Equation-of-state of Neutron Stars from Gravitational Waves and Astrophysical Observations

and

How to Build a Self-Consistent Hierarchical Inference

Reed Essick

essick@cita.utoronto.ca

Topics

Hierarchical Bayesian Inference

- What is hierarchical Bayesian inference?
- How is it connected to Directed (Acyclic) Graphs?
- How to account for selection effects?

Neutron Star Equation of State

- Problems with phenomenological parametric EoS models
- How to
    - build (flexible) EoS models
    - condition flexible models on physics
    - discover (i.e., deal with) exotic behavior
    - extract physics without parameters

References

Hierarchical Inference

- Essick & Fishbach, ApJ (2024)
- Essick, PRD (2023)
- Essick & Holz, arXiv (2024)
- Essick & Farr, arXiv (2022)

Neutron Star Equation of State

- Landry & Essick, PRD (2019)
- Essick, Landry, & Holz, PRD (2020)
- Landry, Essick, & Chatziioannou (2020)
- Legred *et al* (including Essick), PRD 2021)
- Legred *et al* (including Essick), PRD (2022)
- Essick *et al*, PRD (2023)
- Essick *et al*, PRC (2020)
- Essick *et al*, PRL (2021)
- Essick *et al*, PRC (2021)
- Essick, ApJL (2024) *in press*

Parametric EoS Models

Inference of the NS EoS: systematics from parametric models

PRD (2022)

consider a toy model:
→ fitting a 1D function (pressure vs. energy density) without constraints

linear parameterizations

point+slope

$$p(\varepsilon) = p_a + c_s^2(\varepsilon - \varepsilon_a)$$

two-point

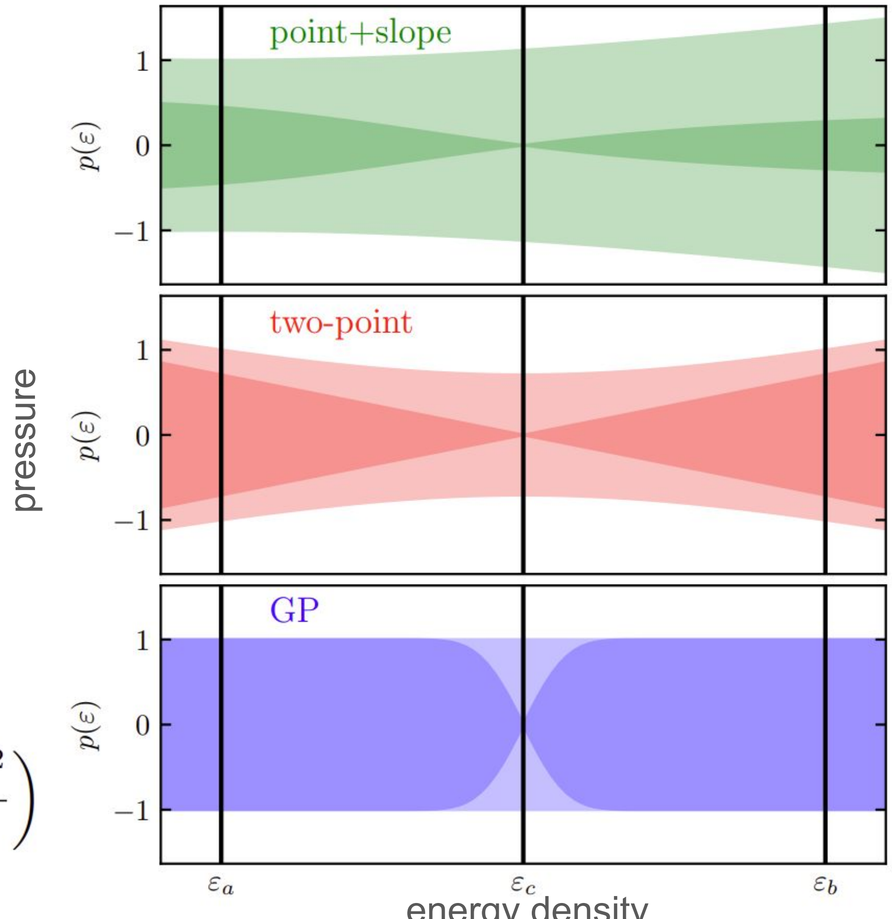
$$p(\varepsilon) = p_a + \frac{p_b - p_a}{\varepsilon_b - \varepsilon_a}(\varepsilon - \varepsilon_a)$$

Gaussian Process

$$\vec{p} \sim \mathcal{N}(\vec{\mu}, \Sigma)$$

$$\Sigma_{ij} = \mathrm{Cov}(p_i, p_j)$$

$$= K_{\mathrm{se}}(\varepsilon_i, \varepsilon_j) = \sigma^2 \exp\left(-\frac{(\varepsilon_i - \varepsilon_j)^2}{l^2}\right)$$



Inference of the NS EoS: systematics from parametric models

PRD (2022)

consider a toy model:
→ fitting a 1D function (pressure vs. energy density) without constraints

point+slope

$$p(\varepsilon) = p_a + c_s^2(\varepsilon - \varepsilon_a)$$

two-point

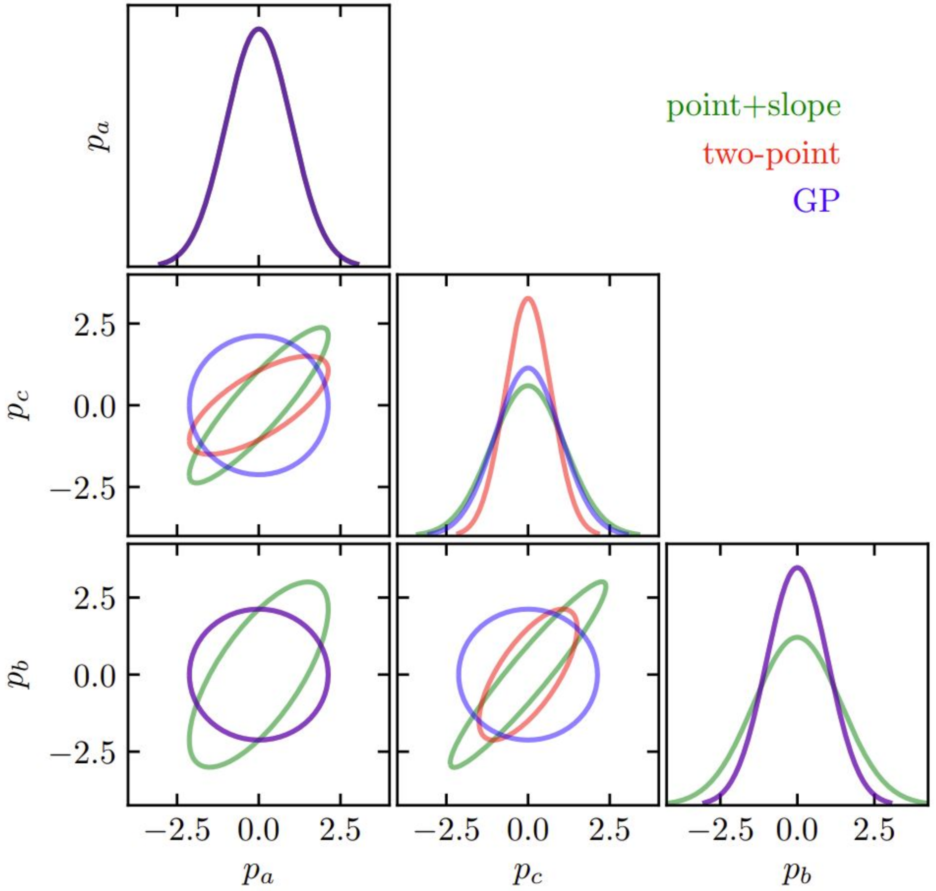
$$p(\varepsilon) = p_a + \frac{p_b - p_a}{\varepsilon_b - \varepsilon_a}(\varepsilon - \varepsilon_a)$$

GP

$$\vec{p} \sim \mathcal{N}(\vec{\mu}, \Sigma)$$

$$\Sigma_{ij} = \mathrm{Cov}(p_i, p_j)$$

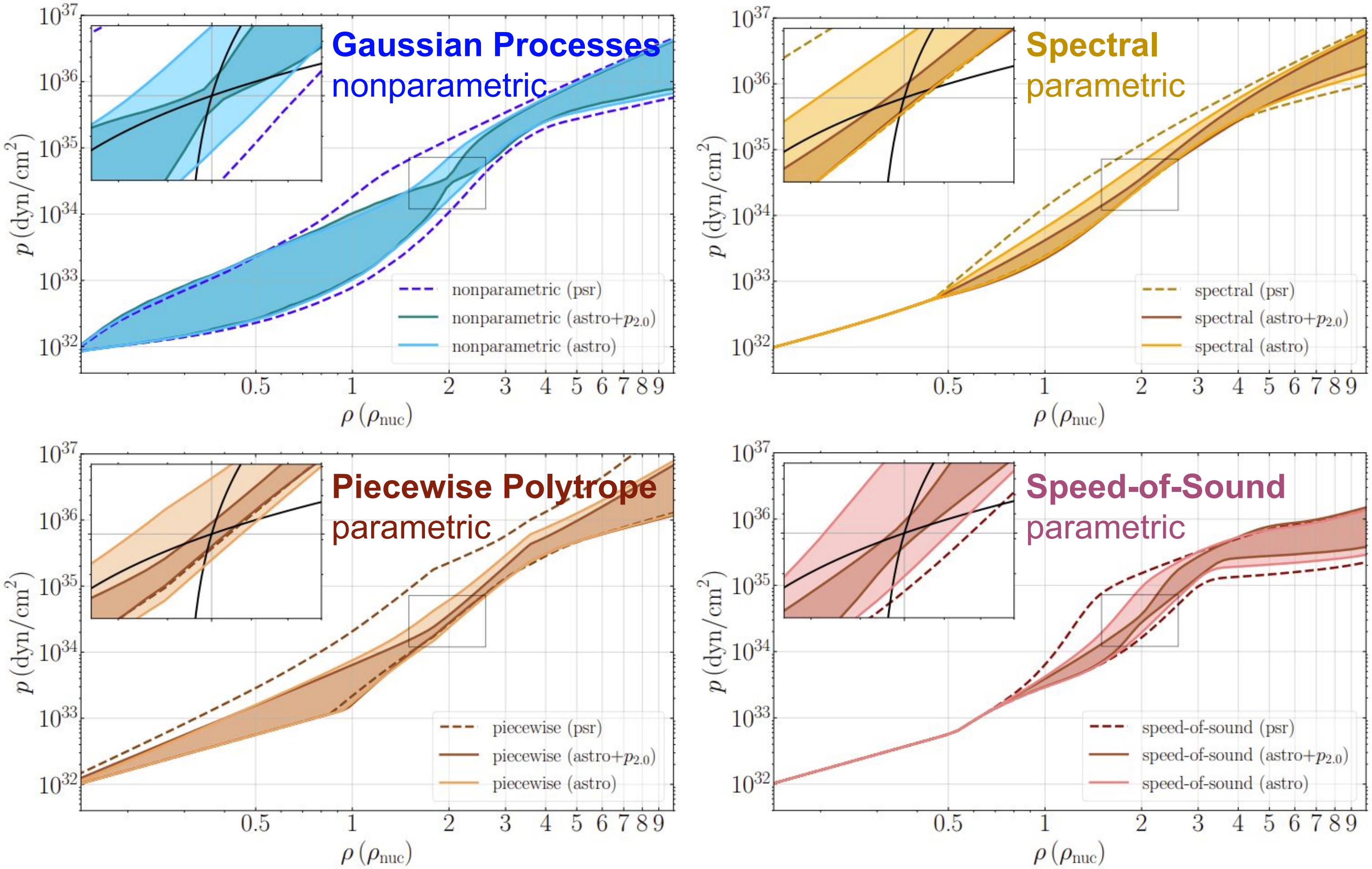
$$= K_{\mathrm{se}}(\varepsilon_i, \varepsilon_j) = \sigma^2 \exp\left(-\frac{(\varepsilon_i - \varepsilon_j)^2}{l^2}\right)$$



only the **GP** has ***independent marginal distributions*** for all pressures

Inference of the NS EoS: systematics from parametric models

PRD (2022)



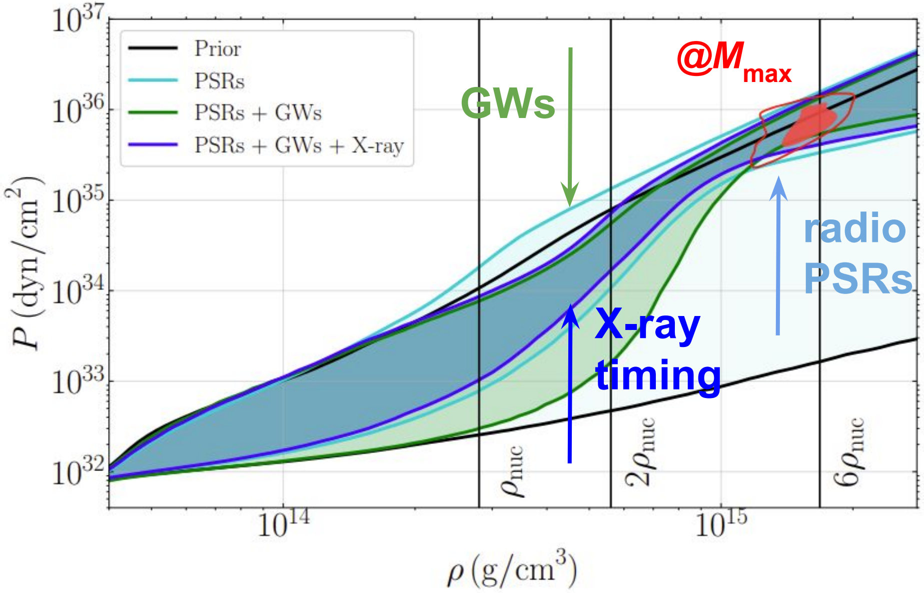
Non-Parametric EoS Models

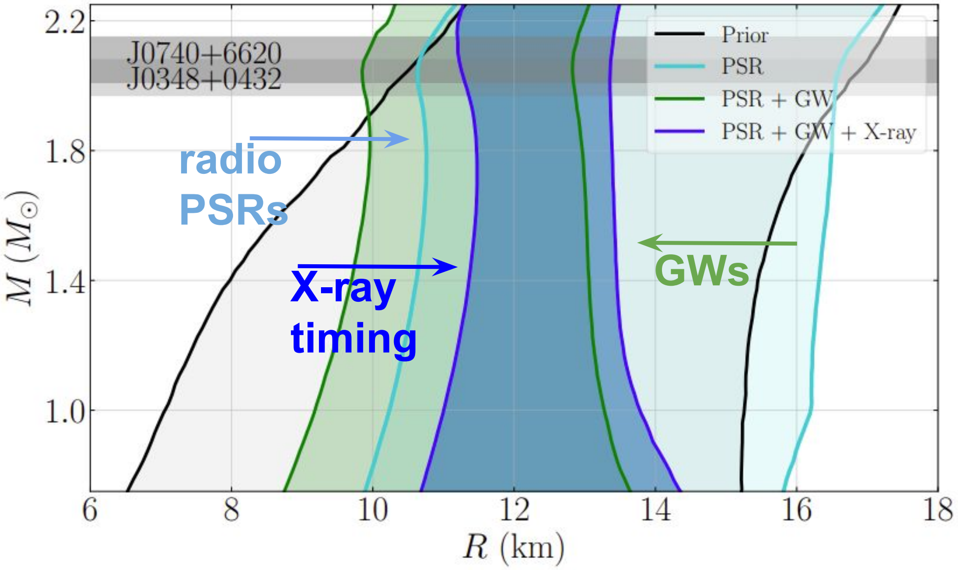
Inference of the NS EoS: nonparametric results

PRD (2021)

Current *Theory Agnostic* Constraints

phase transitions





Inference of the NS EoS: nonparametric results

PRD (2021)

	Observable	Prior	w/ PSRs	w/o J0740+6620	w/J0740+6620 Miller+	w/J0740+6620 Riley+
Properties of the EoS	$M_{\max}$ [M$_\odot$]	$1.47^{+0.71}_{-1.37}$	$2.24^{+0.48}_{-0.24}$	$2.20^{+0.30}_{-0.19}$	$2.21^{+0.31}_{-0.21}$	$2.19^{+0.27}_{-0.19}$
	$p(\rho_{\rm nuc})$ [10^{33}dyn/cm^2]	$2.25^{+5.81}_{-2.15}$	$6.07^{+7.53}_{-5.93}$	$4.05^{+3.59}_{-3.74}$	$4.30^{+3.37}_{-3.80}$	$4.15^{+3.50}_{-3.76}$
	$p(2\rho_{\rm nuc})$ [10^{34}dyn/cm^2]	$1.22^{+4.86}_{-1.21}$	$6.00^{+4.79}_{-5.99}$	$3.75^{+2.36}_{-2.98}$	$4.38^{+2.46}_{-2.96}$	$3.90^{+2.11}_{-2.88}$
	$p(6\rho_{\rm nuc})$ [10^{35}dyn/cm^2]	$2.43^{+4.70}_{-2.43}$	$7.51^{+6.77}_{-5.15}$	$8.33^{+5.22}_{-4.14}$	$7.41^{+5.87}_{-4.18}$	$7.82^{+5.47}_{-3.53}$
	$\max\{c_s^2/c^2\} \mid \rho \leq \rho_c(M_{\max})$	$0.76^{+0.24}_{-0.37}$	$0.72^{+0.28}_{-0.26}$	$0.84^{+0.16}_{-0.28}$	$0.75^{+0.25}_{-0.24}$	$0.80^{+0.20}_{-0.26}$
	$\rho(\max\{c_s^2/c^2\})$ [10^{15}g/cm^3]	$1.38^{+1.65}_{-1.34}$	$0.97^{+0.64}_{-0.70}$	$1.13^{+0.64}_{-0.63}$	$1.01^{+0.63}_{-0.53}$	$1.10^{+0.63}_{-0.58}$
	$p(\max\{c_s^2/c^2\})$ [10^{35}dyn/cm^2]	$1.65^{+8.16}_{-1.65}$	$2.68^{+5.18}_{-2.68}$	$3.52^{+6.90}_{-3.48}$	$2.77^{+5.81}_{-2.70}$	$3.26^{+6.51}_{-3.15}$
Properties defined for both NSs and BHs	$R_{1.4}$ [km]	$8.09^{+5.68}_{-3.96}$	$13.54^{+2.61}_{-3.13}$	$12.25^{+1.13}_{-1.33}$	$12.56^{+1.00}_{-1.07}$	$12.34^{+1.01}_{-1.25}$
	$R_{2.0}$ [km]	$5.90^{+6.97}_{-0.00}$	$13.18^{+3.02}_{-2.90}$	$12.05^{+1.18}_{-1.45}$	$12.41^{+1.00}_{-1.10}$	$12.09^{+1.07}_{-1.17}$
	$\Delta R \equiv R_{2.0} - R_{1.4}$ [km]	$0.48^{+1.28}_{-6.67}$	$-0.07^{+1.00}_{-1.04}$	$-0.17^{+0.85}_{-0.83}$	$-0.12^{+0.83}_{-0.85}$	$-0.20^{+0.82}_{-0.88}$
	$\Lambda_{1.4}$	24^{+841}_{-24}	795^{+1262}_{-708}	442^{+235}_{-274}	507^{+234}_{-242}	457^{+219}_{-256}
	$\Lambda_{2.0}$	0^{+54}_{-0}	66^{+184}_{-66}	34^{+35}_{-27}	44^{+34}_{-30}	35^{+32}_{-24}
Properties defined only for NSs	$\rho_c(1.4\,{\rm M}_\odot)$ [10^{14}g/cm^3]	$8.4^{+12.5}_{-6.0}$	$5.7^{+3.2}_{-3.1}$	$7.2^{+2.6}_{-1.7}$	$6.7^{+1.7}_{-1.3}$	$7.1^{+2.1}_{-1.5}$
	$\rho_c(2.0\,{\rm M}_\odot)$ [10^{14}g/cm^3]	$9.0^{+5.7}_{-6.3}$	$8.5^{+4.8}_{-5.3}$	$10.5^{+4.1}_{-3.8}$	$9.7^{+3.6}_{-3.1}$	$10.4^{+3.6}_{-3.5}$
	$\rho_c(M_{\max})$ [10^{15}g/cm^3]	$2.4^{+0.9}_{-2.0}$	$1.4^{+0.5}_{-0.6}$	$1.6^{+0.3}_{-0.4}$	$1.5^{+0.3}_{-0.4}$	$1.6^{+0.3}_{-0.3}$

