Draft version July 1, 2024
Typeset using LaTeX default style in AASTeX631

Detection of Stochastic Gravitational Wave Background Based on Topological Data Analysis

Ali Salehi [1] and S. M. Sadegh Movahed [1, 2]

[1] Department of Physics, Shahid Beheshti University
P.O.Box 19839-69411, Tehran, Iran
[2] School of Astronomy
Institute for Research in Fundamental Sciences (IPM)
P.O.Box 19395-5531, Tehran, Iran

ABSTRACT

In this paper we will talk about stochastic gravitational wave background and introduce ways to detect it. Then we will introduce a novel approach called topological data analysis to calculate and detect it.

Keywords: IDK

1. INTRODUCTION

Most of the gravitational wave producing events are either too faint or too distant to be individually detected by the current interferometers, but the superposition of their gravitational radiation gives rise to a stochastic background that has a higher chance of detection. On top of that, there are other processes theorized to produce gravitational waves which are inherently stochastic, such as inflation (ref), cosmic strings (ref), and phase transitions (ref), etc.

Just by looking at the time series data recorded in the detectors, the various processes that give rise to a SGWB are essentially indistinguishable from oneanother and from noise. In order to detect and charactrize them, one needs to look at things such as their frequency dependency, their spacial distribution, etc. Or in our case, look at topological features of the data and compare them with the same feature of pure noise.

The study of topological features of data sets, is often called Topological Data Analysis (TDA) and is a rapidly growing area of applied math which is also finding more and more applications in physics each day. The general idea behind TDA is simple: each data (if represented correctly, more on that later) has a paticular ”shape” which can be charactrize with a set of topological variables. The topological variables are quite ressiliant against noise, so studying them can help us ”find” signals among a lot of noise and essentially increase our Signal to Noise Ratio (SNR).

In this paper, we first introduce our SGWB model and show the specral energy density plots. Then, we inject the simulated signals in detector noise and show the parameter estimation results with conventional ways. After that, we introduce TDA and build our model and show the results. In the last section we compare the results between the two methods and conclude.

The standard precedure to try and detect the SGWB is to cross-correlate the gravitational wave data from different interferometers to supress the noise each detector has. Then a search for a SGWB signal is run on the remaning data to either find the paramters (e.g. intensity, powerlaw index, ...) of the SGWB or put a restraint on it. So far, the only signs of a SGWB detected are in the nano hertz frequency band by the NANOGrav team Agazie et al. (2023).

No SGWB has been detected in the LIGO band (Abbott et al. (2021a,b, 2022)). The upper limits on Ω_{SGWB} is of order $\sim 10^{-9}$

2. MODEL

An isotropic background of gravitational waves is usually described by its spectral eneregy density and os explicitly dependent of frequency via

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

in which $\rho_{GW}(f)$ is the energy density of gravitational waves at ovserved frequency f, and $\rho_c = 3H_0^2/8\pi G$ is the critical density today. From this equation, we can also calculate that

$$\Omega_{SGWB}(f) = \frac{2\pi^2}{3H_0^2} f^3 S_h(f), \tag{2}$$

in which $S_h(f)$ is the one-sided power spectral density and can be calculted by averaging over fourier components of strain

$$\langle h_A(f, \hat{\mathbf{n}}) h^*_{A'}(f', \hat{\mathbf{n}}') \rangle_{\rm ens} = \frac{\delta_{AA'}}{2} \frac{\delta^{(2)}(\hat{\mathbf{n}}, \hat{\mathbf{n}}')}{4\pi} \frac{\delta(f - f')}{2} S_h(f), \tag{3}$$

where A represents the two polariztions of gravitational waves, $\hat{\mathbf{n}}$ is the spatial direction of the wave and δ is dirac delta.

