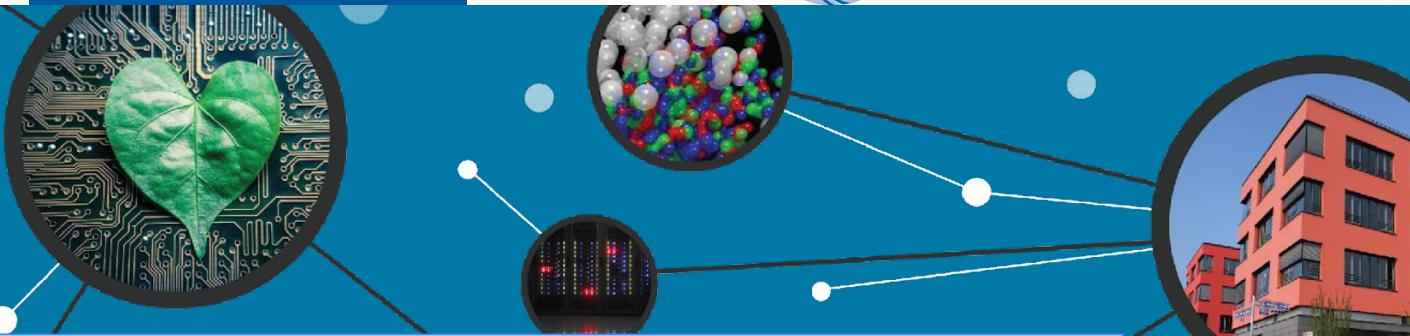
DAAD Deutscher Akademischer Austauschdienst
German Academic Exchange Service
HGS-HIRe for FAIR
Helmholtz Graduate School for Hadron and Ion Research
GOETHE UNIVERSITÄT FRANKFURT AM MAIN
FIAS Frankfurt Institute for Advanced Studies



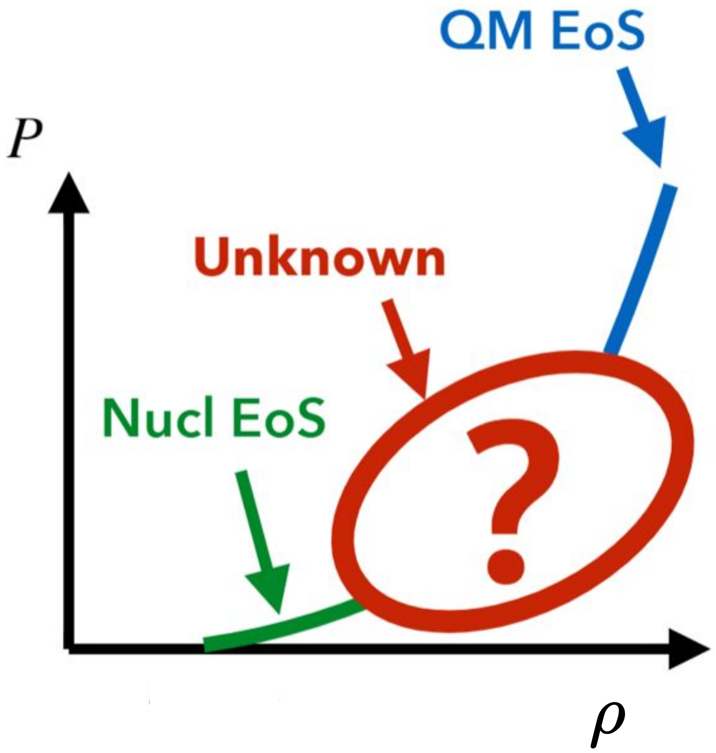
Deducing the neutron star equation of state using deep learning for inverse problem-solving

Workshop on the QCD equation of state in dense matter HIC and astrophysics, Kerala

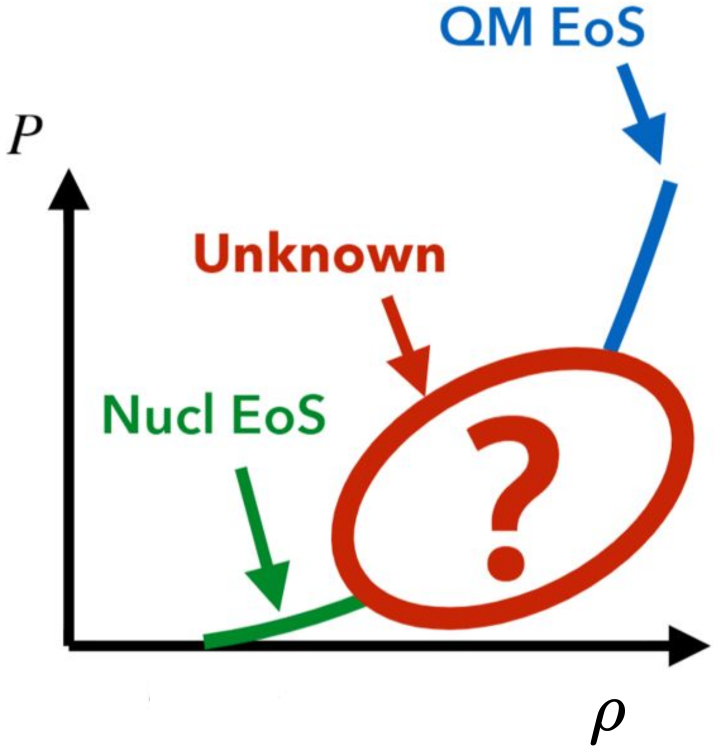
Shriya Soma

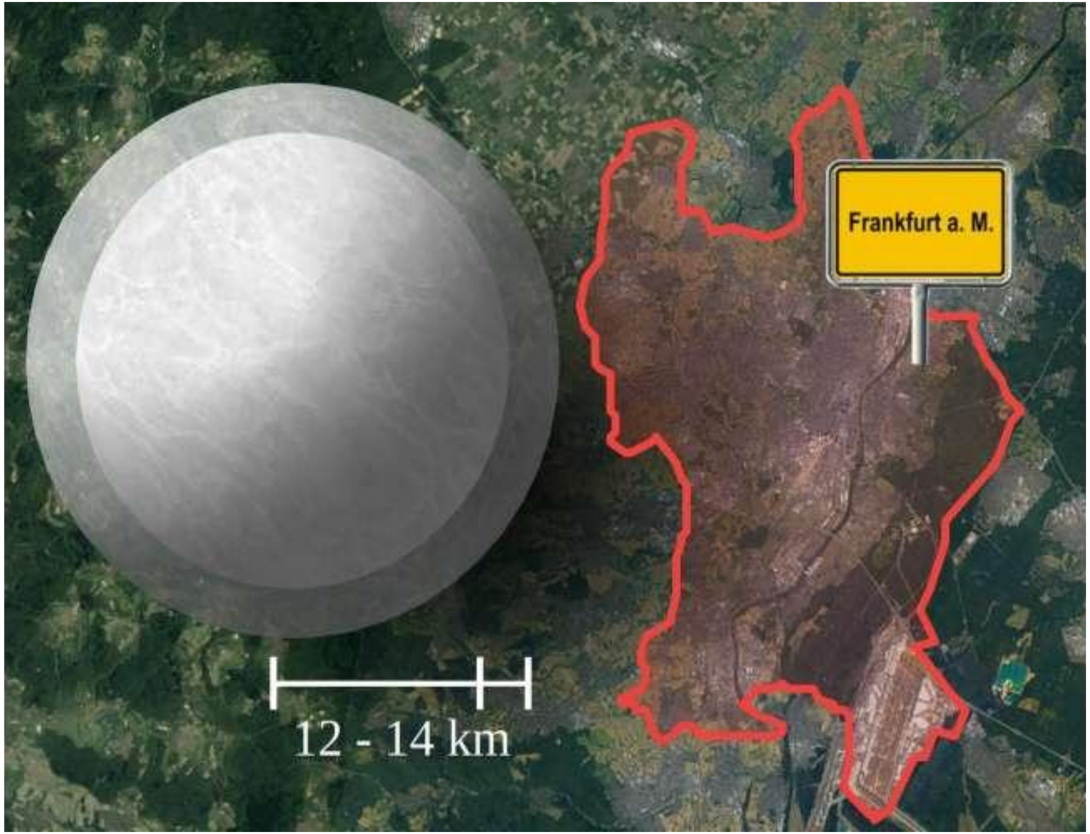
30 March, 2023

Dense Matter EoS



Dense Matter EoS

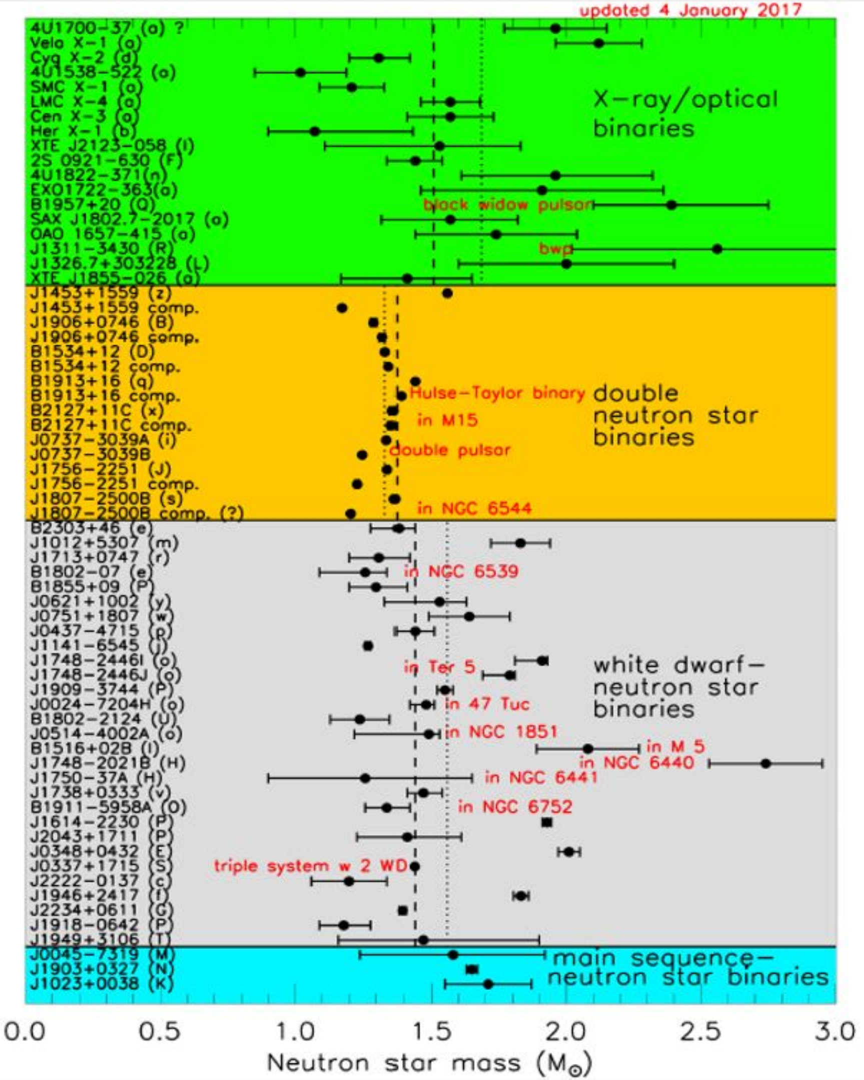




NS Observables

- Mass
- Radius
- Tidal Deformability

Antoniadis *et al.*, Science **340** (2013)
Cromartie *et al.*, NatAs **4** (2019) 72
Riley *et al.*, ApJL **887** (2019) L21
Riley *et al.*, ApJL **918** (2021) L27
Abbott *et al.*, PRX **9** (2019) 011001
Coughlin *et al.*, **480** (2018) 3



[Lattimer]

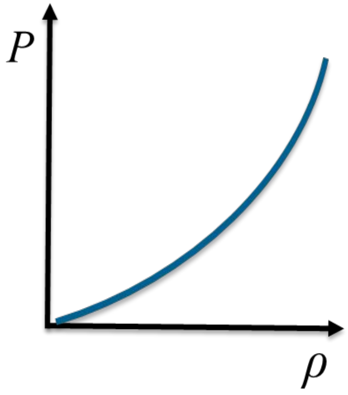
Inverse Problems

- Reconstructing the dense matter EoS using mass-radius observations of neutron stars (NSs)

- Analyzing GWs for inferring the properties of NSs

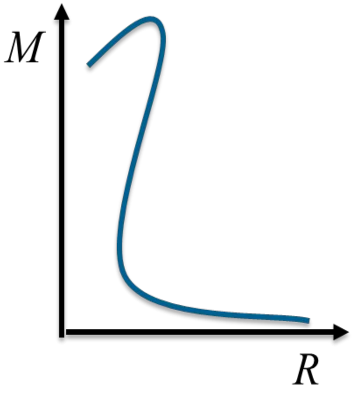
TOV Equations: From EoS to Stellar Structure

EoS → TOV Equations → MR Curve

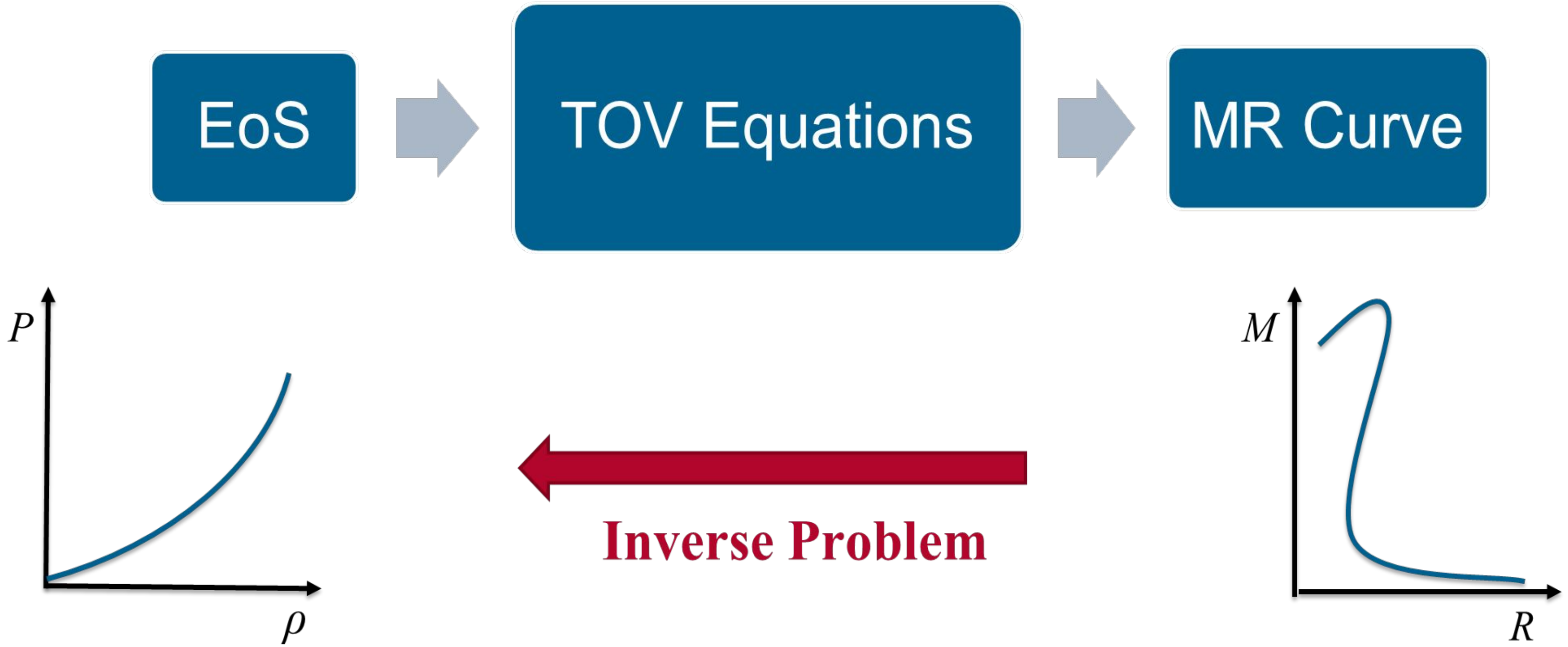


$$-\frac{dP}{dr} = \frac{\left[\epsilon(r) + P(r)\right]\left[M(r) + 4\pi r^3 P(r)\right]}{r[r - 2M(r)]}$$

$$\frac{dM(r)}{dr} = 4\pi r^2 \epsilon(r),$$



MR Observables to EoS: An Inverse Problem



The Bayesian Approach

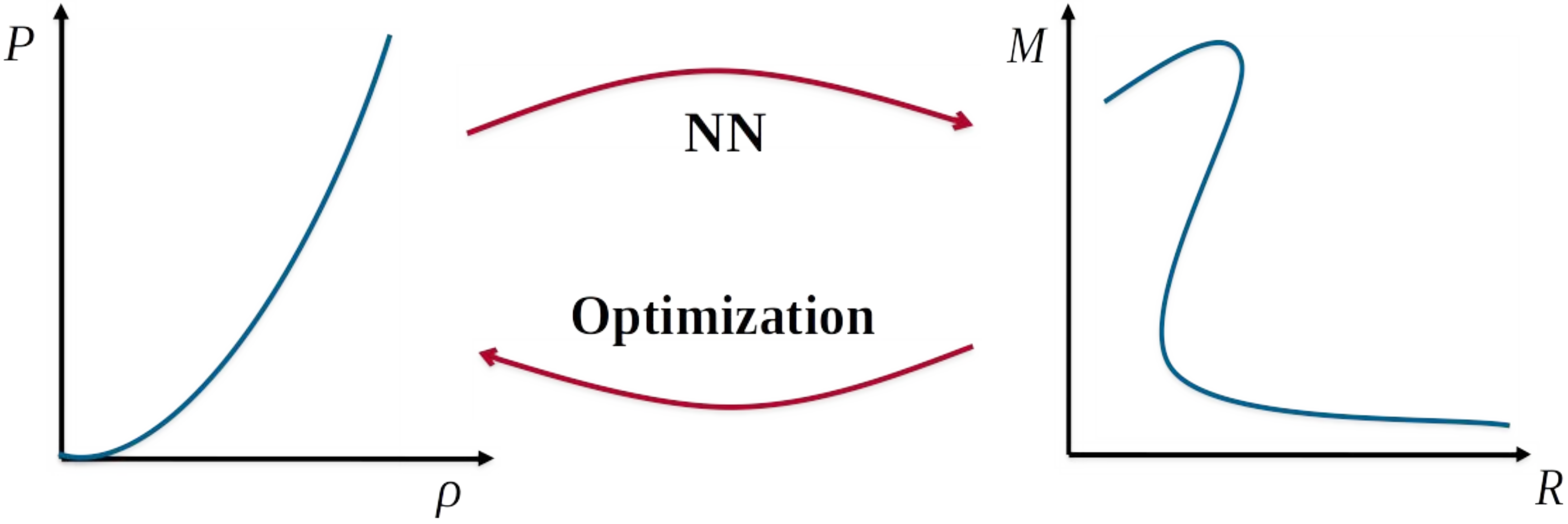
$$P(\text{EoS} \mid M\text{-}R) = \frac{P(M\text{-}R \mid \text{EoS})\, P(\text{EoS})}{P(M\text{-}R)}$$

Steiner *et al.*, ApJL **765** (2013) L5

Raithel *et al.*, ApJ **844** (2017) 156

Automatic Differentiation

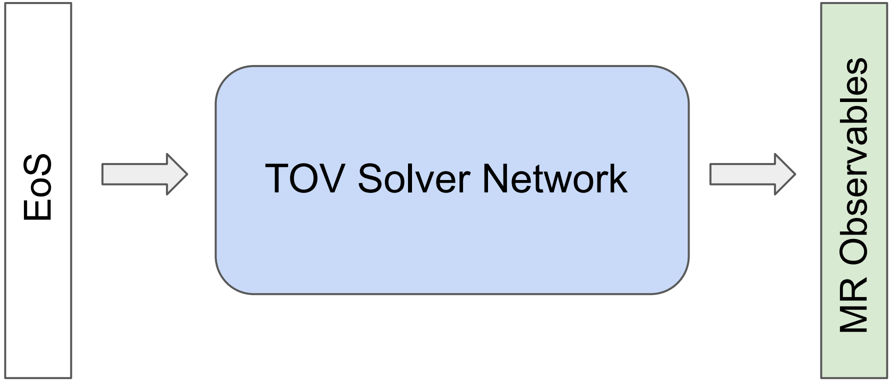
- Train a neural network (NN) to output the MR curve from an EoS



- Optimize the input (EoS) to obtain the desired output (MR curve)

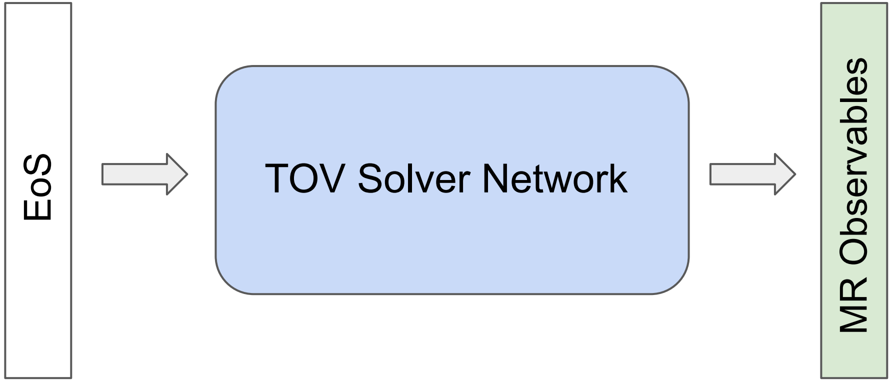
Procedure

1. Train a NN model to solve TOV Equations – TOV Solver Network



Procedure

1. Train a NN model to solve TOV Equations – TOV Solver Network

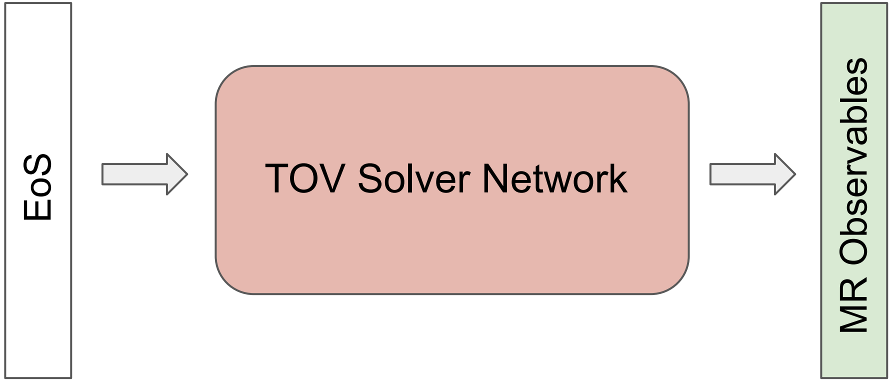


$$P(\text{EoS}\,|\,M\text{-}R) = \frac{\boxed{P(M\text{-}R\,|\,\text{EoS})}\,P(\text{EoS})}{P(M\text{-}R)}$$

Procedure

1. Train a NN model to solve TOV Equations – TOV Solver Network

2. Fix the weights of TOV Solver (freeze training)

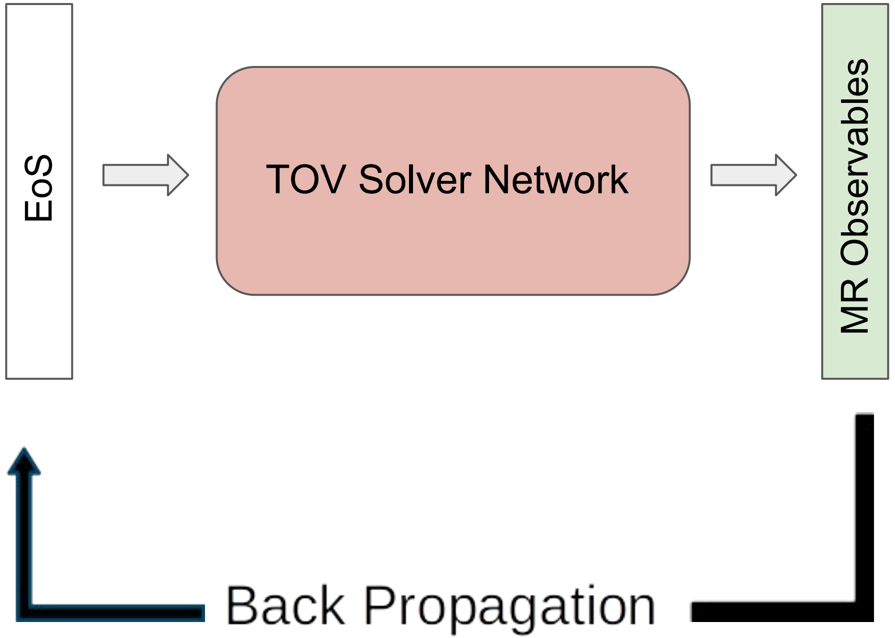


Procedure

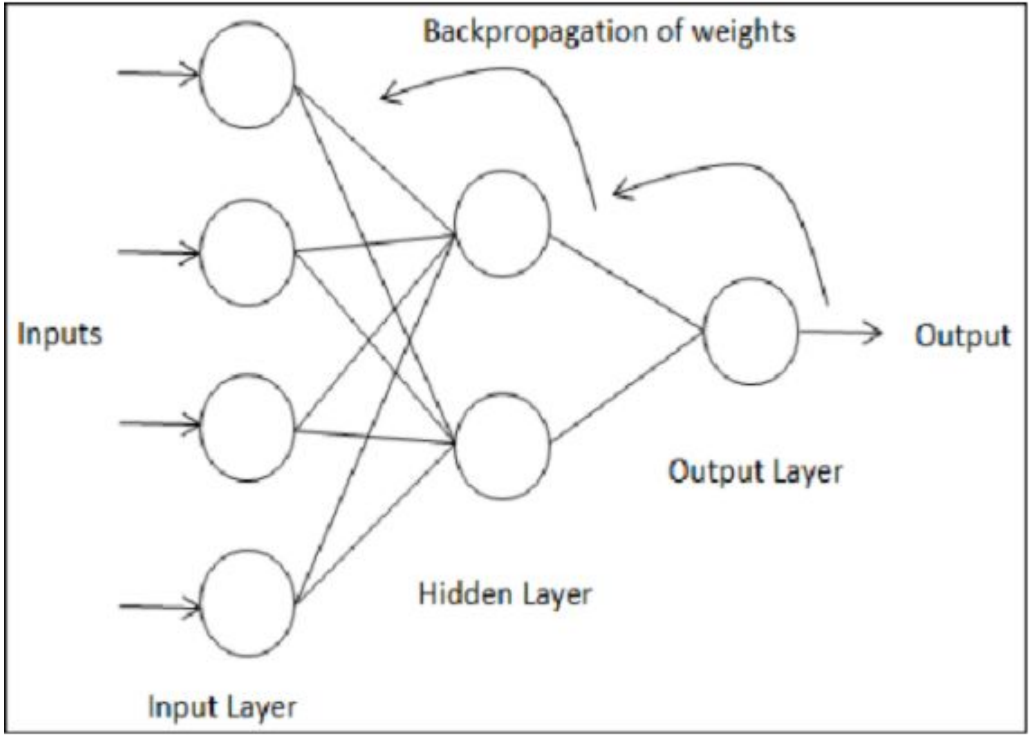
1. Train a NN model to solve TOV Equations – TOV Solver Network

