6-10-2018

Coincident Detection Significance in Multimessenger Astronomy

G. Ashton
Max Planck Institute for Gravitational Physics (Albert Einstein Institute)

E. Burns
NASA Goddard Space Flight Center

T. Dal Canton
NASA Goddard Space Flight Center

T. Dent
Max Planck Institute for Gravitational Physics (Albert Einstein Institute)

H. B. Eggenstein
Max Planck Institute for Gravitational Physics (Albert Einstein Institute)

See next page for additional authors

Recommended Citation

Ashton, G., Burns, E., Canton, T., Dent, T., Eggenstein, H., Nielsen, A., Prix, R., Was, M., & Zhu, S. (2018). Coincident Detection Significance in Multimessenger Astronomy. *Astrophysical Journal, 860* (1) https://doi.org/10.3847/1538-4357/aabfd2

The Astrophysical Journal, 860:6 (7pp), 2018 June 10 https://doi.org/10.3847/1538-4357/aabfd2

Coincident Detection Significance in Multimessenger Astronomy

G. Ashton[1], E. Burns[2], T. Dal Canton[2], T. Dent[1], H.-B. Eggenstein[1], A. B. Nielsen[1], R. Prix[1], M. Was[3], and S. J. Zhu[1,4]
[1] Max Planck Institute for Gravitational Physics (Albert Einstein Institute), D-30167 Hannover, Germany; gregory.ashton@ligo.org
[2] NASA Postdoctoral Program Fellow, Goddard Space Flight Center, Greenbelt, MD 20771, USA
[3] Laboratoire d'Annecy-le-Vieux de Physique des Particules (LAPP), Université Savoie Mont Blanc, CNRS/IN2P3, F-74941 Annecy, France
[4] Max Planck Institute for Gravitational Physics (Albert Einstein Institute), D-14476 Potsdam-Golm, Germany
Received 2017 December 14; revised 2018 March 21; accepted 2018 April 21; published 2018 June 6

Abstract

We derive a Bayesian criterion for assessing whether signals observed in two separate data sets originate from a common source. The Bayes factor for a common versus unrelated origin of signals includes an overlap integral of the posterior distributions over the common-source parameters. Focusing on multimessenger gravitational-wave astronomy, we apply the method to the spatial and temporal association of independent gravitational-wave and electromagnetic (or neutrino) observations. As an example, we consider the coincidence between the recently discovered gravitational-wave signal GW170817 from a binary neutron star merger and the gamma-ray burst GRB 170817A: we find that the common-source model is enormously favored over a model describing them as unrelated signals.

Key words: gamma-ray burst: general – gravitational waves – methods: statistical – neutrinos – stars: neutron

1. Introduction

On 2017 August 17, the observation by LIGO-Virgo of GW170817, a binary neutron star coalescence (BNS; Abbott et al. 2017b, 2017a), and by *Fermi* and *INTEGRAL* of GRB 170817A, a short gamma-ray burst (GRB; Goldstein et al. 2017; Savchenko et al. 2017), began an unprecedented multimessenger observing campaign (Abbott et al. 2017c). Detections and nondetections across the electromagnetic (EM) spectrum and by neutrino observatories have already produced new insights and will continue to do so for some time.

Many of these insights critically depend on the significance of the association between the independent observations. Often, such significance is established by estimating a p-value, the probability of such an event or a more extreme event occurring under the null hypothesis that the observations originate from unrelated distinct sources. Specific applications include, e.g., Abbott et al. (2017a), Coulter et al. (2017), and Soares-Santos et al. (2017) for GW170817 and its counterparts, Baret et al. (2012), Aartsen et al. (2014), and Keivani et al. (2015) for offline triggered search methods, and Urban (2016) for online rapid identification. A small p-value demonstrates that the data is inconsistent with the null hypothesis. The p-value cannot, however, be interpreted as the probability of the null hypothesis itself (Gelman et al. 2013). On the other hand, a large p-value does not necessarily imply that the null hypothesis has to be accepted, only that it cannot be rejected (Gregory 2005).

We introduce a different, generic model comparison method to determine whether two events in separate data sets are produced by a common source or by unrelated phenomena. This Bayesian measure of significance asks fundamentally different questions compared to the Frequentist p-value approach: it quantifies a *degree of belief* or *confidence* when comparing two hypotheses, given a particular nonrepeatable observation, while the p-value determines the consistency of the null hypothesis with the data and the error rate of determining significance, which is important for initial identification. (See Finn 1998 for a related discussion in the context of detection itself). The method is a direct comparison of the probabilities of alternative models and does not require empirical estimates of a background distribution for the interpretation of its result (although this may be necessary if the assumptions about the background are not trusted). Moreover, the framework requires explicit statements of the necessary assumptions; in particular, prior distributions on the relevant parameters and conditions for which significance can be factorized for different common model parameters (discussed later in Section 2.3). Finally, such a method naturally accounts for source uncertainties by marginalization (Loredo 2005).

Similar Bayesian approaches to this problem already exist in the literature: for example, Budavári & Szalay (2008) establish a method for cross-identification of point sources between catalogs, Soiaporn et al. (2012) apply hierarchical Bayesian clustering, and Naylor et al. (2013) develop a Bayesian method for cross-identification when there is a high rate of background events. For a review, see Budavári & Loredo (2015). The method presented here builds on these works and investigates many of the fundamental assumptions made by any such method. However, the approach is distinct from that of Kelley et al. (2013) in which the EM data is used as prior information to understand improvements in sensitivity for triggered searches.

In Section 2, we introduce the method in a general context; Equation (16) is our primary result and describes how to calculate the Bayes factor for a common-source origin of two signals seen in separate data streams. In Section 3, we focus on the application of the method to multimessenger astronomy, considering a calculation of spatial and temporal significance. As an example, we apply it to the gravitational-wave and gamma-ray events GW170817 and GRB 170817A, showing that it strongly supports the hypothesis that they originate from a common source.

