ET France workshop, Orsay, 31/03/26

Populations and cosmology in the 3G era

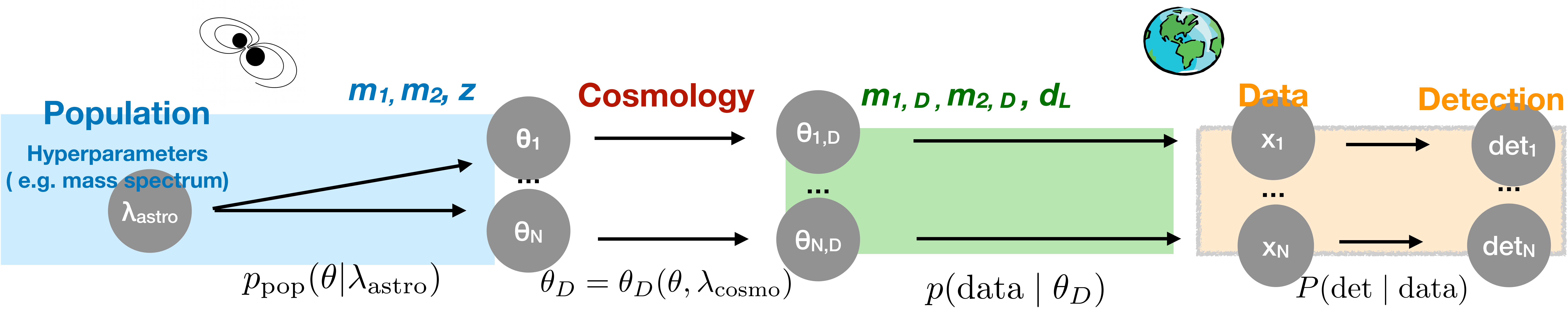
A brief introduction

Michele Mancarella

Centre de Physique Théorique, Aix-Marseille Université

Hierarchical models

Mandel, Farr, Gair, 1809.02063



PRIOR

Astrophysical model of mass-redshift distribution

$$p(\lambda|\{\text{data}\}) \propto \frac{\pi(\lambda)}{\xi(\lambda)^{N_{\rm obs}}} \prod_{i=1}^{N_{\rm obs}} \int \mathrm{d}\theta_{D,i}\, p(\text{data}_i \mid \theta_{D,i})\, p_{\rm pop}\big(\theta_i(\theta_D, \lambda_{\rm c}) \mid \lambda_{\rm a}\big)\, \mathrm{J}(\lambda_{\rm c})$$

