Data analysis and Source modelling for LISA

Von der Fakultät für Mathematik und Physik

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

DOKTOR DER NATURWISSENSCHAFTEN

Dr. rer.nat.

genehmigte Dissertation

von

Yu Shang

geboren am 15.09.1980 in Kai Feng He Nan, China

2014

Referentin/Referent: Prof. Bernard F. Schutz

Korreferentin/Korreferent: Prof. Olaf Lechtenfeld, Dr. John Gair

Tag der Promotion: 19.12.2013

Data Analysis and Source Modelling for LISA

Yu Shang

Department of Physics

University of Hanover

A thesis submitted for the degree of

Doctor of Philosophy

Yet to be decided

I would like to dedicate this thesis to my loving family.

Acknowledgements

And I would like to acknowledge my advisors Prof. Bernard F. Schutz, Dr. Stanislav Babak, my colleagues Dr. Antoine Petiteau, my friend Yan Wang, and all the other members in the astrophysical relativity group. Without their kindly help, I can not finish my work successfully.

Abstract

The gravitational waves are one of the most important predictions in general relativity. Besides of the directly proof of the existence of GWs, there are already several ground based detectors (such as LIGO, GEO, etc) and the planed future space mission (such as: LISA) which are aim to detect the GWs directly. GW contain a large amount of information of its source, extracting these information can help us dig out the physical property of the source, even open a new window for understanding the Universe. Hence, GW data analysis will be a challenging task in seeking the GWs. In this thesis, I present two works about the data analysis for LISA.

In the first work, we introduce an extended multimodal genetic algorithm which utilizes the properties of the signal and the detector response function to analyze the data from the third round of mock LISA data challenge. We have found all five sources present in the data and recovered the coalescence time, chirp mass, mass ratio and sky location with reasonable accuracy. As for the orbital angular momentum and two spins of the Black Holes, we have found a large number of widely separated modes in the parameter space with similar maximum likelihood values. The performance of this method is comparable, if not better, to already existing algorithms.

In the second work, we introduce an new phenomenological waveform model for the extreme mass ratio inspiral system. This waveform consists of a set of harmonics with constant amplitude and slowly evolving phase which we decompose in a Taylor series. We use these phenomenological templates to detect the signal in the simulated data, and then, assuming a particular EMRI model, estimate the physical

parameters of the binary with high precision. The results show that our phenomenological waveform is very feasible in the data analysis of EMRI signal.

Abstract

Gravitationswellen sind eine der wichtigsten Vorhersagen in der Allgemeinen Relativitätstheorie. Der direkte Nachweis von GWs steht noch aus. Es existieren bereits erdgebundene Gravitationswellendetektoren wie LIGO und GEO. Außerdem sind Raummissionen wie LISA geplant, um Gravitationswellen nachzuweisen. GWs enthalten eine große Menge an Informationen über ihre Quelle. Diese Informationen können uns helfen, die physikalischen Eigenschaften der Quelle zu verstehen und sogar ein neues Fenster ins Universum zu öffnen. Die Daten-Analyse von GWs ist dabei die große Herausforderung. In dieser Thesis präsentiere ich zwei Arbeiten über Daten-Analyse für LISA. In der ersten Arbeit führen wir einen ausgedehnten multimodalen genetischen Algorithmus ein, der die Eigenschaften des Signals und der Detektor- Response- Funktion verwendet, um die Daten der MOCK-LISA-Data- Challenge zu analysieren. Wir haben alle fnf Quellen, die in den Daten vorhanden waren, gefunden und die Zeit der Verschmelzung, die Chirp-Masse, das Massenverhältnis und die Himmelsrichtung mit vernünftiger Genauigkeit detektiert. Genauso wie für den gemeinsamen Drehimpuls und die beiden Spins der Schwarzen Löcher haben wir eine große Zahl von weit separierten Moden im Parameterraum mit ähnlichen maximalen Likelihood-Werten gefunden. Die Effizienz der Methode ist vergleichbar, wenn nicht besser, als bereits existierende Methoden. In der zweiten Arbeit führen wir ein neues phänomenologische Wellenformmodell ein für EMRI Systeme (zwei Schwarze Löcher mit extremem Massenverhältnis). Diese Wellenform besteht aus einer Menge von Harmonischen mit konstanter Amplitude und sich langsam ändernder Phase, die wir in eine Taylor-Reihe zerlegen. Wir benutzen diese phänomenologische Vorlage, um das

Signal in den simulierten Daten zu detektieren und dann unter der Annahme eines speziellen EMRI-Modells die physikalischen Parameter des Binärsystems mit hoher Genauigkeit abzuschätzen. Die Resultate zeigen, dass unsere phänomenologische Wellenform sehr praktisch ist für die Daten-Analyse von EMRI Signalen.

Keywords: General relativity, Gravitational Waves, Data analysis,

Schlüsselwort: Allgemeinen Relativitätstheorie, Gravitationswellen, Daten-Analyse

Contents

Contents viii

List of Figures xii

1 Introduction 1

1.1 Basis of gravitational radiation . . . 1
1.2 GW detection . . . 4
1.2.1 Ground base detectors . . . 4
1.2.2 Space detector . . . 5
1.3 The GW sources . . . 9
1.3.1 The Source for The Ground Base Detector . . . 9
1.3.2 The Source for LISA . . . 9
1.3.2.1 Galactic binaries . . . 9
1.3.2.2 Spinning super massive black hole binaries . . . 11
1.3.2.3 Capture of the compact stellar mass by the super massive black hole . . . 14
1.3.2.4 Stochastic GW background . . . 17
1.4 Gravitational waves data analysis . . . 17
1.4.0.5 Hypothesis Test . . . 17
1.4.1 Matched filter . . . 19
1.4.2 Overlap . . . 21
1.4.3 Parameter estimation . . . 22
1.4.3.1 Algorithm for searching on the parameter space . 25
1.4.4 LISA data analysis and MLDC . . . 26
1.5 Structure of the thesis . . . 27

Nomenclature **1**

2 Basic Techniques of GW data analysis **32**
2.1 LISA Measurement 32
2.2 TDI technique 34
2.3 LISA response in low frequency limit 39
2.4 $\mathcal{F}$-statistic 41
2.5 Doppler modulation 44

3 The search for spinning black hole binaries in mock LISA data using a genetic algorithm **46**
3.1 My contribution 47
3.2 Formulation of the problem 47
3.2.1 Model of the template 47
3.2.1.1 Maximization of the likelihood 51
3.2.1.2 Maximization over the time of coalescence 53
3.2.1.3 The waveform termination 54
3.2.1.4 A-statistic 56
3.3 Genetic Algorithm 56
3.3.1 The basic principle 56
3.3.2 Code of the gene 59
3.3.3 Selection 60
3.3.4 Breeding 61
3.3.5 Mutation 61
3.3.6 Tuning the algorithm using code, breeding and mutation . 63
3.4 Acceleration of Genetic Algorithm 63
3.4.1 Standard accelerators 64
3.4.1.1 Elitism 64
3.4.1.2 Simulated annealing 64
3.4.1.3 Evolution of PMR 66
3.4.2 Accelerators specific for SMBH search 67
3.4.2.1 Brother 67
3.4.2.2 Local mutation 68

3.4.2.3 Fixing the significant bits . . . 69
3.4.2.4 Specific breeding and mutation . . . 70
3.4.2.5 Change of environment . . . 71
3.5 Multimodal search . . . 72
3.6 Pipeline . . . 75
3.6.1 Pre-analysis by a time frequency method . . . 76
3.6.2 Multiple steps MGA search . . . 77
3.7 Results . . . 79
3.7.1 MLDC results . . . 79
3.7.1.1 Coalescence within the observational time . . . 79
3.7.1.2 Coalescence beyond the observational time . . . 82
3.7.2 Post-MLDC results . . . 83
3.8 Summary . . . 86

4 EMRI Data Analysis With A Phenomenological Waveform 89
4.1 My Contribution . . . 89
4.2 Motivation . . . 90
4.3 Numerical kludge model . . . 93
4.3.1 The Kerr geodesic in observation time . . . 97
4.3.2 Changes of variables for radial and polar coordinates . . . 104
4.3.3 Adiabatic approximation . . . 105
4.3.4 Evolution of the constants of motion . . . 106
4.3.5 The truncation of the waveform . . . 110
4.4 Harmonic Analysis of the Numerical Kludge Waveform . . . 111
4.4.1 Fourier expansion of arbitrary function $f(r, \theta, \phi)$ . . . 112
4.4.2 Fourier expansion of $h^{+,\times}$ . . . 113
4.4.3 Truncation of the harmonic decomposition . . . 118
4.4.4 Exploring the harmonic content as function of the parameters 119
4.4.4.1 The influence of eccentricity . . . 120
4.4.4.2 The influence of a, p,ι . . . 120
4.4.4.3 The influence of Θ, Φ . . . 121
4.4.5 The class of the dominant harmonics . . . 122
4.4.6 The evolution of the dominant harmonics . . . 123

4.5 Phenomenological waveform family 125
4.5.1 $\mathcal{F}$-statistic in the case of phenomenological waveform . . . 127
4.6 Data analysis with phenomenological waveform 129
4.6.1 Search for Dominant harmonics 130
4.6.1.1 Review of MCMC and simulated annealing . . . 131
4.6.1.2 Analysis of MCMC results 134
4.6.2 Physical parameter searching 141
4.6.2.1 Define the quantity to assess the goodness of fit 141
4.6.2.2 Searching algorithm in time frequency plane . . . 143
4.7 Summary 146

5 Final Summary **150**

References **151**

List of Figures

1.1 Two polarization of GWs. The left picture show the effect on proper separations of particles in a circular ring due to a plus-polarized wave traveling in the perpendicular direction. The right picture show the case in a cross-polarized wave. The ring continuously gets deformed into one of the ellipses and back during the first half of a gravitational wave period and the same effect of deformation in the other orthogonal ellipse and back during the next half period. 3

1.2 The constellation of LISA. The center of this LISA constellation moves around the Sun in an earth-like orbit (R = 1AU), 20° behind the Earth. The constellation plane is inclined by 60° to the ecliptic. The triangular constellation revolves around its center as it completes one orbit in the course of one year. 6

1.3 LISA sensitivity curve. The red line is the instrumental threshold corresponding to sources with SNR = 1. The yellow region which is below the instrumental threshold embody the strength of the noise. Above the red line, it is the LISA discovery space. The broad areas marked indicate what types of sources are expected to populate each region of the discovery space. 8

2.1 Each spacecraft is identified by number, 1, 2, 3, and the opposing arm length is identified by the same index, $\{L_1, L_2, L_3\}$, it is often useful to identify the geometric point o as guiding center, which is equidistant from each of the three spacecraft. The distance l is usually called the guiding center distance. 35

3.1 Distribution over M_c and η, of the Maximized Likelihood (quality) computed with the full waveform on left panel and with the waveform cut at $f_{max} = 0.26$ on the right panel. This example corresponds to a signal with the following parameters: $\theta = -0.38896\ rad$, $\phi = 3.28992\ rad$, $t_c = 19706568.3273\ sec$, $M_c = 1589213.34\ M_\odot$, $\eta = 0.23647$, $\theta_L = 2.78243\ rad$, $\phi_L = 1.53286\ rad$, $\chi_1 = 0.24115$, $\chi_2 = 0.16145$, $\theta_{S1} = 1.20839\ rad$, $\phi_{S1} = 5.61808\ rad$, $\theta_{S2} = 0.39487\ rad$, $\phi_{S2} = 5.82937\ rad$, $D_L = 6856164697.8\ parsec$, $\phi_c = 4.96746\ rad$. The arrow points to the true parameters. . . . 55

3.2 Distribution of A-statistic over M_c and η. This example corresponds to the same signal as in Figure 3.1. The arrow points to the true location of parameters of the signal. 57

3.3 Examples of used breeding: cross-over one random point on the left, cross-over two fixed points in the middle and random on the right panels. 61

3.4 Example of distribution of 100 organisms in two dimensions space which is the sky position (θ,ϕ). The left panel shows the case of large PMR value (0.5) which corresponds to the beginning of the search. The right-panel shows the case of small PMR value (0.01) which corresponds to the end of the search. The best organism as well as the true solution is at the center of the blue circle. 62

3.5 Example of the chirp mass exploration by the colony of organisms. The green points correspond to the position of the best organism. The separations shows the structure of the binary representation. The numbers on the right are values at bit positions listed on top. 70

3.6 Example of accumulation of SNR (MBH-1) as a function of frequency. The points with attached numbers show the time until the coalescence in seconds. 71

3.7 Example of the distribution of the best organisms from 196 runs of GA applied to the search for source MBH-3 of MLDC 3.2 (the third signal). The left upper panel shows the initial direction of spin 1, the initial direction of spin 2 is in the right upper panel and we plot the initial direction of the orbital angular momentum in the bottom panel. The color scale corresponds to the value of A-statisitc. 73
3.8 Time-frequency representation of channel TDI A of MLDC 3.2. We plot the norm of the Morlet wavelets transformation. 77
3.9 Ratio between Maximized likelihood of modes and true Maximized Likelihood for source srcMC2. The blue dotted line corresponds to a ratio equal to one. The green dashed lines correspond to $\Delta SNR_i < 1$ 85

4.1 The cumulative F-statistic of 30 dominant harmonics with true parameters without noise. Since there is no noise, the F-statistic is not normalized. 135
4.2 Time-frequency plot of harmonics. The black and green tracks are two strong harmonics of the EMRI signal (black being stronger). The blue track corresponds to a harmonic of PW that accumulates a significant $\mathcal{F}$-statistic. It intersects the true harmonics at the pink segments, those correspond to times of increase of $\mathcal{F}$-statistic, see 4.3. 136
4.3 Cumulative $\mathcal{F}$-statistic corresponding to the situation depicted in Fig. 4.2. The green and red squares mark the maxima and minima of the curve, thus distinguishing between the increasing and the decreasing slopes. 137
4.4 Time-frequency plot of all patches corresponding to strong accumulation of F-statistic. We can identify parts of frequency tracks of 13 EMRI harmonics. Each track in this plot has a finite width coming from different solutions of MCMC search which have different precision of matching the signal. 138

4.5 Zoom at two harmonics at a specific instance of time. The red stems denote the frequencies of the true harmonics of a signal, while the blue histogram shows the detected frequencies at this instant. The green curves are the Gaussian fit to the frequency data with re-scaled amplitudes. The vertical axis of pink points indicates the relative time over which we have observed strong accumulation of $\mathcal{F}$-statistic for each solution. 139

Abbreviation

- GR: general relativity
- GW: gravitational waves
- TT: transverse traceless
- LISA: laser Interferometer space antenna
- MLDC: mock LISA data challenge
- WD: white dwarf
- SMBH: super massive black holes
- EMRI: extreme mass ratio inspiral
- IMRI: intermediate mass ratio inspiral
- SNR: signal noise ratio
- GA: genetic algorithm
- PW: phenomenological waveform
- MCMC: Markov chain Monte Carlo
- PSO: particle swarm optimization
- TDI: time delay interferometry
- PMR: probability mutation rate
- FIM: Fisher information matrix
- MGA: multimodal genetic algorithm
- AK: analytical kludge
- NK: numerical kludge
- PN: post Newtonian
- LSO: last stable orbit

Chapter 1

Introduction

1.1 Basis of gravitational radiation

In general relativity (GR), gravity is equivalent to the curvature of the space time. The changes of the gravitational field can be treated as the ripples in the curvature, in other words: gravitational waves (GW). The principle within GR of nothing travels faster than light means that these changes can not be felt everywhere instantaneously, the GWs should propagate at exactly the same speed as vacuum electromagnetic waves: the speed of light.

GR is a nonlinear theory and there is, in general, no sharp distinction between the part of the metric that represents the waves and the rest of the metric. Only in certain approximation such as: linearized theory, small perturbation of a smooth, time-independent background metric, we can clearly define gravitational radiation Maggiore [2007].

We follow the description of linearized theory as in Maggiore [2007]; Misner *et al.* [1974], and the Newton constant G and speed of light C are set to be 1. The linearized theory is a weak field approximation to GR. The filed equations are written and solved in a flat space time background. The weak filed can be treated as a small perturbation on the flat space time, and its metric tensor can be written as

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \quad |h_{\alpha\beta} \ll 1|, \alpha, \beta = 0, \cdots, 3 \tag{1.1}$$

where $\eta_{\alpha\beta}$ is the Minkowski metric, and $h_{\alpha\beta}$ is perturbation. Linearized theory is an approximation to GR, and is correct to first order in the case of this perturbation. There is considerable coordinate freedom in the components $h_{\alpha\beta}$, and it is called "gauge freedom".

We use this freedom to enforce the harmonic gauge

$$h^{\alpha\beta}{}_{,\beta} = 0. \tag{1.2}$$

In this gauge, the Einstein field equations can be written as

$$[\nabla^2 - \frac{\partial^2}{\partial t^2}]h_{\alpha\beta} = (\text{source}), \tag{1.3}$$

where "(source)" represents the various energy densities and stresses that can create the field. Tensor $h_{\alpha\beta}$ is symmetric, so it has ten independent components, and the harmonic gauge applies four independent condition for these, reducing the freedom to six Maggiore [2007].

In the region far away from the source, the GWs can be treated as propagating in the vacuum spacetime, the field equations in that region are source free. For linearized Einstein equation in vacuum, the Lorentz gauge does not fully fix the coordinate. In fact if we perform another infinitesimal coordinate transformation $(x^\mu \to x^\mu) + \xi^\mu$ with $\xi^\mu{}_{,\nu} = \mathcal{O}(h)$, and impose $\Box\xi^\mu = 0$, the Lorentz gauge is remained. We can use this freedom to demand :

$$\begin{aligned} h_{\alpha 0} &= 0 \quad \text{there are no time components,} \\ h_k{}^k &= 0 \quad \text{spatial components are trace free.} \end{aligned} \tag{1.4}$$

These conditions are called the transverse-traceless (TT) gauge. The components of the metric perturbation in TT-gauge are found by using the projection $\mathcal{P}$:

$$h^{TT}_{jk} = P_{jl}h_{lm}P_{mk} - \frac{1}{2}P_{jk}(P_{ml}h_{ml}), \tag{1.5}$$

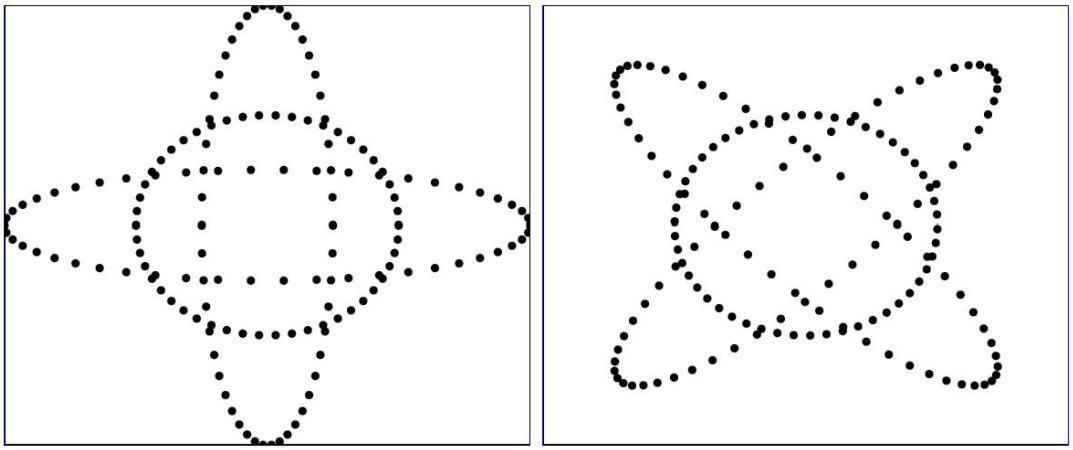

Figure 1.1: Two polarization of GWs. The left picture show the effect on proper separations of particles in a circular ring due to a plus-polarized wave traveling in the perpendicular direction. The right picture show the case in a cross-polarized wave. The ring continuously gets deformed into one of the ellipses and back during the first half of a gravitational wave period and the same effect of deformation in the other orthogonal ellipse and back during the next half period.

or in matrix notation

$$\mathbf{h}^{TT} = \mathbf{PhP} - \frac{1}{2}\mathbf{P}\mathrm{Tr}(\mathbf{Ph}), \tag{1.6}$$

where the projection operator $\mathcal{P}$ is determined by the wave vector $\vec{k}$ of GW according to $P_{ij} = \delta_{ij} - k_i k_j / k^2$, and δ_{ij} are the Kronecker symbol.

Under the TT gauge there are only two independent polarizations out of the original ten. This means that the spatial components of the metric perturbation are determined by two functions $h^{+} = h^{TT}_{xx} = -h^{TT}_{yy}$, $h^{\times} = h^{TT}_{xy} = h^{TT}_{yx}$. These two (physical) degrees of freedom are the two polarizations of gravitational waves.

Let us consider the effect of GW on the test mass in the transverse plane, if a wave increases the proper distance between two free masses that lie along a given direction, it will simultaneously decrease the distance between two free masses lying along the perpendicular direction in the transverse plane. This is illustrated in the Figure 1.1.

1.2 GW detection

Since the prediction of the existence of GW, the detection of GW has been a big challenge. In 1974, the first indirect proof came from a systematic survey of the pulsar PSR 1913+16, (or Hulse-Taylor binary system) at the Arecibo observatory in Puerto Rico Hulse & Taylor [1975]. The shirking of the binary orbit and decreasing of the orbital period imply the loss of the energy. In late 1970s and early 1980s, Damour demonstrated that this should be corresponding to the emission of gravitational radiation. That is to say, the GWs is really exist, they carry away the system's energy and orbital angular momentum. Although, it is not the direct detection of the GW, this brilliant work was awarded the Nobel prize in 1993, "for the discovery of the new type of pulsar, a discovery that has open up new possibilities for the study of gravitation " Hulse [1994]; Taylor [1994].

GWs manifest themselves as time varying tidal force as we discussed at the end of section 1.1. The strength h of the GW is measured by the relative change of the distance between test masses Misner *et al.* [1974].

$$\frac{\delta L}{L} = \frac{h}{2}, \tag{1.7}$$

where L is the distance between the test mass before GW is coming, δL is the change of this distance when the GW arrived. The problem is, nearby the Earth, even for the most intense evolution of the celestial bodies, the strength of GW can be extremely small which is around 10^{-21} or much smaller. Equation (1.7) implies that for measurements need the large L, also need to reduce all the environmental influence. There are ground based that are already in operation and space borne detectors are planed in the near future.

1.2.1 Ground base detectors

There are two kind of ground based detector: resonant-mass detectors and laser interferometers.

The resonance bar detector was firstly designed by Joseph Weber in late 1960s Weber [1967]. These detectors are constituted by a large aluminium bar, which is

put in the vacuum environment to shield from the earthquake and other acoustics noise, and it is cooled to low temperature to reduce the thermal noise. The detectors are equipped with transducers that monitor the complex amplitudes of one or several of the bar's vibrational modes. A passing gravitational wave changes these amplitudes due to its frequency content near the normal mode frequencies. The frequency band of bar detectors are very narrow, its bandwidth is only a few Hz around center-frequencies in the kHz range and the sensitivity is around $h = 10^{-19}$ Sathyaprakash & Schutz [2009].

The laser interferometer detector was suggested in the middle of 1970s G.E. Moss [1971]; M.E. Gertsenshtein [1963]; Pirani [1956]. It is Michelson interferometer with two orthogonal Fabry-Perot laser cavities. When there are no GWs, the laser frequency is exactly the resonance frequency of the cavity. The outputs from the two cavities can cancel each other. When the GW comes through the detector, the changing length of the cavity change the original resonance frequency, this cause a phase difference between two arms. They can not cancel each other, lead to a signal which appears at the output photodiod. The longer the armlength of the cavity, the more sensitive the detector is. Currently used arm length is several kilometers which leads to the best sensitivity around $h = 10^{-22}$. The noise of the ground based interferometer detector compose mainly from the seismic noise, thermal noise of the mirrors, shot noise from the laser. The existing detectors are, two LIGO in USA Abramovici *et al.* [1992], Virgo (joined French and Italian project) Bradaschia *et al.* [1990], GEO600 (joined German and UK project) Danzmann *et al.* [1994] and TAMA in Japan. These detectors are operating within frequencies $30Hz - 2kHz$. Currently, most of these detectors are under upgrade.

1.2.2 Space detector

Laser Interferometer Space Antenna (LISA) was design jointly by NASA and ESA in 1990s as GW detector in space Danzman *et al.* [1998]. Due to financial problem on both side of Atlantic this project was not realized and the partnership was broken in 2011. There is a fundamental difference between LISA and the ground-based interferometers: LISA will search for the distinctively low-frequency

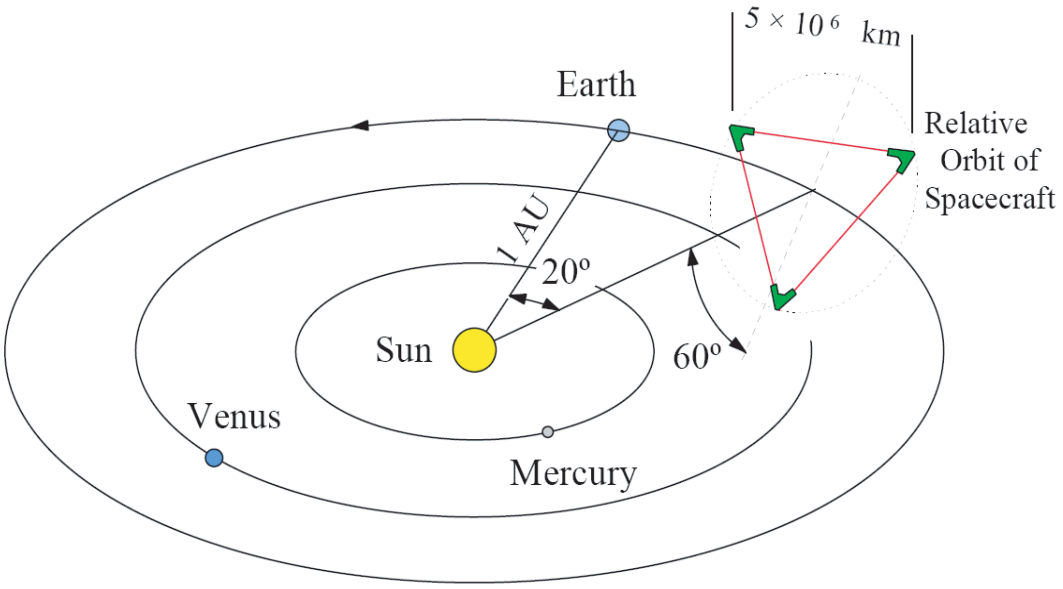

Figure 1.2: The constellation of LISA. The center of this LISA constellation moves around the Sun in an earth-like orbit (R = 1AU), 20° behind the Earth. The constellation plane is inclined by 60° to the ecliptic. The triangular constellation revolves around its center as it completes one orbit in the course of one year.

(milli-hertz) gravitational waves which will never be detectable by any ground based detectors. LISA is a space-borne implementation of a Michelson laser interferometer with the purpose of measuring the fluctuations in the distance between widely separated test masses. They are arranged to be floating almost freely in the center of the spacecraft by using the "Drag-free control" technique. This technique protects the test masses from the external disturbances such as solar wind, only gravitational waves would perturb their relative motion. The Figure 1.2 show the orbit of LISA constellation LISATeam [2000].

The constellation of LISA is comprised of three spacecrafts which are free flying and arranged in an approximately equilateral triangle. Each spacecraft exchange laser signals with two adjacent spacecrafts. The separation between any two spacecrafts is $L = 5 \times 10^9$m. As a single entity, the guiding center of the three spacecrafts lies on the Earth orbit, trailing 20° behind the Earth. The plane of the triangle is inclined to the ecliptic plane. As viewed from the Sun, the

constellation appears to rotate clockwise around the guiding center, in a so-called "reverse cartwheel" motion Danzman *et al.* [1998]; LISATeam [2000].

Each spacecraft moves on elliptical Keplerian orbits with semi-major axis approximately the same as the Earth's orbital radius R, eccentricity of $e = L/(2\sqrt{3}R)$ and inclinations of $\iota = e\sqrt{3}$. The shape of the constellation is not perfectly stable, and the arm lengths will slowly change in time. The maximum change in arm lengths is only a few percent, but this will have important consequences in the performance of the interferometry.

Inside each craft, there are two optical systems responsible for sending and receiving laser signals with adjacent spacecraft. Each optical bench contain a 30 cm Cassegraine telescope which is used for exchanging of laser signal between spacecrafts. The proof mass is cubic in shape, and made of an alloy of gold and platinum, chosen to reduce the magnetic susceptibility. The noise of LISA can be mainly divided into two types: acceleration noise and position noise. The acceleration noise is due to forces (or accelerations) acting on the proof masses, such as thermal deformation, remanent gas impacts on proof mass etc. The position noise contain the shot noise, laser frequency noise, and the instability of the laser beam orientation. The three crafts constitute three separated, but not independent interferometers. Every two interferometer share one arm , so, the noise in them are not independent.

Compared to the ground base detector, the space mission operates in the low frequency band $10^{-4}Hz \sim 10^{-1}Hz$. The Figure 1.3 is the LISA threshold sensitivity curve which shows the typical strength and various GW sources Larson [2005].

eLISA project was recently suggested to replace LISA, as a potential low-cost alternator Amaro-Seoane *et al.* [2012a]. All the methods presented in this thesis are applicable to any space borne detectors. The only difference is on the event rate and on the expected strength of the signals.

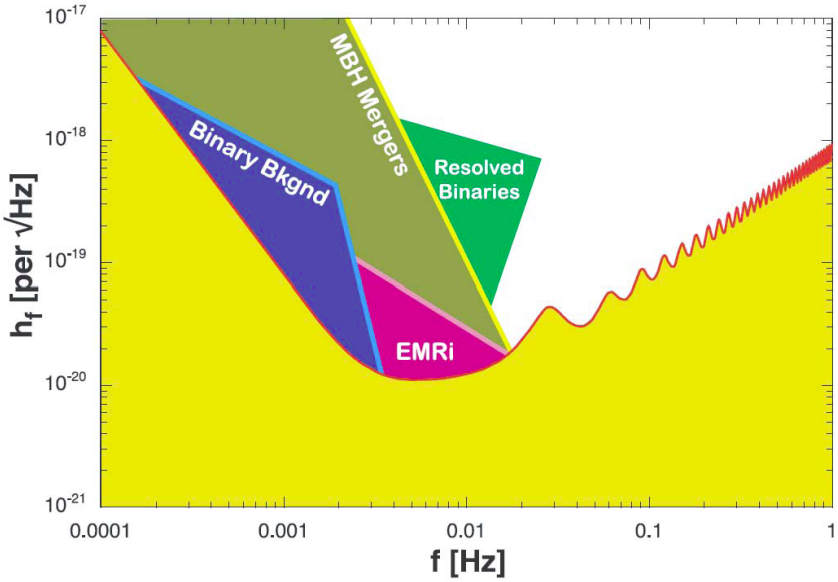

Figure 1.3: LISA sensitivity curve. The red line is the instrumental threshold corresponding to sources with SNR = 1. The yellow region which is below the instrumental threshold embody the strength of the noise. Above the red line, it is the LISA discovery space. The broad areas marked indicate what types of sources are expected to populate each region of the discovery space.

1.3 The GW sources

1.3.1 The Source for The Ground Base Detector

The main sources for the ground base detector are: the burst signal from the supernova collapse; the coalescence of the binary black holes; slightly deformed spinning neutron star; the inspiral of neutron star-neutron star binaries; and neutron star black hole binaries; stochastic background of GWs from the early Universe evolution. Since our work is related to the space detector, we will not discuss further these sources, the details can be found in Cutler & Thorne [2002].

1.3.2 The Source for LISA

The main GW sources for LISA are Galactic white dwarf binary systems, super massive black hole binaries, extreme or intermediate mass ratio inspiraling systems, stochastic GW background LISATeam [2000]. The following subsections will explain them in details respectively.

1.3.2.1 Galactic binaries

Galactic binaries will populate the whole frequency band of the LISA detector Amaro-Seoane *et al.* [2012a]; LISATeam [2009]. The most common sources are white-dwarf / white-dwarf (WD/WD) binaries emitting gravitational wave signals of nearly constant frequency and amplitude. There are detectable known sources from electromagnetic observation, for example the shortest-period interacting binaries HM Cnc (RX J0806.3+1527), V407 Vul, ES Cet and the recently discovered 12 minutes period detached system SDSS J0651+28 Amaro-Seoane *et al.* [2012a]; Brown *et al.* [2011], There are also some interacting systems, such as AM CVn stars, a white dwarf (WD) that accretes from a low-mass helium-star companion; the known detectable example is RXJ1914+245: a $0.07M_{\odot}$ helium star is orbiting a $0.6M_{\odot}$ WD at 100pc from Earth. For interacting systems, the radiation reaction competes with tides and mass transfer, and the orbital period can either increase or decrease. The case of increase in frequency, which is expected for high-frequency sources, is a consequence of the inevitable inspiral of binaries due to loss of the orbital energy and angular momentum Marsh [2011];

Racine *et al.* [2006]; Stroeer & Nelemans [2009]. In some systems the second derivative of the frequency will be measurable, and this extra information can be used to disentangle the various mechanisms that drive the frequency evolution. It has been estimated that interacting systems will contribute 13% of the total galactic binary signal at around 1mHz, with the contribution rising to 26% at $10mHz$ LISATeam [2009]. The total number of detectable binary sources can be around 20000 Blaut *et al.* [2010]; Crowder & Cornish [2007].

From data analysis perspective, galactic binary systems fall into three categories: Resolved Systems, Verification Binaries, and the Confusion-limited Foreground.

The majority of the resolvable systems will be found above 3mHz. This is because LISA's resolving power increases with frequency, and the number density of sources decreases with frequency Bender & Hils [1997]; Hils & Bender [2000]; Hils *et al.* [1990]. It has been estimated that LISA will resolve 10^3 to 10^4 individual with one year of observations Arnaud *et al.* [2007a]; Babak *et al.* [2008a,b, 2010]; Bender & Hils [1997]; Nelemans [2009]; Nelemans *et al.* [2001, 2009].

Verification binaries are already mentioned above, those are the systems which are known from previous astronomical observations. We know sky positions and approximate periods, distances, and masses for many of these sources. Observation of these known binaries will provide a check of the operation of the instrument Stroeer & Nelemans [2009].

The resolvable binaries are just the tip of the iceberg. There are about 10^8-10^9 galactic binaries with GW frequencies $f > 0.1$mHz Ruiter *et al.* [2010]. Majority of them can not be resolved individually, so they constitute a source of "confusion noise" that will dominate over the instrumental noise at $10^{-1}\text{Hz} - 10^{-3}\text{Hz}$. The confusion background is characterized as a source of noise which must be taken into account when we are considering observations of other sources of gravitational waves.

The signals from galactic binary are the simplest and best understood of all the LISA sources. The orbital velocity is much less than the speed of light, and the emitted signals are expected to be very accurately modelled by low-order Post-Newtonian waveforms Arnaud *et al.* [2006b]. The waveform depends on a number of parameters, some of which are intrinsic to the source, like the frequency f at

which the signal is emitted, and rate of change of this frequency $\dot{f}$, the constant amplitude, and the inclination of the orbital plane of the binary to the line of sight.

1.3.2.2 Spinning super massive black hole binaries

The existence of super massive black holes (SMBH) with masses ranging from $10^5 M_\odot$ to $10^9 M_\odot$ in majority of galactic centers is confirmed by several observations (see Ferrarese & Ford [2005] and reference therein) with the SMBH in our Galaxy, SgrA* Schodel *et al.* [2002] being the best example. Mergers of galaxies are common events in the Universe, it is believed that each galaxy has had at least one merger event during its life. During these mergers, the SMBH of each galaxy is driven to the center of the remnant full of stars and gas via dynamical friction Callegari *et al.* [2009]. The pairs of SMBH separated by about kiloparsec are observed in some active galactic nuclei such as NGC6240 Komossa *et al.* [2003], Arp299 Ballo *et al.* [2004], ESO509-IG066 Guainazzi *et al.* [2005], Mrk463 Bianchi *et al.* [2008] and J100043.15+020637.2 Comerford *et al.* [2009]. The interaction with the gas disc can bring the binary on a tighter orbit down to a few parsecs in a reasonable amount of time (few Myr) Dotti *et al.* [2009a]. There are few candidates of SMBH binaries on the sub-parsec scale: the quasars OJ287 Valtonen *et al.* [2008] ($\sim$ 0.05 pc) and SDSSJ092712.65+294344.0 Bogdanovic *et al.* [2007]; Dotti *et al.* [2008]($\sim$ 0.1-0.3 pc). To overcome the last parsec separating the SMBHs and bring them to the efficient gravitational wave (GW) driven inspiral several scenario have been proposed. Here are few possibilities: rotation of the merging galaxies and triaxial potential Berentzen *et al.* [2008], processes involving gas Cuadra *et al.* [2008], resonant relaxation Hopman & Alexander [2006], massive perturber Perets & Alexander [2007], young compact stars cluster McMillan & Portegies Zwart [2003], effect from IMBH Portegies Zwart *et al.* [2005], etc. When the separation is less than 10^{-3} pc, the binary evolution is efficiently driven by the gravitational radiation and can reach the coalescence in less than 10^9 years.

Due to a very high signal-to-noise ratio (SNR) (The definition of SNR will given in the section of data analysis), the GWs emitted by the SMBH binary at the end of the inspiral phase will be the strongest source and detected by the

future space born mission LISA. The merger of two super massive black holes (SMBHs; $10^4 \sim 10^7 M_\odot$) at $z = 1$ could be seen with SNR of thousands. It is at least 10 times greater than we can expect with the ground based detectors.

The rate estimate for LISA's SMBH mergers in the range $10^{-1} - 10^2$ per year. Even if the coalescence rate is only 0.1/yr, then LISA should still see ~ 3 BH/BH binaries with $3000 M_\odot \lesssim M \lesssim 10^5 M_\odot$ that are ~ 30 years away from their final merger. Because of the high SNR of the SMBH-merger, LISA could detect such mergers beyond $z = 20$.

It is believed that almost all the SMBHs are spinning, so spin-spin and spin orbital interaction adds complexity to the signal. However the predictions for the magnitudes and directions of the spins of SMBHs in the binary systems differ largely depending on the models, the environment of the binary and the physical processes involved (coherent accretion, alignment of spins with the gas disc Bogdanovic *et al.* [2007]; Dotti *et al.* [2009b]; Perego *et al.* [2009], sequence of randomly oriented accretion events King *et al.* [2008], etc).

The SMBH could have significant spin as a result of two physical processes. One is the coherent accretion, the gas falls onto black hole from a massive accretion disk, and transfer the angular momentum to black hole causing spin up. This process could last a long time. The other case is a result from the merger of two spinning MBHs with aligned spins. During the evolution, the spins alignment could happen when the binary system have circum-nuclear disk which causes this alignment of spins with angular momentum of a disk.

Several techniques to measure the spins using electromagnetic radiation Brenneman & Reynolds [2006]; Murphy *et al.* [2009] have been proposed, but uncertainties are still very large. The gravitational wave observations of SMBH binaries with LISA should enable us to measure masses and spins of SMBHs in the binary with unprecedented accuracy Lang & Hughes [2006]. The knowledge of spins could give us a lot of information about the kick velocity of remnant SMBH, the engines of active galactic nuclei, the mechanisms involved in galactic centers, etc. Finally, the spin measurements combined with a precise estimation of masses and sky position made with LISA will increase our understanding of the origin of SMBHs, the galactic evolution, the galactic center, cosmology, etc.

The gravitational wave signal from the SMBH binaries can be conventionally

split into three parts: inspiral, merger and the ringdown. The inspiral part of the waveform is also referred to as a chirp. Due to the loss in energy and angular momentum, its amplitude and frequency increase when the two body spiral toward to each other. The signal depends on masses, source's location on the sky and inclination of the orbital plane to the line of sight. The duration of the inspiral in the LISA's band decreases with increasing of the total mass. For a fixed total mass, the duration is longer for unequal mass binaries: $t \sim M^{-5/3}\eta^{-1}$. Here, $M = m_1 + m_2$ is the total mass, and $\eta = \frac{m_1 m_2}{M^2}$ is the symmetric mass ratio. If SMBHs are spinning, then spin-spin and spin-orbital interactions will cause both amplitude and phase modulation of the signal. The merger (i. e., the last few cycles of the orbital evolution) of two compact objects and the associated final burst of gravitational radiation is a broadband signal of a very short duration Kesden *et al.* [2010]; Schnittman [2007]; Schnittman & Buonanno [2007]. This part of the signal is very strong and could be detected alone Babak *et al.* [2010]; Baker *et al.* [2007].

The merger is followed by a "ringdown" of the final black hole, through which it loses its initial distortions to become a stationary Kerr black hole. If the final mass of the black hole is greater than $\sim 10^6 M_\odot$, the SNR from this phase $(10^3 - 10^4)$ will outshine the inspiral and merger of the system.

In the case of low total mass (less than 10^5 solar masses), the ringdown signal falls outside the LISA band or has a low SNR. In the case of large total mass of binary system (more than few million solar masses), the ringdown signal lies in the LISA band, and give significant SNR (comparable or larger than the SNR from the inspiral signal), therefore they could be detectable by LISA (Berti *et al.* [2006]).

The ringdown signal consists of a superposition of quasi-normal modes where each mode has a complex frequency, whose real part is the oscillation frequency and whose imaginary part is the inverse of the damping time. The damping time is uniquely determined by the mass and angular momentum of the newly born black hole. Because of the nature and the strength of the sources, there will be almost no confusion between multiple MBHB signals in the same data stream, making them the cleanest sources that we have to deal with Amaro-Seoane *et al.* [2012a]; Cornish & Porter [2007].

1.3.2.3 Capture of the compact stellar mass by the super massive black hole

SMBHs in the centers of galaxies regularly capture objects from the clouds of stars surrounding them. Many of these capture happened when the compact body have, by chance, been deflected onto orbits close enough to the SMBH, due to the loss of the energy by the gravitational radiation, the compact body is placed into a bound orbit Gurkan & Hopman [2007]; Hopman & Alexander [2006]; Ivanov [2002]. The other mechanics of capture is the tidal effects: compact-object binaries that have come too close to the hole and been disrupted by tidal forces Miller [2004]. The resulting compact objects (neutron stars, black holes, white dwarfs) would then survive as point-like objects orbiting the central hole until their losses to gravitational radiation send them across the horizon. They would radiate waves in the LISA band for a long time, many of them emitting approximately 10^6 cycles or more of radiation before falling into the hole. These objects are interesting because they tell us about the astrophysics of the stellar populations near central galactic black holes and describe the geometry of space time around massive black holes with unprecedented precision. A review of the sources, the astrophysics, and detection of EMRIs using LISA can be found in Amaro-Seoane *et al.* [2007, 2010].

There are two principal kinds of systems: EMRIs and IMRIs. EMRIs are Extreme Mass Ratio Inspiral sources, where the mass ratio is 10^{-4} or smaller, the typical expected system is a stellar black hole with a mass of $10M_\odot$ falling into a $10^6 M_\odot$ supermassive black hole. IMRIs are Intermediate Mass Ratio Inspiral sources, where the ratio is somewhat larger, between 10^{-4} and 10^{-2}. The inspiraling object here is expected to be an intermediate mass black hole, such as remnant of the first generation of star formation (Population III) Miller [2004]. The canonical expected source in the LISA band is an intermediate-mass black hole with a mass of $10^3 M_\odot$ falling into a $10^6 M_\odot$ supermassive black hole.

The event rates for EMRIs are not well known. It is shown (relying in part on simulations of star-cluster evolution around a SMBH by Freitag [2001]) that the detection rate is likely to be dominated by $\sim 10M_\odot$ BH's spiraling into $\sim 10^6 M_\odot$ SMBH's Cutler & Thorne [2002]; Gair *et al.* [2004, 2010]. SMBH's $\sim 10^6 M_\odot$

are likely to have larger relaxed stellar cusps around them than either smaller or bigger BH's; coincidentally $10^6 M_\odot$ being the optimum SMBH mass for LISA. And while inspirals of $\sim 10 M_\odot$ BH's probably have a smaller space density than inspirals of $\sim 1.4 M_\odot$ NS's, because of their larger SNR, they can be detected throughout a volume that is $(\sqrt{10/1.4})^3 \sim 20$ times larger. The estimated inspiral rate for $\sim 10 M_\odot$ BH's is $\sim 1-10/yr$ out to $1 Gpc$ Sigurdsson & Rees [1996], where the S/N's would be ~ 60 to ~ 200 if optimal signal processing is applied

The frequency of EMRI and IMRI signals is determined by the mass of the central black hole, but the amplitude is determined by the mass of the infalling object. This is because the square of the SNR for signals will be proportional to the total energy radiated, which is expected to be around 6% of the mass of the infalling object. The SNR of a "standard" EMRI consisting of a $10 M_\odot$ black hole falling into a $10^6 M_\odot$ black hole, observed during its whole transit through the LISA band, will be of order 30 if the source is at red shift $z = 1$.

