Estimation of spectrum and parameters of relic gravitational waves using space-borne interferometers

Bo Wang and Yang Zhang

Key Laboratory for Researches in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei 230026, China; ymwangbo@mail.ustc.edu.cn, yzh@ustc.edu.cn

Received 2018 June 19; accepted 2018 August 8

Abstract We present a study of spectrum estimation of relic gravitational waves (RGWs) as a Gaussian stochastic background from output signals of future space-borne interferometers, like LISA and ASTROD. As the target of detection, the analytical spectrum of RGWs generated during inflation is described by three parameters: the tensor-scalar ratio, the spectral index and the running index. The Michelson interferometer is shown to have a better sensitivity than Sagnac and symmetrized Sagnac. For RGW detection, we analyze the auto-correlated signals for a single interferometer, and the cross-correlated, integrated as well as un-integrated signals for a pair of interferometers, and give the signal-to-noise ratio (SNR) for RGW, and obtain lower limits of the RGW parameters that can be detected. By suppressing noise level, a pair has a sensitivity 2 orders better than a single for one year observation. SNR of LISA will be 4–5 orders higher than that of Advanced LIGO for the default RGW. To estimate the spectrum, we adopt the maximum likelihood (ML) estimation, calculate the mean and covariance of signals, obtain the Gaussian probability density function (PDF) and the likelihood function, and derive expressions for the Fisher matrix and the equation of the ML estimate for the spectrum. The Newton-Raphson method is used to solve the equation by iteration. When the noise is dominantly large, a single LISA is not effective for estimating the RGW spectrum as the actual noise in signals is not known accurately. For cross-correlating a pair, the spectrum cannot be estimated from the integrated output signals either, and only one parameter can be estimated with the other two being either fixed or marginalized. We use the ensemble averaging method to estimate the RGW spectrum from the un-integrated output signals. We also adopt a correlation of un-integrated signals to estimate the spectrum and three parameters of RGW in a Bayesian approach. For all three methods, we provide simulations to illustrate their feasibility.

Key words: gravitational waves — cosmological parameters — instrumentation: detectors — early universe

1 INTRODUCTION

Gravitational waves (GWs) are a prediction of Einstein’s theory of general relativity, and have been the subject of theoretical study and continuous detection hunting. There are two kinds of GWs, i.e., the first includes those generated by astrophysical processes such as inspiral of compact binaries, merging of massive black holes, supermassive black hole binaries (Sesana et al. 2008; Janssen et al. 2015), etc. The frequencies of these sources are typically in the range $f \sim 10^{-9} - 10^{3}$ Hz. Examples are GW150914, GW151226 and GW170104 from merging of binary black holes and GW170817 from a binary neutron star inspiral that was recently reported by Advanced LIGO and Advanced Virgo as the first direct detections (Abbott et al. 2016c,b,a, 2017b,d,c,a).

Another kind is the relic gravitational wave (RGW), which is generated during the inflation stage of cosmic expansion, as generically predicted by inflation models (Grishchuk 1975, 1997, 2001; Ford & Parker 1977; Starobinskiï 1979; Starobinskiï 1985; Fabbri & Pollock 1983; Abbott & Wise 1984; Allen 1988; Sahni 1990; Giovannini 1999). RGW carries crucial information about the very early Universe, such as the energy scale and slope
of inflation potential, the initial quantum states during inflation (Zhang et al. 2005, 2006; Zhao & Zhang 2006b; Wang et al. 2016), as well as the reheating process (Tong et al. 2014). This is because, to the linear level of metric perturbations, RGW is independent of other matter components and its propagation is almost free. The influences due to neutrinos free-streaming (Weinberg 2004; Miao & Zhang 2007), quark-hadron transition and $e^+e^-$ annihilation are minor modifications (Wang et al. 2008). This is in contrast to the scalar metric perturbation, which is always coupled to cosmic matters and whose short wavelength modes have gone into nonlinear evolution at present. The second-order perturbation beyond the linear perturbation has been also studied for RGW (Ananda et al. 2007; Baumann et al. 2007; Wang & Zhang 2017; Zhang et al. 2017).

RGW has several interesting properties that are quite valuable for GW detection. It is a stochastic background of spacetime fluctuations distributed everywhere in the present Universe, just like the cosmic microwave background (CMB). Moreover, RGW also exists all the present Universe, just like the cosmic microwave background (CMB). Moreover, RGW is independent of other matter components and whose short wavelength modes have gone into nonlinear evolution at present. The second-order perturbation beyond the linear perturbation has been also studied for RGW (Ananda et al. 2007; Baumann et al. 2007; Wang & Zhang 2017; Zhang et al. 2017).

A primary feature of the RGW spectrum is that it has higher amplitude at lower frequencies (Zhang et al. 2005, 2006; Zhao & Zhang 2006b). The highest amplitude is located around $(10^{-18} - 10^{-16})$ Hz which is the target of CMB measurements. So far, the magnetic polarization $C_{BB}^i$ induced by RGW (Basko & Polnarev 1980b,a; Polnarev 1985; Zaldarriaga & Harari 1995; Kosowsky 1996; Kamionkowski et al. 1997; Zhao & Zhang 2006a; Xia & Zhang 2008, 2009) has not yet been detected, and only some constraint is given in terms of the tensor-scalar ratio of metric perturbations $r < 0.1$ (Planck Collaboration et al. 2016a,b,c,d). On the other hand, Advanced LIGO-Virgo so far has not detected RGW, but rather has only been applied to predict a total stochastic GW background with amplitude $1.8^{+2.7}_{-1.3} \times 10^{-9}$ near 25 Hz contributed together by unresolved binaries, RGW, etc (Abbott et al. 2018). In between is the band of the space-borne facilities, LISA and ASTROD, where the amplitude of RGW is higher by 5–6 orders than that in the LIGO band. This great enhancement increases the chance for space-borne interferometers to detect RGWs if their sensitivity level is comparable to LIGO.

In this paper, we shall study RGW detection by space-borne interferometers, such as LISA and ASTROD and the like, and show how to estimate the spectrum and parameters of RGW from output signals of future observations. For this purpose, we shall first briefly introduce the theoretical RGW spectrum as a scientific target, resulting from an analytical solution that covers from inflation to the present acceleration (Zhang et al. 2005, 2006; Zhao & Zhang 2006b). Accurate estimation of this spectrum will also confirm the details of inflation for the very early Universe. In this sense, this will be a direct detection of inflation. For the RGW spectrum in this paper, we focus on three parameters determined by inflation: the tensor-scalar ratio $r$, the spectral index $\beta$ and the spectral running index $\alpha_t$ (Tong & Zhang 2009; Wang et al. 2016). Small modifications of RGW in Weinberg (2004), Miao & Zhang (2007) and Wang et al. (2008) are not considered. We do not consider the Doppler modulation due to orbital motion or related causes (Cornish & Larson 2003a,b; Hellings 2003; Timpano et al. 2006; Ungarelli & Vecchio 2001b). One of the main obstacles to detecting RGW is the stochastic foreground of a GW resulting from the superposition of a
large number of unresolved astrophysical sources. To have a
definite model of the power spectrum of the stochastic
foreground, one has to know the spectrum for each type of
source, as well as the evolution of each type. There are
several categories of these source types. Due to their large
abundance, Galactic white dwarf binaries are generally
considered one of the main components of the foreground
in the $\sim 10^{-5}$ Hz frequency band (Evans et al. 1987; Hils
et al. 1990; Bender & Hils 1997). Nelemans et al. (2001a),
Nelemans et al. (2001b), Ruiter et al. (2009), Seto &
Cooray (2004), Adams et al. (2012) and Adams & Cornish
(2014) provide several models of the foreground gener-
ated from distribution of these binaries. Sesana et al. (2004),
Sesana et al. (2005) have studied a stochastic foreground
from massive black hole binaries and its contribution to the
LISA data stream. Hils & Bender (2000), Nelemans et al.
(2004), Solheim (2010) show the possibility of a fore-
ground generated by an AM CVn binary system. To ex-
plor the effects of these foreground models on a spaced-
based detector, simulation methods to generate a fore-
ground data stream for LISA have been studied by the
Mock LISA data challenge project (Arnaud et al. 2007;
Babak et al. 2008; https://astrogravs.nasa.gov/docs/mldc/)
and other groups (Cornish & Crowder 2005; Cornish &
Robson 2017). Based on these dummy data streams, sev-
eral techniques for model selection and parameter estima-
tion have been developed (Seto & Cooray 2004; Robson
& Cornish 2017; Cornish & Littenberg 2007; Cornish
& Robson 2017; Adams & Cornish 2014). Crowder &
Cornish (2007) and Robson & Cornish (2017) have pro-
vided methods to detect resolvable sources in a foreground
of unresolved sources. Adams & Cornish (2014) inves-
tigated approaches to discriminate the GW background
from a stochastic foreground according to the differences
in spectral shapes and time modulation of the signal.
Currently, the foreground is still under intensive study but
is not fully understood. At this stage of our study we do
not include the foreground in this paper.

GW radiation from a finite source usually has a def-
inite waveform (fixed direction, amplitude, etc) and the
match-filter method (Helstrom 1960; Jaranowski & Królak
2005) is commonly used to estimate the waveform against
certain theoretical templates. Cutler & Flanagan (1994)
and Finn (1992) studied the methods of parameter estima-
tion for ground-based LIGO detectors. Cutler (1998) and
Moore & Hellings (2002) studied detection of a GW radi-
ated from merging compact binaries using LISA detectors.
In contrast, RGW is of stochastic nature, incident from
all directions, containing modes of all possible frequen-
cies and amplitudes. Flanagan (1993), Allen (1997) and
Allen & Romano (1999) systematically studied detection
of RGW using LIGO, and obtained a formula for signal-
to-noise ratio (SNR) as a criterion for detection. Binétruy
et al. (2012) and Caprini et al. (2016) discussed possible
detection by eLISA of GW backgrounds due to first-
order phase transitions, cosmic strings, bubble collision,
etc. Ungarelli & Vecchio (2001a) discussed the possibility
of RGW detection by LISA. So far in the literature, how-
ever, RGW detection by space-based interferometers has
not been systematically studied, in particular, estimation of
the RGW spectrum has not been analyzed. We shall derive
formulations for estimation of the RGW spectrum, using
a single or a pair of space-based interferometers like LISA
and ASTROD.

For this purpose, we shall briefly examine the three
types of interferometers: Michelson, Sagnac and sym-
metrized Sagnac (de Vine et al. 2010; Cornish & Rubbo
2003; Shaddock 2004; Estabrook et al. 2000; Cornish &
Hellings 2003; Schilling 1997; Vallisneri 2008; Larson
et al. 2000; Cornish 2001; Cornish & Robson 2001;
Robinson et al. 2008), whose sensitivity depends on both
the noise and transfer function, which in turn depends on
the detector geometry. We shall show explicitly that the
Michelson has the best sensitivity, which will be taken
as a default interferometer. For a single interferometer
in space, we give SNR and a criterion to detect RGW.
As a Gaussian stochastic background, RGW is similar to
CMB anisotropies, and the statistical methods employed
in CMB studies can be used (Gorski et al. 1994; Jungman
et al. 1996; Tegmark et al. 1997; Oh et al. 1999; Hinshaw
et al. 2003). We shall apply the maximum likelihood (ML)
method (Kay 1993a,b) to estimate the RGW spectrum.
We give the probability density function (PDF) explicitly,
and derive the estimation equation of an RGW spectrum.
However, in practice, our knowledge of the spectrum of
noise that is actually occurring in the detector is not suf-
cient so that a single case is not effective to estimate the
RGW spectrum when the noise is dominantly large.

For a pair, the noise level will be suppressed by cross-
correlation. We shall introduce cross-correlated, integrated
output of the pair, in a fashion similar to the ground-based
LIGO (Allen 1997; Allen & Romano 1999), calculate the
overlapping reduction function, give the sensitivity and
compare with that of a single case, and analyze possible
detection and constraints on RGW parameters. However,
the spectrum as a function of frequency cannot be es-
timated from the integrated output, since the frequency-
dependence has been lost in integration. One can esti-
mate only one parameter in the Bayesian approach by ML-estimation, using the Newton-Raphson method (Oh et al. 1999; Hinshaw et al. 2003; Press et al. 1992). To estimate the spectrum, we propose the ensemble averaging method, and directly take the cross-product of un-integrated output signals from a pair. The method does not depend on precise knowledge of the noise spectrum. We estimate the spectrum using simulated data for illustration, but one cannot estimate the parameters. Seto (2006) suggested a method of correlation for un-integrated signals, by which the whole frequency range of the data is to be divided into many small segments of frequency, and the mean value of a correlation variable over each small segment is taken as the representative point for the segment. In this way, as an approximation, the correlation variable as a function of frequency is defined on the whole range. Seto (2006) considered a simple power-law spectrum of stochastic GW, analyzing the resolution of parameter estimation, but did not give an estimation of the spectrum. We adopt this as the third method to estimate the RGW spectrum by ML-estimation, as well as the three parameters (\(\alpha, \beta, \alpha_i\)) simultaneously in a Bayesian approach. For all these three methods for a pair, we shall provide numerical simulations, demonstrating their feasibility.

The outline of the paper is as follows. In Section 2, we give a short review of the theoretical RGW spectrum. In Section 3 we compare briefly the sensitivity of three types of interferometers, and give a constraint on RGW by a single Michelson in space. In Section 4, we examine signals from a single by the ML method and show that it is not effective to estimate the RGW spectrum when the noise is dominantly large. Section 5 is about the cross-correlated, integrated output signals for a pair. In Section 6, we show that the integrated output signals from a pair can be used to estimate one parameter, but not the spectrum. In Section 7 we use the ensemble averaging method to estimate the spectrum directly. In Section 8, we use the correlation method for un-integrated output signals to estimate the spectrum and parameters of RGW. Appendix A gives the derivation of the Fisher matrix for a pair.

2 RELIC GRAVITATIONAL WAVE

This section reviews the main properties of RGWs relevant to detection by LISA. RGW as the tensor metric perturbations of spacetime is generated during inflation and exists as a stochastic background of fluctuations in the Universe. It has an extremely broad spectrum, ranging from \(10^{-18}\) Hz to \(10^{11}\) Hz. In particular, it has a characteristic amplitude of \(10^{-22} \sim 10^{-24}\) around \(f \sim 10^{-3}\) Hz (see Fig. 1) and can serve as a target for LISA. The exact solution and corresponding analytical spectrum of RGW have been obtained (Zhang et al. 2005, 2006; Zhao & Zhang 2006b; Wang et al. 2016) that cover the whole course of expansion, from inflation, reheating, radiation, matter, to the present accelerating stage.

For a spatially flat Robertson-Walker spacetime, the metric with tensor perturbation is

\[
\begin{align*}
\text{ds}^2 &= a(\tau)^2[-d\tau^2 + (\delta_{ij} + h_{ij})dx^i dx^j],
\end{align*}
\]

where \(h_{ij}\) is the tensor perturbation and \(\tau\) is the conformal time. From inflation to the accelerating expansion, there are five stages, with each stage being described by a power-law scale factor \(a(\tau) \propto \tau^d\) where \(d\) is a constant (Zhang et al. 2005, 2006; Zhao & Zhang 2006b). The particularly interesting stage is inflation with

\[
\begin{align*}
a(\tau) &= l_0|\tau|^{1+\beta}, \quad -\infty < \tau \leq \tau_1,
\end{align*}
\]

where \(\beta\) is the expansion index. For the exact de Sitter, \(\beta = -2\), and for generic inflation models, \(\beta\) can deviate slightly from \(-2\) (Zhang et al. 2005, 2006; Zhao & Zhang 2006b). The present accelerating stage has

\[
\begin{align*}
a(\tau) &= l_H|\tau - \tau_0|^{-\gamma}, \quad \tau_E \leq \tau \leq \tau_H,
\end{align*}
\]

where \(\gamma = 2.018\) is taken for \(\Omega_\Lambda = 0.71\) (Wang et al. 2016). The normalization is taken as \(a(\tau_H) = l_H = \gamma/H_0\), where \(H_0\) is the present Hubble constant.

