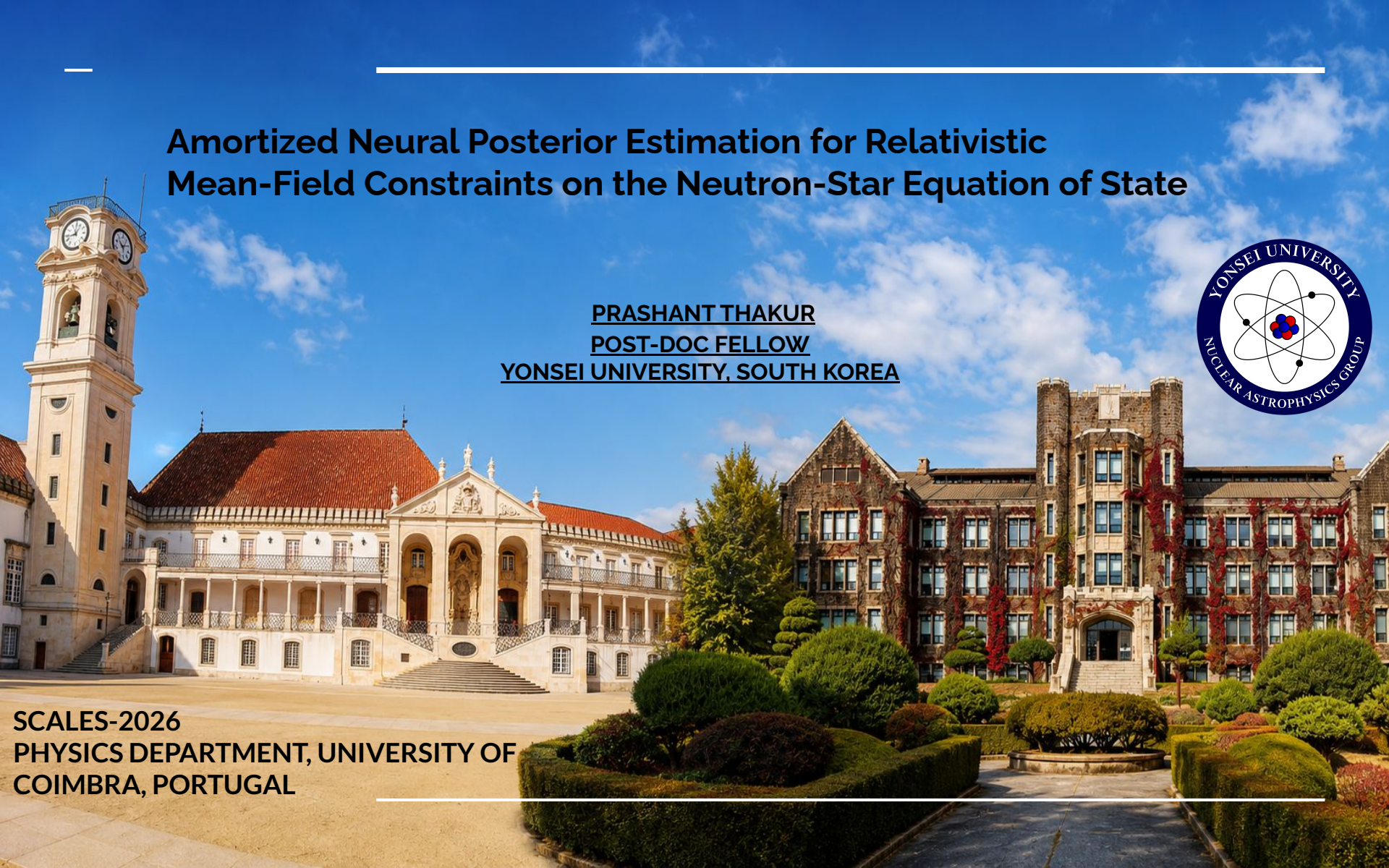
Amortized Neural Posterior Estimation for Relativistic Mean-Field Constraints on the Neutron-Star Equation of State

PRASHANT THAKUR
POST-DOC FELLOW
YONSEI UNIVERSITY, SOUTH KOREA



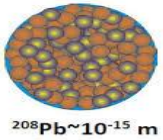
SCALES-2026
PHYSICS DEPARTMENT, UNIVERSITY OF COIMBRA, PORTUGAL

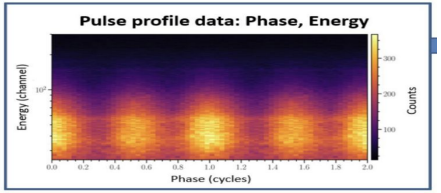
Constraints on dense-matter EOS

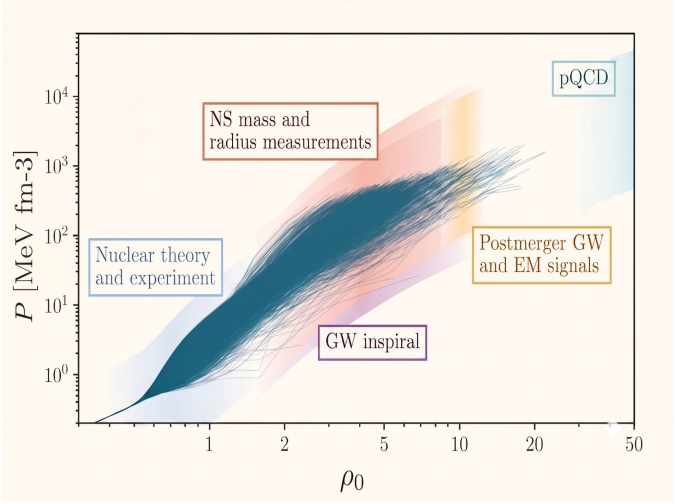
General physical constraints
- Causality: sound speed cannot exceed light speed
- Stability: ε > 0 and ε increases monotonically with pressure P

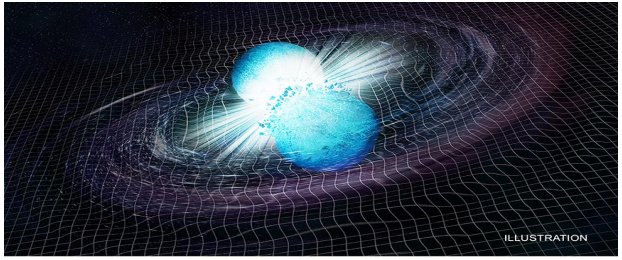
Laboratory constraints
- Nuclear binding energies
- Neutron-skin thickness
- Dipole polarizabilities
- Giant dipole resonances

Astrophysical constraints
- Massive pulsars
- NICER mass–radius measurements
- LIGO/Virgo tidal deformabilities
- Spin limits and moment of inertia estimates









Small footer:
This slide summarizes the broader constraint landscape; not all listed inputs are used in the present analysis.

Image credit: Google

Bayesian Inference Architecture

Systematic integration of prior knowledge, model likelihood, and observational data.

INPUTS & REQUIREMENTS

Model Likelihood

The statistical function mapping theoretical parameters to the observable data space.

Observational Data

The empirical evidence or measurements used to constrain the theoretical model.

Prior Constraints

The predefined parameter space and probability distributions before data is applied.

THE INFERENCE ENGINE

POSTERIOR DISTRIBUTION

$$P(\theta \mid x) = \frac{P(x \mid \theta)\,P(\theta)}{Z}$$

Posterior · Likelihood · Prior · Bayes evidence / Normalization

BAYESIAN EVIDENCE (MARGINAL LIKELIHOOD)

$$Z = \int P(x \mid \theta)\,P(\theta)\,d\theta$$

Evidence / marginal likelihood · Likelihood · Prior · Integrate over parameters θ

COMPUTATIONAL SAMPLERS

MCMC (Metropolis/Gibbs)

Generates samples via a sequential random walk; can struggle to traverse multiple isolated probability peaks.

