Doctoral Thesis

New method of all-sky searches for continuous gravitational waves

Takahiro S. Yamamoto

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

April 20, 2021

Abstract

Continuous gravitational waves have a longer duration than the observation period, an almost monochromatic frequency, and a constant amplitude. In the theory of gravitational waves, various sources (e.g., rotating neutron stars and axion clouds around black holes) are expected to emit continuous gravitational waves. After the three observation runs of the advanced LIGO and the advanced Virgo, continuous gravitational waves have not been detected. In the data analysis context, the detection of continuous gravitational waves is a still challenging task. The coherent matched filtering is the optimal detection method because the continuous gravitational waves can be modeled by relatively simple waveforms. The waveforms depend on the source location through the Doppler modulation caused by the detector's motion associated with the rotation and the orbital motion of the Earth. This Doppler modulation increases the number of templates, and, as a result, the computational cost of the coherent matched filtering becomes desperately expensive. On the other hand, semi-coherent methods employed for the real data analysis by the LIGO/Virgo collaboration are computationally efficient. However, their performances are significantly degraded from that of the coherent matched filtering. We proposed a new method by combining the excess power method and the deep learning method, which are computationally cheap. The workflow is as follows: 1) preprocessing the data strain, 2) candidate selection by the excess power method, 3) predicting the source location by the deep learning method, 4) the follow-up analysis by the coherent matched filtering. We assess the ability of our method with the assumptions that the sources are monochromatic, the extrinsic parameters are fixed to be $(\psi, \cos\iota, \phi_0) = (0, 1, 0)$ and a priori known, the 100% duty cycle single detector, and the detector noise is stationary and Gaussian. We show that our method is worthy for further study for analyzing the signal with the duration of $O(10^7)$sec.

Contents

Abstract **1**

1 Introduction **5**
1.1 Gravitational wave astronomy 5
1.2 Continuous gravitational waves and difficulty of its detection 6
1.3 Brief summary of the thesis 7
1.4 Notation 7

2 Gravitational waves **9**
2.1 Linearized Einstein equation 9
2.1.1 Linearized Einstein equation 9
2.1.2 Gauge transformation 10
2.2 Propagation of gravitational waves 10
2.2.1 Transverse-traceless gauge 10
2.2.2 Monochromatic, plane wave 11
2.3 Generation of gravitational waves 13
2.3.1 Quadrupole formula 13
2.3.2 Energy-momentum tensor of gravitational waves 15
2.3.3 Example 1: Rotating triaxial ellipsoid 16
2.3.4 Example 2: Compact binary inspiral 18
2.4 Continuous gravitational waves 20
2.4.1 Rotating neutron stars 20
2.4.2 Axion clouds around black holes 22
2.4.3 Sub-solar mass black hole binaries 23
2.5 Gravitational wave detectors 23
2.5.1 Michelson interferometer 23

3 Matched filtering and continuous gravitational waves **27**
3.1 Matched filtering 27
3.1.1 Statistical properties of detector noise 27
3.1.2 Matched filter as optimal statistic 28
3.1.3 Matched filtering in time domain 31
3.2 Continuous gravitational waves 31

3.2.1 Waveform model . . . 31
3.2.2 Time resampling technique . . . 33
3.2.3 Difficulty in all-sky search of continuous gravitational waves . . . 33

4 New method of all-sky searches for continuous gravitational waves 35
4.1 Overview . . . 35
4.2 Subtracting the effect due to the Earth's rotation . . . 35
4.3 Double Fourier transform . . . 41
4.4 Excess power method . . . 46
4.5 Deep learning . . . 47
4.5.1 Fundamentals . . . 47
4.5.2 Application to gravitational wave data . . . 50
4.5.3 ANNs in our work . . . 51
4.6 Follow-up analysis by the matched filtering . . . 53
4.7 Results . . . 54
4.7.1 Computational cost . . . 54
4.7.2 False alarm probability . . . 56
4.7.3 Detection probability . . . 57
4.7.4 Applicability of our method . . . 59

5 Summary and future work 63
5.1 Summary . . . 63
5.2 Future work . . . 64

Acknowledgement 65

A Derivation of antenna pattern function 67
A.1 Detected strain . . . 67
A.2 From source frame to wave frame . . . 68
A.3 Celestial sphere frame . . . 69
A.4 Detector frame . . . 70
A.5 The antenna pattern function . . . 73

Chapter 1

Introduction

1.1 Gravitational wave astronomy

From the first direct detection of GW150914 [1], the advanced Laser Interferometer Gravitational-wave Observatory (aLIGO) and the advanced Virgo (AdV) have detected about fifty events which are confirmed as the signal originating from compact binary mergers [2, 3]. These events have given us tremendous information about astrophysics, cosmology, and theories of gravity. GW150914 is the coalescence of two orbiting black holes whose masses are $36M_\odot$ and $29M_\odot$, respectively. The formation scenario of such a stellar mass black hole binary is under discussion [4]. Also, by using detected waveform, various types of tests of general relativity have been done [5]. So far, all results are consistent with the predictions of general relativity. Other exotic phenomena like black hole echoes also are explored. GW170817 is the first gravitational wave event from a binary neutron star merger [6]. It is crucial to observe the same source by gravitational waves and electromagnetic waves [7, 8]. As a result, the unified picture connecting the binary neutron star merger, a short gamma-ray burst, and a kilonova is established. Comparing the arrival times of the gamma-ray photons and the gravitational waves, we obtain a strict limit of the propagation speed of gravitational waves [8]. By identifying the host galaxy, the Hubble parameters can be estimated [9]. Recently, the LIGO and Virgo collaboration reports a new catalog of the gravitational wave events [2, 3]. With this catalog, the population of black holes is started to be discussed statistically [10]. In near future, new ground-based gravitational wave detectors like KAGRA [11, 12] and LIGO India [13] will join the gravitational wave detector network [14]. The observation by the multiple detectors is useful to improve the precision of estimation of source location and to test the polarization of gravitational waves. As a future plan, Einstein Telescope [15, 16] and Cosmic Explorer [17] are proposed.

The ground-based detectors are sensitive to gravitational waves of the frequency band from 10 to 10^3Hz. Besides, the space-borne detectors, which are proposed as future observation plans, would be sensitive to a frequency band from 1 to 100 mHz. The Laser Interferometer Space Antenna (LISA) [18, 19] is planned to be launched by ESA in the early 2030s. LISA consists of three satellites. They are placed at the vertex of a triangle

of the arm length 2,500,000 km and their center of mass orbits around the Sun. Similar projects named Taiji [20] and TianQin [21] are proposed by China. The correlation analysis with LISA and Taiji is also investigated actively (e.g., [22]). These detectors' targets are mergers of supermassive black hole binaries, extreme mass ratio inspirals (EMRIs), the stochastic signals generated in the Early universe, and so on. A Japanese research group has proposed a project named B-DECIGO [23]. Similarly to LISA, it consists of three satellites and form a triangle. But, the size of the triangle is 1,000 km and the satellites will orbit the Earth. B-DECIGO has sensitivity at a frequency band from 100 mHz to 10Hz.

The pulsar timing array (PTA) [24] uses pulsars as accurate clocks and taking correlations between the shifts of the pulse arrival timing of different pulsars. PTA is sensitive to gravitational waves whose frequency is few nHz. Recently, the NANOGrav collaboration reports that the stochastic signal which is not consistent with noise is found [25]. Although this stochastic signal is not confirmed as an astrophysical event, further investigation will elucidate the astrophysics behind the signal.

Gravitational waves become an essential tool in astronomy. The gravitational wave astronomy is believed to open the window to understanding our Universe.

1.2 Continuous gravitational waves and difficulty of its detection

All events detected by the ground-based interferometers so far are the signals originating from compact binary mergers. The signals have short durations and rapidly increasing frequencies and are called chirp signals. On the other hand, various astrophysical and cosmological origins are theoretically expected to be the sources of gravitational waves, and, correspondingly, the waveforms of the expected signals are varied (Ref. [26] as review). For example, rotating deformed neutron stars are possible sources of gravitational waves (Ref. [27] as review). Rotating neutron stars are already discovered as pulsars by radio astronomy [28]. Their rotation frequencies and spin-down rates are precisely estimated by the arrival times of pulse signals. If neutron stars are deformed to break the rotational symmetry, they emit gravitational waves of frequencies twice of the rotation frequency. The wave emission is expected to last longer than the observation period. Such kind of signal is called continuous gravitational waves. They are characterized by 1) longer duration than the observation period, 2) small change rate of frequencies, 3) almost constant amplitudes. Not only rotating neutron stars but also some exotic sources are expected to radiate continuous gravitational waves.

In the context of data analysis, the detection of continuous gravitational waves is still a challenging research area [29]. The difficulty in data analysis of continuous gravitational waves arises from the Doppler modulation due to the detector's motion. Even if the source is monochromatic, the observed signal is modulated because of the relative motion between the detector and the sources. The detector's motion is modeled with the assumption that

it is associated with the rotation and the orbital motion of the Earth. Therefore, the matched filtering, which is the optimal method when we can model the waveforms theoretically, can be applied to detect continuous gravitational waves. The waveform models are characterized not only by the source properties (e.g., gravitational wave frequency at the reference time) but also by the source location on the Celestial sphere. Because the signal strength (signal-to-noise ratio) is susceptible to the source location, we need many templates to keep the maximum loss of the significance of the signal by the mismatch between the source location and the predicted location below an acceptable level. As a result, it takes an unrealistically long computational time to coherently analyze the entire strain data of $O(1)$ year duration. Usually, the strain data is divided into many small segments, and SNRs of these segments are appropriately integrated. This method is called a semi-coherent method. There are various semi-coherent methods, e.g., stack slide [30], Hough transform [31]. They are computationally reasonable methods, but the detection efficiency is significantly reduced compared with the coherent matched filtering.

The semi-coherent analysis is carried out for the strain data obtained during Observation run one (O1) and two (O2) of the collaboration of aLIGO and AdV [32–34]. Despite tremendous efforts to detect continuous gravitational waves, continuous gravitational waves have not been detected so far.

1.3 Brief summary of the thesis

In this thesis, we propose a new method for detecting continuous gravitational waves and show the proof-of-principle of our method using the monochromatic waves. Our method consists of several stages, 1) preprocessing the strain data by using the time resampling technique, the short-time Fourier transform, and another Fourier transform, 2) candidate selection by the excess power method, 3) narrowing down the possible locations of sources by the deep learning method, 4) following-up analysis by the coherent matched filtering.

In Chap. 2, we briefly review the fundamental facts of gravitational waves. Chapter. 3 is devoted to the review of the matched filtering method. We will see the difficulty of the matched filtering in the detection of continuous gravitational waves. In Chap. 4, we explain our new method in detail. It includes a brief review of the deep learning. Comparing with the coherent matched filtering, we also show the performance of our method. Chapter. 5 is devoted to summarizing our work and future works.

This thesis is mainly based on our work Ref. [35].

1.4 Notation

Greek indices $\mu, \nu \ldots$ indicate spacetime coordinates and run over 0, 1, 2, and 3. Latin indices $i, j, \ldots$ indicate sptial coordinates and run over 1,2, and 3. We work with a mostly positive signature. The Cartesian coordinates are

$$x^0 = ct\,,\ x^1 = x\,,\ x^2 = y\,,\ x^3 = z\,. \tag{1.1}$$

The metric of Minkowski spacetime on the Cartesian coordinates is written as

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.2}$$

Einstein's contraction rule is employed. The partial derivative with respect to a coordinate x^μ:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} .$$

The inner product of 4-vectors p and q is denoted by

$$p \cdot q := g_{\mu\nu} p^\mu q^\nu = p_\mu q^\mu .$$

The inner product of 3-vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is denoted by

$$\boldsymbol{a} \cdot \boldsymbol{b} := \delta_{ij} a^i b^j = a_i b^i .$$

The Fourier transform of the function $s(t)$ is defined by

$$\tilde{s}(f) := \int_{-\infty}^{\infty} dt \ s(t) e^{-2\pi i f t} .$$

The inverse Fourier transform is defined by

$$s(t) = \int_{-\infty}^{\infty} df \ \tilde{s}(f) e^{2\pi i f t} .$$

Chapter 2

Gravitational waves

2.1 Linearized Einstein equation

2.1.1 Linearized Einstein equation

In general relativity, the gravitational field is expressed by the metric tensor $g_{\mu\nu}$ which is directly related with the geometry of the spacetime. The line element is written as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \,. \tag{2.1}$$

The dynamics of the gravitational field is governed by the Einstein equation

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu} \,, \tag{2.2}$$

where $R_{\mu\nu}$ is the Ricci tensor, $R := g^{\mu\nu} R_{\mu\nu}$ is the Ricci scalar and $T_{\mu\nu}$ is the energy-momentum tensor of the matter.

Let us consider the metric perturbation $h_{\mu\nu}$ on the flat spacetime $\eta_{\mu\nu}$. The metric tensor is written as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \,. \tag{2.3}$$

In the Cartesian coordinate, the condition,

$$|h_{\mu\nu}| \ll 1 \,, \tag{2.4}$$

is valid for every component of $h_{\mu\nu}$. The meaning of the condition Eq. (2.4) is that there exists the reference frame in which Eq. (2.4) is valid. General relativity has the covariance with respect to the general coordinate transformation and it can be employed to move to such a reference frame.

Expanding the Einstein equation up to the first order of the perturbation $h_{\mu\nu}$, we get the linearized Einstein equation,

$$\Box \psi_{\mu\nu} + \eta_{\mu\nu} \partial^\rho \partial^\sigma \psi_{\rho\sigma} - \partial_\mu \partial^\rho \psi_{\rho\nu} - \partial_\nu \partial^\rho \psi_{\mu\rho} = -\frac{16\pi G}{c^4} T_{\mu\nu} \,. \tag{2.5}$$

Here, $\Box$ is d'Alambertian in the flat spacetime and

$$\psi_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h\,, \tag{2.6}$$

with

$$h := \eta^{\mu\nu}h_{\mu\nu}\,. \tag{2.7}$$

2.1.2 Gauge transformation

By using the degrees of freedom originating from the general covariance, we can move to a reference frame in which Eq. (2.4) holds. However, there is a residual gauge degrees of freedom. Let us consider the following infinitesimal coordinate transformation,

$$x^\mu \to x'^\mu = x^\mu + \xi^\mu(x)\,, \tag{2.8}$$

with

$$|\partial_\mu \xi_\nu| \sim |h_{\mu\nu}| \ll 1\,. \tag{2.9}$$

Under the transformation given in Eq. (2.8), the metric perturbation $h_{\mu\nu}$ is transformed as

$$h_{\mu\nu}(x) \to h'_{\mu\nu}(x') = h_{\mu\nu}(x) - (\partial_\mu \xi_\nu + \partial_\nu \xi_\mu)\,. \tag{2.10}$$

By virtue of this gauge degrees of freedom, we can impose the additional gauge condition, which is called Lorentz gauge,

$$\partial^\nu \psi_{\mu\nu} = 0\,. \tag{2.11}$$

With the Lorentz gauge, the linearized Einstein equation (Eq. (2.5)) becomes

$$\Box\psi_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu}\,. \tag{2.12}$$

This is nothing but the wave equation with a source. Thus, the wave-like function can exists as the solution of the linearlized Einstein equation. This solution is called gravitational waves. Considering that the metric tensor has the geometrical meaning, the gravitational waves can be interpreted as "the ripple of the spacetime".

2.2 Propagation of gravitational waves

2.2.1 Transverse-traceless gauge

Let us consider the gravitational waves propagating in vacuum, hence we set $T_{\mu\nu} = 0$ in Eq. 2.12. The Lorentz gauge (2.11) does not completely fix the degree of freedom. In fact, the Lorentz gauge condition is valid as long as the infinitesimal vector ξ_μ satisfies $\Box\xi_\mu = 0$. Therefore, a vector $\tilde{\xi}_\mu$ that satisfies $\Box\tilde{\xi}_\mu = 0$ can be added to the vector ξ_μ. Additional

conditions can be imposed to $\psi_{\mu\nu}$ by using the residual gauge. In order to figure out the properties of the gravitational waves, we impose the following conditions,

$$\psi := \eta^{\mu\nu}\psi_{\mu\nu} = 0\,, \tag{2.13}$$

$$h^{0i} = 0\,. \tag{2.14}$$

These conditions and the Lorentz gauge conditions are collectively called *transverse-traceless gauge* (TT gauge). The number of the TT gauge conditions is eight and the metric perturbation $h_{\mu\nu}$ has ten components. Therefore, the gravitational waves have two independent degrees of freedom. We refer $h^{\rm TT}_{\mu\nu}$ to the gravitational wave metric in the TT gauge.

For later convenience, let us rewrite the conditions of TT gauge. First of all, the equality $\psi^{\rm TT}_{\mu\nu} = h^{\rm TT}_{\mu\nu}$ holds because of Eq. (2.13). By using Eq. (2.14), the $\mu = 0$ components of the Lorentz gauge can be transformed as

$$0 = \partial^\nu h^{\rm TT}_{0\nu} = \partial^0 h^{\rm TT}_{00} + \partial^j h^{\rm TT}_{0j} = \partial^0 h^{\rm TT}_{00}\,. \tag{2.15}$$

Therefore, $h^{\rm TT}_{00}$ is constant in time. The static part of the perturbation corresponds to the Newton potential generated by the matter. The gravitational waves we pursuing is non-static part of the perturbation. From these discussions, we can set the constant to be zero, i.e.,

$$h^{\rm TT}_{00} = 0\,. \tag{2.16}$$

By substituting Eq. (2.16) to Eqs. (2.13) and (2.11), we obtain

$$\delta^{ij} h^{\rm TT}_{ij} = 0\,, \tag{2.17}$$

and

$$\partial^j h^{\rm TT}_{ij} = 0\,. \tag{2.18}$$

The meanings of Eqs. (2.17) and (2.18) are the traceless condition and the transverse condition. The independent gauge conditions are Eqs. (2.14), (2.16), (2.17) and (2.18).

2.2.2 Monochromatic, plane wave

In the TT gauge, the linearized Einstein equation becomes the wave equation, i.e.,

$$\Box h^{\rm TT}_{ij} = 0\,. \tag{2.19}$$

Let us impose an ansatz representing the monochromatic plane wave, that is,

$$h^{\rm TT}_{ij}(t, \boldsymbol{x}) = A_{ij} e^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}}\,. \tag{2.20}$$

with the constant tensor A_{ij}, the wave vector $\boldsymbol{k}$ and the angular frequency ω. For simplicity, we assume that the gravitational waves propagate along with the z-axis. Then, the wave vector $\boldsymbol{k}$ becomes $\boldsymbol{k} = (0, 0, k_z)$ and

$$h^{\rm TT}_{ij}(t, \boldsymbol{x}) = A_{ij} e^{-i\omega t + ik_z z}\,. \tag{2.21}$$

The TT gauge condition becomes

$$A_{xz} = A_{yz} = A_{zz} = 0\,, \tag{2.22}$$

and

$$A_{xx} + A_{yy} = 0\,. \tag{2.23}$$

Corresponding to the two degrees of freedom, the polarization tensors are defined by

$$e^{+}_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij}\,, \tag{2.24}$$

and

$$e^{\times}_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij}\,. \tag{2.25}$$

The amplitude A_{ij} can be written as the linear combination of two polarizations,

$$A_{ij} = h_{+} e^{+}_{ij} + h_{\times} e^{\times}_{ij}\,, \tag{2.26}$$

and the monochromatic plane wave solution becomes

$$h_{ij} = \begin{pmatrix} h_{+} & h_{\times} & 0 \\ h_{\times} & -h_{+} & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij} \times e^{-i\omega t + i k_z z}\,. \tag{2.27}$$

The similar discussion can be extended to the case of the gravitational waves propagating along with an arbitrary direction. We define the unit vector $\boldsymbol{n}$ by

$$\boldsymbol{n} := \frac{\boldsymbol{k}}{|\boldsymbol{k}|}\,. \tag{2.28}$$

For imposing the TT gauge conditions, it is useful to introduce a projection operator

$$\Lambda_{ij,kl}(\boldsymbol{n}) = \mathcal{P}_{ik}\mathcal{P}_{jl} - \frac{1}{2}\mathcal{P}_{ij}\mathcal{P}_{kl}\,, \tag{2.29}$$

with

$$\mathcal{P}_{ij}(\boldsymbol{n}) = \delta_{ij} - n_i n_j\,. \tag{2.30}$$

The operator $\Lambda_{ij,kl}$ has the following properties.

- $\Lambda_{ij,kl}\Lambda_{kl,mn} = \Lambda_{ij,mn}$,
- transverse, $n^i \Lambda_{ij,kl} = n^j \Lambda_{ij,kl} = 0$,
- traceless, $\Lambda_{ii,kl} = \Lambda_{ij,kk} = 0$,
- symmetric, $\Lambda_{ij,kl} = \Lambda_{kl,ij}$.

Using this projection operator, we can easily move to TT gauge like

$$h^{\mathrm{TT}}_{ij}(t, \boldsymbol{x}) = \Lambda_{ij,kl}(\boldsymbol{n}) h_{kl}(t, \boldsymbol{x})\,. \tag{2.31}$$

2.3 Generation of gravitational waves

2.3.1 Quadrupole formula

In this section, we solve the linearized Einstein equation with the source

$$\Box\psi_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu}\,, \tag{2.32}$$

under the Lorentz gauge

$$\partial_\mu\psi^{\mu\nu} = 0\,. \tag{2.33}$$

Due to the Lorentz gauge condition, the energy-momentum tensor automatically satisfies the conservation law in the flat-space time,

$$\partial_\mu T^{\mu\nu} = 0\,. \tag{2.34}$$

By defining a Green function by

$$\Box_x G(x-x') = \delta^4(x-x')\,, \tag{2.35}$$

the solution of Eq. (2.32) is

$$\psi_{\mu\nu}(x) = -\frac{16\pi G}{c^4}\int d^4x'\; G(x-x')T_{\mu\nu}(x')\,. \tag{2.36}$$

As a solution of Eq. (2.35), we choose the retarded Green function,

$$G(x-x') = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}\delta(x^0_{\rm ret}-x'^0)\,, \tag{2.37}$$

with the retarded time defined by

$$t_{\rm ret} = t - \frac{|\boldsymbol{x}-\boldsymbol{x}'|}{c}\,. \tag{2.38}$$

Using the projection operator Eq. (2.29), we obtain the solution

$$h^{\rm TT}_{ij}(t,\boldsymbol{x}) = \frac{4G}{c^4}\Lambda_{ij,kl}(\boldsymbol{n})\int dx'^3\;\frac{1}{|\boldsymbol{x}-\boldsymbol{x}'|}T_{kl}\left(t-\frac{|\boldsymbol{x}-\boldsymbol{x}'|}{c},\boldsymbol{x}'\right)\,. \tag{2.39}$$

We denote the typical size of the source by d. Assuming the distance between the source and the detector r is longer than d, we expand the solution Eq. (2.39) as

$$h^{\rm TT}_{ij}(t,\boldsymbol{x}) = \frac{1}{r}\frac{4G}{c^4}\Lambda_{ij,kl}(\boldsymbol{n})\int d^3x'\;T_{kl}\left(t-\frac{r}{c}+\frac{\boldsymbol{x}'\cdot\hat{\boldsymbol{n}}}{c},\boldsymbol{x}'\right)\,. \tag{2.40}$$

Here, we introduce the low-velocity approximation, that is,

$$\frac{v}{c} \ll 1\,, \tag{2.41}$$

with the typical velocity of the source v. This assumption is equivalent to the condition

$$d \ll \frac{\lambda}{2\pi}\,, \tag{2.42}$$

where $\omega = c/2\pi\lambda$ is the typical angular frequency of the source motion and the source velocity v is roughly $v \sim \omega d$. Under these assumptions, the energy-momentum tensor can be expanded as

$$T_{kl}\left(t - \frac{r}{c} + \frac{\boldsymbol{x}'\cdot\boldsymbol{n}}{c}, \boldsymbol{x}'\right) \simeq T_{kl}\left(t - \frac{r}{c}, \boldsymbol{x}'\right) + \frac{x'^i n^i}{c}\partial_0 T_{kl} + \frac{1}{2c^2}x'^i x'^j n^i n^j \partial_0^2 T_{kl} + \cdots\,. \tag{2.43}$$

Substituting this into Eq. (2.40) and taking only the leading term, we obtain

$$h_{ij}^{\mathrm{TT}}(t,\boldsymbol{x}) \simeq \frac{1}{r}\frac{4G}{c^4}\Lambda_{ij,kl}(\boldsymbol{n})S_{kl}(t-r/c)\,, \tag{2.44}$$

where

$$S^{ij}(t) := \int d^3x\; T^{ij}(t,\boldsymbol{x})\,. \tag{2.45}$$

We define

$$M^{ij}(t) := \frac{1}{c^2}\int d^3x\; T^{00}(t,\boldsymbol{x})x^i x^j\,. \tag{2.46}$$

Then, we rewrite S_{ij} as

$$S_{ij}(t) = \frac{1}{2}\ddot{M}_{ij}(t)\,, \tag{2.47}$$

by virtue of the integral by part and the conservation law of the energy-momentum tensor. Using this relation, we obtain

$$h_{ij}^{\mathrm{TT}}(t,\boldsymbol{x}) = \frac{1}{r}\frac{2G}{c^4}\Lambda_{ij,kl}(\boldsymbol{n})\ddot{M}^{kl}(t-r/c)\,. \tag{2.48}$$

Because the projection tensor $\Lambda_{ij,kl}$ has the traceless property, i.e,. $\Lambda_{ij,kl}\delta^{kl} = 0$, the trace component of M^{kl} vanishes. Defining the quadrupole moment by

$$Q^{ij}(t) := M^{ij}(t) - \frac{1}{3}\delta^{ij}M_{kk}(t) = \int d^3x\;\rho(t,\boldsymbol{x})\left(x^i x^j - \frac{1}{3}r^2\delta^{ij}\right)\,, \tag{2.49}$$

with $\rho := T^{00}/c^2$, we obtain the quadrupole formula

$$h_{ij}^{\mathrm{TT}}(t,\boldsymbol{x}) = \frac{1}{r}\frac{2G}{c^4}\Lambda_{ij,kl}(\boldsymbol{n})\ddot{Q}_{kl}(t-r/c)\,. \tag{2.50}$$

If the gravitational wave propagates to the direction of z-axis,

$$h_+(t,\boldsymbol{x}) = \frac{1}{r}\frac{G}{c^4}\left\{\ddot{Q}_{11}(t-r/c) - \ddot{Q}_{22}(t-r/c)\right\}\,, \tag{2.51}$$

$$h_\times(t,\boldsymbol{x}) = \frac{2}{r}\frac{G}{c^4}\ddot{Q}_{12}(t-r/c)\,. \tag{2.52}$$

2.3.2 Energy-momentum tensor of gravitational waves

To discuss the energy of the gravitational waves, we move to the situation where the background metric $\bar{g}_{\mu\nu}$ is curved. The metric tensor is written as

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu} \,. \tag{2.53}$$

Generally speaking, we cannot separate the metric tensor into the background and the perturbation. Splitting the background and the perturbation can be done only when the length scale of these curvatures have hierarchy. We refer $\bar{L}$ and $\lambda_{\rm gw}$ to the length scales of the curvatures of the background and the perturbations, respectively. We consider the situation where the hierarchy

$$\lambda_{\rm gw} \ll \bar{L} \,, \tag{2.54}$$

is satisfied.

