Study of the fast localization of coalescing binaries with a hierarchical network of gravitational wave detectors

April 10, 2017

National Astronomical Observatory of Japan

Department of Astronomy
Graduate School of Science
The University of Tokyo

Yoshinori Fujii

Table of Contents

Overview

This part describes the expected fast sky localization of coalescing binaries with a hierarchical network of three gravitational wave (GW) detectors, HLV (Hanford/Livingston/Virgo), and four detectors, HLVK (Hanford/Livingston/Virgo/KAGRA). A hierarchical network is used for detection of objects with different sensitivity GW detectors, and aims to make effective use of the least sensitive detector's information. In this method, the less sensitive detectors participate only when a trigger is found with the more sensitive detectors' coincidences. However, the signal to noise ratio (SNR) thresholds of the less sensitive detectors are set lower than that of the higher sensitivity ones. The main target of this work is to estimate the benefit of this detection network with 3 or 4 different sensitivity detectors.

In this part, the localization with two higher sensitivity LIGO detectors and the less sensitive Virgo detector and KAGRA detector, is demonstrated by using the previous Mock Data Challenge (MDC) result. In the demonstration, it supposes that this method is implemented into a pipeline for GW and Electromagnetic-wave (EM) follow-up observation. The target GW sources of the pipeline is coalescing compact binary systems. The performances of the localization are investigated in terms of sky maps generated from a pipeline for detection of GWs from compact binary coalescences (CBCs), called Multi-Band Template Analysis (MBTA), and a sky localization pipeline, called BAYESian Tri-Angulation and Rapid localization (BAYESTAR). This part presents the expected localization with HLV and HLVK hierarchical search, and the optimization of threshold SNR for the Virgo and KAGRA detectors.

Source localization and a hierarchical network

Chapter 1

This chapter describes the background on low-latency analysis in GW detection, and motivation of the hierarchical search.

1.1 Background on GW data analysis

Some of the background quantities, which are used in the field of the GW data analysis, will be explained in this chapter.

1.1.1 Antenna function

The amplitude of a GW $h(t)$ incidents into a detector is given by

$$h(t) = F_+(\theta,\phi,\psi)h_+(t) + F_\times(\theta,\phi,\psi)h_\times(t), \tag{1.1}$$

where $F_+,\ F_\times$ are the detector response functions, called antenna function, and $h_+(t),\ h_\times(t)$ are the two polarization of the GW. The antenna functions depend on the sky location $(\theta,\ \phi)$, and also the polarization of the source ψ, relative to the detector. The antenna functions are the parameters that depend on the geometric of the detector and the source. In a coordinate system with the x, y axes aligned with the detector-arm, its antenna functions are given by

$$F_+(\theta,\phi,\psi) = \frac{1}{2}(1+\cos^2\theta)\cos 2\phi\cos 2\psi - \cos\theta\sin 2\phi\sin 2\psi, \tag{1.2}$$

$$F_\times(\theta,\phi,\psi) = \frac{1}{2}(1+\cos^2\theta)\cos 2\phi\sin 2\psi + \cos\theta\sin 2\phi\cos 2\psi. \tag{1.3}$$

The antenna functions $F_+,\ F_\times$ describe how the two GW polarizations are combined at the detector sensing. More details are explained in [7].

1.1.2 Effective distance, horizontal distance and detection range

Effective distance

The effective distance of a source inducing a given signal at the detection output is the distance of an optimally located and oriented source that would produce the same signal strength. Thence, the effective distance denotes the distance to the source when the source is located at just above the detector. Here, substitute a binary inspiral waveform obtained from the post-Newtonian approximation into the equation(1.1). If one focuses on either $h_+(t)$ or $h_\times(t)$ term, the strain amplitude $h(t)$ can be described as

$$\begin{aligned} h(t) &= \frac{D(t)}{D_{\rm eff}}\cos\left(\Phi(t)+\Phi_0\right) \\ &= \frac{1}{D_{\rm eff}}\left(D_c(t)\cos\Phi_0 + D_s(t)\sin\Phi_0\right), \end{aligned} \tag{1.4}$$

where $D(t)$ is a certain distance, $D_c,\ D_s$ are the cosine and sine part of the waveform at some reference distance, $D_{\rm eff}$ is known as the effective distance to the source, and Φ_0 is the coalescence phase as observed at the detector. D_c and D_s have information of the intrinsic parameters, such as the chirp mass and spin of the binary. Both of the effective distance and the coalescence phase, depends on the location and orientation of the binary respect to the detector. Especially, the effective distance $D_{\rm eff}$ is defined as

$$D_{\rm eff} = \frac{R}{\sqrt{\frac{1}{4}{F_+}^2(1+\cos^2\iota)^2 + {F_\times}^2\cos^2\iota}}, \tag{1.5}$$

where R is the actual physical distance to the source, ι is the inclination angle of the source.

Horizontal distance

The horizontal distance $R_{\rm H}$ is the distance at which an optimally located and oriented source would produce a certain SNR ρ_0 in a detector. Usually, the ρ_0 is set at 8 [8]. The horizontal distance is defined by using sensitivity of a detector, as follows. First, the output from a detector $s(t)$ is described as

$$s(t) \;=\; n(t) \;+\; h(t), \tag{1.6}$$

where $n(t)$ is the instrumental and environmental noise, $h(t)$ is the GW signal which can be present or absent. Here, the ensemble average of the Fourier components of the noise satisfies

$$\langle \tilde{n}(f)\tilde{n^*}(f')\rangle \;=\; \frac{1}{2}\delta(f-f')S(f), \tag{1.7}$$

where angle $\langle ...\rangle$ denotes an ensemble average. The one-sided power spectra density of the detector noise $S(f)$ is defined by eq (1.7). $S(f)$ has dimension of Hz^{-1}, and satisfies $S(-f) = S(f)$. The factor of 1/2 is inserted in the definition so that the total noise power is calculated by integrating over the range $0 \le f < \infty$:

$$\langle n^2(t)\rangle \;=\; \int_0^\infty df S(f). \tag{1.8}$$

Then, if an inner product is defined as

$$(a|b) = 4Re\left[\int_0^\infty df\; \frac{\tilde{a}(f)\tilde{b^*}(f)}{S(|f|)}\right], \tag{1.9}$$

the SNR of the detector ρ for a given signal h is calculated by

$$\rho^2(t) \;=\; (h|h)\,. \tag{1.10}$$

The corresponding detector sensitivity σ in dimension of length is expressed by

$$\sigma^2 \;= \rho^2 \times {D_{\rm eff}}^2. \tag{1.11}$$

Thence, the signal amplitude ρ is proportional to the ${D_{\rm eff}}^{-1}$. Note that σ can be also expressed by $\sigma^2 = (D_c|D_c)$ according to eq (1.4). Then, the horizontal distance, which is described as the effective distance at the SNR $\rho_0 = 8$, is described by

$$R_{\rm H} \;=\; D_{\rm eff}(\rho=\rho_0) \;=\; \frac{\sigma}{\rho_0} \;=\; \frac{\sigma}{8}. \tag{1.12}$$

Detection range

The detection range $R_{\rm det}$ is the spherical radius, whose volume is equal to the volume surrounded by detection radius $r_{\rm det}$. The detection radius $r_{\rm det}$ is the distance that the detector can sense the signals in each direction for an orientation of the source. The detection radius $r_{\rm det}$ is defined as

$$r_{\rm det} = R_{\rm H}\sqrt{\frac{1}{4}{F_+}^2(1+\cos^2\iota)^2 + {F_\times}^2\cos^2\iota}, \tag{1.13}$$

where $R_{\rm H}$ is the horizontal distance. By using the detection radius $r_{\rm det}$, the detection range $R_{\rm det}$ is obtained as follows.

$$\begin{aligned} R_{\rm det}(\iota,\psi) &= \left(\frac{3}{4\pi}V(\iota,\psi)\right)^{1/3} \\ &= \left(\frac{3}{4\pi}\int_0^{r_{\rm det}(\theta,\phi,\iota,\psi)} dr \int_0^{\pi} d\theta \int_0^{2\pi} d\phi\ r^2\sin\theta\right)^{1/3} \\ &= \left(\frac{1}{4\pi}\int_0^{\pi} d\theta \int_0^{2\pi} d\phi\ \left(r_{\rm det}(\theta,\phi,\iota,\psi)\right)^3\sin\theta\right)^{1/3}. \end{aligned} \tag{1.14}$$

Due to the directional sensitivity or antenna pattern of the detectors, the detection range $R_{\rm det}$ is smaller than the horizontal distance $R_{\rm H}$ by a factor of 2.26, for the same SNR threshold[1],[2]. Then, the relation is given by

$$\frac{R_{\rm H}}{R_{\rm det}} = 2.26. \tag{1.15}$$

Expected SNR

The SNR of a detector is deduced from the detection range and the effective distance. By combining (1.11), (1.12), (1.15), the SNR ρ is given by

$$\rho = \frac{\sigma}{D_{\rm eff}} = \frac{\rho_0\ R_{\rm H}}{D_{\rm eff}} = \frac{\rho_0\ \times\ 2.26\ \times R_{\rm det}}{D_{\rm eff}} = \frac{8\ \times\ 2.26\ \times R_{\rm det}}{D_{\rm eff}}. \tag{1.16}$$

1.1.3 Matched filtering

Matched filtering is the analysis method used in searching GW signals from compact binary coalescences. The matched filter compares the measured signal with the theoretical signal by taking into account of the detector noise, and outputs the SNR $\rho(t)$. Even though the theoretical GW signals from CBC are well known, the wave forms depend on the masses and the spins of the binaries. Thus, the matched filtering needs theoretical wave forms corresponding to all the possible masses and the spins. These waveforms are called templates. The output SNR $\rho(t)$ is given by [1]:

$$\rho(t) = \int_{f_{\rm low}}^{f_{\rm high}} \frac{Measure(f)\times Template^*(f)}{Noise(f)}\,df \tag{1.17}$$

$$= 4Re\left[\int_{f_{\rm low}}^{f_{\rm high}} \frac{\tilde{s}(f)\ \tilde{h}^*_{\rm template}(f)}{S(f)}\ \exp(2\pi i f t)\ df\right], \tag{1.18}$$

where $f_{\rm low}$ and $f_{\rm high}$ are the low and high frequency cut offs. The use of a Fast Fourier Transformation (FFT) allows extraction of the signals for all possible arrival times. Then, the sets which give higher SNR than a threshold SNR, are to be searched. The maximum SNR above the threshold is called a trigger, and the corresponding event is considered as a candidate event. The generated triggers are also used to make the detection network by several detectors.

1.1.4 Source localization

When a GW event is detected by several GW detectors, the source position can be obtained from the detection time lags of each detector, by the triangulation. The principle is describes as

$$\theta = \arccos\left(\frac{c}{d}\,\Delta t\right), \tag{1.19}$$

where θ is an incident angle of the GW, c is speed of light, d is distance of detectors, and Δt is detection time lag between the detectors. Figure 1.1 describes the principle of the triangulation.

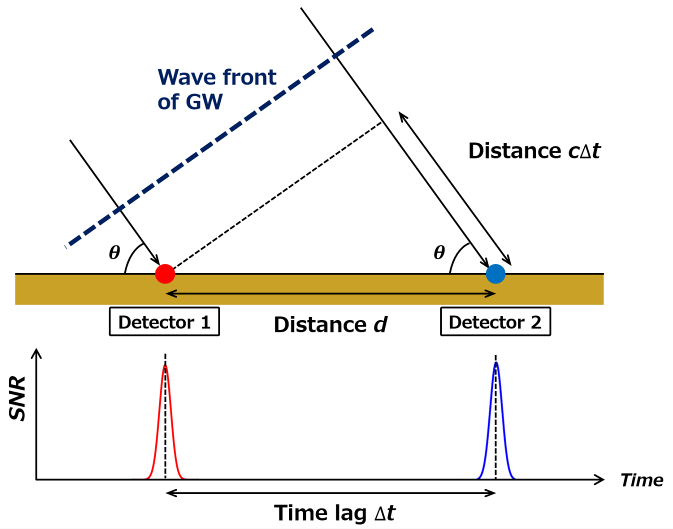

Figure 1.1: Principle of triangulation.

1.2 Hierarchical network search

In order to enhance preciseness of rapid source localization by GW detectors, several GW detectors should be operated together, and thus observation with several GW detectors is important for opening GW astronomy. However, the sensitivities of those detectors, which are operated together, can be different from each other, even if they are operated at the same time. For instance, just after completion of construction of a detector, its sensitivities is expected to be lower than its target sensitivity, and also lower than the already operating detectors' sensitivities. Actually after the LIGO observation O2[1], operation of the newly made Virgo detector should be started, and it is expected that the sensitivity of the Virgo will be lower than the sensitivities of the LIGO Hanford (H1) and LIGO Livingstone (L1) detector[2]. In such situation, the less sensitive detectors can still contribute to the improvement of the GW source localization.

Here, for example, consider an observation by three GW detectors and the two of them have higher sensitivity. In this situation, if one set a same detection SNR threshold for all the detectors, it is expected that the chances to detect an GW event in all the three detectors are rare. This is because the GW signals can be easily buried into the noise in lower sensitive detectors, compared to that in higher sensitivity ones, and thus lower sensitivity detector generates fewer triggers which are based on

[1] O2 is a scientific operation by the two LIGO detectors which is started from end of 2016.

[2] The sensitivity of the V1 is expected at around 20 Mpc, while for the two LIGOs are expected at around 70 Mpc for 1.4 $M_\odot$ binary neutron star.

GW signals than the higher sensitivity ones. Consequently, even if the two higher sensitivity detectors are able to detect an event, the lower sensitivity one might miss to detect it. Thence, in order to use the information of less sensitive detector more, one wants to set a lower SNR threshold for the less sensitive one. However, if one lowers the threshold, one has to handle a lot of triggers based on background noise, and the large numbers of those triggers can easily increase the computational and time cost. This situation is not suitable for low-latency search. Thence, one has to lower the SNR threshold for less sensitive detector so that too may background triggers are not generated. The hierarchical search is one method which realizes this aim.

In the hierarchical search, less sensitive detectors are included into the network with a lower SNR threshold than the SNR of higher sensitive detectors, only when a trigger, which is generated from higher sensitive detector's coincidences, is searched. By using this method, one can set the threshold lower for the less sensitive detectors without handling a lot of the triggers. This searching method would increase chances which less sensitive detectors are used as the network detectors, and also increase chances to have more precise localization. It is thus expected that opportunities of GW and electromagnetic (EM) wave follow-up observation would be increased by using this hierarchical searching method. This is why the hierarchical search would be useful for including detectors which are newly constructed and have not achieved their target sensitivities yet. The hierarchical analysis is expected to be one approach to include low sensitive detectors into the network.

1.3 Research target and outline

The goal of this work is to study GW source localization with a hierarchical network of 3 or 4 detectors, when the search is implemented into a low latency searching pipeline for GWs from CBC. In particular, there are two targets. First is to optimize the SNR threshold for less sensitive detectors, and the second is to get the expected localization at the threshold. In this work, hierarchical network with 3 or 4 detectors is assumed to be implemented into a GW-EM follow up pipeline. In order to study the localization of the hierarchical search, previous mock data challenge (MDC) results are used.

The outlines of this report is as follows. Chapter 2 explains the elements of the GW-EM follow up pipeline, in which the hierarchical network is assumed to be implemented. Chapter 3 explains the localization quantities and the set up for this work. Then, in chapters 4 and 5, the optimization of the SNR threshold for the low sensitivity detectors in the hierarchical search with three or four detectors is described. Chapter 6 summarizes the expected fast localization performance with the hierarchical network search.

GW-EM follow up pipeline

Chapter 2

This chapter highlights the main elements of the GW-EM follow up pipeline. This pipeline is used for the detection of gravitational waves (GWs) from compact binary coalescences (CBCs). The signals output from GW detectors will be processed in the pipeline as follows. First, the signals enter the low latency coincident analysis pipeline called Multi-Band Template Analysis (MBTA)[3]. Then, GW candidate events detected by the MBTA are uploaded onto the gravitational wave candidate event database (GraceDb). Finally, the uploaded events are analyzed by a Bayesian triAngulation and rapid localization, called BAYESTAR[4], whereby the probability sky maps for the sky localization of GW candidate events is generated. Events which pass validation process by human monitors are distributed to astronomical partners for EM follow up. Figure 2.1 shows the conceptual diagram of the GW-EM follow up pipeline. Chapter 2.1, 2.2 describes the key elements of the MBTA and BAYESTAR pipeline, respectively.

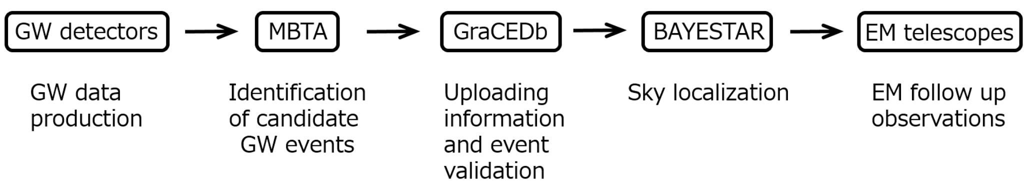

Figure 2.1: GW-EM pipeline.

2.1 Multi-Band Template Analysis : MBTA

The MBTA is a low latency, low computational coincidence analysis pipeline used for detecting GWs from CBCs. In the MBTA, the standard matched filter is implemented to extract CBC signals from GW channel data of each detector in the network independently, before the outputs are combined to find GW candidate events. MBTA is mainly used for the on-line detection of GW candidate events with sub-minute latency, and also for the data quality studies due to the low computational cost. MBTA determines significance of GW candidate events by calculating the false alarm rate (FAR) using data just before the event, to evaluate the detector background noise at the time of the event.

