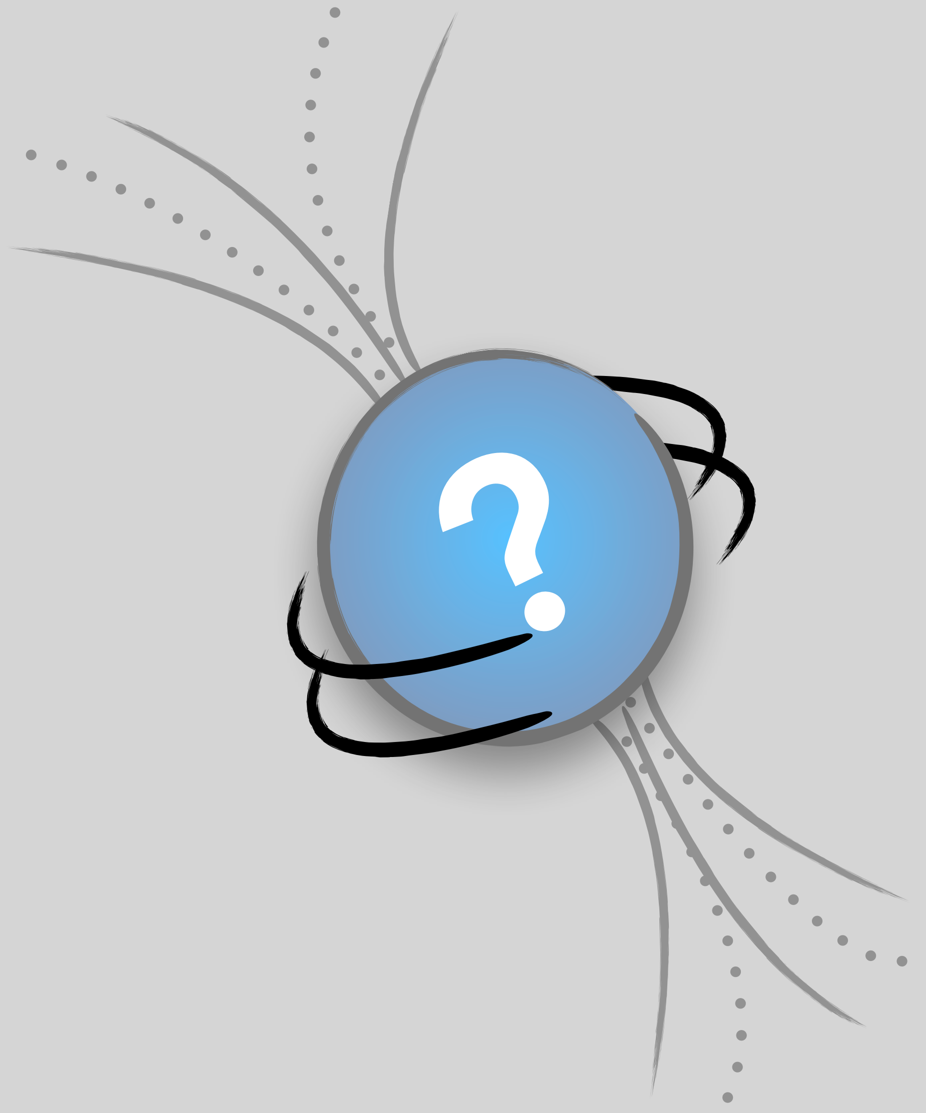


How can deep learning help us decipher neutron star composition

Valéria Carvalho
Constança Providencia, Michał Bejger, Márcio Ferreira

6 November 2025

Outline

How can deep learning help us decipher neutron star composition ?

- Motivation: The Unsolved Puzzle of Neutron Stars

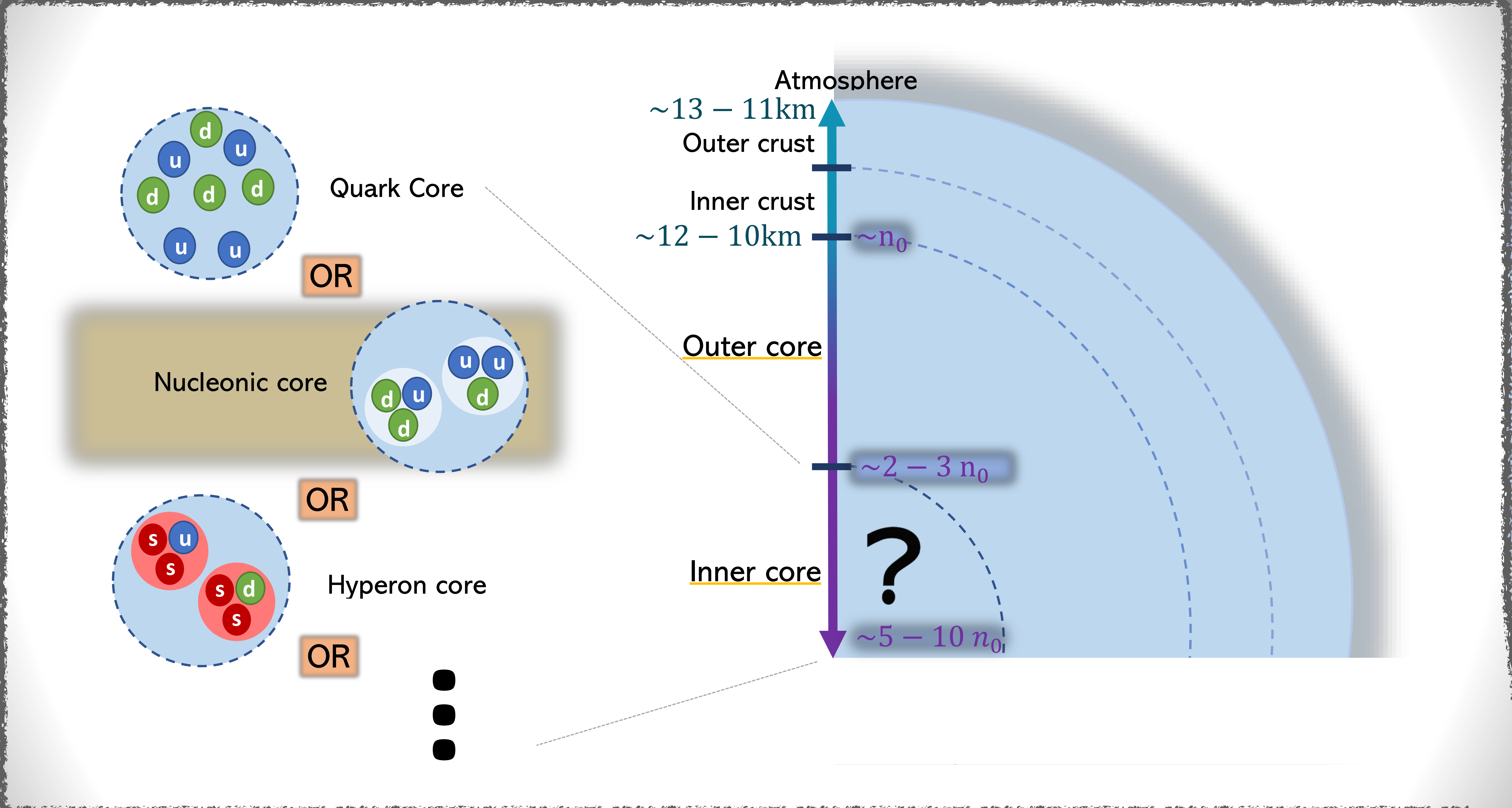
Examples:

- a) Inference of proton fraction and speed of sound with uncertainty estimation V.C, M. Ferreira, T. Malik, C. Providência (PRD **108**.043031)

- b) Neural posterior estimation of neutron star equations of state

V.C, M. Ferreira, M. Bejger, C. Providência (PRD **112.**083044)

What's inside a neutron star?



$M = 1 \sim 2M_\odot$

$R \approx 10^{-5} R_\odot$

$T=0$

The equation of state: A bridge between micro and macro

Equation of State(EoS)

$$p(\varepsilon)$$

Speed of Sound

$$c_s^2(n) = \frac{dp(n)}{d\varepsilon(n)}$$

Trace Anomaly

$$\Delta(n) = \frac{1}{3} - \frac{p(n)}{\varepsilon(n)}$$

Tolman-Oppenheimer-Volkoff equations

$$\frac{dP(r)}{dr} = -\frac{\varepsilon(r)m(r)}{r^2}\left(1 + \frac{P(r)}{\varepsilon(r)}\right)\left(1 + \frac{4\pi r^3 P(r)}{m(r)}\right)\left(1 - \frac{2m(r)}{r}\right)^{-1}$$

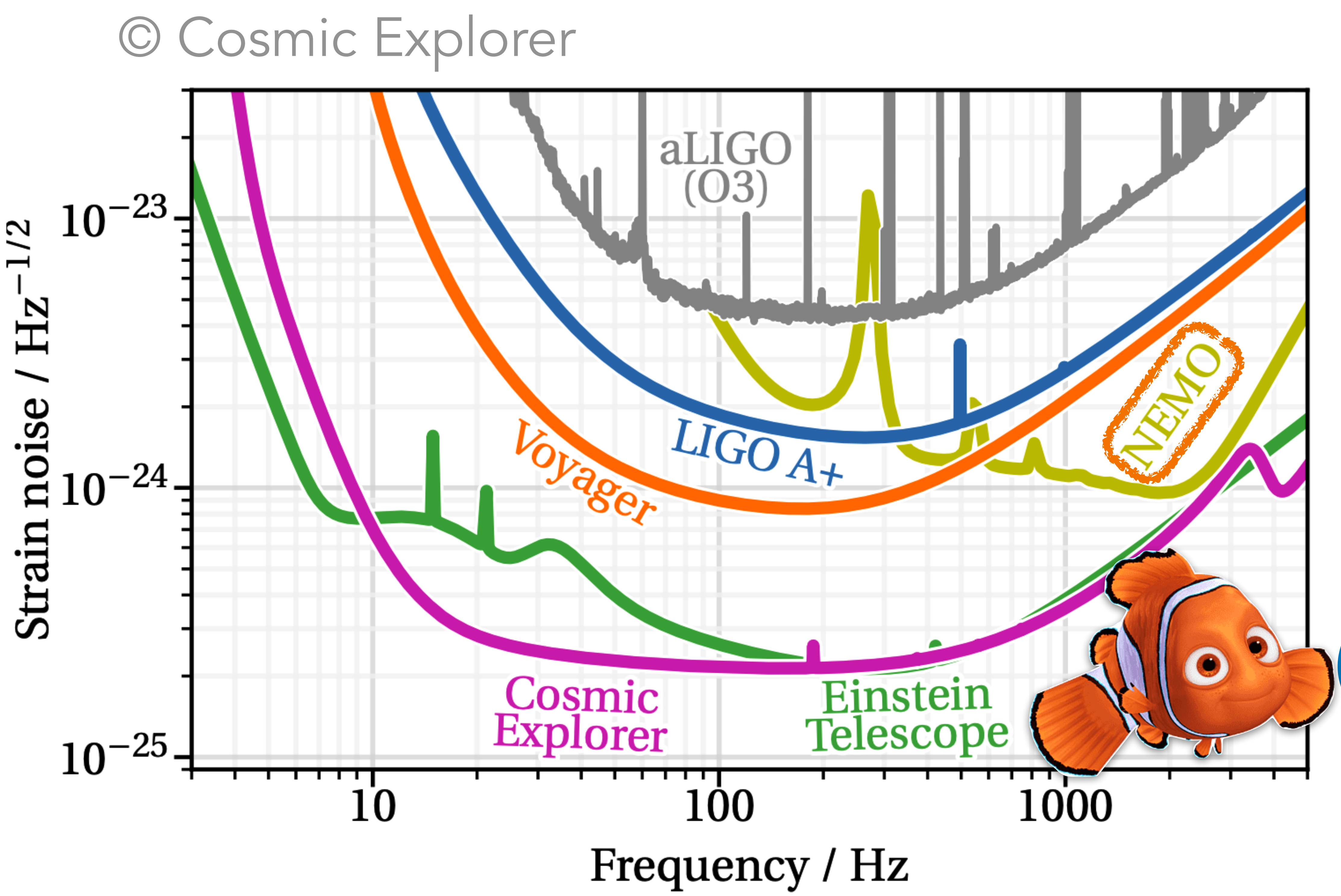
$$\frac{dm(r)}{dr} = 4\pi r^2 \varepsilon(r)$$

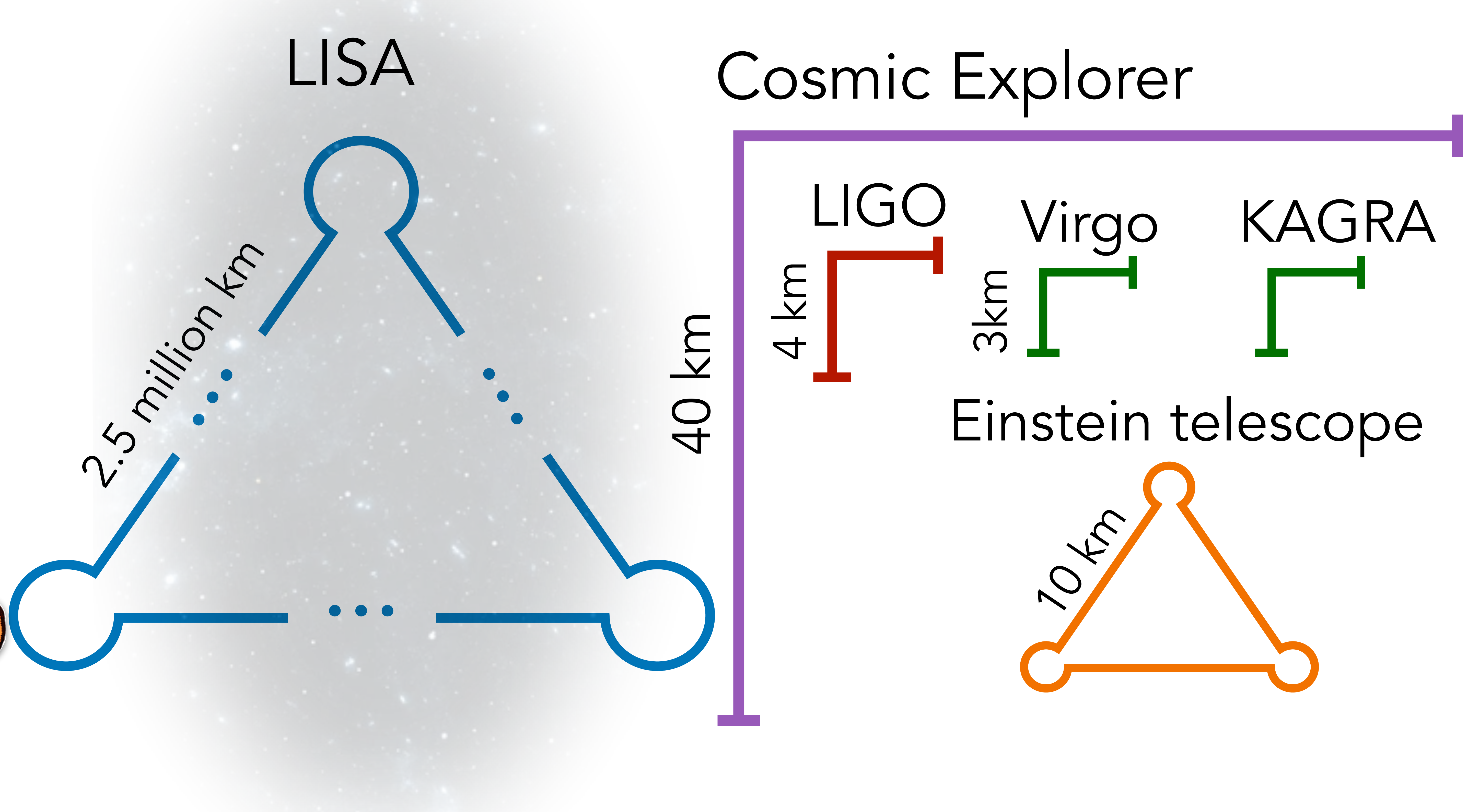
Tidal deformability

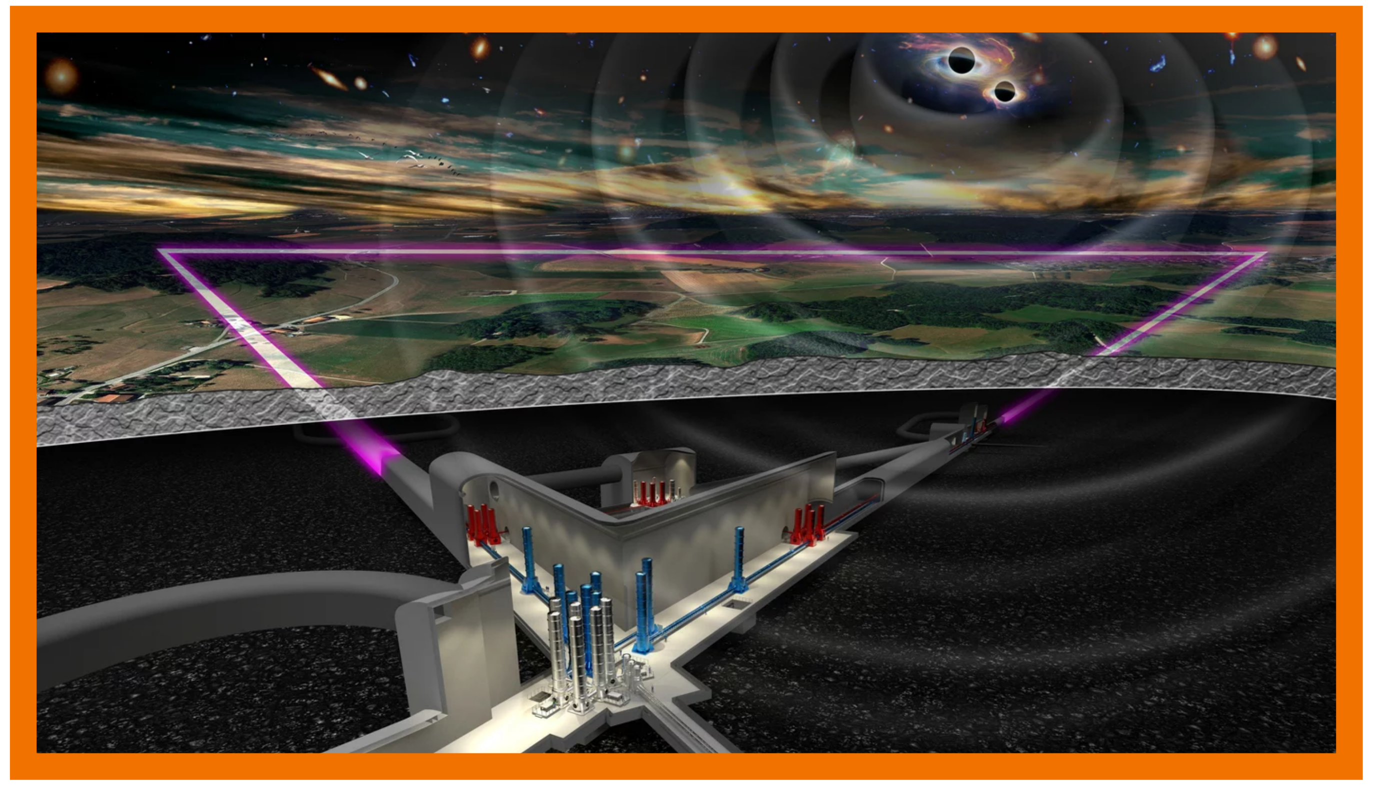
$$\Lambda = \frac{2}{3}k_2 C^{-5}, \qquad C = \frac{M}{R}$$

