Remnant Black Hole Kicks and Implications for Hierarchical Mergers

Parthapratim Mahapatra,[1] Anuradha Gupta,[2] Marc Favata,[3] K. G. Arun,[1] and B. S. Sathyaprakash[4, 5, 6]

[1]*Chennai Mathematical Institute, Siruseri, 603103, India*
[2]*Department of Physics and Astronomy, The University of Mississippi, Oxford MS 38677, USA*
[3]*Department of Physics & Astronomy, Montclair State University, 1 Normal Avenue, Montclair, NJ 07043, USA*
[4]*Institute for Gravitation and the Cosmos, Department of Physics, Penn State University, University Park, PA 16802, USA*
[5]*Department of Astronomy and Astrophysics, Penn State University, University Park, PA 16802, USA*
[6]*School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, United Kingdom*

ABSTRACT

When binary black holes merge in dense star clusters, their remnants can pair up with other black holes in the cluster, forming heavier and heavier black holes in a process called *hierarchical merger*. The most important condition for hierarchical merger to occur is that remnants formed by mergers are retained by the host star cluster. Using the publicly available gravitational-wave event database, we infer the magnitudes of kick velocities imparted to the remnant black holes due to anisotropic emission of gravitational waves and use that to quantify the retention probability of each event as a function of the escape speed of the star cluster. Among the second gravitational-wave transient catalog (GWTC-2) events, GW190814 provides the tightest constraint on the kick magnitude with $V_{\text{kick}} = 74^{+10}_{-7}$ km/s at the 90% credible level. We find that star clusters with escape speeds of 200 km/s can retain about 50% of the events in the GWTC-2. Using the escape speed distributions of nuclear star clusters and globular clusters, we find that ~ 17 (2) remnants of GWTC-2 may be retained by the host star cluster if all GWTC-2 events occurred in nuclear (globular) clusters. Our study demonstrates the importance of folding in kick velocity inferences in future studies of hierarchical mergers.

Keywords: Gravitational Waves; Astrophysics

1. INTRODUCTION

Several of the stellar-mass compact binary mergers observed to date could have occurred in gravitationally bound environments, such as globular clusters (GCs) or nuclear star clusters (NSCs) (Portegies Zwart & McMillan 2000; Miller & Lauburg 2009; Downing et al. 2010; Rodriguez et al. 2015, 2016, 2018; Petrovich & Antonini 2017; Arca-Sedda & Gualandris 2018; Samsing 2018; Fragione & Kocsis 2018; Zevin et al. 2019; Antonini & Gieles 2020). If the star cluster is dense enough, the remnant black hole formed by a merger can subsequently pair with another black hole (BH) via dynamical interactions, merging again under gravitational-wave radiation reaction (Doctor et al. 2020; Antonini et al. 2019; Fragione & Silk 2020; Mapelli et al. 2020). This can repeat, producing increasingly more massive black holes at each step—a process called *hierarchical mergers* (Fra-

gione et al. 2018a,b). Evidence of significant numbers of black holes formed via hierarchical mergers would strongly suggest that compact object binaries form dynamically in dense star clusters. This process can also lead to the formation of intermediate mass black holes (Miller & Hamilton 2002), which could in turn seed the formation of supermassive black holes (Bellovary et al. 2019).

The presence of *heavy* BHs (those with masses $> 50 M_\odot$) in the second LIGO/Virgo Collaboration (LVC) Gravitational-Wave Transient Catalog (GWTC-2) (Abbott et al. 2021a) has led to detailed investigation into hierarchical mergers (Gerosa et al. 2020; Kimball et al. 2020; Tiwari & Fairhurst 2021). Recent analyses (Kimball et al. 2021) have found evidence of hierarchical mergers in six GWTC-2 events. This inference relies on assessing the consistency of the mass and spin parameter distributions with those expected from hierarchical mergers (Fishbach et al. 2017; Gerosa & Berti 2017).

It is well known that the anisotropic emission of gravitational waves (GWs) during the end-stages of binary

coalescence carries away linear momentum and produces a recoil or "kick" of the merger remnant (Fitchett 1983; Favata et al. 2004). This plays a decisive role in the hierarchical merger process, which cannot occur if a merger remnant is ejected from its dense stellar environment. Black hole kicks can reach hundreds to thousands of kilometers per second (Campanelli et al. 2007a; Gonzalez et al. 2007b; Tichy & Marronetti 2007; Lousto & Zlochower 2011; Lousto & Healy 2019). In comparison, globular clusters have escape speeds $\sim$ 2–180 km/s, and nuclear star clusters have larger escape speeds $\sim$ 10–1000 km/s (Antonini & Rasio 2016). For the hierarchical merger process to operate, the BH kick should not exceed the cluster escape speed (Merritt et al. 2004).

A direct detection of the GW kicks is very challenging as it would entail measuring either the very small Doppler shift in the frequencies of the quasi-normal modes of the remnant BHs (Favata 2009; Gerosa & Moore 2016), the rich structure in the higher-order modes of gravitational waves (Calderón Bustillo et al. 2018), or tracking certain post-merger features in the gravitational waveforms (Calderon Bustillo et al. 2020). However, numerical relativity (NR) provides useful fitting formulas (Campanelli et al. 2007b), expressing the remnant kick as a function of the mass ratio and spin configuration of the binary components. This permits inference of GW kicks for a binary black hole (BBH) merger, provided there are reasonably precise measurements of the mass and spin parameters (Varma et al. 2020).

Recently, Fragione & Loeb (2021) computed the kick velocity distributions of GWTC-2 events as a function of the spins of the binary components and studied the probability of retaining these mergers as a function of spin magnitudes. They found that only nuclear star clusters with escape speeds higher than 100 km/s could retain the GWTC-2 merger remnants even for dimensionless spins as small as 0.1. Note that binaries containing higher-spinning BHs typically experience larger kicks as compared to binaries with low-spin BHs, and can thus be more easily ejected from clusters.

Doctor et al. (2021) mapped the properties of GWTC-2 BH remnants to the statistical properties of remnant BHs in our universe. They obtained the distribution of kick velocities of GWTC-2 events using NR surrogate waveform models (Varma et al. 2019b,a) and found that globular clusters and nuclear clusters can retain $\sim$ 4% and 45% of the remnants, respectively. They assumed the escape speeds of all globular clusters to be 50 km/s and that of all nuclear clusters to be 250 km/s.

Our primary goal is to compute the posterior distributions of the GWTC-2 remnant kicks using NR fitting formulas and deduce the probability that remnant BHs are retained by their host star clusters. This, in turn, will help us determine if GWTC-2 remnants can facilitate hierarchical mergers.