Apart from frequency dependence, Ω_{SGWB} is implicitly dependent on many cosmological, astrophysical and source parameters which in the following we show with Θ in curly brakets

$$\Omega_{SGWB} = \Omega_{SGWB}(f, \{\Theta_{\rm source}, \Theta_{\rm astrophysics}, \Theta_{\rm cosmology}, \dots\}). \tag{4}$$

Generally, there are three ways to calculate Ω_{SGWB}: (1) For some simple situations, it is possible to find a mathematical expression for Ω (ref). (2) One could also cut out the middle man and run an accurate simulation for a large number of events and then superpose them and calculate the background (ref). (3) Lastly, there's another way which we dub the **Semi-Analytical** approach first laid out in (this paper). According to this approach, Ω_{SGWB} is represented by an integral over redshift and other implicit variables and one needs to take the integral to find the spectral density. The good thing about this approach is that one can add an arbitrary number of complications to it and since the resulting integral often needs to be taken numerically, we name it the semi-analytical approach.

There are a number of ways to show this semi-analytical integral. Here, we use a combination of (ref1, and ref2):

$$\Omega_{SGWB}(f, \{\Theta\}_{i=0}^n) = \frac{f}{\rho_c c^2} \int {\rm d}^n\Theta_i \int {\rm d}z \quad \overbrace{\frac{{\rm d}t_r}{{\rm d}z}}^{\text{cosmology}} \quad \underbrace{\frac{{\rm d}^m \tau_{\rm merg}(z, \Theta_j)}{{\rm d}\Theta_0 {\rm d}\Theta_1 \dots {\rm d}\Theta_m}}_{\substack{\text{astrophysics} \\ \text{\& cosmology}}} \overbrace{\frac{{\rm d}E_{\rm GW}(f_r, \Theta_k)}{{\rm d}f_r}}^{\text{GW source}}, \tag{5}$$

In which, $\{\Theta\}_{i=0}^n$ represets all the implicit variables which are integrated over and the indicies j and k run over the implicit parameters of merger rate and source energy respectivly. f_r is the frequency of emitted gravitational waves in the redshift of the source:

$$f_r = (1 + z) f. \tag{6}$$

2.1. *cosmology*

Additionally, the term $\frac{{\rm d}t_r}{{\rm d}z}$ in the 5, quanifies the amount of cosmic time which passes in an infinesimal redshift interval and is related to the Hubble parameter as

$$\frac{{\rm d}t_r}{{\rm d}z} = \frac{1}{(1 + z) H(z)}, \tag{7}$$

which itself is related to comological parameters as

$$H^2(z) = H_0^2 \left[\Omega_r (1 + z)^4 + (\Omega_{DM} + \Omega_b)(1 + z)^3 + \Omega_k (1 + z)^2 + \Omega_\Lambda \right]. \tag{8}$$

2.2. *merger rate*

$$\dot{N} = \frac{{\rm d}^2 \tau_{\rm merg}(z, m_1, m_2, \{\alpha_z, \alpha_c\})}{{\rm d}\log_{10} m_1 {\rm d}\log_{10} m_2} =$$

$$\underbrace{R_{\rm clust} \frac{(m_1 + m_2)^{10/7}}{(m_1 m_2)^{5/7}} \frac{(1 + z)^{\alpha_z}}{[1 + (M_{tot}/M_\star)]^{\alpha_c}}}_{\text{PBH clustering/binary formation}} \quad \overbrace{f_{PBH}(m_1) f_{PBH}(m_2)}^{\textit{primordial/thermal evolution}} \quad . \tag{9}$$

$$f_{PBH}(m, \{\sigma_{PBH}, \mu\}) = F_0 \frac{1}{\sqrt{2\pi} \sigma_{PBH} m} \exp\left[-\frac{(\log_{10} \frac{m}{\mu})^2}{2\sigma_{PBH}^2} \right], \tag{10}$$