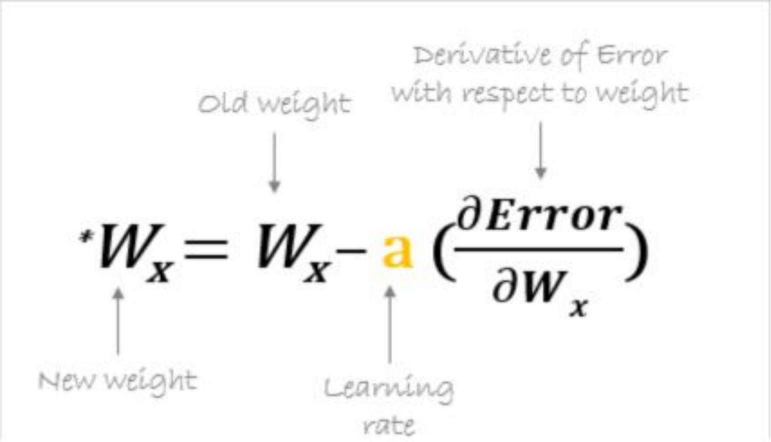
2. Fix the weights of TOV Solver (freeze training)

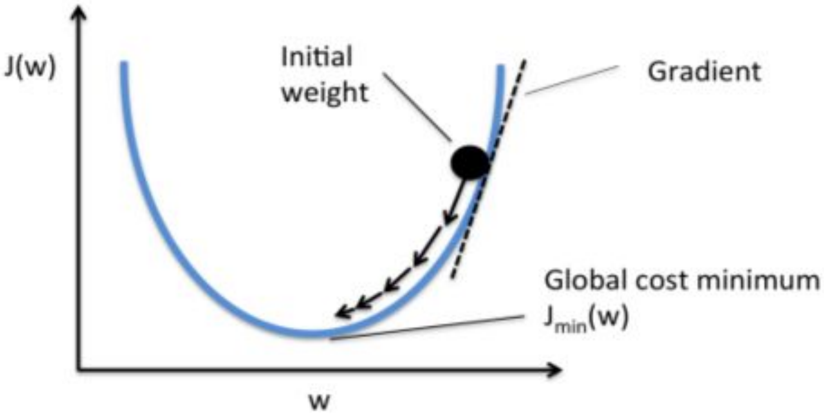
3. Optimize the input layer (EoS)



Back Propagation





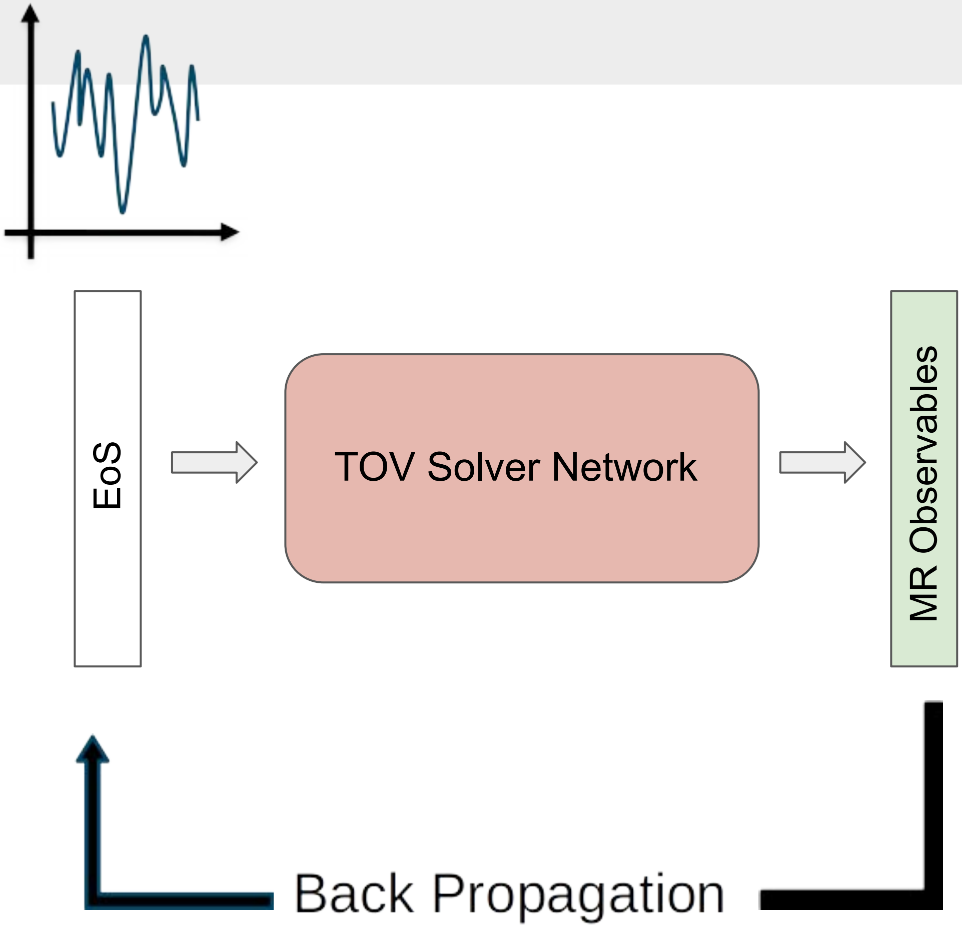


Procedure

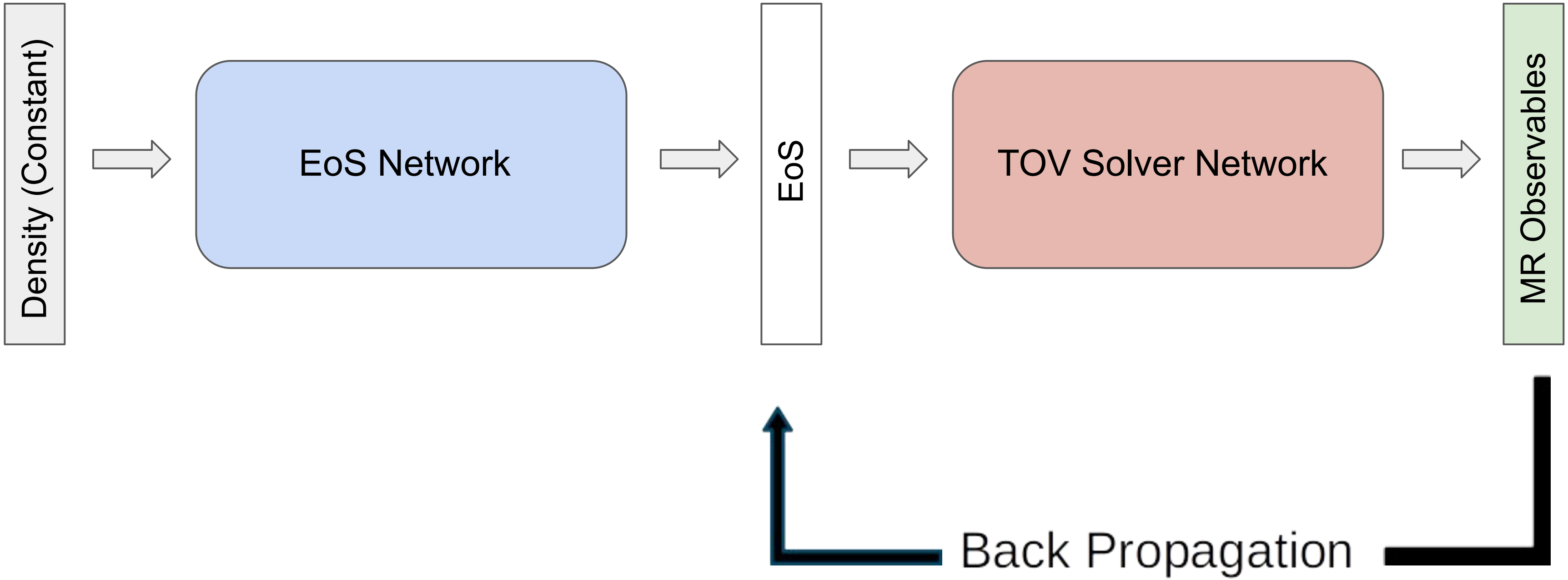
1. Train a NN model to solve TOV Equations – TOV Solver Network

2. Fix the weights of TOV Solver (freeze training)

3. Optimize the input layer (EoS)



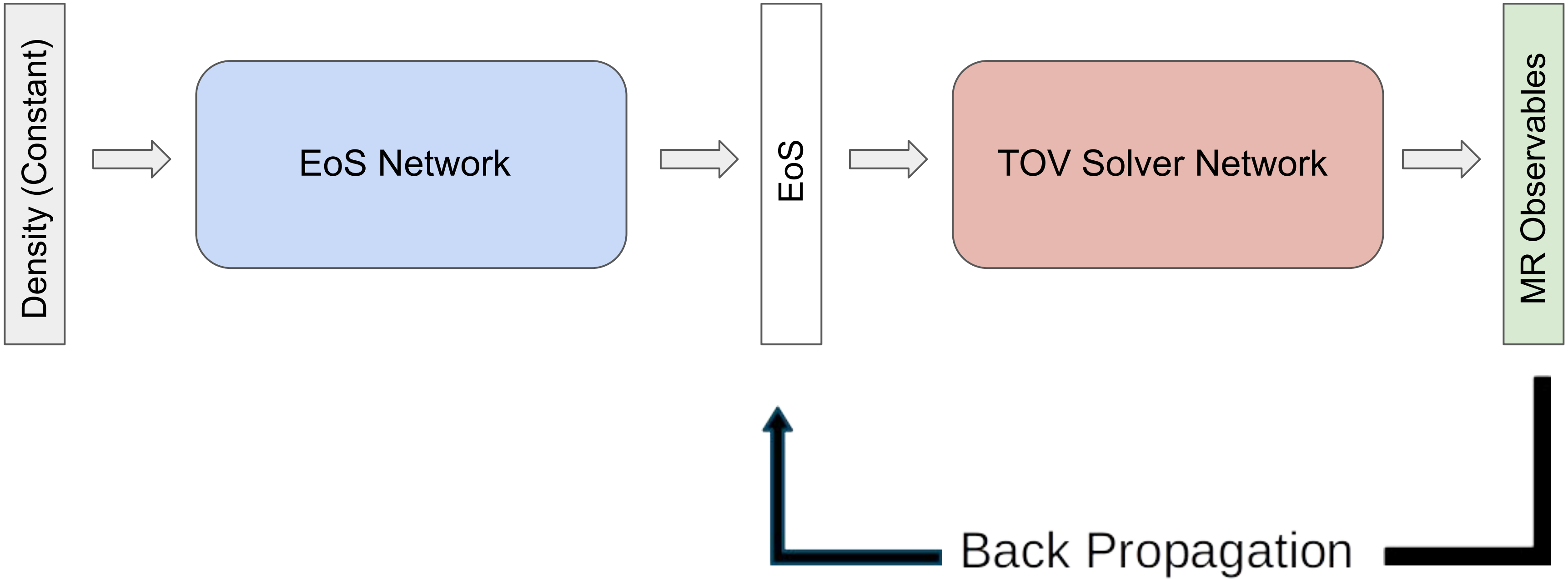
Procedure



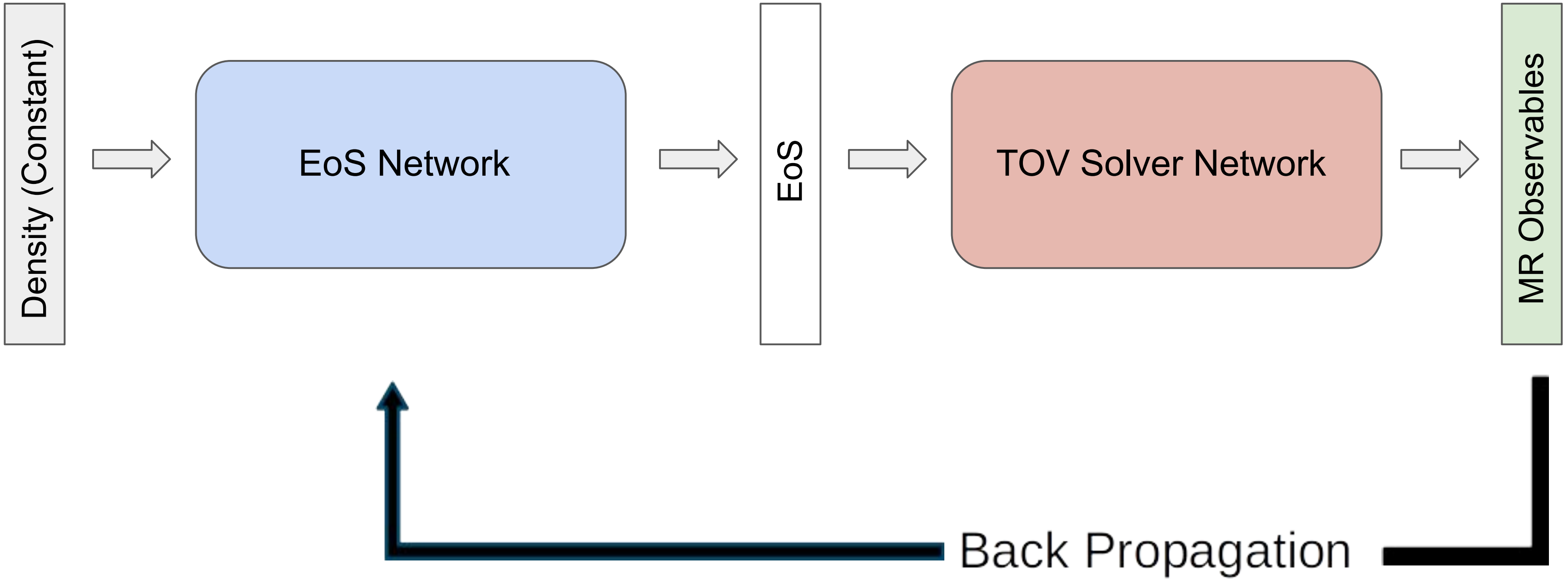
Density (Constant)
EoS Network
EoS
TOV Solver Network
MR Observables
Back Propagation

Procedure

$$P(\mathrm{EoS} \mid M\text{-}R) = \frac{P(M\text{-}R \mid \mathrm{EoS})\,\boxed{P(\mathrm{EoS})}}{P(M\text{-}R)}$$



Procedure


Density (Constant)
EoS Network
EoS
TOV Solver Network
MR Observables
Back Propagation

Data Preparation

- $\rho < \rho_0$: SLy / PS / DD2
- $\rho > \rho_0$: Piecewise Polytropes at (1.0, 1.4, 2.2, 3.3, 4.9, 7.4) ρ_0 [Raithel *et al.*, ApJ **831** (2016) 44]

$$P = K_i \rho^{\Gamma_i} \quad ; \quad d\frac{\epsilon}{\rho} = -P d\frac{1}{\rho}$$

where,

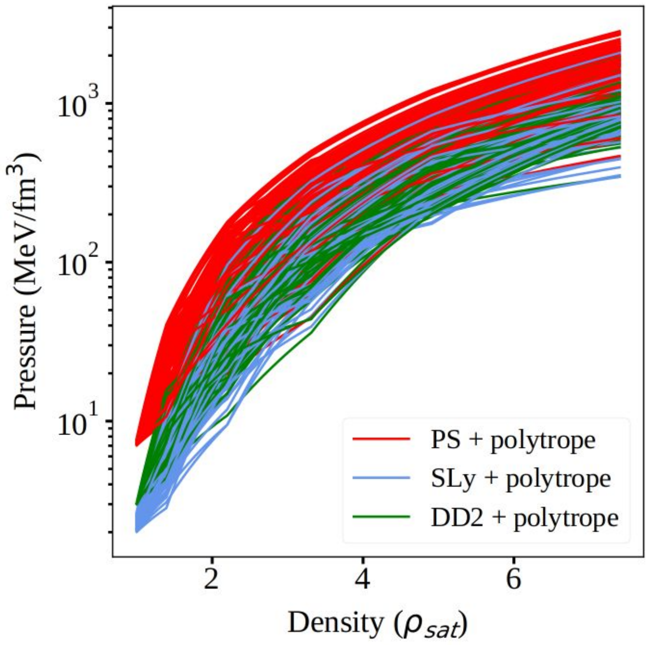
$$K_i = \frac{P_{i-1}}{\rho_{i-1}^{\Gamma_i}}$$

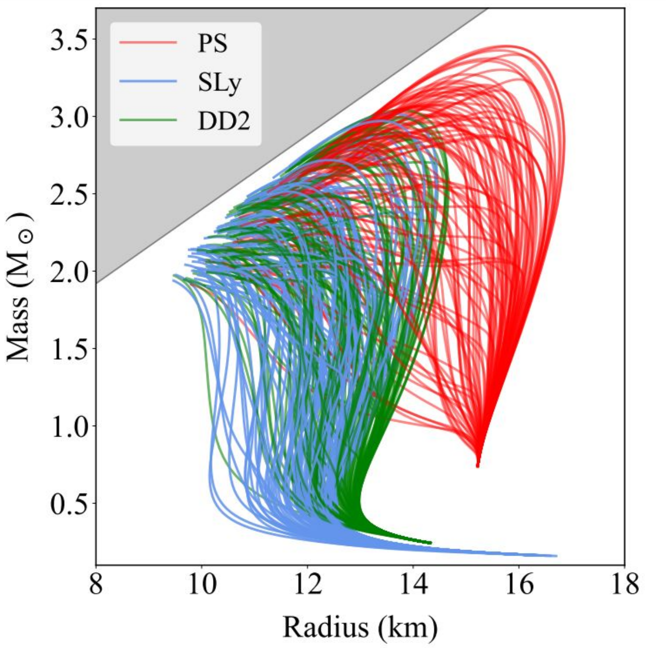
and $\Gamma_i \in [1, \min\{5, \Gamma_{luminal}\}]$;

$$\frac{dP}{d\epsilon} \leq 1 \quad ; \quad \Gamma = \Gamma_{luminal} \quad \text{when} \quad \frac{dP}{d\epsilon} = 1$$

TOV Solver Network: Training

- $\rho < \rho_0$: SLy / PS / DD2
- $\rho > \rho_0$: Piecewise Polytropes at (1.0, 1.4, 2.2, 3.3, 4.9, 7.4) ρ_0 [Raithel *et al.*, *ApJ* **831** (2016) 44]

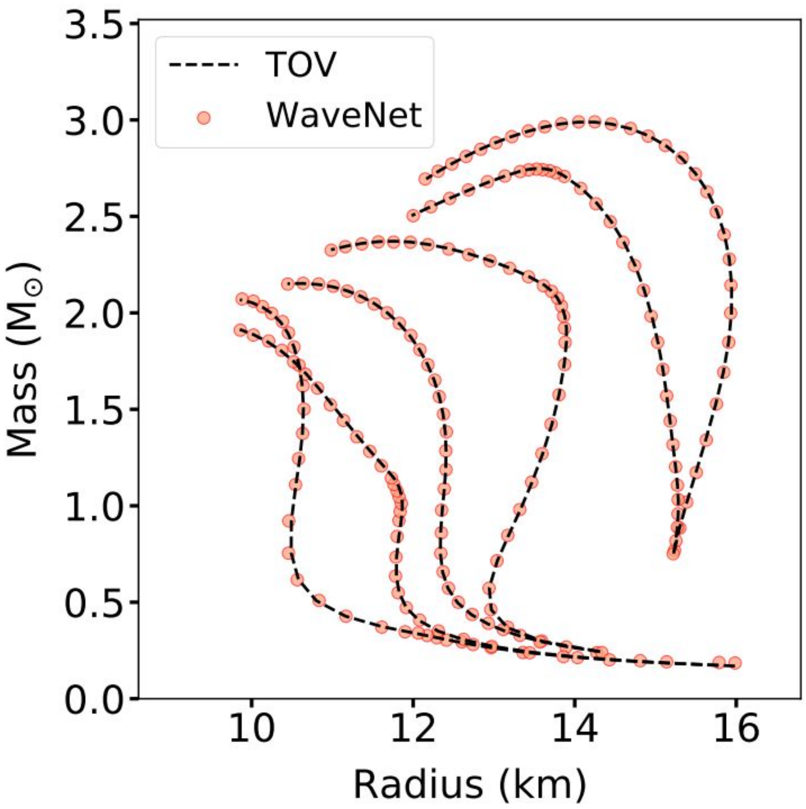




EoSs Generated: 3 x 100,000

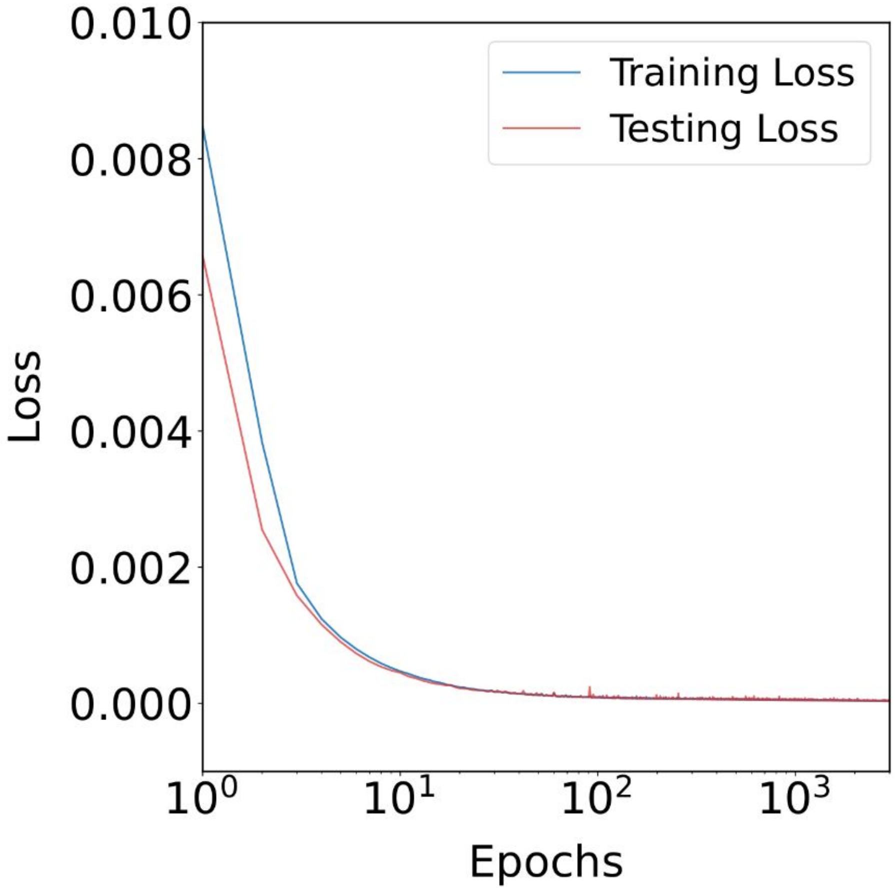
On exclusion of MR curves with maximum mass < 1.9 Solar Mass: 228,569 EoSs

