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[4, 1]

[1] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), D-30167 Hannover, Germany*
[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*

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 vs. 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.

1. INTRODUCTION

On August 17th 2017, the observation by LIGO-Virgo of GW170817, a binary neutron star coalescence (BNS) (Abbott et al. 2017b,a), 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 non-detections across the electromagnetic (EM) spectrum and by neutrino observatories have already produced new insights and will continue to do so for some time yet.

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); Soares-Santos et al. (2017) for GW170817 and its counterparts, Baret et al. (2012); Aartsen et al. (2014); Keivani et al. (2015) for offline triggered search methods, and Urban (2016) for online rapid identification. A small *p*-value demonstrates the data is inconsistent with the null hypothesis. The *p*-value cannot, though, be interpreted as the probability of the null hypothesis itself (Gelman et al. 2013).

Corresponding author: Gregory Ashton
gregory.ashton@ligo.org

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 non-repeatable 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 Sec. 2.3). This 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 Sec. 2 we introduce the method in a general context; Eq. (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 Sec. 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 $\mathcal{H}_a^{\mathrm{S}}$ and $\mathcal{H}_b^{\mathrm{S}}$, respectively. Each signal model will imply a likelihood, a set of parameters and an associated prior for those parameters. To assess whether they originate from a common source, the models must share a common set of parameters $\theta \in \Theta$.

We'll use notation where $\mathcal{H}(\theta) \equiv [\mathcal{H} \text{ and } \theta]$ 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 $\mathcal{H} \equiv [\mathcal{H}(\theta) \text{ for any } \theta]$.

Then, we define the common-source hypothesis:

$$\mathcal{H}^{\mathrm{C}} \equiv \left\{ \left[\mathcal{H}_a^{\mathrm{S}}(\theta) \text{ and } \mathcal{H}_b^{\mathrm{S}}(\theta)\right] \text{ for any } \theta \right\} . \qquad (1)$$

We also define $\mathcal{H}_{a/b}^{\mathrm{N}}$ as the noise hypotheses (by which we mean any non-signal) for each data set. Then we can define any alternative hypothesis for which the observed detections in a and b are unrelated:

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

where $X, Y \in \{\mathrm{N, S}\}$. 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 $\mathcal{H}^{XY}$, which we consider in detail below. However, $\mathcal{H}^{\mathrm{SS}}$ is of particular interest in this work, being two unrelated signals from distinct sources.

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

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

The probability of the common-source hypothesis is given by

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

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

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

where

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

is the *Bayes factor* and $P(\mathcal{H}^{\mathrm{C}})/P(\mathcal{H}^{XY})$ is the *prior odds*. In Sec. 2.2 we discuss the calculation of the Bayes factor in general. The prior odds will depend on the context, but in Sec. 3.1.3 we calculate the prior odds modeling $\mathcal{H}^{\mathrm{C}}$ and $\mathcal{H}^{\mathrm{SS}}$ 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: $P(D_a, D_b|\mathcal{H}^{\mathrm{C}}) \neq P(D_a|\mathcal{H}^{\mathrm{C}})P(D_b|\mathcal{H}^{\mathrm{C}})$. Instead we must compute

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

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

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

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

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

where in the second step we have used 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 that $P(\theta|\mathcal{H}^{\mathrm{C}}) > 0$ within the integration interval. A subtle point is that $P(\theta|D_{a/b}, \mathcal{H}^{\mathrm{C}})$ 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 $\mathcal{H}^{\mathrm{C}}$.

Substituting Eq. (9) into Eq. (7)

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

where the *posterior overlap integral*

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

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.

Eq. (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

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

So from Eq. (6),

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

We now specify three particular cases of interest for the alternative hypothesis. First, consider $\mathcal{H}^{\mathrm{NN}}$: both a and b are caused by noise. Then Eq. (13) specializes to

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

where $\mathcal{B}_{\mathrm{C}/\mathrm{N}}$, in analogy with Eq. (6), is the common-source against noise Bayes factor. In the special case of Eq. (3), it can be shown that $\mathcal{B}_{\mathrm{C}/\mathrm{N}}(D_{a/b}) = \mathcal{B}_{\mathrm{S}/\mathrm{N}}(D_{a/b})$, i.e. the independent signal against noise Bayes factor 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 the integral), we believe they originate from a common event. This is a powerful result as one can compute the joint Bayes factor from the common-source against noise Bayes factor for each detection individually, and the posterior overlap integral of θ. Eq. (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, Eq. (13) gives

