# Measuring the Star Formation Rate with Gravitational Waves from Binary Black Holes

Salvatore Vitale[1,2] , Will M. Farr[3,4,5] , Ken K. Y. Ng[1,2] , and Carl L. Rodriguez[2,6]

[1] LIGO, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; salvatore.vitale@ligo.org
[2] Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3] Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA; will.farr@stonybrook.edu
[4] Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA
[5] Birmingham Institute for Gravitational Wave Astronomy, University of Birmingham, Birmingham, B15 2TT, UK
[6] Harvard Institute for Theory and Computation, 60 Garden Street, Cambridge, MA 02138, USA

## Abstract

A measurement of the history of cosmic star formation is central to understanding the origin and evolution of galaxies. The measurement is extremely challenging using electromagnetic radiation: significant modeling is required to convert luminosity to mass, and to properly account for dust attenuation, for example. Here we show how detections of gravitational waves from inspiraling binary black holes made by proposed third-generation detectors can be used to measure the star formation rate (SFR) of massive stars with high precision up to redshifts of $\sim$10. Depending on the time-delay model, the predicted detection rates ranges from $\sim$2310 to $\sim$56,740 per month with the current measurement of local merger rate density. With 30,000 detections, parameters describing the volumetric SFR can be constrained at the few percent level, and the volumetric merger rate can be directly measured to 3% at $z \sim 2$. Given a parameterized SFR, the characteristic delay time between binary formation and merger can be measured to $\sim$60%.

*Unified Astronomy Thesaurus concepts:* Gravitational waves (678); Gravitational wave sources (677); Gravitational wave astronomy (675); Gravitational wave detectors (676); Star formation (1569); Metallicity (1031); Astrophysical black holes (98); Stellar mass black holes (1611)

## 1. Introduction

The binary black holes (BBHs) detected by the ground-based gravitational-wave (GW) detectors LIGO (Harry 2010) and Virgo (Acernese et al. 2015) all merged in the local universe (Abbott et al. 2016b, 2016c, 2016d, 2017a, 2017b, 2017c; The LIGO Scientific Collaboration et al. 2019). These detections have allowed us to measure the *local* merger rate of BBHs at [24.4–111.7] Gpc$^{-3}$ yr$^{-1}$ (90% credible interval; The LIGO Scientific Collaboration et al. 2019). The sensitivity of advanced detectors limits to $z \sim 1$ the maximum redshift at which a heavy BBH, with total mass of about 60 $M_\odot$, such as GW150914 can be detected, whereas heavier systems, including intermediate-mass black hole binaries, could be observed farther away (Abbott et al. 2016b, 2016c, 2016d, 2017a, 2017b, 2017c, 2017d; Chen et al. 2017; The LIGO Scientific Collaboration et al. 2018, 2019; Hall & Evans 2019).

As the LIGO and Virgo instruments progress toward their design sensitivity (Abbott et al. 2016e), and the network of ground-based detectors grows, it will be possible to detect BBHs at redshifts greater than 1 (the exact value depending on the BBH mass). This can potentially allow us to probe the merger rate of BBHs through a significant distance range, and check how it varies with redshift (Fishbach et al. 2018).

While this might provide precious information on the evolution of the merger rate, it would be interesting to access sources at even higher redshifts. Since compact binaries are constituted of neutron stars and black holes, leftovers of main-sequence stars, a measurement of their abundance at different stages of cosmic history can potentially tell us something about the star formation rate (SFR). This latter is currently measured using various electromagnetic probes (Behroozi et al. 2013;

Madau & Dickinson 2014). However, electromagnetic probes do not directly track the amount of matter being formed on a galaxy. Instead, they track the luminosity, which then is linked to the mass production through several steps of modeling (e.g., on the initial mass function). Furthermore, dust extinction can significantly reduce the bolometric luminosity of a galaxy, or alter the its spectral content, which is a key ingredient to infer the SFR from light. These limitations are particularly severe at redshifts above 3 where, additionally, fewer data points are available from electromagnetic observations.[7] It would thus be valuable to have an independent way of measuring the star formation at high redshifts, possibly by directly tracking masses, rather than light. Gravitational-wave signals can be used for that goal, as they directly encode information about the mass of the source. Two proposals for third-generation (3G) ground-based detectors are currently being pursued, which would allow us to detect BBHs at large redshifts: the Einstein Telescope (ET; Punturo et al. 2010) and Cosmic Explorer (CE; Abbott et al. 2017e). Using the local merger rate calculated by the LIGO and Virgo collaborations it has been estimated that [1–40] $\times$ 10$^4$ BBHs merge in the universe per year (Regimbau et al. 2017). Vitale & Evans (2017) have shown how BBHs can be detected all the way to redshift of $\sim$15 by networks of 3G detectors. Since that is a significant fraction of the volume of the universe, one would thus expect that a large fraction of merging BBHs would be detectable. Indeed, Regimbau et al. (2017) estimate that 99.9% of the BBH mergers will be detectable by 3G detectors.[8] In this Letter we show how, under quite generic hypotheses, accessing BBHs with 3G

---

[7] We notice that this might become less true as future telescope get online, in the timescale relevant for the realization of third-generation GW detectors.
[8] In this Letter we solely focus on BBHs. Previous work exists for binary neutron stars (Van Den Broeck 2010; Safarzadeh et al. 2019).

gravitational-wave detectors allows for a direct inference of the merger rate and the SFR all the way to redshifts of $\sim 10$.

