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, Spain

[2] University of Wisconsin—Milwaukee, Milwaukee, Wisconsin, USA

Abstract

The LIGO Scientific, Virgo and KAGRA Collaborations recently released the third gravitational wave transient catalog or GWTC-3, significantly expanding the number of gravitational wave (GW) signals from binary compact object mergers. 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. Here **we propose a fast and flexible adaptive bandwidth KDE method to reconstruct the population of LIGO–Virgo merging black hole binaries**. We propose this in combination with a fast polynomial fit (to injection results) of the mass-dependent sensitivity as **a simple non-parametric method** for comparison with 'established' Bayesian hierarchical models. **We also demonstrate a robust peak detection algorithm** based on awKDE and use it to calculate the significance of the apparent peak in the black hole (BH) mass distribution around $35M_{\odot}$.

1. Introduction

- Recent studies [1, 2] used specific functional forms and Bayesian hierarchical analysis for mass distribution of binary mergers from the observed GW events
- We propose **adaptive Kernel Density Estimation** (awKDE) method [3] to answer the same question, but without assuming any functional form of the distribution [4]
- KDE with a single global bandwidth unlikely to give good representation of Binary Black Hole (BBH) distribution

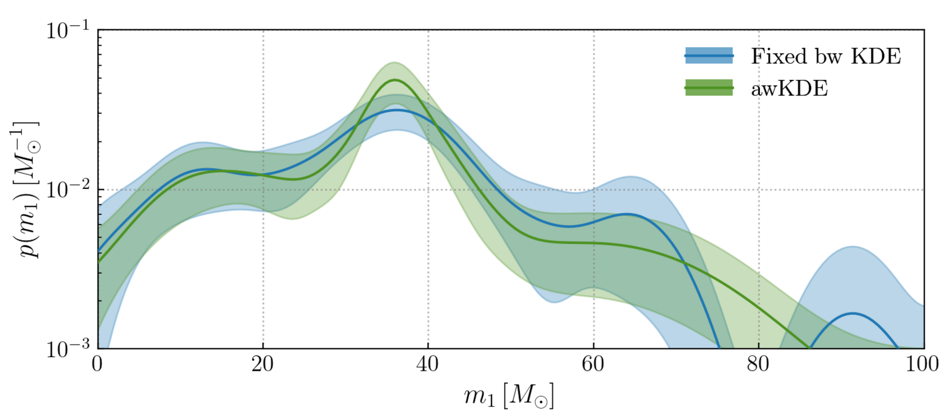

Figure 1: *Fixed global bandwidth versus awKDE reconstruction of primary BH masses using GWTC-2 detected events. The fixed bandwidth KDE overestimates the width (underestimates the height) of the observed peak around $35M_{\odot}$, but also yields a probably unphysical gap in the estimated distribution around $80M_{\odot}$*

- Observed gravitational wave events are sufficient to use in an awKDE to address the bandwidth problem and estimate the mass distribution of binary mergers
- This method will **validate/check the model assumptions** of the standard Bayesian analysis
- Potentially **discover features** not described by models
- **Fast** and computational **efficient** method

2. Method

- Use an open source AWKDE [3] code and modify it.
- Using an initial global bandwidth choice to make a pilot **Gaussian KDE**,
$$\hat{f}(x) = n^{-1}\sum_{i=1}^{n}\frac{1}{h\lambda_i}K\left(\frac{x-X_i}{h\lambda_i}\right), \quad \lambda_i = 1$$
- Derive local (adaptive) bandwidth from pilot KDE & sensitivity parameter, $0 \leq \alpha \leq 1$
$$\lambda_i = \left(\frac{\hat{f}_0(X_i)}{g}\right)^{-\alpha}, \quad \log g = n^{-1}\sum_{i=1}^{n}\log \hat{f}_0(X_i)$$
- Final KDE uses the local bandwidth
- Determine optimum initial bandwidth and α by **leave-one-out cross validation** with
$$\log \mathcal{L}_{\rm LOO} = \sum_{i=1}^{n}\log \hat{f}_{{\rm LOO},i}(X_i),$$
where $\hat{f}_{{\rm LOO},i}$ is the KDE constructed from all samples except X_i
- Compute awKDE for primary mass m_1 and chirp mass $\mathcal{M}$ using samples of PE results
- Compute **uncertainty estimates** using **bootstrap technique**

