Doctorate Dissertation
博士論文

New parameter estimation method being free from the
bias depending on sky region for Gravitational wave
from compact binary coalescence
(正則化手法を用いたコンパクト連星からの重力波の
パラメター推定精度改善の研究)

A Dissertation Submitted for Degree of Doctor of
Science
December 2018

平成30年12月博士(理学)申請
Department of Physics, Graduate School of Science,
The University of Tokyo

東京大学大学院理学系研究科物理学専攻

Kenji Ono

小野 謙次

Abstract

Since LIGO Scientific collaboration and Virgo collaboration has successfully detected Gravitational waves (GWs) directly from a compact binary coalescence (CBC) comprising two black holes, new era of the astronomy - GW astronomy - has begun. Astronomers and physicists expect that Gravitational wave astronomy will reveal phenomena that have not been previously clarified via electromagnetic astronomy. The precise discussion of the GW astronomy requires the precise and homogeneous estimation of the GW parameters inserting from all over the sky. Detection of GWs involves using the network of GW telescopes to observe a large sky region. However, owing to the arrangement of the GW telescopes, parameter estimation accuracy deteriorates depending on the sky region of the GW source due to the ill-posed nature of the inverse operator. The instability of the solution of the ill-posed inverse problem causes the amplifying the amount of error in the result of the inverse problem even if the error in the given data from GW telescopes is small.

One key method that the deteriorated accuracy of the parameters makes improve is called a regularization method. A regularization method provides the mathematical framework to solve the ill-posed inverse problem stably by adding an appropriate correction term to the ill-posed operator. To avoid the ill-posed problem, certain solutions are suggested for detecting GWs by the network of GW telescopes. However, conventional regularization methods suggested for a GW data analysis focus on reducing the residual noise and ignore the fact that estimated GW parameters can exceed the value range of the actual parameters because the regulator adds bias noise. In other words, conventional regularization methods cannot optimize all regulator parameters completely.

Our study propose the new parameter estimation method to minimize a whole of the noise including the bias noise and amplified noise due to ill-posed inverse problem by optimizing all regulator parameters. To obtain the optimized parameters for the estimation of amplitude parameters of GWs,

the residual of the amplitude parameters of GWs, which is expressed by the norm of the difference between the actual amplitudes and the estimated amplitudes evaluated using regularized data analysis, must be minimized. The problem of the minimization is that the actual GW amplitudes cannot be predetermined and the estimated point of GW parameters using a regularization method frequently lying outside of the residuals when the bias error exceeds the amplified error. To resolve these problems, we propose a Lagrange multiplier method with KKT condition for the norm of the difference between the amplitude parameters estimated by the regularized data analysis and amplitude parameters estimated by the non-regularization method, i.e., the a-posteriori parameter choice rule provides optimized regulator values.

The data analysis based on a Bayesian analysis is implemented by using MultiNest software, which is the Bayesian inference software based on the nested sampling algorithm. The data analysis results indicate that the regularization method with the type 2 regulator reduces the credible region of the accuracy of the amplitude parameters. For approximately 90% of the sky region, the credible region of inclination-distance is reduced by approximately 1.5 times and that of the polarization-initial phase is reduced by approximately 3.0 times. The shrinkage rate of the credible region increases with a decreasing determinant value of the inverse operator So we demonstrate that the proposed method can shrink the credible regions of inclination vs luminosity distance and polarization vs initial phase significantly in the sky wherein the accuracy of the amplitude parameters of GWs has been deteriorated.

The proposed method suppresses the systematic error of GWs depending on the sky region and allows us investigating the cosmological information more precisely.

Contents

Chapter 1

Introduction

Gravitational waves(GWs) have been detected using extremely large laser interferometers. i.e., GW telescopes. Km-scale GW telescopes, such as LIGO[1] and Virgo[2] have been constructed and the KAGRA[3] is currently under construction in Japan. LIGO Scientific collaboration and Virgo collaboration had successfully detected GWs directly from a compact binary coalescence (CBC) comprising two black holes[4]. Astronomers and physicists expect that GW astronomy will reveal phenomena that have not been previously clarified via electromagnetic astronomy.

The detection of the GW event GW170817 from the coalescence of NS-NS binary at 2017 August 17 by the network of the GW telescopes composed of LIGO Hanford, LIGO Livingston and Virgo[5][6], short gamma-ray burst event GRB 170817A observed by Fermi-GBM[7][8] and any other electromagnetic counterparts[9] opened the new era of multi-messenger astrophysics, which is the joint observation of gravitational wave and electromagnetic radiation from single source[9]. The multi-messenger astrophysics is expected to reveal the detail of the astronomical and cosmological phenomena.

In the astronomical viewpoint, exotic X-ray transient events, for instance, whose the nature mechanism of the emission remains unclear have been reported[10][11]. In particular, CDF-S XT1 is one of the latest exotic X-ray transient events whose cannot detect associated multiwavelength transient emission expected to be able to observe according to conventional researches[12]. While there are models to explain the properties of the source[12][13][14], the inclination parameter, which can be estimated directly by the precise determination of the amplitude parameters of GWs, is crucial to classify wheather the model is correct or not.

The measurement of the GW amplitude from CBC provides the method to determine the distance from the source of a transient to the earth directly,

which is called "standard siren"[15][16]. In the cosmological viewpoint, the measurement of the distance and redshift, which is determine by using electromagnetic observation, of the source is important to determine a Hubble constant[15][17][18]. The estimation of the Hubble constant has been conducted actually by using GW170817 data; the distance of the source is determined by amplitude parameter of GWs and the redshift is measured by the host galaxy NGC4993 determined by the celestial coordinate of the transient measured by electromagnetic radiations[19]. The measurement of the Hubble constant has been conducted by electromagnetic observation such as the cosmic microwave background(CMB)[20] and cosmic distance ladder[21]. However, these independent measurements indicate that the 3.4σ(99.9% confidence) difference of the measured Hubble constants between the result of CMB and cosmic distance ladder. Such differences suggest that the necessity of additional independent measurements of the Hubble constant by using standard siren, which doesn't depend on the electromagnetic observation.

For implementing these important study, it is expected that more precise estimation of GW waveforms or parameters will be achieved using a network of GW telescopes[22], i.e., using multiple telescopes simultaneously to detect GWs. With a network of GW telescopes, the SNR of GWs detection can be increased and the independent mode of GWs can be determined. An analysis method based on Bayesian statistics has been proposed[23] to estimate GW waveforms detected by a network of GW telescopes. To estimate the amplitude parameters of a GW, maximizing the likelihood of the output of GW telescopes and the GW model are equivalent to solving an inverse problem whose inverse operator considers the parameters of antenna-beam pattern functions and detecting the SNR of a GW[23]. However, the solution of an inverse problem is unstable owing to the rank deficiency of the inverse operator. The instability of the solution, i.e., an ill-posed problem, makes it impossible to distinguish the independent modes of GW due to the degeneration of these modes. In order to avoid the ill-posed problem, certain solutions are suggested for detecting GWs by the network of GW telescopes, especially for burst search of GWs. For instance, Rakhmanov formulated the Maximum likelihood method with Tikonov regularization[24] and Mohanty propose to constract a regulator by a variability of the SNR as the source is displaced on the sky[25].

However, conventional regularization methods for a GW analysis focus on reducing the residual noise and ignore the fact that estimated GW parameters can exceed the value range of the actual parameters because the regulator adds bias noise[26]. This is a significant problem because it is highly likely that the actual GW parameters are outside the credible region

estimated using a regularization method. The bias noise introduced by a regulator can cause inaccurate estimation of parameter values. To implement precise estimation of GW parameters, a regularization method is required to reduce the amplified noise as small as possible while maintaining the bias noise such that the estimation points of the GW parameters are not affected. Thus, we propose a method to optimize regulator parameters for minimizing the influence of amplitude parameters amplified by the ill-posed inverse operator in the analysis of a targeted CBC search. The proposed method attempts to minimize the residual of the amplitude parameters expressed by the sum of amplified and bias noise. The residual of the amplitude parameters of GWs, which is expressed by the norm of the difference between the actual amplitudes and the estimated amplitudes evaluated using regularized data analysis, must be minimized to obtain all of the optimized parameters of a regulator. However, the actual GW amplitudes cannot be predetermined. Thus, the optimized regulator parameters are selected based on an a-posteriori parameter choice rule[26]. The estimated amplitude allows us to evaluate the value of the norm. Then, it becomes possible to determine full optimized parameters of a regulator. We implement data analysis with the optimized regulator to improve the accuracy of the amplitude parameters of GWs.

The remainder of this paper is organized as follows. In Chapter.2, we describe the fundamental formulation of a targeted coherent CBC search. The data analysis method using this thesis is discussed in Chapter.3. We explain the mathematical framework of the regularization method in Chapter.4 and how the optimized regulator in the coherent search method is determined using an a-posteirori parameter choice rule in Chapter.5. In Chapter.6, we review the regularized data analysis algorithm and describe MultiNest[27][28], which is Bayesian inference software. In addition, we present the results of regularized data analysis and describe the behavior of the reduction of the amplified noise. Conclusions are given in Chapter.7.

Chapter 2

Fundamental theory of Gravitational waves

2.1 A brief review of Gravitational waves

In this section, we describe the brief summarize of a GW theory. We refer to [29] and [30] to describe the details of the section.

All the classical gravitational theory including GWs is based on the Einstein Equation such that

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{2.1}$$

where $g_{\mu\nu}$ is the metric tensor of space-time, which is the function of the space-time coordiantes x^α. It also defined by using the invariant squared distance element ds^2 and the infinitisimal coordinate difference between the two points dx^α as

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu. \tag{2.2}$$

The metric tensor is invariant under transformation of an infinitisimal coordinate, any physical quantity doesn't depend on the choice of the coordinate system.

Furthermore, The Ricci tensor $R_{\mu\nu}$ is defined as

$$R_{\mu\nu} = \partial_\rho \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\mu\rho} + \Gamma^\rho_{\mu\nu}\Gamma^\sigma_{\rho\sigma} - \Gamma^\rho_{\mu\sigma}\Gamma^\sigma_{\nu\rho}, \tag{2.3}$$

where $\Gamma^\rho_{\mu\nu}$ is a Christoffel symbol define as

$$\Gamma^\mu_{\rho\sigma} = \frac{1}{2}g^{\mu\nu}\left[\partial_\rho g_{\sigma\nu} + \partial_\sigma g_{\rho\nu} - \partial_\mu g_{\rho\sigma}\right]. \tag{2.4}$$

The Ricci scalar R is defined as

$$R = g^{\mu\nu} R_{\mu\nu}. \tag{2.5}$$

A GW theory is based on the linear perturbed theory of the gravitational field. Here, we consider the background gravitational field is flat, which can be expressed by the Minkowski space whose metric is such that

$$g_{\text{back}\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.6}$$

Since a GW metric denoted by $h_{\mu\nu}$ is the perturbation on the background metric such that $\|\overline{h}_{\mu\nu}\| \ll \|g_{\text{back}\mu\nu}\|$, the metric on the space is

$$g_{\mu\nu} = \eta_{\mu\nu} + \overline{h}_{\mu\nu}. \tag{2.7}$$

By substituting the metric Eq.(2.7) for the Einstein theory Eq.(2.1), we obtain the linear Einstein equation such that

$$\Box h_{\mu\nu} + \eta_{\mu\nu}\partial^\rho \partial^\sigma h_{\rho\sigma} - \partial_\nu \partial^\rho h_{\mu\rho} - \partial_\mu \partial^\rho h_{\nu\rho} = -\frac{16\pi G}{c^4} T_{\mu\nu}, \tag{2.8}$$

where

$$h_{\mu\nu} = \overline{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\eta^{\alpha\beta}\overline{h}_{\alpha\beta}, \tag{2.9}$$

$\Box$ is the d'Alembert operator on the flat space such that

$$\Box = -\frac{\partial^2}{\partial t^2} + \Delta, \tag{2.10}$$

Δ is the laplacian operator on the flat space and $\partial/\partial t \equiv c\partial_0$, $\partial/\partial \mathbf{x}_i \equiv \partial_i$, where $i = 1, 2, 3$ is the index corresponding to the spatial coordinate of the metric. To remove the degree of freedom, we impose the Lorentz gauge condition such that $\partial^\nu h_{\mu\nu} = 0$ on the Eq.(2.8), we obtain the simple wave equation of the GWs

$$\Box h_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu} \tag{2.11}$$

The general solution of the linear Einstein equation with the Lorentz gauge condition Eq.(2.11) is obtained by using the retarded Green function

$$\Box G^+(x - x') = \delta^4(x - x'), \tag{2.12}$$

5

and the result is

$$h_{\mu\nu}(x) = -\frac{16\pi G}{c^4} \int d^4 x' G^+(x-x') T_{\mu\nu}(x'), \qquad (2.13)$$

where x is the coordinate of the observer and x' is the coordinate of the source of gravity. The solution of the retarded Green function, which is obtained by the theory of a complex variables function, is

$$G^+(x-x') = -\frac{1}{4\pi|\mathbf{x}-\mathbf{x}'|} \delta(t - |\mathbf{x}-\mathbf{x}'| - t'). \qquad (2.14)$$

By substituting Eq.(2.14) for Eq.(2.13) and integrating it with respect to the time of the coordinate of the source of gravity t', we obtain the general solution of the GW function such that

$$h_{\mu\nu}(x) = \frac{4G}{c^4} \int d^3 x' \frac{1}{|\mathbf{x}-\mathbf{x}'|} T_{\mu\nu}(t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}, \mathbf{x}'). \qquad (2.15)$$

Here, we impose the hypothesis, which is hold all over the source of the GWs, that the size of the source d is far smaller than in scale than the distance between the observer and the source $|\mathbf{x}| \equiv r$ such that

$$r \gg d. \qquad (2.16)$$

By performing Taylor expansion of $|\mathbf{x}-\mathbf{x}'|$ around the position of the source $x' = 0$, we obtain

$$|\mathbf{x}-\mathbf{x}'| = r - \mathbf{x}' \cdot \hat{\mathbf{n}} + \mathcal{O}\left(\frac{d^2}{r}\right), \qquad (2.17)$$

where $\hat{\mathbf{n}}$ is the spatial unit vector propagating the GWs in the direction outside the source. Eq.(2.17) provides the simple and appropriate approximation formulation of the GW Eq.(2.15) such that

$$h_{\mu\nu}(x) = \frac{4G}{c^4 r} \int d^3 x' T_{\mu\nu}(t - \frac{r}{c} + \frac{\mathbf{x}'\hat{\mathbf{n}}}{c}, \mathbf{x}') + \mathcal{O}(1/r). \qquad (2.18)$$

While we have reduced 10 degrees of freedom of the GWs $h_{\mu\nu}$ to 6 degree of freedom by imposing the Lorentz gauge condition which reduces the degree of freedom with respect to the first order partial derivation of the perturbation of gravity, we have rest the degree of freedom under the transformation of the coordinate system. To reduce that, we impose the TT(Transverse-traceless) gauge whose operator is such that

$$\Lambda_{ij,kl}(\hat{\mathbf{n}}) = P_k^i P_l^j - \frac{1}{2} P^{ij} P_{kl}, \qquad (2.19)$$

$$P^{ij} \equiv \delta^{ij} - \hat{n}^i \hat{n}^j \qquad (2.20)$$

on the coordinate system. The physical property of the TT gauge(or TT flame) is that particle in the flame is at rest before the arrival of the GWs remains at rest even after the arrival of the GWs. In other words, the free falling flame of the test masses remains at rest while the coordinate of the TT flame is stretch themselves by responding the perturbation of the space time as the GWs passes through. Therefore the GWs imposing the Lorentz gauge condition and TT gauge condition $h_{\mu\nu}^{TT}$ is expressed by

$$h_{ij}^{TT}(x) = \frac{4G}{c^4 r}\Lambda_{ij,kl}(\hat{\mathbf{n}}) \int d^3x' T^{kl}(t - \frac{r}{c} + \frac{\mathbf{x}'\hat{\mathbf{n}}}{c}, \mathbf{x}'). \quad (2.21)$$

Since imposing TT gauge removes 4 degree of freedom, h_{ij}^{TT} has the 2 degree of freedom. It means that the GWs have 2 independent wave modes, which are denoted h_+ and $h_\times$. These mode are called the amplitude of the "plus" and "cross" polarization of the wave. Fig. 2.1 shows the deformation of a ring of test masses due to the plus and cross mode polarization.
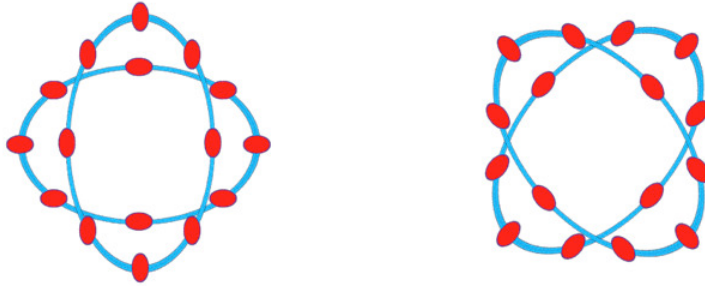
Figure 2.1: The response of the test masses by plus and cross mode of the GWs.

To apply Eq.(2.18) to the GW emission from the CBC system, we use another approximation method called Low-velocity expansion. In a non-relativistic system of the source, the typical velocity of the source v is always much smaller than light speed c such that $v \ll c$. The frequency of the GWs ω we aim to observe is of order the rotational velocity of the source $\omega \approx v/d$. Then typical wavelength of the GWs $\lambda \approx cd/v$ has the relationship such that

$$\begin{aligned} \lambda &\gg d, \\ \frac{\omega}{c}\mathbf{x}' \cdot \hat{\mathbf{n}} &< \frac{\omega_s d}{c} \ll 1, \end{aligned} \quad (2.22)$$

where ω_s is the rotational velocity of the system. Therefore, the energy-momentum tensor $T_{\mu\nu}$ in the Eq.(2.21) can be expanded with respect to

Eq.(2.22), we can find

$$T_{\mu\nu}\left(t - \frac{r}{c} + \frac{\mathbf{x}' \cdot \hat{\mathbf{n}}}{c}, \mathbf{x}'\right) = \sum_n \frac{(\mathbf{x}' \cdot \hat{\mathbf{n}})^n}{c^n n!} \frac{\partial^n}{\partial t^n} T_{\mu\nu}\left(t - \frac{r}{c}, \mathbf{x}'\right). \qquad (2.23)$$

Here, we define the momentum of the stress tensor T^{ij}, where the index $i, j = 1, 2, 3$ is the spatial component of the metric space, as a following:

$$S^{kl, i_1 i_2 \cdots i_n}(t) = \int d^3x \, T^{kl}(t, \mathbf{x}) x^{i_1} x^{i_2} \cdots x^{i_n}, \qquad (2.24)$$

where the first and second indexes separated by comma commute respectively, but we cannot exchange the first and second indexes each other. By substituting Eq.(2.24) for Eq.(2.23) and Eq.(2.21), we obtain the GW function imposing the Low-velocity expansion such that

$$h_{ij}^{TT}(x) = \frac{4G}{c^4 r} \Lambda_{ij,kl}(\hat{\mathbf{n}}) \sum_\alpha \frac{1}{c^\alpha \alpha!} (\partial_0^\alpha S^{kl, i_1 \cdots i_\alpha}) n_{i_1} \cdots n_{i_\alpha}. \qquad (2.25)$$

From the definition of the momentum of the stress tensor Eq.(2.24), the size of the n-th order moment can be estimated as $\mathcal{O}(d^n)$ since the x^{i_k} in the moment correspond to the coordinate of the source of GWs. Furthermore, the n-th order derivative with respect to time in Eq.(2.25) introduces the n-th order of the rotational frequency of the source. In summarize, n-th order of the component of Eq.(2.25) is of order $\mathcal{O}(v^n/c^n)$. Because of $v \ll c$ on the Low-velocity expansion condition, the appropriate approximation of the GW formalism can be estimated by evaluating low degree of the component of the moment.

The appropriate explanation of the physical properties of the moment is provided by a tensor spherical harmonics. First, we define that $\mathbf{L}$ denotes an orbital angular momentum operator, $\mathbf{S}$ denotes a spin operator and $\mathbf{J} = \mathbf{L} + \mathbf{S}$ denotes a total angular momentum. Second, we use the eigenfunctions of the tensor spherical harmonics denoted by $Y_{jj_z}^{ls}(\theta, \phi)$. The eigenfunctions satisfy

$$\mathbf{J}^2 Y_{jj_z}^{ls} = j(j+1) Y_{jj_z}^{ls}, \qquad (2.26)$$
$$J_z^2 Y_{jj_z}^{ls} = j_z Y_{jj_z}^{ls}, \qquad (2.27)$$
$$\mathbf{L}^2 Y_{jj_z}^{ls} = l(l+1) Y_{jj_z}^{ls}, \qquad (2.28)$$
$$\mathbf{S}^2 Y_{jj_z}^{ls} = s(s+1) Y_{jj_z}^{ls}, \qquad (2.29)$$

where l is an azimuthal quantum number, s is a spin quantum number, j is a total angular momentum quantum number and j_z is a magnetic momentum quantum number.

8

The explicit form of the tensor spherical harmonics can be obtained coupling the spherical harmonics Y_{ll_z} to the spin function χ_{ss_z} with the Clebsch-Gordan coefficient such that

$$Y_{jj_z}^{ls}(\theta, \phi) = \langle sls_z l_z | j j_z \rangle Y_{ll_z}(\theta, \phi) \chi_{ss_z} \tag{2.30}$$

Note that the solution of the spin function is

$$\mathbf{S}\chi_{ss_z} = s(s+1)\chi_{ss_z}. \tag{2.31}$$

We are very interested in spin 2 state of the tensor spherical harmonics to describe the behavior of the GWs. The spin function of wave for $s = 2$ with definite value of s_z denoted by $t_{ik}^{(s_z)}$ is obtained combining the two spin-1 function of wave $\xi_i^{m_1}$ and $\xi_k^{m_2}$ with appropriate Clebsch-Gordan coefficients such that

$$t_{ik}^{(s_z)} = \sum_{m_1=-1}^{1} \sum_{m_2=-1}^{1} \langle 11 m_1 m_2 | 2 s_z \rangle \xi_i^{m_1} \xi_k^{m_2}, \tag{2.32}$$

where

$$\xi^{\pm 1} = \mp \frac{1}{\sqrt{2}}(\mathbf{e}_x \pm i\mathbf{e}_y), \qquad \xi^0 = \mathbf{e}_z, \tag{2.33}$$

and $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ are independent unit vector respectively.

The tensor $t_{ik}^{s_{[z]}}$ has 5 tensors depending on $s_z = 0, \pm 1, \pm 2$. These tensors are symmetric and traceless. By substituting Eq.(2.32) for Eq.(2.30), we obtain the spin-2 tensor spherical harmonics:

$$\begin{aligned}
(\mathbf{T}_{jj_z}^l)_{ik} &\equiv (Y_{jj_z}^{l2})_{ik} \\
&= \sum_{l_z=-l}^{l} \sum_{s_z=-2}^{2} \langle 2l s_z l_z | j j_z \rangle Y_{ll_z}(\theta, \phi) t_{ik}^{(s_z)}
\end{aligned} \tag{2.34}$$

From the classification of the Thorne[31], the five tensor spherical harmonics are given by

$$\mathbf{T}_{jj_z}^{S0} = a_{11} \mathbf{T}_{jj_z}^{j+2} + a_{12} \mathbf{T}_{jj_z}^{j} + a_{13} \mathbf{T}_{jj_z}^{j-2}, \tag{2.35}$$

$$\mathbf{T}_{jj_z}^{E1} = a_{21} \mathbf{T}_{jj_z}^{j+2} + a_{22} \mathbf{T}_{jj_z}^{j} + a_{23} \mathbf{T}_{jj_z}^{j-2}, \tag{2.36}$$

$$\mathbf{T}_{jj_z}^{E2} = a_{31} \mathbf{T}_{jj_z}^{j+2} + a_{32} \mathbf{T}_{jj_z}^{j} + a_{33} \mathbf{T}_{jj_z}^{j-2}, \tag{2.37}$$

$$\mathbf{T}_{jj_z}^{B1} = b_{11} i \mathbf{T}_{jj_z}^{j+1} + b_{12} i \mathbf{T}_{jj_z}^{j-1}, \tag{2.38}$$

$$\mathbf{T}_{jj_z}^{B2} = b_{21} i \mathbf{T}_{jj_z}^{j+1} + b_{22} i \mathbf{T}_{jj_z}^{j-1}, \tag{2.39}$$

for $j \geq 2$. The coefficient included by the classification is in Table. 2.1.

a_{11}	$\left[\dfrac{(j+1)(j+2)}{(2j+1)(2j+3)}\right]^{1/2}$	a_{12}	$-\left[\dfrac{2j(j+1)}{3(2j-1)(2j+3)}\right]^{1/2}$
a_{13}	$\left[\dfrac{2j(j-1)}{(2j-1)(2j+1)}\right]^{1/2}$	a_{21}	$-\left[\dfrac{2j(j+2)}{(2j+1)(2j+3)}\right]^{1/2}$
a_{22}	$-\left[\dfrac{3}{(2j-1)(2j+3)}\right]^{1/2}$	a_{23}	$\left[\dfrac{2(j-1)(j+1)}{(2j-1)(2j+1)}\right]^{1/2}$
a_{31}	$\left[\dfrac{j(j-1)}{2(2j+1)(2j+3)}\right]^{1/2}$	a_{32}	$\left[\dfrac{3(j-1)(j+2)}{(2j-1)(2j+3)}\right]^{1/2}$
a_{33}	$\left[\dfrac{(j+1)(j+2)}{2(2j-1)(2j+3)}\right]^{1/2}$		
b_{11}	$\left[\dfrac{j+2}{2j+1}\right]^{1/2}$	b_{12}	$-\left[\dfrac{j-1}{2j+1}\right]^{1/2}$
b_{13}	$-\left[\dfrac{j-1}{2j+1}\right]^{1/2}$	b_{14}	$-\left[\dfrac{j+2}{2j+1}\right]^{1/2}$

Table 2.1: The coefficients of the Thorne's classification.

These tensors can be expressed in terms of the scalar spherical harmonics as follows:

$$(\mathbf{T}_{lm}^{S0})_{ik} = [n_i n_j - (1/3)\delta_{ij}]Y_{lm}, \qquad (2.40)$$

$$(\mathbf{T}_{lm}^{E1})_{ik} = c_l^{(1)}(r/2)(n_i\partial_j + n_j\partial_i)Y_{lm}, \qquad (2.41)$$

$$(\mathbf{T}_{lm}^{B1})_{ik} = c_l^{(1)}(i/2)(n_iL_j + n_jL_i)Y_{lm}, \qquad (2.42)$$

$$(\mathbf{T}_{lm}^{E2})_{ik} = c_l^{(2)}r^2\Lambda_{ij,i'j'}(\hat{\mathbf{n}})\partial_{i'}\partial_{j'}Y_{lm}, \qquad (2.43)$$

$$(\mathbf{T}_{lm}^{B2})_{ik} = c_l^{(2)}r\Lambda_{ij,i'j'}(\partial_{i'}L_{j'} + \partial_{j'}L_{i'})Y_{lm}, \qquad (2.44)$$

where

$$c_l^{(1)} = \left[\frac{2}{l(l+1)}\right]^{1/2}, \qquad c_l^{(2)} = \left[2\frac{(l-2)!}{(l+2)!}\right]^{1/2}. \qquad (2.45)$$

In the classification of the tensor spherical harmonics, $\mathbf{T}_{lm}^{S0}$ has $l \geq 0$, $\mathbf{T}_{lm}^{E1}$ and $\mathbf{T}_{lm}^{B1}$ have $l \geq 1$ and $\mathbf{T}_{lm}^{E2}$ and $\mathbf{T}_{lm}^{B2}$ have $l \geq 2$. Furthermore, the transversality of the GWs $n^i h_{ij}^{TT} = 0$, which comes from the hypothesis that the graviton is massless, eliminates 3 degree of freedom from the 5 tensors. It allows us to remove $\mathbf{T}^{S0}$, $\mathbf{T}^{E1}$ and $\mathbf{T}^{B1}$ from the basis of the GWs since these don't have the transversality. As a result of the discussion, we obtain another general solution of the GWs such as

$$h_{ij}^{TT} = \frac{1}{r}\frac{G}{c^4}\sum_{l=2}^{\infty}\sum_{m=-m}^{l}\left[u_{lm}(\mathbf{T}_{lm}^{E2})_{ij} + v_{lm}(\mathbf{T}_{lm}^{B2})_{ij}\right]. \qquad (2.46)$$

10

Comparing Eq.(2.25) and Eq.(2.46), we can determine the coefficient u_{lm} and v_{lm} of the Eq.(2.46). To multiply both side of Eq.(2.25) and Eq.(2.46) and integrate over the surface of the sphere, we obtain

$$u_{lm} = \sum_\alpha \frac{4}{\alpha!} (\partial_0^\alpha S^{ij,i_1\cdots i_\alpha}) \int d\Omega (\mathbf{T}_{lm}^{E2})^*_{ij} n_{i_1} \cdots n_{i_\alpha}, \qquad (2.47)$$

and

$$v_{lm} = \sum_\alpha \frac{4}{\alpha!} (\partial_0^\alpha S^{ij,i_1\cdots i_\alpha}) \int d\Omega (\mathbf{T}_{lm}^{B2})^*_{ij} n_{i_1} \cdots n_{i_\alpha}. \qquad (2.48)$$

Since Eq.(2.34) indicates $\mathbf{T}_{jj_z}^l \propto Y_{lm}$ and the constant STF tensor of the spherical harmonics(see Appendix.A) is expressed by Eq.(A.24), we obtain

$$\int d\Omega \mathbf{T}_{jj_z}^l n_{i_1} \cdots n_{i_\alpha} \quad \propto \quad \int d\Omega \mathbf{T}_{lm} n_{i_1} \cdots n_{i_\alpha},$$

$$= \quad \mathcal{Y}_{n_{j_1}\cdots n_{j_l}}^{lm} \int d\Omega n_{j_1} \cdots n_{j_l} n_{i_1} \cdots n_{i_\alpha}. \qquad (2.49)$$

Furthermore, the integral of the unit vectors over polar coordinate is

$$\frac{1}{4\pi} \int d\Omega n_{i_1} \cdots n_{i_{2l}} = \frac{1}{(2l+1)!!} \left(\delta_{i_1 i_2} \cdots \delta_{i_{(2l-1)} i_{2l}} + \text{sym}[(l-1)!!] \right). \qquad (2.50)$$

If $\alpha < l$, Eq.(2.49) is vanished because there is a inner product of $\mathcal{Y}$, which is traceless term, and at least one Kronecker delta involving two indices of the group $j_1 \cdots j_l$. Then Eq.(2.43) and Eq.(2.44) is also vanished when $\alpha < l$. Furthermore, Eq.(2.43) has the $l+2$ and $l-2$ angular momentum quantum number and Eq.(2.44) has the $l+1$ and $l-1$ angular momentum quantum number. While the condition of $\alpha < l$ requires the angular momentum quantum number should be larger than α, The Low-velocity expansion tells us that Eq.(2.47) and Eq.(2.48) are of order $\mathcal{O}(v^\alpha/c^\alpha)$. Then we can obtain appropriate approximation of Eq.(2.47) and Eq.(2.48) by using the lowest order $\alpha = l-2$ and $\lambda = l-1$. Then we find

$$u_{lm} \approx a_{33} \frac{4}{(l-2)!} \left(\partial_0^{l-2} S^{kli_1\cdots i_{l-2}} \right) \int d\Omega (\mathbf{T}_{lm}^{l-2})^*_{kl} n_{i_1} \cdots n_{i_{l-2}}, \qquad (2.51)$$

$$v_{lm} \approx -ib_{22} \frac{4}{(l-1)!} \left(\partial_0^{l-1} S^{kli_1\cdots i_{l-1}} \right) \int d\Omega (\mathbf{T}_{lm}^{l-1})^*_{kl} n_{i_1} \cdots n_{i_{l-1}}. \qquad (2.52)$$

Since Eq.(2.40c) and Eq.(2.40e) of Thorne[31] indicates

$$(\mathbf{T}_{lm}^{l-2})_{ij} = \left[\frac{l(l-1)}{(2l-1)(2l+1)} \right]^{1/2} \mathcal{Y}_{iji_1\cdots i_{l-2}}^{lm} n_{i_1} \cdots n_{i_{l-2}}, \qquad (2.53)$$

11

and

$$(\mathbf{T}_{lm}^{l-1})_{ij} = i \left[\frac{2l(l-1)}{(l+1)(2l+1)} \right]^{1/2} \epsilon_{pq(i}\mathcal{Y}_{j)qi_1\cdots i_{l-2}}^{lm} n_p n_{i_1} \cdots n_{i_{l-2}}, \qquad (2.54)$$

where () in the index of the STF tensor indicates that the index surrounded by the parenthesis is commute. Then we have

$$
\begin{aligned}
u_{lm} &\approx \left[\frac{(l+1)(l+2)}{2(2l-1)(2l+1)} \right]^{1/2} \frac{4}{(l-2)!} \left[\frac{l(l-1)}{(2l-1)(2l+1)} \right]^{1/2} \\
&\times \partial_0^{l-2} S^{iji_1\cdots i_{l-2}} \mathcal{Y}_{ijj_1\cdots j_{l-2}}^{lm*} \int d\Omega n_{i_1} \cdots n_{i_{l-2}} n_{j_1} \cdots n_{j_{l-2}}, \\
&= \frac{4}{(l-2!)} \left[\frac{l(l-1)(l+1)(l+2)}{2[(2l-1)(2l+1)]^2} \right]^{1/2} \frac{4\pi}{(2l-3)!!}(l-2)! \partial_0^{l-2} S^{iji_1\cdots i_{l-2}} \mathcal{Y}_{iji_1\cdots i_{l-2}}^{lm*}, \\
&= \frac{16\pi}{(2l+1)!!} \left[\frac{1}{2}l(l-1)(l+1)(l+2) \right]^{1/2} \partial_0^{l-2} S^{iji_1\cdots i_{l-2}} \mathcal{Y}_{iji_1\cdots i_{l-2}}^{lm*}, \qquad (2.55)
\end{aligned}
$$

$$
\begin{aligned}
v_{lm} &\approx -\left[\frac{l+2}{2l+1} \right]^{1/2} \frac{4}{(l-1)!} \left[\frac{2l(l-1)}{(l+1)(2l+1)} \right]^{1/2} \\
&\times \partial_0^{l-1} S^{iji_1\cdots i_{l-1}} \epsilon_{pq(i}\mathcal{Y}_{j)qj_1\cdots j_{l-2}}^{lm*} \int d\Omega n_p n_{j_1} \cdots n_{j_{l-2}} n_{i_1} \cdots n_{i_{l-1}}, \\
&= -\frac{4}{(l-1)!} \left[\frac{2l(l-1)(l+2)}{(l+1)(2l+1)^2} \right]^{1/2} \frac{4\pi}{(2l-1)!!}(l-1)! \partial_0^{l-1} S^{iji_1\cdots i_{l-1}} \epsilon_{i_1q(i}\mathcal{Y}_{j)qi_2\cdots i_{l-1}}^{lm*}, \\
&= -\frac{16\pi}{(2l+1)!!} \left[\frac{2l(l-1)(l+2)}{(l+1)} \right]^{1/2} \partial_0^{l-1} S^{iji_1\cdots i_{l-1}} \epsilon_{i_1q(i}\mathcal{Y}_{j)qi_2\cdots i_{l-1}}^{lm*}. \qquad (2.56)
\end{aligned}
$$

The two modes of the GWs Eq.(2.46) are called "mass mode" and "current mode", respectively, due to the physical properties of these modes.

