Draft version November 7, 2021
Typeset using LATEX **twocolumn** style in AASTeX62

# Flexible and Fast Estimation of Binary Merger Population Distributions with Adaptive KDE

Jam Sadiq,[1] Thomas Dent,[1] and Daniel Wysocki[2]

[1] Instituto Galego de Física de Altas Enerxías, Universidade de Santiago de Compostela, Santiago de Compostela, Galicia, Spain
[2] University of Wisconsin–Milwaukee

## ABSTRACT

A total of more than 50 binary compact object mergers has been detected by the LIGO-Virgo gravitational wave (GW) detector network in their most recent observing run. The formation channels of these compact objects are still highly uncertain: to make progress we require a range of methods to characterize their population properties. The Bayesian hierarchical methods mainly employed so far become more computationally expensive as the size of event catalogs increases, and assume simple functional forms for the source distribution. Here we propose a fast and flexible method to reconstruct the population of LIGO-Virgo binary mergers without such assumptions. Under some conditions (sufficiently high statistics, sufficiently low individual event measurement error relative to width of population features) a kernel density estimator (KDE) reconstruction of the mass distribution from parameter estimation (PE) median masses will be sufficiently accurate. The method we are proposing improves the accuracy and flexibility of kernel density estimation (KDE) by using adaptive bandwidth (awKDE) and cross-validation to obtain an optimal bandwidth. We apply awKDE to publicly released parameter estimates for binary mergers in O1, O2, and O3a, in combination with a fast polynomial fit of search sensitivity, to obtain a non-parametric estimate of the mass distribution for comparison with established Bayesian hierarchical methods.

*Keywords:* adaptive-kde — compact binary stars

## 1. INTRODUCTION

Advanced LIGO's and Virgo's first three observing runs have produced tens of confident detections of binary compact object mergers published in GWTC-1 Abbott et al. (2019) and GWTC-2 Abbott et al. (2021). A detailed investigation of the population properties of binary black hole (BBH) mergers, the most commonly detected source type, has been published in Abbott et al. (2020) focusing on several population characteristics including their component masses and spin.

One objective has been to reconstruct the primary mass distribution of merging BBH, in order to address questions in stellar evolution, BH formation and binary formation channels. The primary BH masses have lower measurement uncertainty compared to other parameters such as binary spins, thus their distribution is expected to yield significant information, which may eventually be compared with astrophysical model predictions. In Abbott et al. (2020) specific functional forms and Bayesian hierarchical analysis are used to infer the mass distribution of binary mergers from these observed gravitational wave events: see Figure 1. These methods are well established but they are computationally extensive and with the growing observations, in future observing runs, can be more complicated.

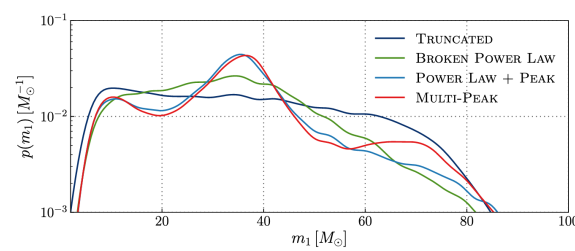


**Figure 1.** Observed primary black hole mass distribution predicted by each mass model in Abbott et al. (2020).

In this paper we are proposing a fast and flexible method for reconstructing binary merger population distributions via adaptive **Kernel Density Estimation** (KDE), using a publically available code (Menne 2020).

Corresponding author: Jam Sadiq
jam.sadiq@usc.es

This method addresses the same goal as the parametric hierarchical analyses, but without assuming any functional form of the distribution. A standard KDE with a fixed (global) bandwidth is unlikely to give a good representation of the BBH mass distribution. Available observed gravitational wave detected events are sufficient to use in an adaptive KDE (awKDE) to address the bandwidth problem and estimate the mass distribution of binary mergers. This method will validate as well as check the model assumptions of the standard Bayesian analyses and can potentially discover features not described by models in a fast, flexible, and computationally efficient way. We can further use this awKDE to compute rate estimates using an estimate of search sensitivity estimate as given in Wysocki et al. (2019).

The paper is organized this way. In section 2 we describe awKDE code and how we produce results that are comparable with published results. In section 3 we applied our awKDE onto observed LVK events and also show the procedure we follow for rate estimates. In section 4 we summarize and conclude.

