Explainable Deep-learning: Monte Carlo methods for Gravitational-Wave Inference

Project No: 628
SUPA, School of Physics and Astronomy
University of Glasgow
Glasgow G12 8QQ, United Kingdom

(Dated: March 24, 2021)

My 250 word abstract goes here...

TABLE I.

parameter name	symbol	status	value	units
mass 1	m_1	inferred	35-80	$M_\odot$
mass 2	m_2	inferred	35-80	$M_\odot$
luminosity distance	d_L	inferred	1-3	Gpc
time of coalescence	t_0	inferred	0.65-0.85	s
inclination	Θ_{jn}	inferred	0-π	rad
polarisation	ψ	inferred	0-π	rad
right ascension	α	fixed	0	rad
declination	δ	fixed	0	rad
spins	-	fixed	0	-
phase at coalescence	ϕ_0	marginalised	0-2π	rad

I. INTRODUCTION

Remember to signpost rest of paper at end of this section!

A. Parameter Estimation

$$p(x|y) = \frac{\mathcal{L}(y|x)p(x)}{\mathcal{Z}}, \tag{1}$$

$$p(x|y) \propto \mathcal{L}(y|x)p(x), \tag{2}$$

$$\mathcal{Z} = \int dx \mathcal{L}(y|x)p(x), \tag{3}$$

Need to know if equation is at end of sentence then swap comma for full stop

B. Deep-learning Approaches

C. VItamin: User-friendly Inference

$$H(p,r) = -\int dx\, p(x|y) \log r_\theta(x|y), \tag{4}$$

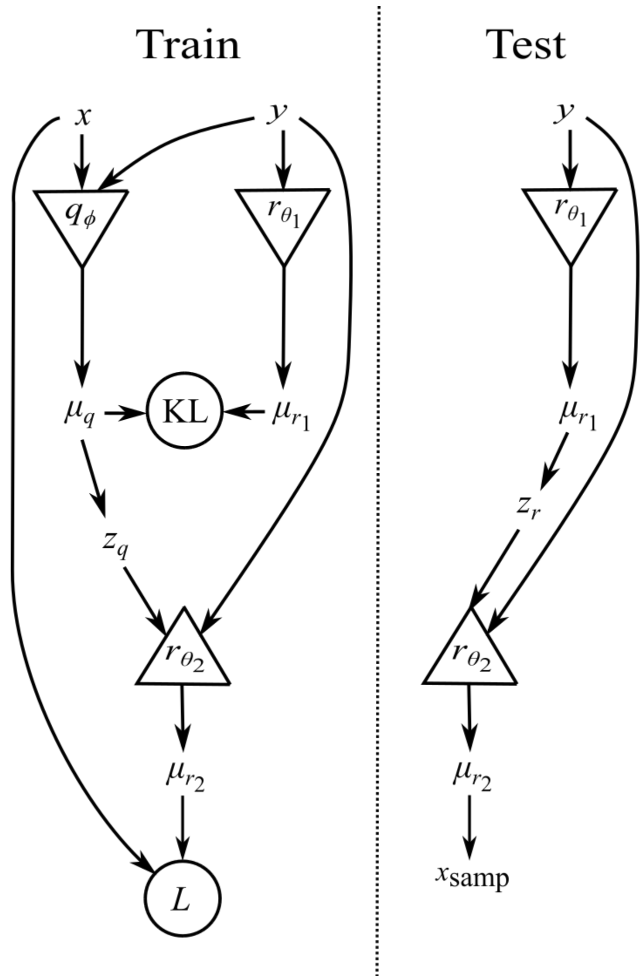


FIG. 1.

$$r_\theta(x|y) = \int dz\, r_{\theta_1}(z|y) r_{\theta_2}(x|y,z). \tag{5}$$

Use gen pap to intro CVAE in context, CONTEXT IS KEY HERE

Need to mention metropolis hastings it seems!

Introduce equations directly to our specifics, we don't have space to intro them blind then again to specifics...