First, we present the mass mode corresponding to $\mathbf{T}^{E2}$. Here, we introduce the new moment called mass moment which is the moment of the energy density T_{00} of the source such as

$$M^{i_1\cdots i_l} \equiv \frac{1}{c^2} \int d^3x T^{00} x^{i_1} \cdots x^{i_l}. \qquad (2.57)$$

From the energy conservation low of the energy-momentum tensor in a flat space $\partial_\mu T^{\mu\nu} = 0$, we obtain $\partial_0 T^{00} = \partial_i T^{0i}$. Then 2 order derivatives with

respect to time for the mass moment $\ddot{M} \equiv c^2 \partial_0^2 M$ is given by

$$
\begin{aligned}
\ddot{M}^{i_1 \cdots i_l} &= \int d^3 x \partial_0^2 T^{00} x^{i_1} \cdots x^{i_l}, \\
&= \int d^3 x \partial_i \partial_j T^{ij} x^{i_1} \cdots x^{i_l}, \\
&= \int d^3 x T^{ij} \partial_i \partial_j (x^{i_1} \cdots x^{i_l}),
\end{aligned} \tag{2.58}
$$

where we suppose that the energy momentum tensor vanishes on the boundary of the integration.

For $l \geq 2$, we have

$$
\begin{aligned}
\partial_i \partial_j (x^{i_1} \cdots x^{i_l}) &= (\partial_i \partial_j x^{i_1} x^{i_2}) x^{i_3} \cdots x^{i_l} + \cdots, \\
&= (\delta_i^{i_1} \delta_j^{i_2} + \delta_i^{i_2} \delta_j^{i_1}) x^{i_3} \cdots x^{i_l} + \cdots,
\end{aligned} \tag{2.59}
$$

where the number of the terms on the right side of Eq.(2.59) is $_lC_2 = l(l-1)/2$. By using a constant STF tensor, we obtain

$$
\mathcal{Y}_{i_1 \cdots i_l}^{lm*} S^{i_1 i_2, i_3 \cdots i_l} = \frac{1}{l(l-1)} \mathcal{Y}_{i_1 \cdots i_l}^{lm*} \ddot{M}^{i_1 \cdots i_l}. \tag{2.60}
$$

By substituting Eq.(2.60) for Eq.(2.55), we have

$$
\begin{aligned}
u_{lm} &\sim \frac{16\pi}{(2l+1)!!} \left[\frac{1}{2} l(l-1)(l+1)(l+2) \right]^{1/2} \partial_0^{l-2} S^{ij,i_1 \cdots i_{l-2}} \mathcal{Y}_{ij i_1 \cdots i_{l-2}}^{lm*}, \\
&= \frac{16\pi}{(2l+1)!!} \left[\frac{1}{2} l(l-1)(l+1)(l+2) \right]^{1/2} \frac{1}{l(l-1)} \mathcal{Y}_{i_1 \cdots i_l}^{lm*} \partial_0^l M^{i_1 \cdots i_l}, \\
&= \frac{1}{c^{l-2}} \frac{16\pi}{(2l+1)!!} \left[\frac{(l+1)(l+2)}{2l(l-1)} \right]^{1/2} \mathcal{Y}_{i_1 \cdots i_l}^{lm*} \frac{d^l}{dt^l} M^{i_1 \cdots i_l} \equiv \frac{d^l}{dt^l} I_{lm}, \tag{2.61}
\end{aligned}
$$

where

$$
\begin{aligned}
I_{lm} &\equiv \frac{1}{c^{l-2}} \frac{16\pi}{(2l+1)!!} \left[\frac{(l+1)(l+2)}{2l(l-1)} \right]^{1/2} \mathcal{Y}_{i_1 \cdots i_l}^{lm*} M^{i_1 \cdots i_l}, \tag{2.62} \\
&= \frac{1}{c^{l-2}} \frac{16\pi}{(2l+1)!!} \left[\frac{(l+1)(l+2)}{2l(l-1)} \right]^{1/2} \int d^3 x r^l T^{00} Y_{lm}^*. \tag{2.63}
\end{aligned}
$$

Eq.(A.24) brings Eq.(2.63) from Eq.(2.62). Therefore, we conclude that the E2 mode of the GWs corresponds to a mass mode arisen from the mass density of the source.

13

Meanwhile, $\mathbf{T}^{B2}$ is connected to the concept of the current mode. Here, we also introduce the new moment called angular moment which is the moment of the density T^{0i} of the source such as

$$P^{i,i_1\cdots i_l} = \frac{1}{c}\int d^3x T^{0i} x^{i_1\cdots i_l}. \tag{2.64}$$

In the same way deriving Eq.(2.58), we have the derivative of Eq.(2.64) with respect to time such as

$$
\begin{aligned}
\dot{P}^{i,i_1\cdots i_l} &= \int d^3x \partial_0 T^{0i} x^{i_1\cdots i_l}, \\
&= -\int d^3x \partial_j T^{ij} x^{i_1\cdots i_l}, \\
&= -\int d^3x T^{ij} \partial_j (x^{i_1\cdots i_l}), \\
&= -\int d^3x T^{ij} (\delta_j^{i_1} x^{i_2\cdots i_l} + \cdots x^{i_1\cdots i_{l-1}} \delta_j^{i_l}), \\
&= -(S^{ii_1,i_2\cdots i_l} + S^{ii_2,i_1i_3\cdots i_l} + \cdots), \tag{2.65}
\end{aligned}
$$

where $\dot{P}^{i,i_1}$ consists of the lth number of moment with respect to S. Then, by using Levi-Civita symbol ϵ_{ijk}, the number of terms of the $\epsilon_{ii_1q}\dot{P}^{i,i_1\cdots i_l}$ is $l-1$ since i and i_1 in $S^{ii_1,ji_2\cdots i_{l-1}}$ are commute. So we have

$$\mathcal{Y}^{lm*}_{jqi_2\cdots i_{l-1}} \epsilon_{ii_1q} S^{ij,i_1\cdots i_{l-1}} = -\frac{1}{l-1} \mathcal{Y}^{lm*}_{jqi_2\cdots i_{l-1}} \epsilon_{ii_1q} \dot{P}^{i,ji_1\cdots i_{l-1}}. \tag{2.66}$$

By substituting Eq.(2.66) for Eq.(2.56), we obtain

$$
\begin{aligned}
v_{lm} &= -\frac{16\pi}{(2l+1)!!}\left[\frac{2l(l-1)(l+2)}{(l+1)}\right]^{1/2} \partial_0^{l-1} S^{iji_1\cdots i_{l-1}} \epsilon_{i_1q(i} \mathcal{Y}^{lm*}_{j)qi_2\cdots i_{l-1}}, \\
&= -\frac{16\pi}{(2l+1)!!}\left[\frac{2l(l-1)(l+2)}{(l+1)}\right]^{1/2} \partial_0^{l-1} S^{iji_1\cdots i_{l-1}} \epsilon_{ii_1q} \mathcal{Y}^{lm*}_{jqi_2\cdots i_{l-1}}, \\
&= \frac{1}{c^{l-1}}\frac{32\pi}{(2l+1)!!}\left[\frac{l(l+2)}{2(l-1)(l+1)}\right]^{1/2} \frac{d^l}{dt^l} \mathcal{Y}^{lm*}_{ii_1i_2\cdots i_{l-1}} \epsilon_{ijk} P^{j,ki_1\cdots i_{l-1}}, \\
&\equiv \frac{d^l}{dt^l} S_{lm}, \tag{2.67}
\end{aligned}
$$

where

$$S_{lm} \equiv \frac{1}{c^l}\frac{32\pi}{(2l+1)!!}\left[\frac{l(l+2)}{2(l-1)(l+1)}\right]^{1/2} \mathcal{Y}^{lm*}_{ii_1i_2\cdots i_{l-1}} \epsilon_{ijk} P^{j,ki_1\cdots i_{l-1}}. \tag{2.68}$$

The physical meaning of Eq.(2.67) is explained by the angular moment of the source. Since T^{0i} indicates the momentum density p^i in non-relativistic scale, which also satisfies the condition of the Low-velocity expansion. we have

$$
\begin{aligned}
\epsilon_{ijk}P^{j,ki_1\cdots i_{l-1}} &= \frac{1}{c}\epsilon_{ijk}\int T^{0j}x^k x^{i_1\cdots i_{l-1}}d^3x, \\
&= \frac{1}{c}\int (\mathbf{p}\times\mathbf{x})x^{i_1\cdots i_{l-1}}d^3x.
\end{aligned}
\tag{2.69}
$$

Then the physical meaning of $\epsilon_{ijk}P^{j,ki_1}$ is a angular momentum of the source. Therefore, we conclude that the B2 mode of the GWs corresponds to a current mode arisen from angular momentum of the source.

The motion of the source including the revolution of the trajectory of the CBC classifies the mass mode discussed above because the motion of the binary system can be interpreted as kinetic motion of the two point particles. Then we discuss the Eq.(2.61) and $E2$ mode to describe the formulation of the GWs from CBC. Since Eq.(2.61) is of order $\mathcal{O}(v^l/c^l)$ because of the result of the Low-velocity expansion and the lowest order of the angular momentum quantum number is $l=2$ due to Eq.(2.51), which implies the angular momentum of the E2 mode must be $l-2\geq 0$, the coefficient of the lowest order of the mass mode can be described as

$$
u_{2m} = \frac{16\pi}{5\sqrt{3}}\mathcal{Y}_{ij}^{2m*}\dddot{M}_{ij}.
\tag{2.70}
$$

By substituting Eq.(2.70) for Eq.(2.46) with $v_{lm}=0$ and using Eq.(A.20), we obtain

$$
\begin{aligned}
h_{ij}^{TT} &= \frac{1}{r}\frac{G}{c^4}\sum_{m=-2}^{2}\left[\frac{16\pi}{5\sqrt{3}}\mathcal{Y}_{ab}^{2m*}\dddot{M}_{ab}\left(\frac{1}{12}\right)^{1/2}r^2\Lambda_{ij,i'j'}\partial_{i'}\partial_{j'}Y_{2m}(\theta,\phi)\right], \\
&= \frac{1}{r}\frac{G}{c^4}\frac{8\pi}{15}\dddot{M}_{ab}r^2\Lambda_{ij,i'j'}\frac{15}{8\pi}\partial_{i'}\partial_{j'}\left(n_a n_b - \frac{1}{3}\delta_{ab}\right).
\end{aligned}
\tag{2.71}
$$

Here, the derivation of the unit vector is

$$
\begin{aligned}
\partial_{i'}\partial_{j'}\left(n_a n_b - \frac{1}{3}\delta_{ab}\right) &= r^{-2}[\delta_{i'a}\delta_{j'b} + \delta_{i'b}\delta_{j'a} - 2(\delta_{i'a}n_{j'}n_b + \delta_{i'b}n_{j'}n_a \\
&\quad + \delta_{j'a}n_{i'}n_b + \delta_{j'b}n_{i'}n_a + \delta_{i'j'}n_a n_b) \\
&\quad + 8n_{i'}n_{j'}n_a n_b].
\end{aligned}
\tag{2.72}
$$

However, these terms of the unit vector in Eq.(2.72) are vanished when we take the inner product by the unit vector TT operator $\Lambda_{ij,i'j'}$ because the TT

15

operator is a tracefree tensor with respect to the pairs of the index ij or $i'j'$. Finally, we obtain the simple formalization of the GWs for the lowest order of mass mode such as

$$
\begin{aligned}
h_{ij}^{TT} &= \frac{2}{r}\frac{G}{c^4}\Lambda_{ij,ab}\ddot{M}^{ab}, & (2.73) \\
&= \frac{2}{r}\frac{G}{c^4}(P\ddot{M}P)_{ij} - \frac{1}{2}P_{ij}\mathrm{Tr}(P\ddot{M}). & (2.74)
\end{aligned}
$$

The TT operator allows us to know where the GW signal is injected from the celestial coordinate to the observer. First, we consider the case that the GW signal is injected along with the z-axis of the observer's coordinate such that $\hat{\mathbf{n}} = (0,0,1)$ and the corresponding matrix of Eq.(2.20) is given by

$$
P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.75}
$$

. Therefore, we obtain

$$
\begin{aligned}
\Lambda_{ij,ab}\ddot{M}_{ab} &= \begin{pmatrix} \ddot{M}_{11} & \ddot{M}_{12} & 0 \\ \ddot{M}_{21} & \ddot{M}_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij} - \frac{\ddot{M}_{11} + \ddot{M}_{22}}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij}, \\
&= \begin{pmatrix} (\ddot{M}_{11} - \ddot{M}_{22})/2 & \ddot{M}_{12} & 0 \\ \ddot{M}_{21} & -(\ddot{M}_{11} - \ddot{M}_{22})/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij}. & (2.76)
\end{aligned}
$$

The two independent modes named plus or cross mode are then obtained such that

$$
\begin{aligned}
h_+ &= \frac{1}{r}\frac{G}{c^4}(\ddot{M}_{11} - \ddot{M}_{22}), & (2.77) \\
h_\times &= \frac{1}{r}\frac{G}{c^4}\ddot{M}_{12}. & (2.78)
\end{aligned}
$$

Next, we consider the case that the GW signal is injected along with the arbitrary axis of the observer's coordinate such that $\hat{\mathbf{n}} = (\sin\theta\sin\phi, \sin\theta\cos\phi, \cos\theta)$. Here, we consider that the rotational transformation of the coordinate of the source. In other words, we transform the matrix of the mass moment in Eq.(2.73) into

$$
M_{ij} = \mathcal{R}_{ik}\mathcal{R}_{jl}M'_{kl}, \tag{2.79}
$$

16

where

$$\mathcal{R}_{ij} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}_{ij} \qquad (2.80)$$

is the operator of the rotational transformation. As a result of the transformation, we obtain the transformed modes of the GWs such that

$$
\begin{aligned}
h_+ &= \frac{1}{r}\frac{G}{c^4}[\ddot{M}_{11}(\cos^2\phi - \sin^2\phi\cos^2\theta) + \ddot{M}_{22}(\sin^2\phi - \cos^2\phi\cos^2\theta) \\
&\quad - \ddot{M}_{33}\sin^2\theta - \ddot{M}_{12}\sin 2\phi(1 + \cos^2\theta) + \ddot{M}_{13}\sin\phi\sin 2\theta + \ddot{M}_{23}\cos\phi\sin 2\theta], \quad (2.81) \\
h_\times &= \frac{1}{r}\frac{G}{c^4}[(\ddot{M}_{11} - \ddot{M}_{22})\sin 2\phi\cos\theta + 2\ddot{M}_{12}\cos 2\phi\cos\theta \\
&\quad - 2\ddot{M}_{13}\cos\phi\sin\theta + 2\ddot{M}_{23}\sin\phi\sin\theta].
\end{aligned}
\qquad (2.82)
$$

2.2 Gravitational waves from compact binary coalescence

In the application of the GW theory, the GWs emitted from Compact Binary Coalescence(CBC) are important source of the GW emission. In this section, we describe the details of the formalism of the GWs emitted from CBC. We refer to [29] and [30] to describe the details of the section.

The mass moment M_{ij} can be evaluated by using the orbital motion of the binary system when we consider the CBC. Here, we choose the cartesian coordinate system (x, y, z) in such a way that the orbit of the binary system lies in the 2d plane (x, y). In this case, the trajectory of the binary system $\mathbf{x}_0(t)$ is given by

$$
\begin{aligned}
x_0(t) &= R\cos\left(\omega_s t + \frac{\pi}{2}\right), & (2.83) \\
y_0(t) &= R\sin\left(\omega_s t + \frac{\pi}{2}\right), & (2.84) \\
z_0(t) &= 0, & (2.85)
\end{aligned}
$$

where R is the orbital radius and ω_s is the rotational frequency of the binary system. Since the 00 component of the energy momentum tensor of particles[32] is expressed by

$$T^{00}(t, \mathbf{x}) = \sum_N \gamma_N m_N c^2 \delta(\mathbf{x} - \mathbf{x}_A(t)), \qquad (2.86)$$

17

where $\gamma_A = (1 - v_N^2/c^2)^{-1/2}$ and m_N is the mass of the Nth particle. By substituting Eq.(2.86) for Eq.(2.57), the second mass moment for a non-relativistic two particles is given by

$$M^{ij}(t) = m_1 x_1^i x_1^j + m_2 x_2^i x_2^j. \tag{2.87}$$

The coordinate system of the trajectory expressed by Eq.(2.83) to Eq.(2.85) is in the center of mass frame. Then Eq.(2.87) is transformed into

$$M^{ij} = \mu x_0^i(t) x_0^j(t), \tag{2.88}$$

where $\mu = m_1 m_2/(m_1 + m_2)$ is the reduced mass. So we find

$$\ddot{M}_{11} = -\ddot{M}_{22} = -2\mu R^2 \omega_s^2 \cos 2\omega_s t, \tag{2.89}$$
$$\ddot{M}_{12} = 2\mu R^2 \omega_s^2 \sin 2\omega_s t. \tag{2.90}$$

By substituting these equation for Eq.(2.81) and Eq.(2.82) and using the Kepler's law

$$\omega_s^2 = \frac{Gm}{R^3}, \tag{2.91}$$

we obtain

$$h_+(t) = \frac{4}{r} \left(\frac{GM_c}{c^2} \right)^{5/3} \left(\frac{\pi f}{c} \right)^{2/3} \frac{1 + \cos^2 \theta}{2} \cos(2\pi f t + 2\phi), \tag{2.92}$$

$$h_\times(t) = \frac{4}{r} \left(\frac{GM_c}{c^2} \right)^{5/3} \left(\frac{\pi f}{c} \right)^{2/3} cos\theta \sin(2\pi f t + 2\phi), \tag{2.93}$$

where

$$M_c \equiv \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \tag{2.94}$$

is a chirp mass. Note that the orbital phases can be expressed in terms of the time and the orbital velocity v:

$$2\pi f t = \left(\frac{v}{c} \right)^3 \frac{c^3 t}{GM}, \tag{2.95}$$

where $M = m_1 + m_2$.

Next, we discuss the phase evolution of the GWs from CBC[33][34]. Because of the emission of GWs, the energy of the rotation of the binary system reduces. The decaying of the orbit of the binary system causes that the evolution of the phases of the GWs doesn't increase uniformly. Calculating the evolution of the phase requires the energy flux of the orbital system. Here,

we define two dimensionless functions: the energy function $\mathcal{E}(v)$ is defined by

$$E(v) - Mc^2 \equiv Mc^2 \mathcal{E}(v), \tag{2.96}$$

where $E(v)$ is the total energy of the binary system and the flux function $\mathcal{F}(v)$ is defined by

$$L_{\text{GW}}(v) \equiv \frac{c^5}{G} \mathcal{F}(v), \tag{2.97}$$

where $L_{\text{GW}}(v)$ is the luminosity of GWs. Since $L_{\text{GW}} = -dE/dt$, we have

$$\frac{dt}{dv} = -\frac{GM}{c^3} \frac{1}{\mathcal{F}} \frac{d\mathcal{E}}{dv}. \tag{2.98}$$

For the Newtonian case, these two quantities are

$$\mathcal{E} = -\frac{1}{2}\eta \left(\frac{v}{c}\right)^2, \tag{2.99}$$

$$\mathcal{F} = \frac{32}{5}\eta^2 \left(\frac{v}{c}\right)^{10}, \tag{2.100}$$

where $\eta = \mu/M$. By substituting Eq.(2.99) and Eq.(2.100) for Eq.(2.98) and taking the integral, we obtain

$$t(v) = t_{\text{coal}} - \frac{5}{256\eta} \frac{GM}{c^3} \left(\frac{v}{c}\right)^{-8}, \tag{2.101}$$

where t_{coal} is the time when the coalescence of the binary system occurs. In summarize, we obtain

$$h_+(t) = \frac{1}{r} \left(\frac{GM_c}{c^2}\right)^{5/4} \left(\frac{5}{c\tau}\right)^{1/4} \left(\frac{1 + \cos^2 \iota}{2}\right) \cos(\Psi(\tau) + \Phi), \tag{2.102}$$

$$h_\times(t) = \frac{1}{r} \left(\frac{GM_c}{c^2}\right)^{5/4} \left(\frac{5}{c\tau}\right)^{1/4} \cos \iota \sin(\Psi(\tau) + \Phi), \tag{2.103}$$

where

$$\Psi(\tau) = -2 \left(\frac{5GM_c}{c^3}\right)^{-5/8} \tau^{5/8}, \tag{2.104}$$

and $\tau = t_{\text{coal}} - t$.

According to [35], a response function $h(t)$ can be written in the time domain as

$$h(t) = F_+(t)h_+(t) + F_\times(t)h_\times(t), \tag{2.105}$$

19

where $F_+(t)$ and $F_\times(t)$ are called beam-pattern functions. These functions depend the geometry of a GW telescope and the local sidereal time(LST) at a GW telescope:

$$
\begin{aligned}
F_+(t) &= \sin\zeta\left[F_{+,0}(t)\cos 2\Phi + F_{\times,0}(t)\sin 2\Phi\right], &(2.106)\\
F_\times(t) &= \sin\zeta\left[-F_{\times,0}(t)\cos 2\Phi + F_{+,0}(t)\sin 2\Phi\right], &(2.107)
\end{aligned}
$$

where ζ is the angle between the arms of a GW telescope, and Φ is the polarization angle. $F_{+,0}(t)$ and $F_{\times,0}(t)$ are the functions when the polarization angle is set to 0[35].

According to [36], the response function of a GW telescope can be decomposed into 4 independent terms as follows:

$$
\begin{aligned}
h(t) &= A^1 F_{+,0}(t_{\mathrm{coal}})h_0(t) + A^2 F_{\times,0}(t_{\mathrm{coal}})h_0(t)\\
&+ A^3 F_{+,0}(t_{\mathrm{coal}})h_{\pi/2}(t) + A^4 F_{\times,0}(t_{\mathrm{coal}})h_{\pi/2}(t)\\
&\equiv A^\mu h_\mu(t). &(2.108)
\end{aligned}
$$

The meaning of the parameters in the function are as follows: the mutually independent phase h_0 and $h_{\pi/2}$ are given by

$$
h_0(t) = \frac{1}{c}\left(\frac{GM_c}{c^2}\right)^{5/4}\left(\frac{5}{c\tau}\right)^{1/4}\cos\Psi(\tau), \qquad (2.109)
$$

$$
h_{\pi/2}(t) = \frac{1}{c}\left(\frac{GM_c}{c^2}\right)^{5/4}\left(\frac{5}{c\tau}\right)^{1/4}\sin\Psi(\tau), \qquad (2.110)
$$

The GW amplitudes A^μ ($\mu = 0,1,2,3$) are given by

$$
\begin{aligned}
A^0 &= \frac{c}{r}\left[\frac{1+\cos^2\iota}{2}\cos 2\Phi\cos\phi_{\mathrm{coal}} - \cos\iota\sin 2\Phi\sin\phi_{\mathrm{coal}}\right],\\
A^1 &= \frac{c}{r}\left[\frac{1+\cos^2\iota}{2}\sin 2\Phi\cos\phi_{\mathrm{coal}} + \cos\iota\cos 2\Phi\sin\phi_{\mathrm{coal}}\right],\\
A^2 &= \frac{c}{r}\left[-\frac{1+\cos^2\iota}{2}\cos 2\Phi\sin\phi_{\mathrm{coal}} - \cos\iota\sin 2\Phi\cos\phi_{\mathrm{coal}}\right],\\
A^3 &= \frac{c}{r}\left[-\frac{1+\cos^2\iota}{2}\sin 2\Phi\sin\phi_{\mathrm{coal}} + \cos\iota\cos 2\Phi\cos\phi_{\mathrm{coal}}\right],
\end{aligned}
$$

where ι is the inclination of a binary system, and an initial phase ϕ_{coal} is the phase at t_{coal}. Note that we suppose the beam-pattern functions are constants during detecting GWs by the use of the network of GW telescopes. The approximation is reasonable because the time over detecting GWs is too short to change the value of beam-pattern functions.

Chapter 3

The method of Bayesian inference

In this chapter, we introduce the basic theory and one of the algorithm of Bayesian inference. On the Bayesian inference, Markov-Chain Monte Carlo(MCMC) method[37][38] is frequently used as the algorithm of Bayesian inference. The nested sampling algorithm, described by Skilling[39] is a reversal of typical Bayesian inference such as MCMC aimed to compute the evidence integral directly and produce the posterior PDF. In this thesis, the MultiNest[27][28] software, which is the implementation of the Nested sampling algorithm and modified to be able to calculate multimodal likelihood function by using clustering method, is used to estimate the parameters of CBC. Therefore, we explain the fundamental theory and functions of the Bayesian inference at first. Second, we discuss the Nested sampling algorithm and next, we show the details of the MultiNest software.

3.1 The basic theory of Bayesian inference

We refer to[40][34][41][42] to describe the fundamental theory and functions of Bayesian inference.

Bayesian statistical conclusions about parameters θ_i we want to estimate are made in terms of probability statement which is conditional on the observed value denoted by y. Here, $p(A)$ denotes a probability distribution of the event A and we define the conditional probability distribution $p(A|B)$ such as the probability of the event A when the event B is true. In addition, we also define the joint probability distribution $p(A \cap B)$ such as the probability when the event A and B are true simultaneously. Since A and B in

21

the joint probability distribution is commute, the statement is correct:

$$p(A|B)p(B) = p(A \cap B) = p(B \cap A) = p(B|A)p(A). \qquad (3.1)$$

By using these terms and definitions, we can construct the relationship between the conditional probability distribution and the joint probability distribution on the observed data and parameters:

$$p(\theta_i \cap y) = p(\theta_i)p(y|\theta_i), \qquad (3.2)$$

and then we also define the conditional probability distribution by using these terms:

$$p(y|\theta_i) \equiv \frac{p(\theta_i \cap y)}{p(\theta_i)}. \qquad (3.3)$$

By using Eq.(3.1) and Eq.(3.2), we get a important function of the Bayesian statistics named *Bayes' theorem*:

$$p(\theta_i|y) = \frac{p(y|\theta_i)p(\theta_i)}{\sum_{j=1}^{k} p(y|\theta_j)p(\theta_j)}, \qquad (3.4)$$

where k is the number of parameters we want to estimate and $p(y) = \sum_{j=1}^{k} p(y|\theta_j)p(\theta_j)$.

Eq.(3.4) plays a central role in probability and statistics. It means we can calculate the probability distribution of the parameters θ_i from the given, observed data y then the prediction of the parameters is available. Appeared terms in the Bayes' theorem have names and roles respectively:

$p(\theta_i)$ prior probability distribution: it is the probability distribution of the parameters without knowing any observed data. Then it means the knowledge of the model of the parameters in advance.

$p(y|\theta_i)$ likelihood function: it is the probability distribution of the observed data given the specific parameters. The data y affect the posterior distribution only through the likelihood.

$p(y)$ evidence: it is the probability distribution of the observed data. In practice, this probability acts as a normalization constant.

$p(\theta_i|y)$ posterior probability distribution: it is the probability distribution of the parameters after the observed data have been collected. Obtaining the posterior probability distribution is the ultimate target of a statistical analysis.

Therefore, the nature of the Bayes' theorem concludes the following relationship:

$$(\text{posterior probability}) \propto (\text{likelihood}) \times (\text{prior probability}). \qquad (3.5)$$

In other words, Bayes' theorem provides a way to update the prior knowledge on the model parameters by using given data and to estimate the posterior probability distribution, which means the estimation of the parameters by obtained data, as a result of the update.

The other expression of the posterior distribution, named odds ratio, is also useful for a Bayesian statistics. A odds ratio is defined that the ratio of the posterior distribution evaluated at the points θ_1 and θ_2 under a given model:

$$O(\theta_2|y) = O(\theta_2)\Lambda(\theta_2|y), \qquad (3.6)$$

where,

$$O(\theta_2|y) \equiv p(\theta_2|y)/p(\theta_1|y), \qquad (3.7)$$
$$O(\theta_2) \equiv p(\theta_2)/p(\theta_1), \qquad (3.8)$$

and $\Lambda(\theta_2|y) \equiv p(y|\theta_2)/p(y|\theta_1)$ is called *likelihood ratio*.

3.2 Bayesian analysis of the GWs from CBC and Matched filtering method

We refer to [34] [43] and [23] to describe the functional properties of the bayesian analysis for detecting and estimating the GWs from CBC.

In this section, we describe a coherent search method[23] with a network of GW telescopes. One fundamental detection statistics of a coherent search method in this thesis is the matched filtering; it is an optimal detection statistics testing a statistical hypothesis. The key idea of the filtering is to calculate the correlation between the output signal and a modeled GW waveform. For simplification, we suppose the terms of the GW phase have Newtonian approximation because we focus on the improvement of the accuracy the parameters of GW amplitude which include the inclination of a binary system, the luminosity distance, the polarization and the initial phase. The accuracy of these parameters is mainly improved by application of the regularization method to the inverse operator into a likelihood. It will be discussed later chapter in details.

First, we discuss the method to apply the Bayesian statistics to the estimation of the parameters of the GWs from CBC. Here, we consider that the time-series data obtained by an observation using GW telescope is

$$s(t) = h(t) + n(t), \tag{3.9}$$

where h is a GW signal and n is a noise of a GW telescope. We suppose that the noise is a stationary random process.

The odds ratio defined by Eq.(3.6) requires two kind of events θ_1 and θ_2 in their function. Here, we consider two kind of events H_0 and H_1 with respect to detect GWs as follows:

H_0 null hypothesis: there is no GW signal in the data $s(t) = n(t)$.

H_1 alternative hypothesis: these is GW signal in the data $s(t) = h(t)+n(t)$.

Then, the odds ratio Eq.(3.6) for GWs can be described as

$$O(H_1|s) = O(H_1)\Lambda(H_1|s), \tag{3.10}$$

where $\Lambda(H_1|s) \equiv p(s|H_1)/p(s|H_0)$.

Since the noise of a GW telescope is supposed to be stationary random process, the probability density distribution of the null hypothesis and alternative hypothesis is given by

$$p(s|H_0) \propto e^{-(s,s)/2}, \tag{3.11}$$

and

$$p(s|H_1) \propto e^{-(s-h,s-h)/2} = e^{-(s,s)/2+(s,h)-(h,h)/2}, \tag{3.12}$$

where the inner product of the probability distribution of a gaussian is given by Eq.(B.18). The details of the properties of the time-series data with stationary random process is described by Appendix.B. Then the likelihood ratio of the GWs is

$$
\begin{aligned}
\Lambda(H_1|s) &= \frac{e^{-(s-h,s-h)/2}}{e^{-(s,s)/2}}, \\
&= e^{(s,h)-(h,h)/2}.
\end{aligned} \tag{3.13}
$$

We call (s, h) a *matched filter*.

The concept of the matched filter is connected to the optimal detection statistics[44], [45], which realizes the optimal signal to noise ratio of the GWs.

First, we define

$$\hat{s} = \int_{-\infty}^{\infty} dt\, s(t) K(t), \tag{3.14}$$

where $K(t)$ is an arbitrary filter function. The ensemble average of Eq.(3.14), which is corresponding to expected value of $\hat{s}$ denoted by S, is

$$\begin{aligned}
S &= \int_{-\infty}^{\infty} dt \langle s(t) \rangle K(t), \\
&= \int_{-\infty}^{\infty} dt \langle h(t) \rangle K(t), \\
&= \int_{-\infty}^{\infty} df \langle \hat{h}(f) \rangle \hat{K}^*(f), \tag{3.15}
\end{aligned}$$

where f is frequency and $\langle n(t) \rangle = 0$. While the expected value of the noise is corresponding to the variance of the null hypothesis such that

$$\begin{aligned}
N^2 &= \left[\langle s(t)^2 \rangle - \langle s(t) \rangle^2 \right], \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dt\, dt' \langle n(t) n(t') \rangle K(t) K(t'), \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dt\, dt'\, K(t) K(t') \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} df\, df'\, \mathrm{e}^{2\pi i (f't' - ft)} \langle \tilde{n}^*(f) \tilde{n}^*(f') \rangle, \\
&= \int_{-\infty}^{\infty} df\, \frac{1}{2} S_n(f) \left| \tilde{K}(f) \right|^2. \tag{3.16}
\end{aligned}$$

The definition of the power spectral density of the noise S_n is given by Eq.(B.2). Then the signal to noise ratio S/N is obtained as

$$\frac{S}{N} = \frac{\int_{-\infty}^{\infty} df \langle \hat{h}(t) \rangle \hat{K}^*(t)}{\left[\int_{-\infty}^{\infty} df\, \frac{1}{2} S_n(f) \left| \tilde{K}(f) \right|^2 \right]^{1/2}}, \tag{3.17}$$

$$= \frac{S}{N} = \frac{(u, h)}{(u, u)^{1/2}}, \tag{3.18}$$

where

$$\tilde{u}(f) = \frac{1}{2} S_n(f) \tilde{K}(f). \tag{3.19}$$

The optimal signal to noise ratio is obtained by searching the unit vector $\hat{n} = u/(u, u)^{1/2}$ such that (u, h) is maximum, which corresponds to the case that $\hat{n}$ and h are parallel. Thus we obtain the optimized filter such that

$$\tilde{K}(f) = \mathrm{const} \times \frac{\tilde{h}(f)}{S_n(f)}. \tag{3.20}$$

By substituting Eq.(3.20) for Eq.(3.14), we obtain the matched filter (s, h). So we conclude that the matched filter provides the method to maximum the signal to noise ratio of the detecting GWs.

Next, we describe the concrete formalism of a coherent search method for the GWs from a CBC by using multiple GW telescopes. The index X indicates the X-th GW telescope belonging to the network of GW telescopes:

$$h^X(t) = A^\mu h_\mu^X(t). \tag{3.21}$$

We suppose that the output data from each of the GW telescopes is

$$s^X(t) = h^X(t) + n^X(t), \tag{3.22}$$

where $n^X(t)$ is the noise of the X-th GW telescope. The noise of a GW telescope is assumed to be stationary and gaussian. It is characterized by the noise power spectral density(PSD) $S_n^X(f)$ defined by using Eq.(B.3) such as

$$\langle \tilde{n}^X(f)\tilde{n}^{Y*}(f') \rangle \equiv \delta^{XY}\delta(f - f')S_h^X(f). \tag{3.23}$$

A likelihood ratio, the ratio of the hypothesis that GW signal h contains in the output of a GW telescope to null hypothesis, provides the output of the matched filtering. The likelihood ratio is defined by

$$\Lambda(h) = \frac{p(s|h)}{p(s|0)} = \frac{e^{-(s^X - h^X, s^X - h^X)/2}}{e^{-(s^X, s^X)}}, \tag{3.24}$$

A log-likelihood ratio, the logarithmic form of the likelihood ratio, is expressed by

$$\ln\Lambda = (s, h) - \frac{1}{2}(h, h). \tag{3.25}$$

The estimation of the GW parameters are evaluated by maximizing Eq.(3.25) over GW parameters.