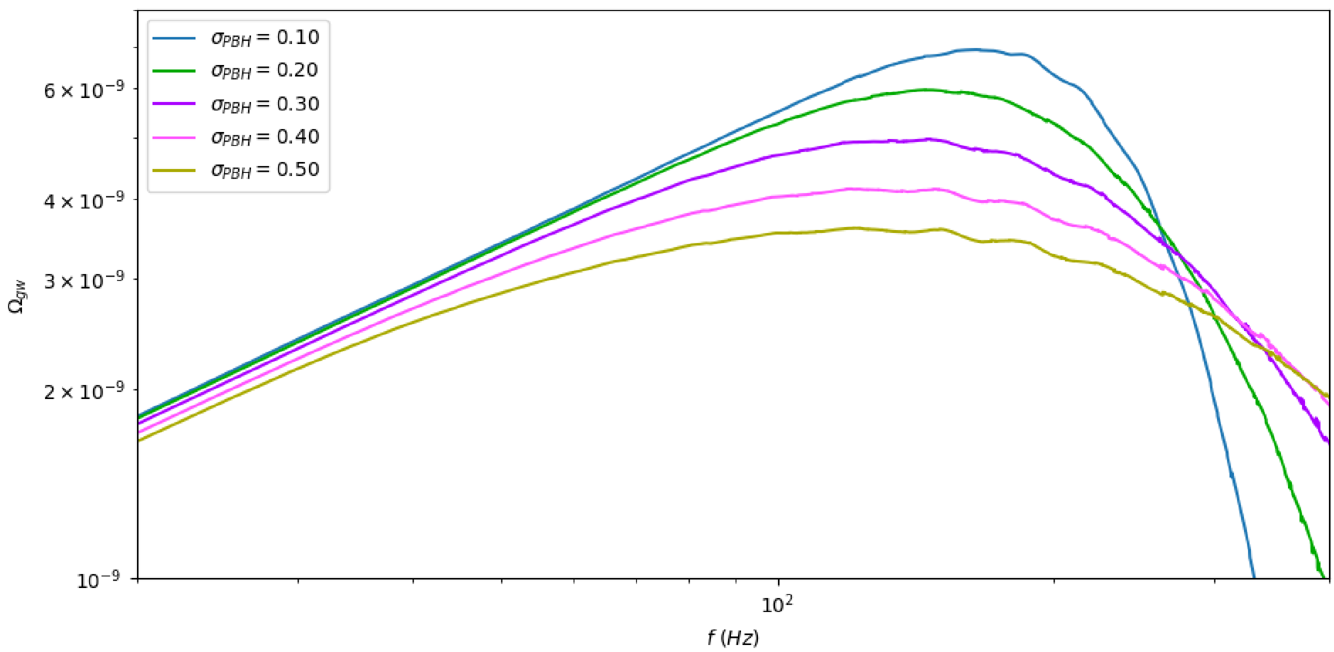


Figure 1. This plot shows Ω_{SGWB} versus f between the frequencies of $20 - 400$ Hz. All the plots follow a power law at first and then deviate from it based on σ_{PBH}. The smaller the parameter, the bigger the initial peak and the sharper the fall.

f_x	$a_1(\times 10^{-1})$	$a_2(\times 10^{-2})$	$a_3(\times 10^{-2})$
f_{merg}	2.9740	4.4819	9.5560
f_{ring}	5.9411	8.9794	19.111
f_{cut}	8.4845	12.848	27.299
f_w	5.0801	7.7515	2.2369

Table 1. parameters used at equation 13

2.3. *Source*

$$\frac{\mathrm{d}E_{\text{GW}}}{\mathrm{d}f_r} = \frac{(G\pi)^{2/3}\mathcal{M}_c^{5/3}}{3}\mathcal{G}(f_r), \tag{11}$$

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}. \tag{12}$$

$$\mathcal{G}(f_r) = \begin{cases} f_r^{-1/3} & f_r < f_{\text{merg}}, \\ \frac{f_r^{2/3}}{f_{\text{merg}}} & f_{\text{merg}} \leq f_r < f_{\text{ring}}, \\ \frac{1}{f_{\text{merg}} f_{\text{ring}}^{4/3}} \left(\frac{f_r}{1 + \left(\frac{f_r - f_{\text{ring}}}{f_w/2} \right)^2} \right)^2 & f_{\text{ring}} \leq f_r < f_{\text{cut}}. \end{cases} \tag{13}$$

