



## COMPAS Star-formation LnL GP surrogate

AVI VAJPEYI,<sup>1,2,3</sup> JEFF RILEY,<sup>4,3</sup> ILYA MANDEL,<sup>4,3</sup> AND CHAYAN CHATTERJEE<sup>5</sup>

<sup>1</sup>*UoA*

<sup>2</sup>*NZ Gravity*

<sup>3</sup>*ARC Centre of Excellence for Gravitational Wave Discovery – OzGrav, Australia*

<sup>4</sup>*School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia*

<sup>5</sup>

### ABSTRACT

We’re using GPs to build surrogate likelihoods for COMPAS star formation parameters, given the LVK observed dataset.

*Keywords:* surrogate model – inference – binaries – stars: evolution – gravitational waves – machine learning – Gaussian Processes

### 1. INTRODUCTION

Over the past decade, improvements in detector sensitivity have led to an exponential increase in the rate of GW detections. With next-generation detectors on the horizon (such as CE, ET, and LISA), the number of GW detections will increase even more. The most recent catalogue of GW contains XX events, the majority of them being Binary Black Holes (BBHs), with XX public alerts for GW that have yet to be confirmed as astrophysical events. As more GW events are observed, more features in the distribution of GW events await to be discovered. One of the most common ways to understand the observed GW population is through phenomenological modelling. However, as the size of the catalogues increases, so does the catalogue complexity, making it challenging to develop interpretable, parameterisable phenomenological models, flexible enough to explain the features in the catalogue.

Alternatively, we can consider using forward models to simulate catalogues of GW events and compare the simulations to observed catalogues. The forward modelling approach would provide astronomers with constraints on easy-to-interpret physical parameters that describe the initial conditions of the population of merging binaries, such as distributions describing the initial stellar mass, metallicity, and binary separation. These param-

eters also govern simulated physical processes like wind mass loss, supernova characteristics, and binary interactions such as mass transfer and common envelope efficiency. However, the generation of synthetic catalogues from the various rapid population synthesis software packages can be expensive, making brute force methods that explore all possible parameter values expensive (for example, see Broekgaarden et al. (2022)). Another alternative, as proposed by Riley et al. (2023) involves building surrogate tools to bridge the computational complexities of simulating a state space large enough to allow us to infer astrophysical constraints from the observed catalogs.

Riley et al. (2023) generated a small number of “training” states using a targeted selection of initial conditions and evolutionary parameters. The training states were then used to teach an interpolant how to map initial conditions and evolutionary parameters to the final state, thus avoiding the computational time of calculating the intermediate steps.

In their work, they explored four parameters governing the cosmic metallicity-specific star formation rate (MSSFR) in the model of Neijssel et al. (2019).

We build upon (Riley & Mandel 2023)’s work, and

The remainder of this paper is organised as follows. Section 2 presents a description of the tool we construct for this proof-of-concept study, and the method used to train the tool. We present and discuss our results in Section 3. We provide some concluding remarks in Section 4

## 2. METHOD

### 2.1. *MSSFR parameters and $\mathcal{M} - z$ rates*

Following (Riley & Mandel 2023), we vary four free parameters for the calculation of the metallicity-specific star formation rate (MSSFR) in the phenomenological model of Neijssel et al. (2019). In this model, the MSSFR is split into two parts, the star formation rate (SFR) and the metallicity distribution:

$$\frac{d^3 M_{\text{SFR}}}{dt_s dV_c dZ}(z) = \frac{d^2 M_{\text{SFR}}}{dt_s dV_c}(z) \times \frac{dP}{dZ}(z), \quad (1)$$

where the SFR is given by:

$$\frac{d^2 M_{\text{SFR}}}{dt_s dV_c} = a \frac{(1+z)^b}{1 + \left(\frac{1+z}{c}\right)^d} M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}, \quad (2)$$

and we use a log-normal distribution in metallicity at each redshift (cf. the skewed log-normal model proposed by van Son et al. (2022)):

$$\frac{dP}{dZ}(z) = \frac{1}{Z\sigma\sqrt{2\pi}} \exp \left[ -\frac{(\ln Z - \ln \langle Z \rangle + \sigma^2/2)^2}{2\sigma^2} \right], \quad (3)$$

with a redshift-independent standard deviation  $\sigma$  in  $\ln Z$  space around a redshift-dependent mean  $\mu$  of  $\ln Z$  given by:

$$\langle Z \rangle = \exp[\mu + \sigma^2/2], \quad (4)$$

with mean metallicity parametrised as in Langer & Norman (2006):

$$\langle Z(z) \rangle = Z_0 10^{\alpha z} \quad (5)$$

We vary SFR parameters  $a$  and  $d$ , and the metallicity distribution parameters  $\alpha$  and  $\sigma$ . We fix  $b = 2.77$ ,  $c = 2.9$ , and  $Z_0 = 0.035$ .

Varying these parameters in COMPAS results in different merger rate distributions (see (Riley & Mandel 2023; Neijssel et al. 2019)).

Plots of the SFR, metallicity distributions and resultant merger rate distributions for different  $\mathcal{M}$  and  $z$  are shown in Figure 1 for three different choices of parameters.

