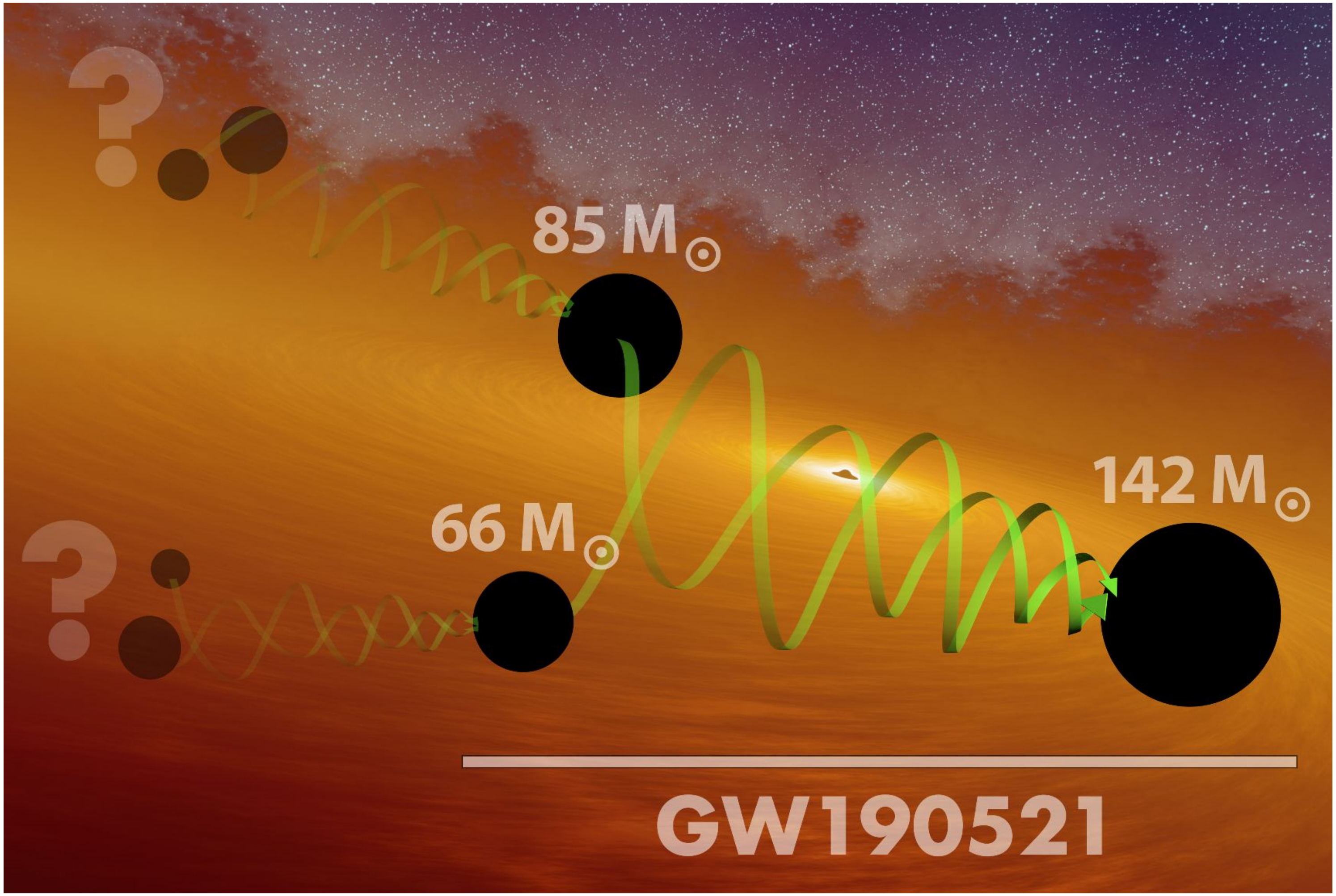
Deep learning and Bayesian inference for gravitational-wave population analysis

Matthew Mould
University of Birmingham, UK
mmould@star.sr.bham.ac.uk
arXiv:2203.03651

Davide Gerosa (Milano-Bicocca)
Stephen R. Taylor (Vanderbilt)

Milano-Bicocca
27/04/2022

Where do black holes merge?



LIGO/Caltech/MIT/R. Hurt (IPAC)

From individual sources to populations

When we detect a gravitational-wave event, we want to ask...

How heavy are the component black holes? How rapidly did they spin?

$$p(\boldsymbol{\theta}_n|\boldsymbol{d}_n) = \frac{\pi(\boldsymbol{\theta}_n)\mathcal{L}(\boldsymbol{d}_n|\boldsymbol{\theta}_n)}{\mathcal{Z}(\boldsymbol{d}_n)}, \quad \boldsymbol{\theta}_n = (\mathrm{masses}, \mathrm{spins}, \ldots)_n$$

We make many detections and then combine them: $n = 1, \ldots, N$

How likely is it that we made those draws from the population?

$$\mathcal{L}(\{\boldsymbol{d}_n\}|\boldsymbol{\lambda}) = \prod_{n=1}^{N} \frac{\int \mathcal{L}(\boldsymbol{d}_n|\boldsymbol{\theta}_n)\pi(\boldsymbol{\theta}_n|\boldsymbol{\lambda})\mathrm{d}\boldsymbol{\theta}_n}{\int P(\mathrm{detect}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{\lambda})\mathrm{d}\boldsymbol{\theta}}$$

We love Bayes' theorem!