This chapter describes the main elements of MBTA as follows. Chapter 2.1.1 describes what is performed in single detector analysis, before information from each detector is combined. Chapter 2.1.2 highlights the functionality to identify coincidences event.

2.1.1 Single detector analysis

The main functionalities in the part of single detector analysis of MBTA is Multi-band matched filtering, χ^2 cut, Matched filter output shape cut, as follows.

Multi-band matched filter

In the MBTA, the matched filter is split into two or more frequency bands. The boundary frequency f_c between the low and high frequency bands is selected so that the SNR is roughly equal between both of the frequency bands. The typical f_c is set at around 100 Hz. This multi-band method allows a computational reduction, without losing SNR on average compared to a matched filter performed with a single band analysis. This computational reduction is obtained by using shorter templates in each frequency band. Since the phase of the GW signal is tracked over fewer cycles, the multi-band method reduces the number of templates which are required to cover the equivalent parameter space of a single band analysis. Another merit of the multi-band method is that a reduced sampling rate can be set for the low frequency band. Due to the sampling rate reduction, the cost of the Fast Fourier Transforms (FFTs) related with the filtering is reduced.

χ^2 cut

The MBTA has a functionality of χ^2 cut, which checks that the repartition of the SNR of a trigger between the low and high frequency band is consistent with what is expected for a true signal. The χ^2 cut is set at $\chi^2 < (A + B \times \mathrm{SNR}^2)$, for the two band analysis. These values are empirically set at $A = 6$, $B = 0.075$ based on measurements.

Matched filter output shape cut

To reject noise events more efficiently, another signal based consistency test is included. This function is called matched filter output shape cut (MFO shape cut). The MFO shape cut is based on the fact that the SNR time series of the MFO for a GW signal will have one steep peak, while for noise the MFO will have a broader peak with multiple maximums around the central structure. First, the MFO shape cut calculates a ratio R between integrated SNR in the 100 ms surrounding a trigger excluding a small central window of 7.5 ms, and the integrated SNR inside the central window. Then, triggers which do not behave as GW signals, are rejected according to the following formula,

$$R \;=\; \frac{\text{surrounding around central window}}{\text{central window}} \;>\; \frac{A}{\mathrm{SNR}^2} + B, \tag{2.1}$$

where typically $A = 65$, $B = 0.4$. They are the empirical values based on measurements.

2.1.2 Identification of coincidences event

The list of the single detector triggers are combined to find coincidence events among GW detectors. The false alarm rate (FAR) and the combined SNR are computed. MBTA identifies the coincident events by using following time coincidence test and the exact match coincidence test.

Time coincidence test

This test checks the time coincidence of the events between the detectors, considering not only GW time flight between the detectors but also the experimental uncertainty in arrival timing measurement.

Exact match coincidence test

This test requires for the events to be found in all detectors with the same template parameters. The tested template parameters are the component masses and the spins.

2.1.3 Upload to GraceDb

Low FAR events generated by MBTA, which pass a given threshold FAR are uploaded onto the GraceDb. The typical threshold is 1.0×10^{-4} Hz. The submitted information includes the GW arrival times, the arrival timing, SNR amplitudes, the GW phase at coalescence, and the noise curve for each detector at the time of detection of the event.

2.1.4 Template region

Templates of MBTA are divided into three parts on the parameter space depending on masses of binary components: the region 1, the region 2, and the region 3. The region 1 has templates calculated by using low mass binaries such as neutron star binaries. Templates of the region 3 are generated by using high mass binaries composed by such as black holes. The region 2 includes templates obtained from middle mass binaries. The number of templates for the region 1, th region 2, and the region 3 are set at 2×10^5, 4414, and 25604 in this research. Basically, trigger collecting time window is set shorter down to the region 1, whereas, the time window is longer up to the region 3. The region 2 is analyzed by the middle time length.

2.2 BAYESian TriAngulation and Rapid localization : BAYESTAR

The BAYESTAR is a rapid Bayesian sky localization method that takes a few minutes however achieves nearly the same accuracy as a full parameter estimation such as the LALINFERENCE stochastic samplers [5],[6]. Using the information of the GW candidate event, BAYESTAR generates a probability sky map. The main feature of BAYESTAR are highlighted in section 2.2.1, while the section 2.2.2 describes the effect of the arrival time of each detector.

2.2.1 Probability sky localization

BAYESTAR uses the Bayesian approach to make use of the GW observations and prior distributions across the source parameters. It also considers the phase consistency across all detectors in the network. The important characteristic of BAYESTAR is to treat the matched filter detection pipeline, such as MBTA, as a measurement system, treating the parameter estimates that it provides as the experimental input data. It does not use the full GW time series data as in the case of the rigorous full parameter estimated analysis. This approach not only reduces the dimensionality of the data and the signal model, but also allows to avoid directly computing the post-Newtonian model waveforms, making the likelihood much faster to evaluate.

The BAYESTAR likelihood function depends only on trigger information, namely, the time, the phase and the SNR amplitude of the GW event at each detector. BAYESTAR uses the leading order independences of errors in the extrinsic and intrinsic parameters[1] by keeping the masses fixed at the values estimated by the detection pipeline. The simplified probability distribution for the sky positions of sources are generated by numerically integrating the probability in each pixel of the sky map.

All detected GW candidate events are followed up with sky localization probability by BAYESTAR. An example of the sky map generated by BAYESTAR is shown in figure 2.2.

[1] A compact binary coalescence (CBC) source is characterized by a vector of extrinsic and intrinsic parameters. The extrinsic parameters describe the source position and orientation, while the intrinsic parameters describe the physical properties of the binary components. For instance, right ascension, declination, distance, arrival time at geocenter, inclination angle, and coalescence phase are regarded as the extrinsic parameters. On the other hand, each component's mass and spin are regarded as the intrinsic parameters in some simplifying assumptions.

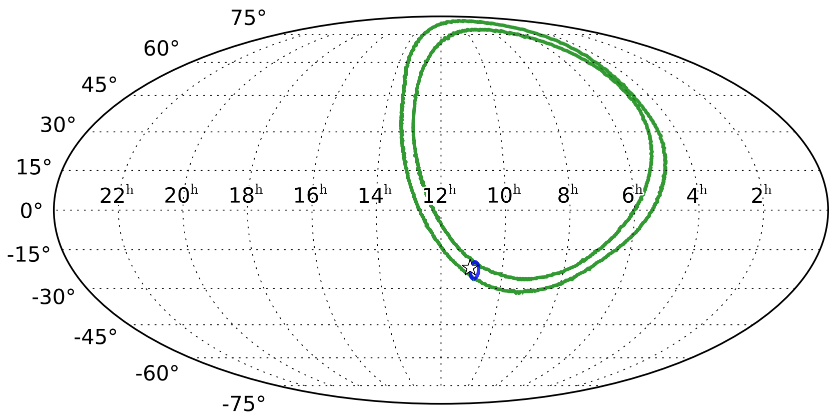

Figure 2.2: Sky map probability generated by BAYESTAR for a simulated GW event detected by MBTA with H1-L1 network (green), H1-L1-V1 network (blue). Here H1, L1, and V1 denote LIGO Hanford, LIGO Livingston, and Virgo detectors respectively. Astro-hours-Mollweide projection is used as geographic coordinates in this plot. The star shows the position of the injected GW signal and the projected area describes the 90 % confidence area.

2.2.2 Effect of arrival timing to sky map

The BAYESTAR handles two kinds of timing errors. First one is the timing error due to the noise curves, and the second is the timing error depending on the arrival timing among the detectors. The first timing error is used to calculate the resolution of the sky map. The second is for pointing the location of the predicted position. Figure 2.3 shows how the BAYESTAR uses the two kinds of timing errors. The blue and red circled sky maps are obtained from triple coincidence search with H1, L1, and V1. The blue map is calculated with using only the timing error from the noise curve. According to the map, the error from noise curve gives only the resolution of the predicted map. On the other hand, the red map is from both of the timing errors, so that, the error from the noise curve and the arrival timing error are considered. In this case, the prediction position is shifted while keeping its resolution for the most part.

Consequently, the timing error calculated from the noise curve is used to obtain only the uncertainty of the maximum probability. On the other hand, the timing error depending on the arrival timing among the detectors is used for pointing the location of the predicted position.

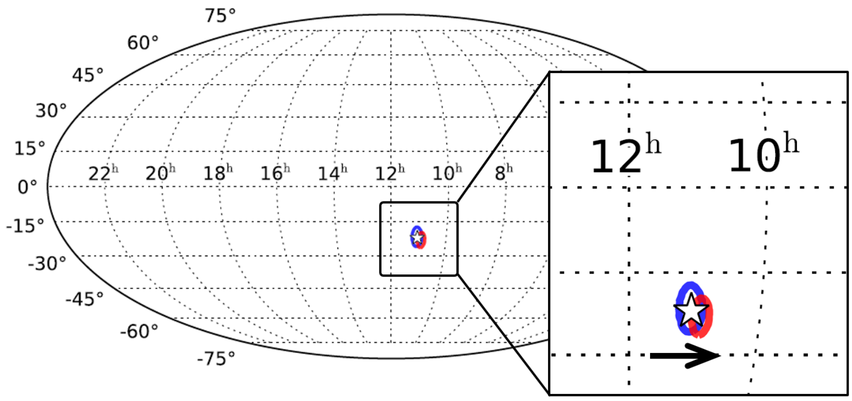

Figure 2.3: Effect of noise curve on its prediction. Blue circled map is generated with using only timing error based on noise curve. Red map is from both of the timing error based on the noise curve and arrival timing.

Calculation setup

Chapter 3

This chapter explains the quantities which measure the localization performance, and how the simulation is conducted. The main flow of this work is shown in figure 3.1. The target is to get the expected sky localization performance with the hierarchical network of three or four detectors, when this method is implemented into GW-EM follow up pipeline. This work focuses on a condition that the two LIGO detectors have higher sensitivity than the Virgo and the KAGRA detector, and investigates the measures of the localization, which are generated from detector coincidence analysis through the MBTA and the BAYESTAR.

Necessary information for this investigation is SNR, arrival time, phase of gravitational waves, and noise power spectral density (PSD) of each detector as the inputs for the BAYESTAR. The information for the two LIGO detectors are obtained from previous MDC result, while the information for the less sensitive detectors are produced artificially.

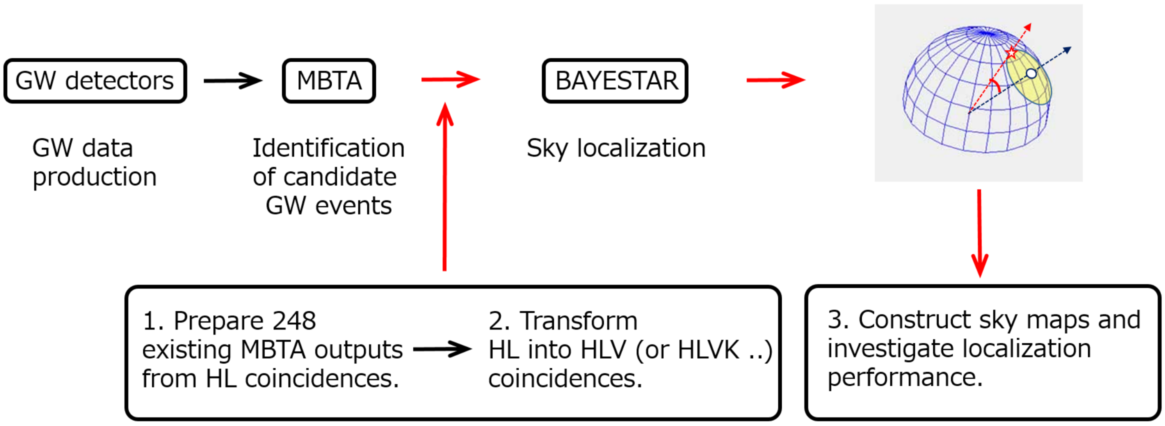

Figure 3.1: Main flow of this work.

This chapter describes the sky localization quantities and how these quantities are generated in this work, including how the information of the low sensitivity detectors is produced. Chapter 3.1 explains sky localization quantities to be investigated. chapter 3.2, 3.3, 3.4 describe the used injection data set, how the information of SNR is generated, and which noise curve is used in this work, respectively. Then, chapter 3.5 summarizes the expected localization performance of the two LIGO detectors search. Finally, chapter 3.6 describes how the lower sensitivity detector's information is produced.

3.1 Sky localization quantities

The quantities which describe the Sky localization performances are summarized in this chapter. Two quantities, called the offset angle and the searched area, are especially used to determine the optimal thresholds. This is because both the offset angle and the searched area describe how far the localization is from the true injected position.

3.1.1 Offset angle

The offset angle is the angle between the sky location of the injected GW signal and the reconstructed pixel which has the maximum probability. This is also shown in the schematic diagram in figure 3.2.

3.1.2 Searched area

The searched area is the smallest area of the highest confidence region around the pixel of maximum probability, that includes the sky location of the injected GW signal. For instance, if the injected position is at a pixel of probability p, the searched area is calculated from the sum of all the pixels, which have larger probabilities than p. According to this definition, the searched area should have a similar behavior as the offset angles. The schematic diagram is also in figure 3.2.

3.1.3 Confidence area

The confidence area describes how a sky map is spread or concentrated. This area is calculated as follows. First, pixels are ranked by descending posterior probability. Then, the cumulative sum of the probabilities of the pixels is computed in order. Finally, one obtains the confidence area by finding the index of the pixels of which the cumulative sum is equal to a given value. For example, in order to find the 90 % confidence area, this value is set 0.9.

3.1.4 Probability-Probability plot (P-P plot)

The Probability-Probability plot is used to check the self-consistency of the Sky localization performances. As described above, the searched area explains the accuracy of the sky map, whereas a certain confidence area shows the precision. However, these two quantities should be correlated. For instance, 90 % of events should have their true locations contained within their respective 90 % confidence area, on average. More generally, if we make a cumulative histogram of the confidence levels corresponding to the searched areas of all the events, it should be found close to a diagonal line in the plot, with small deviations due to finite sample size. The plot is called Probability-Probability plot (P-P plot). Particularly, the horizontal axis corresponds to certain confidence area, while the vertical axis is the function of the injections whose true locations are within the corresponding confidence areas. This self-consistency test is necessary but not sufficient condition for the validity of any Bayesian parameter estimation scheme.

Basically, if the curve is far below the diagonal line, the obtained statistics are interpreted as optimistic result, according to the sky localization. This is because, for example, even if we search for 90 % confidence area, we cannot find 90 % of the injections. On the other hand, if the curve is far above the diagonal line, the result is regarded as pessimistic.

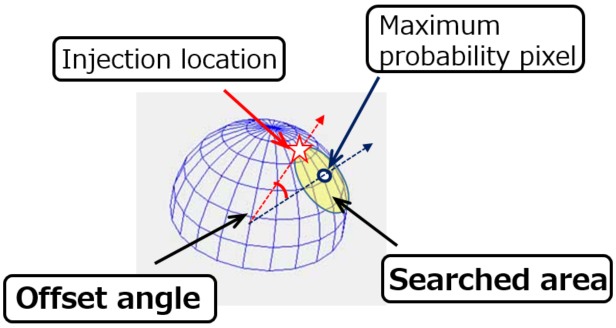

Figure 3.2: Overview of offset angle and searched area. Offset angle and searched area describes the accuracy of the sky map.

3.2 Injection meta data set

248 injection sets, which are selected for a previous MDC, are used in this work[1]. All the waveforms of the injections are Binary Neutron Star (BNS). The waveforms are 'SpinTaylorT4', which are accurate to 3.5 PN order in phase and 1.5 PN order in amplitude. The component masses are distributed between 1.2 $M_{\odot}$ and 1.6 $M_{\odot}$ uniformly, and each component has random spin with maximum magnitude 0.05. The 248 sources are located uniformly over the sky as shown in figure 3.3.

[1]The investigation and the injection file are available from here :
https://www.lsc-group.phys.uwm.edu/ligovirgo/cbcnote/MBTAlocalisationMDC/2015MDC

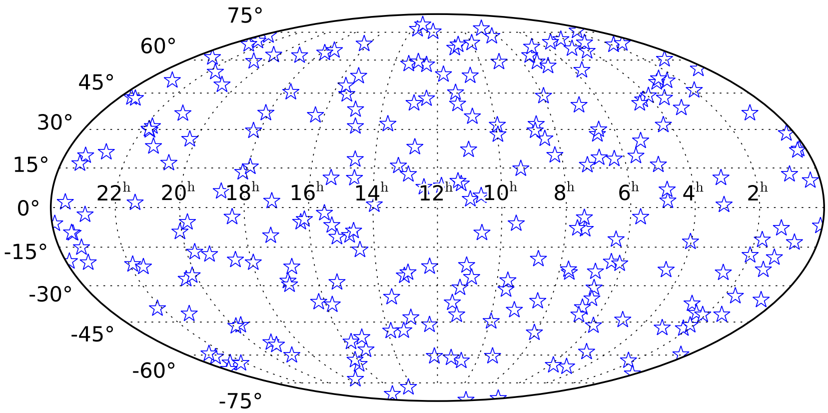

Figure 3.3: Injected source positions. The blue stars describe the location of each injection.