Pyro (NUTS / HMC)

Leverages gradient information to efficiently explore high-dimensional spaces without random walk behavior.

FlowMC

Combines Normalizing Flows with JAX to provide accelerated, global proposal sampling on hardware.

Nested Sampling

Iteratively shrinks the prior volume along iso-likelihood contours to calculate posteriors and evidence.

The Nested Sampling Edge

Evidence Calculation

Directly integrates the Bayesian Evidence (Z), which is strictly required for rigorous Bayes Factor model comparison.

Multi-Modality

Inherently robust at discovering and sampling from multiple disconnected posterior peaks that standard walkers miss.

Prior Shrinkage

Operates by compressing from broad prior boundaries inward, making it highly effective for complex, constrained volumes.

Simulation-Based Inference: From ABC to DINGO

ABC

Approximate Bayesian Computation

Likelihood-free beginnings

Neural SBI

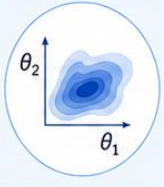
Neural density / ratio estimation

Backbone: normalizing or autoregressive flows; classifiers

Amortized inference

Train once -> rapid inference for new data

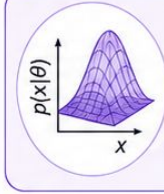
NPE



Learns the posterior directly

Data in -> posterior samples out

NLE



Learns the likelihood

Neural likelihood for standard Bayesian sampling

NRE

Learns the likelihood-to-evidence ratio

Inference through classification

Already well established

Used in cosmology, particle physics, neuroscience, and biological modelling

Already in gravitational-wave physics

- **DINGO** uses amortized NPE with normalizing flows
- **DINGO-IS** adds importance-sampling correction for rapid and reliable parameter estimation

SBI is no longer only promising — it already delivers fast, validated gravitational-wave inference through DINGO.

ABC: Beaumont et al. 2002 | SBI overview: Cranmer et al. 2020; Papamakarios et al. 2021 | DINGO / DINGO-IS: Dax et al. 2023

Astrophysics > High Energy Astrophysical Phenomena

[Submitted on 24 Jun 2026]

Amortized Simulation-Based Inference of Relativistic Mean-Field Couplings for Neutron-Star Equations of State

Prashant Thakur, Tuhin Malik

We present a simulation-based inference framework for constraining microscopic relativistic mean-field parameters of neutron-star equations of state. Neural posterior estimation is applied to two representative RMF families, a density-dependent DDB model and a nonlinear RMF-NL model, using nuclear saturation properties, chiral effective-field-theory pure-neutron-matter pressures, and the maximum-mass constraint as conditioning observables. The inferred posteriors are validated against the conventional nested sampler (PyMultiNest) calculations and tested with the TARP coverage diagnostic. For both RMF parametrizations, the neural posterior reproduces the nested-sampling constraints on model couplings, nuclear-matter properties, and neutron-star observables with no significant bias. The amortized estimator generates 3×10^4 posterior samples in about 2.5 s on a CPU, enabling a rapid inference workflow without the need for retraining for updated data. This constitutes a proof of concept that NPE-emulated RMF models, once validated, can be safely used for superfast exploratory inference. As an additional mock-observation test, imposing $R_{1.4} = 12\,\mathrm{km}$ and $M_{\max} > 1.97\,M_\odot$ leads to consistent predictions for the maximum-mass configuration, with DDB giving $M_{\max} = 2.10^{+0.09}_{-0.07}\,M_\odot$, $R_{\max} = 10.71^{+0.14}_{-0.21}$ km and RMF-NL giving $M_{\max} = 2.05^{+0.10}_{-0.06}\,M_\odot$, $R_{\max} = 10.69^{+0.18}_{-0.19}$ km; although fixing $R_{1.4}$ confines both families to a narrow EOS region, RMF-NL remains marginally softer than DDB at high density, consistent with its slightly lower maximum mass.

Comments: 7 figures, 3 Tables

Subjects: **High Energy Astrophysical Phenomena (astro-ph.HE)**; Nuclear Theory (nucl-th)

Cite as: arXiv:2606.25446 **[astro-ph.HE]**

(or arXiv:2606.25446v1 **[astro-ph.HE]** for this version)

https://doi.org/10.48550/arXiv.2606.25446

Submission history

From: Prashant Thakur [view email]

[v1] Wed, 24 Jun 2026 06:15:28 UTC (1,124 KB)

EOS MODELS

$$\begin{aligned}\mathcal{L} = \bar{\Psi}\Big[\gamma^{\mu}\Big(i\partial_{\mu} - \Gamma_{\omega}\omega_{\mu} - \Gamma_{\varrho}\,\boldsymbol{t}\cdot\boldsymbol{\varrho}_{\mu}\Big) \\ - \Big(m - \Gamma_{\sigma}\sigma\Big)\Big]\Psi + \frac{1}{2}\Big(\partial_{\mu}\sigma\,\partial^{\mu}\sigma - m_{\sigma}^{2}\sigma^{2}\Big) \\ - \frac{1}{4}F_{\mu\nu}^{(\omega)}F^{(\omega)\mu\nu} + \frac{1}{2}m_{\omega}^{2}\omega_{\mu}\omega^{\mu} \\ - \frac{1}{4}\boldsymbol{F}_{\mu\nu}^{(\varrho)}\cdot\boldsymbol{F}^{(\varrho)\mu\nu} + \frac{1}{2}m_{\varrho}^{2}\,\boldsymbol{\varrho}_{\mu}\cdot\boldsymbol{\varrho}^{\mu} \\ + \mathcal{L}_{NL}.\end{aligned}$$

DDB:

$$h_i(x) = a_M\frac{1+b_i(x+d_i)^2}{1+c_i(x+d_i)^2},$$

RMF-NL:

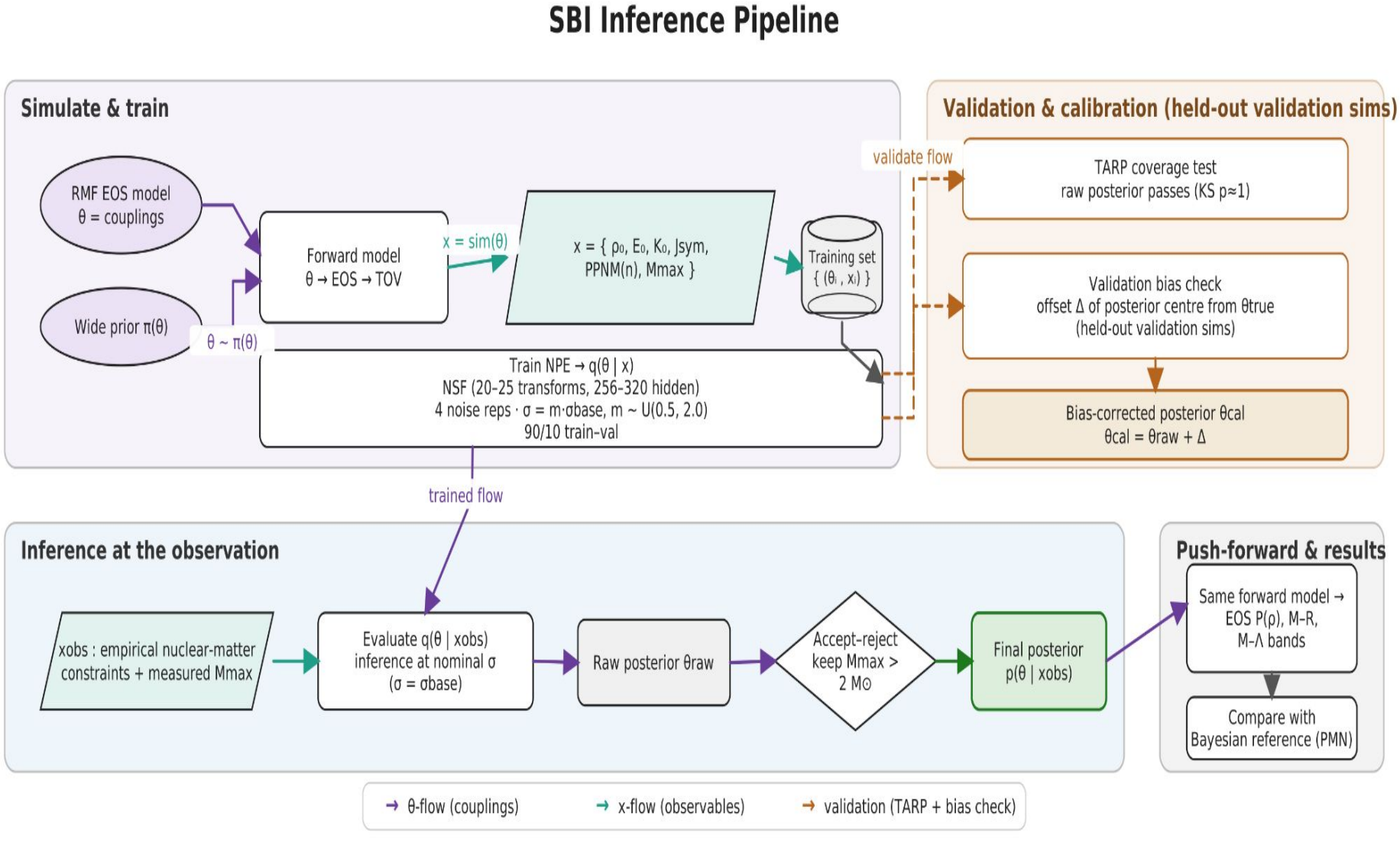
$$\begin{aligned}\mathcal{L}_{NL} = &-\frac{\kappa}{3!}(g_{\sigma}\sigma)^3 - \frac{\lambda}{4!}(g_{\sigma}\sigma)^4 \\ &+ \frac{\xi}{4!}\left(g_{\omega}^2\,\omega_{\mu}\omega^{\mu}\right)^2 + \Lambda_{\omega}\,g_{\varrho}^2\,\boldsymbol{\varrho}_{\mu}\cdot\boldsymbol{\varrho}^{\mu}\,g_{\omega}^2\,\omega_{\mu}\omega^{\mu}.\end{aligned}$$

Table 1: Prior and prior-predictive ranges used to construct the representative DDB and RMF-NL training banks.

DDB			**RMF-NL**		
Quantity	Min.	Max.	Quantity	Min.	Max.
Prior			*Prior*		
a_σ	0	0.300	Γ_σ	6.50	15.50
a_ω	0	0.298	Γ_ω	6.50	15.50
a_ρ	0	1.300	Γ_ρ	5.50	16.50
$\Gamma_{\sigma,0}$	6.50	13.49	κ	0.0048	0.086
$\Gamma_{\omega,0}$	7.50	14.50	λ_0	-0.05	0.05
$\Gamma_{\rho,0}$	5.00	12.55	ζ	0.0	0.04
			Λ_ω	0.0	0.15
Outputs			*Outputs*		
ρ_0 [fm^{-3}]	0.140	0.170	ρ_0 [fm^{-3}]	0.140	0.170
E_0 [MeV]	-100	100	E_0 [MeV]	-99.96	100.0
K_0 [MeV]	~ 0	800	K_0 [MeV]	~ 0	800.0
$J_{\rm sym}$ [MeV]	19.2	50.0	$J_{\rm sym}$ [MeV]	16.58	50.0
$P_{\rm PNM}(0.08)$	~ 0	7.60	$P_{\rm PNM}(0.08)$	0.0	7.71
$P_{\rm PNM}(0.12)$	~ 0	18.45	$P_{\rm PNM}(0.12)$	0.41	17.95
$P_{\rm PNM}(0.16)$	~ 0	34.67	$P_{\rm PNM}(0.16)$	1.25	33.19
$M_{\rm max}$ [$M_\odot$]	1.40	3.56	$M_{\rm max}$ [$M_\odot$]	1.40	3.49

Table 2: Nuclear-matter and astrophysical inputs used in this study. For each quantity written as $\mu \pm \sigma_{\rm base}$, μ denotes the central value used in the fiducial observation vector $\mathbf{x}_{\rm obs}$, while $\sigma_{\rm base}$ denotes the corresponding 1σ uncertainty. During NPE training, noise-augmented replicas are generated using scaled uncertainties $\boldsymbol{\sigma} = m\,\boldsymbol{\sigma}_{\rm base}$, with $m \sim \mathcal{U}(0.5, 2.0)$. The quantities E_0, K_0, and $J_{\rm sym}$ are given in MeV, $P_{\rm PNM}$ in MeV fm^{-3}, ρ_0 in fm^{-3}, and $M_{\rm max}$ in $M_\odot$. The PNM pressure uncertainties correspond to $2 \times$ N^{3}LO chiral-EFT errors.

Quantity	Input	Source / use
E_0	-16.1 ± 0.2	Dutra et al. (2014) [35]
K_0	230 ± 40	Todd-Rutel and Piekarewicz (2005); Shlomo et al. (2006) [36]
$J_{\rm sym}$	32.5 ± 1.8	Essick et al. (2021) [37]
$P_{\rm PNM}(0.08)$	0.505714 ± 0.194286	Hebeler et al. (2013) [10]
$P_{\rm PNM}(0.12)$	1.241429 ± 0.608571	
$P_{\rm PNM}(0.16)$	2.485714 ± 1.382857	
ρ_0	0.153 ± 0.005	Typel and Wolter (1999) [38]
$M_{\rm max}$	> 2.0	Fonseca et al. (2016) [39]


SBI Inference Pipeline
Simulate & train
RMF EOS model
θ = couplings
Wide prior π(θ)
θ ~ π(θ)
Forward model
θ → EOS → TOV
x = sim(θ)
x = { ρ₀, E₀, K₀, Jsym,
PPNM(n), Mmax }
Training set
{ (θᵢ , xᵢ) }
Train NPE → q(θ | x)
NSF (20–25 transforms, 256–320 hidden)
4 noise reps · σ = m·σbase, m ~ U(0.5, 2.0)
90/10 train–val
validate flow
Validation & calibration (held-out validation sims)
TARP coverage test
raw posterior passes (KS p≈1)
Validation bias check
offset Δ of posterior centre from θtrue
(held-out validation sims)
Bias-corrected posterior θcal
θcal = θraw + Δ
trained flow
Inference at the observation
xobs : empirical nuclear-matter
constraints + measured Mmax
Evaluate q(θ | xobs)
inference at nominal σ
(σ = σbase)
Raw posterior θraw
Accept-reject
keep Mmax >
2 M⊙
Final posterior
p(θ | xobs)
Push-forward & results
Same forward model →
EOS P(ρ), M–R,
M–Λ bands
Compare with
Bayesian reference (PMN)
θ-flow (couplings)
x-flow (observables)
validation (TARP + bias check)

Neural Posterior Estimation to infer RMF model couplings from simulated observables

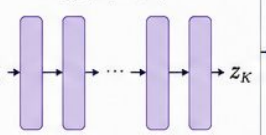
ARCHITECTURE (NPE with Normalizing Flow)

Inputs (conditioning)
- Observable vector $\boldsymbol{x}$
- Noise scale $\boldsymbol{\sigma}$

(see right panel)