Someone has already done the hard work for us to infer the Bayesian posteriors for the individual sources, so let's use that information…

We use Bayes' theorem *twice* to find the population posterior:

$$p(\boldsymbol{\lambda}|\{\boldsymbol{d}_n\}) = \pi(\boldsymbol{\lambda}) \prod_{n=1}^{N} \frac{1}{\sigma(\boldsymbol{\lambda})} \int \frac{\pi(\boldsymbol{\theta}_n|\boldsymbol{\lambda})}{\pi(\boldsymbol{\theta}_n)} p(\boldsymbol{\theta}_n|\boldsymbol{d}_n) \mathrm{d}\boldsymbol{\theta}_n$$

Now we need to **model the population** and **model detection biases**.

$$\pi(\boldsymbol{\theta}|\boldsymbol{\lambda}) \qquad \sigma(\boldsymbol{\lambda})$$

They depend on the **population parameters**, also called **hyperparameters**.

Population models head-to-head

An example **phenomenological** model: $\theta = m, \lambda = \alpha$

- Mass gap between neutron stars and black holes
- Mass gap above the PISN threshold
- IMF is a power law

$$p(m|\alpha) \propto m^{\alpha} \quad \text{if} \quad m_{\min} < m < m_{\max}$$

Simulation-based model: $\theta = m, \lambda = \alpha_{\mathrm{CE}}$

Interpolation

$$\begin{bmatrix} \alpha_{\mathrm{CE},1} \\ \vdots \\ \alpha_{\mathrm{CE},n} \end{bmatrix} \Rightarrow$$

Simulator e.g., pop. synth.

$$\Rightarrow \begin{bmatrix} \{m\}_1 \\ \vdots \\ \{m\}_n \end{bmatrix} \sim \begin{bmatrix} p_{\mathrm{pop}}(m|\alpha_{\mathrm{CE},1}) \\ \vdots \\ p_{\mathrm{pop}}(m|\alpha_{\mathrm{CE},n}) \end{bmatrix}$$

Parents

(Gerosa+Berti 2019)

- Generate 500 *clusters* with **escape speeds**

$$p(v_{\rm esc}) \propto v_{\rm esc}^{\boxed{\delta}}$$
$$v_{\rm esc} \in [0, 500]\,{\rm km\,s^{-1}}$$

- In each cluster, generate 5000 **first-generation** black holes

$$p(m) \propto m^{\boxed{\gamma}}, \quad m \in [5, \boxed{m_{\rm max}}]M_\odot$$
$$p(\chi) = \mathcal{U}(0, \boxed{\chi_{\rm max})}$$

- Spin directions are isotropic

Results in 1000 complex numerical distributions…

Marriage

- Draw two black holes and **pair** them into a binary

$$p(m_1) \propto m_1^{\boxed{\alpha}}$$
$$p(m_2|m_1) \propto m_2^{\boxed{\beta}}\,(m_2 \leq m_1)$$

Babies

- Estimate the **remnant** properties (Gerosa+Kesden 2016)
- **Retain** the remnant BH with mass and spin magnitude if

$$v_{\rm kick} < v_{\rm esc}$$

Filling the gaps in the population

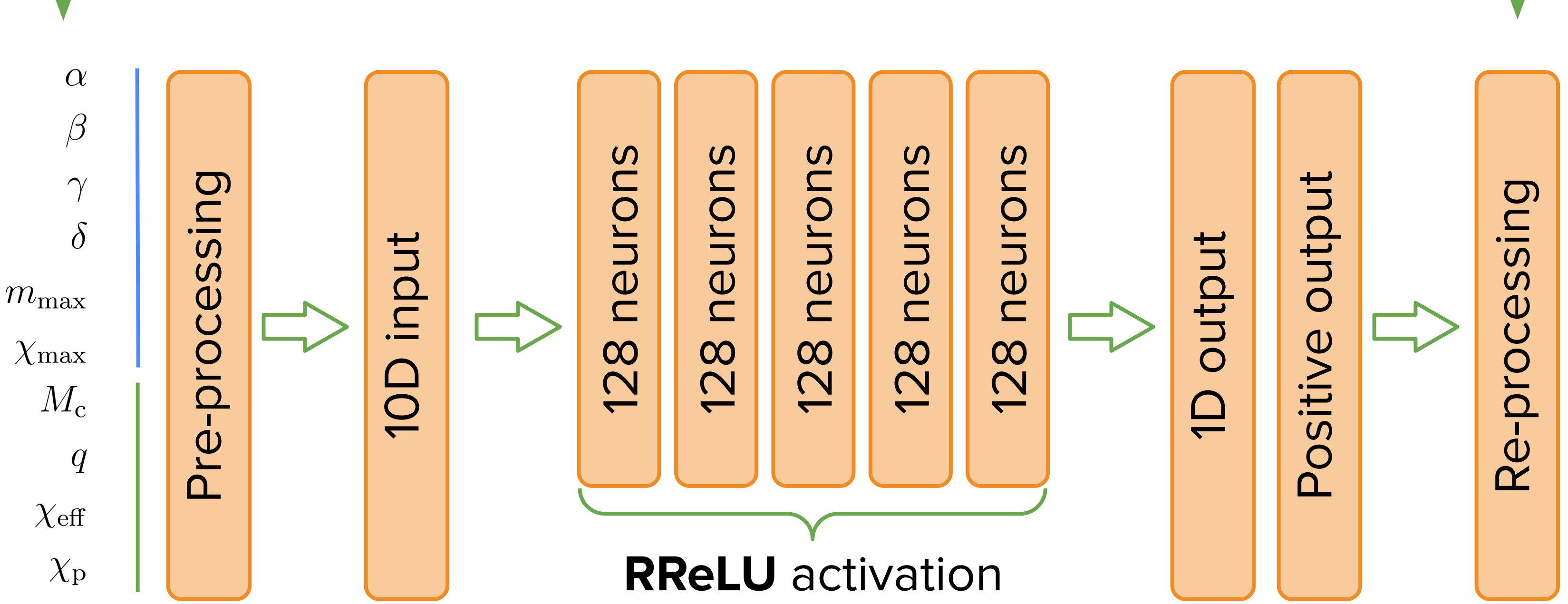
- Estimate each population density with FFT **kernel density estimation**