The tensorial perturbation \(h_{ij}\) as a quantum field is decomposed into Fourier modes,

\[
\begin{align*}
h_{ij}(x, \tau) &= \int \frac{d^3k}{(2\pi)^3/2} \sum_{A=++} \epsilon_{ij}^A(k) \\
&\quad \times \left[ a_k^A h_k^A(\tau)e^{ik \cdot x} + a_k^{-A} h_{k}^{-A}(\tau)e^{-ik \cdot x} \right], \\
&\quad k = \hat{k} \hat{k},
\end{align*}
\]

where \(a_k^A\) and \(a_k^{-A}\) are the annihilation and creation operators respectively of a graviton with wavevector \(k\) and polarization \(A\), satisfying the canonical commutation relation

\[
\begin{align*}
[a_k^A, a_{k'}^{-A}] = \delta_{AA'} \delta^3(k - k').
\end{align*}
\]

Two polarization tensors satisfy

\[
\begin{align*}
\epsilon_{ij}^A(k)\delta_{ij} &= 0, \quad \epsilon_{ij}^A(k)k^i = 0, \quad \epsilon_{ij}^A(k)\epsilon_{ij}^{A'}(k) = 2\delta_{AA'},
\end{align*}
\]

and can be taken as

\[
\begin{align*}
\epsilon_{ij}^+(k) &= (l_i l_j - m_im_j), \quad \epsilon_{ij}^x(k) = (l_i m_j + m_il_j),
\end{align*}
\]
where \( l, m \) are mutually orthogonal unit vectors normal to \( k \). In fact, as an observed quantity for LISA, RGW can be also treated as a classical, stochastic field

\[
h_{ij}(\tau, x) = \int \frac{d^3k}{(2\pi)^{3/2}} \sum_{A=+,-} c_{ij}^A(k) h^A_k(\tau) e^{ik \cdot x} ,
\]

(7)

where the \( k \)-mode \( h^A_k \) is stochastic, independent of other modes. The physical frequency at present is related to the conformal wavenumber via \( f = c k / 2 \pi a(\tau_H) \) (Zhang et al. 2005, 2006; Zhao & Zhang 2006b). For RGW, the two polarization modes \( h_+^k \) and \( h_\times^k \) are assumed to be independent and statistically equivalent, so that the superscript +, \( \times \) can be dropped, and the wave equation is

\[
h''_k(\tau) + 2 \frac{a'(\tau)}{a(\tau)} h'_k(\tau) + k^2 h_k(\tau) = 0 .
\]

(8)

The quantum state during inflation is taken to be \( |0\rangle \) such that

\[
a^k_0 |0\rangle = 0 ,
\]

(9)
i.e., only the vacuum fluctuations of RGW are present during inflation, and the solution of RGW is

\[
h_k(\tau) = \sqrt{\frac{32 \pi G}{a(\tau)}} \frac{\pi k|\tau|}{2k} \left( -i e^{-i \pi \beta / 2} \right) H^{(2)}_{\beta + \frac{1}{2}}(k|\tau|) , \quad -\infty < \tau \leq \tau_1 ,
\]

(10)

where is the positive-frequency mode \( h_k \rightarrow \frac{32 \pi G}{a(\tau)} \frac{1}{\sqrt{k}} e^{-ik \tau} \) and gives a zero point energy \( \frac{1}{2} h_\omega \) in each \( k \)-mode and each polarization in the high frequency limit. The wave equation, Equation (8), has been solved also for other subsequent stages, i.e., reheating, radiation dominant, matter dominant and accelerating. The solution of Equation (8) is simply a combination of two Hankel functions, \( \tau H^{d-1/2} H^{(1)(1)}_{d-1/2} \) and \( \tau^{d-1/2} H^{(2)}_{d+1/2} \). By continuously joining these stages, the full analytical solution \( h_k(\tau) \) has been obtained, which covers the whole course of evolution, in particular, for the present stage of acceleration, it is given by (Wang et al. 2016)

\[
h_k(\tau) = \sqrt{\frac{32 \pi G}{a(\tau)}} \frac{\pi s}{2k} \left[ e^{-i \pi \gamma / 2} \beta_k H^{(1)}_{\gamma - \frac{1}{2}}(s) + e^{i \pi \gamma / 2} \alpha_k H^{(2)}_{\gamma - \frac{1}{2}}(s) \right] , \quad \tau_E < \tau \leq \tau_H ,
\]

(11)

where \( s \equiv k(\tau - \tau_0) \) and the coefficients \( \beta_k, \alpha_k \) are Bogoliubov coefficients (Parker & Toms 2009; Birrell & Davies 1982) satisfying \( |\alpha_k|^2 - |\beta_k|^2 = 1 \). \( |\beta_k|^2 \) is the number of gravitons at the present stage, and the expressions \( \beta_k, \alpha_k \) are explicitly given by Wang et al. (2016). The frequency range of space-borne interferometers is much higher than the Hubble frequency \( H_0 \approx 2 \times 10^{-18} \text{ Hz} \), so that Equation (11) for these modes becomes

\[
h_k(\tau) \approx \sqrt{\frac{32 \pi G}{a(\tau)}} \frac{1}{\sqrt{k}} e^{-ik \tau} \quad \text{for} \; k \gg 1/|\tau| .
\]

(12)

Hence, for space-borne interferometers, RGW is practically a superposition of stochastic plane waves.

The auto-correlation function of RGW is defined by the following expected value

\[
\langle 0 | h^{ij}(x, \tau) h^{ij}(x, \tau) | 0 \rangle = \frac{1}{(2\pi)^3} \int d^3k \langle |h_k|^2 \rangle ,
\]

(13)

one reads off the power spectrum

\[
\Delta^2_k(k, \tau) = \frac{k^3}{2 \pi^2} \langle |h_k(\tau)|^2 \rangle ,
\]

(15)

which is dimensionless. We also use a notation \( h(f, \tau_H) \equiv \sqrt{\Delta^2_k(k, \tau_H)} \). In the literature on GW detection, the characteristic amplitude

\[
h_{\chi}(f) = h_k(\tau_H) / 2 \sqrt{f}
\]

(16)

is often used (Maggiore 2000; Zhang et al. 2010), which has dimension \( \text{Hz}^{-1/2} \). The definition (15) holds for any time \( \tau \), from inflation to the accelerating stage.

Figure 1 shows the evolution of the RGW spectrum from inflation to the present acceleration stage. Equivalently, one can also use the spectral energy density \( \Omega_g \equiv \rho_g / \rho_c \), where

\[
\rho_g = \frac{1}{32 \pi G a^2} \langle 0 | h^{ij}_0 h^{ij}_0 | 0 \rangle
\]
is the energy density of RGW (Brill & Hartle 1964; Su & Zhang 2012; Weinberg 1972; Wang et al. 2016) and \( \rho_c = 3H^2_0/8\pi G \) is the critical density. The spectral energy density \( \Omega_g(f) \) is defined by \( \Omega_g \equiv \int \Omega_g(f)df/f \), and given by

\[
\Omega_g(f) = \frac{\pi^2}{3} h^2(f, \tau_H) \left( \frac{f}{H_0} \right)^2,
\]

which holds for all wavelengths shorter than the horizon.

From the spectrum (15) during inflation, the analytic expressions of spectral and running spectral indices have been obtained (Wang et al. 2016)

\[
n_t \equiv \frac{d \ln \Delta_t^2}{d \ln k} \approx 2 \beta + 4 - \frac{2}{2 \beta + 3} x^2,
\]

\[
\alpha_t \equiv \frac{d^2 \ln \Delta_t^2}{d (\ln k)^2} \approx -\frac{4}{2 \beta + 3} x^2
\]

at \( x \approx |k\tau| \ll 1 \), i.e., at far outside horizon, both related to the inflation index \( \beta \). In the limit \( k \to 0 \), one has the default values

\[
n_t = 2 \beta + 4, \quad \alpha_t = 0
\]

which hold for the inflation models with \( a(\tau) \propto |\tau|^{1+\beta} \). It is incorrect to use \( n_t \) and \( \alpha_t \) evaluated at the horizon-crossing \( |k\tau| = 1 \) (Kosowsky & Turner 1995). With these definitions, the primordial spectrum in the limit \( k \to 0 \) is written as

\[
\Delta_t(k) = \Delta_R r^{1/2} \left( \frac{k}{k_0} \right)^{1/n_t + 1/\alpha_t} \ln(\frac{k}{k_0}),
\]

where \( k_0 \) is a pivot conformal wavenumber corresponding to a physical wavenumber \( k_0/a(\tau_H) = 0.002 \text{ Mpc}^{-1} \), \( \Delta_R \) is the value of curvature perturbation determined by observations \( \Delta^2_R = (2.464 \pm 0.072) \times 10^{-9} \) and \( r \equiv \Delta^2_t(k_0)/\Delta^2_R(k_0) \) is the tensor-scalar ratio, with \( r < 0.1 \) by CMB observations (Planck Collaboration et al. 2016b,c,d).

The primordial spectrum (19) describes the upper curve (red) during inflation in Figure 1. The present spectrum \( \Delta_t(f, \tau_H) \) and the primordial spectrum \( \Delta_t(k) \) are overlapped at very low frequencies \( f < 10^{-18} \text{ Hz} \), with both being \( \propto r^{1/2} \) as in (19). At \( f > 10^{11} \text{ Hz} \), \( \Delta_t(f, \tau_H) \) rises up and has an ultraviolet (UV) divergence, due to vacuum fluctuations. In Wang et al. (2016), the UV divergence has been adiabatically regularized. Higher values of \( (r, n_t, \alpha_t) \) give rise to higher amplitude of RGW. In particular, a slight increase in \( \alpha_t \) will enhance greatly the amplitude of RGW in the relevant band. In this paper, we take \( (r, n_t, \alpha_t) \) as the major parameters of RGW.

![Fig. 2](image_url)

**Fig. 2** The three spacecraft are located at points 1, 2 and 3, and the vectors \( a, b \) and \( c \) label the directions of the three arms.

### 3 Sensitivity of One Interferometer and RGW Detection

We briefly review detection of RGW by a single interferometer, which has been studied before and will be used in this paper later. Figure 2 shows three identical spacecrafts that are placed in space, forming an equilateral triangle. The three arms are of equal length, taken to be \( L = 5 \times 10^9 \text{ m} \) by the original design of LISA (Bender et al. 1998) when no GW is passing by. This value is taken as an example in our paper. In recent years, the designed arm-length has been modified to be \( L = 1 \times 10^9 \text{ m} \) (Amaro-Seoane et al. 2012) or \( L = 2.5 \times 10^9 \text{ m} \) (Amaro-Seoane et al. 2017). Recently-proposed projects, like Tianqin and Taiji, also will have \( L \) around this value. ASTROD has a longer value of \( L = 260 \times 10^9 \text{ m} \) (Ni et al. 2004; Ni 2008, 2010, 2013). Spacecraft 1 can shoot laser beams, which are phase-locked, regenerated with the same phase at spacecrafts 2 and 3, and then sent back (Bender et al. 1998; Amaro-Seoane et al. 2017). This forms one interferometer. In the presence of a GW, the arm lengths and phases of the beams will fluctuate. Combining the optical paths will produce different interferometers (Vallisneri et al. 2008). Here we only discuss three combinations, the Michelson, the Sagnac and the symmetrized Sagnac (Cornish 2001; Cornish & Larson 2001). To focus on the main issue of spectral estimation for RGW, we do not consider the spacecraft orbital effects, Shapiro delay, etc., caused by the Newtonian potential of the solar system (Rubbo et al. 2004; Cornish & Rubbo 2003; Tinto & Armstrong 1999; Tinto et al. 2004, 2005).
3.1 The Response Tensors for Three Kinds of Interferometers

First, the Michelson interferometer (Cornish 2001; Cornish & Larson 2001; Maggiore 2000; Cutler 1998) is considered. The optical path difference is proportional to the strain

\[ h_\alpha = \frac{1}{2L} [l_{12} + l_{21} - l_{13} - l_{31}] , \quad (20) \]

where \( l_{12} \) is the optical path of a photon emitted by spacecraft 1 traveling along arm 1-2, which has arrived at 2, \( l_{21} \) is the one reflected at 2 and back to 1, etc. This is the dimensionless strain measured by the interferometer, also called the output response.

Next, the Sagnac interferometer (Shaddock 2004) is described. One optical path is 1–2–3–1, and the other is 1–3–2–1. The strain is proportional to their difference

\[ h_{\text{cos}1} = \frac{1}{3L} [l_{13} + l_{32} + l_{21} - l_{12} - l_{23} - l_{31}] , \quad (21) \]

where the subscript “1” in \( h_{\text{cos}1} \) refers to vertex 1. In a similar fashion, one can get the output responses at vertex 2 and vertex 3. Last, the symmetrized Sagnac interferometer is examined. Its output response is defined as the average of three Sagnac signals for three vertices, 1, 2 and 3, respectively (Armstrong et al. 1999; Prince et al. 2002; Robinson et al. 2008)

\[ h_{\text{cos}3} = \frac{1}{3} (h_{\text{cos}1} + h_{\text{cos}2} + h_{\text{cos}3}) . \quad (22) \]

We shall study the ability to detect RGW with these three kinds of interferometers when implemented in space.

Now we consider the responses of the three kinds of interferometers to a GW. Let a plane GW, with frequency \( f \) from a direction \( \hat{\Omega} \), pass through the detector located at \( r \) and at time \( t \). The GW is denoted by \( h(\hat{\Omega}, f, t, r) \) as a tensor. The output response of an interferometer is a product

\[ h_\alpha(\hat{\Omega}, f, t, r) = D(\hat{\Omega}, f) : h(\hat{\Omega}, f, t, r) , \quad (23) \]

where \( D(\hat{\Omega}, f) \) is the response tensor, depending on the orientation and geometry of LISA and the operating frequency \( f \). For the Michelson, the response tensor is (Cornish & Larson 2001; Cornish 2001),

\[ D_m(\hat{\Omega}, f) = \frac{1}{2} ((a \otimes a) T_m(a \cdot \hat{\Omega}, f) - (c \otimes c) T_m(-c \cdot \hat{\Omega}, f)) , \quad (24) \]

with \( a \) and \( c \) being the vectors of arms shown in Figure 2, and \( T_m \) being the single-arm transfer function

\[ T_m(a \cdot \hat{\Omega}, f) = \frac{1}{2} \left[ \sin \left( \frac{f}{2f_s} (1 - a \cdot \hat{\Omega}) \right) \right. \]

\[ \exp \left( -i \frac{f}{2f_s} (3 + a \cdot \hat{\Omega}) \right) + \sin \left( \frac{f}{2f_s} (1 + a \cdot \hat{\Omega}) \right) \]

\[ \exp \left( -i \frac{f}{2f_s} (1 + a \cdot \hat{\Omega}) \right) \right] , \quad (25) \]

where \( \sin(x) \equiv \frac{\sin x}{x} \), and \( f_s \equiv c/(2\pi L) \approx 0.0095 \text{ Hz} \) is the characteristic frequency of LISA. In expression (23), “\( : \)" denotes a tensor product defined by

\[ (a \otimes a) : e = a_i a_j e_{ij} , \]

where \( e_{ij} \) is the polarization tensor satisfying the conditions of (6). The output response (23) indicates how the interferometer transfers an incident GW into an output signal through a specific geometric setup. When passing, for the ground-based case of LIGO, the working frequency range is \((10^3 \sim 10^5) \text{ Hz} \) (Abbott et al. 2016a, b, 2017b), and \( f_s \approx 1.2 \times 10^4 \text{ Hz} \) for 4 km-long arms, so one can take the low frequency limit \( f \ll f_s \). \( T_m \approx 1 \), then Equation (24) reduces to that of Allen (1997) and Allen & Romano (1999).