## 2. METHOD

We propose the use of a kernel density estimate (KDE) with adaptive bandwidth selection for reconstructing the probability distribution of source parameters for compact binary mergers observed via GW. This method is non-parametric, straightforward to apply, and enables the identification of general features in the distributions that may be an important input in the astrophysical interpretation of the merging binary population. Other, more sophisticated and computationally demanding non-parametric methods have recently been described in (Tiwari & Fairhurst 2021; Tiwari 2021; Veske et al. 2021; Rinaldi & Del Pozzo 2021; Edelman et al. 2021). While a KDE with fixed (global) bandwidth is known to be a suitable method to estimate a distribution close to a single Gaussian, the mass (and possibly also spin) distributions of BBH mergers appear to show a more complex structure, which we do not expect to be well reconstructed by a simple KDE. In particular, for the primary and secondary masses the distribution may be composed of several components including power laws and one or more Gaussian peaks. Thus, we consider an extension of the KDE to allow the bandwidth to vary locally according to an initial estimate of the density of sample points Terrell & Scott (1992); Sain & Scott (1996).

This 'adaptive bandwidth KDE' (Wang & Wang 2011) is implemented in the open source code `awkde` (Menne 2020). We first describe the construction of a KDE from sample values $X_i$, $i = 1 \ldots n$: for instance, $X_i$ may be a measured property of a binary merger. Later, we will describe how measurement uncertainties in these individual event properties are incorporated in the estimate.

The algorithm computes a density estimate $\hat{f}$ using a Gaussian kernel density estimator

$$\hat{f}(x) = n^{-1} \sum_{i=1}^{n} \frac{1}{h\lambda_i} K\left(\frac{x - X_i}{h\lambda_i}\right) \tag{1}$$

where $K(.)$ is the standard Gaussian kernel, $n$ is the total number of samples and, in general, the product $h\lambda_i$ takes the role of a local bandwidth with $h$ being the global bandwidth. The first step in the algorithm is computation of a pilot estimate $\hat{f}_0$ setting $\lambda_i = 1$ for all $i$, which is a standard fixed bandwidth KDE. Based on the pilot density $\hat{f}_0$, the local bandwidth accounting for variations in the density of samples is obtained via

$$\lambda_i = \left(\frac{\hat{f}_0(X_i)}{g}\right)^{\alpha} \tag{2}$$

where $\alpha$ is the local bandwidth sensitivity parameter ($0 \leq \alpha \leq 1$) and $g$ is a normalization factor

$$\log g = n^{-1} \sum_{i=1}^{n} \log \hat{f}_0(X_i). \tag{3}$$

Finally the adaptive KDE $\hat{f}(x)$ is obtained by evaluating (1) with the variable (local) bandwidth $h\lambda_i$.

The above method assumes one-dimensional data $X_i$; the method may also be applied to two- or more-dimensional data by transforming the data to have zero mean and unit covariance matrix.

The method requires a choice of the initial global bandwidth and sensitivity parameter $\alpha$: we use the *leave-one-out cross-validation* method Hastie et al. (2001) to determine these values. As a figure of merit for the cross-validation we use the (log) likelihood,

$$\mathcal{L}_{\mathrm{LOO}} = \sum_{i=1}^{n} \log \hat{f}_{\mathrm{LOO,i}}(X_i), \tag{4}$$

where $\hat{f}_{\mathrm{LOO,i}}$ is the KDE constructed from all samples *except* $X_i$. We use a grid search over different global bandwidth and $\alpha$ values and select those with the highest $\mathcal{L}_{\mathrm{LOO}}$. We often find that likelihood is maximized for $\alpha$ at or close to 1, thus in some cases we will impose $\alpha = 1$ rather than conduct a full 2d grid search.

### 2.1. *Application to GW observations*

The component masses of observed GW binary mergers have significant uncertainties, and we wish to incorporate these uncertainties in reconstructions of the mass distribution. The parameters of each binary are obtained by Bayesian inference using models of the emitted GW waveform (e.g. Veitch et al. (2015)). We consider detected mergers labelled by $i = 1 \ldots n$: uncertainties in a given parameter $X$ are quantified via MCMC sample values $X_i^k$, whose density is proportional to the posterior PDF for the $i$'th merger. Typically, some thousands of parameter samples are available per merger event LVC (2020).