### 2.2. *Bayesian Model*

We use the Bayesian model described in (Riley & Mandel 2023). Given  $N_{\text{obs}}$  observations from data  $\mathcal{D} \equiv \{D_1, D_2, \dots, D_{N_{\text{obs}}}\}$ , we define the likelihood  $\mathcal{L}$  that we will see data  $\mathcal{D}$  for a particular set of parameters  $\lambda$  as

$$\log \mathcal{L}(\mathcal{D}|\lambda) = \log \mathcal{L}_p(N_{\text{obs}}|\lambda) + \sum_{i=1}^{N_{\text{obs}}} \log \mathcal{L}_{\text{event}}(D_i|\lambda). \quad (6)$$

The first term of Equation 6 is a Poisson likelihood on there being  $N_{\text{obs}}$  detections,

$$\log \mathcal{L}_p(N_{\text{obs}}|\lambda) = -\mu(\lambda) + N_{\text{obs}} \log \mu(\lambda), \quad (7)$$

where  $\mu(\lambda)$  is the expected number of observations and we omit terms that depend on the data only and therefore disappear on normalisation, such as  $\log(N_{\text{obs}}!)$  and permutation coefficients.

The second term of Equation 6 is comprised of a product of the probabilities of individual event detections,

$$\mathcal{L}_{\text{event}}(D_i|\lambda) = p_{\text{grid}}(z = z_i, \mathcal{M} = \mathcal{M}_i|\lambda), \quad (8)$$

where  $p_{\text{grid}}(z = z_i, \mathcal{M} = \mathcal{M}_i|\lambda)$  is the entry in the probability distribution matrix  $p_{\text{grid}}(z, \mathcal{M}|\lambda)$  in the  $z$  and  $\mathcal{M}$  bin of the observed  $i^{\text{th}}$  event.

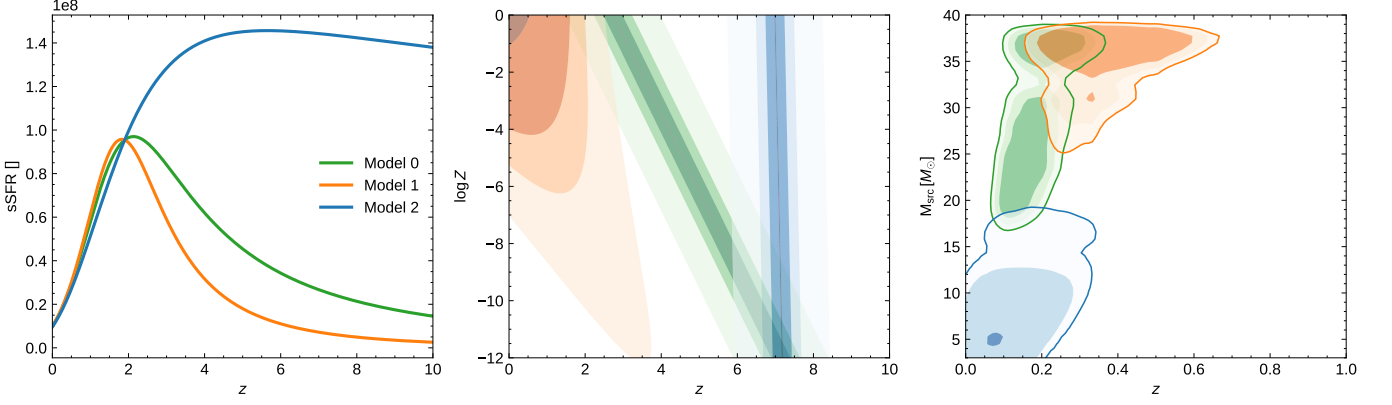
The probability distribution matrix  $p_{\text{grid}}(z, \mathcal{M}|\lambda)$  is constructed by dividing the entries in the matrix of expected detections for a given  $\lambda$ , generated by multiplying the surrogate model detection rate matrix by the run duration, by the sum of the entries in the matrix,  $\mu$ .

In practice, we may not know the values  $z_i, \mathcal{M}_i$  perfectly, but may only have  $K$  samples from a posterior  $p(z_i, \mathcal{M}_i|D_i)$ . Hence the probability  $\mathcal{L}_{\text{event}}(z = z_i, \mathcal{M} = \mathcal{M}_i|\lambda)$  for the  $i^{\text{th}}$  event in Equation 8 can be re-written as

$$\mathcal{L}_{\text{event}}(D_i|\lambda) = \frac{1}{K} \sum_{k=1}^K \frac{p_{\text{grid}}(z = z_i^k, \mathcal{M} = \mathcal{M}_i^k|\lambda)}{\pi(z_i^k, \mathcal{M}_i^k)}, \quad (9)$$

where the subscript  $k$  refers to the  $k^{\text{th}}$  posterior sample among  $K$  available samples, and  $\pi(z, \mathcal{M})$  is the prior used in the original inference for the event (see, e.g. Mandel et al. 2019).

If the number of bins where an event has posterior support is much smaller than the number of samples  $K$ , it may be more efficient to pre-compute the weight that



**Figure 1.** McZ grid given different SFR parameters. Fiducial model is in X,

each observed event  $i$  contributes to each of the bins  $b$ ,  $w_{i,b}$ , which is independent of  $\lambda$ :

$$w_{i,b} = \frac{1}{K} \sum_{k=1}^K \frac{I(\{z_i^k, \mathcal{M}_i^k\} \in b)}{\pi(z_i^k, \mathcal{M}_i^k)}, \quad (10)$$

where  $I(\{z_i^k, \mathcal{M}_i^k\} \in b)$  is an indicator function that evaluates to 1 if sample  $\{z_i^k, \mathcal{M}_i^k\}$  falls into bin  $b$  and 0 otherwise. Then the probability  $\mathcal{L}_{\text{event}}(D_i|\lambda)$  can be evaluated as a sum over bins,

$$\mathcal{L}_{\text{event}}(D_i|\lambda) = \sum_b p_{\text{grid}}(z^b, \mathcal{M}^b|\lambda) w_{i,b}. \quad (11)$$

Unfortunately,  $\mathcal{L}(\mathcal{D}|\lambda)$  is an expensive function, making it hard to sample with. In the next section, we discuss a surrogate model for  $\mathcal{L}(\mathcal{D}|\lambda)$  that is cheaper to sample with.

