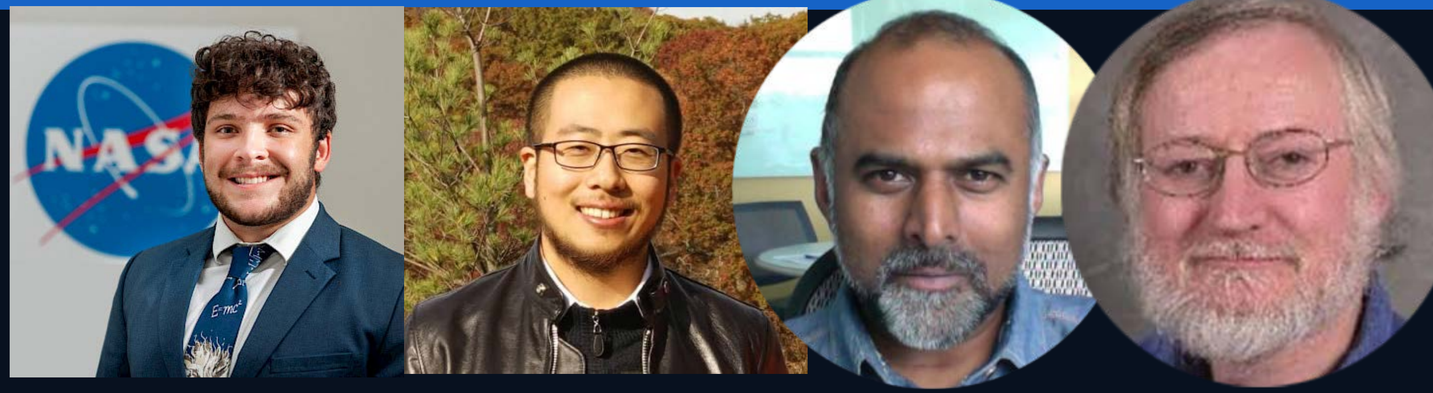


Modeling the Neutron Star EOS with Variational Autoencoder

Presenter: Tianqi Zhao 趙天奇 (UC, Berkeley & UW, Seattle)

Colaborators: Alex Ross, Fanglida Yan, Sanjay Reddy, James Lattimer

arXiv:2601.21231

N3AS Network for Neutrinos, Nuclear Astrophysics, and Symmetries
PHYSICS FRONTIER CENTER

INSTITUTE for NUCLEAR THEORY

Why this problem matters

Scientific bottleneck

- EOS controls the full mass–radius relation through the TOV equations.
- Theory spaces are high-dimensional; direct exploration is expensive.
- We want a compact generator that preserves causality, stability, and astrophysical constraints.

Strategy: Latent controls [$M_{\max}$, $R_{1.4}$, z_v]

Train an PSL-VAE on Skyrme EOSs, supervise the latent space with $M_{\max}$ and $R_{1.4}$, and let stochastic variables z_v capture the remaining EOS freedom.

Multimessenger anchors

- **Radio timing** — precise masses → lower bound on $M_{\max}$
- **NICER X-ray** — X-ray pulse profiles → radius information, e.g. $R_{1.4}$
- **LIGO/Virgo** — Gravitational waves of binary mergers → tidal deformability complementary to $R_{1.4}$

Goal for the talk

Show a 3-parameter representation of 7-parameter Skyrme EOSs that can reconstruct and generate causal, stable NS EOS consistent with target observables.

Table 1. Sampling ranges for the nuclear properties used to determine the Skyrme EDF. The neutron and proton bare masses are fixed to $m_n = m_p = m = 939$ MeV.

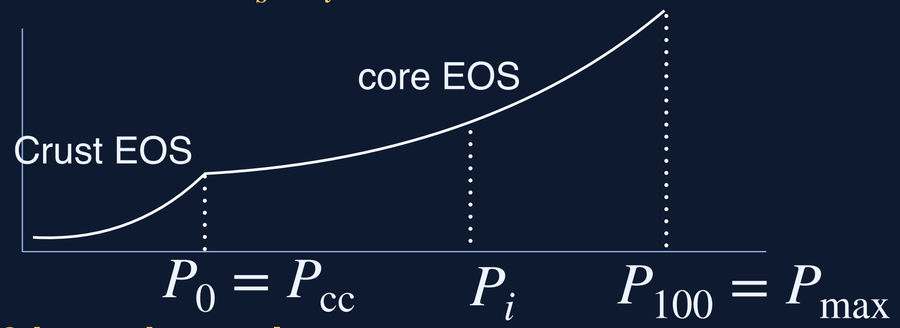
n_S [fm^{-3}]	B [MeV]	K [MeV]	J [MeV]	L [MeV]	m_S^*/m	$\Delta m_S^*/m$
(0.155, 0.165)	(-16.2, -15.8)	(180, 280)	(28, 36)	(0, 150)	(0.7, 1)	(-0.5, 0.5)

Training representation and EOS reconstruction

The input vector stores a sound-speed profile on a log-pressure grid plus six boundary quantities.

Input vector $x \in \mathbf{R}^{107}$

101 array of $c_s^2(P_i)$



6 boundary values

n_{cc} $\varepsilon_{\mathrm{cc}}$ P_{cc}

$n_{\max}$ $\varepsilon_{\max}$ $P_{\max}$

All 107 inputs are log-scaled and standardized before training.

EOS reconstruction from x

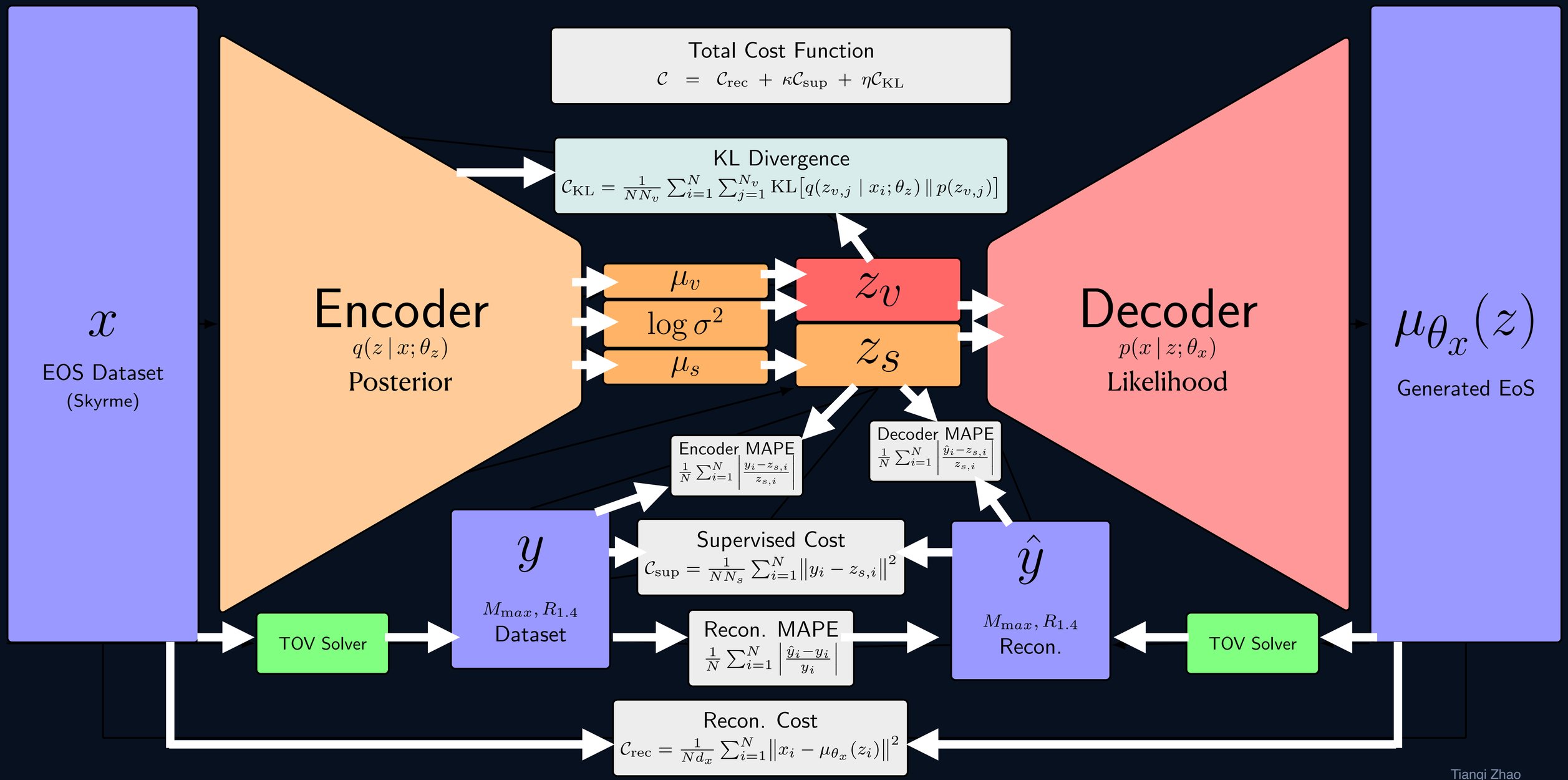
Spline the sound speed $c_s^2(P)$, then integrate towards the core from the crust–core transition.

$$\varepsilon(P) = \varepsilon_{\mathrm{cc}} + \int_{P_{\mathrm{cc}}}^{P} \frac{dP'}{c_s^2(P')}$$

Physical post-processing

- Join the reconstructed core EOS to the SLy4 crust at P_{cc}.
- Thermodynamic stability is satisfied by using log-scaled x.
- Require causal sound speeds, $\log c_s^2(P_i) \leq 0$.
- Solve TOV equations for the supervised observable $y \in [M_{\max}, R_{1.4}]$.

Framework of Partially Supervised Latent-Space Variational Autoencode (PSL-VAE)



Tianqi Zhao

Model selection and latent-space structure

The selected configuration balances low decoder error with a clean, interpretable 1D variational space.

Chosen model

$N_v = 1, \quad \kappa = 5, \quad \eta = 0.01$

Decoder fidelity

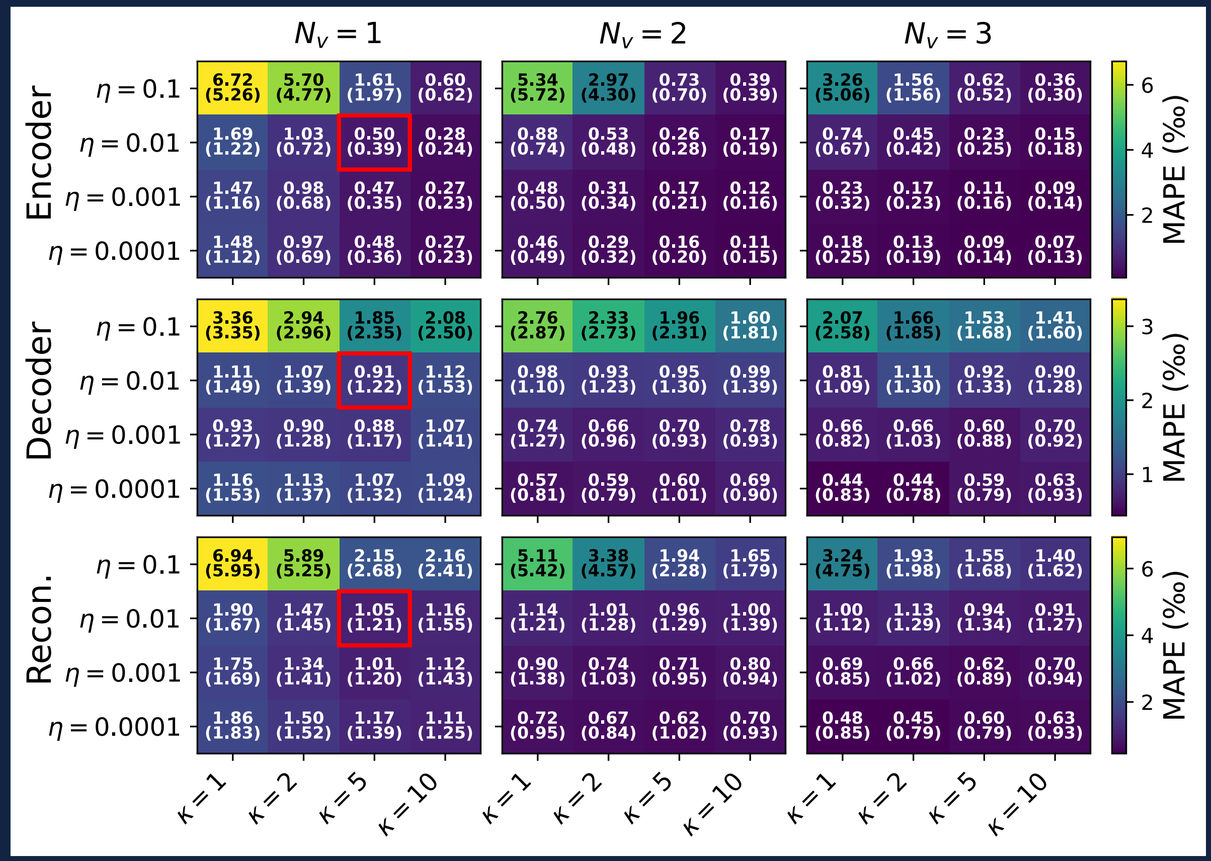
≈ 0.09%
MAPE for $M_{\max}$

≈ 0.12%
MAPE for $R_{1.4}$

Interpretation

Two supervised coordinates already explain most of the Skyrme EOS variation; one extra variational stochastic dimension is enough to capture the residual structure.