2. Generic Derivation

2.1. Model Comparisons

Given two detections a and b in different data sets D_a and D_b, we would like to assess the hypothesis that they originate

from a common source. In general, the two detections will be described by different physical signal models $\mathcal{H}_a^{\rm S}$ and $\mathcal{H}_b^{\rm S}$, respectively. Each signal model will imply a likelihood, a set of parameters and an associated prior for those parameters. To quantify whether they originate from a common source, the models must share a common set of parameters $\theta \in \Theta$.

We will use notation where $\mathcal{H}(\theta) \equiv [\mathcal{H}$ and $\theta]$ denotes a hypothesis $\mathcal{H}$ with a particular choice for the parameters θ, while $\mathcal{H}$ by itself denotes a hypothesis with unknown parameters, i.e., "for any choice of parameters θ." We can formally write this as $\mathcal{H} \equiv [\mathcal{H}(\theta)$ for any $\theta]$.

Then, we define the common-source hypothesis:

$$\mathcal{H}^{\rm C} \equiv \{[\mathcal{H}_a^{\rm S}(\theta) \text{ and } \mathcal{H}_b^{\rm S}(\theta)] \text{ for any } \theta\}. \quad (1)$$

We also define $\mathcal{H}_{a/b}^{\rm N}$ as the noise hypotheses (by which we mean any event not originating from an astrophysical source) for each data set. Then we can define any alternative hypothesis for which the observed detections in a and b are unrelated:

$$\mathcal{H}^{\rm XY} \equiv \{[\mathcal{H}_a^{X}(\theta_a) \text{ for any } \theta_a] \text{ and } [\mathcal{H}_b^{Y}(\theta_b) \text{ for any } \theta_b]\}, \quad (2)$$

where $X, Y \in \{\rm N, S\}$. We write this in a general form, but note that the noise hypothesis will not have any common model parameters. In total, there are four possible realizations of $\mathcal{H}^{\rm XY}$, which we consider in detail below. However, $\mathcal{H}^{\rm SS}$ is of particular interest in this work, being two unrelated signals from distinct sources.

These hypotheses imply priors on θ that, in general, differ from those implied by $\mathcal{H}_{a/b}^{\rm S}$ individually: if a common source can only be detected in some subset of θ, then $\mathcal{H}^{\rm C}$ can only have prior support restricted to this subset. If this is not true, we identify the special case

$$P(\theta|\mathcal{H}^{\rm C}) = P(\theta|\mathcal{H}_a^{\rm S}) = P(\theta|\mathcal{H}_b^{\rm S}). \quad (3)$$

The probability of the common-source hypothesis is given by

$$P(\mathcal{H}^{\rm C}|D_a, D_b) = \frac{P(D_a, D_b|\mathcal{H}^{\rm C})P(\mathcal{H}^{\rm C})}{P(D_a, D_b)}. \quad (4)$$

In this work, we will calculate the *odds* between $\mathcal{H}^{\rm C}$ and different choices of $\mathcal{H}^{\rm XY}$

$$\begin{aligned}\mathcal{O}_{\rm C/XY}(D_a, D_b) &\equiv \frac{P(\mathcal{H}^{\rm C}|D_a, D_b)}{P(\mathcal{H}^{\rm XY}|D_a, D_b)} \\ &= \mathcal{B}_{\rm C/XY}(D_a, D_b)\frac{P(\mathcal{H}^{\rm C})}{P(\mathcal{H}^{\rm XY})},\end{aligned} \quad (5)$$

where

$$\mathcal{B}_{\rm C/XY}(D_a, D_b) \equiv \frac{P(D_a, D_b|\mathcal{H}^{\rm C})}{P(D_a, D_b|\mathcal{H}^{\rm XY})} \quad (6)$$

is the *Bayes factor* and $P(\mathcal{H}^{\rm C})/P(\mathcal{H}^{\rm XY})$ is the *prior odds*. In Section 2.2, we discuss the calculation of the Bayes factor in general. The prior odds will depend on the context, but in Section 3.1.3 we calculate the prior odds modeling $\mathcal{H}^{\rm C}$ and $\mathcal{H}^{\rm SS}$ as realizations of a Poisson point process.

2.2. Derivation of the Bayes Factor

If both data sets contain a signal from the same event, then they are not independent: $P(D_a, D_b|\mathcal{H}^{\rm C}) \neq P(D_a|\mathcal{H}^{\rm C})P(D_b|\mathcal{H}^{\rm C})$. Instead, we must compute

$$\begin{aligned}P(D_a, D_b|\mathcal{H}^{\rm C}) &= \int_\Theta P(D_a, D_b, \theta|\mathcal{H}^{\rm C})d\theta \\ &= \int_{\Theta^{\rm S}} P(D_a, D_b|\theta, \mathcal{H}^{\rm C})P(\theta|\mathcal{H}^{\rm C})d\theta,\end{aligned} \quad (7)$$

where the domain of the integral in the second line is restricted to the prior support of $\mathcal{H}^{\rm C}$, namely

$$\Theta^{\rm S} \equiv \{\theta \in \Theta \text{ where } P(\theta|\mathcal{H}^{\rm C}) > 0\}. \quad (8)$$

