Simulation Based Inference of BNS Kilonova Properties: A Case Study with AT2017gfo

Phelipe A. Darc*
Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brazil.

Clecio R. Bom
Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brazil.
Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET-RJ),
Av. Maracanã, 229,Maracanã, Rio de Janeiro, Brazil.

Bernardo M. O. Fraga
Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brazil.

Charlie D. Kilpatrick
Center for Interdisciplinary Exploration and Research in Astro- physics (CIERA)
and Department of Physics and Astronomy
North-western University, Evanston, IL 60208, USA

Abstract

Kilonovae are a class of astronomical transients observed as counterparts to mergers of compact binary systems, such as a binary neutron star (BNS) or black hole-neutron star (BHNS) inspirals. They serve as probes for heavy-element nucleosynthesis in astrophysical environments, while together with gravitational wave emission constraining the distance to the merger itself, they can place constraints on the Hubble constant. Obtaining the physical parameters (e.g. ejecta mass, velocity, composition) of a kilonova from observations is a complex inverse problem, usually tackled by sampling-based inference methods such as Markov-chain Monte Carlo (MCMC) or nested sampling techniques. These methods often rely on computing approximate likelihoods, since a full simulation of compact object mergers involve expensive computations such as integrals, the calculation of likelihood of the observed data given parameters can become intractable, rendering the likelihood-based inference approaches inapplicable. We propose here to use Simulation-based Inference (SBI) techniques to infer the physical parameters of BNS kilonovae from their spectra, using simulations produced with *KilonovaNet*. Our model uses Amortized Neural Posterior Estimation (ANPE) together with an embedding neural network to accurately predict posterior distributions from simulated spectra. We further test our model with real observations from AT 2017gfo, the only kilonova with multi-messenger data, and show that our estimates agree with previous likelihood-based approaches.

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

1 Introduction

The coalescence of binary neutron stars (BNS) or black hole-neutron star (BHNS) systems are prime sources of gravitational waves (GW) observed by Advanced LIGO/Virgo (Collaboration et al. 2015; Acernese et al. 2014), and also power bright electromagnetic (EM) transients. The so called multi-messenger observations offer insights into the properties of matter under extreme conditions and can be used to infer parameters of the BNS merger and constrain the neutron star equation of state (Radice et al. 2018; Margalit and Metzger 2017), as well as the Hubble constant both from standard sirens or dark sirens perspectives (Coughlin et al. 2020; Bom and Antonella Palmese 2023; Abbott et al. 2017; Soares-Santos, A. Palmese, et al. 2019). In particular, the study of the spectra of these EM counterparts can help us to understand the process behind the mass ejection mechanism (Metzger 2019), the physical conditions during the merger and its aftermath. Kilonovae are powered by the radioactive decay of rare heavy elements produced in the merger, and are primarily observed in the ultraviolet, optical and infrared, reaching peak brightness two to three days after the merger (Li and Paczyń ski 1998).

Kilonovae are rare events, with only a few having been observed so far, primarily as counterparts to short gamma-ray bursts at redshifts $z > 0.1$ where they are difficult to observe at optical wavelengths (Rastinejad et al. 2022). Therefore, investigating the parameter space of kilonova and their effects on their observable properties is done with simulations, which provide detailed models of ejected matter during mergers (Dietrich and Ujevic 2017; Lukoš iute et al. 2022; Kasen et al. 2017). However, these simulations involve detailed particle physics, general relativity, hydrodynamics, and radiative transfer and are therefore computationally complex, taking several hours to produce observables for one parameter set (Bulla 2019; Lukoš iute et al. 2022).

Obtaining information about a kilonova's physical parameters from observational data is a complex inverse problem, which are usually approached using sampling-based inference methods such as Markov-chain Monte Carlo (MCMC) and nested sampling techniques (Cranmer, Brehmer, and Louppe 2020; Montel, Alvey, and Weniger 2023). However, the models that generated the data are complex, involving several variables, often rely on approximate likelihoods, and the time needed to reach convergence scales poorly with the dimensionality of the explored parameter space. Likelihood-free Inference (LFI), or Simulation-based Inference (SBI; Cranmer, Brehmer, and Louppe 2020) are a set of algorithms that bypass the need for an explicit likelihood by training an estimator on simulated data, "learning" to approximate the likelihood or directly estimating the posterior distribution. These models can leverage the use of powerful density estimators such as normalizing flows.