$$\theta \mapsto p_{\text{pop}}(\theta|\lambda^i),\ i = 1, ..., 1000$$

(KDEpy)
(Botev+ 2010)

- Learn the full mapping with a **deep neural network** (TensorFlow)

$$(\theta, \lambda) \mapsto p_{\text{pop}}(\theta|\lambda)$$



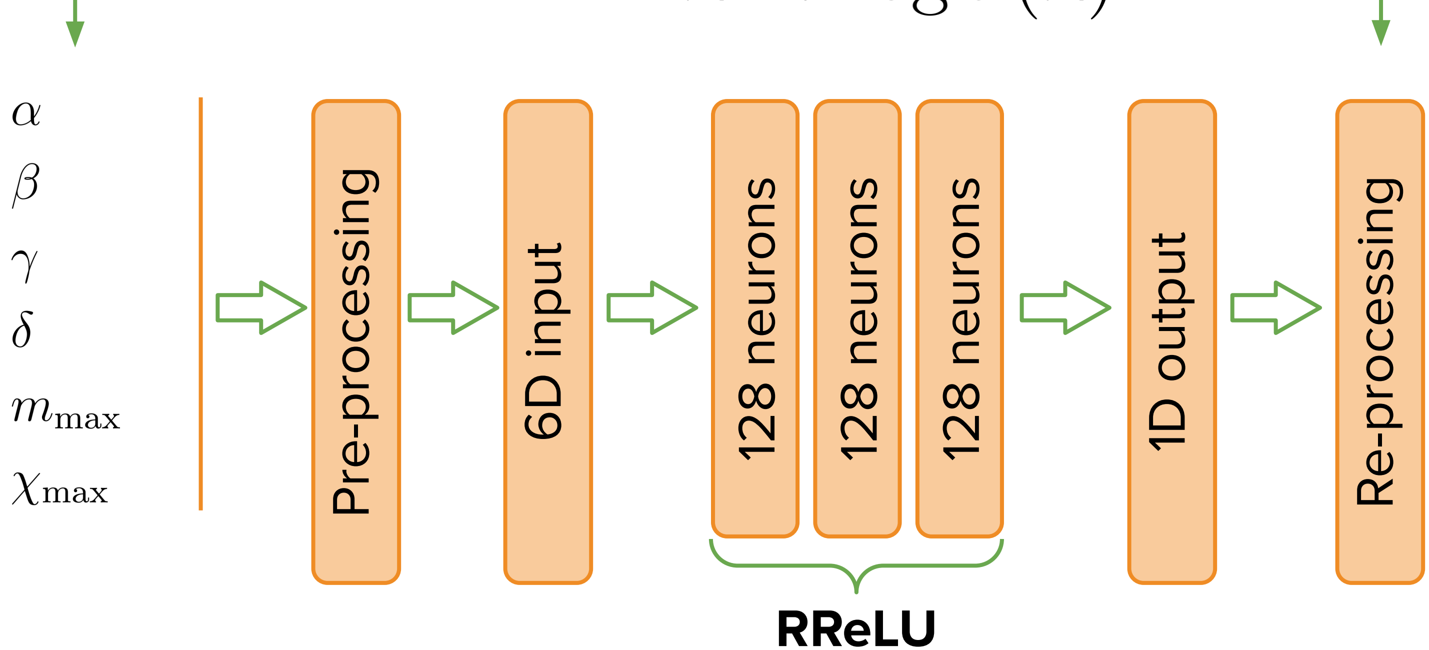
Filling the gaps in the selection function

- Estimate detection efficiency for each simulation (Finn+Chernoff 1993)

$$\sigma(\lambda^i) = \frac{1}{N^i} \sum_{j=1}^{N^i} P_{\mathrm{det}}(\theta_j^i), \;\; i = 1, \ldots, 1000$$

- Learn the full mapping with a deep neural network (TensorFlow)

$$\lambda \mapsto \log \sigma(\lambda)$$



- Ultimately work with the *log*-likelihood
- Also results in better learning

Filling the gaps in the selection function

- Compute fraction of sources in each merger generation $\sum_g f_g(\lambda) = 1$

$$f_g(\lambda^i) = N_g^i / N^i,\; i = 1, \ldots, 1000$$

- Learn the full mapping with a deep neural network [(TensorFlow)]