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

For us to believe that detection b is a real signal *and* originates from the same source as a, we require that the product of the Bayes factor for common-source against noise in b and the posterior overlap be large. The case

$\mathcal{B}_{\mathrm{C}/\mathrm{NS}}$ 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,

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

This equation and the posterior overlap integral of Eq. (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 Eq. (3), the prefactor to the posterior overlap integral is unity, such that

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

On the other hand, when Eq. (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 Eq. (17) was obtained independently by Haris et al. (2017) in the context of strongly lensed gravitational wave signals from binary black hole mergers.

2.3. *Factorization of the posterior overlap integral*

When calculating $\mathcal{I}_\theta$, 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 $\psi = \theta \setminus \phi$. 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. 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 the next Section.

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 gravitational wave (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 $\mathcal{O}_{\mathrm{C}/\mathrm{SS}}(D_{\mathrm{GW}}, D_{\mathrm{EM}})$, 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 (Troja et al. 2017; Margutti 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 Sec. 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 $\boldsymbol{\Omega}$ 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 non-isotropic and non-stationary sensitivity. The Bayes factor can then be calculated from Eq. (17) since the special case of Eq. (3) applies.

3.1. Example

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

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

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, $P(\boldsymbol{\Omega}, t_\mathrm{c}|\mathcal{H}^\mathrm{S}) = P(\boldsymbol{\Omega}|\mathcal{H}^\mathrm{S})P(t_\mathrm{c}|\mathcal{H}^\mathrm{S})$. Second, that t_c inferred from the GW data is exactly determined, i.e.

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

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

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

and

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

This factorization is exact under the two assumptions made. However, the coalescence time is typically known with a nonzero uncertainty. For this case, taking $\widehat{t_\mathrm{c}}$ to be a point estimate (the mean for example),

the factorization is approximate, but applicable provided that over the uncertainty in t_c, $P(t_\mathrm{c}|D_\mathrm{EM}, \mathcal{H}_\mathrm{EM}^\mathrm{S})$, $P(\boldsymbol{\Omega}|t_\mathrm{c}, D_\mathrm{GW}, \mathcal{H}_\mathrm{GW}^\mathrm{S})$, and $P(\boldsymbol{\Omega}|t_\mathrm{c}, D_\mathrm{EM}, \mathcal{H}_\mathrm{EM}^\mathrm{S})$ do not vary substantially. In Sec. 3.1.1 and Sec. 3.1.2, we will provide approximations for Eq. (20) and (21) under some reasonable assumptions and illustrate some of the subtleties in their calculation.

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

3.1.1. Temporal overlap

To evaluate Eq. (20), the temporal overlap, we first need to consider how to compute $P(t_\mathrm{c}|D_\mathrm{EM}, \mathcal{H}_\mathrm{EM}^\mathrm{S})$, the coalescence time given the EM observations. Typically, EM observations do not directly infer t_c, but some other well defined time t_EM, e.g. the time of peak luminosity. 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 $\Delta t = t_\mathrm{EM} - t_\mathrm{c}$ between the coalescence time and the EM emission which 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_EM is defined. To fold these predictions into the analysis, we must specify $P(\Delta t|\mathcal{H}^\mathrm{S})$, a prior distribution on the delay-time (at the Earth), given the model. Assuming Δt and t_EM are independent, the posterior can be transformed as

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

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

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

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

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,

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

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 Eq. (22), we obtain

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

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

The prior on the coalescence time, given $\mathcal{H}^{\mathrm{S}}$ and a co-observing time of duration T, is $P(t_{\mathrm{c}}|\mathcal{H}^{\mathrm{S}}) = \mathrm{U}_0^T(t_{\mathrm{c}})$, where t_{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, Eq. (19) and (25)) have nonzero support from the data, but is otherwise an arbitrary normalization of the time prior. Then, Eq. (20) gives

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

where $[\Delta t] \equiv \Delta t^{\mathrm{max}} - \Delta t^{\mathrm{min}}$. 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 Sec. 3.1.3, this factor cancels with the prior odds $P(\mathcal{H}^{\mathrm{C}})/P(\mathcal{H}^{\mathrm{SS}})$, which depends on both T and the rate of events, such that the odds themselves are T-independent. 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 Eq. (21), the spatial posterior overlap integral. The EM counterparts to GW events are expected to originate from the same source direction and hence $\mathcal{I}_{\mathbf{\Omega}}$ can be directly computed from Eq. (21).