An IMRI event generated by an infalling $10^3 M_\odot$ black hole will be correspondingly stronger, visible with similar SNR out to $z = 20$. Rate estimates are very uncertain. LISA has unique opportunity for a first discovery of an IMBH.

The waveforms are quite complex. If the source object falls toward a spinning Kerr black hole from a direction not in the equatorial plane, it will move along a complex three-dimensional orbit Barack [2009]. To lowest order the orbit will be a geodesic, but we must take into account the first-order effects of the object's mass and spin, which cause the orbit to evolve and eventually to force the object inside the horizon Poisson *et al.* [2011]. The phase evolution of the radiation from such an orbit contains detailed information about the trajectory of the radiating object, and hence about the geometry of the space time. The amplitude of the signal coupled with its phase evolution contains enough information to estimate its luminosity distance. In order to understand the expected waveform as well as possible, we need to solve equations of motion of a small object on a Kerr background, with corrections at least to first order in the mass and spin of the object.

The gravitational field of the compact body interacts with the background which results in the so called self force. The orbit of the compact body is not the geodesic motion in the background space time anymore, the compact body spirals

toward the central black hole under the action of the self force. Inspiral proceeds until the compact body plunges into the central black hole. In the chapter 4, there is a detailed description of the self force.

For EMRIs the mass ratio is small enough that the first-order perturbation analysis might be valid over the entire orbit. Even so, this is a challenging theoretical problem in general relativity, it has attracted considerable attention during the last ten years Barack [2000, 2001]; Barack & Burko [2000]; Barack & Ori [2000, 2003]; Barack & Sago [2007, 2010, 2011]; Barack *et al.* [2002, 2007]; Drasco & Hughes [2004]; Glampedakis & Kennefick [2002]; Hughes [2000, 2001]; Sago *et al.* [2008]; Sundararajan *et al.* [2007, 2008]; Warburton *et al.* [2012]. It is now known in principle how to solve the problem, and researchers are currently developing implementations that will allow efficient computations of predicted waveforms Babak *et al.* [2007]; Barack *et al.* [2010]; Barausse & Rezzolla [2008]; Carlos F. Sopuerta [2011, 2012]; Yu & Jeffery [2010]; Yunes *et al.* [2011]. For IMRIs it may be necessary to supplement perturbation theory with numerical integrations or post-Newtonian approximations in order to get a complete predicted orbit. A feature of EMRI and IMRI signals is the very large parameter space (17 dimensions, some of them very densely sampled) that characterizes them. The emitted waveform depends on the masses and spins (three-dimensional) of the central hole and the infalling object and on the orientation and eccentricity of the orbit. The orientation of infalling object's spin and eccentricity of the orbit will change with time. In addition, the received signal is modulated in phase and amplitude by LISAs motion around the Sun.

A remarkable fact about EMRIs is the high number of harmonics contributing to the GW emission. This reflects in detail the near-geodesic orbits they follow around the central black hole; while the orbits in turn map the space-time Ryan [1995]. The phase evolution of a signal from a compact object in a pure Kerr geometry allows to look for deviations from the geometry predicted by general relativity, thereby testing one of the most fundamental theorems of Einstein's theory, that black holes are uniquely determined by their mass and their angular momentum. This means that EMRI provide information about the geometry of the SMBH in strong field region, and allow to test certain aspects of general relativity, astrophysics , cosmology to high precision Amaro-Seoane *et al.* [2007].

1.3.2.4 Stochastic GW background

In addition to radiation from discrete sources, there should be a stochastic GW backgrounds which potentially might be also detected by LISA. Several mechanisms producing stochastic GW in early Universe are proposed, such as, inflation after the Big Bang, first order phase transition of quantum field, excitation of our universe as a membrane in higher dimension, the cosmic strings. The detail of these sources can be find in Amaro-Seoane *et al.* [2012a]; Cutler & Thorne [2002]; LISATeam [2000, 2009].

1.4 Gravitational waves data analysis

The objective of the gravitational waves data analysis is to reconstruct as well as possible the incoming gravitational wave. Due to the present of the noise, the problem of extracting the signal from the data is a statistical one. The basic idea behind signal detection is that the presence of the signal changes the statistical characteristics of the data, in particular its probability distribution. Facing the long time observational data, what we need to do is using the statistic inference method to find whether the signal is present or not, and how does the signal looks like.

1.4.0.5 Hypothesis Test

In this subsection, we following the results from Jaranowski & Krolak [2005]. The problem of detecting the signal in the noise can be posed as a statistical hypothesis testing. A signal can be detected with only a certain probability called the detection probability. In statistic realm, this belong to a statistical hypothesis testing problem.

The null hypothesis H_0 is that the signal is absent from the data and the alternative hypothesis H_1 is that the signal is present. A hypothesis test (or decision rule) δ is a partition of the observation set into two subsets, R and the rest is R'. If data are in R, we accept the null hypothesis, otherwise we reject it. There are two kinds of errors that we can make. A type I error is choosing hypothesis H_1 when H_0 is true and a type II error is choosing H_0 when H_1 is

true. In signal detection theory the probability of a type I error is called the false alarm probability, whereas the probability of a type II error is called the false dismissal probability.

For given data, in order to determine which hypothesis we should choose, we need to assign the costs to our decision. With the introduced positive number $C_{ij}, i, j = 0, 1$ where C_{ij} are the cost incurred by choosing hypothesis H_i when H_j is true, we define the conditional risk R for each hypothesis:

$$R_j(\delta) = C_{0j}P_j(R) + C_{1j}P_j(R'), \quad j = 0, 1, \tag{1.8}$$

where P_j is the probability distribution of the data when H_j is true. Next, with the assign probabilities π_0 and $\pi_1 = 1 - \pi_0$ to the occurrences of H_0 and H_1, the Bayes risk as the overall average cost incurred by the decision rule is:

$$r(\delta) = \pi_0 R_0(\delta) + \pi_1 R_1(\delta). \tag{1.9}$$

Then, we define the Bayes rule as the rule that minimizes the Bayes risk $r(\delta)$.

If we are not able to assign priors to various hypotheses, then we seek a decision rule that minimizes, over all δ, the maximum of the conditional risks, $R_0(\delta)$ and $R_1(\delta)$. A decision rule that fulfills that criterion is called a minimax rule.

The imposition of a specific cost structure on the decisions is also not possible or desirable. The Neyman-Pearson approach involves a trade-off between the two types of errors that one can make in choosing a particular hypothesis. The Neyman-Pearson design criterion is to maximize the power of the test (probability of detection) subject to a chosen significance of the test (false alarm probability).

It is remarkable that all three very different approaches Bayesian, minimax, and Neyman Pearson lead to the same test called the likelihood ratio test. The likelihood ratio Λ is the ratio of p_1 and p_0:

$$\Lambda(x) = \frac{p_1(x)}{p_0(x)}. \tag{1.10}$$

where p_1 is the probability density functions when the signal is present, and p_0 is the probability density functions when the signal is absent. We accept the hypothesis H_1 if $\Lambda > k$, where k is the threshold that is calculated from the costs C_{ij}, priors π_i or the significance of the test depending on which approach is being used.

1.4.1 Matched filter

Matched filtering is a data analysis technique that efficiently searches for a signal of known shape buried in the noisy data Helstrom [1968]. This technique consists of correlating the noisy output of each interferometer data with a set of theoretical waveform templates. Following the description in Maggiore [2000], we use $s(t)$ to denote the detector output, which is assumed to consist of a stationary noise $n(t)$ and a gravitational wave signal $h(t)$.

$$s(t) = h(t) + n(t). \tag{1.11}$$

The Fourier transform of a function $x(t)$ will be denoted $\tilde{x}(f)$ and is defined as

$$\tilde{x}(f) = \int_{-\infty}^{\infty} x(t) e^{2\pi i f t} dt. \tag{1.12}$$

The autocorrelation function of the noise n is defined as

$$C_n(t, t') = E[n(t)n(t')], \tag{1.13}$$

where E denotes the expectation value. Let us further assume that the detector's noise n is a zero-mean and Gaussian random process.

For stationary noise, its autocorrelation function depends on times t and t' only through the difference $t - t'$. It implies that there exists then an even function κ_n of one variable such that

$$E[n(t)n(t')] = \kappa_n(t - t'). \tag{1.14}$$

Spectral properties of stationary noise are described by its one-sided spectral density, defined for non-negative frequencies $f \geq 0$ through relation

$$S_n(f) = 2\int_0^\infty \kappa_n(t)e^{2\pi i f t}dt. \tag{1.15}$$

The spectral density S_n can also be determined by correlations between the Fourier components of the detector's noise

$$E[\tilde{n}(f)\tilde{n}^*(f')] = \frac{1}{2}S_n(|f|)\delta(f-f'), \quad -\infty < f, f' < \infty. \tag{1.16}$$

A natural inner product between any two waveforms x and y is defined as

$$(x|y) = 4\mathfrak{R}\int_0^\infty \frac{\tilde{x}(f)\tilde{y}^*(f)}{S_n(f)}df, \tag{1.17}$$

for the case of stationary noise with one-sided spectral density S_n, where $\mathfrak{R}$ denotes the real part of a complex expression, and the asterisk is complex conjugation.

Let us show that the matched filtering define an optimal detection statistic. Define :

$$\hat{s} = \int_{-\infty}^{\infty} s(t)K(t)dt, \tag{1.18}$$

where $K(t)$ is called the filter function. The SNR is defined as S/N, where S is the expected value of $\hat{s}$ when the signal is present, and N is the rms value of $\hat{s}$ when the signal is absent. Since $< n(t) >= 0$,

$$S = \int_{-\infty}^{\infty} < s(t) > K(t)dt = \int_{-\infty}^{\infty} h(t)K(t)dt = \int_{-\infty}^{\infty} \hat{h}(f)\hat{K}^*(f)df, \tag{1.19}$$

$$N^2 = [< \hat{s}^2(t) > - < \hat{s}(t) >^2]_{h=0} = \int_{-\infty}^{\infty} \frac{1}{2}S_n(f)|\tilde{K}(f)|^2dt, \tag{1.20}$$

and therefore

$$\frac{S}{N} = \frac{\int_{-\infty}^{\infty} \hat{h}(f)\hat{K}^*(f)df}{[\int_{-\infty}^{\infty} \frac{1}{2}S_n(f)|\tilde{K}(f)|^2]^{1/2}}. \tag{1.21}$$

We want to find an "optimal" filter function that would produce, on the average, the largest SNR. The filter that maximizes equation (1.21) is given by:

$$\tilde{K}(f) = const\frac{\tilde{h}(f)}{S_n(f)}. \tag{1.22}$$

The constant is arbitrary, since rescaling $\hat{s}$ by an overall factor does not change its signal-to noise ratios. The filter function $\tilde{K}(f)$ defines the matched filter. So, the optimal signal-to-noise ratio is:

$$\frac{S}{N} = \sqrt{(h|h)}, \tag{1.23}$$

that is

$$\left(\frac{S}{N}\right)^2 = 4\int_0^{\infty} \frac{|\tilde{h}(f)|^2}{S_n(f)}df. \tag{1.24}$$

Generally, the signal model $\tilde{h}(f)$ also depends on multiple parameters characterizing the emitting system. We do not know a prior of these parameters, we need to apply matched filtering multiple times to determine them.

1.4.2 Overlap

In order to determine similarity of two signals in shape, we will define the overlap between them. By using the inner product defined in equation (1.17), for two template h_1 and h_2, we define the overlap as :

$$\mathcal{O} = \frac{(h_1|h_2)}{\sqrt{(h_1|h_1)(h_2|h_2)}}. \tag{1.25}$$

The overlap is equal to 1 if the two waveforms are identical, and it equals zero if the two waveforms are orthogonal. The overlap is widely used in finding out the mismatch between the signal and the model, this will be seen more clearly in Chapter 4.

1.4.3 Parameter estimation

When we introduced the matched filtering technique, we assumed that the form of $h(t)$ is know. In practice, $h(t)$ will necessarily depends on a number of free parameters. Thus, we must consider a family of possible waveforms, or templates, that we denote generically as $h(t, \lambda)$, where $\lambda = \{\lambda_1, \cdots \lambda_N\}$ is the collection of parameters. How do we recover the most probable value characterizing the parameters of the source?

The rest of this subsection is mainly based on Jaranowski & Krolak [2005]; Jaranowski *et al.* [1998]. From the property of the noise, we see that the variance of the Fourier amplitude of the noise with frequency f is proportional to $\frac{1}{2}S_n(f)$, so the corresponding Gaussian probability distribution for the noise is

$$p(n_0) = \mathcal{N} \exp\left\{-\frac{1}{2}\int_{-\infty}^{+\infty} \frac{|\tilde{n}_0(f)|^2}{(1/2)S_n(f)} df\right\}, \tag{1.26}$$

where $\mathcal{N}$ is a normalization constant. This is the probability that the noise $n(t)$, which is a random Gaussian with zero mean, has a given realization $n_0(t)$. In terms of inner product, it is rewritten as:

$$p(n_0) = \mathcal{N} \exp\{-(n_0|n_0)/2\}. \tag{1.27}$$

Assume that, the signal is present in the data, i.e. it is of the form $s(t) = h(t; \lambda_t) + n_0(t)$, where λ_t is the (unknown) true value of the parameters λ. The likelihood function $\Lambda(s|\lambda_t)$ for the observed output $s(t)$, given the hypothesis that there is a GW signal corresponding to the parameters λ_t, is obtained by plugging $n_0 = s - h(\lambda_t)$ into equation (1.27),

$$\begin{aligned}\Lambda(s|\lambda_t) &= \mathcal{N}\exp\{-\frac{1}{2}(s-h(\lambda_t)|s-h(\lambda_t))\} \\ &= \mathcal{N}\exp\{(h(\lambda_t)|s)-\frac{1}{2}(h(\lambda_t)|h(\lambda_t))-\frac{1}{2}(s|s)\},\end{aligned} \tag{1.28}$$

since, we are working at fixed data s, then the term $-\frac{1}{2}(s|s)$ is constant which does not depend on λ, and it can be absorbed in the normalization factor $\mathcal{N}$. The final form of the likelihood function is

$$\Lambda(s|\lambda_t) = \mathcal{N}\exp\{(h(\lambda_t)|s)-\frac{1}{2}(h(\lambda_t)|h(\lambda_t))\}. \tag{1.29}$$

From this expression, we see immediately that the likelihood test is essentially correlating the data s with the signal h that is present in the noise and comparing the correlation to a threshold Jaranowski & Krolak [2005]. The $\lambda-$ space will in general be a multi-dimensional space of large dimension. For example, for a binary coalescence the parameters λ^i that determine the waveform, are the distance, the source's location (two angle), the orientation of the normal to the orbit (two more angle), the initial orbital phase, the two masses of the stars, their spin, and the spin orientation. Parameters of the signal $\tilde{\lambda}$ which maximize the likelihood are called maximum likelihood estimation or estimators.

A rule for assigning the most probable value is called an estimator. It is a function that assigns to each data set the "best" guess of the true value of this parameter. Ideally we would like our estimator to be (i) unbiased, i.e., its expectation value to be equal to the true value of the parameter, and (ii) of minimum variance. Such estimators are rare and in general difficult to find. But, usually, people are more concerned about the estimator being unbiased, that is why the maximum likelihood estimator is widely used Jaranowski & Krolak [2005].

The value of λ that maximize $\Lambda(s|\lambda_)$ defines the maximum likelihood estimator, and we denote it by $\hat{\lambda}(s)$. It is usually simpler to maximize the log likelihood:

$$\log\Lambda(s|\lambda) = (h(\lambda)|s)-\frac{1}{2}(h(\lambda)|h(\lambda)), \tag{1.30}$$

the value of $\hat{\lambda}$ is found by solving the equation

$$\frac{\partial(\log \Lambda(s|\lambda))}{\partial \lambda_i} = 0, \tag{1.31}$$

that is

$$\left(\frac{\partial h(\lambda)}{\partial \lambda_i}|s\right) - \left(\frac{\partial h(\lambda)}{\partial \lambda_i}|h(\lambda)\right) = 0. \tag{1.32}$$

A natural question is what is the relation between $\hat{\lambda}$ and the value of λ that provide the highest SNR in the matched filtering Maggiore [2007]. In fact, they are the same. To show it, write the template as $h(t, \lambda) = Ah_A(t, \zeta)$, where A is a constant amplitude, and is a free parameter, while the normalization of h_A has been fixed. Separate parameters λ into A and the remaining parameters ζ. The maximization with respect to A of $\log \Lambda$ can be performed analytically. The maximum likelihood estimate for A

$$\hat{A}(s) = \frac{(h_A|s)}{(h_A|h_A)}. \tag{1.33}$$

Using $\hat{A}$ in $\log \lambda$, we obtain

$$\log \Lambda(s|\eta) = \frac{1}{2}\frac{(h_A|s)^2}{(h_A|h_A)}. \tag{1.34}$$

The maximization of this quantity with respect to ζ, amounts to maximizing the inner product of the s with the normalized template $h_A/(h_A|h_A)^{1/2}$.

For the strong signals, $SNR \gg 1$, the expected errors on parameters could be estimated by using the Fisher information matrix Jaranowski & Krolak [2005]; Jaranowski *et al.* [1998]; Vallisneri [2008]. For large SNR, the parameter estimation errors $\Delta\lambda^i$ have the Gaussian probability distribution

$$p(\Delta\lambda^i) = \mathcal{N} e^{-\frac{1}{2}\Gamma_{ij}\Delta\lambda^i\Delta\lambda^j}. \tag{1.35}$$

here Γ_{ij} is the so called Fisher information matrix (FIM) defined by:

$$\Gamma_{ij} = \left(\frac{\partial h}{\partial \lambda^i} | \frac{\partial h}{\partial \lambda^j} \right), \tag{1.36}$$

and $\mathcal{N} = \sqrt{\det(\Gamma/2\pi)}$ is the appropriate normalization factor, and Γ is the determination of Γ_{ij}. The variance covariance matrix is given by:

$$< \Delta\lambda^i \Delta\lambda^j >= (\Gamma^{-1})^{ij}. \tag{1.37}$$

The diagonal elements give the lower bounds on the variances of the estimators of the parameters of the signal and can be used to assess the quality of astrophysical information that can be obtained from detections of gravitational-wave signals. The off-diagonal elements tell us the correlation between these parameters.

1.4.3.1 Algorithm for searching on the parameter space

We have introduced the matched filtering as the optimal technique that maximizes the likelihood function. But, the complex structure of the signals leads to a large number of local maxima in the likelihood function which are widely distributed in the parameter space. Searching for the global maximum in this forest of local maxima becomes a computationally expensive task. The computational efficiency of the search for the global maximum is, thus, an important issue in LISA data analysis. The various search strategies proposed in the GWs literature so far can be broadly divided into those based on sampling the function on predetermined grids of points in the parameter space, and those that use stochastic optimization methods Porter [2009].

In the class of grid-based methods, significant savings in computational costs have been demonstrated with a hierarchy of grids. A nice feature of grid-based methods is that they are easy to characterize statistically. Stochastic methods do not use pre-determined grids but employ some form of pseudo-random walk through the parameter space. The probabilistic rules of the random walk are tuned to maximize the chances of its terminating close to the global maximum.

In general, no matter what kind of stochastic method is used, they should have a common character: in the early stage, the area of exploration should be as large as possible, this is to avoid missing the global maximum. With the progress of the searching, the pseudo random work could have got some information about distribution of maxima, and then the hunting strategy should be changed to the local exploration around the best found maxima. Obviously, any kind of local searching have the same problem: getting trapped into the secondary maxima. This is especially often happens in the case of large amount of comparable local maxima which are spread widely in the parameter space.

At present, there are already many algorithms that fall under the class of stochastic methods: a hybrid of simulated annealing and Metropolis-Hastings Markov Chain Monte Carlo (MCMC) Cornish & Crowder [2005]; Cornish & Porter [2006, 2007], Genetic Algorithm Crowder *et al.* [2006] Petiteau *et al.* [2010], Multi-nest Feroz *et al.* [2009], Particle swarm Optimization (PSO) Wang & Mohanty [2010]. The simulated annealing is widely used in smoothing the likelihood function surface to help the pseudo random walk to explore larger area in the likelihood surface. MCMC is a very efficient stochastic process in finding the local maxima and it also serves for the parameter estimation. Genetic algorithm and PSO method have similar feature, they are all belong to the population optimization. Summary of efficiency of various algorithm for LISA data analysis is given in Arnaud *et al.* [2007a]; Babak *et al.* [2008a,b, 2010], results show that all of them can have the comparable efficacy which partially depends on the GW source type.

1.4.4 LISA data analysis and MLDC

In the ground based detectors, the data stream is acquired at rate of several megabytes per second, and the GW signals which are mainly the transient signals are buried deeply in the noises. Due to the low SNR, the detection of the signal is the most important problem Abramovici *et al.* [1992].

In the case of LISA, the data rate will be 10^3 times less than from a ground-based detector, because of the lower frequencies band. The massive data-handling problems faced by ground-based interferometers will not exist. LISA observes pri-

marily long-lived sources, and the high SNR imply that making decision between H_1 and H_0 is not a problem LISATeam [2000]. The main difficulty is that the GW signals overlap in time and frequency in the LISA data, so we need to disentangle them. For each individual signal, we want to recover its physical parameters like sky location, mass, polarization angle, frequency evolution, etc. as precisely as possible. All of these make the LISA data analysis a big challenge.

Because the space mission has not been launched, the data is not available now. In order to develop the data analysis techniques to support the future mission LISA (or similar), a project named Mock LISA Data Challenges (MLDC) have been carried out for several years. Since 2006, there have been already 4 rounds of challenge with the increasing complexity which are described in Arnaud *et al.* [2007a,b, 2006a,b].

In the first three rounds, each challenge consists of several kinds of 2 year long data sets Arnaud *et al.* [2007a, 2006a]. Each simulated data contain one or several signals from different kind of sources which are constructed by using various simulators. The noise in the data is simulated as Gaussian colored noise. Each round contains two separate data sets: the blind challenge data set where the source parameters are unknown and the training data set with the publicly available parameters. In the latest round, the data set consists of a single challenge that includes all the kinds of sources in the same data. The data contain more than 60 million chirping Galactic binaries, few MBH binaries, several EMRIs, number of cosmic-string bursts, an isotropic stochastic background, and of course instrument noise Babak *et al.* [2010].

The data was analysed by several groups around the world, the detail of their work can be found in the series of publications Arnaud *et al.* [2007a]; Babak *et al.* [2008a,b, 2010].

1.5 Structure of the thesis

In chapter 2, we review the basic techniques of GW dada analysis which are used in LISA project, including: the LISA measurement, TDI technique, LISA response function, $\mathcal{F}$-statistic.

In chapter 3, I present my published work on searching for spinning black hole

binaries in mock LISA data using a Genetic Algorithm Petiteau *et al.* [2009b, 2010]. Several algorithms for detecting non-spinning MBH binaries in simulated LISA data have already been demonstrated Brown *et al.* [2007]; Cornish & Porter [2007]; Feroz *et al.* [2009]; Gair & Porter [2009]; Petiteau *et al.* [2009b]. In our work, we considered inspiralling spinning MBH binaries and we presented a particular adaptation of the Genetic Algorithm (GA) to search for GW signals from those systems. Genetic Algorithms belong to the family of optimization methods, i.e. they look for extremum. The first application of GA in LISA data analysis was proposed in Crowder *et al.* [2006] for Galactic binaries. In this method the waveform template is associated with an organism, and parameters play the role of the set of genes defining this organism. The logarithm of likelihood obtained with a given template defines the quality of the organism. A set (colony) of organisms is then evolved through breeding, mutation and custom designed accelerators with the aim of finding the genotype with the highest quality. This corresponds to the standard Darwin's principle: weak perishes, strong survives, or, translated into the conventional data analysis language: by evolving a set of templates, we are searching for the parameter set that maximizes the likelihood. We have applied the GA to the analysis of the third round of mock LISA data challenge. The mock data set 3.2 consisted of the Gaussian instrumental noise, galactic background and between four to six signals from the inspiralling spinning SMBH binaries in a quasi-circular orbit Babak *et al.* [2008a]. We have found several maxima which are almost equal in value of the likelihood and widely separated in the parameter space. We search for each such strong maximum, which we call *mode*, and then explore it by a designated set of organisms. We refer to this extension of the standard GA as a multimodal GA. The mutlimodal GA applied to the blind search has shown an excellent performance: we have detected all present signals with a very accurate estimation of the parameters (were possible).

In chapter 4, I present my published work Wang *et al.* [2012] on modeling of the phenomenological waveform (PW) of EMRI and its application in data analysis. The signals emitted from EMRIs can be very complex due to the fact that the inspiraling object explores the metric close to the central black hole and experiences the full wealth of relativistic effects, which makes it difficult

and numerically costly to create precise waveforms. In the waveform modeling for EMRI system, the standard post-Newtonian approximation has the problem of poor convergence, the numerical relativity has the problem in dealing with the two extremely different timescale. So, it is more appropriate to employ the perturbation theory to calculate at least approximate waveform. In the perturbation theory, the inspiralling object can be regarded as a small perturbation to the known background space-time of the central black hole. While the mathematical theory of the self-force has been developed in the last decade, the computation of a full generic inspiral trajectory and a waveform from a particle on that trajectory is computationally expensive Barack & Sago [2010]; Poisson *et al.* [2011]; Pound [2010a,b, 2012]; Pound & Poisson [2008a,b]; Pound *et al.* [2005].

In our work, we choose the numerical kludge approach as our physical model of EMRI Babak *et al.* [2007]. In the adiabatic approximation, at each instant, the orbit of the compact body can be treated as a geodesic in Kerr space time Drasco & Hughes [2004]. It can be also labeled by the compact body's energy, projection of orbital angular momentum on the spin of SMBH, and Carter constant. During the inspiral, these three constants of motion change slowly due to the radiation reaction, and the orbit evolves from one geodesic to another Cutler *et al.* [1994]; Drasco & Hughes [2006]; Gair & Glampedakis [2006]; Gair *et al.* [2011]. Based on these, the orbital motion can be described by three fundamental orbital frequencies f_r, f_θ, f_ϕ which correspond to the radial, polar, azimuthal motion independently Schmidt [2002]. Then the phenomenological waveform (PW) can be constructed by harmonics of these three evolving frequencies with slowly changing amplitudes. We use the Taylor expansion in time to express the orbital phase of each harmonics in polynomial form, with the expansion coefficients: initial phase, frequency, first and higher order derivatives of frequency at some fiducial moment. The amplitudes and the expansion coefficients of each harmonics are treated as our phenomenological parameters.

In order to test our waveform model, we first generated the three months of noiseless data which is represented by the numerical kludge waveform, and used the standard matched filtering approach to analyse these data. We used the Markov Chain Monte Carlo (MCMC) method to search the maximum of the log likelihood surface. Due to the complexity of the waveform, MCMC found a

large number of local maxima. From these found maxima, we can extract the information of dominant harmonics by calculating the accumulative log likelihood. With these found harmonics (in the noiseless case we can find about 13 strongest harmonics), we can plot the frequency evolution for each of them, and we need to determine the underlying physical parameters $\{e, p, \iota, a, \mu\}$. For each set of these physical parameters, we calculated the frequency evolution, then used the standard χ^2 test to assess the match with the found frequency tracks of the dominant harmonics. In the searching on the physical parameters space, we used two different optimization algorithms for searching, one is the particle swarm optimization (PSO), the other is the genetic algorithm.

In the PSO method, waveform template is treated as a freely moving particle, with the physical parameters denote its position in the parameter space Wang & Mohanty [2010]. The essential idea behind PSO is to compute the fitness function simultaneously at several locations and use these samples to influence the position of the set of samples in the next iteration. Every particle has a velocity which is related to the global optimal position and its own historic best position to determine its direction of moving and the future position. Every particle can influence other particles by sharing its search information. This process continues iteratively until some stopping rule is satisfied. The process can be visualized by treating the sample locations as a swarm of particles that moves in the parameter space, hence the name of the algorithm.

In the genetic algorithm, as before the waveform template is associated with an organism, and the physical parameters play the role of the set of genes defining this organism. The χ^2 statistic defines the quality of the organism. A set (colony) of organisms is then evolved through breeding, mutation in the iterative manner.

Our results showed that, both of these two methods can recover the physical parameters with errors about $2\% - 3\%$. For testing our waveform further, we also did the same analysis in the noisy data with the Gaussian white noise. The total SNR of the signal in the data was 50. Due to the influence of the noise, we clearly found 5 dominant harmonics, and we could still recover the physical parameters through these five harmonics with the same precision as in the noiseless case.

Finally, in the last chapter, a summary of my work is presented. This part not only gives a conclusion of my present work, but also highlights the future

projects .

Chapter 2

Basic Techniques of GW data analysis

2.1 LISA Measurement

In LISA constellation, the six independent signals comprise the primary laser signals in the three arms that will be used to create the main interferometric data streams. The basic idea is that you measure the frequency of incoming signal and compare to the local oscillator LISATeam [2000]. The treatment of their response to an incident gravitational wave has been shown to be a Doppler shift in the frequency of the received signal relative to the outgoing signal $\frac{\Delta\nu}{\nu_0}$ by Estabrook & Wahlquist [1975]. Here, ν_o is the fundamental frequency of the electromagnetic tracking signal.

It is shown in Cornish & Rubbo [2003]; Sathyaprakash & Schutz [2009], the detector response to a gravitational wave source located in the $\vec{n}$ direction can be found using Barycentric coordinates (t, x, y, z) and the transverse-traceless gauge to describe a plane gravitational wave propagating in the $\vec{k} = -\vec{n}$ direction . The surfaces of constant phase are given by $\xi = t + \vec{n} \cdot \vec{r}$, where r is the radial vector. A general gravitational wave can be decomposed into two polarization states:

$$h(\xi, \vec{n}) = h_+(\xi)\mathbf{e}_+ + h_\times(\xi)\mathbf{e}_\times, \tag{2.1}$$

where $\mathbf{e}_+$ and $\mathbf{e}_\times$ are the polarization tensors associated with these basis vectors

$$\begin{aligned} \mathbf{e}_+ &= \vec{u}\otimes\vec{u} - \vec{v}\otimes\vec{v}, \\ \mathbf{e}_\times &= \vec{u}\otimes\vec{v} + \vec{v}\otimes\vec{u}. \end{aligned} \tag{2.2}$$

In the SSB coordinate system, the basis vector $\vec{u}$, $\vec{v}$ and the source location $\vec{n}$ can be expressed in term of the sky location of the source (θ, ϕ) according to

$$\begin{aligned} \vec{n} &= \sin\theta\cos\phi\vec{x} + \sin\theta\sin\phi\vec{y} + \cos\theta\vec{z}, \\ \vec{u} &= \frac{\partial\vec{n}}{\partial\theta} = \cos\theta\cos\phi\vec{x} + \cos\theta\sin\phi\vec{y} - \sin\theta\vec{z}, \\ \vec{v} &= \frac{\partial\vec{n}}{\partial\phi} = \sin\phi\vec{x} - \cos\phi\vec{y}. \end{aligned} \tag{2.3}$$

The path length variation for a photon propagating from spacecraft i to spacecraft j is given by

$$\delta l_{ij}(t) = \frac{1}{2}\frac{\vec{r}_{ij}(t)\otimes\vec{r}_{ij}(t)}{1-\vec{k}\cdot\vec{r}_{ij}(t)} : \int_{\xi_i}^{\xi_j} h(\xi)d\xi, \tag{2.4}$$

where $\xi_{i,j}$ are the surfaces of constant phase for the spacecraft i and j, and the vectors $\vec{r}_{ij}(t)$ are:

$$\vec{r}_{ij} = \frac{x_j(t_j) - x_i(t_i)}{l_{ij}(t_i)}, \tag{2.5}$$

point from spacecraft i at the time of emission t_i to spacecraft j at the time of reception t_j. $l_{ij}(t)$ is the distance between spacecrafts i and j. The colon here denote a double contraction $\vec{a} : \vec{b} = a^{ij}b_{ij}$.

We can use GW signal in the frequency domain,

$$h(\xi) = \int_{-\infty}^{+\infty} \tilde{h}(f)e^{2\pi i f\xi}df, \tag{2.6}$$

such a decomposition leads equation (2.4) to be

$$\delta_{ij}(t) = l_{ij}(t) \int_{-\infty}^{\infty} \mathbf{D}(f,t,\vec{k}) : \tilde{h}(f) e^{2\pi i f \xi} df, \tag{2.7}$$

where the detector tensor is given by

$$\mathbf{D}(f,t,\vec{k}) = \frac{1}{2}[\vec{r}_{ij}(t) \otimes \vec{r}_{ij}(t)] \mathcal{T}(f,t,\vec{k}), \tag{2.8}$$

and the transfer function is given by

$$\mathcal{T}(f,t,\vec{k}) = sinc\left(\frac{f}{2f^*_{ij}}[1 - \vec{k}\vec{r}_{ij}(t)]\right) \cdot \exp\left(i\frac{f}{2f^*_{ij}}[1 - \vec{k}\vec{r}_{ij}(t)]\right), \tag{2.9}$$

here $f^*_{ij} = 1/(2\pi l_{ij})$ is the transfer frequency for the arm l_{ij}. The transfer functions arise from the interaction of the gravitational wave with the detector.

2.2 TDI technique

LISA will not be a static instrument as compared to the duration of GWs signals. Each individual spacecraft is on its own independent Keplerian orbit, and is constantly in motion relative to the other spacecrafts in the constellation. In Rubbo *et al.* [2004], the Cartesian coordinates of the spacecraft are given to first order in the eccentricity by the following expressions

$$\begin{aligned} x(t) &= R\cos(\alpha) + \frac{1}{2}eR\left[\cos(2\alpha - \beta) - 3\cos\beta\right], \\ y(t) &= R\sin(\alpha) + \frac{1}{2}eR\left[\sin(2\alpha - \beta) - 3\sin\beta\right], \\ z(t) &= -\sqrt{3}eR\cos(\alpha - \beta), \end{aligned} \tag{2.10}$$

where $R = 1$ AU is the radial distance to the guiding center, $\beta = 2n\pi/3 + \lambda$ $(n = 0, 1, 2)$ is the relative orbital phase of each spacecraft, and $\alpha = 2\pi f_m t + \kappa$ is the orbital phase of the guiding center. The parameters κ and λ give the initial ecliptic

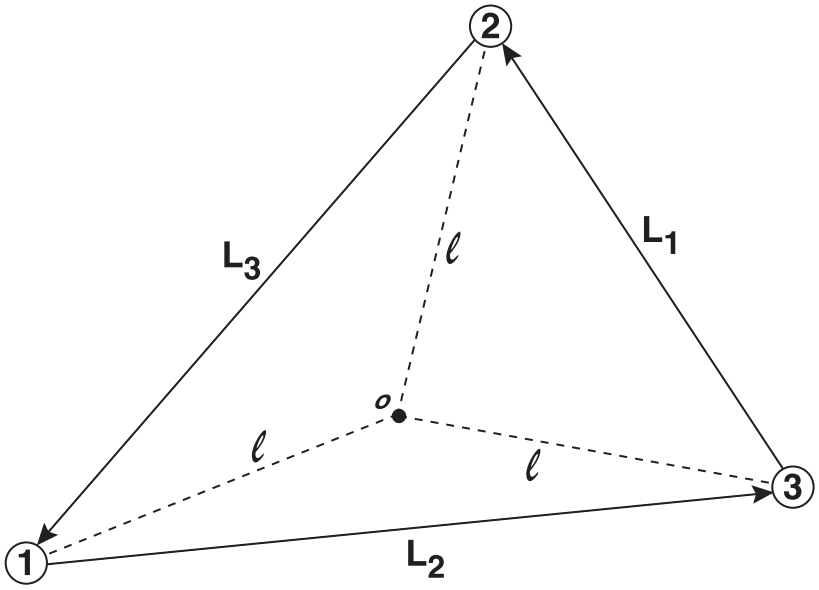

Figure 2.1: Each spacecraft is identified by number, 1, 2, 3, and the opposing arm length is identified by the same index, $\{L_1, L_2, L_3\}$, it is often useful to identify the geometric point o as guiding center, which is equidistant from each of the three spacecraft. The distance l is usually called the guiding center distance.

longitude and orientation of the constellation. $f_m = \frac{1}{\text{yr}}$ is the orbital (modulation) frequency. As a result, the arm lengths are continuously changing in time; the arms "breathe" in an oscillatory way as a function of time. The constellation schematic is shown in Figure 2.1, where each of the three spacecrafts is identified by a number $\{1, 2, 3\}$, and the arm opposite a given spacecraft bears the same numeric index $\{L_1, L_2, L_3\}$.

It is described in Tinto & Dhurandhar [2005] that, in a classic Michelson interferometer, the two arms have the equivalent armlength. The data is constructed as the phase difference between the laser signals in two arms, each measured at the same time t. If the phase of signal arriving along arm L_i is denoted by $s_i(t)$, the traditional Michelson signal $\Delta_3(t)$, which is measured at spacecraft 3, can be written as: $\Delta_3(t) = s_1(t) - s_2(t)$. Because the laser propagates along the two arms and return back to the beam splitter at the same time, the interference of

the two signals cancels out noise associated with fluctuation in the laser frequency .

If the armslength are unequal and changing in time, the fluctuations will arrive back at different time and do not cancel out. Comparing to the magnitude of the other instrumental noise , the expected noise associated with fluctuation is several order larger than them. Fortunately, there is way to solve this serious problem, called time delay interferometry (TDI) J.W. Armstrong & Tinto [1999].

Consider the unequal-arm Michelson TDI combination, which is conventionally called X(t). The one way light travel time along the arm i is denoted as $\tau_i = L_i/c$, then the unequal arm Michelson TDI variable is written as

$$X(t) = s_1(t) - s_1(t - 2\tau_2) - s_2(t) + s_2(t - 2\tau_1), \tag{2.11}$$

which looks like the standard equal-arm Michelson signal with additional time-delayed bits subtracted off. TDI takes the measured phase information from each arm, and delays the linear combinations in such a way to make the optical path lengths in the two phase signals equal. By using this linear combination of the delayed signals, the particular types of fluctuation noise in the laser frequency cancel out.

TDI data streams come in related triples for any given possible combination, there are two related variables which are obtained by exploiting a "permutation symmetry". It is realized by permuting the indices in the manner of $1 \to 2 \to 3 \to 1$. These two companion data streams are called $Y(t)$ and $Z(t)$, they can be written as

$$\begin{aligned} Y(t) &= s_2(t) - s_2(t - 2\tau_3) - s_3(t) + s_3(t - 2\tau_2), \\ Z(t) &= s_3(t) - s_3(t - 2\tau_1) - s_1(t) + s_1(t - 2\tau_3). \end{aligned} \tag{2.12}$$

It is shown in Cornish & Rubbo [2003]; Rubbo *et al.* [2004], in the LISA constellation, the signal along the arm L_i have the contribution from both of spacecrafts at the two ends of this arm, the signal transmitted from spacecraft i that is received at spacecraft j at time t_j has its phase difference to give the output $\Phi_{ij}(t)$. The phase difference has contributions from the laser phase noise,

$C(t)$, optical path length variations, shot noise $n^s(t)$ and acceleration noise $n^a(t)$

$$\begin{aligned}\Phi_{ij}(t_j) &= C_i(t_i) - C_j(t_j) + 2\pi\nu_0[\delta l_{ij}(t_i) + \Delta l_{ij}(t_i)] + n^s_{ij}(t_j) \\ &\quad -\vec{r_{ij}}(t_i)\cdot(\vec{n}^a_{ij}(t_j) - \vec{n}^a_{ji}(t_i)).\end{aligned} \tag{2.13}$$

Here t_i is given implicitly by $t_i = t_j - l_{ij}(t_i)$ and ν_0 is the laser frequency. We have included the variations in the optical path length caused by gravitational waves, $\delta l_{ij}(t_i)$, and those caused by orbital effects, $\Delta l_{ij}(t_i)$. The subscripts on the noise sources identify the particular component that is responsible: C_i is the phase noise introduced by the laser on spacecraft i, n^s_{ij} denotes the shot noise in the photo detector on spacecraft j used to measure the phase of the signal from spacecraft i, and $\vec{n}^a_{ij}$ denotes the acceleration noise on spacecraft j that are mounted on the optical assembly that points toward spacecraft i.

The traditional Michelson signal extracted from vertex 1 has the form

$$\Delta_1(t) = \Phi_{12}(t_2) + \Phi_{21}(t) - \Phi_{13}(t_3) - \Phi_{31}(t), \tag{2.14}$$

where t_2, t_3 are given by

$$\begin{aligned}t_2 &= t - l_{21}(t_2), \\ t_3 &= t - l_{31}(t_3).\end{aligned} \tag{2.15}$$

The unequal arm Michelson TDI combination signal $X_1(t)$ which cancels laser frequency noise in the case of LISA is given as

$$\begin{aligned}X_1(t) &= \Phi_{12}(t_2) - \Phi_{12}(t_1 + t_2 - t) + \Phi_{21}(t) - \Phi_{21}(t_1) \\ &\quad -\Phi_{13}(t_3) + \Phi_{13}(t_1 + t_3 - t) - \Phi_{31}(t) + \Phi_{31}(t_1).\end{aligned} \tag{2.16}$$

The time t_2, t_3 are defined above implicitly and t_1 is

$$t_1 = t - l_{12}(t_1) - l_{21}(t_2)$$

$$= \quad t - l_{13}(t_1) - l_{31}(t_3).$$

The cancelation of frequency noise can be deduced algebraically by direct substitute equation 2.13 into equation 2.16.

Other important TDI triplets which appear in the literature Tinto & Dhurandhar [2005] are $\{A, E, T\}$ (the so-called optimal combinations), $\{\alpha, \beta, \gamma, \}$ (six pulse Combinations) and ζ(the so-called Fully Symmetric Sagnac combinations). The combinations $\{A, E, T\}$ are constructed with uncorrelated noise Tinto & Dhurandhar [2005], they are related to Michelson TDIs as

$$\begin{aligned} A &= \frac{Z - X}{\sqrt{2}}, \\ E &= \frac{X - 2Y + Z}{\sqrt{6}}, \\ T &= \frac{X + Y + Z}{\sqrt{3}}. \end{aligned}$$

The Sagnac combinations ζ and the six pulse combination $\{\alpha, \beta, \gamma, \zeta\}$ have the relationship with Michelson TDIs as follows J.W. Armstrong & Tinto [1999]:

$$\begin{aligned} X(t - L_1) &= \alpha(t - L_2 - L_3) - \beta(t - L_2) - \gamma(t - L_3) + \zeta, \\ Y(t - L_2) &= \beta(t - L_1 - L_3) - \gamma(t - L_3) - \alpha(t - L_1) + \zeta, \\ Z(t - L_3) &= \gamma(t - L_2 - L_1) - \alpha(t - L_1) - \beta(t - L_2) + \zeta. \end{aligned}$$

What we should notice is that the sensitivity curves for different TDI variables have different shapes across the LISA frequency band. Overall, the sensitivities are all very similar, but it raises the intriguing possibility that if a compelling astrophysical argument could be made to reshape the sensitivity curve to provide observational access to a particularly interesting source, the choice of appropriate TDI variables can provide some flexibility.

2.3 LISA response in low frequency limit

The full response of LISA is complicated due to the arm-length fluctuations, pointing ahead, and the transfer functions. Following the approximation which is originally given by Cutler [1998], we work to linear order in the eccentricity for the motion of the constellation, evaluate all spacecraft locations at common time. The transfer function $\mathcal{T}(f,t,\vec{k})$ can be set equal to unity when $f \ll f*$. A LISA mission with $L = 5 \times 10^9$ m arms has a transfer frequency of $f_* = 0.00954 \approx 10^{-2} Hz$. The gravitational wave is taken to be propagating in the $\vec{k}$ direction. From the point of view of observer, we need to express the signal in the detector frame or in the the SSB frame. The theoretical waveform of the GW signal is given in the source or in the radiative frame. The radiative frame is constructed by taking $\vec{n} = -\vec{k}$ as one axis, the other two axes which are called the principle polarization axes $\vec{p}$ and $\vec{q}$, can be chosen as in the paper Finn & Chernoff [1993]. There, $\vec{p}$ and $\vec{q}$ are defined by using the orbital angular momentum $\vec{L}$ as:

$$\vec{p} = \frac{\vec{n} \times \vec{L}}{|\vec{n} \times \vec{L}|}, \vec{q} = \vec{p} \times \vec{n}. \tag{2.17}$$

Instead of $\vec{L}$, we could use some other physical vectors, such as: total angular momentum $\vec{J}$ of the system, or the spin $\vec{S}$ of the MBH . Actually, $\vec{p}$ and $\vec{q}$ are rotated by an angle ψ with respect to the basis $\vec{u}$ and $\vec{v}$. The polarization tensors $\epsilon_+, \epsilon_\times$ in radiation frame which are given in the following expressions

$$\begin{aligned} \epsilon_+ &= \vec{p} \otimes \vec{p} - \vec{q} \otimes \vec{q}, \\ \epsilon_\times &= \vec{p} \otimes \vec{q} + \vec{q} \otimes \vec{p}. \end{aligned} \tag{2.18}$$

have the following relationship with the polarization tensor in SSB frame

$$\begin{aligned} \epsilon_+ &= \mathbf{e}_+ \cos 2\psi + \mathbf{e}_\times \sin 2\psi, \\ \epsilon_\times &= -\mathbf{e}_+ \sin 2\psi + \mathbf{e}_\times \cos 2\psi, \end{aligned} \tag{2.19}$$

and the GW signal

$$h = h_+(t)\epsilon_+ + h_\times\epsilon_\times. \tag{2.20}$$

If the radiative frame is defined as in equation 2.17, then the polarization angle ψ is a function of time if there is an orbital precession (due to spin-orbital coupling). In the case of EMRI, the principle polarization axes can be defined with respect to the spin vector $\vec{S}$ of the central black hole (Babak *et al.* [2007])

$$\vec{p} = \frac{\vec{n}\times\vec{S}}{|\vec{n}\times\vec{S}|}, \vec{q} = \vec{p}\times\vec{n}. \tag{2.21}$$

The advantage of using $\vec{S}$ instead of $\vec{L}$ is that the polarization angle ψ is a constant which could be convenient in the data analysis. In the section 4.3, we use the definition given in equation 2.21.

