

Localization of transient gravitational wave sources: beyond triangulation

Stephen Fairhurst

Cardiff School of Physics and Astronomy, Cardiff University, Queens Buildings,
The Parade, Cardiff. CF24 3AA

E-mail: fairhursts@cardiff.ac.uk

Abstract.

Rapid, accurate localization of gravitational wave transient events has proved critical to successful electromagnetic followup. In previous papers we have shown that localization estimates can be obtained through triangulation based on timing information at the detector sites. In practice, detailed parameter estimation routines use additional information and provide better localization than is possible based on timing information alone. In this paper, we extend the timing based localization approximation to incorporate consistency of observed signals with two gravitational wave polarizations, and an astrophysically motivated distribution of sources. Both of these provide significant improvements to source localization, allowing many sources to be restricted to a single sky region, with an area 40% smaller than predicted by timing information alone. Furthermore, we show that the vast majority of sources will be reconstructed to be circularly polarized or, equivalently, indistinguishable from face-on.

1. Introduction

Numerous gravitational wave signals from binary mergers have now been observed [1, 2, 3, 4]. The sky location of these mergers have been provided with minimal delay, based on low latency gravitational wave (GW) searches [5, 6, 7] and rapid sky localization routines [8]. Subsequently, more detailed parameter estimation routines have provided updated localizations [9]. The rapid analysis and localization enabled the electromagnetic (EM) followup of all GW events and, most spectacularly, the observation of an EM counterpart to GW170817-GRB170817A [10] across the electromagnetic spectrum [11].

In [12, 13], it was argued that timing information provides a good estimate of the localization accuracy of a signal observed in a gravitational wave network. Nonetheless, the localization achieved by full parameter estimation analyses [8, 9, 14, 15] is significantly better than the timing approximation and makes use of the fact that the observed amplitudes and phases of the signal at different detectors must be correlated [8]. A detailed comparison of the results from timing-triangulation and full parameter estimation studies [16] demonstrated that there are important factors above and beyond timing which can improve the localization of sources. Here, we investigate the physical origin of these improvements and provide an intuitive explanation of the most important additional factors in localization. This is achieved by extending the localization calculation to incorporate correlations between the detectors, due to the fact that there are only two gravitational wave polarizations, and astrophysical priors on the source distributions. Inclusion of these effects leads to significant improvements in localization.

The fact that a gravitational wave is comprised of two polarizations guarantees the observed amplitude and phase of the signal at three or more detectors is correlated. Taking this into account improves the localization capabilities of the network. At sky locations other than the correct location, the requirement of signal consistency with two GW polarizations, reduces the reconstructed network signal-to-noise ratio (SNR) which improves the localization [17, 18]. For a three site network, timing triangulation gives two sky locations which are mirror images in the plane defined by the three detector locations. Including amplitude and phase consistency can, in principle, break this degeneracy and allow us to neglect one of the patches. In many cases, the recovered SNR at the mirror position will be of lower amplitude than the original, allowing us to reject the mirror position for around half of the sources.

Gravitational wave sources are expected to be distributed roughly uniformly in volume, at least at the sensitive distances of the advanced GW detectors — large enough to neglect local inhomogeneities and small enough that rate variations are not dominated by cosmological effects or the variation of the star formation rates during the evolution of the universe. Furthermore, the orientation of merging binaries is expected to be uniformly distributed. Consequently, the majority of sources will be observed close to face-on, as the strongest GW signal is emitted in this direction, and at large distances. In a three detector network, the signal at the mirror sky location will typically be reconstructed at a much smaller distance than the true signal, and oriented close to edge on, providing another method of rejecting the mirror sky location.

These effects also improve localization within a single sky patch. At positions offset from the true location, the observed SNRs will not be entirely consistent with a signal, allowing for improved localization. Additionally, the reconstructed orientation of the source will be increasingly edge-on and at a smaller distance as we move away

from the true position. Incorporating amplitude and phase consistency between sites and imposing reasonable astrophysical priors on the source distribution allows us to localize the vast majority of sources to a single patch in the sky. In addition, the error region for most sources is more accurately identified by considering the localization of a circularly polarized signal. For a binary neutron star system observed in three detectors, the cumulative effect is to provide a factor of three improvement in localization over the timing triangulation results.

The majority of the paper deals with the three site advanced LIGO-Virgo network. However, we also consider future four and five detector networks incorporating LIGO India [19, 20] and KAGRA [21]. Sources observed in four or more detectors are localized to a single patch in the sky based on timing alone. Nonetheless, requiring a consistent signal between the sites and favoring distant, approximately face-on sources improves localization significantly.

The paper is laid out as follows. In section 2, we briefly recap the timing triangulation results and then present the extensions to a coherent analysis and the inclusion of astrophysical source distribution. In section 3, we provide the expected localizations for the advanced LIGO-Virgo network at design sensitivity as well as four and five detector networks incorporating KAGRA and LIGO India. Finally, in section 4 we provide a discussion of the results and indicate future directions of research. We provide two appendices: Appendix A which describes in detail the process of marginalizing over the astrophysical priors and Appendix B which gives a detailed derivation of the localization calculation.

2. Localization beyond triangulation

In this section, we present the new features which allow for improved localization. We begin with a brief review of the timing triangulation [12, 13], before extending the formalism to incorporate correlations between detectors due to a coherent analysis and the impact of astrophysically motivated priors on the source parameters, particularly location and orientation of binaries.

2.1. Timing

The accuracy with which the time of arrival of a gravitational wave can be measured in a single detector can be approximated as

$$\sigma_t = \frac{1}{2\pi\rho\sigma_f} \quad (1)$$

where σ_f is the *bandwidth* of the signal in the detector and ρ is the SNR [12]. The SNR is defined as

$$\rho := \frac{(s|h)}{\sqrt{(h|h)}}, \quad (2)$$

where $h(t)$ is the template waveform, $s(t)$ is the detector data,

$$(a|b) := 4 \operatorname{Re} \int \frac{\tilde{a}(f)\tilde{b}^*(f)}{S(f)} df, \quad (3)$$

and $S(f)$ is the power spectral density (PSD). The bandwidth σ_f is given by

$$\sigma_f^2 = \overline{f^2} - \bar{f}^2 \quad \text{where} \quad \bar{f}^n := \frac{(h|f^n h)}{(h|h)}. \quad (4)$$

Defining dt_i as the time difference between the actual time of arrival in a detector and the recovered arrival time, we can obtain the probability distribution for $\mathbf{dt} = (dt_1, \dots, dt_N)$ which is proportional to the likelihood evaluated at those time offsets. Considering timing alone, the time delay distributions in each detector are independent and we obtain the distribution for the time-delay vector \mathbf{dt} as,

$$p(\mathbf{dt}) \propto \exp \left\{ -\frac{1}{2} \left[\sum_i 4\pi^2 \sigma_{f,i}^2 \rho_i^2 dt_i^2 \right] \right\}. \quad (5)$$

This is converted to localization accuracy by re-expressing the relative time delays in terms of sky locations, as discussed in detail in [12, 13] and in Appendix B. We are ignoring the correlation of the arrival time with other parameters, such as the component masses of a binary merger. Since these correlations are similar in all detectors, they will have minimal effect on the measured time delay between sites, which is the important quantity for localization [12, 14].

2.2. Coherent Network Analysis

Since a gravitational wave signal is comprised of only two polarizations, with a network of three or more detectors there are consistency requirements between the observed SNRs in the detectors that comprise the network. As discussed in [22], we can express the network SNR as the projection of the individual detector SNRs onto the two-dimensional signal space, namely the two dimensional physical subspace of signals which is consistent with the two polarizations of gravitational waves.

To do so, we must first introduce the *complex* SNR Z . This is the output of the matched filter of a template waveform h placed at some fiducial distance D_0 , with two orthogonal phases h_0 and $h_{\frac{\pi}{2}}$, against the data s ,

$$Z = \frac{(s|h_0) + i(s|h_{\frac{\pi}{2}})}{\sigma_h} \quad \text{where} \quad \sigma_h^2 = (h_0|h_0). \quad (6)$$

Then $\rho = |Z|$ gives the SNR maximized over the signal amplitude and phase. The network SNR can be expressed as

$$\rho_{\text{net}}^2 = Z_i^* P^{ij} Z_j. \quad (7)$$

where P^{ij} is a rank 2 projection matrix that projects the SNR onto the two gravitational wave polarizations and there is an implicit sum over i and j in (7). When the observed SNRs are consistent with a signal, $P^{ij} Z_j = Z^i$. It is often convenient to work in the dominant polarization frame, where the network is maximally sensitive to the $+$ polarization,[‡] and the projection operator is then given by [22]

$$P^{ij} = \left(\frac{w_+^i w_+^j}{|w_+|^2} + \frac{w_\times^i w_\times^j}{|w_\times|^2} \right) \quad \text{where} \quad w_{+, \times}^i = \sigma_h^i F_{+, \times}^i. \quad (8)$$

The quantities $w_{+, \times}^i$ are the sensitivity weighted detector response functions, and $F_{+, \times}^i$ are the well known detector response functions, see e.g. [24] for the definition.

Consider the situation where the template matches the signal in the detector data, apart from a time and phase offset, and neglect the effect of noise contributions. To

[‡] Specifically, we choose a polarization angle that maximizes $|w_+|$. It can be shown that as a consequence $w_+ \cdot w_\times = 0$ and $|w_\times|$ is minimized. See e.g. [22, 23] for details.

calculate the SNR, we simply note that a time offset can be expressed in the frequency domain as $\exp[2\pi i f dt]$ and expand to second order in dt . Following [12], we have

$$Z_i(dt_i) = Z_i \left[1 + 2\pi i \bar{f}_i dt - 2\pi^2 \bar{f}_i^2 dt_i^2 \right]. \quad (9)$$

Substituting this expansion into (7), we obtain an expression for the network SNR for a given set of time delays \mathbf{dt} . The likelihood, maximized over the amplitude and phase of the signal in the two polarizations, is given by

$$\Lambda_{\max} = \exp \left[\frac{\rho_{\text{net}}^2}{2} \right] \quad (10)$$

as discussed in detail in Appendix A. The associated posterior probability distribution for \mathbf{dt} is proportional to the likelihood so that

$$p(\mathbf{dt}) \propto \exp \left\{ -\frac{1}{2} \left[\sum_i 4\pi^2 \sigma_{f,i}^2 \rho_i^2 dt_i^2 + \sum_{i,j} 4\pi^2 \bar{f}_i dt_i Z_i^* (\delta^{ij} - P^{ij}) \bar{f}_j dt_j Z_j \right] \right\} \quad (11)$$

The first term is identical to the timing triangulation result (5), while the second term imposes amplitude and phase consistency between the sites. Indeed, the timing expression is recovered, as expected, by setting $P^{ij} = \delta^{ij}$. We can intuitively understand the origin of the additional term as follows: when localizing using only time delays, we are free to maximize individually the amplitude and phase of the waveform in each of the detectors; in the coherent network analysis, we are restricted to waveforms consistent with two GW polarizations — any violation of this will cause a loss in SNR as power is removed from the signal space P^{ij} and deposited in the null space, $N^{ij} = \delta^{ij} - P^{ij}$, which, physically, is not expected to contain any GW power [25, 18].