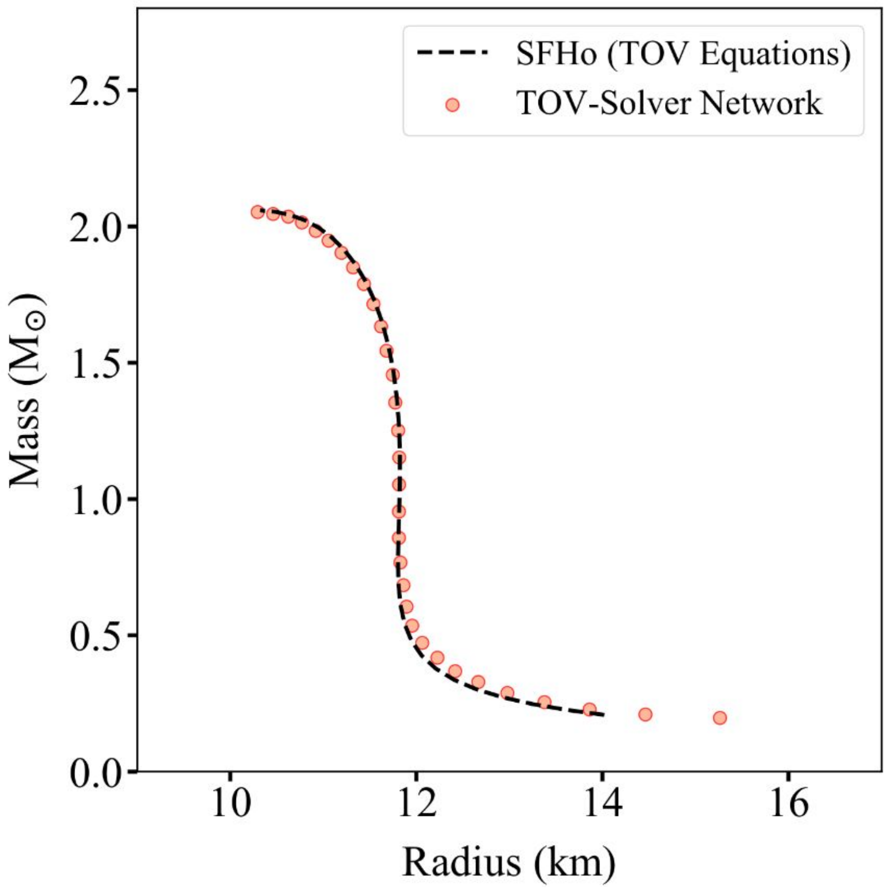
TOV Solver Network: Performance



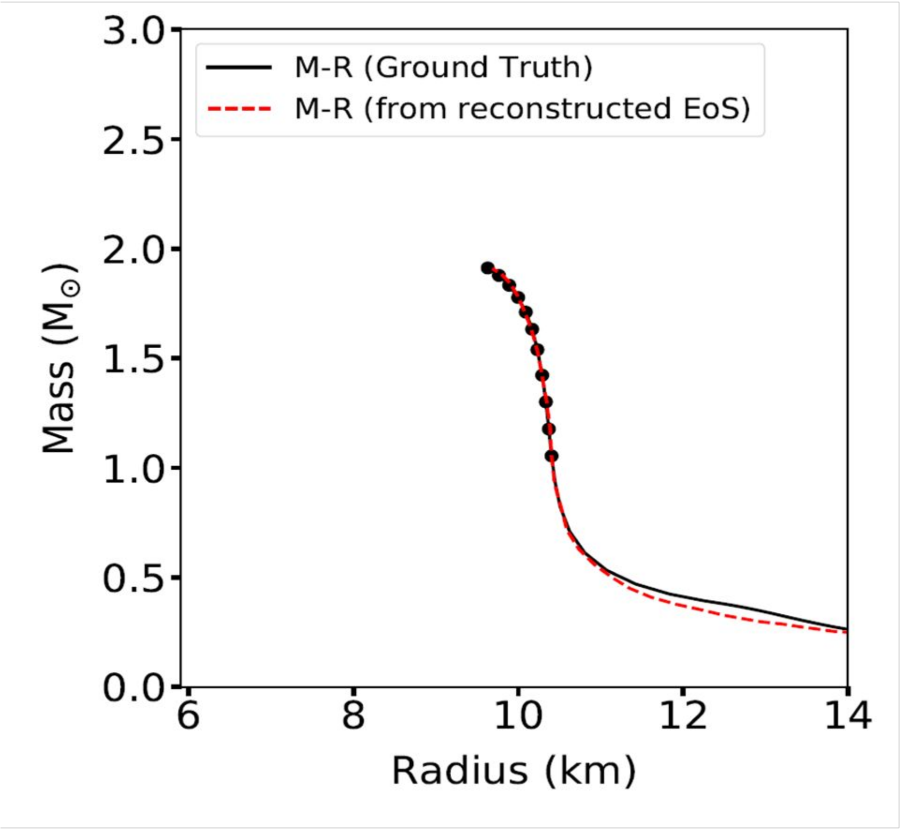
Accuracy : 99.9%

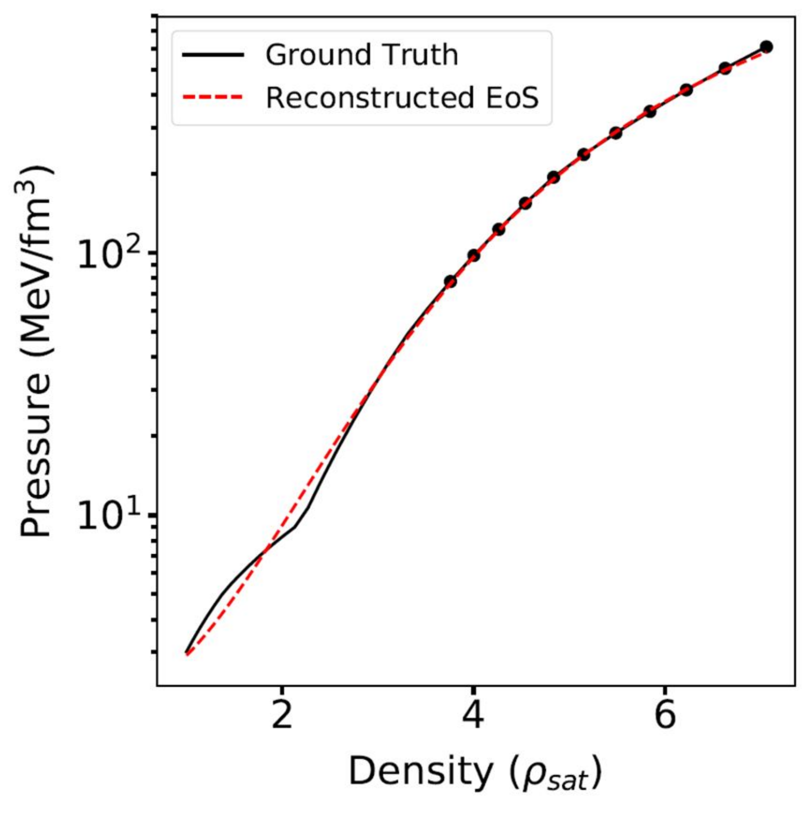
TOV Solver Network: Performance on SFHo



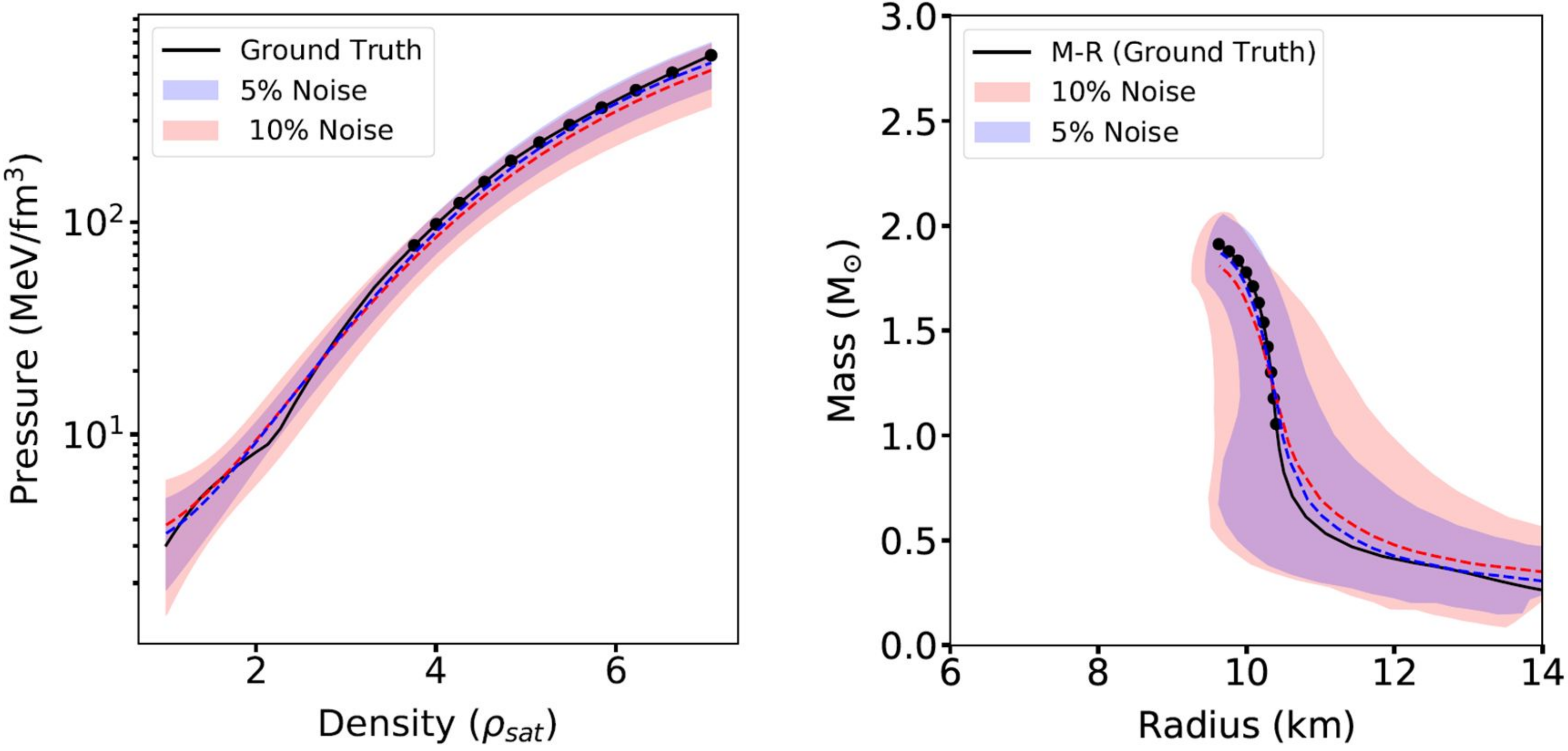


Mock data: An Ideal Case

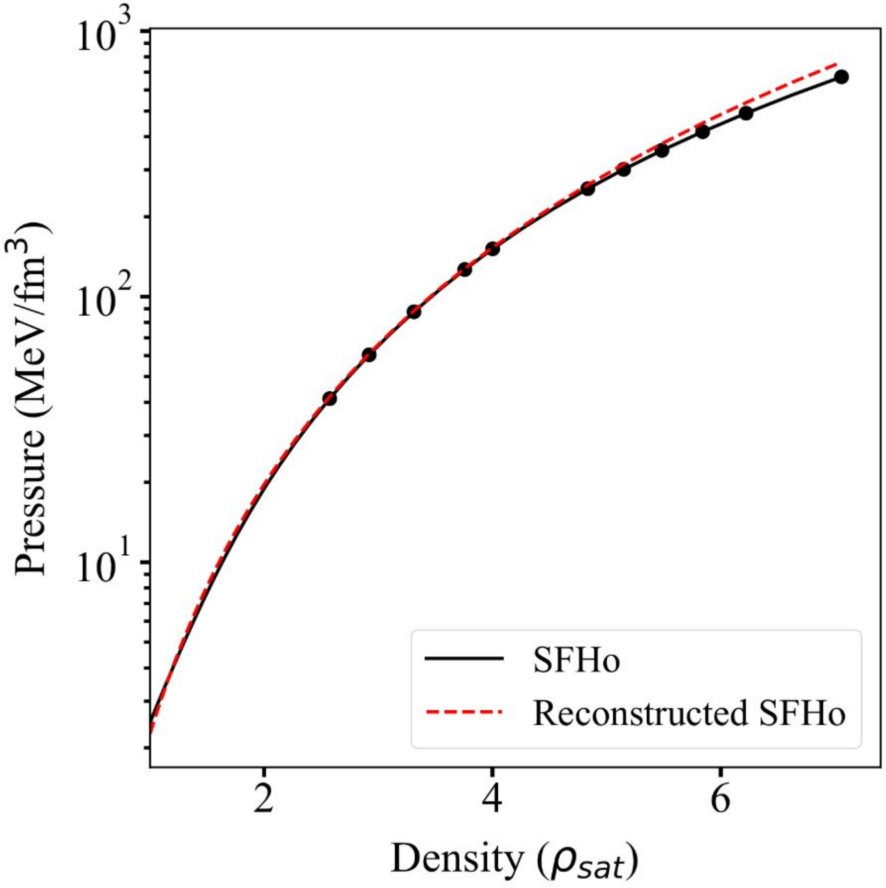


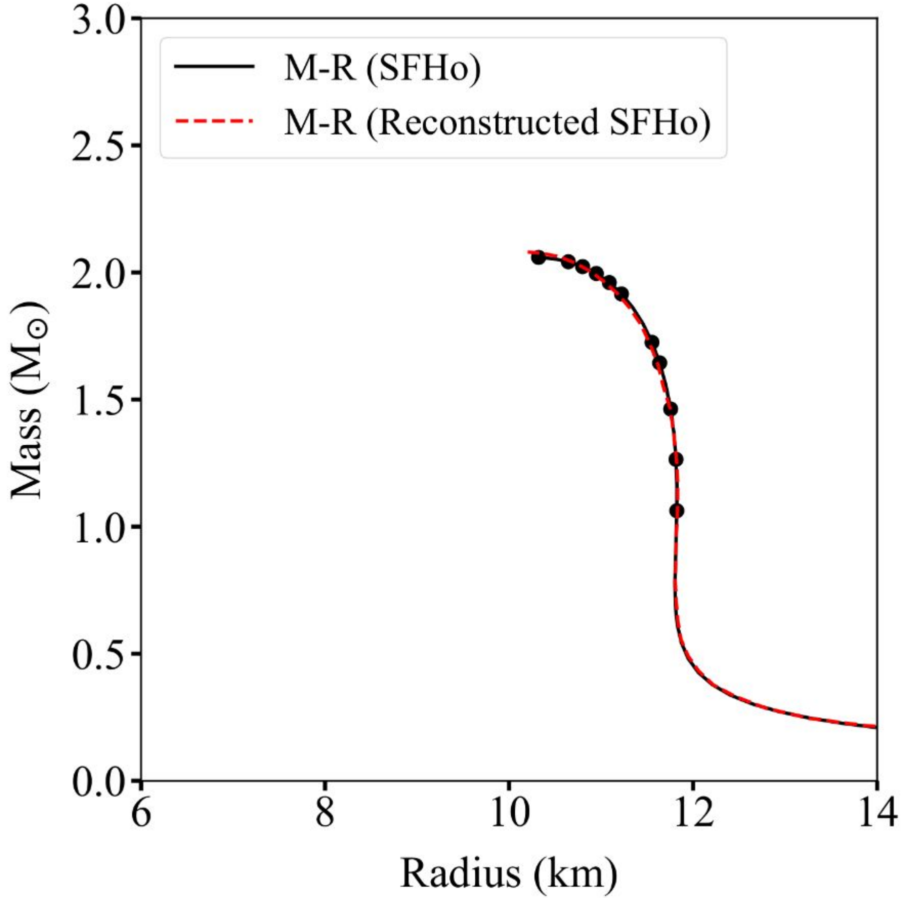


Mock data: A Realistic Scenario



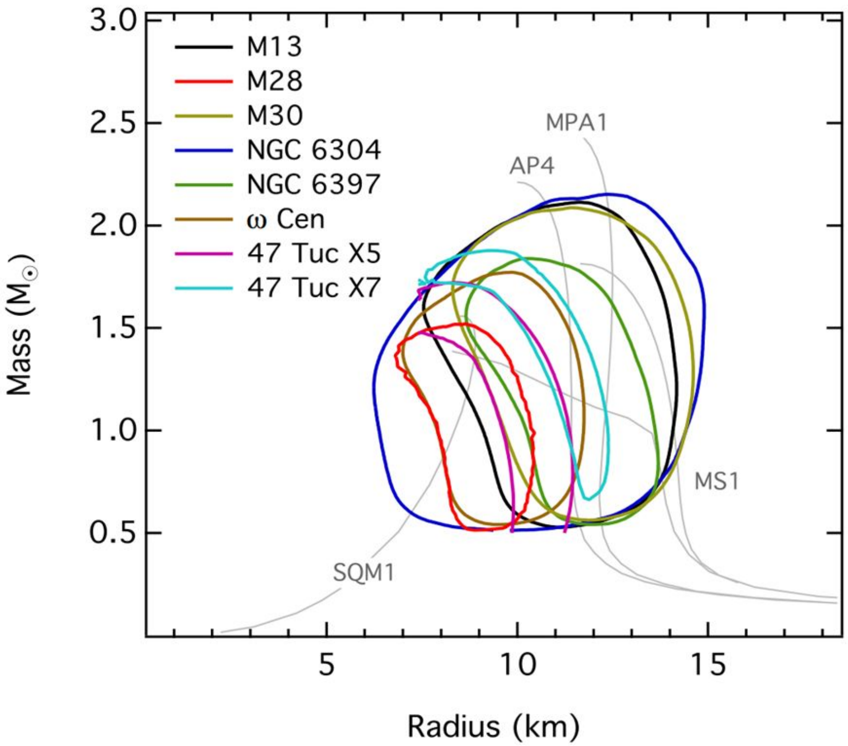
EoS Network: Performance on SFHo



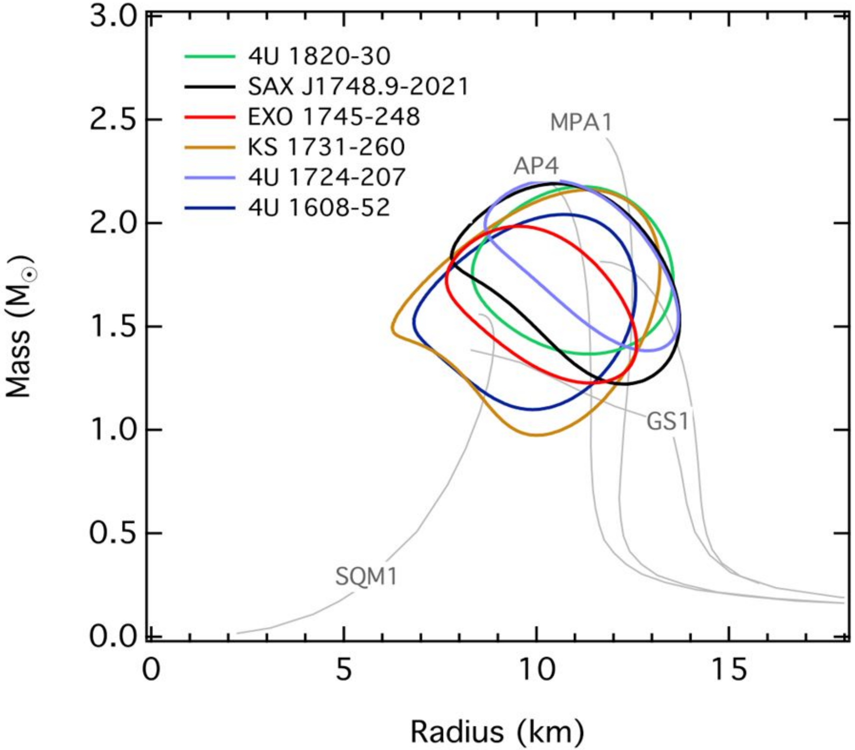


NS Radius and Mass measurements

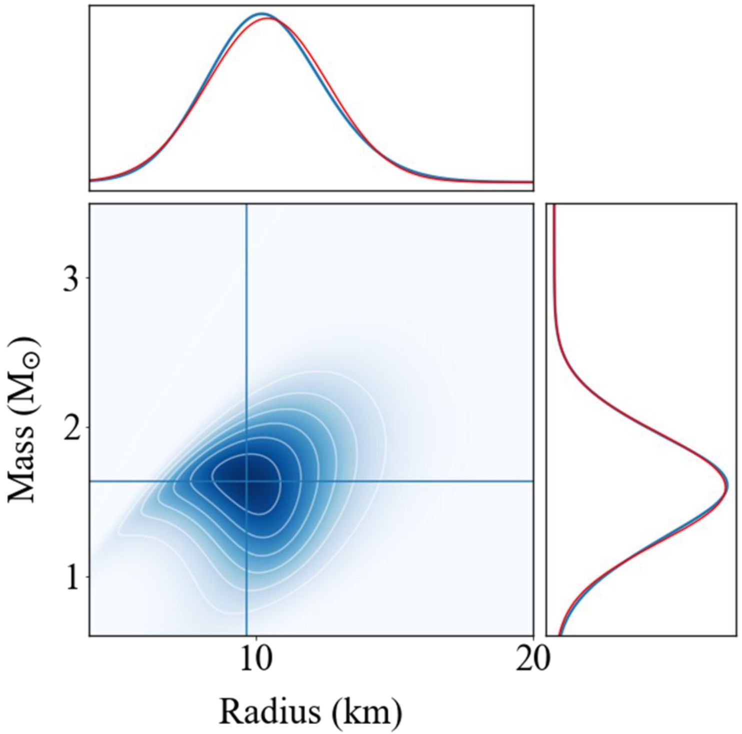
Özel *et al.*, ApJ **820** (2016) 28
Bogdanov *et al.*, ApJ **831** (2016) 184