The background metric is assumed to be low-frequency components and the high-frequency components are regarded as the perturbation. The Ricci tensor is expanded up to the second order of h,

$$R_{\mu\nu} = \bar{R}_{\mu\nu} + R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + \cdots \,, \tag{2.55}$$

where $\bar{R}_{\mu\nu}$ does not contain h and $R^{(i)}_{\mu\nu}$ is the i-th order terms of h. The zero-th order term $\bar{R}_{\mu\nu}$ consists of only low frequency modes and the first order term $R^{(1)}_{\mu\nu}$ has only high frequency modes. The higher order term $R^{(i)}_{\mu\nu}$ for $i \leq 2$ contain both low and high frequency modes. Splitting the Ricci tensor into the high frequency and the low frequency modes, we get

$$\bar{R}_{\mu\nu} = -\left[R^{(2)}_{\mu\nu}\right]^{\rm low} \,, \; R^{(1)}_{\mu\nu} = -\left[R^{(2)}_{\mu\nu}\right]^{\rm high} \,, \tag{2.56}$$

where $[\cdots]^{\rm low/high}$ is the quantity having low/high frequency terms. The energy of the gravitational waves is measured by the effect to the background. Therefore, the low frequency modes gives us the energy of the gravitational waves.

To extract the low frequency modes by the systematic way, we introduce a length scale $L_{\rm avg}$ which satisfies,

$$\lambda_{\rm gw} \ll L_{\rm avg} \ll \bar{L} \,, \tag{2.57}$$

and integrate out the high frequency modes by taking the spatial average over the region of $L_{\rm avg}$. Then we obtain

$$\bar{R}_{\mu\nu} = -\langle R^{(2)}_{\mu\nu} \rangle \,, \tag{2.58}$$

where the bracket $\langle \cdots \rangle$ means the spatial average. We rewrite it as

$$\bar{R}_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}\bar{R} = \frac{8\pi G}{c^4} t^{\rm GW}_{\mu\nu}, \tag{2.59}$$

where

$$t^{\rm GW}_{\mu\nu} := -\frac{c^4}{8\pi G}\left\langle R^{(2)}_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}R^{(2)} \right\rangle \,, \tag{2.60}$$

where

$$R^{(2)} := \bar{g}^{\mu\nu} R^{(2)}_{\mu\nu} \,. \tag{2.61}$$

This can be interpreted as the energy-momentum tensor of the gravitational waves.

In the distant region from the source (i.e., the region where $T_{\mu\nu}$ vanishes), we can approximate $\bar{g}_{\mu\nu}$ is flat. After taking the Lorentz gauge, the energy-momentum tensor $t^{\rm GW}_{\mu\nu}$ is

$$t^{\rm GW}_{\mu\nu} = \frac{c^4}{32\pi G} \langle \partial_\mu h_{\alpha\beta} \partial_\nu h^{\alpha\beta} \rangle \,. \tag{2.62}$$

We can take TT gauge without changing the form of $t^{\rm GW}_{\mu\nu}$. Thus, the energy-momentum tensor of the gravitational waves is rewritten as

$$t^{\rm GW}_{\mu\nu} = \frac{c^4}{32\pi G} \langle \partial_\mu h^{\rm TT}_{ij} \partial_\nu h^{ij}_{\rm TT} \rangle \,. \tag{2.63}$$

Especially, the $(0j)$-component corresponds to the energy flux. We obtain the luminosity of the gravitational waves as

$$\mathcal{L}_{\rm GW} = r^2 c \oint d\Omega \; n^k t^{\rm GW}_{0k} = \frac{r^2 c^5}{32\pi G} \oint d\Omega \; n^k \langle \partial_0 h^{\rm TT}_{ij} \partial_k h^{ij}_{\rm TT} \rangle \,. \tag{2.64}$$

where the surface integral is done over a spherical surface of the radius r. Assuming that the spherical wave propagating outward, the spatial derivative can be recast into

$$n^k \partial_k h^{\rm TT}_{ij}(t - r/c) = -\partial_0 h^{\rm TT}_{ij}(t - r/c) \,. \tag{2.65}$$

Then, the luminosity can be written as

$$\mathcal{L}_{\rm GW} = \frac{c^3 r^2}{32\pi G} \oint d\Omega \; \langle \dot{h}^{\rm TT}_{ij} \dot{h}^{\rm TT}_{ij} \rangle \,. \tag{2.66}$$

Due to the energy conservation, the luminosity should be balanced the change rate of the energy of the source, i.e.,

$$\frac{dE_{\rm source}}{dt} = -\mathcal{L}_{\rm GW} \,. \tag{2.67}$$

2.3.3 Example 1: Rotating triaxial ellipsoid

Let us consider the gravitational waves emitted from a rotating triaxial ellipsoid as a model of rotating neutron stars. We take the coordinates where the moment of inertia is diagonalized. The rotation axis is assumed to coincide with one of the principal axes of inertia. In the co-rotating frame where the ellipsoid is at rest, the quadrupole moment can be written in the form

$$Q'_{ij} = \begin{pmatrix} Q_x & 0 & 0 \\ 0 & Q_y & 0 \\ 0 & 0 & Q_z \end{pmatrix} \,. \tag{2.68}$$

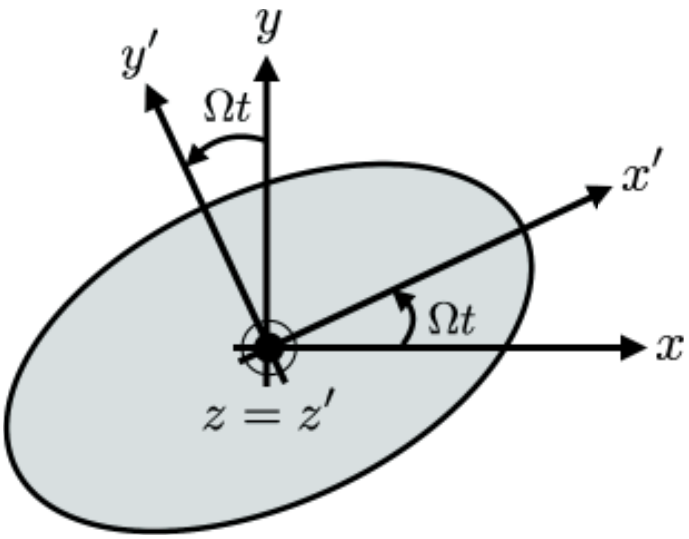

Figure 2.1: The schematic picture of a rotating triaxial ellipsoid. The coordinate $\{x', y', z'\}$ is the co-rotating frame where the ellipsoid is at rest. Also, each axis of the co-rotating frame coincides with each principal axis of inertia. The ellipsoid rotates around the z'-axis with the angular frequency Ω.

Using the matrix

$$R_{ij}(t) = \begin{pmatrix} \cos\Omega t & -\sin\Omega t & 0 \\ \sin\Omega t & \cos\Omega t & 0 \\ 0 & 0 & 1 \end{pmatrix} , \tag{2.69}$$

we obtain the quadrupole moment in the laboratory frame as

$$Q_{ij}(t) = R_{ik}(t)R_{j\ell}(t)Q'_{k\ell} \, . \tag{2.70}$$

Because of the quadrupole formula, the gravitational waves from rotating triaxial unequal ellipsoid can be calculated as

$$h_+(t) = \frac{1}{r}\frac{G}{c^4}4\Omega^2(Q_y - Q_x)\cos 2\Omega\tilde{t} \, , \tag{2.71}$$

$$h_\times(t) = \frac{1}{r}\frac{G}{c^4}4\Omega^2(Q_y - Q_x)\sin 2\Omega\tilde{t} \, , \tag{2.72}$$

with the retarded time $\tilde{t}$ defined by

$$\tilde{t} := t - \frac{r}{c} \, . \tag{2.73}$$

The ellipticity ϵ is defined by

$$\epsilon := \frac{Q_y - Q_x}{Q_z} \, , \tag{2.74}$$

and characterizes the asymmetry between each axis of inertia. The gravitational waveform can be rewritten in the form

$$h_+(t) = h_0 \cos 2\pi f_{\rm gw}\tilde{t}\,, \tag{2.75}$$

$$h_\times(t) = h_0 \sin 2\pi f_{\rm gw}\tilde{t}\,, \tag{2.76}$$

with the gravitational wave frequency $2\pi f_{\rm gw} = 2\Omega$ and the amplitude of the gravitational wave

$$h_0 := \frac{4\pi^2 G}{c^4}\frac{Q_z f_{\rm gw}^2}{r}\epsilon\,. \tag{2.77}$$

Using typical parameters of neutron stars, we can estimate the gravitational wave amplitude to be

$$h_0 \simeq 10^{-27}\left(\frac{Q_z}{1.1\times 10^{45}{\rm g\ cm^2}}\right)\left(\frac{r}{10{\rm kpc}}\right)^{-1}\left(\frac{f_{\rm gw}}{100{\rm Hz}}\right)^2\left(\frac{\epsilon}{10^{-6}}\right)\,. \tag{2.78}$$

2.3.4 Example 2: Compact binary inspiral

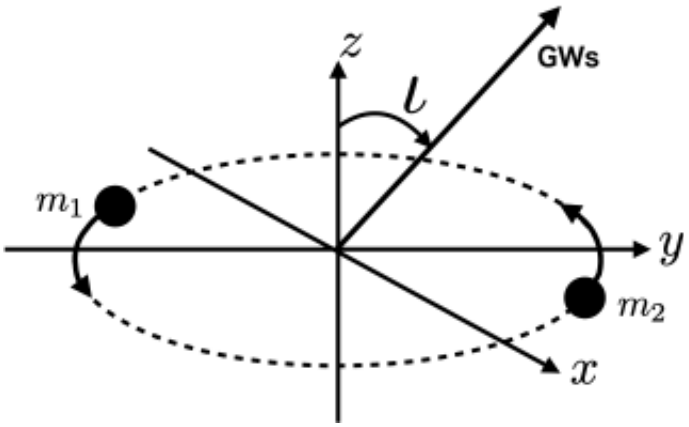

Figure 2.2: The schematic picture of an inspiraling compact binary. The orbital plane is located within the xy-plane. Thus, the orbital angular momentum is directed along the z-axis. The inclination angle ι is the angle between the z-axis and the propagation direction of the gravitational waves.

Second example is two compact objects orbiting around each other. The masses of two compact objects are denoted by m_1 and m_2, respectively. The (x, y)-plane is set on the orbital plane and the z-axis is directed along the orbital angular momentum (see Fig. 2.2). The coordinate origin is set to the center of mass. We approximate the orbit is Keplerian and the compact objects are regarded as point particles. The quadrupole moment of such system is

$$Q_{ij} = \frac{2m_1m_2a}{M_{\rm tot}}\begin{pmatrix} \frac{1}{6}+\frac{1}{2}\cos 2\Omega t & \frac{1}{2}\sin 2\Omega t & 0 \\ \frac{1}{2}\sin 2\Omega t & \frac{1}{6}-\frac{1}{2}\cos 2\Omega t & 0 \\ 0 & 0 & -\frac{1}{3}\end{pmatrix}\,, \tag{2.79}$$

where Ω is an angular velocity of the binary, a is the separation and $M_{\rm tot} := m_1 + m_2$ is the total mass of the binary. Therefore, the gravitational waves from the inspiraling binary can be calculated as

$$h_{ij}(t, \boldsymbol{x}) = \frac{4G\mu}{c^4 r}[\pi G M_{\rm tot} f_{\rm gw}]^{2/3} \begin{pmatrix} \cos 2\pi f_{\rm gw}\tilde{t} & \sin 2\pi f_{\rm gw}\tilde{t} & 0 \\ \sin 2\pi f_{\rm gw}\tilde{t} & -\cos 2\pi f_{\rm gw}\tilde{t} & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij} , \tag{2.80}$$

with $f_{\rm gw} = \Omega/\pi$ and the reduced mass $\mu := m_1 m_2/M_{\rm tot}$. The waveform depends on the masses through their specific combination. The chirp mass is defined by

$$\mathcal{M}_{\rm c} := \mu^{3/5} M_{\rm tot}^{2/5} , \tag{2.81}$$

and the waveform can be written as

$$h_{ij}(t, \boldsymbol{x}) = \frac{4}{r}\left(\frac{G\mathcal{M}_{\rm c}}{c^2}\right)^{5/3}\left(\frac{\pi f_{\rm gw}}{c}\right)^{2/3} \begin{pmatrix} \cos 2\pi f_{\rm gw}\tilde{t} & \sin 2\pi f_{\rm gw}\tilde{t} & 0 \\ \sin 2\pi f_{\rm gw}\tilde{t} & -\cos 2\pi f_{\rm gw}\tilde{t} & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij} . \tag{2.82}$$

This expression is the model of the gravitational waves propagating to the direction perpendicular to the orbital plane. We introduce the inclination angle ι, which is the angle between the propagation direction and the orbital angular momentum of the binary. Taking into account of the angular dependence, we obtain

$$h_{ij}(\iota) = \left[R_2(\iota) \cdot h \cdot R_2^{\rm T}(\iota)\right]_{ij} , \tag{2.83}$$

where $R_2(\iota)$ is the matrix representing the rotation around the y-axis by the angle ι, i.e.,

$$R_2(\iota) = \begin{pmatrix} \cos\iota & 0 & -\sin\iota \\ 0 & 1 & 0 \\ \sin\iota & 0 & \cos\iota \end{pmatrix} . \tag{2.84}$$

Therefore, the waveform becomes

$$h_+(t, \boldsymbol{x}) = h_0 \frac{1 + \cos^2\iota}{2} \cos(2\pi f_{\rm gw}\tilde{t}) , h_\times(t, \boldsymbol{x}) = h_0 \cos\iota \sin(2\pi f_{\rm gw}\tilde{t}) , \tag{2.85}$$

with

$$h_0 = \frac{4}{r}\left(\frac{G\mathcal{M}_{\rm c}}{c^2}\right)^{5/3}\left(\frac{\pi f_{\rm gw}}{c}\right)^{2/3} . \tag{2.86}$$

Assuming that the binary is Keplerian, the binary has the energy

$$E_{\rm binary} = -\frac{G m_1 m_2}{2a} = -\frac{1}{2}\nu^{2/3}\mathcal{M}_{\rm c}c^2 , \tag{2.87}$$

where

$$\nu := \frac{\pi G \mathcal{M}_{\rm c} f_{\rm gw}}{c^3} . \tag{2.88}$$

is the dimensionless parameter. The average $\langle \cdots \rangle$ in the expression of the luminosity $\mathcal{L}_{\rm GW}$ is replaced with the time-average. Then the luminosity of the gravitational waves is

$$\mathcal{L}_{\rm GW} = \frac{32}{5}\left(\frac{c^5}{G}\right)\nu^{10/3}\,. \tag{2.89}$$

Rewriting the energy balance equation as

$$\frac{dE_{\rm binary}}{dt} = \frac{d\nu}{dt}\frac{dE_{\rm binary}}{d\nu} = \frac{\dot{f}_{\rm gw}}{f_{\rm gw}}\cdot\nu\cdot\frac{dE_{\rm binary}}{d\nu} = -\mathcal{L}_{\rm GW}\,, \tag{2.90}$$

we obtain the change rate of the frequency

$$\dot{f}_{\rm gw} = \frac{96}{5}\left(\frac{c^3}{G\mathcal{M}_{\rm c}}\right)\nu^{8/3} f_{\rm gw} = \frac{96}{5}\pi^{8/3}\left(\frac{G\mathcal{M}_{\rm c}}{c^3}\right)^{5/3} f_{\rm gw}^{11/3}\,. \tag{2.91}$$

2.4 Continuous gravitational waves

The *continuous gravitational waves* is the type of gravitational waves having the following properties,

- duration longer than the observational period,
- small change rate of the frequency,
- almost constant amplitude.

In this section, we briefly review the expected sources of continuous gravitational waves.

2.4.1 Rotating neutron stars

Rotating neutron stars are the main target of the continuous wave search (Ref. [27] as a review). We already know that a lot of rotating neutron stars exist in our Galaxy from radio observations. They are known as pulsars which periodically emit radio pulses. The pulsars detected so far are shown in Fig. 2.3. Especially, tens of millisecond pulsars, which rotate with a period $P \sim 10^2$ sec and the change rate of the frequency is about $|\dot{P}| \sim 10^{-20}$, are the typical sources of continuous gravitational waves. The change rate of the gravitational wave frequency can be estimated by

$$\dot{f_{\rm gw}} \sim f_{\rm gw}^2|\dot{P}| \sim 10^{-16}{\rm Hz/sec}\left(\frac{|\dot{P}|}{10^{-20}{\rm sec/sec}}\right)\left(\frac{f_{\rm gw}}{10^2{\rm Hz}}\right)^2\,. \tag{2.92}$$

As we have seen in the previous section, a rotating deformed rigid body emits gravitational waves. The amplitude of the gravitational waves depends on the deformation parameter ϵ. There are many scenarios to deform neutron stars.

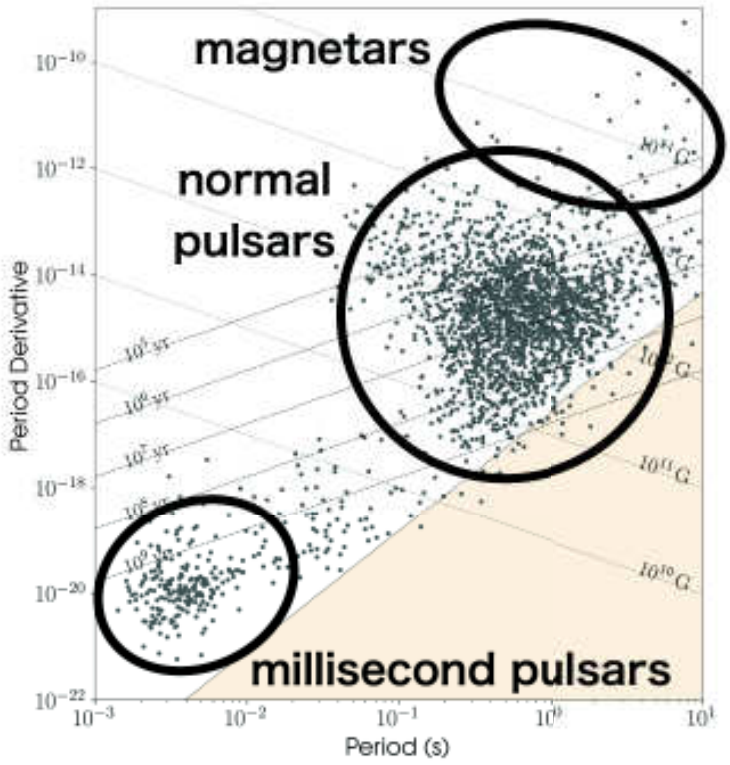

Figure 2.3: The P-$\dot{P}$ plot of the pulsars. From this figure, we can roughly classify all pulsars into three types, 1) normal pulsars, 2) millisecond pulsars, and 3) magnetars. Possible sources of gravitational waves that can be detected by the ground-based detectors are millisecond pulsars, which have a period $\sim O(10^{-2})$sec. The pulsar catalog is obtained from ATNF catalog. (`https://www.atnf.csiro.au/research/pulsar/psrcat/`) The figure is plotted with the module `psrqpy` and slightly modified.

- "Mountain" on neutron stars. If there is a mountain on the neutron star, it can produce quadrupole moment. The ellipticity ϵ is estimated by

$$\epsilon \lesssim 2 \times 10^{-5} \left(\frac{u_{\rm break}}{0.1} \right) ,$$

  where $u_{\rm break}$ is the crustal breaking strain.

- Magnetic fields. Neutron stars are believed to have strong magnetic field. It causes the deformation of neutron stars. For the case of toroidal magnetic field, the ellipticity can achieve

$$\epsilon \sim 3 \times 10^{-11} \left(\frac{R_{\rm NS}}{10{\rm km}} \right)^4 \left(\frac{M_{\rm NS}}{1.4 M_\odot} \right)^{-2} \left(\frac{B}{10^{12}{\rm G}} \right)^2 .$$

2.4.2 Axion clouds around black holes

According to the QCD or the string theory, ultra-light bosons can exist in the Universe while they have not been detected. These bosons are called QCD/string axions. These scalar fields may extract the energy and the angular momentum from the Kerr black hole by the superradiance, and the bound state can be formed around the black hole. This bound state is called "axion cloud", and the coupled system of the black hole and the cloud is regarded as the "gravitational atoms". Because the cloud is formed by a number of boson particles, the reactions of small event rates can be occurred in the cloud. In BH-cloud systems, the pair annihilation of two axion particles and the level transition of the clouds can emit the gravitational waves [36].

The system is characterized by the parameter α which is defined by the ratio of the Schwarzschild radius of the black hole and the Compton wavelength of the axion, i.e,

$$\alpha := \frac{R_{\rm Sch}}{\lambda_{\rm axion}} \sim 7.5 \times 10^{-2} \left(\frac{M_{\rm BH}}{10 M_\odot}\right) \left(\frac{m_{\rm axion}}{10^{-12}\ {\rm eV}}\right) . \tag{2.93}$$

where $M_{\rm BH}$ is the black hole mass and $m_{\rm axion}$ is the axion mass. The cloud is expected to emit gravitational waves via various processes. Roughly speaking, the leading order of the gravitational wave frequency is

$$f_{\rm gw} \sim \frac{2 m_{\rm axion} c^2}{h} \sim 2.5 \times 10^2\ {\rm Hz} \left(\frac{m_{\rm axion}}{10^{-12}\ {\rm eV}}\right) . \tag{2.94}$$

For $\alpha \ll 1$, the luminosity of the gravitaional waves from the axion cloud is estimated by [37]

$$\mathcal{L}_{\rm GW} \sim 0.025 \frac{c^5}{G} \alpha^{14} \left(\frac{M_{\rm cloud}}{M_{\rm BH}}\right)^2 . \tag{2.95}$$

Here, $M_{\rm cloud}$ is the mass of the axion cloud around the black hole and is estimated by

$$M_{\rm cloud} \sim \alpha M_{\rm BH} . \tag{2.96}$$

By roughly rewriting Eq. (2.66) as

$$\mathcal{L}_{\rm GW} \sim \frac{c^3}{G} r^2 \omega_{\rm GW}^2 h_0^2 , \tag{2.97}$$

the characteristic strain is estimated by

$$h_0 \sim 2.1 \times 10^{-25} \left(\frac{\alpha}{0.075}\right)^7 \left(\frac{M_{\rm BH}}{10 M_\odot}\right) \left(\frac{1{\rm kpc}}{r}\right) . \tag{2.98}$$

The duration $\tau_{\rm GW}$ is estimated by

$$\tau_{\rm GW} \sim \frac{M_{\rm cloud} c^2}{\mathcal{L}_{\rm GW}} \sim 1.9 \times 10^{11}\ {\rm sec} \left(\frac{\alpha}{0.075}\right)^{-15} \left(\frac{M_{\rm BH}}{10 M_\odot}\right) . \tag{2.99}$$

Therefore, the gravitational waves lasts much longer than the observation period.

2.4.3 Sub-solar mass black hole binaries

Stellar mass black holes are believed to be formed by supernovae of massive stars. On the other hand, it is theoretically expected that black holes are also generated at over density regions in the early era of the Universe. Black holes that are formed in the early Universe are called primordial black holes (PBHs) (Ref. [38] as a review). Theoretically, PBHs can become flyweight while black holes originating from supernovae of massive stars cannot be less massive than several $M_\odot$. Using Eq. (2.91), we obtain

$$\dot{f}_{\rm gw} = 1.3 \times 10^{-14}{\rm Hz/sec} \left(\frac{\mathcal{M}_{\rm c}}{10^{-9}M_\odot}\right)^{5/3} \left(\frac{f_{\rm gw}}{100{\rm Hz}}\right)^{11/3} . \tag{2.100}$$

Therefore, the frequency is almost constant during the observation period.

2.5 Gravitational wave detectors

2.5.1 Michelson interferometer

We can directly detect gravitational waves with two free-falling test masses. In this section, we consider the detection of the gravitational waves with a Michelson interferometer (see Figure 2.4). The Michelson interferometer consists of a photon emitter, a photo detector, a beam splitter and two mirrors. After two photons propagate different paths in the existence of gravitational waves, they would have different phases even if initially their phases coincide. In order to prepare two photons having a same phases, a photon emitted by a photon emitter is separated into two photons by a beam splitter. These photons are reflected by mirrors and collected into a photo detector.

The distance between the beam splitter and the mirror A (B) is denoted by $L_{\rm A}$ ($L_{\rm B}$). In this thesis, we assume that two arms have an equal length, i.e.,

$$L_{\rm A} = L_{\rm B} =: L\,. \tag{2.101}$$

For simplicity, we here assume that gravitational waves arriving at the detector are approximated by a plane wave.