3.3 SNR distribution and False Alarm Probability

This investigation uses the SNR distribution, and the False Alarm Probability (FAP) which is generated from the SNR distribution. They are based on the measurements in the O1 run[2]. The measured SNR distribution is used to produce V1 and K1 SNR, while the FAP is used to decide if a trigger is generated from either a noise, such as a glitch, or a real GW signal. Here V1 and K1 denote Virgo detector and KAGRA detector. This chapter describes how the SNR distribution and the FAP are constructed.

3.3.1 Measured SNR distribution in O1 run

The SNR for the low sensitivity detectors is generated by using random generator following the SNR distribution. In this work, one SNR distribution based on the O1 measurements, is constructed as follows.

First, some of the measured SNR distributions from about 30 hours data with H1 or L1 detectors are fitted. Here H1 and L1 denote LIGO Hanford detector and LIGO Livingstone detector. An example of the measured SNR distribution and the fitted distributions are shown in figure 3.4 and figure 3.5 respectively. These triggers are collected in the period when the H1 and the L1 detector are stationary, namely, when they were not generating many glitches. The curves are extrapolated below SNR 6, because only the data above SNR 6 was recorded. Then, one typical curve is selected from those fitted distributions. In this work, the worst distribution is selected for one below the SNR of 6, while arbitrary one for above the SNR of 6. This is because the random generator should be more affected from lower SNR region of the distribution, rather than the higher SNR region of the distribution.

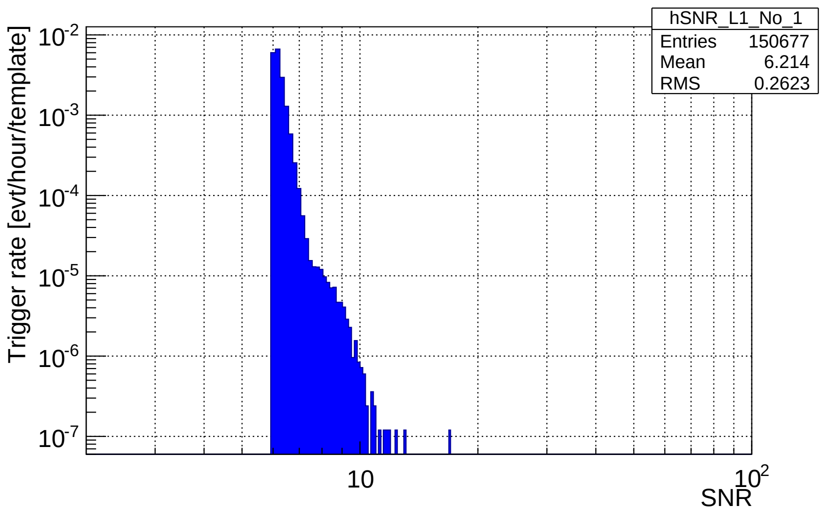

Figure 3.4: An example of the background SNR distribution measured at LIGO Livingston detector (L1) in O1 run.

The selected SNR distribution and its cumulative distribution are plotted in figure 3.6. The cumulative SNR distribution is obtained by counting the cumulative number of the trigger rate in all the bins down to the specific bin. The SNR which holds the lowermost counted bin is called SNR threshold in this theses. The vertical axis is normalized into trigger rate, which is the number of the triggers divided by the analyzed time and by the number of the templates. In the O1 running, only

[2] O1 denotes observation 1 and it was a scientific operation by the two LIGO detectors which was conducted in 2015 and 2016.

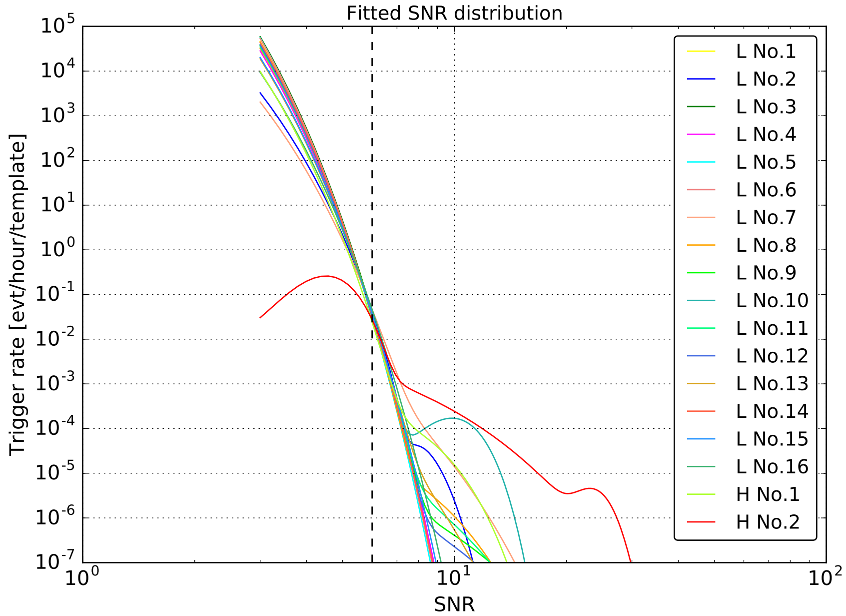

Figure 3.5: Examples of fitted background SNR distribution at H1 and L1 detector in O1 run.

the information of the region 1 of the MBTA was used, and it had about 2×10^5 templates. Thus, all the trigger rates in figure 3.5, figure 3.6 are divided by 2×10^5 templates.

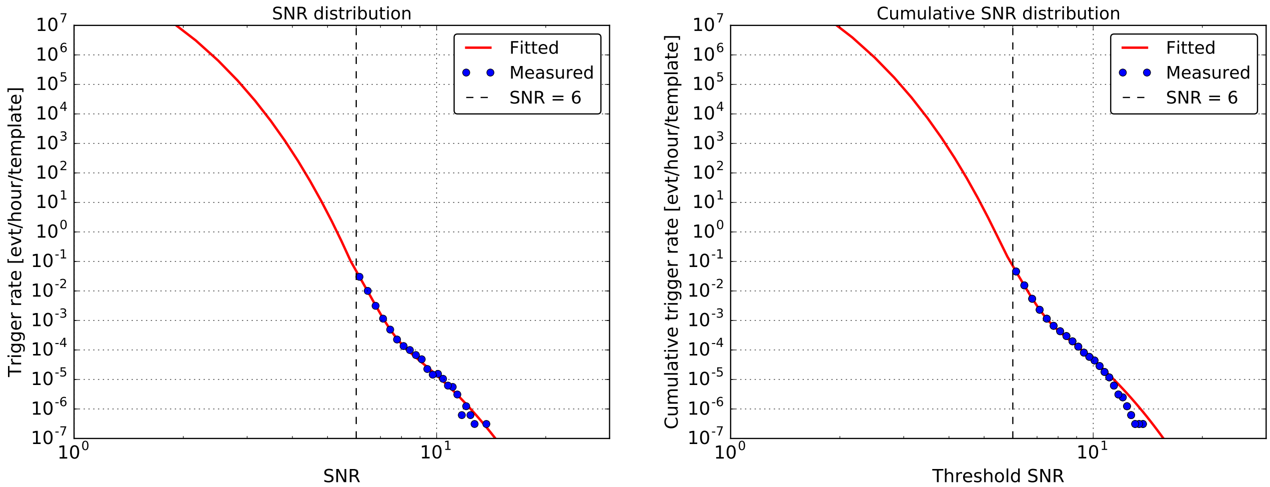

Figure 3.6: Fitted background SNR distribution (*Left*). Fitted cumulative background SNR distribution per template (*Right*). The blue dots are the measured points, while the red curves are obtained by fitting. The vertical axis is normalized into trigger rate, which is the number of the triggers divided by the analyzed time.

3.3.2 Validation of extrapolation

The behavior of other SNR distributions, which have information of lower SNR region is investigated in this section. The purpose is to confirm if the extrapolation at low SNR, which is exerted in above section, is reasonable or not. Then, some SNR distributions from different MBTA searching method[3] are investigated. The obtained distributions and their cumulative distributions are shown in figure 3.7. The curves, except for the red one in figure 3.7, are outputs from off-line analysis. The difference among the curves is the collecting time period to record the triggers. In figure 3.7, the trigger collecting time period becomes shorter and shorter, in the order of the color of green, magenta, yellow, blue and cyan. All triggers except for the red curves are collected in a quiet period of H1 detector, in this plot.

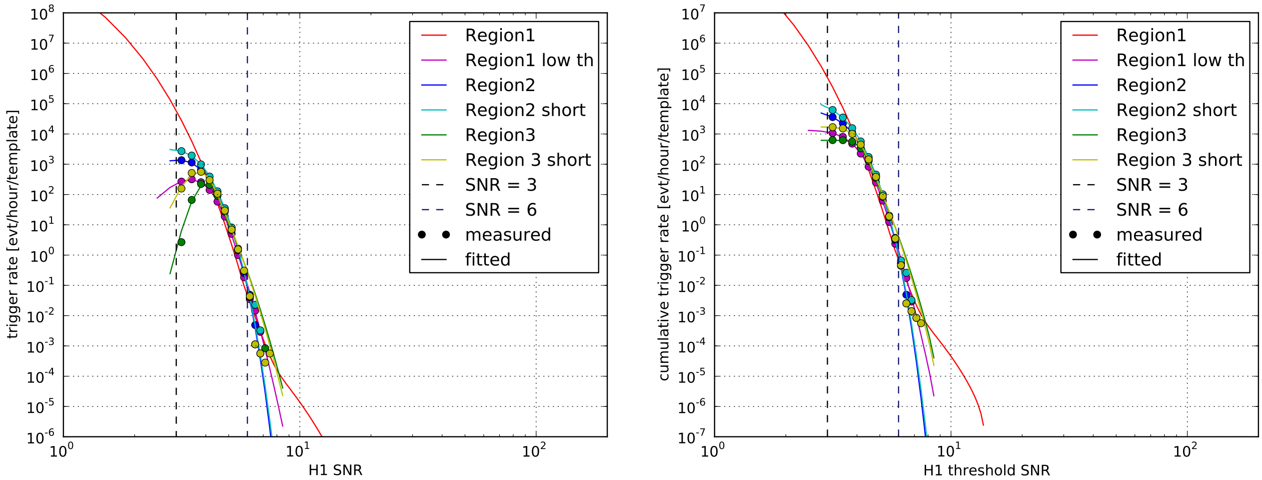

Figure 3.7: SNR distribution (*Left*), cumulative SNR distribution (*Right*). The red curve are extrapolated below SNR 6. Regions denote the template region of MBTA. The 'low th' shows the SNR distribution obtained by using data with lower SNR threshold. Note that this 'Region 1 low th' is analyzed by using 4933 templates. The 'short' represents the SNR distribution calculated by shorter collecting time data. The collecting time becomes longer in the order of the region 1, region 2 and the region 3. The cumulative distribution of the 'region 1 low th' saturates relatively at lower trigger rate. This is because there might be a lot of triggers in its collecting time and miss to count large number of the triggers, compared to the other analysis except for the 'region 3' in the plot.

According to the plots, above the SNR of around 4, all the off-line analyzed distributions follow in accordance with the red line, and behave mostly in the same manner. The differences between the red curve and the others at SNR of 5 are around a factor of 3. It means that the extrapolation in the red line between SNR of 4 and 6, is close to the realistic distribution. In addition, the behavior of the distributions does not depend on the template regions. This fact tells us that the SNR distribution does not depend on how the analysis is done, at high SNR.

On the other hand, below around SNR of 4, all the behavior seem to be different, and the number of triggers is decreased at low SNR, as shown in figure 3.7 *Left*. The cumulative distribution have a saturation below SNR 4, because of the behavior, as seen in figure 3.7 *Right*. However, this behavior seems to be from an artificial problem, due to the following reason. This is because when the time period to collect the triggers is changed, the behavior of the distribution is also changed. For

[3] In O1 running, the MBTA has only one region template (region 1), thus, to produce the SNR distributions from another regions, off-line analysis is necessary.

instance, if the time period to collect the triggers is set shorter, the decrease in the distribution or the saturation in the cumulative distribution at low SNR becomes less effective, and becomes closer to the extrapolated red curve. Due to this behavior, the decrease or saturation of the distributions seems to be caused by a method to collect the triggers in the MBTA. The MBTA records only one trigger in a certain time length, above a threshold SNR. The conceptual diagram is in figure 3.8. According to the functionality, if the threshold is not high enough, or if the collecting time is not short enough, some triggers can be missed, without being recorded. This collecting time effect seems to lead the behavior shown in figure 3.7.

Consequently, based on the above result, it is confirmed that the selected red color SNR distribution is available. This is because the extrapolation is valid at high SNR, especially above SNR of 4, and also even below the SNR of 4, the proper distribution seems to be closer to the red colored curve, rather that the others.

Note that apart from the issue about validation of the SNR distribution, it is found that MBTA can not output correct SNR distribution at low SNR, if the collecting time is set too long.

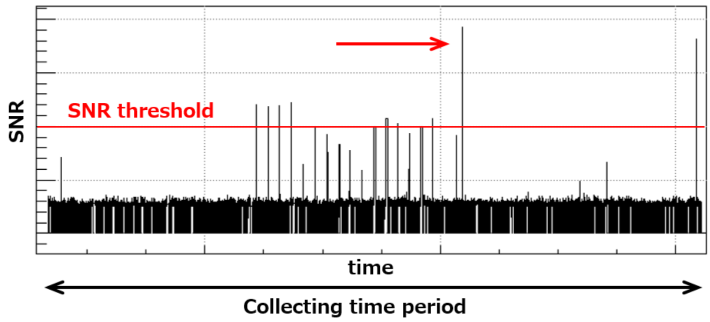

Figure 3.8: Conceptual diagram of collecting triggers in the MBTA. Only one trigger above a given threshold (marked in red), is recorded within a certain time period. If the collecting time period is shortened, the chance to get correct number of triggers is increased.

3.3.3 FAP based on measured SNR distribution

The FAP is calculated from the measured SNR distribution constructed above, via the cumulative distribution, by using the following formula:

$$\mathrm{FAP} \equiv 1 - \exp(-R \times T), \tag{3.1}$$

where R is a cumulative rate of background triggers above a given threshold SNR per one template, which is obtained from the measured O1 data, and T is the time used in the analysis of low sensitivity detectors. The number of the template is set at 2×10^5 for both the Virgo and KAGRA detectors. On the other hand, T is at 70 ms for Virgo, while 80 ms for KAGRA. These numbers come from the differences of typical arrival timing of each detector. In this process, the cumulative distribution is normalized to trigger rate per template. This is because the low sensitivity detectors are analyzed for only one template in this work and there is no need to look for all of the templates in the hierarchical search. The acquired FAP is shown in figure 3.9. In the plot, the horizontal axis is the SNR threshold of one detector, while the vertical axis denotes the probability. According to the plot, the FAP reaches to 1 at around SNR threshold of 3. The shown curves are the FAPs for Virgo and KAGRA detector, which are used in following calculation.

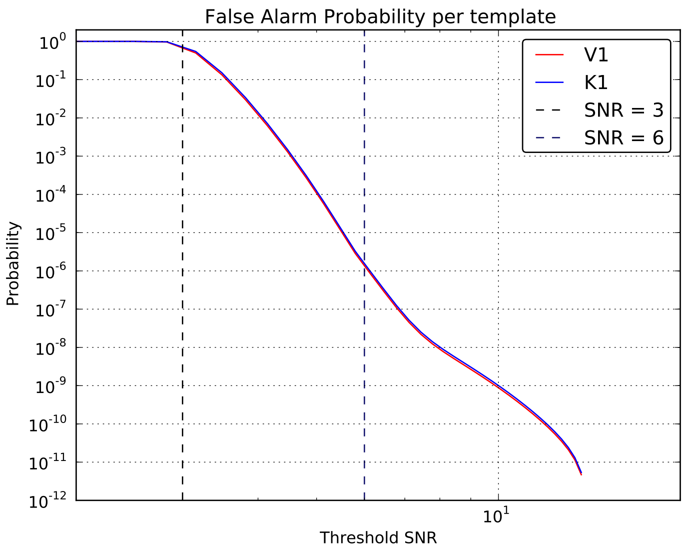

Figure 3.9: False alarm probability (FAP) for V1 (red) and K1 (blue) detector, which are used in this investigation. Here, V1 and K1 denote Virgo detector and KAGRA detector. They are obtained from measured SNR distribution on O1 running, described in chapter 3.3.1. Note that O1 is a scientific operation by the two LIGO detectors conducted in 2015 and 2016.

3.4 Noise curves

The BAYESTAR uses shapes of noise curves[4] to calculate an uncertainty of the maximum probability, via timing uncertainties obtained from the noise curves. This section describes information of used noise curves in this work. Figure 3.10 shows some previously generated noise curves. The curves are based on the expected configuration for the first advanced era observation run in 2015 and 2016 [9]. In the *psd* 2015, the LIGO Hanford (H1) and LIGO Livingston (L1) are assumed to be operating with range of a 54 Mpc for averaged 1.4 $M_\odot$ - 1.4 $M_\odot$ BNS. On the other hand, In the *psd* 2016, the H and L are assumed to be operating with an averaged 1.4 $M_\odot$ - 1.4 $M_\odot$ BNS range 108 Mpc, and the addition of Virgo (V1) with a range of 36 Mpc. The plotted noise curves are Gaussian noise colored to match the expected noise amplitude spectral density curves of these H1, L1, and V1 detectors obtained by mimicking the expected configuration for the first advanced era observation run in 2015 and 2016[5]. The noise curves, labeled as *psd* 2016 in figure 3.10, are used for this calculation.