3. AwKDE Reconstruction Results

- The **GWTC-2** includes events from the **O1, O2 and O3a** observing runs [5, 6]
- The **GWTC-3** includes events from **O3b** [7] runs and an **updated** release for data and **events up to O3a**, called GWTC-2.1 [8]
- Use **samples values** of observed **BBH events**, **44 for GWTC-2 and 69 for GWTC-3** set, and compute primary mass, m_1, and chirp mass, $\mathcal{M}$, distribution using awKDE

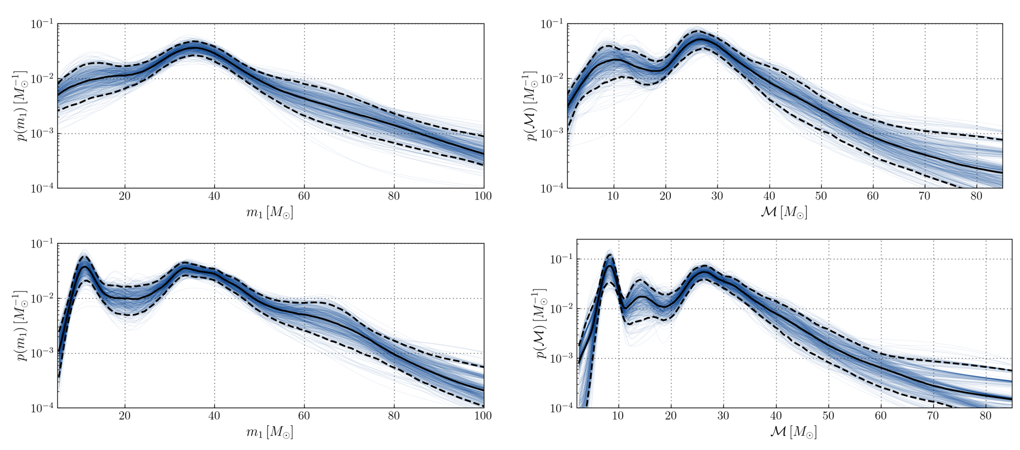

Figure 2: *AwKDE for m_1 and $\mathcal{M}$ constructed using 100 random samples from each BH observed event. The median (black solid), 90% confidence interval (black dashed lines) and blue curves are constructed using a bootstrap. Top plots are the results for detected BBH events in GWTC-2, match closely with the results in [1]. Bottom plots are KDE reconstructions from GWTC-3 events, showing additional features at lower masses consistent with the results in [2].*

4. Merger Rate Estimation from AwKDE

- **Estimate BBH differential merger rate** over m_1 using our awKDE results and a fast polynomial fit of the mass-dependent sensitivity to injection results
- Quantify the **sensitive volume** V, within which a source is detectable based on an approximation for the sensitivity of GW detectors, with corrections to account for the actual behaviour of searches in real data

