

Are all Binary Black Holes Detected by LIGO-Virgo-KAGRA Following the Universal Time-Delay Distributions? Probably Not

SAMSUZZAMAN AFROZ ¹, NAVDHA ², AND SUVODIP MUKHERJEE ¹

¹*Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, Mumbai 400005, India*

²*Indian Institute of Technology Bombay, Mumbai-400076, India*

ABSTRACT

The delay time distribution (DTD) of binary black hole (BBH) mergers encodes the evolutionary link between the formation history and gravitational-wave (GW) emission. We present a non-parametric reconstruction of the mass-dependent DTD using the BBHs from the GWTC-4 that avoids restrictive assumptions of only power-law forms. Our analysis reveals for the first time the signature for mass-dependent evolutionary pathways: lower-mass systems ($20\text{--}40 M_{\odot}$) are consistent with a scale-invariant DTD, whereas higher-mass BBHs ($40\text{--}100 M_{\odot}$) provide the first direct tentative evidence of DTD that deviate from simple power laws, with a pronounced preference for rapid mergers around $2 - 6$ Gyrs. These findings reveal the advantage of the non-parametric technique in reconstructing the mass-dependent DTD and discovering for the first-time the presence of a potential time-scale associated with high-mass GW events.

1. INTRODUCTION

The delay time distribution (DTD) of binary black hole (BBH) mergers represents one of the most fundamental quantities connecting stellar evolution to gravitational-wave (GW) observations. It encodes the probability distribution of times between the formation of massive stellar progenitors and their eventual coalescence as compact binaries, directly linking the cosmic star formation history to the observable merger rate across redshift. Understanding these evolutionary timescales is crucial for constraining binary evolution physics, supernova dynamics, and the efficiency of compact object formation channels (Bailyn et al. 1998; Dominik et al. 2012; Barrett et al. 2018; Mapelli 2021; Franciolini et al. 2022; Bouffanais et al. 2021; Antonelli et al. 2023; Cheng et al. 2023; Ye & Fishbach 2024; Afroz & Mukherjee 2025a,b,c,d).

Gravitational-wave astronomy has opened unprecedented opportunities to constrain DTDs through direct observations of BBH mergers across cosmic time (Abbott et al. 2016a, 2019, 2021, 2023; LIGO-VIRGO-KAGRA Collaboration 2025a). The relationship between stellar formation and merger observations is fundamentally described by a convolution integral that re-

lates the present-day merger rate to the historical star formation rate weighted by the DTD. This connection enables inference of evolutionary timescales from the mass and redshift distribution of detected GW events (Dominik et al. 2012, 2013; Barrett et al. 2018; Mapelli 2016; Mukherjee 2022; Bailes et al. 2021; Arimoto et al. 2021; Iorio et al. 2022; Sicilia et al. 2022; Schiebelbein-Zwack & Fishbach 2024; Rinaldi et al. 2024, 2025; Genari et al. 2025; Lalleman et al. 2025).

Most of the previous approaches to DTD inference from GWTC-3 have not found any evidence of departure from a simple power-law form (Fishbach et al. 2018; Safarzadeh et al. 2020; Mukherjee et al. 2021; Fishbach & Kalogera 2021; Karathanasis et al. 2023a; Smith & Kaplinghat 2024). While the simple power-law forms have provided valuable constraints on characteristic timescales, they inherently restrict the range of evolutionary behaviors that can be discovered. More flexible or semi-agnostic approaches have also been developed (Edelman et al. 2023; Callister & Farr 2024; Stevenson & Clarke 2022; Heinzl et al. 2024). Currently there are limitations both theoretical modeling as well as from data analysis techniques to capture complex DTDs arising from multiple formation channels, environmental dependencies, or mass-dependent evolutionary pathways.

The mass dependence of DTDs is particularly important from both theoretical and observational perspectives. Massive BBH progenitors experience different stellar evolution pathways, including altered wind

samsuzzaman.afroz@tifr.res.in

suvodip@tifr.res.in

mass loss rates, distinct supernova dynamics, and potentially different formation environments. Systems forming through isolated binary evolution may follow different delay time characteristics compared to those assembled through dynamical interactions in dense stellar environments. Additionally, the efficiency of common-envelope evolution and the survival probability of wide binaries may depend sensitively on component masses (Belczynski et al. 2016; Rodriguez et al. 2016; Mapelli 2016; Dominik et al. 2012, 2013; Mandel & Farmer 2022).

In this work, we introduce a novel grid-based approach for non-parametric inference of DTD that addresses the limitations of traditional parametric methods, and allows to infer mass-dependent DTD and local merger rate R_0 in a model-independent way from GW observations. Our framework divides the delay time-probability space into a structured grid and systematically explores all physically allowed evolutionary trajectories. The key innovation is the incorporation of causality constraints: evolutionary trajectories can progress forward in delay time with varying merger probabilities but cannot violate the fundamental requirement that delay times increase monotonically.

This grid-based methodology provides several advantages over conventional approaches. First, it enables discovery of complex, non-monotonic DTD shapes without assuming specific functional forms. Second, it naturally incorporates physical constraints while maintaining maximum flexibility for capturing unexpected features. Third, it provides a systematic framework for model comparison through Bayesian evidence calculations, enabling objective selection among competing evolutionary scenarios.

We apply this framework to the GWTC-4 (LIGO-VIRGO-KAGRA Collaboration 2025b) catalog (including GWTC-3 (Abbott et al. 2023)), detected by the LIGO (Abbott et al. 2016b), Virgo (Acernese et al. 2015), and KAGRA (Akutsu et al. 2021) (LVK) collaboration. The analysis reveals a tentative evidence of mass dependence of the DTD: lower-mass binaries ($20\text{--}40 M_\odot$) favor DTDs that appear broadly flat in the $(t_d, \log P(t_d))$ space across $\sim 1\text{--}10$ Gyr, consistent with mergers occurring over a wide range of evolutionary timescales. By contrast, higher-mass binaries ($40\text{--}100 M_\odot$) exhibit a tentative evidence of non-power law behavior with a sharply peaked distributions at short delays ($\sim 2\text{--}6$ Gyr), indicative of rapid post-formation mergers. These findings provide new insights into mass-dependent binary evolution and demonstrate the power of non-parametric approaches for extracting detailed astrophysical information from GW observations.