$M_{max} \sim 2.21 \pm 0.25\, M_\odot$

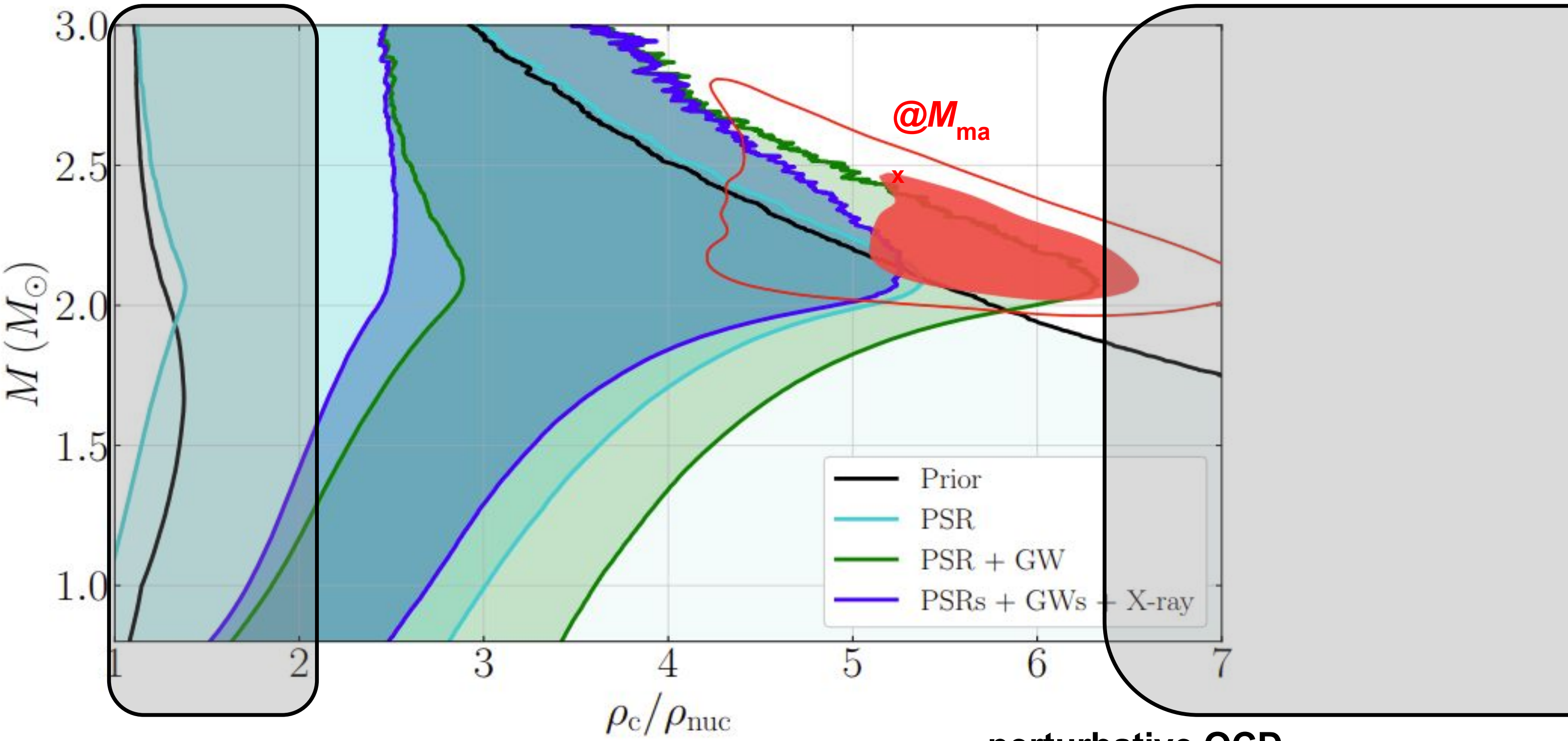
$R(1.4M_\odot) \sim 12.5 \pm 1$ km

at 90% credibility

Future Prospects: EoS constraints

PRD (2021)

nuclear experiment



perturbative QCD

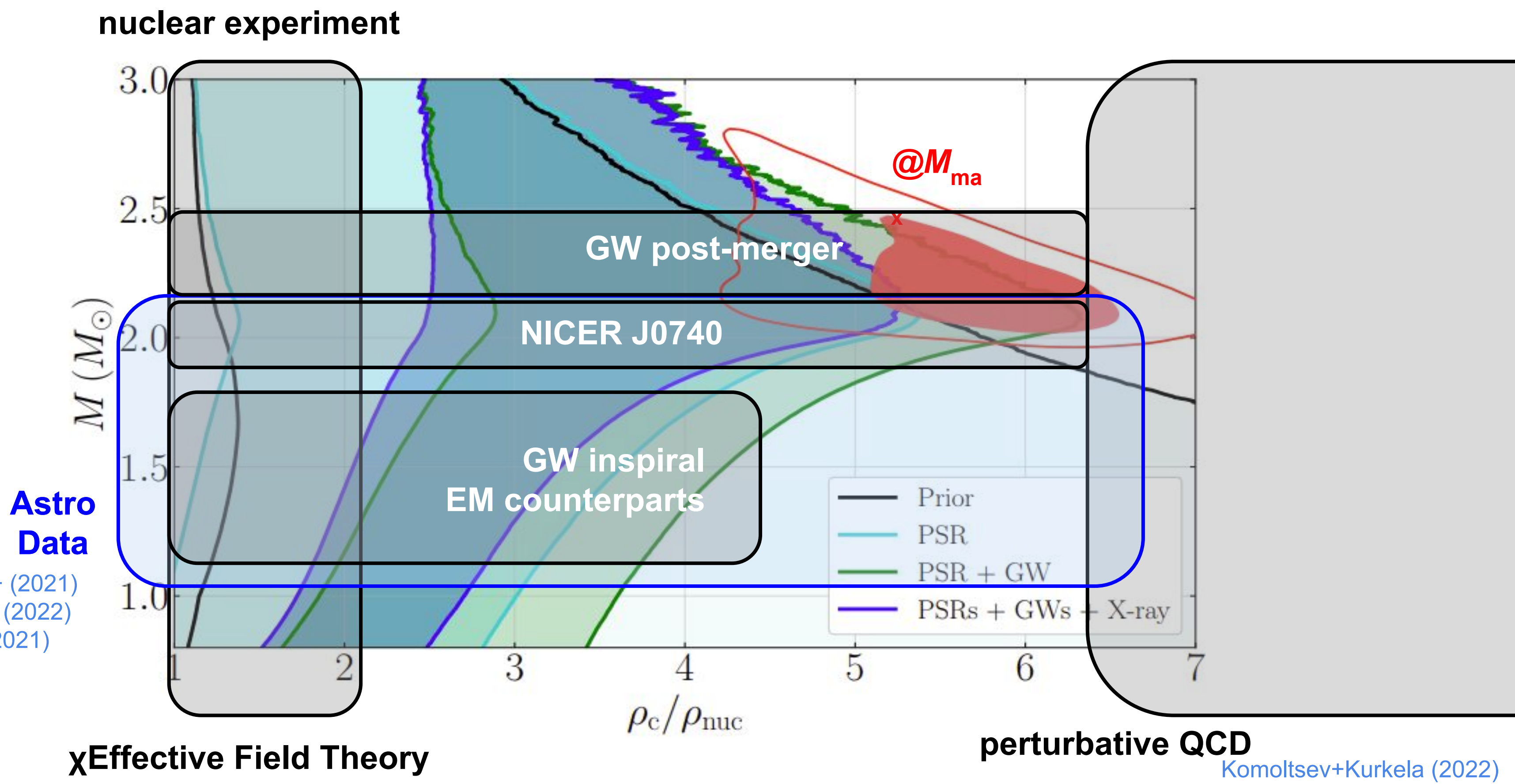
Komoltsev+Kurkela (2022)
Gorda+ (2022)
Komoltsev+ (2023)

χEffective Field Theory

Lynn+ (2016)
Drischler+ (2020), Drischler+ (2020)

Future Prospects: EoS constraints

PRD (2021)



Abbott+**Essick**+ (2021)
Farah+**Essick**+ (2022)
Landry+Read (2021)

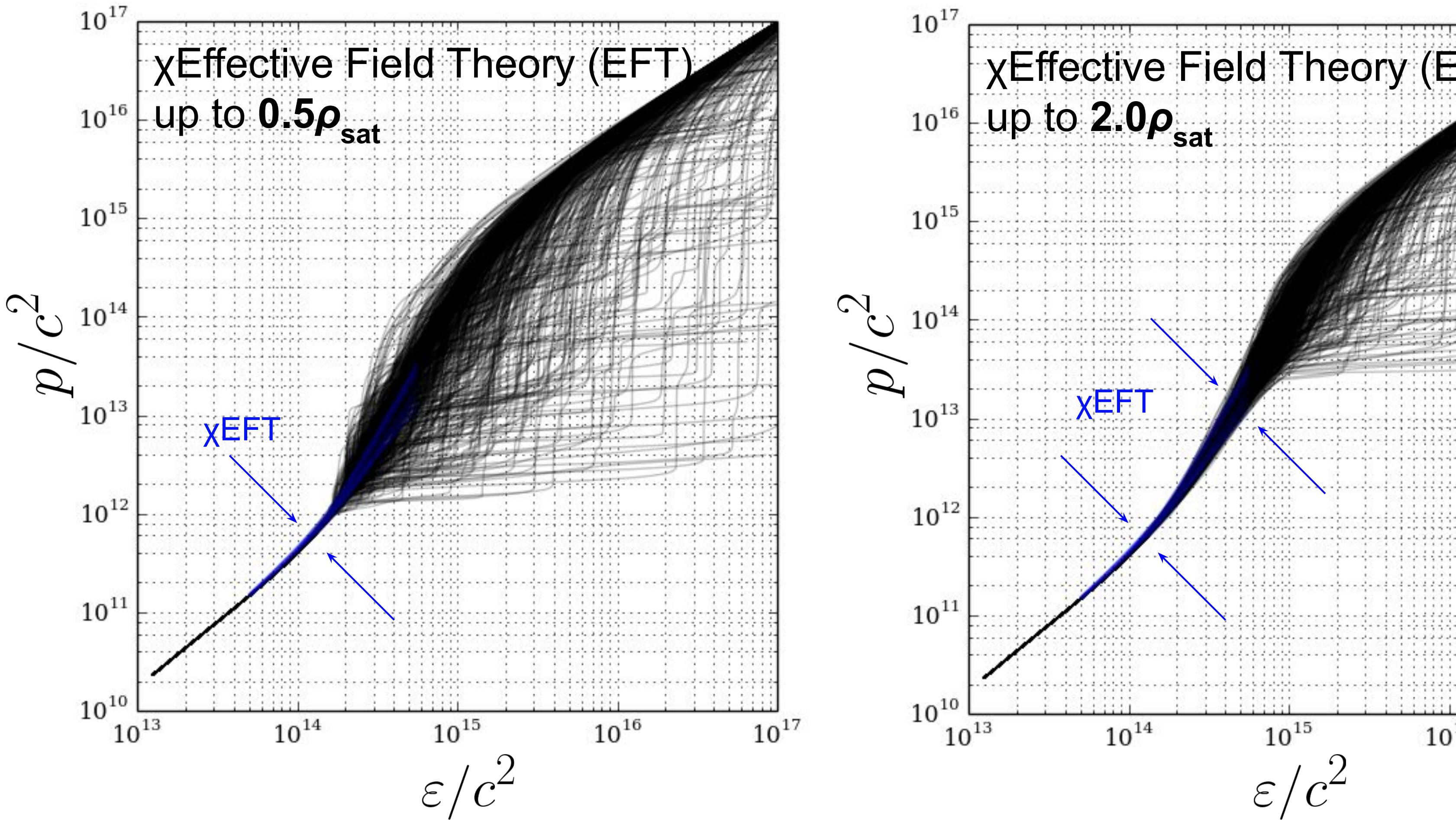
Lynn+ (2016)
Drischler+ (2020), Drischler+ (2020)

Komoltsev+Kurkela (2022)
Gorda+ (2022)
Komoltsev+ (2023)

Non-Parametric theory-informed EoS Models

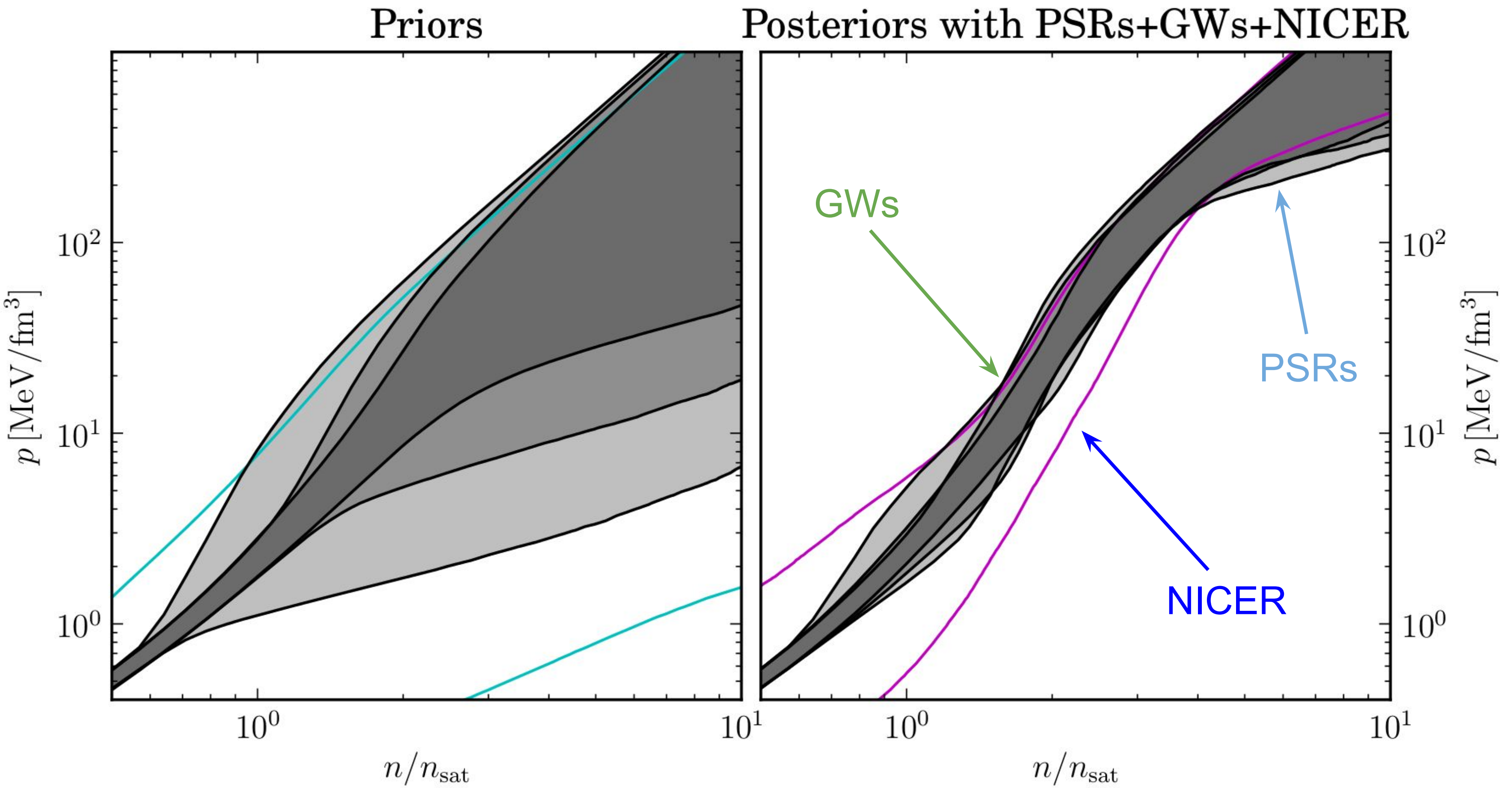
Inference of the NS EoS: incorporating low-density nuclear theory

PRC (2020)



