LIGO-Virgo-KAGRA's Oldest Black Holes

Probing Star Formation at Cosmic Noon With GWTC-3

Fishbach, M.; van Son, L.

DOI
10.3847/2041-8213/ad0560

Publication date
2023

Document Version
Final published version

Published in
Astrophysical Journal Letters

The Astrophysical Journal Letters, 957:L31 (11pp), 2023 November 10 https://doi.org/10.3847/2041-8213/ad0560

LIGO–Virgo–KAGRA's Oldest Black Holes: Probing Star Formation at Cosmic Noon With GWTC-3

Maya Fishbach[1] and Lieke van Son[2,3,4]
[1] Canadian Institute for Theoretical Astrophysics, David A. Dunlap Department of Astronomy and Astrophysics, and Department of Physics, 60 St George St, University of Toronto, Toronto, ON M5S 3H8, Canada; fishbach@cita.utoronto.ca
[2] Center for Astrophysics |Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[3] Anton Pannekoek Institute for Astronomy and GRAPPA, University of Amsterdam, NL-1090 GE Amsterdam, The Netherlands
[4] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85741 Garching, Germany
Received 2023 July 31; revised 2023 October 12; accepted 2023 October 18; published 2023 November 9

Abstract

In their third observing run, the LIGO–Virgo–KAGRA gravitational-wave (GW) observatory was sensitive to binary black hole (BBH) mergers out to redshifts $z_{\rm merge} \approx 1$. Because GWs are inefficient at shrinking the binary orbit, some of these BBH systems likely experienced long delay times τ between the formation of their progenitor stars at $z_{\rm form}$ and their GW merger at $z_{\rm merge}$. In fact, the distribution of delay times predicted by isolated binary evolution resembles a power law $p(\tau) \propto \tau^{\alpha_\tau}$ with slope $-1 \lesssim \alpha_\tau \lesssim -0.35$ and a minimum delay time of $\tau_{\rm min} = 10$ Myr. We use these predicted delay time distributions to infer the formation redshifts of the ∼70 BBH events reported in the third GW transient catalog GWTC-3 and the formation rate of BBH progenitors. For our default $\alpha_\tau = -1$ delay time distribution, we find that GWTC-3 contains at least one system (with 90% credibility) that formed earlier than $z_{\rm form} > 4.4$. Comparing our inferred BBH progenitor formation rate to the star formation rate, we find that at $z_{\rm form} = 4$, the number of BBH progenitor systems formed per stellar mass was $6.4^{+9.4}_{-5.5} \times 10^{-6}\, M_\odot^{-1}$ and this yield dropped to $0.134^{+1.6}_{-0.127} \times 10^{-6}\, M_\odot^{-1}$ by $z_{\rm form} = 0$. We discuss implications of this measurement for the cosmic metallicity evolution, finding that for typical assumptions about the metallicity dependence of the BBH yield, the average metallicity at $z_{\rm form} = 4$ was $\langle \log_{10}(Z/Z_\odot) \rangle = -0.3^{+0.3}_{-0.4}$, although the inferred metallicity can vary by a factor of ≈3 for different assumptions about the BBH yield. Our results highlight the promise of current GW observatories to probe high-redshift star formation.

Unified Astronomy Thesaurus concepts: Massive stars (732); Astrophysical black holes (98); Stellar mass black holes (1611); Star formation (1569); Metallicity (1031); Gravitational wave astronomy (675)

1. Introduction

The gravitational-wave (GW) detector network consisting of Advanced LIGO (LIGO Scientific Collaboration et al. 2015), Advanced Virgo (Acernese et al. 2015), and KAGRA (Akutsu et al. 2021) has observed GW radiation from binary black holes (BBHs) that merge at redshifts $z_{\rm merge} \lesssim 1$ (e.g., Abbott et al. 2019a, 2021a; Nitz et al. 2021; The LIGO Scientific Collaboration et al. 2021; Olsen et al. 2022), with planned detector upgrades expanding the network sensitivity to $z_{\rm merge} \lesssim 2$ (miller et al. 2015; Abbott et al. 2018; Weizmann Kiendrebeogo et al. 2023). Observing stellar-mass binaries that merge beyond $z_{\rm merge} \gtrsim 2$, an era known as "cosmic noon" when the Universe formed most of its stars, requires a next generation of GW detectors proposed for the 2030s, such as the Einstein Telescope (Maggiore et al. 2020) and Cosmic Explorer (Evans et al. 2021). These next-generation observatories would be sensitive to compact binary mergers past $z_{\rm merge} > 50$ (Hall & Evans 2019; Kalogera et al. 2021; Borhanian & Sathyaprakash 2022), likely observing every BBH merger in the Universe and opening up an enormous discovery potential for learning about high-redshift star formation, the first generation (Population III) stars, the assembly and growth of the first galaxies, cosmic metallicity evolution, and the formation histories of BBH systems (Vitale et al. 2019; Ng et al. 2021; Chruślińska 2022; Ng et al. 2022).

Although the current LIGO–Virgo–KAGRA (LVK) network cannot observe mergers that happened earlier than $z_{\rm merge} \approx 1$, many of the observed low-redshift mergers may have formed at significantly higher redshifts. Merging binaries experience a delay time between the formation of the progenitor stars and the binary compact object merger. Because massive stars are short lived, the bulk of this delay time typically consists of the GW-driven binary inspiral. The GW inspiral time $\tau_{\rm inspiral}$ is very sensitive to the orbital separation a, scaling as $\tau_{\rm inspiral} \propto a^4$ (Peters 1964). In fact, a binary needs to reach extremely short separations $\mathcal{O}(10R_\odot)$ just to merge within the age of the Universe (Peters 1964). Small increases in the orbital separation drastically increase the GW inspiral time, implying that even a narrow distribution of initial orbital separations will cause the distribution of delay times to extend out to very long delays.

Indeed, different BBH formation channels (see, e.g., Mapelli 2020; Mandel & Farmer 2022 for a review) typically predict delay time distributions with a long tail that extends beyond a Hubble time. This applies to dynamically assembled BBH systems or stellar triples in various environments (e.g., Rodriguez et al. 2016; Antonini et al. 2017; Di Carlo et al. 2020; Yang et al. 2020; Michaely & Naoz 2022) as well as those resulting from isolated binary evolution (e.g., Dominik et al. 2013; Lamberts et al. 2016; Mapelli et al. 2017). Here we focus on predictions from isolated binary evolution.

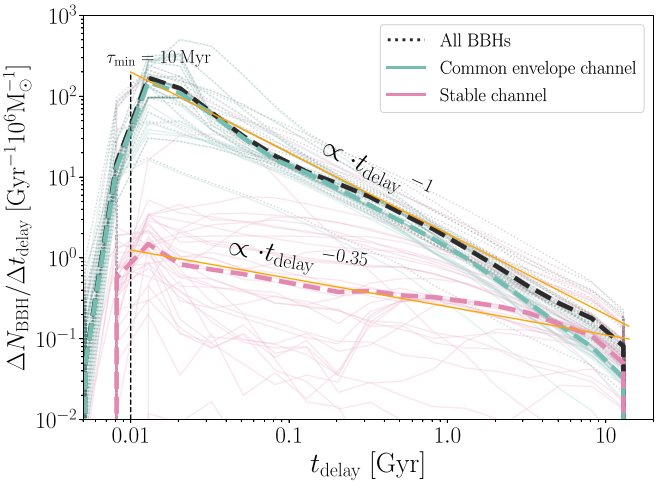

Figure 1. Delay time distributions derived from all physics variations explored in van Son et al. (2022a). We show histograms of the delay times per $10^6\,M_\odot$ of stellar mass formed, assuming a flat-in-log distribution of metallicities, and bin sizes $\Delta \log t_{\rm delay} = 0.2$. Thick dashed lines show the median for each formation channel and orange thin lines show the power laws used in this work.

Different isolated binary formation channels generally predict distinct distributions for the separations at BBH formation, and thus distinct delay time distributions. For example, the common envelope channel leads to shorter final separations than the stable mass transfer channel, because common envelopes shrink the binary orbit more efficiently than stable mass transfer (see Figure 1, and, e.g., Bavera et al. 2021; Gallegos-Garcia et al. 2021; van Son et al. 2022b).[5] Similarly, chemically homogeneously evolving stars are expected to stay compact and tend toward shorter delay time distributions at sufficiently low metallicities (e.g., Marchant et al. 2016; du Buisson et al. 2020; Riley et al. 2021).

Because of these distinct predictions, directly measuring the delay time distribution probes the formation channels of GW sources. For mergers involving neutron stars, the delay time distribution can be inferred from a population of their host galaxies (Safarzadeh et al. 2019; Adhikari et al. 2020), the host galaxy properties of short gamma-ray bursts (Zevin et al. 2022), or the *r*-process enrichment history (e.g., Naidu et al. 2022). For BBH sources without uniquely identified host galaxies, measuring the redshift evolution of the merger rate (Fishbach et al. 2018; Safarzadeh et al. 2019; Vitale et al. 2019; Callister et al. 2020; Callister & Farr 2023; Edelman & Farr 2023) can inform the delay time distribution (Vitale et al. 2019; Fishbach & Kalogera 2021; Karathanasis et al. 2023). Fishbach & Kalogera (2021) found that the relatively steep redshift evolution of the BBH merger rate between $z_{\rm merge} = 0$ and $z_{\rm merge} = 1$, compared to the low-metallicity star formation rate (SFR) model from Madau & Fragos (2017), favors short delay times, ruling out delay time distributions with typical delays $\gtrsim 3$ Gyr. In the absence of a direct model for the SFR, Mukherjee & Dizgah (2022) showed that the distribution of delay times between star formation and BBH mergers can be inferred by cross-correlating the redshift distributions of BBH mergers with electromagnetic tracers of star formation (e.g., line intensity mapping).

[5] This difference is almost "by design" since the energy formalism typically used in population synthesis codes to approximate common envelopes was first formulated to explain very short period double white dwarfs as the progenitors of Type Ia supernovae (Iben & Tutukov 1984; Webbink 1984).