The result (11) has been obtained under two significant approximations. First, we have kept terms only up to quadratic order in dt . This quadratic approximation is accurate for moderate or large SNRs as discussed in [12]. Second, we have neglected the fact that the detector response $F_{+, \times}$ varies across the sky and just used the value at the actual signal location. This can be justified by considering a point offset from the true location by an angle $\delta\phi$. Since the antenna patterns are quadrupolar, the response will characteristically change at a rate $2\delta\phi$. Meanwhile, the time offset at a detector will be $\delta t \sim (\Delta t)\delta\phi$ where Δt is the separation between detectors (on the order of 10s of ms). Therefore, the fractional change in SNR will be $(2\pi\bar{f}\Delta t)\delta\phi \sim 10\delta\phi$, for a typical value of $\bar{f} = 100$ Hz. The effect on phasing is several times larger than the effect of the antenna response, so we ignore the latter in the calculation that follows. For points on the sky where signals are poorly localized, in particular those close to the plane defined by the detectors, the changing antenna response can be important, although the localization will still be poor. The importance of the changing antenna response has been considered in detail in, e.g., [26, 17].

To get a sense of the expected improvement from coherent localization, it is useful to look at a simple example. Consider the situation where we have three detectors which are equally sensitive to a signal and that two detectors are sensitive to the $+$ polarization while the third is sensitive to \times . To a reasonable approximation, the two LIGO detectors are aligned, and consequently sensitive to one GW polarization

while, for many sky locations, Virgo is sensitive to the other GW polarization. Thus, we assume $w_+^H = w_+^L = w_\times^V$ (and $w_\times^H = w_\times^L = w_+^V = 0$). In this case the projection matrix P_{ij} (where the row/column order is H, L, V) is given by

$$P_{ij} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (12)$$

We are interested in the localization of a source, so we want to see how $p(\mathbf{dt})$ falls off as we move away from the true location. The geometry of the LIGO-Virgo network can be approximated by an isosceles triangle, with the LIGO detectors separated by 10ms and the light travel time between each LIGO site and Virgo around 26ms. First, consider recovering the source with time offsets $dt_H = dt_L = -dt_V$, i.e. at a position offset from the true location along a line connecting the midpoint between the LIGO detectors to Virgo. Since $dt_H = dt_L$ and $Z_H = Z_L$ the second line of (11) vanishes and the localization matches the timing result. If, instead, we have $dt_H = -dt_L$ with $dt_V = 0$, i.e. offset from the true position along a line connecting the two LIGO detectors, then the probability distribution is

$$p(\mathbf{dt}) \propto \exp \left\{ -\frac{1}{2} \left[\sum_{i=H,L} 4\pi^2 (\sigma_{f,i}^2 + \bar{f}_i^2) \rho_i^2 dt_i^2 \right] \right\}. \quad (13)$$

This leads to a characteristic width of $(2\pi\rho f_{\text{rms}})^{-1}$, where $f_{\text{rms}} = (\sigma_{f,i}^2 + \bar{f}_i^2)^{1/2}$, rather than $(2\pi\rho\sigma_f)^{-1}$. The improvement arises from the fact that the two LIGO detectors are required to observe the same polarization of GW and hence the same amplitude and phase. For the advanced LIGO detectors at design, $\sigma_f \approx 120\text{Hz}$, $\bar{f} \approx 100\text{Hz}$ and $f_{\text{rms}} \approx 155\text{Hz}$ for binary neutron star (BNS) signals. For this example, the localization area will be reduced by around 25%.

2.3. Mirror Sky Locations

For a gravitational wave detector network comprising three detectors at different sites, localization based on timing will lead to two possible locations, one above and the other below the plane defined by the locations of the detectors. From timing information alone, it is not possible to distinguish the correct location from the *mirror* location. However, once we move to coherent localization, it is, in principle, possible to distinguish between these positions. In the mirror location, the observed amplitudes and phases of the signal will generally not be consistent with a gravitational wave signal comprising only two polarizations — a fraction of the power will be in the null stream [25], thereby reducing the network SNR and allowing us to reject the mirror position.

At the mirror sky location, the measured network SNR will be

$$\rho_{\text{mirror}}^2 = Z_i^* P_{\text{mirror}}^{ij} Z_j \quad (14)$$

where Z_i is the complex SNR in detector i and P_{mirror}^{ij} is the projection onto the signal space appropriate for for the mirror position. Typically, the observed SNRs will not be consistent with the mirror location so $\rho_{\text{mirror}} < \rho_{\text{net}}$.[§]

[§] When the observed SNRs in the detectors at the mirror location are not consistent with a gravitational wave, a nearby sky location may give a larger network SNR even though the SNRs in the single detectors will be lower. This is discussed in detail in Appendix B. In the following discussion, we use ρ_{mirror} to denote the maximum SNR near the mirror location.

There are two peaks in the sky location distribution — one at the true location and the other at the mirror location. However, since the SNR associated to the mirror location is smaller than at the correct location, the peak will be lower and this may allow us to reject it. To identify a confidence region containing a fraction p of the posterior, we typically find the minimum area that contains the required probability. Equivalently, start with the peak of the distribution and add points with the highest available probability until we have incorporated the appropriate fraction of the probability. Thus, if the second peak is low enough, it will not be included in the confidence region.

In many cases, the width of the localization distribution, obtained from (11), at the true and mirror locations is comparable. If we assume that to be the case, then the probability distribution is well approximated by two 2-dimensional Gaussian distributions of equal widths, and the relative height of the second peak is given by

$$\exp\left[-\frac{\Delta\rho^2}{2}\right] \quad \text{where} \quad \Delta\rho^2 = \rho_{\text{net}}^2 - \rho_{\text{mirror}}^2. \quad (15)$$

Then, the second peak will not appear in the localization area provided a fraction p of the total probability lies in the first region, with a likelihood greater than the second peak, i.e

$$1 - \exp\left[-\frac{\Delta\rho^2}{2}\right] > p \left(1 + \exp\left[-\frac{\Delta\rho^2}{2}\right]\right) \quad (16)$$

Here, the left hand side is the integrated probability in the first peak down to the level of the second peak; the right hand side is a p times the total integrated probability. Thus, if

$$\Delta\rho^2 > 2[\ln(1+p) - \ln(1-p)] \quad (17)$$

the source will be localized to a single sky patch.

To obtain a single sky patch at 90% confidence requires $\Delta\rho^2 > 5.9$. For a signal at SNR of 12, this equates to a difference in SNR of 0.25 or greater between the true and mirror position. At 50% confidence, this reduces to $\Delta\rho^2 > 2.2$ or a difference in SNR of less than 0.1. For the advanced LIGO-Virgo detector network at design, about 60% of sources are localized to a single patch at 90% confidence based on this criterion.

2.4. Distribution of Sources

Gravitational wave events from binary mergers are reasonably assumed to be uniformly distributed in volume, at least at distances on the scale of hundreds of mega parsecs. Binary mergers are also expected to be uniformly distributed in source orientation. So far, we have not taken these factors into account when calculating the localization areas. The expression in equation (7) is the maximized SNR — with maximization performed over the amplitude and phase of the two GW polarizations or, equivalently, the distance to and orientation of the binary. In recovering estimates of parameters, it is typical to marginalize (rather than maximize) over any unwanted, or nuisance, parameters. This can have a significant impact upon the localization results.

|| For smaller distances, the discrete structure of galaxies in the nearby universe will affect the distribution [27] while at larger distances, cosmological evolution and the variation of the merger rate with redshift have to be taken into account [28].

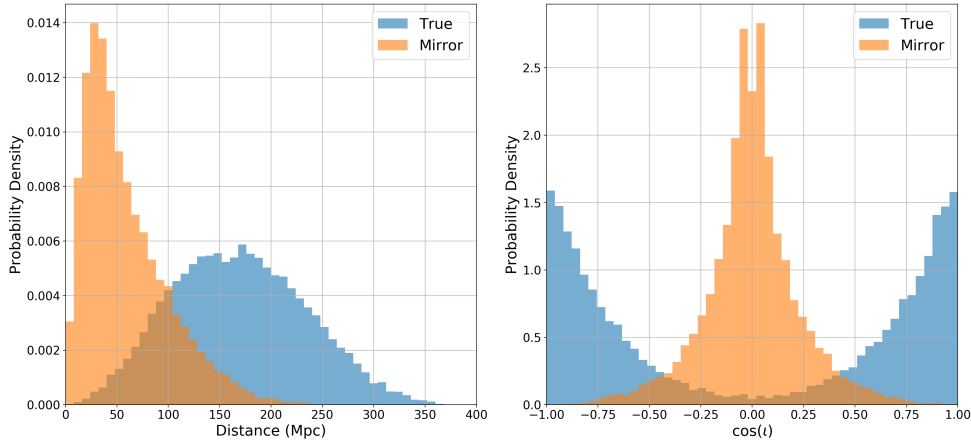


Figure 1. Distribution of inferred distance and inclination, i , of observed sources at the true sky location and the mirror position. The majority of sources at the true location are observed to be face-on and at distances between 100 and 200 Mpc. The majority of sources reconstructed at the mirror location are close to edge on and at distances between 10 and 100 Mpc.

First, consider our ability to exclude the mirror location. For a given sky location, the observed SNRs and detector sensitivities encoded in $w_{+,x}$ can be used to infer the distance to, and orientation of, the source. The distribution of inferred distances for a population of BNS signals that might be observed by the Advanced LIGO-Virgo network operating at design sensitivity is shown in figure 1. This is obtained by generating sources uniform in volume and binary orientation and declaring an event to be observed if the expected network SNR is greater than 12 and the SNR in at least two detectors is greater than 5. The distributions for the correct and mirror locations differ significantly, with the true distance distribution peaking around 150 Mpc and the majority of sources close to face-on. For the mirror location, the inferred distances are, on average, a factor of three smaller, and the majority of sources are recovered edge on.