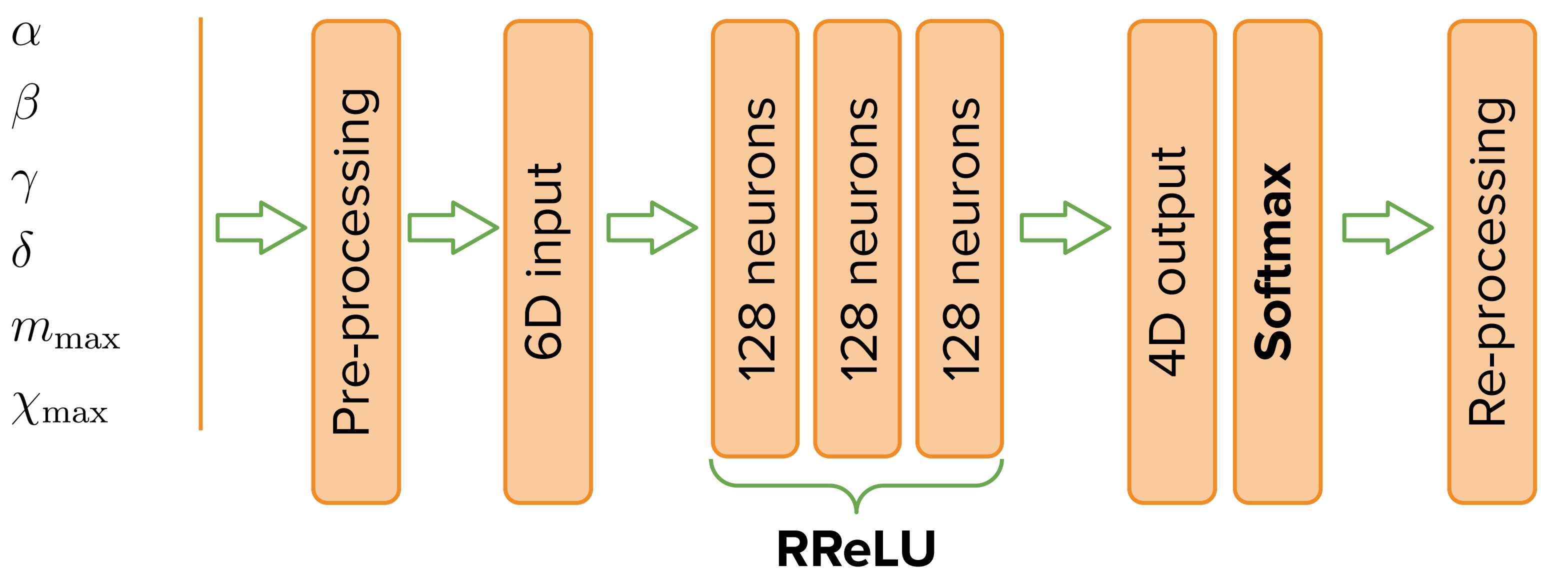
$$\lambda \mapsto f_g(\lambda)$$

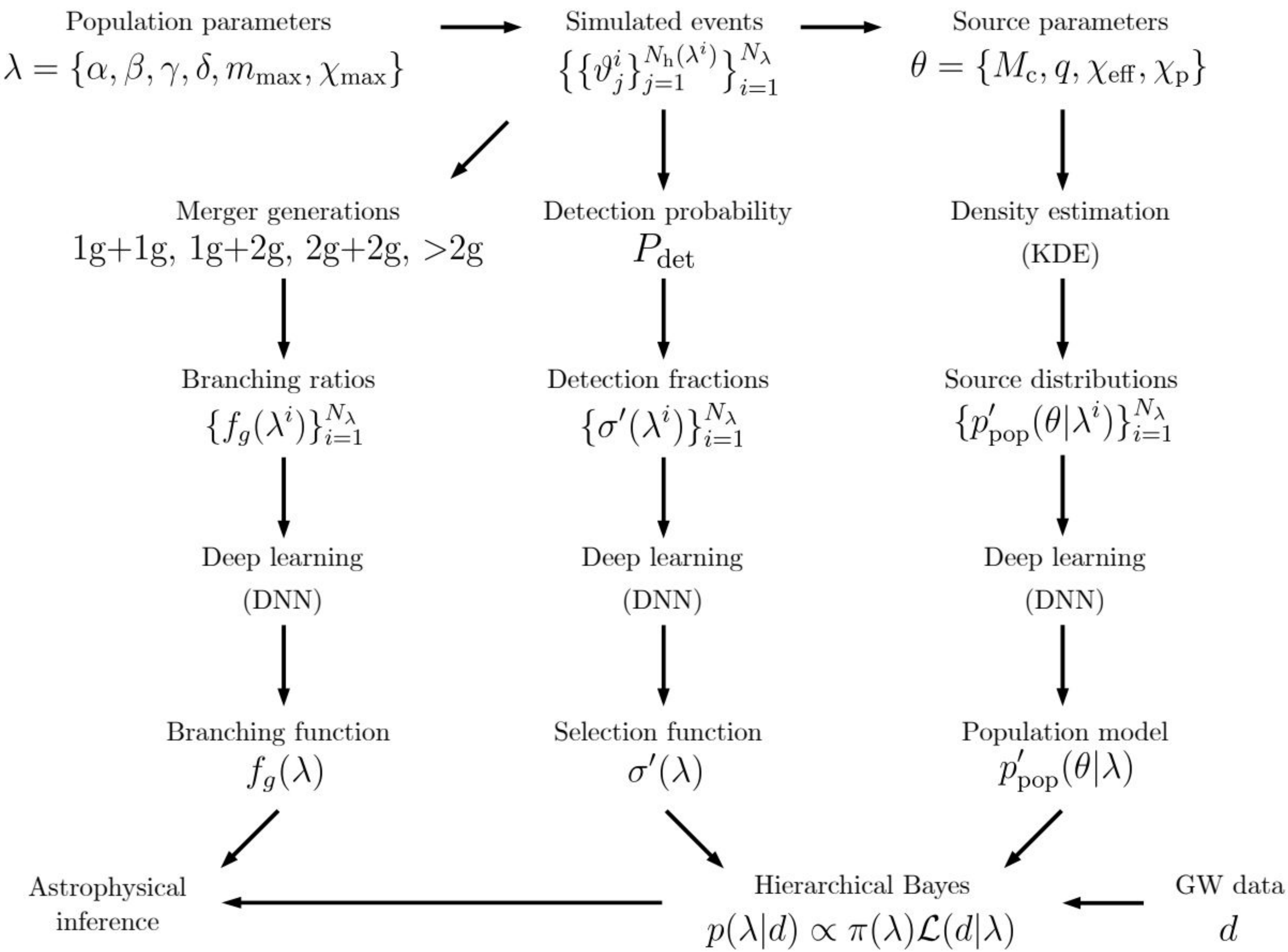
$1\mathrm{g} + 1\mathrm{g}$

$1\mathrm{g} + 2\mathrm{g}$

$2\mathrm{g} + 2\mathrm{g}$

$> 2\mathrm{g}$

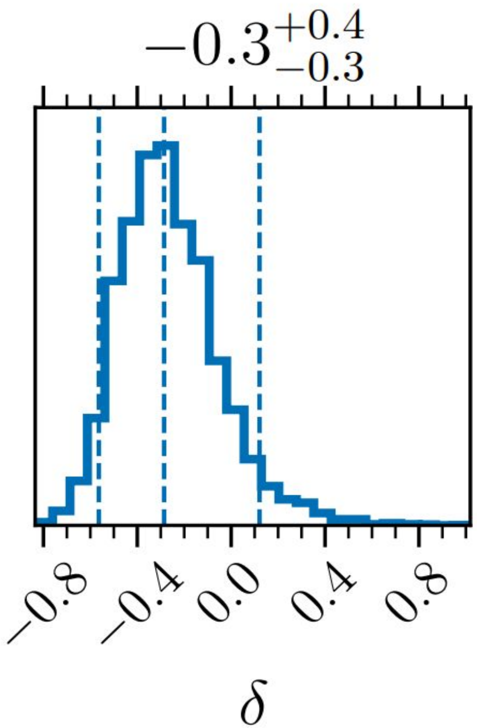


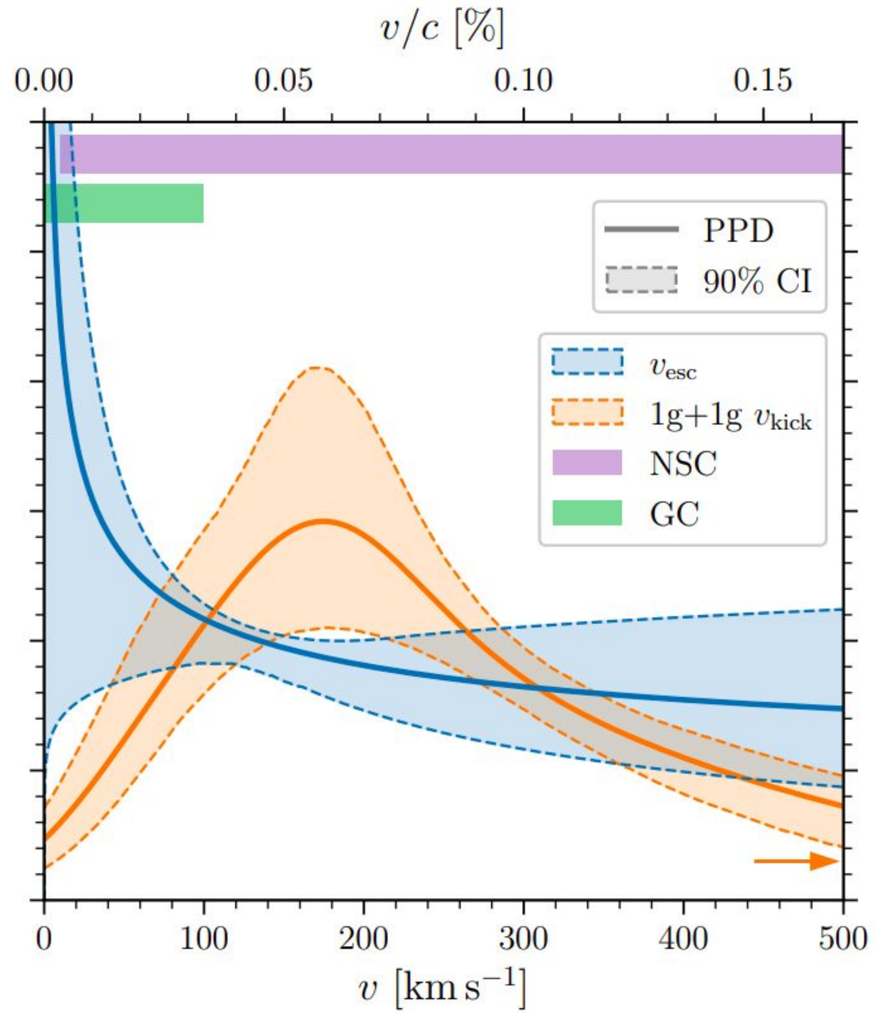

Population parameters
$\lambda = \{\alpha, \beta, \gamma, \delta, m_{\max}, \chi_{\max}\}$
Simulated events
$\{\{\vartheta^i_j\}_{j=1}^{N_{\mathrm{h}}(\lambda^i)}\}_{i=1}^{N_\lambda}$
Source parameters
$\theta = \{M_{\mathrm{c}}, q, \chi_{\mathrm{eff}}, \chi_{\mathrm{p}}\}$
Merger generations
1g+1g, 1g+2g, 2g+2g, >2g
Detection probability
P_{det}
Density estimation
(KDE)
Branching ratios
$\{f_g(\lambda^i)\}_{i=1}^{N_\lambda}$
Detection fractions
$\{\sigma'(\lambda^i)\}_{i=1}^{N_\lambda}$
Source distributions
$\{p'_{\mathrm{pop}}(\theta|\lambda^i)\}_{i=1}^{N_\lambda}$
Deep learning
(DNN)
Deep learning
(DNN)
Deep learning
(DNN)
Branching function
$f_g(\lambda)$
Selection function
$\sigma'(\lambda)$
Population model
$p'_{\mathrm{pop}}(\theta|\lambda)$
Astrophysical
inference
Hierarchical Bayes
$p(\lambda|d) \propto \pi(\lambda)\mathcal{L}(d|\lambda)$
GW data
d