Conversely, if the delay time distribution is known, the redshift evolution of the BBH merger population directly constrains the BBH progenitor formation rate. The progenitor formation rate depends on the SFR, stellar initial mass function (IMF), and the cosmic metallicity as a function of redshift. These are highly important yet uncertain processes, particularly at high redshifts (Madau & Dickinson 2014; Maiolino & Mannucci 2019; Smith 2020). Therefore, GW mergers present an exciting opportunity to probe star-forming conditions in the high-redshift Universe (see, e.g., Chruślińska 2022, for a review).

Previous studies have studied the BBH progenitor formation rate within a population synthesis framework, simulating BBH merger events for a range of parameters that govern the metallicity-specific SFR. By comparing to LVK observations, one can place constraints on the input parameters (e.g., Riley & Mandel 2023). Alternatively, one can simultaneously fit the delay time distribution and the progenitor formation rate to the GW observations in a data-driven approach, but the two will be highly degenerate with each other unless one restricts the model flexibility by placing some physical priors (e.g., Vitale et al. 2019).

In this work, we assume a simplified form of the delay time distribution motivated by theoretical predictions, and use it to propagate the observed merger redshift of each BBH event in GWTC-3 backward to a probability distribution of its progenitor's formation redshift $z_{\rm form}$, while simultaneously inferring the progenitor formation rate. Our approach sits between a population synthesis forward model and a data-driven inference, and can be thought of as a highly simplified version of "backward population synthesis" (Andrews et al. 2018, 2021; Wong et al. 2023). Unlike full backward population synthesis, we apply only a couple of predictions from population synthesis and combine them with GW observations, starting with a delay time distribution to infer BBH progenitor formation redshifts. Indeed, one application of the backward population synthesis approach is providing a straightforward consistency check that a given population synthesis simulation predicts delay times and metallicity-specific BBH rates that match both GW and SFR observations.

The remainder of this paper is structured as follows. In Section 2, we introduce the theoretically motivated delay time distributions and derive the relationship between delay times, merger redshifts, and formation redshifts. In Section 3, we model the BBH merger rate in terms of a fixed delay time distribution and an unknown progenitor formation rate, and infer the progenitor formation rate from the GWTC-3 data. We then adopt a metallicity-dependent yield $dN_{\rm BBH}/dM_{\rm SF}(Z)$ motivated by population synthesis and an SFR model from galaxy observations to turn our inference of the BBH progenitor rate into a constraint on the cosmic metallicity evolution (Section 4). In Section 5, we discuss what our population fit implies for the merger and formation redshifts of individual GWTC-3 BBH events, showing that GWTC-3 likely contains several systems that formed before cosmic noon. We discuss limitations of our method and future directions in Section 6 before concluding in Section 7.

2. Delay Time Distributions

Before inferring the BBH formation rate in the following sections, we first review the relationship between the delay time distribution $p(\tau)$, the merger rate R_m, and the formation rate R_f,

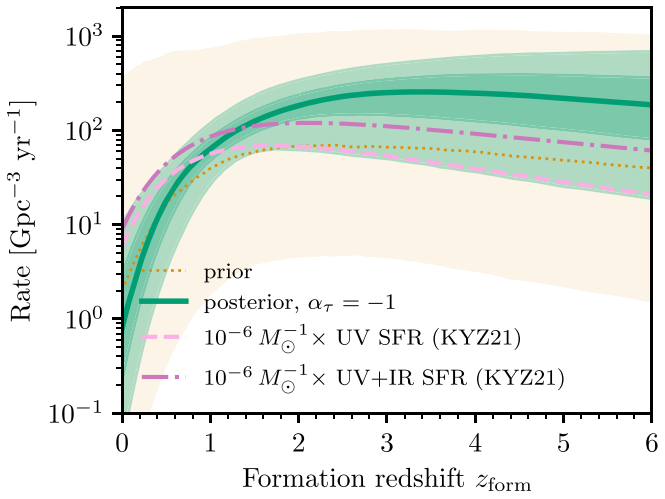

Figure 2. Inferred formation rate assuming the default $\alpha_\tau = -1$ delay time distribution (the green line shows the median a posteriori formation rate as a function of redshift, while the shaded bands enclose the 50% and 90% posterior probabilities). We adopt the formation rate parameterization from Katsianis et al. (2021) with two parameters controlling the shape and one parameter controlling the amplitude (see Equation (6)); the prior is shown in yellow (the dotted line shows median a priori formation rate, while the shaded band encloses the 90% prior probability). For comparison, we show the UV SFR (dashed light pink line) and UV+IR SFR (dotted–dashed dark pink line) from Katsianis et al. (2021) normalized by $10^{-6}\,M_\odot$.

and explain our choice of delay time distribution models. Assuming a delay time τ, a binary that forms at lookback time $t_{\rm form}$ merges at the lookback time:

$$t_{\rm merge} = t_{\rm form} - \tau. \tag{1}$$

Assuming a distribution of delay times $p_d(\tau)$, a binary that formed at $t_{\rm form}$ will merge at a time $t_{\rm merge}$ drawn from the distribution:

$$p(t_{\rm merge}|t_{\rm form}) = p_d(\tau = t_{\rm form} - t_{\rm merge}). \tag{2}$$

The rate of mergers at a given time $t_{\rm merge}$ depends on the formation rate in addition to the delay time distribution (e.g., Nakar 2007; Abbott et al. 2016):

$$\begin{aligned} R_m(t_{\rm merge}) &= \int p(t_{\rm merge}|t_{\rm form}) R_f(t_{\rm form}) dt_{\rm form} \\ &= \int p_d(\tau) R_f(t_{\rm form} = t_{\rm merge} + \tau) d\tau, \end{aligned} \tag{3}$$

where in the second line we changed variables from $t_{\rm form}$ to τ using Equation (2). Meanwhile, by Bayes theorem, given a binary that merged at time $t_{\rm merge}$, the probability it formed at time $t_{\rm form}$ is given by:

$$p(t_{\rm form}|t_{\rm merge}) = \frac{p(t_{\rm merge}|t_{\rm form}) R_f(t_{\rm form})}{R_m(t_{\rm merge})}. \tag{4}$$

Therefore, the BBH merger rate, which we can measure with GW observations, informs a combination of the progenitor formation rate and the delay time distribution (Equation (3)) and allows us to infer the formation times of individual BBH events from their merger times (Equation (4)).

Although the above equations are expressed in terms of lookback time, we assume a cosmological model (in this case, flat Λ cold dark matter (ΛCDM) with parameters from Planck Collaboration et al. 2016) to convert between lookback time t and redshift z (Astropy Collaboration et al. 2018), so that $z_{\rm merge}$ ($z_{\rm form}$) is the cosmological redshift at the lookback time $t_{\rm merge}$ ($t_{\rm form}$).

We assume a power-law parameterization for the delay time distribution with slope α_τ, minimum delay time $\tau_{\rm min}$, and maximum delay time $\tau_{\rm max}$:

$$p_d(\tau) \propto \tau^{\alpha_\tau} \Theta(\tau_{\rm min} < \tau < \tau_{\rm max}), \tag{5}$$

where Θ denotes the indicator function. In reality, the delay time distribution varies depending on the BBH mass and formation metallicity. For simplicity, we neglect this dependence and assign the delay time distribution of Equation (5) to every BBH system regardless of its mass and any assumptions about its formation metallicity, but see Section 6 for a discussion on this assumption. We bound our uncertainty in the delay time distribution by varying the power-law slope α_τ.

In Figure 1 we show variations of the delay time distribution inferred from the population synthesis predictions of van Son et al. (2022a). These results assume a universal Kroupa (2001) IMF and average over a flat-in-log distribution of metallicities. Motivated by these models, we explore three slopes: $\alpha_\tau = -1$ (our default model; similar to the prediction for "All BBHs," which is typically dominated by the common envelope channel), $\alpha_\tau = -0.35$ (similar to the prediction from the stable mass transfer channel), and an intermediate slope of $\alpha_\tau = -0.7$. We further fix $\tau_{\rm min} = 10$ Myr to represent the typical lifetime of a massive star, and fix $\tau_{\rm max}$ to the maximum lookback time of progenitor formation, which we take to be $t_0 = 13.62$ Gyr (corresponding to a maximum progenitor formation redshift $z_0 = 20$). Choices of $\tau_{\rm max} > t_0$ are equivalent to changing the normalization of the progenitor formation rate. As we will see in the following, we are not yet sensitive to the choice of z_0, as GWTC-3 is unlikely to contain any events that formed earlier than $z_{\rm form} \gtrsim 10$ for these choices of delay time distributions.

3. Progenitor Formation Histories

We now use the GWTC-3 observations to fit the BBH merger rate, which we model as a convolution of an unknown progenitor formation rate and a theoretically motivated delay time distribution (Equation (3)). The result is a measurement of the progenitor formation rate as a function of redshift (i.e., the formation rate of systems that will merge as BBHs within $\tau_{\rm max} = 13.62$ Gyr).