Initially, let us only consider the distance. The fractional uncertainty in the distance measurement is inversely proportional to the observed SNR, i.e.

$$\frac{\Delta D}{D} \propto \frac{1}{\rho_{\text{net}}} . \quad (18)$$

Assuming that signals are uniformly distributed in volume, the likelihood of observing a signal at a distance D is proportional to $D^2 \Delta D \propto D^3 / \rho_{\text{net}}$. Thus, a simple marginalization over the distance leads to a re-weighting of the likelihood as

$$\Lambda_D \propto \left(\frac{D^3}{\rho_{\text{net}}} \right) \Lambda_{\text{max}} . \quad (19)$$

Then, following a similar argument to the one presented in section 2.3, it follows that if the inferred distance at the mirror sky location is a factor of $8/3$ lower than the true location, then it will be excluded from the 90% localization, even if the maximized likelihood is identical.¶

¶ In addition, signals are statistically more likely to come from regions with greater sensitivity and,

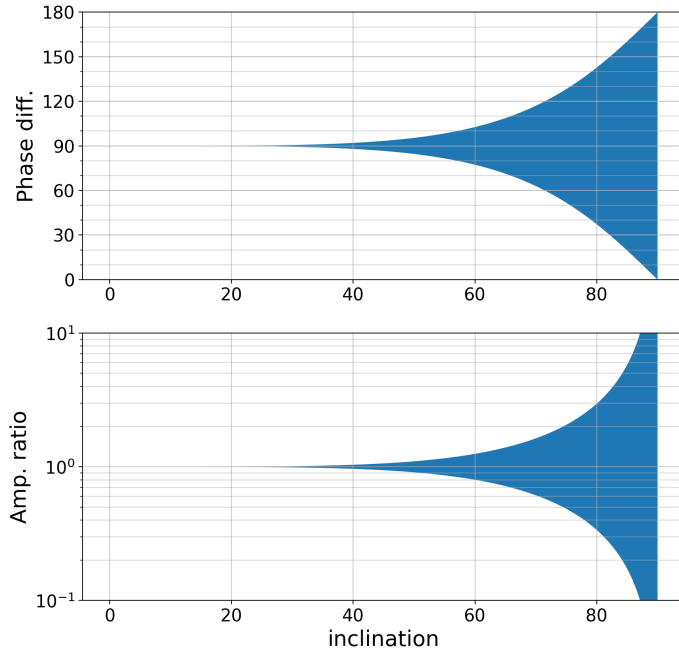


Figure 2. The allowed phase difference between and amplitude ratio of the + and \times gravitational wave polarizations as a function of the binary inclination. For approximately face-on signals, the phase difference is restricted to be very close to 90° while the amplitude ratio is essentially unity. A 10° difference in phase or a 10% difference in amplitude, requires an inclination angle greater than 50° . At an inclination of 90° any amplitude ratio and phase difference is possible.

with the reduced SNR at the mirror location discussed in section 2.2, increases the fraction of sources localized to a single patch with 90% confidence from 60% to 75%.

2.5. Face-on sources

Assuming reasonable parameter priors can also improve the localization accuracy for a single sky location. To illustrate this, consider again a network with three detectors equally sensitive to a given source, with two (H and L) sensitive to the + polarization and the third (V) to \times , so that $w_+^H = w_+^L = w_\times^V$. In addition, assume that the source is face-on, so that the observed gravitational wave is circularly polarized and $Z_H = Z_L = -iZ_V$. If the source is reconstructed at a sky location corresponding to time offsets $dt_H = dt_L = -dt_V$ then, it follows from (9) that, the recovered phase in the H and L detectors will be $2\pi\bar{f}dt$ while in V it will be $(\frac{\pi}{2} - 2\pi\bar{f}dt)$.

In figure 2 we show the range of possible phase and amplitude differences between the observed + and \times polarizations as a function of the binary inclination. For edge on binaries, the emitted GW is linearly polarized so, depending upon the polarization angle, any amplitude ratio and phase difference is possible between + and \times . For face-on signals, the waveform is circularly polarized so the phase difference is required to be

consequently, on average, the mirror location will be less sensitive. Thus, simply weighting the two positions based on the detectors' sensitivity can increase the fraction of sources localized to a single region. However, this is only a minor effect.

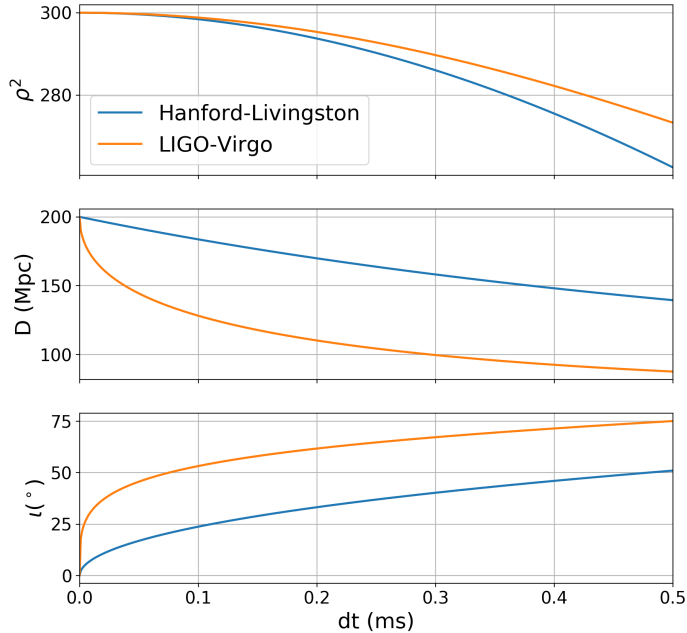


Figure 3. The top panel shows the falloff in SNR as a function of timing offset. We assume a face-on source which gives SNR 10 in all detectors with the two LIGO detectors sensitive to the $+$ polarization, and Virgo to the \times polarization. The middle panel shows the reconstructed distance and bottom panel the inclination angle ι as a function of the timing offset. For time offsets between LIGO-Virgo, the SNR falls off more slowly but, for even a small timing error, the maximum SNR corresponds to a binary at a smaller distance which is significantly inclined.

$\pm 90^\circ$, with equal signal amplitude in the two polarizations. For sources close to face-on, the amplitudes of the two polarizations must be similar and the phase difference close to 90° . Interestingly, the inclination must be large before any significant difference from the face-on case is visible. In particular, a 10° phase difference or 10% amplitude difference is only possible if the binary inclination is greater than 50° .

In figure 3, we show the maximized SNR as a function of the timing offset between the true and reconstructed location, as well as the inferred distance and inclination. This is plotted for both an offset between the two LIGO sites, and an offset between LIGO and Virgo. As expected from section 2.2, the SNR falls off more rapidly with a timing error between the two LIGO sites, as they are sensitive to the same polarization. For changes in the time delay between the LIGO sites, the source is recovered at a slightly reduced distance and not exactly face-on. The inferred amplitude of GW in the $+$ polarization will be slightly reduced due to the phase offset between the signal in the two LIGO detectors. For changes in the LIGO-Virgo time delay, the maximum SNR corresponds to a signal which is significantly closer than the true source and highly inclined. Already with a 0.05 ms offset the reconstructed distance is reduced by 25% and the reconstructed inclination is 40° . As discussed in section 2.4, a lower inferred distance reduces the likelihood and this serves to improve localization.

2.6. Astrophysical Priors

To provide a more accurate assessment of the impact of distance and orientation priors on the source reconstruction, we need to calculate the likelihood marginalized over these degrees of freedom. As discussed in section 2.4, this will give higher weight to sources at a greater distances. It will also give higher weight to sources which are close to face-on or face-off, as the observed waveform varies only slowly with inclination, and the polarization and phase of the signal are degenerate for face-on sources.

In Appendix A, we derive an expression for the approximate marginalized likelihood of a signal based on the observed complex SNRs and detector sensitivities encoded in $w_{+, \times}$. The marginalized likelihood can be approximated by two contributions. The first is simply the area of the peak around the maximum likelihood [29, 30]. However, when the binary is close to circularly polarized, this does *not* give a good approximation to the marginalized result as there is a significant volume of parameter space associated with approximately face-on binaries, for which the waveform is circularly polarized, that provides a large contribution to the likelihood. Thus, we incorporate a second contribution based on the likelihood for approximately circularly polarized waveforms. The total marginalized likelihood is the sum of these two contributions.

To illustrate the result, we return to the example discussed previously, namely a network of three detectors equally sensitive to the signal, all observing with SNR 10. We vary the inclination angle of the source, while keeping the network SNR fixed, and calculate the marginalized likelihood. To make a connection with the previous results, we re-express the marginalized likelihood in terms of a re-weighted SNR as [7]

$$\tilde{\rho}^2(\iota) = \rho^2 + 2 \log \left(\frac{\Lambda_{\text{marg}}(\iota)}{\hat{\Lambda}_{\text{marg}}} \right) \quad (20)$$

where $\Lambda_{\text{marg}}(\iota)$ is the marginalized likelihood for a source with inclination ι , and $\hat{\Lambda}_{\text{marg}}$ is the greatest value of the marginalized likelihood. This allows us to equate an unfavorable orientation with an effective reduction of SNR.

In figure 4 we show the re-weighted SNR computed both numerically and with the approximation based on maximum likelihood and circular polarized signals described above. For face-on, or close to face-on signals, the re-weighted SNR is unchanged. These are the signals which are consistent with the largest volume of parameter space. For edge on systems, $\tilde{\rho}^2$ reduces by almost 20, which is equivalent to reducing the SNR in each detector from 10 to 9.7. Equivalently, the chance of observing a face-on system is over a thousand times higher than an edge on system.

For each binary, we can evaluate whether the likelihood is predominantly accumulated around the maximum SNR or is associated with circularly polarized emission. If it is the former, then the coherent localization presented in section 2.2 is appropriate. However, if the likelihood is dominated by circularly polarized signals, the localization can be improved by imposing this restriction. To do so, we use the previous localization results but now project onto the space of circularly polarized signals using

$$P_{\text{circ}}^{ij} = \left[\frac{(w_+^i \pm iw_\times^i)^*(w_+^j \pm iw_\times^j)}{|w_+|^2 + |w_\times|^2} \right], \quad (21)$$

where the $+/-$ signs give the left and right circular polarizations, respectively. This, in effect, restricts the signal to a single, overall free phase, as conjectured in [16]. In the

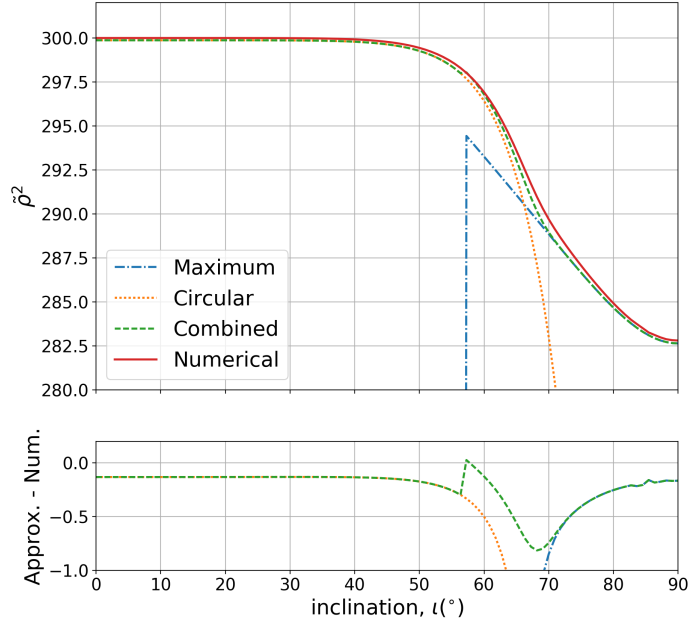


Figure 4. The top panel shows $\tilde{\rho}^2$, the square of the network SNR re-weighted by marginalizing over the distance and orientation parameters, as a function of the binary inclination. The maximum is obtained for signals which are face-on, while for edge on systems $\tilde{\rho}^2$ is reduced by over 5%. The re-weighted SNR is calculated by numerical integration and also by calculating the approximate contributions from around the maximum likelihood and by restricting to circularly polarized signals. For signals with inclination less than 60° it is only the circular polarization contribution that matters, for inclination above 70° only the maximum likelihood is important and between 60° and 70° both contribute. The bottom panel shows the difference between the approximate and the numerical result. The approximate result gives an error in $\tilde{\rho}^2$ less than unity for all values of inclination.

space of time-delays, there is now a single direction — the one that leaves the relative phase between all pairs of detectors unchanged — where the localization is the same as the timing result. For all other directions, the localization width is $(2\pi\rho f_{\text{rms}})^{-1}$.