In this work we propose to use the Neural Posterior Estimation (SNPE-C/NPE; Greenberg, Nonnenmacher, and Macke 2019), a method of posterior estimation, to infer the physical parameters of simulated BNS kilonovae.

2 The Astrophysical Parameters of Kilonovae

Detailed radiative transfer simulations of kilonova spectra play an essential role in multi-messenger astrophysics, providing basic information about the physical conditions, elemental abundances, and velocities in kilonova ejecta. Choosing a suitable simulator for parameter inference studies is therefore crucial. Kilonova spectra and light curves depend strongly on the nuclear yields, neutrino flux, geometric orientations, mass, and velocity of the ejecta (Metzger 2019; Kawaguchi, Shibata, and Tanaka 2020).

In this work, we made use of *KilonovaNet*, a conditional variational autoencoder (cVAE) for generating surrogate kilonova spectra on tens-of-millisecond timescales (Lukoš iute et al. 2022). Given a set of parameters, it's also capable of generating light curves. It was developed to greatly reduce the time required during parameter inference.

KilonovaNet was trained on three different datasets of simulated kilonova spectra from BNS or BHNS mergers (Dietrich, Coughlin, et al. 2020; Kasen et al. 2017; Anand et al. 2020). In this work, we focused on the simulations by Dietrich, Coughlin, et al. 2020, since they come directly from BNS merger simulations and are more realistic than the simpler parameterization of ejecta mass, velocity, and abundance in Kasen et al. 2017. The parameter sets consist of the mass of the dynamical ejecta ($M_{ej,dyn}$), the mass of the post-merger ejecta ($M_{ej,pm}$), the half-opening angle of the lanthanide-rich tidal dynamical ejecta ϕ, and the cosine of the observer viewing angle $\cos(\theta_{obs})$.

3 Simulations and Training

KilonovaNet takes a set of parameters [$M_{ej,dyn}$, $M_{ej,pm}$, ϕ, $\cos(\theta_{obs})$] and a list of days after the merger to produce a spectrum. Initial tests showed that the parameter estimates using single-day spectra were either too broad or wrong. We therefore combined spectra from three different times after merger (1.5, 2.5 and 3.5 days) to serve as input. It is important to mention that an ensemble of models was trained to encompass various time intervals, specifically $[1.0+x, 2.0+x, 3.0+x]$, where x ranged from 0.00 to 0.99, in increments of 0.2. This approach was adopted to address potential variations in observation times. Our analysis will primarily center on the model trained for three distinct post-merger time points: 1.5, 2.5, and 3.5 day, this choice will become evident in the next section.

We generate $100,000$ triplets ($300,000$ spectra in total) and cut them to wavelengths between 5000 and 8000 Å. Each spectrum was interpolated to have 550 points, and the flux was normalized to zero mean and unit variance. We use only the flux value as input, since the wavelengths for all spectra are the same after interpolating. A Gaussian smoothing was applied to reduce the noise from the simulator, and Gaussian noise was added corresponding to 5% of the flux, to make the simulated data more similar to real observations.

We choose uniform priors for all the parameters in the simulations, with the minimum and maximum corresponding to the smaller and larger value coming from the original simulated data (Dietrich and Ujevic 2017). We use the PyTorch-based SBI library (Tejero-Cantero et al. 2020)[1], and their implementation of Amortized Neural posterior estimation (ANPE), with a mixture density network (MDN) as the density estimator.

The basic idea behind an Amortized Neural posterior estimation is to first train a model (training phase) – specifically, a *density estimator* – that is not focused on any particular observation. Instead, it learns to be a versatile estimator that attempts to approximate all posteriors supported by the prior. Once trained, the density estimator can, in the second phase (inference), quickly and continually infer parameters of BNS kilonovae from their spectra. To train the Network, we can simulate using prior-draw parameters to build a dataset, of kilonova parameters (θ_j) and their respective spectra (X_j), and minimize the Loss function over the weights. Once the density estimator has been trained on simulated data x, it can then be applied to empirical data X_0 (ATF2017gfo) to compute the posterior.