Our approach is similar to the progenitor formation rate calculation by Fishbach & Kalogera (2021) with GWTC-2 (see their Figure 7). However, while Fishbach & Kalogera (2021) assumed that the progenitor formation rate follows the low-metallicity SFR given by Madau & Fragos (2017) and fit for the metallicity threshold and scatter in the mean metallicity–redshift relation, here we adopt a more agnostic model, simply assuming that the progenitor formation rate can be described by the functional form from Katsianis et al. (2021) with free parameters $\mathcal{N}$, a, and b:

$$R_f(z_{\rm form}) = \mathcal{N} T(z_{\rm form})^a \exp(-bT(z_{\rm form})), \tag{6}$$

where $\mathcal{N}$ is a normalization in units of Gpc^{-3} yr^{-1}, $T(z)$ is the age (in units of Gyr) of the Universe at redshift z, b has units of Gyr^{-1}, and a is unitless. Although this function naturally falls to zero as $z_{\rm form}$ approaches infinity (when $T = 0$), we set R_f to zero for $z_{\rm form} > z_0$ with $z_0 = 20$. For each of the three delay time distributions we consider, we fit $\mathcal{N}$, a, and b, along with the

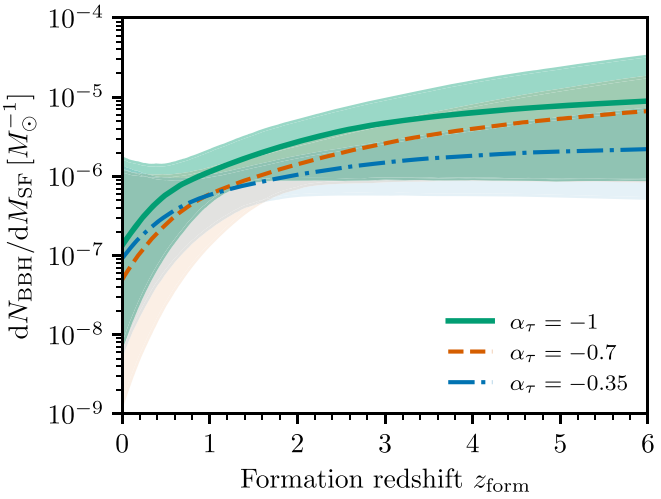

Figure 3. Number of merging BBH systems formed per star-forming mass as a function of redshift z_{form}, as inferred under three different delay time distributions with power-law slopes $\alpha_\tau = -1$ (default, the green solid line), $\alpha_\tau = -0.7$ (orange dashed line), and $\alpha_\tau = -0.35$ (blue dotted–dashed line) and the Katsianis et al. (2021) UV SFR. The lines denote the median BBH yield at each formation redshift while the shaded bands enclose the 90% posterior probability.

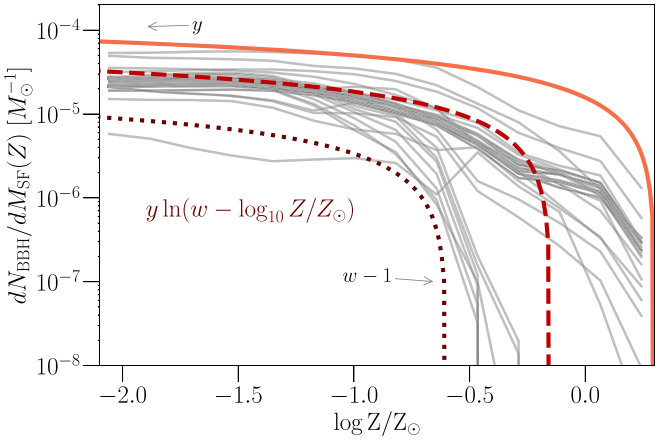

Figure 4. Number of merging BBH systems formed per star-forming mass as a function of metallicity (i.e., the BBH yield) for the physics variations explored in van Son et al. (2022a) shown as the gray lines. We show analytical approximations of low (dotted), medium (dashed), and high (solid line) BBH yields following Equation (8).

BBH mass and spin distribution in a hierarchical Bayesian framework (Loredo 2004; Mandel et al. 2019). Additional analysis details can be found in Appendix.

Figure 2 shows the fit to the progenitor formation history (green) assuming the default $\alpha_\tau = -1$ delay time model. Although we fit the formation rate out to $z_0 = 20$, we only plot the formation rate for $z_{\mathrm{form}} \leqslant 6$, because we expect the GWTC-3 events to have formed within this redshift range (see Section 5 and Figure 6). For reference, we show the prior in yellow, the Katsianis et al. (2021) UV fit to the SFR as the dashed light pink line, and the Katsianis et al. (2021) UV+IR fit to the SFR as the dotted–dashed dark pink line. Both SFRs are normalized by $10^{-6}\ M_\odot$.

In the following, we use the UV SFR as the reference SFR because Katsianis et al. (2021) find that it better matches the observed stellar mass density evolution, whereas they note that the UV+IR SFR overestimates the observed stellar mass density by up to ∼0.5 dex. However, we caution that observational uncertainties affect both the SFR and the stellar mass density (e.g., Narayanan et al. 2023). These uncertainties are currently not significant compared to the uncertainties associated with the GW inference and the choice of delay time model, so we adopt a fixed SFR model for the following proof-of-principle calculations.

We find that the inferred BBH progenitor formation rate rises more steeply with increasing redshift than the SFR, in particular over the range $z_{\mathrm{form}} < 4$ where we expect to get the most meaningful constraints with GWTC-3. This is illustrated in Figure 3, where we show the yield $dN_{\mathrm{BBH}}/dM_{\mathrm{SF}}$ inferred under the three different delay time models. The yield is defined as the number of BBH progenitors that will merge within $\tau_{\mathrm{max}} = 13.62$ Gyr formed per stellar mass. For power-law delay time distributions with slopes $\alpha_\tau = -1$ or shallower, we infer that the BBH yield decreases with decreasing z_{form}, from $6.4^{+9.4}_{-5.5} \times 10^{-6}\ M_\odot^{-1}$ at $z_{\mathrm{form}} = 4$ to $0.134^{+1.6}_{-0.127} \times 10^{-6}\ M_\odot^{-1}$ at $z_{\mathrm{form}} = 0$ under the default $\alpha_\tau = -1$ delay time model.

We find that compared to the default $\alpha_\tau = -1$ delay time distribution, the shallower $\alpha_\tau = -0.7$ and $\alpha_\tau = -0.35$ models provide a worse fit to the GW data in combination with our prior on the progenitor formation history. Under these shallow delay time models, our assumed formation rate model a priori limits the merger rate to evolve only by a factor of ≈2 between $z_{\mathrm{merge}} = 0$ and $z_{\mathrm{merge}} = 1$, whereas the data prefer a factor of ≈7 (Abbott et al. 2023). This worse fit is further illustrated by the maximum likelihood values of the fits under the different delay time assumptions, which differ by a factor of 26.5 (3.3) in favor of the $\alpha_\tau = -1$ assumption relative to the $\alpha_\tau = -0.35$ ($\alpha_\tau = -0.7$) assumption. Either our prior for the progenitor formation history is not sufficiently flexible (as we discuss in Section 6), or, if our prior captures the range of reasonable BBH progenitor formation rates, the $\alpha_\tau = -1$ model is a better description of the true BBH delay time distribution.

4. Metallicity Evolution

We can use our measurement of the BBH yield $dN_{\mathrm{BBH}}/dM_{\mathrm{SF}}(z_{\mathrm{form}})$ presented in Figure 3 to infer the cosmic metallicity evolution with redshift. This is because we expect the BBH yield $dN_{\mathrm{BBH}}/dM_{\mathrm{SF}}$ to depend strongly on the formation metallicity Z (Belczynski et al. 2010). Given a theoretically motivated assumption for $dN_{\mathrm{BBH}}/dM_{\mathrm{SF}}(Z)$, we can use our measurement of $dN_{\mathrm{BBH}}/dM_{\mathrm{SF}}(z_{\mathrm{form}})$ to infer the metallicity distribution as a function of redshift $p(Z|z_{\mathrm{form}})$. These quantities are related as:

$$dN_{\mathrm{BBH}}/dM_{\mathrm{SF}}(z_{\mathrm{form}}) = \int dN_{\mathrm{BBH}}/dM_{\mathrm{SF}}(Z)\, p(Z|z_{\mathrm{form}})\, dZ. \quad (7)$$

In Figure 4 we show predictions for the yield as a function of metallicity under different physics variations from van Son et al. (2022a) in gray. These predictions assume a solar metallicity value of $Z_\odot = 0.0142$ (Asplund et al. 2009), the default value in COMPAS (Riley et al. 2022); the remainder of our calculations therefore adopt this $Z_\odot$ value. The flat behavior at low metallicities followed by a steep drop off toward higher metallicities is characteristic for the BBH yield and has been found before; see, e.g., Figure 6 from Klencki et al. (2018), Figure 1 from Neijssel et al. (2019), Figure 1 from Santoliquido

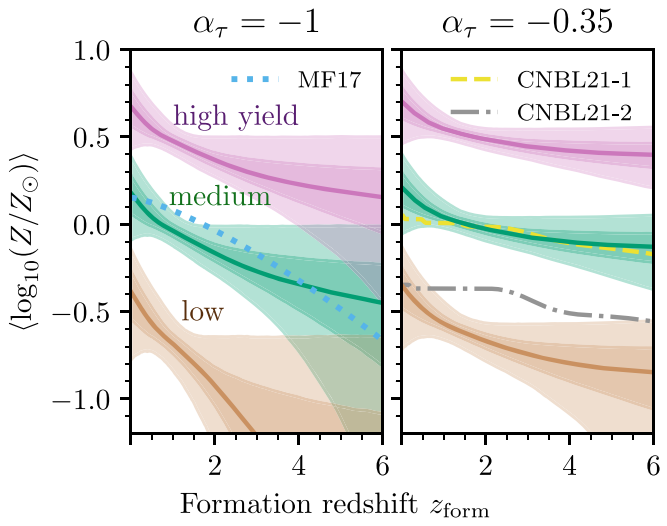

Figure 5. Inferred mean $\log_{10}$ metallicity as a function of redshift, assuming the default delay time distribution ($\alpha_\tau = -1$; left panel) or a shallow delay time distribution ($\alpha_\tau = -0.35$; right panel). We fix the standard deviation $\sigma_{\log_{10}}$ of the metallicity distribution to 0.2 dex at each redshift, and assume three different functions for the BBH yield as a function of metallicity: high yield (pink), medium yield (green), and low yield (brown) as defined in Figure 4. The solid lines show the median $\langle \log_{10}(Z/Z_\odot) \rangle$ at each $z_{\rm form}$ and the shaded bands show the 50% and 90% credible regions. For comparison, we overplot three different measurements from the literature: Madau & Fragos (2017) as the dotted, light blue curve (MF17), and two of the variations from Chruślińska et al. (2021) as the dashed, yellow curve (CNBL21-1) and dotted–dashed gray curve (CNBL21-2).

et al. (2021), Figure 1 from Broekgaarden et al. (2022), and Figures 17 and 18 from Iorio et al. (2023). Motivated by these predictions we take $dN_{\rm BBH}/dM_{\rm SF}(Z)$ to be of the form:

$$\frac{dN_{\rm BBH}}{dM_{\rm SF}}(Z) = y \ln(w - \log_{10} Z)\Theta(\log_{10} Z < w - 1), \quad (8)$$

so that y sets the BBH yield at low metallicities, and as $\log_{10} Z$ approaches $w - 1$ from below, the BBH yield rapidly falls to zero. Typical predictions have $10^{-5} < y < 6 \times 10^{-5}$ and $0.4 < w < 1.3$. To bound these possibilities, we consider a "high-yield" case with $(y, w) = (6 \times 10^{-5}, 1.3)$, a "medium-yield" case with $(y, w) = (3.5 \times 10^{-5}, 0.85)$ and a "low-yield" case with $(y, w) = (10^{-5}, 0.4)$, as shown by the red lines in Figure 4.

