

A self-consistent method to estimate the rate of compact binary coalescences with a Poisson mixture model

Shasvath J. Kapadia,¹ Sarah Caudill,² Jolien D. E. Creighton,¹ Will M. Farr,³
 Gregory Mendell,⁴ Alan Weinstein,⁵ Kipp Cannon,^{6,7} Heather Fong,^{6,8,7} Patrick
 Godwin,^{9,10} Rico K. L. Lo,^{5,11} Ryan Magee,^{9,10} Duncan Meacher,¹ Cody Messick,^{9,10}
 Siddharth R. Mohite,^{1,12} Debnandini Mukherjee,¹ and Surabhi Sachdev^{9,10}

¹*Leonard E. Parker Center for Gravitation, Cosmology, and Astrophysics,
 University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA*

²*Nikhef, Science Park, 1098 XG Amsterdam, Netherlands*

³*Department of Physics and Astronomy, Stony Brook University, Stony Brook NY 11794, USA*

⁴*LIGO Hanford Observatory, Richland, WA 99352, USA*

⁵*LIGO Laboratory, California Institute of Technology, MS 100-36, Pasadena, California 91125, USA*

⁶*Canadian Institute for Theoretical Astrophysics, 60 St. George Street,
 University of Toronto, Toronto, Ontario, M5S 3H8, Canada*

⁷*RESCEU, The University of Tokyo, Tokyo, 113-0033, Japan*

⁸*Department of Physics, 60 St. George Street, University of Toronto, Toronto, Ontario, M5S 3H8, Canada*

⁹*Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA*

¹⁰*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*

¹¹*Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong*

¹²*LSSTC Data Science Fellow*

The recently published GWTC-1 [1] - a journal article summarizing the search for gravitational waves (GWs) from coalescing compact binaries in data produced by the LIGO-Virgo network of ground-based detectors during their first and second observing runs - quoted estimates for the rates of binary neutron star, neutron star black hole binary, and binary black hole mergers, as well as assigned probabilities of astrophysical origin for various significant and marginal GW candidate events. In this paper, we delineate the formalism used to compute these rates and probabilities, which assumes that triggers above a low ranking statistic threshold, whether of terrestrial or astrophysical origin, occur as independent Poisson processes. In particular, we include an arbitrary number of astrophysical categories by redistributing, via mass-based template weighting, the foreground probabilities of candidate events, across source classes. We evaluate this formalism on synthetic GW data, and demonstrate that this method works well for the kind of GW signals observed during the first and second observing runs.

I. INTRODUCTION

The detection of gravitational waves from a coalescing black hole binary, on September 14, 2015, by the Advanced LIGO (Laser Interferometer Gravitational-Wave Observatory) network of ground based detectors, announced the arrival of gravitational-wave astronomy [2, 3]. Since then, a number of additional gravitational-wave detections have been made. These include 10 from merging binary black holes (BBHs) [1, 4–7] and 1 from a coalescing binary neutron star [8] that also produced an electromagnetic counterpart amply identified by a number of telescopes worldwide [9]. All of these detections are reported in “GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs” [1]. We will refer to these observing runs as O1 and O2.

A plethora of interesting astrophysical questions can now be asked, and answered with appropriately quantified uncertainties. Among them is the question: how many compact binary coalescence (CBC) events, of a specific astrophysical source class, occur per unit spacetime volume? The determination of these coalescence rates is a high-profile science target of the LIGO-Virgo collaboration, and for good reason. Prior to the first gravitational-wave detection, estimates of BBH merger-rates spanned many orders of magnitude [10]. These were based on population models that remained unconstrained due to a paucity of electromagnetic observations. GW events allow these rates to be better constrained in a largely model-independent manner. Using the results from O1 and O2, credible intervals and upper limits on the rate of binary black hole (BBH), neutron star - black hole (NSBH) binary, and binary neutron star (BNS) mergers have been previously published in [5, 8, 11–14].

The rates (at 90% confidence) for the same astrophysical source classes were re-estimated after O2, details of which may be found in [1], which we shall hereafter call the “catalog”. The BNS merger rate was updated to $110 - 3840 \text{ Gpc}^{-3}\text{yr}^{-1}$, the BBH merger rate to $9.7 - 101 \text{ Gpc}^{-3}\text{yr}^{-1}$, and a rate upper limit for NSBH mergers was placed at $610 \text{ Gpc}^{-3}\text{yr}^{-1}$. The goal of this write-up is to serve as a complement to the catalog. In particular, we delineate the formalism and method used to determine the rates of compact binary mergers, and probabilities of astrophysical origin for various highly significant events in O1 and O2, quoted in the catalog for the GstLAL search [15].

The method we present in this paper is a non-trivial extension of one that was developed in the context of gravitational wave data analysis by W. Farr, J. Gair, I. Mandel, and C. Cutler (henceforth FGMC) [16]. (Note that a related formalism was developed earlier, in the context of gamma-ray burst data analysis [17]). The FGMC formalism itself is an application of Poisson statistics to determine the rate of astrophysical events when supplied with a mixture of foreground and background events, given that the foreground and background models are known (or assumed) even though the membership of each event to either one of these classes is unknown.

The FGMC formalism was employed in the determination of BBH rates using data from O1 [11]. The scarcity of confirmed BBH detections prompted the use of a method prescribed by Kim et al [18] where each BBH detected was assumed to be a distinct astrophysical source. On the other hand, the rate upper limits of BNS and NSBH mergers were computed by formally assuming that exactly zero BNS and NSBH events occurred during observation time, and then employing the Poisson distribution for zero events [14]. This assumption is based on the fact that no BNS and NSBH events, with a false alarm rate (FAR) of less than 1/100 years, were found. While changing the threshold by a few orders of magnitude does not vastly alter the results [14], the choice of threshold is in itself somewhat arbitrary.

What we propose, as an alternative to what was done in O1, is a self consistent, threshold-independent, counting method that simultaneously estimates the rates of BNS, NSBH and BBH mergers. In essence, this method extends the FGMC formalism by constructing a joint posterior on the Poisson expected counts for an arbitrary choice of foreground categories by redistributing, via mass-based template-weighting, the foreground probabilities of candidate events across astrophysical source classes. Thus, the method presented here, while used in the catalog paper to handle the three astrophysical categories mentioned above, can handle any number of categories. For example, it could set limits on the rates of BBH mergers in the proposed mass gaps [19] and possibly including a higher mass black hole region too.

The template-weights themselves are computed from simulation runs (software injection campaigns), each targeted at a distinct astrophysical source type. An injection campaign involves adding synthetic signals pertaining to a source class with clearly defined mass and spin distributions, into the detector data, and recovering them via a detection pipeline. To construct the weights, we count how many times injections of a given category are recovered in a given template bin and divide this by the total number of recovered injections pertaining to that category.

In the following sections, we describe this “multi-component” extension of the FGMC method, by first constructing the joint posterior on the Poisson expected counts for each source category, followed by details on how to estimate the spacetime volume sensitivities of the detectors to each of the source categories and how to incorporate uncertainties in their measurements into the rates posterior. We then apply this extension to synthetic data, in order to assess its ability to accurately categorize and count a mixture of CBC signals. We end by summarizing the multi-component FGMC method, evaluating its performance, and suggesting other applications of the method.

II. CONSTRUCTING THE MULTI-COMPONENT COUNTS POSTERIOR FROM CANDIDATE EVENTS

A. Posterior on Terrestrial and Astrophysical counts

The original FGMC method [16] constructs a two-component likelihood on the expected number of astrophysical (Λ_1) and terrestrial (Λ_0) counts, per experiment, assuming that the foreground and background triggers above a low ranking statistic threshold where background triggers¹ dominate, occur as independent Poisson processes. More specifically, the expected counts (Λ) referred to here are the Poisson means for the duration of the experiment (the total observing time), given which one can compute the discrete probability distribution on the number k of occurrences of events :

$$p(k|\Lambda) \propto \Lambda^k \exp(-\Lambda). \quad (1)$$

The two-component FGMC likelihood has the following form [16]:

$$p(\vec{x}|\Lambda_0, \Lambda_1) = e^{-\Lambda_0 - \Lambda_1} \prod_{j=1}^N [\Lambda_0 b(x_j) + \Lambda_1 f(x_j)] \quad (2)$$

where $\vec{x} = \{x_j\}$, $j = 1, 2, 3, \dots, N$, are the ranking statistics of the triggers above threshold, and $b(x_j), f(x_j)$ are the background and foreground distributions (normalized density functions, also called the background and foreground models), evaluated at x_j ($b(x_j) \equiv p(x_j|\text{noise})$ and $f(x_j) \equiv p(x_j|\text{signal})$). It is worth noting here that the foreground count Λ_1 is directly proportional to the astrophysical rate of mergers R , which can be determined if the population averaged spacetime volume sensitivity $\langle VT \rangle$ of the detector is known:

$$\Lambda_1 = R \cdot \langle VT \rangle. \quad (3)$$

The FGMC likelihood may therefore also be written in terms of R and $\langle VT \rangle$. One is thus at liberty to choose a prior, either on Λ_1 , or on R and VT jointly. We will come back to this in an upcoming section where we discuss incorporating uncertainties in the measurement of the spacetime volume sensitivity into the rates posteriors. For the moment, we proceed by choosing a prior on Λ_0, Λ_1 , and writing the FGMC posterior [11, 16]:

$$p(\Lambda_0, \Lambda_1|\vec{x}) \propto p(\Lambda_0, \Lambda_1) e^{-\Lambda_0 - \Lambda_1} \prod_{j=1}^N [\Lambda_0 b(x_j) + \Lambda_1 f(x_j)] \quad (4)$$

where $p(\Lambda_0, \Lambda_1)$ is the prior on the counts. We wish to extend this method to include an arbitrary number of astrophysical components (BNS, NSBH, BBH, possibly others), in place of a single aggregated astrophysical component.