First, we calculate the frequency shift of the photon traveling between the emitter and the mirror A. The mirror is treated as the test mass and is rest with respect to the emitter. The unit vector pointing from the emitter to the mirror A is denoted by $\hat{\boldsymbol{p}}_{\rm A}$. The gravitational waves propagate along with the direction of the unit vector $\hat{\boldsymbol{n}}$. We take the Cartesian coordinate where $\hat{\boldsymbol{p}}_{\rm A}$ lies on the xy plane and the origin is set to be the position of the emitter. The unperturbed 4-vector of the photon traveling from the emitter to the mirror A is written as

$$\bar{k}^\mu_{\rm A} = \frac{\nu}{c}(1, \hat{\boldsymbol{p}}_{\rm A})\,. \tag{2.102}$$

The perturbed vector is denoted by $k^\mu_{\rm A} = \bar{k}^\mu_{\rm A} + \delta k^\mu_{\rm A}$. In the following calculation, we neglect higher-order terms of $h_{\mu\nu}$ and $\delta k^\mu_{\rm A}$. The change of the photon frequency ν is determined

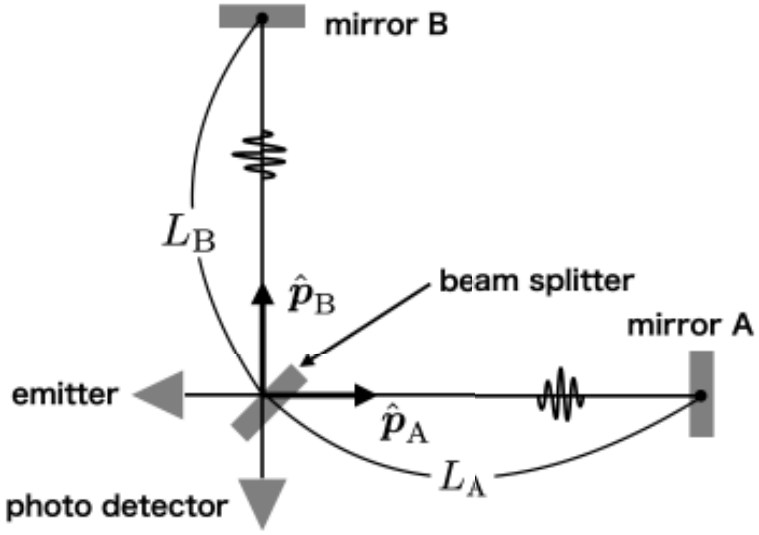

Figure 2.4: The rough sketch of a Michelson interferometer. It consists of a photon emitter, a photo detector, a beam splitter, and two mirrors. The unit vector pointing from the beam splitter to the mirror A (B) is denoted by $\hat{\boldsymbol{p}}_{\rm A}$ ($\hat{\boldsymbol{p}}_{\rm B}$). The distance between the beam splitter and the mirror A (B) is denoted by $L_{\rm A}$ ($L_{\rm B}$). A photon is emitted from the photo emitter and splitted into two photons by the beam splitter. Each photon propagates to each mirror and reflected photons are collected by the photo detector. The phase deviation between two photons caused by the gravitational waves is detected by the photo detector.

by the geodesic equation of $k^\mu_{\rm A}$. The zeroth component of the geodesic equation is

$$\frac{dk^0_{\rm A}}{d\xi} = \frac{1}{2}\bar{k}^i_{\rm A}\bar{k}^j_{\rm A}\partial_0 h_{ij}\,, \tag{2.103}$$

where ξ is the affine parameter. From

$$\frac{dh_{ij}}{d\xi} = \frac{dx^0}{d\xi}\partial_0 h_{ij} + \frac{dx^k}{d\xi}\partial_k h_{ij} = \frac{\nu(1-\hat{\boldsymbol{p}}_{\rm A}\cdot\hat{\boldsymbol{n}})}{c}\partial_0 h_{ij}\,, \tag{2.104}$$

we get

$$\frac{dk^0_{\rm A}}{d\xi} = \frac{c}{2\nu(1-\hat{\boldsymbol{p}}_{\rm A}\cdot\hat{\boldsymbol{n}})}\frac{dh_{ij}}{d\xi}\left(\frac{\nu}{c}\right)^2\hat{p}^i_{\rm A}\hat{p}^j_{\rm A}\,. \tag{2.105}$$

The redshift of the frequency is obtained by integrating the geodesic equation,

$$\frac{\nu_{\rm A}-\nu_0}{\nu_0} \simeq \ln(\nu_{\rm A}/\nu_0) = \int_{\xi_{\rm A}}^{\xi_0} d\xi\,\frac{1}{\nu}\frac{d\nu}{d\xi} = \frac{1}{2}\frac{\hat{p}^i_{\rm A}\hat{p}^j_{\rm A}}{1-\hat{\boldsymbol{p}}_{\rm A}\cdot\hat{\boldsymbol{n}}}\left[h^{\rm TT}_{ij}(x_0) - h^{\rm TT}_{ij}(x_{\rm A})\right]\,, \tag{2.106}$$

where

$$h^{\rm TT}_{ij}(x_0) := h^{\rm TT}_{ij}(t_0)\,, \tag{2.107}$$

and

$$h_{ij}^{\rm TT}(x_{\rm A}) := h_{ij}^{\rm TT}(t_0 + L(1 - \hat{\boldsymbol{p}}_{\rm A} \cdot \hat{\boldsymbol{n}})/c)\,. \tag{2.108}$$

with some reference time t_0. Therefore, the redshift is

The photon is reflected by the mirror A and goes back to the emitter. The reflected photon has the frequency $\nu_{0\rm A}$ and it can be calculated as

$$\ln(\nu_{0\rm A}/\nu_{\rm A}) = \int_{\xi_{0\rm A}}^{\xi_{\rm A}} d\xi\ \frac{1}{\nu}\frac{d\nu}{d\xi} = \frac{1}{2}\frac{\hat{p}_{\rm A}^i \hat{p}_{\rm A}^j}{1 + \hat{\boldsymbol{p}}_{\rm A} \cdot \hat{\boldsymbol{n}}} \left[h_{ij}^{\rm TT}(x_{\rm A}) - h_{ij}^{\rm TT}(x_{0\rm A})\right]\,, \tag{2.109}$$

where

$$h_{ij}^{\rm TT}(x_{0\rm A}) = h_{ij}^{\rm TT}(t_0 + 2L/c)\,. \tag{2.110}$$

We impose the long-wavelength approximation that the detector size is much smaller than the wavelength of the gravitational waves. This assumption is justified for the case of the ground-based detectors. The typical size of the ground-based interferometer is $L \sim O(1)$ km and the ground-based detector is sensitive to the gravitational waves of the wavelength $\lambda_{\rm gw} \sim c \cdot (100{\rm Hz})^{-1} \sim 3 \times 10^3$ km. Therefore,

$$L \ll \lambda_{\rm gw} \tag{2.111}$$

is valid for ground-based interferometers. This condition can be rewritten as

$$\frac{L}{c} \cdot \frac{1}{\omega_{\rm gw}} \ll 1\,, \tag{2.112}$$

where $\omega_{\rm gw} = 2\pi c/\lambda_{\rm gw}$. The Taylor expansion of $h_{ij}^{\rm TT}$ is carried out, and we obtain

$$h_{ij}^{\rm TT}(t_0 + L(1 - \hat{\boldsymbol{p}}_{\rm A} \cdot \hat{\boldsymbol{n}})/c) = h_{ij}^{\rm TT}(t_0) + \frac{L(1 - \hat{\boldsymbol{p}}_{\rm A} \cdot \hat{\boldsymbol{n}})}{c}\dot{h}_{ij}^{\rm TT}(t_0) + \cdots\,. \tag{2.113}$$

This expansion is interpreted as the expansion by the small parameter $L/c\omega_{\rm gw}$. From Eq. (2.106), the redshift can be written as

$$\frac{\nu_{\rm A} - \nu_0}{\nu_0} \simeq -\frac{L}{2c}\hat{p}_{\rm A}^i \hat{p}_{\rm A}^j \dot{h}_{ij}^{\rm TT}(t_0)\,, \tag{2.114}$$

up to the first order of h. The redshift of the reflected photon is computed in the same manner as

$$\frac{\nu_{\rm A0} - \nu_{\rm A}}{\nu_0} \simeq -\frac{L}{2c}\hat{p}_{\rm A}^i \hat{p}_{\rm A}^j \dot{h}_{ij}^{\rm TT}(t_0)\,. \tag{2.115}$$

The phase of the photon at the time when the photon gets back to the emitter is calculated by

$$\phi_{\rm A}(t) = 2\pi\nu_0 \int dt\ \frac{\nu_{\rm A0}}{\nu_0} = 2\pi\nu_0 \left(1 - \frac{L}{2c}\hat{p}_{\rm A}^i \hat{p}_{\rm A}^j h_{ij}^{\rm TT}(t)\right)\,. \tag{2.116}$$

Similarly, the phase of the photon reflected by the mirror B is obtained as

$$\phi_{\rm B}(t) = 2\pi\nu_0 \left(1 - \frac{L}{2c}\hat{p}_{\rm B}^i \hat{p}_{\rm B}^j h_{ij}^{\rm TT}(t)\right)\,. \tag{2.117}$$

The difference of the photon phases is

$$\Delta\phi(t) := \phi_{\rm B} - \phi_{\rm A} = \frac{2\pi\nu_0 L}{c} \cdot \frac{1}{2}(\hat{p}^i_{\rm A}\hat{p}^j_{\rm A} - \hat{p}^i_{\rm B}\hat{p}^j_{\rm B})h^{\rm TT}_{ij}(t)\,. \tag{2.118}$$

The phase fluctuation can be detected by a photo detector. In this sense, what we directly observe is not $h^{\rm TT}_{ij}$ itself but

$$h_{\rm obs}(t) = \frac{1}{2}(\hat{p}^i_{\rm A}\hat{p}^j_{\rm A} - \hat{p}^i_{\rm B}\hat{p}^j_{\rm B})h^{\rm TT}_{ij}(t)\,. \tag{2.119}$$

The detector tensor is defined by

$$d^{ij} := \frac{1}{2}(\hat{p}^i_{\rm A}\hat{p}^j_{\rm A} - \hat{p}^i_{\rm B}\hat{p}^j_{\rm B})\,, \tag{2.120}$$

and the detected strain is written as

$$h_{\rm obs}(t) = d^{ij}h^{\rm TT}_{ij}(t)\,. \tag{2.121}$$

Using the polarization tensor, the gravitational waves can be decomposed into

$$h^{\rm TT}_{ij}(t) = h_+(t)e^+_{ij} + h_\times(t)e^\times_{ij}\,. \tag{2.122}$$

In general situation, the vectors $\hat{\boldsymbol{p}}_{\rm A}$ and $\hat{\boldsymbol{p}}_{\rm B}$ change in time. Then, the observed signal is written as

$$h_{\rm obs}(t) = F_+(t)h_+(t) + F_\times(t)h_\times(t)\,, \tag{2.123}$$

with the antenna pattern functions defined by

$$F_+(t) := d^{ij}(t)e^+_{ij}\,,\ F_\times(t) := d^{ij}(t)e^\times_{ij}\,. \tag{2.124}$$

The detail of the antenna pattern functions is shown in Appendix. A.

Chapter 3

Matched filtering and continuous gravitational waves

Gravitational wave detectors have complicated structures consisting of many components. All of these components are possible sources of the noise. Improving data analysis techniques is an important research area for the gravitational wave astronomy. When the gravitational waveform expected to be observed is a priori known and the detector noise is stationary and Gaussian, the matched filtering is the optimal method for detection and the parameter estimation. In this chapter, we explain the fundamentals of the matched filtering and its application to the continuous gravitational waves.

3.1 Matched filtering

3.1.1 Statistical properties of detector noise

Let us assume that the detector noise $n(t)$ is stationary and Gaussian. The average of noise is assumed to be zero, i.e.,

$$\langle n(t)\rangle = 0\,,\ \langle \tilde{n}(f)\rangle = 0\,, \tag{3.1}$$

where $\langle\cdots\rangle$ is the ensemble average. The probability density function of the noise $n(t)$ can be written as

$$p_{\rm noise}[n(t)] \propto \exp\left[-\frac{1}{2}\cdot 4\int_0^\infty df \frac{|\tilde{n}(f)|^2}{S_{\rm n}(f)}\right]\,. \tag{3.2}$$

Here, $S_{\rm n}(f)$ is called the power spectral density of the noise and defined by

$$\langle \tilde{n}(f)\tilde{n}^*(f')\rangle = \frac{1}{2}S_{\rm n}(f)\delta(f-f')\,. \tag{3.3}$$

Defining the noise weighted inner product by

$$(a|b) = 4{\rm Re}\left[\int_0^\infty df \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_{\rm n}(f)}\right]\,, \tag{3.4}$$

we can rewrite $p_{\rm noise}[n(t)]$ as

$$p_{\rm noise}[n(t)] \propto \exp\left[-\frac{(n|n)}{2}\right] . \tag{3.5}$$

3.1.2 Matched filter as optimal statistic

Let us consider to test two hypotheses,

$$\begin{cases} \mathcal{H}_0 : s(t) = n(t) , & \text{(only detector noise exist)}, \\ \mathcal{H}_1 : s(t) = n(t) + h(t) , & \text{(signal exists)}, \end{cases} \tag{3.6}$$

for an observed strain data $s(t)$. The match is defined by the inner product between the observed strain $s(t)$ and the waveform $h(t)$, i.e.,

$$(s|h) = 4{\rm Re}\left[\int_0^\infty df \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_{\rm n}(f)}\right] . \tag{3.7}$$

The match is employed as a detection statistic, i.e., we regard the hypothesis $\mathcal{H}_1$ is correct when the match $(s|h)$ exceeds a threshold value.

When the hypothesis $\mathcal{H}_0$ is true, from Eq. (3.5), we can derive a likelihood

$$p[s(t)|\mathcal{H}_0] = p_{\rm noise}[s(t)] \propto \exp\left[-\frac{(s|s)}{2}\right] . \tag{3.8}$$

Similarly, when the hypothesis $\mathcal{H}_1$ is true, a likelihood can be calculated as

$$p[s(t)|\mathcal{H}_1] = p_{\rm noise}[s(t) - h(t)] \propto \exp\left[-\frac{(s-h|s-h)}{2}\right] \tag{3.9}$$

Generally speaking, the likelihood ratio gives us the most powerful unbiased test under a given significant level (the Neyman-Pearson lemma). Using Eqs. (3.8) and (3.9), the likelihood ratio is obtained as

$$\Lambda[s(t)] := \frac{p[s(t)|\mathcal{H}_1]}{p[s(t)|\mathcal{H}_0]} = \exp\left[(s|h) - \frac{1}{2}(h|h)\right] . \tag{3.10}$$

Because the likelihood ratio monotonically increases with $(s|h)$, the likelihood ratio test is equivalent to the test where we use the match $(s|h)$ as the detection statistic.

Let us examine the statistical properties of the match $(s|h)$. When the signal $h(t)$ is absent, the expected value and the variance of $(s|h)$ are calculated as

$$\langle (s|h) \rangle = 0 , \tag{3.11}$$

$${\rm Var}[(s|h)] = \langle (s|h)^2 \rangle - \langle (s|h) \rangle^2 = (h|h) . \tag{3.12}$$

Here, we use

$$\begin{aligned}\langle (s|h)^2\rangle &= \langle (n|h)^2\rangle \\ &= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \, \frac{\left\langle (\tilde{n}(f)\tilde{h}^*(f)+\tilde{h}(f)\tilde{n}^*(f))(\tilde{n}(f')\tilde{h}^*(f')+\tilde{h}(f')\tilde{n}^*(f'))\right\rangle}{S_{\rm n}(|f|)S_{\rm n}(|f'|)} \\ &= 4\cdot\int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \, \frac{1}{2}\frac{S_{\rm n}(f)\delta(f+f')\tilde{h}(f)\tilde{h}^*(f')}{S_{\rm n}(|f|)S_{\rm n}(|f'|)} \\ &= 2\int_{-\infty}^{\infty} df \, \frac{|\tilde{h}(f)|^2}{S_{\rm n}(f)} \\ &= (h|h)\,. \end{aligned} \tag{3.13}$$

Because we assume that the detector noise $n(t)$ is Gaussian and the match is linear in $n(t)$, the match $(s|h)$ follows a Gaussian distribution with mean 0 and variance $(h|h)$ when the strain data contains only detector noise. When the signal exists, we obtain

$$\langle (s|h)\rangle = (h|h)\,, \tag{3.14}$$

$$\mathrm{Var}[(s|h)] = \langle (s|h)^2\rangle - \langle (s|h)\rangle^2 = (h|h)\,. \tag{3.15}$$

Therefore, when the signal exists in the strain data, the match $(s|h)$ follows a Gaussian distribution with mean $(h|h)$ and variance $(h|h)$. Here, we define the signal-to-noise ratio (SNR) $\rho_{\rm MF}$ by

$$\rho_{\rm MF} := \frac{(s|h)}{\sqrt{(h|h)}}\,. \tag{3.16}$$

From the above discussion, $\rho_{\rm MF}$ follows a Gaussian distribution with mean 0 and variance 1 when the signal is absent, and a Gaussian distribution with mean $\sqrt{(h|h)}$ and variance 1 when the signal exists. In literature, $\langle \rho_{\rm MF}\rangle$ is also referred to as a signal-to-noise ratio and implies how large SNR can be expected with the given detector configuration.

Generally speaking, gravitational wave signals depend on various parameters (e.g., the masses of the compact binary). We denote them by $\boldsymbol{\theta}$. Correspondingly, we need to set the alternative hypothesis as

$$\mathcal{H}_{\boldsymbol{\theta}} : s(t) = n(t) + h(t;\boldsymbol{\theta})\,. \tag{3.17}$$

The maximum likelihood statistic is defined by

$$\Lambda[s(t)] := \max_{\boldsymbol{\theta}} \tilde{\Lambda}[s(t);\boldsymbol{\theta}] \tag{3.18}$$

where

$$\tilde{\Lambda}[s(t);\boldsymbol{\theta}] := \frac{p[s(t)|\mathcal{H}_{\boldsymbol{\theta}}]}{p[s(t)|\mathcal{H}_0]} = \frac{p_{\rm noise}[s(t)-h(t;\boldsymbol{\theta})]}{p_{\rm noise}[s(t)]}\,. \tag{3.19}$$

We refer $\boldsymbol{\theta}_{\rm ML}$ to the value of $\boldsymbol{\theta}$ which maximizes $\tilde{\Lambda}$, i.e.,

$$\boldsymbol{\theta}_{\rm ML} := \arg\max_{\boldsymbol{\theta}} \tilde{\Lambda}[s(t);\boldsymbol{\theta}] = \arg\max_{\boldsymbol{\theta}} p[s(t)|\mathcal{H}_{\boldsymbol{\theta}}]\,. \tag{3.20}$$

It is called the maximum likelihood estimator and is used as the predicted value of the parameter. Here, we treat the amplitude A separately. The waveform model is rewritten as

$$h(t;\boldsymbol{\theta}) = Ah_0(t;\boldsymbol{\Theta})\,, \tag{3.21}$$

with $\boldsymbol{\theta} = \boldsymbol{\Theta} \cup \{A\}$. From Eq. (3.20), the maximum likelihood estimator of the amplitude is calculated by

$$\frac{\partial}{\partial A}\tilde{\Lambda}[s(t);A,\boldsymbol{\Theta}]\bigg|_{A_{\rm ML}} = 0\,. \tag{3.22}$$

Solving this for $A_{\rm ML}$, we get

$$A_{\rm ML} = \frac{(s|h_0(\boldsymbol{\Theta}))}{(h_0(\boldsymbol{\Theta})|h_0(\boldsymbol{\Theta}))}\,. \tag{3.23}$$

Thus, the maximum likelihood estimator of the amplitude is derived analytically. Substituting $A_{\rm ML}$ into Eq. (3.19), we obtain the log likelihood ratio marginalized over the amplitude as

$$\log\tilde{\Lambda}[s(t);A_{\rm ML},\boldsymbol{\Theta}] = A_{\rm ML}(s|h_0(\boldsymbol{\Theta})) - \frac{A_{\rm ML}^2}{2}(h_0(\boldsymbol{\Theta})|h_0(\boldsymbol{\Theta})) = \frac{1}{2}\cdot\frac{(s|h_0(\boldsymbol{\Theta}))^2}{(h_0(\boldsymbol{\Theta})|h_0(\boldsymbol{\Theta}))}\,. \tag{3.24}$$

A marginalized signal-to-noise ratio $\rho_{\rm MF}$ is defined by

$$\rho_{\rm MF}(\boldsymbol{\Theta}) = \frac{(s|h_0(\boldsymbol{\Theta}))}{\sqrt{(h_0(\boldsymbol{\Theta})|h_0(\boldsymbol{\Theta}))}}\,. \tag{3.25}$$

Then, the log likelihood can be written as

$$\log\tilde{\Lambda}[s(t);A_{\rm ML},\boldsymbol{\Theta}] = \frac{1}{2}\rho_{\rm MF}^2(\boldsymbol{\Theta})\,. \tag{3.26}$$

The maximum likelihood estimator $\boldsymbol{\Theta}_{\rm ML}$, which is defined by Eq. (3.20), can also be obtained by

$$\boldsymbol{\Theta}_{\rm ML} = \arg\max_{\boldsymbol{\Theta}} \rho_{\rm MF}^2(\boldsymbol{\Theta}). \tag{3.27}$$

As a remark, we comment the difference between the frequentist approach and the Bayesian approach. The formalism we explained is so-called frequentist approach. In this approach, the parameter $\boldsymbol{\theta}$ is a deterministic number. On the other hand, in Bayesian approach, the parameter $\boldsymbol{\theta}$ is assumed to follow some distribution. Given a prior distribution $\pi(\mathcal{H}_{\boldsymbol{\theta}})$, we can predict the posterior distribution as

$$p[\mathcal{H}_{\boldsymbol{\theta}}|s(t)] = \frac{p[s(t)|\mathcal{H}_{\boldsymbol{\theta}}]\pi(\mathcal{H}_{\boldsymbol{\theta}})}{\int d\boldsymbol{\theta}\; p[s(t)|\mathcal{H}_{\boldsymbol{\theta}}]\pi(\mathcal{H}_{\boldsymbol{\theta}})}\,. \tag{3.28}$$

The parameter value at which the posterior distribution (3.28) has its maximum,

$$\boldsymbol{\theta}_{\rm MAP} := \arg\max_{\boldsymbol{\theta}} p[\mathcal{H}_{\boldsymbol{\theta}}|s(t)]\,, \tag{3.29}$$

is called the maximum posterior estimator. When the signal is large, two estimators, (3.20) and (3.29), are equivalent.

3.1.3 Matched filtering in time domain

Let us assume that the strain data $s(t)$ has a duration T. Over the duration T, the template $h(t)$ is assumed to be almost monochromatic. The inner product Eq. (3.7) can be calculated without carrying out the Fourier transform to the strain $s(t)$ and template $h(t)$ as

$$(s|h) \simeq \frac{2}{S_{\rm n}(f_{\rm temp})} \int_{-T/2}^{T/2} dt\ s(t)h^*(t)\,. \tag{3.30}$$

In the following, the expression (3.30) is used as the inner product.