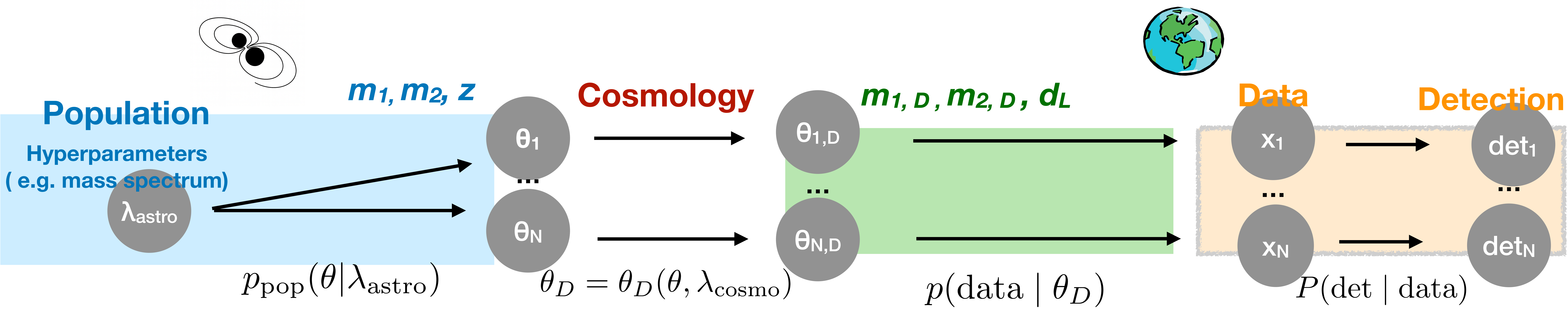
GW mass-distance measurement

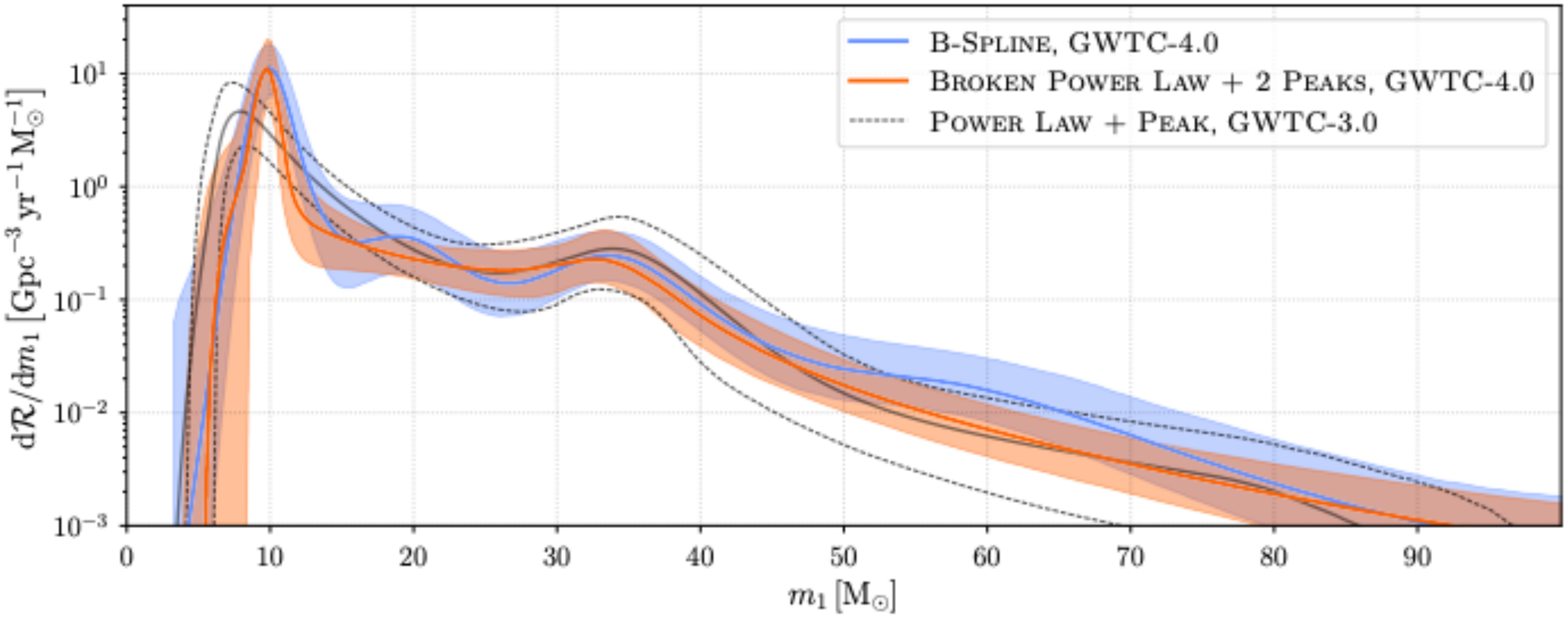
"SELECTION BIAS":
Expected fraction of observed events

$$\xi(\lambda) = \int \mathrm{d}\theta_D\, P(\det|\theta_D)\, p_{\rm pop}\big(\theta(\theta_D, \lambda_{\rm c})|\lambda_{\rm a}\big)\mathrm{J}(\lambda_{\rm c})$$

Hierarchical models

Mandel, Farr, Gair, 1809.02063





arXiv:2508.18080
(GWTC-4.0 population)

Hierarchical models

Prior

$$p(\lambda|\{\mathrm{data}\}) \propto \frac{\pi(\lambda)}{\xi(\lambda)^{N_{\mathrm{obs}}}} \prod_{i=1}^{N_{\mathrm{obs}}} \int \mathrm{d}\theta_{D,i}\, p(\mathrm{data}_i \mid \theta_{D,i})\, p_{\mathrm{pop}}\big(\theta_i(\theta_D, \lambda_{\mathrm{c}}) \mid \lambda_{\mathrm{a}}\big)\, \mathrm{J}(\lambda_{\mathrm{c}})$$

H_0 here

Selection bias (depends on astrophysical assumptions)