### 2.3. GP surrogate using Bayesian optimisation

Given  $M$  “observations” of  $\mathcal{L}(\mathcal{D}|\lambda)$ ,  $\{\mathcal{L}(\lambda_0), \dots, \mathcal{L}(\lambda_M)\}$ , we define a conditional probability distribution

$$\begin{aligned} \mathcal{L}^*(\lambda) &= \mathcal{GP}(\mathcal{L}(\lambda)|\mathcal{L}(\lambda_0), \dots, \mathcal{L}(\lambda_M)) \\ &= \mathcal{GP}(\mu(\lambda), \Sigma(\lambda)) \end{aligned} \quad (12)$$

where the conditional probability distribution is given by a Gaussian process  $\mathcal{GP}$ , of mean  $\mu$  and covariance  $\Sigma$ . A Gaussian process, provides a Bayesian posterior probability distribution that describes potential values for  $\mathcal{L}(\mathcal{D}|\lambda)$  at a candidate point  $\lambda$ . Each time we observe a new  $\mathcal{L}(\lambda')$ , this posterior distribution is updated.

The acquisition function measures the value that would be generated by evaluation of the objective function at a new point  $\lambda'$ , based on the current posterior distribution over  $\mathcal{L}(\mathcal{D}|\lambda)$ . We discuss expected improvement, the most commonly used ac

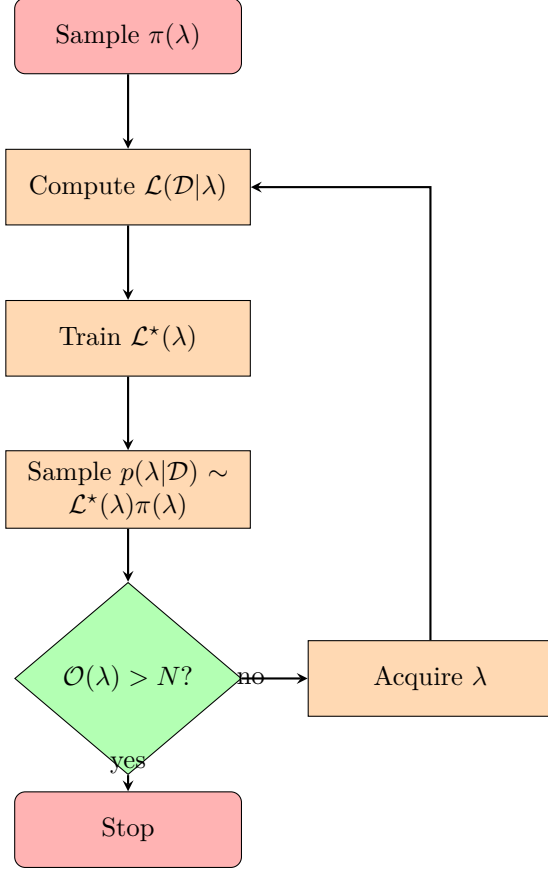
Bayesian optimisation modeling using Gaussian processes is discussed in detail in .

---

#### Algorithm1 Basic pseudo-code for Bayesian optimization

---

- 1: Place Gaussian process  $\mathcal{GP}$  prior on  $\mathcal{L}(\mathcal{D}|\lambda)$
  - 2: Observe  $\mathcal{L}(\mathcal{D}|\lambda)$  at  $N_0$  points randomly sampled.
  - 3: **while**  $n \leq N$  **do**
  - 4:   Update the posterior probability distribution on  $\mathcal{L}(\mathcal{D}|\lambda)$  using all available data
  - 5:   Let  $\lambda_{\text{best}}$  be a maximizer of the acquisition function over  $\lambda$ , where the acquisition function is computed using the current posterior distribution.
  - 6:   Observe  $\mathcal{L}_{\text{best}} = \mathcal{L}(\mathcal{D}|\lambda_{\text{best}})$ .
  - 7:   Increment  $n$
  - 8: **end while**
  - 9: Return a solution:  $\mathcal{L}^*(\lambda) \sim \mathcal{GP}(\lambda)$
-



#### 2.4. Acquisition Functions

An acquisition function is a function that maps the posterior predictive mean and variance to a scalar value. The scalar value is used to select the next point to observe from the model.

In general, we will write the acquisition function as  $\alpha(\mu, \sigma)$ , where  $\mu$  is the posterior predictive mean and  $\sigma$  is the posterior predictive variance.

Each criterion performs a trade-off between exploration i.e. investigating regions where the variance is large and exploitation i.e. investigating regions where the prediction is minimized.

*Uncertainty-based exploration (UE)*—This acquisition function aims at reducing the uncertainty of the GP model. It is defined as

$$\alpha_{\text{UE}}^i(\sigma^i) = \sigma^i$$

and helps to explore the function space in regions of high uncertainty.

One of the most popular BO algorithms is the Efficient Global Optimization (EGO). It uses GP as surrogate model and the Expected Improvement as the infill sampling criterion.