B. Multi-component counts posterior

To a very good approximation, the foreground distribution of ranking statistics is independent of source category [12, 20, 21]. Symbolically:

$$p(x|\alpha) \approx p(x|\text{signal}) \quad (5)$$

¹ A trigger is a gravitational wave candidate event acquired during a templated matched-filtering based analysis of detector strain data. A background trigger is one that was most likely produced by terrestrial processes.

where α is an astrophysical source category. However, it is necessary to re-weight the foreground distribution with source-category specific weights $W_\alpha(x)$. This re-weighting would allow us to split the foreground distribution into multiple, source-specific foreground distributions $f_\alpha(x)$, where x could now encapsulate multiple properties of a trigger, and not necessarily only the ranking statistic. The general mathematical form of the posterior becomes:

$$p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) \propto p(\Lambda_0, \vec{\Lambda}_1) e^{-\Lambda_0 - \vec{\Lambda}_1 \cdot \vec{u}} \cdot \prod_{j=1}^N [\Lambda_0 b(x_j) + \vec{f}(x_j) \cdot \vec{\Lambda}_1] \quad (6)$$

where $\vec{f}(x) \equiv \{f_\alpha(x)\}$ (for $\alpha = \{\text{BNS}, \text{NSBH}, \dots\}$) is a vector of source-specific foreground distributions, $\vec{\Lambda}_1 \equiv \{\Lambda_\alpha\}$ (for $\alpha = \{\text{BNS}, \text{NSBH}, \dots\}$) is a vector of source-specific Poisson expected counts, and $\vec{u}$ is a vector of 1s (See Appendix A).

Gravitational waves from different CBC sources are expected to activate (ring-up) templates from different (though not necessarily disjoint) regions of a template-bank's parameter space. By dividing the template bank into multiple bins (which we denote as m), we can assign to each trigger source-specific template-weights based on the bin in which the template lives. As derived in Appendix A, the source-specific foreground distributions are constructed using template-weights and bin-dependent foreground distributions:

$$f_\alpha(x) \equiv p(L, m | \alpha) \approx p(L | m, \text{signal}) \cdot W_\alpha(m). \quad (7)$$

Assuming that the detector data was analyzed with the GstLAL detection pipeline [15, 22], L is the log-likelihood-ratio ranking statistic [23], m is the bin number, $W_\alpha(m) \equiv p(m | \alpha)$ are the bin-dependent template weights, and $p(L | m, \text{signal})$ are the bin-dependent foreground distributions. On the other hand, the background distribution is given by:

$$b(x) \equiv p(L, m | \text{noise}) = p(L | m, \text{noise}) \cdot W_0(m) \quad (8)$$

where $W_0(m) \equiv p(m | \text{noise})$.

It is convenient to define source-specific Bayes-factors for a trigger x , using the foreground and background distributions evaluated at x 's ranking statistic value L , as well as the template weights:

$$\vec{K}(x) \equiv \frac{\vec{f}(x)}{b(x)} = \frac{p(L | m, \text{signal})}{p(L | m, \text{noise})} \frac{\vec{W}_1(m)}{W_0(m)} \quad (9)$$

where $\vec{W}_1(m) = \{W_{\text{BNS}}(m), W_{\text{NSBH}}(m), \dots\}$. The multi-component counts posterior can now be written more compactly as:

$$p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) \propto p(\Lambda_0, \vec{\Lambda}_1) e^{-\Lambda_0 - \vec{\Lambda}_1 \cdot \vec{u}} \cdot \prod_{j=1}^N [\Lambda_0 + \vec{\Lambda}_1 \cdot \vec{K}(x_j)]. \quad (10)$$

C. Useful Approximations to the Multi-component Counts Posterior

In this section, we recast the counts posterior ((10)) in approximate forms that make it computationally efficient to marginalize out the terrestrial count Λ_0 .

If the number of candidate events is sufficiently large ($N \gg 1$), and the number of background events vastly exceeds the number of foreground events, then, using the method of Laplace:

$$\Lambda_0^N e^{-\Lambda_0} \approx N^N e^{-N} e^{-(\Lambda_0 - N)^2 / (2N)} \quad (11)$$

The multi-component counts posterior then assumes the form:

$$p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) \propto N^N e^{-N} e^{-(\Lambda_0 - N)^2 / (2N)} p(\Lambda_0, \vec{\Lambda}_1) e^{-\vec{\Lambda}_1 \cdot \vec{u}} \cdot \prod_{j=1}^N \left[1 + \frac{\vec{\Lambda}_1}{\Lambda_0} \cdot \vec{K}(x_j) \right] \quad (12)$$

From the perspective of determining astrophysical rates, the posterior on the terrestrial counts is generally not of much interest, and is usually marginalized out. Writing the multi-component posterior in the above form makes the marginalization over terrestrial counts amenable to Gauss-Hermite quadrature.

A further simplification to the counts posterior can be written down, again in the limit of large N dominated by background events. In this simplification, the posterior on the terrestrial count is modeled as a Dirac-delta function centered on N . Thus, when marginalizing out the terrestrial count, the multi-component counts posterior assumes the form:

$$p(\vec{\Lambda}_1|\vec{x}) \propto p(N, \vec{\Lambda}_1) e^{-\vec{\Lambda}_1 \cdot \vec{u}} \cdot \prod_{j=1}^N [1 + \vec{\Lambda}_1 \cdot \vec{k}(x_j)] \quad (13)$$

where $\vec{k}(x) \equiv \frac{\vec{K}(x)}{N}$ is the reduced Bayes factor for trigger x .

Note that $k_\alpha(x_j) \ll 1$ for the majority of triggers, since the majority are background events. Conversely, $k_\alpha(x_j) \gg 1$ for certain highly significant foreground events, i.e, events that are almost unambiguously of astrophysical category α .

D. Bin-dependent template weights

The key to constructing the multi-component counts posterior is to determine the weights $W_\alpha(m) \equiv p(m|\alpha)$, which is a measure of how the astrophysical signals of a specific source category distribute themselves in the template bank. This subsection outlines how these weights are approximately computed for the GstLAL detection pipeline.

The GstLAL pipeline splits the template bank into sub-banks (which we simply refer to as bins) [24], in the “ $\mathcal{M} - \chi_{\text{eff}}$ ” space, where $\mathcal{M}$ is the template’s chirp mass, and χ_{eff} is the template’s effective spin parameter. The chirp mass is defined as:

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \quad (14)$$

and the effective spin parameter is defined as:

$$\chi_{\text{eff}} = \frac{m_1 \chi_1 + m_2 \chi_2}{m_1 + m_2} \quad (15)$$

with m_1, m_2, χ_1, χ_2 as the component masses and spin angular momenta (or more precisely, their components parallel to the orbital angular momentum) of the binary.

While this binning was originally designed to speed-up the extraction of GW signals from detector data, one can also exploit it for the construction of template-weights and the multi-component counts posterior.

These bins can be thought of as coarse-grained templates; when a template is “rung-up”, the corresponding bin in which it lives is said to be “activated”. Thus, during a run, we can count the number of times each bin gets activated, and thus determine an “activation count” for each bin. Intuitively, one can see that signals from a specific astrophysical source-class will tend to predominantly activate only a subset of all the bins. BNS signals for example will tend to activate low-mass bins, whereas BBH signals will tend to activate high-mass bins.

Now, suppose we run distinct injection campaigns targeted at specific source categories (BNS, NSBH, BBH, ...). In other words, we inject simulated waveforms of a specific source class, and recover these injections using the binned template bank. Considering only those injections that were recovered with false alarm rate (FAR) of less than 1/30 days, we determine activation counts A_α , ($\alpha = 1, 2, \dots, Q$, where Q is the total number of astrophysical categories) corresponding to these injections, and construct a set of weights as follows:

$$W_\alpha(m) = \frac{A_\alpha(m)}{\sum_{m=0}^{N_{\text{bins}}-1} A_\alpha(m)} \quad (16)$$

where m is the bin number and N_{bins} is the total number of bins. (Note that GstLAL-based CBC rate estimations quoted in the catalog paper were computed with analyses that split the template bank into 686 bins. A visual representation of these weights is shown in Figure 1 .)