A coherent search method employs the value summed up with each of the log-likelihood ratios calculated by the use of the output of corresponding GW telescopes[23]. The log-likelihood ratio for the multiple telescopes is given by

$$\ln\Lambda = A^\mu(\boldsymbol{s}, \boldsymbol{h}_\mu) - \frac{1}{2}A^\mu M_{\mu\nu}A^\nu, \tag{3.26}$$

where the matrix $M_{\mu\nu}$ is defined by

$$M_{\mu\nu} \equiv (\boldsymbol{h}_\mu, \boldsymbol{h}_\nu). \tag{3.27}$$

The inner product of the multiple data is expressed by the sum of each of the inner products:

$$(\boldsymbol{a}, \boldsymbol{b}) \equiv \sum_X \left(a^X, b^X\right). \tag{3.28}$$

Since the oscillation of GWs in the sensitive frequency band of a GW telescope is fast enough, the phases h_0 and $h_{\pi/2}$ is regarded as orthogonal each other. The orthogonality of phases leads to express the inner product of phases as followings:

$$\left(h_0^X, h_{\pi/2}^X\right) = 0, \tag{3.29}$$

$$\left(h_0^X, h_0^X\right) = \left(h_{\pi/2}^X, h_{\pi/2}^X\right) \equiv (\sigma^X)^2. \tag{3.30}$$

Therefore, Eq.(3.27) can be expressed in terms of the inner product of phases and beam-pattern functions:

$$M_{\mu\nu} = \begin{pmatrix} A & C & 0 & 0 \\ C & B & 0 & 0 \\ 0 & 0 & A & C \\ 0 & 0 & C & B \end{pmatrix}, \tag{3.31}$$

where

$$A = \sum_X (\sigma^X F_+^X)^2, \tag{3.32}$$

$$B = \sum_X (\sigma^X F_\times^X)^2, \tag{3.33}$$

$$C = \sum_X (\sigma^X F_+^X)(\sigma^X F_\times^X). \tag{3.34}$$

According to Eq.(3.31), the log-likelihood ratio Eq.(3.26) can be decomposed into two parts;

$$\ln\Lambda_1 = A^i(\boldsymbol{s}, \boldsymbol{h}_i) - \frac{1}{2}A^i M_{ij} A^j, \tag{3.35}$$

$$\ln\Lambda_2 = A^k(\boldsymbol{s}, \boldsymbol{h}_k) - \frac{1}{2}A^k M_{kl} A^l, \tag{3.36}$$

where $i, j = 0, 1$ and $k, l = 2, 3$.

3.3 Nested sampling algorithm

We refer to [39][43] and [46] to describe the concept of the nested sampling algorithm to implement the Bayesian inference in numerical analysis.

The key concept of the nested sampling algorithm is to calculate the evidence directory producing the posterior distribution. Here, we put the terms of a Bayesian inference as followings:

- $\pi(\theta)$: prior distribution

- $L(\theta)$: likelihood function

- Z: evidence

- $p(\theta)$: posterior distribution

, where θ denotes vector space of the parameters we want to estimate. By using Bayes' theorem Eq.(3.4), we have

$$L(\theta)\pi(\theta)d\theta = Zp(\theta)d\theta. \tag{3.37}$$

The integration of the posterior distribution over vector space of parameters Ω_θ is 1 since the posterior distribution has been normalized. It leads to the following equation:

$$Z = \int_{\Omega_\theta} L(\theta)\pi(\theta)d\theta. \tag{3.38}$$

We should calculate Eq.(3.38), but it is computationally expensive in general since Eq.(3.38) needs multiple integration of the parameter space. Nested sampling algorithm resolves the problem to replace the multiple integration with single integration by sorting the likelihood values. Here, we define the cumulant prior mass X as

$$X(\lambda) \equiv \int_{L(\theta)>\lambda} \pi(\theta)d\theta, \tag{3.39}$$

and we also define the element of the prior mass dX as

$$dX = \pi(\theta)d\theta. \tag{3.40}$$

The cumulant prior mass covers all likelihood values greater than λ. The enclosed mass X decreases from 1 to 0 depending on increasing λ. Then the evidence becomes a one-dimensional integral over unit range such that

$$Z = \int_0^1 L(X)dX. \tag{3.41}$$

Accomplishing the transformation from multiple parameters θ to single parameter X involves dividing the unit prior mass into tiny elements, and sorting them by likelihood.

To compute Eq.(3.41) on a numerical analysis, we need to express Eq.(3.41) as a discrete approximation form. Here, We use a discrete prior mass X_i which is the fraction of the prior volume enclosed by the ith contour of the equal likelihood in the parameter space such that $L_i = L(X_i)$. The right-to-left sequence of N points of the discrete prior mass is

$$0 < X_N < \cdots < X_2 < X_1 < X_0 = 1. \tag{3.42}$$

The relation between the parameter space and the discrete prior volume is illustrated as Fig. 3.1.
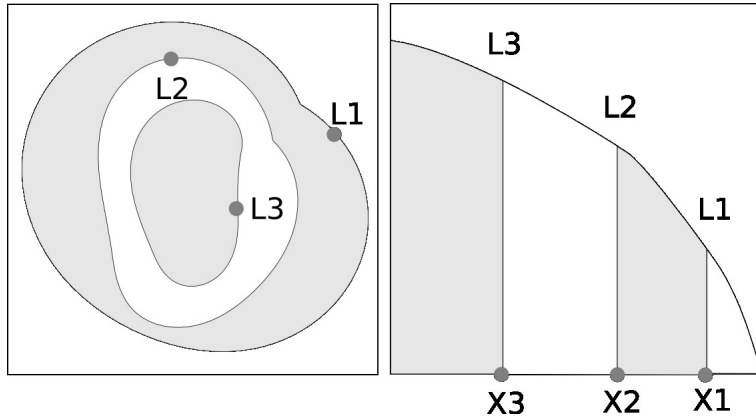

Figure 3.1: Nested likelihood contours are sorted to enclosed prior mass. The left panel is the contour line of the equal likelihood in a parameter space. The right panel is the sorted prior mass and corresponding likelihood. The each points corresponds to the live points.

Here, we define $\Delta X_i = X_i - X_{i-1}$. The evidence evaluated by the discrete prior mass is approximated by

$$Z \approx \sum_{i=1}^{N} L_i \Delta X_i. \tag{3.43}$$

The way to choose the discrete prior volume is the key of the nested sampling algorithm. Since the prior volume is sorted in dependence order, the next step of the prior volume is set stochastically such that

$$X_i = t_i X_{i-1}, \tag{3.44}$$

where $t_i \in U(0,1)$, denoted by *shrinkage ratio*, is drown from the uniform distribution. Such a stochastic sampling of the prior volume to generate N

29

sample is called *live point*. The probability of the shrinkage ratio is

$$p(t_i) = Nt_i^{N-1} \tag{3.45}$$

since $t_i \propto (t_{i-1} \cdots t_1)^{-1}$. The volume enclosed at each iteration shrinks geometrically, ensuring the speedy convergence of the integral. The mean decrease in the volume at each iteration and estimated statistical standard deviation are as follows:

$$
\begin{aligned}
E[\log(t)] &= \int_0^1 \log(t)p(t)dt \\
&= \int_0^1 \log(t)Nt^{N-1}dt \\
&= \left[log(t)t^N\right]_0^1 - \int_0^1 t^{N-1}dt \\
&= -N^{-1}, \tag{3.46}
\end{aligned}
$$

$$
\begin{aligned}
\sigma[\log(t)] &= \sqrt{\int_0^1 \left(\log(t) - E[\log(t)]^2\right)p(t)dt} \\
&= N^{-1}. \tag{3.47}
\end{aligned}
$$

Eq.(3.46) and Eq.(3.47) shows that the prior mass is expected to shrink depending on increasing the iteration numbers. Since individual $\log t$ is independent, can write the fractional prior volumes

$$
\begin{aligned}
\log X_i &\sim -(i \pm \sqrt{i})/N, \tag{3.48} \\
X_i &\sim \exp(-i/N). \tag{3.49}
\end{aligned}
$$

Therefore, the procedure of the Nested sampling algorithm when there are to be j iterative steps is as a following:

The last step fills in the missing band $0 < X < X_j$ of the desired integral with weight $w = N^{-1}X_j$ for each surviving point. Note that the algorithm tells us it isn't necessary to select N points at each iteration because $N-1$ of them remaining live points after removing one indicated the worst likelihood are already available.

The termination of the main loop is determined that the amount of the increment of the evidence in the iteration of the algorithm, which decreases with increasing the number of iteration, is less than the use-define value

---
Algorithm 1 Nested sampling procedure

Draw N points $\theta_i \cdots \theta_N$ from prior $\pi(\theta)$ and calculate these likelihood L_i;
initialize $Z = 0$, $X_0 = 1$ and $i = 0$;
while i = j **do**
   $i = i + 1$;
   record the lowest of the current likelihood values as L_i;
   set $X_i = \exp(-i/N)$;
   set $\Delta X_i = X_i - X_{i-1}$;
   set $Z_i = Z_{i-1} + L_i \Delta X_i$;
   replace the live point of the lowest likelihood by new one drawn from
   the prior $\pi(\theta)$ within $L(\theta) > L_i$;
end while
$Z = Z_j + N^{-1}(L(\theta_1) + \cdots + L(\theta_N))X_j$.

---

named *tolerance*. Such a increment of the evidence in a final step is easily written by

$$\Delta Z_i = L_{\max} X_i, \tag{3.50}$$

where $L_{\max}$ is the largest likelihood in live points of the current step.

To evaluate the posterior distribution from the Nested sampling algorithm is the most interest one. Nested sampling algorithm provides the sequence of the parameters and its likelihood from the stored live points. By using Eq.(3.37), the posterior distribution is generated by the stored live points and produced evidence. Therefore we can write the posterior distribution of the ith point from the nested sampling output as

$$p_i(\theta) = \frac{L_i \Delta X_i}{Z}. \tag{3.51}$$

By collecting these posterior distributions, the joint posterior PDF can be easily calculated by post processing the obtained data from nested sampling algorithm.

The error of the evidence evaluated by nested sampling algorithm, which is introduced by the democratization of the integral such as Eq.(3.43), is expressed by the negative *relative entropy* defined by

$$H = \int \log \left(\frac{dP}{dX}\right) dX \approx \sum_{i=1}^{N} \frac{L_i w_i}{Z} \log \left(\frac{L_i}{Z}\right), \tag{3.52}$$

where P denotes the posterior. Most of the contribution to the final evidence value comes from the final step of the iteration. Typically, the contribution

occupies a small fraction such that

$$X \approx e^{-H}.$$ (3.53)

Since Eq.(3.48), the number of necessary step i of the procedure to shrink down to the bulk of the posterior is

$$
\begin{aligned}
-H &\sim \log X_i, \\
-H &\sim -(i \pm \sqrt{i})/N, \\
i &\sim NH \pm \sqrt{NH}.
\end{aligned}
$$ (3.54)

Substituting Eq.(3.54) into Eq.(3.48), the standard deviation of the logarithm prior mass is $\sqrt{H/N}$. This uncertainty is transmitted to the evidence Z then the evidence with uncertainty is expressed by

$$\log Z = \log \left(\sum_{i=1}^{N} L_i w_i \right) \pm \sqrt{\frac{H}{N}}.$$ (3.55)

3.4   MultiNest software

We refer to [27][28][47] and [46] to describe key concept of the MultiNest software which implement a Bayesian inference based on nested sampling algorithm described in Sec.3.3 briefly.

The main challenges in implementing the computational nested sampling algorithm is to draw unbiased samples efficiently from the prior constrained by likelihood enclosed by the outermost live point and to evaluate the multimodal posterior modes with high dimensional estimation parameter space efficiently in one algorithm. The MultiNest software made by F. Feroz, M.P. Hobson and M. Bridges resolves the problems by using an ellipsoidal sampling method and Recursive clustering method in principle.

Ellipsoidal sampling method[48] provides to approximate the likelihood contour constituted by the same likelihood value in parameter space to be replaced by multi-dimensional ellipsoid determined from the covariance matrix of the current set of live points. In general, exact contour of the likelihood is an intricate construction then it is inefficient to replace the live point indicating the lowest likelihood in current set of the live points and the next live point being constrained by likelihood increasing the previous point. However, approximating the likelihood as a ellipsoid provides the simple construction of the contour of the likelihood to select the next step of live point from the prior within this ellipsoidal bound. The problem of the ellipsoid sampling

method is that the ellipse does not always restrict the sampling region within the lowest likelihood in current live points to avoid overestimating the evidence. To resolve the problem, generated ellipsoid is enlarged by a fixed enlargement factor which is larger than 1. This technical treatment allows the ellipse to encompass all live points at any steps in most case.

The ellipsoidal sampling method provides an genius solution for efficient implementation of a nested sampling algorithm, but it isn't well approximation when there are multimodal posterior modes since it cannot express one ellipsoid. To resolve the problem and be able to implement the MultiNest for the multimodal posterior distribution problem, recursive clustering method is the useful extension.

The key concept of the recursive clustering method is to split one ellipsoid into multiple ellipsoids centered on individual isolated peaks in the posterior. The division of the ellipsoid and live points begins when two conditions are accepted:

1. the total volume of the two ellipsoids is less than some fraction of the original pre-clustering ellipsoid

2. clusters are sufficiently separated by some distance to avoid overlapping regions

The k-mean clustering is used the clustering of the individual ellipsoid. It can greatly reduce the region of sampling region and increases sampling efficiency.

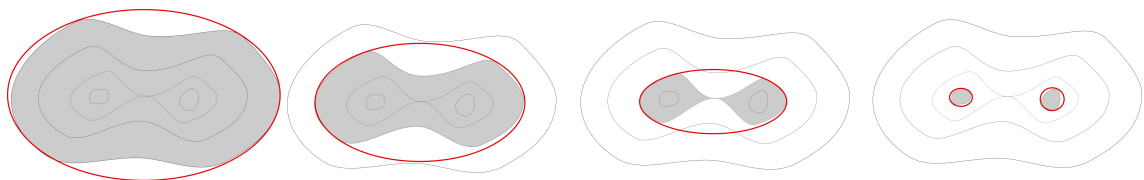The summarize picture of the MultiNest algorithm is shown in Fig. 3.2.

Figure 3.2: The overview picture of ellipsoidal nested sampling from a bimodal distribution. The left panel indicates that the ellipsoid represents a good bound to the region of the live points. The two center panels indicate that the acceptance rate will rapidly decrease as the bound steadily worsens. The right panel indicates that the increase in efficiency obtained by sampling from each clustered region separately.

Chapter 4

A regularization method

4.1 Introduction: ill-posedness of the inverse problem

To obtain the solution of a linear problem such as $Tx = y$, we solve an inverse problem such that $x = T^{-1}y$, where T^{-1} is a inverse operator of T. This solution provides the quantity of x, which is not able to measure directly, by using the given data y and the operator T. The estimation of the unknown quantity through solving the inverse problem should fulfill the Hadamard's definition of *well-posedness* conditions[49]:

- a solution exists;

- the solution is unique;

- the solution depends continuously on the data.

The first and second condition imply that the inverse operator T^{-1} exists and the third condition implies that the inverse operator is continuous and stability of the solution, which indicates that the small function of the given data y give rise to a small change in the solution x. The stability of the solution is important for physics because it ensures that if the error in the given data or model is small, the error in the result of the inverse problem will also be small. However, the inverse problem frequently violates the condition of the well-posedness and it called *ill-posedness* of the inverse problem or *ill-posed problem*.

One simple example of the ill-posed problem is the solution of the simultaneous equation with noisy data. The solution can be estimated by finding the

point of the intersection of the lines on graph, which corresponds to the solution of the inverse problem. However, the estimated region of the equation is amplified when the two lines on graph is almost parallel respectively, which implies the third condition of the well-posedness is violated. The figure. 4.1 shows the picture of the above situation - the estimated region corresponding to the red circle is amplified because of the overlap area of the two lines with noise is larger than the case when the two lines on graph is not parallel.
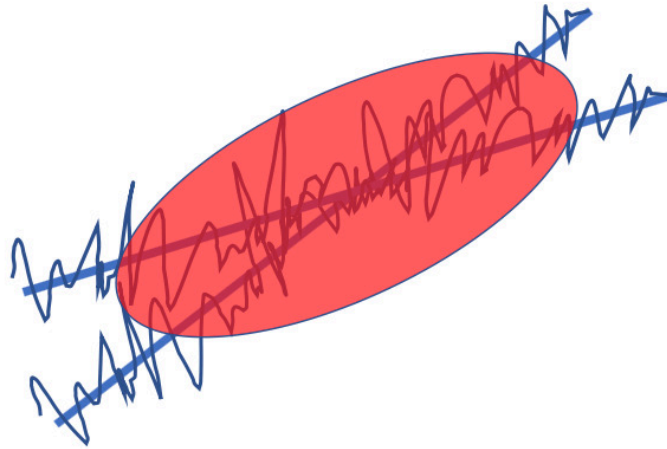

Figure 4.1: The example of the ill-posed problem.

Therefore, the ill-posed problem faces the difficulty of physics - even if the error in the given data or model is small, the error in the result of the inverse problem will be large. The regularization method introduced in this chapter provides the way to reduces the effect of the amplification of the estimation result due to the ill-posedness of the inverse problem. We present the general framework of the inverse problem named *generalized inverse problem* as a preparation of the explanation of the regularization method in Sec.4.2 at first. Next, we present the mathematical framework of the regularization method in Sec.4.3 to Sec.4.5. The application of the regularization method for the parameter estimation of a GW provides the next chapter.

4.2   Generalized inverse problem

The non-singular system of linear equations expressed by

$$Ax = b, \tag{4.1}$$

where $A \in \mathbb{C}^{n \times n}$ and $b \in \mathbb{C}^n$, has a unique solution with respect to x by taking inverse of the operator A, which is denoted by A^{-1}, such that

$$x = A^{-1}b. \tag{4.2}$$

However, such a inverse problem denoted by Eq.(4.2) cannot always solve uniquely. Especially for the inconsistent system of linear equations

$$Ax = b, \tag{4.3}$$

where $A \in \mathbb{C}^{n \times m}$ with $n \neq m$ and $b \notin \mathcal{R}(A)$, has no longer unique solution with respect to x.

Can we find or "estimate" an appropriate matrix X that $x = Xb$ indicates the solution of a general linear equation $Ax = b$? In general, the X fulfilling above condition is called the operator of a generalized inverse. The most popular operator of the generalized inverse is a Moor-Penrose inverse (abbreviated as the M-P inverse). In this section, we discuss the properties of a M-P inverse as the method to obtain the solution for a general inverse problem. The description of this section is based on [50].

First, we introduce the definition the M-P inverse:

Definition 4.2.1. A Moor-Penrose inverse
Let $A \in \mathbb{C}^{m \times n}$. The matrix $X \in \mathbb{C}^{n \times m}$ satisfying the *Penrose conditions* such that

- $AXA = A$,

- $XAX = X$,

- $(AX)^* = AX$,

- $(XA)^* = XA$,

is called the *Moor-Penrose inverse* of A denoted by $A^\dagger$.

To investigate the properties of the M-P inverse, we introduce some useful definitions with respect to the subspaces of a matrix at first.

Definition 4.2.2. Range and null space of a linear operator
For $A \in \mathbb{C}^{m \times n}$, we denote the range of A as

$$\mathcal{R}(A) = \{y \in \mathbb{C}^m | y = Ax, \ x \in \mathbb{C}^n\}. \tag{4.4}$$

The null space of A as

$$\mathcal{N}(A) = \{x \in \mathbb{C}^n | Ax = 0\}. \tag{4.5}$$

From the definitions, we can construct the following relationship between the range and the null space of a linear operator:

Lemma 4.2.3. *From the Definition.4.2.2, the following formula is hold:*

$$\mathcal{R}(A)^\perp = \mathcal{N}(A^*), \tag{4.6}$$

where $\mathcal{R}(A)^\perp$ is the orthogonal complementary subspace of $\mathcal{R}(A)$

Proof. Eq.(4.5) allows us to define the null space of A^* such that

$$\mathcal{N}(A^*) = \{y \in \mathbb{C}^m | A^*y = 0\}. \tag{4.7}$$

The orthogonal complementary subspace means the set of all vectors which are orthogonal to every vector in a subspace. So we can express as a following:

$$\mathcal{R}(A)^\perp = \{y \in \mathbb{C}^m | A'y = 0, \ A' \in \mathbb{C}^{n \times m}\}. \tag{4.8}$$

Taking $A' = A^*$ in Eq.(4.8), we have the desired result. $\qquad\square$

Remark 4.2.4. *From Lemma.4.2.3, every vector $x \in \mathbb{C}^m$ can be expressed uniquely as the sum*

$$x = y + z, \quad y \in \mathcal{R}(A), \quad z \in \mathcal{R}(A)^\perp. \tag{4.9}$$

Thus, the domain of a M-P inverse operator $D(A^\dagger)$ is limited in such as

$$D(A^\dagger) = \mathcal{R}(A) + \mathcal{R}(A)^\perp. \tag{4.10}$$

The expression of the $A^\dagger$ is determined by a least-squares solution of $Ax = b$ which provides a best-approximate solutions of linear operator equations by minimizing the residual $\|Ax - b\|$.

Definition 4.2.5. A least-squares solution and minimum-norm least squares solution

Let $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$.

(1) A vector $u \in \mathbb{C}^n$ is called a *least squares solution* of $Ax = b$ if $\|Au - b\| \le \|Av - b\|$ for all $v \in \mathbb{C}^n$.

(2) A vector $u \in \mathbb{C}^n$ is called a *minimum-norm least squares solution* of $Ax = b$ if u is a least-squares solution of $Ax = b$ and $\|u\| < \|v\|$ for any other least-squares solution w

The next two theorems show the M-P inverse $A^\dagger$ is derived from the least-squares solution and we get the expression of $A^\dagger$.

Theorem 4.2.6. a minimum-norm least squares solution and M-P inverse

Let $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$, then $A^\dagger b$ is the minimum-norm least-squares solution of Eq.(4.3).

Proof. Let $b = b_1 + b_2$, where

$$
\begin{align}
b_1 &= AA^\dagger b \in R(A) \tag{4.11} \\
b_2 &= (I - AA^\dagger)b \in R(A)^\perp, \tag{4.12}
\end{align}
$$

then

$$\|Ax - b\|^2 = \|Ax - b_1\|^2 + \|b_2\|^2. \tag{4.13}$$

From Eq.(4.13), x is a least-squares solution iff x is a solution of the consistent system $Ax = AA^\dagger b$. From the Definition of Eq.(4.7), Eq.(4.8) and Lemma.4.2.3, we obtain the expression of the null space of A such that

$$\mathcal{N}(A) = \mathcal{N}(A^\dagger A) = \mathcal{R}(I - A^\dagger A) = \left\{ (I - A^\dagger A)h \,|\, h \in \mathcal{C}^n \right\}. \tag{4.14}$$

Since $x = A^\dagger b \in \mathcal{R}(A)$ is a trivial particular solution, the general solution of the consistent system $Ax = AA^\dagger b$ is given by

$$x = A^\dagger b + (I - A^\dagger A)h. \tag{4.15}$$

Since

$$
\begin{align}
\|x\|^2 &= \|A^\dagger b + (I - A^\dagger A)h\|^2 = \|A^\dagger b\|^2 + \|(I - A^\dagger A)h\|^2 \\
&> \|A^\dagger b\|^2, \tag{4.16}
\end{align}
$$

$x = A^\dagger b$ is the minimum-norm least-squares solution. $\qquad \square$

Theorem 4.2.7. *Let $A \in \mathcal{C}^{m \times n}$ and $b \in \mathcal{C}^m$, then the following statements are equivalent:*

1. *u is a least-squares solution of $Ax = b$;*

2. *u is a solution of $Ax = AA^\dagger b$;*

3. *u is a solution of $A^* Ax = A^* b$;*

4. *u is of the form $A^\dagger b + h$, where $h \in \mathcal{N}(A)$.*

Proof. From Theorem.4.2.6, the equivalent relations between 1,2 and 4 has been proofed. So we discuss the equivalent relation about 3. If 1 holds, then $A^*Ax = A^*b$ also holds. On the other hand, multiplying $A^*Au = A^*b$ with $A^{*\dagger}$, the following equation holds because of a Penrose condition:

$$
\begin{aligned}
A^{*\dagger}A^*Au &= A^{*\dagger}A^*b \\
(AA^\dagger)^*Au &= (AA^\dagger)^*b \\
AA^\dagger Au &= AA^\dagger b \\
Au &= AA^\dagger b.
\end{aligned}
\tag{4.17}
$$

Consequently,

$$
u = A^\dagger(AA^\dagger b) + h = A^*b + h.
\tag{4.18}
$$

Thus 4 holds. $\qquad\square$

Therefore, we get a expression of the M-P inverse operator $A^\dagger$ as

$$
A^\dagger = (A^*A)^{-1}A^*.
\tag{4.19}
$$

Remark 4.2.8. *Let $A \in \mathbb{C}^{m \times n}$ and $m = n = \operatorname{rank}(A)$, then we have*

$$
A^\dagger = (A^*A)^{-1}A^* = A^{-1}(A^*)^{-1}A^* = A^{-1}.
\tag{4.20}
$$

This remark indicates $A^\dagger = A^{-1}$ when the operator A is a regular matrix. Note that we discuss the theory of regularization method by using M-P inverse operator, but the same discussion is able to apply to the invertible matrix.

4.3 The mathematical framework of a regularization method

This section introduces the basic theory of an regularization method. The explanation of the theory of the regularization method refers to [26] and [51]. The mathematical framework of the regularization method is based on the Hilbert space theory. A brief summarize of the Hilbert space is provided by Appendix.C.

As mentioned before, a regularization method provides the way to obtain the stable solution of the ill-posed problem. First, we introduce the definition of a regularization method. Here, $x^\dagger = T^\dagger y$ is a solution of the M-P inverse $Tx = y$ discussed Sec.4.2 and y is a exact data, which is the given data

eliminating noise so it is not known precisely in general. Moreover, y^δ is a noisy data and δ is a noise level. The noisy data corresponds to an output data from the observation of an experiment.

Definition 4.3.1. A regularization method

Let $T : \mathcal{X} \to \mathcal{Y}$ be a bounded linear operator between the Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$. For every $\alpha \in (0, \alpha_0)$, where $\alpha_0 \in (0, \infty]$, let

$$R_\alpha : \mathcal{Y} \to \mathcal{X} \tag{4.21}$$

be a continuous linear operator. The family R_α is called a *regularization operator* if there exists a parameter choice rule $\alpha = \alpha(\delta, y^\delta)$ such that

$$\lim_{\delta \to 0} \sup \left\{ \left\| R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y \right\| | y^\delta \in \mathcal{Y}, \left\| y^\delta - y \right\| \leq \delta \right\} = 0 \tag{4.22}$$

holds for every $y \in \mathcal{D}(T^\dagger)$.Here,

$$\alpha : \mathcal{R} \times \mathcal{Y} \to (0, \alpha_0) \tag{4.23}$$

is determined by

$$\lim_{\delta \to 0} \sup \left\{ \alpha(\delta, y^\delta) | y^\delta \in \mathcal{Y}, \left\| y^\delta - y \right\| \leq \delta \right\} = 0. \tag{4.24}$$

As discussed Sec.4.1, the solution of the inverse problem with noisy data $x^\delta = T^\dagger x^\delta$ is unstable when T is an ill-posed operator. The key concept of a regularization method is to replace the ill-posed operator $T^\dagger$ by a parameter-dependent family $\{R_\alpha\}$ of continuous operator. In other words, a regularization method provides the mathematical framework to solve the ill-posed inverse problem stably by adding an appropriate correction term to the ill-posed operator. Eq.(4.22) and Eq.(4.24) ensure that the solution of the regularization method converses to the $x^\dagger$ when the noise level tends to 0.

Here, we reconsider the case of the simultaneous equation discussed Sec.4.1. In this case, a regularization method makes the lines on graph be not parallel respectively and reduce the overlap area of the two lines with noise such as figure. 4.2. The amount of change is determined by the parameter α.

It indicates that the choice of the regularization parameter α is very important to optimize the error amplified by the ill-posed problem. The way to determine the parameter α is called *parameter choice rule*. The definition of the parameter choice rule is as following:

Definition 4.3.2. A parameter choice rule

Let α be a parameter choice rule. If α only depend on the noise level δ so that $\alpha = \alpha(\delta)$, we call α an *a-priori parameter choice rule*. Otherwise, especially for $\alpha = \alpha(\delta, y^\delta)$, we call α an *a-posteriori parameter choice rule*.
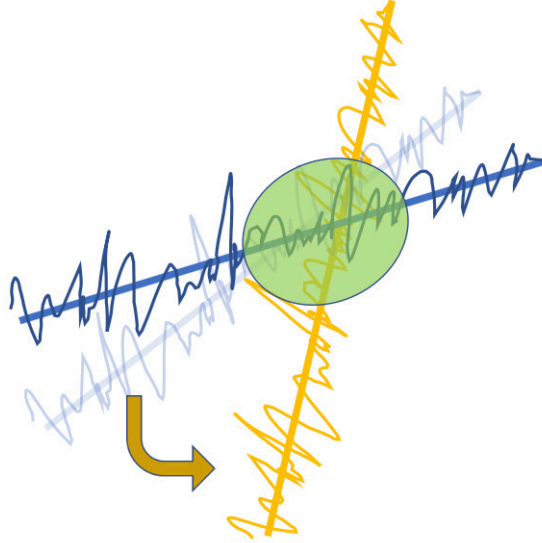
Figure 4.2: The concept of the regularization method.

The fundamental explanation of the convergence properties is in common both of the parameter choice rules. Thus we construct regularization operators by an a-priori parameter choice rule at first and next we construct the same one by an a-posteriori parameter choice rule based on a common discussion. Next, we define some of the important words and formulas for a discussion of a regularization method.

Definition 4.3.3. A convergence rate

Let x_α be a solution of the exact data y by a regularization operator such that

$$x_\alpha \equiv R_\alpha y \tag{4.25}$$

and let $x^\delta_{\alpha(\delta,y^\delta)}$ be a solution of the noisy data y^δ by a regularization operator such that

$$x^\delta_{\alpha(\delta,y^\delta)} \equiv R_{\alpha(\delta,y^\delta)} y^\delta. \tag{4.26}$$

Then we can define a *convergence rate* as a difference between $x^\delta_{\alpha(\delta,y^\delta)}$ and $x^\dagger$ such that

$$\left\| x^\delta_{\alpha(\delta,y^\delta)} - x^\dagger \right\| \leq \left\| x_{\alpha(\delta,y^\delta)} - x^\dagger \right\| + \left\| x^\delta_{\alpha(\delta,y^\delta)} - x_{\alpha(\delta,y^\delta)} \right\|. \tag{4.27}$$

In the right hand side of the triangle inequality,

$$\| x_{\alpha(\delta,y^\delta)} - x^\dagger \| \tag{4.28}$$

indicates the residuals noise introduced by a regularization method and we call it a *bias noise*. Moreover,

$$\|x_{\alpha(\delta,y^\delta)} - x^\delta_{\alpha(\delta,y^\delta)}\| \tag{4.29}$$

indicates the residual noise decreased by a regularization method and we call it a *diminished noise*.

The convergence rate of the inverse problem applying a regularization method is specified two such a convergence rate like Eq.(4.27). Here, we will introduce the properties of the convergence, especially for the inverse operator is non-closed at first. Next, we will show the explicit representation of Eq.(4.28) and we explore the same of Eq.(4.29) and then we indicate the general convergence rate of a regularization method. To prepare to archive these targets, we will define some useful formula.

Definition 4.3.4. A modulus of continuity
Let $\mathcal{X}$ be a subspace of a Hilbert space. For $\mathcal{M} \subset \mathcal{X}, \delta > 0$, let

$$\Omega(\delta,\mathcal{M}) \equiv \sup_{\|Tx\| \le \delta} \{\|x\| | x \in \mathcal{M}\} \tag{4.30}$$

be a *modulus of continuity*.

In general, $\Omega(\delta,\mathcal{M})$ has a infinite values, but when $\mathcal{M}$ is compact and $\mathcal{M} \cap \mathcal{N}(T) = 0$, then the modulus of continuity has a finite value.

Then one more important definition of the a-priori parameter choice rule, which is a "worst case error" of the convergence rate under an a-priori parameter choice rule and assumption $x^\dagger \in \mathcal{M}$, is introduced as a following.

Definition 4.3.5. A worst case error
Let $\mathcal{X}, \mathcal{Y}$ be a subspace of a Hilbert space. For $\mathcal{M} \subset \mathcal{X}, \delta > 0$, let

$$\Delta(\delta,\mathcal{M},R) \equiv \sup_{\|Tx-y^\delta\| \le \delta} \left\{\|Ry^\delta - x\| | x \in \mathcal{M}, y^\delta \in \mathcal{Y}\right\} \tag{4.31}$$

be a *worst case error*.

The two definitions have to do with each other describing below Lemma: the worst case error is bounded by the modulus of continuity.

Lemma 4.3.6. *Let $\mathcal{M} \subset \mathcal{X}, \delta > 0$ and $R : \mathcal{Y} \to \mathcal{X}$ be an arbitrary map with $R(0) = 0$. Then*

$$\Delta(\delta,\mathcal{M},R) \ge \Omega(\delta,\mathcal{M}) \tag{4.32}$$

is hold.

Proof. From Eq.(4.30) and Eq.(4.31) with the case $y^\delta = 0$, these two equations are equal with each other because of $R(0) = 0$. When $y^\delta = 0$, Eq.(4.31) takes a lowest value then we have the desired result. □

Investigating of the convergence property of Eq.(4.27) is very important for a regularization method. If $\mathcal{R}(T)$ is closed, it is guaranteed that Eq.(4.27) converges uniformly[52]. However, If $\mathcal{R}(T)$ is non-closed, there can be no uniform convergence rate for a regularization method.

Theorem 4.3.7. *Let $\mathcal{R}(T)$ be non-closed, $\{R_\alpha\}$ be a regularization operator for $T^\dagger$ with $R_\alpha(0) = 0$, $\alpha = \alpha(\delta, y^\delta)$ be a parameter choice rule. Then there can be NO function $f : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{\delta \to 0} f(\delta) = 0$ such that*

$$\|R_{\alpha(\delta,y^\delta)}y^\delta - T^\dagger y\| \le f(\delta) \tag{4.33}$$

holds for all $y \in \mathcal{D}(T^\dagger)$ with $\|y\| \le 1$ and all $\delta > 0$.

Proof. Assume that a function with $\lim_{\delta \to 0} f = 0$ exists such that Eq.(4.33) holds. Let $B_1 \equiv \{y \in \mathcal{Y} | \|y\| \le 1\}$ be a unit ball in $\mathbb{R}$. we consider that Eq.(4.33) holds for any $y \in \mathcal{D}(T^\dagger)$ with $\|y\| \le 1$. Under an a-priori assumption $x^\dagger \in \mathcal{M}$, we can choose $\mathcal{M} = \mathcal{R}(T^\dagger) \cap T^{-1}B_1 = \mathcal{N}(T)^\perp \cap T^{-1}B_1$ (see Definition.4.2.2). Therefore, from Eq.(4.31) and Eq.(4.30), we have

$$\Delta(\delta, \mathcal{N}(T)^\perp \cap T^{-1}B_1, R^\delta_{\alpha(\delta,y^\delta)}) \le f(\delta) \tag{4.34}$$

and

$$\Omega(\delta, \mathcal{N}(T)^\perp \cap T^{-1}B_1) \le f(\delta). \tag{4.35}$$

Let $\{y_k\}$ be any sequence in $\mathcal{R}(T) \cap B_1$ converging to a $y \in \mathcal{R}(T) \cap B_1$ From Eq.(4.30), if $k \in \mathbb{N}$ is such that $\|y_k - y\| < \delta$, then

$$\|T^\delta y_k - T^\dagger y\| \le \Omega(\delta, \mathcal{N}(T)^\perp \cap T^{-1}B_1). \tag{4.36}$$

The left hand side of Eq.(4.36) converges 0 as $\delta \to 0$. Then the right hand side of Eq.(4.36) also converges to 0. Hence $T^\dagger$ is bounded since $T^\dagger$ is continuous on bounded region, which contradicts the non-closedness of $\mathcal{R}(T)$. □

Theorem.4.3.7 implies that the convergence with non-closed operator is an *arbitrarily slow*, which means that there is NO uniform convergence rate – there is a solution with worse convergence rate in general, but there may exist some solutions for which the approximations have a good convergence rate. These properties emerges in the estimation of the convergence rate. It will be shown in the next passage.

