Deep Multimessenger Search for Compact Binary Mergers in LIGO, Virgo, and Fermi/GBM Data from 2016–2017

Marion Pillas[1] ![ORCID], Tito Dal Canton[1] ![ORCID], Cosmin Stachie[2] ![ORCID], Brandon Piotrzkowski[3] ![ORCID], Fergus Hayes[4] ![ORCID], Rachel Hamburg[1] ![ORCID],
Eric Burns[5] ![ORCID], Joshua Wood[6] ![ORCID], Pierre-Alexandre Duverne[1,7] ![ORCID], and Nelson Christensen[2] ![ORCID]

[1] Université Paris-Saclay, CNRS/IN2P3, IJCLab, F-91405 Orsay, France; marion.pillas@ligo.org
[2] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Artemis, F-06304 Nice, France
[3] University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
[4] SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK
[5] Department of Physics & Astronomy, Louisiana State University, Baton Rouge, LA 70803, USA
[6] NASA Marshall Space Flight Center, Huntsville, AL 35805, USA
[7] Université Paris Cité, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France
Received 2023 June 6; revised 2023 August 25; accepted 2023 August 25; published 2023 October 6

Abstract

GW170817–GRB 170817A provided the first observation of gravitational waves from a neutron star merger with associated transient counterparts across the entire electromagnetic spectrum. This discovery demonstrated the long-hypothesized association between short gamma-ray bursts and neutron star mergers. More joint detections are needed to explore the relation between the parameters inferred from the gravitational wave and the properties of the gamma-ray burst signal. We developed a joint multimessenger analysis of LIGO, Virgo, and Fermi/GBM data designed for detecting weak gravitational-wave transients associated with weak gamma-ray bursts. As such, it does not start from confident (GWTC-1) events only. Instead, we take the full list of existing compact binary coalescence triggers generated with the PyCBC pipeline from the second Gravitational-Wave Observing Run (O2), and reanalyze the entire set of public Fermi/GBM data covering this observing run to generate a corresponding set of gamma-ray burst candidate triggers. We then search for coincidences between the gravitational-wave and gamma-ray burst triggers without requiring a confident detection in any channel. The candidate coincidences are ranked according to a statistic combining each candidate's strength in gravitational-wave and gamma-ray data, their time proximity, and the overlap of their sky localization. The ranking is then converted to a false alarm rate using time shifts between the gravitational-wave and gamma-ray burst triggers. We present the results using O2 triggers, which allowed us to check the validity of our method against GW170817–GRB 170817A. We also discuss the different configurations tested to maximize the significance of the joint detection.

Unified Astronomy Thesaurus concepts: Gamma-ray bursts (629); Gravitational waves (678); Neutron stars (1108)

1. Introduction

Since 2015, the year of the first gravitational-wave (GW) detection GW150914 (Abbott et al. 2016), the Advanced LIGO (Aasi et al. 2015) and Advanced Virgo (Acernese et al. 2015), L-shaped kilometer-scale Michelson laser interferometers, have continued their searches to detect GW events. Sources such as compact binary coalescences (CBCs) are candidates for detection. Possible counterparts of CBCs involving at least one neutron star, for example binary neutron stars (BNSs) and neutron star–black hole (NSBH) mergers, are short gamma-ray bursts (sGRBs) (Eichler et al. 1989; Nakar 2007). GRBs are bursts of highly energetic gamma rays with a duration ranging from less than a second to several minutes. When a BNS or an NSBH merger occurs, an accretion disk may be formed around the resulting black hole and an sGRB jet can be ejected. The fraction of BNS and NSBH mergers leading to GRB emission is currently not precisely known. GRBs are usually classified into two categories based on their duration, represented by the T_{90} statistic, which is instrument-dependent and defined as the time interval in which the integrated photon count increases

from 5% to 95% of the total counts above the background. The spectral hardness of the prompt emission also contributes to the classification: long GRBs tend to be soft while short GRBs are hard (Kouveliotou et al. 1993). Long GRBs, having a duration $T_{90} \geqslant 2$ s, are typically associated with a subclass of core-collapse supernovae (CCSNe) (Woosley 1993; Cano et al. 2017). sGRBs are believed instead to originate mostly in CBC systems containing at least one neutron star, as demonstrated by the joint detection GW170817–GRB 170817A (Abbott et al. 2017a, 2017b, 2017d). There are, however, important exceptions to this classification (Ahumada et al. 2021; Rastinejad et al. 2022).

To answer fundamental questions about sources of GWs and GRBs, for instance to determine the fraction of sGRBs associated with neutron star mergers, or the mechanisms responsible for the formation of a jet, a significantly larger sample of joint GW/GRB observations is needed. To this end, there are currently several independent approaches to searching for joint GW/GRB associations, such as PyGRB (Harry & Fairhurst 2011; Williamson et al. 2014; Abbott et al. 2017c) and X-pipeline (Sutton et al. 2010; Was et al. 2012; Abbott et al. 2017c), which assume that there is a GRB detection and perform a deep search for a nearby GW signal (or more generally GW transients in the case of X-pipeline). On the low-latency timescale, RAVEN (Urban 2016; Abbott et al. 2017c;

Piotrzkowski & LIGO Team 2022) assumes a GW and an electromagnetic detection and checks for their compatibility in time and sky location. By contrast, in the analysis presented in Fletcher et al. (2022), they start from the GW events from the GWTC-3 catalog and they look for potential associations with Fermi-GBM and Swift-BAT data. Other analyses have been performed for specific events, such as Blackburn's method (Blackburn et al. 2015) to search for a counterpart to GW150914 (Connaughton et al. 2016). The LIGO/Virgo O1 —Fermi/GBM (Burns et al. 2019) analysis also computes a false alarm probability assuming a GW detection and checks for a nearby GBM signal. Moreover, a method to search for associations between insignificant GW and GRB candidates during the first GW observing run was proposed in Nitz et al. (2019), and a potential association, named "1-OGC 151030," was found.

Finally, during O2, a joint analysis calculates a *p*-value for the association of GW170817 and GRB 170817A, assuming a GW detection and a GBM detection and checking for compatibility in time and sky localization. However, many of these searches have statistical or computational limitations that prevent their application to a large number of weak candidate events. For example, both X-pipeline and PyGRB require hours to days (running on a CPU cluster) to analyze the GW data around the time of a single GRB and start from the assumption that the GRB is a robust detection. With RAVEN the joint false alarm rate (FAR $[yr^{-1}]$) of the foreground association is proportional to the single FAR in each channel, leading to statistical limitations in the case of sub-threshold events. However, recent versions of RAVEN can also deal with events of not particularly high confidence (Piotrzkowski & LIGO Team 2022) by changing its calculation of the joint FAR.

Motivated by the possibility that sGRBs come from CBCs, and by the joint observation of GW170817 (Abbott et al. 2017d) and GRB 170817A (Goldstein et al. 2017; Savchenko et al. 2017) coming from a common BNS source, we develop a method to search for coincidences between transient events in the Fermi/GBM data and LIGO O2 CBC triggers, in order to increase the number of joint GW/GRB detections. This is a deep search in the sense that it does not consider confident events only. Instead, for the analysis presented in this paper we use the full list of Fermi/GBM triggers (here 780,206 triggers) and GW triggers (here 15,270) coming from O2. The goal is to find pairs of Fermi/GBM and GW triggers that could possibly have a common origin, rank the pairs with a ranking statistic, and assign a FAR to them. The statistic we use here is similar to the ranking statistic in Stachie et al. (2020), which we describe in Section 2.3, but different configurations are tested to compute the terms of this statistic and rank the associations.

The organization of the paper is as follows. In Section 2 we describe the method used to search for GW/GBM pairs. In Section 3 we detail the different configurations tested in this analysis and the results we obtained with each configuration. Section 4 contains a discussion about the main results of this search, and proposes improvements that we can implement in the method for future searches.

2. Method

In the following section, we describe how we rank the potential associations. First the triggers are generated from the GW data and the gamma-ray data independently. Then we find all the possible associations and calculate a ranking statistic based on the properties we have on these triggers (GW and GBM triggers' significance, time and sky proximity between them). Finally, we estimate the statistical significance of each pair. The following subsections present how we generate the GW and gamma-ray candidates and detail the computation of the different terms of the ranking statistic.

2.1. Gravitational-wave Candidate Generation

The search for GW signals produced by a CBC is carried out by several independent pipelines using different techniques to improve detection efficiency and to compare the recorded data with theoretical signals derived from general relativity. Since the parameters of each GW source are not known in advance and must be inferred from the data, the search explores the CBC source parameter space, usually composed of the masses and spins of the components of the binary, by covering it with a grid of model waveforms called a *template bank* (Dal Canton & Harry 2017). In the following analysis we use O2 triggers from the PyCBC pipeline (Usman et al. 2016), which is a modeled matched-filtering-based (Allen et al. 2012) analysis pipeline that identifies CBC events by correlating data with a template bank of waveforms. There are several steps in the matched-filtering techniques used by the pipeline to distinguish noise from signals and measure their significance. The data are collected from all the available detectors and are then scanned to find matches with the waveforms in the template bank. A signal-to-noise ratio (S/N or ρ) time series is computed to find times when the S/N exceeds a predetermined threshold, and a CBC trigger is generated at each maximum in the time series. One major problem that can interfere with GW detection is the instrumental or environmental noise that severely limits the sensitivity of the GW interferometric detectors. This pipeline implements veto techniques to reject these transient non-Gaussian noise *glitches* that create large S/N values but can be easily discriminated from CBC signals. Many other strategies, such as the χ^2 test or the reweighted S/N defined in Equation (6) in Abbott et al. (2016), are used to distinguish signals from these glitches. Finally, to compute the significance of the resulting triggers, which are described by a FAR value representing how often we expect noise to produce a trigger with the same ranking statistic as the candidate in question, a background sample is generated using a time-shift method (Was et al. 2010) between the data of the available detectors. The background generation in our analysis is based on the same time-shift method.

In this analysis, we focus on O2 PyCBC coincident triggers from the two LIGO detectors (LIGO Hanford (H1) and LIGO Livingston (L1)) in which potential binary black hole, BNS, and NSBH signals can be found. For each GW trigger, we compute the posterior distribution for the sky location of the source, under the assumption that the trigger is of astrophysical origin. We use `Bayestar` (Singer & Price 2016) due to its simplicity and low computational cost, through the `pycbc_make_skymap` tool. Although the search triggers are produced from LIGO data only, we also use Virgo data (when available) for the sky localization, as they can lead to significantly narrower posterior distributions.

2.2. Gamma-Ray Candidate Generation

The Fermi Gamma-ray Space Telescope (Michelson et al. 2010) is a space observatory launched on 2008 June 11 and dedicated to the detection of the most energetic phenomena taking place in the Universe through observations of gamma-ray radiation. Aboard Fermi, the Gamma-ray Burst Monitor (GBM; Meegan et al. 2009) instrument is used to observe GRBs coming from the entire sky not occulted by the Earth within the energy range 8 keV–40 MeV. The GBM flight hardware comprises 14 NaI(Tl) detectors, which are used to measure the low-energy spectrum from 8 keV to 1 MeV, and two bismuth germanate detectors with an energy range of ∼200 keV to ∼40 MeV. An onboard "trigger" occurs when there is an increase in the count rates of two or more NaI detectors above an adjustable threshold. The GBM flight software contains algorithms to compute the location of trigger events based on the relative rates in the NaI detectors and to evaluate the probability that a trigger arises from a GRB. In the analysis presented here, we do not use the onboard triggers; instead we use the time-tagged event (TTE) data candidate gamma-ray signals from ground searches of GBM data. The TTE data are time series of photon "counts" for which the time and energy are recorded.

The GBM trigger generation is realized by the so-called GBM Targeted Search (Burns et al. 2019; Goldstein et al. 2019). This produces triggers by searching for an excess of photon counts compatible with GRBs over a variety of overlapping time windows, covering durations from 64 ms to 8.192 s. A log-likelihood ratio (LLR) is computed for each time window, allowing for the generation of GBM triggers labeled with a duration, a mission elapsed time (MET), and an LLR. The LLR calculation uses photon rates produced by a GRB in the GBM detectors that depend on the energy channels in a way that can be predicted after a particular spectral shape has been assumed for the GRB. Finally, the GBM Targeted Search also contains a clustering algorithm that only keeps the trigger in regions of the duration–time plane that have the highest LLR if LLR ⩾ 5. This threshold is justified in Stachie et al. (2020).

We use the clustered GBM Targeted Search results in the analysis presented here. This is justified by our desire to remove the correlation between the triggers and also by the high cost of dealing with the unfiltered triggers. The result of the Targeted Search also gives us the sky localization of the GBM triggers with a sky grid resolution of 5°. The search produces maps of the posterior probability distribution $P(\Omega|D_\gamma)$, which quantifies the probability that the GBM trigger comes from a specific location Ω given the data D_γ. Note that $P(\Omega|D_\gamma)$ may be nonzero over the portion of sky occulted by the Earth; a large $P(\Omega|D_\gamma)$ over the Earth indicates that the GBM trigger is likely to be terrestrial.