To define specific source categories, injection campaigns were designed to reflect the choices made for parameter boundaries of the astrophysical sources in the catalog paper. The BNS source category included neutron stars with component masses m_i distributed uniformly in $\ln m_i$ between $1 \leq m_i/M_\odot \leq 3.0$ such that the total mass M was less than $6.0M_\odot$. The lower mass limit is motivated by a 6σ deviation from masses of components in double neutron star systems [25] while the upper mass limit is motivated by certain models and observations which allow neutron stars to form up to $3M_\odot$ [26–30]. Spin vectors for BNS components were allowed to be isotropic in direction and uniform in magnitude, with a maximum allowed spin magnitude of 0.4. This maximum magnitude constraint is motivated by observations of the fastest spinning pulsar with $\chi \lesssim 0.4$ [31]. The BBH source category included black holes with component masses distributed uniformly in $\ln m_i$ between $5 \leq m_i/M_\odot \leq 50$ such that $M \leq 100M_\odot$. The lower mass limit is motivated by the possible existence of a minimum black hole mass [32–34] while the upper mass limit is motivated by evidence of an upper cutoff in the BBH mass spectrum based on the first few LIGO detections [35–37]. Spin vectors for BBH components were allowed to be isotropic on the sphere with a maximum allowed spin magnitude of 0.99. The relativistic Kerr bound provides a theoretical constraint on black hole spin magnitude of 1.0 although we are also constrained by the limit of the waveform approximant. The NSBH source category included neutron stars with component masses distributed uniformly in $\ln m_1$ between $1.0 \leq m_1/M_\odot \leq 3.0$ and black holes with component masses distributed uniformly in $\ln m_2$ between $5.0 \leq m_2/M_\odot \leq 100.0$. The total mass for the NSBH category was constrained to $M \leq 103.0M_\odot$. Both the NS and BH components were allowed to be isotropic on the sphere with maximum allowed spin magnitudes of 0.4 and 0.99, respectively. These mass and spin limits are motivated by the NS and BH observations and theoretical constraints already mentioned.²

Injectons in each source category were distributed uniformly in co-moving volume out to redshift of 0.2 for BNS and NSBH and out to 0.7 for BBH. In order to maximize the number of recoverable injections included in the injection campaign, an initial cut on expected signal-to-noise ratio less than 3.0 was applied to exclude injections that would be too far away or in a poor sky location for either of the Hanford or Livingston detectors. The parameters of these injections were tabulated and stored as unrecoverable.

E. Probability of Astrophysical Origin

From the original, two-component, FGMC counts posterior, one can compute the posterior probability that an event, with foreground and background distribution values $f(x)$ and $b(x)$, evaluated at the event's ranking statistic x , is of astrophysical origin, given the data $\vec{x}$ [11]:

$$P_1(x | \vec{x}) = \int_0^\infty p(\Lambda_0, \Lambda_1 | \vec{x}) \frac{\Lambda_1 f(x)}{\Lambda_0 b(x) + \Lambda_1 f(x)} d\Lambda_0 d\Lambda_1. \quad (17)$$

Its complementary quantity is the posterior probability $P_0(x | \vec{x})$ that the same event originated from the Earth, with $P_0(x) + P_1(x) = 1$.

These posterior probabilities can be straightforwardly extended to the case when we have a multi-component counts posterior. The source-specific foreground distributions of the multi-component posterior allow one to compute posterior probabilities pertaining to specific astrophysical source categories:

$$P_\alpha(x | \vec{x}) = \int_0^\infty p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) \frac{\Lambda_\alpha K_\alpha(x)}{\Lambda_0 + \vec{\Lambda}_1 \cdot \vec{K}(x)} d\Lambda_0 d\vec{\Lambda}_1. \quad (18)$$

The complementary terrestrial posterior probability is once again related to the astrophysical probabilities via $P_0(x | \vec{x}) + \sum_\alpha P_\alpha(x | \vec{x}) = 1$, where α is summed over all astrophysical source categories corresponding to the multi-component posterior.

Astrophysical probabilities of candidate events are of considerable interest, from the perspective of following up gravitational wave events with telescopes sensitive to various parts of the electromagnetic spectrum. For instance, if P_{BNS} were high, the probability that this event would produce an electromagnetic

² To ensure appropriate coverage of the component mass space, an additional injection set where at least one of the components lies in the range $3 - 5M_\odot$, was constructed. The masses were distributed uniformly in $\ln m_i$, with the other component spanning $1 - 100M_\odot$. The spins were assumed to be isotropic, with a maximum value of 0.4 for the first component, and 0.99 for the latter.

counterpart also becomes high; such information could therefore be of great value to astronomers interested in hunting for electromagnetic signals associated with GWs, if reported with a sufficiently low latency from the time of occurrence of a GW candidate event.

With low-latency in mind, it is possible to re-write astrophysical probabilities as a function of the mean values of the Poisson expected counts. We define the mean values in the standard way:

$$\langle \Lambda_\alpha \rangle = \int_0^\infty \Lambda_\alpha p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) d\Lambda_0 d\vec{\Lambda}_1 \quad (19)$$

Suppose now we have a set of N candidate events $\vec{x}_N = \{x_0, x_1, \dots, x_{N-1}\}$, from which we compute the mean value for the terrestrial Poisson count $\langle \Lambda_0 \rangle_N$ and astrophysical Poisson counts $\langle \vec{\Lambda}_1 \rangle_N = \{\langle \Lambda_1 \rangle, \langle \Lambda_2 \rangle, \dots, \langle \Lambda_Q \rangle\}$. The astrophysical probability, for category α , of a new candidate event x_{N+1} can now be computed as (See Appendix B):

$$P_\alpha(x_{N+1} | \vec{x}_{N+1}) = \frac{\langle \Lambda_\alpha \rangle_N K_\alpha(x_{N+1})}{\langle \Lambda_0 \rangle_N + \langle \vec{\Lambda}_1 \rangle_N \cdot \vec{K}(x_{N+1})} \quad (20)$$

The above expression can be readily derived using Eq. (18) and Eq. (B7). Thus, if the mean values $\langle \Lambda_0 \rangle_N$ and $\langle \vec{\Lambda}_1 \rangle_N$ are pre-computed, then the the astrophysical probabilities $P_\alpha(x_{N+1} | \vec{x}_{N+1})$ for a new candidate event can be computed almost instantaneously using only a handful of floating point operations. (Of course, the mean values would then need to be updated using Eq. (B12)).

III. DETERMINING THE SPACETIME-VOLUME SENSITIVITY $\langle VT \rangle$

In order convert the posterior on counts to a posterior on rates, we need to determine the population averaged spacetime volume sensitivity $\langle VT \rangle_\alpha$ of the detectors, for every astrophysical source category α . This sensitivity is written as [1]:

$$\langle VT \rangle_\alpha = T \int dz d\theta \frac{dV_c}{dz} \frac{1}{1+z} p_\alpha(\theta) f(z, \theta) \quad (21)$$

where T is the duration over which the sensitivity is estimated, z is the redshift, V_c is the co-moving volume, $p_\alpha(\theta)$ is an assumed distribution of source-parameters θ for source-category α , and $f(z, \theta)$ is a selection function that estimates how likely it is to detect sources with parameters θ at redshift z .

The above quantity is typically estimated using Monte-Carlo integration. Simulated signals (injections), with parameters drawn from $p_\alpha(\theta)$ and placed in redshift assuming standard cosmology, are added to the detector data. They are then searched for by the GstLAL pipeline, and assigned Bayes factors $\vec{K}(x)$. An estimate is then made of the number of injections that were recovered, N_{rec} . Since the number that was injected, N_α^{inj} , and the spacetime volume into which the injections were made, $\langle VT \rangle_\alpha^{\text{inj}}$, are both known, the measured spacetime volume is simply the injected spacetime volume re-scaled by the ratio of number-recovered to number-injected [38]:

$$\langle VT \rangle_\alpha = \frac{N_{\text{rec}}}{N_\alpha^{\text{inj}}} \langle VT \rangle_\alpha^{\text{inj}}. \quad (22)$$

The crucial step then in the evaluation of the sensitive spacetime volume is the determination of N_{rec} . In order to be consistent with the way the multicomponent counts posterior was constructed, we choose the same low-ranking statistic threshold when counting the number of recovered injections. However, it is not trivial to map a given trigger to an injection, and count that trigger as a recovered injection, given the low ranking statistic threshold used and the resulting preponderance of background triggers. We therefore propose the following method. (Note that a similar method, using the two-component counts posterior, was used for the determination of $\langle VT \rangle$, in [5].)

Let $[x_1, \dots, x_N]$ be the list of triggers produced during a search, and let $[\gamma_1, \gamma_2, \dots, \gamma_I]$ be the list of triggers produced during the injection campaign. Suppose now we include one trigger from the injection campaign, to the list of triggers from the search. We now have the following list $[x_1, x_2, \dots, x_N, \gamma_i]$. The contribution

of γ_i to the mean value of the count for category α can be computed using Eq. B10. We also compute the contributions to the mean values pertaining to all the other astrophysical source categories, due to γ_i , and determine an aggregated contribution:

$$\Delta\Lambda_{\gamma_i} = \sum_{\alpha} \frac{\text{cov}_N(\Lambda_0, \Lambda_{\alpha}) + \sum_{\beta} \text{cov}_N(\Lambda_{\alpha}, \Lambda_{\beta}) K_{\beta}(\gamma_i)}{\langle\Lambda_0\rangle_N + \sum_{\beta} K_{\beta}(\gamma_i) \langle\Lambda_{\beta}\rangle_N} \quad (23)$$

where both α and β are summed over astrophysical source categories, and cov_N is defined in Appendix B.

We repeat this procedure for all triggers from the injection campaign, adding each one separately (with replacement) to the triggers from the search, and determining the change in the mean due to each addition. N_{rec} is estimated as the sum of the contributions to the mean due to all γ_i s in $[\gamma_1, \gamma_2, \dots, \gamma_I]$:

$$N_{\text{rec}} = \sum_{i=1}^I \Delta\Lambda_{\gamma_i} \quad (24)$$

IV. RATES POSTERIOR WITH UNCERTAINTIES

As such, going from the counts posterior to the rates posterior is a trivial change of variables. Thought of in another way, rates and counts are essentially the same quantity expressed in different units:

$$\vec{R} = \left\{ \frac{\Lambda_{\alpha}}{(VT)_{\alpha}} \right\}, \alpha = 1, 2, \dots, Q \quad (25)$$

where, as before, Q is the total number of astrophysical components.

