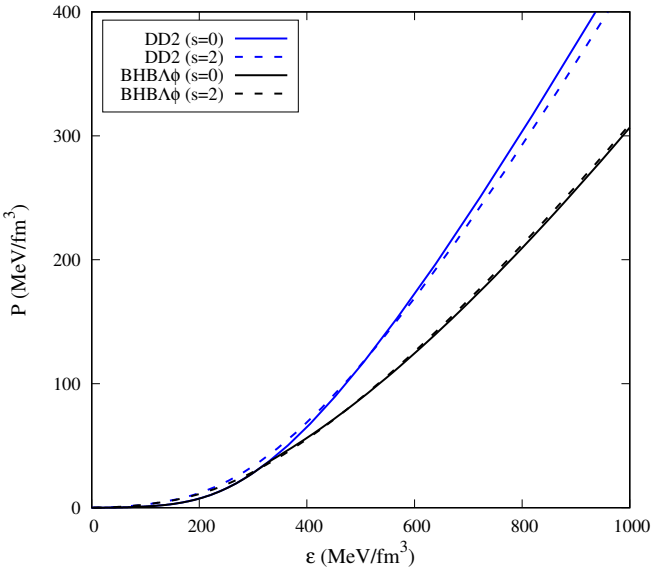

Figure 6.1: Pressure as a function of energy density for the DD2 (blue) and BHB$\Lambda\phi$ (black) EoSs are shown at entropy per baryon $s = 0$ and $s = 2$ as solid and dashed lines, respectively. Figure taken from Ref. [1] (DOI:10.3847/1538-4357/ab6a9e).

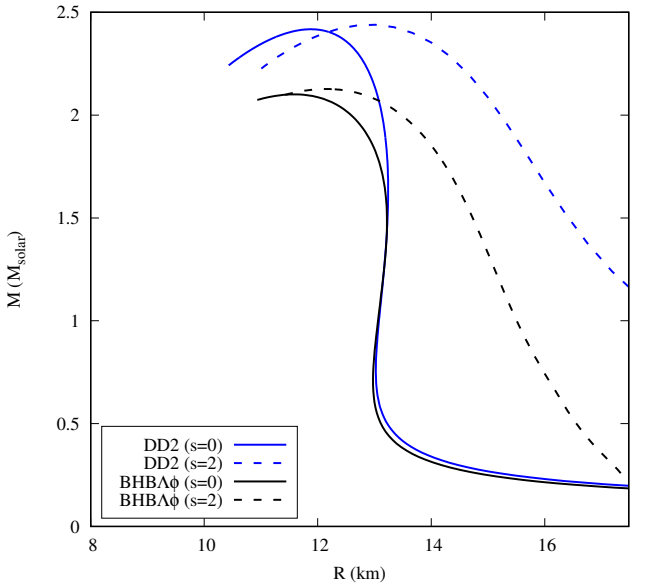

Figure 6.2: Mass–Radius ($M - R$) sequences of the DD2 (blue) and BHB$\Lambda\phi$ (black) EoSs are shown at entropy per baryon $s = 0$ and $s = 2$ as solid and dashed lines, respectively. The thermal effects produce much larger radii, compared to the higher maximum masses. Figure taken from Ref. [1] (DOI:10.3847/1538-4357/ab6a9e).

sequences correspond to the DD2 and BHB$\Lambda\phi$ EoSs under two different conditions, namely, at zero temperature ($s = 0$) and at entropy per baryon, $s = 2$.

It can be observed from figure 6.2 that the EoSs with $s = 2$ produce higher maximum masses of non-rotating neutron stars due to thermal pressure. The thermal effects on the maximum mass are however much smaller compared to the radii. The thermal effects produce much larger radii, and the radii are significantly pronounced for the DD2 EoS. This demonstrates the importance of thermal effects on non-rotating stars.

We further extend the study to explore the thermal effects on the properties of the remnant, a rapidly rotating neutron star. With the assumption that the remnant in GW170817 did not immediately collapse to a black hole, we compute the maximum rotation rate and moment of inertia using EoSs at non-zero temperatures. These values were obtained with the numerical library, LORENE [336], which is best suited for rapidly rotating compact stars at a fixed entropy per baryon. The remnant properties then include thermal effects as opposed to the calculations by the LVC [7] carried out using a large numbers of cold and β-equilibrated EoSs. Note that the rotation rate of the remnant could exceed the Keplerian limit of the uniformly rotating neutron star due to differential rotation. Nonetheless, the mass-shedding limit was set as the upper limit of the remnant's rotation [7]. The upper bound of the remnant's initial baryonic mass was estimated to be $3.05\mathrm{M}_\odot$ [7]. Therefore, in our analysis, we perform the calculations on the remnant with baryonic mass that does not exceed $\sim 3\mathrm{M}_\odot$ at the mass-shedding limit. Note: We find that the maximum masses at the Keplerian speed for some EoSs are much higher than $3.05\mathrm{M}_\odot$. For those EoSs, we restrict calculations to the baryonic mass of $3\mathrm{M}_\odot$. We compute the properties of the remnant using the microphysical EoSs described in chapter 2. More specifically, we use the DD2 [131], BHB$\Lambda\phi$ [144], SFHo [133], SFHx [133], TM1 [128], and TMA [129] EoSs for these calculations. The baryonic mass, moment of inertia, and Keplerian frequency for each EoS with a fixed entropy per baryon ($s = 2$) is listed in table 6.1.

These values are then compared to those obtained with zero temperature EoSs ($s = 0$). We find that the Keplerian frequencies at $s = 2$ for all EoSs listed in table 6.1 are appreciably lower than those with $s = 0$. As illustrated in figure 6.2, the impact of thermal effects on the radii of neutron stars is significant, i.e., the stars are bigger in size compared to their cold counterparts. This increase in size effectively compensates the increase in mass due to thermal effects, thereby keeping the total baryonic mass close to that of the cold remnant [335]. Additionally, this demonstrates that the Keplerian frequencies of the remnant using zero temperature EoSs are grossly overestimated in Ref. [7].

This work can also be compared to the calculations carried out by Kaplan et al. in Ref. [335], using EoSs at finite constant temperatures. The calculations in Ref. [335] were carried out using a numerical method, known as the Cook, Shapiro, and Teukolsky solver. We observe a consistency with Ref. [335] in the qualitative outcome of an extended remnant. A lower Keplerian frequency is observed in both works due to thermal effects.

In the next section, we present a few results on the binary component NSs in GW170817.

Table 6.1: Gravitational mass of the rigidly rotating remnant at Kepler Frequency, the corresponding baryonic mass and moment of inertia listed for several EoSs at entropy per baryon $s = 0$, and $s = 2$. These results show that the Keplerian frequencies at $s = 2$ for all EoSs are considerably lower than those with $s = 0$. See Ref. [1] for original publication.

EoS	$s = 0$				$s = 2$			
	M_G^{Rot} ($M_\odot$)	M_B^{Rot} ($M_\odot$)	I (10^{38} kg m^2)	$f_{\rm Kep}$ (kHz)	M_G^{Rot} ($M_\odot$)	M_B^{Rot} ($M_\odot$)	I (10^{38} kg m^2)	$f_{\rm Kep}$ (kHz)
DD2	2.606	3.004	5.439	1236	2.657	2.998	5.400	1109
BHB$\Lambda\phi$	2.525	2.914	4.204	1424	2.427	2.717	3.755	1269
SFHo	2.444	2.856	3.214	1763	2.447	2.807	3.346	1606
SFHx	2.556	3.000	4.051	1581	2.492	2.832	3.715	1425
TM1	2.623	3.003	5.241	1228	2.634	3.001	6.767	1011
TMA	2.439	2.785	4.191	1315	2.448	2.728	4.460	1099

6.3 Binary components in GW170817 at $s = 0$

In this section, we focus on calculating the properties of the binary components in GW170817 using zero temperature EoSs. The range of combined tidal deformability extracted from GW170817, i.e., $70 \leq \tilde{\Lambda} \leq 720$ [105], offers significant insights into the radii of the binary components in the binary neutron star merger [262, 283, 284, 338–340]. This follows from the known fact that tidal deformability is closely related to the radius of a neutron star, as evident from the expression,

$$\Lambda = \frac{2}{3} k_2 \Big(\frac{R}{M}\Big)^5 , \tag{6.7}$$

where k_2 is the Love number [200]. Several research groups have exploited the knowledge of the tidal deformability estimate from GW170817 to estimate the radii of the binary components [283, 284]. An analytical approach that relates the value of tidal deformability obtained from GW170817 to the radius of a 1.4 $M_\odot$ neutron star was introduced in Ref. [284]. We extend this analytical prescription to estimate the radius of neutron stars in the mass range $1.1M_\odot \lesssim M \lesssim 1.6M_\odot$, as described in our previous work [8]. Note that in this prescription, we assume that for a particular EoS, the change in radii of NSs in the mass range $1.1M_\odot \lesssim M \lesssim 1.6M_\odot$ is insignificant. This may change for EoSs with a strong first order phase transitions from hadronic to deconfined quark matter [141], but such EoSs are not studied in this chapter.

It is known that $k_2 \propto C$, where $C = M/R$ is defined as the compactness of the star. It then follows from Eq. (6.7) that $\Lambda \propto C^{-6}$, for a large collection of EoSs. This relation is plotted in Fig. 6.3 for the different EoSs introduced earlier in chapter 2. In this work, we apply

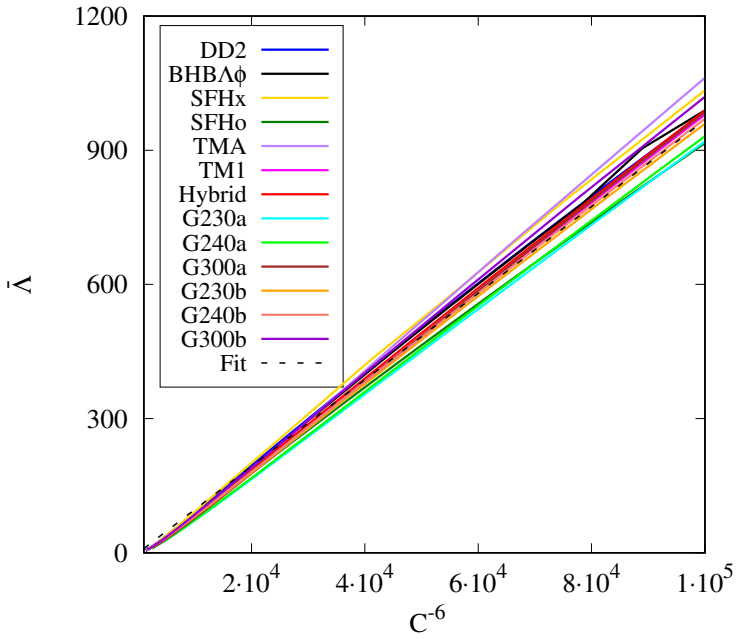

Figure 6.3: Dimensionless tidal deformability is plotted as a function of C^{-6}, for several EoSs. Here, C $(= M/R)$ is the compactness of the neutron star. The dashed black line depicts the fitting curve $(\Lambda = aC^{-6})$ to the EoSs. Figure taken from Ref. [1] (DOI:10.3847/1538-4357/ab6a9e).

slight modifications to the methods described in Ref. [8]. This is done by introducing a larger set of EoSs to estimate the fit parameter, a, that is used to fit the relation, $\Lambda = aC^{-6}$. Using the larger set of EoSs as shown in Fig. 6.3, we obtain an estimate of $a \approxeq 0.00967$.

The combined tidal deformability is defined as

$$\tilde{\Lambda} = \frac{16}{13}\frac{[(m_1 + 12m_2)m_1^4\Lambda_1 + (m_2 + 12m_1)m_2^4\Lambda_2]}{(m_1 + m_2)^5} \ , \tag{6.8}$$

where Λ_1 and Λ_2 are the dimensionless tidal deformabilities of the binary neutron star components, m_1 and m_2, respectively. Using the fit relation, $\Lambda = aC^{-6}$ in Eq. (6.8), and assuming that the radii of binary components in the mass range $1.1M_\odot \lesssim M \lesssim 1.6M_\odot$ are nearly equal, i.e., $R_1 \simeq R_2 \simeq R$, we obtain,

$$\tilde{\Lambda} = \frac{16a}{13} \times \frac{1}{(1+q)^5} \times \left(\frac{R}{m_1}\right)^6 \left\{\frac{(1+12q)}{m_1^6} + \frac{(q+12)}{q^2 m_1^6}\right\}, \tag{6.9}$$

where q is the mass ratio of the binary components. Rewriting the expression in terms of chirp mass yields the expression,

$$\tilde{\Lambda} = \frac{16a}{13} \times \left(\frac{R}{\mathcal{M}}\right)^6 \times \frac{q^{8/5}}{(1+q)^{26/5}} \times \left[12 - 11q + 12q^2\right] \ . \tag{6.10}$$

Since it was noted from numerical simulations that the combined tidal deformability had no dependence on the mass ratio, q [99,341,342], we investigate the extrema of $\tilde{\Lambda}$ by taking the derivative with respect to q at a fixed chirp mass and obtain,

$$\Big(\frac{\partial\tilde{\Lambda}}{\partial q}\Big)_{\mathcal{M}} = \tilde{\Lambda} \times \frac{1-q}{5q(1+q)} \times \Big[\frac{96-263q+96q^2}{12-11q+12q^2}\Big] \ . \tag{6.11}$$

The roots of this derivative are then obtained as, $q = 1$ and $q = 0.43346$. Following the same procedure in Ref. [8], we consider $m_2^{\rm min} \sim 1M_\odot$ and $m_1^{\rm max} \sim 2M_\odot$. This implies a mass ratio, $q \geq 0.5$. Moreover, the mass ratio in GW170817 was estimated as $q \geq 0.7$ [104]. Rewriting Eq. (6.10) as,

$$\tilde{\Lambda} = a'\Big(\frac{R}{\mathcal{M}}\Big)^6 \ . \tag{6.12}$$

where,

$$a' = a \times \frac{16}{13} \times \frac{q^{8/5}}{(1+q)^{26/5}} \times \big[12-11q+12q^2\big], \tag{6.13}$$

we estimate the value of a' for the range of q estimated from GW170817. For $q = 0.7$, we obtain $a' = 0.0043$ and $a' = 0.0042$ for $q = 1$, further proving the independence of $\tilde{\Lambda}$ on mass ratio.