2. THEORETICAL FRAMEWORK AND GRID CONSTRUCTION

The theoretical foundation for merger rate density inference rests on the fundamental relationship between BBH formation, evolution, and the observable GW merger rates across cosmic time. For a population of BBHs with component masses m_1 and m_2 , the differential merger rate density at redshift z can be expressed as

$$\frac{dN_{\text{GW}}}{dz}(z, m) = T_{\text{obs}} \frac{dV_c}{dz} \frac{R(z, m)}{(1+z)} \cdot p(m|z) \cdot \mathcal{S}(m, z), \quad (1)$$

where $R(z, m)$ represents the intrinsic merger rate density we seek to infer, $p(m|z)$ is the normalized mass distribution at redshift z , T_{obs} is the observation time, dV_c/dz is the comoving volume element, $(1+z)^{-1}$ accounts for cosmological time dilation, and $\mathcal{S}(m, z)$ accounts for detector selection effects derived from injection campaigns.

The merger rate density can be linked to the cosmic star formation history through a convolution with the DTD. In normalized form, it can be written as (Dominik et al. 2015; Karathanasis et al. 2023b; Mukherjee 2022; Karathanasis et al. 2023a)

$$R(z, m) = R_0(m) \frac{\int_z^\infty p_t(t_d(m)) R_{\text{SFR}}(z_f) \frac{dt}{dz_f} dz_f}{\int_0^\infty p_t(t_d(m)) R_{\text{SFR}}(z_f) \frac{dt}{dz_f} dz_f}, \quad (2)$$

where $R_{\text{SFR}}(z_f)$ denotes the cosmic star formation rate density at formation redshift z_f , $p_t(t_d(m))$ is the DTD (potentially mass dependent), and $t_d(m) = t(z) - t(z_f(m))$ is the delay between formation and merger. The denominator provides normalization, ensuring that $R(z, m)$ has a proper probabilistic interpretation.

A widely used parameterization of the star formation rate is the Madau-Dickinson model (Madau & Dickinson 2014):

$$R_{\text{SFR}}(z) = \frac{0.015(1+z)^{2.7}}{1 + \left[\frac{1+z}{2.9}\right]^{5.6}} M_\odot \text{ Mpc}^{-3} \text{ yr}^{-1}. \quad (3)$$

It is important to note that in this work we assume the merger rate follows the cosmic star formation history convolved with a DTD. However, since the birth metallicity distribution evolves with redshift and the formation of massive BBHs is strongly suppressed at high metallicities (Belczynski et al. 2010; Mapelli 2016; Lamberts 2018; Mukherjee 2022; Afroz & Mukherjee 2025a; Neijssel et al. 2019), this simplification does not explicitly account for metallicity-dependent formation

efficiency. Our main goal in this paper is to identify whether there is any deviation from an universal power-law form of the delay-time distribution in a mass-dependent way. Moreover, it is also important to note that the metallicity-dependent star formation rate as a function of redshift is not well measured from observations. As a result, in this analysis which is the first-of-its-kind from GWTC-4, we do not allow a metallicity dependent star formation rate. Given the limited number of GW events currently available, we restrict our analysis to this effective approximation.

2.1. Grid-Based Delay Time Distribution Construction

The key innovation of our approach lies in the non-parametric construction of DTDs through a structured grid methodology. Rather than assuming specific functional forms for $p_t(t_d)$, we discretize the delay time-probability space and systematically explore all physically allowed evolutionary trajectories. We construct a structured grid in the $(t_d, \log p_t(t_d))$ plane, where t_d represents delay time and $p_t(t_d)$ denotes the probability density of the DTD. The logarithmic transformation of the probability axis ensures numerical stability while capturing distributions spanning multiple orders of magnitude.

The delay time axis is divided from $t_{d,\min}$ to $t_{d,\max}$ using N_t equally spaced grid points in logarithmic space:

$$t_d^{(i)} = t_{d,\min} \left(\frac{t_{d,\max}}{t_{d,\min}} \right)^{(i-1)/(N_t-1)}, \quad i = 1, 2, \dots, N_t. \quad (4)$$

The probability density values are divided using N_p grid points in logarithmic space:

$$\log p_t^{(j)} = \log p_{t,\min} + (j-1)\Delta \log p_t, \quad j = 1, 2, \dots, N_p, \quad (5)$$

where $\Delta \log p_t = (\log p_{t,\max} - \log p_{t,\min})/(N_p - 1)$ provides adequate dynamic range for diverse DTD shapes.

The fundamental physical constraint governing our grid exploration is causality: delay times cannot decrease as we move forward through the evolutionary sequence. Each trajectory begins from a single starting point in the grid and can only progress to delay times that are greater than the current position, while the probability values can move freely upward or downward to any accessible grid point. This constraint ensures that the evolutionary pathways respect the fundamental arrow of time while allowing maximum flexibility in exploring different DTD shapes. Any valid trajectory through the grid must satisfy:

$$\mathcal{T} = \{(t_d^{(i_1)}, \log p_t^{(j_1)}), (t_d^{(i_2)}, \log p_t^{(j_2)}), \dots, (t_d^{(i_{N_t})}, \log p_t^{(j_{N_t})})\}, \quad (6)$$

with the causality constraint:

$$i_1 \leq i_2 \leq i_3 \leq \dots \leq i_{N_t} \quad (\text{no backward time evolution}). \quad (7)$$

The probability values $p_t^{(j)}$ can increase or decrease along the trajectory without restriction, enabling discovery of complex DTD shapes including multi-modal structures, peaked distributions, and broad evolutionary timescales. This flexibility allows trajectories to capture sharp transitions between evolutionary phases, gradual changes in formation efficiency, or any combination of evolutionary behaviors that might characterize different mass ranges or formation channels.

For each valid trajectory $\mathcal{T}$, we construct a continuous DTD by interpolating between grid points using cubic splines:

$$p_t^{\text{raw}}(t_d|\mathcal{T}) = \text{CubicSpline} \left(\{t_d^{(i_k)}, p_t^{(j_k)}\}_{k=1}^{N_t} \right). \quad (8)$$

The interpolated function is then normalized to ensure proper probability density normalization:

$$p_t(t_d|\mathcal{T}) = \frac{p_t^{\text{raw}}(t_d|\mathcal{T})}{\int_0^\infty p_t^{\text{raw}}(t_d|\mathcal{T}) dt_d}. \quad (9)$$

This procedure converts each discrete trajectory into a continuous, normalized DTD suitable for convolution with the cosmic star formation history.