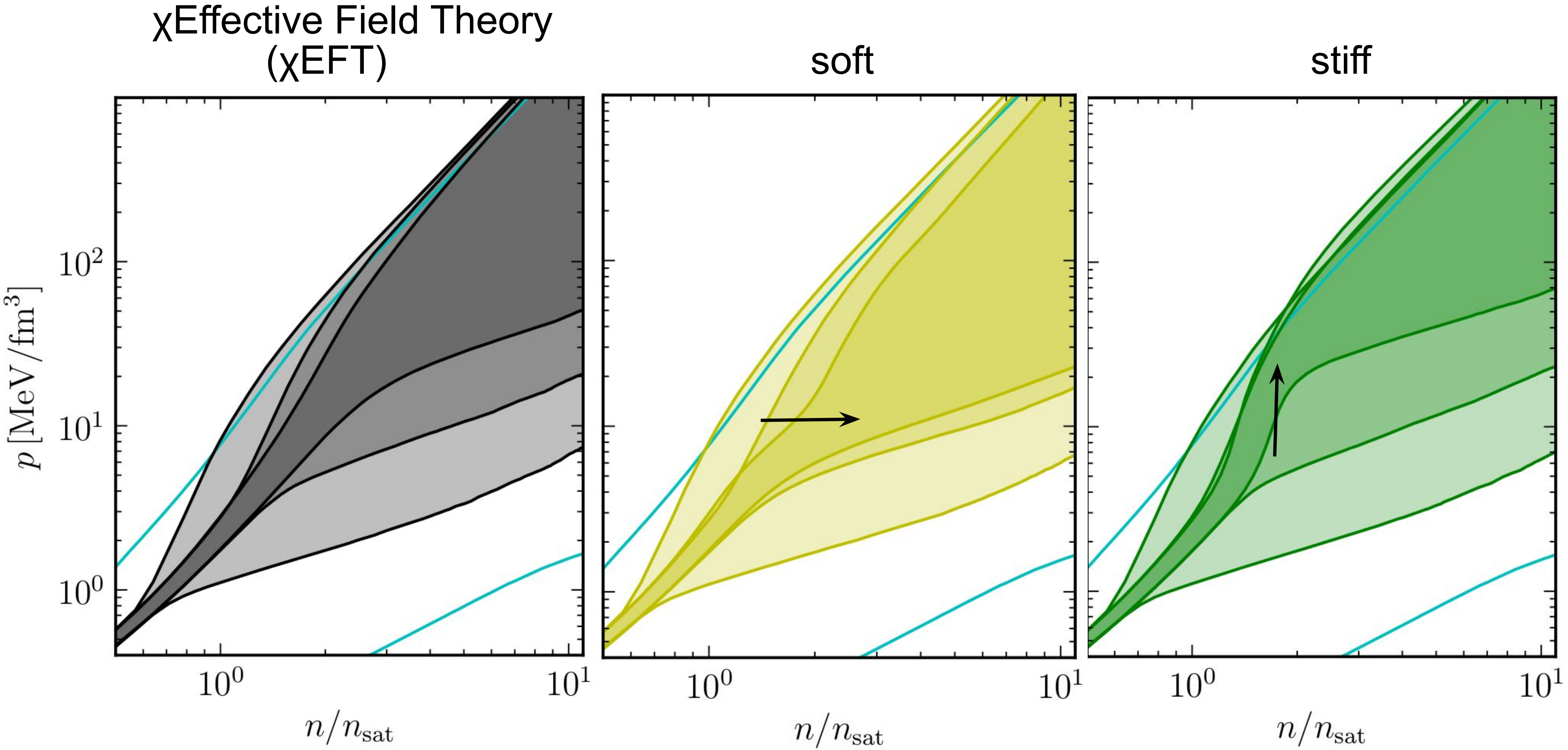
Inference of the NS EoS: incorporating low-density nuclear theory

PRC (2020)



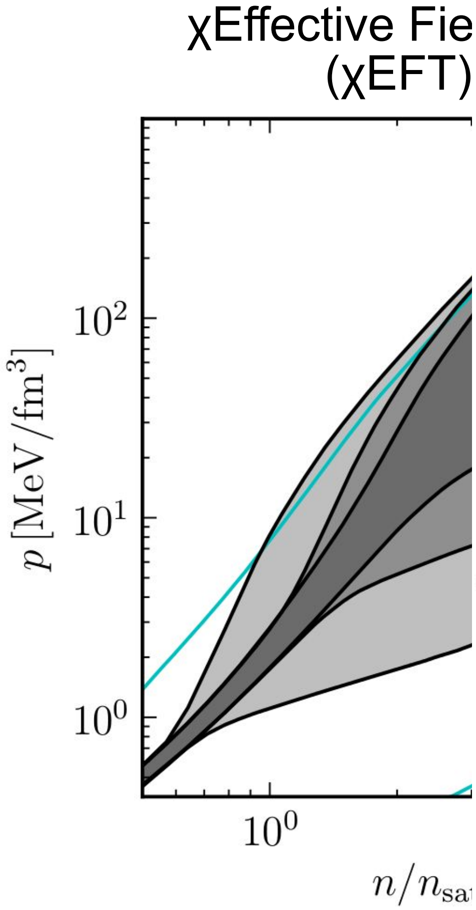
Inference of the NS EoS: comparing low-density theories

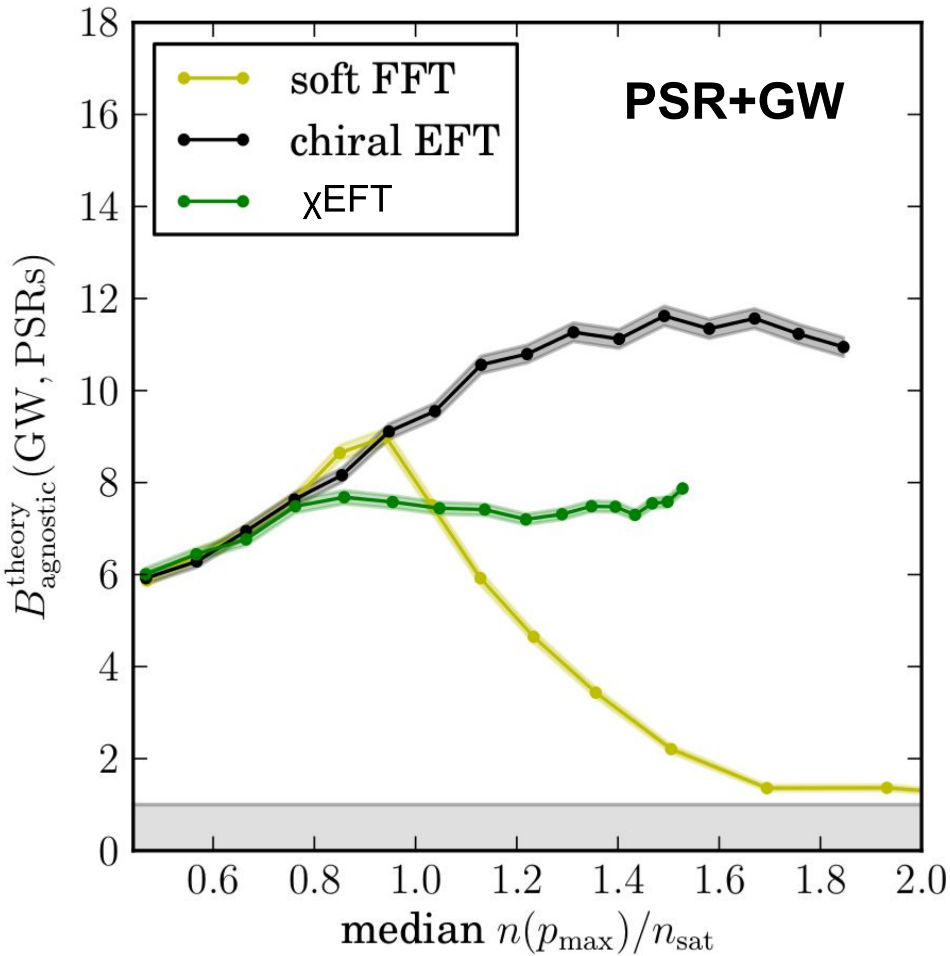
PRC (2020)

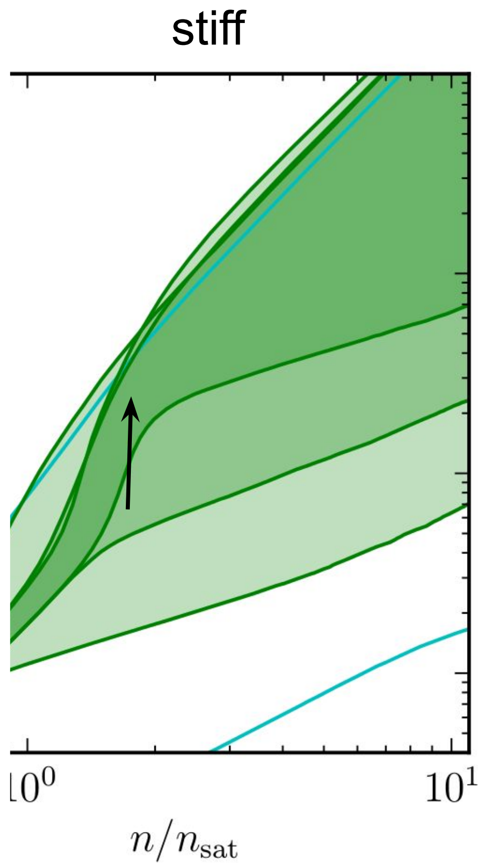


Inference of the NS EoS: comparing low-density theories

PRC (2020)



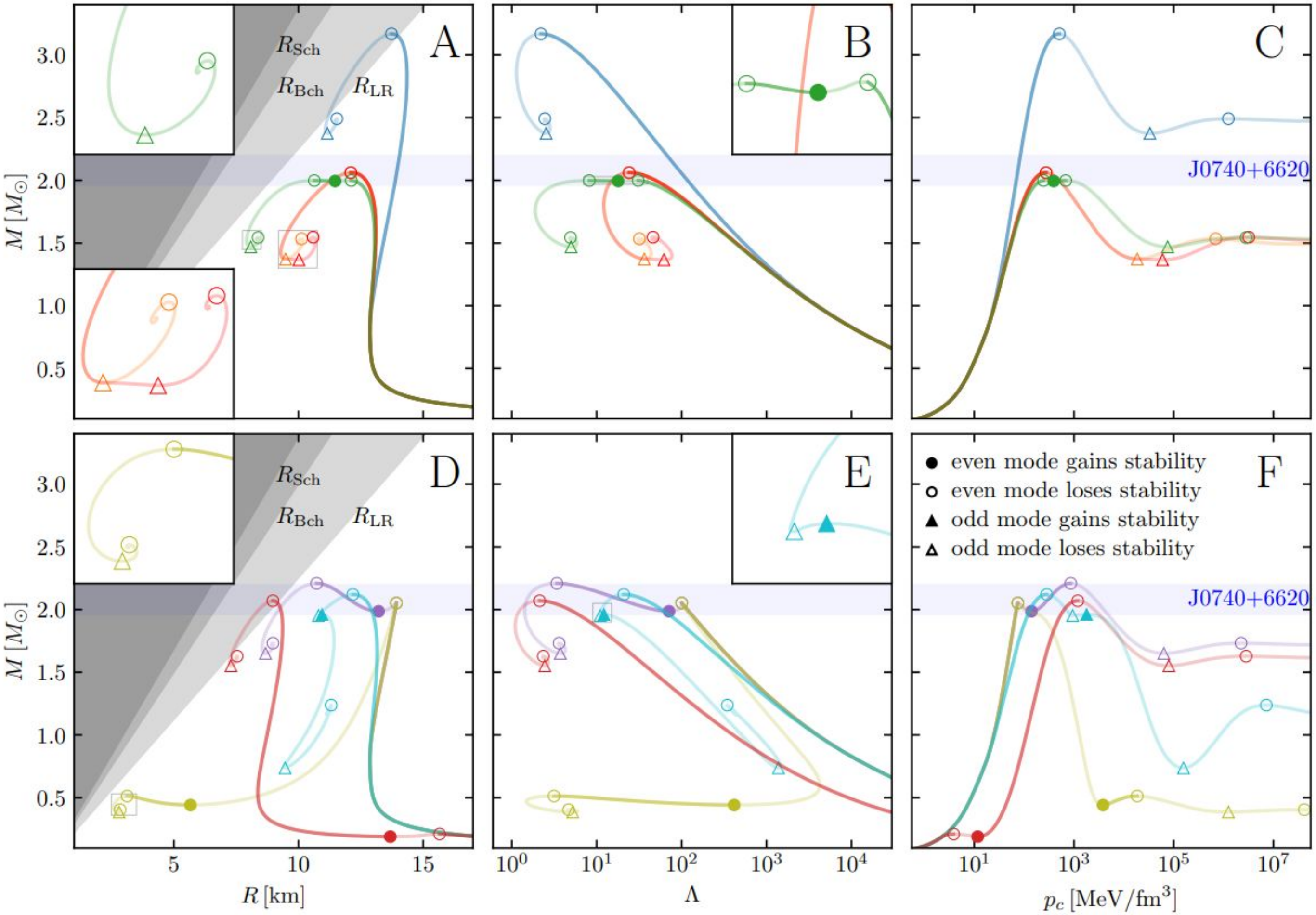




Exotic TOV Sequences

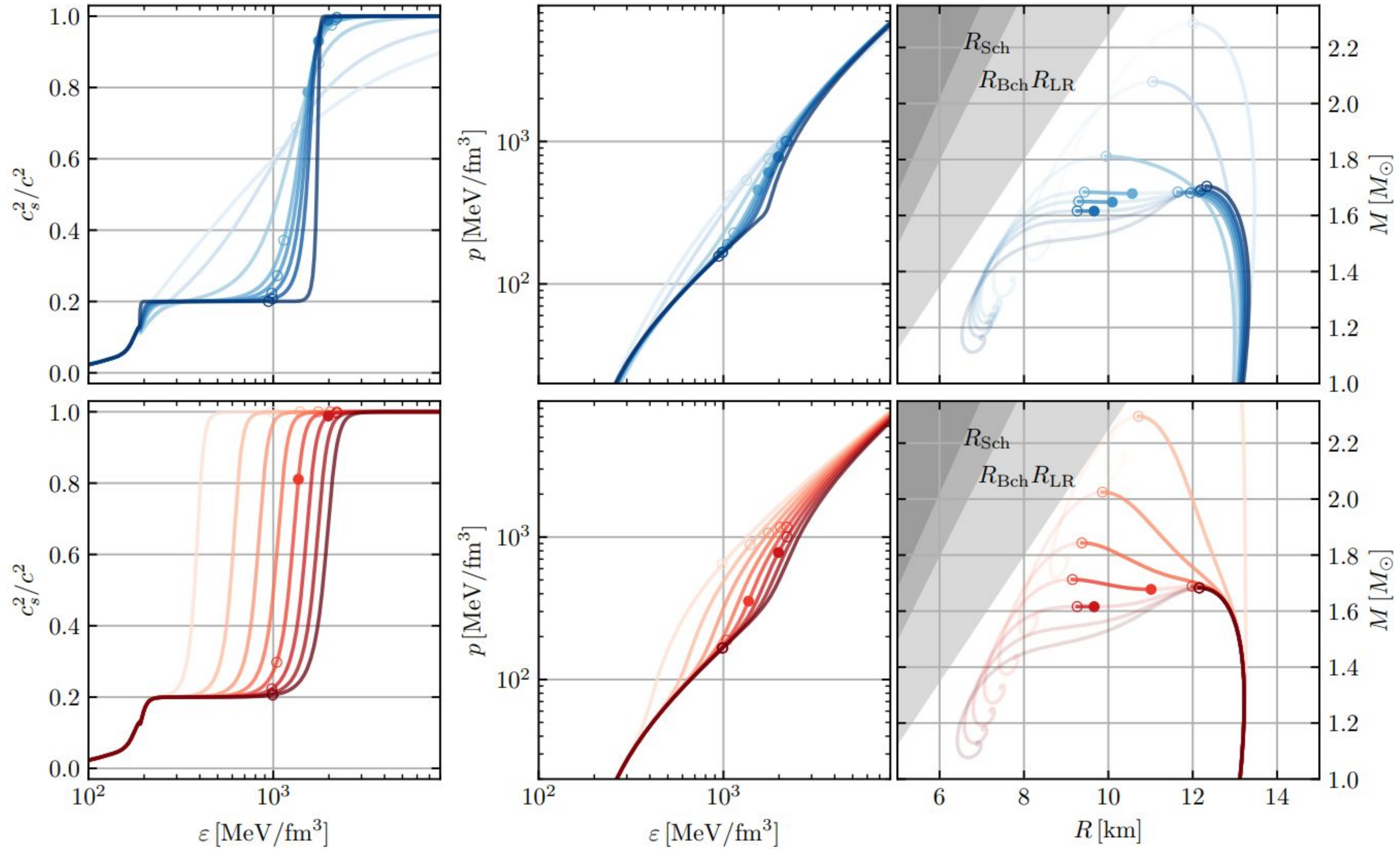
Exotic Behavior with Efficient TOV Sequences