Although we use a random selection of these PE samples when performing the optimized KDE including measurement uncertainty, to obtain the optimum global bandwidth and sensitivity parameter $\alpha$ we first evaluate the likelihood $\mathcal{L}_{\rm LOO}$ using as data $X_i$ only the *median* parameter value for each merger. This choice reduces computational cost as well as avoiding over-fitting of random fluctuations: note that these medians are independent values, whereas the PE samples for a given event $i$, $X_i^k$, are not independent of one other. Having obtained the optimal $h$ and $\alpha$ choices, we construct the population KDE using 100 randomly chosen samples for each event. As noted above, in most cases the optimum $\alpha$ is found to be at or close to 1, thus we often fix $\alpha = 1$ for convenience. We also verify that increasing the number of PE samples in constructing the population KDE did not change our results significantly. We note that for large measurement errors relative to the scale of structures in the underlying distribution, the KDE result is likely to be over-dispersed, i.e. any sudden variations in the actual density will be smoothed out.[1]

We also compute an uncertainty estimate (confidence interval) for the population KDE using the *bootstrap* technique (Efron 1979). Our major source of uncertainty lies in the finite number of binary merger events and the resulting count fluctuations in the estimate at a given parameter value. We account for this uncertainty by bootstrap resampling over the *merger events* $i$, i.e. each bootstrap iteration contains some number of copies of (the PE samples for) each event $i$, with the number of copies approximately following a Poisson distribution. We conduct 1000 bootstrap iterations and then extract the median, 5th and 95th percentiles as an estimate of the confidence interval at any given parameter value. By plotting all bootstrap KDEs we can visually identify regions of high and low uncertainties in our population KDE.

[1] In principle this bias can be tackled by the computationally demanding hierarchical methods, if they model functional forms that are sufficiently close to the real distribution.

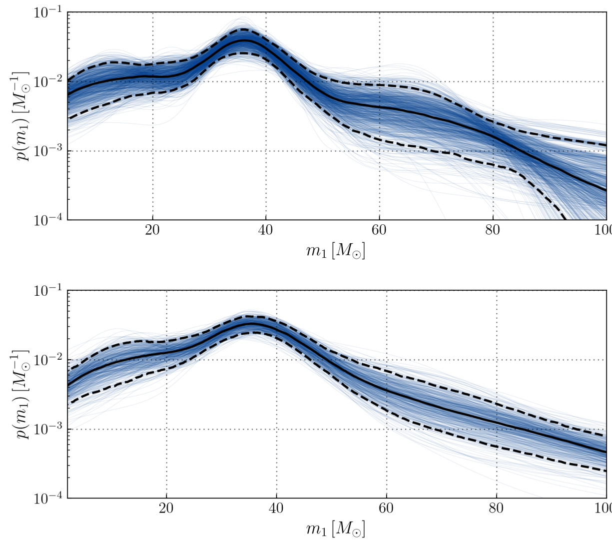


**Figure 2.** KDE with uncertainty estimates of source frame primary mass, $m_1$, for detected events in the O1, O2 and O3a observing runs. Top: KDE using the PE sample median for each event. Bottom: KDE constructed using 100 random samples from each observed event. The median (black solid), 90% confidence interval (black dashed lines) and blue curves are constructed using a bootstrap. Our KDE results match closely the Power Law + Peak and Multi Peak model results in Abbott et al. (2020), see Fig. 1.

## 3. RESULTS OF AWKDE FOR LIGO-VIRGO DETECTIONS FROM O1, O2 AND O3A OBSERVING RUNS

Here we apply these methods to reconstruct the distribution of detected primary mass $m_1$ (in the binary source frame) for the BBH observed in the O1, O2 and O3a runs (Abbott et al. 2019, 2021). The resulting KDE bootstrap samples, median estimate and 90% confidence intervals are shown in Figure 2, where the top panel uses only the median $m_1$ value for each binary merger and the lower panel accounts for measurement uncertainties via the PE samples. These primary mass distributions match well those inferred using parametric models in Abbott et al. (2020), specifically with the Power Law + Peak and Multi Peak models most preferred by the GWTC-2 data. We note that there is a clear global maximum in the observed distribution between $\sim 35-40\,\mathrm{M}_\odot$and that the uncertainty in the mass distribution is smallest here. The awKDE is also able to reconstruct a wide dynamic range of densities without excessive statistical uncertainties.

As an alternative to the primary mass, we also consider the chirp mass $\mathcal{M}$ for observed BBH mergers, which is also measured with relatively small uncertainty. The corresponding KDEs are shown in Figure 3.