There are, however, uncertainties associated with the determination of the spacetime volume sensitivity, arising from calibration errors when measuring the detector strain h ³:

$$S_{\text{cal}} \approx 3 \frac{\Delta h}{h} \quad (26)$$

as well as statistical Monte Carlo errors when evaluating Eq. 21:

$$S_{\text{stat}} = \frac{1}{\sqrt{N_{\text{rec}}}}. \quad (27)$$

Let S be the total fractional error associated with these uncertainties:

$$S = \sqrt{S_{\text{stat}}^2 + S_{\text{cal}}^2}. \quad (28)$$

One can then model the measurement distribution on VT as a log-normal:

$$p(VT | \langle VT \rangle, S) = \frac{1}{VT \cdot S \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\ln VT - \ln \langle VT \rangle}{S} \right)^2 \right] \quad (29)$$

where $\langle VT \rangle$ is assumed to be the same as the measured population-averaged spacetime volume in Eq. (21).

There are now two ways in which the distributions on the Poisson counts, and the distribution on the spacetime volume sensitivity, can be combined to evaluate the posterior on the rates.

The first method is a direct application of the ratio distribution: given two positive random variables, y_1 and y_2 , with joint distribution $f(y_1, y_2)$, the distribution on the ratio of these two variables goes as:

$$p(u \equiv y_2/y_1) = \int_0^{\infty} y_1 f(y_1, u y_1) dy_1. \quad (30)$$

³ The uncertainty in the GW amplitude h measured from detector data gets translated to an uncertainty in the detector range d at leading order ($d \sim h$), or equivalently, an uncertainty in volume ($V \sim h^3 \Rightarrow \frac{dV}{V} \sim 3 \frac{dh}{h}$); (see [12], specifically section 5 therein.)

Using the ratio in Eq. (25), we identify y_1 as VT and y_2 as Λ , with y_1 distributed as given in Eq. (29) and y_2 distributed as given in Eq. (10). And assuming the joint distribution $f(y_1, y_2)$ is the product of Eqns.(29) and (10), we can put down a distribution on the rate:

$$p(R_\alpha | \vec{x}, \langle VT \rangle_\alpha, S_\alpha) = \frac{1}{S_\alpha \sqrt{2\pi}} \int_{-\infty}^{\infty} p(R_\alpha \cdot e^{v_\alpha} | \vec{x}, \langle VT \rangle_\alpha) \exp \left[-\frac{1}{2} \left(\frac{v_\alpha}{S_\alpha} \right)^2 + v_\alpha \right] dv_\alpha \quad (31)$$

where $v_\alpha = \ln \left(\frac{\langle VT \rangle_\alpha}{\langle VT \rangle_\alpha} \right)$, and $p(R_\alpha \cdot e^{v_\alpha} | \vec{x}, \langle VT \rangle_\alpha)$ is the marginalized counts posterior (cf. Eq.10) for source-category α with the change of variable $\Lambda_\alpha \rightarrow \Lambda_\alpha / \langle VT \rangle_\alpha = R_\alpha e^{v_\alpha}$, given a measured spacetime volume sensitivity $\langle VT \rangle_\alpha$.

The second method starts by asserting that R and VT are independent random variables, and the joint probability distribution on these two variables (we have dropped the subscript α for notational simplicity):

$$p(R, VT | \langle VT \rangle, S) = p(R) \cdot p(VT | \langle VT \rangle, S) \quad (32)$$

acts as the prior for the joint posterior on R and VT :

$$p(R, VT | \vec{x}, \langle VT \rangle, S) \propto p(R, VT | \langle VT \rangle, S) \cdot p(\vec{x} | R, VT). \quad (33)$$

Here, $p(\vec{x} | R, VT)$ is the FGMC likelihood, and $p(VT | \langle VT \rangle, S)$ is modeled with the log-normal distribution as in Eq. (29). Marginalizing out VT from the above equation gives the sought after rate posterior. Assuming further that the prior on the rate $p(R)$ follows a power law, the rate posterior becomes up to a normalization constant (re-introducing the subscript α to simplify comparing the equation below with (31)):

$$p(R_\alpha | \vec{x}, \langle VT \rangle_\alpha, S_\alpha) \propto \int_{-\infty}^{+\infty} p(R_\alpha \cdot e^{v_\alpha} | \vec{x}, \langle VT \rangle_\alpha) \exp \left[-\frac{1}{2} \left(\frac{v_\alpha}{S_\alpha} \right)^2 - a v_\alpha \right] dv_\alpha \quad (34)$$

where a is the exponent of the power-law priors on R . The rate posteriors in the catalog either use a uniform prior ($a = 0$) for categories with no confirmed detections (viz. NSBH), and a Jeffreys prior ($a = -0.5$) for the other categories (viz. BNS, BBH, Terrestrial).

V. ILLUSTRATIVE RESULTS

As a proof of principle, we apply this multi-component extension of the FGMC method to a mixture of synthetic BNS, NSBH and BBH signals added to real detector data from O1 and O2 devoid of astrophysical GW signals. Making the data free from GW signals is achieved by a method referred to as “time-sliding”: the detector strain time series from one of the detectors is translated in time, with respect to the strain time series in another detector, by an amount greater than the light-travel time between the detectors; coincident events post this time-shifting are then used.

The BNS, NSBH and BBH signals were injected into O1 and O2 detector data, and recovered with the GstLAL pipeline, separately. From those injections that were recovered, as determined by whether a trigger exists within a 1 second time window of the injections, 30 BNS injections, 30 NSBH injections, and 100 BBH injections were selected at random. Their corresponding triggers were then added to the list of triggers produced when analyzing the time-slid data.

The set of CBC signals used for the injection campaigns were selected as follows:

BNS signals were drawn at random from a “broad” distribution of synthetic BNS signals. The component masses m_i were distributed uniformly between $0.8 \leq m_i/M_\odot \leq 2.3$. The component spins were isotropically distributed on the sphere with a maximum spin magnitude of 0.4.

NSBH signals were drawn at random from each of three delta-function distributions of synthetic NSBH signals. The delta-functions were centered at component masses ($1.4M_\odot, 5M_\odot$) (low-mass NSBHs), ($1.4M_\odot, 10M_\odot$), ($1.4M_\odot, 30M_\odot$) (high-mass NSBHs), and the spins for each of these sets were aligned with the orbital angular momentum of the binary, with a maximum allowed spin magnitude of 0.05 for the neutron star component, and 0.99 for the black hole component.

BBH signals were drawn at random from a “broad” distribution of synthetic BBH signals. The component masses m_i were distributed uniformly in $\ln m_i$, between $5 \leq m_i/M_\odot \leq 100$, and a total mass cutoff at $100M_\odot$.

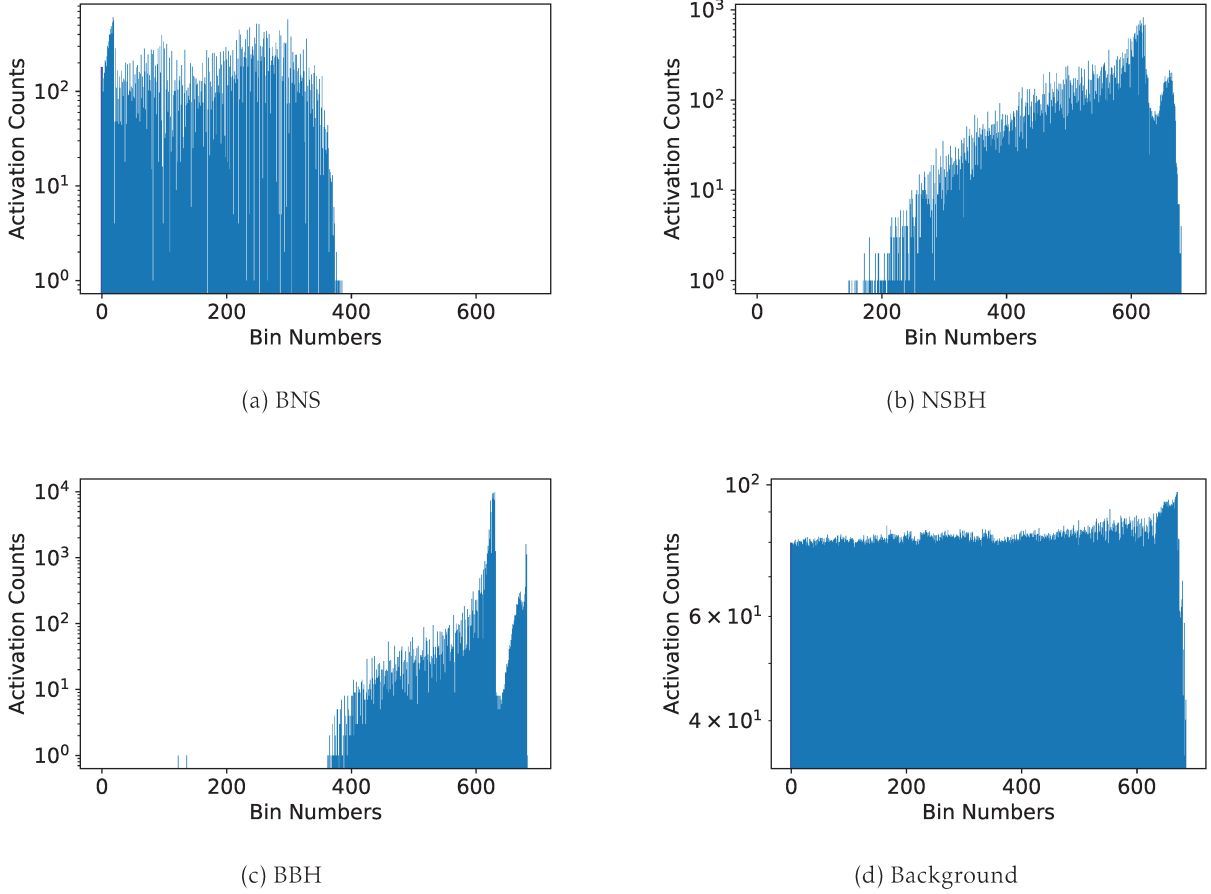


FIG. 1: The panels in this figure show how often each of the 686 non-overlapping regions of the template bank, denoted by a bin number, get “rung up” when recovering BNS, NSBH, and BBH signals. An estimate of how background events distribute themselves in the template bank is also given; this estimate is evaluated by GstLAL when analyzing real detector data devoid of synthetic signals. The weights $W_\alpha(m)$, $\alpha = \{ \text{BNS, NSBH, BBH, 0} \}$, $m = \{0, 1, \dots, 685\}$ (cf. Eq. (16)) are determined from these histograms.