## 2. Event Rates

As sources are detected in a gravitational-wave detector network, one can estimate their redshifts (Veitch et al. 2015; Farr et al. 2016; Vitale & Evans 2017) and measure their detection rate in the local frame. Let[9] $R_m(z_m) \equiv \frac{dN_m}{dt_d dz}$ be the total redshift rate density of mergers in the detector frame (the number of mergers per detector time per redshift). The shape of this function, given the uncertainty in the observed redshift of the detected sources, can be inferred with hierarchical analysis (Hogg et al. 2010; Mandel 2010; Farr et al. 2011; Youdin 2011).

The redshift rate density can be written in terms of the volumetric total merger rate in the source frame $\mathcal{R}_m(z_m) \equiv \frac{dN_m}{dV_c dt_s}$ as

$$R_m(z_m) = \frac{1}{1 + z_m} \frac{dV_c}{dz} \mathcal{R}_m(z_m), \tag{1}$$

where the $1 + z_m$ term arises from converting source-frame time to detector-frame time (Dominik et al. 2013).

The volumetric merger rate in galactic fields depends on the SFR, the metallicity, and the delay between the formation of the BBH progenitors and their eventual merger. All the systems that merge at a lookback time $t_m$ (or, equivalently, at a redshift $z_m = z(t_m)$) are systems that formed at $z_f > z_m$ (or $t_f > t_m$). The delay time distribution, $p(t_m|t_f, \lambda)$, is the probability density that a system that formed at time $t_f$ will merge at time $t_m$. This function may depend on an (unknown) timescale, the parameters of the system that is merging, and possibly other parameters. We capture this dependence using parameters $\lambda$.

We can write the merger rate at redshift $z_m$ as a function of the black hole binary volumetric formation rate, $\mathcal{R}_f(z_f)$:

$$\mathcal{R}_m(z_m) = \int_{z_m}^{\infty} dz_f \frac{dt_f}{dz_f} \mathcal{R}_f(z_f) p(t_m|t_f, \lambda). \tag{2}$$

Here we assume that volumetric formation rate $\mathcal{R}_f(z_f)$ is simply proportional to the SFR density at the same redshift, $\psi(z)$ (see Equation (5)) and to the efficiency $\eta(z)$:

$$\mathcal{R}_f(z_f) \equiv \frac{dN_{\text{form}}}{dV_C dt_f} \propto \eta(z_f) \psi(z_f). \tag{3}$$

The fact that the merger rate is proportional to the SFR *at the same redshift* is a reasonable assumption (Madau & Dickinson 2014; Abbott et al. 2016a), since the lifetime of massive stars that will become black holes is of the order of tens of Myr and hence negligible when compared to the other timescales of interest. The efficiency $\eta(z)$ takes into account the fact that of the star formation at a given redshift, only the fraction $\eta(z)$ with low metallicity will result in heavy black hole formation. Following Belczynski et al. (2017) we define $\eta(z)$ as the fraction of star formation that has metallicity below 10% of the solar metallicity, and calculate it as

$$\eta(z) \equiv \int_{-\infty}^{\log(0.1 Z_{\odot})} \Phi(\log Z_{\text{mean}}(z), 0.5) \, dZ, \tag{4}$$

where $\Phi$ is the cumulative distribution function of the metallicity at redshift $z$, assumed to be a Gaussian distribution with mean $Z_{\text{mean}}$ and 0.5 dex of uncertainty (Equation (2) of Belczynski et al. 2017).

We do not account here for eventual contributions to the formation rate arising from binaries that do not form in galactic fields (e.g., binaries from globular clusters or from Population III stars). The methods we use can be extended to account for multiple formation channels; we discuss this possibility further below.

Both the formation rate and the time-delay distribution might depend on some intrinsic properties of the binary being formed, e.g., the component masses (Dominik et al. 2013). These dependencies can be included in an extension of our analysis in a straightforward manner, by adding the masses and other parameters to $\lambda$ and marginalizing them in Equation (2). However, for this proof-of-principle study we will assume these details can be neglected.

In this work we will follow two different approaches. First, we will assume that nothing is known about the true functional form of the SFR and the time-delay distribution. In this case, we use a nonparametric Gaussian-process algorithm to directly measure the volumetric rate density in the detector frame, $\mathcal{R}_m(z)$. Next, we will show that assuming the parameterized functional form of both the SFR and the time-delay distribution, the parameters on which they depend can be measured from the GW detections.