Specifically, we develop a framework to compute the probability that a BBH merger is retained by its environment (assuming the merger takes place inside a star cluster and that dynamical N-body interactions do not play a significant role in ejecting heavy BHs). This is applied to all BBHs in the GWTC-2 catalog, obtaining their individual retention probabilities as a function of the cluster escape speed. We also discuss what the kick velocities of the GWTC-2 population imply for different types of star clusters and their efficiency to retain BH remnants.

We find that BH kicks will not be a major obstacle for hierarchical mergers if star clusters with escape speeds $\gtrsim$ 200 km/s are abundant in the universe and BBHs with masses characteristic of those in GWTC-2 primarily form in such clusters. We also find that among the six GWTC-2 events identified by Kimball et al. (2021) as systems showing evidence for hierarchical mergers, all except GW190517_055101 may be retained by their host clusters with a probability $\sim$ 50%, paving the path for participating in further binary formation, provided they merged in clusters with escape speeds larger than $\sim$ 700 km/s. With our present knowledge of the escape speed distribution of star clusters, $\sim$ 17 (2) of the GWTC-2 remnants may be retained by nuclear (globular) clusters. *We stress the importance of using kick velocity measurements as an ingredient in future studies of hierarchical black hole formation.*

2. INFERRING RETENTION PROBABILITY FROM KICK POSTERIOR DISTRIBUTION

We first discuss how we infer the kick velocity probability distribution for a given GWTC-2 event. We then compute the retention probability of each BBH event as a function of the cluster escape speed V_{esc}.

2.1. *Kick inference*

The kick imparted by the GW linear momentum loss is described by fitting formulas based on NR simulations of BBHs (Gonzalez et al. 2007b; Campanelli et al. 2007b; Gonzalez et al. 2007a; Lousto & Zlochower 2008; Lousto et al. 2012a; Lousto & Zlochower 2013). We apply the kick magnitude of Campanelli et al. (2007b), summarized in Appendix A.

The kick velocity depends on the following parameters: the mass ratio q ($\leq$ 1); the dimensionless spin magnitudes χ_i; the angles between the spin vectors and the total orbital angular momentum θ_i; the difference
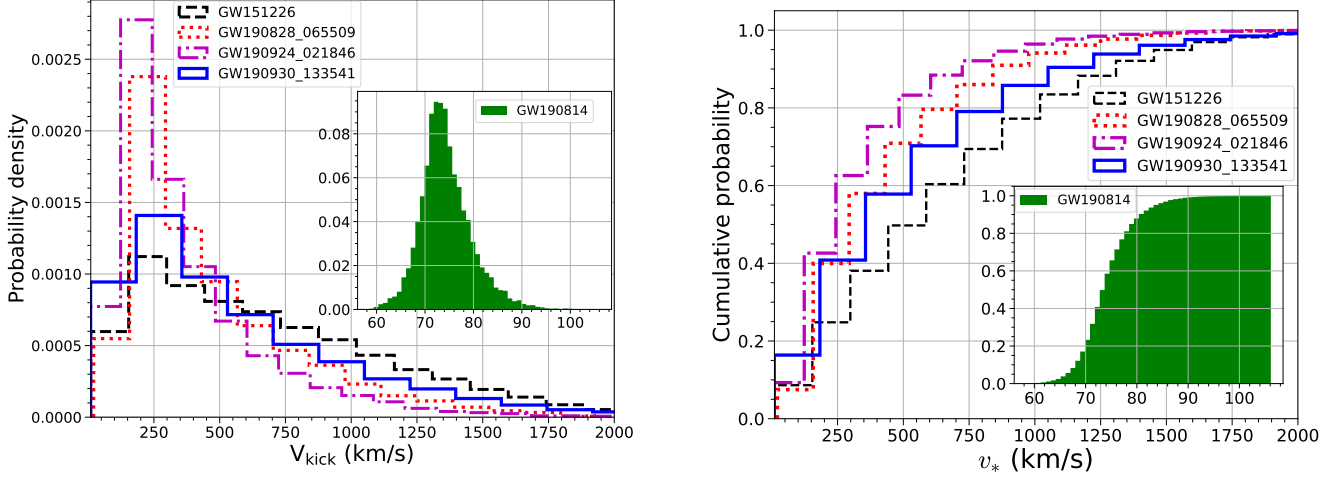
Figure 1. The probability density function (PDF) $p(\mathrm{V_{kick}})$ (left), and the cumulative distribution function (CDF) $F(v_*)$ (right) of the kick magnitude for representative GW events in GWTC-2. The PDF is constructed by combining the parameter estimation posterior samples for the binary parameters with the fitting formula for the kick. The CDF $F(v_*)$ is computed by integrating $p(\mathrm{V_{kick}})$ from 0 to v_*. Shown in the insets are the PDF and the CDF of GW190814, which has the most precisely inferred kick magnitude and the lowest kick of all GWTC-2 events.

between the azimuthal angles of the two spin vectors, ϕ_{12}; and an additional parameter Θ, which is the angle between $\boldsymbol{\Delta} \times \hat{\mathbf{L}}$ and a fiducial infall direction of the two BHs at merger (see Appendix A for the definition of $\boldsymbol{\Delta}$ and $\hat{\mathbf{L}}$). Here $\mathrm{i} = 1, 2$ labels the primary or secondary BH. Currently the posterior samples of θ_{i} for GWTC-2 events are not available, but the posterior samples of $\theta_{LS_{\mathrm{i}}}$ (the angles between the spin vectors and the direction of the Newtonian orbital angular momentum $\hat{\mathbf{L}}_N$) are accessible. Here we use the posterior samples for $\theta_{LS_{\mathrm{i}}}$, assuming $\hat{\mathbf{L}} \approx \hat{\mathbf{L}}_N$ and $\theta_{\mathrm{i}} \approx \theta_{LS_{\mathrm{i}}}$. Deviations in the directions of $\hat{\mathbf{L}}$ and $\hat{\mathbf{L}}_N$ only enter as relative 1.5 post-Newtonian-order corrections (i.e., corrections cubic in the relative orbital speed) and are proportional to the components of the spin perpendicular to $\hat{\mathbf{L}}_N$ (Kidder 1995); since there is little evidence for large spins or precession in GWTC-2, $\hat{\mathbf{L}} \approx \hat{\mathbf{L}}_N$ is a reasonable approximation.

In addition to feeding samples of q, χ_{i}, $\theta_{LS_{\mathrm{i}}}$, and ϕ_{12}, into the kick fitting formula, we also need to know the parameter Θ. As we cannot directly infer Θ from the posterior samples, we assume it is uniformly distributed between $[0, 2\pi]$ (Gerosa & Sesana 2015; Gerosa & Kesden 2016) when computing the posteriors of the kick magnitude via Equation (A1).

Restricting to the 47 BBHs in GWTC-2 (Abbott et al. 2021a), we construct probability density functions (PDFs) for the kick magnitude, using as input the posterior PDFs for the mass ratio and spin parameters of each BBH event. (See Appendix A for additional details on constructing the kick magnitude PDFs.)