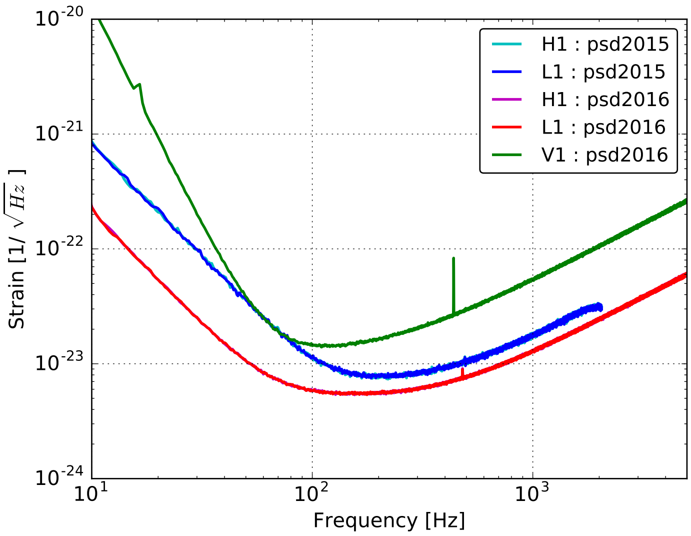

Figure 3.10: Modeled detector noise curves.

[4]The sensitivities and noise curves can be input separately in the BAYESTAR.

[5]The noise curves of the *psd* 2016, which are used in this study, are different from what is used in previous investigation. In the previous test, *psd* 2015 was used.

3.5 Sky localization with 2 detectors

This section describes the localization performance of the double coincidences with the H1 and the L1 detector. As seen above, the necessary information to generate a sky map is detected SNR, arrival timing, phase of GW, and a noise curve of each detector. In this investigation of double detectors search, the input information for the BAYESTAR, namely the output of MBTA, was previously generated in a software injection test using injection meta data set introduced in chapter 3.2. The localization performances with the two detector search using such information, are summarized in figure 3.11. In this calculation, 248 sky maps are generated and their statistics of the offset angle, the searched area, 90 % confidence area, and the P-P plot are in the plots. According to the results, the median values of the offset angle, searched area, and 90 % confidence area are 21 deg, 137 deg^2, and 840 deg^2 respectively.

Uncertainties of arrival timing and SNR are shown in figure 3.12. These uncertainties are the ones which are expected to be included during the passing through the MBTA. Figure 3.13 shows the relation between the uncertainties of timing and the expected detection SNR for each H1 and L1 detector. According to them, the absolute value of the timing uncertainties becomes smaller if the detected SNR becomes larger. The SNR dependences of the arrival timing are fitted as follows,

$$\begin{aligned}\delta t_{\mathrm{H1}} &= 0.65\ \mathrm{ms} \times \frac{6}{\mathrm{SNR}_{\mathrm{H1}}^{\mathrm{expected}}},\\ \delta t_{\mathrm{L1}} &= 0.67\ \mathrm{ms} \times \frac{6}{\mathrm{SNR}_{\mathrm{L1}}^{\mathrm{expected}}},\end{aligned}$$

where the δt_{H1}, δt_{L1} are the absolute values of arrival timing error of each H1, L1 detector.

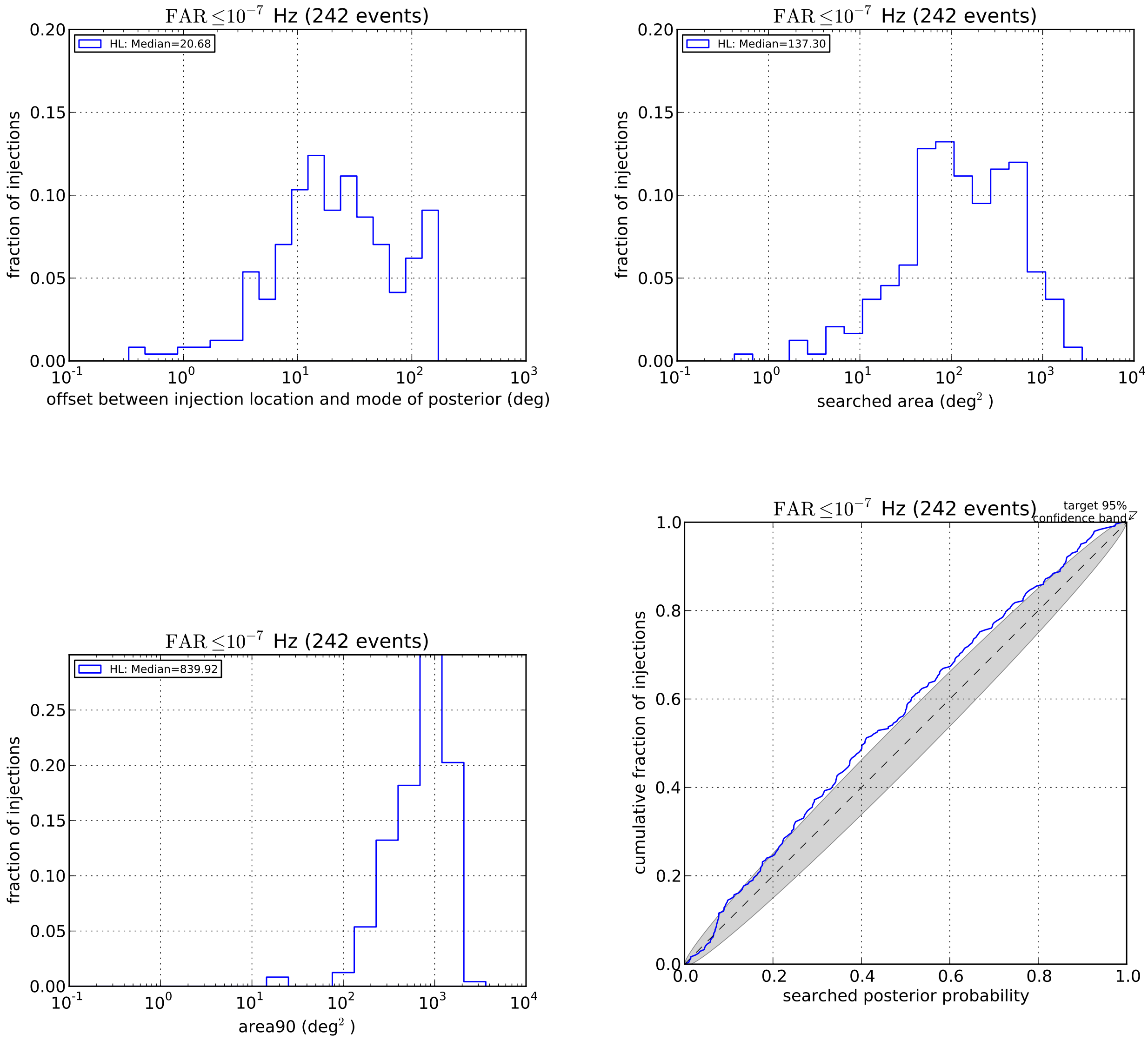

Figure 3.11: Statistics of sky localization performances. Offset angle (*Upper left*), searched area (*Upper right*), 90 % confidence area (*Lower left*), and the P-P plot (*Lower right*) are illustrated. The gray colored area, in the 90 % confidence area plot, denotes the 95 % confidence band. In this thesis, it is regarded as self-consistent if the curve is plotted within or close to the gray area. Details about the P-P plot will be described in following subsection.

	Offset angle	searched area	90 % confidence area
HL (psdHLV_2016MDC)	21 deg	137 deg^2	840 deg^2

Table 3.1: Sky localization performance with HL double coincidence search. The median values of the offset angle, searched area, 90 % confidence area are described.

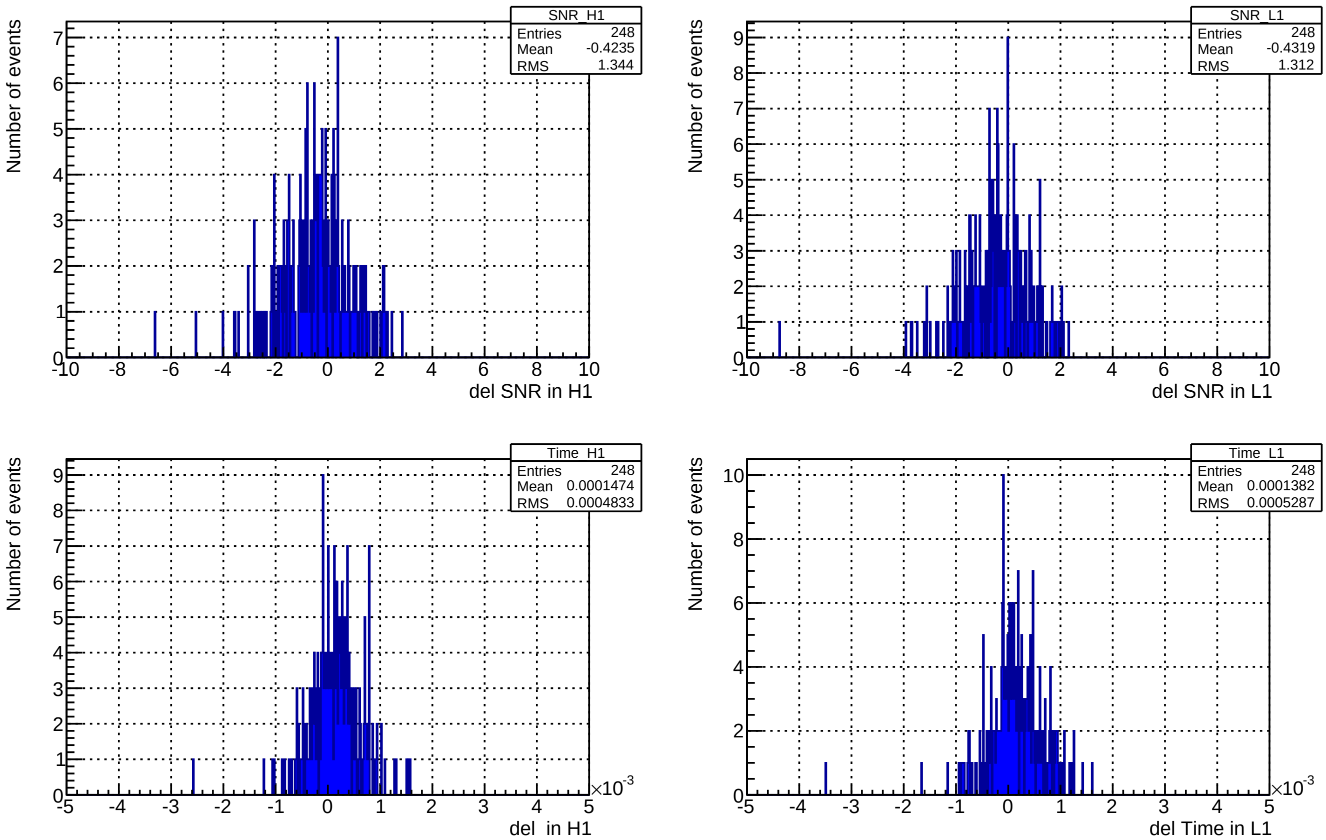

Figure 3.12: Added uncertainties of SNR and timing for H1, L1 detectors.

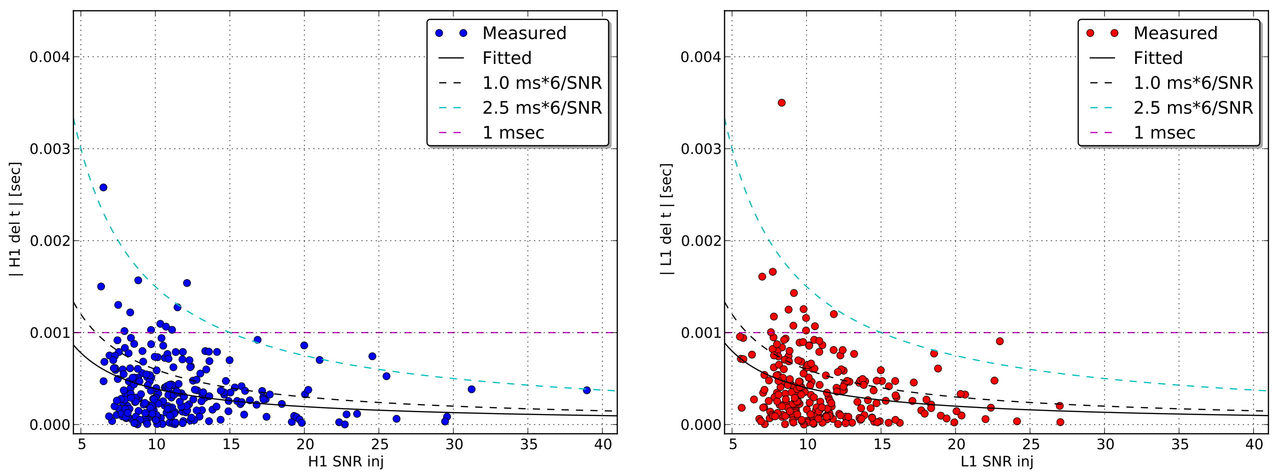

Figure 3.13: SNR dependence of arrival timing error for H1 (*Left*), L1 (*Right*).

3.6 Generating triggers for low sensitivity detectors

The trigger information of the low sensitivity detectors, such as Virgo or KAGRA, is generated as described in this chapter. This is because those detectors are still under construction, and real information is yet to be known. Necessary information for this work is SNR, arrival timing, phase of gravitational waves, and noise curve of each detectors. In this work, the effective distance D_{eff} and the expected SNR is related as described in the 1.1.2. In this simulation, detection ranges for two LIGO detectors (H1, L1) are assumed at 70 Mpc, while 20 Mpc for the Virgo, KAGRA detectors, to consider realistic condition when newly constructed detectors entered the detection network.

The trigger information for Virgo or KAGRA is added to the previous triggers from the double coincidences with H1, L1. Then, information is converted into sky maps via the BAYESTAR.

There are two ways of generating a trigger, and three procedures, as follows.

3.6.1 Triple detector search with HLV

This section describes transforming a set of HL double coincidences, detected in the sky localization MDC into HLV coincidences.

Generating random triggers : $\mathrm{V}_{\text{random}}$

This is for a case if we want to make a trigger, which is generated from noise such as glitches:

- SNR = Random above a threshold SNR, following measured O1 SNR distribution.
- Time = $t_0 + \Delta t$
  ◦ $t_0 = t_{\text{H1}}$ if $\text{SNR}_{\text{H1}} > \text{SNR}_{\text{L1}}$, otherwise $t_0 = t_{\text{L1}}$
  ◦ Δt = random uniform number:
  from -35 ms to 35 ms, for V1.
- Phase = random uniform number from 0 rad to 2π rad.
- Effective distance $D_{\text{eff}} = 2.26 \times$ detection range $\times$ 8 / SNR

Generating triggers based on injection parameters : V_{inj}

This is for a case if we want to make a trigger, which is generated from a GW signal :

- SNR = $\text{SNR}^{\text{expected}} + \Delta\text{SNR}$,
  ◦ $\text{SNR}^{\text{expected}} = 2.26 \times$ detection range $\times$ 8 / D_{eff},
  ◦ $\Delta\text{SNR} =$ random Gaussian(0, 1),
  ◦ $D_{\text{eff}} =$ injection meta data,
  ◦ detection range for V1 = 20 Mpc.
- Time = $t^{\text{expected}} + \Delta t$,
  ◦ $t^{\text{expected}} =$ injection meta data,
  ◦ $\Delta t =$ random Gaussian(0, 0.66 ms $\times$ 6 / $\text{SNR}^{\text{expected}}$).
- Phase = $\phi_0 + \Delta\phi$,
  ◦ $\phi_0 = \phi_{\text{H1}} - \Delta\phi^{\text{expected}}_{\text{HV}}$ if $\text{SNR}_{\text{H1}} > \text{SNR}_{\text{L1}}$, otherwise $\phi_0 = \phi_{\text{L1}} - \Delta\phi^{\text{expected}}_{\text{LV}}$, for V1,
  $\phi_{\text{H1}}, \phi_{\text{L1}} =$ measured data by MBTA in the previous injection test,
  $\Delta\phi^{\text{expected}}_{\text{HV}}, \Delta\phi^{\text{expected}}_{\text{LV}}$ are generated from injection meta data,
  ◦ $\Delta\phi =$ random Gaussian(0, 0.25 rad).

Note that the Gaussian(μ, σ) corresponds to this function:

$$\text{Gaussian}(\mu, \sigma) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \tag{3.2}$$

Here, the $\sigma = 1$ in ΔSNR and $\sigma = 0.25$ rad in $\Delta\phi$ come from measured data by MBTA in the previous injection test.

Mixing procedure

By using $\mathrm{V}_{\mathrm{random}}$ and $\mathrm{V}_{\mathrm{inj}}$ triggers which are generated as described above, the coincidences of the detectors are constructed as follows. In this procedure, the curve in figure 3.9 is used for the FAP for V1. In the following, p_{V1} is a random uniform number from 0 to 1.

Case 1: V1 triggers are random

This is for the worst case with three detector search.

Conditions	Generated coincidences
$p_{V1} < FAP_{\mathrm{V1}}$	HL + $\mathrm{V}_{\mathrm{random}}$
$p_{V1} > FAP_{\mathrm{V1}}$	HL

Table 3.2: Generated coincidences in case 1.

Case 2: V1 triggers are based on injection parameters

This is for the best case with three detector search.

Conditions	Generated coincidences
$\mathrm{SNR}_{V1} > \mathrm{Threshold}_{\mathrm{V1}}$	HL + $\mathrm{V}_{\mathrm{inj}}$
$\mathrm{SNR}_{V1} < \mathrm{Threshold}_{\mathrm{V1}}$	HL

Table 3.3: Generated coincidences in case 2.