Do theory on normal IS and then say that SIR is an monte carlo approach/approx to normal IS then give equations for bot (talk about the NEW IMPROVED SIR method (link to Section **??**))

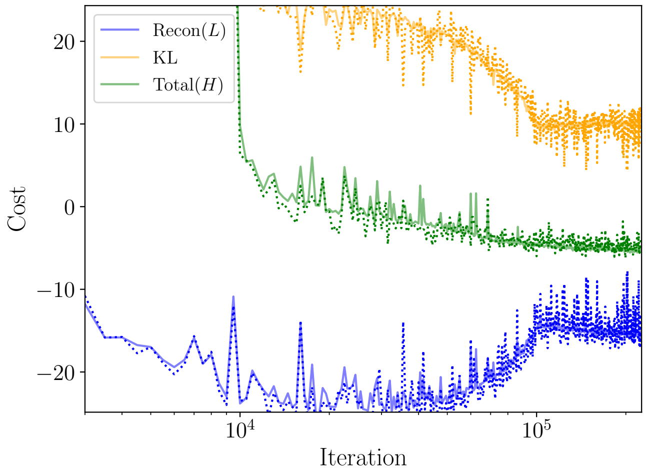


FIG. 2.

II. METHODOLOGY

Apply the intro/theory mateiral to our case, JUSTIFY scientific decisions like number of samples, batch size, npars!!

A. Model Training

$$H \lesssim \frac{1}{N}\sum_{n=1}^{N_\mathrm{b}}\Big[\overbrace{-\log r_{\theta_2}(x_n|z_n, y_n)}^{L} + \overbrace{\mathrm{KL}\left[q_\phi(z|x_n, y_n)||r_{\theta_1}(z|y_n)\right]}^{\mathrm{KL}}\Big], \tag{6}$$

B. Likelihood Estimates

$$\begin{aligned} r_\theta(x|y) =& \mathbb{E}_{r_{\theta_1}(z|y)} r_{\theta_2}(x|y,z) \\ \approx& \frac{1}{N}\sum_{j=1}^{N} r_{\theta_{2,j}}(x_i|y, z_j)\big|_{z_j\sim r_{\theta_1}(z_j|y)}, \end{aligned} \tag{7}$$

$$r_\theta(x|y) \sim \mathcal{L}_\theta(y|x), \tag{8}$$

Figs: Monte flowchart

C. Likelihood Reweighting

Start with eq 2 then multiply by unity

$$\begin{aligned} p(x|y) &\propto \frac{\mathcal{L}_\theta(y|x)}{\mathcal{L}_\theta(y|x)}\mathcal{L}(y|x)p(x) \\ &\propto w(y|x)\,\overbrace{\mathcal{L}_\theta(y|x)p(x)}^{r_\theta(x|y)}. \end{aligned} \tag{9}$$

Here,

$$w(y|x) \equiv \frac{\mathcal{L}(y|x)}{\mathcal{L}_\theta(y|x)}, \tag{10}$$

is the weight function...

III. RESULTS

A. Self-consistency

B. Reproducibility

$$\sigma \propto \frac{1}{\sqrt{N}}, \tag{11}$$

where N is monte carlo batch size

$$\mathrm{Error} = \frac{\sigma}{\sqrt{n_\mathrm{samples}}}, \tag{12}$$

C. Importance Resampling

IV. CONCLUSIONS

This is section has to encapsulate everything we did so that after the abstract a reader can go here and see if they want to buy the paper or not!

As we find ourself in a proof-of-concept mode, there is justification of a section dedicated to the next steps leading towards production of this code.

ACKNOWLEDGEMENTS

Thanks to Chris and Hunter and Michael and Daniel.