The component spins were aligned with the orbital angular momentum of the binary, with a maximum allowed spin magnitude of 0.99.

All sets from which the various synthetic signals were drawn, ensured that the binary systems were distributed uniformly in co-moving volume.

Employing the GstLAL detection pipeline, this time-slid data added with synthetic signals was analyzed and candidate events were assigned foreground ($p(L|m, \text{signal})$) and background ($p(L|m, \text{noise})$) distribution values. Furthermore, using software injection campaigns described in section II D, bin-dependent template weights $W_\alpha(m)$ ($\alpha = \text{BNS, NSBH, BBH}$) were estimated (see Equation (16) and Figure 1).

We subsequently computed the Bayes factors $\vec{K}_\alpha(x)$ and constructed the multi-component joint posterior on the Poisson counts for each of the categories. Via a Python implementation of Markov-Chain-Monte-Carlo described in [39], we sampled the multi-component posterior. The corner plot in Figure 2, and the median values (with a 90% symmetric confidence interval as error bar) of the Poisson counts reveal that while most BNS and BBH signals are being categorized, on average, as BNS and BBH signals, respectively, some of the NSBH signals are being at least partially confused as BNS or BBH signals. This is perhaps not so surprising if Figure 2 is viewed in light of Figure 1. The set of bins in the template bank activated by BNS signals is not disjoint from the set activated by NSBH signals. The same is also true for NSBH and BBH signals. We would therefore expect to be correlation between BNS and NSBH categories, and NSBH and BBH categories, resulting in signals being partially counted in the wrong category.

We also computed categorical astrophysical probabilities (cf. section II E) for each of the ~ 4000 triggers in

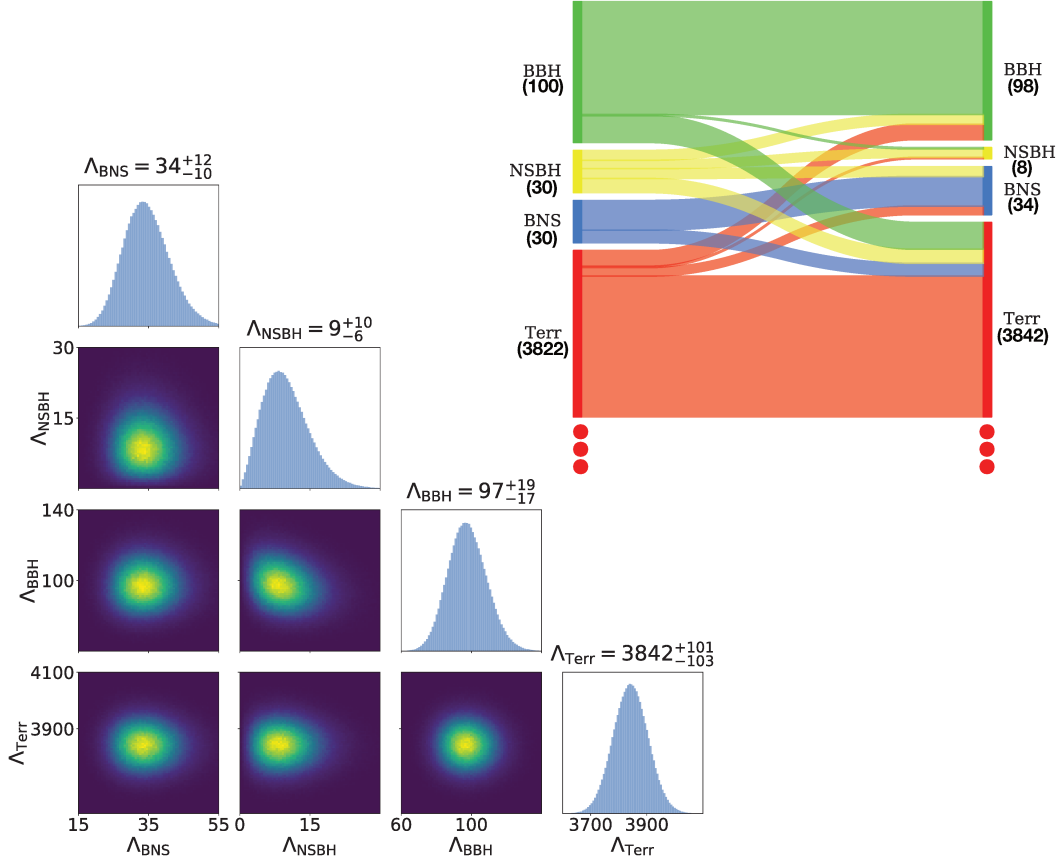


FIG. 2: The **corner plot** provides a visual representation of the multi-component posterior on the Poisson counts for BNS, NSBH, BBH, Terr (Background/Terrestrial) events; this plot also contains the median values (with 90% symmetric confidence intervals as error bars) of the Poisson counts. Time-slid O1 and O2 data, injected with 30 BNS, 30 NSBH and 100 BBH signals, were used. The median BNS and BBH counts are approximately equal to the numbers of BNS and BBH signals injected in the data, and these numbers are well within the error bars. However, the posterior on the NSBH has a median value much lower than the number of injected NSBH signals. This results from a combination of factors: low-mass NSBH signals being partially confused as BNS signals, high mass NSBH signals being partially confused as BBH signals, and NSBH signals that were not recovered by the pipeline with high significance being confused as background.

The **Sankey diagram** - with the number of injected signals depicted on the left hand side, and the aggregated categorical astrophysical probabilities on the right hand side - also illustrates this. It further reveals that the few low-significance BNS and BBH signals classified as noise are being compensated for by high-significance NSBH signals and terrestrial events counted as BNS and BBH signals, resulting in BNS and BBH signals being counted as their appropriate categories on average.

the data set. Rather than tabulate the results for this big list, we depict, via the Sankey diagram in Figure 2, how the multi-component FGMC formalism is categorizing and counting these triggers. A priori, we know which ones of these triggers correspond to a time-slide event, and which ones correspond to an injection. The thickness of the bands on the left hand side of the Sankey diagram is proportional to these numbers. The aggregated astrophysical probabilities, by source category, are also represented by a proportionately thick band on the right hand side of the Sankey diagram. Figure 2 shows that most BBHs and BNSs are

counted as BBHs and BNSs, respectively, and there is no confusion between these two categories. This can again be understood with the help of Figure 1: the regions of the template bank activated by BNS and BBH signals are almost disjoint. On the other hand, the overlap in the regions activated by BNS and NSBH signals, and BBH and NSBH signals, causes low-mass NSBH signals to be partially counted as BNS signals, and high-mass NSBH signals to be partially counted as BBH signals. There is also some correlation between the Terrestrial category and the astrophysical categories, resulting in weaker signals being partially counted as noise, and vice-versa. These misclassifications are discussed further in the next section.

Note that the aggregated categorical astrophysical probabilities are approximately equal to the mean values of the Poisson counts ⁴:

$$\sum_{j=1}^N P_{\alpha}(x_j | \vec{x}) = \langle \Lambda_{\alpha} \rangle - a - 1 \quad (35)$$

where, like in section IV, a is the exponent on the Poisson count for category α in the joint prior for the multi-component FGMC posterior, assuming that the prior has the form $p(\Lambda_0, \vec{\Lambda}_1) = p(\Lambda_0, \Lambda_{\alpha} \notin \vec{\Lambda}_1) \Lambda_{\alpha}^a$ (see appendix D). An identical formula also exists connecting the terrestrial probabilities with the mean of the terrestrial Poisson count. Thus, the right-hand side values (aggregated astrophysical and terrestrial probabilities) in the Sankey diagram of Figure 2 may be thought of as mean values on the Poisson counts.

VI. SUMMARY AND OUTLOOK

Inferring rates of compact binary mergers from gravitational-wave data is an important science goal of the LIGO-Virgo Collaboration (LVC). GWTC-1 [1] (or the “catalog”, as we’ve been calling it in this paper), quoted rate estimates for the mergers of BNSs, NSBHs, and BBHs, using GW data from the LIGO-Virgo network of ground-based interferometric detectors, and analyzed by two separate matched-filter based pipelines (GstLAL and PyCBC). This paper serves as a supplement to the catalog, by describing the formalism used to produce rate estimates from data analyzed by GstLAL. The formalism itself is inspired by a Poisson-statistics based counting method developed by W. Farr, J. Gair, I. Mandel, and C. Cutler (abbreviated to “FGMC” in this paper) in the context of GW data analysis [16] (although similar work in non-GW contexts was done earlier, e.g: [17]).