To express Eq.(4.30) and Eq.(4.31) to a singular system, we formulate the exact solution $x \in \mathcal{X}$ as another representation:

Definition 4.3.8. A source set and source representation

Let B be a bounded linear operator from a Hilbert space into $\mathcal{X}$ and be given by

$$B = (T^*T)^\mu. \tag{4.37}$$

We define the exact solution $x \in \mathcal{X}$ by the use of a bounded linear operator such that

$$\mathcal{X}_{\mu,\rho} \equiv \{x \in \mathcal{X} | x = (T^*T)^\mu w, \|w\| \le \rho\} \tag{4.38}$$

and also define

$$\mathcal{X}_\mu \equiv \cup_{\rho>0} \mathcal{X}_{\mu,\rho} = \mathcal{R}((T^*T)^\mu). \tag{4.39}$$

The sets of the exact solution are called *source sets* and $x \in \mathcal{X}_{\mu,\rho}$ is said to have a *source representation*.

In the case that the operator T into Eq.(4.37) is compact, the source set can be represented by a singular system (the details of a singular system is given by Sec.C.2):

Lemma 4.3.9. *Let K be a compact with singular system (σ_n, u_n, v_n), for $\mu > 0$ we have*

$$K^\dagger y \in \mathcal{R}((K^*K)^\mu) \iff \sum_{n=1}^\infty \frac{|\langle y, u_n \rangle|^2}{\sigma_n^{2+4\mu}} < \infty \tag{4.40}$$

Proof. From Eq.(C.36), Eq.(C.35) and Eq.(C.44), we get

$$\sum_{n=1}^\infty \frac{\langle y, u_n \rangle}{\sigma_n} v_n = K^\dagger y = (K^*K)^\mu w = \sum_{n=1}^\infty \sigma_n^{2\mu} \langle w, v_n \rangle v_n \tag{4.41}$$

and also get

$$\frac{\langle y, u_n \rangle}{\sigma_n} = \sigma_n^{2\mu} \langle w, v_n \rangle. \tag{4.42}$$

By repeating the discussion of Lemma.C.2.11, we have the desired result. $\square$

Theorem 4.3.10. *For any $\mu, \rho > 0$, let $\mathcal{X}_{\mu,\rho}$ be defined by Eq.(4.38). Then the worst case error defined by Eq.(4.30) on the concrete set $\mathcal{M} = \mathcal{X}_{\mu,\rho}$ such that*

$$\Omega(\delta, \mathcal{X}_{\mu,\rho}) \le \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}. \tag{4.43}$$

holds for any $\delta > 0$.

Proof. From Eq.(4.38), let $x = (T^*T)w$, where $\|w\| \leq \rho$ and let $\{E_\lambda\}$ be the spectral family for T^*T (see Appendix.C.3 about a spectral family). The Hölder's inequality[53] argues that

$$\|(T^*T)^r x\| \leq \|(T^*T)^q x\|^{\frac{r}{q}} \|x\|^{1-\frac{r}{q}}. \tag{4.44}$$

Then, with $r = \mu$ and $q = \mu + 1/2$,

$$\begin{aligned} \|x\| &= \|T^*T)^\mu w\|, \\ &\leq \|(T^*T)^{\mu+\frac{1}{2}} w\|^{\frac{2\mu}{2\mu+1}} \|w\|^{\frac{1}{2\mu+1}}, \\ &= \|(T^*T)^{\frac{1}{2}} x\|^{\frac{2\mu}{2\mu+1}} \|w\|^{\frac{1}{2\mu+1}}, \\ &= \|Tx\|^{\frac{2\mu}{2\mu+1}} \|w\|^{\frac{1}{2\mu+1}} \end{aligned} \tag{4.45}$$

$$\tag{4.46}$$

From the definition of $\Omega(\delta, \mathcal{M})$, we take the supreme of $\|x\|$ under the condition $\|Tx\| \leq \delta$. Then we have

$$\Omega(\delta, \mathcal{X}_{\mu,\rho}) \leq \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}. \tag{4.47}$$

$\square$

Theorem.4.3.10 is a general case which doesn't distinguish whether the adjoint operator is closed or not. It indicates that the worst case error converges to 0 "at least" as fast as $\mathcal{O}(\delta^{\frac{2\mu}{2\mu+1}})$. Then it may converge more faster than the limit we obtain. However, we have different result in the case the operator is non-closed.

Theorem 4.3.11. *Let K be a compact with non-closed range. Then, for any $\mu, \rho > 0$, there is a sequence $\{\delta_k\}$ converging to 0 such that*

$$\Omega(\delta_k, \mathcal{X}_{\mu,\rho}) = \delta_k^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}} \tag{4.48}$$

Proof. Let $\{\sigma_k\}$ be the singular values of K (see Appendix.C.2 for a singular value) and suppose that the singular value converge to 0. Here we define $\delta_k \equiv \rho \sigma_k^{2\mu+1}$. Then $(\delta_k/\rho)^{\frac{2}{2\mu+1}} = \sigma_k^2$ is an eigenvalue of K^*K with eigenvector v_k. Let $x_k \equiv \rho(K^*K)^\mu v_k$, so that $x_k \in \mathcal{X}$ since Eq.(4.38). From these relations of the equation, we have

$$x_k = \rho \sigma_k^{2\mu} v_k = \delta_k^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}} v_k, \tag{4.49}$$

$$\|x_k\| = \delta_k^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}. \tag{4.50}$$

45

Therefore, we have

$$K^*Kx_k = \delta_k^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}}\sigma_k^2 v_k = \delta_k^{\frac{2\mu+2}{2\mu+1}}\rho^{-\frac{1}{2\mu+1}}v_k, \tag{4.51}$$

so that

$$\|Kx_k\|^2 = \delta_k^2. \tag{4.52}$$

From the definition of Eq.(4.30), we have

$$\begin{aligned}
\Omega(\delta_k, \mathcal{X}_{\mu,\rho}) &= \sup_{\|Kx_k\|=\delta_k}\{\|x_k\|\,|x_k \in \mathcal{X}_{\mu,\rho}\}, \\
&\geq \delta_k^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}}. \tag{4.53}
\end{aligned}$$

Together with Eq.(4.43), we have the desired result. $\qquad\square$

From Theorem.4.3.10 and Theorem.4.3.9, we conclude an important role for a regularization method: a regularization algorithm cannot converge not faster than $\mathcal{O}(\delta^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}})$ if the operator is non-closed. In other words, the optimal convergence rate we can choose is $\mathcal{O}(\delta^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}})$. Therefore, we can define the optimal condition for condition for a regularization method in virtue of Eq.(4.32).

Definition 4.3.12. Let $\mathcal{R}(T)$ be non-closed, $\{R_\alpha\}$ be a regularization operator for $T^\dagger$. For $\mu, \rho > 0$ and $y \in T\mathcal{X}_{\mu,\rho}$, let α be a parameter choice rule. We call (R_α, α) optimal in $\mathcal{X}_{\mu,\rho}$ if

$$\Delta(\delta, \mathcal{X}_{\mu,\rho}, R_\alpha) = \delta^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}} \tag{4.54}$$

holds for all $\delta > 0$. We call (R_α, α) of optimal order in $\mathcal{X}_{\mu,\rho}$ if there exist a constant $c \leq 1$ such that

$$\Delta(\delta, \mathcal{X}_{\mu,\rho}, R_\alpha) \leq c\delta^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}} \tag{4.55}$$

holds for all $\delta > 0$.

4.4 An a-priori parameter choice rule

Definition.4.3.12 implies that we should choose the parameter choice rule α so as to fulfill the Eq.(4.54) or Eq.(4.55) when we need to optimize a regularization method. In this section, we discuss the construction of the optimal convergence rate for a-priori parameter choice rule in order to construct the

optimal convergence rate for a-posteriori parameter choice rule, which is used to optimize the accuracy for estimating the GW parameters.

Now, we reconsider that what problem occur for solving the inverse problem. Here, let $\{E_\lambda\}$ be a spectral family for T^*T. If T^*T is continuously invertible, then $(T^*T)^{-1} = \int \lambda^{-1}dE_\lambda$ since Eq.(C.79) holds. To estimate the best-approximation solution $x^\dagger = T^\dagger y$, we use Eq.(4.19) and we obtain

$$x^\dagger = \int \frac{1}{\lambda}dE_\lambda T^* y. \tag{4.56}$$

However, if $\mathcal{R}(T)$ is non-closed and $y \notin \mathcal{D}(T^\dagger)$, which is called ill-posed problem, the solution Eq.(4.56) doesn't exits since $1/\lambda$ has a pole in 0. To avoid the ill-posed problem, the idea is to replace a $g_\alpha(\lambda)$ which depends on parameter-choice rule α from $1/\lambda$ such that

$$x_\alpha \equiv \int g_\alpha(\lambda)dE_\lambda T^* y. \tag{4.57}$$

For noisy data y^δ with $\|y - y^\delta\| \leq \delta$, Eq.(4.57) is expanded as

$$x_\alpha^\delta \equiv \int g_\alpha(\lambda)dE_\lambda T^* y^\delta. \tag{4.58}$$

By using these conditions, a bias noise Eq.(4.28) is represented by

$$x^\dagger - x_\alpha = x^\dagger - g_\alpha(T^*T)T^* y \;\; = \;\; (I - g_\alpha(T^*T)T^*T)x^\dagger, \tag{4.59}$$

$$= \;\; \int (1 - \lambda g_\alpha(\lambda))dE_\lambda x^\dagger. \tag{4.60}$$

Now, we define

$$\gamma(\lambda) \equiv (1 - \lambda g_\alpha(\lambda)) \tag{4.61}$$

for all α and λ and hence

$$\gamma_\alpha(0) = 1 \tag{4.62}$$

then we have

$$x^\dagger - x_\alpha = \gamma_\alpha(T^*T)x^\dagger. \tag{4.63}$$

In these cases, the regularization operator for $T^\dagger$ can be defined as

$$R_\alpha \equiv \int g_\alpha(\lambda)dE_\lambda T^* \tag{4.64}$$

and the $g_\alpha(\lambda)$ converges to $1/\lambda$ as $\alpha \to 0$.

By using these symbols, the approximation solution of $Ty = x$ can be discussed as a following:

47

Theorem 4.4.1. *For $\alpha > 0$, let $g_\alpha : [0, \|T\|^2] \rightarrow \mathbb{R}$ fulfill the following assumptions: g_α is piecewise continuous and there is a $C \in \mathbb{R}_+$ such that*

$$|\lambda g_\alpha(\lambda)| \leq C, \tag{4.65}$$

$$\lim_{\alpha \to 0} g_\alpha(\lambda) = \frac{1}{\lambda} \tag{4.66}$$

for all $\lambda \in (0, \|T\|^2]$. Then, for all $y \in \mathcal{D}(T^\dagger)$,

$$\lim_{\alpha \to 0} g_\alpha(T^*T)T^*y = x^\dagger \tag{4.67}$$

holds with $x^\dagger = T^\dagger y$. If $y \notin \mathcal{D}(T^\dagger)$, then

$$\lim_{\alpha \to 0} g_\alpha(T^*T)T^*y = +\infty. \tag{4.68}$$

Proof. Since Eq.(4.63), Eq.(4.57) and $\lambda \in (0, \|T\|^2]$ by assumption, we have

$$\|x^\dagger - x_\alpha\|^2 = \int_0^{\|T\|^2+} \gamma_\alpha^2(\lambda)d\|E_\lambda x^\dagger\|^2 < (C+1)^2. \tag{4.69}$$

Hence, by the Lebesgue dominated convergence theorem[54],

$$\lim_{\alpha \to 0} \int_0^{\|T\|^2+} \gamma_\alpha^2(\lambda)d\|E_\lambda x^\dagger\|^2 = \int_0^{\|T\|^2+} \lim_{\alpha \to 0} \gamma_\alpha^2(\lambda)d\|E_\lambda x^\dagger\|^2 \tag{4.70}$$

holds. From the definition and assumption of the Eq.(4.61), there exists a value of $\lim_{\alpha \to 0} \gamma_\alpha(\lambda)$ when $\lambda = 0$ such that

$$\lim_{\lambda \to 0+} \|E_\lambda x^\dagger\| - \|E_0 x^\dagger\| = \|Px^\dagger\|, \tag{4.71}$$

where P is a orthogonal projector onto $\mathcal{N}(T)$. Since $x^\dagger \in \mathcal{N}(T)^\perp$ (see Sec.4.2), $Px^\dagger = 0$. Therefore we have the convergence of

$$\lim_{\alpha \to 0} \|x^\dagger - x_\alpha\| = 0 \tag{4.72}$$

and Eq.(4.67) holds. Here $\|x_\alpha\| \to +\infty$ as $\alpha \to 0$ if $y \notin D(T^\dagger)$ then Eq.(4.68) holds. $\qquad\square$

Next theorem allows us to realize the properties of diminished noise Eq.(4.29) by using symbols defined this section.

Theorem 4.4.2. *Let g_α and C be in as Theorem.4.4.1, for $\alpha > 0$, let*

$$G_\alpha \equiv \sup\{|g_\alpha(\lambda)| \mid \lambda \in [0 \in \|T\|^2]\}. \tag{4.73}$$

Then,

$$\|Tx_\alpha - Tx_\alpha^\delta\| \leq C\delta \tag{4.74}$$

$$\|x_\alpha - x_\alpha^\delta\| \leq \delta\sqrt{CG_\alpha} \tag{4.75}$$

hold.

Proof. From Eq.(C.84), we have

$$g_\alpha(T^*T)T^* = T^*g_\alpha(T^*T). \tag{4.76}$$

Then we also have

$$\|Tx_\alpha - Tx_\alpha^\delta\| \leq \|TT^*g_\alpha(T^*T)\|\|y - y^\delta\|, \tag{4.77}$$

$$\leq \delta\|TT^*g_\alpha(T^*T)\|. \tag{4.78}$$

Let $\{F_\lambda\}$ be the spectral family of TT^*. Then, for all $y \in \mathcal{Y}$ with $\|y\| = 1$,

$$\|TT^*g_\alpha(T^*T)y\|^2 = \int_0^{\|T\|^2+} (\lambda g_\alpha(\lambda))^2 d\|F_\lambda y\|^2$$

$$\leq \int_0^{\|T\|^2+} C^2 d\|F_\lambda y\|^2 = C^2\|y\|^2 \tag{4.79}$$

holds and then

$$\|TT^*g_\alpha(T^*T)y\|^2 \leq C^2 \tag{4.80}$$

holds. Eq.(4.77) and Eq.(4.80) imply Eq.(4.74).
Next, we have

$$\|x_\alpha - x_\alpha^\delta\|^2 = \langle x_\alpha - x_\alpha^\delta, T^*g(TT^*)(y - y^\delta)\rangle,$$

$$= \langle Tx_\alpha - Tx_\alpha^\delta, g(TT^*)(y - y^\delta)\rangle,$$

$$\leq \|Tx_\alpha - Tx_\alpha^\delta\|\|g(TT^*)\|\delta. \tag{4.81}$$

By assumption of G_α by Eq.(4.73), we have

$$\|g_\alpha(TT^*)\| \leq G_\alpha \tag{4.82}$$

so we prove Eq.(4.75). $\qquad\square$

Next, we discuss the expression of a bias noise defined as Eq.(4.28) by using our symbols.

Theorem 4.4.3. *Let g_α be as in Theorem.4.4.1 and $\gamma_\alpha(\lambda)$ be defined by Eq.(4.61). For μ, ρ and for all $\alpha \in (0, \alpha_0)$ and $\lambda \in [0, \|T\|^2]$, let $\omega_\mu : (0, \alpha_0) \to \mathbb{R}$ be the function fulfilled*

$$\lambda^\mu |\gamma_\alpha(\lambda)| \leq \omega_\mu(\alpha). \tag{4.83}$$

Then, for $x^\dagger \in \mathcal{X}_{\mu,\rho}$,

$$\|x_\alpha - x^\dagger\| \leq \omega_\mu(\alpha)\rho, \tag{4.84}$$
$$\|Tx_\alpha - Tx^\dagger\| \leq \omega_{\mu+\frac{1}{2}}(\alpha) \tag{4.85}$$

hold.

Proof. From the definition of the exact solution Eq.(4.38), let $w \in \mathcal{X}$ be such that $x^\dagger = (T^*T)^\mu w$, $\|w\| \leq \rho$. Then we have

$$x^\dagger - x_\alpha = \gamma_\alpha(T^*T)(T^*T)^\mu w, \tag{4.86}$$
$$Tx^\dagger - Tx_\alpha = T\gamma_\alpha(T^*T)(T^*T)^\mu w. \tag{4.87}$$

The norm of Eq.(4.84) indicates Eq.(4.84). Suppose that $z = x^\dagger - x_\alpha$, we have

$$\|Tz\| = \langle Tz, Tz \rangle = \langle T^*Tz, z \rangle = \langle (T^*T)^{\frac{1}{2}}z, (T^*T)^{\frac{1}{2}}z \rangle = \|(T^*T)^{\frac{1}{2}}z\|. \tag{4.88}$$

Therefore, we have Eq.(4.85). $\qquad\square$

Theorem.4.4.2 and Theorem.4.4.3 allow us to express and investigate the convergence rate for a-priori parameter choice rule. It is given by next Corollary:

Corollary 4.4.4. *Let the assumptions of Theorem.4.4.3 hold with*

$$\omega_\mu(\alpha) = c\alpha^\mu \tag{4.89}$$

for $c > 0$. and assume that G_α as defined in Eq.(4.73) fulfills

$$G_\alpha = O\left(\frac{1}{\alpha}\right) \quad as \quad \alpha \to 0. \tag{4.90}$$

Then, with the parameter choice rule

$$\alpha \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{2\mu+1}}, \tag{4.91}$$

the regularization method (R_α, α) is optimal order in $\mathcal{X}_{\mu,\rho}$.

Proof. From Eq.(4.75), Eq.(4.84) and the assumptions, we have

$$\|x_\alpha^\delta - x^\dagger\| \le c\alpha^\mu\rho + \delta\sqrt{\frac{C}{\alpha}}. \tag{4.92}$$

Now we suppose that the parameter choice rule α is expressed by Eq.(4.91). Then we have

$$\|x_\alpha^\delta - x^\dagger\| \propto \delta^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}}. \tag{4.93}$$

Therefore, we have the desired result. $\qquad\square$

Note that the assumption of Eq.(4.90) comes from the fact that $g_\alpha(\lambda)$ converges to $1/\lambda$ as $\alpha \to 0$. Eq.(4.92) implies the properties of convergence via α. While the first term of the right hand side of Eq.(4.92) (or the bias noise) converges 0 as $\alpha \to 0$, the second term (or the diminished noise) diverges for fixed $\delta > 0$. It can be interpreted as follows: The bias term is the noise which is generated by applying a regularization method. $\alpha \to 0$ indicates that the regularization methods doesn't apply for the inverse problem, then the bias term is no longer exist. On the contrary, the diminished noise doesn't converge when $\alpha \to 0$ since it is a ill-posed inverse problem which has no solution. We realize that the convergence rate constitutes the two parts and the value can be determined by the relationship of these terms and the strength of the α.

4.5   An a-posteriori parameter choice rule

Now we discuss the another parameter choice strategy for a regularization method called "a-posteriori parameter choice rule". On the a-priori parameter choice rule discussed previous section, we cannot give the method to determine the value of the parameter choice rule α. The a-posteriori parameter choice rule provides the method to determine the α directly by estimating the minimum value of the convergence rate constituted by the bias noise that is a decreasing function via α and the diminished noise which is a increasing function via α.

In virtue of Eq.(4.75) and an inequality of arithmetic and geometric means, the convergence rate is given by

$$\begin{aligned}
\|x_\alpha^\delta - x^\dagger\|^2 &= \|x_\alpha - x^\dagger\|^2 + \|x_\alpha^\delta - x_\alpha\|^2 + 2\langle(x_\alpha - x^\dagger),(x_\alpha^\delta - x_\alpha)\rangle, \\
&\le 2\|x_\alpha - x^\dagger\|^2 + 2\|x_\alpha^\delta - x_\alpha\|^2, \\
&\le 2(\|\gamma_\alpha(T^*T)x^\dagger\|^2 + CG_\alpha\delta^2), \tag{4.94}
\end{aligned}$$

where symbols emerged in Eq.(4.94) are defined in Sec.4.4. The tragedy of the parameter choice rule is as a following: Minimizing the right hand side of Eq.(4.94) over α should certainly yield a good parameter choice for α. A necessary condition for such a minimum is given by

$$\frac{\partial}{\partial \alpha}\left(\|\gamma_\alpha(T^*T)x^\dagger\|^2 + CG_\alpha\delta^2\right) = 0. \tag{4.95}$$

Lemma 4.5.1. *In addition to the assumption made about g_α in Theorem.4.4.1, we assume that g_α and G_α are continuously differentiable with respect to α and that $\frac{\partial G_\alpha}{\partial \alpha} \neq 0$. Moreover, let*

$$f(\alpha, w) \equiv 2\left(\frac{\partial G_\alpha}{\partial \alpha}\right)^{-1}\left\langle\frac{\partial g_\alpha}{\partial \alpha}(T^*T)\gamma_\alpha(T^*T)Qw, Qw\right\rangle \tag{4.96}$$

be a auxiliary function for $\alpha > 0$ and for all $w \in \mathcal{Y}$. Then, for $w \in \mathcal{D}(T^\dagger)$,

$$\frac{\partial}{\partial \alpha}\|\gamma_\alpha(T^*T)T^\dagger w\|^2 = -\frac{\partial G_\alpha}{\partial \alpha}f(\alpha, \omega). \tag{4.97}$$

Proof. By definitions of γ_α and g_α,

$$\frac{\partial \gamma_\alpha}{\partial \alpha} = -\lambda\frac{\partial g_\alpha}{\partial \alpha} \tag{4.98}$$

holds. Then we have

$$\begin{aligned}
\frac{\partial}{\partial \alpha}\|\gamma_\alpha(T^*T)T^\dagger w\|^2 &= 2\left\langle\gamma_\alpha(T^*T)T^\dagger w, \frac{\partial \gamma_\alpha(T^*T)}{\partial \alpha}T^\dagger w\right\rangle, \\
&= 2\left\langle\gamma_\alpha(T^*T)T^\dagger w, -(T^*T)g_\alpha(T^*T)T^\dagger w\right\rangle, \\
&= -2\left\langle\gamma_\alpha(T^*T)TT^\dagger w, g_\alpha(T^*T)TT^\dagger w\right\rangle, \\
&= -\frac{\partial G_\alpha}{\partial \alpha}f(a, w). \tag{4.99}
\end{aligned}$$

$\square$

From Eq.(4.95) and Eq.(4.97), we have the simple equation such that

$$f(\alpha, y) = C\delta^2. \tag{4.100}$$

If we could find a solution $\alpha = \alpha(\delta, y)$, this would yield a good parameter choice strategy because we can determine α directly with the convergence rate minimum. However, Eq.(4.100) cannot be solve in principle because

there exist unknown exact data y in the Eq.(4.100). We never know the value of y since it's a value that is buried in noise and should be estimated. The idea of the a-posteriori parameter choice rule is to replace y by y^δ in Eq.(4.100). To analyze the strategy for a regularization method, we suppose one assumption for g_α.

Assumption 4.5.2. *Let g_α fulfill the assumption made in Theorem.4.4.1 and assume that $\alpha \mapsto g_\alpha$ and $\alpha \mapsto G_\alpha$ is continuously differentiable. Then there exists a constant $K > 0$ such that*

$$\left| \frac{\partial g_\alpha}{\partial \alpha}(\lambda) \left(\frac{\partial G_\alpha}{\partial \alpha} \right)^{-1} \right| \leq K \tag{4.101}$$

holds for $\alpha > 0$ and $\lambda \leq 0$. Moreover, the function

$$\alpha \mapsto \frac{\partial g_\alpha}{\partial \alpha}(\lambda) \left(\frac{\partial G_\alpha}{\partial \alpha} \right)^{-1} \gamma_\alpha(\lambda) \tag{4.102}$$

is strictly increasing for $\lambda > 0$

Lemma 4.5.3. *Under Assumption.4.5.2, the function f defined by Eq.(4.96) is continuous and strictly increasing in α for $Qw \neq 0$. Furthermore, we define*

$$h(w) \equiv 2 \int \lim_{\alpha \to +\infty} \left[\frac{\partial g_\alpha}{\partial \alpha}(\lambda) \left(\frac{\partial G_\alpha}{\partial \alpha} \right)^{-1} \gamma_\alpha(\lambda) \right] d\|F_\lambda Qw\|^2, \tag{4.103}$$

where $\{F_\lambda\}$ is the spectral family generated by TT^*. Then we have

$$\lim_{\alpha \to 0} f(\alpha, w) = 0, \tag{4.104}$$

$$\lim_{\alpha \to \infty} f(\alpha, w) = h(w). \tag{4.105}$$

Proof. By Eq.(4.96) and assumptions, we have

$$f(\alpha, w) = 2 \int \left[\frac{\partial g_\alpha}{\partial \alpha}(\lambda) \left(\frac{\partial G_\alpha}{\partial \alpha} \right)^{-1} \gamma_\alpha(\lambda) \right] d\|F_\lambda Qw\|^2. \tag{4.106}$$

By assumption, f is a continuous and strictly increasing. From Eq.(4.69) and Eq.(4.101), Eq.(4.106) is bounded by $K(C+1)$. Then, by the Lebesgue dominated convergence theorem[54], we have Eq.(4.105). γ_α has the properties of the convergence such that $\lim_{\alpha \to 0} \gamma_\alpha(\lambda) = 0$ for $\lambda \neq 0$ and $\lim_{\alpha \to 0} \gamma_\alpha(0) = 1$, Eq.(4.106) is in proportion to $\|P_{\mathcal{N}(TT^*)}Qw\|$ as $\alpha \to 0$, where $P_{\mathcal{N}(TT^*)}$ is a orthogonal projector onto $\mathcal{N}(TT^*)$. Since $Qw \in N(TT^*)^\perp$, we have the desired result for Eq.(4.104). $\qquad\square$

By using these preparations, we discuss the expression of the function f with noisy data y^δ.

Theorem 4.5.4. *Under Assumption.4.5.2, for any $\delta > 0$ and $y^\delta \in \mathcal{Y}$ with $Qy^\delta \neq 0$, there is a unique $\alpha = \alpha(\delta, y^\delta) > 0$ such that*

$$f(\alpha, y^\delta) = \tau\delta^2 \tag{4.107}$$

holds, where

$$\tau \in (0, h(y^\delta)\delta^{-2}) \tag{4.108}$$

and h is defined by Eq.(4.103).

Proof. From Lemma.4.5.3 and the intermediate value theorem[55], we have the desired result. □

Theorem.4.5.4 shows that the parameter choice strategy for y^δ exists uniquely. Next, we see that such a replacement doesn't change the asymptotic behavior.

Theorem 4.5.5. *With the assumption and notation of Theorem.4.5.4, let*

$$L \equiv 2\sup\left\{\left|\frac{\partial g_\alpha}{\partial\alpha}(\lambda)\left[\frac{\partial G_\alpha}{\partial\alpha}\right]^{-1}\gamma_\alpha(\lambda)\right| \;\Big|\; \alpha > 0, \lambda \geq 0\right\}, \tag{4.109}$$

and let $\tau > L$, $Qy \neq 0$, For any $\delta > 0$, let $y^\delta \in \mathcal{Y}$ be such that $\|y - y^\delta\| \leq \delta$. Furthermore, let

$$\tau_1 \equiv (\sqrt{\tau} - \sqrt{L})^2, \tag{4.110}$$
$$\tau_2 \equiv (\sqrt{\tau} + \sqrt{L})^2 \tag{4.111}$$

hold. If $Qy^\delta \neq 0$ and Theorem.4.5.4 hold, then there exists a $\tau = \tau(\delta, y^\delta) \in [\tau_1, \tau_2]$ such that

$$f(\alpha(\delta, y^\delta), y) = \tau(\delta, y^\delta)\delta^2 \tag{4.112}$$

holds.

Proof. By Assumption.4.5.2, the operator

$$H_\alpha \equiv \left[\frac{\partial g_\alpha}{\partial\alpha}(T^*T)\left(\frac{\partial G_\alpha}{\partial\alpha}\right)^{-1}\gamma_\alpha(T^*T)\right]^{\frac{1}{2}} \tag{4.113}$$

is well-defined for any $\alpha > 0$ since Eq.(4.102) is positive for any $\alpha > 0$ and $\lambda > 0$ and

$$\|H_\alpha\| \leq \sqrt{L} \tag{4.114}$$

holds. From Eq.(4.106), we have

$$\sqrt{f(\alpha, w)} = \|H_\alpha Q w\|. \tag{4.115}$$

Hence, for any $\alpha > 0$,

$$
\begin{aligned}
\sqrt{f(\alpha, y^\alpha)} &= \|H_\alpha y - H_\alpha(Qy - Qy^\delta)\|, \\
&\geq \|H_\alpha y\| - \|H_\alpha y - H_\alpha(Qy - Qy^\delta)\|, \\
&\geq \sqrt{f(\alpha, y)} - \sqrt{L}\delta.
\end{aligned}
\tag{4.116}
$$

Here, assume the case that

$$\delta^2 < \frac{h(y)}{\tau_2}. \tag{4.117}$$

Now, if Eq.(4.117) holds, Eq.(4.116) indicates that

$$\sqrt{h(y^\delta)} > (\sqrt{\tau_2} - \sqrt{L})\delta = \sqrt{\tau}\delta, \tag{4.118}$$

which implies Theorem.4.5.4 and then $Qy^\delta \neq 0$.

Now we consider to insert the exact data y to the Eq.(4.106). Then we have the two result for the inequality relations such that

$$
\begin{aligned}
\sqrt{f(\alpha, y)} &= \|H_\alpha y^\delta + H_\alpha(y - y^\delta)\| \\
&\leq \|H_\alpha y^\delta\| + \sqrt{L}\delta \\
&= \sqrt{f(\alpha, y^\delta)} + \sqrt{L}\delta = \sqrt{\tau}\delta + \sqrt{L}\delta = \sqrt{\tau_2}\delta \quad (4.119)
\end{aligned}
$$

and

$$
\begin{aligned}
\sqrt{f(\alpha, y)} &= \|H_\alpha y^\delta - H_\alpha(-y + y^\delta)\| \\
&\geq \|H_\alpha y^\delta\| - \sqrt{L}\delta \\
&= \sqrt{f(\alpha, y^\delta)} - \sqrt{L}\delta = \sqrt{\tau}\delta - \sqrt{L}\delta = \sqrt{\tau_1}\delta. \quad (4.120)
\end{aligned}
$$

Therefore, we have the desired result with $\tau(\delta, y^\delta) \in [\tau_1, \tau_2]$. $\qquad\square$

Comparing Eq.(4.100) and Eq.(4.112), we can see that two things: the boundary condition C is replaced by $[\tau_1, \tau_2]$ and the parameter choice strategy $\alpha = \alpha(\delta, y)$, which never calculate in principle, is replaced by $\alpha = \alpha(\delta, y^\delta)$, which is always able to evaluate. Then the regularization method obtained by the parameter choice strategy $\alpha = \alpha(\delta, y^\delta)$ has the same optimal convergence properties as that with $\alpha = \alpha(\delta, y)$. The details of the explanation can be shown to the next theorem.

Theorem 4.5.6. *Let Assumption.4.5.2 holds, let $y \in \mathcal{R}(T)$ with $y \neq 0$, $\tau > L$ where L is defined by Eq.(4.109), and $y \in \mathcal{Y}$ fulfill $\|y^\delta - y\| \leq \delta$ and Theorem.4.5.4 for any $\delta > 0$. Then, if $\alpha = \alpha(\delta, y^\delta)$ is determined as the unique solution of Eq.(4.107), there is a constant η such that*

$$\|x^\delta_{\alpha(\delta,y^\delta)} - T^\dagger y\| \leq \eta \inf\{\|x_\alpha - x^\dagger\| + \delta\sqrt{CG_\alpha}\}. \tag{4.121}$$

Proof. Due to Theorem.4.5.4 and Theorem.4.5.5, Eq.(4.112) has a unique solution $\alpha = \alpha(\delta, y^\delta)$. Eq.(4.97) implies that Eq.(4.107) is the first-order necessary condition for minimizing the functional

$$\alpha \mapsto \|\gamma_\alpha(T^*T)x^\dagger\|^2 + \tau(\delta, y^\delta)\delta^2 G_\alpha \tag{4.122}$$

In virtue of the strict monotonicity of $f(\cdot, y)$, Eq.(4.122) has one pole of the local minima. Thus, with

$$a \equiv \max\{1, C/\tau(\delta, y^\delta)\}, \tag{4.123}$$
$$b \equiv \max\{1, \tau(\delta, y^\delta)/C\}, \tag{4.124}$$

we obtain

$$
\begin{aligned}
\frac{1}{2}\|x^\delta_{\alpha(\delta,y^\delta)} - x^\dagger\|^2 &\leq \|\gamma_{\alpha(\delta,y^\delta)}(T^*T)x^\dagger\|^2 + CG_{\alpha(\delta,y^\delta)}\delta^2, \\
&\leq a\left[\|\gamma_{\alpha(\delta,y^\delta)}(T^*T)x^\dagger\|^2 + \tau(\delta, y^\delta)G_{\alpha(\delta,y^\delta)}\delta^2\right], \\
&= a\min_{\alpha>0}\{\left[\|\gamma_\alpha(T^*T)x^\dagger\|^2 + \tau(\delta, y^\delta)G_\alpha\delta^2\right]\}, \\
&\leq ab\inf_{\alpha>0}\{\left[\|\gamma_\alpha(T^*T)x^\dagger\|^2 + CG_\alpha\delta^2\right]\} \\
&\leq \frac{1}{2}\eta^2 \inf_{\alpha>0}\{\left[\|\gamma_\alpha(T^*T)x^\dagger\| + \sqrt{CG_\alpha}\delta\right]\}.
\end{aligned} \tag{4.125}
$$

Here, we define

$$\eta^2 \equiv 2\max\left\{\frac{C}{\tau_1}, \frac{\tau_2}{C}\right\} \tag{4.126}$$

and then, we have the desired result. $\qquad\square$

Chapter 5

A regularization method for GW observation

The parameter estimation requires to solve a inverse problem with a maximum likelihood method. The estimation of a GW amplitude is obtained by maximizing Eq.(3.26) over parameter space of GWs:

$$\hat{A}^\mu = M^{\mu\nu} \left(\boldsymbol{s} | \boldsymbol{h}_\nu \right), \tag{5.1}$$

where $\hat{A}^\mu$ are the estimated parameters of a GW amplitude and $M^{\mu\nu}$ is the inverse of the operator $M_{\mu\nu}$. The solution of the inverse problem, however, often deteriorates the accuracy of a parameter space because of the degeneracy of the parameters to estimate. The degeneracy is caused by rank deficiency of the inverse operator, which is corresponding to the ill-posedness of the inverse operator. In the literature of the estimation of GW parameters, the rank deficiency of the inverse operator happens when the array of the beam-pattern functions of $+$ mode $\mathbf{F}_+ \equiv [F_+^1, \cdots, F_+^n]$ is approximately in proportion to these of the $\times$ mode $\mathbf{F}_\times \equiv [F_\times^1, \cdots, F_\times^n]$, where n is the number of GW telescopes. As a result, similar response functions is detected over GW telescopes and the GW modes are degenerated each other. In case that the GWs observe on November 01, 2006 at 15:45:12 UTC by using the network of the GW telescopes constituted by LIGO Hanford and LIGO Livingston, whose situation is as same as the simulated data analysis discussed in Chapter.6, for instance, the sky region where the ratio of the beam-pattern functions is $\left(F_+^{\text{Hanford}} / F_+^{\text{Livingston}} \right) / \left(F_\times^{\text{Hanford}} / F_\times^{\text{Livingston}} \right) \in [1/4 : 4]$, which corresponds to $\mathbf{F}_+ \propto \mathbf{F}_\times$ approximately, is shown by Figure.5.1.

