CARDIFF UNIVERSITY

Mapping the anisotropy of a stochastic gravitational wave background: spherical harmonic decomposition

by

Jack Yu

A thesis submitted in partial fulfillment for the degree of Master of Philosophy

in the

School of Physics and Astronomy

March 2014

Declaration

1. This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

   Signed: _____

   Date: March 2014 _____

2. This thesis is being submitted in partial fulfillment of the requirements for the degree of MPhil.

   Signed: _____

   Date: March 2014 _____

3. This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

   Signed: _____

   Date: March 2014 _____

4. I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

   Signed: _____

   Date: March 2014 _____

5. I hereby give consent for my thesis, if accpeted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards and Quality Committee.**

   Signed: _____

   Date: March 2014

CARDIFF UNIVERSITY

Abstract

School of Physics and Astronomy

Master of Philosophy

by Jack Yu

Gravitational wave sources so numerous that they become unresolvable to detectors produce stochastic backgrounds. The Laser Interferometer Space Antenna (LISA) was a proposed space detector. In this thesis, the way with which it can be used to detect gravitational waves is explained. Then, various methods for estimating the anisotropy of a stochastic background are briefly reviewed. Of these, the *spherical harmonic decomposition algorithm* is implemented to simulate the signal from backgrounds with arbitrary anisotropy, and to compute estimates of a background's anisotropy from data, as well as providing uncertainties of these estimates. The capability of the implementation is demonstrated through the simulation and analysis of of example backgrounds. It is seen that the implementation is able to simulate the signal from a background with the anisotropy of a point source, and recover the anisotropy from simulated signal in the presence of instrumental noise reasonably well. The same thing is also demonstrated for an isotropic background. At a closer look however, it is seen that the overall normalisation of the anisotropy estimates are not consistent with the uncertainty estimates in general. The performance of the implementation is then discussed in the context of the isotropic background. As a step towards resolving this issue, simplified detector responses and noise models are also considered. Several unexplained observations from the simulation of these are described as possible hints for future improvements of the implementation.

Acknowledgements

I would like to thank: Gerald Davies, for helping me get started with almost everything computer-related that is used in this thesis; Michele Vallisneri, for his help and advice, especially with Synthetic LISA; Atsushi Taruya, for showing me how to manipulate the complicated analytical expressions of the overlap-reduction functions using *Mathematica*; Matt Benacquista, for helpful discussions, and for providing simulated stochastic background signals from galactic binaries; David McKechan, my brother and my parents for their 'encouragements'; Sathya, for helping me with the submission of this thesis.

Foremost, I would like to thank my project supervisor, Joe Romano, for his generous guidance, assistance and patience throughout my studies.

Contents

List of Figures

List of Tables

Abbreviations

LAH **L**ist **A**bbreviations **H**ere

Physical Constants

Speed of Light $c = 2.997\ 924\ 58 \times 10^8\ \mathrm{ms}^{-\mathrm{S}}$ (exact)

Symbols

a	distance	m
P	power	W (Js^{-1})
ω	angular frequency	$rads^{-1}$

For/Dedicated to/To my...

Chapter 1

Introduction

In 1905, Einstein published his *general theory of relativity*, a theory of gravity that integrated *special relativity* and Newton's theory of gravity, which has now become *the* theory of gravity in modern physics.[1] Rather than treat gravity as a force, it describes gravity as a *curvature* in the geometry of spacetime, where, in the absence of non-gravitational interactions, massive bodies move along *geodesics*-i.e., the straighest possible paths in a curved spacetime. The theory also says that any mass, or energy, contributes to the curvature of spacetime, and this is summed up in Einstein's field equations.

Many interesting predictions have come out of general relativity, such as the existence of black holes, gravitational lensing and modern cosmology etc. One of the earliest predictions of general relativity was the existence of [2] *gravitational waves*. Generally speaking, a wave can be produced by any accelerating mass, or time-varying distribution of energy, so gravitational wave sources can range from a bike in the velodrome to a pair of inspiralling black holes. As promisingly wide as this range might sound, there has been no direct detection of gravitational waves, mainly because it takes a lot of energy to produce a wave strong enough to be detected with available technologies. There is only indirect evidence of the existence of gravitational waves from the monitoring of binary systems, such as the Hulse-Taylor[3] binary (PSR1913+16). The first binary pulsar system to be discovered, its orbital period was found to be decreasing at a rate that was expected if energy was carried away by gravitational waves. This is the first

[1] Early experiments that test the foundations of general relativity include famous experiments such as the Michelson-Morley experiment, which suggested the invariance of the speed of light, and the Eotvos experiments, which tested the weak equivalence principle. Immediate confirmations of general relativity include the observation of deflection of light by Eddington and the perihelion shift of Mercury. For two decades starting in the 60s, due to certain advances in technology and new astronomical discoveries, many tests were carried out to test the theory where gravity is weak to higher and higher precisions, such as the measurement of the perhelion shift of Mercury and the Pound-Rebka experiment, etc. General relativity has passed all these tests along with some other alternative theories of gravity, but, because general relativity is the simplest, it is the most preferred.

[2] Gravitational waves are perturbations in the metric of spacetime, which satisfy the wave equation in the linear approximation (see Chapter 2).

[3] Russell Hulse and Joseph Taylor discovered the binary pulsar in 1974. They were later awarded the Nobel prize for work on it.

observational evidence of gravitational waves. With a new generation of gravitational wave detectors (such as the LIGO project in the U.S, GEO and Virgo in Europe, and TAMA300 in Japan) currently in operation, scientists around the world are looking to directly detect gravitational waves produced by highly energetic and relativistic astrophysical systems. Detectors like the LIGO are interferometer detectors, in which magnetic radiation is used to monitor the effects of gravitational waves as they change the proper distance between test masses. There are also resonance bar detectors, which measure the oscillations induces on a massive body by gravitational wave as it transfers energy to it. They have been used extensively in the past, but they are not as sensitive as current laser interferometers. Whilst gravitational waves can have a wide range of frequencies, gravitational wave detectors, because of limitation due to noise and detector response, can only look for waves in certain frequency bands. Ground-based interferometric detectors are designed to search at a few hundred Hertzs, being limited at lower frequencies by seismic noise, and at higher frequencies by photon shot noise which is related to the laser power in the interferometer. The detection of gravitational waves would provide another confirmation of general relativity, enable further tests in strong gravity, and provide a whole new way of doing astronomy.

The LIGO project, for example, consists of three laser interferometers: two coaligned detectors, one with 2km-long arms and one with 4km-long arms, are at Hanford, Washington; one other with 4km-long arms is at Livingston, Louisiana. These detectors achieved design sensitivity at the end of 2005 and took data for about 2 years, after which they were shut down for enhancements to improve the sensitivity by a factor of 2. The *enhanced* LIGO detectors came online in 2009 for some observing runs. They were shut down again in order for a major upgrade to be carried out. The resulting *advanced* LIGO, expected in 2014, is expected to have a further factor of 5 in sensitivity, effectively increasing the observable volume in space by a factor of 1000, making regular detections highly probable. In the Virgo project, there is a 3km-arms interferometer at Cascina. In the GEO project, there is a 600-m interferometer, which has been running since 2001 at Hanover, Germany. The 'second-generation' technology that has been developed there has been transfered to LIGO and VIRGO for their upgrades. It will also act as a cover for the larger detectors when they are down during the upgrade process. There is also the 300m-arms inteferometer TAMA300 in Japan, which is used for developing advanced techniques that can be used in larger detectors.

At relatively low frequencies ($\sim$ 1mHz), ground-based detectors do not work, because terrestrial disturbances close to the detector can potentially induce on the test-masses gravity gradient effects which are larger than the effect produced by gravitational waves. Therefore, detection strategies for waves at these frequencies are devised in space, where such disturbances are minimal. These include spacecraft ranging, pulsar timing and interferometry. All of these rely on the accurate measurements of the light travel-time between freely falling reference objects, with the aim of measuring a change in these travel-times due to the effects of

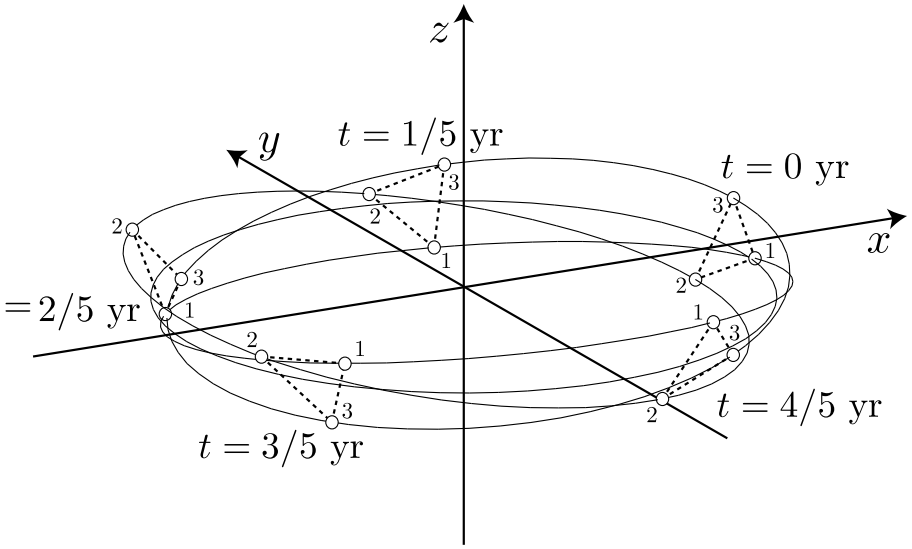

FIGURE 1.1: Orbits of the spacecrafts in LISA. The Sun is located at the origin. The distance between the spacecrafts themselves and the distance between LISA's guiding centre and the Sun are not to scale. The LISA triangle rotates once as its guiding centre orbits around the Sun once. (Figure copied from [1])

gravitational waves.

There have been numerous ideas of space-borne detectors, and one of these was the proposed Laser Interferometer Space Antenna (LISA) [3][4]. It consists of three spacecrafts, in orbits around the Sun, forming a quasi-equilateral triangle configuration that is inclined with respect to the ecliptic plane by 60 ° and trails the Earth in its orbit by 20 °. (See Fig. 1.1 . for an illustration of its configuration and its movement around the Sun.) The distance between the spacecrafts is about 5×10^6km. Inside each spacecraft, there are two proof-masses which are kept as freely falling as possible with the shielding from the spacecrafts, which also fire tiny retro-rockets to keep themselves as steady as possible under external forces such as solar radiation and solar wind. Laser transmission is enabled in both directions between any two proof-masses using transponders, and changes in these transmission times is measured to look for signatures of gravitational waves. The details of this transmission For an in-depth description of LISA's design, see [3]. Operating between 10^{-4} and 10^{-1}Hz, LISA is expected to detect gravitational waves from many different kinds of astrophysical systems, including *supermassive black hole inspirals* and *extreme mass ratio inspirals*, and possibly other unknown sources [4]. However, of these, the galactic compact binaries are so numerous that LISA would not be able to resolve all of them individually. In fact, waves from these unresolvable sources arrive at LISA as a confusion noise forming what is called a *stochastic gravitational wave background*. And because the galaxy is not uniformly distributed across the sky (just look at the Milky Way at night), this galactic background is *anisotropic*, emitting more gravitational waves in some direction than others. If this anisotropy can be estimated, the information might shed more light on how the compact binaries population are distributed in the galaxy.

[4]Video summary of the mission: http://www.youtube.com/watch?v=MuHE8x8nq7U

1.1 Outline

I have chosen to organise the thesis as follows:

Chapter 2: An introduction to gravitational waves is given. This includes a description of some general properties of gravitational waves, astrophysical sources and the detection of gravitational waves. The later part of this chapter contains a concise introduction of gravitational waves as considered in the linearised theory. Several useful expressions that will be used in later chapters are given.

Chapter 3: The way with which LISA can be used to optimally detect gravitational waves is explained. It starts by introducing LISA's basic one-way Doppler measurements, followed by a description of their responses to gravitational wave and instrumental noises. Then, following a short introduction on laser noise-cancellation using Time Delay Interferometry (TDI), the response to gravitational waves and to noises of TDI observables are described.

Chapter 4: Different types of stochastic gravitational wave background are introduced and assumptions about their nature are discussed. Making these assumptions, these backgrounds are then characterised by various ensemble averages.

Chapter 5: Brief reviews of several existing algorithms for mapping an anisotropic stochastic gravitational background are given. Particular attention is paid to the spherical harmonic decomposition algorithm currently being used in LIGO searches.

Chapter 6: The implementation of the *spherical harmonic decomposition* algorithm for searches with LISA is described. The general workflow and elements of the pipeline are explained. Important or difficult technical details are emphasised. Crucial test results and illustrations of the pipeline working are presented.

Chapter 8: Discussion on the performance of the data-simulation and data-analysis pipeline, issues related to their usage, and suggestions for possible improvements.

1.2 History

After the majority of the work in thesis was completed, it was decided that the LISA mission was too expensive. A modified version of it, called the New Gravitational Wave Observatory, has since been proposed, which consists of two arms instead of three.

Chapter 2

Gravitational waves

In Newton's theory of gravity, changes in the gravitational field due to motions of massive bodies reach all observers and affect everything that respond to gravity *instananeuosly*. However, in a relativistic theory of gravity, they propagate at a finite and limiting speed $v_g \leq c$ and reach different observers at different times. These propagating changes in the gravitational field are known as *gravitational waves*. Analogies of these propagating oscillations exist in many other physical systems: water waves, sound waves, electromagnetic waves etc. In this work, gravitational waves predicted by the general theory of relativity are considered.

As will be shown in the next section, in general relativity, gravitational waves are transverse waves, producing a 'tidal' force in the plane perpendicular to their direction of propagation. This force acts on matter in a way such that, if a set of test masses are arranged in a ring in the plane normal to the direction of wave propagation, its shape oscillates between a circle and an ellipse[1]. (See Fig. 2.1.) The tidal force squeezes the ring in one direction and stretches it in another. And since the force is tidal, a ring of a different size still deforms to the same extent. (The ellipse at the extremum of the oscillations has the same eccentricity.) In fact, the amount of this fractional change (fractional change in the radius of the ring for example) is the measure of the wave's strength and it is called the *strain* induced by the wave, often denoted by $\delta l/l$. The amplitude of the wave, denoted by h, is defined to be twice the strain, $h \approx 2\delta l/l$.

The type of motion of masses that generates a gravitational wave tends to be the same as the one which the gravitational wave itself would induce on the masses. For example, the kind of motion that gives rise to the waves in Fig. 2.1 are similar to the motions of those test masses. Therefore, since GWs are transverse, the gravitational wave generated in a given direction is from the effective motion of the source in the plane *perpendicular* to that given direction. Also, similar to the circles turning into ellipses in Fig. 2.1 , this motion needs to have certain

[1] [5] In general relativity, the ellipse at any phase has the same area as the circle. In some scalar theories of gravity, gravitational waves are not *area-preserving* however.

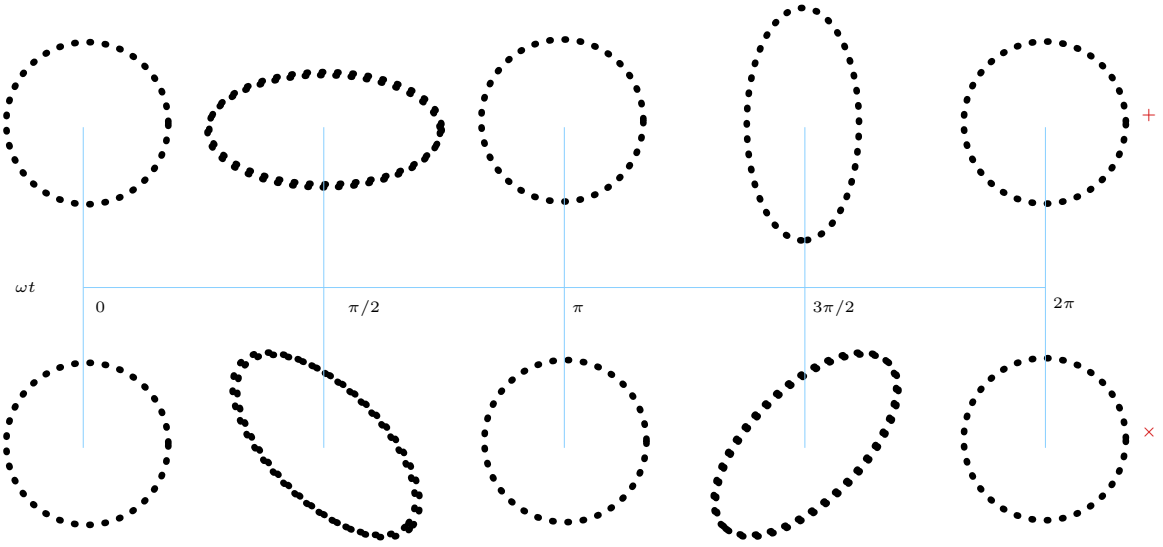

FIGURE 2.1: Gravitational wave action on a ring of test particles. (Top row) 'plus' polarisation. (Bottom row) 'cross' polarisation. Each dotted circle or ellipse is the ring of test particles at phases $\omega t = 0$, $\pi/2$, π, $3\pi/2$ or 2π. The two polarisations differ by a rotation of 45 ° about the direction of wave propagation (which is into the page).

irregularity, or asymmetry. A circular ring of test masses that simply expands and shrinks repeatedly would not generate any gravitational wave.

The frequency of a gravitational wave usually corresponds to the time-scale on which phsyical processes take place inside the source. By considering astronomical systems as sources of gravitational waves, the upper bound for the waves' frequency is about 10^4 Hz, because at this scale it is difficult for the most widely separated parts of the system to be causally connected on time-scales shorter than 10^{-4} s.

In analogy to waves in ordinary materials [5], gravitational waves, travelling at the fastest possible speed c and in the 'stiffest material' that is spacetime, have small amplitudes, even when they are created by very energetic processes. As a result, it is virtually impossible to manufacture gravitational waves in the laboratory that are anywhere strong enough to be detectable with today's technology. However, it is expected that some astrophysical systems, with the large masses and velocities involved, produce gravitational waves which are strong enough to be detected here on Earth. In binary star systems (such as white dwarf binaries, neutron star binaries and black hole binaries, etc.), as the two stars inspiral towards each other and coalesce, part of the energy lost is through the emission of gravitational waves. As compact objects or black holes are formed during supernova explosions, non-spherical parts of the gravitational collapse produce gravitational waves which carry some of the star's binding energy and angular momentum, and this is another type of GW source. Gravitational waves can also be emitted by spinning neutron stars with a non-uniform density. It is also expected that there will be stochastic backgrounds of gravitational waves. An *astrophysical* stochastic background can be formed as the waves emitted by many astrophysical systems become

unresolvable. Random processes in the early universe can also form a stochastic background that is analogous to the Cosmic Microwave Background Radiation. This *cosmological* GW background is thought to have been produced within about 10^{-35} s after the Big Bang.

While the detection of gravitational waves is an important test of general relativity by itself, their potential existence can also provide further stringent tests of the theory in the strong field (for example, gravitational waves from black holes.). In addition to these important roles in physics, gravitational waves would also have applications in astrophysics and cosmology [4]. Electromagnetic waves are emitted primarily by charged elementary particles, and because opposite charges have the tendency to cancel each other, electromagnetic waves are only emitted over small regions, so they only give direct information about the physical conditions of small regions of astronomical systems. Gravitational waves, however, being generated by motions of mass of entire systems, give direct information about large-scale regions. Also, because the electromagnetic interaction is strong, electromagnetic waves tend to be scattered or absorbed by the matter between the source and the observer. On the other hand, as mentioned above, gravitational waves tend to have small amplitudes and interact weakly with matter, so they arrive at the observer virtually unaffected by the matter in between. Hence, considering that most of the matter-energy in the universe is electrically neutral, gravitational waves would provide astronomers with a new tool to study a wide range of astrophsyical systems (like the ones mentioned above), and potentially reveal new exotic objects in the Universe.

Ground-based beam detectors' designs are based on the Michelson interferometers, where a light beam is split into two, which are then sent down two different paths; they are then reflected off mirrors and brought back together to interfere with each other. However, inside a interferometric gravitational wave detector, the mirrors at the end of the interferometer's two arms act as test masses which can be freely falling in the gravitational wave's field. For a wave perpendicular to the plane of the detector, these move in the same way as those test masses in Fig. 2.1 , so as one arm of the detector is stretched, the other is squeezed. This causes a difference in the time it takes the laser beam to travel down the arms and back to the photo-detector, resulting in a change in the observed inteference pattern.

2.1 Gravitational waves in linearised theory

In the linearised theory of general relativity, it is assumed that there exists a reference frame in which the metric tensor can be taken to consist of a background flat-spacetime metric $\eta_{\mu\nu}$ and a small expansion, $h_{\mu\nu}$, around it:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \ , \tag{2.1}$$

where $h_{\mu\nu} = h_{\nu\mu}$ and $|h_{\mu\nu}| \ll 1$. Only under certain coordinate transformations, called *Poincare transformations*, can one get to another reference frame in which (2.1) is also true. These include translations, Lorentz transformations and infinitesimal local transformations (also known as gauge transformations). This is in constrast with the full theory of general relativity, which is invariant under all coordinate transformations. Reference frames in which the condition $\partial^\mu h_{\mu\nu} = 0$ is imposed through infinitesimal local transformations, are called the *Lorentz gauge*. In this gauge, the Einstein equations become

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu} \ , \tag{2.2}$$

where $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu}$ (and $h = h^{\mu\nu} h_{\mu\nu}$) is the trace-reverse of $h_{\mu\nu}$. By considering the energy-momentum tensor $T_{\mu\nu}$ of the source, this equation forms the basis for computing the generation of gravitational waves.

When studying the propagation of gravitational waves and its interaction with test masses outside the source, $T_{\mu\nu} = 0$, (2.2) becomes

$$\Box \bar{h}_{\mu\nu} = 0 \ . \tag{2.3}$$

In this case, it is possible to further use infinitesimal local transformations to impose additional constraints, giving, all together, the following conditions on $h_{\mu\nu}$:

$$h^{0\mu} = 0 \ , \quad h^i_i = 0 \ , \quad \partial^j h_{ij} = 0 \ . \tag{2.4}$$

A reference frame, in which (2.4) is true, is called a *transverse-traceless* gauge, and $h_{\mu\nu}$ in this frame is often denoted by $h^{TT}_{\mu\nu}$ to indicate that they are the components in this gauge.

As a symmetric tensor, $h_{\mu\nu}$ has 10 independent components. $\partial^\mu h_{\mu\nu} = 0$ imposes four constraints on $h_{\mu\nu}$, so in the Lorentz gauge, there are only 6 indepedent components. Since each infinitesimal local transformation is done through the choice of 4 independent arbitrary functions, each such transformation on $h_{\mu\nu}$ reduces the number of independent components it has by 4. Therefore, there are only 2 independent $h^{TT}_{\mu\nu}$s in the transverse-traceless gauge.

Using (2.3) and (2.4), gravitational waves can be further described. Equation (2.3) admits plane wave solutions which, because $h^{TT}_{0\mu} = 0$, can essentially be written in the form $h^{TT}_{ij}(t, \vec{x}) = e_{ij}(\hat{k}) e^{ik^\mu x_\mu}$. Since $\Box = -(1/c^2)\partial_t^2 + \partial_j \partial^j$, the wave speed is c. The wave vector, $\vec{k}$, having components $(\omega/c, \vec{\Omega})$, is therefore a null vector (i.e. $k^\mu k_\mu$=0). This implies that the 3-vector, $\vec{\Omega}$, completely specifies the frequency and the direction of propagation of the wave. Applying the Lorentz condition from (2.4) to this solution, it can be shown that $k^i h^{TT}_{ij} = 0$; the only non-zero components of h^{TT}_{ij} are those in the direction perpendicular to the direction of propagation, so gravitational waves are transverse waves. For example, for a wave travelling in the z-direction, $h^{TT}_{zi} = 0$. If the traceless condition in (2.4) and $h_{\mu\nu} = h_{\nu\mu} = 0$ are also applied, the remaining

components of h_{ij}^{TT} are sorted as:

$$h_{\mu\nu}^{TT} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{2.5}$$

h_+ and $h_\times$ are called the wave amplitudes of the + polarisation and × polarisation, respectively, and are the two independent components of $h_{\mu\nu}^{TT}$ as mentioned above. Using (2.5), the plane wave solution can be written as

$$h_{ij}^{TT}(t,\bar{x}) = h_+(t,\bar{x})e_{ij}^+(\hat{z}) + h_\times(t,\bar{x})e_{ij}^\times(\hat{z}) , \tag{2.6}$$

where $e_{ij}^+(\hat{z}) = \hat{x}_i\hat{x}_j - \hat{y}_i\hat{y}_j$ and $e_{ij}^\times(\hat{z}) = \hat{x}_i\hat{y}_j + \hat{y}_i\hat{x}_j$ are the tensor basis for + polarisation and the × polarisation, respectively; the x-axis and y-axis are called the principal axes of polarisation, as the unit vectors along them form the polarisation tensor bases. It is important to note that the direction of the wave ($\hat{z}$ in this case) and its principal axes of polarisation ($\hat{x}$ and $\hat{y}$ here) form a right-handed system, and that it can be obtained from that of the $(\hat{x},\hat{y},\hat{z})$ through Euler rotations. In general, for a wave from direction $\hat{k} = -\hat{\Omega}$ (at polar angle θ and azimuthal angle ϕ), and having principal polarisation axes $\hat{m}$ and $\hat{n}$ (described by an angle of rotation ψ in the plane perpendicular to $\hat{k}$), its right-handed system $(\hat{k},\hat{m},\hat{n})$ can be written as[2]

$$\begin{pmatrix} \hat{m} \\ \hat{n} \\ \hat{k} \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\phi\cos\psi - \sin\phi\sin\psi & \cos\theta\cos\psi\sin\phi + \cos\phi\sin\psi & -\cos\psi\sin\theta \\ -\cos\psi\sin\phi - \cos\theta\cos\phi\sin\psi & \cos\phi\cos\psi - \cos\theta\sin\phi\sin\psi & \sin\theta\sin\psi \\ \cos\phi\sin\theta & \sin\theta\sin\phi & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} \tag{2.7}$$

ψ is called the polarisation angle. Using these, it is possible to generalise the plane wave solution in (2.6) and write it as[3]:

$$h_{ab}(t,\bar{x}) = h_+(t,\bar{x})e_{ab}^+(\hat{k},\psi) + h_\times(t,\bar{x})e_{ab}^\times(\hat{k},\psi) , \tag{2.8}$$

where

$$e_{ab}^+(\hat{k},\psi) = \hat{m}_a\hat{m}_b - \hat{n}_a\hat{n}_b , \tag{2.9}$$

$$e_{ab}^\times(\hat{k},\psi) = \hat{m}_a\hat{n}_b + \hat{n}_a\hat{m}_b . \tag{2.10}$$

[2] $(\hat{k},\hat{m},\hat{m})$ is basically a new set of axes found from $(\hat{x},\hat{y},\hat{z})$ by first rotating it by ϕ about $\hat{z}$, then by θ about $\hat{y}$, and then by ψ about $\hat{z}$ again.

[3] The superscript TT is dropped, but the using of ab intead of ij indicates that h_{ab} are the components in the transverse-traceless gauge.

When there are plane waves of different frequencies and directions, the total metric perturbation can be written in a plane wave expansion:

$$h_{ab}(t,\bar{x}) = \int_{-\infty}^{+\infty} df \int_{S^2} d\hat{k}\; e^{A}_{ab}(\hat{k})\; \mathcal{H}_A(f,\hat{k})\; e^{i2\pi f(t+\hat{k}\cdot\bar{x})}\;, \tag{2.11}$$

where $A = +, \times$.

The interaction of gravitational waves with test masses are often described with the use of the *geodesic equation* and the *equation of geodesic deviation* (see Chapter 1.3 in [6]) in the frame of interest. In the transvere-traceless frame, test masses initially at rest will remain at rest with the passing of gravitational waves. Coordinate separations between them remain constant. However, the proper distances between them change, so it is possible to measure the effects of gravitational waves (for example, by monitoring the time it takes for light to travel across these proper distances). There will be more details on this in the next chapter. In the *proper detector frame*, by considering a plane wave transverse to the x-y plane, the geodesic equation says that the coordinates of a test mass at an initial position (x_0, y_0) change as

$$\delta x(t) = \frac{h_{+/\times}}{2} x_0 \sin\omega t\;, \tag{2.12}$$

$$\delta y(t) = -\frac{h_{+/\times}}{2} y_0 \sin\omega t\;. \tag{2.13}$$

For a ring of test masses, these expressions describe exactly the oscillatory movement illustrated in Fig. 2.1. The two test masses separated by 90 ° from the centre of the ring are like the test masses in the two arms of an interferometer gravitational wave detector.

Chapter 3

Detecting Gravitational Waves with LISA

3.1 LISA's basic measurements

Fig. 3.1 is a schematic that illustrates the LISA configuration. The three spacecrafts are represented here by the three black dots at the vertices of the triangle. The six arrows point in the direction of the laser beam transmission along the six links (or one-way Doppler measurements), two in opposite directions along each arm. In this diagram, each link is enumerated by the spacecraft opposite it. If it points in the clock-wise direction, the link number is primed. There are two optical benches on each spacecraft, and on each optical bench, there is a proof-mass. The optical benches and their proof-masses are indicated by the numbers beside the spacecrafts in the schematic. The number associated with an optical bench is starred, if a

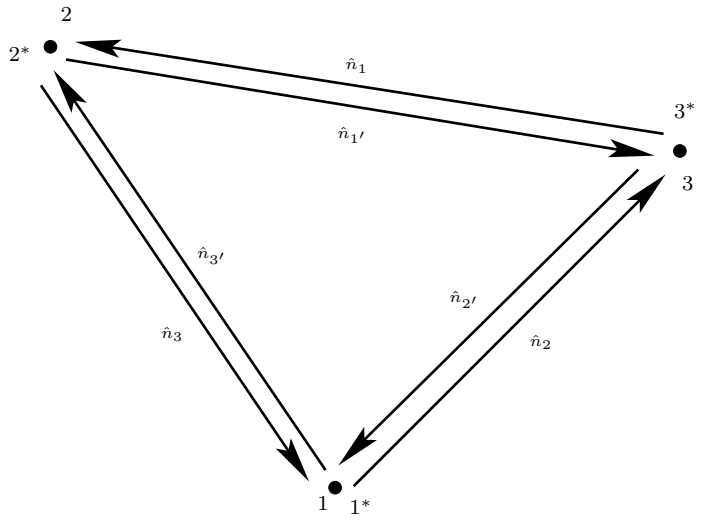

FIGURE 3.1: Schematic LISA configuration. There are *three* spacecrafts: $\{1, 2, 3\}$, *six* links: $\{1, 2, 3, 1', 2'3'\}$ and *six* optical benches and proof-masses: $\{1, 2, 3, 1^*, 2^*, 3^*\}$.

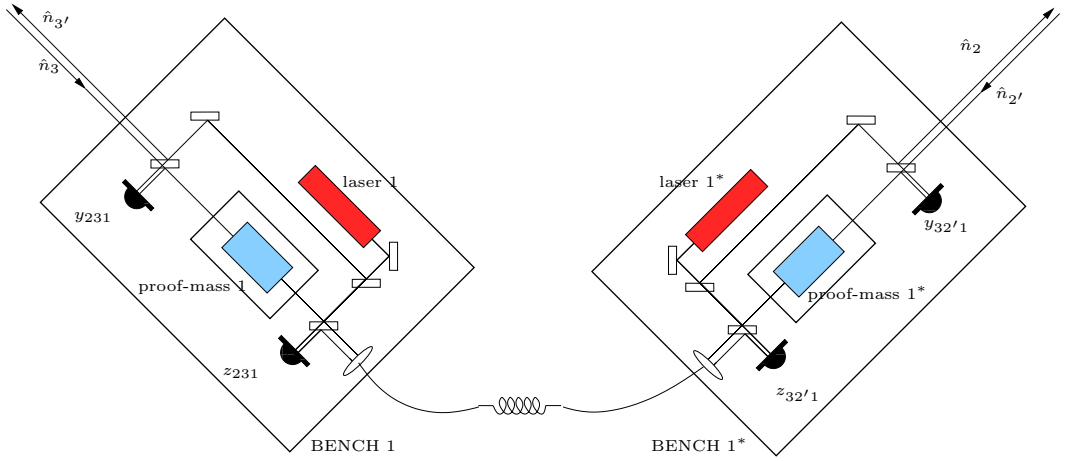

FIGURE 3.2: Schematic diagram of the proof-mass and optical-bench within LISA spacecraft 1. The black semi-circular shapes are *phasemeters* which take the basic measurements.

clock-wise link arrives at it. Each interspacecraft measurement (its data or time-series) is here denoted by y_{slr}; s (for sender) is the number of the spacecraft from which the laser beam of the link is emitted; l (for link) is the number indicating the link along which the laser beam is sent, and r (for receiver) is the number of the spacecraft at which the laser beam of the link is received. For example, the interspacecraft measurement consisting of the transmission of the laser from spacecraft 1 to spacecraft 3 is denoted by y_{123}. More precisely, this transmission goes from proof-mass 1* to proof-mass 3. Fig. 3.2 zooms in on the optical benches on board spacecraft 1. The schematic shows two optical benches, each carrying a proof-mass, mirrors, and the paths along which the laser beams travel. It also shows that between the two optical benches, there are two intraspacecraft links, one going from the bench 1 to bench 1*, and another in the opposite direction. Each intraspacecraft measurement(its data or time-series) is denoted by z_{slr}; where l is the link whose receiving optical bench is the one that z_{slr} arrives at. For example, the intraspacecraft measurement consisting of the transmission from optical bench 1 to 1* is named $z_{32'1}$, because the interspacecraft link $32'1$ arrives, and is the only one that does, at optical bench 1*. Looking at bench 1 in Fig. 3.2, the *incoming* laser travels in the direction of $\hat{n}_3$. Part of it goes through the first beam splitter, hits proof-mass 1, reflects off the first beam splitter and goes into the phasemeter of the interspacecraft measurement y_{231}. The *outgoing* laser originates from laser 1, reflects off a mirror, a beam splitter, another mirror, and finally a beam splitter before travelling in the direction of $\hat{n}_{3'}$.

3.2 Gravitational wave response

The change in the time it takes for the laser beams to travel between the spacecrafts is essentially the response of LISA to gravitational waves. As mentioned above, the lasers on board the spacecrafts have the same nominal frequency. In the absence of gravitational waves, the incoming laser on board the spacecrafts always has the same phase difference relative to

the local laser, because the time of travel for the photons stays constant. In fact, if the lasers are phase-locked, then the phase difference is always zero. In the presence of gravitational waves, this phase difference changes. This is what is measured by the *phasemeters* on board LISA's spacecrafts. The rate of change of the phase difference is the fractional change in frequency between the incoming and local laser, and this will be the definition used to here for the gravitational wave response. [1] For details of its derivation, see Appendix A, but the result is [1][2][9]

$$y_{slr}^{GW} = (1 - \hat{k} \cdot \hat{n}_l) \left[\Psi_l(t_s + \hat{k} \cdot \vec{p}_s(t_s)) - \Psi_l(t_r + \hat{k} \cdot \vec{p}_r(t_r)) \right] \quad , \tag{3.1}$$

where

$$\Psi_l(t + \hat{k} \cdot \bar{x}) = \frac{1}{2} \frac{\hat{n}_l^a \hat{n}_l^b h_{ab}(t, \bar{x})}{1 - (\hat{k} \cdot \hat{n}_l)^2} \tag{3.2}$$

$(t_s, \vec{p}_s(t_s))$ and $(t_r, \vec{p}_r(t_r))$ are the coordinates of the event of the photon leaving the sender and the event of the photon being received by the receiver, respectively. $\hat{n}_l$ is the unit vector along the spatial path of the photon, defined by

$$\hat{n}_l = \frac{\vec{p}_r(t_r) - \vec{p}_s(t_s)}{|\vec{p}_r(t_r) - \vec{p}_s(t_s)|} = \frac{\vec{p}_r(t_r) - \vec{p}_s(t_s)}{L_l(t_r)} \tag{3.3}$$

Implicitly,

$$t_s = t_r - L_l(t_r) \ , \tag{3.4}$$

where $L_l(t_r)$ is the time (setting $c = 1$) it takes for a photon to travel along the link l to be received by the receiver at time t_r. This Doppler measurement is the reponse of the basic observable, denoted by slr, to a plane gravitational wave described by (**??**). This is also known as the Estabrook-Wahlquist two-pulse response. Note as this name suggests, the response depends on both the time and location at which the photon is sent and received. In order to assign a time-stamp to the response , an event of reference is usually chosen, whose time is used as the time-stamp of the time-series of the response. In some papers, like [7], the emission event is chosen, so $y_{slr}^{GW} = y_{slr}^{GW}(t_s)$. In others, like [2], the reception event is used, so $y_{slr}^{GW} = y_{slr}^{GW}(t_r)$. In fact, the event of reference can be chosen quite arbitrarily. Taking it at some $(t_0, \vec{x}_0)$, and substituting in h_{ab} from (**??**), (3.1) can be written as

$$y_{slr}^{GW}(t_0, \vec{x}_0) = F_l^A(f, \hat{k}, \psi, t_r; t_0, \vec{x}_0) \ \mathcal{H}_A(f, \hat{k}, \psi) \ e^{i2\pi f(t_0 + \vec{k} \cdot \vec{x}_0)} \tag{3.5}$$

[1] Some authors use the fractional phase as the response. For example, [7] and [8]. Some use the fractional frequency. For example, [9] and [1].

, where the following definitions are used:

$$F_l^A(f,\hat{k},\psi,t_r;t_0,\vec{x}_0) = \mathrm{D}_l^{ab}(f,\hat{k},t_r;t_0,\vec{x}_0)\ \mathrm{e}_{ab}^A(\hat{k},\psi) \tag{3.6}$$

$$\mathrm{D}_l^{ab}(f,\hat{k},t_r;t_0,\vec{x}_0) = \hat{n}_l^a(t_r)\hat{n}_l^b(t_r)\ \mathcal{T}_l(f,\hat{k},t_r)\ e^{i2\pi f(t_r-t_0+\hat{k}\cdot(\vec{p}_r(t_r)-\vec{x}_0))} \tag{3.7}$$

$$\mathcal{T}_l(f,\hat{k},t_r) = -i\frac{f}{2f_l(t_r)}\ \mathrm{sinc}\left(\frac{f}{2f_l(t_r)}(1+\hat{k}\cdot\hat{n}_l(t_r))\right) e^{-i\frac{f}{2f_l(t_r)}(1+\hat{k}\cdot\hat{n}_l(t_r))}\ , \tag{3.8}$$

and $\mathcal{H}_A(f,\hat{k},\psi)$ is given by ?(So far it's not been given by anything.) $F_l^A(f,\hat{k},\psi,t_r;t_0,\vec{x}_0)$, $\mathrm{D}_l^{ab}(f,\hat{k},t_r;t_0,\vec{x}_0)$ and $\mathcal{T}_l(f,\hat{k},t_r)$ are called the *response function*, the *detector tensor* and the *transfer function*, respectively. From (3.8), note that the transfer function does not depend on the choice of the reference event, whereas the detector tensor and the response function do depend on this choice. The t_r dependence of $\hat{n}_l$, L_l and $f_l = 1/(2\pi L_l)$, the *the transfer frequency*, are also shown explicitly here. Therefore, the time dependence of the detector tensor, or of the response function, does not solely come from the $e^{i2\pi f(t_r-t_0+\hat{k}\cdot(\vec{p}_r(t_r)-\vec{x}_0))}$ term.[2]

In (3.5), y_{slr}^{GW} is labeled by the chosen reference event $(t_0,\vec{x}_0)$. As shown above, it equal to the product of the response function and the GW amplitude at the reference event. It needs to be emphasised though that the choice of this event is arbitrary; what are not are the actual emission and reception events of the two-pulse response. Their time stamps and locations, are implicitly indicated by the subscript *slr* in the LHS, and by t_r and $\vec{p}_r$ in the RHS. When $(t_0,\vec{x}_0) = (t_r,\vec{p}_r(t_r))$,

$$\begin{aligned} y_{slr}^{GW}(t_r,\vec{p}_r(t_r)) &\equiv y_{slr}^{GW}(t_r) \\ &= \mathcal{T}_l(f,\hat{k},t_r)\ \hat{n}_l^a\hat{n}_l^b\ \mathrm{e}_{ab}^A(\hat{k},\psi)\ \mathcal{H}_A(f,\hat{k},\psi)\ e^{i2\pi f(t_r+\hat{k}\cdot\vec{p}_r)} \end{aligned} \tag{3.9}$$

In what follows, whenever there is a need to select a time stamp for a link measurement, unless otherwise stated, the time of reception of the link will be used.