## 3. Simulated Signals

To demonstrate how the cosmic BBH merger rate can be measured, we generate 30,000 synthetic BBH detections in each time-delay model with realistic redshift uncertainty (see below; Vitale & Evans 2017). We assume that the SFR is the Madau–Dickinson (MD) SFR, which can be written

$$\psi_{\text{MD}}(z) = \psi_0 \frac{(1 + z)^{\alpha}}{1 + \left(\frac{1+z}{C}\right)^{\beta}}, \tag{5}$$

with parameters $\alpha = 2.7$, $\beta = 5.6$, $C = 2.9$, and $\psi_0 = 0.015 \, M_{\odot} \, \text{Mpc}^{-3} \, \text{yr}^{-1}$ (Madau & Dickinson 2014). The proportionality coefficient in Equation (3) is chosen such that the local BBH merger rate $R_m(0)$ is equal to 50 Gpc$^{-3}$ yr$^{-1}$, consistent with LIGO and Virgo's measurements. We notice that, since the SFR also affects the mean metallicity, and hence the efficiency $\eta(z)$ (Belczynski et al. 2017), the local merger rate is not simply proportional to $\psi_0$. This is shown if Figure 1. Different panels show how $\eta(z)$ varies when the parameters describing the MD SFR are varied, one at the time. The range of variability is taken to be representative of the uncertainties we find in their measurement in Section 4. We see that $\psi_0$ and $\alpha$ have similar effects in the efficiency, and we should thus expect them to be anticorrelated. C and $\beta$ instead have a milder effect on $\eta(z)$.

We consider two different functional forms for the distribution of time delays between formation and merger: an exponential function with timescale parameter $\tau$,

$$p(t_m|t_f, \tau) = \frac{1}{\tau} \exp\left\{\left[-\frac{(t_f - t_m)}{\tau}\right]\right\}, \tag{6}$$

---

[9] We will use the subscript "f" for quantities related to the formation of binaries, and "m" for quantities related to their merger.
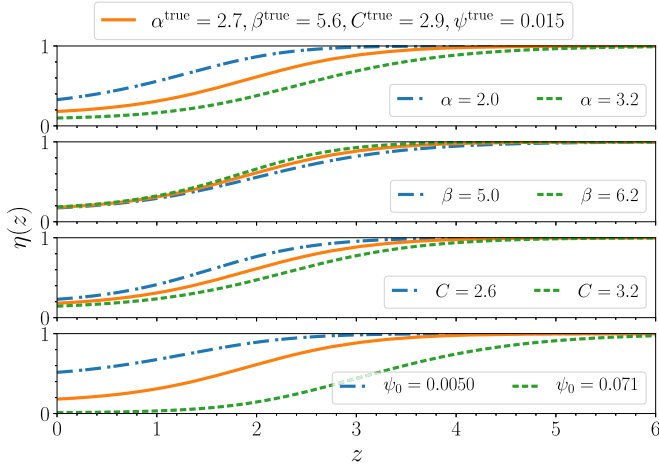
**Figure 1.** Efficiency $\eta(z)$ plotted as function of the redshift. In each panel, the orange curve is obtained using the nominal MD SFR. The other curves are obtained by varying, in turn, each of the four parameters controlling the MD SFR. The range of variation is taken to be representative of the uncertainties we find in Section 4.

and a distribution uniform in the logarithm of the time delay,

$$p(\log(t_m - t_f)) \propto \begin{cases} 1 & 10\,\text{Myr} < t_m - t_f < 10\,\text{Gyr} \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

The true redshifts of the sources under both delay assumptions are randomly drawn from Equation (1), after normalizing it to unity in the redshift range $z \in [0, 15]$.

In Figure 2 we show the redshift distribution of the simulated BBH merger events using the exponential time delay with $\tau = 0.1, 1, 10$ Gyr,[10] and with the flat-in-log distribution at a fixed local merger rate density of 50 Gpc$^{-3}$ yr$^{-1}$. The estimated number of events in one month is $M = 56{,}740,\ 33{,}900,\ 2310,$ and 25,980, respectively. The corresponding time to observe 30,000 events is $T = 16$ days, 27 days, 13 months, and 35 days, respectively.

The redshift of detected BBHs cannot be perfectly measured using GW detectors. We approximate the results of a full analysis of a three-detector 3G network (Vitale & Evans 2017) by assuming that the likelihood function for the true redshift follows a log-normal distribution conditioned on the true redshift with standard deviation $\sigma_{LN}$ $(z_{\text{true}}) = 0.017 z_{\text{true}} + 0.012$.