Figure 1 shows some typical examples of the resulting kick PDFs $p(\mathrm{V_{kick}})$ for a few selected GW events. GW190814 provides the best inference of the kick and has the lowest inferred kick. The former is due to the stringent constraint on the spin of its primary ($\chi_1 \leq 0.007$), and the latter is due to the small spin magnitude.

As the spins for several GW events are poorly constrained, our analysis risks returning posteriors that are prior dominated. We use the Jensen-Shannon (JS) divergence (Lin 1991) to quantify how informative are the posteriors compared to the priors, considering only those events that cross a JS divergence threshold of 0.007. This leaves 42 events for further analysis. See Appendix B for further discussion.

Given a posterior PDF $p(\mathrm{V_{kick}})$ for the kick magnitude of each event, the corresponding cumulative distribution function (CDF) $F(v_*) = \int_0^{v_*} p(\mathrm{V_{kick}})\, d\mathrm{V_{kick}}$ quantifies the probability that $\mathrm{V_{kick}} \leq v_*$ and therefore gives the retention probability $F(\mathrm{V_{esc}})$ given a cluster with an escape velocity $\mathrm{V_{esc}} = v_*$. Example CDFs are shown in the right panel of Figure 1.

2.2. *Retention probability of GWTC-2 events*

Operating under the assumption that all reported GWTC-2 BBHs occur in the dense core of star clusters with unknown escape speeds, the retention probability $\mathrm{P_{ret},_j}(\mathrm{V_{esc}})$ of the jth event in a cluster with es-
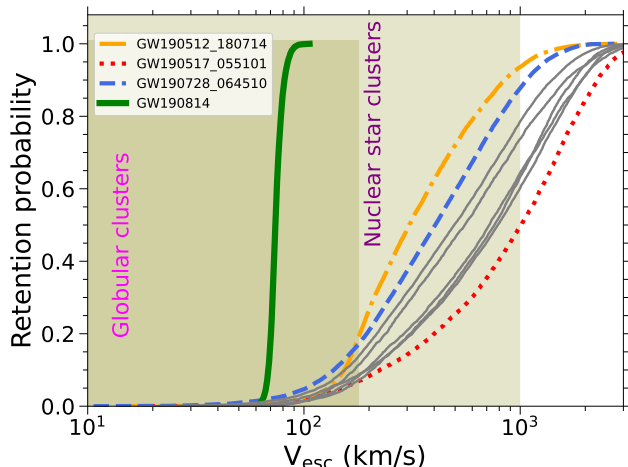
Figure 2. Retention probability of a few representative GWTC-2 events as a function of the cluster escape speed (thick curves, listed in the legend). The gray curves correspond to five of the six events reported by Kimball et al. (2021) to be of hierarchical origin (the sixth one is the red dotted curve; see Section 2.3 for discussion). The retention probability is computed directly from the kick CDF (Figure 1) as discussed in the text. The shaded regions show the range of escape speeds for globular clusters and nuclear star clusters (Antonini & Rasio 2016). For example, a nuclear star cluster with escape speed ~ 500 km/s can retain GW190512_180714-like systems with probability ~ 0.70, while a globular cluster with escape speed ~ 74 km/s can retain a GW190814-like system with probability ~ 0.50. Higher retention probability increases the likelihood that merger remnants can participate in subsequent BBH mergers within the cluster.

cape speed V_{esc} is simply found by replacing $v_* \rightarrow V_{esc}$ in the CDF; i.e., $P_{ret,j}(V_{esc}) = F(V_{esc})$. For example, because the CDF for the GW190814 kick distribution indicates that there is $\approx 50\%$ probability that $V_{kick} < 74$ km/s (Figure 1 insets), this implies that a cluster with $V_{esc} = 74$ km/s would retain this binary with $\approx 50\%$ probability.

We apply this observation to the 42 BBHs in GWTC-2 with informative kick posteriors, mapping the kick magnitude CDFs to their retention probabilities. This result is shown in Figure 2 for a selected sample of the 42 GWTC-2 events. (The four events highlighted in thick colored lines are representative examples of the range of kick velocity distributions; the thin gray curves are discussed below.) These retention probabilities help to quantify, as a function of the cluster escape speed, the ability of these events to take part in future BBH mergers.

It is evident that GW190814 has the highest retention probability among the GWTC-2 events with informative

kick posteriors. It would almost surely be retained by a cluster with an escape speed as low as 80 km/s. The lowest retention probability is for GW190517_055101, which would be retained at 50% probability by clusters with an escape speed of ~ 1000 km/s.

As these conclusions depend only weakly on the details of the clusters, our method provides a powerful diagnostic of the ability of star clusters to retain the BBH merger remnants. Retaining these remnants is an important prerequisite if they are to participate in subsequent BBH mergers.

2.3. *Retention of likely hierarchical merger events*

We now assess the retention probabilities of the six events reported in Kimball et al. (2021) and discussed earlier. Using the method of the previous section, we compute the kick PDFs for these six events and determine their retention probabilities $P_{ret,j}(V_{esc})$ as a function of cluster escape speed. The result is shown in Figure 2, where the gray lines represent five out of the six events that Kimball et al. (2021) indicate as likely to be hierarchical in origin. The sixth event (GW190517_055101) is shown in dotted red and has the lowest retention probability (30% for a cluster with escape speed 600 km/s). The highest retention probability is for GW190602_175927: its remnant is retained with 60% probability if the cluster has an escape speed of 600 km/s. GW190521, the most massive BBH detected to date, also has a significant retention probability. A dense star cluster with escape speed ~ 700 km/s will be able to retain GW190521 with 50% probability, potentially allowing for the future formation of an intermediate mass BBH.

We note that our method complements that of Kimball et al. (2021). While their approach determines the hierarchical merger history of the binary constituents by measuring the masses and spins, our method projects these systems into the future and quantifies their ability to further participate in such mergers.