The search results can be displayed with a waterfall plot. The top panel of Figure 1 shows the display of GRB 170817A. The x-axis represents the GBM MET and the y-axis shows the different duration windows. The color of each rectangle denotes the LLR value. An sGRB shows up as a particular waterfall shape, due to the progressive drop of the LLR as the search window deviates more and more from the best-fitting time and duration. Far from the sGRB, displayed in the bottom panel of Figure 1, one can also see the statistical fluctuations of the persistent, slowly varying, and ever-present GBM

background. This background can be divided into a photon component and a charged particle component. The former is produced by the actual overall photon flux coming from the ensemble of sources in the celestial sphere around Fermi, such as the Sun, the Earth, the cosmic gamma-ray background, and gamma-ray pulsars. The latter includes cosmic rays and any charged particle interacting with the satellite materials or with the photomultipliers.

2.3. Ranking Statistic

Addressing joint detections, there are many statistics to rank the potential associations. We first consider two simple ranking statistics, inspired by Nitz et al. (2019):

1. *Naive time statistic*: a naive product of GW and GBM likelihoods, including the time overlap, and summarized by

$$\ln(\Lambda) = \frac{\hat{\rho}_g^2}{2} + \mathrm{LLR} + \ln(I_{\Delta t}). \tag{1}$$

2. *Naive time and sky statistic*: a naive product of GW and GBM likelihoods, including the time and skymap overlaps and summarized by

$$\ln(\Lambda) = \frac{\hat{\rho}_g^2}{2} + \mathrm{LLR} + \ln(I_{\Delta t}) + \ln(I_{\Omega}). \tag{2}$$

Here $\hat{\rho}_g$ is the network-reweighted S/N (Usman et al. 2016) described by Equation (3):

$$\hat{\rho}_g = \sqrt{\sum_d \hat{\rho}_d^2} \tag{3}$$

where d represents the available GW detectors. The term $\hat{\rho}_g^2/2$ represents the log-likelihood ratio $\frac{\mathbb{P}(\mathrm{data \mid noise, \ signal})}{\mathbb{P}(\mathrm{data \mid noise})}$ on the GW side, following the computation from Nitz et al. (2017). $I_{\Delta t}$ and I_{Ω} are Bayes factors that quantify the overlap of the posterior distributions for the arrival times (time offset) and sky locations (skymap overlap) inferred separately from the GW and GBM data. These ranking statistics are straightforward to implement, and their results are presented here in Section 3.1. We also justify our choice to not use them in this work and to compute the Bayesian ranking statistic described in the following paragraphs.

When searching for signals in GW and GBM data, the two simple possibilities of noise and noise+signal become a complicated multimessenger rainbow of mutually exclusive hypotheses:

1. (H_{NN}) hypothesis: both GW and GBM data contain their respective noise only.
2. (H_{NS}) hypothesis: GW data contain noise and signal, and GBM contains only noise.
3. (H_{SN}) hypothesis: GW data contain noise, and GBM contains noise and signal.
4. (H_{SS}) hypothesis: both GW and GBM contain noise and signal, but the signals are unrelated.
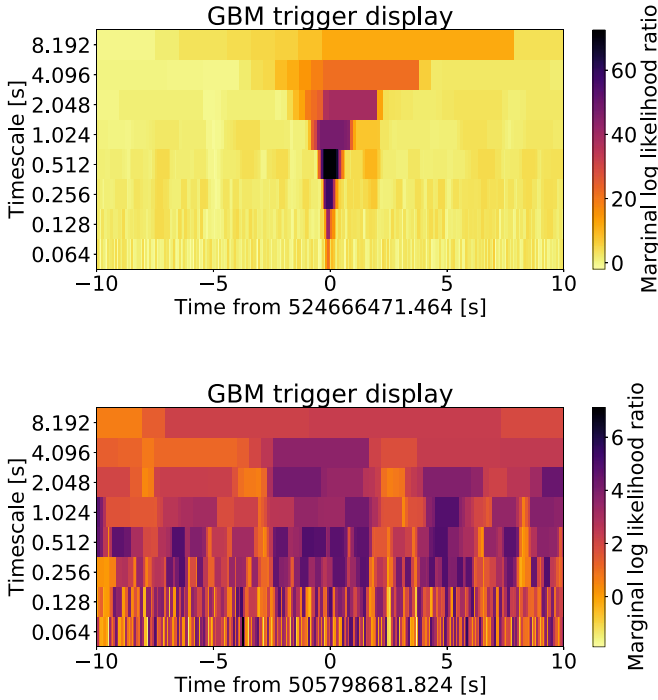5. (H_C) hypothesis: both GW and GBM contain noise and signal, and the signals come from the same source.

GBM trigger display

GBM trigger display

Figure 1. Difference in the display of 20 s windows over the search result for GRB 170817A (top) and background only (bottom).

Here by "noise" in the GW data we also consider the possibility of Gaussian noise with a glitch on top. To make a frequentist multimessenger detection in the usual way, we combine cases 1–4 into a null hypothesis and then test hypothesis 5 versus null. As done in many astrophysical searches (Piotrzkowski et al. 2022), we then define a joint ranking statistic from the GW and GBM data, calculate the corresponding background distribution and finally produce a p-value or FAR. Here hypotheses 1–4 in the list above can be defined as situations that all represent the "background."

When ranking the different GW–GBM candidate combinations, we make the following considerations:

1. A larger GW candidate significance corresponds to a more likely candidate.
2. A larger GBM candidate significance corresponds to a more likely candidate in the GBM channel.
3. The closer in time the GBM candidate is to the GW merger time, the more likely the candidate is.
4. The more the sky localizations overlap, the more confident the candidate is.

Going from the list of different hypotheses, and by a Bayesian derivation, we arrive at the association ranking statistic calculated in Stachie et al. (2020) and derived from Ashton et al. (2018), which can be summarized with the following formula:

$$\Lambda = \frac{P(D_g, D_\gamma | H_C)}{P(D_g, D_\gamma | H_{NN} \bigvee H_{SN} \bigvee H_{NS} \bigvee H_{SS})}, \quad (4)$$

where D_g and D_γ are the data sets from the GW and gamma-ray channels, and (H_i) are the different hypotheses listed above.

Finally, if we assume complete ignorance about prior probabilities, this expression can be simplified as

$$\Lambda = \frac{I_{\Delta t} I_\Omega}{1 + Q_g + Q_\gamma + Q_g Q_\gamma}, \quad (5)$$

where

$$Q_g = Q_g(D_g) = \frac{P(D_g | \text{noise})}{P(D_g | \text{signal})} \quad (6)$$

and

$$Q_\gamma = Q_\gamma(D_\gamma) = \frac{P(D_\gamma | \text{noise})}{P(D_\gamma | \text{signal})} \quad (7)$$

are the single-instrument Bayes factors comparing the noise-only and noise+signal hypotheses in the GW and GBM data respectively. We discuss the calculation of each term in this ranking statistic in Sections 2.4–2.7.

Equation (5) has intuitive limits:

1. If both signals are very marginal, deep into their noise distribution, then $Q_g Q_\gamma \gg Q_g$, $Q_\gamma \gg 1$, and $\Lambda \approx \frac{P(D_g | H_S)}{P(D_g | H_N)} \times \frac{P(D_\gamma | H_S)}{P(D_\gamma | H_N)} \times I_{\Delta t} I_\Omega \ll I_{\Delta t} I_\Omega$. The association is suppressed by both single-instrument Bayes factors, and the parameter overlap must be exceptional for the association to be significant.
2. If one signal is very significant (say the GW) then $Q_g \ll 1$ and Λ only depends on the significance in the other instrument (and the parameter overlap).
3. If we have very significant signals in both GW and GBM data, then $Q_g Q_\gamma \ll Q_g$, $Q_\gamma \ll 1$, and $\Lambda \approx I_{\Delta t} I_\Omega$: the single-instrument significances become irrelevant, and the only thing that can prove the association between the triggers is the overlap in their inferred parameters.

2.4. GBM Bayes Factor Q_γ

Calculating Q_γ accurately would in principle involve a complete parametric model for the photon flux of an sGRB, for the detailed response of GBM to that photon flux, and for the GBM background. This would then allow us to write a likelihood function and marginalize it over the relevant parameter space to compute the Bayes factor. Although attempts have been made to solve parts of this modeling task (see, e.g., Hayashi et al. 2022), this is clearly a major challenge. Instead, here we use an approximate phenomenological model based on a sample of sGRBs observed by GBM and confirmed by other gamma-ray observatories, and a sample of triggers unlikely to be associated with sGRBs.

Specifically, we approximate Q_γ following a 2D kernel density estimation (KDE; Weglarczyk 2018) in the log (duration)–log(LLR) space. The KDE is fit on a sample of positive triggers. These include a sample of 61 GBM-triggered Targeted Search triggers within 3 s of confirmed events (i.e., identified with the GBM untargeted search that has confirmation in other instruments such as INTEGRAL SPI-ACS (Rau et al. 2005) or Swift (Gehrels et al. 2004)) with $T_{90} < 2$ s and >90% probability of being an sGRB during O2, during the third GW observing run (O3), and between observing runs, and on a negative sample. Conservatively, every non-positive trigger (including detector noise) is considered to be negative. The negative sample counts 1,897,956 triggers. The KDE is
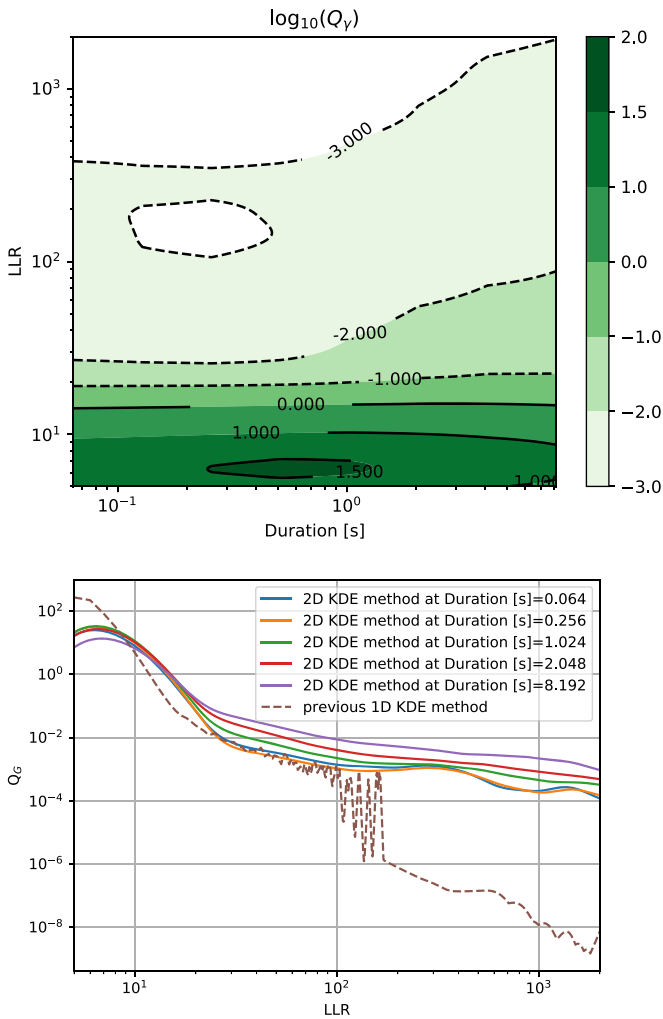
Figure 2. GBM Bayes factor computed with a 2D KDE training with real data over $\log_{10}$(LLR) and $\log_{10}$(duration (s)). Top: behavior of the GBM Bayes factor Q_γ computed with a 2D KDE fitted with real data. Bottom: comparison between the GBM Bayes factor Q_γ computed with the 2D KDE (solid lines) and the 1D KDE trained on LLR and used in Stachie et al. (2020) (dashed line).

then evaluated on the GBM triggers and the ratio of the probability functions presented in Equation (7) is computed. The resulting Q_γ distribution is shown in the top panel of Figure 2. We can observe that the signal-like region is associated with the highest LLR values. Moreover, at fixed LLR, triggers with a shorter duration have a smaller (i.e., signal-like) Q_γ than those with a longer duration.

The bottom panel of Figure 2 presents the present approximation with an earlier study (Stachie et al. 2020), which used a simpler 1D KDE on the LLR only, coupled with some assumptions to extrapolate the distributions beyond the range of the available training set. The 2D KDE method described here produces a similar behavior, but the rapid oscillations in the curve of Bayes factor versus LLR, which were artifacts of the previous method, have disappeared. We can also see that most of the curves with the 2D KDE Bayes factor are above the 1D KDE curve, meaning that with our 2D KDE-based Bayes factor it is harder for a candidate to be signal-like.