Neural Posterior Estimator

$q_\phi(\boldsymbol{\theta} \mid \boldsymbol{x}, \boldsymbol{\sigma})$



Neural Spline Flow (NSF)

Output
Posterior samples
$p(\boldsymbol{\theta} \mid \boldsymbol{x}, \boldsymbol{\sigma})$

- **Normalizing flow:** Neural Spline Flow (NSF)
- **Coupling transforms:** 20 – 25
- **Hidden features:** 256 – 320
- **Parameter vector:** 6 couplings (DDB) / 7 couplings (RMF-NL)
- **Trained density-estimator state:** 554 tensors

TRAINING SETUP

Objective (Maximum Likelihood)

Minimize negative log-probability:

$$\mathcal{L}(\phi) = -\sum_{i}^{N_{\text{sim}}} \log q_\phi(\boldsymbol{\theta}_i \mid \boldsymbol{x}_i, \boldsymbol{\sigma}_i)$$

Training Hyperparameters

- Batch size: 2048
- Learning rate: 3×10^{-4}
- Max epochs: 500
- Validation every: 10 epochs
- Early stopping patience: 120 epochs
- Optimizer: Adam (default in sbi)

Model Selection (Validation Score)

$$S = \text{NLPD} + 5\,|C_{68} - 0.68| + 5\,|C_{90} - 0.90| + 0.2\,\text{MAE}$$

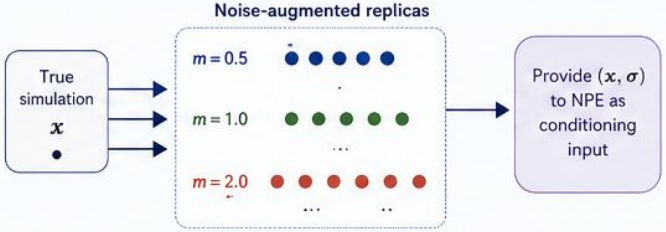
- C_{68}, C_{90} : empirical 68% and 90% coverage
- NLPD : negative log-posterior density
- MAE : mean absolute error of posterior location

NOISE CONDITIONING

For each simulated observable vector $\boldsymbol{x}$, create noise-augmented replicas with a random scale m :

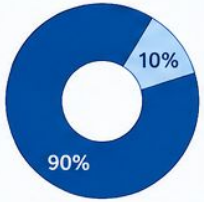
$$\boldsymbol{\sigma} = m\,\boldsymbol{\sigma}_{\text{base}}, \qquad m \sim \mathcal{U}(0.5, 2.0)$$

$\boldsymbol{\sigma}_{\text{base}}$: fiducial observational uncertainty vector (Table 2 in paper)



- Exposes the flow to a continuous range of uncertainties.
- Learned posteriors remain robust and well-calibrated across different observational precisions.
- Inference is amortized over both $\boldsymbol{x}$ and $\boldsymbol{\sigma}$ — one trained network works for the full noise range.

TRAIN–VALIDATION SPLIT

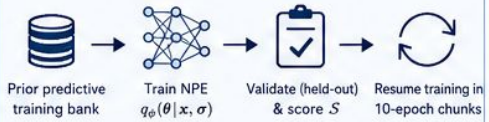


Training set (90%)
Used for NPE training

Validation set (10%)
Held-out for evaluation and calibration only

Statistical calibration (e.g., TARP diagnostic) performed only on the held-out validation set.

Training Procedure



Simulation pairs $(\boldsymbol{\theta}_i, \boldsymbol{x}_i, \boldsymbol{\sigma}_i)$ are used for training.

HARDWARE (GPU) & RUNTIME

NVIDIA

GPU Specification

- **GPU:** NVIDIA GeForce RTX 3060 Ti
- **Memory:** 8 GB GDDR6
- **Compute Capability:** 8.6 (Ampere)
- **CUDA:** 12.4 | **cuDNN:** 9.2.1
- **NVIDIA Driver:** 580.159.03

