

# Classifying the generation and formation channels of individual LIGO-Virgo-KAGRA observations from dynamically formed binaries

Andrea Antonelli<sup>1,2,\*</sup>, Konstantinos Kritos, Ken K. Y. Ng<sup>1</sup>, Roberto Cotesta, and Emanuele Berti

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

(Received 19 June 2023; accepted 25 September 2023; published 23 October 2023)

We address two important questions in gravitational-wave astronomy. What is the astrophysical formation scenario leading to black-hole binary mergers? Did some of the merging black holes form hierarchically through previous generations of mergers? Leveraging fast-to-generate astrophysical simulations from the RAPSTER code and a random forest algorithm, we develop a pipeline to accurately classify the most likely generation and formation scenario of dynamically formed black holes on an event-by-event basis. We test our framework on four merger events with features suggesting a dynamical origin: the large total mass event GW190521, GW190412 (with large mass asymmetry), and two events with effective spins antialigned with the orbital angular momentum (GW191109 and GW200225). Within the models we consider, and assuming these events to be formed dynamically, we find that one of the component black holes in GW190521 formed from a previous merger with high probability ( $\gtrsim 85\%$ ). GW190521, GW191109, and GW200225 are compatible with formation through three-body interactions, while the most likely formation channel for GW190412 is two-body captures. We also rule out that GW191109 contains only first-generation black holes with a probability of 97%. Our pipeline could be useful to identify the evolutionary path of individual gravitational-wave observations once it is trained on more comprehensive sets of binary formation simulations.

DOI: 10.1103/PhysRevD.108.084044

## I. INTRODUCTION

The growing catalog of transient gravitational-wave (GW) events from the LIGO-Virgo-KAGRA (LVK) Collaboration has opened up the possibility to discern the evolutionary paths that lead to the formation of compact object binaries. The two most common scenarios for black hole (BH) binaries involve either field formation, or dynamical formation in dense stellar environments (see, e.g., [1,2] for reviews). Each scenario may involve sub-channels, such as common envelope evolution [3,4], stable mass transfer [5], and chemically homogeneous evolution [6] for isolated binaries in the field or formation in young massive star clusters [7–13], globular clusters [14–20], nuclear star clusters [21–28], open clusters [29–31], and active galactic nuclei [32,33] in the dynamical case. Moreover, a binary BH (BBH) that has formed dynamically in a collisional system may have been assembled in various dynamical subchannels, such as two-body [34], three-body [35], or binary-single exchange interactions [36].

One of the main goals of GW astrophysics is to extract information about formation scenarios from LVK data. Population synthesis codes for binaries formed in the field [37–39] and in dynamical environments [19,40–44]

typically involve a large number of hyperparameters. These are population-level parameters that must be distinguished from the source parameters that describe the binary itself, such as the masses and the spins. Examples of hyperparameters for field formation scenarios are the metallicity and efficiency of the common envelope, while for dynamical formation scenarios examples are the generation of the observed binary or the masses and radii of the clusters they form in. The common approach is to predict the properties of the population of observable mergers from astrophysical simulations with a specified set of hyperparameters, and identify features that can help to infer the hyperparameters by comparison with LVK catalogs [45–47].

Predicting datasets to match observations has the advantage of giving broad insights into the stages of the formation of LVK binaries and their host environments. The necessary tuning of the initial conditions is however not amenable to hyperparameter inference. The inverse problem (inferring the hyperparameters from data) allows one instead to gain direct insight from observations [48–50].

This kind of hyperparameter inference is only possible when the underlying astrophysical simulations are very fast to generate, either because the astrophysical simulations are fast to evaluate, or through the use of emulators. In this paper, we follow the first approach. We use the RAPSTER

\*aantonelli3@jh.edu; andrico@hotmail.it

code [44], which can compute the dynamical evolution of BHs in clusters in a matter of minutes, while retaining information from the hyperparameters that describe the cluster evolution at various stages of the simulations. We train a classification algorithm on a host of such RAPSTER datasets. Our goal is to use the individual masses and spins of the BBHs and compare with individual observation in the LVK catalog to quantify (in a statistical sense) whether the binary components formed hierarchically, and infer their most likely formation subchannel.

As RAPSTER only describes dynamically formed BHs, at present we can only apply the model to observations of binaries that we assume are formed dynamically. For this reason we test the pipeline with real GW data from events that are likely to have a dynamical origin: GW190412 (with asymmetric masses) [51,52], GW190521 (with components in the mass gap) [53], and two events with effective spins antialigned with the orbital angular momentum, GW191109 and GW200225 [54]. However the approach we describe in this paper can be applied to other population synthesis codes (whether the binaries are formed in the field or dynamically), as long as the event catalogs are reasonably fast to generate. By including more information in the training stage, the present pipeline could be fruitfully applied to LVK data to discern hyperparameters of binary systems on an individual-event basis.

The paper is structured as follows. In Sec. II we describe the RAPSTER datasets. In Sec. III we discuss the classifier’s underlying model and validation. In Sec. IV we apply the model to selected events from the LVK catalog, and we discuss our main results. In Sec. V we present some conclusions and directions for future work.

## II. SIMULATIONS

In this section we describe the model and code used to perform our simulations, as well as the initial conditions for the population of clusters we simulate (Sec. II A). Then we establish some notation for hierarchical mergers (Sec. II B) and dynamical BBH formation channels (Sec. II C) that will be useful in the rest of the paper.

### A. The model

We use RAPSTER to generate the simulations needed to train our pipeline. RAPSTER is a population synthesis code for dynamical formation of BBHs in star clusters. The code simulates the BH subsystem in the core of the cluster in a semianalytic way. It uses the SEVN code to compute BH masses given the initial mass function of stars assuming a metallicity value [55], and the PRECESSION code to estimate the properties (mass and spin) of the merger remnants [56]. The RAPSTER code requires as input the initial conditions of a star cluster, such as the initial cluster mass, initial half-mass radius, metallicity, and the redshift of cluster

formation. It also assumes a model for the spin of first-generation (or 1g) BHs, i.e., BHs that form from the collapse of massive stars. For this work, the relevant output from RAPSTER is the “mergers” file, which contains information about the dynamical mergers that occurred throughout the history of a single cluster simulation (see the documentation in Ref. [44]).

