Title

November 7, 2025

Abstract

1 Introduction

{sec:intro}

The current generation of ground-based gravitational wave (GW) detectors, the LIGO/Virgo/KAGRA (LVK) network, constrains the redshift distribution of binary black hole (BBH) mergers up to redshifts $z \sim 1.5$ (The LIGO Scientific Collaboration et al., 2025). Thanks to their unprecedented sensitivity, next generation detectors (Cosmic Explorer, Einstein Telescope) will be able to detect BBH mergers beyond $z \sim 20$ (Abac et al., 2025; Evans et al., 2021), unveiling a whole new segment of the BBH population.

One of the main goals of GW astrophysics is to uncover the properties of merging black holes (BH) and neutron stars (NS), shedding light on the origin of the observed population. With next generation detectors we will be able to map the BBH merger rate beyond the current limits, accessing the unexplored epoch of the formation of the first stars. Not only the overall mass distribution, but we will reconstruct the distribution as a function of redshift, allowing us to investigate how BHs progenitors form and evolve (Abac et al., 2025). But **how well do we need to measure BBH mergers to infer the population properties at a given *target redshift (z_t)* with sufficient precision to learn conclusively if and how they differ from what we infer at redshift ~ 0?**

Compact binary coalescences (CBC) detectable by ground-based detectors originate from the evolution of massive stars ($\gtrsim 8\ M_\odot$). Due to the short evolutionary timescales of such stars, the formation rate of stellar BHs follows the overall cosmic star formation history (SFH). However, the time delay between the formation of a black hole binary (BHB) and a merger spans a wide range of values, from few Myrs to the Hubble time (Boesky et al., 2024). Those BHs originate from different formation channels and across cosmic history. The efficiency of different formation channels is environment dependent, especially on metallicity (van Son et al., 2025). Population synthesis models reveal a correlation between the formation channel, the final BH mass, and the delay time distribution (DTD) (van Son et al., 2022). This prediceted correlation unlocks the scientific potential of mapping the BBHs population in the exotic environments of the high redshift Universe.

The population observed in the local universe, however, is composed by a mixture of BBH coming from both the short-delay and the long-delay channels. Therefore it contains systems coming from very different environments across the cosmic history. Because of that, selecting sources at z_t is the most direct way to ensure that the observed double compact objects (DCO) originate from stars formed at redshifts $> z_t$. By making appropriate choices for z_t we can select sources that formed in environments significantly different from the local one, for example in terms of metallicity. By comparing the BBH population at z_t with that at $z \sim 0$, we can begin to disentangle the cosmic history and the relative contribution of different stellar evolutionary pathways.

We developed a toy model to give precise predictions of what we need to measure in order to infer a certain level of variation of the CBC population across different redshifts. The final purpose of the model is understanding wheter and to what level a network of third generation detectors (Einstein Telescope and Cosmic Explorer) will allow us to make conclusive statements on the evolution and origins of DCO. More specifically, *we explore the measurement requirements to infer a variation of*

the merger rate density between the target redshift (z_t) and a reference redshift z_{ref} in a fixed mass bin. We explain the reasons behind our choices of the target redshifts and the mass bins in sections 1.2 and 1.3.

1.1 Merger rate density

The key feature of our toy model is the possibility to infer the confidence interval of the merger rate density ratio between the BBH mergers subpopulation in two different mass × redshift bins (MZbin), given a set of detected events falling inside such bins. This requires to connect the actual number of observations N to the intrinsic merger rate density $\mathcal{R}$ of the subpopulation. That is because there are selection effects (Gerosa et al., 2020; Gerosa & Bellotti, 2024; Chen et al., 2017) that mask the true population of mergers in the universe. Among these effects are: luminosity distance, chirp mass, binary spins, source orinetation, sky location, detector sensitivity.