Case 3: V1 triggers are either random or based on injection parameters

This is for a more realistic case with three detector search. The coincidences are generated according to the FAP and a given threshold. Here, $FAP_{\mathrm{V1}} = FAP_{\mathrm{V1}}(\mathrm{SNR})$ if $\mathrm{SNR}_{\mathrm{V1}} > \mathrm{Threshold}_{\mathrm{V1}}$, otherwise $FAP_{\mathrm{V1}} = FAP_{\mathrm{V1}}(\mathrm{Threshold}_{\mathrm{V1}})$.

Conditions	Generated coincidences
If $p_{V1} < FAP_{\mathrm{V1}}$	
$\mathrm{SNR}_{\mathrm{V1}} > \mathrm{Threshold}_{\mathrm{V1}}$	HL + $\mathrm{V}_{\mathrm{random}}(\mathrm{SNR}_{\mathrm{th}} = \mathrm{SNR}_{\mathrm{V1}})$
$\mathrm{SNR}_{\mathrm{V1}} < \mathrm{Threshold}_{\mathrm{V1}}$	HL + $\mathrm{V}_{\mathrm{random}}(\mathrm{SNR}_{\mathrm{th}} = \mathrm{Threshold}_{\mathrm{V1}})$
If $p_{V1} > FAP_{\mathrm{V1}}$	
$\mathrm{SNR}_{\mathrm{V1}} > \mathrm{Threshold}_{\mathrm{V1}}$	HL + $\mathrm{V}_{\mathrm{inj}}$
$\mathrm{SNR}_{\mathrm{V1}} < \mathrm{Threshold}_{\mathrm{V1}}$	HL

Table 3.4: Generated coincidences in case 3.

3.6.2 Quadruple detector search with HLVK

This section describes transforming the set of HL double coincidences detected in the sky localization MDC into HLV, HLK, or HLVK coincidences. The concept is same as the case with three detectors search.

Generating random triggers : $\mathrm{V_{random}}$, $\mathrm{K_{random}}$

- SNR = Random above a threshold SNR, following measured O1 SNR distribution.
- Time = $t_0 + \Delta t$
  ◦ $t_0 = t_{\mathrm{H1}}$ if $\mathrm{SNR_{H1}} > \mathrm{SNR_{L1}}$, otherwise $t_0 = t_{\mathrm{L1}}$
  ◦ Δt = random uniform number:
  from -35 ms to 35 ms, for V1.
  from -30 ms to 30 ms, for K1 to H1.
  from -40 ms to 40 ms, for K1 to L1.
- Phase = random uniform number from 0 rad to 2π rad.
- Effective distance D_{eff} = 2.26 × detection range × 8 / SNR

Generating triggers based on injection parameters : $\mathrm{V_{inj}}$, $\mathrm{K_{inj}}$

- SNR = $\mathrm{SNR^{expected}} + \Delta\mathrm{SNR}$
  ◦ $\mathrm{SNR^{expected}}$ = 2.26 × detection range × 8 / D_{eff}
  ◦ $\Delta\mathrm{SNR}$ = random Gaussian(0, 1).
  ◦ D_{eff} = injection meta data
  ◦ detection range for V1 = 20 Mpc
  ◦ detection range for K1 = 20 Mpc

- Time = $t^{\mathrm{expected}} + \Delta t$
  ◦ t^{expected} = injection meta data
  ◦ Δt = random Gaussian(0, 0.66 ms × 6 / $\mathrm{SNR^{expected}}$).
- Phase = $\phi_0 + \Delta\phi$
  ◦ $\phi_0 = \phi_{\mathrm{H1}} - \Delta\phi^{\mathrm{expected}}_{\mathrm{HV}}$ if $\mathrm{SNR_{H1}} > \mathrm{SNR_{L1}}$, otherwise $\phi_0 = \phi_{\mathrm{L1}} - \Delta\phi^{\mathrm{expected}}_{\mathrm{LV}}$, for V1
  ◦ $\phi_0 = \phi_{\mathrm{H1}} - \Delta\phi^{\mathrm{expected}}_{\mathrm{HK}}$ if $\mathrm{SNR_{H1}} > \mathrm{SNR_{L1}}$, otherwise $\phi_0 = \phi_{\mathrm{L1}} - \Delta\phi^{\mathrm{expected}}_{\mathrm{LK}}$, for K1
  $\phi_{\mathrm{H1}}, \phi_{\mathrm{L1}}$ = measured data by MBTA in the previous injection test
  $\Delta\phi^{\mathrm{expected}}_{\mathrm{HV}}$, $\Delta\phi^{\mathrm{expected}}_{\mathrm{LV}}$, $\Delta\phi^{\mathrm{expected}}_{\mathrm{HK}}$, $\Delta\phi^{\mathrm{expected}}_{\mathrm{LK}}$ are generated from injection meta data.
  ◦ $\Delta\phi$ = random Gaussian(0, 0.25 rad).

Note that the Gaussian(μ, σ) corresponds to this function:

$$\mathrm{Gaussian}(\mu,\sigma) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{3.3}$$

Mixing procedure

$p_{V1},\ p_{K1}$ = random uniform number from 0 to 1.

Case 1: V1, K1 triggers are random

Conditions	Generated coincidences
$p_{V1} < FAP_{\mathrm{V1}}$ and $p_{\mathrm{K1}} < FAP_{K1}$	HL + $\mathrm{V_{random}}$ + $\mathrm{K_{random}}$
$p_{V1} > FAP_{\mathrm{V1}}$ and $p_{\mathrm{K1}} < FAP_{K1}$	HL + $\mathrm{K_{random}}$
$p_{V1} < FAP_{\mathrm{V1}}$ and $p_{\mathrm{K1}} > FAP_{K1}$	HL + $\mathrm{V_{random}}$
$p_{V1} > FAP_{\mathrm{V1}}$ and $p_{\mathrm{K1}} > FAP_{K1}$	HL

Table 3.5: Generated coincidences in case 1.

Case 2: V1, K1 triggers are based on injection parameters

Conditions	Generated coincidences
$\mathrm{SNR}_{V1} > \mathrm{Threshold_{V1}}$ and $\mathrm{SNR}_{K1} > \mathrm{Threshold_{K1}}$	HL + $\mathrm{V_{inj}}$ + $\mathrm{K_{inj}}$
$\mathrm{SNR}_{V1} < \mathrm{Threshold_{V1}}$ and $\mathrm{SNR}_{K1} > \mathrm{Threshold_{K1}}$	HL + $\mathrm{K_{inj}}$
$\mathrm{SNR}_{V1} > \mathrm{Threshold_{V1}}$ and $\mathrm{SNR}_{K1} < \mathrm{Threshold_{K1}}$	HL + $\mathrm{V_{inj}}$
$\mathrm{SNR}_{V1} < \mathrm{Threshold_{V1}}$ and $\mathrm{SNR}_{K1} < \mathrm{Threshold_{K1}}$	HL

Table 3.6: Generated coincidences in case 2.

Case 3: V1, K1 triggers are either random or based on injection parameters

Here, $FAP = FAP(\mathrm{SNR})$ if SNR > Threshold, otherwise $FAP = FAP(\mathrm{Threshold})$.

Conditions	Generated coincidences
If $p_{V1} < FAP_{\mathrm{V1}}$ and $p_{K1} < FAP_{\mathrm{K1}}$	
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{SNR_{V1}}) + \mathrm{K_{ran}}(\mathrm{SNR_{K1}})$
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{SNR_{V1}}) + \mathrm{K_{ran}}(\mathrm{Threshold_{K1}})$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{Threshold_{V1}}) + \mathrm{K_{ran}}(\mathrm{SNR_{K1}})$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{Threshold_{V1}}) + \mathrm{K_{ran}}(\mathrm{Threshold_{K1}})$
If $p_{V1} < FAP_{\mathrm{V1}}$ and $p_{K1} > FAP_{\mathrm{K1}}$	
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{SNR_{V1}}) + \mathrm{K_{inj}}$
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{SNR_{V1}})$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{Threshold_{V1}}) + \mathrm{K_{inj}}$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{ran}}(\mathrm{Threshold_{V1}})$
If $p_{V1} > FAP_{\mathrm{V1}}$ and $p_{K1} < FAP_{\mathrm{K1}}$	
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{inj}} + \mathrm{K_{ran}}(\mathrm{SNR_{K1}})$
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{inj}} + \mathrm{K_{ran}}(\mathrm{Threshold_{K1}})$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} \quad + \mathrm{K_{ran}}(\mathrm{SNR_{K1}})$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	$\mathrm{HL} \quad + \mathrm{K_{ran}}(\mathrm{Threshold_{K1}})$
If $p_{V1} > FAP_{\mathrm{V1}}$ and $p_{K1} > FAP_{\mathrm{K1}}$	
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{inj}} + \mathrm{K_{inj}}$
$\mathrm{SNR_{V1}} > \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	$\mathrm{HL} + \mathrm{V_{inj}}$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} > \mathrm{Threshold_{K1}}$	$\mathrm{HL} \quad + \mathrm{K_{inj}}$
$\mathrm{SNR_{V1}} < \mathrm{Threshold_{V1}}$ & $\mathrm{SNR_{K1}} < \mathrm{Threshold_{K1}}$	HL

Table 3.7: Generated coincidences in case 3. Here $\mathrm{V_{ran}}$ denotes $\mathrm{V_{random}}$.

Sky localization with hierarchical search of 3 detectors

Chapter 4

This chapter describes the expected fast localization with the hierarchical search of three GW detectors of H1, L1, and V1 network. Results shown in this chapter are obtained by following the procedure described in chapter 3.6.1. Chapter 4.1 describes the typical sky maps, and the typical localization performances with HLV hierarchical search. Chapter 4.2 shows the optimization of the SNR threshold for the V1 detector, according to the localization performances. Chapter 4.3 tests if the artificial trigger information is reasonable. The localization for cases where there are a lot of background triggers, is investigated in chapter 4.4. Chapter 4.5 describes the localization performance of HLV hierarchical search compared to HL double coincidence search.

4.1 Sky localization performance

Typical sky maps generated from MBTA outputs through BAYESTAR are shown in figure 4.1 and figure 4.2. They are from HL double coincidences, and HLV triple coincidences. Figure 4.1 shows the sky map generated from triple coincidences with V1 trigger based on noise, while figure 4.2 shows the sky map from triple coincidences with V1 trigger based on GW signal. Table 4.1 describes the localization performance of the two events shown in figure 4.1 and figure 4.2. These events are found when the threshold SNRs are set at 5 for H1 and L1, while 3.5 for V1 detector.

Item	Offset [deg]	Searched area [deg^2]	90 % area [deg^2]
Contribution of V_{ran}			
HLV_{ran}	45	429	247
HL	28	315	916
Contribution of V_{inj}			
HLV_{inj}	2	8	61
HL	132	1134	1357

Table 4.1: Examples of sky localization with HLV hierarchical network by BAYESTAR. The upper row event is found with SNRs of 9.5, 10.7, 4.1 for H1, L1, V1. This event is generated from HLV hierarchical network with V1 trigger based on noise. While the lower event is found with SNR of 8.2, 7.2, 5.5 for H1, L1, V1. This event is from HLV hierarchical network with V1 trigger based on GW signal. The corresponding performance of HL search is also described.

According to the two events, the V_{ran} does not contribute to the improvement of the sky map accuracy, even though the precision of the prediction is improved. On the other hand, V_{inj} improves both the precision and the accuracy of the map.

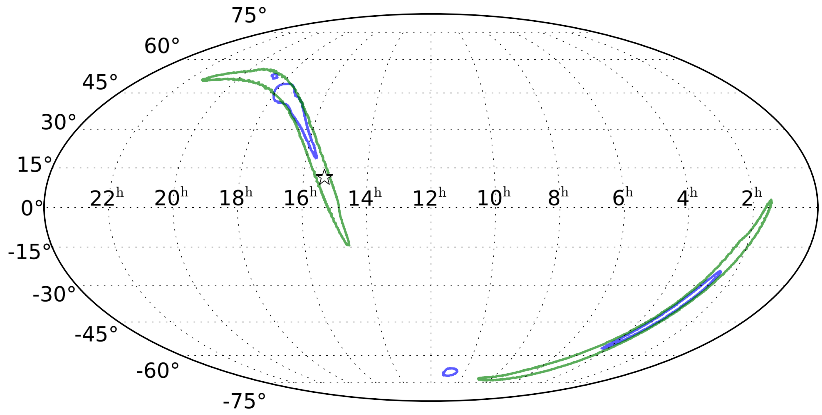

Figure 4.1: Examples of sky maps from the trigger HL double coincidences (green) and from $\mathrm{HLV_{ran}}$ triple coincidences (blue). The star indicates the location of the injected source. Even though the precision of the prediction is improved, the injected location is found out of the 90 % confidence area.

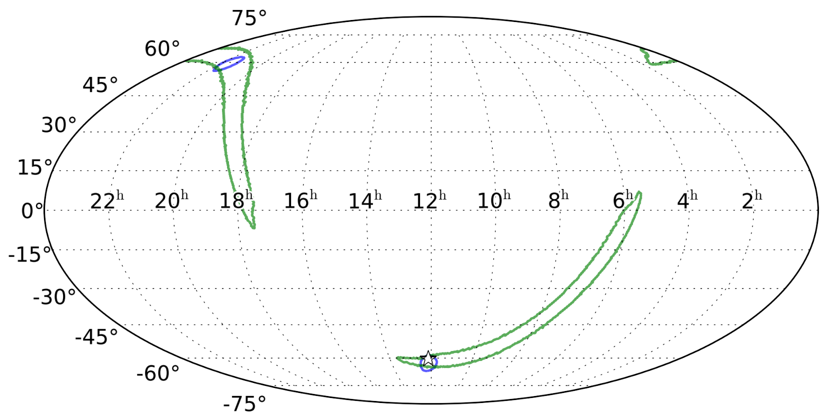

Figure 4.2: Examples of sky maps from the trigger HL double coincidences (green) and from $\mathrm{HLV_{inj}}$ triple coincidences (blue). The star indicates the location of the injected source. Both precision of the prediction and accuracy of the sky map are improved.

To confirm localization performance more statistically, offset angle, searched area, 90 % confidence area are collected into the histograms shown in figure 4.3 for all 248 events, and for each case. The color indicates the information of the cases which are defined in the chapter 3.6.1. All performances shown in figure 4.3 are obtained when the SNR threshold is set at 3.5 for V1, while 5 for the H1 and L1.

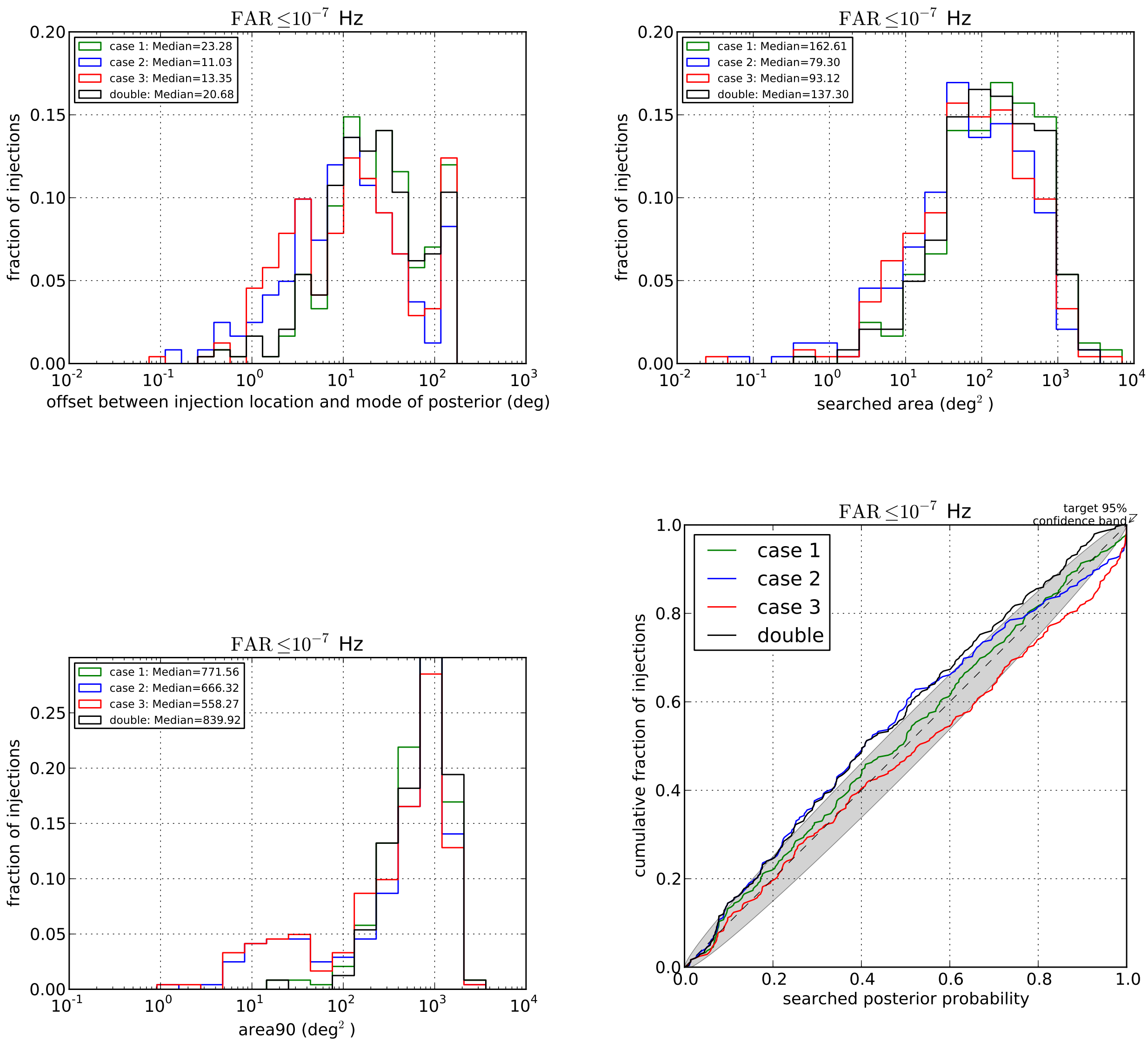

Figure 4.3: Offset angle (*upper left*), serched area (*upper right*), cumulative 90 % confidence area (*lower left*), P-P plot (*lower right*). The green, blue, red, black curves correspond to case 1, case 2, case 3 with 3-detector search with HLV, and 2-detector search with HL, respectively.

