High Speed Source Localization in Searches for Gravitational Waves from Compact Object Collisions

Takuya Tsutsui, Kipp Cannon, and Leo Tsukada

(Introduction: May 19, 2020)

Multi-messenger astronomy is of great interest since the success of the electromagnetic follow up of the neutron star merger GW170817. However, the information that was learned from GW170817 was limited by the long delay in finding the optical transient. Even in the best-case scenario, the current gravitational-wave source localization method is not sufficient for some frequency bands. Therefore, one needs a more rapid localization method even if it is less accurate. Building upon an Excess power method, we describe a new localization method for compact object collisions that produces posterior probability maps in only a few hundred milliseconds. Some accuracy is lost, with the searched sky areas being approximately 10 times larger. We imagine this new technique playing a role in a hierarchical scheme were fast early location estimates are iteratively improved upon as better analyses complete on longer time scales.

I. INTRODUCTION

In August 17 2017, the advanced LIGO [1, 2] and the advanced Virgo [3] observed a gravitational wave (GW) from binary neutron star (BNS) merger, dubbed as GW170817 [4]. Then, many electromagnetic (EM) telescopes followed it to find the EM counterpart with multi-wavelength from radio wave to gamma ray [5]. By these observations, BNS merger was corroborated to be the origin of short gamma ray burst (sGRB), which had been discussed for a long time [6]. The coordinated observation by different means of astronomical signals, for example, GW and EM wave is so-called multi-messenger astronomy. GW170817 is one of the successful cases of multi-messenger astronomy. By multi-wavelength observations, information of systems is much more increased. Furthermore, the third observing run (O3) with improved sensitivity was started in 2019 and many observations with higher sensitivity have been already planned. GW observation is expected to play a more important role in physics.

So far, all compact binary coalescence (CBC) events have been detected after the merger of the two bodies. Unlike EM waves, GW is emitted during inspiral. Thus if CBC signals are sufficiently loud, one can detect them before the merger by accumulating enough signal to noise ratio (SNR) to detect, which is called early warning [7]. This could bring scientific benefits for multi-messenger astronomy because one can prepare for transient events and observe precursor events. For example, there are prompt optical flash from BNS [8], and characteristic EM emission from tidal disruption of neutron star-black hole (NSBH) before merger (precursor) [9]. There should be many unexpected events over multi-wavelength.

In an early warning context, location estimates can be iteratively improved. There are currently two stages of refinement, BAYESTAR [10, 11] which takes about 3 s, and LALInference [12] which takes hours to days. However, there is a need for a still faster location estimate even at the expense of localization accuracy. A rough location estimate available in $\mathcal{O}(100 \text{ ms})$ could trigger the slewing of fast facilities like Cherenkov telescopes, allow better data retention decisions by low-frequency radio facilities, and it might even be used to inform the ranking statistic and improve GW signal identification.

II. NOTATION

We have GW detectors, LIGO-Hanford, LIGO-Livingston, Virgo, KAGRA [14, 15] and so on. Each detector outputs time series data. From here, those are written as a vector:

$$d[j] = (d_1[j], d_2[j], \ldots, d_D[j])^T$$

where $j$ is an integer index enumerating discrete time, and $D$ is the number of detectors such that $D \geq 2$. Also corresponding antenna responses [16] with polarization angle $\psi$ for GW source direction $\Omega$ are written as a matrix:

$$F(\Omega, \psi) = \begin{pmatrix} F_{1, +}(\Omega, \psi) & \cdots & F_{D, +}(\Omega, \psi) \\ F_{1, \times}(\Omega, \psi) & \cdots & F_{D, \times}(\Omega, \psi) \end{pmatrix}^T$$ (2)
We assume that noise is uncorrelated between any pair of the detectors:

\[ \langle \hat{n}_t[k] \hat{n}_t'[k'] \rangle := \frac{1}{2} \delta_{t,t'} \delta_{k,k'} S_{n,t}[k] \tag{3} \]

where \( S_{n,t} \) is noise power spectral density (PSD) for an \( I \)-th detector. Using this, complex SNR is defined as follows,