*Expected Improvement*—This also focuses on the maximisation/minimisation of the GP’s posterior predictive mean. It looks for the best value of  $\mu$  so far and then looks for the next point that has a higher probability of being better than the best value so far.

$$\alpha_{\text{EI}}^i(\mu^i, \sigma^i) = \sigma^i (u^i \mathcal{N}_{\text{CDF}}(u^i) \mathcal{N}_{\text{PDF}}(u^i)),$$

where  $u^i = \frac{\mu^i - \mu_{\text{best}}}{\sigma^i}$  and  $\mu_{\text{best}}$  is the best value of  $\mu$  so far.

where  $\mathcal{N}_{\text{PDF}}$  denotes the Probability Density function (PDF) of the univariate Gaussian probability distribution.

While we would like to choose  $x$  so that this improvement is large,  $f(x)$  is unknown until after the evaluation. What we can do, however, is to take the expected value of this improvement and choose  $x$  to maximize it.

The Expected Improvement (EI) takes into account the improvement induced by a candidate that is defined as:  $I(\mathbf{x}) = \max\{0, y_{\min} - f(\mathbf{x})\}$ .

#### 2.5. Diagnostics

We run an MCMC sampler using

$$\mathcal{L}^*(\lambda) = \mu_{\mathcal{GP}}(\lambda) \quad (13)$$

$$\mathcal{L}_V^*(\lambda) = \mathcal{N}(\mu_{\mathcal{GP}}(\lambda), \sigma_{\mathcal{GP}}(\lambda)) \quad (14)$$

The KL divergence using the two  $\mathcal{L}^*(\lambda)$  should converge once the model is trained. Additionally, the KL divergence using a reference posterior trained with a lot of points should converge. This will indicate that the surrogate likelihood distribution has been adequately modelled in regions of parameter space.

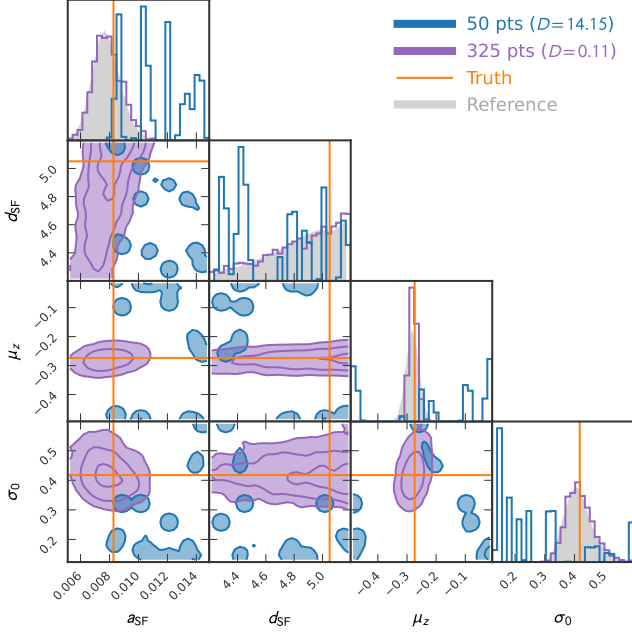
### 3. INFERENCE WITH SURROGATE MODEL

#### 3.1. Simulation study

To validate our method, we conduct the study described in Section 2.5 of (Riley & Mandel 2023).

We simulate the evolution of 32 million binary systems using COMPAS (settings specified in Appendix A), resulting in a population of approximately 68 thousand DCOs that merge within the age of the Universe,  $\sim 13.8\text{Gyr}$ . We sample 100  $\lambda$  from a  $\pi(\lambda)$ , and generate  $\vec{\mathcal{D}}$ , unique mock COMPAS populations for each  $\lambda$ . We build surrogate likelihoods for each dataset, and use them to approximate  $p(\lambda|\mathcal{D})$ . Figure 2 displays the  $p(\lambda|\mathcal{D})$  for one mock population. The gray “reference” posterior is obtained using a surrogate likelihood  $\mathcal{L}^*$  trained with 950 points. The blue posterior is obtained using a surrogate likelihood  $\mathcal{L}_V^*$  trained with 50 points, while the purple posterior’s surrogate is built with 325

points. The blue posterior has a  $D_{KL} > 10$  when compared with the gray reference posterior. The purple posterior has a  $D_{KL} < 1$  when compared with the gray reference posterior, indicating that the two posteriors are similar. The orange line marks the true  $\lambda$  used to generate the dataset.



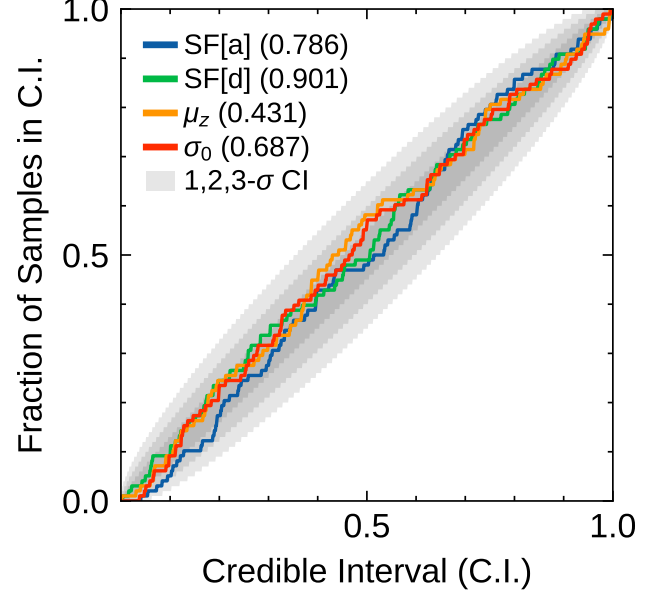
**Figure 2.** Simulation posteriors

The PP-plot in Figure 3 shows that the true value lies within the correct region of the credible intervals, indicating that the results are not biased.