The combination of all the selection effects determines the GW detectability—or detection probability, $p_{\rm det}$. The detectability is then directly related to another key quantity, the so-called effective spacetime volume $\langle VT\rangle$ (Gerosa et al., 2020; Abbott et al., 2016; Kapadia et al., 2020). $\langle VT\rangle$ connects the number of detections to the intrinsic merger rate density by combining $p_{\rm det}$, the observed comoving volume, and the time transformation between the detector frame and the source rest-frame (more details in **??**). Finally, these quantities are related by:

{eq:R-N_VT}
$$\mathcal{R} = \frac{N}{\langle VT\rangle}. \tag{1}$$

Equation (1) plays a crucial role when we our task is the comparison of two different MZbins. For instance, consider two MZbins, both centered at 40 $msun$ but at different redshifts, let's say $z \sim 1$ and $z \sim 5$. The detectability would be higher for the low redshift bin, but, depending on the width of the bins in redshift, also the associated comoving volume spherical shells will differ by some factor (of order unity in this case). These informations are encoded into the $\langle VT\rangle$ value for each bin (in general, for an MZbin, an higher $\langle VT\rangle$ reflects a higher sensitivity of the detector for events in that bin). The informations contained into $\langle VT\rangle$ are then propagated through eq. (1) when we take the ratio of the merger rate denisities between the two bins. Let's (irrealistically) say that in the same observation time we detect the same number of events in both bins; then the merger rate density ratio we could infer would not be 1, but the ratio of the $\langle VT\rangle$s of the two bins. This concept is expanded in our toy model to also take into account the poissonian uncertainty of the measurment in each bin. The details of the model are reported in **??**.

1.2 Target redshift

{sec:z_t_intro

As we already mentioned, we want to define measurment requirements of BBH mergers for a significantly different astrophysical population compared to the one we observe locally (at $z \sim 0$). However, the observable CBC population contains systems formed through the entire cosmic history. That is a consequence of the delay time distribution (DTD) between double compact objects (DCO) formation and merger (Mennekens & Vanbeveren, 2016; Abac et al., 2025). In that sense, selecting sources at a given target redshift (z_t) is the most straightforward way to ensure that the observed DCO mergers originated from stellar progenitors that formed and evolved at $z_{\rm form} > z_t$. But how do we choose z_t so that the selected progenitor population is significantly different from the local one?

Only second to mass, metallicity is fundamental for stellar evolution and DCO formation. The most significant and direct effect of metallicity on massive stars is its control over mass loss rates via stellar winds (Chowdhury & Santra, 2024). Stars in high metallicity environments can lose a substantial fraction of their initial mass before they undergo core collapse. Conversly, stars in low metallicity environments, typical of the early Universe, experience much weaker winds and therefore retain a significantly larger fraction of their mass throughout their evolution (Vink et al., 2001). This directly translates to the masses of the resulting BHs. The dependence of the maximum BH mass on metallicity is a robust prediction of stellar evolution models, with more massive remnants being sistematically produced in more metal poor environments (Belczynski et al., 2010; Spera et al., 2015). Metallicity also influences the opacity and the nuclear burning processes inside the star, which in turns affect its radius at varius stages of its evolution, especially

during post-main sequence stages (Xin et al., 2022; Romagnolo et al., 2023).

The complex interplay of metallicity-dependent stellar winds and radii propagate through the evolutionary processes that form merging BBHs. Different formation channels are characterized by distinct initial conditions, but the overall trend is a significant increase in formation efficiency at lower metallicities (Chruslinska et al., 2018; van Son et al., 2025; Klencki et al., 2018; Giacobbo & Mapelli, 2018; Neijssel et al., 2019).

The choice of the target redshift has to reflect our need of disentangling the environmental effects and the BBHs properties across the cosmic history. Based on that, we operatively define z_t as the redshift at which the majority of stars ($\gtrsim 90\%$) formed with a metallicity Z significantly lower than the solar one ($\lesssim 0.1 Z_\odot$). Despite our toy model does not make any assumption about how the CBC properties evolves, the choice of z_t inevitably needs to rely on how the SFH depends on redshift and metallicity. The metallicity-dependent SFH can be obtained combining the distribution of galaxy stellar masses, with the distributions describing the star formation rates and metallicities of galaxies at fixed stellar mass. A detailed discussion can be found in Chruslinska & Nelemans (2019); Chruślińska et al. (2021). Nevertheless, the final outcome is affected by uncertainties related to the assumed empirical scaling relations. That is, the choice of z_t is model dependent. We show in **??** that in the optimistic case the target redshift can be set at $z_t \sim 4$, while in the pessimistic case it can be as high as $z_t \sim 8$.

1.3 Mass distribution

{sec:m_dist}

Stellar evolution, binary interaction and metallicity, all leave an imprint on the BBH merger mass distribution. The mass spectrum observed by the LIGO/Virgo/KAGRA (LVK) collaboration exhibits several structures that are linked to the underlying formation physics (The LIGO Scientific Collaboration et al., 2025). The origin of these features provides the astrophysical context for selecting specific mass bins to probe a population evolution.