\[ \rho_I[j] := (d_I[j + j']|h[j']) \tag{4} \]
\[ = 4 \sum_{j=0}^{N-1} \frac{\hat{d}_I[j] \hat{h}[j]}{S_{n,t}[k]} \exp \left[ \frac{2 \pi i j k}{N} \Delta f \right] \tag{5} \]

where \( \hat{h} = h_f + \text{i} h_x \) is a template, that is, a theoretical waveform normalized by \( (h[j]|h[j]) = 2 \). Fourier transformation is given by

\[ \hat{d}_a[k] := \sum_{j=0}^{N-1} d_a[j] \exp \left[ -2 \pi i \frac{k j}{N} \right] \Delta t \tag{6} \]

and

\[ d_a[j] := \sum_{k=0}^{N-1} \hat{d}_a[k] \exp \left[ 2 \pi i \frac{k j}{N} \right] \Delta f \tag{7} \]

where \( \Delta t \Delta f = 1/N \). Then, SNR PSD is defined as noise PSD of the output of the matched filter (See Appendix [A]).

Later, the data of \( I \)-th detector on time domain is shifted to represent the data at geocenter. If GW comes from \( \Omega \), the time delay on discrete time is \( \tau_I(\Omega) := r_I \cdot \Omega / (c \Delta t) \) between \( I \)-th detector at \( r_I \) and geocenter. When this delay is applied, the time shifted data on frequency domain should be written as

\[ \hat{d}[k; \Omega] = \begin{pmatrix} \hat{T}_I[k; \Omega] \hat{d}_I[k] \\ \vdots \\ \hat{T}_D[k; \Omega] \hat{d}_D[k] \end{pmatrix} \tag{10} \]

where

\[ \hat{T}_I[k; \Omega] := \exp \left[ 2 \pi i \frac{k}{N} \tau_I(\Omega) \right] \tag{11} \]

is time delay operator. For this data, time shifted SNR series are written as \( \hat{\rho}[j; \Omega] := (d_I[j + j']|h[j']) \).

Similarly, whitened SNR and whitened time shifted SNR are introduced

\[ \hat{\rho}[k] := \begin{pmatrix} \hat{\rho}_I[k] / \sqrt{S_{\rho,1}[k]} \\ \vdots \\ \hat{\rho}_D[k] / \sqrt{S_{\rho,D}[k]} \end{pmatrix} \tag{12} \]

and whitened antenna response:

\[ \hat{F}[k; \Omega, \psi] := \left( \hat{F}_+[k; \Omega, \psi], \hat{F}_\times[k; \Omega, \psi] \right) \tag{14} \]
\[ := \left( \frac{F_{1,+}(\Omega, \psi)/S_{n,1}[k] \sqrt{S_{\rho,1}[k]}, F_{1,\times}(\Omega, \psi)/S_{n,1}[k] \sqrt{S_{\rho,1}[k]}, \ldots, F_{D,+}(\Omega, \psi)/S_{n,D}[k] \sqrt{S_{\rho,D}[k]} \right)^T \tag{15} \]

Under this notation, if GW is contained in data, a whitened SNR frequency series is written as

\[ \hat{\rho}_I[k; \Omega] = \left( \hat{F}_{I,+}[k; \Omega, \psi], \hat{F}_{I,\times}[k; \Omega, \psi] \right) \left( \hat{h}_+[k] + \hat{\rho}_I[k; \Omega] \right) \hat{h}_G[k] \]
\[ + \frac{\hat{n}_I[k](\hat{h}_+[k] + \hat{h}_x[k])}{S_{n,I}[k]} \exp \left[ 2 \pi i \frac{k}{N} \tau_I(\Omega) \right] \sqrt{S_{\rho,I}} \tag{16} \]

where \( h_G = h_{GW,+} + h_{GW,\times} \) is the true GW strain in nature with \( (h_+, h_{GW,\times}) = (h_+, (h_{GW,\times}) = 0 \). We will use this series for the localization instead of the strain data.