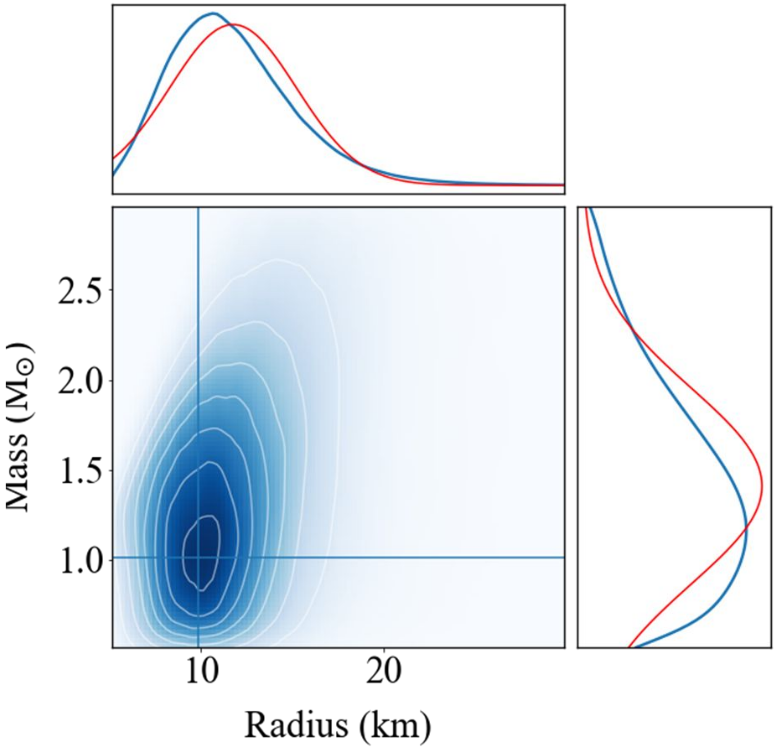


Riley *et al.*, ApJL **887** (2019) L21
Riley *et al.*, ApJL **918** (2021) L27

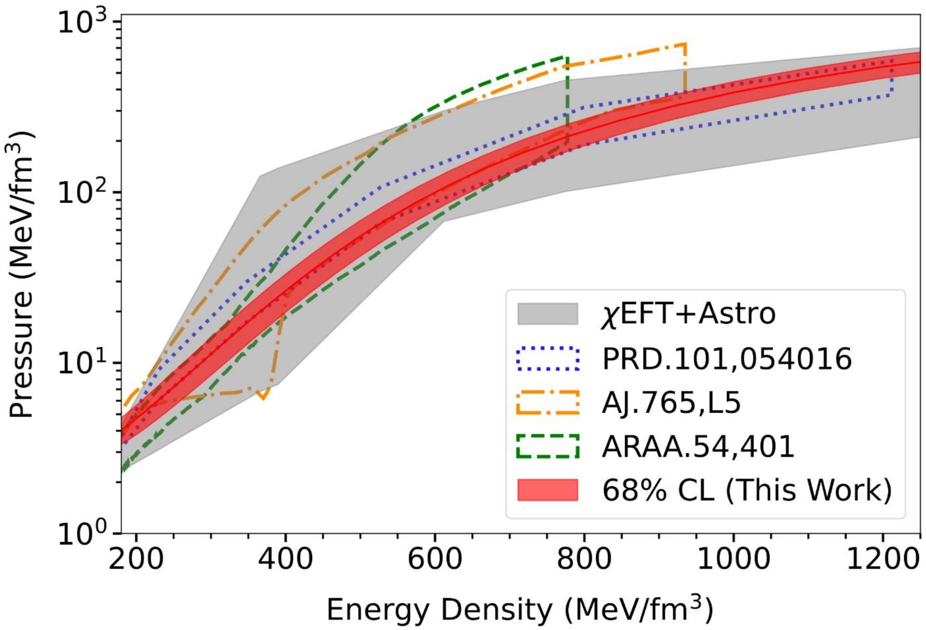


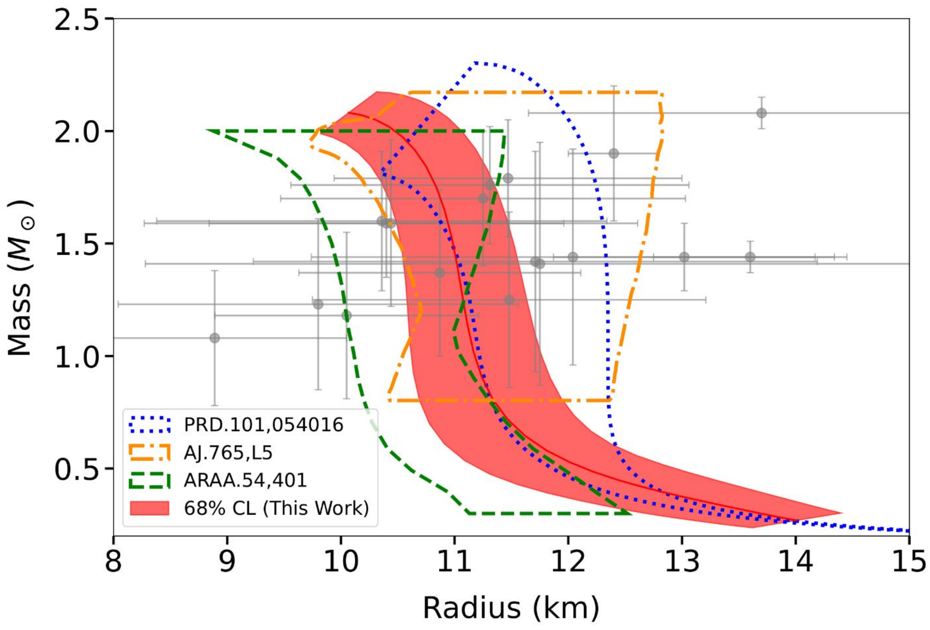
Normal Distribution Fits to MR data





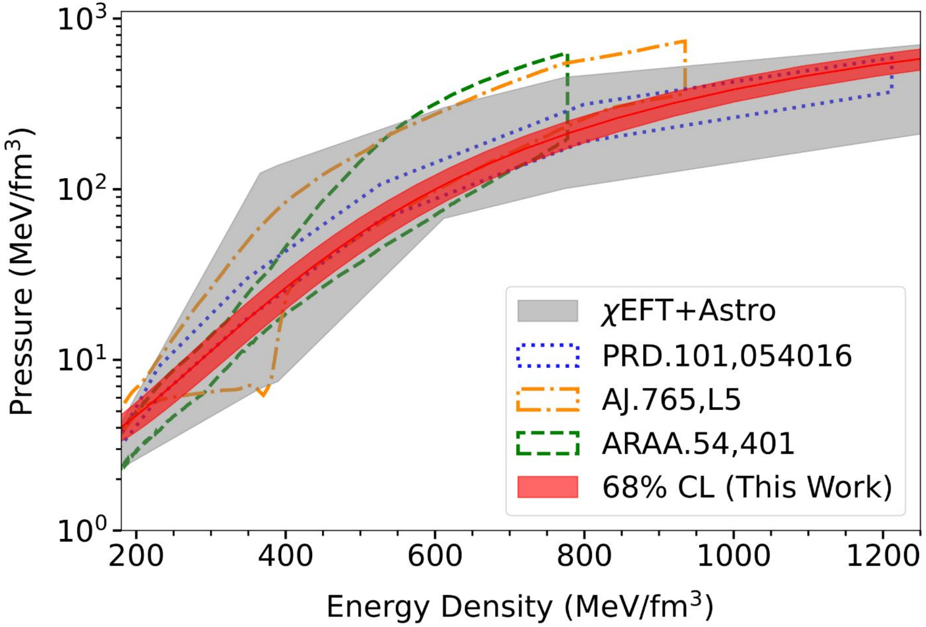
Comparison with Previous Works

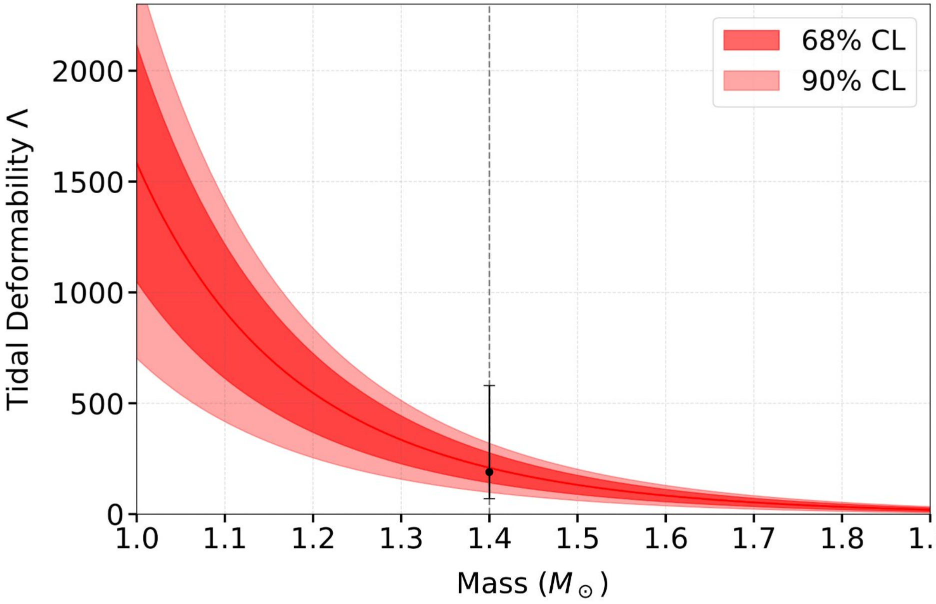




PRD : Fujimoto *et al.*
AJ : Steiner *et al.*
ARAA : Özel *et al.*

Comparison with Previous Works





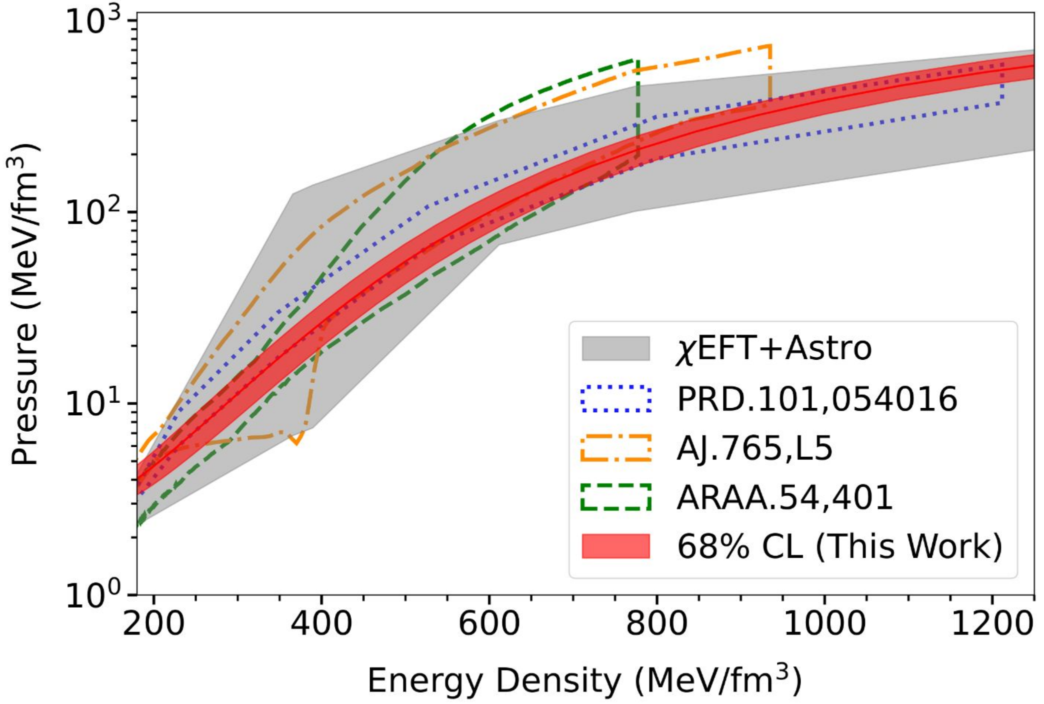
PRD : Fujimoto *et al.*
AJ : Steiner *et al.*
ARAA : Özel *et al.*

$$\Lambda_{1.4} = 190^{+390}_{-120} \quad \text{at the 90\% level}$$

Abbott *et al.*, PRL **121** (2018) 161101

Summary

S.S, K. Zhou, L. Wang, S. Shi, H. Stöcker: JCAP 08(2022)071
S.S, K. Zhou, L. Wang, S. Shi, H. Stöcker: arXiv: 2209.08883



- Trained a NN to replace the TOV Equations
- Inverted the NN to optimize the input layer (EoS)
- Reconstructed the EoS from Real Observations (post successful tests on mock data)
- Consistent with Λ limits from GW170817

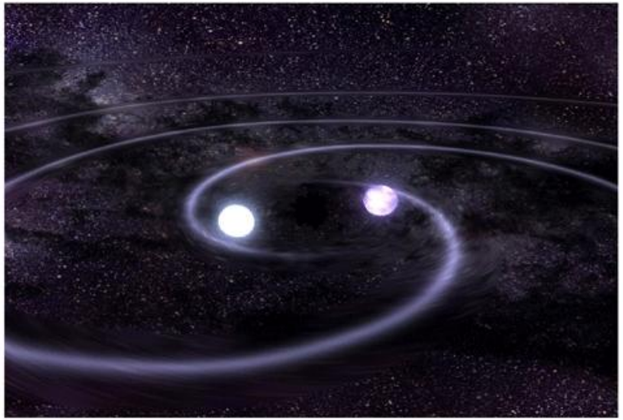
Inverse Problems

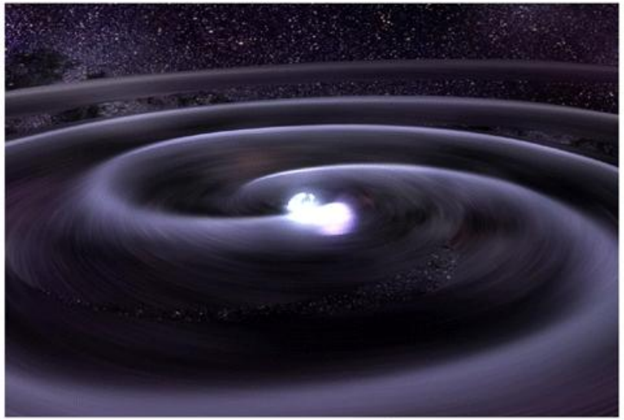
- Reconstructing the dense matter EoS using mass-radius observations of neutron stars (NSs)

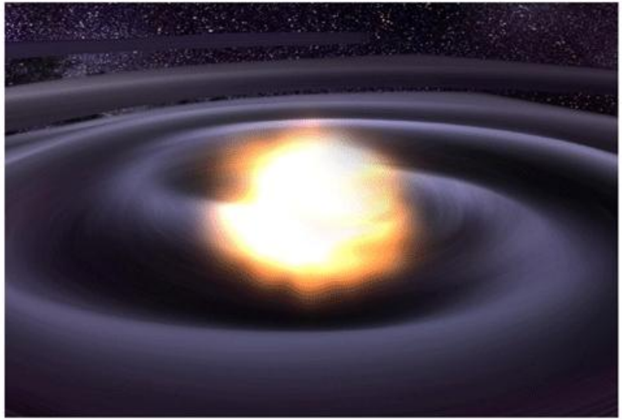
- Analyzing GWs for inferring the properties of NSs

GWs : Ripples in the fabric of space-time

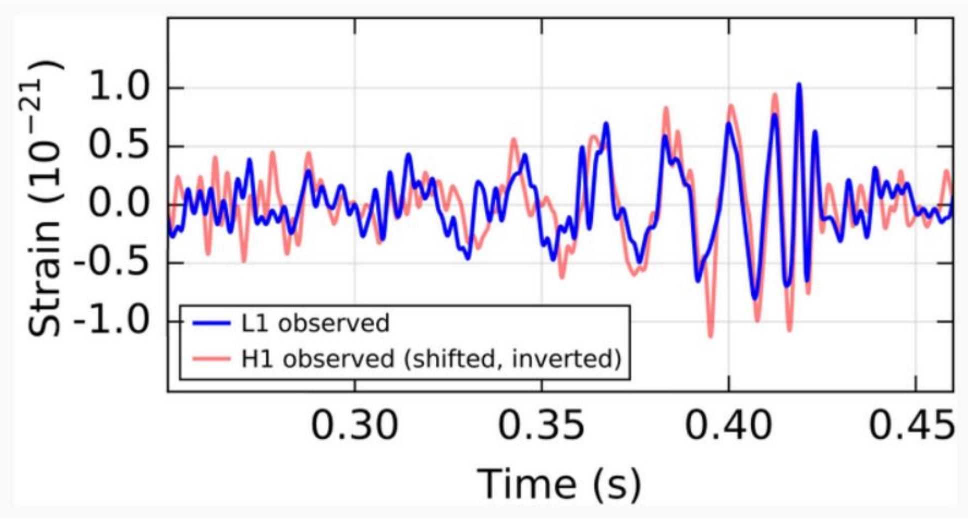
Massive accelerating objects disrupt space-time and emit Gravitational waves.





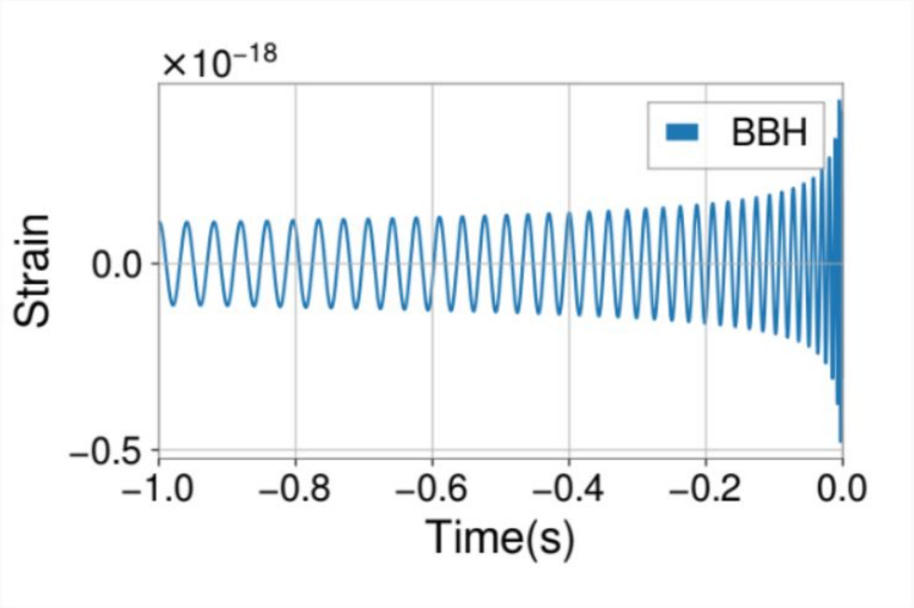


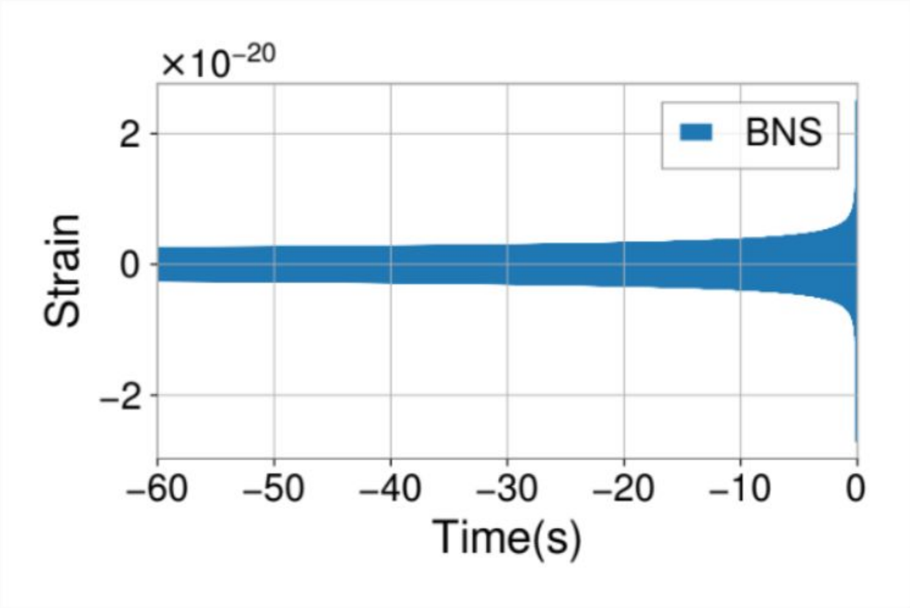
Listening to Cosmic Whispers - aLIGO & Virgo



LIGO Livingston Credit: Caltech/MIT/LIGO Lab

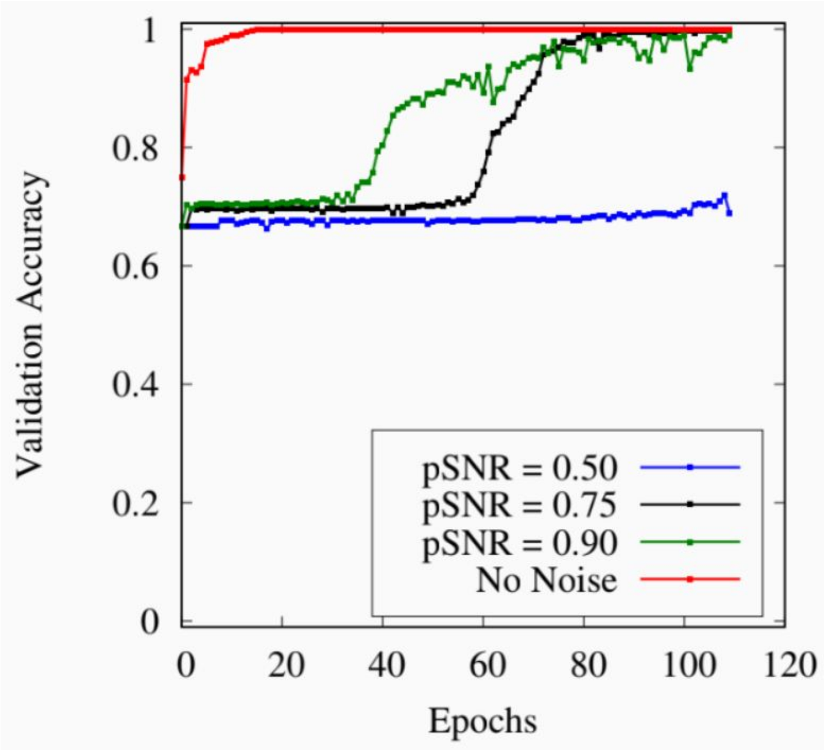
Simulation of GW Waveforms from mergers of BBHs and BNSs





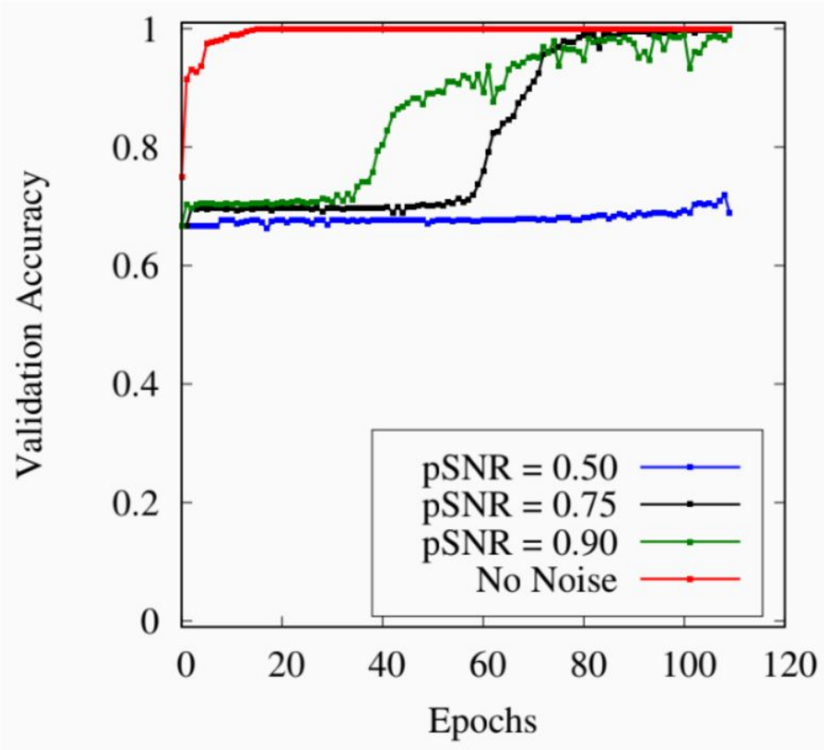
- LALSuite Library to generate simulated signals
- Model: IMRPhenomPv2_NRTidal (Frequency Domain)
- Inputs: m1, m2, $\Lambda 1$, $\Lambda 2$ (Note: For BHs, $\Lambda = 0$)

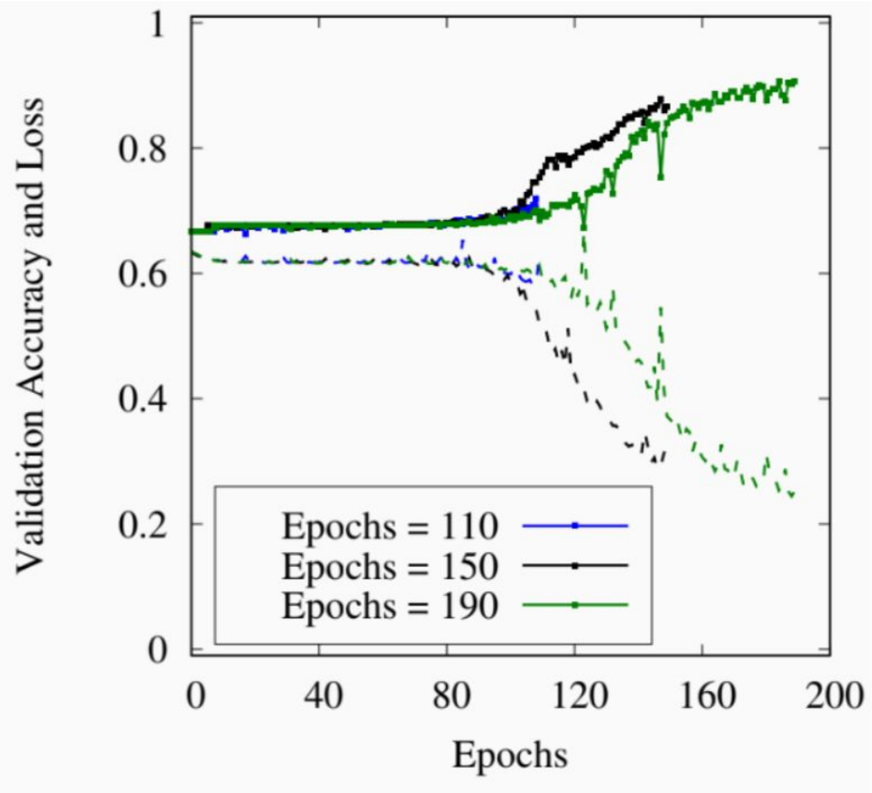
Classification of BBH signals, BNS signals, and Noise



- For pSNR ≥ 0.75, accuracy ≥ 98%
- For pSNR = 0.50, train longer?

Classification of BBH signals, BNS signals, and Noise.

