Searching for Echoes of Gravitational Waves from the Coalescence of Exotic Compact Objects: A Bayesian Approach

Rico Ka Lok Lo
Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR

Mentor: Alan Weinstein
LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, US
(LIGO Scientific Collaboration)
(Dated: September 20, 2017)

The ringdown part of the gravitational wave from the merger of two black holes was suggested as a probe of the internal structure of the remnant compact object, which may be more exotic than a black hole. Cardoso et al. pointed out that there would be a train of echoes in the late-time ringdown phase for different types of exotic compact objects (ECOs). Abedi et al. claimed that they have found evidence of echoes in binary black hole mergers detected by LIGO. In this project, we aim to search for echoes of gravitational waves in the three detections LIGO made to date, and verify their results using their phenomenological model with Bayesian analysis instead. We perform a Bayesian parameter estimation on the parameters related to echoes and Bayesian model selection of presence of echoes versus their absence, to provide stronger evidence for the presence or absence of echoes in these detections. The analysis technique developed in this project could be repeated with different models to provide even more robust evidence of the existence of echoes from ECOs.

LIGO DCC Number: LIGO-T1700322

I. BACKGROUND

As of this writing, the Laser Interferometer Gravitational-wave Observatory (LIGO) has successfully detected three compact binary coalescence events: GW150914, GW151226, GW170104 [1–3], and one candidate event LVT151012 [4]. These discoveries mark the beginning of a new era of gravitational wave astronomy and astrophysics, where we can infer and probe the properties and structure of astronomical objects using gravitational-wave signals.

During the inspiral phase of gravitational wave emission from the coalescence of a compact binary system, for instance a binary black hole system, the two black holes spiral towards each other with an increasing orbital frequency. Eventually, they coalesce in the merger phase to form one single black hole. The final black hole then relaxes to a stationary Kerr black hole during the ringdown phase. Figure 1 shows the numerical relativity simulation and the reconstructed template of the first detection.

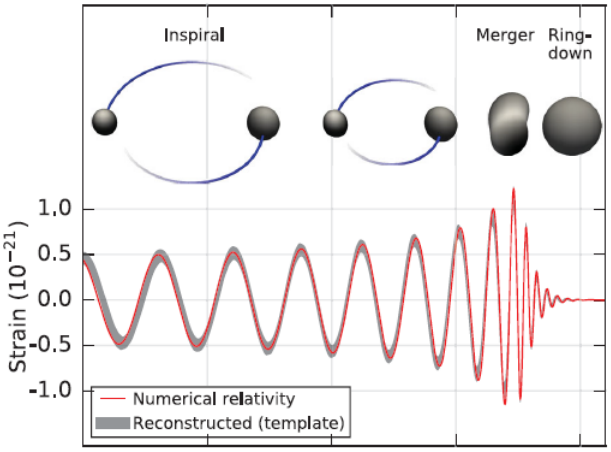


Figure 1: The numerical relativity simulation (in red) and reconstructed waveform (in gray) of GW150914. Figure taken from [2].

Cardoso, Franzin and Pani [5] first pointed out that the ringdown part of the gravitational wave can be used as a probe of the structure of a compact object. A very compact object, not necessary a black hole, with light ring will also exhibit similar ringdown as that of a black hole. Cardoso, Hopper, Macedo, Palenzuela and Pani [6] further showed that similar ringdown stage will also be exhibited for different types of **exotic compact objects** (ECOs) with light ring (or photon sphere), and there will be a train of echoes in the late-time ringdown phase associated with the photon sphere. Examples of ECOs are theoretical alternatives to black holes, such as gravastars[7] and fuzzballs. A common feature of these alternatives is that there is some kind of structure near the would-be event horizon. The echoes in the late-time ringdown phase are caused by repeated and damped reflections between the angular momentum barrier and the reflective structure. Figure 2 shows an example of the aforementioned echoes from a radially infalling particle into a wormhole in the ringdown phase. Throughout the project, we will use geometrized units, with $c = G = 1$.

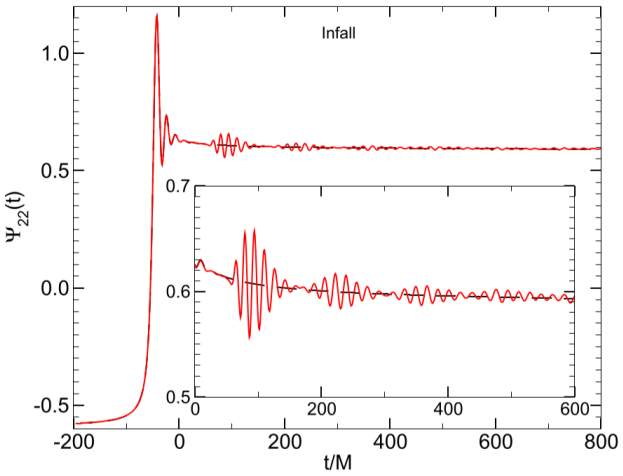


Figure 2: The waveform of a radially infalling particle into a wormhole (in red) compared to that into a black hole (in black). We can clearly see the train of echoes for the wormhole case, whereas this feature is absent for black hole. Figure taken from [6].