### III. COMPACT BINARY COALESCENCE PARAMETERIZED LIKELIHOOD

Here, we assume that the source is CBC, therefore the two polarizations are related by \( \hat{h}_x = i \beta \hat{h}_+ \) where \( \beta = \frac{2 \cos \iota}{1 + \cos \iota} \) with the inclination \( \iota \). Then if the noise is Gaussian, the probability of obtaining \( \hat{\rho} \) in the presence of \( h_G \) with given parameters \( \Omega, \beta, \psi \) is

\[ p(\hat{\rho} | \Omega, \hat{h}_G, \beta, \psi) \propto \exp \left[ -2 \sum_{k=0}^{N-1} \left| \hat{\rho}_I[k; \Omega] \right|^2 \Delta f \right] \tag{17} \]

Since \( \hat{h}_G \) is not known a priori, this probability should be maximized with respect to \( \hat{h}_G \). This was solved by Sutton et al. in [17] for the case of general GWs. This probability is maximized by \( \hat{h}_+ \hat{h}_G = \left( \hat{F}_+ [k; \Omega, \psi] + \beta^2 \hat{F}_\times[k; \Omega, \psi] \right) \hat{h}_+[k] \hat{h}_G[k] \) which is in effect maximizing the probability over the distance to the source:

\[ p(\hat{\rho} | \Omega, \beta, \psi) \propto \exp \left[ \sum_{k=0}^{N-1} \left| \hat{\rho}_I[k; \Omega] \hat{P}[k; \Omega, \beta, \psi] \hat{\rho}_I[k; \Omega] \right|^2 \right] \tag{18} \]

where
FIG. 1. Schematic representation of the concept of $\hat{P}$ for a three detectors case, which is in a SNR data space spanned by those detectors. Since $\hat{P}$ is complex, SNR data space has twice dimensions of the number of detectors, although the half dimensions are omitted in the above picture. Red vector $\hat{P}$ is an observed data in the data space. Blue dashed line is the GW space where $P$ project data onto is one (see Fig. 4).

\[ \hat{P}[k; \Omega, \beta, \psi] := \frac{\left( \hat{F}_+ [k; \Omega, \psi] + i \beta \hat{F}_x [k; \Omega, \psi] \right) \otimes (\hat{F}_+ [k; \Omega, \psi] - i \beta \hat{F}_x [k; \Omega, \psi])}{|\hat{F}_+ [k; \Omega, \psi]|^2 + \beta^2 |\hat{F}_x [k; \Omega, \psi]|^2} \]  

(19)

where the projection operator $\hat{P}$ for $\beta = \pm 1$ can neglect a dependence of $\psi$, so that $\psi = 0$ is set.

Each of the exponential terms in the sum in (21) contains factors that depend on data from single detectors ($I = J$) and factors that depend on data from pairs of detectors ($I \neq J$). We call these auto- and cross-correlation terms, respectively, and the probability can be written $p(\hat{P}|\Omega) = p^{\text{auto}}(\hat{P}|\Omega)p^{\text{cross}}(\hat{P}|\Omega)$. From Appendix D $p^{\text{cross}}$ is the probability of obtaining GW signal energy $E_{\text{gw}}$ when total signal energy is $E_{\text{gw}} + E_{\text{noise}}$, and $p^{\text{auto}}$ is that of obtaining the total signal energy $E_{\text{gw}} + E_{\text{noise}}$ when the GW signal energy is $E_{\text{gw}}$ in the meaning of expectation value. We are not interested in noise signal energy so that $p^{\text{cross}}$ should be used to make the sky map[3].

\[ p^{\text{cross}}(\hat{P}|\Omega) \propto \sum_{\beta = \pm 1} \exp \left[ 4R \sum_{l \neq J \in \text{IFO}} \sum_{k=0}^{N-1} \hat{P}^*_l[k] \hat{P}_J[k] \times \hat{P}_l[k; \Omega, \beta, \psi = 0] \hat{T}^*_l[k; \Omega] \hat{T}_J[k; \Omega] \right] \]

(22)

Here, if $\hat{P}$ is replaced with $P(\Omega, \beta, \psi = 0) := \hat{P} |_{\beta = F}$, $P_l \hat{T}^*_l \hat{T}_J$ becomes independent of the SNR data, allowing it to be pre-computed for speed[4]. Following the approach presented in [18], we expand the $\rho$-independent factor in spherical harmonics $Y_{lm}$:

\[ p^{\text{cross}}(\hat{P}|\Omega) \propto \sum_{\beta = \pm 1} \exp \left[ 4R \sum_{l \neq J \in \text{IFO}} \sum_{l m} \left\{ \sum_{k=0}^{N-1} \left( P \hat{T}^*_l \hat{T}_J \right)^{l m} [k; \beta] \hat{P}^*_l[k] \hat{P}_J[k] \right\} Y_{lm}(\Omega) \right] \]