The response tensor of the Sagnac is (Cornish 2001; Kudoh & Tsuruya 2005)

\[ D_s = \frac{1}{6} \left( (a \otimes a) T_a(f) + (b \otimes b) T_b(f) + (c \otimes c) T_c(f) \right) , \quad (26) \]

depending also on the vector arm \( b \) and the transfer functions

\[ T_a(f) = \sin \left( \frac{f}{2f_s} (1 + a \cdot \hat{\Omega}) \right) \exp \left( -i \frac{f}{2f_s} (1 + a \cdot \hat{\Omega}) \right) \]

\[ - \sin \left( \frac{f}{2f_s} (1 - a \cdot \hat{\Omega}) \right) \times \exp \left( -i \frac{f}{2f_s} (5 + a \cdot \hat{\Omega}) \right) , \quad (27) \]

\[ T_b(f) = \left[ \sin \left( \frac{f}{2f_s} (1 + b \cdot \hat{\Omega}) \right) - \sin \left( \frac{f}{2f_s} (1 - b \cdot \hat{\Omega}) \right) \right] \]

\[ \times \exp \left( -i \frac{f}{2f_s} (3 + a \cdot \hat{\Omega} - c \cdot \hat{\Omega}) \right) , \quad (28) \]

\[ T_c(f) = \sin \left( \frac{f}{2f_s} (1 + c \cdot \hat{\Omega}) \right) \exp \left( -i \frac{f}{2f_s} (5 - c \cdot \hat{\Omega}) \right) \]

\[ - \sin \left( \frac{f}{2f_s} (1 - c \cdot \hat{\Omega}) \right) \times \exp \left( -i \frac{f}{2f_s} (1 - c \cdot \hat{\Omega}) \right) . \quad (29) \]

In expression (23), the output response of an interferometer is a product

\[ h_\alpha(\hat{\Omega}, f, t, r) = D(\hat{\Omega}, f) : h(\hat{\Omega}, f, t, r) , \quad (23) \]

where \( D(\hat{\Omega}, f) \) is the response tensor, depending on the orientation and geometry of LISA and the operating frequency \( f \). For the Michelson, the response tensor is (Cornish & Larson 2001; Cornish 2001),

\[ D_m(\hat{\Omega}, f) = \frac{1}{2} ((a \otimes a) T_m(a \cdot \hat{\Omega}, f) - (c \otimes c) T_m(-c \cdot \hat{\Omega}, f)) , \quad (24) \]
Notice that a factor of $\frac{1}{2}$ is missed in the exponent function in (9) of Cornish (2001). The response tensor of the symmetrized Sagnac is (Cornish 2001; Kudoh & Taruya 2005)

$$D_{ss}(\hat{\Omega}, f) = \frac{1}{6} \left( (a \otimes a) T_{ss}(a \cdot \hat{\Omega}, f) + (b \otimes b) T_{ss}(b \cdot \hat{\Omega}, f) + (c \otimes c) T_{ss}(c \cdot \hat{\Omega}, f) \right),$$

(30)

and

$$T_{ss}(u \cdot \hat{\Omega}, f) = \left( 1 + 2 \cos \frac{f}{f_s} \right) \exp \left( -i \frac{f}{2f_s} (3 + u \cdot \hat{\Omega}) \right) \times \left[ \text{sinc} \left( \frac{f}{2f_s} (1 + u \cdot \hat{\Omega}) \right) - \text{sinc} \left( \frac{f}{2f_s} (1 - u \cdot \hat{\Omega}) \right) \right].$$

(31)

Thus, for a given incident GW, these three kinds of interferometers will yield different output responses due to response tensors.

The above output response (23) applies to GW emitted by a fixed source far away from the detector. The matched filter technique is usually used to search for a GW embedded in the noise (Thorne 1987; Cutler & Flanagan 1994; Finn 1992).

### 3.2 The Output Response to RGW and the Transfer Function

RGW as a stochastic background contains a mixture of all independent $k$ modes of plane waves. In regards to its detection by space-based interferometers, RGW in (7) can be also written as a sum over frequencies and directions (Allen 1997; Allen & Romano 1999; Cornish 2001)

$$h_{ij}(t, \vec{x}) = \sum_{A, A', +} \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} \tilde{h}_{A}(f, \hat{\Omega}) \tilde{h}_{A'}^{*}(f, \hat{\Omega'}) e^{-i2\pi ft} e^{i2\pi f\hat{\Omega} \cdot \vec{r} / c},$$

(32)

where $\hat{\Omega} = k / k$, $\vec{r} = a(\tau_H) \vec{x}$ and

$$\tilde{h}_{A}(f, \hat{\Omega}) = \frac{2\pi a(\tau_H) / c}{(2\pi)^{3/2}} k^2 h^{A}(\tau) e^{i2\pi ft}$$

(33)

with the mode $h^{A}(\tau)$ given in Equation (7). By its stochastic nature, each mode of frequency $f$ and in direction $\hat{\Omega}$ is random. Statistically, RGW can be assumed to be a Gaussian random process, and the ensemble averages are given by (Allen 1997; Allen & Romano 1999)

$$\langle \tilde{h}_{A}(f, \hat{\Omega}) \rangle = 0,$$

(34)

$$\langle \tilde{h}_{A}(f, \hat{\Omega}) \tilde{h}_{A'}^{*}(f', \hat{\Omega'}) \rangle = \frac{1}{2} \delta(f - f') \delta^2(\hat{\Omega}, \hat{\Omega}') \frac{\delta A A'}{4\pi} S_{h}(f),$$

(35)

where $\delta^2(\hat{\Omega}, \hat{\Omega}') = \delta(\phi - \phi')\delta(c \theta - c \theta')$, and $S_{h}(f)$ is the spectral density, also referred to as the spectrum, in the unit of Hz$^{-1}$ satisfying $S_{h}(f) = S_{h}(-f)$. The factor $\frac{1}{2}$ is introduced considering that the variable $f$ of $S_{h}$ ranges between $-\infty$ and $+\infty$. Equations (34) and (35) specify fully the statistical properties of RGW. The normalization of $S_{h}(f)$ is chosen such that

$$\langle \tilde{h}_{A}^{*}(f) \tilde{h}_{A'}(f') \rangle = \frac{1}{2} \delta(f - f') \delta A A' S_{h}(f).$$

(36)

From Equations (32) and (35), the auto-correlation function of RGW can be written as

$$\langle h_{ij}(t) \tilde{h}_{ij}^{*}(t) \rangle = 2 \int_{-\infty}^{+\infty} df S_{h}(f) = 4 \int_{0}^{\infty} d(\log f) f S_{h}(f),$$

(37)

so that the spectral density is related to the characteristic amplitude (16) as the following

$$S_{h}(f) = h_{c}^{2}(f) = \frac{3H_{0}^{2} \Omega_{g}(f)}{4\pi^{2} f^{3}}.$$

(38)

where $\Omega_{g}(f)$ is the spectral energy density (17).

Let us consider the output response of an interferometer to RGW. Substituting $h_{ij}$ of Equation (32) into Equation (23) yields the output response

$$h_{c}(t) = \sum_{A, A', +} \int_{-\infty}^{+\infty} df \int_{-\infty}^{\infty} \tilde{h}_{A}(f, \hat{\Omega}) e^{-i2\pi ft} e^{i2\pi f\hat{\Omega} \cdot \vec{r} / c} D(\hat{\Omega}, f) : e^{A}(\hat{\Omega}),$$

(39)

which is valid for all three kinds of interferometers with their respective response tensor $D$. Since the RGW background is isotropic in the Universe, we are free to take the detector location at $\vec{r} = 0$ (Cornish & Larson 2001), so that Equation (39) becomes

$$h_{c}(t) = \sum_{A, A', +} \int_{-\infty}^{+\infty} df \int_{-\infty}^{\infty} \tilde{h}_{A}(f, \hat{\Omega}) e^{-i2\pi ft} D(\hat{\Omega}, f) : e^{A}(\hat{\Omega}),$$

(40)
which is a summation over all frequencies, directions and polarizations, in contrast to a GW from a fixed source. Its Fourier transform is

\[ \tilde{h}_o(f) = \sum_A \int d\hat{\Omega} \tilde{h}_A(f, \hat{\Omega}) D(\hat{\Omega}, f) : e^A(\hat{\Omega}) . \tag{41} \]

The ensemble averages (34) and (35) lead to

\[ \langle \tilde{h}_o(f) \rangle = 0 , \]
\[ \langle \tilde{h}_o(f) \tilde{h}_o(f') \rangle = \frac{1}{2} \delta(f - f') S_h(f) R(f) . \tag{42} \]

The auto-correlation of output response is

\[ \langle h^2_o(t) \rangle = \int_{-\infty}^{+\infty} df \frac{1}{2} S_h(f) R(f) = \int_{-\infty}^{+\infty} df S_h(f) R(f) , \tag{43} \]

where the transfer function

\[ R(f) = \int_4 d\hat{\Omega} \sum_{A=x,+,y} F^{A*}(\hat{\Omega}, f) F^A(\hat{\Omega}, f) , \tag{44} \]

is a sum over all directions and polarizations, and the detector response function

\[ F^A(\hat{\Omega}, f) = D(\hat{\Omega}, f) : e^A(\hat{\Omega}) . \tag{45} \]

\[ R(f) \] is determined by the geometry of the interferometer and transfers the stochastic RGW signal into the output signal. A greater value of \( R(f) \) means a stronger ability to transfer RGW into the output signal. Formula (44) applies to the Michelson, Sagnac and symmetrized Sagnac interferometers. Using \( D_m, D_s, D_sss \) of (24), (26) and (30) yields the transfer functions \( R_m(f), R_s(f) \) and \( R_{sss}(f) \), respectively. They are plotted in the top of Figure 3. For the Michelson, \( R(f) \approx 0.3 \) in low frequencies, which is much greater than the corresponding value for the Sagnac and symmetrized Sagnac, so the Michelson has a stronger ability to transfer incident RGW into the output signals.

3.3 The Sensitivity of one Interferometer and Detection of RGW

Including the noise, the total output signal of an interferometer is a sum

\[ s(t) = h_o(t) + n(t) , \tag{46} \]

where \( h_o(t) \) is the output response of (40) and \( n(t) \) is a Gaussian noise signal with a zero mean \( \langle n(t) \rangle = 0 \), uncorrelated to \( h_o \). Define

\[ \langle n(t)n(t') \rangle = \frac{1}{2} \int_{-\infty}^{+\infty} df e^{i2\pi f(t-t')} S_n(f) , \tag{47} \]

where \( S_n(f) \) is the noise spectral density. It satisfies

\[ \langle n^2(t) \rangle = \frac{1}{2} \int_{-\infty}^{+\infty} df S_n(f) = \int_{-\infty}^{+\infty} df S_n(f) . \tag{48} \]

The noise in the frequency domain can be equivalently specified by

\[ \langle \tilde{n}(f) \rangle = 0 , \]
\[ \langle \tilde{n}^*(f)\tilde{n}(f') \rangle = \frac{1}{2} \delta(f - f') S_n(f) . \tag{49} \]

There are two major kinds of noise (Bender et al. 1998; Bender 2003; Cornish 2001). The first kind is called the optical-path noise, which includes shot noise, beam pointing instabilities, thermal vibrations, etc. Among these, shot noise is the most important and its noise spectral density is given by (Cornish 2001)

\[ S_s(f) = \frac{1.21 \times 10^{-22} \text{m}^2 \text{Hz}^{-1}}{(5 \times 10^9 \text{m})^2} \]
\[ = 4.84 \times 10^{-42} \text{Hz}^{-1} . \tag{50} \]

The other kind of noise is the acceleration noise with a spectral density

\[ S_a(f) = \frac{9 \times 10^{-30} \text{m}^2 \text{s}^{-4} \text{Hz}^{-1}}{(5 \times 10^9 \text{m})^2} \]
\[ = 2.31 \times 10^{-40} \left( \frac{\text{mHz}}{f} \right)^4 \text{Hz}^{-1} . \tag{51} \]

From these follow the noise spectral density of the Michelson

\[ S_m^m(f) = 8S_a(f) \left( 1 + \cos^2 \left( \frac{f}{f_s} \right) \right) + 4S_s(f) , \tag{52} \]

the noise spectral density of the Sagnac

\[ S_n^s(f) = 6S_s(f) + 8 \left( \sin^2 \frac{3f}{2f_s} + 2 \sin^2 \frac{f}{2f_s} \right) S_a(f) , \tag{53} \]

and the noise spectral density of the symmetrized Sagnac

\[ S_n^{ss}(f) = \frac{2}{3} \left( 1 + 2 \cos \left( \frac{f}{f_s} \right) \right)^2 \]
\[ \times \left( S_s(f) + 4 \sin^2 \left( \frac{f}{2f_s} \right) S_a(f) \right) . \tag{54} \]

These are plotted in the bottom of Figure 3. It is seen that the noise of the Michelson is larger than those of Sagnac and symmetrized Sagnac, and the symmetrized Sagnac has the least noise. Higher symmetries in the optical path designs of Sagnac and symmetrized Sagnac cancel more noise. The symmetric Sagnac has a transfer function several orders of magnitude lower than Michelson of around \( 10^{-3} \) Hz, and can be used to monitor noise level in practice (Tinto et al. 2001; Cornish 2001).
To detect RGW by signals from one interferometer, one considers the auto-correlation of the total output signal
\[ \langle s^2(t) \rangle = \langle h_n^2(t) \rangle + \langle n^2(t) \rangle = \int_0^{\infty} df S(f), \quad (55) \]
where Equations (43) and (48) are used, and the total spectral density
\[ S(f) \equiv S_h(f) R(f) + S_n(f). \quad (56) \]
Equation (55) is equivalently written in the frequency domain
\[ \langle \tilde{s}^*(f) \tilde{s}(f') \rangle = \frac{1}{2} \delta(f - f') S(f). \quad (57) \]
Since both the RGW signal and noise occur in Equation (56), the SNR for a single interferometer which is denoted as SNR\_1 can be naturally defined as
\[ \text{SNR}_1 \equiv \frac{h_c(f)}{h(f)}, \quad (58) \]
where \( h_c(f) \) is related to \( S_h(f) \) by Equation (38), and the sensitivity is introduced by
\[ \tilde{h}(f) \equiv \frac{1}{\sqrt{\mathcal{R}(f)}}, \quad (59) \]
which reflects the detection capability of one interferometer. A smaller \( \tilde{h}(f) \) indicates a better sensitivity, which requires a lower \( S_n \) and a greater \( \mathcal{R} \). Figure 4 shows the sensitivity curves of three interferometers, which is similar to the result of Cornish (2001). It is seen that the Michelson has the best sensitivity level, \( \tilde{h}(f) \sim 10^{-20} \text{Hz}^{-1/2} \) around \( f_\star = c/(2\pi L) \approx 10^{-2} \text{Hz} \) for the arm-length \( L = 5 \times 10^9 \text{m} \). This is because the transfer function \( \mathcal{R}(f) \) of the Michelson is greatest, giving rise to a lowest value for \( \tilde{h}(f) \), even though its \( S_n(f) \) is slightly higher than the other two. Therefore, we shall use the Michelson in the subsequent sections. As a preliminary criterion, a single interferometer will detect RGW when \( \text{SNR}_1 > 1 \), i.e.,
\[ h_c(f) > \tilde{h}(f). \quad (60) \]
This criterion was used to constrain the RGW parameters from the data of LIGO S5 (Abbott et al. 2009; Zhang et al. 2010). We plot \( h_c(f) \) and \( \tilde{h}(f) \) in Figure 5 for a single interferometer. When the data of space-borne interferometers are available in the future, Equation (60) will put a constraint on the parameters. The interferometer will be able to detect RGW of \( \alpha_t > 0.016 \) at fixed \( r = 0.1 \) and \( \beta = -2.016 \).

4 ESTIMATION OF RGW SPECTRUM BY ONE INTERFEROMETER

Now we try to determine the RGW spectrum from the output signals of one interferometer. This is a typical estimation problem of statistical signals, which can be studied by statistical methods. From the view of statistics, RGW and CMB anisotropies share some similar properties; both of them form a stochastic background in the Universe and can be modeled by a Gaussian random field (Gorski et al. 1994; Jungman et al. 1996; Tegmark et al. 1997; Oh et al. 1999; Hinshaw et al. 2003).

The time series of the output signal (46) can be put into the Fourier form in frequency domain
\[ \tilde{s}(f) = \tilde{h}_o(f) + \tilde{n}(f). \]
For practical computation, the data set can be divided into the following sample vector
\[ \tilde{\mathbf{s}} = [\tilde{s}(f_1),...,\tilde{s}(f_N)], \quad (61) \]
with \( f_{i+1} - f_i = \Delta f, i = 1,2,\cdots,N \), where \( N \) is a sufficiently large number. Since both \( \tilde{h}_o(f) \) and \( \tilde{n}(f) \) are Gaussian and independent, \( \tilde{s}(f) \) is a Gaussian random variable and \( \tilde{\mathbf{s}} \) consists of \( N \) statistically independent Gaussian data points, having zero mean
\[ \langle \tilde{s}(f_i) \rangle = 0. \quad (62) \]
The covariance matrix is (57), which is written in the discrete form
\[ \Sigma_{ij} = \delta_{ij} \frac{1}{2\Delta f} S(f_i), \quad i,j = 1,2,\cdots,N \quad (63) \]
where the Dirac delta function \( \delta(f) \) has been replaced by its discrete form (Finn 1992).
\[ \delta(f_i - f_j) = \lim_{\Delta f \to 0} \frac{1}{\Delta f} \delta_{ij}. \quad (64) \]
The total spectral density in (56) is also written in the discrete form
\[ S(f_i) \equiv [S_h(f_i) \mathcal{R}(f_i) + S_n(f_i)]. \quad (65) \]
The inverse covariance matrix is
\[ (\Sigma^{-1})_{ij} = \frac{2\Delta f \delta_{ij}}{S(f_i)}, \quad (66) \]
depending on the RGW signal \( S_h \), the noise \( S_n \) and the transfer function \( \mathcal{R} \) of the interferometer. Note that here \( \Sigma_{ij} \) is diagonal since \( \tilde{s}(f_i) \) and \( \tilde{s}(f_j) \) are independent for
We present a pair of Michelson interferometers in space which has been studied in Cornish & Larson (2001); Cornish (2001). Based on the three spacecraft forming an equilateral triangle in Figure 2, we consider two configurations of a pair. Config. 1 consists of the three spacecraft, equipped with six detection equipment, forming two interferometers. Config. 2 consists of two triangles, equipped with six spacecraft, forming two interferometers in space (Cornish & Larson 2001). The second configuration in one triangle, but with each spacecraft carrying two sets of independent detection equipment. Then two independent Michelson interferometers form (Cutler 1998): the first having the point 1 as the vertex, and the second having the point 2 as the vertex, differing from the first by a rotation of 120°, as in Figure 2. This configuration is economically favored, however, a possible problem is that the equipment on one craft may have dependent noise. For simplicity we assume that the noises are independent by better setup in the design. Config. 2 consists of two triangles, equipped with six spacecraft, forming two interferometers in space (Cornish & Larson 2001). The second configuration is an equilateral triangle, as in Figure 6. This configuration can ensure independent noise since the six spacecraft have been located far away from each other in space. For both configurations, we assume that the noises from the two interferometers are independent of each other, and independent of RGW.