3.3 LISA noise sources

There are several sources of noise in LISA. (See [3] for a detailed description of them.) The most dominant source of these comes from the fluctuation of the lasers' frequency, called *laser noise*. The rest can be divided into two main categories. In the first category, the optical path along LISA's links fluctuates due to a combination of shot noise and beam-pointing instabilities, etc. These are collectively called the *optical path noise*. Then there are the random fluctuations of the positions of the proof-masses on the optical benches. These also

[2] t_r is used here delibrately to label the time-dependence of $\hat{n}_l$, L_l and $f_l = 1/(2\pi L_l)$, because t_r is used in this exponential term. Had expressions been written in a different way where t_s appears in the exponential term, then it would be more tidy to use t_s for the time-dependence. But either is fine since the two are related through (3.4).

distort the measurements, and are collected in the *proof-mass noises*. In most of the literature and in most simulators of LISA's scientific processes, these three noises are used, and they are assumed to be stationary, and either white or coloured with a power-law power spectrum [9].

Laser Frequency Noise There is a laser noise associated with each of the six lasers. For example, y_1^{laser} denotes the noise due to the fluctuation of the nominal frequency of the laser on optical bench 1, and y_{1*}^{laser} is that of the laser on optical bench 1^*. In LISA, all laser noises are expected to be white, uncorrelated with each other, and all having an expected power spectral density of fractional frequency fluctuations of

$$P^{laser}(f) = 1.1 \times 10^{-26}\ \mathrm{Hz}^{-1} \tag{3.10}$$

Proof Mass Noise There are six proof-mass noises. For example, y_2^{pm} denotes the proof-mass noise of the proof-mass on optical bench 2, and y_{2*}^{pm} is the proof-mass noise of the proof-mass on optical bench 2^*. Each proof-mass noise contains a white acceleration noise of $3 \times 10^{15} \mathrm{m\,s^{-2}\,Hz^{-1/2}}$ along the line-of-sight. This translates to a power spectral density of fractional frequency fluctuations of

$$P^{pm}(f) = 2.5 \times 10^{-48} f^{-2}\ \mathrm{Hz}^{-1} \tag{3.11}$$

All proof-mass noises are uncorrelated with each other, and they all have the above power spectral density.

Optical Path Noise There is an optical path noise associated with each of the six links. For example, y_{slr}^{op} denotes the optical path noise of the link slr. This consists of shot noise and beam-pointing noise on the optical bench r, if l is unprimed, or r^*, if l is primed. All the optical path noises are uncorrelated with each other, and have an expected power spectral density of fractional frequency fluctuations of

$$P^{op}(f) = 1.8 \times 10^{-37} f^{2}\ \mathrm{Hz}^{-1} \tag{3.12}$$

These three different types of noises all enter into the basic one-way Doppler measurements as

$$y_{slr}^{noise}(t) = \begin{cases} y_s^{laser*}(t - L_l(t)) - y_r^{laser}(t) + y_{slr}^{op}(t) - 2y_r^{pm}(t) & \text{if } l \text{ is unprimed}\ , \\ y_s^{laser}(t - L_l(t)) - y_{r*}^{laser}(t) + y_{slr}^{op}(t) - 2y_{r*}^{pm}(t) & \text{if } l \text{ is primed}\ . \end{cases} \tag{3.13}$$

$$z_{slr}^{noise}(t) = \begin{cases} y_r^{laser*}(t) - y_r^{laser}(t) + 2y_{r*}^{pm}(t) & \text{if } l \text{ is unprimed}\ , \\ y_r^{laser}(t) - y_{r*}^{laser}(t) + 2y_r^{pm}(t) & \text{if } l' \text{ is primed}\ . \end{cases} \tag{3.14}$$

$L_l(t)$ measures the light travel time along the link l, as defined in (3.4). It can be seen in (3.13) and (3.14) that the laser frequency noises have the same form in the interspacecraft

and the intraspacecraft measurements of the same name, except that in the interspacecraft measurement, the laser noise from the 'sender' is delayed by the light-travel time along the link. As the laser frequency noise is a considerably stronger source of noise than the others, the laser noise is a primary noise, and the rest are secondary noises.

3.4 Time delay interferometry

In section 3.2, it is shown that the effect of gravitational waves on two freely falling test masses(spacecrafts) can be measured by sending a laser beam from one to another. The gravitational wave response is essentially the change in the light travel time between the two test masses. In practice, however, the laser's nominal frequency fluctuates and introduces a laser noise into the measurements, and because this is of several orders of magnitude larger than the weakest gravitational waves detectable by LISA, gravitational wave signals get buried in the noise. *Time delay interferometry*, or TDI, is the technique for combining the available one-way Doppler measurements into virtual measurements which contain laser noise at the level of the secondary noises or lower. These combinations are called the *TDI observables*.

In ground-based interferometric GW detectors, which, by design, have equal arms, the same laser noise goes into the two arms at the beam-splitter, and experience the same amount of delay, cancelling each other out at the photodetector. For LISA, the laser noises from two different 'arms' do not cancel because LISA's armlengths vary with time, with differences as large as $0.5s$. [10] Therefore, although TDI is not needed for ground-based interferemetric detectors, it is a necessity for LISA.

It is useful to make a few simplifying assumptions first before considering the cancellation of laser noise in more detail. We will first concentrate on laser noise, ignoring the secondary noises in (3.13) and in (3.14). Also, for the purpose of explaining laser noise-cancellation, the intraspacecraft measurements can be ignored, since their laser noises, as it will be seen, cancel in the same way as those in interspacecraft measurements. This is equivalent to assuming that there is only one optical bench and one proof-mass in each spacecraft. With these assumptions, (3.13) and (3.14) reduce to

$$y_{slr}^{noise}(t) = y_s^{laser}(t - L_l(t)) - y_r^{laser}(t) \ . \tag{3.15}$$

Often, models of LISA of varying complexities are considered. Since the armlengths, or light travel time, along links is an important factor in laser noise-cancellation, the different complexities below are closely associated with different models of LISA:

Equal-arm, non-rotating and static. The simplest of these is an equal-arm, non-rotating and static LISA. In this model, the three spacecrafts form an equilateral triangular configuration, which lies at the centre of the source frame, and does not rotate. The armlengths

remain the same, and constant. The light travel time along any link is taken to be the nominal armlength $L_l(t) = L = 5 \times 10^6 \text{km} \approx 16.67\text{s}$.

Unequal-arm, non-rotating and static. At the next level of complexity is the unequal-arm, non-rotating and static LISA. In this model, the LISA triangle also lies at the centre of the source frame, and does not rotate. Each armlength still remains constant, but the armlengths of different arms are no longer equal to each other. This means that $L_{l_1} \neq L_{l_2}$, if l_1 and l_2 are not the same link.

Unequal-arm, rotating and static. In this model, in addition to having unequal armlengths, the LISA triangle also rotates about some fixed direction(usually the z-axis) in the source frame. The resulting *Sagnac effect* causes the light travel time along links of opposite directions to be different. This implies that $L_l \neq L_{l'}$.

Unequal-arm, rotating and non-static. Finally, in the most complicated model of LISA, the spacecrafts follow their realistic orbits around the Sun, forming a nearly equilateral configuration. As this orbits around the Sun at a distance of 1 AU, it is inclined to the ecliptic plane, and rotates about its centre. Due to orbital dynamics, the armlengths change with time, so, here, the light travel time along the links are functions of time, denoted by $L_l(t)$.

One of the most simple TDI observables is the $0th$ generation Michelson TDI observable about spacecraft 1, denoted $X0$:

$$y_{X0} = y_{13'2,3} + y_{231} - y_{32'1} - y_{123,2'} \ . \tag{3.16}$$

This combination of one-way Doppler measurements gives zero laser noise in a LISA that is equal-arm, non-rotating and static. A notation is used here in which each number to the right of a comma in the subscript denotes a time-independent time-delay along the link indicated by the number. Since the delays are time-independent here, the order in which they are written after the comma does not matter. For example, if $y_{13'2}$ is evaluated at time t, then $y_{13'2,32'}$ is evaluated at time $t, 32' \equiv t - L_{2'} - L_3 = t - L_3 - L_{2'} \equiv t, 2'3$. It can be seen explicitly that there is no laser noise in $X0$:

$$\begin{aligned}
y^{laser}_{X0} &= y^{laser}_{13'2}(t - L_3) + y^{laser}_{231}(t) - y^{laser}_{32'1}(t) - y^{laser}_{123}(t - L_{2'}) && (3.17)\\
&= y^{laser}_1(t - L_3 - L_{3'}) - y^{laser}_2(t - L_3) + y^{laser}_2(t - L_3) - y^{laser}_1(t) \\
&\quad - \left[y^{laser}_3(t - L_{2'}) - y^{laser}_1(t)\right] - \left[y^{laser}_1(t - L_{2'} - L_2) - y^{laser}_3(t - L_{2'})\right] && (3.18)\\
&= y^{laser}_1(t - 2L) - y^{laser}_2(t - L) + y^{laser}_2(t - L) - y^{laser}_1(t) \\
&\quad - \left[y^{laser}_3(t - L) - y^{laser}_1(t)\right] - \left[y^{laser}_1(t - 2L) - y^{laser}_3(t - L)\right] && (3.19)\\
&= 0
\end{aligned}$$

In (3.17), the time arguments of the links in (3.16) are written out explicitly. Eq.(3.18) is obtained by substituting (3.15) in (3.17). In (3.19), the equal-arm and non-rotating assumptions are used.

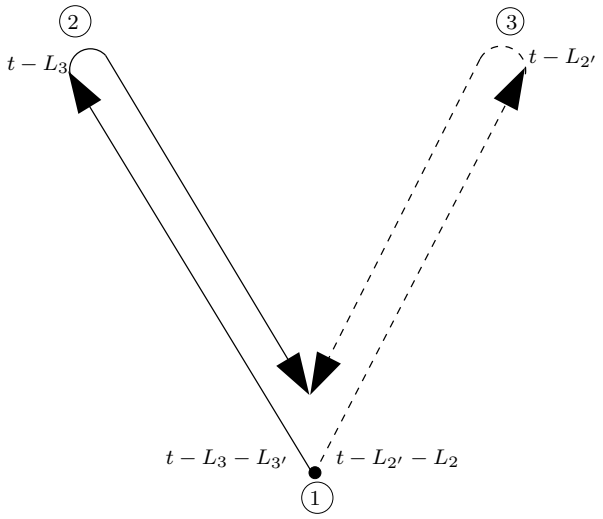

FIGURE 3.3: Pictorial representation of the TDI observable X0. Each arrow represents a interspacecraft measurement. The numbers at 1 2 and 3 denote the spacecrafts, or optical benches. Each arrow represents a link measurement, where the arrowhead is the event of reception, and the tail is the event of emission of the laser transmission of the link. The times at these events are also indicated. The dashed arrows are summed with the opposite sign to the solid arrows.

Another way of representing the TDI observable $X0$ is to draw a picture of the spacecrafts and the links that are involved, as shown in Fig.(3.3). In this geometrical interpretation, each link, or its laser beam, is represented by an arrow, with the tail matching the emission event, and the head matching the reception event. Every emission and reception event is given a time stamp. For those arrows that are dashed, the algebraic symbols of the links they represent take on an opposite sign to those represented by solid arrows in the total sum. Therefore, to translate Fig.(3.3) into (3.16), write down the corresponding y_{slr} for each arrow, evaluate each at time t_r, the reception time, and then gather them in a sum, with the solid arrows taking a plus sign, and the dashed arrows taking a minus sign, or vice versa. $X0$, as its name suggests, is analogous to a Michelson interferometer. The two solid arrows join to form a virtual beam which travels down one arm and comes back to meet the virtual beam made up of the two dashed arrows, at the 'photodectector', spacecraft 1 in this case.

In comparing the arrow, as a link measurement, with (3.15), it is possible to attach the laser noise at the emission event, $y_{slr}^{laser}(t - L_l(t))$, to the arrowtail, and the laser noise at the reception event with a *minus* sign, $-y_{slr}^{laser}(t)$, to the arrowhead. Taking this sign difference into account , it is clear that there are three occasions where two laser noise terms cancel each other:

Head-tail When the head of a solid arrow meets the tail of a soild arrow, or when the head of a dashed arrow meets the tail of a dashed arrow.

Head-head When the head of a solid arrow meets the head of a dashed arrow.

Tail-tail When the tail of a solid arrow meets the tail of a dashed arrow.

Note that in order to meet, the heads or tails need to correspond to events with the same time stamp. In Fig.(3.3), there are four such encounters at which a pair of laser noise terms cancel each other: tail-tail at time $t - 2L$, head-tail(solid arrows) at time $t - L$, head-tail(dashed arrows) at time $t - L$ and head-head at time t. It is this observation of laser noise cancelling in pairs that forms the basis of the *Geometric TDI Principle* [2] which says that for a TDI observable, the arrows must line up head to head, tail to tail, or head to tail, forming a closed loop in a way such that laser noise terms cancel each other at all emission and reception events. There can be no arrowhead, or tail, that is left open, because there would be no other laser noise term to cancel it with. Another important observation to be made is that, by tracing along the closed loop once in any given direction, one goes through a sequence of light travel time-intervals of the links. In order for the events at the starting point and end point of such a round trip to have a matching time stamp(so that laser noises cancel), it is necessary that these light travel time-intervals add up to zero. So the loops have to be closed in time as well. This forms the basis for the *Feynman-Wheeler Geometric TDI Principle* [2], which says that 'any $2N$-beam Geometric-TDI closed loop can be seen as a single beam that travel forward and backward in time to meet itself back at its origin.' It is easy to see that $X0$ obeys this principle. In Fig. 3.3, suppose the starting point of the virtual beam is at time $t - 2L$ at spacecraft 1. Then, if the beam travels into the future as it goes along the solid arrows and travels backward in time as it goes along the dashed arrows, the net time lapse, when it returns to the starting point, is $L + L + (-L) + (-L) = 0$. The two $-L$ terms correspond to the dashed arrows, against which the beam travels backward in time. Note that the terms in the sum in (3.19) can be rearranged into a different order, so any point in the loop can be the starting point.

There is another way to represent TDI observables, which is more general and concise. It is also very helpful in explaining the method with which TDI observables are obtained. In a *combinatorial enumeration* [2], each link measurement, or arrow, is either denoted by a $\overrightarrow{l}$ or $\overleftarrow{l}$, instead of a solid arrow or a dashed arrow in the geometrical representation. For example, $X0$ is denoted, in this enumeration, by

$$X0 \ : \ \overrightarrow{3'3}\,\overleftarrow{2'2} \tag{3.20}$$

An expression like this, $\overrightarrow{l}$ or $\overleftarrow{l}$ is often called a *string*. Since it is normal to read from left to right, the virtual TDI beam is taken to travel in that direction by default; overhead right arrows make up the part of the TDI beam that travels forward in time; the overhead left arrows make up the part that travels backward in time. As it is, (3.20) only encodes, for the TDI observable, the links that are involved, their direction in time, and their time stamps *relative* to one another. It *does not* say what time stamp each link has. However, the string contains information about how their time stamps relate to each other. Once a time stamp is assigned to a particular event along the beam, the time stamp of every other event can be

found by adding up the sequence of light travel time intervals along the virtual beam which originates from the first event to be given a time stamp to itself. For example, in (3.20), if the reception time of $\overleftarrow{2}$ is at time t, then, to find the emission time of $\overrightarrow{3'}$, a virtual beam is drawn from the former event to the latter, going from right to left. The first string it goes past is $\overleftarrow{2'}$, and since it is in the same direction as the beam, a time interval of $L_{2'}$ is accumulated, giving $\overleftarrow{2'}$ a reception time of $t + L_{2'}$. Next, the beam goes past the string $\overrightarrow{3}$, and since this string is in the opposite direction as the beam, a time interval of $-L_3$ is accumulated, giving $\overrightarrow{3}$ an emission time of $t + L_{2'} - L_3$. By applying this once more to the string $\overrightarrow{3'}$, one finds that its emission time is $t + L_{2'} - L_3 - L_{3'}$.[3]

Using the Feynman-Wheeler Geometric TDI principle, it is possible to set criteria for identifying strings which are closed in time for LISA models of different complexities described above. These strings are the combinatorial enumerations of TDI observables. There are four different generations of TDI observables, each corresponding to a LISA model of a different complexity, and their criteria are justified as follows:

G0 LISA is equal-arm, non-rotating and static. This is the simplest model of LISA. In this case, the light travel time along all links is equal to the nominal LISA arm length L. If there are as many links that travel forward in time as there are that travel backward in time, there is time-closure. Any string of links that obeys this criterion is said to be *closed*, and is a TDI observable of 0th Generation.

G1 LISA is unequal-arm, non-rotating and static. Because the arm lengths are different, light travel times along different links are different. It is now necessary to distinguish different link numbers. Given a string of links, if there are equal numbers of l , whether primed or unprimed, travelling forward in time as there are travelling backward in time for all l in $\{1, 2, 3\}$, there is time-closure. The string is said to be $|L|$-*closed*, and is a TDI observable of 1st Generation.

Gm LISA is unequal-arm, rotating and static. Now, due to the Sagnac effect from the rotation, different link numbers and as well as their primedness need to be distinguished. Given a string of links, if there are equal numbers of l travelling forward in time as there are travelling backward in time for all l in $\{1, 2, 3, 1', 2', 3'\}$, there is time-closure. The string is said to be L-*closed*, and is a TDI observable of the *Modified* Generation.

G2 LISA is unequal-arm, rotating and non-static. As well as being unequal, the arm lengths now change with time, so $L_l(t)$ are now functions of time. There is no perfect time-closure for such a model. But, it is possible to obtain time-closure to first order/degree[4] in $L_l(t)$, which brings the laser noise down to a sufficiently low level. The closure criterion for this is simply stated here. (For a more detailed explanation, see Appendix F.) Consider

[3] Notice the time stamps of events along the way are found as well.
[4] This means that any $\dot{L}^2$ and $\ddot{L}$ terms are ignored.

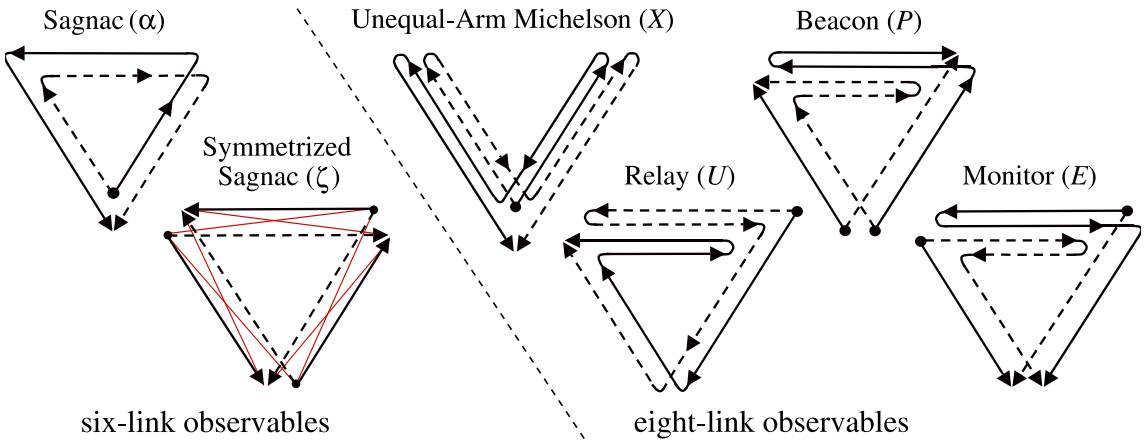

FIGURE 3.4: Different types of TDI observables. In this thesis we will only work with the Michelson and Sagnac observables. (Taken from Ref. [2])

a TDI beam that goes from the leftmost point to the rightmost point of a given string $l_1 l_2 \cdots l_k \cdots l_N$; the overhead arrows are omitted since they can be either left-pointing or right-pointing in general. For each link $\overrightarrow{l_k}$ that travels forward in time, collect the set $\{l_k l_k\ ,\ l_k l_{k+1}\ ,\ \cdots\ ,\ l_k l_N\}$. For each link $\overleftarrow{l_k}$ that travels backward in time, collect the set $\{l_k l_{k+1}\ ,\ \cdots\ ,\ l l_N\}$. After this is done for every link, collect the sets obtained together, and add the overhead arrows to each link as they appear in the string. Now take $\overrightarrow{l}\,\overrightarrow{l}$ and $\overleftarrow{l}\,\overleftarrow{l}$ to be equivalent, and that each cancels either $\overrightarrow{l}\,\overleftarrow{l}$ or $\overleftarrow{l}\,\overrightarrow{l}$, which are also to be taken as equivalent. If elements of the above collected set all pair up and cancel out, the string is said to be $\dot{L}$-closed, and is a TDI observable of the 2nd Generation.

In practice, TDI observables of a given length N can be found by searching through the space of all strings of length N, and identifying those which satisfy the various criteria listed above.[5]

There are several different types of TDI observables, and they are shown in Fig. 3.4. The most often used are the Michelson and the Sagnac types. Michelson type observables are analogous to the Michelson interferometer. They can be regarded as consisting of two virtual beams, which originate from, and meet back at some spacecraft. Higher generations of them are like Michelson interferometers in which the split beams are sent down the two arms alternately one or more times. The spacecraft at which the two virtual beams meet is analagous to the photodetector in a Michelson interferometer. Conventionally, if it is spacecraft 1, the TDI observable is named X, if it is spacecraft 2, Y, and if it is spacecraft 3, Z. In this work, the number following the TDI observable's name is its generation number. For example, $Z2$ is a second generation Michelson TDI observable centred about spacecraft 3. $X2$'s combinatorial enumeration is

$$X2\ :\ \overrightarrow{3'322'22'3'3}\ \overleftarrow{2'233'33'2'2} \tag{3.21}$$

[5] This method can find TDI observables of all the generations, but was only used in [2] to find the second generation observables. Those of lower generations were already found in other algebraic studies of TDI observables, e.g, [10].

Being $\dot{L}$-closed, this, along with $Y2$ and $Z2$, are important observables in realistic LISA models.

Sagnac observables also consist of two virtual beams which originate from a spacecraft and come back to meet, but they circle LISA's triangle in opposite directions, one or more times, before coming back to meet at the spacecraft. Conventionally, α, β and γ are used to denote Sagnac TDI observables whose two virtual beams originate from spacecraft 1, spacecraft 2 and spacecraft 3, respectively. There is a special Sagnac observable called the *symmetrised Sagnac*, often denoted ζ, which has a small GW response at low frequencies. α, β, γ and ζ form a basis for TDI observables [10]. The modified generation α is combinatorially enumerated as

$$\alpha\mathrm{m} \ : \ \overrightarrow{3'1'2'213} \ \overleftarrow{2'1'3'212} \tag{3.22}$$

This is actually not a $\dot{L}$-closed TDI observable.[6] However, due to their specific combination, the residual $\dot{L}_l(t)$ terms' contribution is small for the LISA orbit. Therefore, αm, βm and γm are also important observables for realistic LISA models.

There are several useful properties of the TDI observables which are worth mentioning. (i) Firstly, it is seen that TDI observables X, Y and Z differ only in the spacecraft about which they are defined. Because of this, given the combinatorial enumeration of one of them, the combinatorial enumerations of the rest can be found from cyclic permutations of all the spacecraft numbers and all the link numbers: $1 \to 2 \to 3 \to 1$. The same goes for α, β and γ.[7] (ii) Often, a TDI observable can be expressed in terms of those of a lower generation. For example, for $X1$: $\overrightarrow{3'322'}\overleftarrow{33'2'2}$, an algebraic form of it is

$$y_{X1} = y_{X0;22'} - y_{X0} \tag{3.23}$$

It can be seen here that $X1$ is basically a *self-splicing* of $X0$ with its time-shifted reversal. This is in fact true for all Michelson and Sagnac observables. In an equal-arm, non-rotating and static LISA, this property greatly simplifies the calculation of the GW response and noise response of higher generation TDI observables. Taking the Fourier transform of (3.23),

$$\tilde{y}_{X1}(f) = \left(e^{i2\pi f(-2L)} - 1\right) \ \tilde{y}_{X0}(f) \tag{3.24}$$

As shown in (3.24), the Fourier transform of $y_{X1}(t)$ is simply the Fourier transform of $y_{X0}(t)$ multiplied by a factor determined by the amount of time-shift in the self-splicing. So, once the GW or noise response is known for $X0$, (3.24) can be used to quickly obtain those of higher TDI generations.

[6] Even 2nd generation analogs of these do not cancel laser noise to first order/degree in $\dot{L}$.

[7] This is useful for checking whether a combinatorial enumeration has been correctly translated into its algebraic expression.

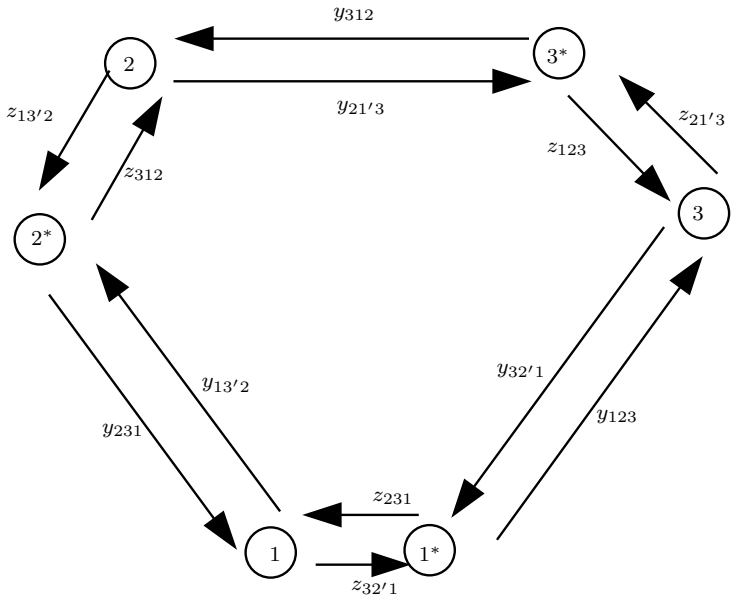

FIGURE 3.5: Schematic showing all of LISA's basic measurements. The long arrows are the interspacecraft measurements, and the short arrows are the intraspacecraft measurements.

3.4.1 Extension of TDI observables to intraspacecraft measurement

Perhaps the easiest way to see how what has been obtained so far can be generalised to the case where there are two, instead of one, optical benches in each spacecraft, is to view the intraspacecraft measurements as arrows in the geometrical interpretation. In Fig. 3.5 are shown all six interspacecraft and intraspacecraft measurements. The arrows of the intraspacecraft measurements are the same as those of the interspacecraft measurements in that the laser noise term at the arrowhead and the laser noise term at the arrowtail have opposite signs.[8] Because of this, pairs of laser noise terms still cancel each other for the same situations described above for interspacecraft measurements, and so intraspacecraft measurements' arrows can be used in lining up the closed loops of a TDI observable in the same way as the interspacecraft measurements' arrows can be used. They have to be used whenever the TDI observable's virtual beam traverses a spacecraft, otherwise, there would be interspacecraft arrowheads and arrowtails which are left open, leaving laser noise terms un-canceled.[9] In the $X0$ example in Fig. 3.3, at spacecraft 2 and 3, the virtual beam is simply 'reflected' off optical bench 2^* and 3, respectively, so intraspacecraft links are not involved in either case. At time $t - 2L$, two virtual beams diverge from spacecraft 1, but because one goes from optical bench 1 and the other goes from optical bench $1*$, intraspacecraft measurements need to be inserted. Similarly, intraspacecraft links are also required at time t, in order for the two virtual beams to meet.

[8] The only difference is that they do not represent a light travel time interval; the transmission time from one optical bench to its adjacent optical bench is negligible.

[9] This is equivalent to saying that a pair of laser noise terms that used to cancel each other no longer do, because where they used to be from the same, and only, optical bench in some spacecraft, they are now from two adjacent, hence different, optical benches.

The resulting algebraic expression for $X0$ now becomes[10]

$$y_{X0} = y_{13'2,3} + y_{231} - y_{32'1} - y_{123,2'} + \frac{1}{2}(z_{32'1} - z_{231}) + \frac{1}{2}(z_{231,3'3} - z_{32'1,22'}) \tag{3.25}$$

3.4.2 GW response of TDI observables

We saw in the previous section that a TDI observable can be expressed algebraically as a sum of interspacecraft measurements and intraspacecraft measurements, each appropriately time-delayed, and multiplied by either a $+1$ or -1. Its GW response can be found by subsituting in the GW response of each interespacecraft measurement in the sum. The intraspacecraft measurements can be ignored because they have negligible GW response.[11] Hence, if I denotes a TDI observable, then its GW response can be written as

$$y_I^{GW} = \sum_{(l)}^{I} \epsilon_{(l)}\, y_{(l)}^{GW} \tag{3.26}$$

The (l)'s here denote the links that make up the TDI observable. For each link (l), $\epsilon_{(l)} = \pm 1$; $y_{(l)}^{GW} = y_{s(l)r}^{GW}$, with its own emission time $t_{s,(l)}$ and reception time $t_{r,(l)}$. By choosing a common reference event[12], $(t_I, \vec{p}_I)$, for the links, (3.26) can be written as

$$y_I^{GW}(t_I, \vec{p}_I) = \sum_{(l)}^{I} \epsilon_{(l)}\, y_{(l)}^{GW}(t_I, \vec{p}_I) \tag{3.27}$$

$$= F_I^A(f, \hat{k}, \psi; t_I, \vec{p}_I)\; h_A(f, \hat{k}, \psi)\; e^{i2\pi f(t_I + \hat{k}\cdot\vec{p}_I)}\ , \tag{3.28}$$

where I's response function is given by

$$F_I^A(f, \hat{k}, \psi; t_I, \vec{p}_I) = \sum_{(l)}^{I} \epsilon_{(l)}\; F_{(l)}^A(f, \hat{k}, \psi,\, t_{r,(l)}\,; t_I, \vec{p}_I)\ . \tag{3.29}$$

In obtaining (3.28), (3.5) is substituted into (3.27). The response functions of the links in the RHS of (3.29) are given by (3.6). As an example, by applying (3.29) to $X0$, as defined in (3.25), its response function for an equal-arm, non-rotating and static LISA, when $(t, \vec{p}_1(t))$ is

[10] In practice, whenever an intraspacecraft measurement is used, the other one in the same spacecraft(in the opposite direction) is also used. Their total contribution is halved to produce essentially the same laser noise cancellation. The reason for using this is that this combination also cancels proof-mass noises for a head-head meeting between two interspacecraft measurements.

[11] An intraspacecraft measurement is like an interspacecraft measurement with emission and reception event the same, so there is zero light travel time between the two.

[12] Usually, for Sagnac(α,β,γ) and Michelson(X,Y,Z) observables, the reference event is chosen to be the event at which the two virtual beams(one going forward in time and one going backward in time) meet. So, if the reference time is t, then the reference event for α and X is $(t, \vec{p}_1(t))$; for β and Y, it is $(t, \vec{p}_2(t))$; for γ and Z, $(t, \vec{p}_3(t))$. $\vec{p}_1$, $\vec{p}_2$ and $\vec{p}_3$ are the positions of spacecraft 1, 2 and 3, respectively.

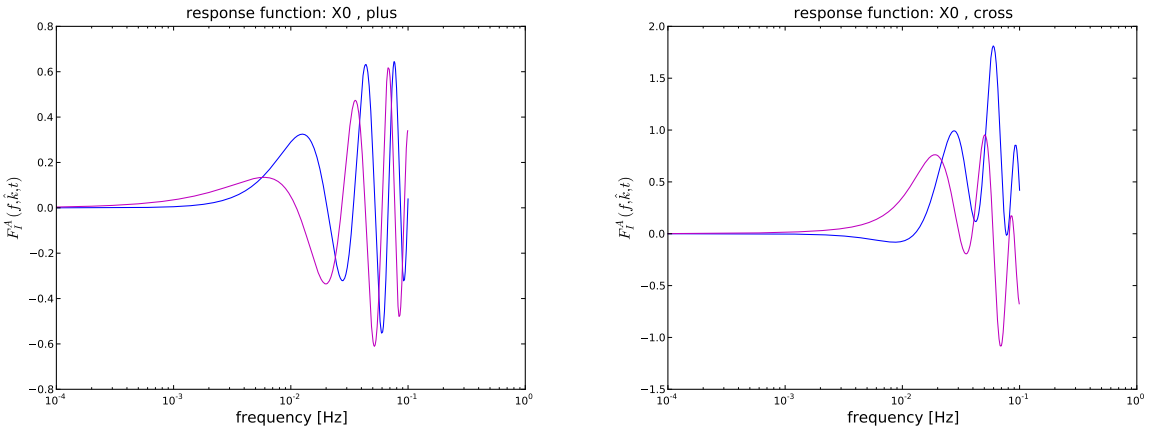

FIGURE 3.6: GW response functions for $X0$ to a plane GW with $(\phi, \theta, \psi) = (89\,°, 120\,°, -53\,°)$ for plus polarisation (left) and cross polarisation (right). The blue line is the real part, and the pink line is the imaginary part.

the reference event, is

$$F^A_{X0}(f, \hat{k}; t, \vec{p}_1(t)) = \mathrm{e}^A_{ab}(\hat{k}, \psi) \left(\hat{n}^a_{3'} \hat{n}^b_{3'} \, \mathcal{T}(f, \hat{k} \cdot \hat{n}_{3'}) - \hat{n}^a_2 \hat{n}^b_2 \, \mathcal{T}(f, \hat{k} \cdot \hat{n}_2) \right) , \tag{3.30}$$

where

$$\begin{aligned} \mathcal{T}(f, \hat{k} \cdot \hat{n}_l) = -i \frac{f}{2f_l} \, e^{-i \frac{f}{f_l}} \Bigg[& \mathrm{sinc} \left(\frac{f}{2f_l} (1 + \hat{k} \cdot \hat{n}_l) \right) e^{-i \frac{f}{2f_l} (1 - \hat{k} \cdot \hat{n}_l)} \\ & + \mathrm{sinc} \left(\frac{f}{2f_l} (1 - \hat{k} \cdot \hat{n}_l) \right) e^{-i \frac{f}{2f_l} (-1 - \hat{k} \cdot \hat{n}_l)} \Bigg] . \end{aligned} \tag{3.31}$$

$f_l = 1/(2\pi L)$ is LISA's transfer frequency; it is the same for all links in this case. Note that in the RHS, there is no time-dependence, because LISA is static here, so all the spacecraft positions and all the unit link vectors remain constant. Keeping this LISA model, F^A_{Xm} and F^A_{X2} can be found by simply using (3.24). It can be seen from (3.30) that, even for a simple TDI observable, the response function is a rather complicated expression. In the low frequency limit where $f \ll f_l$, $\mathcal{T}$ in (3.31) tends to 1, and the response function in (3.30) becomes

$$F^A_{X0}(f, \hat{k}; t, \vec{p}_1(t)) = \mathrm{e}^A_{ab}(\hat{k}, \psi) \left(\hat{n}^a_{3'} \hat{n}^b_{3'} - \hat{n}^a_2 \hat{n}^b_2 \right) . \tag{3.32}$$

In Fig. 3.6, the response function for $X0$ in (3.30) is plotted as a function of frequency. In Fig. 3.7, the same response function is plotted as a function of directions in the sky, at $f = 10^{-2}$Hz.

3.4.3 Noise response of TDI observables

The noise response of TDI observables are found by substituting the noise responses of inter-spacecraft links and intraspacecraft links, (3.13) and (3.14), into the TDI observable's algebraic

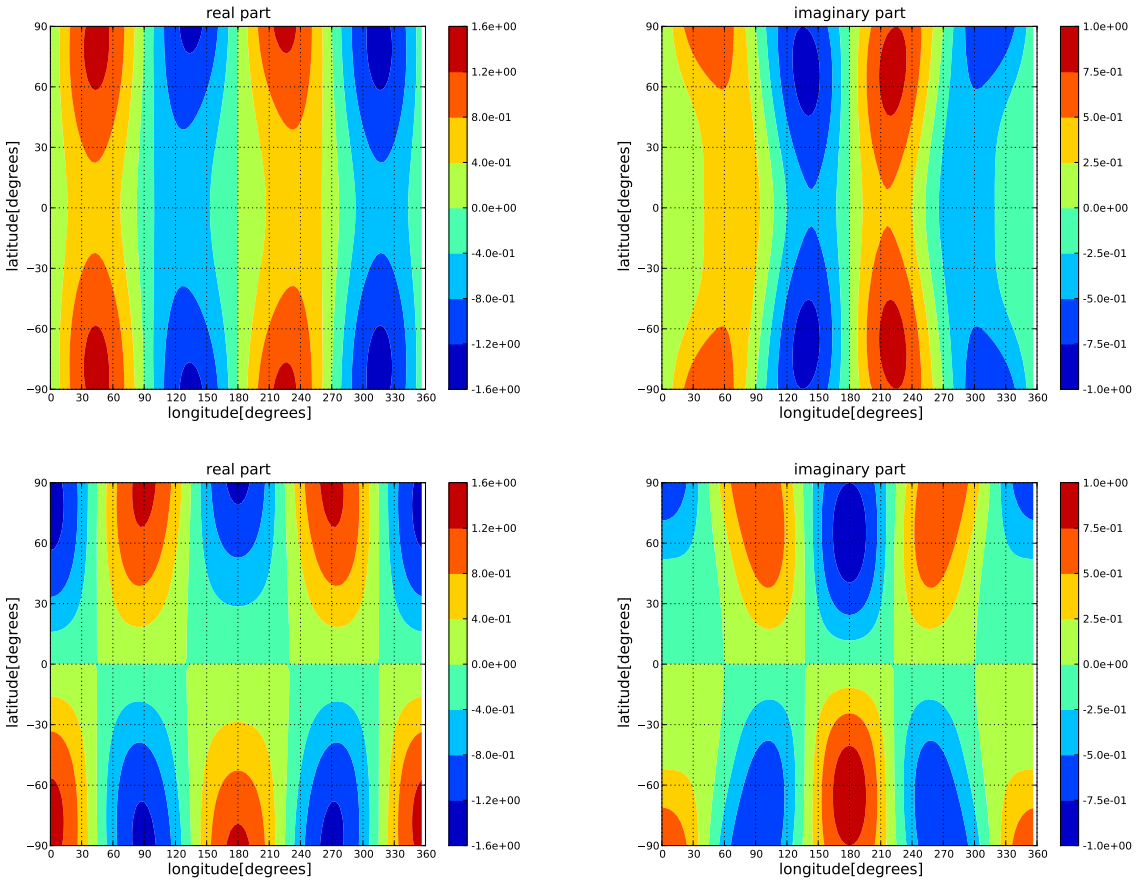

FIGURE 3.7: GW response function of $X0$ to plane GW, as a function of ecliptic longitude and ecliptic latitude, at $f = 10^{-2}$Hz for plus polarisation(top row) and cross polarisation(bottom row).

expression. For example, the noise response of $X0$ in the frequency domain, found by substituting (3.13) and (3.14) into (3.25), is

$$\begin{aligned}\tilde{y}^{noise}_{X0}(f) = e^{-i2\pi fL}\,\tilde{y}^{op}_{13'2} - e^{-i2\pi fL}\,\tilde{y}^{op}_{123} + \tilde{y}^{op}_{231} - \tilde{y}^{op}_{32'1} \\ -2e^{-i2\pi fL}\,\tilde{y}^{pm}_{2*} + 2e^{-i2\pi fL}\,\tilde{y}^{pm}_{3} - \left(1 + e^{-i4\pi fL}\right)\tilde{y}^{pm}_{1} - \left(1 + e^{-i4\pi fL}\right)\,\tilde{y}^{pm}_{1*}\end{aligned} \tag{3.33}$$

As mentioned at the end of Sec. 3.4, $\tilde{y}^{noise}_{Y0}$ and $\tilde{y}^{noise}_{Z0}$ can be found from (3.33) via cyclic permutations of the spacecrafts' and links' indices. Since $\langle \tilde{y}^*_I(f)\tilde{y}_J(f)\rangle = \delta(f - f')P_{IJ}(f)$, the auto and cross-spectral densities of $X0$, $Y0$ and $Z0$ can be found from these:

$$P^{noise}_{X0X0}(f) = P^{noise}_{Y0Y0}(f) = P^{noise}_{Z0Z0}(f) \tag{3.34}$$

$$= 4\,[P^{op}(f) + (3 + \cos 4\pi fL)\,P^{pm}(f)] \tag{3.35}$$

$$P^{noise}_{X0Y0}(f) = P^{noise}_{Y0Z0}(f) = P^{noise}_{Z0X0}(f) \tag{3.36}$$

$$= -2\cos(2\pi fL)\,[P^{op}(f) + 4P^{pm}(f)]\ , \tag{3.37}$$

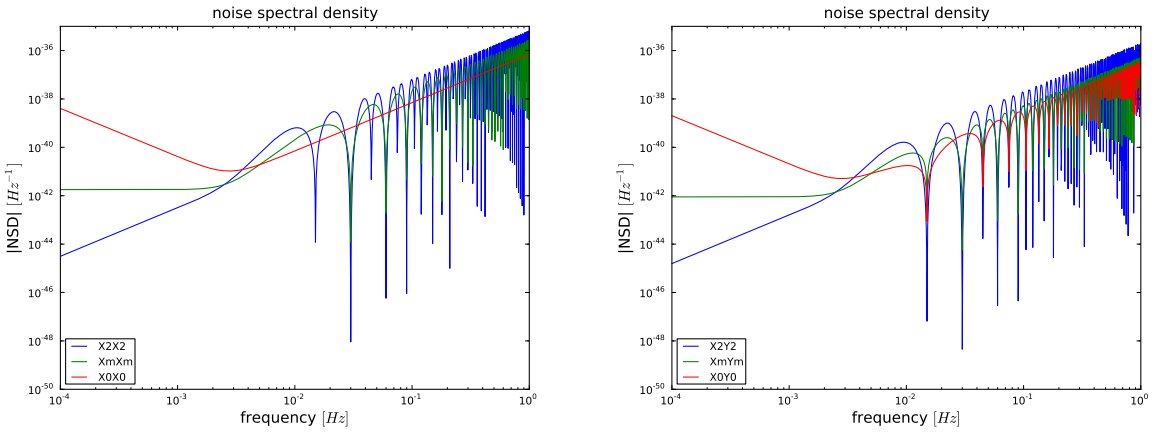

FIGURE 3.8: Noise spectral densities of the Michelson observables of generations $G0$, Gm and $G2$. Left panel: auto-spectral densities as given by (3.35) Right panel: cross-spectral densities as given by (3.37)

where P^{op} and P^{pm} are the power spectral densities of the optical path noise and the proof-mass noise in (3.11) and (3.12). In obtaining (3.34) and (3.36), the symmetry between $X0$, $Y0$ and $Z0$ is used. By using the fact that higher generation TDI observables can be interpreted as self-splicings of lower generation ones, it can be shown, by using expressions like (3.24), that

$$P^{noise}_{ImJm}(f) = 4\,\sin^2(2\pi f L)\,P^{noise}_{I0J0}(f) \tag{3.38}$$

$$P^{noise}_{I2J2}(f) = 16\,\sin^2(2\pi f L)\sin^2(2\pi f L)\,P^{noise}_{I0J0}(f)\ , \tag{3.39}$$

where I and J can be any of of Michelson observables X, Y or Z.