We approximate $p(Z|z_{\rm form})$ with a log-normal distribution with width $\sigma_{\log_{10} Z} = 0.2$ dex (though see Chruślińska 2022; van Son et al. 2023; we discuss the impact of the log-normal approximation in Section 6). Using Equation (7) with $dN_{\rm BBH}/dM_{\rm SF}(Z)$ given by Equation (8) and $dN_{\rm BBH}/dM_{\rm SF}(z_{\rm form})$ shown in Figure 3, we infer the mean log-metallicity $\langle \log_{10}(Z/Z_\odot) \rangle$ as a function of redshift.

The results are shown in Figure 5. The left panel shows $\langle \log_{10}(Z/Z_\odot) \rangle$ inferred under the default $\alpha_\tau = -1$ delay time distribution, while the right panel shows the inference under a shallow $\alpha_\tau = -0.35$ delay time distribution. Each panel displays the inferred metallicity under the three different $dN_{\rm BBH}/dM_{\rm SF}(Z)$ assumptions: high yield (pink), medium yield (green), and low yield (brown).

These measurements can be compared against various metallicity evolution results in the literature (Maiolino & Mannucci 2019; Chruślińska 2022). For reference, we overplot three examples: the mean log-metallicity from Madau & Fragos (2017; "MF17" or dotted, light blue line) and the peak metallicity curves from the left and right panels of Figure 7 of Chruślińska et al. (2021), calculated under two different assumptions ("CNBL21-1" and "CNBL21-2" in the yellow dashed and gray dotted–dashed lines, respectively). While we can directly compare our results to Madau & Fragos (2017) who also adopt a log-normal metallicity distribution, the comparison to Chruślińska et al. (2021) is less straightforward, since their metallicity distributions have an extended low-metallicity tail, resulting in a peak log-metallicity that is higher than the average. We expect this distinction to be less significant than the various uncertain assumptions illustrated in Figure 5, but we discuss the log-normal assumption further in Section 6. We also note that our results assume a constant scatter in the log-metallicity of 0.2 dex. If we instead assumed a larger (smaller) scatter, our uncertainty in the inferred $\langle \log_{10}(Z/Z_\odot) \rangle(z_{\rm form})$ would increase (decrease).

Assuming the default delay time model with $\alpha_\tau = -1$ (left panel of Figure 5) and a medium BBH yield (green), we find a mean log-metallicity today of $\langle \log_{10}(Z/Z_\odot) \rangle(z_{\rm form} = 0) = 0.2^{+0.2}_{-0.3}$, and at redshift 4, $\langle \log_{10}(Z/Z_\odot) \rangle(z_{\rm form} = 4) = -0.3^{+0.3}_{-0.4}$. (Here and in the following, we quote median values and 90% credibility intervals.) This is generally consistent with the Madau & Fragos (2017) metallicity evolution (MF17) and the first peak metallicity curve from Chruślińska et al. (2021; CNBL21-1), although our average log-metallicity may evolve more steeply than these models over the well-constrained range $z_{\rm form} \lesssim 3$.[6] If we assume a higher or lower BBH yield from Figure 4, the inferred metallicities increase or decrease by a factor of $\sim$3 (0.5 dex).

Meanwhile, different assumptions about the delay time distribution also affect the inferred metallicity evolution. A delay time model that favors longer delays, such as $\alpha_\tau = -0.35$, implies lower BBH yields across the plotted range $0 < z_{\rm form} < 6$ in order to avoid overpredicting the local merger rate, which largely consists of systems that formed at high redshifts. This lower BBH yield translates to higher inferred metallicities (right panel of Figure 5). For example, under the medium-yield assumption and the $\alpha_\tau = -0.35$ delay time model, we infer $\langle \log_{10}(Z/Z_\odot) \rangle(z_{\rm form} = 3) = -0.1^{+0.1}_{-0.1}$. However, as discussed in the previous section, the $\alpha_\tau = -0.35$ model provides a poor fit to the GW data compared to the $\alpha_\tau = -1$ model.

Our results illustrates the promise of BBH mergers as probes of the cosmic metallicity evolution, especially as GWs start providing improved constraints at formation redshifts past $z_{\rm form} > 3$ where the metallicity distribution inferred from other probes is highly uncertain (see, e.g., Figure 7 in Chruślińska et al. 2021). Our inferred metallicity evolution is based on models of the delay time distribution, the BBH yield, and SFR. If we instead assumed that the metallicity evolution is known, we could use our measurement of $dN_{\rm BBH}/dM_{\rm SF}(z_{\rm form})$ from Figure 3 and apply Equation (7) to infer the parameters governing the BBH yield as a function of metallicity, similar to the analysis in Fishbach & Kalogera (2021). In this study, the delay time distribution and BBH yield are inferred from population synthesis predictions, while we have used constraints from UV observations for the total SFR. Although each of these components is plagued by uncertainties, our inference of the metallicity evolution can be seen as a consistency check

[6] Note, however, that the steeper $\langle \log_{10}(Z/Z_\odot) \rangle(z_{\rm form})$ evolution may instead indicate a breakdown of the log-normal assumption; see Section 6.

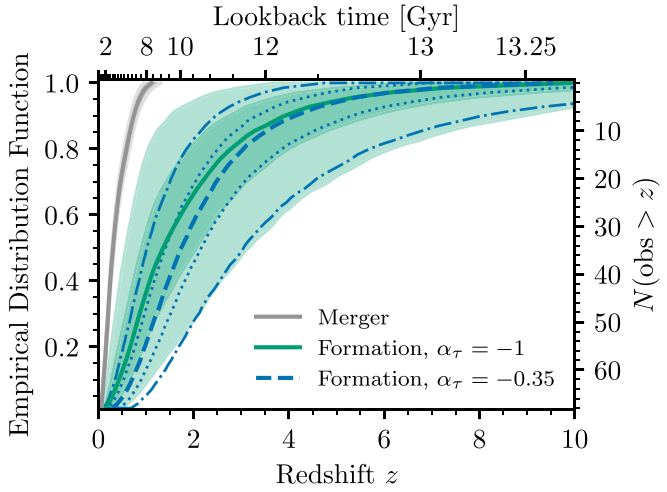

Figure 6. EDFs of the merger redshifts (gray) and formation redshifts (green and blue) of the 69 BBH events in GWTC-3 for fixed delay time distributions. The green filled band assumes a delay time distribution with power-law slope $\alpha_\tau = -1$, and the blue, unfilled band assumes $\alpha_\tau = -0.35$. The solid green and dashed blue lines denote the median EDF, and bands enclose the 50% and 90% credibilities. Both delay time distributions assume $\tau_{\rm min} = 10$ Myr. The number of BBH events in GWTC-3 with a (merger or formation) redshift above a given value is denoted in the right-hand y-axis.

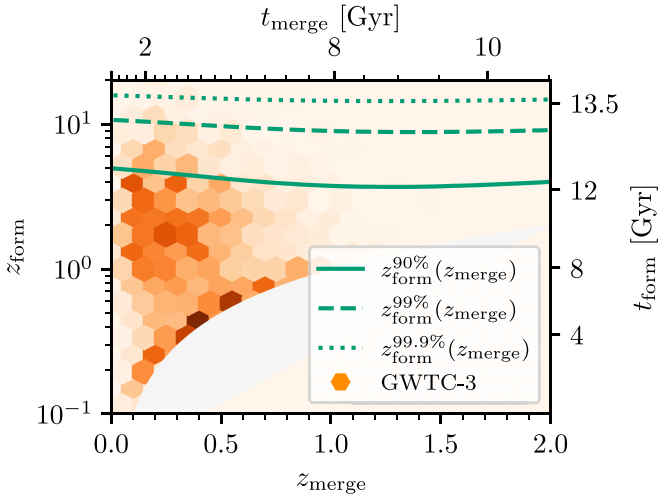

Figure 7. Orange: joint distribution of the merger and formation redshifts of detected BBH events at GWTC-3 sensitivity, inferred under the default population model (with the $\alpha_\tau = -1$ delay time distribution). We average over the population posterior uncertainty. The orange shading represents the probability density in a given ($z_{\rm merge}$, $z_{\rm form}$) hexagonal bin. For reference, the top and right axes show the corresponding lookback times. Green: at each $z_{\rm merge}$, the 90, 99, and 99.9 percentiles of the conditional $z_{\rm form}$ distribution (see Equation (4)). These represent the maximum formation redshift that we expect to probe given $\mathcal{O}(10)$, $\mathcal{O}(100)$, and $\mathcal{O}(1000)$ observations at a fixed $z_{\rm merge}$.

between these different model components. We find that our default delay time distribution (from the "All BBHs" prediction in Figure 1), the medium-yield assumption (dashed red line in Figure 4) and the Katsianis et al. (2021) SFR imply a metallicity evolution that is consistent with external measurements in the literature (e.g., Madau & Fragos 2017). On the other hand, the high-yield and low-yield predictions, coupled with our other assumptions, require metallicities that are significantly different from the external measurements shown in Figure 5 in order to match the GW merger rates.