It may therefore be concluded that the effective dependence of the combined dimensionless tidal deformability, $\tilde{\Lambda}$, on chirp mass, $\mathcal{M}$, is similar to that of Λ on M, i.e., as given by Eq. (6.12). Based on this relation, we can provide estimates on the radii of the binary masses involved in GW170817. The components were estimated to have masses in the range $1.17 - 1.6M_\odot$ [262], but the radii are assumed to have negligible differences. Substituting the chirp mass, $\mathcal{M} = 1.188$ for GW170817 [262], we obtain,

$$R = 4.36 \times \tilde{\Lambda}^{1/6} \ . \tag{6.14}$$

The upper bound on $\tilde{\Lambda} = 720$, results in a radius estimate of ~ 13.04 km. Similarly, the lower bound of $\tilde{\Lambda} = 70$, provides us with a radius estimate of 8.85 km. Furthermore, if we assume the lower limit obtained from electromagnetic observation, $\tilde{\Lambda} = 197$ [315], we obtain a radius of 10.52 km.

Apart from the analytical relation described above, we propose an alternative approach that exploits the information on tidal deformability, to constrain the radii of the binary components in GW170817. This method follows from Ref. [8], where improvisations are implemented by using a larger set of EoSs. For this method, we assume a scenario where $m_1 = 1.55M_\odot$, and $m_2 = 1.2M_\odot$, satisfying $\mathcal{M} = 1.18M_\odot$. This way, we obtain the radii of the stars with a mass difference greater than $0.3M_\odot$ (a rather asymmetric mass ratio is chosen here to obtain values close to the minimum and maximum possible radii of the binary components in GW170817). Using several different EoSs, we plot the combined tidal deformabilities and radii of the primary and secondary neutron stars in the right and left panels of figure 6.4.

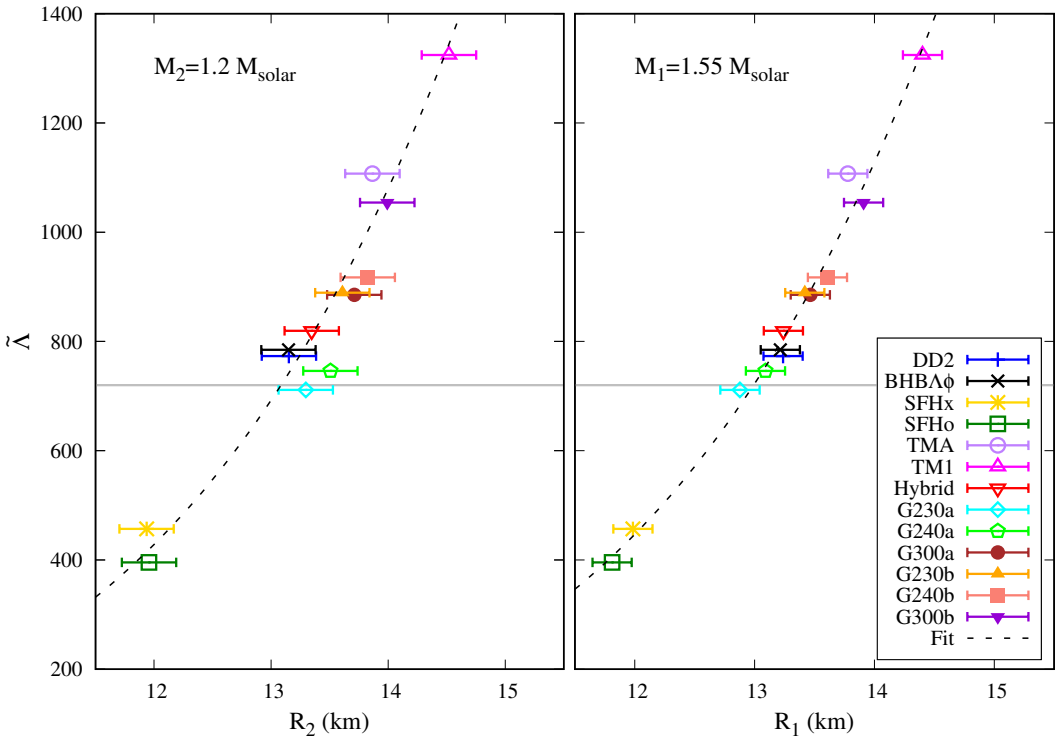

Figure 6.4: Combined tidal deformability, $\tilde{\Lambda}$, as a function of radius, for the primary mass ($m_1 = 1.55M_\odot$) and secondary mass ($m_2 = 1.2M_\odot$) in the right and left panels, respectively. The symbols with colors correspond to different EoSs. The horizontal grey lines represent the curve $\tilde{\Lambda} = 720$, i.e., the upper bound obtained from GW170817 [272]. The dashed black lines are fit functions ($\tilde{\Lambda} \propto R^6$) to the data. Figure taken from Ref. [1] (DOI:10.3847/1538-4357/ab6a9e).

We fit the points using the relation $\tilde{\Lambda} \propto R^6$, as seen from Eq. (6.12). This was seen as a better fit as opposed to $\tilde{\Lambda} \propto R^5$ from Ref. [8], where a smaller sample size of EoSs was used. The fitting curves are depicted as dashed black lines. We plot the error bars which are estimated with respect to the fitting curve. The upper bound on $\tilde{\Lambda}$, obtained from GW170817 [272], is shown as a grey horizontal line. With this upper bound, we are able to constrain the radii of $1.2M_\odot$ and $1.55M_\odot$ neutron stars at a maximum of 13.08 km, and 12.99 km, respectively. These estimates are in good agreement with the values obtained from the analytical relation obtained in Eq. (6.14).

6.4 Conclusions

In conclusion, the post-merger gravitational wave signal provides valuable insights into the EoS of neutron star matter, but challenges in detecting high-frequency signals limit our understanding. By considering a uniformly rotating object, we have made predictions for

certain properties of the remnant, therefore advancing our understanding of these enigmatic cosmic phenomena. Electromagnetic observations also provide crucial information about the merger remnant, which opens up intriguing possibilities regarding its fate based on its mass and the maximum possible mass of a TOV NS. While there have been several studies on how high the maximum mass of non-rotating NSs could be [252–255], this issue has not been settled.

In particular, we investigated the moment of inertia and Kepler frequency of the rigidly rotating remnant of GW170817 using EoSs at fixed entropy per baryon, $s = 2$. These values were compared with those calculated at zero temperature i.e., $s = 0$. As estimated in Ref. [7], we assumed a baryonic mass of the remnant $\lesssim 3M_\odot$. Furthermore, we assumed that the remnant is rigidly rotating at the mass-shedding limit after the differential rotation is eased out due to viscous effect. We observe that the thermal effects have negligible impact on the remnant mass. However, the radius of the remnant increases significantly at finite entropy. Consequently, we find that the Kepler frequencies for the sample EoSs with $s = 2$ are significantly lower than those with $s = 0$. We conclude that Kepler frequencies calculated with zero temperature EoSs are grossly overestimated in Ref. [7]. One can further estimate the properties of the hypermassive neutron star before it settles into uniform rotation as done in Refs. [343, 344].

Furthermore, we study the properties of the binary components in GW170817 before the merger, using zero temperature EoSs. Using the analytic relations from Ref. [8], we increase the EoS sample size to provide better estimates on the radii of the binary components in GW170817. With the knowledge of the effective tidal deformability, we provide upper and lower bounds for the radii estimates as ~ 13 and ~ 8.85 km, respectively. Using improved fit relations on $R - \tilde{\Lambda}$, we also extract the radii of $1.2M_\odot$ and $1.55M_\odot$ NSs as ~ 13.08 and ~ 12.99 km, respectively. The choice of these masses reflects a possible combination of m_1 and m_2 for GW170817 that is rather asymmetric, therefore providing us with estimates of the minimum and maximum possible radii values of the binary components.

Chapter 7

Summary and Outlook

In this thesis we focus on developing deep learning (DL) algorithms to constrain the equation of state (EoS) of dense matter. These algorithms rely on observations of neutron stars (NSs), like masses (M), radii (R) and tidal deformabilities (Λ). We also provide an overview of how the developed algorithms in this thesis can be employed in future research endeavours. We discuss possible extensions of these methods and scopes for improvement.

In chapter 4, we begin by presenting the novel physics-based DL method designed to reconstruct the EoS of strongly interacting dense matter based on $M-R$ observations of NSs. This unique method employs a model-independent and flexible representation of the dense matter EoS, with neural networks (NNs) that output individual pressure (P) points as a function of density (ρ), i.e., $P(\rho)$. The algorithm incorporates a Bayesian perspective, in optimizing the EoS curve and calculating the associated uncertainties via importance sampling. This way, the computational inefficiencies related to Bayesian analyses in a multi-dimensional parameter space can be evaded. We introduce two NNs, namely the `EoS Network` and the `TOV-Solver Network`. The `EoS Network` is used to represent the EoS, in a model-independent way. The `TOV-Solver Network`, as the name suggests is trained to solve the Tolman–Oppenheimer–Volkoff (TOV) equations using the traditional supervised learning techniques. The EoS Network is then combined with the pre-trained `TOV-Solver Network` to form the basis of the designed pipeline. A gradient-based approach is implemented in the Automatic Differentiation framework to optimize the weights of the EoS Network such that the resulting EoS produces an $M-R$ curve (through the `TOV-Solver Network`) that best fits the observation, in an unsupervised manner. We perform several tests on simulated mock data, i.e., from piece-wise polytropic EoSs as well as from microphysical EoSs. We show that the performance of the algorithm improves as noise levels in the mock data are reduced. A higher precision on future measurements of NS global properties with next-generation telescopes can therefore provide the scope for a fine reconstruction of the EoS of dense matter. We apply the developed scheme on available $M-R$ data of NSs from different instruments and present the reconstructed EoS in the density range (1-7ρ_0). The results are compatible with the EoSs reconstructed from earlier works

that used conventional methods, and are also in agreement with the limits of tidal deformability obtained from the gravitational wave event, GW170817. Additional priors can be added to the EoS from perturbative Quantum Chromodynamics (pQCD) constraints. It has been shown that pQCD constraints the EoS from $2n_s$ and at $5n_s$ excludes at least 65% of the otherwise allowed area in the $\epsilon - P$ space [345]. Furthermore, the upcoming eXTP mission [346] and the SKA telescope [347] are presumed to observe a large number of neutron stars providing valuable information on the underlying EoS. It is also important to note that the data used in this chapter to reconstruct the dense matter EoS is restricted to mass and radius observations of NSs. With a rise in the number of GW detections from BNS mergers, this study can be extended to incorporate the $M - \Lambda$ relationships of NSs in addition to the $M - R$ relations. Current estimates of Λ from GW170817 bear larger uncertainties. However, improved measurements of tidal deformability are expected in future observations of GWs from next generation GW detectors, like LIGO India, KAGRA, Einstein Telescope and Cosmic Explorer, which will have better sensitivities. Another aspect that can be covered in future is the explicit inclusion of EoSs that undergo phase transitions from hadronic to deconfined quark matter in the data. Moreover, the current framework can be upgraded to include Bayesian Neural Networks, and also restrict the usage of emulators for solving the TOV-equations for higher accuracies. Subsequently, one would have to calculate the linear response of an $M - R$ curve to a change in the EoS, and use this gradient in physics-informed neural networks. Any bias introduced by the `TOV-Solver Network` can also be eradicated by using the TOV-equations instead. The scheme devised here can be generalized to and implemented in a number of fields that face challenges with inverse-problems.