GW signals are correlated in the two interferometers. By cross-correlating the output data of the pair, the detection capability will be enhanced. Consider the output signals from a pair of two interferometers

\[ s_1(t) = h_1(t) + n_1(t) , \]
\[ s_2(t) = h_2(t) + n_2(t) , \]

in which each is similar to Equation (46). It is assumed that

\[ \langle n_i(t)n_j(t) \rangle = 0 , \quad \langle n_1(t)n_2(t) \rangle = 0 , \]
and

\[ \langle n_1^2(t) \rangle = \int_0^\infty df S_{n1}(f) , \]
\[ \langle n_2^2(t) \rangle = \int_0^\infty df S_{n2}(f) , \]

(80)

where \( S_{n1}, S_{n2} \) are the noise spectral densities of the two interferometers as in (48), and specified as in Equation (52) for the Michelson. When the two interferometers are identical, one can take \( S_{n1} \approx S_{n2} \).

Using the output response of (39) for each interferometer and formula (35), the ensemble average of the correlation of two output responses is

\[ \langle h_1(t)h_2(t') \rangle = \frac{1}{2} \int_{-\infty}^{+\infty} df S_h(f) \times \int \frac{d\tilde{\Omega}}{4\pi} \sum_{A=x,+} F_1^{A+}(\tilde{\Omega}, f) F_2^{A+}(\tilde{\Omega}, f) \times e^{-i2\pi f\tilde{\Omega}(r_1-r_2)} e^{i2\pi f(t-t')} \]
\[ = \int_0^\infty df S_h(f) R_{12}(f) e^{i2\pi f(t-t')} , \]

(81)

similarly, by (41) and (35),

\[ \langle \hat{h}_1(f)\hat{h}_2(f') \rangle = \frac{1}{2} \delta(f-f')S_h(f)R_{12}(f) , \]

(82)

where the transfer function \( R_{12}(f) \) is

\[ R_{12}(f) \equiv \int \frac{d\tilde{\Omega}}{4\pi} \sum_{A=x,+} F_1^{A+}(\tilde{\Omega}, f) \times F_2^{A+}(\tilde{\Omega}, f) e^{-i2\pi f\tilde{\Omega}(r_1-r_2)} , \]

(83)

where \( r_1 \) and \( r_2 \) represent the respective positions of the vertices of the two interferometers. A higher value of \( R_{12}(f) \) means a better capability to transfer incoming RGW into signals from the detector.

In Figure 7 we plot the transfer function of a single and of a pair with conf.1 and conf.2. One introduces the overlapping reduction function \( \gamma(f) \) (Flanagan 1993; Allen 1997; Allen & Romano 1999) by normalizing \( R_{12}(f) \) as the following

\[ \gamma(f) \equiv \frac{5}{\sin^2 \beta_0} R_{12}(f) = \frac{20}{3} R_{12}(f) \quad \text{for conf.1} , \]
\[ \gamma(f) \equiv \frac{5}{2\sin^2 \beta_0} R_{12}(f) = \frac{10}{3} R_{12}(f) \quad \text{for conf.2} , \]

(84)
(85)

where \( \beta_0 = \pi/3 \) is the angle between arms of one interferometer, \( \sin^2 \beta_0 = 3/4 \). Clearly, \( \gamma(f) \) depends on the geometry of the pair and transfers the incident RGW from all the directions (by integration over angle \( \tilde{\Omega} \)) into the output signal. We compute \( \gamma(f) \) numerically and plot it in Figure 8 for configs. 1 (top) and 2 (bottom). At high frequencies, \( \gamma(f) \) oscillates around zero. At low frequencies, \( \gamma(f) \to 1 \) for both configurations. To a high accuracy, it can be fitted by the following formula:

\[ \gamma(f) = \begin{cases} 
1 - 0.811508 \left( \frac{f}{f_*} \right)^2 + 0.241292 \left( \frac{f}{f_*} \right)^4 - 0.0374118 \left( \frac{f}{f_*} \right)^6, & f < f_* \\
0.43636 + 2.12337 \left( \frac{f}{f_*} \right) - 4.00143 \left( \frac{f}{f_*} \right)^2 + 2.37321 \left( \frac{f}{f_*} \right)^3 - 0.588745 \left( \frac{f}{f_*} \right)^4 + 0.0528759 \left( \frac{f}{f_*} \right)^5, & f_* \leq f < 3.3f_* \\
0.000311254 - 0.11762 e^{-0.37176 f/f_*} \left[ 0.0849126 - \sin \left( 2.89649 - 2.40829 \left( \frac{f}{f_*} \right) e^{0.0097865 f/f_*} \right) \right], & 3.3f_* \leq f 
\end{cases} \]

(86)
for config. 1, and

\[
\gamma(f) = \begin{cases} 
1 - \frac{383}{504} \left( \frac{f}{f_*} \right)^2 + \frac{383}{588} \left( \frac{f}{f_*} \right)^4 - \frac{5414989}{14300480} \left( \frac{f}{f_*} \right)^6, & f < f_* \\
0.629524 + 1.52435 \left( \frac{f}{f_*} \right) - 3.23303 \left( \frac{f}{f_*} \right)^2 + 1.96633 \left( \frac{f}{f_*} \right)^3 - 0.496608 \left( \frac{f}{f_*} \right)^4 + 0.0454382 \left( \frac{f}{f_*} \right)^5, & f_* \leq f < 3.3f_* \\
0.000190192 - 0.157835 \times 10^{-5} e^{-0.570049f/f_*}\left[ 0.264226 - \sin \left( 1.66131 - 2.01502 \left( \frac{f}{f_*} \right) e^{0.03493779f/f_*} \right) \right], & 3.3f_* \leq f
\end{cases}
\]

for config. 2. \( \gamma(f) \) of configs. 1 and 2 looks the same except around \( f \approx (2 \sim 8) \times 10^{-2} \) Hz. Formula (87) at \( f < f_* \) agrees with that in Cornish & Larson (2001), which does not list the expression for the part \( f > f_* \). See also Ungarelli & Vecchio (2001a). In the following we shall mainly use config. 2 for demonstration.

### 5.2 SNR for a Pair Case for LISA

To suppress the noise of a pair, one defines the cross-correlated, integrated signal of \( s_1(t) \) and \( s_2(t') \) as the following (Allen 1997; Allen & Romano 1999)

\[
C = \int_{-T/2}^{T/2} dt \int_{-\infty}^{+\infty} dt' s_1(t)s_2(t')Q(t-t') = \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_T(f-f')\tilde{s}_1(f)\tilde{s}_2(f')\tilde{Q}(f'),
\]

where \( T \) is the observation time, \( \tilde{s}(t) = \tilde{s}(-t) \) is the Fourier transform of \( s(t) \) and \( \tilde{s}_i(f) = \tilde{s}_i(-f) \), \( Q(t-t') \) is a filter function to be determined by maximizing SNR\(_{12} \) (SNR for the pair). Its Fourier transform is \( \tilde{Q}(f) = \tilde{Q}(-f) \), and

\[
\delta_T(f) = \int_{-T/2}^{T/2} dt e^{-i\pi f t} = \frac{\sin(\pi f T)}{\pi f}.
\]

is the finite-time Dirac delta function. For a finite \( T \), one has \( \delta_T(0) = T \), and in the limit \( T \rightarrow \infty \), \( \delta_T(f) \) reduces to the Dirac delta function \( \delta(f) \). Given the frequency band of \((10^{-4} - 10^{-1}) \) Hz, one can take the length of each segment, say, \( T \approx 3 \) h \( \sim 10^4 \) s. When \( T \) is large enough, \( \delta_T(\Delta f) \) is sharply peaked over a narrow region of width \( \sim 1/T \). Thus, in the integration (88), the product of \( \tilde{s}_1(f)\tilde{s}(f') \) contributes only in the region \( |f-f'| < 1/T \sim 10^{-4} \) Hz. The frequency band contains \( \sim 10^3 \) of these regions. By the central limit theorem, \( C \) is well-approximated by a Gaussian random variable.

In actual computations, the cross-correlated signal \( C \) can be expressed either in the time domain, or equivalently, the frequency domain. In the following we shall use the frequency domain. By Equations (79) and (82), the mean of \( C \) is

\[
\mu = \langle C \rangle = \frac{3T}{10} \int_{-\infty}^{\infty} df S_h(f)\gamma(f)\tilde{Q}(f).
\]

Notice that the mean \( \mu \) is non-zero, in contrast to (62) for one interferometer. Furthermore, the noise terms disappear in the above since they are removed by cross-correlation, and only the RGW signal accumulates with the observation time \( T \). This feature of a pair is the advantage over a single case. A greater value of \( \mu \) is desired for RGW detection. The covariance \( C \) is

\[
\sigma^2 = \langle C^2 \rangle - \langle C \rangle^2
\]

\[
= \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dk' \delta(f-f')\delta_T(k-k')\tilde{Q}(f')\tilde{Q}^*(k')
\]

\[
\times \left[ \tilde{h}_1^*(f) + \tilde{n}_1^*(f) \right] \left[ \tilde{h}_2(f') + \tilde{n}_2(f') \right] \left[ \tilde{h}_1(k) + \tilde{n}_1(k) \right] \left[ \tilde{h}_2^*(k') + \tilde{n}_2^*(k') \right]
\]

\[
- \left( \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_T(f-f')\frac{1}{2}\tilde{s}(f-f')S_h(f')R_{12}(f)\tilde{Q}(f') \right)^2.
\]

Using the “factorization” property (Allen 1997; Allen & Romano 1999)

\[
\langle x_1 x_2 x_3 x_4 \rangle = \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle,
\]

...
valid for Gaussian random variables \(x_1, x_2, x_3, x_4\), each having zero mean, and using Equations (42), (79) and (82), one obtains
\[
\sigma^2 = \frac{1}{4} \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_T^2(f - f') \left[ \left| \hat{Q}(f') \right|^2 \left[ \mathcal{R}(f') S_{1n}(f') + \mathcal{R}(f') S_{2n}(f') \right] \right. \\
+ \left. \mathcal{R}(f) S_{1n}(f) S_{2n}(f') + \mathcal{R}(f) S_{1n}(f') S_{2n}(f) + \mathcal{R}_{12}(f) S_{1n}(f) \mathcal{R}_{12}(f') S_{2n}(f') \right],
\]
where \(\mathcal{R}(f)\) is the transfer function for a single case in (44), and \(\mathcal{R}_{12}(f)\) is the transfer function for a pair in (83). For \(T\) sufficiently long, one \(\delta_T(f - f')\) can be set to be the Dirac function \(\delta(f - f')\), yielding
\[
\sigma^2 = \frac{T}{2} \int_{-\infty}^{\infty} df |\hat{Q}(f)|^2 M(f),
\]
where the function
\[
M(f) \equiv S_{1n}(f) S_{2n}(f) + \mathcal{R}(f) \left[ S_{1n}(f) + S_{2n}(f) \right] S_{h}(f) + \left[ \mathcal{R}^2(f) + \mathcal{R}_{12}^2(f) \right] S_h^2(f),
\]
which reduces to
\[
M(f) \simeq S_{2n}^2(f) + 2 \mathcal{R}(f) S_{h}(f) S_{n}(f) + \left[ \mathcal{R}^2(f) + \mathcal{R}_{12}^2(f) \right] S_h^2(f)
\]
when \(S_{1n} \simeq S_{2n} = S_n\). We plot the functions \(M, S_{2n}^2\) and \((\mathcal{R}^2 + \mathcal{R}_{12}^2) S_h^2\) of (96) for different values of parameters in Figure 9 with SNR\(_{12}\) given by (101). \(M\) is dominated by \(S_n^2\) at reasonable values of SNR\(_{12}\), so that one can take \(M(f) \simeq S_{1n}(f) S_{2n}(f)\) as a good approximation.

The SNR of the pair is defined as (Allen 1997; Allen & Romano 1999)
\[
\text{SNR}_{12} = \frac{\mu}{\sigma} = \frac{3\sqrt{T}}{10} \left[ \int_{-\infty}^{\infty} df S_h(f) \gamma(f) \hat{Q}(f) \right] \left[ \int_{-\infty}^{\infty} df |\hat{Q}(f)|^2 M(f) \right]^{1/2},
\]
which describes the detection capability of a pair. To maximize SNR\(_{12}\), one chooses the filter function (Allen 1997; Allen & Romano 1999)
\[
\hat{Q}(f) = \frac{S_h(f) \gamma(f)}{M(f)},
\]
for which the mean is
\[
\mu = \frac{3T}{10} \int_{-\infty}^{\infty} df \frac{S_h^2(f) \gamma^2(f)}{M(f)},
\]
the covariance is
\[
\sigma^2 = \frac{T}{2} \int_{-\infty}^{\infty} df \frac{S_h^2(f) \gamma^2(f)}{M(f)} = \frac{5}{3} \mu,
\]
and
\[
\text{SNR}_{12} = \frac{3\sqrt{T}}{10} \left[ \int_{-\infty}^{\infty} df \frac{\gamma^2(f) S_h^2(f)}{M(f)} \right]^{1/2}.
\]
SNR\(_{12}\) \(\propto r\) for large noise, and its dependences on \(\beta\) and \(\alpha_t\) are shown in Figure 10. When the noise is dominant, Equation (99) becomes
\[
\mu = \frac{3T}{10} \int_{0}^{\infty} df \frac{S_h^2(f) \gamma^2(f)}{S_{1n}(f) S_{2n}(f)},
\]
and (101) becomes
\[
\text{SNR}_{12} = \frac{3\sqrt{T}}{10} \left[ \int_{0}^{\infty} df \frac{\gamma^2(f) S_h^2(f)}{S_{1n}(f) S_{2n}(f)} \right]^{1/2}.
\]
Formula (103) is similar to that of the ground-based LIGO (Allen 1997; Allen & Romano 1999; Flanagan 1993; Cornish & Larson 2001). Clearly, the dependence of SNR\(_{12}\) on \(r, \beta\) and \(\alpha_t\) is implicitly contained in \(S_h(f)\) and SNR\(_{12}\) \(\propto r\) for large noise. There is a growing factor \(\sqrt{T}\) of SNR\(_{12}\) in (103), because the noise gets suppressed by cross-correlation and only the RGW signals accumulate with time.
To demonstrate the capability of a pair of space interferometers to detect RGWs, we compute the values of $\text{SNR}_{12}$ using Equation (101). The result is in Table 1, with an observation time $T = 1 \text{yr}$ and $r = 0.1$. For comparison, we have also attached the result for the pair of ground-based LIGO S6 (The LIGO Scientific Collaboration & The Virgo Collaboration 2012; https://losc.ligo.org/archive/S6/; https://losc.ligo.org/timeline/) and LIGO O1 and Advanced LIGO as well (Abbott et al. 2016a,b, 2017b), for which we use the formulae of $\text{SNR}_{12}$ and $\gamma(f)$ in Allen (1997) and Allen & Romano (1999). It is seen that $\text{SNR}_{12}$ of LISA is higher than that of LIGO S6 by 7 orders of magnitude for the default ($\alpha_t = 0, \beta = -2$), and by 8 orders of magnitude for the observed-inferred ($\alpha_t = 0, \beta = -2.016$). Therefore, LISA will have a much stronger capability than LIGO to detect RGWs.