Paragraph on the software used Bilby [1] [2]help please dont fuck up now , lets just keep typing and see what happens, im really shitting ymself now... and dont want to have to spend precious time on fucking type setting and bibliography hunting [3]

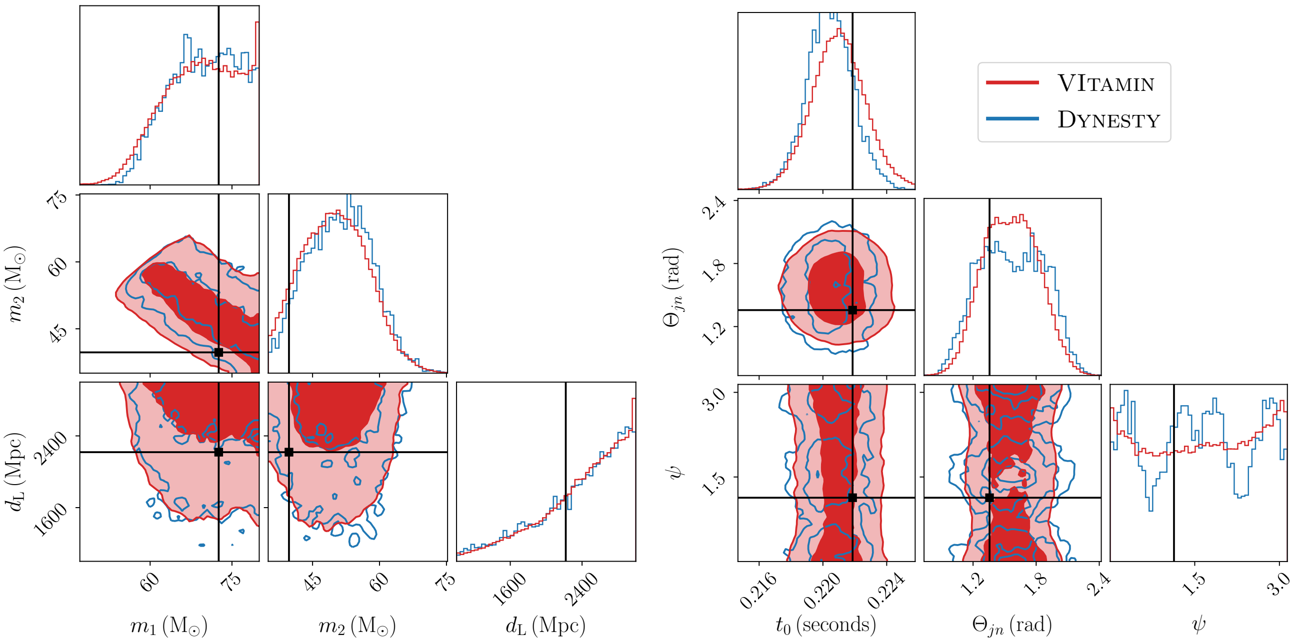


FIG. 3. Probability-probability (P-P) plot showing the confidence interval versus the fraction of the events within that confidence interval for the posterior distributions obtained using our analysis Nessai for 128 simulated compact binary coalescence signals produced with Bilby and bilby_pipe. The 1-, 2- and 3-σ confidence intervals are indicated by the shaded regions and p-values are shown for each of the parameters and the combined p-value is also shown.