We do not explicitly draw mass values or calculate a signal-to-noise ratio. As long as one works with BBHs of total mass above $\sim 15\,M_\odot$, all sources are detectable by 3G networks including the CE up to redshifts where the merger rate becomes negligible (Regimbau et al. 2017; Vitale & Evans 2017).

Once the catalog of simulated events and the corresponding redshift likelihoods have been generated, our analysis proceeds hierarchically (Hogg et al. 2010; Mandel 2010; Farr et al. 2011; Youdin 2011). We assume that the production of gravitational-wave sources is an (inhomogeneous) Poisson process, with rate density

$$\mathcal{R}_m(z|\lambda),$$



**Figure 2.** Merger redshift distribution of the simulated population of BBHs. We assume a Madau–Dickinson SFR, and four different prescriptions for the time delay between formation and merger: an exponential time delay with e-fold times of 100 Myr, 1, and 10 Gyr, and a uniform-in-log distribution, with a minimum of 10 Myr and a maximum of 10 Gyr.

depending on some parameters $\lambda$. Therefore, the posterior for the population-level parameters given (synthetic) data for 30,000 events, $\boldsymbol{d} \equiv \{d_i\}_{i=1}^M$, is (Youdin 2011; Foreman-Mackey et al. 2014; Farr et al. 2015)

$$p(\lambda|\boldsymbol{d}) \propto \left[\prod_{i=1}^M \int dz_i\, p(d_i|z_i) R_m(z_i|\lambda)\right] e^{-\chi}\, p(\lambda)$$

$$\simeq \left[\prod_{i=1}^M \frac{1}{M_i} \sum_{j=1}^{M_i} R_m(z_{ij}|\lambda)\right] e^{-\chi}\, p(\lambda), \quad (8)$$

where $\chi \equiv \int dz\, dt_d\, R_m(z|\lambda)$, $z_i$ is the redshift of event $i$, $p(\lambda)$ is a prior imposed on the parameters describing the merger rate density, and we use $M_i$ samples, $\{z_{ij}\}_{j=1}^{M_i}$, drawn from a density proportional to the likelihood, $z_{ij} \sim p(d_i|z_{ij})dz_{ij}$, to approximate the marginalization integral over $z_i$.

## 4. Results

We desire to understand how well we can expect to constrain the merger rate density and the time-delay distribution from our synthetic data set of 30,000 observations.

We first consider an unmodeled approach, where nothing is assumed about the underlying SFR function and time-delay distribution other than that it is relatively smooth (Foreman-Mackey et al. 2014). We assume that the log of the merger rate can be described by a piecewise-constant function over $K = 29$ redshift bins. To ensure there are enough samples in each bin, we choose the bins in the following way: $0 \leqslant z < 0.32$ for the first bin, while the remaining bins are uniformly distributed in $\log(1 + z)$ with $z \in [0.32, 15)$ so that the log of merger rate is

$$\log \mathcal{R}_m = \begin{cases} n_1 & 0 \leqslant z < z_1 \\ \cdots & \\ n_i & z_{i-1} \leqslant z < z_i \\ \cdots & \\ n_K & z_{K-1} \leqslant z < z_K, \end{cases} \quad (9)$$

and we treat the per-bin merger rates, $n_i$, as parameters, $\lambda$, in Equation (8). We apply a squared-exponential Gaussian-

---

[10] These three values, as well as the minimum and maximum time delay in Equation (7), are chosen to cover a reasonable range of characteristic time delays (Belczynski et al. 2001, 2006; Berger et al. 2007; Nakar 2007; Dominik et al. 2012, 2013).
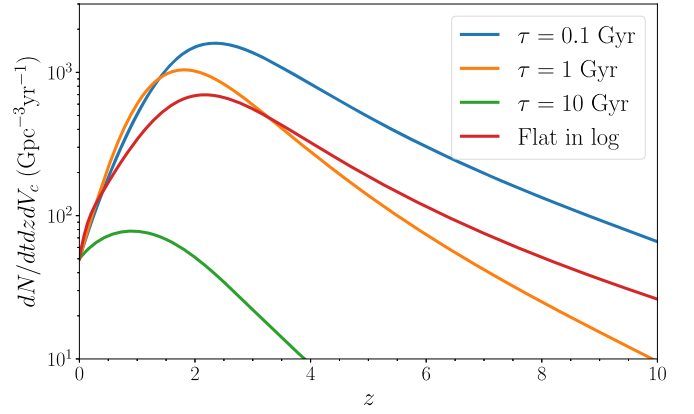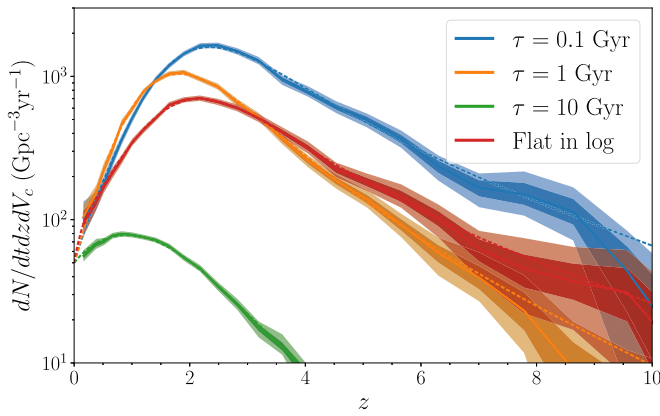
**Figure 3.** Posterior on the volumetric merger rate density calculated using an unmodeled approach. The dashed lines are the true rates under the four possible time-delay distributions we consider. Full lines give the median measurement, while the bands report the 68% and 95% credible intervals. Near the peak $z \sim 2$ the uncertainty in the rate estimate is $\sim$3% for $\tau = 0.1$, 1 Gyr and flat-in-log models. The uncertainty rises to 10% in the $\tau = 10$ Gyr model around the peak $z \sim 1$, as the total number of events is 10 times smaller than the numbers in other models. The small systematic offset for the flat-in-log and prompt data sets is likely due to a 100 Myr lower limit on the delay time imposed for numerical stability; see the corresponding discussion in the parameterized model results.