5.3 The Sensitivity of a Pair Compared with a Single

To describe the sensitivity of a pair, we can extend the expression of (103) and allow SNR to vary with frequency. Consider the averaged $\text{SNR}_{12}$ over a frequency band of width $\Delta f$, centered at $f$ as the following (Cornish 2001)

$$
\text{SNR}_{12}(f) \simeq \sqrt{T} \left[ \int_{f-\Delta f/2}^{f+\Delta f/2} df \frac{R_{12}^2(f)S_h^2(f)}{(S_n(f))^2} \right]^{1/2}
$$

$$
\simeq \sqrt{T} \Delta f S_h(f) \left( \frac{R_{12}^2(f)}{(S_n(f))^2} \right)^{1/2},
$$

(104)

In analogy to (59), we define the effective sensitivity of the pair

$$
\tilde{h}_{12}(f) \equiv \sqrt{\frac{S_h(f)}{\text{SNR}_{12}(f)}} = \frac{1}{(T\Delta f)^{1/4}} \left( \frac{R_{12}^2(f)}{(S_n(f))^2} \right)^{-1/4},
$$

(105)

which depends on $T$ and frequency resolution $\Delta f$, in contrast to that of a single in (59). A longer $T$ increases the sensitivity of (105). The sensitivities of a single and a pair are plotted in Figure 11, where $\Delta f = f/10$ and $T = 1 \text{yr}$ are taken. Clearly, a pair has a better sensitivity than a single by $\sim 100$ times around $f \sim 10^{-2} \text{Hz}$.

5.4 Constraints on the RGW Parameters by a Pair

By (103), a constraint on $\text{SNR}_{12}$ will transfer into a constraint on $(r, \beta, \alpha_t)$. One such constraint on $\text{SNR}_{12}$ is given by (Allen 1997; Allen & Romano 1999)

$$
\text{SNR}_{12} \geq \sqrt{2} \left( \text{erfc}^{-1}(2\alpha) - \text{erfc}^{-1}(2\gamma) \right),
$$

(106)

where $\text{erfc}^{-1}(\alpha)$ is the inverse function of the complementary error function $\text{erfc}(z) \equiv \frac{2}{\sqrt{\pi}} \int_{z}^{\infty} dx e^{-x^2}$, $\alpha$ is called the false alarm rate and $\gamma$ is called the detection rate. Taking $\alpha = 5\%$ and $\gamma = 95\%$, Equation (106) gives

$$
\text{SNR}_{12} \geq 3.29.
$$

(107)

Thus, fixing two parameters out of $(r, \alpha_t, \beta)$, we can convert (107) into a lower limit on the remaining parameter of RGW. Table 2 shows the lower limits of $\alpha_t$ with the other two being fixed for $T = 1 \text{yr}$.

6 ESTIMATION BY INTEGRATED SIGNALS FROM A PAIR

6.1 The Integrated Output Signals

We now try to determine the RGW spectrum by a pair. Let the sample vector of the cross-correlated signals

$$
C = [C_1, C_2, \ldots, C_N],
$$

(108)

where each $C_i$ is the cross-correlated, integrated output signal of (88),

$$
C_i = \int_{-\infty}^{\infty} df \int_{-\infty}^{\infty} df' \delta_{TT}(f - f')s_1^*(f)s_2(f')\tilde{Q}(f') , \quad i = 1, 2, \ldots, N,
$$

(109)
and $N$ is the number of segments and is sufficiently large. When $T \gg$ the light travel time $L/c \sim 2\, s$ between the two detectors of LISA, non-overlapping $C_i$ and $C_j$ for $j \neq i$ are statistically independent (Allen 1997; Allen & Romano 1999). For each $i$, $C_i$ has the mean $\mu_i = \langle C_i \rangle$ given by (99) and the variance $\sigma_i^2 = \langle C_i^2 \rangle - \mu_i^2$ given by (100). In general, $\mu_i$ varies for different $i$, as does $\sigma_i^2$. We denote the mean of $C$ by $\mu = [\mu_1, \mu_2, \ldots, \mu_N]$ and the covariance matrix by $\Sigma = (\Sigma_{ij})$ with

$$\Sigma_{ij} = \langle (C_i - \mu_i)(C_j - \mu_j) \rangle, \quad i, j = 1, 2, \ldots, N,$$

which is diagonal, $\Sigma_{ij} = \delta_{ij} \sigma_j^2$, by independence. (Here $\Sigma$ for a pair should not be confused with that in Section 4 for a single.) Explicitly,

$$\mu_i = \langle C_i \rangle = \frac{3T_i}{10} m,$$

$$\Sigma_{ij} = \delta_{ij} b \mu_j,$$

where $b \equiv 5/3$ and

$$m \equiv \int_0^\infty df \frac{S_i^2(f) \gamma^2(f)}{M(f)},$$

is a function of $S_h(f)$. We assume that the PDF of $C$ is a multivariate Gaussian

$$f(C) = \frac{1}{(2\pi)^{\frac{N}{2}} \det[\Sigma]} \times \exp\left\{ -\frac{1}{2} (C - \mu) \Sigma^{-1} (C - \mu)^T \right\},$$

which, by (112), is

$$f(C) = \frac{1}{(2\pi)^{\frac{N}{2}} (\Pi_i^N b \mu_i)^\frac{1}{2}} \times \exp\left\{ -\frac{1}{2b} \sum_{i} (C_i - \mu_i)^2 \right\} .$$

The likelihood function is, after dropping an irrelevant constant $\frac{1}{2}N \ln 2\pi$,

$$\mathcal{L} \equiv -\ln f = \frac{1}{2} \sum_{i} \ln(b \mu_i) + \frac{1}{2b} \sum_{i} \frac{(C_i - \mu_i)^2}{\mu_i},$$

which is a function of the spectrum $S_h$ through $\mu_i$. Once the PDF is chosen, an estimator of the spectrum is a specification to give the value $S_h$ for the given data set $C$. For this, we shall adopt the ML method. In general, $\mathcal{L}$ can be expanded in a neighborhood of some spectrum $\bar{S}_h(f)$,

$$\mathcal{L} = \mathcal{L} + \sum_{k=1}^{N} \frac{\partial \mathcal{L}}{\partial S_h(f_k)} (S_h(f_k) - \bar{S}_h(f_k)) + \frac{1}{2} \sum_{k, l=1}^{N} \frac{\partial^2 \mathcal{L}}{\partial S_h(f_k) \partial S_h(f_l)} (S_h(f_k) - \bar{S}_h(f_k)) (S_h(f_l) - \bar{S}_h(f_l)).$$

We look for the most likely spectrum $\bar{S}_h(f)$ at which $\mathcal{L}$ is minimized

$$\frac{\delta \mathcal{L}}{\delta S_h} \bigg|_{\bar{S}_h} = 0.$$

The first order derivative is (see Appendix A for details)

$$\frac{\delta \mathcal{L}}{\delta S_h} = \frac{1}{2} \sum_{i} \left[ \frac{1}{\mu_i} - \frac{C_i^2}{b \mu_i^2} + \frac{1}{b} \frac{\delta \mu_i}{\delta S_h} \right] \delta S_h = \frac{1}{2} S_h(f) \gamma^2(f) \left( \frac{1}{M(f)} \right) \times \left( \frac{N}{m} - \frac{2}{m^2} \sum_{i} \frac{C_i^2}{T_i} + \frac{1}{2b^2} \sum_{i} T_i \right),$$

where $m$ is given by (113) and $N(f)$ is given by (A.9). The analytical expression of the solution for (118) is not available, and one needs to use a numerical method. The Newton-Raphson method (Oh et al. 1999; Hinshaw et al. 2003; Press et al. 1992) is generally used to find the ML-estimate of the spectrum. In many applications, the Newton-Raphson method is known to converge quadratically in the neighborhood of the root. For instance, in the spectral estimation of CMB anisotropies (Oh et al. 1999; Hinshaw et al. 2003), typically 3–4 iterations will be sufficient. Let $\bar{S}_h^{(0)}(f)$ be a trial
power spectrum, which can be tentatively chosen as the analytical spectrum (38) with some values of parameters. In the neighborhood of $S_h^{(0)}(f)$, the first order derivative of the likelihood is expanded as the following

$$\frac{\delta \mathcal{L}}{\delta S_h(f)} \bigg|_{S_h(f)} \approx \frac{\delta \mathcal{L}}{\delta S_h^{(0)}(f)} + \int df' \frac{\delta^2 \mathcal{L}}{\delta S_h(f) \delta S_h(f')} \bigg|_{S_h^{(0)}(f)} (S_h(f') - S_h^{(0)}(f')) = 0. \tag{120}$$

As an approximation, $\frac{\delta^2 \mathcal{L}}{\delta S_h \delta S_h}$ is replaced by its expected value, i.e., the Fisher matrix,

$$\mathcal{F}(f, f') = \left[ \frac{S_h(f') \gamma^2(f')}{M(f')} \left( 1 - \frac{N(f')}{M(f')} \right) \right] \left[ \frac{S_h(f) \gamma^2(f)}{M(f)} \left( 1 - \frac{N(f)}{M(f)} \right) \right] \frac{1}{2} \frac{N}{m^2} + \frac{9}{25m} \sum_{i} T_i, \tag{121}$$

(see Appendix A for the derivation). However, this Fisher matrix is degenerate and has no inverse, and one will not be able to invert Equation (120) to get an estimated spectrum. This is because the signal $C_i$ constructed in (109) is an integration over frequency, as is $\mu_i$. On the other hand, for spectrum estimation, one needs to assign a value $S_h(f_j)$ at each frequency $f_j$. Thus, we conclude that $C_i$ will not help to estimate the RGW spectrum by a pair, even though it is useful for detection of an RGW signal.

### 6.2 Parameter Estimation in a Bayesian Approach

We shall be able to use $C_i$ to estimate one parameter of RGW in a Bayesian approach. Consider the PDF as in Equation (115),

$$f(C; \theta) = \frac{1}{(2\pi)^{N/2} \left( \prod_i b_{\mu_i}(\theta) \right)^{1/2}} \times \exp \left\{ -\frac{1}{2b} \sum_i N_i (C_i - \mu_i(\theta))^2 \right\}, \tag{122}$$

where $\mu(\theta)$ and $\Sigma(\theta)$ as in (111) and (112) respectively now depend on the RGW parameters through the theoretical spectrum $S_h$, and $\theta$ denotes the RGW parameters which are random variables since they are some functions of the data set (Kay 1993a,b). We adopt the unbiased estimation, which assumes that the average value of an estimator of the parameters $\theta$ is regarded as its true value. Using the ML method, the likelihood function $L = -\ln f(\theta)$ can also be Taylor expanded around certain values $\bar{\theta}$

$$L = \bar{L} + \sum_a \frac{\partial \mathcal{L}}{\partial \theta_a} |_{\bar{\theta}} (\theta_a - \bar{\theta}_a) + \frac{1}{2} \sum_{a,b} \frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_b} |_{\bar{\theta}} (\theta_a - \bar{\theta}_a)(\theta_b - \bar{\theta}_b) + \cdots.$$ 

Now we require $\bar{\theta}$ to be the ML estimator, at which

$$\frac{\partial \mathcal{L}}{\partial \theta_a} |_{\bar{\theta}} = 0, \quad a = 1, 2, 3. \tag{123}$$

As is known (Kay 1993a,b), when $N$ is large enough, the second order derivative at $\bar{\theta}$ is equal to its average value,

$$\mathcal{F}_{ab} \equiv \left( \frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_b} \bigg|_{\bar{\theta}} \right) = \frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_b} |_{\bar{\theta}}, \quad a, b = 1, 2, 3,$n

so that in the neighborhood of $\bar{\theta}$, the PDF of (122) becomes the following Bayesian PDF in the parameter space

$$f(\theta) \propto \exp [-\mathcal{L}] \propto \exp \left[ -\frac{1}{2} (\theta - \bar{\theta}) \mathcal{F}(\theta - \bar{\theta})^T \right], \tag{124}$$

which is approximately Gaussian in a neighborhood of $\bar{\theta}$. For detailed derivation see appendix 7B of Kay (1993a,b).

The likelihood function follows (122) as

$$\mathcal{L}(C; \theta) \equiv -\ln f(C; \theta) = \frac{1}{2} \sum_i \ln (b_{\mu_i}(\theta)) + \frac{1}{2b} \sum_i N_i (C_i - \mu_i(\theta))^2. \tag{125}$$

To estimate $\theta$, one needs the first order derivative (see Appendix A),

$$\frac{\partial \mathcal{L}}{\partial \theta_a} = \int df \frac{S_h(f) \gamma^2(f)}{M(f)} \left( 1 - \frac{N(f)}{M(f)} \right) \times \frac{\partial S_h(f)}{\partial \theta_a} \left( \frac{N}{m^2} - \frac{2}{m^2} \sum_i \frac{C_i^2}{T_i} + \frac{1}{2b^2} \sum_i T_i \right), \tag{126}$$
and the $3 \times 3$ Fisher matrix
\[ F_{ab} = \left( \int_0^\infty df \frac{S_b(f') \gamma^2(f')}{M(f')} \left( 1 - \frac{N(f')}{M(f')} \right) \partial S_b(f') \right) \times \left( \int_0^\infty df \frac{S_a(f) \gamma^2(f)}{M(f)} \left( 1 - \frac{N(f)}{M(f)} \right) \partial S_a(f) \right) \times 2 \left( \frac{N}{m^2} + \frac{9}{25m} \sum_i T_i \right) \] (127)

However, this Fisher matrix is degenerate and has no inverse. Thus, one cannot determine the whole set $(r, \beta, \alpha_t)$ simultaneously. What one can do is to estimate only one of the RGW parameters, while the other two parameters are fixed at certain values or marginalized. Note that this method cannot determine the correlation between two parameters, which will be given by another method in Section 8.3 later. For the former case, one gets the conditional PDF, and for the latter, one integrates the PDF of (122) over $\theta_b$ and $\theta_c$, and gets the marginal PDF for $\theta_a$
\[ f(C; \theta_a) = \int \int f(C; \theta)d\theta_b d\theta_c \] (128)

and the marginal likelihood function $\mathcal{L}(C; \theta_a) \equiv -\ln f(C; \theta_a)$. With these specifications, one can estimate the parameter $\theta_a$. Let $\theta_a^{(0)}$ be a trial parameter. We expand the first order derivative of $\mathcal{L}$, conditional or marginalizing, around $\theta_a^{(0)}$
\[ \frac{\partial \mathcal{L}}{\partial \theta_a} |_{\theta_a^{(0)}} \approx \frac{\partial \mathcal{L}}{\partial \theta_a} |_{\theta_a} \bigg|_{\theta_a^{(0)}} \left( \theta_a - \theta_a^{(0)} \right) = 0 \] (129)

Replacing $\frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_a}$ by the $(aa)$ element of Fisher matrix $\mathcal{F}_{aa} \equiv \left\langle \frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_a} \mathcal{L} \right\rangle$, one gets
\[ \theta_a = \theta_a^{(0)} - \mathcal{F}_{aa}^{-1} \frac{\partial \mathcal{L}}{\partial \theta_a} |_{\theta_a^{(0)}} \] (130)

By iteration, one will obtain the estimate of $\theta_a$. This is a general formula to estimate one parameter. The Fisher matrix $\mathcal{F}_{ab}$ also gives the standard error of the estimated spectrum. When the data sample is sufficiently large, one can take the equality in the Cramer-Rao lower bound (Kay 1993a,b)
\[ \sigma_{\theta_a}^2 = \mathcal{F}_{aa}^{-1}(f) \] (131)

where $\mathcal{F}_{aa}^{-1}$ is evaluated at the ML-estimate parameter $\theta_a$ that has been obtained.

In fact, when noise is dominant over the RGW signal, the estimate of $r$ can be also obtained analytically. By the property $m \propto r^2$ implied by (113), one can write $m(r) = r^2 m(r = 1)$ with $\beta$ and $\alpha_t$ being fixed in $m(r = 1)$. Setting (119) to zero, and solving for $m$, one has the positive root
\[ \bar{m} = \frac{25}{9} \frac{1}{N \sum_i T_i} \left( -1 + \frac{1}{1 + \frac{36}{25} \frac{1}{N^2} \left( \sum_j T_j \right) \left( \sum_i \frac{C_i^2}{T_i} \right) } \right) \] (132)

from which one obtains the analytical ML-estimate of $r$ as the following
\[ r = \frac{1}{\sqrt{m(r = 1)}} \times \frac{25/9}{\frac{1}{N \sum_i T_i} \left( -1 + \frac{1}{1 + \frac{36}{25} \frac{1}{N^2} \left( \sum_j T_j \right) \sum_i \frac{C_i^2}{T_i} \right) } \] (133)

However, no analytical ML-estimates are available for $\beta$ and $\alpha_t$. Still one can estimate $\beta$ and $\alpha_t$ in a manner simpler than (130). Let $\theta_a$ be $\beta$ and
\[ \mathcal{M}[\theta_a] = \int_0^\infty df \frac{S_a(f) \gamma^2(f)}{M(f)} - \bar{m} = 0 \] (134)

in which $r$ and $\alpha_t$ are fixed. We use the Newton-Raphson method by iterations as before. Write $\mathcal{M}[\theta_a]$ as
\[ \mathcal{M}[\theta_a] \approx \mathcal{M}[\theta_a^{(0)}] + \frac{\partial \mathcal{M}}{\partial \theta_a} \bigg|_{\theta_a^{(0)}} \left( \theta_a - \theta_a^{(0)} \right) = 0 \] (135)
where \( \theta_{a}^{(0)} \) is a trial value and
\[
\frac{\partial \mathcal{M}}{\partial \theta_a} = \int_0^\infty df \frac{2 S_h(f) \gamma^2(f)}{M(f)} \left(1 - \frac{N(f)}{M(f)}\right) \frac{\partial S_h(f)}{\partial \theta_a}.
\]
(136)

One solves (135) and gets the estimate
\[
\theta_a = \theta_{a}^{(0)} - \frac{\partial \mathcal{M}}{\partial \theta_a} \bigg|_{\theta_a^{(0)}}^{-1} \mathcal{M}[\theta_{a}^{(0)}].
\]
(137)

Similarly, the estimation of \( \alpha_t \) can be also done. Note that since the filter function \( \hat{Q} \) in (98) contains the theoretical spectrum \( S_h(f) \), this ML-estimation method is essentially a technique of matched filter (Gair et al. 2013).