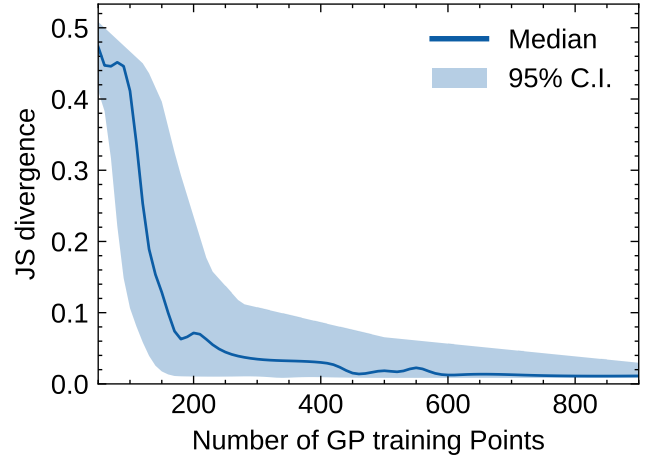
Finally, Figure 4 shows the JS divergence

### 3.2. LVK analysis

*Data*—The catalogue of LIGO, Virgo, and KAGRA (LVK) gravitational wave observations (LIGO Scientific and Virgo Collaborations (2022), LIGO Scientific, Virgo and KAGRA Collaborations (2021)) provides us with a number of merger detections from a population of double compact objects (DCOs). For each LVK event we know, among other things, the chirp mass of the merging binary and the merger redshift, up to a measurement uncertainty. We get the data from the 4-OGC. Plots of the event  $\mathcal{M}, z$  median estimates and  $X - \sigma$  uncertainty can be found in Figure 5. Filtering the events based on the OGC definition of  $p(bbh) \geq 0.95$  and ensuring that the median value of the  $p(\mathcal{M}, z|d_i)$  lies within  $\mathcal{M} \in [4, 45]$ ,  $z \in [0, 1]$ , we get XX events. Less confident events could be included (e.g. Farr et al. 2015), they will typically contribute little information due to



**Figure 3.** PP-plots from simulation study.



**Figure 4.** KL-Distances of the posteriors obtained with XX, compared against to those from YY.

greater measurement uncertainties. These events that we use in our analysis are colored green in Fig 5.

The priors for the events are in  $\pi(\mathcal{M}, V_c)$ , however, we need  $\pi(\mathcal{M}, z)$ . To transform the priors, we use the following equation

$$\pi(\mathcal{M}, z) = \pi(\mathcal{M}, V_c(z)) \frac{dV_c(z)}{dz}. \quad (15)$$

We use the definition of  $dV_c(z)/dz$  from (Hogg 1999), given by

$$dV_C = D_H \frac{(1+z)^2 D_A^2}{E(z)} d\Omega dz \quad (16)$$

Using Equation 10, we obtain the weights for each posterior. A plot of each event’s weights summed together is shown in Figure 6.

*Posteriors*—pass

#### 4. DISCUSSION

Future work

- Do hyper-parameter optimisation for the GP (kernel choices, kernel settings)
- Do hyper-parameter optimisation for the acquisition function
- Explore range of acquisition functions

#### DATA AVAILABILITY

The data underlying this article is available in the Zenodo deposit for this work. [LINK TO ZENODO](#).

#### ACKNOWLEDGEMENTS

We gratefully acknowledge the Swinburne Supercomputing OzSTAR Facility for computational resources. All analyses (including test and failed analyses) performed for this study used  $XXK$  core hours on OzSTAR.

This would have amounted to a carbon footprint of  $\sim XXt$  CO<sub>2</sub> (??). Thankfully, as OzSTAR is powered by wind energy from Iberdrola Australia; the electricity for computations produces negligible carbon waste. A.V. is supported by the Australian Research Council (ARC) Centre of Excellence CE170100004.

4-OGC. IM is a recipient of the Australian Research Council Future Fellowship FT190100574. AV is XX. Discussions with Renate, Matt, Kate.

*Facilities:* LIGO-VIRGO-KAGRA

*Software:* Python (Oliphant 2007; Millman & Aivazis 2011), astropy (Price-Whelan et al. 2018), COMPAS (Team COMPAS: Riley, J. et al. 2022a), PyCBC (), emcee (), trieste (), TensorFlow (), numpy (Harris et al. 2020), scipy (Virtanen et al. 2020), pandas (Reback et al. 2020), matplotlib (Hunter 2007), corner (Foreman-Mackey 2016), sphinx (Brandl 2021), Jupyter-book (Kluyver et al. 2016; Executable Books Community 2020).