2.5. GW Bayes Factor Q_g

In addition, Bayestar also provides two Bayes factors:

1. BSN: Bayes factor signal versus noise compares the probability for a trigger to be signal-like and noise-like. This Bayes factor is qualitatively similar to an S/N and defined as

$$\frac{P(D_g \mid \text{signal} + \text{Gaussian noise})}{P(D_g \mid \text{Gaussian noise})} \quad (8)$$

2. BCI: Bayes factor coherent versus incoherent compares the probability of being a coherent signal versus an incoherent signal. It is defined as

$$\frac{P(D_g \mid C + \text{Gaussian noise})}{P(D_g \mid I + \text{Gaussian noise})} \quad (9)$$

where C and I are respectively the coherent and incoherent signals.

Intuitively, we expect GW astrophysical signals to have both high $\log_{10}$(BSN) and $\log_{10}$(BCI), Gaussian noise to have both small $\log_{10}$(BSN) and $\log_{10}$(BCI), and glitches to have a high $\log_{10}$(BSN) and small $\log_{10}$(BCI).

Two configurations are tested to compute the GW Bayes factor Q_g. The first one is used in the analysis and the second one is presented in Appendix A.

1. We use the logarithm of the BCI, which is the result of a model comparison between a coherent GW signal in the entire network and a signal that is not coherent (the most likely occurrence being a single-detector signal).
2. We compute a KDE-based GW Bayes factor in the $\log_{10}$(BCI) − $\log_{10}$(BSN) plane. We train the KDE on a positive sample composed of GW injections and a background sample made of GW triggers with an FAR above two per day; this is detailed in Appendix A.

The BCI is robust to glitches, contrary to the BSN, which is why we only extract the BCI from the skymaps in the main method presented here. We explore a KDE-based method using both the BCI and the BSN in Appendix A. Indeed, a KDE requires a training sample, so GW injection skymaps should be generated in addition to skymaps for noise and glitches, which is much more complicated than considering only the BCI.

2.6. Time Overlap

The goal of the time offset term $I_{\Delta t}$ is to quantify how probable it is for a pair formed by a GW trigger and a GBM trigger to be separated by a certain amount of time $\Delta t = t_{\text{GBM}} - t_{\text{GW}}$, where t_{GW} is the merger time of the GW candidate estimated by the GW searches, and t_{GBM} is the central time of the GBM trigger with the maximum LLR.

Here we define the time term as

$$I_{\Delta t} = \begin{cases} 1 - \dfrac{|\Delta t|}{30} & \text{if } |\Delta t| < 30 \text{ s,} \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

This is similar to what has been used in Stachie et al. (2020) under the same assumptions. Since we currently have only one joint detection, this choice of prior is essentially arbitrary. The time delay between GW170817 and GRB 170817A was about 1.7 s, so we make the prior assumption that the closer in time the two messengers are, the more likely they are to come from the same source. Our prior extends to much larger delays than 1.7 s, which, although unlikely according to some prompt

emission models, is large enough to allow us to detect so-far-unconfirmed phenomena, such as sGRB precursors (Tsang et al. 2012; Stachie et al. 2022).

2.7. Sky Overlap

The general form of the Bayes factor for the sky proximity is given in Ashton et al. (2018) as

$$I_\Omega = \int \frac{P(\Omega|D_g)P(\Omega|D_\gamma)}{P(\Omega)}d\Omega. \tag{11}$$

This sky overlap Bayes factor I_Ω should be large (strongly positive) when the GW and the GBM triggers are overlapping and well localized (small uncertainty region). It will be close to 0 when the two triggers are localized far from each other. The sky overlap is $I_\Omega \approx 1$ even if the two triggers are overlapping but the skymaps are uninformative (large uncertainty regions on both sides). In that case the posterior $P(\Omega|D_i)$ (with $i = g$ or γ) is almost equal to the prior $P(\Omega)$ in Equation (11), leading to $I_\Omega \approx 1$.

Moreover, when looking for a CBC producing coincident signals in GW data and Fermi/GBM, we know a priori that it cannot be located behind the Earth. Therefore, a more realistic prior (compared to a uniform prior on the whole sky) for the sky location is zero over the Earth, and uniform over the portion of the sky not occulted by the Earth. This leads to

$$P(\Omega) = \begin{cases} 1/\int_{\oplus-} d\Omega = 1/f_{\rm vis}4\pi & \text{if } \Omega \notin \oplus \\ 0 & \text{if } \Omega \in \oplus \end{cases} \tag{12}$$

where

$$f_{\rm vis} = \frac{1}{4\pi}\int_{\bar{\oplus}} d\Omega \tag{13}$$

is the fraction of the sky not occulted by the Earth. Finally, one can rewrite Equation (11) as

$$I_\Omega^{\rm EA} = 4\pi f_{\rm vis}\int_{\bar{\oplus}} P(\Omega|D_g)P(\Omega|D_\gamma)\, d\Omega. \tag{14}$$

In contrast to an all-sky prior, the Earth-avoiding prior systematically reduces the Bayes factor when the events being compared are well separated from the Earth, regardless of how much their skymaps overlap.

When the sky localization of the joint association is most likely behind the Earth, the difference between the sky term with a prior avoiding the Earth and considering a uniform prior over the whole sky (ignoring the presence of the Earth) becomes important. An example is shown in Figure 3. Here, the morphology seen in the spectrogram in the bottom panel of Figure 3 points out that the GW trigger is unlikely to be a compact binary merger signal but rather a glitch (triggering the high mass templates, S/N ≈ 73, and $\chi^2 \approx 259$) and the GBM trigger has a probability of being occulted by the Earth of 83%. Hence, it is unlikely to be a GRB but rather noise coming from behind the Earth or from the Earth itself (e.g., terrestrial gamma-ray flashes; Roberts et al. 2018) triggering the detector. This association is highly unlikely to be an astrophysical association and should be suppressed by its sky overlap term. When we set to zero the probability behind the Earth, the Earth-
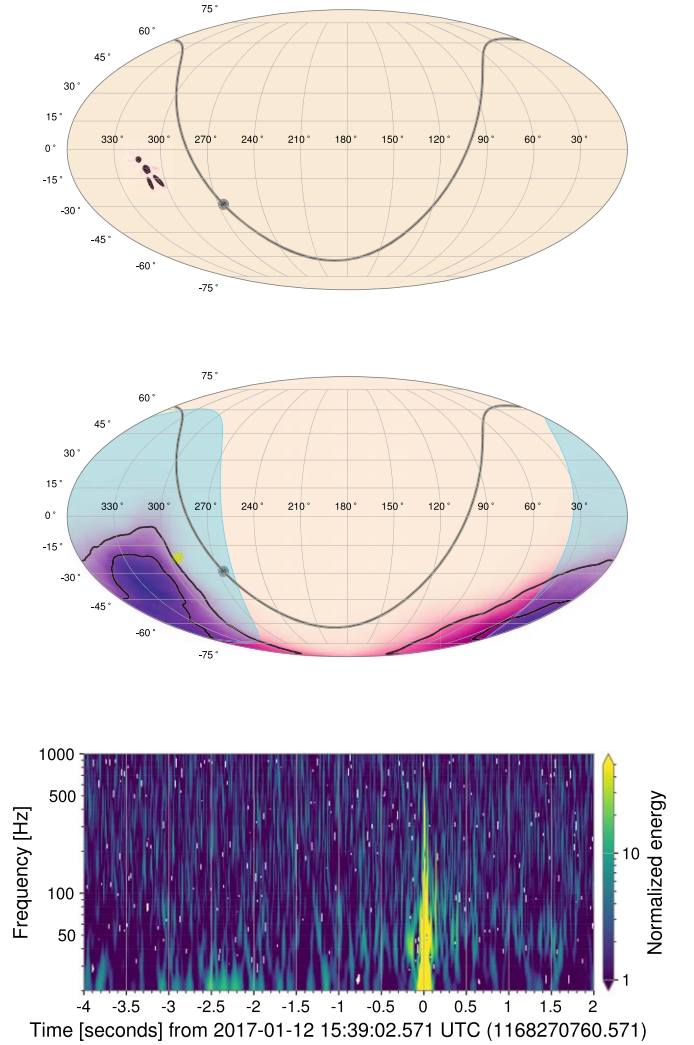

Figure 3. Example of a GW–GBM association that is strongly suppressed by the Earth-avoiding prior. Top: sky localization of the GW trigger. Middle: sky localization of the GBM trigger. The gray solid line represents the Galactic plane, and the Sun is represented by the yellow star in the GBM skymap. The blue region is the Earth's location. The two contour levels represent the 90% and 50% credible regions. The darker the purple, the higher the probability. Bottom: spectrogram of L1 data around the GW trigger.

avoiding sky term is $I_\Omega^{\rm EA} \approx 3.08 \times 10^{-10}$, which indicates that the association is more likely to be an accidental coincidence. In contrast, with the all-sky prior, it becomes $I_\Omega \approx 4.06$. Since both triggers have most of their posterior distribution behind the Earth, the sky overlap term of the association should be very close to zero, which is the case with $I_\Omega^{\rm EA}$. The discrete computation of the Earth-avoiding prior is detailed in Appendix B.

2.8. Calculation of the Significance

Eventually, we want to assign a statistical significance to each foreground pair. To do this, we need to compare the foreground to the background. The FAR (or the inverse false alarm rate, IFAR [yr]) of a candidate gives a quantitative idea of how often noise generates a candidate. It is defined as the rate of background with ranking statistic equal to or higher than the one observed for the candidate in question
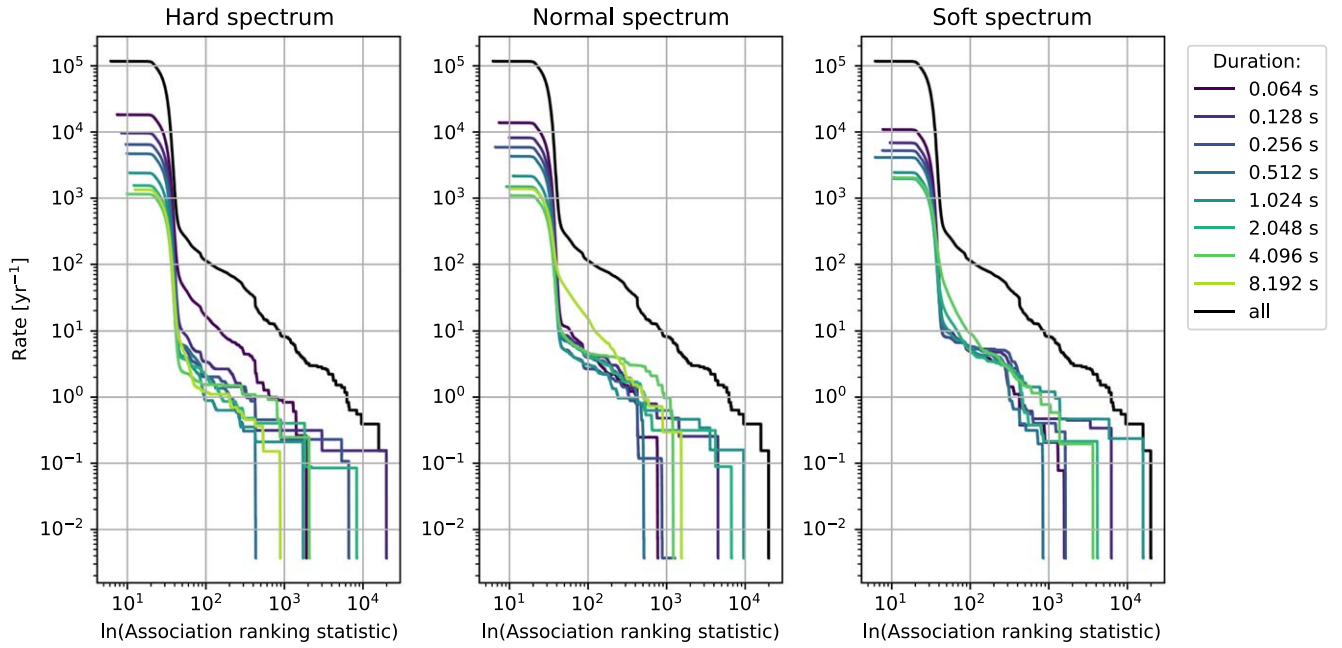
Figure 4. Background rate as a function of the naive time ranking statistic. Left: associations with a hard GBM spectrum. Middle: associations with a normal GBM spectrum. Right: associations with a soft GBM spectrum. The black curve represents all the associations regardless of the duration and spectral hardness of their GBM triggers.