Assuming that candidate events triggered by terrestrial phenomena or astrophysical GWs occur as independent Poisson processes, the original FGMC formalism constructs a joint posterior conditional probability distribution from the significance of triggers assigned by a detection pipeline (e.g: GstLAL), as measured by the Bayes-Factor (see Section II), above a low-ranking statistic threshold, on the Poisson expected counts for astrophysical and terrestrial CBC events.

The multi-component extension of the FGMC formalism delineated in this paper constructs, from the foreground distribution of ranking statistics estimated by GstLAL, an arbitrary number of foreground probability distributions, each targeted at a specific astrophysical source class. This is achieved by a mass-based template-weighting method. The weights are determined via injection campaigns, by studying the distribution of templates, across the template bank, that got “rung-up” during the recovery of injections.

A proof-of-principle application of this multi-component extension was conducted on time-slid O1 and O2 data, added with synthetic BNS, NSBH and BBH signals, and analyzed with the GstLAL detection pipeline. We find that BNS and BBH signals are for the most part counted appropriately as BNS and BBH. On the other hand, there exist correlations between low-mass NSBH signals and BNS signals, as well as between high mass NSBH signals, and BBH signals, causing partial mis-categorization of NSBH signals. There are also correlations between the astrophysical categories, and the terrestrial category.

Is the confusion in counting and labeling of triggers a cause for concern? It is worth pointing out here that the partial mis-classification of the astrophysical signals as terrestrial signals, and vice-versa, is not a limitation of this method per se, but rather a consequence of the significance assigned to these events by the detection pipeline. Nevertheless, from the perspective of computing rates, the partial mis-classification of

⁴ Equation 35 suggests that the choice of prior becomes increasingly irrelevant as $\langle \Lambda_{\alpha} \rangle$ becomes sufficiently large. This is simply an indication that, when we’ve accumulated many events of category α , the counts posterior is informed primarily by the data, and not any prior assumption on the distribution of the counts.

astrophysical and terrestrial signals is somewhat compensated for by the fact that the same threshold on the ranking statistic is used when determining the counts posterior, and the spacetime volume sensitivity. The *ratio* of these two quantities, which gives the astrophysical rate, is largely unaffected by this choice of threshold, provided the threshold ensures a preponderance of terrestrial triggers [16].

However, the confusion between low (high) mass NSBH signals and BNS (BBH) signals, would be a concern, if this significantly changed the counts of these types of signals. Nevertheless, the extensive follow-up analysis using the parameter estimation given in Sec. V and Table III of the catalog show that there were no NSBH signals of any sort detected. Thus, the misclassification of NSBH events at the border between source categories, is not a concern for the data set being analyzed, viz. O1 and O2 data.

Parameter estimation studies show that 10 significant BBH detections and 1 significant BNS detection, were uncovered from the data. These findings are consistent with the multi-component rates results given in the catalog and computed using the methods described in this paper. The multi-component rates are also consistent with separate (albeit simpler) rates analyses described in the catalog, which serves as a nice sanity check for the method described in this paper.

However, we will need to extend the methods described here in the future to better identify astrophysical signals at the boundaries between source-categories. For example, a computationally intensive method involving a synergy of the formalism delineated in this paper and the samples provided by accurate parameter estimation studies, such as the one described in [40], might be required.

It is worth noting here that the application of the multi-component extension of FGMC used in the catalog is by no means unique. One need not restrict oneself to just BNSs, NSBHs, and BBHs. Additional astrophysical source classes could be added, like, for example, CBCs from the putative “mass-gaps” (with binary component mass range $3M_{\odot} - 5M_{\odot}$ and above $45M_{\odot}$ [19]), and different categories of black holes. Moreover, the redistribution of foreground events need not involve mass-based weighting alone; for example, work on constructing redshift dependent foreground distributions, to study redshift dependent rates of BBHs, is expected to be published soon [41].

VII. ACKNOWLEDGEMENTS

We thank the LIGO-Virgo Scientific Collaboration for access to data. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation (NSF) and operates under cooperative agreement PHY-0757058. We gratefully acknowledge the support by NSF grant PHY-1626190 for the UWM computer cluster. We also thank Deep Chatterjee, Shaon Ghosh and Chad Hanna for illuminating discussions. SJK gratefully acknowledges support through NSF grant PHY-1607585. SRM thanks the LSSTC Data Science Fellowship Program, which is funded by LSSTC, NSF Cybertraining Grant #1829740, the Brinson Foundation, and the Moore Foundation.

-
- [1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. X **9**, 031040 (2019).
 - [2] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].
 - [3] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), Phys. Rev. D **93**, 122003 (2016), arXiv:1602.03839 [gr-qc].
 - [4] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), Phys. Rev. Lett. **116**, 241103 (2016), arXiv:1606.04855 [gr-qc].
 - [5] B. P. Abbott *et al.* (VIRGO, LIGO Scientific), Phys. Rev. Lett. **118**, 221101 (2017), arXiv:1706.01812 [gr-qc].
 - [6] B. P. Abbott *et al.* (Virgo, LIGO Scientific), Astrophys. J. **851**, L35 (2017), arXiv:1711.05578 [astro-ph.HE].
 - [7] B. P. Abbott *et al.* (Virgo, LIGO Scientific), Phys. Rev. Lett. **119**, 141101 (2017), arXiv:1709.09660 [gr-qc].
 - [8] B. P. Abbott *et al.* (Virgo, LIGO Scientific), Phys. Rev. Lett. **119**, 161101 (2017).
 - [9] B. P. Abbott *et al.* (GROND, SALT Group, OzGrav, DFN, INTEGRAL, Virgo, Insight-Hxmt, MAXI Team, Fermi-LAT, J-GEM, RATIR, IceCube, CAASTRO, LWA, ePESSTO, GRAWITA, RIMAS, SKA South Africa/MeerKAT, H.E.S.S., 1M2H Team, IKI-GW Follow-up, Fermi GBM, Pi of Sky, DWF (Deeper Wider Faster Program), Dark Energy Survey, MASTER, AstroSat Cadmium Zinc Telluride Imager Team, Swift, Pierre Auger, ASKAP, VINROUGE, JAGWAR, Chandra Team at McGill University, TTU-NRAO, GROWTH, AGILE Team, MWA, ATCA, AST3, TOROS, Pan-STARRS, NuSTAR, ATLAS Telescopes, BOOTES, CaltechNRAO, LIGO Scientific, High Time Resolution Universe Survey, Nordic Optical Telescope, Las Cumbres Observatory Group, TZAC Consortium, LOFAR, IPN, DLT40,

- Texas Tech University, HAWC, ANTARES, KU, Dark Energy Camera GW-EM, CALET, Euro VLBI Team, ALMA), *Astrophys. J.* **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE].
- [10] J. Abadie *et al.*, *Classical and Quantum Gravity* **27**, 173001 (2010).
 - [11] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), *Astrophys. J. Lett.* **833**, L1 (2016), arXiv:1602.03842 [astro-ph.HE].
 - [12] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), *Astrophys. J. Suppl.* **227**, 14 (2016), arXiv:1606.03939 [astro-ph.HE].
 - [13] B. P. Abbott *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), *Phys. Rev. X* **6**, 041015 (2016), arXiv:1606.04856 [gr-qc].
 - [14] B. P. Abbott *et al.*, *The Astrophysical Journal Letters* **832**, L21 (2016).
 - [15] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. E. Creighton, R. Everett, C. Hanna, D. Keppel, R. N. Lang, T. G. F. Li, D. Meacher, A. Nielsen, C. Pankow, S. Privitera, H. Qi, S. Sachdev, L. Sadeghian, L. Singer, E. G. Thomas, L. Wade, M. Wade, A. Weinstein, and K. Wiesner, *Phys. Rev. D* **95**, 042001 (2017).
 - [16] W. M. Farr, J. R. Gair, I. Mandel, and C. Cutler, *Phys. Rev. D* **91**, 023005 (2015).
 - [17] T. J. Loredo and I. M. Wasserman, *The Astrophysical Journal Supplement* **96**, 261 (1995).
 - [18] C. Kim, V. Kalogera, and D. R. Lorimer, *Astrophys. J.* **584**, 985 (2003), astro-ph/0207408.
 - [19] B. P. Abbott *et al.*, arXiv e-prints, arXiv:1811.12940 (2018), arXiv:1811.12940 [astro-ph.HE].
 - [20] B. F. Schutz, *Classical and Quantum Gravity* **28**, 125023 (2011), arXiv:1102.5421 [astro-ph.IM].
 - [21] H.-Y. Chen and D. E. Holz, arXiv e-prints, arXiv:1409.0522 (2014), arXiv:1409.0522 [gr-qc].
 - [22] S. Sachdev, S. Caudill, H. Fong, R. K. L. Lo, C. Messick, D. Mukherjee, R. Magee, L. Tsukada, K. Blackburn, P. Brady, P. Brockill, K. Cannon, S. J. Chamberlin, D. Chatterjee, J. D. E. Creighton, P. Godwin, A. Gupta, C. Hanna, S. Kapadia, R. N. Lang, T. G. F. Li, D. Meacher, A. Pace, S. Privitera, L. Sadeghian, L. Wade, M. Wade, A. Weinstein, and S. Liting Xiao, arXiv e-prints, arXiv:1901.08580 (2019), arXiv:1901.08580 [gr-qc].
 - [23] K. Cannon, C. Hanna, and J. Peoples, arXiv e-prints, arXiv:1504.04632 (2015), arXiv:1504.04632 [astro-ph.IM].
 - [24] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, *Phys. Rev. D* **82**, 044025 (2010).
 - [25] F. Özel and P. Freire, *Annual Review of Astronomy and Astrophysics* **54**, 401 (2016), <https://doi.org/10.1146/annurev-astro-081915-023322>.
 - [26] C. E. Rhoades and R. Ruffini, *Phys. Rev. Lett.* **32**, 324 (1974).
 - [27] V. Kalogera and G. Baym, *The Astrophysical Journal Letters* **470**, L61 (1996).
 - [28] F. Özel, D. Psaltis, R. Narayan, and A. S. Villarreal, *The Astrophysical Journal* **757**, 55 (2012).
 - [29] J. M. Lattimer, *Annual Review of Nuclear and Particle Science* **62**, 485 (2012), <https://doi.org/10.1146/annurev-nucl-102711-095018>.
 - [30] B. Kiziltan, A. Kottas, M. D. Yoreo, and S. E. Thorsett, *The Astrophysical Journal* **778**, 66 (2013).
 - [31] J. W. T. Hessels, S. M. Ransom, I. H. Stairs, P. C. C. Freire, V. M. Kaspi, and F. Camilo, *Science* **311**, 1901 (2006), astro-ph/0601337.
 - [32] C. D. Bailyn, R. K. Jain, P. Coppi, and J. A. Orosz, *The Astrophysical Journal* **499**, 367 (1998).
 - [33] F. Özel, D. Psaltis, R. Narayan, and J. E. McClintock, *The Astrophysical Journal* **725**, 1918 (2010).
 - [34] W. M. Farr, N. Sravan, A. Cantrell, L. Kreidberg, C. D. Bailyn, I. Mandel, and V. Kalogera, *The Astrophysical Journal* **741**, 103 (2011).
 - [35] M. Fishbach and D. E. Holz, *The Astrophysical Journal Letters* **851**, L25 (2017).
 - [36] C. Talbot and E. Thrane, *The Astrophysical Journal* **856**, 173 (2018).
 - [37] D. Wysocki, J. Lange, and R. O'Shaughnessy, arXiv e-prints, arXiv:1805.06442 (2018), arXiv:1805.06442 [gr-qc].
 - [38] V. Tiwari, *Classical and Quantum Gravity* **35**, 145009 (2018), arXiv:1712.00482 [astro-ph.HE].
 - [39] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, *Publications of the Astronomical Society of the Pacific* **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].
 - [40] S. M. Gaebel, J. Veitch, T. Dent, and W. M. Farr, *Monthly Notices of the Royal Astronomical Society* **484**, 4008 (2019), arXiv:1809.03815 [astro-ph.IM].
 - [41] E. Chase, C. Pankow, and V. Kalogera, (in preparation).