To illustrate the subtleties of $\mathcal{I}_{\mathbf{\Omega}}$ and provide some intuition, in Fig. 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 Fig. 1(a), the means of both posteriors are aligned, but the uncertainty on both is large with respect to the all-sky prior; therefore, $\mathcal{I}_{\mathbf{\Omega}}$ is greater than one, but not large enough to be of note. For Fig. 1(b), $\mathcal{I}_{\mathbf{\Omega}}$ 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 Fig. 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 $\mathcal{I}_{\mathbf{\Omega}}$ for the simplified case where the posterior distributions on the sky are uniform distributions, i.e. constant inside the sets Π_{GW} and Π_{EM} and zero outside. Labeling $\Delta(\Pi)$ the area of set Π in square radians, we obtain

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

The 4π prefactor comes from the all-sky prior and acts as a metric to compare the size of the overlap. For example, if Π_{GW} is entirely contained within Π_{EM}, then $\mathcal{I}_{\mathbf{\Omega}}(D_{\mathrm{GW}}, D_{\mathrm{EM}}) = 4\pi/\Delta(\Pi_{\mathrm{EM}})$: 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).
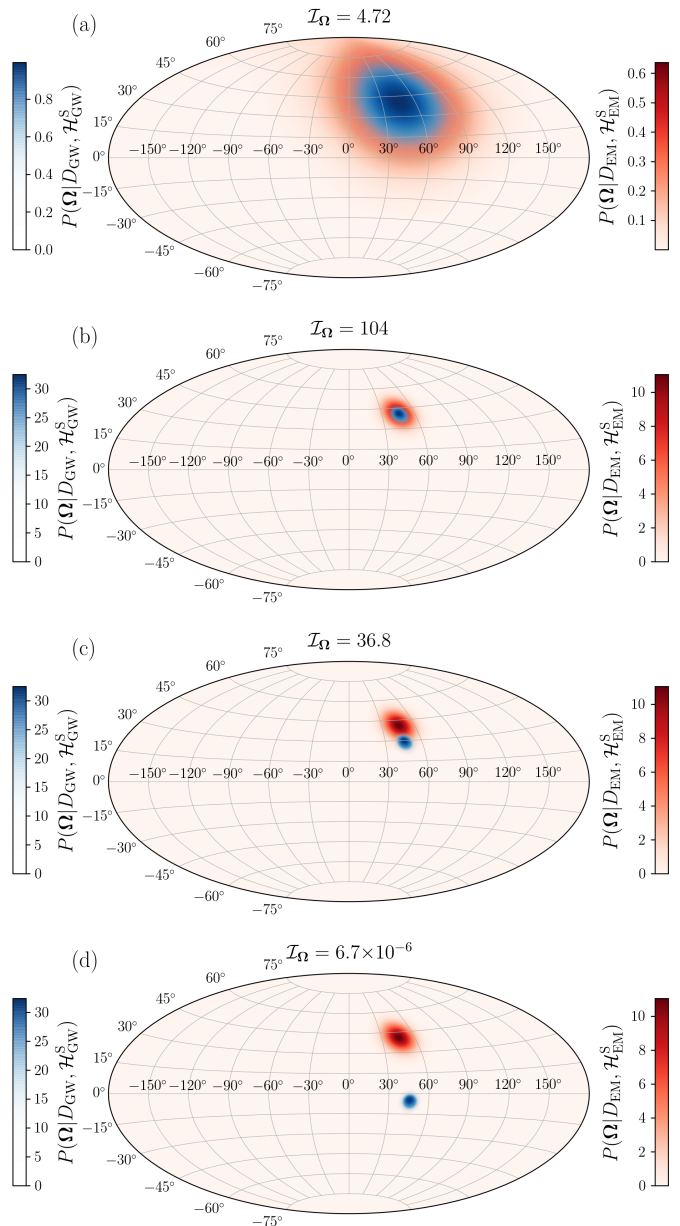