3. CHARACTERIZING CLUSTERS VIA THEIR RETENTION OF BH MERGER REMNANTS

Having examined the retention probability of a given GWTC-2 event j, we now extend our results to the entire population of all informative GWTC-2 events. Let us suppose that all these informative events are characteristic of the population of BBHs that merge in a single star cluster that is solely parameterized by its escape speed V_{esc}. Over time this entire population merges in this cluster, and individual merger remnants are retained if $V_{kick}^j < V_{esc}$. If the escape speed of the cluster V_{esc} is large compared to typical values of $\{V_{kick}^j\}$,
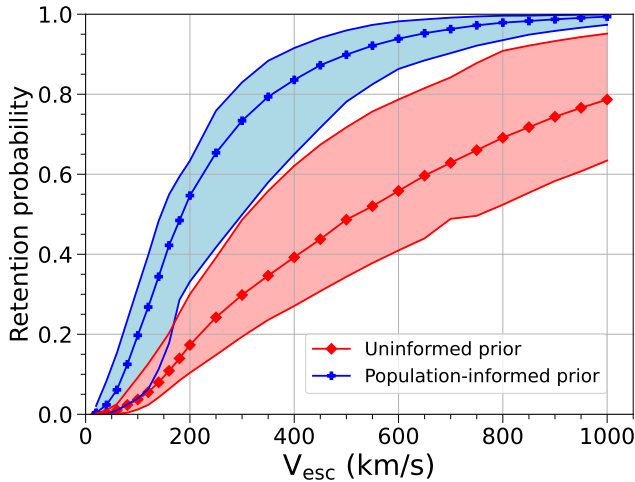
Figure 3. Distribution of retention probability, with and without population weighted samples, computed from all the informative events in GWTC-2 as a function of the escape speed of the star cluster V_{esc}. The solid lines with markers show the median values while the shaded regions show the 90% credible intervals of the distributions. These curves are obtained without any assumptions about the cluster type or properties (besides being characterized by a single value, V_{esc}).

a large fraction of the population is retained and can go on to seed the next generation of mergers. If not, then clusters with a particular value of V_{esc} (or lower) are likely to be poor sites for subsequent generations of BBH mergers. Here we will attempt to characterize if a particular cluster with escape speed V_{esc} is a good site for subsequent BBH mergers.

It is a strong assumption that the GWTC-2 BBHs form a sample reflective of the BBH population in a typical cluster parameterized only by its escape speed, and this assumption needs to be informed by future GW observations and cluster modeling. Besides, posteriors from the individual events do suffer from selection effects associated with gravitational wave observations (Fishbach & Holz 2020; Moore & Gerosa 2021). Such selection effects may be mitigated on individual event posteriors with a careful reweighting via priors that are population-informed, as opposed to uniform, on different parameters (Abbott et al. 2021b). This is achieved by invoking the population model inferred from the GWTC-2 and using the corresponding priors on various parameters for the analysis. The population-reweighted parameter estimation samples for different events are drawn from the *GWTC-2 Data Release* (LIGO Scientific Collaboration and Virgo Collaboration 2020) corresponding to the Power-law + Peak mass model (Abbott et al. 2021b).

To characterize a particular cluster's suitability as a site of future hierarchical growth via its retention probability, we consider posteriors from individual events with uninformed priors and population-informed priors. We look only into the 42 informative events based on the JS divergence criterion mentioned earlier, of which population-informed priors are available only for 40 of them (GW190719_215514 and GW190909_114149 are excluded). Hence, the results for population-informed priors are based on 40 events, whereas those for uninformative priors are based on 42 events.

For a cluster with escape speed V_{esc}, Section 2 showed that a given event j has a retention probability $P_{ret,j}(V_{esc})$ (which, e.g., can be read off from Figure 2). If we consider that the entire set $\{j = 1, \dots, j_{max}\}$ of informative GWTC-2 events is in a single cluster with V_{esc}, and that the set is representative of the cluster's BBH population, then the samples for the distribution of retention probability for that cluster are given by the set $[P_{ret,j}(V_{esc}), \; j = 1, \dots, j_{max}]$, where j_{max} is 40 or 42 as discussed above. This defines our distribution of retention probability, which is readily calculated from data like that shown in Figure 2.

Using our calculated values for $P_{ret,j}(V_{esc})$, we proceed to calculate the distribution of retention probability as a function of the cluster escape speed. In doing so, we make no assumption about the nature of the cluster. The result is shown in Figure 3. The dots, which are joined, show the median retention probability and the shaded regions denote the 90% credible interval.

With the uninformed priors, we find that clusters with escape speeds of 200 km/s and 500 km/s can retain $17^{+13}_{-7}\%$ and $49^{+23}_{-14}\%$ of the informative GWTC-2 population, while a cluster with an escape speed of about 800 km/s can retain $69^{+22}_{-17}\%$ of the remnants. When we use population-informed priors, we see that clusters with escape speeds of 200 km/s and 500 km/s can retain $55^{+9}_{-22}\%$ and $90^{+6}_{-12}\%$ of the informative GWTC-2 population, while a cluster with an escape speed of about 800 km/s can retain $98^{+2}_{-4}\%$ of the remnants. The numbers arising from the population-weighted analysis point to a significantly higher retention probability, thereby increasing the prospects of hierarchical mergers. Our estimates with population weighted samples broadly agree with those in Doctor et al. (2021), which obtained retention fractions of 4% and 40% for clusters with escape speeds of 50 km/s and 250 km/s.

An important implication of this result concerns the role of globular clusters, which have $V_{esc} \lesssim 80 - 180$ km/s (with 180 km/s at the high end of the escape speed distribution and most clusters under 80 km/s; see Figure 3 of Antonini & Rasio (2016)). These have median

retention probabilities $\lesssim 14\%$ ($\lesssim 48\%$ with population weighted analysis). While a small subset of globular clusters could facilitate hierarchical mergers, the overall probability of globular clusters serving as major sites of hierarchical mergers is low. However, nuclear star clusters, which have escape speeds $\gtrsim 200$ km/s, can readily host hierarchical mergers as they retain more than half of the merger remnants.

4. RETENTION PROBABILITIES OF GWTC-2 EVENTS BY GLOBULAR AND NUCLEAR STAR CLUSTERS

We now map our "agnostic cluster" curves in Figure 3 to globular and nuclear clusters, using their expected escape speed distributions (Harris 1996; Antonini & Rasio 2016; Georgiev et al. 2016) to determine the probability that a particular GW event can be retained by a GC or NSC. (Note that these estimates of escape speeds are from the center of the cluster for GCs, whereas for NSCs these are defined at the half-mass radius of the cluster.)

We first construct a cluster escape speed PDF $p_k(\mathrm{V_{esc}})$, where the index k denotes GC or NSC. This PDF is constructed using log-normal fits to the data in Figure 3 of Antonini & Rasio (2016), which have means $\left\langle \log_{10}\left(\frac{\mathrm{V_{esc}}}{\mathrm{km/s}}\right)\right\rangle = (1.5, 2.2)$ and standard deviations $\sigma_{\log_{10}[\mathrm{V_{esc}}/(\mathrm{km/s})]} = (0.3, 0.36)$ for GCs and NSCs (respectively). The corresponding CDF (giving the probability that $\mathrm{V_{esc}} \leq v_*$) is denoted by $F_k(v_*)$. Then the probability for a BH with kick $\mathrm{V_{kick}}$ to be retained by a given cluster is $1 - F_k(\mathrm{V_{kick}})$.