To demonstrate the utility of this approximation, figure 5 shows the falloff of the SNR under the timing, coherent and circular polarization approximations and compares these to the re-weighted SNR obtained from incorporating astrophysical priors and marginalizing over the likelihood. For a relative time delay between the two LIGO detectors, the coherent and circular polarized results are essentially identical, and match the re-weighted SNR result. Using timing only leads to a slower falloff of SNR and consequently poorer localization. For an offset between LIGO and Virgo, the timing and coherent results are identical, as Virgo is sensitive to a different polarization. Restricting to circular polarization causes the SNR to fall off more quickly, and provides a good match to the re-weighted SNR, at least for time offsets $\lesssim 0.3\text{ms}$. For larger time offsets, the re-weighted SNR falls off more slowly, as at these locations the signal is no longer well approximated as circularly polarized. We note, however, that the 90% source localization is determined by points with $\Delta\rho^2 \leq 4.6$ and in this region the circular polarization approximation works well.

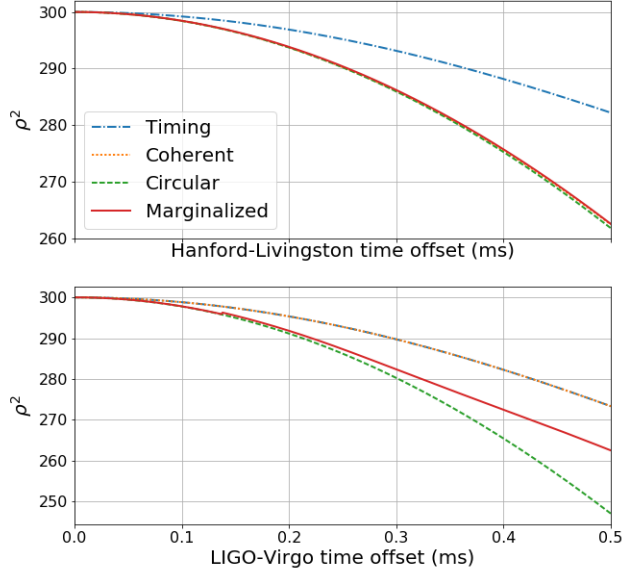


Figure 5. The falloff in SNR as a function of timing error based on timing, coherent and circular polarization approximations. As before, we consider a source with SNR 10 in each of three detectors and introduce a timing offset between the two LIGO detectors (top panel) or LIGO and Virgo (bottom panel). We also show the re-weighted SNR based on marginalizing over astrophysical priors. For a time offset in LIGO, the re-weighted SNR matches well with both the coherent and circular polarization approximations, while for Virgo it matches the circular polarization approximation for time offsets up to 0.3ms.

3. Localization Results

3.1. Advanced LIGO-Virgo network

A realistic scenario for the evolution of the advanced LIGO and Virgo detectors towards their design sensitivities, and the expected localization of sources is given in [31]. We will not reproduce the full observing scenario here, but instead present results for the advanced LIGO-Virgo network operating at design sensitivity: advanced LIGO with a sky averaged sensitivity to BNS mergers of 200 Mpc, mean frequency of 100 Hz and bandwidth of 120 Hz; Virgo with a sensitivity of 130 Mpc, mean frequency of 130 Hz and bandwidth of 150 Hz. In addition, we assume that all detectors have an 80% duty cycle and that the times they operate are independent.

As with previous studies [12, 13, 14], we distribute sources uniformly in volume and binary orientation and deem sources detectable if the expected network SNR is above 12 with an SNR above 5 in at least two detectors. Furthermore, we assume that a detector will contribute to localization provided the SNR in that detector is greater than 4, and neglect entirely contributions from detectors with SNR less than 4. For each source that can be localized by three detectors, we calculate the localization area. We do this by first calculating the marginalized likelihood, as

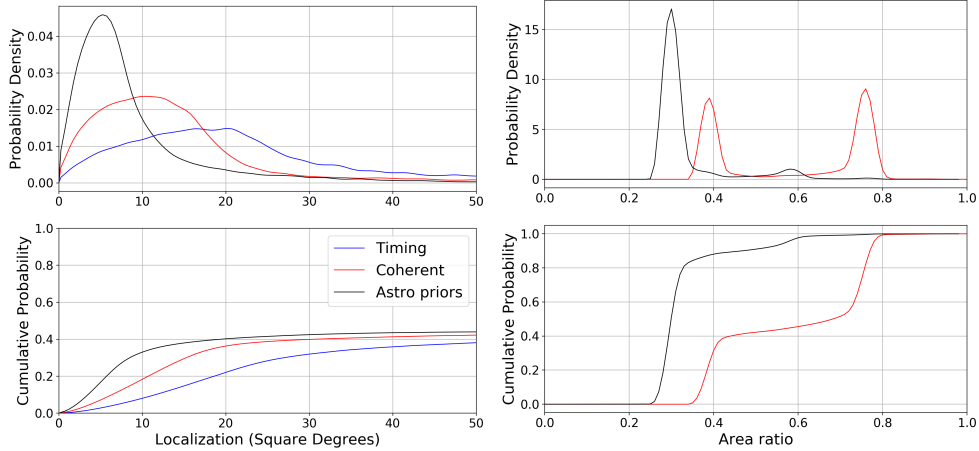


Figure 6. Left: The distribution of localization areas for timing and coherent localizations and the use of astrophysical priors. The top panel shows the area distribution while the bottom panel gives the cumulative distribution. Right: the fractional improvement, on an event by event basis, of the localization area over the timing result.

discussed in section 2.6 and presented in detail in Appendix A. For sources which are essentially face-on, we compute the localization from (11) by first restricting attention to circularly polarized signals using (21). When the signal is not circularly polarized, we project onto the coherent signal space with (8) when calculating the localization. We repeat the calculation for the mirror sky location and combine the distributions from the two locations as discussed in section 2.3. The procedure for going from a timing distribution (11) to an area is presented in Appendix B.

In figure 6, we show the distribution of localization areas for found events, using the three methods described in section 2: timing only, coherent analysis and a coherent analysis incorporating astrophysically motivated priors. We show both the distribution of areas and the fractional improvements for each event. The mode of the sky area distribution is reduced from 20 deg^2 for timing to 10 deg^2 for coherent analysis and 6 deg^2 when we incorporate the expected astrophysical distribution. In all cases only around 40% of sources are localized within 50 deg^2 . The majority of events which are not well localized are seen in only two detectors, either because only 2-detectors are operating due to the assumed duty cycle (about 25%) or the event does not have sufficient SNR in one detector for it to contribute to localization (about 25%) with the final 10% seen with sufficient SNR in all detectors to be localized, but just having poor localization.

The second panel in figure 6 shows the ratio of the coherent or marginalized area to the value obtained with timing triangulation alone on an event-by-event basis. This shows two clear peaks for coherent localization. The peak at around 0.4 corresponds to events which can be localized to a single sky patch using coherent localization, while the second peak at around 0.75 includes those events which still give two sky patches. In both cases, there is a 25% reduction in the area of the sky region due to the requirement of amplitude and phase consistency across the three detectors. By incorporating astrophysical priors, the localization area for the majority of events is reduced to 30% of its original value. This corresponds to events which are localized

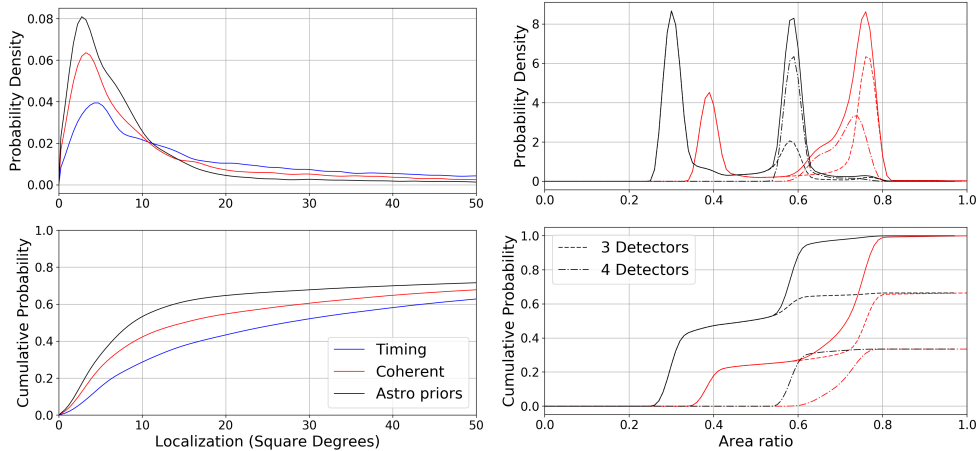


Figure 7. Localization of events in the LIGO Hanford–Livingston–India and Virgo network. Left: The distribution of localization areas for timing and coherent localizations and the use of astrophysical priors. The top panel shows the area distribution while the bottom panel gives the cumulative distribution. Right: the fractional improvement, on an event by event basis, of the localization area over the timing result. Here, the improvement is sub-divided into events which are observed in 3 or 4 detectors as there are different factors which affect the localization in these cases.

in a single sky patch and are effectively face-on. The other, smaller peaks can be explained in a similar way: events with 40% of the original area are localized to a single patch, but not face-on; events at 60% are effectively face-on but from two sky regions; events at 75% are not face-on and come from two sky regions.

Finally, we note that in the above plots, we *never* assume that a source can be localized to a single sky location for the timing results. This is in contrast to the original work [12, 13] which gave timing areas under the assumption that sources would be localized to a single sky patch based on amplitude information. Thus, the distributions for timing-based localization above give twice the area as those presented in e.g. [12, 13]. The coherent localization results presented here are comparable to the timing results in the original papers. However, for each source, the localization is either 25% better due to coherent localization or 50% worse as the event is localized to two sky patches. When we incorporate astrophysical priors, the average localization is 40% smaller than the single patch timing triangulation result.

3.2. Networks of four and five detectors

The longer term plan calls for a global network of five detectors, with the addition of KAGRA and LIGO India [21, 19, 31]. The analysis above can easily be extended to these networks. Sky location degeneracy is not a problem with four or more detectors — a unique sky patch can be determined from the time delays observed between four detectors. However, we can still require consistency with two gravitational wave polarizations, and incorporate astrophysical priors to improve the localization.