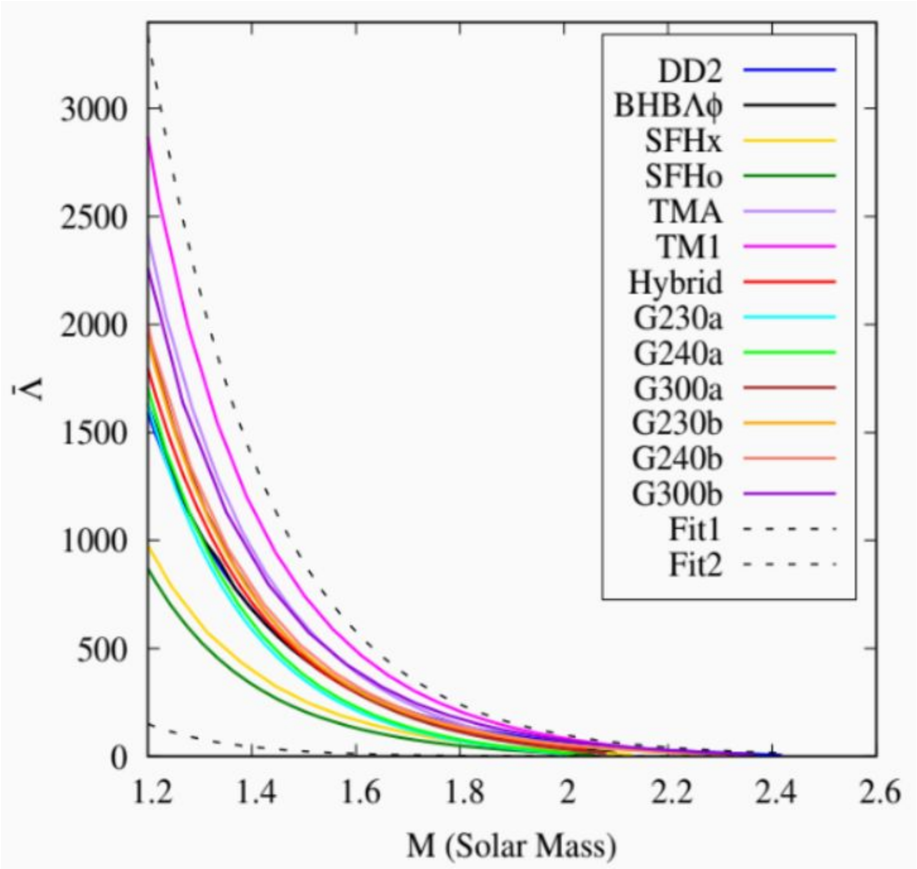




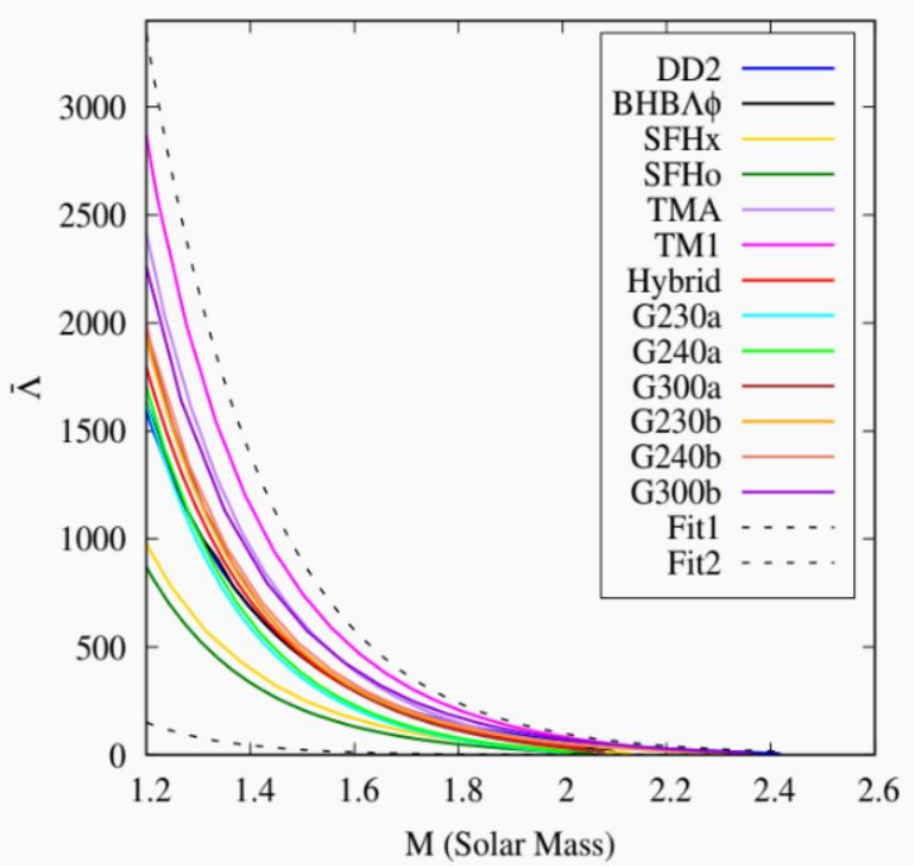
- **Classification**

- **Regression?**

Regression - mass & tidal deformation



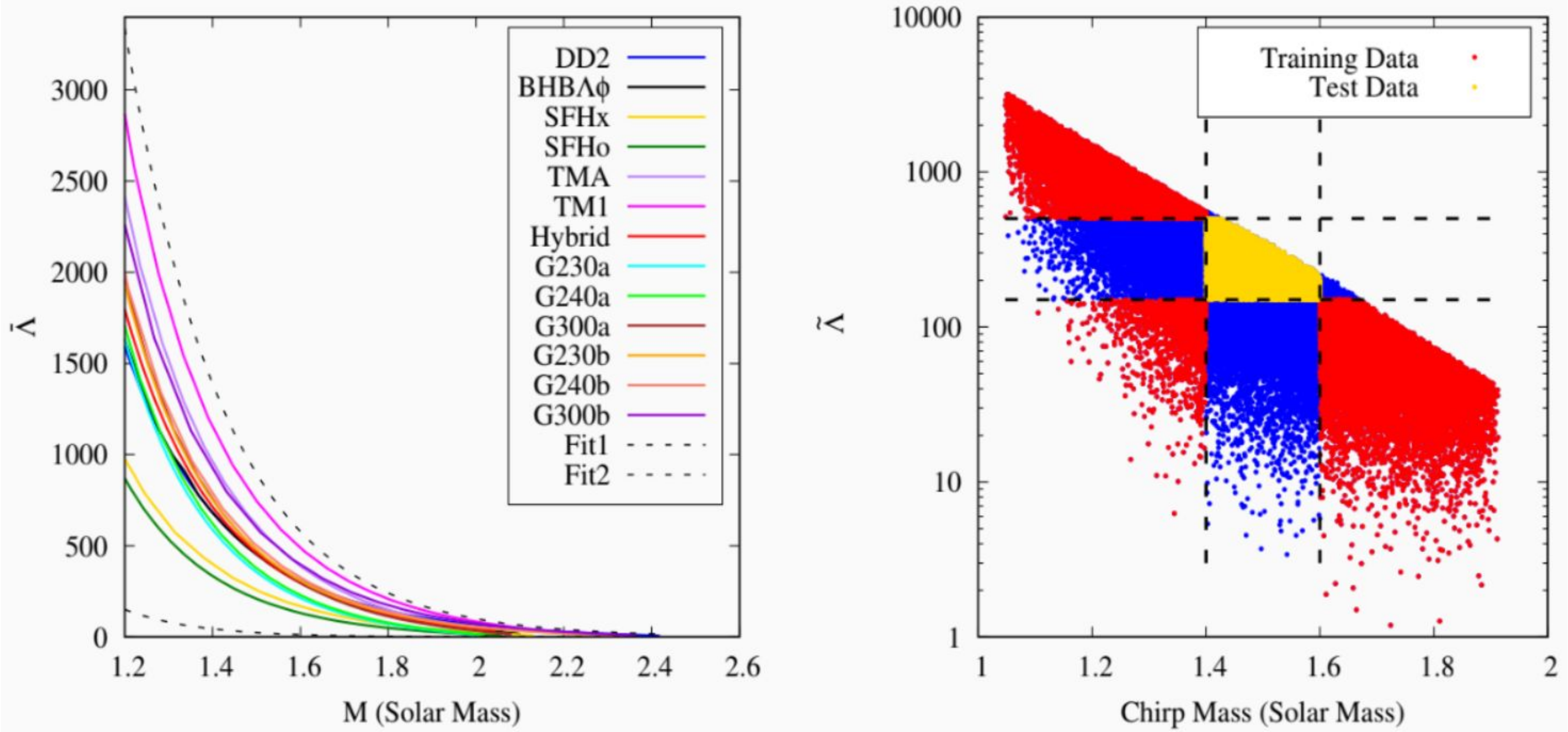
Regression - chirp mass & combined tidal deformation



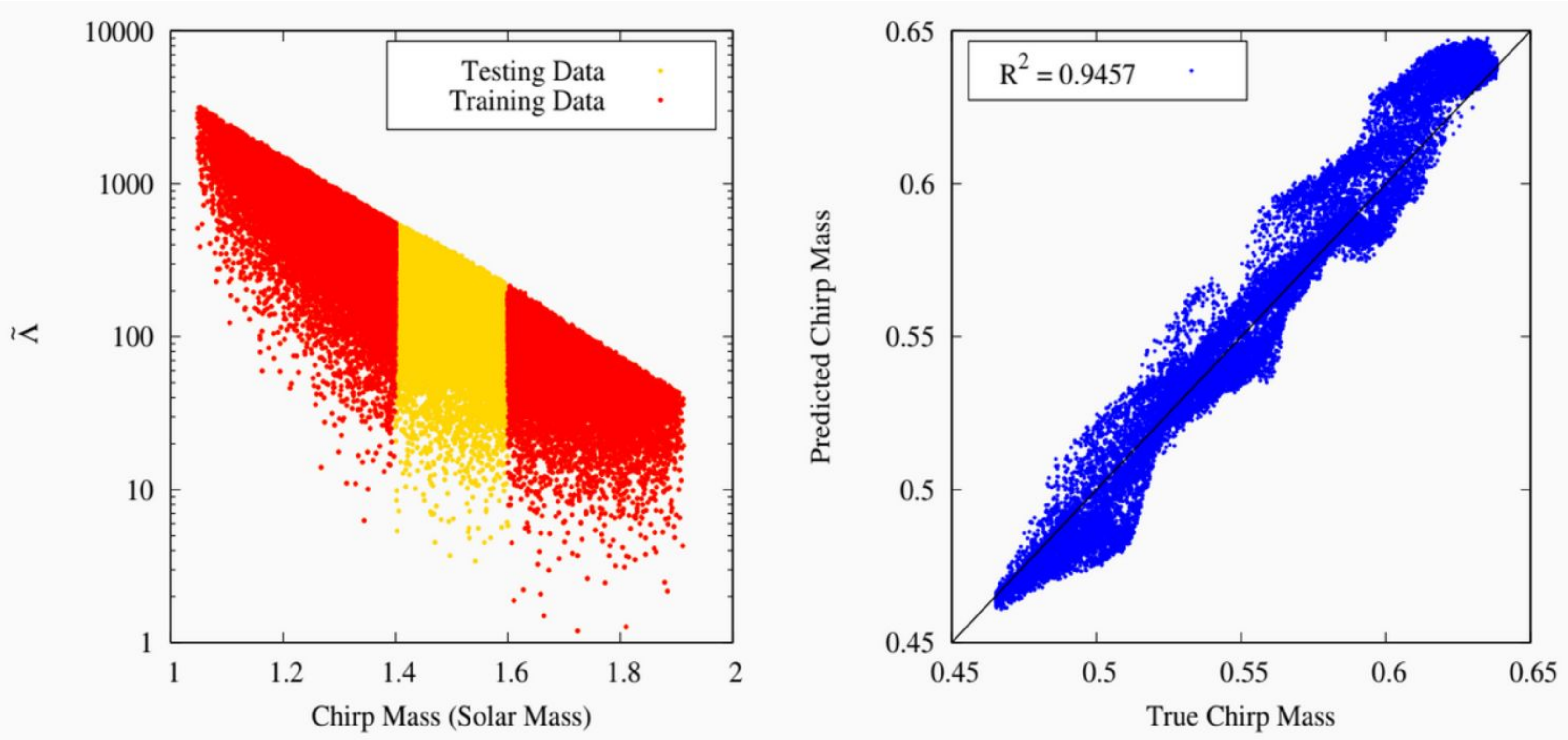
$$\textit{Chirp Mass}, \ \mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$$

$$\widetilde{\Lambda} = \frac{16}{13} \frac{[(M_1 + 12M_2)M_1^4\bar{\Lambda}_1 + (M_2 + 12M_1)M_2^4\bar{\Lambda}_2]}{(M_1 + M_2)^5}$$

Regression - chirp mass & combined tidal deformation



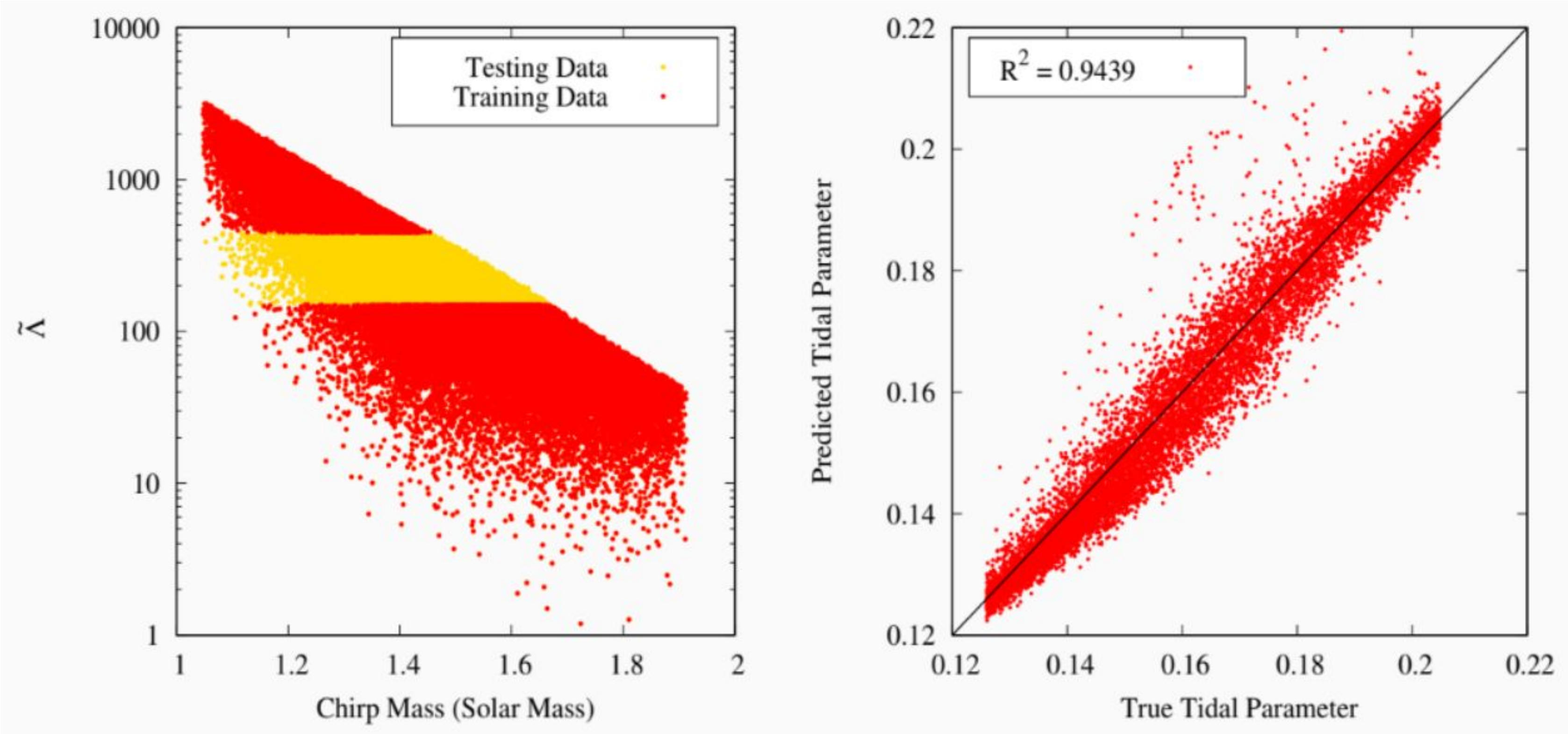
Predicting Chirp Mass



Channel 1: Absolute Value, Channel 2: Argument

Training Examples: 55055, Test Examples: 19945

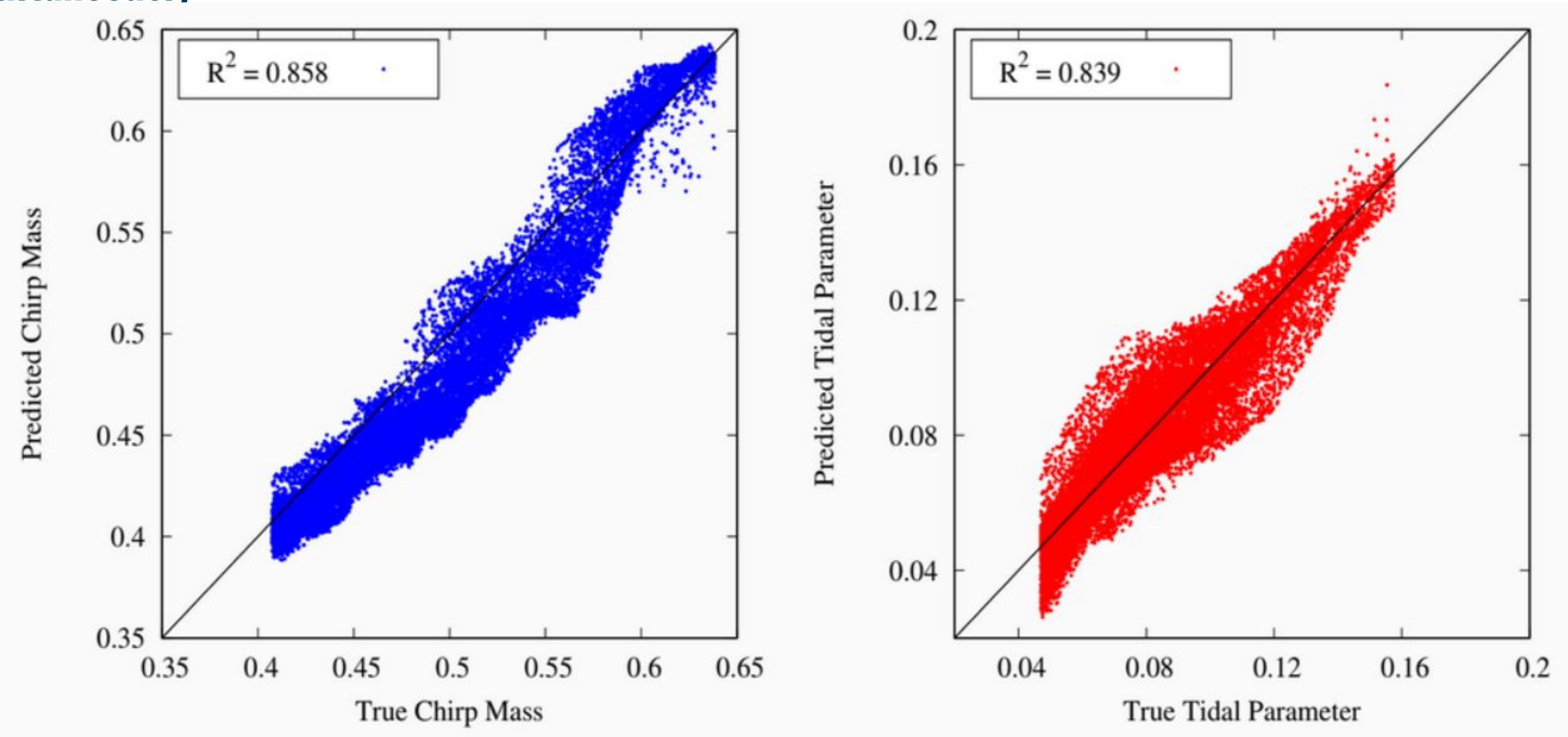
Predicting Combined Λ



Channel 1: Absolute Value, Channel 2: Argument

Training Examples: 64532, Test Examples: 10468

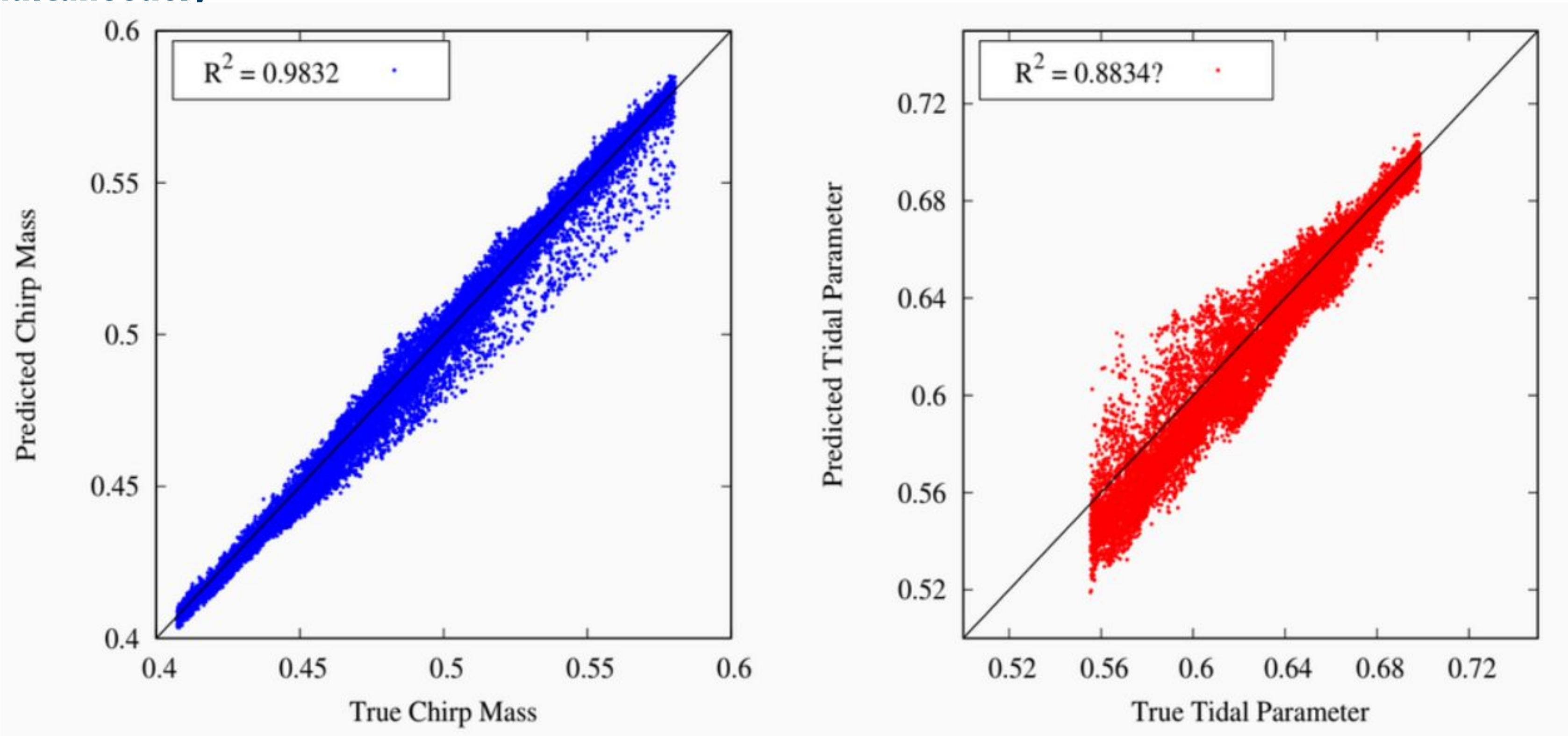
Predicting chirp mass and combined Λ simultaneously



Channel 1: Absolute Value, Channel 2: Argument

Training Examples: 33686, Test Examples: 18718

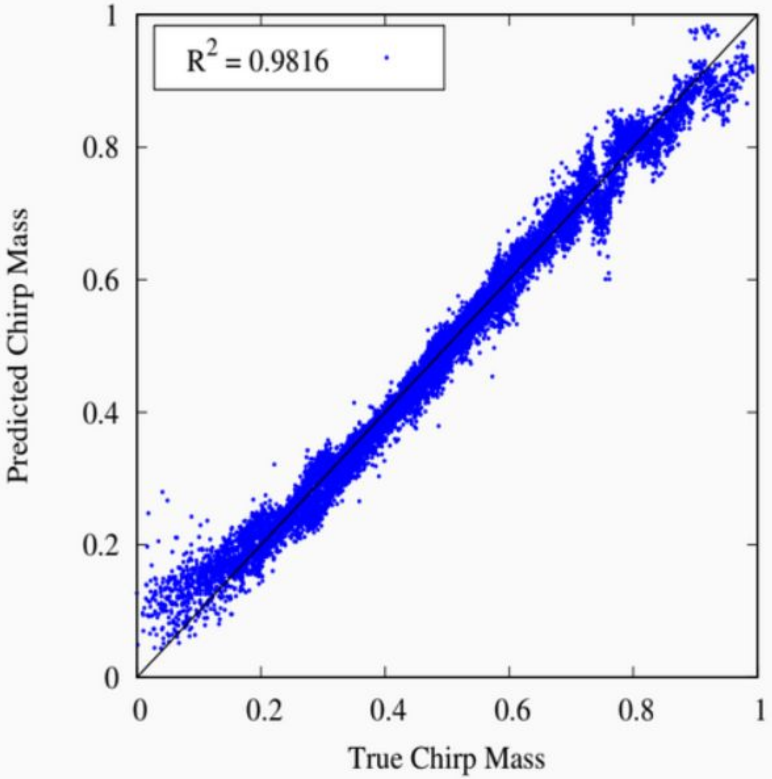
Predicting chirp mass and combined Λ simultaneously