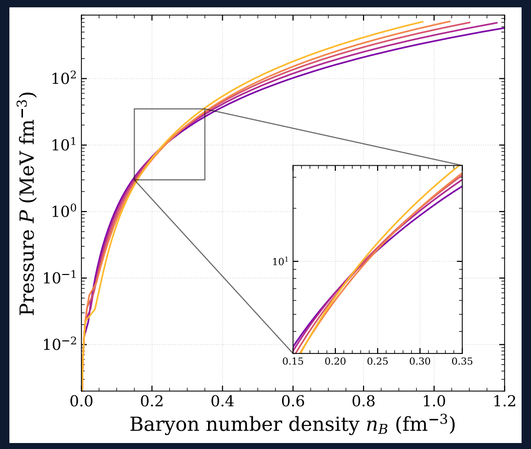
3 MAPEs of supervised observable y

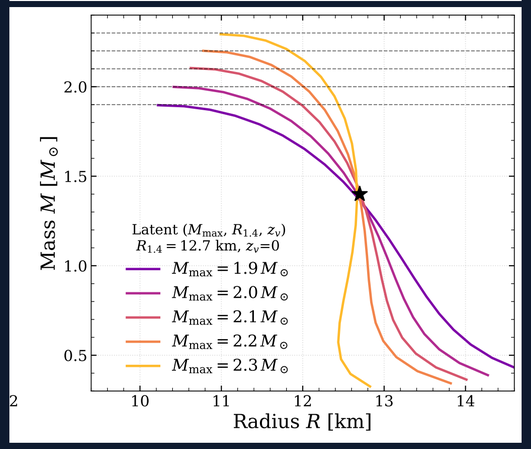


What each latent coordinate controls

Vary one latent parameter at a time, decode the EOS, then propagate it through the TOV solver.

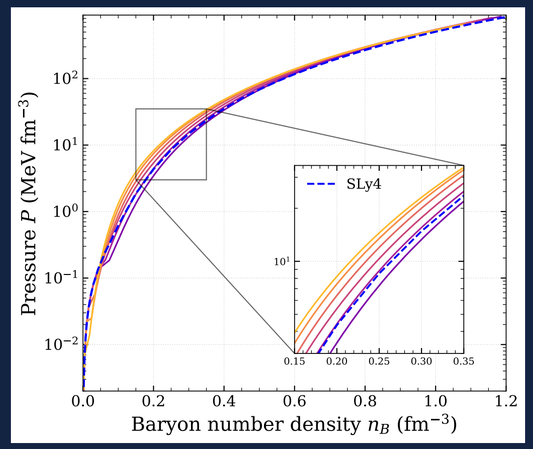
Vary $M_{\max}$

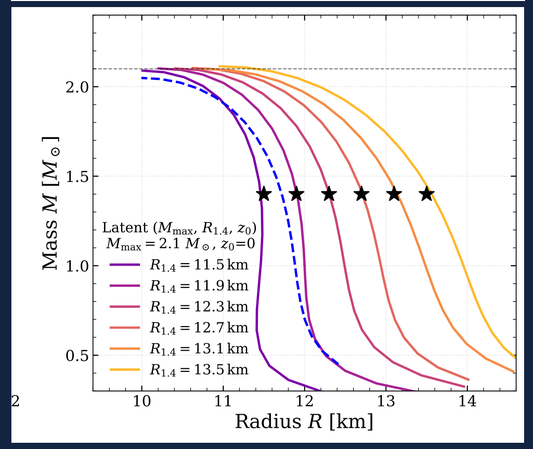




Higher M_max stiffens the high-density EOS and raises mass support.

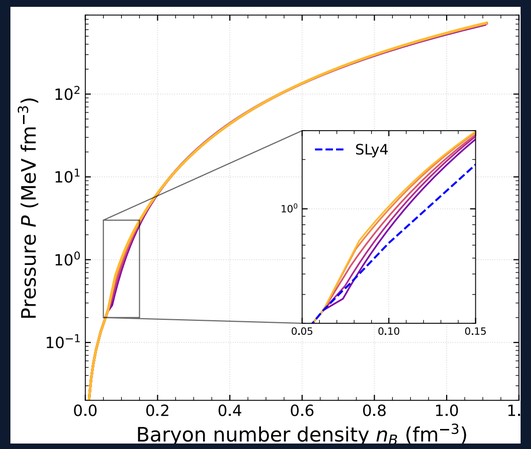
Vary $R_{1.4}$

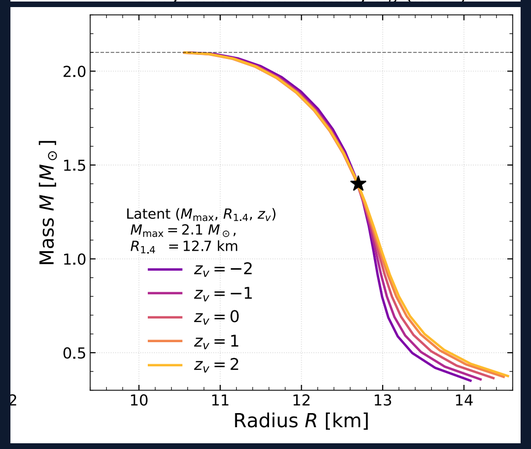




R_1.4 changes the pressure near saturation and shifts the MR curve horizontally.

Vary z_ν



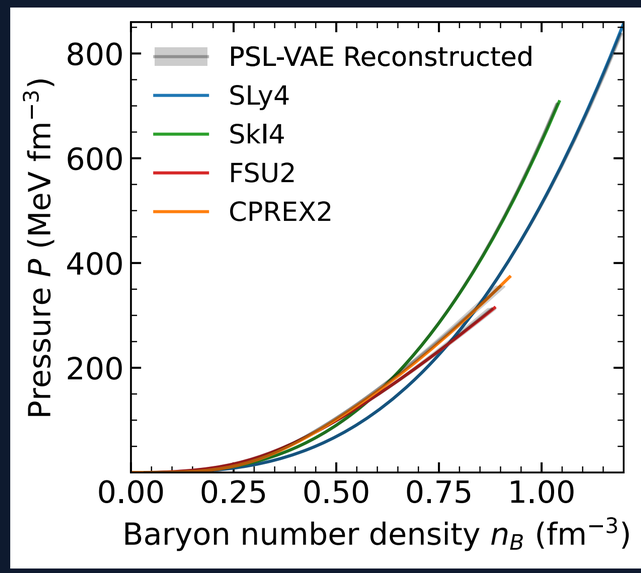


z_ν produces differences at low-density, strongest in lower-mass radii.

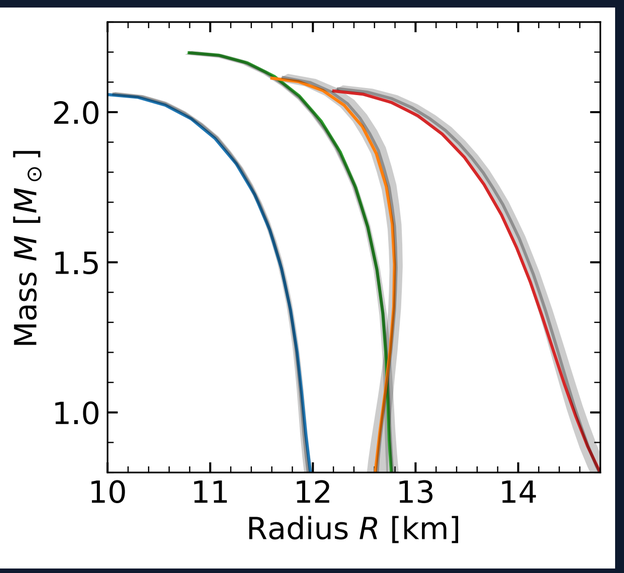
Accurate reconstruction with a 3-parameter latent space

Benchmark Skyrme EOSs vs.decoded reconstructed EOS

EOS $P(n_B)$



MR relation



Compression payoff

9 Skyrme parameters → **3 latent space**

$[M_{\max}, R_{1.4}, z_v]$

Target observables (SLy4):

- $M_{\max}$ = 2.046 M⊙
  $R_{1.4}$ = 11.717 km.

Latent parameters (SLy4):

- $M_{\max}$ = 2.058 ± 0.004 M⊙
  $R_{1.4}$ =11.74 ± 0.023 km
  z_v = 0.560 ± 0.049

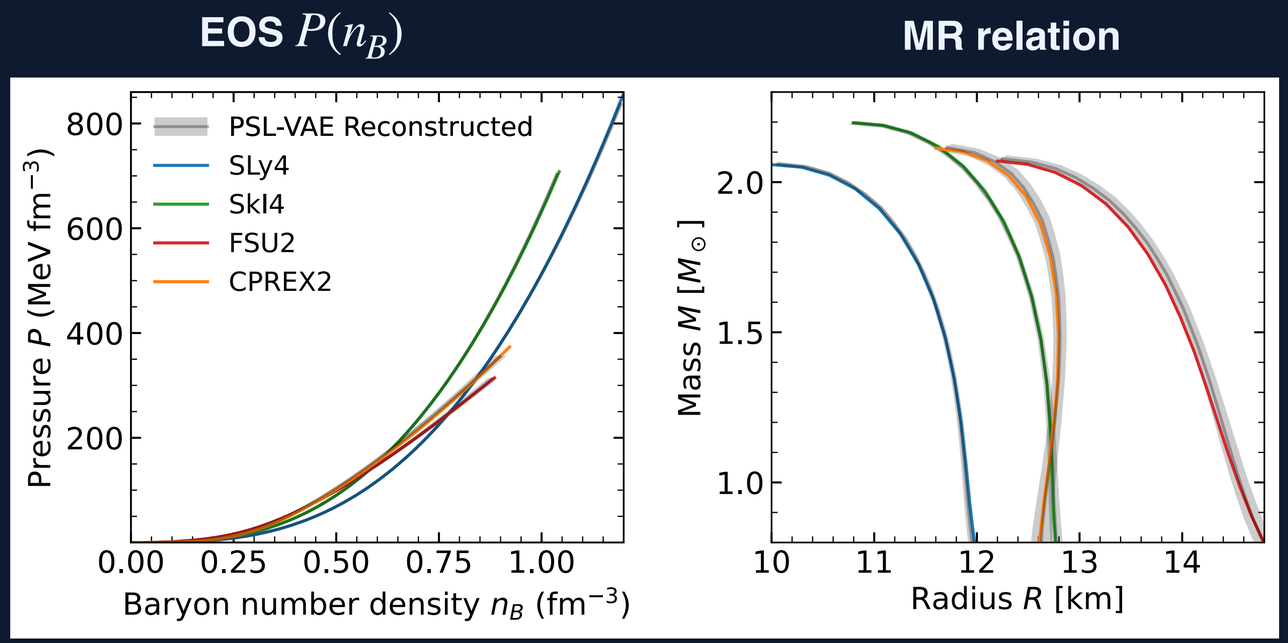
Summary:

PSL-VAE is a practical surrogate for fast EOS generation and later Bayesian inference with multimessenger data.