The Michelson signal considered by Cutler [1998] takes the form

$$h_1(t) = \frac{\delta l_{12}(t-2L) + \delta l_{21}(t-L)}{2L} - \frac{\delta l_{13}(t-2L) + \delta l_{31}(t-L)}{2L}. \tag{2.22}$$

This expression ignores the time variation of the arm lengths due to higher order terms in the orbital eccentricity or perturbations from other Solar System bodies. Using the formula of variation of optical path in Cornish & Rubbo [2003], h_1 can be written as:

$$h_1 = F_+(\theta,\phi,\psi)h_+(t) + F_\times(\theta,\phi,\psi)h_\times(t), \tag{2.23}$$

where the antenna pattern functions are :

$$\begin{aligned} F_+(\theta,\phi,\psi) &= \frac{1}{2}(\vec{r}_{12}\otimes\vec{r}_{12} - \vec{r}_{13}\otimes\vec{r}_{13}) : \epsilon_+ \\ F_\times(\theta,\phi,\psi) &= \frac{1}{2}(\vec{r}_{12}\otimes\vec{r}_{12} - \vec{r}_{13}\otimes\vec{r}_{13}) : \epsilon_\times, \end{aligned} \tag{2.24}$$

or, explicitly:

$$\begin{aligned} F^{+}(t) &= \frac{1}{2}(\cos(2\psi)D^{+}(t) - \sin(2\psi)D^{\times}(t)), \\ F^{+}(t) &= \frac{1}{2}(\sin(2\psi)D^{+}(t) + \cos(2\psi)D^{\times}(t)). \end{aligned} \tag{2.25}$$

The explicit form of $D^{+,\times}$ is derived in Cornish & Rubbo [2003]

$$\begin{aligned} D^{+}(t) &= \frac{\sqrt{3}}{64}\Big[-36\sin^2\theta\sin(2\alpha(t)-2\lambda) \\ &\quad +(3+\cos 3\theta)(\cos 2\phi(9\sin 2\lambda - \sin(4\alpha(t)-2\lambda))) \\ &\quad +\sin 2\phi(\cos(4\alpha(t)-2\lambda) - 9\cos 2\lambda) \\ &\quad -4\sqrt{3}\sin 2\theta(\sin(3\alpha(t)-2\lambda-\phi) - 3\sin(\alpha(t)-2\lambda+\phi))\Big], \\ D^{\times}(t) &= \frac{1}{16}\Big[\sqrt{3}\cos\theta(9\cos(2\lambda-2\phi) - \cos(4\alpha(t)-2\lambda-2\phi)) \\ &\quad -6\sin\theta(\cos(3\alpha(t)-2\lambda-\phi) + 3\cos(\alpha(t)-2\lambda+\phi))\Big]. \end{aligned} \tag{2.26}$$

2.4 $\mathcal{F}$-statistic

In this section, we review the $\mathcal{F}$-statistic based on Jaranowski & Krolak [2005]. In the detector frame, the response function of the detector to a plane GW can be written as a linear combination of n waveforms $h_k(t,\zeta)$ with constant amplitudes $\mathbf{a} = \{a_k, k = 1, \cdots n\}$ and the remaining time dependent parameters called as ζ,

$$h(t,\lambda) = \sum_{k=1}^{n} a_k h_k(t,\zeta). \tag{2.27}$$

The parameters a_k are called extrinsic (or amplitude) parameters whereas the ζ are called intrinsic parameters. This form of GW template is very useful in analysing maximum log likelihood function.

The log likelihood equation (1.30) can be written as:

$$\log \Lambda(s; \mathbf{a}, \zeta) = \sum_{k=1}^{n} (a_k(s|h_k(t,\zeta))) - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i(h_i(t,\zeta)|h_j(t,\zeta))a_j. \quad (2.28)$$

Maximization of $\log \Lambda(s; \mathbf{a}, \zeta)$ over $\mathbf{a}$

$$\frac{\partial \log \Lambda(s; \mathbf{a}, \zeta)}{\partial \mathbf{a}} = 0, \quad (2.29)$$

can be solved explicitly and as a results we obtain the ML estimator $\hat{\mathbf{a}}$ of the parameters $\mathbf{a}$:

$$\hat{\mathbf{a}}(s, \zeta) = \mathbf{M}(\zeta)^{-1}\mathbf{N}(s, \zeta), , \quad (2.30)$$

where the column $n \times 1$ matrix $\mathbf{N}$ and $n \times n$ matrix $\mathbf{M}$ are defined as:

$$\mathbf{N}_k(s, \zeta) = (s|h_k(t, \zeta)), \quad \mathbf{M}_{ij} = (h_i(t, \zeta)|h_j(t, \zeta)). \quad (2.31)$$

Replacing the extrinsic parameters $\mathbf{a}$ in equation (2.28) by their ML estimators $\hat{\mathbf{a}}$, we obtain the log likelihood function maximized over extrinsic parameters

$$\mathcal{F}(s, \zeta) = \log \Lambda(s, \hat{\mathbf{a}}, \zeta) = \frac{1}{2}\mathbf{N}\mathbf{M}^{-1}\mathbf{N}, \quad (2.32)$$

that is called the $\mathcal{F}$-statistic. The $\mathcal{F}$-statistic depends nonlinearly on the intrinsic parameters ζ.

Let us take the white dwarf binaries as an example to show this procedure. In the source frame, the GW template is given as:

$$\begin{aligned} h_+ &= \frac{1}{2}h_0(1 + \cos^2 \iota) \cos(\varphi_0 + \varphi_t) \equiv A_+ \cos(\varphi_0 + \varphi_t), \\ h_\times &= h_0 \cos \iota \sin(\varphi_0 + \varphi_t) \equiv A_\times \sin(\varphi_0 + \varphi_t). \end{aligned} \quad (2.33)$$

where h_0 is the constant amplitude, ι is the inclination of the orbital plane of the binaries to the line of sight, and φ_0 is the initial phase. With the response function, the output in the detector frame should be written as:

$$\begin{aligned} h &= F_+(\theta,\phi,\psi)h_+ + F_\times(\theta,\phi,\psi)h_\times \\ &= \frac{1}{2}\sum_{\mu=1}^{4} a^\mu h_\mu(t), \end{aligned} \tag{2.34}$$

where a^μ are described by the extrinsic parameters $h_0, \iota, \psi, \varphi_0$:

$$\begin{aligned} a^1 &= A_+ \cos 2\psi \cos\varphi_0 + A_\times \sin 2\psi \sin\varphi_0, \\ a^2 &= -A_+ \sin 2\psi \cos\varphi_0 + A_\times \cos 2\psi \sin\varphi_0, \\ a^3 &= -A_+ \cos 2\psi \sin\varphi_0 + A_\times \sin 2\psi \cos\varphi_0, \\ a^4 &= A_+ \sin 2\psi \sin\varphi_0 + A_\times \cos 2\psi \cos\varphi_0, \end{aligned} \tag{2.35}$$

and $h^\mu(t)$ are given by:

$$\begin{aligned} h_1(t) &= D^+(t)\cos\varphi_t, \quad h_2(t) = D^\times(t)\cos\varphi_t, \\ h_3(t) &= D^+(t)\sin\varphi_t, \quad h_4(t) = D^\times(t)\sin\varphi_t. \end{aligned} \tag{2.36}$$

For each set of intrinsic parameters, we maximize $\log\Lambda$ over a_i to get the $\mathcal{F}$-statistic. After finding the maxima of $\mathcal{F}$ in the intrinsic parameters space, we can use the analytic formula equation (2.30) to compute a^μ.

From the expression 2.35, we can see that the value of a^μ are connected to the physical parameters, therefore they have some physical prior distribution. Equation 2.30 does not guaranty that, the parameters deduced from a^μ always correspond to the physical values. In our analysis, we did not check whether each set of a^μ is within the physical range or not. However, when the search converge to a solution, we get the physical parameters by inverting expression 2.35, and we check that they are all within the physical ranges.

Utilizing the expression of equation (2.35), we can derive explicitly the extrinsic parameters of binaries Prix & Whelan [2007].

$$a_s^2 = \sum_{\mu=1}^{4} (a^\mu)^2 = A_+^2 + A_\times^2, \quad D_a = a^1 a^4 - a^2 a^3 = A_+ A_\times. \tag{2.37}$$

Defining the following notations

$$B = A_\times / A_+, \quad b_1 = a^4 - B a^1, \quad b_2 = a^3 + B a^2, \quad b_3 = B a^4 - a^1. \tag{2.38}$$

we have

$$\psi = \frac{1}{2} \tan^{-1} \left(\frac{b_1}{b_2} \right), \varphi_0 = \tan^{-1} \left(\frac{b_2}{b_3} \right), h_0 = A_+ + \sqrt{A_+^2 - A_\times^2}, \cos \iota = \frac{A_\times}{h_0} \tag{2.39}$$

For spinning SMBH binaries, we can apply this procedure to maximize over only two extrinsic parameters, because inclination angle is not a constant due to the orbital precession (see Chapter 3). For phenomenological waveform of EMRIs (Chapter 4) we apply this maximization to each harmonics of the signal.

2.5 Doppler modulation

The motion of the detector relative to the source leads to the phase modulation of the measured GW signal. In Maggiore [2007], it is shown that this modulation can be calculated by the coordinate transformation between time of arrival at the solar system barycenter and the time at the detector. Let S_d and S_b be signals in the detector and the barycenter frame respectively, they are related as

$$S_d(t_d) = S_b(t_b[t_d, \theta, \phi]), \tag{2.40}$$

where (θ, ϕ) is the angular position of the source in SSB frame LISATeam [2000]. We need to measure the t_d which is related to t_b as:

$$t_d = t_b[t_d, \theta, \phi] - \frac{\vec{n}(\theta, \phi) \cdot \vec{r}(t_d)}{c}, \tag{2.41}$$

with $\vec{n}$ being a unit vector pointing towards the source , and $\vec{r}$ is a vector connecting the guiding center of LISA and the sun.

$$\vec{n} = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta), \quad \vec{r} = R\left(\cos\frac{2\pi t}{T}, \sin\frac{2\pi t}{T}, 0\right). \tag{2.42}$$

where R is the distance from the guiding center to the sun.

Explicitly, the relation between the two signals S_d and S_b as function of time is

$$S_d(t_d) = S_b\left(t_d + R\sin\theta\cos(\frac{2\pi t}{T} - \phi)\right). \tag{2.43}$$

This implies that, the phase of signal as observed by the detector is modulated due to the detector motion. Take the white dwarf binaries as an example, the signal with the doppler shift in the detector frame is assuming constant frequency:

$$h_+(t) = A_+\cos\left[\phi_0 + 2\pi f_{GW}t_b + 2\pi f_{GW}R\sin\theta\cos\left(\frac{2\pi t}{T} - \phi\right)\right]. \tag{2.44}$$

The detector motion also cause the amplitude modulation as explicitly given as time dependented terms $D^+, D^\times$ equations (2.26).

Chapter 3

The search for spinning black hole binaries in mock LISA data using a genetic algorithm

This chapter is based on the my published work Petiteau *et al.* [2009b, 2010].

In section 1.3.2.2, we have mentioned that coalescing massive Black Hole binaries are the strongest and probably the most important gravitatinal wave sources in the LISA band. The spin and orbital precessions bring complexity in the waveform and make the likelihood surface richer in structure as compared to the non-spinning case. We introduced an extended multimodal Genetic Algorithm which utilizes the properties of the signal and the detector response functions to analyze the data from the third round of mock LISA data challenge. The third round of MLDC consisted of five challenges but in this work we focused our attention on the data set 3.2 which contained GW signals from 4-6 binaries of spinning SMBHs (exact number was not revealed to the participants), on top of the confusion galactic binaries background and the instrumental noise. These data was an improvement upon the MLDC challenges 1.2 and 2.2 by adding spins to SMBHs. The spin-spin and spin-orbital coupling causes the orbital and spins precession which results in the modulation of the amplitude and phase of the GW signal. The prior range on the parameters and the detailed set up of the challenge can be found in Babak *et al.* [2008a]. We have found all the five sources present

in MLDC 3.2 and recovered the coalescence time, chirp mass, mass ratio and sky location with reasonable accuracy. As for the orbital angular momentum and two spins of the Black Holes, we have found a large number of widely separated modes in the parameter space with similar maximum likelihood values.

3.1 My contribution

During this work, I wrote the Matlab code based on the Genetic Algorithm to test its performance and efficiency in searching the signal from non-spinning SMBH. Due to the poor efficiency of Matlab itself, the search was implemented in C++ code by Dr. Antoine Petiteau with my participation. I have participated in the discussion on the materials given in the section 3.2, namely, the detection statistic, termination of the waveform, and A-statistic. Some standard accelerators which are used in the Genetic Algorithm implementations could be found in the literature, but they do not give sufficient performance in our problem. Therefore, I suggested and implemented some new accelerators in our search: brother system, alternation of Gray code and Binary code, change of environment, (the detailed descriptions are given in section 3.4.2). I have participated in the discussion with my colleagues on the design of the multi step Multimodal Genetic Algorithm which is explained in section 3.5. I have run the multi-steps MGA code, tuning the parameter which are used to control the search, and analyzed the results which are shown in the section 3.7. I have concentrated on searching the weakest signal with time of coalescence outside the observation period, and found it successfully. We are the only one group who found all five signals, no other groups recover the fifth one which was found by me during the runs.

3.2 Formulation of the problem

3.2.1 Model of the template

The signal used in MLDC is modeled as the amplitude-restricted waveform (i.e. only dominant harmonic at the leading order is used) with the phase taken up to the second Post-Newtonian (PN) order which includes the leading order con-

tributions from the spin-orbital and spin-spin coupling. The binary evolution is described as a quasi-circular adiabatic inspiral following by the merger and ringdown process.

The waveform is described by fifteen parameters which are: the two masses m_1 and m_2, the time at coalescence t_c, the sky location of the source in ecliptic coordinates, latitude θ and longitude ϕ, the dimensionless spin parameters, χ_1 and χ_2, the initial direction of the spins, polar angles θ_{S_1} and θ_{S_2} and azimuthal angles ϕ_{S_1} and ϕ_{S_2}, the initial direction of the orbital angular momentum, polar angle θ_L and azimuthal angle ϕ_L, the phase at coalescence Φ_c, and the luminosity distance D_L.

We denote the unit vector in the direction of the orbital angular momentum as $\hat{L}$ and two spins are $\vec{S_1} = \chi_1 m_1^2 \hat{S}_1$, $\vec{S_2} = \chi_2 m_2^2 \hat{S}_2$, where $\hat{S}_{1,2}$ are unit vectors and $0 < \chi_{1,2} < 1$. The precession equations are given in Apostolatos *et al.* [1994]

$$\dot{\vec{S}_1} = \frac{(M\omega)^2}{2M}\left\{\eta(M\omega)^{-1/3}\left(4+\frac{3m_2}{m_1}\right)\hat{L} + \frac{1}{M^2}\left[\vec{S_2} - 3(\vec{S_2}\cdot\hat{L})\hat{L}\right]\right\} \times \vec{S_1}, \quad (3.1)$$

$$\dot{\vec{S}_2} = \frac{(M\omega)^2}{2M}\left\{\eta(M\omega)^{-1/3}\left(4+\frac{3m_1}{m_2}\right)\hat{L} + \frac{1}{M^2}\left[\vec{S_1} - 3(\vec{S_1}\cdot\hat{L})\hat{L}\right]\right\} \times \vec{S_2}, \quad (3.2)$$

$$\begin{aligned}\dot{\hat{L}} = -\frac{(M\omega)^{1/3}}{\eta M^2}\left(\dot{\vec{S}_1}+\dot{\vec{S}_2}\right) \;=\; & \frac{\omega^2}{2M}\left\{\left[\left(4+\frac{3m_2}{m_1}\right)\vec{S_1} + \left(4+\frac{3m_1}{m_2}\right)\vec{S_2}\right]\times\hat{L}\right. \\ & \left. -\frac{3\omega^{1/3}}{\eta M^{5/3}}\left[\left(\vec{S_2}\cdot\hat{L}\right)\vec{S_1} + \left(\vec{S_1}\cdot\hat{L}\right)\vec{S_2}\right]\times\hat{L}\right\} \quad (3.3)\end{aligned}$$

The modulation of the waveform due to the presence of spins is taken at the leading order.

The orbital angular frequency with spin effect up to 2 PN order is given by

$$\begin{aligned}M\omega \;=\; & \frac{1}{8}\tau^{-3/8}\left[1 + \left(\frac{743}{2688}+\frac{11}{32}\eta\right)\tau^{-1/4} - \frac{3}{10}\left(\pi - \frac{\beta}{4}\right)\tau^{-3/8}\right. \\ & \left. + \left(\frac{1855099}{14450688}+\frac{56975}{258048}\eta + \frac{371}{2048}\eta^2 - \frac{3}{64}\sigma\right)\tau^{-1/2}\right], \quad (3.4)\end{aligned}$$

where $M = m_1 + m_2$ is total mass, $\eta = \frac{m_1 m_2}{M^2}$ is the symmetric mass ratio and

$$\tau = \frac{\eta}{5M}(t_c - t), \tag{3.5}$$

$$\beta = \frac{1}{12}\sum_{i=1,2}\left[\chi_i\left(\hat{L}\cdot\hat{S}_i\right)\left(113\frac{m_i^2}{M^2} + 75\eta\right)\right], \tag{3.6}$$

$$\sigma = -\frac{1}{48}\eta\chi_1\chi_2\left[247\left(\hat{S}_1\cdot\hat{S}_2\right) - 721\left(\hat{L}\cdot\hat{S}_1\right)\left(\hat{L}\cdot\hat{S}_2\right)\right]. \tag{3.7}$$

In our following analysis, we used η and the chirp mass $M_c = \frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}}$ as independent parameters instead of m_1 and m_2.

The intrinsic phase is

$$\begin{aligned}\Phi_{\rm orb} = \Phi_C - \frac{\tau^{5/8}}{\eta}\Bigg[1 + \left(\frac{3715}{8064} + \frac{55}{96}\eta\right)\tau^{-1/4} - \frac{3}{16}\left(4\pi - \beta\right)\tau^{-3/8} \\ + \left(\frac{9275495}{14450688} + \frac{284875}{258048}\eta + \frac{1855}{2048}\eta^2 - \frac{15}{64}\sigma\right)\tau^{-1/2}\Bigg].\end{aligned} \tag{3.8}$$

The phase is defined with respect to the ascending node, however the spin-orbital coupling causes precession of the orbit, therefore we need to introduce precessional correction to the phase, $\delta\Phi(t)$. It depends on the unit vector $\hat{n}$ pointing from the solar system barycenter to the source:

$$\Phi(t) = \Phi_{\rm orb}(t) + \delta\Phi(t), \tag{3.9}$$

$$\dot{\Phi}(t) = \omega + \delta\dot{\Phi} = \omega + \frac{(\hat{L}\cdot\hat{n})(\hat{L}\times\hat{n})\cdot\dot{\hat{L}}}{1 - (\hat{L}\cdot\hat{n})^2}, \tag{3.10}$$

$$\delta\Phi(t) = -\int_t^{t_c}\left(\frac{\hat{L}\cdot\hat{n}}{1 - (\hat{L}\cdot\hat{n})}\right)(\hat{L}\times\hat{n})\cdot\dot{\hat{L}}dt. \tag{3.11}$$

The gravitational wave polarization components in the source frame are given by

$$h_+ = h_+^0\cos 2\Phi = -\frac{2M\eta}{D_L}(1 + \cos^2\iota)(M\omega)^{2/3}\cos 2\Phi,$$

$$h_\times \;=\; h^0_\times \sin 2\Phi = \frac{4M\eta}{D_L}\cos\iota (M\omega)^{2/3}\sin 2\Phi, \tag{3.12}$$

where $\cos\iota = \hat{L}\cdot\hat{n}$.

Based on the discussion in the section 2.3, the strain $h(t)$ that the GW exerts on the LISA detector is the following linear combination of h_+ and $h_\times$

$$h(t) = F_+(\theta,\phi)h^S_+(t) + F_\times(\theta,\phi)h^S_\times(t). \tag{3.13}$$

The polarization components in the radiation frame, h^S_+ and $h^S_\times$, are expressed as

$$\begin{aligned} h^S_+ &= -h_+\cos 2\psi - h_\times \sin 2\psi, \\ h^S_\times &= h_+\sin 2\psi - h_\times\cos 2\psi, \end{aligned} \tag{3.14}$$

where the polarization angle ψ is defined by

$$\tan\psi = \frac{\sin\theta\cos(\lambda-\phi_L)\sin\theta_L - \cos\theta_L\cos\theta}{\sin\theta_L\sin(\phi-\phi_L)}. \tag{3.15}$$

Note, due to the precession of the orbital plane, this polarization angle varies during the evolution of the binary.

The data sets in MLDC are the TDI variables which are already explained in section 2.2. In our search, we adopted the two orthogonal (i.e. with uncorrelated noise) TDI channels, A and E, in the phase domain (i.e. strain). In our template, we considered a long wavelength approximation to these signals Cornish & Rubbo [2003]; Petiteau [2008]. This approximation ($L\omega \ll 1$, where L is LISA' armlength and ω is an instantaneous frequency of GW) works pretty well below approximately 5 mHz Cutler [1998]. Assuming rigid LISA with equal arms, the waveform sampled at discrete times takes the following form Cornish & Rubbo [2003]; Petiteau [2008]

$$\begin{aligned} h_I(t_k) \;\simeq\; & 2L\;\sin\Delta\phi_{2L}(t_k)\times \\ & \{-h_{+0}(t_k)\left[\cos\left(2\psi(t_k)\right)F_{+I}(t) - \sin\left(2\psi(t_k)\right)F_{\times I}(t)\right]\sin\phi'(t_k) \\ & +h_{\times 0}(t_k)\left[\sin\left(2\psi(t_k)\right)F_{+I}(t) + \cos\left(2\psi(t_k)\right)F_{\times I}(t)\right]\cos\phi'(t_k)\}, \end{aligned} \tag{3.16}$$

where $I = \{A, E\}$, $\Delta\phi_{2L}(t) = (\phi(t_k) - \phi(t_k - 2L))/2$, $\phi'(t) = (\phi(t_k) + \phi(t_k - 2L))/2$ with $\phi(t)$ being the phase of GW and $t_k = t + \widehat{n} \cdot \vec{r}$ is the time in LISA frame with $\vec{r}$ the vector from the Sun to LISA barycenter. Based on the formula in the section 2.3, the antenna pattern functions F_{+I} and $F_{\times I}$ corresponding to the TDI channel, have the following expressions

$$\begin{aligned} F_+(\theta_d, \lambda_d; t, \Omega) &= \frac{1}{32}\left[6\sin(2\theta_d)\left(3\sin\left(\Phi_T(t) + \lambda_d + \Omega\right) - \sin\left(3\Phi_T(t) - \lambda_d + \Omega\right)\right)\right. \\ &\quad -18\sqrt{3}\sin^2\theta_d \sin\left(2\Phi_T(t) + \Omega\right) - \sqrt{3}\left(1 + \cos^2\theta_d\right) \times \\ &\quad \left.\left(\sin\left(4\Phi_T(t) - 2\lambda_d + \Omega\right) + 9\sin\left(2\lambda_d + \Omega\right)\right)\right], \qquad (3.17) \\ F_\times(\theta_d, \lambda_d; t, \Omega) &= \frac{1}{16}\left[\sqrt{3}\cos\theta_d(\cos(4\Phi_T(t) - 2\lambda_d + \Omega) - \right. \\ &\quad 9\cos(2\lambda_d + \Omega)) + 6\sin\theta_d(\cos(3\Phi_T(t) - \lambda_d + \Omega) + \\ &\quad \left. 3\cos(\Phi_T(t) + \lambda_d + \Omega))\right], \qquad (3.18) \end{aligned}$$

with $\theta_d = \beta + \pi/2$, $\lambda_d = \lambda + \pi$, $\Phi_T(t) = 2\pi t/Year$ and $\Omega = 0, -\pi/2$ for channels A and E respectively. Note that this is the long wavelength approximation to the signal injected in the simulate data, we found it to be a reasonably accurate representation until the last $1 - 2$ cycles before the merger. The end of the signal is discussed in more detail later.

3.2.1.1 Maximization of the likelihood

As we mentioned in the section 1.4, the signal from one detector is

$$s_i(t) = h_i(t, \hat{\lambda}) + n_i(t), \qquad (3.19)$$

where $h_i(t, \hat{\lambda})$ is a signal described by a set of parameters $\hat{\lambda}$ and $n_i(t)$ is the stationary Gaussian noise characterized by the power spectral density (PSD) $S_n(f)$.

The noise is the sum of instrumental noise $S_n^{\text{inst.}}(f)$ and the GW confusion noise from Galactic binaries $S_n^{\text{Gal. Bin.}}(f)$. In strain data (i.e. phase measurements), the

instrumental noise for TDI variables A and E is described by the following PSD

$$\begin{aligned} S_n^{\text{inst.}}(f) &= 16\sin^2(\phi_L)\left(S_n^{\text{OPN}}(f) + (\cos(2\phi_L))\, S_n^{\text{acc.}}(f)\right) \\ &\quad -4\sin(2\phi_L)\sin(\phi_L)\left(4S_n^{\text{OPN}}(f) + S_n^{\text{acc.}}(f)\right), \end{aligned} \tag{3.20}$$

where the acceleration noise is $S_n^{\text{acc.}}(f) = 5.75 \times 10^{-53}(f^{-4} + 10^{-8} f^{-6})\ Hz^{-1}$ and the optical path noise and the shot noise are $S_n^{\text{OPN}}(f) = 3.675 \times 10^{-42}\ Hz^{-1}$.

In section 1.3.2.1, we have explained the Galactic GW confusion noise is a combination of the unresolved signals from ~ 30 millions of white dwarf binaries. This noise is modeled by the following function, in units of Hz^{-1}, Nelemans *et al.* [2004]; Timpano *et al.* [2006]

$$S_n^{\text{Gal. Bin.}}(f) = \begin{cases} 10^{-44.62} f^{-2.3} & 10^{-4} \leq f \leq 10^{-3} \\ 10^{-50.92} f^{-4.4} & 10^{-3} \leq f \leq 10^{-2.7} \\ 10^{-62.8} f^{-8.8} & 10^{-2.7} \leq f \leq 10^{-2.4} \\ 10^{-89.68} f^{-20} & 10^{-2.4} \leq f \leq 10^{-2} \end{cases} \tag{3.21}$$

The goal of the method presented in this chapter is to find the maximum of the likelihood in the 15-dimensional parameter space, and, thus, obtain the maximum likelihood estimation of the parameters. The value of the likelihood tells us also about the statistical significance of the detected event. In the case of LISA data, the signals usually have high SNR, so the probability of the false detection is rather low. However, the data is signal dominated and several GW signals of one type (say, Galactic binaries) could conspire and produce significantly high SNR at the output of the matched filtering during the search for another type of signal (say, SMBH binary) Racine & Cutler [2007].

By using the maximum likelihood method we have discussed in the chapter 2, we maximized the log likelihood over two parameters: the luminosity distance D_L and the phase at coalescence Φ_c. We call the resulting function Maximized Likelihood (or quality).

Following the formalism described in section 2.4, the GW template equation (3.16) can be express as

$$h_I(t) = a_1\ h_{1I}(t) + a_2\ h_{2I}(t), \tag{3.22}$$

with

$$\begin{cases} a_1 = \cos(2\phi_c)/D_L \\ a_2 = \sin(2\phi_c)/D_L \end{cases}, \quad \begin{cases} h_{1I} = h_I(D_L = 1Gpc;\ \phi_c = 0) \\ h_{2I} = h_I(D_L = 1Gpc;\ \phi_c = \pi/4). \end{cases} \tag{3.23}$$

Here, $I = \{A, E\}$ stand for the two pipeline.

Using the above expressions and the orthogonality $\tilde{h}_{2I} \simeq i\ \tilde{h}_{1I}$ we obtained the maximized log likelihood.

$$\mathcal{F} \simeq \frac{\left(\sum_I < s_I|h_{1I} >\right)^2 + \left(\sum_I < s_I|h_{2I} >\right)^2}{\left(\sum_I < h_{1I}|h_{1I} >\right)^2}. \tag{3.24}$$

This is not the full F-statistic which is derived in the case of white dwarf binaries in Chapter 2. Here, we just maximize the log-likelihood with respect to only two extrinsic parameters: D_L and Φ_c.

3.2.1.2 Maximization over the time of coalescence

In order to efficiently find the time of coalescence, we use correlation in place of the inner products. Given a template h which is constructed with the initial value (usually taken at the lower edge of the prior) $t_{c,0}$ and using the inverse Fourier transform, we find the value of $\tau_{\max}$ which maximizes equation (3.24) or which is almost equivalent to maximizing

$$c(\tau) = 2\int_0^\infty df\ \frac{\widetilde{h}(f)\ \widetilde{s}^*(f) + \widetilde{h}^*(f)\ \widetilde{s}(f)}{S_n(f)} e^{-2i\pi f\tau}. \tag{3.25}$$

The equation (3.25) help us to evaluate matched filtering of the time-shifted template with the data for multiple lags in the very efficient way by using the fast Fourier transformation. If the intrinsic parameters are similar to the true value, the maximum of $c(\tau)$ will be near the true time of coalescence. This is because the frequency of the shift waveform matches very well with the signal when the shift time is close to the true t_c.

Note that the amplitude of the signal depends on the choice of t_c via annual modulation caused by LISA's orbital motion, therefore the new value $t_{c,1} = t_{c,0} + \tau_{\max}$ is not necessarily the final answer. The time of coalescence which max-

imizes the quality equation (3.24) for given other parameters should correspond to maximum of equation (3.25) at zero (or almost zero) lag. Using the new value of t_c we repeat the maximization, and we stop iterations when the difference $|t_{c,i} - t_{c,i-1}|$ is sufficiently small. Usually few iterations are sufficient to find t_c which maximizes the quality ,it is much faster than the random jump when we just take t_c as a waveform parameter. We have performed tests on these two methods, the results show that the use of equation (3.25) is much more efficient to significantly reduce the search range of t_c, and converge to the true value. In practice, we start the search using the Fourier transform trick, but once the value of t_c does not change significantly from the generation to generation, we treated it as a search parameter within GA and refine it.

3.2.1.3 The waveform termination

The signal from SMBH binaries is band limited, the lower frequency limit is defined approximately by twice the orbital frequency at $t = 0$. The upper frequency is introduced somewhat arbitrarily. To terminate both the signal and the template smoothly, an exponential taper is applied. The taper affects the data when two black holes are separated by a distance $R = 7M$ and kills the signal completely around $R = 6M$ (which is the last stable orbit for the test mass in Schwarzschild space-time). Therefore, in computing the overlaps, we used the maximum frequency in the integration corresponding to the orbital separation $6M$:

$$f_{\max} = \frac{1}{\pi M (R/M)^{3/2}} = \frac{\eta^{3/5}}{\pi (R/M)^{3/2} M_c}. \tag{3.26}$$

The exponential taper causes problems for the long-wavelength approximation, and our template deviates from the signal during the last cycle. Unfortunately these small deviations fall in the most sensitive part of the LISA band and are further enhanced by high SNR. This causes a significant problem: the bias caused by this deviation is unacceptably large because there is a large region of the parameter space that produces templates which fit the end part of the signal perfectly (using incorrect parameters) but fail to reproduce the low frequency part of the signal.

In order to solve this problem we terminate the template waveform few cycles earlier by fixing cutoff frequency which corresponds to the orbital separation $R > 7M$. Our approximation becomes better as we go to lower frequencies, however we start losing power of the signal (SNR) which is highly undesirable. We automatically readjust the frequency cut-off if the SNR of truncated template drops below a certain threshold ($SNR = 20$). It is accomplished by starting with a template terminated initially at the orbital separation $R = 60M$, if its SNR is smaller than the threshold, we terminate it at smaller orbital separation, decreasing it step by step until the template gives sufficient SNR.

We want to emphasize a very important feature which accompanies the earlier termination of the waveform. The map of the maximized log likelihood changes: in the Figure 3.1 we show the map of the quality in the "chirp mass" - "eta" plane keeping other parameters fixed to their true values. On the left panel we show $\mathcal{F}_{full}$ (we use no frequency cut off other that introduced by the taper), and, on the right panel, we plot $\mathcal{F}_{cut}$ with template cut at $f_{max} = f_{cut} = 0.26\,\mathrm{mHz}$. One can see multiple maxima in both plots, but(!) the position of the secondary maxima are different whereas the location of the true (global) maximum (indicated by an arrow) is the same. It can also be seen that the size of the secondary maxima on the right panel is smaller. We will use these features later in our search.

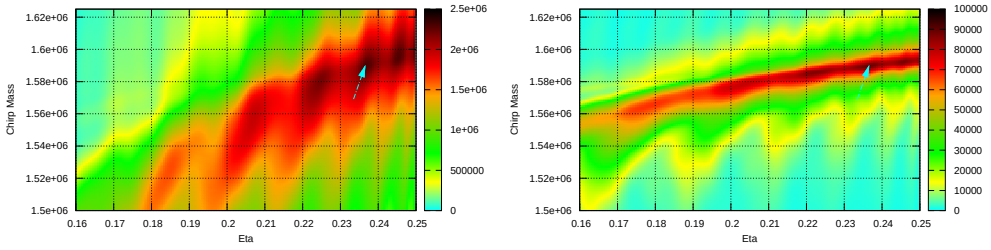

Figure 3.1: Distribution over M_c and η, of the Maximized Likelihood (quality) computed with the full waveform on left panel and with the waveform cut at $f_{max} = 0.26$ on the right panel. This example corresponds to a signal with the following parameters: $\theta = -0.38896\ rad$, $\phi = 3.28992\ rad$, $t_c = 19706568.3273\ sec$, $M_c = 1589213.34\ M_\odot$, $\eta = 0.23647$, $\theta_L = 2.78243\ rad$, $\phi_L = 1.53286\ rad$, $\chi_1 = 0.24115$, $\chi_2 = 0.16145$, $\theta_{S1} = 1.20839\ rad$, $\phi_{S1} = 5.61808\ rad$, $\theta_{S2} = 0.39487\ rad$, $\phi_{S2} = 5.82937\ rad$, $D_L = 6856164697.8\ parsec$, $\phi_c = 4.96746\ rad$. The arrow points to the true parameters.

3.2.1.4 *A*-statistic

Chopping the template at lower frequency solves the problems mentioned above but is not completely satisfactory. We lose some SNR and consequently some accuracy in the parameter estimation, we also lose information stored at the end of the signal which is especially important to recover spin-related parameters. In order to reduce the impact of the coalescence part, without killing it completely, we introduce a new function, called *A-statistic* which is simply a geometrical mean of the Maximized Likelihood of the cut waveform and the Maximized Likelihood of the full waveform:

$$\mathcal{A} = \sqrt{\mathcal{F}_{\text{cut}} \times \mathcal{F}_{\text{full}}}. \tag{3.27}$$

A-statistic is not log likelihood anymore, but one of its advantages is that it keeps the information from the full waveform including the coalescence but at the same time it enhances the information coming from the low-frequency part. *A*-statistic also reduces the number of local maxima as can be seen in the Figure 3.2. In this example we have reduced the size and number of maxima from five to three.

3.3 Genetic Algorithm

3.3.1 The basic principle

In order to find all the parameters of the signal, we need an effective algorithm to search over the 13 dimensional parameter space. Building the grid in the multi-dimensional parameter space is a highly non-trivial problem. The use of the stochastic/random bank Babak [2008]; Harry *et al.* [2008]; Manca & Vallisneri [2010]; Messenger *et al.* [2009] is a feasible method for the template placement, however a full grid scan over the whole parameter space would be prohibitively computationally expensive.

Alternative would be to use variations of the Markov chain Monte-Carlo Cornish & Porter [2007] or nested sampling Feroz *et al.* [2009] methods. Here we have chosen to use Genetic Algorithm (adjusted to our needs) to search for the global maximum of the likelihood in multi-dimensional parameter space.

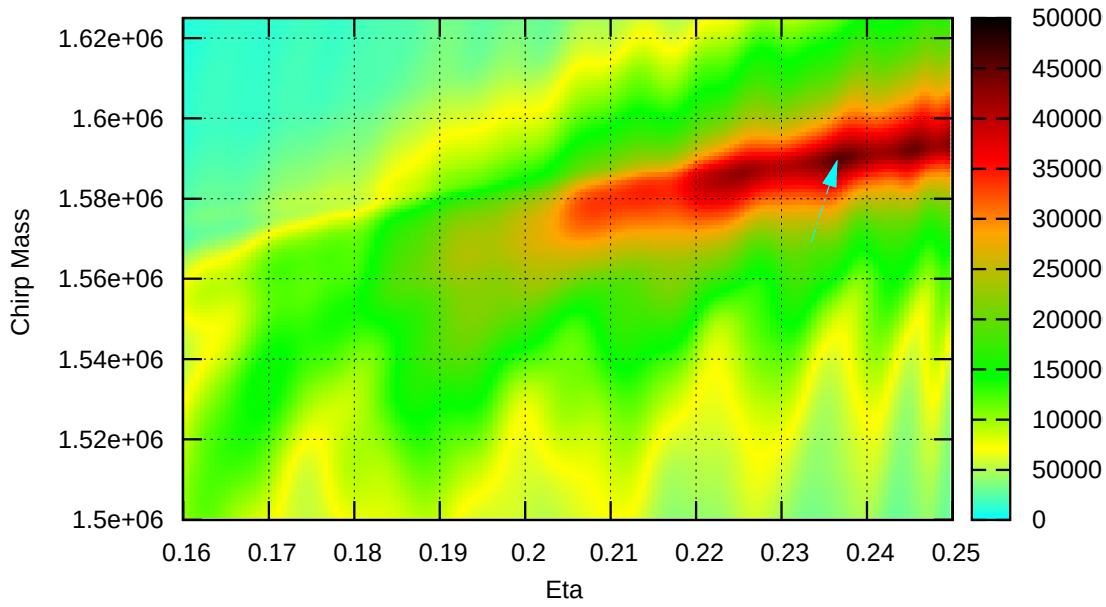

Figure 3.2: Distribution of A-statistic over M_c and η. This example corresponds to the same signal as in Figure 3.1. The arrow points to the true location of parameters of the signal.

The GA is derived from the computer simulations of the biological system, which were originally introduced by Professor Holland and his students in Michigan University. It is a method for the global search (optimization method) based on the natural selection principle - the basis for the evolution theory established by C. Darwin. In the nature, organisms adapt themselves to their environment: the smartest/strongest/healthiest organisms are more likely to survive and participate in the breeding to produce the offsprings. These two processes, selection and breeding, are used in Genetic Algorithms to produce a new generation of organisms. Since the best organisms are more likely to participate in breeding, the new generation should be better than the previous one (at least no worse). So this procedure induces the evolution of the organism, just like in the nature, the good qualities of the parents can be transferred to their offsprings.

In the biological world, besides these two basic operations, among every generation, there are always few individuals which have better characteristics to adopt to the environment, produced as a result of a positive mutation. By introducing the new genotype into the population, mutation can potentially improves the

forthcoming generations and consequently accelerates the evolution towards the global maximum.

Some measure of "goodness" needs to be associated with each organism.

In the case of gravitational wave search, it is natural to associate the logarithm of the likelihood (or any other equivalent detection statistic e.g. Maximized Likelihood or A-statistic discussed in the previous section) with the "goodness" which needs to be "improved" through the evolution of the organisms. We will call the chosen measure of "goodnees", the *quality* of an organism Q.

Following is a brief description of how a typical GA works. We start with a randomly chosen group of organisms (templates), we evaluate the quality of each organism (log likelihood). We select set of pairs (parents) based on their qualities, the organisms with better quality (templates with higher likelihood) are chosen more often than weak organisms. We combine genotype of two parents to produce a child (we combine parameters of two chosen templates to produce a new one). Number of produced children is equal to the number of parents (we keep number of evolving organisms (generation) fixed). Next we allow with a certain probability of a random mutation in the children's genes (with some probability we randomly change the parameters of the new templates, exploring a larger area of the parameter space). The parents are discarded and the resulting children form a new generation. We repeat the procedure until we reach steady state (maximum in the quality). In this simple example we keep only one generation active (one group of templates).

A list of (biological) GA terms with the equivalent terms in GW data analysis is given in the Table 3.1.

Genetic Algorithm		GW search
organism	$\Longleftrightarrow$	template
gene (of an organism)	$\Longleftrightarrow$	parameter (of a template)
allele (of a gene)	$\Longleftrightarrow$	bits (of the value of the parameter)
quality Q	$\Longleftrightarrow$	Maximized Likelihood or A-statistic
colony of organisms	$\Longleftrightarrow$	evolving group of templates
n-th generation	$\Longleftrightarrow$	the state of colony at n-th step of evolution
(selection + breeding) + mutation	$\Longleftrightarrow$	way of exploring the parameter space

Table 3.1: Relation between GA and GW notions.

In the following subsections we give a detailed description of each element of the basic GA and then we introduce the specific modifications to speed up the search.

3.3.2 Code of the gene

As we have discussed above, every organism is associated with a template and the parameters of the template play the role of genes. So each organism is described by 15 genes, two of them are chosen optimally (maximization of the log likelihood, see Section 3.2.1.1) and the gene corresponding to the time of coalescence is efficiently found using correlation (see Section 3.2.1.2). We imitate the DNA structure by describing the gene (parameter value) by a set of alleles. In our implementation we adopt a binary representation of the gene (parameter) which means that each allele (bit) has two possible values: 0 or 1. In practice we first fix the precision of each parameter (by fixing the number of significant digits in the decimal format) and then we translate it to standard binary and/or in Gray form. In our method we use both representations, the reason will be explained later when we discuss quantization issue.

Let us show how this is done in practice. Consider a parameter λ_k with the uniform prior range $[x_{k,min}, x_{k,max}]$. First we convert a value x_k of λ_k into an integer $c_k = (x_k - x_{k,min})/\Delta x_k$ where $\Delta x_k = (x_{k,max} - x_{k,min})/2^{N_k}$ is the resolution of θ_k and N_k is the number of bits. Then, we convert c_k into the set of bits $b_k[i]$ using the coding rule of the chosen representation. As we see, the resolution for each parameter depends on the number of bits N_k used for describing it and is the same for both representations. The importance of the bit is determined by its position. A change of a bit in a higher position (significant bit) corresponds to a big change in the parameter value. In our convention, the first bit, $b_k[0]$, is the lowest significant bit and the last bit, $b_k[N-1]$, is the highest significant bit.

There is a close relationship between two gene representations. We can transform the binary representation to Gray representation by the following procedure: given a string of binary code with N bits $\{B[0], B[1], \cdots, B[N-1]\}$, we set

$B[N] = 0$ then the Gray code with the same N bits is

$$G[i] = B[i+1] \wedge B[i], \tag{3.28}$$

where the operator "$\wedge$" corresponds to the XOR operator in programming languages. Other way round, by setting again $B[N] = 0$, we can get the binary representation with N bits from the Gray representation as

$$B[i] = B[i+1] \wedge G[i]. \tag{3.29}$$

3.3.3 Selection

The selection process chooses the parents for breeding. The probability of selecting an organism is defined by its quality. Organisms with higher quality have a better chance of being chosen to participate in the breeding.

First the quality Q_i, i.e. the Maximized Likelihood or A-statistic, for all organisms is computed (index i refers to the i-th organism).

Then each organism is assigned the probability of being chosen for breeding as $p_i = Q_i / \sum_j^N Q_j$. The selection is made by the roulette selection method: we choose a random number uniformly from [0,1]; if it is bigger than p_i and smaller than p_{i+1}, then the i^{th} organism is selected.