We demonstrate the capability and performance of neural networks for the analysis of simulated gravitational wave (GW) data in chapter 5, by performing two distinct tasks: (1) classification of GW signals from binary black hole (BBH) mergers, from binary neutron star (BNS) mergers and signals which contain only noise, and (2) regression of chirp mass ($\mathcal{M}$) and the combined tidal deformability ($\tilde{\Lambda}$) from simulated GW signals of binary neutron star (BNS) mergers. The DL classifier developed in this chapter is capable of identifying the different kinds of signals with macro-averaged precisions of 0.99 and 0.96 for oSNR=25 and 20, respectively. In addition to the classifier, we design a few DL regressors to estimate the $\mathcal{M}$ and $\tilde{\Lambda}$ from BNS merger signals. We show that the network achieves high $\mathcal{R}^2$ values, 0.98 and 0.97, for regressing $\mathcal{M}$ and $\tilde{\Lambda}$, respectively, using test samples of whitened simulated BNS merger signals injected into aLIGO's colored noise. Note that the prediction of the tidal parameter is vital for constraining the underlying NS EoS. In this work, however, the regression network does not output the individual tidal deformabilities. Nor does it unravel the individual masses in the binary system from the estimation of $\mathcal{M}$. Hence, we assume a currently achievable scenario, i.e., where we are likely to obtain estimates only on the $\mathcal{M}$ and $\tilde{\Lambda}$ values to constrain the dense matter NS EoS. We show that with sufficient number of observations, this information can also add constraints on the NS EoS. The classifiers and regressors described here can be integrated to establish a comprehensive pipeline for future GW signal analyses. Upon detecting a

BNS merger using the classification network, the identified event can be passed on to the regression network for parameter estimation. This seamless integration forms a holistic pipeline for GW analysis in future, where GW detectors are expected to have sensitivities that are an order of magnitude better than those from the current generation. Another aspect to note is that the framework developed in this work aims to constrain the NS EoS using only mass and tidal parameters. However, unlike the one-to-one correspondence between the mass-radius ($M - R$) curve and the underlying EoS, the $M - \Lambda$ curves do not necessarily have a direct relation with the EoS [348, 349]. This has been demonstrated with different EoS models that undergo a first-order phase transition (FOPT) at significantly different densities, but result in identical $M - \Lambda$ curves. The effect of this degeneracy can be scrutinized in future work. Independent radius measurements from NICER could also help break this ambiguity. Furthermore, the classification and regression networks designed in this work are elementary. They can be further developed to incorporate additional parameters of the binary coalescing system like spins, inclinations and distance. It's essential to note that this work is based on a specific waveform approximant, i.e., 'IMRPhenomPv2_NRTidalv2' for BNS merger simulations, which could introduce model dependence into the trained network. Full general relativistic magnetohydrodynamic (GRMHD) simulations of binary neutron star mergers in 3D are considerably more accurate for modelling gravitational waveforms. However, a single simulation increases the computational cost by several orders of magnitude. Therefore, full GRMHD simulations are currently a less favourable alternative, albeit being more accurate and model-independent. If more computational power is available in future, one could replace the approximate waveform models with full GRMHD simulations. Another crucial open question which can be considered in this study is again the possibility of a phase transition from hadronic to deconfined quark matter at high densities and temperatures. Mergers of binary neutron stars can potentially harbour such extreme conditions, rendering possibilities to study these effects [268, 320, 350]. Despite employing a waveform approximant that models the inspiral, merger, and ringdown components of a GW event, we limit our analyses in this study to a maximum GW frequency of 2048 Hz. This restricted frequency range excludes information from the ringdown or post-merger phase of a BNS merger. In order to address this, the current work can be extended to include post-merger frequencies that reach as high as 4096 Hz. A comprehensive and extensive analysis of the post-merger GW signals of BNS mergers can enhance our understanding of the possible existence of a first-order phase transition. Additionally, post-merger GW analyses could benefit from deep learning techniques with exponential boosts in the associated computational costs. Similar algorithms might also prove useful in analyzing GWs from detectors like the upcoming space probes, LISA (Laser Interferometer Space Antenna) [351] and TianQin [352]. These missions are designed to detect gravitational waves in the low-frequency range. The gravitational wave sources that LISA would discover include ultra-compact binaries in our Galaxy, supermassive black hole mergers, and extreme mass ratio inspirals. TainQin on the other hand, is used to detect GWs in the mHz range from known sources. This aspect can also be explored in future.

In chapter 6, we demonstrate the effects of finite temperature on BNS merger remnants. We use different isentropic EoSs (fixed entropy per baryon, $s = 2$) to investigate the moment of inertia and Kepler frequency of the rigidly rotating remnant in the event, GW170817. When compared to the same quantities calculated at zero temperature i.e., $s = 0$, we note a significant decrease in the Kepler frequencies. We conclude that Kepler frequencies calculated with zero temperature EoSs are grossly overestimated in Ref. [7]. Additionally, we provide radii estimates of the binary components in GW170817. We utilize analytic relations from Ref. [8] with an increased sample size of zero temperature EoSs. Using the effective tidal deformability estimated for the event, we calculate upper and lower bounds as $\sim$ 13 km and $\sim$ 8.85 km, respectively, for the radii of the components. Furthermore, we employ enhanced fit relations for $R - \tilde{\Lambda}$ to derive the radii of $1.2M_{\odot}$ and $1.55M_{\odot}$ NSs as $\sim$ 13.08 km and $\sim$ 12.99 km, respectively. Note that we provide estimates for the radii of the binary components in GW170817 by making assumptions that the radii of the NSs in the mass range $1.17 - 1.6\mathrm{M}_{\odot}$ have insignificant differences. However, this assumption is likely to break down in situations where strong first order phase transitions occur. The fit relations for $R - \tilde{\Lambda}$ employed in this study are based majorly on nucleonic EoSs. This implies that the study is possibly invalid for hybrid EoSs and remains open for further investigation.

In summary, this thesis presents novel deep learning methods that utilize observational data of neutron stars to add constraints on the underlying equation of state. These algorithms will prove useful in future, as more observational data is available with time.

Appendix A

Automatic Differentiation on a Test function

We demonstrate the efficiency of the Automatic Differentiation (AD) technique along with the importance sampling and the uncertainty estimation used in chapter 4, on a simple test case. We provide an example using a test function, which is a second-degree polynomial, given as,

$$y = a + b_1 x_1 + b_2 x_2^2. \tag{A.1}$$

In this example, we aim to infer the coefficients (a, b_1, b_2) of the polynomial using 18 data samples (similar to the number of $M - R$ observations used in chapter 4). The samples are created by using random values for x_1 and x_2 from a univariate normal (Gaussian) distribution. These data points are analogous to the mean values of the $M - R$ data. For each of these 18 data points, we generate 1000 samples from fixed normal distribution that resembles its noise or uncertainty.

With this simple example, we infer the coefficients of the polynomial using the concept of automatic differentiation (see section 4.1 of chapter 4 for a detailed description). In addition, we use importance sampling [258] to assign proper weights to the inferred values corresponding to each of the 1000 samples generated from the noise distribution (see section 4.3.3 on uncertainty estimation). These results are then compared to the posterior distributions of the coefficients obtained using the Bayesian Inference (BI). The comparison is depicted in figure A.1, for the polynomial coefficients, a, b_1, and b_2 in red, green and blue, respectively. It can be observed that the posterior distributions obtained using the two different inference methods are rather similar. This simple test can be used to demonstrate the utility of the uncertainty estimation used on the Automatic differentiation (AD) inference here, i.e., importance sampling.

Note that importance sampling has also been successfully used in the inference of parameters from gravitational wave analyses that utilize deep learning techniques [296]. This method can therefore be applied to all techniques that utilize 'proposal posterior distri-

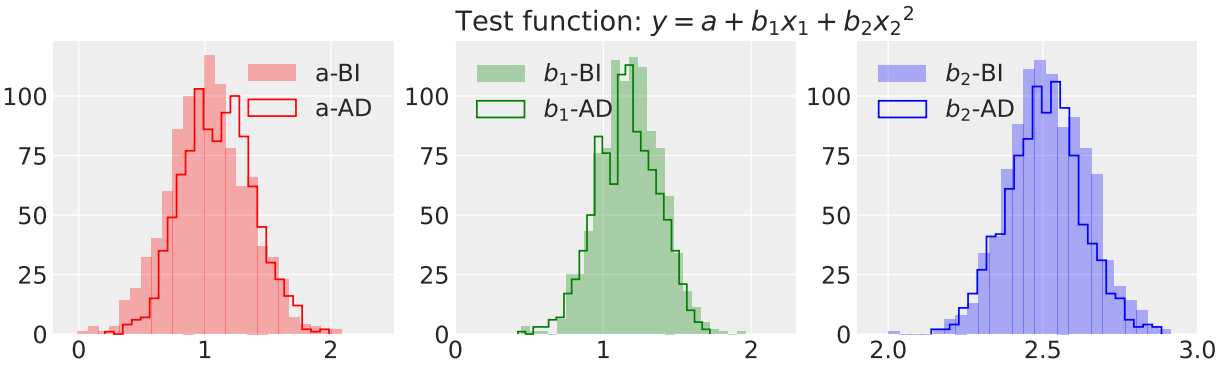

Figure A.1: Posterior distributions obtained for the coefficients of a polynomial using (i) the automatic differentiation method with importance sampling, denoted AD and (ii) the Bayesian Inference (BI) method.

butions' rather than using the exact ones that are achieved via Bayesian inference. An application of IS to the results from chapter 5 is left for future work.

Bibliography

[1] S. Soma and D. Bandyopadhyay, "Properties of binary components and remnant in GW170817 using equations of state in finite temperature field theory models," *Astrophys. J.*, vol. 890, p. 139, 2020.

[2] S. Soma, L. Wang, S. Shi, H. Stöcker, and K. Zhou, "Neural network reconstruction of the dense matter equation of state from neutron star observables," *JCAP*, vol. 08, p. 071, 2022.

[3] S. Soma, L. Wang, S. Shi, H. Stöcker, and K. Zhou, "Reconstructing the neutron star equation of state from observational data via automatic differentiation," *Phys. Rev. D*, vol. 107, no. 8, p. 083028, 2023.

[4] S. Soma, L. Wang, S. Shi, H. Stöcker, and K. Zhou, "A physics-based neural network reconstruction of the dense matter equation of state from neutron star observables," *EPJ Web Conf.*, vol. 276, p. 06007, 2023.

[5] S. Soma, L. Wang, S. Shi, H. Stöcker, and K. Zhou, "A neural network reconstruction of the neutron star equation of state via automatic differentiation," *PoS*, vol. FAIRness2022, p. 055, 2023.

[6] S. Soma, H. Stöcker, and K. Zhou, "Mass and tidal parameter extraction from gravitational waves of binary neutron stars mergers using deep learning," *JCAP*, vol. 01, p. 009, 2024.

[7] B. P. Abbott *et al.*, "Model comparison from LIGO–Virgo data on GW170817's binary components and consequences for the merger remnant," *Class. Quant. Grav.*, vol. 37, no. 4, p. 045006, 2020.

[8] S. Soma, "Neutron Stars: Celestial Laboratories for Dense Matter." `http://eprints.iiserkol.ac.in/id/eprint/932`, 2019.

[9] `https://www.gsi.de/work/forschung/theorie/theory-new/hot-and-dense-qcd`. Accessed: 2023-08-18.

[10] M. A. Stephanov, "QCD phase diagram: An Overview," *PoS*, vol. LAT2006, p. 024, 2006.

[11] H. Stoecker and W. Greiner, "High-Energy Heavy Ion Collisions: Probing the Equation of State of Highly Excited Hadronic Matter," *Phys. Rept.*, vol. 137, pp. 277–392, 1986.

[12] W. Busza, K. Rajagopal, and W. van der Schee, "Heavy Ion Collisions: The Big Picture, and the Big Questions," *Ann. Rev. Nucl. Part. Sci.*, vol. 68, pp. 339–376, 2018.

[13] K. Fukushima, B. Mohanty, and N. Xu, "Little-Bang and Femto-Nova in Nucleus-Nucleus Collisions," *AAPPS Bull.*, vol. 31, p. 1, 2021.

[14] S. Chatrchyan *et al.*, "The CMS Experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[15] P. Spiller and G. Franchetti, "The FAIR accelerator project at GSI," *Nucl. Instrum. Meth. A*, vol. 561, pp. 305–309, 2006.

[16] P. Spiller *et al.*, "The FAIR Heavy Ion Synchrotron SIS100," *JINST*, vol. 15, no. 12, p. T12013, 2020.

[17] H. Stoecker and C. Sturm, "The facility for antiproton and ion research FAIR cosmic matter in the Laboratory," *Nucl. Phys. A*, vol. 862-863, pp. 92–97, 2011.

[18] M. Durante *et al.*, "All the fun of the FAIR: fundamental physics at the facility for antiproton and ion research," *Phys. Scripta*, vol. 94, no. 3, p. 033001, 2019.

[19] B. Friman, C. Hohne, J. Knoll, S. Leupold, J. Randrup, R. Rapp, and P. Senger, eds., *The CBM physics book: Compressed baryonic matter in laboratory experiments*, vol. 814. 2011.

[20] K. C. Meehan, "The fixed-target experiment at STAR," *J. Phys. Conf. Ser.*, vol. 742, no. 1, p. 012022, 2016.

[21] G. Odyniec, "The RHIC Beam Energy Scan program in STAR and what's next ...," *J. Phys. Conf. Ser.*, vol. 455, p. 012037, 2013.

[22] V. Friese, "The CBM experiment at GSI/FAIR," *Nucl. Phys. A*, vol. 774, pp. 377–386, 2006.

[23] R. Krucken, "The NuSTAR facility at FAIR," *J. Phys. G*, vol. 31, pp. S1807–S1811, 2005.

[24] W. Erni, I. Keshelashvili, B. Krusche, M. Steinacher, Y. Heng, Z. Liu, H. Liu, X. Shen, O. Wang, H. Xu, *et al.*, "Physics performance report for PANDA: strong interaction studies with antiprotons," *arXiv:0903.3905*, 2009.

[25] B. B. Abelev *et al.*, "Performance of the ALICE Experiment at the CERN LHC," *Int. J. Mod. Phys. A*, vol. 29, p. 1430044, 2014.

[26] K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC," *JINST*, vol. 3, p. S08002, 2008.

[27] M. Gazdzicki, "NA49/NA61: results and plans on beam energy and system size scan at the CERN SPS," *J. Phys. G*, vol. 38, p. 124024, 2011.

[28] N. Abgrall *et al.*, "NA61/SHINE facility at the CERN SPS: beams and detector system," *JINST*, vol. 9, p. P06005, 2014.

[29] C. P. Welsch and J. Ullrich, "FLAIR: A facility for low-energy antiproton and ion research," *Hyperfine Interact.*, vol. 172, pp. 71–80, 2006.