5. Oldest Black Holes in GWTC-3

Within a hierarchical Bayesian population analysis, we simultaneously infer the population properties and update our inference of individual-event parameters (Fishbach et al. 2020; Galaudage et al. 2020; Miller et al. 2020; Essick & Fishbach 2021; Moore & Gerosa 2021). Therefore, our fit to the BBH merger rate in the previous sections, modeled in terms of the progenitor formation rate and delay time distribution, allows us to infer the probable merger redshift, delay time, and formation redshift jointly for each BBH event in GWTC-3.

Figure 6 shows the empirical distribution functions (EDFs) of the population-informed merger redshifts and formation redshifts of the 69 BBH events. To construct the EDFs, we draw one sample from each of the 69 population-informed redshift ($z_{\rm merge}$ or $z_{\rm form}$) posteriors and order the samples from smallest to largest. The EDF at the K-th ordered z sample takes the value $K/69$. Repeating this over 4000 draws, we calculate the median EDF of GWTC-3 redshifts and its 50% and 90% uncertainty bands.

The gray band in Figure 6 shows the EDF of the merger redshifts $z_{\rm merge}$. We use the population-informed $z_{\rm merge}$ posteriors inferred under our default population model (with $\alpha_\tau = -1$), but the single-event $z_{\rm merge}$ posteriors are similar across the different models we consider, because they are largely informed by the GW data. The EDF of the formation redshifts $z_{\rm form}$ inferred under this default $\alpha_\tau = -1$ population model is plotted in green (filled bands) in Figure 6, while the formation redshifts inferred under the $\alpha_\tau = -0.35$ model is shown in blue (unfilled bands). The population-informed formation redshift inferred for each event depends on the assumed delay time distribution, which favors longer delay times for the $\alpha_\tau = -0.35$ model compared to the default $\alpha_\tau = -1$ model, together with the inferred formation rate, which acts as a prior over the formation redshifts (see Equation (4)). Under the $\alpha_\tau = -0.35$ delay time distribution, events tends to experience longer delay times and therefore have higher formation redshifts, although the two EDFs overlap within statistical uncertainties due to the remaining uncertainty in the inferred formation histories.

We can see from Figure 6 that although the observed BBH systems all merged at $z_{\rm merge} \lesssim 1$ (lookback times $t_L^{\rm merge} \lesssim 8$ Gyr), some of them experienced long delays of several gigayears between formation and merger according to either delay time distribution. The maximum merger redshift in GWTC-3 is $\max(z_{\rm merge}) = 1.1^{+0.2}_{-0.2}$ (median and 90% credible interval). Nevertheless, for our default delay time distribution, more than 21 events in GWTC-3 formed earlier than $z_{\rm form} > 1$, 24^{+26}_{-20} events formed earlier than $z_{\rm form} > 2$, and 13^{+24}_{-12} events formed earlier than $z_{\rm form} > 3$ (with 90% credibility). Furthermore, at least one event in GWTC-3 formed earlier than $z_{\rm form} > 4.4$ (90% credibility), when the Universe had only formed $\sim$4% of its stellar mass (Katsianis et al. 2021).

To understand which merger redshifts in GWTC-3 are responsible for the highest formation redshifts (i.e., the oldest BHs), we plot the joint distribution of merger and formation redshifts in Figure 7. The orange histogram shows the detected distribution of ($z_{\rm merge}$, $z_{\rm form}$) according to our population inference under the default delay time model, with "detected" referring to BBH events observable at GWTC-3 sensitivity. We average over the population uncertainty. This population-averaged, detected distribution is often referred to as a posterior population distribution (PPD) in the GW literature (e.g., Abbott et al. 2019b).

Figure 7 shows that, thanks to the delay time distribution that peaks at $\tau_{\rm min} = 10$ Myr, many BBH events experienced short delay times, leading to a region of high probability density at $z_{\rm form} \approx z_{\rm merge}$ in the detected distribution. However, there is a second region of high density centered at $z_{\rm merge} \approx 0.2$ and $z_{\rm form} \approx 2$. This is caused by the fact that most observable BBH events merge at low redshifts $z_{\rm merge} \approx 0.2$, but we infer that the formation rate is very small at similarly low $z_{\rm form}$. Therefore, many of these low-redshift mergers experienced long delay times and formed at much higher redshifts. This is broadly consistent with theoretical predictions for the formation redshifts of the observed, low-$z_{\rm merge}$ BBH systems (Belczynski et al. 2016; Mapelli et al. 2018, 2019; Boco et al. 2021). Our simplified "backward population synthesis" approach therefore provides a straightforward test of whether predictions for the delay time distribution (which we fixed here to our default $\alpha_\tau = -1$, $\tau_{\rm min} = 10$ Myr model) yield reasonable BBH formation redshifts.

Meanwhile, the green solid, dashed, and dotted lines in Figure 7 show the 0.9, 0.99, and 0.999 quantiles of the conditional probability distribution $p(t_{\rm form}|t_{\rm merge})$ (given by Equation (4)), converting between lookback times to cosmological redshifts as necessary. With $\mathcal{O}(N)$ events at a given merger redshift, we expect one of them to have formed at the $\frac{N-1}{N}$ quantile of the $z_{\rm form}$ distribution. Therefore, we expect to probe the 0.9, 0.99, and 0.999 quantiles with $\mathcal{O}(10)$, $\mathcal{O}(100)$, and $\mathcal{O}(1000)$ events, respectively. Among the 69 BBH events in GWTC-3, the maximum formation redshift is $\max(z_{\rm form}) = 9.9^{+9.1}_{-6.2}$, which lies between the solid and dashed green lines. Interestingly, the largest $z_{\rm form}$ in a GW sample is unlikely to correspond to the largest $z_{\rm merge}$. As the GW detectors' sensitivity improves, we will probe higher formation redshifts $z_{\rm form}$ not directly because the detection horizon will increase to higher $z_{\rm merge}$, but because we will increase the BBH sample size, which allows us to probe higher quantiles of the $z_{\rm form}$ distribution. The highest $z_{\rm form}$ in the GW sample most likely corresponds to the $z_{\rm merge}$ that hosts the majority of detections ($z_{\rm merge} \approx 0.2$ for GWTC-3).

With thousands of observations at $z_{\rm merge} < 1$ (to be expected within the next couple of observing runs) we expect to probe progenitor formation past $z_{\rm form} \gtrsim 15$ according to our default delay time model and our inferred progenitor formation history (see the dotted green line in Figure 7). This suggests that with enough observing time, the current generation of GW observatories will already be sensitive to the first star formation in the Universe. However, this prediction depends on the assumed delay time distribution. This corroborates the need for next-generation GW detectors, which will provide a direct measurement of the merger rate at high redshifts.

6. Discussion on Simplifying Assumptions

In this work, we have made several simplifying assumptions that can be relaxed in future work. Here we briefly discuss the validity of these assumptions.

Mass and metallicity dependence of the delay time distribution. In Figure 8 we show the delay time distribution from van Son et al. (2022a) split between BBH systems formed at low (red), medium (yellow), and high (blue) metallicities. This shows a shift toward longer delay times for the highest metallicities. We suspect that this is caused by metallicity-dependent winds, in particular during the Wolf–Rayet (W-R) stage (e.g., Vink & de Koter 2005). When the system consists of a W-R+BH (the final stage before BBH formation), stronger winds at higher metallicities lead to wider separations, and thus longer delay times (this was also concluded by, e.g., Riley et al. 2021; Bavera et al. 2022).

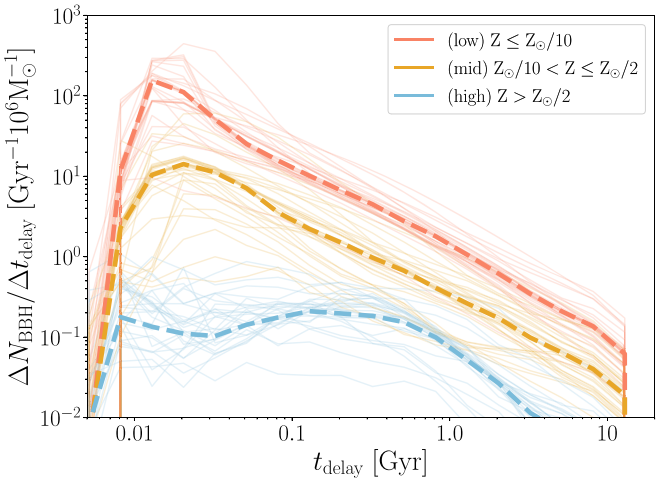

Figure 8. Similar to Figure 1, but binned by formation metallicity. This shows that the delay time distribution shifts toward longer delay times for high metallicities.

The metallicity dependence will cause the delay time distribution to correlate with the formation redshift $z_{\rm form}$, so that systems forming at high $z_{\rm form}$ likely experience shorter delays relative to systems forming at low $z_{\rm form}$. For a fixed BBH formation rate, this leads to a larger BBH merger rate at high redshifts compared to low redshifts, and therefore mimics the effect of changing the BBH formation rate. Neglecting this effect, we are likely overestimating the degree to which the BBH yield evolves with redshift. In other words, if we took into account the metallicity dependence of the delay time distribution, we would infer a more gradual evolution of the BBH formation yield with redshift (Figure 3), and correspondingly a weaker dependence of metallicity on redshift (Figure 5). However, we expect the impact on our conclusions to be mild, because (a) the delay time distribution depends only weakly on metallicity over the relevant (low-metallicity) range for BBH formation, as seen by the small difference between the $Z \leqslant Z_\odot/10$ and $Z_\odot/10 < Z \leqslant Z_\odot/2$ delay time distributions in Figure 8; and (b) even neglecting the metallicity dependence of the delay time distribution, we infer relatively mild evolution of the average metallicity with redshift over the well-constrained range $0 < z_{\rm form} < 3$ in Figure 5.