The initial conditions for our simulated populations of clusters are chosen as follows. Motivated by astronomical observations [57], we draw the initial cluster mass  $M_{\text{cl},0}$  from a Schechter-type distribution of the form  $p(M_{\text{cl},0}) \propto M_{\text{cl},0}^{-2} \exp(-M_{\text{cl},0}/M_t)$  in the range  $[10^4, 10^7] M_\odot$ . Given current uncertainties, we set the truncation mass  $M_t$  to  $10^9, 10^6, 10^{5.3} M_\odot$ , a choice also adopted in Ref. [58]. We also set the initial half-mass radius to 0.5, 1.0 or 2.0 pc. The initial central stellar density is set to the Plummer value:  $\rho_{\star,0} = 3M_{\text{cl},0}/(4\pi(r_{\text{h},0}/1.3)^3)$ , while the initial galactocentric radius of the cluster is sampled between 0 and 20 kpc assuming a Sérsic profile with Sérsic index  $n = 1$  and scale radius of 1 kpc. We sample the redshift of cluster formation from the Madau-Fragos distribution [59], and for the metallicity we use the fit in Eq. (6) of the same reference. Finally, we sample the spin of 1g BHs from a uniform distribution  $U[0, \chi_{\text{max}}]$ , with  $\chi_{\text{max}} \in \{0.0, 0.2, 0.5\}$ . All other initial conditions for RAPSTER are set to their default values. These choices result in a grid of  $3 \times 3 \times 3 = 27$  simulations for the cluster population. However, for illustrative purposes we chose to focus on only seven of these models, that will be discussed in the figures below and in Table I. We denote these models by  $S_x M_y r_z$ , where the three subscripts  $(x, y, z)$  refer to the values of  $\chi_{\text{max}}$ ,  $\log_{10} M_t$ , and  $r_{\text{h},0}/\text{pc}$ , respectively. The reason for this choice is to explore the sensitivity of our results to the choice of initial conditions of our cluster population, and not to exhaustively simulate every possible astrophysical situation.

### B. Hierarchical mergers

After the merger of two BHs, the remnant typically has a nonvanishing recoil velocity (or “kick”) due to asymmetric GW emission [60]. If the GW kick does not exceed the escape velocity of the host cluster, then the merger remnant is retained in the cluster, and it can potentially merge again. We will refer to the merger remnant of two 1g BHs as a “second generation” (or 2g) BH. The 2g BH, if retained, may merge with either a 1g or another 2g BH. Thus we broadly identify three categories of mergers: “1g + 1g,” “1g + 2g,” and “2g + 2g.” In principle, it is possible that mergers involving even higher-generation BHs could occur. However we find that the fraction of events with the next most popular flavor (the “1g + 3g” mergers) only represent  $\lesssim 0.1\%$  of our populations. Thus, in this study we will neglect 3g or higher-generation BHs.

TABLE I. Classification results of the analysis for the GW190521, GW190412, GW191109, and GW200225 events. To guide the eye, we mark with a superscript \* those binary configurations or formation subchannels that are most favored in our analysis, as well as the highest probabilities in each case we studied. We find GW190521 to be most likely composed of two second-generation BHs formed through three body interactions; GW190412 to be most likely a result of two-body captures (though with a smaller probability  $\sim 50\%$ ); and GW191109 and GW200225 to be most likely formed through three-body interactions. GW191109 is most likely a 2g + 2g binary, with a very small chance that it contains a 1g BH. Our analysis is unable to uniquely identify the generations of the BHs in GW190412 and GW200225, even though for the latter it seems plausible that no 1g BH is present in the binary.

Model	GW190521						GW190412							
	$\hat{p}_{1g+1g}$	$\hat{p}_{1g+2g}$	$\hat{p}_{2g+2g}$	$\hat{p}_{3bb}$	$\hat{p}_{exc}$	$\hat{p}_{cap}$	$\hat{p}_{ZLK}$	$\hat{p}_{1g+1g}$	$\hat{p}_{1g+2g}$	$\hat{p}_{2g+2g}$	$\hat{p}_{3bb}$	$\hat{p}_{exc}$	$\hat{p}_{cap}$	$\hat{p}_{ZLK}$
$S_{0.0}M_{9.0}r_{1.0}$	1%	24%	75%*	82%*	13%	2%	3%	0%	20%	80%*	2%	49%*	38%	11%
$S_{0.2}M_{5.3}r_{1.0}$	11%	44%	45%*	73%*	27%	0%	0%	31%	69%*	0%	40%	1%	58%*	0%
$S_{0.2}M_{6.0}r_{1.0}$	11%	43%	46%*	72%*	18%	2%	8%	29%	56%*	15%	39%	0%	61%*	0%
$S_{0.2}M_{9.0}r_{0.5}$	3%	28%	69%*	88%*	10%	2%	0%	0%	26%	73%*	34%	2%	64%*	0%
$S_{0.2}M_{9.0}r_{1.0}$	4%	34%	62%*	83%*	15%	2%	0%	1%	38%	61%*	35%	0%	65%*	0%
$S_{0.2}M_{9.0}r_{2.0}$	10%	37%	53%*	91%*	7%	2%	0%	4%	81%*	15%	38%	0%	61%*	0%
$S_{0.5}M_{9.0}r_{1.0}$	15%	49%*	36%	91%*	6%	3%	0%	52%*	46%	2%	36%	0%	64%*	0%