The need for this restriction arises because assuming that $(\theta = \theta')$ and $\mathcal{H}^{\rm C}$ are both true would be a contradiction if $P(\theta'|\mathcal{H}^{\rm C}) = 0$, and so $P(D|\theta', \mathcal{H}^{\rm C})$ would be undefined. Rearranging the likelihood in the integrand

$$\begin{aligned}&P(D_a, D_b|\theta, \mathcal{H}^{\rm C}) \\ &= P(D_a|D_b, \theta, \mathcal{H}^{\rm C})P(D_b|\theta, \mathcal{H}^{\rm C}) \\ &= P(D_a|\theta, \mathcal{H}^{\rm C})P(D_b|\theta, \mathcal{H}^{\rm C}) \\ &= \frac{P(D_a|\mathcal{H}^{\rm C})P(\theta|D_a, \mathcal{H}^{\rm C})}{P(\theta|\mathcal{H}^{\rm C})}\frac{P(D_b|\mathcal{H}^{\rm C})P(\theta|D_b, \mathcal{H}^{\rm C})}{P(\theta|\mathcal{H}^{\rm C})},\end{aligned} \quad (9)$$

where, in the second step, we have assumed that the likelihoods conditional on θ can be separated for the two data sets, provided that θ is the set of all model parameters common between the two likelihoods. In the last step, we again used $P(\theta|\mathcal{H}^{\rm C}) > 0$ within the integration interval. A subtle point is that $P(\theta|D_{a/b}, \mathcal{H}^{\rm C})$ is the posterior distribution for the common model parameters (given either $D_{a/b}$) marginalized over all other model parameters and using the prior implied by $\mathcal{H}^{\rm C}$.

Substituting Equation (9) into Equation (7)

$$P(D_a, D_b|\mathcal{H}^{\rm C}) = P(D_a|\mathcal{H}^{\rm C})P(D_b|\mathcal{H}^{\rm C})\mathcal{I}_\theta(D_a, D_b), \quad (10)$$

where the *posterior overlap integral*

$$\mathcal{I}_\theta(D_a, D_b) \equiv \int_{\Theta^{\rm S}} \frac{P(\theta|D_a, \mathcal{H}^{\rm C})P(\theta|D_b, \mathcal{H}^{\rm C})}{P(\theta|\mathcal{H}^{\rm C})}d\theta \quad (11)$$

quantifies the agreement between the posterior distributions of θ derived independently. In this integral, the prior has the effect of setting a scale against which the degree of overlap can be compared.

If the special case of Equation (3) holds, then the posterior overlap integral can be computed directly from the individual posteriors for each detection (i.e., $P(\theta|D_{a/b}, \mathcal{H}^{\rm C}) \to P(\theta|D_{a/b}, \mathcal{H}_{a/b}^{\rm S})$ in Equation (11)). In this case, the integrand is proportional to the posterior distribution for θ conditioned on both data sets (see Fan et al. 2014).

Equations (10)–(11) demonstrate how probabilities from separate data sets combine when each provides independent inferences about a common model parameter.

Returning to the Bayes factor, by our definition of the alternative hypothesis

$$P(D_a, D_b|\mathcal{H}^{\rm XY}) = P(D_a|\mathcal{H}_a^{X})P(D_b|\mathcal{H}_b^{Y}). \quad (12)$$

So from Equation (6),

$$\mathcal{B}_{\rm C/XY}(D_a, D_b) = \frac{P(D_a|\mathcal{H}^{\rm C})P(D_b|\mathcal{H}^{\rm C})}{P(D_a|\mathcal{H}_a^{X})P(D_b|\mathcal{H}_b^{Y})}\mathcal{I}_\theta(D_a, D_b). \quad (13)$$

We now specify three particular cases of interest for the alternative hypothesis. First, consider $\mathcal{H}^{\rm NN}$: both a and b are

caused by noise. Then Equation (13) specializes to

$$\mathcal{B}_{\mathrm{C/NN}}(D_a, D_b) = \mathcal{B}_{\mathrm{C/N}}(D_a)\mathcal{B}_{\mathrm{C/N}}(D_b)\mathcal{I}_\theta(D_a, D_b). \quad (14)$$

where $\mathcal{B}_{\mathrm{C/N}}$, in analogy with Equation (6), is the Bayes factor for common-source versus noise hypotheses. In the special case of Equation (3), it can be shown that $\mathcal{B}_{\mathrm{C/N}}(D_{a/b}) = \mathcal{B}_{\mathrm{S/N}}(D_{a/b})$, i.e., the Bayes factor for an independent signal versus noise for each data stream.

This agrees with our intuition: if both signals are strong compared to the background noise *and* there is a good overlap of their common model parameters (quantified by $\mathcal{I}_\theta$), we believe they originate from a common event. This is a powerful result, as one can compute the joint Bayes factor from the individual Bayes factors, and the posterior overlap integral of θ. Equation (14) has analogous applications to the Fisher combined probability test used in Aartsen et al. (2014).

Second, consider $\mathcal{H}^{\mathrm{SN}}$: a was due to a signal, but b was due to noise. For this case, Equation (13) gives

$$\mathcal{B}_{\mathrm{C/SN}}(D_a, D_b) = \mathcal{B}_{\mathrm{C/N}}(D_b)\mathcal{I}_\theta(D_a, D_b). \quad (15)$$

For us to believe that detection b is a real signal *and* originates from the same source as a, we require the product of the Bayes factor for the common-source against noise in b and the posterior overlap to be large. The case $\mathcal{B}_{\mathrm{C/NS}}$ is analogous and the same special cases apply as mentioned previously.

Finally, consider $\mathcal{H}^{\mathrm{SS}}$, the distinct-source hypothesis: both a and b are of the same nature as in the common-source hypothesis $\mathcal{H}^{\mathrm{C}}$, but they are physically distinct (i.e., they belong to unrelated sources with different parameters $\theta_a \neq \theta_b$). Then,

$$\mathcal{B}_{\mathrm{C/SS}}(D_a, D_b) \equiv \frac{P(D_a|\mathcal{H}^{\mathrm{C}})P(D_b|\mathcal{H}^{\mathrm{C}})}{P(D_a|\mathcal{H}^{\mathrm{S}}_a)P(D_b|\mathcal{H}^{\mathrm{S}}_b)}\mathcal{I}_\theta(D_a, D_b). \quad (16)$$

This equation and the posterior overlap integral of Equation (11) are the main results of this paper. This provides a simple and intuitive way to assess whether two detections originate from the same event, based on the posterior overlap of their common model parameters.