We perform a numerical simulation to estimate \( r \), using (130). For \( r = 0.1 \), \( \alpha_t = 0.016 \) and \( \beta = -2.016 \) with \( \text{SNR}_{12}=179 \), we use the PDF (115) to generate a cross-correlated data stream numerically. We take \( T \approx 3 \, \text{h} \approx 10^4 \, \text{s} \) for one segment, with total observation duration \( \sim 1 \, \text{yr} \), and number of segments \( n \sim 3 \times 10^3 \). Then we estimate \( r \) numerically by (130), and after five steps of iteration, \( r \) converges to \( r_{\text{ML}} = 0.1011 \).

According to (124) and (131), the standard deviation is \( \sigma_{\theta_a} \equiv 1/\sqrt{\mathcal{J}_{\theta a}} \). If the estimation is required to be at the 95\% confidence level (cl), \( 0.95 = \frac{2}{\sqrt{\pi}} \int_0^{\Delta \theta_a/(\sqrt{2\pi}\sigma_{\theta_a})} e^{-t^2} \, dt \), the resolution of the estimated parameter \( \theta_a \) will be
\[
\Delta \theta_a = 1.96\sigma_{\theta_a} \quad \text{at 95\% cl}.
\]
(138)

Table 3 lists the resolution of \( r, \alpha_t \) and \( \beta \), separately, and the corresponding values of \( \text{SNR}_{12} \) (≥ 3.29).

### 7 SPECTRAL ESTIMATION BY ENSEMBLE AVERAGING OF A PAIR

To estimate the RGW spectrum, we turn to the method of ensemble averaging of data from a pair. Consider the output signals \( \hat{s}_1(f), \hat{s}_2(f) \) in frequency space from (77) and (78) respectively. Since the noises are uncorrelated, the ensemble average \( \langle s_1(f)\hat{s}_2(f') \rangle \) is given by (82). In practice, when there are \( n \) independent sets of observational data, each being \( (s_1(f)\hat{s}_2(f'))_n \), they can form the sample mean,
\[
\langle s_1(f)\hat{s}_2(f') \rangle_t \equiv \frac{(s_1(f)\hat{s}_2(f'))_1 + \cdots + (s_1(f)\hat{s}_2(f'))_n}{n},
\]
(139)

which represents the ensemble average when the independent sets of data are large enough. Thus (82) becomes
\[
\langle \hat{s}_1(f)\hat{s}_2(f') \rangle_t = \frac{1}{2} \delta(f - f') S_h(f) \mathcal{R}_{12}(f).
\]
(140)

In practical analysis, we can replace \( \delta(f - f') \) by its discrete form in (64),
\[
\langle \hat{s}_1(f_i)\hat{s}_2(f_j) \rangle_t = \frac{\delta_{ij}}{2\Delta f} S_h(f_i) \mathcal{R}_{12}(f_i).
\]
(141)

Solving Equation (141), one obtains an estimate of the RGW spectrum by a pair
\[
\hat{S}_h(f_i) = \frac{2\Delta f}{\mathcal{R}_{12}(f_i)} \langle \hat{s}_1(f_i)\hat{s}_2(f_i) \rangle_t.
\]
(142)

(142) is the main formula in our paper to estimate the spectrum from a pair. As an advantage, it does not require a priori knowledge of the noise spectrum, in contrast to (76). In the ensemble averaging method, \( \langle \hat{s}_1(f_i)\hat{s}_2(f_j) \rangle_t \) as the basic quantity does not involve integration over frequency.

Using (142), we conduct a numerical simulation to examine its feasibility. First, we construct the vector of RGW output response \( \hat{h}_o(f_i) \equiv [\hat{h}_1(f_i), \hat{h}_2(f_i)] \) with \( i = 1, 2, \cdots, N \), where each of \( \hat{h}_1(f_i) \) and \( \hat{h}_2(f_i) \) is defined as in Equation (41). The mean and variance are given in Equations (42) and (82), and the corresponding PDF is
\[
f\left(\hat{h}_o(f_i)\right) = \frac{1}{(2\pi)^{\frac{D}{2}} \det \Sigma_{\hat{h}}(f_i)} \exp\left\{-\frac{1}{2} \hat{h}_o(f_i) \Sigma_{\hat{h}}(f_i)^{-1} \left[\hat{h}_o(f_i)\right]^T\right\}, \quad i = 1, 2, \cdots, N,
\]
(143)

where the covariance matrix is
\[
\Sigma_{\hat{h}}(f_i) = \frac{1}{2\Delta f} S_h(f_i) \left(\begin{array}{cc} \mathcal{R}(f_i)_1 & \mathcal{R}_{12}(f_i) \\ \mathcal{R}_{12}(f_i) & \mathcal{R}(f_i)_2 \end{array}\right), \quad i = 1, 2, \cdots, N.
\]
(144)
Here $\mathcal{R}(f_i)_1$ and $\mathcal{R}(f_i)_2$ are transfer functions of interferometers 1 and 2 respectively, and we can assume $R_1 \simeq R_2$, and $\mathcal{R}_{12}$ is the transfer function for the pair defined in (83). The inverse matrix of (144) is

$$
[\Sigma_{(k)}(f_i)]^{-1} = \frac{2\Delta f}{S_h(f_i) R^2(f_i) - R^2_{12}(f_i)} \begin{pmatrix}
\mathcal{R}(f_i) & -R_{12}(f_i) \\
-R_{12}(f_i) & \mathcal{R}(f_i)
\end{pmatrix}.
$$

Similarly, for the noise in the pair, we write the noise vector $\tilde{n}(f_i) \equiv [\tilde{n}_1(f_i), \tilde{n}_2(f_i)]$. The mean and covariance are in (49) and (80), and the PDF is

$$
f(\tilde{n}(f_i)) = \frac{1}{(2\pi)^N \det[\Sigma_{(n)}(f_i)]} \times \exp \left\{-\frac{1}{2} \tilde{n}(f_i) [\Sigma_{(n)}(f_i)]^{-1} [\tilde{n}(f_i)]^T \right\},
$$

where the covariance matrix is

$$
\Sigma_{(n)}(f_i) = \frac{S_n(f_i)}{2\Delta f} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},
$$

which is diagonal since the noise in the pair is uncorrelated. Based on the above construction, the joint PDF for the total output signal is given by (67) with the covariance matrix $\Sigma(f_i) = \Sigma_{(k)}(f_i) + \Sigma_{(n)}(f_i)$.

We use PDFs (143) and (146) to numerically generate the output response and noise of a pair. We need the data set of $n$ segments $(\tilde{s}(f_i)_1), (\tilde{s}(f_i)_n)$. One can take one typical segment of the data stream with a period of $T \sim 3 \text{h} \sim 10^4 \text{s}$, total observation duration $\sim 1 \text{yr}$ and number of segments $n \sim 3 \times 10^3$. According to the PDF of (143) and (146) with the specific variance $\Sigma$ by taking RGW with $r = 0.1, \alpha_i = 0.016, \beta = -2.016$ with SNR$_{12} = 179$, we numerically generate $3 \times 10^3$ independent sets of random data streams $\tilde{s}(f_i)_i (i = 1, \cdots, N)$. Substituting these generated data streams into Equation (142), we obtain the estimated RGW spectrum $\tilde{S}_h(f_i)$ for $n = 30$ and $n = 3000$, as shown in Figure 12. For illustration, the theoretical spectrum and the simulated noise are also shown. The estimation depends on the length of simulated data. A longer length $n$ of data gives a better estimate. In an ideal case of infinitely long data length, the off-diagonal elements of (147) would be 0 and the RGW signal at all relevant frequencies could be detected. But with a finite size of data, the estimation will be limited when the noise is large. As Figure 12 shows, the estimate (142) at high frequencies is actually contributed by the nonzero off-diagonal elements of noise, i.e., $\tilde{S}_h(f) \sim \frac{2\Delta f}{R_{12}(f_i)}$ at $f \sim 10^{-2} \text{Hz}$, because of a finite length of data.

8 ESTIMATIONS BY CORRELATION OF UN-INTEGRATED SIGNALS FROM A PAIR

In this section, we adopt a method of correlation of un-integrated signals to estimate the spectrum and parameters of RGW as suggested by Seto (2006). By dividing the whole frequency range of the data into many small segments, the mean value of the correlation variable over each segment is taken as the representative point for the segment. As an approximation, the method is able to give the estimate of the RGW spectrum, as well as the three parameters of RGW, improving that in Section 6. We assume that the output data are sufficient for this purpose.

8.1 A Correlation Variable of Un-integrated Signals

In analogy with (3.4) in Seto (2006), we divide a positive frequency range into $N$ segments, and the i-th segment $F_i (i = 1, 2, \ldots, N)$ of width $\delta f_i$ has a center frequency $f_i$. For instance, the frequency range is taken as $(10^{-4} \sim 1) \text{Hz}$ for LISA, $N \sim 10^4$ and $\delta f \sim 10^{-4} \text{Hz}$. A correlation variable is defined in each segment $F_i$ as

$$
Z_i = \sum_{f \in F_i} \Delta f \tilde{s}_1(f)\tilde{s}_2(f), \quad i = 1, 2, \ldots, N
$$

where the frequency resolution $\Delta f = 1/T_i \ll \delta f_i$ with $T_i$ being the observation period, say $\Delta f \sim 10^{-6} \text{Hz}$, so that each segment contains a large number of Fourier modes. (Notice that (3.4) in Seto (2006) should have a factor of $\Delta f$ for consistency of dimension.) The mean of $Z_i$ is

$$
\mu_i = \langle Z_i \rangle = \sum_{f \in F_i} \Delta f \frac{1}{2} \delta(f - f)S_h(f)\mathcal{R}_{12}(f),
$$
where (79) and (82) are used. Using the formula (64) to replace the Dirac delta function by its discrete form, (149) can be written as the following

\[
\mu_i = \sum_{f \in F_i} \frac{\delta f_i}{2\Delta f} S_h(f) R_{12}(f). \tag{150}
\]

The summation can be approximately replaced as

\[
\mu_i = \frac{\delta f_i}{2\Delta f} S_h(f_i) R_{12}(f_i), \quad i = 1, 2, ..., N \tag{151}
\]

where \(S_h(f_i) R_{12}(f_i)\) is the mean value over the \(i\)-th segment, as suggested in Seto (2006). To keep the error small in this approximation, \(\delta f_i\) should be sufficiently small so that the mean value of the function represents the summation function within \(\delta f_i\). (We note that the overlapping reduction function \(\gamma_{12}(f)\) defined in Seto (2006) is related to our \(R_{12}(f)\) by

\[
\gamma_{12}(f) = \frac{5}{2} R_{12}(f),
\]

which together with (38) leads to \(\mu_i = \frac{\delta f_i}{\Delta f} \frac{3H_0^2 \Omega_s(f)}{\pi^2} \gamma_{12}(f)\), the same as (3.5) in Seto (2006).

The variance of \(Z_i\) is

\[
\sigma_i^2 = \left \langle \left \{ Z_i - \langle Z_i \rangle \right \}^2 \right \rangle = \langle Z_i^2 \rangle - \langle Z_i \rangle^2. \tag{152}
\]

By noting that \(\int_{-\infty}^{\infty} df' \delta(f-f')\) is equivalent to \(2 \sum_{f' \in F_j} \Delta f \delta(f-f')\), (148) can be written as

\[
Z_i = \sum_{f \in F_i} \Delta f \left[ 2 \sum_{f' \in F_j} \Delta f \delta(f-f') \right] \tilde{s}_1(f) \tilde{s}_2(f'), \tag{153}
\]

and the variance is written as

\[
\sigma_i^2 = 4 \left \langle \sum_{f \in F_i} \Delta f \sum_{f' \in F_j} \Delta f \delta(f-f') \tilde{s}_1^*(f) \tilde{s}_2(f') \times \sum_{k \in F_i} \Delta f \sum_{k' \in F_j} \Delta f \delta(k-k') \tilde{s}_1^*(k) \tilde{s}_2(k') \right \rangle - \mu_i^2. \tag{154}
\]

By similar calculations leading to (94), we obtain the following result

\[
\sigma_i^2 = \frac{1}{8 \Delta f} M(f_i), \quad i = 1, 2, ..., N \tag{155}
\]

where \(M(f)\) is defined in (95). For large noise, \(M(f) \simeq S_{1n}(f) S_{2n}(f)\),

\[
\sigma_i^2 \simeq \frac{1}{8 \Delta f} S_{1n}(f_i) S_{2n}(f_i), \tag{156}
\]

which is the same as (3.6) in Seto (2006).

SNR of each segment for this correlation is defined as in Seto (2006)

\[
\text{SNR}_i^2 = \frac{\mu_i^2}{\sigma_i^2} = 2 \frac{\delta f_i}{\Delta f} \frac{R_{12}^2(f_i) S_{2n}^2(f_i)}{M(f_i)}, \tag{157}
\]

and summing up all segments yields the total SNR

\[
\text{SNR}^2_c = 2 \sum_{i=1}^{N} \frac{\delta f_i}{\Delta f} \frac{R_{12}^2(f_i) S_{2n}^2(f_i)}{M(f_i)}. \tag{158}
\]

Replacing \(\Delta f\) with \(1/T_1\), the summation \(\sum_{i=1}^{N} \delta f\) with integration \(\int_0^\infty df\), one has the SNR over an observation period \(T_1\) as

\[
\text{SNR}_c = \sqrt{2T_1 \left[ \int_0^\infty df \frac{R_{12}^2(f) S_{2n}^2(f)}{M(f)} \right]^{1/2}}. \tag{159}
\]

When the whole observation duration \(T\) consists of many observation periods, we can use \(T\) to replace \(T_1\) in the above formula. This result is consistent with (101) by noting that \(R_{12}(f) = \frac{\delta f}{10} \gamma(f)\).
8.2 Spectrum Estimation

Since \( \delta f_i / \Delta f \gg 1 \), according to the central limit theorem, \( Z_i \) can be described by a Gaussian distribution, and the PDF for \( Z = [Z_1, Z_2, \ldots, Z_N] \) is

\[
f(Z) = \frac{1}{(2\pi)^{N/2} \left( \Pi_i \sigma_i^2 \right)^{\frac{N}{2}}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{N} (Z_i - \mu_i)^2 / \sigma_i^2 \right\}.
\] (160)

The likelihood function is, after dropping an irrelevant constant \( \frac{1}{2} N \ln 2\pi \),

\[
\mathcal{L} \equiv -\ln f(Z) = \frac{1}{2} \sum_{i=1}^{N} \ln(\sigma_i^2) + \frac{1}{2} \sum_{i=1}^{N} (Z_i - \mu_i)^2 / \sigma_i^2,
\] (161)

which is a function of the spectrum \( S_h \) through \( \mu_i \). We look for the most likely power spectrum \( \bar{S}_h \) at which \( \frac{\partial \mathcal{L}}{\partial S_h} \bigg|_{S_h} = 0 \). The first order derivative is

\[
\frac{\partial \mathcal{L}}{\partial S_h(f_i)} = \frac{N(f_i)}{S_h(f_i) M(f_i)} - \frac{8 \left( Z_i - \left( \frac{\delta f_i}{2\Delta f} S_h(f_i) \mathcal{R}_{12}(f_i) \right) \right)^2}{\delta S_h(f_i) [M(f_i)]^2} \frac{N(f_i)}{\mathcal{R}_{12}(f_i) \left( Z_i - \left( \frac{\delta f_i}{2\Delta f} S_h(f_i) \mathcal{R}_{12}(f_i) \right) \right)}
\] (163)

where \( N(f) \) is defined in (A.9). In the neighborhood of a trial spectrum \( S_h^{(0)}(f) \), the first order derivative is expanded as the following

\[
\frac{\partial \mathcal{L}}{\partial S_h(f_i)} \bigg|_{S_h^{(0)}(f_i)} \approx \frac{\partial \mathcal{L}}{\partial S_h(f_i)} \bigg|_{S_h^{(0)}(f_i)} + \frac{N(f_i)}{S_h^{(0)}(f_i) M^{(0)}(f_i)} \sum_{k=1}^{N} \partial^2 \mathcal{L} \bigg|_{S_h^{(0)}(f_i)} \left( S_h(f_k) - S_h^{(0)}(f_k) \right) = 0 \]. (164)