process prior on the $n_i$, which has a covariance kernel of

$$\mathrm{Cov}(n_i, n_j) = \sigma^2 \exp\left[-\frac{1}{2}\left(\frac{z_{i-1/2} - z_{j-1/2}}{l}\right)^2\right], \quad (10)$$

with $z_{i-1/2} = (z_i - z_{i-1})/2$ as the midpoint of the $i$th redshift bin. We treat the variance of the $n_i$, $\sigma^2$, and the correlation length in redshift space, $l$, as additional parameters in the fit. The squared-exponential Gaussian-process prior enforces the smoothness of the merger rate on scales that are comparable to or larger than $l$ (which may be much larger than the bin spacing if the data support it), and guards against overfitting when $K$ is large (Foreman-Mackey et al. 2014).

The results for this fit are shown in Figure 3, where for each true synthetic population we show the median posterior on the piecewise-constant $dN/dV_c dt_d$, together with 68% and 95% ($1\sigma$ and $2\sigma$) credible intervals. We see that the unmodeled Gaussian-process (GP) method pinpoints the merger rates so precisely that all four distributions are clearly distinguishable; near $z \sim 2$ the uncertainty in the measured merger rate is $\sim$3%. At moderate redshifts, $z < 4$, the uncertainties are smaller than the separation between different populations. At larger red-shifts, the measurement becomes more uncertain, and overlaps exist. This is due to a combination of two effects: from one side, fewer sources merge, and hence are detected, at those redshifts; from the other, the uncertainty in their measured redshift is higher. The advantage of this approach over a more rigid parameterization of the merger rate is that it can fit *any* sufficiently smooth merger rate; a disadvantage is that we learn nothing individually about the time-delay distribution or the SFR, since they are completely degenerate in this flexible model.

Next, we want to verify how well we can measure the characteristic parameters of the SFR and time-delay distribu-tion *assuming* we know their functional forms.

For this analysis, we take the MD SFR and the exponential time-delay distribution as models, treating the parameters $\alpha$, $\beta$, $C$, $\psi_0$, as well as the time-delay scale $\tau$ as unknowns. We then
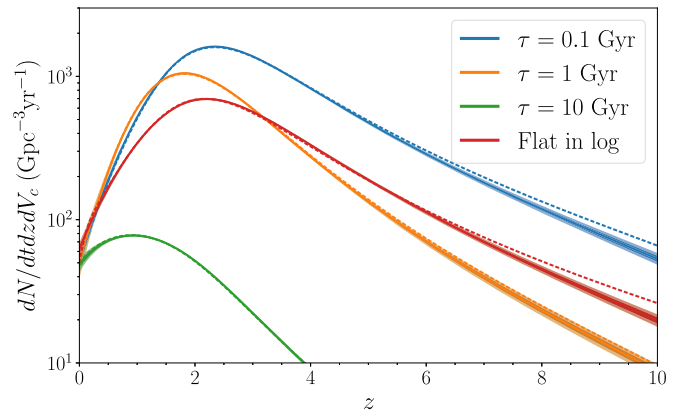


**Figure 4.** Posterior on the merger rate density calculated from the parameterized fits described in the text. Dashed lines show the true merger rate distributions for our models. Solid lines give the posterior median and dark and light bands the 68% and 95% credible intervals. See the text for more details.

calculate the posterior for $\lambda_{\mathrm{MD}} = \{\alpha, \beta, C, \psi_0, \tau\}$ with Equation (8). Note that the parameterized model with an exponential time delay cannot perfectly match the flat-in-log data-generating model, no matter what value of $\tau$ is used.