3.2 Continuous gravitational waves

3.2.1 Waveform model

In this section, we explain the geometrical configuration of the source and the detector and its effect on the observed gravitational wave signal. Because the detector moves with respect to the source, the phase of the gravitational wave we observe gets modulated. If the source is at rest with respect to the solar system barycenter (SSB), the relative motion between the source and the detector is caused by the Earth's rotation and the orbital motion around the Sun. It is convenient to introduce two time coordinates, the SSB time τ and the detector time t, for describing the phase modulation. The SSB time is used for describing the waveform emitted from the source. The waveform is modeled by

$$h^{\rm source}(\tau) = h_0 e^{2\pi i f_{\rm gw}\tau + i\phi_0}\,, \tag{3.31}$$

with the amplitude h_0 and the initial phase ϕ_0. We use the SSB reference frame as a spatial coordinate. In the SSB reference frame, the x-axis points toward the vernal equinox, and the z-axis is parallel to the orbital angular momentum of the Earth. In the SSB reference frame, the unit vector $\boldsymbol{n}(\alpha,\delta)$ pointing towards a point indicating the right ascension α and the declination δ can be written as

$$\boldsymbol{n}(\alpha,\delta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\epsilon & \sin\epsilon \\ 0 & -\sin\epsilon & \cos\epsilon \end{pmatrix} \begin{pmatrix} \cos\alpha\cos\delta \\ \sin\alpha\cos\delta \\ \sin\delta \end{pmatrix}\,, \tag{3.32}$$

where $\epsilon \simeq 23.44^\circ$ is the tilt angle between the ecliptic and the celestial equator. When the source is located at the sky position $(\alpha_{\rm s},\delta_{\rm s})$, the detector time t is defined by

$$\tau =: t + \frac{\boldsymbol{r}(t)\cdot\boldsymbol{n}_{\rm s}}{c}\,, \tag{3.33}$$

where the vector $\boldsymbol{n}_{\rm s} := \boldsymbol{n}(\alpha_{\rm s},\delta_{\rm s})$ is the normal vector pointing from the SSB to the source, and the vector $\boldsymbol{r}(t)$ is the detector's location with respect to the SSB. The vector $\boldsymbol{r}(t)$ can be decomposed into

$$\boldsymbol{r}(t) = \boldsymbol{r}_\odot(t) + \boldsymbol{r}_\oplus(t)\,, \tag{3.34}$$

where $\boldsymbol{r}_\odot(t)$ points from the SSB to the center of the Earth, and $\boldsymbol{r}_\oplus(t)$ is the location of the detector with respect to the center of the Earth. In the SSB reference frame, $\boldsymbol{r}_\odot(t)$ is written as

$$\boldsymbol{r}_\odot(t) = R_{\mathrm{ES}} \begin{pmatrix} \cos(\varphi_\odot + \Omega_\odot t) \\ \sin(\varphi_\odot + \Omega_\odot t) \\ 0 \end{pmatrix} , \tag{3.35}$$

where R_{ES}, $\Omega_\odot$, and $\varphi_\odot$ are the distance between the Earth and the Sun, the angular velocity of the orbital motion, and the initial phase, respectively. The vector $\boldsymbol{r}_\oplus(t)$ can be described by

$$\boldsymbol{r}_\oplus(t) = R_{\mathrm{E}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\epsilon & \sin\epsilon \\ 0 & -\sin\epsilon & \cos\epsilon \end{pmatrix} \cdot \begin{pmatrix} \cos\lambda\cos(\varphi_\oplus + \Omega_\oplus t) \\ \cos\lambda\sin(\varphi_\oplus + \Omega_\oplus t) \\ \sin\lambda \end{pmatrix} , \tag{3.36}$$

with the Earth's radius R_{E}, the latitude of the detector's cite λ, the initial phase $\varphi_\oplus$, and the rotational angular frequency of the Earth $\Omega_\oplus$. From Eqs. (3.31) and (3.33), the modulated waveform can be described as

$$h(t) = h_0 e^{i\Phi(t)} , \tag{3.37}$$

with

$$\Phi(t) = 2\pi f_{\mathrm{gw}} t + 2\pi f_{\mathrm{gw}} \frac{\boldsymbol{r}(t)\cdot\boldsymbol{n}_{\mathrm{s}}}{c} . \tag{3.38}$$

Although the waveform Eq. (3.37) has a complex value, the observed signal has a real value and is affected by the antenna pattern functions. In this thesis, the inclination angle and the polarization angle are set to be zero, i.e.,

$$\iota = 0 , \qquad \psi = 0 , \tag{3.39}$$

for simplicity. The complex-valued antenna pattern function is defined by

$$G(t) := \frac{F_+(t) + iF_\times(t)}{2} , \tag{3.40}$$

where

$$\begin{aligned} F_+(t) = {} & \frac{1}{16}\sin 2\gamma(3-\cos 2\lambda)(3-\cos 2\delta)\cos(2(\alpha-\varphi_\oplus-\Omega_\oplus t)) \\ & - \frac{1}{4}\cos 2\gamma\sin\lambda(3-\cos 2\delta)\sin(2(\alpha-\varphi_\oplus-\Omega_\oplus t)) \\ & + \frac{1}{4}\sin 2\gamma\sin 2\lambda\sin 2\delta\cos(\alpha-\varphi_\oplus-\Omega_\oplus t) \\ & - \frac{1}{2}\cos 2\gamma\cos\lambda\sin 2\delta\sin(\alpha-\varphi_\oplus-\Omega_\oplus t) \\ & + \frac{3}{4}\sin 2\gamma\cos^2\lambda\cos^2\delta , \end{aligned} \tag{3.41}$$

and

$$\begin{aligned} F_\times(t) = \; & \cos 2\gamma \sin\lambda \sin\delta \cos(2(\alpha - \varphi_\oplus - \Omega_\oplus t)) \\ & + \frac{1}{4}\sin 2\gamma (3 - \cos 2\lambda)\sin\delta \sin(2(\alpha - \varphi_\oplus - \Omega_\oplus t)) \\ & + \cos 2\gamma \cos\lambda \cos\delta \cos(\alpha - \varphi_\oplus - \Omega_\oplus t) \\ & + \frac{1}{2}\sin 2\gamma \sin 2\lambda \cos\delta \sin(\alpha - \varphi_\oplus - \Omega_\oplus t)\,. \end{aligned} \tag{3.42}$$

(see Appendix. A). The observed signal $h_{\rm obs}(t)$ (Eq. (2.123)) is obtained as

$$h_{\rm obs}(t) = G(t)h(t) + G^*(t)h^*(t)\,. \tag{3.43}$$

Taking into account the detector noise $n(t)$, we can describe the strain data by

$$s(t) = h_{\rm obs}(t) + n(t)\,. \tag{3.44}$$

We assume that the detector noise is stationary and Gaussian and the strain data has no gaps in time.

3.2.2 Time resampling technique

The time resampling technique [39] enables us to rapidly search the predicted value of the continuous gravitational wave frequency $f_{\rm gw}$, when the source location is specified. Here, we assume that the source location is known a priori. By inverting the relation Eq. (3.33), the demodulation of the signal can be done precisely enough that the monochromatic waveform (Eq. (3.31)) can be used as the template of the matched filtering. For the time-domain matched filtering (3.30), the SNR is obtained by

$$\rho_{\rm MF}(f_{\rm gw}) = \frac{(s|h(f_{\rm gw}))}{\sqrt{(h(f_{\rm gw})|h(f_{\rm gw}))}} \simeq \sqrt{\frac{2h_0}{S_{\rm n}(f_{\rm gw})T}} \int_{-T/2}^{T/2} d\tau \; s(t(\tau))e^{-2\pi i f_{\rm gw}\tau}\,. \tag{3.45}$$

This has the same form as that of the Fourier transform of $s(t(\tau))$. Therefore, the rapid algorithm, i.e., the Fast Fourier Transform (FFT), can be employed, and the search over the gravitational wave frequency $f_{\rm gw}$ can be performed by a single Fourier transformation.

3.2.3 Difficulty in all-sky search of continuous gravitational waves

Let us consider the all-sky search of monochromatic gravitational waves. Because we don't know the source location, we need to distribute a lot of grid points to cover the whole sky. Before applying the matched filtering, we introduce the new time coordinate ζ as

$$\zeta := t + \frac{\boldsymbol{r}(t)\cdot\boldsymbol{n}_{\rm g}}{c}\,, \tag{3.46}$$

with a chosen grid point $\boldsymbol{n}_{\rm g}$. If the grid point $\boldsymbol{n}_{\rm g}$ is sufficiently close to the source direction, the signal is almost demodulated. Then, the matched filtering with the monochromatic waveforms is applicable. On the other hand, if the grid point $\boldsymbol{n}_{\rm g}$ largely deviates from the source location $\boldsymbol{n}_{\rm s}$, the SNR will be drastically reduced. In order to reduce the false dismissal probability, a lot of grid points need to be placed so that the maximum loss of the SNR is suppressed below a certain acceptable level. The number of grid points $N_{\rm grid,coh}$ is roughly determined by the angular resolution $(\delta\theta)_{\rm coh}$. The angular resolution is estimated by the ratio between the wavelength of the gravitational waves $\lambda_{\rm gw}$ and the diameter of the Earth's orbit around the Sun. Therefore, we obtain

$$(\delta\theta)_{\rm coh} \sim \frac{\lambda_{\rm gw}}{2R_{\rm ES}} \sim 10^{-5}\ [\rm rad]\,, \tag{3.47}$$

where the frequency of the gravitational wave is assumed to be 100 Hz. Hence, the necessary number of the grid points is

$$N_{\rm grid,coh} \sim \frac{4\pi}{(\delta\theta)^2_{\rm coh}} \sim 1.3 \times 10^{11}\,. \tag{3.48}$$

As we see the previous subsection, the time-domain matched filtering between the time-resampled strain and a monochromatic waveform is equivalent to the Fourier transform of the time-resampled strain. Therefore, the computational cost of the all-sky coherent search is dominated by the cost of the Fourier transform. The required number of floating point operations is estimated by

$$\mathcal{N}_{\rm coh} \sim N_{\rm grid,coh} \cdot \mathcal{N}_{\rm FFT,coh}\,, \tag{3.49}$$

where $\mathcal{N}_{\rm FFT,coh}$ is the number of floating point operations required for the Fourier transfrom. For the strain data of the duration 10^7 sec and the sampling rate 1024 Hz, the number of data can be estimated by

$$N_{\rm data,coh} \sim 10^{10}\,, \tag{3.50}$$

and the number of the floating point operations required for carrying out FFT is

$$\mathcal{N}_{\rm FFT,coh} \sim 5N_{\rm data,coh}\log_2 N_{\rm data,coh} \sim 1.7 \times 10^{12}\,. \tag{3.51}$$

Therefore, the required number of floating point operations is estimated by

$$\mathcal{N}_{\rm coh} \sim 2.2 \times 10^{23}\,. \tag{3.52}$$

Even if computational resource with 1TFlops is available, the computational time becomes $\sim 2.2 \times 10^{11}$ sec. Therefore, it is unrealistic to employ a naive coherent method with the time resampling even for the simplest case of monochromatic sources (see [***] for alternative coherent method).

Chapter 4

New method of all-sky searches for continuous gravitational waves

4.1 Overview

In this chapter, we explain and assess our method of all-sky search of continuous gravitational waves. Figure. 4.1 shows the workflow of our detection method. In the first step, we remove the phase modulation due to the Earth's rotation by the time resampling with the grid points placed coarser than that of the coherent matched filtering (Sec. 4.2). The short-time Fourier transform is applied to the partially demodulated strains. Regarding the short-time Fourier transform of each frequency bin as a time series data, we take the Fourier transform. By doing this, we can prevent the dispersion of the signal power (Sec. 4.3). Although a similar idea of the double Fourier transform is already proposed in Refs. [40, 41], this is the first time to incorporate it in a detection algorithm. The excess power method is employed to select the candidates (Sec. 4.4), and the deep learning method restricts the source's possible location (Sec. 4.5). As a follow-up analysis, the coherent matched filtering is carried out for the selected candidates (Sec. 4.6). We examine the computational cost and the detection ability of our method in Sec. 4.7.

4.2 Subtracting the effect due to the Earth's rotation

We employ the time resampling technique to subtract the modulation caused by the Earth's rotation. By using a resampled time coordinate ζ with a representative grid point $\boldsymbol{n}_{\rm g}$, the phase can be rewritten as

$$\Phi(t) = 2\pi f_{\rm gw} t + \Phi_{\oplus}(t) + \Phi_{\odot}(t) = 2\pi f_{\rm gw}\zeta + \delta\Phi_{\oplus}(t) + \delta\Phi_{\odot}(t)\,, \tag{4.1}$$

where

$$\delta\Phi_{\oplus}(t) := 2\pi f_{\rm gw}\frac{\boldsymbol{r}_{\oplus}(t)\cdot\Delta\boldsymbol{n}}{c}\,, \tag{4.2}$$

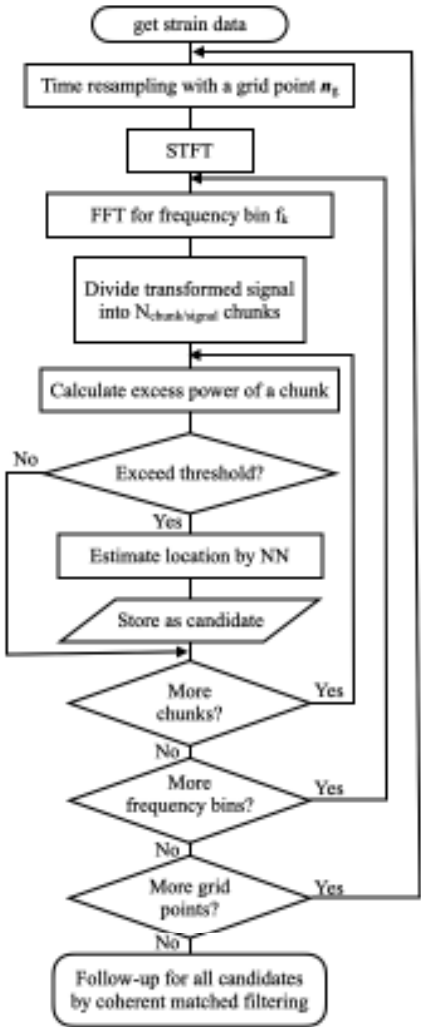

Figure 4.1: Flowchart of our method.

$$\delta\Phi_\odot(t) := 2\pi f_{\rm gw} \frac{\boldsymbol{r}_\odot(t)\cdot\Delta\boldsymbol{n}}{c}\,, \tag{4.3}$$

and $\Delta\boldsymbol{n} := \boldsymbol{n}_{\rm s} - \boldsymbol{n}_{\rm g}$ is the deviation between the directions of the source location and the representative grid point. Setting a threshold value $\delta\Phi_\varepsilon$, we place the grid points so that the condition,

$$\min_{\boldsymbol{n}_{\rm g}} \max_t |\delta\Phi_\oplus(t)| < \delta\Phi_\varepsilon\,, \tag{4.4}$$

is satisfied for any arbitrary source location. Here, we expand the residual phase up to the first order of $\Delta\alpha := \alpha_{\rm s} - \alpha_{\rm g}$ and $\Delta\delta := \delta_{\rm s} - \delta_{\rm g}$. Then, we get

$$\delta\Phi_\oplus(t) \simeq \frac{2\pi f_{\rm gw}}{c} R_{\rm E}\cos\lambda\,[-\Delta\delta\sin\delta_{\rm g}\cos(\alpha_{\rm g}-\varphi_\oplus-\Omega_\oplus t) - \Delta\alpha\cos\delta_{\rm g}\sin(\alpha_{\rm g}-\varphi_\oplus-\Omega_\oplus t)]\,, \tag{4.5}$$

where the constant term independent of $\Delta\alpha$ and $\Delta\delta$ is neglected because it degenerates with the initial phase ϕ_0. Therefore,

$$\max_t |\delta\Phi_\oplus(t)| = \frac{2\pi f_{\rm gw}}{c} R_{\rm E} \cos\lambda \sqrt{(\Delta\delta)^2 \sin^2\delta_{\rm g} + (\Delta\alpha)^2 \cos^2\delta_{\rm g}}\,. \tag{4.6}$$

The residual phase (4.6) is invariant under the transformation $\delta_{\rm g} \to -\delta_{\rm g}$. Therefore, the grid points on the plane of negative $\delta_{\rm g}$ can be generated by flipping the sign of the grid points on the plane of positive $\delta_{\rm g}$. In the following, we consider how to place the grid points covering only a half plane of $0 \le \delta_{\rm g} \le \pi/2$.

In the vicinity of the pole $\delta_{\rm g} = \pi/2$, the residual phase (4.6) does not depend on $\alpha_{\rm g}$. Owing to this fact, the vicinity of the pole can be covered by a single grid. At $\delta_{\rm g} = \pi/2$, the residual phase (4.6) becomes

$$\max_t |\delta\Phi_\oplus(t)| = \frac{2\pi f_{\rm gw}}{c} R_{\rm E} |\Delta\delta| \cos\lambda\,. \tag{4.7}$$

Because the condition (4.4) gives

$$|\Delta\delta| = \left|\delta_{\rm s} - \frac{\pi}{2}\right| \le \delta\Phi_\varepsilon \times \frac{c}{2\pi f_{\rm gw}} \frac{1}{R_{\rm E}\cos\lambda}\,, \tag{4.8}$$

the grid point $(\alpha_{\rm g}, \delta_{\rm g}) = (0, \pi/2)$ can cover the region

$$\{(\alpha,\delta)| -\pi \le \alpha \le \pi,\ \delta_1 \le \delta \le \pi/2\}\,,$$

where

$$\delta_1 := \frac{\pi}{2} - \delta\Phi_\varepsilon \times \frac{c}{2\pi f_{\rm gw}} \frac{1}{R_{\rm E}\cos\lambda}\,. \tag{4.9}$$

The region

$$\{(\alpha,\delta)| -\pi \le \alpha \le \pi,\ 0 \le \delta \le \delta_1\} \tag{4.10}$$

is covered by plural patches. The 2-dimensional metric $d\sigma^2$ corresponding to the residual phase (4.6) is introduced by

$$d\sigma^2 = \cos^2\delta d\alpha^2 + \sin^2\delta d\delta^2\,. \tag{4.11}$$

Defining the new variables

$$X := \alpha,\ Y := -\log|\cos\delta|\,, \tag{4.12}$$

we can transform the metric (4.11) into a conformally flat metric as

$$d\sigma^2 = e^{-2Y}(dX^2 + dY^2)\,. \tag{4.13}$$

In the space that admits a conformally flat metric, the contour becomes a circle. Ref. [42] provides an efficient placement of grid points on 2-dimensional parameter space. By using

the coordinate transformation (4.12), the region given by Eq. (4.10) is transformed into the rectangular region on (X, Y)-plane,

$$\{(X, Y) \mid X_{\rm min} \leq X \leq X_{\rm max},\ Y_{\rm min} \leq Y \leq Y_{\rm max}\}\ , \tag{4.14}$$

with $X_{\rm min} = -\pi$, $X_{\rm max} = \pi$, $Y_{\rm min} = -\log\cos\delta_1$, and $Y_{\rm max} = 0$. Because the conformal factor of the metric (4.13) depends only on Y, circles placed on the line $Y =$ constant have the same radius. At first, we consider the placement of circles along the line $Y =$ constant and start from the vicinity of the corner point $(X_{\rm max}, Y_{\rm max})$. Figure 4.2 shows the schematic picture of template placement in the vicinity of the line $Y = Y_{\rm max}$. Two circles centered at the point (x_{11}, y_1) and the point $(x_{12}, y_1) = (x_{11} - \Delta x_1, y_1)$ will be placed to cover a part of the region around the line $Y = Y_{\rm max}$. These circles have a radius

$$r_1^2 = e^{2y_1} d\sigma_{\rm max}^2\,, \tag{4.15}$$

where

$$d\sigma_{\rm max}^2 = \left(\frac{c}{2\pi f_{\rm gw}} \cdot \frac{\delta\Phi_\epsilon}{R_{\rm E}\cos\lambda}\right)^2\,, \tag{4.16}$$

(see Eq. (4.6).) The intersections of two circles are denoted by $(x_{11} - \Delta x_1/2, y_1 \pm d)$ with

$$d(r_1, \Delta x_1) = \sqrt{r_1^2 - \left(\frac{\Delta x_1}{2}\right)^2}\,. \tag{4.17}$$

Two circles are placed so that the upper intersection $(x_{11} - \Delta x_1/2, y_1 + d)$ lies on the line $Y = Y_{\rm max}$. Therefore,

$$y_1 = Y_{\rm max} - d(r_1, \Delta x_1)\,. \tag{4.18}$$

To determine the parameter Δx_1, we impose the condition that the area of a square formed by the two intersections and the origins of the neighbouring circles are maximized. The condition can be written as

$$\Delta x_1 = \arg\max_{\Delta x'} \Delta x' \cdot d(r_1, \Delta x')\,. \tag{4.19}$$

Using Eq. (4.17), we obtain

$$\Delta x_1 = \sqrt{2} r_1\,, \tag{4.20}$$

and

$$d(r_1, \Delta x_1) = \frac{r_1}{\sqrt{2}}\,. \tag{4.21}$$

With Eqs. (4.15) and (4.18), the relation

$$r_1^2 = \exp\left[2Y_{\rm max} - \sqrt{2} r_1\right] d\sigma_{\rm max}^2\,, \tag{4.22}$$

is valid, and we determine r_1 by solving this relation numerically. As a result, the location of two circles are determined by

$$(x_{11}, y_1) = (X_{\rm max} - \Delta x_1/2,\ Y_{\rm max} - d(r_1, \Delta x_1)) = \left(X_{\rm max} - \frac{r_1}{\sqrt{2}},\ Y_{\rm max} - \frac{r_1}{\sqrt{2}}\right)\,, \tag{4.23}$$

and

$$(x_{12}, y_1) = \left(X_{\rm max} - \frac{3r_1}{\sqrt{2}},\ Y_{\rm max} - \frac{r_1}{\sqrt{2}} \right) . \tag{4.24}$$

Carrying out the above procedures recursively, we can obtain the origin of the n-th circle as

$$(x_{1n}, y_1) = \left(X_{\rm max} - \frac{(2n-1)r_1}{\sqrt{2}},\ Y_{\rm max} - \frac{r_1}{\sqrt{2}} \right) . \tag{4.25}$$

The number of circles lying along the line $Y = Y_{\rm max}$ is

$$N_1 = \left\lceil \frac{X_{\rm max} - X_{\rm min}}{\Delta x_1} \right\rceil , \tag{4.26}$$

where $\lceil x \rceil$ is the minimum integer larger than x.

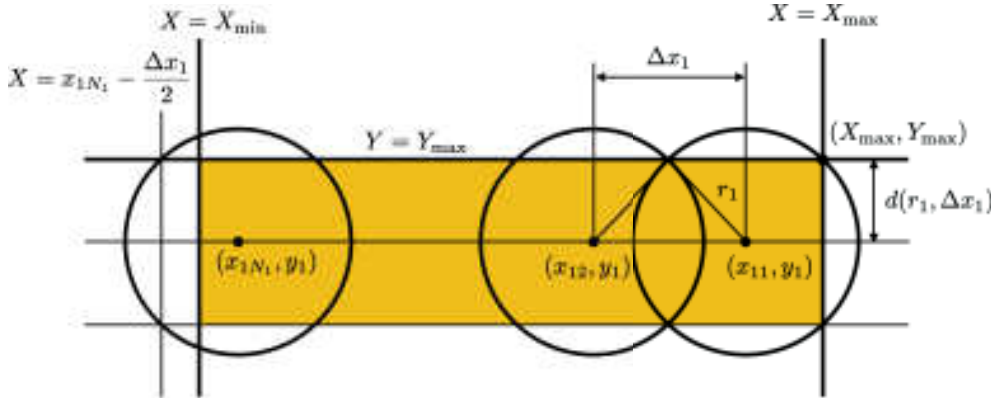

Figure 4.2: Image of grids placement on 2-dimensional parameter space. Each origin of a circle corresponds to each grid point. The orange region indicates the whole region covered by N_1 circles.

Next, we consider the second line $Y = Y_1 =$ constant (see Fig. 4.3). The constant Y_1 should be set to

$$Y_1 := Y_{\rm max} - \frac{r_1}{\sqrt{2}} - d(r_1, \Delta x_1) = Y_{\rm max} - \sqrt{2} r_1 . \tag{4.27}$$

The first circle on the second line is centered at

$$(x_{21}, y_2) = \left(X_{\rm max} - \frac{r_2}{\sqrt{2}},\ Y_1 - \frac{r_2}{\sqrt{2}} \right) , \tag{4.28}$$

where the radius of the circle r_2 is determined by

$$r_2^2 = \exp\left[2Y_1 - \sqrt{2} r_2 \right] d\sigma_{\rm max}^2 . \tag{4.29}$$

The distance between the origins of neighbouring circles is determined by

$$\Delta x_2 = \sqrt{2} r_2 . \tag{4.30}$$

The number of circles lying along the line $Y = Y_1$ is

$$N_2 = \left\lceil \frac{X_{\max} - X_{\min}}{\Delta x_2} \right\rceil . \tag{4.31}$$

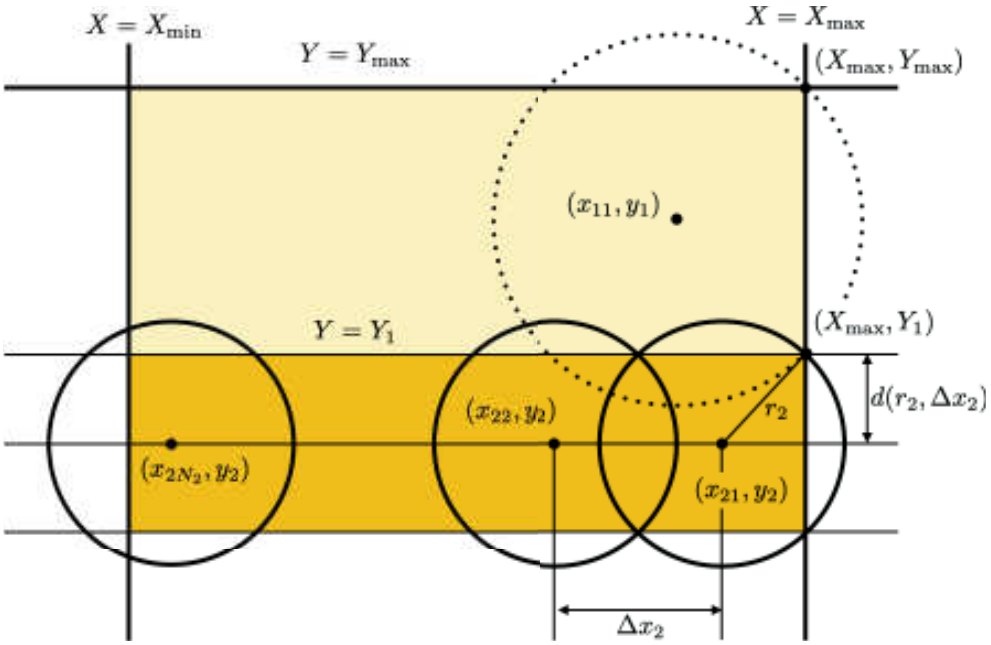

Figure 4.3: Template spacing along the line $Y = Y_1$. The bright orange region indicates the region already covered by circles lying on the line $Y = Y_{\max}$. The dark orange region is the region to be covered by circles lying on the line $Y = Y_1$.

These procedures are repeated until the whole region shown in Eq. (4.14) are completely covered. The total number of circles to cover the half of the sky except for the vicinities of two poles is

$$N_{\text{grid/half}} = \sum_{m=1}^{M} N_m \,, \tag{4.32}$$

where M is the minimum integer satisfying the condition

$$y_M - d(r_M, \Delta x_M) = y_M - \frac{r_M}{\sqrt{2}} \leq Y_{\min} \,. \tag{4.33}$$

Finally, the total number of grid points on the sky is obtained as

$$N_{\text{grid}} = 2 \times N_{\text{grid/half}} + 2 \,. \tag{4.34}$$

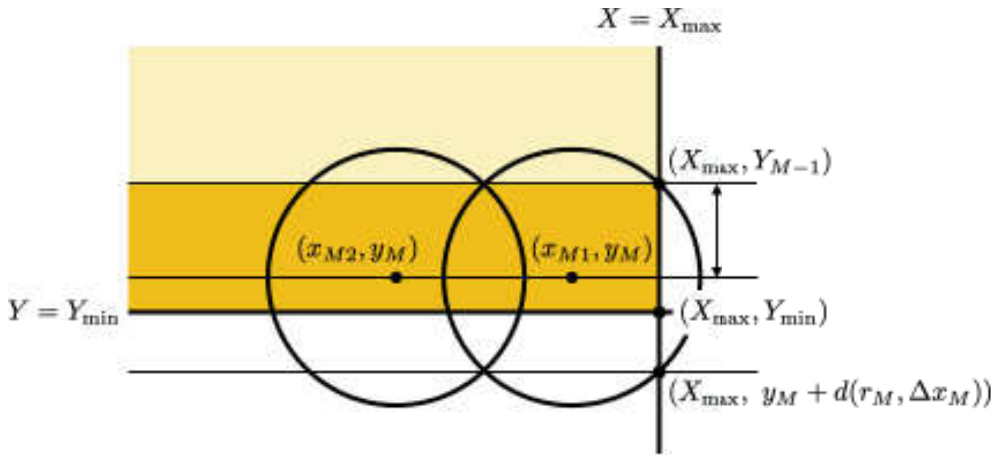

Figure 4.4: M-th sequence of grid points. The thick region is covered by circles centered at (x_{Mi}, y_M) for $i = 1, 2, \cdots, N_M$. The brighter region is already covered by other grid points. No more grid points is required because the rectangular region is completely covered.