The most robustly detected feature is a peak in the primary mass distribution at ~ 10 M$_\odot$. This low-mass peak is thought to be predominantly produced by the evolution of isolated massive binaries (Giacobbo & Mapelli, 2018; Neijssel et al., 2019; Wiktorowicz et al., 2019). In addition to this, a second, broader, significant feature is present at ~ 35 M$_\odot$. At higher masses the merger rate appears to decline more steeply. The origin of this structure is object of debate, being linked to the evolution of the most massive stellar progenitors.

Broadly speaking we can distinguish two main classes of formation channels, the isolated channel and the dynamical channel. Within the isolated scenario we can further identify three main subchannels: the common envelope (CE) channel, the stable mass transfer (MT) channel, and chemically homogeneous evolution (CHE) (Chruślińska, 2022; Abac et al., 2025; Neijssel et al., 2019; Li et al., 2025). Population synthesis models suggest that the CE and stable MT channels are the major contributors to the low-mass population (Neijssel et al., 2019; Li et al., 2025). The high-mass peak could originate from the "pile-up" of masses caused by pulsation pair-instability supernovae (PPISNe) (Woosley, 2017), but this explanation is in tension with most recent stellar evolution simulations (Stevenson et al., 2019; Woosley & Heger, 2021; Hendriks et al., 2023; Tong et al., 2025). CHE might be the major conntributor to the $\sim 35\ M_\odot$ mass spectrum feature (Riley et al., 2021; Li et al., 2025). At even higher masses (65 M$_\odot \lesssim M_{\rm He} \lesssim 160$ M$_\odot$), stellar evolution models predict a dearth of BHs, known as the pair instability supernovae (PISN) mass gap (Woosley & Heger, 2021; Tong et al., 2025). The metallicity-dependent efficiency of these processes must be carefully reminded to perform physically informative subpopulations comparisons across different redshifts (metallicities).

The mass ratio distribution of BBHs mergers should also be considered in a population analysis. While we are not discussing the astrophysical origin of the mass ratio distribution, we make use of the most recent results from The LIGO Scientific Collaboration et al. (2025) to select reasonable values for our work. Specifically, it seems that BHs with masses ~ 10 M$_\odot$ preferentially merge with lighter BHs, pointing to a value of the mass ratio around $q \sim 0.8$. BHs with masses ~ 35 M$_\odot$, on the other hand, may preferentially merge with other BHs with more equal masses, reason why we prefer a value of the mass ratio $q \sim 1$.

maybe elaborate a bit more? We propose an analysis of the "low" ($\sim$ 10 $M_\odot$) and "high" ($\sim$ 35 $M_\odot$) mass bins (using fixed mass ratios) by means of our merger rate density toy model. Can we infer a population variation from future measurements? At what level? We present our results in **??**.

2 Merger rate density toy model

Our framework does not rely on any assumption on the underlying true astrophysical population of DCOs. Our goal is infering the merger rate denisity of BBHs in a given MZbin given: 1. A detector network; 2. A set of measurments from the network in the chosen bin. At surface level the main operation performed is set by eq. (1), where $\langle VT \rangle$ is defined by point 1, and N by point 2. To this we need to integrate the poisson uncertainty associated with a counting process. Moreover, we are interested in the ratio between merger rates between different bins. In the next sections we explain how we treat the number of detected events N (and its uncertainty), the spacetime volume sensitivity $\langle VT \rangle$, and crucially how their information content propagates to the merger rate densities ratio. The final result consists in a straightforward way to retrieve the confidence interval of the merger rate densities ratio between two MZbin. From now on we will refer to the two bins as the "target" and the "reference" bins, and our main goal will be estimating the value of $\mathcal{R}_t/\mathcal{R}_{\rm ref}$ from observations.

2.1 Poisson uncertainty

Given the number of detections inside the target bin (N_t) and the reference bin ($N_{\rm ref}$), can we conclude that $N_t/N_{\rm ref} \neq N_0$? The answer cleary depends on the values of N_t and $N_{\rm ref}$ through the fact that the number of detections inside each bin is poisson distributed. We developed two different approaches to answer the question: one is powered by a likelihood-ratio test (LRT), and the other one by bayesian inference methods. In **??** we show that the two approaches give similar results (in some specific regime? in general?) and we prefer the latter for its semplicity.