A. Previous attempts to search for echoes of gravitational waves

1. *Search and phenomenological model for echoes by Abedi et al.*

Abedi, Dykaar and Afshordi published a paper on December 2016, claiming that they have found evidence of Planck-scale structure near the black hole event horizons at a combined 2.9σ significance level [8] on GW150914, LVT151012 and GW151226 using matched filtering technique.

In their search, they proposed a phenomenological 'toy' model of echoes with 5 free parameters, with the phase change between each echo being π. The description of these 5 parameters are tabulated in Table I.

Parameter	Description
$\Delta t_{\rm echo}$	The time interval between each echo
$t_{\rm echo}$	The time of arrival of the first echo
t_0	The time of truncation of the GW IMR template $\mathcal{M}_{\rm I}(t)$ to produce the echo template $\mathcal{M}_{\rm TE,I}(t)$
γ	The damping factor
A	The amplitude of the echoes relative to the IMR template

Table I: The five free parameters of templates used by Abedi et al. [8].

Using the notations in [8], the echo template $\mathcal{M}_{\rm TE,I}(t)$ in time-domain is given by

$$\mathcal{M}_{\rm TE,I}(t) \equiv A\sum_{n=0}^{\infty}(-1)^{n+1}\gamma^n \mathcal{M}_{\rm T,I}(t)(t+t_{\rm merger}-t_{\rm echo}-n\Delta t_{\rm echo}, t_0), \tag{1}$$

where $t_{\rm merger}$ is the time of merger [9] and $\mathcal{M}_{\rm T,I}(t)$ is a smooth activation of the GW IMR[10] template given by

$$\begin{aligned}\mathcal{M}_{\rm T,I}(t) &\equiv \Theta(t,t_0)\mathcal{M}_{\rm I}(t)\\ &\equiv \frac{1}{2}\left\{1+\tanh\left[\frac{1}{2}\omega_{\rm I}(t)(t-t_{\rm merger}-t_0)\right]\right\}\mathcal{M}_{\rm I}(t).\end{aligned}$$

Here $\omega_{\rm I}(t)$ denotes the angular frequency evolution of the IMR waveform as a function of time, and $\mathcal{M}_{\rm I}(t)$ is the IMR waveform without echo. The smooth activation essentially selects the ringdown, which is the part of the waveform that one might expect to survive in the echoes [8].

A team of physicists in Max Planck Institute for Gravitational Physics (AEI) questioned the data analysis method in Abedi et al.'s paper[11], especially the method to estimate the background and hence the statistical significance of the evidence they claimed. Abedi et al. responded that the shortcomings pointed by them would only change the significance by less than 0.3σ [12].

2. *Phenomenological model by Nakano et al.*

Motivated by solving the linear perturbation with a completely reflective boundary condition around a Kerr black hole, Nakano, Sago, Tagoshi and Tanaka suggested a refined phenomenological template of echoes of gravitational wave [13]. In frequency-domain, the template of the $n^{\rm th}$ echo $\tilde{h}_n(f)$ is given by

$$\tilde{h}_n(f) = e^{-i(2\pi f\Delta t+\phi(f))(n-1)}(\sqrt{R(f)})^{n-1}\sqrt{1-R(f)}\quad \tilde{h}(f), \tag{2}$$

where $\tilde{h}(f)$ is the Fourier transform of the waveform that reflects off the surface of ECO and $\phi(f)$ is the phase change when the gravitational wave was reflected at the potential barrier and at the boundary near horizon. The reflection rate, denoted by $\sqrt{R}$ in Eq.2, or reflection amplitude $\tilde{\mathcal{R}}_{\rm BH}$ (later defined in Eq.7), can be estimated using the fitting formulas in [13].

They also proposed two simple models of $h(t)$, the waveform that reflects off the surface of ECO, that we have just mentioned. The first model is given by

$$h_1(t) \propto \frac{e^{-i\omega_{\rm QNM}\tilde{t}}}{1+e^{-2\beta|\Im(\omega_{\rm QNM})|t}}, \tag{3}$$

where $\beta \ll 1$ is a model parameter to eliminate the high frequency tail [13] due to the implicit periodic assumption when using Discrete Fourier Transform, and $\tilde{t}$ gives

the frequency evolution, which is given by

$$\frac{d\tilde{t}}{dt} = \frac{1}{1 + e^{-2\alpha|\Im(\omega_{\rm QNM})|t}}. \quad (4)$$