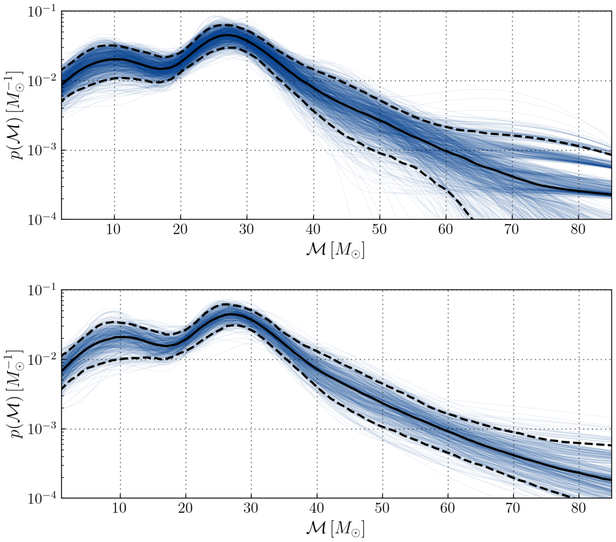


**Figure 3.** KDE with uncertainty estimates of source chirp mass, $\mathcal{M}$for detected events in the O1, O2 and O3a observing runs. Top: KDE using the PE sample median for each event. Bottom: KDE constructed using 100 random samples from each observed event. The median (black solid), 90% confidence interval (black dashed lines) and blue curves are constructed using a bootstrap. In addition to the principal peak in the distribution around 25–30 $M_\odot$, there is a hint of further structure around 10 $M_\odot$.

### 3.1. *Merger rate estimation from awKDE*

We also consider estimation of BBH merger rates using our awKDE results. The additional ingredient in this analysis is the sensitive volume $V$ within which a source is detectable: we quantify this based on an approximation for sensitivity of GW detectors, with corrections to account for the actual behaviour of searches in real data (Wysocki et al. 2019). The idea is to calibrate a semi-analytic function $VT_{\rm analytic}$ against the re-weighted result of injections (i.e. simulated signals added to real data and analyzed by search pipelines) $VT_{\rm inj}$, assuming a parameterized relationship between them that can be expressed via basis functions.

For the O1-O2 and O3a observing runs, the 'uncalibrated' semi-analytic sensitivity is estimated using the criterion that the signal-to-noise ratio (SNR) of a signal in the second most sensitive detector, using specific reference power spectral densities, should be greater than 8. This semi-analytic estimate $VT_{\rm analytic}$ is obtained for a given observing time $T$ over a grid of intrinsic source parameters such as component masses. For the injection $VT_{\rm inj}$, one performs a set of injections and counts the number detected by search pipelines to obtain an average $VT_{\rm inj}$ for the injected population at given intrinsic parameters. The calibrated estimate of $VT$ is then obtained by fitting a correction function to the deviations

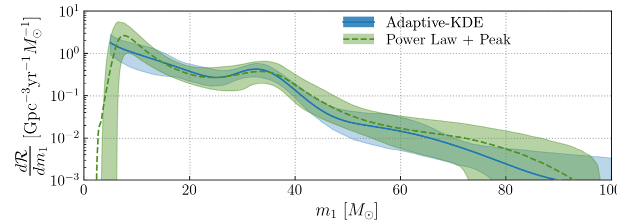


**Figure 4.** Rate estimates using sensitive volume and adaptive KDE. Green curve is the rate estimate using the Power Law + Peak method from Abbott et al. (2020) where blue curve is the rate estimate using KDE results from median PE values and using VT from Wysocki et al. (2019); Wysocki (2020)

of $VT_{\rm inj}$ from $VT_{\rm analytic}$, as described in Wysocki (2020) and applied in Abbott et al. (2020).

We compute the merger rate density using this corrected $VT$ and our KDE results for the primary mass $\hat{f}(m_1)$ from observations up to O3a. We obtain $VT(m_1)$ assuming a power law mass distribution for the secondary $m_2$, taking the median value of the power from Abbott et al. (2020), and find the merger rate density via

$$\frac{dR}{dm_1} = N\frac{\hat{f}(m_1)}{VT(m_1)}. \quad (5)$$

Our estimate is in good agreement with the results in Abbott et al. (2020), as shown in Fig. 4.