We use log-normal priors with a width of $\simeq$0.25 in the log for $\alpha$, $\beta$, and $C$, reflecting an approximation to the uncertainty in the determination of the SFR (Madau & Dickinson 2014). We also use a log-normal prior for $\psi_0$, with a prior large enough that the posterior is not truncated. For $\tau$, we use a width of 2 in log to cover the whole dynamical range from 0.1 to 10 Gyr. The uncertainties are large enough that the posterior distributions are not truncated by the prior; with 30,000 simulated detections we obtain meaningful constraints on the SFR parameters at the few percent level and the time delay at a few tens of percent in all models. We place a lower bound on the time-delay parameter $\tau \geqslant 100$ Myr in order to ensure numerical stability in our computation of the integral in Equation (2). This results in some discrepancy between the fit and the data-generating distribution for the "prompt" data set; the prompt data are recovered in the limit $\tau \to 0$, but as this is excluded by our prior there is a bias in the fit, particularly at high redshift where timescales of 100 Myr are a significant fraction of the age of the universe. The inferred posterior on the merger rate redshift density is shown in Figure 4. In Figure 5 we show posteriors for the parameters $\lambda_{\mathrm{MD}}$ for the set of events with $\tau = 1$ Gyr.

After 30,000 detections in the 1 Gyr scenario, the scale factor of the time-delay distribution can be measured with relative uncertainty of 60% (at the 90% credible interval): $\tau = 0.93^{+0.35}_{-0.31}$. The parameters controlling the peak and high-$z$ slope of the MD SFR can also be measured with precision of $\sim$20% or better, and we obtain $\beta = 5.57^{+0.59}_{-0.54}$ and $C = 2.90^{+0.23}_{-0.27}$. On the other hand, $\alpha = 2.54^{+0.68}_{-0.60}$ and $\log_{10}\psi_0 = -1.72^{+0.56}_{-0.60}$ are only marginally narrower than their priors. Many correlations are visible in Figure 5, which is worth discussing, as they arise from different astrophysical factors. First, $\tau$ and $C$ show a clear correlation, which can be understood as follows. If $C$ increases then the peak of the SFR moves to higher redshift. In order to keep the *observed* merger rate fixed the delay time must increase. On the other hand, $\psi_0$ and $\alpha$ are anticorrelated as they both affect the efficiency $\eta(z)$, and hence the merger rate, in a similar way. When $\psi_0$ increases, the total star formation at each redshift, as
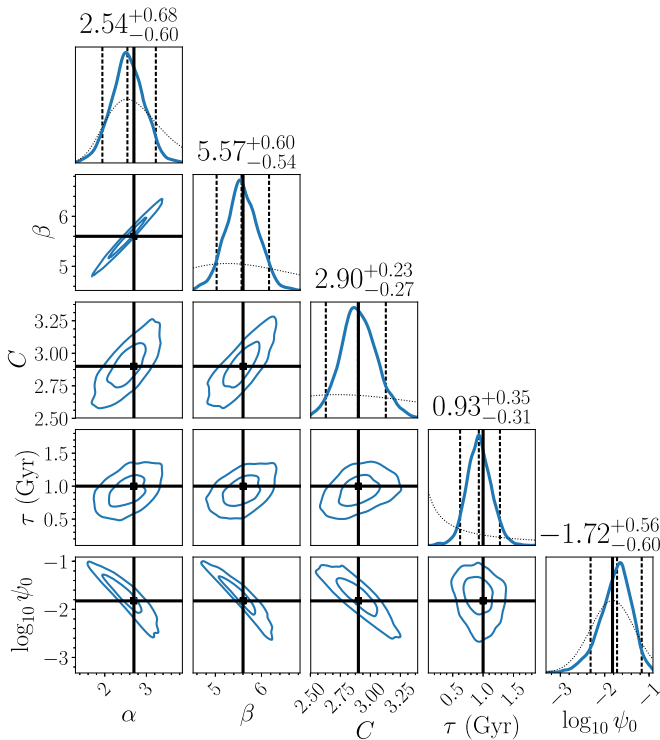
**Figure 5.** Posterior distribution for the time-delay timescale and the MD SFR parameters after 30,000 detections in the 1 Gyr delay timescale scenario. Truth is indicated by blue lines. $\psi_0$ is in units of $M_\odot$ Mpc$^{-3}$ yr$^{-1}$. Dashed lines indicate the highest posterior density 90% credible interval; star formation rate parameters are measured to few percent precision, and the delay timescale is measured to ~60%. Plot labels give the median and the highest posterior density 90% credible interval for each parameter.

well as the metallicity, increase, which reduces the overall efficiency; bottom panel of Figure 1. To compensate the loss of efficiency, $\alpha$ (and to a smaller extent, $C$) need to be decrease, as shown by the blue dotted–dashed curves in Figure 1. This explains the anticorrelations seen for the pairs ($\alpha$, $\psi_0$) and ($C$, $\psi_0$) in Figure 5. Finally, $\beta$ does not affect $\eta(z)$ much, Figure 1, and the correlation seen in Figure 5 for the pair ($\beta$, $\psi_0$) is really only a consequence of the fact that $\alpha$ and $\beta$ are strongly correlated.