Given our kick PDF $p_j(\mathrm{V_{kick}})$ for GWTC-2 event j, the probability that event j is retained by a cluster of type k is then

$$\mathrm{P}_{\mathrm{ret},j}^{k} = \int_0^{\mathrm{V_{max}}} p_j(\mathrm{V_{kick}})[1 - F_k(\mathrm{V_{kick}})]\, d\mathrm{V_{kick}} . \quad (1)$$

This simply amounts to weighting each bin of the kick PDF by the probability $[1-F_k(\mathrm{V_{kick}})]$ that a cluster k retains a kicked BH with $\mathrm{V_{kick}}$. We set $\mathrm{V_{max}} = 5000$ km/s (any choice $\mathrm{V_{max}} \gtrsim 1000$ km/s would make only a negligible difference). It should be noted that Equation (1) assumes a uniform merger rate across the different clusters. In reality massive clusters, which have larger escape speeds, may have a higher merger rate. This may be folded into this estimate in the future.

Applying Equation (1), the retention probability of GWTC-2 event j by a GC or NSC is computed and given in Table 1. These estimates are also given for both uninformed and population-informed priors. The highest retention probability is for GW190814 which is $\approx 82\%$ ($\approx 84\%$) for uninformed (population-informed) priors for NSCs. This is due to the precise estimation of the low
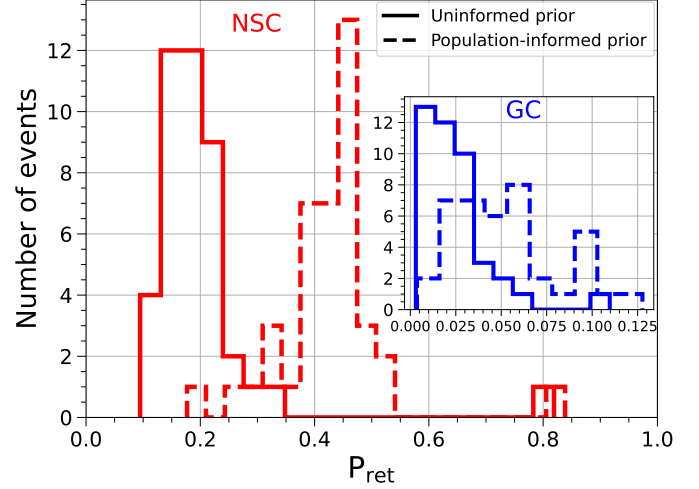
Figure 4. Distribution of retention probabilities for nuclear star clusters (red curves) and globular clusters (blue curves, inset) obtained from the informative events of GWTC-2. The distribution is obtained by binning the data in Table 1. With uninformed priors (solid curves) 90% of the events have retention probability between 0.121-0.287 (0.006-0.056) for nuclear (globular) clusters. With population-informed priors (dashed curves), 90% of the events have a retention probability between 0.304-0.514 (0.016-0.102) for nuclear (globular) clusters.

kick magnitude for this event. With uninformed priors, five events have retention probabilities higher than 25% for NSCs, whereas with population-informed samples 15 events have retention probabilities higher than 45% for NSCs. Similarly, the probability that GW190814 is retained by a GC, assuming the uninformed (population-informed prior) is $\approx 11\%$ ($\approx 13\%$). With the uninformed priors, four events have retention probabilities in GCs higher than 4%, while the population-informed analysis yields 17 events with retention probabilities in GCs higher than 6%.

Using the numbers in Table 1, we construct in Figure 4 the distribution of retention probabilities for NSCs and GCs. If all the informative GWTC-2 mergers happened in NSCs, we find that the retention probabilities of 90% of those events would be between 0.121 and 0.287 (0.304–0.514) based on uninformed (population-informed) priors. Similarly, if all informative GWTC-2 events are associated with globular clusters, 90% of them will have retention probabilities between 0.006 and 0.056 (0.016–0.102) using uninformed (population-informed) priors. The outlier bump at large $\mathrm{P_{ret}}$ in Figure 4 is due to GW190814; with the smallest (and best-measured) kick magnitude, it yields the largest retention probability.

Table 1. Probability of retention for all informative GWTC-2 events by GCs or NSCs using uninformed and population-informed (inside the parentheses) priors; see the text for details. This table uses Equation (1) to compute the probability that a BH merger remnant is retained by a cluster, given the distribution of cluster escape speeds and the inferred kick posterior distribution for each GWTC-2 event.

| Event name | Retention probability | | Event name | Retention probability | |
	GC	NSC		GC	NSC
GW150914	0.0563 (0.0967)	0.2880 (0.4780)	GW190602_175927	0.0097 (0.0455)	0.1795 (0.4139)
GW151226	0.0216 (0.0273)	0.1645 (0.3069)	GW190620_030421	0.0110 (0.0237)	0.1276 (0.3150)
GW170104	0.0331 (0.0313)	0.2192 (0.4190)	GW190630_185205	0.0357 (0.0320)	0.2385 (0.4004)
GW170729	0.0182 (0.0335)	0.1413 (0.3298)	GW190701_203306	0.0161 (0.0749)	0.1875 (0.4745)
GW170814	0.0345 (0.0440)	0.1980 (0.4429)	GW190706_222641	0.0124 (0.0226)	0.1603 (0.3712)
GW170818	0.0271 (0.0373)	0.1724 (0.3936)	GW190708_232457	0.0489 (0.1058)	0.2728 (0.5189)
GW170823	0.0291 (0.0638)	0.1897 (0.4570)	GW190719_215514	0.0136	0.1432
GW190408_181802	0.0383 (0.0655)	0.2247 (0.4700)	GW190720_000836	0.0250 (0.0413767)	0.2125 (0.4069)
GW190412	0.0031 (0.0037)	0.1903 (0.2442)	GW190727_060333	0.0238 (0.0916)	0.1768 (0.4929)
GW190413_052954	0.0274 (0.0940)	0.1939 (0.4540)	GW190728_064510	0.0311 (0.0478)	0.2278 (0.4444)
GW190413_134308	0.0202 (0.0464)	0.1658 (0.4104)	GW190731_140936	0.0252 (0.0652)	0.1757 (0.4567)
GW190421_213856	0.0130 (0.0770)	0.1662 (0.4706)	GW190803_022701	0.0145 (0.0612)	0.1961 (0.4463)
GW190424_180648	0.0192 (0.0585)	0.1534 (0.4371)	GW190814	0.1097 (0.1276)	0.8190 (0.8384)
GW190503_185404	0.0112 (0.0642)	0.1930 (0.4757)	GW190828_063405	0.0181 (0.0322)	0.1614 (0.3811)
GW190512_180714	0.0344 (0.0261)	0.2735 (0.4255)	GW190828_065509	0.0204 (0.0216)	0.2226 (0.3905)
GW190514_065416	0.0117 (0.0656)	0.1647 (0.4674)	GW190909_114149	0.0098	0.1658
GW190517_055101	0.0092 (0.0168)	0.0947 (0.1768)	GW190910_112807	0.0361 (0.0992)	0.2217 (0.5133)
GW190519_153544	0.0060 (0.0090)	0.1096 (0.3138)	GW190915_235702	0.0138 (0.0790)	0.1519 (0.4534)
GW190521	0.0071 (0.1023)	0.1201 (0.4732)	GW190924_021846	0.0583 (0.0387)	0.3216 (0.4456)
GW190521_074359	0.0174 (0.0620)	0.1854 (0.4278)	GW190929_012149	0.0049 (0.0241)	0.2080 (0.3928)
GW190527_092055	0.0142 (0.0399)	0.1615 (0.4041)	GW190930_133541	0.0299 (0.0506)	0.2130 (0.4208)