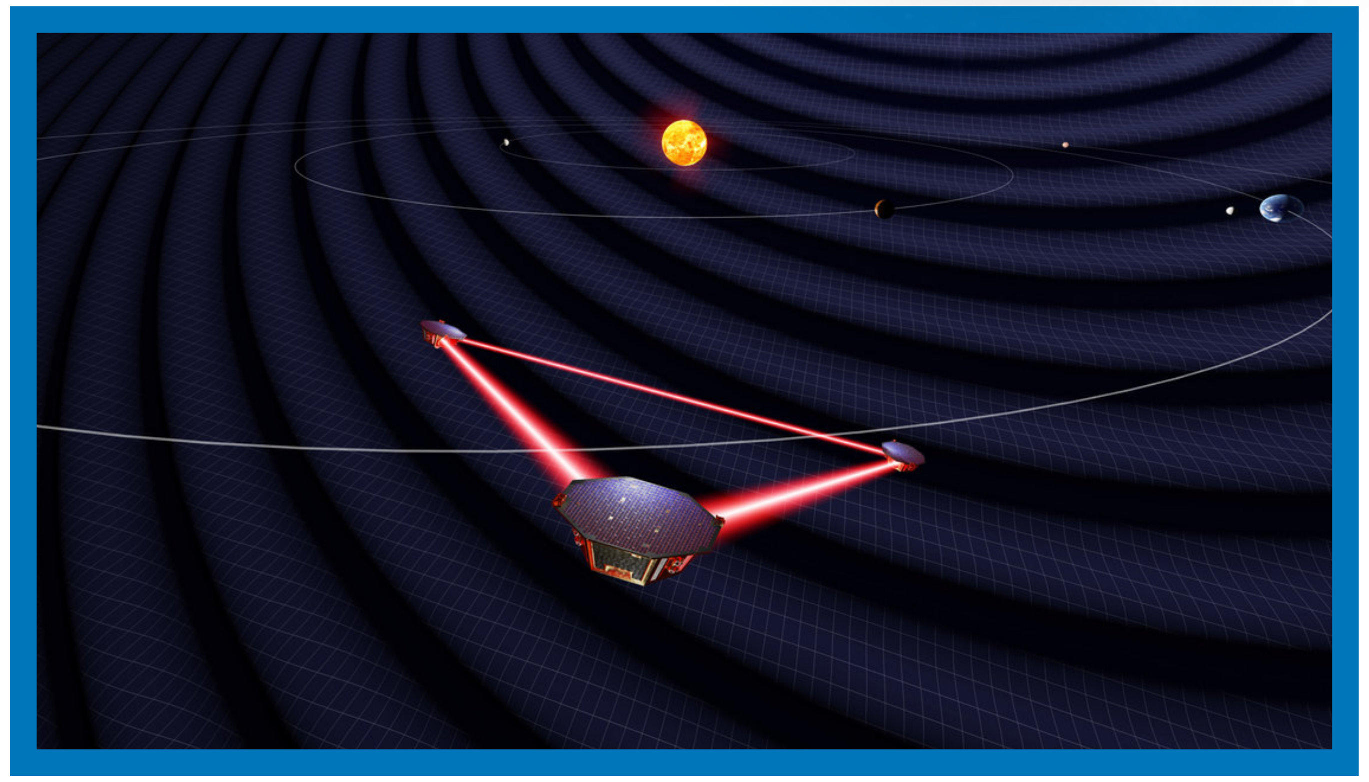
Current observations

A golden age of data is coming- GW future detectors









The challenge: A sparse and noisy inverse problem

For current observations:

- Sparse coverage of observables,
- Uncertainties and degeneracies.

For future observations:

- Big amount of data,
- Smaller uncertainties, less degeneracies.

Deep Learning pipeline

Deep Learning ⊂ Machine Learning ⊂ AI

Benefits :

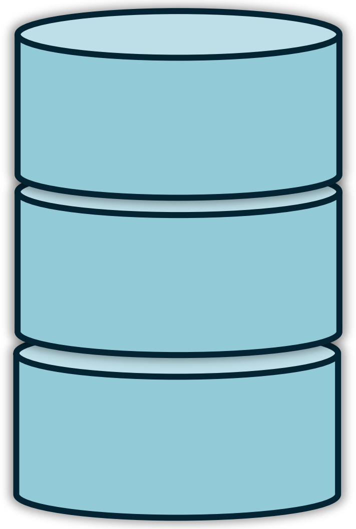
- Handles complexity,
- Extremely Fast,
- Quantifies Uncertainty.

≈ **70%** Prepare data | **28%** Build and train models | **2%** Deploy and predict

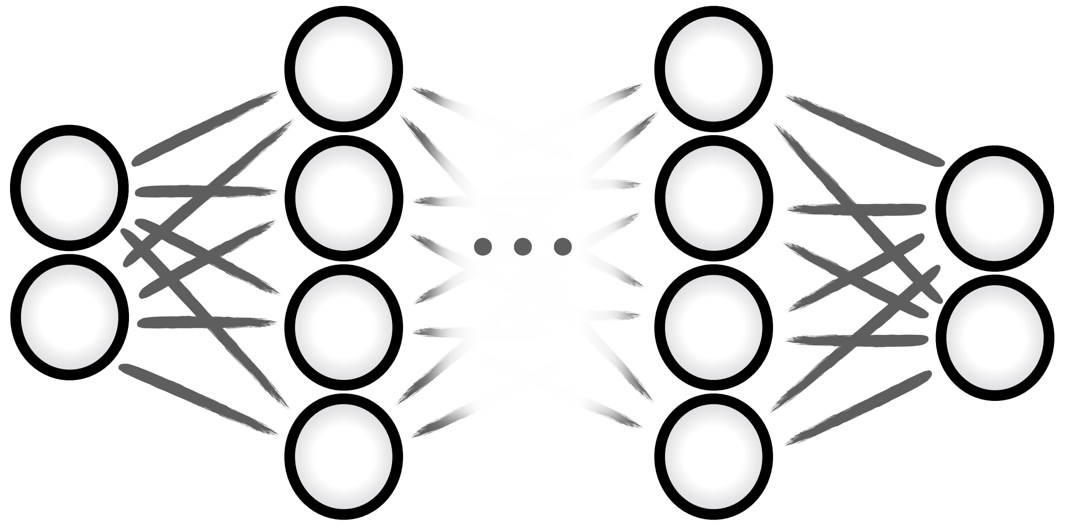
Train

Validation

Test



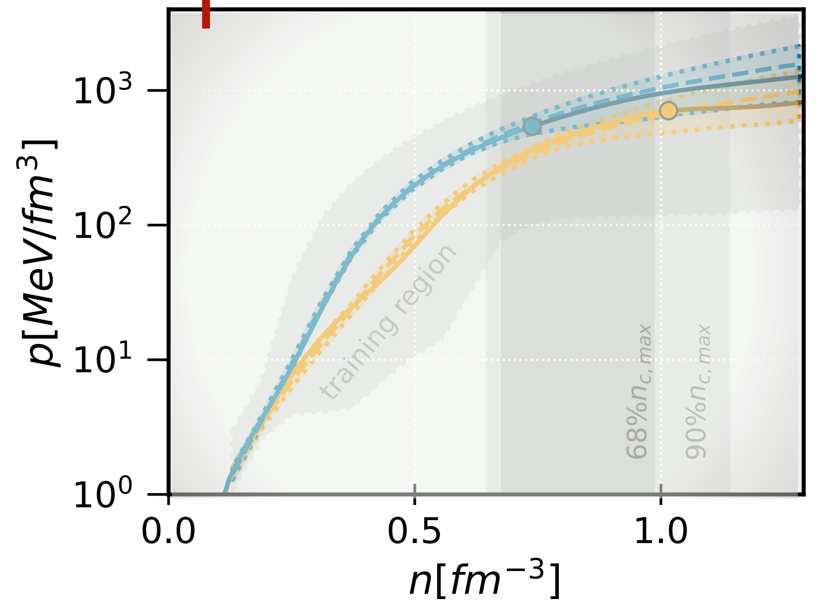
X



Y

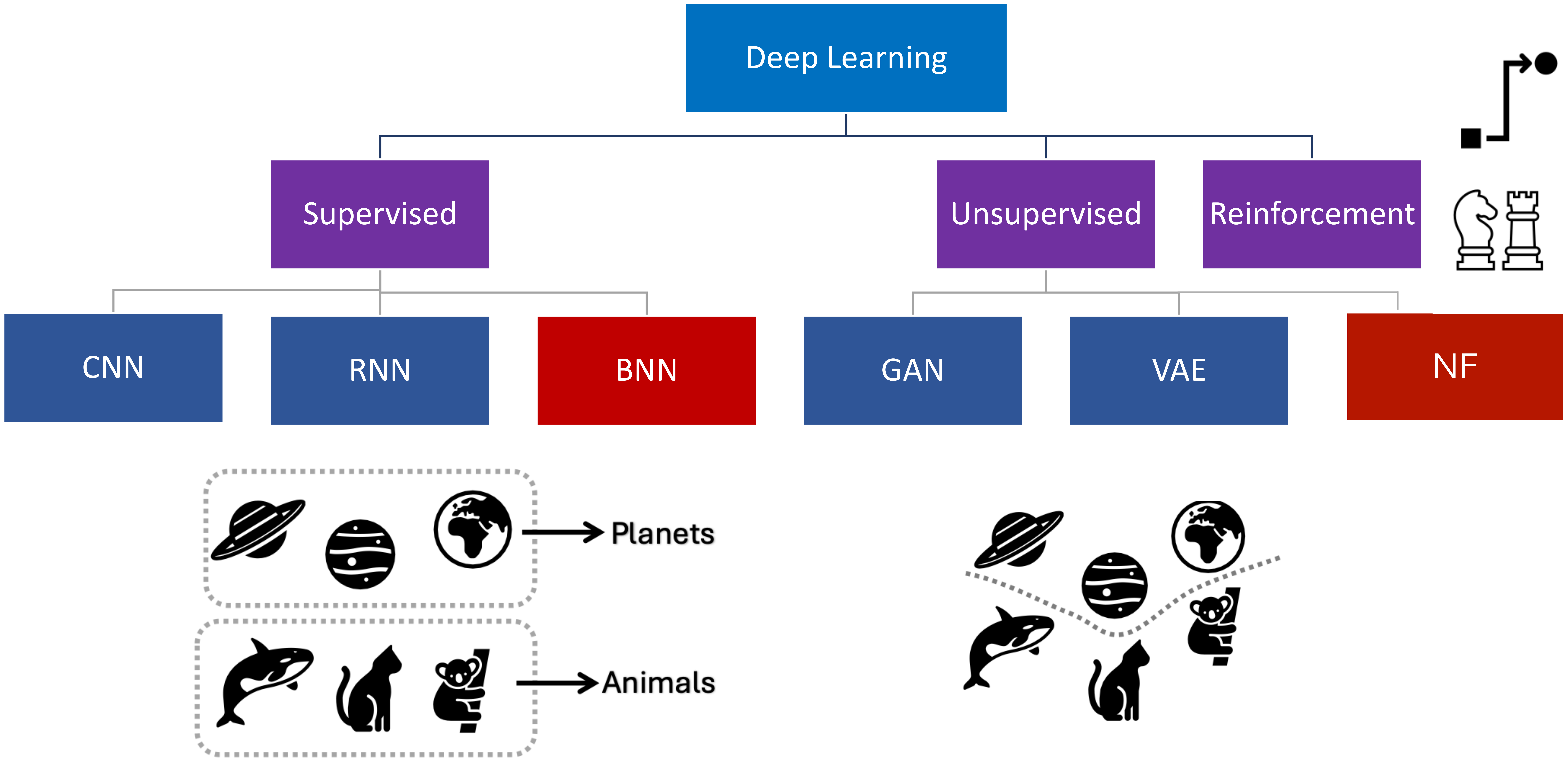
$$f_\theta(x) : X \to Y$$

Spoiler alert !

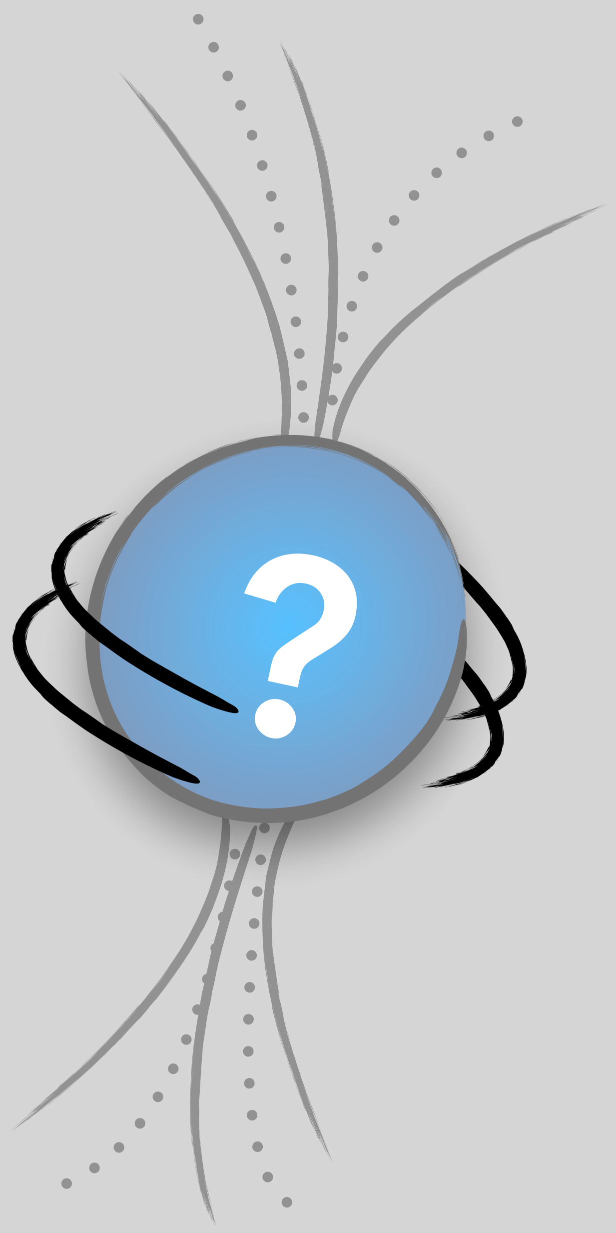


Deep Learning quick recap

Deep Learning taxonomy



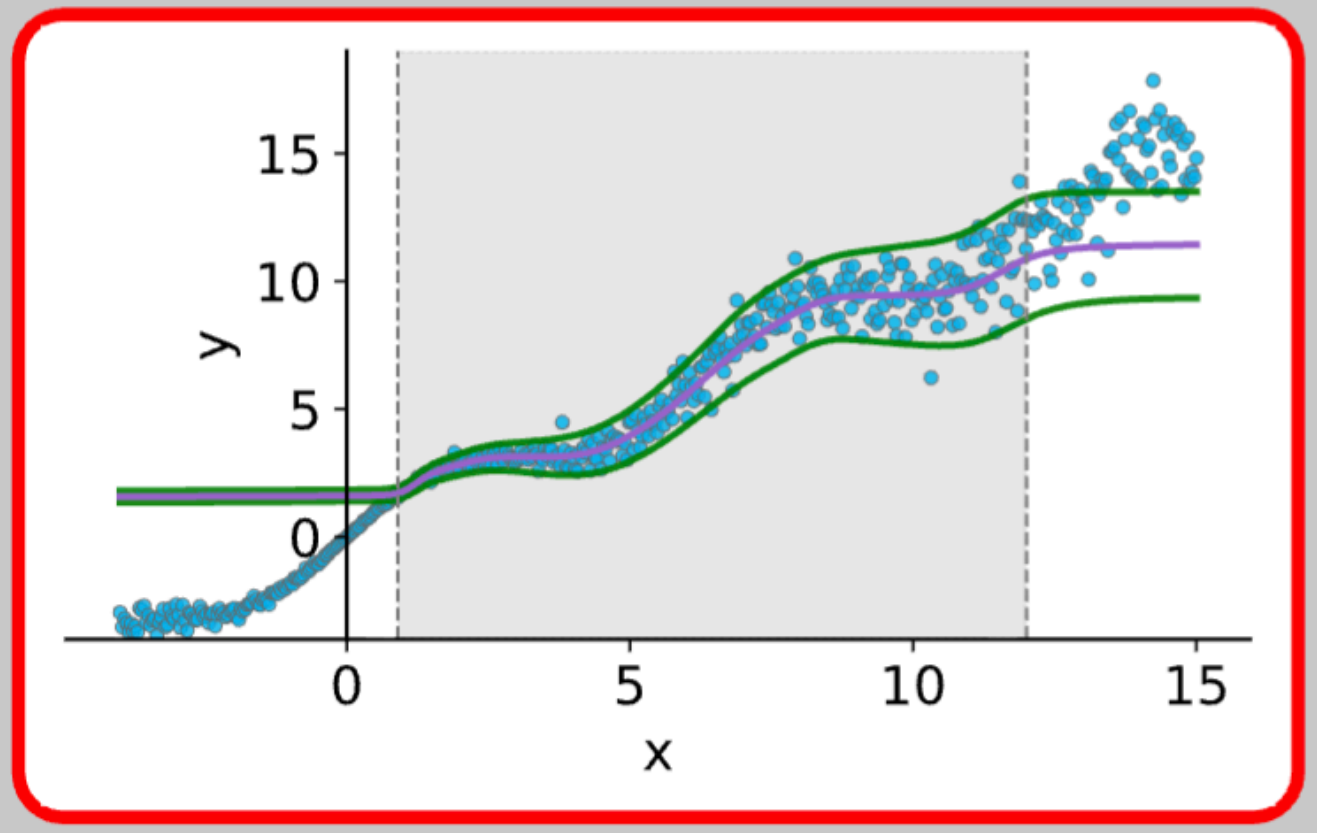
a) Can we infer the proton fraction and speed of sound with uncertainty estimation