In this work, we set the value of the threshold $\delta\Phi_\varepsilon$ as

$$\delta\Phi_\varepsilon = 0.058. \tag{4.35}$$

For $f_{\rm gw} = 100$ Hz, the total number of grid points to cover the entire sky is

$$N_{\rm grid} = 352,436\,. \tag{4.36}$$

Figure. 4.5 shows an enlarged view of the grid points.

4.3 Double Fourier transform

After subtracting the modulation due to the Earth's rotation, only the effect of the Earth's orbital motion remains. Assuming the subtraction error is negligibly small, we obtain

$$\Phi(t) = 2\pi f_{\rm gw}\zeta + \delta\Phi_\odot(t)\,, \tag{4.37}$$

with an optimal grid point. After the time resampling, we obtain the time-resampled strain data

$$s(\zeta) = h_{\rm obs}(\zeta) + n(\zeta)\,. \tag{4.38}$$

In the following, we consider only the time-resampled data. Therefore, we can denote each term in Eq. (4.38) by the same character without confusion. Before moving to the next process, we evaluate how large the phase modulation $\delta\Phi_\odot$ is. With the set of grid points explained in the previous section, the norm of the deviation vector $|\Delta\boldsymbol{n}|$ is estimated by

$$|\Delta\boldsymbol{n}| \sim \frac{c}{2\pi f_{\rm gw} R_{\rm E}}\delta\Phi_\epsilon\,, \tag{4.39}$$

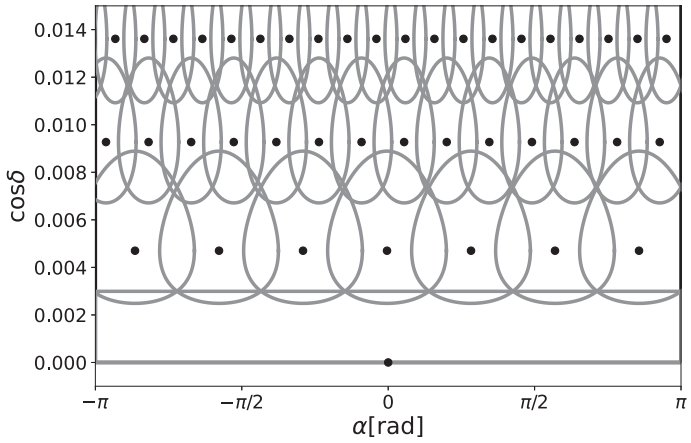

Figure 4.5: Enlarged view of the grid points placed on the sky. Black dots are grid points and gray lines are contours of $\delta\Phi_\oplus$. The region $\{(\alpha,\delta) \mid \delta_1 \le \delta \le \pi/2\}$ is covered by the single grid point $(\alpha,\delta) = (0.0, \pi/2)$.

when the source location is near the reference grid point. Therefore, the

$$\delta\Phi_\odot(t) \sim \frac{2\pi f_{\rm gw} R_{\rm ES}}{c} |\Delta \boldsymbol{n}| \cos(\Omega_\odot t) \sim \frac{R_{\rm ES}}{R_{\rm E}} \delta\Phi_\epsilon \cos(\Omega_\odot t)\,. \tag{4.40}$$

This phase modulation induces the frequency modulation as

$$\delta f_\odot \sim \frac{1}{2\pi}\left|\frac{d\delta\Phi_\odot}{dt}\right| \sim \frac{R_{\rm ES}\Omega_\odot}{2\pi R_{\rm E}}\delta\Phi_\epsilon \sim 3.9\times10^{-5}\ {\rm Hz}\,. \tag{4.41}$$

First, we carry out the short-time Fourier transform (STFT) to the resampled strain (4.38). The resampled strain data $s(\zeta)$ is divided into $N_{\rm seg}$ segments having the duration $T_{\rm seg}$. The start time of each segment is $\zeta_j := jT_{\rm seg}$ for $j = 0, 1, \cdots N_{\rm seg} - 1$. Before carrying out the Fourier transform to each segment, the window function $w(\zeta)$ is applied to suppress fictitious high frequency components. In this thesis, the tukey window function, which is defined by

$$w(\zeta) = \begin{cases} \frac{1}{2} - \frac{1}{2}\cos\left(\frac{2\pi\zeta}{\alpha_{\rm tw}T_{\rm seg}}\right)\,, & \left(0 \le \frac{\zeta}{T_{\rm seg}} < \frac{\alpha_{\rm tw}}{2}\right)\,, \\ 1\,, & \left(\frac{\alpha_{\rm tw}}{2} \le \frac{\zeta}{T_{\rm seg}} \le 1 - \frac{\alpha_{\rm tw}}{2}\right)\,, \\ \frac{1}{2} - \frac{1}{2}\cos\left(\frac{2\pi(T_{\rm seg}-\zeta)}{\alpha_{\rm tw}T_{\rm seg}}\right)\,, & \left(1 - \frac{\alpha_{\rm tw}}{2} < \frac{\zeta}{T_{\rm seg}} \le 1\right)\,, \end{cases} \tag{4.42}$$

is employed. Here, the parameter $\alpha_{\rm tw}$ characterizes the steepness of the edge of the window, and we set $\alpha_{\rm tw} = 1/8$. The output of the STFT is denoted by

$$s^{\rm STFT}_{j,k} = h^{\rm STFT}_{j,k} + n^{\rm STFT}_{j,k}\,, \tag{4.43}$$

where

$$h_{j,k}^{\mathrm{STFT}} = \frac{1}{T_{\mathrm{seg}}} \int_{\zeta_j}^{\zeta_j + T_{\mathrm{seg}}} d\zeta' \; w(\zeta' - \zeta_j) h_{\mathrm{obs}}(\zeta') e^{-2\pi i f_k \zeta'} \,, \tag{4.44}$$

$$n_{j,k}^{\mathrm{STFT}} = \frac{1}{T_{\mathrm{seg}}} \int_{\zeta_j}^{\zeta_j + T_{\mathrm{seg}}} d\zeta' \; w(\zeta' - \zeta_j) n(\zeta') e^{-2\pi i f_k \zeta'} \,, \tag{4.45}$$

and $f_k := k\Delta f = k/T_{\mathrm{seg}}$ is the frequency of the k-th bin. We focus on the positive frequency modes, $f_k > 0$. Then, the second term of Eq. (3.43) can be neglected and the STFT of $h(\zeta)$ can be approximated by

$$h_{j,k}^{\mathrm{STFT}} \simeq \frac{1}{T_{\mathrm{seg}}} \int_{\zeta_j}^{\zeta_j + T_{\mathrm{seg}}} d\zeta' \; w(\zeta' - \zeta_j) G(t(\zeta')) e^{2\pi i \delta f_k \zeta'} e^{i\delta\Phi_\odot(\zeta')} \,. \tag{4.46}$$

Because the difference between $\Omega_\oplus \zeta$ and $\Omega_\oplus t(\zeta)$ is negligibly small, we can replace $G(t(\zeta'))$ with $G(\zeta')$. The duration of each segment T_{seg} is chosen so that $G(\zeta)$ can be regarded as a constant in each segment. In this work, we choose

$$T_{\mathrm{seg}} = 32 \text{ sec} \,. \tag{4.47}$$

With this choice of T_{seg}, the factor related with the residual phase, $e^{i\delta\Phi_\odot}$, can also be treated as a constant in each segment. Therefore, Eq (4.46) can be approximated by

$$h_{j,k}^{\mathrm{STFT}} \simeq h_0 e^{i\delta\Phi_\odot(\zeta_j)} G(\zeta_j) W_k(\zeta_j) \,, \tag{4.48}$$

with

$$W_k(\zeta_j) := \frac{1}{T_{\mathrm{seg}}} \int_{\zeta_j}^{\zeta_j + T_{\mathrm{seg}}} d\zeta' \; w(\zeta' - \zeta_j) e^{2\pi i \delta f_k \zeta'} \,. \tag{4.49}$$

Substituting Eq. (4.42), we obtain

$$W_k(\zeta_j) = e^{2\pi i \delta f_k \zeta_j} \times \frac{(1 + e^{i\pi\alpha\beta_k})(1 - e^{2\pi i \beta_k (1-\alpha/2)})}{4\pi i \beta_k (\alpha^2 \beta_k^2 - 1)} \,, \tag{4.50}$$

with

$$\beta_k := \delta f_k T_{\mathrm{seg}} \,. \tag{4.51}$$

The factor $e^{i\delta\Phi_\odot(\zeta_j)}$ can be expanded by virtue of the Jacobi-Anger expansion,

$$e^{iz\cos\theta} = \sum_{\ell=-\infty}^{\infty} i^\ell J_\ell(z) e^{i\ell\theta} \,, \tag{4.52}$$

where $J_\ell(z)$ is the Bessel function of the first kind. Using Eq. (4.52), we obtain

$$e^{i\delta\Phi_\odot(\zeta_j)} = \sum_{\ell=-\infty}^{\infty} i^\ell J_\ell(X) e^{i\ell\Omega_\odot\zeta_j} e^{i\ell(\varphi_\odot - \phi_X)} \,, \tag{4.53}$$

with

$$X := \frac{2\pi f_{\rm gw}}{c} R_{\rm E} \sqrt{(\Delta n_x)^2 + (\Delta n_y)^2}\,, \tag{4.54}$$

and

$$e^{i\phi_X} := \frac{1}{\sqrt{(\Delta n_x)^2 + (\Delta n_y)^2}} (\Delta n_x + i \Delta n_y)\,. \tag{4.55}$$

Equation (4.46) can be expressed as

$$h^{\rm STFT}_{j,k} \simeq h_0 G(\zeta_j) W_k(0) e^{2\pi i \delta f_k \zeta_j} \sum_{\ell=-\infty}^{\infty} i^\ell J_\ell(X) e^{i\ell(\varphi_\odot - \phi_X)} e^{i\ell\Omega_\odot \zeta_j}\,. \tag{4.56}$$

Because the frequency modulation $\delta f_\odot$ given in Eq. (4.41) is smaller than the frequency resolution $\Delta f = 1/T_{\rm seg} \sim 3.1 \times 10^{-2}$Hz, the signal is expected to be confined into a single frequency bin. Therefore, we can analyze each frequency bin of STFT independently. After applying STFT, we take the Fourier transform to the STFT output $s^{\rm STFT}_{j,k}$ with a fixed k and we refer to the output of second Fourier transform as *ℓ-domain strain*. The Fourier transform of $s^{\rm STFT}_{j,k}$ is denoted by $\mathsf{S}_{\ell,k}$, i.e.,

$$\mathsf{S}_{\ell,k} = \frac{1}{N_{\rm seg}} \sum_{j=0}^{N_{\rm seg}-1} s^{\rm STFT}_{j,k} e^{-2\pi i j \ell / N_{\rm seg}}\,. \tag{4.57}$$

In the same manner, $\mathsf{H}_{\ell,k}$ and $\mathsf{N}_{\ell,k}$ are defined by the followings,

$$\mathsf{H}_{\ell,k} = \frac{1}{N_{\rm seg}} \sum_{j=0}^{N_{\rm seg}-1} h^{\rm STFT}_{j,k} e^{-2\pi i j \ell / N_{\rm seg}}\,, \tag{4.58}$$

and

$$\mathsf{N}_{\ell,k} = \frac{1}{N_{\rm seg}} \sum_{j=0}^{N_{\rm seg}-1} n^{\rm STFT}_{j,k} e^{-2\pi i j \ell / N_{\rm seg}}\,. \tag{4.59}$$

They are clearly related by

$$\mathsf{S}_{\ell,k} = \mathsf{H}_{\ell,k} + \mathsf{N}_{\ell,k}\,. \tag{4.60}$$

The advantage of taking another Fourier transform is that the signal power can be made concentrated in a limited number of data points. From Eq. (4.56), the ℓ-domain signal $\mathsf{H}_{\ell,k}$ is approximated by

$$\mathsf{H}_{\ell,k} \simeq h_0 W_k(0) i^{\ell'} J_{\ell'}(X) e^{i\ell'(\varphi_\odot - \phi_X)}\,, \tag{4.61}$$

with $\ell' \sim \ell + 2\pi\Omega_\odot^{-1}\delta f_k$. Assuming that the duration of the original strain data is 1 year, an ℓ-domain strain has $N_{\rm seg} \sim 1{\rm yr}/T_{\rm seg} \sim O(10^6)$ data points (see Eq. (4.47)). Because $J_\ell(X) \simeq 0$ for $|X| \lesssim |\ell|$ and $X \lesssim O(10^3)$ for $f_{\rm gw} = 100$ Hz, the gravitational wave signal is concentrated into only few thousand bins in ℓ-domain. Although the antenna pattern function blurs the localization of the signal power, the qualitative feature does not change.

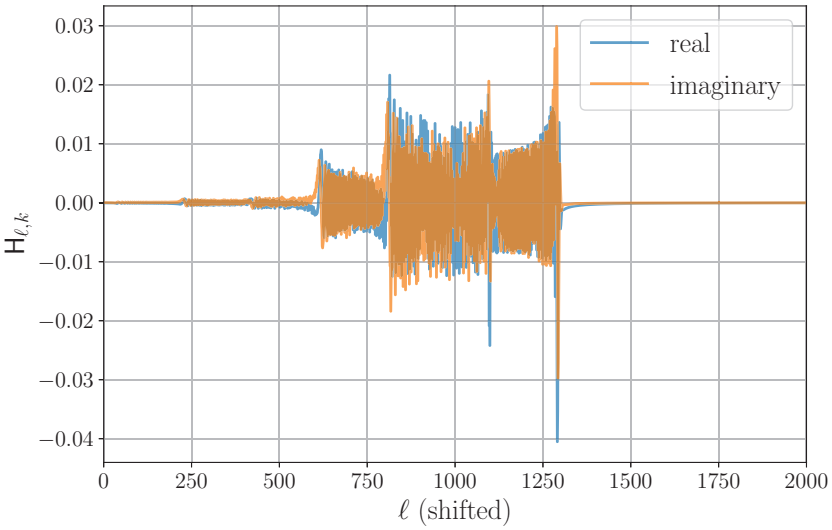

Figure 4.6: Example of an ℓ-domain signal. The width of the signal is almost $O(10^3)$.

Figure. 4.6 shows an example of an ℓ-domain signal. One can find an ℓ-domain signal is localized within a few thousand bins of ℓ-domain.

Before moving to the next section, let us examine the statistical properties of noise in ℓ-domain. The variance of the STFT of the noise can be calculated from the definition of the spectral density (3.3). Taking into account the window function, we obtain

$$\langle (n^{\mathrm{STFT}}_{j,k})(n^{\mathrm{STFT}}_{j',k'})^* \rangle = \frac{1}{2T_{\mathrm{seg}}} \sum_{r=0}^{N-1} S_{\mathrm{n}}(f_r) \tilde{w}[k-r] \tilde{w}^*[k'-r] \delta_{jj'} \,, \tag{4.62}$$

where

$$\tilde{w}[k] := \frac{1}{N} \sum_{m=0}^{N-1} w(m\Delta T) e^{-2\pi i k m/N} \,, \tag{4.63}$$

and N is the number of the data points contained in one segment. Here, we need care about the factor T_{seg}^{-1} existing in the definition of $n^{\mathrm{STFT}}_{j,k}$ (see Eq. (4.45)). The variance of $\mathsf{N}_{\ell,k}$ is

$$\langle \mathsf{N}_{\ell,k} \mathsf{N}^*_{\ell',k'} \rangle = \frac{1}{2T_{\mathrm{seg}} N_{\mathrm{seg}}} \sum_{r=0}^{N-1} S_{\mathrm{n}}(f_r) \tilde{w}[k-r] \tilde{w}^*[k'-r] \delta_{\ell\ell'} \,. \tag{4.64}$$

Because the Fourier transform of the window function has a sharp peak at zero frequency, we can approximate

$$\langle \mathsf{N}_{\ell,k} \mathsf{N}^*_{\ell',k'} \rangle = \frac{S_{\mathrm{n}}(f_k)}{2T_{\mathrm{seg}} N_{\mathrm{seg}}} \sum_{r=0}^{N-1} |\tilde{w}[k-r]|^2 \delta_{kk'} \delta_{\ell\ell'} =: \frac{S_{\mathrm{n}}(f_k)}{2T_{\mathrm{seg}} N_{\mathrm{seg}}} \mathcal{W} \delta_{kk'} \delta_{\ell\ell'} \,, \tag{4.65}$$

where $\mathcal{W}$ is the power of the window function. For the case of the tukey window, we get

$$\mathcal{W} = 0.921875\,. \tag{4.66}$$

For convenience, we introduce $\tilde{\sigma}_k$ by

$$\tilde{\sigma}_k^2 := \frac{S_{\rm n}(f_k)}{T_{\rm seg} N_{\rm seg}} \mathcal{W}\,, \tag{4.67}$$

and obtain

$$\langle \mathsf{N}_{\ell,k} \mathsf{N}^*_{\ell',k} \rangle = \frac{1}{2} \tilde{\sigma}_k^2 \delta_{kk'} \delta_{\ell\ell'}\,. \tag{4.68}$$

4.4 Excess power method

The excess power method [43] is one of the unmodeled detection methods, i.e., the precise predictions of the waveforms are not required. When the signal we want to detect is localized to a narrower region than the total data points, the vast data points contain only noise. Therefore, the signal power can be collected efficiently by summing up limited data points where the signal is expected to exist. As we see in the previous section, the ℓ-domain signal is well localized. Therefore, the excess power method is suitable to select candidates rapidly.

An ℓ-domain strain is divided into short chunks. The length of each chunk is denoted by $\delta\ell$, and neighboring chunks are overlapped by $(\delta\ell)_{\rm slide}$. In this work, we set $\delta\ell = 2048$, and $(\delta\ell)_{\rm slide} = 128$. The number of chunks obtained from one ℓ-domain strain is $N_{\rm chunk/signal} = (N_{\rm seg} - \delta\ell)/(\delta\ell)_{\rm slide}$. For the c-th chunk in a ℓ-domain strain data obtained from the k-th frequency bin of STFT and the grid point $\boldsymbol{n}_{\rm g}$, the detection statistic $\mathcal{E}$ is defined by

$$\mathcal{E}(\boldsymbol{n}_{\rm g}, f_k, c) := 4 \sum_{\ell=0}^{\delta\ell-1} \frac{|\mathsf{S}_{\ell + c(\delta\ell)_{\rm slide}, k}|^2}{\tilde{\sigma}_k^2}\,, \tag{4.69}$$

where c is the index specifying a chunk ($c = 0, 1, \cdots, N_{\rm chunk/signal} - 1$).

When the signal is absent, $\mathcal{E}$ becomes

$$\mathcal{E} = 4 \sum_{\ell=0}^{\delta\ell-1} \frac{|\mathsf{N}_{\ell + c(\delta\ell)_{\rm slide}, k}|^2}{\tilde{\sigma}_k^2}\,, \tag{4.70}$$

where we omit the arguments of $\mathcal{E}$, for brevity. It coincides with the sum of $2\delta\ell$ variables that independently follow a Gaussian distribution with mean 0 and variance 1. Rigorously speaking, the distribution of $\mathcal{E}$ in the absence of a signal is a chi-square distribution with $2\delta\ell$ degrees of freedom. However, it can be approximated by a Gaussian distribution with mean $2\delta\ell$ and standard deviation $2\sqrt{\delta\ell}$ because here we choose $\delta\ell$ to be large. The average and the standard deviation of $\mathcal{E}$ in the case when only the noise exists are denoted by $\langle \mathcal{E} \rangle_{\rm n}$ and $\sigma_{\rm n}(\mathcal{E})$, respectively. Namely,

$$\langle \mathcal{E} \rangle_{\rm n} = 2\delta\ell,\ \sigma_{\rm n}(\mathcal{E}) = 2\sqrt{\delta\ell}\,. \tag{4.71}$$

The SNR of the excess power method $\rho_{\rm EP}$ is defined by

$$\rho_{\rm EP} := \frac{\mathcal{E} - \langle \mathcal{E} \rangle_{\rm n}}{\sigma_{\rm n}(\mathcal{E})} \,. \tag{4.72}$$

From Eq. (4.71), the distribution of $\rho_{\rm EP}$ when the signal is absent is a standard normal distribution, i.e.,

$$\langle \rho_{\rm EP} \rangle = 0 \,, \;\; \sigma_{\rm n}(\rho_{\rm EP}) = 1 \,. \tag{4.73}$$

Setting the threshold value $\hat{\rho}_{\rm EP}$ so that the number of the false alarm events is suppressed below a certain level, we select the candidate set of parameter values $\{\boldsymbol{n}_{\rm g}, f_k, c\}$ when

$$\rho_{\rm EP}(\boldsymbol{n}_{\rm g}, f_k, c) > \hat{\rho}_{\rm EP} \,. \tag{4.74}$$

In the same manner, the distribution of the SNR under the condition that the signal also exists can be approximated by a Gaussian distribution. Its mean and standard deviation are derived as

$$\mu_{\rm EP}(\boldsymbol{\xi}) = \frac{2P_k(\boldsymbol{\xi})}{\tilde{\sigma}_k^2 \sqrt{\delta\ell}} \,, \tag{4.75}$$

and

$$\sigma_{\rm EP}(\boldsymbol{\xi}) = \sqrt{1 + \frac{4P_k(\boldsymbol{\xi})}{\tilde{\sigma}_k^2 \delta\ell}} \,, \tag{4.76}$$

with

$$P_k(\boldsymbol{\xi}) := \sum_{\ell} |\mathsf{H}_{\ell,k}(\boldsymbol{\xi})|^2 \,. \tag{4.77}$$

Here, $\boldsymbol{\xi}$ denotes a set of parameters $\boldsymbol{\xi} := (h_0, \vec{\xi}) = (h_0, f_{\rm gw}, \alpha_{\rm s}, \delta_{\rm s})$.

4.5 Deep learning

4.5.1 Fundamentals

Deep learning (Ref. [44] as a textbook) is one of the approaches for extracting features being hidden in the data. Artificial neural networks (ANNs) are the architectures playing the central role in deep learning. Originally, an ANN is invented as a mathematical model of the human brain, which is formed from a lot of neurons interconnected with each other [45]. Nowadays, an ANN is mainly used as a mathematical tool for analyzing big data, and its structure has less relation with neuroscience.

An ANN consists of consecutive layers, and each layer is formed by a lot of units ("neurons"). Each layer takes inputs from the previous layer and passes the processed data to the next layer. As a simple example, the process occurring in each layer can be written as the combination of affine transformation and a non-linear transformation, i.e.,

$$x_i^{(l)} = g^{(l)} \left(\sum_{j=1}^{N^{(l-1)}} w_{ij}^{(l)} x_j^{(l-1)} + b^{(l)} \right) \;\; (i = 1, 2, \cdots, N^{(l)}) \,, \tag{4.78}$$

where $x^{(l)}$ is a set of input data on the l-th layer, $N^{(l)}$ is the number of units in the l-th layer and $g^{(l)}$ is a nonlinear function called an activation function. Each unit (neuron) corresponds to each element of the vector $x_i^{(l)}$. A whole structure of ANN is formed by accumulated layers and ANN transforms the input data $\boldsymbol{x}_{\rm in}$ sequentially,

$$\boldsymbol{x}_{\rm in} = \boldsymbol{x}^{(0)} \to \boldsymbol{x}^{(1)} \to \cdots \to \boldsymbol{x}^{(L-1)} \to \boldsymbol{x}^{(L)} = \boldsymbol{y}\,, \tag{4.79}$$

where $\boldsymbol{y}$ is the output of the ANN. The parameters w and b are, respectively, called weights and biases. They are tunable parameters and optimized to capture the features of data. The optimization of the weights and the biases is called training. Frequently, the affine transformation and the non-linear transformation are divided into two layers, called a linear layer and a non-linear transformation layer, respectively. Typically, the activation functions except for that of the last layer are set to be the same. We use a rectified linear unit function (ReLU) [46], which is defined by

$$g^{(l)}(z) = \max[z, 0]\,,\ (l = 1, 2, \cdots, L-1)\,. \tag{4.80}$$

The activation function of the last layer is properly chosen so that the output $\boldsymbol{y}$ is suitable for the problem we want to solve. For the example of a regression problem, the identity function

$$g^{(L)}(z) = z\,, \tag{4.81}$$

is usually employed. The model (4.78) is called a "multi-layer perceptron" (MPL). With appropriately adjusted parameters, an MPL model can approximate arbitrary functions in principle. This fact is called "universal approximation theory" [47, 48] and is believed to be one reason why an ANN works well.

In practice, various types of layers are employed depending on the properties of the data we want to analyze. In this work, we use also one-dimensional convolutional layers [49] and max-pooling layers [50]. The input of a convolutional layer, denoted by x_i^c, is a set of vectors. For example, in a color image, each pixel has three channels corresponding to three primary colors of light. Therefore, the input data is a set of three two-dimensional arrays. In our case, the real part and the imaginary part of ℓ-domain signals consist of two channels. Therefore, the input data is regarded as a set of two one-dimensional arrays. In a convolutional layer, the discrete convolution, which is represented as

$$o_i^{c'} = \sum_{c=1}^{C} \sum_{k=1}^{K} f_k^{c,c'} x_{i+k}^c + b^{c'}\,, \tag{4.82}$$

is calculated. Here, x is the input, and o is the output data of the layer. C and K are the number of channels and the width of the kernel, respectively. Each pixel of the data is specified by an index i. The parameters f and b are optimized during the training. A max pooling layer is defined by

$$o_i^c = \max_{k=1,2,\cdots,K} [x_{si+k}^c]\,, \tag{4.83}$$

with the stride s and the kernel size K. This layer reduces the length of the data and hence the computational cost.