In the special case of Equation (3), the prefactor to the posterior overlap integral is unity, such that

$$\mathcal{B}_{\mathrm{C/SS}}(D_a, D_b) = \mathcal{I}_\theta(D_a, D_b). \quad (17)$$

On the other hand, when Equation (3) does not apply, the prefactor plays an important role in quantifying how the restricted prior implied by $\mathcal{H}^{\mathrm{C}}$ affects the Bayes factor.

A similar result to Equation (17) was obtained independently by K. Haris et al. (2017, private communication) in the context of strongly lensed gravitational wave (GW) signals from binary black hole mergers. This Bayes factor can also be formulated as in Budavári & Szalay (2008); however, the ensuing normality assumptions are not generally applicable.

In this derivation, we have not explicitly discussed the selection effects due to considering events that are triggered on properties of one or another data set. Such selection procedures result in a normalization of the likelihood over all data sets that would have passed the threshold to be analyzed in a given detector (Loredo 2005). The parameter posteriors $P(\theta|D_a)$ and $P(\theta|D_b)$ are, however, unaffected by this consideration.

2.3. Factorization of the Posterior Overlap Integral

When calculating the Bayes factor, it is often convenient to factorize the posterior overlap integral, e.g., $\mathcal{I}_\theta = \mathcal{I}_\phi\mathcal{I}_\psi$, where $\phi \subsetneq \theta$ and $\psi = \theta \backslash \phi$. The total Bayes factor could then be expressed as the product of individual Bayes factors for each set of parameters. This factorization can only be performed, however, if $P(\phi|\psi, D_{A/B}, \mathcal{H}^{\mathrm{S}}) = P(\phi|D_{A/B}, \mathcal{H}^{\mathrm{S}})$. There are situations in which this is the case, for example, if the joint posterior distribution is an uncorrelated multivariate normal distribution (see, e.g., Budavári & Szalay 2008). But, generally, this will not be the case and the posterior over the full common parameter space must be used. There are, however, cases where, under certain assumptions, the integral can be approximately factorized. We will explore one such setting in Section 3.

3. Application to Multimessenger Transient Astronomy

We now focus on the application of the above formalism to multimessenger transient astronomy. To guide our intuition, we consider a transient GW candidate and a detection made by an EM instrument, although, we could just as well consider any pair of EM, GW, or neutrino detectors. Assuming that detections are made in both the GW and EM detectors and are independently significant, we aim to calculate $\mathcal{O}_{\mathrm{C/SS}}(D_{\mathrm{GW}}, D_{\mathrm{EM}})$, the odds quantifying the probability of the common-source hypothesis to a distinct-source hypothesis.

The Bayes factor should be calculated from all common-source parameters; typically, this will involve parameters such as a characteristic time of the event, source direction, luminosity distance, and source orientation (Margutti et al. 2017; Troja et al. 2017). Ideally, the posterior overlap integral should be computed over the complete joint distribution of parameters since it will not generally factorize (see Section 2.3).

However, to illustrate the utility of the method, we will calculate the result considering only the spatial and temporal common parameters (specifically, the source direction $\mathbf{\Omega}$ and coalescence time of the BNS system t_{c}) and make assumptions under which the posterior overlap integral may be factorized. We also consider both observatories to be all-sky, neglecting nonisotropic and nonstationary sensitivity. The Bayes factor can then be calculated from Equation (17) since the special case of Equation (3) applies.

3.1. Example

To calculate the Bayes factor, Equation (17), we first write down the posterior overlap integral over the conditional joint distribution of the spatial and temporal parameters

$$\begin{aligned}\mathcal{I}_{\mathbf{\Omega},t_{\mathrm{c}}}(D_{\mathrm{GW}}, D_{\mathrm{EM}}) &= \iint \frac{P(\mathbf{\Omega}, t_{\mathrm{c}}|D_{\mathrm{GW}}, \mathcal{H}^{\mathrm{S}}_{\mathrm{GW}})P(\mathbf{\Omega}, t_{\mathrm{c}}|D_{\mathrm{EM}}, \mathcal{H}^{\mathrm{S}}_{\mathrm{EM}})}{P(\mathbf{\Omega}, t_{\mathrm{c}}|\mathcal{H}^{\mathrm{S}})} d\mathbf{\Omega}dt_{\mathrm{c}} \\ &= \iint \frac{P(\mathbf{\Omega}|t_{\mathrm{c}}, D_{\mathrm{GW}}, \mathcal{H}^{\mathrm{S}}_{\mathrm{GW}})P(\mathbf{\Omega}|t_{\mathrm{c}}, D_{\mathrm{EM}}, \mathcal{H}^{\mathrm{S}}_{\mathrm{EM}})}{P(\mathbf{\Omega}, t_{\mathrm{c}}|\mathcal{H}^{\mathrm{S}})} \\ &\quad \times P(t_{\mathrm{c}}|D_{\mathrm{GW}}, \mathcal{H}^{\mathrm{S}}_{\mathrm{GW}})P(t_{\mathrm{c}}|D_{\mathrm{EM}}, \mathcal{H}^{\mathrm{S}}_{\mathrm{EM}})d\mathbf{\Omega}dt_{\mathrm{c}}.\end{aligned} \quad (18)$$

We will now show that this can be factorized into a spatial and temporal overlap under the following assumptions. First, that the prior itself factors, $P(\mathbf{\Omega}, t_{\mathrm{c}}|\mathcal{H}^{\mathrm{S}}) = P(\mathbf{\Omega}|\mathcal{H}^{\mathrm{S}})P(t_{\mathrm{c}}|\mathcal{H}^{\mathrm{S}})$.