### 3.2. *2D KDE for cosmological evolution of BH mass distribution*

We apply our adaptive KDE in a two-dimensional space using source frame median PE values for $m_1$ and luminosity distance $D_L$ for all events in the O1-O2 and O3a (GWTC-2) observing runs, following the same procedure of leave-one-out cross-validation to determine the optimal global bandwidth and sensitivity parameter. As shown in Figure 5, the resulting two dimensional KDE shows peaks around 10 and 35–40 $M_\odot$, similarly to the one-dimensional $m_1$ KDE; these over densities appear to be present consistently over different distances.

## 4. CONCLUSIONS

We introduce a fast and efficient method of using adaptive KDE to get binary merger population distributions. This method can be used for sanity checks for computational extensive models for binary merger population distributions. This methods can use PE sample values of each events and can take into account uncertainty estimates. Results of KDE can be used for rate estimates. This methods can be extended for higher dimensions given more observational results. A possible application of this method can be in cosmology.

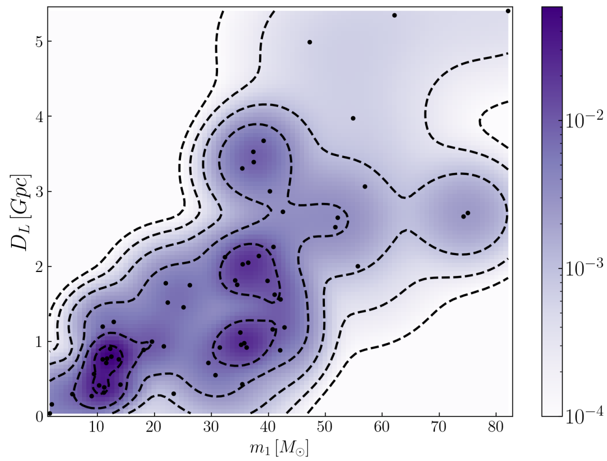


**Figure 5.** Using median values of parameters for observed BBH events adaptive KDE is computed in two dimensions using source frame primary mass and luminosity distance. The features around $10M_\odot$ and $40M_\odot$ seem to be present consistently at different distances.


## ACKNOWLEDGEMENT

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org/ ), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain.


## REFERENCES


Abbott, B. P., et al. 2019, Phys. Rev. X, 9, 031040, doi: 10.1103/PhysRevX.9.031040

Abbott, R., et al. 2020. https://arxiv.org/abs/2010.14533

—. 2021, Phys. Rev. X, 11, 021053, doi: 10.1103/PhysRevX.11.021053

Edelman, B., Doctor, Z., Godfrey, J., & Farr, B. 2021. https://arxiv.org/abs/2109.06137

Efron, B. 1979, The Annals of Statistics, 7, 1 . https://doi.org/10.1214/aos/1176344552

Hastie, T., Tibshirani, R., & Friedman, J. 2001, The Elements of Statistical Learning, Springer Series in Statistics (New York, NY, USA: Springer New York Inc.)

LVC. 2020, GWTC-2 Data Release: Parameter Estimation Samples and Skymaps, https://dcc.ligo.org/LIGO-P2000223/public/

Menne, T. 2020, awKDE code, https://github.com/mennthor/awkde

Rinaldi, S., & Del Pozzo, W. 2021. https://arxiv.org/abs/2109.05960

Sain, S. R., & Scott, D. W. 1996, Journal of the American Statistical Association, 91, 1525. https://www.tandfonline.com/doi/abs/10.1080/01621459.1996.10476720

Terrell, G. R., & Scott, D. W. 1992, The Annals of Statistics, 20, 1236 . https://doi.org/10.1214/aos/1176348768

Tiwari, V. 2021, Class. Quant. Grav., 38, 155007, doi: 10.1088/1361-6382/ac0b54

Tiwari, V., & Fairhurst, S. 2021, Astrophys. J. Lett., 913, L19, doi: 10.3847/2041-8213/abfbe7

Veitch, J., et al. 2015, Phys. Rev. D, 91, 042003, doi: 10.1103/PhysRevD.91.042003

Veske, D., Bartos, I., Márka, Z., & Márka, S. 2021. https://arxiv.org/abs/2105.13983

Wang, B., & Wang, X. 2011, Bandwidth Selection for Weighted Kernel Density Estimation. https://arxiv.org/abs/0709.1616

Wysocki, D. 2020, Calibrating semi-analytic VT's to injections in O3a, https://dcc.ligo.org/LIGO-T2000432/public

Wysocki, D., Lange, J., & O'Shaughnessy, R. 2019, Physical Review D, 100. http://dx.doi.org/10.1103/PhysRevD.100.043012