In figure 7 we show the expected localizations for events observed by the Advanced LIGO–Virgo network, where we assume that Advanced LIGO incorporates three detectors, with one in India. As before, we show the localization distribution as

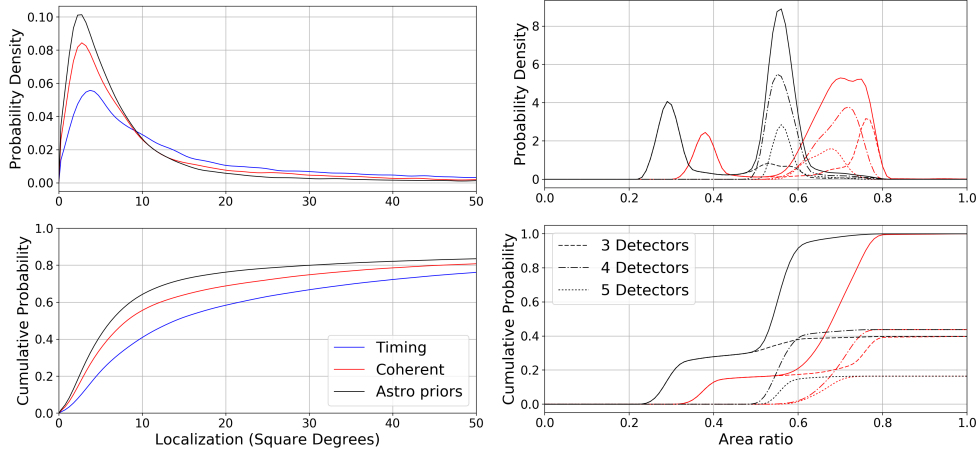


Figure 8. Localization of events in the LIGO Hanford–Livingston–India, Virgo, KAGRA network. Plots are in figure 7 with the fractional improvement in area now divided between 3, 4 and 5 detector observations.

well as the event-by-event improvement of localization. The mode of the localization distribution decreases from 6 deg^2 for timing triangulation to 4 deg^2 for coherent and 3 deg^2 when astrophysical priors are incorporated. Around 70% of found events are localized within 50 deg^2 , with the remainder either observed when only two detectors were operational (7%), only seen in two detectors (11%) or observed in three detectors and poorly localized (12%). Of the events which are well localized, one third are observed in four detectors, due to live time and SNR thresholds, with the remainder observed in three. Since events observed in four detectors are localized to a single sky patch based on timing alone, the majority of these events see a 25 – 35% reduction in area from coherent localization and a 40% reduction when we incorporate an appropriate astrophysical distribution. The fact that coherent localization has a bigger impact for events observed in four detectors (25 – 35% improvement rather than 25% with three detectors) is expected as this leads to a reduction from four data streams, rather than three, to the two GW polarizations.

In figure 8 we show the expected localizations for events observed by the Advanced LIGO–Virgo–KAGRA network, where we assume that Advanced LIGO incorporates three detectors, with one in India. The results are similar to the four-detector network, with the mode of the area distribution decreasing from 4 deg^2 for timing triangulation to 3 deg^2 for coherent and 2 deg^2 when astrophysical priors are incorporated. Around 80% of observed events are localized within 50 deg^2 with 2% of those not localized observed when only two detectors are operational; 5% observed in only two detectors and 13% localized more poorly than 50 deg^2 . For the 5-detector network, around 20% of well-localized events are seen in all five detectors, with 40% in four and the remaining 40% in three. Qualitatively, the 5-detector events are similar to those seen in 4-detectors: they are already localized to a single sky patch so coherent localization serves to reduce the area by 30 – 40%, where again the localization improvement is better than with 3 or 4 detectors from consistency with two GW polarizations. For events seen in four or five detectors, inclusion of astrophysical priors reduces the area by between 40% and 50%.

4. Discussion

In this paper, we have examined in detail two effects that improve source localization over and above what is achieved with the timing triangulation approximation. The first is the requirement of a coherent signal, comprised of two gravitational wave polarizations, observed consistently in all detectors in the network. The second is the incorporation of an astrophysically motivated distribution of events, which are used as priors in the distance and orientation distributions, and lead to the expectation that the majority of observed signals will be distant and close to face-on. Both of these lead to considerable improvements in localization over timing, in the ability to restrict signals to a single sky patch and also to reduce the area of the sky patch. In previous timing triangulation studies, we assumed that sources could be localized to a single sky patch using additional information. Here, we have justified that assumption by showing that the vast majority, over 90%, of sources will be localized to a single sky patch. Furthermore, the area of the sky region will be around 40% smaller than predicted using timing information alone. Finally, we note that the improvement in localization is not uniform for all events — it varies depending upon whether events can, indeed, be localized to a single region and also whether they are recovered as circularly polarized.

In [16] a detailed comparison between timing triangulation and full parameter estimation localization was performed. The results presented here help to explain the discrepancies observed. In a future work, we plan to repeat the detailed comparison with full localization to verify that the new formalism provides a better approximation to the full results, and identify cases where the approximations introduced here are inadequate to capture the full parameter recovery results.

In this paper, we have not considered the localization of events observed in only two detectors. When only two detectors are operating, timing triangulation can only give localization to a ring on the sky. The coherent analysis does not improve this — the observed signal is consistent with a binary at a given distance and orientation at every sky point in the ring [32]. However, for many of these points, the binary will be inferred to be nearby and close to edge-on. By incorporating the astrophysical priors, those points are rejected [33, 34, 8]. In addition, even when three detectors are operating, a reasonable fraction of events will be observed in only two of them. As demonstrated by the observation of GW170817 [4], the lack of a clear signal in one detector can serve to significantly reduce the localization uncertainty of the event, as the source can be restricted to originate from a location close to one of the dead spots for the that detector. The techniques presented here can be extended to the case where two detectors observe a signal, and we will investigate localization with two detectors in a future work.

The analysis presented here is not intended to replace the detailed parameter estimation and localization techniques [8, 9, 15, 35] as it still makes use of a number of approximations. Nonetheless there are useful applications of an approximate method. The calculation of the localization is computationally straightforward, which allows for extensive studies to be run in a straightforward manner. This may identify regions of parameter space which are interesting for more detailed follow-up studies. The method presented here also provides some intuition into the key factors contributing to localization and provides a simple sanity check to the results obtained from complex analysis pipelines. The insights obtained from a simple understanding can also be used to inform development of detailed analyses. For example, it may be possible to

adapt the jump proposals in an MCMC routine to follow the observed degeneracies in the approximations presented here [36]. Alternatively, the approximate likelihood is a natural candidate for the network detection statistic in an analysis pipeline as it naturally folds in the astrophysical weighting of the signal. Indeed, the improved detection statistic presented in [7] was independently motivated by investigations of calculating the marginalized likelihood for two detector observations.

Another application of this analysis is the study of calibration accuracy requirements for current and future detectors. In [12] we investigated the impact of calibration uncertainty on timing-based localization. It would be relatively straightforward to repeat that analysis using the improved localization techniques discussed in this paper. As an example, figure 2 makes it clear that for approximately face-on sources the amplitude of the two gravitational wave polarizations must be nearly equal and the phase difference essentially 90° . And, if the source is well localized, the detector response is known with good accuracy. Then, we can test whether the observed amplitude and phase in the detectors in the network are consistent with this expectation. For example, with GW170817, the sky location is known from EM observations [11] and the GW signal was consistent with circular polarization [37]. From this information, it is possible to calculate the expected relative amplitude and phase differences between the signal observed at the Hanford and Livingston detectors. At the SNR that the event was observed with, this would allow a test of the relative calibration on the order of a few percent in amplitude and a few degrees in phase. A related method of using astrophysical observations to calibrate the detectors has been discussed in [38].

A natural application is to the study of localization capabilities of third generation gravitational wave detector networks. An initial investigation into the performance of homogeneous and heterogeneous networks was performed in [39]. These networks are expected to be sensitive to binary neutron star mergers to $z \sim 1$ and binary black hole mergers to $z \sim 20$. The merger rate will vary significantly over this range and the approximation of a uniform volumetric rate no longer holds. Indeed, as has been pointed out in [40], the bias towards face-on sources will not exist for third generation networks. Consequently, in [39] we used coherent localization, without assuming an astrophysical distribution. However, it should be relatively straightforward to incorporate changes in the astrophysical priors due to redshift effects and the evolution of the merger rate over the history of the universe to obtain more accurate estimates of source localization in third generation gravitational wave networks.

Acknowledgments

We would like to thank many people for interesting discussions on this topic over the years, in particular Duncan Brown, Kipp Cannon, Ilya Mandel, Cameron Mills, Larry Price, Frank Ohme, Bangalore Sathyaprakash, Leo Singer, Patrick Sutton, Vaibhav Tiwari and John Veitch. This research was made possible thanks to support from the Royal Society and the UK Science and Technology Facilities Council (STFC).

Appendix A. Marginalized Likelihood

In this Appendix, we derive an approximate expression for the likelihood, marginalized over distance and orientation parameters. In particular, we calculate the likelihood from a generic orientation and also from a face-on, or face-off, signal and show that the

combined likelihood is well approximated by the sum of these. We begin in Appendix A.1 by briefly reviewing the properties of the waveform and its parametrization. In Appendix A.2, we derive the marginalized likelihood around the maximum SNR and in Appendix A.3 we obtain the result when restricted to circularly polarized signals (face-on or face-off). Finally, we give the combined result in Appendix A.4.

Appendix A.1. The binary merger waveform

The response of a gravitational wave detector to a signal can be expressed as [24]

$$h(t) = F_+(\theta, \phi, \psi)h_+(t) + F_\times(\theta, \phi, \psi)h_\times(t) \quad (\text{A.1})$$

where h_+ and h_\times are the two polarizations of the gravitational wave and F_+ and F_\times are the detector response functions that depend upon location of the source and its polarization.

The merger waveform can be calculated to high accuracy using the post-Newtonian framework [41] during the inspiral phase and numerical relativity simulations when the binary components are close to merger. For a generic binary, the waveform can be complicated due to both the precession of the binary [42] and also the contribution of multiple signal harmonics to the waveform [43]. In this paper, we will not consider either of these effects, and restrict our attention to non-precessing systems that have either no spin or spins aligned with the binary's orbital angular momentum [44] and for which the higher signal harmonics do not contribute significantly. This is appropriate for binary neutron star systems, where the spins of the neutron star do not have a significant effect on the waveform and also for systems of comparable masses where the spins are close to aligned [45, 46, 47, 43]

In this case we can express the waveforms as

$$\begin{aligned} h_+(t) &= \mathcal{A}^1 h_0(t) + \mathcal{A}^3 h_{\frac{\pi}{2}}(t) \\ h_\times(t) &= \mathcal{A}^2 h_0(t) + \mathcal{A}^4 h_{\frac{\pi}{2}}(t), \end{aligned} \quad (\text{A.2})$$

where h_0 and $h_{\frac{\pi}{2}}$ denote the two phases of the waveform and will depend upon the masses and (aligned) spins of the binary. Furthermore, the two phases of the waveform satisfy

$$\tilde{h}_{\frac{\pi}{2}}(f) = i\tilde{h}_0(f). \quad (\text{A.3})$$

The functional form of the \mathcal{A}^i is well known, and depends upon four free parameters: the distance D , inclination ι , polarization ψ and coalescence phase ϕ_0 of the system [48, 32]:

$$\begin{aligned} \mathcal{A}^1 &= \frac{D_0}{D} \frac{(1 + \cos^2 \iota)}{2} \cos 2\phi_0 \cos 2\psi - \frac{D_0}{D} \cos \iota \sin 2\phi_0 \sin 2\psi, \\ \mathcal{A}^2 &= \frac{D_0}{D} \frac{(1 + \cos^2 \iota)}{2} \cos 2\phi_0 \sin 2\psi + \frac{D_0}{D} \cos \iota \sin 2\phi_0 \cos 2\psi, \\ \mathcal{A}^3 &= -\frac{D_0}{D} \frac{(1 + \cos^2 \iota)}{2} \sin 2\phi_0 \cos 2\psi - \frac{D_0}{D} \cos \iota \cos 2\phi_0 \sin 2\psi, \\ \mathcal{A}^4 &= -\frac{D_0}{D} \frac{(1 + \cos^2 \iota)}{2} \sin 2\phi_0 \sin 2\psi + \frac{D_0}{D} \cos \iota \cos 2\phi_0 \cos 2\psi. \end{aligned} \quad (\text{A.4})$$

Here, D_0 is a fiducial distance used in generating the waveforms h_0 and $h_{\frac{\pi}{2}}$. The choice of D_0 is arbitrary and does not affect any of the results presented below. Furthermore,

every set of values for \mathcal{A}^μ corresponds to a set of physical parameters, unique up to an overall reflection symmetry, $\psi \rightarrow \psi + \frac{\pi}{2}$ and $\phi \rightarrow \phi + \frac{\pi}{2}$.