As an approximation, \( \frac{\partial^2 \mathcal{L}}{\partial S_h \partial S_h} \) is replaced by its expected value, i.e., the Fisher matrix which by a formula similar to (A.5) is given by

\[
F_{ij} = \left\{ \frac{\partial^2 \mathcal{L}}{\partial S_h(f_i) \partial S_h(f_j)} \right\} = \sum_{k,l=1}^{N} \frac{\partial \mu_k}{\partial S_h(f_i)} \frac{\partial \mu_l}{\partial S_h(f_j)} \frac{\delta_{kl}}{\sigma_k \sigma_l} + \frac{1}{2} \sum_{k,l,m,r=1}^{N} \left( \frac{\delta_{kl}}{\sigma_k \sigma_l} \frac{\partial (\sigma_i^2)}{\partial S_h(f_m)} \frac{\delta_{mr}}{\sigma_m} \frac{\delta_{kl}}{\sigma_k \sigma_l} \frac{\partial (\sigma_i^2)}{\partial S_h(f_r)} \right) \]. (165)

Substituting (151) and (155) into the above yields

\[
F_{ij} = \left[ \frac{\delta f_i}{\Delta f} \frac{2 \mathcal{R}_{12}(f_i)}{M(f_i)} + \frac{2 \mathcal{N}^2(f_i)}{S^2_h(f_i) M^2(f_i)} \right] \delta_{ij}.
\] (166)

It is remarked that the Fisher matrix is not degenerate, in contrast to (121) which is degenerate. In the approximation of large noise, one has \( M(f) \simeq S_{1n}(f) S_{2n}(f) \), \( N(f) = \frac{1}{2} S_h(f) \frac{\partial M(f)}{\partial S_h(f)} \simeq 0 \) and (166) reduces to that used in Seto (2006)

\[
F_{ij} \simeq \left[ \frac{\delta f_i}{\Delta f} \frac{2 \mathcal{R}_{12}(f_i)}{S_{1n}(f) S_{2n}(f_i)} \right] \delta_{ij}.
\] (167)

We plot \( F_{ij}(f) \) of (166) and of (167) in Figure 13. They differ significantly at high frequencies. Thus, we shall use the full expression (166) in computations later.

Given \( F_{ij} \), one solves Equation (164) for the estimated spectrum

\[
S_h(f_i) = S_h^{(0)}(f_i) - \sum_{j=1}^{N} \left( F^{-1} \right)_{ij} \frac{\partial \mathcal{L}}{\partial S_h(f_j)} \bigg|_{S_h^{(0)}}.
\] (168)
To avoid random outcomes from one set of data streams, similar to (139), we shall replace \( \frac{\partial L}{\partial S_n(f)} \) by its sample mean in practical computation

\[
S_h(f_i) = S_h^{(0)}(f_i) - \left( \frac{\delta f_i}{\Delta f} \frac{2R^2_{12}(f_i)}{M(f_i)} + \frac{2N^2(f_i)}{S_h^{(0)}(f_i)} \right)^{-1} \left. \left( \frac{\partial L}{\partial S_h(f_i)} \right) \right|_{f_i},
\]

(169)

where the expression (166) has been used. This equation will be used to estimate the spectrum of RGW numerically by Newton-Raphson iteration (Oh et al. 1999; Hinshaw et al. 2003; Press et al. 1992).

We perform a simulation to estimate \( S_h(f) \). We divide a one-year duration into 100 periods, which are regarded as \( n = 100 \) realizations of data output. Thus, one observation period \( T_1 \approx 1 \text{ yr} / 100 \approx 3.2 \times 10^5 \text{ s} \). The working frequency range \((10^{-4} \sim 1) \text{ Hz} \) is divided into \( N \approx 10 \times 637 \) segments and the width of each segment is \( \delta f \approx 9.4 \times 10^{-8} \text{ Hz} \). Thus, each segment contains \( \delta f / \Delta f \approx 30 \) frequency points. For RGW we take \( r = 0.1, \alpha_1 = 0.016 \) and \( \beta = -2.016 \), and the formula (159) yields SNR = 179. We numerically generate the output response \([\tilde{h}_{1,1}(f_i), \cdots, \tilde{h}_{1,n}(f_i); \tilde{h}_{2,1}(f_i), \cdots, \tilde{h}_{2,n}(f_i)]\) according to (143) and the noise \([\tilde{n}_{1,1}(f_i), \cdots, \tilde{n}_{1,n}(f_i); \tilde{n}_{2,1}(f_i), \cdots, \tilde{n}_{2,n}(f_i)]\) according to (146) of a pair, for \( i = 1, \cdots, 10 \times 637 \) and for \( n = 100 \) realizations. We use Equation (148) to calculate the correlated signal \([Z_{i,1}, \cdots, Z_{i,n}]\) for \( n = 100 \) realizations. Using these generated data streams, we apply Newton-Raphson iterations to Equation (169) to estimate the spectrum of RGW numerically.

Figure 14 shows the resulting estimator of spectrum \( S_h^{(n)}(f) \) after each iterative step. It is seen that after three iterations, \( S_h^{(n)}(f) \) converges. The estimated RGW spectrum \( S_h(f_i) \) is shown in Figure 15. For illustration, the theoretical spectrum is also shown. It is seen that the rapid oscillations in the estimated spectrum are smoother than the theoretical one. This is because the estimated spectrum is actually a mean \( \bar{\theta} \), given by (151) and (155) respectively, are now regarded as functions of parameters \( \theta \) through the theoretical spectrum \( S_h(f) \). The likelihood function \( L = -\ln f(\theta) \) can also be Taylor expanded around certain values \( \bar{\theta} \)

\[
L = \bar{\mathcal{L}} + \sum_a \frac{\partial \mathcal{L}}{\partial \theta_a} \bigg|_{\bar{\theta}} (\theta_a - \bar{\theta}_a) + \frac{1}{2} \sum_{a,b} \frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_b} \bigg|_{\bar{\theta}} (\theta_a - \bar{\theta}_a)(\theta_b - \bar{\theta}_b) + \cdots.
\]

(170)

Now we require \( \bar{\theta} \) to be the ML estimator, at which

\[
\frac{\partial \mathcal{L}}{\partial \theta_a} \bigg|_{\bar{\theta}} = 0, \quad a = 1, 2, 3.
\]

(171)

Based on (171), we use Newton-Raphson method (Oh et al. 1999; Hinshaw et al. 2003; Press et al. 1992) to estimate \( \bar{\theta} \). The first order derivative is expanded around the trial \( \theta^{(0)} \) as the following

\[
\frac{\partial \mathcal{L}}{\partial \theta_a} \bigg|_{\theta^{(0)}} \approx \frac{\partial \mathcal{L}}{\partial \theta_a} \bigg|_{\bar{\theta}} + \sum_{b=1}^{3} \frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_b} \bigg|_{\theta^{(0)}} (\theta_b - \theta_b^{(0)}) = 0, \quad a = 1, 2, 3.
\]

(172)

The second order derivative in the above is approximately replaced by the Fisher matrix, leading to

\[
\frac{\partial \mathcal{L}}{\partial \theta_a} \bigg|_{\theta^{(0)}} + \sum_{b=1}^{3} F_{ab} \bigg|_{\theta^{(0)}} (\theta_b - \theta_b^{(0)}) = 0, \quad a = 1, 2, 3,
\]

(173)
from which we obtain
\[ \theta_a = \theta_a^{(0)} - \sum_{b=1}^{3} F_{ab}^{-1} \left\{ \frac{\partial L}{\partial \theta_b} \right\}_{\theta^{(0)}} , \quad a = 1, 2, 3 . \] (174)

By iteration, one can obtain an estimate of the parameters \{\theta_a\}. The explicit expressions of derivatives in the above are given by the chain rule by using (163),
\[
\frac{\partial L}{\partial \theta_a} = \sum_{i=1}^{N} \frac{\partial \theta_a}{\partial s_h(f_i)} \frac{\partial s_h(f_i)}{\partial \theta_a} = \sum_{i=1}^{N} \left[ N(f_i) \frac{\partial s_h(f_i)}{s_h(f_i)} M(f_i) - 8 \left( Z_i - (\frac{\partial f_i}{\partial f_i} S_h(f_i) R_{12}(f_i)) \frac{\partial s_h(f_i)}{s_h(f_i)} \frac{\partial M(f_i)}{M(f_i)} \right)^2 \right] \frac{\partial s_h(f_i)}{\partial \theta_a} , \quad a = 1, 2, 3 .
\] (175)

The Fisher matrix is provided by the following
\[
F_{ab} = \left\{ \frac{\partial^2 L}{\partial \theta_a \partial \theta_b} \right\} = \sum_{k,l,m,r=1}^{N} \frac{\partial \mu_k}{\partial \mu_l} \frac{\partial \mu_l}{\partial \mu_k} + \frac{1}{2} \sum_{k,l,m,r=1}^{N} \left( \frac{\partial f_i}{\partial f_i} \frac{\partial f_i}{\partial f_i} \frac{\partial f_i}{\partial f_i} \frac{\partial f_i}{\partial f_i} \right) \frac{\partial s_h(f_k)}{s_h(f_k)} \frac{\partial s_h(f_k)}{s_h(f_k)} , \quad a = 1, 2, 3 .
\] (176)

Using (151), (155) and the chain rule, one has
\[
F_{ab} = \sum_{k}^{N} \frac{\partial f_k}{\partial f_k} \frac{\partial s_h(f_k)}{s_h(f_k)} \frac{\partial s_h(f_k)}{s_h(f_k)} + \sum_{k=1}^{N} \left( \frac{2 N^2(f_k)}{S_k^2(f_k)} \frac{\partial s_h(f_k)}{s_h(f_k)} \frac{\partial s_h(f_k)}{s_h(f_k)} \right) , \quad (177)
\] which is not degenerate, since \(\mu_i\) and \(\sigma_i^2\) in (151) and (156) respectively, carry information on frequency \(f_i\). Thus, (174) can be used to estimate the three parameters at the same time. In the limit of dominating noise, one has \(M(f) \approx S_{n1}(f) S_{n2}(f)\) and \(N(f) \approx 0\), and (177) reduces to
\[
F_{ab} = \sum_{k}^{N} \frac{\partial f_k}{\partial f_k} \frac{\partial s_h(f_k)}{s_h(f_k)} \frac{\partial s_h(f_k)}{s_h(f_k)} , \quad (178)
\] Replacing \(\Delta f\) with \(1/T_1\), and sum \(\sum_{i=1}^{N} \Delta f \) with \(f_{\infty} \), the above becomes
\[
F_{ab} = T_1 \int_{0}^{\infty} df \frac{2 R_{12}(f_k)}{S_{n1}(f) S_{n2}(f_k)} \frac{\partial s_h(f_k)}{s_h(f_k)} \frac{\partial s_h(f_k)}{s_h(f_k)} , \quad (179)
\] which agrees with (3.11) in Seto (2006).

The element \(F_{ab}\) can be viewed as an “inner” product of two vectors \(\frac{\partial s_h(f)}{\partial \theta_a}\) and \(\frac{\partial s_h(f)}{\partial \theta_b}\). When the vectors \(\frac{\partial s_h(f)}{\partial \theta_a}\) with \(a = 1, 2, 3\) are orthogonal to each other, \(F_{ab}\) will be diagonal, and the errors in estimates of different parameters will be uncorrelated. On the other hand, when two \(\frac{\partial s_h(f)}{\partial \theta_a}\) and \(\frac{\partial s_h(f)}{\partial \theta_b}\) have similar shapes, \(\theta_a\) and \(\theta_b\) will be degenerate and their effects are difficult to distinguish in estimation.

In Figure 16, the three curves \(\frac{\partial s_h(f)}{\partial \theta_a}\) based on the theoretical spectrum are plotted, showing that the three parameters of RGW have strong degeneracy within a small frequency range.

For \(r = 0.1, \alpha_t = 0.016\) and \(\beta = -2.016\) with \(\text{SNR_C}=179\), we use the data streams generated in Section 8.2 and use (174) to numerically estimate \(r\), and find that \(r\) converges to \(r_{\text{ML}} = 0.1070\) after nine iterations.

As discussed before, in the neighborhood of \(\bar{\theta}\), one has the following Bayesian PDF in the parameter space
\[
f(\theta) \propto \exp \left[ -L \right] \propto \exp \left[ -\frac{1}{2} (\theta - \bar{\theta}) F(\theta - \bar{\theta}) \right].
\] (180)

The resolution of the parameters will be
\[
\Delta \theta_a = 1.96 \sigma_{\theta_a} = 1.96 \sqrt{(F^{-1})_{aa}} \text{ at 95% cl}, \quad (181)
\]
Fig. 3 Left: the transfer functions $R(f)$. Right: the noise spectrum $S_n(f)$.

Fig. 4 The sensitivity curves $\tilde{h}(f)$ of Michelson, Sagnac and symmetrized Sagnac.

Fig. 5 Comparison of $h_c(f)$ for RGW with the sensitivity $\tilde{h}(f)$ of a single interferometer.

Fig. 6 A pair of interferometers in two triangles for config. 2.

Fig. 7 A comparison between the transfer function of a single and a pair.
Fig. 8 The overlap reduction function for a pair. Top: Config. 1; Bottom: Config. 2. Solid lines are plotted numerically and dotted lines by the fitted formulae (86) and (87).

Fig. 9 $M, S_n^2, (R^2 + R_{12}^2) S_n^2.$ Left: for SNR$_{12} = 8.62$; Right: for SNR$_{12} = 395$.

Table 1 SNR$_{12}$ for a Pair Case for LISA and for a Pair Case for LIGO with $r = 0.1$

| $\alpha, \beta$ | $\alpha = 0.01, \beta = 0.2016$ | $\alpha = 0.02, \beta = 0.2016$ |
|---|---|---|
| LIGO S6 | $1.6 \times 10^{-13}$ | $5.0 \times 10^{-10}$ |
| LIGO O1 | $3.0 \times 10^{-11}$ | $7.3 \times 10^{-8}$ |
| Advanced LIGO | $3.3 \times 10^{-10}$ | $7.3 \times 10^{-7}$ |
| A pair case for LISA | $1.1 \times 10^{-4}$ | $2.2 \times 10^{-2}$ |
Fig. 10  SNR changes with $\beta$ (top) and $\alpha_t$ (bottom).

Fig. 11  The sensitivity curves of a single and a pair.

Table 2  The Lower Limits of $\alpha_t$ with the Other Two Being Fixed

| $r$ | $\beta$ | $-1.94$ | $-1.96$ | $-1.98$ | $-2$ | $-2.02$ | $-2.04$ | $-2.06$ | $-2.08$ |
|-----|---------|--------|--------|--------|------|--------|--------|--------|--------|
| 0.1 |         | $-0.00041$ | $0.00190$ | $0.00421$ | $0.00653$ | $0.00884$ | $0.1115$ | $0.01346$ | $0.01577$ |
| 0.05|         | $0.00074$ | $0.00306$ | $0.00537$ | $0.00768$ | $0.00999$ | $0.01230$ | $0.01461$ | $0.01692$ |
Fig. 12 The estimated spectrum by a pair. The cases with only the theoretical spectrum and simulated noise are also shown.

Fig. 13 The diagonal element $F_{ii}$ to estimate spectrum for both the full expression (166) and the noise dominant (167) cases.

Fig. 14 The estimator of spectrum $S_{\alpha}^{(n)}(f)$ in each iterative step for a pair.
Fig. 15 The estimated spectrum by the correlation method for a pair. The theoretical spectrum is also shown.