Appendix A: Derivation of the multicomponent FGMC counts posterior

The derivation follows [16, 17].

Consider a time interval T during which gravitational-wave (GW) observations are made. We divide this interval into N_t time-fragments:

$$\delta t = \frac{T}{N_t}. \quad (\text{A1})$$

We make N_t large enough to ensure that at most 1 event occurs over time interval δt .

Let us assume that there are N intervals containing exactly 1 event and $N_t - N$ intervals containing exactly 0 events. We denote by x the properties of an event. Previously (i.e, in [16]), this was simply the ranking statistic. Here, $x \rightarrow \{L, m\}$, where L is the ranking statistic and m is a set of template parameters. Thus, x_i is the only trigger in the i th time interval, with detection statistic value L_i and template parameters m_i . We denote by $\emptyset$ a time interval that contains no events. Thus, $\emptyset_j$ is the j th time interval that contains no events.

The joint probability of these N propositions gives us the likelihood up to a combinatorial constant:

$$\mathcal{L} \propto \prod_{i=1}^N p(x_i | \Lambda_0, \vec{\Lambda}_1) \times \prod_{j=N+1}^{N_t} p(\emptyset_j | \Lambda_0, \vec{\Lambda}_1). \quad (\text{A2})$$

Computing $p(\emptyset_j | \Lambda_0, \vec{\Lambda}_1)$ is a straightforward application of the Poisson distribution for 0 counts in a time interval δt :

$$p(\emptyset_j | \Lambda_0, \vec{\Lambda}_1) \propto \exp(-\lambda). \quad (\text{A3})$$

where λ is the expected number of counts (terrestrial and astrophysical combined) in interval δt . Meanwhile:

$$p(x_i | \Lambda_0, \vec{\Lambda}_1) = p(x_i | 1, \Lambda_0, \vec{\Lambda}_1) p(1 | \Lambda_0, \vec{\Lambda}_1). \quad (\text{A4})$$

where $p(1 | \Lambda_0, \vec{\Lambda}_1)$ is the probability of acquiring exactly 1 event in an interval of duration δt , which in turn can be computed via straightforward application of the Poisson distribution:

$$p(1 | \Lambda_0, \vec{\Lambda}_1) = \lambda \exp(-\lambda) \quad (\text{A5})$$

and λ , the mean number of triggers in interval δt is:

$$\lambda = \frac{1}{N_t} \cdot \left(\Lambda_0 + \sum_{\alpha} \Lambda_{\alpha} \right). \quad (\text{A6})$$

On the other hand, $p(x_i | 1, \Lambda, \vec{\Lambda}_1)$ is the fraction of triggers with detection statistic L_i in bin m_i ,

$$p(x_i | 1, \Lambda_0, \vec{\Lambda}_1) = \frac{\Lambda_0 p(L_i, m_i | \text{noise}) + \sum_{\alpha} \Lambda_{\alpha} p(L_i, m_i | \alpha)}{\Lambda_0 + \sum_{\alpha} \Lambda_{\alpha}}. \quad (\text{A7})$$

Therefore, the likelihood becomes:

$$\mathcal{L} \propto \exp \left(- \left(\Lambda_0 + \sum_{\alpha} \Lambda_{\alpha} \right) \right) \cdot \prod_{i=1}^N \left(\Lambda_0 \cdot p(L_i, m_i | \text{noise}) + \sum_{\alpha} \Lambda_{\alpha} \cdot p(L_i, m_i | \alpha) \right). \quad (\text{A8})$$

The multicomponent counts posterior, up to a normalization constant, thus has the following general form:

$$p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) = p(\Lambda_0, \vec{\Lambda}_1) \cdot \exp \left(- \left(\Lambda_0 + \sum_{\alpha} \Lambda_{\alpha} \right) \right) \cdot \prod_{i=1}^N \left(\Lambda_0 \cdot p(L_i, m_i | \text{noise}) + \sum_{\alpha} \Lambda_{\alpha} \cdot p(L_i, m_i | \alpha) \right). \quad (\text{A9})$$

The distributions (normalized density functions), $p(L_i, m_i | \text{noise})$ and $p(L_i, m_i | \alpha)$ in Equation A9 can be divided into two pieces each:

$$p(L_i, m_i | \text{noise}) = p(L_i | m_i, \text{noise}) \cdot W_0(m_i) \quad (\text{A10})$$

$$p(L_i, m_i | \alpha) = p(L_i | m_i, \alpha) \cdot p(m_i | \alpha) \approx p(L_i | m_i, \text{signal}) \cdot W_{\alpha}(m_i). \quad (\text{A11})$$

where the weights W are defined as:

$$W_0(m_i) \equiv p(m_i | \text{noise}), \quad (\text{A12})$$

$$W_{\alpha}(m_i) \equiv p(m_i | \alpha). \quad (\text{A13})$$

$p(L_i | m_i, \alpha) \approx p(L_i | m_i, \text{signal})$ is a statement of the universality of the ranking statistic distribution L under the foreground model [12, 20, 21]; in other words, the foreground model is not expected to change appreciably for different classes of astrophysical signals.

Appendix B: Evolution of the Counts posterior with the addition of candidate events

1. Updating the Multicomponent counts posterior

Having constructed the multi-component counts posterior, it is useful to investigate how the posterior evolves with the addition of candidate events.

Let $p_N(\vec{\Lambda}_1 | \vec{x})$ be the counts posterior constructed from a set of N candidate events, appropriately normalized. Suppose we now acquire an additional trigger, x_{N+1} . We wish to determine how the inclusion of this trigger modifies the counts posterior. It is straightforward to see, from equation 10, that:

$$p_{N+1}(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) = A \cdot p_N(\vec{\Lambda}_1 | \vec{x}) \cdot [\Lambda_0 + \vec{\Lambda}_1 \cdot \vec{K}(x_{N+1})] \quad (B1)$$

where A is a constant that we can determine via normalization:

$$\int_0^\infty p_{N+1}(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) d\Lambda_0 d\vec{\Lambda}_1 = 1. \quad (B2)$$

Writing $p_{N+1}(\Lambda_0, \vec{\Lambda}_1 | \vec{x})$ in terms of $p_N(\vec{\Lambda}_1 | \vec{x})$ and carrying out the above integral yields:

$$\int_0^\infty p_{N+1}(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) d\Lambda_0 d\vec{\Lambda}_1 = \int_0^\infty A \cdot p_N(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) \cdot [\Lambda_0 + \vec{\Lambda}_1 \cdot \vec{K}(x_{N+1})] d\Lambda_0 d\vec{\Lambda}_1 \quad (B3)$$

$$= A \left[\langle \Lambda_0 \rangle_N + \sum_\alpha K_\alpha(x_{N+1}) \langle \Lambda_\alpha \rangle_N \right] \quad (B4)$$

where the quantities in the angular brackets are the mean values of the counts (whether of terrestrial or astrophysical origin), defined as:

$$\langle \Lambda_\beta \rangle_N \equiv \int_0^\infty \Lambda_\beta p_N(\Lambda_0, \vec{\Lambda}_1) d\Lambda_0 d\vec{\Lambda}_1. \quad (B5)$$

Thus, from equation B2:

$$A(x_{N+1}) = \frac{1}{\langle \Lambda_0 \rangle_N + \sum_\alpha K_\alpha(x_{N+1}) \langle \Lambda_\alpha \rangle_N}. \quad (B6)$$

Therefore, the multicomponent counts posterior is updated to be:

$$p_{N+1}(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) = p_N(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) \cdot \frac{\Lambda_0 + \vec{\Lambda}_1 \cdot \vec{K}(x_{N+1})}{\langle \Lambda_0 \rangle_N + \sum_\alpha K_\alpha(x_{N+1}) \langle \Lambda_\alpha \rangle_N}. \quad (B7)$$

2. Updating the mean value of the counts

The change in the posterior due to the addition of an event leads to changes in the mean values of the Poisson expected counts for each source category.