Table 1
Properties of the First Four Most Significant Background Associations with the Naive Time Statistic

	GW Properties		GBM Properties			Joint Properties		
Rank	Merger Time	$\hat{\rho}_g$	Duration (s)	Spectrum	LLR	Δt (s)	Time Shift (s)	$\ln(\Lambda)$
1	1,169,534,855.804	7.694	0.128	hard	19,827.204	10.060	18,700	19,856.394
2	1,169,552,972.042	7.602	0.128	hard	19,827.204	−6.178	36,800	19,855.867
3	1,169,529,460.304	7.406	0.128	hard	19,827.204	5.560	13,300	19,854.422
4	1,169,558,255.479	7.406	0.128	hard	19,827.204	10.385	42,100	19,854.202

Note. $\hat{\rho}_g$: network-reweighted S/N. Δt: time delay between the GBM and the GW triggers. $\ln(\Lambda)$: natural logarithm of the association rank value.

(Usman et al. 2016; Morrás et al. 2023). Therefore, it gives a frequentist interpretation to the ranking statistic. We want to generate the foreground sample, i.e., identify pairs of GW–GBM triggers that could come from a common astrophysical source, sort these pairs with our ranking statistic, and compute their FAR. A background sample is needed to assign to each foreground event. The same set of triggers is used to generate both the foreground and the background. To compute the background sample, we use a time-slide method (Was et al. 2010) in which we time-shift the GBM triggers by a predetermined offset, and we look for (fake) coincidences between the GW triggers and the time-shifted GBM triggers.

To choose an optimal background interval (i.e., the number of time slides and the time difference between two time slides) one must consider some requirements. First, keep the interval to a minimum in order to have "local" estimates of the background, so the detector is in the same state during the background time interval to avoid bias. Second, the interval chosen should not be extremely large, to avoid a high computational cost. Third, the interval needs to be large enough to reach an interesting FAR ($<1/[1000$ yr]) for claiming a discovery. All these requirements lead us to start with a shorter time interval of $\pm 5 \times 10^4$ s ($\sim$27 hours around the GW trigger) in 100 s steps for testing, and then

use the larger interval from $−1.80070 \times 10^5$ s to $+1.8 \times 10^5$ s ($\sim$4 days around the GW trigger) in 70 s steps for final estimates of significance. The lack of symmetry between the lower and upper boundaries is chosen for computational reasons.

The time offset chosen to compute the background must be greater than twice the maximum time offset of 30 s, considered in this analysis to be a nonphysical time delay between a CBC and any possible GRB emission resulting from it. We use a ± 70 s offset to accumulate background associations. We repeat this process multiple times, each with a different nonzero integer multiple of 70 s, and accumulate the background distribution of Λ values. For the foreground, we rank the pairs with the same statistic but without time-shifted GBM triggers in order to find the potential GW–GBM candidate pairs.

3. Configurations and Results

To check the validity of our method against GW170817–GRB 170817A we apply it to the data from O2. We first present in Section 3.1 the results obtained using the naive time ranking statistic described in Section 2.3 and computing the background associations with time slides from $−5 \times 10^4$ to $+5 \times 10^4$ with steps of 100 s. We then focus on the main

analysis of this paper using the Bayesian ranking statistic defined above. The background associations are computed by shifting with time slides from -1.80070×10^5 to $+1.8 \times 10^5$ with steps of 70 s and several configurations are tested to maximize the significance of GW170817–GRB 170817A:

1. In Section 3.2 we present configuration $n°1$ consisting in the computation of the significance of the foreground associations by separating the associations by GBM spectral value and GBM duration. The Targeted Search uses with three spectral models (Goldstein et al. 2020): a "soft" Band function (Band et al. 1993), a "normal" Band function, and a "hard" exponentially cutoff power law. By separating by spectral value and duration we will compare GW170817–GRB 170817A to associations with the same characteristics on the GBM trigger side. This prevents us from comparing GW170817–GRB 170817A to loud associations whose properties correspond less to GRBs coming from neutron star mergers. So we treat each different (spectral value–duration) pair as an independent search and we apply a final trial factor to the FARs to account for the number of searches.
2. Second, we describe in Section 3.3 the computation of the significance of the foreground associations without the separation by GBM spectral value and GBM duration (configuration $n°2$).
3. Finally, in Section 3.4 we discuss configuration $n°3$: no separation by GBM spectral value and GBM duration and application of a preliminary cut of the GW triggers based on the FAR. We remove the triggers with a FAR above 2 per day, threshold inspired by GWTC-3 (Abbott et al. 2021).

3.1. Naive Ranking Statistics

We first run this analysis on O2 data with the naive time and naive time and sky ranking statistics summarized by Equation (1) and Equation (2). Background associations are computed using time shifts going from $-50,000$ to $+50,000$ s with a step of 100 s. The results of the ranking statistic including only the time proximity and the one including the time and sky proximity are very similar, so we only present the naive time ranking statistic.

Figure 4 represents how the background behaves. The associations are shown separated by GBM spectral value (the three panels) and GBM duration (the colored curves). Here, the rate of some curves is an order of magnitude smaller than the rate of others. In the right panel of Figure 4 the curve representing the associations containing an 8.192 s–soft GBM trigger (lightest green curve) does not appear. The GBM Targeted Search intentionally removed such GBM triggers because they contaminate the background. In these plots we can see that the background associations go to extremely high values of association rank. Indeed, as shown in Table 1 and Figure 5, the naive time and naive time and sky statistics are limited by random coincidences between actual bright short GRBs and random noise in the GW detectors. Moreover, we can see that adding the sky overlap here would not affect the results because of the extremely high values of association rank we get. Since the GBM LLR for bright events is several orders of magnitude larger than any of the other quantities in the ranking statistic, the ranking statistic just reduces to the LLR
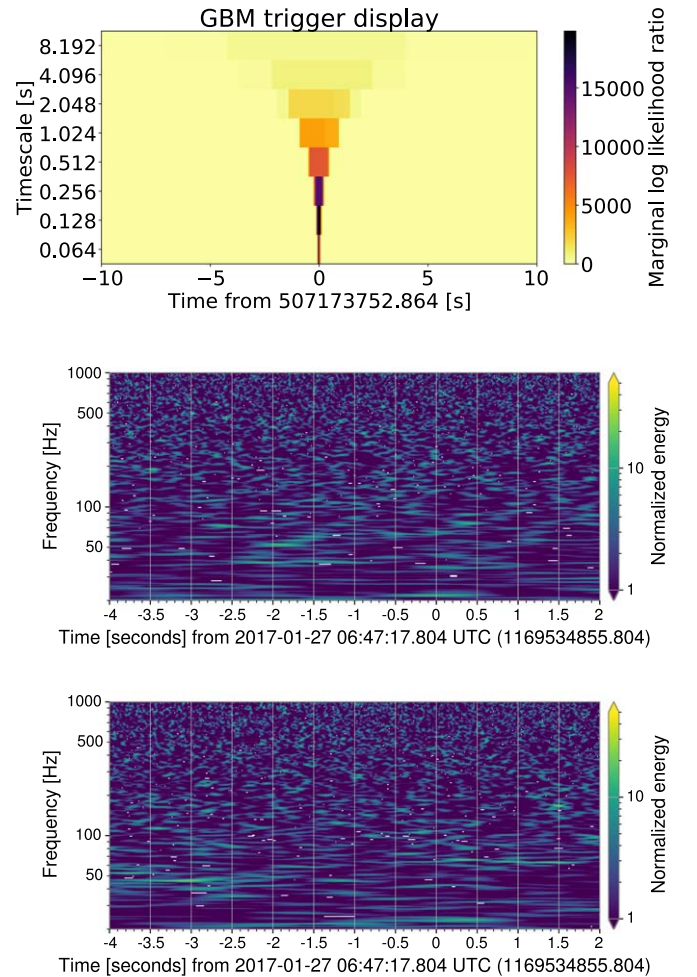

Figure 5. Display of the first most significant background association with the naive time statistic. The GBM trigger here is related to GRB 170127C and no GW confident event has been found at the GPS time of the GW trigger. Top: waterfall plot representing the GBM trigger. Middle: spectrogram representing the GW trigger in the H1 detector. Bottom: spectrogram representing the GW trigger in the L1 detector.

for bright GRBs and the other properties become irrelevant. Therefore, the naive statistics cannot be used unless the various quantities summed together have similar magnitudes.

Finally, if one computes Λ in the case of GW170817–GRB 170817 with the naive time statistic, a value of $\ln(\Lambda) \approx 580$ is found, which will not give a significant FAR compared to the most significant background associations from Table 1 (with $\ln(\Lambda) \approx 19,856$) that are contaminated by extremely loud triggers in the GBM channel.

3.2. Bayesian Ranking Statistic and Separation of the Associations by GBM Spectral Value and GBM Duration

We now switch to the Bayesian statistic (Equation (5)) and present the background behavior, then focus on the top background and foreground associations.

3.2.1. Background Associations

Figure 6 represents how the statistic behaves for the background associations. The associations are separated by GBM spectral value (the three panels) and duration. As for the
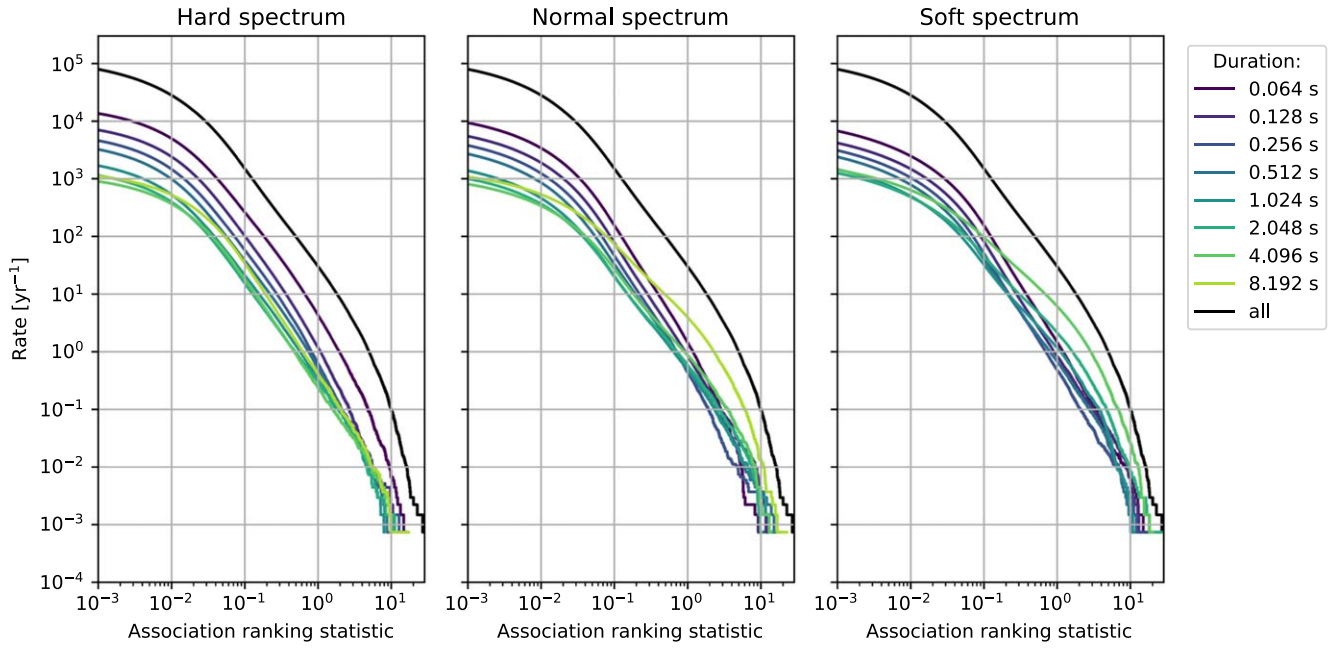
Figure 6. Background rate as a function of the Bayesian ranking statistic (configuration 1). Left: associations with a hard GBM spectrum. Middle: associations with a normal GBM spectrum. Right: associations with a soft GBM spectrum. The black curve represents all the associations regardless of the duration and spectral hardness of their GBM triggers.