The area of the two distributions in Figure 4 (or summing the values in Table 1) gives us an estimate of the number of remnants that will be retained out of the 42 (40) events, corresponding to the uninformed (population-informed) priors, we considered in the GWTC-2 population. There are ~ 9 (~ 1) that could participate in further mergers if those events happened in NSCs (GCs), using uninformed priors. The same numbers for the population-informed priors are ~ 17 (~ 2). The estimates based on standard single event analysis are in good agreement with the projections of Mapelli et al. (2020), which indicate that roughly 15% (0.6%) of mergers in NSCs (GCs) would be hierarchical in nature. Note that these are present-day ($z = 0$) retention probabilities; given the limited redshift reach of GWTC-2, we have not considered the redshift evolution of retention probability.

5. CONCLUSION

Formation and growth of black holes at different scales is an important open problem in astrophysics. Hierarchical mergers are one of the channels considered for black hole growth, although its efficiency is largely un-

known. Observations of binary black hole mergers by LIGO/Virgo provide one of the most important tools to probe the hierarchical growth of black holes at the stellar mass scale. Predicted mass and spin distributions from hierarchical mergers were compared with GWTC-2 events in Kimball et al. (2021) to assess the presence of hierarchical mergers. We discussed another approach for evaluating the prospects of hierarchical growth, by computing the probability that GWTC-2 merger remnants are retained by cluster environments—a necessary condition for hierarchical growth. This involved computing the gravitational-wave kick probability distribution for each GWTC-2 event and comparing with cluster escape speeds. Using population-informed posterior samples, we found that among the GWTC-2 merger remnants, 17 remnants are expected to be retained by their host-clusters if they all merged in a nuclear star cluster, or about two events if all GWTC-2 mergers occurred in globular clusters. These retained remnants could potentially participate in a subsequent generation of binary black hole mergers, provided favorable circumstances for binary formation exist near these black hole remnants.

As hierarchical mergers are assisted by larger cluster escape speeds and a higher frequency of in-cluster mergers, understanding the redshift evolution of these two quantities will be important in determining the efficiency of hierarchical mergers as a function of cosmic time.

Our conclusions assume that all GWTC-2 mergers happened in star clusters and not in galactic fields, and this is hard to verify at present. As enhanced detector sensitivities allow the discovery of more binary black holes with improved parameter accuracy and precision, the methods discussed here will provide a unique probe of the hierarchical growth of stellar-mass black holes.

APPENDIX

A. KICK FITTING FORMULA

Here we summarize the employed model for inferring the magnitude of the kick velocity. We directly apply fitting formulas based on numerical relativity simulations of binary black holes (Gonzalez et al. 2007b; Campanelli et al. 2007b; Gonzalez et al. 2007a; Lousto & Zlochower 2008; Lousto et al. 2012a; Lousto & Zlochower 2013). Following Campanelli et al. (2007b), the kick velocity magnitude is decomposed as

$$V_{\rm kick} = \sqrt{V_m^2 + 2V_m V_{s\perp}\cos\xi + V_{s\perp}^2 + V_{s\parallel}^2}\,, \tag{A1}$$

where V_m, $V_{s\perp}$, $V_{s\parallel}$ are expressed in terms of

$$\boldsymbol{\Delta} = \frac{\boldsymbol{\chi_1} - q\boldsymbol{\chi_2}}{1+q}\,, \tag{A2}$$

$$\widetilde{\boldsymbol{\chi}} = \frac{q^2\boldsymbol{\chi_2} + \boldsymbol{\chi_1}}{(1+q)^2}\,, \tag{A3}$$

where $q \leq 1$ is the mass ratio of the binary and $\boldsymbol{\chi_{1,2}}$ are the dimensionless spin vectors of the progenitor BHs (Lousto et al. 2012a; Lousto & Zlochower 2013; Gerosa & Kesden 2016). Contributions to the kick velocity listed in Equation (A1) are given by (Lousto et al. 2012a; Lousto & Zlochower 2013):

$$V_m = A\,\eta^2\frac{1-q}{1+q}(1 + B\,\eta)\,, \tag{A4}$$

$$V_{s\perp} = H\,\eta^2\,\Delta_{\parallel}\,, \tag{A5}$$

$$V_{s\parallel} = 16\,\eta^2[\Delta_{\perp}(V_{11} + 2V_A\widetilde{\chi}_{\parallel} + 4V_B\widetilde{\chi}_{\parallel}^2$$
$$+ 8V_C\widetilde{\chi}_{\parallel}^3) + 2\widetilde{\chi}_{\perp}\Delta_{\parallel}(C_2 + 2C_3\widetilde{\chi}_{\parallel})]\cos\Theta\,, \tag{A6}$$

where

$$\eta = \frac{q}{(1+q)^2}\,, \tag{A7}$$

$$\Delta_{\parallel} = |\boldsymbol{\Delta}\cdot\hat{\mathbf{L}}| = \frac{\chi_1\cos\theta_1 - q\chi_2\cos\theta_2}{1+q}\,, \tag{A8}$$

$$\Delta_{\perp} = |\boldsymbol{\Delta}\times\hat{\mathbf{L}}|$$
$$= \frac{1}{1+q}[\chi_1^2\sin^2\theta_1 + q^2{\chi_2}^2\sin^2\theta_2$$
$$-2q\chi_1\chi_2\sin\theta_1\sin\theta_2\cos\phi_{12}]^{1/2}\,, \tag{A9}$$

$$\widetilde{\chi}_{\parallel} = \widetilde{\boldsymbol{\chi}}\cdot\hat{\mathbf{L}} = \frac{\chi_1\cos\theta_1 + q^2\chi_2\cos\theta_2}{(1+q)^2}\,, \quad \text{and} \tag{A10}$$

$$\widetilde{\chi}_{\perp} = |\boldsymbol{\chi}\times\hat{\mathbf{L}}|$$
$$= \frac{1}{(1+q)^2}[{\chi_1}^2\sin^2\theta_1 + q^4{\chi_2}^2\sin^2\theta_2$$
$$+2q^2\chi_1\chi_2\sin\theta_1\sin\theta_2\cos\phi_{12}]^{1/2}\,. \tag{A11}$$