Accurate reconstruction with a 4-parameter latent space

Thank You!

Benchmark Skyrme EOSs vs.decoded reconstructed EOS



Compression payoff

9 RMF parameters → **4 latent space**

$[M_{\max}, R_{1.4}, z_{v1}, z_{v2}]$

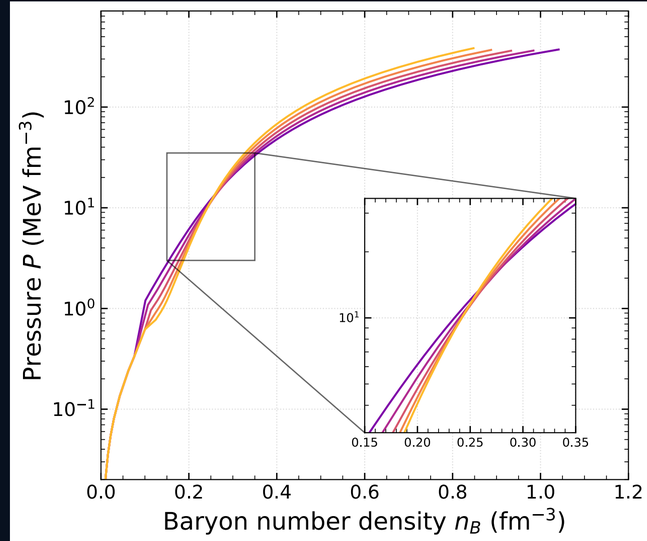
Summary:

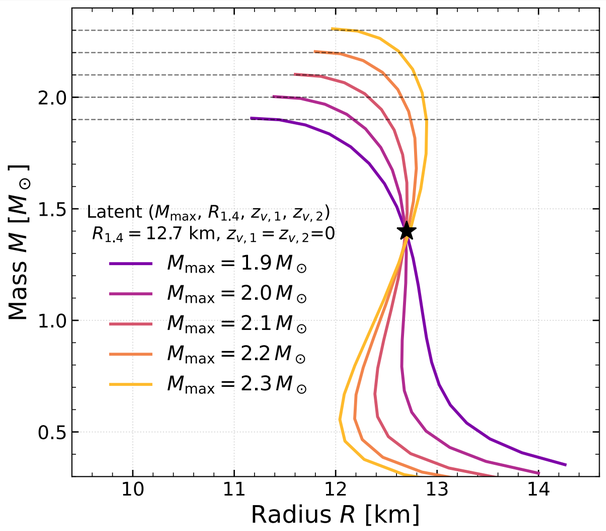
PSL-VAE is a practical surrogate for fast EOS generation and later Bayesian inference with multimessenger data.

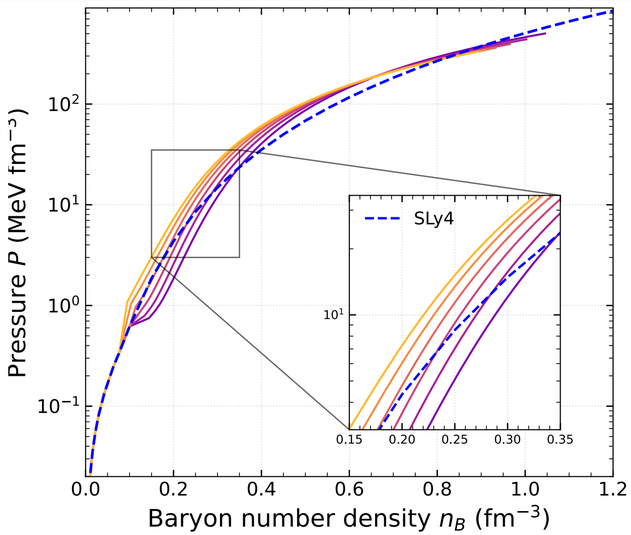
Future:

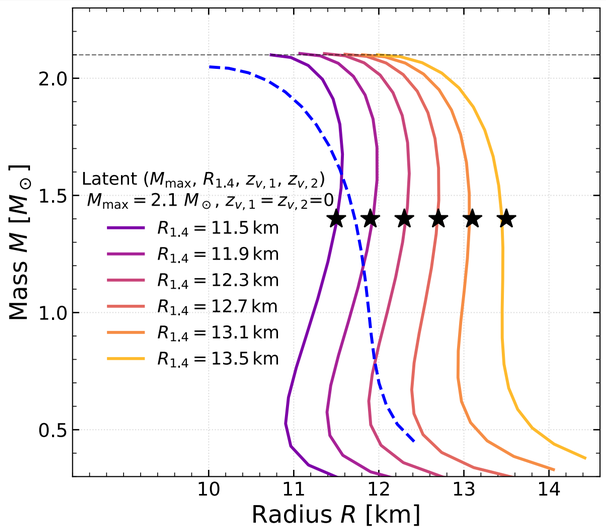
Expand the training set beyond Skyrme to RMF, phase transitions, and exotic EOSs.

RMF

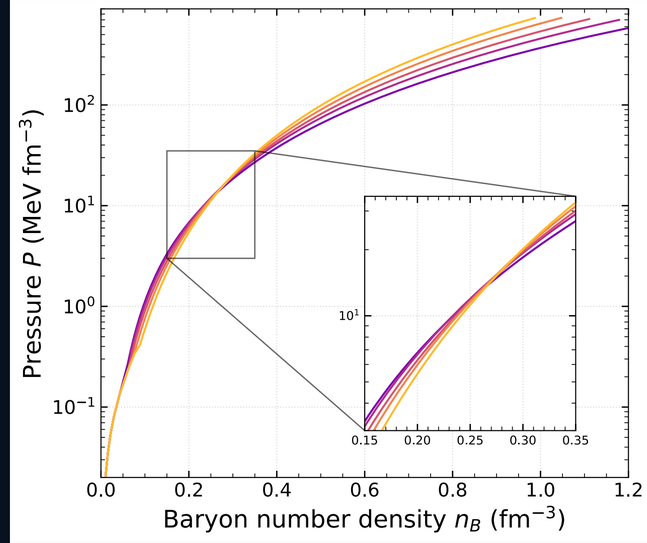

Pressure P (MeV fm^{-3})
Baryon number density n_B (fm^{-3})

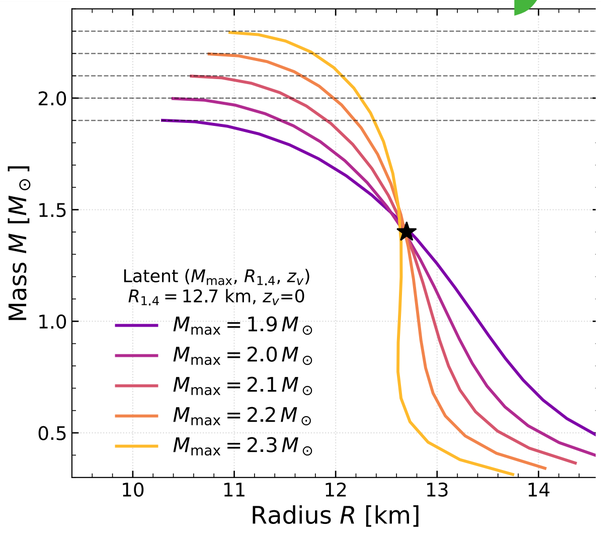

Latent (M_{max}, $R_{1.4}$, $z_{v,1}$, $z_{v,2}$)
$R_{1.4}$ = 12.7 km, $z_{v,1} = z_{v,2}$=0
$M_{max} = 1.9\,M_\odot$
$M_{max} = 2.0\,M_\odot$
$M_{max} = 2.1\,M_\odot$
$M_{max} = 2.2\,M_\odot$
$M_{max} = 2.3\,M_\odot$
Mass M [$M_\odot$]
Radius R [km]

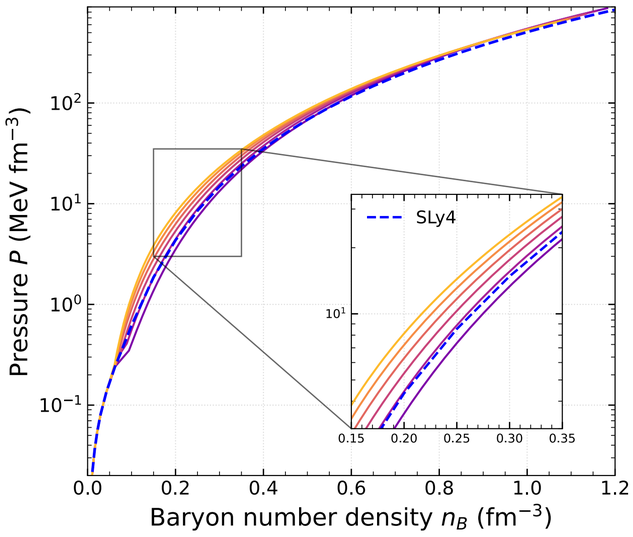

SLy4
Pressure P (MeV fm^{-3})
Baryon number density n_B (fm^{-3})

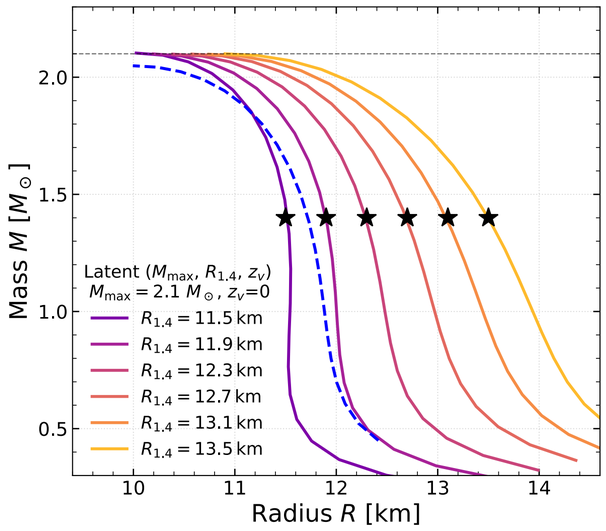

Latent (M_{max}, $R_{1.4}$, $z_{v,1}$, $z_{v,2}$)
M_{max} = 2.1 $M_\odot$, $z_{v,1} = z_{v,2}$=0
$R_{1.4}$ = 11.5 km
$R_{1.4}$ = 11.9 km
$R_{1.4}$ = 12.3 km
$R_{1.4}$ = 12.7 km
$R_{1.4}$ = 13.1 km
$R_{1.4}$ = 13.5 km
Mass M [$M_\odot$]
Radius R [km]

Skyrme

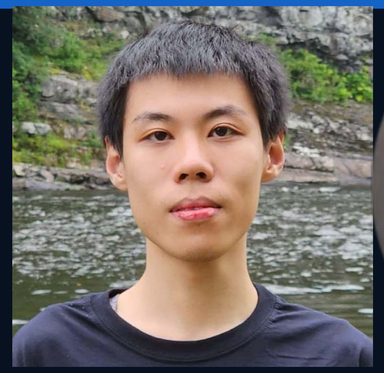

Pressure P (MeV fm^{-3})
Baryon number density n_B (fm^{-3})


Latent (M_{max}, $R_{1.4}$, z_v)
$R_{1.4}$ = 12.7 km, z_v=0
$M_{max} = 1.9\,M_\odot$
$M_{max} = 2.0\,M_\odot$
$M_{max} = 2.1\,M_\odot$
$M_{max} = 2.2\,M_\odot$
$M_{max} = 2.3\,M_\odot$
Mass M [$M_\odot$]
Radius R [km]