#### REFERENCES

- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481, doi: [10.1146/annurev.astro.46.060407.145222](#)
- Barrett, J. W., Gaebel, S. M., Neijssel, C. J., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 477, 4685, doi: [10.1093/mnras/sty908](#)
- Belczynski, K., Bulik, T., Fryer, C. L., et al. 2010, *ApJ*, 714, 1217, doi: [10.1088/0004-637X/714/2/1217](#)
- Bhattacharya, D., & van den Heuvel, E. P. J. 1991, *PhR*, 203, 1, doi: [10.1016/0370-1573\(91\)90064-S](#)
- Brandl, G. 2021, URL <http://sphinx-doc.org/sphinx.pdf>
- Broekgaarden, F. S., Justham, S., de Mink, S. E., et al. 2019, *MNRAS*, 490, 5228, doi: [10.1093/mnras/stz2558](#)
- de Kool, M. 1990, *ApJ*, 358, 189, doi: [10.1086/168974](#)
- Dominik, M., Belczynski, K., Fryer, C., et al. 2012, *ApJ*, 759, 52, doi: [10.1088/0004-637X/759/1/52](#)
- Executable Books Community. 2020, *Jupyter Book*, v0.10, Zenodo, doi: [10.5281/zenodo.4539666](#)
- Farr, W. M., Gair, J. R., Mandel, I., & Cutler, C. 2015, *PhRvD*, 91, 023005, doi: [10.1103/PhysRevD.91.023005](#)
- Foreman-Mackey, D. 2016, *The Journal of Open Source Software*, 1, 24, doi: [10.21105/joss.00024](#)
- Fryer, C. L., Belczynski, K., Wiktorowicz, G., et al. 2012, *ApJ*, 749, 91, doi: [10.1088/0004-637X/749/1/91](#)
- Hamann, W. R., & Koesterke, L. 1998, *A&A*, 335, 1003
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357, doi: [10.1038/s41586-020-2649-2](#)
- Hobbs, G., Lorimer, D. R., Lyne, A. G., & Kramer, M. 2005, *MNRAS*, 360, 974, doi: [10.1111/j.1365-2966.2005.09087.x](#)
- Hogg, D. W. 1999, *arXiv e-prints, astro*, doi: [10.48550/arXiv.astro-ph/9905116](#)
- Hunter, J. D. 2007, *Computing in science & engineering*, 9, 90
- Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, *MNRAS*, 329, 897, doi: [10.1046/j.1365-8711.2002.05038.x](#)
- Kluyver, T., Ragan-Kelley, B., Pérez, F., et al. 2016, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, ed. F. Loizides & B. Schmidt, IOS Press, 87 – 90
- Kroupa, P. 2001, *MNRAS*, 322, 231, doi: [10.1046/j.1365-8711.2001.04022.x](#)
- Langer, N., & Norman, C. A. 2006, *The Astrophysical Journal*, 638, L63, doi: [10.1086/500363](#)



- LIGO Scientific and Virgo Collaborations. 2022, GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run - Parameter Estimation Data Release, v2, Zenodo, doi: [10.5281/zenodo.6513631](https://doi.org/10.5281/zenodo.6513631)
- LIGO Scientific, Virgo and KAGRA Collaborations. 2021, GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run – Parameter estimation data release, Zenodo, doi: [10.5281/zenodo.5546663](https://doi.org/10.5281/zenodo.5546663)
- Mandel, I., Farr, W. M., & Gair, J. R. 2019, MNRAS, 486, 1086, doi: [10.1093/mnras/stz896](https://doi.org/10.1093/mnras/stz896)
- Marchant, P., Renzo, M., Farmer, R., et al. 2019, ApJ, 882, 36, doi: [10.3847/1538-4357/ab3426](https://doi.org/10.3847/1538-4357/ab3426)
- Massevitch, A., & Yungelson, L. 1975, Mem. Soc. Astron. Italiana, 46, 217
- Millman, K. J., & Aivazis, M. 2011, Computing in Science Engineering, 13, 9, doi: [10.1109/MCSE.2011.36](https://doi.org/10.1109/MCSE.2011.36)
- Neijssel, C. J., Vigna-Gómez, A., Stevenson, S., et al. 2019, MNRAS, 490, 3740, doi: [10.1093/mnras/stz2840](https://doi.org/10.1093/mnras/stz2840)
- Oliphant, T. E. 2007, Computing in Science Engineering, 9, 10, doi: [10.1109/MCSE.2007.58](https://doi.org/10.1109/MCSE.2007.58)
- Pfahl, E., Rappaport, S., & Podsiadlowski, P. 2002, ApJ, 573, 283, doi: [10.1086/340494](https://doi.org/10.1086/340494)
- Podsiadlowski, P., Langer, N., Poelarends, A. J. T., et al. 2004, ApJ, 612, 1044, doi: [10.1086/421713](https://doi.org/10.1086/421713)
- Price-Whelan, A. M., Sipőcz, B., Günther, H., et al. 2018, The Astronomical Journal, 156, 123
- Reback, J., McKinney, W., jbrockmendel, et al. 2020, pandas-dev/pandas: Pandas 1.0.2, v1.0.2, Zenodo, doi: [10.5281/zenodo.3708035](https://doi.org/10.5281/zenodo.3708035)
- Riley, J., & Mandel, I. 2023, ApJ, 950, 80, doi: [10.3847/1538-4357/accf90](https://doi.org/10.3847/1538-4357/accf90)
- Soberman, G. E., Phinney, E. S., & van den Heuvel, E. P. J. 1997, A&A, 327, 620
- Stevenson, S., Sampson, M., Powell, J., et al. 2019, ApJ, 882, 121, doi: [10.3847/1538-4357/ab3981](https://doi.org/10.3847/1538-4357/ab3981)
- Stevenson, S., Vigna-Gómez, A., Mandel, I., et al. 2017, Nature Communications, 8, 14906, doi: [10.1038/ncomms14906](https://doi.org/10.1038/ncomms14906)
- Tauris, T. M., Langer, N., & Podsiadlowski, P. 2015, MNRAS, 451, 2123, doi: [10.1093/mnras/stv990](https://doi.org/10.1093/mnras/stv990)
- Tauris, T. M., & van den Heuvel, E. P. J. 2006, Formation and evolution of compact stellar X-ray sources, Vol. 39, 623–665
- Tauris, T. M., Kramer, M., Freire, P. C. C., et al. 2017, ApJ, 846, 170, doi: [10.3847/1538-4357/aa7e89](https://doi.org/10.3847/1538-4357/aa7e89)
- Team COMPAS: Riley, J., Agrawal, P., Barrett, J. W., et al. 2022a, Journal of Open Source Software, 7, 3838, doi: [10.21105/joss.03838](https://doi.org/10.21105/joss.03838)
- . 2022b, ApJS, 258, 34, doi: [10.3847/1538-4365/ac416c](https://doi.org/10.3847/1538-4365/ac416c)
- Timmes, F. X., Woosley, S. E., & Weaver, T. A. 1996, ApJ, 457, 834, doi: [10.1086/176778](https://doi.org/10.1086/176778)
- van Son, L. A. C., de Mink, S. E., Chruslinska, M., et al. 2022, arXiv e-prints, arXiv:2209.03385. <https://arxiv.org/abs/2209.03385>
- Vigna-Gómez, A., Neijssel, C. J., Stevenson, S., et al. 2018, MNRAS, 481, 4009, doi: [10.1093/mnras/sty2463](https://doi.org/10.1093/mnras/sty2463)
- Vinciguerra, S., Neijssel, C. J., Vigna-Gómez, A., et al. 2020, MNRAS, 498, 4705, doi: [10.1093/mnras/staa2177](https://doi.org/10.1093/mnras/staa2177)
- Vink, J. S., & de Koter, A. 2005, A&A, 442, 587, doi: [10.1051/0004-6361:20052862](https://doi.org/10.1051/0004-6361:20052862)
- Vink, J. S., de Koter, A., & Lamers, H. J. G. L. M. 2000, A&A, 362, 295
- . 2001, A&A, 369, 574, doi: [10.1051/0004-6361:20010127](https://doi.org/10.1051/0004-6361:20010127)
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
- Webbink, R. F. 1984, ApJ, 277, 355, doi: [10.1086/161701](https://doi.org/10.1086/161701)
- Xu, X.-J., & Li, X.-D. 2010a, ApJ, 716, 114, doi: [10.1088/0004-637X/716/1/114](https://doi.org/10.1088/0004-637X/716/1/114)
- . 2010b, ApJ, 722, 1985, doi: [10.1088/0004-637X/722/2/1985](https://doi.org/10.1088/0004-637X/722/2/1985)