GW mass-distance measurement

Astrophysical model of mass-redshift distribution

- Need N_{obs} (+1) Monte Carlo estimators of N_θ-d integrals. Accuracy/efficiency balance
- Sensitivity via “injections” in real noise. Very expensive (human+machine time)
- Computing cost/accuracy vs scalability
- Assumes events are NON OVERLAPPING

(i) Monte Carlo integrations: the problem(s)

Prior

$$p(\lambda|\{\text{data}\}) \propto \frac{\pi(\lambda)}{\xi(\lambda)^{N_{\rm obs}}} \prod_{i=1}^{N_{\rm obs}} \underbrace{\int \mathrm{d}\theta_{D,i}}_{\approx \frac{1}{N_{\mathrm{s},i}} \sum_{k=1}^{N_{\mathrm{s},i}} \frac{p_{\rm pop}(\theta_{k,i}|\lambda)}{\pi_{\rm PE}(\theta_{k,i})}} p(\text{data}_i \mid \theta_{D,i})\, p_{\rm pop}(\theta_i(\theta_D, \lambda_{\rm c}) \mid \lambda_{\rm a})\, \mathrm{J}(\lambda_{\rm c})$$

H_0 here

Selection bias (depends on astrophysical assumptions)

GW mass-distance measurement

Astrophysical model of mass-redshift distribution

$$\frac{1}{N_{\mathrm{s},i}} \sum_{k=1}^{N_{\mathrm{s},i}} \frac{p_{\rm pop}(\theta_{k,i}|\lambda)}{\pi_{\rm PE}(\theta_{k,i})}$$

- Up to 7-D MC integrals. Recycle single-event PE posterior samples
- Accuracy of MC integration is a known issue

Cut "effective" points

$$N_{\rm eff} \equiv N_{\rm draw} \frac{\langle f \rangle^2}{\langle f^2 \rangle}$$

$> N_{\rm obs}$ ... or 10

Farr 2019

Data-dependent "prior"

(ii) "Injections"

Prior

$$p(\lambda|\{\text{data}\}) \propto \frac{\pi(\lambda)}{\xi(\lambda)^{N_{\rm obs}}} \prod_{i=1}^{N_{\rm obs}} \int \mathrm{d}\theta_{D,i}\, p(\text{data}_i \mid \theta_{D,i})\, p_{\rm pop}\big(\theta_i(\theta_D, \lambda_{\rm c}) \mid \lambda_{\rm a}\big)\, \mathrm{J}(\lambda_{\rm c})$$

H_0 here

Selection bias (depends on astrophysical assumptions)

GW mass-distance measurement

Astrophysical model of mass-redshift distribution

$$\xi(\lambda) = \int \mathrm{d}\theta_D\, P(\det|\theta_D)\, p_{\rm pop}\big(\theta(\theta_D, \lambda_{\rm c})|\lambda_{\rm a}\big)\mathrm{J}(\lambda_{\rm c}) \quad \approx \frac{1}{N_{inj}} \sum_{i \in \text{detected inj.}} \frac{p_{pop}(\theta_i|\lambda)}{p_{draw,i}}$$

Need to know P(det|θ)

re-weighted MC integration

Tiwari 2019, Farr 2019

- Sensitivity via "injections" in real noise + re-weighted Monte Carlo integration.