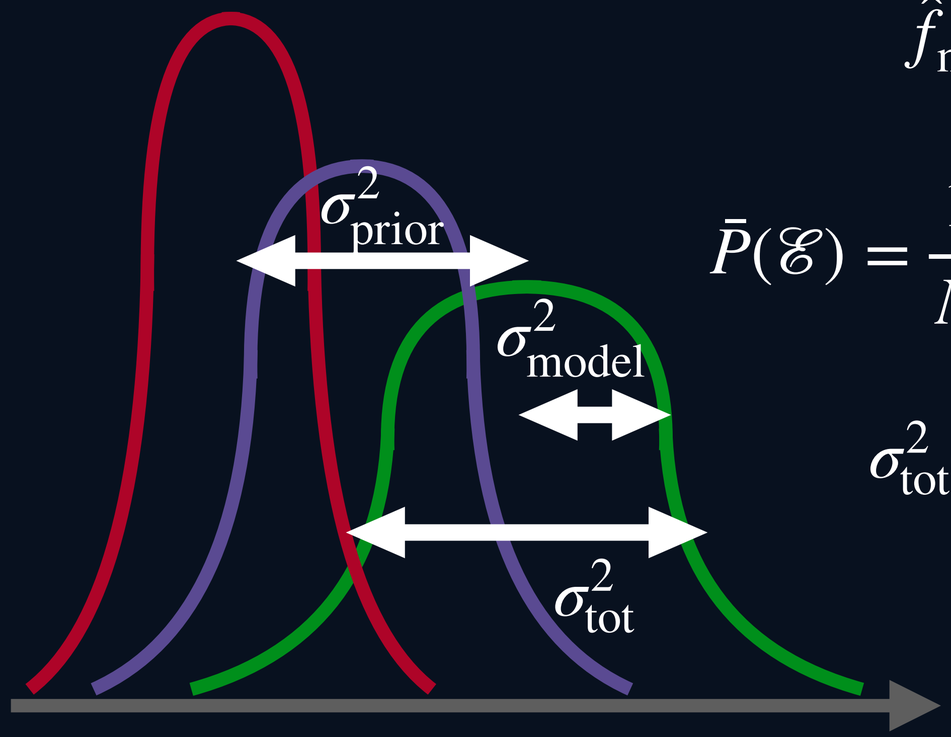

SLy4
Pressure P (MeV fm^{-3})
Baryon number density n_B (fm^{-3})


Latent (M_{max}, $R_{1.4}$, z_v)
M_{max} = 2.1 $M_\odot$, z_v=0
$R_{1.4}$ = 11.5 km
$R_{1.4}$ = 11.9 km
$R_{1.4}$ = 12.3 km
$R_{1.4}$ = 12.7 km
$R_{1.4}$ = 13.1 km
$R_{1.4}$ = 13.5 km
Mass M [$M_\odot$]
Radius R [km]

Investigate predictions with different prior choices



$$\hat{f}_{\rm mix}(P \mid \mathscr{E}) = \frac{1}{N}\sum_{i=1}^{N} \hat{f}_i(P \mid \mathscr{E})$$



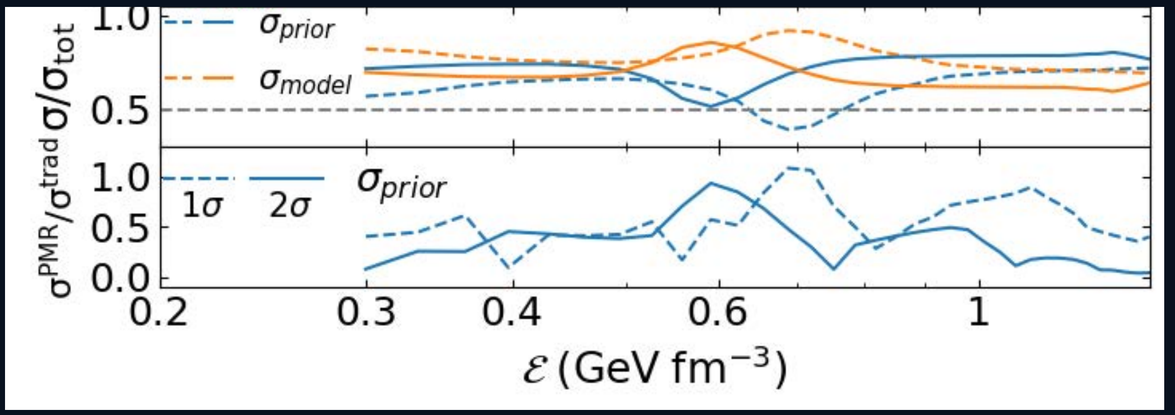
$$\bar{P}(\mathscr{E}) = \frac{1}{N}\sum_{i=1}^{N} \bar{P}_i(\mathscr{E}) \qquad \bar{P}_i(\mathscr{E}) = \int P\, \hat{f}_i(P \mid \mathscr{E})\, dP$$

$$\sigma^2_{\rm tot}(\mathscr{E}) = \int \left[P - \bar{P}(\mathscr{E})\right]^2 \hat{f}_{\rm mix}(P \mid \mathscr{E})\, dP \;= \sigma^2_{\rm model}(\mathscr{E}) + \sigma^2_{\rm prior}(\mathscr{E})$$

$P(\mathscr{E} = 150\ \text{MeV fm}^{-3})$

Due to model constraints:

$$\sigma^2_{\rm model}(\mathscr{E}) = \frac{1}{N}\sum_{i=1}^{N} \int \left[P - \bar{P}_i(\mathscr{E})\right]^2 \hat{f}_i(P \mid \mathscr{E})\, dP$$



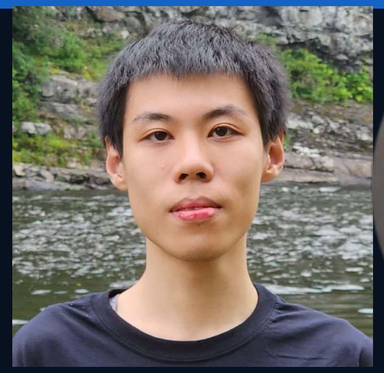
Due prior choice:

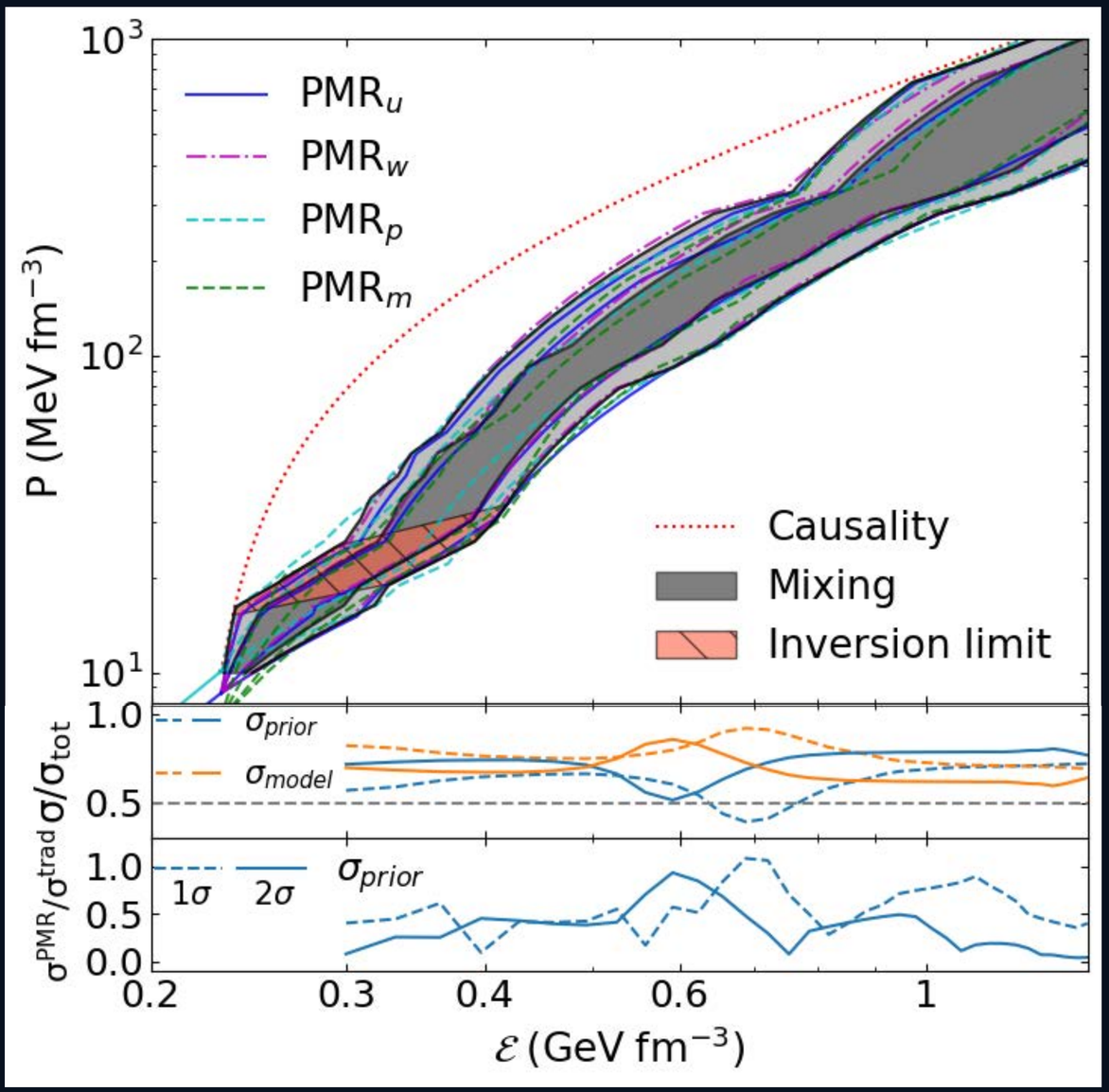
$$\sigma^2_{\rm prior}(\mathscr{E}) = \frac{1}{N}\sum_{i=1}^{N} \left[\bar{P}_i(\mathscr{E}) - \bar{P}(\mathscr{E})\right]^2$$