In Fig. 3.8, the auto- and cross- spectral densities in (3.35) and (3.37) are plotted. The higher generation ones are obtained from (3.38) and (3.39).

3.4.4 Optimal TDI observables

It turns out that the TDI observables mentioned so far can be further combined into what are called *optimal TDI observables*. The advantage of these optimal observables is that their noises are *uncorrelated*. Such observables can be formed from Michelson type observables or Sagnac type observables.[13] For the Michelson observables, they are typically chosen to be

$$A = \frac{1}{3}\,(2X - Y - Z) \tag{3.40}$$

$$E = \frac{1}{\sqrt{3}}\,(Z - Y) \tag{3.41}$$

$$T = \frac{1}{3}\,(X + Y + Z)\ . \tag{3.42}$$

[13] The optimal observables are found by diagonalising the original noise covariance matrix in terms of the Sagnacs (or the Michelsons). The eigenvectors found this way correspond to the optimal observables.

And for the Sagnac observables, they are

$$\bar{A} = \frac{1}{\sqrt{2}}(\gamma - \alpha) \tag{3.43}$$

$$\bar{E} = \frac{1}{\sqrt{6}}(\alpha - 2\beta + \gamma) \tag{3.44}$$

$$\bar{T} = \frac{1}{\sqrt{3}}(\alpha + \beta + \gamma)\ . \tag{3.45}$$

The definition[14] applies to all generations. Because only TDI observables are involved in the combination, the optimal TDI observables are also laser noise-free. Implicitly, but importantly, the above definitions imply that the time-stamps of the events, at which the two virtual beams(one going forward in time and one going backward in time) meet in all Michelson and Sagnac observables, should be the same.

3.4.5 GW response of optimal TDI observables

Similar to sec. 3.4.2, by choosing a common reference event for all the constituent TDI observables, the GW response of an optimal TDI observable can be written in the same form as that in (3.28). Even though the choice of a reference event is still arbitrary, there is not a natural choice here, like there is for the Sagnac and Michelson observables, because an optimal TDI observable cannot be viewed as consisting of only two virtual beams, with one going forward in time and one going backward in time. Suppose $(t, \vec{p}_1(t))$ is chosen as the reference event for A, then its GW response to a plane wave of (θ, ϕ, ψ) can be written as

$$y_A^{GW}(t, \vec{p}_1(t)) = \frac{1}{3}\left[\, 2\, y_X^{GW}(t, \vec{p}_1(t)) - y_Y^{GW}(t, \vec{p}_1(t)) - y_Z^{GW}(t, \vec{p}_1(t)) \,\right] \tag{3.46}$$

$$= F_A^A(f, \hat{k}, \psi; t, \vec{p}_1(t))\, h_A(f, \hat{k}, \psi)\, e^{i2\pi f(t + \hat{k}\cdot\vec{p}_1(t))}\ , \tag{3.47}$$

where its response function is

$$F_A^A(f, \hat{k}, \psi; t, \vec{p}_1(t)) = \frac{1}{3}\left[\, 2\, F_X^A(f, \hat{k}, \psi; t, \vec{p}_1(t)) - F_Y^A(f, \hat{k}, \psi; t, \vec{p}_1(t)) - F_Z^A(f, \hat{k}, \psi; t, \vec{p}_1(t)) \,\right]\ . \tag{3.48}$$

Notice here that, except for X, the response functions of other TDI observables in the RHS are *not* evaluated at their natural reference events. As an example, with this event chosen as reference, the GW response function for $A2$ is plotted in Fig. 3.9 as a function of frequency, and projected in the sky in Fig. **??**.

[14] There are variations in these definitions in the literature, but in all of them, the optimal TDI observables have uncorrelated noise.

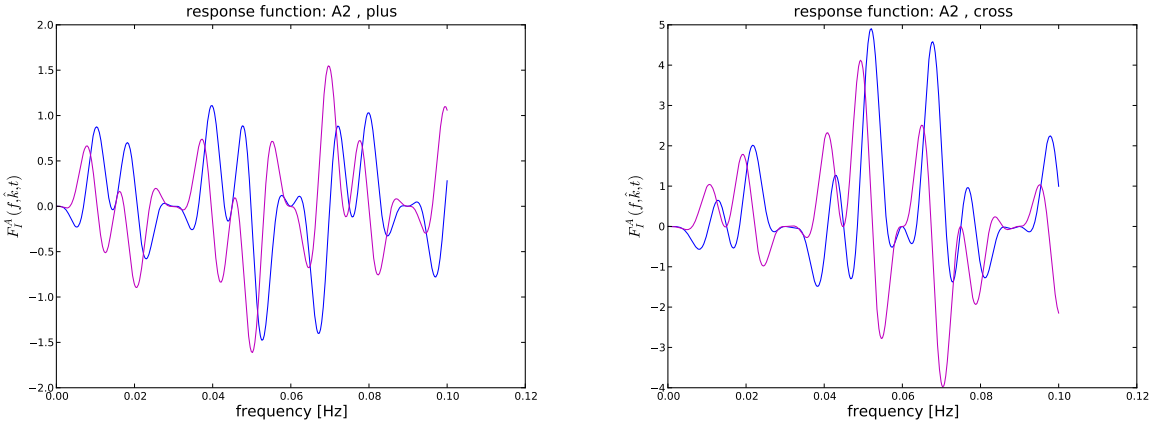

FIGURE 3.9: GW response function of $A2$ to a plane GW of $(\phi, \theta, \psi) = (89\,^\circ, 120\,^\circ, -53\,^\circ)$ for plus polarisation(left) and cross polarisation(right). (The blue line is the real part, and the pink line is the imaginary part.)

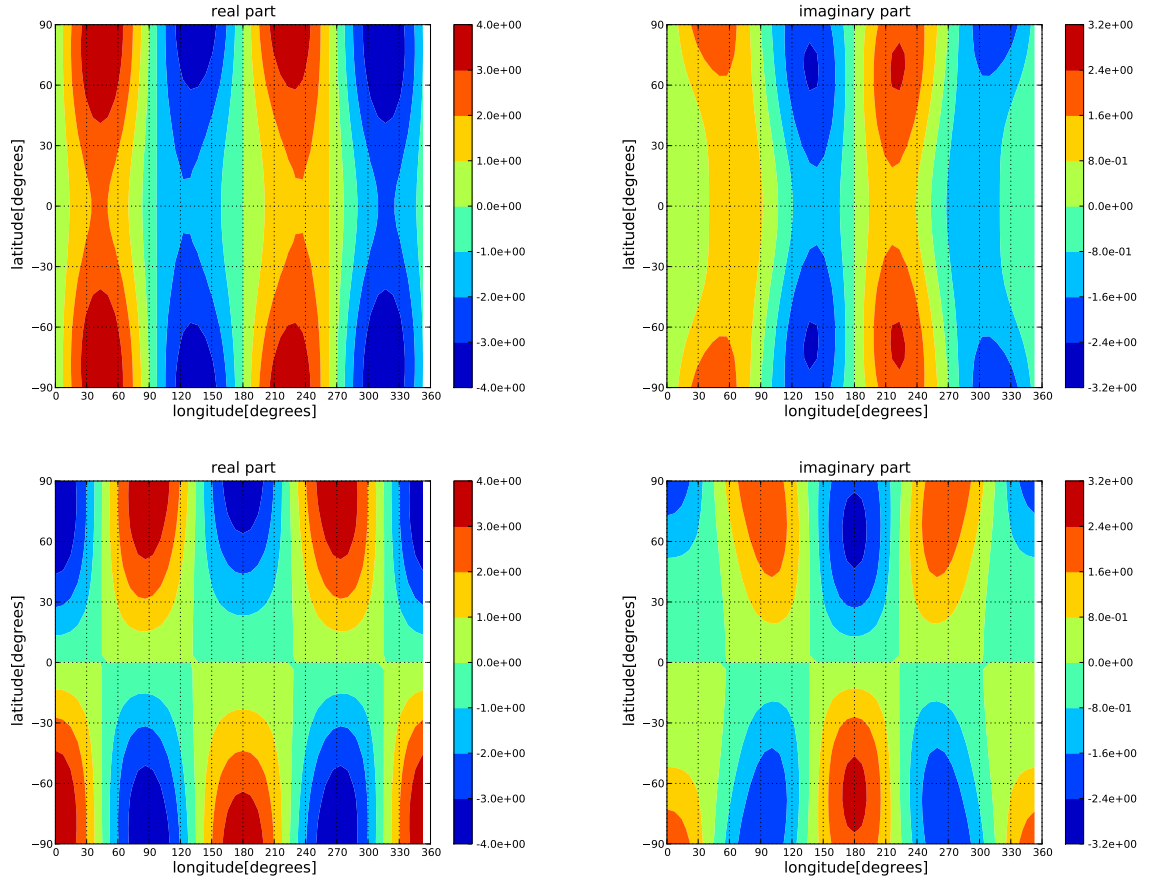

FIGURE 3.10: GW response function of $A2$ to plane GW, as a function of ecliptic longitude and ecliptic latitude, at $f = 10^{-2}$Hz for plus polarisation(top row) and cross polarisation(bottom row).

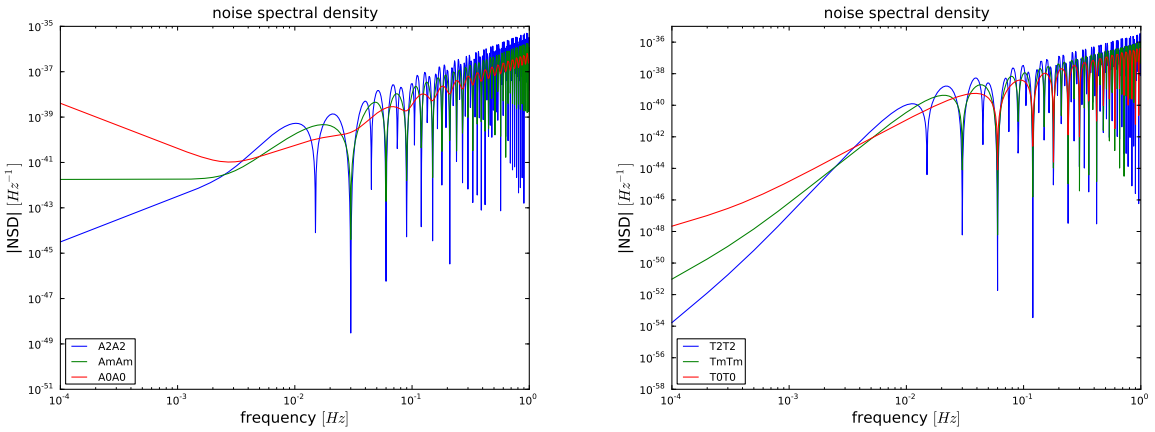

FIGURE 3.11: Noise spectral densities of A and T, for $G0$, Gm and $G2$ in the left and right panel, respectively. The noise spectral density of E is the same as that for A.

3.4.6 Noise response of optimal TDI observables

The noise response of the optimal TDI observables can be found from (3.40) to (3.45) by substituting in the noise responses of the TDI observables in the RHS of these. For the Michelson observables, by using (3.34) to (3.37), it can be shown that the noise auto- and cross-spectral densities for A, E and T are

$$P^{noise}_{A0A0}(f) = P^{noise}_{E0E0}(f) \tag{3.49}$$

$$= \frac{4}{3}\left((2+\cos 2\pi fL)P^{op}(f) + 2(3+2\cos(2\pi fL)+\cos(4\pi fL))P^{pm}(f)\right) \tag{3.50}$$

$$P^{noise}_{T0T0}(f) = \frac{8}{3}\sin^2(\pi fL)\,\left(P^{op}(f) + 2\left(1-\cos(2\pi fL)\right)P^{pm}(f)\right) \tag{3.51}$$

$$P^{noise}_{A0E0}(f) = P^{noise}_{E0T0}(f) = P^{noise}_{A0T0}(f) \tag{3.52}$$

Notice that the power spectral densities of A and E are equal. This is also true for $\bar{A}$ and $\bar{E}$. The cross-spectral densities are zero as expected. Because A, E and T are linear combinations of the Michelson observables, higher generations of them are related to lower generations of them through self-splicing in the same way as among the Michelson observables. Therefore, (3.38) and (3.39) can also be used to obtain power spectral densities of A, E and T of Gm and $G2$ generations.[15] In Fig. 3.11 are plotted the power spectral densities described (3.50) and (3.51) and their higher generation counterparts.

[15] There are similar expressions for the Sagnac observables and their optimal observables. But because the self-splicing is different from the Michelson observables, these expressions are also different.

Chapter 4

Stochastic Gravitational Wave Backgrounds

A *stochastic gravitational wave background* is usually described as the superposition of gravitational waves from a large number of sources. In fact, without selection, this would be from *all* the sources that have emitted waves that reach us today. These sources can be divided into different types, such as coalescing binary systems, pulsars and supernovae, etc., which can differ in their evolution, and in the general characteristics of their waves. Different systems of the same type can also differ, for example, in the frequency, duration and direction of wave propagation, etc. The most general stochastic background consists of all these different waves.

A gravitational wave detector is designed to be able to 'see' gravitational waves only within some frequency band, and some types of sources up to certain distances (or redshifts). In the language of wave detection, this frequency band can be divided into bins of width equal to the frequency resolution of the detector. A source may start off emitting gravitational waves at some frequency, which puts its signal in the corresponding frequency bin. As it evolves and the frequency changes, the signal moves to different frequency bins. For example, as an inspiralling binary system evolves, so its signal spends more time in lower frequency bins than in higher frequency bins. With many sources contributing to a stochastic background, more than one signal can be present in the frequency bin at a given time. If there is only one signal present in a particular bin at some given time, the signal is said to be *resolvable* by the detector at the frequency of that bin, at that given time. In this case, it is possible to infer some properties (parameters) of the source, such as how far away it is, its location in the sky, or the angle of inclination (for a binary system). On the other hand, waves whose signals occupy the same frequency bin cannot be resolved, and in this case, it is only possible obtain information about these sources as whole, and not as individual systems. As the number of unresolvable sources increases, the superposition of their waves, and hence their total signal, become *noise-like*, and can only be characterised through their statstical properties. The sources that can be resolved

form the *resolvable* part of the stochastic background, and those that cannot be resolved form the *unresolvable* part of the stochastic background, which is the part that is considered in this work. For a more formal explanation of resolvability, see [11].

4.1 Cosmological background

The Cosmic Microwave Background (discovered in 1965 by Penzias and Wilson [12]) is composed of photons which decoupled from the primordial plasmas. It is a stochastic background of *electromagnetic* waves because of the random processes from which it was produced in the early Universe. There are good reasons to believe the existence of a stochastic background of gravitational waves is similar to this, often called the *cosmological* gravitational wave background. Such a background is expected to have statistical properties similar to those of the Cosmic Microwave Background. A background created early in the universe takes a time-scale of the order of the age of the Universe to change substantially, so it can be assumed to be *stationary*, especially over the duration of a typical expreiment, which is about a few years. According to the central limit theorem, the background can also assumed to be *Gaussian*, because the background can be viewed as made up of a large number of independent random processes. Like the Cosmic Microwave Background, a cosmological graviational wave background should also be *isotropic*, reflecting the highly isotropic early Universe. Of course, like the Cosmic Microwave Background, there may also be small anisotropies which might give interesting information about the early Universe. Finally, it is natural to take the background to be *unpolarised*, as there is no reason to believe otherwise. Whilst the Cosmic Microwave Background was produced about $100,000$ years after the Big Bang, the cosmological gravitational wave background would have been produced as early as 10^{-30} s after the Big Bang.

4.2 Astrophysical background

In addition, there exist *astrophysical* stochastic gravitational wave backgrounds. In [11], the name *contemporary background* is used. The sources of this background are astrophysical systems which were formed in the past, and can still form today. As was mentioned in chapter 1, there are a large number of compact binary systems (about $130,000$ [13]) in our galaxy, whose waves alone form an unresolvable background. This background is expected to be stationary, Gaussian and unpolarised. However, the background will be *anisotropic*, since the binaries are confined to our galaxy.

4.3 Characterisation

A stochastic background is conveniently described by a superposition of plane waves of different frequencies and directions in the sky. So, using (2.11), the spatial part of the metric pertubation of a stochastic background can be written as

$$h_{ab}(t,\vec{x}) = \int_{-\infty}^{\infty} df \int_{S^2} d\hat{k}\; e^A_{ab}(\hat{k}) h_A(f,\hat{k})\; e^{i2\pi f(t+\hat{k}\cdot\vec{x})} \tag{4.1}$$

It does not depend on ψ, because it is assumed that the background is unpolarised. (See next section for details on this.) Here, the Fourier amplitudes $h_A(f,\hat{k})$ are a random field. Since $h_{ab}(t,\vec{x})$ is real, $h_A(-f,\hat{k}) = h^*_A(f,\hat{k})$.

The statistical properties of $h_A(f,\hat{k})$ characterises the stochastic background. As discussed above, it will be assumed that the stochastic background is Gaussian, stationary, unpolarised and anisotropic. These assumptions can be summerized in the following expectation values:

$$\langle h_A(f,\hat{k})\rangle = 0 \tag{4.2}$$

$$\langle h^*_A(f,\hat{k}) h_{A'}(f',\hat{k}')\rangle = \frac{1}{4}\delta(f-f')\delta_{AA'}\delta(\hat{k}-\hat{k}')\mathcal{P}(f,\hat{k}) \tag{4.3}$$

Because of Gaussianity, these two expectation values of h completely determine the stochastic background. $\delta(f-f')$ corresponds to the stationarity assumption; $\delta_{AA'}$ corresponds to the "unpolarised" assumption; $\delta(\hat{k}-\hat{k}')$ corresponds to the assumption that the power coming from two different directions are independent of each other. $\mathcal{P}(f,\hat{k})$ describes the spectral and spatial distribution of the stochastic background.

The *spectral function*, $\Omega_{GW}(f)$, of a stochastic background is defined as

$$\Omega_{GW}(f) = \frac{1}{\rho_c}\frac{d\rho_{GW}}{d\ln f}\,, \tag{4.4}$$

where $d\rho_{GW}$ is the energy density of the background in a logarithmic interval of frequency, f/df; ρ_c is the current critical energy density for closing the Universe, $(3c^2H_0^2)/(8\pi G)$. It can be shown that

$$\Omega_{GW}(f) = \frac{2\pi^2}{3H_0^2} f^3 \int_{S^2} d\hat{k}\mathcal{P}(f,\hat{k}) \tag{4.5}$$

Notice that there a difference of f^3 between the frequency-depedence of the spectral function and $\mathcal{P}(f,\hat{k})$.

Chapter 5

Mapping An Anisotropic Background

One of the main objectives of this work is to estimate how the intensity of a stochastic gravitational wave background varies across the sky (or more precisely, to determine the spatial dependence of the function $\mathcal{P}(f, \hat{k})$). This process will often be referred to as map-making or mapping, and the spatial distribution to be determined will often be simply called the skymap.

One of the very first steps in the mapping of a stochastic background is to obtain data, or data products, that only contain contribution from gravitational wave signals. There are several ways to distinguish the noise-like random signals from stochastic gravitational wave backgrounds from intrumental and environmental noises in the data stream. If there is good knowledge of the instrumental noise, it might be possible to subtract it from the observed data, leaving only the background's signal [1]. Another way of separating a random signal from instrumental and environmental noise is to form a data product by cross-correlating the outputs of two noise-orthogonal observables. Since the noise in the two observables are uncorrelated, the cross-correlated output essentially contains signal contributions. This is in fact an important technique that is often used in the detection of stochastic backgrounds, and will be used in this work for map-making.

As seen from the description of LISA's gravitational wave response functions in chapter 3, LISA is more sensitive in some directions than in others. This directional sensitivity is often described by a function called the *antenna pattern*[2]. At a given time, the antenna pattern measures the entire skymap with varying levels of sensitivity, and as LISA moves and rotates relative to the sky, it measures a given part of the skymap with changing sensitivity.

[1] In fact, it has been suggested that, because the optimal TDI observable T is relatively insensitive to gravitational waves (especially at low frequencies), it can be used as a noise monitor, and be used for subtracting out noises from the observables' outputs.

[2] The exact definition of the antenna pattern depends on the data product considered, be it a TDI observable's GW signal or a cross-correlated signal.

Since the signal (or data product) considered contains in general contributions from all directions, the following observations can bnnne made. If the skymap is isotropic, then the average magnitude of the signal is constant in time, since each part of the antenna pattern points at a spot of the same intensity, at all times. If the skymap is anisotropic, the average magnitude of the signal is expected to change with time. When there are more insenstive directions of the antenna pattern pointing at the non-zero intensity parts of the skymap, the magnitude is low; when more sensitive directions point at the non-zero intensity parts of the skymap, the signal magnitude is high. If the movement of the antenna pattern relative to the skymap is regular, as is almost the case with LISA, the signal will appear to be modulated in time. In fact this is demonstrated for the background from galactic white dwarf-white dwarf binaries in [14], where the signal is described as *cyclostationary*. Therefore, time-modulations of the average magnitude of the signal indicate an anisotropy of the background. Below, it will be explained how this information can be used to reconstruct the skymap.

In this chapter, methods for estimating the anisotropy of a stochastic gravitational wave background, their limitations, and the assumptions used are discussed. These can be considered as *cross-correlation* methods, which take the cross-correlated outputs of TDI observables as the starting data product, and, using knowledge of the relevant antenna pattern, estimate the anisotropy of the background. There are two different approaches that exist in the literature: one concentrates on map-making at particular frequencies, and one takes into account signals measured at all frequencies. An overview of these methods are given here.

5.1 Basic assumptions

The following assumptions are made for all the methods considered here. First, it will be assumed that the function $\mathcal{P}(f,\hat{k})$ is separable:

$$\mathcal{P}(f,\hat{k}) = H(f)\mathcal{P}(\hat{k})\ . \tag{5.1}$$

This means that the anisotropy of the gravitational wave background is the same at all frequencies. Of course, there is no reason for this to be true in gneral, but it is a good approximation for a small frequency range. It will be assumed that the frequency band of LISA is small enough for this assumption to be made. The anisotropy of the background is therefore purely described by the function $\mathcal{P}(\hat{k})$.

In most of the methods, it is assumed that the function $\mathcal{P}(\hat{k})$ can be well-approximated in a spherical harmonic expansion up to some degree $l = l_{max}$:

$$\mathcal{P}(\hat{k}) = \sum_{l=0}^{l_{max}} \sum_{m=-l}^{m=+l} \mathcal{P}_{lm} Y_{lm}(\hat{k}) \tag{5.2}$$

For a diffuse source whose intensity varies slowly across the sky, l_{max} could be small; for sharply varying spatial distributions such as a point source, l_{max} is expected to be large. Here, l_{max} is simply chosen somewhat arbitrarily.[3]

For a physical background, it is expected that

$$H(-f) = H(f) \tag{5.3}$$

and $\mathcal{P}(\hat{k}) \in \Re$, implying

$$\mathcal{P}_{l-m} = (-1)^m \mathcal{P}^*_{lm} \ . \tag{5.4}$$

Next, the gravitational wave spectrum $H(f)$ will be taken as a power-law function,

$$H(f) = \left(\frac{f}{f_R}\right)^{\beta} \ , \tag{5.5}$$

where f_R is some reference frequency. β is called the *gravitational wave spectral index*. In general, in some methods, this parameter needs to be chosen before the analysis of data, in some it is one of the unknown parameters to be determined by the data analysis. As will be seen later for one of the methods, it is only when the chosen spectral index matches the true spectral index of the power-law, that the estimators for the skyamp become unbiased.

While the above general assumptions are made about the gravitational wave background, the following are made about LISA. There are three important time-scales to be considered here. The LISA configuration rotates about its own centre while moving in an orbit around the Sun tailing the Earth. This motion can be taken to be periodic with a period of $T_{detector} = 1\,year$. Importantly, the antenna pattern of LISA also sweeps across the sky with this period. In this work, this continuous motion is broken up into segments of an appropriate duration, τ, over which the LISA configuration can be assumed to be stationary and static, so that the signal can be assumed to be stationary during this period. The duration also needs to be short compared to $T_{detector}$ so that not too much information about the antenna pattern's motion is lost. In addition, in order to minimise the amount of correlation between the signals from adjacent segments of τ, τ needs to much greater than a time interval of about $T_{corr} = 128\,s$[4]. These requirements imply that

$$T_{corr} \ll \tau \ll T_{detector} \ . \tag{5.6}$$

This contraint is obeyed in this work by choosing $\tau = 1\,day$. In this way, the potentially cyclo-stationary signal (over the LISA mission lifetime) produced by a gravitational wave background

[3] It has not been possible to choose an 'optimal' l_{max}, which gives the lowest the sky-averaged standard deviation of the point estimates, based on the injection and analysis of the standard LISA noise.

[4] T_{corr} should be a minimum time interval, beyond which the gravitational wave correlation goes to zero. The second generation 16-link TDI observables used in this work have a temporal footprint of about $8 \times (5 \times 10^9 m)/c = 128\,s$. Therefore, for two adjacent segments of duration τ with boundary at time t, there will be correlation between the gravitational wave responses from the two segments in the interval $[t - T_{corr}/2, t + T_{corr}/2]$. However, as long as $T_{corr} \ll \tau$, the contribution of this correlation to the entire τ is small.

is divided into small stationary sections. It will be seen below that certain data analysis procedures common to the different methods are applied to these stationary sections that last a day.

5.2 Cross-correlation and the overlap-reduction

In Chapter 3, we showed that the TDI observables provide data streams free of the otherwise dominating laser frequency noise, so they are generally considered to be the 'detector outputs' of LISA. For a TDI observable I, its output at time t is

$$y_I(t) = y_I^{GW}(t) + y_I^{noise}(t) \ , \tag{5.7}$$

y_I^{noise} is the noise response of the TDI observable. It includes the secondary instrumental noises from the links of the observable. Here, y_I^{GW} is the GW response of the observable to a stochastic gravitational wave background. Using the response to a plane wave (3.28) and the plane wave expansion of a gravitational wave background (4.1),

$$y_I^{GW}(t) = \int_{-\infty}^{\infty} df \int_{S^2} d\hat{k}\ F_I^A(f,\hat{k},t)\ h_A(f,\hat{k})\ e^{i2\pi f(t+\hat{k}\cdot\vec{p}_I(t))} \ . \tag{5.8}$$

Here, the event of reference is chosen to be $(t, \vec{p}_I(t))$. Implicitly, the angle of polarisation ψ has been set to zero; there is no problem with this, since waves from all direction are considered and the background is assumed to be unpolarised.

For a segment of duration τ, centred at time t, its Fourier transform is

$$\tilde{y}(f,t) = \int_{t-\tau/2}^{t+\tau/2} dt'\ e^{-i2\pi f t'} y(t') \tag{5.9}$$

This is sometimes referred to as the *short-term Fourier transform* of the time-series y at time t.

In general, the outputs of two observables I and J are cross-correlated. In this work, the cross-correlated output used is defined as

$$C_{IJ}(f,t) \equiv \frac{2}{\tau}\tilde{y}_I^*(f,t)\ \tilde{y}_J(f,t) \tag{5.10}$$

The factor of 2 here indicates that $C_{IJ}(f,t)$ corresponds to a *one-sided* power spectral density. Since no correlation is expected between GW responses and instrumental noise, using (5.7), (5.9) and (5.10), the expectation value of the cross-correlated output is simply

$$\langle C_{IJ}(f,t)\rangle = \frac{2}{\tau}\langle\tilde{y}_I^{GW*}(f,t)\tilde{y}_J^{GW}(f,t)\rangle + \frac{2}{\tau}\langle\tilde{y}_I^{noise*}(f,t)\tilde{y}_J^{noise}(f,t)\rangle \tag{5.11}$$

In Appendix B, it is shown that

$$\frac{2}{\tau}\langle \tilde{y}_I^{GW*}(f,t)\tilde{y}_J^{GW}(f,t)\rangle = \int_{S^2} d\hat{k}\ \gamma_{IJ}(f,\hat{k};t)\mathcal{P}(f,\hat{k})\ , \tag{5.12}$$

where

$$\gamma_{IJ}(f,\hat{k},t) = \frac{1}{2}F_I^{A*}(f,\hat{k},t)F_J^A(f,\hat{k},t)e^{-i2\pi f\hat{k}\cdot(\vec{p}_I(t)-\vec{p}_J(t))}\ . \tag{5.13}$$

If the TDI observables I and J have uncorrelated noise, then (5.11) becomes

$$\langle C_{IJ}(f,t)\rangle = \int_{S^2} d\hat{k}\ \gamma_{IJ}(f,\hat{k};t)\mathcal{P}(f,\hat{k})\ . \tag{5.14}$$

Using this, the main problem to be tackled can be formulated here. Referring first to the picture described at the beginning of the chapter, (5.14) can be interpreted as follows: during a segment of duration τ centred at time t, LISA, assumed to be static and stationary, has an antenna pattern $\gamma_{IJ}(f,t)$, which measures the gravitational wave background $\mathcal{P}(f,\hat{k})$ over the whole sky and produces the data $C_{IJ}(f,t)$. The task then is to determine the unknown $\mathcal{P}(f,\hat{k})$, using the information about $\gamma^{IJ}(f,\hat{k},t)$ and $C_{IJ}(f,t)$.

The function $\gamma^{IJ}(f,\hat{k},t)$ is called the *overlap-reduction*. It describes the frequency and directional sensitivity of the cross-correlation of observables I and J to a stochastic gravitational wave background. In (5.13), it is defined in terms of the gravitational wave response functions of the two observables and the difference between their reference locations. When working in the spherical harmonic basis up to l_{max}, it is expanded here as

$$\gamma_{IJ}(f,\hat{k},t) = \sum_{l=0}^{l_{max}}\sum_{m=-l}^{l}\gamma_{lm}^{IJ}(f,t)Y_{lm}^*(\hat{k})\ . \tag{5.15}$$

In [8], a study of the overlap-reduction's geometric properties is carried out for an equal-arm LISA[5]. Even though a more complicated model of LISA is used in this work, the results in [8] nevertheless provide a picture of what to expect in general. They show that sensitivity for higher multipole moments increases with frequency. For Michelson, Saganc and their optimal TDI observables, the sensitivity is zero for multipole moments with odd l. They also shows that at low frequencies, the overlap-reduction is simplified and can be written out analytically for basic TDI observables. (As expected, it lacks sensitivity for higher multipole moments.) In general however, at arbitrary frequencies and in a realistic LISA model, the overlap-reduction is a complicated function, and it is only practical to compute it numerically. As an example, in Fig. 5.1, the real and imaginary parts of an overlap-reduction for a given direction in the sky are plotted as a function of frequency. In Fig. 5.2, an overlap-reduction at a selected frequency plotted as a function of ecliptic longitudue and latitude.

[5]The overlap-reduction is called the *antenna-pattern* in reference [8], denoted by $\mathcal{F}^{IJ}(f,\hat{k},t)$

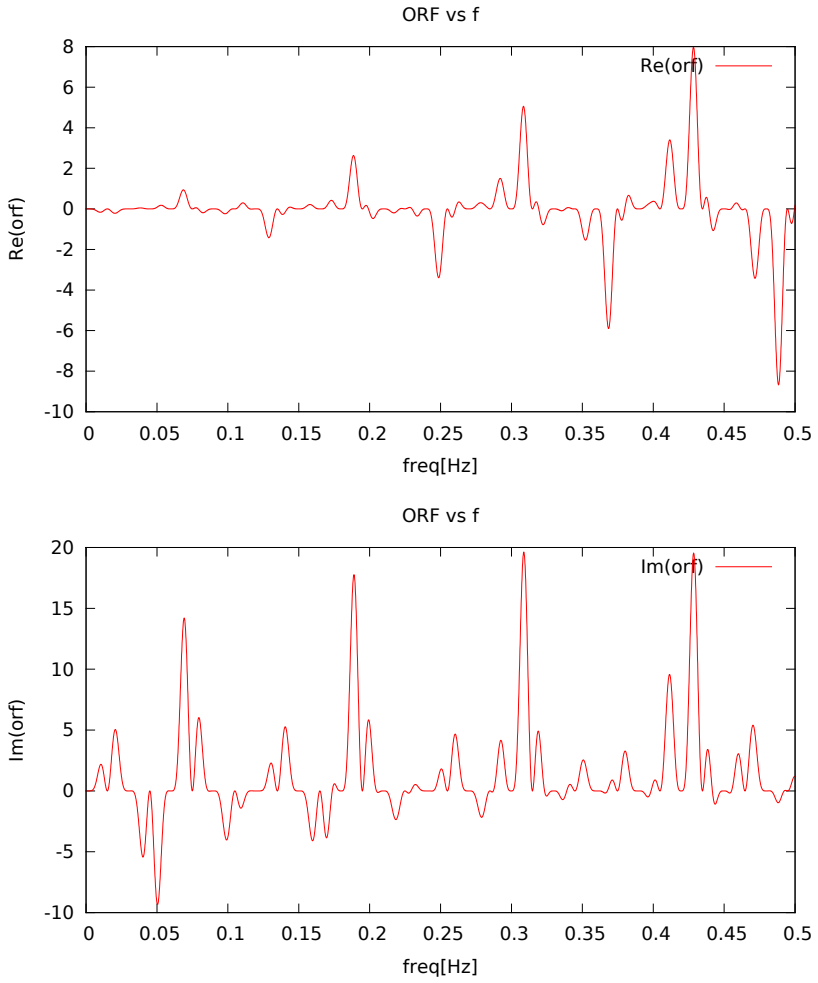

FIGURE 5.1: The overlap-reduction for Michelson-defined, second generation TDI observables A and E, in the direction $(\theta, \phi) = (,)$ and $t_0 =$. (left) Real part. (right) Imaginary part.

In the rest of the chapter, two methods which make use of the cross-correlated outputs of TDI observables are presented.

5.3 Cornish/Taruya and Kudoh

In this first method [8, 15, 16, 17], by making use of spherical harmonic expansions of the overlap-reduction and of the gravitational wave background, a system of linear equations is formed and solved to obtain estimates of the multipole moments of the background. By using (5.1), (5.2) and (5.15) in (5.14), the expectation value of cross-correlated output in the absence of cross-correlated noise becomes

$$S_{IJ}(f,t) \equiv \langle C_{IJ}(f,t) \rangle = H(f) \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} \gamma_{lm}^{IJ}(f,t) \mathcal{P}_{lm} \ . \tag{5.16}$$

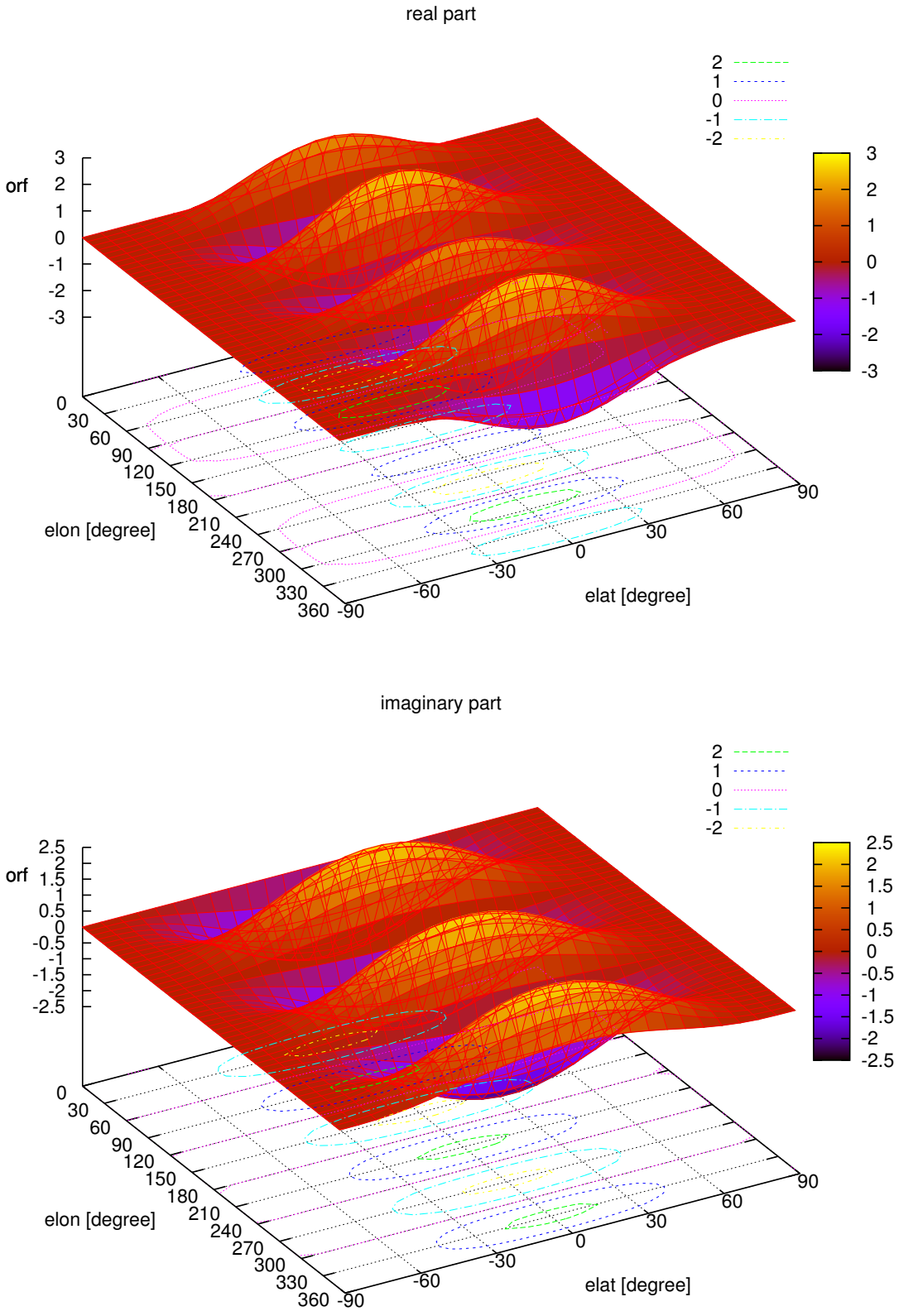

FIGURE 5.2: The overlap-reduction for Michelson-defined, second generation TDI observables A and E, at frequency $f = 0.02Hz$ and $t_0 = s$. (left) Real part. (right) Imaginary part.