Cut lik. variance

Talbot, Golomb 2023

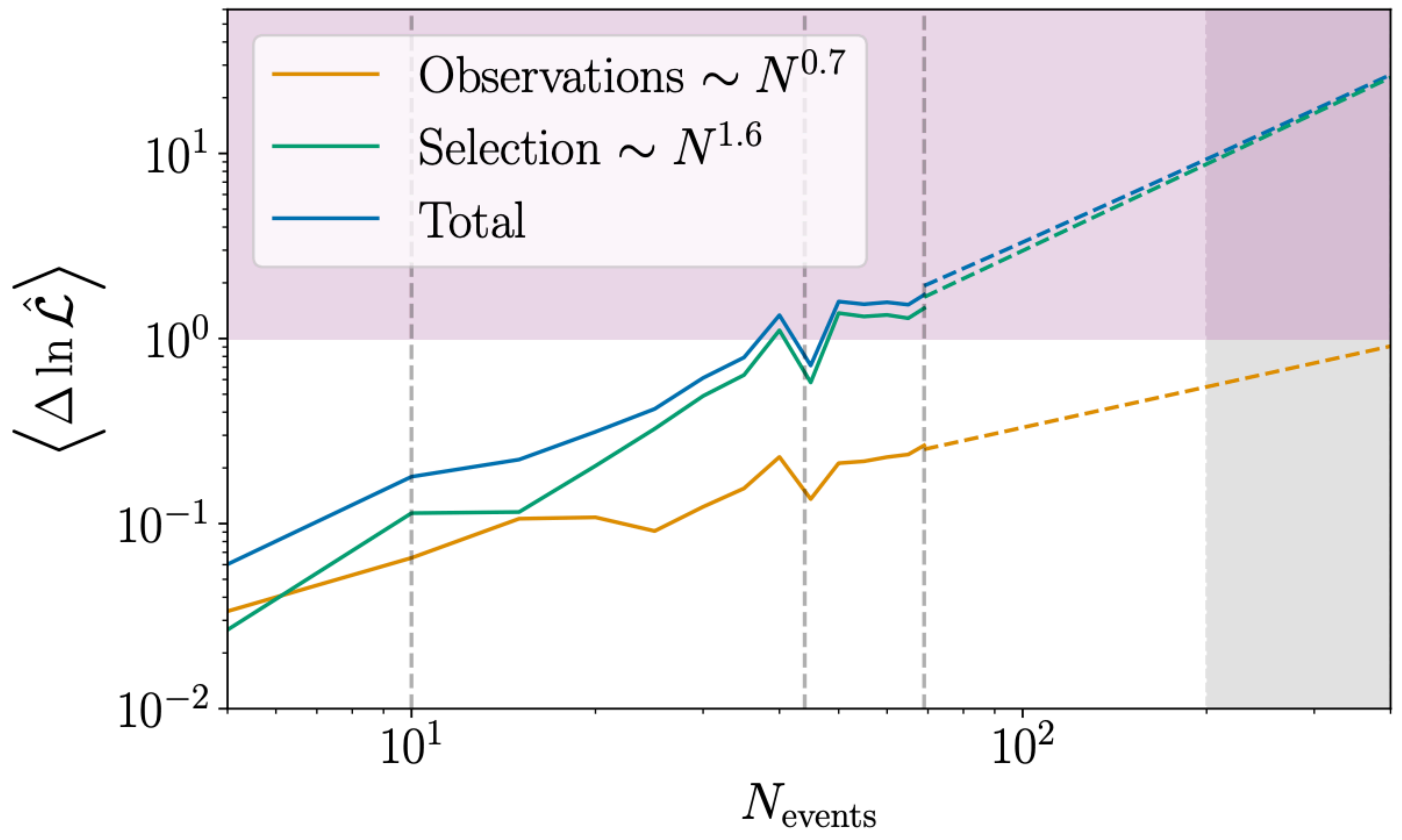
$$\mathrm{var}(\log \mathcal{L}) = \sum_{i=1}^{N_{\rm obs}} \frac{\mathrm{var}\mathcal{L}_i}{\hat{\mathcal{L}}_i^2} + N_{\rm obs}^2 \frac{\mathrm{var}\,\xi}{\hat{\xi}^2} \quad < 1$$

Can prevent exploring some models (spin distributions)