[30] V. Dexheimer, J. Noronha, J. Noronha-Hostler, C. Ratti, and N. Yunes, "Future physics perspectives on the equation of state from heavy ion collisions to neutron stars," *J. Phys. G*, vol. 48, no. 7, p. 073001, 2021.

[31] A. Motornenko, J. Steinheimer, V. Vovchenko, S. Schramm, and H. Stoecker, "Equation of state for hot QCD and compact stars from a mean field approach," *Phys. Rev. C*, vol. 101, no. 3, p. 034904, 2020.

[32] L. D. Landau, "On the theory of stars," *Phys. Z. Sowjetunion*, vol. 1, p. 285, 1932.

[33] D. G. Yakovlev, P. Haensel, G. Baym, and C. J. Pethick, "Lev Landau and the concept of neutron stars," *Phys. Usp.*, vol. 56, pp. 289–295, 2013.

[34] J. Chadwick, "Possible Existence of a Neutron," *Nature*, vol. 129, p. 312, 1932.

[35] J. Chadwick, "The Existence of a Neutron," *Proc. Roy. Soc. Lond. A*, vol. 136, no. 830, pp. 692–708, 1932.

[36] W. Baade and F. Zwicky, "Supernovae and cosmic rays," *Phys. Rev*, vol. 45, no. 138, p. 20, 1934.

[37] W. Baade and F. Zwicky, "On Super-Novae," *Proc. Nat. Acad. Sci.*, vol. 20, no. 5, pp. 254–259, 1934.

[38] W. Baade and F. Zwicky, "Cosmic rays from super-novae," *Proceedings of the National Academy of Sciences*, vol. 20, no. 5, pp. 259–263, 1934.

[39] J. R. Oppenheimer and G. M. Volkoff, "On massive neutron cores," *Phys. Rev.*, vol. 55, pp. 374–381, 1939.

[40] J. Schaffner-Bielich, *Compact Star Physics*. Cambridge University Press, 8 2020.

[41] R. C. Tolman, "Static solutions of Einstein's field equations for spheres of fluid," *Phys. Rev.*, vol. 55, pp. 364–373, 1939.

[42] A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins, "Observation of a rapidly pulsating radio source," *Nature*, vol. 217, pp. 709–713, 1968.

[43] A. Burrows and D. Vartanyan, "Core-Collapse Supernova Explosion Theory," *Nature*, vol. 589, no. 7840, pp. 29–39, 2021.

[44] S. Chandrasekhar, "An introduction to the study of stellar structure," *Chicago*, 1939.

[45] L. Woltjer, "X-Rays and Type I Supernova Remnants.," *The Astrophysical Journal*, vol. 140, pp. 1309–1313, 1964.

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

[47] J. M. Lattimer, "Neutron stars," *Gen. Rel. Grav.*, vol. 46, p. 1713, 2014.

[48] L. Rezzolla, P. Pizzochero, D. I. Jones, N. Rea, and I. Vidaña, eds., *The Physics and Astrophysics of Neutron Stars*, vol. 457. Springer, 2018.

[49] D. C. Backer, S. R. Kulkarni, C. Heiles, M. Davis, and W. Goss, "A millisecond pulsar," *Nature*, vol. 300, no. 5893, pp. 615–618, 1982.

[50] M. M. Davis, J. H. Taylor, J. M. Weisberg, and D. C. Backer, "HIGH PRECISION TIMING OBSERVATIONS OF THE MILLISECOND PULSAR PSR 1937 + 21," *Nature*, vol. 315, pp. 547–550, 1985.

[51] D. R. Lorimer, "Binary and millisecond pulsars at the new millennium," *Living Rev. Rel.*, vol. 4, p. 5, 2001.

[52] I. H. Stairs, "Testing general relativity with pulsar timing," *Living Rev. Rel.*, vol. 6, p. 5, 2003.

[53] E. Fonseca *et al.*, "Refined Mass and Geometric Measurements of the High-mass PSR J0740+6620," *Astrophys. J. Lett.*, vol. 915, no. 1, p. L12, 2021.

[54] J. Antoniadis *et al.*, "A Massive Pulsar in a Compact Relativistic Binary," *Science*, vol. 340, p. 6131, 2013.

[55] R. Manchester and J. Taylor, *Pulsars*. A Series of books in astronomy and astrophysics, W. H. Freeman, 1977.

[56] R. Blandford and S. A. Teukolsky, "Arrival-time analysis for a pulsar in a binary system.," *Astrophysical Journal, Vol. 205, p. 580-591*, vol. 205, pp. 580–591, 1976.

[57] A. Einstein, "Zur allgemeinen Relativitätstheorie," *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften*, pp. 778–785, 1915.

[58] A. Einstein, "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie," *Sitzungsberichte der königlich preussischen Akademie der Wissenschaften*, pp. 831–839, 1915.

[59] A. Einstein, "Die Feldgleichungen der Gravitation," *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften*, pp. 844–847, 1915.

[60] I. I. Shapiro, "Fourth Test of General Relativity," *Phys. Rev. Lett.*, vol. 13, pp. 789–791, 1964.

[61] T. Damour and N. Deruelle, "General relativistic celestial mechanics of binary systems. II. The post-Newtonian timing formula," in *Annales de l'IHP Physique théorique*, vol. 44, pp. 263–292, 1986.

[62] J. M. Lattimer, "The nuclear equation of state and neutron star masses," *Ann. Rev. Nucl. Part. Sci.*, vol. 62, pp. 485–515, 2012.

[63] J. M. Lattimer and A. W. Steiner, "Neutron Star Masses and Radii from Quiescent Low-Mass X-ray Binaries," *Astrophys. J.*, vol. 784, p. 123, 2014.

[64] `https://stellarcollapse.org/nsmasses.html`. Accessed: 2023-08-23.

[65] M. Prakash, T. L. Ainsworth, and J. M. Lattimer, "Equation of state and the maximum mass of neutron stars," *Phys. Rev. Lett.*, vol. 61, pp. 2518–2521, 1988.

[66] R. B. Wiringa, V. Fiks, and A. Fabrocini, "Equation of state for dense nucleon matter," *Phys. Rev. C*, vol. 38, pp. 1010–1037, 1988.

[67] N. K. Glendenning and S. A. Moszkowski, "Reconciliation of neutron star masses and binding of the lambda in hypernuclei," *Phys. Rev. Lett.*, vol. 67, pp. 2414–2417, 1991.

[68] M. Prakash, J. R. Cooke, and J. M. Lattimer, "Quark - hadron phase transition in protoneutron stars," *Phys. Rev. D*, vol. 52, pp. 661–665, 1995.

[69] H. Mueller and B. D. Serot, "Relativistic mean field theory and the high density nuclear equation of state," *Nucl. Phys. A*, vol. 606, pp. 508–537, 1996.

[70] N. K. Glendenning and J. Schaffner-Bielich, "First order kaon condensate," *Phys. Rev. C*, vol. 60, p. 025803, 1999.

[71] W. C. Ho, C. M. Espinoza, D. Antonopoulou, and N. Andersson, "Pinning down the superfluid and measuring masses using pulsar glitches," *Science advances*, vol. 1, no. 9, p. e1500578, 2015.

[72] 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.*, vol. 908, no. 2, p. L46, 2021.

[73] R. W. Romani, D. Kandel, A. V. Filippenko, T. G. Brink, and W. Zheng, "PSR J0952−0607: The Fastest and Heaviest Known Galactic Neutron Star," *Astrophys. J. Lett.*, vol. 934, no. 2, p. L18, 2022.

[74] K. C. Gendreau, Z. Arzoumanian, P. W. Adkins, C. L. Albert, J. F. Anders, A. T. Aylward, C. L. Baker, E. R. Balsamo, W. A. Bamford, S. S. Benegalrao, *et al.*, "The neutron star interior composition explorer (NICER): design and development,"

in *Space telescopes and instrumentation 2016: Ultraviolet to gamma ray*, vol. 9905, pp. 420–435, SPIE, 2016.

[75] 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.*, vol. 820, no. 1, p. 28, 2016.

[76] W. Lewin and M. Van der Klis, *Compact stellar X-ray sources*, vol. 39. Cambridge University Press, 2006.

[77] F. Ozel, "Surface Emission from Neutron Stars and Implications for the Physics of their Interiors," *Rept. Prog. Phys.*, vol. 76, p. 016901, 2013.

[78] R. E. Rutledge, L. Bildsten, E. F. Brown, G. G. Pavlov, and V. E. Zavlin, "The thermal X-Ray spectra of Centaurus X-4, Aquila X-1, and 4U 1608-522 in quiescence," *The Astrophysical Journal*, vol. 514, no. 2, p. 945, 1999.

[79] P. Haensel and J. Zdunik, "Non-equilibrium processes in the crust of an accreting neutron star," *Astronomy and Astrophysics (ISSN 0004-6361), vol. 227, no. 2, Jan. 1990, p. 431-436.*, vol. 227, pp. 431–436, 1990.

[80] P. Haensel and J. Zdunik, "Models of crustal heating in accreting neutron stars," *Astronomy & Astrophysics*, vol. 480, no. 2, pp. 459–464, 2008.

[81] E. F. Brown, L. Bildsten, and R. E. Rutledge, "Crustal heating and quiescent emission from transiently accreting neutron stars," *The Astrophysical Journal*, vol. 504, no. 2, p. L95, 1998.

[82] S. Bogdanov, C. O. Heinke, F. Özel, and T. Güver, "Neutron Star Mass-Radius Constraints of the Quiescent Low-mass X-ray Binaries X7 and X5 in the Globular Cluster 47 Tuc," *Astrophys. J.*, vol. 831, no. 2, p. 184, 2016.

[83] K. C. Gendreau, Z. Arzoumanian, and T. Okajima, "The Neutron star Interior Composition ExploreR (NICER): an Explorer mission of opportunity for soft x-ray timing spectroscopy," in *Space Telescopes and Instrumentation 2012: Ultraviolet to Gamma Ray*, vol. 8443, pp. 322–329, SPIE, 2012.

[84] Z. Arzoumanian, K. Gendreau, C. Baker, T. Cazeau, P. Hestnes, J. Kellogg, S. Kenyon, R. Kozon, K.-C. Liu, S. Manthripragada, *et al.*, "The neutron star interior composition explorer (NICER): mission definition," in *Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray*, vol. 9144, pp. 579–587, SPIE, 2014.

[85] A. L. Watts *et al.*, "Colloquium : Measuring the neutron star equation of state using x-ray timing," *Rev. Mod. Phys.*, vol. 88, no. 2, p. 021001, 2016.

[86] T. E. Riley *et al.*, "A *NICER* View of PSR J0030+0451: Millisecond Pulsar Parameter Estimation," *Astrophys. J. Lett.*, vol. 887, no. 1, p. L21, 2019.

[87] 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.*, vol. 887, no. 1, p. L24, 2019.

[88] S. Vinciguerra, T. Salmi, A. L. Watts, D. Choudhury, T. E. Riley, P. S. Ray, S. Bogdanov, Y. Kini, S. Guillot, D. Chakrabarty, *et al.*, "An updated mass-radius analysis of the 2017-2018 NICER data set of PSR J0030+ 0451," *arXiv:2308.09469*, 2023.

[89] 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.*, vol. 918, no. 2, p. L27, 2021.

[90] M. C. Miller *et al.*, "The Radius of PSR J0740+6620 from NICER and XMM-Newton Data," *Astrophys. J. Lett.*, vol. 918, no. 2, p. L28, 2021.

[91] T. Salmi *et al.*, "The Radius of PSR J0740+6620 from NICER with NICER Background Estimates," *Astrophys. J.*, vol. 941, no. 2, p. 150, 2022.

[92] R. A. Hulse and J. H. Taylor, "Discovery of a pulsar in a binary system," *Astrophys. J. Lett.*, vol. 195, pp. L51–L53, 1975.

[93] J. H. Taylor, L. A. Fowler, and P. M. McCulloch, "Measurements of general relativistic effects in the binary pulsar PSR 1913+16," *Nature*, vol. 277, pp. 437–440, 1979.

[94] J. H. Taylor and J. M. Weisberg, "A new test of general relativity: Gravitational radiation and the binary pulsar PS R 1913+16," *Astrophys. J.*, vol. 253, pp. 908–920, 1982.

[95] J. H. Taylor and J. M. Weisberg, "Further experimental tests of relativistic gravity using the binary pulsar PSR 1913+16," *Astrophys. J.*, vol. 345, pp. 434–450, 1989.

[96] J. M. Weisberg and J. H. Taylor, "Relativistic binary pulsar B1913+16: Thirty years of observations and analysis," *ASP Conf. Ser.*, vol. 328, p. 25, 2005.

[97] P. C. Peters and J. Mathews, "Gravitational radiation from point masses in a Keplerian orbit," *Phys. Rev.*, vol. 131, pp. 435–439, 1963.

[98] P. C. Peters, "Gravitational Radiation and the Motion of Two Point Masses," *Phys. Rev.*, vol. 136, pp. B1224–B1232, 1964.

[99] K. Chatziioannou, "Neutron star tidal deformability and equation of state constraints," *Gen. Rel. Grav.*, vol. 52, no. 11, p. 109, 2020.

[100] E. E. Flanagan and T. Hinderer, "Constraining neutron star tidal Love numbers with gravitational wave detectors," *Phys. Rev. D*, vol. 77, p. 021502, 2008.

[101] L. Blanchet, "Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries," *Living Rev. Rel.*, vol. 17, p. 2, 2014.

[102] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, "Systematic and statistical errors in a bayesian approach to the estimation of the neutron-star equation of state using advanced gravitational wave detectors," *Phys. Rev. D*, vol. 89, no. 10, p. 103012, 2014.