The delay time distribution could furthermore depend on BBH mass. In particular, van Son et al. (2022b) found that more massive BHs are formed exclusively by the stable mass transfer channel, while the common envelope channel dominates the formation of lower-mass systems. This result was confirmed by Belczynski et al. (2022) and Briel et al. (2023). Given the distinct delay time distributions of these two channels (Figure 1), this implies a mass dependence of the delay time distribution. Currently, the contribution of each of these channels is an active area of research (Neijssel et al. 2019; Bavera et al. 2021; Gallegos-Garcia et al. 2021; Marchant et al. 2021). If lighter BBH systems tend to experience shorter delay times compared to the more massive systems, we expect a correlation between BBH mass and merger redshift, in which more massive BBH systems are more likely to merge at lower redshifts compared to low-mass BBH systems. With current

GW data, there is no evidence for such a correlation, although it is not ruled out (Fishbach et al. 2021; van Son et al. 2022b; Abbott et al. 2023). Because the evolution of the mass distribution with redshift is degenerate with the evolution of the overall merger rate, if a negative correlation between BBH mass and merger redshift exists, we are likely underestimating the BBH merger rate at high redshifts, which is dominated by light systems that are harder to detect. We may therefore also be underestimating the progenitor formation rate at high redshifts, although we expect this effect to be subdominant to other uncertainties given that GWTC-3 does not clearly display a correlation between BBH mass and merger redshift.

In reality, the observed population of merging BBHs is most likely a mixture of multiple formation channels with a corresponding mixture of delay time distributions (e.g., Bouffanais et al. 2021; Zevin et al. 2021; Stevenson & Clarke 2022; Godfrey et al. 2023). Future work should simultaneously incorporate delay time distributions from several formation channels, including predictions from non-isolated binary evolution channels. The mass (or spin; e.g., Bavera et al. 2022) dependence of the delay time distribution might also provide a solution to differentiate between formation channels: if a formation channel predicts unique observable properties (such as massive and highly spinning black holes), then our analysis can be repeated for a subset of the observations that meets these criteria. For example, Fishbach & Fragione (2023) leveraged predicted delay time distributions from dynamical assembly of BBHs in globular clusters to infer the globular cluster formation history from the observed merger rate of systems with misaligned spins, because misaligned spins indicate a possible dynamical origin.

Flexible progenitor formation rate models. In this work, we used a simple three-parameter model for the progenitor formation rate, which follows a power law in time at early times (high redshift) and exponentially decays at late times (low redshifts). Future work should allow for more flexible formation rate models, such as Gaussian processes (Vitale et al. 2019), splines (Edelman & Farr 2023), or autoregressive processes (Callister & Farr 2023). In fact, there is already indication that our formation rate model provides a poor fit to the GWTC-3 data when we assume shallow delay time models ($\alpha_\tau = -0.7$ or $\alpha_\tau = -0.35$). This is consistent with the results of Fishbach & Kalogera (2021), who also found that short delay times were favored under the different models they considered for the BBH progenitor formation rate. However, because of the degeneracy between the delay time distribution and the progenitor formation rate, this may instead indicate that our formation rate parameterization fails to fit the data adequately. If the delay time distribution is indeed shallower than the default assumption of $\alpha_\tau = -1$, the formation rate likely deviates from our simple parameterization at formation redshifts $z_{\rm form} > 2$.

Deviations from a log-normal metallicity distribution. When inferring the cosmic metallicity evolution in Section 4, we assumed a log-normal metallicity distribution at each redshift. However, there are observational and theoretical indications that the metallicity distribution may deviate from the log-normal approximation, featuring an extended low-metallicity tail that grows with redshift (see, e.g., the discussion in Chruślińska 2022). Depending on the BBH yield as a function of metallicity (see Figure 4), we expect GW observations to be predominantly sensitive to the amount of low-metallicity star formation as a function of redshift. Therefore, current GW data cannot distinguish between a log-normal metallicity distribution that peaks at lower metallicities and a distribution with an extended low-metallicity tail that peaks at higher metallicities. Future work should fit for more flexible metallicity distributions (e.g., van Son et al. 2023) in order to account for this degeneracy in the location and shape of the metallicity distribution as a function of redshift.

Universality of the IMF. Our calculations have assumed a universal Kroupa (2001) IMF. However, there are many indications that the IMF varies with metallicity (see, e.g., the reviews by Kroupa et al. 2013; Hopkins 2018). The formation rate of BBH progenitors is unlikely to be strongly affected by variations in the high-redshift IMF within observational uncertainties, because the formation rate of high-mass stars, which is a combination of the SFR and the IMF, is better measured than the SFR or IMF independently (Klencki et al. 2018; Chruślińska et al. 2020). Nevertheless, a nonuniversal IMF may affect the delay time distribution and the BBH mass distribution. Our current model only considers the delay time distribution, the BBH yield, and the SFR derived under a fixed IMF. Adding the IMF to this equation also implies that we can in principle use a similar method to measure the IMF at high redshift, in particular the transition between the IMF of Population III stars and the IMF today.

7. Summary

The preceding sections have demonstrated that existing GW observations by LVK are starting to reveal the conditions of BBH formation beyond cosmic noon. This argument stems from the fact that regardless of the precise formation scenario, some BBH systems are predicted to experience long delay times between the formation of their progenitor stars and their GW-driven merger. For this proof-of-principle study, we have adopted a few fixed delay time distributions motivated by predictions from isolated binary evolution. We then applied the assumed delay time distributions to GWTC-3 events to infer their progenitor formation redshifts, while simultaneously measuring the progenitor formation history out to $z_{\rm form} \gtrsim 4$ and providing implications for the cosmic metallicity evolution. Our main results are as follows.

1. Fixing the delay time distribution to our default distribution with power-law slope $\alpha_\tau = -1$ and minimum delay of $\tau_{\rm min} = 10\,{\rm Myr}$, we infer the BBH progenitor formation rate as a function of redshift $z_{\rm form}$. Assuming a star formation history model from Katsianis et al. (2021), we find that the number of BBH progenitor systems formed per stellar mass was likely higher in the past than today (for our default delay time model, the BBH yield was $6.4^{+9.4}_{-5.5} \times 10^{-6}\ M_\odot$ at $z_{\rm form} = 4$ compared to $0.134^{+1.6}_{-0.127} \times 10^{-6}\ M_\odot$ at $z_{\rm form} = 0$).
2. Combining our inferred BBH yield as a function of redshift with population synthesis predictions for the BBH yield as a function of metallicity, we measure the average metallicity as a function of redshift. For our default delay time, BBH yield, and star formation history models, we find that the mean log-metallicity $\langle \log_{10}(Z/Z_\odot) \rangle$ today is $0.2^{+0.2}_{-0.3}$ and was $-0.3^{+0.3}_{-0.4}$ at $z_{\rm form} = 4$. This is consistent with the metallicity evolution from Madau & Fragos (2017), highlighting that a simplified form of "backward population synthesis"

(Andrews et al. 2018; Wong et al. 2023) can provide a powerful self-consistency check on the various components of BBH population modeling.

3. Our fit to the BBH population allows us to infer the merger and formation redshifts of the 69 confident BBH events in GWTC-3 jointly. We find that GWTC-3 likely contains at least one BBH system that formed earlier than $z_{\rm form} > 4.4$.

The BBH merger rate is shaped by the SFR, IMF, BBH yield, and the delay time distribution. By fixing a few of these ingredients at a time and inferring the rest from the GW data, we can cross-check the various predictions of population synthesis with external measurements of the cosmic SFR and metallicity evolution. As the GW catalog grows and population synthesis simulations improve, the connection between these pieces will provide valuable insights into high-redshift and low-metallicity star formation, complementary to electromagnetic observations.

8. Software and Data

Posterior samples for the hyperparameters used in this work and data behind the figures are publicly available on Zenodo (Fishbach & van Son 2023). The population synthesis data used in this work is also available on Zenodo (van Son et al. 2022a, 2022b). This research made use of Astropy (http://www.astropy.org; Astropy Collaboration et al. 2013, 2018; Price-Whelan et al. 2018; Astropy Collaboration et al. 2022), jupyter (https://jupyter.org; Kluyver et al. 2016), matplotlib (https://matplotlib.org; Hunter 2007), numpy (https://numpy.org; Harris et al. 2020), scipy (https://scipy.org; Virtanen et al. 2020), jax (https://jax.readthedocs.io; Bradbury et al. 2018), and numpyro (https://num.pyro.ai; Bingham et al. 2019; Phan et al. 2019).

Acknowledgments

We are grateful to an anonymous referee for pointing out a flaw in the calculations that appeared in an earlier draft of this work. We furthermore thank Will Farr for his insightful comments on the manuscript and Reed Essick and Daniel Holz for inspiring discussions. M.F. is grateful for the hospitality of Perimeter Institute where part of this work was carried out. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Economic Development, Job Creation and Trade. L.v.S. acknowledges partial financial support from the National Science Foundation under grant No. (NSF grant No. 2009131), the Netherlands Organisation for Scientific Research (NWO) as part of the Vidi research program BinWaves with project No. 639.042.728, and the European Union's Horizon 2020 research and innovation program from the European Research Council (ERC, grant agreement No. 715063). This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. This is LIGO Document LIGO-P2300241.