By using data collected in segments at time $t = t_1, t_2, \ldots, t_N$ at some frequency f, (5.16) can be written as matrix equation

$$\mathbf{s} = \mathbf{A} \cdot \mathbf{p} \,, \tag{5.17}$$

where $\mathbf{s}$ is a $N \times 1$ matrix, containing the measured cross-correlated outputs $\{S_{IJ}(f, t_i)\}_{i=1,2,\ldots,N}$. $\mathbf{p}$ is a $(l_{max}+1)^2 \times 1$ matrix that contains all the $\mathcal{P}_{lm}$ up to l_{max}. $\mathbf{A}$ is a $N \times (l_{max}+1)^2$ matrix containing the $\gamma^{IJ}_{lm}(f, t_i)$s. In general, it is found that this system of equations is underdetermined. However, using the singular value decomposition of $\mathbf{A}$, a pseudo-inverse can be found and used to give an estimates of $\mathbf{p}$:

$$\hat{\mathbf{p}} = \mathbf{A}^{+} \cdot \mathbf{s} \,, \tag{5.18}$$

where $\mathbf{A}^{+}$ is a pseudo-inverse of $\mathbf{A}$, and $\hat{\mathbf{p}}$ is a $(l_{max}+1)^2 \times 1$ matrix whose elements are estimates of the multipole moments of $\mathcal{P}(\hat{k})$ up to $l = l_{max}$. In the presence of cross-correlated noise, there are additional unwanted terms to these estimates from $\mathbf{A}^{+} \cdot \mathbf{n}$, where $\mathbf{n}$ is a matrix containing the cross-correlated noise terms in (5.11).

For a simplified model of LISA, as is shown in [15] and [16], the implementation of this method is relatively straightforward. In this case, LISA is equal-arm and rigid, and, instead of trailing Earth in its orbit, it rotates and spins with its centre of the (sun-centred) ecliptic frame. If the axes of its rest frame are lined up with those of the ecliptic frame inititially, they can be found, at a later time t, by a set of Euler rotations: first a rotation of $\psi = -(2\pi/T_{detector})t$ about the z-axis, followed by a rotation of $\theta = -\pi/3$ about the y-axis, and then a rotation of $\phi = (2\pi/T_{detector})t$ about the z-axis. The multipole moments of the overlap-reduction in the ecliptic frame at time t can be found from those in LISA's rest frame using:

$$\gamma^{IJ}_{lm}(f, t) = \sum_{n=-l}^{l} e^{-in\psi} d^{l}_{nm}(\theta) e^{-im\phi} \bar{\gamma}^{IJ}_{ln}(f) \,, \tag{5.19}$$

where $d^{l}_{nm}(\theta)$ are the Wigner D matrices for $n \geq m$. $\bar{\gamma}^{IJ}_{lm}(f)$ are the overlap-reduction's multipole moments in its rest-frame, and, knowing them, the multipole moments in the ecliptic frame at time t can be found by using the transformation (5.19). Because LISA is rigid (as is reflected in the time-independence of $\bar{\gamma}^{IJ}_{lm}(f)$), the cross-correlated output is periodic in t with a period of $T_{detector}$, and can be written in terms of a Fourier series with coefficients:

$$\tilde{S}_{IJ,k}(f) = \frac{1}{T_{detector}} \int_{0}^{T_{detector}} dt \; e^{-ik2\pi t/T_{detector}} S_{IJ}(f, t) \,. \tag{5.20}$$

Substituting (5.19) in (5.16), and then (5.16) in (5.20), the Fourier series' coefficients of the cross-correlated output can be written as

$$\tilde{S}_{IJ,k}(f) = H(f) \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l-k} (-1)^m \, d^{l}_{(m+k),m}(-\pi/3) \, \bar{\gamma}^{IJ}_{l,-(m+k)}(f) \, \mathcal{P}_{lm} \,. \tag{5.21}$$

This can be cast in the form of (5.17), from which estimates for $\mathcal{P}_{lm}$ can be obtained in a similar fashion.

Another variation to (5.21) comes about when the gravitational wavelengths λ considered are longer than the distance between LISA's spacecrafts. The *transfer frequency* of LISA is defined to be

$$f^* = \frac{1}{2\pi}\frac{c}{L}\,, \tag{5.22}$$

where L is the nominal LISA armlength, which is about $5 \times 10^6 km$. It is possible to expand (5.21) in powers of $\hat{f} \equiv f/f^*$[6]:

$$\tilde{S}^{(i)}_{IJ,k} = \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l-k} (-1)^m \, d^l_{(m+k),m}(-\pi/3)\, \bar{\gamma}^{IJ,(i)}_{l,-(m+k)}\, \mathcal{P}_{lm}\,, \tag{5.23}$$

where $\tilde{S}^{(i)}_{IJ,k}$ and $\bar{\gamma}^{IJ,(i)}_{l,-(m+k)}$ are the coefficients of the ith power of $\hat{f}$ in the expansion. In this *low frequency* limit, $\lambda \gtrsim L$, $\hat{f} \lesssim 1$, the analysis using (5.23) then amounts to choosing the lowest i with a non-zero coefficient, and forming systems of equations like (5.17). As is more precisely shown in [16], the estimation of $\mathcal{P}_{lm}$ for odd l and even l require different sets of observable pairs (IJs); and the lowest i for which $\bar{\gamma}^{IJ,(i)}_{lm}$ is non-zero is different for different observable pairs.

Whilst (5.21) and (5.23) are versions of this method in which the multipole moments of the overlap-reduction can be obtained analytically because of simplifying assumptions, in general (5.16) has to be used for mapping at arbitrarily high frequencies. For this, the overlap-reduction and its multipole moments have to be obtained numerically.

As stated in the conclusion in [17], this method conentrates on map-making with data at some chosen frequency, and the estimated maps from this method should be considered rough estimates like the 'dirty' map in radio astronomy. Also, this method does not provide estimates of uncertainties for the estimated $\mathcal{P}_{lm}$ through statistical inference. In implementations demonstrated in [16, 17], the quality of the estimates are quantified in terms of the difference between the estimated map and the 'true' map. The effect of the presence of instrumental noise is studied with a definition of a signal-to-noise ratio, in terms of the expectation value of cross-correlated signals and the noise spectral densities. By choosing a value for the signal-to-noise ratio as the threshold above which the cross-correlated signal can be used, P_{lm} estimates are obtained according to the descriptions above, and their qualities determined. Because the 'true' map is not known in practice, this way of quantifying the quality of the estimated map can only be applied in the testing of the method, where the 'true' map is the injected map used to simulate the signals.

[6] The factor $H(f)$, which is assumed to be known, is first implicitly absorbed in $\tilde{S}_{IJ,k}(f)$.

5.4 Spherical harmonic decomposition algorithm

In this method, using statistical inference, a likelihood function is written for $C_{IJ}(f,t)$. By maximising this function with respect to $\mathcal{P}_{lm}$, an estimator for $\mathcal{P}_{lm}$ is obtained. It will also be shown below that it is possible to assign an uncertainty to the estimated $\mathcal{P}_{lm}$ from their covariances.

Collected at frequencies $f = f_1, f_2, \ldots, f_N$ from segments of duration τ, centred at times $t = t_1, t_2, \ldots, t_M$, $C_{IJ}(f,t)$ can be arranged in a $NM \times 1$ matrix:

$$\mathbf{c}_{IJ} = \begin{pmatrix} C_{IJ}(f_1,t_1) \\ C_{IJ}(f_2,t_1) \\ \vdots \\ C_{IJ}(f_N,t_1) \\ C_{IJ}(f_1,t_2) \\ \vdots \\ C_{IJ}(f_{N-1},t_M) \\ C_{IJ}(f_N,t_M) \end{pmatrix} \tag{5.24}$$

These data are distributed according to a multivariate Gaussian, because the stochastic gravitational wave background and LISA noise are assumed to be Gaussian. If it is also assumed that the noises in observables I and J are uncorrelated, the expectation value of the cross-correlated output is the same as in (5.16):

$$\langle C_{IJ}(f,t)\rangle = H(f)\sum_{l=0}^{l_{max}}\sum_{m=-l}^{l}\gamma_{lm}^{IJ}(f,t)\mathcal{P}_{lm} = H(f)\gamma_{\alpha}^{IJ}(f,t)\mathcal{P}_{\alpha}\,, \tag{5.25}$$

where in the second equality, the sum over l and m is abbreviated with the repeated index α[7]. Using (5.24) and (5.25), the likelihood function, which is the probability of obtaining the multivariate Gaussian data $\mathbf{c}_{IJ}$ given $\mathcal{P}_{\alpha}$, can be written as:

$$\mathcal{L}(\mathbf{c}_{IJ}|\{\mathcal{P}_{\alpha}\},\mathcal{I}) = \frac{1}{(2\pi)^{(NM)/2}\sqrt{|\mathbb{C}|}}e^{-\frac{1}{2}(\mathbf{c}_{IJ}-\langle\mathbf{c}_{IJ}\rangle)\cdot\mathbb{C}^{-1}\cdot(\mathbf{c}_{IJ}-\langle\mathbf{c}_{IJ}\rangle)}\,. \tag{5.26}$$

$\mathbb{C}$ is the $MN \times MN$ covariance matrix of $C_{IJ}(f,t)$, with elements

$$\mathbb{C}_{ft,f't'} = \delta_{tt'}\delta_{ff'}P_I(f,t)P_J(f,t) + \delta_{tt'}\delta_{f-f'}(H(f)\gamma_{\alpha}^{IJ}(f,t)\mathcal{P}_{\alpha})^2\,, \tag{5.27}$$

as shown in Appendix C). P_I is the power spectral density of the output of observable I. Notice from (5.27) and (5.25) that $\mathcal{P}_{\alpha}$ appears in the likelihood function in the variables $\mathbb{C}$ and $\langle\mathbf{c}_{IJ}\rangle$. For gravirational wave signals that are weak relative to the noise, the second term

[7] This will be used if the indices l and m do not need to be displayed explicitly, to make expressions tidier.

in (5.27) can be ignored so that only the total power spectral densities of the data contribute:

$$\mathbb{C}_{ft,f't'} \approx \delta_{tt'}\delta_{ff'}P_I(f,t)P_J(f,t)\ . \tag{5.28}$$

This makes the extremising of the likelihood function with respect to $\mathcal{P}_\alpha$ easier, as they only appear in the $\langle \mathbf{c}_{IJ} \rangle$ terms. It will be seen later in the section that this weak-signal assumption does not invalidate the $\mathcal{P}_{lm}$ estimates when the signal is not weak compared to the noise. The estimates remain unbiased, but the uncertainty in the estimate will be underestimated for this case.

With (5.25) and (5.28) in (5.26), the maximising of the likelihood function with respect to $\mathcal{P}_{lm}$ reduces to maximising the function

$$\chi^2(\mathcal{P}_\alpha) = \sum_t \sum_f \frac{(C^*_{IJ}(f,t) - H(f)\gamma^*_\alpha(f,t)\mathcal{P}^*_\alpha)(C_{IJ}(f,t) - H(f)\gamma_\beta(f,t)\mathcal{P}_\beta)}{P_1(f,t)P_2(f,t)}\ . \tag{5.29}$$

Doing this, *maximum likelihood estimators* $\mathcal{P}_{lm}$ are obtained[8]:

$$\hat{\mathcal{P}}^{IJ}_\alpha = \Gamma^{IJ\,-1}_{\alpha\beta} X^{IJ}_\beta\ , \tag{5.30}$$

where

$$X^{IJ}_{lm} = \sum_{f,t} \gamma^{IJ*}_{lm}(f,t) \frac{H(f)}{P_I(f,t)P_J(f,t)} C_{IJ}(f,t) \tag{5.31}$$

$$\Gamma^{IJ}_{lm,l'm'} = \sum_{f,t} \gamma^{IJ*}_{lm}(f,t) \frac{H^2(f)}{P_I(f,t)P_J(f,t)} \gamma^{IJ}_{l'm'}(f,t)\ . \tag{5.32}$$

It can be shown that this is an *unbiased* estimator of $\mathcal{P}_{lm}$:

$$\langle \hat{\mathcal{P}}_\alpha \rangle = \mathcal{P}_\alpha\ . \tag{5.33}$$

Therefore, by estimating P_I, P_J and C_{IJ} from the data to obtain X^{IJ}_α and $\Gamma^{IJ}_{\alpha\beta}$, (5.30) gives an estimtate for the $\mathcal{P}_{lm}$ that describe the anisotropy of the gravitational wave background.

Despite weak-signal assumption used in the derivation, it is evident from (5.33) that (5.30) can be applied to the case where the signal is not weak, and still give an unbiased estimate of $\mathcal{P}_{lm}$; regardless of the strength of the signal, it is the power spectral densities of the outputs that are required in (5.31) and (5.32). The true $H(f)$ is generally not known before the analysis, so, in practice, several different $H(f)$ (or their spectral indices) are tried, with the one giving the best results taken in the end.

Notice that while $\Gamma^{IJ}_{\alpha\beta}$ can be computed, $\hat{\mathcal{P}}_\alpha$ and X^{IJ}_α are random variables due to their dependence on the random data $C_{IJ}(f,t)$. X^{IJ}_α and $\Gamma^{IJ}_{\alpha\beta}$ are often referred to as the *dirty map*

[8] Again, here the greek indices denote all (l,m)s up to l_{max}. Repeated indices are summed over.

and the *beam matrix* respectively because the of relationship:

$$\langle X_\alpha^{IJ} \rangle = \Gamma_{\alpha\beta} \mathcal{P}_\beta \ . \tag{5.34}$$

In Appendix D, the covariances of X_α^{IJ} and $\hat{\mathcal{P}}_\alpha$ are calculated. It is shown that the variance of $\hat{\mathcal{P}}_\alpha$ is given by[9]

$$\sigma^2_{\hat{\mathcal{P}}_\alpha} = \Gamma^{IJ-1}_{\alpha\alpha} + \Gamma^{IJ-1}_{\alpha\mu} \left(\Psi^{IJ}_{\mu\nu}\right) \Gamma^{IJ-1*}_{\nu\alpha} \ , \tag{5.35}$$

where

$$\Psi^{IJ}_{\alpha\beta} = \sum_t \sum_f \gamma^{IJ*}_\alpha(f,t) \frac{H^2(f) P^{GW^2}_{IJ}(f,t)}{P^2_I(f,t) P^2_J(f,t)} \gamma^{IJ}_\beta(-f,t) \ . \tag{5.36}$$

Since in practice, the cross-spectral density of the signals is not known, it is not possible to estimate the contribution to the variance due to $\Psi^{IJ}_{\alpha\beta}$. Therefore, the estimated variance of $\hat{\mathcal{P}}_\alpha$ is here simply taken as

$$\sigma^2_{\hat{\mathcal{P}}_\alpha} \approx \Gamma^{IJ-1}_{\alpha\alpha} \ . \tag{5.37}$$

This is equivalent to assuming that the signal is weak. In Appendix D, it is discussed that this approximation is ok in general, as the contribution due to $\Psi^{IJ}_{\alpha\beta}$ is relatively small.

Because $\sigma^2_{X^{IJ}_\alpha} \approx \Gamma^{IJ}_{\alpha\alpha}$ (see Appendix D) and $\sigma^2_{\hat{\mathcal{P}}_\alpha} \approx \Gamma^{IJ-1}_{\alpha\alpha}$, $\Gamma^{IJ}_{\alpha\beta}$ is a *Fisher* matrix[10].

The Fisher matix, from its definition in (5.32), is Hermitian, so in a singular value decomposition, it can be written as

$$\Gamma^{IJ}_{\alpha\beta} = USU^\dagger \ , \tag{5.38}$$

where U is a unitary matrix, and S is a diagonal matrix, whose non-zero elements are the positive and real eigenvalues of $\Gamma^{IJ}_{\alpha\beta}$, arranged in descending order. Larger eigenvalues correspond to those components of the skymap to which the overlap-reduction is more sensitive, while smaller eigenvalues correspond to those to which the overlap-reduction is less sensitive. Taking the inverse of (5.38),

$$\Gamma^{IJ-1}_{\alpha\beta} = US^{-1}U^\dagger \ , \tag{5.39}$$

where S^{-1} is S with all the diagonal elements replaced by their reciprocals. By considering this in (5.37), it can be seen that those components with smaller eigenvalues have larger variances in their estimates. In practice, some of the eigenvalues are even close to, if not equal to, zero[11]. This makes the Fisher matrix *ill-conditioned*[12] and not most useful in the estimation of the skymap using (5.30), since the components with extremely large uncertainities are included. To improve the situation, often a value between the largest and the smallest eigenvalue is

[9] Here the repetition of the index α is not a summation.

[10] It will often be referred to as the Fisher matrix in what follows

[11] Remember from 5.2 that, for example, certain overlap-reductions are only sensitive to multipole moments of even degree l

[12] An ill-conditioned matrix has a very large condition number ($> 10^6$), which is the ratio of its largest eigenvalue to the smallest.

arbitrarily selected, and any eigenvalue below this cut-off is considered *too* small, corrsponding to a problematic component. Then, by replacing the eigenvalues of these problematic components in S with infinity, their contribution to the estimate in (5.30) can be excluded. Alternatively, they can all be replaced with the smallest eigenvalue above the cut-off. This process of changing the eigenvalues of S is called *regularisation*, and the changed matrix is often denoted by S'. By replacing S with S' in (5.38) and (5.39), the *regularised* Fisher matrix and Fisher matrix inverse are obtained, denoted by $\Gamma'^{IJ}_{\alpha\beta}$ and $\Gamma'^{IJ-1}_{\alpha\beta}$, respectively. Using them, (5.30) becomes

$$\hat{\mathcal{P}}^{IJ}_{\alpha} = \Gamma'^{IJ-1}_{\alpha\beta} X^{IJ}_{\beta} \ , \tag{5.40}$$

while (5.37) becomes (see (D.4))

$$\sigma^2_{\hat{\mathcal{P}}_\alpha} = \Gamma'^{IJ-1}_{\alpha\mu} \left(\Gamma^{IJ-1}_{\mu\nu} + \Psi^{IJ}_{\mu\nu} \right) \Gamma'^{IJ-1*}_{\nu\alpha} \ . \tag{5.41}$$

While it is not good to select a cut-off that includes too many eigenvalues, especially those which are too small, it is not useful either select one that excludes too many. For example, by selecting the largest eigenvalue as the cut-off and setting all remaining eigenvalues equal to it, S', and hence $\Gamma'^{IJ-1}_{\alpha\beta}$, becomes proportional to the identity matrix, implying that the estimate $\hat{\mathcal{P}}_\alpha$ is simply proportional to the dirty map X^{IJ}_{α}.

Chapter 6

Implementation

In the work, the spherical harmonic algorithm described in chapter 5 is used to estimate the anisotropy of a stochastic background. In order to test the method, gravitational wave signal from some given stochastic background is first simulated, added to simulated instrumental noise, and then passed onto the algorithm to be analysed. This whole process is sketched out in Fig. 6.1. It displays the main stages of data-simulation in the green boxes, of the main analysis in the blue boxes, and of the post-analysis in the pink boxes. The data-simulation and the main analysis both require knowledge of the overlap-reduction, which is here computed separately. In the sections below, details about how these are implemented and tested are given. The resultant implementation is in the form of a collection of Python (`www.python.org`) modules and scripts, available at `https://github.com/qAp/LisaMapp`.

6.1 Computation of overlap-reduction

Using (5.13), (3.29), (3.6), (3.7) and (3.8), the overlap-reduction for a pair of TDI observables, I and J, can be expressed explicitly in terms of LISA's armlengths, unit vectors along the arms, and the spacecrafts' positions, etc. Knowing these parameters, it is then possible to compute overlap-reduction's value for a given frequency f, direction $\hat{k}$ and time t. However, for TDI observables with several links, the overlap-reduction quickly becomes a very complicated function, making the writing out of its analytical expression impractical. Instead, the expression of the response function of a single link in (3.6) is used directly to compute values. With repeated application of this to all the links in a TDI observable, the value of the TDI observable's response functione is found according to (3.29). Then, using (5.13), the value of the overlap-reduction between two TDI observables can be found. In this way, by simply providing information such as its constituent links and the reference spacecraft, etc., the value of the response function (or of the overlap-reduction) can be computed for any TDI observable (or observables) using the same set of routines.

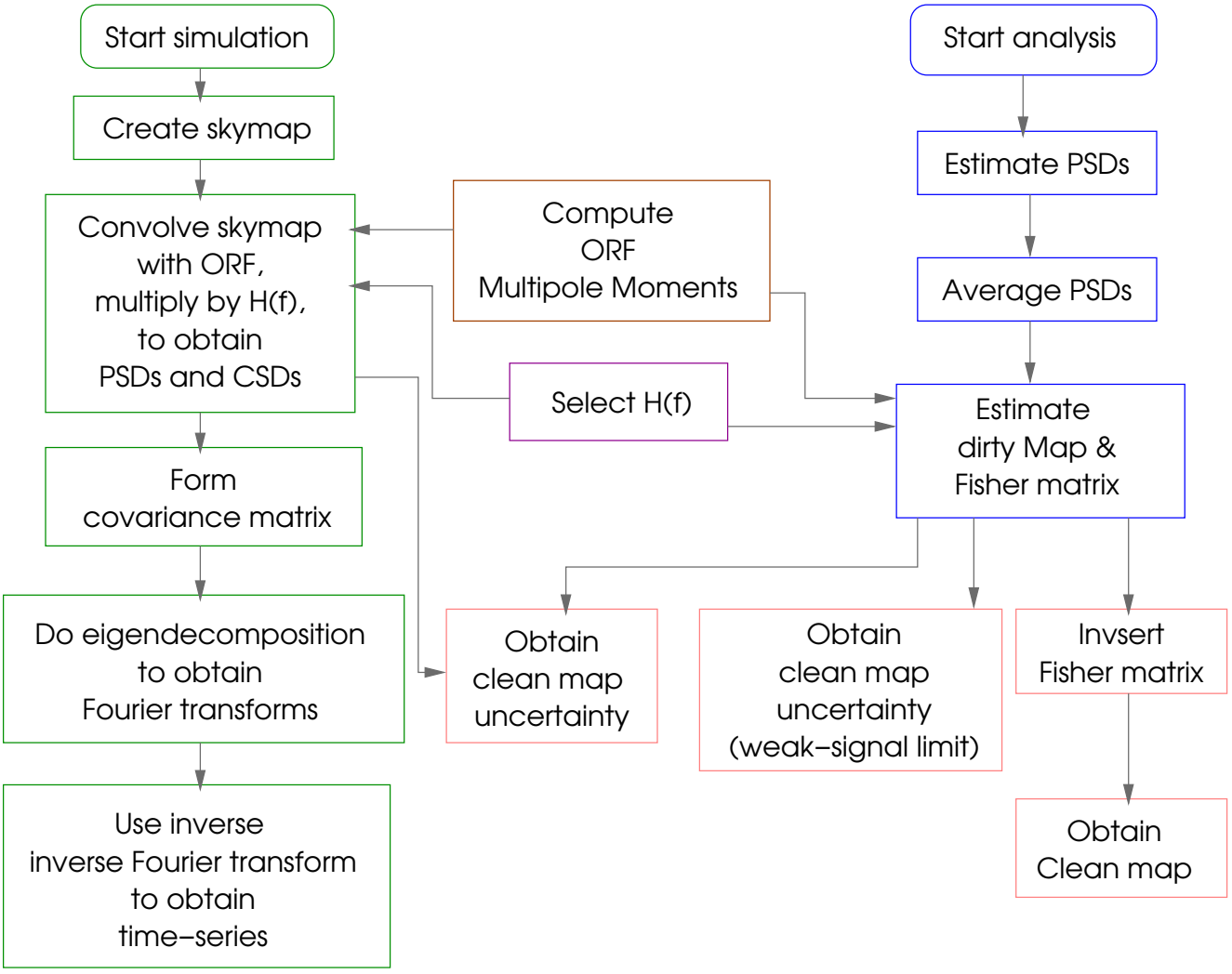

FIGURE 6.1: Flow chart showing main parts of the implemented pipelines. (*green*) Signal-simulation pipeline. (*blue*) Analysis pipeline. (*pink*) Post-analysis pipeline. (*brown*) Computation of the multipole moments of overlap-reduction $\gamma_{lm}^{IJ}(f,t)$. (*purple*) Selection of gravitational wave spectrum $H(f)$.

Parameters mentioned above (such as LISA's armlengths, unit vectors along the arms, and the spacecrafts' positions), which are needed to compute the values of the response functions, are in this work obtained with Synthetic LISA [9]. This package provides models of LISA with different levels of complexity, ranging from a static and non-rotating LISA to a non-static and rotating LISA. If the a static LISA is selected, then its geometric parameters are constant in time, whereas in a non-static model, these change with time.

By computing the overlap-reduction at all points, $\hat{k}$, on a suitably chosen grid over the sky, its multipole moments can be computed with

$$\gamma_{lm}^{IJ}(f,t) = \int_{s^2} d\hat{k} Y_{lm}(\hat{k}) \gamma_{IJ}(f,\hat{k},t) \; . \tag{6.1}$$

In this work, this is done using the Python package PYSPHARM (`http://code.google.com/p/pyspharm`). Appendix G gives more details about this. As it is time-consuming to compute

the multipole moments of the overlap-reductions, these are normally saved to disk for repeated use during the data-simulation and the data-analysis process.

In order to have confidence in the values of the overlap-reductions computed by the codes, they are compared with those evaluated from explicit analytical expressions in special cases. Even though it is in general difficult to write out whole analytical expressions of overlap-reductions as mentioned above, it is possible when the model of LISA assumed and the TDI observables considered are simple. In [8], the analytical expression is written out in full for the response function of the TDI observable $X0$ (3.16), for a static, equal-arm and non-rotationg LISA. From this, it is possible to use (3.24) to obtain expressions for higher TDI generations. Expressions for Y and Z can be quickly found from those of X through cyclic permutations. Once the response functions are obtained, the analytical expression for the overlap-reduction can be found with (5.13). The obtaining these analytical expressions have been carried out in the Mathematica (`http://www.wolfram.com/mathematica/`) notebook and modules attached in Appendix. By using these to provide the expected values, for a selected range of frequencies and direction in the sky, the computed values of the response function for TDI observables $X0$, $Y0$, $Z0$, $A0$, $E0$, $T0$, Em, Tm, $X2$, $A2$, $E2$ and $T2$, and the computed values of the overlap-reduction for TDI observable pairs $A0A0$, $A0E0$, $A2A2$, $A2E2$, $A2T2$, $E2A2$, $E2E2$, $E2T2$, $T2A2$, $T2E2$ and $T2T2$, have been verified. At a selected frequency and polarisation angle, the sky-projections of the response functions for y_{123}, $X0$, $Y0$, $Z0$, $A0$, $E0$, $T0$, and of the overlap-reductions for $A2A2$, $A2E2$, $A2T2$, $E2A2$, $E2E2$, $E2T2$, $T2A2$, $T2E2$ and $T2T2$, are compared and found to agree[1].

6.2 Data simulation

In order to test the performance of the data-analysis pipeline, it is applied to simulated data containing signals from known stochastic backgrounds $\mathcal{P}(\hat{k})$ and instrumental noise. In this work, a frequency-domain method is implemented, where a stationary time-series, or a set of stationary time-series, are constructed from its assigned power spectral density or covariance matrix. Consider the simulation of a set of N_{ts} time-series, $\{y_1(t), y_2(t), \ldots, y_{N_{ts}}(t)\}$, with a given covariance matrix:

$$\mathbf{C}(f) = \begin{pmatrix} P_{11}(f) & P_{12}(f) & \ldots & P_{1N_{ts}}(f) \\ P_{21}(f) & P_{22}(f) & \ldots & P_{1N_{ts}}(f) \\ \vdots & \vdots & \ddots & \vdots \\ P_{N_{ts}1}(f) & P_{N_{ts}2}(f) & \ldots & P_{N_{ts}N_{ts}}(f) \end{pmatrix}, \tag{6.2}$$

[1] Unfortunately, it is not possible to test further by obtaining values from analytical expressions for the multipole moments $\gamma^{IJ}_{\alpha\beta}$ (in [8], this is done for the low-frequency limit, in which $\gamma^{IJ}(f, \hat{k}, t)$ is simplified to allow this.). These can only be obtained numerically, which is done with PYSPHARM.

where $P_{II}(f)$ is the *one-sided* power spectral density of $y_I(t)$, and $P_{IJ}(f)$ is the *one-sided* cross-spectral density of $y_I(t)$ and $y_J(t)$. In terms of the Fourier transforms of the time-series,

$$\mathbf{C}(f) = \frac{2}{\tau}\langle \mathbf{Y}^\dagger(f)\mathbf{Y}(f)\rangle \ , \tag{6.3}$$

where

$$\mathbf{Y}(f) = \begin{pmatrix} \tilde{y}_1(f) & \tilde{y}_2(f) & \dots & \tilde{y}_{N_{ts}}(f) \end{pmatrix} \ . \tag{6.4}$$

Since covariance matrices are *normal* matrices, $\mathbf{C}(f)$ can be diagonalised as

$$\mathbf{C}(f) = \mathbf{X}(f)\boldsymbol{\lambda}^{\frac{1}{2}}(f)\,\mathbf{I}\,\boldsymbol{\lambda}^{\frac{1}{2}}(f)\mathbf{X}^\dagger(f) \ . \tag{6.5}$$

$\boldsymbol{\lambda}^{\frac{1}{2}}(f)$ is the $N_{ts} \times N_{ts}$ diagonal matrix whose diagonal elements are the square-roots of the real and non-negative eigenvalues of $\mathbf{C}(f)$. $\mathbf{X}(f)$ is the $N_{ts} \times N_{ts}$ *unitary*[2] matrix whose columns are the orthonormal eigenvectors of $\mathbf{C}(f)$, corrsponding to the eigenvalues in $\boldsymbol{\lambda}(f)$. The identity matrix, $\mathbf{I}$, can be written as

$$\mathbf{I} = \langle \mathbf{z}^\dagger(f)\mathbf{z}(f)\rangle \ , \tag{6.6}$$

where

$$\mathbf{z}(f) = \begin{pmatrix} z_1(f) & z_2(f) & \dots & z_{N_{ts}}(f) \end{pmatrix} \tag{6.7}$$

$\{z_i(f)\}_{i=1,\dots,N_{ts}}$ are unocrrelated and complex frequency-series with zero mean and unit variance. By inserting (6.6) in (6.5), and then comparing it with (6.3), it can be seen that the Fourier transforms of the time-series can be found with

$$\mathbf{Y}(f) = \sqrt{\frac{\tau}{2}}\mathbf{z}(f)\,\boldsymbol{\lambda}^{\frac{1}{2}}(f)\,\mathbf{X}^\dagger(f) \ . \tag{6.8}$$

Then, by using the inverse Fourier transform, the time-series can be found[3]:

$$y_I(t) = \int_{-\infty}^{+\infty} df\ e^{i2\pi ft}\,\tilde{y}_I(f) \ . \tag{6.9}$$

In some situations, it is desirable to simulate a set of time-series lasting a fraction of the total duration, and then 'stitch' them together. Using some given covariance matrix (6.2), these shorter time-series are simulated from different draws of random numbers in (6.6) so that they are uncorrelated with each other. Each of them is windowed with half of a sine wave:

$$w[i] = \sin\left(\pi\frac{i}{n}\right) \ , \tag{6.10}$$

[2] $\mathbf{X}\mathbf{X}^\dagger = \mathbf{X}^\dagger\mathbf{X} = \mathbf{I}$

[3] In practice, the Fourier transform is converted to the discrete Fourier transform with $\tilde{y}_{I,k} = \tilde{y}_I(f)/\Delta t$, before the inverse fast Fourier transform is applied.

where n is the number of samples in a short time-series, and i is the sample index. Then, they are overlapped by 50% before being added together. The overlapping and the windowing ensure that the short time-series are joined together smoothly with the same amount of power in the resulting time-series.

6.2.1 signal simulation

The expected power spectral densities of signals from a stochastic background are found by convolving the overlap-reductions with the background's spatial distribution. For a period of τ centered at time t_0 over which LISA can be assumed stationary, by using (5.25) in (6.2), the covariance matrix for simulating the signals of the optimal TDI observables A, E and T can be shown to be:

$$\mathbf{C}(f,t_0) = H(f) \begin{pmatrix} \gamma_\alpha^{AA}(f,t_0)\mathcal{P}_\alpha & \gamma_\alpha^{AE}(f,t_0)\mathcal{P}_\alpha & \gamma_\alpha^{AT}(f,t_0)\mathcal{P}_\alpha \\ \gamma_\alpha^{EA}(f,t_0)\mathcal{P}_\alpha & \gamma_\alpha^{EE}(f,t_0)\mathcal{P}_\alpha & \gamma_\alpha^{ET}(f,t_0)\mathcal{P}_\alpha \\ \gamma_\alpha^{TA}(f,t_0)\mathcal{P}_\alpha & \gamma_\alpha^{TE}(f,t_0)\mathcal{P}_\alpha & \gamma_\alpha^{TT}(f,t_0)\mathcal{P}_\alpha \end{pmatrix} . \tag{6.11}$$

Since the repeated index α indicates summation over all multipole moments, if the spatial distribution of the background is initially described by a function $\mathcal{P}(\hat{k})$, it needs to be first expanded in the spherical harmonic basis. The spectral function $H(f)$ is described by (5.5), where the spectral index is assigned by the user. Using this covariance matrix, the signals $y_A^{GW}(t,t_0)$, $y_E^{GW}(t,t_0)$ and $y_T^{GW}(t,t_0)$ are then found as described in (6.8) and (6.9).

As mentioned in chapter 5, τ is here taken to be a day ($T_{day} = 86400s$). LISA's geometric parameters (hence the overlap-reductions) are therefore evaluated at

$$t_0 = \frac{T_{day}}{2} \, , \, \frac{3T_{day}}{2} \, , \, \frac{5T_{day}}{2} \, , \, \ldots \, , \left(k - \frac{1}{2}\right)\frac{T_{day}}{2} \, , \, \ldots \, , \left(N_{days} - \frac{1}{2}\right)\frac{T_{day}}{2} \, , \tag{6.12}$$

where $k = 1 \ , \ 2 \ , \ \ldots \ , \ N_{days}$. The frequency resolution corresponding to T_{day} is $\Delta f \approx 10^{-5} Hz$. It is found that computing γ_{lm}^{IJ} at this frequency resolution upto a desired upper limit of about $0.1Hz$ is too time-consuming. Therefore, time-series of shorter duration are simulated before they are stitched together to make up the one day's worth of data.

As a demonstration of the performance of the simulation method, in FIGURE 6.2, the spectral densities of an example of simulated signal are plotted with the expected spectral densities.

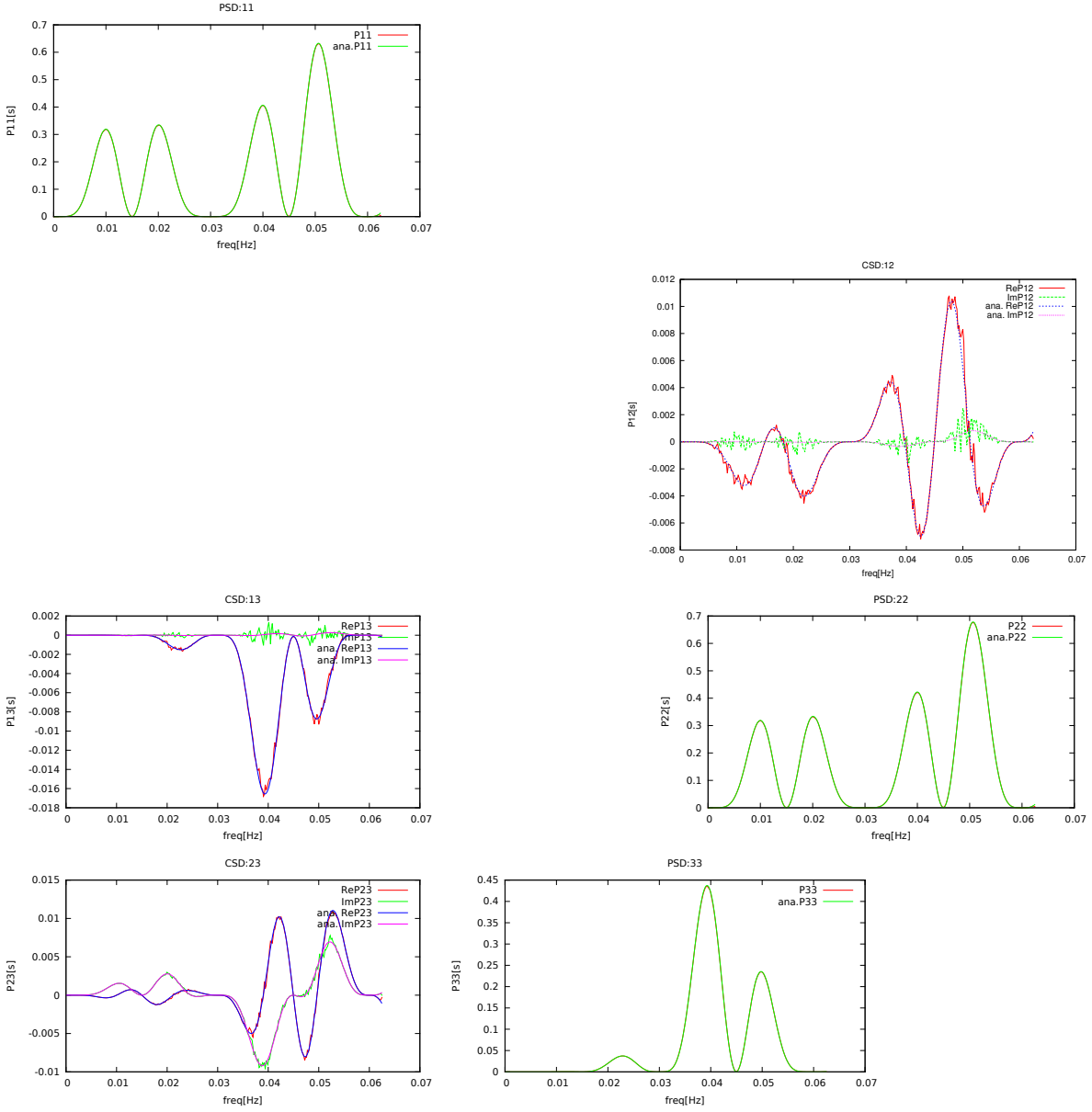

FIGURE 6.2: Spectral densities of simulated $A2$, $E2$, $T2$ signals on day 1 ($t_0 = T_{day}/2$) from a monopole source ($\mathcal{P}_{00} = 1/90$), a spectral function of $H(f) = 1$, averaged over 9075 realisations. (top left) $P_{AA}(f)$; (top right) $P_{AE}(f)$; (middle left) $P_{AT}(f)$; (middle right) $P_{EE}(f)$; (bottom left) $P_{ET}(f)$; (bottom right) $P_{TT}(f)$. Time-series sampled at $8s$, made up of 44 segments of $4000s$ in duration.