In the above $\hat{\mathbf{L}}$ is the unit vector along the total orbital angular momentum, $\theta_{1,2}$ are the angles between $\hat{\mathbf{L}}$ and $\boldsymbol{\chi_{1,2}}$, ϕ_{12} is the difference between the azimuthal angles of the two spin vectors, and $\chi_{1,2}$ are the spin magnitudes. The values of the different numerical fitting coefficients are $A = 1.2\times10^4$ km/s, $B = -0.93$ (Gonzalez et al. 2007b), $H = 6.9\times10^3$ km/s (Lousto & Zlochower 2008), $V_{11} = 3677.76$ km/s, $V_A = 2481.21$ km/s, $V_B = 1792.45$ km/s, $V_C = 1506.52$ km/s (Lousto et al. 2012b), $C_2 = 1140$ km/s, and $C_3 = 2481$ km/s (Lousto & Zlochower 2013). The angle $\xi \sim 145°$ for a wide range of quasi-circular configurations (Lousto & Zlochower 2008), whereas Θ is defined as the angle between $\boldsymbol{\Delta}\times\hat{\boldsymbol{L}}$ and a fiducial infall direction of the two holes at merger. (See Gerosa & Kesden (2016) for a more precise definition of Θ, which depends on the orbital phase of the binary at an arbitrary reference time.)

Since the spin vectors can evolve considerably throughout the binary inspiral, the kick velocity fitting formula can only be employed close to the merger, at separations $\sim 10M$ where NR simulations typically start (see Barausse & Rezzolla (2009) and Section 2.1 of Gerosa & Sesana (2015)). To account for this, we evolve the posterior samples (which are specified at a reference frequency of 20 Hz) to a separation of $10M$ using orbit-averaged post-Newtonian equations of motion (Apostolatos et al. 1994; Kidder 1995; Racine 2008) contained in the LALSimulation libraries of LALSuite (LIGO Scientific Collaboration 2018). (For GW190521, the reference frequency is chosen to be 11 Hz.)

For BBHs from the first two observing runs (GWTC-1), we use the posterior samples produced by the Bayesian inference package `Bilby` (Ashton et al. 2019; Romero-Shaw et al. 2020) using the `IMRPhenomPv2` waveform model (Hannam et al. 2014). For events observed during the first half of the third observing run (reported in GWTC-2), posterior samples are drawn from the *GWTC-2 Data Release* (LIGO Scientific Collaboration and Virgo Collaboration 2020) corresponding to the `SEOBNRv4PHM` waveform model (Cotesta et al. 2018); `SEOBNRv4P` (Ossokine et al. 2020) samples are used for those events for which `SEOBNRv4PHM` samples are not available.

B. JENSEN-SHANNON DIVERGENCE FOR THE KICK POSTERIORS

The issue of prior-dominated posteriors in the context of kick magnitude inference was pointed out by Varma et al. (2020) while inferring the kicks for GW150914 and GW170729. Priors on the kick magnitude are constructed by combining priors on the mass ratio and spin parameters with relevant NR fitting formulas for the kick. It was specifically noted that the difference between the inferred kick posteriors and the corresponding priors is very small, as evidenced by the Kullback–Leibler (KL) divergence (Kullback & Leibler 1951) between them. This was attributed to the lack of stringent constraints on the spin parameters of the binary. Here, in order to assess the information content of our kick posteriors relative to our priors, we employ a similar measure, the Jensen–Shannon (JS) divergence (Lin 1991). This is a symmetric (with respect to the two distributions that are being compared) version of the KL divergence and always has a finite value. We set a threshold of 0.007 for the JS divergence (the same threshold used in Abbott et al. (2021a) for comparing posterior distributions between different waveform approximants); above this threshold we consider the posteriors to be informative. For technical details of JS divergence, see Appendix A of Abbott et al. (2021a).

This criterion reduces our GWTC-2 sample to 42 informative binaries. We restrict this study to these informative events. Note that we do not quote the JS divergence for the population-informed analysis as the 40 events we analyze under this assumption will, by definition, cross this threshold as they carry more information. Table 2 lists the JS divergence for the events that crossed the 0.007 threshold. GW190814 has the highest JS divergence and also has the most informative kick posterior in the entire GWTC-2 catalog.

REFERENCES

Abbott, R., et al. 2021a, Phys. Rev. X, 11, 021053

—. 2021b, Astrophys. J. Lett., 913, L7

Antonini, F., & Gieles, M. 2020, Mon. Not. Roy. Astron. Soc., 492, 2936

Antonini, F., Gieles, M., & Gualandris, A. 2019, Mon. Not. Roy. Astron. Soc., 486, 5008

Antonini, F., & Rasio, F. A. 2016, Astrophys. J., 831, 187

Apostolatos, T. A., Cutler, C., Sussman, G. J., & Thorne, K. S. 1994, Phys. Rev. D, 49, 6274

Arca-Sedda, M., & Gualandris, A. 2018, Mon. Not. R. Astron. Soc., 477, 4423

Ashton, G., et al. 2019, Astrophys. J. Suppl., 241, 27

Barausse, E., & Rezzolla, L. 2009, Astrophys. J. Lett., 704, L40

Bellovary, J., Brooks, A., Colpi, M., et al. 2019, BAAS, 51, 175

Calderón Bustillo, J., Clark, J. A., Laguna, P., & Shoemaker, D. 2018, Phys. Rev. Lett., 121, 191102

Calderon Bustillo, J., Evans, C., Clark, J. A., et al. 2020, Communications Physics, 3, 176

Campanelli, M., Lousto, C. O., Zlochower, Y., & Merritt, D. 2007a, Phys. Rev. Lett., 98, 231102

—. 2007b, Astrophys. J. Lett., 659, L5

Cotesta, R., Buonanno, A., Bohé, A., et al. 2018, Phys. Rev. D, 98, 084028

Doctor, Z., Farr, B., & Holz, D. E. 2021, Astrophys. J. Lett., 914, L18

Doctor, Z., Wysocki, D., O'Shaughnessy, R., Holz, D. E., & Farr, B. 2020, ApJ, 893, 35

Downing, J. M. B., Benacquista, M. J., Giersz, M., & Spurzem, R. 2010, Mon. Not. Roy. Astron. Soc., 407, 1946

Favata, M. 2009, J. Phys. Conf. Ser., 154, 012043

Table 2. JS divergence between the posterior and prior distributions of the kick magnitude for different events in GWTC-2 using uninformed priors. Higher values of JS divergence imply more informative kick posteriors, with GW190814 being the most informative one in GWTC-2. Following Abbott et al. (2021a), a threshold of 0.007 on the JS divergence is imposed, and only those events that pass the threshold are presented in the table. For a Gaussian distribution, this threshold corresponds to a 20% shift in its mean.