According to the offset angle and the searched area, in case 1 (green), the V1 triggers from noise do not contribute to the improvement of the localization. This is because, for instance, in the offset angle plot (Fig 4.3 *upper left*), the population at the lower angle becomes smaller, while the the population at the higher angle becomes larger, compared to the result of the HL double coincidences search. Then, the V1 triggers generated from noises give only worse localize performances, and thus the behavior of case 1 tells the worst performances of the analysis method at the threshold SNR. This effect can also be seen in the median values of the offset angle and the searched area. In the investigation of the threshold 3.5, case 1 has 13 % of the V1 triggers from the noise, while about 87 % are HL double triggers. On the other hand, in both case 2 (blue) and case 3 (red), the population of lower angle becomes larger while that of higher angle becomes smaller, compared to the HL double coincidence search. According to this, the V1 triggers based on real GW signals, improve the localize performances. In case 2, all V1 triggers are based on the GW signals, even if one sets a low threshold SNR for the V1, such as 2. While in case 3, the V1 triggers are a mixture of both triggers from noises and the GW based triggers. This is why case 2 investigation gives better results than the case 3 search at the threshold. Thus, the results of case 2 give the best performances of the localization. However, in a more realistic situation, some triggers generated from noises should be included, if one sets low threshold SNR around 2. Thus, the results of case 3 provide more realistic localize performances.

The 90 % confidence area (Fig 4.3 $lower\ left$) describes that the size of the predicted confidence area tends to shrink by adding V1 information. This happens if one uses a V1 trigger from noise.

The P-P plot describes if the obtained histograms are realistic or not, as described previously. If the calculation is done with proper settings, the curve should follow the diagonal line. This is because in such situation, 90 % of the injections should be found in the 90 % confidence area, for instance. Thus, if the curve is below the diagonal line, then too much timing uncertainties are added into the V1 triggers. Then according to the blue line (case 2) in figure 4.3 $lower\ right$, the added timing error into the V1 triggers is adequate, because the most of the curve follows the diagonal line. Other curves except for case 2, which are generated by using some of V1 triggers from noises, and thus the P-P plots tends to stay below the diagonal line or the curve of the HL double search. This is because the V1 triggers from noise should have large timing errors, compared to the timing error expected from the noise curve.

Note that in case 2, there are 28 % V1 triggers based on the GW signals and 72 % of HL double coincidences. While case 3 has 11 % of triggers from noises, 29 % of GW based triggers, and 60 % of HL double coincidences. This confirms that even if the number of HLV triple coincidence triggers with V1 triggers based on GW signals is small, the sky localization performance can be improved in this method.

Consequently, in the realistic case, the median value of the offset angle, searched area, and 90 % confidence area are expected to become about 70 % times smaller than that of the HL double coincidence search by using this method, when the V1 threshold is set at 3.5 in the configuration.

4.2 Optimization of V1 threshold SNR

This chapter describes the optimization of the SNR threshold for the V1 detector using localization performances. The median values of the sky localization statistics shown in figure 4.3, are collected for each case by changing the V1 SNR threshold. The purpose of this investigation is to obtain the optimal SNR threshold for the V1 detector. Table 4.2 describes the population of the triggers, offset angle, and the searched area, 90% confidence area of each case at each V1 threshold. Note that in this calculation, threshold SNR of H1, L1 are set at 5, and the mean values of the SNR for H1, L1 are 11. In addition, the median values of the offset angle and the searched area by double coincidences HL are 21 deg, 137 deg^2, respectively.

V1 threshold	HLV_{ran}	HLV_{inj}	HL	Offset [deg]	Searched area [deg^2]	90 % area [deg^2]
Case 1						
3.0	78 %	0 %	22 %	28	249	489
3.5	13 %	0 %	87 %	23	163	772
4.0	1 %	0 %	99 %	21	137	834
5.5	0 %	0 %	100 %	21	137	840
Case 2						
2.0	0 %	59 %	41 %	7	61	511
3.0	0 %	37 %	63 %	10	94	645
3.5	0 %	28 %	72 %	12	95	666
4.0	0 %	22 %	78 %	13	102	723
5.5	0 %	10 %	90 %	17	107	791
7.0	0 %	5 %	95 %	19	120	812
8.0	0 %	4 %	96 %	19	124	820
10.0	0 %	1 %	99 %	21	133	835
Case 3						
2.0	66 %	34 %	0 %	14	112	353
2.5	67 %	33 %	0 %	15	124	358
3.0	52 %	33 %	16 %	14	117	382
3.1	41 %	33 %	26 %	13	103	391
3.2	32 %	32 %	36 %	13	134	434
3.3	25 %	30 %	45 %	13	97	494
3.4	19 %	30 %	51 %	14	103	544
3.5	11 %	29 %	61 %	13	95	599
3.6	8 %	28 %	64 %	13	94	628
3.7	5 %	27 %	68 %	13	94	650
3.8	3 %	25 %	72 %	14	97	673
3.9	2 %	24 %	74 %	14	96	681
4.0	1 %	23 %	76 %	14	103	714
4.5	0 %	18 %	82 %	15	103	729
5.0	0 %	15 %	85 %	15	107	766
5.5	0.4 %	10.5 %	89.1 %	16	107	746
7.0	0 %	5 %	95 %	19	126	830
8.0	0 %	3 %	97 %	20	129	832
10.0	0 %	1 %	99 %	21	136	835

Table 4.2: Trigger population, offset angle, searched area, 90% confidence area of HLV hierarchical search.

Figure 4.4 shows the median values of the offset angle and the searched area at each threshold, compared to the results of HL search. In case 3 investigation, calculation is repeated for three times, using different set of random numbers, to estimate the uncertainties of these median values. Thence, the averaged values of case 3 results are shown in table 4.2. Figure 4.4 gives the following information.

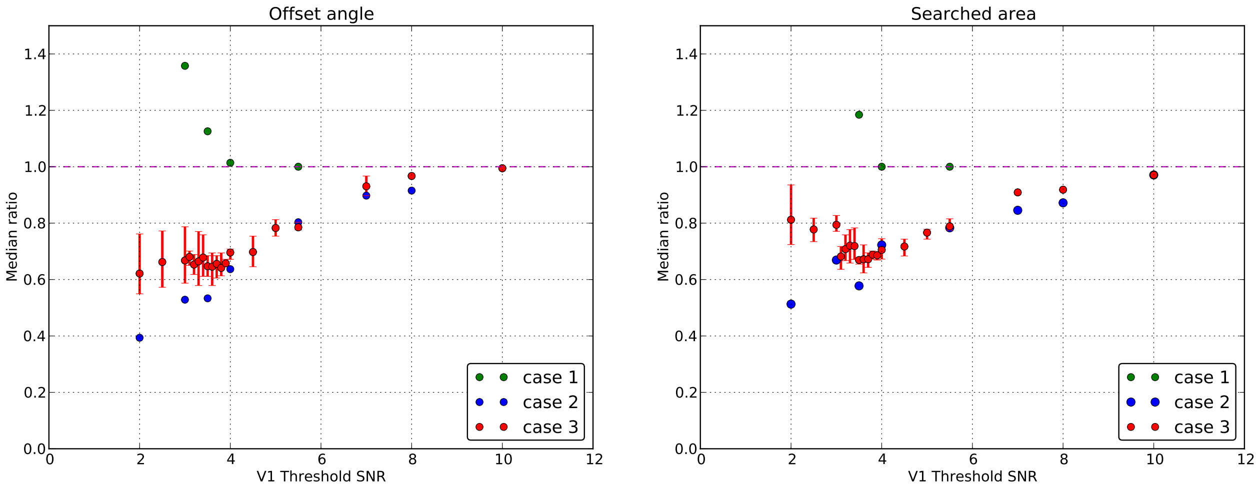

Figure 4.4: Median ratio of Offset angle ($Left$), searched area ($Right$). Both of the vertical axes are normalized with the values of the double coincidence search. In case 3 investigation, calculation is repeated for three times, using different set of random numbers.

First, in all the cases, if the V1 threshold is set at higher SNR, such as 10, the performance becomes closer to that of double coincidences, because the number of the HL double coincidence triggers are increased.

Second, according to case 1, the V1 triggers generated from noise do not contribute to the improvement of the Sky localization performances. If the V1 threshold is lowered in this case, the localization performance worsens, because of an increase in the number of the V1 triggers from noise. This is the worst performance with the hierarchical search using H1,L1,V1.

Third, in case 2, the lower the V1 SNR threshold is set, the better the localization becomes. Case 2 investigation gives the best performance at each threshold, because all V1 triggers are GW signals. If the MBTA could find the triggers generated only from GW signals from V1 detector signal, the offset angle and the searched area are reduced from 21 deg, 137 deg^2 to 7 deg, 61 deg^2 at threshold 2.

Fourth, in case 3, a more realistic behavior of the localization is shown, because the triggers in this case include all the triggers. The triggers consist of HL and V1 triggers from noise, HL and V1 triggers based on GW signals, and HL double coincidence triggers. They are mixed depending on the FAP and the threshold SNR. The uncertainties of the median values become larger below threshold SNR of 4, because the population of the triggers can be changed easier, compared to the higher threshold case as shown in figure 3.9. Since the FAP becomes larger than 0.01 at those thresholds, the FAP become to comparable value to the generated random numbers. According to the figure 4.4, the searched area has a local minimum value. This is because, if the threshold is set too low, the number of the V1 triggers from noise increases and effect of the noises become more dominant than the effect of the V1 triggers from the signals.

Finally, according to the two statistics, the optimal threshold for the V1 detector in the hierarchical search is between 3.5 and 4. At the threshold, the offset angle, the searched area are around 0.7 times smaller than the values with double coincidence search.

On the other hand, figure 4.5 gives 90 % confidence area as a function of the V1 SNR threshold of the HLV hierarchical search. As described in the previous section, a confidence area gives information of precision of the predicted region, while offset angle and searched area give information of accuracy of the predicted sky map. According to the plot, the areas for all cases are reduced when the SNR threshold is lowered, due to the increase in number of the participating detectors in the network. The saturation seen in the case 3 curve happenes because all the triggers are generated from triple coincidences. The expected minimum value of the 90 % confidence area with HLV hierarchical search is 380 deg^2 below the threshold SNR 3 in this investigation.

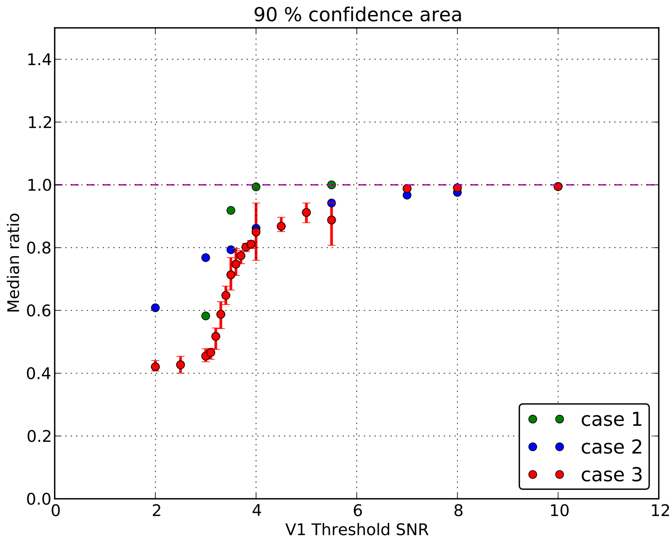

Figure 4.5: 90 % confidence area with HLV hierarchical search. The vertical axis is normalized with the value obtained in HL double search.

Thus, from the point of view of 90 % confidence area, setting SNR threshold as low as possible, would give better localization. In the end, based on all of the offset angle, searched area, and the confidence area, the optimal threshold for the V1 detector in this search is around 3.5. At this threshold, expected offset angle, searched area, and 90 % confidence area are about 13 deg, 93 deg^2, 600 deg^2, respectively. At the threshold, about 40 % of all the triggers are from HLV triple coincidences, while others are from HL double coincidences. Also the HLV$_{\mathrm{ran}}$ triggers comprise 11 % of all the triggers at the threshold of 3.5.

In conclusion, it is confirmed that the hierarchical search allows making effective use of the low sensitivity detector. The expected fast localization of the HLV hierarchical network at the optimal SNR threshold is summarized in table 4.3.

SNR threshold	Offset [deg]	Searched area [deg^2]	90 % area [deg^2]
5 for H1, L1, 3.5 for V1	13 deg	93 deg^2	600 deg^2

Table 4.3: Expected fast sky localization performance with HLV hierarchical network.

This V1 optimal threshold, however, largely depends on a lot of assumptions, especially the sensitivities of each detector, the SNR distribution, and FAP. Thus, if more exact optimal threshold is necessary, this investigation should be repeated by using more realistic SNR distribution, FAP based on a measurement.

4.3 Self-consistency test

As described above, the optimal threshold is obtained at around 3.5, based on the statistics of offset angle, searched area, and 90 % confidence area in the calculation of case 3. This section tests the feasibility of the results according to the P-P plot. This plot describes whether or not the added timing error for V1 detector is reasonable. As described in chapter 3.1.4, the curve should follow the diagonal line if the performance is feasible. In other words, curves in P-P plot becomes closer to the diagonal line if the added timing uncertainties are adequate. For example, if an added timing error is too large, the number of the injections detected within a certain confidence area tends to become smaller, and thus the curve tends to shift towards a lower direction, compared to the diagonal line.

Figure 4.6 shows the P-P plots, obtained in the optimization work. All those curves are from case 2 investigation. This is because if case 3 is investigated, one has to treat V1 triggers based on noise. Since the purpose of this self-consistency test is to verify if the added timing error is reasonable or not, one has to remove V1 triggers based on noise in this test. The $Left$ plot shows the P-P plot when the timing uncertainty Δt is generated from following Gaussian function, $\Delta t \equiv \text{Gaussian}(\mu, \sigma) = \text{Gaussian}(0, 0.66 \text{ ms} \times 6/\text{SNR}^{\text{expected}})$. While, the *Right* plot shows the dependence of the RMS value of the timing uncertainty σ when the V1 SNR threshold is set at 2.0. The left plot implies that the added uncertainties into the V1 timing in this work is appropriate, because all the curves are not so far from the target 95 % confidence band.

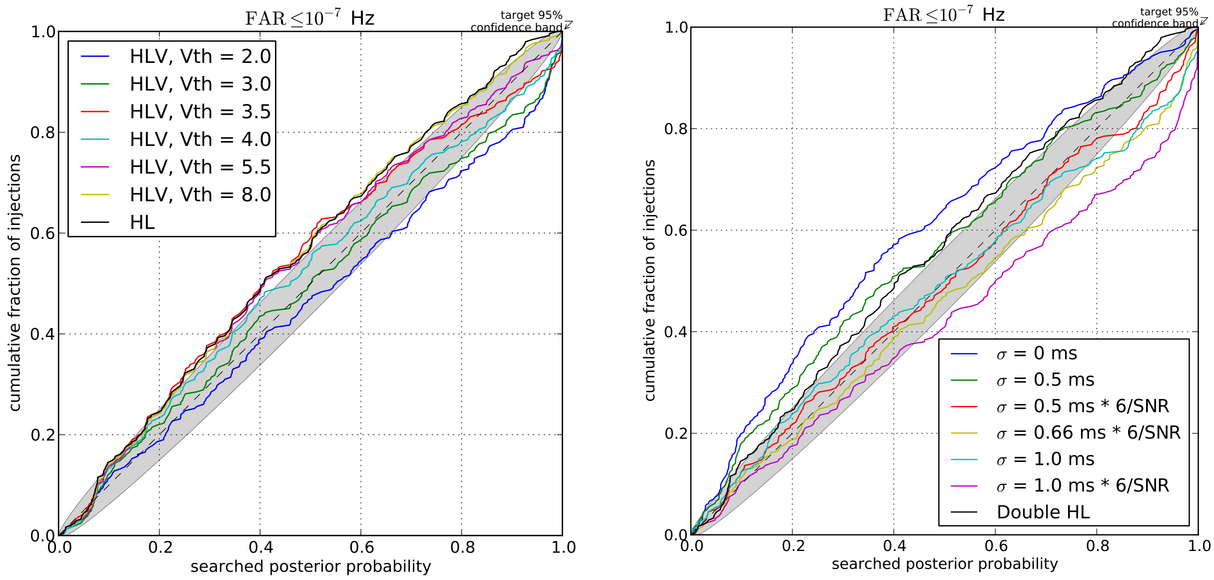

Figure 4.6: P-P plot in the case 2 with various V1 SNR thresholds ($Left$), P-P plot in the case 2 at the SNR threshold 2 with various timing error σ (*Right*)

However, when the threshold is set at a lower value than around 3, the added timing error seems to slightly increase. While the right plot implies that if σ is set at small value, the curve shifts upwards. It seems to be better to set σ as $\sim 0.5 \text{ ms} \times 6/\text{SNR}^{\text{expected}}$, at the SNR threshold 2. Thence, the model of the timing error which was used in this investigation works around the optimal SNR threshold. However it is not the best model for investigation of low threshold behavior such as below around 2.