arXiv:2601.04294

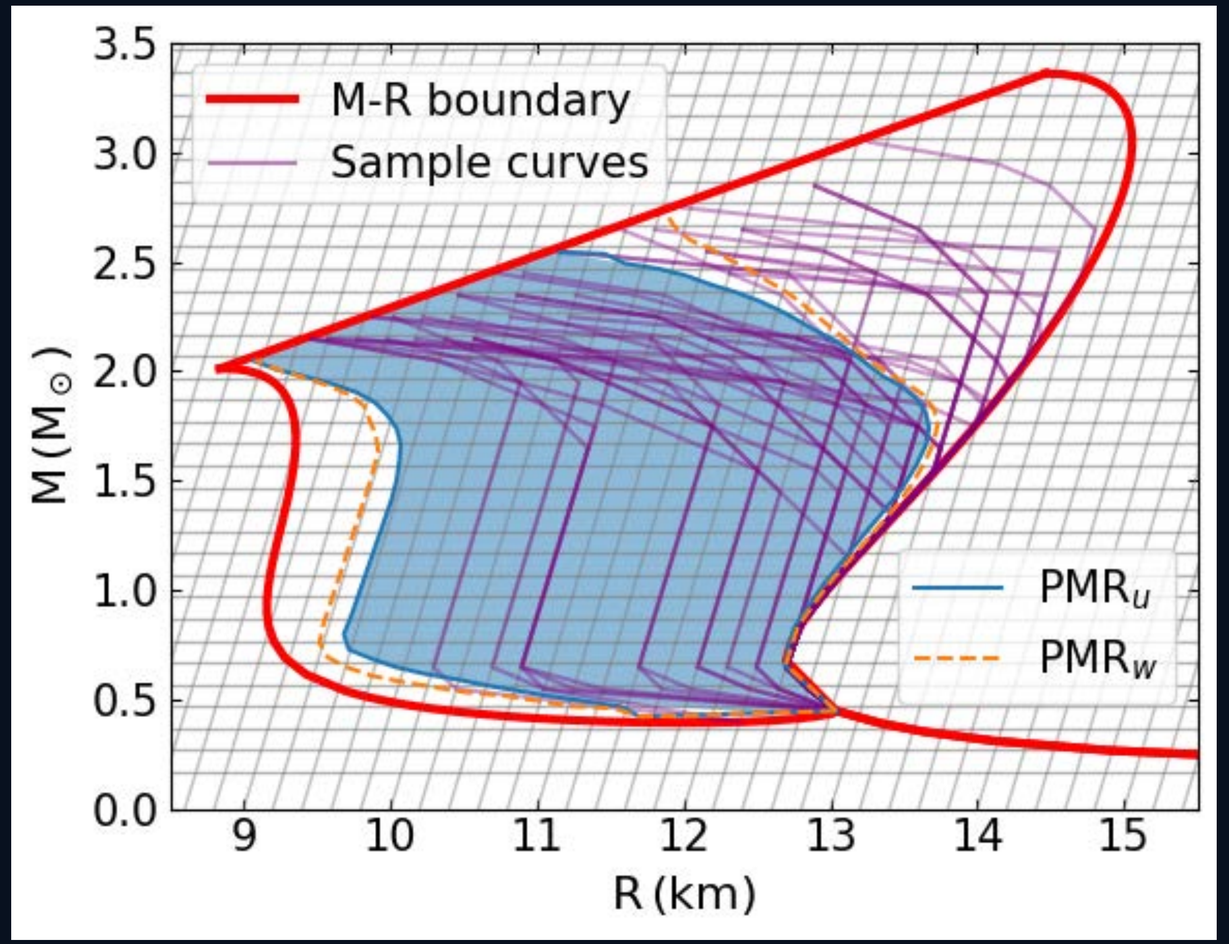
Tianqi Zhao

Investigate predictions with different prior choices





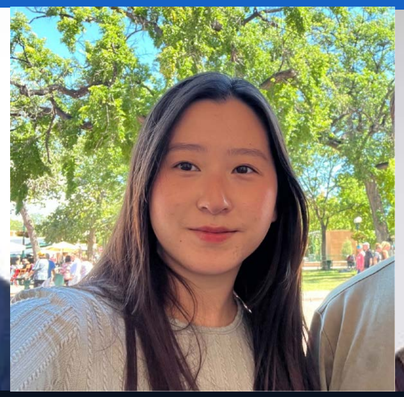
Prior in Mass-Radius (PMR)



arXiv:2601.04294

Tianqi Zhao

Modeling the EOS with Gaussian Process



ASSUMING the energies of matter at two arbitrary densities, proton fractions, and temperatures follow a joint Gaussian distribution:

$$p\begin{pmatrix}\mathscr{E}(n_i,x_j,T_k)\\ \mathscr{E}(n_l,x_m,T_n)\end{pmatrix}\sim\mathcal{N}\left(\begin{pmatrix}\mu_{ijk}\\ \mu_{lmn}\end{pmatrix},\begin{pmatrix}K_{ijk,ijk} & K_{ijk,lmn}\\ K_{lmn,ijk} & K_{lmn,lmn}\end{pmatrix}\right),$$

GIVEN a subset of these points is measured with covariance.

Only 6 Sampled EOS !!!

THEN, the conditional distribution of other points also jointly Gaussian.

Their derivatives is also jointly Gaussian !!!

In PRACTICE, Squared exponential Kernels:

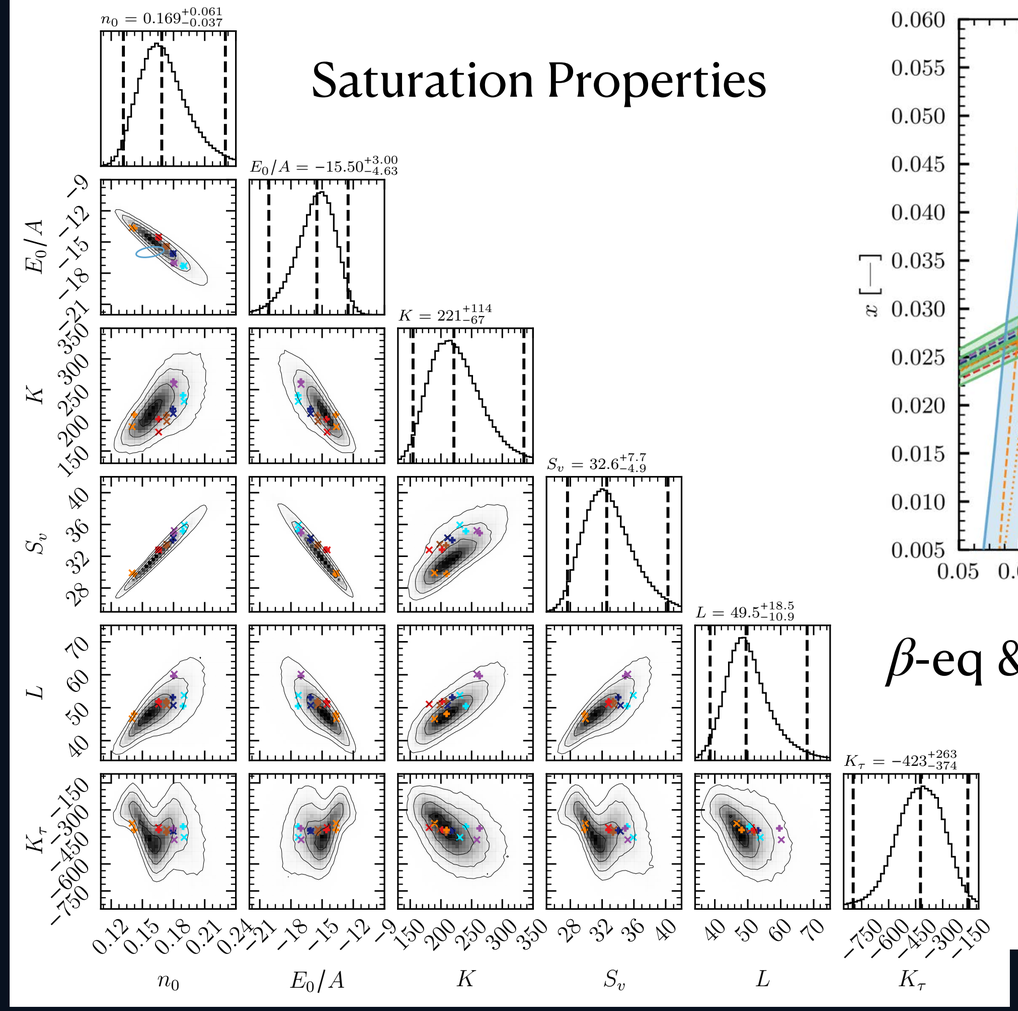
Kernels can be density and iso-spin dependent !!!

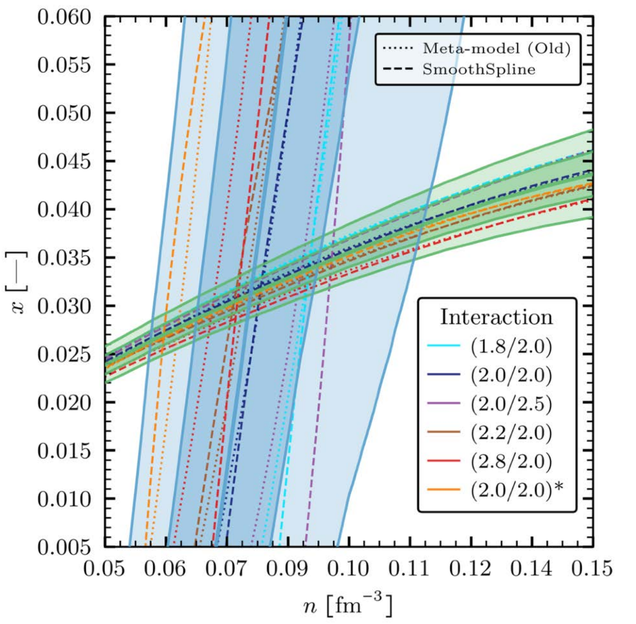
$$K_{ijk,lmn}=\sigma_{\mathscr{E}}^2\exp\left[-\frac{(n_i-n_l)^2}{2\ell_n^2}-\frac{(x_j-x_m)^2}{2\ell_x^2}-\frac{(T_k-T_n)^2}{2\ell_T^2}\right]$$