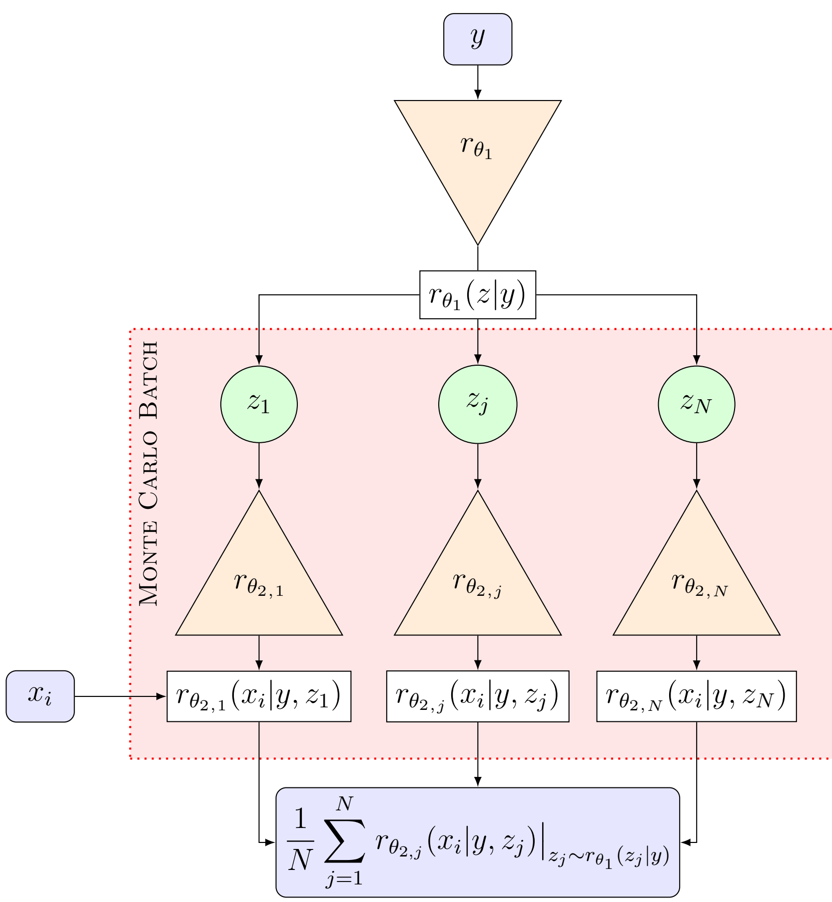


FIG. 4.