3. SIGNAL INJECTION

To correctly asses the detection possibility of a simulated SGWB, one needs to have a realistic model for detector noise and embed signal into it before trying to test his method. Such a noise profile is called interferometers baseline noise and the method to embed a signal into it is often called signal injection. In order to do that, we use `pygwb` Renzini et al. (2023) which a python package heavily based on `Bilby` (Ashton et al. (2019)) that is optimized to help search for SGWB signals.

The way that we can use `pygwb` to our advantage is as follows: First we need some sort of signal to inject. Here we have the simulated spectral energy data that we explaind how to produce in the previous section. After that, we need to load interferometer objects from `Bilby`. these objects contain information about a particular gw detector such as: its location, response functions and most importantly for us, its baseline noise. The interferometers that we loaded for our analyses are: LIGO's detectors located at Hanford, Washington (H1) and Livingston, Louisiana (L1), Virgo detector located at (folan) Italy and finally a designed cosmic explorer detector located at H1.

When we load these detector, we add our desired SED to their baseline power and after choosing which combinations of detector we want for our analysis, we add them to a `network` object. We can then proceed to calculate cross correlation power spectrums for each pair of detectors and their point estimates. Finally, we run a parameter estimation pipeline to try and detect the injected SGWB.

Now that we have given a brief review on how `pygwb` package works, let us now lay out the mathematical framework upon which an estimate of Ω_{SGWB} is calculated for a network of interferometers. Here we show the formulae for two detectors I and J, these results can be generalized for more detectors as well. The estimation of Ω_{SGWB} and its varience at a frequency bin f are given by,

$$\hat{\Omega}_{\mathrm{GW},f} = \frac{\mathrm{Re}[C_{IJ,f}]}{\gamma_{IJ}(f) S_0(f)}, \tag{14}$$

and

$$\sigma_{\mathrm{GW},f} = \frac{1}{2T\Delta f} \frac{P_{I,f} P_{J,f}}{\gamma^2_{IJ}(f) S^2_0(f)}. \tag{15}$$

Following the notation in Renzini et al. (2023), in the equations above, the subscription f shows that the function in question is a discrete function of frequency as opposed to the continious functions shown in prenthesis. In the above also, $P_{I,f}$ are the discrete version of $S_h(f)$, $C_{IJ,f}$ is the cross-correlation spectrum, T is the total time of ovservation and Δf is the frequency resolution. S_0 is a conversion factor between strain and energy density defined as

$$S_0(f) = \frac{3H_0^2}{10\pi^2} \frac{1}{f^3}. \tag{16}$$

Finally, γ_{IJ} is called the overlap reduction function and is related to the detectors' response functions F_I^A as

$$\gamma_{IJ}(f) = \frac{5}{8\pi} \sum_A \int_{S^2} \mathrm{d}\hat{\mathbf{n}} F_I^A(f, \hat{\mathbf{n}}) F_J^A(f, \hat{\mathbf{n}}) e^{-i2\pi f, \hat{\mathbf{n}} \cdot (\mathbf{x}_I - \mathbf{x}_J)}, \tag{17}$$

where $\mathbf{x}_I - \mathbf{x}_J$ is the distance between the two detectors.