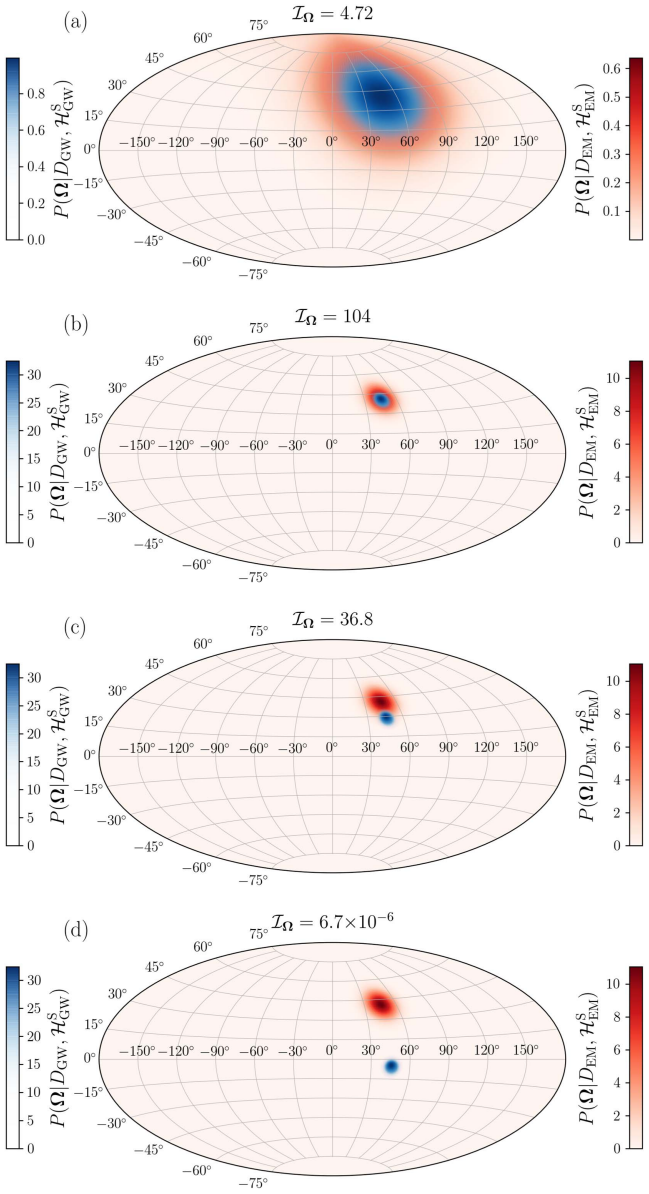

Figure 1. Examples of the spatial overlap $\mathcal{I}_\Omega$ (numerical values for each figure are given in the individual figure title). A blue (red) density map shows the probability per pixel for the GW (EM) detection. $\mathcal{I}_\Omega$ is calculated by numerical integration over these pixels with an all-sky uniform prior. The posteriors are computed over an array of pixels, each with equal area, using the HEALPix projection (Gorski et al. 2005).

Second, that $t_{\rm c}$ inferred from the GW data is exactly determined, i.e.,

$$P(t_{\rm c}|D_{\rm GW}, \mathcal{H}^{\rm S}_{\rm GW}) = \delta(t_{\rm c} - \widehat{t_{\rm c}}), \tag{19}$$

where the "hat" indicates the observed value. Then, Equation (18) can be factorized as $\mathcal{I}_{\mathbf{\Omega},t_{\rm c}} = \mathcal{I}_{t_{\rm c}}\mathcal{I}_{\mathbf{\Omega}}$, where

$$\mathcal{I}_{t_{\rm c}} = \frac{P(t_{\rm c} = \widehat{t_{\rm c}}|D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM})}{P(t_{\rm c} = \widehat{t_{\rm c}}|\mathcal{H}^{\rm S})} \tag{20}$$

and

$$\mathcal{I}_{\mathbf{\Omega}} = \int \frac{P(\mathbf{\Omega}|\widehat{t_{\rm c}}, D_{\rm GW}, \mathcal{H}^{\rm S}_{\rm GW})P(\mathbf{\Omega}|\widehat{t_{\rm c}}, D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM})}{P(\mathbf{\Omega}|\mathcal{H}^{\rm S})}\, d\mathbf{\Omega}. \tag{21}$$

This factorization is exact under the two assumptions made. However, the coalescence time is typically known with a nonzero uncertainty. For this case, taking $\widehat{t_{\rm c}}$ to be a point estimate (the mean, for example), the factorization is approximate, but applicable provided that over the uncertainty in $t_{\rm c}$, $P(t_{\rm c}|D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM})$, $P(\mathbf{\Omega}|t_{\rm c}, D_{\rm GW}, \mathcal{H}^{\rm S}_{\rm GW})$, and $P(\mathbf{\Omega}|t_{\rm c}, D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM})$ do not vary substantially. In Sections 3.1.1 and 3.1.2, we will provide approximations for Equations (20) and (21) under some reasonable assumptions and illustrate some of the subtleties in their calculation.

We note that a similar result to Equation (21) was previously derived in Urban (2016); in particular, Equation (3.6) of that work is equivalent to Equation (21) assuming an isotropic prior. Then the resulting joint likelihood ratio is defined using the alternative hypothesis that went into Equation (15).