ApJL (2024) *in press*



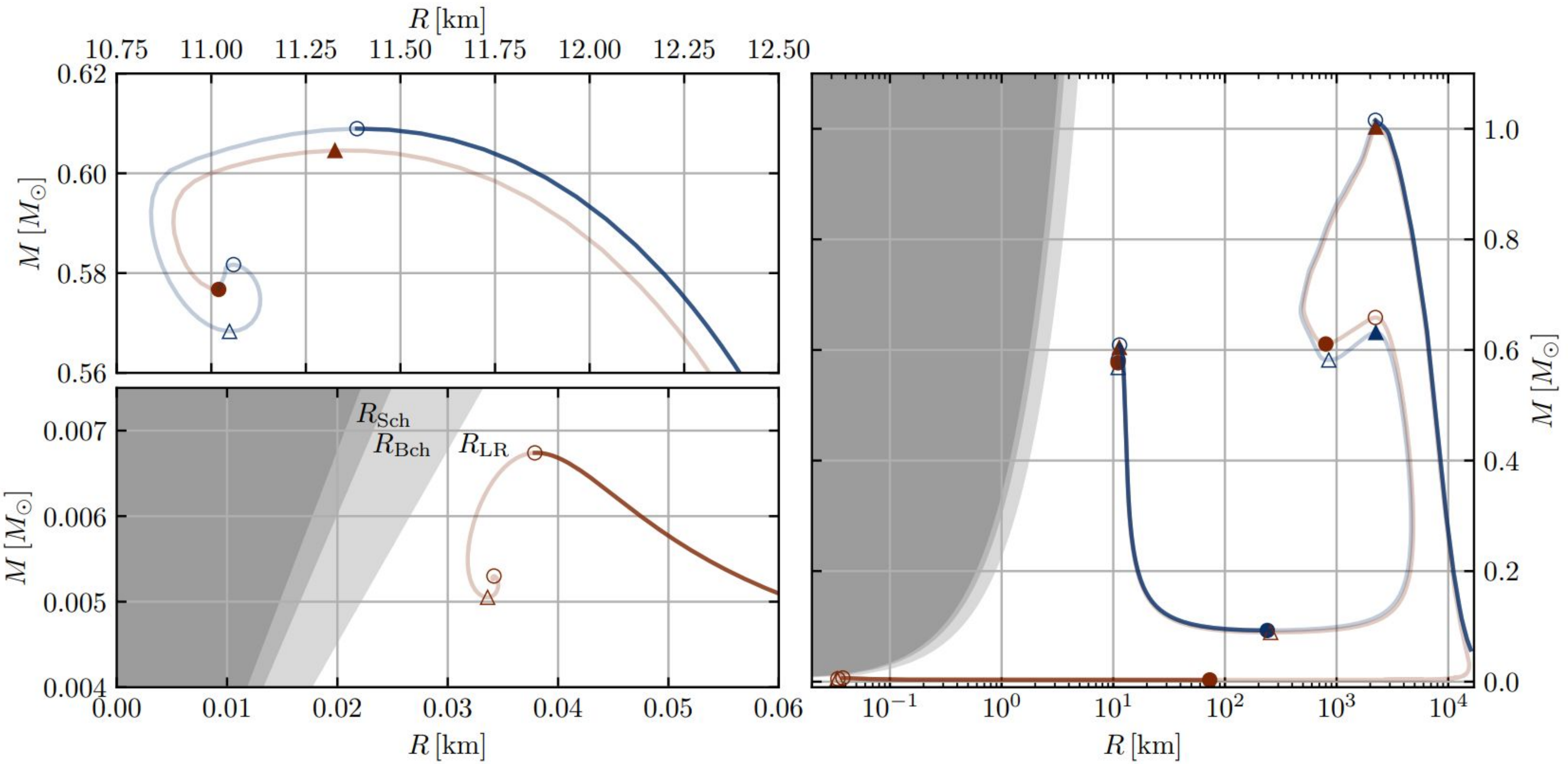
Exotic Behavior with Efficient TOV Sequences

ApJL (2024) *in press*



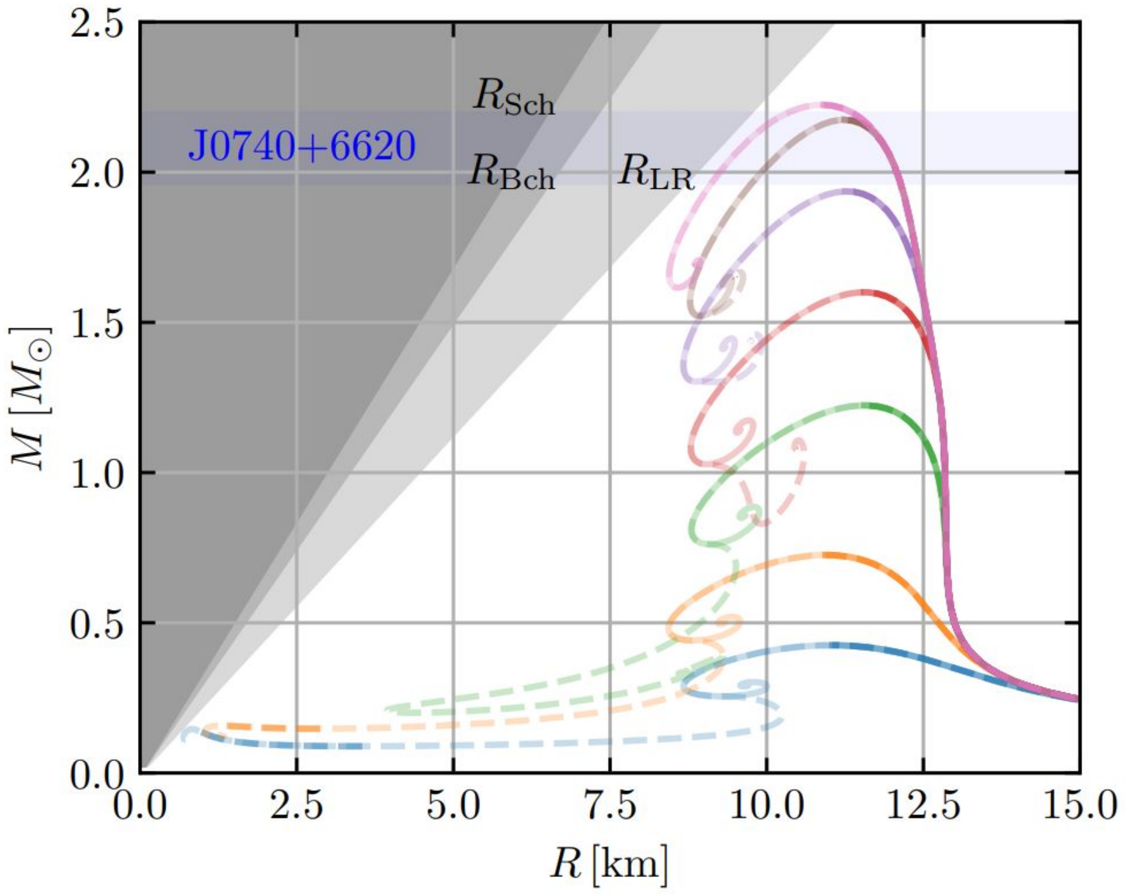
Exotic Behavior with Efficient TOV Sequences

ApJL (2024) *in press*



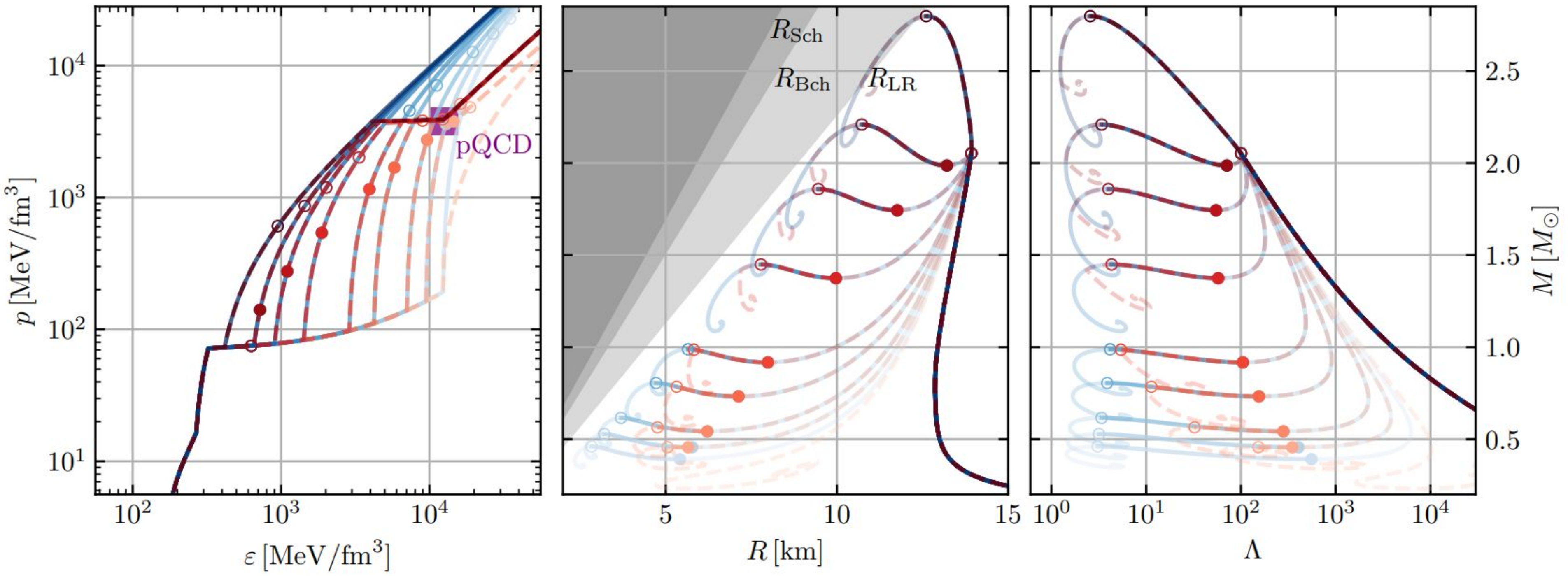
Exotic Behavior with Efficient TOV Sequences

ApJL (2024) *in press*



Exotic Behavior with Efficient TOV Sequences

ApJL (2024) *in press*

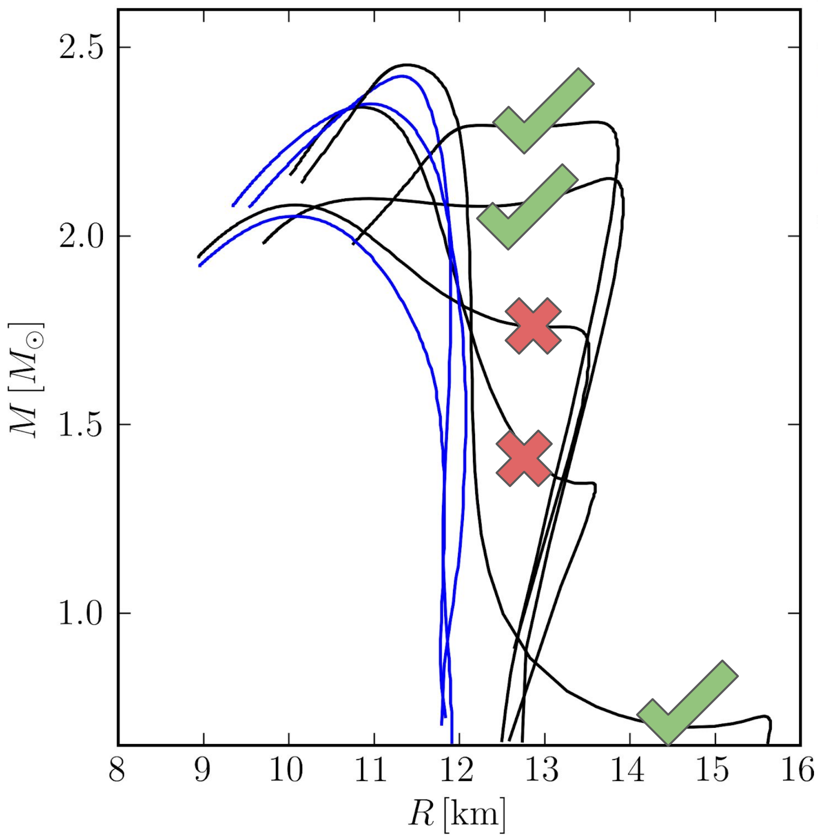


Extracting Physics without Parameters

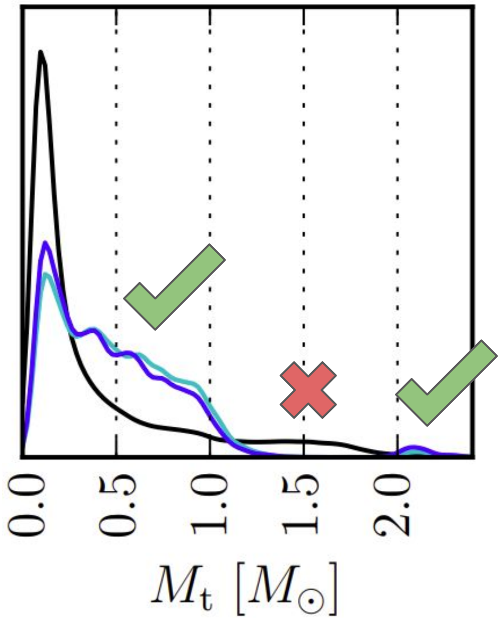
Phase Transitions

Inference of the NS EoS: phase transitions

PRD (2021)



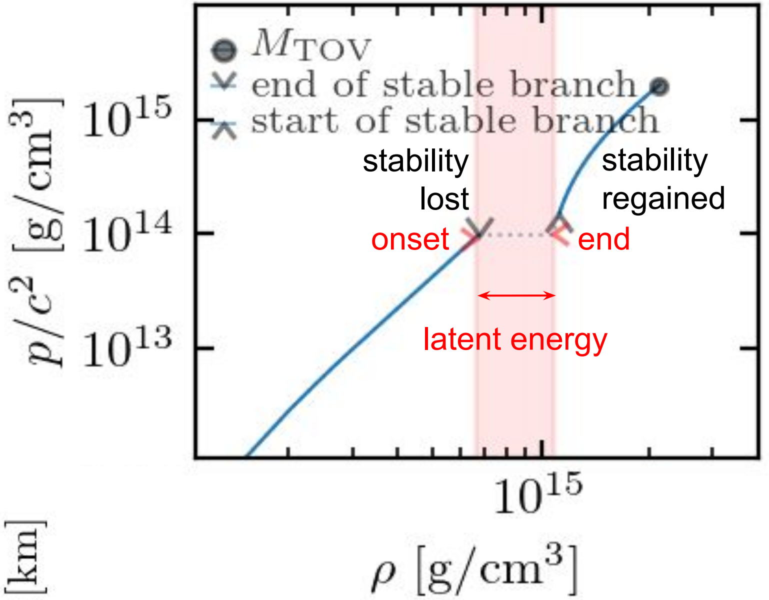
Data		$\max \mathcal{L}^{n>1}_{n=1}$	$\mathcal{B}^{n>1}_{n=1}$	$\max \mathcal{L}^{c_s^2>c^2/3}_{c_s^2\leq c^2/3}$	$\mathcal{B}^{c_s^2>c^2/3}_{c_s^2\leq c^2/3}$
w/PSRs		1.00	0.120 ± 0.002	1.0	10.2 ± 0.5
w/o J0740+6620		0.97	0.220 ± 0.007	50.8	2220 ± 790
w/J0740+6620	Miller+	0.60	0.146 ± 0.005	26.7	1000 ± 340
	Riley+	0.94	0.185 ± 0.006	72.7	2450 ± 1820

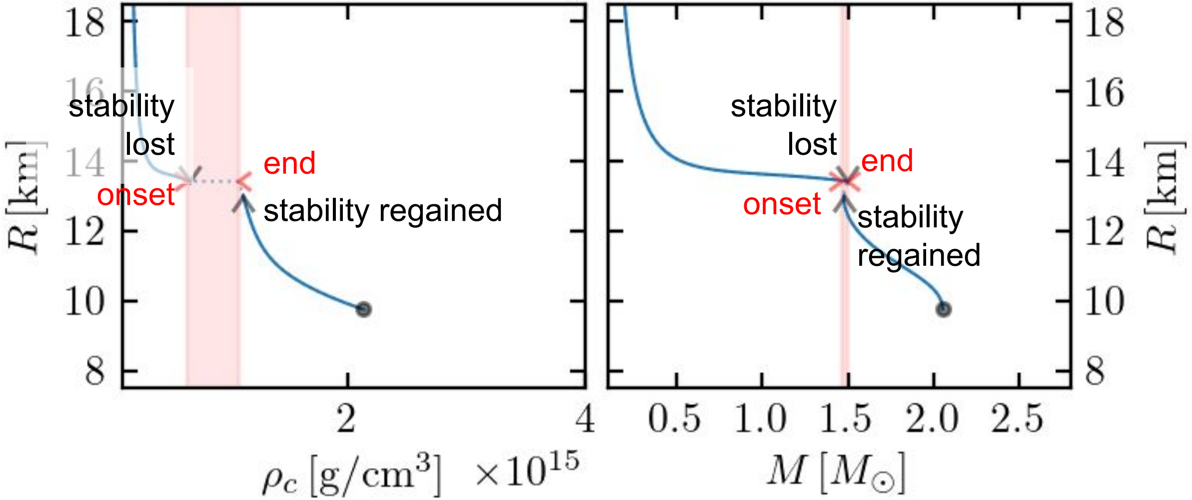


Inference of the NS EoS: phase transitions

the sudden appearance of new (degenerate) degrees of freedom produces a sharp drop in the sound speed (c_s)

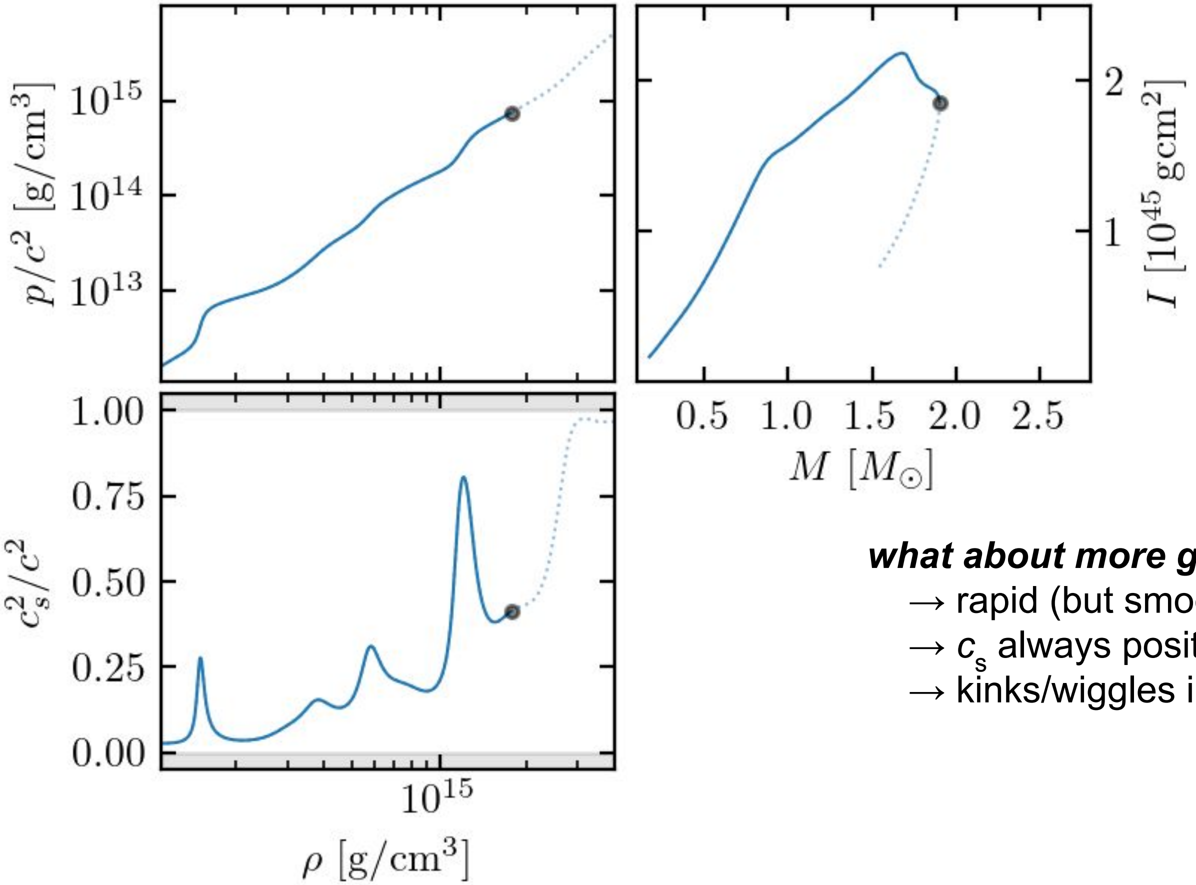
→ *"loss of pressure support"*





Inference of the NS EoS: phase transitions

PRD (2023)



what about more general phenomenology?