[103] M. Favata, "Systematic parameter errors in inspiraling neutron star binaries," *Phys. Rev. Lett.*, vol. 112, p. 101101, 2014.

[104] B. P. Abbott *et al.*, "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral," *Phys. Rev. Lett.*, vol. 119, no. 16, p. 161101, 2017.

[105] B. P. Abbott *et al.*, "Properties of the binary neutron star merger GW170817," *Phys. Rev. X*, vol. 9, no. 1, p. 011001, 2019.

[106] B. P. Abbott *et al.*, "Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA," *Living Rev. Rel.*, vol. 21, no. 1, p. 3, 2018.

[107] F. Iacovelli, M. Mancarella, S. Foffa, and M. Maggiore, "Forecasting the Detection Capabilities of Third-generation Gravitational-wave Detectors Using GWFAST," *Astrophys. J.*, vol. 941, no. 2, p. 208, 2022.

[108] B. M. Barker and R. F. O'Connell, "Gravitational Two-Body Problem with Arbitrary Masses, Spins, and Quadrupole Moments," *Phys. Rev. D*, vol. 12, pp. 329–335, 1975.

[109] D. Lai, L. Bildsten, and V. M. Kaspi, "Spin orbit interaction in neutron star / main sequence binaries and implications for pulsar timing," *Astrophys. J.*, vol. 452, p. 819, 1995.

[110] J. M. Lattimer and B. F. Schutz, "Constraining the equation of state with moment of inertia measurements," *Astrophys. J.*, vol. 629, pp. 979–984, 2005.

[111] T. Damour and G. Schafer, "Higher-order relativistic periastron advances and binary pulsars," *Nuovo Cimento, B*, vol. 101, no. 2, pp. 127–176, 1988.

[112] H. Hu, M. Kramer, N. Wex, D. J. Champion, and M. S. Kehl, "Constraining the dense matter equation-of-state with radio pulsars," *Mon. Not. Roy. Astron. Soc.*, vol. 497, no. 3, pp. 3118–3130, 2020.

[113] P. W. Anderson and N. Itoh, "Pulsar glitches and restlessness as a hard superfluidity phenomenon," *Nature*, vol. 256, pp. 25–27, 1975.

[114] D. Pines and M. A. Alpar, "Superfluidity in neutron stars," *Nature*, vol. 316, no. 6023, pp. 27–32, 1985.

[115] B. Link, R. I. Epstein, and J. M. Lattimer, "Pulsar constraints on neutron star structure and equation of state," *Phys. Rev. Lett.*, vol. 83, pp. 3362–3365, 1999.

[116] L. Samuelsson and N. Andersson, "Neutron Star Asteroseismology. Axial Crust Oscillations in the Cowling Approximation," *Mon. Not. Roy. Astron. Soc.*, vol. 374, pp. 256–268, 2007.

[117] P. Möller, W. D. Myers, W. J. Swiatecki, and J. Treiner, "Nuclear mass formula with a finite-range droplet model and a folded-Yukawa single-particle potential," *Atom. Data Nucl. Data Tabl.*, vol. 39, pp. 225–233, 1988.

[118] J. M. Lattimer and Y. Lim, "Constraining the Symmetry Parameters of the Nuclear Interaction," *Astrophys. J.*, vol. 771, p. 51, 2013.

[119] M. Oertel, M. Hempel, T. Klähn, and S. Typel, "Equations of state for supernovae and compact stars," *Rev. Mod. Phys.*, vol. 89, no. 1, p. 015007, 2017.

[120] E. Epelbaum, H.-W. Hammer, and U.-G. Meissner, "Modern Theory of Nuclear Forces," *Rev. Mod. Phys.*, vol. 81, pp. 1773–1825, 2009.

[121] R. Machleidt and D. R. Entem, "Chiral effective field theory and nuclear forces," *Phys. Rept.*, vol. 503, pp. 1–75, 2011.

[122] 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.*, vol. 773, p. 11, 2013.

[123] C. Drischler, A. Carbone, K. Hebeler, and A. Schwenk, "Neutron matter from chiral two- and three-nucleon calculations up to N^3LO," *Phys. Rev. C*, vol. 94, no. 5, p. 054307, 2016.

[124] C. Drischler, K. Hebeler, and A. Schwenk, "Chiral interactions up to next-to-next-to-next-to-leading order and nuclear saturation," *Phys. Rev. Lett.*, vol. 122, no. 4, p. 042501, 2019.

[125] C. Drischler, J. W. Holt, and C. Wellenhofer, "Chiral Effective Field Theory and the High-Density Nuclear Equation of State," *Ann. Rev. Nucl. Part. Sci.*, vol. 71, pp. 403–432, 2021.

[126] V. R. Pandharipande and R. A. Smith, "A model neutron solid with π 0 condensate," *Nucl. Phys. A*, vol. 237, pp. 507–532, 1975.

[127] J. M. Lattimer and F. D. Swesty, "A Generalized equation of state for hot, dense matter," *Nucl. Phys. A*, vol. 535, pp. 331–376, 1991.

[128] Y. Sugahara and H. Toki, "Relativistic mean field theory for unstable nuclei with nonlinear sigma and omega terms," *Nucl. Phys. A*, vol. 579, pp. 557–572, 1994.

[129] H. Toki, D. Hirata, Y. Sugahara, K. Sumiyoshi, and I. Tanihata, "Relativistic many body approach for unstable nuclei and supernova," *Nucl. Phys. A*, vol. 588, no. 1, pp. c357–c363, 1995.

[130] F. Douchin and P. Haensel, "A unified equation of state of dense matter and neutron star structure," *Astron. Astrophys.*, vol. 380, p. 151, 2001.

[131] S. Typel, G. Ropke, T. Klahn, D. Blaschke, and H. H. Wolter, "Composition and thermodynamics of nuclear matter with light clusters," *Phys. Rev. C*, vol. 81, p. 015803, 2010.

[132] G. Shen, C. J. Horowitz, and S. Teige, "A New Equation of State for Astrophysical Simulations," *Phys. Rev. C*, vol. 83, p. 035802, 2011.

[133] A. W. Steiner, M. Hempel, and T. Fischer, "Core-collapse supernova equations of state based on neutron star observations," *Astrophys. J.*, vol. 774, p. 17, 2013.

[134] A. S. Schneider, L. F. Roberts, and C. D. Ott, "Open-source nuclear equation of state framework based on the liquid-drop model with Skyrme interaction," *Phys. Rev. C*, vol. 96, no. 6, p. 065802, 2017.

[135] X. Du, A. W. Steiner, and J. W. Holt, "Hot and dense matter equation of state probability distributions for astrophysical simulations," *Phys. Rev. C*, vol. 105, no. 3, p. 035803, 2022.

[136] M. Alford, M. Braby, M. W. Paris, and S. Reddy, "Hybrid stars that masquerade as neutron stars," *Astrophys. J.*, vol. 629, pp. 969–978, 2005.

[137] J. M. Lattimer and M. Prakash, "Neutron star structure and the equation of state," *Astrophys. J.*, vol. 550, p. 426, 2001.

[138] I. Sagert, T. Fischer, M. Hempel, G. Pagliara, J. Schaffner-Bielich, A. Mezzacappa, F. K. Thielemann, and M. Liebendorfer, "Signals of the QCD phase transition in core-collapse supernovae," *Phys. Rev. Lett.*, vol. 102, p. 081101, 2009.

[139] D. Blaschke, D. E. Alvarez-Castillo, and S. Benic, "Mass-radius constraints for compact stars and a critical endpoint," *PoS*, vol. CPOD2013, p. 063, 2013.

[140] R. D. Mellinger, F. Weber, W. Spinella, G. A. Contrera, and M. G. Orsaria, "Quark Deconfinement in Rotating Neutron Stars," *Universe*, vol. 3, no. 1, p. 5, 2017.

[141] N.-U. F. Bastian, "Phenomenological quark-hadron equations of state with first-order phase transitions for astrophysical applications," *Phys. Rev. D*, vol. 103, no. 2, p. 023001, 2021.

[142] T. Demircik, C. Ecker, and M. Järvinen, "Dense and Hot QCD at Strong Coupling," *Phys. Rev. X*, vol. 12, no. 4, p. 041012, 2022.

[143] S. Balberg and A. Gal, "An Effective equation of state for dense matter with strangeness," *Nucl. Phys. A*, vol. 625, pp. 435–472, 1997.

[144] S. Banik, M. Hempel, and D. Bandyopadhyay, "New Hyperon Equations of State for Supernovae and Neutron Stars in Density-dependent Hadron Field Theory," *Astrophys. J. Suppl.*, vol. 214, no. 2, p. 22, 2014.

[145] V. Dexheimer, R. Negreiros, and S. Schramm, "Reconciling Nuclear and Astrophysical Constraints," *Phys. Rev. C*, vol. 92, no. 1, p. 012801, 2015.

[146] V. Dexheimer, "Tabulated Neutron Star Equations of State Modeled within the Chiral Mean Field Model," *Publ. Astron. Soc. Austral.*, vol. 34, p. E006, 2017.

[147] M. Marques, M. Oertel, M. Hempel, and J. Novak, "New temperature dependent hyperonic equation of state: Application to rotating neutron star models and *I-Q* relations," *Phys. Rev. C*, vol. 96, no. 4, p. 045806, 2017.

[148] M. Fortin, M. Oertel, and C. Providência, "Hyperons in hot dense matter: what do the constraints tell us for equation of state?," *Publ. Astron. Soc. Austral.*, vol. 35, p. 44, 2018.

[149] H. Kochankovski, A. Ramos, and L. Tolos, "Equation of state for hot hyperonic neutron star matter," *Mon. Not. Roy. Astron. Soc.*, vol. 517, no. 1, pp. 507–517, 2022. [Erratum: Mon.Not.Roy.Astron.Soc. 518, 6376].

[150] T. Malik, S. Banik, and D. Bandyopadhyay, "Equation-of-state Table with Hyperon and Antikaon for Supernova and Neutron Star Merger," *Astrophys. J.*, vol. 910, no. 2, p. 96, 2021.

[151] N. Jokela, M. Järvinen, and J. Remes, "Holographic QCD in the Veneziano limit and neutron stars," *JHEP*, vol. 03, p. 041, 2019.

[152] C. Hoyos, N. Jokela, M. Jarvinen, J. G. Subils, J. Tarrio, and A. Vuorinen, "Transport in strongly coupled quark matter," *Phys. Rev. Lett.*, vol. 125, p. 241601, 2020.

[153] 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*, vol. 79, p. 124032, 2009.

[154] F. Ozel and D. Psaltis, "Reconstructing the Neutron-Star Equation of State from Astrophysical Measurements," *Phys. Rev. D*, vol. 80, p. 103003, 2009.

[155] K. Hebeler, J. M. Lattimer, C. J. Pethick, and A. Schwenk, "Constraints on neutron star radii based on chiral effective field theory interactions," *Phys. Rev. Lett.*, vol. 105, p. 161102, 2010.

[156] A. W. Steiner, J. M. Lattimer, and E. F. Brown, "The Equation of State from Observed Masses and Radii of Neutron Stars," *Astrophys. J.*, vol. 722, pp. 33–54, 2010.

[157] C. A. Raithel, F. Ozel, and D. Psaltis, "From Neutron Star Observables to the Equation of State: An Optimal Parametrization," *Astrophys. J.*, vol. 831, no. 1, p. 44, 2016.

[158] S. K. Greif, G. Raaijmakers, K. Hebeler, A. Schwenk, and A. L. Watts, "Equation of state sensitivities when inferring neutron star and dense matter properties," *Mon. Not. Roy. Astron. Soc.*, vol. 485, no. 4, pp. 5363–5376, 2019.

[159] Y. Fujimoto, K. Fukushima, and K. Murase, "Methodology study of machine learning for the neutron star equation of state," *Phys. Rev. D*, vol. 98, no. 2, p. 023019, 2018.

[160] Y. Fujimoto, K. Fukushima, and K. Murase, "Mapping neutron star data to the equation of state using the deep neural network," *Phys. Rev. D*, vol. 101, no. 5, p. 054016, 2020.

[161] Y. Fujimoto, K. Fukushima, and K. Murase, "Extensive Studies of the Neutron Star Equation of State from the Deep Learning Inference with the Observational Data Augmentation," *JHEP*, vol. 03, p. 273, 2021.

[162] L. Lindblom, "Spectral Representations of Neutron-Star Equations of State," *Phys. Rev. D*, vol. 82, p. 103011, 2010.

[163] F. Foucart, M. D. Duez, A. Gudinas, F. Hebert, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, "Smooth Equations of State for High-Accuracy Simulations of Neutron Star Binaries," *Phys. Rev. D*, vol. 100, no. 10, p. 104048, 2019.

[164] A. Knight, F. Foucart, M. D. Duez, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, "Gravitational Waves from Binary Neutron Star Mergers with a Spectral Equation of State," *arXiv:2307.03250*, 2023.

[165] P. Landry and R. Essick, "Nonparametric inference of the neutron star equation of state from gravitational wave observations," *Phys. Rev. D*, vol. 99, no. 8, p. 084049, 2019.

[166] 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*, vol. 101, no. 6, p. 063007, 2020.

[167] 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*, vol. 101, no. 12, p. 123007, 2020.

[168] 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.*, vol. 919, no. 1, p. 11, 2021.

[169] I. Legred, K. Chatziioannou, R. Essick, and P. Landry, "Implicit correlations within phenomenological parametric models of the neutron star equation of state," *Phys. Rev. D*, vol. 105, no. 4, p. 043016, 2022.