(23)

IV. REGULATOR

In the definition of the whitened SNR in (12) the ratio $\hat{P}[k]/\sqrt{S_\rho[k]}$ is not well defined for all frequency bins

\[ \hat{F}_+ \pm i \hat{F}_x \rightarrow \left( \hat{F}_+ \pm i \hat{F}_x \right) \exp(\pm \psi) \text{ is satisfied by rotating } \psi; \]

\[ \hat{F}_+ \rightarrow \hat{F}_+ \cos 2 \psi + \hat{F}_x \sin 2 \psi; \hat{F}_x \rightarrow -\hat{F}_+ \sin 2 \psi + \hat{F}_x \cos 2 \psi. \]

This phase factor is canceled in $\hat{P}$.

\[ \hat{F}_+ \rightarrow \hat{P} \rightarrow P \text{ corresponds to an assumption that all detector have same PSD because, if so, } \sqrt{S_\rho} \text{ in denominator and numerator are canceled.} \]
In particular, because inspiral templates have 0 signal energy above some high-frequency cutoff $S_\rho$ is 0 for some $k$ and the whitened SNR is undefined. In future work a more careful treatment of this problem will be presented, but at present we have found it is sufficient to regulate the instability by multiplying each term in $p^{\text{cross}}$ by $2\sqrt{S_{\rho I}[k]}\sqrt{S_{\rho J}[k]}^5$

$$p^{\text{cross}}(\tilde{\rho}|\Omega) \propto \sum_{\beta=\pm 1} \exp \left[ \Re \sum_{lm} \left\{ \sum_{I>J}^{N-1} \sum_{k=0}^N \right. \left. \left( \tilde{P}^* \tilde{T} \right)^{lm}_{IJ}[k; \beta] \tilde{\rho}_I[k] \tilde{\rho}_J[k] \right\} Y_{lm}(\Omega) \right]$$

This regulator corresponds to no whitening process (or flat SNR PSD), that is, SNR frequency series is not normalized, which acts as non-Gaussian noise.

V. POSTERIOR

In this paper, uniform prior is assumed to get posterior:

$$p(\delta, \alpha) = \frac{1}{4\pi} \cos \delta$$

(25)

Hence one gets a below posterior from Bayes’ theorem:

$$p^{\text{cross}}(\Omega|\rho) \propto p^{\text{cross}}(\rho|\Omega)p(\delta, \alpha)$$

(26)

We use this probability to produce sky maps.

VI. RESULTS & DISCUSSION

We will compare the new method with current methods, BAYESTAR [10, 11].

A. Injection test

We evaluated the performance from an injection test. The setup is below:

- TaylorT4threePointFivePN was injected into second observing run (O2) data from 1186624181s to 1187312718s in GPS time, that is, August 13-21 in 2017.
- The three detectors, LIGO-Hanford, LIGO-Livingston and Virgo were used.

The component masses are randomly sampled for $m_1, m_2 \in [1.08 M_\odot, 1.58 M_\odot]$ with the mean of 1.33$M_\odot$ and the standard deviation of 0.05$M_\odot$.
- no component spins.
- The distance was randomly sampled from a log-uniform distribution for $r \in [20 \text{ Mpc}, 200 \text{ Mpc}]$.
- Candidates were selected with satisfying:
  - Those are contained within 1 s around injected time.
  - The SNRs of more than two detectors are exceeded over 8.
  - All detectors are worked on Science mode.
  - The network SNR $\sqrt{\sum_{I \in \text{IFO}} \text{SNR}_I^2}$ is maximized in the candidates.
- 935 injections were used.

Under the above setting, complex SNR time series are generated in 0.17s around the triggered time when detecting the candidate. Fig. 2 is an example of the localization of the injections.