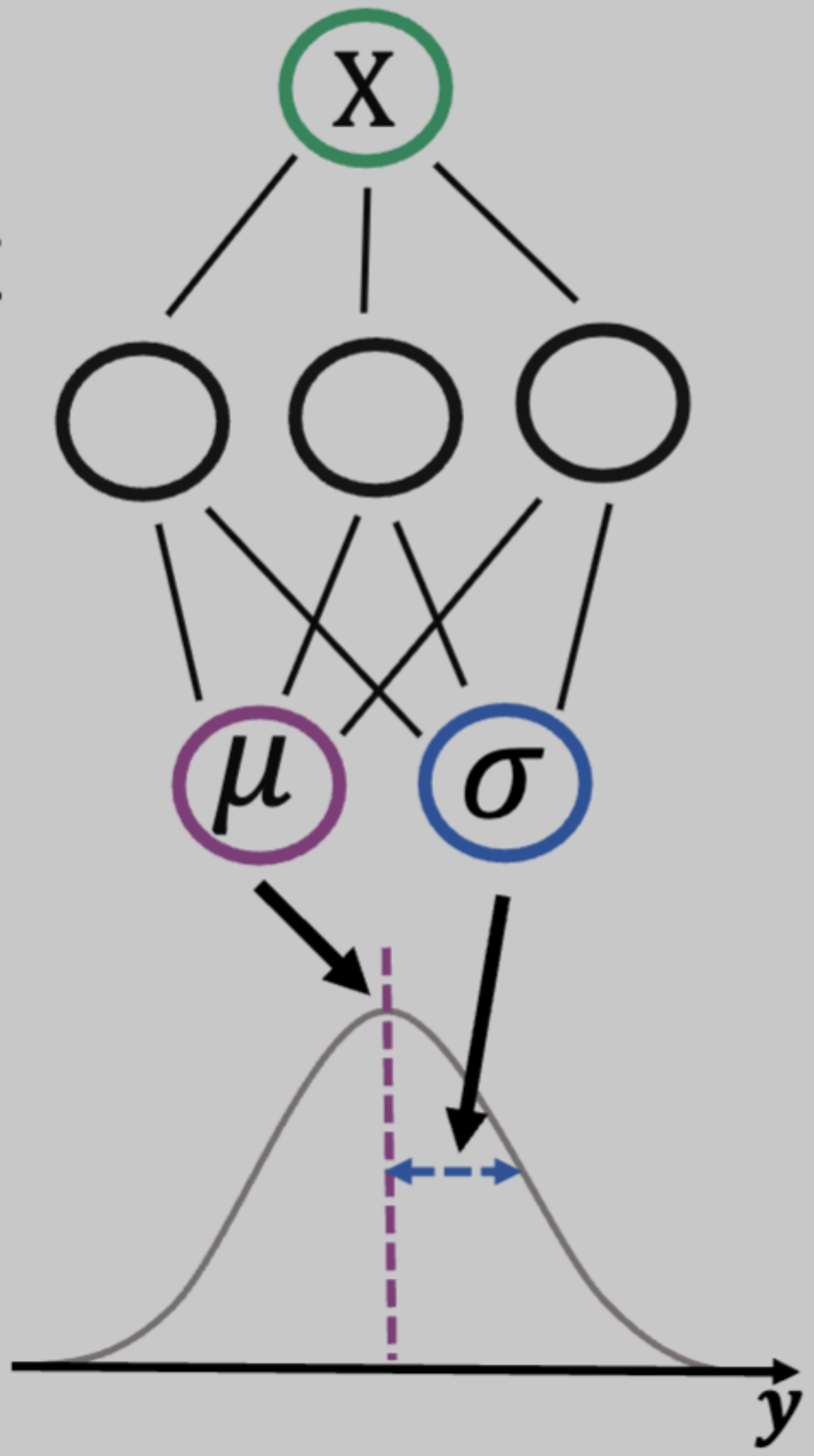


Exploring uncertainty

Aleatoric uncertainty

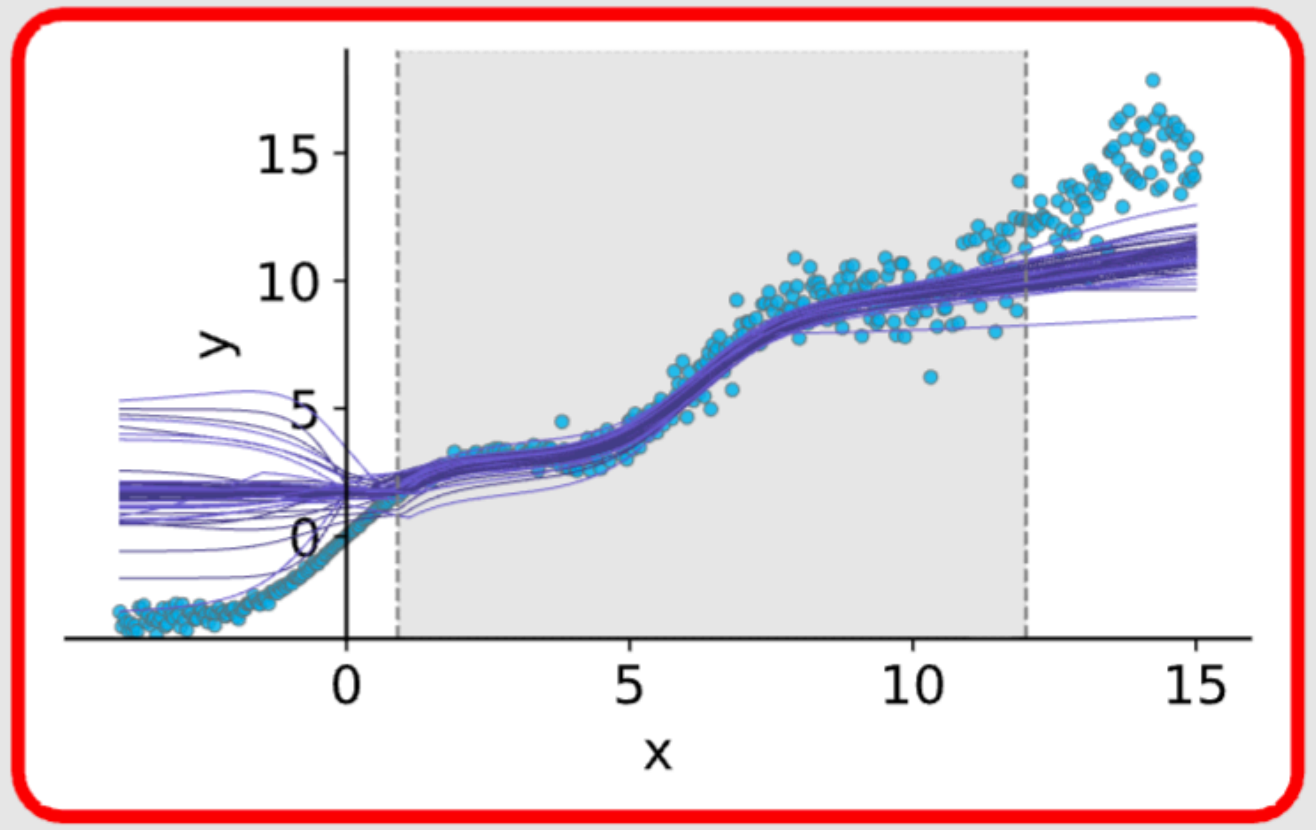
Uncertainty on the dataset
Irreducible

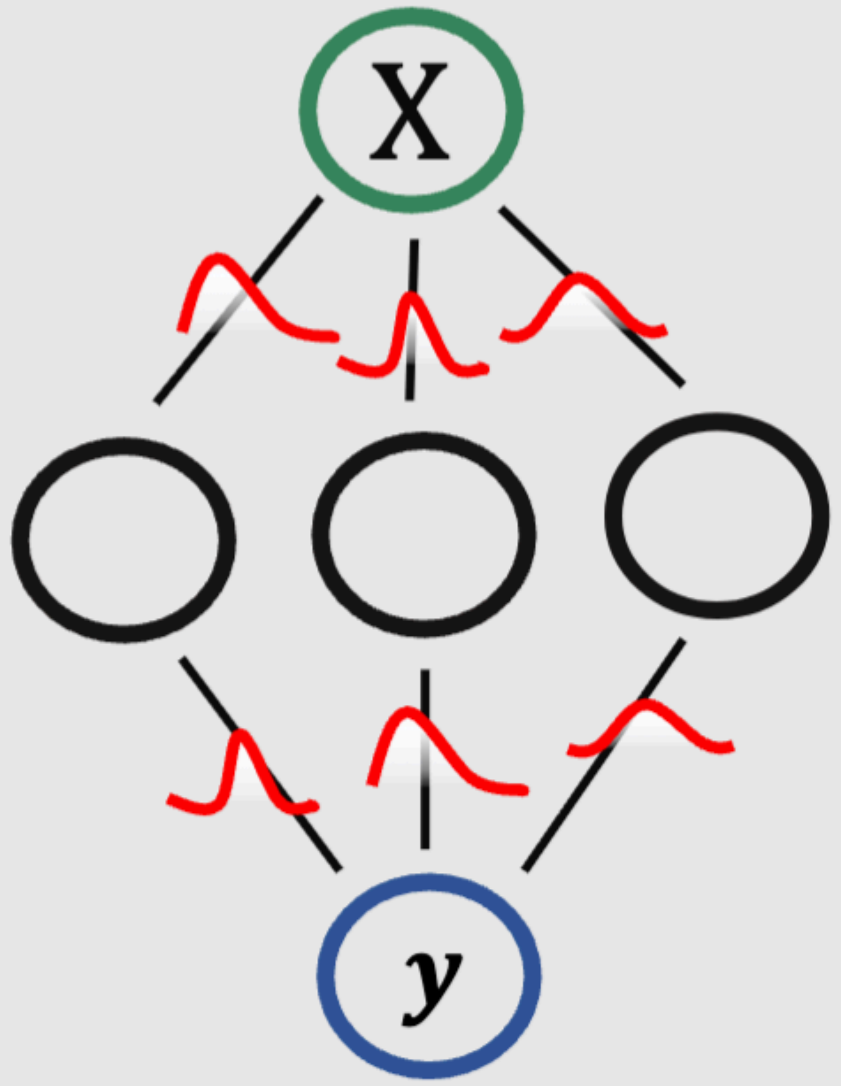




Epistemic uncertainty

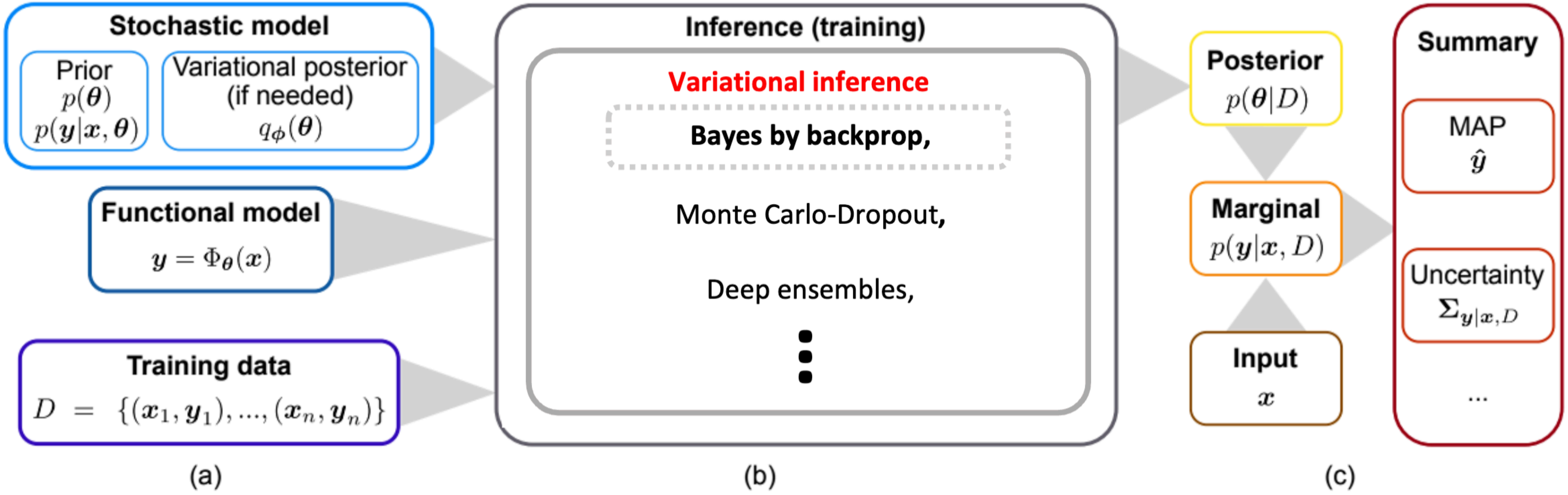
Uncertainty on the underlying function
Reducible





$$y = sin(x) + x(1 + 0.1\epsilon(x)), \quad \epsilon(x) \sim \mathcal{N}(0,1)$$

Bayesian Neural Networks



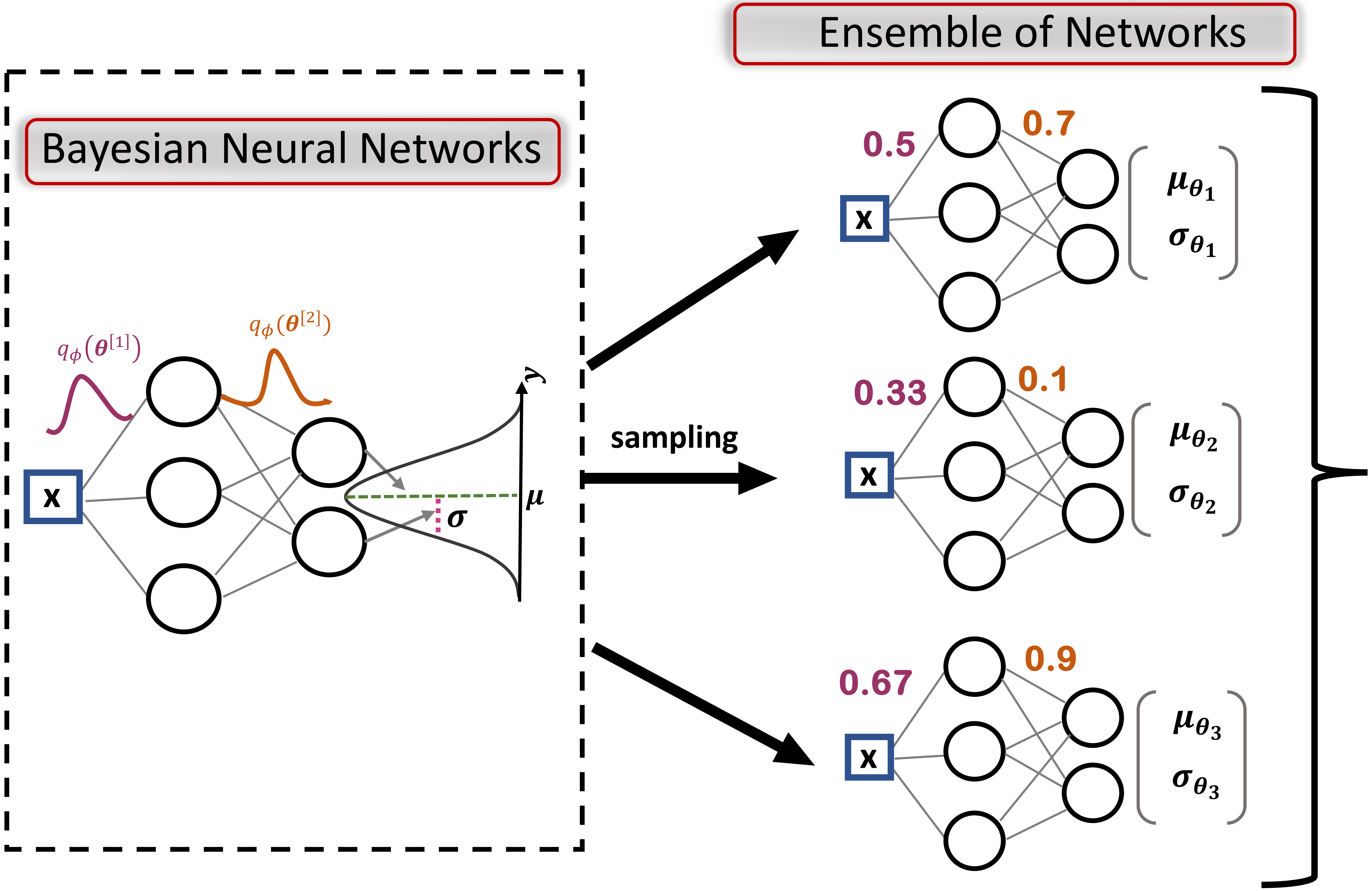
$$q_{\phi^*} = \arg\min_{q_\phi} \mathrm{KL}(q_\phi(\boldsymbol{\theta}) || P(\boldsymbol{\theta}|D)) = \arg\min_{q_\phi} F(D, \phi)$$

$$F(D, \phi) = \mathrm{KL}(q_\phi(\boldsymbol{\theta}) || P(\boldsymbol{\theta})) - \frac{1}{N} \sum_{n=1}^{N} \log P(D|\boldsymbol{\theta}_n)$$

Bayes rule

$$\overbrace{P(\boldsymbol{\theta}|D)}^{\text{Posterior}} = \frac{\overbrace{P(\theta)}^{\text{Prior}}\,\overbrace{P(D|\theta)}^{\text{Likelihood}}}{\underbrace{P(D)}_{\text{Evidence}}}$$

Bayesian Neural Networks prediction



Output

$$\hat{\mu} = \frac{1}{N}\sum_{n=1}^{N} \mu_{\theta_n}$$

$$\hat{\sigma}^2_{alea} = \frac{1}{N}\sum_{n=1}^{N} \sigma^2_{\theta_n}$$

$$\hat{\sigma}^2_{epis} = \frac{1}{N}\sum_{n=1}^{N} \left(\mu_{\theta_n} - \hat{\mu}\right)^2$$