Table 2
Properties of the First Four Most Significant Background Associations with Configuration 1

	GW Properties		GBM Properties				Joint Properties			
Rank	Merger Time	Q_g	Duration (s)	Spectrum	LLR	Q_γ	Δt (s)	I_Ω^{EA}	Time Shift (s)	Λ
1	1,172,327,682.418	7.47×10^{-1}	2.048	soft	328.1	1.84×10^{-3}	0.158	48.4	−134,640	27.5
2	1,177,306,450.914	1.13×10^{-1}	0.256	soft	1595.4	2.15×10^{-4}	13.854	55.4	22,650	26.8
3	1,165,069,774.280	2.05×10^{-1}	8.192	normal	64.28	1.53×10^{-2}	1.216	28.8	−104,540	22.6
4	1,165,070,949.873	1.63×10^{-1}	4.096	soft	27.66	5.23×10^{-2}	0.679	25.2	−16,410	20.1

Note. Q_g: GW Bayes factor. Q_γ: GBM Bayes factor. Δt: time delay between the GBM and the GW triggers. I_Ω^{EA}: sky overlap value. Λ: association rank value.

naive time ranking statistic (Figure 4), the rate of some curves is an order of magnitude smaller than the rate of others. It goes above 1 only for a small rate with a maximum association rank of ~27 at a background rate of almost 10^{-3} yr^{-1}; in other words, we would expect one fake (background) association to have such an association rank per 10^3 yr. Otherwise, most of the background has an association rank smaller than 1. Table 2 shows the properties of the first four top ranked background associations and Figure 7 displays the first one. In the first row, in the GBM channel, the candidate has a signal-like Q_γ. It is reported on the Gamma-ray Coordinates Network (GCN; Barthelmy et al. 2000) as GRB $n°510119909$ (Fermi GBM Team 2017). On the GW side, the candidate has a signal-like Q_g but it is more likely to be noise when looking at the H1 and L1 spectrograms and considering that we did not find any GW confident event at the GPS time of this trigger. More generally, the top background comprises associations with signal-like GBM candidates and GW triggers that are more likely to be noise. It is composed of very diverse GBM triggers, bright and less bright GBM triggers (large and small LLR), a large range of duration (going from 0.064 s to 8.192s), and all spectral values (hard, normal, and soft). This diversity reassures us that they are accidental coincidences.

3.2.2. Foreground Associations and Significance

The FAR is computed by counting the number of associations in the background with a higher association ranking statistic than the foreground association in question. One can then build the cumulative rate as a function of the IFAR (defined as 1/FAR) for the foreground associations as shown in Figure 8. Here, if we have a joint detection we will see the foreground (blue curve) deviating from the expectation (orange region).

The significance of the foreground presented in Figure 8 has been computed by separating the associations by pairs of GBM spectral value and GBM duration. In Figure 8, one association deviates at 3σ from the background with an IFAR ~28 yr. Table 3 presents the first four most significant foreground associations with this configuration. Here, the top two are related to GW170817–GRB 170817A. The second association corresponds to the so-called "soft tail" of GRB 170817A (Abbott et al. 2017a, 2017b; Goldstein et al. 2017). The display of GW170817–GRB 170817A is shown in Figure 9. With this configuration, GW170817–GRB 170817A is discovered at barely 3σ, which is not highly significant. Here, the L1 spectrogram has been displayed using the data after subtraction
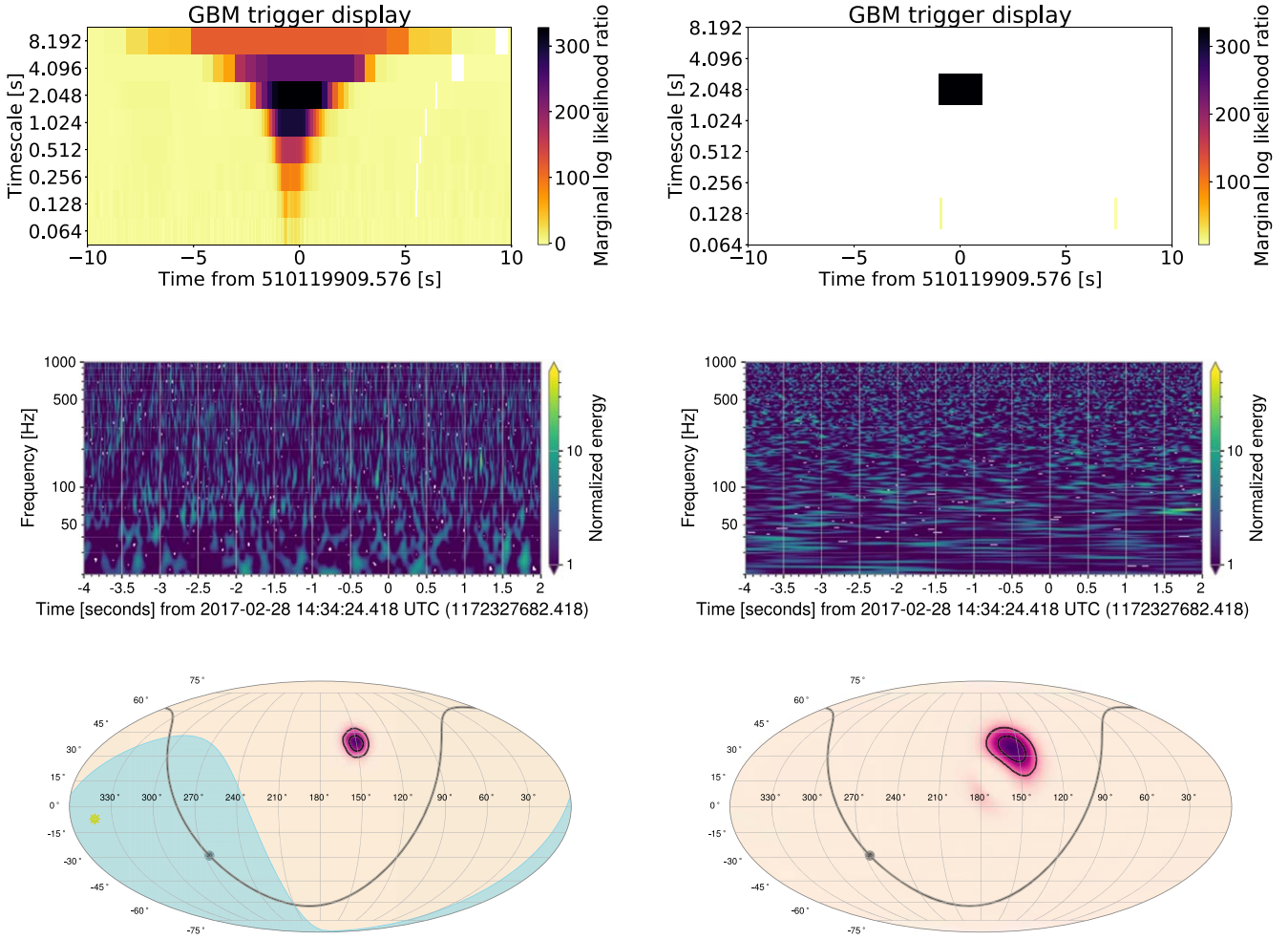
Figure 7. Display of the most significant background association with configuration 1. Top: GBM trigger display with waterfall plots of the GBM trigger (left) and clustered GBM trigger (right). Middle: GW trigger display of spectrograms from H1 (left) and L1 (right). Bottom: sky localization of the association showing the GBM trigger (left) and GW trigger (right). The gray solid line represents the Galactic plane and the yellow star in the GBM skymap represents the Sun. The blue region is the Earth's location. The two contour levels represent the 90% and 50% credible regions. The darker the purple, the higher the probability.
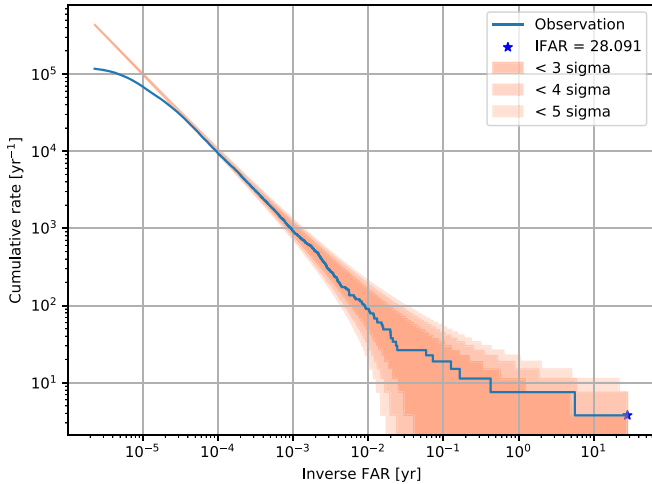

Figure 8. Cumulative rate as a function of the IFAR for foreground (solid line) with configuration 1. The foreground represents associations between Fermi/GBM candidates and LIGO candidates with no time shift.

of the glitch. In Table 3, the next associations after GW170817–GRB 170817A are more likely composed of noise in both GBM and GW channels despite their signal-like Q_g.

Their spectrograms, S/N time series, and the O2 confident GW events list have been checked. These associations are not significant since they have a small IFAR, so they are more likely to be accidental coincidence. Finally, the third most significant foreground association in Table 3 appears to have a negative delay, meaning that the GBM trigger is detected before the GW trigger, so it would be surprising if they have a common origin when we put it in the context of astrophysics, in addition to the low IFAR of this association. We stop at the fourth top foreground association because the rest of the table does not contain association with IFAR large enough to be investigated. Moreover, the beginning of the foreground curve deviates from the expectation computed with the background. This behavior is understood and explanations are given in Appendix C.

3.3. Bayesian Ranking Statistic without the Separation of the Associations by GBM Spectral Value and GBM Duration

When calculating the FAR of an association, we can compare the associations with either the entire background sample or with only a subset of it, and then apply a trial factor based on the number of "bins" into which we split the background sample. In Section 3.2, we have considered splitting the background by both
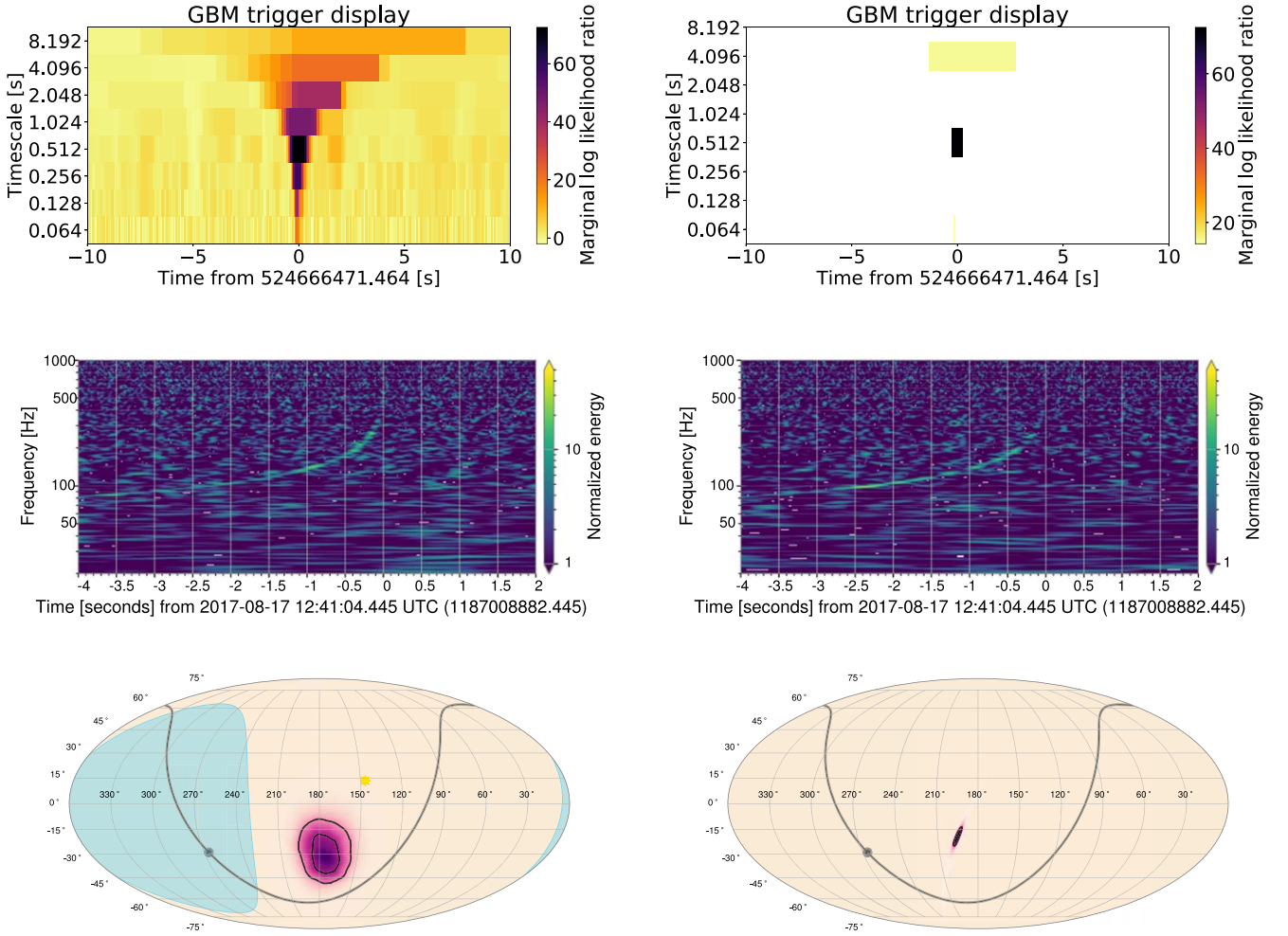
Figure 9. Display of the most significant foreground association with configuration 1 corresponding to GW170817–GRB 170817A. Top: GBM trigger display with waterfall plots of the GBM trigger (left) and clustered GBM trigger (right). Middle: GW trigger display of spectrograms from H1 (left) and L1 (right) in which we can see a glitch on the top of the signal. (However, in the data we use for the analysis, this glitch has been gated.) Bottom: sky localization of the association showing the GBM trigger (left) and GW trigger (right). The gray solid line represents the Galactic plane, and the Sun is represented by the yellow star in the GBM skymap. The blue region is the Earth's location. The two contour levels represent the 90% and 50% credible regions. The darker the purple, the higher the probability.