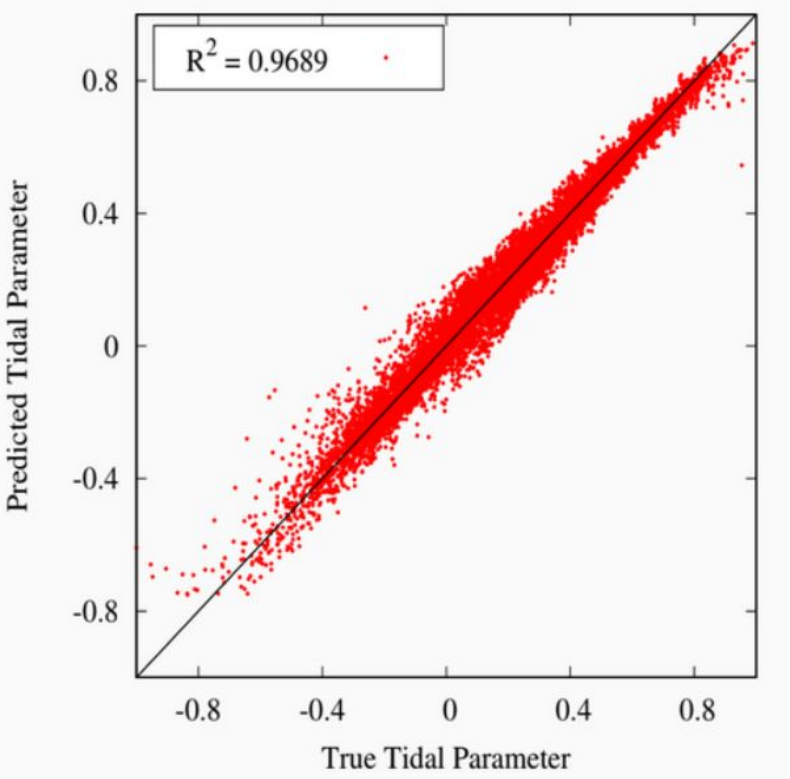


Channel 1: Real part, Channel 2: Imaginary part

Training Examples: 41588, Test Examples: 14808

Predicting chirp mass and combined Λ simultaneously (Signal + Noise)

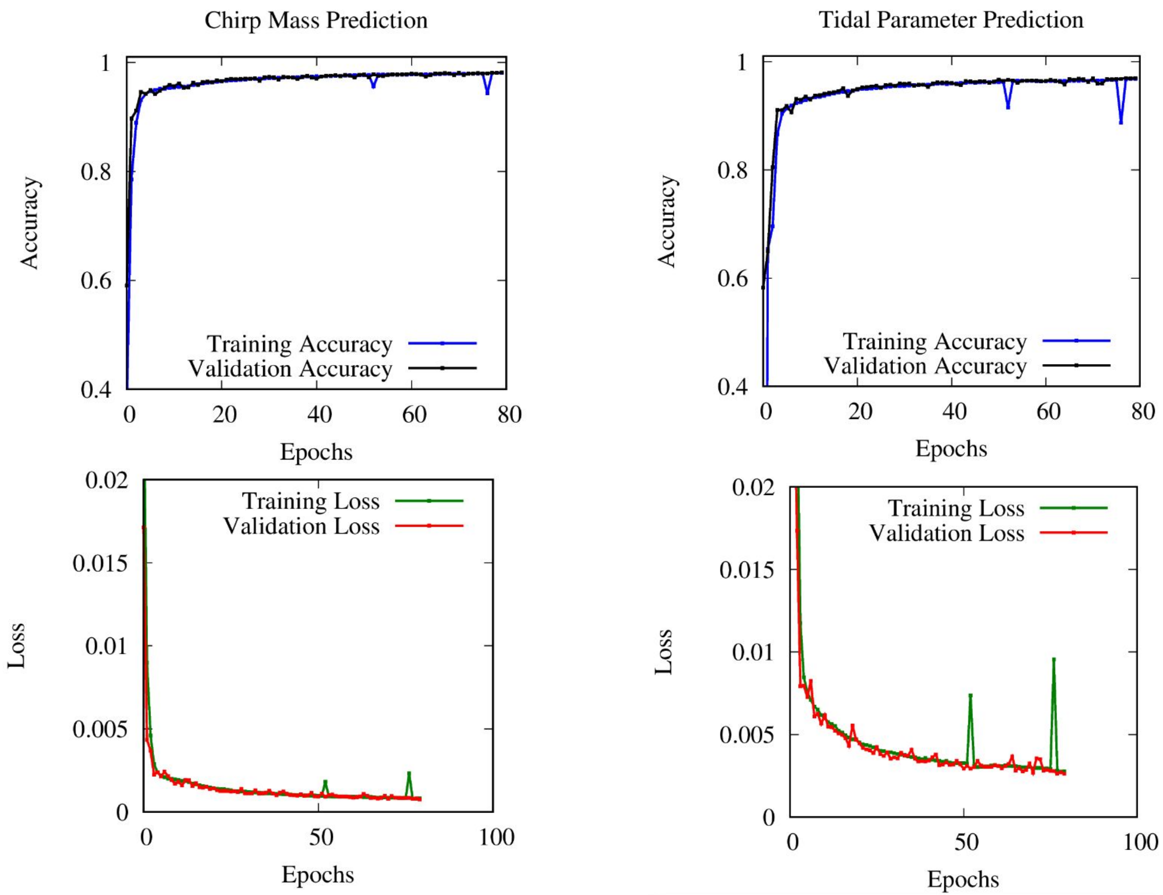




Channel 1: Real part, Channel 2: Imaginary part

Training Examples: 36000, Test Examples: 12000

Learning Curves (Signal + Noise)



Summary

- Classification of GW signals from mergers of BBHs, BNSs and noise.

- Regression of mass and tidal parameters from GW signals of BNSMs.
    - Without noise
    - With white noise
    - With detector noise?

S.S, K. Zhou, J. Steinheimer, H. Stöcker: (in preparation)

Thank you.

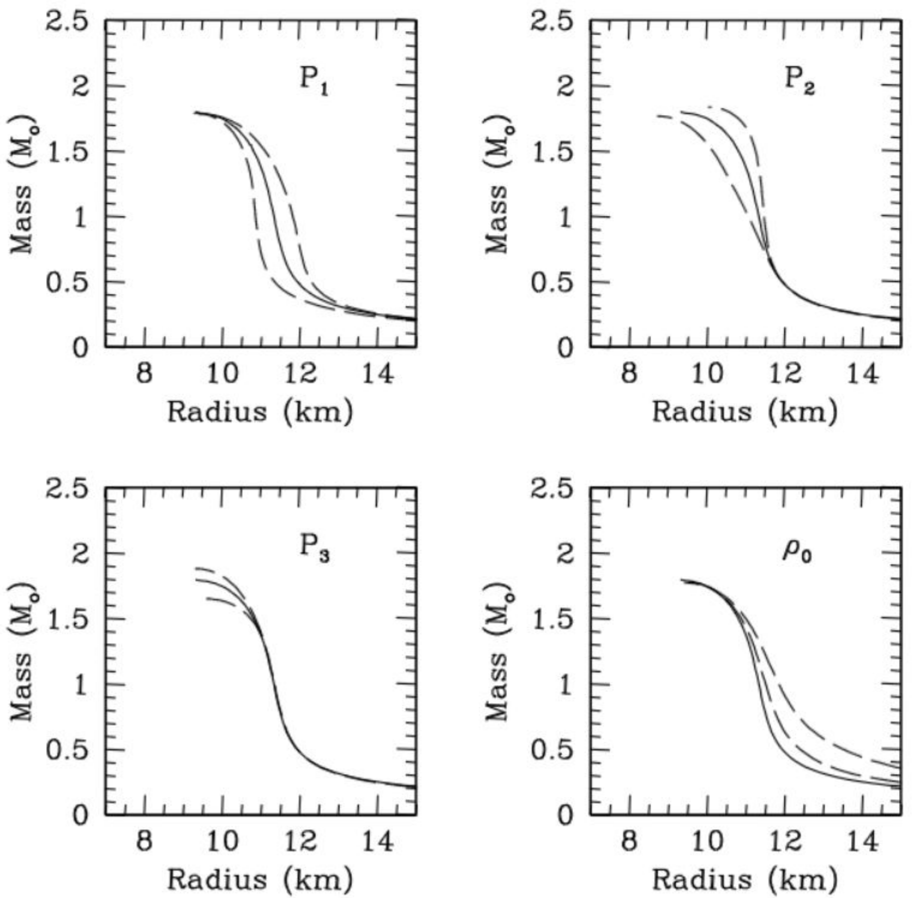
Email: soma@fias.uni-frankfurt.de

Why 7.4ρ_0?

We set the last point, $\rho = 7.4\ \rho_0$, following the results of Read et al. (2009) and Özel & Psaltis (2009) who found that the pressure at this density determines the NS maximum mass and that pressures at higher densities do not significantly affect the overall shape of the resulting MR curve.

[Raithel *et al.*, *ApJ* **831** (2016) 44]

How do P affect the MR curve?

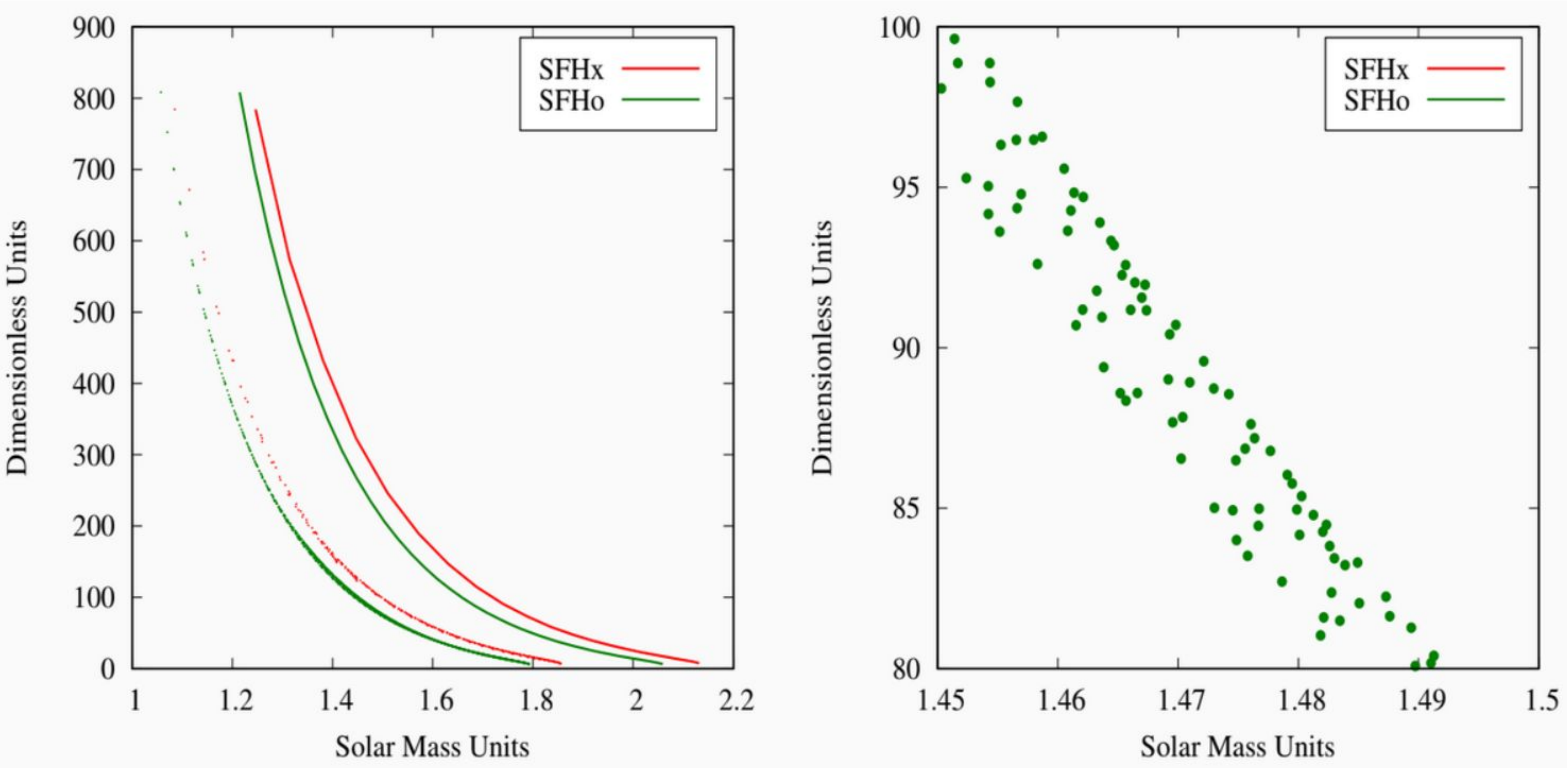


$P_1 = P(\rho = 1.85\rho_0)$
$P_2 = P(\rho = 2\rho_1 = 3.7\rho_0)$
$P_3 = P(\rho = 2\rho_2 = 7.4\rho_0)$

The first three panels show the change in the predicted relation when the values of the parameters P_1, P_2, and P_3 are varied by 25% in each direction away from the best-fit values for the equation of state.

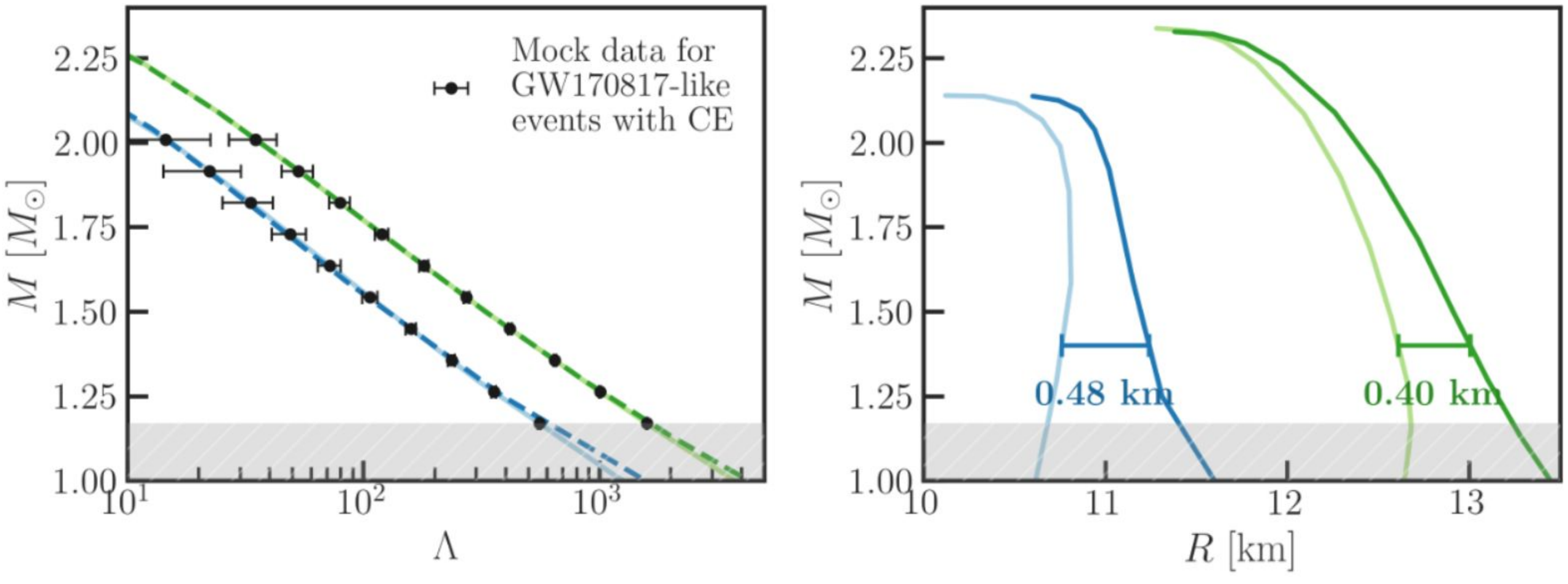
[Özel and Psaltis 009, PhRvD, 80, 103003]

Comparing M-Λ and chirp mass - combined Λ



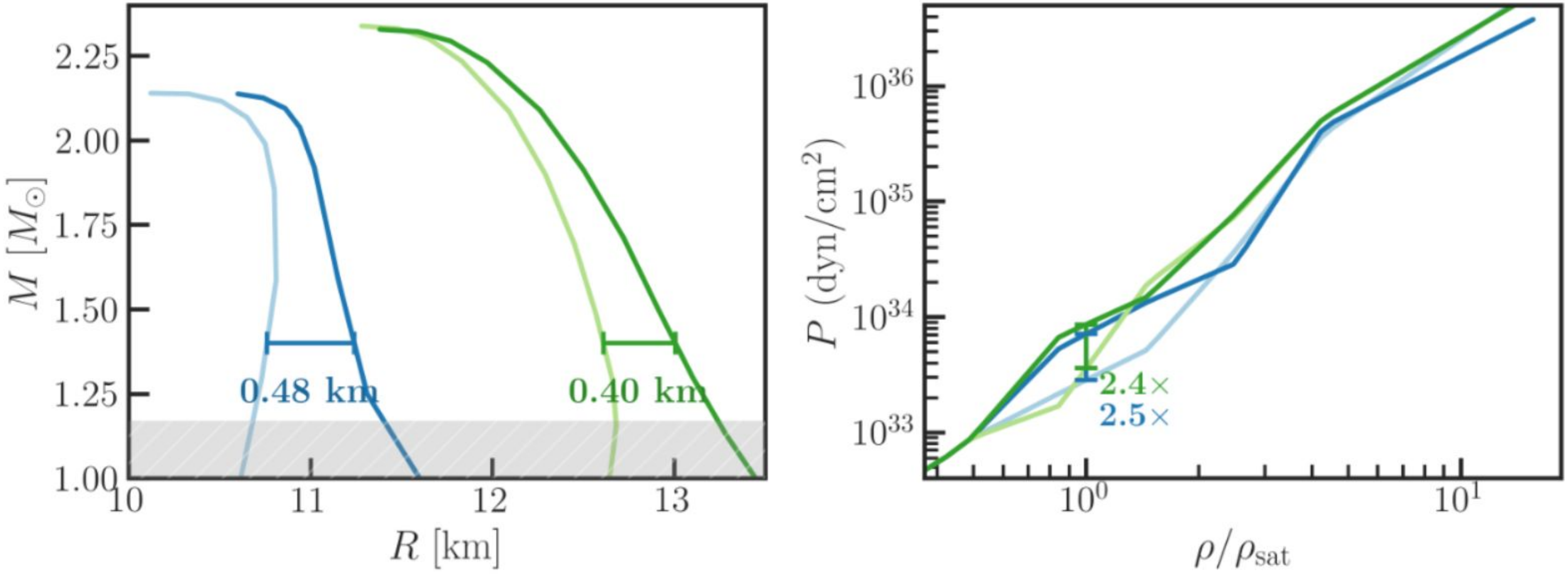
Tidal Deformability Doppelgängers

arXiv:2208.04294

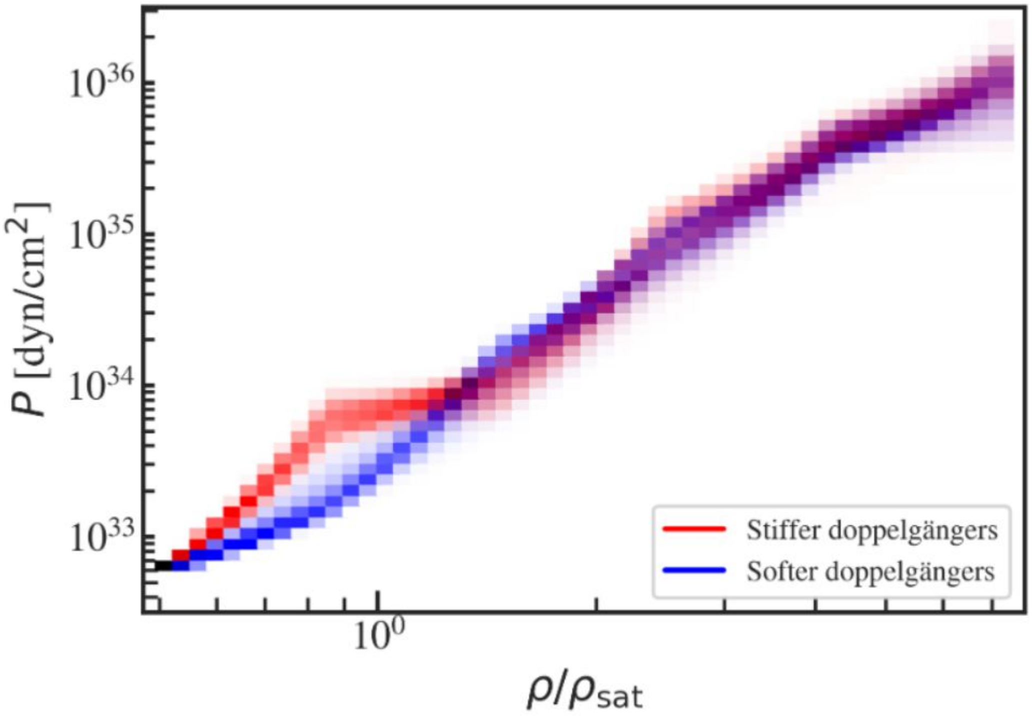


Tidal Deformability Doppelgängers

arXiv:2208.04294



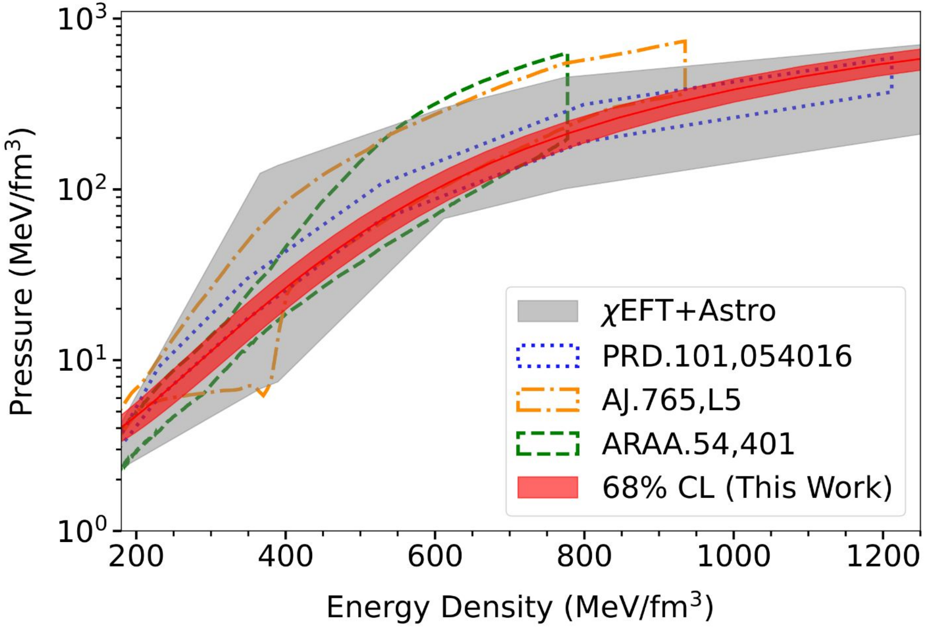
Stiff and soft doppelgängers

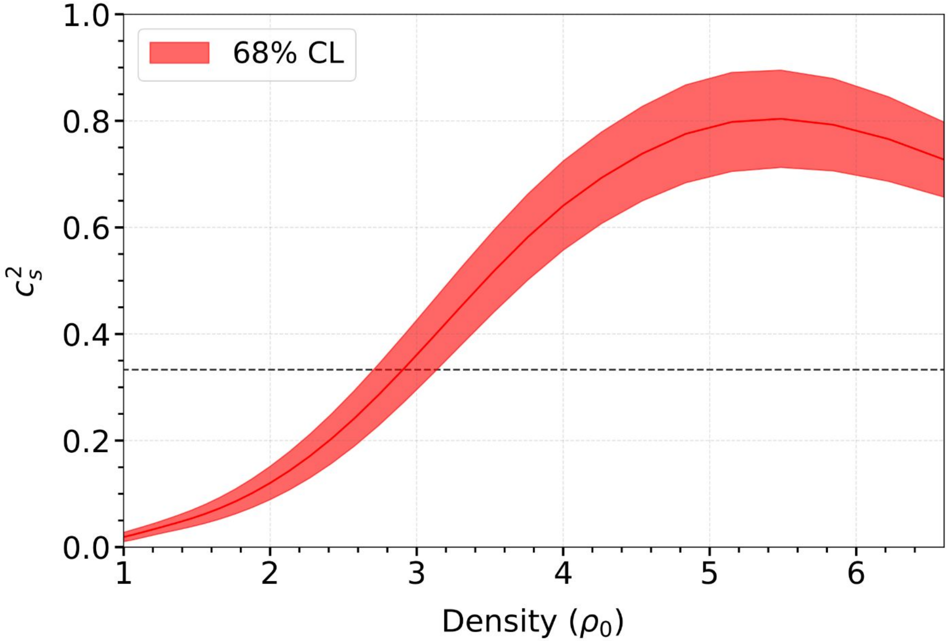


Pressure-density histograms for the set of doppelgängers identified from the randomly-generated sample of PWP EoSs, for which the parametrization starts at 0.5ρ_sat. We classify each EoS in a given pair of doppelgängers as "stiff" or "soft" based on the pressure at the first fiducial density, and we plot the 2D histograms for each subclass in red and blue respectively. The general doppelgänger behavior is caused by allowing for a phase transition at densities near the nuclear saturation density ρ_sat. The onset of the phase transition can be pushed to higher densities by adopting more restrictive nuclear input.