In supervised learning, the trainable parameters of the ANN (weights, biases and filters) are adjusted by using a given dataset $\mathcal{D}$ consisting of many pairs of an input data and a target value,

$$\mathcal{D} = \{(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2), \cdots (\boldsymbol{x}_{N_{\rm train}}, \boldsymbol{y}_{N_{\rm train}})\} \,. \tag{4.84}$$

The dataset used for optimizing the adjustable parameters and the biases is called the "training dataset". An ANN learns the relation between input data and target values from the training dataset and predicts the target values corresponding to newly given input data. In order to train an ANN, the deviation between the target value and the predicted value is quantified by a loss function. For a regression problem, the square loss,

$$\mathcal{L}[\boldsymbol{y}(\boldsymbol{x}; w), \boldsymbol{t}] = \frac{1}{2} \sum_{i=1}^{N^{(L)}} \{y_i(\boldsymbol{x}; w) - t_i\}^2 \,, \tag{4.85}$$

is often used. The predicted value $\boldsymbol{y}$ depends on the adjustable parameters of the neural network, which are collectively denoted by w, for brevity. An ANN is optimized so as to minimize the average of the loss function taken over the training dataset $\mathcal{D}$,

$$\mathcal{L}[w|\mathcal{D}] = \frac{1}{|\mathcal{D}|} \sum_{(\boldsymbol{x},\boldsymbol{t})\in\mathcal{D}} \mathcal{L}[\boldsymbol{y}(\boldsymbol{x}; w), \boldsymbol{t}] \,. \tag{4.86}$$

Because an ANN is a highly complicated function of w, the minimization of $\mathcal{L}[w]$ cannot be done analytically. Therefore, the optimization of w is done by an iterative method. In the simplest way, the adjustable parameters w are updated by the algorithm

$$w \to w - \eta \nabla_w \mathcal{L}[w|\mathcal{D}] \,, \tag{4.87}$$

where the parameter η characterizing the strength of each update is called a learning rate. The algorithm (4.87) seems reasonable, but in practice, it does not work well because the loss function $\mathcal{L}[w]$ has a lot of local minima. The mini-batch training is usually employed to prevent to stack the training. In the mini-batch training, a subset of the training dataset ("mini-batch") is randomly chosen at the beginning of each iteration. We denote a mini-batch by

$$\mathcal{D}_{\rm batch} = \{(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{t}}_1), (\tilde{\boldsymbol{x}}_2, \tilde{\boldsymbol{t}}_2), \cdots, (\tilde{\boldsymbol{x}}_{\tilde{N}}, \tilde{\boldsymbol{t}}_{\tilde{N}})\} \,. \tag{4.88}$$

It must be satisfies $\mathcal{D}_{\rm batch} \subset \mathcal{D}$. The update algorithm (4.87) is replaced with the one using the average of the loss function over a mini-batch. It is called stochastic gradient descent and can be written as

$$w \to w - \eta \nabla_w \mathcal{L}[w|\mathcal{D}_{\rm batch}] \,, \tag{4.89}$$

So far, many variants (e.g., momentum [51], RMSProp [52], Adam [53]) are proposed. Regardless of the update algorithm, the gradients of a loss function are required, and they can be quickly calculated by the backpropagation scheme [54].

In the training process, we optimize a neural network so that the loss function is minimized for a dataset. However, there are two problems which need to be concerned seriously. The first one is a problem called overfitting. The trained ANN fits well with the training dataset, but the predictions for newly given inputs are not accurate. This situation is called overfitting. We cannot find the neural network is in overfitting by monitoring only the loss function of the training data. Second, we do not know a priori which structure of ANN nor which update procedure is plausible. The performance of the trained neural network depends not only on the adjustable parameters but also on the number of neurons of each layer and the learning rate, and so on. These parameters, so-called hyperparameters, are not optimized during the training process. In order to overcome these problems, another dataset that is independent of the training dataset is needed to be prepared. We refer to the additional dataset as a validation dataset. In order to prevent the overfitting, the training process should be stopped when the loss for the validation dataset tends to deviate from that for the training dataset (early stopping). In order to optimize the hyperparameters, many neural network models having various structures are trained with different training schemes. The validation data is used to monitor the training process and assess which model is better for the problem that the user wants to solve. Among these ANNs, we choose the one performing with the smallest loss for the validation dataset.

4.5.2 Application to gravitational wave data

Before we explain our neural network setup, the recent studies of its application to the gravitational wave data analysis are briefly reviewed.

Since the pioneering work done by George and Huerta, deep learning methods are widely applied to gravitational wave data analysis. In Ref. [55], the ANN optimized for the detection and the parameter estimation of the binary black hole coalescence. Their ANN is trained with the mock signals generated by injecting the waveforms of non-spinning BBH into simulated Gaussian noise data. Each data has a duration of 1 second and a sampling frequency of 8192 Hz. Employing the matched filtering as the benchmark, they examined the detection efficiency and the parameter estimation accuracy. Their ANN exhibits comparable performance to the matched filtering. The incredible result is that the ANN can complete the parameter estimation and the detection for one data only within 535 microseconds with a GPU, in contrast to the matched filtering that takes 1.1 seconds with a CPU. Their result indicates that the deep learning method will promote the real-time analysis of the gravitational waves. This work is extended to a wider class of binaries, e.g., spinning binaries [56, 57], eccentric binaries [58].

The deep learning methods are also applied to classify glitch noises [59], model waveforms rapidly [60], and predict the posterior distributions [61–64]. As for the continuous gravitational waves, several groups already applied deep learning. Dreissigacker *et al.* [65, 66] applied neural networks to the all-sky search of signals with the duration of 10^5 and 10^6 seconds. For the case of a duration of 10^5 sec, their neural network works well. On the other hand, for 10^6 sec case, their ANN cannot show the competitive performance

to the matched filtering. In [67, 68], it is shown that the combination of deep learning and the existing methods is useful.

4.5.3 ANNs in our work

Table 4.1 shows the architecture of the neural network we used. The real part and the imaginary part of a short chunk are used as input data of our ANN. It is empirically known that the input data should be normalized to take the value of $O(1)$ for successful training the neural network. We normalized each input data by subtracting its average and dividing by its standard deviation. The output of the neural network is to be interpreted as the predicted sky position. The ℓ-domain waveform, $\mathsf{H}_{\ell,k}$, depends mainly on the residual phase $\delta\Phi_\odot$, and it depends on the sky position $(\alpha_\mathrm{s}, \delta_\mathrm{s})$ through the vector $\Delta\boldsymbol{n}$. Because we approximate the orbital motion of the Earth so that $z_\odot = 0$, only x and y components of $\Delta\boldsymbol{n}$ affect the residual phase $\delta\Phi_\odot$. Therefore, it is reasonable to label ℓ-domain waveforms with the values of Δn_x and Δn_y. As well as the input data, the target values are also normalized. The absolute values of Δn_x and Δn_y is $\lesssim O(10^{-3})$. Therefore, we use $\Delta n'_x = 10^3 \Delta n_x$ and $\Delta n'_y = 10^3 \Delta n_y$ as the label. The neural network outputs the predicted values of $\Delta n'_x$ and $\Delta n'_y$, and they are converted into the predicted values of the right ascension and the declination angle of the source after multiplying 10^{-3}. The neural network is applied to each candidate, which is selected by the excess power method. By narrowing down the possible region where the source is likely to be located, we aim to reduce the computational cost of the follow-up analysis. We introduce a parameter r_NN to quantify how large the area to be followed up by the coherent matched filtering is. At this time, we leave r_NN as a free parameter, and it will be determined by considering the computational time and the detection probability.

To generate datasets, we use the Fourier transform of Eq. (4.48) as the waveform model. We assume that the observation period is 2^{24} seconds, and the frequency of the gravitational waves is less than 100Hz. We assume that the strain data is obtained from a single detector. The geometrical information (e.g., the latitude of the detector's site) of the LIGO Hanford is used for calculating the antenna pattern function. In this work, the parameter space is restricted to a narrow region, for simplicity. Specifically, the gravitational wave frequency f_gw is uniformly distributed on

$$f_\mathrm{gw} \in \left[100 - \frac{1}{2T_\mathrm{seg}}, 100 + \frac{1}{2T_\mathrm{seg}}\right] \mathrm{Hz}\,. \tag{4.90}$$

Also the source location is sampled only from the vicinity of a reference grid point shown in Table. 4.2. We introduce the normalized amplitude $\hat{h}_0$ by

$$\hat{h}_0 := h_0 \left(\frac{S_\mathrm{n}(f_k)}{1\mathrm{Hz}^{-1}}\right)^{-1/2}\,. \tag{4.91}$$

We have to train the neural network with signals having various amplitudes, but it is not efficient to store all signals. Thus, the amplitudes of waveforms are normalized to be

Table 4.1: Table of the architecture of the neural network we used. For convolutional layers and max-pooling layers, the input and the output are characterized by a pair of two values (C, N) where C is the number of channels, and N is the data length. The kernel sizes of the convolutional layers are 16, 16, 8, 8, 4, and 4 from the earlier to the later layer. For all of the max-pooling layers, the kernel sizes are set to be 4.

Layer	Input	output
1-d convolution	(2, 2048)	(64, 2033)
ReLU	(64, 2033)	(64, 2033)
1-d convolution	(64, 2033)	(64, 2018)
ReLU	(64, 2018)	(64, 2018)
max pooling	(64, 2018)	(64, 504)
1-d convolution	(64, 504)	(128, 497)
ReLU	(128, 497)	(128, 497)
1-d convolution	(128, 497)	(128, 490)
ReLU	(128, 490)	(128, 490)
max pooling	(128, 490)	(128, 122)
1-d convolution	(128, 122)	(256, 119)
ReLU	(256, 119)	(256, 119)
1-d convolution	(256, 119)	(256, 116)
ReLU	(256, 116)	(256, 116)
max pooling	(256, 116)	(256, 29)
Dense	256×29	64
ReLU	64	64
Dense	64	64
ReLU	64	64
Dense	64	2

$\hat{h}_0 = 1$, and they are used as the reference waveforms. In the process of the training, we randomly sample the amplitudes $\{\hat{h}_0\}$ for each iteration step and multiply them to the reference waveforms. The amplitudes $\{\hat{h}_0\}$ are sampled from the uniform distribution on

$$-2.1 \leq \log_{10} \hat{h}_0 \leq -1.0\,. \tag{4.92}$$

Thus, the injected waveform is the ℓ-domain signal normalized by the square-root of the power spectral density of the noise with fixed frequency which we set $f = f_k = 100$Hz. The variance of the real part and the imaginary part of the Gaussian noise to which the ℓ-domain waveforms are injected should be set to

$$\sigma^2 = \frac{\mathcal{W}}{4T_{\rm seg}N_{\rm seg}}\,, \tag{4.93}$$

because the variance of the noise in ℓ-domain is given by Eq. (4.68) and the noise is also normalized by $S_{\rm n}(f_k)$. The numbers of waveforms in the training dataset and the

validation dataset are 200,000 and 10,000, respectively. The auxiliary setups are listed in the following.

- We employ the mini-batch training with a batch size 256.
- A GPU GeForce 1080Ti is used for the training.
- The updating algorithm is the Adam with the learning rate 10^{-4}.
- PyTorch [69], the python library for deep learning, is used for implementing our code.

Table. 4.2 summarizes parameter values we used.

Table 4.2: List of the parameter values we used in this work.

Symbol	Parameters	Value
$T_{\rm obs}$	Observation period	2^{24} sec
$f_{\rm s}$	Sampling frequency	1024 Hz
$N_{\rm grid}$	The number of grids	352436
$N_{\rm bin}$	The number of frequency bins of STFT	3200
$T_{\rm seg}$	Duration of a STFT segment	32 sec
$\delta\ell$	Length of chunk	2048
$(\delta\ell)_{\rm slide}$	Dilation of chunk	128
f_k	Fixed frequency bin	100 Hz
$f_{\rm s,coh}$	Sampling frequency of the heterodyned strain	$1/T_{\rm seg}$
$\alpha_{\rm g}$	Right ascension of grid	-0.158649 rad
$\delta_{\rm g}$	Declination of grid	1.02631 rad

4.6 Follow-up analysis by the matched filtering

After the selection of the candidates by the excess power method and the specification of the possible source location by the neural network, we take the matched filtering to each candidate as a follow-up analysis. For the follow-up analysis, we refine the grid points so that the possible region of the candidate is entirely covered by the new grid points. The time-resampling process is applied to the original detector strain with each grid point. Each candidate is assigned by the frequency bin of STFT. It means that, if the candidate is the real event, the gravitational wave frequency $f_{\rm gw}$ satisfies

$$f_k - \frac{1}{2T_{\rm seg}} \leq f_{\rm gw} \leq f_k + \frac{1}{2T_{\rm seg}}\,. \tag{4.94}$$

Using this information, we can reduce the computational cost of the matched filtering. Heterodyning is the widely used method in the signal processing and is already applied

to the search of continuous gravitational waves from known sources [70]. The resampled strain Eq. (4.38) is multiplied by the factor $e^{-2\pi i f_k \zeta}$ and we obtain heterodyned strain as

$$s(\zeta) \cdot e^{-2\pi i f_k \zeta} = h_0 e^{2\pi i (f_{\rm gw} - f_k)\zeta + i\delta\Phi_\odot} + n(\zeta) \cdot e^{-2\pi i f_k \zeta} \tag{4.95}$$

With Eq. (4.41) and (4.94), the frequency of the heterodyned gravitational wave signal is obtained as

$$|f_{\rm het}| \lesssim |f_{\rm gw} - f_k| + \left| \frac{1}{2\pi} \frac{d\delta\Phi_\odot}{dt} \right| \lesssim \frac{1}{T_{\rm seg}} + \delta f_\odot \sim \frac{1}{T_{\rm seg}} \,. \tag{4.96}$$

This means that the gravitational waves become the near DC components after the heterodyning. Thus, we can reduce the data points by the downsampling without the loss of the signal power. It is reasonable to set the sampling frequency of the downsampled strain to be $\sim 1/T_{\rm seg}$.

4.7 Results

4.7.1 Computational cost

First of all, the computational cost of the excess power method is estimated. To calculate the excess power statistics $\mathcal{E}$ for all chunks obtained from one ℓ-domain signal, we need $2N_{\rm seg}$ multiplications and $2N_{\rm seg}$ additions of real numbers. Because we have $N_{\rm bin} \times N_{\rm grid}$ ℓ-domain signals, the computational cost of the excess power method is evaluated as

$$\mathcal{N}_{\rm EP} = 4N_{\rm seg} \times N_{\rm bin} \times N_{\rm grid} \sim 4.7 \times 10^{15} \,, \tag{4.97}$$

in the unit of the number of floating point operations. As shown in later, this cost is negligibly small comparing with the computational costs of the preprocess and the follow-up.

Next, we assess the computational time of the neural network. A GPU is used for the calculation of the neural network. Therefore, it is not straightforward to compare the computational cost of the neural network with others. Here, we estimate the computational time of the neural network by measuring the elapsed time for analyzing test data. Using a single GPU (GeForce 1080Ti), it takes 1.4sec to process ten thousand chunks. The neural network is applied to all candidates selected by the excess power method. The number of candidates $N_{\rm candidate}$ is estimated by

$$N_{\rm candidate} = N_{\rm chunk} \times {\rm FAP}_{\rm EP} = N_{\rm grid} \cdot N_{\rm bin} \cdot \frac{N_{\rm seg}}{(\delta\ell)_{\rm slide}} \times {\rm FAP}_{\rm EP} \simeq 4.6 \times 10^{10} \left(\frac{{\rm FAP}_{\rm EP}}{10^{-2}} \right) \,. \tag{4.98}$$

Therefore, the computational time of the neural network is estimated by

$$\mathcal{T}_{\rm NN} = N_{\rm candidate} \times \frac{1.4{\rm sec}}{10^4 {\rm data}} \simeq 6.4 \times 10^6 {\rm sec} \left(\frac{{\rm FAP}_{\rm EP}}{10^{-2}} \right) \,. \tag{4.99}$$

Here, we focus on $\mathrm{FAP}_{\mathrm{EP}} \leq 10^{-2}$. Then, the computational time of the neural network is suppressed to an acceptable level. In addition, if multiple GPUs are available, the neural network analysis of many candidates can be parallelized. In this thesis, we, therefore, neglect the computational time of the neural network.

The computational cost of the preprocess is originated mainly from STFTs and FFTs. STFTs are carried out for all grid points on the sky, and FFTs are performed for all STFT frequency bins lower than 100Hz and all grid points. Therefore, we can estimate

$$\mathcal{N}_{\mathrm{preprocess}} = N_{\mathrm{grid}} \times (\mathcal{N}_{\mathrm{STFT}} + \mathcal{N}_{\mathrm{FFT}} \times N_{\mathrm{bin}}) \,. \tag{4.100}$$

We stress that $\mathcal{N}_{\mathrm{preprocess}}$ is constant. With the values listed in Table. 4.2, we obtain

$$\mathcal{N}_{\mathrm{preprocess}} \simeq 2.3 \times 10^{18} \,. \tag{4.101}$$

Comparing it with $\mathcal{N}_{\mathrm{EP}}$ given in Eq. (4.97), we can neglect $\mathcal{N}_{\mathrm{EP}}$.

At last, we estimate $\mathcal{N}_{\mathrm{follow-up}}$, i.e., the computational cost of the follow-up analysis. The cost of the follow-up analysis is dominated by that of taking match by the Fourier transform. Therefore, we can estimate $\mathcal{N}_{\mathrm{follow-up}}$ as

$$\mathcal{N}_{\mathrm{follow-up}} = N_{\mathrm{candidate}} \times \frac{\pi (r_{\mathrm{NN}})^2}{(\delta\theta)^2_{\mathrm{coh}}} \times \mathcal{N}_{\mathrm{FFT,coh}} \,. \tag{4.102}$$

Let us remind that the typical size of the region where each grid point covers is given by $(\delta\theta)^2_{\mathrm{coh}}$ (see Eq. (3.47)). Thus, the second factor on the right-hand side is the required number of grid points to cover the region predicted by the neural network for each candidate. Multiplying it and $N_{\mathrm{candidate}}$, we obtain the total number of grid points for the coherent matched filtering. The third factor is the computational cost of FFT for each candidate and can be estimated by

$$\mathcal{N}_{\mathrm{FFT,coh}} = 5(T_{\mathrm{obs}} f_{\mathrm{s,coh}}) \log_2 (T_{\mathrm{obs}} f_{\mathrm{s,coh}}) \simeq 5.0 \times 10^7 \,. \tag{4.103}$$

Finally, we get the estimated value of $\mathcal{N}_{\mathrm{follow-up}}$ as

$$\mathcal{N}_{\mathrm{follow-up}} \simeq 7.2 \times 10^{22} \left(\frac{r_{\mathrm{NN}}}{10^{-3}\mathrm{rad}} \right)^2 \left(\frac{\mathrm{FAP}_{\mathrm{EP}}}{10^{-2}} \right) \,. \tag{4.104}$$

Using Eqs. (4.101) and (4.104), we obtain the estimation of the total computational cost as

$$\mathcal{N}_{\mathrm{comp}} = \mathcal{N}_{\mathrm{preprocess}} + \mathcal{N}_{\mathrm{follow-up}} \,, \tag{4.105}$$

as a function of $(\mathrm{FAP}_{\mathrm{EP}}, r_{\mathrm{NN}})$. Figure. 4.7 shows the computational cost of our entire analysis. Depending on an available computational resources, we can read the suitable combination of $(\mathrm{FAP}_{\mathrm{EP}}, r_{\mathrm{NN}})$.

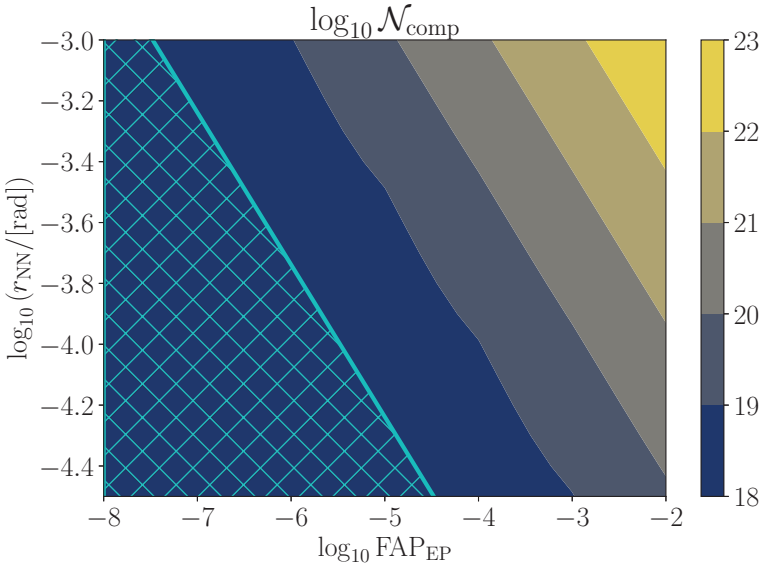

Figure 4.7: Logarithm of the computational cost of the entire analysis. In the cyan hatched region, the preprocess dominates the computational cost, and $\mathcal{N}_{\rm comp}$ is almost constant.

4.7.2 False alarm probability

The false alarm probability of the entire analysis is given by multiplying the false alarm probabilities of each analysis stage (see [71, 72]), that is,

$$\begin{aligned} p_{\rm fa}(\boldsymbol{\rho}) = &\left\{1-\left(\mathrm{Prob}\left[\rho_{\rm EP} < \hat{\rho}_{\rm EP} \mid \rho_{\rm EP} \sim \mathcal{N}(0,1)\right]\right)^{N_{\rm chunk}}\right\} \\ &\times \left\{1-\left(\mathrm{Prob}\left[\rho_{\rm MF} < \hat{\rho}_{\rm MF} \mid \rho_{\rm MF} \sim \mathcal{N}(0,1)\right]\right)^{N_{\rm t}}\right\}, \end{aligned} \tag{4.106}$$

with $\boldsymbol{\rho} := (\mathrm{FAP}_{\rm EP}, r_{\rm NN}, \hat{\rho}_{\rm MF})$. Here, $N_{\rm t}$ denotes the number of templates required for the coherent matched filtering, and it can be calculated by

$$N_{\rm t} = N_{\rm candidate} \times \frac{\pi (r_{\rm NN})^2}{(\delta\theta)^2_{\rm coh}} \times N_{\rm bin,coh}\,, \tag{4.107}$$

where $N_{\rm bin,coh}$ is the number of frequency bins appeared in the coherent matched filtering. Substituting the values shown in Table. 4.2, we obtain

$$N_{\rm t} \simeq 5.6 \times 10^{21} \left(\frac{r_{\rm NN}}{10^{-3}\rm rad}\right)^2 \left(\frac{\mathrm{FAP}_{\rm EP}}{10^{-2}}\right). \tag{4.108}$$

The false alarm probability of the entire process is determined by that of the follow-up stage. Therefore, we can approximate the false alarm probability as

$$p_{\rm fa}(\boldsymbol{\rho}) \simeq \left\{1-\left(\,\mathrm{Prob}[\rho_{\rm MF} < \hat{\rho}_{\rm MF} \mid \rho_{\rm MF} \sim \mathcal{N}(0,1)]\,\right)^{N_{\rm t}}\right\}. \tag{4.109}$$

In addition, because $N_{\rm t} \ll 1$, the false alarm probability is

$$p_{\rm fa}(\boldsymbol{\rho}) \simeq N_{\rm t} \cdot {\rm Prob}\left[\rho_{\rm MF} > \hat{\rho}_{\rm MF} \mid \rho_{\rm MF} \sim \mathcal{N}(0,1)\right] . \tag{4.110}$$

In this thesis, $\hat{\rho}_{\rm MF}$ is chosen so that $p_{\rm fa}(\boldsymbol{\rho}) = 0.01$. As a remark, $N_{\rm t}$ is a function of ${\rm FAP}_{\rm EP}$ and $r_{\rm NN}$, and , hence, $\hat{\rho}_{\rm MF}$ also depends on them.

Before discussing the detection probability, the statistical properties of the excess power statistic are checked. Because it is difficult to simultaneously treat a whole signal of a duration $T_{\rm obs} \sim O(10^7)$sec, we generate $N_{\rm seg}$ strain data with the duration of $T_{\rm seg}$ which contain only noise with assuming $S_{\rm n}(f) = 1$. We carry out the Fourier transform to each short strain after applying the window function. The k-th frequency bin is picked up and regarded as an element of $\left\{n_{j,k}^{\rm STFT}\right\}_{j=1}^{N_{\rm seg}}$. We divide them into $N_{\rm seg}/\delta\ell$ chunks. Repeating this procedure 80 times, we obtain 10,240 chunks. The excess power statistics $\mathcal{E}$ and the SNR $\rho_{\rm EP}$ are calculated for each chunk. Figure. 4.8 shows the histogram of $\rho_{\rm EP}$. As stated in Sec. 4.4, $\rho_{\rm EP}$ should follow a standard normal distribution. From Fig. 4.8, this property seems to be valid. In order to confirm it more quantitatively, the Kolmogorov-Smirnov (KS) test is carried out. The p-value of the KS test is 0.753254 and means the SNR of noise data follows the standard normal distribution.

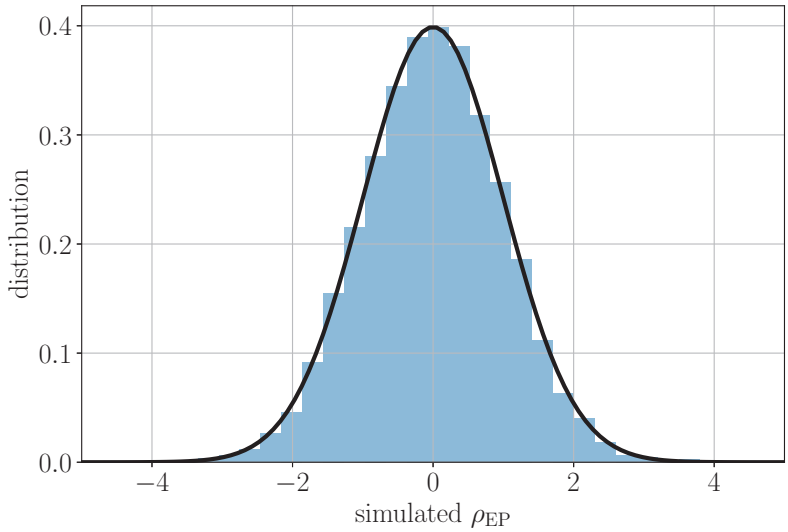

Figure 4.8: Histogram of the SNR of the excess power method. Black solid line indicates the standard normal distribution. The theoretical prediction and the recovered distribution match well.