2.2. Bayesian Inference Framework

For each trajectory $\mathcal{T}$, we calculate the expected number of GW detections in redshift bin j and mass bin k :

$$N_{\text{exp}}(z_j, m_k|\mathcal{T}) = T_{\text{obs}} \int_{z_j}^{z_{j+1}} R(z, m_k|\mathcal{T}) \frac{dV_c}{dz} \frac{\mathcal{S}(m_k, z)}{1+z} dz, \quad (10)$$

where $R(z, m_k|\mathcal{T})$ is computed via Equation 2 and $\mathcal{S}(m_k, z)$ is the detector selection function from injection campaigns.

We model the detection process using Poisson statistics, yielding the likelihood:

$$\mathcal{L}_k(\mathcal{T}) = \prod_{j=1}^{N_z} \frac{[N_{\text{exp}}(z_j, m_k|\mathcal{T})]^{N_{\text{obs}}(z_j, m_k)}}{N_{\text{obs}}(z_j, m_k)!} \times \exp[-N_{\text{exp}}(z_j, m_k|\mathcal{T})]. \quad (11)$$

where $N_{\text{obs}}(z_j, m_k)$ is the observed number of events in each bin. This framework uses the Dynesty sampler (Speagle 2020) to calculate the posterior on the delay-time distribution using a normalised prior on all the allowed trajectories, along with the evidence for all the

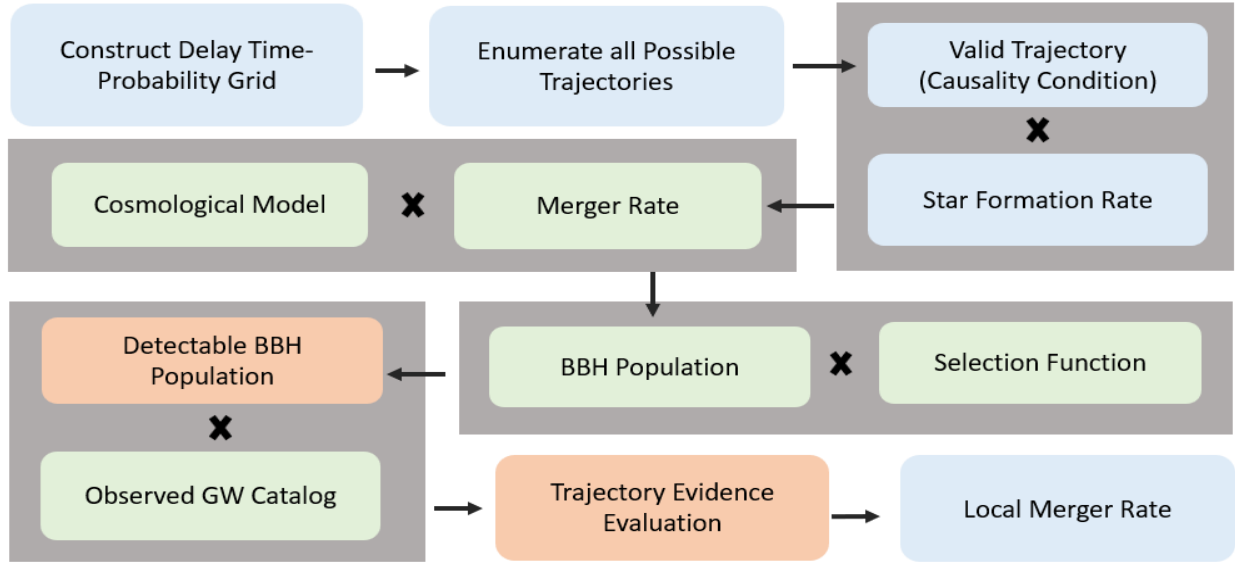


Figure 1. Flowchart summarizing the non-parametric grid-based framework for inferring the BBHs DTD and merger rate. The framework begins with the construction of a delay time-probability grid, followed by the enumeration of all possible trajectories and the selection of causality-constrained ones. These trajectories are convolved with the cosmic star formation rate and cosmological model to obtain the merger rate evolution. Incorporating the BBH population model and detector selection function yields the detectable BBH population, which is then compared with the observed GW catalog. Bayesian evidence evaluation across trajectories enables inference of both the DTD shape and the local merger rate.

trajectories. The best-fit trajectory is defined as the one which has highest evidence. The framework provides a model-independent method for discovering complex evolutionary pathways that would be difficult to capture with traditional parametric approaches.

2.3. Inference of Local Merger Rate Density $R_0(m)$

Once the best-fitting DTD shape has been identified for each mass bin through the Bayesian framework described above, we proceed to infer the local merger rate density $R_0(m)$. For this inference, we note that the shape of $p_t(t_d)$ determines the redshift evolution of the merger rate, while $R_0(m)$ acts as a normalization parameter that scales the overall amplitude. With the best-fitting trajectory fixed, the inference of $R_0(m)$ becomes a single-parameter problem for each mass bin. We adopt a flat prior on ranging from 1 to 100 $\text{Gpc}^{-3} \text{ yr}^{-1}$. The likelihood function follows the same Poisson formulation, with the expected counts now explicitly dependent on $R_0(m)$. This yields both point estimates and uncertainties for the local BBH merger rate density across different mass bins.

The overall workflow of our framework is summarized in Figure 1. The flowchart illustrates the construction of delay time-probability grids, enforcement of causal trajectories, convolution with the star formation history, incorporation of cosmological and selection effects, and

the final Bayesian inference pipeline leading to the estimation of the local BBH merger rate density

3. DATA CHARACTERIZATION

3.1. Mass Binning and Redshift Distribution

We analyze BBH mergers from the GWTC-3 and GWTC-4 catalogs, the most comprehensive set of GW detections to date. For each event, we draw posterior samples of the source-frame component masses (m_1, m_2) and luminosity distance d_L from the released parameter-estimation data products. Adopting a spatially flat Λ CDM cosmology with parameters from PLANCK 2018 (Aghanim et al. 2020)¹, we convert the luminosity distance into redshift z . By combining the component-mass samples across all events, weighted by the number of posterior draws per event, we construct population-level summaries in the (m, z) plane that account for measurement uncertainty.

To investigate mass-dependent formation channels, we divide the component mass distribution into two bins: low-mass systems ($20\text{--}40 M_\odot$) and high-mass systems ($40\text{--}100 M_\odot$). This choice balances statistical power with the ability to probe distinct evolutionary pathways

¹ As the error on the luminosity distance is currently more than 10% on the individual sources, the change in the result due to the change in the Hubble constant from Planck-2018 to SH0ES (Riess et al. 2022), will not have any substantial impact on the result.

across different mass regimes. For each bin, we compute the normalized redshift distribution, which highlights the relative merger occurrence as a function of cosmic time independent of absolute event rates.