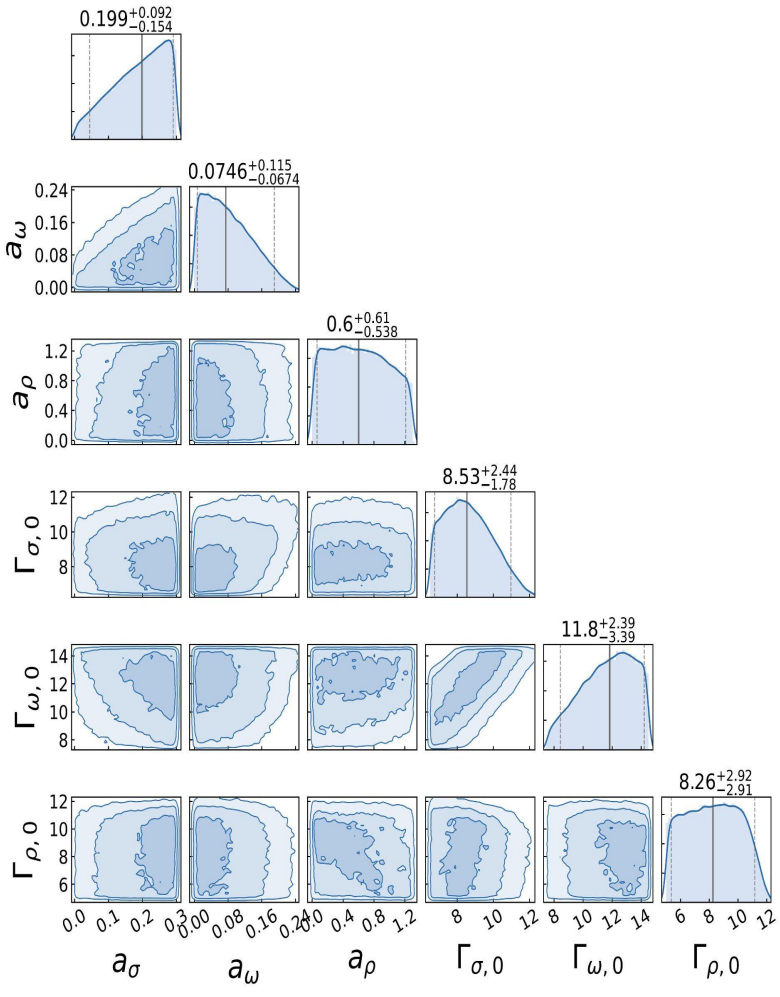
Runtime Performance

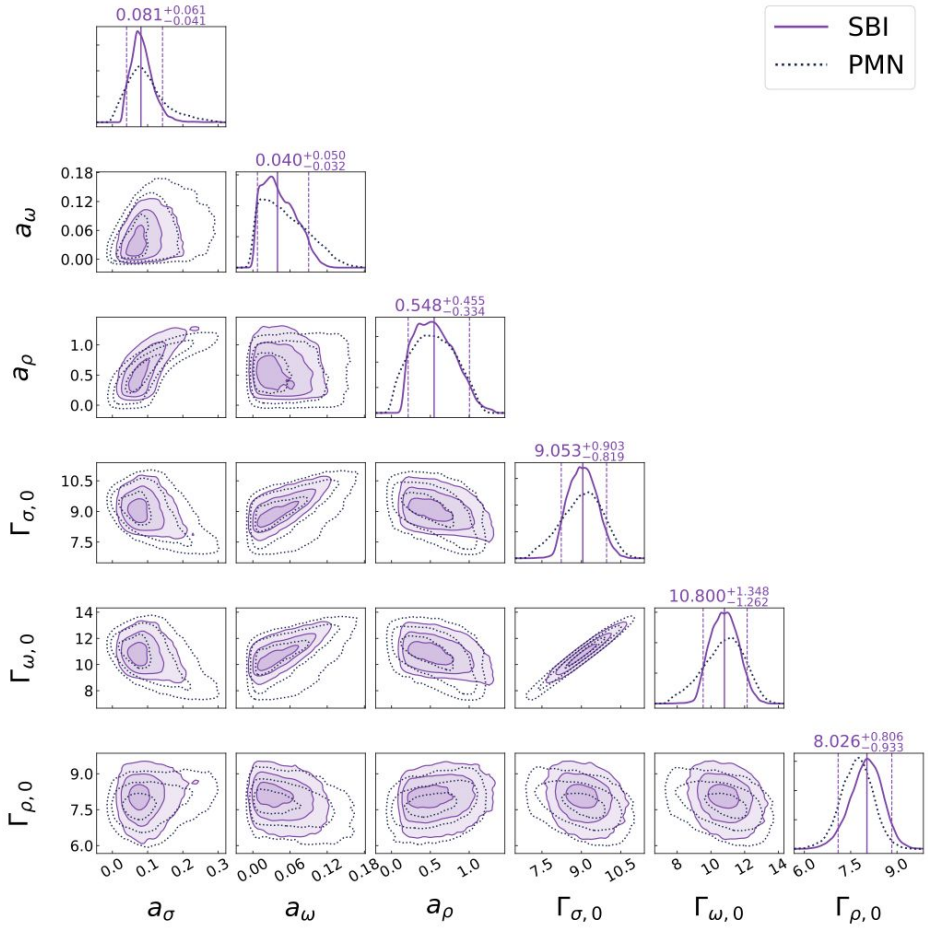
One-time training cost (500 epochs): $\sim$ **2.2 hours** ($\sim$ 15.8 s per epoch)

Posterior sampling speed 4×10^5 NPE samples in $\sim$ **5.6 s** (GPU)

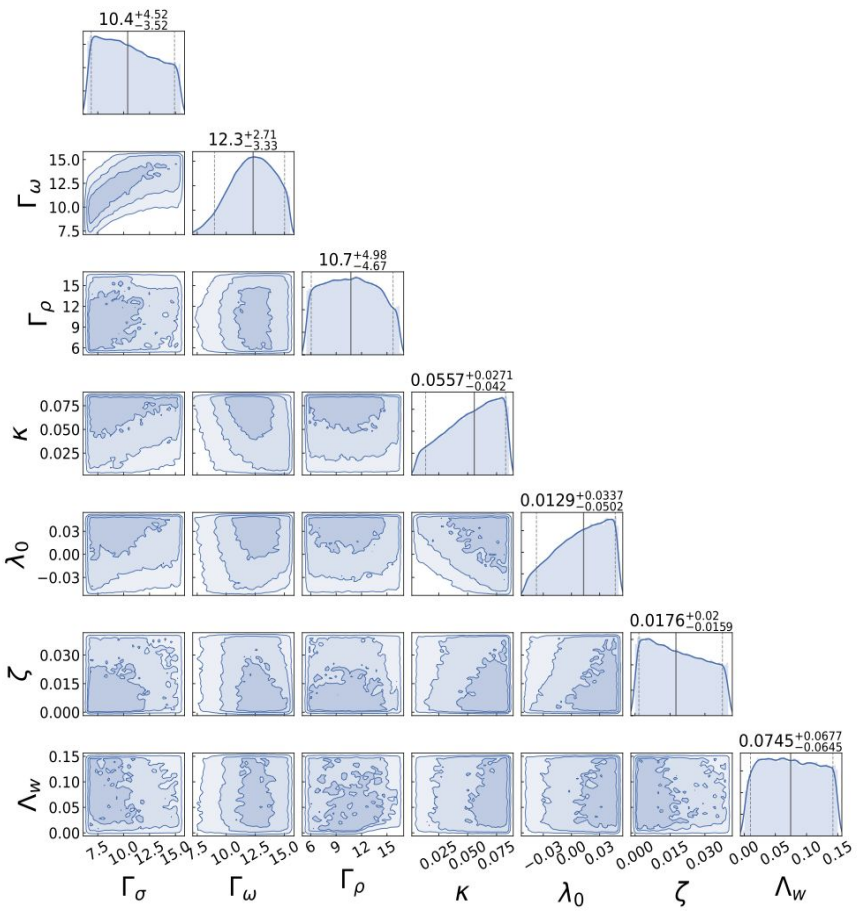
Once trained, the NPE provides calibrated posterior samples in seconds for any new input or uncertainty model within the trained noise range— without retraining.