Model	GW191109						GW200225							
	$\hat{p}_{1g+1g}$	$\hat{p}_{1g+2g}$	$\hat{p}_{2g+2g}$	$\hat{p}_{3bb}$	$\hat{p}_{exc}$	$\hat{p}_{cap}$	$\hat{p}_{ZLK}$	$\hat{p}_{1g+1g}$	$\hat{p}_{1g+2g}$	$\hat{p}_{2g+2g}$	$\hat{p}_{3bb}$	$\hat{p}_{exc}$	$\hat{p}_{cap}$	$\hat{p}_{ZLK}$
$S_{0.0}M_{9.0}r_{1.0}$	0%	14%	86%*	92%*	7%	0%	1%	2%	25%	73%*	57%*	3%	40%	0%
$S_{0.2}M_{5.3}r_{1.0}$	2%	37%	61%*	69%*	31%	0%	0%	12%	88%*	0%	61%*	9%	30%	0%
$S_{0.2}M_{6.0}r_{1.0}$	2%	36%	62%*	73%*	21%	0%	5%	11%	49%*	40%	55%*	4%	41%	0%
$S_{0.2}M_{9.0}r_{0.5}$	0%	19%	81%*	94%*	5%	1%	0%	4%	30%	66%*	49%*	13%	38%	0%
$S_{0.2}M_{9.0}r_{1.0}$	0%	24%	76%*	85%*	13%	0%	2%	5%	39%	56%*	53%*	4%	43%	0%
$S_{0.2}M_{9.0}r_{2.0}$	2%	30%	68%*	92%*	8%	0%	0%	6%	55%*	39%	52%*	3%	43%	2%
$S_{0.5}M_{9.0}r_{1.0}$	3%	42%	55%*	91%*	7%	1%	0%	21%	53%*	26%	48%	3%	49%*	0%

### C. Dynamical formation subchannels

BHBs may form in RAPSTER in one of the following three ways:

- Three-body binary formation (3bb): the close interaction of three single BHs can result in the formation of a BBH if the third BH extracts enough kinetic energy from the system.
- Dynamical capture (cap): if the GW energy released during the hyperbolic encounter of two single BHs exceeds the kinetic energy of the two-body system at infinity, this may result in a highly eccentric and compact BBH.
- Exchanges (exc): the formation of a BBH occurs through a pair of exchange interactions of the form star-star  $\rightarrow$  BH-star  $\rightarrow$  BH-BH pair, where each BH substitutes each stellar component in succession.

We denote a merger occurring from these three formation channels by “3bb,” “cap,” or “exc,” respectively. We also include von Zeipel-Lidov-Kozai (ZLK) mergers, which involve the formation of hierarchical triples from BBH-BBH interactions. In the ZLK channel, the inner binary merges as a result of angular momentum and inclination oscillations in the orbit. This channel can lead to values of the eccentricity close to unity [61].

### III. MODEL TRAINING AND VALIDATION

We train a machine-learning algorithm with the datasets from RAPSTER described above. The goal is to (i) approximate a function  $\mathbf{y}_{(i)} = f(\mathbf{X}_{(i)})$  with predictors  $\mathbf{X}_{(i)}$  that can be readily linked to samples from GW observations, and (ii) predict variables  $\mathbf{y}_{(i)} = \{y_{gen}, y_{fc}\}$  corresponding to the binary components’ generation and to the formation subchannels. The full RAPSTER datasets contain 19 cluster parameters and nine BBH parameters. The simulations keep track of the generation of the individual black holes, from which we create a binary parameter associated to the variable  $y_{gen} = \{1g + 1g, 1g + 2g, 2g + 2g\}$  that we aim to predict. Similarly, one of the cluster parameters from RAPSTER is the formation subchannels, which we pick as the second variable to predict, i.e.,  $y_{fc} = \{3bb, exc, cap, ZLK\}$ . From the remaining RAPSTER parameters, not all of them are necessary to accurately train the function. By trial and error and inspection of the correlation matrix in search for high correlations, we find that the component masses  $m_{1,2}$  and dimensionless spin parameters  $\chi_{1,2}$  are the most important parameters to include in the training datasets. Both sets of parameters are in fact measured by the LVK Collaboration, which makes the connection to observations direct. In the evaluation of the classifier, we use the posterior samples on

$m_{1,2}$  and  $\chi_{1,2}$ . When compared to the likelihood, the posterior on  $m_{1,2}$  has a factor of  $(1+z)^2$  difference, due to the conversion of the flat prior on detector-frame masses to that on source-frame masses. This discrepancy is negligible in our analysis since most of the sources are local ( $z \lesssim 1$ ), but it must be accounted for as the detectors improve, allowing for the detection of farther sources. As an illustration, we only calculate the portions of the posterior samples that fall into each subclass. The probabilities presented in the following sections should not be interpreted as posterior probabilities, because the prior of the source parameters used to infer the LVK posteriors is not the same as the prior based on the RAPSTER simulations used by the classifier. We will comment more on this in the conclusions.

The predictor set  $\mathbf{X}_{\text{gen}} = \{m_1, m_2, \chi_1, \chi_2\}$  is sufficient to train accurate models for the binary components' generations. For the formation channel, we reparametrize the set to  $\{\log M_{\text{tot}}, q, \chi_1, \chi_2\}$  to aid the numerical fitting. Here  $M_{\text{tot}} = m_1 + m_2$  is the total mass of the binary, and  $q = m_1/m_2 \geq 1$  is the mass ratio. We find that two more hyperparameters must be added to the predictor set for  $\mathbf{y}_{\text{fc}}$ : the cluster formation redshift  $z_{\text{cl}}$  and initial cluster mass  $M_{\text{cl}}$ . The cluster mass  $M_{\text{cl}}$  anticorrelates most strongly with  $\mathbf{y}_{\text{fc}}$ , with Pearson coefficient  $\sigma_{\mathbf{y}_{\text{fc}}, M_{\text{cl}}} = -0.35$ , while  $z_{\text{cl}}$  correlates most strongly with the log of the total mass, with  $\sigma_{\log M_{\text{tot}}, M_{\text{cl}}} = 0.21$ . The predictor set then reads as  $\mathbf{X}_{\text{fc}} = \{\log M_{\text{tot}}, q, \chi_1, \chi_2, z_{\text{cl}}, M_{\text{cl}}\}$ .