2.1.1 Likelihood-ratio test

The LRT is a hypothesis test that involves comparing the likelihoods of two nested models: a full model, and a reduced model that is a subset of the full model where we impose some restrictions. The restricted model makes the null hypothesis $\mathcal{H}_0$. The testing procedure involves obtaining the maximum likelihood estimators (MLE) for the parameters under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively $\hat{\theta}_0$ and $\hat{\theta}$, and computing the test statistic $\lambda_{\rm LR} = -2\Big[\ell(\theta_0) - \ell(\hat{\theta})\Big]$ (Koch, 1988), where ℓ denotes the log-likelihood. By Wilks' theorem (Wilks, 1938), $\lambda_{\rm LR}$ converges asymptotically to being χ^2-distributed under $\mathcal{H}_0$.

We define our likelihood to be the product of two poissonians with expectation values λ_t and $\lambda_{\rm ref}$: $\mathcal{L}(\lambda_t, \lambda_{\rm ref}) = {\rm Pois}(\lambda_t) \times {\rm Pois}(\lambda_{\rm ref})$. We reparametrize $\mathcal{L}$ as a function of $\lambda_{\rm ref}$ and $\alpha = \lambda_t/\lambda_{\rm ref}$ and we compute the test statistic for the null hypothesis $\alpha = \alpha_0$. It turns out that:

$$\lambda_{\rm LR} = 2N_{\rm ref}\left[\alpha \ln\left(\frac{\alpha}{\alpha_0}\right) - (1+\alpha)\ln\left(\frac{1+\alpha}{1+\alpha_0}\right)\right]. \quad (2) \quad \{\texttt{eq:R-N\_VT}\}$$

Mathematical details can be found in section A.

Our likelihood has two parameters, one of which is fixed. Therefore $\lambda_{\rm LR}$ is asymptotically χ^2-distributed with 1 degree of freedom. The 95% confidence interval for $\alpha = \lambda_t/\lambda_{\rm ref}$ can be found with a root finding algorithm for a fixed choice of α_0 and $N_{\rm ref}$.

3 Results {sec:results}

3.1 How likely can we detect BHB?

Figure 1 shows the detecatibility of BBH mergers in the primary mass-redshift space for equal mass ratio BHBs. Thanks to the high sensitivity of third generation detectors, most BBH mergers happening at $z \lesssim 1$ should be virtually detectable, with the exception of the lowest mass ratio binaries.

A Likelihood-ratio test {apx:lrt}

References

Abac A., et al., 2025, arXiv e-prints, p. arXiv:2503.12263

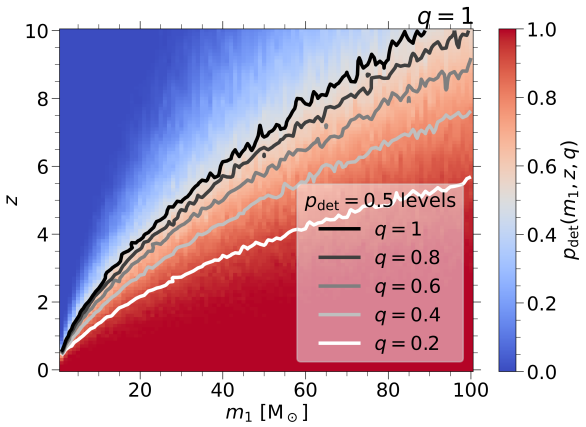

Figure 1: Detection probability colormap as a function of the primary BH mass and redhsift, for an equal mass ratio BHB. Colored lines mark the 0.5 probability levels (white band in the colormap for $q = 1$) for different choices of the mass ratio. $p_{\rm det}$ is computed for an ET + CE network assuming an isotropic distribution of sources and averaging over spins distribution.