This selection ensures that the "good" organisms are chosen more often than the "bad" ones and guarantees that the genotype responsible for a high quality propagates in generations approaching the optimal value.

In our selection rule we do not take into account the geographical proximity between parents (in other words possible correlation between templates in the same generation). By forbidding the breeding between the correlated parents, it might be more efficient to explore a larger region of parameter space, but the overall resolution of the method will be reduced. We therefore do the selection based only on the quality.

3.3.4 Breeding

After selecting the parents, we need to produce a new generation, this can be achieved through "breeding". Breeding is the rule according to which a child is produced from the selected parents. The genes of the child are constructed by mixing the corresponding genes of each parent. We take one part from the first parent and the other part from the second one. Depending on which parts are chosen, there are several types of breeding. We usually use three different types: cross-over one random point, cross-over two fixed points, and random. For the cross-over one point, we choose one bit (denoted by i) randomly as the cross-over point and the child's genes are created by combining the first i bits of the genes from the first parent with the last $N - i$ bits of the genes of the second parent (see the left panel of the Figure 3.3). For the cross-over two fixed points, the genes of the child are built from three equal parental parts (see the middle panel of the Figure 3.3). In the random breeding, each child's bit is chosen randomly from the corresponding bits of the parents (see the right panel of the Figure 3.3).

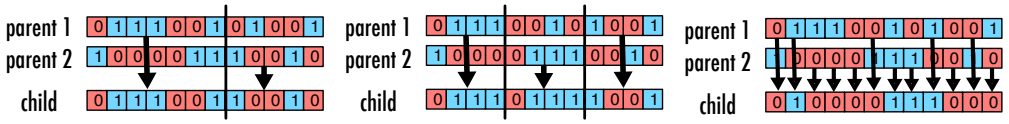

Figure 3.3: Examples of used breeding: cross-over one random point on the left, cross-over two fixed points in the middle and random on the right panels.

3.3.5 Mutation

The first generation is chosen randomly by drawing parameters uniformly within the priors specified in Babak *et al.* [2008a]. The chosen selection implies that the best quality of our organisms is likely to be increased with each generation. But, if we use only these two processes, the range of resulting genes is quite restricted: it totally depends on the initial random state and is just a combination of the parts from the first generation. The combination of genes and therefore the exploration of the parameter space is very limited and completely dependent on the initial choice. This undesirable feature can be cured by introducing mutation.

Mutation in GA works in a way similar to how it operates in the nature. Mutation is a random change of few alleles in a gene of an organism; in our algorithm it corresponds to changing few bits in a representation of a parameter value of a template. The probability of mutation is called the probability mutation rate (PMR). We mutate each gene of each child independently and there are several types of mutation. First we need to decide whether we mutate a gene or not, and, if yes, we need to decide on the mutation rule (how we do it). The first possibility is that we always mutate the gene and mutation is applied to each bit of gene independently. Each bit is flipped with probability PMR. The second possibility is to mutate a gene with probability PMR. In this case we have used two different rules to mutate the gene: (i) we flip N randomly chosen bits (ii) we flip N adjacent bits. In the case of (ii), the start point is chosen randomly. The value of N and PMR depends on the parameters, we usually choose N around 6 or 7, PMR decreases from 0.7 to 0.01.

Different types of mutations together with the value of PMR define the exploration area of the parameter space. An example is shown in Figure 3.4, in which we start with PMR = 0.5 at the beginning of the search (left panel, one can see that the templates are scattered all over the space) and then slowly reduce it to PMR = 0.01 (the right panel). The true solution is located in the center of the blue circle. We will come back to PMR again in the Section 3.4.1.3.

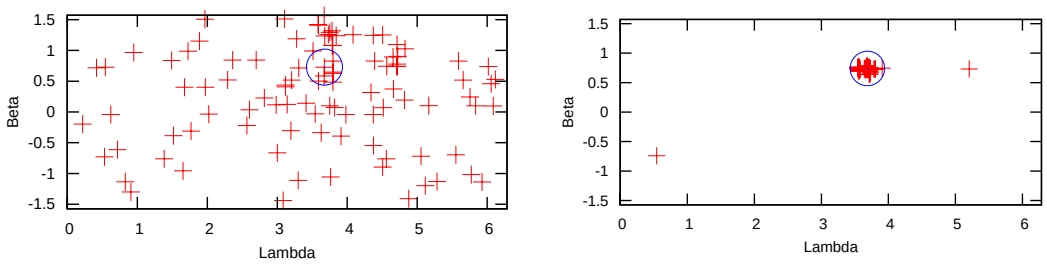

Figure 3.4: Example of distribution of 100 organisms in two dimensions space which is the sky position (θ,ϕ). The left panel shows the case of large PMR value (0.5) which corresponds to the beginning of the search. The right-panel shows the case of small PMR value (0.01) which corresponds to the end of the search. The best organism as well as the true solution is at the center of the blue circle.

3.3.6 Tuning the algorithm using code, breeding and mutation

In order to get comprehensive understanding of all kinds of representation, breeding and mutation to tune our search, we did the systematic experiments to test one degree of freedom at the time. We fixed GA configuration and allow only one parameter to vary and analyzed the results of the search. We have tested PMR, number of organisms in the generation, gene representation (binary, Gray, alternating) type of breeding and mutation. Here, we will just give a small example below.

We found that alternating the binary and Gray representation is more effective than using only one of them. We characterized different types of breeding and mutation according to the resulting exploration area of the parameter space. The result of three such combinations is summarized in the Table 3.2. Based on this result we decided to start the evolution with exploration of the large part of the parameter space (BCO1R-MNR8), then continued with BCO1R-MA, and finally, as our algorithm converged to the solution we explored small area around the best point intensively (refinement with BR-MNA8).

Name	Combination		Width of Exploration	
	Breeding	Mutation	large area	local area
BCO1R-MNR8	Cross-over one random point	N bits randomly	Greatly	No
BCO1R-MA	Cross-over one random point	Each bits independently	Yes	Yes
BR-MNA8	Random	N adjacent bits	No	Greatly

Table 3.2: Impact of different types of breeding and mutation on exploration of parameter space.

3.4 Acceleration of Genetic Algorithm

We have introduced above three fundamental concepts used in any GA (selection, breeding and mutation), which might be sufficient for a simple search. However, in our case (multi-dimensional parameter space with many local maxima) it might

require a large number of iterations with the possibility that the end results might correspond to a local maximum. To reduce the required number of iterations and to increase the stability and efficiency of the algorithm, we introduce several accelerators which are used in our search.

3.4.1 Standard accelerators

In this part, we describe the standard accelerators used in GA.

3.4.1.1 Elitism

Selection and breeding do not guarantee that the next generation will be better than previous one. If we completely replace the old generation with the new one, it is possible that we might lose the organisms with best qualities. The overall tendency (trend of the evolution) is to increase the quality, but it can go down over some short period of time.

The elitism (or cloning) is a simple way to maintain the quality across generations. If the best quality of the new generation is lower than the best quality of the current one, the best organism is propagated to the new generation. It is possible to clone one or several best organisms into the new generation. We take the best organism to be the only one which we clone in the next generation during the evolution. In the next generation the cloned organism breeds with others, improving overall qualities. The elitism stabilizes the GA and guarantees the convergence of the algorithm.

3.4.1.2 Simulated annealing

The simulated annealing method has been already employed in LISA data analysis Cornish & Porter [2007] and proven to be very useful. In this method the smoothness of the quality surface is controlled by the introduced temperature parameter. If the temperature is high, the quality surface is very smooth and nearly all the organisms (good and bad) can be selected for breeding with a similar probability. If the temperature is low, the quality surface is highly peaked around the maxima and only the best organisms can be selected. Usually, a high temperature is selected at the beginning of the search to have a large area of

exploration and to avoid non-significant maxima. The temperature is decreased as we progress the search converging to a maximum of high likelihood.

Temperature is introduced in the selection process through the probability of selecting an organism. We set this probability according to $p_i = q_i / \sum_j^N q_j$ where we have the quality of each organism redefined by the introduction the temperature parameter T as follows:

$$q_i = \exp \frac{(Q_i - Q_{\text{best}})}{T}, \tag{3.30}$$

where T is the temperature, Q_i is the quality of i-th organism and Q_{best} is the quality of the best organism. One can see that all q_i are similar if temperature is high.

When $T = 1$, equation 3.30 reduce to $q_i = \frac{\exp(Q_i)}{\exp(Q_{\text{best}})}$ which is the likelihood ratio between the current organism and the best organism. The introduced temperature is exactly the same as the temperature in the standard simulated annealing, and it plays here a similar role. If the temperature is high, both "good" and "bad" organisms have similar q_i which is close to 1 (analogue of smoothening of the likelihood surface in the simulated annealing), and all organisms can be selected for breeding with the similar probabilities. If the temperature is low, the probability of choosing the "good" organism is much higher than the probability of selecting the "bad" one.

We devised several kinds of annealing. A standard type is the cooling: the temperature evolves from the initial temperature T_i to the final temperature T_f as follows:

$$T(g) = \begin{cases} T_i(\frac{T_f}{T_i})^{\frac{n}{n_c}}, & n < n_c \\ T_f, & n \geq n_c \end{cases} \tag{3.31}$$

where n is a generation number and n_c is the duration of the cooling (in number of generations). The values of T_i and T_f are not known a priori. An alternative approach to control the temperature evolution is to relate it with the quality of the current generation. The temperature is then evolved according to

$$T = \left(\frac{\rho}{\rho_{th}}\right)^g \quad \text{with } \rho_{th} = \rho \quad \text{if } \rho < \rho_{th}, \tag{3.32}$$

where $\rho = \sqrt{2Q_{best}}$ (which is approximately the SNR of the best organism if we use log likelihood as a quality Q) and g and ρ_{th} are two additional parameters. Usually we used $g = 2$ which corresponds to the thermostated annealing introduced in Cornish & Porter [2007]. In the beginning we kept the temperature equal to unity, and a high PMR was used to explore the parameter space and build up the SNR. On reaching ρ_{th}, heating is switched on to increase the exploration area by smoothing the likelihood surface and to force the colony to search for a higher maximum. Periods of high temperature are alternated with periods of low temperature (in a periodic manner), this way the regions around the local maximum and the global parameter space are explored in turn.

3.4.1.3 Evolution of PMR

As mentioned above, another way to control the volume of exploration is by varying the PMR (see Section 3.3.5).

Usually we start with a large value for the PMR (about 0.2), which is then gradually decreased to give more importance to the breeding. In the end, the search becomes stationary and approaches the true solution, so the PMR needs to be quite low (usually we decrease it down to 0.01). The typical spread of the organisms in the beginning of the search is depicted in the left panel of the Figure 3.4 and we slowly evolve it towards the right panel by decreasing PMR.

The three most frequently used types of PMR evolution in our analysis are (i) cooling, (ii) fixing and cooling and (iii) genetic Genetic Algorithm with PMR.

In the first case of cooling, the PMR evolves from the initial value PMR_i at generation $n = 0$ to the final value PMR_f at generation $n = n_c$ according to

$$PMR = \begin{cases} \mathrm{PMR}_f \left(\frac{\mathrm{PMR}_i}{\mathrm{PMR}_f}\right)^{\frac{n_c - n}{n_c}}, & n \leq n_c \\ \mathrm{PMR}_f, & \text{otherwise} \end{cases} \tag{3.33}$$

In the second case of fixing and cooling, at the beginning, the PMR is fixed as $\mathrm{PMR} = \mathrm{PMR}_i$ for n_i generations, then it is cooled to PMR_f in the next n_c

generations as follows:

$$\mathrm{PMR} = \begin{cases} \mathrm{PMR}_i, & \text{if } n < n_i \\ \mathrm{PMR}_f \left(\frac{\mathrm{PMR}_i}{\mathrm{PMR}_f}\right)^{\frac{n_c+n_i-n}{n_c}}, & \text{if } n \geq n_i. \end{cases} \tag{3.34}$$

In the last case of genetic Genetic Algorithm, the PMR is treated as an additional parameter of each organism. The PMR parameter evolves (we search for an optimal value) by the genetic operations in the specified range $[\mathrm{PMR}_{\mathrm{min}}, \mathrm{PMR}_{\mathrm{max}}]$.

We used all the above types of PMR, for each gene we specified its own evolution path. Some parameters converge to the true solution faster than other, and some spin related parameters have multiple solutions. We used the PMR evolution scheme which reflects the convergence of the parameter and uniqueness of the solution.

Note that we controled the exploration area by using both simulated annealing and PMR. Each of these performs somewhat differently. Simulated annealing acts on the quality of the organism and affects the selection procedure for breeding, thus it uses the combination of the initial genes without adding new. On the other hand the PMR changes the structure of each gene and therefore brings in "new blood" into the generation (creates new combinations). The best result is usually achieved by combining together PMR with simulated annealing.

3.4.2 Accelerators specific for SMBH search

In this part, we describe the non-standard acceleration processes introduced by us and which utilize the properties of the signal and/or of the antenna beam pattern.

3.4.2.1 Brother

As explained in section 2.5, the source sky position is encoded in our model of the signal in the phase and in the amplitude through the antenna pattern function. For low frequencies the Doppler term is weak and majority of the information is stored in the directional sensitivity of the detector. However the antenna pattern functions given in expressions equation (3.17) and equation (3.18) are symmetric

with respect to the transformation $\beta \to -\beta, \quad \lambda \to \lambda + \pi$ (mirrored/antipodal sky position). This implies a possible degeneracy in the parameter space, and, indeed we observed a high value of the quality at the antipodal sky position, making it very difficult to distinguish between those two.

In order to overcome this problem, we introduced what we call the brother of the clone. With each clone we associated one organism "brother" created by copying the parameters values from the clone and then changing a few of these value by following particular rules. In our application of the GA for black hole binaries, the brother explores the parameter space around the mirrored (antipodal) sky position of each clone. In a particular search, the best organism usually jumps between these two sky positions until it settles on the best solution in terms of the quality. We have computed the "brother" only for the best organism in each generation. If the "brother" have better quality, the original best organism is not cloned anymore, and the "brother" take its place.

3.4.2.2 Local mutation

What benefit one can have from using binary and Gray representations of the same parameter? The reason lies in representation of two adjacent integers in the binary representation. Two close decimal values of θ_k which differ only by Δ_x (i.e. corresponding integers differ by 1), may differ by several bits in their binary code. For example, in the standard binary representation, the separation between the gene value 011111 and 100000 is equal to the resolution Δx (i.e. minimal distance), but, as one can see, it is necessary to flip all the bits for making this small change in the parameter value. This problem can be solved in two ways: (i) by alternating the Gray representation (where two adjacent integers differ by one bit) with binary, and (i) by introducing the "local mutation". Local mutation is a small (of order few Δx) random change in the parameter value which can push it across the boundary. Note that the binary and Gray codes have different bit boundaries, so the alternation between them helps in the global exploration of the parameter space, whereas the local mutation helps the organisms to cross a particular bit boundary and works locally.

3.4.2.3 Fixing the significant bits

During the test runs of the GA, we noticed that some parameters are very well estimated already after few hundred generations. For example, the time of coalescence t_c can be found with high precision in less than 200 generations. By restricting the search range of these well estimated parameters, the search efficiency can be improved. We achieved this by fixing (freezing) the most significant bits of such parameters which reduces the allowed dynamical range. This significantly speeds up the search. Note that we might still keep the PMR for this gene high in order to have an efficient exploration of the restricted parameter space.

Let us give an example how it works in practice. The Figure 3.5 shows a typical example of the chirp mass, M_c, exploration in our search. This parameter is encoded using 20 bits. First 200 generations have no restriction and a large PMR is used so the colony explores the whole of the prior range. However the higher concentration of the organisms around the best one (depicted by a green line) can be noticed which reflects its high quality and, therefore proximity to the true solution.

After the 200^{th} generation we fixed the bits at a position higher than a randomly chosen number between 14 and 16. It means that the bits $b_{Mc}[16]$, $b_{Mc}[17]$, $b_{Mc}[18]$, $b_{Mc}[19]$ (and sometimes $b_{Mc}[14]$, $b_{Mc}[15]$) of all the organisms are fixed to the value of the best organism (1,1,1 and 0 here). It shrinks the search area to $[2138483.938, 2384509.746]M_\odot$ which corresponds to

$$\begin{aligned} \text{lower boundary} &= M_{c,min} + \Delta M_c \left(0 \times 2^{19} + 1 \times 2^{18} + 1 \times 2^{17} + 1 \times 2^{16}\right), \\ \text{upper boundary} &= M_{c,min} + \Delta M_c \left(0 \times 2^{19} + 1 \times 2^{18} + 1 \times 2^{17} + 1 \times 2^{16} + \left(2^{16} - 1\right)\right). \end{aligned}$$

After the 600^{th} generation we try to restrict the range further by fixing all the bits starting at the position 8^{th} or 9^{th} (again randomly chosen), which corresponds to narrowing down the range $\Delta M_c \times 2^9 = 1922.106 \sim 2000 M_\odot$.

Note that, we can still release the bits (or change the random range) during the evolution to check the robustness of the found solution.

Each parameter have different rate of convergence to the true value. Therefore, we restrict the exploration range of the these parameters at different instance during the search by fixing the significant bits. The chirp mass, symmetric mass

ratio, and time of coalescence can approach the true value much faster than the other parameters. After several hundreds of generations, the search converges to the true values, at this point we restrict the range of these parameters. The number of the significant bits we have to fix is determined by the spread of these parameters from different organisms. As for the other parameters, they may spend much longer time to approach the true value, and we start fixing their bits only after several thousands generations.

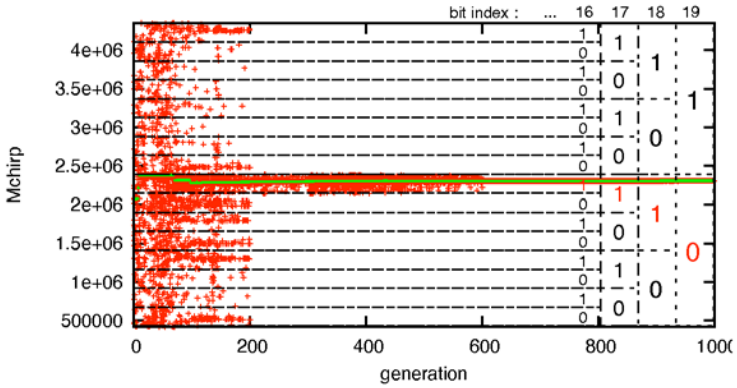

Figure 3.5: Example of the chirp mass exploration by the colony of organisms. The green points correspond to the position of the best organism. The separations shows the structure of the binary representation. The numbers on the right are values at bit positions listed on top.

3.4.2.4 Specific breeding and mutation

As mentioned above in Sections 3.3.4 and 3.3.5, different types of breeding and mutation have different properties (main difference is in the exploration area around the best organism). The genes (i.e. parameters) do not have the same rate of evolution during the search. For example, the time of coalescence and the chirp mass converge to their true values quicker than other parameters. We customized the evolution of each gene by fixing the significant bits in a similar manner to the example discussed in the previous section. We also altered the type of breeding and mutation of each gene, forcing the exploration range to be large at the beginning of the search and changing to the types which are more suitable for more intensive local exploration close to the end.

3.4.2.5 Change of environment

While mapping the log-likelihood we have noticed that the binaries that coalesce within the observational time have more local maxima than the binaries coalescing outside the observational time (this also was mentioned in Babak [2008] for non-spinning BHs). This can be explained by the accumulation of SNR. Due to the shape of the LISA's sensitivity and the evolution of the signal's amplitude, the largest part of SNR comes from the last month of inspiral. In the Figure 3.6 we give accumulation of SNR (scaled by the total SNR) for one of the signals analyzed in MLDC3.2. We plot the SNR as a function of frequency, we also show (numbers

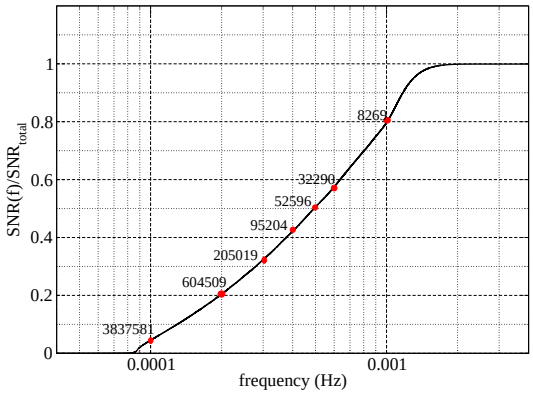

Figure 3.6: Example of accumulation of SNR (MBH-1) as a function of frequency. The points with attached numbers show the time until the coalescence in seconds.

attached to the circles) the time left to coalescence. As one can see 60% of SNR comes from the last day and a half of inspiral. The above implies that we need to fit only the last day of the signal in order to get a large SNR (in case where we see the coalescence). This is obviously can be done in many ways and this results in multiple maxima of the likelihood. If the coalescence is not observed then we need to fit a large number of cycles to accumulate the appreciable SNR, this is harder to achieve unless we are close to the correct solution. In this sense it is easier to find the weaker signal with the time of coalescence outside the observational time. We have utilized this fact in introducing the A-statistic which enhances the low-frequency part of the signal.

We have also implemented the accelerator which we call "change of envi-

ronment". We put the colony in different environments and expect the fitter organisms to survive in a variety of circumstances. In practice we terminate the template earlier in frequency and evolve the colony for some time with chopped templates. By changing the frequency range we change the likelihood surface, the secondary maxima change the size and position, but the global maximum remains at the same position (location of the true parameters, as shown in the Figure 3.1). We use this property to alternate between different environments. It helps to move the search away from the local maxima where it has a tendency to get stuck, and guides the best organism to the true solution. It forces the search to seek a better choice of parameters and can also be used to check for the convergence of the algorithm to the global maximum.

A typical scheme used in our search is as follows: we start off with a full template and use the Maximized Likelihood, $\mathcal{F}_{full}$, as the quality, then we alternate the evolution between full and chopped templates ("change of environment") still using $Q = \mathcal{F}$. We finish the evolution of colony using A-statistic.

We should mention that the frequency annealing introduced in Cornish & Porter [2007] helps not only to speed up the search, but also assists in moving between local maxima. The structure of the likelihood surface changes as the duration of the signal increases.

3.5 Multimodal search

In this Section we explain how we modify the GA to explore multi-modality of the likelihood surface. As discussed in the previous sections, the quality surface have many local maxima. Several techniques (simulated annealing, PMR evolution, change environment, etc.) introduced above, help in finding the global maximum, but they all assume a single solution and, therefore, cannot help if there are several maxima of almost equal heights/amplitudes.

Five spin-unrelated parameters (time of coalescence, chirp mass, mass ratio and sky position) can be estimated using the GA implementation described in previous sections with very high accuracy. The magnitudes of spins can be also determined in some cases quite well. However other parameters corresponding to the initial orientation of spins and of the orbital angular momentum are quite

problematic. A typical situation is presented in the Figure 3.7. We color-code the quality corresponding to each initial orientation of vectors, it varies within 12% of maximum while the points are scattered over the whole range. One can see several solutions which are very close in quality to the true one (depicted by a circle). The search for a single maximum will miss other peaks. Instead we want to explore all of them and, based on the likelihood of each peak, we can make a claim about possible multiple solutions.

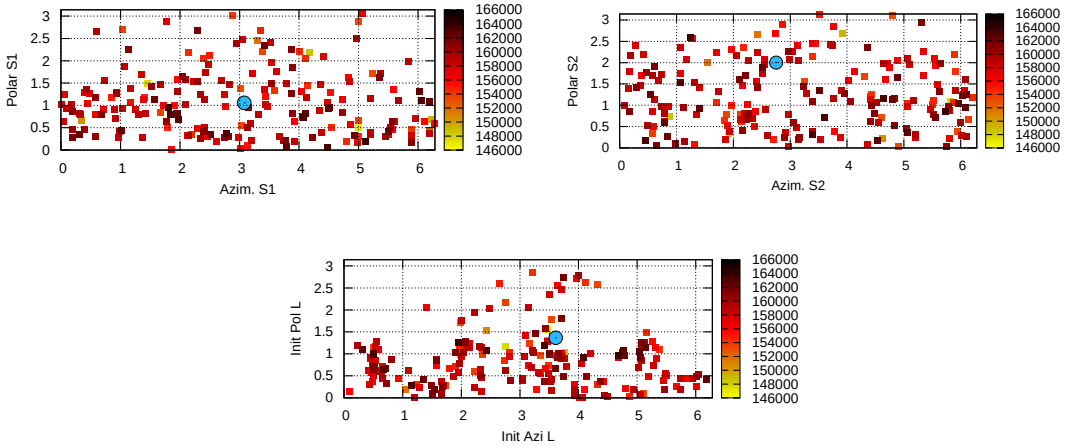

Figure 3.7: Example of the distribution of the best organisms from 196 runs of GA applied to the search for source MBH-3 of MLDC 3.2 (the third signal). The left upper panel shows the initial direction of spin 1, the initial direction of spin 2 is in the right upper panel and we plot the initial direction of the orbital angular momentum in the bottom panel. The color scale corresponds to the value of A-statisitc.

The reason for such a degeneracy lies in the nature of the waveform itself. First of all these parameters are highly correlated, second, they enter the expression for the GW phase at higher post-Newtonian orders, and affect the phase and amplitude rather weakly. The later can also explain that we can determine the spins better if we observe the end of inspiral, where the contribution from the high order terms is appreciable.

Another, and most natural, reason for multi-modality of the likelihood is the presence of multiple signals in the data. In the analyzed data set there were between 4 and 6 signals, but exact number was not disclosed. The signals usually have different SNR, the search converges to the signal with the largest SNR

and explores the modes of this signal, other signals appear at the initial stage of the search (up to the point at which accumulated SNR of different signals is comparable). The main hint that we are looking at the multiple signals is different values for t_c and M_c: parameters which are determined most accurately. The strongest signal can be removed from the data to recover weaker ones. It is desirable at the end to refine all the parameters by using a super-template formed by combining of several signals.

We want to define the mode associated with each local maximum and explore the parameter space in its vicinity.

The basic idea of our Multimodal Genetic Algorithm (MGA) is to put a cluster of organism in each mode, to do so we use several clones.

Each clone corresponds to a mode and all modes should have comparably high qualities. We also increase the size of the colony so that we keep the number of organisms per clone constant. The clones participate in the breeding often and attract other organisms and consequently exploring the neighborhood of each mode. We describe the implementation of the evolution later.

The crucial point of the MGA is the choice of the clones, two conditions are necessary for an organism to become a clone. First, it should have a quality higher than a certain level. This level can be fixed arbitrary or defined relative to the best organism (for example $Q_{clone} \geq Q^{level} = 0.8 \times Q_{best}$). The second condition is that there should not be another clone on the same mode. For that, we define boundaries around a mode (i.e. the rule to separate the different modes) using the variances $\sigma_k^2\left(\widehat{\theta}_{clone}\right)$ of each parameter θ_k at the clone position $\widehat{\theta}_{clone}$. These variances correspond to the diagonal terms of the inverse Fisher Information Matrix defined as equation (1.36) in section 1.4.3.

For each generation we choose all organisms with $Q_i \geq Q^{level}$ as candidates to be cloned. Among these we select only the ones which form the independent modes:

$$\left|\theta_{k,i} - \theta_{k,clone_j}\right| > F_{\theta_k}\sqrt{\sigma_k^2(\widehat{\theta}_{clone_j})}, \tag{3.35}$$

where index i refers to the candidates to be cloned, j to the selected clone and

F_{θ_k} is a factor to control how large should be the distance between the two modes along the parameter θ_k. We choose F_{θ_k} individually for each parameter and it varies between 15 and 50. This way we define the volume of each mode.

There are two ways to evolve a colony with multiple clones. The first one is mentioned above, where we increase the number of organisms proportional to the number of selected clones (modes) and evolve the system using the GA described in previous sections. If qualities of modes are comparable we expect to have a fraction of organisms in the close vicinity of each clone, while the remaining organisms explore the space in between the modes. Once we have started the evolution, we keep the number of clones fixed. If another independent mode is found and its quality is higher than the lowest quality among the clones, than the weakest clone (lowest quality) is moved to the new found location. Note that we always attach a brother to each clone.

The MGA described above requires a large number of organisms (we need to use at least 10 organisms per clone). This requires a specific implementation if we want to use a computer cluster. We use this algorithm but with a small (less than 10) number of clones.

The second approach, which we used the most, disallows continuous communication between the modes. We perform several independents runs (evolutions) with a single clone. Then we analyze the end results of all runs and identify independent modes among them. We use these modes as clones for the next set of independent runs (evolutions). We iterate this procedure until no new modes are found. In this approach the modes exchange information discretely, after each single run. This multi-run MGA is described in detail below in section 3.6.2.

3.6 Pipeline

In this section we describe the chain of algorithms used to arrive at the final result presented in the following section.

3.6.1 Pre-analysis by a time frequency method

The GW signals from the MBH binaries are usually very strong and do not need very sophisticated methods to detect them, especially if we observe the end part of the binary evolution. However, it is more complicated if we observe only early part of the inspiral. Before analyzing the data with GA we looked at the time-frequency map of the data constructed using the Morlet wavelet transform (see Figure 3.8).

From this map, we can clearly identify three strong signals with the time of coalescence within the observational time and one weak signal with time of coalescence about 3 months after the end of observations (signals are pointed by arrows).

As we have mentioned earlier and will discuss in detail later, there is one more weak signal which coalesces even later. This signal is low frequency and too weak to be seen by eye in the data.

In producing submitted MLDC results, we split the data in three parts, based on the time-frequency analysis discussed in the previous paragraph. The first part contains the strongest signal completely and low frequency parts (few first months) of all other signals. The second part contains two other coalescing binaries and parts of remaining weak signals. Finally the third part contains only the signals coalescing outside the observational time. An iterative approach could also have been employed where the strongest signal is found and then removed from the data which is then analysed to detect other signals. This process is repeated until no more signals can be found. Estimating the residuals after subtracting the detected signals presents a particular problem in this incremental approach. A disadvantage of our chosen approach is the lose of some SNR, but we can be sure of avoiding the corruption of the weak signals by residuals of strong ones. However, it turns out that in order to find the fifth (the weakest) signal we had to remove the fourth signal (the right most one in Figure 3.8) due to the strong interference with the secondary maxima.

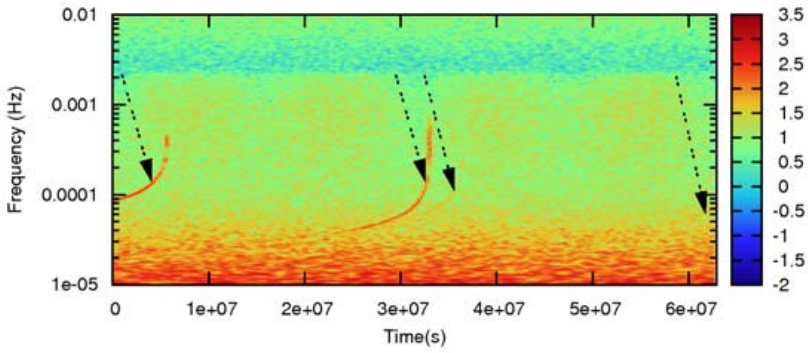

Figure 3.8: Time-frequency representation of channel TDI A of MLDC 3.2. We plot the norm of the Morlet wavelets transformation.

3.6.2 Multiple steps MGA search

We should not forget that GA is a stochastic search method. We can be sure about the convergence to the true solution if the likelihood surface is smooth and uni-modal. Unfortunately it is not the case, we have implemented many tricks to get through the forest of the local maxima to find the highest peaks. As mentioned earlier, our algorithm found several solutions with similar values of the likelihood. The evolution can still end up on one or another of these maxima, depending on the initial state and the seed of the sequence of random numbers. This is the reason behind implementing MGA.

We have briefly introduced the multi-runs MGA in Section 3.5, here we give a bit more detailed description.

In this implementation of MGA, the modes exchange information discretely. We start with N_{run}^{std} runs of a "single clone + brother + 20 organism" evolution. We use all accelerators introduced in Section 3.4. We call each of these runs "standard" (as opposed to the global MGA run/search). In the first step we explore the parameter space trying to find as many maxima as possible. We evolve each colony for 2500 generations, then we collect the results of all these evolutions and identify the modes associated with the best organisms as described in the Section 3.5. In our search we followed 50 best modes. We attach a colony and start another standard run for each mode, in other words we start $N_{run}^{mode} = 50$ inde-

pendent evolutions for a single clone plus colony. In this step each mode is either refined or migrates to a new location outside its boundary, with a higher quality. In addition we restart N_{run}^{std} standard evolutions searching for more modes. At the end of this step, the results from the $N_{run}^{mode} + N_{run}^{std}$ runs are collected and a new set of modes is identified. We iterate the process until the 50 best modes do not change anymore.

We found that number of strong modes depends on the parameters of the signal and therefore keeping the number of modes to be explored fixed is unreasonable. In the post-MLDC exploration we used a variable number of modes: keeping all the modes with the quality $\geqslant 98\%$ of the best one.

Both standard and mode exploration runs have similar evolution and differ mainly by an initial state of the evolution. We always use simulated annealing and the temperature alternates between hot and cold phases, the threshold ρ_{th} (see part 3.4.1.2), which regulates the temperature, decreases with the number of generation for both phases. The number of organisms in each run is kept fixed at 20.

For the standard run, the evolution of the PMR and of the type of breeding and mutation are chosen such that the exploration volume is high at the beginning of the search and low towards the end (i.e. the volume near the vicinity of the best organism is search more intensively).

As evolution progresses we gradually fix the range of parameters in the following order: t_c, M_c, η, sky position, amplitude of spins, and finally, the initial direction of the orbital angular momentum and spins.

As mentioned above there is a significant difference in likelihood maps for the signals coalescing within and outside the observational time. This is reflected in the search strategy for those two types: due to fewer number of secondary maxima for signals with coalescence outside observational time, we used only moderate simulated annealing and change of environment and fixed the ranges of M_c, t_c much later in the evolution.

After some iterations the modes reach the stable state: we do not see any new modes and the existing modes are settled at stationary positions (maxima). At this point we stop the run, and all the modes found constitute our solution.

One can use a Bayesian approach to assign a probability to each mode by

calculating the Bayesian evidence.

3.7 Results

We have previously shown Petiteau *et al.* [2009a] that GA works very well in the case of the non-spinning SMBH even without using the multimodal search. In this section we discuss the results of the search for spinning SMBH binaries. We present here the outcome of the analysis of MLDC3.2. By the deadline we did not implemented the MGA in full and therefore we have below two subsections: In the subsection 3.7.1 we give the results submitted by the deadline, and in the subsection 3.7.2 we present the results of the full scale MGA analysis (obtained after the deadline). The main difference is in the number of recovered modes and switching to the full LISA response at the end of the search to reduce the bias due to mismatch between response function used in the signal generation and the one used in our analysis.

3.7.1 MLDC results

The signals present in the data can be split in two types: the binaries with the time of coalescence inside the observational time and others whose coalescence happened outside the observations. The difference between these two types is in the number of local maxima, SNR and consequently in the accuracy of the recovered parameters.

3.7.1.1 Coalescence within the observational time

We have found three signals of this kind in MLDC3.2.

In the MGA we restricted the search to only 50 best modes selected at each step. Among 50 explored modes for each signal, we have identified a small number of distinctly strong and comparable modes for the submission. After 14th, 8th and 7th iterations respectively of the multi-runs MGA search, we obtained five modes for the strongest signal with the shortest t_c (srcMC1 which is SMBH-1 in MLDC notation), four modes for the second one (srcMC2 or SMBH-3 in MLDC) and six modes for the weakest signal (srcMC3 or SMBH-4 in MLDC).

The results are presented in the first half (first three rows) of the Table 3.3 which lists the relative/absolute errors, global overlaps and quality for modes submitted in MLDC 3.2 for each signal (without the direction of the spins and of the orbital angular momentum). These errors should be compared to the corresponding predictions from FIM which are also given in the Table 3.3 in the row labelled as "True". For the chirp mass, the errors for all the modes are similar to the ones estimated from the FIM. For others parameters, the errors are generally few times higher than predicted by the FIM. At least part of this discrepancy comes from the bias caused by the signal approximation – we have used the long wavelength limit which is valid for the low frequency part and breaks down near the coalescence. The mode with the error for the sky position higher than 175 degrees corresponds to the antipodal location on the sky. Taking this as a genuine degeneracy, we see that the source location is found with the precision better than 10 degrees for srcMC1, 5 degrees for srcMC2 and one degree for srcMC3.

We found a strong degeneracy in the initial directions of the orbital angular momentum and spins, so we decided to submit several well separated modes. Only for srcMC2, one of these modes corresponds to the true parameter set. For srcMC1 and srcMC3, the true mode was missed, however we found it in the full scale MGA analysis conducted after the deadline (see subsection 3.7.2).

The last two columns of the Table 3.3 show the value of A-statistic (quality column) for each mode and the multi-stream overlap defined as

$$\mathcal{O}(\widehat{\theta}_e) = \frac{\langle h_A(\widehat{\theta}_e) \mid h_A(\widehat{\theta}_t)\rangle + \langle h_E(\widehat{\theta}_e) \mid h_E(\widehat{\theta}_t)\rangle}{\mathcal{N}[h(\widehat{\theta}_t)]\ \ \mathcal{N}[h(\widehat{\theta}_e)]}, \tag{3.36}$$

where $\widehat{\theta}_t$ corresponds to the true parameters, $\widehat{\theta}_e$ are our estimated parameters and $\mathcal{N}[h(\widehat{\theta})]$ is the norm of the template $h(\widehat{\theta})$ defined as

$$\mathcal{N}[h(\widehat{\theta})] = \sqrt{\langle h_A(\widehat{\theta}) \mid h_A(\widehat{\theta})\rangle + \langle h_E(\widehat{\theta}) \mid h_E(\widehat{\theta})\rangle}. \tag{3.37}$$

This is an extension of the overlap introduced in Chapter 1. The overlap $\mathcal{O}$ varies between -1 and 1 (from perfect anti-correlation to perfect correlation) and it tells us the loss in the SNR $= \mathcal{N}[h(\widehat{\theta}_t)]$. In the Table 3.3, the overlaps are

Source	mode	$\Delta M_c/M_c$ $\times 10^{-5}$	$\Delta\eta/\eta$ $\times 10^{-4}$	Δt_c (sec)	Δ Sky (deg)	Δa_1 $\times 10^{-3}$	Δa_2 $\times 10^{-3}$	$\Delta D/D$ $\times 10^{-3}$	$\mathcal{O}$	quality
	True	1.3	4.4	6.1	1.18	2.7	6.2	6.2	1.0	392171
	1	4.3	39.8	7.2	4.15	8.6	83.0	8.7	0.99189	392628
srcMC1	2	7.8	58.6	631.4	177.54	4.7	64.0	0.6	0.99236	392595
MBH-1	3	2.7	15.1	0.7	5.39	5.2	84.9	3.8	0.99198	392589
	4	0.2	62.5	33.7	1.43	1.9	87.1	14.2	0.99174	392533
	5	2.4	6.1	62.9	11.67	7.6	47.5	80.4	0.99235	392385
	True	4.3	7.2	9.1	0.82	2.9	5.3	7.2	1.0	164559
srcMC2	1	8.9	5.1	100.8	175.94	6.2	18.7	27.2	0.98965	164626
MBH-3	2	6.4	106.6	164.4	178.49	46.2	31.8	9.4	0.99800	164608
	3	1.4	106.6	39.0	3.65	42.4	39.9	27.9	0.99592	164589
	4	0.5	113.9	179.9	176.53	45.1	11.1	20.5	0.99754	164583
	True	10.1	55.3	26.7	0.47	29.2	151.4	138.7	1.0	5823.92
	1	22.0	126.2	362.2	179.82	57.4	93.1	337.9	0.99403	5845.22
srcMC3	2	16.8	153.9	337.7	179.84	51.6	373.4	252.2	0.99752	5832.83
MBH-4	3	1.1	166.8	30.7	0.19	51.9	385.2	252.2	0.99463	5832.07
	4	29.2	303.4	349.5	179.79	28.8	401.7	220.8	0.99686	5832.01
	5	4.5	75.2	31.4	0.12	47.1	173.7	90.8	0.99935	5832.01
	6	29.2	138.5	258.9	179.27	226.4	184.0	125.3	0.99710	5830.61
	True	167.3	702.8	14641.2	10.37	725.6	902.6	167.9	1.0	184.99
srcMC4	1	251.3	778.7	18562.2	45.07	38.3	150.6	93.3	0.936389	197.77
MBH-2	2	270.2	118.3	8405.7	10.47	218.5	251.4	178.4	0.965423	197.31
	3	1114.1	952.2	38160.8	171.10	331.7	409.0	153.0	0.943096	197.08
	4	714.4	1104	7942.4	141.59	11.7	665.2	169.3	0.935997	196.00
srcMC5	True	315.1	670.3	73890.8	6.40	453.7	699.0	321.4	1.0	38.75
MBH-6	1	1042.3	1235.6	82343.2	2.11	258.2	191.6	260.5	0.929130	47.41
	2	293.7	618.8	43456.8	173.94	89.6	122.9	430.5	0.729048	41.78

Table 3.3: Relative/absolute errors, global overlap, $\mathcal{O}$, and quality for the modes submitted in MLDC 3.2. All parameters are defined in section 3.2.1. ΔSky is the angular distance in the sky between the true and the estimated positions. The second column gives the mode number. The errors for true parameters are obtained using the FIM. For the tree first sources (srcMC1, srcMC2 and srcMC3) which coalesce during the observational time the quality corresponds to A-statistic and the two others (srcMC4 and srcMC5) which coalesce after the end of observation the quality is the Maximized Likelihood $\mathcal{F}$.

computed by using the low-frequency approximation.

Due to the noise, we could expect the recovered SNR to deviate from the true SNR within ± 1. The SNRs of the injected signals srcMC1, srcMC2 and srcMC3 are 1670.58, 847.61 and 160.51 respectively. The best recovered mode for the source srcMC1 has the SNR equal to 1657.71, it is far away from the true one. The best modes of srcMC2 and srcMC3 have the SNRs equal to 846.96 and 160.05 respectively, they are clearly close to the true value. However, the FIM errors rows in the table 3.3 indicate that the found solutions are quite distinct from the true values.

All of these modes have an overlap with the true solution higher than 99%. The value of A-statistic as well that of Maximized Likelihood for the recovered modes is higher than the corresponding values for the true parameter set. This is a manifestation of the mismatch between the signal and the template and indicates the importance of using the full response towards the end of the search. We were aware of this but did not have time to implement it completely before the MLDC submission deadline. Nevertheless, given this bias in the search, our results are still quite accurate.

3.7.1.2 Coalescence beyond the observational time

During the search we found two signals of this kind. The results are presented in the second half (last two rows) of the Table 3.3. Those are low frequency signals, so our long wavelength approximation works very well resulting in very small or no bias in the parameter estimation due to the mismatch between the response functions as discussed in the previous section.

First we identified the source with SNR 18.63 (which is srcMC4 or SMBH-2 in MLDC notations). For this signal we found several modes after 8 steps of MGA search, out of which we selected four modes with highest quality for the submission. From the Table 3.3 it can bee seen that the errors in the spin independent parameters are similar to the errors predicted by the FIM. Spin-orbital and spin-spin couplings enter the phase at 1.5 and 2 PN orders respectively, and since we do not observe the end of the inspiral, these terms contribute very little to the phase as well as to the amplitude modulation (see the orbital frequency

dependent term in equations (3.1-3.3)). Therefore the spin related parameters are intrinsically poorly identified for these sources which is reflected in our results.

The fifth and the last source is the weakest. In fact it was completely contaminated by the secondaries of the srcMC4. In order to identify this source we had to remove the fourth signal. We identified the srcMC4 with the best (highest quality) recovered mode, generated the signal and subtracted it from the time series. After that we repeated the search and already the first standard run found the mode with SNR > 7 which was a positive detection. Before the deadline we could perform only 3 steps of the MGA, however this turned out to be sufficient, as is indicated by the overlap column in the Table 3.3. We have clearly identified two modes with the opposite sky positions. The SNR for this signal was 12.82 and consequently the parameters have large uncertainties. The initial directions of the spins and the orbital angular momentum could not be identified at all. Other uncertainties are consistent with the FIM. This, fifth signal, was correctly identified only by us among all the participants of MLDC3.2 (at least with the precision which gave an overlap of 0.92).

The SNRs of the injected signals srcMC4 and srcMC5 are 18.95 and 12.82 respectively. The SNRs of the best modes of srcMC4 and srcMC5 are equal to 20.54 and 13.69 respectively. The recovered parameters are within few FIM sigmas away from the true.

3.7.2 Post-MLDC results

After the MLDC submission deadline, we finalized the implementation of the MGA (this time we have kept all the modes within some fraction of the maximum) and have performed the search to completion. We have also incorporated the full LISA response in our template using LISACode LISACode [2000]; Petiteau *et al.* [2008] to refine the final solutions. We discuss the details below.