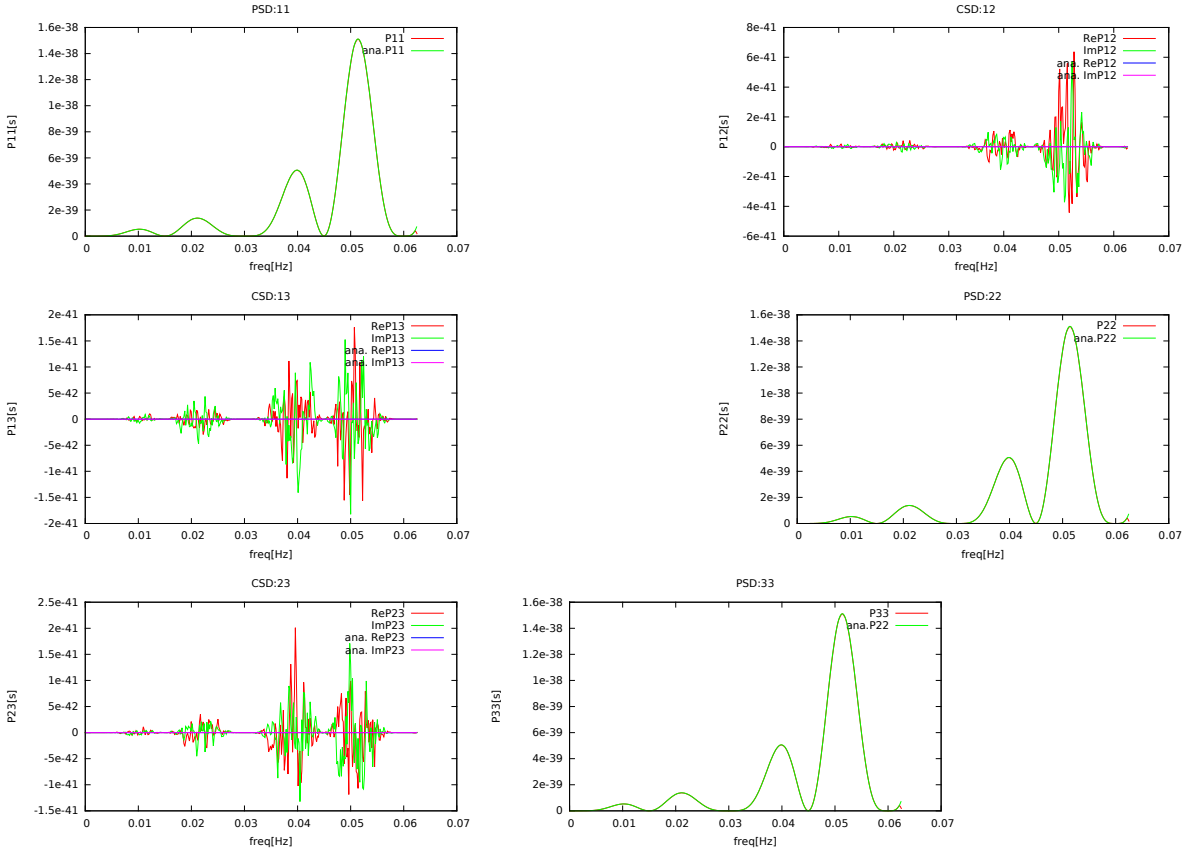

FIGURE 6.3: Spectral densities of simulated $A2$, $E2$, $T2$ noise, averaged over 9075 realisations. (top left) $P_{AA}(f)$; (top right) $P_{AE}(f)$; (middle left) $P_{AT}(f)$; (middle right) $P_{EE}(f)$; (bottom left) $P_{ET}(f)$; (bottom right) $P_{TT}(f)$. Time-series sampled at $8s$, made up of 44 segments of $4000s$ in duration.

6.2.2 noise simulation

Noise is simulated by using the appropriate noise spectral densities in (6.2). The noise spectral densities of optimal TDI observables can be found in section 3.4.6 for example. Since noise-simulation does not involve the time-consuming computation of γ_{lm}^{IJ} as is the case with signal-simulation, time-series of length τ can be simulated in one go. Since instrumental noise is assumed to be uncorrelated with the signals, it is important to make sure that a different seed from signal-simulation is used in the random generator.

6.3 Main analysis

In this section, details are given about the different stages (see Fig. 6.1) of the implemented data analysis pipeline based on the method described in section 5.4.

6.3.1 power spectral density estimation

The *one-sided* power spectral densities of the data time-series $y_A(t)$, $y_E(t)$ and $y_T(t)$ over an interval of τ are estimated using the *pwelch* method. In this method, the time-series is divided into segments of a shorter duration chosen here by the user. This PSD estimation is carried out for all the intervals of τ specified by the user, and the results from each interval of τ is written to disk in its own file. For example, in the file that results from the PSD estimation of data in the interval of τ centred at t_0, the PSDs $P_A(f, t_0)$, $P_E(f, t_0)$ and $P_T(f, t_0)$ are saved.

In this work, for $\tau = T_{day}$, the duration of the segments for the pwelch method is chosen to be $5760\,s$ (corresponding to a frequency resolution of $\Delta f \approx 1.7361 \times 10^{-4}\,Hz$), with the segments overlapping each other by $50\,\%$. The estimation is carried out for all 365 adjacent intervals of $\tau = T_{day}$ of the year.

6.3.2 average power spectral density

For an interval of τ, an 'average' PSD estimate is obtained by averaging those from its adjacent intervals of τ. For example, for an interval centred on the k^{th} day, this average PSD of $y_I(t)$ is

$$P_{I,avg}(f, k^{th}\text{day}) = \frac{1}{2}\left[P_I(f, (k-1)^{th}\text{day}) + P_J(f, (k+1)^{th}\text{day})\right] \ . \tag{6.13}$$

As it is defined above, the average PSDs have the same frequency resolution as the PSDs. Since the first and the last intervals of τ do not have two adjacent intervals, the average PSD is cannot obtained according to (6.13), and they will not be considered further in the analysis. The average PSD for each interval of τ is written to disk in its own file.

6.3.3 cross-correlated output

In practice, a window (such as a hanning window) is applied to a time-series before it is Fourier transformed in the process of obtaining the cross-correlated output. Therefore, if windows $w_I(t)$ and $w_J(t)$ are applied to $y_I(t)$ and $y_J(t)$ from the interval of τ centred at t_0, respectively, the cross-correlated output between them is estimated with

$$C_{IJ}(f, t_0) \equiv \frac{2}{\tau \overline{w_I w_J}} \tilde{y}_I^*(f, t_0)\, \tilde{y}_J(f, t_0) \ . \tag{6.14}$$

The factor $\overline{w_I w_J}$ is defined by

$$\overline{w_I w_J} = \frac{1}{N} \sum_{i=0}^{N} w_I[i] w_J[i] \ , \tag{6.15}$$

where N is the number of samples in the time-series $y_I(t)$, or $y_J(t)$. This factor compensates for the loss, or gain, of power due to the windowing in the Fourier transforms.

Since the Fourier tranforms that make up the cross-correlated output are those of time-series which have not been divided into shorter segments, the cross-correlated outputs have a finer frequency resolution than the PSD estimates in general. For $\tau = T_{day}$, thay have a frequency resolution of $\Delta f \approx 1.7540 \times 10^{-5}\, Hz$.

6.3.4 dirty map

By using (E.5) and (5.31), the dirty map for each interval of τ, centred at t_0 can be estimated with:

$$X^{IJ}_{lm}(t_0) = \frac{\overline{w_I w_J}^2}{\overline{w_I^2 w_J^2}} N\Delta t \Delta f \sum_{f=f_{low}}^{f_{high}} \frac{H(f)}{P_I(f,t_0)P_J(f,t_0)} \times \left[\gamma^{IJ*}_{lm}(f,t_0)C^{IJ}(f,t_0) + (-1)^m \gamma^{IJ}_{l-m}(f,t_0)C^{IJ*}(f,t_0)\right] . \tag{6.16}$$

N is the number of samples in the time-series; Δt is the sampling time; $\tau \approx N\Delta t$. The spectral function $H(f)$ again takes the form in (5.5), with the spectral index β and f_R chosen by the user prior to the start of the analysis. The overlap-reduction's multipole moments γ^{IJ}_{lm}, the (average) power spectral densities and the cross-correlated outputs are obtained as have been described above.

$$\frac{\overline{w_I w_J}^2}{\overline{w_I^2 w_J^2}} = \left(\frac{1}{N}\sum_{i=0}^{N} w_I[i] w_J[i]\right)^2 \tag{6.17}$$

is a factor due to the windowing of the time-series in the construction of the cross-correlated output. f_{low} and f_{high} are the lower and higher frequency limits between which the data is to be analysed. The estimated dirty map X^{IJ}_{lm} for each interval of τ is written to disk in its own file.

6.3.5 Fisher matrix estimation

By using (E.5) and (5.31), the Fisher matrix for each interval of τ, centred at t_0 can be estimated with:

$$\Gamma^{IJ}_{lm,l'm'}(t_0) = \frac{\overline{w_I w_J}^2}{\overline{w_I^2 w_J^2}} N\Delta t \Delta f \sum_{f=f_{min}}^{f_{max}} \frac{H^2(f)}{P_I(f,t_0)P_J(f,t_0)} \times \left[\gamma^{IJ*}_{lm}(f,t_0)\gamma^{IJ}_{l'm'}(f,t_0) + (-1)^{m+m'} \gamma^{*}_{l'-m'}(f,t_0)\gamma_{l-m}(f,t_0)\right] , \tag{6.18}$$

with all the variables defined the same way as in (6.16). The estimated Fisher matrix $\Gamma^{IJ}_{lm,l'm'}$ for each interval of τ is written to disk in its own file.

6.3.6 'strong-signal bias matrix'

By using (E.5) and (5.36), the 'strong-signal bias matrix' for each interval of τ, centred at t_0, can be estimated with:

$$\Psi^{IJ}_{lm,l'm'}(t_0) = \frac{\overline{w_I w_J}^2}{\overline{w_I^2 w_J^2}} N\Delta t\Delta f \sum_{f=f_{low}}^{f_{high}} \frac{H^2(f)}{P_I^2(f,t_0)P_J^2(f,t_0)} \Big[P_{IJ}^{GW2}(f,t_0)\gamma^{IJ*}_{lm}(f,t_0)(-1)^{m'}\gamma^{IJ*}_{l'-m'}(f,t_0) + P_{IJ}^{GW2*}(f,t_0)\gamma^{IJ*}_{l'm'}(f,t_0)(-1)^{m}\gamma^{IJ*}_{l-m}(f,t_0)\Big] \ . \tag{6.19}$$

As is shown in appendix D, $Psi^{IJ}_{lm,l'm'}$ needs to be taken into consideration in order to estimate the uncertainties of the $\mathcal{P}_{00}$ estimates, when the signal cannot be considered weak compared to the noise. The estimated $\Psi^{IJ}_{lm,l'm'}$ for each interval of τ is written to disk in its own file.

6.3.7 coarse-graining

As have already been seen above, the overlap-reduction $\gamma^{IJ}_{lm}(f,t)$, the power spectral densities $P_I(f,t)$ and the cross-correlated outputs $C_{IJ}(f,t)$ are obtained and evaluated at different frequencies. They have different start and end frequencies, and different frequency resolutions. In order to combine them in the sum over frequencies in (6.16), (6.18) and (6.19), they first need to be at the same set of frequencies. For this, the largest (or the coarsest) frequency resolution, Δf_{coarse}, is picked out. The highest of start frequencies is taken as the common start frequency, and the lowest of the end frequencies is taken to be the common end frequency. For a variable having an initial frequency resolution finer than Δf_{coarse}, its value in a given coarse frequency bin Δf_{coarse} is defined to be the average of the values in its original frequency bins that overlap with the coarse bin, weighted by the amount frequency overlapped. The Δf that appears in (6.16), (6.18) and (6.19) is Δf_{coarse}. Once all the variables, have been coarsegrained, they can then be combined to obtain the dirty map and the Fisher matrix.

6.3.8 Networks of cross-correlated observables

So far the cross-correlation between two TDI observables (I and J) have been discussed. However, provided with more than two, there are more than one possible cross-correlations, and the results from them can be combined to provide estimates of the stochastic background's anisotropy as well. This can be done by simply summing up the dirty maps and Fisher matrices from all the cross-correlated pairs of TDI observables in the network [18]:

$$X^{\mathcal{N}}_{\alpha} = \sum_{I}\sum_{J>I} X^{IJ}_{\alpha} \tag{6.20}$$

$$\Gamma^{\mathcal{N}}_{\alpha\beta} = \sum_{I}\sum_{J>I} \Gamma^{IJ}_{\alpha\beta}, \tag{6.21}$$

where the superscript $\mathcal{N}$ denotes a network/collection of cross-correlated pairs of observables. For example, given the optimal observables A, E and T, the possible indepedent cross-correlated pairs are: AE, AT, ET, and the possible networks are: $\{AE, AT\}$, $\{AE, ET\}$, $\{AT, ET\}$ and $\{AE, AT, ET\}$. An advantage of networks is that they make use of the sensitivity of all their memebers. This means that there are fewer 'blind spots', hence making $\Gamma^{\mathcal{N}}_{\alpha\beta}$ better conditioned. This ultimately reduces the uncertainty in the clean map.

6.4 Post-analysis

The main products of the main analysis described in the above section are the dirty map (6.16), the Fisher matrix (6.18) and the strong-signal bias matrix (6.19). After being computed and saved to disk for all intervals of τ analysed, these are then used in the post-analysis to obtain an estimate of the anisotropy of the stochastic background and its uncertainty.

6.4.1 Inversion of the Fisher matrix

As is seen in Fig. 6.1, the first step towards this is the computation of the Fisher matrix inverse, which is described at the end of section 5.4. The user can choose whether to regularise the Fisher matrix. There are three available regularisation schemes:

1. The user specifies how many of the largest eigevalues to keep.
2. The user specifies a number, and an eigenvalue larger than it is kept.
3. The user specifies a number between 0 and 1. If the ratio of an eigenvalue to the largest eigenvalue is larger than this number, it is kept.

The user also chooses whether to replace those discarded eigenvalues with infinity or the smallest eigevalue kept. Once the regularised diagonal matrix of eigenvalues S' is found using one of the schemes above, the regularised Fisher matrix inverse is obtained with:

$$\Gamma'^{IJ-1}_{\alpha\beta} = US'^{-1}U^{\dagger} \,, \tag{6.22}$$

where U is the unitary matrix of orthonormal eigenvectors. It is possible to restrict the Fisher matrix inverse to an l_{max} smaller than that used in the main analysis, by first restricting the Fisher matrix to this smaller l_{max} and the inverting as above[4].

In this work, the regularisation scheme most often used is (3), with the cut-off set to 0.001.

[4]It is erroneous to simply truncate the Fisher matrix inverse from a larger l_{max} to a smaller l_{max}.

6.4.2 clean map

By left-multiplying the Fisher matrix inverse to the dirty map (5.40), the clean map, $\hat{\mathcal{P}}^{IJ}_{lm}$, is obtained. This estimate of the true map, $\mathcal{P}_{lm}$, is the main result of the whole analysis pipeline. A projection of it gives a picture of how the intensity of the graviational wave background varies across the sky. When the signal is simulated, the performance simulation and/or analysis pipeline can evaluated by comparing $\hat{\mathcal{P}}^{IJ}_{lm}$ with the injected map, $\mathcal{P}_{lm}$. Notice that, as indicated by the superscript IJ, this estimate is dependent on the pair of cross-correlated TDI observables used in the analysis. In our implementation, the clean map is represented by a `AnisotropySearch.SkyMap` object. Its l_{max} is at most equal to the l_{max} chosen for the computation of the dirty map, X^{IJ}_{α}, and the Fisher matrix, $\Gamma^{IJ}_{\alpha\beta}$, in the main analysis.

6.4.3 Uncertainty of $\hat{\mathcal{P}}_{lm}$

Using the Fisher matrix, $\Gamma^{IJ}_{\alpha\beta}$, $\Psi^{IJ}_{\alpha\beta}$ and the Fisher matrix inverse, $\Gamma^{IJ-1}_{\alpha\beta}$, the variance of the clean map can be computed using (5.41). The standard deviation is computed and taken to be the measure of uncertainity of the clean map:

$$\sigma_{\hat{\mathcal{P}}^{IJ}_{\alpha}} = \left(\Gamma'^{IJ-1}_{\alpha\mu}\left(\Gamma^{IJ}_{\mu\nu} + \Psi^{IJ}_{\mu\nu}\right)\Gamma'^{IJ-1*}_{\nu\alpha}\right)^{\frac{1}{2}} . \tag{6.23}$$

The implementation offers the option of the inclusion of $\Psi^{IJ}_{\alpha\beta}$ in the computation. When dealing with real data, because the gravitational wave auto/cross-spectral density is not known, $\Psi^{IJ}_{\alpha\beta}$ cannot be computed and it would not be possible to include it[5] in (6.23). However, when the data is simulated, it is possible to include it, since the gravitational wave auto/cross-spectral density can be computed from the injected map. In both cases, $\Psi^{IJ}_{\alpha\beta}$ can be safely neglected when the GW spectral densities are small compared to the noise spectral densities.

Just as $\mathcal{P}_{lm}$ can be transformed into $\mathcal{P}(\hat{k})$ in the pixel basis, by first transforming the Fisher matrix $\Gamma^{IJ}_{lm,l'm'}$ to the pixel basis with

$$\Gamma^{IJ}_{\hat{k},\hat{k}'} = Y_{\alpha}(\hat{k})\,\Gamma^{IJ}_{\alpha\beta}Y^{*}_{\beta}(\hat{k}) , \tag{6.24}$$

the standard deviation of the clean map at sky location $\hat{k}$ is

$$\sigma_{\hat{\mathcal{P}}(\hat{k})} = \sqrt{\Gamma^{IJ-1}_{\hat{k},\hat{k}}} . \tag{6.25}$$

[5] However, if the TDI observables I and J have uncorrelated noise, the cross-spectral density of their time-series should give an estimate of their GW signals' cross-spectral density.

Chapter 7

Demonstration

In this chapter, the developed data-simulation and data-analysis pipelines described in the previous chapter are used on a selection of made-up sources. The spatial distribution of the source, $\mathcal{P}(\hat{k})$, is first created either in the pixel or spherical harmonic basis. Then, using the signal-simulation pipeline, the GW signals are obtained for the optimal TDI observables of the second generation Michelson TDI observables. In some of the cases below, only GW signals are analysed. In cases where LISA noise is taken into account, the noise is simulated for the corresponding TDI observables, and combined with the signals. The resulting time-series are then passed into the analysis pipeline for anisotropy estimation. This produces various data products, among which is the clean map, which is the main result of the analysis. The results of the analysis are mainly presented below in graphs or projections of sky-maps.

The anisotropy of a stochastic background in general can consist of spatial features from very small scales to very large scales. Below, results for a point source (the smallest feature possible) and for an isotropic background (the largest feature possible) are presented.

7.1 Point source

As a demonstration of the mapping capability of the implementation, the sky-mapping of a point source[1] with and without instrumental noise is carried out. A point source is by definition a source that occupies an infinitesimally small area in the sky, and so should have multipole moments up to $l_{max} = \infty$. In practice, $\mathcal{P}_{lm}$ only up to a finite l_{max} is considered, and are obtained as follows in this demonstration. The sky is first divided into pixels of 1 square degree (360 ecliptic longitudes and 181 ecliptic latitudes). Then, one of these pixels is picked out to represent the point source. Out of all the pixels in the sky, only this one is assigned a non-zero number equal to the strength of the point source. These pixels describe the

[1] Why point source? From experience of doing a similar analysis in LIGO, point sources are the easiest pick out, and so they offer a good first test.

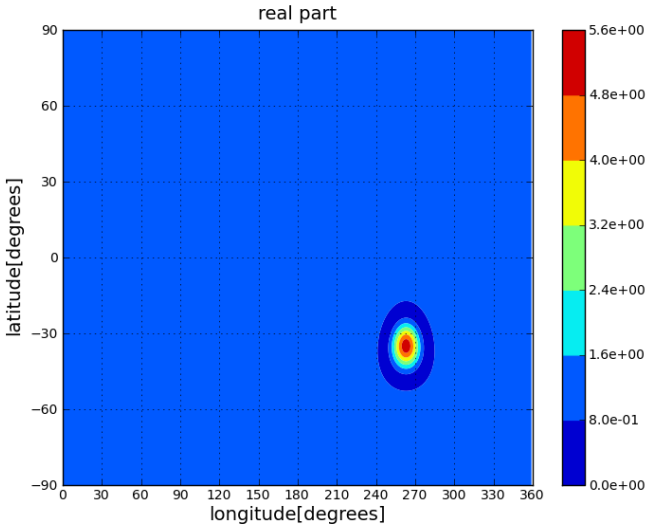

FIGURE 7.1: Projection of a point source, at ecliptic longitude and latitude of 263° and −35°, respectively, on top of a source uniform over the sky. $\mathcal{A}_{\hat{k}_0} = 100$; $\mathcal{A}_{uniform} = 1$. $l_{max} = 20$.

spatial distributation of the point source in the pixel basis, $\mathcal{P}(\hat{k})$. By applying the spherical harmonics transform to this (say using the `PYSPHARM` package), the multipole moments $\mathcal{P}_{lm}$ can be computed up to $l_{max} = N_{lat} - 1$, where N_{lat} is the number of latitudes. However, it is observed that in this case, the $\mathcal{P}(\hat{k})$ resulting from the inverse spherical harmonic transform of $\mathcal{P}_{lm}$ with a finite l_{max} can have negative values in some pixels. This is because the $\mathcal{P}_{lm}$ here, with its finite l_{max}, is only an approximation of the original pixel map $\mathcal{P}(\hat{k})$. Since $\mathcal{P}(\hat{k})$ is proportional to the *power* of the GW background, it should be semi-positive definite over the sky. To ensure this, it is necessary to add, to the point source's spatial distribution, a uniform source with an amplitude sufficiently large to make $\mathcal{P}_{lm}$ inverse-transform into a semi-positive definite $\mathcal{P}(\hat{k})$ over the sky.

An example of such a spatial distribution, $\mathcal{P}(\hat{k})$, consists of a point source located at $\hat{k}_0$,

$$\mathcal{P}_{\hat{k}_0}(\hat{k}) = \mathcal{A}_{\hat{k}_0}\delta(\hat{k}, \hat{k}_0) \ , \tag{7.1}$$

and a uniform source,

$$\mathcal{P}_{uniform}(\hat{k}) = \mathcal{A}_{uniform} \ , \tag{7.2}$$

where $\mathcal{A}_{\hat{k}_0} = 100$ and $\mathcal{A}_{uniform} = 1$. With the point source located at ecliptic longitude of 263° and latitude of −35° (or $\hat{k}_0 = (\phi, \theta) = (263°, 125°)$) and $l_{max} = 20$, the projection of $\mathcal{P}(\hat{k})$ is shown in Fig. 7.1. Notice that because of the finite $l_{max} = 20$, the map drops sharply around the location of the point source (the contour with the darkest blue), before coming back up to the level of the uniform background. If the uniform source used is too weak, the map would drop below zero. It is sufficiently strong here as shown by the non-negative markers on the colorbar in the figure.

By taking $H(f) = 1$, the GW signals of this source are simulated for the optimal TDI observables A, E and T of the second generation Michelson TDI observables $X2$, $Y2$ and $Z2$, as defined in (3.40), (3.41) and (3.42). They are sampled at 1 s. The overlap reduction functions are computed with a frequency resolution of $\Delta f = 10^{-4}$ Hz. Since a Fourier transform with this Δf corresponds to a time-series of duration of $T = 10^4$ s, by using stitching with 50% overlap, it takes a total of 18 time-series to make up a time-series of duration of a day. In total, 365 days of data are simulated. As demonstration of how well the covariance matrix used in simulation is recovered, the estimated auto and cross spectral densities of the simulated signals on three days of the year are plotted together with their analytical counterparts in Fig. 7.2. It can ben seen that they agree reasonably well.

7.1.1 In the absence of LISA noise

For the main analysis of this data set, $H(f)$ is chosen to be the same as that used in the simulation. Hanning windows are applied to the time-series prior to obtaining the cross-correlated output. The lower and higher frequency limit for the integral in the dirty map, the Fisher matrix and $\Psi^{IJ}_{\alpha\beta}$ are set to 1.5×10^{-4} Hz and 4.998×10^{-1} Hz, respectively. The analysis l_{max} is set to 20, the same as the injected l_{max}.

For the post analysis, the Fisher matrix is regularised with scheme number 3 with a cut-off of 10^{-3}. i.e Only those eigenvalues that are a thousandth of the largest eigenvalue, or larger, are retained, while the rest are set to the smallest eigenvalue kept. Fig. 7.3 shows the clean maps from all cross-correlations between A, E and T, and all their networks. Comparing these with the injected map in Fig. 7.1, it can be seen that the point source's location is recovered, and the normalisation of the map is not that far off.

7.1.2 In the presence of LISA noise

In this demonstration, the point source above is used again to simulate a signal. This signal is scaled with three different appropriate factors to mimic point sources of three different powers. The scaled signal time-series is then injected into simulated LISA noise, and the output passed through the analysis pipeline. From Fig. 6.3, it can be seen that the maximum value in the LISA noise spectral density is $\sim 10^{-38}$, whilst the power of the injected map in Fig. 7.1 is ~ 1. By scaling the signal time-series simulated from this map by 10^{-18}, 10^{-20} and 10^{21}, the power of the point source (in the injected map) is effectively brought down to $\sim 10^{-36}$, $\sim 10^{-40}$ and $\sim 10^{-42}$, respectively. In the first case, the point source's power is larger than LISA's noise; in the second case, its power is 'comparable' to LISA's noise; in the third case, its power is smaller than LISA's noise. In Fig. 7.4, the clean map from the AE cross-correlation for these three cases are shown. It is seen that only in the first case, where the injected map's power is

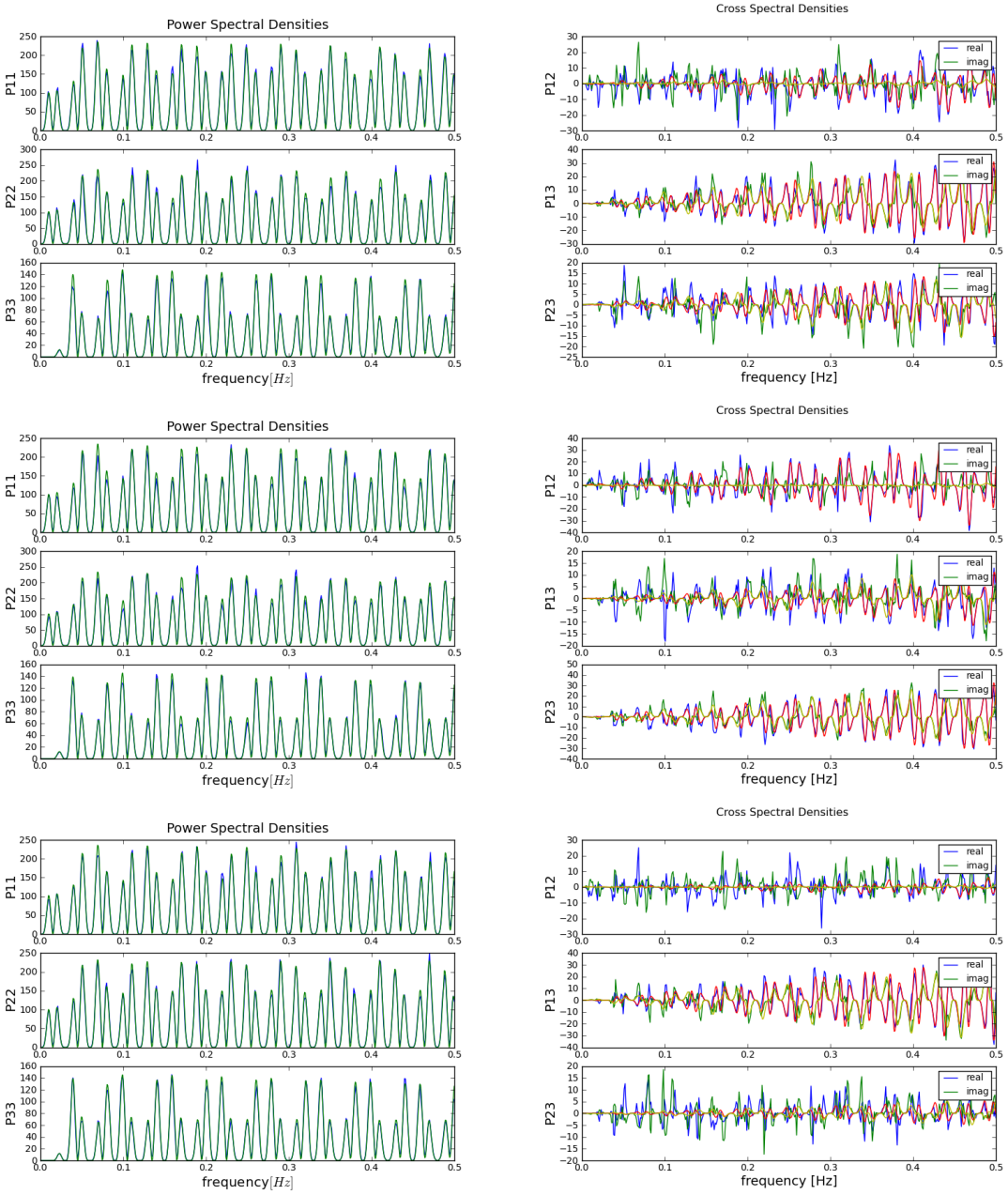

FIGURE 7.2: Auto and cross spectral densities of simulated GW signals. The source is the point source-plus-uniform distribution in Fig. 7.1. The three signals are the optimal A, E and T TDI observables defined (3.40), (3.41) and (3.42) using $X2$, $Y2$ and $Z2$. In the y-axis labels, the numbers 1, 2 and 3 correspond to A, E and T, respecively. (*Left*)Power spectral densities; (*green*)analytical PSDs; (*blue*)estimated PSDs. (*Right*)Cross spectral densities; (*red*)analytical real part; (*yellow*)analytical imaginary part; (*blue*)estimated real part; (*green*)estimated imaginary part. (*Top*)Day 66. (*Middle*)Day 166. (*Bottom*)Day 266. The Pwelch periodogram method is used for PSD and CSD estimation; segments of 2 hour duration with 50% overlap are used.

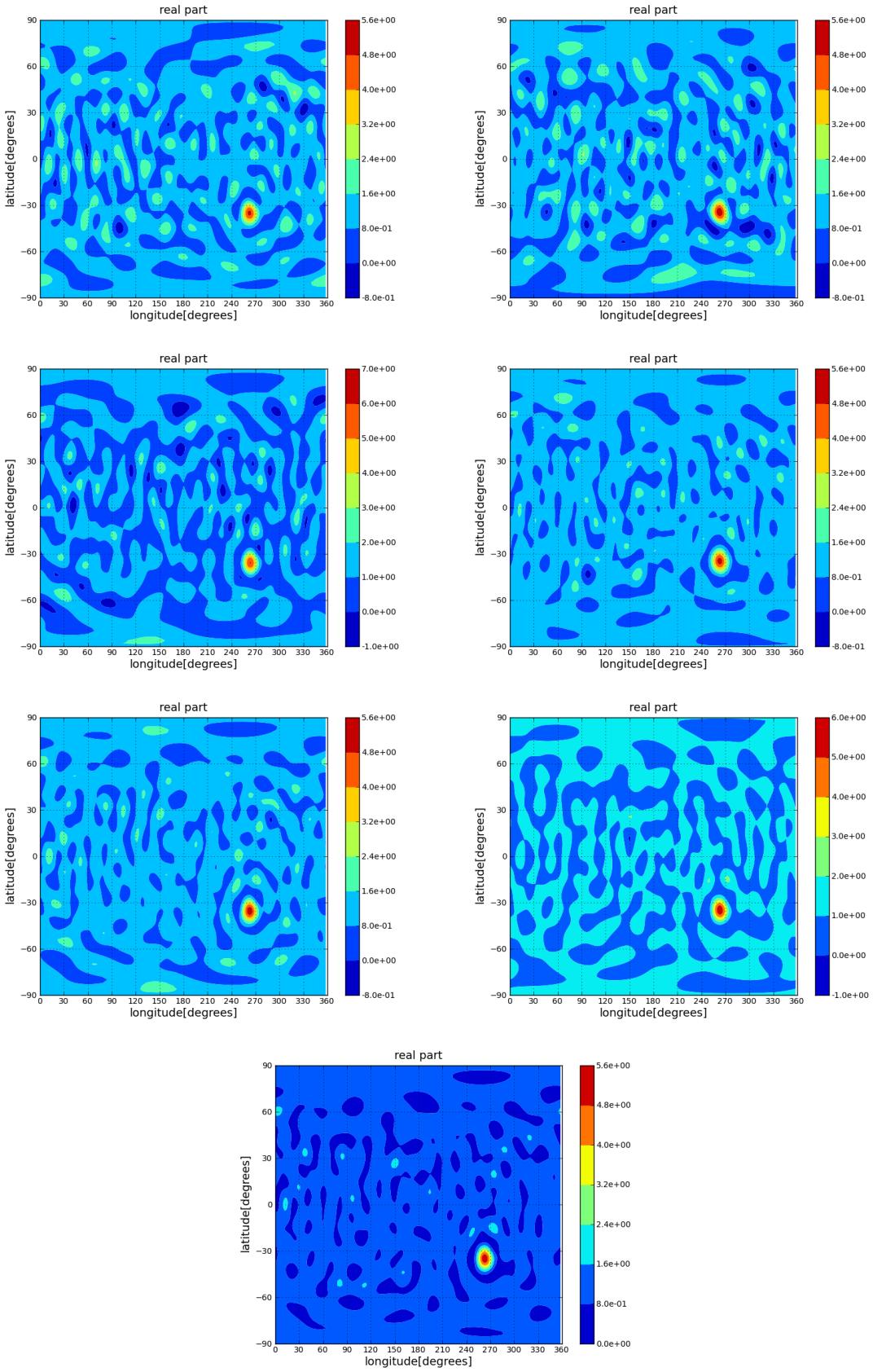

FIGURE 7.3: Clean map from cross-correlations and networks AE, AT, ET, $\{AE, AT\}$, $\{AE, ET\}$, $\{AT, ET\}$ and $\{AE, AT, ET\}$. Fisher matrix is regularised with scheme number 3 with a cut-off of 10^{-3}.

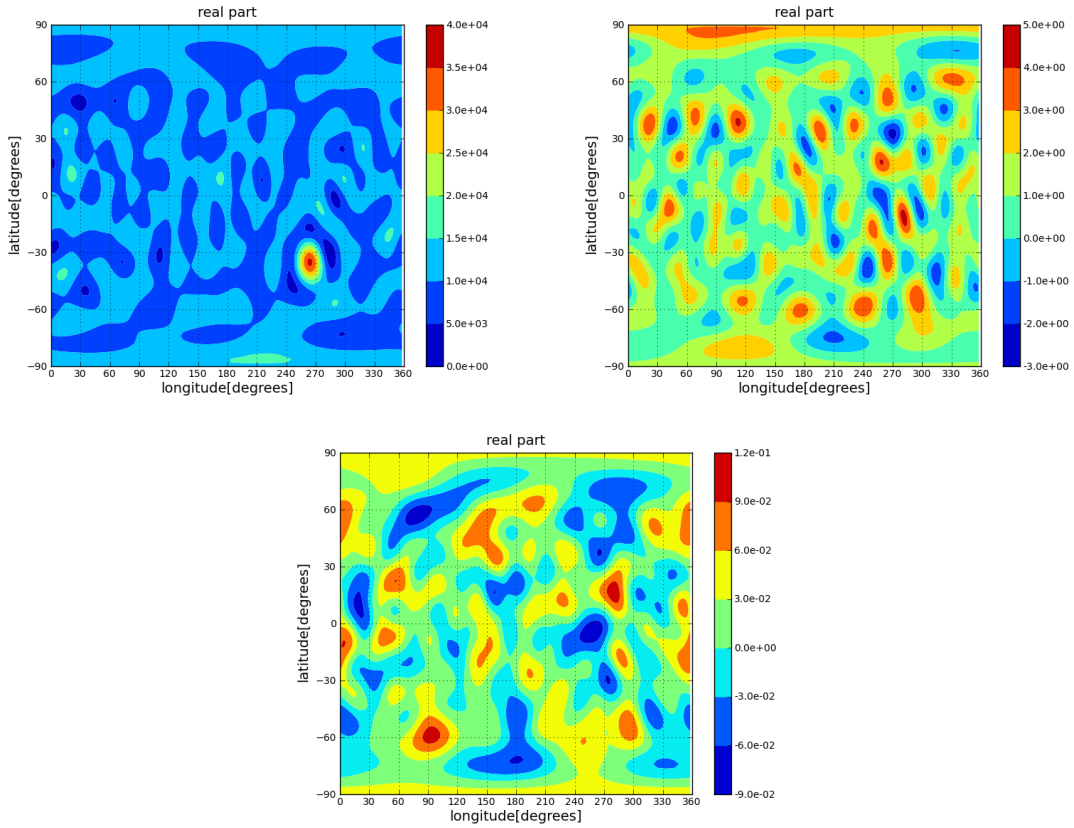

FIGURE 7.4: Clean maps from the cross-correlation of AE, $\hat{P}^{AE}_{lm}$, for different strengths of the point source relative to LISA's noise. Fisher matrix is regularised with scheme number 3 with a cut-off of 10^{-3}. *top left*: $\mathcal{P}/P^{noise} \sim 10^{2}$; *top right*: $\mathcal{P}/P^{noise} \sim 10^{-2}$. *bottom*: $\mathcal{P}/P^{noise} \sim 10^{-4}$.

about 100 times that of LISA's noise, is the point source visible. In the other two cases, the point source is not visible. Prior to the analyses carried out here, the time-series is first scaled by a factor of 10^{20}[2]. This factor effectively scales up the power of the injected map by 10^{40} in each of the above cases, so they become 10^{4}, 1 and 10^{-2}, in the same order as above. These values are reflected in the normalisation of the clean maps, even in the ones where the point source is not visible.

In the second case above, even though the point source is not visible in the clean map from the cross-correlation of AE, it is in the network $\{AE, AT\}$ and the network $\{AE, AT, ET\}$, as shown in Fig. 7.5

7.2 Isotropic background

In the following demonstrations, an isotropic background is considered. The injected map $\mathcal{P}(\hat{k})$ is created by dividing the sky into 360 ecliptic longitudes and 180 ecliptic latitudes. For

[2]This is so that the samples have values close to unity, in order to avoid the implementation working with extremely small and extremely large numbers.

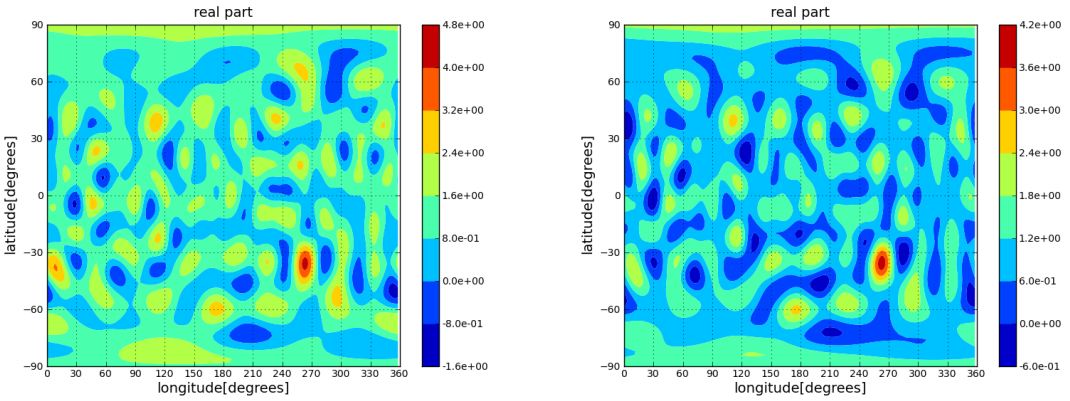

FIGURE 7.5: Clean maps from $\{AE, AT\}$ and $\{AE, AT, ET\}$ for $\mathcal{P}/P^{noise} \sim 10^{-2}$. Fisher matrix is regularised with scheme number 3 with a cut-off of 10^{-3}.

an isotropic background, each pixel is assigned the same value of power (7.2).

7.2.1 In the absence of LISA noise

In this example, $\mathcal{A}_{uniform} = 10^{-34}$ s. By taking $H(f) = 1$, the signals of this source are simulated for the optimal TDI observables A, E and T (of the second generation Michelson TDI observables $X2$, $Y2$ and $Z2$) with $l_{max} = 20$ and $\Delta t = 1$ s. The same set of computed overlap-reduction values as before are used, ranging from 10^{-4} Hz to 10^{-1} Hz, with a frequency resolution of 10^{-4} Hz. Again, time-series of duration $\tau = 1$ day are simulated for 365 days of the year.