From equations B5 and B7:

$$\langle \Lambda_\beta \rangle_{N+1} \equiv \int_0^\infty \Lambda_\beta p_{N+1}(\Lambda_0, \vec{\Lambda}_1 | \vec{x}_{N+1}) d\Lambda_0 d\vec{\Lambda}_1 \quad (B8)$$

$$= \frac{\langle \Lambda_0 \Lambda_\beta \rangle_N + \sum_\alpha \langle \Lambda_\alpha \Lambda_\beta \rangle_N K_\alpha(x_{N+1})}{\langle \Lambda_0 \rangle_N + \sum_\alpha K_\alpha(x_{N+1}) \langle \Lambda_\alpha \rangle_N}. \quad (B9)$$

The change in the mean value due to the addition of the $N + 1$ th trigger is:

$$\begin{aligned} \langle \Lambda_\beta \rangle_{N+1} - \langle \Lambda_\beta \rangle_N &= \frac{\langle \Lambda_0 \Lambda_\beta \rangle_N + \sum_\alpha \langle \Lambda_\alpha \Lambda_\beta \rangle_N K_\alpha(x_{N+1})}{\langle \Lambda_0 \rangle_N + \sum_\alpha K_\alpha(x_{N+1}) \langle \Lambda_\alpha \rangle_N} - \langle \Lambda_\beta \rangle_N \\ &= \frac{\text{cov}_N(\Lambda_0, \Lambda_\beta) + \sum_\alpha \text{cov}_N(\Lambda_\alpha, \Lambda_\beta) K_\alpha(x_{N+1})}{\langle \Lambda_0 \rangle_N + \sum_\alpha K_\alpha(x_{N+1}) \langle \Lambda_\alpha \rangle_N} \end{aligned} \quad (B10)$$

where:

$$\text{cov}_N(\Lambda_A, \Lambda_B) \equiv \langle \Lambda_A \Lambda_B \rangle_N - \langle \Lambda_A \rangle_N \langle \Lambda_B \rangle_N \quad (\text{B11})$$

is the covariance. The updated mean may be written in terms of the original mean as:

$$\langle \Lambda_\beta \rangle_{N+1} = \langle \Lambda_\beta \rangle_N + \frac{\text{cov}_N(\Lambda_0, \Lambda_\beta) + \sum_\alpha K_\alpha(x_{N+1}) \text{cov}_N(\Lambda_\alpha, \Lambda_\beta)}{\langle \Lambda_0 \rangle_N + \sum_\alpha K_\alpha(x_{N+1}) \langle \Lambda_\alpha \rangle_N}. \quad (\text{B12})$$

It is interesting to note here that the addition of a highly significant candidate event of a certain astrophysical source category could increase the mean values of the count for that category by more than unity. Indeed, if $K_\beta(x_{N+1}) \gg \langle \Lambda_0 \rangle_N$, and $K_\beta(x_{N+1}) \gg K_{\alpha \neq \beta}(x_{N+1})$, as is the case for certain highly significant events of category β , then Eq. B10 is approximated as:

$$\langle \Lambda_\beta \rangle_{N+1} - \langle \Lambda_\beta \rangle_N \approx \frac{\text{var}_N(\Lambda_\beta)}{\langle \Lambda_\beta \rangle} \quad (\text{B13})$$

where $\text{var}_N(\Lambda_\beta) \equiv \text{cov}_N(\Lambda_\beta, \Lambda_\beta)$. If the variance exceeds the mean, then the addition of that highly significant event would cause an increase in the mean of more than one. This is not so surprising, though perhaps counter intuitive at first. In effect, the addition of a highly significant event of a certain source-category informs the posterior that the rate of events of that source-category is higher than was inferred from the previously available set of triggers.

Appendix C: Recursive Counts Posterior

Given distinct chunks of GW data over which the spacetime volume sensitivity is assumed to be constant, we can construct FGMC counts posteriors for each chunk separately. However, comparing candidate events directly between chunks is not meaningful, because in general the spacetime volume sensitivity will differ from chunk to chunk. Therefore, in order to construct a posterior with all the events from all the chunks, we require some form of weighting involving the spacetime volume sensitivities of the chunks.

The key idea here is that while the Poisson expected counts (the Λ s), and the spacetime volume sensitivities (the $\langle VT \rangle$ s), will change from chunk to chunk, what is assumed to remain constant between chunks is the astrophysical rate of compact binary mergers that we seek to determine.

In the following, we put superscripts to various quantities to identify the chunk of data they correspond to, and the subscripts label the source-category type, as earlier.

We define the total spacetime-volume sensitivity, across all chunks c , for an astrophysical source-category α , as:

$$\langle VT \rangle_\alpha^{\text{tot}} = \sum_c \langle VT \rangle_\alpha^c \quad (\text{C1})$$

and the total Poisson expected counts for category α as:

$$\Lambda_\alpha^{\text{tot}} = \sum_c \Lambda_\alpha^c \quad (\text{C2})$$

Working with the “astrophysical-rate is time-independent” assumption, the following must be true:

$$R_\alpha = \frac{\Lambda_\alpha^{\text{tot}}}{\langle VT \rangle_\alpha^{\text{tot}}} = \frac{\Lambda_\alpha^c}{\langle VT \rangle_\alpha^c}, \forall c \quad (\text{C3})$$

where R_α is the astrophysical rate for category α .

An important point to make here is that the background counts Λ_0^c across chunks are not connected in any way: there is no unchanging R_0 corresponding to these counts. We can, however, resort to the delta-function approximation, and fix the terrestrial counts, for each chunk c , to the total number of candidate events for that chunk: $\Lambda_0^c \rightarrow N^c$.

One can now combine the likelihoods from different chunks to construct a single posterior on the total counts Λ_α^{tot} of each astrophysical source category by making the following change of variables:

$$\Lambda_\alpha^c = \frac{\langle VT \rangle_\alpha^c}{\langle VT \rangle_\alpha^{tot}} \Lambda_\alpha^{tot} \quad (C4)$$

We can now write the combined counts-posterior as:

$$p(\Lambda_0, \vec{\Lambda}^{tot} | \vec{x}) = p(\vec{\Lambda}^{tot}) \left\{ \prod_c \prod_{j=1}^{N^c} [N^c + \vec{\Lambda}^{tot} \cdot \vec{K}^c(x_j^c)] \right\} e^{-\vec{\Lambda}^{tot} \cdot \vec{u}} \quad (C5)$$

where:

$$\vec{K}^c = \left\{ K_\alpha \times \frac{\langle VT \rangle_\alpha^c}{\langle VT \rangle_\alpha^{tot}} \right\}, \alpha = 1, 2, 3, \dots, Q \quad (C6)$$

and:

$$\vec{\Lambda}^{tot} = \{ \Lambda_\alpha^{tot} \}, \alpha = 1, 2, 3, \dots, Q \quad (C7)$$

with Q as the number of astrophysical source categories.

It is worth noting here that the Gstlal pipeline incorporates spacetime volume sensitivity into its ranking statistic, and therefore the need to employ a method to weight candidate events by the sensitive spacetime volume sensitivity, as is done here, is redundant for the Gstlal pipeline. This however may not be true for all gravitational wave detection pipelines.

Appendix D: Connecting terrestrial and astrophysical probabilities with mean values of the Poisson counts

Suppose the prior on the multi-component FGMC posterior (cf. Eq. (10)) has the form: $p(\Lambda_0, \vec{\Lambda}_1) = p(\Lambda_0, \Lambda_\alpha \notin \vec{\Lambda}_1) \Lambda_\alpha^a$. Taking the derivative of $\Lambda_\alpha p(\Lambda_0, \vec{\Lambda}_1 | \vec{x})$ with respect to Λ_α yields:

$$\frac{d}{d\Lambda_\alpha} (\Lambda_\alpha p(\Lambda_0, \vec{\Lambda}_1 | \vec{x})) = (a+1) p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) - \Lambda_\alpha p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) + p(\Lambda_0, \vec{\Lambda}_1 | \vec{x}) \sum_{j=1}^N \frac{\Lambda_\alpha K_\alpha(x_j)}{\Lambda_0 + \vec{K}_\alpha(x_j) \cdot \vec{\Lambda}_1} \quad (D1)$$

The derivative was chosen so that its antiderivative vanishes at $\Lambda_\alpha = 0$ and ∞ when marginalizing.

If the multi-component posterior is appropriately normalized, then, marginalizing out both sides of the above equation with respect to Λ_0 and $\vec{\Lambda}_1$, yields:

$$0 = a + 1 - \langle \Lambda_\alpha \rangle + \sum_{j=1}^N P_\alpha(x_j | \vec{x}) \quad (D2)$$

which, upon rearranging, gives Eq (35).