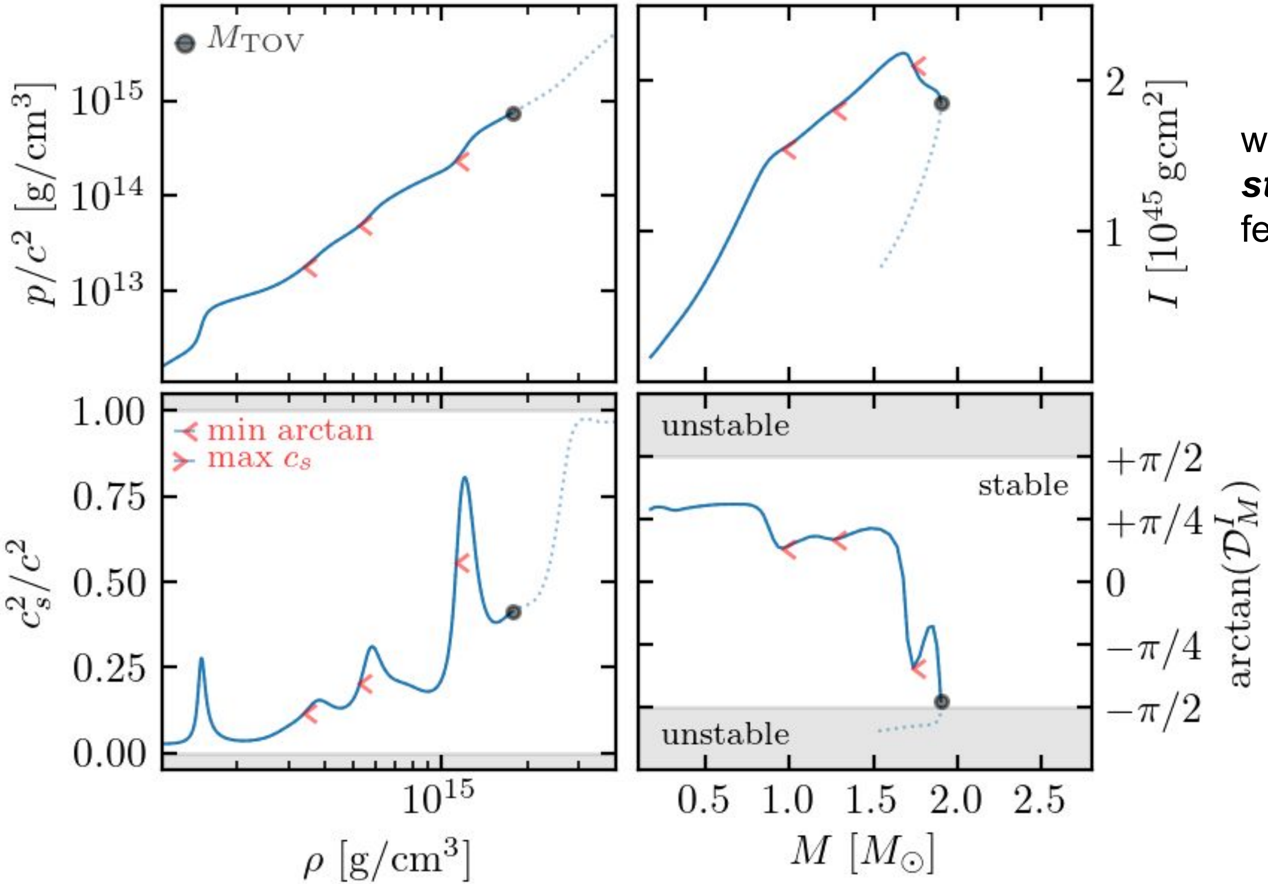
→ rapid (but smooth) changes in c_s?

→ c_s always positive definite?

→ kinks/wiggles in stellar sequence but no loss of stability?

Inference of the NS EoS: phase transitions

PRD (2023)

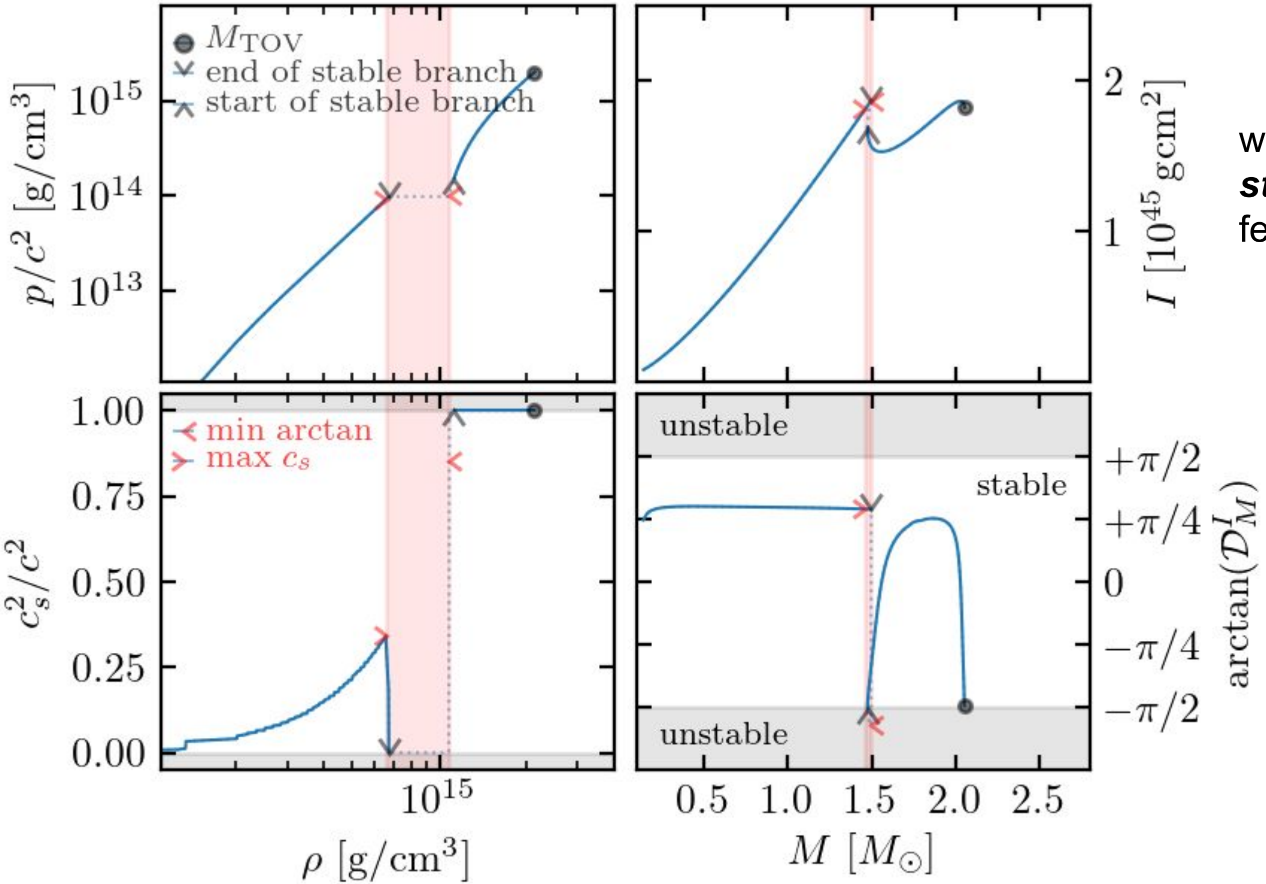


we search for ***features based on stellar properties*** and connect these to features in c_s

$$\mathcal{D}_M^I \equiv \frac{d\log I/d\log p_c}{d\log M/d\log p_c}$$

Inference of the NS EoS: phase transitions

PRD (2023)



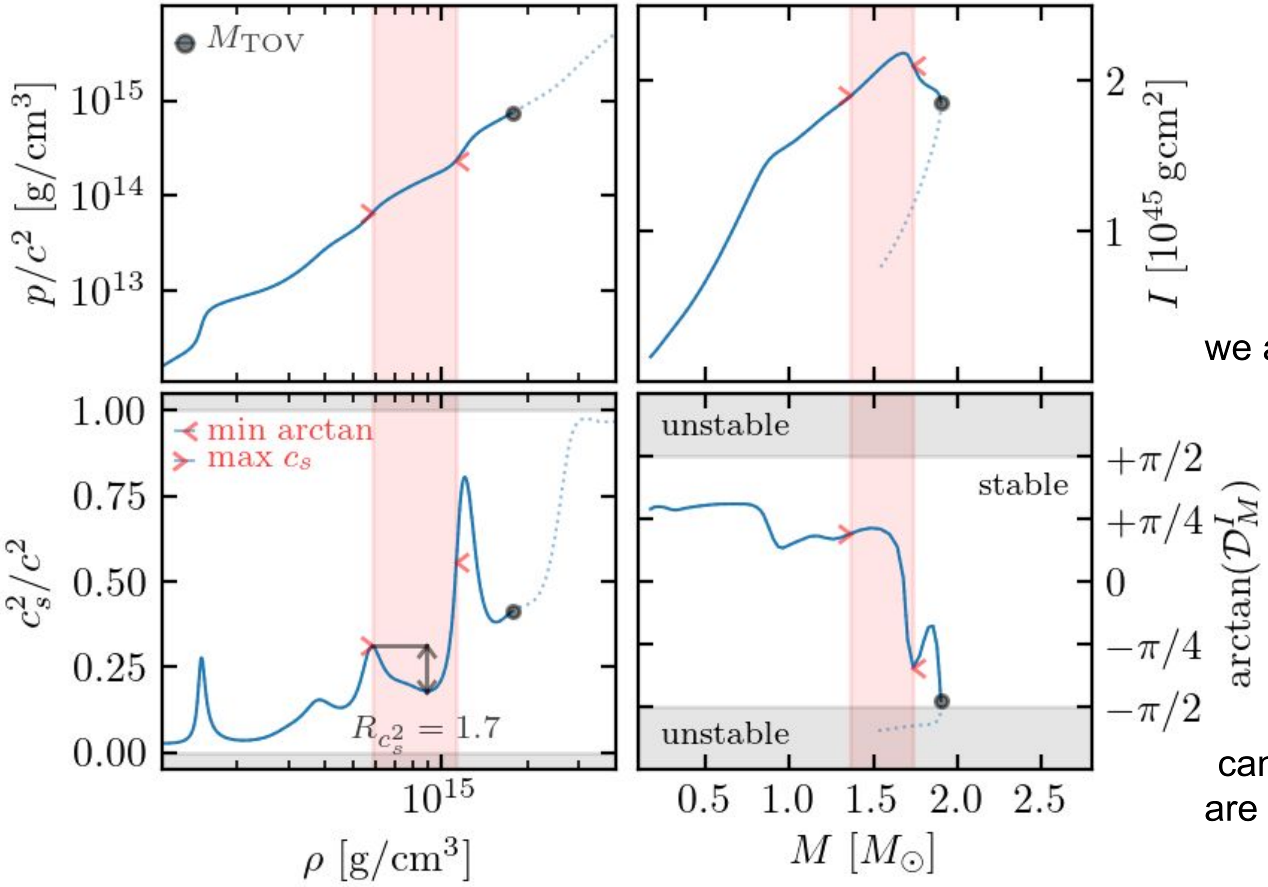
we search for ***features based on stellar properties*** and connect these to features in c_s

local minima → ***end*** of phase trans

and do so more precisely than where stability is regained

Inference of the NS EoS: phase transitions

PRD (2023)

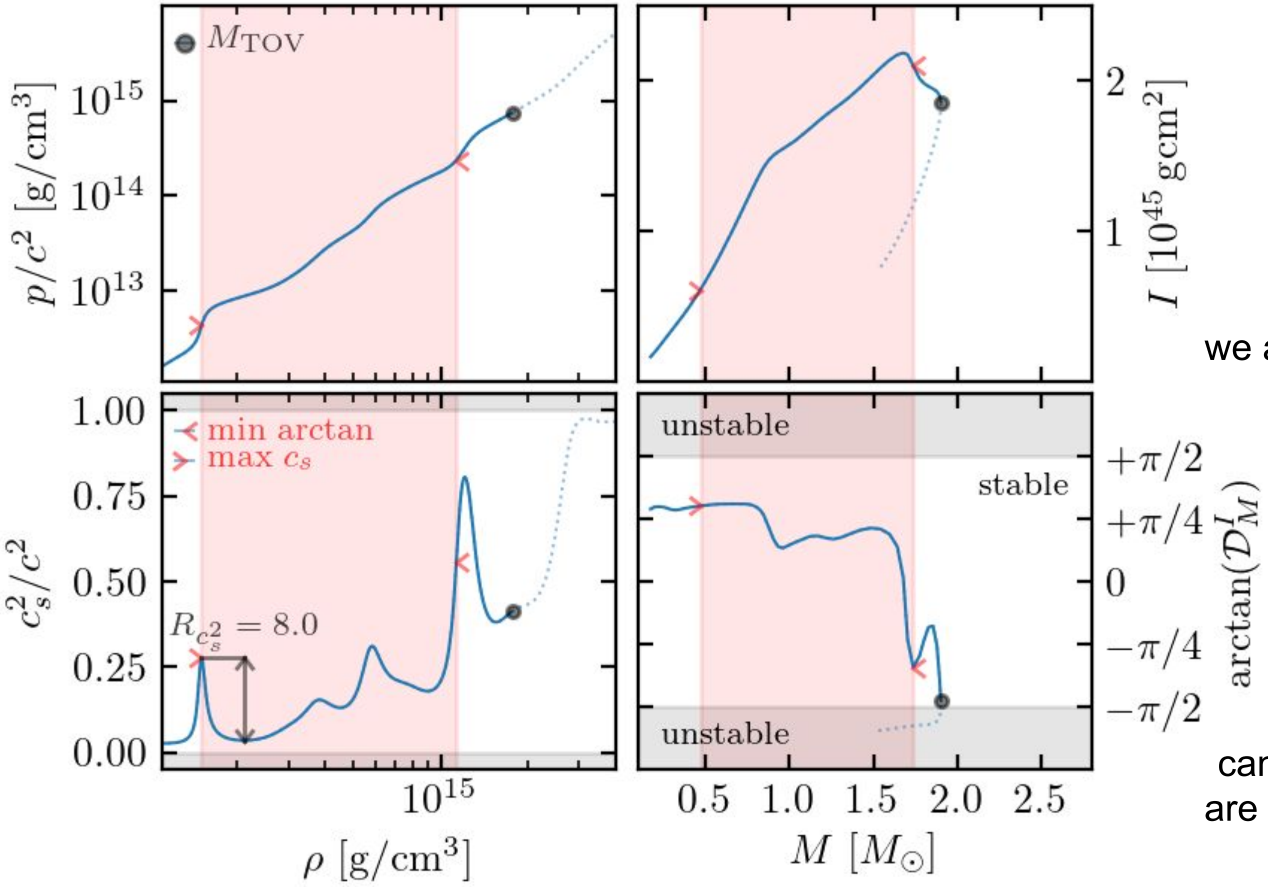


we associate a ***local minimum*** (***end***) with the most recent preceding ***"running maximum" in c_s*** (***onset***)

candidate ***running maxima in c_s*** (***onset***) are accepted only if they are followed by a ***large drop in c_s***

Inference of the NS EoS: phase transitions

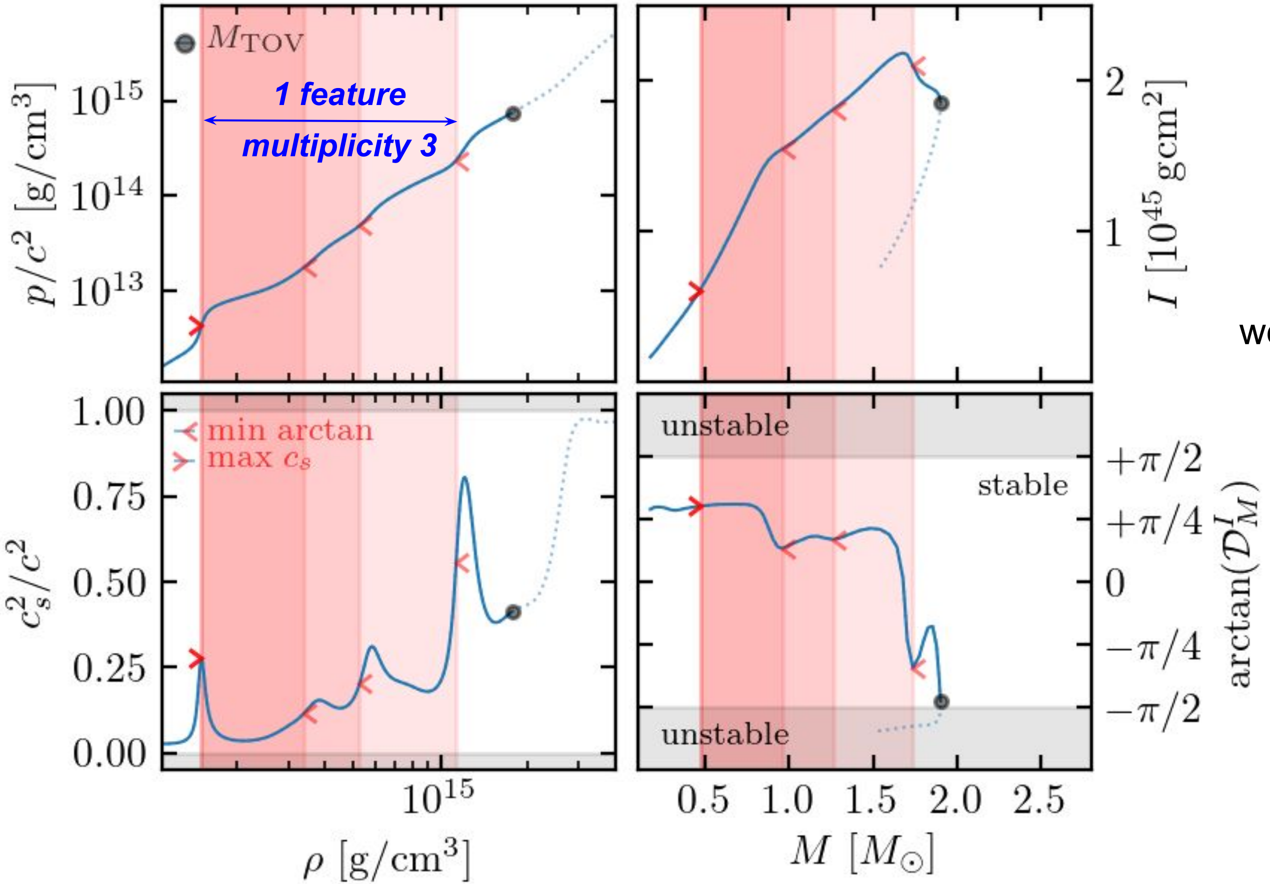
PRD (2023)



we associate a ***local minimum*** (***end***) with the most recent preceding ***"running maximum" in c_s*** (***onset***)

candidate ***running maxima in c_s*** (***onset***) are accepted only if they are followed by a ***large drop in c_s***