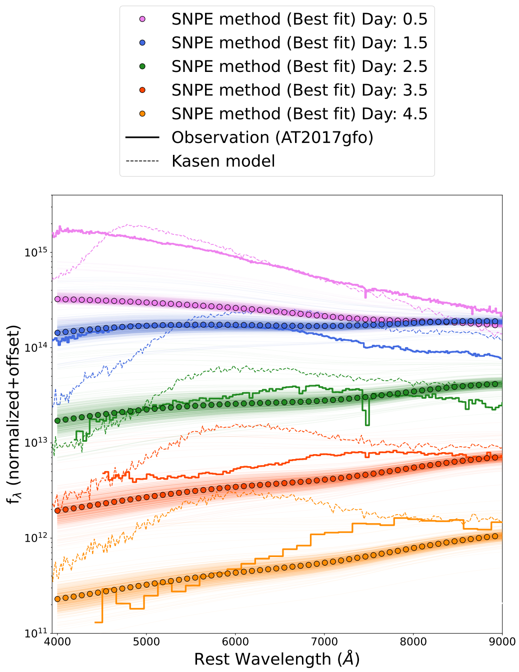
We implemented an embedding network in our model, composed of three convolutional layers followed by a max pooling layer, two Long Short-Term memory (LSTM) layers and a dense layer with 100 neurons. The use of Convolutional layers and LSTM layers are essential to extract specific patterns or trends in the spectral data, which can be indicative of certain physical properties, such as the lanthanide composition of the ejecta and the viewing angle (Metzger 2019). We use the ADAM optimizer with a learning rate of 0.0001 to perform the parameter updates during training.

4 Results

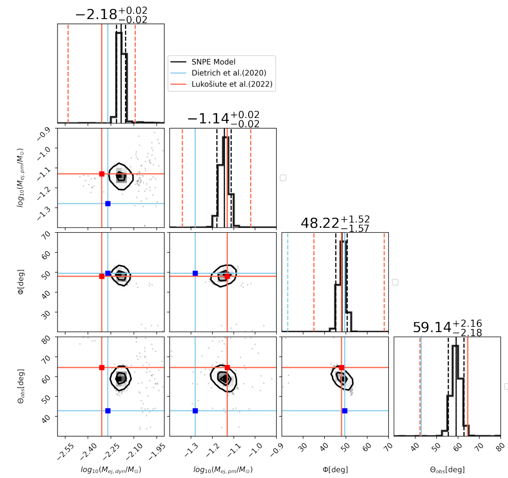
We validated the model's training by applying it to a new set of synthetic data. We obtain coefficient of correlation for the parameters $M_{ej,dyn}$, $M_{ej,pm}$, ϕ, $cos(\theta_{obs})$ of 0.968, 0.974, 0.991, and 0.815, respectively. Additionally, we made use of a well known metrics in SBI to access the quality of our posterior: Simulation-based calibration (SBC), it provides a qualitative view and a quantitative measure to check, whether the uncertainties of the posterior are balanced, i.e., neither over-confident nor under-confident. And it can be summarized by the Classifier 2-Sample Tests (C2ST; Lueckmann et al. 2021) score, we obtain the C2ST score for the parameters $M_{ej,dyn}$, $M_{ej,pm}$, ϕ, $cos(\theta_{obs})$ of 0.4946, 0.5034, 0.5072 and 0.5018, respectively. These scores demonstrate that our model is reliable and capable of inferring the physical parameters of kilonova on simulated data.

Ours main goals are to demonstrate that SBI is suited for kilonova parameter retrieval and to apply this model to real spectra of kilonova; so far, only one was identified as an optical counterpart to a GW event: GW170817 and AT 2017gfo (Soares-Santos, Holz, et al. 2017). We use spectral data collated in Shappee et al. 2017, taken at 1.477, 2.476 and 3.479 days after the merger, passing a uniform filter of size 10 spaxels (corresponding to a step of 54.45 Å) in order to be visually similar to the simulated data. Fig. 1b displays the parameter posteriors and the best-fit values obtained by our

[1] https://www.mackelab.org/sbi/



(a) Spectroscopic time series of AT2017gfo, the spectra generated by *KilonovaNet* using the Best fit of our NPE model and the Kasen model best fit made by Shappee et al. 2017, the vertical axis is observed flux (f_λ).



(b) Inferred posteriors for the model parameters at 10%, 32%, 68% and 95% confidence intervals. The median and 90% interval are shown in vertical solid and dashed lines, respectively, and reported above each column. Results from Dietrich, Coughlin, et al. 2020 (blue) and Lukoš iute et al. 2022 (orange) are also shown for comparison.