To give an idea on what is the response function, let's look at this equation:

$$d(t) = F(t) \star h(t) + n(t). \tag{18}$$

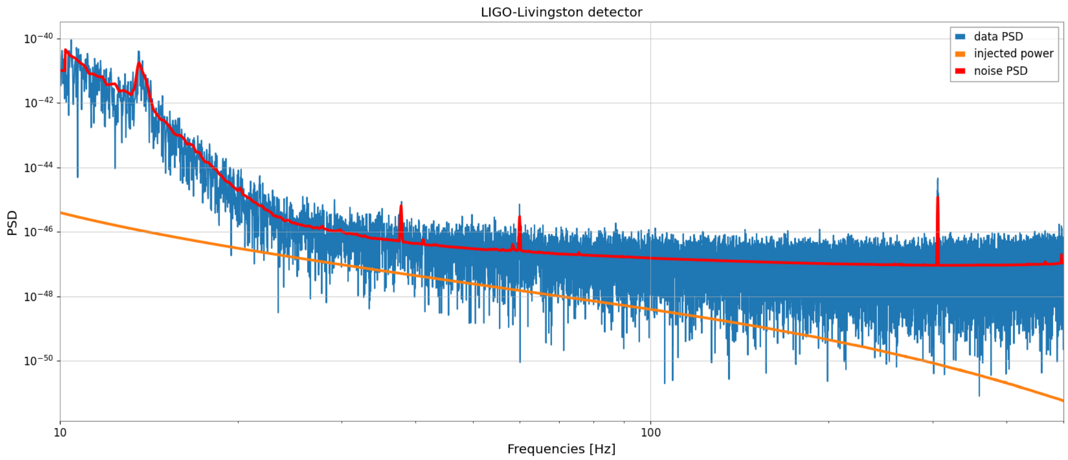
Here, $d(t)$ is the recorded data on our interferometers, $n(t)$ is the noise and $h(t)$ is the strain signal of incoming gravitational waves and $\star$ shows a convolution. In order to read $h(t)$ then, we need to first remove the noise which here is done by cross-correlating between the detectors, and then mathematically remove the effect of detectors response function which is done using the overlap reduction function.

4. TDA

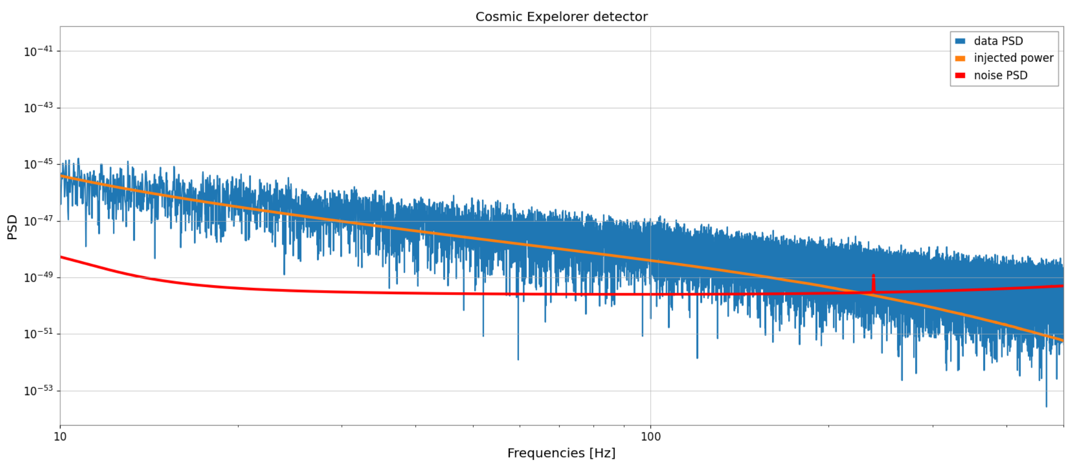
Here we propose a way of detecting SGWB which bypasses all the steps shown in section 3. Let's look back at equation 18: in the past section, we were trying to find algorithms to extract $h(t)$ from $d(t)$. Here we are using a more indirect approach. We are aiming to find another type of algorithm which can look for a topological stucture inside the data which is fed into it i.e. $d(t)$. If the noise is not too much, the method of persistence homology will be able to detect the traces of that topological structure and return us parameters which then we can use to recustruct the signal.