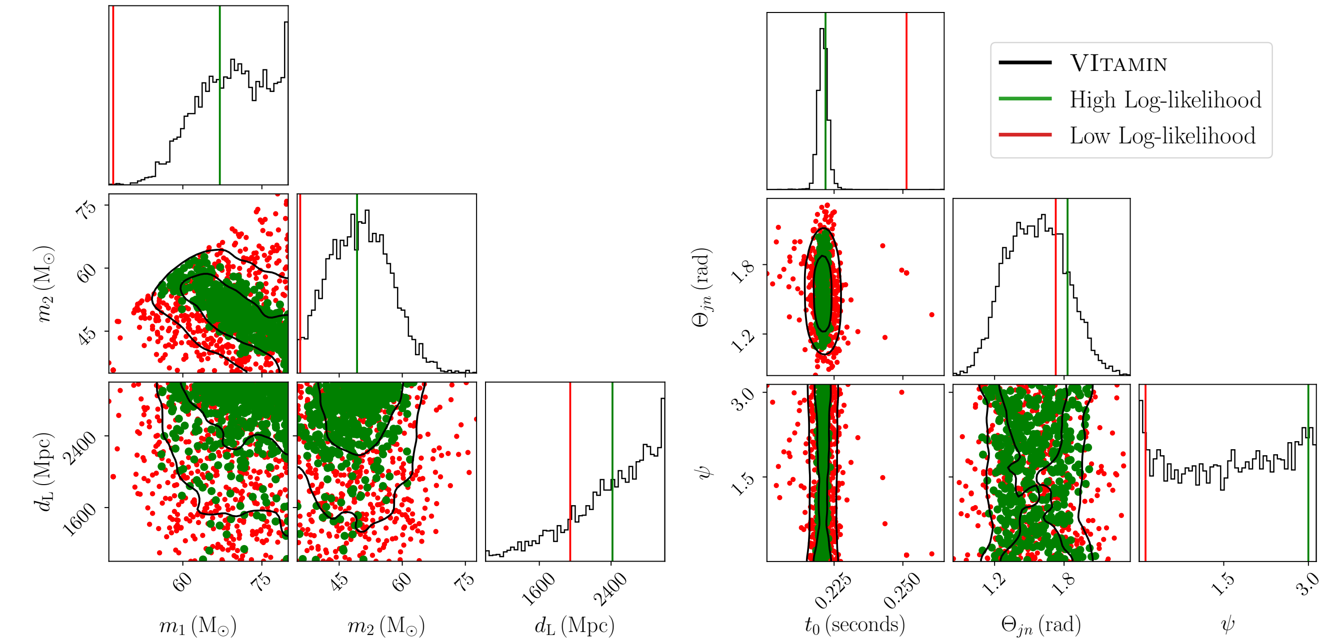


FIG. 5. Probability-probability (P-P) plot showing the confidence interval versus the fraction of the events within that confidence interval for the posterior distributions obtained using our analysis Nessai for 128 simulated compact binary coalescence signals produced with Bilby and bilby_pipe. The 1-, 2- and 3-σ confidence intervals are indicated by the shaded regions and p-values are shown for each of the parameters and the combined p-value is also shown.

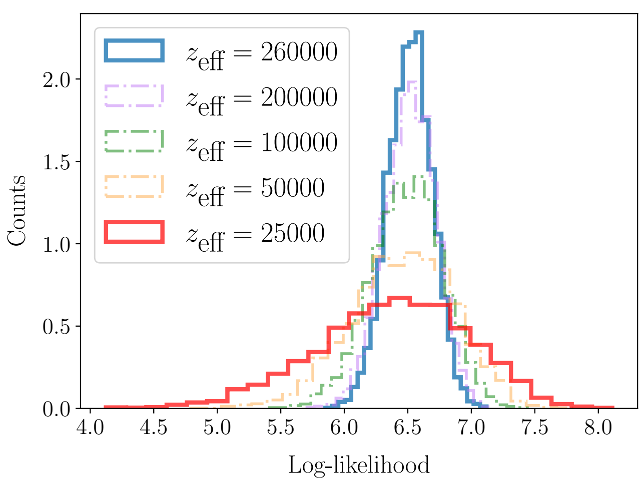


FIG. 6.

[1] G. Ashton, M. Huebner, P. D. Lasky, Colm Talbot, K. Ackley, Sylvia Biscoveanu, Q. Chu, A. Divarkala, P. J. Easter, Boris Goncharov, Francisco Hernandez Vivanco, J. Harms, M. E. Lower, Grant D. Meadors, D. Melchor, E. Payne, M. D. Pitkin, J. Powell, N. Sarin, Rory J. E. Smith, and E. Thrane, Astrophys. J. Supp. **241**, 27 (2019).