Consequently, it is confirmed that the added timing error for V1 is reasonable at around the optimal threshold. If one wants to study the behavior at the SNR threshold of around 2 more in detail, it is better to use more realistic SNR dependence of σ.

4.4 Effect of noisy background triggers

Above investigations focuse on the case where the less sensitive V1 detector is stable. However, GW detectors sometimes generate a lot of background triggers called glitches. The condition which a detector generates glitches is called 'noisy' condition in this section. In this situation, the background SNR distribution differs from the one seen in previous section. In the actual low latency search, the data from noisy background can be included into the analysis because veto system in the latency search is not set so strictly. This chapter describes the localization performance when the low sensitivity detector generates a lot of background triggers. The purpose of this investigation is to know how the noisy condition contributes to the sky localization.

4.4.1 SNR distribution and FAP in noisy case

In this noisy case investigation, the SNR distribution and the FAP of the V1 detector is changed. Figure 4.7 describes the SNR distribution and the FAP in this investigation. The red colored curves are for the case when the detector is stable, while the blue curves correspond to the noisy condition. Thus, the main effect of the noisy SNR distribution would occur above the SNR 4, and the optimal threshold would be the same as the previous result.

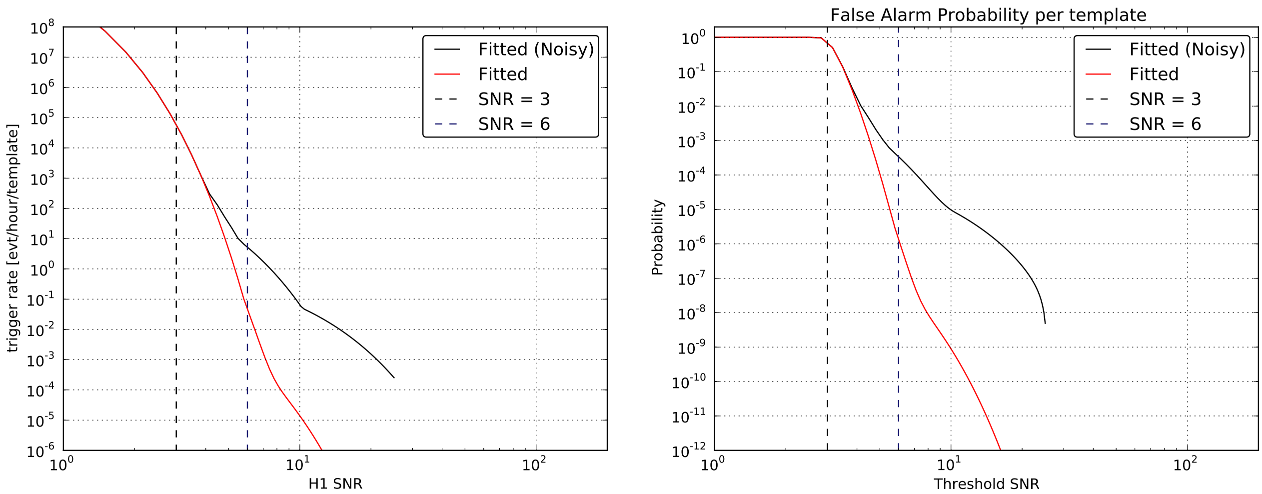

Figure 4.7: SNR distribution in a noisy case ($Left$), FAP in a noisy case ($Right$)

4.4.2 Localization in noisy case

Table 4.4 summarizes the performance in a noisy case, and figure 4.8 shows the obtained sky localization performance using noisy SNR distribution and FAP. The V1 triggers are generated in accordance with the procedure described in the chapter 3.6.1. This investigation is done only for case 1 and case 3, which depend on the SNR distribution and the FAP. The yellow and black dots mean the results from case 1 and case 3, respectively. The behavior of case 1, above the threshold SNR 4, becomes closer to that of the HL double detector search.

V1 threshold	HLV_{ran}	HLV_{inj}	HL	Offset [deg]	Searched area [deg^2]	90 % area [deg^2]
case 1						
3.0	78 %	0 %	22 %	29	245	440
3.5	14 %	0 %	86 %	21	151	764
4.0	2 %	0 %	98 %	21	145	834
5.5	0.4 %	0 %	99.6 %	21	138	841
case 3						
2.0	66 %	34 %	0 %	13	103	365
3.0	51 %	33 %	16 %	14	109	375
3.3	24 %	32 %	44 %	13	99	495
3.5	11 %	29 %	60 %	13	95	597
3.8	3 %	25 %	72 %	13	92	673
4.0	2 %	23 %	76 %	14	95	694
4.5	1 %	18 %	81 %	15	99	726
5.0	0.1 %	13 %	87 %	16	105	758
5.5	0.1 %	10 %	89.9 %	17	107	789
7.0	0 %	5 %	95 %	19	126	829
8.0	0 %	3 %	97 %	20	129	835
10.0	0 %	1 %	99 %	21	133	835

Table 4.4: Simulation summary in a noisy case

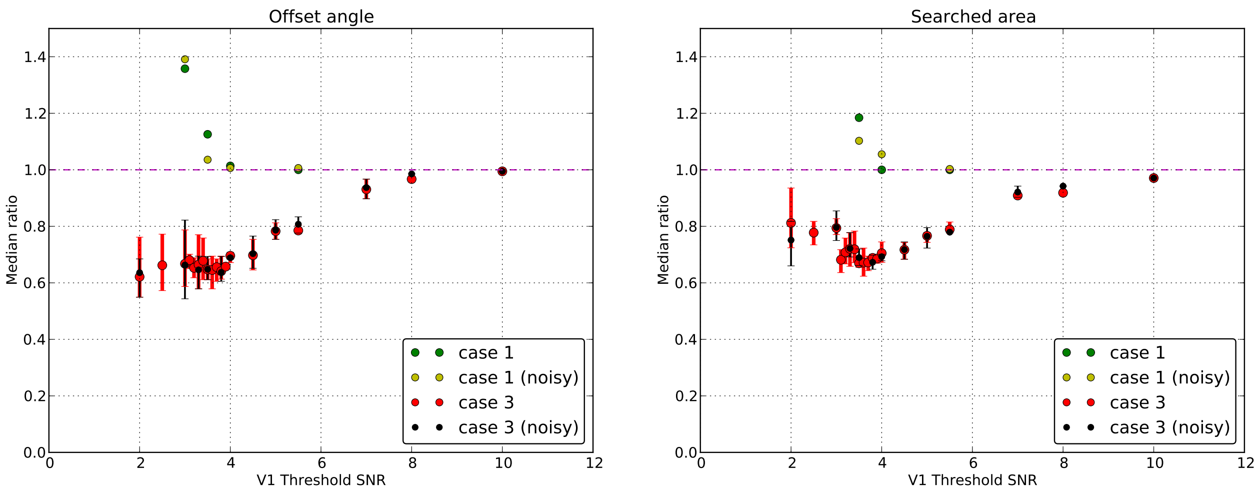

Figure 4.8: Offset angle (*Left*) and searched area (*Right*), when the V1 detector generates a lot of background triggers. In case 3 investigation, calculation is repeated for three times, using different set of random numbers.

According to the offset angle, searched area, and 90 % confidence area, the localization performance does not become worse even if the low sensitivity detector generates a lot of background triggers. The saturation shown in offset angle and the minimum in the searched area depend on the accuracy of the arrival timing and also the number of the HLV_{ran} triggers. The number of the HLV_{ran} triggers depends on the FAP, and thence it is found that the more critical part for the localization performance is the behavior at the region where the FAP per template is larger than around 10^{-2}.

In conclusion, the HLV hierarchical search works even in consideration of the noisy background triggers from the low sensitivity detector.

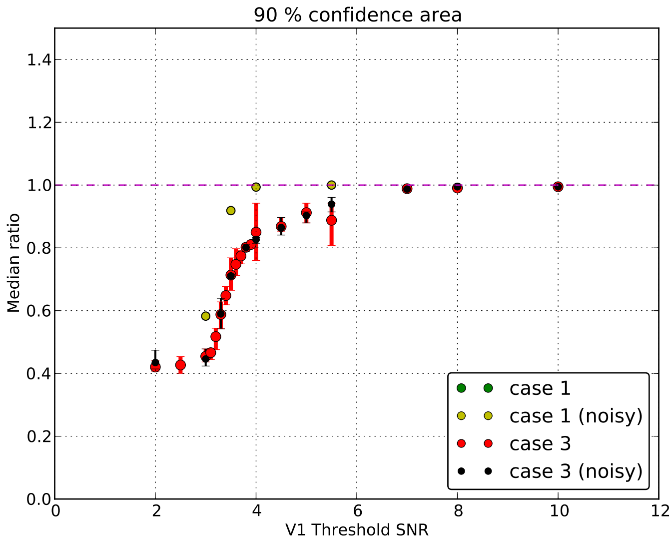

Figure 4.9: 90 % confidence area with HLV hierarchical search, when the V1 detector generates a lot of background triggers. The vertical axis is normalized with the value obtained in HL double search.

4.5 Typical error of sky maps

Above sections explain the statistical merit of the HLV hierarchical network by investigating the discrepancies between the predicted position and the source location. According to the above results, the prediction has a systematic error, such as 13 deg in the offset angle at the optimal SNR threshold. However, it is also important to investigate how the 90 % confidence area and the source position appear in the celestial sphere. This section describes the relation between the predicted 90 % confidence area and the injected source location.

In the hierarchical search, it is found that BAYESTAR can fail to point the injection position within the 90 % confidence area, when it constructs the sky maps from HLV triple coincidence triggers. This situation sometimes happens, especially if the V1 triggers are generated from noise, as shown in figure 4.10. However, this error also happens even if the sky map is generated from triple coincidence trigger with the V1 triggers based on GW signal, as seen in figure 4.11. Table 4.5 summarizes the number of events which the injected location is inside or outside the 90 % confidence area. These numbers are collected when the V1 SNR threshold is set at optimal one.

Trigger type	Inside HL Inside HLV	Inside HL Outside HLV	Outside HL Intside HLV	Outside HL Outside HLV	SUM
HLV_{ran}	22 event	6 events	0 events	1 events	29 events
HLV_{inj}	59 event	10 events	3 events	1 events	73 events

Table 4.5: Number of the events in which the injected location is inside/outside the 90 % confidence area.

According to this table, when the EM counterpart is searched based on the sky map generated from the HLV triple coincidence triggers of hierarchical network, around 20 % of events are found outside its 90 % confidence area. However, even if the EM counterpart does not appear in the 90 % confidence area predicted by HLV triple coincidence trigger, the source is expected to appear within the confidence area from HL double coincidences. Also the source location appears next to the HLV 90 % confidence area in most cases.

Thence following method would help faster identification of the counterpart using EM telescopes. First, one looks for the source based on the sky map predicted by HLV hierarchical network. Then if the source does not appear within the HLV map, one searches the counterpart based on the HL map by expanding the HLV map along the HL prediction.

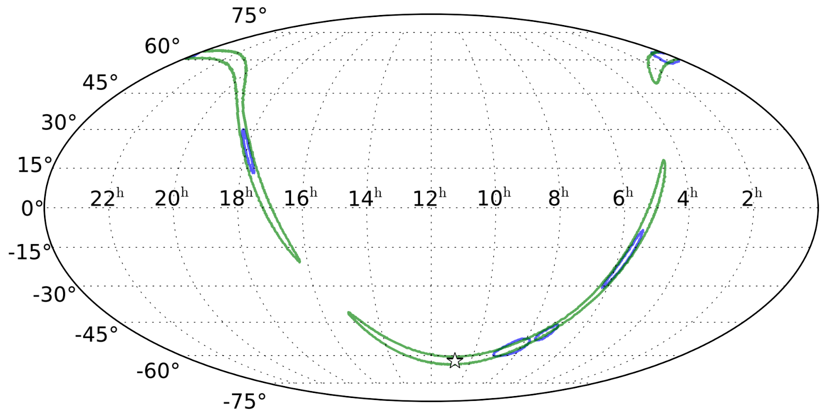

Figure 4.10: Example of sky map error generated from the trigger of H,L and V_{ran} coincidences. The green and blue map are constructed from the trigger HL double coincidences, and the HLV_{ran} triple coincidences.The star indicates the location of the injected source.

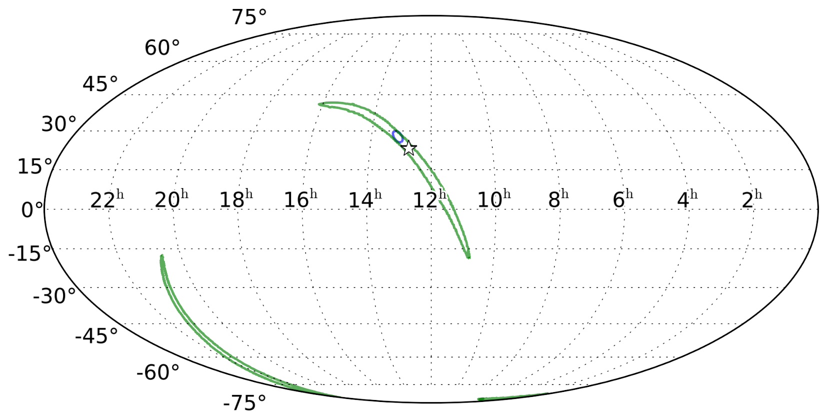

Figure 4.11: Example of sky map error generated from the trigger of H,L and V_{inj} coincidences.The green and blue map are constructed from the trigger HL double coincidences, and the HLV_{inj} triple coincidences.The star indicates the location of the injected source.

Sky localization with hierarchical search of 4 detectors

Chapter 5

Rapid sky localization with the hierarchical search of 4 detector network of H1, L1, V1, and K1 are investigated in this chapter. In this calculation, the sensitivities of each detector are set at 70 Mpc for H1 and L1, while 20 Mpc for V1 and K1 detector. The noise curve for K1 detector is the same as the V1 curve described in chapter 3.3.3. In all cases, the K1 threshold is set at the same value of the V1 threshold for simple investigation.

Chapter 5.1 describes the expected localization with 4 detector hierarchical search, and the optimization of V1 and K1 SNR threshold. In chapter 5.2, contribution of the fourth detector for the localization is explained.

5.1 Optimization of V1, K1 threshold SNR

The investigation of 4 detector hierarchical search is described in this section. This calculation is done in accordance with the procedure explained in chapter 3.6.2. The results of trigger population and expected localization are summarized in table 5.1, 5.2. Figure 5.1 shows the median values of the offset angle and the searched area as a function of the SNR threshold, which are obtained from investigation of 4 detector hierarchical search. The corresponding 90 % confidence area is shown in figure 5.2.

According to figure 5.1 and figure 5.2, the behavior of the offset angle, searched area, and 90 % confidence area are the same as those of 3 detector hierarchical search. The optimal SNR threshold is around 3.5 according to the searched area. The expected typical offset angle, the searched area, and 90 % confidence area at the threshold are 12 deg, 87 deg^2, and 500 deg^2, respectively.

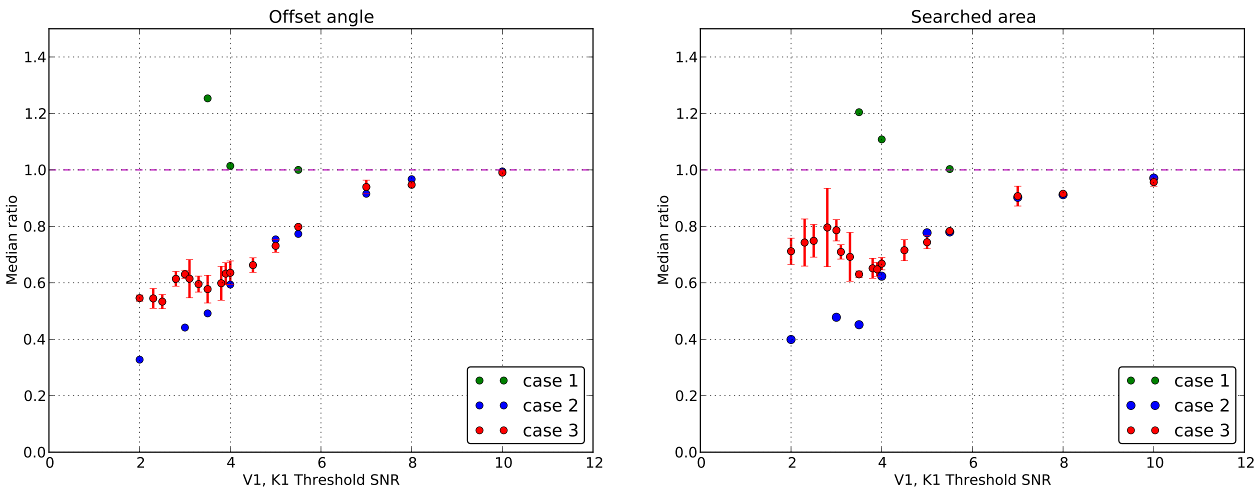

Figure 5.1: Offset angle (*Left*), searched area (*Right*) with 4 detector hierarchical search. In case 3 investigation, calculation is repeated for two times, using different set of random numbers.

