

Single-event likelihood of star cluster properties with LIGO-Virgo-Kagra binary black hole observations

Ken K. Y. Ng^{1,2,*}, Konstantinos Kritos, Andrea Antonelli, Roberto Cotesa, and Emanuele Berti^{1,2}

William H. Miller III Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA

(Received 17 July 2023; accepted 12 September 2023; published 12 October 2023)

The population of binary black hole mergers observed in gravitational waves, together with astrophysical simulations, can help us to understand the properties of the progenitors and the binary formation mechanisms in different astrophysical scenarios. Here we focus on dynamical formation in star clusters. We use the third gravitational-wave transient catalog (GWTC-3) and RAPSTER, a rapid code to simulate cluster dynamics, to show that it is possible to construct the single-event likelihood of star cluster properties from individual observations. We find that the measured primary mass in a binary black hole merger correlates with the measured star cluster mass, because the mass spectrum of the primary component increases with the mass of the cluster. This trend may be caused by two physical mechanisms: (i) the more efficient production of hierarchical mergers with primary mass above $\sim 40M_{\odot}$ for cluster masses of $\gtrsim 10^6M_{\odot}$, and (ii) the suppression of more massive first-generation binaries, which happens because ejected binaries do not merge within the lookback time for cluster masses of $\lesssim 10^5M_{\odot}$. The formalism presented here can be generalized to infer the population properties of binary progenitors in more realistic scenarios involving multiple formation channels.

DOI: 10.1103/PhysRevD.108.083015

I. INTRODUCTION

The recent catalogs of gravitational-wave (GW) transients released by the LIGO-Virgo-Kagra Collaboration [1–6] motivated efforts to investigate the properties of the binary black hole (BBH) population and their possible formation channels. There are various ways to address this problem.

One approach is to construct *phenomenological models* that reproduce the main distinctive features of astrophysical formation channels [7–9]. This is a sensible approach because it requires minimal astrophysical modeling.

For instance, certain broad features of the population—such as the BH spin alignment with the orbital angular momentum [10–18] or the measurement of binary component masses populating the pair instability supernova (PISN) mass gap, above $\sim 40M_{\odot}$ [19–23]—may provide evidence for multiple formation channels. This is because isolated binary evolution is expected to produce mostly binaries with aligned spins and masses below the mass gap [24–38], while the spins of BHs produced through dynamical formation in star clusters should be isotropically oriented, and hierarchical mergers can populate the PISN mass gap [39–44,44–53]. Other features that can be captured by phenomenological models include the time

(or redshift) evolution of the merger rate density [54–59], or the presence of peaks and tails in the redshift evolution of merger rate densities due to putative Population III or primordial BBH components, which could be observable with next-generation GW detectors [58,60].

One drawback of phenomenological models is that they are affected by modeling systematics: for example, the class of parametrized functions used to reproduce the data may be too restrictive, leading to an erroneous mapping between the parametrized models and detailed astrophysical simulations.

A second approach is to infer the empirical distribution using *data-driven models*, leaving the interpretation of the resulting distribution to the postprocessing stage [61–64]. Even within this approach, finding a suitable statistical metric connecting data with astrophysical simulations could be problematic.

A third approach (and one that we follow in this paper) is to build a *direct mapping* between the measured parameters of a BBH merger event and the observables predicted by astrophysical simulations [65–72]. While the inference is still limited by our incomplete knowledge of astrophysical formation scenarios, this approach allows for in-depth studies of the astrophysical mechanisms that correspond to certain features seen in the populations. There have been many attempts to infer some of the key astrophysical parameters affecting the isolated binary evolution scenario, as well as the relative contribution (or branching ratios)

*kng15@jhu.edu

of multiple formation channels: see, e.g., [73–78] for an incomplete list.

In this paper we avoid the complications related to multiple formation channels, and we focus on the dynamical formation scenario in dense star clusters. Each cluster in the star cluster population has different properties, and therefore it produces a different BBH subpopulation. Here we develop a two-level hierarchical Bayesian framework that can ultimately infer the properties of the star cluster population from BBH merger observations (Sec. II). We focus on the first step in this framework, which consists in constructing the single-event likelihood of star cluster properties: i.e., we aim to identify clusters with parameters which are more likely to generate a particular BBH observed in the third GW transient catalog (GWTC-3) [3,79]. To this end, we use a code for rapid simulations of cluster dynamics, RAPSTER [44], to build a statistical mapping between the BBH parameters and the star cluster parameters. By analyzing these simulations we observe a positive correlation between the measured BBH primary mass and the inferred cluster mass. As we discuss in Sec. III, this correlation may be explained by the cluster mass scaling of the efficiency in the production of hierarchical mergers and by the inspiral timescale of the ejected binaries. In Sec. IV we discuss some technical aspects and future prospects to interpret the observed BBH population using astrophysical simulations. In Appendix we give details on the cluster simulations and on the kernel density estimation (KDE) we use to approximate the joint distribution from the simulated mergers.

II. TWO-LEVEL HIERARCHICAL BAYESIAN FRAMEWORK

As the Universe evolves, numerous star clusters form with redshift-dependent rates and with different physical properties (such as total mass, radius, and metallicity) [80]. Each cluster evolves dynamically and produces an ensemble of BBHs whose statistical distribution depends on the properties of the host cluster. Therefore, the distribution of BBH properties observed by LVK in the cluster scenario should be modeled by considering the population of BBHs originating from *a population of star clusters*. The “inverse problem” consists of inferring the properties of the star cluster population that can host BBH mergers observable in GW detectors.

One may attempt to perform the full hierarchical analysis by simulating BBHs drawing from different realizations of cluster populations. However, it is more beneficial to consider a two-level hierarchical model that can break down the inference, as follows.

In the first level of hierarchy, we map the single-event likelihood of BBH parameters to the likelihood of parameters of individual clusters that are likely to produce them, using the BBH properties predicted by star cluster simulations. In the second level of hierarchy, we combine these

single-event cluster likelihoods and infer the distribution of the cluster properties.