3.1.1. Temporal Overlap

To evaluate Equation (20), the temporal overlap, we first need to consider how to compute $P(t_{\rm c}|D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM})$, the coalescence time given the EM observations. Typically, EM observations do not directly infer $t_{\rm c}$, but some other well defined time $t_{\rm EM}$, e.g., the time of peak flux. We therefore need to specify a model that relates these two times. One simple model is that both signals travel at the speed of light, but there is a delay $\Delta t = t_{\rm EM} - t_{\rm c}$ between the coalescence time and the EM emission that will depend on the physics (see, e.g., Finn et al. 1999; Abadie et al. 2012 for GRB delay time predictions), but also on how $t_{\rm EM}$ is defined. To fold these predictions into the analysis, we must specify $P(\Delta t|\mathcal{H}^{\rm S})$, a prior distribution on the delay time (at the Earth), given the model. Assuming Δt and $t_{\rm EM}$ are independent, the posterior can be transformed as

$$P(t_{\rm c}|D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM}) = \int p_{t_{\rm EM}}(t_{\rm c} + \Delta t)P(\Delta t)d\Delta t, \tag{22}$$

where $p_{t_{\rm EM}}(t_{\rm EM}) \equiv P(t_{\rm EM}|D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM})$ denotes the posterior distribution of $t_{\rm EM}$.

Having defined how to relate the time inferred by the EM data to the coalescence time with a suitable model, we now calculate Equation (20) under some simple assumptions. Equation (19) was the first of these assumptions and was already applied in factorizing the full posterior overlap integral. In addition, let

$$P(t_{\rm EM}|D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM}) \equiv p_{t_{\rm EM}}(t_{\rm EM}) = \delta(t_{\rm EM} - \widehat{t_{\rm EM}}). \tag{23}$$

Next we need a prior for the delay in the GW-EM arrival time, which could be due to differences in emission time or propagation speed of GW and EM radiation. For simplicity, we take a uniform distribution,

$$P(\Delta t) = \mathrm{U}^{\Delta t^{\rm max}}_{\Delta t^{\rm min}}(\Delta t). \tag{24}$$

That is, the EM emission can arrive any time between a minimum and maximum value compared to the GW-inferred coalescence time; outside of that interval, we are certain the two events are not related. Inserting these definitions into Equation (22), we obtain

$$P(t_{\rm c}|D_{\rm EM}, \mathcal{H}^{\rm S}_{\rm EM}) = \mathrm{U}^{\Delta t^{\rm max}}_{\Delta t^{\rm min}}(\widehat{t_{\rm EM}} - t_{\rm c}), \tag{25}$$

from which the numerator of Equation (20) can be calculated.

The prior on the coalescence time, given $\mathcal{H}^{\rm S}$ and a co-observing time of duration T, is $P(t_{\rm c}|\mathcal{H}^{\rm S}) = \mathrm{U}^{T}_{0}(t_{\rm c})$, where $t_{\rm c}$ is chosen to be zero at the start of the co-observing time. The

period T should cover the entire range of t_c for which the t_c posteriors (in this example, Equations (19) and (25)) have nonzero support from the data, but is otherwise an arbitrary normalization of the time prior. Then, Equation (20) gives

$$\mathcal{I}_{t_c} = \begin{cases} \frac{T}{[\Delta t]} & \text{if } (\widehat{t_c} - \widehat{t_{\rm EM}}) \in [\Delta t^{\rm min}, \Delta t^{\rm max}] \\ 0 & \text{otherwise} \end{cases}, \tag{26}$$

where $[\Delta t] \equiv \Delta t^{\rm max} - \Delta t^{\rm min}$. The dependence on T would suggest that we can arbitrarily change the significance through the Bayes factor by adjusting T. However, as will be shown in Section 3.1.3, this factor cancels with the prior odds $P(\mathcal{H}^{\rm C})/P(\mathcal{H}^{\rm SS})$, which depends on both T and the rate of events, such that the odds themselves are T-independent. This cancellation relies on the assumption that T is short compared to the average time between events, or equivalently that a single detection exists in each of the data sets. We also will see that in a particular class of cases the temporal odds can be well approximated by this Bayes factor, setting T to the average interval between signals detectable in EM, (i.e., the inverse of the rate of such signals).

3.1.2. Spatial Overlap

We now discuss calculating Equation (21), the spatial posterior overlap integral. The EM counterparts to GW events are expected to originate from the same source direction and hence $\mathcal{I}_\Omega$ can be directly computed from Equation (21).

To illustrate the subtleties of $\mathcal{I}_\Omega$ and provide some intuition, in Figure 1, we show four examples varying the size of the uncertainty region and angular separation of the means of the EM and GW sky localizations. For all examples, a uniform all-sky prior is used. In Figure 1(a), the means of both posteriors are aligned, but the uncertainty on both is large with respect to the all-sky prior; therefore, $\mathcal{I}_\Omega$ (given in the individual figure title) is greater than one, but not large enough to be of note. For Figure 1(b), $\mathcal{I}_\Omega$ strongly indicates the two detections are from the same event: the means are aligned and the uncertainties are small with respect to the all-sky prior. In Figures 1(c) and (d), the means of the distributions are not aligned. While in (c) this results in modest evidence in favor of a common event, the separation is sufficiently wide in (d) to strongly disfavor a common source.

To help guide our intuition, we can also calculate $\mathcal{I}_\Omega$ for the simplified case where the posterior distributions on the sky are uniform distributions, i.e., constant inside the sets $\Pi_{\rm GW}$ and $\Pi_{\rm EM}$ and zero outside. Labeling $\Delta(\Pi)$ the area of set Π in square radians, we obtain

$$\mathcal{I}_\Omega(D_{\rm GW}, D_{\rm EM}) = 4\pi \frac{\Delta(\Pi_{\rm GW} \cap \Pi_{\rm EM})}{\Delta(\Pi_{\rm GW})\Delta(\Pi_{\rm EM})}. \tag{27}$$

The 4π prefactor comes from the all-sky prior and acts as a metric to compare the size of the overlap. For example, if $\Pi_{\rm GW}$ is entirely contained within $\Pi_{\rm EM}$, then $\mathcal{I}_\Omega(D_{\rm GW}, D_{\rm EM}) = 4\pi/\Delta(\Pi_{\rm EM})$: the Bayes factor is determined entirely by the fraction of the sky covered by the uncertainty on the EM detections (or vice versa if the EM posterior is contained within the GW posterior).