Inference of the NS EoS: phase transitions

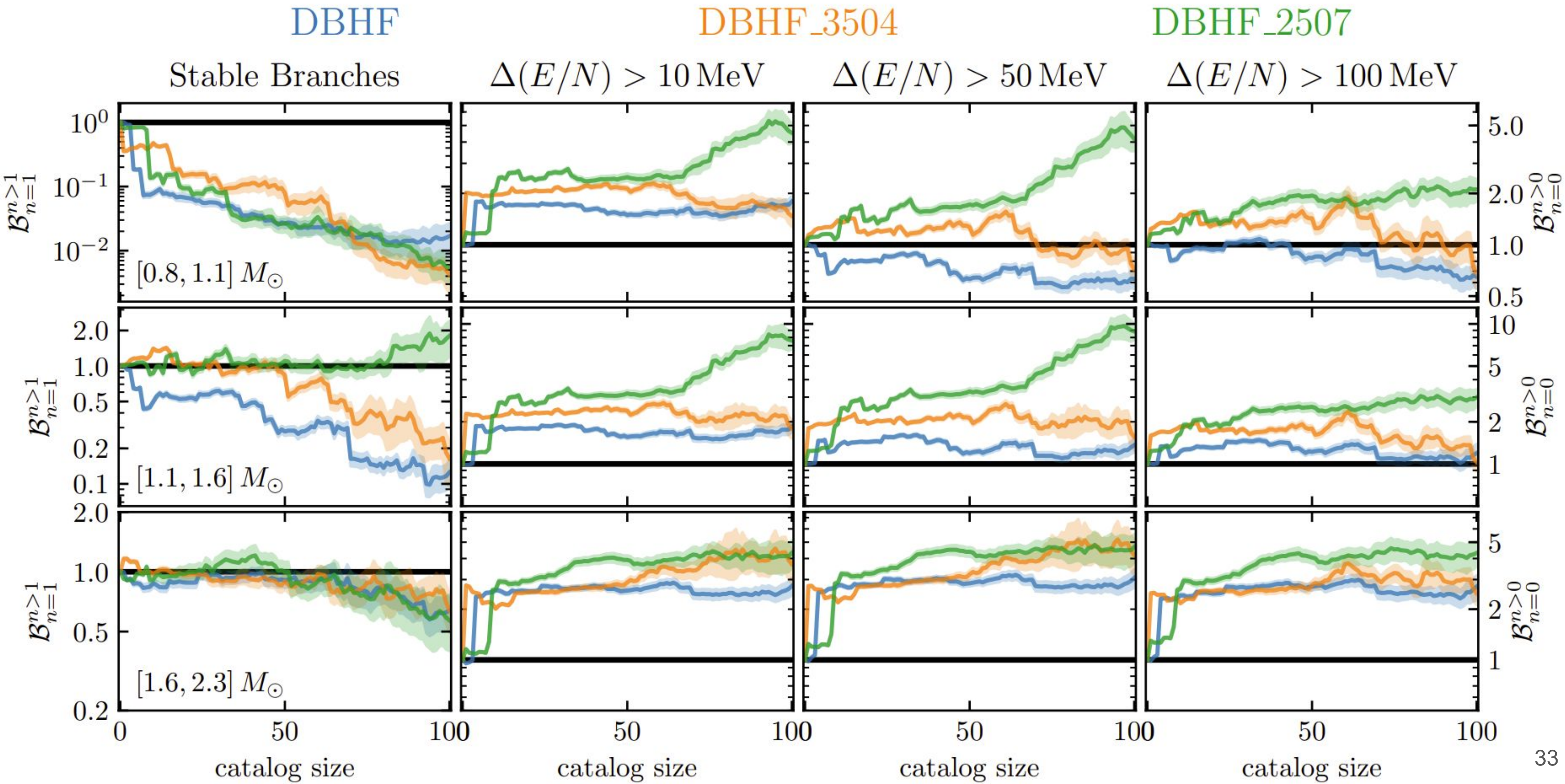


we iterate until all ***local minimum*** (***end***) *have an associated* ***"running maximum" in c_s*** (***onset***)

this EoS has
1 feature
with
multiplicity 3

Future Prospects: phase transitions

PRD (2023)



Extracting Physics without Parameters

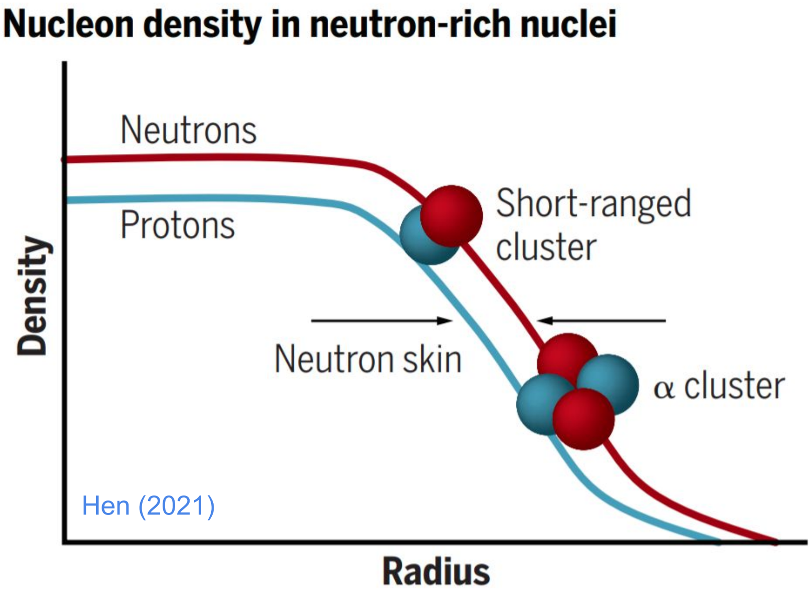
Nuclear Symmetry Energy

Inference of the NS EoS: low-density nuclear experiment

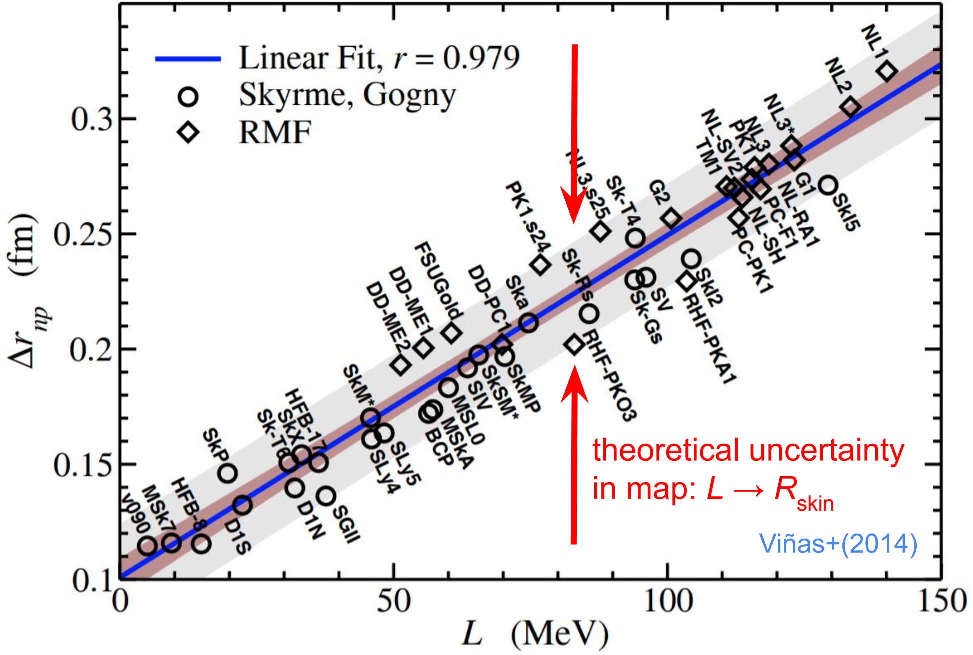
PRL (2021)
PRC (2021)

Connection to "new" experimental probes: Neutron Skin Thickness (R_{skin})

Reed+(2021) infer $L \gtrsim 100$ MeV based on $R_{skin} = 0.29 \pm 0.07$ fm. Suggest this implies $R_{1.4} \gtrsim 14$ km.



Scattering experiments (PREX) measure this



we constrain this with astrophysical observations

Inference of the NS EoS: low-density nuclear experiment

Map from nonparametric EoS in β-equilibrium to nuclear params describing the energy per particle near nuclear saturation (n_0: minimum of E_{SNM})

$$x = n_p/n$$

proton fraction

$$E_{\mathrm{nuc}}(n,x) = E_{\mathrm{SNM}}(n) + (1-2x)^2 S_0(n) + \mathcal{O}(x^4) = \frac{\varepsilon_\beta(n) - \varepsilon_e(n,x)}{n} - m_N$$

nuclear energy per particle

$$E_{\mathrm{SNM}}(n) = E_0 + \frac{1}{2} K_0 \left(\frac{n-n_0}{3n_0}\right)^2 + \cdots$$

symmetric-nuclear-matter energy per particle (local min at n_0)

$$\mu_n = \mu_p + \mu_e$$

condition for β-equilib

solve these self-consistently to obtain $S_0(n)$ and then compute

$$L = 3n\left(\frac{dS_0}{dn}\right)$$

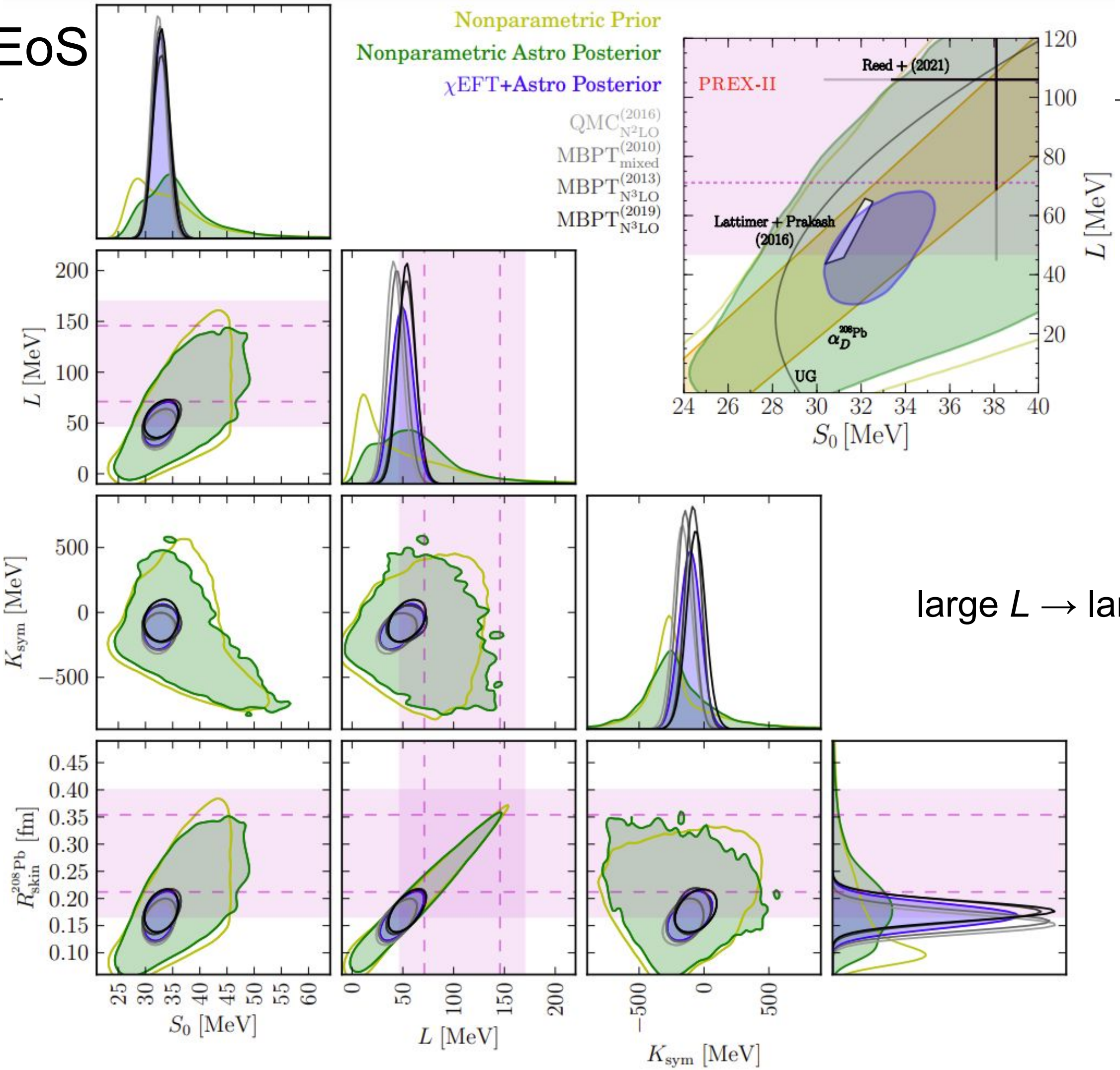
$$K_{\mathrm{sym}} = 9n^2\left(\frac{d^2 S_0}{dn^2}\right)$$

constrained by astro observations (input from nonparametric analysis)
measured in the lab (input from terrestrial experiment)
modeled as degenerate Fermi gas (input from theory)
expressed in terms of derivatives of E_{nuc}

$$\mu_i = \frac{dE}{dN_i}$$

Inference of the NS EoS

PRL (2021)
PRC (2021)



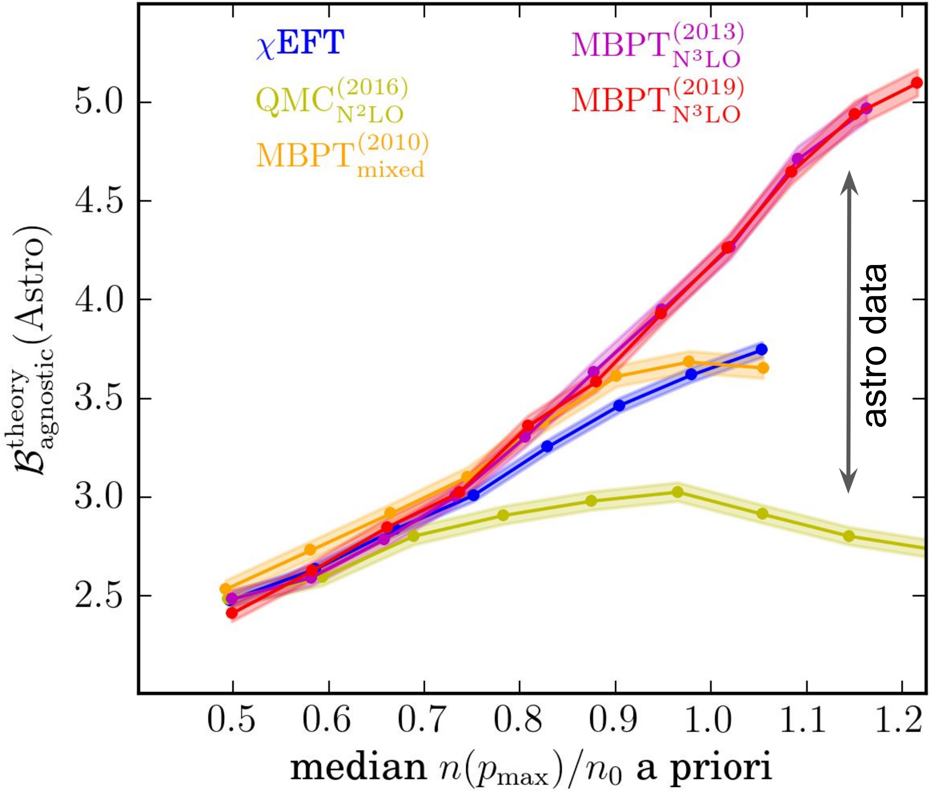
We can also extract "nuclear parameters" directly from nonparametric EoS without the need for parametrized EoS models

large L → large $p(n_{sat})$

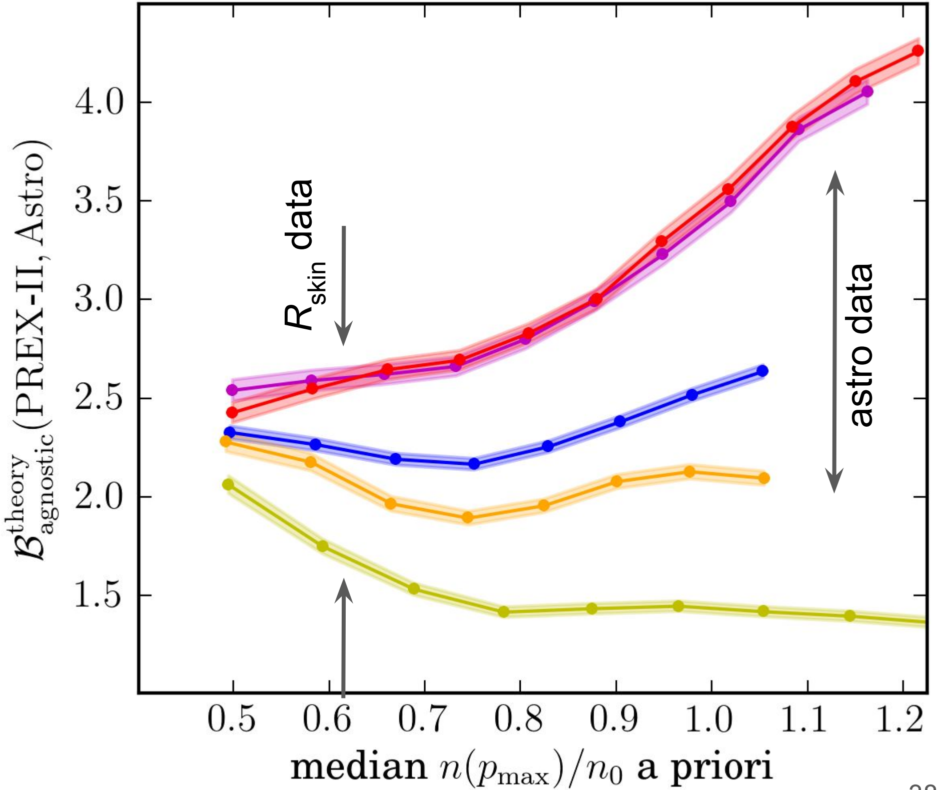
Inference of the NS EoS: low-density nuclear experiment