To see how the single-event cluster likelihood enters the hierarchical framework, we derive it using a top-down approach. The full hierarchical likelihood based on a Poisson process of data generation is given by [81,82]

$$p(\Lambda^{\text{cl}} | \mathbf{d}) \propto e^{-N_{\text{det}}^{\text{B}}(\Lambda^{\text{cl}})} \prod_{i=1}^N \int p(d_i | \theta_i^{\text{B}}) \frac{dN^{\text{B}}}{d\theta^{\text{B}}}(\theta_i^{\text{B}} | \Lambda^{\text{cl}}) d\theta_i^{\text{B}}, \quad (1)$$

where $\mathbf{d} = \{d_i\}_{i=1}^N$ is the dataset of N BBH observations, $p(d_i | \theta_i^{\text{B}})$ is the individual likelihood of the i th BBH characterized by parameters θ_i^{B} such as component masses and spins, $dN^{\text{B}}/d\theta^{\text{B}}$ is the differential number of BBHs expected for a given cluster population characterized by hyperparameters Λ^{cl} , and $N_{\text{det}}^{\text{B}}$ is the number of detectable BBHs:

$$N_{\text{det}}^{\text{B}}(\Lambda^{\text{cl}}) = \int \frac{dN^{\text{B}}}{d\theta^{\text{B}}}(\theta^{\text{B}} | \Lambda^{\text{cl}}) \epsilon_{\text{det}}(\theta^{\text{B}}) d\theta^{\text{B}}, \quad (2)$$

where $0 \leq \epsilon_{\text{det}}(\theta^{\text{B}}) \leq 1$ is the detection efficiency for a BBH merger with binary parameters θ^{B} .

The differential rate can be written as

$$\begin{aligned} \frac{dN^{\text{B}}}{d\theta^{\text{B}}}(\theta^{\text{B}} | \Lambda^{\text{cl}}) &= \int \frac{d^2 N^{\text{B}}}{d\theta^{\text{B}} d\lambda^{\text{cl}}}(\theta^{\text{B}}, \lambda^{\text{cl}} | \Lambda^{\text{cl}}) d\lambda^{\text{cl}} \\ &= \int p(\theta^{\text{B}} | \lambda^{\text{cl}}) \eta(\lambda^{\text{cl}}) \frac{dN^{\text{cl}}}{d\lambda^{\text{cl}}}(\lambda^{\text{cl}} | \Lambda^{\text{cl}}) d\lambda^{\text{cl}}, \end{aligned} \quad (3)$$

where $p(\theta^{\text{B}} | \lambda^{\text{cl}})$ is the distribution of θ^{B} originating from a single cluster characterized by some λ^{cl} , $\eta(\lambda^{\text{cl}})$ is the number of BBHs produced by the cluster, and N^{cl} is the total number of clusters. For example, λ^{cl} could be the mass of a single cluster, and Λ^{cl} the power law index of the cluster mass function. Both $p(\theta^{\text{B}} | \lambda^{\text{cl}})$ and $\eta(\lambda^{\text{cl}})$ are predicted by the simulation, while $p(d_i | \theta_i^{\text{B}})$ is obtained from GW observations.

The integral in Eq. (1) is equivalent to the expected number of BBHs averaged over the individual BBH likelihoods. It can be rewritten as

$$\begin{aligned} \langle N^{\text{B}} \rangle_i(\Lambda^{\text{cl}}) &= \iint p(d_i | \theta_i^{\text{B}}) p(\theta_i^{\text{B}} | \lambda_i^{\text{cl}}) \eta(\lambda_i^{\text{cl}}) \frac{dN^{\text{cl}}}{d\lambda^{\text{cl}}}(\lambda_i^{\text{cl}} | \Lambda^{\text{cl}}) d\lambda_i^{\text{cl}} d\theta_i^{\text{B}} \\ &= \int p(d_i | \lambda_i^{\text{cl}}) \eta(\lambda_i^{\text{cl}}) \frac{dN^{\text{cl}}}{d\lambda^{\text{cl}}}(\lambda_i^{\text{cl}} | \Lambda^{\text{cl}}) d\lambda_i^{\text{cl}}, \end{aligned} \quad (4)$$

where the individual cluster likelihood of the i th observation, $p(d_i | \lambda_i^{\text{cl}})$, is the marginalization of $p(\theta_i^{\text{B}} | \lambda_i^{\text{cl}})$ over the individual BBH likelihoods:

$$p(d_i|\lambda_i^{\text{cl}}) \equiv \int p(d_i|\theta_i^{\text{B}})p(\theta_i^{\text{B}}|\lambda_i^{\text{cl}})d\theta_i^{\text{B}}. \quad (5)$$

This result may also be obtained by applying Bayes' theorem and marginalizing over θ_i^{B} on the joint distribution $p(d_i, \theta_i^{\text{B}}|\lambda_i^{\text{cl}})$ "directly" in the bottom-up approach.

This procedure is practically advantageous as we only need to approximate $p(\theta^{\text{B}}|\lambda^{\text{cl}})$ and $\eta(\lambda^{\text{cl}})$ once with the finite samples produced by the simulations. On the contrary, the emulation of the entire BBH population $dN^{\text{B}}/d\theta^{\text{B}}$ originating from all possible cluster populations is limited to the choices of the prior functions used in the training set, and thus hinders the use of the more flexible nonparametric models for the cluster population that we are interested in. In the following, we will study solely the single-event cluster likelihood in Eq. (5) for selected events in GWTC-3.

III. INDIVIDUAL LIKELIHOOD OF CLUSTER PROPERTIES

The best-measured BBH parameters in current GW observations, and therefore the parameters that are most informative in the inference of formation channels, are the (source-frame) masses of the primary, m_1 , and secondary, m_2 ; the effective spin projected along the orbital angular momentum, χ_{eff} ; and the redshift, z .