For the few first steps we kept all the modes with quality higher than 50% of the best one, then we increased the mode selection threshold to 90% (or higher, depending on the number of modes detected for a given signal). We also improved the mode separation criteria by adjusting F_{θ_k} based on the detailed study of the quality distribution. Finally, we have also added two final search steps using the

templates with the full TDI response.

Here we used a lite version of LISACode simulatorLISACode [2000]; Petiteau *et al.* [2008]. The lite version contains some fine-tuned trick which allowed us to compute the two-years long template in less than 15 seconds.

The final steps with the full TDI response are required only for the signals which coalesce within the observational time. Only those signals propagate to high frequency where the long wavelength approximation is not accurate any more, and the SNR is high enough for this to matter. Including the full response also helped for srcMC1 to promote the mode (increase its quality) closest to the true solution and slightly suppress the others.

For the last, full response search, we selected the modes within 98% of the best one. This results in the selection of 26 modes for source srcMC1 after 13 steps of MGA, and 175 and 17 modes for the sources srcMC2 and srcMC3 respectively after 9 steps of MGA. Figure 3.9 shows the distribution of the Maximized Likelihood with the full LISA response relative to the true one, $\mathcal{F}_{Full,i}/\mathcal{F}_{Full,true}$, for the modes of the source srcMC2. For this source, 36 modes have SNR within one sigma deviation form the SNR_{true}: $|\mathrm{SNR}_i - \mathrm{SNR}_{true}| < 1$, and 4 of them have $\mathcal{F}_{Full,i}$ higher than for the true waveform. Deviation in SNR by unity can be easily produced by noise, note that besides the stationary gaussian instrumental noise we also had cyclo-stationary Galactic confusion noise. There are similar results for other sources. For srcMC1 we have identified 21 modes with $\Delta SNR_i < 1$ and 6 of them have $\mathcal{F}_{Full,i} > \mathcal{F}_{Full,True}$; for srcMC3, 21 modes have $\Delta SNR_i < 1$ and 6 of them with $\mathcal{F}_{Full,i} > \mathcal{F}_{Full,True}$. We have confirmed those results also with the signals generated using syntheticLISA Vallisneri [2005], another simulator used to produce the data set. This is to avoid possible error coming from the use of two different simulators.

The difference between identified modes is within the fluctuation that can be caused by the noise, therefore there can not be a unique solution. However, all the modes are single valued in the non-spinning parameters and they split for the initial directions of the spins and the orbital angular momentum (sometimes also antipodal sky location).

Our post-MLDC results are presented in Table 3.4 in which we list the parameter estimates for three modes found for each source. These modes are described

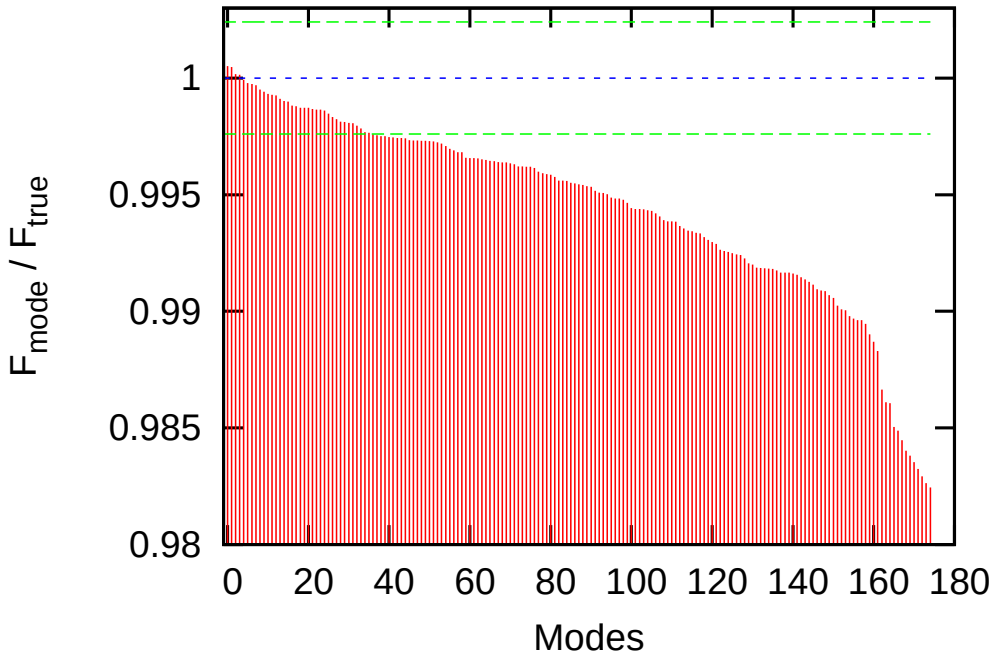

Figure 3.9: Ratio between Maximized likelihood of modes and true Maximized Likelihood for source srcMC2. The blue dotted line corresponds to a ratio equal to one. The green dashed lines correspond to $\Delta SNR_i < 1$

as follows:

1. 'B'-mode ('B'-est) is the the mode with the highest Maximized Likelihood value using the template with the full response.

2. 'C'-mode ('C'-losest) is the mode closest to the injected signal in all parameters.

3. 'A'-mode ('A'-strophysically relevant) is the mode with the smallest error in the most relevant parameters from the astrophysical point of view (sky position, distance, masses, spin amplitude and time of coalescence).

We estimate the "closeness" to the true parameters as

$$d \equiv \max_k \left\{ \frac{\sigma^{\theta_k}}{\sigma^{\theta_k}_{FIM}} \right\},$$

where σ^{θ_k} is an error in the estimation of parameter θ_k, which is determined by computing the absolute difference between the value of the modes and the true value of the signal: $\sigma^{\theta_k} = |\theta^k_{\mathbf{Mode}} - \theta^k_{\mathbf{true}}|$. In the table 3.4, the values which are corresponding to rows of 'A,B,C' are the σ^{θ_k} for each parameter. $\sigma^{\theta_k}_{FIM}$ is a corresponding prediction from the FIM. Both of σ^{θ_k} and $\sigma^{\theta_k}_{FIM}$ have the same unit as the parameter θ^k, hence, the distance d is dimensionless, and gives measure of a distance in unites of FIM.

We have eliminated the systematic error by using the full LISA response function. All the modes presented in the table 3.4 have likelihood higher than the true signal, but the difference in the SNR is still less than 1. The FIM errors in the table tell us how much deviation from the true parameters we expect due to the influence of noise. The differences in some parameters have the values comparable to the FIM errors, other parameters are quite off. This indicates the strong degeneracy in these parameters, and the found solutions do not correspond to the true one.

As for the computational expense, it mainly depends on the duration time of the signal. When we searched the strongest signal "srcMC1", the waveform duration is several month. A single PC can run 2000 generations for one individual genetic algorithm in less than 8 hours. When it come to the weaker signal, or the longer waveform, the cost of the computation increase correspondingly. In the full multi step MGA search, the number of steps is also related to the signal itself. Taking "srcMC1" as an example, we performed 8 steps by using 10 CPU and the full run takes 5 days. For the weakest signal, we need around 20 steps and more than 10 days.

3.8 Summary

To conclude this chapter, we have described the application of the Genetic Algorithm to the the problem of detecting gravitational wave signals from inspiralling spinning SMBH binaries and estimating their parameters. We described how GA can be translated to the problem of GW data analysis, and introduced some custom-designed accelerators of the evolution which allow us to efficiently explore the 13-dimentional parameter space. In addition to the standard $\mathcal{F}$-statistic which

Source	mode	$\Delta\frac{M_c}{M_c}$ $\times10^{-5}$	$\frac{\Delta\eta}{\eta}$ $\times10^{-4}$	Δt_c (sec)	Δ Sky (deg)	Δa_1 $\times10^{-3}$	Δa_2 $\times10^{-3}$	$\Delta D/D$ $\times10^{-3}$	$\Delta\widehat{S}_1$ (deg)	$\Delta\widehat{S}_2$ (deg)	$\Delta\widehat{L}$ (deg)	$\Delta\phi_c$ $\times10^{-2}$	$\mathcal{O}$	$\mathcal{F}_{Full}$
	True	1.3	4.4	6.1	1.18	2.7	6.2	6.2	8.30	6.06	1.07	7.5	1.0	1387732
srcMC1	A	10.1	9.2	25.7	1.92	0.7	44.3	13.1	64.39	79.59	84.02	259.6	0.999870	1387772
SMBH-1	B, C	13.5	8.6	24.6	8.04	6.7	28.9	22.4	9.39	15.70	7.70	14.7	0.999944	1387946
		10.8	6.3	24.2	7.04	7.1	21.8	20.8	48.18	19.07	6.20	40.6	0.999952	1387914
	True	4.3	7.2	9.1	0.82	2.9	5.3	7.2	1.52	3.29	0.95	2.9	1.0	355588
srcMC2	A	13.6	4.8	29.0	1.33	2.8	28.3	19.8	104.65	155.92	138.55	4.1	0.999939	355755
SMBH-3	B	11.0	92.5	154.6	176.51	24.1	3.5	24.7	55.37	54.86	81.31	135.7	0.999827	355769
	C	15.6	44.6	158.9	169.3	52.4	15.0	66.1	16.49	64.82	14.68	25.0	0.997845	354301
	True	10.1	55.3	26.7	0.47	29.2	151.4	138.7	22.90	65.30	16.17	102.2	1.0	12814.2
srcMC3	A, B, C	9.2	139.0	40.5	0.34	55.9	390.1	181.3	159.55	74.93	63.38	7.3	0.999311	12834.4
SMBH-4		17.7	8.5	234.0	179.48	96.5	506.7	319.6	60.55	87.22	42.61	413.2	0.998723	12818.8

Table 3.4: Post-MLDC results. We present relative/absolute errors, global overlap, $\mathcal{O}$, and Maximized Likelihood using full response $\mathcal{F}_{Full}$ for selected modes for sources srcMC1, srcMC2 and srcMC3. All parameters are defined in section 3.2.1. ΔSky, $\Delta\widehat{S}_1$, $\Delta\widehat{S}_2$ and $\Delta\widehat{L}$ are the angular (geodesic on a sphere) distance between the true and the estimated direction for, respectively the sky position, the spin of MBH_1, the spin of MBH_2 and the orbital angular momentum. The selected modes are given in several rows for each source and marked as: 'A' corresponding to the mode closest to the true one in the most relevant astrophysical parameters (sky position, distance, masses, spin amplitude and time at coalescence); 'C' corresponding to the mode closest to the true one in all parameters; 'B' is the mode with the best $\mathcal{F}_{Full}$. The mode without any mark is the second closest.

is popular in the GW data analysis, we introduced a new detection statistic called A-statistic, which enhances the low frequency part of the signal. Use of A-statistic allows us to partially compensate for the mismatch between the template and the true signal and to change the structure of the quality surface eliminating some of the secondary maxima.

We have found that the likelihood surface is highly multimodal with several modes having very high amplitudes. In order to incorporate this in our search we have extended the standard GA to the multimodal Genetic Algorithm. We cluster strong local likelihood maxima in the parameter space within the volume defined by the slightly enlarged error boxes predicted by inverse of the Fisher information matrix.

To each cluster or "mode" identified in such a manner, we attach a colony of the organisms. The colonies explore their local regions intensively and exchange the information after every 2500 generations.

We apply this method for the analysis of MLDC3.2 data set. In the blind search, we have successfully found all 5 signals and the recovered solutions have overlap higher than 99.2% for the strong (high SNR) signals and higher than 93% for the weak signals. The results submitted by the deadline did not fully reflect the capability of our search method, as the implementation was not complete. We have completed the search after the deadline by allowing MGA to reach the stable solution. We have also used the full TDI response during the last two steps of our post-MLDC analysis. This has allowed us to recover all modes and reduce the bias in the parameter estimation due to use of the long wavelength approximation in our search template. We have achieved a remarkable accuracy in estimating non-spinning parameters, as well as reasonably accurate estimation of the spin magnitudes if binary coalesces within the observational window. Our method is at least comparable, if not better, to other very successful algorithms such as MCMC with parallel tempering and MultiNest Feroz *et al.* [2009]. The success of the MGA in case of the inspiralling spinning SMBH binaries gives us the confidence that it should also prove to be highly efficient in many other data analysis problem. In the next chapter, we use GA (simplified version) for analysing GW signals from EMRIs.

Chapter 4

EMRI Data Analysis With A Phenomenological Waveform

This chapter is based on the my published paper Wang *et al.* [2012].

In gravitational wave astronomy, EMRIs are among the most interesting sources. Due to the large amount of local maxima on the likelihood surface, it is very challenging to detect those sources with the accurate recovered parameters of binaries. Because the matched filtering requires computation of the gravitational waveform hundreds thousand of times, the most accurate (faithful) model of EMRIs is not feasible. In this chapter, we introduce a phenomenological template family which covers a large range of EMRIs parameter space, and used them to detect the signal in the simulated data. Then, assuming a particular EMRI model, we estimate the physical parameters of the binaries.

4.1 My Contribution

In the section 4.3, I have used the results from Babak *et al.* [2007] to code up the numerical kludge model both in time and frequency domain using matlab. Following the methods which are introduced by Drasco [2009]; Drasco & Hughes [2004, 2006], I have done the harmonic analysis of the numerical kludge waveform, and the results (which are the basis of our phenomenological template family) are shown in the section 4.4. I have participated in the discussion with my colleagues

on the MCMC results which are obtained by Yan Wang in section 4.6.1. Then, I have used the multi-step genetic algorithm to recover the physical parameters of the GW signal, and the results are given in the section 4.6.2.2. My colleague Wang Yan used another algorithm: particle swarm optimization to do the same search in parallel. The results of these two methods are comparable and presented in the table 4.1.

4.2 Motivation

As mentioned in the section 1.3.2.3, due to the N-body interaction, the stellar compact object like a black hole, neutron star or white dwarf located in the cusp surrounding SMBH in the galactic nuclei could be thrown on a very eccentric orbit. The compact body could either plunge (directly or after few orbits) into SMBH or form an EMRI. In the case of EMRI, it will spends $\sim 10^6$ orbits, and the information of this orbital evolution will be encoded in the phase of emitted gravitational waves. LISA will observes those sources few years before the plunge and we can extract extremely accurate parameters of a binary system Barack & Cutler [2004] (like mass and spin of SMBH M, a, mass of a small object m, inclination of the orbital plane (to the spin of SMBH), orbital eccentricity and semi-latus rectum (ι_0, e_0, p_0) at some fiducial moment of time t_0, location of the source on the sky (Θ, Φ) and more) by fitting precisely the GW phase.

We can also test the nature of the central massive object through this precise tracking of GW phase. Generally, it is believed that it should be a SMBH with surrounding spacetime described by a Kerr metric. The orbital evolution of the compact object is affected by the nature of the spacetime which could be extracted from the GW phase. As stated by the "no-hair" theorem, Kerr spacetime is described by only two parameters: black hole's mass M and spin a. The spacetime have the multipole moments structure of the central massive object, and, for Kerr BH, all the moments depend only on M, a: $M_l + iS_l = (ia)^l M$, where where $S_1 = J$ is the spin and $a = J/M$ is usual Kerr spin parameter. By measuring the first three moments (mass, spin and quadrupole moment) Barack & Cutler [2007], we can check the "Kerrness" of a spacetime. In general, the deviations from Kerr could be caused by several reasons: (i) it is Kerr BH but with an additional

perturber (gas disk, another SMBH) (ii) it is not Kerr BH but some other object satisfying GR (boson star, gravastar), (iii) there are deviations from GR. For discussion on the this topic we refer reader to Amaro-Seoane *et al.* [2010, 2012b]; Babak *et al.* [2011] and references therein.

So far, even in GR, the full modeling of orbital evolution has not yet been completed. The large mass ratio allows us to treat the small compact object as a perturbation of the Kerr background spacetime, and the problem can be described perturbatively in orders of the mass ratio. In the zero order approximation, the compact object moves in a pure geodesic orbit. However, once we take into account its mass, it creates its own gravitational field interacting with the background which results in the so called self force, and system emits gravitational radiation. This leads to the motion of the compact object which could be treated as the forced geodesic motion. Alternative interpretation is that the motion is governed by a geodesic motion but in the perturbed spacetime. Calculation of the self force is a complicated task which is only accomplished for the orbits around Schwarzschild BH Barack & Sago [2010]; Warburton *et al.* [2012], the Kerr spacetime is underway. To compute the motion under the self force one can use the osculating elements approach Gair *et al.* [2011], or self-consistent approach of direct integration of the regularized equations Diener *et al.* [2012]. For more details on this subject we refer to Poisson *et al.* [2011].

All in all, the modeling of the orbital evolution and the GW signal is a complex task which requires significant theoretical and computational developments. The later prevents us currently from using the state-of-art GW models of EMRIs in our data analysis explorations. The most used model is called "analytic kludge" (AK) which is a phenomenological model suggested in Barack & Cutler [2004]. It is based on Post-Newtonian expressions and puts all relevant physics of EMRIs together. However the defects of this model are the restrictions in the number of harmonics and in their strength. The recovered parameters could be very different from the actual ones. Any search algorithm which relies on its specific harmonic content will not work for a more realistic model of GW signal. The main motivation of our work is to create the phenomenological search template family which would fit a very large range of EMRI-like signals. The typical EMRI signal consist of set of harmonics of three (slowly evolving) orbital frequencies,

and we will use it as a basis of our template.

The present work is an extension of the previously published work by Gair *et al.* [2008] and Babak *et al.* [2009]. The authors in Gair *et al.* [2008] produced the spectrogram from the data, and applied the threshold based on the false alarm with the assumption of Gaussian noise to identify the pixels with excess of power. As a result, they obtained the time frequency tracks of the signal. Then they run matched filtering based on MCMC search on the time frequency plane to extract the physical parameters using the analytical kludge model for EMRI signal. The authors in Babak *et al.* [2009] have conducted search on the parameters space to find the local maxima of the likelihood using the analytical kludge model. They computed the accumulative likelihood of each found local maxima to identify the time and frequencies when templates match some harmonics of the signal. We use this technique and described it in details in section 4.6.1. Form the found time frequency tracks, they fixed approximately the three fundamental frequencies at some reference time. Then they fix found frequencies of harmonics and use it as a constrain in the next step of MCMC search (requiring that each template passes at the reference time through the fixed frequencies). Finally they release this constrain, and evolve the frequency to refine the parameter estimation. In this work, we follow Babak *et al.* [2009] to find time frequency tracks by using the phenomenological waveform which is constructed in a model independent way. Then using the numerical kludge model, we search the physical parameter in a similar way to Gair *et al.* [2008]. We choose the Genetic Algorithm and particle swarm optimization to do this search.

The phenomenological template is constructed by N_h harmonics with constant amplitude and slowly evolving phase which is expanded in a Taylor series. Truncation of the Taylor series and the assumption about constant amplitude lead to the restrictions on the duration over which the phenomenological template can fit an EMRI signal. There are several factors that lead to change in the amplitude of EMRI's harmonics: (1) shrinking of the orbit lead to the overall amplitude increases, (2) circularization of the orbit results in the shifting power to the lower harmonics, (3) slight change in the inclination of the orbit to the spin of MBH. Because the phase modulation is more important than the accurate description of the amplitude, we should use more terms in the Taylor series which can track

phase of the EMRI signal as accurate as possible. Finally, we decide on the number harmonics to use in the template (and their indices) depending on the analysis of the harmonic structure of the Numerical Kludge (NK) model Babak *et al.* [2007] of EMRI in different parts of the parameters space. The restriction that the phenomenological waveform (PW) is valid only for a limited period of time is very weak since we can fit the signal piecewise, as long as accumulated signal-to-noise ratio (SNR) over that time is significant to claim presence of the signal. In this work we consider only those parts of the EMRI signal where the orbital frequencies are not decreasing which is true over almost all time of the inspiral and breaks quite close to the plunge. However this is not really necessary since we did not restrict the values of frequency derivatives to positive values during the search.

In fact, the PW family is quite generic and does not depend on particular model for the orbital evolution, or, in other words, the orbital evolution of the binary is encoded in the Taylor coefficients of phase of each harmonic. This allows us to detect EMRI signal in a model independent way. Once the harmonics of the signal are recovered we can analyze them using a specific EMRI model to recover physical parameters of the system.

It is at this point we need the orbital evolution with high accuracy. Besides the orbital precision, the computational cost is also important. In this sense, the Nk model are quite suitable for our purpose. First, it is much faster to be generated compare to the other high precision model, such as the Teukolsky-based numerical waveforms Cutler *et al.* [1994]; Drasco & Hughes [2006]; Glampedakis & Kennefick [2002]; Hughes [2001]; Poisson [2004]. Second the NK waveform has more realistic orbital than the AK model. With these two benefits, we can do the harmonics analysis very effectively. In the next section, we will review the NK waveform construction and our results of harmonics analysis of NK waveform.

4.3 Numerical kludge model

We describe NK waveform closely following the paper Babak *et al.* [2007]. The NK waveform is constructed by using weak field perturbation theory. We keep trajectory as Kerr geodesic plus evolution under the radiation reaction, and use

this trajectory to compute waveform by assuming the flat spacetime background. By accurately modeling the inspiral trajectory we achieve a good phase match between the true waveform and the approximate one. The present trajectory is described in the Boyer-Lindquist coordinates in the Kerr space time, and we need to map these coordinates onto the Cartesian coordinate in the flat space. The most natural way is to treat the (r, θ, ϕ) in B-L coordinate as the spherical polar coordinate in the flat space Babak *et al.* [2007]. The pseudo-Cartesian coordinate can be constructed as

$$x = r \sin\theta \cos\phi, \tag{4.1}$$
$$y = r \sin\theta \sin\phi, \tag{4.2}$$
$$z = r \cos\theta. \tag{4.3}$$

The trajectory is expressed as $Z(t) = \{x(t), y(t), z(t)\}$. For waveform generation, we can consider the weak-field approximation and write down the space time metric as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{4.4}$$

where $\eta_{\mu\nu}$ is the flat metric and $h^{\mu\nu}$ is small perturbations. In terms of the trace reversed metric $\bar{h}_{\mu\nu}$ and the Lorentz gauge condition $\bar{h}^{\mu\alpha}{}_{,\alpha} = 0$, the linearized Einstein field equations can be written as

$$\Box \bar{h}^{\mu\nu} = -16\pi \mathcal{T}^{\mu\nu}, \tag{4.5}$$

in which $\Box$ denotes the usual flat-space wave operator and the effective energy-momentum tensor $\mathcal{T}$ satisfies $\mathcal{T}^{\mu\nu}{}_{,\nu} = 0$. At the field point $(t, \vec{x})$, the solution $\bar{h}^{jk}$ $(j, k = \{x, y, z\})$ is the quadrupole-octupole formula which is under the point particle limit,:

$$\bar{h}^{jk} = \frac{2}{r}[\ddot{I}^{jk} - 2n_i\ddot{S}^{ijk} + n_i\frac{d^2M^{ijk}}{dt^3}], \tag{4.6}$$

where $r = |\vec{x}|$, $\vec{n} = \frac{\vec{x}}{r}$, the over dot is the time derivative with respect to observation time t Bekenstein [1973]; Press [1977], and we have neglected higher multi-moment. The mass quadrupole moment I^{jk}, mass octupole moments M^{ijk} and current quadrupole moments S^{ijk} are defined as

$$I^{jk} = \mu x^j x^k, \tag{4.7}$$
$$S^{ijk} = v^i I^{jk}, \tag{4.8}$$
$$M^{ijk} = x^i I^{jk}, \tag{4.9}$$

here $v^i = dx^i/dt$ is the three velocity of the particle.

In order to simplify our work, we start with only the mass quadrupole I^{ij} to construct the waveform.

$$h^{ij} = \ddot{I}^{ij}, \tag{4.10}$$

with the components of I^{jk} expressed in the pseudo-Cartesian coordinate

$$I^{jk} = \begin{pmatrix} r^2\sin^2\theta\cos^2\phi & r^2\sin^2\theta\sin\phi\cos\phi & r^2\sin\theta\cos\theta\cos\phi \\ r^2\sin^2\theta\sin\phi\cos\phi & r^2\sin^2\theta\cos^2\phi & r^2\sin\theta\cos\theta\sin\phi \\ r^2\sin\theta\cos\theta\cos\phi & r^2\sin\theta\cos\theta\sin\phi & r^2\cos^2\phi \end{pmatrix}. \tag{4.11}$$

Actually, the M^{ijk} and S^{ijk} will bring more harmonics, we will discuss them later in subsection 4.4.5.

In order to get the waveform in the TT gauge. We define an orthogonal spherical coordinate system via

$$\mathbf{e_r} = \frac{\partial}{\partial r}, \tag{4.12}$$
$$\mathbf{e_\Theta} = \frac{1}{r}\frac{\partial}{\partial\Theta}, \tag{4.13}$$
$$\mathbf{e_\Phi} = \frac{1}{r\sin\Theta}\frac{\partial}{\partial\Phi}. \tag{4.14}$$

The angles $\{\Theta, \Phi\}$ denote the observation point's latitude and azimuth, respectively. The waveform in transverse-traceless gauge is then given by

$$h^{jk}_{TT} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & h^{\Theta\Theta} - h^{\Phi\Phi} & 2h^{\Theta\Phi} \\ 0 & 2h^{\Theta\Phi} & -(h^{\Theta\Theta} - h^{\Phi\Phi}) \end{pmatrix}, \tag{4.15}$$

with

$$\begin{aligned} h^{\Theta\Theta} &= \cos^2\Theta[h^{xx}\cos^2\Phi + h^{xy}\sin 2\Phi + h^{yy}\sin^2\Phi] + h^{zz}\sin^2\Theta \\ &\quad - \sin 2\Theta[h^{xz}\cos\Phi + h^{yz}\sin\Phi], \\ h^{\Theta\Phi} &= \cos\Phi[-\frac{1}{2}h^{xx}\sin 2\Phi + h^{xy}\cos 2\Phi + \frac{1}{2}h^{yy}\sin 2\Phi] + \sin\Theta[h^{xz}\sin\Phi - h^{yz}\cos\Phi], \\ h^{\Phi\Phi} &= h^{xx}\sin^2\Phi - h^{xy}\sin 2\Phi + h^{yy}\cos^2\Phi. \end{aligned}$$

The "plus" and "cross" waveform polarizations are given by $h^{\Theta\Theta} - h^{\Phi\Phi}$ and $2h^{\Theta\Phi}$ respectively Babak *et al.* [2007].

We must emphasize that the NK prescription is inconsistent Babak *et al.* [2007]. Due to the gravitational potential of the background, the propagation of GW can not be as freely as in the case of flat space background. Just like the quantum mechanics, the barrier of the gravitational potential could reflect the GW back and forth in the "potential well". We use the quotes, because there is indeed no real well in the background, the gravitational potential is unbound. Therefore, the field at a point (t, x) depends not only on the surface of the past light cone, but also the interior of the past light cone, in other words, the whole history of the particle from the retarded point of (t, x). This effect is called the backscattering, which appears at 1.5PN order in Blanchet *et al.* [1995]. This bring a tail term in the full waveform construction. Such "tail" of waves are particularly predominant in the strong field Glampedakis & Kennefick [2002]. The assumption of the flat background implies the kludge waveforms are unable to capture any features related to backscattering. However, the backscattering becomes important for the orbits very close to the central black hole , which we do not consider in our work.

The spirit of this calculation is not a formal and consistent approximation to

EMRI waveforms; it is rather a "phenomenological" approach which takes into account those pieces of physics that are the most crucial in particular, the exact Kerr geodesic motion. By including the exact source trajectory, we ensure that the spectral components of the kludge waveforms are at the correct frequencies, although their relative amplitudes might not be exact, but good enough for our purpose.

4.3.1 The Kerr geodesic in observation time

The first step in the NK waveform modeling is to describe the trajectory. In this section we will review the geodesic in Kerr spacetime in the Boyer-Lindquist coordinate Bardeen *et al.* [1972]; Chandrasekhar [1983]. The metric in these coordinates is given by

$$\begin{aligned} ds^2 &= -\left(1-\frac{2Mr}{\Sigma}\right)dt^2+\frac{\Sigma}{\Delta}dr^2+\Sigma d\theta^2 \\ &\quad +\left(r^2+a^2+\frac{2Ma^2r}{\Sigma}\sin^2\theta\right)\sin^2\theta d\phi^2 \\ &\quad -\frac{4Mar}{\Sigma}\sin^2\theta dt d\theta, \end{aligned} \tag{4.16}$$

where

$$\Sigma = r^2+a^2\cos^2\theta, \qquad \Delta = r^2-2Mr+a^2. \tag{4.17}$$

The geodesic equations are the second order differential equations

$$\frac{d^2x^\alpha}{d\tau^2}+\Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx\gamma}{d\tau}=0, \tag{4.18}$$

where τ is the proper time, and $x^\alpha(\tau)$ is the trajectory of the geodesic. A geodesic is uniquely characterized by the initial value of particle's position $x^\alpha(0)$ and velocity $\frac{dx^\alpha}{d\tau}|_{\tau=0}$. Among this eight constants, $\frac{dx^\alpha}{d\tau}|_{\tau=0}$ correspond to the four constant of the motion which describe the dynamics of the bound orbit. Three of these constants are: the rest mass m; the conserved energy E per unit rest mass; the

conserved z-component of the angular momentum L_z per unit rest mass . There is a fourth constant Q which is also scaled by unit rest mass squared, called Carter constant Chandrasekhar [1983]; Misner *et al.* [1974].

Using these four scaled constants, the geodesic equations can be written as a system of first order differential equations

$$\begin{aligned}
\Sigma\frac{dr}{d\tau} &= \pm\sqrt{V_r}, \\
\Sigma\frac{d\theta}{d\tau} &= \pm\sqrt{V_\theta}, \\
\Sigma\frac{d\phi}{d\tau} &= \sqrt{V_\phi}, \\
\Sigma\frac{dt}{d\tau} &= \sqrt{V_t},
\end{aligned} \tag{4.19}$$

the various potentials are defined by

$$\begin{aligned}
V_r &= [E(r^2+a^2) - L_z a]^2 - \Delta[r^2 + (L_z - aE)^2 + Q], \\
V_\theta &= Q - \cos^2\theta\left[a^2(1-E^2) + \frac{L_z^2}{\sin^2\theta}\right], \\
V_\phi &= \frac{L_z}{\sin^2\theta} - aE + \frac{a}{\Delta}\left[E(r^2+a^2) - L_z a\right], \\
V_t &= a(L_z - aE\sin^2\theta) + \frac{r^2+a^2}{\Delta}\left[E(r^2+a^2) - L_z a\right].
\end{aligned} \tag{4.20}$$

For given E, L_z, Q, the roots of V_r determine the turning points of the radial motion: the periapsis r_p, and apoapsis r_a. From these, one can define an orbital eccentricity e, and semilatus rectum p, using the conventional Keplerian definitions:

$$\begin{aligned}
& r_p = \frac{p}{1+e}, \quad r_a = \frac{p}{1-e}, \\
& \Rightarrow p = \frac{2r_a r_p}{r_a + r_p}, \quad e = \frac{r_a - r_p}{r_a + r_p}.
\end{aligned} \tag{4.21}$$

For convenience, we replace the Carter constant by an "inclination angle" which

stands for the inclination of the spin to the orbital plane, defined by

$$\tan^2 \iota = \frac{Q}{L_z^2}. \tag{4.22}$$

From these, we can realize that $\{E, L_z, Q\}$ are equivalent to $\{e, p, \iota\}$. Theoretically, given any one set, we can find another one analytically.

The Boyer-Lindquist time coordinate t reduces at large radius to time as measured by a distant observer. From this point of view, it is more convenient to use t rather than τ in the geodesic equations Drasco & Hughes [2006]. It can be achieved by changing time variable in the above equation:

$$\frac{dr}{dt} = \frac{dr}{d\tau}\left(\frac{dt}{d\tau}\right)^{-1}, \tag{4.23}$$

and likewise for $d\theta/dt$ and $d\phi/dt$. Thus, with the initial conditions of the particle's position and an allowed set of orbital constants (E, L_z, Q), we can integrate the geodesic equation to find the trajectory $z(t) = \{r(t), \theta(t), \phi(t)\}$.

The author in Schmidt [2002] used the elegant Hamiltonian-Jacobi technique to show that bound orbits are characterized by multy-periodic motions in r, θ and ϕ with three fundamental frequencies Ω_r, Ω_θ, and Ω_ϕ respectively. The r and θ motions belong to the oscillatory or libration-type, but the ϕ orbital motion corresponds to a rotation-type periodic motion, so, Ω_ϕ is considered less fundamental than Ω_r and Ω_θ Drasco & Hughes [2004]. Unfortunately, the term $\Sigma = r^2 + a^2 \cos^2\theta$ on the left hand side of the geodesic equation leads to the coupling of the r and θ motions. To solve this problem, another time, λ, which absorbs the coupling is introduced by Mino [2003].

The Mino time parameter λ which decouples the radial and polar motions is related to the test mass's proper time τ by

$$\frac{d\tau}{d\lambda} = \Sigma. \tag{4.24}$$

The geodesic equations in λ - time are

$$\left(\frac{dr}{d\lambda}\right)^2 = V_r(r),$$

$$\left(\frac{d\theta}{d\lambda}\right)^2 = V_\theta(\theta),$$
$$\left(\frac{d\phi}{d\lambda}\right)^2 = V_\phi(r,\theta),$$
$$\left(\frac{dt}{d\lambda}\right)^2 = V_t(r,\theta). \tag{4.25}$$

Now, radial and polar motions are pure periodic in λ.

$$r(\lambda) = r(\lambda + n\Lambda_r), \tag{4.26}$$
$$\theta(\lambda) = \theta(\lambda + n\Lambda_\theta), \tag{4.27}$$

where the periods are given by

$$\Lambda_r = 2\int_{r_p}^{r_a} \frac{dr}{V_r(r)^{1/2}}, \tag{4.28}$$
$$\Lambda_\theta = 4\int_{\theta_{min}}^{\pi/2} \frac{d\theta}{V_\theta(\theta)^{1/2}}. \tag{4.29}$$
$$\tag{4.30}$$

The numerical factors in the front of the integrals come from the symmetry of the r, θ motion. The r motion is taken to range between periaps, r_p, and apoapsis, r_a; the θ motion ranges from a minimum θ_{min} to a maximum $\pi - \theta_{min}$ Drasco & Hughes [2006]. During one period, the r motion undergo $r_a \to r_p \to r_a$; the θ motion undergo $\theta_{min} \to \frac{\pi}{2} \to \pi - \theta_{min} \to \frac{\pi}{2} \to \theta_{min}$. The ϕ motion will be discussed later.

The frequencies in Mino time can be defined as

$$\Upsilon_{r,\theta} = \frac{2\pi}{\Lambda_{r,\theta}}. \tag{4.31}$$

Fourier series of $r(\lambda)$ and $\theta(\lambda)$ in these new frequencies can be expressed as:

$$r(\lambda) = \sum_{n=-\infty}^{\infty} r_n e^{-in\Upsilon_r\lambda}, \tag{4.32}$$

$$\theta(\lambda) = = \sum_{k=-\infty}^{\infty} \theta_k e^{-ik\Upsilon_\theta\lambda}, \tag{4.33}$$

where θ_k, r_n are constants, k, n are harmonic indices.

In terms of the angle variables: $w_r = \Upsilon_r\lambda, \quad w_\theta = \Upsilon_\theta\lambda$, these series can be simplified as

$$r(w_r) = \sum_{n=-\infty}^{\infty} r_n e^{-inw_r}, \tag{4.34}$$

$$\theta(w_\theta) = = \sum_{k=-\infty}^{\infty} \theta_k e^{-ikw_\theta}. \tag{4.35}$$

As mentioned above, $dt/d\lambda$ and $d\phi/d\lambda$ are biperiodic functions of r and θ. This means that they can be expanded in a two-dimensional Fourier series Drasco & Hughes [2004]:

$$\frac{dt}{d\lambda} = T(r,\theta) = \sum_{kn} T_{kn} e^{-i(k\Upsilon_\theta + n\Upsilon_r)\lambda}, \tag{4.36}$$

$$\frac{d\phi}{d\lambda} = \Phi(r,\theta) = \sum_{kn} \Phi_{kn} e^{-i(k\Upsilon_\theta + n\Upsilon_r)\lambda}, \tag{4.37}$$

with the expansion coefficients are given by

$$T_{kn} = \frac{1}{(2\pi)^2}\int_0^{2\pi} dw^r \int_0^{2\pi} dw^\theta T[r(w^r),\theta(w^\theta)]e^{i(kw^\theta + nw^r)}, \tag{4.38}$$

$$\Phi_{kn} = \frac{1}{(2\pi)^2}\int_0^{2\pi} dw^r \int_0^{2\pi} dw^\theta \Phi[r(w^r),\theta(w^\theta)]e^{i(kw^\theta + nw^r)}. \tag{4.39}$$

Because $V_t(r,\theta)$ and $V_\phi(r,\theta)$ are real, we have

$$T_{-k,-n} = \bar{T}_{kn}, \tag{4.40}$$
$$\Phi_{-k,-n} = \bar{\Phi}_{kn}. \tag{4.41}$$

In addition to this, the matrix T_{kn}, Φ_{kn} have the following properties: T_{k0} and T_{0n} are nonzero, but $T_{kn} = 0$, if $k \neq 0$ and $n \neq 0$ (and likewise for Φ_{kn}). This lack of "crosstalk" between the θ and r harmonics is because $T(r,\theta)$ and $\Phi(r,\theta)$ have the form $f(r) + g(\theta)$ Drasco & Hughes [2004]. To take advantage of this property, we define

$$\begin{aligned} T^{\theta}_{k} &= T_{k0}, \quad T^{r}_{n} = T_{0n}, \\ \Phi^{\theta}_{k} &= \Phi_{k0}, \quad \Phi^{r}_{n} = \Phi_{0n}, \end{aligned} \tag{4.42}$$

with this, we can rewrite above expressions as

$$\begin{aligned} \frac{dt}{d\lambda} &= \Gamma + \sum_{k=1}^{\infty}(T^{\theta}_{k}e^{-ik\Upsilon_{\theta}\lambda} + \bar{T}^{\theta}_{k}e^{ik\Upsilon_{\theta}\lambda}) \\ &\quad + \sum_{n=1}^{\infty}(T^{r}_{n}e^{-in\Upsilon_{r}\lambda} + \bar{T}^{r}_{n}e^{in\Upsilon_{r}\lambda}), \end{aligned} \tag{4.43}$$
$$\begin{aligned} \frac{d\phi}{d\lambda} &= \Upsilon_{\phi} + \sum_{k=1}^{\infty}(\Phi^{\theta}_{k}e^{-ik\Upsilon_{\theta}\lambda} + \bar{\Phi}^{\theta}_{k}e^{ik\Upsilon_{\theta}\lambda}) \\ &\quad + \sum_{n=1}^{\infty}(\Phi^{r}_{n}e^{-in\Upsilon_{r}\lambda} + \bar{\Phi}^{r}_{n}e^{in\Upsilon_{r}\lambda}). \end{aligned} \tag{4.44}$$

Here,

$$\Gamma = T_{00}, \tag{4.45}$$
$$\Upsilon_{\phi} = \Phi_{00}. \tag{4.46}$$

These two variables tell us about the secular average rate at which ϕ and t accumulate with respect to λ, the frequency Υ_{ϕ} has the same meaning as equations

(4.31). Therefore integrating these equations leads to

$$t(\lambda) = \Gamma\lambda + \delta t(\lambda), \tag{4.47}$$
$$\phi(\lambda) = \Upsilon_\phi\lambda + \delta\phi(\lambda). \tag{4.48}$$

we have chosen $t(\lambda = 0) = 0 = \phi(\lambda = 0)$, and defined oscillatory part as

$$\delta t(\lambda) = \sum_{k=1}^{\infty}(\delta t_k^\theta e^{-ik\Upsilon_\theta\lambda} + c.c.) + \sum_{n=1}^{\infty}(\delta t_n^r e^{-in\Upsilon_r\lambda} + c.c.), \tag{4.49}$$
$$\delta\phi(\lambda) = \sum_{k=1}^{\infty}(\delta\phi_k^\theta e^{-ik\Upsilon_\theta\lambda} + c.c.) + \sum_{n=1}^{\infty}(\delta\phi_n^r e^{-in\Upsilon_r\lambda} + c.c.), \tag{4.50}$$

where

$$\delta t_j^{r,\theta} = \frac{iT_j^{r,\theta}}{j\Upsilon_{r,\theta}}, \tag{4.51}$$
$$\delta\phi_j^{r,\theta} = \frac{i\Phi_j^{r,\theta}}{j\Upsilon_{r,\theta}}. \tag{4.52}$$

With this definition, $t(\lambda)$ and $\phi(\lambda)$ are decomposed into pieces which accumulate secularly with λ plus pieces $\delta t(\lambda)$ and $\delta\phi(\lambda)$ that oscillate at harmonics of Υ_r and Υ_θ.

Since Ω_ϕ is the average rate at which ϕ accumulates as a function of t and since Γ and Υ_ϕ are the average rates at which t and ϕ accumulate as a functions of λ, we have the following relations between frequencies in different time parameters Drasco & Hughes [2004]

$$\Omega_\phi = \frac{\Upsilon_\phi}{\Gamma}, \tag{4.53}$$
$$\Omega_r = \frac{\Upsilon_r}{\Gamma}, \tag{4.54}$$
$$\Omega_\theta = \frac{\Upsilon_\theta}{\Gamma}. \tag{4.55}$$

4.3.2 Changes of variables for radial and polar coordinates

From the point of vies of the numerical calculation, evaluating the integral containing the function of $r(w^r)$ and $\theta(w^\theta)$ would require a direct integration of the radial and polar geodesic equations. Unfortunately this is somewhat difficult because the potentials V_r and V_θ vanish at the orbital turning points Babak *et al.* [2007]; Drasco & Hughes [2004]. This means, for example, the integral in the direct expression for $r(\lambda)$

$$\lambda = \int_{r_{min}}^{r(\lambda)} \frac{dr'}{\pm\sqrt{V_r(r')}}, \tag{4.56}$$

contains apparent singularities which complicate the numerical implementation. In order to regularize the integral at the turning points we introduce new variable ψ through the equation

$$r = \frac{p}{1 + e\cos\psi}, \tag{4.57}$$

and define χ by the equation $z = \cos^2\theta = z_- \cos^2\chi$ for the θ motion, where z_- is given by

$$\beta(z_+ - z)(z - z_-) = \beta z^2 - z[Q + L_z^2 + a^2(1 - E^2)] + Q,, \tag{4.58}$$

with $\beta = a^2(1-E^2)$. $z_\pm$ are the two roots of the quadratic equation (4.58) Babak *et al.* [2007]; Drasco & Hughes [2004].

The radial potential V_r has four roots, two of them are r_p and r_a, the other two roots are denoted as r_3 and r_4. By using these four roots, we can expand the radial potential as

$$V_r = (1 - E^2)(r_a - r)(r - r_p)(r - r_3)(r - r_4). \tag{4.59}$$

The radii r_3 and r_4 do not correspond to the turning point, but still represent

zeros of V_r. It is useful to rewrite these radii by

$$r_3 = \frac{p_3}{1-e}, \tag{4.60}$$
$$r_4 = \frac{p_4}{1+e}. \tag{4.61}$$

In fact, r_3 is used to find the innermost stable orbit, and r_4 is always fall into the event horizon, we return to this discussion later in the subsection 4.3.4 on truncation of the orbital evolution.

In the Mino time, the geodesic equations for χ and ψ are

$$\begin{aligned}
\frac{d\chi}{d\lambda} &= \sqrt{\beta(z_+ - z)} = \sqrt{\beta(z_+ - z_- \cos^2\chi)}, \\
\frac{d\psi}{d\lambda} &= \frac{M\sqrt{1-E^2}[(p-p_3) - e(p+p_3\cos\psi)]^{1/2}[(p-p_4) - e(p-p_4\cos\psi)]^{1/2}}{1-e^2}.
\end{aligned}$$

In the coordinate time t, the evolution equations for ψ and χ have the form:

$$\begin{aligned}
\frac{d\psi}{dt} &= \frac{M\sqrt{1-E^2}[(p-r_3(1-e)) - e(p+r_3(1-e)\cos\psi)]^{1/2}}{[\gamma + a^2 E z(\chi)](1-e^2)} \\
&\quad \cdot[(p-r_4(1+e)) - e(p+r_4(1+e)\cos\psi)]^{1/2}, \\
\frac{d\chi}{dt} &= \frac{\sqrt{\beta[z_+ - z(\chi)]}}{\gamma + a^2 E z(\chi)},
\end{aligned} \tag{4.62}$$

where $\gamma = e\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\right] - \frac{2MraL_z}{\Delta}$ Babak *et al.* [2007].