$$\hat{\sigma}^2 = \hat{\sigma}^2_{epis} + \hat{\sigma}^2_{alea}$$

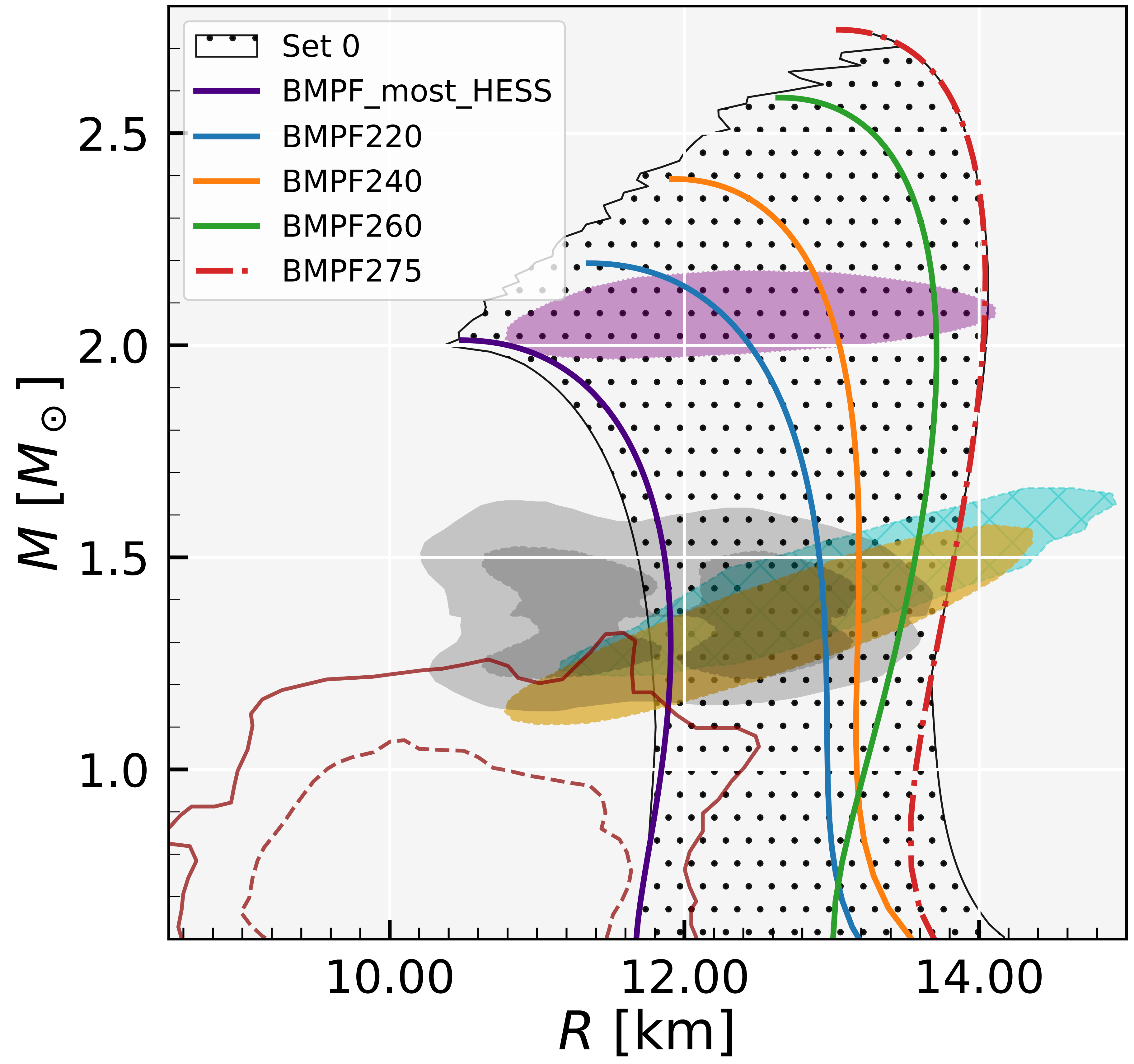
Dataset - source being used

GW170817
HESS J1731-347
PSR J0030 + 0451
PSR J0740 + 6620

Relativistic mean field approximation + Bayesian approach

Fit data

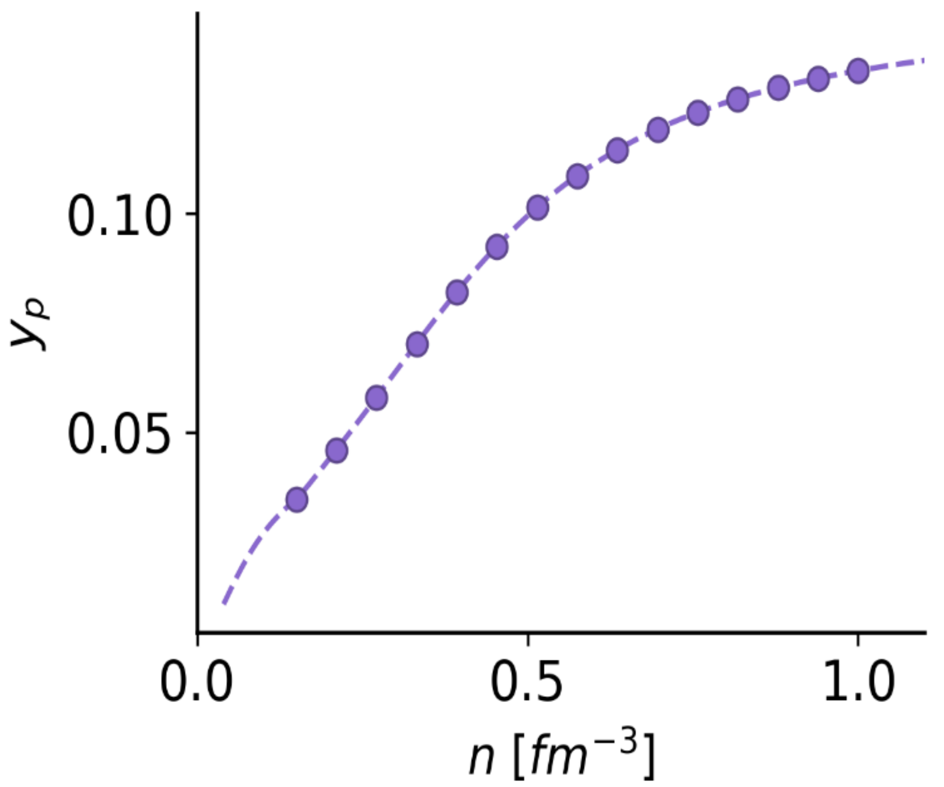
Quantity			Value/Band [Ref.]
NMP	[fm^{-3}]	n_0	0.153 ± 0.005 [Nucl. Phys. A 656]
		ϵ_0	-16.1 ± 0.2 [PRC 90.5]
	[MeV]	K_0	230 ± 40 [Eur. Phys. J. A 30.1 , PR95.122501]
		$J_{\mathrm{sym},0}$	32.5 ± 1.8 [PRC 104.6]
PNM [MeV fm^{-3}]		$P(n)$	2× N^{3}LO [ApJ 773 11]
		dP/dn	> 0
NS mass [$M_\odot$]		M_{max}	> 2.0 [ApJL 915 L12]



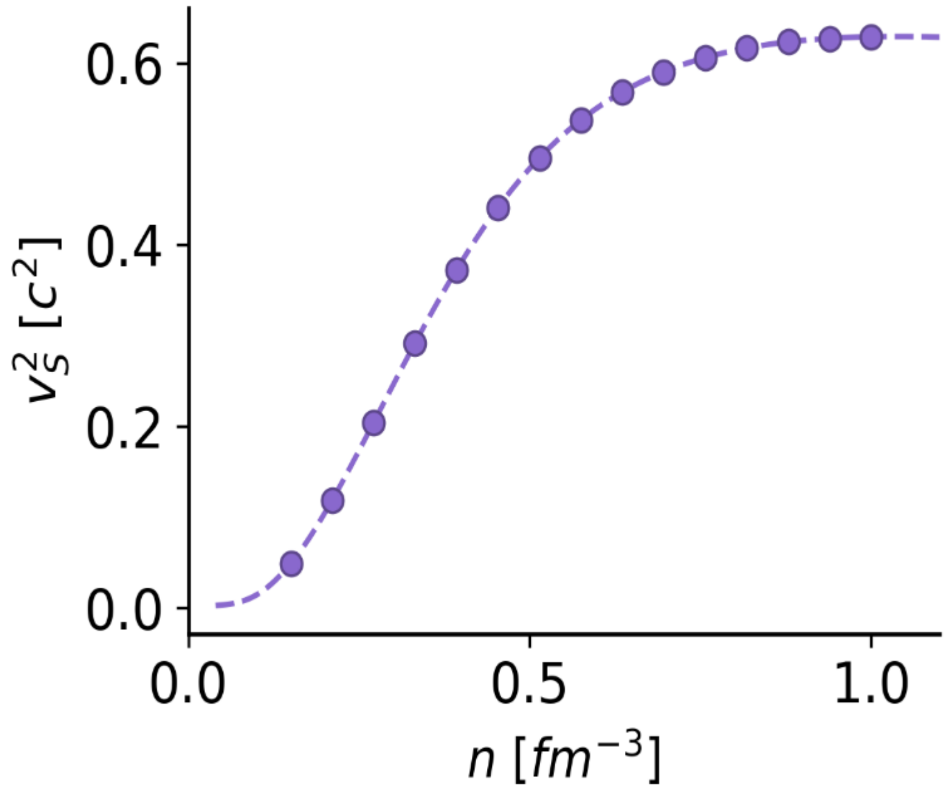
PRD 107.103018

Dataset Creation

Output



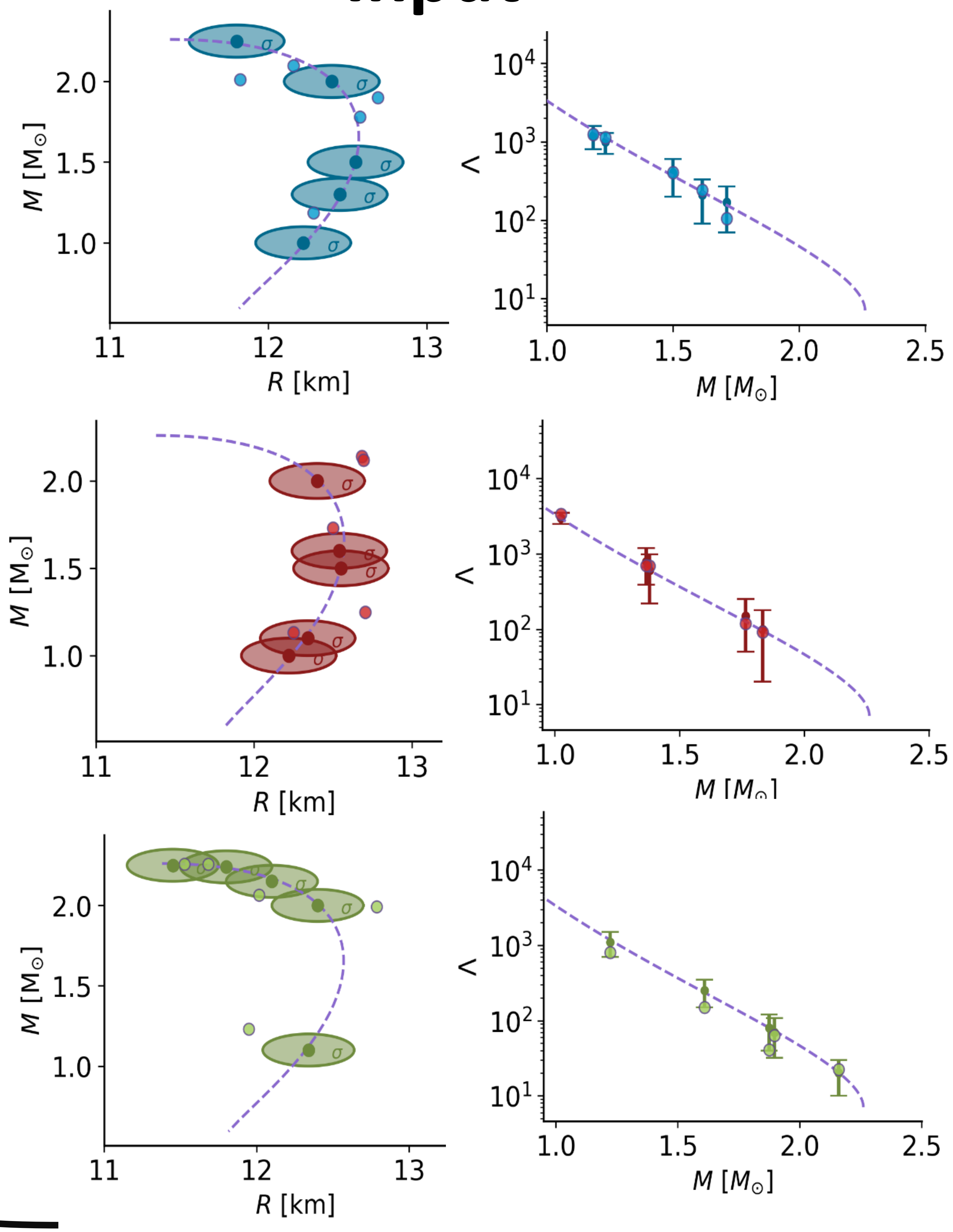
or



$n_s = 3$

X3

Input

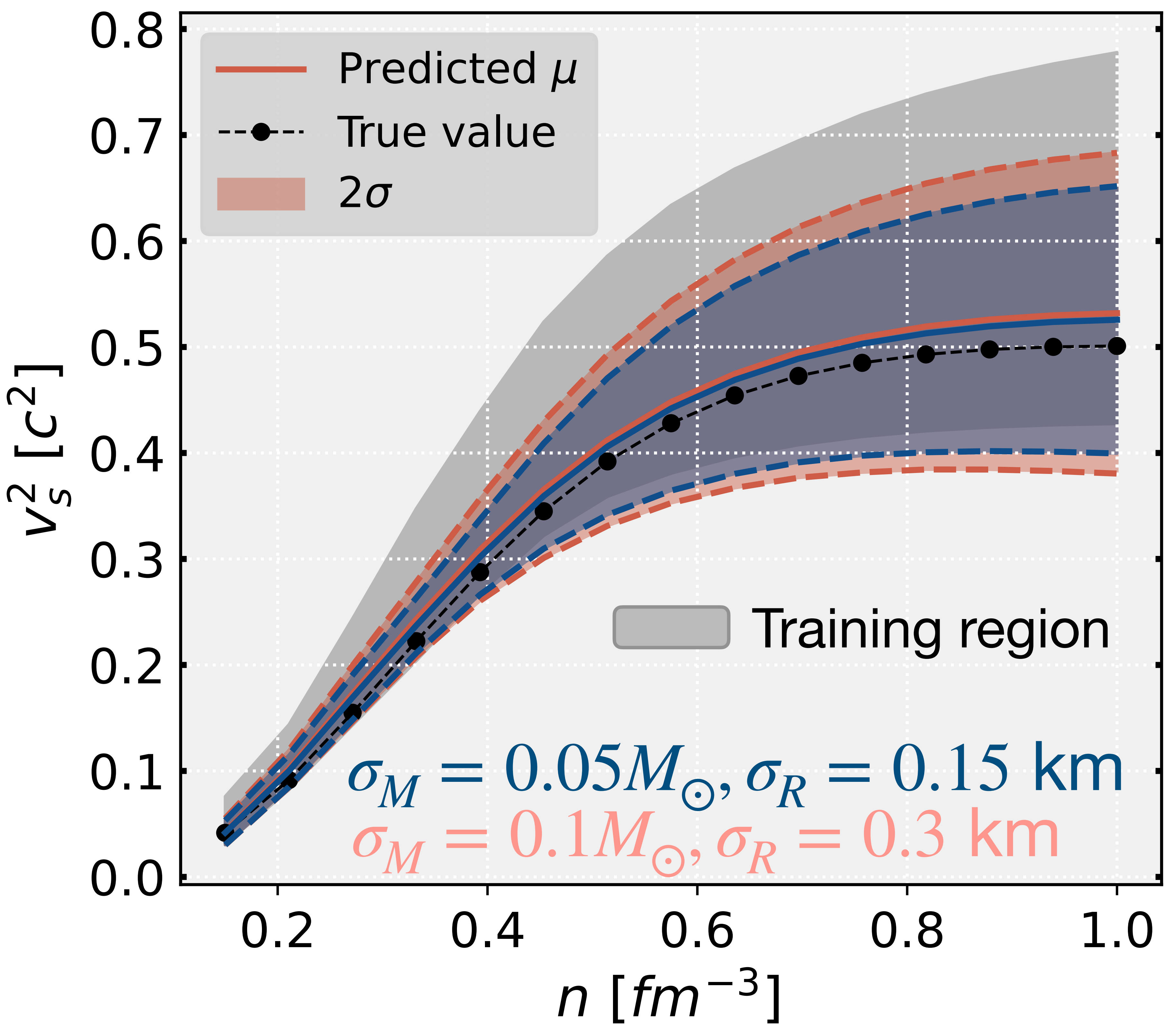


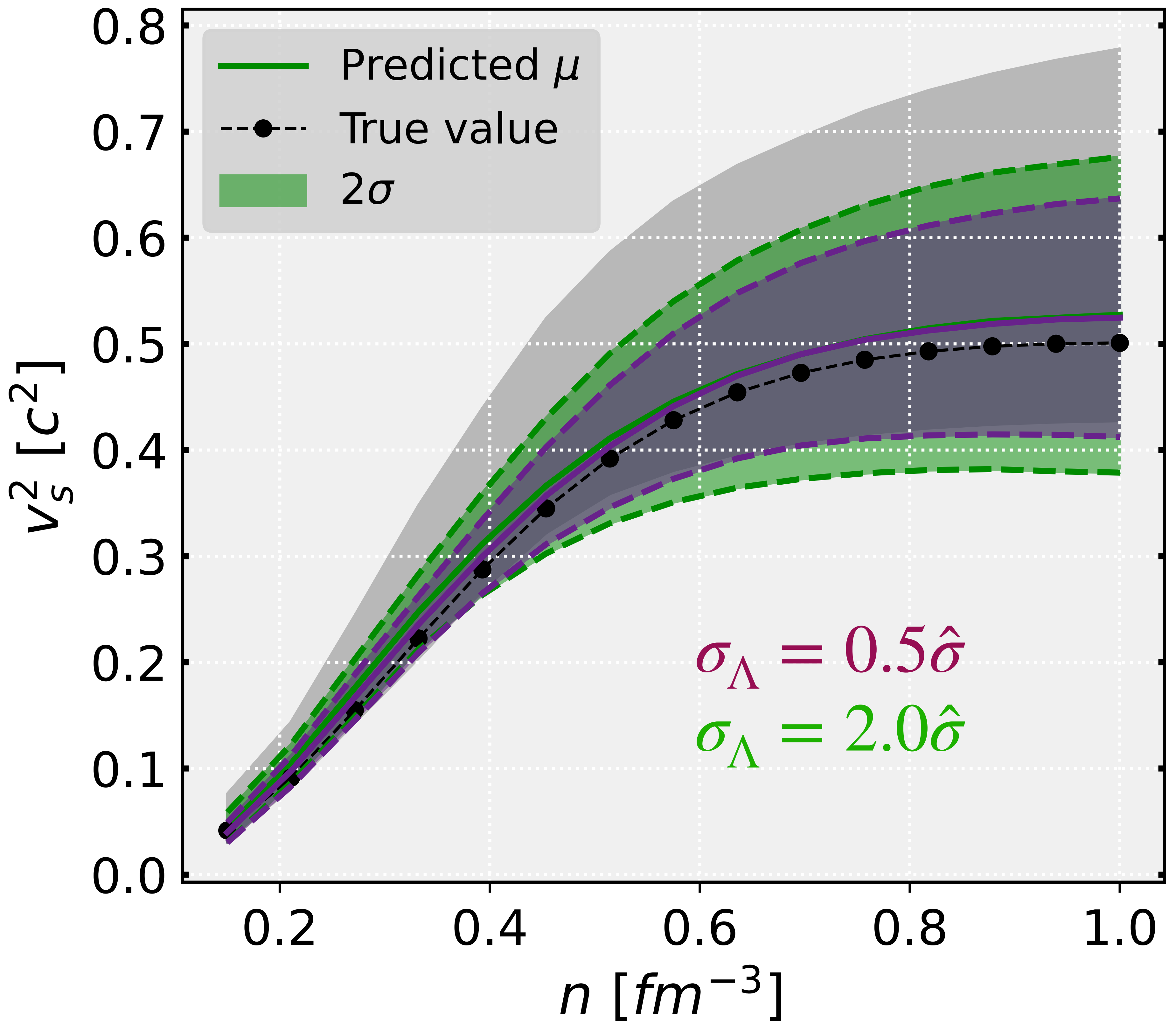
Dataset	$\sigma_M\ [M_\odot]$	σ_R [km]	$\sigma_\Lambda(M_j)$
1	0.05	0.15	—
2	0.1	0.3	—
3	0.1	0.3	$0.5\hat{\sigma}(M_j)$
4	0.1	0.3	$2\hat{\sigma}(M_j)$