1. Consistency

From the above complex SNR time series, We produce skymaps and a p-p plot (Fig. 3) for the new method and
FIG. 3. p-p plot\textsuperscript{19} of the new method and BAYESTAR. Cumulative fractions of the injections are a ratio included in a \( p \) value. Gray region is error region in 95%.

TABLE I. Cherenkov Telescope Array has three size telescopes, SST, MST and LST \textsuperscript{20}.

| Name | Field of view | Target energy | Slew speed |
|------|---------------|---------------|------------|
| SST  | 8.8 deg       | \( 1 \)–\( 300 \) TeV | \( \approx 1 \) min |
| MST  | 7.5 – 7.7 deg | 80 GeV – 50 TeV | \( < 90 \) s |
| LST  | 4.5 deg       | 20 GeV – 3 TeV | \( < 20 \) s |

BAYESTAR \textsuperscript{10, 11}. From the definition of \( p \) value, the fraction of the injections with a \( p \) from the peak of maps to the injected direction should be equal to the \( p \), that is, the cumulative lines should be on the diagonal. From Fig. 3 the average of the cumulative line of the new method is on the diagonal. Then, the average of the new method is statistically consistent. Nevertheless, parts of the cumulative line are out of the 95\% error region. The origin should be from the approximation of \( \hat{P} \rightarrow P \) (see Sec. III), because both methods assumed Gaussian noise and CBC waveform, that is, the difference was from the other. That approximation is the sole one to be able to shift the peak of maps.

2. Accuracy

Fig. 4 is the area size distribution recognized as accuracy. Then, the square root of it can be recognized as the opening angle which the telescopes require. From Fig. 4, the new method is about 10 times less accurate than BAYESTAR \textsuperscript{10, 11}. Since the area size is \( \approx 65 \text{ deg}^2 \), the opening angle is \( \approx 8 \) deg. This opening angle is comparable with the field of view of the cherenkov telescope array (CTA) (see Table I), so that it is sufficiently accurate for early warning. This worse accuracy than BAYESTAR should be due to the regulator, that is, no whitening approximation (see Sec. III). Thus the new method and BAYESTAR have complementary relation with each other in terms of speed and accuracy. BAYESTAR is more robust than the new method because the new method compares the data of detectors but BAYESTAR does that with the reconstructed waveforms. Therefore BAYESTAR should have better accuracy even if all approximations are removed.

3. Computational Cost

The main advantage of the new method is its reduced computational cost and its speed. We measured the relative computational cost of this algorithm and BAYESTAR in single-threaded mode on an Intel Core i7-7600U CPU \( @ 2.80 \) GHz, and also measured the relative run-times of BAYESTAR in that configuration to a fully parallel configuration on an Intel Xeon Gold 6136 CPU \( @ 3.00 \) GHz. Taking the single-threaded run times to be dominated by arithmetic operations (I/O is not significant) then this comparison provides an estimate of the ratio of arithmetic operation count required by the two techniques to produce a location estimate. BAYESTAR is a mature code that has been optimized for the highly parallel Xeon hardware, so we also report a speed comparison of the BAYESTAR code in its production configuration.

|           | New method | BAYESTAR |
|-----------|------------|----------|
| single-threaded | 0.73 s     | 47 s     |
| parallelized  | -          | 3.3 s    |
VII. SUMMARY & FUTURE WORK

We developed a rapid localization method which is 100 times faster than BAYESTAR [10, 11] at the cost of accuracy by an order of magnitude.

Our method assumes the Gaussian noise. To estimate the direction, the new method takes into account the time delays and the amplitude ratios between SNR time series from different detectors. By maximizing or marginalizing the probability model [18] and extracting precalculated factors, the number of parameters to estimate during the calculation is reduced, which leads to speeding up the localization.

The new method has three differences from Excess power method [17] and BAYESTAR [10, 11] as follows:

1. Compared to BAYESTAR which marginalizes the posterior sky map over distance to source and source orbit inclination, the new method maximizes the posterior with respect to these two parameters. This sacrifices some accuracy in the map, but allows for some expressions to be factored into terms that depend only on data and terms that do not, which can then be pre-computed for greater speed.

2. SNR time series are used instead of strain data. By this, one can generate sky maps optimized for CBC templates, and suppress the noise contamination which is orthogonal to the template. This is the difference from Excess power method.