4.7.3 Detection probability

The detection probability of the entire process is also estimated by multiplying those of all stages. If the signal with the amplitude of h_0 exists, the detection probability can be

estimated as

$$\begin{aligned} p_{\rm det}(h_0;\boldsymbol{\rho}) &= \langle p_{\rm det}(\boldsymbol{\xi};\boldsymbol{\rho})\rangle_{\vec{\xi}} \\ &= \Bigg\langle {\rm Prob}\left[\rho_{\rm EP} > \hat{\rho}_{\rm EP} \mid \rho_{\rm EP} \sim \mathcal{N}(\mu_{\rm EP}(\boldsymbol{\xi}), \sigma_{\rm EP}(\boldsymbol{\xi}))\right] \\ &\times {\rm Prob}\left[\text{source is located in a predicted region} \mid \boldsymbol{\xi}, r_{\rm NN}\right] \\ &\times {\rm Prob}\left[\rho_{\rm MF} > \hat{\rho}_{\rm MF} \mid \rho_{\rm MF} \sim \mathcal{N}(\mu_{\rm MF}(\boldsymbol{\xi}), 1)\right] \Bigg\rangle_{\vec{\xi}}, \end{aligned} \tag{4.111}$$

where $\langle\cdots\rangle_{\vec{\xi}}$ is the average over the parameter $\vec{\xi}$, i.e.,

$$\langle\cdots\rangle_{\vec{\xi}} := \int d\vec{\xi}(\cdots)\pi(\vec{\xi})\,, \tag{4.112}$$

with a prior distribution $\pi(\vec{\xi})$. In principle, we need to take an average over the possible parameter regions, e.g., source position on the entire sky. However, we trained the neural network with the dataset containing the signals that are located only in the vicinity of the reference grid point. The gravitational wave frequencies are also sampled from the narrow frequency band (see Eq. (4.90)). It is reasonably envisaged that the trained neural network does not work well for signals outside of the parameter region where the training data are sampled. Therefore, in this thesis, we only take the average over the same parameter region as the training data.

In order to quantify the detection power, we define the *sensitivity depth* $\mathcal{D}^{95\%}$ by

$$\mathcal{D}^{95\%} := \frac{1}{h_0^{95\%}}\left(\frac{S_{\rm n}(f_k)}{1{\rm Hz}^{-1}}\right)^{1/2}, \tag{4.113}$$

where $h_0^{95\%}$ satisfies the following relation,

$$p_{\rm det}(h_0^{95\%};\boldsymbol{\rho}) = 0.95\,. \tag{4.114}$$

Eq. (4.114) means that the signal having the amplitude $h_0^{95\%}$ is expected to be detected with the detection probability of 95%. We optimize $({\rm FAP}_{\rm EP}, r_{\rm NN})$ by comparing the values of $\mathcal{D}^{95\%}$ for each combination of $({\rm FAP}_{\rm EP}, r_{\rm NN})$. In order to explore the parameter space of $({\rm FAP}_{\rm EP}, r_{\rm NN})$, the regular grid is placed on $\log_{10}{\rm FAP}_{\rm EP}$ from -8 to -2 by a step of 1, and on $\log_{10} r_{\rm NN}$ from -4.5 to -3.0 by a step of 0.05. The values of $\mathcal{D}_0^{95\%}$ are calculated by the following procedure. First of all, a regular grid on $\log_{10}\hat{h}_0$ is placed from -2.3 to -1.0 by a step of 0.05. For each sample of $\hat{h}_0$, the parameters $\vec{\xi}$ are randomly sampled for 1,024 times. These samples are denoted by $\{\vec{\xi}^{(i)}\}_{i=1}^{1024}$. We analytically estimate the detection probabilities of the excess power method and the follow-up analysis by assuming the Gaussian distribution. The detection probability of the neural network is estimated by applying it to simulated signals. To prepare simulated signals, we generate 1024 ℓ-domain waveforms corresponding to the parameters $\{\vec{\xi}^{(i)}\}_{i=1}^{1024}$. In order to estimate the detection probability

for each parameter $\vec{\xi}^{(i)}$, 512 noise realizations are generated by assuming Gaussian noise, same as the generation of the training data set. The detection probability of the neural network is estimated by the fraction of events which the source position is located within the predicted region. Multiplying the detection probabilities of all stages, we obtain the estimated values of $p_{\rm det}(h_0, \vec{\xi}^{(i)}; \boldsymbol{\rho})$. Finally, the detection probability of the entire process is approximated by the ensemble average, i.e.,

$$p_{\rm det}(h_0; \boldsymbol{\rho}) \simeq \frac{1}{1024} \sum_{i=1}^{1024} p_{\rm det}(h_0, \vec{\xi}^{(i)}; \boldsymbol{\rho}) \,. \tag{4.115}$$

Repeating these procedure with changing the values of $\hat{h}_0$, or correspondingly, $\mathcal{D} := \hat{h}_0^{-1}$, we obtain the estimated detection probability for each value of $\mathcal{D}$. Using the sigmoid-like function,

$$\varsigma(\mathcal{D}; a, b) := \frac{1}{1 + \exp[(\mathcal{D} - a)/b]} \,, \tag{4.116}$$

we carry out the curve fitting to the detection probability. By inverting the fitted curve, we obtain $\mathcal{D}^{95\%}$ as

$$\mathcal{D}^{95\%}({\rm FAP}_{\rm EP}, r_{\rm NN}) = a^* - b^* \ln \frac{1 - 0.95}{0.95} \,, \tag{4.117}$$

with the optimized parameters a^* and b^*. An example of the fitting procedure is shown in Figure. 4.9. Estimated values of $\mathcal{D}^{95\%}$ for each combination $({\rm FAP}_{\rm EP}, r_{\rm NN})$ are shown in Figure. 4.10. For example, let us assume that a 1TFlops computer is available and the computational time of 10^7sec is acceptable. Then, the total computational cost should be below 10^{19}. From Fig. 4.7, we can read a feasible parameter space where the computational cost is suppressed to 10^{19}. Among them, we can pick up the combination $({\rm FAP}_{\rm EP}, r_{\rm NN})$ with which our method is the most sensitive. Figure 4.10 gives us that when we set $({\rm FAP}_{\rm EP}, r_{\rm NN}) = (10^{-5}, 10^{-3.6}{\rm rad})$ the sensitivity depth reaches $\mathcal{D}^{95\%} \sim 90$.

To check our estimation of the detection probability is appropriate, the injection test is performed. The excess power method and the neural network stage are tested with ℓ-domain signals injected into simulated Gaussian noise data. For reducing the computational cost, the detection probability of the coherent matched filtering is estimated analytically. We prepare ten thousand signals for each combination of $({\rm FAP}_{\rm EP}, r_{\rm NN})$, and count the number of detected events. Figure. 4.11 shows the recovered values of $p_{\rm det}$. They match with the error of a few percent, and we, therefore, conclude that our estimation of $\mathcal{D}^{95\%}$ is reliable.

4.7.4 Applicability of our method

In this thesis, the deep learning methods are demonstrated with a bit simplified setup. The training data, which is used for the training of the neural network, consists of chunks containing the signals. When the signals are generated, we pick up a grid point as a reference. We uniformly sample the locations of signals from the region where the reference

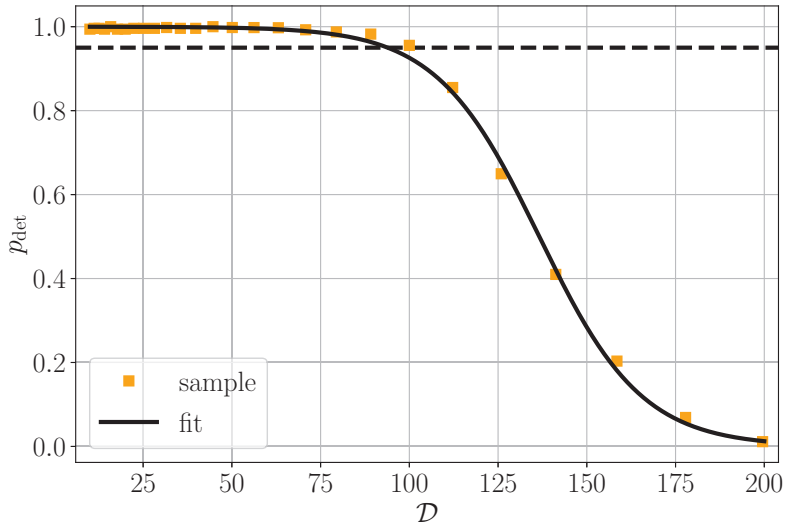

Figure 4.9: An example of the fitting procedure. Orange squares are the samples of the detection probabilities. Black solid line is the fitted curve. Black dashed line indicates $p_{\mathrm{det}} = 0.95$. The intersection between the black solid line and the black dashed line corresponds to the point of $\mathcal{D}^{95\%}$.

grid point covers. The gravitational wave frequencies of our signals are uniformly sampled from

$$f_{\mathrm{gw}} \in \left[100 - \frac{1}{2T_{\mathrm{seg}}}, 100 + \frac{1}{2T_{\mathrm{seg}}}\right] \ \mathrm{Hz}\,. \tag{4.118}$$

These two assumptions seem to make the applicable situation is strictly limited. Nevertheless, we expect our method can be applied for all-sky searches of monochromatic sources of $f_{\mathrm{gw}} \lesssim 100$ Hz for the following reasons. The important point of our method is that the signal power is concentrated into a few thousand bins in an ℓ-domain signal of $O(10^6)$ data points. The width of the localization does not become worse as long as we consider lower frequency. The width of the localization depends on the strength of the residual phase X, which can be written as

$$X \sim 2\pi f_{\mathrm{gw}} \frac{R_{\mathrm{ES}}}{c} \delta\theta\,. \tag{4.119}$$

This expression leads that if the gravitational wave frequency becomes lower, the signal tends to localize within a narrower region in ℓ-domain. Therefore, lowering the gravitational wave frequency does not affect the efficiency of the excess power method. As for the source location, the phase modulation and the antenna pattern depend on it. But the phase modulation is partly subtracted by the time resampling procedure, and typical values of $\delta\theta$ appearing in X are almost the same for any source locations. The effect on the antenna pattern function changes only the shape of the ℓ-domain signal and does not affect the localization of the signal power in the ℓ-domain. Because of these reasons, the

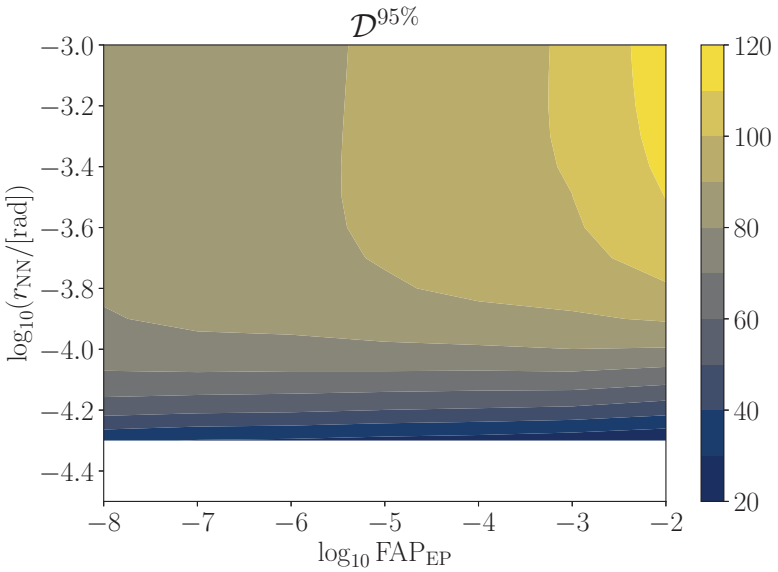

Figure 4.10: Estimated values of $\mathcal{D}^{95\%}$. In white blank region, the detection probability does not exceed 95% with the range of the amplitude we varied. Combining with Fig 4.7, one can estimate plausible values of $(\rm FAP_{EP}, r_{NN})$.

demonstration of the deep learning methods can be extrapolated into the all-sky search of monochromatic sources of $f_{\rm gw} \lesssim 100$ Hz. Of course, the neural network used in this demonstration cannot be directly applied to extended situations. The model of the neural network is needed to be fine-tuned with more reasonably generated datasets. Moreover, the structure of the neural network may be required to be changed. Nevertheless, our setup in this thesis is enough to provide the proof-of-principle.

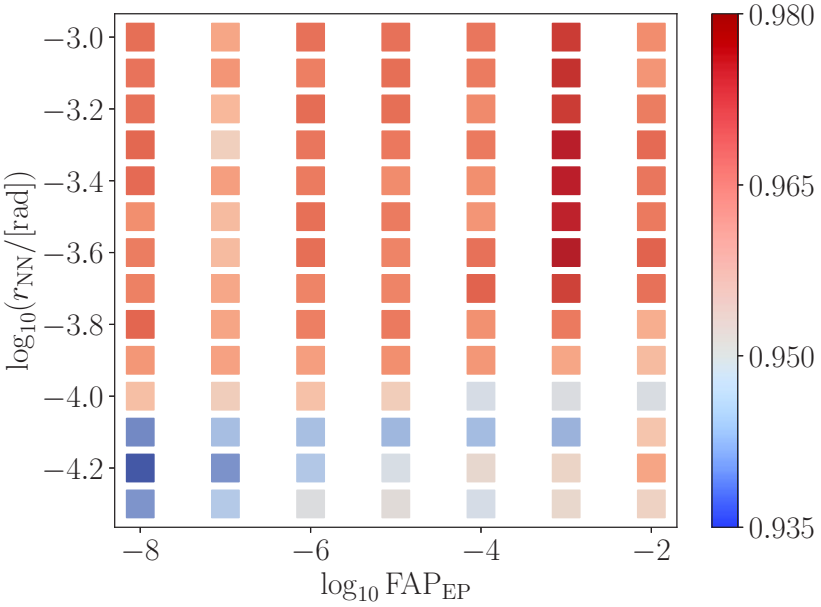

Figure 4.11: Recovered detection probabilities. For every combinations of $(\mathrm{FAP}_{\mathrm{EP}}, r_{\mathrm{NN}})$, the recovered detection probability coincides with 95% in a few percent error.

Chapter 5

Summary and future work

5.1 Summary

We proposed a new detection method of continuous gravitational waves by combining the excess power method and the deep learning method. First, the strain data is resampled with the new time coordinate, which is defined for each grid point on the sky. Then, a set of time-resampled strain is obtained. Second, the short-time Fourier transform is applied to each resampled strain, and the Fourier transform of the time series data in each frequency bin is taken. We obtain an ℓ-domain signal for each grid point for each frequency of STFT. Third, the excess power method is applied to each ℓ-domain signal. Forth, the deep learning method is employed to constrain the possible location of the source further. The chunk whose excess power exceeds the threshold value is used as the input of the artificial neural network, and the ANN estimates the source location. Finally, as the follow-up analysis, the coherent matched filtering is applied to all candidates.

We examined the computational cost and the detection power of our method. We assume that the single detector with a duty cycle of 100%, and the stationary and Gaussian noise. The sources are assumed to be monochromatic, located within the only vicinity of the reference grid point, and have the frequency is within a narrow frequency band. Also, the extrinsic parameters are assumed to have specific values and a priori known, that is, $\psi = 0$, $\cos\iota = 1$, and ϕ_0. The false alarm probability and the detection probability of the excess power method, and the coherent follow-up search are analytically estimated. As for the deep learning method, the artificial neural network is applied to the mock data generated by modeling ℓ-domain signals and injected them into Gaussian noise data. The expected computational time to analyze all candidates is estimated from the elapsed time to process ten thousand mock data. As a result, the computational time for the neural network is negligibly small if we appropriately set the false alarm probability of the excess power. The total computational cost is dominated by those of the preprocessing and the follow-up analysis. The parameters characterizing our method, $(\mathrm{FAP}_{\mathrm{EP}}, r_{\mathrm{NN}})$, is restricted by the total computational cost and the available computational power. The detection power is estimated in terms of the sensitivity depth. We calculate the sensitivity depth

with the detection probability of 95% for various pairs of $(\mathrm{FAP}_{\mathrm{EP}}, r_{\mathrm{NN}})$. If we assume that a 1TFlops machine is available, the sensitivity depth achieves ~ 90. Therefore, the new method we proposed can become an alternative method for the all-sky search of continuous gravitational waves.

5.2 Future work

Besides employing more realistic conditions of the detectors (e.g., non-Gaussianity and non-stationarity of the detector noise, the detector's duty cycle, use of strains from multiple detectors), there are two works remaining as future works.

In our work, the artificial neural network is trained by using the dataset generated by the limited parameter space. Namely, their source locations are uniformly sampled from the vicinity of the reference grid point, and their gravitational wave frequencies are uniformly sampled from a very narrow band (see Eq. (4.118)). Therefore, without some modifications, the neural network cannot be applied to the all-sky and broader frequency band. One naive modification is preparing many neural networks and training each network by a dataset corresponding to the reference grid point and the reference frequency. However, it is computationally too expensive. A less expensive way is to prepare one trained neural network as the reference and train the network with a new dataset. This process is called fine-tuning and can reduce the computational cost to train the neural network. Another way is to modify the neural network structure to make it applicable to all-sky and a wide frequency band. This approach requires only one neural network, while this requires more training data than the presented case.

In this work, the source is assumed to be monochromatic. In order to apply our method to more comprehensive types of waveforms, the time derivative of the frequency $\dot{f}$ should be taken into account in the waveform model. With such a waveform model, the track on the spectrogram that indicates the time evolution of the gravitational wave frequency would sweep inside the frequency bin. For a worse case, the frequency track would cross several bins. In such a case, our method loses its detection power. It remains as future works to extend our method to the situation where the frequency change is not negligible.

Acknowledgement

First, I would like to express my sincere gratitude to Takahiro Tanaka for his invaluable advice, continuous support, and patience during my Ph.D. study. I am grateful to the staff in the Theoretical Astrophysics group (TAP), Takashi Hosokawa, Koutarou Kyutoku, Naoki Seto, and Yoshiyuki Yamada, for their mentorship. Furthermore, I would also like to thank all members of the TAP group and the astrophysics group in the Yukawa Institute for Theoretical Physics (YITP) for having fruitful discussions and suggestions. A GPU machine at the Nagaoka University of Technology partly does the calculation of this work. I am thankful to Katsuko Nakahira for technical support.

I am deeply grateful to my parents, Hidehiro and Fumiko, for their great support and for allowing me to study at Kyoto University. I also thank the Yoshida dormitory for providing me with inexpensive housing and a chance to interact with people with various thoughts, beliefs, and careers. Lastly, I would like to say thank you to Manami Iwamoto for making my life happy.

Appendix A

Derivation of antenna pattern function

A.1 Detected strain

Gravitational waves have two polarization, plus and cross modes. The metric perturbation tensor can be written as the sum of these modes, i.e.,

$$h_{ij} = h_+ e^+_{ij} + h_\times e^\times_{ij}\,, \tag{A.1}$$

where e^+_{ij} and $e^\times_{ij}$ are the polarization tensors of the plus and the cross modes, respectively. Here, we assume that the gravitational waves propagate along with the direction of the vector k_i. The tensors e^+_{ij} and $e^\times_{ij}$ are symmetric,

$$e^+_{ij} = e^+_{ji}\,,\ e^\times_{ij} = e^\times_{ji}\,, \tag{A.2}$$

traceless,

$$\delta^{ij} e^+_{ij} = \delta^{ij} e^\times_{ij} = 0\,, \tag{A.3}$$

and transverse to the wave vector $\boldsymbol{k}$,

$$k^i e^+_{ij} = k^i e^\times_{ij} = 0\,. \tag{A.4}$$

The projection tensor $\Lambda_{ij,kl}$ can be rewritten as

$$\Lambda_{ij,kl} = \frac{1}{2}(e^+_{ij}e^+_{kl} + e^\times_{ij}e^\times_{kl})\,. \tag{A.5}$$

Therefore, the quadrupole formula becomes

$$h_{ij} = \frac{1}{2}(e^+_{ij}e^+_{kl} + e^\times_{ij}e^\times_{kl})\frac{2G}{rc^4}\ddot{Q}^{kl}\,, \tag{A.6}$$

with the quadrupole moment Q_{kl}. With Eq. (A.1), the plus and cross mode of gravitational waves are

$$h_+ = \frac{G}{rc^4} e^+_{ij}\ddot{Q}^{ij}\,,\ h_\times = \frac{G}{rc^4} e^\times_{ij}\ddot{Q}^{ij}\,. \tag{A.7}$$

We define two unit vectors $\hat{p}^i_{\rm A}$ and $\hat{p}^i_{\rm B}$ to represent the directions of the two-arms of the interferometer. The vector $w_i := \epsilon_{ijk}\hat{p}^j_{\rm A}\hat{p}^k_{\rm B}$ directed to the outward from the Earth's surface. The detector tensor d^{ij} is defined by

$$d^{ij} := \frac{1}{2}(\hat{p}^i_{\rm A}\hat{p}^j_{\rm A} - \hat{p}^i_{\rm B}\hat{p}^j_{\rm B})\,. \tag{A.8}$$

The observed signal h is obtained by the contraction between the perturbation tensor h_{ij} and the detector tensor d_{ij},

$$h = h_{ij}d^{ij} = h_+ e^+_{ij}d^{ij} + h_\times e^\times_{ij}d^{ij}\,. \tag{A.9}$$

The antenna pattern functions, F_+ and $F_\times$, are given by

$$F_+ := e^+_{ij}d^{ij}\,,\ F_\times := e^\times_{ij}d^{ij}\,. \tag{A.10}$$

Here, we introduce three frames. *The source frame* is the frame where it is easy to calculate the quadrupole tensor. *The wave frame* is the frame where the polarization tensors can be written as

$$e^+_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,,\ e^\times_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,. \tag{A.11}$$

Therefore, in the wave frame, the metric perturbation is expressed by

$$h_{ij} = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix}\,. \tag{A.12}$$

The detector frame is the frame where the vectors $\hat{p}^i_{\rm A}$ and $\hat{p}^i_{\rm B}$ can be written as

$$\hat{p}^i_{\rm A} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\,,\ \hat{p}^i_{\rm B} = \begin{pmatrix} \cos\zeta \\ \sin\zeta \\ 0 \end{pmatrix}\,, \tag{A.13}$$

where ζ is the angle between two arms of the detector. To calculate Eqs. (A.9) and (A.10), we need to carry out the coordinate transformation appropriately.

A.2 From source frame to wave frame

The Cartesian coordinates of the source frame and the wave frame are denoted by $\{x_{\rm s}, y_{\rm s}, z_{\rm s}\}$ and $\{x_{\rm w}, y_{\rm w}, z_{\rm w}\}$, respectively. The transformation from the source frame to the wave frame is characterized by the initial phase of the gravitational waves ϕ_0 and the inclination angle ι. The relation between two coordinates, $\{x_{\rm s}, y_{\rm s}, z_{\rm s}\}$ and $\{x_{\rm w}, y_{\rm w}, z_{\rm w}\}$, is written as

$$\begin{pmatrix} x_{\rm w} \\ y_{\rm w} \\ z_{\rm w} \end{pmatrix} = R_1(\iota)\cdot R_3(\phi_0)\cdot \begin{pmatrix} x_{\rm s} \\ y_{\rm s} \\ z_{\rm s} \end{pmatrix}\,, \tag{A.14}$$

where $R_I(\theta)$ for $I = 1, 2, 3$ is the rotation matrix representing the rotation around the I-th axis by the angle θ. Namely,

$$R_1(\theta) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} , \tag{A.15}$$

$$R_2(\theta) := \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} , \tag{A.16}$$

$$R_3(\theta) := \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{A.17}$$

The inclination angle is the angle between the line of sight and the $z_{\rm s}$-axis. The source frame is taken so that the calculation of the (time derivative of the) quadrupole moment can be easily done. Assuming that the time derivative of the quadrupole moment in the source frame is written as

$$\ddot{Q} = -(2\pi f)^2 \mathcal{Q} \begin{pmatrix} \cos 2\pi ft & \sin 2\pi ft & 0 \\ \sin 2\pi ft & -\cos 2\pi ft & 0 \\ 0 & 0 & 0 \end{pmatrix} , \tag{A.18}$$

the plus and the cross modes of the gravitational waves are

$$h_+ = h_0 \frac{1+\cos^2\iota}{2} \cos(2\pi ft - 2\phi_0) , \ h_\times = h_0 \cos\iota \sin(2\pi ft - 2\phi_0) , \tag{A.19}$$

with

$$h_0 := \frac{2G\mathcal{Q}}{rc^4} (2\pi f)^2 . \tag{A.20}$$

A.3 Celestial sphere frame

The celestial sphere frame $\{x_{\rm c}, y_{\rm c}, z_{\rm c}\}$ is useful to take into account the geometrical information of the source location. Therefore, from the wave frame, we go to the detector frame after moving to the celestial sphere frame. In the celestial sphere frame, the $x_{\rm c}$-axis points towards the vernal equinox and the $z_{\rm c}$-axis is directed to the spin angular momentum of the Earth.