arXiv:2608.09678

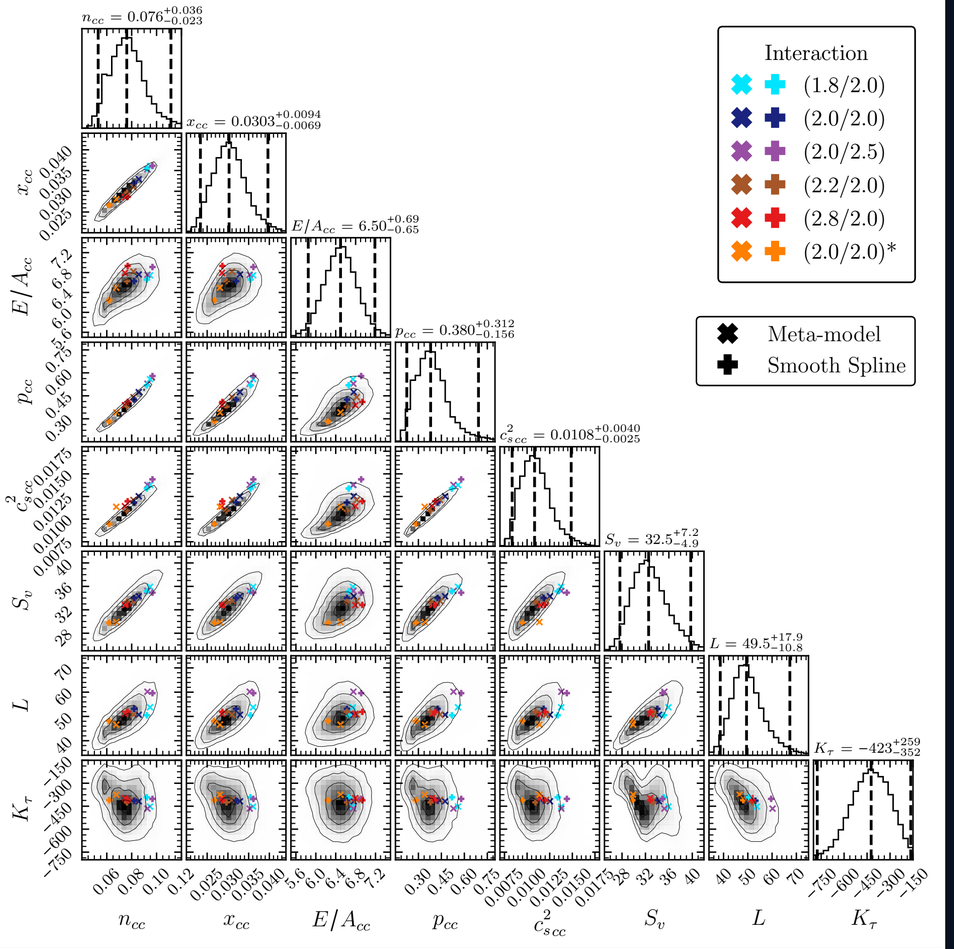
Tianqi Zhao

Modeling the EOS with Gaussian Process





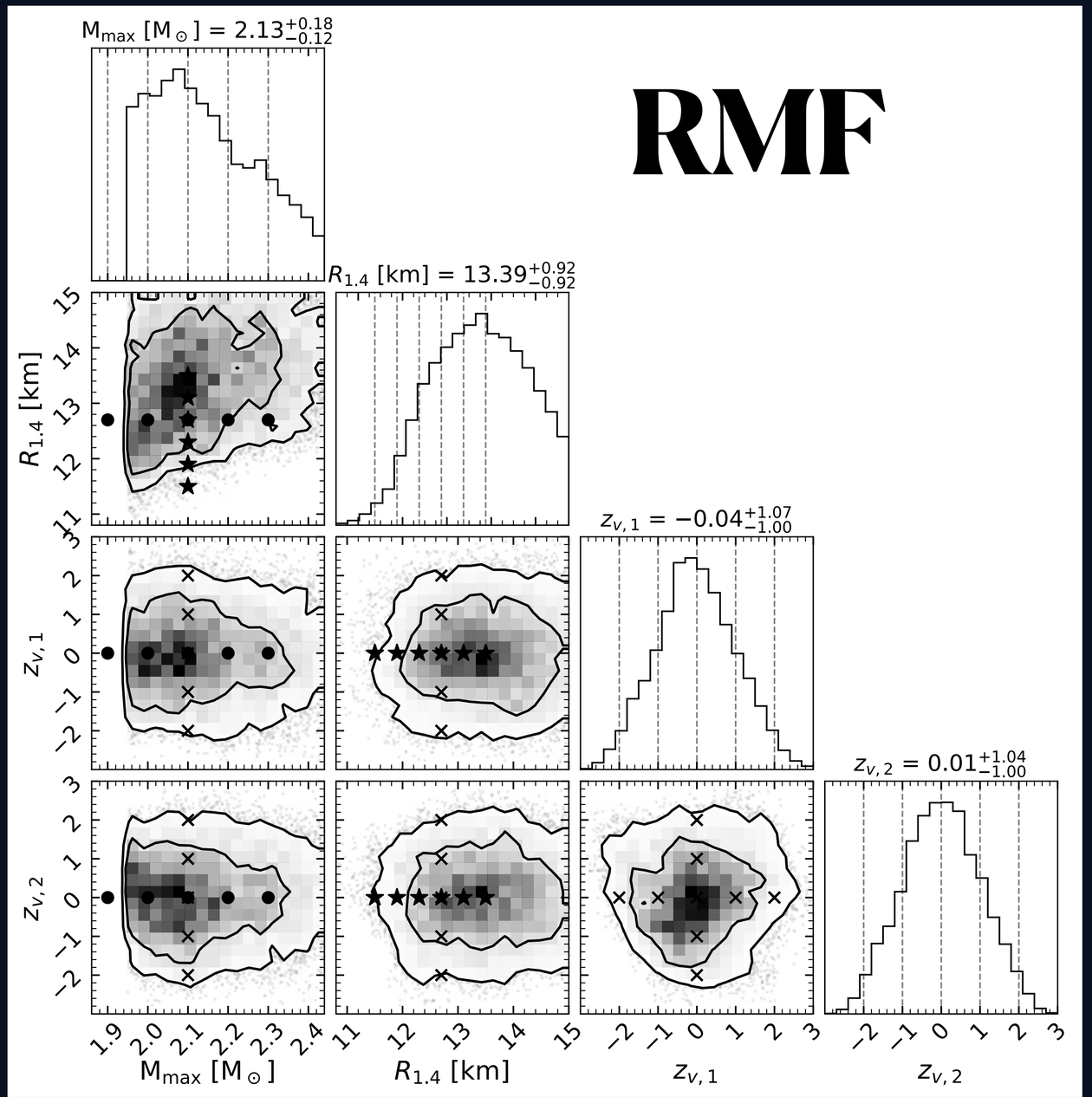
β-eq & crust-core instability

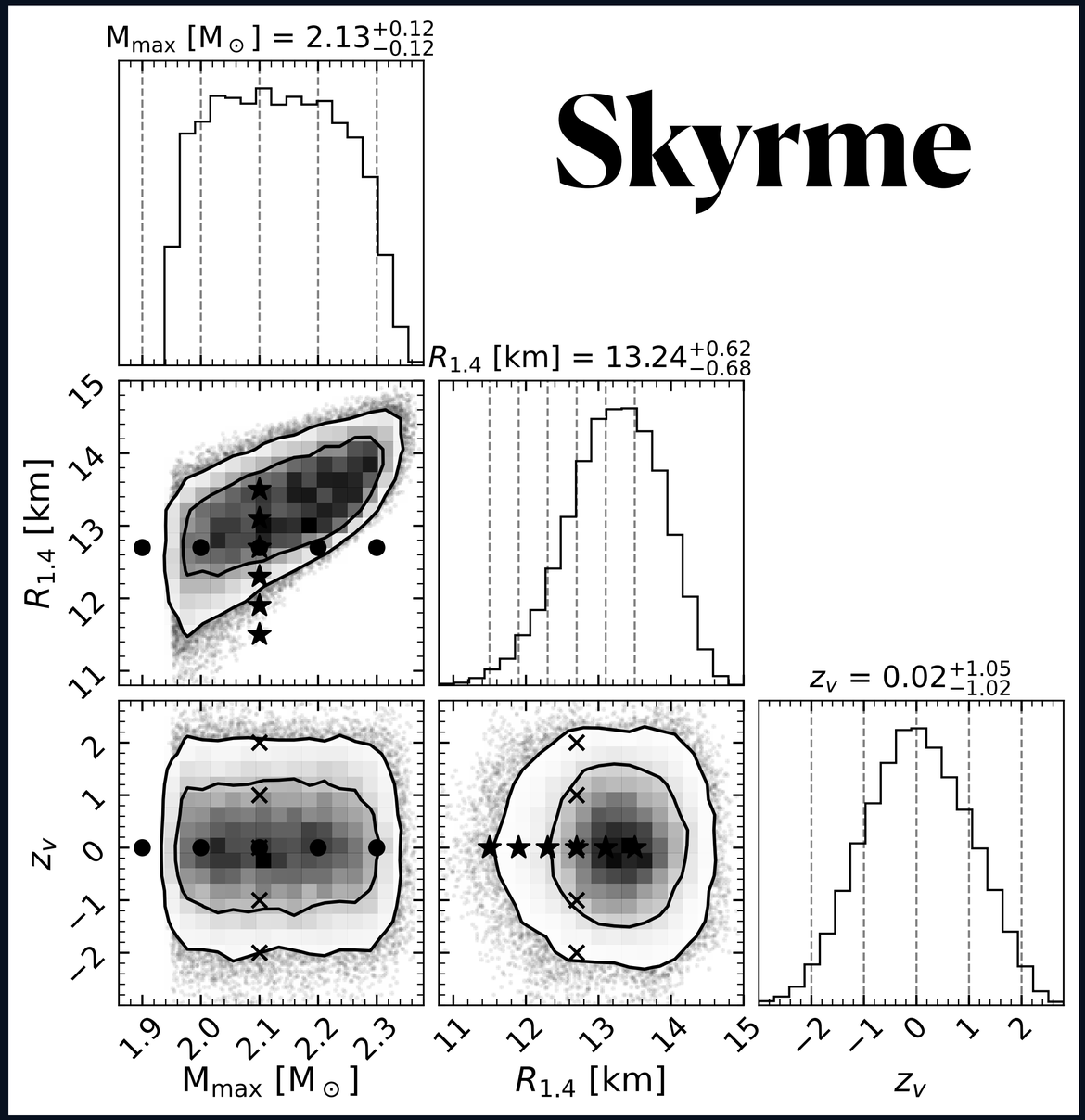


arXiv:2608.09678

Tianqi Zhao

Back up slides


RMF
$M_{max}\ [M_\odot] = 2.13^{+0.18}_{-0.12}$
$R_{1.4}$ [km] = $13.39^{+0.92}_{-0.92}$
$z_{v,1} = -0.04^{+1.07}_{-1.00}$
$z_{v,2} = 0.01^{+1.04}_{-1.00}$
$M_{max}\ [M_\odot]$
$R_{1.4}$ [km]
$z_{v,1}$
$z_{v,2}$


Skyrme
$M_{max}\ [M_\odot] = 2.13^{+0.12}_{-0.12}$
$R_{1.4}$ [km] = $13.24^{+0.62}_{-0.68}$
$z_v = 0.02^{+1.05}_{-1.02}$
$M_{max}\ [M_\odot]$
$R_{1.4}$ [km]
z_v

VAE: a framework to find latent parameters and their distributions

Goal:
Maximize the marginalized likelihood of N samples of data x

$$\mathcal{J}(\theta_x) = \sum_{i=1}^{N} \log \int p(x_i \mid z; \theta_x)\, p(z)\, dz\,.$$

Liklihood of sample x_i

Prior of latent z

neural network weights & biases

Step 1:
Take the posterior $q(z \mid x; \theta_z)$ as a variational distribution

$$= \sum_{i=1}^{N} \log \int q(z \mid x_i; \theta_z) \frac{p(x_i \mid z; \theta_x)\, p(z)}{q(z \mid x_i; \theta_z)} dz$$

$$\geq \sum_{i=1}^{N} \int q(z \mid x_i; \theta_z) \log \frac{p(x_i \mid z; \theta_x)\, p(z)}{q(z \mid x_i; \theta_z)} dz$$

The evidence lower bound($\mathcal{J}_{ELBO}$) from Jensen's inequality

Step 2:
Assume prior, likelihood, and posterior all follow normal distributions.

$$p(z) = \mathcal{N}(0, I)$$

$$q(z \mid x; \theta_z) = \mathcal{N}\left(\mu_{\theta_z}(x), \mathrm{diag}[\sigma^2_{\theta_z}(x)]\right)$$

$$p(x \mid z; \theta_x) = \mathcal{N}\left(\mu_{\theta_x}(z), \sigma^2_x I\right)$$

Neural networks enter here !

$$= \sum_{i=1}^{N} \int q(z \mid x_i; \theta_z) \log p(x_i \mid z; \theta_x) dz + \mathrm{KL}\left(q(z \mid x_i; \theta_z) \,\|\, p(z)\right)$$

Semi-supervised VAE: VAE with supervised latent parameter y

VAE part: $p(z_v) = \mathcal{N}(0, I)$

$$q(z_v \mid x; \theta_{z_v}) = \mathcal{N}\left(\mu_{\theta_{z_v}}(x), \mathrm{diag}[\sigma^2_{\theta_{z_v}}(x)]\right)$$

$$p(x \mid z; \theta_x) = \mathcal{N}\left(x \mid \mu_{\theta_x}(z), \sigma_x^2 I\right)$$

Supervised part: $p(z_s) = \mathrm{Uniform}(\mathcal{S})$

$$q(z_s \mid x; \theta_{z_s}) = \delta\left(z_s - \mu_{\theta_{z_s}}(x)\right)$$

$$p(y \mid z_s) = \mathcal{N}\left(y \mid z_s, \sigma_y^2 I\right)$$

The PSL-VAE cost function: $\mathscr{C}_{\mathrm{SSVAE}} = -\frac{2\sigma_x^2}{d_x}\mathscr{J}_{ELBO} = \mathscr{C}_{rec} + \kappa\mathscr{C}_{sup} + \eta\mathscr{C}_{KL}$ $\quad \kappa = \frac{\sigma_x^2 N_s}{\sigma_y^2 d_x} \quad \eta = \frac{2\sigma_x^2 N_v}{d_x}$

$$= \frac{1}{N d_x}\sum_{i=1}^{N} \left\| x_i - \mu_{\theta_x}(z_i) \right\|^2 + \frac{\kappa}{N N_s}\sum_{i=1}^{N} \left\| y_i - \mu_{\theta_{z_s}}(x_i) \right\|^2 + \frac{\eta}{N N_v}\sum_{i=1}^{N}\sum_{j=1}^{N_v}\left[-\frac{1}{2}\left(1 + \log\sigma^2_{z_v,i,j} - \mu^2_{z_v,i,j} - \sigma^2_{z_v,i,j}\right)\right]$$

Latent z:

z_v, z_s **have dimentsion** N_v, N_s.

$$z = \{z_v, z_s\}$$

$$z_{v,i} = \mu_{\theta_{z_v}}(x_i) + \sigma_{\theta_{z_v}}(x_i) \odot \varepsilon$$

$$z_{s,i} = \mu_{\theta_{z_s}}(x_i) \qquad \varepsilon \sim \mathcal{N}(0, I)$$

Data x **and** y**:**

x **has dimension** d_x, y **has dimension** N_s **like** z_s

Table 2. VAE boundary-value inputs at the core–crust transition and at the maximum-density endpoint.