3. The CBC parameterization is used instead of the general parameterization used by Excess power method. By this, our target is only CBC, which is same as BAYESTAR. Then, the new method is more accurate than Excess power method, but not BAYESTAR. Also the new method can localize GW sources for more than single detector working case but Excess power method can localize for more than double detector working case.

As a potential of further improvements, the approximations applied in Chapter III are enumerated:

1. All detectors have the same PSD, that is, neglecting frequency dependence of Projection operator to correct distortions from the antenna responses and extract the GW components from data: $P[k; \Omega, \beta, \psi = 0] \to P(\Omega, \beta, \psi = 0)$.

2. The PSDs of SNR time series are flat, that is, no whitening approximation: $\hat{\rho}[k] \to \hat{\rho}[k]$.

The both approximations are meant to avoid numerical instability. Removing these approximations is future work. First one could shift the peak of maps to the correct peak because our probability should be more affected from the detector with higher sensitivities (more likely). Second one could make error region of sky maps wavy (smaller) because it makes complex phase variation fast, and our probability picks up just real part from the correlations.

VIII. ACKNOWLEDGMENTS

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes. This work was supported by the International Graduate Program for Excellence in Earth-Space Science (IGPEES). We would like to thank Heather Fong, Duncan Meacher, Cody Messick and Leo Pound Singer for teaching us how to use the analyzing software. Also, we are grateful for LIGO-Virgo’s computational resources and the data [21], because the injection data sets are produced with those.

Appendix A: Derivation of the Power Spectral Density of Signal to Noise Ratio

We derive the PSD of SNR with no GW. For uncorrelated noise,

$$\frac{1}{2} \delta_{IJ} \delta(f - f') S_{\rho I}(f) = \langle \hat{\rho}_J(f') \rangle \delta(f - f') S_{\rho J}(f')$$

(A1)

$$\int dt dt' \langle \rho_I(t) \rho_J(t') \rangle e^{-2 \pi i (f t - f' t')}$$

(A2)

$$= 4 \int dt dt' \int_\infty^{\infty} df dg \langle \hat{\rho}_I(g) \hat{\rho}_J(g') \rangle \hat{h}(g) \hat{h}^*(g') \frac{S_{n I}(g) S_{n J}(g')}{S_{n I}(g) S_{n J}(g')}
\times e^{-2 \pi i (f t - f' t')} + 2 \pi i (g t - g' t')$$

(A3)

$$= \frac{\delta_{IJ}}{2} \int_\infty^{\infty} df dg \langle \hat{h}(g) \hat{h}^*(g) \rangle \delta(f - g) \delta(f' - g)$$

(A4)

$$= \frac{\delta_{IJ}}{2} \int_\infty^{\infty} df dg \langle \hat{h}(f) \hat{h}^*(f) \rangle \frac{S_{n I}(g)}{S_{n I}(g)}$$

(A5)

$$= \frac{1}{2} \delta_{IJ} \delta(f - f') \int_\infty^{\infty} df \langle \hat{h}(f) \hat{h}^*(f) \rangle \frac{S_{n I}(f)}{S_{n I}(f)}$$

(A6)

Comparing LHS with RHS, (9) is obtained.

Appendix B: Prior of $\beta$

The probability of detecting GWs with an inclination $\iota$ should be proportional to an observable volume if the number density of CBC is uniform.

$$p(\text{detect} | \iota) \propto D_{\text{range}}^3(\iota) \propto g^3(\iota)$$

(B1)

$$g(\iota) := \left(1 + \cos^2 \iota \right)^2 + \cos^2 \iota$$

(B2)

where $D_{\text{range}}$ is the range of detectors [10, 22]. Since our universe should not have special direction, the prior is
This prior does not mean that one always observes the system from head-on or -off although $\beta = \pm 1$ ($\iota = 0, \pi$). The interpretation is discussed in Appendix C.