Figure 1. Examples of the spatial overlap $\mathcal{I}_{\mathbf{\Omega}}$. A blue (red) density map shows the probability per pixel for the GW (EM) detection. $\mathcal{I}_{\mathbf{\Omega}}$ is calculated by numerical integration over these pixels with an all-sky uniform prior. The posteriors are computed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

3.1.3. *The spatial and temporal odds*

To calculate the odds in this example via Eq. (5), we require the prior odds. We consider a Poisson point process which 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. Further, 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 other, but not both. Choosing T such that $RT \ll 1$, and defining $\mathrm{Poisson}(1;\lambda)$ to be the probability of one event given an expected number of signals λ, we obtain

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

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

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

which is not dependent on the co-observing time. 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 Eq. (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.

3.2. *Application to GW170817 and GRB 170817A*

We now apply the example calculated in Sec. 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 Fig. 1 of Abbott et al. (2017a). Using the published localization FITS files (Singer 2017; Goldstein et al. 2017) and a uniform prior distribution on the whole sky, Eq. (21) yields $\mathcal{I}_{\boldsymbol\Omega} = 32.4$. 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 90% confidence intervals for the sky localization of GW170817 and GRB 170817A and take the localization posterior distributions to be uniform within those intervals. The intervals cover respectively $28\,\mathrm{deg}^2$ (Abbott et al. 2017c) and $1100\,\mathrm{deg}^2$ (Goldstein et al. 2017) and the GW170817 interval is entirely contained within GRB 170817A. A straightforward application of Eq. (27) then yields an approximate spatial Bayes factor of $\mathcal{I}_{\boldsymbol\Omega} = 37.5$, which is close to the exact value. Since we do have the full posteriors, however, we can repeat this calculation with different confidence levels. We find that $\mathcal{I}_{\boldsymbol\Omega}$ can be biased 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.

In calculating the odds from Eq. (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 approximately $R_{Fermi} = 0.124$ per-day (Abbott et al. 2017a). Then using $[\Delta t^{\min}, \Delta t^{\max}] = [-1, 5]$ s (as used in the Abadie et al. (2012) search) in Eq. (29), the odds, including the spatial overlap, are $\mathcal{O}_{\rm C/SS}(D_{\rm GW}, D_{\rm EM}) \gtrsim 10^6$: 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 Sec. 3.1 and the p-value approach of (Abbott et al. 2017a); namely, the spatial overlap Eq. (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 is exactly our relative degrees of belief for the common- vs. 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 Eq. (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 Eq. (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, consistent with Abbott et al. (2017a).

REFERENCES

Aartsen, M. G., Ackermann, M., Adams, J., et al. 2014, Phys. Rev. D, 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

—. 2017b, PhRvL, 119, 161101

—. 2017c, ApJL, 848, L12

Baret, B., Bartos, I., Bouhou, B., et al. 2012, PhRvD, 85, 103004

Coulter, D. A., Foley, R. J., Kilpatrick, C. D., et al. 2017, Science

Finn, L. S. 1998, in Second Edoardo Amaldi Conference on Gravitational Wave Experiments, ed. E. Coccia, G. Veneziano, & G. Pizzella, 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 (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 University Press)

Haris, K., Mehta, A. K., Kumar, S., & Ajith, P. 2017, private communication

Keivani, A., Fox, D. B., Tešić, G., Cowen, D. F., & Fixelle, J. 2015, ArXiv e-prints, arXiv:1508.01315 [astro-ph.HE]

Kelley, L. Z., Mandel, I., & Ramirez-Ruiz, E. 2013, Phys. Rev. D, 87, 123004

Margutti, R., Berger, E., Fong, W., et al. 2017, ApJL, 848, L20

Savchenko, V., Ferrigno, C., Kuulkers, E., et al. 2017, ApJL, 848, L15

Siellez, K., Boer, M., Gendre, B., & Regimbau, T. 2016, ArXiv e-prints, arXiv:1606.03043 [astro-ph.HE]

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

Troja, E., Piro, L., van Eerten, H., et al. 2017, Nature

Urban, A. L. 2016, Ph.D. dissertation, University of Wisconsin-Milwaukee, https://search.proquest.com/docview/1806791084