At this stage, we need an appropriate algorithm for the classification problem. We decided to choose Random Forest algorithms for their interpretability, as they can be thought of as ensemble averages of decision trees. To avoid overfitting, we make sure that the individual decision trees do not “grow too large” by requiring each tree to have at most three nodes. We use the implementation from SKLEARN, and train our models on 70% of a randomly selected subset of RAPSTER simulated binaries (with sample size  $N \sim 10^5$ ). We compare the predictions of this emulator against a test set consisting of the remaining 30% of the binaries. For a reference training set  $S_{0.0}M_{9.0}r_{1.0}$ , we find that the pipeline can accurately identify all 28252 “1g + 1g,” 1444 “1g + 2g,” and 295 “2g + 2g” test binaries, while correctly identifying the formation subchannel in 74% of the cases. Similar numbers are obtained when varying the training set.

Such a remarkable accuracy for the binary's generations should not be a surprise, as RAPSTER clearly separates the various configurations of parent generations in the  $(\chi_1, \chi_2)$  plane. The decision boundaries drawn by the classifier are such that for  $\chi_1 \lesssim 0.5$  and  $\chi_2 \lesssim 0.5$ , the classifier predicts 1g + 1g configurations; for  $\chi_1 \gtrsim 0.5$  and  $\chi_2 \gtrsim 0.5$ , it predicts 2g + 2g; and it predicts 1g + 2g otherwise.

#### IV. APPLICATION TO GW DATA

Our approach to connect cluster simulations with LVK data is exemplified in Fig. 1 for the case in which we want

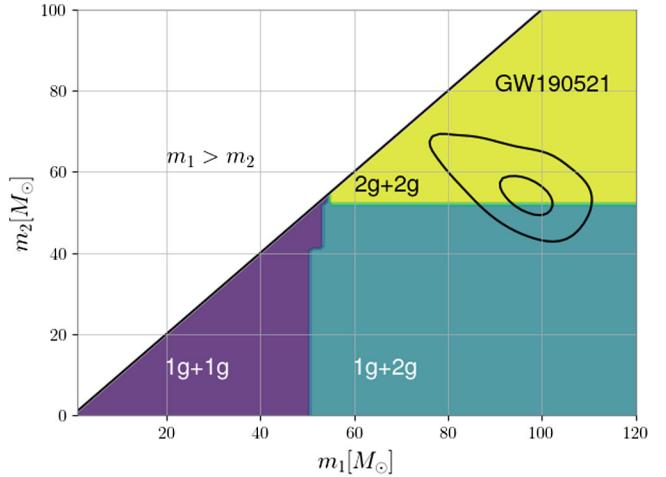


FIG. 1. We plot the partitioning of the  $(m_1, m_2)$  parameter space by the binary configurations under the fiducial model  $S_{0.0}M_{9.0}r_{1.0}$ . We also show the  $1\sigma$  and  $2\sigma$  contour plots for GW190521 [62]. By using the event samples as input in the fit of Sec. II, we can statistically establish the most likely binary configuration for GW190521.

to predict the binary configuration. Here we show the partitioning by the classifier's decision boundaries, restricting the focus on the  $(m_1, m_2)$  parameter space for illustrative purposes. We also plot the  $1\sigma$  and  $2\sigma$  contours for the GW190521 event, chosen here as an example. By inputting the event's samples into the fit, we can establish what binary configuration it most likely pertains to by checking how many samples fall within each decision boundary. To predict the formation subchannel we follow a similar approach. In this case, we must further specify some values for  $z_{\text{cl}}$  and  $M_{\text{cl}}$ , both of which appear in  $\mathbf{X}_{\text{fc}}$ . We choose  $M_{\text{cl}} = \{10^4, 3 \times 10^4, 10^5, 10^6, 10^7\} M_{\odot}$  and  $z_{\text{cl}} = \{0.5, 1, 1.5, 2, 3\}$ . These values cover the parameter space of our population of star clusters, the majority of which form at  $z_{\text{cl}} \sim 2$  (see Sec. II A). The final probabilities are obtained by averaging over pairs of  $(z_{\text{cl}}, M_{\text{cl}})$ . For each event, we repeat the procedure varying the model hyperparameters that specify the training dataset.

The models presented above are trained on datasets that only include binary formation in clusters. The probabilities are in fact conditioned on the hypothesis that the input samples pertain to a dynamically formed BH. This current limitation of the procedure can be bypassed by training the fits to data from other formation channels. However, there are reasons to believe that certain GW events have indeed formed dynamically (though we do not claim with certainty that the ones reported are). In the analysis below we will illustrate the power of the approach by focusing on these specific events. We will start by discussing GW190521, arguably the most plausible event of dynamical origin.

#### A. GW190521

The GW190521 BBH components have estimated masses  $m_1 = 85^{+21}_{-14} M_{\odot}$  and  $m_2 = 66^{+17}_{-18} M_{\odot}$ ; hence both

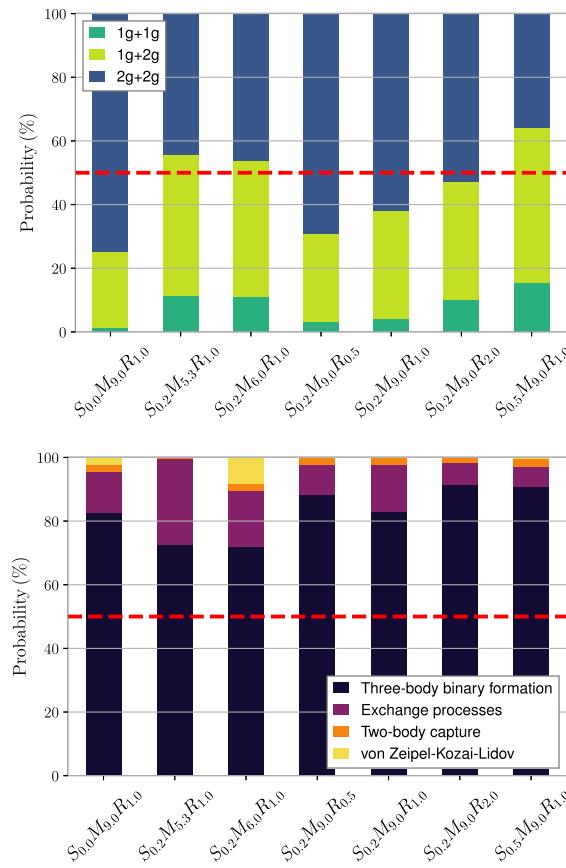


FIG. 2. Top panel: probability that the GW190521 event is a 1g + 1g, 1g + 2g, or 2g + 2g binary. The binary is unlikely to contain only 1g BHs. Bottom panel: probability that it was formed through three-body interactions, exchange processes, two-body capture, or the ZLK mechanisms. Three-body interactions are the most likely subchannel among those considered in this paper. In each panel we show the 50% probability as a dashed red line.

components have high probability of being within the upper mass gap [53]. This suggests that the binary could have formed dynamically through hierarchical mergers in clusters [63,64] (though it could also have formed from isolated binaries [65] or even in a primordial BH scenario [66]).