Table 3
Properties of the First Four Most Significant Foreground Associations with Configuration 1

	GW Properties		GBM Properties				Joint Properties			
Rank	Merger Time	Q_g	Duration (s)	Spectrum	LLR	Q_γ	Δt (s)	I_Ω^{EA}	Λ	IFAR (yr)
1	1,187,008,882.445	6.31×10^{-6}	0.512	normal	72.51	1.91×10^{-3}	2.02	17.2	16.0	28.091
2	1,187,008,882.445	6.31×10^{-6}	4.096	soft	15.38	8.67×10^{-1}	2.72	27.9	13.6	5.618
3	1,168,226,845.160	3.76×10^{-2}	4.096	hard	16.61	5.31×10^{-1}	−6.62	3.31	1.63	0.426
4	1,185,721,264.338	7.29×10^{-3}	0.064	soft	13.05	1.75	0.638	6.22	2.20	0.165

Note. The two first associations correspond to GW170817–GRB 170817A and its soft tail. Q_g: GW Bayes factor. Q_γ: GBM Bayes factor. Δt: time delay between the GBM and the GW triggers. I_Ω^{EA}: sky overlap value. Λ: association rank value.

GBM duration and GBM spectral index. One can also compute the significance of the foreground without separating by GBM spectral value and GBM duration. The results when removing this separation are shown in Figure 10. By comparing the joint associations GW170817–GRB 170817A to many more background associations, all GBM spectral values and GBM duration included, we still have a 3σ deviation from the expected background; however, we increase the IFAR of the joint detection from 28 yr to 90 yr. Indeed, when we compute the FAR by separating the associations in configuration n°1 and thus treat each search as an independent search, we need to multiply the significance by a trial factor to account for the number of searches. Removing this trial factor increases the value of the IFAR of GW170817–GRB 170817A in configuration n°2. Later we decide not to separate the associations by GBM spectral value and GBM duration.

Table 4
Properties of the Background Association with the Same GBM Duration and Spectral Value as GRB 170817A That Limits the Significance of GW170817–GRB 170817A in Configuration $n°1$

	GW Properties		GBM Properties				Joint Properties		
Rank	Merger Time	Q_g	Duration (s)	Spectrum	LLR	Q_γ	Δt (s)	I_Ω^{EA}	Λ
1	1,176,213,122.254	2.98×10^{-1}	0.512	normal	176.6	1.05×10^{-3}	−6.78	17.2	17.8

Note. Q_g: GW Bayes factor. Q_γ: GBM Bayes factor. Δt: time delay between the GBM and GW triggers. I_Ω^{EA}: sky overlap value. Λ: association rank value.
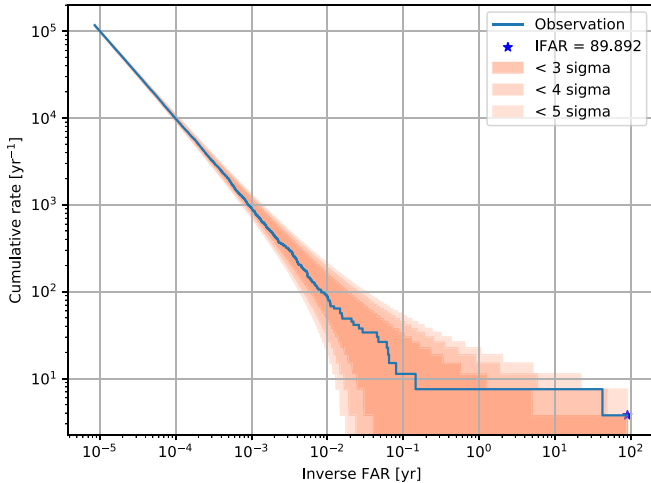
Figure 10. Cumulative rate as a function of the IFAR for the foreground (solid line). The foreground represents associations between Fermi/GBM candidates and LIGO triggers with no time shift. Here the significance is computed with configuration 2 without separating the associations by pairs of GBM spectral value and GBM duration.

3.4. Bayesian Ranking Statistic and Preselection of GW Triggers Based on Their FAR

3.4.1. Limitations of GW170817–GRB 170817A Significance

The poor significance of GW170817–GRB 170817A presented in Sections 3.2 and 3.3 demonstrates that our configuration is not as satisfactory as it should be, and investigation of what limits the joint detection's significance shows that some background associations have a higher association ranking statistic than GW170817–GRB 170817A. As an example, Table 4 describes a background association with the same GBM duration and spectral value as GRB 170817A and a higher association ranking statistic. The GBM trigger in this association corresponds to a real GRB and, on the GW side, the IFAR of the GW trigger is about 4.265×10^{-5} yr, which is noise-like.

So in configuration $n°1$, when we separate the associations by GBM spectral hardness and duration, such association limits the significance of the joint detection. However, this association is more likely to be an accidental coincidence and should be suppressed by the FAR of its GW candidate. One can conclude that its poor significance is mainly due to the large number of GW triggers we have to consider (mainly composed of noise), and to maximize the significance we decide to apply a cut on the GW triggers based on their FAR value. We choose a threshold of 2 per day, inspired by the choices made in GWTC-3.

3.4.2. Background Associations

Figure 11 shows a lower maximum association ranking statistic than Figure 6 (from 27.5 in configuration $n°1$ to 13.5 in configuration $n°3$) with a rate of almost $10^{-3}\,\mathrm{yr}^{-1}$ and it is dominated by soft 4.096 s and hard 0.064 s GBM triggers. Although this configuration does not separate by GBM spectral hardness and duration we still display the background associations separated for illustration. This justifies our wish to remove this separation since the background associations in the three panels of Figure 11 seem to behave similarly. The background comprises associations with signal-like GBM candidates and noise-like GW candidates. It is composed of very diverse GBM triggers, a large range of LLR and durations (going from 0.064 to 8.192 s) on the GBM trigger side, and all kinds of spectral values (hard, normal, and soft). As previously, a diverse background reassures us that they are more likely accidental coincidences.

3.4.3. Foreground Associations and Significance

Figure 12 shows that we now have a discovery at more than 4σ contrary to the results without preselecting the GW triggers. This joint association has an IFAR that is higher than 1348 yr. Since we do not have background associations ranked high enough to compare this foreground association with, this IFAR of the joint detection is now a limit, and the actual IFAR value of this association is greater than this.

Table 5 shows the first four most significant foreground associations. Now the top three are related to GW170817–GRB 170817A. The cut of the GW trigger has brought a third weak extra GBM trigger connected to GRB 170817A and coupled with GW170817. Figure 13 shows the three GBM triggers of GRB 170817A that correspond to the three most significant associations. The fourth ranked foreground association does not contain significant candidates in both channels. Its Q_γ is noise-like and its Q_g is signal-like but L1 and H1 data do not show any excess of power at the time of the GW trigger. Finally, its joint IFAR is also not extremely high, so this association is more likely to be accidental.

Discarding GW triggers highly likely to be noise before running this analysis allows us to remove loud background (time-shifted) associations containing noise in the GW channel and limiting the significance of GW170817–GRB 170817A such as those in Table 4. We finally found an optimal configuration that enables the maximization of GW170817–GRB 170817A's significance. Table 5 also demonstrates that GW170817–GRB 170817A was the only GW–GRB astrophysical association that has been observed in the GW and GBM data during O2 because the fourth most significant foreground association does not have a sufficiently high IFAR to be investigated as an interesting association.
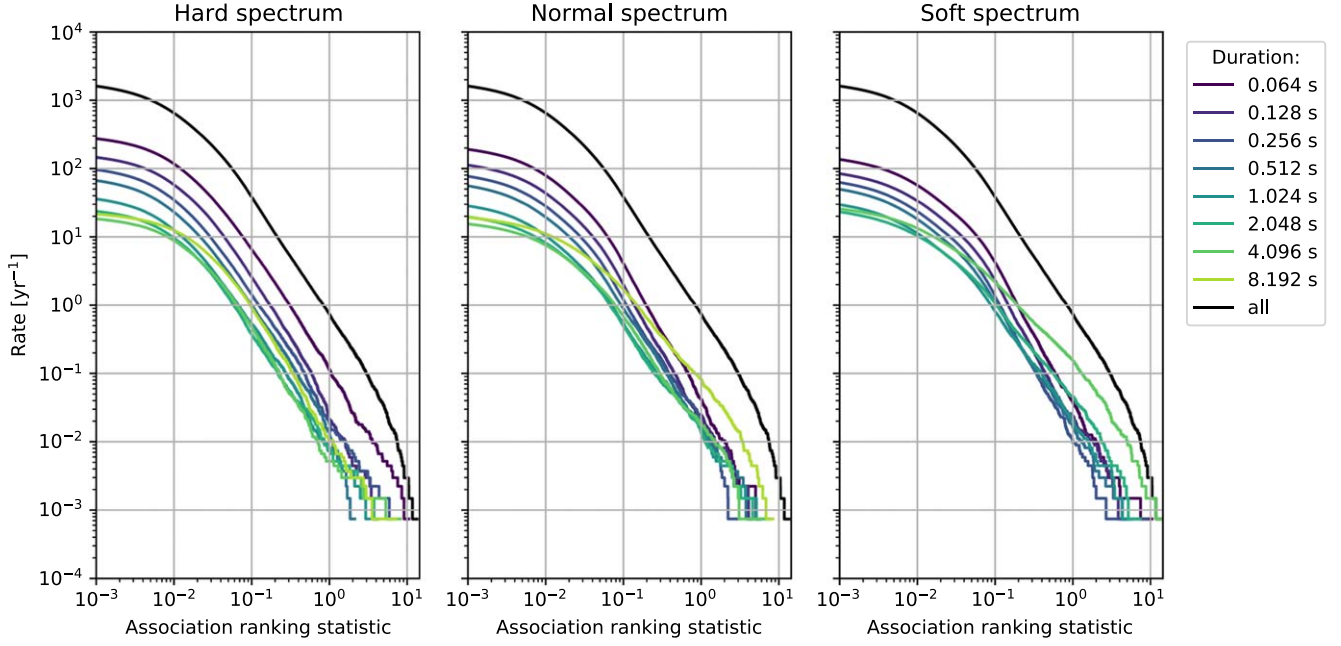
Figure 11. Background association when we cut the GW triggers on their FAR before the analysis (configuration 3). Although this configuration does not separate by GBM spectral hardness and duration we still display the associations separated for illustration. Background rate is shown as a function of the Bayesian ranking statistic. Left: associations with a hard GBM spectrum. Middle: associations with a normal GBM spectrum. Right: associations with a soft GBM spectrum. The black curve represents all the associations regardless of the duration and spectral hardness of their GBM triggers.
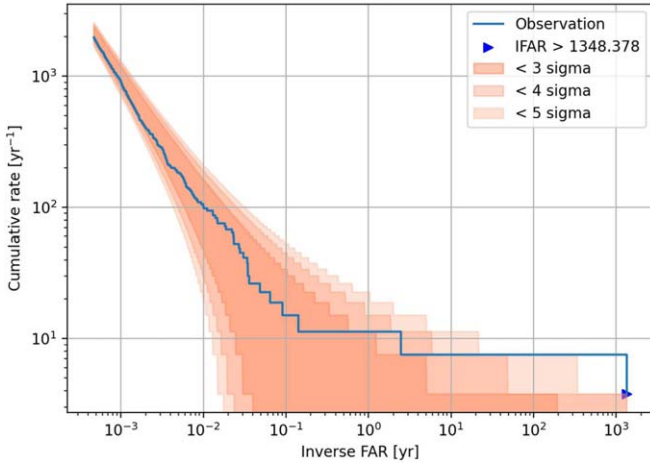
Figure 12. Cumulative rate as a function of the IFAR for the foreground (solid line). The foreground represents associations between Fermi/GBM candidates and LIGO triggers with no time shift. Here a preliminary cut is applied to the GW triggers based on their FAR (configuration n°3).