Figure 2 shows the normalized redshift distributions for the two mass bins. The low-mass population peaks at $z \sim 0.2$ and exhibits a relatively narrow distribution, consistent with more recent mergers produced by longer delay times. In contrast, high-mass systems peak at $z \sim 0.5$ and display a broader distribution extending to earlier cosmic epochs, consistent with shorter delay times. The distinct shapes of the two distributions suggest differences in the underlying time-delay functions that govern binary evolution and merger timescales. In future with more number of sources, finer mass-bins will be possible to explore.

The mass-dependent redshift distributions provide valuable qualitative constraints on delay time evolution, but disentangling selection effects from intrinsic population properties requires quantitative modeling. To distinguish between genuine DTDs and observational biases, we employ the LVK injection campaigns from Observing Runs O1-O4². These datasets contain simulated signals with known source parameters, analyzed through the same search pipelines as the real data, thereby providing an empirical estimate of detection efficiency across the mass-redshift plane. For consistency with the observed catalog, we adopt a conservative network signal-to-noise threshold of $\rho_{\text{thresh}} > 10$ when defining detections in both the injections and the real sample. The selection function, i.e. the probability of detecting a binary with component mass m at redshift z , is defined as

$$S(m, z) = \int P(\text{det} | m, z, \Omega) p(\Omega) d\Omega \approx \frac{N_{\text{det}}(m, z)}{N_{\text{inj}}(m, z)}, \quad (12)$$

where Ω denotes extrinsic parameters (sky location, orientation, etc.), and N_{inj} and N_{det} are the number of injected and recovered signals in each (m, z) bin (Abbott et al. 2023; Essick & Holz 2024; Talbot & Thrane 2018; Gerosa & Bellotti 2024; Essick et al. 2025). This selection function, together with the posterior samples of the observed events, is incorporated into our hierarchical Bayesian framework (Section 2). The results of this analysis are presented in Section 4.

² The injection sets and recovered selection functions for GWTC-3 and GWTC-4 are publicly available at <https://zenodo.org/record/5636816> and <https://zenodo.org/records/16740128> respectively.

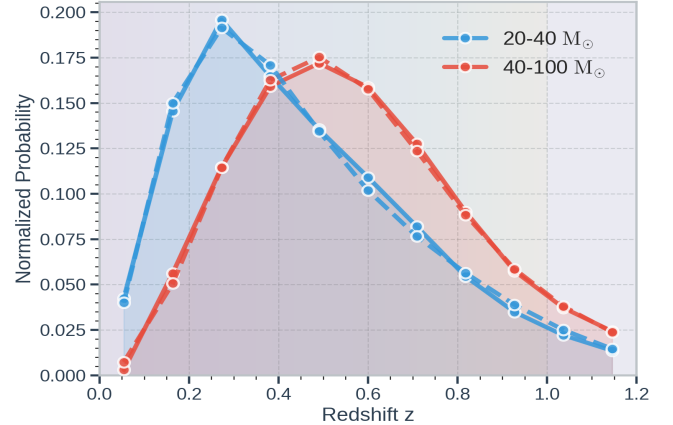


Figure 2. Normalized redshift distributions of BBHs from the combined GWTC-3 and GWTC-4 catalogs, binned by source-frame primary mass: $20\text{--}40M_{\odot}$ (blue) and $40\text{--}100M_{\odot}$ (red). Solid curves represent the observed distributions, while dashed curves show the reconstructed trajectories from the best-fit delay-time distribution model obtained using our grid-based technique (see Sec. 4 for more details on this.). The lower-mass systems peak at $z \sim 0.2$, reflecting more recent mergers with longer delay times, whereas higher-mass systems peak at $z \sim 0.5$ and exhibit broader distributions extending to earlier cosmic epochs, consistent with shorter delay times. The close agreement between the reconstructed trajectories and the observations supports the inferred form of the underlying time-delay function governing BBH formation channels.

3.2. Grid Configuration and Trajectory Space from data

We implement a 5×5 grid structure in $(t_d, \log p_t)$ space to systematically explore DTDs while maintaining computational tractability. The delay time axis spans from $t_{d,\text{min}} = 500$ Myr to $t_{d,\text{max}} = 13$ Gyr, encompassing the full range of astrophysically plausible delay times from prompt dynamical mergers in dense stellar environments to extended delays from wide binaries or metallicity-dependent formation suppression. A higher number of bins in the $(t_d, \log p_t)$ space increases the computational cost significantly without much gain in our findings, due to limited number sources. However, the technique can be scaled up to a larger number of bins with more number of detected GW sources in the future. The grid nodes are positioned to provide adequate resolution across the full dynamic range:

$$\{t_d^{(i)}\}_{i=1}^5 = \{0.5, 2, 5, 10, 13\} \text{ Gyr}, \quad (13)$$

$$\{\log_{10} p_t^{(j)}\}_{j=1}^5 = \{2, 1.5, 1.0, 0.5, 0\}. \quad (14)$$

The probability density range spans two orders of magnitude, sufficient to represent diverse DTD shapes from broad power-law forms ($p_t \sim t_d^{-1}$) to sharply

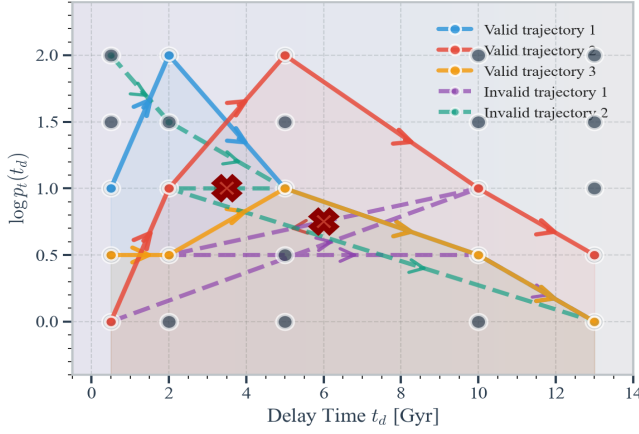


Figure 3. Grid-based exploration of DTDs in $(t_d, \log_{10} p_t)$ space. The 5×5 grid spans delay times from 0.5 to 13 Gyr and probability densities over six orders of magnitude. Solid lines show three example valid trajectories that satisfy the causality constraint $t_d^{(i)} \leq t_d^{(i+1)}$, while dashed lines illustrate invalid trajectories that violate causality by decreasing in delay time (marked with $\times$). Arrows indicate the direction of temporal evolution. Each trajectory through the grid defines a unique DTD via cubic spline interpolation between grid points, enabling systematic exploration of all physically allowed evolutionary pathways without assuming specific parametric forms.

peaked distributions characteristic of specific formation channels with the limited number of data. This logarithmic spacing ensures numerical stability in likelihood calculations while capturing both gradual evolutionary trends and rapid transitions between formation regimes.