In Fig. 1 we plot the LVK samples for this event over the simulation data points of model  $S_{0.0}M_{9.0}r_{1.0}$ . The binary generation classifier predicts probabilities of 1%, 24%, and 75% that the merger resulted from a 1g + 1g, 1g + 2g, and 2g + 2g binary, respectively. This quantifies what is visually apparent from Fig. 1, i.e., that most configuration data points (and all those within  $2\sigma$ ) are compatible with at least one 2g BH overlapping with the LVK samples. These conclusions can change as we vary the underlying model, as we show in Fig. 2 (and in Table I, where we list results for all of the events we consider).

In many cases the 2g + 2g configuration has a probability larger than 50%, but in some cases the probability

drops to  $\sim 30\%$  (top panel of Fig. 2). This suggests that we cannot make a robust prediction within the context of the RAPSTER models. For example, the 2g + 2g probability is quite low for the model  $S_{0.5}M_{9.0}R_{1.0}$ , because in that case 1g BHs are born with higher spins. Higher spins result in large recoil velocities of the merger remnant, and therefore in lower retention probability of 2g BHs. Note, however, that there is still a high probability ( $\gtrsim 85\%$ ) that *at least one* of the merging BHs in GW190521 is 2g.

A similar analysis for the formation channels is shown in the bottom panel of Fig. 2. In this case, a comparison with the RAPSTER models implies that the most likely formation channel (with a probability  $> 70\%$ ) is three-body interactions. This is because GW190521 has a “typical” formation path in the RAPSTER datasets, and 3bb is the most likely dynamical channel according to our RAPSTER models.

## B. GW190412

The GW190412 event exhibits a high asymmetry in the masses ( $q \sim 1/3$ ) that points to a possible dynamical formation [51,52]. However, running the event’s samples through the pipeline described above leads to inconclusive results on the hierarchical character of the binary components. The predictions in Table I are very sensitive to changes in the underlying simulations: in some models the presence of a 1g BH is strongly disfavored (with  $P_{1g+1g} \lesssim 1\%$ ), while in others it is the most likely scenario (with  $P_{1g+1g} \gtrsim 50\%$ ). This suggests that, under the assumptions used in the RAPSTER models, slightly asymmetric mass ratios *alone* are not a good discriminant to determine whether one BH results from a previous merger. The loophole in this case is simulations of the “ $S_{0.5}M_yR_z$ ” class. In these simulations, a smaller fraction of merger remnants is retained in the cluster, and the low-mass ratio region is still dominated by 1g + 1g mergers (see, e.g., Fig. 11 in Ref. [44]). In these scenarios, the predictor will favor 1g + 1g over 1g + 2g despite the low-mass ratio, simply because the 1g + 1g rate is higher.

The formation subchannel analysis is more insightful. Across all the models we explored, the probability that the binary is formed through two-body capture is  $\sim 60\%$ . Again, RAPSTER can help us gain insight on the physics of the problem. In the capture scenario the lighter component is typically exchanged, but on a timescale much *longer* than the merger timescale. Exchanges typically lead to the ejection of the lightest component, and to a binary mass ratio closer to unity. In two-body captures there is a small probability to equalize the masses, and the binaries tend to have some residual mass asymmetry. Conversely, in the second most favored scenario (3bb), exchanges that tend to equalize the masses happen on timescales *shorter* than the merger timescale, leading to fewer configurations exhibiting mass asymmetry. This explains why the 3bb scenario is disfavored, but not negligibly so.

### C. GW191109 and GW200225

Reference [54] suggested that the GW191109 and GW200225 events may be formed through hierarchical mergers, as suggested by their mostly negative effective spins. Moreover, GW191109 is massive enough to suggest it was formed hierarchically. Our analysis excludes that GW191109 contains a 1g BH ( $p_{1g+1g} < 3\%$ ), with a slight preference for a 2g + 2g configuration ( $p_{2g+2g} > 50\%$ ).

There is a strong preference for a 3bb formation channel for this event, with  $p_{3bb}$  ranging from  $\sim 70\%$  to  $90\%$ . Again, this is consistent with our picture of cluster dynamics, where the heavier 2g BHs sink into the core of the cluster (if retained) and pair up with another BH through a three-body interaction.

Our predictions on the hierarchical origin of the GW200225 event are more inconclusive. In particular, the  $S_{0.2}M_{5.3}r_{1.0}$  model appears to skew the predictions. Our cluster population in this case is dominated by low-mass clusters with smaller escape velocities, that do not easily retain their 2g products. The most likely formation channel is still the three-body subchannel, although only marginally: we predict  $p_{3bb} \sim 50\%$ , which is comparable to the probabilities for the two-body capture scenario.

This high sensitivity to the formation models is a reminder that detailed analyses with more comprehensive astrophysical models are needed to draw more definite conclusions on the evolutionary history of each of the GW events analyzed in this section.

## V. CONCLUSIONS

Finding out the hierarchical origin and formation scenarios of BHs remains one of the key challenges in GW astrophysics. In this paper we have presented a pipeline that, leveraging machine-learning classification algorithms and fast-to-generate astrophysical simulations, can be applied to GW data to establish the most likely formation scenario of a given event, and the generation of the binary components. Our pipeline is trained on cluster simulation datasets from RAPSTER. It uses mass and spin samples from GW observations as input, and returns key metrics for the binary evolutionary history as output.