Event name	JS divergence	Event name	JS divergence	Event name	JS divergence
GW190814	0.4329	GW190521_074359	0.0460	GW170729	0.0203
GW190412	0.1579	GW151226	0.0408	GW190512_180714	0.0190
GW190924_021846	0.1158	GW190620_030421	0.0405	GW190408_181802	0.0159
GW190521	0.1092	GW190602_175927	0.0401	GW190708_232457	0.0159
GW190517_055101	0.0917	GW190720_000836	0.0364	GW190727_060333	0.0153
GW190514_065416	0.0913	GW190929_012149	0.0361	GW190527_092055	0.0133
GW190930_133541	0.0900	GW170818	0.0301	GW190701_203306	0.0114
GW190909_114149	0.0605	GW190424_180648	0.0300	GW190910_112807	0.0114
GW190728_064510	0.0588	GW190731_140936	0.0293	GW190503_185404	0.0108
GW190828_065509	0.0560	GW190421_213856	0.0288	GW190706_222641	0.0100
GW190413_134308	0.0521	GW190828_063405	0.0281	GW150914	0.0090
GW190719_215514	0.0471	GW190803_022701	0.0261	GW170104	0.0086
GW190519_153544	0.0471	GW190630_185205	0.0242	GW190413_052954	0.0084
GW190915_235702	0.0467	GW170814	0.0224	GW170823	0.0082

Favata, M., Hughes, S. A., & Holz, D. E. 2004, Astrophys. J. Lett., 607, L5

Fishbach, M., & Holz, D. E. 2020, Astrophys. J. Lett., 891, L27

Fishbach, M., Holz, D. E., & Farr, B. 2017, Astrophys. J. Lett., 840, L24

Fitchett, M. J. 1983, Mon. Not. R. Astron. Soc., 203, 1049

Fragione, G., Ginsburg, I., & Kocsis, B. 2018a, Astrophys. J., 856, 92

Fragione, G., & Kocsis, B. 2018, Phys. Rev. Lett., 121, 161103

Fragione, G., Leigh, N., Ginsburg, I., & Kocsis, B. 2018b, Astrophys. J., 867, 119

Fragione, G., & Loeb, A. 2021, Mon. Not. Roy. Astron. Soc., 502, 3879

Fragione, G., & Silk, J. 2020, Mon. Not. Roy. Astron. Soc., 498, 4591

Georgiev, I. Y., Böker, T., Leigh, N., Lützgendorf, N., & Neumayer, N. 2016, MNRAS, 457, 2122

Gerosa, D., & Berti, E. 2017, Phys. Rev. D, 95, 124046

Gerosa, D., & Kesden, M. 2016, Phys. Rev. D, 93, 124066

Gerosa, D., & Moore, C. J. 2016, Phys. Rev. Lett., 117, 011101

Gerosa, D., & Sesana, A. 2015, Mon. Not. Roy. Astron. Soc., 446, 38

Gerosa, D., Vitale, S., & Berti, E. 2020, Phys. Rev. Lett., 125, 101103

Gonzalez, J., Hannam, M., Sperhake, U., Bruegmann, B., & Husa, S. 2007a, Phys. Rev. Lett., 98, 231101

Gonzalez, J. A., Sperhake, U., Bruegmann, B., Hannam, M., & Husa, S. 2007b, Phys. Rev. Lett., 98, 091101

Hannam, M., Schmidt, P., Bohé, A., et al. 2014, Phys. Rev. Lett., 113, 151101

Harris, W. E. 1996, AJ, 112, 1487

Kidder, L. 1995, Phys. Rev. D, 52, 821

Kimball, C., Talbot, C., L. Berry, C. P., et al. 2020, Astrophys. J., 900, 177

Kimball, C., et al. 2021, Astrophys. J. Lett., 915, L35

Kullback, S., & Leibler, R. A. 1951, The Annals of Mathematical Statistics, 22, 79 . https://doi.org/10.1214/aoms/1177729694

LIGO Scientific Collaboration. 2018, LIGO Algorithm Library - LALSuite, doi:10.7935/GT1W-FZ16. https://doi.org/10.7935/GT1W-FZ16

LIGO Scientific Collaboration and Virgo Collaboration. 2020, GWTC-2 data release, doi:10.7935/99gf-ax93. https://doi.org/10.7935/99gf-ax93

Lin, J. 1991, IEEE Transactions on Information Theory, 37, 145

Lousto, C. O., & Healy, J. 2019, Phys. Rev. D, 100, 104039

Lousto, C. O., & Zlochower, Y. 2008, Phys. Rev. D, 77, 044028

—. 2011, Phys. Rev. Lett., 107, 231102

—. 2013, Phys. Rev. D, 87, 084027

Lousto, C. O., Zlochower, Y., Dotti, M., & Volonteri, M. 2012a, Phys. Rev. D, 85, 084015

—. 2012b, Phys. Rev. D, 85, 084015

Mapelli, M., Santoliquido, F., Bouffanais, Y., et al. 2020, arXiv:2007.15022

Merritt, D., Milosavljević, M., Favata, M., Hughes, S. A., & Holz, D. E. 2004, Astrophys. J, 607, L9

Miller, M. C., & Hamilton, D. P. 2002, Mon. Not. Roy. Astron. Soc., 330, 232

Miller, M. C., & Lauburg, V. M. 2009, Astrophys. J., 692, 917

Moore, C. J., & Gerosa, D. 2021, arXiv:2108.02462

Ossokine, S., et al. 2020, Phys. Rev. D, 102, 044055

Petrovich, C., & Antonini, F. 2017, Astrophys. J., 846, 146

Portegies Zwart, S. F., & McMillan, S. 2000, Astrophys. J. Lett., 528, L17

Racine, E. 2008, Phys. Rev. D, 78, 044021

Rodriguez, C. L., Amaro-Seoane, P., Chatterjee, S., et al. 2018, Phys. Rev. D, 98, 123005

Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016, Phys. Rev. D, 93, 084029

Rodriguez, C. L., Morscher, M., Pattabiraman, B., et al. 2015, Phys. Rev. Lett., 115, 051101, [Erratum: Phys.Rev.Lett. 116, 029901 (2016)]

Romero-Shaw, I. M., et al. 2020, Mon. Not. Roy. Astron. Soc., 499, 3295

Samsing, J. 2018, Phys. Rev. D, 97, 103014

Tichy, W., & Marronetti, P. 2007, Phys. Rev. D, 76, 061502

Tiwari, V., & Fairhurst, S. 2021, Astrophys. J. Lett., 913, L19

Varma, V., Field, S. E., Scheel, M. A., et al. 2019a, Phys. Rev. Research., 1, 033015

Varma, V., Gerosa, D., Stein, L. C., Hébert, F., & Zhang, H. 2019b, Phys. Rev. Lett., 122, 011101

Varma, V., Isi, M., & Biscoveanu, S. 2020, Phys. Rev. Lett., 124, 101104

Zevin, M., Samsing, J., Rodriguez, C., Haster, C.-J., & Ramirez-Ruiz, E. 2019, Astrophys. J., 871, 91