arXiv:2208.04294

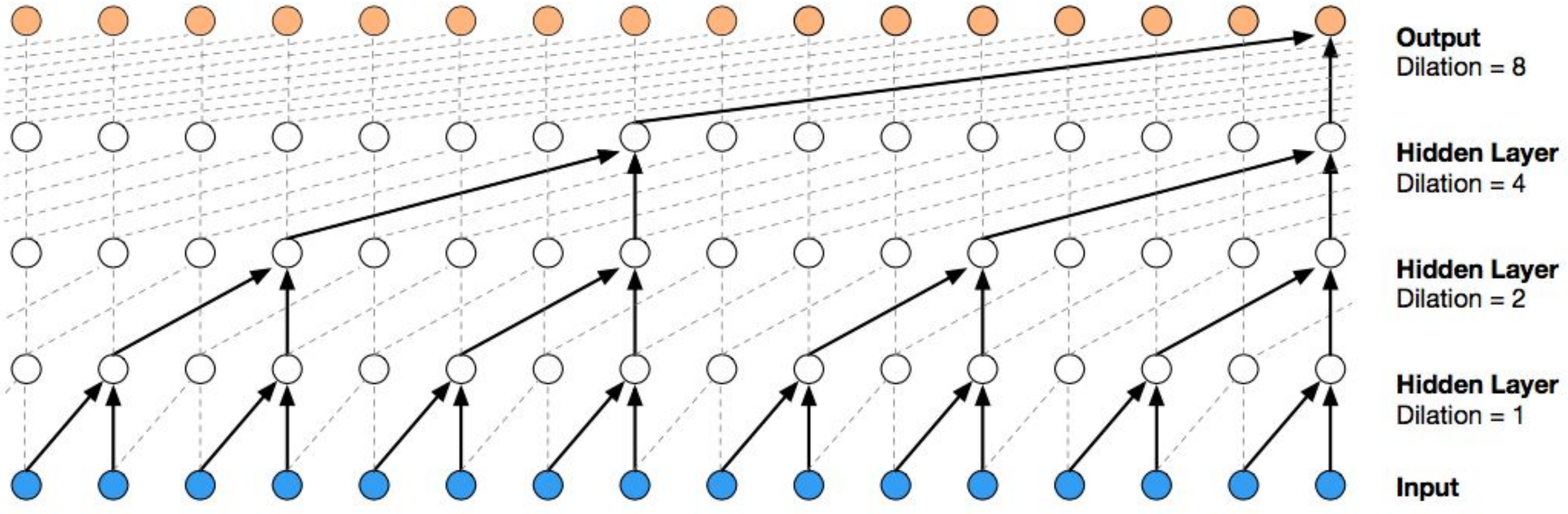
Speed of Sound



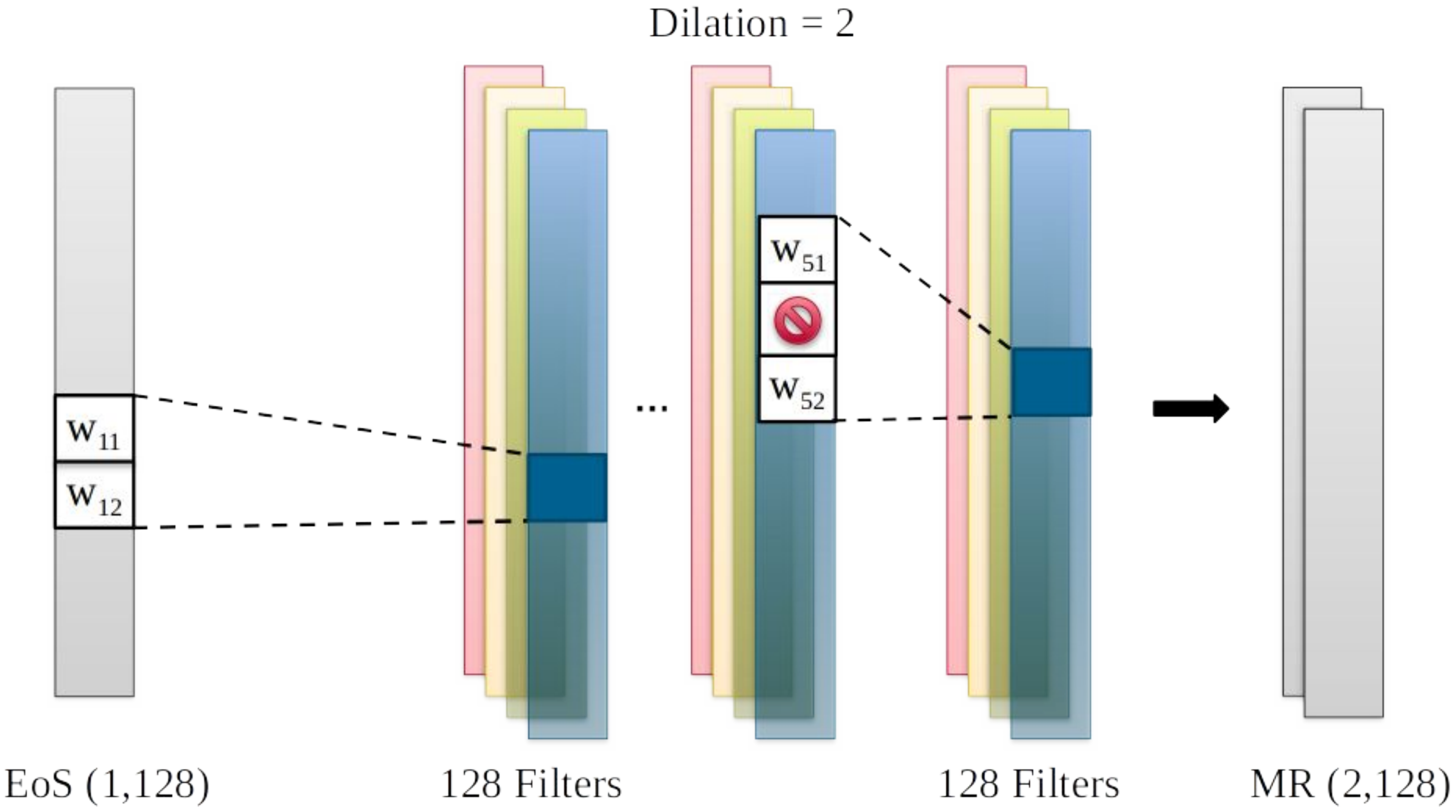


WaveNet - An Autoregressive Network

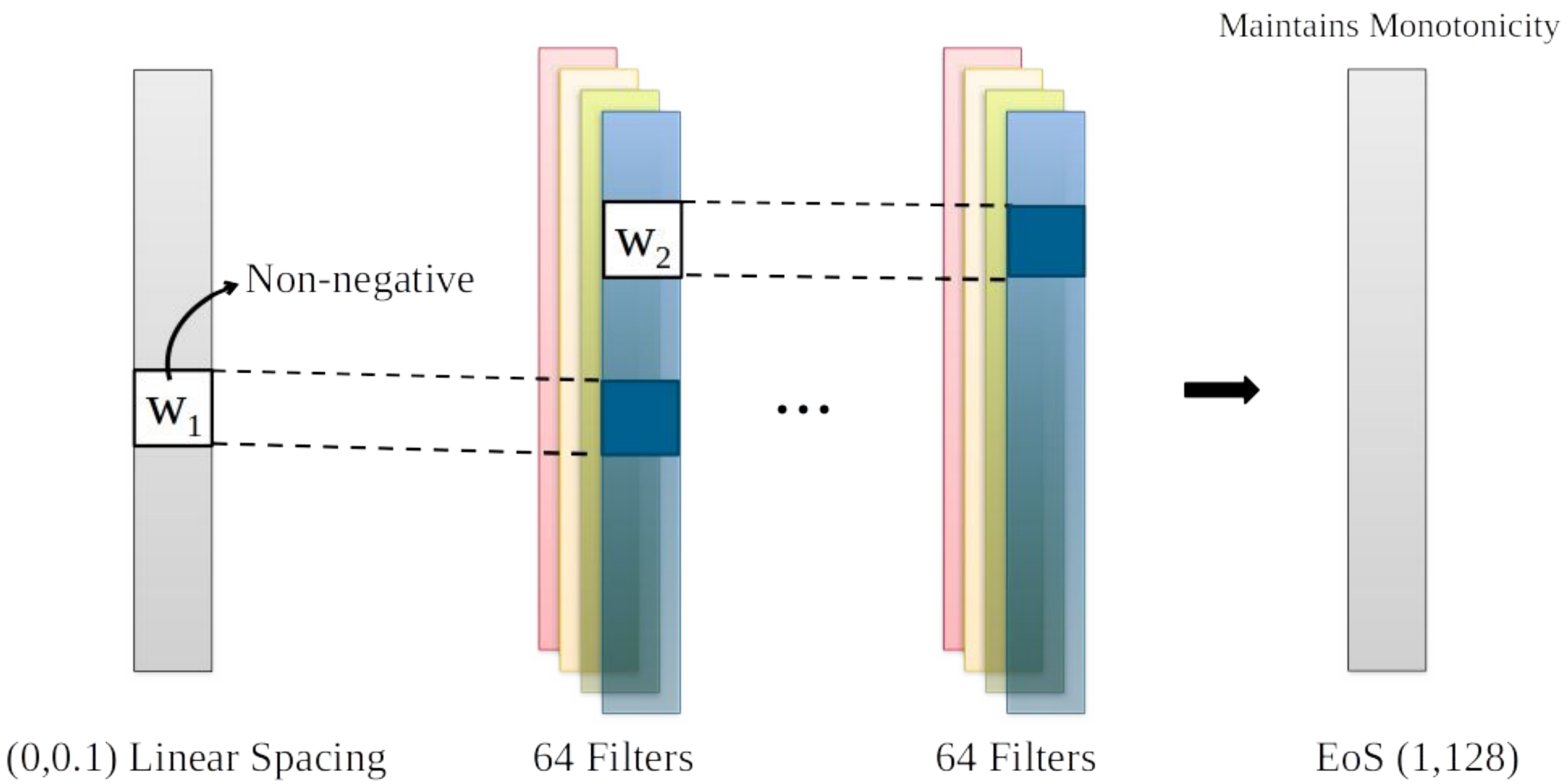
Causal Series



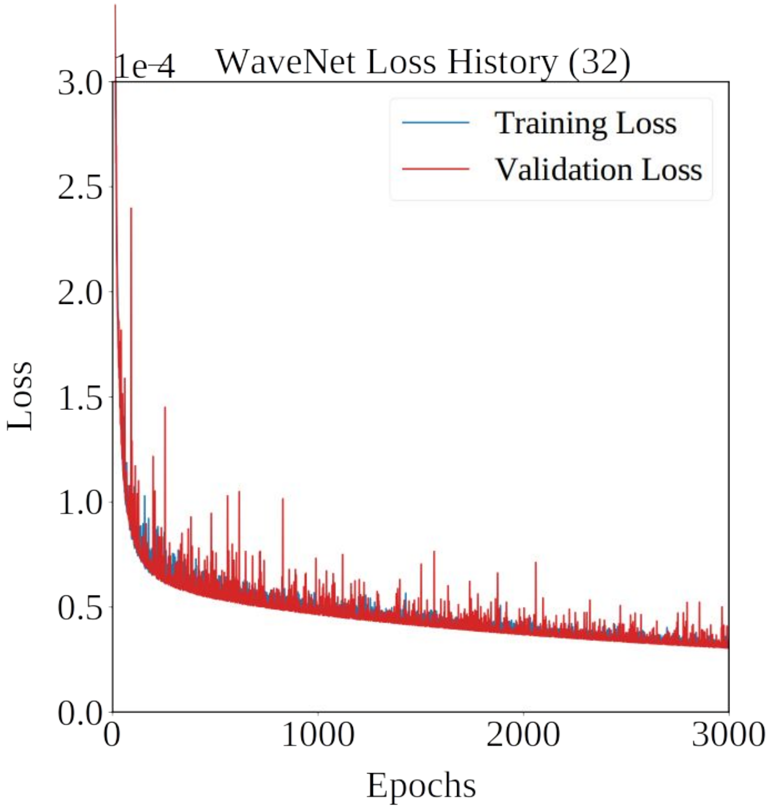
WaveNet - An Autoregressive Network



1D Convolutions - Preserving Order



Learning Curves

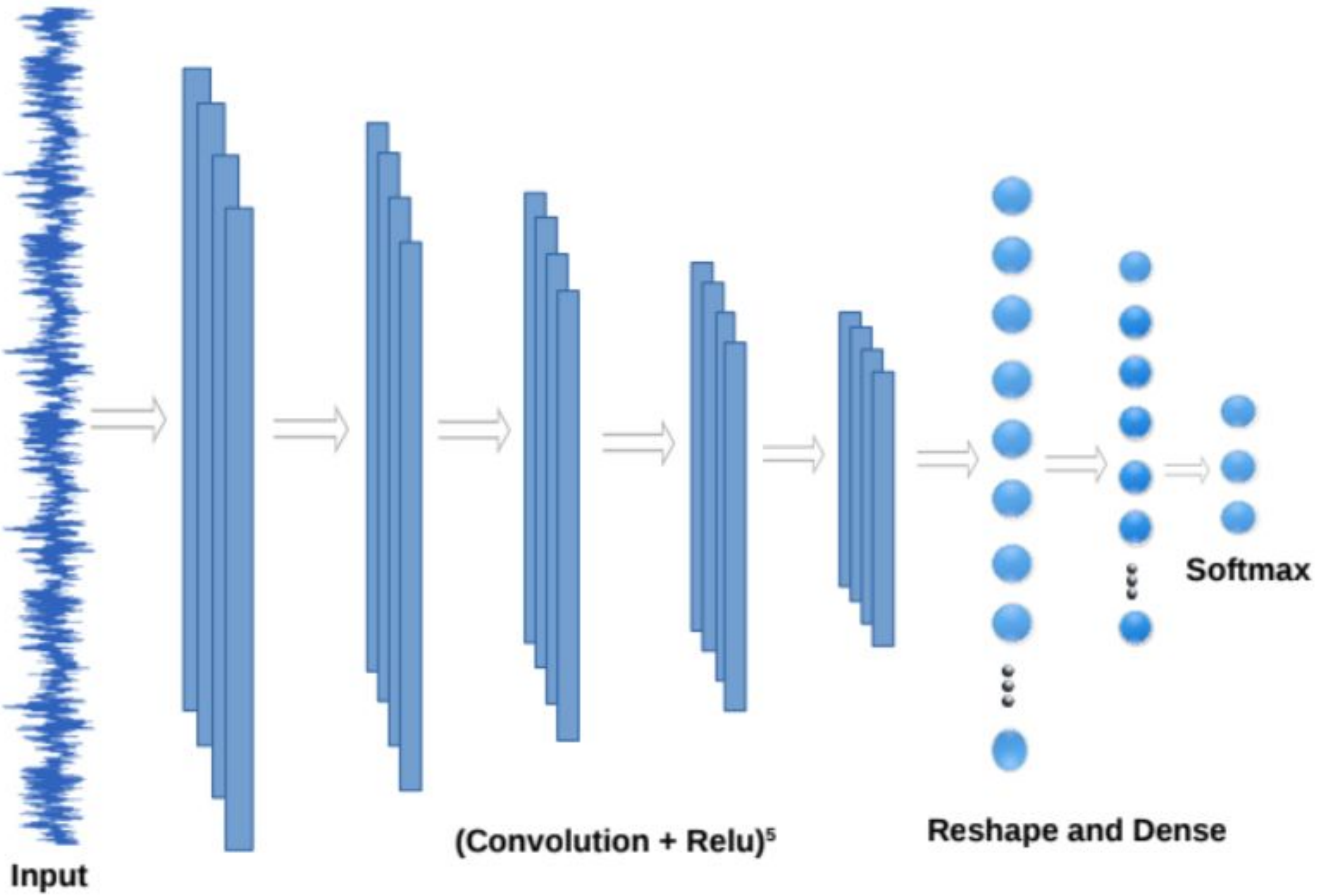


Number of Epochs : 3000

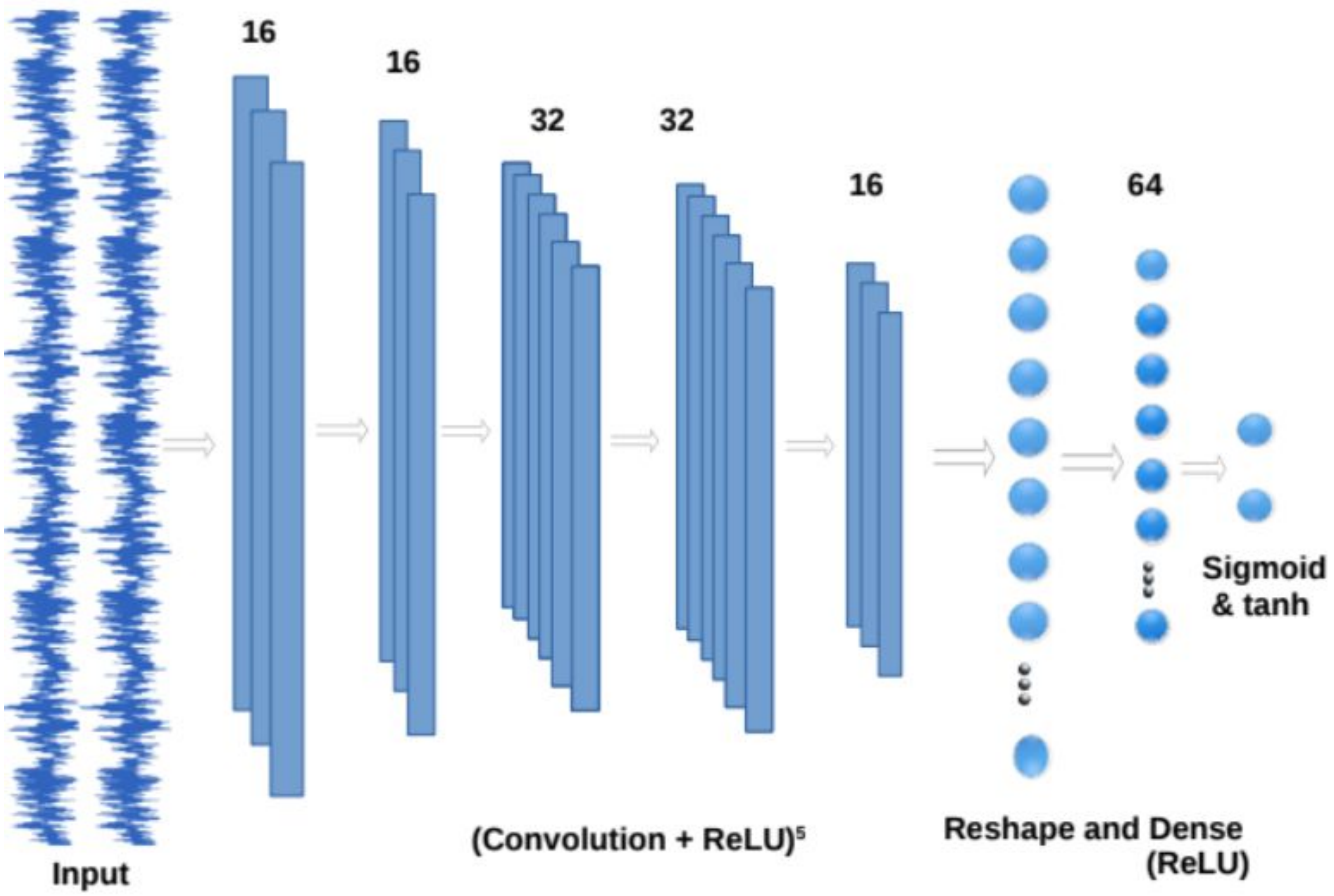
Number of Layers : 10
Dilations : 1, 2, 4, 8, 16, 32, 16, 32, 64

Padding : ‘causal’
Activation function : ‘elu’ (last layer - sigmoid)

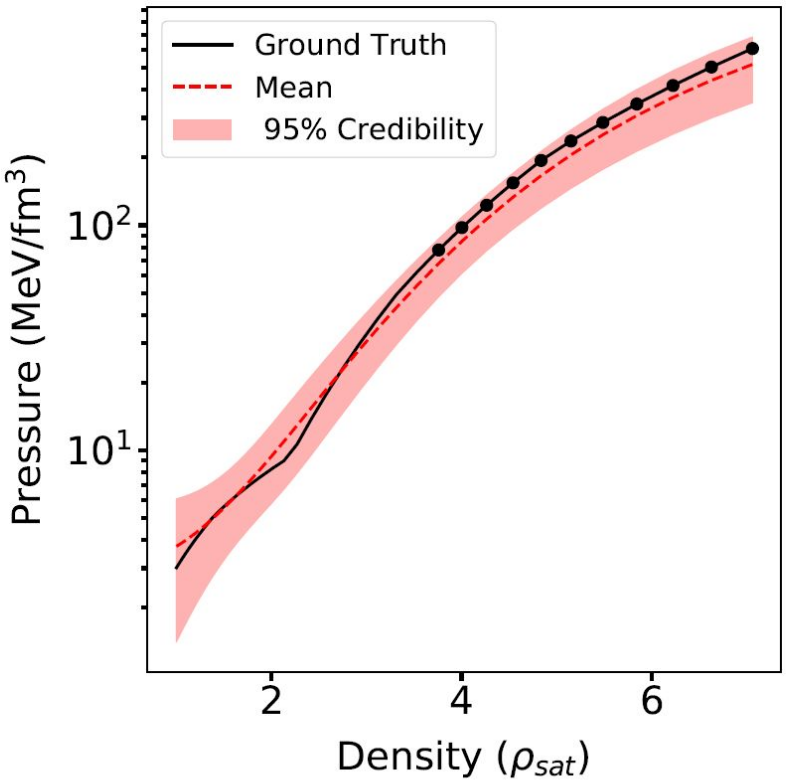
Classification Network



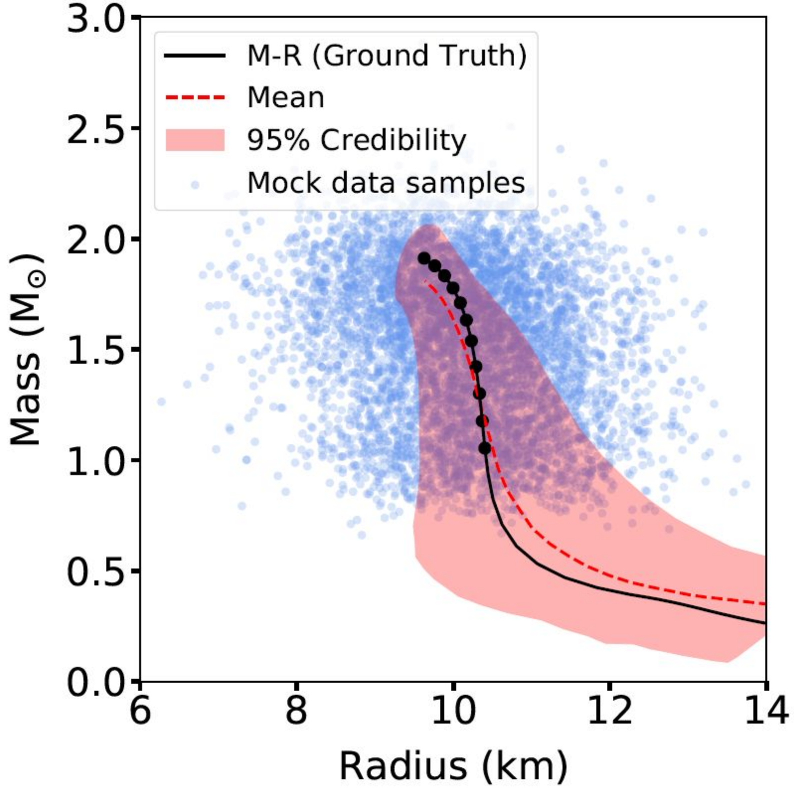
Regression Network



Mock data

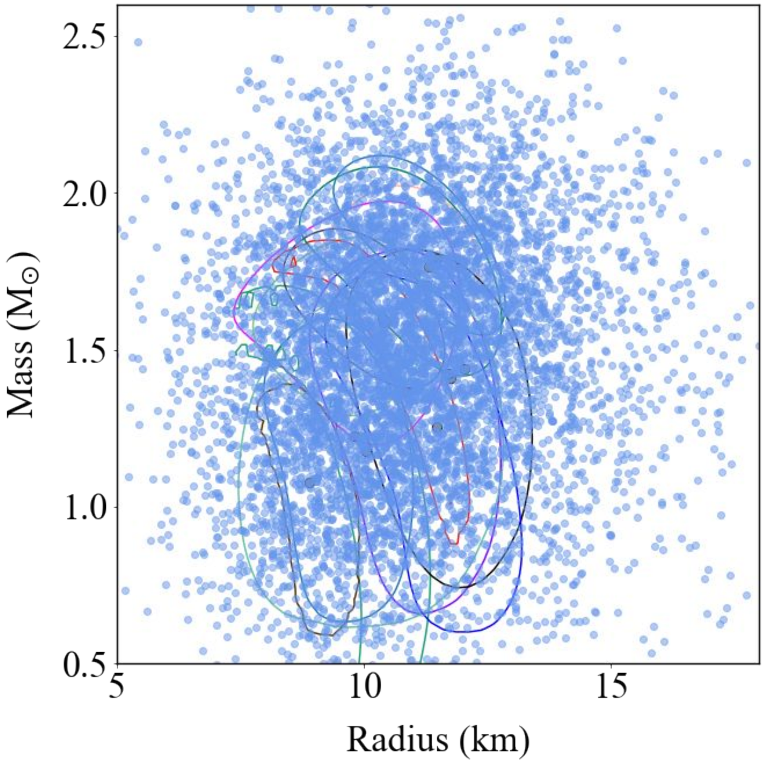


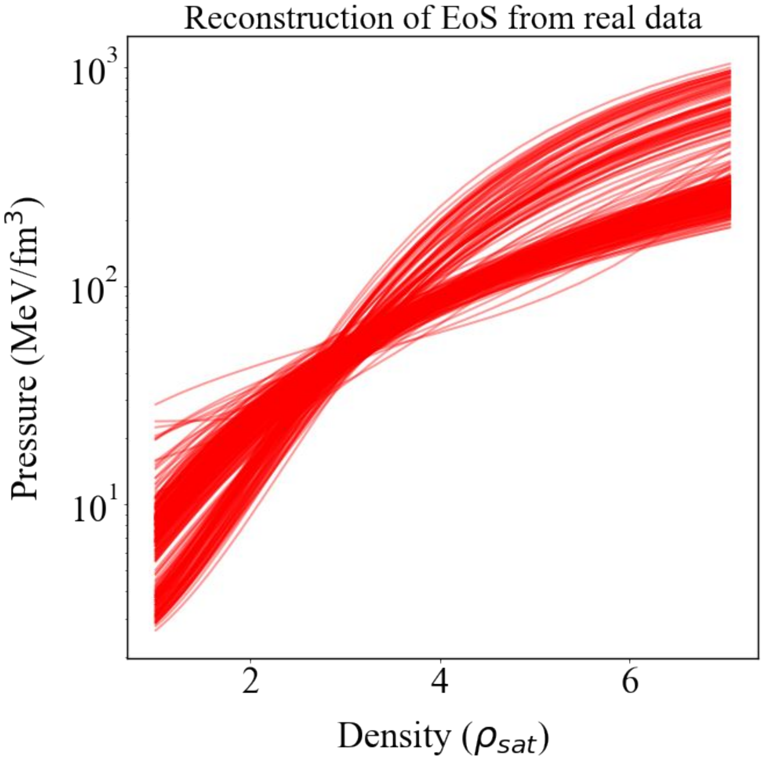
Ground Truth
Mean
95% Credibility
Pressure (MeV/fm^3)
10^2
10^1
2
4
6
Density (ρ_{sat})