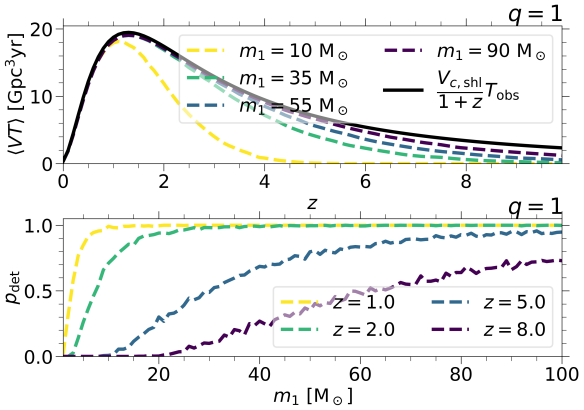

Figure 2: *Top*: $\langle VT \rangle$ evolution with redshift for different choices of the mass bins for equal mass ratio BHBs. Position and value of the peak depend on the choice of the bin widths, both in the mass and redshift dimensions. The solid line represents the (redshift integrated) cosmology-dependent factor in the expression for $\langle VT \rangle$. It is the comoving volume spherical shell covered by a redshift bin divided by $1 + z$ times the observation time of the detector ($T_{\rm obs} = 1$ yr in the plot). The faster $\langle VT \rangle$ decline for lower masses after the peak is determined by lower detection probabilitiy at high redshifts (see fig. 1). *Bottom*: detection probability dependence with primary mass for equal mass ratio BHBs for different fixed redhsifts. {fig:pdet_and_

Abbott B. P., et al., 2016, The Astrophysical Journal Letters, 833, L1

Belczynski K., Bulik T., Fryer C. L., Ruiter A., Valsecchi F., Vink J. S., Hurley J. R., 2010, The Astrophysical Journal, 714, 1217

Boesky A. P., Broekgaarden F. S., Berger E., 2024, The Astrophysical Journal, 976, 24

Chen H.-Y., Essick R., Vitale S., Holz D. E., Katsavounidis E., 2017, The Astrophysical Journal, 835, 31

Chowdhury S. R., Santra D., 2024, A population study on the effect of metallicity on ZAMS to the merger, doi:10.48550/ARXIV.2411.11902

Chruslinska M., Nelemans G., 2019, Monthly Notices of the Royal Astronomical Society, 488, 5300

Chruslinska M., Nelemans G., Belczynski K., 2018, Monthly Notices of the Royal Astronomical Society, 482, 5012