In terms of the variables ψ, and χ, the geodesic equations are well behaved at the turning points, which facilitates numerical integration. A geodesic can be described by a set of six constants $I = \{E, L_z, Q, \psi_0, \phi_0, \chi_0\}$, where $\psi_0, \chi_0, \varphi_0$ are the initial position in terms of the introduced above phase angles.

4.3.3 Adiabatic approximation

Till now, we have considered particle moving in a pure geodesic. But in reality, the particle also has its own gravitational field which will disturb the background,

this disturbance cause the emission of GWs. Due to the loss of gravitational energy and orbital angular momentum, there will be a "back reaction" effect on the particle's motion. We can consider this effect as a self force, which causes it to deviated from its geodesic path Barack & Sago [2010]; Gair *et al.* [2011]; Poisson *et al.* [2011]. Because of the difficulty in calculating the whole self force, it is usually assumed that the evolution is adiabatic: the time scale τ_{RR} for the orbital parameters to change is much larger than the orbital periods of the r, θ, ϕ motion. In this limit, at each instance, the trajectory can be treated as nearly a geodesic orbit. Suppose all the possible Kerr geodesic orbit constitute a space, and each set of (E, L_z, Q) corresponds to a geodesic which can be treated as a point in this space. When orbits evolve due to the radiation reaction, they correspond to a trajectory $[E(t), L_z(t), Q(t)]$ through this space. The orbital motion can be treated as slowly passing from one geodesic configuration to another, the effect of this force can be understand as a slow change in all the six constants $(E, L_z, Q, \psi_0, \phi_0, \chi_0)$.

4.3.4 Evolution of the constants of motion

It is necessary to use the self force to deduce the rate of change for the constants of motion Mino [2003]. The self force has two components: a dissipative force which corresponds to carrying energy and angular momentum away from the binary, and it can be used to derive the rate of change of E, L_z, Q; and, a conservative part of the force which is related to the change of the initial position χ_0, ψ_0, ϕ_0 Barack [2009]; Barack & Sago [2010, 2011]; Hinderer & Flanagan [2008]; Warburton *et al.* [2012]. There is no analytical expression of the self force, it is not trivial to obtain $\dot{\chi}_0, \dot{\psi}_0, \dot{\phi}_0$ and to evolve the motion. We neglect the change of χ_0, ψ_0, ϕ_0, the instantaneous geodesic orbit is updated by evolving the three constants E, L_z, Q. In the future, if we can get the accurate expressions of the $\dot{\chi}_0, \dot{\psi}_0, \dot{\phi}_0$, we can also apply them in our method to obtain the accurate orbital motion. This implies that we do not model the phase of GW signal well enough, but, we do not intend to use NK waveform for detection purpose. We will use NK model to mimic the signal for testing phenomenological waveform described in detail below.

The energy and the orbital angular momentum can be read out from the

gravitational wave flux, but the Carter constant can not Hughes [2000]; Kennefick & Ori [1996]. This is because the energy and angular momentum are linearly constructed from the particle's momentum p^μ, while the Carter constant is quadratically constructed from the momentum. Sago *et al.* [2005, 2006] shows that all the change of these motion constants can be found from the self force. Ganz *et al.* [2007] provide an adiabatic expressions of $\dot{E}, \dot{L}_z, \dot{Q}$ by solving the Teukolsky equations Teukolsky [1973]. For generic orbit, it is necessary to consider all these three rates, but there are two special cases which need evolution of only two components. The first one is the equatorial orbit: Q for equatorial orbit is zero, and since the test mass does not leave the equatorial plane during the evolution, Q remains zero all the time. The system evolution is govern by the two quantities $\dot{E}, \dot{L}_z$. The other one is the circular, non-equatorial orbit: such an orbit has non-zero Cartar constant, but it turn out that, in the adiabatic approximation, its evolution is entirely determined by the radiated energy and orbital angular momentum, this is because circular orbit remain circular as they adiabatically evolve due to the radiation reaction and it is enough to use the measurement of $\dot{E}, \dot{L}_z$ to entirely specify the evolution of the system Hughes [2000].

Our analysis does not depend on precise description of the orbit. The aim is to show how our method works in practice. No matter what kind of approximation formula for $\dot{E}, \dot{L}_z, \dot{Q}$ is used, the analysis described in this thesis remains unchanged. Just instead of equations (4.63) one would use other (more precise and more complicated expressions). In order to simplify our work, we mimic the toy gravitational signal by choosing the 2PN approximation formula for $\dot{E}, \dot{L}_z, \dot{Q}$ derived in Gair & Glampedakis [2006] for the orbital evolution. Those are

$$\begin{aligned}
\frac{dE}{dt}_{2PN} &= -\frac{32}{5}\frac{m^2}{M^2}\left(\frac{M}{p}\right)^5(1-e^2)^{\frac{3}{2}}\Big[g_1(e) - \frac{a}{M}\left(\frac{M}{p}\right)^{\frac{3}{2}}g_2(e)\cos\iota - \left(\frac{M}{p}\right)g_3(e) + \pi\left(\frac{M}{p}\right)^{\frac{3}{2}}g_4(e) \\
&\quad -\left(\frac{M}{p}\right)^2 g_5(e) + \left(\frac{a}{M}\right)^2\left(\frac{M}{p}\right)^2 g_6(e) - \frac{527}{96}\left(\frac{a}{M}\right)^2\left(\frac{M}{p}\right)^2\sin^2\iota + \mathcal{O}(v^5)\Big], \\
\frac{dL_z}{dt}_{2PN} &= -\frac{32}{5}\frac{m^2}{M}\left(\frac{M}{p}\right)^{\frac{7}{2}}(1-e^2)^{\frac{3}{2}}\Big[g_7(e)\cos\iota + \frac{a}{M}\left(\frac{M}{p}\right)^{\frac{3}{2}}\{g_8(e) - \cos^2\iota g_9(e)\}
\end{aligned}$$

$$
\begin{aligned}
& -(\frac{M}{p})g_{10}(e)\cos\iota + \pi(\frac{M}{p})^{\frac{3}{2}}g_{11}(e)\cos\iota - (\frac{M}{p})^2 g_{12}(e)\cos\iota \\
& +(\frac{a}{M})^2(\frac{M}{p})^2\cos\iota(g_{13}(e) - \frac{45}{8}\sin^2\iota) + \mathcal{O}(v^5)], \\
\frac{dQ}{dt}_{2PN} &= -\frac{64}{5}\frac{m^2}{M}(\frac{M}{p})^{\frac{7}{2}}\sqrt{Q}\sin\iota(1-e^2)^{\frac{3}{2}}[g_7(e) - \frac{a}{M}(\frac{M}{p})^{\frac{3}{2}}\cos\iota g_9(e) - (\frac{M}{p})g_{10}(e) \\
& +\pi(\frac{M}{p})^{\frac{3}{2}}g_{11}(e) - (\frac{M}{p})^2 g_{12}(e) + (\frac{a}{M})(\frac{M}{p})(g_{13}(e) - \frac{45}{8}\sin^2\iota)]. \qquad (4.63)
\end{aligned}
$$

Here :

$$
\begin{aligned}
g_1(e) &= 1 + \frac{73}{24}e^2 + \frac{37}{96}e^4, \\
g_2(e) &= \frac{73}{12} + \frac{823}{24}e^2 + \frac{949}{32}e^4 + \frac{491}{192}, \\
g_3(e) &= \frac{1247}{336} + \frac{9181}{672}e^2, \\
g_4(e) &= 4 + \frac{1375}{48}e^2, \\
g_5(e) &= \frac{44711}{9072} + \frac{172157}{2592}e^2, \\
g_6(e) &= \frac{33}{16} + \frac{359}{32}e^2, \\
g_7(e) &= 1 + \frac{7}{8}e^2, \\
g_8(e) &= \frac{61}{24} + \frac{63}{8}e^2 + \frac{95}{64}e^4, \\
g_9(e) &= \frac{61}{8} + \frac{91}{4}e^2 + \frac{461}{64}e^4, \\
g_{10}(e) &= \frac{1247}{336} + \frac{425}{336}e^2, \\
g_{11}(e) &= 4 + \frac{97}{8}e^2, \\
g_{12}(e) &= \frac{44711}{9072} + \frac{302893}{6048}e^2, \\
g_{10}(e) &= \frac{33}{16} + \frac{95}{16}e^2. \qquad (4.64)
\end{aligned}
$$

We must emphasize that these 2PN formula are not sufficiently accurate to deduce the rate of change for E, L_z, Q. However, the comparison between the NK waveform which constructed by using the 2PN formula to evolve the particle

orbit and the Teukolsky-based numerical waveform Cutler *et al.* [1994]; Drasco & Hughes [2006]; Glampedakis & Kennefick [2002]; Hughes [2001]; Poisson [2004] show that, it is accurate enough in the mild relativistic region Babak *et al.* [2007]. When accurate description of the orbital evolution is available, it can be used to improve the NK waveform. We can use orbital evolution under the numerically computed self force, and use it instead of PN expression, then, we again can use the weak field approximation to compute the gravitational waveform.

With these rate of change for the constants of motion, it is still not convenient to understand the orbital evolution, we will rewrite the orbital evolution in terms of $e(t), p(t), \iota(t)$ and the three fundamental frequencies.

The rate of change for e, p, ι can be deduced from their definitions,

$$\frac{dp}{dt} = \frac{1}{2}\left[(1+e)^2\frac{dr_p}{dt} + (1-e)^2\frac{dr_a}{dt}\right], \tag{4.65}$$

$$\frac{de}{dt} = \frac{1}{2p}(1+e)(1-e)\left[(1-e)\frac{dr_a}{dt} - (1+e)\frac{dr_p}{dt}\right], \tag{4.66}$$

$$\frac{d\iota}{dt} = \frac{\sin\iota\cos\iota\frac{dQ}{dt}}{2Q} - \frac{\sin^2\iota\frac{dL_z}{dt}}{\sqrt{Q}}. \tag{4.67}$$

Next, we use the fact that r_a and r_p are the turning points:

$$[E(r_a^2+a^2) - L_z a]^2 - \Delta(r_a)[r_a^2 + (L_z - aE)^2 + Q] = 0, \tag{4.68}$$

$$[E(r_p^2+a^2) - L_z a]^2 - \Delta(r_p)[r_p^2 + (L_z - aE)^2 + Q] = 0. \tag{4.69}$$

For convenience, we use the variable y instead of r_a or r_p. Using the above expressions, Gair & Glampedakis [2006] derived the following formula:

$$\frac{dy}{dE} = -\frac{N1}{D}, \quad \frac{dy}{dL_z} = -\frac{N2}{D}, \quad \frac{dy}{dQ} = -\frac{N3}{D}, \tag{4.70}$$

where

$$
\begin{aligned}
N1 &= 2Ey^4 + 2Ea^2y^2 - 4aMyL_z + 4Mya^2E, && (4.71)\\
N2 &= -2y^2L_z + 4MyL_z - 4MyaE, && (4.72)\\
N3 &= -y^2 + 2My - a^2, && (4.73)\\
D &= 2E^2(y^2 + a^2)2y - 2EL_z a2y - 2y(y^2 - 2My + a^2) \\
&\quad -y^2(2y - 2M) - (L_z - aE)^2(2y - 2M) - (2y - 2M)Q. && (4.74)
\end{aligned}
$$

Then substitute r_a and r_p instead of y, we can get $\frac{dr_a}{dt}$, $\frac{dr_p}{dt}$. Instead of the rates $\frac{dE}{dt}$, $\frac{dL_z}{dt}$, and $\frac{dQ}{dt}$, we can have $\frac{de}{dt}$, $\frac{dp}{dt}$, $\frac{d\iota}{dt}$. With these rates of change and the initial conditions, we can calculate the value of e, p, ι, E, L_z, Q at each moment of time, and use these instant values to compute the three fundamental frequencies $\Omega_r, \Omega_\theta, \Omega_\phi$. Next we need to formulate conditions for termination of the orbital evolution.

4.3.5 The truncation of the waveform

The NK waveform becomes increasingly inaccurate as a small object moves deeper into the central black hole's potential well, and we need to define conditions for termination of orbital evolution. The natural termination come from the last stable orbit (LSO). In Kerr space time, it is fulfilled by analysing the four real roots $r_a > r_p > r_3 > r_4$ of the radial potential V_r. During the early stage of the inspiral, the orbit is in the weak field with large p. These four roots are different from each other. The orbit is approximately a Keplerian eccentric orbit with particle's radial motion between r_a and r_p. As inspiral proceeds, both the size of the orbit and eccentricity decrease . The turning points r_a and r_p are gradually approach the values for circular orbit, $r_a = r_p$. Moreover, the roots r_p approached r_3, when $r_p = r_3$, the orbit will be marginally stable, or reaches LSO (the triple root $r_a = r_p = r_3$ denote a marginally stable circular orbit), below this orbit, the test particle will rapidly plunge into the black hole. Thus, $r_p = r_3$ condition is called the separatrix which is used to distinguish stable and unstable orbit, and the inspiral evolution should be terminated when this condition satisfied Cutler *et al.* [1994]; Glampedakis & Kennefick [2002].

The radial frequency Ω_r is gradually increase during the early stage of the inspiral. As the orbit approaches the separatrix, the particle spends a considerable amount of its orbital life close to the periapsis, this leads to increase in the orbit period of radial motion. The evolution of Ω_r should have a turning point, before attaining this turning point, Ω_r increase slowly to the maximum, after that, Ω_r began to decrease. Near the separatirx, the particle will trace a quasi-circular path before being reflected back to the apoapsis Cutler *et al.* [1994]. Such behavior will be particularly prominent for eccentric orbits: the particle will "zoom in" from its apoapsis position, and perform a certain number of quasi-circular revolutions "whirls" close to the periapsis which should have a value close to the radius of LSO Glampedakis & Kennefick [2002]. Finally, the particle will be reflected and "zoom out" towards the apoapsis again. This directly leads to the large increase in the orbital periods, and bring down the radial frequency Ω_r.

Near the separatrix, besides the decreasing of the Ω_r, the eccentricity will slightly increase Barack & Sago [2010]; Drasco [2009]; Glampedakis & Kennefick [2002]. So there must exist a critical condition $\dot{e} = 0$.

In our work, we do not consider the region which is close to the separatrix. The NK waveform becomes increasingly inaccurate in this region, and we use combination of conditions: separatrix, Ω_r decrease, $\dot{e} < 0$ as truncation conditions (whatever happens first) for the orbital evolution.

4.4 Harmonic Analysis of the Numerical Kludge Waveform

As mentioned above, the orbital motion can be characterized by the three slowly changing fundamental frequencies $\Omega_r(t), \Omega_\theta(t), \Omega_\phi(t)$ or $\Upsilon_r(\lambda), \Upsilon_\theta(\lambda), \Upsilon_\phi(\lambda)$. From the quadrupole formula, we can see that, the waveform h^{ij} is constructed directly from the orbital motion. This implies that we can decompose h^+, $h^\times$ into sets of harmonics Drasco [2009]; Drasco & Hughes [2004, 2006]. Due to the adiabatic approximation, the orbit can be treated as a geodesic in each instance. We analyze the harmonics structure of NK in the next subsection

4.4.1 Fourier expansion of arbitrary function $f(r, \theta, \phi)$

In this section, we review the results of Fourier expansion of arbitrary function $f(r, \theta, \phi)$ which is derived by Drasco [2009]; Drasco & Hughes [2004, 2006].

For a bound geodesic, the r, θ motions are periodic. Any function $f(r, \theta)$ can be decomposed as the Fourier expansion in both observational time t and Mino time λ.

$$f[r(t), \theta(t)] = \sum_{kn} f_{kn} e^{-i\Omega_{kn} t}, \tag{4.75}$$

$$f[r(\lambda), \theta(\lambda)] = \sum_{kn} \tilde{f}_{kn} e^{-i\Upsilon_{kn} \lambda}, \tag{4.76}$$

here, we have defined:

$$\Omega_{kn} = k\Omega_\theta + n\Omega_r, \tag{4.77}$$

$$\Upsilon_{kn} = k\Upsilon_\theta + n\Upsilon_r. \tag{4.78}$$

The coefficients f_{kn} tell us about the harmonic structure of the function $f(r(t), \theta(t))$ as seen by distant observer. From computational point of view, it is easier to obtain $\tilde{f}_{kn}$, and because these two kind of expansions are essentially equivalent, there is a relationship between these set of coefficients. The Fourier coefficients are

$$f_{kn} = \frac{\mathcal{F}_{kn}(\Omega_{kn})}{\Gamma}, \tag{4.79}$$

here Γ if already defined in equation (4.46), and $\mathcal{F}_{kn}(\omega)$ is the Fourier expansion coefficients of $\mathcal{F}(\omega^\theta, \omega^r, \omega)$ which are defined as:

$$\begin{aligned} \mathcal{F}(w^\theta, w^r, \omega) &= T[r(w^r), \theta(w^\theta)] e^{i\omega \delta t(w^r, w^\theta)} f[r(w^r), \theta(w^\theta)], \\ &= \sum_{kn} \mathcal{F}_{kn}(\omega) e^{-i\Upsilon_{kn}\lambda}, \end{aligned} \tag{4.80}$$

with coefficients given by

$$\mathcal{F}_{kn}(\omega) = \frac{1}{(2\pi)^2}\int_0^{2\pi} dw^r \int_0^{2\pi} dw^\theta \mathcal{F}[w^r, w^\theta, \omega] e^{ikw^\theta} e^{inw^r}. \tag{4.81}$$

Later, we will also need to take into account the first and higher order derivative of $f[r(t), \theta(t)]$. Set $\frac{d^m f}{dt^m} = g[r(t), \theta(t)] = \sum_{kn} g_{kn} e^{-i\Omega_{kn} t}$. The Fourier expansion is

$$g[r(t), \theta(t)] = \sum_{kn} g_{kn} e^{-i\Omega_{kn} t}. \tag{4.82}$$

Therefore the Fourier coefficients of derivative are

$$g_{kn} = (-i\Omega_{kn})^m f_{kn}. \tag{4.83}$$

4.4.2 Fourier expansion of $h^{+,\times}$

In this subsection, we still consider the orbit as the pure geodesic, where the three fundamental frequencies are constants. In order to analyse influence of Ω_ϕ in the Fourier expansion of the function $f[r(t), \theta(t), \phi(t)]$, we need to take into account the following expansion:

$$f(r(t), \theta(t), \phi(t)) = \sum_{knm} f_{knm} e^{-i\Omega_{knm} t}, \tag{4.84}$$

where $\Omega_{knm} = k\Omega_\theta + n\Omega_r + m\Omega_\phi$.

In the quadrupole formula, each component of I^{ij} are in the form of trigonometric functions of θ, ϕ. In order to find the harmonics decomposition of I^{ij}, it is necessary to understand the harmonic decomposition of the trigonometric function of ϕ. First, let us have a look at $\phi(\lambda)$ and $t(\lambda)$.

Take into account the decomposition of equation (4.48), we can decompose $\phi(\lambda)$ into several parts:

$$\begin{aligned} \phi(\lambda) &= \Upsilon_\phi \lambda + \delta\phi(\lambda) \\ &= \frac{\Upsilon_\phi T_{00} \lambda}{T_{00}} + \delta\phi(w^r, w^\theta) \end{aligned}$$

$$
\begin{aligned}
&= \Omega_\phi T_{00}\lambda + \delta\phi_r(w^r) + \delta\phi_\theta(w^\theta) \\
&= \Omega_\phi[t(\lambda) - \delta t_r(w^r) - \delta t_\theta(w^\theta)] + \delta\phi_r(w^r) + \delta\phi_\theta(w^\theta) \\
&= \Omega_\phi t(\lambda) - \Omega_\phi[\delta t_r(w^r) + \delta t_\theta(w^\theta)] + \delta\phi_r(w^r) + \delta\phi_\theta(w^\theta) \\
&\equiv \Omega_\phi t(\lambda) + \phi_r(w^r) + \phi_\theta(w^\theta) \qquad (4.85)
\end{aligned}
$$

Here, $\phi_r(w^r), \phi_\theta(w^\theta)$ are defined as

$$
\begin{aligned}
\phi_r(w^r) &= -\Omega_\phi \delta t_r(w^r) + \delta\phi_r(w^r), \qquad (4.86) \\
\phi_\theta(w^\theta) &= -\Omega_\phi \delta t_\theta(w^\theta) + \delta\phi_\theta(w^\theta). \qquad (4.87)
\end{aligned}
$$

The first part is directly related to $\Omega_\phi t$ and it is an average accumulation of the phase with the angular frequency Ω_ϕ. The second, and third parts are the oscillatory terms with frequencies Ω_r Ω_θ respectively.

Next, we consider the trigonometric function of ϕ. In the following example of $\sin(\phi(\lambda))$, we can see that the decomposition of $\phi(\lambda)$ will paly an important role in the harmonics analysis. Consider $\sin(\phi(\lambda))$ as an example:

$$
\begin{aligned}
\sin(\phi(\lambda)) &= \sin(\Omega_\phi t(\lambda))\cos(\phi_r + \phi_\theta) + \cos(\Omega_\phi t(\lambda))\sin(\phi_r + \phi_\theta) \\
&= \sin(\Omega_\phi t)\cos\phi_r\cos\phi_\theta - \sin(\Omega_\phi t)\sin\phi_r\sin\phi_\theta \\
&+ \cos(\Omega_\phi t)\sin\phi_r\cos\phi_\theta + \cos(\Omega_\phi t)\cos\phi_r\sin\phi_\theta, \qquad (4.88)
\end{aligned}
$$

and using :

$$
\sin(\Omega_\phi t) = \frac{e^{i\Omega_\phi t} - e^{-i\Omega_\phi t}}{2i}, \quad \cos(\Omega_\phi t) = \frac{e^{i\Omega_\phi t} + e^{-i\Omega_\phi t}}{2}. \qquad (4.89)
$$

The term $\cos\phi_r\cos\phi_\theta$ can be treated as a function of r and θ, they can be expressed as the harmonic decomposition as $\sum_{kn} A_{kn} e^{-i(k\Omega_\theta + n\Omega_r)t}$. Together with the decomposition of $\sin(\Omega_\phi t) = \sum_{m=\pm1} \frac{m}{2i} e^{im\Omega_\phi}$, we have the harmonic expansion of $\sin(\phi(\lambda))\cos\phi_r\cos\phi_\theta$ as

$$
\sum_{kn,m=\pm1} \frac{mA_{kn}}{2i} e^{-i(k\Omega_\theta + n\Omega_r + m\Omega_\phi)t}.
$$

The other three terms in the equation (4.88) can also be handled similarly. Com-

bine the decomposition of these four terms, we can have the final harmonic decomposition of $\sin(\phi(\lambda))$.

This method can also be used for other combinations of trigonometric function. The idea is to decompose those combination into the terms which contain $\Omega_\phi t$, ϕ_r, ϕ_θ,

$$\begin{aligned}
&\cos(m\Omega_\phi t)\cos(n\phi_r)\cos(k\phi_\theta),\\
&\cos(m\Omega_\phi t)\sin(n\phi_r)\sin(k\phi_\theta),\\
&\sin(m\Omega_\phi t)\sin(n\phi_r)\cos(k\phi_\theta),\\
&\sin(m\Omega_\phi t)\cos(n\phi_r)\sin(k\phi_\theta),
\end{aligned} \tag{4.90}$$

and the other terms which are just the function of r and θ. Let us take the $I^{11} = r^2\sin^2\theta\sin^2\phi$ as an example:

$$\begin{aligned}
I^{xx} &= \frac{r^2\sin^2\theta}{2} + \frac{r^2\sin^2\theta\cos(2\phi)}{2}\\
&= \cos(2\Omega_\phi t)\left[\frac{\cos(2\phi_r)\cos(2\phi_\theta)r^2\sin^2\theta}{2}\right]\\
&\quad + \cos(2\Omega_\phi t)\left[\frac{-\sin(2\phi_r)\sin(2\phi_\theta)r^2\sin^2\theta}{2}\right]\\
&\quad + \sin(2\Omega_\phi t)\left[\frac{-\sin(2\phi_r)\cos(2\phi_\theta)r^2\sin^2\theta}{2}\right]\\
&\quad + \sin(2\Omega_\phi t)\left[\frac{-\cos(2\phi_r)\sin(2\phi_\theta)r^2\sin^2\theta}{2}\right]\\
&\quad + \left[\frac{r^2\sin^2\theta}{2}\right]\\
&= \cos(2\Omega_\phi t)\cdot V^1 + \cos(2\Omega_\phi t)\cdot V^2 + \sin(2\Omega_\phi t)\cdot V^3 + \sin(2\Omega_\phi t)\cdot V^4 + V^5.
\end{aligned}$$

Here V_i stand for the corresponding terms in the bracket, they are the function of r and θ. Just like what we have done in the case $\sin(\phi)$, the $\cos(2\Omega_\phi t)$ and $\sin(2\Omega_\phi t)$ give the contributions on $e^{-im\Omega_\phi t}$ with $m = \pm 2$ in the harmonics decomposition. The contribution of $e^{-i(k\Omega_\theta + n\Omega_r)t}$ will come from the terms V_i. Combine them together, we can obtain the final harmonics decomposition of I^{11}.

Having the harmonics decomposition of I^{ij}, we can go further to the case of h^{ij} which are the second derivative of I^{ij}. The idea is the same, by using the algebric calculation, decompose h^{ij} into a linear combination of terms in the form of $\cos(2\Omega_\phi t)\cdot f(r,\theta)$. We continue to use h^{xx} as an example, the second derivative is:

$$\begin{aligned}\frac{d^2I^{xx}}{dt^2} &= -(2\Omega_\phi)^2\cos(2\Omega_\phi t)\cdot V^1 - 2(2\Omega_\phi)\sin(2\Omega_\phi t)\cdot \dot{V}^1 + \cos(2\Omega_\phi t)\cdot \ddot{V}^1 \\ &\quad -(2\Omega_\phi)^2\cos(2\Omega_\phi t)\cdot V^2 - 2(2\Omega_\phi)\sin(2\Omega_\phi t)\cdot \dot{V}^2 + \cos(2\Omega_\phi t)\cdot \ddot{V}^2 \\ &\quad -(2\Omega_\phi)^2\sin(2\Omega_\phi t)\cdot V^3 + 2(2\Omega_\phi)\cos(2\Omega_\phi t)\cdot \dot{V}^3 + \sin(2\Omega_\phi t)\cdot \ddot{V}^3 \\ &\quad -(2\Omega_\phi)^2\sin(2\Omega_\phi t)\cdot V^4 + 2(2\Omega_\phi)\cos(2\Omega_\phi t)\cdot \dot{V}^4 + \sin(2\Omega_\phi t)\cdot \ddot{V}^4 \\ &\quad +\ddot{V}^5. \end{aligned} \tag{4.91}$$

Again, V^i $\dot{V}^i$ and $\ddot{V}^i$ are just the function of r and θ which give the contribution to $e^{-i(k\Omega_\theta+n\Omega_r)t}$. The harmonic decomposition of h^{xx} can be expressed as

$$h^{xx} = \sum_{knm} h^{xx}_{knm} e^{-i\Omega_{knm}t} \tag{4.92}$$

here

$$\Omega_{knm} = k\Omega_\theta + n\Omega_r + m\Omega_\phi \tag{4.93}$$

and $m = 0, \pm 2$. The terms with $m = 0$ come form $\ddot{V}^5$, and the $m = \pm 2$ arise from $\sin(2\Omega_\phi t)$ and $\cos(2\Omega_\phi t)$. For the other components h^{jk} we can use the same method to express them into the Fourier expansion.

Using the above techniques, $h^{\Theta\Theta}$ $h^{\Theta\Phi}$ and $h^{\Phi\Phi}$ have the following harmonics decomposition:

$$\begin{aligned} h^{\Theta\Theta} &= \sum_{knm} h^{\Theta\Theta}_{knm} e^{-(ik\Omega_\theta+n\Omega_r+m\Omega_\phi)t}, \\ h^{\Theta\Theta}_{kn,-2} &= \cos^2\Theta\cos^2\Phi h^{xx}_{kn,-2} + \cos^2\Theta\sin 2\Phi h^{xy}_{kn,-2} + \cos^2\Theta\sin^2\Phi h^{yy}_{kn,-2}, \\ h^{\Theta\Theta}_{kn,-1} &= -\sin 2\Theta\cos\Phi h^{xz}_{kn,-1} - \sin 2\Theta\sin\Phi h^{yz}_{kn,-1}, \\ h^{\Theta\Theta}_{kn,0} &= \cos^2\Theta\cos^2\Phi h^{xx}_{kn,0} + \cos^2\Theta\sin^2\Phi h^{yy}_{kn,0} + \sin^2\Theta h^{zz}_{kn,0}, \\ h^{\Theta\Theta}_{kn,1} &= -\sin 2\Theta\cos\Phi h^{xz}_{kn,1} - \sin 2\Theta\sin\Phi h^{yz}_{kn,1}, \end{aligned}$$

$$h^{\Theta\Theta}_{kn,2} = \cos^2\Theta\cos^2\Phi h^{xx}_{kn,2} + \cos^2\Theta\sin 2\Phi h^{xy}_{kn,2} + \cos^2\Theta\sin^2\Phi h^{yy}_{kn,2}.$$

$$\begin{aligned}
h^{\Theta\Phi} &= \sum_{knm} h^{\Theta\Phi}_{knm} e^{-(ik\Omega_\theta + n\Omega_r + m\Omega_\phi)t}, \\
h^{\Theta\Theta}_{kn,-2} &= -\frac{1}{2}\cos\Theta\sin 2\Phi h^{xx}_{kn,-2} + \cos\Theta\cos 2\Phi h^{xy}_{kn,-2} + \frac{1}{2}\cos\Theta\sin 2\Phi h^{yy}_{kn,-2}, \\
h^{\Theta\Phi}_{kn,-1} &= -\sin\Theta\sin\Phi h^{xz}_{kn,-1} - \sin\Theta\cos\Phi h^{yz}_{kn,-1}, \\
h^{\Theta\Phi}_{kn,0} &= -\frac{1}{2}\cos\Theta\sin 2\Phi h^{xx}_{kn,0} + \frac{1}{2}\cos\Theta\sin 2\Phi h^{yy}_{kn,0}, \\
h^{\Theta\Phi}_{kn,1} &= \sin\Theta\sin\Phi h^{xz}_{kn,1} - \sin\Theta\cos\Phi h^{yz}_{kn,1}, \\
h^{\Theta\Theta}_{kn,2} &= -\frac{1}{2}\cos\Theta\sin 2\Phi h^{xx}_{kn,2} + \cos\Theta\cos 2\Phi h^{xy}_{kn,2} + \frac{1}{2}\cos\Theta\sin 2\Phi h^{yy}_{kn,2}.
\end{aligned}$$

$$\begin{aligned}
h^{\Phi\Phi} &= \sum_{knm} h^{\Phi\Phi}_{knm} e^{-(ik\Omega_\theta + n\Omega_r + m\Omega_\phi)t}, \\
h^{\Phi\Phi}_{kn,-2} &= \sin^2\Phi h^{xx}_{kn,-2} - \sin 2\Phi h^{xy}_{kn,-2} + \cos^2\Phi h^{yy}_{kn,-2}, \\
h^{\Phi\Phi}_{kn,0} &= \sin^2\Phi h^{xx}_{kn,0} + \cos^2\Phi h^{yy}_{kn,0}, \\
h^{\Phi\Phi}_{kn,2} &= \sin^2\Phi h^{xx}_{kn,2} - \sin 2\Phi h^{xy}_{kn,2} + \cos^2\Phi h^{yy}_{kn,2}.
\end{aligned}$$

Then the harmonics decomposition of h^+ and $h^\times$ can be expressed as:

$$h^+ = \frac{1}{2}(h^{\Theta\Theta} - h^{\Phi\Phi}) = \sum_{knm} h^+_{knm} e^{-(ik\Omega_\theta + n\Omega_r + m\Omega_\phi)t}, \tag{4.94}$$

with the expansion coefficients:

$$\begin{aligned}
h^+_{kn,-2} &= \frac{1}{2}h^{\Theta\Theta}_{kn,-2} - \frac{1}{2}h^{\Phi\Phi}_{kn,-2}, \\
h^+_{kn,-1} &= \frac{1}{2}h^{\Theta\Theta}_{kn,-1}, \\
h^+_{kn,0} &= \frac{1}{2}h^{\Theta\Theta}_{kn,0} - \frac{1}{2}h^{\Phi\Phi}_{kn,0},
\end{aligned}$$

$$
\begin{aligned}
h^{+}_{kn,1} &= \frac{1}{2} h^{\Theta\Theta}_{kn,1}, \\
h^{+}_{kn,2} &= \frac{1}{2} h^{\Theta\Theta}_{kn,2} - \frac{1}{2} h^{\Phi\Phi}_{kn,2}.
\end{aligned}
\tag{4.95}
$$

$$
h^{\times} = h^{\Theta\Phi} = \sum_{knm} h^{\times}_{knm} e^{-(ik\Omega_\theta + n\Omega_r + m\Omega_\phi)t},
\tag{4.96}
$$

with the expansion coefficients:

$$
\begin{aligned}
h^{\times}_{kn,-2} &= h^{\Theta\Phi}_{kn,-2}, \\
h^{\times}_{kn,-1} &= h^{\Theta\Phi}_{kn,-1}, \\
h^{\times}_{kn,0} &= h^{\Theta\Phi}_{kn,0}, \\
h^{\times}_{kn,1} &= h^{\Theta\Phi}_{kn,1}, \\
h^{\times}_{kn,2} &= h^{\Theta\Phi}_{kn,2}.
\end{aligned}
\tag{4.97}
$$

The next step is to analyse the properties of these expansion coefficients.

4.4.3 Truncation of the harmonic decomposition

Till now, the Fourier decomposition of the waveform is constructed from a pure geodesic and represented as infinite sums, for practical purpose, we need to truncate these expansions. First we generate the waveform in time domain, we integrate the geodesic equation numerically, and substitute the obtained numerical orbital motion into the quadrupole formula to construct the waveform s. Then we analyse the s in frequency domain by using the harmonic analysis, and construct $h = \sum_{knm} h_{knm}$ as the harmonic decomposition. By calculating the overlap $\mathcal{O}_{knm}$ between each individual harmonic h_{knm} and s, we can sort these harmonics in order of their strength: the harmonics with higher overlap are dominant and give larger contribution to the SNR. Starting form the strongest one, we add to the

sum subdominant and compute overlap with s, we truncate the accumulation when $\mathcal{O}$ exceed the threshold, 99%. We call the chosen harmonics the dominant harmonics, and we can collect their indices together as an index table which will be very convenient for understanding the harmonics structure of the signal.

Different sets of parameters have lead to the waveforms with different dominant harmonics. We conducted extensive Monte carlo simulations for different parts of the parameters space, and identified dominant harmonics (index table). We find that e has the largest influence on the table as expected, orbits with high e require more dominant harmonics for accurate description.

4.4.4 Exploring the harmonic content as function of the parameters

In this subsection, we describe how does the harmonic index table (k, n, m) of the dominant harmonics change according to the various sets of parameters $(e, p, \iota, a, \Theta, \Phi)$, here indices k, n, m correspond to θ, r, ϕ respectively.

For the convenience of explanation, we start with the parameters $\{e = 0.1, p = 7, \iota = 5^\circ, a = 0.9, \Theta = \pi/3, \Phi = \pi/4\}$. We find 4 dominant harmonics given as the following indices table,

$$\begin{array}{rcccc} m: & 2 & 2 & 2 & 1 \\ k: & 0 & 0 & 0 & 1 \\ n: & 0 & 1 & -1 & 0 \end{array} \tag{4.98}$$

Each column represents a harmonics with corresponding indices.

In the simulations, I have investigated rather large range of each parameter: eccentricity $e = \{0.01 \sim 0.99\}$ spin $a = \{0.01 \sim 0.99\}$, semilatus rectum $p = \{5 \sim 20\}$, and the inclination angle $\iota = \{5^\circ \sim 175^\circ\}$. I found that, the eccentricity has the main effect on the dominant harmonics, the influence of the other three parameters depends on the particular value of the eccentricity. In the following subsections, I will chose few examples to illustrate these results.

4.4.4.1 The influence of eccentricity

We fix the value of $\{p = 7, \iota = 5°, a = 0.9, \Theta = \pi/3, \Phi = \pi/4\}$, and vary the value of eccentricity in the range $\{0.3, 0.5\}$. The dominant harmonics are given as following tables:

$$\begin{array}{rcccccccc} m: & 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 \\ (e=0.3)\quad k: & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ n: & 1 & 0 & -1 & 2 & 3 & 0 & 1 & -1 \end{array} \tag{4.99}$$

$$\begin{array}{rcccccccc} m: & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ (e=0.5)\quad k: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ n: & -1 & 2 & 3 & 1 & 4 & 5 & 6 & -2 \end{array} \tag{4.100}$$

We find that the dominant harmonics, even the most dominant ones, can change due to the different e. In order to confirm this, we have performed more tests, where we keep $\{p, \iota, a, \Theta, \Phi\}$ to a new set of values, and vary e. The results show the similar behavior.

4.4.4.2 The influence of a, p,ι

We keep e, p, a unchanged, and vary ι, the result shows that, in the case of lower eccentricity $e : 0 - 0.4$, the dominant harmonics do not change. But, in the case of higher eccentricity ($e \geq 0.5$), the situation is more complicate. Let us give few examples:

$$\begin{array}{rcccccccc} m: & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ (e=0.7, \iota=10°)\quad k: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ n: & 7 & 6 & -1 & 8 & 0 & 5 & 9 & 10 \end{array} \tag{4.101}$$

$$\begin{array}{rcccccccc} m: & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ (e=0.7, \iota=20°)\quad k: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ n: & 7 & 6 & 8 & 0 & -1 & 5 & 9 & 10 \end{array} \tag{4.102}$$

$$
\begin{array}{cccccccccc}
 & m: & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\
(e=0.7, \iota=30^\circ) & k: & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & n: & 7 & 8 & 6 & 0 & 9 & 5 & -1 & 10
\end{array}
\tag{4.103}
$$

We can see that, these eight dominant harmonics are the same for three different values of ι. But, the harmonic $(2, 0, -1)$ which is the third strongest in the case of $\iota = 10^\circ$ become the fifth in the case of $\iota = 20^\circ$, and seventh in the case of $\iota = 30^\circ$. On the contrary, the harmonics $(2, 0, 8)$ which is the fourth in the case of $\iota = 10^\circ$ become the third in the case of $\iota = 20^\circ$, and the second in the case of $\iota = 30^\circ$.

That is to say, the variation of ι can change the strength of the dominant harmonics: the weak harmonics can become strong and vice verse, but the table itself does not change, just the order of harmonics. The similar situation is observed for varying p and spin.

4.4.4.3 The influence of Θ, Φ

As for the parameters Θ and Φ (position of observer as seen from the source), they have very week influence on the dominant harmonics. From our results, the case of $\Theta = \pi/2$ can bring additional one or two very weak harmonics in $h_\times$ compare to the other value of Θ. What is interesting is that, the parameter Φ does not change the strength of the harmonics: $|h^{+,\times}_{knm}|$, and this can be explain analytically.

From the expression of h^{ij}_{knm}, we have the following properties:

$$
\begin{aligned}
h^{xx}_{kn,-2} &= -h^{yy}_{kn,-2}, \\
h^{xx}_{kn,-2} &= -ih^{xy}_{kn,-2}, \\
h^{xz}_{kn,-1} &= ih^{yz}_{kn,-1}, \\
h^{xx}_{kn,0} &= h^{yy}_{kn0}.
\end{aligned}
\tag{4.104}
$$

Using these expressions, we can rewrite the h^+_{knm} as:

$$\begin{aligned}
h^+_{kn,-2} &= \frac{1}{2}(\cos^2\Theta\cos^2\Phi h^{xx}_{kn,-2} + \cos^2\Theta\sin 2\Phi h^{xy}_{kn,-2} + \cos^2\Theta\sin^2\Phi h^{yy}_{kn,-2}) \\
&\quad -\frac{1}{2}(\sin^2\Phi h^{xx}_{kn,-2} - \sin 2\Phi h^{xy}_{kn,-2} + \cos^2\Phi h^{yy}_{kn,-2}) \\
&\quad = \frac{1}{2}[\cos^2\Theta(\cos 2\Phi - i\sin 2\phi)h^{xx}_{kn,-2} - (-\cos 2\Phi + i\sin 2\Phi)h^{xx}_{kn,-2}] \\
&\quad = \frac{1}{2}(\cos^2\Theta + 1)(\cos 2\Phi - i\sin 2\Phi)h^{xx}_{kn,-2}, \qquad (4.105) \\
h^+_{kn,-1} &= \frac{1}{2}h^{\Theta\Theta}_{kn,-1} \\
&\quad = -\frac{1}{2}(\sin 2\Theta)(\cos\Phi - i\sin\Phi)h^{xx}_{kn,-1} \qquad (4.106) \\
h^+_{kn,0} &= \frac{1}{2}(\cos^2\Theta\cos^2\Phi h^{xx}_{kn0} + \cos^2\Theta\sin^2\Phi h^{yy}_{kn0} + \sin^2\Theta h^{zz}_{kn0}), \\
&\quad -\frac{1}{2}(\sin^2\Phi h^{xx}_{kn0} + \cos^2\Phi h^{yy}_{kn0}) \\
&\quad = \frac{1}{2}((\cos^2\Theta - 1)h^{xx}_{kn0} + \sin^2\Theta h^{zz}_{kn0}), \qquad (4.107)
\end{aligned}$$

this means the amplitude of the harmonics $|h^+_{knm}|$ does not depend on Φ, and similar for the other two terms $h^+_{kn,1}$ and h^+_{kn2}. This holds also for $h^\times$.

4.4.5 The class of the dominant harmonics

From the Monte Carlo simulation on the different parameter sets, we find that, the frequency of the dominant harmonic can be described by the following combined table:

$$n\Omega_r + k\Omega_\theta + m\Omega_\phi = \begin{cases} I: & 2\Omega_\phi + n_1\Omega_r \\ II: & \Omega_\phi + \Omega_\theta + n_2\Omega_r \\ III: & 2\Omega_\theta + n_3\Omega_r \\ IV: & -\Omega_\phi + \Omega_\theta + n_4\Omega_r \\ V: & -\Omega_\phi + 3\Omega_\theta + n_5\Omega_r \\ VI: & -2\Omega_\phi + 4\Omega_\theta + n_6\Omega_r \end{cases}, \qquad (4.108)$$

where $n_{i=1,2,\cdots 6} = \{-4, -3, \cdots, 10\}$ are the integers. Among these six classes of harmonics, I, II, III are the stronger ones comparing to IV, V, VI. Generally, we can divide the eccentricity into two group, one is low eccentricity with $e \leq 0.4$, the other is the high eccentricity with $e > 0.4$. For the low eccentricity, the dominant harmonics are belong to the classes $\{I, II, IV, V, VI\}$. For the high eccentricity, the dominant harmonics belong to the classes $\{II, III, IV, V, VI\}$. Although, the harmonics from classes $\{IV, V, VI\}$ appear in both cases, their strength is much weaker than the harmonics in classes $\{I, II, III\}$, and for weak signal, are probably not detected.

4.4.6 The evolution of the dominant harmonics

So far, we have considered instantaneous dominant harmonics based on geodesic motion. In order to analyse the whole waveform, we need to understand how these dominant harmonics change during the evolution due to the radiation reaction.

Starting from a set of physical parameters $\{e, p, \iota, a, \mu\}|_{t=0}$ as initial state, we can calculate the dominant harmonics of NK waveform for the geodesic at initial time: $\{h_{knm}\}$. We collect the index combination (k, n, m) of all the dominant harmonics together as a index table. We need to find out whether this index table change during the evolution. We choose several instants of time spread in the three month as our test points.

From the discussion about changing of harmonics index for the different set of physical parameters, we know that, the change in harmonics index is mostly related to the eccentricity e. As mentioned above, we do not consider the zoom-whirl orbit near the separatrix where e can grow. If the initial eccentricity below 0.4, it can only decrease. Therefore , starting from $e < 0.4$, we can always keep the same set of dominant harmonics throughout the whole evolution. Hence, we can concluded that, in the case of low eccentricity, the harmonics index table constructed by the initial value of the physical parameters nearly does not change. In the subsection 4.4.5 with classification of the dominant harmonics, we have found that, for the low eccentricity, the dominant harmonics indices are mostly belong to the classes $\{I, II, VI, V, VI\}$.

If the initial eccentricity is higher than 0.4, the situation is a little more

complicated. From the classification of the dominant harmonics, we know that, the strongest harmonics for high eccentricity belong to the class of II, III, and the weaker harmonics are class VI,V,VI. This mean that, the harmonics index table which is constructed by the initial stage is only suitable when eccentricity is still larger than 0.4. When the eccentricity decrease to less than 0.4, the new dominant harmonics of class I should be also taken into account. However, if we enlarge the index table to include all these dominant harmonics at the initial stage, during the whole evolution, the index table can be kept the same. In the beginning of the evolution, the harmonics of class I are unnecessary, they can become stronger during the future evolution. Enlarging the index table in the beginning makes it valid throughout the evolution.