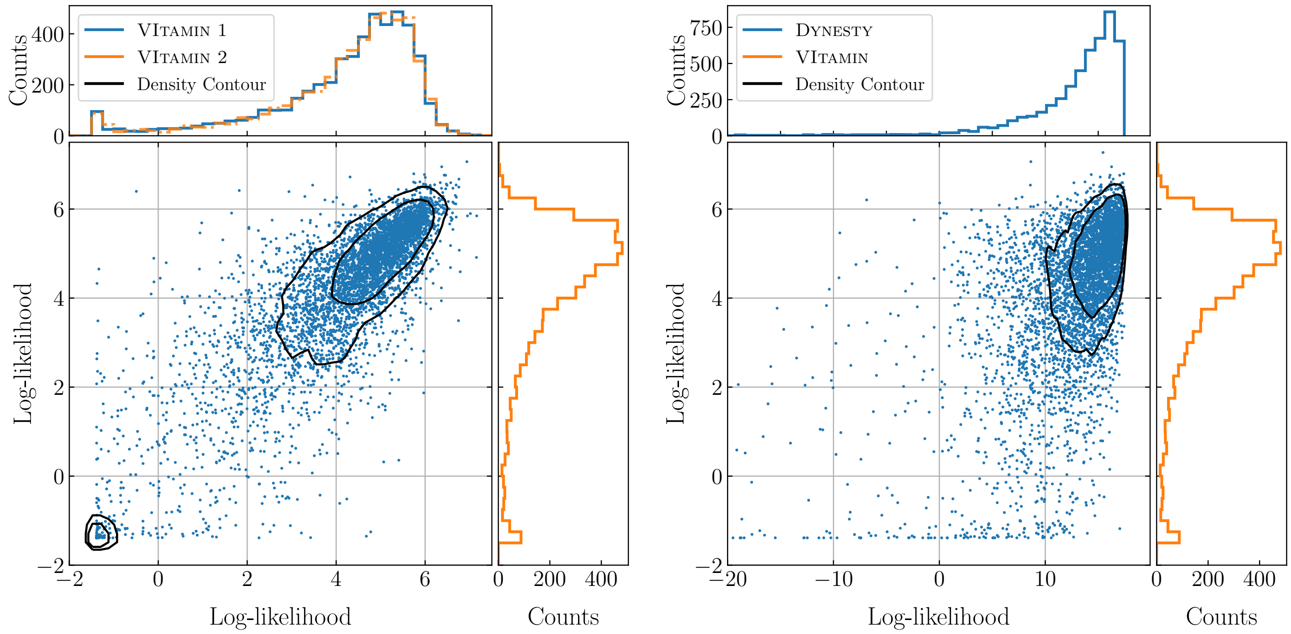
[2] K. Cannon, R. Cariou, A. Chapman, M. Crispin-Ortuzar, N. Fotopoulos, M. Frei, C. Hanna, E. Kara, D. Keppel, L. Liao, S. Privitera, A. Searle, L. Singer, and A. Weinstein, The Astrophysical Journal **748**, 136 (2012).

[3] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, PASP **125**, 306 (2013), 1202.3665.

[4] W. K. Hastings, (1970).

[5] W. Vousden, W. M. Farr, and I. Mandel, (2015), 10.1093/mnras/stv2422, arXiv:1501.05823.

[6] J. S. Speagle, "dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evi-



(a) (b)

FIG. 7. Probability-probability (P-P) plot showing the confidence interval versus the fraction of the events within that confidence interval for the posterior distributions obtained using our analysis Nessai for 128 simulated compact binary coalescence signals produced with Bilby and bilby_pipe. The 1-, 2- and 3-σ confidence intervals are indicated by the shaded regions and p-values are shown for each of the parameters and the combined p-value is also shown.

TABLE II. Generic Caption like in vitamin paper!

sampler	algorithm	deep-learning	run-time (s)[a]
emcee [3]	MCMC [4]	X	32070
ptemcee [5]	MCMC	X	24372
dynesty [6]	NS [7]	X	19400
cpnest [8]	NS	X	26202
Nessai [9]	NS	✓	9372
VItamin [10]	VI [11]	✓	1×10^{-1}

[a] The benchmark samplers all produced $\mathcal{O}(10000)$ samples dependent on the default sampling parameters used, run-time for depp-learning algorihms does not include trainign time

dences," (2019).