The exploration of this grid space is constrained by fundamental causality requirements: any valid evolutionary trajectory must satisfy $t_d^{(i)} \leq t_d^{(i+1)}$, ensuring that delay times are non-decreasing along the temporal sequence. Figure 3 illustrates this constraint through example trajectories that demonstrate both valid evolutionary pathways and violations of the causality principle. Valid trajectories (solid lines) represent physically allowable delay time evolution, where the system can remain at the same delay time or progress to longer delays, but cannot evolve backward in time. The directional arrows clearly indicate the temporal evolution along each pathway. Invalid trajectories (dashed lines) violate the causality constraint by attempting to decrease delay times during evolution, marked with violation symbols ($\times$) at the problematic transitions. These violations represent nonphysical scenarios where a binary system would have multi-valued function at a fixed t_d .

4. RESULTS

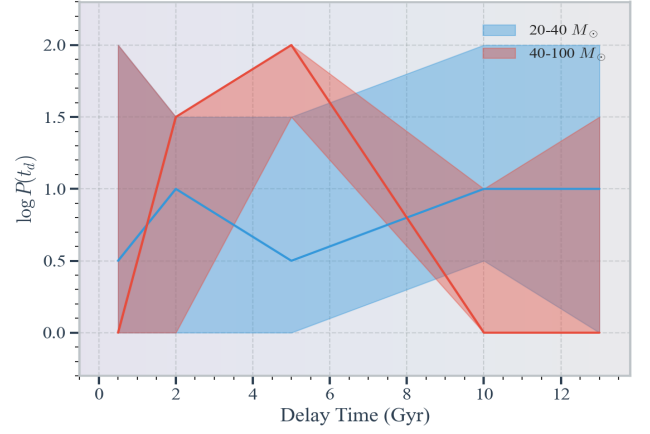


Figure 4. Envelope of the top 50 highest-evidence delay-time distribution trajectories reconstructed from the GWTC-3 and GWTC-4 catalogs. The shaded regions show the allowed range of $\log P(t_d)$ as a function of delay time for binaries in two mass intervals: 20–40 $M_\odot$ (pink) and 40–100 $M_\odot$ (gold), while the dark solid curves denote the single best-evidence trajectory in each mass bin. Lower-mass systems exhibit broadly flat distributions across the delay-time range, whereas higher-mass binaries show a pronounced concentration toward short delays ($\lesssim 5$ Gyr), providing direct non-power-law evidence for merger timescales.

We apply our grid-based DTD reconstruction framework described in Section 2 to the combined GWTC-3 and GWTC-4 catalogs of BBH mergers, incorporating the corresponding injection campaigns to correct for selection biases (Talbot & Thrane 2018; Gerosa & Bellotti 2024; Essick et al. 2025). The analysis is performed in the (m, z) plane with multiple grid resolutions to test bin-size systematics. For robustness, we restrict to systems with component masses above $20 M_\odot$, where detection efficiency is well characterized, and focus on two broad mass intervals: 20–40 $M_\odot$ and 40–100 $M_\odot$.

Figure 4 shows the DTDs for the two mass ranges. For lower-mass binaries (20–40 $M_\odot$), the inferred $P(t_d)$ is broad and nearly flat across ~ 1 –10 Gyr, indicating that both prompt and significantly delayed mergers occur with comparable likelihood. This shape suggests a diverse mixture of evolutionary pathways and environments, including both isolated binary evolution with long delays and dynamical channels that can accelerate coalescence.

In contrast, higher-mass binaries (40–100 $M_\odot$) exhibit peaked DTDs, with merger probabilities strongly concentrated at short to intermediate delays (~ 2 –6 Gyr) and a sharp suppression of long delays. The decay in the high t_d for the high mass bin is coming from the fact that the merger rate is increasing as one moves from lower redshifts to higher redshift (see figure 2 red solid line.).

Whereas at the lower delay time (for $t_d < 2$ Gyrs), the constraints are weak (exhibiting a large error) due to non-detection of high redshift sources impacted by the selection function. In figure 2 we show the normalized redshift distributions of BBH mergers, comparing the observed GW catalog data (solid line) with the reconstructed trajectories (dashed line) derived from the best-evidence DTD inferred using our grid-based technique. The remarkable agreement between the two provides a strong self-consistency check on our framework, demonstrating that the inferred DTD not only yield statistically favored trajectories in $(t_d, \log P(t_d))$ space but also naturally reproduce the observed cosmic evolution of the merger population.

This represents the first direct evidence towards a nonparametric DTD showing that massive BBH mergers do not follow a simple power-law delay distribution, but instead preferentially occur on specific evolutionary timescales. The peaked structure points toward more uniform and rapid evolutionary pathways for delay time between 2-6 Gyrs, potentially a typical scale in the initial spatial distribution of the parent stars for different formation channel. Taken together, these results reveal a mass-dependent evolutionary dichotomy: low-mass BBH mergers are distributed across the full range of delay time with nearly a scale invariant shape, while high-mass BBHs predominantly merge at some typical timescale. This nonparametric, selection-corrected reconstruction provides strong constraints on BBH formation channels, challenges the commonly assumed power-law delay models, and opens new opportunities for probing the astrophysical environments that drive black hole binary evolution. If this result strengthens with more observations in the future, understanding of this finding from astrophysical modeling for different formation channel will be crucial.