The RAPSTER code has a total of 19 input parameters. While it is hopeless to constrain all of these parameters, as a proof of principle we explore two of the most important intrinsic properties of individual clusters: the total mass of the cluster, M_{cl} , and the half-mass radius at the time of cluster formation, r_h . In other words, we set $\theta^{\text{B}} = (m_1, m_2, \chi_{\text{eff}}, z)$ and $\lambda^{\text{cl}} = (M_{\text{cl}}, r_h)$ in the formalism of Sec. II. We reweigh the LVK posterior samples and obtain the likelihood samples of θ^{B} .

We limit the cluster parameter space to the ranges $M_{\text{cl}} \in [10^4, 10^7]M_{\odot}$ and $r_h \in [0.5, 3]$ pc, respectively, based on current observations of young star clusters [80]. The simulation samples for constructing the KDE are generated by the following settings. The initial cluster masses M_{cl} are drawn from a power-law distribution with a spectral index -2 in the range $[10^{3.7}, 10^{7.3}]M_{\odot}$. To avoid hard cutoffs in the range $[10^4, 10^7]M_{\odot}$ where we construct our KDE, we taper the distribution using a Tukey window function with shape parameter 0.18. The initial half-mass radius r_h is drawn from a linear distribution in the range $[0.3, 3]$ pc. This choice is to balance the number of mergers per cluster in the simulation set, which scales with the inverse of the cluster radius. In the inference, we obtain the likelihood of M_{cl} and r_h by reweighing the chosen priors. We note that the above initial conditions are reweighed out eventually and that they do not affect the evaluation of the likelihood, as shown in Eq. (5).

The other cluster parameters and the initial cluster conditions are fixed as follows. We use SEVN to compute

the initial mass function of BHs so that the PISN cut-off is at $\sim 40M_{\odot}$, with the exact value depending on the metallicity [83]. The dimensionless natal spin of first-generation BHs is sampled from a uniform distribution in the range $[0, 0.2]$, as in Ref. [84]. The masses and spins of BBH merger remnants are computed using the PRECESSION code [85]. The initial central stellar density is calculated as $3M_{\text{cl}}/(4\pi(r_h/1.3)^3)$, assuming a Plummer profile [86]. This choice of mass-radius relation is motivated by observations of star clusters in the local Universe [80]. A detailed cross-validation of the cluster population resulting from this assumption with other surveys would be interesting, but is beyond the scope of this paper.

Moreover, we assume a young massive cluster population with the redshift of cluster formation and mean metallicity sampled from the Madau-Fragos distribution [87]. We also apply a log-normal spread with a variance of 0.3 in the metallicity distribution at each redshift. The rest of the input parameters are set to their default RAPSTER values, as listed in Table I of Ref. [44]. With one node (48 processors) at the Maryland Advanced Research Computing Center at Johns Hopkins, we can simulate $\sim 10^6$ star clusters within 2 days.

To approximate the conditional probability distribution $p(\theta^{\text{B}}|\lambda^{\text{cl}})$, we employ Gaussian KDE on a set of $\sim 7 \times 10^5$ simulated BBH mergers generated by the synthesis code RAPSTER. Given $p(\theta^{\text{B}}|\lambda^{\text{cl}})$, we can evaluate Eq. (5) for each BBH observation in GWTC-3 released by the LVK Collaboration. Since the integral is generally intractable, we sample the likelihood in Eq. (5) by Monte Carlo methods. Technical details on the KDE and on the integration are given in Appendix.

In Fig. 1 we show the joint likelihoods of (M_{cl}, r_h) for two events: GW190521 [88], and GW151226 [89]. These events were chosen because they have very different primary masses: GW190521 has $m_1 \sim 100M_{\odot}$, suggestive of a hierarchical merger origin [90], while GW151226 has $m_1 \sim 14M_{\odot}$, a more typical value for events in the GWTC-3 catalog [9].

We find that the likelihood for r_h is almost uninformative even for GW190521. This is because, in the RAPSTER simulations [44], the compactness of the cluster mostly affects the number of BBHs produced in the cluster, i.e. η , rather than the shape of the BBH mass distribution. As η is not involved in the single-event cluster likelihood, we do not extract any new information about r_h from the single-event analysis. However, we note that the current version of RAPSTER does not include stellar mergers, which would allow for the possibility to form initial BHs within the PISN mass gap. This mechanism may skew the distribution of m_1 to higher values for very compact clusters [52].

The key feature of Fig. 1 is that the likelihood of M_{cl} in the two systems is very different: GW190521 favors $M_{\text{cl}} \gtrsim 10^6M_{\odot}$, and GW151226 favors $M_{\text{cl}} \lesssim 10^6M_{\odot}$. This may hint at a positive correlation between the primary BH mass

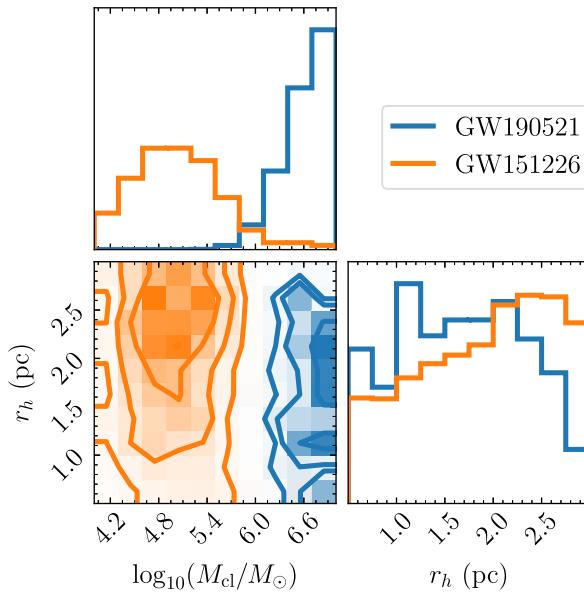


FIG. 1. The joint likelihoods of (M_{cl}, r_h) for GW190521 (blue) and GW191226 (orange). The marginalized M_{cl} likelihood for GW190521 (GW191226) favors values above (below) $M_{\text{cl}} \sim 10^6 M_{\odot}$, while the r_h likelihoods are mostly flat.

m_1 and the probable cluster mass M_{cl} that produces the corresponding BBH merger event.