8 different datasets

8 different Models

Out-of-sample accuracy : speed of sound

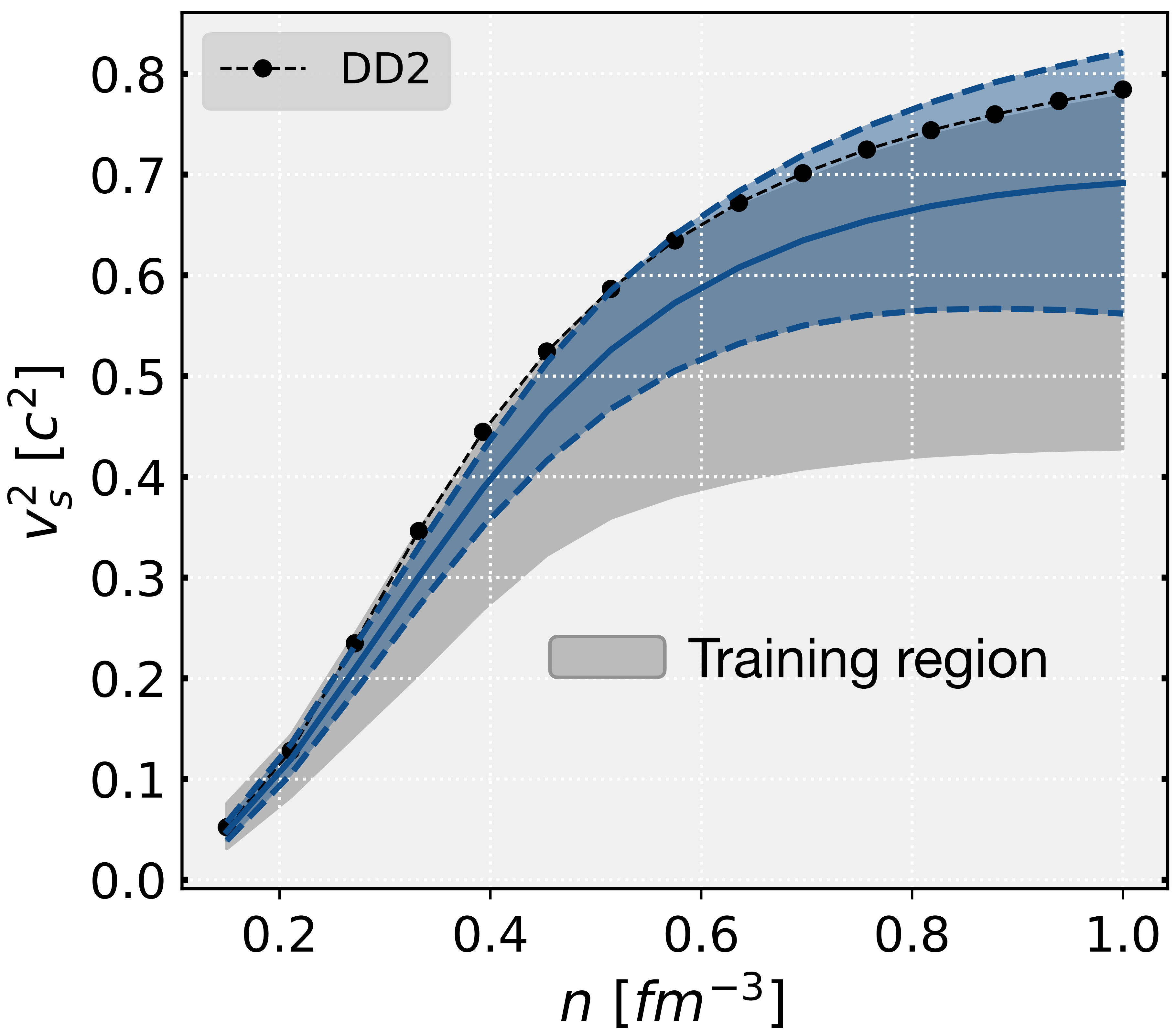


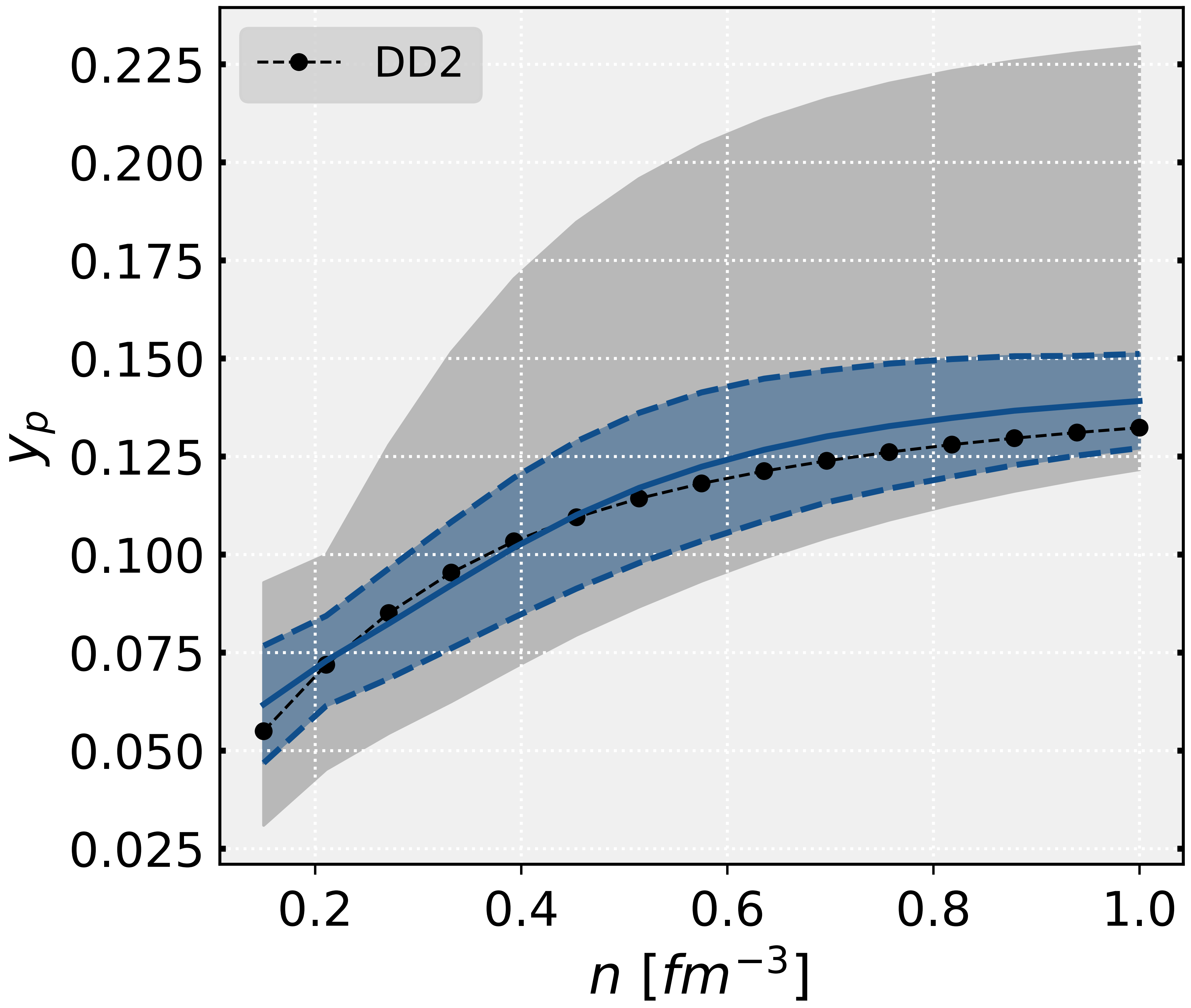


Larger uncertainty prediction with increased input noise

Out-of-sample accuracy : different framework

Model trained with Dataset 1



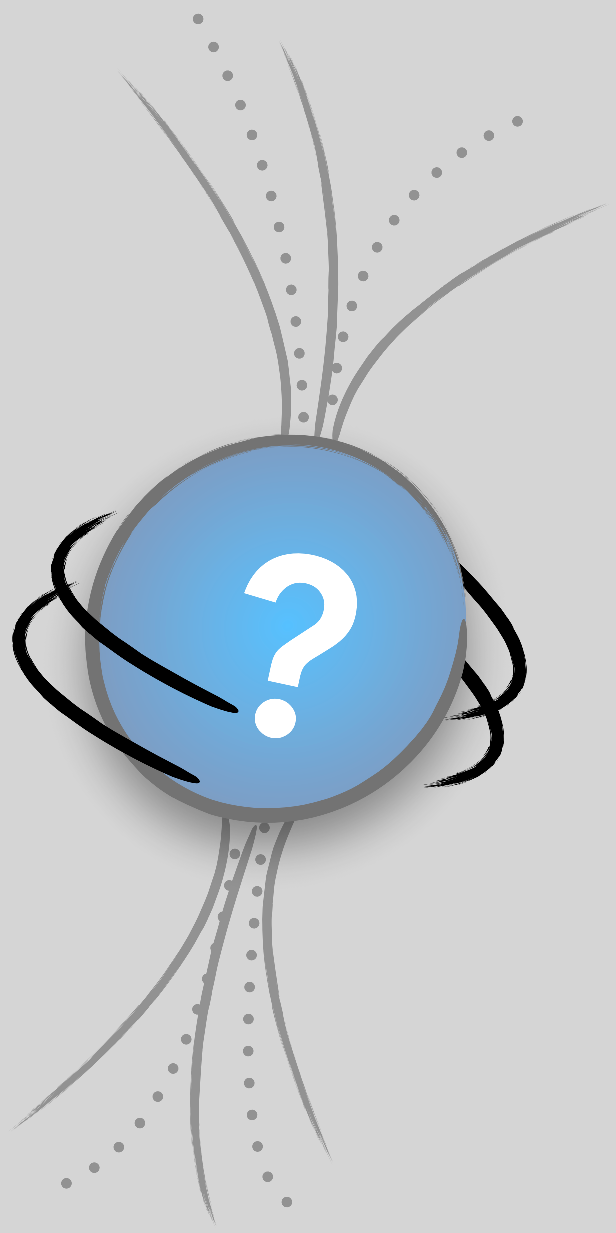


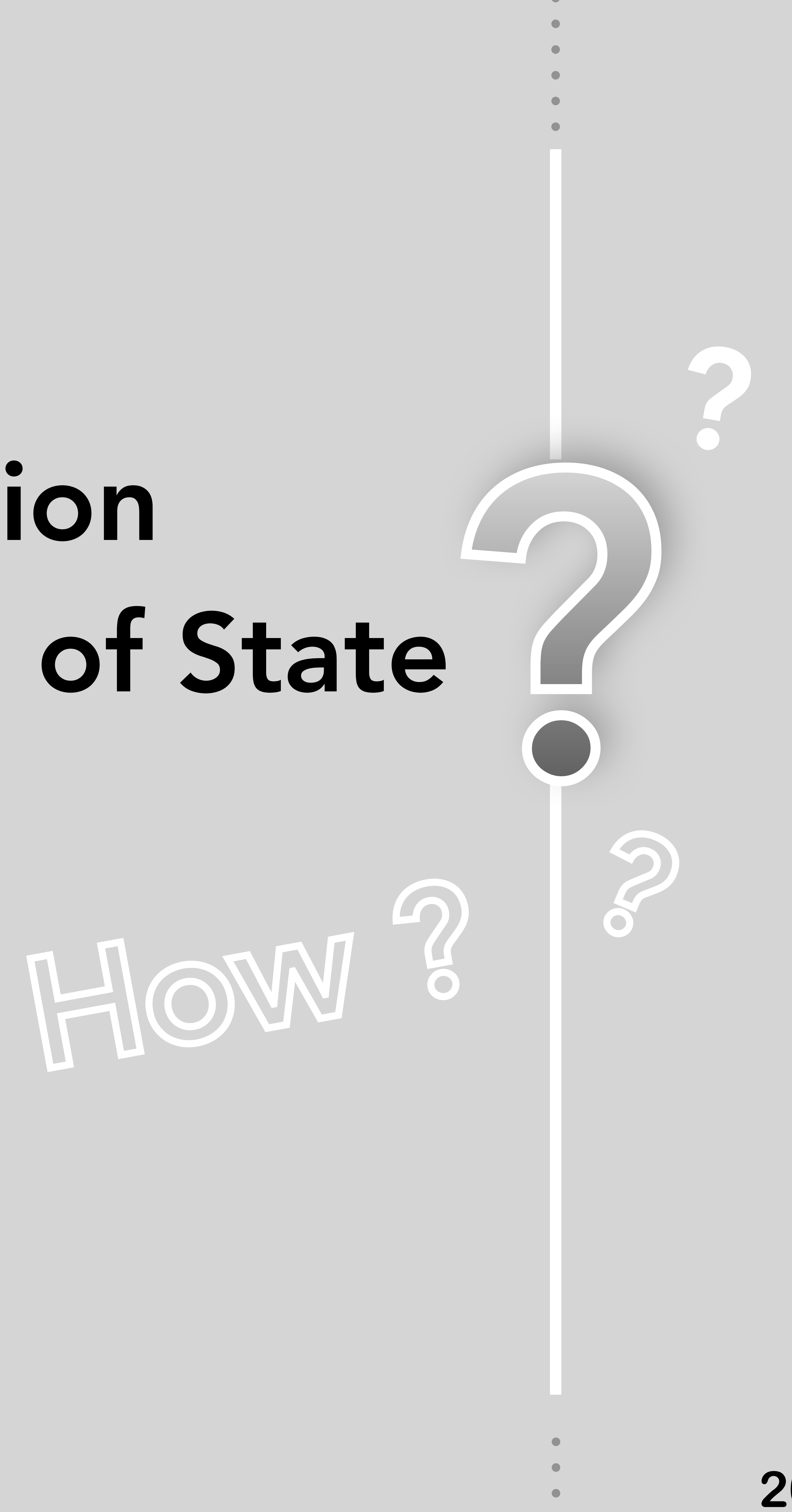
DD2 : Generalized RMF model with density-dependent couplings

Summary

- Bayesian neural networks accurately connect observations to v_s^2 and y_p while measuring their uncertainty
- They are able to distinguish how Λ affects the uncertainty of different outputs
- Effectively capture both types of uncertainty: epistemic and aleatoric
- Correctly predicts the output for a distinct dataset

b) Can Neural Posterior Estimation Infer the Neutron Star Equation of State

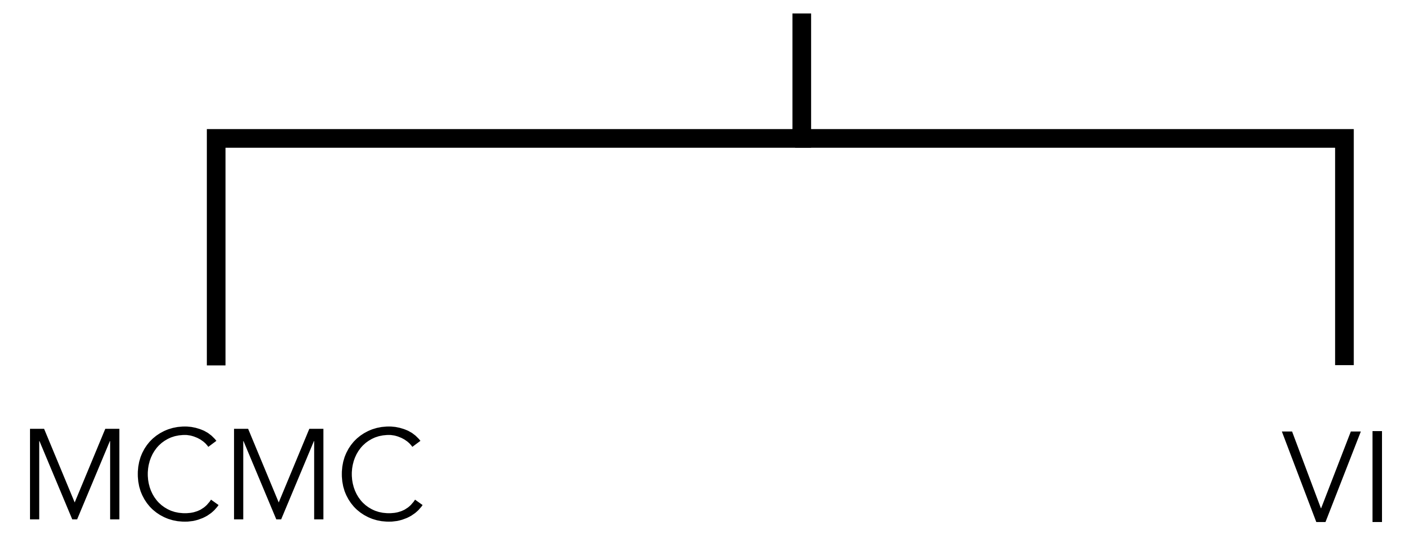




Likelihood-based or simulation-based?