Figure 1: Results of the performance of our ANPE model on the AT 2017gfo (counterpart to the event GW170817) with the data collated in Shappee et al. 2017.

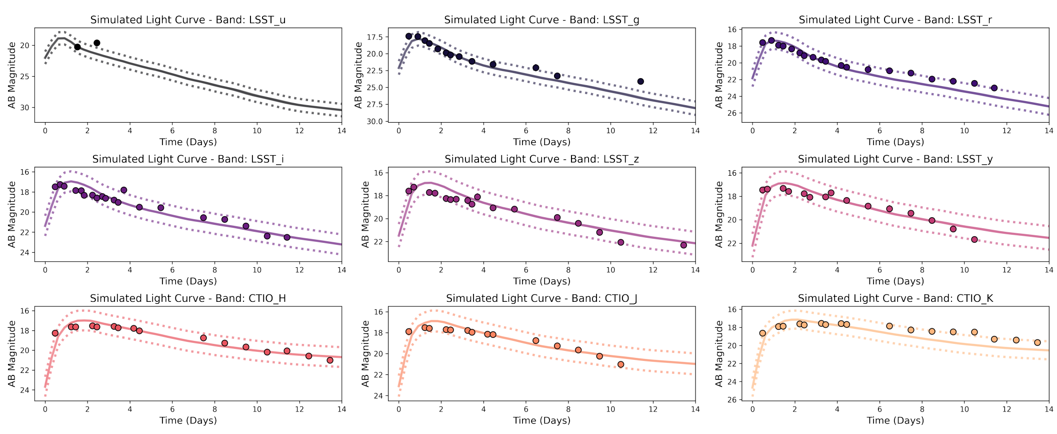


Figure 2: Light curves for AT 2017gfo. Observed values (points) and the prediction based inferred parameters using the NPE model (solid lines). The shaded bands represent the 90% confidence interval of light curves constructed from the posterior samples. The dashed lines represent the 1 mag tolerance typically used to represent modelling error of kilonova light curves (Lukoš iute et al. 2022)

model, compared to the values obtained by Dietrich, Coughlin, et al. 2020 (blue) and Lukoš iute et al. 2022 (orange). The marginal distributions show the most probable value as a solid line, and the 90% interval as dashed lines. Some of the intervals are beyond the plot range, and are omitted for the sake of clarity. While the results are consistent, with our median values lying inside their 90% confidence interval, the data and method of obtaining the parameters are different. Dietrich, Coughlin, et al. 2020 uses light curve data, and incorporates priors coming from GW and pulsar observations; Lukoš iute et al. 2022 used the same dataset, and a dynamic nested-sampling algorithm to obtain the best-fit parameters in a likelihood-based approach. Here, we combined spectral data for three different times after the merger, and used a trained ANPE model on them to obtain the parameter posteriors.

Fig. 1a displays a comparison of spectral data reconstructed using *KilonovaNet* surrogate model with our best parameters, the real spectra, and the fitted spectra from Shappee et al. 2017, for five different times after the merger. Our model reconstructs the spectra from AT 2017gfo for all times after 0.5 days post-merger following the with reasonable accuracy. The worst results are for 0.5 and 4.5 days after merger, times that the model was not trained on. At 0.5 days, however, the error reported for the spectra generated by *Kilonovanet* is large, and could be at least partly responsible for the difference between the predicted and the observed spectra.

Fig. 2 shows the light curves derived from the fits, for different bands. Solid lines represent the inferred light curve using the best-fit parameters, and the shaded region the reconstructed light curves using the 90% confidence interval from the inferred parameter distributions. Our model is able to accurately predict light curves for all filters up to 14 days after the merger, considering a tolerance of one magnitude.

5 Conclusion

Astrophysicists have a suite of simulators at their disposal that can model observables from kilonovae. Driven by the upcoming Large Synoptic Survey of Space and Time (LSST) (Ivezić et al. 2019), the number of events will increase, and the need for a fast and reliable inference method based on simulation will grow, as the complexity of these simulators leads to an increasing number of parameters to be inferred and high dimensionality of parameter spaces. In this work, we tested a simulation-based (or likelihood-free) inference method - an amortized neural posterior estimator - that allows us to infer the parameters' posterior within seconds and bypasses the conventional likelihood.