(i+ii) towards O5 and 3G

- Current approach scales bad with number of cataloged events

Talbot & Golomb 2023



- Area of active research

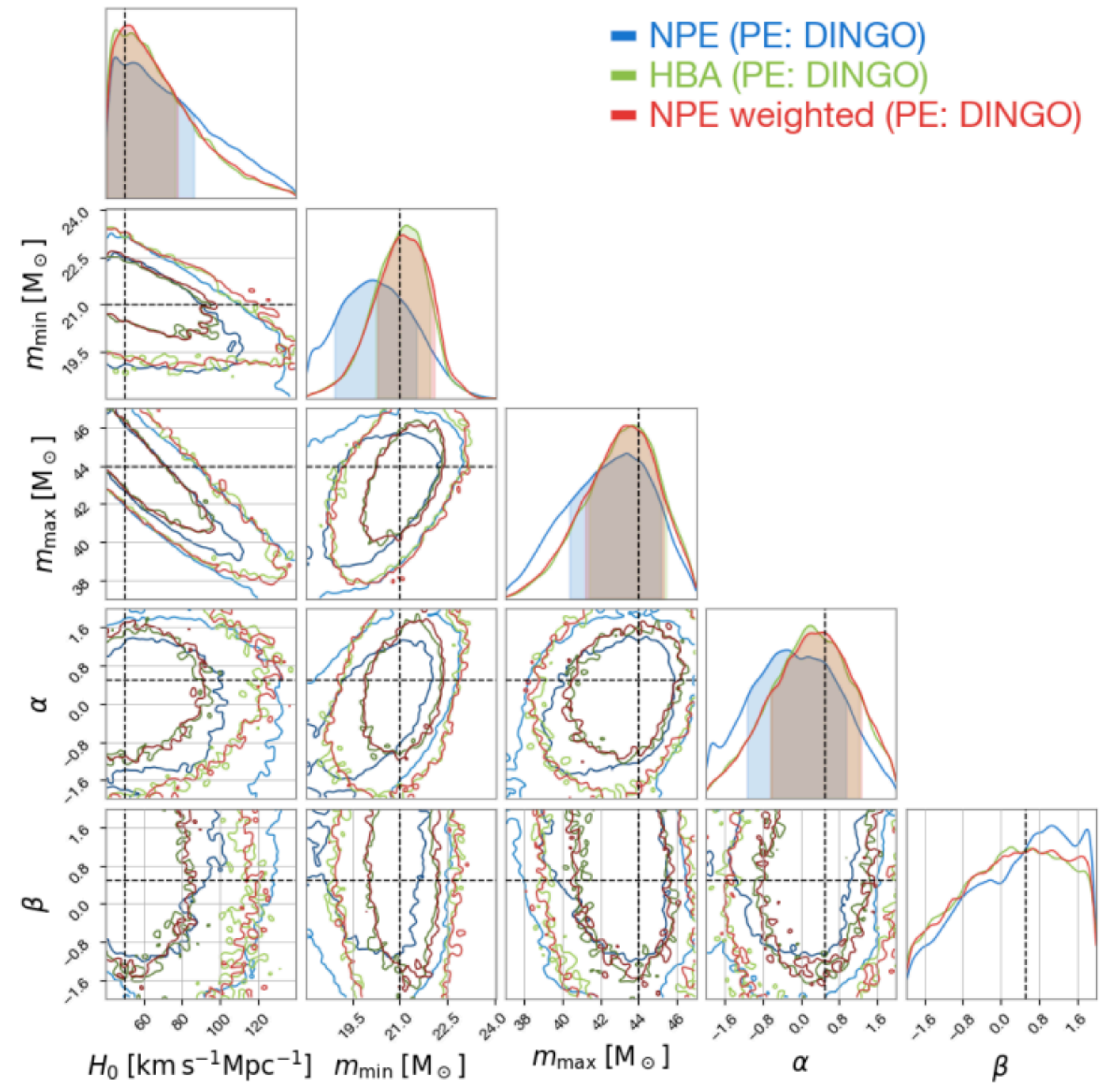
  - Do an insanely high number of injections — currently done
  - Mitigate MC variance with targeted strategies — Hussain + 2024,25
  - Machine-learning emulators for selection effect — Gerosa+ 22, Talbot&Thrane22,Callister+ 2023, …
  - A “full”, high dimensional approach (no MC variance from single events) — Mancarella & Gerosa 2025
  - Skip selection effect correction: compare observed populations (does not work for cosmology) — Toubiana + 2025

What’s best for ET ? Work for Div 3, 10, 9

(i+ii) Machine-learning strategies

- A tantalising possibility: skip the likelihood evaluation. <u>Simulation-based inference</u> for population

  - Only few prototypes exist
  - Need to emulate single-event PE in <seconds : DINGO currently used, but not calibrated everywhere
  - Scalability not clear