To test this hypothesis, we have analyzed all BBH events in GWTC-3. The results of this analysis are shown in Fig. 2, where we show the inferred value of M_{cl} as a function of m_1 for the GWTC-3 catalog. Most events are in gray, but a selected subset (listed in the legend) is highlighted in color. The highlighted subset is chosen to cover three different ranges of m_1 : values in the PISN mass gap (GW190521 and GW191109, with $m_1 \gtrsim 40 M_{\odot}$), events

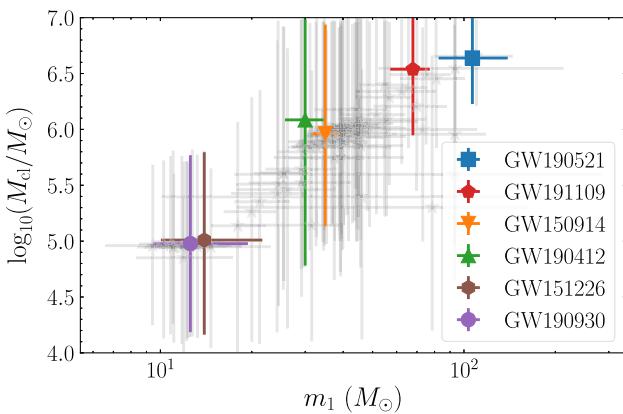


FIG. 2. The 90% credible intervals of the likelihood of m_1 (horizontal bars) and M_{cl} (vertical bars) for all BBH events in GWTC-3, with markers indicating the median values of m_1 and M_{cl} . Six events (GW190521, GW191109, GW190412, GW150914, GW151226, and GW190930) are highlighted in color to better illustrate the correlation between m_1 and M_{cl} .

with $40 \gtrsim m_1 \gtrsim 20 M_{\odot}$ (GW150914 and GW190412), and low-primary mass events with $m_1 \lesssim 20 M_{\odot}$ (GW151226 and GW190930). As anticipated, we observe that the inferred values of M_{cl} (within the 90% credible interval) tend to increase as a function of the measured values of m_1 . We have also checked that the correlation of M_{cl} with other binary parameters (such as χ_{eff} and q) is not as significant as the correlation with m_1 .

To understand this correlation, in Fig. 3 we compare the primary mass distributions $p(m_1)$ generated by clusters having masses M_{cl} in different ranges (highlighted by histograms in different colors). These mass distribution histograms have two major features.

First of all, the relative fraction of BBHs above the PISN mass gap is larger when $M_{\text{cl}} \geq 10^5 M_{\odot}$ (i.e., for the orange and green histograms). Hierarchical mergers within the mass gap occur more frequently in more massive clusters, because these clusters have larger escape velocities and thus they are more likely to retain the merger remnants despite their gravitational recoils. This is compatible with the correlation between M_{cl} and primary BHs having $m_1 \gtrsim 40 M_{\odot}$ observed in Fig. 2.

Second, the mass distribution of first-generation mergers below the PISN mass gap (those with $m_1 \lesssim 40 M_{\odot}$) is skewed toward lower values for $M_{\text{cl}} \leq 10^5 M_{\odot}$ (blue histogram): for these light clusters, the peak in m_1 decreases from $\sim 35 M_{\odot}$ to $\sim 15 M_{\odot}$. This trend may be qualitatively explained by a combination of the ejection mechanism discussed above, and the finite merging time window. In a star cluster, first-generation mergers are typically formed by a combination of mass segregation and exchange interactions. The majority of first-generation mergers are nearly equal-mass systems, whose critical semimajor axis for getting ejected out of the cluster after

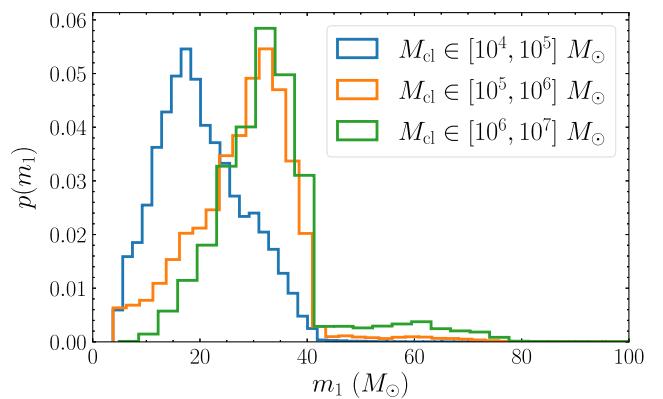


FIG. 3. Histograms of the primary mass distributions generated by clusters with masses in the range $M_{\text{cl}} \in [10^4, 10^5] M_{\odot}$ (blue), $[10^5, 10^6] M_{\odot}$ (orange), and $[10^6, 10^7] M_{\odot}$ (green), using a prior $p(M_{\text{cl}}) \propto M_{\text{cl}}^{-2}$. The peak location of the m_1 distribution, which represents the most probable first-generation mergers, shifts from $m_1 \sim 15 M_{\odot}$ to $m_1 \sim 35 M_{\odot}$ as M_{cl} increases from $10^4 M_{\odot}$ to $10^6 M_{\odot}$.

a binary-single interaction scales with $a_{\text{ej}} \propto m_1/M_{\text{cl}}$: see e.g. Eq. (8) in Ref. [44], or Eq. (8) in Ref. [91]. Therefore, in less massive clusters the more massive BBHs are ejected at an earlier stage of their inspiral evolution. Since the GW inspiral timescale $\tau \propto a_{\text{ej}}^4/m_1^3$ [92], the typical inspiral time for ejected mergers has the scaling $\tau \propto m_1/M_{\text{cl}}^4$. As the binaries can only merge within the (finite) cosmic time since their formation, the critical m_1 below which BBHs can merge scales with $m_1 \propto M_{\text{cl}}^4$. This leads to the observed shift in the primary mass distribution as M_{cl} decreases. Note that this is only a qualitative explanation, and the quantitative correlation between M_{cl} and m_1 for first-generation BHs is very likely model-dependent. The ejection efficiency and the resulting merging timescale are sensitive to nonlinear effects in cluster dynamics, to the formation redshift and to the cluster metallicity. All of these effects may modify the shape of the primary mass distribution at different merger redshifts.