SBI is particularly useful when the likelihood function is intractable or computationally expensive to evaluate. Given the speed with which the ANPE model performs during parameter inference, it will serve as a useful tool in future gravitational wave observing runs to quickly analyze potential kilonova candidates. The speed-up provided in the parameter inference and spectra retrieval also enables the exploration of several different simulation models over limited observations in a reasonable time.

Another important driver of these algorithms is the rapidly increasing capabilities of machine learning, which enable us to analyze high-dimensional data efficiently and automatically extract features from the data, allowing for faster inference. We demonstrate the capacity of our model to reproduce observables, spectra, and light curves of the synthetic data and the AT 2017gfo event, exhibiting the capability and reliability of using an SBI approach to constrain the parameters of kilonova models.

In summary, our exploration into the merger of machine learning and physics, specifically applied to kilonova events, has yielded promising results. The accuracy of our Posterior estimator model, when applied to synthetic and real data, represents a significant stride forward in the study of these explosive phenomena.

References

Abbott, B. P. et al. (Nov. 2017). "A gravitational-wave standard siren measurement of the Hubble constant". In: 551, pp. 85–88. DOI: 10.1038/nature24471. arXiv: 1710.05835.

Acernese, F et al. (Dec. 2014). "Advanced Virgo: a second-generation interferometric gravitational wave detector". In: *Classical and Quantum Gravity* 32.2, p. 024001. DOI: 10.1088/0264-9381/32/2/024001. URL: https://dx.doi.org/10.1088/0264-9381/32/2/024001.

Anand, Shreya et al. (Sept. 2020). "Optical follow-up of the neutron star–black hole mergers S200105ae and S200115j". In: *Nature Astronomy* 5.1, pp. 46–53. DOI: 10.1038/s41550-020-1183-3. URL: https://doi.org/10.1038%2Fs41550-020-1183-3.

Bom, Clecio R and Antonella Palmese (2023). "Standard Siren Cosmology with Gravitational Waves from Binary Black Hole Mergers in Active Galaxy Nuclei". In: *arXiv preprint arXiv:2307.01330*.

Bulla, M (Sept. 2019). "possis: predicting spectra, light curves, and polarization for multidimensional models of supernovae and kilonovae". In: *Monthly Notices of the Royal Astronomical Society* 489.4, pp. 5037–5045. ISSN: 0035-8711. DOI: 10.1093/mnras/stz2495. eprint: https://academic.oup.com/mnras/article-pdf/489/4/5037/30739278/stz2495.pdf. URL: https://doi.org/10.1093/mnras/stz2495.

Collaboration, The LIGO Scientific et al. (Mar. 2015). "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7, p. 074001. DOI: 10.1088/0264-9381/32/7/074001. URL: https://dx.doi.org/10.1088/0264-9381/32/7/074001.

Coughlin, Michael W. et al. (Aug. 2020). "Measuring the Hubble constant with a sample of kilonovae". In: *Nature Communications* 11, 4129, p. 4129. DOI: 10.1038/s41467-020-17998-5. arXiv: 2008.07420 [astro-ph.HE].

Cranmer, Kyle, Johann Brehmer, and Gilles Louppe (May 2020). "The frontier of simulation-based inference". In: *Proceedings of the National Academy of Sciences* 117.48, pp. 30055–30062. DOI: 10.1073/pnas.1912789117. URL: https://doi.org/10.1073%2Fpnas.1912789117.

Dietrich, Tim, Michael W. Coughlin, et al. (Dec. 2020). "Multimessenger constraints on the neutron-star equation of state and the Hubble constant". In: *Science* 370.6523, pp. 1450–1453. DOI: 10.1126/science.abb4317. URL: https://doi.org/10.1126%2Fscience.abb4317.

Dietrich, Tim and Maximiliano Ujevic (Apr. 2017). "Modeling dynamical ejecta from binary neutron star mergers and implications for electromagnetic counterparts". In: *Classical and Quantum Gravity* 34.10, p. 105014. DOI: 10.1088/1361-6382/aa6bb0. URL: https://dx.doi.org/10.1088/1361-6382/aa6bb0.

Greenberg, David, Marcel Nonnenmacher, and Jakob Macke (Sept. 2019). "Automatic Posterior Transformation for Likelihood-Free Inference". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. Long Beach, California, USA: PMLR, pp. 2404–2414. URL: http://proceedings.mlr.press/v97/greenberg19a.html.