Appendix
Details of the Population Fit

In Section 3, we fit the population distributions of masses, spins, and merger redshifts of the GWTC-3 BBH events. We use the 69 confident BBH events with false alarm rate (FAR) $< 1\ {\rm yr}^{-1}$ and both component masses heavier than 3 $M_\odot$, as selected by Abbott et al. (2023). Our population model consists of independent distributions for component masses m_1 and m_2, effective inspiral spin parameter $\chi_{\rm eff}$, and redshift $z_{\rm merge}$ with corresponding hyperparameters Λ_m, Λ_χ, and Λ_z. For the mass distribution $p(m_1,\ m_2|\Lambda_m)$, we adopt the slightly modified version of the POWER LAW+PEAK model (Talbot & Thrane 2018) from Fishbach & Fragione (2023); see their Equations (A2)–(A6). For the $\chi_{\rm eff}$ distribution $p(\chi_{\rm eff}|\Lambda_\chi)$, we assume the TRUNCATED GAUSSIAN model, truncated to the physical range $-1 < \chi_{\rm eff} < 1$ (Roulet & Zaldarriaga 2019; Miller et al. 2020). The merger redshift distribution is given by the formation rate $R_f(z_{\rm form})$ of Equation (6) convolved with an assumed delay time distribution $p_d(\tau)$ of Equation (5). Given hyperparameters Λ_z, the merger rate as a function of redshift is (see Equation (3)):

$$R_m(z_{\rm merge}|\Lambda_z) = \int_{\tau_{\rm min}}^{\tau_{\rm max}} R_f(z_{\rm form}(z_{\rm merge},\ \tau)|\Lambda_z) p_d(\tau) d\tau, \quad \text{(A1)}$$

where the formation redshift corresponding to a given merger redshift and delay time, $z_{\rm form}(z_{\rm merge},\ \tau)$, is given by converting the lookback time $t_f = t_{\rm merge} + \tau$. Fixing the delay time distribution $p(\tau)$, Equation (A1) depends on the hyperparameters $\Lambda_z = \{\mathcal{N},\ a,\ b\ \}$ from Equation (6). The merger rate is defined as the differential number density of mergers per comoving volume and source-frame time, $R_m \equiv dN/dV_c dt_s$. Converting this to a probability density over merger redshifts, i.e., the differential BBH number density per merger redshift and detector-frame time, we have:

$$\frac{dN}{dz_{\rm merge}}(\Lambda_z) = N \frac{R_m(z_{\rm merge}|\Lambda_z)\frac{dV_c}{dz_{\rm merge}}(1+z_{\rm merge})^{-1}}{\int_0^{z_{\rm merge}^{\rm max}} R_m(z_{\rm merge}|\Lambda_z)\frac{dV_c}{dz_{\rm merge}}(1+z_{\rm merge})^{-1} dz_{\rm merge}}. \quad \text{(A2)}$$

The normalization N in the above equation represents the total number of BBH sources that merge between redshift 0 and $z_{\rm merge}^{\rm max}$ in a given observation time. We are free to choose this maximum merger redshift as long as it is higher than the BBH horizon for GWTC-3; we choose $z_{\rm merge}^{\rm max} = 3$. Specifying this normalization N is equivalent to specifying the amplitude $\mathcal{N}$ of the formation rate $R_f(z_{\rm form})$ of Equation (6). For consistency with other GW population analyses, we sample over N in our population likelihood with a flat-in-log prior.

The full population distribution in terms of source parameters $\theta = \{m_1,\ m_2,\ \chi_{\rm eff},\ z_{\rm merge}\}$ and hyperparamters $\Lambda = \{\Lambda_m,\ \Lambda_\chi,\ \Lambda_z,\ N\}$ is:

$$\frac{dN}{d\theta}(\Lambda) = \frac{dN}{dz_{\rm merge}}(\Lambda_z) p(m_1,\ m_2|\Lambda_m) p(\chi_{\rm eff}|\Lambda_\chi), \quad \text{(A3)}$$

where $p(\ldots)$ denotes a normalized probability distribution.

Table 1
Summary of Hyperparameters Λ Describing the Phenomenological BBH Population Model in Equation (A3)

BBH Hyperparameter	Description	Prior
$m_{\rm min}/M_\odot$	Low-mass end of the primary mass spectrum	U(3, 12)
$m_{\rm max}/M_\odot$	High-mass end of the primary mass spectrum	U(30, 80)
η_h	Smoothing parameter for the low-mass end of the primary mass spectrum	U(2, 20)
η_l	Smoothing parameter for the high-mass end of the primary mass spectrum	U(2, 20)
α	Power-law slope of the primary mass distribution	U(−6, −0.5)
$f_{\rm peak}$	Height of the Gaussian peak in the primary mass distribution	U(0.0001, 0.3)
$m_{\rm peak}/M_\odot$	Location of the Gaussian peak	U(25, 50)
$w_{\rm peak}/M_\odot$	Width of the Gaussian peak	U(2, 8)
γ	Power-law slope of the secondary mass distribution	U(−2, 6)
μ	Center of the $\chi_{\rm eff}$ distribution	U(−0.5, 0.5)
σ	Width of the $\chi_{\rm eff}$ distribution	U(0.03, 0.5)
a	Power-law slope in $T(z_{\rm form})$ of the formation rate evolution	U(0.3, 2.6)
$b/{\rm Gyr}^{-1}$	Characteristic timescale for the exponential decline in $T(z_{\rm form})$	U(0.0001, 0.001)
$\log N$	Normalization constant; the total number of BBH mergers at $0 < z < 3$	U(7, 13)

Note. The notation U(a, b) denotes a uniform distribution between a and b.

We use an inhomogeneous Poisson process to model the likelihood for the GWTC-3 data consisting of $N_{\rm obs}$ independent observations $x = \{x_i\}_{i=1}^{N_{\rm obs}}$ given hyperparameters Λ, marginalizing over the source parameters θ (see, e.g., Loredo 2004; Mandel et al. 2019; Thrane & Talbot 2019; Vitale et al. 2022, for reviews):

$$p(x|\Lambda) = \exp\left(-\int \frac{dN}{d\theta}(\Lambda) P_{\rm det}(\theta) d\theta\right) \prod_{i=1}^{N_{\rm obs}} \int p(x_i|\theta) \frac{dN}{d\theta}(\Lambda) d\theta. \quad \text{(A4)}$$

We approximate both integrals in this likelihood with importance sampling (Monte Carlo) averages. For the first integral over $P_{\rm det}(\theta)d\theta$, we use the simulated signals from the GWTC-3 sensitivity estimates (LIGO Scientific Collaboration et al. 2021a), applying the appropriate weights to account for the simulated draw probability (Tiwari 2018; Essick & Farr 2022). For the second integral over $p(x_i|\theta)d\theta$, we use the same GWTC-1, GWTC-2, GWTC-2.1, and GWTC-3 parameter estimation samples (Abbott et al. 2021b; LIGO Scientific Collaboration 2021b, 2021c) used in the GWTC-3 population analysis (Abbott et al. 2023), applying the appropriate weights to account for the parameter estimation prior (Callister 2021).

Recall that there is an additional integral in the likelihood of Equation (A4), because the population model $dN/d\theta$ is defined in terms of an integral over the delay time distribution τ (Equation (A1)). However, $dN/d\theta$ always appears in the likelihood within another integral over θ, either with $p(x|\theta)$ or $P_{\rm det}(\theta)$. For computational efficiency, we approximate both the integral over τ and the integral over θ with a single Monte Carlo integral over the m_1, m_2, $\chi_{\rm eff}$, and $z_{\rm form}(z_{\rm merge}, \tau)$ samples. The $z_{\rm form}(z_{\rm merge}, \tau)$ samples are drawn by assigning a τ sample drawn from $p(\tau)$ to each of the $z_{\rm merge}$ samples.

The posterior for the hyperparameters Λ is given by the likelihood of Equation (A4) with a choice of prior. We choose broad priors over all hyperparameters, given in Table 1. We sample from the posterior with NUMPYRO (Bingham et al. 2019; Phan et al. 2019).

ORCID iDs

Maya Fishbach https://orcid.org/0000-0002-1980-5293
Lieke van Son https://orcid.org/0000-0001-5484-4987

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, PhRvL, 116, 131102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2018, LRR, 21, 3
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019a, PhRvX, 9, 031040
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019b, ApJL, 882, L24
Abbott, R., Abbott, T. D., Abraham, S., et al. 2021a, PhRvX, 11, 021053
Abbott, R., Abbott, T. D., Abraham, S., et al. 2021b, SoftX, 13, 100658
Abbott, R., Abbott, T. D., Acernese, F., et al. 2023, PhRvX, 13, 011048
Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, CQGra, 32, 024001
Adhikari, S., Fishbach, M., Holz, D. E., Wechsler, R. H., & Fang, Z. 2020, ApJ, 905, 21
Akutsu, T., Ando, M., Arai, K., et al. 2021, PTEP, 2021, 05A101
Andrews, J. J., Cronin, J., Kalogera, V., Berry, C. P. L., & Zezas, A. 2021, ApJL, 914, L32
Andrews, J. J., Zezas, A., & Fragos, T. 2018, ApJS, 237, 1
Antonini, F., Toonen, S., & Hamers, A. S. 2017, ApJ, 841, 77
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bavera, S. S., Fishbach, M., Zevin, M., Zapartas, E., & Fragos, T. 2022, A&A, 665, A59
Bavera, S. S., Fragos, T., Zevin, M., et al. 2021, A&A, 647, A153
Belczynski, K., Doctor, Z., Zevin, M., et al. 2022, ApJ, 935, 126
Belczynski, K., Dominik, M., Bulik, T., et al. 2010, ApJL, 715, L138
Belczynski, K., Holz, D. E., Bulik, T., & O'Shaughnessy, R. 2016, Natur, 534, 512
Bingham, E., Chen, J. P., Jankowiak, M., et al. 2019, J. Mach. Learn. Res., 20, 1, http://jmlr.org/papers/v20/18-403.html
Boco, L., Lapi, A., Chruslinska, M., et al. 2021, ApJ, 907, 110
Borhanian, S., & Sathyaprakash, B. S. 2022, arXiv:2202.11048
Bouffanais, Y., Mapelli, M., Santoliquido, F., et al. 2021, MNRAS, 507, 5224
Bradbury, J., Frostig, R., Hawkins, P., et al., 2018 JAX: Composable Transformations of Python+NumPy Programs, v0.3.13,, http://github.com/google/jax
Briel, M. M., Stevance, H. F., & Eldridge, J. J. 2023, MNRAS, 520, 5724
Broekgaarden, F. S., Berger, E., Stevenson, S., et al. 2022, MNRAS, 516, 5737
Callister, T., Fishbach, M., Holz, D. E., & Farr, W. M. 2020, ApJL, 896, L32
Callister, T. A. 2021, arXiv:2104.09508
Callister, T. A., & Farr, W. M. 2023, arXiv:2302.07289
Chruślińska, M. 2022, arXiv:2206.10622
Chruślińska, M., Jeřábková, T., Nelemans, G., & Yan, Z. 2020, A&A, 636, A10
Chruślińska, M., Nelemans, G., Boco, L., & Lapi, A. 2021, MNRAS, 508, 4994
Di Carlo, U. N., Mapelli, M., Giacobbo, N., et al. 2020, MNRAS, 498, 495
Dominik, M., Belczynski, K., Fryer, C., et al. 2013, ApJ, 779, 72