Host escape speeds

$$p(v_{\mathrm{esc}}|\delta) \propto v_{\mathrm{esc}}^{\delta}$$

$$v_{\mathrm{esc}} \in (0, 500)\ \mathrm{km\,s^{-1}}$$





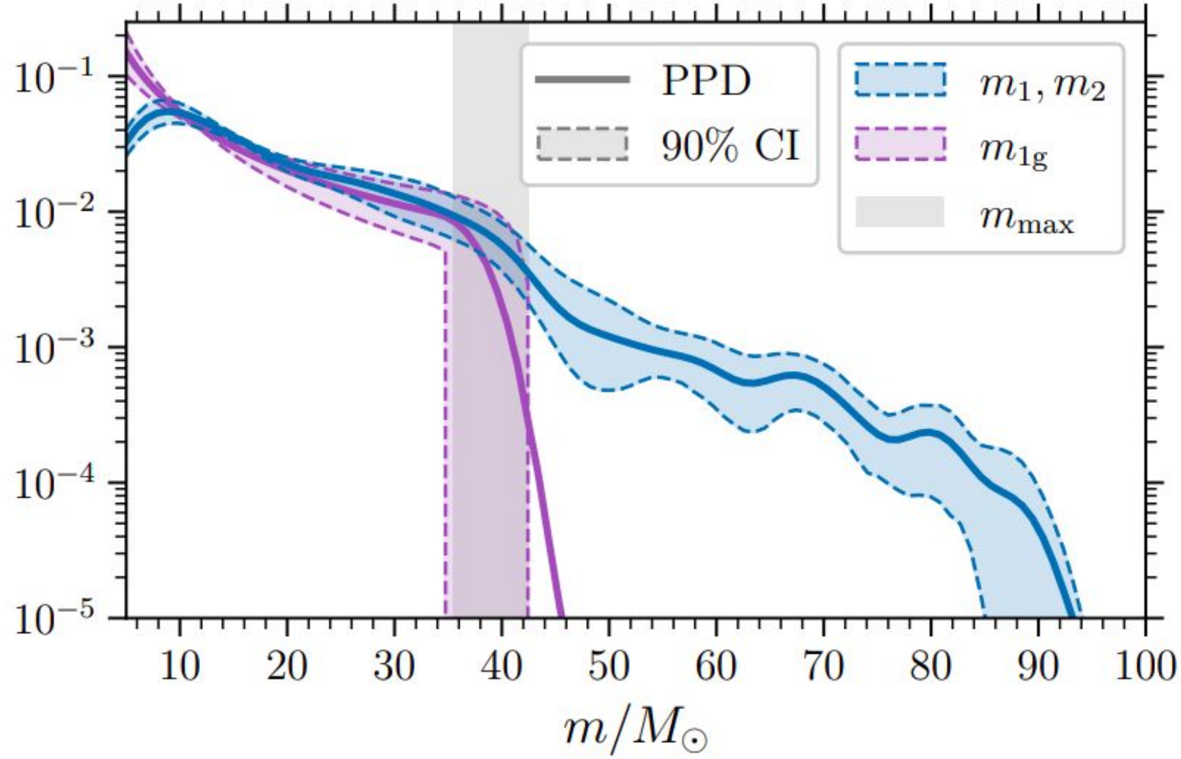
$$1\mathrm{g} + 1\mathrm{g}:$$

$$P(v_{\mathrm{kick}} < 500\ \mathrm{km\,s^{-1}}) = 0.85^{+0.06}_{-0.07}$$

$$P(v_{\mathrm{kick}} < v_{\mathrm{esc}}) = 0.39^{+0.12}_{-0.12}$$

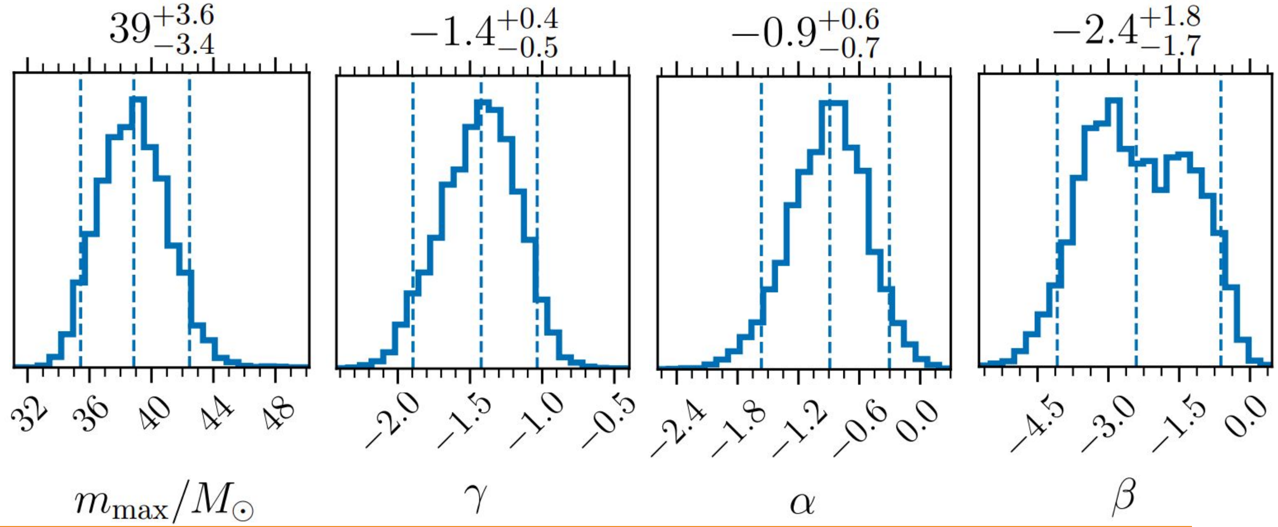
Masses

$$\text{99\% quantile}$$