IV. DISCUSSION

In this paper we propose a two-level hierarchical framework to analyze the BBH population observed in GWs under the assumption that the merger events are produced dynamically in star clusters. The two-level hierarchy is based on the idea that each cluster in the cluster population has different physical properties, and therefore produces a different BBH population. The implication is that we can not only characterize the BBH population produced by an assumed cluster population, but (vice versa) we can also infer the population properties of the clusters, given a physical model of cluster dynamics.

To carry out this hierarchical inference, we first need to perform single-event inference—that is, we need to identify the cluster properties that are most likely to produce any observed BBH merger event. Estimating the single-event likelihood of cluster properties requires a knowledge of how the distribution of BBH parameters (such as binary masses, spins, and redshift) depends on the individual cluster properties, such as the cluster mass and radius. In this paper, we carried out a proof-of-principle demonstration that this hierarchical inference is possible using, for illustrative purposes, astrophysical models built with the RAPSTER code. With RAPSTER we can simulate the formation of BBHs in star clusters, and then approximate the joint distribution of BBH parameters and cluster parameters from the simulated BBH samples by KDE methods.

We find that the inferred cluster mass is correlated with the measured BBH primary mass, as shown in Fig. 2. The correlation is a result of the variation of the primary mass distribution as a function of cluster mass observed in Fig. 3: more massive clusters enhance the production of hierarchical mergers above the PISN mass gap, and less massive clusters eject more massive first-generation binaries at semimajor axes that are too large to efficiently produce

mergers. As emphasized in the main text, this is a mostly qualitative explanation: the extent of the shift of the primary mass distribution in first-generation BBHs is sensitive to the details of cluster dynamics and to the initial conditions of cluster formation (including redshift and metallicity).

For cluster radii $r_h \in [0.5, 3]$ pc, the radius only affects the overall number of BBHs produced in each cluster (because stellar mergers are not been included in the model), and it has a negligible effect on the primary mass distribution. In a more realistic scenario, the radius should play an important role in the inference of cluster population properties, because the number of BBHs produced per cluster η entering the next hierarchy depends on r_h , and affects the expected BBH merger rate. Our results are again model-dependent, and therefore it would be useful to validate this trend by comparing against other existing codes, especially for low-mass clusters with $M_{\text{cl}} \lesssim 10^5 M_{\odot}$, whose evolution is more sensitive to the details of every dynamical process [39,43,44,47,93]. Another source of uncertainty is the initial binary fraction, which we assumed to be 10% in the simulations. While differences in the initial binary star population have an impact on the exchange channel in RAPSTER, the effect of this population on the black hole mass-cluster mass correlation is probably more dependent on the physics of binary star evolution. The study of this effect would require complementing RAPSTER with input from a binary star evolution code to simulate common envelope and stable mass transfer for the original binary stars. This is an interesting problem, but it is beyond the scope of our work.

While the KDE method we employed suffices to capture the broad features discussed above, it may not be robust enough to proceed to the next hierarchy and infer the properties of the cluster population. Performing the full hierarchical analysis may require more advanced techniques, such as deep generative modeling, to better approximate the multidimensional probability density functions involved in Eq. (5). For example, one may learn $p(\theta^B | \lambda^{\text{cl}})$ and $\eta(\lambda^{\text{cl}})$ separately by simulating the BBH populations given a set of cluster properties $\{\lambda^{\text{cl}}\}$, or work with the joint distribution $p(\theta^B, \lambda^{\text{cl}} | \tilde{\Lambda}^{\text{cl}})$ at a chosen cluster population characterized directly by $\tilde{\Lambda}^{\text{cl}}$, and obtain $p(\theta^B | \lambda^{\text{cl}}) \eta(\lambda^{\text{cl}}) \propto p(\theta^B, \lambda^{\text{cl}} | \tilde{\Lambda}^{\text{cl}}) / p(\lambda^{\text{cl}} | \tilde{\Lambda}^{\text{cl}})$ by reweighting the chosen prior of the cluster population.

Finally, we note that the two-level hierarchy may be generalized to include contributions from multiple formation channels. This may be achieved by using the relevant parametrization to set up Eq. (4) for each channel and build a mixture model in Eq. (1). For example, one may obtain the likelihood of progenitor redshift and metallicity based on binary evolution simulations for galactic field binaries (either by backpropagation or though other numerical techniques, as in Refs. [67,68]); combine with the similar likelihood of cluster binaries; and trace the evolution of star formation rate, cluster formation rate, and stellar metallicity

all at once. The full hierarchical inference will be presented in future work.

ACKNOWLEDGMENTS

We thank Davide Gerosa, Will Farr, Chase Kimball, Sharan Banagiri, and Vicky Kalogera for fruitful discussions. K. K. Y. N. is supported by the Miller Fellowship at Johns Hopkins University and the Croucher Fellowship by the Croucher Foundation. K. K., A. A., R. C., and E. B. are supported by NSF Grants No. AST-2006538, No. PHY-2207502, No. PHY-090003, and No. PHY-20043, by NASA Grants No. 20-LPS20-0011 and No. 21-ATP21-0010, and by the John Templeton Foundation Grant 62840. K. K. Y. N., K. K., A. A., R. C., and E. B. acknowledge support from the ITA-USA Science and Technology Cooperation program, supported by the Ministry of Foreign Affairs of Italy (MAECl), and from the Indo-US Science and Technology Forum through the Indo-US Centre for Gravitational-Physics and Astronomy, grant IUSSTF/JC-142/2019. This research project was conducted using computational resources at the Maryland Advanced Research Computing Center (MARCC). This research has made use of data or software obtained from the Gravitational Wave Open Science Center (gwosc.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. K. A. G. R. A. is supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS) in Japan; National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea;

Academia Sinica (AS) and National Science and Technology Council (NSTC) in Taiwan.