We have applied the pipeline to GW190412, GW190521, GW191109, and GW200225, all of which have a reasonable probability (based on physical considerations) of being formed dynamically [51,52,54,63,64].

Within the models we consider, and assuming that these events are indeed formed dynamically, we draw the following conclusions:

- (1) We exclude that GW190521 is composed of only 1g BHs ( $p_{1g+1g} \lesssim 15\%$ ), while it is likely to be formed through three-body interactions ( $p_{3bb} \gtrsim 70\%$ ).
- (2) We find that GW190412 is formed through two-body captures with  $p_{cap} \sim 60\%$ .

- (3) We rule out that GW191109 contains only 1g BHs ( $p_{1g+1g} \lesssim 3\%$ ), with three-body interactions being the most likely scenario ( $p_{3bb} \gtrsim 70\%$ ).
- (4) We find that GW200225 may result from three-body interactions with  $p_{3bb} \gtrsim 50\%$ .

Note that our method accounts for the relative rate of different classes of formation scenarios. For example, the classifier favored a 1g + 1g merger for GW190412 when we assumed a higher spinning population of 1g BHs, because in that case the tail of the 1g + 1g distribution overcomes the relative rate of 1g + 2g binaries in the low-mass ratio regime. Moreover, the 3bb subchannel has a high probability of being associated to the events because it is typically the most likely mechanism for the formation of a BBH in a star cluster, even though there is a degeneracy where both subchannels populate roughly the same region in the  $(m_1, m_2)$  parameter space. Note that the probabilities we present in this paper are not Bayesian posterior probabilities, because the LVK posteriors do not use the same simulation-based priors that we use here. One may train an emulator to approximate the astrophysical prior generated by the simulations to fix this problem. However, the use and validation of the emulator may then become the bottleneck of the inference. An optimized way to compute both the prior weight and the classification of each subchannel should be explored further.

As pointed out in Sec. II B, in this analysis we have neglected third- or higher-generation BHs, because they represent a small fraction of our population. However, since 3g BHs are typically heavier than 2g BHs, most mergers with a 3g component would populate the region above  $m_1 \sim 70M_\odot$ , that is currently underpopulated. Our fit is effectively an extrapolation when we consider LVK samples in those underpopulated regions. More simulation points and the inclusion of higher-generation BHs would be required to probe the high-mass region with greater accuracy.

Our results depend on the astrophysical models used to train the classification algorithm, and (more specifically) on the choice of the RAPSTER model to generate the cluster simulations. However, our approach should be considered as a proof of principle, and it is modular. Follow-up studies considering other simulation codes, such as CMC [16] and MOCCA [19], are needed before drawing definite conclusions on the origin of the binaries we analyzed above. Our method is flexible enough to accommodate these improvements, as long as the underlying simulations are fast to generate. We also plan to extend the method to include field-formation scenarios. By coupling the classifier to a larger set of astrophysical simulations, it should be possible to draw more solid conclusions on the astrophysical origin of GW observations on an event-by-event basis. Finally, it is important to investigate the systematics involved in the choice of the classification algorithm. For instance, it would be interesting to see if a neural network infrastructure would yield similar results.

The supporting data and code for this paper are openly available in [67].

## ACKNOWLEDGMENTS

We thank Floor Broekgaarden, Davide Gerosa, Matt Mould, Veome Kapil, and Mike Zevin for very useful discussions. K. K. Y. N. is supported by a Miller Fellowship at Johns Hopkins University and Croucher Fellowship by the Croucher Foundation. A. A., K. K., R. C., and E. B. are supported by NSF Grants No. AST-2006538, PHY-2207502, PHY-090003, and PHY-20043; by NASA Grants No. 20-LPS20-0011 and 21-ATP21-0010; and by the John Templeton Foundation Grant No. 62840. We acknowledge support from the ITA-USA Science and Technology Cooperation program, supported by the Ministry of Foreign Affairs of Italy (MAEAI), 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

[68], 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. KAGRA 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.

---

[1] J. Abadie *et al.* (LIGO Scientific and Virgo Collaborations), Predictions for the rates of compact binary coalescences observable by ground-based gravitational-wave detectors, *Classical Quantum Gravity* **27**, 173001 (2010).

[2] I. Mandel and F. S. Broekgaarden, Rates of compact object coalescences, *Living Rev. Relativity* **25**, 1 (2022).

[3] M. Dominik, K. Belczynski, C. Fryer, D. Holz, E. Berti, T. Bulik, I. Mandel, and R. O’Shaughnessy, Double compact objects I: The significance of the common envelope on merger rates, *Astrophys. J.* **759**, 52 (2012).

[4] F. S. Broekgaarden *et al.*, Impact of massive binary star and cosmic evolution on gravitational wave observations II: Double compact object rates and properties, *Mon. Not. R. Astron. Soc.* **516**, 5737 (2022).

[5] A. Olejak, C. L. Fryer, K. Belczynski, and V. Baibhav, The role of supernova convection for the lower mass gap in the isolated binary formation of gravitational wave sources, *Mon. Not. R. Astron. Soc.* **516**, 2252 (2022).

[6] J. Riley, I. Mandel, P. Marchant, E. Butler, K. Nathaniel, C. Neijssel, S. Shortt, and A. Vigna-Gomez, Chemically homogeneous evolution: A rapid population synthesis approach, *Mon. Not. R. Astron. Soc.* **505**, 663 (2021).

[7] U. N. Di Carlo, N. Giacobbo, M. Mapelli, M. Pasquato, M. Spera, L. Wang, and F. Haardt, Merging black holes in young star clusters, *Mon. Not. R. Astron. Soc.* **487**, 2947 (2019).

[8] S. Banerjee, Stellar-mass black holes in young massive and open stellar clusters and their role in gravitational-wave generation—II, *Mon. Not. R. Astron. Soc.* **473**, 909 (2018).

[9] S. Banerjee, On the formation of exotic, massive, stellar-remnant black holes at solar and sub-solar metallicities through evolution of massive binaries, [arXiv:1912.06022](https://arxiv.org/abs/1912.06022).