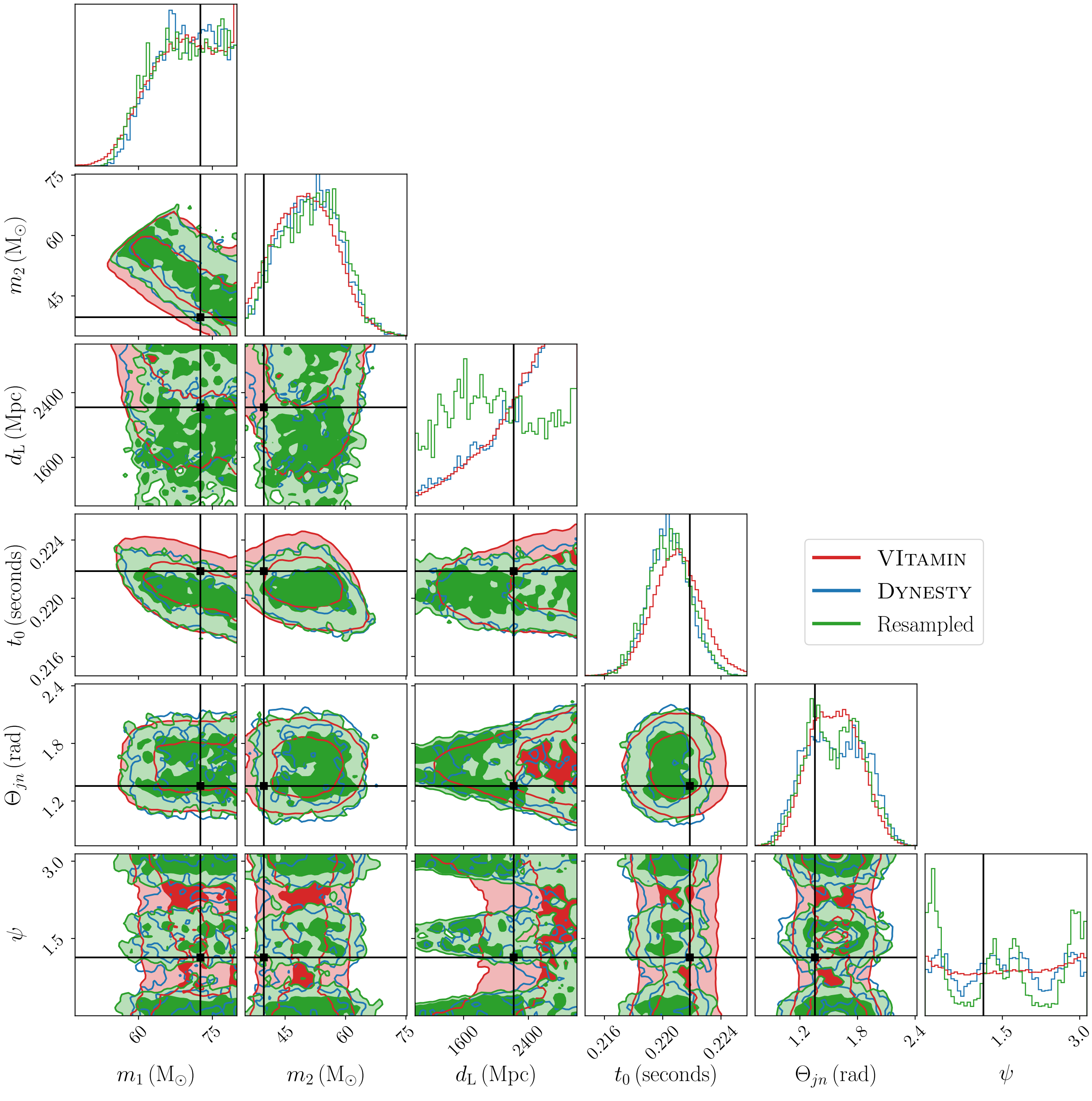
[7] J. Skilling, Bayesian Anal. **1**, 833 (2006).

[8] W. D. Pozzo and J. Veitch, "Cpnest," (2019).

[9] M. J. Williams, J. Veitch, and C. Messenger, arXiv preprint arXiv:2102.11056 (2021).

[10] H. Gabbard, C. Messenger, I. S. Heng, F. Tonolini, and R. Murray-Smith, arXiv preprint arXiv:1909.06296 (2019).

[11] F. Tonolini, A. Lyons, P. Caramazza, D. Faccio, and R. Murray-Smith, "Variational inference for computational imaging inverse problems," (2019), to appear in JMLR, arXiv:1904.06264.



VItamin
Dynesty
Resampled
$m_1\,(\mathrm{M}_\odot)$
$m_2\,(\mathrm{M}_\odot)$
$d_\mathrm{L}\,(\mathrm{Mpc})$
$t_0\,(\mathrm{seconds})$
$\Theta_{jn}\,(\mathrm{rad})$
ψ

FIG. 8.