Likelihood-based Inference

Evaluate $p(d\,|\,\theta)$

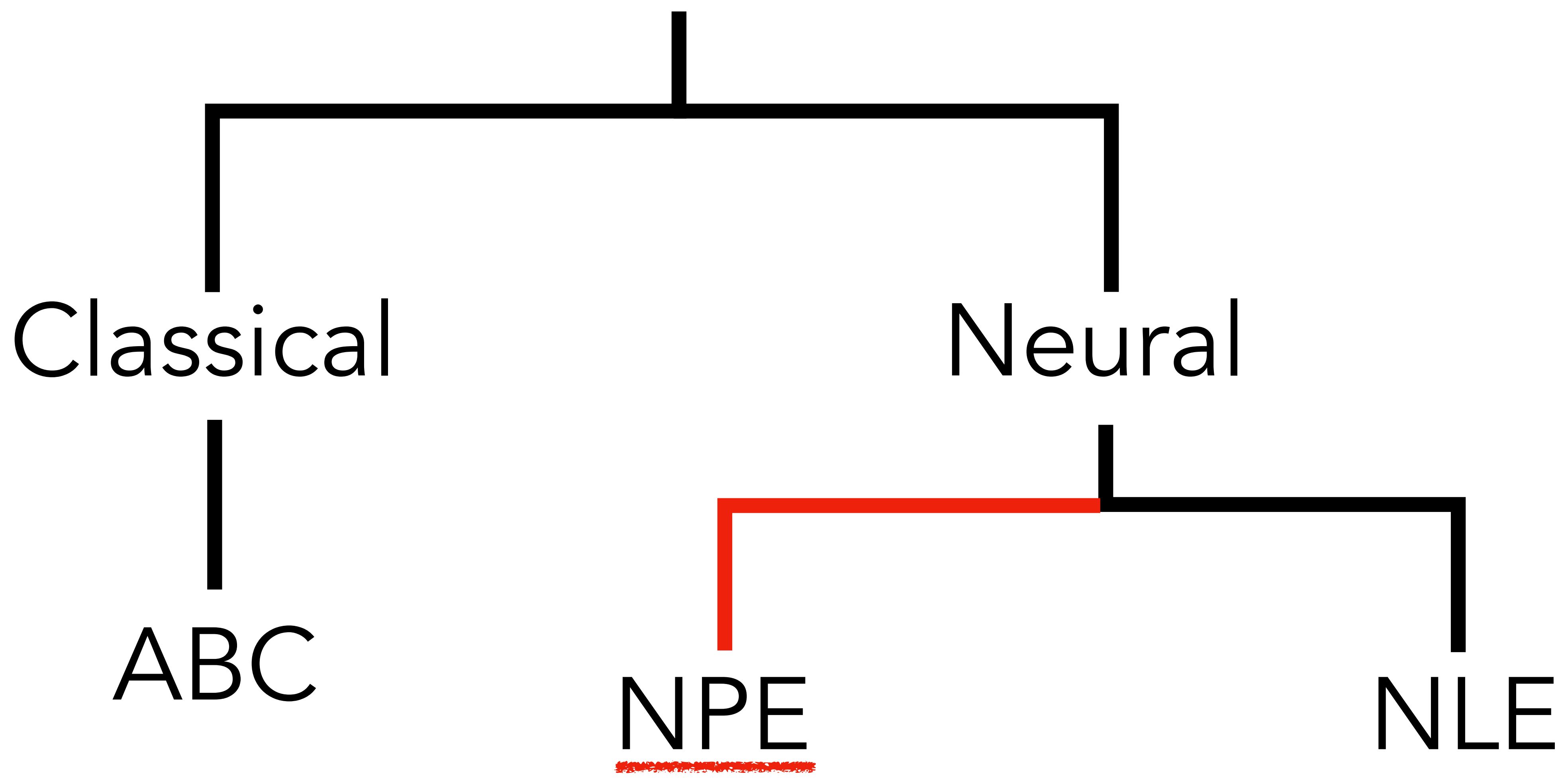


Requires a likelihood

Slow inference

Simulation-based Inference

Sample $d \sim p(d\,|\,\theta)$



Doesn't require a likelihood

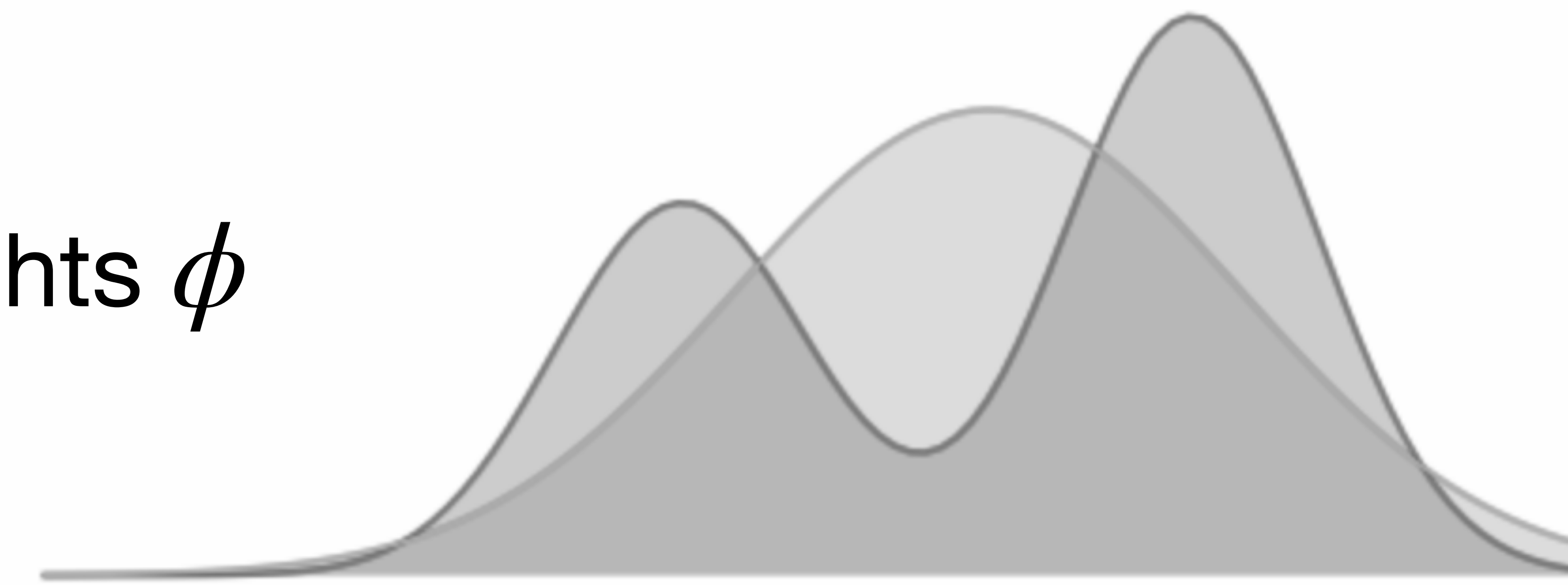
Slow training

Fast Inference

Neural posterior estimation

Learning an amortised posterior estimation $q_\phi(\theta | d)$

Forward



Trains with the goal:

$$q_\phi(\theta | d) \approx p(\theta | d)$$ w.r.t the weights ϕ

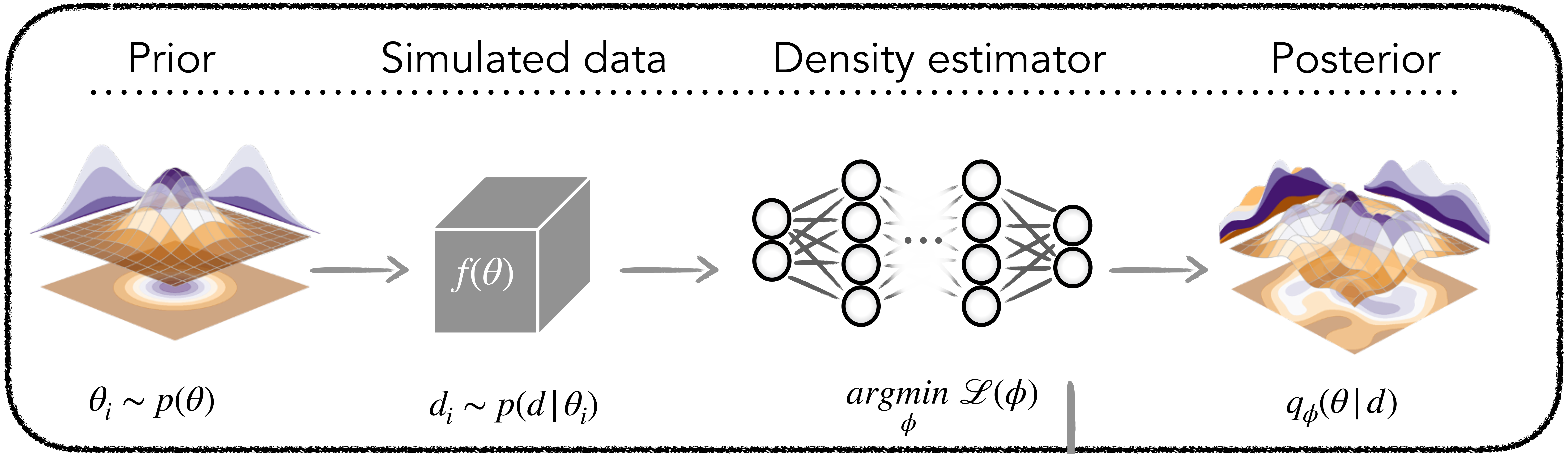
Which translates to:

$$\phi^* = \underset{\phi}{argmin}\ \mathbb{E}_{d \sim p(d)} KL(p(\theta | d) | q_\phi(\theta | d))$$

$$= \underset{\phi}{argmin}\ - \mathbb{E}_{\theta \approx p(\theta)} \mathbb{E}_{d \approx p(d|\theta)} [\log q_\phi(\theta | d)]$$

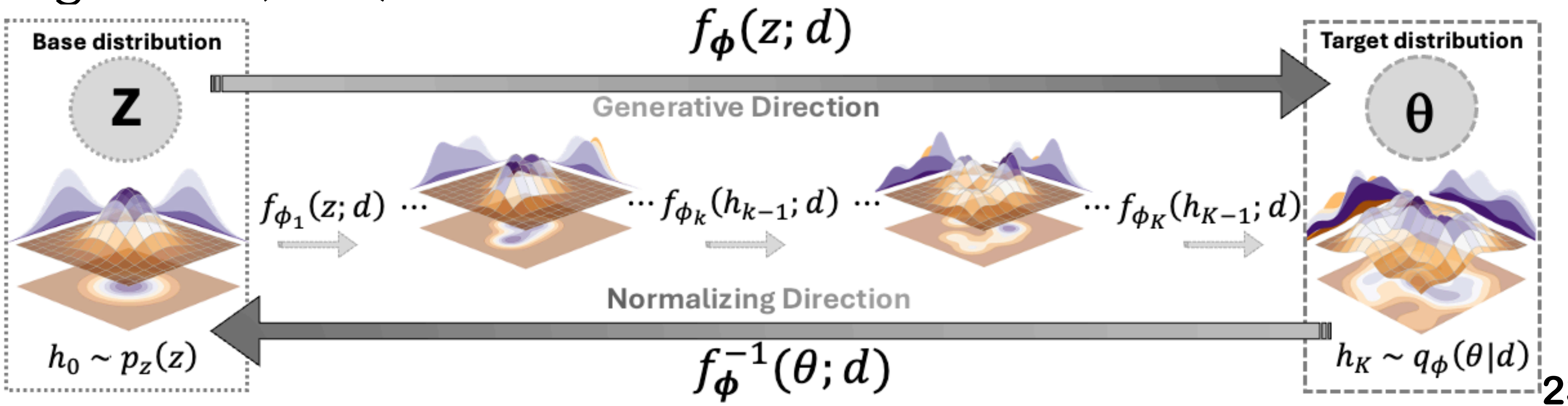
$$KL(p(x) | q(x)) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} dx = \begin{cases} 0 & \text{if } p(x) = q(x) \\ ]0,\infty[& \text{if } p(x) \neq q(x) \end{cases}$$

Neural posterior estimation



Conditional Normalizing Flows (CNF)

- ☑ Invertible
- ☑ Flexible
- ☑ Bijective



Dataset - Quantities we aim to predict

ation can recover the neutron star EoS from

ng the posterior $p(EoS\,|\,O)$.

The predicted physical quantities :

$$\begin{cases} \boldsymbol{p}(\boldsymbol{n}) = [p(n_1), p(n_2), \cdots, p(n_{20})], \\ \boldsymbol{c_s^2}(\boldsymbol{n}) = [c_s^2(n_1), c_s^2(n_2), \cdots, c_s^2(n_{20})], \\ \boldsymbol{\Delta}(\boldsymbol{n}) = [\Delta(n_1), \Delta(n_2), \cdots, \Delta(n_{20})], \end{cases}$$

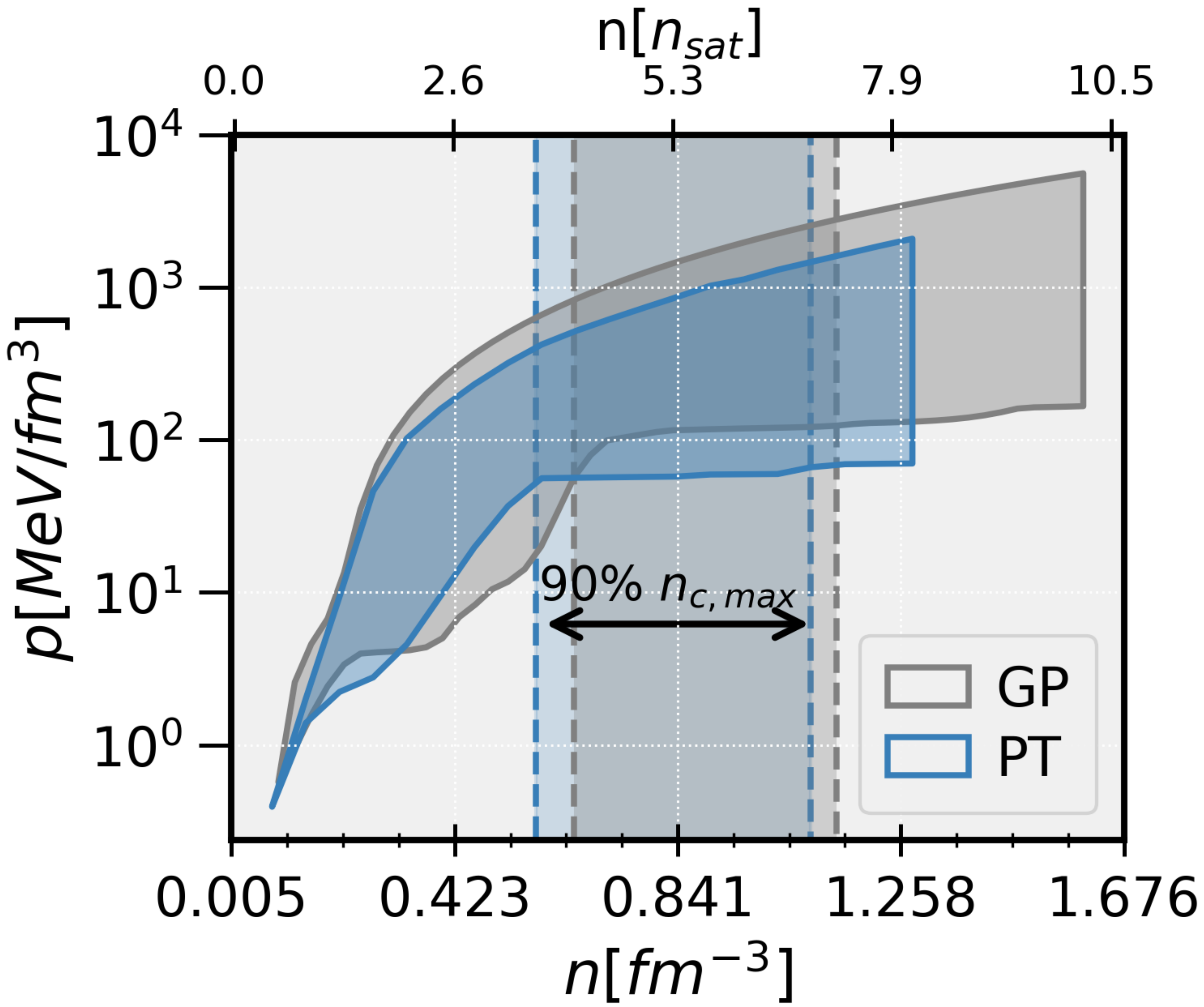
Two agnostic models :

Piece-wise Polytropics (PT)

PRD **111**,023035 (2025)

Gaussian Processes (GP)

Nat Commun **14**, 8451 (2023)



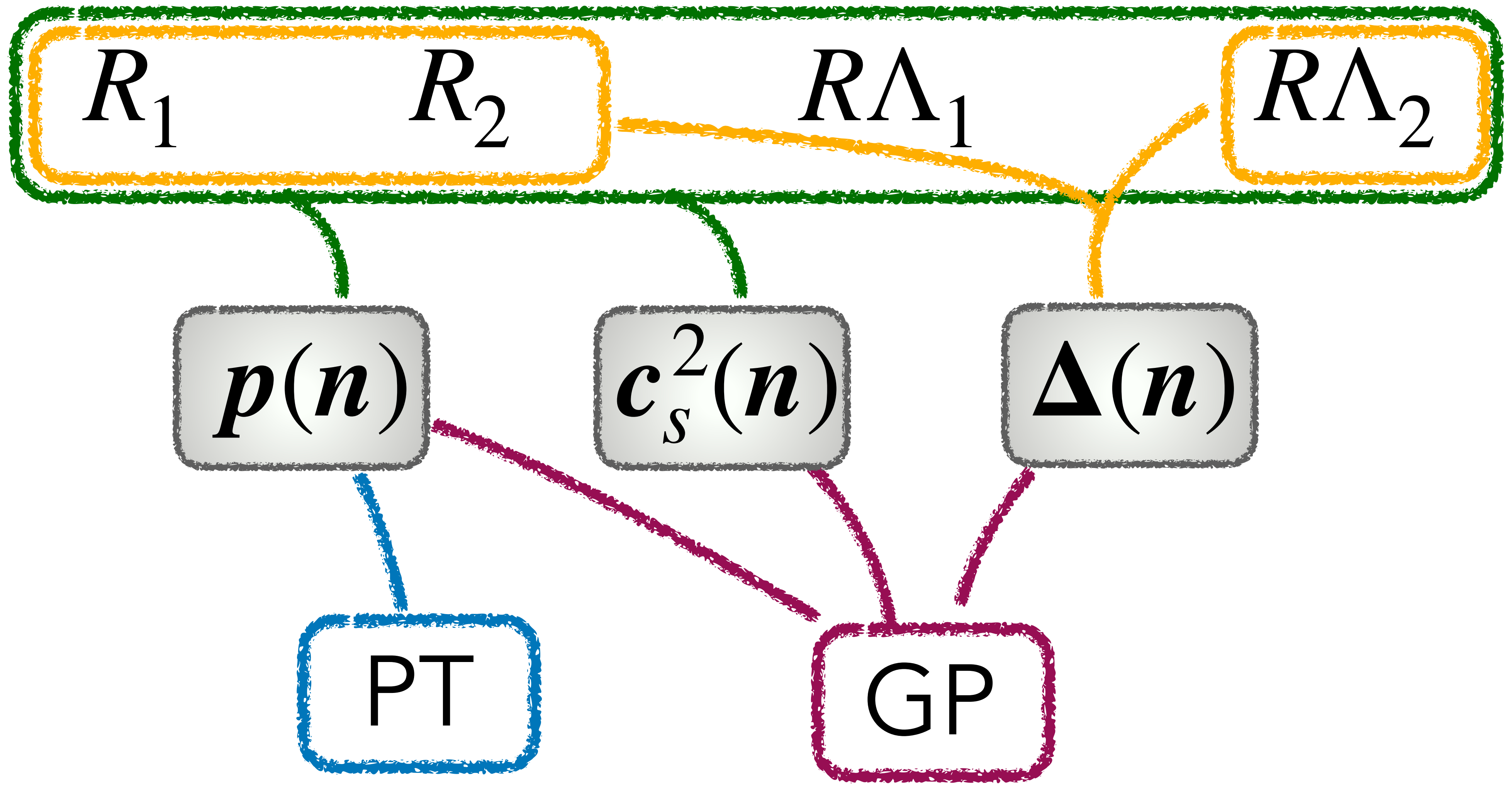
Dataset - Quantities we condition

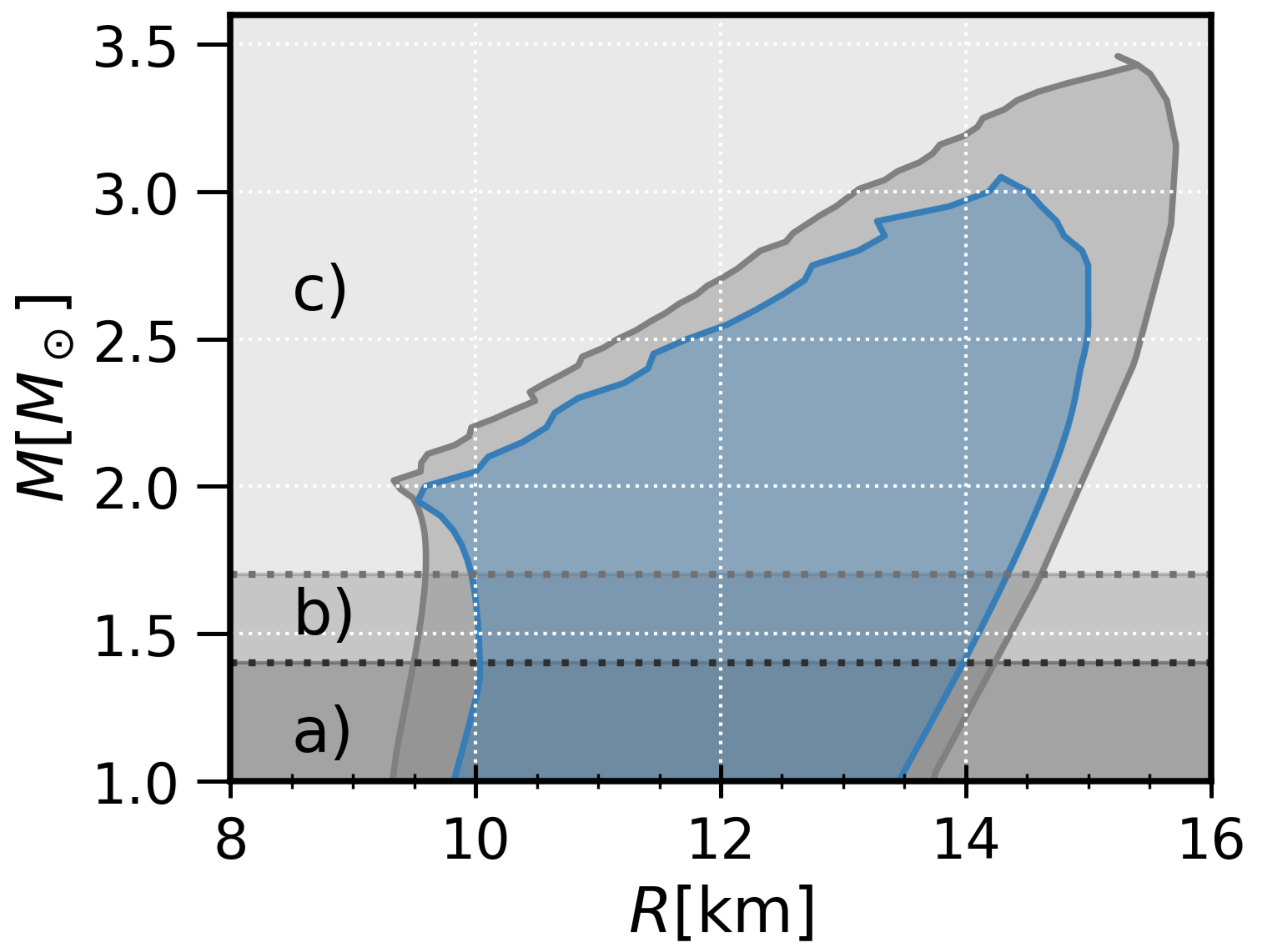
The conditioned quantity :