The parameter recovery for the other scenarios is similar, but for the flat-in-log scenario the systematic bias from model mismatch is significantly larger than the statistical uncertainty. The parameter estimates obtained from all scenarios are given in Table 1. Determination of the time-delay distribution and the parameters of the SFR also allow measurement of the total number of BBH mergers per solar mass of star formation (not shown).

## 5. Discussion and Outlook

In this Letter we have shown how next-generation ground-based detectors will enable using gravitational waves from BBHs to infer their merger rate throughout cosmic history, even in absence of any model for the star formation history. On the other hand, if a modeled template is available for the SFR and for the time-delay distribution between formation and merger, we have shown how their characteristic parameters can be measured with 30,000 simulated signals.

We have simulated four different "Universes," assuming the BBH formation rate is proportional to the MD SFR. The

coefficient of proportionality is a redshift- and SFR-dependent function that accounts for the fraction of the SFR with metallicity below 10% of the solar metallicity (Belczynski et al. 2006). The four data sets use four different prescriptions for the delay between formation and merger: flat in the logarithm of the time delay, or exponential, with e-fold times of 0.1, 1, or 10 Gyr.

The unmodeled approach yields a direct measurement of the volumetric merger rate $\mathcal{R}_m \equiv dN/dV_c dt_d$. Figure 3 shows the measurement obtained with 30,000 simulated signals. The four models are clearly distinguishable, and have uncertainties much smaller than their separation for redshifts below ~6. At larger redshifts, the uncertainties increase due to the smaller number of sources, and the larger uncertainty on their redshifts.

Including a model for the star formation history and the time-delay distribution dramatically increases the power of the method, and the expense of its generality. Using the MD SFR, Equation (5), and an exponential time-delay distribution with an unknown e-fold time $\tau$ as templates, we have shown how all unknowns can be measured with good precision after 30,000 simulated signals. The measurement of the SFR parameters is not accurate for the universe with flat-in-log time delays, as one would have expected given the mismatch between the time-delay template and the actual time-delay distribution. This kind of issue can be mitigated using templates with more parameters. The number of parameters will increase the computational cost of the analysis, and the uncertainty in the measurement. However, the number of detectable BBHs is in the hundreds of thousand per year, which will compensate for the extra complexity of the model.

In this work we have made a few simplifying assumptions to keep the computational cost under control. First, we have assumed that the time-delay distribution is the same for all sources at all redshifts, while in reality it will depend on the redshift of the source through the metallicity of the environment (Chruslinska et al. 2019). This limitation can be lifted, introducing a functional form that relates time delay to redshift and possible other parameters, which will eventually be marginalized over. Relatedly, we have neglected the dependence of the SFR and time-delay distribution on the mass and spins of the sources. This is not an intrinsic limitation of the method, and can be easily folded in the analysis. As these extra parameters are accounted for, we would expect that more sources will be required to achieve the same precision. But, as mentioned above, in this work we have considered 30,000 simulated signals, which correspond to a few weeks to one year of observing time, depending on the actual time-delay distribution. More detections will be available for these tests, hence compensating for the increased complexity of the model.

Finally, while generating the simulated signals, we have assumed that all sources come from galactic fields. There is growing evidence that at least a fraction of BBHs detected by LIGO and Virgo have been formed in globular clusters (Rodriguez et al. 2015, 2016). These sources would show a very different evolution with redshift, with a peak of the merger rate at higher redshift. If black holes from Population III stars merge, they could also contribute to the total merger rate, probably with a peak above $z \sim 10$ (Belczynski et al. 2017; Kinugawa et al. 2016). Depending on the relative abundance of mergers in these channels, one could be able to calculate their branching ratios as a function of redshift. This would give information that is complementary to what can be obtained

**Table 1**
Median and 90% Credible Intervals for the Posterior of the MD and Time-delay Scale

| True Time-delay | $\alpha$ | $\beta$ | $C$ | $\tau$ (Gyr) | $\log_{10} \psi_0$ |
|---|---|---|---|---|---|
| Exp. $\tau = 0.1$ Gyr | $2.83^{+0.45}_{-0.49}$ | $5.62^{+0.41}_{-0.44}$ | $3.10^{+0.23}_{-0.24}$ | $0.21^{+0.13}_{-0.11}$ | $-2.08^{+0.56}_{-0.63}$ |
| Exp. $\tau = 1.0$ Gyr | $2.54^{+0.68}_{-0.60}$ | $5.57^{+0.59}_{-0.54}$ | $2.90^{+0.23}_{-0.27}$ | $0.93^{+0.35}_{-0.31}$ | $-1.72^{+0.56}_{-0.60}$ |
| Exp. $\tau = 10$ Gyr | $2.62^{+0.71}_{-0.77}$ | $5.67^{+0.65}_{-0.65}$ | $3.03^{+0.44}_{-0.49}$ | $9.46^{+3.60}_{-3.33}$ | $-1.88^{+0.70}_{-0.65}$ |
| Flat Log | $2.09^{+0.36}_{-0.38}$ | $5.09^{+0.40}_{-0.40}$ | $3.39^{+0.13}_{-0.14}$ | $0.19^{+0.12}_{-0.09}$ | $-2.18^{+0.49}_{-0.55}$ |