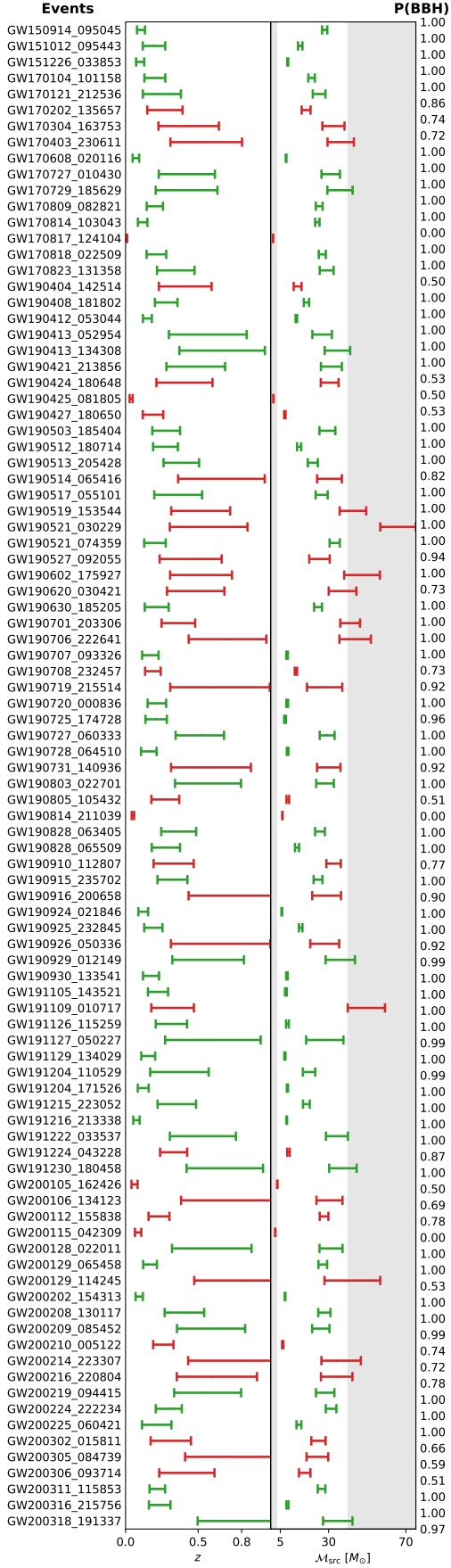
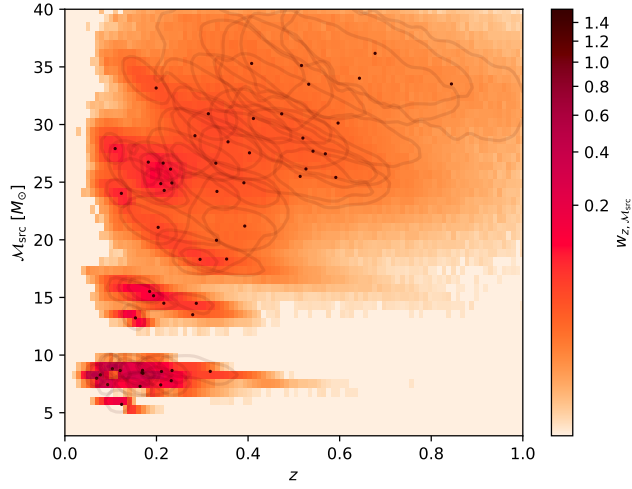