APPENDIX: CONSTRUCTION OF KDE AND IMPORTANCE SAMPLING

RAPSTER enables rapid simulations of BBHs generated from a population of clusters. Therefore, it is more convenient to (i) choose a cluster population (parametrized by $\tilde{\Lambda}_{\text{cl}}$) that produces enough simulated BBHs in the range of the BBH parameter space θ^B (our simulated samples), (ii) perform kernel density estimation (KDE) to approximate the joint distribution $\tilde{p}(\theta^B, \lambda^{\text{cl}} | \tilde{\Lambda}_{\text{cl}})$ from the simulated mergers, and (iii) obtain $p(\theta^B | \lambda^{\text{cl}}) = \tilde{p}(\theta^B, \lambda^{\text{cl}} | \tilde{\Lambda}_{\text{cl}}) / \tilde{p}(\lambda^{\text{cl}} | \tilde{\Lambda}_{\text{cl}})$ from Bayes' theorem.

Here, $\tilde{p}(\lambda^{\text{cl}} | \tilde{\Lambda}_{\text{cl}})$ contains a factor of the differential merger rate, $\eta(\lambda^{\text{cl}})$, on top of the chosen prior of the cluster population put into the simulations, since each input cluster may produce a different number of mergers. To ensure that $p(\theta^B | \lambda^{\text{cl}})$ is properly normalized to unity for each λ^{cl} , we require a second KDE for $\tilde{p}(\lambda^{\text{cl}} | \tilde{\Lambda}_{\text{cl}}) \propto \eta(\lambda^{\text{cl}}) p(\lambda^{\text{cl}} | \tilde{\Lambda}_{\text{cl}})$, which can also be constructed from the simulated samples, because the count of λ^{cl} is proportional to the differential merger rate. We employ `gaussian_kde` from `scipy`, written in the `jax` infrastructure, to speed up the KDE. We use $\sim 7 \times 10^5$ simulation points, with a bandwidth of ~ 0.25 set by Silverman's rule.

One may attempt to first approximate the integral of Eq. (5) by an importance sum over the θ^B samples of individual BBH likelihoods, and then draw the λ^{cl} sample from the approximated λ^{cl} likelihood by Monte Carlo methods. Since the parameter space of λ^{cl} has a relatively low dimension, we simplify the sampling procedures further by importance sampling of the joint distribution $p(d_i, \theta_i^B | \lambda^{\text{cl}}) \equiv p(d_i | \theta_i^B) p(\theta_i^B | \lambda^{\text{cl}})$. In practice, we append an additional set of λ_i^{cl} samples, $\{\lambda_{i,j}^{\text{cl}}\}_{j=1}^K$, drawn from a uniform distribution $U(\lambda_i^{\text{cl}}) \propto 1$ to the set of BBH likelihood samples, $\{\theta_{i,j}^B\}_{j=1}^K$. The aggregated set $\{\theta_{i,j}^B, \lambda_{i,j}^{\text{cl}}\}_{j=1}^K$ follows the joint distribution $p(d_i | \theta_i^B) U(\lambda_i^{\text{cl}})$. The desired set of λ_i^{cl} samples that are representative of the marginalized likelihood $p(d_i | \lambda_i^{\text{cl}})$ is equivalent to the set of $\{\lambda_{i,j}^{\text{cl}}\}_{j=1}^K$ weighed by $\{w_{i,j}^{\text{cl}} \propto p(\theta_{i,j}^B | \lambda_{i,j}^{\text{cl}})\}_{j=1}^K$.

[1] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
[2] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Phys. Rev. X* **11**, 021053 (2021).
[3] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), [arXiv:2111.03606](https://arxiv.org/abs/2111.03606) [Phys. Rev. X (to be published)].
[4] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).

[5] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).

[6] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto (KAGRA Collaboration), *Phys. Rev. D* **88**, 043007 (2013).

[7] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **882**, L24 (2019).

[8] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Astrophys. J. Lett.* **913**, L7 (2021).

[9] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Phys. Rev. X* **13**, 011048 (2023).

[10] D. Gerosa, M. Kesden, E. Berti, R. O’Shaughnessy, and U. Sperhake, *Phys. Rev. D* **87**, 104028 (2013).

[11] S. Vitale, R. Lynch, R. Sturani, and P. Graff, *Classical Quantum Gravity* **34**, 03LT01 (2017).

[12] B. Farr, D. E. Holz, and W. M. Farr, *Astrophys. J. Lett.* **854**, L9 (2018).

[13] W. M. Farr, S. Stevenson, M. Coleman Miller, I. Mandel, B. Farr, and A. Vecchio, *Nature (London)* **548**, 426 (2017).

[14] K. K. Y. Ng, S. Vitale, A. Zimmerman, K. Chatzioannou, D. Gerosa, and C.-J. Haster, *Phys. Rev. D* **98**, 083007 (2018).

[15] C. Talbot and E. Thrane, *Phys. Rev. D* **96**, 023012 (2017).

[16] D. Wysocki, D. Gerosa, R. O’Shaughnessy, K. Belczynski, W. Gladysz, E. Berti, M. Kesden, and D. E. Holz, *Phys. Rev. D* **97**, 043014 (2018).

[17] M. Fishbach, C. Kimball, and V. Kalogera, *Astrophys. J. Lett.* **935**, L26 (2022).

[18] M. Fishbach and G. Fragione, *Mon. Not. R. Astron. Soc.* **522**, 5546 (2023).