Since the degeneracy of GW modes amplifies the noise, called amplified noise, the accuracy of the parameters of GW amplitude deteriorates. As
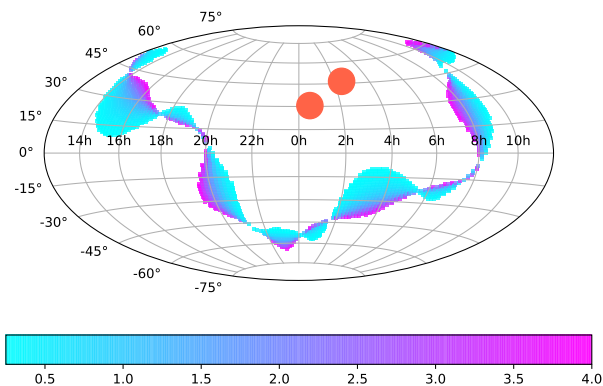
Figure 5.1: The ratio of the beam-pattern function on the sky region. The red circles are the celestial point where corresponds to the zenith of the GW telescopes based on horizontal coordinate system at the observed time.

mentioned previous Chapter, a regularization method provides the effective solution to recover the rank deficiency of the inverse operator by applying a correction term to a likelihood. As mentioned Chapter.1, M. Rakhmanov has indicated the possibility that the reconstruction of the GWs can be improved by applying Tikhonov regularization method[24]. His work, however, mainly aim to reduce the residual noise for GW burst, then the accuracy of the reconstruction of the GW waveform is ignored. Moreover, his work couldn't optimize full of the parameters introduced by the regularization method, named regulator parameters. To optimize all of the regulator parameters and apply for the parameter estimation for the GWs from CBC, we develop the new regularization method by using two methods; the method of Lagrange multiplier with Karush-Kuhn-Tucker (KKT) conditions[56], which guarantees that the distortion of the waveform by introducing regularization method suppresses and the parameter choice method to obtain the full optimization of the regulator parameters. In this Chapter, we discuss the method to optimize full of regulator parameters to suppress the amplified noise caused by a ill-posed problem.

5.1 Maximizing the likelihood with a regularization method

To suppress the amplified noise, the key idea is to add a correction term to the log-likelihood ratio:

$$\ln\Lambda_{1g} = A^i(\boldsymbol{s}|\boldsymbol{h}_i) - \frac{1}{2}A^i M_{ij} A^j - \|A^i \Omega_{ij}(\omega[k])A^j\|, \qquad (5.2)$$

where the regulator Ω_{ij} is a 2×2 matrix of regulator parameters $\omega[k]$, the k is the number of regulator parameters. Index g indicates the regularized log-likelihood ratio. The estimation $\hat{A}_g^i$, which are the parameters of a GW amplitude with regularization method, is evaluated from maximizing Eq.(5.2) over parameters of a GW amplitude

$$\hat{A}_g^i = M_g^{ij}(\boldsymbol{s}|\boldsymbol{h}_i), \qquad (5.3)$$

where

$$M_{g,ij} \equiv M_{ij} + 2\Omega_{ij}(\omega[k]). \qquad (5.4)$$

Eq.(5.4) clearly shows the inverse operator M_{ij} is corrected by the regulator and prevents the inverse operator from reducing the number of rank. The numerical expression of matched filtering of output can be decomposed into a GW signal and the noise of GW telescopes:

$$(\boldsymbol{s}|\boldsymbol{h}_i) = A^j M_{ij} + (\boldsymbol{n}|\boldsymbol{h}_i) \qquad (5.5)$$

The distinct decomposition into a GW signal and a noise couldn't be calculated because the GW signal in output of GW telescopes is never known. Note that the noise of the matched filtering $(\boldsymbol{n}|\boldsymbol{h}_i)$ is calculated by the variance of the values of the matched filtering under the bayesian inference. To make it clear, we write the mismatching as $(\boldsymbol{n}|\boldsymbol{h}_i) \equiv \Delta(\boldsymbol{s}|\boldsymbol{h}_i)$.

The numerical expression of the residual noise of a GW amplitude is obtained as a following:

$$\begin{aligned}
\left\|\hat{A}_g^i - A^i\right\|^2 &= 4\left\|M_g^{ij}\Omega_{ji'}(\omega([k]))A^{i'}\right\|^2 \\
&+ \sum_{i,i'} M_g^{ij} M_g^{i'j'} \langle \Delta(\boldsymbol{s}|\boldsymbol{h}_j)\Delta(\boldsymbol{s}|\boldsymbol{h}_{j'})\rangle \qquad (5.6) \\
&\equiv B(\omega[k], A^i) + N(\omega[k], \Delta(\boldsymbol{s}|\boldsymbol{h}_i)), \qquad (5.7)
\end{aligned}$$

where $i', j' = 0, 1$, $\langle\rangle$ is the ensemble average of the number of iterations performed in the procedure of the bayesian inference, and $\|\cdot^i\|^2 \equiv \sum_i \langle(\cdot^i)^2\rangle$.

The first term $B(\omega[k], A^i)$ corresponds to the bias noise introduced by the regulator and the second term $N(\omega[k], \Delta(\boldsymbol{s}|\boldsymbol{h}_j))$ corresponds to the amplified noise suppressed by the use of a regularization method, called reduced noise. With an appropriate choice of regulator parameters, the sum of the two noise is possible to be smaller than the amplified noise even if the bias noise, which behaves additional noise of a estimation, is introduced.

To find the optimized regulator parameters, we must consider the method to minimize the residual noise Eq.(5.6). The minimization is archived by finding the local minimum of the residual noise over the regulator parameter space. One should be paid attention: in the case that the amount of the bias noise is larger than the reduced noise as a result of the optimization of regulator parameters, the amplitude parameters of GWs lies outside of the credible regions of the probability distribution. It indicates that an appropriate restriction for determining the regulator parameters is required. We propose the requirement as a following:

$$\frac{B(\omega[k], A^i)}{N(\omega[k], \Delta(\boldsymbol{s}|\boldsymbol{h}_j))} \leq C, \tag{5.8}$$

where C is the real and positive user-defined value that is less than 1.

The minimization of the residual noise under the condition of Eq.(5.8) is performed by the method of Lagrange multiplier with Karush-Kuhn-Tucker (KKT) conditions[56]:

$$\frac{\partial}{\partial \omega[k]} \left[\left\| \hat{A}_g^i - A^i \right\|^2 + \lambda \left(B(\omega[k], A^i) - CN(\omega[k], \Delta(\boldsymbol{s}|\boldsymbol{h}_j)) \right) \right] = 0, \tag{5.9}$$

$$\frac{B(\omega[k], A^i)}{N(\omega[k], \Delta(\boldsymbol{s}|\boldsymbol{h}_j))} \leq C,$$

where the real positive number λ is a Lagrange multiplier.

5.2 An a-posteriori parameter choice rule

The evaluation of Eq.(5.9) is essentially impossible in the bayesian inference because the method to determine optimized regulator parameters includes a fatal defect; the actual parameters of a GW amplitude A^μ, which needs to be known for evaluating the bias noise of Eq.(5.6), are never known a-priori. The key idea to avoid this problem is to replace the actual parameters of GWs amplitude in the residual noise with the estimated parameters:

$$A^i \to M^{ij} \left(\boldsymbol{s}|\boldsymbol{h}_j \right). \tag{5.10}$$

This is an a-posteriori parameter choice rule discussed Sec.4.5 applying to GW observation. By adapting the parameter choice rule, the optimized regulator parameters can be calculated without using actual parameters of a GW amplitude. The estimated parameters are obtained in each of the iterations of the non-regularized bayesian inference. The ensemble average of the residual noise calculated by the actual parameters are essentially equivalent to the one calculated by estimated parameters because the difference between the actual parameters and estimated parameters is canceled out over the iterative calculation of the bayesian inference. As a result, the optimized regulator parameters are obtained by solving following equations:

$$\frac{\partial}{\partial \omega[k]}\left[\left\|\hat{A}_g^i - M^{ij}\left(\boldsymbol{s}|\boldsymbol{h}_j\right)\right\|^2 + \lambda(B(\omega[k], M^{ij}\left(\boldsymbol{s}|\boldsymbol{h}_j\right)) \quad - \quad CN(\omega[k], \Delta(\boldsymbol{s}|\boldsymbol{h}_i)))\right]$$
$$= \quad 0 \tag{5.11}$$
$$\frac{B(\omega[k], M^{ij}\left(\boldsymbol{s}|\boldsymbol{h}_j\right))}{N(\omega[k], \Delta(\boldsymbol{s}|\boldsymbol{h}_i))} \quad \leq \quad C.$$

The flowchart of the calculation of the regulator parameters $\omega[k]$ is shown as Fig. 5.2.
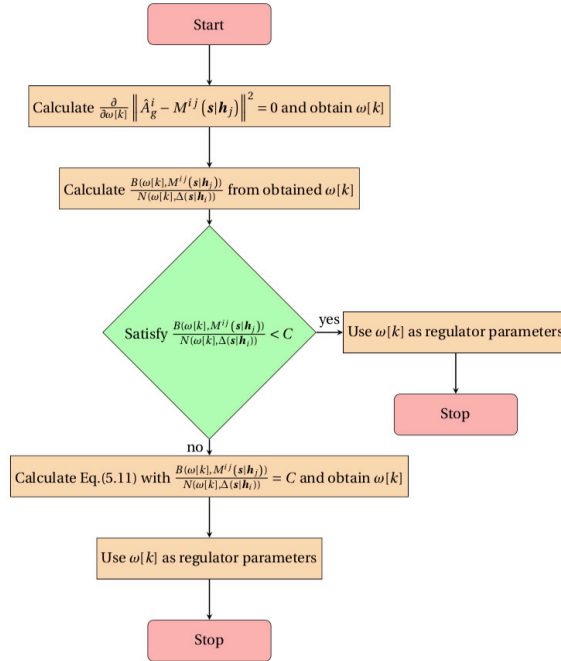

Figure 5.2: The flowchart of the calculation of the regulator parameters

While Eq.(5.11) has a lot of solutions in regulator parameters, the optimized regulator parameters are able to be determined uniquely by the use of following criteria: (i) The complex values of the regulator parameters must be excluded. (ii) the regulator parameters in which Eq.(5.11) indicates 0 or infinity by substituting must be excluded. (iii) the regulator parameters are chosen so that the real value of the Lagrange multiplier is the highest one satisfying the criteria (i) and (ii).

5.3   Regulator matrix

In this section, we will give the concrete form of the regulator matrix. As mentioned before, the rank deficiency of an inverse operator happens when one of the proper vectors of the inverse operator is approximately vanished. To compensates the lack of the proper vector, two kinds of regulator matrices are considered in this thesis. We call the formar regulator as type 1 and the latter one as type 2. The symmetricity of the regulator matrix is required since a mathematical framework of the regularization method is based on a Hilbert space.

(i) Type 1:

$$\Omega = \begin{pmatrix} w & 0 \\ 0 & w \end{pmatrix}, \tag{5.12}$$

where w is the regulator parameter which has real value. This is a traditional regulator matrix form being used for a lot of ill-posed problem.

(ii) Type 2:

$$\Omega = \begin{pmatrix} w & we \\ we & we^2 \end{pmatrix}, \tag{5.13}$$

where w and e are the regulator parameters. The rank of the regulator matrix is set to 1 ($\det \Omega = 0$). This condition comes from the fact that we should only compensate the lack of the proper vector of the inverse operator with ill-posedness whose the number of the lack of the proper vector is 1.

Chapter 6

Data analysis and result

6.1 Data analysis

This section first presents an overview of our regularized data analysis and the Bayesian inference algorithm. We also discuss the states of a software-injected signal of GWs from CBC; then the results of our regularized data analysis are presented and discussed.

Bayesian inference for a software-injected GW signal is performed using the nested sampling algorithm, which has been described by Skilling[39] as a reversal of typical bayesian interface such as Markov chain Monte Carlo methods[40]. Note that a detailed study into the parameters estimation of GWs using the nested sampling algorithm has been performed previously[43]. MultiNest implements Bayesian inference based on the nested sampling algorithm discussed Chapter.3 In addition to MultiNest, PyMultiNest software[20], a Python interface for Multinest, is also used in the data analysis. Note that MultiNest users must specify a few proper parameters, i.e., the number of live points and the tolerance, and set the prior and log-likelihood functions. In this study, the number of live points is 1000 for all data analyses and the tolerance is set to 0.0001 for the data analysis without a regulator (non-regularized data analysis) and 0.05 for the data analysis with regulator-optimized analysis (regularized data analysis).

We must evaluate Eq.(5.11) for the regularized data analysis to determine optimized regulator parameters. However, the evaluation of Eq.(5.11) requires another algorithm in addition to the nested sampling algorithm. As discussed previously, the degree of mismatch of the matched filtering evaluated relative to the accuracy of the phase parameters, such as a chirp mass and a coalescence time when GWs pass though the center of the earth, is re-

quires to determine optimized regulator parameters. These mismatch values are obtained by evaluating the accuracy of the phase parameters obtained by the non-regularized data analysis. The regularized data analysis is performed as follows. First, non-regularized data analysis is performed to obtain the covariance matrix of the phase parameters. Here, we also obtain estimated amplitude parameter values and corresponding credible regions to confirm the accuracy improvement gained by the amplitude parameters obtained in the regularized data analysis. Second, we obtain the distribution of h_μ by substituting the distribution of phase parameters generated by multivariate random distribution of the covariance matrix of the phase parameters into the matched filtering. Here, the number of samples in the distribution set of h_μ are 10000 in this thesis. Third, the variance of the h_μ is calculated using the distribution set.

The details the data analysis are as follows: the number of search parameters of the GWs is six, inclination ι, luminosity distance r, polarization Φ, initial phase ϕ_{coal}, chirp mass M_c, and coalescence time t_{coal} when GWs pass through the center of the earth. To reduce data analysis time, we ignore the post-Newtonian parameters of GWs. Note that all of the prior distribution of the search parameters are flat distributions. In addition, the data analysis result is expressed in terms of physical parameters: however, the data analysis is implemented using the terms of amplitude parameters $A^1 - A^4$. We convert these amplitude parameters into physical parameters after data analysis is performed. The range of the amplitude parameters is $A^i \in [-10^{-15}, 10^{-15}]$. The range of the inclination expressed as a cosine is $\cos \iota \in [-1, 1]$, luminosity distance is $r \in [10 : 1000]$Mpc, polarization is $\Phi \in [0, \pi]$ and initial phase is $\phi_{\mathrm{coal}} \in [0, \pi]$. To reduce data analysis time, the range of a chirp mass is set to $\pm 1 M_{solar}$ of the input value, and the coalescence time range is set to $\pm 0.1s$ of the input value. These range settings are appropriate because the regularization method helps in reducing the accuracy of amplitude parameters, and the accuracy of the phase parameters is not affected significantly.

Both regulators (Eq.(5.12) and Eq.(5.13)) are applied to the regularized data analysis to compare the differences in the reduction of the accuracy of the amplitude parameters. Here, the user-defined value of C in Eq.(5.8) is set to 0.1. To reduce the computational cost of the type 2 regulator (Eq.(5.13)), we exclude the regulator parameter e to evaluate the requirement in Eq.(5.8). The validity of the exclusion of the calculation can be explained by the fact that the variable e in the type 2 regulator does not have strong dependency on the strength of the regulator. The strength of the regulator primarily affects the amount of Eq.(5.8).

6.2 Software-injected data

The software-injected GW signal is examined herein. In this study, the data analysis is implemented using two types of GW signals. The different status of these signals reflects the amplitude parameters because we are interested in the difference in the reduction of the amplified noise relative to the amplitude parameters of GWs. In the following sections, we first describe the common states of the software-injected GW signal; then, we discuss the difference in the amplitude parameters.

6.2.1 Common states

We consider that the network of GW telescopes comprises the LIGO Livingston and the LIGO Hanford[57]. The location of the GW telescopes is the actual geographic state[58] and the noise power spectrum of the GW telescopes uses the theoretical function of aLIGO[45]. The properties of the noise of GW telescopes are assumed be both stationary and Gaussian, and the integrated SNR of the software-injected GW signal is set to 20 ± 0.5. The common states of the software-injected GW signal are listed in Table.6.1. The details of these parameters are explained in the caption.

6.2.2 Different states

Two different states are involved in the amplitude parameters of the GW signal. We are interested in reducing each amplitude parameter because the accuracy of the amplitude parameters of GWs is reduced by the regularized data analysis. The states of the amplitude parameters of the software-injected data are given in Table.6.2. The luminosity distance is set to the value which the integrated SNR indicates $20\pm$. The location of the target CBC in celestial coordinates is shown in Figure. 6.1, which also illustrates the determinant value of the beam-pattern functions because the influence of ill-posed of the inverse operator increases relative to the reduction of the determinant value.

6.3 Result and discussion

The data analysis results are expressed as the joint-probability distribution of the posterior as represented by the profile likelihood. The joint-probability distributions for all data analyses are shown in Figs. 6.2 - Fig. 6.13. In these

Parameters	Values	Unit
Observation UTC	2006/11/01	yyyy/mm/dd
	15:45:12	hh:mm:ss
chirp mass	7	solar mass
coalescence time	8	second
integrated SNR	20±0.5	
time duration	10	sec
sampling frequency	2048	Hz
upper frequency	1024	Hz
cutoff frequency	20	Hz

Table 6.1: Common state of the parameters of the software-injected GW signal. Observation UTC is the time at which the GWs pass through in the Earth's center of gravity. Chirp mass and coalescence time indicate the phase parameters of the GWs. Integrated SNR is the SNR in the GW signal detected by the network of GW telescopes (LIGO Hanford and LIGO Livingston.) simultaneously Time duration is the length of time of the output data of the GW telescopes. Sampling, upper and cutoff frequencies indicate the configuration of the calibrated output data.

figures, the area enclosed by the bold line is the 1σ (67%) credible region and the dashed line is the 2σ (95%) credible region. The shrinkage rate of the credible region obtained by comparing the posterior distribution of the non-regularized and regularized data analysis using the type 1 and type 2 regulators is shown in Table.6.3 and 6.4, respectively. Here, state 1 and 2 of the software-injected data were used Table.6.3 and 6.4, respectively. An average of 100000 posterior samples for the non-regularized data analysis and 200000 posterior samples for regularized data analysis were evaluated. The implementation time of the algorithm was approximately one CPU hour for the non-regularized data analysis and approximately 50 CPU hour for the regularized data analysis.

The results clearly indicate that the accuracy of the amplitude parameters was improved by the regularized data analysis with the type 2 regulator in the most of the sky region; however, the accuracy of the amplitude parameters was primarily unimproved by the regularized data analysis with the type 1 regulator. Here, most actual values of the amplitude parameters were outside the credible region of the joint-posterior distribution. The average shrinkage rate of the credible region for the inclination vs luminosity distance all over

Parameters	State 1	State 2	Unit
inclination ($\cos \iota$)	0.400	-0.750	
initial phase	-0.448	-1.023	Radian
polarization	-0.480	0.462	Radian

Table 6.2: States of amplitude parameters of software-injected GW signal. State 1 indicates the values of the first software-injected GW signal, and state 2 indicates the values of the second signal. The inclination data are expressed by cosine function.
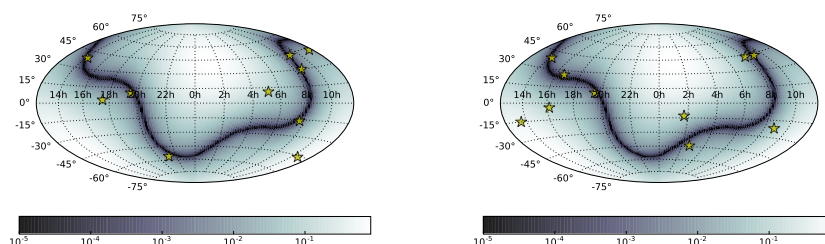

Figure 6.1: The sky location of the software-injected data and the determinant values of the antenna-beam pattern function of GW telescopes. The degree of shade in both figures indicates the amount of determinant of beam-pattern function all over the sky region at the UTC specified in Table.6.1. The stars indicates the location of the target CBC of the software-injected GW signal. The first panel shows the state 1 of software-injected data and the second panels shows the same of state 2.

the sky was approximately 1.5 times and that of the polarization vs initial phase was approximately 3.0 times for the regularized data analysis with the type 2 regulator. The shrinkage rate tends to be greater depending on the strength of the influence of the ill-posed inverse operator. However, in some sky region where the determinant value of the inverse matrix was greater than a certain threshold, the actual value was outside the credible region estimated by the regularized data analysis with the type 2 regulator. The root-mean-square(rms) of the accuracy of the GW amplitude $(\ln A)_{\mathrm{rms}}$ for every data analysis is shown by Table. 6.5 and Table. 6.6. Note that the theoretically optimal value of the rms of the accuracy is expressed by

$$(\Delta \ln A)_{\mathrm{rms}} = \frac{1}{\mathrm{SNR}} = \frac{1}{20} = 0.05, \tag{6.1}$$

where SNR is the signal-to-noise ratio of the detected GW signal[34].

Here, we discuss the results of the data analysis for states 1 and 2 in detail. For the state 1, the regularized data analysis with the type 2 regulator improved the accuracy of the amplitude parameters in most of the sky region. In contrast, the actual value was outside the estimated credible region because the influence of the ill-posed inverse operator was significant in this case. This failure implies that the range of application of the regularized data analysis, where valid estimation can be implemented, is limited depending on the degree of influence the ill-posed inverse operator, even if the restriction in Eq.(5.8) is imposed to prevent the credible region from lying outside of the actual parameter values. Fortunately, the sky region where the misestimation occurred is small and does not affect data analysis result significantly. The sky region wherein such a misestimation occurred accounts for $< 1\%$ for all over the sky in the case of state 1. Note that the significant reduction to the accuracy of luminosity distance, which is an important GW observation parameter, was realized at the determinant value, i.e., less than approximately 1×10^{-3}.

In contrast to state 1, the result of the regularized data analysis for state 2 exhibits specific characteristics. The nature of the reduction of the accuracy of the amplitude parameters via the regularized data analysis is similar to the data analysis of state 1. However, the actual amplitude parameter values are outside the credible region of the inclination vs luminosity distance when the determinant value of the inverse operator is less than 5×10^{-4}, which threshold to the misestimation is larger than state 1. The main difference between states 1 and 2 is the absolute value of inclination, i.e., the absolute inclination value of state 2 expressed in cosine is greater than state 1, which may be related to the difficulty associated with correct estimation of the inclination-distance when the absolute inclination converted by cosine is relatively large. The reason of the misestimation is expressed by the another degeneracy having nothing to do with the ill-posed problem due to the GW telescopes: the nature of the degeneracy of the plus and cross polarization of the GWs (h_+ and $h_\times$) for observing relatively face-on/away CBC event, which corresponds to $|\cos\iota| \approx 1$[59][60]. While the amplitude parameters of the GWs is estimated by the absolute value of the amplitude of the GWs and the difference value between the plus polarization and cross polarization, the difference of the polarization is small when we observe the face-on/away CBC. For instance, the relative difference between plus and cross polarization is less than 1% for inclination value less than $30°$ or greater than $150°$ and 5% for inclination less than $45°$ or greater than $135°$. It means that we cannot distinguish the contribution of the plus and cross polarization for $|\cos\iota| \approx 1$

due to the degeneracy of the plus and cross polarization for estimating these values. However, in spite of the degeneracy of these polarization, the sky region wherein such a misestimation occurred accounts for $< 1\%$ for all over the sky in the case of state 2.

The rms of the accuracy of the GW amplitude can be used for evaluating whether the accuracy of amplitude estimated by using regularization method is approaching to the optimal value or not. From Table. 6.5 and Table. 6.6, we can point out that the rms of the accuracy of the GW amplitude estimated by regularization method by using type 2 regulator for state 1 and state 2 when the actual amplitude parameters of inclination vs luminosity distance are inside the credible region is within the several times of the optimal rms value expressed in Eq.(6.1).

		Inclination vs luminosity distance					
sky	det	type 1 1σ	type 1 2σ	y/n	type 2 1σ	type 2 2σ	y/n
90%	5×10^{-1}	0.94	1.00	y	1.46	1.38	y
48%	1×10^{-1}	1.09	0.95	y	1.78	1.64	y
29%	5×10^{-2}	0.77	0.82	y	1.43	1.44	y
18%	1×10^{-2}	1.35	1.25	y	1.43	1.36	y
12%	5×10^{-3}	1.03	1.26	y	1.81	2.51	y
6%	1×10^{-3}	1.37	1.90	y	3.10	3.26	y
4%	5×10^{-4}	1.06	1.12	y	1.36	1.59	y
2%	1×10^{-4}	4.11	7.73	n	3.18	4.24	y
2%	5×10^{-5}	3.95	6.57	n	7.07	10.2	y
1%	1×10^{-5}	0.96	1.29	y	3.55	6.59	n

		Polarization vs initial phase					
sky	det	type 1 1σ	type 1 2σ	y/n	type 2 1σ	type 2 2σ	y/n
90%	5×10^{-1}	1.02	0.92	y	3.71	3.00	y
48%	1×10^{-1}	1.29	1.09	y	3.07	2.42	y
29%	5×10^{-2}	0.88	0.80	y	2.68	2.17	y
18%	1×10^{-2}	1.43	1.49	y	3.05	2.47	y
12%	5×10^{-3}	1.28	1.45	y	4.64	4.30	y
6%	1×10^{-3}	1.66	2.13	y	4.35	3.19	y
4%	5×10^{-4}	1.04	1.15	y	2.97	3.49	y
2%	1×10^{-4}	5.90	9.25	n	8.60	9.51	n
2%	5×10^{-5}	4.72	7.36	n	10.4	12.5	y
1%	1×10^{-5}	1.17	1.74	n	6.80	11.1	n

Table 6.3: Rate of shrinkage of credible region from the joint-posterior distribution and indication of whether the 2σ credible region contains the actual point of the amplitude parameters. State 1 is used for the software-injected GW signal. The left-most column shows the quantile of the sky region depending on decreasing the determinant value of the beam-pattern matrix corresponding to the software-injected data. The next column is the determinant value of the beam-pattern function corresponding to the quantile of the sky region. y/n indicates whether the actual points of the amplitude parameters are inside the 2σ credible region or not.

		Inclination vs luminosity distance					
sky	det	type 1 1σ	type 1 2σ	y/n	type 2 1σ	type 2 2σ	y/n
90%	5×10^{-1}	1.06	1.09	y	1.23	1.50	y
48%	1×10^{-1}	1.00	1.00	y	1.18	1.52	y
29%	5×10^{-2}	0.95	0.99	y	1.01	1.05	y
18%	1×10^{-2}	1.66	2.81	n	1.36	1.57	y
12%	5×10^{-3}	1.52	2.52	n	1.11	1.15	y
6%	1×10^{-3}	1.12	1.31	y	1.19	1.27	y
4%	5×10^{-4}	4.54	6.33	n	5.11	5.97	n
2%	1×10^{-4}	1.20	1.83	y	4.06	6.67	n
2%	5×10^{-5}	3.96	7.85	n	6.69	10.1	n
1%	1×10^{-5}	1.92	4.48	n	13.6	26.7	n

		Polarization vs initial phase					
sky	det	type 1 1σ	type 1 2σ	y/n	type 2 1σ	type 2 2σ	y/n
90%	5×10^{-1}	1.14	1.40	y	4.04	4.48	y
48%	1×10^{-1}	1.00	0.98	y	3.92	4.30	y
29%	5×10^{-2}	1.07	1.23	y	2.78	2.41	y
18%	1×10^{-2}	2.51	3.90	y	3.14	3.48	y
12%	5×10^{-3}	2.65	4.53	n	2.75	4.54	y
6%	1×10^{-3}	1.24	1.93	y	1.97	2.10	y
4%	5×10^{-4}	5.10	7.80	y	8.15	8.69	y
2%	1×10^{-4}	1.63	2.35	y	11.5	19.5	n
2%	5×10^{-5}	4.68	7.28	y	9.96	12.3	n
1%	1×10^{-5}	3.04	6.60	y	24.5	44.9	n

Table 6.4: Rate of shrinkage of credible region from the joint-posterior distribution and indication of whether the 2σ credible region contains the actual point of the amplitude parameters. State 1 is used for the software-injected GW signal. State 2 was used for the software-injected GW signal. The explanation of the table is as same as the table. 6.3.

sky	det	w/o regulator	type 1	y/n	type 2	y/n
90%	5×10^{-1}	0.169361	0.167728	y	0.074887	y
48%	1×10^{-1}	0.185151	0.149698	y	0.085490	y
29%	5×10^{-2}	0.177168	0.161167	y	0.083451	y
18%	1×10^{-2}	0.193302	0.155744	y	0.091900	y
12%	5×10^{-3}	0.322854	0.294635	y	0.184962	y
6%	1×10^{-3}	0.407348	0.335222	y	0.242069	y
4%	5×10^{-4}	0.650330	0.590825	y	0.391114	y
2%	1×10^{-4}	1.834264	0.284385	n	0.330782	y
2%	5×10^{-5}	1.653140	0.254334	n	0.197014	y
1%	1×10^{-5}	5.887165	3.059262	y	0.789518	n

Table 6.5: The root-mean-square of the accuracy of the amplitude of the GWs from CBC. State 1 is used for the software-injected GW signal. The left-most column shows the quantile of the sky region depending on decreasing the determinant value of the beam-pattern matrix corresponding to the software-injected data. The next column is the determinant value of the beam-pattern function corresponding to the quantile of the sky region. The values of the next column indicate the value of the root-mean-square of the accuracy of the amplitude of the GWs for each data analysis method. y/n indicates whether the actual parameter points of the inclination-distance relationship are inside the 2σ credible region or not.

sky	det	w/o regulator	type 1	y/n	type 2	y/n
90%	5×10^{-1}	0.178684	0.165999	y	0.084327	y
48%	1×10^{-1}	0.194792	0.179279	y	0.102924	y
29%	5×10^{-2}	0.193852	0.188414	y	0.096755	y
18%	1×10^{-2}	0.273980	0.216209	n	0.115307	y
12%	5×10^{-3}	0.235382	0.187376	n	0.119479	y
6%	1×10^{-3}	0.389403	0.339281	y	0.211415	y
4%	5×10^{-4}	0.688487	0.223673	n	0.143216	n
2%	1×10^{-4}	1.863904	1.336913	y	0.262836	n
2%	5×10^{-5}	2.084366	0.311180	n	0.405315	n
1%	1×10^{-5}	5.901212	0.525819	n	0.215656	n

Table 6.6: The root-mean-square of the accuracy of the amplitude of the GWs from CBC. State 2 is used for the software-injected GW signal. The explanation of the table is as same as Table. 6.4.
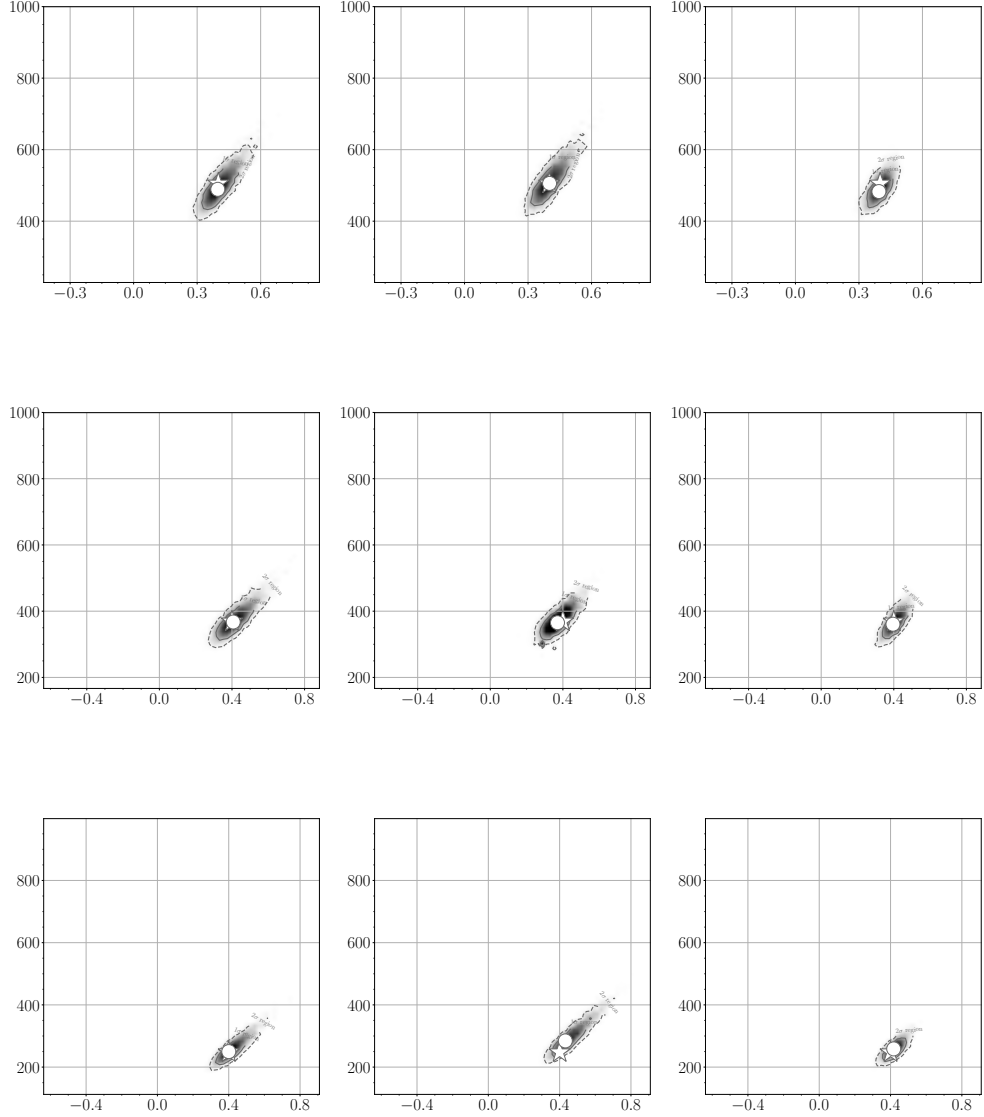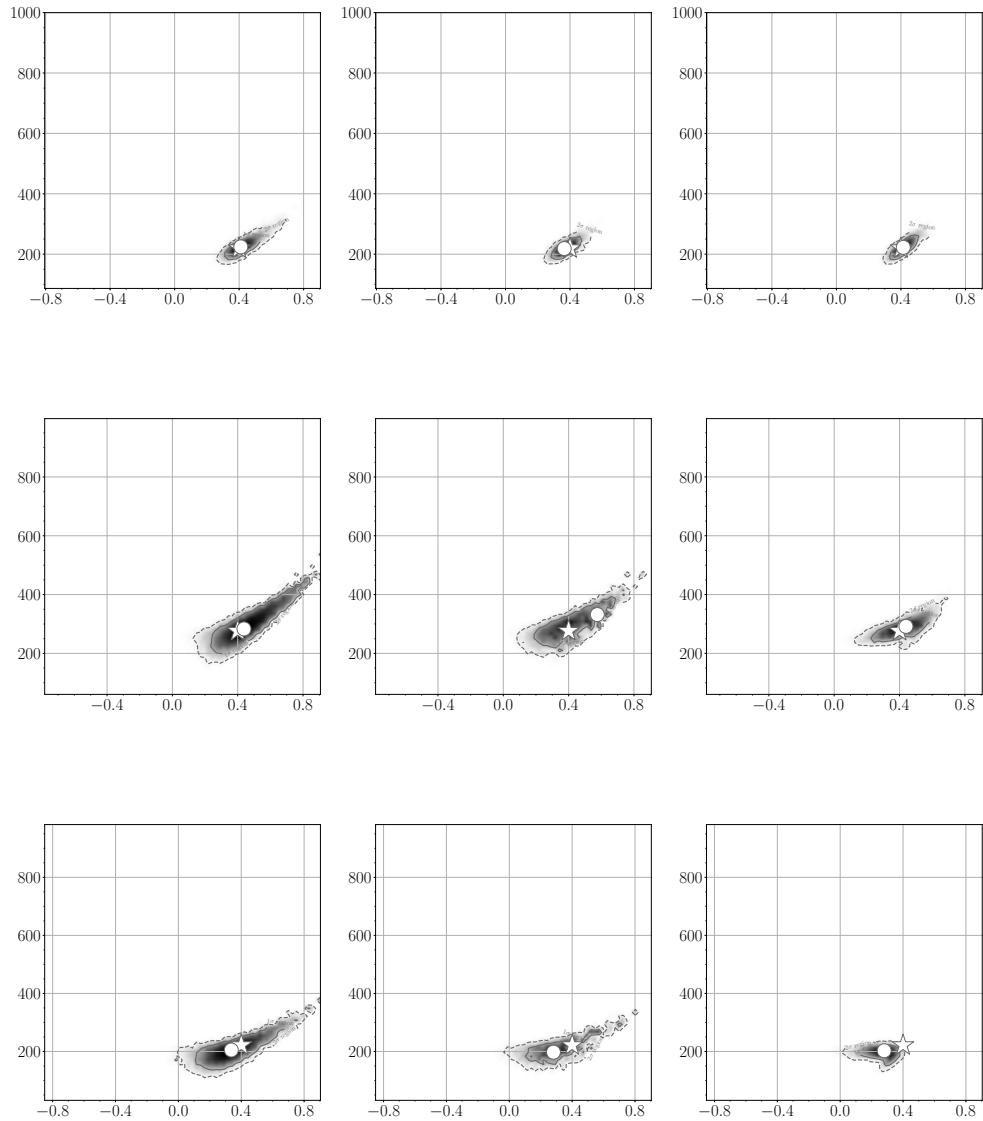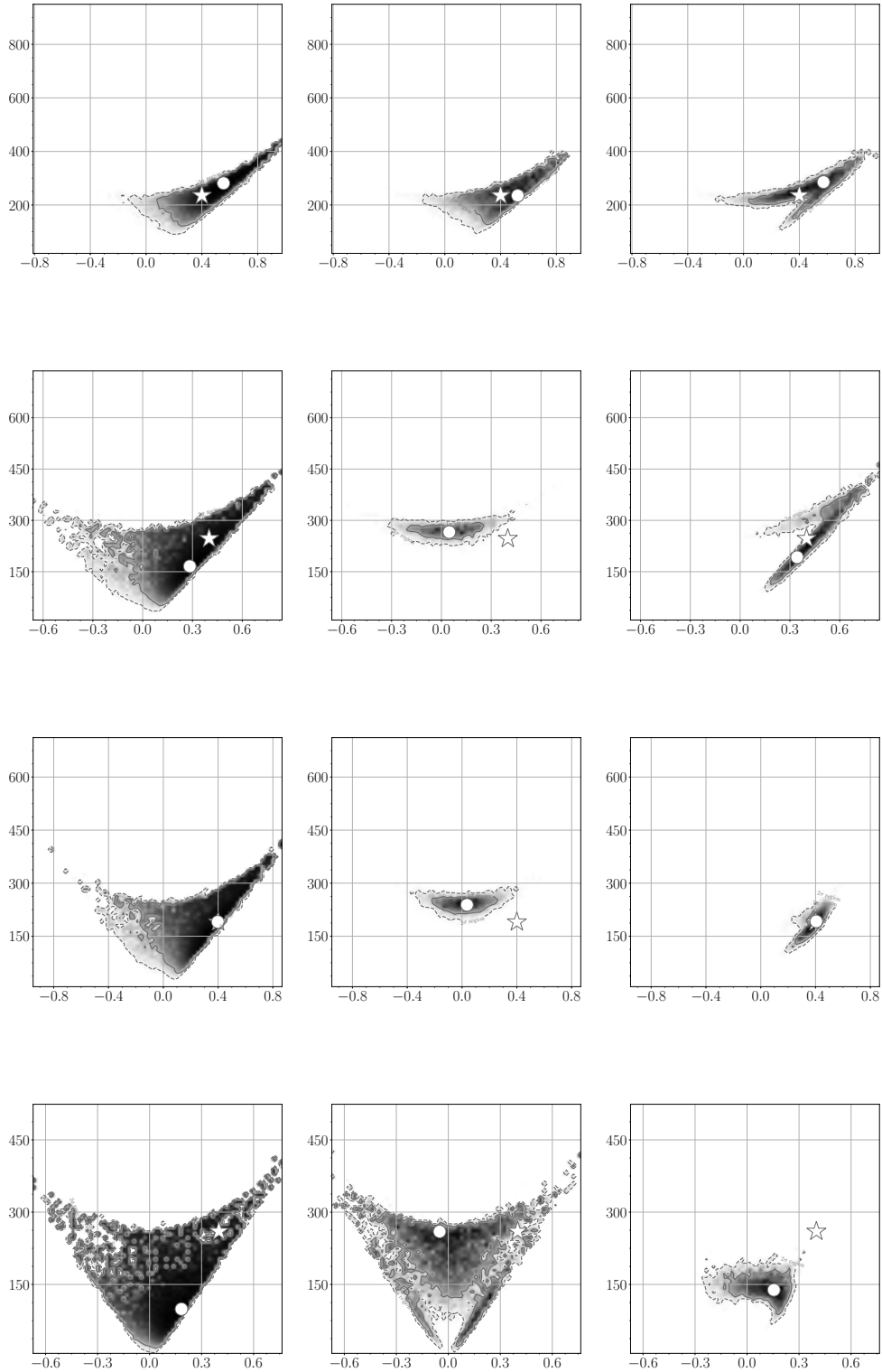
72