To sum up, the harmonics index table which is given by the initial stage can be enough to construct the dominant harmonics during the whole evolution. For the chosen three month data, the amplitude of the dominant harmonics do not change appreciably (less than 10% or 20%). So based on the that table, and the slowly changing of the amplitude of $|h_{knm}|$ the NK waveform for the whole evolution can be expressed as the harmonics decomposition which constructed from the index table as an approximation

$$h = \sum_{knm \in \mathbf{indextable}} h_{knm}(t) e^{i(k\Omega_\theta + n\Omega_r + m\Omega_\phi)}. \tag{4.109}$$

It is already mentioned that, the index table 4.108 is based on the quadrople formula. If we also take into account the current quadrople and mass octupole, the procedure is the same, the index table of the dominant harmonics will change. Based on the paper Babak *et al.* [2007], we believe that, there must be more weak dominant harmonics in the case of quadrupole-octupole formula (4.6) which should be added if necessary, and enhance the strength of classes $\{VI, V, VI\}$.

The index table of dominant harmonics helps us to compute efficiently the GW signal from EMRI without much loss in accuracy.

4.5 Phenomenological waveform family

In this section we introduce the phenomenological waveform family of the EMRI system. The first idea of using phenomenological waveform was suggested for detecting GWs from inspiraling BHs in the LIGO band Buonanno *et al.* [2005]. The main goal was to construct phenomenological waveform which fits different PN models of the signal. By using the stationary phase approximation on the PN formalism, the phenomenological waveforms are expressed in Fourier domain and given by a single analytical expression:

$$h(f) = A_{eff}(f)e^{i\varphi_{eff}(f)} \tag{4.110}$$

The parameter including the amplitude parameters, and the phase parameters which can be found in Buonanno *et al.* [2005].

Another example of phenomenological waveform is a templates family which resembles inspiral, merge and ring down stage of the black hole coalescence. It combines the PN waveforms with results from NR simulations, and its implementation in the search for signatures of gravitational waves in the data of ground-based interferometers is given in Ajith *et al.* [2007].

As we have shown in previous section, the full waveform construction for EMRI is very complicated. From the point of view of data analysis, even the NK waveforms are not fast enough to be generated for employing in the matched filtering. We want to construct a phenomenological waveform family for EMRIs by using the following assumptions about GW signal:

1. The orbital motion can be described by six slowly changing quantities; three time-dependent initial phases governed by the conservative part of the self force; three fundamental time-dependent frequencies governed by the radiative part of the self force.

2. The waveform is represented by harmonics of three slowly evolving phases with slowly changing intrinsic amplitudes:

$$h(t) = \sum_{k,n,m} h_{knm}(t)$$

$$
\begin{aligned}
&= Re\left(\sum_{k,n,m} A_{knm}(t) e^{i\varphi_{knm}(t)}\right) \\
&= Re\left(\sum_{k,n,m} A_{knm}(t) e^{i(n\varphi_r + k\varphi_\theta + m\varphi_\phi)}\right), \qquad (4.111)
\end{aligned}
$$

where $\varphi_r, \varphi_\theta, \varphi_\varphi$ are the phases corresponding to the three fundamental motions. Here we omitted the tensorial spatial indices for simplicity.

The point 1 basically says that the motion is described by a slow drift from one geodesic to another. The initial phases correspond to the initial position of a compact object on a given geodesic and the orbital frequencies are functions of the energy, azimuthal component of the orbital momentum and Carter constant. The slow drift ensures that phases φ_{knm} are slowly changing functions of time.

Both amplitudes and the phases are slowly varying functions of time, thus we can make the Taylor expansion:

$$
\begin{aligned}
\varphi_r(t) &= \varphi_r(t_0) + \omega_r(t_0)(t-t_0) + \frac{1}{2}\dot{\omega}_r(t-t_0)^2 + ... \\
&= \varphi_r(t_0) + 2\pi f_r(t_0)(t-t_0) + \pi \dot{f}_r(t-t_0)^2 + ..., \qquad (4.112)
\end{aligned}
$$

$$
\begin{aligned}
\varphi_\theta(t) &= \varphi_\theta(t_0) + \omega_\theta(t_0)(t-t_0) + \frac{1}{2}\dot{\omega}_\theta(t-t_0)^2 + ... \\
&= \varphi_\theta(t_0) + 2\pi f_\theta(t_0)(t-t_0) + \pi \dot{f}_\theta(t-t_0)^2 + ..., \qquad (4.113)
\end{aligned}
$$

$$
\begin{aligned}
\varphi_\varphi(t) &= \varphi_\phi(t_0) + \omega_\phi(t_0)(t-t_0) + \frac{1}{2}\dot{\omega}_\phi(t-t_0)^2 + ... \\
&= \varphi_\phi(t_0) + 2\pi f_\phi(t_0)(t-t_0) + \pi \dot{f}_\phi(t-t_0)^2 + ..., \qquad (4.114)
\end{aligned}
$$

$$
A_{knm}(t) = A_{knm}(t_0) + \dot{A}_{knm}(t_0)(t-t_0) + ... \qquad (4.115)
$$

Since the amplitudes A_{lmn} are even smoother than the phase over extended period of time, and because the detection techniques are more sensitive to mismatch in the phase than in the amplitude, we can neglect the time evolution in the amplitudes and treat all of them as constants. It is a very good assumption over three months of the simulated data which we analyze in this paper. Although, it

is quite short compare to the LISA observation, three month data is still sufficient for testing our method. We can recover the parameters of a signal very well even with these short data. As for the data which is longer than three months, we can analyse the data piecewise following the same techniques. In the reality, the SNR of the signal may be lower than in our simulated data, three month might not be enough for definite detection, and we must use longer segment. In this case, we need to modify phenomenological waveform by extending the Taylor expansion of the phase in equations 4.115 to higher orders.

As for the phase expansion, in the numerical calculation, the Taylor expansion of the phase should be truncated at some order. The termination is determined by calculating the so-called fitting factor (FF) Apostolatos [1995]; Jaranowski & Krolak [2005] for the different orders of polynomial approximations of the phase to check the fidelity of the PW. Numerical results show that the Taylor expansion for three months data, up to $\ddot{f}$ order, gives the FF around 0.9, and up to $\frac{d^3f}{dt^3}$ order the FF is larger than 0.999. So it is sufficient to expand the phase to $\frac{d^3f}{dt^3}$ order. This is the phenomenological waveform family which we propose to analyze EMRI signal.

To summarize, the phenomenological waveform is a summation of individual harmonics with constant (or linear in time) amplitudes and polynomial (in time) phases. In practice, we need to truncate this infinite expansion. The results of harmonics analysis help us to determine how many dominant harmonics are required, which harmonic index are needed to construct the templates.

4.5.1 $\mathcal{F}$-statistic in the case of phenomenological waveform

As mentioned in the section 2.3, the GW waveform in the frame attached to LISA (or LISA-like space based observatory) is given by

$$h_I = F_I^+(t)h^+(t) + F_I^\times(t)h^\times(t), \quad h_{II} = F_{II}^+(t)h^+(t) + F_{II}^\times(t)h^\times(t), \quad (4.116)$$

where the response function $F^{+}_{I,II}$ and $F^{\times}_{I,II}$ are already given in equation (2.25), with the harmonics decomposition of waveforms $h^{+,\times}$:

$$h^{+} = \sum_{knm} A^{+}_{knm} \cos\varphi_{knm}(t) = \sum_{knm} A^{+}_{knm}(t) \cos(\varphi^{0}_{knm} + \tilde{\varphi}_{knm}(t)), \quad (4.117)$$

$$h^{\times} = \sum_{knm} A^{\times}_{knm} \sin\varphi_{knm}(t) = \sum_{knm} A^{\times}_{knm}(t) \sin(\varphi^{0}_{knm} + \tilde{\varphi}_{knm}(t)). \quad (4.118)$$

$$(4.119)$$

The waveform in the detector I can be written as:

$$h_I = \sum_{knm} a^{(1)}_{knm} \hat{h}^{I(1)}_{knm}(t) + a^{(2)}_{knm} \hat{h}^{I(2)}_{knm}(t) + a^{(3)}_{knm} \hat{h}^{I(3)}_{knm}(t) + a^{(4)}_{knm} \hat{h}^{I(4)}_{knm}(t), \quad (4.120)$$

where

$$\begin{aligned}
a^{(1)}_{knm} &= A^{+}_{knm} \cos(2\psi) \cos\varphi^{0}_{knm} + A^{\times}_{knm} \sin(2\psi) \sin\varphi^{0}_{knm}, \\
a^{(2)}_{knm} &= -A^{+}_{knm} \sin(2\psi) \cos\varphi^{0}_{knm} + A^{\times}_{knm} \cos(2\psi) \sin\varphi^{0}_{knm}, \\
a^{(3)}_{knm} &= -A^{+}_{knm} \cos(2\psi) \sin\varphi^{0}_{knm} + A^{\times}_{knm} \sin(2\psi) \cos\varphi^{0}_{knm}, \\
a^{(4)}_{knm} &= A^{+}_{knm} \sin(2\psi) \sin\varphi^{0}_{knm} + A^{\times}_{knm} \cos(2\psi) \cos\varphi^{0}_{knm},
\end{aligned} \quad (4.121)$$

$$\begin{aligned}
\hat{h}^{I(1)}_{knm}(t) &= D^{+}_{I}(t) \cos\tilde{\varphi}_{knm}(t), \\
\hat{h}^{I(2)}_{knm}(t) &= D^{\times}_{I}(t) \cos\tilde{\varphi}_{knm}(t), \\
\hat{h}^{I(3)}_{knm}(t) &= D^{+}_{I}(t) \sin\tilde{\varphi}_{knm}(t), \\
\hat{h}^{I(4)}_{knm}(t) &= D^{\times}_{I}(t) \sin\tilde{\varphi}_{knm}(t).
\end{aligned} \quad (4.122)$$

Similar expression we have for detector II, there, we should use $D^{+}_{II}, D^{\times}_{II}$ instead of $D^{+}_{I}, D^{\times}_{I}$.

The amplitude parameters $\hat{a}^{\mu}_{knm}$ depend only on $(A_{+}, A_{\times}, \varphi^{0}_{knm}, \psi)$, which

are usually called extrinsic parameters, while $\hat{h}^{\mu}_{knm}(t), \hat{h}^{\mu}_{knm}(t)$ are functions of $(\Theta, \Phi, f, \frac{df}{dt}, \frac{d^2f}{dt^2}, \frac{d^3f}{dt^3})$, which are usually called intrinsic parameters.

We have found that over three months we can consider all harmonics as being completely independent with virtually zero overlap between them, $< h_{knm}|h_{k'n'm'} >= \delta_{kk'}\delta_{nn'}\delta_{mm'}$. Denote the measured data with noise corresponding to h_I, h_{II} by s_I, s_{II}. The total likelihood for the two detectors I, II is:

$$\begin{aligned}[\log \Lambda] &= [\log \lambda_I] + [\log \lambda_{II}] \\ &= (s_I|h_I) - \frac{1}{2}(h_I|h_I) + (s_{II}|h_{II}) - \frac{1}{2}(h_{II}|h_{II}). \end{aligned} \tag{4.123}$$

Following the method of $\mathcal{F}$-statistic which we have introduced in the section 2.4, we can maximized the total log likelihood over the extrinsic parameters a^j_{knm} to obtain the total $\mathcal{F}$-statistic as:

$$\mathcal{F} = \sum_{knm} \sum_{i,j=1}^{4} \frac{1}{2}(s^{I(i)}_{knm} + s^{II(i)}_{knm})[(M^{I}_{knm} + M^{II}_{knm})^{-1}]^{ij}(s^{I(j)}_{knm} + s^{II(j)}_{knm}), \tag{4.124}$$

where arrays $s^{I,II(i)}_{knm}$ and the matrixes $M^{I,II}_{knm}$ are defined as:

$$\begin{aligned} s^{I,II(i)}_{knm} &= (s^{I,II}|\hat{h}^{I,II(i)}_{knm}(t)), \\ (M^{I,II}_{knm})^{ij} &= (\hat{h}^{I,II(i)}_{knm}|\hat{h}^{I,II(j)}_{knm}). \end{aligned} \tag{4.125}$$

4.6 Data analysis with phenomenological waveform

In this section, we use the PW described above as the template to analyse the signal of EMRI in two 3 months simulated data sets: one is constructed without noise and the other is a noisy data with the total SNR of the signal 50. We use the same GW signal (based on NK model) in both cases. We have taken the following parameters for the EMRI: the mass of the MBH $M = 10^6 M_\odot$, the mass of the compact object (stellar mass BH) $m = 10 M_\odot$, the initial orbital eccentricity $e = 0.4$, the semi-latus rectum $p = 8M$, the inclination angle $\iota = \pi/9$, the spin

of the MBH $a = 0.9M$, the sky position of the source $(\Theta, \Phi) = (\pi/4, \pi/4)$, the polarization angle $\psi = 0$. The noiseless case is used to avoid any possible bias in the final result due to stochastic nature of the noise, and we want to asses only possible restrictions of our search technique and PW family. Then, we apply the same search method to the same GW signal buried in the noise, which would justify its effectiveness in practice.

In the practical implementation of NK, the harmonics decomposition should be chosen from the most dominant ones, this is realized by using the index table which is given in the subsection 4.4.5:

$$h^{+,\times} = \sum_{knm \in [index\ \ table]} h^{+,\times}_{knm} e^{i\varphi_{knm}(t)}. \tag{4.126}$$

We need to search the intrinsic parameter space: sky location of the source, and the phase parameters to find the set of parameters which maximizes the log likelihood. To simplify our problem, we treat the sky location of the source as known, and search only for the intrinsic parameters $f_{knm}, \dot{f}_{knm}, \ddot{f}_{knm}, \cdots$.

4.6.1 Search for Dominant harmonics

Generally, the best estimation of the intrinsic parameters corresponds to the global maximum in the likelihood surface. In order to find the global maximum, we use the standard MCMC to explore the intrinsic parameter space. However, due to the complexity of the waveform, there is a large amount of local maxima. Although these local maxima do not match all the intrinsic parameter, some of them (with the high value of likelihood), must contain useful information about the dominant harmonics. Harmonics of a template correspond to local maxima, match parts of the signal for some period of time. By using the information stored in the local maxima, we can reconstruct signal's dominant harmonics during the evolution. In the time-frequency plot, we can mark these parts of dominant harmonics as a time frequency track within the matched span of time and frequencies evolution. We are trying to reconstruct the signal in time-frequency plane by these patches from local maxima.

4.6.1.1 Review of MCMC and simulated annealing

In this subsection, we review the basic knowledge of MCMC and simulated annealing which are used in our approach based on Cornish & Crowder [2005]; Gamerman [1997]. Given the data s which consists of the noise and signal parameterized by $\vec{\lambda} = \{\lambda_1, \lambda_2, \cdots, \}$, the parameter estimation can be obtained trough the posterior distribution, $p(\lambda|s)$ by using the Bayesian estimator:

$$\lambda^i_{\mathbf{B}}(s) = \int \lambda^i p(\lambda|s) d\lambda. \tag{4.127}$$

By Bayes theorem, the posterior distribution is related to the prior distribution $\pi(\vec{\lambda})$ and the likelihood $p(s|\vec{\lambda})$ as

$$p(\vec{\lambda}|s) = \frac{\pi(\vec{\lambda}) p(s|\vec{\lambda})}{\int \pi(\vec{\lambda}') p(s|\vec{\lambda}') d\vec{\lambda}'}. \tag{4.128}$$

The numerical calculation of this integral can be achieved by using the Monte carlo methods, MCMC is nothing but Monte Carlo integration with the Markov Chain. MCMC can be used for sampling from probability distributions of $p(\lambda|s)$ based on constructing a Markov chain that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the distribution $p(\lambda|s)$.

Denote the equilibrium distribution as $p(\vec{\lambda})$. A Markov Chain, $\vec{\lambda}^0, \vec{\lambda}^1, \cdots, \vec{\lambda}^n$, can be thought of as a discrete stochastic process in which knowledge of the past, $\vec{\lambda}^0, \vec{\lambda}^1, \cdots, \vec{\lambda}^{n-1}$, is irrelevant to the future, $\vec{\lambda}^{(n+1)}$, given knowledge of the present, $\vec{\lambda}^{(n)}$. In other words, the conditional probability that chain transitions to state $\vec{y}$ satisfies:

$$P(\vec{\lambda}^{(n+1)} = \vec{y}|\vec{\lambda}^{(0)}, \vec{\lambda}^{(1)}, \cdots, \vec{\lambda}^{(n)}) = P(\vec{\lambda}_{n+1} = \vec{y}|\vec{\lambda}^{(n)}). \tag{4.129}$$

This means there exist a transition probability $P(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)})$ depending only on the current state $\vec{\lambda}^k$. If the Markov chain satisfies the following balance equa-

tion

$$p(\vec{\lambda}^{(k)})P(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)}) = p(\vec{\lambda}^{(k+1)})P(\vec{\lambda}^{(k)}|\vec{\lambda}^{(k+1)}), \tag{4.130}$$

Then, it samples from the equilibrium distribution $p(\vec{\lambda})$ after some time (burn in stage). The distribution of the parameters $p(\vec{\lambda})$ can be estimated by using the Markov chain samples.

For constructing a Markov chain with desired properties, the Metropolis-Hastings (MH) scheme would be the most general one. The MH algorithm replace the unknown transition probability $P(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)})$ with three quantities that can be calculated: the likelihood $p(s|\vec{\lambda})$; the prior $\pi(\vec{\lambda})$; and the proposal distribution $q(\cdot|\vec{\lambda}^{(k)})$. To do this, the chain starts with a randomly chosen point in the parameters space $\vec{\lambda}^0$. The proposal distribution $q(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)})$ is picked to guide transition from the current state $\vec{\lambda}^k$ to the a candidate state $\vec{\lambda}^{k+1}$. This transition is either accepted or rejected by calculating the acceptance probability (Metropolis-Hastings ratio) defined by the following formula:

$$\alpha(\vec{\lambda}^{(k)}, \vec{\lambda}^{(k+1)}) = \min(1, \frac{\pi(\vec{\lambda}^{(k+1)})p(s|\vec{\lambda}^{(k+1)})q(\vec{\lambda}^{(k)}|\vec{\lambda}^{(k+1)})}{\pi(\vec{\lambda}^{(k)})p(s|\vec{\lambda}^{(k)})q(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)})}), \tag{4.131}$$

This acceptance rule corresponds to the transition probability

$$P(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)}) = q(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)})\alpha(\vec{\lambda}^{(k)}, \vec{\lambda}^{(k+1)}) \tag{4.132}$$

The above transition probability implies that the Markov chain satisfies the detailed balance equation.

The best performance is achieved if the proposal probability resembles closely the target distribution $p(\vec{\lambda})$. Without prior knowledge about the probability distribution around the true parameter location, it is naturally to choose it as a multivariate normal distribution centered at the present point $\vec{\lambda}^{(k)}$. Due to the multimodal structure of the likelihood surface, a single multivariate normal distribution cannot describe the probability density over the entire parameters

space but only a very small region around $\vec{\lambda}^{(k)}$. Thus the Markov chain runs from one local maximum to another.

In the early stage of the search, we want the chain to explore as large area of the parameters space as possible. But, due to the multimodal structure of the likelihood function surface, some local maxima are quite strong, and the Markov chain could be trapped in one of these peaky local maxima. This means that the chain fall into a local area very quickly, does not explore various maxima in a reasonable time. Thus, we use the simulated annealing to soften the likelihood function, and make it easier for the chain to move between the modes Cornish & Porter [2007]. The basic idea of simulated annealing is introducing a parameter β which stand for the inverse temperature to control the acceptance probability at k step:

$$\alpha = \min(1, \beta \frac{\pi(\vec{\lambda}^{(k+1)}) p(s|\vec{\lambda}^{(k+1)}) q(\vec{\lambda}^{(k)}|\vec{\lambda}^{(k+1)})}{\pi(\vec{\lambda}^{(k)}) p(s|\vec{\lambda}^{(k)}) q(\vec{\lambda}^{(k+1)}|\vec{\lambda}^{(k)})}). \tag{4.133}$$

In our approach, we choose a decreasing function $\beta = \exp(\frac{N-k}{N})$, where N is the length of the chain and the priors are chosen as the uniform distribution. In the beginning, β is small, the temperature is high. Even if the chain move to a point with small likelihood, the acceptance probability can be still large. This means that, it accepts this point at the new stage of the chain frequently. In other words, the likelihood function surface becomes smother, the chain can move much easier exploring a large area. As the chain process ends, the search falls into some areas which contain the local maxima. When β is large, the temperature is low, the chain remains around the local maxima it has found. At the end of the search, the chain performs local exploration.

We use 5000 chains searching in the parameters space, to find a large number of local maxima, one of them could be the global one. In the next subsection, we will discuss how to use these local maxima to extract the useful information about the GW signal.

4.6.1.2 Analysis of MCMC results

In this subsection we extract information from the local maxima detected by MCMC search. We first focus on the noiseless data to explain our methods, then modify it a bit and apply it to the noisy data.

Although the best estimation of the intrinsic parameters generally corresponds to global maximum of the log likelihood, the strong secondary local maxima can also be used to dig out the useful information about the dominant harmonics of the true signal. The dominant harmonics of PW might match parts of the harmonics of the true signal and results in local maxima. If we can extract these part of matched true harmonics, we can get information about the signal, multiple secondary maxima might provide us enough clues to recover all the dominant harmonics of the signal.

To attain this goal, we collect the local maxima: distinct points in the parameters space with significant likelihood. We compute the accumulative (in frequency or time) likelihood for harmonics of PW which give largest contribution to the likelihood. Let us define precisely the accumulative likelihood. During the whole observation time $[0, T_{\rm ob}]$, the frequency of each harmonics changes from $f_{\rm min}$ to $f_{\rm max}$, the full log likelihood is defined as before

$$\log \Lambda = (h|s) - \frac{1}{2}(h|h), \tag{4.134}$$

with the inner product $(a|b)$

$$(a|b) = 4\mathfrak{R} \int_{f_{\rm min}}^{f_{\rm max}} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} df. \tag{4.135}$$

If we focus on the part of the observation time $[0, t]$, the change of frequency during this time period is $[f_{\rm min}, f(t)]$. The cumulative log likelihood is defined by replacing $f_{\rm max}$ by $f(t)$ in the inner product calculation

$$\log \Lambda_{\rm [acc]}(t) = (h|s)_{\rm acc}(t) - \frac{1}{2}(h|h)_{\rm acc}(t), \tag{4.136}$$

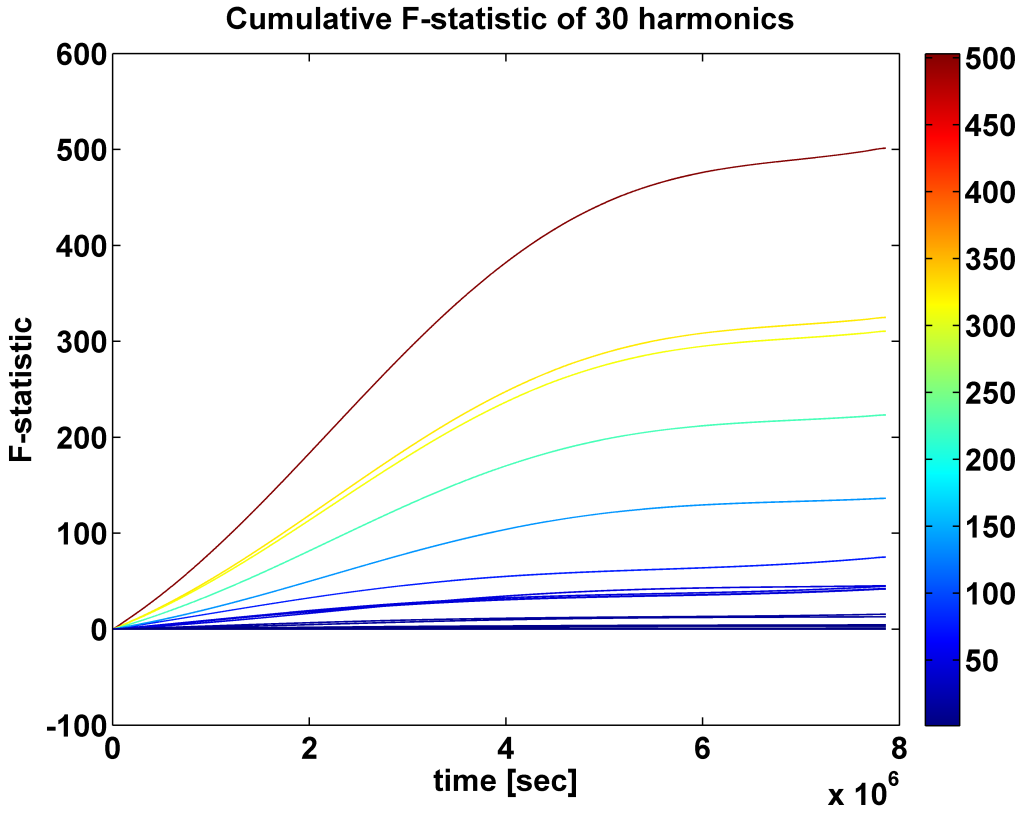

Figure 4.1: The cumulative F-statistic of 30 dominant harmonics with true parameters without noise. Since there is no noise, the F-statistic is not normalized.

where the inner product $(a, b)_{\rm acc}$

$$(a|b)_{\rm acc}(t) = 4\Re \int_{f_{\rm min}}^{f(t)} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} df. \tag{4.137}$$

With the increase of t, the term $\log \Lambda_{[\rm acc]}$ goes from 0 to $\log \Lambda$. The similar implementation can be also used in the maximized likelihood (or $\mathcal{F}$-statistic), that will be the cumulative maximized likelihood or cumulative $\mathcal{F}$-statistic. Take the true signal as an example, in the noseless case, the picture 4.1 show the cumulative $\mathcal{F}$-statistic for 30 dominant harmonics.

We compute accumulative likelihood only for those harmonics which give significant contribution to the total $\mathcal{F}$-statistic. These PW harmonics detect a part of signal with frequency and its derivatives close to the true parameters, but not

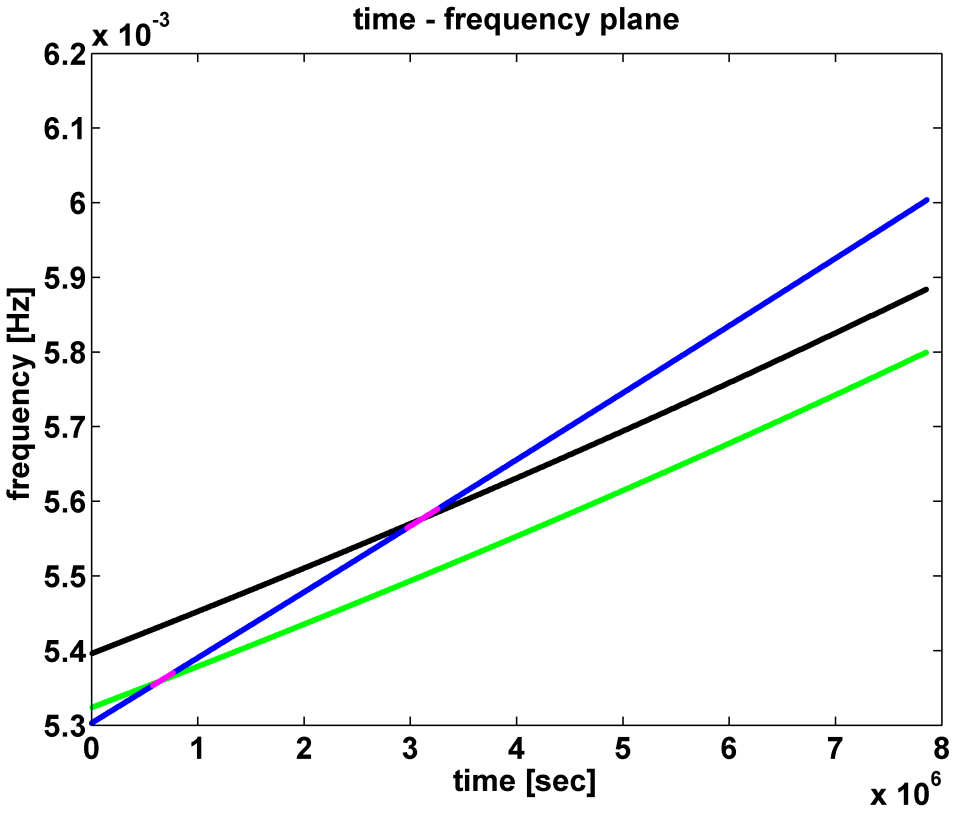

Figure 4.2: Time-frequency plot of harmonics. The black and green tracks are two strong harmonics of the EMRI signal (black being stronger). The blue track corresponds to a harmonic of PW that accumulates a significant $\mathcal{F}$-statistic. It intersects the true harmonics at the pink segments, those correspond to times of increase of $\mathcal{F}$-statistic, see 4.3.

necessarily exact. Single PW harmonic can match different parts of the signal at different instance, it is showed in Figure 4.2. where, black and green are two strong true harmonics of the signal (black being stronger), and the blue line is a harmonic of PW. The pink are regions where the PW matches frequency of the true harmonics for a short period of time. The corresponding cumulative $\mathcal{F}$-statistic of the blue harmonics is shown in the Figure 4.3. From this figure, we can see that there are two positive jumps in the accumulation of the $\mathcal{F}$-statistic which correspond to two instances of intersection. The positive slope in the cumulative $\mathcal{F}$-statistic (if it happens over a significant duration) corresponds to the part of the frequency and time where the particular harmonic of PW matches

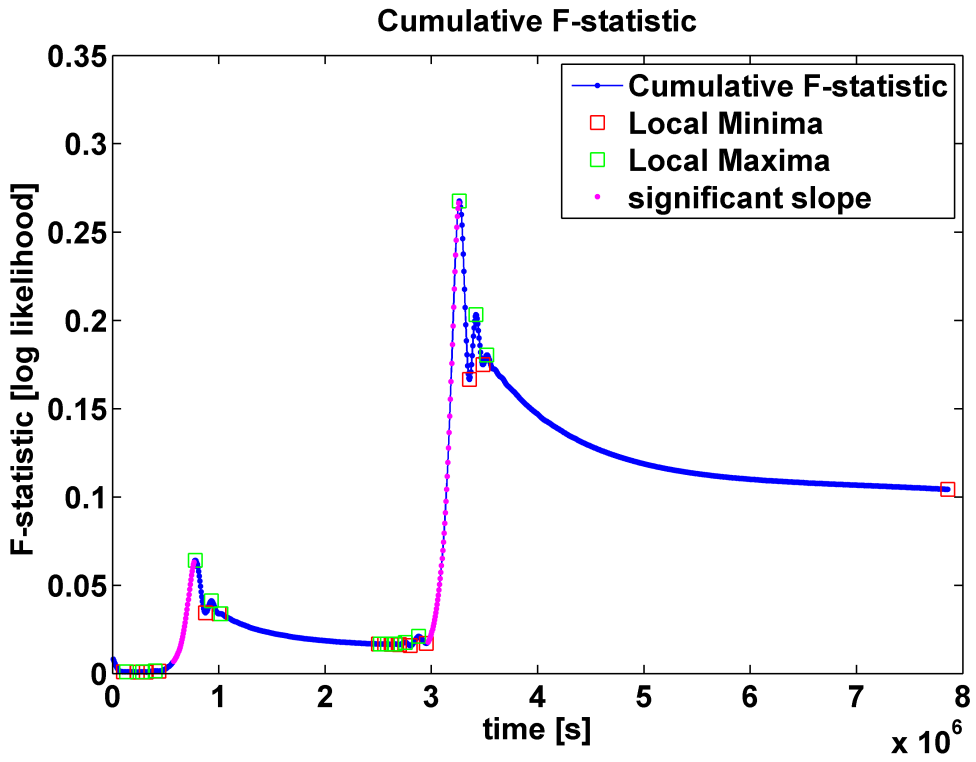

Figure 4.3: Cumulative $\mathcal{F}$-statistic corresponding to the situation depicted in Fig. 4.2. The green and red squares mark the maxima and minima of the curve, thus distinguishing between the increasing and the decreasing slopes.

parts of some true harmonics of a signal. In practice, we utilize two criteria for selecting the significant $\mathcal{F}$-statistic accumulation: (i) the slope must be larger than a certain threshold; (ii) the accumulation time must be longer than a certain period. In our search we have made the following choice for these two criteria. In the case of noiseless data, we require the slope to be larger than one-tenth of the largest slope of the cumulative $\mathcal{F}$-statistic of that trial harmonic, and the cumulative time (over which we observe steep positive slope) to be longer than three days.

We collect all patches where the harmonics match parts of the signal together and display them on the time-frequency plane in the Figure 4.4, we can identify by eye 13 strong harmonics. Among these 13 time-frequency tracks, there are several

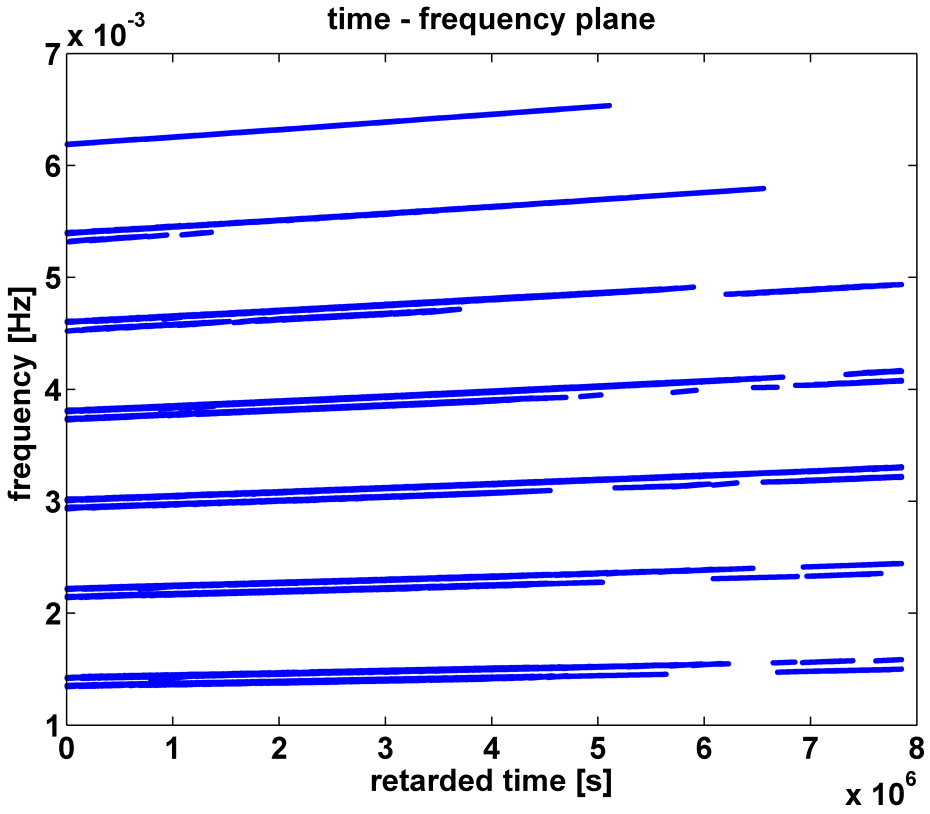

Figure 4.4: Time-frequency plot of all patches corresponding to strong accumulation of F-statistic. We can identify parts of frequency tracks of 13 EMRI harmonics. Each track in this plot has a finite width coming from different solutions of MCMC search which have different precision of matching the signal.

pairs of nearly overlapping tracks. The value of three fundamental frequencies f_r, f_θ, f_ϕ ($\sim 10^{-3}$ Hz) are close to each other, if two harmonics have the same value in the sum of their indices $k + n + m$, their frequencies will have small difference and they appear as nearly overlapped in the time frequency plot. For example, the pair of tracks in the bottom have the indices $\{m = 2, k = 0, n = -1\}, \{m = 1, k = 0, n = 0\}$ respectively, hence, the difference is $f_\phi - f_r$ which is around 10^{-5} Hz.

Although the weaker harmonics are lost, the strong ones contain enough information about the system evolution which can be used to recover the physical parameters. At each instance, there are many found patches from the different solutions of MCMC which give us a spread in the frequency for a given har-

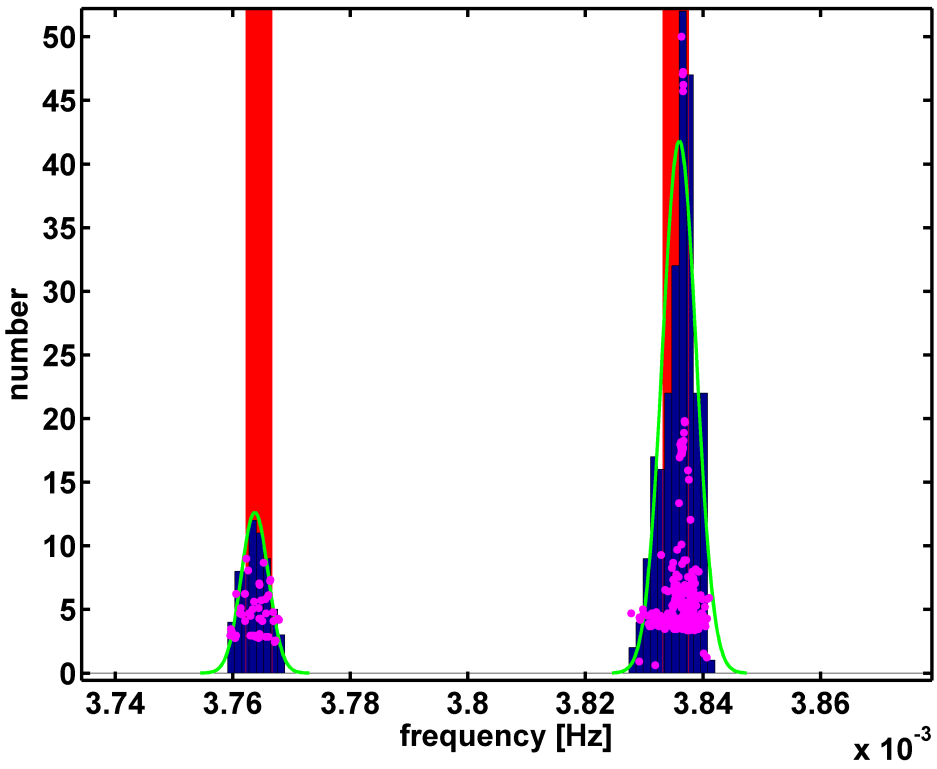

Figure 4.5: Zoom at two harmonics at a specific instance of time. The red stems denote the frequencies of the true harmonics of a signal, while the blue histogram shows the detected frequencies at this instant. The green curves are the Gaussian fit to the frequency data with re-scaled amplitudes. The vertical axis of pink points indicates the relative time over which we have observed strong accumulation of $\mathcal{F}$-statistic for each solution.

monic. This is because various solutions from MCMC search matched a given true harmonic of the signal with different precision, however we expect that the distribution of found frequencies at each instant of time will be centered close to the true frequency of signal's harmonic. As an example we show the distribution of found frequencies at a particular instance of time for two harmonics in the Figure 4.5.

In this plot, the blue histogram stands for the number of found solutions around that frequency at a given instance, the smooth green curve gives the Gaussian fit, and the frequencies of two harmonics from the signal are presented

in red. As mentioned above, different solutions of MCMC results vary in precision of matching the signal, and we can use accumulation time as the measure of goodness of match. The relative accumulation time of different solutions shown as pink points in the Figure 4.5. First, one can see that Gaussian fit lies on the top of the true frequency, and second, that the distribution of pink points is similar to the blue histogram, so either can be taken to characterize the found harmonics of a signal. Similarly we can do at each instance of time for all found tracks in the time-frequency plane.

For the noiseless search, we have picked uniformly distributed instances $\{t_j\}_{j=1,2,...,K}$ and made a Gaussian fit around each harmonic. We identify the mean value ρ_i of the Gaussian fit as the most likely frequency of a signal's harmonics at that instance, and we identify the variance σ_i as an error in our estimation of a frequency.

In the case of noisy data, the techniques are roughly the same as in the noiseless case with a minor alteration. First, we choose the local maxima with SNR greater than 4.5. Then, we select the significantly steep slopes of the cumulative F-statistic depending on three requirements: (i) the maximum $\mathcal{F}$-statistic along the cumulative $\mathcal{F}$-statistic curve should be larger than 50, (ii) the slope is larger than $4 \times 10^{-6} s^{-1}$, (iii). the monotonic increasing duration is longer than about a week. These values should be fixed based on the desired false alarm probability. Instead we have varied these parameters until the time-frequency tracks are clearly visible "by eye". This implies that our search could be further optimized. From this selection, we identified 5 strong harmonics in the noisy case. After that the procedure is similar to the noiseless case.

The number of found time-frequency tracks is less than the number of harmonics given in the index table. What if we had found more time-frequency than the number of expected (dominant) harmonics? By using the harmonics from the index table to truncate the harmonics expansion, the waveform template could recover 99.99% SNR. If the number of found tracks is less than the number in the index table, it means we have not find all the weak harmonics in the index table, probably, they are too weak to be detected. There could be two cases when the number of found tracks is larger than the number in the index table. One possibility is that there is an unusually stronger signal, so that even the weak

harmonics which are not included in the index table could be detected. Or, a false detection, in other words, we have found some fake (noise generated) signal. In both cases, we can tune the parameters which are used to control the selection of the dominant harmonics to decrease the number of found tracks, until the number of found tracks does not exceed the number of harmonics from the index table. For example, we could raise the slope of accumulative likelihood curves, or increase the threshold of SNR.

Here I must acknowledge that the Figures 4.1, 4.2, 4.3, 4.4, 4.5 are made by my colleague Yan Wang, and are used here with his permission.

4.6.2 Physical parameter searching

As explained in the section of the waveform modeling 4.3, given the initial value of the physical parameter of $\{e, p, \iota, a, \mu, M\}$, the numerical kludge waveform is determined by solving the differential equations for the orbital parameter evolution. Now, we have found some dominant harmonics of the signal as the tracks in the time-frequency plane. Next we need to identify the physical parameters of the signal whose harmonics fit found time-frequency tracks. The index table can guide us to recover the physical parameters of the signal from the found frequency tracks of the dominant harmonics. When searching on the physical parameter space, we need to choose the indices to generate the tracks of the template, and compare to the found tracks. The index table could tell us which dominant harmonic might contribute to the found frequency tracks, and which track need to be generated. What we need to do is to simulate the frequency tracks by using the indices from the index table, then calculate their degree of matching with the found frequency tracks to find the best fit. And our results show that it is very efficient to recover the physical parameters by using this index table. This procedure can be accomplished by utilizing the genetic algorithm on the physical parameter space, which will be described clearly in section 4.6.2.2.

4.6.2.1 Define the quantity to assess the goodness of fit

We have found sets of tracks corresponding to the strong harmonics of the signal, but we do not know the corresponding index of these harmonics. So far, we have

just tracks in the time-frequency plane. For each e, p, ι, a, μ we can construct the tracks of dominant harmonics of the signal. The idea is to find the set of e, p, ι, a, μ which corresponding harmonics match the best to the found tracks. This procedure is similar to described in Gair *et al.* [2008]. We need to define some quantity which can measure goodness of fit .

When we use the physical parameters to construct the dominant harmonics, all combinations of indices from the index table have the equivalent possibility of matching the found tracks. With the initial physical parameters $\{e, p, \iota, a, \mu\}$, we can calculate three fundamental frequencies f_r, f_θ, f_ϕ at this moment of time. At each instance t_j, for each harmonic h_{knm} from the index table, we compute the frequency $f^{(j)}_{knm}(t) = kf^{(j)}_\theta(t) + nf^{(j)}_r(t) + mf^{j}_\phi(t)$, then compare to the mean value $\rho^{(j)}_i$ which we have discussed in the previous subsection, and chose the closest one $(f^{(j)}_i)$. Each harmonic can be chosen only once. We compute the χ^2 (weighted) goodness of fit test:

$$\chi^2 = \sum_{j=1}^{K}\sum_{i=1}^{N}\frac{(f^{(j)}_i - \rho^{(j)}_i)^2}{\sigma^{(j)}_i}, \tag{4.138}$$

here, index $j = 1, 2, \cdots, K$ stand for the representative time instance t_j, index $i = 1, 2, \cdots, N$ stand for the harmonics index.