PRL (2021)
PRC (2021)

astro data can distinguish between nuclear theories at high densities



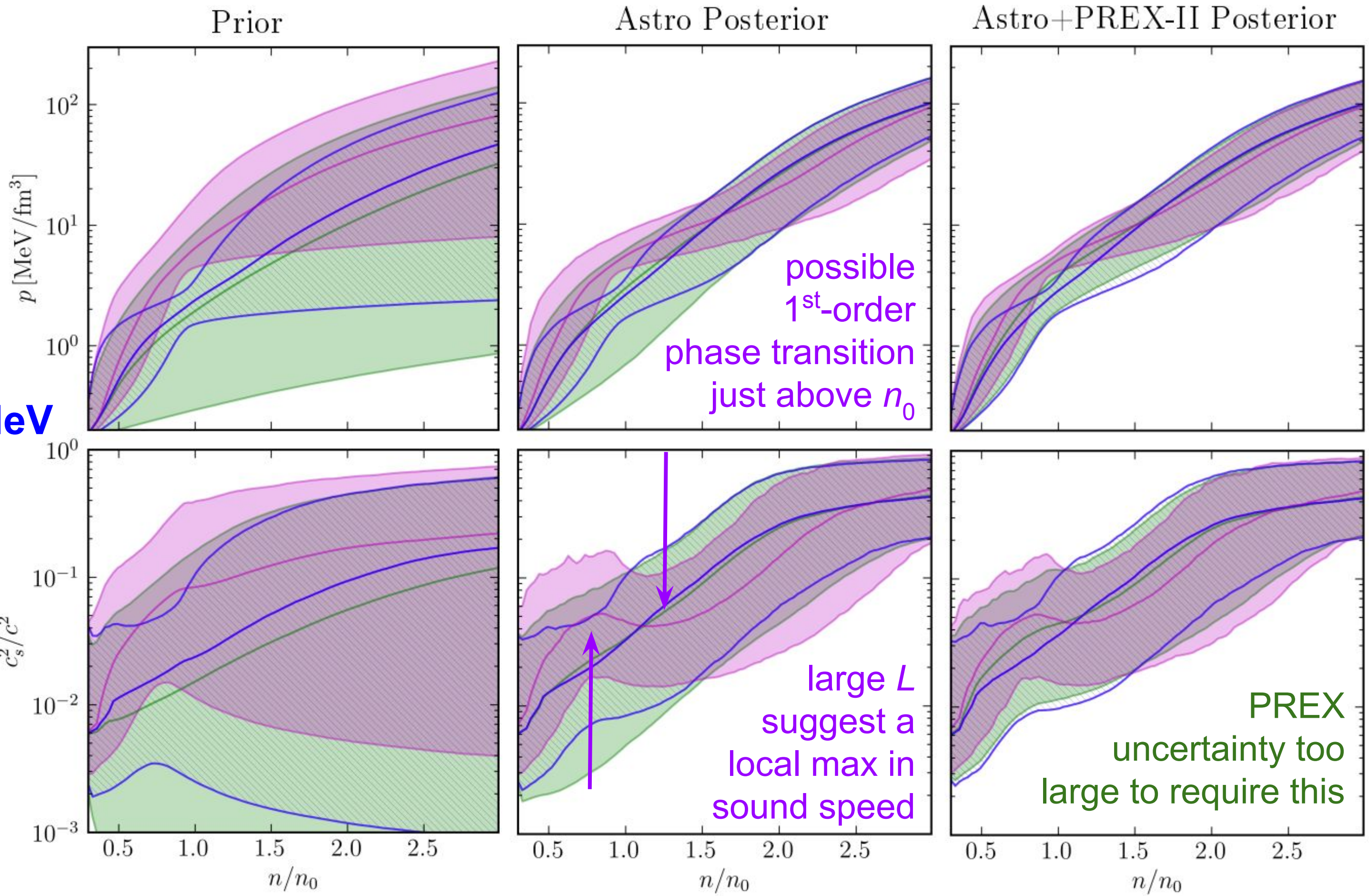
nuclear experiments probe lower densities



Inference of the NS EoS: low-density nuclear experiment

PRL (2021)
PRC (2021)

100 MeV < *L*
30 MeV < *L* < 70 MeV
All *L*



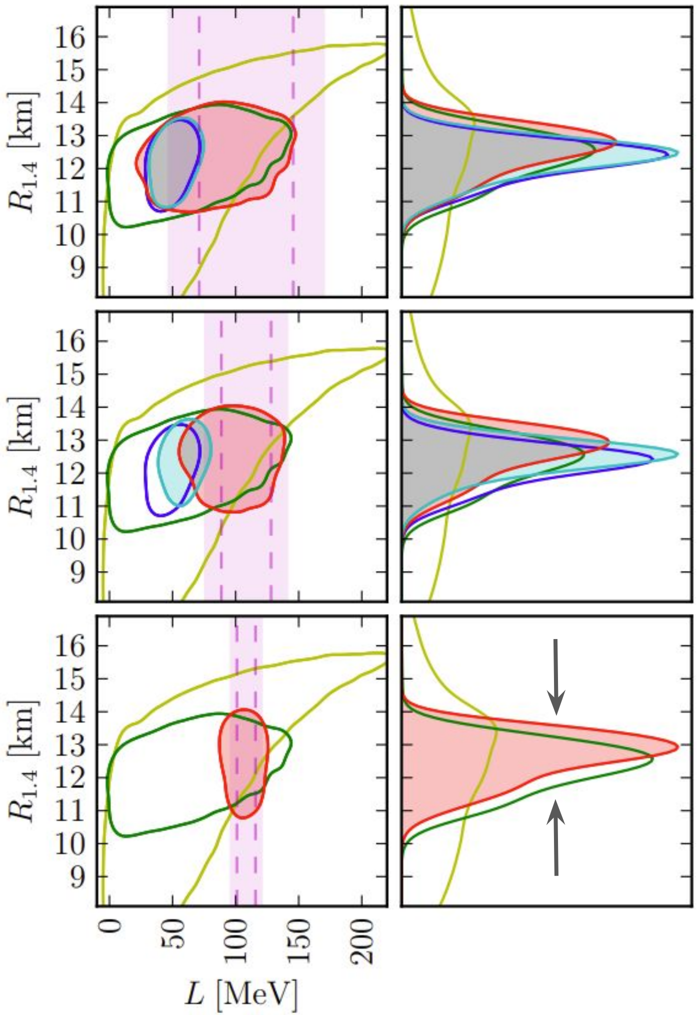
Inference of the NS EoS

PRL (2021)
PRC (2021)

current R_{skin} uncertainty

R_{skin} uncertainty improved by a factor of 2

hypothetical perfect R_{skin} measurement



nonparametric prior
nonparametric astro-only posterior
χEFT+astro posterior
nonparametric astro+R_{skin} posterior
χEFT+astro+R_{skin} posterior

improved precision in nuclear experiments is unlikely to affect our knowledge of NS radii without improved theoretical calculations

Self-Consistent Hierarchical Bayesian Inference

Introduction

ApJ (2024)

Construct a hierarchical generative model that relates the differential astrophysical rate of events

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

to the construction of an observed catalog of discrete i.i.d. events (i.e., detected data)

$$\{d_i, \mathbb{D}_i\}_{i=1,\ldots,N}$$

each of which has several latent (unobserved) parameters

$$\{\theta_i\}_{i=1,\ldots,N}$$

(Λ) population parameters (e.g., minimum and maximum mass, etc.), the

(θ_i) single-event parameters for the i^{th} event (e.g., masses, spins, etc.), the

(d_i) $= n_i + h(\theta_i)$ observed data for the i^{th} event assuming additive noise n_i and a signal model $h(\theta_i)$, and an

($\mathbb{D}_i$) indicator signifying that the i^{th} event was detected.

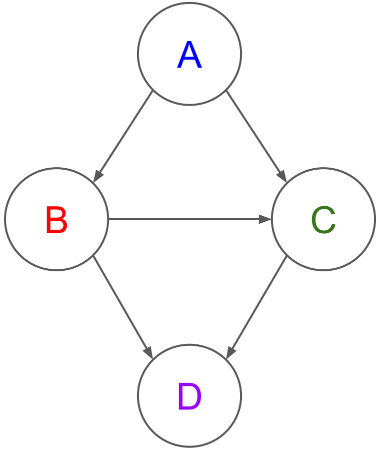
($\mathcal{K}$) expected number of astrophysical events within the past light-cone spanning the duration of the experiment or, alternatively,

(K) $= \mathcal{K} P(\mathbb{D}|\Lambda)$ the expected number of detected events

Directed Acyclic Graphs

We can represent this model as a Directed Acyclic Graph (DAG), which encodes conditional (in)dependencies between variables

graph

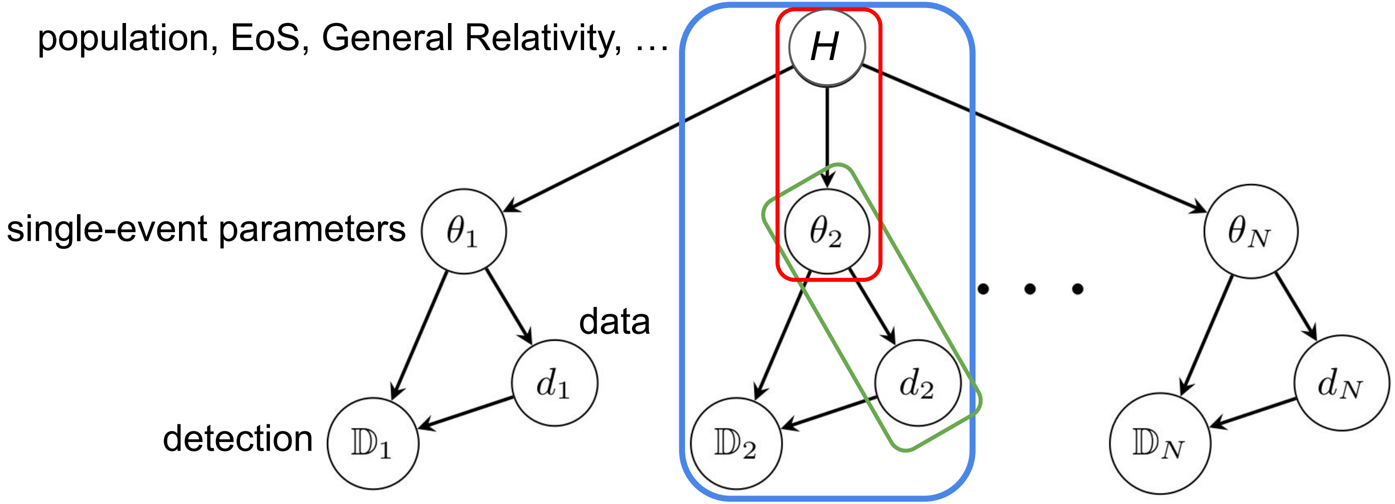


equation

$$p(D|B,C)\ p(C|A,B)\ p(B|A)\ p(A)$$

Directed Acyclic Graphs

We use a hierarchical model of the data-generation process



or, equivalently,

$$p(\{\mathbb{D}_i, d_i, \theta_i\}|N, H) = \prod_i^N P(\mathbb{D}_i|d_i, \theta_i)\, p(d_i|\theta_i)\, p(\theta_i|H)$$

likelihood prior

Generative Model (bottom-up approach)

For an astronomical survey, the number of events is uncertain (Poisson distributed)

$$p(N|K(\Lambda, \mathcal{K})) = \frac{K^N}{N!}e^{-K}$$

Generative Model (bottom-up approach)

ApJ (2024)

For an astronomical survey, the number of events is uncertain (Poisson distributed)

$$p(N|K(\Lambda, \mathcal{K})) = \frac{K^N}{N!} e^{-K}$$

and the data are i.i.d.

$$p(\{d_i\}|\{\mathbb{D}_i\}, \Lambda, N) = \prod_i^N p(d_i|\mathbb{D}_i, \Lambda)$$

Generative Model (bottom-up approach)

ApJ (2024)

For an astronomical survey, the number of events is uncertain (Poisson distributed)

$$p(N|K(\Lambda, \mathcal{K})) = \frac{K^N}{N!} e^{-K}$$

and the data are i.i.d.

$$p(\{d_i\}|\{\mathbb{D}_i\}, \Lambda, N) = \prod_i^N p(d_i|\mathbb{D}_i, \Lambda)$$

so that we can construct a joint distribution

$$p(\{d_i\}, K, \Lambda|\{\mathbb{D}_i\}, N) \propto p(K, \Lambda) p(\{d_i\}|\{\mathbb{D}_i\}, \Lambda, N) p(N|K)$$

Generative Model (bottom-up approach)

ApJ (2024)

For an astronomical survey, the number of events is uncertain (Poisson distributed)

$$p(N|K(\Lambda,\mathcal{K})) = \frac{K^N}{N!}e^{-K}$$

and the data are i.i.d.

$$p(\{d_i\}|\{\mathbb{D}_i\},\Lambda,N) = \prod_i^N \boxed{p(d_i|\mathbb{D}_i,\Lambda)}$$

so that we can construct a joint distribution

$$p(\{d_i\},K,\Lambda|\{\mathbb{D}_i\},N) \propto p(K,\Lambda)p(\{d_i\}|\{\mathbb{D}_i\},\Lambda,N)p(N|K)$$

from which we obtain the (hyper)posterior

$$p(K,\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(K,\Lambda)K^N e^{-K}\prod_i^N \boxed{\left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i,\theta)p(d_i|\theta)p(\theta|\Lambda)\right]}}{\int dK d\Lambda\, p(K,\Lambda)K^N e^{-K}\prod_i^N\left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i,\theta)p(d_i|\theta)p(\theta|\Lambda)\right]}$$

which is ***a mess***…

Simplifying Assumptions

ApJ (2024)

$p(K, \Lambda) = p(K)p(\Lambda)$

The number of detected events is independent of the shape of the astro population

Simplifying Assumptions

ApJ (2024)

$p(K, \Lambda) = p(K)p(\Lambda)$ The number of detected events is independent of the shape of the astro population

The inference then factors

$$p(K, \Lambda|\{d_i, \mathbb{D}_i\}, N) = p(K|N)p(\Lambda|\{d_i, \mathbb{D}_i\}, N)$$

where

$$p(K|N) = \frac{p(K)K^N e^{-K}}{\int dK\, p(K)K^N e^{-K}}$$

$$p(\Lambda|\{d_i, \mathbb{D}_i\}, N) = \frac{p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i, \theta)p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i, \theta)p(d_i|\theta)p(\theta|\Lambda)\right]}$$

Simplifying Assumptions

ApJ (2024)

$p(K, \Lambda) = p(K)p(\Lambda)$ The number of detected events is independent of the shape of the astro population

The inference then factors

$$p(K, \Lambda | \{d_i, \mathbb{D}_i\}, N) = p(K|N)p(\Lambda|\{d_i, \mathbb{D}_i\}, N)$$

where

$$p(K|N) = \frac{p(K)K^N e^{-K}}{\int dK\, p(K)K^N e^{-K}}$$

$$p(\Lambda|\{d_i, \mathbb{D}_i\}, N) = \frac{p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1} \int d\theta\, P(\mathbb{D}_i|d_i,\theta)p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1} \int d\theta\, P(\mathbb{D}_i|d_i,\theta)p(d_i|\theta)p(\theta|\Lambda)\right]}$$

We will focus on this in what follows.