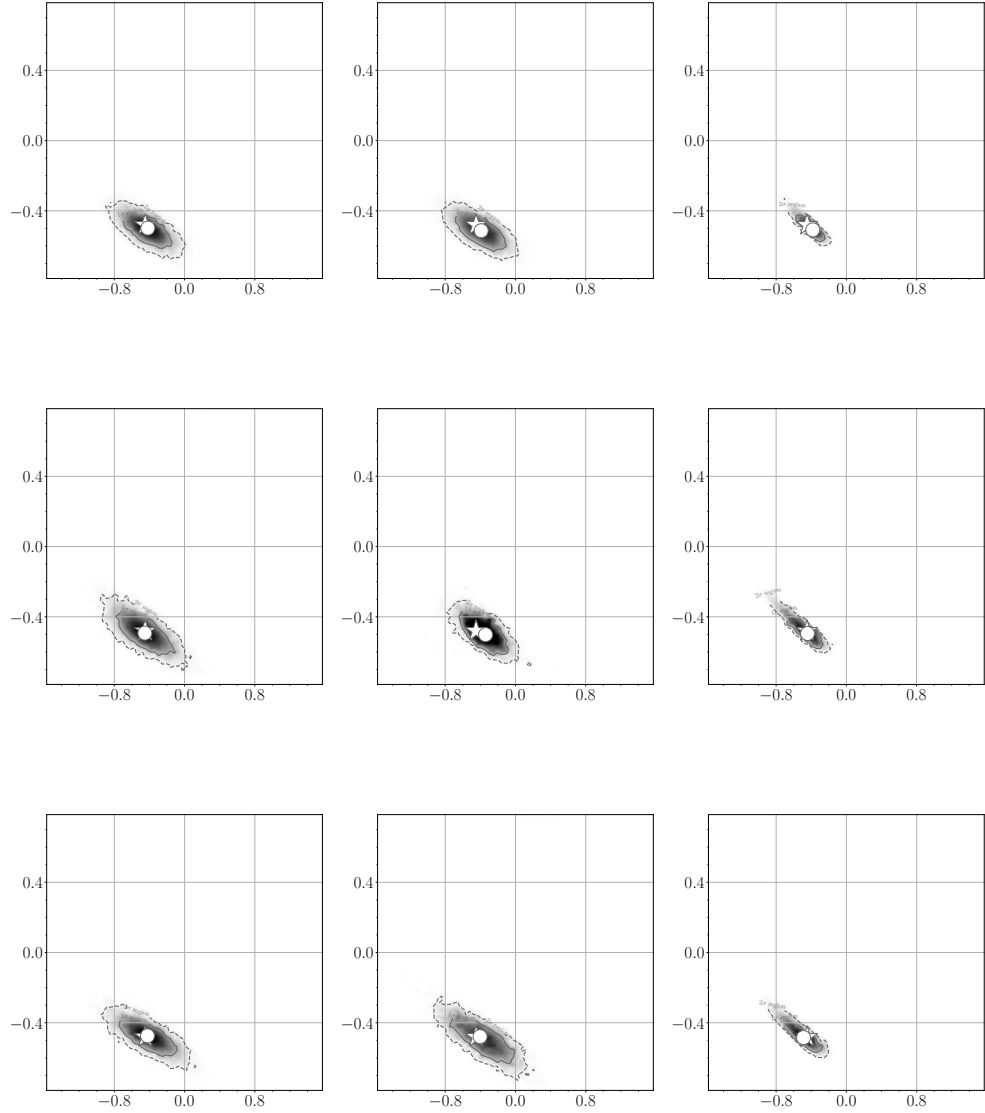Figure 6.2: The credible region of the joint-probability distribution represented by profile likelihood. The state 1 of the software-injected GW signal is used. The vertical scale indicates the luminosity distance and the horizontal scale indicates the inclination expressed in cosine. The left panel of all figure*s shows the result of non-regularized data analysis, the center panel of all figure*s shows the same of regularized data analysis with the type 1 regulator and the left panel of all figure*s shows the same of regularized data analysis with the type 2 regulator. The star indicate the actual parameters of the software-injected GW signal and the circle indicate the maximum likelihood parameters evaluated by the data analysis respectively. The determinant values of beam-pattern function matrix are arranged $5 \times 10^{-1}, 1 \times 10^{-1}, 5 \times 10^{-2}$ from top to bottom.
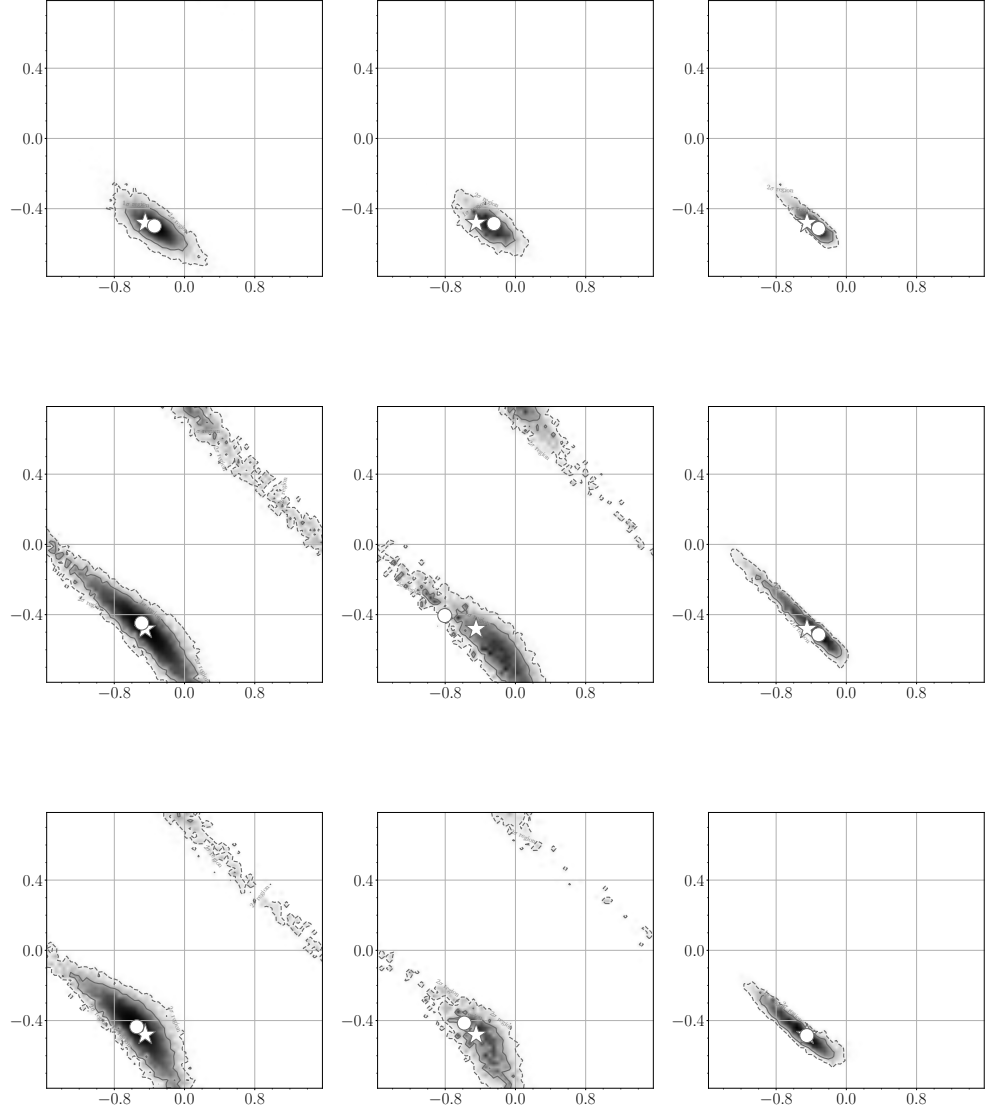
Figure 6.3: The probability distribution region represented by profile likelihood. The state 1 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.2. The determinant values of beam-pattern function matrix arrange $1 \times 10^{-2}, 5 \times 10^{-3}, 1 \times 10^{-3}$ from top to bottom.

Figure 6.4: The probability distribution region represented by profile likelihood. The state 1 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.2. The determinant values of beam-pattern function matrix are arranged $5\times10^{-4}, 1\times10^{-4}, 5\times10^{-5}, 1\times10^{-5}$ from top to bottom.
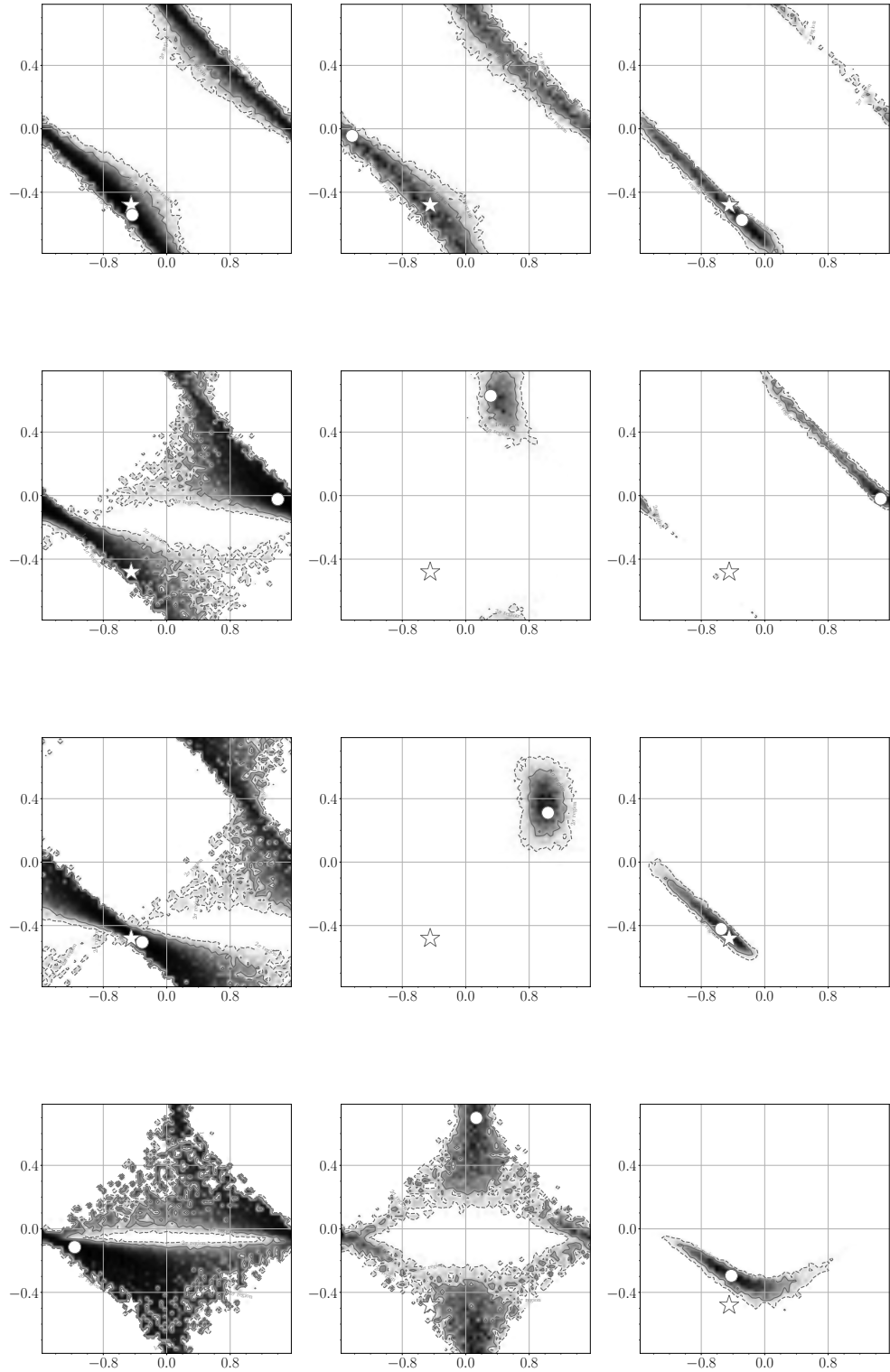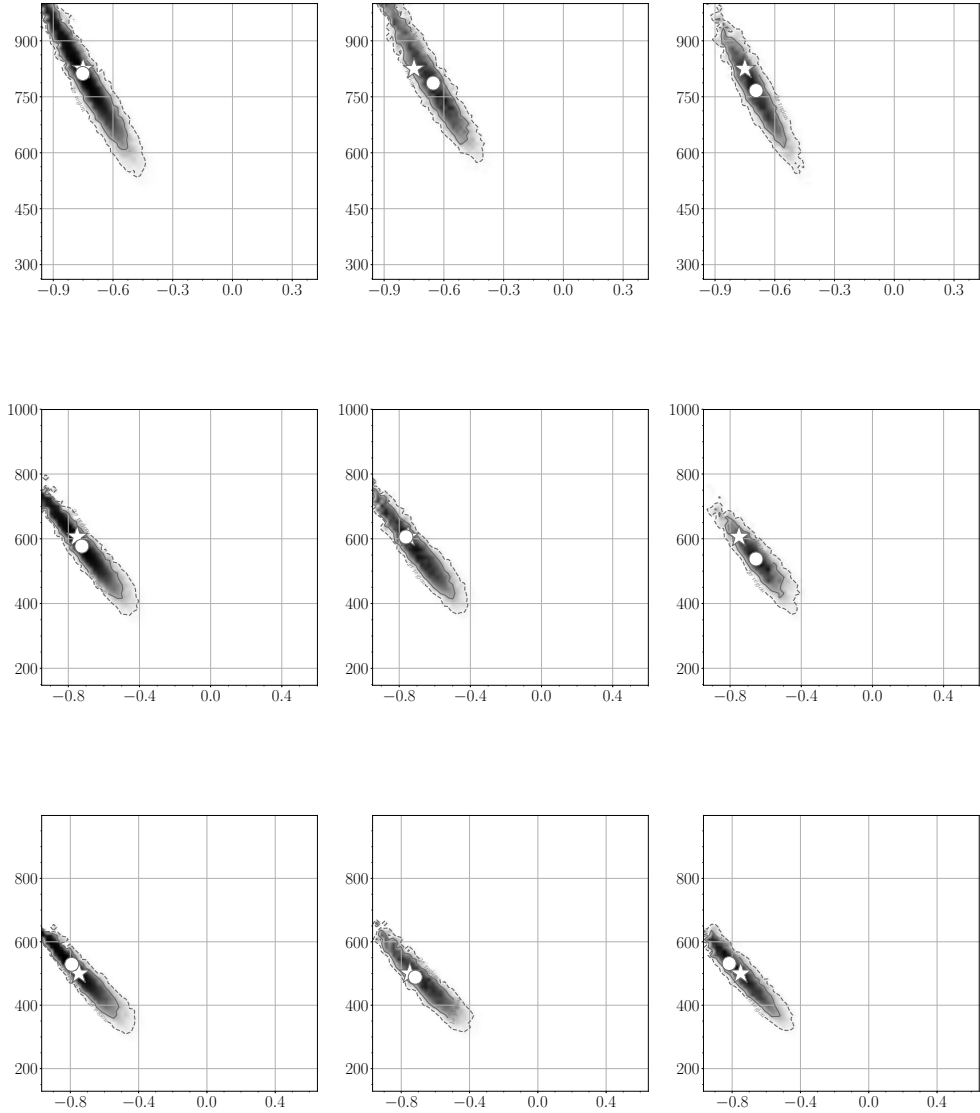
Figure 6.5: The probability distribution region represented by profile likelihood. The state 1 of the software-injected GW signal is used. The vertical scale the indicates initial phase and the horizontal scale indicates the polarization. The left panel of all figure*s shows the result of non-regularized data analysis, the center panel of all figure*s shows the same of regularized data analysis with the type 1 regulator and the left panel of all figure*s shows the same of regularized data analysis with the type 2 regulator. The star indicate the actual parameters of the software-injected GW signal and the circle indicate the maximum likelihood parameters evaluated by the data analysis respectively. The determinant values of beam-pattern function matrix are arranged $5 \times 10^{-1}, 1 \times 10^{-1}, 5 \times 10^{-2}$ from top to bottom.

Figure 6.6: The probability distribution region represented by profile likelihood. The state 1 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.5. The determinant values of beam-pattern function matrix are arranged $1 \times 10^{-2}, 5 \times 10^{-3}, 1 \times 10^{-3}$ from top to bottom.
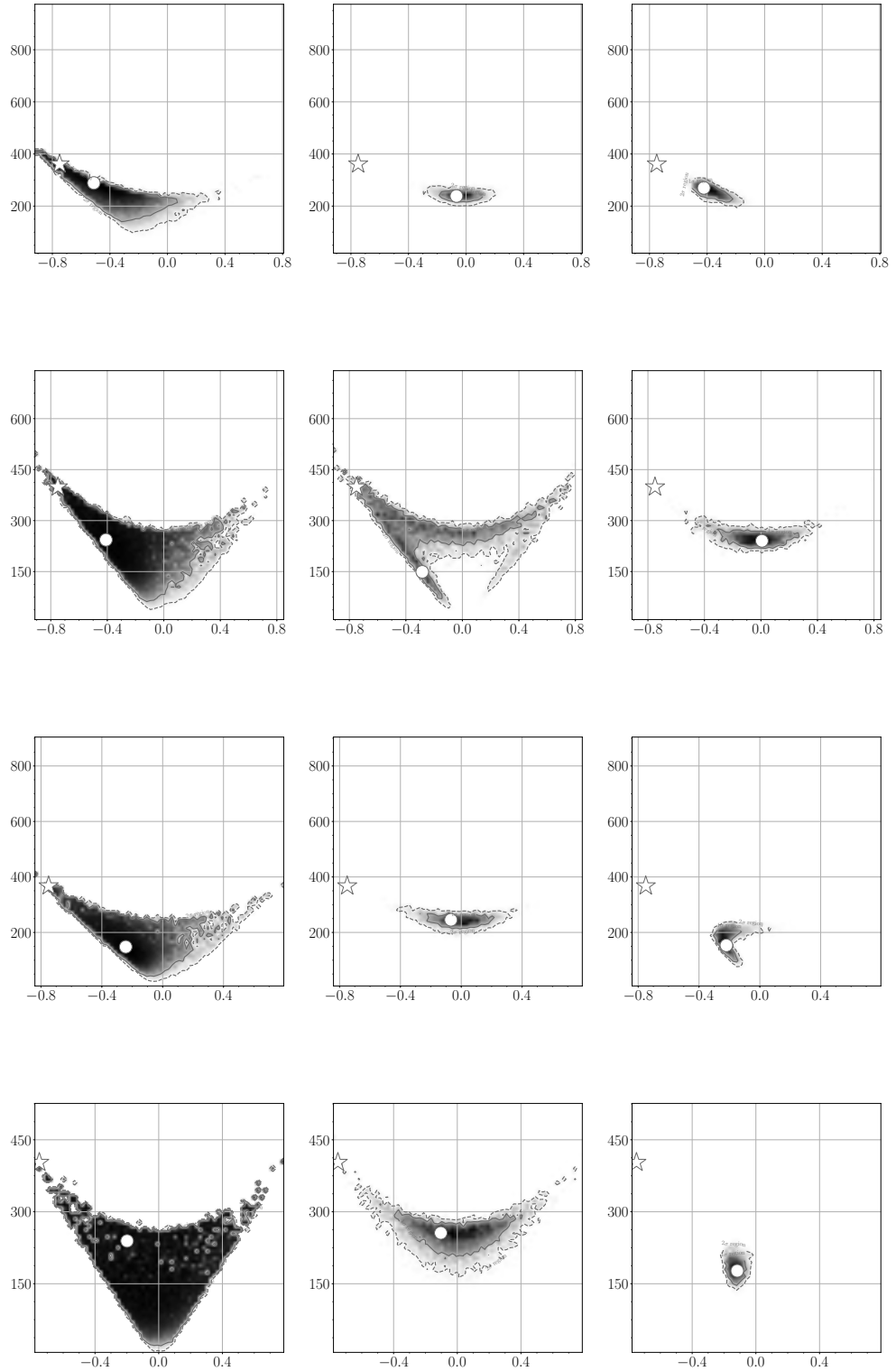
Figure 6.7: The probability distribution region represented by profile likelihood. The state 1 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.5. The determinant values of beam-pattern function matrix are arranged $5\times10^{-4}, 1\times10^{-4}, 5\times10^{-5}, 1\times10^{-5}$ from top to bottom.
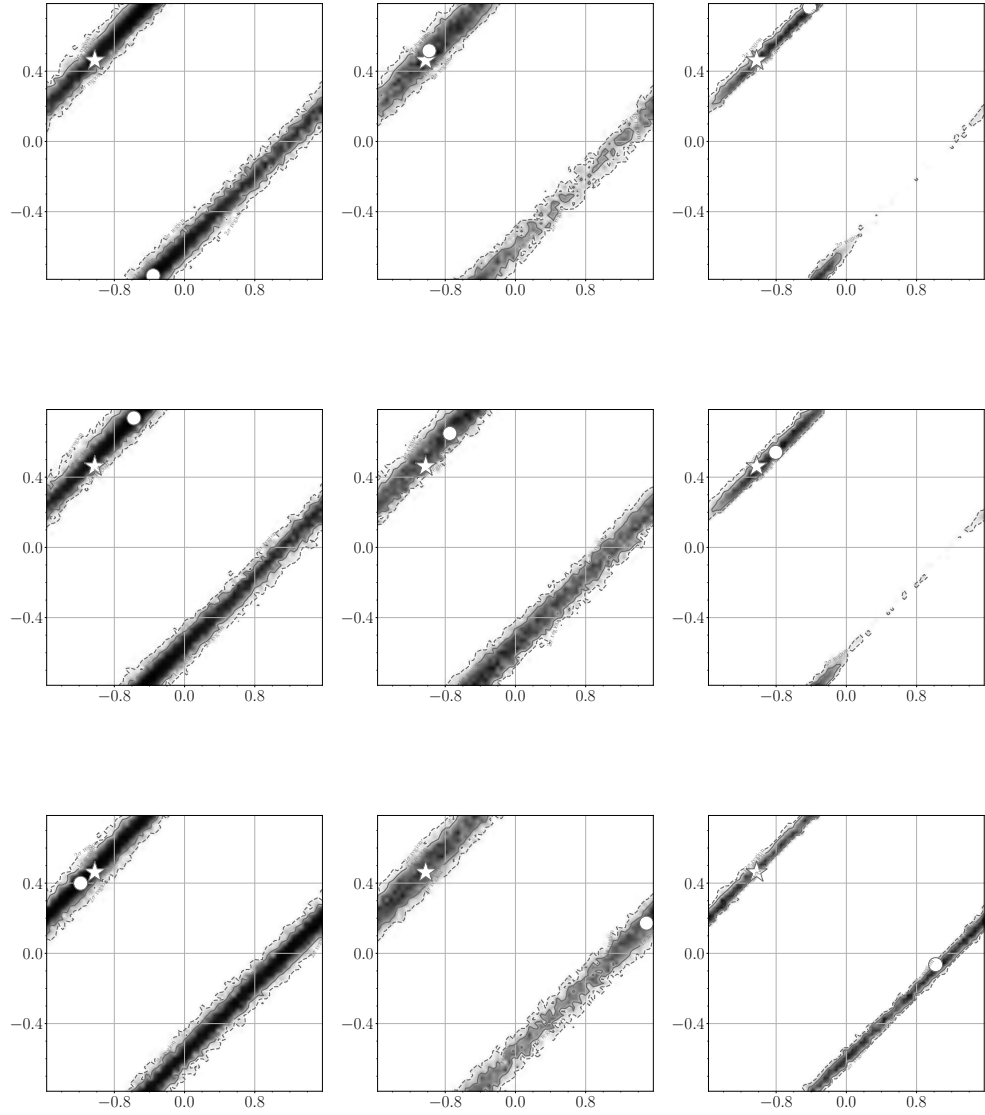
Figure 6.8: The probability distribution region represented by profile likelihood. The state 2 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.2. The determinant values of beam-pattern function matrix are arranged $5 \times 10^{-1}, 1 \times 10^{-1}, 5 \times 10^{-2}$ from top to bottom.
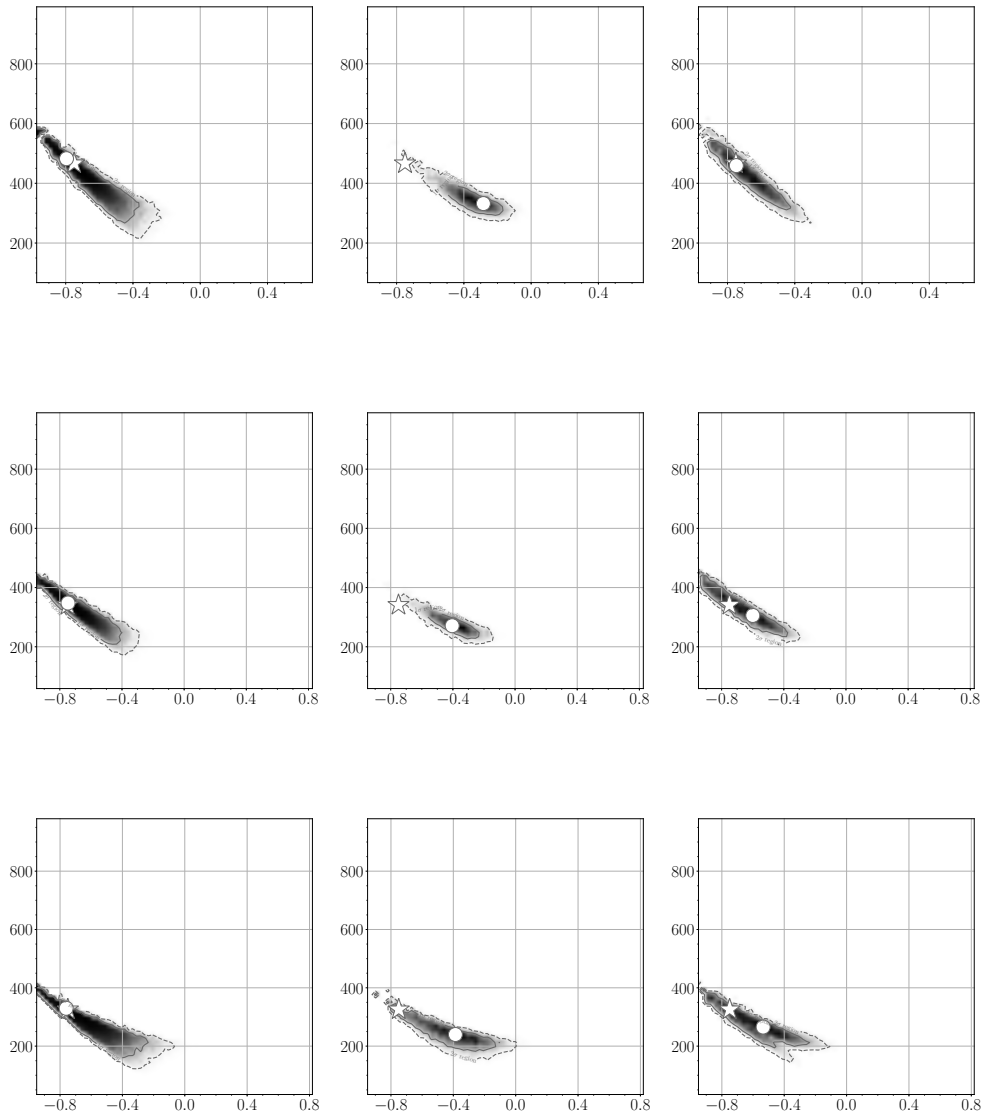
Figure 6.9: The probability distribution region represented by profile likelihood at the second data analysis. The explanation of these figure*s are as same as Fig. 6.2. The determinant values of beam-pattern function matrix are arranged $1 \times 10^{-2}, 5 \times 10^{-3}, 1 \times 10^{-3}$ from top to bottom.
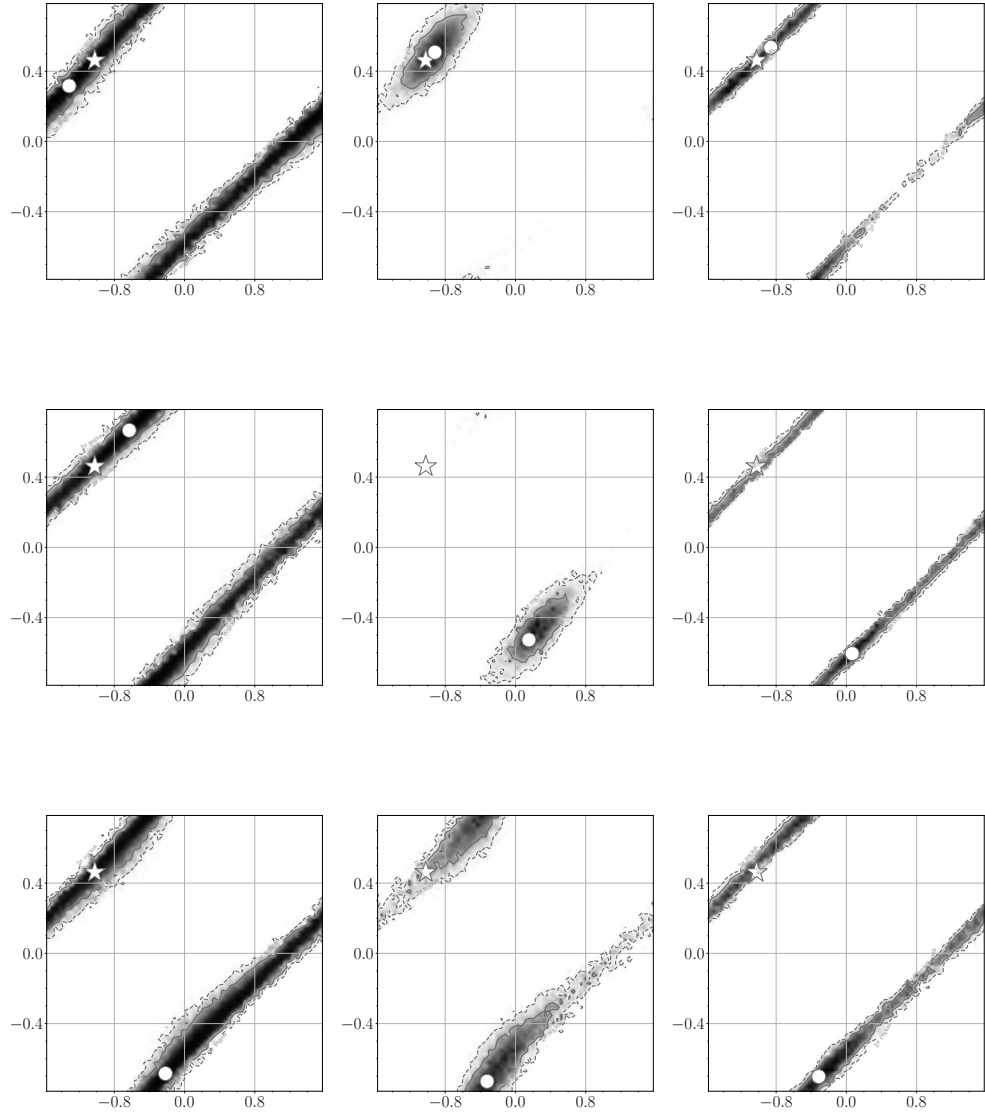
Figure 6.10: The probability distribution region represented by profile like-lihood. The state 2 of the software-injected GW signal is used. The expla-nation of these figure*s are as same as Fig. 6.2. The determinant values of beam-pattern function matrix are arranged $5 \times 10^{-4}, 1 \times 10^{-4}, 5 \times 10^{-5}, 1 \times 10^{-5}$ from top to bottom.