$$m = 60^{+7.2}_{-6.5} M_\odot$$



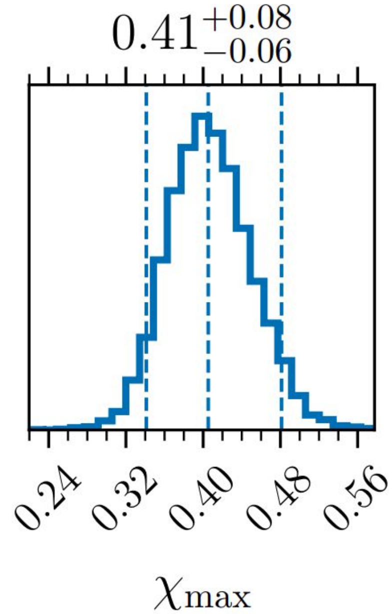
$$p(m_{1g}|\gamma, m_{max}) \propto m_{1g}^{\gamma}$$
$$m_{1g} \in (5M_\odot, m_{max})$$

$$p(m_1|\alpha) \propto m_1^{\alpha}$$
$$p(m_2|\beta, m_1) \propto m_2^{\beta}$$
$$m_2 \leq m_1$$



Spins

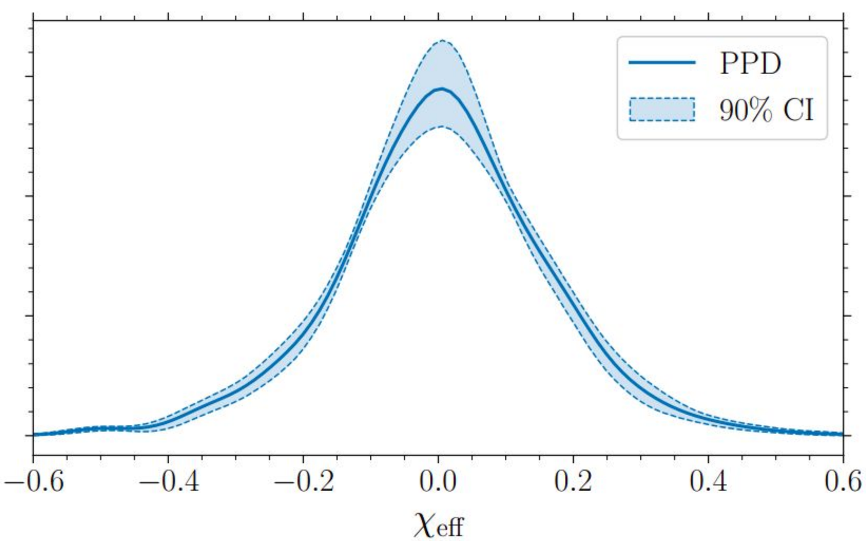
$$p(\chi_{1g}|\chi_{\max}) = 1/\chi_{\max}$$
$$\chi \in [0, \chi_{\max}]$$
$$p(\cos\theta_{1,2}) = 1/2$$