In addition to the qualitative differences in DTD, we also infer distinct local merger rates ($R_0(m)$) for the two mass ranges illustrated in Figure 5. For the low-mass bin ($20\text{--}40 M_\odot$), the estimated local merger rate is $21^{+7}_{-6.5} \text{ Gpc}^{-3} \text{ yr}^{-1}$, whereas for the high-mass bin ($40\text{--}100 M_\odot$) it is significantly lower, $9.5^{+10}_{-5} \text{ Gpc}^{-3} \text{ yr}^{-1}$. The suppression of the high-mass rate can be understood in terms of two complementary effects. First, the intrinsic number of high-mass systems is smaller compared to their lower-mass counterparts. Second, the reconstructed delay-time distribution for high-mass binaries favors systematically shorter delays. Since shorter delays shift mergers to earlier cosmic epochs where the star-formation rate is higher, the local merger rate is naturally reduced. Together, these effects explain why

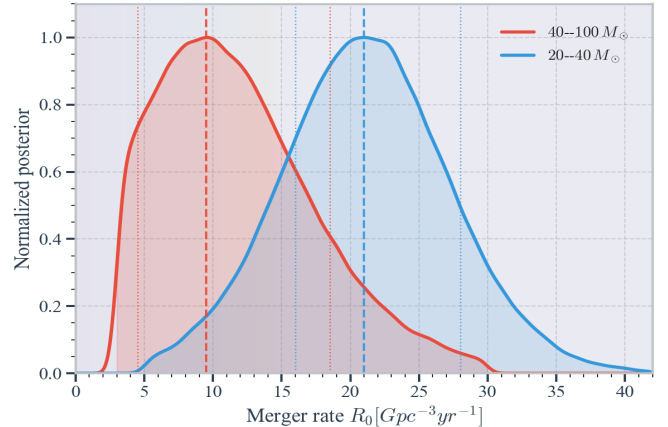


Figure 5. Posterior distributions of the local binary-black-hole merger rate R_0 for two mass bins, $20\text{--}40 M_\odot$ and $40\text{--}100 M_\odot$. Solid curves show the marginalized one-dimensional posteriors; vertical dashed lines mark the posterior medians and dotted lines indicate the 68% credible bounds.

the local rate of high-mass BBH mergers is suppressed relative to low-mass systems.

5. CONCLUSIONS

We have developed a novel grid-based, non-parametric framework for inferring mass-dependent DTDs of BBH mergers from GW observations. Unlike traditional parametric approaches, our method discretizes the delay time-probability space and systematically explores all causally valid evolutionary trajectories, enabling the reconstruction of complex DTD shapes without imposing restrictive functional forms.

Applied to the GWTC-3 and GWTC-4 catalogs, our analysis reveals clear mass-dependent trends. Lower-mass systems ($20\text{--}40 M_\odot$) are consistent with broad, nearly flat distributions across a wide range of evolutionary timescales, indicating diverse formation environments and channels. By contrast, higher-mass BBHs ($40\text{--}100 M_\odot$) exhibit sharply peaked DTDs at short delays ($\sim 2\text{--}6$ Gyr), providing the first direct non-parametric evidence that massive systems deviate from simple power-law behavior and preferentially undergo rapid post-formation mergers. In addition, we find distinct local merger rates: $21^{+7}_{-6.5} \text{ Gpc}^{-3} \text{ yr}^{-1}$ for the low-mass bin and $9.5^{+10}_{-5} \text{ Gpc}^{-3} \text{ yr}^{-1}$ for the high-mass bin, the latter suppressed due to both fewer systems and shorter delays shifting mergers to earlier epochs.

These findings align with theoretical expectations that more massive binaries may experience different formation scenarios than the lighter ones. At the same time, the observed broad distributions for lower-mass systems highlight the coexistence of both short and long delay time, with nearly a scale invariant spectrum, as ex-

pected from a flat in log-space initial separation of the binaries (Barrett et al. 2018; Safarzadeh & Berger 2019; Mennekens et al. 2010; Ruiter et al. 2011; Beniamini & Piran 2019). The detection of mass-dependent DTDs represents a key milestone in GW population science, demonstrating that current catalogs already contain discriminating power to probe the detailed physics of stellar evolution and compact-object formation. As GW astronomy enters an era of precision cosmology and astrophysics, non-parametric, evidence-based frameworks such as the one presented here will be essential for extracting the full astrophysical information encoded in the rapidly growing population of detected mergers. In future with more observation, a non-parametric reconstruction of the DTD using the grid-based technique with more number of bins and for finer mass range can shed line in the long standing question on the formation channel of BBHs. Moreover, with observation of more binary neutron stars and neutron star black hole systems, the non-parametric reconstruction of the DTD will be possible for these sources as well from GW observations.

6. DATA AVAILABILITY

The gravitational-wave catalogs used in this study are publicly available on Zenodo as part of the LIGO-Virgo-KAGRA data releases: GWTC-4 (link), GWTC-2.1 (link), and GWTC-3 (link).

ACKNOWLEDGMENTS

The authors express their gratitude to Sergio Andrés Vallejo-Peña for reviewing the manuscript and providing useful comments as a part of the LIGO publication policy. This work is part of the `<data|theory>`

Universe Lab, supported by TIFR and the Department of Atomic Energy, Government of India. The authors express gratitude to the system administrator of the computer cluster of `<data|theory>` Universe Lab. This research is supported by the Prime Minister Early Career Research Award, Anusandhan National Research Foundation, Government of India. Special thanks to the LIGO Virgo KAGRA Scientific Collaboration for providing noise curves. LIGO, funded by the U.S. National Science Foundation (NSF), and Virgo, supported by the French CNRS, Italian INFN, and Dutch Nikhef, along with contributions from Polish and Hungarian institutes. This collaborative effort is backed by the NSF’s LIGO Laboratory, a major facility fully funded by the National Science Foundation. The research leverages data and software from the Gravitational Wave Open Science Center, a service provided by LIGO Laboratory, the LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA. Advanced LIGO’s construction and operation receive support from STFC of the UK, Max Planck Society (MPS), and the State of Niedersachsen/Germany, with additional backing from the Australian Research Council. Virgo, affiliated with the European Gravitational Observatory (EGO), secures funding through contributions from various European institutions. Meanwhile, KAGRA’s construction and operation are funded by MEXT, JSPS, NRF, MSIT, AS, and MoST. This material is based upon work supported by NSF’s LIGO Laboratory which is a major facility fully funded by the National Science Foundation. We acknowledge the use of the following packages in this work: Numpy (Van Der Walt et al. 2011), Scipy (Jones et al. 2001–), Matplotlib (Hunter 2007), Astropy (Robitaille et al. 2013), Dynesty (Speagle 2020) and emcee (Foreman-Mackey et al. 2013).