For the analysis, $H(f) = 1$ (same as in simulation). Hanning windows are applied to the time-series prior to computing the cross-correlated output, and the lower and higher limits for the frequency integration are $f_{low} = 1.5 \times 10^{-14}$ Hz and $f_{high} = 4.998 \times 10^{-1}$ Hz. In the post analysis, the Fisher matrix is regularised by retaining the eigenvalues which are equal or larger than a thousanth the largest eigenvalue, and setting the rest to the smallest eigenvalue retained. In Fig. 7.6, the clean maps from all cross-correlations between A, E and T, and all their networks, are shown. The clean maps in general display a randomly fluctuating pattern across the sky. It can be seen, from the colour bars to the side of each projection, that the median of this fluctuation is about 10^6 s, which, taking into account that the time-series are scaled by 10^{20} prior to analysis, is consistent with the injected $\mathcal{A}_{uniform}$ of 10^{-34} s.

7.2.2 In the presence of LISA noise

In this demonstration, the above signal time-series are appropriately scaled to obtain signals for which the isotropic background has effectively $\mathcal{A}_{uniform} = 10^{-50}$ s, $\mathcal{A}_{uniform} = 10^{-40}$ s and $\mathcal{A}_{uniform} = 10^{-36}$ s.

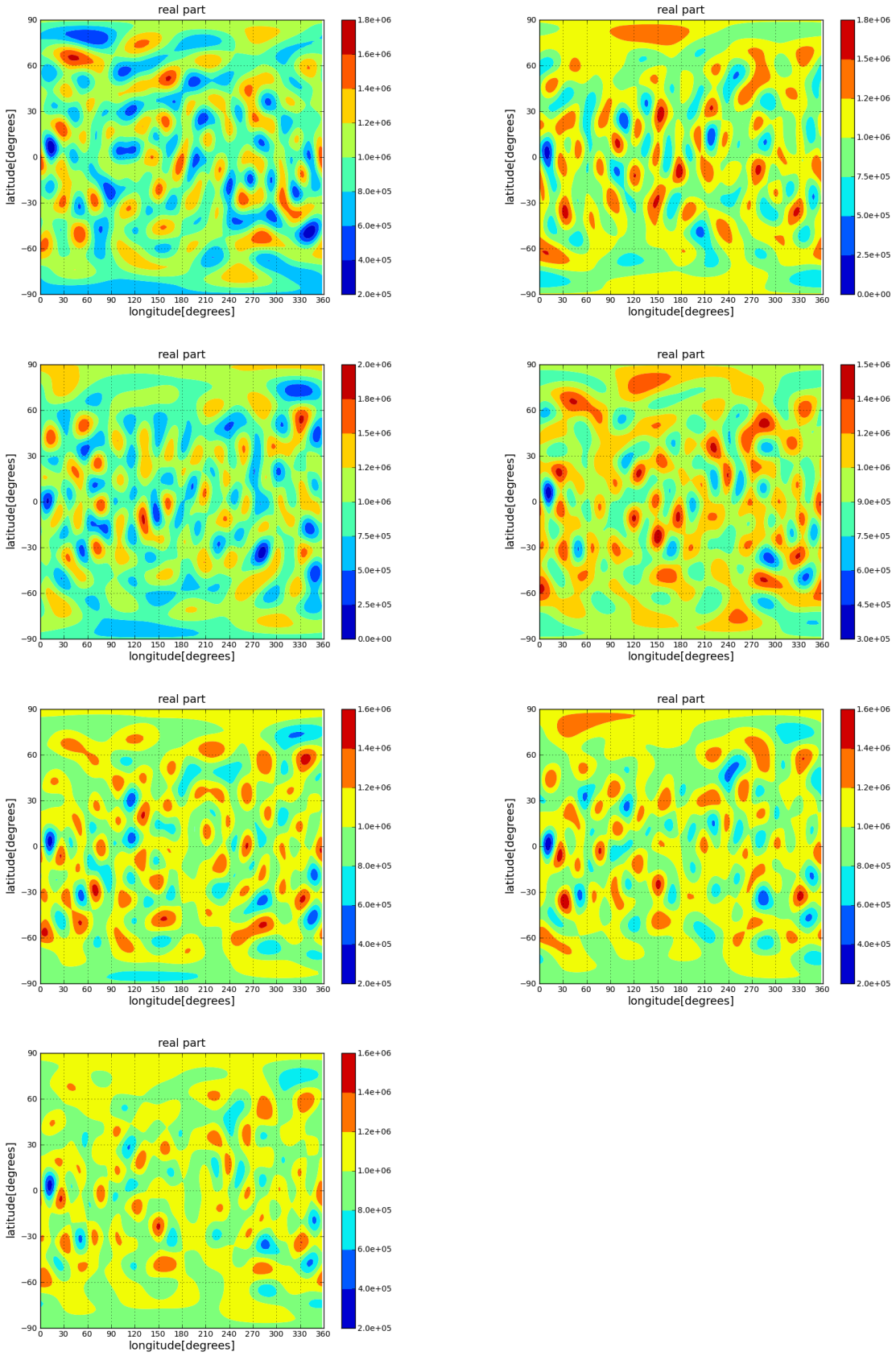

FIGURE 7.6: Clean map from cross-correlations and networks AE, AT, ET, $\{AE, AT\}$, $\{AE, ET\}$, $\{AT, ET\}$ and $\{AE, AT, ET\}$. Fisher matrix is regularised with scheme number 3 with a cut-off of 10^{-3}.

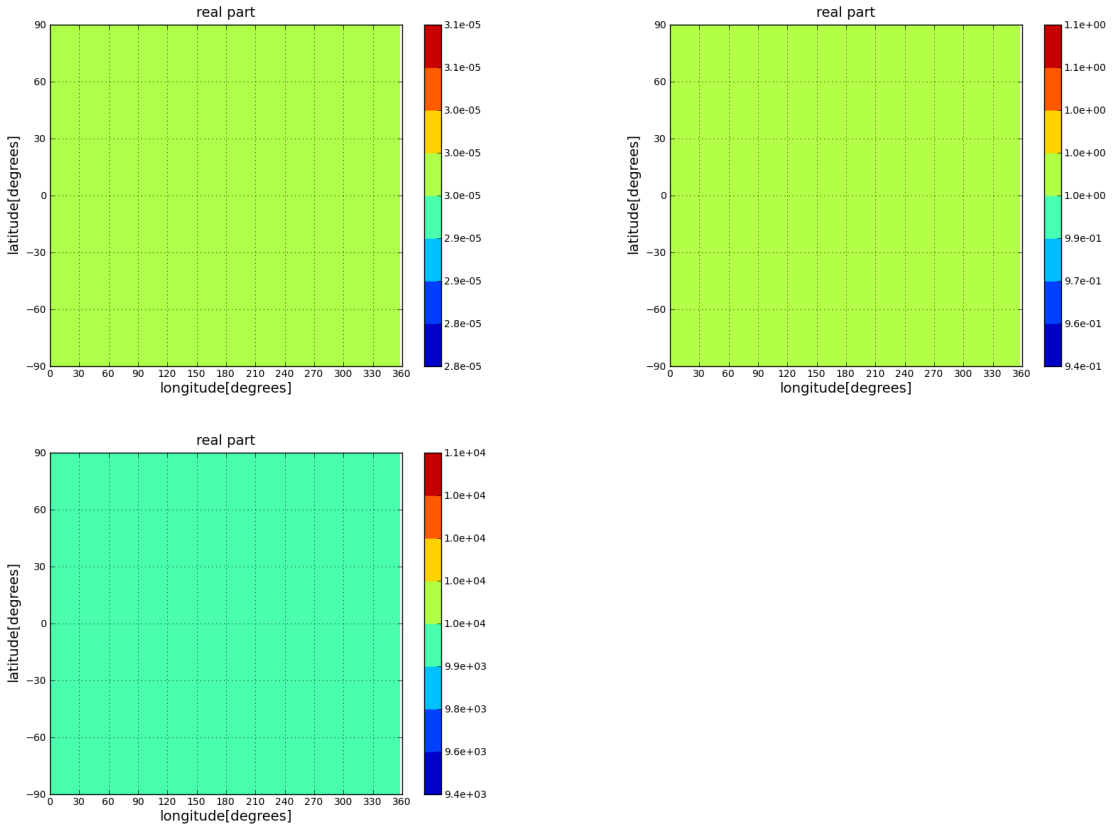

FIGURE 7.7: Clean map from cross-correlation AE for different values of injected monopole. (*top*): $\mathcal{A}_{uniform} = 10^{-50}$. (*middle*): $\mathcal{A}_{uniform} = 10^{-40}$. (*bottom*): $\mathcal{A}_{uniform} = 10^{-36}$

All the analysis parameters are the same as before, except that here $l_{max} = 0$. In Fig. 7.7, the clean maps from the cross-correlation of AE are shown. By restricting l to 0 in the analysis, the resultant clean map necessary contains only the monopole moment, so there are no fluctuations as in Fig. 7.6. By taking into account that the time-series are scaled by 10^{40} prior to analysis, it is seen here from the colour bars beside the projections that the clean maps' normalisations agree reasonably well with the injected $\mathcal{A}_{uniform}$s.

Chapter 8

Discussion

8.1 Which l_{max}?

For any source background, the choice of l_{max}, whether in the simulation or in the subsequnt analysis, is somewhat arbitrary. A point source contains multipole moments up to $l = \infty$ by definition. In this case, the choice of l_{max} can depend on how fine a spatial resolution is desired. It can also be limited by the amount of computer resources, or time, available. For diffuse sources, l_{max} might not have to be that large at all. For an isotropic background, there is really no need for l_{max} to be larger than zero. In fact, by having too large an l_{max} than is needed to describe a background, additional uncertainties can be introduced to the estimates from correlations with the extra multipole moments [18]. In practice, there might not be enough knowledge about the source to choose a sufficient l_{max}.

One way of choosing l_{max} that has been suggested is to look at the sky-averaged standard deviation of $\mathcal{P}(\hat{k})$, per number of independent spots in the sky, per steradian, formally defined as:

$$\mathcal{K}(l_{max}) = \frac{1}{4\pi} \left(\int_{S^2} \mathrm{d}\hat{k}\, \sigma_{\hat{\mathcal{P}}(\hat{k})} \right) \frac{4\pi}{(l_{max}+1)^2} \,, \tag{8.1}$$

where $\sigma_{\hat{\mathcal{P}}(\hat{k})}$ is the standard deviation of the clean map in pixel basis, given by (6.25); since it is computed from the Fisher matrix, it is also dependent on l_{max}. It has been observed in the context of LIGO that $\mathcal{K}(l_{max})$ exhibits a minimum turning point between $l_{max} = 0$ and $l_{max} = 20$ for a noise-only injection. Since at this point, there is least uncertainty in the estimate of $\mathcal{P}(\hat{k})$, l_{max} is said to be optimal. In this way, it can be justified that this l_{max} is chosen for the implementation.

Similarly, by simulating LISA noise, analysing it, and computing $\mathcal{K}(l_{max})$, an optimal l_{max} is looked for. In Fig. 8.1, $\mathcal{K}(l_{max})$ is plotted between $l_{max} = 0$ and $l_{max} = 15$ for all the cross-correlations and networks. It is seen that there are *no* obvious minimum turning points in these plots, so an optimal l_{max} cannot be identified.

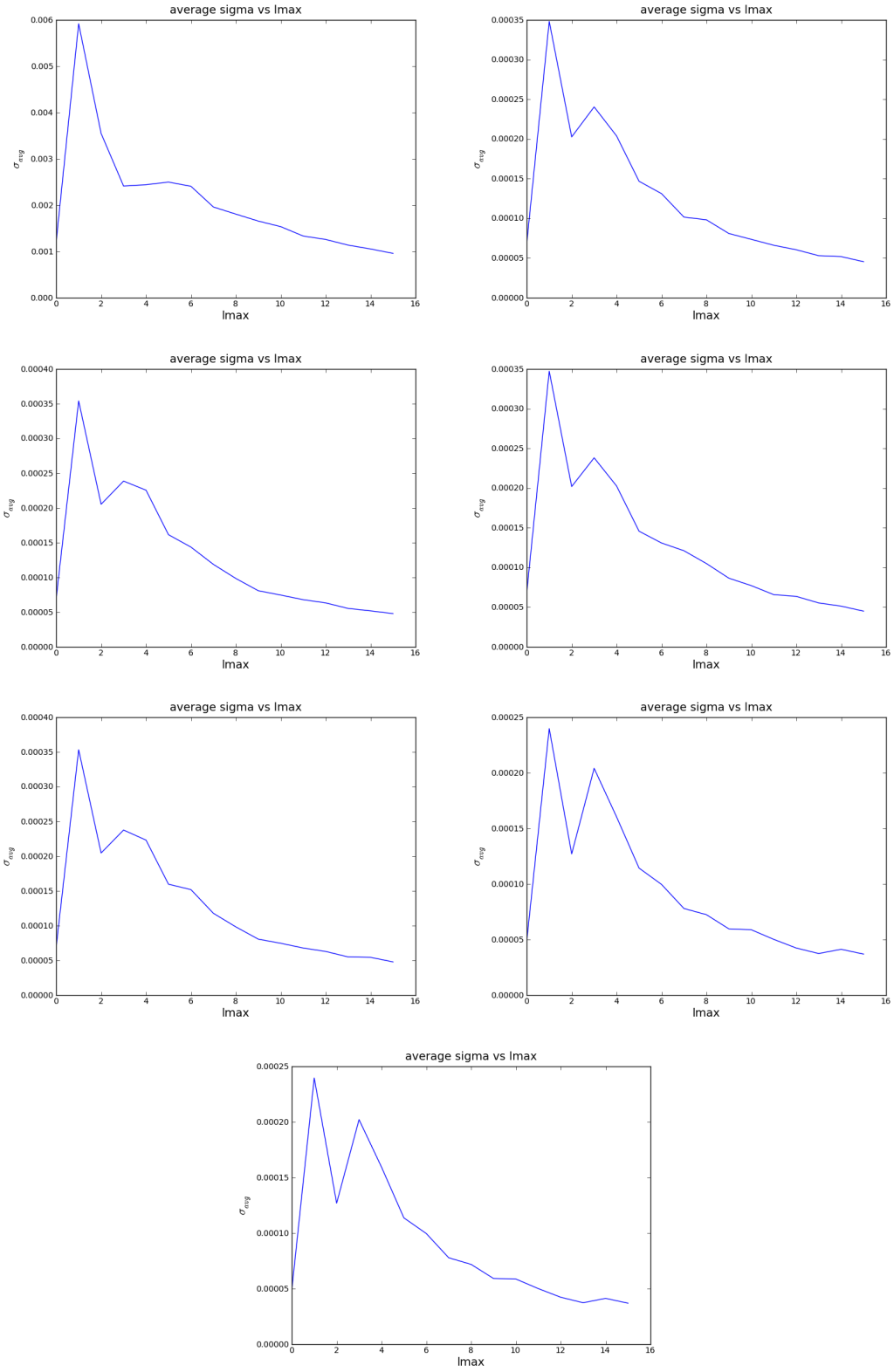

FIGURE 8.1: Sky-averaged standard deviation of pixel estimates of LISA's noise-only background from cross-correlations and networks: AE, AT, ET, $\{AE, AT\}$, $\{AE, ET\}$, $\{AT, ET\}$ and $\{AE, AT, ET\}$.

8.2 which $H(f)$?

Two other parameters that need to be chosen prior to analysis are f_R and β which describe the power-law GW power spectrum $H(f)$ in (5.5). In chapter 7, the $H(f)$ used in analysis is always the same as in simulation, in order for the Fisher matrix inverse to *invert* the one used in simulation (see (5.30) and (5.34)). In practice however, the $H(f)$ of the source background is not known, and a helpful thing to do might be to use Bayseian model selection to select it. In fact, this also applies to l_{max}. A given $H(f)$ and l_{max} specifies the signal model used in the analysis (described by (5.1) and (5.2)). It might be possible to compute the probability density of different models given the data, $p(\, M(H(f), l_{max}) \,|\, C_{IJ}(f,t)\,)$ with $(l_{max}+1)^2$, and use the $H(f)$ and l_{max} of the model with the highest probability [18].

8.3 Which regularisation?

In Fig. 8.2, eigenvalues of Fisher matrices obtained from time-series containing only LISA noise are shown for all the cross-correlations and networks. In the plots, it can be seen that the ratio of the largest to the smallest eigenvalues is about 10^5 for all cross-correlations. This is expected, as the overlap-reductions functions are distinctively blind to multipole moments of either even or odd l. Note also that, because of this, truncating a Fisher matrix to a smaller l_{max} does not really decrease the condition number. The large condition number of 10^5 implies that the system of linear equations (5.34) is *ill-conditioned*. As mentioned in chapter 6, this can be improved by regularising the Fisher matrix. How to do the regularisation is also down to choice. It is about finding the right balance between reducing the uncertainty due to ill-condition, by throwing away enough of the smallest eigenvalues, and not throwing away too many that the beam (of the Fisher matrix) loses its charateristic sensitivity[1]. Here, through trial-and-error, it is decided that, in general, eigenvalues which are less than a thousandth of the largest are cut off. If the type of source is known, as is in the cases here where the signal is simulated, the best thing to do perhaps is to adjust the regularisation cut-off, and pick the one that gives the best estimates. In practice, where it is not known what type of source to expect, it is not clear how this can be done.

8.4 Normalisation problem

In the examples demonstrated in chapter 7, it is seen from their projections that the clean maps resemble their injected sources pretty well, both in the spatial features revealed and in

[1] In the extreme, if only the largest eigenvalue is kept with the rest set equal to it, the regularised Fisher matrix becomes proportional to the identity matrix. In this case, there is no difference between the spatial pattern of the clean map and the dirty map.

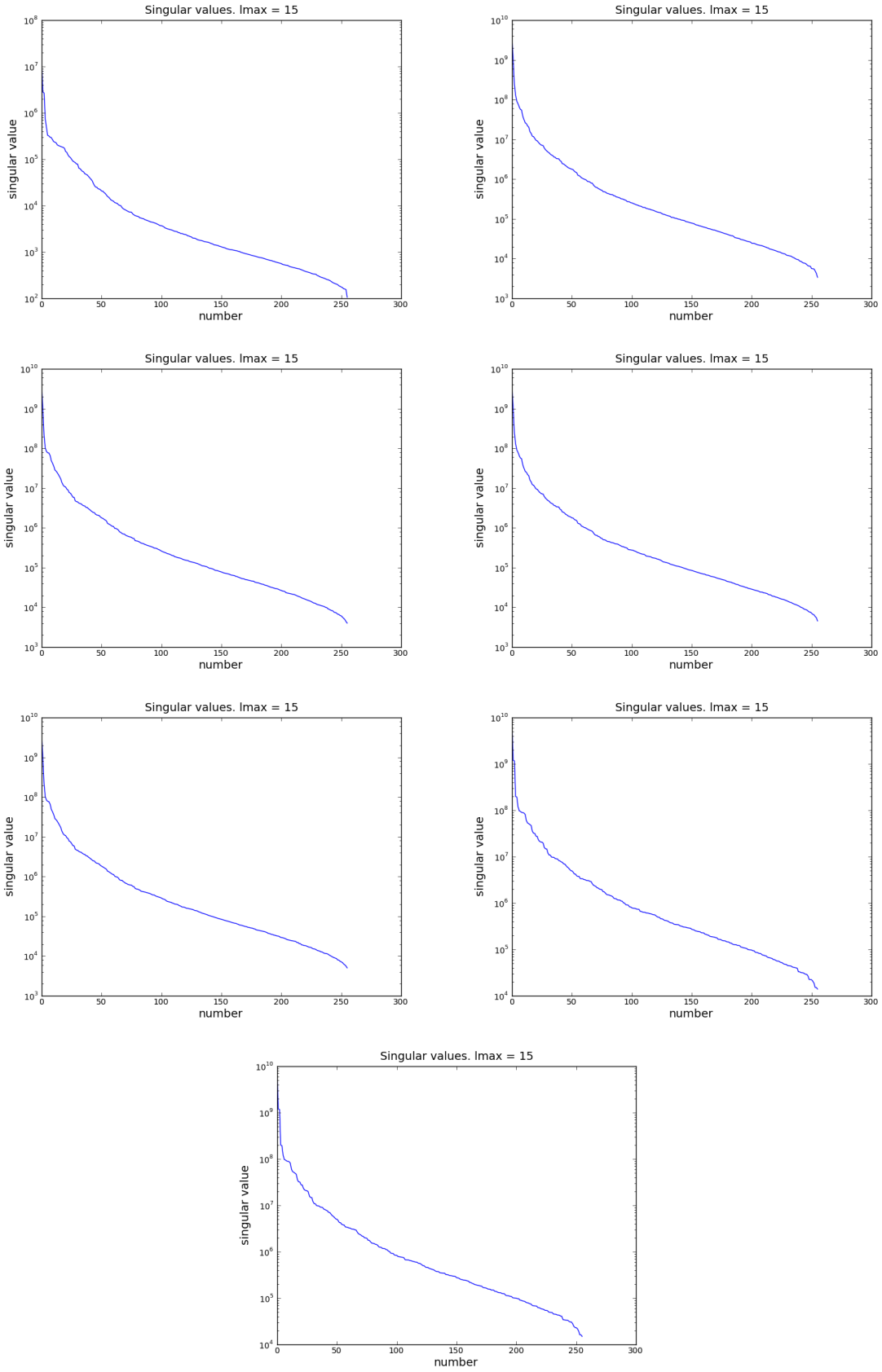

FIGURE 8.2: Eigenvalues of Fisher matrices from LISA noise-only outputs for cross-correlations and networks: AE, AT, ET, $\{AE, AT\}$, $\{AE, ET\}$, $\{AT, ET\}$ and $\{AE, AT, ET\}$.

the normalisation. This is true for both the point source and for the isotropic source, whose simulation and analysis are demonstrated in the previous chapter.

As a formal verification of the overall normalisation of the clean map, a 'close enough' agreement is looked for between the monopole moment of the clean map, $\hat{\mathcal{P}}_{00}$, and the monopole moment of the injected map, $\mathcal{P}_{00}$. By simulating multiple realisations, if about 69% of the $\hat{\mathcal{P}}_{00}$s fall within one standard deviation of $\mathcal{P}_{00}$, then the overall noramlisation is said to be satisfactory. From simulation of each realisation, a clean map and standard deviation estimate for the monopole moment, $\hat{\mathcal{P}}_{00}$ and $\hat{\sigma}_{00}$ respectively, are obtained. If $|\hat{\mathcal{P}}_{00} - \mathcal{P}_{00}| \leq \hat{\sigma}_{00}$, then $\hat{\mathcal{P}}_{00}$ from this realisation is said to fall within the one standard deviation. Whilst it is expected that the $\hat{\mathcal{P}}_{00}$s from different realisations will fluctuate about $\mathcal{P}_{00}$, the $\hat{\sigma}_{00}$s are expected to remain fairly constant, with a value close to the analytical standard deviation, σ_{00}.

Since the comparison is only made between monopole moments, it is natural to consider the isotropic source here, which, by definition, contains only the monopole moment. In this case, l_{max} can be set to zero. This means that the signal-simulation and subsequent analysis can be done in a relatively short time, allowing multiple realisations to be considered within a realistic time frame. For the covariance matrix computation in signal-simulation (6.11), the convolution of the overlap-reduction function with the injected map now only involves one multipole moment, instead of $(l_{max}+1)^2$ multipole moments. In the main analysis, the main data products become

$$X_{00}^{IJ}(t) = \frac{\overline{w_I w_J}^2}{\overline{w_I^2 w_J^2}} \tau \Delta f \sum_f \gamma_{00}^{IJ*}(f,t) \frac{H(f)}{P_I(f,t)P_J(f,t)} C^{IJ}(f,t) \,, \tag{8.2}$$

$$\Gamma_{00,00}^{IJ}(t) = \frac{\overline{w_I w_J}^2}{\overline{w_I^2 w_J^2}} \tau \Delta f \sum_f \gamma_{00}^{IJ*}(f,t) \frac{H^2(f)}{P_I(f,t)P_J(f,t)} \gamma_{00}^{IJ}(f,t) \,, \tag{8.3}$$

$$\Psi_{00,00}^{IJ}(t) = \frac{\overline{w_I w_J}^2}{\overline{w_I^2 w_J^2}} \tau \Delta f \sum_f \gamma_{00}^{IJ*}(f,t) \frac{H^2(f) P_{IJ}^{GW2}(f,t)}{P_I^2(f,t)P_J^2(f,t)} \gamma_{00}^{IJ}(-f,t) \,, \tag{8.4}$$

where $X_{00}^{IJ}(t)$, $\Gamma_{00,00}^{IJ}(t)$ and $\Psi_{00,00}^{IJ}(t)$ are the *daily* dirty map, Fisher matrix and 'strong-signal bias' matrix[2].

From simulation of multiple realisations of the isotropic source with real LISA noise, analysed in the previous chapter, the overall normalisation of the implementation *cannot* be verified in the sense described above. In fact, not only is there not a high enough percentage of $\hat{\mathcal{P}}_{00}$s falling within one standard deviation, the results also show that the estimated standard deviations, $\hat{\sigma}_{00}$, are smaller than the expected σ_{00}, even with the inclusion of the 'strong-signal bias matrix', $\Psi_{\alpha\beta}^{IJ}$.

[2]Notice that t denotes a day, and there is no sum over it like in (6.16), (6.18) and (6.19). The implementation saves to disk both these *daily* estimates as well as their sums over the number of days during the observation period τ, X_{α}^{IJ}, $\Gamma_{\alpha\beta}^{IJ}$ and $\Psi_{\alpha\beta}^{IJ}$

This issue remains unresolved because the exact cause of the discrepancy is difficult to track down. In (8.2), (8.3) and (8.4), if the expected $P_I(f,t)$, $P_J(f,t)$ and $C^{IJ}(f,t)$ (based on analytical expressions) are used instead of the estimated ones, $\hat{\mathcal{P}}_{00} = \mathcal{P}_{00}$. This shows at least that the normalisation is correct in expressions (8.2), (8.3) and (8.4) which are used in the analysis pipeline, and suggests that the problem could lie within the simulation pipeine.

In a more systematic approach to finding the cause, the simple example of cross-correlating two detector outputs with uncorrelated white noise and same white signal is examined and understood. By gradually generalising this example towards the realistic case above, it is hoped that the normalisation of the implementation fails at some point along the way, and that the source of the failure can be pinpointed at the same time. Several of such examples are here described and reviewed:

8.4.1 Case 0: uncorrelated white noise + same white signal

Suppose $y_I(t)$ and $y_J(t)$ are the outputs of two detectors, I and J, with uncorrelated white noise and the same signal:

$$y_I(t) = n_I(t) + h(t) \,, \tag{8.5}$$

$$y_J(t) = n_J(t) + h(t) \,, \tag{8.6}$$

where n_I, n_J and h are all white and uncorrelated with each other, with variances $\sigma^2_{n_I}$, $\sigma^2_{n_J}$ and σ^2_h, respectively. This is akin to having two coinicident and coaligned gravitational wave detectors. The aim here is to simulate the time-series y_I and y_J, and then estimate from these the variance of the signal, σ^2_h. Since the noise and signal are all white, they can be simulated simply by drawing sequences of random numbers from the standard normal distribution, and then appropriately scaling them with either σ_{n_I}, σ_{n_J}, or σ_h. The data used in this example are therefore *independent* of the entire signal-simulation pipeline, and since the samples in the time-series are obtained directly, this type of simulation will be referred to as simulation in the *time-domain*. Two methods of estimation of σ^2_h are used here. The first one is to form a simple cross-correlation of y_I and y_J. The estimator in this case is

$$\hat{\sigma^2_h} = \frac{1}{N} \sum_{i=1}^{N} y_I[i]\, y_J[i] \,, \tag{8.7}$$

where $y_I[i]$ and $y_J[i]$ are the ith sample of y_I and y_J, respectively, and N is the number of samples in them. The variance of this estimator, according to the definition $\langle \hat{\sigma^2_h} \hat{\sigma^2_h}^* \rangle - \langle \hat{\sigma^2_h} \rangle \langle \hat{\sigma^2_h} \rangle^*$, is

$$\sigma^2_{\hat{\sigma^2_h}} = \frac{1}{N} \left((\sigma^2_{n_I} + \sigma^2_h)(\sigma^2_{n_J} + \sigma^2_h) + \sigma^4_h \right) \,. \tag{8.8}$$

It is found that the sample variance of $\hat{\sigma_h^2}$ from multiple realisations agree closely with the analytical variance given by (8.8), and that $\sim 69\%$ of $\hat{\sigma_h^2}$ lie within the sample variance (or the analytical variance $\sigma^2_{\hat{\sigma_h^2}}$). So, the normalisation is consistent here.

The second method uses the sum of the cross-power weighted with some filter to estimate σ_h^2. In addition, it provides an expression from which the variance of $\hat{\sigma_h^2}$ can be estimated from the data. By considering the dimensions of various quantities (σ_h^2, a variance, is dimensionless), it can be shown that

$$\hat{\sigma_h^2} = \frac{X}{G} \,, \tag{8.9}$$

where

$$X = \tau\Delta f \sum_f \frac{H(f)C_{IJ}(f)}{P_I(f)P_J(f)} \,, \tag{8.10}$$

$$G = \tau\Delta f \sum_f \frac{H^2(f)}{P_I(f)P_J(f)} \,, \tag{8.11}$$

is an unbiased estimator of σ_h^2. $C_{IJ}(f)$ is the (*one-sided*) cross-correlated output, $(2/\tau)\tilde{y}_I^*(f)\tilde{y}_J(f)$, as defined in (5.10). Because the noise are uncorrelated, $\langle C_{IJ}(f)\rangle = P_h(f)$, where $P_h(f)$ is the power spectral density of the signal. $\langle\hat{\sigma_h^2}\rangle = \sigma_h^2$ also implies that $P_h(f) = H(f)\sigma_h^2$, so $H(f)$ describes the shape of the GW spectrum. And because $\hat{\sigma_h^2}$ is dimensionless, X and G necessarily have the same unit, and $H(f)$ has the unit of *time*. $P_i(f)$ is the *one-sided* power spectral density of y_i, estimated using pwelch methods. The factor $\tau\Delta f$ is used in the definitions here so that $1/G$ gives the variance of $\hat{\sigma_h^2}$. Note that these definitions closely resemble the maximum likelihood estimators used in the implementation ((**??**), (5.32) and (**??**)). The expected variance of the estimator in (8.9) can be shown to be

$$\sigma^2_{\hat{\sigma_h^2}} = \frac{1}{G}\left(1 + \frac{\tau\Delta f(2\Delta t)^4\sigma_h^4 \sum_f \frac{1}{P_I^2(f)P_J^2(f)}}{G}\right) \,, \tag{8.12}$$

where Δf and Δt are the frequency and time cadences of the power spectral densities and the time-series, respectively. This expression can be used to estimate the variance of $\hat{\sigma_h^2}$ from data. By further using the fact $P_I(f) = 2\Delta\sigma_I^2$ for white processes, the expected (analytical) variance is found to be

$$\sigma^2_{\hat{\sigma_h^2}} = \frac{1}{N}\left((\sigma^2_{n_I} + \sigma_h^2)(\sigma^2_{n_J} + \sigma_h^2) + \sigma_h^4\right) \,. \tag{8.13}$$

Note that it is the same as (8.8)[3]. If τ is taken to be the duration of a single 'measurement', like the *one*-day in the implementation, then, by repeating (8.9) to (8.12) for multiple measurements, a set of σ_h^2 estimates, $\{\hat{\sigma_{h,i}^2}\}_{i=1,\ldots,M}$, and their variances, $\{\sigma^2_{\hat{\sigma_h^2},i}\}$, are collected.

[3] If there was not the factor $\tau\Delta f$ in (8.11), there would have only been a factor of $1/N_f$ (N_f is the number of frequencies in the sum over frequencies), instead of $1/N$, in (8.13). $N_f = N/(2N_{avg}) = N/(2\tau\Delta f)$.

From these, one can form *optimal* estimates with

$$\hat{\sigma^2_h}_{,opt} = \frac{\sum_{i=1}^{M} \left(\hat{\sigma^2_h}_{,i} \;/\; \sigma^2_{\hat{\sigma^2_h},i} \right)}{\sum_{i=1}^{M} \left(1 \;/\; \sigma^2_{\hat{\sigma^2_h},i} \right)} \,, \tag{8.14}$$

$$\sigma^2_{\hat{\sigma^2_h},opt} = \frac{1}{\sum_{i=1}^{M} \left(1 \;/\; \sigma^2_{\hat{\sigma^2_h},i} \right)} \,. \tag{8.15}$$

In the weak-signal limit, the second term in the bracket in the RHS of (8.12) can be ignored, and $\sigma^2_{\hat{\sigma^2_h}} = 1/G$. Substituting this and (8.9) in (8.16) and (8.17), they become

$$\hat{\sigma^2_h}_{,opt} = \frac{\sum_{i=1}^{M} X_i}{\sum_{i=1}^{M} G_i} \,, \tag{8.16}$$

$$\sigma^2_{\hat{\sigma^2_h},opt} = \frac{1}{\sum_{i=1}^{M} G_i} \,, \tag{8.17}$$

where X_i and G_i are the X and G in (8.10) and (8.11) from the ith measurement. Comparing these with (5.30) and (5.37), it can be seen that the nominator in the RHS of (8.16) is analogous to X^{IJ}_{00}, and the denominator is analogous to $\Gamma^{IJ}_{00,00}$. This connection means that this method can be potentially generalised for realistic LISA case. In practice, if the M measurements are interpreted as all the measurements that are available in the whole observation or experiment, then the analysis pretty much ends at (8.16) and (8.17). We can repeat the M measurements, but the results can only be interpreted as optimal estimators from a different realisation. It does not make sense to gather these optimal estimators from different realisations and use (8.16) and (8.17) to form another set of optimal estimators. However, one can look at several of these to see if there is a sufficient number of realisations for which $\hat{\sigma^2_h}_{,opt}$ falls within one $\sigma_{\hat{\sigma^2_h},opt}$ of σ^2_h, in order to verify the normalisation of the method.

To avoid confusion in what follows, the variances evaluated completely independent of data, such as (8.8) and (8.13), will be referred to as the *analytical* variances. These will have the expected values to which our estimations from data will be compared. The sample variance of σ^2_2 estimates from a set of realisations will be referred to as the *empirical* variance[4] Finally, there are the variances obtained by combining several quantities, estimated from data, according to some analytical expression, such as (8.12). These will be referred to as *theoretical* variances. It has been observed that the theoretical variances tend to be smaller than the empirical variances. This is actually due to the bias in PSD estimation. As the number of pwelch averages increases, the theoretical variances (which depends on estimated PSDs) tend

[4]This can often be computed directly using built-in functions in alreay existing computing packages, such as numpy.var().

towards the empirical variances. Roughly, when 20 averaging segments are used in PSD esimtaion, there is about $1/20 = 5\%$ difference between the theoretical variances and the empirical variances.

The results of applying (8.16) and (8.17) to 1000 realisations of y_I and y_J is summarised in Fig. 8.3. Here, σ_n^2 and σ_h^2 are set to 1.0 and 0.1, respectively. It can be seen that the estimated σ_h^2 and $\sigma^2_{\sigma_h^2}$ fluctuate evenly about their expected values, and the hitogram shows that they are consistent with each other.

8.4.2 Case 1: uncorrelated white noise + same white signal

Whilst the above example demonstrates a correctly normalised cross-correlation technique for the particular choice of y_I and y_J, ultimately, we would like to find where the normalisation fails within the implementation described in chapter 6. Therefore, the next thing to try is to accommodate (8.5) and (8.6) in the implementation, and see if the implementation gives the correct normalisation.

Since the same gravitational wave signal $h(t)$ is present in $y_I(t)$ and $y_J(t)$, their signal covariance matrix is expected to be of the form:

$$\mathbf{C}(f) = P_h(f) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} . \tag{8.18}$$

$P_h(f)$ is here taken to be the *one-sided* power spectral density of $h(t)$. For a white $h(t)$, $P_h(f)$ is constant in frequency, and by comparing with (6.11), it can be seen that we can choose $H(f) = H_0$ and $\gamma_\alpha^{IJ}(f) = \gamma_{00}^{IJ}$, where H_0 and γ_{00}^{IJ} are constant values. In the implementation, the real LISA overlap-reduction functions are now replaced with these constant-value overlap-reduction functions. By using the fact that $\sigma_h^2 = \int_0^{+\infty} df P_h(f)$ and (5.25), it can be shown that σ_h^2 can be written as:

$$\sigma_h^2 = \frac{\gamma_{00}^{IJ} H_0}{2\Delta t} \mathcal{P}_{00} = \gamma_{00}^{IJ} H_0 f_{Nyq} \mathcal{P}_{00} , \tag{8.19}$$

where f_{Nyq} is the Nyquist frequency and Δt is the time cadence of the time-series. This allows H_0, γ_{00}^{IJ}, Δt and the injected monopole, $\mathcal{P}_{00}$, to be chosen suitably, given σ_h^2. Similarly, in place of LISA noise spectral densities, constant white noise spectral densities are used in the noise-simulation. Recall from section 6.2.1 that the signal time-series can be simulated by stitching together shorter stetches. In order to test the simulation method progressively, this is *not* done at this stage. Note also that it is possible to do this withouth stitching within reasonable time because the overlap-reduction functions $\gamma_\alpha^{IJ}(f) = \gamma_{00}^{IJ}$ are relatively simple functions, and there is no need to do a spherical harmonic transform.

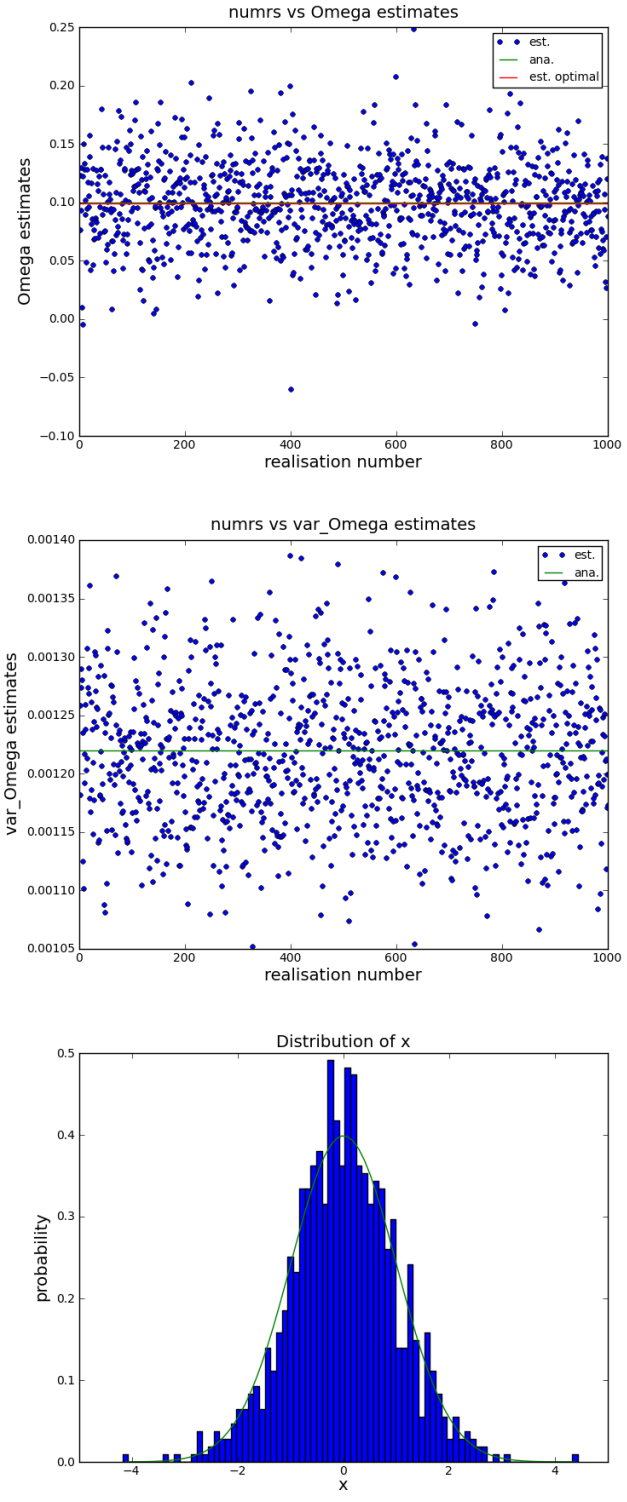

FIGURE 8.3: (*Top*) $\hat{\sigma^2_{h,i}}$ from 1000 realisations, with their sample mean, $\hat{\sigma^2_{h\,opt}}$ and σ^2_h plotted as lines. (*Middle*) Corresponding $\sigma^2_{\hat{\sigma^2_h},i}$'s for the 1000 realisations above, with the analytical σ^2 plotted as a line. (*Bottom*) Normalised histogram of $x_i = (\hat{\sigma^2_{h,opt}} - \hat{\sigma^2_{h,i}})/\sigma_{\hat{\sigma^2_h},i}$.

y_I and y_J are simulated this ways for $\tau = 1\ day$, for 365 days. The main analysis pipeline produces for each of these a dirty map, Fisher matrix and 'strong-signal bias matrix' according to (8.2), (8.3) and (8.4), respectively. By using these in in (5.41), the clean map and its variance, produced by the post-analysis pipeline, can be written as:

$$\hat{\mathcal{P}}_{00}(t) = \Gamma^{IJ\,-1}_{00,00}(t) X^{IJ}_{00}(t)\ , \tag{8.20}$$

$$\sigma^2_{\hat{\mathcal{P}}_{00}(t)} = \frac{1}{\Gamma^{IJ}_{00,00}(t)} \left(1 + \frac{\Psi^{IJ}_{00,00}(t)}{\Gamma^{IJ}_{00,00}(t)}\right)\ . \tag{8.21}$$

By replacing the spectral densities with variances in (8.2), (8.3) and (8.4), and then inserting them into (8.21), the analytical variance of $\hat{\mathcal{P}}_{00}$ is:

$$\sigma^2_{\hat{\mathcal{P}}_{00}(t)} = \left(\frac{1}{H_0 \gamma^{IJ}_{00} f_{Nyq}}\right)^2 \frac{1}{N} \left(\frac{f_{Nyq}}{(f_{high} - f_{low})} \frac{\overline{w_I^2 w_J^2}}{\overline{w_I w_J}^2} \left[(\sigma^2_{n_I} + \sigma^2_h)(\sigma^2_{n_J} + \sigma^2_h) + \sigma^4_h\right]\right)\ , \tag{8.22}$$

where f_{low} and f_{high} are the lower and upper limit of frequency integration in the analysis. Similar to the previous example, (8.16) and (8.17) can be used to combine the estimates in (8.20) and (8.21) to obtain optimal estimates.