[170] I. Legred, B. O. Sy-Garcia, K. Chatziioannou, and R. Essick, "Assessing equation of state-independent relations for neutron stars with nonparametric models," *arXiv:2310.10854*, 2023.

[171] G. Baym, C. Pethick, and P. Sutherland, "The Ground state of matter at high densities: Equation of state and stellar models," *Astrophys. J.*, vol. 170, pp. 299–317, 1971.

[172] J. W. Negele and D. Vautherin, "Neutron star matter at subnuclear densities," *Nucl. Phys. A*, vol. 207, pp. 298–320, 1973.

[173] M. Hempel and J. Schaffner-Bielich, "Statistical Model for a Complete Supernova Equation of State," *Nucl. Phys. A*, vol. 837, pp. 210–254, 2010.

[174] J. Boguta and A. R. Bodmer, "Relativistic Calculation of Nuclear Matter and the Nuclear Surface," *Nucl. Phys. A*, vol. 292, pp. 413–428, 1977.

[175] J. D. Walecka, "A Theory of highly condensed matter," *Annals Phys.*, vol. 83, pp. 491–529, 1974.

[176] S. A. Chin and J. D. Walecka, "An Equation of State for Nuclear and Higher-Density Matter Based on a Relativistic Mean-Field Theory," *Phys. Lett. B*, vol. 52, pp. 24–28, 1974.

[177] A. W. Steiner, M. Prakash, J. M. Lattimer, and P. J. Ellis, "Isospin asymmetry in nuclei and neutron stars," *Phys. Rept.*, vol. 411, pp. 325–375, 2005.

[178] M. Hempel, T. Fischer, J. Schaffner-Bielich, and M. Liebendorfer, "New Equations of State in Simulations of Core-Collapse Supernovae," *Astrophys. J.*, vol. 748, p. 70, 2012.

[179] S. Typel, M. Oertel, and T. Klähn, "CompOSE CompStar online supernova equations of state harmonising the concert of nuclear physics and astrophysics compose.obspm.fr," *Phys. Part. Nucl.*, vol. 46, no. 4, pp. 633–664, 2015.

[180] S. Typel and H. H. Wolter, "Relativistic mean field calculations with density dependent meson nucleon coupling," *Nucl. Phys. A*, vol. 656, pp. 331–364, 1999.

[181] I. Tews, J. M. Lattimer, A. Ohnishi, and E. E. Kolomeitsev, "Symmetry Parameter Constraints from a Lower Bound on Neutron-matter Energy," *Astrophys. J.*, vol. 848, no. 2, p. 105, 2017.

[182] D. Lonardoni, I. Tews, S. Gandolfi, and J. Carlson, "Nuclear and neutron-star matter from local chiral interactions," *Phys. Rev. Res.*, vol. 2, no. 2, p. 022033, 2020.

[183] C. B. Dover and A. Gal, "HYPERON NUCLEUS POTENTIALS," *Prog. Part. Nucl. Phys.*, vol. 12, pp. 171–239, 1985.

[184] J. Schaffner and I. N. Mishustin, "Hyperon rich matter in neutron stars," *Phys. Rev. C*, vol. 53, pp. 1416–1429, 1996.

[185] D. J. Millener, C. B. Dover, and A. Gal, "Lambda Nucleus Single Particle Potentials," *Phys. Rev. C*, vol. 38, pp. 2700–2708, 1988.

[186] J. Mares, E. Friedman, A. Gal, and B. K. Jennings, "Constraints on Sigma nucleus dynamics from Dirac phenomenology of Sigma- atoms," *Nucl. Phys. A*, vol. 594, pp. 311–324, 1995.

[187] J. Schaffner, H. Stoecker, and C. Greiner, "Metastable exotic multihypernuclear objects," *Phys. Rev. C*, vol. 46, pp. 322–329, 1992.

[188] V. Dexheimer and S. Schramm, "Proto-Neutron and Neutron Stars in a Chiral SU(3) Model," *Astrophys. J.*, vol. 683, pp. 943–948, 2008.

[189] V. Dexheimer, "Tabulated neutron star equations of state modelled within the chiral mean field model," *Publications of the Astronomical Society of Australia*, vol. 34, p. e066, 2017.

[190] A. Motornenko, J. Steinheimer, V. Vovchenko, S. Schramm, and H. Stoecker, "QCD equation of state at vanishing and high baryon density: Chiral Mean Field model," *Nucl. Phys. A*, vol. 1005, p. 121836, 2021.

[191] J. R. Stone, N. J. Stone, and S. A. Moszkowski, "Incompressibility in finite nuclei and nuclear matter," *Phys. Rev. C*, vol. 89, no. 4, p. 044316, 2014.

[192] N. Hornick, L. Tolos, A. Zacchi, J.-E. Christian, and J. Schaffner-Bielich, "Relativistic parameterizations of neutron matter and implications for neutron stars," *Phys. Rev. C*, vol. 98, no. 6, p. 065804, 2018. [Erratum: Phys.Rev.C 103, 039902 (2021)].

[193] A. Perego, S. Bernuzzi, and D. Radice, "Thermodynamics conditions of matter in neutron star mergers," *Eur. Phys. J. A*, vol. 55, no. 8, p. 124, 2019.

[194] T. Fischer, M. Hempel, I. Sagert, Y. Suwa, and J. Schaffner-Bielich, "Symmetry energy impact in simulations of core-collapse supernovae," *Eur. Phys. J. A*, vol. 50, p. 46, 2014.

[195] K. Schwarzschild, "Über das Gravitationsfeld eines Massenpunktes nach der EINSTEINschen Theorie," *Sitzungsberichte der königlich preussischen Akademie der Wissenschaften*, pp. 189–196, 1916.

[196] K. Schwarzschild, "Über das Gravitationsfeld einer Kugel aus inkompressibler Flüssigkeit nach der EINSTEINschen Theorie," *Sitzungsberichte der königlich preußischen Akademie der Wissenschaften zu Berlin*, pp. 424–434, 1916.

[197] G. D. Birkhoff and R. E. Langer, *Relativity and Modern Physics*. Harvard University Press, 1927.

[198] J. T. Jebsen, "On the General Spherically Symmetric Solutions of Einstein's Gravitational Equations in Vacuo.," *Arkiv for Matematik, Astronomi och Fysik*, vol. 15, no. 18, p. 18, 1921.

[199] T. Hinderer, "Tidal Love numbers of neutron stars," *Astrophys. J.*, vol. 677, pp. 1216–1220, 2008.

[200] T. Hinderer, B. D. Lackey, R. N. Lang, and J. S. Read, "Tidal deformability of neutron stars with realistic equations of state and their gravitational wave signatures in binary inspiral," *Phys. Rev. D*, vol. 81, p. 123016, 2010.

[201] K. S. Thorne, "Tidal stabilization of rigidly rotating, fully relativistic neutron stars," *Phys. Rev. D*, vol. 58, p. 124031, 1998.

[202] A. Guerra Chaves and T. Hinderer, "Probing the equation of state of neutron star matter with gravitational waves from binary inspirals in light of GW170817: a brief review," *J. Phys. G*, vol. 46, no. 12, p. 123002, 2019.

[203] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4. Springer, 2006.

[204] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[205] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.

[206] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 5, pp. 455–455, 1992.

[207] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural networks*, vol. 6, no. 6, pp. 861–867, 1993.

[208] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[209] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 249–256, JMLR Workshop and Conference Proceedings, 2010.

[210] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034, 2015.

[211] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv:1609.04747*, 2016.

[212] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv:1412.6980*, 2014.

[213] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[214] S. Forte, L. Garrido, J. I. Latorre, and A. Piccione, "Neural network parametrization of deep inelastic structure functions," *JHEP*, vol. 05, p. 062, 2002.

[215] L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, "Neural network determination of parton distributions: The Nonsinglet case," *JHEP*, vol. 03, p. 039, 2007.

[216] L. Kades, J. M. Pawlowski, A. Rothkopf, M. Scherzer, J. M. Urban, S. J. Wetzel, N. Wink, and F. P. G. Ziegler, "Spectral Reconstruction with Deep Neural Networks," *Phys. Rev. D*, vol. 102, no. 9, p. 096001, 2020.

[217] M. Zhou, F. Gao, J. Chao, Y.-X. Liu, and H. Song, "Application of radial basis functions neutral networks in spectral functions," *Phys. Rev. D*, vol. 104, no. 7, p. 076011, 2021.

[218] S.-Y. Chen, H.-T. Ding, F.-Y. Liu, G. Papp, and C.-B. Yang, "Machine learning spectral functions in lattice QCD," *arXiv:2110.13521*, 2021.

[219] L.-G. Pang, K. Zhou, N. Su, H. Petersen, H. Stöcker, and X.-N. Wang, "An equation-of-state-meter of quantum chromodynamics transition from deep learning," *Nature Commun.*, vol. 9, no. 1, p. 210, 2018.

[220] J. Carrasquilla and R. G. Melko, "Machine learning phases of matter," *Nature Physics*, vol. 13, no. 5, pp. 431–434, 2017.

[221] Y.-L. Du, K. Zhou, J. Steinheimer, L.-G. Pang, A. Motornenko, H.-S. Zong, X.-N. Wang, and H. Stöcker, "Identifying the nature of the QCD transition in relativistic collision of heavy nuclei with deep learning," *Eur. Phys. J. C*, vol. 80, no. 6, p. 516, 2020.

[222] L. Wang, Y. Jiang, L. He, and K. Zhou, "Continuous-Mixture Autoregressive Networks Learning the Kosterlitz-Thouless Transition," *Chin. Phys. Lett.*, vol. 39, no. 12, p. 120502, 2022.

[223] R. Wang, Y.-G. Ma, R. Wada, L.-W. Chen, W.-B. He, H.-L. Liu, and K.-J. Sun, "Nuclear liquid-gas phase transition with machine learning," *Phys. Rev. Res.*, vol. 2, no. 4, p. 043202, 2020.

[224] L. Jiang, L. Wang, and K. Zhou, "Deep learning stochastic processes with QCD phase transition," *Phys. Rev. D*, vol. 103, no. 11, p. 116023, 2021.

[225] D. Boyda, G. Kanwar, S. Racanière, D. J. Rezende, M. S. Albergo, K. Cranmer, D. C. Hackett, and P. E. Shanahan, "Sampling using $SU(N)$ gauge equivariant flows," *Phys. Rev. D*, vol. 103, no. 7, p. 074504, 2021.

[226] K. Zhou, G. Endrődi, L.-G. Pang, and H. Stöcker, "Regressive and generative neural networks for scalar field theory," *Phys. Rev. D*, vol. 100, no. 1, p. 011501, 2019.

[227] G. Kanwar, M. S. Albergo, D. Boyda, K. Cranmer, D. C. Hackett, S. Racanière, D. J. Rezende, and P. E. Shanahan, "Equivariant flow-based sampling for lattice gauge theory," *Phys. Rev. Lett.*, vol. 125, no. 12, p. 121601, 2020.

[228] M. S. Albergo, G. Kanwar, and P. E. Shanahan, "Flow-based generative models for Markov chain Monte Carlo in lattice field theory," *Phys. Rev. D*, vol. 100, no. 3, p. 034515, 2019.

[229] M. Omana Kuttan, J. Steinheimer, K. Zhou, A. Redelbach, and H. Stoecker, "A fast centrality-meter for heavy-ion collisions at the CBM experiment," *Phys. Lett. B*, vol. 811, p. 135872, 2020.

[230] P. Thaprasop, K. Zhou, J. Steinheimer, and C. Herold, "Unsupervised Outlier Detection in Heavy-Ion Collisions," *Phys. Scripta*, vol. 96, no. 6, p. 064003, 2021.

[231] F. Li, Y. Wang, H. Lü, P. Li, Q. Li, and F. Liu, "Application of artificial intelligence in the determination of impact parameter in heavy-ion collisions at intermediate energies," *J. Phys. G*, vol. 47, no. 11, p. 115104, 2020.

[232] A. Andreassen, S.-C. Hsu, B. Nachman, N. Suaysom, and A. Suresh, "Parameter estimation using neural networks in the presence of detector effects," *Phys. Rev. D*, vol. 103, no. 3, p. 036001, 2021.

[233] M. Omana Kuttan, K. Zhou, J. Steinheimer, A. Redelbach, and H. Stoecker, "An equation-of-state-meter for CBM using PointNet," *JHEP*, vol. 21, p. 184, 2020.

[234] H. Huang, B. Xiao, Z. Liu, Z. Wu, Y. Mu, and H. Song, "Applications of deep learning to relativistic hydrodynamics," *Phys. Rev. Res.*, vol. 3, no. 2, p. 023256, 2021.

[235] 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.*, vol. 844, no. 2, p. 156, 2017.

[236] S. Traversi, P. Char, and G. Pagliara, "Bayesian Inference of Dense Matter Equation of State within Relativistic Mean Field Models using Astrophysical Measurements," *Astrophys. J.*, vol. 897, p. 165, 2020.

[237] F. Morawski and M. Bejger, "Neural network reconstruction of the dense matter equation of state derived from the parameters of neutron stars," *Astron. Astrophys.*, vol. 642, p. A78, 2020.

[238] M. a. Ferreira and C. Providência, "Unveiling the nuclear matter EoS from neutron star properties: a supervised machine learning approach," *JCAP*, vol. 07, p. 011, 2021.

[239] P. G. Krastev, "Translating Neutron Star Observations to Nuclear Symmetry Energy via Deep Neural Networks," *Galaxies*, vol. 10, no. 1, p. 16, 2022.

[240] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Science advances*, vol. 4, no. 6, p. eaar4206, 2018.