Leyde + 2023

(iii) Scalability

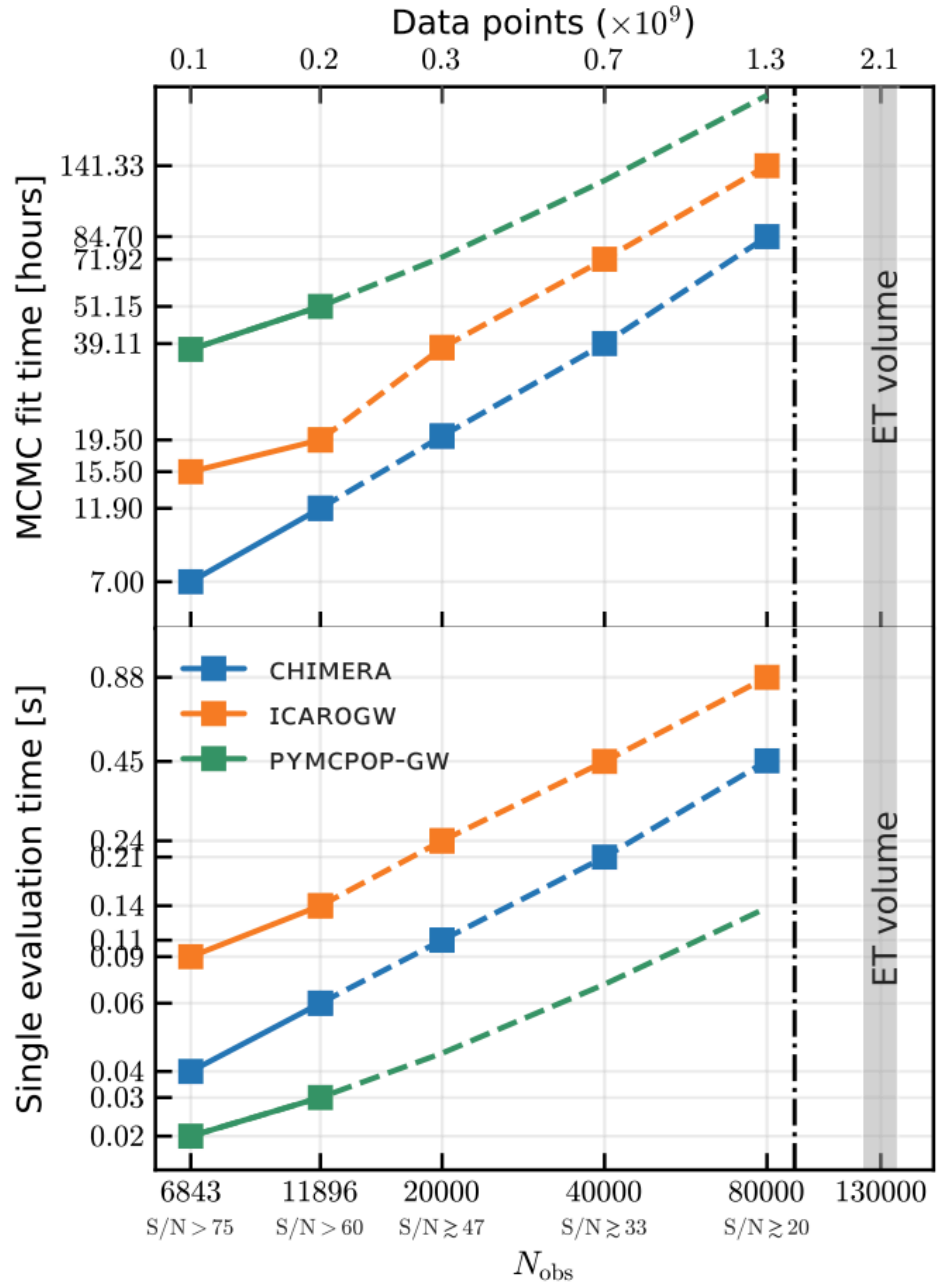
- Some simulations with GPU-compatible code push up to ~10k events

  - GPU memory saturation prevents scaling up to full ET data volume
  - Selection effect accuracy is likely underestimated
  - More work and resources needed to move towards more realism

- **Realistic scenarios needed - work for Div 10, 9 +3**

  - on simulation “whatever works” - we know the generative model, we do not see issues with selection uncertainty, no systematics

  - **connect MDC, population models and hierarchical inference for realistic assessments. CURRENTLY ABSENT FROM ET/3G PLANNING, BUT ARGUABLY THE MOST IMPORTANT AREA FOR 3G ASTRO/COSMO**



Tagliazucchi + 2026

(iv) Core statistical issues/computing issues

- Current models assume events are **non-overlapping** and statistically independent. What is the full likelihood if not ? Is such analysis feasible? Will it limit us to high-SNR sources (not enough) ?

- Do we really need selection bias as we compute it now?

- What about **stochastic background**? Non-resolved sources are useful for the high-redshift end of the population. Full likelihood currently not clearly formulated.

- **Likelihood-based vs simulation-based**? (simulation-based solves the above two issues… but is limited by the next one)

- More realistic assessment of **computational cost?** Parallel GPU seems mandatory

- Can **global-fit-like approaches** be extended to population ? (remember population knowledge re-shapes also single-event knowledge!)