Ivezić, Ž. et al. (Mar. 2019). "LSST: From Science Drivers to Reference Design and Anticipated Data Products". In: *The Astrophysical Journal* 873, 111, p. 111. DOI: 10.3847/1538-4357/ab042c. arXiv: 0805.2366.

Kasen, Daniel et al. (Oct. 2017). "Origin of the heavy elements in binary neutron-star mergers from a gravitational-wave event". In: *Nature* 551.7678, pp. 80–84. DOI: 10.1038/nature24453. URL: https://doi.org/10.1038%2Fnature24453.

Kawaguchi, Kyohei, Masaru Shibata, and Masaomi Tanaka (Feb. 2020). "Diversity of Kilonova Light Curves". In: *The Astrophysical Journal* 889.2, p. 171. DOI: 10.3847/1538-4357/ab61f6. URL: https://doi.org/10.3847%2F1538-4357%2Fab61f6.

Li, Li-Xin and Bohdan Paczyń ski (Nov. 1998). "Transient Events from Neutron Star Mergers". In: *The Astrophysical Journal* 507.1, pp. L59–L62. DOI: 10.1086/311680. URL: https://doi.org/10.1086%2F311680.

Lueckmann, Jan-Matthis et al. (2021). *Benchmarking Simulation-Based Inference*. arXiv: 2101.04653 [stat.ML].

Lukoš iute, K et al. (Aug. 2022). "KilonovaNet: Surrogate models of kilonova spectra with conditional variational autoencoders". In: *Monthly Notices of the Royal Astronomical Society* 516.1, pp. 1137–1148. DOI: 10.1093/mnras/stac2342. URL: https://doi.org/10.1093%2Fmnras%2Fstac2342.

Margalit, Ben and Brian D. Metzger (Nov. 2017). "Constraining the Maximum Mass of Neutron Stars from Multi-messenger Observations of GW170817". In: *The Astrophysical Journal* 850.2, p. L19. DOI: 10.3847/2041-8213/aa991c. URL: https://doi.org/10.3847%2F2041-8213%2Faa991c.

Metzger, Brian D. (Dec. 2019). "Kilonovae". In: *Living Reviews in Relativity* 23.1. DOI: 10.1007/s41114-019-0024-0. URL: https://doi.org/10.1007%2Fs41114-019-0024-0.

Montel, Noemi Anau, James Alvey, and Christoph Weniger (2023). *Scalable inference with Autoregressive Neural Ratio Estimation*. arXiv: 2308.08597 [astro-ph.IM].

Radice, David et al. (Jan. 2018). "GW170817: Joint Constraint on the Neutron Star Equation of State from Multimessenger Observations". In: *The Astrophysical Journal Letters* 852.2, p. L29. DOI: 10.3847/2041-8213/aaa402. URL: https://dx.doi.org/10.3847/2041-8213/aaa402.

Rastinejad, Jillian C. et al. (Dec. 2022). "A kilonova following a long-duration gamma-ray burst at 350 Mpc". In: 612.7939, pp. 223–227. DOI: 10.1038/s41586-022-05390-w. arXiv: 2204.10864 [astro-ph.HE].

Shappee, B. J. et al. (Dec. 2017). "Early spectra of the gravitational wave source GW170817: Evolution of a neutron star merger". In: *Science* 358.6370, pp. 1574–1578. DOI: 10.1126/science.aaq0186. URL: https://doi.org/10.1126%2Fscience.aaq0186.

Soares-Santos, M., D. E. Holz, et al. (Oct. 2017). "The Electromagnetic Counterpart of the Binary Neutron Star Merger LIGO/Virgo GW170817. I. Discovery of the Optical Counterpart Using the Dark Energy Camera". In: *The Astrophysical Journal Letters* 848.2, L16, p. L16. DOI: 10.3847/2041-8213/aa9059. arXiv: 1710.05459 [astro-ph.HE].

Soares-Santos, M., A. Palmese, et al. (May 2019). "First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo BinaryBlack-hole Merger GW170814". In: 876, L7, p. L7. DOI: 10.3847/2041-8213/ab14f1. arXiv: 1901.01540.

Tejero-Cantero, Alvaro et al. (2020). "sbi: A toolkit for simulation-based inference". In: *Journal of Open Source Software* 5.52, p. 2505. DOI: 10.21105/joss.02505. URL: https://doi.org/10.21105/joss.02505.