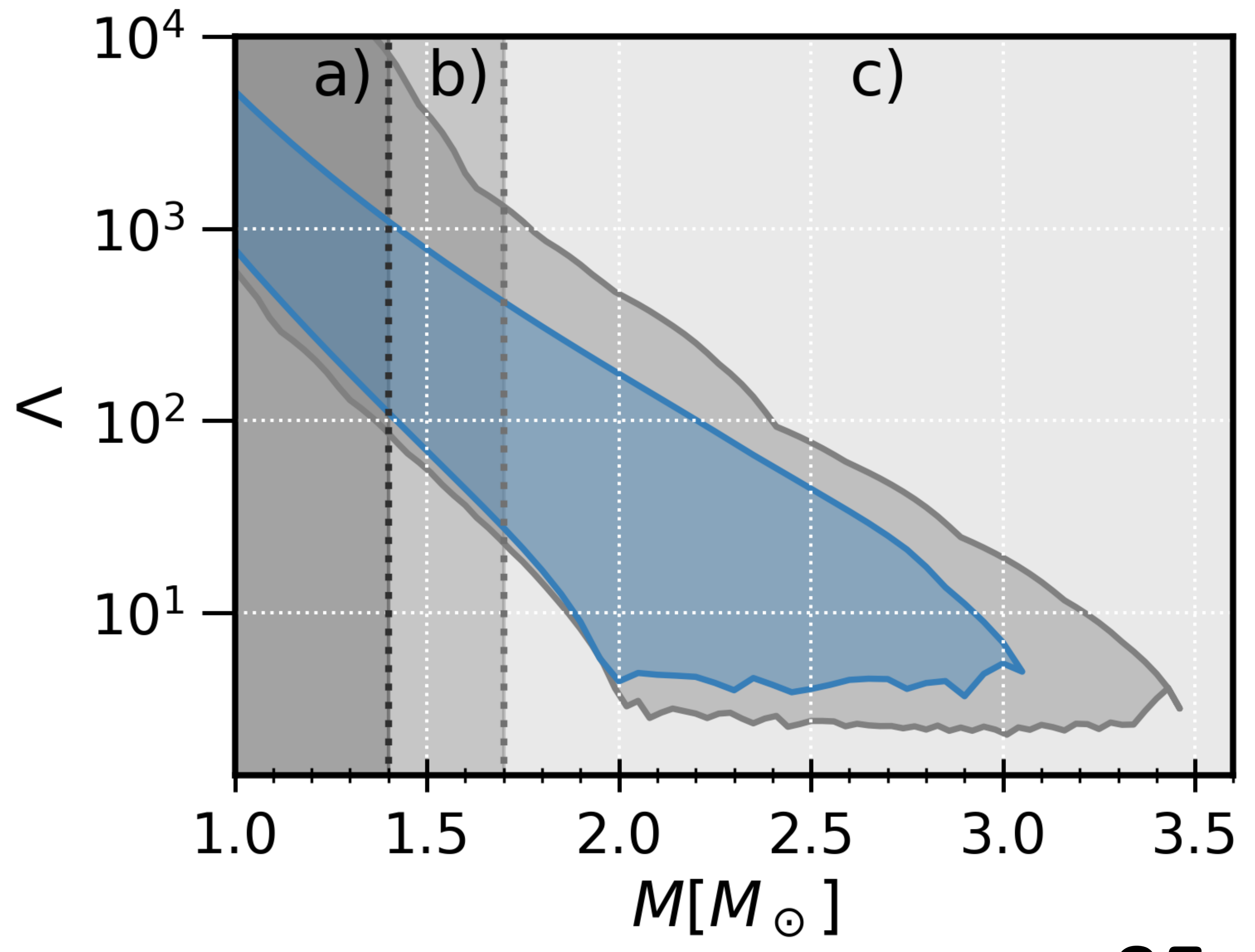
$$\begin{cases} R_x = [M_1, \cdots, M_{15}, R_1, \cdots, R_{15}], x \in [1,2] \\ R\Lambda_x = [M_1, \cdots, M_{15}, R_1, \cdots, R_{15}, M_1^*, \cdots, M_{15}^*, \Lambda_1, \cdots, \Lambda_{15}]\,. \end{cases}$$

$x = 1$ without noise, $x = 2$ with gaussian noise.

The datasets:

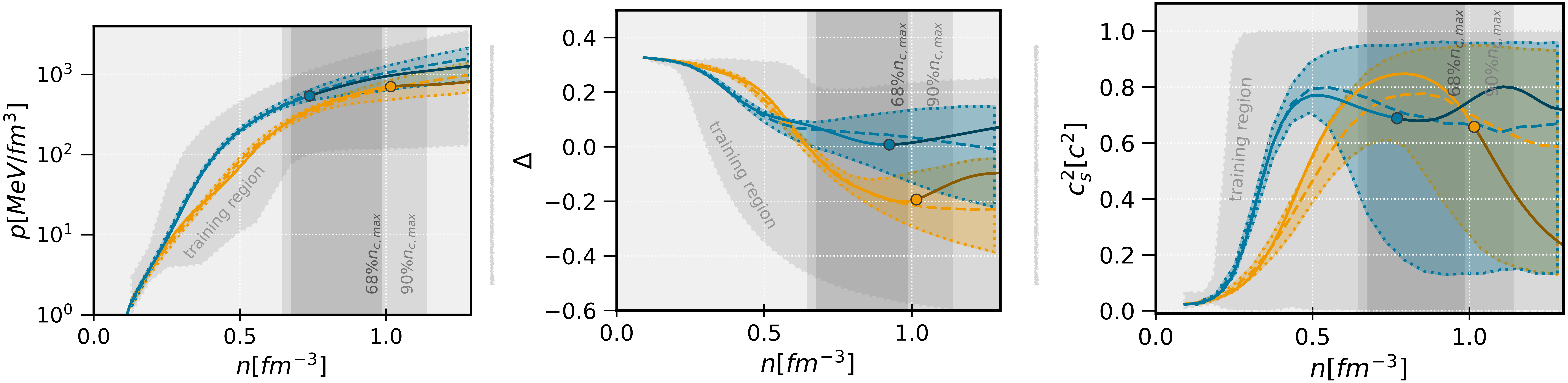






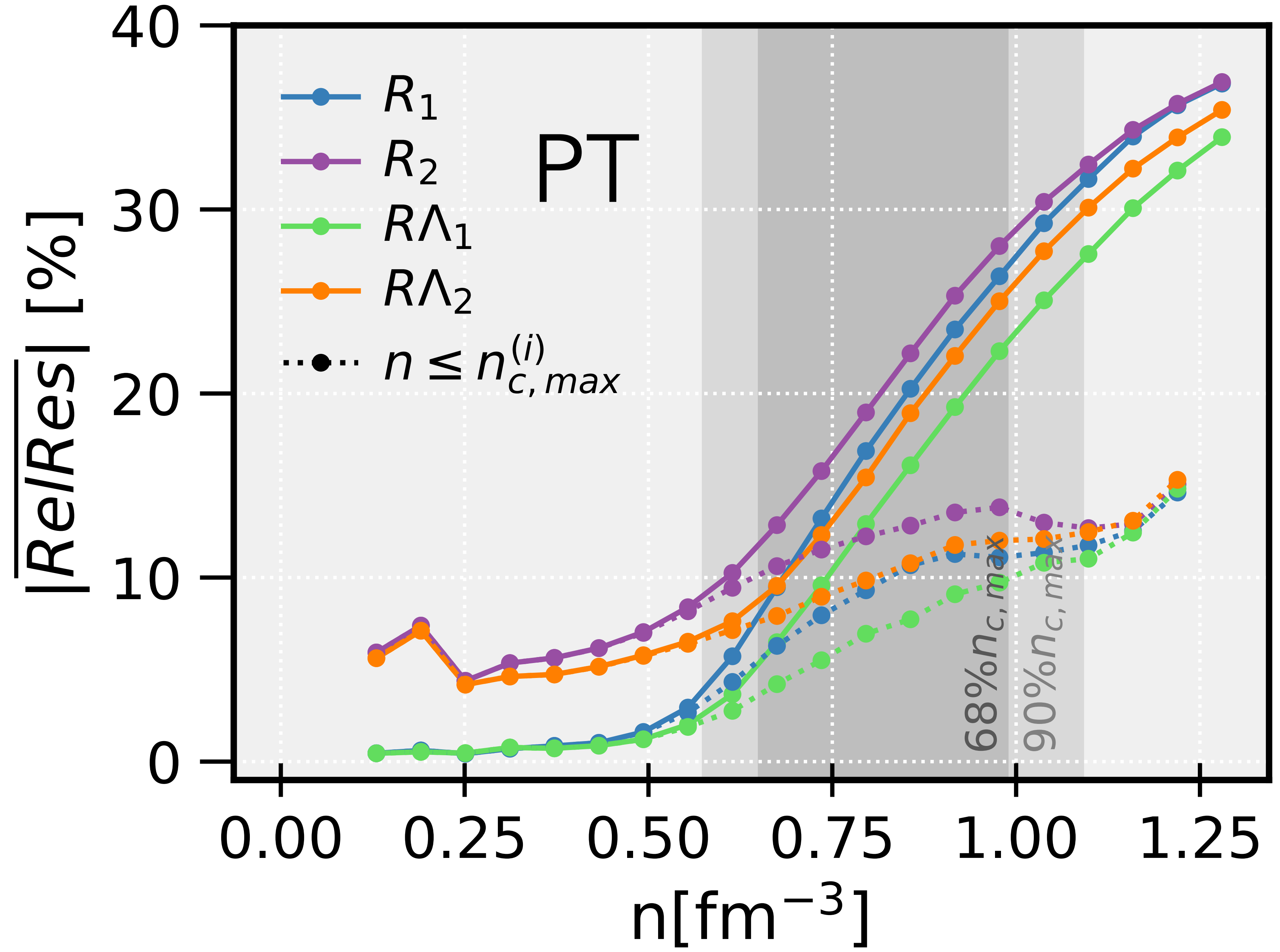
Key result: Accurate EoS reconstruction

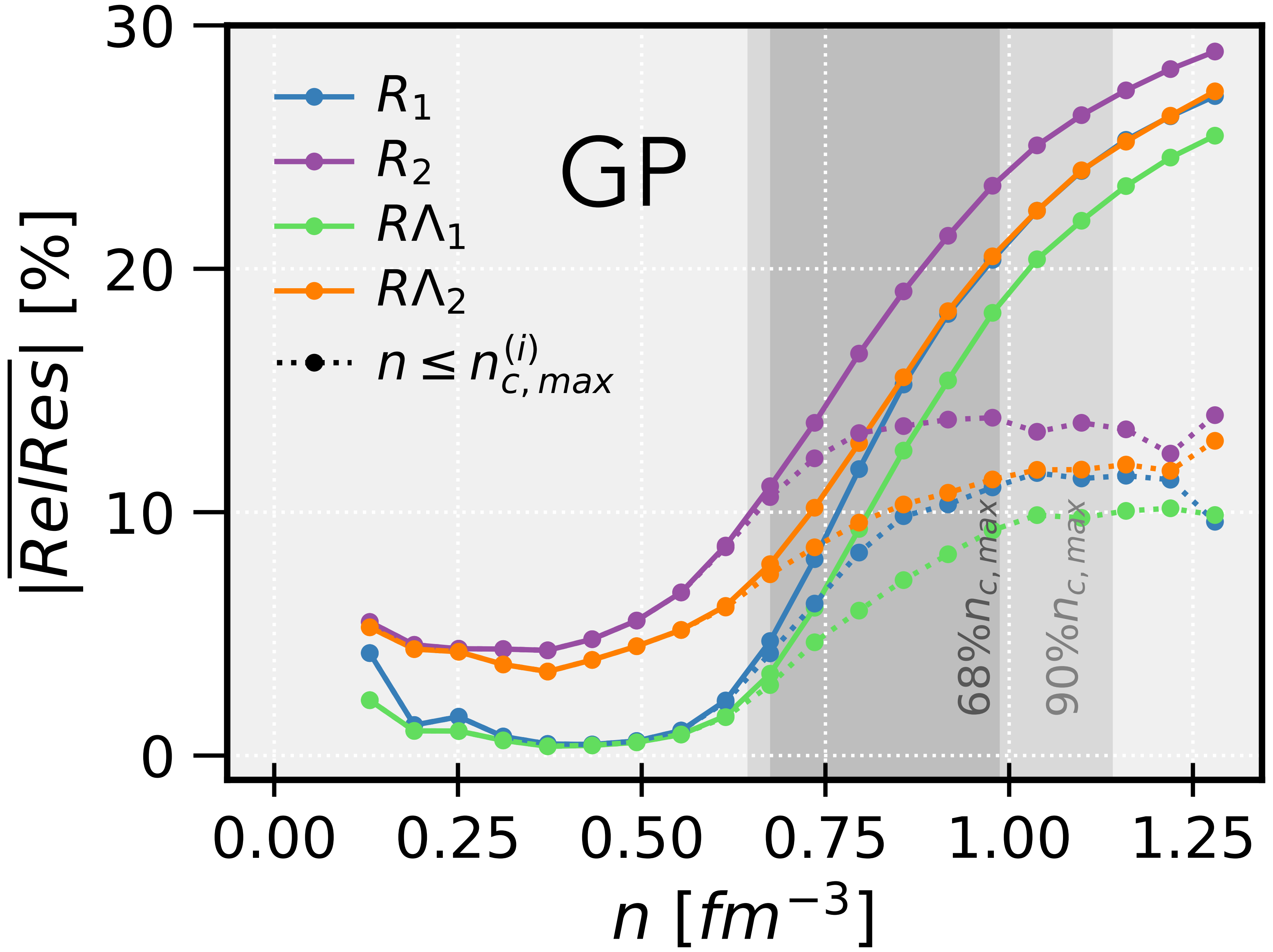
- Prediction for 2 samples of the test set $R\Lambda_2$ —, with a 90 % CI and median - - -,
- Increase in dispersion near maximum central density, represented by ●,
- Predictions always inside the CI.



Just for GP dataset

Relative error for pressure





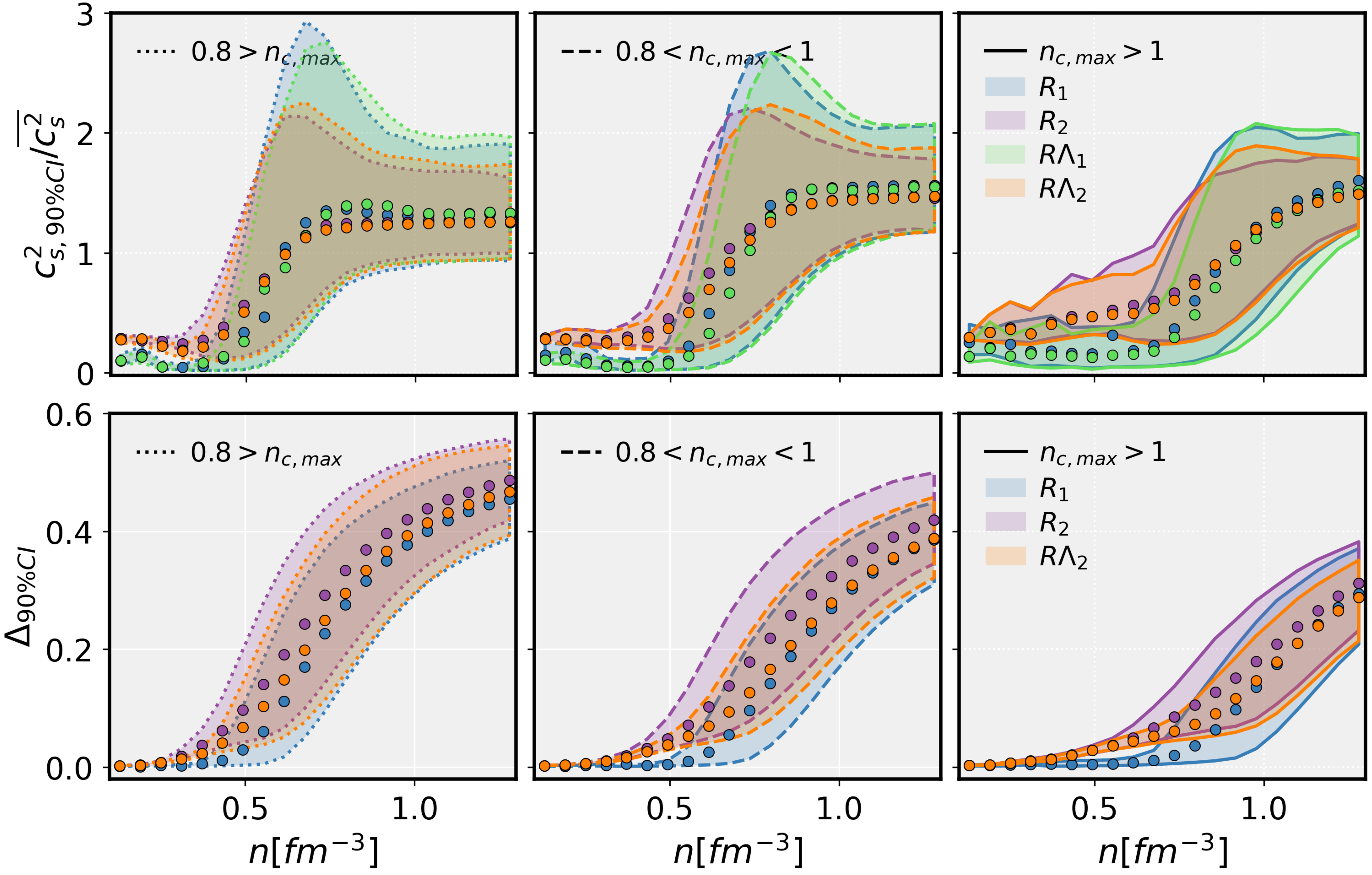
$$\text{RelRes}^{(i)}(n) = \text{Med}_l \left[\frac{X_p^{(i,l)}(n) - X_T^{(i)}(n)}{X_T^{(i)}(n)} \right] \times 100$$

$$\begin{cases} |\overline{\text{RelRes}}|_{\le n_{c,max}} = \text{Med}_i \left| \text{RelRes}^{(i)}(n) \right| \quad \text{for} \quad n \le n^{(i)}_{c,max}\,. \\ |\overline{\text{RelRes}}| = \text{Med}_i \left| \text{RelRes}^{(i)}(n) \right|\,. \end{cases}$$

l = Posterior sample, i = EoS at density n

- Error increases with noise,
- Error decreases with tidal deformability,
- Error decreases when filtered for $n_{c,max}$.