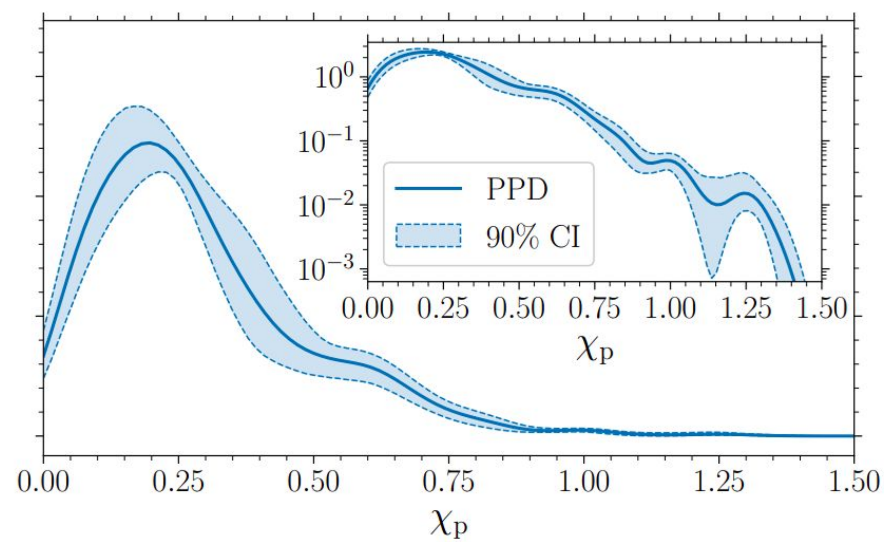


99% quantile
$$\chi_{\mathrm{p}} = 0.88^{+0.13}_{-0.07}$$

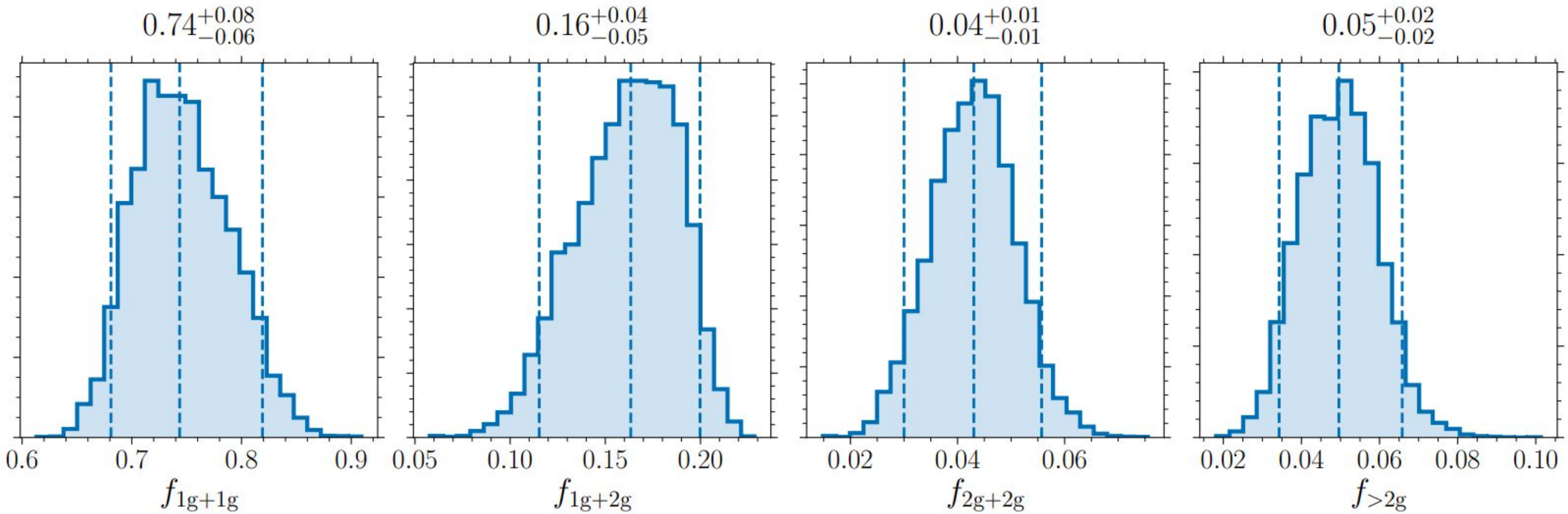
99% quantile
$$|\chi_{\mathrm{eff}}| = 0.46^{+0.04}_{-0.05}$$

$$P(\chi_{\mathrm{p}} > 1) = 0.6^{+0.5}_{-0.3}\%$$





Merger generations



$$99\% \text{ quantile}$$

$$1 - f_{1g+1g} = 0.15$$