[19] D. Gerosa and E. Berti, *Phys. Rev. D* **95**, 124046 (2017).

[20] M. Fishbach, D. E. Holz, and B. Farr, *Astrophys. J. Lett.* **840**, L24 (2017).

[21] C. Talbot and E. Thrane, *Astrophys. J.* **856**, 173 (2018).

[22] A. M. Farah, B. Edelman, M. Zevin, M. Fishbach, J. M. Ezquiaga, B. Farr, and D. E. Holz, [arXiv:2301.00834](https://arxiv.org/abs/2301.00834).

[23] C. Kimball, C. Talbot, C. P. L. Berry, M. Carney, M. Zevin, E. Thrane, and V. Kalogera, *Astrophys. J.* **900**, 177 (2020).

[24] K. Belczynski, V. Kalogera, F. A. Rasio, R. E. Taam, A. Zezas, T. Bulik, T. J. Maccarone, and N. Ivanova, *Astrophys. J. Suppl. Ser.* **174**, 223 (2008).

[25] K. Belczynski, D. E. Holz, T. Bulik, and R. O’Shaughnessy, *Nature (London)* **534**, 512 (2016).

[26] K. Belczynski *et al.*, *Astron. Astrophys.* **636**, A104 (2020).

[27] M. Dominik, K. Belczynski, C. Fryer, D. Holz, E. Berti, T. Bulik, I. Mandel, and R. O’Shaughnessy, *Astrophys. J.* **759**, 52 (2012).

[28] M. Dominik, K. Belczynski, C. Fryer, D. E. Holz, E. Berti, T. Bulik, I. Mandel, and R. O’Shaughnessy, *Astrophys. J.* **779**, 72 (2013).

[29] M. Dominik, E. Berti, R. O’Shaughnessy, I. Mandel, K. Belczynski, C. Fryer, D. E. Holz, T. Bulik, and F. Pannarale, *Astrophys. J.* **806**, 263 (2015).

[30] S. S. Bavera, T. Fragos, Y. Qin, E. Zapartas, C. J. Neijssel, I. Mandel, A. Batta, S. M. Gaebel, C. Kimball, and S. Stevenson, *Astron. Astrophys.* **635**, A97 (2020).

[31] K. Breivik *et al.*, *Astrophys. J.* **898**, 71 (2020).

[32] S. Stevenson, A. Vigna-Gómez, I. Mandel, J. W. Barrett, C. J. Neijssel, D. Perkins, and S. E. de Mink, *Nat. Commun.* **8**, 14906 (2017).

[33] F. S. Broekgaarden, S. Justham, S. E. de Mink, J. Gair, I. Mandel, S. Stevenson, J. W. Barrett, A. Vigna-Gómez, and C. J. Neijssel, *Mon. Not. R. Astron. Soc.* **490**, 5228 (2019).

[34] V. Baibhav, E. Berti, D. Gerosa, M. Mapelli, N. Giacobbo, Y. Bouffanais, and U. N. Di Carlo, *Phys. Rev. D* **100**, 064060 (2019).

[35] S. de Mink and K. Belczynski, *Astrophys. J.* **814**, 58 (2015).

[36] K. Belczynski *et al.*, *Astron. Astrophys.* **594**, A97 (2016).

[37] S. E. Woosley, *Astrophys. J.* **836**, 244 (2017).

[38] J. Riley *et al.* (COMPAS Team, Team COMPAS Collaborations), *Astrophys. J. Suppl. Ser.* **258**, 34 (2022).

[39] A. Askar, M. Szkudlarek, D. Gondek-Rosińska, M. Giersz, and T. Bulik, *Mon. Not. R. Astron. Soc.* **464**, L36 (2017).

[40] C. L. Rodriguez *et al.*, *Astrophys. J. Suppl. Ser.* **258**, 22 (2022).

[41] M. Mapelli, Y. Bouffanais, F. Santoliquido, M. A. Sedda, and M. C. Artale, *Mon. Not. R. Astron. Soc.* **511**, 5797 (2022).

[42] F. Antonini and M. Gieles, *Mon. Not. R. Astron. Soc.* **492**, 2936 (2020).

[43] M. A. Sedda, M. Mapelli, M. Benacquista, and M. Spera, *Mon. Not. R. Astron. Soc.* **520**, 5259 (2023).

[44] K. Kritos, V. Strokov, V. Baibhav, and E. Berti, [arXiv:2210.10055](https://arxiv.org/abs/2210.10055).

[45] C. L. Rodriguez, M. Zevin, P. Amaro-Seoane, S. Chatterjee, K. Kremer, F. A. Rasio, and C. S. Ye, *Phys. Rev. D* **100**, 043027 (2019).

[46] C. L. Rodriguez, M. Morscher, B. Pattabiraman, S. Chatterjee, C.-J. Haster, and F. A. Rasio, *Phys. Rev. Lett.* **115**, 051101 (2015); **116**, 029901(E) (2016).

[47] C. L. Rodriguez, S. Chatterjee, and F. A. Rasio, *Phys. Rev. D* **93**, 084029 (2016).

[48] C. L. Rodriguez, P. Amaro-Seoane, S. Chatterjee, and F. A. Rasio, *Phys. Rev. Lett.* **120**, 151101 (2018).

[49] Z. Doctor, D. Wysocki, R. O’Shaughnessy, D. E. Holz, and B. Farr, *Astrophys. J.* **893**, 35 (2020).

[50] U. N. Di Carlo, N. Giacobbo, M. Mapelli, M. Pasquato, M. Spera, L. Wang, and F. Haardt, *Mon. Not. R. Astron. Soc.* **487**, 2947 (2019).

[51] F. Antonini and M. Gieles, *Phys. Rev. D* **102**, 123016 (2020).

[52] K. Kremer, M. Spera, D. Becker, S. Chatterjee, U. N. Di Carlo, G. Fragione, C. L. Rodriguez, C. S. Ye, and F. A. Rasio, *Astrophys. J.* **903**, 45 (2020).