[10] S. Banerjee, Binary black hole mergers from young massive clusters in the pair-instability supernova mass gap, *Astron. Astrophys.* **665**, A20 (2022).

[11] B. M. Ziosi, M. Mapelli, M. Branchesi, and G. Tormen, Dynamics of stellar black holes in young star clusters with different metallicities—II. Black hole–black hole binaries, *Mon. Not. R. Astron. Soc.* **441**, 3703 (2014).

[12] F. P. Rizzuto, T. Naab, R. Spurzem, M. Arca-Sedda, M. Giersz, J. P. Ostriker, and S. Banerjee, Black hole mergers in compact star clusters and massive black hole formation beyond the mass gap, *Mon. Not. R. Astron. Soc.* **512**, 884 (2022).

[13] M. A. Sedda, F. P. Rizzuto, T. Naab, J. Ostriker, M. Giersz, and R. Spurzem, Breaching the limit: Formation of GW190521-like and IMBH mergers in young massive clusters, *Astrophys. J.* **920**, 128 (2021).

[14] S. F. Kulkarni, S. McMillan, and P. Hut, Stellar black holes in globular clusters, *Nature (London)* **364**, 421 (1993).

[15] S. F. Portegies Zwart and S. McMillan, Black hole mergers in the universe, *Astrophys. J. Lett.* **528**, L17 (2000).

[16] C. L. Rodriguez, S. Chatterjee, and F. A. Rasio, Binary black hole mergers from globular clusters: Masses, merger rates, and the impact of stellar evolution, *Phys. Rev. D* **93**, 084029 (2016).

[17] J. Samsing, Eccentric black hole mergers forming in globular clusters, *Phys. Rev. D* **97**, 103014 (2018).

[18] C. L. Rodriguez, M. Zevin, P. Amaro-Seoane, S. Chatterjee, K. Kremer, F. A. Rasio, and C. S. Ye, Black holes: The next generation—repeated mergers in dense star clusters and their gravitational-wave properties, *Phys. Rev. D* **100**, 043027 (2019).

[19] A. Askar, M. Szkudlarek, D. Gondek-Rosińska, M. Giersz, and T. Bulik, MOCCA-SURVEY database—I. Coalescing binary black holes originating from globular clusters, *Mon. Not. R. Astron. Soc.* **464**, L36 (2017).

[20] M. Arca-Sedda, G. Li, and B. Kocsis, Order in the chaos—Eccentric black hole binary mergers in triples formed via strong binary-binary scatterings, *Astron. Astrophys.* **650**, A189 (2021).

[21] M. C. Miller and V. M. Lauburg, Mergers of stellar-mass black holes in nuclear star clusters, *Astrophys. J.* **692**, 917 (2009).

[22] R. M. O’Leary, B. Kocsis, and A. Loeb, Gravitational waves from scattering of stellar-mass black holes in galactic nuclei, *Mon. Not. R. Astron. Soc.* **395**, 2127 (2009).

[23] F. Antonini and F. A. Rasio, Merging black hole binaries in galactic nuclei: Implications for Advanced-LIGO detections, *Astrophys. J.* **831**, 187 (2016).

[24] B.-M. Hoang, S. Naoz, B. Kocsis, F. A. Rasio, and F. Dosopoulou, Black hole mergers in galactic nuclei induced by the eccentric Kozai–Lidov effect, *Astrophys. J.* **856**, 140 (2018).

[25] G. Fragione, E. Grishin, N. W. C. Leigh, H. B. Perets, and R. Perna, Black hole and neutron star mergers in galactic nuclei, *Mon. Not. R. Astron. Soc.* **488**, 47 (2019).

[26] M. Mapelli *et al.*, Hierarchical black hole mergers in young, globular and nuclear star clusters: The effect of metallicity, spin and cluster properties, *Mon. Not. R. Astron. Soc.* **505**, 339 (2021).

[27] H. Wang, A. P. Stephan, S. Naoz, B.-M. Hoang, and K. Breivik, Gravitational-wave signatures from compact object binaries in the Galactic Center, *Astrophys. J.* **917**, 76 (2021).

[28] M. A. Sedda, Birth, life, and death of black hole binaries around supermassive black holes: Dynamical evolution of gravitational wave sources, *Astrophys. J.* **891**, 47 (2020).

[29] J. Kumamoto, M. S. Fujii, and A. Tanikawa, Gravitational-wave emission from binary black holes formed in open clusters, *Mon. Not. R. Astron. Soc.* **486**, 3942 (2019).

[30] S. Rastello, M. Mapelli, U. N. di Carlo, G. Iorio, A. Ballone, N. Giacobbo, F. Santoliquido, and S. Torniamenti, Dynamics of binary black holes in low-mass young star clusters, *Mon. Not. R. Astron. Soc.* **507**, 3612 (2021).

[31] S. Rastello, P. Amaro-Seoane, M. Arca-Sedda, R. Capuzzo-Dolcetta, G. Fragione, and I. Tosta e Melo, Stellar black hole binary mergers in open clusters, *Mon. Not. R. Astron. Soc.* **483**, 1233 (2019).

[32] J. Samsing, I. Bartos, D. J. D’Orazio, Z. Haiman, B. Kocsis, N. W. C. Leigh, B. Liu, M. E. Pessah, and H. Tagawa, AGN as potential factories for eccentric black hole mergers, *Nature (London)* **603**, 237 (2022).

[33] H. Tagawa, Z. Haiman, and B. Kocsis, Formation and evolution of compact object binaries in AGN disks, *Astrophys. J.* **898**, 25 (2020).

[34] G. D. Quinlan and S. L. Shapiro, Dynamical evolution of dense clusters of compact stars, *Astrophys. J.* **343**, 725 (1989).

[35] J. Goodman and P. Hut, Binary–single-star scattering. V. Steady state binary distribution in a homogeneous static background of single stars, *Astrophys. J.* **403**, 271 (1993).

[36] D. C. Heggie, P. Hut, and S. L. W. McMillan, Binary–single-star scattering. VII. Hard binary exchange cross sections for arbitrary mass ratios: Numerical results and semianalytic FITS, *Astrophys. J.* **467**, 359 (1996).