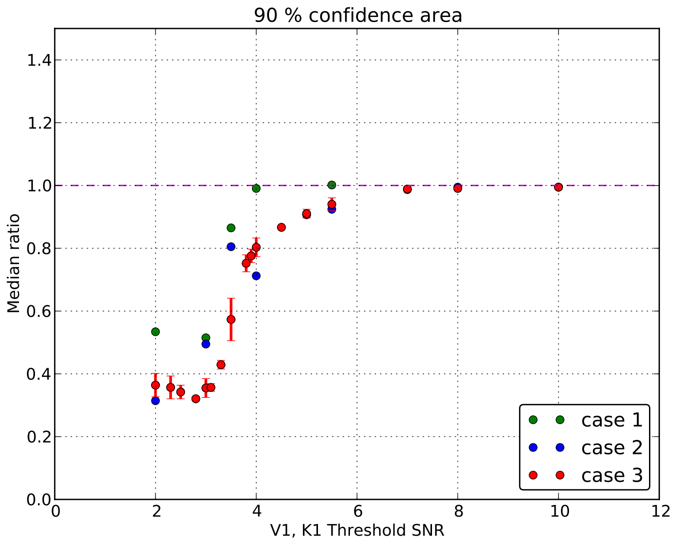

Figure 5.2: 90 % confidence area with 4 detector hierarchical search.

Typical sky maps generated by the triggers from HLVK quadruple coincidences in its hierarchical network are shown in figures 5.3 to 5.6. In these figures, 'ran' and 'inj' denote triggers based noise and injection parameters respectively. If all triggers from each detector in the network are based on GW signals, the maps from the HLVK coincidences are expected to point the location of the source as shown in figure 5.3. Meanwhile, if the trigger from a low sensitivity detector is based on noise, the generated map starts to spread within the 90 % confidence area of HL, in most cases. Also accuracy of the prediction starts to worsen. However, the source is expected to appear next to the confidence area from HLVK trigger. If both of the triggers from the low sensitivity detectors are generated from noise, the HLVK triggered map cannot predict the source position anymore, and this worst situation can be seen in around 5 % of HLVK triggered maps.

This point like prediction[1] is much more useful for the EM counterparts search, compared to the HL search or HLV hierarchical search. This implies adding another low sensitivity detector into the hierarchical network of two more sensitive detectors and one less sensitive detector brings more benefits for fast localization and for looking for EM counterparts.

Table 5.3 describes the number of events in which the injected location is inside or outside the 90 % confidence area from HL or HLVK hierarchical network search.

[1]The red colored map shown in figure 5.3 is the example of the 'point like prediction'.

V1, K1 threshold	Offset [deg]	Searched area [deg^2]	90 % area [deg^2]
case 1			
2.0	35	311	449
3.0	33	271	432
3.5	26	165	726
4.0	21	152	832
5.5	21	138	841
case 2			
2.0	7	55	264
3.0	9	66	416
3.5	10	62	676
4.0	12	86	589
5.0	16	107	761
5.5	16	107	776
7.0	19	124	829
8.0	20	125	835
10.0	21	133	835
case 3			
2.0	11	98	306
2.3	11	102	300
2.5	11	103	287
2.8	13	109	269
3.0	13	108	298
3.1	13	97	300
3.3	12	95	360
3.5	12	87	482
3.8	12	89	632
3.9	13	89	652
4.0	13	92	674
4.5	14	98	728
5.0	15	102	765
5.5	17	108	790
7.0	19	125	831
8.0	20	126	832
10.0	20	131	835

Table 5.1: Simulation summary of HLVK quadruple network

V1, K1 threshold	$HLV_{ran}K_{ran}$	HLV_{ran}	HLK_{ran}	$HLV_{inj}K_{inj}$	HLV_{inj}	HLK_{inj}	$HLV_{ran}K_{inj}$	$HLV_{inj}K_{ran}$	HL
Case 1									
2.0	100	0	0	0	0	0	0	0	0
3.0	59	15	19	0	0	0	0	0	7
3.5	2	11	13	0	0	0	0	0	74
4.0	0	1	2	0	0	0	0	0	97
5.5	0	0	0	0	0	0	0	0	100
Case 2									
2.0	0	0	0	43	16	13	0	0	28
3.0	0	0	0	22	14	15	0	0	48
3.5	0	0	0	15	15	15	0	0	55
4.0	0	0	0	10	12	10	0	0	67
5.5	0	0	0	2	8	7	0	0	83
7.0	0	0	0	2	3	3	0	0	92
8.0	0	0	0	1	1	2	0	0	95
10.0	0	0	0	0	0.5	0.5	0	0	99
Case 3									
2.0	53	0	0	17	0	0	12	18	0
2.3	53	0	0	17	0	0	12	18	0
2.5	53	0	< 0.5	17	0	0	12	18	0
2.8	49	2	2	17	1	1	11	17	0
3.0	31	7	11	16	4	3	10	14	4
3.1	20	11	14	16	6	5	7	11	10
3.3	7	12	15	14	10	8	5	7	22
3.5	1	7	9	12	13	10	2	2	44
3.8	0	3	2	10	12	11	< 0.5	1	61
3.9	0	3	1	9	12	10	< 0.5	< 0.5	64
4.0	0	1	1	9	12	9	< 0.5	< 0.5	66
4.5	0	< 0.5	< 0.5	6	11	9	0	0	74
5.0	0	0	< 0.5	4	9	7	0	0	80
5.5	0	0	< 0.5	4	6	5	0	0	85
7.0	0	0	0	2	2	3	0	0	93
8.0	0	0	0	1	2	2	0	0	95
10.0	0	0	0	1	1	1	0	0	97

Table 5.2: Trigger population of HLVK hierarchical search (in %).

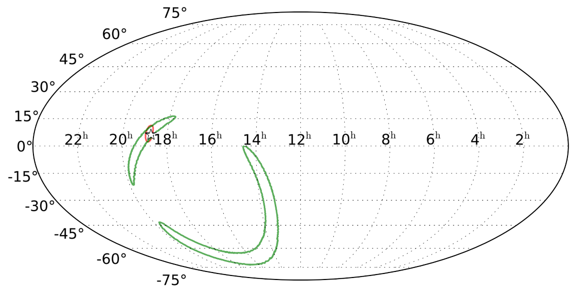

Figure 5.3: $\mathrm{HLV_{inj}K_{inj}}$ sky map(a)

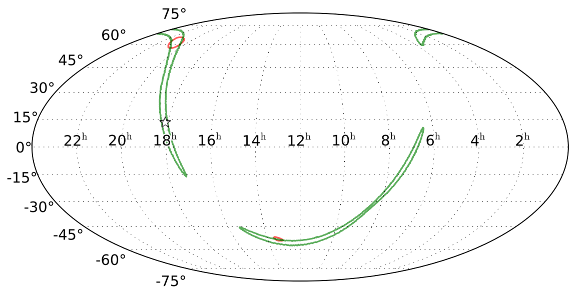

Figure 5.4: $\mathrm{HLV_{ran}K_{ran}}$ sky map(b)

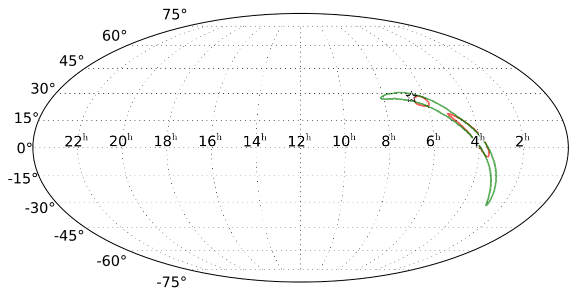

Figure 5.5: $\mathrm{HLV_{ran}K_{inj}}$ sky map(c)

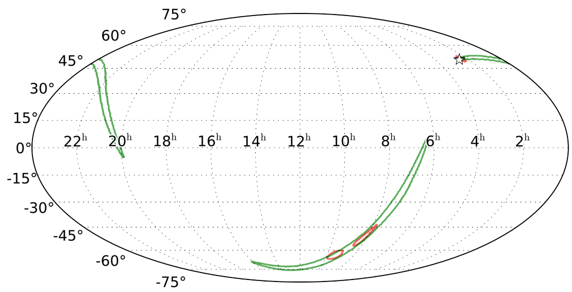

Figure 5.6: $\mathrm{HLV_{inj}K_{ran}}$ sky map(d)

5.2 Contribution of fourth detector

This section describes how the fourth detector contributes to the fast localization with the hierarchical search of 3 detectors. Figure 5.7 and 5.8 show the comparison of the offset angle, searched area, and 90 % confidence area. All of them are obtained from case 3 investigation.

It is found that the optimal thresholds for V1, K1 detectors in the 4 detector search, and the accuracy of the sky maps at the threshold are mostly the same as the 3 detector search. This is because the same parameters as V1 detector are used for KAGRA detector. The accuracy of the sky maps seen in the offset angle and the searched area depend on the number of V1, K1 triggers based on GW signals, and also depend on the added timing uncertainties and sensitivities of each detector. Since same FAP, timing uncertainties, and sensitivity as V1 detector are used for the K1 detector in this investigation, it is expected that the accuracy of the sky maps do not improve so much, compared

Trigger type	Inside HL Inside HLVK	Inside HL Outside HLVK	Outside HL Inside HLVK	Outside HL Outside HLVK	SUM
$HLV_{ran}K_{ran}$	0 event	2 events	0 events	0 events	2 events
$HLV_{rad}K_{inj}$	3 event	2 events	0 events	0 events	5 events
$HLV_{inj}K_{ran}$	5 event	0 events	0 events	0 events	5 events
$HLV_{inj}V_{inj}$	25 event	3 events	1 events	0 events	29 events
HLV_{ran}	15 event	3 events	0 events	0 events	18 events
HLV_{inj}	24 event	3 events	2 events	1 events	30 events
HLK_{ran}	16 event	4 events	0 events	0 events	20 events
HLK_{inj}	17 event	5 events	0 events	0 events	22 events

Table 5.3: Number of events in which the injected location is inside/outside the 90 % confidence area of HLVK, HL search.

to the 3 detector hierarchical search. The obtained result of 4 detector hierarchical search is consistent with this issue.

If different sensitivity is selected for each low sensitivity detector, the optimal threshold for each detector, and the expected localization performance would produce different results.

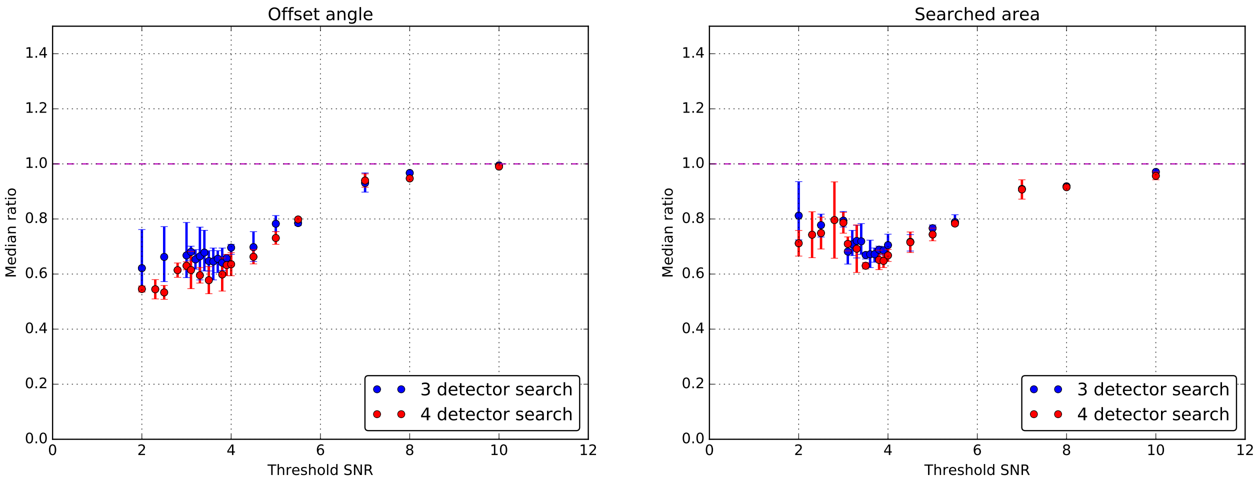

Figure 5.7: Comparison of offset angle (*Left*), searched area (*Right*) between 3 detector hierarchical search and 4 detector search. Offset angle and searched area are reduced from 13 deg, 93 deg^2 to 12 deg, 87 deg^2 at the SNR threshold of 3.5.

On the other hand, 4 detector search improves the 90 % confidence area as shown in figure 5.8. The median value of the confidence area is reduced to around 500 deg^2 at the optimal threshold, which is smaller than HLV hierarchical search by around 100 deg^2.

Even though the systematic prediction errors seen in offset angle and searched area are not improved, the 4 detector search gives more precise sky maps. This would increase opportunities whereby the source objects are found in EM follow up observation.

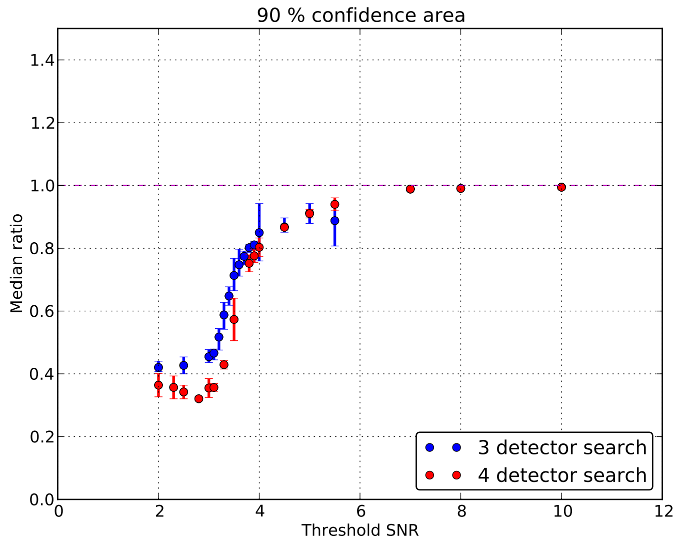

Figure 5.8: Comparison of 90 % confidence area between 3 detector hierarchical search and 4 detector search. At below SNR 4, the area is further reduced due to the contribution of quadruple triggers than 3 detector hierarchical network.

In conclusion, if the hierarchical network is constructed with two high sensitivity detectors, and two low sensitivity ones are implemented into the pipeline, it is expected that same accuracy of the sky maps are obtained as that of three detector hierarchical search with two high sensitivity detectors and one low sensitivity detector. However, the confidence area in the four detector search is reduced compared to the three detector search, and thus the four detector search is more useful for looking for the EM counterpart.

If more accurate sky maps are necessary, the arrival timing uncertainties of the low sensitivity detectors should be improved, so that higher SNR triggers are analyzed. In addition, more stable operation is necessary so that they do not to generate a lot of back ground triggers.

Conclusion and future work

Chapter 6

6.1 Conclusion

The fast localization with hierarchical network is demonstrated in this work. This is the first time to demonstrate this localization method with 3 and 4 detectors, quantitatively. The expected performance of 3 and 4 detector network search at the optimal thresholds are summarized in table 6.1. In the actual situation, the sensitivities of all the four detectors which are considered in this investigation are expected to be improved. Especially, as for the fast localization with 4 detector hierarchical search, more improvement in performance is expected, when the search is implemented into the pipeline.

	Offset [deg]	Searched area [deg^2]	90 % area [deg^2]
HL search	21	137	840
HLV search	13	93	600
HLVK search	12	87	500

Table 6.1: Expected localization performance with the hierarchical network. This table describes median values at the optimal SNR threshold, when the sensitivity of H1, L1 is 70 Mpc, while that of V1, K1 is set at 20 Mpc.

The optimal SNR threshold for the low sensitivity detectors are obtained at around 3.5, according to the accuracy of the sky maps. Based on these results, it is confirmed that the hierarchical network is one effective method to construct the detection network with different sensitivity detectors. This searching method is useful for newly constructed detectors.

6.2 Future work

In this work, the two LIGO detectors are assumed to maintain its work. However, H1 and L1 detector can discontinue their operation for some reason. If GW signals come in this condition, information from double (or more) coincidences is to be generated from different sensitivity detectors. In this case, for making effective use of the low sensitivity detectors, it is expected to set high SNR threshold for higher sensitivity detectors, while lower threshold for the less sensitive ones. This method is to be investigated from the point of view of the sky maps.

The searched area as a function of the SNR threshold for the low sensitivity detector, has a minimum value as shown in figure 4.4. It is expected that it depends on the number of the V1 triggers from noise, and the accuracy of the added timing uncertainties. However, it is still not described theoretically. The theoretical prediction of the localization performance can simplify the finding procedure of the optimal thresholds of less sensitive detectors in the hierarchical search. It can also help making detailed plan for adding newly constructed detectors into the network, and thence this is also to be investigated.

Bibliography

[1] B. Allen, et al., Phys. Rev. D, 85, 122006 (2012).

[2] J. Abadie, et al., arXiv:1203.2674 (2012).

[3] T. Adams, et al., arXiv:1512.02864 (2015).

[4] L. Singer, et al., arXiv:1508.03634 (2016).

[5] J. Abadie et al., A&A, 541, A155 (2012).

[6] J. Aasi et al., Phys. Rev. D, 88, 062001 (2013).

[7] M. Maggiore, Physics Reports 331, 283, (2000).

[8] L. Singer, PhD Thesis, California Institute of Technology, (2015).

[9] L. Singer, et al., The Astrophysical Journal 795, 105, (2014).