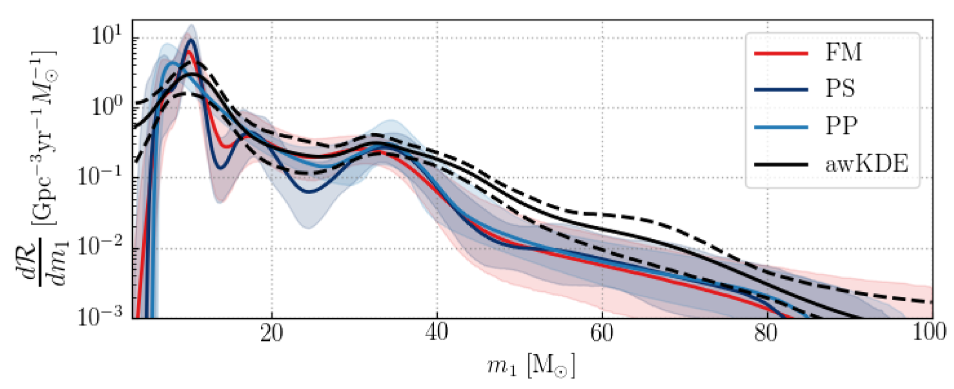

Figure 3: *Rate estimates using awKDE and sensitive volume for BBH events in GWTC-3. Black curves show the rate estimate using the KDE reconstruction in the bottom left panel of Fig. 2 and a fit to search sensitivity, compared with the* FLEXIBLE MIXTURES *(FM),* POWER LAW + SPLINE *(PS), and* POWER LAW + PEAK *(PP) mass models, as reported in [2].*

5. Two Dimensional AwKDE

- **Estimate two dimensional awKDE for cosmological evolution** of BH mass distribution using PE values for m_1 in the source frame and luminosity distance D_L for BBH events in GWTC-3 [7]

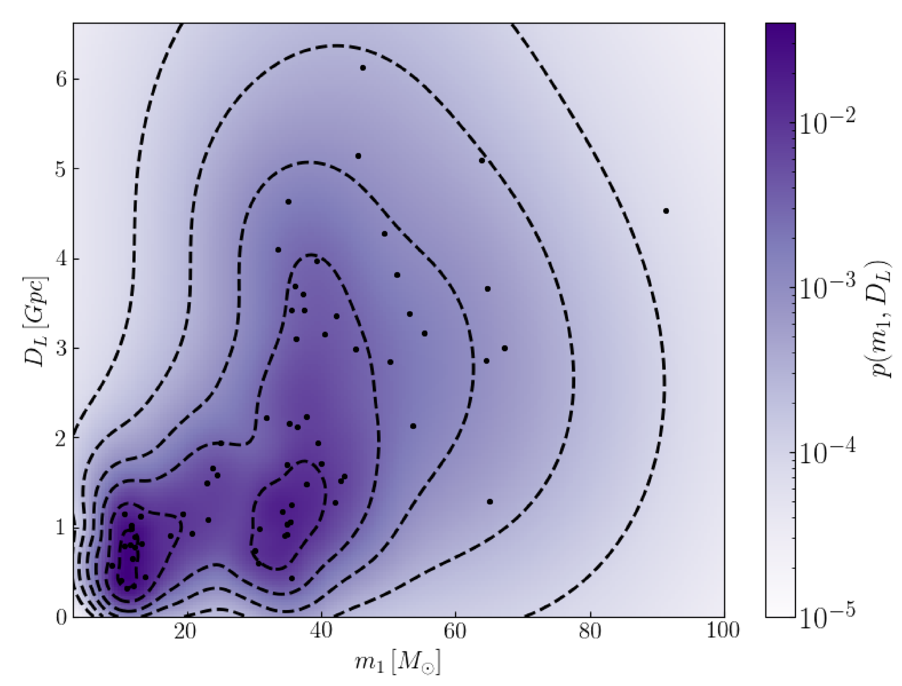

Figure 4: *AwKDE over the source frame primary mass and luminosity distance for BBH events in GWTC-3. Plot shows the distribution reconstructed from 100 PE samples per event with features around $10M_{\odot}$ and $35M_{\odot}$ appear to be present consistently over different distances.*

6. Detection of BH mass Function Peak

- Propose a method to **detect significance of prominent peak** around $35M_{\odot}$ in primary mass distribution
- For a given distribution $\hat{f}(x)$, find all local maxima, denoted $\{x_j^{\rm p}\}$.
- Using a window of given size δ (the choice of δ is discussed below), evaluate the PDF on either side of the peak, $\hat{f}_j^{\pm} = \hat{f}(x_j^{\rm p} \pm \delta)$
- Compute peak heights using $H_j^{\rm p} = \hat{f}(x_j^{\rm p}) - (\hat{f}_j^{-} + \hat{f}_j^{+})/2.$
- Evaluate the estimated uncertainty at each peak, $\epsilon_j^{\rm p}$,
- Determine the **most significant peak** by **maximizing a detection statistic**, $\mathcal{D}$, which combines $H_j^{\rm p}$ and $\epsilon_j^{\rm p}$