Unlike the standard method laid out in the last section, the values that persistence homology returns, are not simple algebraic functions of the input but with a little bit of help from machine learning, we will be able to train a model that can take in the topological variables and recunstruct the signal.

This method however is more resilliant to noise and hence it can help us increase the SNR for gravitational wave searches. A study on how to exactly do that is under way and will be published in another paper. Here we only wish to comapre the to methods and not combine them.



(a)



(b)

Figure 2. Top: Signal as seen in LIGO-Hanford detector. Bottom: Data as seen in a detector with cosmic expelorer design at the place of Hanford detector.

5. RESULTS AND CONCLUSION

REFERENCES

Abbott, R., Abbott, T., Abraham, S., et al. 2021a, Physical Review D, 104, 022004

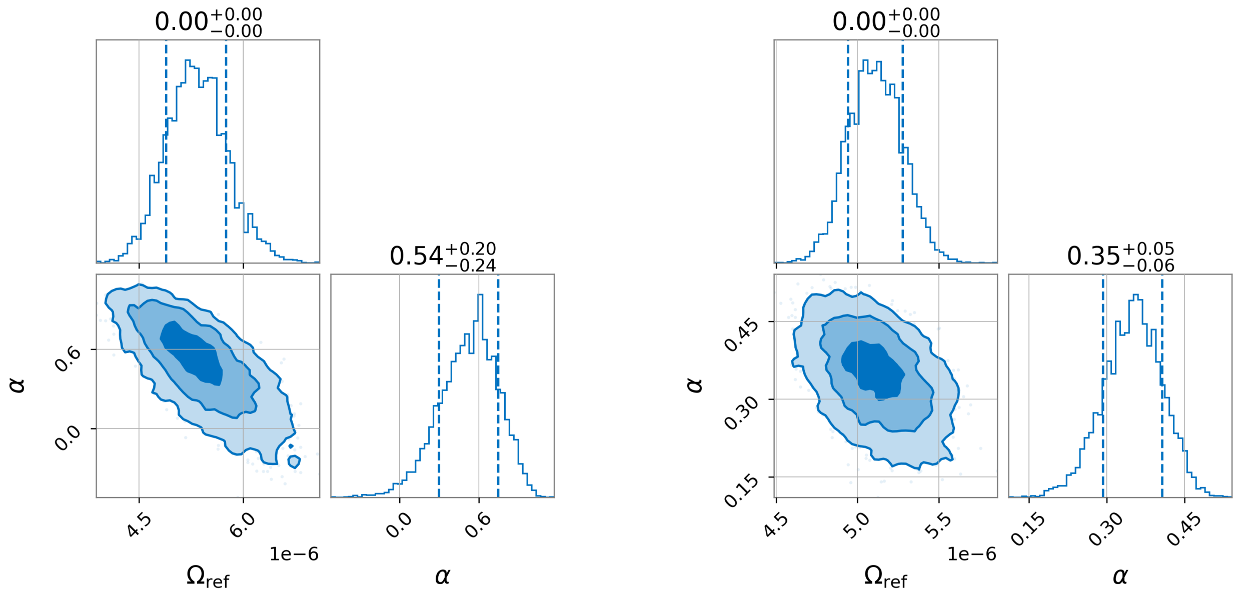
—. 2021b, Physical Review D, 104, 022005

Abbott, R., Abbott, T., Acernese, F., et al. 2022, Physical Review D, 105, 122001

Agazie, G., Anumarlapudi, A., Archibald, A. M., et al. 2023, The Astrophysical Journal Letters, 951, L8

Ashton, G., Hübner, M., Lasky, P. D., et al. 2019, The Astrophysical Journal Supplement Series, 241, 27

Renzini, A. I., Romero-Rodríguez, A., Talbot, C., et al. 2023, The Astrophysical Journal, 952, 25



(a) (b)

Figure 3. (a) Parameter estimation for a network of detectors: Hanford, Livingston and Virgo. (b) Parameter estimation while replacing the Hanford detector with a cosmic explorer detector.