Quantity	Core–crust transition	Maximum
Baryon number density	$n_{\mathrm{cc}} \in [0.039, 0.112]$ fm^{-3}	$n_{\mathrm{max}} \in [0.912, 0.141]$ fm^{-3}
Energy density	$\mathcal{E}_{\mathrm{cc}} \in [37.2, 106.4]$MeV fm^{-3}	$\mathcal{E}_{\mathrm{max}} \in [1160, 1859]$MeV fm^{-3}
Pressure	$P_{\mathrm{cc}} \in [0.1, 0.7]$MeV fm^{-3}	$P_{\mathrm{max}} \in [579, 1099]$MeV fm^{-3}

Table 3. Dataset splits used for training and evaluation.

Split	Size	Purpose	During training
Training	∼60%	Update model weights	Yes
Validation	∼20%	Best model checkpoint selection & early stopping	Yes
Selection	∼10%	Hyperparameter and model selection	No
Test	∼10%	Final unbiased evaluation	No

Table 3. Neural network architecture.

Layer	Type	Neurons	Activation
Input	N/A	107	N/A
Layer 1	Dense	128	SiLU
Layer 2	Dense	64	SiLU
Layer 3	Dense	64	SiLU
Layer 4	Dense	$N_s + 2N_v$	Linear
Latent layer (a): μ_{s}	Slice	N_s	Linear
Latent layer (b_1): μ_{v}	Slice	N_v	Linear
Latent layer (b_2): $\log \sigma_{\mathrm{v}}^2$	Slice	N_v	Linear
Latent layer (b): z_{v}	Sampling	N_v	N/A
Latent layer: $z = [\mu_{\mathrm{s}}, z_{\mathrm{v}}]$	Concatenate	$N_s + N_v$	N/A
Layer 5	Dense	64	SiLU
Layer 6	Dense	64	SiLU
Layer 7	Dense	128	SiLU
Layer 8	Dense	107	Linear
Output (a): c_{s}^2	Slice	101	Softnegative
Output (b): $[n, \varepsilon, p]_{\mathrm{cc,max}}$	Slice	6	Linear
Output	Concatenate	107	N/A

Physics pipeline: from Skyrme functional to observables

The supervised latent variables are derived from the EOS through beta equilibrium and TOV solutions.

Skyrme EDF

10 parameters
mean-field interaction

β equilibrium

minimize E(n_B,δ)
at fixed density

Stability + crust match

spinodal cutoff n_cc
join to SLy4 crust

EOS P(ε)

barotropic relation
for the stellar core

TOV solver

M(R), M_max,
R_1.4

Representative EDF term

$$H_{n,p} = T(n_{n,p}, m^*_{n,p}) + H_{pot}(n_n, n_p)$$

TOV closure

$$dP/dr = -G(m + 4\pi Pr^3/c^2)(\varepsilon + P) / [\, r(rc^2 - 2Gm) \,]$$

$$dm/dr = 4\pi r^2 \varepsilon / c^2$$

What becomes supervised?

Two directly interpretable outputs are used as latent coordinates:

M_max and R_1.4

These anchor the latent space to astrophysically meaningful directions.

Static Neutron Star Properties

- Tolman–Oppenheimer–Volkoff(TOV) equations:

$$\frac{dp}{dr} = -(\varepsilon + p)\frac{d\Phi}{dr}$$

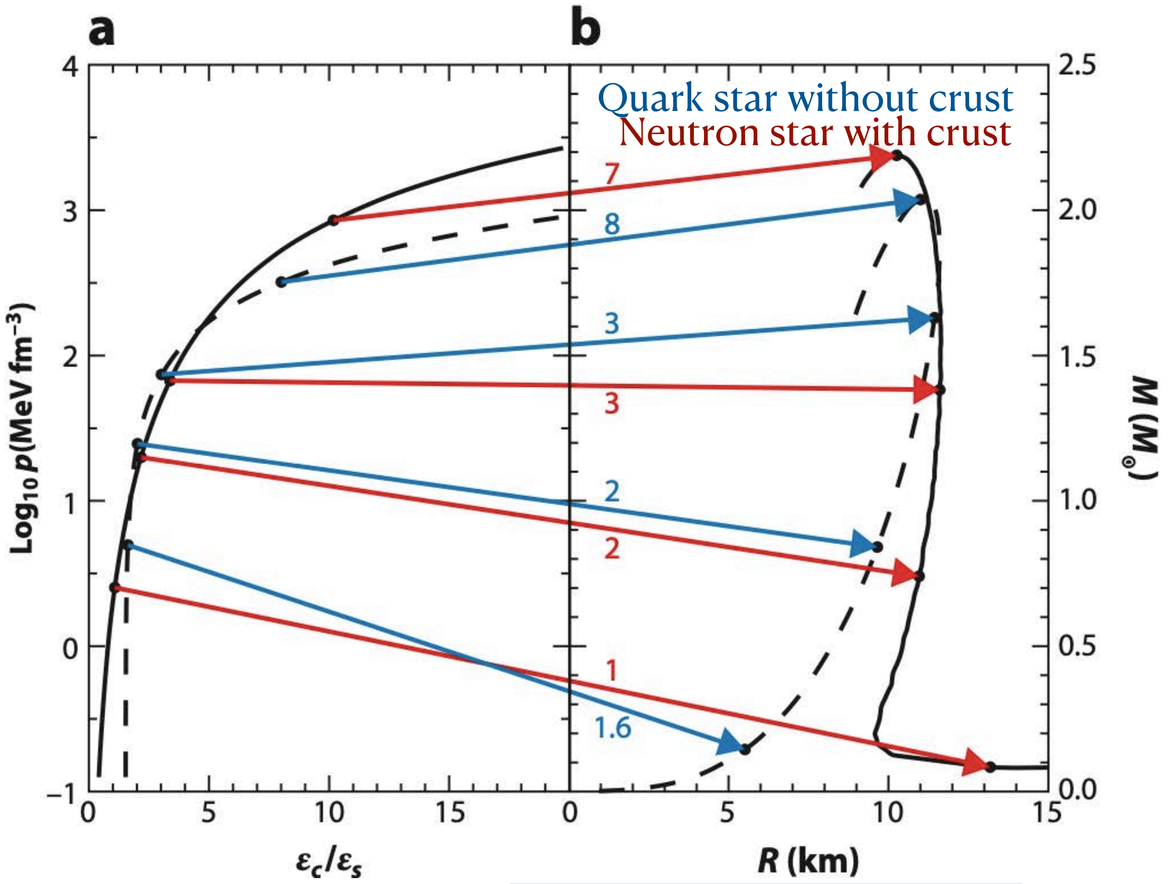
$$\frac{d\Phi}{dr} = -\frac{G(mc^2 + 4\pi r^3 p)}{r^2 c^4 (1 - \frac{2Gm}{rc^2})}$$

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

$\varepsilon(n_p, n_n, T)$ ← $E(n_B = n_p + n_n, x = n_p/n_B, T)$

$p(\varepsilon)$ the EOS → NS mass M and radius R

- Other Properties:
  Binding energy
  Moment of Inertia
  Tidal deformability
  Oscillating frequency
  …

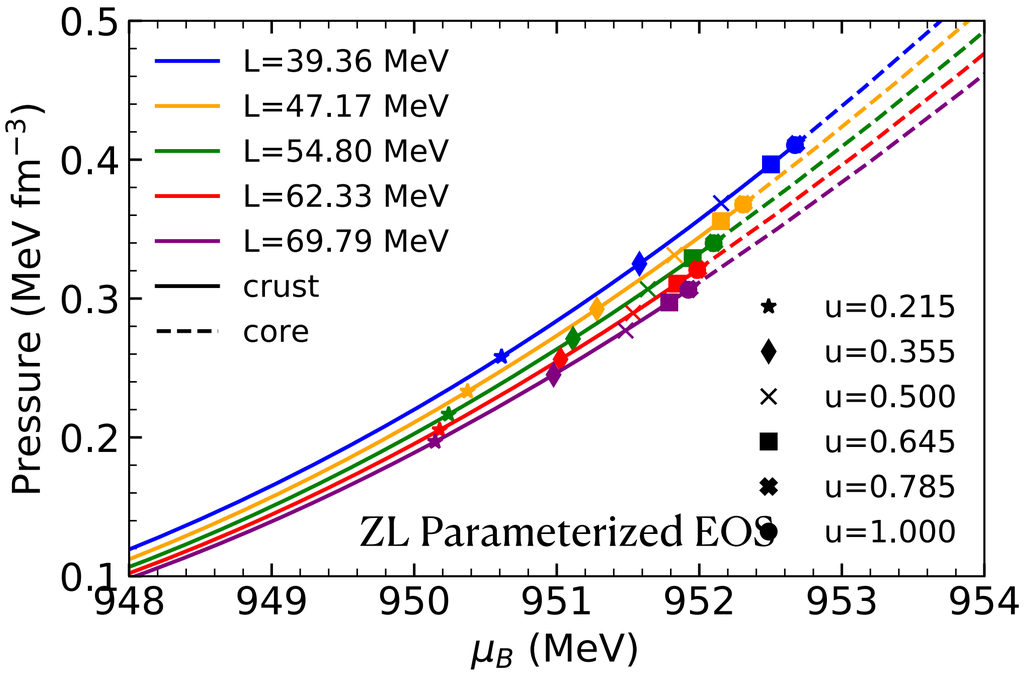


Credit: figure from Lattimer 2012

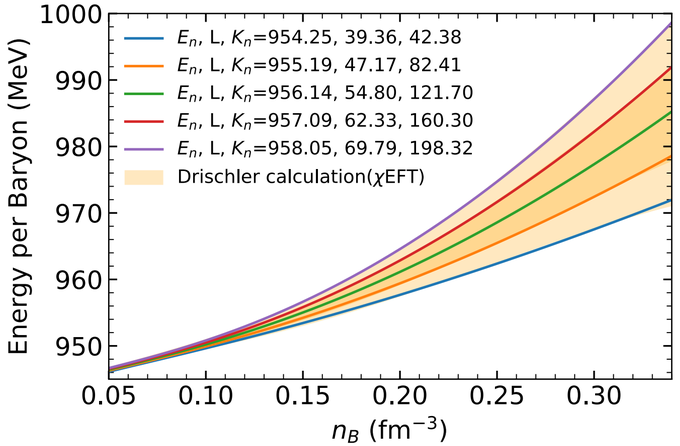
Impacts of Crust EOS and Core EOS on NS Properties

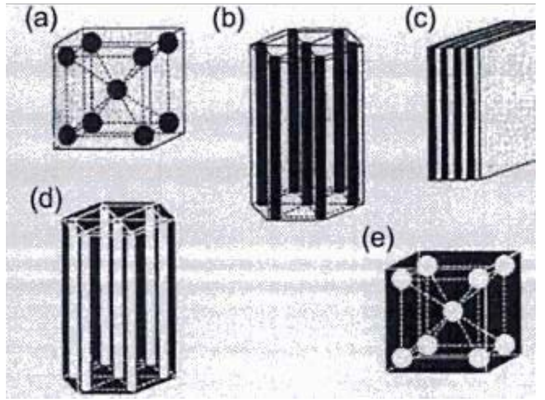
Impact on crust-core transition density

- Transition density decreases with the symmetry energy slope L.
- With the current χEFT interaction, the core-crust transition density is, $n_B^{\text{3DB-UNI}} = 0.074^{+0.004}_{-0.003}(\sigma)^{+0.010}_{-0.006}(2\sigma)\ \text{fm}^{-3}$.