Fig. 16 $\frac{\partial S_h}{\partial \theta}$ for $r=0.1$, $\beta=-2.016$ and $\alpha_t=0.016$.

and the correlation coefficient between two parameters will be

$$CR_{\theta_a, \theta_b} = \frac{\langle (\theta_a - \langle \theta_a \rangle)(\theta_b - \langle \theta_b \rangle) \rangle}{\sigma_{\theta_a} \sigma_{\theta_b}} = \frac{(F^{-1})_{ab}}{\sqrt{(F^{-1})_{aa}(F^{-1})_{bb}}}.$$ (182)

Note that $CR_{\theta_a, \theta_b} = 0$ indicates the independency of $\theta_a$ and $\theta_b$, and $|CR_{\theta_a, \theta_b}| = 1$ indicates the complete correlation of $\theta_a$ and $\theta_b$. Comparing Tables 4 and 3, when $r$, $\alpha_t$, and $\beta$ are estimated at the same time, where Table 4 lists the resolutions, correlations and the corresponding values of SNR_{12} ($\geq 3.29$), we find that when estimating the three parameters simultaneously, the resolution would get worse. This is due to the degeneracy of the three parameters as shown in Figure 16, which can also be seen with $|CR_{\theta_a, \theta_b}| \simeq 1$ in Table 4. Since the amplitude of the spectrum increases with $r$, $\beta$ and $\alpha_t$, when simultaneously estimating the three parameters, a larger estimated value of $\beta$ than its true value will lead to smaller estimated $r$ and $\alpha_t$ than their true values, or vice versa. This is reflected in the negative signs of $CR_{r, \beta}$ and $CR_{\alpha_t, \beta}$ in Table 4. Besides, we find that if estimating only two parameters with the third one fixed, the correlation coefficients between every two parameters are all negative, which is also an expected feature.

9 CONCLUSIONS

We have presented a study of statistical signal processing for RGW detection by space-borne interferometers, using LISA as an example, and have shown how to estimate the RGW spectrum and parameters from the output signals in the future. We have given the relevant formulations of estimations, which apply to LISA, as well as to other
space-borne interferometers with some appropriate modifications.

For a single interferometer, the Michelson is shown to have a better sensitivity than Sagnac and symmetrized Sagnac, due to its greater transfer function \( R \), even though its noise is larger. A pair has the advantage of suppressing the noise level by cross-correlation, so that RGW signal in the cross-correlated output will be accumulating with observation time, leading to a higher sensitivity than a single case. We have given the expressions of SNR for both a single and a pair, which are 4 \sim 6 orders of magnitude higher than those of ground-based ones for the default RGW parameters. We have shown that a single is not practical to estimate the RGW spectrum when noise is dominantly large, because we do not know the precise noise that actually occurs in the data.

For a pair of interferometers, we have used the cross-correlated integrated signals \( C \) in (88) and calculated \( \text{SNR}_{12} \) as in (101) which provides a statistical criterion for detection of RGW. However, one cannot estimate the spectrum by the integrated signals \( C \), because it is an integration over frequency. Assuming Gaussian output signals, we have calculated the covariance of signals, and obtained the Gaussian PDF, the likelihood function \( L \) and the Fisher matrix. By the Bayesian approach, we estimate one parameter and compute the resolution using \( C \). In the second method we have proposed applying the ensemble averaging method to estimate the spectrum, using the unintegrated output signals of a pair. (142) is the main formula. We have demonstrated by simulation that the method will be effective in estimating the spectrum. Besides, the method is simple and does not depend on detailed knowledge of the noise. For the third method, we have also studied the correlation variable of un-integrated signals from (148). We have obtained the formulations for estimation of the spectrum and parameters of RGW by the ML-estimation. Equations (169) and (174) are the main formulæ. We have shown by simulations that the method is feasible when the data set is sufficiently large. This method uses the mean value for each segment and thus loses some fine information on the RGW spectrum, but it is capable of estimating all three parameters.

There are other effects to be analyzed that are not presented in this paper. In particular, other types of GWs that are different from RGW also exist in the Universe, and should be separated in order to estimate the RGW spectrum. GWs from a resolved astrophysical source, either periodic or short-lived, can be distinguished in principle. The real concern is the stochastic foreground that may be mixed with RGW. So far, the theoretical spectrum of this foreground is less known and highly model-dependent. For a definite model of the foreground spectrum, Adams & Cornish (2014) discuss a discrimination method using the spectral shapes and the time modulation of the signal. The estimation of RGW in the presence of foreground will require substantial analysis and will be left for future work.

**Acknowledgements** Y. Zhang is supported by the National Natural Science Foundation of China (Grant Nos. 11421303, 11675165 and 11633001), SRFDP and CAS, the Strategic Priority Research Program “The Emergence of Cosmological Structures” of the Chinese Academy of Sciences (Grant No. XDB09000000).

| Table 3 Resolution of \( r, \alpha_t, \beta \) Separately at 95% cl for a Pair |
|-----------------|--------|--------|--------|--------|--------|
| \( r \)         | \( \alpha_t \) | \( \beta \) | \( \Delta r/r \) | \( \Delta \alpha_t \) | \( \Delta \beta \) | \( \text{SNR}_{12} \) |
| 0.05            | 0.01   | -2.016 | 2.58 \times 10^{-2} | 4.29 \times 10^{-5} | 3.72 \times 10^{-4} | 4.36          |
| 0.05            | 0.015  | -2.016 | 1.46 \times 10^{-2} | 2.42 \times 10^{-5} | 2.10 \times 10^{-4} | 72.8         |
| 0.1             | 0      | -1.93  | 2.56 \times 10^{-2} | 4.28 \times 10^{-5} | 3.70 \times 10^{-4} | 8.34         |
| 0.1             | 0.01   | -2.016 | 2.55 \times 10^{-2} | 4.26 \times 10^{-5} | 3.69 \times 10^{-4} | 8.62         |
| 0.1             | 0.01   | -1.93  | 1.46 \times 10^{-2} | 2.40 \times 10^{-5} | 2.09 \times 10^{-4} | 390          |
| 0.1             | 0.016  | -2.016 | 9.72 \times 10^{-3} | 1.61 \times 10^{-5} | 1.40 \times 10^{-4} | 179          |

| Table 4 Resolution of \( r, \alpha_t, \beta \) at 95% cl for a Pair, and Correlations between Them |
|-----------------|--------|--------|--------|--------|--------|--------|--------|
| \( r \)         | \( \alpha_t \) | \( \beta \) | \( \Delta r/r \) | \( \Delta \alpha_t \) | \( \Delta \beta \) | \( \text{SNR}_{C} \) |
| 0.05            | 0.01   | -2.016 | 1.54 \times 10^{4} | 26.3   | 450   | 0.999697 | -0.999923 | -0.999932 | 4.36        |
| 0.05            | 0.015  | -2.016 | 997    | 1.70   | 29.1  | 0.999724 | -0.999930 | -0.999932 | 72.8        |
| 0.1             | 0      | -1.93  | 6.40 \times 10^{3} | 11.0   | 188   | 0.999634 | -0.999907 | -0.999910 | 8.34        |
| 0.1             | 0.01   | -2.016 | 7.71 \times 10^{3} | 13.2   | 226   | 0.999696 | -0.999923 | -0.999925 | 8.62        |
| 0.1             | 0.01   | -1.93  | 67.5   | 0.117  | 1.99  | 0.999460 | -0.999863 | -0.999867 | 390         |
| 0.1             | 0.016  | -2.016 | 329    | 0.561  | 9.60  | 0.999681 | -0.999919 | -0.999921 | 179         |
Appendix A: THE FISHER MATRIX FOR A PAIR

Given the data of cross-correlated signals \( C \) in (109) for a pair, we assume that the PDF is multivariate Gaussian

\[
f(C) = \frac{1}{(2\pi)^{\frac{N}{2}} \det\Sigma} \exp\left\{ -\frac{1}{2} (C - \mu) \Sigma^{-1} (C - \mu)^T \right\},
\]

where the mean \( \mu_i \) and covariance matrix \( \Sigma_{ij} \) are given by (111) and (112) respectively, both being functions of the spectrum \( S_h(f) \). The likelihood function is (dropping an irrelevant constant \( \frac{1}{2} N \ln 2\pi \))

\[
\mathcal{L} \equiv -\ln f = \frac{1}{2} \ln \det[\Sigma] + \frac{1}{2} (C - \mu) \Sigma^{-1} (C - \mu)^T.  
\]

The first order derivative is (Kay 1993a,b)

\[
\frac{\delta \mathcal{L}}{\delta S_h(f)} = \frac{1}{2}Tr \left( \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f)} \Sigma^{-1} y^T - \frac{\delta \mu}{\delta S_h(f)} \Sigma^{-1} y^T \right)
\]

where \( y \equiv C - \mu \), and the second order derivative is

\[
\frac{\delta^2 \mathcal{L}}{\delta S_h(f) \delta S_h(f')} = - \frac{1}{2} Tr \left( \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f')} \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f)} \right) + \frac{1}{2} Tr \left( \Sigma^{-1} \frac{\delta^2 \Sigma}{\delta S_h(f') \delta S_h(f)} \right)
\]

\[
+ y \left( \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f')} \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f)} + \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f')} \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f)} \right) y^T
\]

\[
- \frac{1}{2} y \Sigma^{-1} \frac{\delta^2 \Sigma}{\delta S_h(f') \delta S_h(f)} \Sigma^{-1} y^T - \frac{\delta \mu}{\delta S_h(f')} \Sigma^{-1} \frac{\delta \mu}{\delta S_h(f)} \Sigma^{-1} y^T
\]

\[
+ \frac{\delta \mu}{\delta S_h(f)} \Sigma^{-1} \frac{\delta \mu}{\delta S_h(f')} \Sigma^{-1} y^T.
\]

Taking the expected value of the above yields the Fisher matrix

\[
\mathcal{F}(f, f') = \left\langle \frac{\delta^2 \mathcal{L}}{\delta S_h(f) \delta S_h(f')} \right\rangle = \frac{\delta \mu}{\delta S_h(f)} \Sigma^{-1} \frac{\delta \mu}{\delta S_h(f')} + \frac{1}{2} Tr \left( \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f)} \Sigma^{-1} \frac{\delta \Sigma}{\delta S_h(f')} \right)
\]

where \( \left\langle y^T \right\rangle = 0 \) and \( \left\langle y^T y \right\rangle = \Sigma \) are used. Using \( \Sigma_{ij} \) of (112), one has

\[
\frac{\delta \mathcal{L}}{\delta S_h(f)} = \sum_{i} \left[ \frac{1}{2 \mu_i} - \frac{C_i^2}{2b} \right] \frac{\delta \mu_i}{\delta S_h(f)}
\]

\[
\mathcal{F}(f, f') = \sum_{i} \left[ \frac{1}{2 \mu_i} + \frac{1}{2b} \right] \frac{\delta \mu_i}{\delta S_h(f)} \frac{\delta \mu_i}{\delta S_h(f')}.
\]

where \( b \equiv \frac{5}{3} \). By (111), the derivative of \( \mu \) is given by

\[
\frac{\delta \mu_i}{\delta S_h(f)} = \frac{T_i}{2b} \frac{\delta m}{\delta S_h(f)} = \frac{T_i S_h(f) \gamma^2(f)}{2b M(f)} \left( 1 - \frac{N(f)}{M(f)} \right), \quad i = 1, \ldots, N,
\]

where \( M(f) \) is defined by (95) and

\[
N(f) \equiv \frac{1}{2} S_h(f) \frac{\delta M(f)}{\delta S_h(f)} = \frac{1}{2} \left[ S_{1n}(f) + S_{2n}(f) \right] \mathcal{R}(f) S_h(f) + \left[ \mathcal{R}^2(f) + \mathcal{R}^2_{12}(f) \right] S_h^2(f).
\]

Using (111) and (A.8), the first order derivative is

\[
\frac{\delta \mathcal{L}}{\delta S_h(f)} = \frac{1}{2} \frac{S_h(f) \gamma^2(f)}{M(f)} \left( 1 - \frac{N(f)}{M(f)} \right) \left( \frac{N}{m} - \frac{2}{m^2} \sum_{i} C_i^2 \frac{1}{T_i} + \frac{1}{2b^2} \sum_{i} T_i \right),
\]
and the Fisher matrix is
\[
\mathcal{F}(f, f') = \left[ \frac{S_h(f') \gamma^2(f')}{M(f')} \left( 1 - \frac{N(f')}{M(f')} \right) \right] \left[ \frac{S_h(f) \gamma^2(f)}{M(f)} \left( 1 - \frac{N(f)}{M(f)} \right) \right] \frac{1}{2} \left( \frac{N}{m^2} + \frac{9}{25m} \sum_i T_l \right). 
\]
(A.11)

From the above formulae, we also derive the Fisher matrix \( F_{ab} \) for parameter estimation. Consider the PDF in (A.1)
\[
f(C; \theta) = \frac{1}{(2\pi)^\frac{3}{2} \text{det}[\Sigma(\theta)\}] \exp \left\{ -\frac{1}{2} (C - \mu(\theta))^\top \Sigma^{-1}(\theta) (C - \mu(\theta)) \right\}, \tag{A.12}
\]
where \( \mu(\theta) \) and \( \Sigma(\theta) \) now depend on the RGW parameters \( \theta = \{r, \beta, \alpha_\ell \} \) through the theoretical spectrum \( S_h(f) \) in (111) and (112). Using the result (A.6) and (A.7), by the chain rule, one obtains the derivatives with respect to the parameters
\[
\frac{\partial \mathcal{L}}{\partial \theta_a} = \frac{1}{2} \sum_i^N \left[ \frac{1}{\mu_i} - \frac{C_i^2}{b \mu_i^2} + 1 \right] \frac{\partial \mu_i}{\partial \theta_a}, \tag{A.13}
\]
\[
F_{ab} = \left\langle \frac{\partial^2 \mathcal{L}}{\partial \theta_a \partial \theta_b} \right\rangle = \sum_i^N \left( \frac{1}{b \mu_i} + \frac{1}{2 \mu_i^2} \right) \frac{\partial \mu_i}{\partial \theta_a} \frac{\partial \mu_i}{\partial \theta_b}. \tag{A.14}
\]

Taking derivative of \( \mu \) in Equation (111) with respect to \( \theta_a \) yields
\[
\frac{\partial \mu_4}{\partial \theta_a} = \frac{T_i}{A} \int_0^\infty df \frac{S_h(f) \gamma^2(f)}{M(f)} \left( 1 - \frac{N(f)}{M(f)} \right) \frac{\partial S_h(f)}{\partial \theta_a}, \quad a = 1, 2, 3, \quad i = 1, \ldots, N. \tag{A.15}
\]

Substituting (A.15) into (A.13) and (A.14) leads to (126) and (127).

References

Abbott, B. P., Abbott, R., Acernese, F., et al. 2009, Nature, 460, 990
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, Physical Review Letters, 116, 131102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016b, Physical Review Letters, 116, 241103
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016c, Physical Review Letters, 116, 061102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, Nature, 551, 85
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, Physical Review Letters, 118, 221101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, Physical Review Letters, 119, 161101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017d, Physical Review Letters, 118, 121101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2018, Physical Review Letters, 120, 091101
Abbott, L. F., & Wise, M. B. 1984, Nuclear Physics B, 244, 541
Abramovici, A., Althouse, W. E., Drever, R. W. P., et al. 1992, Science, 256, 325
Acernese, F., Amico, P., Al-Shourbagy, M., et al. 2005, Classical and Quantum Gravity, 22, S869
Adams, M. R., & Cornish, N. J. 2014, Phys. Rev. D, 89, 022001
Adams, M. R., Cornish, N. J., & Littenberg, T. B. 2012, Phys. Rev. D, 86, 124032
Allen, B. 1988, Phys. Rev. D, 37, 2078
Allen, B. 1997, in Relativistic Gravitation and Gravitational Radiation, eds. J.-A. Marck & J.-P. Lasota, 373 (arXiv:gr-qc/9604033)
Allen, B., & Romano, J. D. 1999, Phys. Rev. D, 59, 102001
Amaro-Seoane, P., Aoudia, S., Babak, S., et al. 2012, Classical and Quantum Gravity, 29, 124016
Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017, arXiv:1702.00786
Ananda, K. N., Clarkson, C., & Wands, D. 2007, Phys. Rev. D, 75, 123518
Armstrong, J. W., Estabrook, F. B., & Tinto, M. 1999, ApJ, 527, 814
Arnaud, K. A., Babak, S., Baker, J. G., et al. 2007, Classical and Quantum Gravity, 24, S551
Aso, Y., Michimura, Y., Somiya, K., et al. 2013, Phys. Rev. D, 88, 043007
Babak, S., Baker, J. G., Benacquista, M. J., et al. 2008, Classical and Quantum Gravity, 25, 114037
Ballantini, R., Bernard, P., Calatroni, S., et al. 2005, arXiv:0502054
Basko, M. M., & Polnarev, A. G. 1980a, Soviet Ast., 24, 268
Baumann, D., Steinhardt, P., Takahashi, K., & Ichiki, K. 2007, Phys. Rev. D, 76, 084019
Bender, P. L., & Hils, D. 1997, Classical and Quantum Gravity, 14, 1439
Bender, P., et al. 1998, LISA Pre-Phase A Report (second edi-