Effect of maximum central density

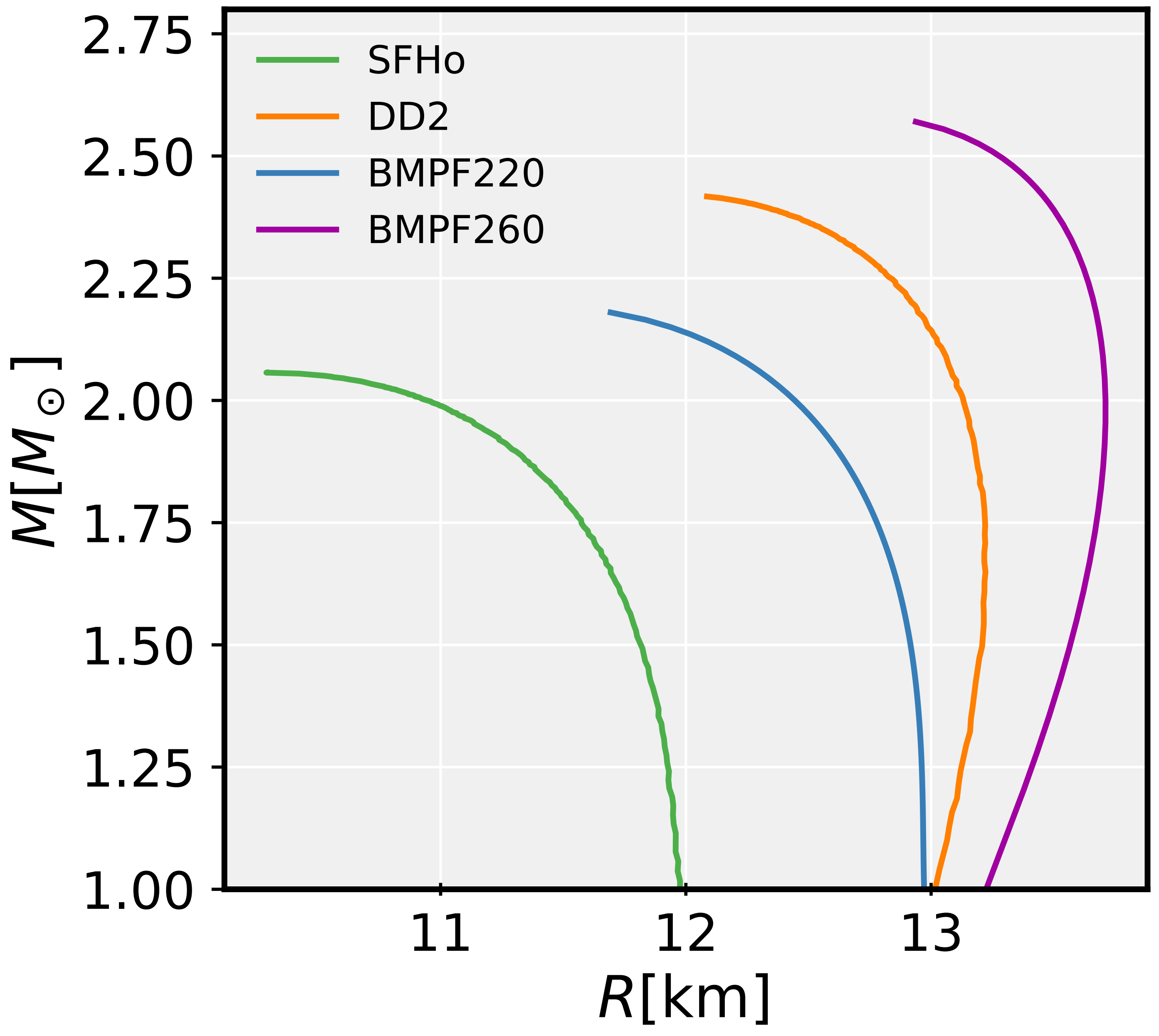


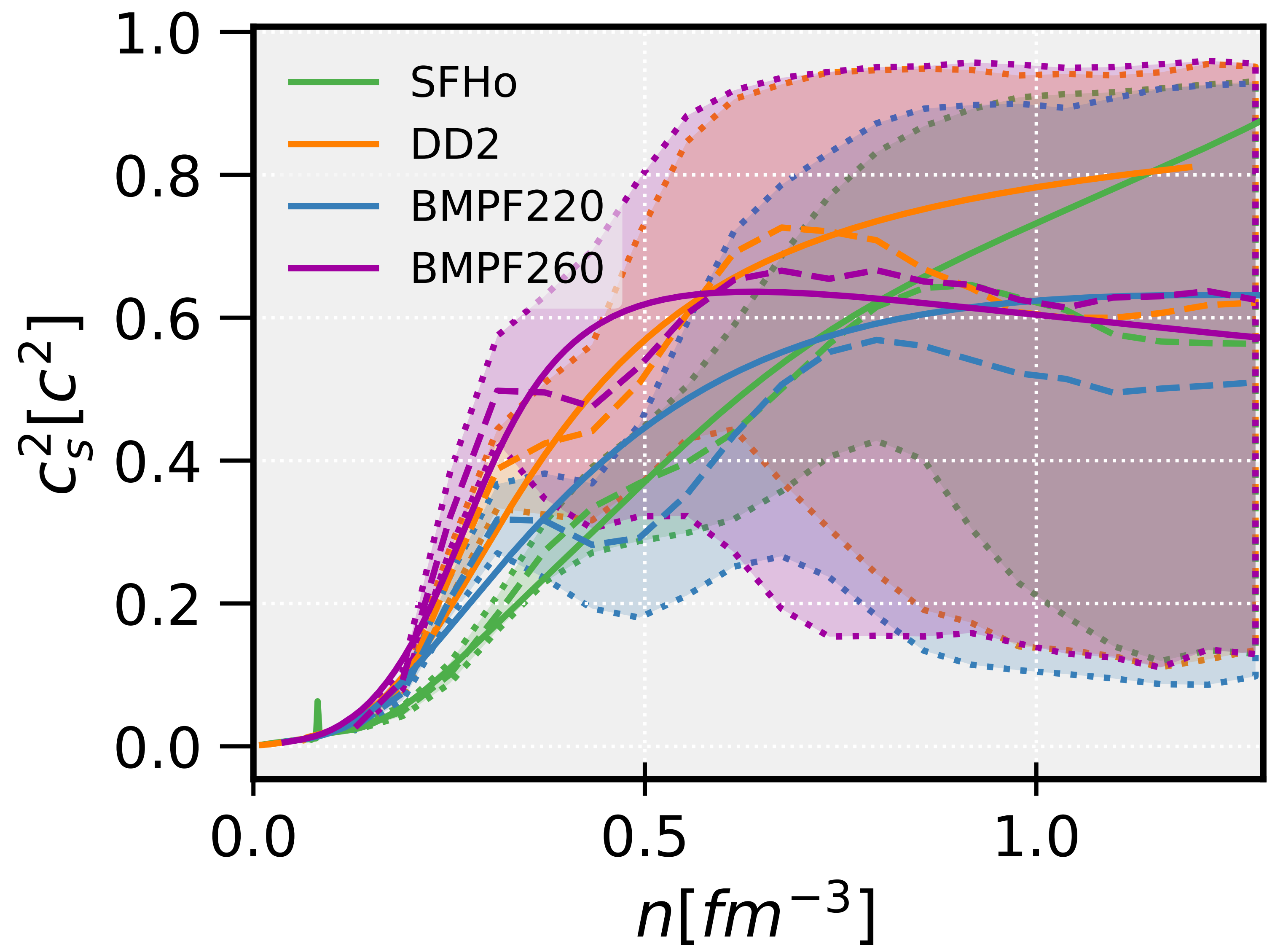
- Bands are the 90% CI,
- Dots are the mean.

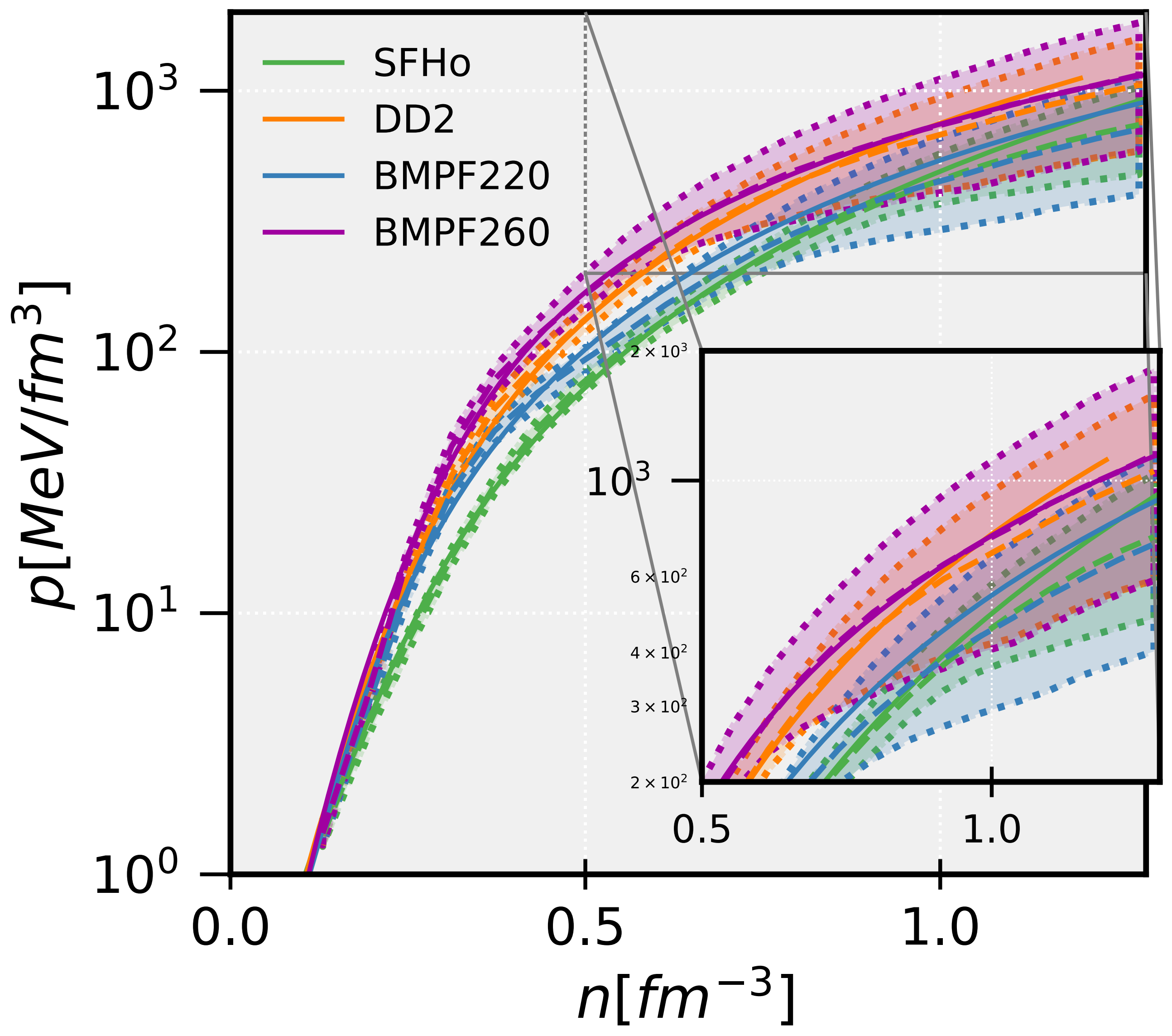
- Predicted dispersion (CI) decrease with the increase of $n_{c,max}$.

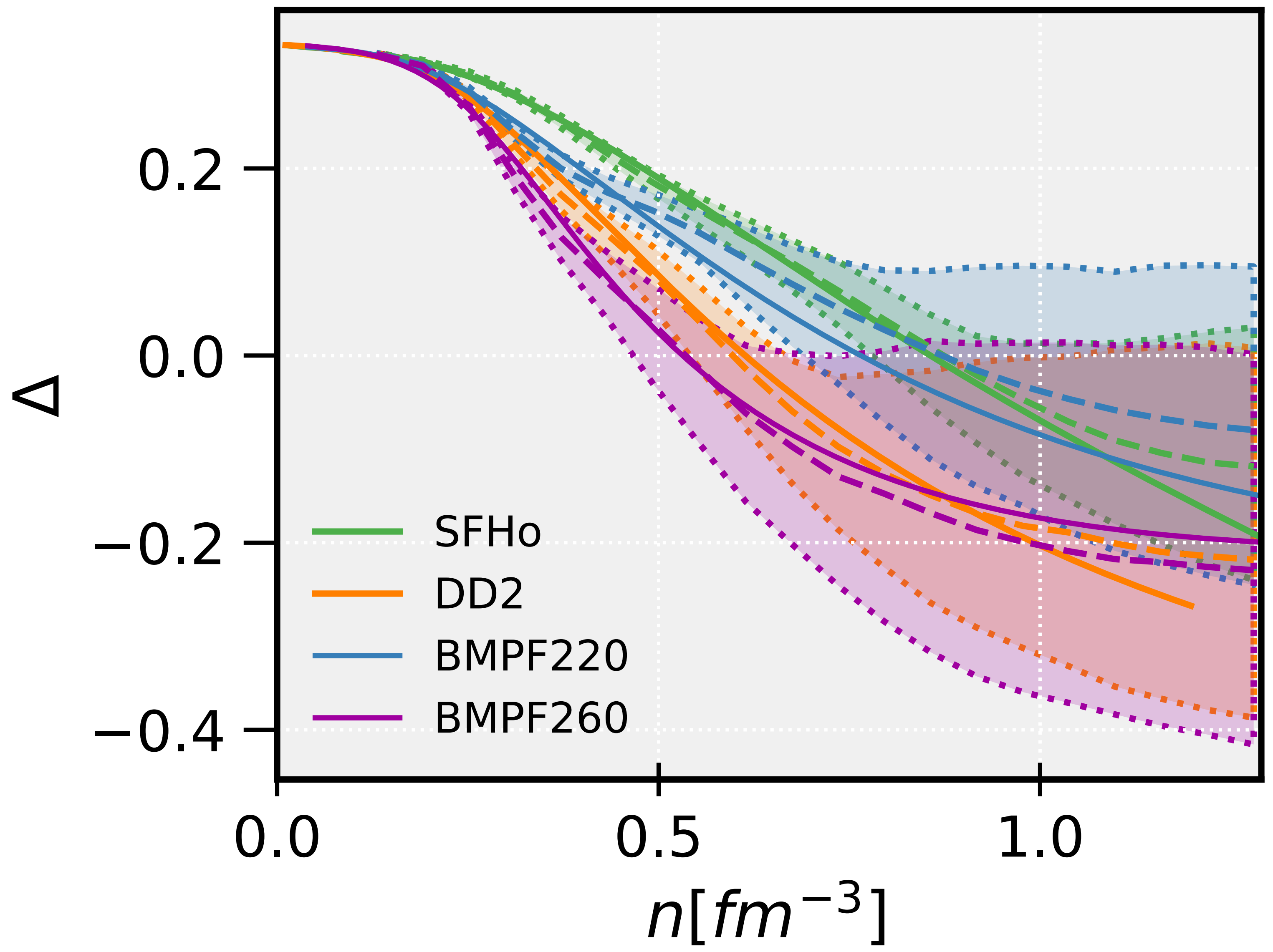
Another dataset inference test

- For the R_2 dataset we tested 4 EoS,
- Very different models.







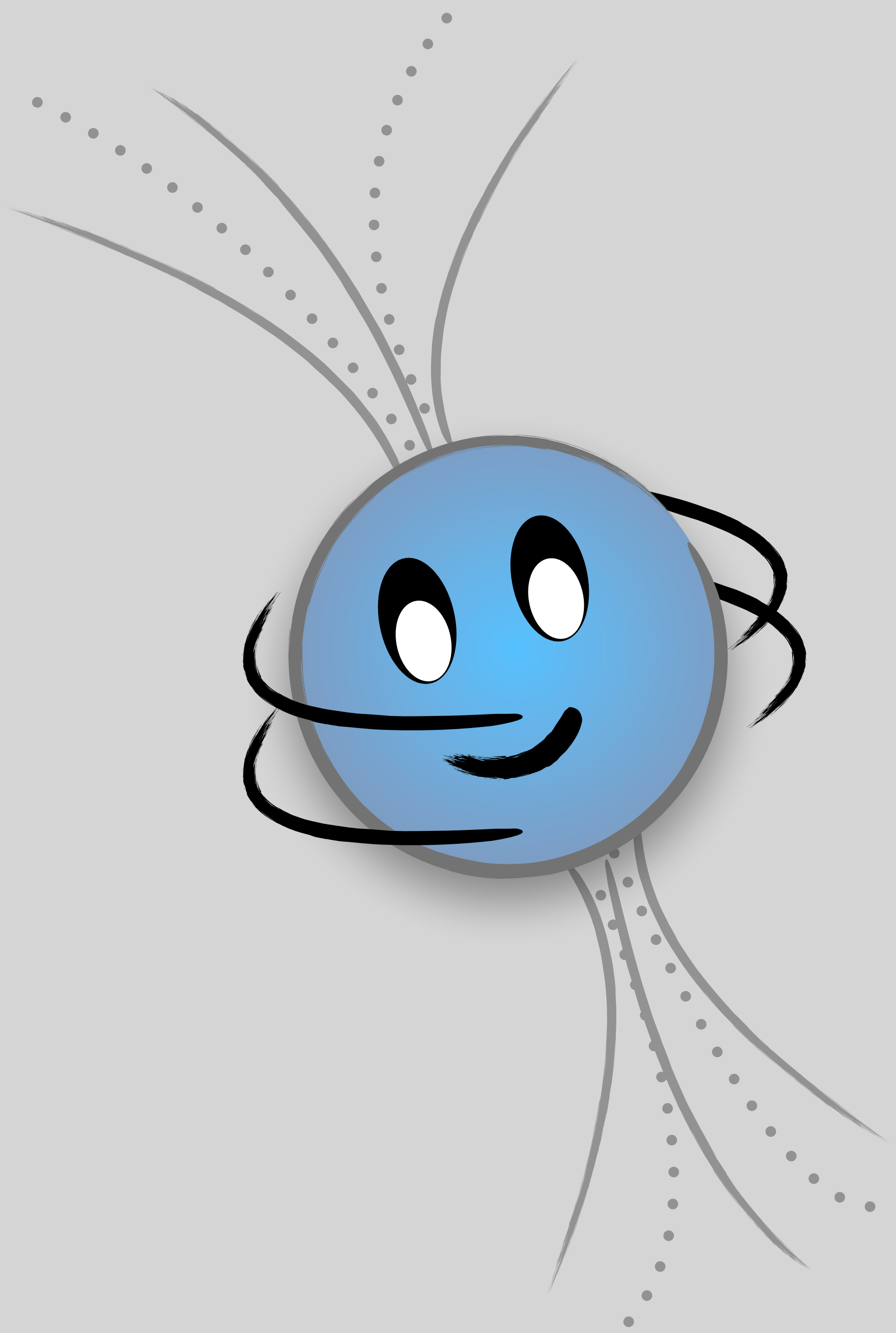


Just for GP dataset

Summary

- **Validation of neural posterior estimation:** Demonstrated that CNF can successfully recover neutron star EoS,
- **Crucial role of tidal deformability:** Including Λ alongside mass–radius data improves predictions,
- **Sensitivity to maximum central density:** The model naturally learns correlations between predictive uncertainty and the $n_{c,max}$,
- **Uncertainty quantification:** NPE provides well-calibrated posteriors.

The outlook: This method is well-suited for the upcoming "golden age" of multimessenger data, offering a promising tool to constrain dense matter physics.



Thank you for your attention

More information available at:

val.mar.dinis@gmail.com