Figure 6.11: The probability distribution region represented by profile like-lihood. The state 2 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.5. The determinant values of beam-pattern function matrix are arranged $5 \times 10^{-1}, 1 \times 10^{-1}, 5 \times 10^{-2}$ from top to bottom.
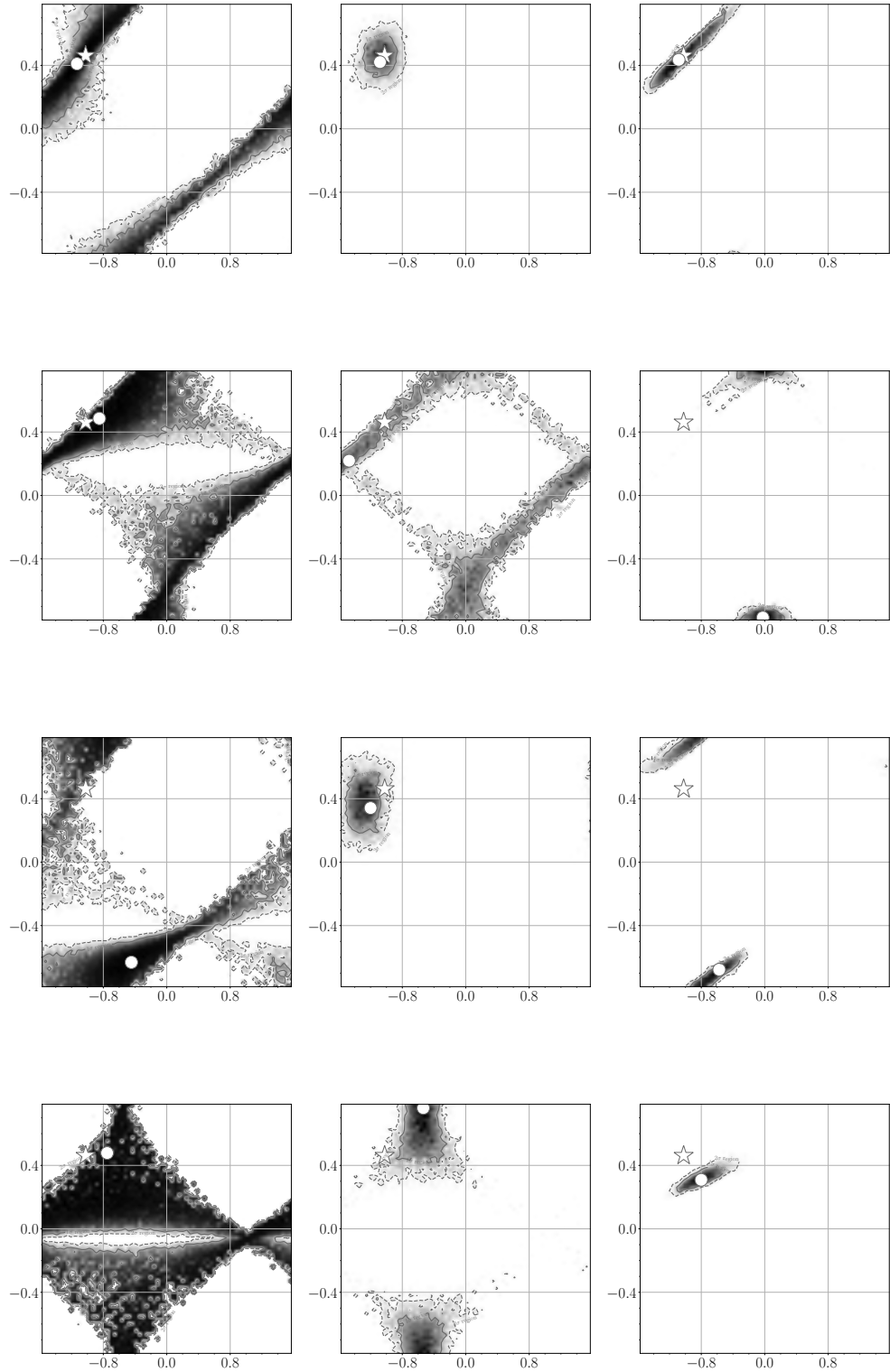
Figure 6.12: The probability distribution region represented by profile like-lihood. The state 2 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.5. The determinant values of beam-pattern function matrix are arranged $1 \times 10^{-2}, 5 \times 10^{-3}, 1 \times 10^{-3}$ from top to bottom.

84

Figure 6.13: The probability distribution region represented by profile like-lihood. The state 2 of the software-injected GW signal is used. The explanation of these figure*s are as same as Fig. 6.5. The determinant values of beam-pattern function matrix are arranged $5 \times 10^{-4}, 1 \times 10^{-4}, 5 \times 10^{-5}, 1 \times 10^{-5}$ from top to bottom.

Chapter 7

Conclusion

In this study, we discuss the method to prevent the accuracy of amplitude parameters of GWs from a CBC from being amplified owing to the ill-posed nature of the inverse operator comprising the beam-pattern functions of a network of GW telescopes. The key idea to resolve this problem is conpensation via an appropriate correction term represented as a matrix called a "regulator" for an inverse operator whose rank is deficient. The Tikhonov type regulator(Eq.(5.12), type 1) and symmetric trace-free type regulator regulator(Eq.(5.13), type 2) are used to express the regulator matrix. A regulator reduce the amplified noise caused by the ill-posed problem, whereas the value of the regulator affects the parameter accuracy as bias noise. To resolve the bias problem, the optimization of the regulator is evaluated by minimizing the sum of these noises. A Lagrange multiplier method with KKT condition for the norm of the difference between the amplitude parameters estimated by the regularized data analysis and amplitude parameters estimated by the non-regularization method (Eq.(5.9)), i.e., the a-posteriori parameter choice rule provides optimized regulator values. The data analysis for two types of software-injected GW signal using the MultiNest software was implemented to evaluate the reduction accuracy of the amplitude parameters using proposed method. The data analysis results indicate that the regularization method with the type 2 regulator reduces the credible region of the accuracy of the amplitude parameters. For approximately 90% of the sky region, the credible region of inclination-distance is reduced by approximately 1.5 times and that of the polarization-initial phase is reduced by approximately 3.0 times. The shrinkage rate of the credible region increases with a decreasing determinant value of the inverse operator; however the actual value lies outside the credible region in the regularized data analysis owing to the significant ill-posed inverse operator. The threshold to maintain

the validity of the regularization method appears to be related the inclination values.

We consider that the proposed method, which employ a regularization method for a targeted GW signal from a CBC and uses an optimized regulator, represents robust and coherent data analysis that enables us to minimize the amplified noise caused by an ill-posed nature of the inverse operator. In other words, for most of the sky region, it is possible that the accuracy of the amplitude parameters estimated by the proposed method is nearly the same as the limit of the accuracy as calculated by the SNR of the signal. Thus, the proposed method is suitable for the parameter estimation of a CBC search.

Appendix A

STF tensor and spherical harmonics

We refer [61] to describe the relationship between a STF tensor and a spherical harmonics. A STF(symmetric tracefree tensor) is the alternative representation of a spherical harmonics. While the decomposition in spherical harmonics depends on the spherical polar coordinate, the STF tensor relies on the Cartesian coordinates. So a STF tensor provides the way to express the spherical harmonics represented by a polar coordinates as a Cartesian coordinate.

First, we introduce the properties of the STF tensor represented by products of the unit vector. One example of such a STF tensor is

$$n^{\langle jk \rangle} = n^j n^k - \frac{1}{3} \delta^{jk}, \tag{A.1}$$

where the bracket indicates the STF label that the tensor with the bracket is STF tensor. You see that the index of Eq.(A.1) is commute and the trace of Eq.(A.1) is 0. The general formula for such STF tensor is

$$n^{\langle L \rangle} = \sum_{p=0}^{[l/2]} (-1)^p \frac{(2l - 2p - 1)!!}{(2l - 1)!!} \left[\delta^{2P} n^{L-2P} + \text{sym}(q) \right], \tag{A.2}$$

where L represents a collective of l individual indices, $[l/2]$ is the largest integer not larger than $l/2$, δ^P stands for a product of p Kronecker deltas, n^{L-2P} stands for a product of $l - 2p$ unit vectors and $\text{sym}(q)$ denotes all distinct terms arising from permuting indices.

To investigate the inner product of STF tensor Eq.(A.2) is useful to discuss a spherical harmonics. For any tensor A_L, and any STF tensor $B^{\langle L \rangle}$,

the product of these tensor indicates

$$A_L B^{\langle L \rangle} = A_{\langle L \rangle} B^{\langle L \rangle}, \tag{A.3}$$

where $A_{\langle L \rangle}$ is the STF tensor obtained from A_L by complete symmetrization and removal of all trace. Then we have

$$
\begin{aligned}
n'_{\langle L \rangle} n^{\langle L \rangle} &= n'_L n^{\langle L \rangle}, \\
&= n'_L \sum_{p=0}^{[l/2]} (-1)^p \frac{(2l-2p-1)!!}{(2l-1)!!} \left[\delta^{2P} n^{L-2P} + \text{sym}(q) \right], \\
&= \sum_{p=0}^{[l/2]} (-1)^p \frac{(2l-2p-1)!!}{(2l-1)!!} \left[\mu^{L-2P} + \text{sym}(q) \right], \\
&= \sum_{p=0}^{[l/2]} (-1)^p \frac{l!(2l-2p-1)!!}{(l-2p)!(2l-1)!!(2p)!!} \mu^{L-2P}, \\
&= \frac{l!}{(2l-1!!)} \frac{1}{2^l} \sum_{p=0}^{[l/2]} (-1)^p \frac{(2l-2p)!}{(l-2p)!p!(l-p)!} \mu^{L-2P}, \\
&= \frac{l!}{(2l-1)!!} P_l(\mu), \tag{A.4}
\end{aligned}
$$

where $\mu \equiv \mathbf{n}' \cdot \mathbf{n}$ and $P_l(\mu)$ is a Legendre function. When $\mathbf{n} \cdot \mathbf{n} = 1 = \mu$, the inner product Eq.(A.4) is

$$n_{\langle L \rangle} n^{\langle L \rangle} = \frac{l!}{(2l-1)!!}. \tag{A.5}$$

The inner product whose the number of indices is different from each tensors can also be established. From Eq.(A.5), we have

$$n_j n_{\langle L \rangle} n^{\langle jL \rangle} = n_{\langle jL \rangle} n^{\langle jL \rangle} = \frac{(l+1)!}{(2l+1)!!}. \tag{A.6}$$

By multiplying $n\langle L \rangle$ by the both side of Eq.(A.6), we have

$$
\begin{aligned}
n_j n_{\langle L \rangle} n^{\langle jL \rangle} n^{\langle L \rangle} &= \frac{(l+1)!}{(2l+1)!!} n^{\langle L \rangle}, \\
n_j n^{\langle jL \rangle} \frac{l!}{(2l-1)!!} &= \frac{(l+1)!}{(2l+1)!!} n^{\langle L \rangle}, \\
n_j n^{\langle jL \rangle} &= \frac{l+1}{2l+1} n^{\langle L \rangle}. \tag{A.7}
\end{aligned}
$$

In the same way, by multiplying $n\langle j \rangle$ by the both side of Eq.(A.6), we have

$$
\begin{aligned}
n_j n_{\langle L \rangle} n^{\langle jL \rangle} n^j &= \frac{(l+1)!}{(2l+1)!!} n^j \\
n_{\langle L \rangle} n^{\langle jL \rangle} &= \frac{(l+1)!}{(2l+1)!!} n^j .
\end{aligned}
\tag{A.8}
$$

Here, we consider the inner product of STF tensors whose coordinate system is different respectively such that

$$
n'_{\langle L \rangle} n^{\langle jL \rangle} = a n'^j + b n^j ,
\tag{A.9}
$$

where a and b are arbitrary numbers. In the same way deriving the Eq.(A.7) and Eq.(A.8), we have

$$
\begin{aligned}
n'_{\langle L \rangle} n^{\langle jL \rangle} &= a n'^j + b n^j , \\
n'_{\langle L \rangle} n^{\langle L \rangle} n^j n_j &= (a n'^j + b n^j) n_j , \\
\frac{(l+1)!}{(2l+1)!!} P_l(\mu) &= (a n'^j + b n^j) n_j , \\
P_l(\mu) &= \frac{(2l+1)!!}{(l+1)!} (a\mu + b) ,
\end{aligned}
\tag{A.10}
$$

and

$$
\begin{aligned}
n'_{\langle L \rangle} n^{\langle jL \rangle} &= a n'^j + b n^j , \\
n'_{\langle jL \rangle} n^{\langle jL \rangle} &= (a n'^j + b n^j) n'_j , \\
\frac{(l+1)!}{(2l+1)!!} P_{l+1}(\mu) &= (a n'^j + b n^j) n'_j , \\
P_{l+1}(\mu) &= \frac{(2l+1)!!}{(l+1)!} (a + b\mu) .
\end{aligned}
\tag{A.11}
$$

Since the derivation of the Legendre function is

$$
\frac{dP_l(\mu)}{d\mu} = \frac{l}{\mu^2 - 1} \left[\mu P_l(\mu) - P_{l-1}(\mu) \right] ,
\tag{A.12}
$$

the arbitrary number a and b are

$$
\begin{aligned}
a &= \frac{l!}{(2l+1)!!} \frac{1}{l+1} \frac{l+1}{\mu^2 - 1} \left[\mu P_l(\mu) - P_{l+1}(\mu) \right] , \\
&= -\frac{l!}{(2l+1)!!} \frac{dP_l(\mu)}{d\mu} ,
\end{aligned}
\tag{A.13}
$$

89

and

$$b = \frac{l!}{(2l+1)!!} \frac{1}{l+1} \frac{l+1}{\mu^2-1} \left[\mu P_{l+1}(\mu) - P_l(\mu) \right],$$

$$= \frac{l!}{(2l+1)!!} \frac{dP_{l+1}(\mu)}{d\mu}. \tag{A.14}$$

Finally, we obtain

$$n'_{\langle L \rangle} n^{\langle jL \rangle} = \frac{l!}{(2l+1)!!} \left[\frac{dP_{l+1}(\mu)}{d\mu} n^j - \frac{dP_l(\mu)}{d\mu} n'^j \right]. \tag{A.15}$$

Next, we wee the method to express a spherical harmonics represented by STF tensor. Here, we consider the arbitrary scalar field denoted by Ψ such as

$$\Psi = A_{\langle L \rangle} \frac{x^{\langle L \rangle}}{r^l}, \tag{A.16}$$

where $A_{\langle L \rangle}$ is arbitrary STF tensor and the unit vector is $n^j = x^j/r$. The derivative of the Eq.(A.16) is

$$\partial_j \Psi = A_{\langle L \rangle} \partial_j \frac{x^L}{r^l},$$

$$= l A_{\langle jL-1 \rangle} \frac{x^{L-1}}{r^l} - l A_{\langle L \rangle} \frac{x^L}{r^{l+1}} n_j. \tag{A.17}$$

Then we also have

$$\Delta \Psi = \partial_j \partial^j \Psi = l(l-1) A_{\langle jjL-2 \rangle} \frac{x^{L-2}}{r^l} - 2l^2 A_{\langle jL-1 \rangle} \frac{x^{L-1}}{r^{l+1}} n_j$$

$$+ l(l+1) A_{\langle L \rangle} \frac{x^{L-1}}{r^{l+2}} - 2l A_{\langle L \rangle} \frac{x^L}{r^{l+2}},$$

$$= -l(l+1)\Psi/r^2. \tag{A.18}$$

Since $A_{\langle L \rangle}$ is arbitrary STF tensor, we obtain the eigenvalue of the STF tensor such that

$$r^2 \Delta n^{\langle L \rangle} = -l(l+1) n^{\langle L \rangle}. \tag{A.19}$$

The eigenvalue of the STF tensor satisfies the same one of the spherical harmonics, we can define the STF tensor by using a spherical harmonics such that

$$n^{\langle L \rangle} \equiv N_l \sum_{m=-l}^{l} \mathcal{Y}_{lm}^{\langle L \rangle *} Y_{lm}(\theta, \phi), \qquad N_l \equiv \frac{4\pi l!}{(2l+1)!!}, \tag{A.20}$$

90

where $\mathcal{Y}_{lm}^{\langle L\rangle*}$ is a constant STF tensor which satisfies $\mathcal{Y}_{lm}^{\langle L\rangle} = \mathcal{Y}_{l,-m}^{\langle L\rangle*}$. The orthogonal property of the spherical function

$$\int Y_{l'm'}^*(\theta, \phi) Y_{lm}(\theta, \phi) d\Omega = \delta_{l'l}\delta_{m'm}, \tag{A.21}$$

allows us to evaluate the constant STF tensor such that

$$\mathcal{Y}_{lm}^{\langle L\rangle*} = \frac{1}{N_l} \int n^{\langle L\rangle} Y_{lm}^*(\theta, \phi) d\Omega. \tag{A.22}$$

Furthermore, since a Legendre function can be expressed by using a spherical harmonics such that

$$P_l(\mu) = \sum_{l=-m}^{m} \frac{4\pi}{2l+1} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi), \tag{A.23}$$

and Eq.(A.4) provides the inner product of the STF tensor, we finally obtain

$$
\begin{aligned}
\mathcal{Y}_{lm}^{\langle L\rangle*} n'_{\langle L\rangle} &= \frac{1}{N_l} \int n^{\langle L\rangle} n'_{\langle L\rangle} Y_{lm}^*(\theta, \phi) d\Omega \\
&= \frac{2l+1}{4\pi} \int P_l(\mu) Y_{lm}^*(\theta, \phi) d\Omega \\
&= \int \sum_{l=-m}^{m} \frac{4\pi}{2l+1} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi) Y_{lm}^*(\theta, \phi) d\Omega \\
&= Y_{lm}^*(\theta, \phi). \tag{A.24}
\end{aligned}
$$

Appendix B

Gaussian noise

In this appendix, we introduce the method to describe the gaussian noise by using stationary time-series data $x(t)$. we refer to [29] and [30] to describe the details of the section.

We suppose that the time-series data $x(t)$ is stationary random process, which is the statistical properties that the random process don't change with time. We define the ensemble average of the time-series such that

$$\langle x \rangle \equiv \int x p_x(x) dx, \tag{B.1}$$

where $p_x(x)$ is the probability distribution function of x. Because the time-series $x(x)$ which is a random process doesn't have the correlation between the data of the another frequency, we can define the power spectral density of the stationary random process $S_x(f)$ such that

$$\langle \tilde{x}^*(f)\tilde{x}(f') \rangle \equiv \frac{1}{2} S_x(f) \delta(f - f'), \tag{B.2}$$

where $\tilde{x}(f)$ is the Fourier transform of $x(t)$. Because of the reality of $x(t)$, negative frequency components are related to positive frequency components such that $\tilde{x}(-f) = \tilde{x}^*(f)$. The symmetricity of the power spectral density with respect to the frequency provides $S_x(-f) = S_x^*(f)$. Then, the variance of the time-series data $x(t)$ can be expressed by using the definition of the

power spectral density such as

$$
\begin{aligned}
\langle x(t)^2 \rangle &= \langle x(t=0)^2 \rangle, \\
&= \int_{-\infty}^{\infty} df\, df' \langle \tilde{x}^*(f)\tilde{x}(f') \rangle, \\
&= \frac{1}{2}\int_{-\infty}^{\infty} df\, S_x(f), \\
&= \int_{0}^{\infty} df\, S_x(f).
\end{aligned}
\tag{B.3}
$$

Furthermore, the definition of the delta function

$$
\delta(f=0) \rightarrow \left[\int_{-T/2}^{T/2} dt\, e^{i2\pi ft} \right]\Bigg|_{f=0} = T
\tag{B.4}
$$

provides the method to describe the expression of the power spectral density by substituting Eq.(B.4) for Eq.(B.2):

$$
\begin{aligned}
\langle |\tilde{x}(f)| \rangle &= \frac{1}{2}S_x(f)T, \tag{B.5} \\
S_x(f) &= 2\Delta f \langle |\tilde{x}(f)| \rangle, \tag{B.6}
\end{aligned}
$$

where $\Delta f = 1/T$.

Since the random process of time is ergodic, we have

$$
\langle x \rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)dt.
\tag{B.7}
$$

By substituting Eq.(B.7) for Eq.(B.6), we obtain the power spectral density with respect to time:

$$
S_x(f) = \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} x(t)e^{-2\pi ift} dt \right|^2.
\tag{B.8}
$$

Here, we define the autocorrelation function such as

$$
R_x(\tau) \equiv \langle x(t)x(t+\tau) \rangle,
\tag{B.9}
$$

and Wiener–Khinchin theorem provides

$$
S_x(f) = 2 \int_{-\infty}^{\infty} R_x(\tau)e^{-2\pi if\tau} d\tau.
\tag{B.10}
$$

93

Now we consider the discrete time-series data $x_j \equiv x(j\Delta t)$, where Δt is the interval of the discrete time and $N = T/\Delta t$ samples contain the discrete time-series data, where T is a integral time. Since the probability density distribution of the discrete-time series $p_x(x_j)$ is gaussian with the variance σ^2, it can be expressed by

$$p_x(x_j) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{j=0}^{N-1} x_j^2\right]. \tag{B.11}$$

The approximation of the Fourier transformation

$$\lim_{\Delta t \to \infty} \sum_{j=0}^{N-1} x_j^2 = \int_0^T x^2(t)dt \sim \int_{-\infty}^{\infty} |\tilde{x}(f)|^2 df \tag{B.12}$$

is valid when the integral time T is sufficiently long. Furthermore, the auto-correlation function can be described by using Kronecker delta

$$R_{ij} = \langle x_i x_j \rangle = \sigma^2 \delta_{ij}. \tag{B.13}$$

By substituting Eq.(B.13) for Eq.(B.10), we obtain

$$S_x(f) = \lim_{\Delta t \to 0} 2\sigma^2 \Delta t. \tag{B.14}$$

In summarize, the gaussian distribution of the discrete time-series data can be described as

$$
\begin{aligned}
\lim_{\Delta t \to 0} \exp\left[-\frac{\sum_{j=0}^{N-1} x_j^2}{2\sigma^2}\right] &= \lim_{\Delta t \to 0} \exp\left[-\frac{\sum_{j=0}^{N-1} x_j^2 \Delta t}{2\sigma^2 \Delta t}\right], \\
&= \exp\left[-\frac{1}{S_x}\int_0^T x(t)^2 dt\right], \\
&\sim \exp\left[-\frac{1}{2}4\int_0^{\infty} \frac{|\tilde{x}(f)|^2}{S_x} df\right]. \tag{B.15}
\end{aligned}
$$

So the proportional relation of the probability density distribution of the stationary time-series data with random process is approximated by

$$p_x[x(t)] \propto \exp\left[-\frac{1}{2}4\int_0^{\infty} \frac{|\tilde{x}(f)|^2}{S_x} df\right], \tag{B.16}$$

which also indicates the probability density distribution of the gaussian. Here, we define the inner product (a, b) between two arbitrary time-series data $a(t)$ and $b(t)$ such as.

$$(a, b) \equiv 4Re \int_0^\infty \frac{\tilde{a}^*(f)\tilde{b}(f)}{S(f)} df, \tag{B.17}$$

where $S(f)$ is a one-sided power spectral density which contains the total power of the signal in the frequency interval to half of the Nyquist rate. Eq.(B.16) provides the probability density distribution of the gaussian such as

$$p_x[x(t)] \propto e^{-(x,x)/2}. \tag{B.18}$$

Appendix C

Spectral theory and Hilbert space

The construction of the theory of Regularization method is based on a Hilbert space and its applications. I introduce some definitions and theorems of a Hilbert space which is important for a regularization method theory briefly in this appendix. I refer to [62] [63] [55] [26] [64] [65] and [66] to show the definitions and theorems.

C.1 Definition of a Hilbert space

First, we introduce the simple definition of a Hilbert space and related spaces of the Hilbert space.

Definition C.1.1. Normed space

A function $x \mapsto \|x\|$ from a vector space E into $\mathbb{R}$ is called a *norm* if it satisfies the following conditions:

1. $\|x\| = 0$ implies $x = 0$;

2. $\|\lambda x\| = |\lambda|\|x\|$ for every $x \in E$ and $\lambda \in \mathbb{R}$ or $\lambda \in \mathbb{C}$;

3. $\|x + y\| \leq \|x\| + \|y\|$ for every $x, y \in E$;

A vector space with a norm is called a *normed space.*

Definition C.1.2. Inner product space

Let E be a complex vector space. A mapping $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{C}$ is called an *inner product* in E if for any $x, y, z \in E$ and $\alpha, \beta \in \mathbb{C}$, the following conditions are satisfied:

1. $\langle x, y \rangle = \overline{\langle y, x \rangle}$;

2. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;

3. $\langle x, x \rangle \geq 0$;

4. $\langle x, x \rangle = 0$ implies $x = 0$;

A vector space with an inner product is called an *inner product space.*

A normed space and inner product space can be related with the following definition:

Definition C.1.3. Norm in an inner product space

In an inner product space E and the vector $x \in E$, the norm induced by the inner product on E, denoted by $\|x\|$, is the nonnegative real number as defined by

$$\|x\| = \sqrt{\langle x, x \rangle}. \tag{C.1}$$

The validity of the definition can indicate simply. The first and second conditions of the Definition.C.1.1 is trivial. The third condition means the Triangle inequality which also holds the inner product space[62]. Then we can say that every inner product space is a normed space defined by Definition.C.1.3. More useful definitions of the construction of a Hilbert space is about the completeness:

Definition C.1.4. Cauchy sequence in norm

A sequence of vectors (x_n) in a normed space is called a *Cauchy sequence* if for every $\epsilon > 0$, thereexists a number M such that $\|x_m - x_n\| < \epsilon$ for all $m, n > M$.

Definition C.1.5. Completeness and a Banach space

A normed space E is called complete if every Cauchy sequence in E converges to an element of E. A complete normed space is called a *Banach space.*

By using above discussions, we can define a Hilbert space as a following:

Definition C.1.6. Hilbert space

A complete inner product space denoted by H is called a *Hilbert space.*

The notable properties of a Hilbert space will be introduced by the below sections.

C.2 Spectral theorem for compact, self-adjoint operator

A compact self-adjoint operators in H are of special interest since it can be construct an eigensystem of the operators and allows us to obtain the fundamental theory for a Spectral theory.

Definition C.2.1. Compact operator

An operator A on a Hilbert space H is called a compact operator if, for every bounded sequence (x_n) in H, the sequence (Ax_n) contains a convergent sequence

Lemma C.2.2. *Let T be a self-adjoint operator on a Hilbert space H, then*

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle|. \tag{C.2}$$

Proof. Let

$$M = \sup_{\|x\|=1} |\langle Tx, x \rangle|. \tag{C.3}$$

Now suppose $\|x\| = 1$, then

$$|\langle Tx, x \rangle| \leq \|T\|\|x\|\|x\| = \|T\|. \tag{C.4}$$

Thus

$$M \leq \|T\|. \tag{C.5}$$

On the other hand, for all $x, y \in H$, we have

$$\langle T(x + z), x + z \rangle - \langle T(x - z), x - z \rangle = 4\mathrm{Re}\,\langle Tx, z \rangle \tag{C.6}$$

Therefore,

$$\mathrm{Re}\,\langle Tx, z \rangle \leq \frac{M}{4}\left(\|x + z\|^2 + \|x - z\|^2\right) = \frac{M}{2}\left(\|x\|^2 + \|z\|^2\right). \tag{C.7}$$

Now suppose $\|x\| = 1$ and $Tx \neq 0$. Let $z = Tx/\|Tx\|$, then

$$\mathrm{Re}\,\langle Tx, z \rangle = \|Tx\| \leq M. \tag{C.8}$$

By Eq.(C.5) and Eq.(C.8), we have the desired result. $\qquad\square$

The existence of at least one nonzero eigenvalues for compact, self-adjoint operator on Hilbert space is guaranteed by following lemma.

Lemma C.2.3. *If A is a compact, self-adjoint operator on a Hilbert space, then at least one of the numbers $\|A\|$ or $-\|A\|$ is an eigenvalue of A*

Proof. $A = 0$ is a trivial case. then assume that A is a non-zero, compact, self-adjoint operator on a Hilbert space H. Because of $\|A\| = \sup_{\|x\|=1} |\langle Ax, x \rangle|$, there exists a sequence $\{x_n \in H | \|x_n\| = 1, |\langle Ax, x \rangle| \to \|A\| \text{ as } n \to \infty\}$. Here, let λ be a eigenvalue of the operator such as $Ax = \lambda x$. It can translate to $\langle Ax, x \rangle = \lambda \langle () x, x \rangle$, then without loss of generality, we can assume that $\langle Ax_n, x_n \rangle \to \lambda$. For every $n \in \mathbb{N}$, we have

$$
\begin{aligned}
\|Ax_n - \lambda x_n\|^2 &\leq \|A\|^2 - 2\lambda \langle Ax_n, x_n \rangle + \lambda^2 \\
&= 2\lambda(\lambda - \langle Ax_n, x_n \rangle) \to 0.
\end{aligned} \tag{C.9}
$$

Thus

$$
Ax_n - \lambda x_n \to 0 \quad \text{as } n \to \infty. \tag{C.10}
$$

Compactness of A implies the sequence (x_n) has a subsequence (x_{p_n}) such that the sequence (Ax_{p_n}) converges. It follows that there exist $u \in H$ such that $x_{p_n} \to u$. Then we obtain the linear equation such that $Au = \lambda u$ and we have the desired result. $\qquad \square$

Next Corollary gives us a useful way to find the eigenvalue by maximizing a quadratic expression.

Corollary C.2.4. *If A is a compact, self-adjoint operator on a Hilbert space, H, then there exists a vector $w \in H$ such that $\|w\| = 1$ and*

$$
|\langle Aw, w \rangle| = \sup_{\|x\| \leq 1} |\langle Ax, x \rangle|. \tag{C.11}
$$

Proof. Let w be an eigenvector corresponding to an eigenvalue λ such that $\delta = \|A\|$. By lemma.C.2.2, we obtain

$$
|\langle Aw, w \rangle| = |\langle \lambda w, w \rangle| = \lambda = \|A\| = \sup_{\|x\| \leq 1} |\langle Ax, x \rangle|. \tag{C.12}
$$

$\qquad \square$

Theorem C.2.5. Hilbert-Schmidt theorem

For every compact, self-adjoint operator A on an infinite-dimensional Hilbert space H, there exist an orthonormal system of eigenvectors (u_n) corresponding to nonzero eigenvalues (λ_n) such that every element $x \in H$ has a unique representation in the form

$$
x = \sum_{n=1}^{\infty} \alpha_n u_n + v, \tag{C.13}
$$

where $\alpha_n \in \mathbb{C}$ and $v \in \mathcal{N}(A)$.

Proof. By Lemma.C.2.3 and Corollary.C.2.4, there exists an eigenvalue λ_1 of A such that

$$|\lambda_1| = \sup_{\|x\| \leq 1} |\langle Ax, x\rangle|. \tag{C.14}$$

Let u_1 be a normalized eigenvector corresponding to λ_1 and

$$Q_1 = \{x \in H | x \perp u_1\}. \tag{C.15}$$

Q_1 is a closed linear subspace of H so that if $x \in Q$, then $Ax \in Q_1$ because of

$$\langle Ax, u\rangle = \langle x, Au\rangle = \lambda_1 \langle x, u\rangle = 0. \tag{C.16}$$

Therefore, we can say that the A maps the Hilbert space Q_1 into itself. It implies that the other eigenvalue λ_2 of A can be given by

$$|\lambda_2| = \left\{ \sup_{\|x\| \leq 1} |\langle Ax, x\rangle| \, | \, x \in Q_1 \right\}. \tag{C.17}$$

Let u_2 be a normalized eigenvector corresponding to λ_2 and it is clear that $u_1 \perp u_2$, then we set

$$Q_2 = \{x \in Q_1 | x \perp u_2\}. \tag{C.18}$$

and repeat the above argument. Having eigenvalues $\lambda_1 \cdots \lambda_n$ and corresponding eigenvectors $u_1 \cdots u_n$, we set

$$Q_n = \{x \in Q_{n-1} | x \perp u_n\}. \tag{C.19}$$

Suppose that there is a $k \in \mathbb{N}$ such that $\langle Ax, x\rangle = 0$ for every $x \in Q_k$. Then every element $x \in H$ has a unique representation given by

$$x = \sum_{n=1}^{k} \alpha_n u_n + v, \tag{C.20}$$

whrere $Av = 0$ and

$$Ax = \sum_{n=1}^{k} \alpha_n \lambda_n u_n, \tag{C.21}$$

which is the desired representation as a case of a finite elements.

Suppose that the described procedure yields an infinite sequence of eigenvalues (λ_n) and eigenvectors u_n. Let S be the closed space spanned by the eigenvectors such that

$$S = \left\{ \sum_{n=1}^{\infty} \alpha_n u_n | \sum_{n=1}^{\infty} \|\alpha_n\|^2 < 0 \right\}. \tag{C.22}$$

100

Every $x \in H$ has a unique decomposition

$$x = \sum_{n=1}^{\infty} \alpha_n u_n + v, \tag{C.23}$$

where $v \in S^{\perp}$.

Let $\left\{ v \in S^{\perp} | v \neq 0 \right\}$ and we set $w = v/\|v\|$. Then we obtain $\langle Av, v \rangle = \|v\|^2 \langle Aw, w \rangle$.

Since $w \in S^{\perp} \subset Q_n$ for every $n \in \mathbb{N}$, we have

$$
\begin{aligned}
|\langle Av, v \rangle| &= \|v\|^2 \|\langle Aw, w \rangle\| \leq \|v\|^2 \left\{ \sup_{\|x\| \leq 1} |\langle Ax, x \rangle| \, | \, x \in Q_n \right\} \\
&= \|v\|^2 |\lambda_{n+1}| \to 0. \tag{C.24}
\end{aligned}
$$

This implies $\langle Av, v \rangle = 0$ for every $v \in S^{\perp}$. The norm of A restricted to $S^{\perp}$ is 0 by Lemma.C.2.3, then $Av = 0$ for all $v \in S^{\perp}$. $\qquad \square$

This theorem leads to a important representation of a eigensystem of a Hilbert space.