Combining equations (A.1) and (A.2), we can express the signal h at a detector i as

$$h_i = (\mathcal{A}^1 F_+^i + \mathcal{A}^2 F_\times^i) h_0(t_i) + (\mathcal{A}^3 F_+^i + \mathcal{A}^4 F_\times^i) h_{\frac{\pi}{2}}(t_i) \quad (\text{A.5})$$

where t_i is the time of arrival of the signal at the detector and \mathcal{A}^μ are the four amplitudes introduced in (A.4).

Appendix A.2. Marginalized Likelihood

The probability that a given signal h is present in the data is proportional to the likelihood. For a network of detectors, the log-likelihood is simply the sum of individual detector contributions:

$$\ln \Lambda = \sum_i \ln \Lambda_i = (\mathbf{s}|\mathbf{h}) - \frac{1}{2}(\mathbf{h}|\mathbf{h}) \quad \text{where} \quad (\mathbf{a}|\mathbf{b}) = \sum_i (a_i|b_i) \quad (\text{A.6})$$

and the inner product is defined in (3).

It has been shown, see e.g. [49], that the likelihood can be maximized over the four amplitude parameters encoded by the \mathcal{A}^μ . Here, we will re-derive this result as a starting point for marginalizing over the physical parameters, using a notation that more closely follows that used in gravitational wave searches [50, 6]. In particular, we work with the complex SNR, Z , defined in equation (6). We combine the \mathcal{A}^μ into two complex parameters

$$\mathcal{A}^+ = \mathcal{A}^1 + i\mathcal{A}^3 \quad \text{and} \quad \mathcal{A}^\times = \mathcal{A}^2 + i\mathcal{A}^4. \quad (\text{A.7})$$

Then, we can write the inner product between signal, s , and template waveform, h , as

$$(s|h) = \text{Re} [(\mathcal{A}^+ w_+ + \mathcal{A}^\times w_\times)^* Z], \quad (\text{A.8})$$

where $w_{+, \times}$ are the sensitivity weighted detector response functions introduced in (8).

The log-likelihood for the network can then be expressed as

$$\ln \Lambda = \text{Re} [(\mathcal{A}^\alpha)^* w_\alpha^i Z_i] - \frac{1}{2}(\mathcal{A}^\alpha)^* \mathcal{M}_{\alpha\beta} \mathcal{A}^\beta \quad (\text{A.9})$$

where there is an implicit sum over both $\alpha \in (+, \times)$ and i , which runs over the detectors in the network. The matrix \mathcal{M} encapsulates the network's sensitivity to the two GW polarizations, and is defined as

$$\mathcal{M}_{\alpha\beta} = \sum_i w_\alpha^i w_\beta^i. \quad (\text{A.10})$$

It is straightforward to differentiate the likelihood and show that the maximum occurs at

$$\hat{\mathcal{A}}^\beta = \mathcal{M}^{\alpha\beta} w_\alpha^i Z_i, \quad (\text{A.11})$$

where $\mathcal{M}^{\alpha\beta}$ is taken to be the inverse of $\mathcal{M}_{\alpha\beta}$ [29]. The maximum likelihood is

$$\ln \Lambda_{\text{max}} = \frac{1}{2}(\hat{\mathcal{A}}^\alpha)^* \mathcal{M}_{\alpha\beta} \hat{\mathcal{A}}^\beta = \frac{1}{2} Z_i^* P^{ij} Z_j \quad (\text{A.12})$$

where P_{ij} is a projection onto the two-dimensional signal space

$$P^{ij} = w_\alpha^i \mathcal{M}^{\alpha\beta} w_\beta^j. \quad (\text{A.13})$$

In the dominant polarization, where $w_+ \cdot w_\times = 0$, $\mathcal{M}_{\alpha\beta}^{\text{DP}} = \text{diag}(w_+^2, w_\times^2, w_+^2, w_\times^2)$ and the projection operator can be expressed as

$$P_{\text{DP}}^{ij} = \frac{w_+^i w_+^j}{|w_+|^2} + \frac{w_\times^i w_\times^j}{|w_\times|^2}. \quad (\text{A.14})$$

The likelihood can be written as

$$\ln \Lambda = \frac{1}{2}(\hat{\mathcal{A}}^\alpha)^* \mathcal{M}_{\alpha\beta} \hat{\mathcal{A}}^\beta - \frac{1}{2} \text{Re} \left[(\hat{\mathcal{A}}^\alpha - \mathcal{A}^\alpha)^* \mathcal{M}_{\alpha\beta} (\hat{\mathcal{A}}^\beta - \mathcal{A}^\beta) \right]. \quad (\text{A.15})$$

The first term is the maximum likelihood while the second term shows the quadratic falloff around the maximum. The benefit of using the \mathcal{A} variables is clear, as the likelihood is a multivariate Gaussian.

We now wish to marginalize (A.15) over the four parameters D, ι, ψ, ϕ . Taking a prior that is uniform in the \mathcal{A}^μ the integral can be easily performed. As discussed in detail in [29], the flat prior on the \mathcal{A}^μ corresponds to a highly non-physical prior on the original parameters. In particular,

$$d\mathcal{A}^1 d\mathcal{A}^2 d\mathcal{A}^3 d\mathcal{A}^4 = \frac{D_0^4 dD}{2D^5} (1 - \cos^2 \iota)^3 d \cos \iota d\psi d\phi, \quad (\text{A.16})$$

where, as before, D_0 is the fiducial distance used in defining $h_0, \frac{\pi}{2}$.

Simple astrophysical considerations lead us, instead, to choose a prior that is uniform in volume, with source orientations uniformly distributed on a two-sphere and a uniform distribution in phase,

$$\frac{3D^2 dD}{D_{\text{max}}^3} \frac{d \cos \iota}{2} \frac{d\psi}{2\pi} \frac{d\phi}{2\pi} = \frac{3D^7}{4\pi^2 D_0^4 D_{\text{max}}^3} (1 - \cos^2 \iota)^{-3} d\mathcal{A}^1 d\mathcal{A}^2 d\mathcal{A}^3 d\mathcal{A}^4, \quad (\text{A.17})$$

where D_{max} is the maximum distance to which the integral is performed. The normalization is chosen so that the prior integrates to unity. Comparing (A.16) to (A.17), it is clear that the flat in amplitudes prior significantly over-weights nearby sources which are close to edge-on, as discussed in [29, 30].

The exact solution for the marginalized likelihood has been calculated for number of special cases in [30] and was recently computed in general [51]. Here, we will not use the exact expression, but rather look to obtain an approximate expression which is accurate enough to be applied to the localization problem. The most straightforward method is to re-weight the likelihood by the value of the prior at the maximum likelihood values $\hat{\mathcal{A}}^\mu$ [30]. Under this approximation, it is straightforward to evaluate the Gaussian integral as

$$\Lambda_{\text{marg}} \approx \frac{24 \hat{D}^7}{D_0^4 D_{\text{max}}^3} \frac{(1 - \cos^2 \hat{\iota})^{-3}}{w_+^2 w_\times^2} \exp \left[\frac{\rho_{\text{net}}^2}{2} \right], \quad (\text{A.18})$$

where the factor of 8 arises due to the fact that the \mathcal{A}^μ are functions of 2ϕ and 2ψ , giving a factor of 4, and the existence of a discrete degeneracy $\phi \rightarrow \phi + \pi/2$ and $\psi \rightarrow \psi + \pi/2$ gives the final factor of two. The Gaussian integral evaluates to $4\pi^2/(w_+^2 w_\times^2)$, as can easily be seen in the dominant polarization frame.

Appendix A.3. Circularly Polarized Signal

Unfortunately, the simple approximation (A.18) breaks down when the system is close to circularly polarized, i.e. $|\cos \iota| \approx 1$, as discussed in detail in [30]. The reason is that the marginalized likelihood is dominated by contributions away from the peak, and so approximating the integral based on the area around the maximum likelihood no

longer works. For a face-on binary, the signal is circularly polarized and there are only two remaining degrees of freedom: the overall amplitude and a single phase $2(\phi \pm \psi)$, with the sign depending upon if the system is left or right circular. Thus, the volume of phase space consistent with a circularly polarized binary is significantly larger than for other systems. Also, for face-on binaries, the inferred distance is the largest, again increasing the marginalized likelihood.