REFERENCES

- Abbott, B. P., et al. 2016a, *Phys. Rev. Lett.*, 116, 061102, doi: [10.1103/PhysRevLett.116.061102](https://doi.org/10.1103/PhysRevLett.116.061102)
- . 2016b, *Phys. Rev. X*, 6, 041015, doi: [10.1103/PhysRevX.6.041015](https://doi.org/10.1103/PhysRevX.6.041015)
- . 2019, *Phys. Rev. X*, 9, 031040, doi: [10.1103/PhysRevX.9.031040](https://doi.org/10.1103/PhysRevX.9.031040)
- Abbott, R., et al. 2021, *Astrophys. J. Lett.*, 913, L7, doi: [10.3847/2041-8213/abe949](https://doi.org/10.3847/2041-8213/abe949)
- . 2023, *Phys. Rev. X*, 13, 011048, doi: [10.1103/PhysRevX.13.011048](https://doi.org/10.1103/PhysRevX.13.011048)
- Acernese, F., et al. 2015, *Class. Quant. Grav.*, 32, 024001, doi: [10.1088/0264-9381/32/2/024001](https://doi.org/10.1088/0264-9381/32/2/024001)
- Afroz, S., & Mukherjee, S. 2025a, *Phys. Rev. D*, 112, 023531, doi: [10.1103/7zc2-g9vq](https://doi.org/10.1103/7zc2-g9vq)
- . 2025b. <https://arxiv.org/abs/2505.22739>
- . 2025c. <https://arxiv.org/abs/2509.09123>
- . 2025d. <https://arxiv.org/abs/2509.25356>
- Aghanim, N., et al. 2020, *Astron. Astrophys.*, 641, A6, doi: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910)
- Akutsu, T., et al. 2021, *PTEP*, 2021, 05A101, doi: [10.1093/ptep/ptaa125](https://doi.org/10.1093/ptep/ptaa125)
- Antonelli, A., Kritos, K., Ng, K. K. Y., Cotesta, R., & Berti, E. 2023, *Phys. Rev. D*, 108, 084044, doi: [10.1103/PhysRevD.108.084044](https://doi.org/10.1103/PhysRevD.108.084044)
- Arimoto, M., et al. 2021, arXiv preprint. <https://arxiv.org/abs/2104.02445>
- Bailes, M., et al. 2021, *Nature Rev. Phys.*, 3, 344, doi: [10.1038/s42254-021-00303-8](https://doi.org/10.1038/s42254-021-00303-8)