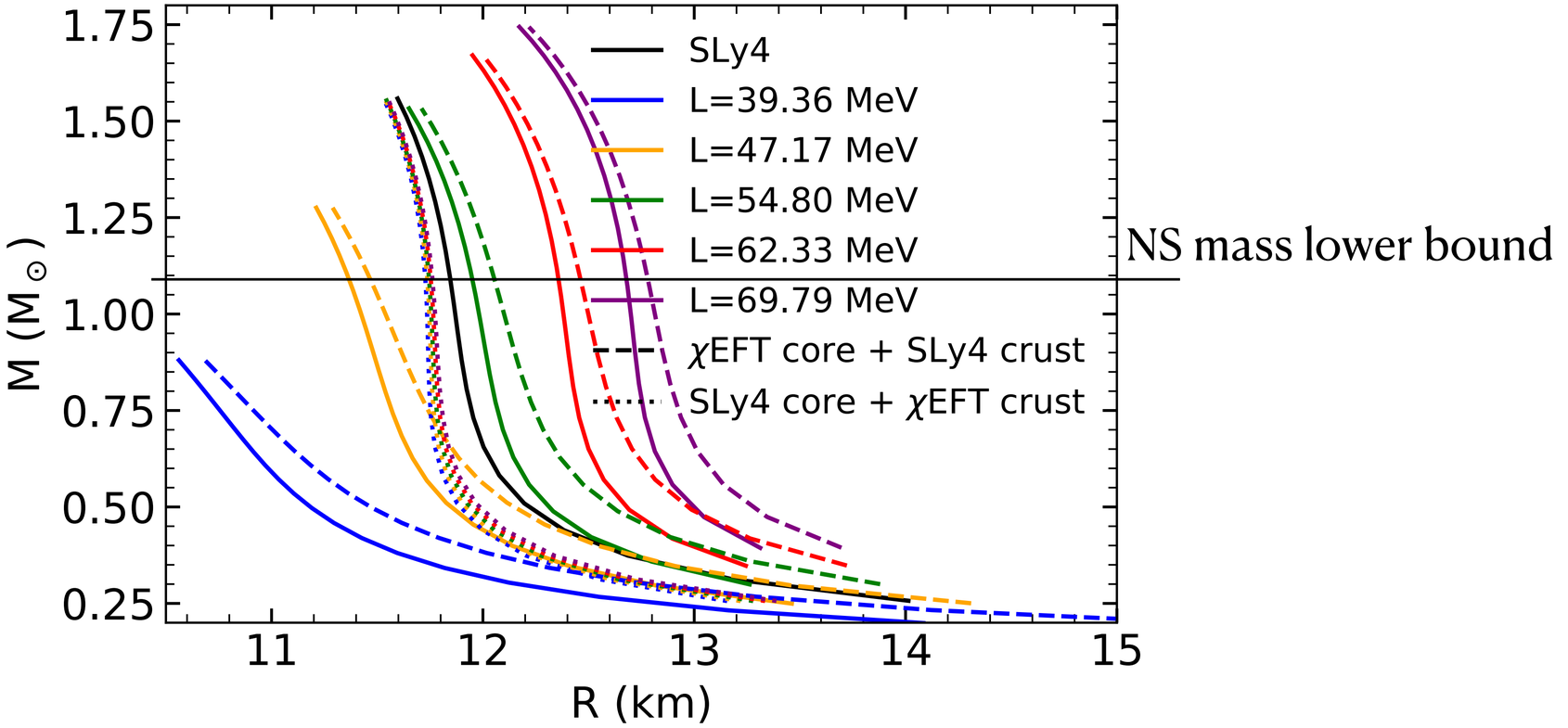
EOS at crust-core transition



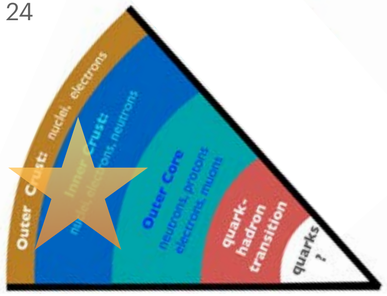


Impacts of Crust EOS and Core EOS on NS Properties

- Theoretical uncertainty in χEFT leads to an 1-2 km deviation, 30% of deviation in moment of inertia and a factor of 3 difference in tidal deformability.
- Different crust EOSs contribute to about 0.2 km deviation in radius, negligible difference in moment of inertia and tidal deformability.



Mass-radius difference with different crust EOSs



Neutron Star Crust EOS

- Outer crust EOS:
  Semi-empirical mass formula

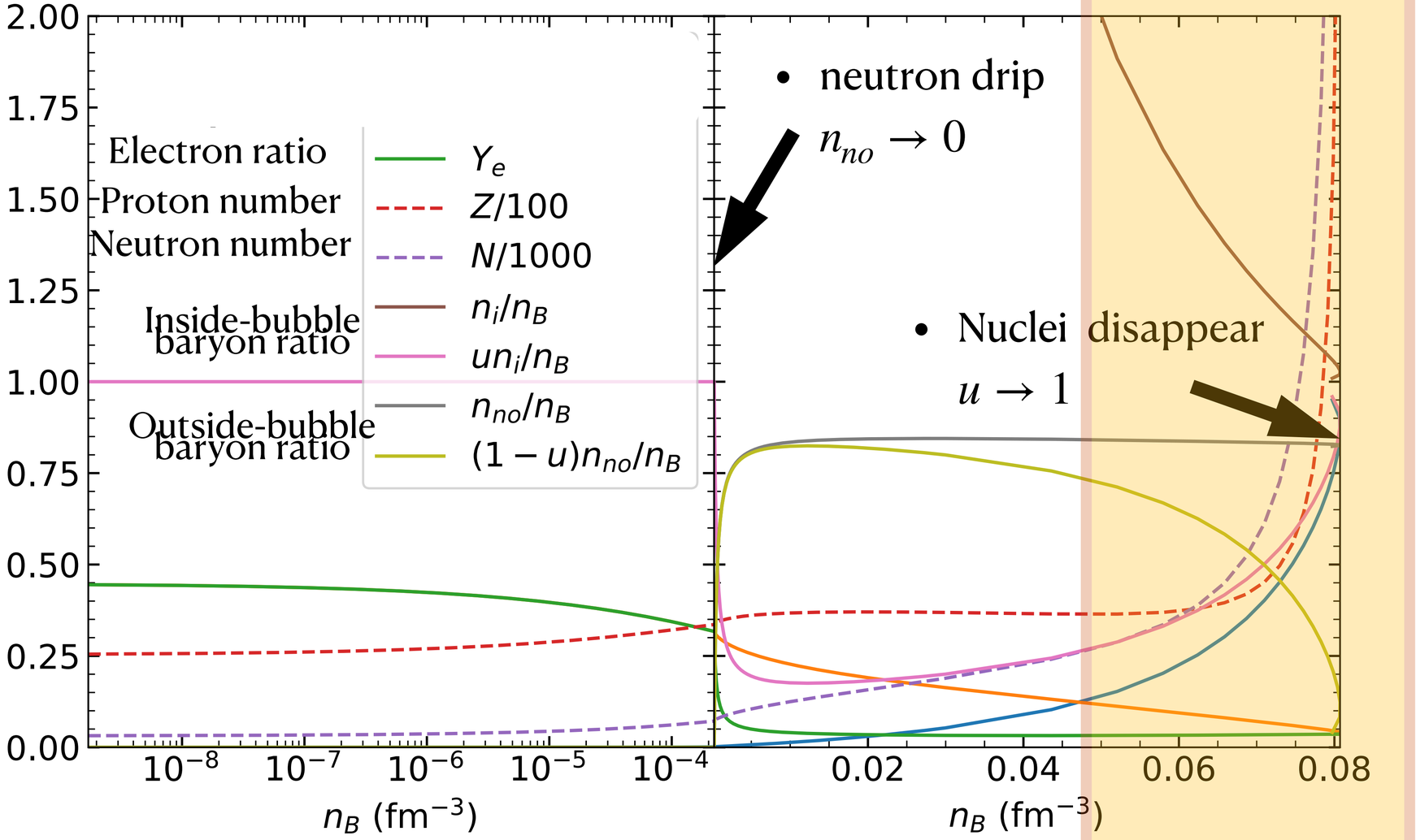
$$e^- \quad Z,\ N,\ M$$

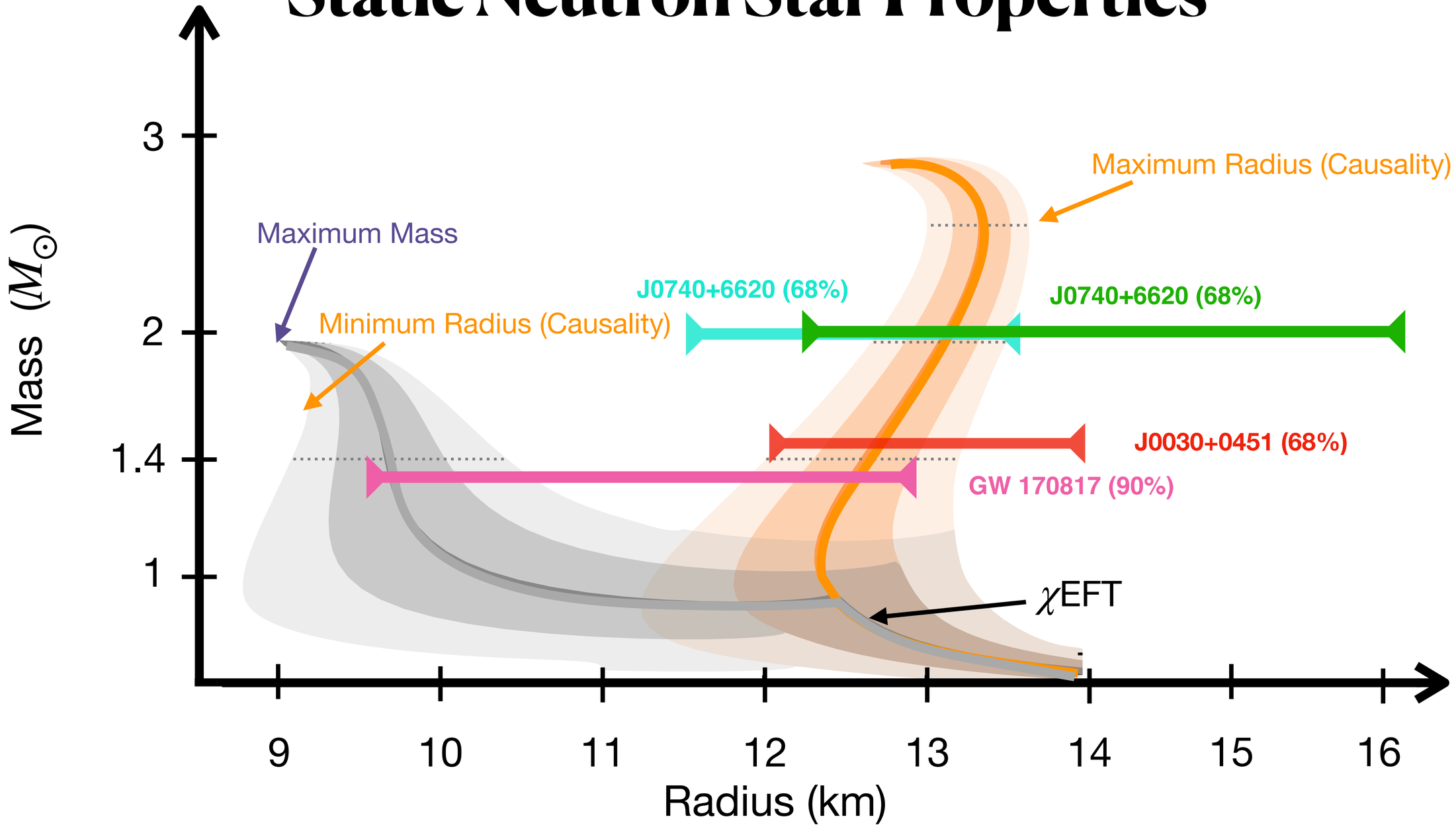
- Inner crust EOS (neutron drip):
  Compressible droplet models

$$e^- \quad Z,\ N,\ M \qquad n_o,\ P,\ \mu \quad n_i,\ P,\ \mu$$

Inter-nuclei space $1-u$ Nuclei fraction u

$$(1-u)n_o + un_i = n_B$$



Compressible droplet crust of Skyrme model SLy4

Static Neutron Star Properties