**Note.** The first column reports which event set is used.

studying the mass, spin, and eccentricity distribution of gravitational-wave detections. The method we developed can be extended to account for multiple populations, which we will explore in a future publication.

### ORCID iDs

Salvatore Vitale https://orcid.org/0000-0003-2700-0767
Will M. Farr https://orcid.org/0000-0003-1540-8562
Ken K. Y. Ng https://orcid.org/0000-0003-3896-2259
Carl L. Rodriguez https://orcid.org/0000-0003-4175-8881

### References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, PhRvL, 116, 131102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016b, PhRvL, 116, 241102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016c, PhRvL, 116, 241103
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016d, PhRvX, 6, 041015
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016e, LRR, 19, 1
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, PhRvL, 118, 221101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, ApJL, 851, L35
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, PhRvL, 119, 141101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017d, PhRvD, 96, 022001
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017e, CQGra, 34, 044001
Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, CQGra, 32, 024001
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57
Belczynski, K., Holz, D. E., Bulik, T., & O'Shaughnessy, R. 2016, Natur, 534, 512
Belczynski, K., Kalogera, V., & Bulik, T. 2001, ApJ, 572, 407
Belczynski, K., Perna, R., Bulik, T., et al. 2006, ApJ, 648, 1110
Belczynski, K., Ryu, T., Perna, R., et al. 2017, MNRAS, 471, 4702
Berger, E., Fox, D. B., Price, P. A., et al. 2007, ApJ, 664, 1000
Chen, H.-Y., Holz, D. E., Miller, J., et al. 2017, arXiv:1709.08079
Chruslinska, M., Nelemans, G., & Belczynski, K. 2019, MNRAS, 482, 5012
Dominik, M., Belczynski, K., Fryer, C., et al. 2012, ApJ, 759, 52
Dominik, M., Belczynski, K., Fryer, C., et al. 2013, ApJ, 779, 72
Farr, B., Berry, C. P. L., Farr, W. M., et al. 2016, ApJ, 825, 116
Farr, W. M., Gair, J. R., Mandel, I., & Cutler, C. 2015, PhRvD, 91, 023005
Farr, W. M., Sravan, N., Cantrell, A., et al. 2011, ApJ, 741, 103
Fishbach, M., Holz, D. E., & Farr, W. M. 2018, ApJL, 863, L41
Foreman-Mackey, D., Hogg, D. W., & Morton, T. D. 2014, ApJ, 795, 64
Hall, E. D., & Evans, M. 2019, CQGra, 36, 225002
Harry, G. M. 2010, CQGra, 27, 084006
Hogg, D. W., Myers, A. D., & Bovy, J. 2010, ApJ, 725, 2166
Kinugawa, T., Miyamoto, A., Kanda, N., & Nakamura, T. 2016, MNRAS, 456, 1093
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Mandel, I. 2010, PhRvD, 81, 084029
Nakar, E. 2007, PhR, 442, 166
Punturo, M., Abernathy, M., Acernese, F., et al. 2010, CQGra, 27, 194002
Regimbau, T., Evans, M., Christensen, N., et al. 2017, PhRvL, 118, 151105
Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016, PhRvD, 93, 084029
Rodriguez, C. L., Morscher, M., Pattabiraman, B., et al. 2015, PhRvL, 115, 051101
Safarzadeh, M., Berger, E., Ng, K. K. Y., et al. 2019, ApJL, 878, L13
The LIGO Scientific Collaboration, the Virgo Collaboration, Abbott, B. P., et al. 2019, PhRvX, 9, 011001
The LIGO Scientific CollaborationThe Virgo Collaboration, Abbott, B. P., et al. 2018, ApJ, in press (arXiv:1811.12940)
Van Den Broeck, C. 2010, in Proc. MG12 Meeting on General Relativity, 12th Marcel Grossmann Meeting on General Relativity, ed. T. Damour, R. Jantzen, & R. Ruffini (Singapore: World Scientific), 1682
Veitch, J., Raymond, V., Farr, B., et al. 2015, PhRvD, 91, 042003
Vitale, S., & Evans, M. 2017, PhRvD, 95, 064052
Youdin, A. N. 2011, ApJ, 742, 38