[241] S. Shi, K. Zhou, J. Zhao, S. Mukherjee, and P. Zhuang, "Heavy quark potential in the quark-gluon plasma: Deep neural network meets lattice quantum chromodynamics," *Phys. Rev. D*, vol. 105, no. 1, p. 014017, 2022.

[242] L. Wang, S. Shi, and K. Zhou, "Reconstructing spectral functions via automatic differentiation," *Phys. Rev. D*, vol. 106, no. 5, p. L051502, 2022.

[243] F. Chollet *et al.*, "Keras." `https://github.com/fchollet/keras`, 2015.

[244] M. Abadi, A. Agarwal, P. Barham, *et al.*, " TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015. Software available from tensorflow.org.

[245] P. Demorest, T. Pennucci, S. Ransom, M. Roberts, and J. Hessels, "Shapiro Delay Measurement of A Two Solar Mass Neutron Star," *Nature*, vol. 467, pp. 1081–1083, 2010.

[246] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "WaveNet: A Generative Model for Raw Audio," 2016.

[247] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets* (J.-M. Combes, A. Grossmann, and P. Tchamitchian, eds.), (Berlin, Heidelberg), pp. 286–297, Springer Berlin Heidelberg, 1990.

[248] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," *arXiv:1511.07122*, 2015.

[249] A. Thete, K. Banerjee, and T. Malik, "Inferring the dense matter equation of state from neutron star observations via artificial neural networks," *arXiv:2208.13163*, 2022.

[250] G. Gonçalves, M. Ferreira, J. Aveiro, A. Onofre, F. F. Freitas, C. Providência, and J. A. Font, "Machine-Learning Love: classifying the equation of state of neutron stars with Transformers," *arXiv:2210.08382*, 2022.

[251] S. Chatterjee, H. Sudhakaran, and R. Mallick, "Analyzing the speed of sound in neutron star with machine learning," *arXiv:2302.13648*, 2023.

[252] L. Rezzolla, E. R. Most, and L. R. Weih, "Using gravitational-wave observations and quasi-universal relations to constrain the maximum mass of neutron stars," *Astrophys. J. Lett.*, vol. 852, no. 2, p. L25, 2018.

[253] M. Shibata, E. Zhou, K. Kiuchi, and S. Fujibayashi, "Constraint on the maximum mass of neutron stars using GW170817 event," *Phys. Rev. D*, vol. 100, no. 2, p. 023015, 2019.

[254] S.-P. Tang, B. Gao, Y.-J. Li, Y.-Z. Fan, and D.-M. Wei, "Measuring Mass and Radius of the Maximum-mass Nonrotating Neutron Star," *arXiv:2309.15441*, 2023.

[255] S. Ai, H. Gao, Y. Yuan, B. Zhang, and L. Lan, "What constraints can one pose on the maximum mass of neutron stars from multi-messenger observations?," *arXiv:2310.07133*, 2023.

[256] J. Nättilä, M. C. Miller, A. W. Steiner, J. J. E. Kajava, V. F. Suleimanov, and J. Poutanen, "Neutron star mass and radius measurements from atmospheric model fits to X-ray burst cooling tail spectra," *Astron. Astrophys.*, vol. 608, p. A31, 2017.

[257] D. Gonzalez-Caniulef, S. Guillot, and A. Reisenegger, "Neutron star radius measurement from the ultraviolet and soft X-ray thermal emission of PSR J0437−4715," *Mon. Not. Roy. Astron. Soc.*, vol. 490, no. 4, pp. 5848–5859, 2019.

[258] S. T. Tokdar and R. E. Kass, "Importance sampling: a review," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 1, pp. 54–60, 2010.

[259] A. W. Steiner, J. M. Lattimer, and E. F. Brown, "The Neutron Star Mass-Radius Relation and the Equation of State of Dense Matter," *Astrophys. J. Lett.*, vol. 765, p. L5, 2013.

[260] F. Özel and P. Freire, "Masses, Radii, and the Equation of State of Neutron Stars," *Ann. Rev. Astron. Astrophys.*, vol. 54, pp. 401–440, 2016.

[261] C. D. Capano, I. Tews, S. M. Brown, B. Margalit, S. De, S. Kumar, D. A. Brown, B. Krishnan, and S. Reddy, "Stringent constraints on neutron-star radii from multi-messenger observations and nuclear theory," *Nature Astron.*, vol. 4, no. 6, pp. 625–632, 2020.

[262] B. P. Abbott *et al.*, "GW170817: Measurements of neutron star radii and equation of state," *Phys. Rev. Lett.*, vol. 121, no. 16, p. 161101, 2018.

[263] E. Annala, T. Gorda, A. Kurkela, and A. Vuorinen, "Gravitational-wave constraints on the neutron-star-matter Equation of State," *Phys. Rev. Lett.*, vol. 120, no. 17, p. 172703, 2018.

[264] I. Tews, J. Carlson, S. Gandolfi, and S. Reddy, "Constraining the speed of sound inside neutron stars with chiral effective field theory interactions and observations," *Astrophys. J.*, vol. 860, no. 2, p. 149, 2018.

[265] K. Fukushima and T. Kojo, "The Quarkyonic Star," *Astrophys. J.*, vol. 817, no. 2, p. 180, 2016.

[266] T. Kojo, "QCD equations of state and speed of sound in neutron stars," *AAPPS Bull.*, vol. 31, no. 1, p. 11, 2021.

[267] M. O. Kuttan, J. Steinheimer, K. Zhou, and H. Stoecker, "The QCD EoS of dense nuclear matter from Bayesian analysis of heavy ion collision data," *arXiv:2211.11670*, 2022.

[268] E. R. Most, L. J. Papenfort, V. Dexheimer, M. Hanauske, S. Schramm, H. Stöcker, and L. Rezzolla, "Signatures of quark-hadron phase transitions in general-relativistic neutron-star mergers," *Phys. Rev. Lett.*, vol. 122, no. 6, p. 061101, 2019.

[269] P. Jakobus, B. Mueller, A. Heger, A. Motornenko, J. Steinheimer, and H. Stoecker, "The role of the hadron-quark phase transition in core-collapse supernovae," *Mon. Not. Roy. Astron. Soc.*, vol. 516, no. 2, pp. 2554–2574, 2022.

[270] A. Figura, J. J. Lu, G. F. Burgio, Z. H. Li, and H. J. Schulze, "Hybrid equation of state approach in binary neutron-star merger simulations," *Phys. Rev. D*, vol. 102, no. 4, p. 043006, 2020.

[271] B. P. Abbott *et al.*, "GW150914: The Advanced LIGO Detectors in the Era of First Discoveries," *Phys. Rev. Lett.*, vol. 116, no. 13, p. 131103, 2016.

[272] B. P. Abbott *et al.*, "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," *Phys. Rev. X*, vol. 9, no. 3, p. 031040, 2019.

[273] R. Abbott *et al.*, "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run," *Phys. Rev. X*, vol. 11, p. 021053, 2021.

[274] R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. Adhikari, V. Adya, C. Affeldt, D. Agarwal, *et al.*, "GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run," *arXiv:2108.01045*, 2021.

[275] R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. Adhikari, V. Adya, C. Affeldt, D. Agarwal, *et al.*, "GWTC-3: compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run," *arXiv:2111.03606*, 2021.

[276] "LIGO, Virgo and KAGRA Observing Run Plans." `https://observing.docs.ligo.org/plan/`, 2023.

[277] R. Abbott *et al.*, "Observation of Gravitational Waves from Two Neutron Star–Black Hole Coalescences," *Astrophys. J. Lett.*, vol. 915, no. 1, p. L5, 2021.

[278] R. Abbott *et al.*, "GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object," *Astrophys. J. Lett.*, vol. 896, no. 2, p. L44, 2020.

[279] B. P. Abbott *et al.*, "Multi-messenger Observations of a Binary Neutron Star Merger," *Astrophys. J. Lett.*, vol. 848, no. 2, p. L12, 2017.

[280] B. D. Metzger, "Kilonovae," *Living Rev. Rel.*, vol. 23, no. 1, p. 1, 2020.

[281] B. P. Abbott *et al.*, "Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A," *Astrophys. J. Lett.*, vol. 848, no. 2, p. L13, 2017.

[282] B. P. Abbott *et al.*, "A gravitational-wave standard siren measurement of the Hubble constant," *Nature*, vol. 551, no. 7678, pp. 85–88, 2017.

[283] C. Raithel, F. Özel, and D. Psaltis, "Tidal deformability from GW170817 as a direct probe of the neutron star radius," *Astrophys. J. Lett.*, vol. 857, no. 2, p. L23, 2018.

[284] T. Zhao and J. M. Lattimer, "Tidal Deformabilities and Neutron Star Mergers," *Phys. Rev. D*, vol. 98, no. 6, p. 063020, 2018.

[285] J.-L. Jiang, S.-P. Tang, Y.-Z. Wang, Y.-Z. Fan, and D.-M. Wei, "PSR J0030+0451, GW170817 and the nuclear data: joint constraints on equation of state and bulk properties of neutron stars," *Astrophys. J.*, vol. 892, p. 1, 2020.

[286] J. M. Lattimer, "Neutron Star Mass and Radius Measurements," *Universe*, vol. 5, no. 7, p. 159, 2019.

[287] E. R. Most, L. J. Papenfort, L. R. Weih, and L. Rezzolla, "A lower bound on the maximum mass if the secondary in GW190814 was once a rapidly spinning neutron star," *Mon. Not. Roy. Astron. Soc.*, vol. 499, no. 1, pp. L82–L86, 2020.

[288] D. George and E. A. Huerta, "Deep Neural Networks to Enable Real-time Multimessenger Astrophysics," *Phys. Rev. D*, vol. 97, no. 4, p. 044039, 2018.

[289] P. G. Krastev, "Real-Time Detection of Gravitational Waves from Binary Neutron Stars using Artificial Neural Networks," *Phys. Lett. B*, vol. 803, p. 135330, 2020.

[290] H. Shen, D. George, E. A. Huerta, and Z. Zhao, "Denoising gravitational waves with enhanced deep recurrent denoising auto-encoders," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3237–3241, IEEE, 2019.

[291] F. Morawski, M. Bejger, and P. Cieciel\ag, "Convolutional neural network classifier for the output of the time-domain F-statistic all-sky search for continuous gravitational waves," *Mach. Learn. Sci. Tech.*, vol. 1, no. 2, p. 025016, 2020.

[292] C. Verma, A. Reza, D. Krishnaswamy, S. Caudill, and G. Gaur, "Employing deep learning for detection of gravitational waves from compact binary coalescences," *AIP Conf. Proc.*, vol. 2555, no. 1, p. 020010, 2022.

[293] S. Jadhav, M. Shrivastava, and S. Mitra, "Towards a robust and reliable deep learning approach for detection of compact binary mergers in gravitational wave data," *arXiv:2306.11797*, 2023.

[294] S. R. Green, C. Simpson, and J. Gair, "Gravitational-wave parameter estimation with autoregressive neural network flows," *Phys. Rev. D*, vol. 102, no. 10, p. 104057, 2020.

[295] M. Dax, S. R. Green, J. Gair, J. H. Macke, A. Buonanno, and B. Schölkopf, "Real-Time Gravitational Wave Science with Neural Posterior Estimation," *Phys. Rev. Lett.*, vol. 127, no. 24, p. 241103, 2021.

[296] M. Dax, S. R. Green, J. Gair, M. Pürrer, J. Wildberger, J. H. Macke, A. Buonanno, and B. Schölkopf, "Neural Importance Sampling for Rapid and Reliable Gravitational-Wave Inference," *Phys. Rev. Lett.*, vol. 130, no. 17, p. 171403, 2023.

[297] W. Wei and E. A. Huerta, "Deep learning for gravitational wave forecasting of neutron star mergers," *Phys. Lett. B*, vol. 816, p. 136185, 2021.

[298] R. Mesuga and B. J. Bayanay, "A Deep Transfer Learning Approach on Identifying Glitch Wave-form in Gravitational Wave Data," *arXiv:2107.01863*, 2021.

[299] R. Qiu, P. G. Krastev, K. Gill, and E. Berger, "Deep learning detection and classification of gravitational waves from neutron star-black hole mergers," *Phys. Lett. B*, vol. 840, p. 137850, 2023.

[300] P. G. Krastev, K. Gill, V. A. Villar, and E. Berger, "Detection and Parameter Estimation of Gravitational Waves from Binary Neutron-Star Mergers in Real LIGO Data using Deep Learning," *Phys. Lett. B*, vol. 815, p. 136161, 2021.

[301] T. Dal Canton *et al.*, "Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors," *Phys. Rev. D*, vol. 90, no. 8, p. 082004, 2014.

[302] LIGO Scientific Collaboration, "LIGO Algorithm Library - LALSuite." free software (GPL), 2018.

[303] L. Santamaria *et al.*, "Matching post-Newtonian and numerical relativity waveforms: systematic errors and a new phenomenological model for non-precessing black hole binaries," *Phys. Rev. D*, vol. 82, p. 064016, 2010.

[304] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, "Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal," *Phys. Rev. D*, vol. 93, no. 4, p. 044006, 2016.

[305] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, "Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era," *Phys. Rev. D*, vol. 93, no. 4, p. 044007, 2016.

[306] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, "Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms," *Phys. Rev. Lett.*, vol. 113, no. 15, p. 151101, 2014.

[307] S. Bernuzzi, A. Nagar, T. Dietrich, and T. Damour, "Modeling the Dynamics of Tidally Interacting Binary Neutron Stars up to the Merger," *Phys. Rev. Lett.*, vol. 114, no. 16, p. 161103, 2015.