3.1.3. The Spatial and Temporal Odds

To calculate the odds in this example via Equation (5), we require the prior odds. We consider a Poisson point process that produces events detectable via either (or both) their GW or EM emission with a total rate R per unit time, acting during the co-observing time T. Furthermore, we let $R = R_{\rm GW} + R_{\rm EM} + R_{\rm GW,EM}$: the total rate is the sum of the rates of events detectable only in GW, events detectable only in EM, and events jointly detectable in both. Then $\mathcal{H}^{\rm C}$ refers to a signal seen by both detectors, and $\mathcal{H}^{\rm SS}$ to signals detected in one or another, but not both. Choosing T such that RT $\ll 1$, and defining Poisson$(1;\,\lambda)$ to be the probability of one event given an expected number of signals λ, we obtain

$$\begin{aligned} \frac{P(\mathcal{H}^{\rm C})}{P(\mathcal{H}^{\rm SS})} &= \frac{\text{Poisson}(1;\, R_{\rm GW,EM}T)}{\text{Poisson}(1;\, R_{\rm GW}T)\text{Poisson}(1;\, R_{\rm EM}T)} \\ &\approx \frac{R_{\rm GW,EM}}{R_{\rm GW}R_{\rm EM}T}. \end{aligned} \tag{28}$$

This prior odds clearly depends on the co-observing time T. Combining this with the spatial and temporal Bayes factor (Equations (21) and (26)) then gives

$$\mathcal{O}_{\rm C/SS}(D_a, D_b) = \frac{R_{\rm GW,EM}}{R_{\rm GW}R_{\rm EM}} \frac{1}{[\Delta t]} \mathcal{I}_\Omega, \tag{29}$$

which is not dependent on the co-observing time (an analogous result was found by Budavári 2011). One special case is when $R_{\rm GW} \simeq R_{\rm GW,EM} \ll R_{\rm EM}$, i.e., if signals detectable in EM only are much more frequent than in GW, but we otherwise have little information on the rates of GW detections with or without EM counterparts. This may typically occur if our estimates of $R_{\rm GW}$ and $R_{\rm GW,EM}$ are based on $\mathcal{O}(1)$ detection. The odds are then proportional to $1/(R_{\rm EM}[\Delta t])$, which reproduces the temporal Bayes factor Equation (26) setting $T = 1/R_{\rm EM}$, i.e., the waiting time between EM detections (where the great majority have no GW counterpart).

As can be expected intuitively, the association becomes less significant if the Δt prior is broader or the prior background rate of signals is higher, but increases with the prior expected rate of joint detections.

A more detailed treatment of the prior odds would include a model of the selection process by which the GW and EM triggers are produced. We do not expect our numerical results to be strongly affected by such modeling; however, it could be useful to generalize the method to other types of observations.

3.2. Application to GW170817 and GRB 170817A

We now apply the example calculated in Section 3.1 to GW170817 and GRB 170817A, the result of which can be compared with Abbott et al. (2017a). We note that, the calculation presented here could be improved by using the full joint distribution without making assumptions that allow the result to be factorized, and including other pertinent model parameters such as the luminosity distance (for *Fermi*-GBM, this may be as simple as estimating the range of conceivable values).

The sky localization for the BNS inspiral and short GRB can be seen in Figure 1 of Abbott et al. (2017a). Using the published localization `FITS` files (Goldstein et al. 2017; Singer 2017) and a uniform prior distribution on the whole sky,

Equation (21) yields $\mathcal{I}_\Omega = 32.4$. The spatial overlap alone provides moderate support for the common-event model, the main limitation being the uncertainty on the localization of GRB 170817A.

If we did not have the actual FITS files, we could still use the published confidence intervals for the sky localization of GW170817 and GRB 170817A, take the localization posterior distributions to be uniform within those intervals, and apply Equation (27). The 90% intervals cover, respectively, 28 deg^2 (Abbott et al. 2017c) and 1100 deg^2 (Goldstein et al. 2017) and the GW170817 interval is entirely contained within GRB 170817A. Applying Equation (27) then yields an approximate spatial Bayes factor of $\mathcal{I}_\Omega = 37.5$, which is close to the exact value. However, since in this case we do have the full posteriors, we can repeat this calculation with different confidence levels. We find that $\mathcal{I}_\Omega$ can be in error by a factor of a few in both directions, depending on what confidence level is used; the 90% interval just happens to produce a particularly close number. We therefore do not recommend a naive application of Equation (27), but instead the full numerical integration of the posteriors (i.e., in this case, the $\mathcal{I}_\Omega = 32.4$ calculated previously).

In calculating the odds from Equation (29), there are large uncertainties on the three rates. However, the rate of short GRB detections by *Fermi*-GBM is well known and must satisfy $R_{Fermi} \approx R_{\rm EM} + R_{\rm GW,EM}$. One could model the uncertainties to produce beaming and volume corrected estimates for these rates (see, e.g., Fong et al. 2015; Siellez et al. 2016). However, for a simple estimate, we assume that $R_{\rm GW,EM}$ and $R_{\rm GW}$ are of similar magnitude, and take $R_{\rm EM}$ to be $R_{Fermi} = 0.124$ per day (Abbott et al. 2017a). We will also assume $[\Delta t^{\rm min}, \Delta t^{\rm max}] = [-1, 5]$ s, the range used in Abadie et al. (2012). Under the assumptions described above, Equation (29) yields

$$\mathcal{O}_{\rm C/SS}(D_{\rm GW}, D_{\rm EM}) = \frac{1}{R_{Fermi}}\frac{1}{[\Delta t]}\mathcal{I}_\Omega \gtrsim 10^6, \qquad (30)$$

the odds provide decisive evidence that the two detections originate from the same event.