Physical Detection Processes
and the
Inconsistency of Unphysical Model Assumptions

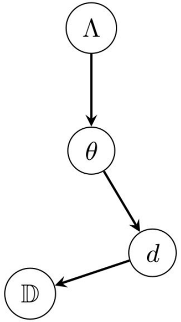
Simplifying Assumptions

ApJ (2024)

Physical $P(\mathbb{D}|d,\theta) = P(\mathbb{D}|d)$

$(\mathbb{D} \perp \theta \mid d)$

$$p(\Lambda|\{d_i, \mathbb{D}_i\}, N) = \frac{p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i)p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i)p(d_i|\theta)p(\theta|\Lambda)\right]}$$



physical detection processes only have access to the data

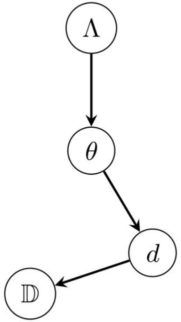
Simplifying Assumptions

Physical $\quad P(\mathbb{D}|d,\theta) = P(\mathbb{D}|d)$

$(\mathbb{D} \perp \theta \mid d)$

term-by-term cancellation

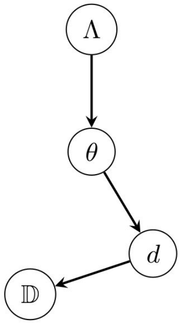
$$p(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i)p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i)p(d_i|\theta)p(\theta|\Lambda)\right]}$$



physical detection processes only have access to the data

Simplifying Assumptions

Physical $(\mathbb{D} \perp \theta \mid d)$ $\quad P(\mathbb{D}|d,\theta) = P(\mathbb{D}|d)$



term-by-term cancellation

$$p(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N\left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i)p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N\left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i)p(d_i|\theta)p(\theta|\Lambda)\right]}$$

$$\propto p(\Lambda)P(\mathbb{D}|\Lambda)^{-N}\prod_i^N\int d\theta\, p(d_i|\theta)p(\theta|\Lambda)$$

physical detection processes only have access to the data

standard expression

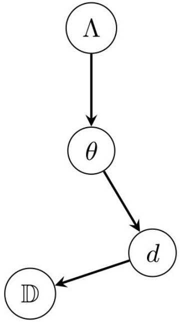
$$P(\mathbb{D}|\Lambda) = \int d\theta\, p(\theta|\Lambda)\int dd P(\mathbb{D}|d)p(d|\theta)$$
$$= \int d\theta\, p(\theta|\Lambda)P(\mathbb{D}|\theta)$$

Simplifying Assumptions

ApJ (2024)

Physical $P(\mathbb{D}|d,\theta) = P(\mathbb{D}|d)$

$(\mathbb{D} \perp \theta \mid d)$



term-by-term cancellation

$$p(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i) p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i) p(d_i|\theta)p(\theta|\Lambda)\right]}$$

$$\propto p(\Lambda)P(\mathbb{D}|\Lambda)^{-N}\prod_i^N \int d\theta\, p(d_i|\theta)p(\theta|\Lambda)$$

standard expression

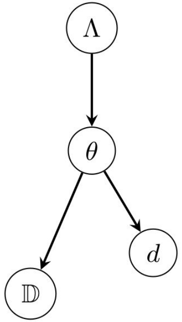
$$P(\mathbb{D}|\Lambda) = \int d\theta\, p(\theta|\Lambda)\int dd P(\mathbb{D}|d)p(d|\theta)$$
$$= \int d\theta\, p(\theta|\Lambda)P(\mathbb{D}|\theta)$$

physical detection processes only have access to the data

Unphysical $P(\mathbb{D}|d,\theta) = Q(\mathbb{D}|\theta)$

$(\mathbb{D} \perp d \mid \theta)$

$$q(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N \left[Q(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, Q(\mathbb{D}_i|\theta)p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[Q(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, Q(\mathbb{D}_i|\theta)p(d_i|\theta)p(\theta|\Lambda)\right]}$$



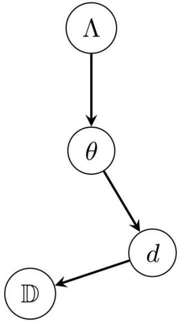
incorrectly models detection as independent of the data given the event's true parameters

Simplifying Assumptions

ApJ (2024)

Physical $P(\mathbb{D}|d,\theta) = P(\mathbb{D}|d)$

$(\mathbb{D} \perp \theta \mid d)$



term-by-term cancellation

$$p(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i) p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i) p(d_i|\theta)p(\theta|\Lambda)\right]}$$

$$\propto p(\Lambda)P(\mathbb{D}|\Lambda)^{-N}\prod_i^N \int d\theta\, p(d_i|\theta)p(\theta|\Lambda)$$

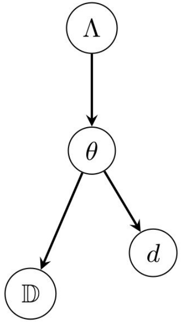
standard expression

$$P(\mathbb{D}|\Lambda) = \int d\theta\, p(\theta|\Lambda)\int dd P(\mathbb{D}|d)p(d|\theta)$$
$$= \int d\theta\, p(\theta|\Lambda)P(\mathbb{D}|\theta)$$

physical detection processes only have access to the data

Unphysical $P(\mathbb{D}|d,\theta) = Q(\mathbb{D}|\theta)$

$(\mathbb{D} \perp d \mid \theta)$



no cancellation

$$q(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N \left[Q(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, Q(\mathbb{D}_i|\theta) p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[Q(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, Q(\mathbb{D}_i|\theta) p(d_i|\theta)p(\theta|\Lambda)\right]}$$

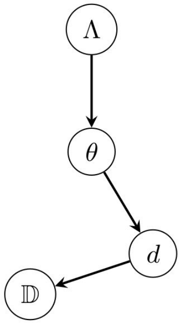
incorrectly models detection as independent of the data given the event's true parameters

Simplifying Assumptions

ApJ (2024)

Physical $P(\mathbb{D}|d,\theta) = P(\mathbb{D}|d)$

$(\mathbb{D} \perp \theta \mid d)$



term-by-term cancellation

$$p(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i) p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[P(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, P(\mathbb{D}_i|d_i) p(d_i|\theta)p(\theta|\Lambda)\right]}$$

$$\propto p(\Lambda)P(\mathbb{D}|\Lambda)^{-N}\prod_i^N \int d\theta\, p(d_i|\theta)p(\theta|\Lambda)$$

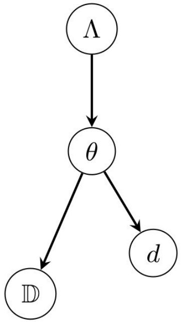
standard expression

$$P(\mathbb{D}|\Lambda) = \int d\theta\, p(\theta|\Lambda)\int dd P(\mathbb{D}|d)p(d|\theta)$$
$$= \int d\theta\, p(\theta|\Lambda)P(\mathbb{D}|\theta)$$

physical detection processes only have access to the data

Unphysical $P(\mathbb{D}|d,\theta) = Q(\mathbb{D}|\theta)$

$(\mathbb{D} \perp d \mid \theta)$



no cancellation

$$q(\Lambda|\{d_i,\mathbb{D}_i\},N) = \frac{p(\Lambda)\prod_i^N \left[Q(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, Q(\mathbb{D}_i|\theta) p(d_i|\theta)p(\theta|\Lambda)\right]}{\int d\Lambda p(\Lambda)\prod_i^N \left[Q(\mathbb{D}_i|\Lambda)^{-1}\int d\theta\, Q(\mathbb{D}_i|\theta) p(d_i|\theta)p(\theta|\Lambda)\right]}$$

$$\propto p(\Lambda)\prod_i^N \int d\theta\, p(d_i|\theta)q(\theta|\mathbb{D},\Lambda)$$

fitting the “detected distribution”

$$q(\theta|\mathbb{D},\Lambda) = \frac{Q(\mathbb{D}|\theta)p(\theta|\Lambda)}{Q(\mathbb{D}|\Lambda)}$$

incorrectly models detection as independent of the data given the event’s true parameters

Common Mistake 1

ApJ (2024)

Physical DAG but unphysical approximation $P(D|\Lambda) \sim Q(D|\Lambda)$

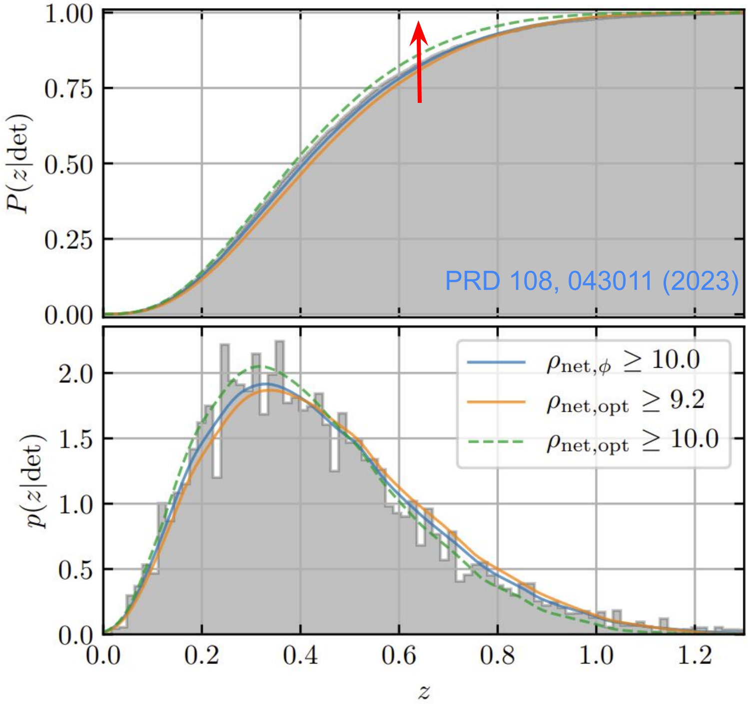
The physical model

$$p(\Lambda|\{d_i, \mathbb{D}_i\}, N) \propto p(\Lambda) P(\mathbb{D}|\Lambda)^{-N} \prod_i^N \int d\theta\, p(d_i|\theta) p(\theta|\Lambda)$$

is replaced by the unphysical approximation

$$q(\Lambda|\{d_i, \mathbb{D}_i\}, N) \propto p(\Lambda) Q(\mathbb{D}|\Lambda)^{-N} \prod_i^N \int d\theta\, p(d_i|\theta) p(\theta|\Lambda)$$

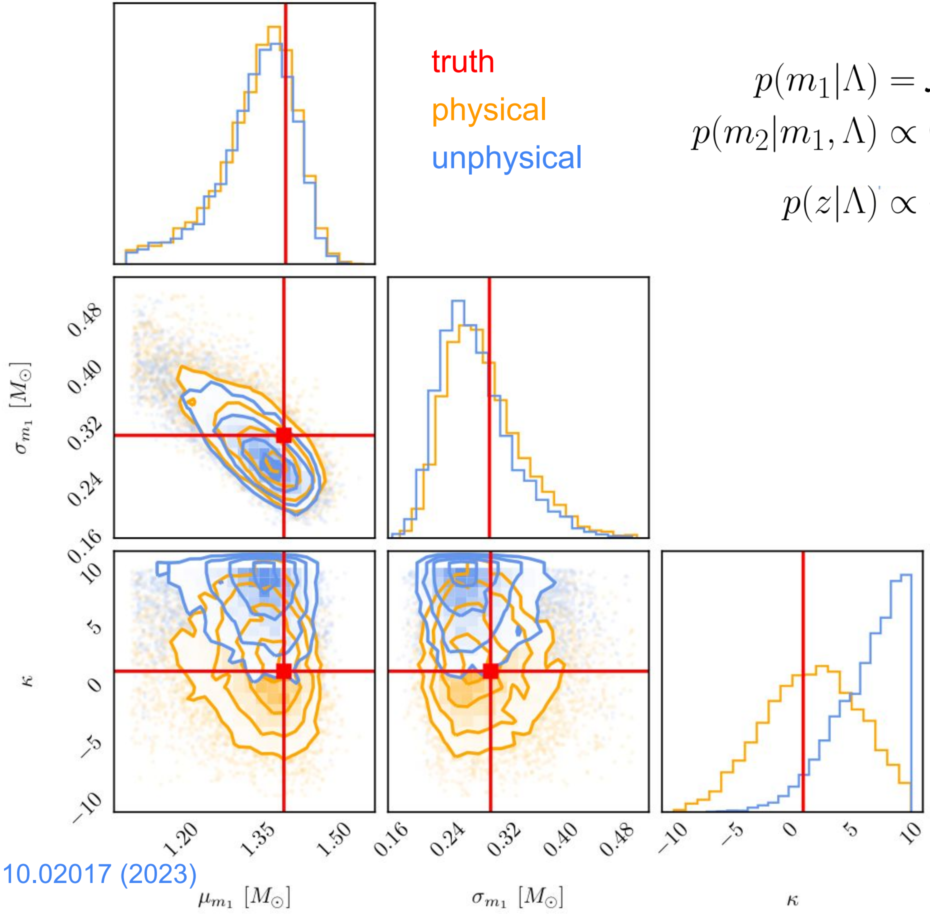
which will tend to systematically underestimate the sensitivity to quiet signals.



Common Mistake 1

ApJ (2024)

Physical DAG but unphysical approximation $P(D|\Lambda) \sim Q(D|\Lambda)$



$$p(m_1|\Lambda) = \mathcal{N}(\mu_{m_1}, \sigma^2_{m_1})\Theta(m_{\min} \leq m_1 \leq m_{\max})$$
$$p(m_2|m_1, \Lambda) \propto \Theta(m_{\min} \leq m_2 \leq m_1)$$
$$p(z|\Lambda) \propto \frac{dV_c}{dz}(1+z)^{\kappa-1}$$

selection is primarily against high-*z* systems, so the bias primarily affects the redshift-evolution

$$Q(\mathbb{D}|z) \ll P(\mathbb{D}|z) \Rightarrow \kappa \uparrow$$

in order to keep

$$K = \mathcal{K}P(\mathbb{D}|\Lambda) = \mathcal{K}\int d\theta\, p(\theta|\Lambda)P(\mathbb{D}|\theta)$$

approximately constant

arXiv:2310.02017 (2023)

Common Mistake 2

ApJ (2024)

Fitting the “detected distribution” and then dividing by $P(\mathrm{D}|\theta)$

Motivated by the observation

$$p(\theta|\Lambda) = \left(\frac{p(\theta|\mathbb{D},\Lambda)}{P(\mathbb{D}|\theta)}\right) P(\mathbb{D}|\Lambda)$$

Common Mistake 2

ApJ (2024)

Fitting the “detected distribution” and then dividing by *P*(D|θ)

Motivated by the observation

$$p(\theta|\Lambda) = \left(\frac{p(\theta|\mathbb{D},\Lambda)}{P(\mathbb{D}|\theta)}\right) P(\mathbb{D}|\Lambda)$$

It is tempting to fit for the distribution of true-parameters for detected events $p(\theta|\mathbb{D},\Lambda)$ from the observed data via

$$q(\{d_i\}|\{D_i\},N,\Lambda) = \prod_i^N \int d\theta\, p(d_i|\theta) q(\theta|\mathbb{D}_i,\Lambda)$$

with a “flexible enough” model for $q(\theta|\mathbb{D},\Lambda)$

Common Mistake 2

ApJ (2024)

Fitting the “detected distribution” and then dividing by *P*(D|θ)

Motivated by the observation

$$p(\theta|\Lambda) = \left(\frac{p(\theta|\mathbb{D},\Lambda)}{P(\mathbb{D}|\theta)}\right) P(\mathbb{D}|\Lambda)$$

It is tempting to fit for the distribution of true-parameters for detected events $p(\theta|D,\Lambda)$ from the observed data via

$$q(\{d_i\}|\{D_i\}, N, \Lambda) = \prod_i^N \int d\theta\, p(d_i|\theta) q(\theta|\mathbb{D}_i, \Lambda)$$

with a “flexible enough” model for $q(\theta|D,\Lambda)$ and then estimate the astro distribution

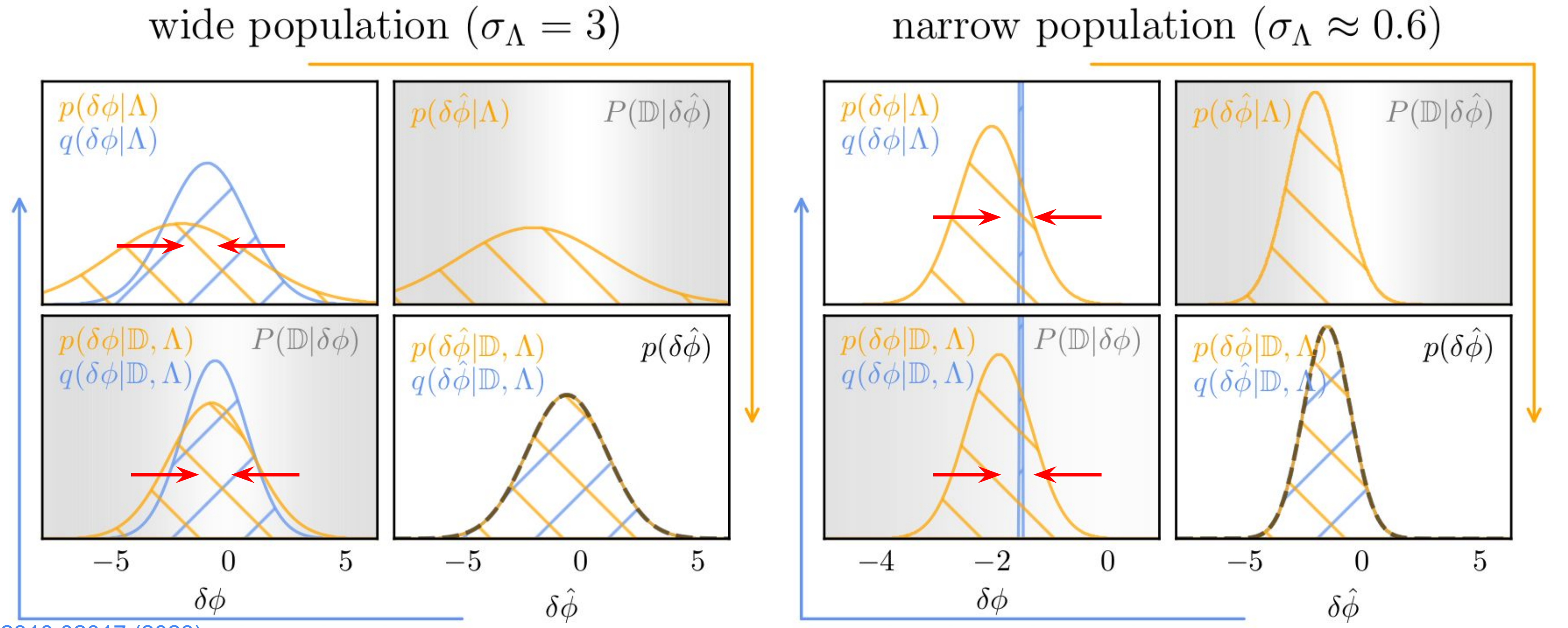
$$q(\theta|\Lambda) \propto \frac{q(\theta|\mathbb{D},\Lambda)}{P(\mathbb{D}|\theta)}$$

However, this procedure yields $q(\theta|D,\Lambda)$ that are too narrow, and therefore biased estimates of $q(\theta|\Lambda)$.

Common Mistake 2

Fitting the "detected distribution" and then dividing by $P(\mathrm{D}|\theta)$

ApJ (2024)



arXiv:2310.02017 (2023)

$$p(\delta\phi|\Lambda) = \mathcal{N}(\mu_\Lambda, \sigma_\Lambda^2) \qquad p(\delta\hat{\phi}|\delta\phi) = \mathcal{N}(\delta\phi, \sigma_o^2) \qquad P(\mathbb{D}|\delta\hat{\phi}) = \exp\left(-\frac{(\delta\hat{\phi} - \mu_D)^2}{\sigma_D^2}\right)$$