[53] D. Gerosa and M. Fishbach, *Nat. Astron.* **5**, 749 (2021).

[54] M. Fishbach, D. E. Holz, and W. M. Farr, *Astrophys. J. Lett.* **863**, L41 (2018).

[55] M. Fishbach and V. Kalogera, *Astrophys. J. Lett.* **914**, L30 (2021).

[56] M. Mapelli, N. Giacobbo, F. Santoliquido, and M. C. Artale, *Mon. Not. R. Astron. Soc.* **487**, 2 (2019).

[57] C. L. Rodriguez and A. Loeb, *Astrophys. J. Lett.* **866**, L5 (2018).

[58] K. K. Y. Ng, S. Vitale, W. M. Farr, and C. L. Rodriguez, *Astrophys. J. Lett.* **913**, L5 (2021).

[59] S. Vitale, W. M. Farr, K. Ng, and C. L. Rodriguez, *Astrophys. J. Lett.* **886**, L1 (2019).

[60] K. K. Y. Ng, G. Franciolini, E. Berti, P. Pani, A. Riotto, and S. Vitale, *Astrophys. J. Lett.* **933**, L41 (2022).

[61] I. Mandel, W. M. Farr, A. Colonna, S. Stevenson, P. Tiňo, and J. Veitch, *Mon. Not. R. Astron. Soc.* **465**, 3254 (2017).

[62] S. Rinaldi and W. Del Pozzo, *Mon. Not. R. Astron. Soc.* **509**, 5454 (2021).

[63] B. Edelman, B. Farr, and Z. Doctor, *Astrophys. J.* **946**, 16 (2023).

[64] K. W. K. Wong and M. Cranmer, in *39th International Conference on Machine Learning Conference* (2022); [arXiv:2207.12409](https://arxiv.org/abs/2207.12409).

[65] M. Zevin, C. Pankow, C. L. Rodriguez, L. Sampson, E. Chase, V. Kalogera, and F. A. Rasio, *Astrophys. J.* **846**, 82 (2017).

[66] K. W. Wong, G. Contardo, and S. Ho, *Phys. Rev. D* **101**, 123005 (2020).

[67] K. W. K. Wong, K. Breivik, W. M. Farr, and R. Luger, *Astrophys. J.* **950**, 181 (2023).

[68] J. J. Andrews, A. Zezas, and T. Fragos, *Astrophys. J. Suppl. Ser.* **237**, 1 (2018).

[69] S. R. Taylor and D. Gerosa, *Phys. Rev. D* **98**, 083017 (2018).

[70] M. Mould, D. Gerosa, and S. R. Taylor, *Phys. Rev. D* **106**, 103013 (2022).

[71] M. Mould, D. Gerosa, M. Dall'Amico, and M. Mapelli, *Mon. Not. R. Astron. Soc.* **525**, 3986 (2023).

[72] A. Antonelli, K. Kritos, K. K. Y. Ng, R. Cotesta, and E. Berti, [arXiv:2306.11088](https://arxiv.org/abs/2306.11088).

[73] S. Stevenson, F. Ohme, and S. Fairhurst, *Astrophys. J.* **810**, 58 (2015).

[74] Y. Bouffanais, M. Mapelli, D. Gerosa, U. N. Di Carlo, N. Giacobbo, E. Berti, and V. Baibhav, *Astrophys. J.* **886**, 25 (2019).

[75] K. W. Wong, K. Breivik, K. Kremer, and T. Callister, *Phys. Rev. D* **103**, 083021 (2021).

[76] K. W. K. Wong, G. Franciolini, V. De Luca, V. Baibhav, E. Berti, P. Pani, and A. Riotto, *Phys. Rev. D* **103**, 023026 (2021).

[77] M. Zevin, S. S. Bavera, C. P. Berry, V. Kalogera, T. Fragos, P. Marchant, C. L. Rodriguez, F. Antonini, D. E. Holz, and C. Pankow, *Astrophys. J.* **910**, 152 (2021).

[78] J. J. Andrews, J. Cronin, V. Kalogera, C. P. L. Berry, and A. Zezas, *Astrophys. J. Lett.* **914**, L32 (2021).

[79] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Astrophys. J. Suppl. Ser.* **267**, 29 (2023).

[80] M. R. Krumholz, C. F. McKee, and J. Bland-Hawthorn, *Annu. Rev. Astron. Astrophys.* **57**, 227 (2019).

[81] I. Mandel, W. M. Farr, and J. R. Gair, *Mon. Not. R. Astron. Soc.* **486**, 1086 (2019).

[82] S. Vitale, D. Gerosa, W. M. Farr, and S. R. Taylor, Inferring the properties of a population of compact binaries in presence of selection effects, in *Handbook of Gravitational Wave Astronomy*, edited by C. Bambi, S. Katsanevas, and K. D. Kokkotas (Springer Singapore, Singapore, 2020), pp. 1–60.

[83] M. Spera and M. Mapelli, *Mon. Not. R. Astron. Soc.* **470**, 4739 (2017).

[84] V. Baibhav, E. Berti, D. Gerosa, M. Mould, and K. W. K. Wong, *Phys. Rev. D* **104**, 084002 (2021).

[85] D. Gerosa and M. Kesden, *Phys. Rev. D* **93**, 124066 (2016).

[86] H. C. Plummer, *Mon. Not. R. Astron. Soc.* **71**, 460 (1911).

[87] P. Madau and T. Fragos, *Astrophys. J.* **840**, 39 (2017).

[88] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **125**, 101102 (2020).

[89] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 241103 (2016).

[90] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **900**, L13 (2020).

[91] F. Antonini and F. A. Rasio, *Astrophys. J.* **831**, 187 (2016).

[92] P. C. Peters, *Phys. Rev.* **136**, B1224 (1964).

[93] K. Joshi, F. Rasio, and S. F. Portegies Zwart, *Astrophys. J.* **540**, 969 (2000).