Besides the χ^2 test defined above, we can also introduce another statistics:

$$W = \sum_{j=1}^{K}\sum_{i=1}^{N}(\max_{\{l=1,2,\cdots,P\}}\{G^{(j)}_i(f^{(j)}_l, \rho^{(j)}_i, \sigma^{(j)}_i)\}). \tag{4.139}$$

Here, again, the index i, j have the same meaning as in the χ^2 test, index $l = 1, 2, \cdots P$ stand for the harmonics from the index table which we have introduced in the subsection 4.4. $G^{(j)}_i(f, \rho^{(j)}_i, \sigma^{(j)}_i)$ are the Gaussian fit functions at the time instance t_j for the harmonics characterized by $\rho^{(j)}_i$ and $\sigma^{(j)}_i$ as the mean value and variance. After getting the frequency $f^{(j)}_l$ for each harmonics from the index table, we substitute it in to all the Gaussian fit function $G^{(j)}_i(f^{(j)}_l, \rho^{(j)}_i, \sigma^{(j)}_i)$, and choose the maximum as the closest between $\rho^{(j)}_i$ and $f^{(j)}_l$. If the $f^{(j)}_l$ are not

close to any $\rho_i^{(j)}$, its value of Gaussian function is very small. The best match corresponds to maxima of W.

Both of these two quantities can be used to define the goodness of fit of each set of physical parameters, and they are applied in our search. Next, we choose search algorithm which maximizing W or minimizing χ^2.

4.6.2.2 Searching algorithm in time frequency plane

Having the statistic χ^2 or W, the searching on the physical parameters space is just minimizing χ^2 or maximizing W. To accomplish this, we use two stochastic search algorithms. One is Genetic Algorithm and the other is the Particle Swarm Optimization method.

In chapter 3, we have explained in details GA as a powerful stochastic search method. Let us give a short summary of a specific implementation of GA utilized in the current search. Each parameter set $\{e, p, \iota, a, \mu\}$ is called an organism, individual parameters are called the genes of this organism and the set of organisms at k-th search iteration step is called k-th generation. As before, we evolve generations according to the prescribed rules called "parents selection", "breading" and "mutation".

We use χ^2 value as a measure of fitness for each organism (smaller value is better). In each generation we use the roulette method with the selection probability proportional to the fitness of each organism. For breeding we have used the one random point crossover rule. The probability mutation rate is monotonically decreasing function of the generation number: we have started with high probability of mutation to explore a large part of the parameter space and decrease it gradually as organisms converge to a particular part of the parameter space. We have used "children" and "parents" sorted in the fitness to make a new generation: we use 50% of the best organisms. We automatically achieve the "elitism" in a way that the best χ^2 value is never increasing from one generation to the next.

We use the multi-step method to accelerate the search. In each step we evolve the colony for 500 generations as described above, but each new step uses the last generation of the previous step as the initial state. We have started

evolution in the first step with completely random distribution of the organisms. The evolution of the colony at each step finishes with a very small mutation probability and with organisms confined to a quite small volume of the parameter space. The consequent search steps ensure that the found solution is a robust solution with respect to the increase of the mutation probability which disperses organisms forcing them to explore the parameter space for presence of a solution with a better fitness. This helps to avoid being trapped in the local minima. The termination condition is the stability of the best solution over several steps of the search.

In the case of noiseless data, we have found that the final results are sensitive to the sampling of time-frequency tracks (t_j). Using only first two month of data provide with smoother surface on χ^2 , and it is easier to find the global minima than the case of three month. Some distinct local minima in the case of three month long data are not local minima any more in the case of two month, or not so prominent. The explanation is in the detector's motion and antenna beam function. From the Figure 4.4, we can see that, in the first two month, the found dominant harmonics are much more prominent comparing the third month (the last month). For first two month, the detector had better orientation with respect to the source (large response) than in the last month, amplitudes are suppressed in the last month, and it leads to less accurately determination. As for how to choose the suitable representative time t_j, we can use the information of time of positive accumulation of the log likelihood in each found patch. At each moment of time, besides the frequency of each found patch, both the accumulative time and the strength of harmonics is stored and can be used. Both the longer accumulative time and larger strength means high probability of good matching to the true harmonics for the found patch, and it is more likely to contain accurate information of the true signal in this patch. Take the accumulative time as an example: at each moment of time, we add the accumulative time of all the found patch together, and use the resulting sum as a measure of quality. Then we choose sampling of time-frequency plane based on these sums. We should also take care that the points do represent the whole time span.

Besides the genetic algorithm, we also use the Particle Swarm Optimization (PSO) to do the search. PSO is a population based stochastic optimization

technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behavior of bird flocking or fish schooling Kennedy & Eberhart [1995]. PSO simulates how do the bird flock find their food. Suppose, there is only a piece of bread in a searching area. All the birds have no idea where the bread is, but, they know how far away from their present position to the food. So, the most simple and efficient tactics is searing nearby the bird which is nearest to the food.

As an evolutionary computation technique, PSO shares many similarities with GA. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solution is called particles. PSO is initialized with a group of random particle and then search for the optima by updating generation. In every iteration, each particle is updated by following two "best value". The first one is the individual best position that each single particle has achieved so far during its search history, denoted as p_I. Another "best " value is the global best position that is obtained so far by all the particles in the whole swarm, denoted as p_g. After finding these two best positions, each particle updates its velocity and position according to the following equations:

$$\begin{aligned} \text{Velocity}_I(new) &= \zeta\text{Velocity}_I(\text{Present}) + c_1\eta_1[p_I(\text{Present}) - \text{Position}I(\text{Present})] \\ &\quad +c_2\eta_2[p_g(\text{Present}) - \text{Position}_I(\text{Present})], \\ \text{Position}_I(new) &= \text{Position}_I(\text{Present}) + \text{Velocity}_I(\text{New}), \end{aligned} \tag{4.140}$$

where ζ is called the inertia weight, c_1 and c_2 are called acceleration constants, and η_1,η_2 are random numbers drawn independently at each step from the uniform distribution on [0,1]. The first term simply moves a particle along a straight line, while the remaining two terms are sources of acceleration, one pulling it towards its individual best position and another pulls it towards the global best position. The last two effects are combined with random weights η_1 and η_2. The random deflections and inertial motion allow a particle to explore the parameter space, while the attractive pulls of p_I and p_g give preferred deflections. With a dynamic inertia weight that decreases in time, the attractive pull eventually wins over.

Base on this basic principle, we use a specific implementation which is de-

Table 4.1: Recovered parameters of EMRi against actual parameters used in simulating data sets.

description	$e(t_0)$	$p(t_0)$	$\iota(t_0)$	a	μ
True parameters	0.4	8.0	0.349	0.9	10^{-5}
Recovered parameters (with noise)	0.395	8.029	0.342	0.891	9.79×10^{-6}
Recovered parameters (no noise)	0.402	7.991	0.360	0.901	1.002×10^{-5}

scribed in Wang & Mohanty [2010] to our work, the final results are comparable to that in GA method .

The recovered parameters are given in the table 4.1. From this table, we can see that, we have recovered the physical parameters with good precision (fractional errors are less than $2-3\%$).

4.7 Summary

In this chapter, we have introduced phenomenological family of waveforms (PW) for detecting EMRI signals in the data from the LISA-like observatory. The template is constructed out of independent (over the time interval we have applied our analysis) harmonics of slowly evolving three orbital frequencies. We have neglected the amplitude evolution and presented the phase as a Taylor series up to the third derivative of frequency. We have restricted in our analysis to the case of monotonically increasing frequencies, this condition will break only close to the plunge. The number of harmonics and range of indices were taken from the analysis of dominant harmonics of our model signal, though we have found at the end that the search only weakly depends on the number of used harmonics (only through the accumulated total SNR, which should be sufficient to claim detection).

Constructed phenomenological templates allows us to search for EMRI signals in a model independent way. This way we avoid complexity of accurate modeling the orbital evolution and gravitational waveform during the search. In addition PW covers also all possible small deviations of the background spacetime from the Kerr solution which would influence the signal's phase and could lead even to loss of the signal if the template assumes pure Kerr background geometry.

We have used MCMC based search to find a large number of local maxima of the likelihood surface. We were not that lucky to find the global maximum. We have analyzed the found solutions by means of cumulative over the time F-statistic and identified the patches of the signal which were match by templates. As a result we have constructed time-frequency map of (parts of) signal's harmonics. Each track could be characterized by the best guess and the error bar at each instance of time (by fitting Gaussian profile into found at that time frequencies around each track). The next step is to assume a model for the binary orbital evolution, and map found time-frequency picture to the strongest harmonics of a signal. In other words we want to find the physical parameters of the binary system which strong GW harmonics could leave the found imprint. We do that by conducting a search using particle swarm optimization techniques and, independently, genetic algorithm. We have used weighted chi-square goodness of fit test to choose the best matching harmonics of the signal. We have assumed the same model as was used in the simulated data, and the recovered parameters are within $2-3\%$ of the true values.

We want to make few final remarks. (i) The found time-frequency tracks of the GW signal from EMRI did not assume any particular model. The mapping these tracks to the physical parameters could be done in post processing using several models. We have chosen on purpose rather short (3 month) duration of the data. The search procedure could be repeated for each three months and then one can check consistency of a given model or further improve accuracy in the recovered parameters (if model gives consistent parameters of the system across different data segments). This could be a powerful method to search deviations from "Kerness". (ii) In the mapping of the time-frequency tracks to the physical parameters of the binary, we have only weakly used information about the strength of each track/harmonic. We have found that the information stored in the frequency evolution is sufficient to recover parameters of EMRI, however, the use, in addition, information about the strength of the recovered harmonics and harmonics of the modeled GW signal could give us additional confidence in the result and/or distinguish between several solution if ambiguity happens. (iv) Mapping from the found time-frequency tracks onto the physical parameters might turn out to be the most computationally intensive task, however one

might use the information about the strength and a number of found harmonics to restrict a volume of the searched parameter space. In addition, to perform mapping we require mainly to compute the orbital evolution, not the full waveform, however, it is then important to know which harmonics are the strongest for a given parameter set. (iv) In the future work we intend to include the sky location and the MBH mass into the search and investigate the possibility to differentiate between different models of EMRIs based on the results of MCMC search with PW (as discussed in (i)).

The aim of this work is to introduce the phenomenological waveform family and show in practice how to use it to detect the GW from EMRIs. Due to lack of time, I have not considered the cases of weak signal and several simultaneous signals . For the weak signal, it is more difficult to identify the dominant harmonics. This might result in smaller number of found time-frequency tracks. We will need to use longer data segments (6 months) or even the whole data set to accumulate sufficient SNR. In the PW model, we will need to extend the Taylor series to the higher orders both in the phase and amplitude evolution. This will lead to a larger parameter space and potentially bring more complexities in the search for time-frequency tracks. Once we have found the dominant harmonics of the signal, the method described in this thesis can still be used to recover the physical parameters. With larger parameter space and longer data segment, it is more likely to identify the time-frequency track more accurately since we require to match signal for longer time to accumulate sufficient SNR. Therefore, the found time-frequency tracks should have smaller deviation from the true signal. At the same time, the low SNR means larger influence of the noise on the parameter estimation. We can also utilize the strength of the found dominant harmonics which we did not use in our present work to refine the parameter estimation. The SNR of the found tracks could tell us which tracks correspond to the strong harmonics. Therefore, we can correlate only strongest harmonics (from index table) with corresponding time-frequency tracks. As for the case of multiple signals, we need to answer the question: how to identify time frequency track with each signal. One possibility is to identify them when some frequency tracks cross each other. Alternatively we can calculate the rate of frequency change, $\dot{f}$, for each found track, and the different slop ($\dot{f}$) could help us to classify these tracks and

attribute them to different signals. All in all, this might be a big challenge for the highly overlap signals, and should be addressed in the future.

Chapter 5

Final Summary

The detailed summaries were given after each chapter, here, I give a brief conclusion of my thesis. In the work of searching the SMBH binaries in mock LISA data, we apply an extended genetic algorithm based on the property of the signal and the response function of the detector. The results show that our method is very efficient in finding the signals and recovering the physical parameters with a high precision. In the future, we can apply our method to the more sophisticated simulated data which contains GW signals from multiple sources of different kind all together, this includes SMBH, EMRIs, Galactic white dwarf binaries, bursts from cosmic cusps.

In the second half of my thesis, we introduce a new family of the phenomenological template, and utilize theses templates to detect the signal in the simulated data by using the MCMC technique. After that, by assuming a particular EMRIs model, we apply two stochastic algorithm: PSO and GA, to estimate the physical parameters of the binary with high accuracy. In the future work, we need to extend our method by including more search parameters, investigate the possibility to distinguish between different models of EMRIs based on the results of MCMC search with PW, and set a practical scheme for using PW in the mapping spacetime analysis, and check "no hair " theorem.

References

Abramovici, A., Althouse, W.E., Drever, R.W., Gursel, Y., Kawamura, S. *et al.* (1992). LIGO: The Laser interferometer gravitational wave observatory. *Science*, **256**, 325–333. 5, 26

Ajith, P., Babak, S., Chen, Y., Hewitson, M., Krishnan, B. *et al.* (2007). Phenomenological template family for black-hole coalescence waveforms. *Class.Quant.Grav.*, **24**, S689–S700. 125

Amaro-Seoane, P., Gair, J.R., Freitag, M., Coleman Miller, M., Mandel, I. *et al.* (2007). Astrophysics, detection and science applications of intermediate- and extreme mass-ratio inspirals. *Class.Quant.Grav.*, **24**, R113–R169. 14, 16

Amaro-Seoane, P., Schutz, B. & Sopuerta, C.F. (2010). Gravitational Waves Notes, Issue 4 : 'A Roadmap to Fundamental Physics from LISA EMRI Observations'. 14, 91

Amaro-Seoane, P., Aoudia, S., Babak, S., Binetruy, P., Berti, E. *et al.* (2012a). eLISA: Astrophysics and cosmology in the millihertz regime. 7, 9, 13, 17

Amaro-Seoane, P., Aoudia, S., Babak, S., Binetruy, P., Berti, E. *et al.* (2012b). Low-frequency gravitational-wave science with eLISA/NGO. *Class.Quant.Grav.*, **29**, 124016. 91

Apostolatos, T. (1995). Search templates for gravitational waves from precessing, inspiraling binaries. *Phys.Rev.*, **D52**, 605–620. 127

Apostolatos, T.A., Cutler, C., Sussman, G.J. & Thorne, K.S. (1994). Spin induced orbital precession and its modulation of the gravitational wave forms from merging binaries. *Phys.Rev.*, **D49**, 6274–6297. 48

Arnaud, K., Auger, G., Babak, S., Baker, J., Benacquista, M. *et al.* (2007a). Report on the first round of the Mock LISA data challenges. *Class.Quant.Grav.*, **24**, S529–S540. 10, 26, 27

Arnaud, K., Babak, S., Baker, J., Benacquista, M., Cornish, N. *et al.* (2007b). An Overview of the second round of the Mock LISA Data Challenges. *Class.Quant.Grav.*, **24**, S551–S564. 27

Arnaud, K.A., Babak, S., Baker, J.G., Benacquista, M.J., Cornish, N.J. *et al.* (2006a). The Mock LISA Data Challenges: An overview. *AIP Conf.Proc.*, **873**, 619–624. 27

Arnaud, K.A. *et al.* (2006b). A How-To for the Mock LISA Data Challenges. *AIP Conf.Proc.*, **873**, 625–632. 10, 27

Babak, S. (2008). Building a stochastic template bank for detecting massive black hole binaries. *Class.Quant.Grav.*, **25**, 195011. 56, 71

Babak, S., Fang, H., Gair, J.R., Glampedakis, K. & Hughes, S.A. (2007). 'Kludge' gravitational waveforms for a test-body orbiting a Kerr black hole. *Phys.Rev.*, **D75**, 024005. 16, 29, 40, 89, 93, 94, 96, 104, 105, 109, 124

Babak, S., Baker, J.G., Benacquista, M.J., Cornish, N.J., Crowder, J. *et al.* (2008a). The Mock LISA Data Challenges: From Challenge 1B to Challenge 3. *Class.Quant.Grav.*, **25**, 184026. 10, 26, 27, 28, 46, 61

Babak, S., Gair, J.R. & Porter, E.K. (2009). An algorithm for detection of extreme mass ratio inspirals in lisa data. *Class.Quant.Grav.*, **26**, 135004. 92

Babak, S., Gair, J.R., Petiteau, A. & Sesana, A. (2011). Fundamental physics and cosmology with LISA. *Class.Quant.Grav.*, **28**, 114001. 91

Babak, S. *et al.* (2008b). Report on the second Mock LISA Data Challenge. *Class.Quant.Grav.*, **25**, 114037. 10, 26, 27

Babak, S. *et al.* (2010). The Mock LISA Data Challenges: From Challenge 3 to Challenge 4. *Class.Quant.Grav.*, **27**, 084009. 10, 13, 26, 27

Baker, J.G., McWilliams, S.T., van Meter, J.R., Centrella, J., Choi, D.I. *et al.* (2007). Binary black hole late inspiral: Simulations for gravitational wave observations. *Phys.Rev.*, **D75**, 124024. 13

Ballo, L., Braito, V., Della Ceca, R., Maraschi, L., Tavecchio, F. *et al.* (2004). Arp 299: A Second merging system with two active nuclei? *Astrophys.J.*, **600**, 634–639. 11

Barack, L. (2000). Selfforce on a scalar particle in spherically symmetric space-time via mode sum regularization: Radial trajectories. *Phys.Rev.*, **D62**, 084027. 16

Barack, L. (2001). Gravitational selfforce by mode sum regularization. *Phys.Rev.*, **D64**, 084021. 16

Barack, L. (2009). Gravitational self force in extreme mass-ratio inspirals. *Class.Quant.Grav.*, **26**, 213001. 15, 106

Barack, L. & Burko, L.M. (2000). Radiation reaction force on a particle plunging into a black hole. *Phys.Rev.*, **D62**, 084040. 16

Barack, L. & Cutler, C. (2004). LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy. *Phys.Rev.*, **D69**, 082005. 90, 91

Barack, L. & Cutler, C. (2007). Using LISA EMRI sources to test off-Kerr deviations in the geometry of massive black holes. *Phys.Rev.*, **D75**, 042003. 90

Barack, L. & Ori, A. (2000). Mode sum regularization approach for the self-force in black hole space-time. *Phys.Rev.*, **D61**, 061502. 16

Barack, L. & Ori, A. (2003). Gravitational selfforce on a particle orbiting a Kerr black hole. *Phys.Rev.Lett.*, **90**, 111101. 16

Barack, L. & Sago, N. (2007). Gravitational self force on a particle in circular orbit around a Schwarzschild black hole. *Phys.Rev.*, **D75**, 064021. 16

Barack, L. & Sago, N. (2010). Gravitational self-force on a particle in eccentric orbit around a Schwarzschild black hole. *Phys.Rev.*, **D81**, 084021. 16, 29, 91, 106, 111

Barack, L. & Sago, N. (2011). Beyond the geodesic approximation: conservative effects of the gravitational self-force in eccentric orbits around a Schwarzschild black hole. *Phys.Rev.*, **D83**, 084023. 16, 106

Barack, L., Mino, Y., Nakano, H., Ori, A. & Sasaki, M. (2002). Calculating the gravitational selfforce in Schwarzschild space-time. *Phys.Rev.Lett.*, **88**, 091101. 16

Barack, L., Golbourn, D.A. & Sago, N. (2007). m-Mode Regularization Scheme for the Self Force in Kerr Spacetime. *Phys.Rev.*, **D76**, 124036. 16

Barack, L., Damour, T. & Sago, N. (2010). Precession effect of the gravitational self-force in a schwarzschild spacetime and the effective one-body formalism. *Phys. Rev. D*, **82**, 084036. 16

Barausse, E. & Rezzolla, L. (2008). The influence of the hydrodynamic drag from an accretion torus on extreme mass-ratio inspirals. *Phys. Rev. D*, **77**, 104027. 16

Bardeen, J.M., Press, W.H. & Teukolsky, S.A. (1972). Rotating black holes: Locally nonrotating frames, energy extraction, and scalar synchrotron radiation. *Astrophys.J.*, **178**, 347. 97

Bekenstein, J.D. (1973). Gravitational-Radiation Recoil and Runaway Black Holes. *Astrophys.J.*, **183**, 657–664. 95

Bender, P. & Hils, D. (1997). Confusion noise level due to galactic and extragalactic binaries. *Class.Quant.Grav.*, **14**, 1439–1444. 10

Berentzen, I., Preto, M., Berczik, P., Merritt, D. & Spurzem, R. (2008). Post-Newtonian simulations of super-massive black hole binaries in galactic nuclei. 11

Berti, E., Cardoso, V. & Will, C.M. (2006). On gravitational-wave spectroscopy of massive black holes with the space interferometer lisa. *Phys. Rev. D*, **73**, 064030. 13

Bianchi, S., Chiaberge, M., Piconcelli, E., Guainazzi, M. & Matt, G. (2008). Chandra unveils a binary Active Galactic Nucleus in Mrk463. 11

Blanchet, L., Damour, T. & Iyer, B.R. (1995). Gravitational waves from inspiralling compact binaries: Energy loss and wave form to second postNewtonian order. *Phys.Rev.*, **D51**, 5360. 96

Blaut, A., Babak, S. & Krolak, A. (2010). Mock LISA Data Challenge for the galactic white dwarf binaries. *Phys.Rev.*, **D81**, 063008. 10

Bogdanovic, T., Reynolds, C.S. & Miller, M.C. (2007). Alignment of the spins of supermassive black holes prior to merger. *Astrophys.J.Lett.*. 11, 12

Bradaschia, C., Calloni, E., Cobal-Grassmann, M., Del Fabbro, R., Di Virgilio, A. *et al.* (1990). Virgo: A Ground based interferometric antenna for gravitational wave detection above 10-Hz. 5

Brenneman, L.W. & Reynolds, C.S. (2006). Constraining Black Hole Spin Via X-ray Spectroscopy. *Astrophys.J.*, **652**, 1028–1043. 12

Brown, D.A., Crowder, J., Cutler, C., Mandel, I. & Vallisneri, M. (2007). A Three-stage search for supermassive black hole binaries in LISA data. *Class.Quant.Grav.*, **24**, S595–S606. 28

Brown, W.R., Kilic, M., Hermes, J., Prieto, C.A., Kenyon, S.J. *et al.* (2011). A 12 minute Orbital Period Detached White Dwarf Eclipsing Binary. 9

Buonanno, A., Chen, Y., Pan, Y., Tagoshi, H. & Vallisneri, M. (2005). Detecting gravitational waves from precessing binaries of spinning compact objects. II. Search implementation for low-mass binaries. *Phys.Rev.*, **D72**, 084027. 125

Callegari, S., Mayer, L., Kazantzidis, S., Colpi, M., Governato, F. *et al.* (2009). Pairing of Supermassive Black Holes in unequal-mass galaxy mergers. *Astrophys.J.*, **696**, L89–L92. 11

Carlos F. Sopuerta, N.Y. (2011). New kludge scheme for the construction of approximate waveforms for extreme-mass-ratio inspirals. *Phys. Rev. D*, **84**, 124060. 16

Carlos F. Sopuerta, N.Y. (2012). Approximate waveforms for extreme-mass-ratio inspirals: The chimera scheme. *arXiv*, **astro-ph.**, 1201.5715. 16

Chandrasekhar, S. (1983). *The mathematical theory of black holes*. Oxford Univ. Press, New York. 97, 98

Comerford, J.M., Griffith, R.L., Gerke, B.F., Cooper, M.C., Newman, J.A. *et al.* (2009). A 1.75 kpc/h Separation Dual AGN at z=0.36 in the COSMOS Field. *Astrophys.J.*, **702**, L82–L86. 11

Cornish, N.J. & Crowder, J. (2005). LISA data analysis using MCMC methods. *Phys.Rev.*, **D72**, 043005. 26, 131

Cornish, N.J. & Porter, E.K. (2006). MCMC exploration of supermassive black hole binary inspirals. *Class.Quant.Grav.*, **23**, S761–S768. 26

Cornish, N.J. & Porter, E.K. (2007). The Search for supermassive black hole binaries with LISA. *Class.Quant.Grav.*, **24**, 5729–5755. 13, 26, 28, 56, 64, 66, 72, 133

Cornish, N.J. & Rubbo, L.J. (2003). The LISA response function. *Phys.Rev.*, **D67**, 022001. 32, 36, 40, 41, 50

Crowder, J. & Cornish, N.J. (2007). Extracting galactic binary signals from the first round of Mock LISA data challenges. *Class.Quant.Grav.*, **24**, S575–S586. 10

Crowder, J., Cornish, N.J. & Reddinger, L. (2006). Darwin meets Einstein: LISA data analysis using genetic algorithms. *Phys.Rev.*, **D73**, 063011. 26, 28

Cuadra, J., Armitage, P., Alexander, R. & Begelman, M. (2008). Massive black hole binary mergers within sub-pc scale gas discs. 11

Cutler, C. (1998). Angular resolution of the LISA gravitational wave detector. *Phys.Rev.*, **D57**, 7089–7102. 39, 40, 50

Cutler, C. & Thorne, K.S. (2002). An Overview of gravitational wave sources. 9, 14, 17

Cutler, C., Kennefick, D. & Poisson, E. (1994). Gravitational radiation reaction for bound motion around a Schwarzschild black hole. *Phys.Rev.*, **D50**, 3816–3835. 29, 93, 109, 110, 111

Danzman, K. *et al.* (1998). *LISA, Laser Interferometer Space Antenna, Pre-Phase A Report*. Max-Planck-Institute fur Quantenoptic, Report MPQ 233. 5, 7

Danzmann, K., Lueck, H., Ruediger, A., Schilling, R., Schrempel, M. *et al.* (1994). GEO 600: A 600-m laser interferometric gravitational wave antenna. 5

Diener, P., Vega, I., Wardell, B. & Detweiler, S. (2012). Self-consistent orbital evolution of a particle around a Schwarzschild black hole. *Phys.Rev.Lett.*, **108**, 191102. 91

Dotti, M., Montuori, C., Decarli, R., Volonteri, M., Colpi, M. *et al.* (2008). SDSSJ092712.65+294344.0: a candidate massive black hole binary. 11

Dotti, M., Ruszkowski, M., Paredi, L., Colpi, M., Volonteri, M. *et al.* (2009a). Dual black holes in merger remnants. I: linking accretion to dynamics. 11

Dotti, M., Volonteri, M., Perego, A., Colpi, M., Ruszkowski, M. *et al.* (2009b). Dual black holes in merger remnants. II: spin evolution and gravitational recoil. 12

Drasco, S. (2009). Verifying black hole orbits with gravitational spectroscopy. *Phys.Rev.*, **D79**, 104016. 89, 111, 112

Drasco, S. & Hughes, S.A. (2004). Rotating black hole orbit functionals in the frequency domain. *Phys.Rev.*, **D69**, 044015. 16, 29, 89, 99, 101, 102, 103, 104, 111, 112

Drasco, S. & Hughes, S.A. (2006). Gravitational wave snapshots of generic extreme mass ratio inspirals. *Phys.Rev.*, **D73**, 024027. 29, 89, 93, 99, 100, 109, 111, 112

Estabrook, F. & Wahlquist, H. (1975). Response of doppler spacecraft tracking to gravitational radiation. *Gen. Relativ. Gravit.*, **6**, 439. 32

Feroz, F., Gair, J.R., Hobson, M.P. & Porter, E.K. (2009). Use of the MultiNest algorithm for gravitational wave data analysis. *Class.Quant.Grav.*, **26**, 215003. 26, 28, 56, 88

Ferrarese, L. & Ford, H. (2005). Supermassive black holes in galactic nuclei: Past, present and future research. *Space Sci.Rev.*, **116**, 523–624. 11

Finn, L.S. & Chernoff, D.F. (1993). Observing binary inspiral in gravitational radiation: One interferometer. *Phys.Rev.*, **D47**, 2198–2219. 39

Freitag, M. (2001). Monte Carlo cluster simulations to determine the rate of compact star inspiraling to a central galactic black hole. 14

Gair, J.R. & Glampedakis, K. (2006). Improved approximate inspirals of test-bodies into Kerr black holes. *Phys.Rev.*, **D73**, 064037. 29, 107, 109

Gair, J.R. & Porter, E.K. (2009). Cosmic Swarms: A Search for Supermassive Black Holes in the LISA data stream with a Hybrid Evolutionary Algorithm. *Class.Quant.Grav.*, **26**, 225004. 28

Gair, J.R., Barack, L., Creighton, T., Cutler, C., Larson, S.L. *et al.* (2004). Event rate estimates for LISA extreme mass ratio capture sources. *Class.Quant.Grav.*, **21**, S1595–S1606. 14

Gair, J.R., Mandel, I. & Wen, L. (2008). Improved time-frequency analysis of extreme-mass-ratio inspiral signals in mock LISA data. *Class.Quant.Grav.*, **25**, 184031. 92, 142

Gair, J.R., Tang, C. & Volonteri, M. (2010). LISA extreme-mass-ratio inspiral events as probes of the black hole mass function. *Phys.Rev.*, **D81**, 104014. 14

Gair, J.R., Flanagan, E.E., Drasco, S., Hinderer, T. & Babak, S. (2011). Forced motion near black holes. *Phys.Rev.*, **D83**, 044037. 29, 91, 106

Gamerman, D. (1997). *Markov Chain Monte Carlo Stochastic simulation for Bayesian inference*. Chapman Hall, England. 131

Ganz, K., Hikida, W., Nakano, H., Sago, N. & Tanaka, T. (2007). Adiabatic Evolution of Three 'Constants' of Motion for Greatly Inclined Orbits in Kerr Spacetime. *Prog.Theor.Phys.*, **117**, 1041–1066. 107

G.E. Moss, R.F., L.R. Miller (1971). Photon-Noise-Limited Laser Transducer for Gravitational Antenna. *Appl. Opt*, **10**, 2495. 5

Glampedakis, K. & Kennefick, D. (2002). Zoom and whirl: Eccentric equatorial orbits around spinning black holes and their evolution under gravitational radiation reaction. *Phys.Rev.*, **D66**, 044002. 16, 93, 96, 109, 110, 111

Guainazzi, M., Piconcelli, E., Jimenez-Bailon, E. & Matt, G. (2005). The Early stage of a cosmic collision? XMM-Newton unveils two obscured AGN in the galaxy pair ESO509-IG066. *Astron.Astrophys.*, **429**, L9–L12. 11

Gurkan, M.A. & Hopman, C. (2007). Resonant relaxation near a massive black hole: the dependence on eccentricity. *Mon.Not.Roy.Astron.Soc.*, **379**, 1083–1088. 14

Harry, I., Fairhurst, S. & Sathyaprakash, B. (2008). A Hierarchical search for gravitational waves from supermassive black hole binary mergers. *Class.Quant.Grav.*, **25**, 184027. 56

Helstrom, C. (1968). *Statistical theory f signal detection, international series of monographs in electronics and instrumentation, vol 9*. Pergamon Press, Oxford, New York. 19

Hils, D. & Bender, P.L. (2000). Gravitational radiation from helium cataclysmics. *Astrophys.J.*, **537**, 334–341. 10

Hils, D., Bender, P. & Webbink, R. (1990). Gravitational radiation from the Galaxy. *Astrophys.J.*, **360**, 75–94. 10

Hinderer, T. & Flanagan, E.E. (2008). Two timescale analysis of extreme mass ratio inspirals in Kerr. I. Orbital Motion. *Phys.Rev.*, **D78**, 064028. 106

Hopman, C. & Alexander, T. (2006). Resonant Relaxation near the Massive Black Hole in the Galactic Center. *J.Phys.Conf.Ser.*, **54**, 321–327. 11, 14

Hughes, S.A. (2000). The Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational wave emission. *Phys.Rev.*, **D61**, 084004. 16, 107

Hughes, S.A. (2001). Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational wave emission. 2. Inspiral trajectories and gravitational wave forms. *Phys.Rev.*, **D64**, 064004. 16, 93, 109

Hulse, R. & Taylor, J. (1975). Discovery of a pulsar in a binary system. *Astrophys.J.*, **195**, L51–L53. 4

Hulse, R.A. (1994). The discovery of the binary pulsar. *Rev.Mod.Phys.*, **66**, 699–710. 4

Ivanov, P. (2002). On formation rate of close binaries consisting of a supermassive black hole and a white dwarf. *Mon.Not.Roy.Astron.Soc.*, **336**, 373. 14

Jaranowski, P. & Krolak, A. (2005). Gravitational-Wave Data Analysis. Formalism and Sample Applications: The Gaussian Case. *Living Rev.Rel.*, **8**, 3. 17, 22, 23, 24, 41, 127

Jaranowski, P., Krolak, A. & Schutz, B.F. (1998). Data analysis of gravitational - wave signals from spinning neutron stars. 1. The Signal and its detection. *Phys.Rev.*, **D58**, 063001. 22, 24

J.W. Armstrong, F.E. & Tinto, M. (1999). Time-delay interferometry for space-based gravitational wave searches. *Astrophys. J.*, **527**, 814–826. 36, 38

Kennedy, J. & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, **IV**, 1942–1948. 145

Kennefick, D. & Ori, A. (1996). Radiation reaction induced evolution of circular orbits of particles around Kerr black holes. *Phys.Rev.*, **D53**, 4319–4326. 107

Kesden, M., Sperhake, U. & Berti, E. (2010). Relativistic Suppression of Black Hole Recoils. *Astrophys.J.*, **715**, 1006–1011. 13

King, A., Pringle, J. & Hofmann, J. (2008). The Evolution of Black Hole Mass and Spin in Active Galactic Nuclei. 12

Komossa, S., Burwitz, V., Hasinger, G., Predehl, P., Kaastra, J.S. *et al.* (2003). Discovery of a binary agn in the ultraluminous infrared galaxy ngc 6240 using chandra. *Astrophys.J.*, **582**, L15–L20. 11

Lang, R.N. & Hughes, S.A. (2006). Measuring coalescing massive binary black holes with gravitational waves: The Impact of spin-induced precession. *Phys.Rev.*, **D74**, 122001. 12

Larson, S. (2005). Lisa: A modern astrophysical observatory. 7

LISACode (2000). Web pages of LISACode, lisa simulator. 83, 84

LISATeam (2000). Lisa: laser interferometer space antenna, a cornerstone mission for the observation of grvitatinal waves. 6, 7, 9, 17, 27, 32, 44

LISATeam (2009). Lisa data analysis status. 9, 10, 17

Maggiore, M. (2000). Gravitational wave experiments and early universe cosmology. *Phys. Rept.*, **331**, 283–367. 19

Maggiore, M. (2007). Gravitational Waves. Vol. 1: Theory and Experiments. 1, 2, 24, 44

Manca, G.M. & Vallisneri, M. (2010). Cover art: Issues in the metric-guided and metric-less placement of random and stochastic template banks. *Phys.Rev.*, **D81**, 024004. 56

Marsh, T. (2011). Double white dwarfs and LISA. *Class.Quant.Grav.*, **28**, 094019. 9

McMillan, S.L. & Portegies Zwart, S.F. (2003). The Fate of star clusters near the Galactic center 1: Analytic considerations. *Astrophys.J.*, **596**, 314–322. 11

M.E. Gertsenshtein, V.P. (1963). The galactic center massive black hole and nuclear star cluster. *JETP*, **16**, 433. 5

Messenger, C., Prix, R. & Papa, M. (2009). Random template banks and relaxed lattice coverings. *Phys.Rev.*, **D79**, 104017. 56

Miller, M.C. (2004). Probing general relativity with mergers of supermassive and intermediate-mass black holes. *Astrophys.J.*, **618**, 426–431. 14

Mino, Y. (2003). Perturbative approach to an orbital evolution around a supermassive black hole. *Phys.Rev.*, **D67**, 084027. 99, 106

Misner, C.W., Thorne, K. & Wheeler, J. (1974). Gravitation. 1, 4, 98

Murphy, K., Yaqoob, T., Dovciak, M. & Karas, V. (2009). On the Prospect of Constraining Black-Hole Spin Through X-ray Spectroscopy of Hotspots. *Astrophys.J.*, **701**, 635–641. 12

Nelemans, G. (2009). The Galactic Gravitational wave foreground. *Class.Quant.Grav.*, **26**, 094030. 10

Nelemans, G., Yungelson, L. & Portegies Zwart, S.F. (2001). The gravitational wave signal from the galactic disk population of binaries containing two compact objects. *Astron.Astrophys.*, **375**, 890–898. 10

Nelemans, G., Yungelson, L. & Portegies Zwart, S. (2004). Short-period AM CVn systems as optical, x-ray and gravitational wave sources. *Mon.Not.Roy.Astron.Soc.*, **349**, 181. 52

Nelemans, G., Wood, M., Groot, P., Anderson, S., Belczynski, K. *et al.* (2009). The astrophysics of ultra-compact binaries. 10

Perego, A., Dotti, M., Colpi, M. & Volonteri, M. (2009). Mass and spin coevolution during the alignment of a black hole in a warped accretion disc. 12

Perets, H.B. & Alexander, T. (2007). Massive perturbers and the efficient merger of binary massive black holes. 11

Petiteau, A. (2008). *PhD thesis*. Université Paris Diderot (Paris 7). 50

Petiteau, A., Auger, G., Halloin, H., Jeannin, O., Plagnol, E. *et al.* (2008). LISACode: A Scientific simulator of LISA. *Phys.Rev.*, **D77**, 023002. 83, 84

Petiteau, A., Shang, Y. & Babak, S. (2009a). The search for spinning black hole binaries in mock lisa data using a genetic algorithm. *Class. Quantum Grav.*, **26**, 204011. 79

Petiteau, A., Yu, S. & Babak, S. (2009b). The Search for spinning black hole binaries using a genetic algorithm. *Class.Quant.Grav.*, **26**, 204011. 28, 46

Petiteau, A., Shang, Y., Babak, S. & Feroz, F. (2010). The search for spinning black hole binaries in mock LISA data using a genetic algorithm. *Phys.Rev.*, **D81**, 104016. 26, 28, 46

Pirani, F. (1956). On the Physical significance of the Riemann tensor. *Acta Phys.Polon.*, **15**, 389–405. 5

Poisson, E. (2004). Absorption of mass and angular momentum by a black hole: Time-domain formalisms for gravitational perturbations, and the small-hole / slow-motion approximation. *Phys.Rev.*, **D70**, 084044. 93, 109

Poisson, E., Pound, A. & Vega, I. (2011). The Motion of point particles in curved spacetime. *Living Rev.Rel.*, **14**, 7. 15, 29, 91, 106

Portegies Zwart, S.F., Baumgardt, H., McMillan, S.L., Makino, J., Hut, P. *et al.* (2005). The ecology of star clusters and intermediate mass black holes in the Galactic bulge. *Astrophys.J.*. 11

Porter, E.K. (2009). An Overview of LISA Data Analysis Algorithms. 25

Pound, A. (2010a). Self-consistent gravitational self-force. *Phys.Rev.*, **D81**, 024023. 29

Pound, A. (2010b). Singular perturbation techniques in the gravitational self-force problem. *Phys.Rev.*, **D81**, 124009. 29

Pound, A. (2012). Second-order gravitational self-force. *Phys.Rev.Lett.*, **109**, 051101. 29

Pound, A. & Poisson, E. (2008a). Multi-scale analysis of the electromagnetic self-force in a weak gravitational field. *Phys.Rev.*, **D77**, 044012. 29

Pound, A. & Poisson, E. (2008b). Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals. *Phys.Rev.*, **D77**, 044013. 29

Pound, A., Poisson, E. & Nickel, B.G. (2005). Limitations of the adiabatic approximation to the gravitational self-force. *Phys.Rev.*, **D72**, 124001. 29

Press, W. (1977). Gravitational Radiation from Sources Which Extend Into their Own Wave Zone. *Phys.Rev.*, **D15**, 965–968. 95

Prix, R. & Whelan, J.T. (2007). F-statistic search for white-dwarf binaries in the first Mock LISA Data Challenge. *Class.Quant.Grav.*, **24**, S565–S574. 44

Racine, E. & Cutler, C. (2007). Gaussianity of LISA's confusion backgrounds. *Phys.Rev.*, **D76**, 124033. 52

Racine, E., Sterl Phinney, E. & Arras, P. (2006). Non-dissipative tidal synchronization in accreting binary white dwarf systems. *Astrophys.J.*. 10

Rubbo, L.J., Cornish, N.J. & Poujade, O. (2004). Forward modeling of space borne gravitational wave detectors. *Phys.Rev.*, **D69**, 082003. 34, 36

Ruiter, A.J., Belczynski, K., Benacquista, M., Larson, S.L. & Williams, G. (2010). The LISA Gravitational Wave Foreground: A Study of Double White Dwarfs. *Astrophys.J.*, **717**, 1006–1021. 10

Ryan, F.D. (1995). Gravitational waves from the inspiral of a compact object into a massive, axisymmetric body with arbitrary multipole moments. *Phys. Rev. D*, **52**, 5707–5718. 16

Sago, N., Tanaka, T., Hikida, W. & Nakano, H. (2005). Adiabatic radiation reaction to the orbits in Kerr spacetime. *Prog.Theor.Phys.*, **114**, 509–514. 107

Sago, N., Tanaka, T., Hikida, W., Ganz, K. & Nakano, H. (2006). The Adiabatic evolution of orbital parameters in the Kerr spacetime. *Prog.Theor.Phys.*, **115**, 873–907. 107

Sago, N., Barack, L. & Detweiler, S.L. (2008). Two approaches for the gravitational self force in black hole spacetime: Comparison of numerical results. *Phys.Rev.*, **D78**, 124024. 16

Sathyaprakash, B. & Schutz, B. (2009). Physics, Astrophysics and Cosmology with Gravitational Waves. *Living Rev.Rel.*, **12**, 2. 5, 32

Schmidt, W. (2002). Celestial mechanics in Kerr space-time. *Class.Quant.Grav.*, **19**, 2743. 29, 99

Schnittman, J.D. (2007). Retaining Black Holes with Very Large Recoil Velocities. 13

Schnittman, J.D. & Buonanno, A. (2007). The Distribution of Recoil Velocities from Merging Black Holes. *Astrophys.J.Lett.*. 13

Schodel, R., Ott, T., Genzel, R., Hofmann, R., Lehnert, M. *et al.* (2002). A Star in a 15.2 year orbit around the supermassive black hole at the center of the Milky Way. *Nature*, **419**, 694–696. 11

Sigurdsson, S. & Rees, M. (1996). Capture of stellar mass compact objects by massive black holes in galactic cusps. *Mon. Not. Roy. Astron. Soc.*. 15

Stroeer, A. & Nelemans, G. (2009). The Influence of short term variations in AM CVn systems on LISA measurements. *Mon.Not.Roy.Astron.Soc.Lett.*. 10

Sundararajan, P.A., Khanna, G. & Hughes, S.A. (2007). Towards adiabatic waveforms for inspiral into Kerr black holes. I. A New model of the source for the time domain perturbation equation. *Phys.Rev.*, **D76**, 104005. 16

Sundararajan, P.A., Khanna, G., Hughes, S.A. & Drasco, S. (2008). Towards adiabatic waveforms for inspiral into Kerr black holes: II. Dynamical sources and generic orbits. *Phys.Rev.*, **D78**, 024022. 16

Taylor, J.H. (1994). Binary pulsars and relativistic gravity. *Rev.Mod.Phys.*, **66**, 711–719. 4

Teukolsky, S.A. (1973). Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations. *Astrophys.J.*, **185**, 635–647. 107

Timpano, S.E., Rubbo, L.J. & Cornish, N.J. (2006). Characterizing the galactic gravitational wave background with LISA. *Phys.Rev.*, **D73**, 122001. 52

Tinto, M. & Dhurandhar, S.V. (2005). Time-delay interferometry. *Living Rev.Rel.*, **8**, 4. 35, 38

Vallisneri, M. (2005). Synthetic LISA: Simulating time delay interferometry in a model LISA. *Phys.Rev.*, **D71**, 022001. 84

Vallisneri, M. (2008). Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects. *Phys.Rev.*, **D77**, 042001. 24

Valtonen, M., Lehto, H., Nilsson, K., Heidt, J., Takalo, L. *et al.* (2008). A massive binary black-hole system in OJ 287 and a test of general relativity. *Nature*, **452**, 851–853. 11

Wang, Y. & Mohanty, S.D. (2010). Particle Swarm Optimization and gravitational wave data analysis: Performance on a binary inspiral testbed. *Phys.Rev.*, **D81**, 063002. 26, 30, 146

Wang, Y., Shang, Y., Babak, S., Shang, Y. & Babak, S. (2012). EMRI data analysis with a phenomenological waveform. *Phys.Rev.*, **D86**, 104050. 28, 89

Warburton, N., Akcay, S., Barack, L., Gair, J.R. & Sago, N. (2012). Evolution of inspiral orbits around a Schwarzschild black hole. *Phys.Rev.*, **D85**, 061501. 16, 91, 106

Weber, J. (1967). Gravitational radiation. *Phys.Rev. Lett*, **18**, 498–501. 4

Yu, S. & Jeffery, C.S. (2010). The gravitational wave signal from diverse populations of double white dwarf binaries in the galaxy. *A.A*, **521**, A85. 16

Yunes, N., Buonanno, A., Hughes, S.A., Pan, Y., Barausse, E. *et al.* (2011). Extreme Mass-Ratio Inspirals in the Effective-One-Body Approach: Quasi-Circular, Equatorial Orbits around a Spinning Black Hole. *Phys.Rev.*, **D83**, 044044. 16