[37] J. Riley *et al.* (COMPAS Team, Team COMPAS Collaborations), Rapid stellar and binary population synthesis with COMPAS, *Astrophys. J. Suppl. Ser.* **258**, 34 (2022).

[38] K. Belczynski, V. Kalogera, F. A. Rasio, R. E. Taam, A. Zezas, T. Bulik, T. J. Maccarone, and N. Ivanova, Compact object modeling with the startrack population synthesis code, *Astrophys. J. Suppl. Ser.* **174**, 223 (2008).

[39] K. Breivik *et al.*, COSMIC variance in binary population synthesis, *Astrophys. J.* **898**, 71 (2020).

[40] C. L. Rodriguez *et al.*, Modeling dense star clusters in the Milky Way and beyond with the cluster Monte Carlo code, *Astrophys. J. Suppl. Ser.* **258**, 22 (2022).

[41] M. Mapelli, Y. Bouffanais, F. Santoliquido, M. A. Sedda, and M. C. Artale, The cosmic evolution of binary black holes in young, globular, and nuclear star clusters: Rates, masses, spins, and mixing fractions, *Mon. Not. R. Astron. Soc.* **511**, 5797 (2022).

[42] F. Antonini and M. Gieles, Population synthesis of black hole binary mergers from star clusters, *Mon. Not. R. Astron. Soc.* **492**, 2936 (2020).

[43] M. A. Sedda, M. Mapelli, M. Benacquista, and M. Spera, Population synthesis of black hole mergers with B-POP: The impact of dynamics, natal spins, and intermediate-mass black holes on the population of gravitational wave sources, *Mon. Not. R. Astron. Soc.* **520**, 5259 (2023).

[44] K. Kritos, V. Strokov, V. Baibhav, and E. Berti, RAPSTER: A fast code for dynamical formation of black-hole binaries in dense star clusters, [arXiv:2210.10055](https://arxiv.org/abs/2210.10055).

[45] M. Zevin and D. E. Holz, Avoiding a cluster catastrophe: Retention efficiency and the binary black hole mass spectrum, *Astrophys. J. Lett.* **935**, L20 (2022).

[46] C. Kimball, C. Talbot, C. P. L. Berry, M. Carney, M. Zevin, E. Thrane, and V. Kalogera, Black hole genealogy: Identifying hierarchical mergers with gravitational waves, *Astrophys. J.* **900**, 177 (2020).

[47] C. Kimball *et al.*, Evidence for hierarchical black hole mergers in the second LIGO–Virgo gravitational wave catalog, *Astrophys. J. Lett.* **915**, L35 (2021).

[48] K. W. K. Wong, K. Breivik, W. M. Farr, and R. Luger, Backward population synthesis: Mapping the evolutionary history of gravitational-wave progenitors, *Astrophys. J.* **950**, 181 (2023).

[49] S. R. Taylor and D. Gerosa, Mining gravitational-wave catalogs to understand binary stellar evolution: A new hierarchical Bayesian framework, *Phys. Rev. D* **98**, 083017 (2018).

[50] M. Mould, D. Gerosa, and S. R. Taylor, Deep learning and Bayesian inference of gravitational-wave populations: Hierarchical black-hole mergers, *Phys. Rev. D* **106**, 103013 (2022).

[51] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190412: Observation of a binary-black-hole coalescence with asymmetric masses, *Phys. Rev. D* **102**, 043015 (2020).

[52] D. Gerosa, S. Vitale, and E. Berti, Astrophysical implications of GW190412 as a remnant of a previous black-hole merger, *Phys. Rev. Lett.* **125**, 101103 (2020).

[53] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190521: A binary black hole merger with a total mass of  $150M_{\odot}$ , *Phys. Rev. Lett.* **125**, 101102 (2020).

[54] M. Fishbach, C. Kimball, and V. Kalogera, Limits on hierarchical black hole mergers from the most negative  $\chi_{\text{eff}}$  systems, *Astrophys. J. Lett.* **935**, L26 (2022).

[55] M. Spera and M. Mapelli, Very massive stars, pair-instability supernovae and intermediate-mass black holes with the SEVN code, *Mon. Not. R. Astron. Soc.* **470**, 4739 (2017).

[56] D. Gerosa and M. Kesden, PRECESSION: Dynamics of spinning black-hole binaries with PYTHON, *Phys. Rev. D* **93**, 124066 (2016).

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

[58] C. L. Rodriguez and A. Loeb, Redshift evolution of the black hole merger rate from globular clusters, *Astrophys. J. Lett.* **866**, L5 (2018).

[59] P. Madau and T. Fragos, Radiation backgrounds at cosmic dawn: X-rays from compact binaries, *Astrophys. J.* **840**, 39 (2017).

[60] M. J. Fitchett, The influence of gravitational wave momentum losses on the centre of mass motion of a Newtonian binary system, *Mon. Not. R. Astron. Soc.* **203**, 1049 (1983).

[61] M. C. Miller and D. P. Hamilton, Four-body effects in globular cluster black hole coalescence, *Astrophys. J.* **576**, 894 (2002).

[62] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo, *SoftwareX* **13**, 100658 (2021).

[63] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Properties and astrophysical implications of the  $150M_{\odot}$  binary black hole merger GW190521, *Astrophys. J. Lett.* **900**, L13 (2020).

[64] M. Zevin, S. S. Bavera, C. P. L. Berry, V. Kalogera, T. Fragos, P. Marchant, C. L. Rodriguez, F. Antonini, D. E. Holz, and C. Pankow, One channel to rule them all? Constraining the origins of binary black holes using multiple formation pathways, *Astrophys. J.* **910**, 152 (2021).

[65] K. Belczynski, The most ordinary formation of the most unusual double black hole merger, *Astrophys. J. Lett.* **905**, L15 (2020).

[66] V. De Luca, V. Desjacques, G. Franciolini, P. Pani, and A. Riotto, GW190521 Mass gap event and the primordial black hole scenario, *Phys. Rev. Lett.* **126**, 051101 (2021).

[67] <https://github.com/aantonelli94/TheBHClassifier>.

[68] <https://gwosc.org/>.