**Appendix C: $\beta = \pm 1$ and helicity**

For $\beta = \pm 1$, we write the projection operator (19) by introducing a new basis which is right- and left-handed one:

$$\hat{P}[k; \Omega, \beta, \psi] = \begin{cases} e^R \otimes e^{R*} & \text{for } \beta = +1 \\ e^L \otimes e^{L*} & \text{for } \beta = -1 \end{cases}$$

where $*$ is complex conjugate. Therefore the projection operators for $\beta = \pm 1$ separate data by chirality or helicity. In this notation, the probability (21) is

$$p(\hat{\beta}|\Omega) \propto \exp \left[ \sum_{k=0}^{N-1} \hat{\beta}^R[k; \Omega]\hat{\beta}^R[k; \Omega] + \hat{\beta}^L[k; \Omega]\hat{\beta}^L[k; \Omega] \right]$$

where

$$\hat{\beta}^R/L[k; \Omega] := \sum_{I \in IFO} e^{R/L*}_I[k; \Omega]\hat{\beta}_{I}[k]$$

$e^{R/L}_I$ is the $I$-th component of the vector $e^{R/L}$. This marginalization can be understood from the physical picture which one takes into account those polarizations. $\beta = +1/-1$ corresponds to $\iota = 0/\pi$ respectively. Nevertheless, it does not mean that there are observers on the azimuthal axis of the CBC system. Since the distance to the CBC and the inclination are degenerate, our probability model has lost the information of the inclination when the probability was extremized with respect to $h_{GW}$. Thus, $\beta = \pm 1$ is just a label of the helicity. One does not have to try to read out anything from the inclination.

**Appendix D: Relation between correlations and signal energy**

SNR is an amplitude ratio between GW and noise. GW signal energy $E_{gw}$ is proportional to the squared amplitude. Thus, introducing a conceptual noise signal energy $E_{\text{noise}}$ from the dimensional analysis, one can recognize SNR as a square root of the signal energy ratio [10].

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig5.png}
\caption{This is a probability of $\beta$ for detected events, that is $\beta$ vs. $p(\beta|\text{detect})$.}
\end{figure}
The expectation values of the correlations with GW are:

\[
\langle |\hat{\rho}_I|^2[k] \rangle = \frac{\hat{h}_{GW}[k]\hat{h}_{GW}^*[k]}{S_n,k[k]} + \frac{N}{2} \quad (D1)
\]

\[
\propto E_{gw} + E_{noise} \quad (D2)
\]

\[
\langle \hat{\rho}_I[k]\hat{\rho}_J[k] \rangle = \frac{\hat{h}_{GW}[k]\hat{h}_{GW}[k]}{\sqrt{S_n,k[k]S_n,j[j]}} \quad (D3)
\]

where \(N\) is the number of bins and the SNR is whitened. Therefore the auto-correlations correspond to the total signal energy and the cross-correlations to the GW signal energy with respect to the expectation values. Then, \(p^{cross}\) is the conditional probability of observing the GW signal energy \(E_{gw}\) when the total signal energy is \(E_{gw} + E_{noise}\):

\[
p^{cross}(\rho(\Omega)) = p(\rho(\Omega))/p(\rho(\Omega) =: p(E_{gw}(\Omega, E = E_{gw} + E_{noise})).
\]