- Bailyn, C. D., Jain, R. K., Coppi, P., & Orosz, J. A. 1998, *Astrophys. J.*, 499, 367, doi: [10.1086/305614](https://doi.org/10.1086/305614)
- Barrett, J. W., Gaebel, S. M., Neijssel, C. J., et al. 2018, *Mon. Not. Roy. Astron. Soc.*, 477, 4685, doi: [10.1093/mnras/sty908](https://doi.org/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](https://doi.org/10.1088/0004-637X/714/2/1217)
- Belczynski, K., et al. 2016, *Astron. Astrophys.*, 594, A97, doi: [10.1051/0004-6361/201628980](https://doi.org/10.1051/0004-6361/201628980)
- Beniamini, P., & Piran, T. 2019, *Mon. Not. Roy. Astron. Soc.*, 487, 4847, doi: [10.1093/mnras/stz1589](https://doi.org/10.1093/mnras/stz1589)
- Bouffanais, Y., Mapelli, M., Santoliquido, F., et al. 2021, *Mon. Not. Roy. Astron. Soc.*, 507, 5224, doi: [10.1093/mnras/stab2438](https://doi.org/10.1093/mnras/stab2438)
- Callister, T. A., & Farr, W. M. 2024, *Phys. Rev. X*, 14, 021005, doi: [10.1103/PhysRevX.14.021005](https://doi.org/10.1103/PhysRevX.14.021005)
- Cheng, A. Q., Zevin, M., & Vitale, S. 2023, *Astrophys. J.*, 955, 127, doi: [10.3847/1538-4357/aced98](https://doi.org/10.3847/1538-4357/aced98)
- Dominik, M., Belczynski, K., Fryer, C., et al. 2012, *Astrophys. J.*, 759, 52, doi: [10.1088/0004-637X/759/1/52](https://doi.org/10.1088/0004-637X/759/1/52)
- . 2013, *Astrophys. J.*, 779, 72, doi: [10.1088/0004-637X/779/1/72](https://doi.org/10.1088/0004-637X/779/1/72)
- Dominik, M., Berti, E., O’Shaughnessy, R., et al. 2015, *Astrophys. J.*, 806, 263, doi: [10.1088/0004-637X/806/2/263](https://doi.org/10.1088/0004-637X/806/2/263)
- Edelman, B., Farr, B., & Doctor, Z. 2023, *Astrophys. J.*, 946, 16, doi: [10.3847/1538-4357/acb5ed](https://doi.org/10.3847/1538-4357/acb5ed)
- Essick, R., & Holz, D. E. 2024, *Phys. Rev. D*, 110, 103018, doi: [10.1103/PhysRevD.110.103018](https://doi.org/10.1103/PhysRevD.110.103018)
- Essick, R., et al. 2025. <https://arxiv.org/abs/2508.10638>
- Fishbach, M., Holz, D. E., & Farr, W. M. 2018, *Astrophys. J. Lett.*, 863, L41, doi: [10.3847/2041-8213/aad800](https://doi.org/10.3847/2041-8213/aad800)
- Fishbach, M., & Kalogera, V. 2021, *Astrophys. J. Lett.*, 914, L30, doi: [10.3847/2041-8213/ac05c4](https://doi.org/10.3847/2041-8213/ac05c4)
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *Publications of the Astronomical Society of the Pacific*, 125, 306
- Franciolini, G., Baibhav, V., De Luca, V., et al. 2022, *Phys. Rev. D*, 105, 083526, doi: [10.1103/PhysRevD.105.083526](https://doi.org/10.1103/PhysRevD.105.083526)
- Gennari, V., Mastrogiovanni, S., Tamanini, N., Marsat, S., & Pierra, G. 2025, *Phys. Rev. D*, 111, 123046, doi: [10.1103/ftw9-7xd5](https://doi.org/10.1103/ftw9-7xd5)
- Gerosa, D., & Bellotti, M. 2024, *Class. Quant. Grav.*, 41, 125002, doi: [10.1088/1361-6382/ad4509](https://doi.org/10.1088/1361-6382/ad4509)
- Heinzel, J., Biscoveanu, S., & Vitale, S. 2024, *Phys. Rev. D*, 109, 103006, doi: [10.1103/PhysRevD.109.103006](https://doi.org/10.1103/PhysRevD.109.103006)
- Hunter, J. D. 2007, *Computing in science & engineering*, 9, 90
- Iorio, G., et al. 2022, *mnras*, doi: [10.1093/mnras/stad1630](https://doi.org/10.1093/mnras/stad1630)
- Jones, E., Oliphant, T., Peterson, P., et al. 2001–, *SciPy: Open source scientific tools for Python*. <http://www.scipy.org/>
- Karathanasis, C., Mukherjee, S., & Mastrogiovanni, S. 2023a, *Mon. Not. Roy. Astron. Soc.*, 523, 4539, doi: [10.1093/mnras/stad1373](https://doi.org/10.1093/mnras/stad1373)
- Karathanasis, C., Revenu, B., Mukherjee, S., & Stachurski, F. 2023b, *Astron. Astrophys.*, 677, A124, doi: [10.1051/0004-6361/202245216](https://doi.org/10.1051/0004-6361/202245216)
- Lalleman, M., Turbang, K., Callister, T., & van Remortel, N. 2025, *Astron. Astrophys.*, 698, A85, doi: [10.1051/0004-6361/202553941](https://doi.org/10.1051/0004-6361/202553941)
- Lamberts, A. 2018, in *IAU Symposium*, Vol. 338, *Gravitational Wave Astrophysics: Early Results from Gravitational Wave Searches and Electromagnetic Counterparts*, ed. G. González & R. Hynes, 40–45, doi: [10.1017/S1743921318000182](https://doi.org/10.1017/S1743921318000182)
- LIGO-VIRGO-KAGRA Collaboration. 2025a. <https://arxiv.org/abs/2508.18083>
- . 2025b. <https://arxiv.org/abs/2508.18082>
- Madau, P., & Dickinson, M. 2014, *Ann. Rev. Astron. Astrophys.*, 52, 415, doi: [10.1146/annurev-astro-081811-125615](https://doi.org/10.1146/annurev-astro-081811-125615)
- Mandel, I., & Farmer, A. 2022, *Phys. Rept.*, 955, 1, doi: [10.1016/j.physrep.2022.01.003](https://doi.org/10.1016/j.physrep.2022.01.003)
- Mapelli, M. 2016, *Mon. Not. Roy. Astron. Soc.*, 459, 3432, doi: [10.1093/mnras/stw869](https://doi.org/10.1093/mnras/stw869)
- . 2021, *Formation Channels of Single and Binary Stellar-Mass Black Holes*, doi: [10.1007/978-981-15-4702-7_16-1](https://doi.org/10.1007/978-981-15-4702-7_16-1)
- Mennekens, N., Vanbeveren, D., De Greve, J. P., & De Donder, E. 2010, *A&A*, 515, A89, doi: [10.1051/0004-6361/201014115](https://doi.org/10.1051/0004-6361/201014115)
- Mukherjee, S. 2022, *Mon. Not. Roy. Astron. Soc.*, 515, 5495, doi: [10.1093/mnras/stac2152](https://doi.org/10.1093/mnras/stac2152)
- Mukherjee, S., Broadhurst, T., Diego, J. M., Silk, J., & Smoot, G. F. 2021, *Mon. Not. Roy. Astron. Soc.*, 506, 3751, doi: [10.1093/mnras/stab1980](https://doi.org/10.1093/mnras/stab1980)
- 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)
- Riess, A. G., et al. 2022, *Astrophys. J. Lett.*, 934, L7, doi: [10.3847/2041-8213/ac5c5b](https://doi.org/10.3847/2041-8213/ac5c5b)
- Rinaldi, S., Del Pozzo, W., Mapelli, M., Lorenzo-Medina, A., & Dent, T. 2024, *Astron. Astrophys.*, 684, A204, doi: [10.1051/0004-6361/202348161](https://doi.org/10.1051/0004-6361/202348161)
- Rinaldi, S., Liang, Y., Demasi, G., Mapelli, M., & Del Pozzo, W. 2025. <https://arxiv.org/abs/2506.05929>
- Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, *Astronomy & Astrophysics*, 558, A33

- Rodriguez, C. L., Haster, C.-J., Chatterjee, S., Kalogera, V., & Rasio, F. A. 2016, *Astrophys. J. Lett.*, 824, L8, doi: [10.3847/2041-8205/824/1/L8](https://doi.org/10.3847/2041-8205/824/1/L8)
- Ruiter, A. J., Belczynski, K., Sim, S. A., et al. 2011, *MNRAS*, 417, 408, doi: [10.1111/j.1365-2966.2011.19276.x](https://doi.org/10.1111/j.1365-2966.2011.19276.x)
- Safarzadeh, M., & Berger, E. 2019, *Astrophys. J. Lett.*, 878, L12, doi: [10.3847/2041-8213/ab24df](https://doi.org/10.3847/2041-8213/ab24df)
- Safarzadeh, M., Biscoveanu, S., & Loeb, A. 2020, *Astrophys. J.*, 901, 137, doi: [10.3847/1538-4357/abb1af](https://doi.org/10.3847/1538-4357/abb1af)
- Schiebelbein-Zwack, A., & Fishbach, M. 2024, *Astrophys. J.*, 970, 128, doi: [10.3847/1538-4357/ad5353](https://doi.org/10.3847/1538-4357/ad5353)
- Sicilia, A., Lapi, A., Boco, L., et al. 2022, *Astrophys. J.*, 924, 56, doi: [10.3847/1538-4357/ac34fb](https://doi.org/10.3847/1538-4357/ac34fb)
- Smith, T. B., & Kaplinghat, M. 2024, <https://arxiv.org/abs/2412.13318>
- Speagle, J. S. 2020, *Monthly Notices of the Royal Astronomical Society*, 493, 3132
- Stevenson, S., & Clarke, T. 2022, *mnras*, doi: [10.1093/mnras/stac2936](https://doi.org/10.1093/mnras/stac2936)
- Talbot, C., & Thrane, E. 2018, *Astrophys. J.*, 856, 173, doi: [10.3847/1538-4357/aab34c](https://doi.org/10.3847/1538-4357/aab34c)
- Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *Computing in science & engineering*, 13, 22
- Ye, C. S., & Fishbach, M. 2024, *Astrophys. J.*, 967, 62, doi: [10.3847/1538-4357/ad3ba8](https://doi.org/10.3847/1538-4357/ad3ba8)