These numbers are consistent with the *p*-values estimated in Abbott et al. (2017a): the time overlap dominates, the spatial part is small but supports the hypothesis, and the overall factor is highly significant (the total *p*-value was found to be 5×10^{-8}; Abbott et al. 2017a).

3.3. Comparison with p-values

There are parallels that can be drawn between the odds calculated in Section 3.1 and the *p*-value approach of Abbott et al. (2017a); namely, the spatial overlap Equation (21) with the $\mathcal{S}$ statistic and the form of the temporal overlap (i.e., inversely proportional to the background rate).

However, the two methods are not equivalent and the numerical values themselves cannot be directly compared, as they answer different questions. The odds are exactly our relative degrees of belief for the common—versus distinct—source hypotheses, given the assumptions made in the calculation; the *p*-value tests whether the data is consistent or not with the null (distinct-source) hypothesis, and is typically interpreted as the rate at which a rule for deciding the significance of a joint detection leads to false positives (Finn et al. 1999).

Finally, a more practical difference is that while Equation (21) can be directly interpreted as a Bayes factor for the spatial overlap, interpreting the $\mathcal{S}$ statistic requires numerical calculation of the background by randomly rotating a set of observed short GRB sky localizations.

4. Conclusions

We introduce a Bayesian model comparison approach to estimating our confidence that two multimessenger observations are due to a common source as opposed to an accidental coincidence of distinct sources. The primary result of this work is Equation (16), which generically allows the calculation of the Bayes factor (and, hence, the odds) from the joint posterior distributions of common model parameters inferred independently from two data sets. This approach forces us to recognize the conditions under which the contributions to the Bayes factor can be factorized.

We provide an example where the spatial and temporal overlap calculation can be approximately factorized for two independent observations with isotropic observatories and apply the result to GW170817 and GRB 170817A. We find decisive evidence in favor of their association, which is consistent with Abbott et al. (2017a).

The authors are grateful to Michael Briggs, Collin Capano, Sebastian Khan, Badri Krishnan, Francesco Pannarale, Yafet Sanchez Sanchez, Karelle Siellez, Grant Meadors, members of the LIGO and Virgo collaborations, and the referee for useful comments during the preparation of this work. E.B. and T.D.C. are supported by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by Universities Space Research Association under contract with NASA.

ORCID iDs

G. Ashton https://orcid.org/0000-0001-7288-2231
T. Dal Canton https://orcid.org/0000-0001-5078-9044
T. Dent https://orcid.org/0000-0003-1354-7809
H.-B. Eggenstein https://orcid.org/0000-0001-5296-7035

References

Aartsen, M. G., Ackermann, M., Adams, J., et al. 2014, PhRvD, 90, 102002
Abadie, J., Abbott, B. P., Abbott, R., et al. 2012, ApJ, 760, 12
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, ApJL, 848, L13
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, PhRvL, 119, 161101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, ApJL, 848, L12
Baret, B., Bartos, I., Bouhou, B., et al. 2012, PhRvD, 85, 103004
Budavári, T. 2011, ApJ, 736, 155
Budavári, T., & Loredo, T. J. 2015, AnRSA, 2, 113
Budavári, T., & Szalay, A. S. 2008, ApJ, 679, 301
Coulter, D. A., Foley, R. J., Kilpatrick, C. D., et al. 2017, Sci, 358, 1556
Fan, X., Messenger, C., & Heng, I. S. 2014, ApJ, 795, 43
Finn, L. S. 1998, in Second Edoardo Amaldi Conference on Gravitational Wave Experiments, ed. E. Coccia, G. Veneziano, & G. Pizzella (Singapore: World Scientific), 180
Finn, L. S., Mohanty, S. D., & Romano, J. D. 1999, PhRvD, 60, 121101
Fong, W., Berger, E., Margutti, R., & Zauderer, B. A. 2015, ApJ, 815, 102
Gelman, A., Carlin, J. B., Stern, H. S., et al. 2013, Bayesian Data Analysis (Boca Raton, FL: CRC Press)
Goldstein, A., Veres, P., Burns, E., et al. 2017, ApJL, 848, L14
Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Gregory, P. 2005, Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support (Cambridge: Cambridge Univ. Press)

Keivani, A., Fox, D. B., Tešić, G., Cowen, D. F., & Fixelle, J. 2015, arXiv:1508.01315
Kelley, L. Z., Mandel, I., & Ramirez-Ruiz, E. 2013, PhRvD, 87, 123004
Loredo, T. J. 2005, in AIP Conf. Proc. 735, 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (Melville, NY: AIP), 195
Margutti, R., Berger, E., Fong, W., et al. 2017, ApJL, 848, L20
Naylor, T., Broos, P. S., & Feigelson, E. D. 2013, ApJS, 209, 30
Savchenko, V., Ferrigno, C., Kuulkers, E., et al. 2017, ApJL, 848, L15
Siellez, K., Boer, M., Gendre, B., & Regimbau, T. 2016, arXiv:1606.03043
Singer, L. 2017, GW170817 sky Localization, https://dcc.ligo.org/LIGO-G1701985/public
Soares-Santos, M., Holz, D. E., Annis, J., et al. 2017, ApJL, 848, L16
Soiaporn, K., Chernoff, D., Loredo, T., et al. 2012, in Statistical Challenges in Modern Astronomy V, ed. E. D. Feigelson & G. J. Babu (New York: Springer), 543
Troja, E., Piro, L., van Eerten, H., et al. 2017, Natur, 551, 7678
Urban, A. L. 2016, Ph.D. Dissertation, Univ. Wisconsin-Milwaukee, https://search.proquest.com/docview/1806791084