Chruślińska M., 2022, Annalen der Physik, 536

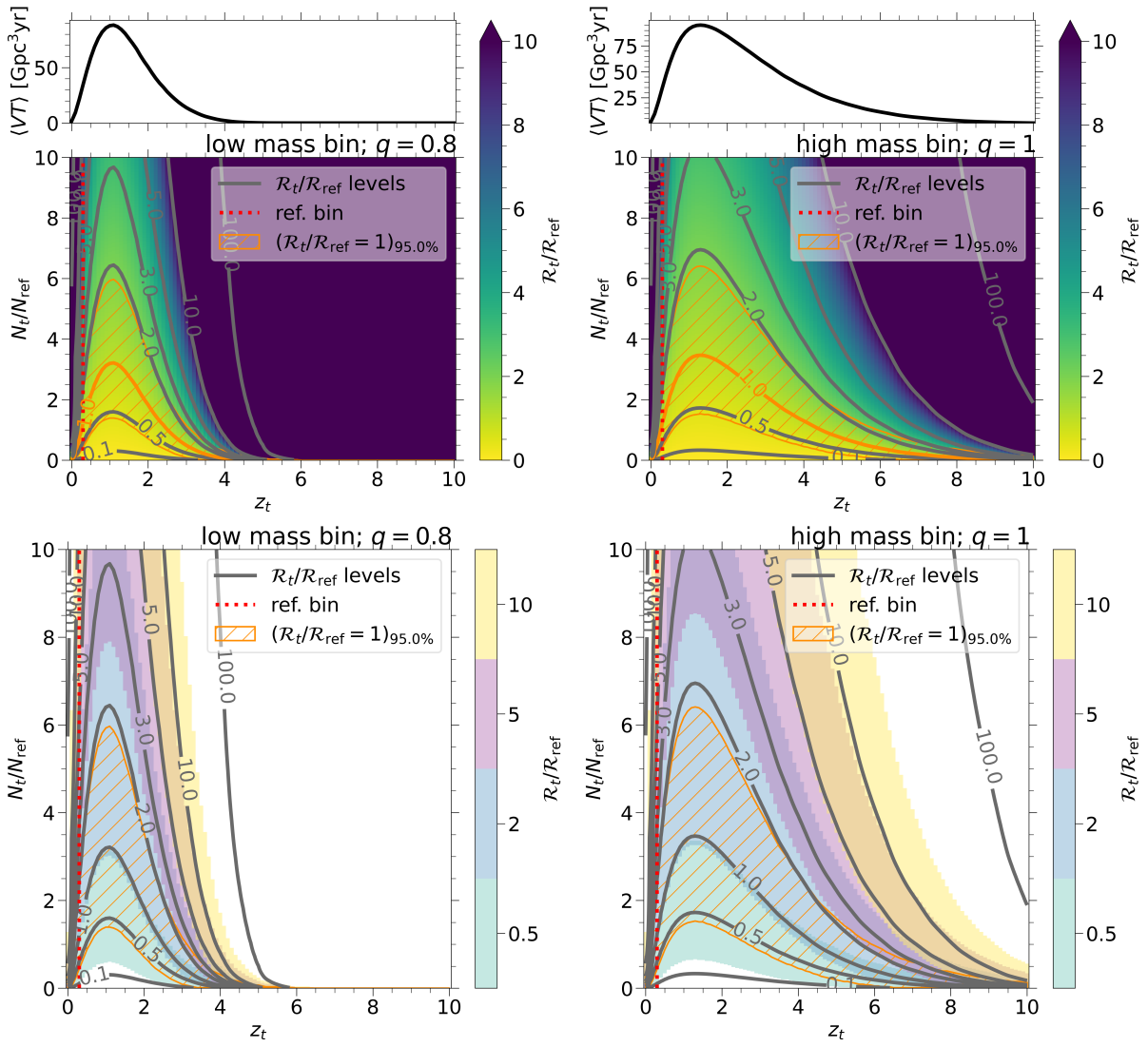

Figure 3: *Top*: merger rate density ratio colormap for fixed mass bins between the "target" and the "reference" redshift bins as a function of the target redshift and the number of detected events ratio. The black solid line in the top-side panel represents the $\langle VT \rangle$ evolution for the presented choice of the z_t axis. *Bottom*: merger rate density ratio 95% confidence intervals where we can infer different values of variation (on the colorbars) for fixed mass bins. *Left*: low mass bin $m_1 \in [inf, sup]$ $M_\odot$ with fixed mass ratio $q = 0.8$. *Right*: high mass bin $m_1 \in [inf, sup]$ $M_\odot$ with fixed mass ratio $q = 1$.
Dashed orange lines mark the 95% confidence interval where we can infer $\mathcal{R}_t/\mathcal{R}_{\rm ref} = 1$. The vertical red dotted line marks the position of the center of the reference redshift bin. Results in these plots are computed using a fixed redshift bin width of $z_w = 0.1$ and fixed choices of the "reference" number of events: $N_{\rm ref} = 10$. {fig:N_zt_map}