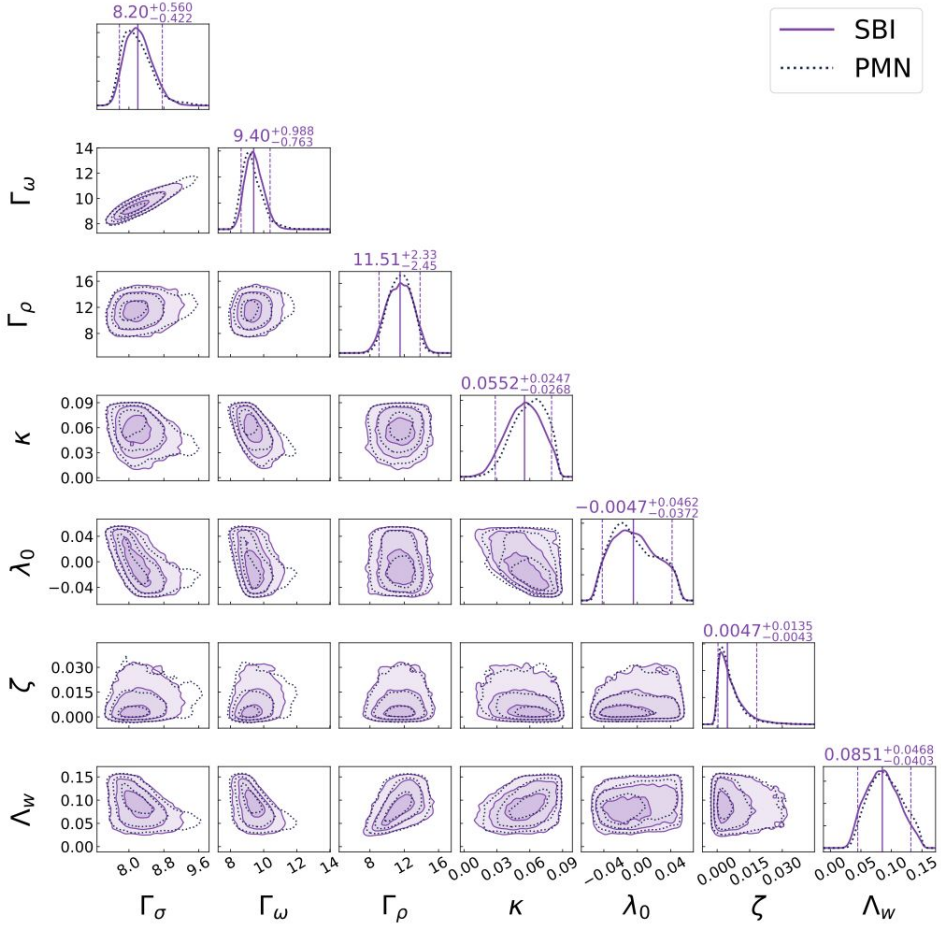
DDB PRIORS & POSTERIORS



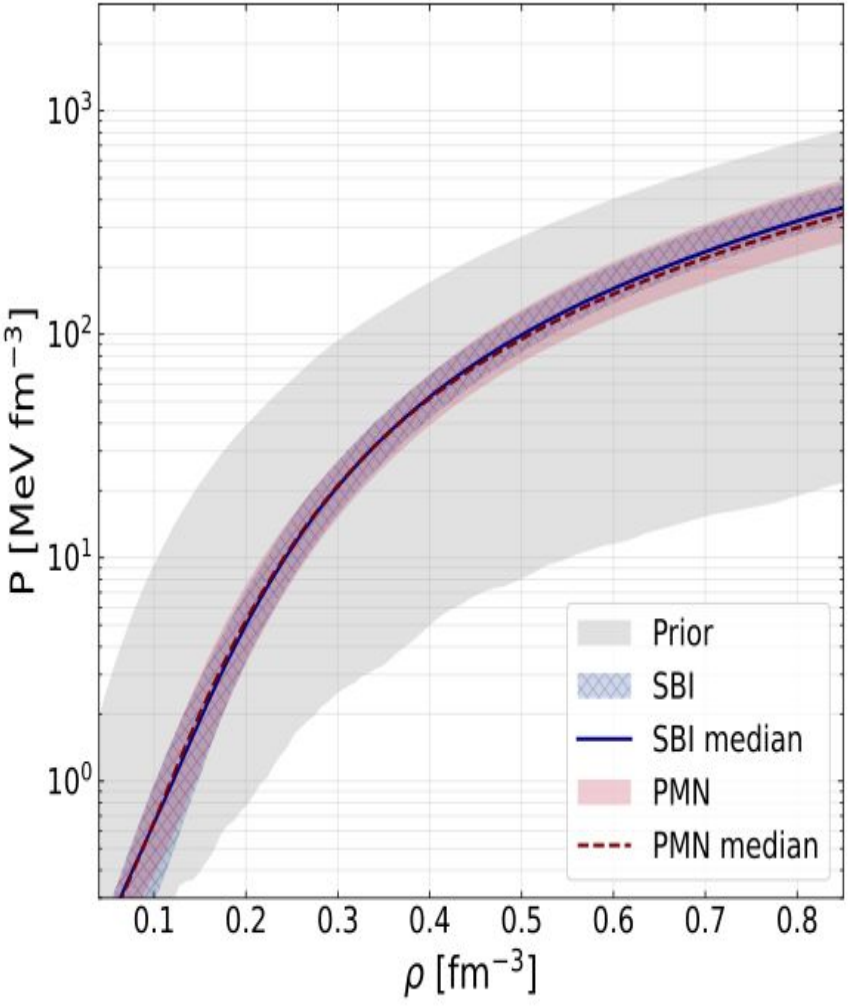


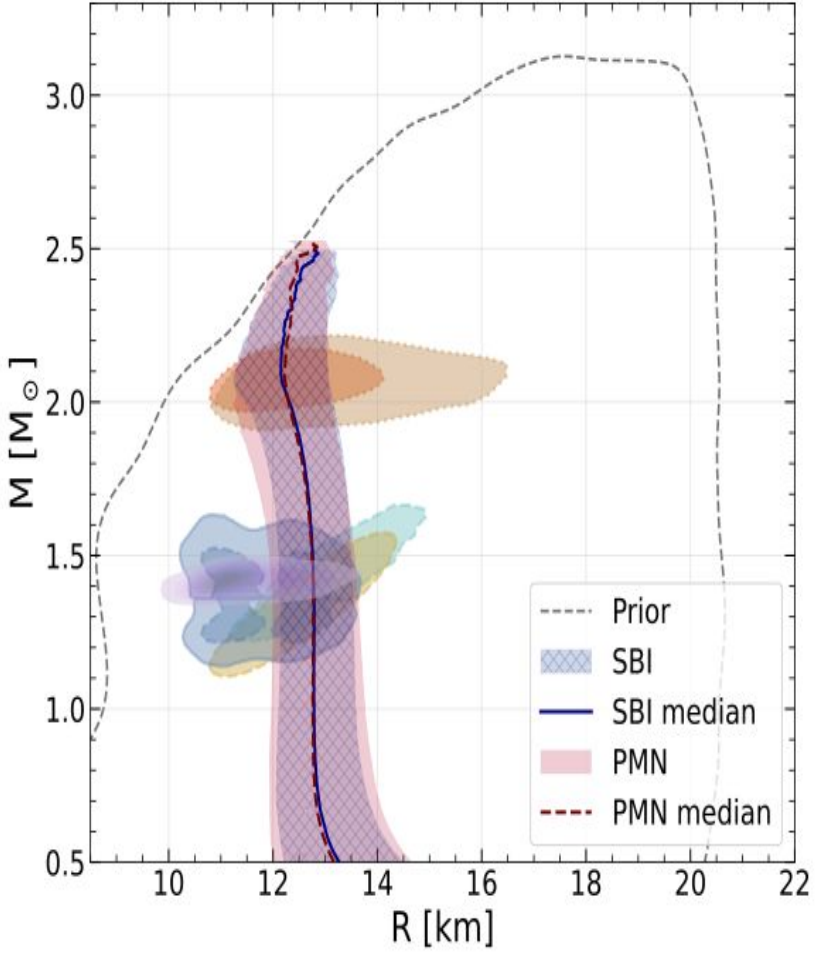
RMF-NL PRIORS & POSTERIORS



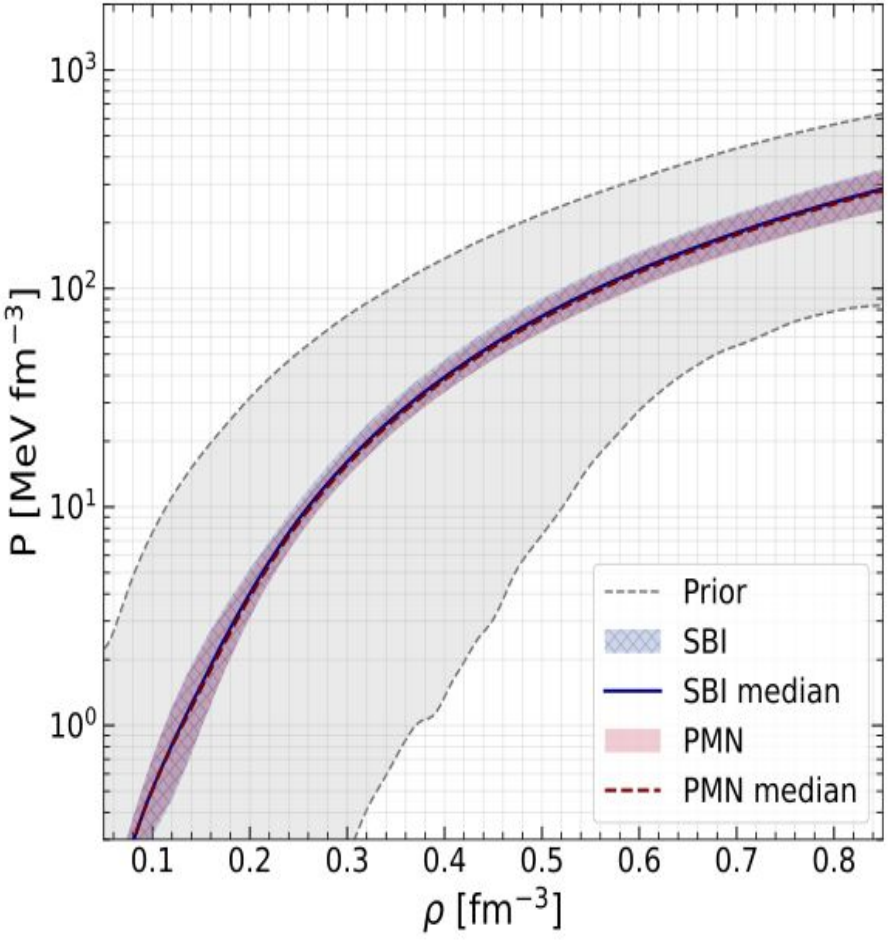


DDB: SBI–PyMultiNest Posterior Consistency





RMF-NL: SBI–PyMultiNest Posterior Consistency



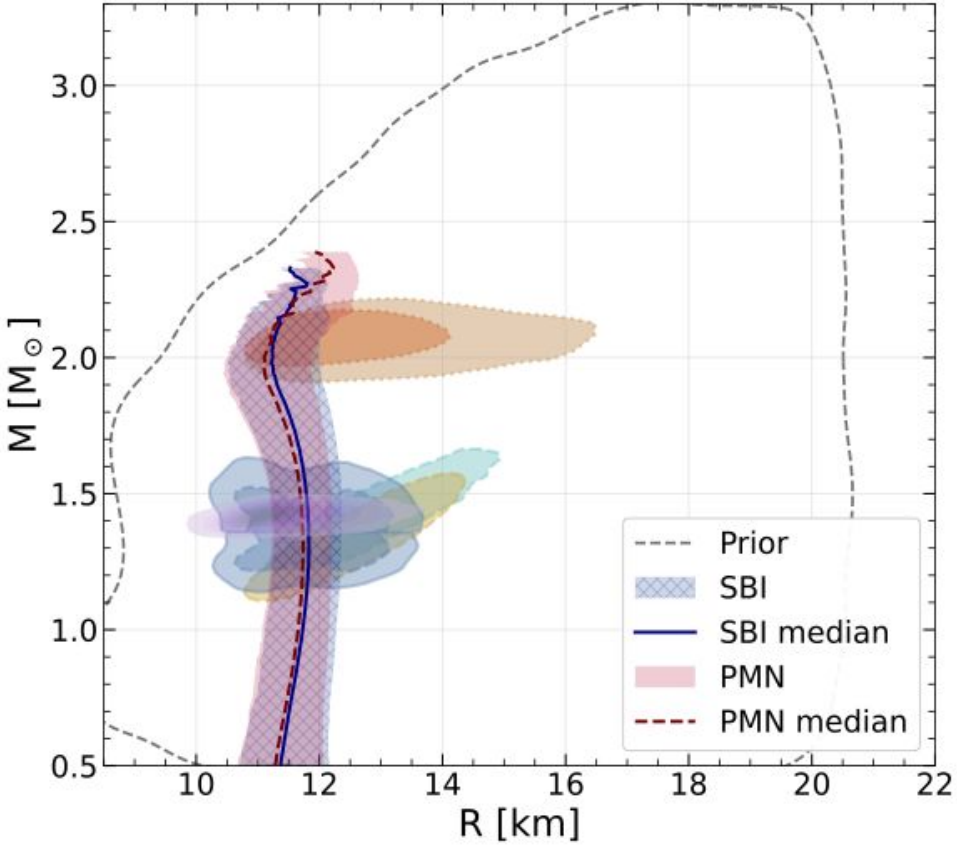


Table 3: Combined comparison of the density-dependent RMF (DDB)and RMF-NL results obtained from SBI and PyMultiNest. Values denote posterior medians with 90% credible intervals.

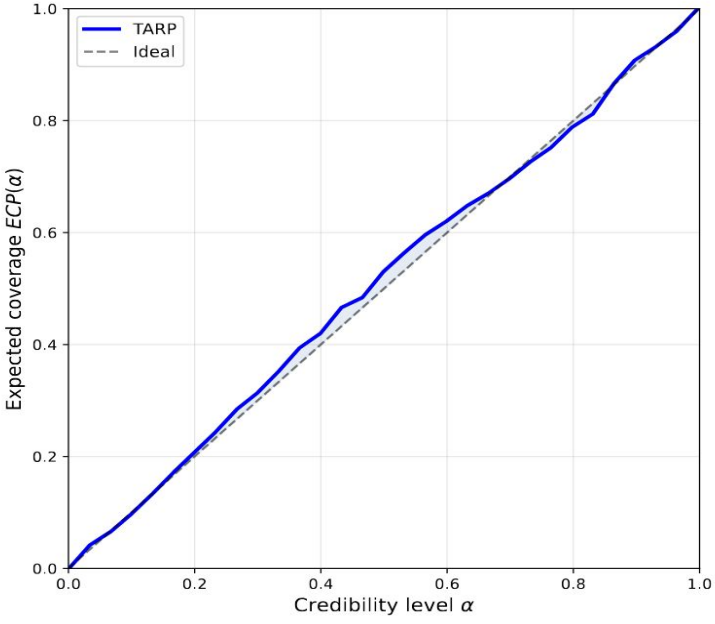
DDB			*RMF-NL*		
Quantity	SBI	PMN	Quantity	SBI	PMN
a_σ	$0.081^{+0.061}_{-0.041}$	$0.087^{+0.121}_{-0.065}$	Γ_σ	$8.20^{+0.560}_{-0.422}$	$8.15^{+0.678}_{-0.407}$