[308] T. Hinderer *et al.*, "Effects of neutron-star dynamic tides on gravitational waveforms within the effective-one-body approach," *Phys. Rev. Lett.*, vol. 116, no. 18, p. 181101, 2016.

[309] M. Pürrer, "Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass-ratios and spins," *Phys. Rev. D*, vol. 93, no. 6, p. 064041, 2016.

[310] T. Dietrich, A. Samajdar, S. Khan, N. K. Johnson-McDaniel, R. Dudi, and W. Tichy, "Improving the NRTidal model for binary neutron star systems," *Phys. Rev. D*, vol. 100, no. 4, p. 044003, 2019.

[311] G. Montana, L. Tolos, M. Hanauske, and L. Rezzolla, "Constraining twin stars with GW170817," *Phys. Rev. D*, vol. 99, no. 10, p. 103009, 2019.

[312] B. S. Sathyaprakash and B. F. Schutz, "Physics, Astrophysics and Cosmology with Gravitational Waves," *Living Rev. Rel.*, vol. 12, p. 2, 2009.

[313] C. Cutler and E. E. Flanagan, "Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form?," *Phys. Rev. D*, vol. 49, pp. 2658–2697, 1994.

[314] T. Damour, A. Nagar, and L. Villain, "Measurability of the tidal polarizability of neutron stars in late-inspiral gravitational-wave signals," *Phys. Rev. D*, vol. 85, p. 123007, 2012.

[315] M. W. Coughlin *et al.*, "Constraints on the neutron star equation of state from AT2017gfo using radiative transfer simulations," *Mon. Not. Roy. Astron. Soc.*, vol. 480, no. 3, pp. 3871–3878, 2018.

[316] K. Yagi and N. Yunes, "I-Love-Q Relations in Neutron Stars and their Applications to Astrophysics, Gravitational Waves and Fundamental Physics," *Phys. Rev. D*, vol. 88, no. 2, p. 023009, 2013.

[317] J. S. Read, L. Baiotti, J. D. E. Creighton, J. L. Friedman, B. Giacomazzo, K. Kyutoku, C. Markakis, L. Rezzolla, M. Shibata, and K. Taniguchi, "Matter effects on binary neutron star waveforms," *Phys. Rev. D*, vol. 88, p. 044042, 2013.

[318] S. Bernuzzi, T. Dietrich, and A. Nagar, "Modeling the complete gravitational wave spectrum of neutron star mergers," *Phys. Rev. Lett.*, vol. 115, no. 9, p. 091101, 2015.

[319] L. Rezzolla and K. Takami, "Gravitational-wave signal from binary neutron stars: a systematic analysis of the spectral properties," *Phys. Rev. D*, vol. 93, no. 12, p. 124051, 2016.

[320] A. Bauswein, N.-U. F. Bastian, D. B. Blaschke, K. Chatziioannou, J. A. Clark, T. Fischer, and M. Oertel, "Identifying a first-order phase transition in neutron star mergers through gravitational waves," *Phys. Rev. Lett.*, vol. 122, no. 6, p. 061102, 2019.

[321] G. Lioutas, A. Bauswein, and N. Stergioulas, "Frequency deviations in universal relations of isolated neutron stars and postmerger remnants," *Phys. Rev. D*, vol. 104, no. 4, p. 043011, 2021.

[322] B. P. Abbott *et al.*, "Search for gravitational waves from a long-lived remnant of the binary neutron star merger GW170817," *Astrophys. J.*, vol. 875, no. 2, p. 160, 2019.

[323] M. Shibata, S. Fujibayashi, K. Hotokezaka, K. Kiuchi, K. Kyutoku, Y. Sekiguchi, and M. Tanaka, "Modeling GW170817 based on numerical relativity and its implications," *Phys. Rev. D*, vol. 96, no. 12, p. 123012, 2017.

[324] B. D. Metzger, "Welcome to the multi-messenger era! Lessons from a neutron star merger and the landscape ahead," *arXiv:1710.05931*, 2017.

[325] Y. Sekiguchi, K. Kiuchi, K. Kyutoku, and M. Shibata, "Gravitational waves and neutrino emission from the merger of binary neutron stars," *Phys. Rev. Lett.*, vol. 107, p. 051102, 2011.

[326] A. Endrizzi, A. Perego, F. M. Fabbri, L. Branca, D. Radice, S. Bernuzzi, B. Giacomazzo, F. Pederiva, and A. Lovato, "Thermodynamics conditions of matter in the neutrino decoupling region during neutron star mergers," *Eur. Phys. J. A*, vol. 56, no. 1, p. 15, 2020.

[327] D. Radice, S. Bernuzzi, W. Del Pozzo, L. F. Roberts, and C. D. Ott, "Probing Extreme-Density Matter with Gravitational Wave Observations of Binary Neutron Star Merger Remnants," *Astrophys. J. Lett.*, vol. 842, no. 2, p. L10, 2017.

[328] D. Radice, A. Perego, S. Bernuzzi, and B. Zhang, "Long-lived Remnants from Binary Neutron Star Mergers," *Mon. Not. Roy. Astron. Soc.*, vol. 481, no. 3, pp. 3670–3682, 2018.

[329] M. Breschi, S. Bernuzzi, F. Zappa, M. Agathos, A. Perego, D. Radice, and A. Nagar, "kiloHertz gravitational waves from binary neutron star remnants: time-domain model and constraints on extreme matter," *Phys. Rev. D*, vol. 100, no. 10, p. 104029, 2019.

[330] R. Ciolfi, W. Kastaun, J. V. Kalinani, and B. Giacomazzo, "First 100 ms of a long-lived magnetized neutron star formed in a binary neutron star merger," *Phys. Rev. D*, vol. 100, no. 2, p. 023005, 2019.

[331] K. Sumiyoshi, S. Fujibayashi, Y. Sekiguchi, and M. Shibata, "Properties of Neutrino Transfer in a Deformed Remnant of a Neutron Star Merger," *Astrophys. J.*, vol. 907, no. 2, p. 92, 2021.

[332] K. Hotokezaka, K. Kiuchi, K. Kyutoku, T. Muranushi, Y.-i. Sekiguchi, M. Shibata, and K. Taniguchi, "Remnant massive neutron stars of binary neutron star mergers: Evolution process and gravitational waveform," *Phys. Rev. D*, vol. 88, p. 044026, 2013.

[333] S. Fujibayashi, K. Kiuchi, N. Nishimura, Y. Sekiguchi, and M. Shibata, "Mass Ejection from the Remnant of a Binary Neutron Star Merger: Viscous-Radiation Hydrodynamics Study," *Astrophys. J.*, vol. 860, no. 1, p. 64, 2018.

[334] K. Kiuchi, K. Kyutoku, Y. Sekiguchi, and M. Shibata, "Global simulations of strongly magnetized remnant massive neutron stars formed in binary neutron star mergers," *Phys. Rev. D*, vol. 97, no. 12, p. 124039, 2018.

[335] J. D. Kaplan, C. D. Ott, E. P. O'Connor, K. Kiuchi, L. Roberts, and M. Duez, "The Influence of Thermal Pressure on Equilibrium Models of Hypermassive Neutron Star Merger Remnants," *Astrophys. J.*, vol. 790, p. 19, 2014.

[336] E. Gourgoulhon, P. Grandclément, J.-A. Marck, J. Novak, and K. Taniguchi, "LORENE: Spectral methods differential equations solver." Astrophysics Source Code Library, record ascl:1608.018, Aug. 2016.

[337] S. Bonazzola, E. Gourgoulhon, M. Salgado, and J. A. Marck, "Axisymmetric rotating relativistic bodies: A new numerical approach for 'exact' solutions," *Astron. Astrophys.*, vol. 278, pp. 421–443, 1993.

[338] S. De, D. Finstad, J. M. Lattimer, D. A. Brown, E. Berger, and C. M. Biwer, "Tidal Deformabilities and Radii of Neutron Stars from the Observation of GW170817," *Phys. Rev. Lett.*, vol. 121, no. 9, p. 091102, 2018. [Erratum: Phys.Rev.Lett. 121, 259902 (2018)].

[339] F. J. Fattoyev, J. Piekarewicz, and C. J. Horowitz, "Neutron Skins and Neutron Stars in the Multimessenger Era," *Phys. Rev. Lett.*, vol. 120, no. 17, p. 172702, 2018.

[340] 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.*, vol. 120, no. 26, p. 261103, 2018.

[341] D. Radice, A. Perego, F. Zappa, and S. Bernuzzi, "GW170817: Joint Constraint on the Neutron Star Equation of State from Multimessenger Observations," *Astrophys. J. Lett.*, vol. 852, no. 2, p. L29, 2018.

[342] S. A. Bhat and D. Bandyopadhyay, "Neutron star equation of state and GW170817," *J. Phys. G*, vol. 46, no. 1, p. 014003, 2019.

[343] P. Szewczyk, D. Gondek-Rosińska, and P. Cerdá-Durán, "Maximum mass and stability of differentially rotating neutrons stars," *arXiv:2302.06007*, 2023.

[344] D. Farrell, F. Weber, J. Jie Li, and A. Sedrakian, "Differential Rotation in Compact Objects with Hyperons and Delta Isobars," *arXiv:2311.09158*, 2023.

[345] O. Komoltsev and A. Kurkela, "How Perturbative QCD Constrains the Equation of State at Neutron-Star Densities," *Phys. Rev. Lett.*, vol. 128, no. 20, p. 202701, 2022.

[346] A. L. Watts *et al.*, "Dense matter with eXTP," *Sci. China Phys. Mech. Astron.*, vol. 62, no. 2, p. 29503, 2019.

[347] S. Konar *et al.*, "Neutron Star Physics in the Square Kilometer Array Era : An Indian Perspective," *J. Astrophys. Astron.*, vol. 37, p. 36, 2016.

[348] C. A. Raithel and E. R. Most, "Degeneracy in the Inference of Phase Transitions in the Neutron Star Equation of State from Gravitational Wave Data," *Phys. Rev. Lett.*, vol. 130, no. 20, p. 201403, 2023.

[349] C. A. Raithel and E. R. Most, "Tidal deformability doppelgänger: Implications of a low-density phase transition in the neutron star equation of state," *Phys. Rev. D*, vol. 108, no. 2, p. 023010, 2023.

[350] S. Blacker, N.-U. F. Bastian, A. Bauswein, D. B. Blaschke, T. Fischer, M. Oertel, T. Soultanis, and S. Typel, "Constraining the onset density of the hadron-quark phase transition with gravitational-wave observations," *Phys. Rev. D*, vol. 102, no. 12, p. 123023, 2020.

[351] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Bortoluzzi, *et al.*, "Laser Interferometer Space Antenna," *arXiv:1702.00786*, 2017.

[352] J. Luo *et al.*, "TianQin: a space-borne gravitational wave detector," *Class. Quant. Grav.*, vol. 33, no. 3, p. 035010, 2016.

Acknowledgements

I would like to express my sincere gratitude to all those who have contributed to the completion of this doctoral thesis.

First and foremost, I thank my supervisor, Prof. Dr. Horst Stöcker for welcoming me to his group at FIAS. His constant encouragement, unparalleled enthusiasm and incredible support kept me motivated throughout the duration of this journey. Ich bin sehr dankbar, dass ich in seine Gruppe arbeiten konnte.

I would like to thank Dr. Kai Zhou for being a wonderful mentor throughout this endeavor. His invaluable help, enduring patience, constructive feedback, competence and expertise on deep learning were instrumental in the successful completion of this thesis. I would also like to express my deepest appreciation to Dr. Shuzhe Shi and Dr. Lingxiao Wang for their dedicated collaboration. I am grateful to Dr. Jan Steinheimer and Dr. Anton Motornenko for helpful discussions at various stages of this journey.

I would like to thank the DAAD and GSI for their financial support during the past three years. I am also grateful to HGS-HIRe for the travel grants to several conferences and for organizing the soft skills courses. My sincere thanks go to Prof. Dr. Christoph Blume, who has been a part of my PhD committee and has taken time to attend regular meetings. I extend my gratitude to Prof. Dr. Luciano Rezzolla for agreeing to read and review this dissertation.

The completion of this thesis would not have been possible without the love and support from Julia, who was there for me through all the lows and highs since the beginning of this journey. I thank her for all the bouldering sessions, the hikes, the walks, the surprise visits, the planned visits, the invitations, the lunches, the sweets, the cakes, the Christmas cookies, the gifts, the chit-chats, and for being the wonderful friend that she is. I further thank Pia for ensuring I had an effortless start to this journey. In addition, I thank all my colleagues from the third floor for the fun and amicable atmosphere at FIAS. I would like to extend a token of gratitude to Manjunath for frequent discussions and for his pleasant company at several conferences. I am ever thankful to have had the opportunity to share the office with wonderful people over the years, including Julia, Kirill and Wei (along with his wife, Xiaoli, and their little daughter, Yingning). Without the IT assistance from Pavithran, this thesis would have been missing the results it has today. I thank him for

looking out not only for our well-being at FIAS, but also for our careers in future.

I am deeply indebted to my friends, Tanmayi, Priyanka, Mythreyi, Shradha and Ittu, most of them who live continents away, yet manage to make their presence felt in my life.

My heartfelt thanks go to Sebastian for everything that he is and has been to me. I thank him for proof-reading this thesis, for translating the abstract of this thesis to German, for the interesting physics discussions, for helping me rehearse talks for several conferences, for motivating me to complete this thesis, for his love, his understanding, his time, and for more than words can say. I also thank Daniela, Beate and Karl for making the last year and a half delightful.

Finally, I am incredibly grateful for my family, Amma, Nana and Shreyas, for their immense love, for their significant sacrifices and for their unconditional support. I would not have come this far without them. I would like to credit this achievement to them.