Figure 5. The events and uncertainty as predicted by Ogc4







**Figure 6.** Weights generated using the OGC4 posteriors 

## APPENDIX

### A. COMPAS CONFIGURATION FIDUCIAL VALUES

### B. SAMPLER SETTINGS

Here we describe the sampler settings we use. We also plot some trace plots.

**Table 1.** Initial values and default settings for the COMPAS fiducial model.

Description and name	Value/range	Note / setting
Initial conditions		
Initial primary mass $m_{1,i}$	$[5, 150] M_{\odot}$	Kroupa (2001) IMF $\propto m_{1,i}^{-\alpha}$ with $\alpha_{\text{IMF}} = 2.3$ for stars above $5 M_{\odot}$
Initial mass ratio $q_i = m_{2,i}/m_{1,i}$	$[0, 1]$	We assume a flat mass ratio distribution $p(q_i) \propto 1$ with $m_{2,i} \geq 0.1 M_{\odot}$
Initial semi-major axis $a_i$	$[0.01, 1000] \text{ AU}$	Distributed flat-in-log $p(a_i) \propto 1/a_i$
Initial metallicity $Z_i$	$[0.0001, 0.03]$	Distributed uniform-in-log
Initial orbital eccentricity $e_i$	0	All binaries are assumed to be circular at birth
Fiducial parameter settings:		
Stellar winds for hydrogen rich stars	Belczynski et al. (2010)	Based on Vink et al. (2000, 2001), including LBV wind mass loss with $f_{\text{LBV}} = 1.5$
Stellar winds for hydrogen-poor helium stars	Belczynski et al. (2010)	Based on Hamann & Koesterke (1998) and Vink & de Koter 2005
Max transfer stability criteria	$\zeta$ -prescription	Based on Vigna-Gómez et al. (2018) and references therein
Mass transfer accretion rate	thermal timescale	Limited by thermal timescale for stars Hurley et al. (2002); Vinciguerra et al. (2020)
	Eddington-limited	Accretion rate is Eddington-limit for compact objects
Non-conservative mass loss	isotropic re-emission	Mashevitch & Yungelson (1975); Bhattacharya & van den Heuvel (1991); Soberman et al. (1997) Tauris & van den Heuvel (2006)
Case BB mass transfer stability	always stable	Based on Tauris et al. (2015, 2017); Vigna-Gómez et al. (2018)
CE prescription	$\alpha - \lambda$	Based on Webbink (1984); de Kool (1990)
CE efficiency $\alpha$ -parameter	1.0	
CE $\lambda$ -parameter	$\lambda_{\text{Nanjing}}$	Based on Xu & Li (2010a,b) and Dominik et al. (2012)
Hertzsprung gap (HG) donor in CE	pessimistic	Defined in Dominik et al. (2012); HG donors don't survive a CE phase
SN natal kick magnitude $v_k$	$[0, \infty) \text{ km s}^{-1}$	Drawn from Maxwellian distribution with standard deviation $\sigma_{\text{rms}}^{\text{1D}}$
SN natal kick polar angle $\theta_k$	$[0, \pi]$	$p(\theta_k) = \sin(\theta_k)/2$
SN natal kick azimuthal angle $\phi_k$	$[0, 2\pi]$	Uniform $p(\phi) = 1/(2\pi)$
SN mean anomaly of the orbit	$[0, 2\pi]$	Uniformly distributed
Core-collapse SN remnant mass prescription	delayed	From Fryer et al. (2012), which has no lower BH mass gap
USSN remnant mass prescription	delayed	From Fryer et al. (2012)
ECSN remnant mass prescription	$m_f = 1.26 M_{\odot}$	Based on Equation 8 in Timmes et al. (1996)
Core-collapse SN velocity dispersion $\sigma_{\text{rms}}^{\text{1D}}$	$265 \text{ km s}^{-1}$	1D rms value based on Hobbs et al. (2005)
USSN and ECSN velocity dispersion $\sigma_{\text{rms}}^{\text{1D}}$	$30 \text{ km s}^{-1}$	1D rms value based on e.g., Pfahl et al. (2002); Podsiadlowski et al. (2004)
PISN / PPISN remnant mass prescription	Marchant et al. (2019)	As implemented in Stevenson et al. (2019)
Maximum NS mass	$\text{maxNS} = 2.5 M_{\odot}$	Following Fryer et al. (2012)
Tides and rotation		We do not include tides and/or rotation in this study
Binary fraction	$f_{\text{bin}} = 0.7$	
Solar metallicity $Z_{\odot}$	$Z_{\odot} = 0.0142$	based on Asplund et al. (2009)
Simulation settings		
Binary population synthesis code	COMPAS	Stevenson et al. (2017); Barrett et al. (2018); Vigna-Gómez et al. (2018); Neijssel et al. (2019) Broekgaarden et al. (2019); Team COMPAS: Riley, J. et al. (2022b).