4. Conclusion

We presented a method to search for associations in a symmetric way between GW triggers from LIGO and Virgo interferometers and data from Fermi/GBM. We sorted the associations thanks to a ranking statistic introduced in Stachie et al. (2020) based on the time and spatial overlap and the significance of the GW and the GBM candidates. The method described here has been applied to PyCBC and Fermi/GBM triggers from O2 to check its validity against GW170817–GRB 170817A. Several configurations have been tested to maximize the significance of this joint detection. Indeed, when we have to analyze a large amount of noise on the GW side, GW170817–GRB 170817A is not highly significant. When we preselect the

GW triggers on their FAR to reduce their rate, we rediscovered GW170817–GRB 170817A with the three most significant foreground associations related to this joint detection. The first two have an IFAR >1348 yr. Other configurations could be tested to maximize the significance of the joint detection such as removing the GW triggers with masses incompatible with neutron stars or using a more realistic prior for the time offset, since currently we allow the GBM trigger to happen long before the GW trigger. In Zhang (2019), the time delay between the GW and the GRB is written as a function of the time for the central engine to launch a relativistic jet Δt_{jet}, the time for the jet to penetrate through and break out from the surrounding medium Δt_{bo}, and the time after breakout for the jet to reach the energy dissipation radius where the observed γ-rays are emitted Δt_{GRB}, thanks to the following formula: $\Delta t_{GW-GRB} = (\Delta t_{jet} + \Delta t_{bo} + \Delta t_{GRB})(1 + z)$. Starting from this review (Zhang 2019) and understanding the jet and GRB physics in BNS systems would lead to an estimation of Δt_{GW-GRB} that could be used as our time offset prior in future analysis.

We are also considering improving the calculation of the GW Bayes factor. Thus, a new calculation of Q_g with a KDE is proposed in Appendix A and we may test it in future work. Finally, we did a search that is optimized for short GRBs since they are believed to be produced by neutron star mergers, contrary to long GRBs that are due to CCSNe. However, this classification is now questioned since Rastinejad et al. (2022) recently observed cases where a kilonova, the remnant of neutron star mergers, has been associated with a long GRB. One other improvement would be to extend the search to all types of GRBs by, for example, taking a sample of short and long GRBs as a signal training sample in the calculation of GBM Bayes factor.

Furthermore, improvements in the filtering and clustering algorithm of the GBM Targeted Search can be explored. Indeed, Figure 13 shows that the clustering generated three

Table 5
Properties of the First Four Most Significant Foreground Associations with Configuration $n°3$

	GW Properties		GBM Properties				Joint Properties			
Rank	Merger Time	Q_g	Duration (s)	Spectrum	LLR	Q_γ	Δt (s)	I_Ω^{EA}	Λ	IFAR (yr)
1	1,187,008,882.445	6.31×10^{-6}	0.512	normal	72.51	1.91×10^{-3}	2.02	17.2	16.0	>1348
2	1,187,008,882.445	6.31×10^{-6}	4.096	soft	15.38	8.67×10^{-1}	2.72	27.9	13.6	>1348
3	1,187,008,882.445	6.31×10^{-6}	0.064	hard	14.32	9.20×10^{-1}	1.86	2.86	1.40	2.474
4	1,185,284,217.254	1.18×10^{-2}	0.064	normal	9.457	10.0	12.71	5.05	0.261	0.142

Note. The top three contain GW170817–GRB 170817A. Q_g: GW Bayes factor. Q_γ: GBM Bayes factor. Δt: time delay between the GBM and GW triggers. I_Ω^{EA}: sky overlap value. Λ: association rank value.
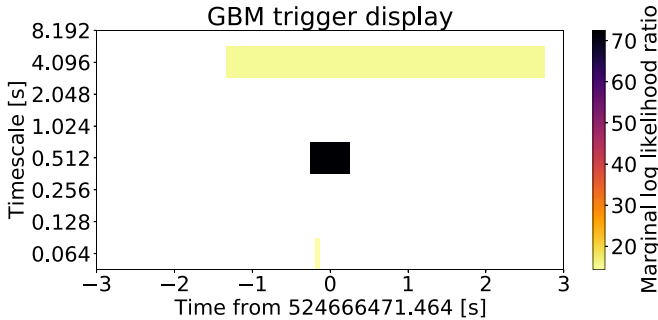

Figure 13. Waterfall plot showing GRB 170817A. The three rectangles at three different timescales are the results of the clustering that produced three GBM triggers for this event. These three are at the top of the table of the most significant foreground associations.

triggers instead of one for the same GRB. This indicates that some residual correlations are present in the GBM triggers.

The method presented with the optimal configuration will be used on the GWTC-3 released PyCBC, GstLAL, and MBTA triggers and Fermi/GBM data covering the duration of O3, in the hope of discovering associations that have not been discovered yet. Finally, the approach we presented in this paper is not limited to Fermi/GBM data. Other missions like the future SVOM mission (Atteia et al. 2022) can use a similar approach if they produce their own form of GBM Bayes factor Q_γ and provide independent sky localizations. Thus, another potential improvement would be considering information from multiple GRB monitors.

Acknowledgments

Appendix A
KDE-based Method to Compute Q_g

The current GW Bayes factor does not allow us to discriminate correctly between noise and an astrophysical GW signal. For instance, Table 2 shows that the four highest ranked background associations have signal-like GW Bayes factor ($Q_g < 1$). Thus, a new way to compute Q_g should be

Table 6
Injection Parameters for the Positive Sample Used to Train the KDE

Parameter	Distribution
Polarization	$\mathcal{U}([0, 2\pi])$
Phase at coalescence	$\mathcal{U}([0, 2\pi])$
$\cos \iota$	$\mathcal{U}([-1, 1])$
Position	$\mathcal{U}(\mathbb{S}(0, [40, 250]\ \text{Mpc}))$
Mass	$\mathcal{N}(\mu = 1.4\ M_\odot, \sigma = 0.1\ M_\odot)$
Spin	$\mathcal{N}(\mu = 0, \sigma = 0.01)$

Note. The ι parameter corresponds to the inclination of the binary. The positions of the simulated signals are generated uniformly in a shell $\mathbb{S}$ with inner radius 40 Mpc and outer radius 250 Mpc.

investigated for future analysis to properly separate the noise-like GW triggers from the interesting GW candidates.

The skymaps produced by `Bayestar` (Singer & Price 2016) provide the BCI (the current Q_g) along with the BSN. Consequently, it should be possible to use a KDE-based method to compute Q_g with a KDE trained in the $\log_{10}(\text{BCI}) - \log_{10}(\text{BSN})$ plane. We also preprocess the data using the logarithm here in order to avoid dealing with values very different from each other. The expected behavior is summarized in the following list:

1. Gaussian noise should have both small $\log_{10}(\text{BCI})$ and $\log_{10}(\text{BSN})$ values.
2. Glitches should have a small $\log_{10}(\text{BCI})$ value and a high $\log_{10}(\text{BSN})$ value.
3. Astrophysical GW signals should in principle have both high $\log_{10}(\text{BCI})$ and high $\log_{10}(\text{BSN})$ values.

Thus, we constructed a GW Bayes factor using these parameters. We built a signal sample with 1000 BNS injections and GW events coming from all runs. The background sample is composed of 2000 O2 GW triggers with a FAR >2/day. The injections of the signal sample were done in Gaussian and stationary noise with the parameters summarized in Table 6. A cut on the optimal S/N has been applied: only signals with a network S/N above 8 and an S/N above 5.5 in at least one of the interferometers were kept. The remaining signals were localized with `Bayestar` to generate the skymaps.

The results of the GW Bayes factor are presented in Figure 14. Here we preprocessed the data with a quantile transform (Pedregosa et al. 2011) that flattens the distributions and moves them between 0 and 1. We can clearly see a separation between the signal-like and the noise-like regions. We also checked that this Bayes factor behaves as expected by applying the KDE on a validation sample (stars in the bottom
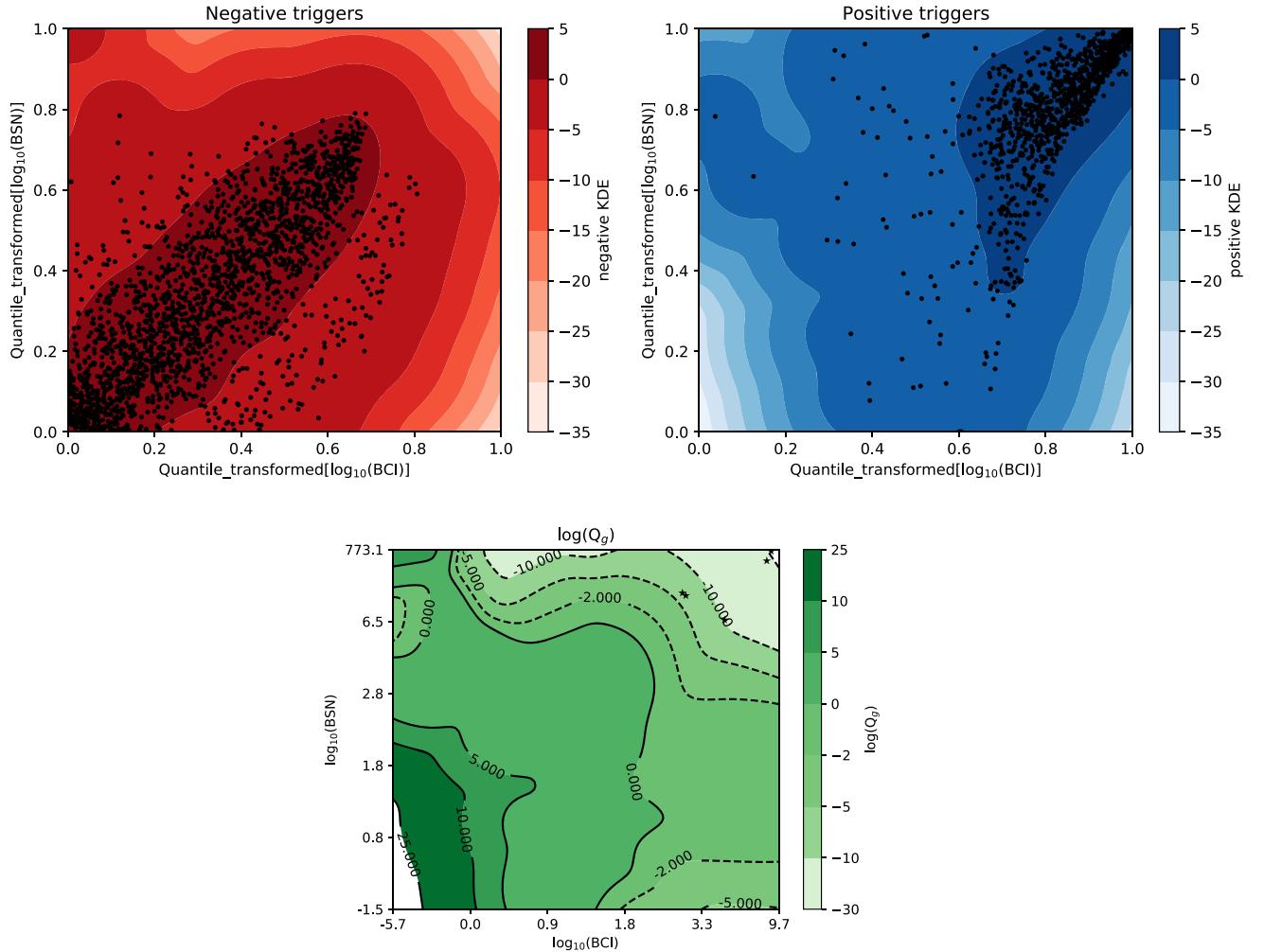
Figure 14. Summary of the KDE training and evaluation on a simulated sample. The signal-like region corresponds to a Q_g close to 0 (i.e., $\log(Q_g)$ negative). Top left: on the negative sample: O2 GW triggers with a FAR > 2/day. Top right: on the positive sample: O3 injections and confident GW events from O2 and O3. Bottom: Bayes factor distribution in the $\log_{10}(\mathrm{BCI}) - \log_{10}(\mathrm{BSN})$ plane. Here the stars are a validation sample composed of four GW events that were not included in the positive sample.

plot of Figure 14). Here, all the GW events from the validation sample are in the signal-like region with strongly negative $\log_{10}(Q_g)$, which is consistent with the expectations. However, the $\ln(Q_g) = 0$ curve in Figure 14 is almost vertical from $\log_{10}(\mathrm{BSN}) \approx -1.5$ to $\log_{10}(\mathrm{BSN}) \approx 6.5$, so except for very high BSN, Q_g seems to depend only on the BCI. As mentioned in Section 2.5, computing a KDE is complicated because some assumptions need to be made. We need to generate skymaps for GW injections, noise, and glitches. We also need to choose a *bandwidth* for the KDE computation, and results can be strongly bandwidth-dependent. Thus, using only the BCI might be an acceptable approximation and reduces both the number of assumptions we need to make and the computational cost of Q_g.

Another method that could be investigated in the future consists in using the calculation of the probability of astrophysical origin, *p-astro* (Abbott et al. 2019), which is based on binary system rate densities. However, for now the estimation of the BNS merger rate is not robust enough because we have not observed enough of these events, but hopefully we will be able to improve this estimation through future GW observing runs.