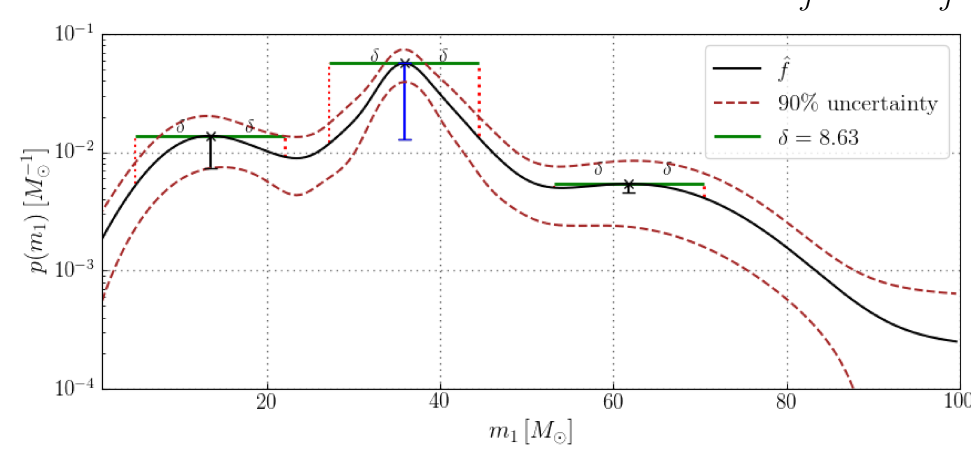

Figure 5: *Peak detection algorithm for KDEs using a window size δ (green lines), uncertainty estimates, ϵ (brown dashed lines), and peak heights (vertical blue/black lines) around each peak.*

- Tune bandwidth, h, δ, and choice of detection statistic $\mathcal{D}$ and use ROCs to get optimized results
$$\delta = 3\,h, \quad \mathcal{D} = H_p/\hat{\epsilon}_p$$
- For real data GWTC-2 (GWTC-3) events we get $\mathcal{D}$ = **2.85 (3.36)**
- Using background data from hyper-parameters of truncated power law in [1] with sensitive volume corrections, we get FAP, $\mathcal{F}$ = **0.00013 , (0.0012)**
- The resulting peak detection **significance** is **3.6σ, (3.0σ)**

7. Conclusion and Future

- Introduce a **fast and efficient method** of using awKDE to get binary merger population distributions
- AwKDE results can be used for **rate estimates**
- AwKDE results can be used for **sanity checks for computational extensive models** for binary merger population distributions
- Method can be **extended for higher dimensions** and can have possible **application in cosmology**
- **Develop a statistic for peak prominence** to probe if it is a real astrophysical feature or a random excess

References

[1] R. Abbott et al. Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. 10 2020.

[2] R. Abbott et al. The population of merging compact binaries inferred using gravitational waves through GWTC-3. 11 2021.

[3] Thorben Menne. awKDE code. https://github.com/mennthor/awkde, with modifications at https://github.com/JAMSADIQ/awkde, 2021. [Online; accessed Jul 2021].

[4] Jam Sadiq, Thomas Dent, and Daniel Wysocki. Flexible and Fast Estimation of Binary Merger Population Distributions with Adaptive KDE. 12 2021.

[5] B. P. Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev. X*, 9(3):031040, 2019.

[6] R. Abbott et al. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. *Phys. Rev. X*, 11:021053, 2021.

[7] R. Abbott et al. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run. 11 2021.

[8] R. Abbott et al. GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. 8 2021.