du Buisson, L., Marchant, P., Podsiadlowski, P., et al. 2020, MNRAS, 499, 5941
Edelman, B., Farr, B., & Doctor, Z. 2023, ApJ, 946, 16
Essick, R., & Farr, W. 2021, LIGO Document T1900895-v2, https://dcc.ligo.org/LIGO-T1900895/public
Essick, R., & Farr, W. 2022, arXiv:2204.00461
Evans, M., Adhikari, R. X., Afle, C., et al. 2021, arXiv:2109.09882
Fishbach, M., Doctor, Z., Callister, T., et al. 2021, ApJ, 912, 98
Fishbach, M., Farr, W. M., & Holz, D. E. 2020, ApJL, 891, L31
Fishbach, M., & Fragione, G. 2023, MNRAS, 522, 5546
Fishbach, M., Holz, D. E., & Farr, W. M. 2018, ApJL, 863, L41
Fishbach, M., & Kalogera, V. 2021, ApJL, 914, L30
Fishbach, M., & van Son, L. 2023, Data Release: LIGO-Virgo-KAGRA's Oldest Black Holes: Probing star formation at cosmic noon with GWTC-3," v1, Zenodo, doi:10.5281/zenodo.8256745
Galaudage, S., Talbot, C., & Thrane, E. 2020, PhRvD, 102, 083026
Gallegos-Garcia, M., Berry, C. P. L., Marchant, P., & Kalogera, V. 2021, ApJ, 922, 110
Godfrey, J., Edelman, B., & Farr, B. 2023, arXiv:2304.01288
Hall, E. D., & Evans, M. 2019, CQGra, 36, 225002
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hopkins, A. M. 2018, PASA, 35, e039
Hunter, J. D. 2007, CSE, 9, 90
Iben, I. J., & Tutukov, A. V. 1984, ApJS, 54, 335
Iorio, G., Mapelli, M., Costa, G., et al. 2023, MNRAS, 524, 426
Kalogera, V., Sathyaprakash, B. S., Bailes, M., et al. 2021, arXiv:2111.06990
Karathanasis, C., Mukherjee, S., & Mastrogiovanni, S. 2023, MNRAS, 523, 4539
Katsianis, A., Yang, X., & Zheng, X. 2021, ApJ, 919, 88
Klencki, J., Moe, M., Gladysz, W., et al. 2018, A&A, 619, A77
Kluyver, T., Ragan-Kelley, B., Pérez, F., et al. 2016, in Positioning and Power in Academic Publishing: Players, Agents and Agendas, ed. F. Loizides & B. Scmidt (Netherlands: IOS Press) https://eprints.soton.ac.uk/403913/
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P., Weidner, C., Pflamm-Altenburg, J., et al. 2013, in Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations, ed. T. D. Oswalt & G. Gilmore (Dordrecht: Springer), 115
Lamberts, A., Garrison-Kimmel, S., Clausen, D. R., & Hopkins, P. F. 2016, MNRAS, 463, L31
LIGO Scientific Collaboration, Aasi, J., Abbott, B. P., et al. 2015, CQGra, 32, 074001
LIGO Scientific CollaborationVirgo Collaboration and KAGRA Collaboration 2021a, GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run ? O1+O2+O3 Search Sensitivity Estimates, v1, Zenodo, doi:10.5281/zenodo.5636816
LIGO Scientific CollaborationVirgo Collaboration and KAGRA Collaboration 2021b, GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run—Parameter Estimation Data Release, v1, Zenodo, doi:10.5281/zenodo.5117703
LIGO Scientific CollaborationVirgo Collaboration and KAGRA Collaboration 2021c, GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run — Parameter estimation data release, v1, Zenodo, doi:10.5281/zenodo.5546663
Loredo, T. J. 2004, in AIP Conf. Ser. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th Int. Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint (Melville, NY: AIP), 195
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Madau, P., & Fragos, T. 2017, ApJ, 840, 39
Maggiore, M., Van Den Broeck, C., Bartolo, N., et al. 2020, JCAP, 2020, 050
Maiolino, R., & Mannucci, F. 2019, A&ARv, 27, 3
Mandel, I., & Farmer, A. 2022, PhR, 955, 1
Mandel, I., Farr, W. M., & Gair, J. R. 2019, MNRAS, 486, 1086
Mapelli, M. 2020, FrASS, 7, 38
Mapelli, M., Giacobbo, N., Ripamonti, E., & Spera, M. 2017, MNRAS, 472, 2422
Mapelli, M., Giacobbo, N., Santoliquido, F., & Artale, M. C. 2019, MNRAS, 487, 2
Mapelli, M., Giacobbo, N., Toffano, M., et al. 2018, MNRAS, 481, 5324
Marchant, P., Langer, N., Podsiadlowski, P., Tauris, T. M., & Moriya, T. J. 2016, A&A, 588, A50
Marchant, P., Pappas, K. M. W., Gallegos-Garcia, M., et al. 2021, A&A, 650, A107
Michaely, E., & Naoz, S. 2022, ApJ, 936, 184
Miller, J., Barsotti, L., Vitale, S., et al. 2015, PhRvD, 91, 062005
Miller, S., Callister, T. A., & Farr, W. M. 2020, ApJ, 895, 128
Moore, C. J., & Gerosa, D. 2021, PhRvD, 104, 083008
Mukherjee, S., & Dizgah, A. M. 2022, ApJL, 937, L27
Naidu, R. P., Ji, A. P., Conroy, C., et al. 2022, ApJL, 926, L36
Nakar, E. 2007, PhR, 442, 166
Narayanan, D., Lower, S., Torrey, P., et al. 2023, arXiv:2306.10118
Neijssel, C. J., Vigna-Gómez, A., Stevenson, S., et al. 2019, MNRAS, 490, 3740
Ng, K. K. Y., Franciolini, G., Berti, E., et al. 2022, ApJL, 933, L41
Ng, K. K. Y., Vitale, S., Farr, W. M., & Rodriguez, C. L. 2021, ApJL, 913, L5
Nitz, A. H., Capano, C. D., Kumar, S., et al. 2021, ApJ, 922, 76
Olsen, S., Venumadhav, T., Mushkin, J., et al. 2022, PhRvD, 106, 043009
Peters, P. C. 1964, PhRv, 136, 1224
Phan, D., Pradhan, N., & Jankowiak, M. 2019, arXiv:1912.11554
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123
Riley, J., Agrawal, P., Barrett, J. W., et al. 2022, ApJS, 258, 34
Riley, J., & Mandel, I. 2023, ApJ, 950, 80
Riley, J., Mandel, I., Marchant, P., et al. 2021, MNRAS, 505, 663
Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016, PhRvD, 93, 084029
Roulet, J., & Zaldarriaga, M. 2019, MNRAS, 484, 4216
Safarzadeh, M., Berger, E., Ng, K. K. Y., et al. 2019, ApJL, 878, L13
Santoliquido, F., Mapelli, M., Giacobbo, N., Bouffanais, Y., & Artale, M. C. 2021, MNRAS, 502, 4877
Smith, R. J. 2020, ARA&A, 58, 577
Stevenson, S., & Clarke, T. A. 2022, MNRAS, 517, 4034
Talbot, C., & Thrane, E. 2018, ApJ, 856, 173
The LIGO Scientific Collaborationthe Virgo Collaborationthe KAGRA Collaboration et al. 2021, arXiv:2111.03606
Thrane, E., & Talbot, C. 2019, PASA, 36, e010
Tiwari, V. 2018, CQGra, 35, 145009
van Son, L., de Mink, S., Renzo, M., et al. 2022a, Dataset 1of 2: No peaks without valleys: The stable mass transfer channel for gravitational- wave sources in light of the neutron star-black hole mass gap., v02.26.03, Zenodo, doi:10.5281/zenodo.7080725
van Son, L., de Mink, S., Renzo, M., et al. 2022b, Dataset 2 of 2: No peaks without valleys: The stable mass transfer channel for gravitational- wave sources in light of the neutron star-black hole mass gap., v02.26.03, Zenodo, doi:10.5281/zenodo.7080164
van Son, L. A. C., de Mink, S. E., Chruślińska, M., et al. 2023, ApJ, 948, 105
van Son, L. A. C., de Mink, S. E., Renzo, M., et al. 2022a, ApJ, 940, 184
van Son, L. A. C., de Mink, S. E., Callister, T., et al. 2022b, ApJ, 931, 17
Vink, J. S., & de Koter, A. 2005, A&A, 442, 587
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Vitale, S., Farr, W. M., Ng, K. K. Y., & Rodriguez, C. L. 2019, ApJL, 886, L1
Vitale, S., Gerosa, D., Farr, W. M., & Taylor, S. R. 2022, Handbook of Gravitational Wave Astronomy (Singapore: Springer), 45
Webbink, R. F. 1984, ApJ, 277, 355
Weizmann Kiendrebeogo, R., Farah, A. M., Foley, E. M., et al. 2023, arXiv:2306.09234
Wong, K. W. K., Breivik, K., Farr, W. M., & Luger, R. 2023, ApJ, 950, 181
Yang, Y., Bartos, I., Haiman, Z., et al. 2020, ApJ, 896, 138
Zevin, M., Bavera, S. S., Berry, C. P. L., et al. 2021, ApJ, 910, 152
Zevin, M., Nugent, A. E., Adhikari, S., et al. 2022, ApJL, 940, L18