Here, we provide a simple approximation to the likelihood for approximately circularly polarized signals. We restrict to circularly polarized waveforms, as described in [52], by requiring

$$\mathcal{A} := \mathcal{A}_+ = \pm i \mathcal{A}_\times, \quad (\text{A.19})$$

where the positive corresponds to left-circular polarization and negative to right. The complex amplitude \mathcal{A} can be written as

$$\mathcal{A} \approx \frac{D_0}{D} \cos \iota e^{-i\phi_{L,R}} \quad \text{where} \quad \phi_{L,R} = 2(\phi_0 \pm \psi), \quad (\text{A.20})$$

and the likelihood as

$$\ln \Lambda = \text{Re} [\mathcal{A}^* (w_+^i \pm i w_\times^i) Z_i] - \frac{1}{2} |\mathcal{A}|^2 (|w_+|^2 + |w_\times|^2) \quad (\text{A.21})$$

As before, it's straightforward to maximize and obtain the value of \mathcal{A} at the peak,

$$\hat{\mathcal{A}} = \frac{(w_+^i \pm i w_\times^i) Z_i}{|w_+|^2 + |w_\times|^2}, \quad (\text{A.22})$$

and the maximum likelihood

$$\ln \Lambda_{\max} = \frac{1}{2} (|w_+|^2 + |w_\times|^2) \hat{\mathcal{A}}^* \hat{\mathcal{A}} = \frac{1}{2} Z_i^* P^{ij} Z_j \quad (\text{A.23})$$

where P_{ij} is a projection onto the one-dimensional signal space

$$P^{ij} = \frac{(w_+^i \pm i w_\times^i)^* (w_+^j \pm i w_\times^j)}{|w_+|^2 + |w_\times|^2}. \quad (\text{A.24})$$

The likelihood can be expressed as

$$\ln \Lambda = \frac{1}{2} (|w_+|^2 + |w_\times|^2) \left[\hat{\mathcal{A}}^* \hat{\mathcal{A}} - (\hat{\mathcal{A}} - \mathcal{A})^* (\hat{\mathcal{A}} - \mathcal{A}) \right]. \quad (\text{A.25})$$

This has the form of a multivariate Gaussian, but here in two, rather than four, dimensions. As before, we wish to rewrite the astrophysical prior in terms of variables appearing in the likelihood. Here, it is natural to use $|\mathcal{A}|$, $\phi_{L,R}$ and $\cos \iota$. Then,⁺

$$\frac{3D^2 dD}{D_{\max}^3} \frac{d \cos \iota}{2} \frac{d\psi}{2\pi} \frac{d\phi}{2\pi} = \frac{3D_0^3 \cos^3 \iota}{2D_{\max}^3 |\mathcal{A}|^5} d \cos \iota |\mathcal{A}| d|\mathcal{A}| \frac{d\phi_L}{2\pi} \frac{d\phi_R}{2\pi}.$$

The likelihood depends upon $|\mathcal{A}|$ and one of $\phi_{L,R}$, but is independent of the second phase angle and $\cos \iota$. We begin by integrating out the second phase angle and ι . While the likelihood is independent of ι , the approximation of circular polarization is only appropriate for a limited range of ι . Consider a generic, not necessarily circularly polarized, signal then the loss in SNR from projecting this onto the circular polarization is

$$\Delta \rho^2 = \rho_{\text{net}}^2 - \rho_{\text{circ}}^2 = \left(\frac{D_0}{D} \right)^2 \frac{w_+^2 w_\times^2}{w_+^2 + w_\times^2} \frac{(1 \mp \cos \iota)^4}{4}. \quad (\text{A.26})$$

⁺ The factor of 8 in the Jacobian of the transformation to $\phi_{L,R}$ is compensated by the fact that the space $\phi_{L,R} \in [0, 2\pi)$ is covered eight times for $\phi_0, \psi \in [0, 2\pi)$ (see [30] for details).

The expected value of ρ^2 due solely to noise contributions is equal to the number of degrees of freedom. Thus, we allow for $\Delta\rho^2 \leq 2$. For values smaller than this, the loss in SNR from projection to circular polarization will be indistinguishable from noise. We then can solve for the minimum value of $\cos \iota$

$$|\cos \iota|_{\min} = 1 - \left[\frac{2\sqrt{2} (w_+^2 + w_\times^2)}{\rho w_+ w_\times} \right]^{1/2}, \quad (\text{A.27})$$

subject to the restriction that $|\cos \iota|_{\min} \geq 0$. This allows us to approximate the ι integral as

$$\frac{1 - |\cos \iota|_{\min}^4}{4}. \quad (\text{A.28})$$

Finally, we can perform the Gaussian integral, and obtain a factor of $2\pi/(w_+^2 + w_\times^2)$. Then, evaluating the prior at the peak, using $|\hat{\mathcal{A}}| = \frac{D_0}{\hat{D}}$ where \hat{D} is the distance associated to a face-on signal, we obtain

$$\Lambda_{\text{marg}} \simeq \frac{3}{8} \frac{\hat{D}^5}{D_0^2 D_{\text{max}}^3} \frac{1 - |\cos \iota|_{\min}^4}{(w_+^2 + w_\times^2)} \exp \left[\frac{\rho_{\text{circ}}^2}{2} \right]. \quad (\text{A.29})$$

Appendix A.4. Combined Likelihood

We approximate the marginalized likelihood by three contributions: one from around the peak likelihood, (A.18), and one from each of the left and right circularly polarized restrictions, (A.29). As is clear from (A.18), the peak likelihood contribution diverges as $\cos \iota \rightarrow \pm 1$ as the approximations used break down. To rectify this, we discard the contribution from the peak any time that

$$\rho_{\text{net}}^2 - \rho_{\text{circ}}^2 \leq 2. \quad (\text{A.30})$$

As discussed above, when the difference between the peak SNR and the circular projection of the SNR is less than two, there is no significant evidence for power in the second polarization so the circular approximation is more appropriate.

In figure 4 in the main text we show the dependence of the marginalized likelihood, expressed as a re-weighted SNR, as a function of the inclination angle for a source with fixed SNR. When the inclination angle is less than 55° the marginalized likelihood, calculated numerically, is in good agreement with the circular polarization and the restriction (A.30) is met so that the peak likelihood doesn't contribute at all. For angles up to 40° the marginalized likelihood is constant, as the signal is circularly polarized to a very good approximation. Above 40° the signal is no longer perfectly circular and projecting onto the circular polarization leads to a loss of SNR, which is seen from the face-on contribution decreasing rapidly. There is a small discontinuity in the likelihood at around 55° as we transition to including the contribution around the maximum likelihood. For angles above 70° the likelihood is well approximated by the contribution around the maximum likelihood. For the edge-on system the re-weighted SNR is reduced from 300 to 283. Since the likelihood, and posterior probability, is proportional to $\exp[\rho^2/2]$, the edge-on signal is down-weighted by a factor of almost 5,000 relative to a face-on signal.

Appendix B. Localization

In this appendix, we present the details of the localization calculation. Furthermore, we show how the posterior distribution on time delays can be converted to a localization distribution over the sky.

Appendix B.1. Timing accuracy

To calculate the leading order localisation expression, we expand network SNR (7)

$$\rho_{\text{net}}^2 = \sum_{i,j} Z_i^* P^{ij} Z_j. \quad (\text{B.1})$$

to quadratic order in dt_i , using the expression in equation (9). As discussed in the main text, we neglect changes in the projection operator P^{ij} as we argue they will be less significant than changes in SNR. When the observed SNRs are consistent with a signal, the projection operator is idempotent, i.e.

$$P^{ij} Z_j = Z^i \quad (\text{B.2})$$

However, in the presence of noise, or when reconstructing the signal from an incorrect sky location, this will no longer be the case. Consequently, we keep the calculation below general, and require only that the projection operator be Hermitian.

Inserting the expansion of for $Z(dt)$ in powers of dt (9) into (7) and keeping terms up to second order in dt , we obtain

$$\begin{aligned} \rho_{\text{net}}^2(\mathbf{dt}) &= \sum_{i,j} Z_i^* P^{ij} Z_j + 4\pi \text{Im} \left[\sum_{i,j} (\bar{f}_i dt_i) Z_i^* P^{ij} Z_j \right] \\ &\quad - 4\pi^2 \text{Re} \left[\sum_{i,j} Z_i^* P^{ij} Z_j (\bar{f}_i^2 dt_i^2 - \bar{f}_i dt_i \bar{f}_j dt_j) \right] \\ &=: \rho_{\text{net}}^2 + B^i dt_i - C^{ij} dt_i dt_j, \end{aligned} \quad (\text{B.3})$$

where we have used the fact that the projection operator is Hermitian to simplify the result. In the above, we have defined

$$\rho_{\text{net}}^2 = \sum_{i,j} Z_i^* P^{ij} Z_j \quad (\text{B.4})$$

$$B^i = 4\pi \text{Im} \left[\sum_j \bar{f}_i Z_i^* P^{ij} Z_j \right] \quad (\text{B.5})$$

$$C^{ij} = 4\pi^2 \text{Re} \left[\left(\sum_k Z_i^* P^{ik} Z_k \right) \bar{f}_i^2 \delta_{ij} - (Z_i^* P^{ij} Z_j) \bar{f}_i \bar{f}_j \right] \quad (\text{B.6})$$

The three quantities in (B.3) have simple interpretations. The first, ρ_{net}^2 is the network SNR for a source at the original position (i.e. with $dt_i = 0$). If the observed SNRs are consistent with a signal, i.e. eq. (B.2) holds, then B^i vanishes as we are taking the imaginary part of a real quantity. If the single detector SNRs are not consistent with a source, then the maximum network SNR can be offset from $dt_i = 0$: a nearby sky location may give a larger coherent SNR *even though the single detector SNRs will be reduced*. The third term, C_{ij} determines the quadratic falloff in SNR away from the peak.

Restricting to the case where the observed SNRs Z_i are consistent with a signal from the given sky location, (B.3) reduces to the expression (11) given in section 2.2.

Appendix B.2. Sky Localization

To convert the timing expression, (B.3) to localization area, we use the same approach as described in [12, 13]. In this paper, we simulate signals from a given sky location \mathbf{r} and at a given geocentric arrival time t_0 . From these, we can construct the arrival time in each detector as

$$t_i = t_0 - \mathbf{r} \cdot \mathbf{D}_i, \quad (\text{B.7})$$

where \mathbf{D}_i is the location of the i th detector.* Furthermore, we use the known location to calculate the projection operator P^{ij} and, in the case of three detector networks, the projection at the mirror location P_{mirror}^{ij} . We then calculate the quantities B_i and C_{ij} from (B.5) and (B.6) which give the distribution for the arrival times T_i as

$$p(T_i|t_i, Z_i) \propto p(T_i) \exp \left[\frac{1}{2} \{B^i(T_i - t_i) - C^{ij}(T_i - t_i)(T_j - t_j)\} \right]. \quad (\text{B.8})$$

We are interested in re-expressing this in terms of a distribution for the the sky location, \mathbf{R} and geocentric arrival time T_0 , which are related to the arrival times by

$$T_i = T_0 - \mathbf{R} \cdot \mathbf{D}_i. \quad (\text{B.9})$$

For a simulated signal, the observed SNRs will be consistent with a signal from the sky location \mathbf{r} . However, they will not necessarily be consistent with either a circularly polarized signal from that sky location, which we will use in approximating localizations, or with a signal from the mirror location $\mathbf{r}_{\text{mirror}}$. In this case, the centre of the reconstructed localization region will be offset from the true (or mirror) location and the peak SNR will be lower than the total SNR observed in the network. In particular, we obtain a maximum coherent SNR with time offsets

$$(\hat{t}_i - t_i) = \frac{1}{2} C_{ij}^{-1} B^j, \quad (\text{B.10})$$

which gives the maximum SNR as

$$\hat{\rho}_{\text{coh}}^2 := \rho_{\text{coh}}^2 + C^{ij}(\hat{t}_i - t_i)(\hat{t}_i - t_i) \quad (\text{B.11})$$

Using the time offsets that maximize the SNR, we construct a best fit sky location and arrival time

$$\hat{t}_i = \hat{t}_0 - \hat{\mathbf{r}} \cdot \mathbf{D}_i \quad (\text{B.12})$$

which allows us to re-express the localization distribution as

$$p(T_0, \mathbf{R}|t_i, Z_i) \propto p(T_0, \mathbf{R}) \exp \left[-\frac{1}{2} C^{ij} dt_i dt_j \right]. \quad (\text{B.13})$$

where

$$dt_i = T_0 - \hat{t}_0 - (\mathbf{R} - \hat{\mathbf{r}}) \cdot \mathbf{D}_i \quad (\text{B.14})$$

Finally, we would like to eliminate T_0 to obtain the distribution over the sky. Since the time of arrival will not be known, we simply marginalize over it using a

* As discussed in the Appendix of [13], given an observed set of SNRs and times of arrival t_i , we can use a constrained maximization to calculate the best fit \mathbf{r} and t_0 and from these the best fit arrival times t_i consistent with a physical source. We then use this location to calculate the projection operator in the following.

flat prior on T_0 . In this case, marginalizing the distribution over T_0 is equivalent to maximizing the SNR over this variable (details in the appendix of [13]). Upon doing so, we find that the localization distribution can be expressed as:

$$p(T_0, \mathbf{R}|t_i, Z_i) \propto p(\mathbf{R}) \exp \left[-\frac{1}{2} (\hat{\mathbf{r}} - \mathbf{R})^T \mathbf{M} (\hat{\mathbf{r}} - \mathbf{R}) \right] \quad (\text{B.15})$$

where the matrix \mathbf{M} is given as

$$\mathbf{M} = \frac{1}{2} \sum_{i,j} \mathbf{D}_{ij} \mathbf{D}_{ij}^T \left[\frac{c^i c^j}{c} - C^{ij} \right] \quad (\text{B.16})$$

and

$$c^i = \sum_j C^{ij} \quad \text{and} \quad c = \sum_i c^i. \quad (\text{B.17})$$

The localization matrix is very similar in form to the one obtained for timing only. The main difference is that there are terms involving correlations between the detectors which were not present before. It is then reassuring to note that the triangulation expression can be recovered by the replacement $P_{ij} \rightarrow \delta_{ij}$. In this case A_i vanishes and $C_{ij} = \delta_{ij} c_i$, and $c^i = (2\pi\rho_i\sigma_{f,i})^2$.