1. J Aasi, B P Abbott, R Abbott, T Abbott, M R Abernathy, K Ackley, C Adams, T Adams, P Addesso, and et al. Advanced LIGO: Classical and Quantum Gravity, 32(7):074001, Mar 2015.
2. Gregory M Harry. Advanced LIGO: the next generation of gravitational wave detectors. Classical and Quantum Gravity, 27(8):084006, apr 2010.
3. F Acernese, M Agathos, K Agatssuma, D Aisa, N Almandou, A Alocco, J Amarni, P Astone, G Balestri, G Ballardini, and et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. Classical and Quantum Gravity, 32(2):024001, Dec 2014.
4. B.P. Abbott, R. Abbott, T.D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R.X. Adhikari, V.B. Adya, and et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. Physical Review Letters, 119(16), Oct 2017.
5. B. P. Abbott et al. Multi-messenger Observations of a Binary Neutron Star Merger. Astrophys. J., 848(2):L12, 2017.
6. William H. Lee, Enrico Ramirez-Ruiz, and Jonathan Granot. A Compact Binary Merger Model for the Short, Hard GRB 050509b. The Astrophysical Journal, 630(2):L165–L168, Aug 2005.
7. Kipp Cannon, Romain Cariou, Adrian Chapman, Minea Crispin-Ortuza, Nickolas Fotopoulos, Melissa Frei, Chad Hanna, Erin Kara, Drew Keppel, Laura Liao, Stephen Poprocki, Jameson Terraner, and Alan Weinstein. TOWARD EARLY-WARNING DETECTION OF GRAVITATIONAL WAVES FROM COMPACT BINARY COALESCENCE. The Astrophysical Journal, 748(2):136, Mar 2012.
8. Ehud Nakar. Short-hard gamma-ray bursts. Physics Reports, 442(1):166 – 236, 2007. The Hans Bethe Centennial Volume 1906-2006.
9. Sean T. McWilliams and Janna Levin. ELECTROMAGNETIC EXTRACTION OF ENERGY FROM BLACKHOLE–NEUTRON-STAR BINARIES. The Astrophysical Journal, 742(2):90, Nov 2011.
10. Leo P. Singer and Larry R. Price. Rapid Bayesian position reconstruction for gravitational-wave transients. Phys. Rev. D, 93:024013, Jan 2016.
11. Leo P. Singer, Hsin-Yu Chen, Daniel E. Holz, Will M. Farr, Larry R. Price, Vivien Raymond, S. Bradley Cenko, Neil Gehrels, John Cannizzo, Mansi M. Kasliwal, Samaya Nissanke, Michael Coughlin, Ben Farr, Alex L. Urban, Salvatore Vitale, John Veitch, Philip Graff, Christopher P. L. Berry, Satya Mohapatra, and Ilya Mandel. GOING THE DISTANCE: MAPPING HOST GALAXIES OF LIGO AND VIRGO SOURCES IN THREE DIMENSIONS USING LOCAL COSMOGRAPHY AND TARGETED FOLLOW-UP. The Astrophysical Journal, 820(1):L15, sep 2016.
12. J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, and et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. Physical Review D, 91(4), Feb 2015.
13. Hsin-Yu Chen and Daniel E. Holz. Facilitating Follow-up of LIGO–Virgo Events Using Rapid Sky Localization. The Astrophysical Journal, 840(2):88, May 2017.
14. Kentaro Somiya. Detector configuration of KAGRA—the Japanese cryogenic gravitational-wave detector. Classical and Quantum Gravity, 29(12):124007, Jun 2012.
15. Yoichi Aso, Yuta Michimura, Kentaro Somiya, Masaki Ando, Osamu Miyakawa, Takanori Sekiguchi, Daiuise Tatsumi, and Hiroaki Yamamoto. Interferometer design of the KAGRA gravitational wave detector. Phys. Rev. D, 88:043007, Aug 2013.
16. M. Maggiore. Gravitational Waves: Volume 1: Theory and Experiments. Gravitational Waves. OUP Oxford, 2008.
17. Patrick J Sutton, Gareth Jones, Shourov Chatterji, Peter Kalmus, Isabel Leonor, Stephen Poprocki, Jameson Rollins, Antony Searle, Leo Stein, Massimo Tinto, and Michal Was. X-Pipeline: an analysis package for autonomous gravitational-wave burst searches. New Journal of Physics, 12(5):053034, may 2010.
18. Kipp C. Cannon. Efficient algorithm for computing the time-resolved full-sky cross power in an interferometer with omnidirectional elements. Phys. Rev. D, 75:123003, Jun 2007.
19. ligo.skymap. https://lscsoft.docs.ligo.org/ligo-skymap/index.html
20. I Bartos, T Di Girolamo, J R Gair, M Hendry, I S Heng, T B Humensky, S Märka, Z Märka, C Messen-ger, R Mukherjee, and et al. Strategies for the follow-up of gravitational wave transients with the Cherenkov Telescope Array. Monthly Notices of the Royal Astronomical Society, 477(1):639–647, Mar 2018.
21. The LIGO Scientific Collaboration, the Virgo Collaboration, and et al. Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo, 2019.
22. Jolien D. E. (Jolien Donald Earl) Creighton and Warren G. Anderson. Gravitational-wave physics and astronomy : an introduction to theory, experiment and data
analysis. Wiley-VCH, 2011.