a_ω	$0.040^{+0.050}_{-0.032}$	$0.047^{+0.073}_{-0.042}$	Γ_ω	$9.40^{+0.988}_{-0.763}$	$9.22^{+1.277}_{-0.735}$
a_ρ	$0.548^{+0.455}_{-0.334}$	$0.534^{+0.469}_{-0.431}$	Γ_ρ	$11.51^{+2.33}_{-2.45}$	$11.53^{+2.10}_{-2.24}$
$\Gamma_{\sigma,0}$	$9.053^{+0.903}_{-0.819}$	$9.068^{+1.156}_{-1.431}$	κ	$0.0552^{+0.0247}_{-0.0268}$	$0.0607^{+0.0207}_{-0.0267}$
$\Gamma_{\omega,0}$	$10.800^{+1.348}_{-1.262}$	$10.835^{+1.733}_{-2.310}$	λ_0	$-0.0047^{+0.0462}_{-0.0372}$	$-0.0073^{+0.0485}_{-0.0338}$
$\Gamma_{\rho,0}$	$8.026^{+0.806}_{-0.933}$	$7.711^{+0.828}_{-0.890}$	ζ	$0.0047^{+0.0135}_{-0.0043}$	$0.0044^{+0.0127}_{-0.0040}$
–	–	–	Λ_ω	$0.0851^{+0.0468}_{-0.0403}$	$0.0896^{+0.0473}_{-0.0406}$