Table. 8.1 summarises the optimal estimators and their standard deviations for 20 realisations using a particular set of simulation and analysis parameters. Note that the results are in terms of σ^2_h instead of $\mathcal{P}_{00}$ for comparison with the previous example. In this case, because, for 14 out of the 20 realisations (70%) simulated , $|\hat{\sigma^2_{h,opt}} - \sigma^2_h| < \sigma_{\hat{\sigma^2_h},opt}$ is true, and that $\sigma^2_{\hat{\sigma^2_h},opt} \approx \sigma^2_{\hat{\sigma^2_h},opt}$, the example is said to have passed the normalisation test.

8.4.3 Other examples

There are many changes that could be made to the above working example to continue the normalisation test further.

For example, one can stay with the same type of noise and signals, but change one of the parameters in the signal-simulation, such as the time cadence; this changes the Nyquist frequency, and f_{low} and f_{high} in the analysis may have to be changed accordingly.

One can also introduce stitching in the signal-simulation, where the time-series lasting τ is made of shorter time-series lasting T_{seg}. Then, different values of T_{seg} can be tested to see the effect of changing the 'amount of stitching' in the simulation.

One can also try a different relative strength of the white noise and signal by changing σ^2_n/σ^2_h. Alternatively, more complicated types of noise and signal can be tested, such as changing the uncorrelated white noise to uncorrelated *colour* noise with a power-law power spectral density. Or, similarly, the signal can be changed to a colour signal. In general, the power spectral

seed	$\hat{\sigma}^2_{h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt,weak}$	$\hat{\sigma}^2_{h,opt} - \sigma^2_h$
201	0.0998257892	0.0002701299	0.0002689355	0.0001742108
202	0.0996736733	0.0002700422	0.0002688474	0.0003263267
203	0.1000506850	0.0002701881	0.0002689939	-0.0000506850
204	0.0998555221	0.0002702081	0.0002690140	0.0001444779
205	0.1003125831	0.0002701552	0.0002689607	-0.0003125831
206	0.0999630010	0.0002700848	0.0002688900	0.0000369990
207	0.0999636798	0.0002701498	0.0002689554	0.0000363202
208	0.0997775613	0.0002701980	0.0002690039	0.0002224387
209	0.1001909316	0.0002701284	0.0002689339	-0.0001909316
210	0.0993610133	0.0002701803	0.0002689861	0.0006389867
211	0.0998917666	0.0002701552	0.0002689608	0.0001082334
212	0.1000347539	0.0002701515	0.0002689570	-0.0000347539
213	0.1000788316	0.0002701318	0.0002689373	-0.0000788316
214	0.1001796599	0.0002700643	0.0002688695	-0.0001796599
215	0.0999776719	0.0002701352	0.0002689407	0.0000223281
216	0.1001034189	0.0002701078	0.0002689131	-0.0001034189
217	0.1004492217	0.0002701441	0.0002689497	-0.0004492217
218	0.1000811479	0.0002701399	0.0002689454	-0.0000811479
219	0.0995958704	0.0002701554	0.0002689609	0.0004041296
220	0.0996692882	0.0002700643	0.0002688692	0.0003307118

TABLE 8.1: Normalisation test results. *Expected values*: $\sigma^2_h = 0.1$; $\sigma^2_n = 1.0$; $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0002751186$. *Simulation parameters*: $\Delta t = 0.5$ s; $H(f) = 1$, $\mathcal{P}_{00} = 1/90$ s. *Analysis parameters*: $f_{low} = 1.74 \times 10^{-4}$ Hz , $f_{high} = 4.998 \times 10^{-1}$ Hz.

densities that make up the covariance matrix used in simulation can be changed to achieve this. For signal-simulation, this can be done by changing $H(f)$, or $\gamma^{IJ}_\alpha(f,t)$, or both.

The number of combinations of changes are therefore numerous, and even though ideally only one change should be made at a time, there is really not a preferred order in which to make them. Table 8.2 summarises examples containing one or more of the changes mentioned here, which have been simulated, and the normalisation examined. On average 20 realisations are simulated and analysed for each example. The percentage of all realisations whose optimal estimate falls within one standard deviation is shown. The values of estimates for individual realisations are in a table to which there is a reference under 'results'. Three methods of simulation have been tested. In the 'time-domain' method (denoted by `td`), random numbers drawn from random number generators are taken directly to be samples of the time-series. In the 'frequency-domain-covariance-matrix' method (denoted by `fd`), the covariance matrix of the time-series is first obtained from analytical expressions of their PSDs and CSDs. In the 'frequency-domain-multipole-moments' method (denoted by `SpH`), the PSDs and CSDs of the signal are obtained by convolving $\mathcal{P}_{lm}$ with $\gamma^{IJ}_{lm}(f,t)$, then the covariance matrix is constructed from these. Under 'T_{seg}', if no stitching is used, there is the entry '-'. If stitching is used, the duration of short time-series stitched together is entered.

Uncorrelated white noise + same white signal
$\sigma_n^2 = 1.0 \quad \sigma_h^2 = 0.1 \quad \Delta t = 0.5\text{s} \quad \mathcal{P}_{00} = 1/90$

noise		signal		normalisation test		
method	T_{seg}	method	T_{seg}	%	results	notes
`td`	-	`td`	-	81%	Table H.1	
`fd`	-	`fd`	-	*pass*		
`fd`	-	`SpH`	-	69%	Table 8.1	
`td`	10^4 s	`td`	10^4 s	60%	Table H.2	*Bug*: incorrect random seed parsing.
`td`	-	`SpH`	10^4 s	69%	Table H.3	
`fd`	10^4 s	`SpH`	10^4 s	69%	Table H.4	

Uncorrelated white noise + same white signal
$\sigma_n^2 = 1.0 \quad \sigma_h^2 = 0.1 \quad \Delta t = 8.0\text{s} \quad \mathcal{P}_{00} = 1/90$

noise		signal		normalisation test		
method	T_{seg}	method	T_{seg}	%	results	notes
`fd`	-	`SpH`	-	*pass*		
`fd`	1.6×10^4 *s*	`SpH`	1.6×10^4 *s*	73%	Table H.5	*Bug*: wrong window definition.

Scaled LISA's A and E noise + same white signal
$\Delta t = 8.0\text{s} \quad \mathcal{P}_{00} = 1/90$

noise		signal		normalisation test		
method	T_{seg}	method	T_{seg}	%	results	notes
`fd`	1.6×10^4 *s*	`SpH`	1.6×10^4 *s*	68%	Table H.6	
`fd`	4×10^3 *s*	`SpH`	1.6×10^4 *s*	64%	Table H.7	
`fd`	4×10^3 *s*	`SpH`	4×10^3 *s*	48%	Table H.8	

Scaled LISA's A and E noise + same colour (power-law) signal
$\Delta t = 8.0\text{s} \quad \mathcal{P}_{00} = 1/90$

noise		signal		normalisation test		
method	T_{seg}	method	T_{seg}	%	results	notes
`fd`	1.6×10^4 *s*	`SpH`	1.6×10^4 *s*	68%	Table H.9	

Scaled LISA's A and E noise + same colour ($y_A^{GW}(t)$) signal
$\Delta t = 8.0\text{s} \quad \mathcal{P}_{00} = 1/90$

noise		signal		normalisation test		
method	T_{seg}	method	T_{seg}	%	results	notes
`fd`	1.6×10^4 *s*	`SpH`	1.6×10^4 *s*	70%	Table H.10	$\gamma_{sim}^{IJ} = \gamma_{ana}^{IJ}$

Scaled LISA's A and E noise + different white signal
$\Delta t = 8.0\text{s} \quad \mathcal{P}_{00} = 1/90$

noise		signal		normalisation test		
method	T_{seg}	method	T_{seg}	%	results	notes
`fd`	1.6×10^4 *s*	`SpH`	1.6×10^4 *s*	68%	Table H.11	

TABLE 8.2: Case for which normalisation test are carried out. For each case, time-series $y_I(t) = n_I(t) + h_I(t)$ and $y_J(t) = n_J(t) + h_J(t)$ are simulated and analysed. The widest boxes in the table give description of the noise and signal used in those cases examples below it. For each case, the method of simulation for the noise and the signal are either *td* (for time-domain), *fd* (for frequency-domain-covariance-matrix) or *SpH* (for frequency-domain-multipole-moments). A T_{seg} entry not equal to - indicates the duration of the stitched time-series, else stitching is not used and $T_{seg} = 1$ day $= 86400$ s. % indicates the percentage of reailisations for which the estimated $\mathcal{P}_{00}$ or σ_h^2 fall within one standard deviation of the true value. *results* provides reference to a table showing estimates from all realisations.

In the process of testing these examples, several bugs in the implementation are found and corrected. In addition, the following observations which have not been understood can perhaps offer clues for further test examples:

1. The sample mean of a white time-series (signal or noise) is different for the 'frequency-domain-covariance-matrix' method. When it is obtained by this method, the sample mean is $\sim 10^{-18}$. When the other two methods are used, it is $\sim 10^{-4}$. It is not clear how this effects the results of the normalisation test. However, if n_I is simulated with this method, while the signal h_I is simulated with another, then the sample mean of the output (signal + noise) is essentially that of the signal.

2. In Table 8.2, three examples have been carried out for '*Scaled LISA's A and E noise + same white signal*', with different T_{seg} in stitching. The smaller the T_{seg} the more stitching there is, since more time-series are needed to make up the total duration of τ. Hence, these three examples show that the noramlisation deteriorates when there is more stitching. In fact, only the first example has a satisfactory normalisation. This suggests that there is still something not right with the part of the implementation that does the stitching.

In summary, the spherical harmonic decomposition algorithm has been implemented to estimate the anisotropy of a stochastic background in the context of LISA. By using the optimal TDI observables of the second generation Michelsons, it is seen that the implementation can produce estimates of the multipole moments of a stochastic background's anisotropy. When projected, these multipole moments display a spatial pattern that resembles that of the injected source, in the case where the source contains a point source, or an isotropic background. At a glance, the normalisation of the estimated backgrounds agrees reasonably well with the injected background. However, a more detailed examination shows that it is not satisfactorily close. The issue remain to be resolved, and there is evidence that the cause is in the data-simulation pipeline, not the analysis pipeline. A systematic way to find the cause is described and carried out, with the latest status and clues summarised at the end.

Appendix A

Derivation of the two-pulse GW response

Using a transverse-traceless gauge, consider the situation in which a plane graviational wave propagates in the $-z$ direction, incident on a photon travelling in the x-direction in the $x-y$ plane. In this case, the components of $h_{\mu\nu}$ is given by (2.5), and, since $dy = dz = z = 0$, along the photon's path:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + [1 + h_+(t)]\, dx^2 \ . \tag{A.1}$$

Since the photon's path is a null path, $ds^2 = 0$. By inserting this into (A.1), it can be shown that

$$dt \approx dx \left[1 + \frac{1}{2}h_+(t)\right] \ , \tag{A.2}$$

where all terms of second order or higher in $h_{\mu\nu}$ are ignored ($|h_{\mu\nu}| \ll 1$). The time it takes the photon to travel between two events (t_1, x_1) and (t_2, x_2) can be found by integrating (A.2):

$$\int_{t_1}^{t_2} dt \approx \int_{x_1}^{x_2} dx \left[1 + \frac{1}{2}h_+(t)\right] \ . \tag{A.3}$$

However, before doing this, a curve to the photon's path must first be defined, in order to express the t in the RHS in terms of x. For this purpose, it is assumed here the photon's path is of zeroth order in $h_{\mu\nu}$, unperturbed in the absence of gravitational waves. In this case, the time it takes for the photon to travel between the events is $t_2 - t_1 = x_2 - x_1 = L$, and the following curve parameterises the path with λ:

$$t = \lambda \tag{A.4}$$

$$x(t) = t - (t_2 - L) \tag{A.5}$$

$$y(t) = 0 \tag{A.6}$$

$$z(t) = 0 \ . \tag{A.7}$$

By substituting (A.5) in (A.3), the total time of travel becomes

$$T(t) = \int_{t_1}^{t} dt' \approx L + \frac{1}{2}\int_0^L dx\; h_+(t - L + x)\;, \tag{A.8}$$

where t_2 has been replaced with t. (A.8) gives the time it takes for the photon to travel between (t_1, x_1) and (t, x_2) in the presence of gravitational waves. The second term in the RHS is the change of this time due to the gravitational wave. If the GW response is defined as the fractioanl change in $T(t)$, its Fourier transform can be shown to be

$$\frac{\tilde{\Delta T}}{L}(f) = \frac{1}{2}\text{sinc}\left(\frac{1}{2}\frac{f}{f_*}\right)\; e^{-i\frac{1}{2}\frac{f}{f_*}}\; \tilde{h}_+(f)\;, \tag{A.9}$$

where $f_* = 1/(2\pi L)$ is the transfer frequency. If the fractional change in the photon frequency is defined as the response instead, it can be found from the fractional change in time by differentiating with respect to time. In this case, the Fourier transform is:

$$\frac{\tilde{\Delta\nu}}{\nu_0}(f) = i\frac{1}{2}\frac{f}{f_*}\text{sinc}\left(\frac{1}{2}\frac{f}{f_*}\right)\; e^{-i\frac{1}{2}\frac{f}{f_*}}\; \tilde{h}_+(f)\;. \tag{A.10}$$

Comparing with (3.8), it can be seen that (A.10) the special case where the direction of wave propagation is perpendicular to the photon's path.

In general, for a GW of frequency f and direction of propagation $-\hat{k}$,

$$h_{ij}(t, \bar{x}) = H_{ij}e^{i2\pi f\left[t + \hat{k}\cdot\bar{x}\right]}\;, \tag{A.11}$$

the derivation of the two pulse response can be done by consdering the geodesic equations of the photon:

$$\frac{dp^\alpha}{d\lambda} + \Gamma^\alpha_{\mu\nu}p^\mu p^\nu = 0\;, \tag{A.12}$$

where $\vec{p}$ is the 4-momentum of the photon. Since it is tangent to the photon's path, if $x^\mu(\lambda)$ is a curve to the path:

$$\vec{p} \to p^\mu = \frac{dx^\mu}{d\lambda} = (h\nu, \bar{p})\;, \tag{A.13}$$

where ν is the photon frequency, and h is Planck's constant. For the photon path unperturbed by gravitational waves, the 4-momentum is written as

$$\vec{p_0} \to h\nu_0(1, \hat{n}) \tag{A.14}$$

where $\hat{n}$ is a unit vector in the direction in which the photon travels. If the parameterisation is changed from λ to $h\nu_0\lambda$, the t-component of a 4-momentum vector $\vec{p}$ now becomes the fractional photon frequency (instead of the photon energy), ν/ν_0. Then, the geodesic equation

(A.12) with $\alpha = t$ describes how this changes along the path of the photon:

$$\frac{d}{d\lambda}\left(\frac{\nu}{\nu_0}\right) = -\Gamma^t_{\mu\nu} p^\mu p^\nu \ . \tag{A.15}$$

Using (2.1) and (2.4) in the definition of Christoffel symbols,

$$\Gamma^t_{ij} = \frac{1}{2}\frac{\partial h_{ij}}{\partial t} + O(h^2) \ . \tag{A.16}$$

It can also be shown that

$$\frac{dh_{ij}}{d\lambda} = \frac{dx^\mu}{d\lambda}\frac{\partial h_{ij}}{\partial x^\mu} = \left(\frac{dt}{d\lambda} + k_k \frac{x^k}{d\lambda}\right)\frac{\partial h_{ij}}{\partial t} = \left(1 + \hat{k}\cdot\hat{n}\right)\frac{\partial h_{ij}}{\partial t} + O(h^2) \ , \tag{A.17}$$

where the fact that λ has been scaled by $h\nu_0$ and (A.14) are used. Substituting (A.16) and (A.17) in (A.15),

$$\frac{d}{d\lambda}\left(\frac{\nu}{\nu_0}\right) = -\frac{1}{2}\frac{n^i n^j}{1 + \hat{k}\cdot\hat{n}}\frac{dh_{ij}}{d\lambda} + O(h^2) \ , \tag{A.18}$$

where p^i has been replaced by n^i, because, as can be seen in (A.16), the lowest order term in Γ^t_{ij} is already $O(h)$. The total change in fractional frequency of the photon between two events λ_1 and λ_2 can now be found by integrating (A.18), giving

$$\frac{\nu(\lambda_1) - \nu(\lambda_2)}{\nu_0} = -\frac{1}{2}\frac{n^i n^j}{1 + \hat{k}\cdot\hat{n}}\left[h_{ij}(\lambda_2) - h_{ij}(\lambda_1)\right] \ . \tag{A.19}$$

This is the two pulse GW response defined in (3.1), and the events at λ_1 and λ_2 can be intepreted as the emission and reception of a photon transmission between two LISA spacecrafts.

Appendix B

Cross-correlation of short-term Fourier transforms

B.1 Signal Correlation

Using the definition of short-term Fourier transform in (5.9) and the response to a stochastic gravitational wave background in (5.8), it can be shown that

$$\tilde{y}_I^{GW}(f,t) = \int_{-\infty}^{+\infty} df'\ e^{i2\pi(f'-f)t}\delta_\tau(f-f') \int_{S^2} d\hat{k}\ F_I^A(f',\hat{k};t)e^{i2\pi f'\hat{k}\cdot\vec{p}_I(t)}\mathcal{H}_A(f',\hat{k}) \tag{B.1}$$

The data, which we will take the short-term Fourier transform of, has a duration of τ. Because this is chosen to be much larger than the correlation time between any pair of detectors [1], we can assume that if t and t' label two different stretches of gravitational response of duration τ, then the two stretches are uncorrelated; this is indicated with the Kronecker delta $\delta_{tt'}$. Using

[1] Not sure if detectors in this instance translate directly to optimal TDI observables

(B.1),

$$\begin{aligned}
&\langle \tilde{y}_I^{GW*}(f,t)\tilde{y}_J^{GW}(f',t')\rangle \\
&= \int_{-\infty}^{+\infty} df_1\, e^{-i2\pi(f_1-f)t}\delta_\tau(f-f_1)\int_{-\infty}^{+\infty} df_2\, e^{i2\pi(f_2-f')t'}\delta_\tau(f'-f_2)\cdot \\
&\int_{S^2} d\hat{k}' \int_{S^2} d\hat{k}\, F_I^{A*}(f_1,\hat{k};t)F_J^{A'}(f_2,\hat{k};t')\, e^{-i2\pi f_1\hat{k}\cdot\vec{p}_I(t)}e^{i2\pi f_2\hat{k}'\cdot\vec{p}_I(t')}\langle \mathcal{H}_A^*(f_1,\hat{k})\mathcal{H}_{A'}(f_2,\hat{k}')\rangle
\end{aligned} \tag{B.2}$$

$$\begin{aligned}
&= \delta_{tt'}e^{i2\pi(f-f')t}\int_{-\infty}^{+\infty} df_1\delta_\tau(f-f_1)\delta_\tau(f'-f_1)\cdot \\
&\int_{S^2} d\hat{k}\, F_I^{A*}(f_1,\hat{k};t)F_J^{A}(f_1,\hat{k};t)e^{-i2\pi f_1\hat{k}\cdot(\vec{p}_I(t)-\vec{p}_J(t))}\cdot\frac{\mathcal{P}(f_1,\hat{k})}{4}
\end{aligned} \tag{B.3}$$

$$\begin{aligned}
&= \delta_{tt'}e^{i2\pi(f-f')t}\int_{-\infty}^{+\infty} df_1\delta(f-f_1)\delta_\tau(f'-f_1)\cdot \\
&\int_{S^2} d\hat{k}\, F_I^{A*}(f_1,\hat{k};t)F_J^{A}(f_1,\hat{k};t)e^{-i2\pi f_1\hat{k}\cdot(\vec{p}_I(t)-\vec{p}_J(t))}\cdot\frac{\mathcal{P}(f_1,\hat{k})}{4}
\end{aligned} \tag{B.4}$$

$$= \delta_{tt'}e^{i2\pi(f-f')t}\delta_\tau(f-f')\int_{S^2} d\hat{k}\, F_I^{A*}(f,\hat{k};t)F_J^{A}(f,\hat{k};t)e^{-i2\pi f\hat{k}\cdot(\vec{p}_I(t)-\vec{p}_J(t))}\frac{\mathcal{P}(f,\hat{k})}{4} \tag{B.5}$$

In (B.3), the expectation value is substituted with (??). And we take $t = t'$ for the cross-correlation to be non-zero. In (B.4), one of the finite-time dirac-delta functions δ_τ is replaced by a dirac-delta function δ. This is a good approximation because the power spectrum $\mathcal{P}(f,\hat{k})$ varies relatively slowly to δ_τ in f. Finally, for discrete frequencies, we make the further approximation

$$\delta(f-f') \approx \tau\delta_{ff'} \tag{B.6}$$

Note that the RHS has the same value as the LHS at $f = f'$, and it also has the same dimension of $[time]$. By substituting (B.6) in (B.5),

$$\langle \tilde{y}_I^{GW*}(f,t)\tilde{y}_J^{GW}(f',t')\rangle = \frac{\tau}{2}\delta_{tt'}\delta_{ff'}\int_{S^2} d\hat{k}\, \gamma_{IJ}(f,\hat{k};t)\mathcal{P}(f,\hat{k}) \tag{B.7}$$

, where

$$\gamma_{IJ}(f,\hat{k},t) = \frac{1}{2}F_I^{A*}(f,\hat{k},t)F_J^{A}(f,\hat{k},t)e^{-i2\pi f\hat{k}\cdot(\vec{p}_I(t)-\vec{p}_J(t))} \tag{B.8}$$

Using (??), (??) and (??),

$$\langle \tilde{y}_I^{GW*}(f,t)\tilde{y}_J^{GW}(f',t')\rangle = \frac{\tau}{2}\delta_{tt'}\delta_{ff'}H(f)\,\gamma_\alpha^{IJ}(f,t)\mathcal{P}_\alpha \tag{B.9}$$

Using (B.9), we find that

$$\begin{aligned}
&\langle \tilde{y}_I^{GW}(f,t)\tilde{y}_J^{GW}(f',t')\rangle \\
=&\langle \tilde{y}_I^{GW*}(-f,t)\tilde{y}_J^{GW}(f',t')\rangle \\
=&\frac{\tau}{2}\,\delta_{tt'}\,\delta_{-ff'}H(-f)\sum_{l,m}\gamma_{lm}^{IJ}(-f,t)\mathcal{P}_{lm} \\
=&\frac{\tau}{2}\,\delta_{tt'}\,\delta_{-ff'}H(f)\sum_{l,m}(-1)^m\gamma_{l-m}^{IJ*}(f,t)(-1)^m\mathcal{P}_{l-m}^* \qquad \text{(B.10)}\\
=&\frac{\tau}{2}\,\delta_{tt'}\,\delta_{f-f'}H(f)\gamma_{\alpha}^{IJ*}(f,t)\mathcal{P}_{\alpha}^* \qquad \text{(B.11)}
\end{aligned}$$

In (B.10), we have used the identities (E.5), $\mathcal{P}_{l-m} = (-1)^m\mathcal{P}_{lm}^*$, since $\mathcal{P}(\hat{k})$ is real, and the fact that $H(-f) = H(f)$.

B.2 Noise Correlation

The noise is assumed to be stationary over a segment of duration τ as given in (**??**). This means that

$$\langle y_I^{noise}(t_1)y_J^{noise}(t_2)\rangle = \delta_{IJ}\,C_I^{noise}(t_1-t_2,t) \qquad \text{(B.12)}$$

, where $t_1, t_2 \in [t-\tau/2, t+\tau/2]$. It will also be assumed that noise from different such segments are uncorrelated, implying

$$\langle \tilde{y}_I^{noise*}(f,t)\tilde{y}_J^{noise}(f',t')\rangle = \delta_{tt'}\int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}}dt_1\int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}}dt_2\,e^{i2\pi ft_1}e^{-i2\pi f't_2}\langle y_I^{noise}(t_1)y_J^{noise}(t_2)\rangle \qquad \text{(B.13)}$$

Subsituting (B.12) in (B.13), and writing the auto-correlation function, $C_I^{noise}(t_1-t_2)$, as the inverse Fourier transform of the power spectral density, $P_I^{nosie}(f,t)$, we obtain

$$\begin{aligned}
&\langle \tilde{y}_I^{noise*}(f,t)\tilde{y}_J^{noise}(f',t')\rangle \\
=&\frac{1}{2}\delta_{IJ}\delta_{tt'}\int_{-\infty}^{+\infty}df_1\int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}}dt_1\,e^{i2\pi(f-f_1)t_1}\int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}}dt_2\,e^{i2\pi(f_1-f')t_2}P_I^{noise}(f_1,t) \\
=&\frac{1}{2}\,\delta_{IJ}\,\delta_{tt'}\,e^{i2\pi(f-f')t}\int_{-\infty}^{+\infty}df_1\,\delta_\tau(f-f_1)\delta_\tau(f'-f_1)P_I^{noise}(f_1,t) \\
=&\frac{1}{2}\,\delta_{IJ}\,\delta_{tt'}\,e^{i2\pi(f-f')t}\int_{-\infty}^{+\infty}df_1\,\delta(f-f_1)\delta_\tau(f'-f_1)P_I^{noise}(f_1,t) \qquad \text{(B.14)}\\
=&\frac{1}{2}\,\delta_{IJ}\,\delta_{tt'}\,e^{i2\pi(f-f')t}\,\delta_\tau(f-f')\,P_I^{noise}(f,t) \\
=&\frac{\tau}{2}\,\delta_{IJ}\,\delta_{tt'}\,\delta_{ff'}\,P_I^{noise}(f,t) \qquad \text{(B.15)}
\end{aligned}$$

We replace one of the finite Dirac delta functions with a a real Dirac delta function in (B.14). Using the approximation in (B.6), the discretised-frequency version is obtained in (B.15).

Again, using (B.15),

$$\begin{aligned}
&\langle \tilde{y}_I^{noise}(f,t)\tilde{y}_J^{noise}(f',t')\rangle \\
=&\langle \tilde{y}_I^{noise*}(-f,t)\tilde{y}_J^{noise}(f',t')\rangle \\
=&\frac{\tau}{2}\,\delta_{IJ}\,\delta_{tt'}\,\delta_{f-f'}\,P_I^{noise}(f,t)
\end{aligned} \tag{B.16}$$

B.3 Output Correlation

Using (B.9), (B.11), (B.15) and (B.16), the cross-correlations of detector outputs can be summerised as follows.

$$\langle \tilde{y}_I^*(f,t)\tilde{y}_J(f',t')\rangle = \frac{\tau}{2}\,\delta_{tt'}\,\delta_{ff'}\,\left(\delta_{IJ}\,P_I^{noise}(f,t) + H(f)\gamma_\alpha^{IJ}(f,t)\mathcal{P}_\alpha\right) \tag{B.17}$$

$$\langle \tilde{y}_I(f,t)\tilde{y}_J(f',t')\rangle = \frac{\tau}{2}\,\delta_{tt'}\,\delta_{f-f'}\,\left(\delta_{IJ}P_I^{noise}(f,t) + H(f)\gamma_\alpha^{IJ*}(f,t)\mathcal{P}_\alpha^*\right) \tag{B.18}$$

Appendix C

Covariance Matrix

Using the definition of the cross-correlated output in (5.10),

$$
\begin{aligned}
&\langle C_{IJ}(f,t)C^*_{IJ}(f',t')\rangle \\
=&\frac{4}{\tau^2}\langle \tilde{y}^*_I(f,t)\tilde{y}_J(f,t)\tilde{y}_I(f',t')\tilde{y}^*_J(f',t')\rangle \\
=&\frac{4}{\tau^2}(\ \langle \tilde{y}^*_I(f,t)\tilde{y}_J(f,t)\rangle\langle \tilde{y}_I(f',t')\tilde{y}^*_J(f',t')\rangle + \langle \tilde{y}^*_I(f,t)\tilde{y}_J(f',t')\rangle\langle \tilde{y}_J(f,t)\tilde{y}^*_J(f',t')\rangle \\
&+ \langle \tilde{y}^*_I(f,t)\tilde{y}^*_J(f',t')\rangle\langle \tilde{y}_I(f',t')\tilde{y}_J(f,t)\rangle\ ) \qquad \text{(C.1)} \\
=&\langle C_{IJ}(f,t)\rangle\langle C^*_{IJ}(f',t')\rangle \\
&+ \langle \frac{2}{\tau}\tilde{y}^*_I(f,t)\tilde{y}_I(f',t')\rangle\langle \frac{2}{\tau}\tilde{y}^*_J(f,t)\tilde{y}_J(f',t')\rangle^* + \langle \frac{2}{\tau}\tilde{y}_I(f,t)\tilde{y}_J(f',t')\rangle^*\langle \frac{2}{\tau}\tilde{y}_I(f',t')\tilde{y}_J(f,t)\rangle \qquad \text{(C.2)}
\end{aligned}
$$

In (C.1), the product rule for expectation values is used; in (C.2), (5.10) is used again.

Now using (**??**), (B.17), and (B.18), elements of the covariance matrix are given by

$$
\mathbb{C}_{ft,f't'} = \delta_{tt'}\delta_{ff'}P_I(f,t)P_J(f,t) + \delta_{tt'}\delta_{f-f'}(\delta_{IJ}P^{noise}_I(f,t) + H(f)\gamma^{IJ}_\alpha(f,t)\mathcal{P}_\alpha)^2 \qquad \text{(C.3)}
$$

In obtaining (C.3), we have used

$$
\sum_{l,m}\gamma^{JI*}_{lm}(f,t)\mathcal{P}^*_{lm} = \sum_{l,m}\gamma^{IJ}_{lm}(f,t)\mathcal{P}_{lm} \qquad \text{(C.4)}
$$

, which can be shown using (E.7), and the reality of $\mathcal{P}(\hat{k})$. $P_I(f,t)$ is the power spectral density of the detector output during the τ-segment at time t:

$$
\mathrm{P}^{noise}_I(f,t) + H(f)\gamma^{II}_\alpha(f,t)\mathcal{P}_\alpha \qquad \text{(C.5)}
$$

Appendix D

Covariances of Estimators

By using the definition of the dirty map in (5.31),

$$\langle X_\alpha^{IJ} X_\beta^{IJ*} \rangle = \sum_{t,f} \sum_{t',f'} \gamma_\alpha^{IJ*}(f,t) \gamma_\beta^{IJ}(f',t') \frac{H(f)H(f')}{P_I(f,t)P_J(f,t)P_I(f',t')P_J(f',t')} \mathbb{C}_{ft,f't'}^{IJ} \tag{D.1}$$

Inserting (C.3) into (D.1) , and taking $I \neq J$, the covariance of the dirty map estimate X_α^{IJ} can be expressed as

$$\langle X_\alpha^{IJ} X_\beta^{IJ*} \rangle - \langle X_\alpha^{IJ} \rangle \langle X_\beta^{IJ*} \rangle = \Gamma_{\alpha\beta}^{IJ} + \Psi_{\alpha\beta}^{IJ} \, , \tag{D.2}$$

where

$$\Psi_{\alpha\beta}^{IJ} = \sum_t \sum_f \gamma_\alpha^{IJ*}(f,t) \frac{H^2(f) P_{IJ}^{GW^2}(f,t)}{P_I^2(f,t) P_J^2(f,t)} \gamma_\beta^{IJ}(-f,t) \, . \tag{D.3}$$

Similiarly, by using the definition of the a regularised clean map estimate, its covariance is

$$\begin{aligned} \langle \hat{\mathcal{P}}_\alpha^{IJ} \hat{\mathcal{P}}_\beta^{IJ*} \rangle - \langle \hat{\mathcal{P}}_\alpha^{IJ} \rangle \langle \hat{\mathcal{P}}_\beta^{IJ*} \rangle &= \Gamma'^{IJ-1}_{\alpha\gamma} \Gamma'^{IJ-1*}_{\beta\delta} \left(\langle X_\gamma^{IJ} X_\delta^{IJ*} \rangle - \langle X_\gamma^{IJ} \rangle \langle X_\delta^{IJ*} \rangle \right) \\ &= \Gamma'^{IJ-1}_{\alpha\gamma} \left(\Gamma_{\alpha\beta}^{IJ} + \Psi_{\alpha\beta}^{IJ} \right) \Gamma'^{IJ-1*}_{\beta\delta} \end{aligned} \tag{D.4}$$

Ψ_α^{IJ} defined in (D.3) can be written as

$$\begin{aligned} \Psi_{\alpha\beta}^{IJ} = \sum_t \sum_f \gamma_\alpha^{IJ*}(f,t) \frac{H^2(f)}{P_I(f,t)P_J(f,t)} \gamma_\beta^{IJ}(-f,t) \times \\ \frac{H(f)\gamma_\gamma^{IJ}(f,t)\mathcal{P}_\gamma}{P_I^{noise}(f,t) + H(f)\gamma_\mu^{II}(f,t)\mathcal{P}_\mu} \frac{H(f)\gamma_\delta^{IJ}(f,t)\mathcal{P}_\delta}{P_J^{noise}(f,t) + H(f)\gamma_\nu^{JJ}(f,t)\mathcal{P}_\nu} \, . \end{aligned} \tag{D.5}$$

When the GW signal is much weaker than the noise, the last two terms in this expression become

$$\begin{aligned}\frac{H(f)\gamma_\gamma^{IJ}(f,t)\mathcal{P}_\gamma}{P_I^{noise}(f,t)+H(f)\gamma_\mu^{II}(f,t)\mathcal{P}_\mu} &\approx \frac{H(f)\gamma_\nu^{IJ}(f,t)\mathcal{P}_\nu}{\mathrm{P}_I^{noise}(f,t)} \ll 1\\ \frac{H(f)\gamma_\delta^{IJ}(f,t)\mathcal{P}_\delta}{P_J^{noise}(f,t)+H(f)\gamma_\nu^{JJ}(f,t)\mathcal{P}_\nu} &\approx \frac{H(f)\gamma_\delta^{IJ}(f,t)\mathcal{P}_\delta}{P_J^{noise}(f,t)} \ll 1\,.\end{aligned} \tag{D.6}$$

This means that, in the *weak-signal limit*, $\Psi_{\alpha\beta}^{IJ}$ can effectively be ignored in the covariances of $\hat{\mathcal{P}}_\alpha$ and $X_{\alpha\beta}^{IJ}$[1]. It can be seen from (D.4) that the covariance of the unregularised clean map is simply the inverse of the unregularised Fisher matrix.

In general, in the *strong-signal limit*, $\Psi_{\alpha\beta}^{IJ}$ can not be ignored. In this case, (D.6) becomes

$$\begin{aligned}\frac{H(f)\gamma_\gamma^{IJ}(f,t)\mathcal{P}_\gamma}{\mathrm{P}_I^{noise}(f,t)+H(f)\gamma_\mu^{II}(f,t)\mathcal{P}_\mu} &\approx \frac{H(f)\gamma_\gamma^{IJ}(f,t)\mathcal{P}_\gamma}{H(f)\gamma_\mu^{II}(f,t)\mathcal{P}_\mu}\\ \frac{H(f)\gamma_\delta^{IJ}(f,t)\mathcal{P}_\delta}{\mathrm{P}_J^{noise}(f,t)+H(f)\gamma_\nu^{JJ}(f,t)\mathcal{P}_\nu} &\approx \frac{H(f)\gamma_\delta^{IJ}(f,t)\mathcal{P}_\delta}{H(f)\gamma_\nu^{JJ}(f,t)\mathcal{P}_\nu}\,.\end{aligned} \tag{D.7}$$

If the overlap reduction between two different observables is much smaller than that between two same observables[2],

$$|\gamma_\alpha^{IJ}(f,t)| \ll |\gamma_\alpha^{II}(f,t)| \approx |\gamma_\alpha^{JJ}(f,t)|\,. \tag{D.8}$$

Substituting this into (D.7), it can be seen that $\Psi_{\alpha\beta}^{IJ}$ also tends to zero, and the covariances are the same as those in the weak-signal limit.

[1] Because of this $\Psi_{\alpha\beta}^{IJ}$ is often referred to as the strong-signal bias matrix for the covariances.

[2] In general, this is only true for detectors at different locations. For LIGO H1 and H2, which are at the same location, the term additional to $\Gamma_{\alpha\beta}^{IJ}$ can cause a difference of about 30%.