Corollary C.2.6. Spectral theorem for compact self-adjoint operators

Let A be a compact, self-adjoint operator on an infinite-dimensional Hilbert space H. Then H has a complete orthonormal system $\{u_1, u_2, \cdots\}$ consisting of A eigenvector of A. Moreover, for every $x \in H$,

$$Ax = \sum_{n=1}^{\infty} \lambda_n \langle x, u_n \rangle u_n, \tag{C.25}$$

whrere λ_n is the eigenvalue corresponding to u_n

Proof. Eq.(C.13) implies that

$$\langle x, u_m \rangle = \left\langle \sum_{n=1}^{\infty} \alpha_n u_n, u_m \right\rangle + \langle v, u_m \rangle = \alpha_m, \tag{C.26}$$

where $n \neq m$. Since Eq.(C.13), Eq.(C.26) and $Av = 0$, we have the desired result. $\qquad \square$

Polynomials of the operator can also be represented by using these discussion.

Corollary C.2.7. *Let A be a compact, self-adjoint operator on an infinite-dimensional Hilbert space H. For any polynomial $p(t) = \alpha_n t^n$, we have*

$$p(A)x = \sum_{n=1}^{\infty} p(\lambda_n) \langle x, u_n \rangle u_n, \tag{C.27}$$

whrere λ_n is the eigenvalue corresponding to a complete orthonormal vector u_n.

Proof. From Eq.(C.25), $A^2 x$ is calculated as

$$A^2 x = A(Ax) = \sum_{m=1}^{\infty} \lambda_m \left\langle \sum_{n=1}^{\infty} \lambda_n \langle x, u_n \rangle u_n, u_m \right\rangle u_m = \sum_{n=1}^{\infty} \lambda_n^2 \langle x, u_n \rangle u_n. \tag{C.28}$$

Similarly, for any $k \in \mathbb{N}$, we have

$$A^k x = \sum_{n=1}^{\infty} \lambda_n^k \langle x, u_n \rangle u_n. \tag{C.29}$$

Therefore we have the desired result. $\square$

This method can be generalized as a following definition:

Definition C.2.8. continuous function of an operator
Let f be a continuous function on $\mathbb{R}$ such that

$$f(\lambda) \quad \to \quad 0 \quad \text{as} \quad \lambda \to 0, \tag{C.30}$$
$$f(0) \quad = \quad 0. \tag{C.31}$$

For a compact, self-adjoint operator A, we define

$$f(A)x = \sum_{n=1}^{\infty} f(\lambda_n) \langle x, u_n \rangle u_n, \tag{C.32}$$

where the symbols are as in Corollary.C.2.6.

It is convenient that we construct the eigensystem of the compact, *non-self-adjoint*(say, adjoint), linear operator of a Hilbert space.

Definition C.2.9. A singular system
Let $\mathcal{X}, \mathcal{Y}$ be subspaces of a Hilbert space H and let $K : \mathcal{X} \to \mathcal{Y}$ is a compact adjoint linear operator on Hilbert space. If $K^* : \mathcal{Y} \to \mathcal{X}$ is a adjoint

of the operator, then the operators $K^*K : \mathcal{X} \to \mathcal{X}$ and $KK^* : \mathcal{Y} \to \mathcal{Y}$ are compact, self-adjoint linear operators. Let an orthonormal system (v_n) be an eigenvectors of the operator K^*K corresponding to nonzero eigenvalues (σ_n^2). We define the orthonormal system (u_n)

$$u_n \equiv \frac{Kv_n}{\|Kv_n\|}, \tag{C.33}$$

corresponding to an eigenvector of the operators KK^*. In these statement, the tuple of (σ_n, v_n, n_n) is called a *singular system* and a following equations hold:

$$Kv_n = \sigma_n u_n \tag{C.34}$$
$$K^*u_n = \sigma_n v_n \tag{C.35}$$

As a analogies of the Corollary.C.2.6, the eigensystem of these adjoint operators can be expanded as a decomposition.

Theorem C.2.10. Singular value expansion

Let $\mathcal{X}, \mathcal{Y}$ be subspaces of a Hilbert space H and $x \in \mathcal{X}$, $y \in \mathcal{Y}$ be a element of these subspaces. Let $K : \mathcal{X} \to \mathcal{Y}$ be a compact adjoint linear operator on Hilbert space and $K^* : \mathcal{Y} \to \mathcal{X}$ be a adjoint of the operator. If the operator has a singular system (σ_n, v_n, u_n),

$$Kx = \sum_{n=1}^{\infty} \sigma_n \langle x, v_n \rangle u_n \tag{C.36}$$

$$K^*y = \sum_{n=1}^{\infty} \sigma_n \langle y, u_n \rangle v_n \tag{C.37}$$

are hold for every elements x and y.

This is a trivial result by using Definition.C.2.9

The construction of the singular value expansion is also applied to a M-P inverse operator defined as Sec.4.2. This expression is very important for a regularization method because to express an arbitrary inverse operator as a function of a singular system allows us to calculate the amount of the error easily.

Lemma C.2.11. Let (σ_n, v_n, u_n) be a singular system for the compact linear operator K and $y \in \mathcal{Y}$ is in a Hilbert space. Then we have

$$y \in \mathcal{D}(K^\dagger) \iff \sum_{n=1}^{\infty} \frac{\langle y, u_n \rangle^2}{\sigma_n^2} < \infty. \tag{C.38}$$

Proof. Let Q be a orthogonal projector onto $R(K)$. Now we have $y \in \mathcal{D}^\dagger(K) = \mathcal{R}(K) + \mathcal{R}(K)^\perp$, then $Qy \in \mathcal{R}(K)$. From Eq.(C.37) and Eq.(C.34), the (u_n) span $\mathcal{R}(KK^*) = \mathcal{R}(K)$. Since $Qy \in \mathcal{R}(K)$, the expression of the Q as a singular system is given by

$$Q = \sum_{n=1}^{\infty} \langle \cdot, u_n \rangle \, u_n. \tag{C.39}$$

From Eq.(C.39) and Eq.(C.36), we have $QKx = Kx$ and also have $Kx = Qy$. Then

$$Kx = \sum_{n=1}^{\infty} \sigma_n \langle x, v_n \rangle \, u_n = \sum_{n=1}^{\infty} \langle y, u_n \rangle \, u_n. \tag{C.40}$$

Therefore, for all $n \in \mathbb{N}$,

$$\langle y, u_n \rangle = \sigma_n \langle x, v_n \rangle \tag{C.41}$$

holds. Since the operator K is compact, the coefficient of the singular value expansion $\sigma_n \langle x, v_n \rangle$ must be finite and $\langle y, u_n \rangle$ is also finite. Then $\sum_{n=1}^{\infty} \langle y, u_n \rangle^2 / \sigma_n^2 < \infty$ holds.

Conversely, assume that $\sum_{n=1}^{\infty} \langle y, u_n \rangle^2 / \sigma_n^2 < \infty$ holds. From the Riesz-Fischer Theorem[65][63], we have

$$x \equiv \sum_{n=1}^{\infty} \frac{\langle y, u_n \rangle}{\sigma_n} v_n \in \mathcal{X}. \tag{C.42}$$

Then we have

$$Kx = \sum_{n=1}^{\infty} \langle y, u_n \rangle \, v_n = Qy, \tag{C.43}$$

where $Qy \in \mathcal{R}(K)$ and hence $y \in \mathcal{D}(K^\dagger)$. $\qquad \square$

Theorem C.2.12. *Let (σ_n, v_n, u_n) be a singular system for the compact linear operator K and $y \in \mathcal{Y}$ is in a Hilbert space. Then for all $y \in \mathcal{D}(K^\dagger)$, we have*

$$K^\dagger y = \sum_{n=1}^{\infty} \frac{\langle y, u_n \rangle}{\sigma_n} v_n. \tag{C.44}$$

Proof. Since the (v_n) span $\mathcal{R}(K^*) = \mathcal{N}(K)^\perp$ as a analogue of (u_n), then, from Eq.(4.6), $x \in \mathcal{N}(K)^\perp = R(K^*)$. From Eq.(4.14) and Eq.(4.15),

$$\{x \in \mathcal{X} | Kz = Qy\} = K^\dagger y + \mathcal{N}(K). \tag{C.45}$$

Since x lies in $\mathcal{N}(K)^\perp$, we have the desired result. $\qquad \square$

C.3 Spectral theory

We have shown that using a self-adjoint operator in a Hilbert space is critical since it can be represented by a singular system. Next, we discuss that how does such a self-adjoint operator and a singular system respond when we apply them to a continuous function.

First, we define the notations of the orthogonal projectors:

Definition C.3.1. A spectral family
A family (E_λ) of orthogonal projectors in $\mathcal{X}$ is called a *spectral family* if it satisfies the following conditions:

1. $E_\lambda E_\mu = E_{\min\{\lambda,\mu\}}, \quad \lambda, \mu \in \mathbb{R}$;

2. $E_{-\infty} = 0, E_{+\infty} = I$;

3. $E_{\lambda-0} = E_\lambda$, where $E_{\lambda-0}x = \lim_{\epsilon \to 0^+} E_{\lambda-\epsilon}$.

Projection operators in a Hilbert space have useful properties as a following:

Theorem C.3.2. *A bounded operator P is a projection iff it is idempotent and self-adjoint $P = P^* = P^2$.*

Proof. Let P be a projection operator on a subspace S in a Hilbert space H, $P^\perp$ be a projection operator on a subspace $S^\perp$ in a Hilbert space and x, y be arbitrary vectors. Since $Px \in S$, then we have

$$P^2 x = P(Px) = Px. \tag{C.46}$$

And since $\langle Pf, P^\perp g \rangle = \langle P^\perp f, Pg \rangle = 0$, then

$$\langle Px, y \rangle = \langle Px, (P + P^\perp)y \rangle = \langle Px, Py \rangle = \langle (P + P^\perp)x, Py \rangle = \langle x, Py \rangle. \tag{C.47}$$

which proves

$$P = P^* = P^2. \tag{C.48}$$

Conversely, suppose Eq.(C.48). Define

$$S = \{x \in H | Px = x\}. \tag{C.49}$$

Since for any sequence $x_n \subset S$ converses to a vector $x = \lim_{n \to \infty} x_n \in \overline{S}$, then

$$Px = \lim_{n \to \infty} Px_n = \lim_{n \to \infty} x_n = x \in \overline{S}. \tag{C.50}$$

so P is bounded operator. To prove P is the projection onto S, we need to show that $Px \in S$ and $x - Px \in S^\perp$. The first statement is obvious since P is idempotent. The next statement is shown by a following equation:

$$\langle x - Px, z \rangle = \langle x, z \rangle - \langle Px, z \rangle = \langle x, z \rangle - \langle x, Pz \rangle = 0, \tag{C.51}$$

where $z \in S$. $\qquad\qquad\square$

Corollary C.3.3. *If P is a projection operator on a Hilbert space H, then $\langle Px, x \rangle = \|Px\|^2$ for all $x \in H$.*

Proof. From Theorem.C.3.2, we have

$$\langle Px, x \rangle = \langle PPx, x \rangle = \langle Px, Px \rangle = \|Px\|^2. \tag{C.52}$$

$\qquad\qquad\square$

By using these statements can define an integral with respect to a spectral theory.

Theorem C.3.4. A spectral theorem for bounded, self-adjoint operators

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function and A be a bounded, self-adjoint operator. For every $\epsilon > 0$ there exist $\delta > 0$ such that

$$\left\| f(A) - \sum_{i=1}^{n} f(\xi_i)(E_{\lambda_i} - E_{\lambda_{i-1}}) \right\| \leq \epsilon I, \tag{C.53}$$

where

$$\lambda_0 \quad < \quad \alpha_1 = \lambda_1 < \cdots < \lambda_{n-1} < \lambda_n = \alpha_2, \tag{C.54}$$
$$\lambda_i - \lambda_{i-1} \quad \leq \quad \delta \quad \text{for} \quad 1 \leq i \leq n, \tag{C.55}$$
$$\xi_i \quad \in \quad (\lambda_{i-1}, \lambda_i] \quad \text{for} \quad 1 \leq i \leq n. \tag{C.56}$$

In other words, we have a Riemann-Stieltjes type integral of a spectral family:

$$f(A) = \int_{\alpha_1 - 0}^{\alpha_2} f(\lambda) dE_\lambda. \tag{C.57}$$

Proof. From the statement, for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|f(\xi) - f(\eta)| \leq \epsilon, \tag{C.58}$$

where $\xi, \eta \in [\alpha_1 - \theta : \alpha_2]$, $\theta > 0$ and $|\xi - \eta| \leq \delta$. We split the interval $[\alpha_1 - \theta : \alpha_2]$ such that

$$\alpha_1 - \theta = \lambda_0 < \lambda_1 < \cdots < \lambda_{n-1} < \lambda_n = \alpha_2 \tag{C.59}$$

, choose any subdivision

$$\max\{\lambda_i - \lambda_{i-1} | 1 \leq i \leq n\} \leq \delta \tag{C.60}$$

and choose any real number $\xi'_i \in [\lambda_i - \lambda_{i-1}]$. We obtain

$$|f(\xi) - f(\xi'_i)| \leq \epsilon \quad \text{for} \quad \xi \in [\lambda_{i-1}, \lambda_i], \ 1 \leq i \leq n, \tag{C.61}$$

$$f(\xi'_i) \left[f_{\lambda_i}(\xi) - f_{\lambda_{i-1}}(\xi) \right] = \begin{cases} f(\xi'_i) & \text{for} \quad \xi \in (\lambda_{i-1}, \lambda_i], \\ 0 & \text{for} \quad \xi \notin (\lambda_{i-1}, \lambda_i], \end{cases} \tag{C.62}$$

where $f(x)_k$ is a step function defined by

$$f_k(x) = \begin{cases} 1 & \text{for} \quad -\infty < x \leq k, \\ 0 & \text{for} \quad k < x < \infty. \end{cases} \tag{C.63}$$

From Eq.(C.61) to Eq.(C.63), we have

$$\left| f(\xi) - \sum_{i=1}^{n} f(\xi'_i) \left[f_{\lambda_i}(\xi) - f_{\lambda_{i-1}}(\xi) \right] \right| \leq \epsilon \quad \text{for} \quad \xi \in (\alpha_1 - \theta, \alpha_2]. \tag{C.64}$$

The difference of the step function of a bounded operator A such that $f_k(A) - f_l(A)$ is a projection operator since it is idempotent and self-adjoint (see Lemma.C.3.3). Therefore Eq.(C.64) for a bounded operator A can be denoted by

$$\left\| f(A) - \sum_{i=1}^{n} f(\xi'_i) \left(E_{\lambda_i} - E_{\lambda_{i-1}} \right) \right\| \leq \epsilon I. \tag{C.65}$$

$\square$

The finite range of the integral can be extended to the infinity one as a following definition:

Definition C.3.5. For any given $x \in \mathcal{X}$ and any continuous function f on $\mathbb{R}$, the integral

$$\int_{-\infty}^{\infty} f(\lambda) dE_\lambda x \tag{C.66}$$

is defined as the limit in $\mathcal{X}$, if it exists, of $\int_{\alpha_1 - 0}^{\alpha_2} f(\lambda) dE_\lambda x$ when $\alpha_2 \to \infty$ and $\alpha_1 \to -\infty$.

The condition that there is able to exist Eq.(C.66) can be shown by the next theorem.

Theorem C.3.6. *For $x \in \mathcal{X}$ and a continuous function f on $\mathbb{R}$, the following conditions are equivalent:*

$$\int_{-\infty}^{\infty} f(\lambda) dE_\lambda x \quad \text{exists}, \tag{C.67}$$

$$\int_{-\infty}^{\infty} f^2(\lambda) d\|E_\lambda x\|^2 < +\infty. \tag{C.68}$$

Proof. Eq.(C.67)$\Rightarrow$ Eq.(C.68):

Let $y \in \mathcal{Y}$ and $\mathcal{Y}$ be in a Hilbert space. Since the spectral family E_λ is an orthogonal projection operator and fulfills Eq.(C.48), the product of y with Eq.(C.67) is as following form:

$$\int_{-\infty}^{\infty} f(\lambda) d \langle y, E_\lambda x \rangle = \int_{-\infty}^{\infty} f(\lambda) d \langle E_\lambda y, x \rangle . \tag{C.69}$$

Here, we define

$$y \mapsto F(y) \equiv \int_{-\infty}^{\infty} f(\lambda) d \langle E_\lambda y, x \rangle , \tag{C.70}$$

where $y \in \mathcal{Y}$ is on a Hilbert space. For each y, we have

$$\|F(y)\| \leq \| \int_{-\infty}^{\infty} f(\lambda) dE_\lambda x \| \|y\|. \tag{C.71}$$

From the assumption, there exist $\int_{-\infty}^{\infty} f(\lambda) dE_\lambda x$ finitely. Therefore, we have the constant value M such that

$$\|F(y)\| \leq M \|y\|. \tag{C.72}$$

In virtue of the Uniform Boundedness principle[54], $F(y)$ is uniformly bounded and then this is a continuous linear form[67].

Suppose that $y = \int_\alpha^\beta f^*(\lambda) dE_\lambda x$. By definition of the spectral family,

$$
\begin{aligned}
(E_\beta - E_\alpha) y &= \sum_{i=1}^{n} f(\xi_i')(E_{\lambda_i} - E_{\lambda_{i-1}}) - \sum_{i=1}^{n} f(\xi_i')(E_\alpha - E_\alpha), \\
&= y.
\end{aligned}
\tag{C.73}
$$

By substituting Eq.(C.73) into the conjugate, we have

$$
\begin{aligned}
F^*(y) &= \lim_{\alpha' \to -\infty} \lim_{\beta' \to \infty} \int_{\alpha'}^{\beta'} f^*(\lambda)d\langle E_\lambda x, y\rangle, \\
&= \lim_{\alpha' \to -\infty} \lim_{\beta' \to \infty} \int_{\alpha'}^{\beta'} f^*(\lambda)d\langle E_\lambda x, (E_\beta - E_\alpha)\,y\rangle, \\
&= \lim_{\alpha' \to -\infty} \lim_{\beta' \to \infty} \int_{\alpha'}^{\beta'} f^*(\lambda)d\langle (E_\beta - E_\alpha)\,E_\lambda x, y\rangle, \\
&= \int_{\alpha}^{\beta} f^*(\lambda)d\langle E_\lambda x, y\rangle = \|y\|^2.
\end{aligned}
\tag{C.74}
$$

Since $F(y)$ is a linear form, we have $\|y\|^2 \le \|F\|\|y\|$[65] and

$$
\|y\| \le \|F\| < +\infty.
\tag{C.75}
$$

Furthermore,

$$
\|y\|^2 = \left| \int_{\alpha}^{\beta} f^*(\lambda)dE_\lambda x \right| = \int_{\alpha}^{\beta} |f(\lambda)|\,d\|E_\lambda x\|^2,
\tag{C.76}
$$

then we have

$$
\int_{\alpha}^{\beta} |f(\lambda)|\,d\|E_\lambda x\|^2 \le \|F\|^2 < +\infty.
\tag{C.77}
$$

Letting $\alpha \to -\infty$ and $\beta \to \infty$, we have the desired result.

Eq.(C.68)$\Rightarrow$Eq.(C.67): For $\alpha' < \alpha < \beta < \beta'$, we have

$$
\left\| \int_{\alpha'}^{\beta'} f^*(\lambda)dE_\lambda x - \int_{\alpha}^{\beta} f^*(\lambda)dE_\lambda x \right\|^2 = \int_{\alpha}^{\alpha'} |f(\lambda)|\,d\|E_\lambda x\|^2 + \int_{\beta}^{\beta'} |f(\lambda)|\,d\|E_\lambda x\|^2,
\tag{C.78}
$$

then we have the desired result. $\qquad\square$

From these theorems and definitions, we can define the functional self-adjoint operator for the spectral theory represented by integral with an infinite range:

Definition C.3.7. Spectral theory represented by integral

Let A be a self-adjoint operator in the subspace of a Hilbert space $\mathbb{X}$ with spectral family $\{E_\lambda\}$ and let f be piecewise continuous functions for all $x \in \mathcal{X}$. Then $f(A)$ is the operator defined by

$$
f(A)x = \int_{-\infty}^{\infty} f(\lambda)dE_\lambda x, \quad x \in \mathcal{D}(f(A)),
\tag{C.79}
$$

where

$$\mathcal{D}(f(A)) = \left\{ x \in \mathcal{X} \mid \int_{-\infty}^{\infty} f^2(\lambda) d\|E_\lambda x\|^2 < +\infty \right\}. \qquad \text{(C.80)}$$

To compare the equation between Eq.(C.32) and Eq.(C.67), we conclude that

$$\int_{-\infty}^{\infty} f(\lambda) dE_\lambda x = \sum_{n=1}^{\infty} f(\sigma_n^2)\langle x, v_n \rangle v_n, \qquad \text{(C.81)}$$

$$\int_{-\infty}^{\infty} f(\lambda) d\langle E_\lambda x, y \rangle = \sum_{n=1}^{\infty} f(\sigma_n^2)\langle x, v_n \rangle \langle y, v_n \rangle, \qquad \text{(C.82)}$$

$$\int_{-\infty}^{\infty} f(\lambda) d\|E_\lambda x\|^2 = \sum_{n=1}^{\infty} f(\sigma_n^2)|\langle x, v_n \rangle|^2 \qquad \text{(C.83)}$$

for a compact, self-adjoint operator.

Here, we introduce some useful applications of the spectral theory.

Theorem C.3.8. *Let $\mathcal{X}, \mathcal{Y}$ be subspaces of a Hilbert space, $K : \mathcal{X} \to \mathcal{Y}$ be an adjoint linear operator on Hilbert space, $K^* : \mathcal{Y} \to \mathcal{X}$ be an adjoint of the operator K, and f be the continuous function for operator defined by Definition.C.3.7. Then we have*

$$f(K^*K)K^* = K^*f(KK^*). \qquad \text{(C.84)}$$

Proof. We define the polynomial of the self-adjoint operator

$$p(K^*K) \equiv \sum_{k=0}^{n} a_k (K^*K)^k. \qquad \text{(C.85)}$$

Since the operator K is adjoint, then we have

$$K^*(KK^*)^k = (K^*(KK^*)^k)^* = (KK^*)^k K. \qquad \text{(C.86)}$$

for all $k \in \mathbb{N}$. From Eq.(C.86) and Eq.(C.85), we obtain

$$K^*p(KK^*) = \sum_{k=0}^{n} a_k K^*(KK^*)^k = \sum_{k=0}^{n} a_k (KK^*)^k K = p(KK^*)K. \qquad \text{(C.87)}$$

Since a polynomial function is continuous at every point, we have the desired result. $\qquad \square$

Acknowledgements

It is a great pleasure to thank all the people who made this thesis possible.

First, I would like to thank Professor Masatake Ohashi. He is not only my reliable supervisor, but also supporting whole of my life during the graduate school. In particular, I'm grateful to him from the bottom of my heart that he supported me with great patience when I couldn't continue to research owing to my illness. Without his effort to support, this thesis would not have been possible.

I also would like to thank Associate Professor Kazuhiro Hayama. He introduced me the concept of the thesis -a regularization method- and advised a whole of the study.

And I also would like to thank Associate Professor Yousuke Itoh. He has taught me the fundamental and applicational theory of general relativity when I was undergraduate school student in Tohoku University. His passion of the GW brings me to finish to write the thesis.

Finally, I would like to thank all of my family for their limitless live, patience and support.

Bibliography

[1] B. P. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, G. Allen, R. S. Amin, S. B. Anderson, W. G. Anderson, M. A. Arain, and et al. Ligo: the laser interferometer gravitational-wave observatory. *Reports on Progress in Physics*, 72(7):076901, 2009.

[2] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin, and et al. Advanced virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, 2015.

[3] Yoichi Aso, Yuta Michimura, Kentaro Somiya, Masaki Ando, Osamu Miyakawa, Takanori Sekiguchi, Daisuke Tatsumi, and Hiroaki Yamamoto. Interferometer design of the kagra gravitational wave detector. *Phys. Rev. D*, 88:043007, Aug 2013.

[4] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.

[5] B. P. Abbott and et. al. Gw170817: Observation of gravitational waves from a binary neutron star inspiral. *Phys. Rev. Lett.*, 119:161101, Oct 2017.

[6] The LIGO Scientific Collaboration and the Virgo Collaboration. Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs, 2018.

[7] A. Goldstein, P. Veres, E. Burns, M. S. Briggs, R. Hamburg, D. Kocevski, C. A. Wilson-Hodge, R. D. Preece, S. Poolakkil, O. J. Roberts, C. M. Hui, V. Connaughton, J. Racusin, A. von Kienlin, T. Dal Canton, N. Christensen, T. Littenberg, K. Siellez, L. Blackburn, J. Broida,

E. Bissaldi, W. H. Cleveland, M. H. Gibby, M. M. Giles, R. M. Kippen, S. McBreen, J. McEnery, C. A. Meegan, W. S. Paciesas, and M. Stanbro. An ordinary short gamma-ray burst with extraordinary implications: Fermi -gbm detection of grb 170817a. *The Astrophysical Journal Letters*, 848(2):L14, 2017.

[8] V. Savchenko, C. Ferrigno, E. Kuulkers, A. Bazzano, E. Bozzo, S. Brandt, J. Chenevez, T. J.-L. Courvoisier, R. Diehl, A. Domingo, L. Hanlon, E. Jourdain, A. von Kienlin, P. Laurent, F. Lebrun, A. Lutovinov, A. Martin-Carrillo, S. Mereghetti, L. Natalucci, J. Rodi, J.-P. Roques, R. Sunyaev, and P. Ubertini. Integral detection of the first prompt gamma-ray signal coincident with the gravitational-wave event gw170817. *The Astrophysical Journal Letters*, 848(2):L15, 2017.

[9] B. P. Abbott and et. al. Multi-messenger observations of a binary neutron star merger. *The Astrophysical Journal Letters*, 848(2):L12, 2017.

[10] Riccardo Ciolfi. X-ray flashes powered by the spindown of long-lived neutron stars. *The Astrophysical Journal*, 829(2):72, 2016.

[11] Jimmy A Irwin, W Peter Maksym, Gregory R Sivakoff, Aaron J Romanowsky, Dacheng Lin, Tyler Speegle, Ian Prado, David Mildebrath, Jay Strader, Jifeng Liu, and Jon M Miller. Ultraluminous X-ray bursts in two ultracompact companions to nearby elliptical galaxies. *Nature*, 538:356, oct 2016.

[12] Franz E. Bauer, Ezequiel Treister, Kevin Schawinski, Steve Schulze, Bin Luo, David M. Alexander, William N. Brandt, Andrea Comastri, Francisco Forster, Roberto Gilli, David Alexander Kann, Keiichi Maeda, Ken'ichi Nomoto, Maurizio Paolillo, Piero Ranalli, Donald P. Schneider, Ohad Shemmer, Masaomi Tanaka, Alexey Tolstov, Nozomu Tominaga, Paolo Tozzi, Cristian Vignali, Junxian Wang, Yongquan Xue, and Guang Yang. A new, faint population of x-ray transients. *Monthly Notices of the Royal Astronomical Society*, 467(4):4841–4857, 2017.

[13] Tatsuya Matsumoto and Shigeo S. Kimura. Delayed jet breakouts from binary neutron star mergers. *The Astrophysical Journal Letters*, 866(2):L16, 2018.

[14] Akihiro Suzuki, Keiichi Maeda, and Toshikazu Shigeyama. Relativistic supernova ejecta colliding with a circumstellar medium: An application to the low-luminosity grb 171205a. *The Astrophysical Journal*, 870(1):38, 2019.

[15] Bernard F. Schutz. Determining the Hubble constant from gravitational wave observations. *Nature*, 323(6086):310–311, sep 1986.

[16] Daniel E. Holz and Scott A. Hughes. Using gravitational-wave standard sirens. *The Astrophysical Journal*, 629(1):15, 2005.

[17] Samaya Nissanke, Daniel E. Holz, Neal Dalal, Scott A. Hughes, Jonathan L. Sievers, and Christopher M. Hirata. Determining the hubble constant from gravitational wave observations of merging compact binaries, 2013.

[18] Walter Del Pozzo. Inference of cosmological parameters from gravitational waves: Applications to second generation interferometers. *Phys. Rev. D*, 86:043011, Aug 2012.

[19] The LIGO Scientific Collaboration Collaboration, The Virgo, The 1M2H Collaboration, The Dark Energy Camera GW-EM Collaboration Collaboration, the D E S, The DLT40 Collaboration, The Las Cumbres Observatory Collaboration, The VINROUGE Collaboration, and The MASTER Collaboration. A gravitational-wave standard siren measurement of the Hubble constant. *Nature*, 551(7678):85–88, nov 2017.

[20] Buchner, J., Georgakakis, A., Nandra, K., Hsu, L., Rangel, C., Brightman, M., Merloni, A., Salvato, M., Donley, J., and Kocevski, D. X-ray spectral modelling of the agn obscuring region in the cdfs: Bayesian model selection and catalogue. *A& A*, 564:A125, 2014.

[21] Adam G. Riess, Lucas M. Macri, Samantha L. Hoffmann, Dan Scolnic, Stefano Casertano, Alexei V. Filippenko, Brad E. Tucker, Mark J. Reid, David O. Jones, Jeffrey M. Silverman, Ryan Chornock, Peter Challis, Wenlong Yuan, Peter J. Brown, and Ryan J. Foley. A 2.4% determination of the local value of the hubble constant. *The Astrophysical Journal*, 826(1):56, 2016.

[22] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, T. Adams, and et al. Parameter estimation for compact binary coalescence signals with the first generation gravitational-wave detector network. *Phys. Rev. D*, 88:062001, Sep 2013.

[23] I. W. Harry and S. Fairhurst. Targeted coherent search for gravitational waves from compact binary coalescences. *Phys. Rev. D*, 83:084002, Apr 2011.

[24] M Rakhmanov. Rank deficiency and tikhonov regularization in the inverse problem for gravitational-wave bursts. *Classical and Quantum Gravity*, 23(19):S673, 2006.

[25] S D Mohanty, M Rakhmanov, S Klimenko, and G Mitselmakher. Variability of signal-to-noise ratio and the network analysis of gravitational wave burst signals. *Classical and Quantum Gravity*, 23(15):4799, 2006.

[26] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Springer Netherlands, 1996.

[27] F. Feroz, M. P. Hobson, and M. Bridges. MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. *ArXiv e-prints*, 398:1601–1614, oct 2009.

[28] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt. Importance Nested Sampling and the MultiNest Algorithm. *ArXiv e-prints*, jun 2013.

[29] M. Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments*. Gravitational Waves. OUP Oxford, 2007.

[30] J. D. E. Creighton and W. G. Anderson. *Gravitational Waves, in Gravitational-Wave Physics and Astronomy: An Introduction to Theory, Experiment and Data Analysis*. Wiley-VCH, 2011.

[31] K. S. Thorne. Multipole expansions of gravitational radiation. *Reviews of Modern Physics*, 52:299–340, April 1980.

[32] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields, Course of Theoretical Physics Volume 2*. 1971.

[33] Konstantin Postnov and Lev Yungelson. The Evolution of Compact Binary Star Systems. *Living Reviews in Relativity*, 17(1):3, jan 2007.

[34] J.D.E. Creighton and W.G. Anderson. *Gravitational-Wave Physics and Astronomy: An Introduction to Theory, Experiment and Data Analysis*. Wiley Series in Cosmology. Wiley, 2012.

[35] Piotr Jaranowski, Andrzej Królak, and Bernard F. Schutz. Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection. *Phys. Rev. D*, 58:063001, Aug 1998.

[36] Neil J Cornish and Edward K Porter. The search for massive black hole binaries with lisa. *Classical and Quantum Gravity*, 24(23):5729, 2007.

[37] PETER J. GREEN. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[38] D. Gamerman. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference.* Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 1997.

[39] John Skilling et al. Nested sampling for general bayesian computation. *Bayesian Analysis*, 1(4):833–859, 2006.

[40] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. *Bayesian Data Analysis, Third Edition.* Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis, 2013.

[41] 繁桝算男. ベイズ統計入門. 東京大学出版会, 1985.

[42] 松原望. 入門ベイズ統計: 意思決定の理論と発展. 東京図書, 2008.

[43] J. Veitch and A. Vecchio. Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network. *Phys. Rev. D*, 81:062003, Mar 2010.

[44] C. W. Helstrom. *Statistical Theory of Signal Detection*, volume 9 of *International Series of Monographs in Electronics and Instrumentation.* Pergamon Press, Oxford; New York, 2nd edition, 1968.

[45] B.S. Sathyaprakash and Bernard F. Schutz. Physics, astrophysics and cosmology with gravitational waves. *Living Reviews in Relativity*, 12(2), 2009.

[46] F. Feroz and M. P. Hobson. Multimodal nested sampling: an efficient and robust alternative to markov chain monte carlo methods for astronomical data analyses. *Monthly Notices of the Royal Astronomical Society*, 384(2):449–463, 2008.

[47] J. R. Shaw, M. Bridges, and M. P. Hobson. Efficient bayesian inference for multimodal problems in cosmology. *Monthly Notices of the Royal Astronomical Society*, 378(4):1365–1370, 2007.

[48] Pia Mukherjee, David Parkinson, and Andrew R. Liddle. A nested sampling algorithm for cosmological model selection. *The Astrophysical Journal Letters*, 638(2):L51, 2006.

[49] V.G. Maz'ja and T.O. Shaposhnikova. *Jacques Hadamard: A Universal Mathematician*. History of mathematics. American Mathematical Society, 1999.

[50] Guorong Wang, Yimin Wei, and Sanzheng Qiao. *Generalized inverses: theory and computations*. Springer, 2018.

[51] Y. Wang, A.G. Yagola, and C. Yang. *Optimization and Regularization for Computational Inverse Problems and Applications*. Springer Berlin Heidelberg, 2011.

[52] Eberhard Schock. Approximate Solution of Ill-Posed Equations: Arbitrarily Slow Convergence vs. Superconvergence BT - Constructive Methods for the Practical Treatment of Integral Equations: Proceedings of the Conference Mathematisches Forschunginstitut Oberwolfach, June . pages 234–243. Birkhäuser Basel, Basel, 1985.

[53] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. Dover Publications, 1964.

[54] M. Haase. *Functional Analysis: An Elementary Introduction*. Graduate Studies in Mathematics. American Mathematical Society, 2014.

[55] James Raymond Munkres. *Topology second edition*. Pearson, 2000.

[56] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[57] The LIGO Scientific Collaboration. Advanced ligo. *Classical and Quantum Gravity*, 32(7):074001, 2015.

[58] B. Allen. Gravitational Wave Detector Sites. *ArXiv General Relativity and Quantum Cosmology e-prints*, July 1996.

[59] Samantha A. Usman, Joseph C. Mills, and Stephen Fairhurst. Constraining the inclination of binary mergers from gravitational-wave observations, 2018.

[60] Piotr Jaranowski and Andrzej Krolak. Optimal solution to the inverse problem for the gravitational wave signal of a coalescing compact binary. *Phys. Rev. D*, 49:1723–1739, Feb 1994.

[61] Eric Poisson and Clifford M. Will. *Gravity: Newtonian, Post-Newtonian, Relativistic*. Cambridge University Press, 2014.

[62] L. Debnath and P. Mikusinski. *Introduction to Hilbert Spaces with Applications*. Elsevier Science, 2005.

[63] Harkrishan Lal Vasudeva. *Elements of Hilbert Spaces and Operator Theory*. Springer Singapore, 2017.

[64] R. Dautray, M. Artola, J.C. Amson, M. Cessenat, and J.L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 3 Spectral Theory and Applications*. Mathematical Analysis and Numerical Methods for Science and Technology. Springer Berlin Heidelberg, 1999.

[65] E. Hewitt and K. Stromberg. *Real and Abstract Analysis*. Springer, 1965.

[66] G. Helmberg, H.A. Lauwerier, and W.T. Koiter. *Introduction to Spectral Theory in Hilbert Space*. North-Holland Series in Applied Mathematics and Mechanics. Elsevier Science, 2014.

[67] R. Dautray, M. Artola, J.C. Amson, M. Cessenat, and J.L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 2 Functional and Variational Methods*. Mathematical Analysis and Numerical Methods for Science and Technology. Springer Berlin Heidelberg, 2000.