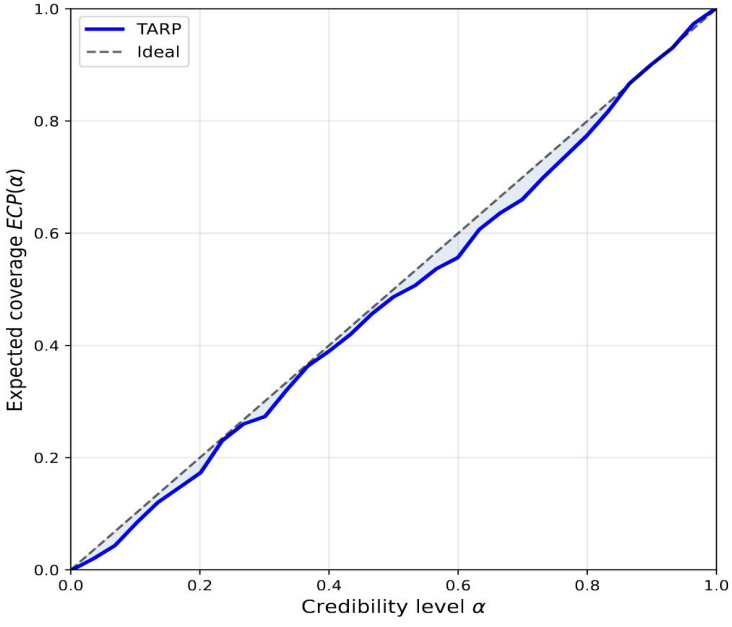
Quantity	*DDB*		*RMF-NL*	
	SBI	PMN	SBI	PMN
ρ_0 [fm^{-3}]	$0.1522^{+0.0080}_{-0.0071}$	$0.1528^{+0.0055}_{-0.0055}$	$0.1519^{+0.0080}_{-0.0083}$	$0.1524^{+0.0068}_{-0.0064}$
E_0 [MeV]	$-16.18^{+0.53}_{-0.50}$	$-16.10^{+0.32}_{-0.32}$	$-16.00^{+0.33}_{-0.33}$	$-16.10^{+0.30}_{-0.30}$
K_0 [MeV]	$243.9^{+53.6}_{-48.9}$	$239.8^{+52.7}_{-57.7}$	$251.1^{+45.6}_{-40.1}$	$247.9^{+37.1}_{-36.1}$
$J_{\rm sym}$ [MeV]	$33.45^{+2.91}_{-3.22}$	$32.30^{+2.79}_{-2.84}$	$32.02^{+2.80}_{-2.79}$	$32.09^{+2.53}_{-2.54}$
$L_{\rm sym}$ [MeV]	$44.3^{+35.1}_{-52.1}$	$46.1^{+42.8}_{-57.2}$	$40.9^{+21.9}_{-15.3}$	$40.5^{+20.5}_{-14.3}$
$K_{\rm sym}$ [MeV]	$-97.1^{+101.0}_{-46.1}$	$-83.8^{+99.4}_{-50.6}$	$-151.5^{+56.4}_{-44.6}$	$-155.8^{+62.3}_{-43.4}$
$M_{\rm max}$ [$M_\odot$]	$2.176^{+0.175}_{-0.104}$	$2.129^{+0.265}_{-0.211}$	$1.996^{+0.152}_{-0.163}$	$1.982^{+0.149}_{-0.151}$
$R(M_{\rm max})$ [km]	$11.32^{+0.61}_{-0.57}$	$11.30^{+0.64}_{-0.76}$	$10.45^{+0.54}_{-0.44}$	$10.36^{+0.53}_{-0.37}$
$R_{1.4}$ [km]	$12.82^{+0.70}_{-0.64}$	$12.77^{+0.85}_{-0.80}$	$11.82^{+0.54}_{-0.55}$	$11.74^{+0.49}_{-0.52}$
$\Lambda_{1.4}$	$516.0^{+209.6}_{-151.9}$	$521.1^{+225.8}_{-183.9}$	$378.0^{+128.8}_{-96.2}$	$360.5^{+112.0}_{-84.6}$

SBI - PyMultiNest differences are small

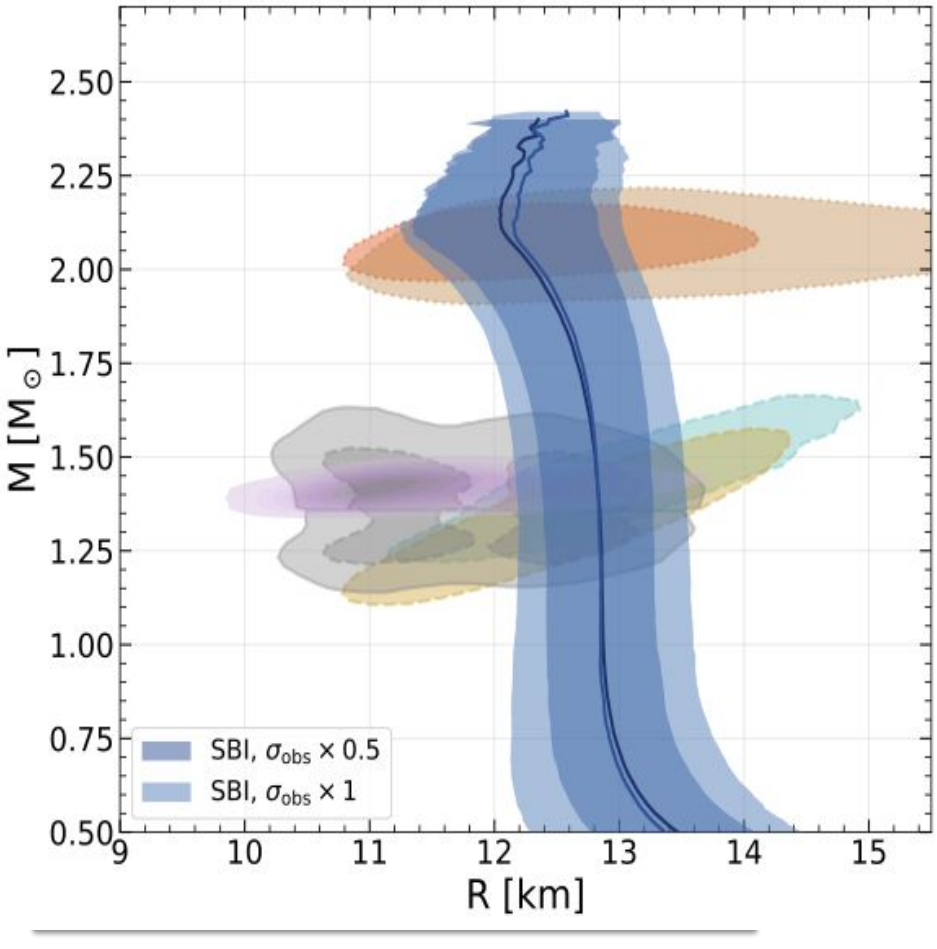
- R1.4 shift: 0.05 km (0.39%) for DDB; 0.08 km (0.68%) for RMF-NL
- Mmax shift: 0.047 Msun (2.21%) for DDB; 0.014 Msun (0.71%) for RMF-NL
- Lambda1.4 shift: 5.1 (0.98%) for DDB; 17.5 (4.85%) for RMF-NL

TARP DIAGNOSTIC TEST





- **DDB:** 68% → **68.2% ECP** (+0.2 pp); 90% → **91.0% ECP** (+1.0 pp)
- **RMF-NL:** 68% → **64.0% ECP** (−4.0 pp); 90% → **90.0% ECP** (0.0 pp)
- **Overall:** TARP curves remain close to ideal coverage, indicating well-calibrated SBI posteriors



Train Once, Re-Infer in Seconds

- Same SBI network tested at **sigma_obs x 1** and **sigma_obs x 0.5 — no retraining**
- Reducing uncertainty tightens posteriors: **R1.4 width shrinks by ~35%**, **R(Mmax) by ~24%**
- Median predictions stay stable: **Mmax = 2.179 Msun**, **R1.4 = 12.83 km**
- Speed advantage: **30,000 posterior samples in ~2.5 s on CPU** vs PyMultiNest needs a new run for each uncertainty setup

Conclusion & Future Outlook

From proof-of-concept to production-scale EOS inference

- **Validated proof-of-concept**
  SBI recovers PyMultiNest-level RMF coupling, EOS, M-R, and tidal posteriors.

- **Speed breakthrough**
  30,000 posterior samples in ~2.5 s on CPU after one-time training.

ET/CE motivation: growing GW catalogues will make repeated event-by-event Bayesian EOS scans prohibitively expensive

- **Next step: real multimessenger likelihoods**
  Directly include NICER M-R posteriors, GW tidal likelihoods, and updated massive-pulsar constraints.

- **Expand physics space**
  Add pQCD constraints, more EOS families, hyperons/quarks/phase transitions, and composition-dependent models.

- **Production goal**
  Fast joint microscopic-EOS inference as new NICER/GW data and revised uncertainties arrive.

Take-home message:

SBI shifts EOS inference from repeated costly scans to reusable, catalogue-scale inference for the ET/CE era.