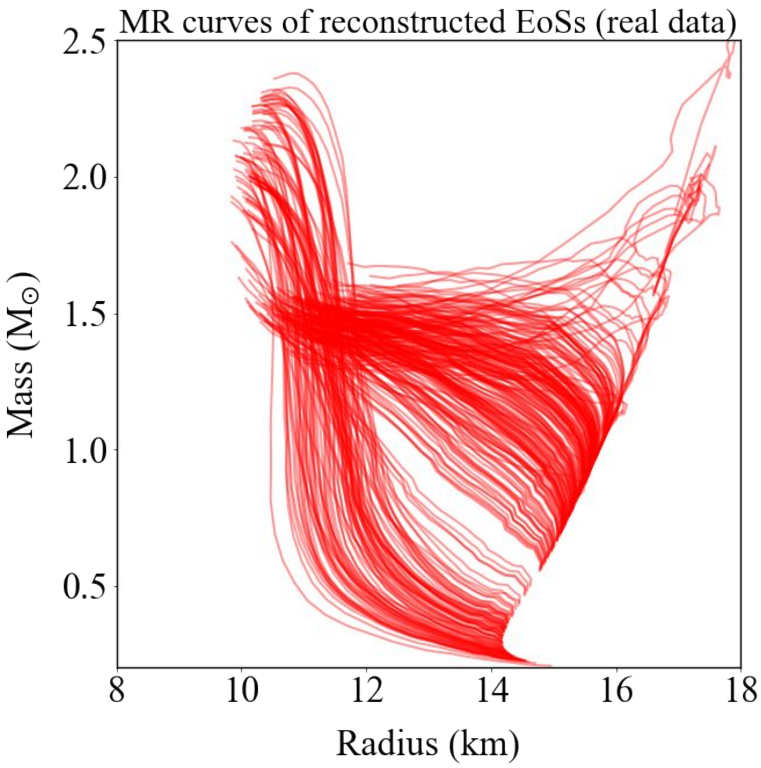
M-R (Ground Truth)
Mean
95% Credibility
Mock data samples
Mass ($M_\odot$)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
6
8
10
12
14
Radius (km)

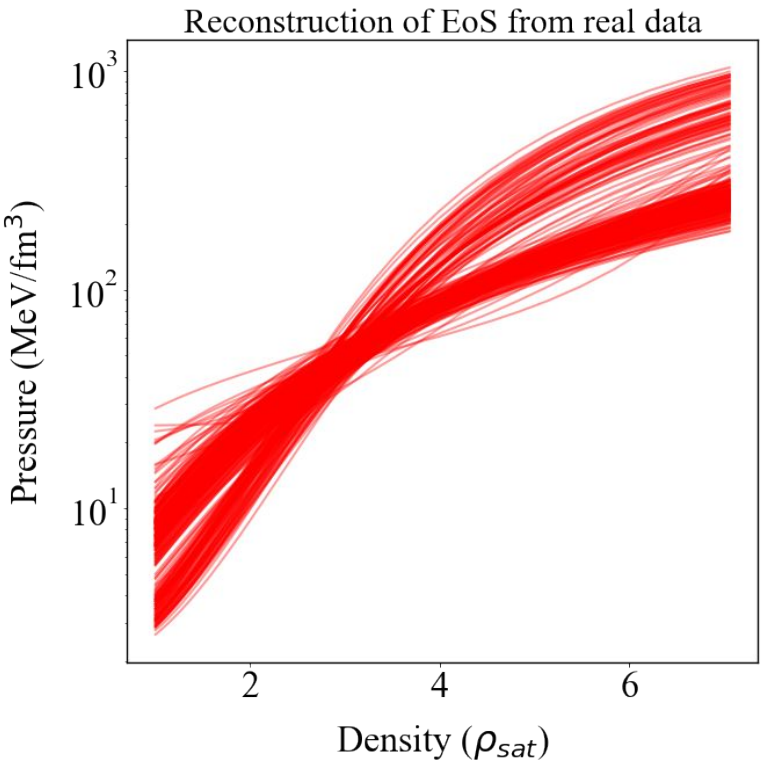
Behind the Scenes



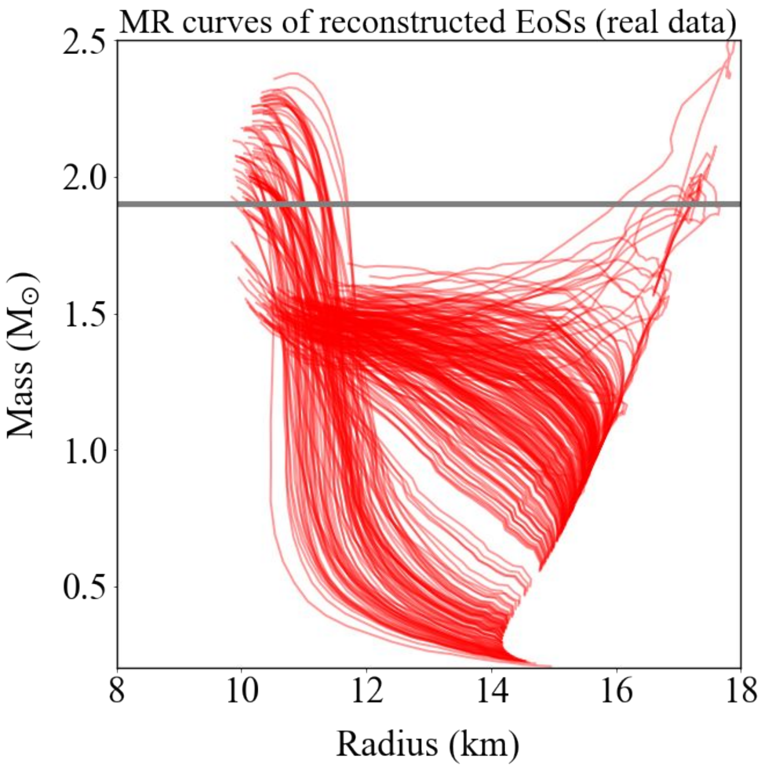


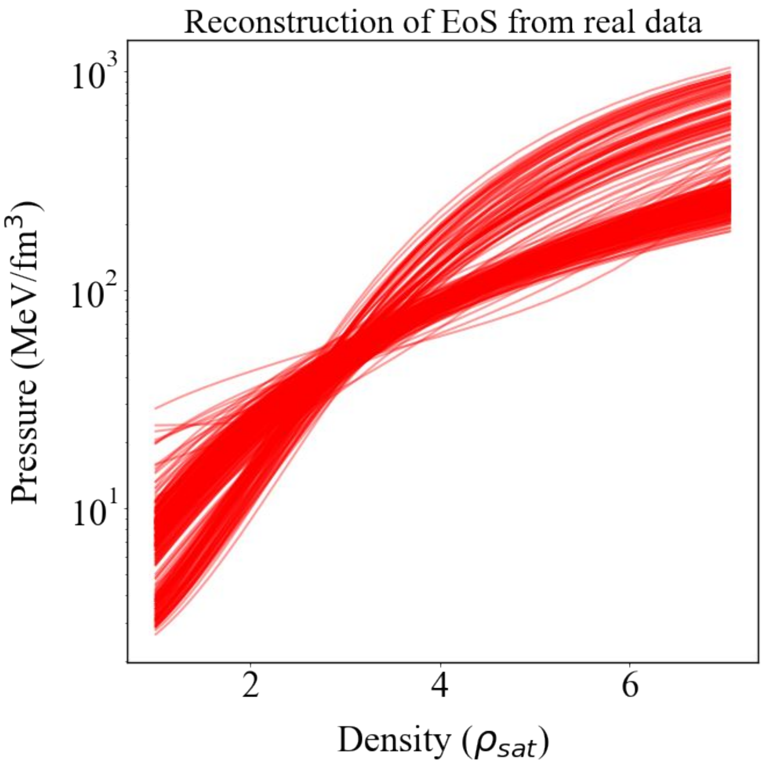
Behind the Scenes



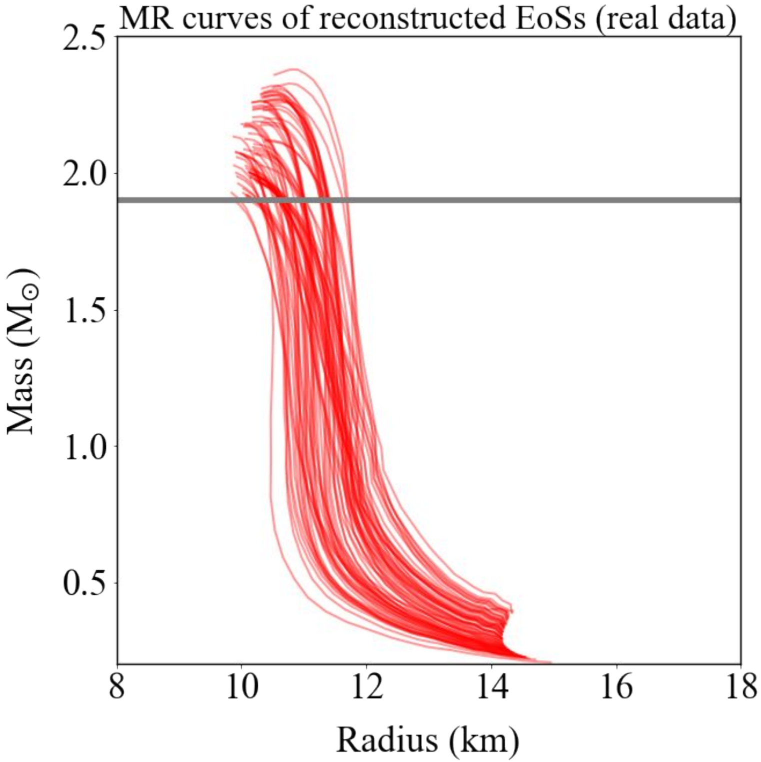


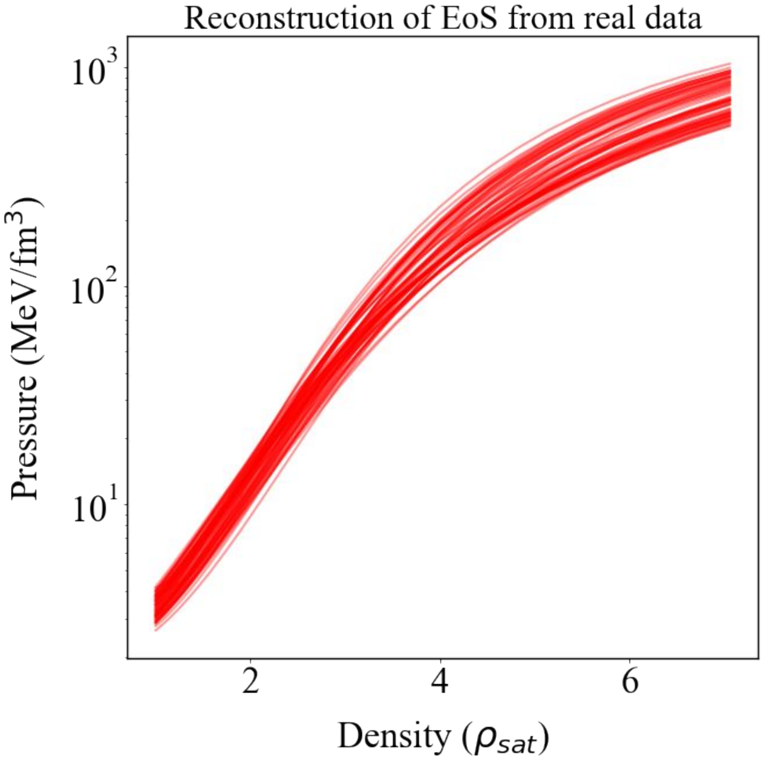
Behind the Scenes



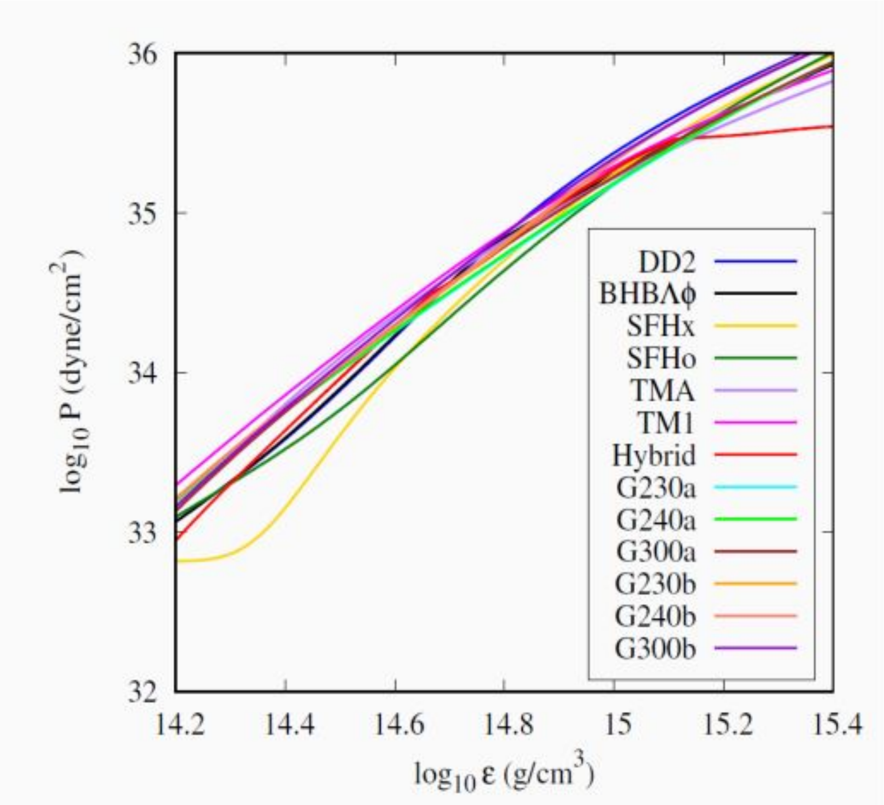


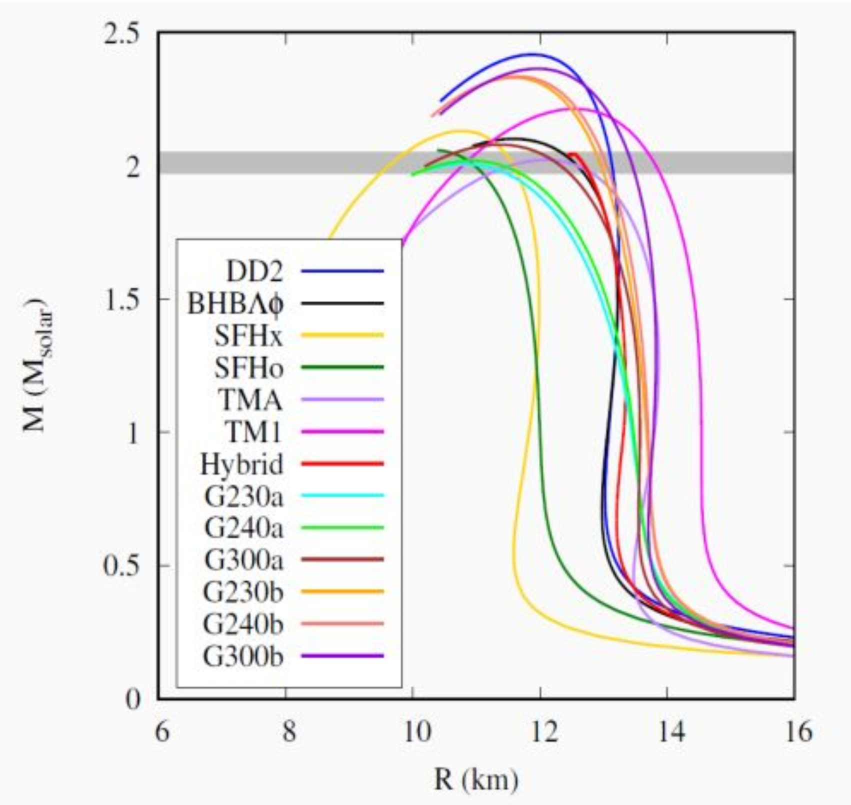
Behind the Scenes



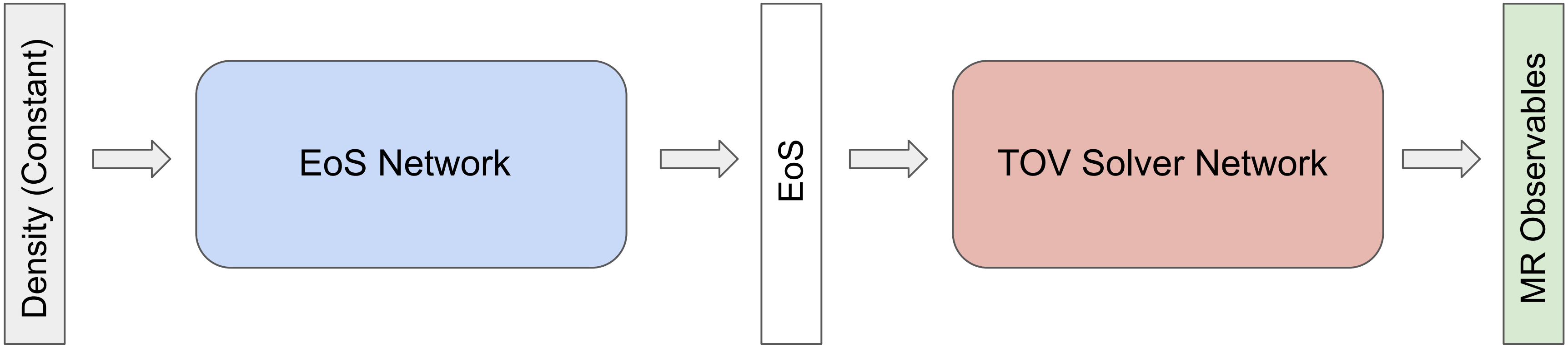


EoSs and corresponding MR curves





Procedure


Density (Constant)
EoS Network
EoS
TOV Solver Network
MR Observables

EoS Parameters

EoS	n_0 (fm^{-3})	m^*/m	BE (MeV)	K (MeV)	S (MeV)	L (MeV)	M_{max} ($M_\odot$)	M_B ($M_\odot$)
DD2	0.1491	0.56	16.02	243.0	31.67	55.04	2.42	2.89
BHB$\Lambda\phi$	0.1491	0.56	16.02	243.0	31.67	55.04	2.1	2.43
SFHo	0.1583	0.76	16.19	245.4	31.57	47.10	2.06	2.43
SFHx	0.1602	0.72	16.16	238.8	28.67	23.18	2.13	2.53
TM1	0.1455	0.63	16.31	281.6	36.95	110.99	2.21	2.30
TMA	0.1472	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	0.1491	0.56	16.02	243.0	31.67	55.04	2.05	2.39
Exp.	0.15-0.16	0.55-0.75	16.00	220-315	29.00-31.70	45.00-61.90	-	-

The Bayesian Posterior

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)}$$

$\text{prior } p(\theta)$

$\text{model likelihood } p(d|\theta)$

Likelihood-based methods

Markov Chain Monte Carlo (MCMC) : Likelihood-based sampler used to draw samples from the posterior.

Nested Sampling, etc.

Likelihood-free inference methods : Deep Neural Networks

If it is possible to sample $d \sim p(d|\theta)$ (i.e., simulate data) one can alternatively use simulation-based (or likelihood-free) inference methods.

For Gravitational Wave (GW) inference, deep neural networks (DNNs) have also been shown to achieve similar accuracy to MCMC.

Deep Neural Networks : Pros?

- When rapid results are desired—for alerts to trigger electromagnetic follow-up of transient phenomena—computational efficiency makes all the difference, by using either fast models or specialized inference algorithms.

Deep Neural Networks : Cons?

- Can it generalize to out-of-distribution (i.e., data inconsistent with the training distribution) data?
- Insufficient training?
- Lack diagnostics to be confident in results.

These powerful approaches are therefore rarely used in applications where accuracy is important.

Importance Sampling

- Start from a collection of *n* samples $\theta_i \sim q(\theta|d)$ (the "proposal")
- Assign to each sample an importance weight $w_i = p(d|\theta_i)p(\theta_i)/q(\theta_i|d)$
- For a perfect proposal, w_i = constant
- Number of effective samples is related to the variance, $n_{eff} = (\sum_i w_i)^2 / \sum_i (w_i^2)$
- The sample efficiency $\epsilon = n_{eff}/n \in (0, 1]$ arises naturally as a quality measure of the proposal

Importance Sampling

- Importance sampling requires evaluation of $p(d|\Theta)p(\Theta)$ rather than the normalized posterior

- The evidence can then be estimated from the normalization of the weights as $p(d) = 1/n \sum_i w_i$

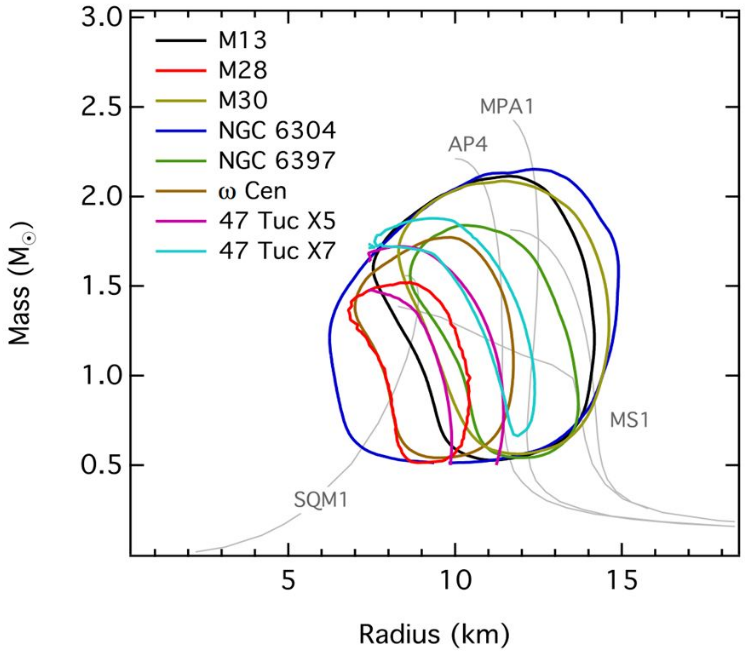
Importance Sampling for EoS reconstruction

Özel *et al.*, ApJ **820** (2016) 28
Bogdanov *et al.*, ApJ **831** (2016) 184
Riley *et al.*, ApJL **887** (2019) L21
Riley *et al.*, ApJL **918** (2021) L27

- TOV Solver trained on 4 low-density EoSs ⟶ Bias at low-densities while reconstructing the EoS

- Several M-R curves sampled from the uncertainty distribution of data. Each reconstructed curve is fitted to these M-R samples rather than the mean M-R curve.

- Importance sampling to correct bias



Importance Sampling for EoS reconstruction

Özel *et al.*, ApJ **820** (2016) 28
Bogdanov *et al.*, ApJ **831** (2016) 184
Riley *et al.*, ApJL **887** (2019) L21
Riley *et al.*, ApJL **918** (2021) L27

Calculating the sample efficiency :

- $n_{eff} = (\sum_i w_i)^2 / \sum_i (w_i^2) \sim 90.2$
- Samples ~ 1759
- $\epsilon = n_{eff}/n \sim 5.13\ \%$

Apply a cut-off and reweight the samples accordingly