Appendix B
Discrete Computation of Overlaps of HEALPix Skymaps

A discrete HEALPix localization array from `Bayestar` or Fermi/GBM, hpa, is normalized such that each pixel contains the probability for the source to be inside that pixel, which means that

$$\sum_{i}^{N} \mathtt{hpa}[i] = 1 \tag{B1}$$

where N is the number of pixels in the HEALPix array. We must compare this expression with the normalization of a continuous probability density,

$$\int P(\Omega|\mathtt{hpa})d\Omega = 1. \tag{B2}$$

Hence, to go from the continuous version to the discrete one, we need the following replacement:

$$\int \rightarrow \sum_{i}^{N} \tag{B3}$$

$$P(\Omega|\mathtt{hpa})d\Omega \rightarrow \mathtt{hpa}[i]. \tag{B4}$$

The discrete form of Equation (11) is then, assuming we have three HEALPix arrays for the prior and posteriors,

$$I_\Omega = \sum_i^N \frac{\texttt{hpa\_gw}[i]\texttt{hpa\_gbm}[i]}{\texttt{hpa\_prior}[i]}. \quad (B5)$$

With a uniform prior on the sky, Equation (11) becomes

$$I_\Omega^{\mathrm{AS}} = 4\pi \int P(\Omega|D_g)P(\Omega|D_\gamma)\, d\Omega. \quad (B6)$$

In the case of Equation (B6), the $d\Omega$'s no longer cancel out: one remains at the denominator and needs to be replaced with the HEALPix pixel area $\Delta\Omega$. Therefore

$$I_\Omega^{\mathrm{AS}} = \frac{4\pi}{\Delta\Omega} \sum_i^N \texttt{hpa\_gw}[i]\texttt{hpa\_gbm}[i]. \quad (B7)$$

The case of Equation (14) is similar, but we can note that the sum must avoid the pixels covered by the Earth:

$$I_\Omega^{\mathrm{EA}} = \frac{4\pi}{\Delta\Omega} \sum_{i\notin\oplus}^N \texttt{hpa\_gw}[i]\texttt{hpa\_gbm}[i]. \quad (B8)$$

Appendix C
Modeling the Curve of the Cumulative Rate as a Function of the Inverse False Alarm Rate

As mentioned in Section 3.2, Figure 6 shows a rate turnover on the left of the plot. This is due to the fact that each pair (GBM temporal–spectral value) does not have the same rate. GRBs with short timescale, for instance, are occurring more often due to their short duration. This behavior is explicit when looking at Figure 6, in which we can see that each curve does not start at the same rate. So when we separate the associations by GBM spectral value and GBM duration each pair does not start to contribute to the cumulative rate at the same IFAR. Investigations are made to model the expectation curve (in orange) to have the same turnover. We consider here the results from configuration $n°1$ (i.e., no cut on GW triggers, before applying the search and separation between GBM spectral value and duration). Figure 15 shows that when we sort by GBM duration, all the curves do not start at the same rate. One
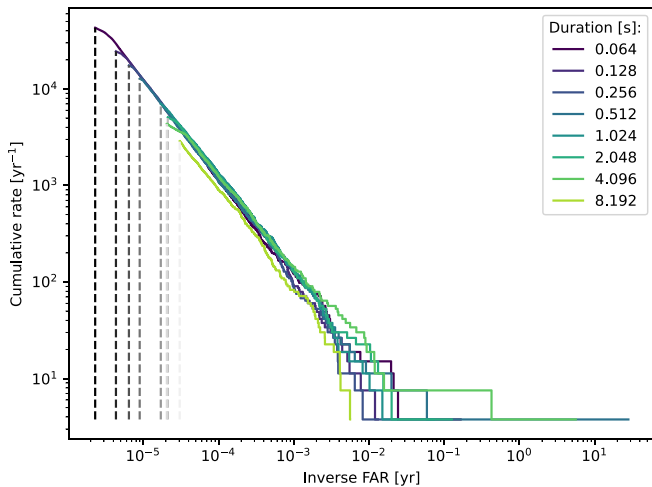

Figure 15. Cumulative rate as a function of IFAR [yr] for different GBM durations. The gray dashed lines show how to bin the curve depending on the duration of the GBM triggers.
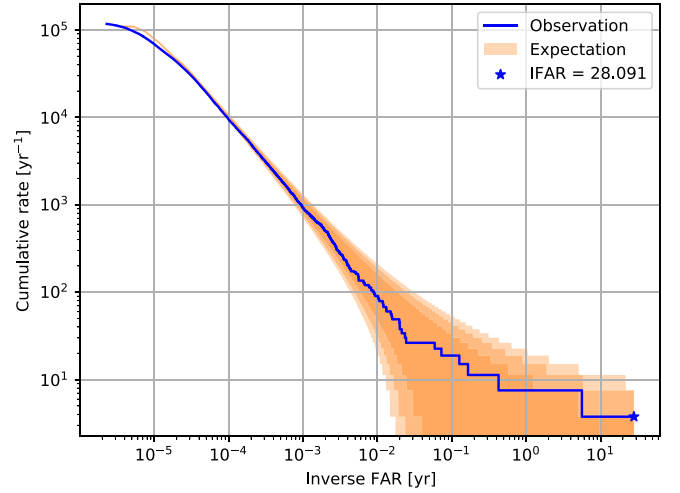

Figure 16. Modeling the turnover of the curve of the cumulative rate as a function of the IFAR in configuration $n°1$ when we separate the associations by GBM spectral hardness and duration value.

can create bins with edges defined by the dashed lines in Figure 15. Each bin is represented by its mean IFAR value. Then, one can weigh the bins by the cumulative number of triggers that contribute to each bin.

This can be summarized by Equations (C1) and (C2), where i represents each bin's index of the lower edge:

$$R_i = \mathrm{Rate}\left(\frac{\texttt{edge}[i] + \texttt{edge}[i+1]}{2}\right) \quad (C1)$$

$$R_{i'} = R_i \times \frac{\sum_0^{i+1} \texttt{nb\_triggers}}{\sum_0^\infty \texttt{nb\_triggers}}. \quad (C2)$$

Here we have represented each bin by the IFAR value in its center, so we take only one value per bin to model the left of the curve.

To check the validity of this modeling, one can compute the significance of the foreground in configuration $n°1$ when we separate the analysis by GBM spectral value and duration. The results are represented in Figure 16. To improve this modeling, one could define bins for each (spectral hardness–duration) pair (and not only duration as presented here) and apply a weight for each of these bins. This could be something to investigate for future analysis.

ORCID iDs

Marion Pillas ⓘ https://orcid.org/0000-0003-3224-2146
Tito Dal Canton ⓘ https://orcid.org/0000-0001-5078-9044
Cosmin Stachie ⓘ https://orcid.org/0000-0002-5723-2342
Brandon Piotrzkowski ⓘ https://orcid.org/0000-0001-8919-0899
Fergus Hayes ⓘ https://orcid.org/0000-0001-7628-3826
Rachel Hamburg ⓘ https://orcid.org/0000-0003-0761-6388
Eric Burns ⓘ https://orcid.org/0000-0002-2942-3379
Joshua Wood ⓘ https://orcid.org/0000-0001-9012-2463
Pierre-Alexandre Duverne ⓘ https://orcid.org/0000-0002-3906-0997
Nelson Christensen ⓘ https://orcid.org/0000-0002-6870-4202

References

Aasi, J., Abbott, B. P., Abbott, R., et al. 2015, CQGra, 32, 074001
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, PhRvD, 93, 122003
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, ApJL, 848, L12
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, ApJL, 848, L13
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, ApJ, 841, 89
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017d, PhRvL, 119, 161101
Abbott, R., Abbott, T. D., Acernese, F., et al. 2021, arXiv:2111.03606
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019, PhRvX, 9, 031040
Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, CQGra, 32, 024001
Ahumada, T., Singer, L. P., Anand, S., et al. 2021, NatAs, 5, 917
Allen, B., Anderson, W. G., Brady, P. R., Brown, D. A., & Creighton, J. D. E. 2012, PhRvD, 85
Ashton, G., Burns, E., Canton, T. D., et al. 2018, ApJ, 860, 6
Atteia, J. L., Cordier, B., & Wei, J. 2022, IJMPD, 31, 2230008
Band, D., Matteson, J., Ford, L., et al. 1993, ApJ, 413, 281
Barthelmy, S. D., Cline, T. L., Butterworth, P., et al. 2000, in AIP Conf. Proc. 526, Gamma-ray Bursts, 5th Huntsville Symp., ed. R. M. Kippen, R. S. Mallozzi, & G. J. Fishman (Melville, NY: AIP), 731
Blackburn, L., Briggs, M. S., Camp, J., et al. 2015, ApJS, 217, 8
Burns, E., Goldstein, A., Hui, C. M., et al. 2019, ApJ, 871, 90
Cano, Z., Wang, S.-Q., Dai, Z.-G., & Wu, X.-F. 2017, AdAst, 2017, 8929054
Connaughton, V., Burns, E., Goldstein, A., et al. 2016, ApJL, 826, L6
Dal Canton, T., & Harry, I. W. 2017, arXiv:1705.01845
Eichler, D., Livio, M., Piran, T., & Schramm, D. N. 1989, Natur, 340, 126
Fermi GBM Team 2017, GCN/FERMI Notice 510119909, NASA, https://gcn.gsfc.nasa.gov/other/510119909.fermi
Fletcher, C., Wood, J., Crnogorcevic, M., Hamburg, R., & Veres, P. 2022, AAS Meeting, 240, 348.01
Gehrels, N., Chincarini, G., Giommi, P., et al. 2004, ApJ, 611, 1005
Goldstein, A., Fletcher, C., Veres, P., et al. 2020, ApJ, 895, 40
Goldstein, A., Hamburg, R., Wood, J., et al. 2019, arXiv:1903.12597
Goldstein, A., Veres, P., Burns, E., et al. 2017, ApJL, 848, L14

Harry, I. W., & Fairhurst, S. 2011, PhRvD, 83, 084002
Hayashi, K., Fujibayashi, S., Kiuchi, K., et al. 2022, PhRvD, 106, 023008
Kouveliotou, C., Meegan, C. A., Fishman, G. J., et al. 1993, ApJL, 413, L101
Meegan, C., Lichti, G., Bhat, P. N., et al. 2009, ApJ, 702, 791
Michelson, P. F., Atwood, W. B., & Ritz, S. 2010, RPPh, 73, 074901
Morrás, G., Siles, J. F. N., García-Bellido, J., & Morales, E. R. 2023, PhRvD, 107, 023027
Nakar, E. 2007, PhR, 442, 166
Nitz, A. H., Dent, T., Canton, T. D., Fairhurst, S., & Brown, D. A. 2017, ApJ, 849, 118
Nitz, A. H., Nielsen, A. B., & Capano, C. D. 2019, ApJL, 876, L4
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825
Piotrzkowski, B. & LIGO Team 2022, APS Meeting Abstracts, 2022, Y13.002
Piotrzkowski, B., Baylor, A., & Hernandez, I. M. 2022, CQGra, 39, 085010
Rastinejad, J. C., Gompertz, B. P., Levan, A. J., et al. 2022, Natur, 612, 223
Rau, A., Kienlin, A. v., Hurley, K., & Lichti, G. G. 2005, A&A, 438, 1175
Roberts, O. J., Fitzpatrick, G., Stanbro, M., et al. 2018, JGRA, 123, 4381
Savchenko, V., Ferrigno, C., Kuulkers, E., et al. 2017, ApJL, 848, L15
Singer, L. P., & Price, L. R. 2016, PhRvD, 93, 024013
Stachie, C., Dal Canton, T., Burns, E., et al. 2020, CQGra, 37, 175001
Stachie, C., Dal Canton, T., Christensen, N., et al. 2022, ApJ, 930, 45
Sutton, P. J., Jones, G., Chatterji, S., et al. 2010, NJPh, 12, 053034
Tsang, D., Read, J. S., Hinderer, T., Piro, A. L., & Bondarescu, R. 2012, PhRvL, 108, 011102
Urban, A. L. 2016, Monsters in the Dark: High Energy Signatures of Black Hole Formation with Multimessenger Astronomy, University of Wisconsin-Milwaukee, https://dc.uwm.edu/etd/121
Usman, S. A., Nitz, A. H., Harry, I. W., et al. 2016, CQGra, 33, 215004
Was, M., Bizouard, M.-A., Brisson, V., et al. 2010, CQGra, 27, 015005
Was, M., Sutton, P. J., Jones, G., & Leonor, I. 2012, PhRvD, 86, 022003
Weglarczyk, S. 2018, ITM Web Conf., 23, 00037
Williamson, A., Biwer, C., Fairhurst, S., et al. 2014, PhRvD, 90, 122004
Woosley, S. E. 1993, ApJ, 405, 273
Zhang, B. 2019, FrPhy, 14, 64402