Appendix E

Overlap Reduction Functions: Identities

E.1 Negative Frequencies: $f \to -f$

From (**??**), it can be shown that, for the response function of a link observable,

$$F_l^A(-f, \hat{k}, \psi) = F_l^{A*}(f, \hat{k}, \psi) \tag{E.1}$$

This implies that if the frequency is inverted for the response function of a TDI observable as given in (**??**),

$$F_X^A(-f, \hat{k}, \psi; t_X, \vec{p}_X) = F_X^{A*}(f, \hat{k}, \psi; t_X, \vec{p}_X) \tag{E.2}$$

By looking at (**??**), it is clear that the response function of an optimal TDI observable is simply proportional to the response functions of the TDI observables which it consists of. So, for an optimal TDI observable I,

$$F_I^A(-f, \hat{k}, \psi; t_I, \vec{p}_I) = F_I^{A*}(f, \hat{k}, \psi; t_I, \vec{p}_I) \tag{E.3}$$

Using this identity, and (**??**), the following identity can be obtained for the overlap reduction function of an optiaml TDI observable

$$\gamma_{IJ}(-f, \hat{k}, t) = \gamma_{IJ}^*(f, \hat{k}, t) \tag{E.4}$$

Because the spherical harmonics obey

$$Y_{l-m} = (-1)^m Y_{lm}^*$$

, (E.4) implies that the multipole moments of the overlap reduction function obey

$$\gamma_{lm}^{IJ}(-f,t) = (-1)^m \, \gamma_{l-m}^{IJ*}(f,t) \tag{E.5}$$

E.2 Reversing order of detectors: $IJ \to JI$

It can be shown from the definition of the overlap reduction function given in (**??**) that

$$\gamma_{JI}(f,\hat{k},t) = \gamma_{IJ}^{*}(f,\hat{k},t) \tag{E.6}$$

Similar to (E.5), this implies

$$\gamma_{lm}^{JI}(f,t) = (-1)^m \gamma_{l-m}^{IJ*}(f,t) \tag{E.7}$$

Appendix F

Time-closure for G2 TDI Observables

In a non-static model of LISA, for which second generation TDI observables are required, the light travel times along the links vary with time. For the purpose of time-keeping along the virtual beam in such observables, it is useful to define two types of time intervals[1] :

1. $L_l(t)$ is the time it has taken for light to travel down the link slr to arrive at time t

2. $\Gamma_l(t)$ is the time it takes for light to travel down the link slr having been emitted at time t

For second generation TDI observables, the aim is to have time-closure up to first order/degree in $\dot{L}_l$. This actually means that all $\dot{L}^2$ and $\ddot{L}$ terms are ignored. Using a simple example string, it will be shown here how the time-lapse accumulated along a virtual beam can be approximated up to first order/degree in $\dot{L}_l$. It will then be seen how these first order/degree terms can be read, or counted, straight off the string. Then, this method will be applied to the string of a $G2$ TDI observable to show that it is indeed $\dot{L}-$closed. The simple example string is that of $X0$, given in (3.21):

$$X0 \ : \ \overrightarrow{3'3}\,\overleftarrow{2'2} \ . \tag{F.1}$$

It is not time-closed up to first degree/order in $L_l(t)$, but being a short string, it is useful for demonstrating the points mentioned above quickly. For the virtual beam which travels from

[1] These two time intervals are defined because, in going along the virtual beam pf a TDI observable, one either goes past a link string that is in the same direction as the beam, in which case a time interval into the future $\Gamma_l(t_s)$ is accumulated, or one goes past a link string that is in the opposite direction, in which case a time interval into the past $-L_l(t_r)$ is accumulated. Note it only makes sense that Γ_l acts on an emission time, and L_l acts on a reception time.

the leftmost point to the rightmost point of the string, if the time stamp at the leftmost point is t, then the time stamp at the rightmost point, t', with the repetitive use of $L_l(t)$ and $\Gamma_l(t)$, is given by

$$\begin{aligned} t' = t \; + \; \Gamma_{3'}(t) \; + \; \Gamma_3[t+\Gamma_{3'}(t)] \; - \; L_{2'}[t+\Gamma_{3'}(t)+\Gamma_3[t+\Gamma_{3'}(t)]] \\ - L_2[t+\Gamma_{3'}(t)+\Gamma_3[t+\Gamma_{3'}(t)] - L_{2'}[t+\Gamma_{3'}(t)+\Gamma_3[t+\Gamma_{3'}(t)]]] \end{aligned} \tag{F.2}$$

If all Γ_l and L_l terms are Taylor-expanded up to first order/degree about t, it becomes

$$\begin{aligned} t' &\approx t + \Gamma_{3'}(t) + \Gamma_3(t) + \dot{\Gamma}_3(t)\Gamma_{3'}(t) \\ &\quad - \Big(L_{2'}(t) + \dot{L}_{2'}(t)[\Gamma_{3'}(t)+\Gamma_3(t)]\Big) - \Big(L_2(t) + \dot{L}_2(t)[\Gamma_{3'}(t)+\Gamma_3(t) - L_{2'}(t)]\Big) \qquad \text{(F.3)} \\ &\approx t + \Gamma_{3'}(t)\Big(1+\dot{\Gamma}_3(t) - \dot{L}_{2'}(t) - \dot{L}_2(t)\Big) + \Gamma_3(t)\Big(1 - \dot{L}_{2'}(t) - \dot{L}_2(t)\Big) \\ &\quad -L_{2'}(t)\Big(1-\dot{L}_2(t)\Big) - L_2(t)\,. \qquad \text{(F.4)} \end{aligned}$$

In (F.4), the first derivatives of all links to the right of each link are grouped together. By their definitions,

$$\Gamma_l(t) = L_l(t+\Gamma_l(t))\,. \tag{F.5}$$

By Taylor-expanding (F.3) recursively about t, and keeping terms only up to first order/degree in $\dot{L}_l(t)$ again,

$$\Gamma_l(t) \approx L_l(t) + \dot{L}_l(t)\Gamma_l(t) \approx L_l(t) + \dot{L}_l(t)L_l(t) \tag{F.6}$$

$$\dot{\Gamma}_l(t) \approx \dot{L}_l(t)\,, \tag{F.7}$$

Substituting (F.6) and (F.7) in (F.4),

$$\begin{aligned} t' = t + L_{3'}(t) + L_{3'}(t)\Big(\dot{L}_{3'}(t) + \dot{L}_3(t) - \dot{L}_{2'}(t) - \dot{L}_2(t)\Big) + L_3(t) + L_3(t)\Big(\dot{L}_3(t) - \dot{L}_{2'}(t) - \dot{L}_2(t)\Big) \\ -L_{2'}(t) - L_{2'}(t)\Big(-\dot{L}_2(t)\Big) - L_2(t)\,. \end{aligned} \tag{F.8}$$

All the terms on the RHS of (F.8), except t, have to add up to zero for time-closure. The zeroth order terms need to cancel each other to give an L-closed string. In addition, the remaining first order terms need to add to zero to give an $\dot{L}$-closed string.

It can be seen in (F.8), that, the contribution of first order terms grouped together for each link is accumulated from those links to its right. More precisely, if the link is a time-advance, then this contribution is the sum of the products of the time-advance with the derivative of itself and with the derivative of each of the remaining time-advances or time-delays to its right in the string. If the link is a time-delay, this contribution is the sum of the products of the time-delay with the derivative of each of the time-advances or time-delays to its right in the string. It can also be seen in (F.8) that the sign in front of each variable has been preserved from the string. For example, the fourth term in the first line of (F.8) is link $\overrightarrow{3}$'s

contribution. The '+' in front of $L_3(t)$ indicates that the link travels forward in time. The first first-order term inside the bracket that it multiplies with, $\dot{L}_3(t)$, is its own, and because it is a time-advance link, it also has a positive sign. The next derivative is from the first link to its right, $\overleftarrow{2'}$. Since it is a time-delay, there is a '$-$' in front of $\dot{L}_{2'}(t)$. This process goes on for all other links to its right. In fact, by making these connections, it is possible to represent all the first order terms in terms of strings, by making the following replacements:

$$[+L_l(t)][+\dot{L}_l(t)] \longrightarrow \overrightarrow{l}\,\overrightarrow{l} \tag{F.9}$$

$$[-L_l(t)][-\dot{L}_l(t)] \longrightarrow \overleftarrow{l}\,\overleftarrow{l} \tag{F.10}$$

$$[+L_l(t)][-\dot{L}_l(t)] \longrightarrow \overrightarrow{l}\,\overleftarrow{l} \tag{F.11}$$

$$[-L_l(t)][+\dot{L}_l(t)] \longrightarrow \overleftarrow{l}\,\overrightarrow{l} \tag{F.12}$$

From (F.9) to (F.12), it is clear that the two-string products in each of the following pairs cancel each other in a sum: $(\overrightarrow{l}\,\overrightarrow{l}, \overrightarrow{l}\,\overleftarrow{l})$, $(\overrightarrow{l}\,\overrightarrow{l}, \overleftarrow{l}\,\overrightarrow{l})$, $(\overleftarrow{l}\,\overleftarrow{l}, \overrightarrow{l}\,\overleftarrow{l})$ and $(\overleftarrow{l}\,\overleftarrow{l}, \overleftarrow{l}\,\overrightarrow{l})$. It can also be seen that if these replacements are made in (F.8), collecting all the first order terms amounts to collecting the following set of two-links strings: $\{\overrightarrow{3'}\,\overrightarrow{3'}, \overrightarrow{3'}\,\overrightarrow{3}, \overrightarrow{3'}\,\overleftarrow{2'}, \overrightarrow{3'}\,\overleftarrow{2}, \overrightarrow{3}\,\overrightarrow{3}, \overrightarrow{3}\,\overleftarrow{2'}, \overrightarrow{3}\,\overleftarrow{2}, \overleftarrow{2'}\,\overleftarrow{2}\}$. Since none of these pair up to cancel each other, $t \neq t'$ to first order/degree in $\dot{L}_l(t)$, and hence the $X0$ is not $\dot{L}$-closed.

The string of the $G2$ Michelson observable $X2$, given in (3.21), is:

$$X2 \;:\; \overrightarrow{3'322'22'3'3}\;\overleftarrow{2'233'33'2'2} \tag{F.13}$$

There are 4 different link numbers in this string: 2, 2′, 3 and 3′. By collecting the two-link strings in the same way as above, one finds that each of all possible ordered pairs of these 4 link numbers appear 8 times.[2] Four of these appearances are in a negative two-link string $\overrightarrow{l}\,\overleftarrow{m}$, two in a positive two-link string $\overrightarrow{l}\,\overrightarrow{l}$ and a further two in a $\overleftarrow{l}\,\overleftarrow{l}$. For example, the two-link strings in which the ordered pair $(3', 2')$ appear are $\{4 \times \overrightarrow{3'}\,\overleftarrow{2'}, 2 \times \overrightarrow{3'}\,\overrightarrow{2'}, 2 \times \overleftarrow{3'}\,\overleftarrow{2'}\}$ Therefore, all 8 two-link strings for each possible ordered pair of link numbers cancel each other, demonstrating that $X2$ is indeed $\dot{L}$-closed.

[2] Since there are $4 \times 4 = 16$ possible ordered pairs, there are in total $16 \times 8 = 128$ two-link strings collected.

Appendix G

Computation of $\gamma_{lm}^{IJ}(f,t)$ using PYSPHARM

Since the overlap-reduction $\gamma_{IJ}(\theta,\phi)$ is in general a complex-valued function over S^2, its multipole moments can be written as

$$\gamma_{lm} = p_{lm} + iq_{lm} \ , \tag{G.1}$$

where p_{lm} and q_{lm} are the multipole moments of the real and the imaginary part of $\gamma(\theta,\phi)$, respectively. In this work, the definition of the spherical harmonic is

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)}{(l+m)}}\, P_l^m(\cos\theta)\, e^{im\phi} \ . \tag{G.2}$$

The overlap-reduction is expanded in spherical harmonics as in (5.15), and its multipole moments are defined as in (6.1). The integral for computing the multipole moments are done numerically with PYSPHARM. In this package, the spherical harmonic, $Y'_{lm}(\hat{k})$, is related to the one in (G.2) by

$$Y'_{lm}(\theta,\phi) = (-1)^m\ \sqrt{2\pi}\ Y_{lm}(\theta,\phi) \ , \tag{G.3}$$

and the multipole moments of a real-valued function over S^2, $F(\hat{k})$, is defined by

$$F_{lm} = \int_{S^2} d\hat{k} Y'^{*}_{lm}(\hat{k}) F(\hat{k}) \ . \tag{G.4}$$

Using (G.1), (G.2), (G.3), (G.4) and (6.1), it can be shown that, if p'_{lm} and q'_{lm} are the multipole moments of the real and imaginary part of $\gamma_{IJ}(\hat{k})$ as computed by PYSPHARM,

$$\gamma_{lm} = \sqrt{2\pi}(-1)^m(p'^{*}_{lm} + iq'^{*}_{lm}) \ . \tag{G.5}$$

Since PYSPHARM only works with real-valued functions, when using it to carry out the inverse transform from the spherical harmonic basis to the pixel basis ($\gamma_{lm}^{IJ} \to \gamma_{IJ}(\hat{k})$), it is necessary to obtain p_{lm} and q_{lm} first, using:

$$p_{lm} = \frac{1}{2}\left(\gamma_{lm} + \gamma_{l-m}^{*}\right) \quad , \quad q_{lm} = \frac{1}{2i}\left(\gamma_{lm} - \gamma_{l-m}^{*}\right) \ . \tag{G.6}$$

If the chosen grid over the sky has $nlat$ number of latitudes, then the maximum degree l_{max} up to which multipole moments can be computed is $nlat - 1$.

Using the implementation in this work, the procedure for computing $\gamma_{lm}^{IJ}(f,t)$ is as follows:

1. Choose the number of longitudes (`nlon`) and latitudes (`nlat`) in the sky at which the overlap-reduction will be evaluated. Use these to create an object `myLISAmodule.mySpharmt(nlon,nlat)`, which represents a gridded sky.

2. Choose a model of LISA and create its `synthlisa.LISA()` object. For example, for the non-static and rotating LISA, create with `synthlisa.EccentricInclined()`. This object contains all the information about the geometrical proporties of LISA, such as spacecraft positions, direction of arms, etc.

3. Use the sky object and the LISA object to create a `myLISAmodule.LISA_in_the_Sky()` object. This allows the computation of various LISA-related functions over the sky, in the pixel and the spherical harmonic basis.

4. To compute the multipole moments of the overlap-reductions between TDI observables, use the method `myLISAmodule.LISA_in_the_Sky.get_tdiORF_SpHs()`. For example, for the overlap-reduction between the Michelson-defined second generation optimal TDI observables I and J, do
   `myLISAmodule.LISA_in_the_Sky.get_tdiORF_SpHs( lmax ,\`
   `('Michelson','G2','I','1'),('Michelson','G2','J','1'),t )`, where `lmax` is the maximum degree of multipole moments, and `t` is the time at which evaluate. This statement returns the multipole moments as computed by PYSPHARM, i.e p'_{lm} and q'_{lm}.

Appendix H

Normalisation Test Results

seed	$\hat{\sigma}^2_{h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt,weak}$	$\hat{\sigma}^2_{h,opt} - \sigma^2_h$
100	0.1000121868	0.0002701616	0.0002689672	-0.0000121868
101	0.1002028924	0.0002701774	0.0002689832	-0.0002028924
102	0.0999468485	0.0002700735	0.0002688786	0.0000531515
103	0.1004459623	0.0002701956	0.0002690015	-0.0004459623
104	0.0998756541	0.0002701961	0.0002690019	0.0001243459
105	0.1001151961	0.0002701340	0.0002689397	-0.0001151961
106	0.1000468014	0.0002700992	0.0002689045	-0.0000468014
107	0.1001673644	0.0002701060	0.0002689114	-0.0001673644
108	0.1002668815	0.0002701543	0.0002689600	-0.0002668815
109	0.1003274655	0.0002700679	0.0002688729	-0.0003274655
110	0.0998033705	0.0002702349	0.0002690409	0.0001966295

TABLE H.1: Normalisation test results. *Expected values*: $\sigma^2_h = 0.1$; $\sigma^2_n = 1.0$; $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0002751186$. *Simulation parameters*: $\Delta t = 0.5$ s. *Analysis parameters*: $f_{low} = 1.74 \times 10^{-4}$ Hz , $f_{high} = 4.998 \times 10^{-1}$ Hz.

seed	$\hat{\sigma}^2_{h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt,weak}$	$\hat{\sigma}^2_{h,opt} - \sigma^2_h$
201	0.1001681726	0.0002701456	0.0002689511	-0.0001681726
202	0.1001396581	0.0002702093	0.0002690151	-0.0001396581
203	0.1006516748	0.0002701294	0.0002689350	-0.0006516748
204	0.0998504792	0.0002701361	0.0002689417	0.0001495208
205	0.1005450540	0.0002701280	0.0002689336	-0.0005450540
206	0.1002498240	0.0002701268	0.0002689323	-0.0002498240
207	0.0996261941	0.0002701036	0.0002689091	0.0003738059
208	0.0999693202	0.0002701342	0.0002689398	0.0000306798
209	0.1002613609	0.0002701055	0.0002689109	-0.0002613609
210	0.1006541285	0.0002700648	0.0002688700	-0.0006541285

TABLE H.2: Normalisation test results. *Expected values*: $\sigma^2_h = 0.1$; $\sigma^2_n = 1.0$; $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0002751186$. *Simulation parameters*: $\Delta t = 0.5$ s; $T_{seg} = 10^4$ s. *Analysis parameters*: $f_{low} = 1.74 \times 10^{-4}$ Hz , $f_{high} = 4.998 \times 10^{-1}$ Hz.

seed	$\hat{\sigma}^2_{h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt,weak}$	$\hat{\sigma}^2_{h,opt} - \sigma^2_h$
800	0.0995620888	0.0002701032	0.0002689086	0.0004379112
801	0.1000074212	0.0002701600	0.0002689657	-0.0000074212
802	0.1001900230	0.0002700530	0.0002688580	-0.0001900230
803	0.0993372496	0.0002701079	0.0002689133	0.0006627504
804	0.1001681209	0.0002700742	0.0002688792	-0.0001681209
805	0.0996781177	0.0002701277	0.0002689332	0.0003218823
806	0.1001524660	0.0002701271	0.0002689327	-0.0001524660
807	0.0995229446	0.0002702100	0.0002690158	0.0004770554
808	0.0997691187	0.0002701434	0.0002689490	0.0002308813
809	0.1004605907	0.0002701175	0.0002689229	-0.0004605907
810	0.1002592572	0.0002700958	0.0002689013	-0.0002592572
811	0.1001017844	0.0002701233	0.0002689288	-0.0001017844
812	0.1001367431	0.0002701326	0.0002689382	-0.0001367431
813	0.1002440989	0.0002702426	0.0002690486	-0.0002440989
814	0.1007050092	0.0002701921	0.0002689979	-0.0007050092
815	0.0997936499	0.0002702100	0.0002690158	0.0002063501
816	0.0998306095	0.0002700699	0.0002688752	0.0001693905
817	0.1000177342	0.0002701777	0.0002689834	-0.0000177342
818	0.1002647423	0.0002702009	0.0002690067	-0.0002647423
819	0.1002015380	0.0002701774	0.0002689829	-0.0002015380

TABLE H.3: Normalisation test results. *Expected values*: $\sigma^2_h = 0.1$; $\sigma^2_n = 1.0$; $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0002751186$. *Simulation parameters*: $\Delta t = 0.5$ s; $T_{seg} = 10^4$ s; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 1.74 \times 10^{-4}$ Hz , $f_{high} = 4.998 \times 10^{-1}$ Hz.

seed	$\hat{\sigma}^2_{h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt,weak}$	$\hat{\sigma}^2_{h,opt} - \sigma^2_h$
701	0.0999967655	0.0002701246	0.0002689300	0.0000032345
702	0.0998928474	0.0002701456	0.0002689511	0.0001071526
703	0.0996880163	0.0002700754	0.0002688804	0.0003119837
704	0.0999618697	0.0002701998	0.0002690056	0.0000381303
705	0.1003690203	0.0002701417	0.0002689473	-0.0003690203
706	0.0998481148	0.0002701498	0.0002689554	0.0001518852
707	0.1001564326	0.0002701340	0.0002689395	-0.0001564326
708	0.1002292564	0.0002701816	0.0002689873	-0.0002292564
709	0.0995635703	0.0002702094	0.0002690153	0.0004364297
710	0.1000405858	0.0002701969	0.0002690028	-0.0000405858
711	0.0997071118	0.0002701880	0.0002689936	0.0002928882
712	0.1001458465	0.0002701189	0.0002689244	-0.0001458465
713	0.0998760295	0.0002701740	0.0002689796	0.0001239705
714	0.0999956217	0.0002701477	0.0002689533	0.0000043783
715	0.1001965806	0.0002701572	0.0002689628	-0.0001965806
716	0.0998541549	0.0002702115	0.0002690172	0.0001458451
717	0.0995567656	0.0002701277	0.0002689330	0.0004432344
718	0.0997665032	0.0002700872	0.0002688925	0.0002334968
719	0.1000385900	0.0002701061	0.0002689115	-0.0000385900
720	0.1002104236	0.0002701363	0.0002689417	-0.0002104236

TABLE H.4: Normalisation test results. *Expected values*: $\sigma^2_h = 0.1$; $\sigma^2_n = 1.0$; $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0002751186$. *Simulation parameters*: $\Delta t = 0.5$ s; $T_{seg} = 10^4$ s; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 1.74 \times 10^{-4}$ Hz , $f_{high} = 4.998 \times 10^{-1}$ Hz.

seed	$\hat{\sigma}^2_{h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt}$	$\sigma^2_{\hat{\sigma}^2_h,opt,weak}$	$\hat{\sigma}^2_{h,opt} - \sigma^2_h$
901	0.1015729451	0.0007736187	0.0007703161	-0.0015729451
902	0.0997030022	0.0007727762	0.0007694702	0.0002969978
903	0.0988896323	0.0007730753	0.0007697708	0.0011103677
904	0.0998277673	0.0007734696	0.0007701665	0.0001722327
905	0.1002811130	0.0007741382	0.0007708394	-0.0002811130
906	0.0993573934	0.0007732066	0.0007699034	0.0006426066
907	0.0990131219	0.0007735950	0.0007702928	0.0009868781
908	0.1013037885	0.0007734480	0.0007701455	-0.0013037885
909	0.1005223582	0.0007729148	0.0007696096	-0.0005223582
910	0.1000973903	0.0007732497	0.0007699456	-0.0000973903
911	0.0989267905	0.0007727417	0.0007694361	0.0010732095
912	0.0998262355	0.0007734895	0.0007701873	0.0001737645
913	0.1001171054	0.0007731355	0.0007698315	-0.0001171054
914	0.1001157782	0.0007724965	0.0007691900	-0.0001157782
915	0.0994138434	0.0007731543	0.0007698507	0.0005861566
916	0.0983763065	0.0007731783	0.0007698750	0.0016236935
917	0.0994222646	0.0007736234	0.0007703214	0.0005777354
918	0.0981205728	0.0007725203	0.0007692147	0.0018794272
919	0.0996507325	0.0007727263	0.0007694203	0.0003492675
920	0.1005400318	0.0007733934	0.0007700905	-0.0005400318
921	0.1000202315	0.0007733636	0.0007700607	-0.0000202315
922	0.1012118173	0.0007734313	0.0007701279	-0.0012118173
923	0.0994585858	0.0007732863	0.0007699821	0.0005414142
924	0.0995635141	0.0007742583	0.0007709595	0.0004364859
925	0.0992601136	0.0007737675	0.0007704659	0.0007398864
926	0.0985199917	0.0007739177	0.0007706174	0.0014800083
927	0.0995307100	0.0007729469	0.0007696421	0.0004692900
928	0.0995114493	0.0007729215	0.0007696161	0.0004885507
929	0.1002986036	0.0007736061	0.0007703038	-0.0002986036
930	0.1005659804	0.0007732971	0.0007699937	-0.0005659804

TABLE H.5: Normalisation test results. *Expected values*: $\sigma^2_h = 0.1$; $\sigma^2_n = 1.0$; $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0007842148$. *Simulation parameters*: $\Delta t = 8.0$ s; $T_{seg} = 1.6 \times 10^4$ s; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 5 \times 10^{-4}$ Hz , $f_{high} = 6.2 \times 10^{-2}$ Hz.

seed	$\hat{\mathcal{P}}_{00,opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt,weak}$	$\hat{\mathcal{P}}_{00,opt} - \mathcal{P}_{00}$
801	0.0111186357	0.0000194608	0.0000143015	0.0000075246
802	0.0111093028	0.0000194687	0.0000143129	-0.0000018083
803	0.0110867344	0.0000194664	0.0000142954	-0.0000243767
804	0.0111243119	0.0000194839	0.0000143370	0.0000132008
805	0.0111026585	0.0000194710	0.0000143263	-0.0000084526
806	0.0110977226	0.0000194653	0.0000143077	-0.0000133885
807	0.0110326497	0.0000194475	0.0000142809	-0.0000784615
808	0.0110883800	0.0000194584	0.0000142949	-0.0000227311
809	0.0111093418	0.0000194851	0.0000143200	-0.0000017693
810	0.0111185353	0.0000194761	0.0000143536	0.0000074242
811	0.0111316879	0.0000194857	0.0000143342	0.0000205768
812	0.0111096033	0.0000194733	0.0000143170	-0.0000015078
813	0.0110979527	0.0000194783	0.0000143133	-0.0000131584
814	0.0110472446	0.0000194440	0.0000142597	-0.0000638665
815	0.0111109781	0.0000194714	0.0000143061	-0.0000001330
816	0.0111205012	0.0000194588	0.0000143145	0.0000093901
817	0.0111361298	0.0000194960	0.0000143474	0.0000250187
818	0.0111214755	0.0000194725	0.0000143149	0.0000103643
819	0.0110997714	0.0000194763	0.0000143286	-0.0000113397
820	0.0110940619	0.0000194794	0.0000143296	-0.0000170492
821	0.0111110459	0.0000194857	0.0000143342	-0.0000000652
822	0.0110988300	0.0000194677	0.0000143134	-0.0000122811
823	0.0111360823	0.0000194911	0.0000143337	0.0000249712
824	0.0110922115	0.0000194715	0.0000143213	-0.0000188996
825	0.0110975396	0.0000194677	0.0000143061	-0.0000135715

TABLE H.6: Normalisation test results. *Expected values*: $\sigma_{\hat{\mathcal{P}}_{00},opt} = 0.0000194006$. *Simulation parameters*: $\Delta t = 8.0$ s; $T_{seg} = 1.6 \times 10^4$ s; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 5 \times 10^{-4}$ Hz , $f_{high} = 6.2 \times 10^{-2}$ Hz.

seed	$\hat{\mathcal{P}}_{00,opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt,weak}$	$\hat{\mathcal{P}}_{00,opt} - \mathcal{P}_{00}$
1	0.0110893917	0.0000195139	0.0000143621	-0.0000217194
2	0.0110958310	0.0000194908	0.0000143502	-0.0000152801
3	0.0110864761	0.0000195008	0.0000143379	-0.0000246350
4	0.0110826922	0.0000194982	0.0000143386	-0.0000284190
5	0.0111038955	0.0000195199	0.0000143747	-0.0000072156
6	0.0110825520	0.0000195056	0.0000143518	-0.0000285591
7	0.0110874143	0.0000194947	0.0000143156	-0.0000236968
8	0.0110936761	0.0000195088	0.0000143776	-0.0000174350
9	0.0111441072	0.0000195121	0.0000143983	0.0000329961
10	0.0110773666	0.0000195011	0.0000143459	-0.0000337445
11	0.0110756843	0.0000195146	0.0000143764	-0.0000354269
12	0.0111120158	0.0000195203	0.0000144007	0.0000009047
13	0.0111232631	0.0000195285	0.0000143815	0.0000121520
14	0.0111342979	0.0000195134	0.0000143964	0.0000231868
15	0.0111057525	0.0000195088	0.0000143601	-0.0000053587
16	0.0111244005	0.0000195124	0.0000143822	0.0000132894
17	0.0110916153	0.0000195074	0.0000143359	-0.0000194958
18	0.0111043402	0.0000195126	0.0000143688	-0.0000067709
19	0.0111198585	0.0000195079	0.0000143605	0.0000087474
20	0.0111197565	0.0000195213	0.0000143761	0.0000086454
21	0.0111159938	0.0000195159	0.0000143852	0.0000048827
22	0.0111113296	0.0000194951	0.0000143603	0.0000002185
23	0.0111140233	0.0000195219	0.0000143877	0.0000029122
24	0.0111175095	0.0000195247	0.0000143824	0.0000063983
25	0.0111102596	0.0000195139	0.0000143808	-0.0000008515

TABLE H.7: Normalisation test results. *Expected values*: $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0000194006$. *Simulation parameters*: $\Delta t = 8.0$ s; $T_{seg} = 1.6 \times 10^4$ s for signal; $T_{seg} = 4 \times 10^3$ s for noise; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 5 \times 10^{-4}$ Hz , $f_{high} = 6.2 \times 10^{-2}$ Hz.

seed	$\hat{\mathcal{P}_{00,opt}}$	$\sigma_{\hat{\mathcal{P}_{00},opt}}$	$\sigma_{\hat{\mathcal{P}_{00},opt,weak}}$	$\hat{\mathcal{P}_{00,opt}} - \mathcal{P}_{00}$
1	0.0110616317	0.0000194649	0.0000144311	-0.0000494794
2	0.0110825206	0.0000194754	0.0000144613	-0.0000285905
3	0.0110527892	0.0000194642	0.0000144018	-0.0000583219
4	0.0110738544	0.0000194606	0.0000144289	-0.0000372567
5	0.0110973769	0.0000194612	0.0000144284	-0.0000137342
6	0.0111220029	0.0000194634	0.0000144325	0.0000108918
7	0.0110858715	0.0000194669	0.0000144219	-0.0000252396
8	0.0111441811	0.0000194745	0.0000144656	0.0000330700
9	0.0110683765	0.0000194573	0.0000144366	-0.0000427346
10	0.0110918240	0.0000194654	0.0000144331	-0.0000192871
11	0.0111074313	0.0000194741	0.0000144487	-0.0000036798
12	0.0110900006	0.0000194657	0.0000144359	-0.0000211105
13	0.0110997401	0.0000194776	0.0000144451	-0.0000113710
14	0.0110969433	0.0000194785	0.0000144526	-0.0000141679
15	0.0110855747	0.0000194706	0.0000144511	-0.0000255364
16	0.0111115322	0.0000194658	0.0000144496	0.0000004211
17	0.0110628483	0.0000194567	0.0000144097	-0.0000482628
18	0.0110964473	0.0000194605	0.0000144291	-0.0000146639
19	0.0110935064	0.0000194843	0.0000144538	-0.0000176047
20	0.0110980166	0.0000194637	0.0000144421	-0.0000130945
21	0.0111274367	0.0000194827	0.0000144557	0.0000163256
22	0.0110735412	0.0000194514	0.0000144202	-0.0000375699
23	0.0110629763	0.0000194454	0.0000143835	-0.0000481348
24	0.0110823278	0.0000194631	0.0000144353	-0.0000287834
25	0.0111036606	0.0000194695	0.0000144369	-0.0000074505

TABLE H.8: Normalisation test results. *Expected values*: $\sigma^2_{\hat{\sigma^2_h},opt} = 0.0000194006$. *Simulation parameters*: $\Delta t = 8.0$ s; $T_{seg} = 4 \times 10^3$ s for signal; $T_{seg} = 4 \times 10^3$ s for noise; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 5 \times 10^{-4}$ Hz , $f_{high} = 6.2 \times 10^{-2}$ Hz.

seed	$\hat{\mathcal{P}}_{00,opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt,weak}$	$\hat{\mathcal{P}}_{00,opt} - \mathcal{P}_{00}$
501	0.0111218295	0.0000207135	0.0000155339	0.0000107183
502	0.0111290972	0.0000207205	0.0000155452	0.0000179861
503	0.0111050288	0.0000207017	0.0000155206	-0.0000060823
504	0.0110866264	0.0000206921	0.0000155070	-0.0000244847
505	0.0110798036	0.0000206943	0.0000155238	-0.0000313075
506	0.0111198169	0.0000207300	0.0000155667	0.0000087058
507	0.0110713133	0.0000207020	0.0000155063	-0.0000397978
508	0.0110975180	0.0000207011	0.0000155148	-0.0000135931
509	0.0111418872	0.0000207247	0.0000155505	0.0000307761
510	0.0111196959	0.0000207275	0.0000155584	0.0000085848
511	0.0111135386	0.0000207170	0.0000155297	0.0000024275
512	0.0110784103	0.0000207037	0.0000155329	-0.0000327008
513	0.0111203821	0.0000207121	0.0000155286	0.0000092710
514	0.0111014618	0.0000207101	0.0000155362	-0.0000096493
515	0.0111095102	0.0000206985	0.0000155158	-0.0000016009
516	0.0111207899	0.0000207160	0.0000155389	0.0000096788
517	0.0111522030	0.0000207238	0.0000155439	0.0000410919
518	0.0111059290	0.0000207153	0.0000155467	-0.0000051822
519	0.0111046505	0.0000207032	0.0000155260	-0.0000064606
520	0.0110857537	0.0000206910	0.0000154968	-0.0000253574
521	0.0111166129	0.0000207079	0.0000155254	0.0000055018
522	0.0111217827	0.0000207118	0.0000155479	0.0000106716
523	0.0111180723	0.0000207205	0.0000155448	0.0000069612
524	0.0111131040	0.0000207280	0.0000155499	0.0000019928
525	0.0110842787	0.0000207156	0.0000155338	-0.0000268325

TABLE H.9: Normalisation test results. *Expected values*: $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0000207793$. *Simulation parameters*: $\Delta t = 8.0$ s; $T_{seg} = 1.6 \times 10^4$ s; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 5 \times 10^{-4}$ Hz , $f_{high} = 6.2 \times 10^{-2}$ Hz.

seed	$\hat{\mathcal{P}}_{00,opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt,weak}$	$\hat{\mathcal{P}}_{00,opt} - \mathcal{P}_{00}$
301	2.2305591339	0.0076004397	0.0070525511	0.0083369117
302	2.2293927204	0.0076078243	0.0070601345	0.0071704982
303	2.2152736659	0.0075955667	0.0070478517	-0.0069485564
304	2.2351533839	0.0076052187	0.0070577541	0.0129311617
305	2.2128773671	0.0075982422	0.0070504428	-0.0093448552
306	2.2252766979	0.0075863746	0.0070368393	0.0030544757
307	2.2246443121	0.0076033411	0.0070555503	0.0024220899
308	2.2238630262	0.0076041268	0.0070577609	0.0016408039
309	2.2172531566	0.0075983803	0.0070508443	-0.0049690656
310	2.2103386163	0.0076025409	0.0070560038	-0.0118836059
311	2.2150544305	0.0075952948	0.0070477348	-0.0071677918
312	2.2381923188	0.0076083016	0.0070635699	0.0159700966
313	2.2250805893	0.0075992738	0.0070520290	0.0028583670
314	2.2305083613	0.0076067313	0.0070597506	0.0082861390
315	2.2261222853	0.0075985951	0.0070511016	0.0039000630
316	2.2193832824	0.0075973510	0.0070502358	-0.0028389398
317	2.2042046113	0.0075996835	0.0070516402	-0.0180176109
318	2.2166324098	0.0075976501	0.0070503552	-0.0055898124
319	2.2191441804	0.0075986116	0.0070512152	-0.0030780418
320	2.2041298778	0.0075862992	0.0070367883	-0.0180923444
321	2.2089420432	0.0075954186	0.0070470237	-0.0132801790
322	2.2250281945	0.0076017435	0.0070544860	0.0028059722
323	2.2072353888	0.0076041181	0.0070567693	-0.0149868334
324	2.2301225515	0.0075922776	0.0070450653	0.0079003293
325	2.2232867632	0.0075978200	0.0070488563	0.0010645409
326	2.2186149218	0.0076016183	0.0070546619	-0.0036073004
327	2.2180076110	0.0075995872	0.0070531477	-0.0042146112
328	2.2183795282	0.0075881171	0.0070392984	-0.0038426940
329	2.2223203745	0.0076603259	0.0071095620	0.0000981523
330	2.2243147377	0.0076035104	0.0070563246	0.0020925155
331	2.2340146724	0.0076021688	0.0070558984	0.0117924502
332	2.2276110782	0.0076013040	0.0070533507	0.0053888560
333	2.2243284026	0.0075985184	0.0070514228	0.0021061803
334	2.2288139516	0.0076032032	0.0070555736	0.0065917294
335	2.2264606869	0.0076015595	0.0070545082	0.0042384647
336	2.2188142765	0.0076054008	0.0070579125	-0.0034079457
337	2.2240354063	0.0076045242	0.0070585732	0.0018131841
338	2.2205902267	0.0076141549	0.0070677814	-0.0016319955
339	2.2176972338	0.0076139568	0.0070693879	-0.0045249884
340	2.2254663919	0.0075959285	0.0070466854	0.0032441697

TABLE H.10: Normalisation test results. *Expected values*: $\sigma^2_{\hat{\sigma}^2_h,opt} = 0.0075400861$. *Simulation parameters*: $\Delta t = 8.0$ s; $T_{seg} = 1.6 \times 10^4$ s; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 5 \times 10^{-4}$ Hz , $f_{high} = 6.2 \times 10^{-2}$ Hz.

seed	$\hat{\mathcal{P}}_{00,opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt}$	$\sigma_{\hat{\mathcal{P}}_{00},opt,weak}$	$\hat{\mathcal{P}}_{00,opt} - \mathcal{P}_{00}$
801	0.0112388678	0.0001447520	0.0001441706	0.0001277567
802	0.0112441473	0.0001446605	0.0001440778	0.0001330361
803	0.0111379832	0.0001446067	0.0001440252	0.0000268721
804	0.0111927169	0.0001448983	0.0001443174	0.0000816058
805	0.0109175705	0.0001449834	0.0001444038	-0.0001935406
806	0.0113206443	0.0001446492	0.0001440663	0.0002095332
807	0.0107948984	0.0001445274	0.0001439446	-0.0003162127
808	0.0110132069	0.0001445691	0.0001439857	-0.0000979042
809	0.0111413250	0.0001448643	0.0001442823	0.0000302139
810	0.0112916326	0.0001448211	0.0001442411	0.0001805215
811	0.0110928497	0.0001450340	0.0001444537	-0.0000182615
812	0.0112007679	0.0001446402	0.0001440581	0.0000896567
813	0.0108961767	0.0001448017	0.0001442197	-0.0002149344
814	0.0107775015	0.0001443412	0.0001437576	-0.0003336096
815	0.0112182157	0.0001447686	0.0001441877	0.0001071046
816	0.0112505430	0.0001448044	0.0001442241	0.0001394319
817	0.0113151693	0.0001449842	0.0001444039	0.0002040582
818	0.0111218694	0.0001447702	0.0001441884	0.0000107583
819	0.0112030525	0.0001447444	0.0001441628	0.0000919414
820	0.0109533648	0.0001448699	0.0001442894	-0.0001577463
821	0.0111869210	0.0001448400	0.0001442596	0.0000758099
822	0.0110853118	0.0001447456	0.0001441643	-0.0000257994
823	0.0111861967	0.0001448998	0.0001443190	0.0000750856
824	0.0111468667	0.0001447274	0.0001441460	0.0000357556
825	0.0110757793	0.0001447909	0.0001442095	-0.0000353318

TABLE H.11: Normalisation test results. *Expected values*: $\sigma^2_{\hat{\sigma^2_h},opt} = 0.0001470824$. *Simulation parameters*: $\Delta t = 8.0$ s; $T_{seg} = 1.6 \times 10^4$ s; $\mathcal{P}_{00} = 1/90$. *Analysis parameters*: $f_{low} = 5 \times 10^{-4}$ Hz , $f_{high} = 6.2 \times 10^{-2}$ Hz.

Bibliography

[1] A. Krolak et al. Optimal filtering of the LISA data. 2009.

[2] M. Vallisneri. Geometric Time Delay Interferometry. 2005.

[3] LISA Study Team. *LISA*: Laser Interferometric Space Antenna for the detectiona and observation of gravitational waves, Pre-Phase A Report. 1998.

[4] B.S Sathyaprakash Bernard Schutz. Physics, Astrophysics and Cosmology with Gravitational Waves. *Living Reviews in Relativity*, 2009.

[5] Bernard Schutz. *Gravity from the ground up.* Cambridge, 2003.

[6] Michele Maggiore. *Gravitational Waves. Volume* 1*: Theory and experiments.* Oxford, 2008.

[7] N.J. Cornish and L. J. Rubbo. The LISA Response Function. 2003.

[8] H. Kudoh and A. Taruya. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer: *Geometric properties of antenna patterns and their angular power*. 2005.

[9] M. Vallisneri. Synthetic LISA:Simulating Time Delay Interferometry in a Model LISA . *Phys. Rev. D*, 2005.

[10] Sanjeev V. Dhurandhar Massimo Tinto. Time-Delay Interferometry. *Living Reviews in Relativity*, 2005.

[11] Pablo A. Rosado. Gravitational wave background from binary systems. 2011.

[12] Robert Wilson Arno Penzias. A measurement of excess antenna temperature at $4080Mc/s$. *Astrophysical journal*, 142:419, 1965.

[13] I.N. Reid I.B Thompson M.H. Siegel, S.R Majewski. Star counts redivivus. IV. Density laws through photometric parallaxes. *Astrophysical journal*, 578:151, 2002.

[14] Andrzej Krolak Gijs Helemans Jeffrey A. Edlund, Massimo Tinto. Simulation of the White Dwarf - White Dwarf galactic background in the LISA data. . *Class. Quantum Grav.*, 2005.

[15] Neil Cornish. Mapping the gravitational wave background. 2001.

[16] H. Kudoh and A. Taruya. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer II: perturbative reconstruction of a low-frequency skymap. 2005.

[17] A. Taruya. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer III: Reconstruction of a high-frequency skymap. 2006.

[18] E. Thrane et al. Probing the anisotropies of a stochastic gravitational-wave background using a network of ground-based laser interferometers. *Phys. Rev. D*, 80:122002, 2009.