References

- [1] Abbott B P *et al.* (Virgo, LIGO Scientific) 2016 *Phys. Rev. Lett.* **116** 061102 (*Preprint* 1602.03837)
- [2] Abbott B P *et al.* (Virgo, LIGO Scientific) 2016 *Phys. Rev. Lett.* **116** 241103 (*Preprint* 1606.04855)
- [3] Abbott B P *et al.* (Virgo, LIGO Scientific) 2016 *Phys. Rev.* **X6** 041015 (*Preprint* 1606.04856)
- [4] Abbott B P *et al.* (Virgo, LIGO Scientific) 2017 *Phys. Rev. Lett.* **119** 161101 (*Preprint* 1710.05832)
- [5] Messick C *et al.* 2017 *Phys. Rev.* **D95** 042001 (*Preprint* 1604.04324)
- [6] Usman S A *et al.* 2016 *Class. Quant. Grav.* **33** 215004 (*Preprint* 1508.02357)
- [7] Nitz A H, Dent T, Dal Canton T, Fairhurst S and Brown D A 2017 *Astrophys. J.* **849** 118 (*Preprint* 1705.01513)
- [8] Singer L P and Price L R 2016 *Phys. Rev.* **D93** 024013 (*Preprint* 1508.03634)
- [9] Veitch J, Raymond V, Farr B, Farr W, Graff P, Vitale S, Aylott B, Blackburn K, Christensen N, Coughlin M, Del Pozzo W, Feroz F, Gair J, Haster C J, Kalogera V, Littenberg T, Mandel I, O’Shaughnessy R, Pitkin M, Rodriguez C, Röver C, Sidery T, Smith R, Van Der Sluys M, Vecchio A, Voudsen W and Wade L 2015 *Phys. Rev. D* **91**(4) 042003 (*Preprint* 1409.7215)
- [10] Abbott B P *et al.* (Virgo, Fermi-GBM, INTEGRAL, LIGO Scientific) 2017 *Astrophys. J.* **848** L13 (*Preprint* 1710.05834)
- [11] Abbott B P *et al.* (GROND, SALT Group, OzGrav, DFN, INTEGRAL, Virgo, Insight-Hxmt, MAXI Team, Fermi-LAT, J-GEM, RATIR, IceCube, CAASTRO, LWA, ePESSTO, GRAWITA, RIMAS, SKA South Africa/MeerKAT, H.E.S.S., 1M2H Team, IKI-GW Follow-up, Fermi GBM, Pi of Sky, DWF (Deeper Wider Faster Program), Dark Energy Survey, MASTER, AstroSat Cadmium Zinc Telluride Imager Team, Swift, Pierre Auger, ASKAP, VINROUGE, JAGWAR, Chandra Team at McGill University, TTU-NRAO, GROWTH, AGILE Team, MWA, ATCA, AST3, TOROS, Pan-STARRS, NuSTAR, ATLAS Telescopes, BOOTES, CaltechNRAO, LIGO Scientific, High Time Resolution Universe Survey, Nordic Optical Telescope, Las Cumbres Observatory Group, TZAC Consortium, LOFAR, IPN, DLT40, Texas Tech University, HAWC, ANTARES, KU, Dark Energy Camera GW-EM, CALET, Euro VLBI Team, ALMA) 2017 *Astrophys. J.* **848** L12 (*Preprint* 1710.05833)
- [12] Fairhurst S 2009 *New Journal of Physics* **11** 123006
- [13] Fairhurst S 2011 *Class. Quantum Grav.* **28** 105021 (*Preprint* 1010.6192)
- [14] Singer L P, Price L R, Farr B, Urban A L, Pankow C *et al.* 2014 *Astrophys. J.* **795** 105 (*Preprint* 1404.5623)
- [15] Berry C P L *et al.* 2015 *Astrophys. J.* **804** 114 (*Preprint* 1411.6934)

- [16] Grover K, Fairhurst S, Farr B F, Mandel I, Rodriguez C *et al.* 2014 *Phys.Rev.* **D89** 042004 (*Preprint* 1310.7454)
- [17] Wen L and Chen Y 2010 *Phys.Rev.* **D81** 082001 (*Preprint* 1003.2504)
- [18] Wen L and Schutz B F 2005 *Class.Quant.Grav.* **22** S1321–S1336 (*Preprint* gr-qc/0508042)
- [19] Iyer B *et al.* 2011 LIGO India Tech. Rep. LIGO-M1100296 <https://dcc.ligo.org/LIGO-M1100296/public>
- [20] Unnikrishnan C S 2013 *International Journal of Modern Physics D* **22** 1341010
- [21] Somiya K (KAGRA) 2012 *Class. Quant. Grav.* **29** 124007 (*Preprint* 1111.7185)
- [22] Harry I W and Fairhurst S 2011 *Phys. Rev.* **D83** 084002 (*Preprint* 1012.4939)
- [23] Klimentenko S, Vedovato G, Drago M, Mazzolo G, Mitselmakher G, Pankow C, Prodi G, Re V, Salemi F and Yakushin I 2011 *Phys. Rev. D* **83**(10) 102001 URL <http://link.aps.org/doi/10.1103/PhysRevD.83.102001>
- [24] Thorne K S 1987 *Three hundred years of gravitation* ed Hawking S W and Israel W (Cambridge: Cambridge University Press) chap 9, pp 330–458
- [25] Guersel Y and Tinto M 1989 *Phys. Rev. D* **40** 3884–3938
- [26] Keppel D 2013 (*Preprint* 1307.4158)
- [27] Kopparapu R K, Hanna C, Kalogera V, O’Shaughnessy R, Gonzalez G, Brady P R and Fairhurst S 2008 *ApJ* **675** 1459–1467
- [28] Abbott B P *et al.* (Virgo, LIGO Scientific) 2016 *Astrophys. J.* **818** L22 (*Preprint* 1602.03846)
- [29] Prix R and Krishnan B 2009 *Class. Quant. Grav.* **26** 204013 (*Preprint* 0907.2569)
- [30] Whelan J T, Prix R, Cutler C J and Willis J L 2014 *Class. Quant. Grav.* **31** 065002 (*Preprint* 1311.0065)
- [31] Abbott B P *et al.* (LIGO Scientific Collaboration, Virgo Collaboration) 2016 *Living Rev. Relat.* **19** 1 (*Preprint* 1304.0670)
- [32] Bose S, Pai A and Dhurandhar S V 2000 *Int. J. Mod. Phys.* **D9** 325–329 (*Preprint* gr-qc/0002010)
- [33] Raymond V, van der Sluys M V, Mandel I, Kalogera V, Rover C and Christensen N 2009 *Class. Quant. Grav.* **26** 114007 (*Preprint* 0812.4302)
- [34] Kasliwal M M and Nissanke S 2014 *Astrophys. J.* **789** L5 (*Preprint* 1309.1554)
- [35] Pankow C, Brady P, Ochsner E and O’Shaughnessy R 2015 *ArXiv e-prints* (*Preprint* 1502.04370)
- [36] Raymond V and Farr W 2014 (*Preprint* 1402.0053)
- [37] Abbott B P *et al.* (LIGO Scientific, VINROUGE, Las Cumbres Observatory, DLT40, Virgo, 1M2H, MASTER) 2017 *Nature* (*Preprint* 1710.05835)
- [38] Pitkin M, Messenger C and Wright L 2016 *Phys. Rev.* **D93** 062002 (*Preprint* 1511.02758)
- [39] Mills J, Tiwari V and Fairhurst S 2017 (*Preprint* 1708.00806)
- [40] Vitale S and Evans M 2017 *Phys. Rev.* **D95** 064052 (*Preprint* 1610.06917)
- [41] Blanchet L 2014 *Living Reviews in Relativity* **17** URL <http://www.livingreviews.org/lrr-2014-2>
- [42] Apostolatos T A, Cutler C, Sussman G J and Thorne K S 1994 *Phys. Rev. D* **49** 6274
- [43] Capano C, Pan Y and Buonanno A 2014 *Phys. Rev.* **D89** 102003 (*Preprint* 1311.1286)
- [44] Ajith P *et al.* 2007 *Class. Quantum Grav.* **24** S689–S699 (*Preprint* arXiv:0704.3764)
- [45] Brown D A, Harry I, Lundgren A and Nitz A H 2012 *Phys. Rev. D* **86** 084017 (*Preprint* 1207.6406)
- [46] Harry I W, Nitz A H, Brown D A, Lundgren A P, Ochsner E and Keppel D 2014 *Phys. Rev. D* **89**(2) 024010 URL <http://link.aps.org/doi/10.1103/PhysRevD.89.024010>
- [47] Brown D A, Kumar P and Nitz A H 2013 *Phys.Rev.* **D87** 082004 (*Preprint* 1211.6184)
- [48] Cornish N J and Porter E K 2007 *Class. Quant. Grav.* **24** 5729–5755 (*Preprint* gr-qc/0612091)
- [49] Jaranowski P, Krolak A and Schutz B F 1998 *Phys.Rev.* **D58** 063001 (*Preprint* gr-qc/9804014)
- [50] Babak S, Biswas R, Brady P, Brown D, Cannon K *et al.* 2013 *Phys. Rev. D* **87** 024033 (*Preprint* 1208.3491)
- [51] Dhurandhar S, Krishnan B and Willis J L 2017 (*Preprint* 1707.08163)
- [52] Williamson A R, Biwer C, Fairhurst S, Harry I W, Macdonald E, Macleod D and Predoi V 2014 *Phys. Rev.* **D90** 122004 (*Preprint* 1410.6042)