The transform from the wave frame to the celestial sphere frame is determined by three angles, the polarization angle ψ, the right ascension of the source α and the declination of the source δ. The polarization angle ψ represents the relation between the wave frame and the Earth. The right ascension α and the declination δ represent the source location

with respect to the Earth's center. The unit vector pointing from the center of the Earth to the position of the source is denoted by $\boldsymbol{n}$ and it can be written in

$$\boldsymbol{n} = \begin{pmatrix} \cos\delta\cos\alpha \\ \cos\delta\sin\alpha \\ \sin\delta \end{pmatrix} , \tag{A.21}$$

in the celestial sphere frame. The transformation matrix from the wave frame to the celestial sphere frame is denoted by $M^{(1)}$ and the metric perturbation in the celestial sphere frame is

$$h_{ij}^{\mathrm{C}} = M_{ik}^{(1)} h_{kl} M_{jl}^{(1)} . \tag{A.22}$$

First, the polarization angle ψ is measured from the line of node to the x_{w}-axis counter-clockwise around z_{w}-axis (see Fig. A.1). This transformation is represented by the matrix

$$R_3(-\psi) = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{A.23}$$

The coordinate after the transformation is denoted by $\{x_1, y_1, z_1\}$. Second, we rotate the axis around the x_1-axis to make the z_1-axis pointing to the North celestial pole. From Fig.A.2, the rotation matrix is written in

$$R_1(3\pi/2-\delta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(3\pi/2-\delta) & \sin(3\pi/2-\delta) \\ 0 & -\sin(3\pi/2-\delta) & \cos(3\pi/2-\delta) \end{pmatrix} . \tag{A.24}$$

The new coordinate is denoted by $\{x_2, y_2, z_2\}$. Third, we rotate the frame around the z_2-axis so that the x_2-axis points towards the vernal equinox (see Fig. A.3). The rotation matrix is

$$R_3(\pi/2-\alpha) = \begin{pmatrix} \cos(\pi/2-\alpha) & \sin(\pi/2-\alpha) & 0 \\ -\sin(\pi/2-\alpha) & \cos(\pi/2-\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{A.25}$$

Finally, the matrix transforming from the wave frame $\{x_{\mathrm{w}}, y_{\mathrm{w}}, z_{\mathrm{w}}\}$ to the celestial sphere frame $\{x_{\mathrm{c}}, y_{\mathrm{c}}, z_{\mathrm{c}}\}$ is obtained as

$$M^{(1)} = R_3(\pi/2-\alpha) \cdot R_1(3\pi/2-\delta) \cdot R_3(\psi) . \tag{A.26}$$

A.4 Detector frame

The detector is located on the Earth and is co-rotating with the Earth's rotation. The longitude L of the detector's site is measured from the Greenwich Meridian to the local

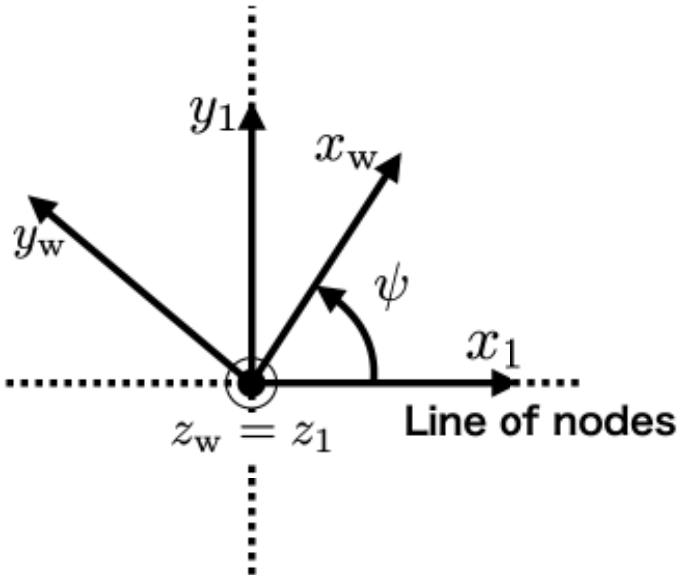

Figure A.1: Definition of the polarization angle ψ.

meridian with counter-clockwise around the North celestial pole. In the celestial sphere coordinate, the right ascension of the detector's site can be obtained by

$$\alpha_{\rm det}(t) = \Omega_{\oplus}(t - t_0) + L\,, \tag{A.27}$$

where t_0 is the reference time when the meridian of longitude 0° is at sidereal midnight and $\Omega_{\oplus}$ is the Earth's rotation frequency. We rotate the frame around the $z_{\rm c}$-axis by the angle $\alpha_{\rm det}(t)$ and the corresponding rotation matrix is

$$R_3(\alpha_{\rm det}(t)) = \begin{pmatrix} \cos(\Omega_{\oplus}(t-t_0)+L) & \sin(\Omega_{\oplus}(t-t_0)+L) & 0 \\ -\sin(\Omega_{\oplus}(t-t_0)+L) & \cos(\Omega_{\oplus}(t-t_0)+L) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{A.28}$$

By this rotation, the celestial sphere coordinate $\{x_{\rm c}, y_{\rm c}, z_{\rm c}\}$ goes to the new coordinate $\{x_3, y_3, z_3\}$ (see Fig. A.4). In the new coordinate, the x_3-axis points to the intersection between the celestial equator and the great circle passing through the poles and the detector's site.

The latitude λ of the detector's site is measured from the celestial equator to the North celestial pole and takes a value in $[-\pi/2, \pi/2]$ (see Fig. A.5. We rotate the frame around the y_3-axis so that the z_3-axis becomes parallel the direction from the Earth's center to the detector's site. This rotation is represented by

$$R_2(\pi/2 - \lambda) = \begin{pmatrix} \cos(\pi/2-\lambda) & 0 & -\sin(\pi/2-\lambda) \\ 0 & 1 & 0 \\ \sin(\pi/2-\lambda) & 0 & \cos(\pi/2-\lambda) \end{pmatrix}. \tag{A.29}$$

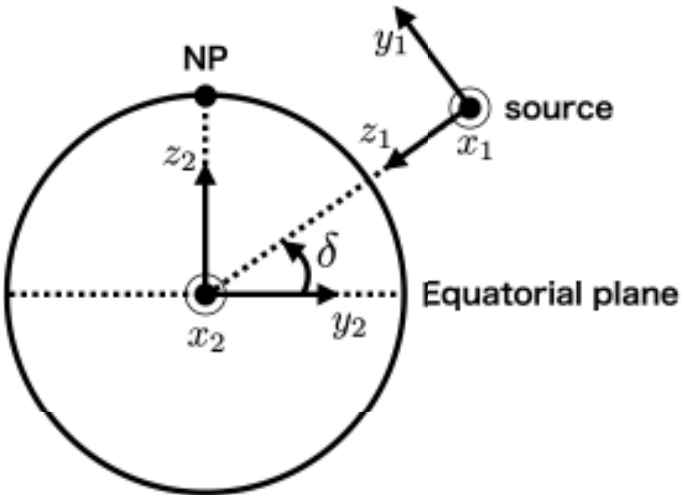

Figure A.2: Definition of the declination angle δ. "NP" is short for the North Pole. The figure is shown in the plane defined by the center of the Earth, the North pole, and the source location.

The new coordinate is called the cardinal coordinate and is denoted by $\{x_4, y_4, z_4\}$. The transformation matrix from the celestial sphere coordinate to the cardinal coordinate is

$$M^{(2)} := R_2(\pi/2 - \lambda) \cdot R_3(\alpha_{\rm det}(t)) \,. \tag{A.30}$$

In the detector coordinate $\{x_{\rm d}, y_{\rm d}, z_{\rm d}\}$, the $(x_{\rm d}, y_{\rm d})$-plane is tangent to the Earth's surface and the $x_{\rm d}$-axis along to one detector's arm. The another arm lies in the first or second quadrant. The unit vectors along with two arms are written in

$$\hat{\boldsymbol{p}}_{\rm A} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \,,\; \hat{\boldsymbol{p}}_{\rm B} = \begin{pmatrix} \cos\zeta \\ \sin\zeta \\ 0 \end{pmatrix} \,, \tag{A.31}$$

where $\zeta \in [0, \pi]$ is the angle between the two arms. The mutual position of the detector's arms are specified by the angle γ which is measured counter-clockwise from East to the bisector of the detector's arms. From Fig. A.6, the transformation from the cardinal coordinate and the detector coordinate is

$$M^{(3)} := R_3(\gamma - \zeta/2 + \pi/2) = \begin{pmatrix} \cos(\gamma - \zeta/2 + \pi/2) & \sin(\gamma - \zeta/2 + \pi/2) & 0 \\ -\sin(\gamma - \zeta/2 + \pi/2) & \cos(\gamma - \zeta/2 + \pi/2) & 0 \\ 0 & 0 & 1 \end{pmatrix} \,. \tag{A.32}$$

Finally, we obtain the matrix transforming from the wave coordinate to the detector coordinate is

$$M := M^{(3)} \cdot M^{(2)} \cdot M^{(1)} \,. \tag{A.33}$$

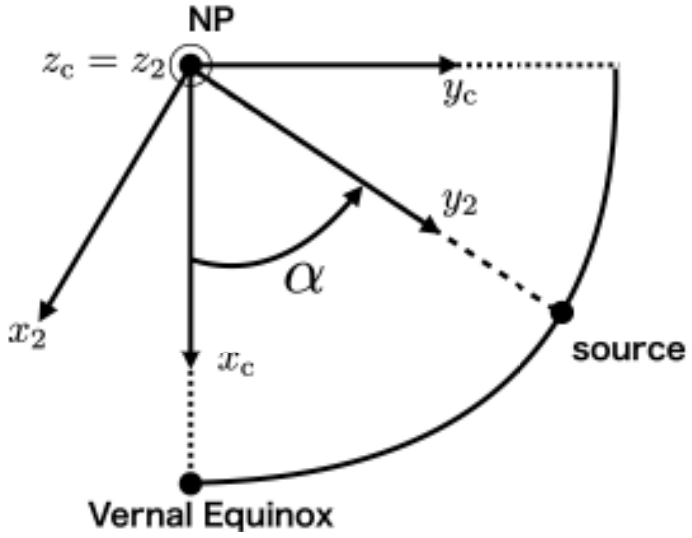

Figure A.3: Definition of the right ascension α.

A.5 The antenna pattern function

The polarization tensor, which is written as Eq. (A.11) in the wave coordinate, is transformed into the detector coordinate by

$$e^+ = M \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot M^{\mathrm{T}}\,,\ e^\times = M \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot M^{\mathrm{T}}\,, \tag{A.34}$$

The explicit forms of the antenna pattern functions are

$$F_+(t) = \sin\zeta[a(t)\cos 2\psi + b(t)\sin 2\psi]\,, \tag{A.35}$$
$$F_\times(t) = \sin\zeta[b(t)\cos 2\psi - a(t)\sin 2\psi]\,, \tag{A.36}$$

with

$$\begin{aligned} a(t) =& \frac{1}{16}\sin 2\gamma(3-\cos 2\lambda)(3-\cos 2\delta)\cos[2(\alpha-\varphi_\oplus-\Omega_\oplus t)] \\ &- \frac{1}{4}\cos 2\gamma\sin\lambda(3-\cos 2\delta)\sin[2(\alpha-\varphi_\oplus-\Omega_\oplus t)] \\ &+ \frac{1}{4}\sin 2\gamma\sin 2\lambda\sin 2\delta\cos[\alpha-\varphi_\oplus-\Omega_\oplus t] \\ &- \frac{1}{2}\cos 2\gamma\cos\lambda\sin 2\delta\sin[\alpha-\varphi_\oplus-\Omega_\oplus t] \\ &+ \frac{3}{4}\sin 2\gamma\cos^2\lambda\cos^2\delta\,, \end{aligned} \tag{A.37}$$

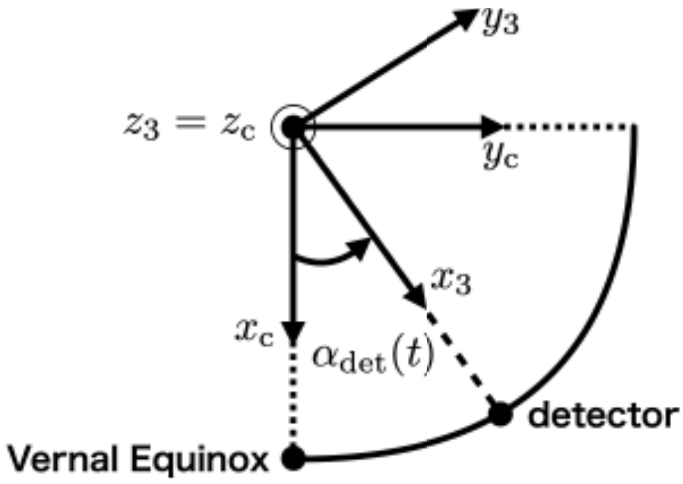

Figure A.4: Definition of the detector longitude L.

and

$$\begin{aligned} b(t) = & \cos 2\gamma \sin\lambda \sin\delta \cos[2(\alpha - \varphi_\oplus - \Omega_\oplus t)] \\ & + \frac{1}{4} \sin 2\gamma (3 - \cos 2\lambda) \sin\delta \sin[2(\alpha - \varphi_\oplus - \Omega_\oplus t)] \\ & + \cos 2\gamma \cos\lambda \cos\delta \cos[\alpha - \varphi_\oplus - \Omega_\oplus t] \\ & + \frac{1}{2} \sin 2\gamma \sin 2\lambda \cos\delta \sin[\alpha - \varphi_\oplus - \Omega_\oplus t] \,. \end{aligned} \tag{A.38}$$

Here, we define $\varphi_\oplus := -\Omega_\oplus t_0 + L$.

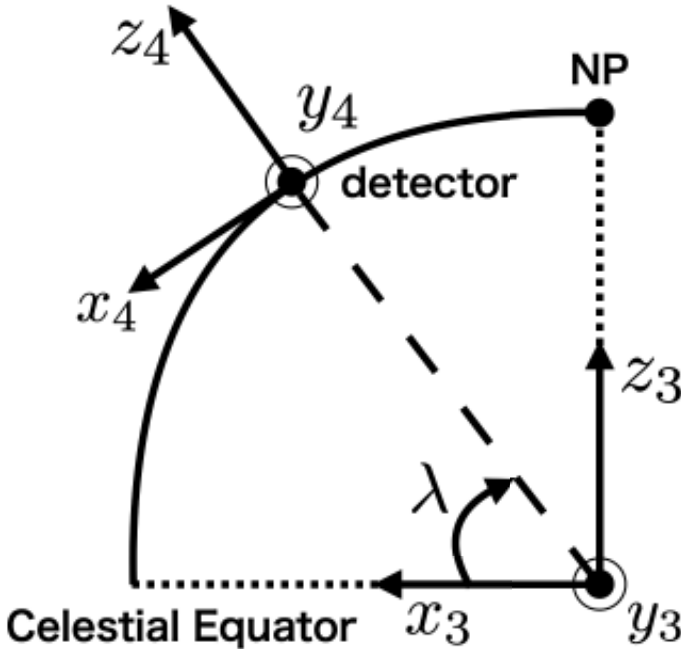

Figure A.5: Definition of the detector's latitude λ

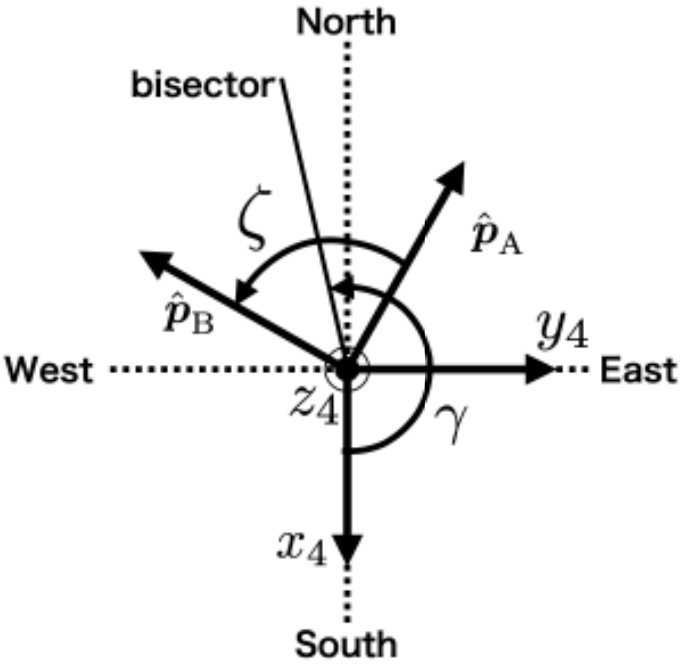

Figure A.6: Definition of the detector frame. The angle γ is the angle measured from the North-South axis to the bisector of the detector's arms. The angle ζ is the angle between two arms.

Bibliography

[1] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc] .

[2] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE] .

[3] R. Abbott et al. (LIGO Scientific, Virgo), "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run," (2020), arXiv:2010.14527 [gr-qc] .

[4] B. P. Abbott et al. (LIGO Scientific, Virgo), Astrophys. J. Lett. **818**, L22 (2016), arXiv:1602.03846 [astro-ph.HE] .

[5] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 221101 (2016), [Erratum: Phys.Rev.Lett. 121, 129902 (2018)], arXiv:1602.03841 [gr-qc] .

[6] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **119**, 161101 (2017), arXiv:1710.05832 [gr-qc] .

[7] B. P. Abbott et al. (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT), Astrophys. J. Lett. **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE] .

[8] B. P. Abbott et al. (LIGO Scientific, Virgo, Fermi-GBM, INTEGRAL), Astrophys. J. Lett. **848**, L13 (2017), arXiv:1710.05834 [astro-ph.HE] .

[9] B. P. Abbott et al. (LIGO Scientific, Virgo, 1M2H, Dark Energy Camera GW-E, DES, DLT40, Las Cumbres Observatory, VINROUGE, MASTER), Nature **551**, 85 (2017), arXiv:1710.05835 [astro-ph.CO] .

[10] R. Abbott et al. (LIGO Scientific, Virgo), "Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog," (2020), arXiv:2010.14533 [astro-ph.HE] .

[11] T. Akutsu et al. (KAGRA), "Overview of KAGRA: Detector design and construction history," (2020), arXiv:2005.05574 [physics.ins-det] .

[12] T. Akutsu et al. (KAGRA), "Overview of KAGRA : KAGRA science," (2020), arXiv:2008.02921 [gr-qc] .

[13] B. Iyer et al., "LIGO-India, Proposal of the Consortium for Indian Initiative in Gravitational-wave Observations (IndIGO)," .

[14] B. P. Abbott et al. (KAGRA, LIGO Scientific, VIRGO), Living Rev. Rel. **21**, 3 (2018), arXiv:1304.0670 [gr-qc] .

[15] M. Punturo et al., Class. Quant. Grav. **27**, 194002 (2010).

[16] M. Punturo et al., Class. Quant. Grav. **27**, 084007 (2010).

[17] D. Reitze et al., Bull. Am. Astron. Soc. **51**, 035 (2019), arXiv:1907.04833 [astro-ph.IM] .

[18] P. Amaro-Seoane et al. (LISA), (2017), arXiv:1702.00786 [astro-ph.IM] .

[19] J. Baker et al., "The Laser Interferometer Space Antenna: Unveiling the Millihertz Gravitational Wave Sky," (2019), arXiv:1907.06482 [astro-ph.IM] .

[20] W.-H. Ruan, Z.-K. Guo, R.-G. Cai, and Y.-Z. Zhang, Int. J. Mod. Phys. A **35**, 2050075 (2020), arXiv:1807.09495 [gr-qc] .

[21] J. Luo et al. (TianQin), Class. Quant. Grav. **33**, 035010 (2016), arXiv:1512.02076 [astro-ph.IM] .

[22] H. Omiya and N. Seto, Phys. Rev. D **102**, 084053 (2020), arXiv:2010.00771 [gr-qc] .

[23] N. Seto, S. Kawamura, and T. Nakamura, Phys. Rev. Lett. **87**, 221103 (2001), arXiv:astro-ph/0108011 .

[24] G. Hobbs et al., Class. Quant. Grav. **27**, 084013 (2010), arXiv:0911.5206 [astro-ph.SR] .

[25] Z. Arzoumanian et al. (NANOGrav), (2020), arXiv:2009.04496 [astro-ph.HE] .

[26] B. S. Sathyaprakash and B. F. Schutz, Living Rev. Rel. **12**, 2 (2009), arXiv:0903.0338 [gr-qc] .

[27] M. Sieniawska and M. Bejger, Universe **5**, 217 (2019), arXiv:1909.12600 [astro-ph.HE] .

[28] D. R. Lorimer and M. Kramer, Handbook of pulsar astronomy (2012).

[29] B. F. Schutz, "The Detection of gravitational waves," (Cambridge University Press, 1991) pp. 406–452.

[30] P. R. Brady and T. Creighton, Phys. Rev. D **61**, 082001 (2000), arXiv:gr-qc/9812014 .

[31] B. Krishnan, A. M. Sintes, M. A. Papa, B. F. Schutz, S. Frasca, and C. Palomba, Phys. Rev. D **70**, 082001 (2004), arXiv:gr-qc/0407001 .

[32] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. D **96**, 062002 (2017), arXiv:1707.02667 [gr-qc] .

[33] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. D **97**, 102003 (2018), arXiv:1802.05241 [gr-qc] .

[34] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. D **100**, 024004 (2019), arXiv:1903.01901 [astro-ph.HE] .

[35] T. S. Yamamoto and T. Tanaka, "Use of Excess Power Method and Convolutional Neural Network in All-Sky Search for Continuous Gravitational Waves," (2020), arXiv:2011.12522 [gr-qc] .

[36] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, Phys. Rev. D **81**, 123530 (2010), arXiv:0905.4720 [hep-th] .

[37] R. Brito, V. Cardoso, and P. Pani, Class. Quant. Grav. **32**, 134001 (2015), arXiv:1411.0686 [gr-qc] .

[38] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, Class. Quant. Grav. **35**, 063001 (2018), arXiv:1801.05235 [astro-ph.CO] .

[39] P. Jaranowski, A. Krolak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998), arXiv:gr-qc/9804014 .

[40] K. Jotania, S. Valluri, and S. V. Dhurandhar, Astron. Astrophys. **306**, 317 (1996).

[41] S. Valluri, V. Dergachev, X. Zhang, and F. Chishtie, "The Fourier Transform of the Continuous Gravitational Wave Signal," (2020), arXiv:2010.13647 [gr-qc] .

[42] H. Nakano, H. Takahashi, H. Tagoshi, and M. Sasaki, Phys. Rev. D **68**, 102003 (2003), arXiv:gr-qc/0306082 .

[43] W. G. Anderson, P. R. Brady, J. D. Creighton, and E. E. Flanagan, Phys. Rev. D **63**, 042003 (2001), arXiv:gr-qc/0008066 .

[44] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning (MIT Press, 2016) `http://www.deeplearningbook.org`.

[45] W. S. McCulloch and W. Pitts, Bulletin of Mathematical Biographysics **5**, 115 (1943).

[46] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun, in 2009 IEEE 12th International Conference on Computer Vision (ICCV) (IEEE Computer Society, Los Alamitos, CA, USA, 2009) pp. 2146–2153.

[47] K. Hornik, Neural Networks **4**, 251 (1991).

[48] G. Cybenko, Mathematics of Control, Signals and Systems **2**, 303 (1989).

[49] Y. LeCun, "Generalization and network design strategies," in Connectionism in perspective, edited by R. Pfeifer, Z. Schreter, F. Fogelman, and L. Steels (Elsevier, 1989).

[50] Y. Zhou and R. Chellappa, in IEEE 1988 International Conference on Neural Networks (1988) pp. 71–78.

[51] B. Polyak, USSR Computational Mathematics and Mathematical Physics **4**, 1 (1964).

[52] G. Hinton, "Neural networks for machine learning," (2012).

[53] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," (2014), arXiv:1412.6980 [cs.LG] .

[54] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature **323**, 533 (1986).

[55] D. George and E. A. Huerta, Phys. Rev. D **97**, 044039 (2018), arXiv:1701.00008 [astro-ph.IM] .

[56] A. Khan, E. A. Huerta, and A. Das, Phys. Lett. B **808**, 0370 (2020), arXiv:2004.09524 [gr-qc] .

[57] W. Wei, A. Khan, E. A. Huerta, X. Huang, and M. Tian, Phys. Lett. B **812**, 136029 (2021), arXiv:2010.15845 [gr-qc] .

[58] W. Wei, E. A. Huerta, M. Yun, N. Loutrel, R. Haas, and V. Kindratenko, "Deep Learning with Quantized Neural Networks for Gravitational Wave Forecasting of Eccentric Compact Binary Coalescence," (2020), arXiv:2012.03963 [gr-qc] .

[59] D. George, H. Shen, and E. A. Huerta, Phys. Rev. D **97**, 101501(R) (2018).

[60] A. J. K. Chua, M. L. Katz, N. Warburton, and S. A. Hughes, "Rapid generation of fully relativistic extreme-mass-ratio-inspiral waveform templates for LISA data analysis," (2020), arXiv:2008.06071 [gr-qc] .

[61] H. Gabbard, C. Messenger, I. S. Heng, F. Tonolini, and R. Murray-Smith, "Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy," (2019), arXiv:1909.06296 [astro-ph.IM] .

[62] A. J. K. Chua and M. Vallisneri, Phys. Rev. Lett. **124**, 041102 (2020), arXiv:1909.05966 [gr-qc] .

[63] H. Nakano, T. Narikawa, K.-i. Oohara, K. Sakai, H.-a. Shinkai, H. Takahashi, T. Tanaka, N. Uchikata, S. Yamamoto, and T. S. Yamamoto, Phys. Rev. D **99**, 124032 (2019), arXiv:1811.06443 [gr-qc] .

[64] T. S. Yamamoto and T. Tanaka, "Use of conditional variational auto encoder to analyze ringdown gravitational waves," (2020), arXiv:2002.12095 [gr-qc] .

[65] C. Dreissigacker, R. Sharma, C. Messenger, R. Zhao, and R. Prix, Phys. Rev. D **100**, 044009 (2019), arXiv:1904.13291 [gr-qc] .

[66] C. Dreissigacker and R. Prix, Phys. Rev. D **102**, 022005 (2020), arXiv:2005.04140 [gr-qc] .

[67] B. Beheshtipour and M. A. Papa, Phys. Rev. D **101**, 064009 (2020), arXiv:2001.03116 [gr-qc] .

[68] F. Morawski, M. Bejger, and P. Ciecieląg, "Convolutional neural network classifier for the output of the time-domain F-statistic all-sky search for continuous gravitational waves," (2019), arXiv:1907.06917 [astro-ph.IM] .

[69] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, in Advances in Neural Information Processing Systems 32, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019) pp. 8024–8035.

[70] R. J. Dupuis and G. Woan, Phys. Rev. D **72**, 102002 (2005), arXiv:gr-qc/0508096 .

[71] K. Wette, Phys. Rev. D **85**, 042003 (2012), arXiv:1111.5650 [gr-qc] .

[72] C. Dreissigacker, R. Prix, and K. Wette, Phys. Rev. D **98**, 084058 (2018), arXiv:1808.02459 [gr-qc] .