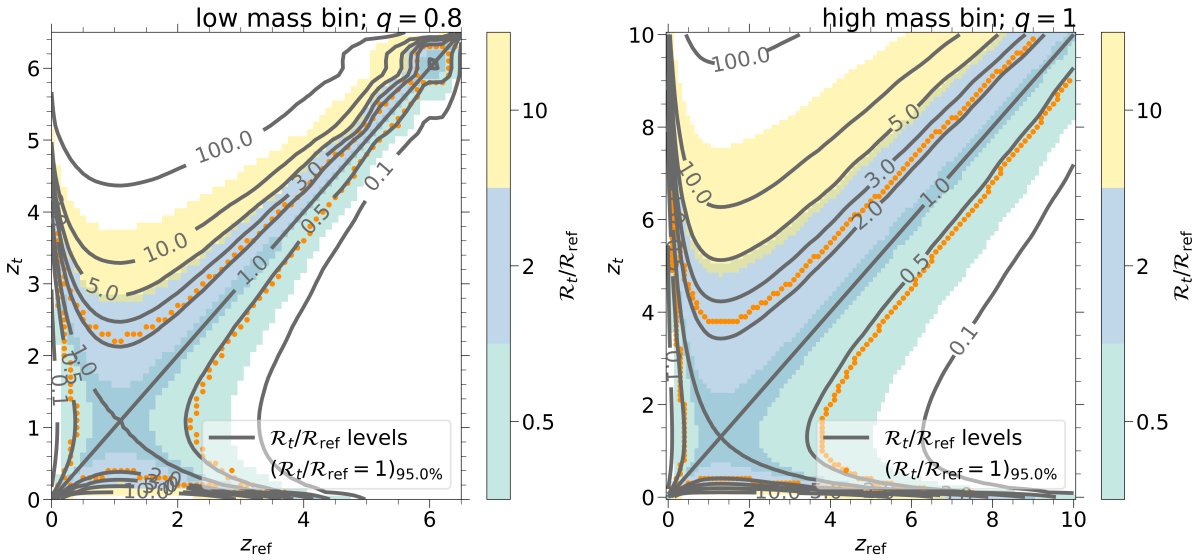

Figure 4: Merger rate density ratio 95% confidence intervals where we can infer different values of variation (on the colorbars) for fixed mass bins: $m_1 \in [inf, sup]$ $M_\odot$ with $q = 0.8$ (*Left*) and $m_1 \in [inf, sup]$ $M_\odot$ with $q = 1$ (*Right*). Dotted orange lines mark the 95% confidence interval for $\mathcal{R}_t/\mathcal{R}_{\rm ref} = 1$. Results are shown in the z_t vs. $z_{\rm ref}$ plane for $N_t/N_{\rm ref} = 1$, redshift bin widths $z_w = 0.1$, and $N_{\rm ref} = 10$. {fig:zt_zref_r

Chruślińska M., Nelemans G., Boco L., Lapi A., 2021, Monthly Notices of the Royal Astronomical Society, 508, 4994

Evans M., et al., 2021, arXiv e-prints, p. arXiv:2109.09882

Gerosa D., Bellotti M., 2024, Classical and Quantum Gravity, 41, 125002

Gerosa D., Pratten G., Vecchio A., 2020, Physical Review D, 102, 103020

Giacobbo N., Mapelli M., 2018, Monthly Notices of the Royal Astronomical Society, 480, 2011

Hendriks D. D., van Son L. A. C., Renzo M., Izzard R. G., Farmer R., 2023, Monthly Notices of the Royal Astronomical Society, 526, 4130

Kapadia S. J., et al., 2020, Classical and Quantum Gravity, 37, 045007

Klencki J., Moe M., Gladysz W., Chruslinska M., Holz D. E., Belczynski K., 2018, Astronomy & Astrophysics, 619, A77

Koch K.-R., 1988, Parameter estimation and hypothesis testing in linear models. Springer, Berlin

Li L., Lv G., Zhu C., Guo S., Ge H., Gu W., Li Z., He X., 2025, Explanation of the Mass Distribution of Binary Black Hole Mergers, doi:10.48550/ARXIV.2510.08231

Mennekens N., Vanbeveren D., 2016, Astronomy & Astrophysics, 589, A64

Neijssel C. J., et al., 2019, Monthly Notices of the Royal Astronomical Society, 490, 3740

Riley J., Mandel I., Marchant P., Butler E., Nathaniel K., Neijssel C., Shortt S., Vigna-Gómez A., 2021, Monthly Notices of the Royal Astronomical Society, 505, 663

Romagnolo A., Belczynski K., Klencki J., Agrawal P., Shenar T., Szécsi D., 2023, Monthly Notices of the Royal Astronomical Society, 525, 706

Spera M., Mapelli M., Bressan A., 2015, Monthly Notices of the Royal Astronomical Society, 451, 4086

Stevenson S., Sampson M., Powell J., Vigna-Gómez A., Neijssel C. J., Szécsi D., Mandel I., 2019, The Astrophysical Journal, 882, 121

The LIGO Scientific Collaboration et al., 2025, arXiv e-prints, p. arXiv:2508.18083

Tong H., et al., 2025, Evidence of the pair instability gap in the distribution of black hole masses, doi:10.48550/ARXIV.2509.04151

Vink J. S., de Koter A., Lamers H. J. G. L. M., 2001, Astronomy & Astrophysics, 369, 574

Wiktorowicz G., Wyrzykowski Ł., Chruslinska M., Klencki J., Rybicki K. A., Belczynski K., 2019, The Astrophysical Journal, 885, 1

Wilks S. S., 1938, The Annals of Mathematical Statistics, 9, 60

Woosley S. E., 2017, The Astrophysical Journal, 836, 244

Woosley S. E., Heger A., 2021, The Astrophysical Journal Letters, 912, L31

Xin C., Renzo M., Metzger B. D., 2022, Monthly Notices of the Royal Astronomical Society, 516, 5816

van Son L. A. C., et al., 2022, The Astrophysical Journal, 931, 17

van Son L. A. C., et al., 2025, The Astrophysical Journal, 979, 209