As for the second model, it is given by

$$h_2(t) \propto \frac{e^{-i\omega_{\rm QNM}t}}{1 + e^{-2\bar{\alpha}|\Im(\omega_{\rm QNM})|t}}. \quad (5)$$

The two parameters α and $\bar{\alpha}$ are for controlling the excitation of quasi-normal mode at QNM frequency $\omega_{\rm QNM}$, in which the real part $\Re(\omega_{\rm QNM})$ is the oscillation frequency, and the imaginary part $\Im(\omega_{\rm QNM}) < 0$ is the inverse of the exponential damping time [13]. It should be noted that Nakano et al. claimed that the templates and reprocessing they proposed in [13] work in the context of gravitational-waves, namly

$$h(t) = h_+(t) + ih_\times(t),$$

where $h_+(t)$ and $h_\times(t)$ are the plus and cross polarization of gravitational waves respectively.

3. Phenomenological model by Mark et al.

Recently, Mark, Zimmerman, Du and Chen published a paper about the echoes from ECOs [14]. They modeled an exotic compact object with a Schwarzschild metric in the exterior and a spherically symmetric metric in the interior, together with a reflective boundary condition on the ECO surface $x = x_0$ with a reflection amplitude $\tilde{\mathcal{R}}(\omega)$. Here x is the tortoise coordinate defined as

$$x = r + 2M\ln(\frac{r-2M}{M}).$$

They numerically calculated the scalar waves as observed by distant observers. They argued that the echo waveforms can be modulated from the waveform on the black hole horizon when there is no reflecting surface [14]. As such, they proposed a phenomenological template $Z_{\rm T}$ of the echoes from ECOs in frequency domain

$$Z_{\rm T} = \tilde{K} Z^{\rm H}_{\rm T}, \quad (6)$$

where $Z^{\rm H}_{\rm T}$ is the horizon waveform template in frequency domain and $\tilde{K}$ is the transfer function, which contains most of the physics in this model. The transfer function $\tilde{K}(\omega)$, as a function of angular frequency ω, is given by

$$\tilde{K}(\omega) = \frac{\tilde{\mathcal{T}}_{\rm BH}\tilde{\mathcal{R}}e^{-2i\omega x_0}}{1 - \tilde{\mathcal{R}}_{\rm BH}\tilde{\mathcal{R}}e^{-2i\omega x_0}}, \quad (7)$$

where $\tilde{\mathcal{T}}_{\rm BH}(\omega)$, $\tilde{\mathcal{R}}_{\rm BH}(\omega)$ are the transmission amplitude and reflection amplitude of the black hole respectively, using the notations in [14]. Note that

$$|\tilde{\mathcal{T}}_{\rm BH}|^2 + |\tilde{\mathcal{R}}_{\rm BH}|^2 = 1.$$

We can expand Eq.7 as a geometric series:

$$\tilde{K}(\omega) = \tilde{\mathcal{T}}_{\rm BH}\tilde{\mathcal{R}}e^{-2i\omega x_0}\sum_{n=1}^{\infty}(\tilde{\mathcal{R}}_{\rm BH}\tilde{\mathcal{R}})^{n-1}e^{-2i(n-1)\omega x_0}. \quad (8)$$

If we define the transfer function for the $n^{\rm th}$ echo as

$$\tilde{K}^{(n)} = (\tilde{\mathcal{T}}_{\rm BH}\tilde{\mathcal{R}})(\tilde{\mathcal{R}}_{\rm BH}\tilde{\mathcal{R}})^{n-1}e^{-2in\omega x_0}, \quad (9)$$

then we can write Eq.6 as a sequence of echoes

$$Z_{\rm T} = \sum_{n=1}^{\infty} Z^{(n)}_{\rm T} = \sum_{n=1}^{\infty} \tilde{K}^{(n)} Z^{\rm H}_{\rm T}.$$

The overall waveform observed by distant observers $Z^{\infty}_{\rm ref}$ is given by

$$Z^{\infty}_{\rm ref} = Z^{\infty}_{\rm BH} + \tilde{K} Z^{\rm H}_{\rm BH}.$$

Therefore, to generate a template for searching echoes, one just needs to add the reprocessed horizon template to the BH template in the frequency-domain.

Mark et al. observed that the ringdown of the horizon waveform determined the shape of the echoes. Hence, to get these features correct, Mark et al. proposed to use the following model [14]:

$$Z^{\rm H}_{\rm T}(\omega) = e^{i\omega t_s}e^{-\omega^2/(2\beta^2)}\left(\frac{\alpha_+}{\omega - \Omega_+} + \frac{\alpha_-}{\omega - \Omega_-}\right), \quad (10)$$

the meaning of the template parameters are tabulated in Table II. Together with the location of the surface of ECO x_0, there are a total of five free parameters in this echo model.

Parameter	Description
α_+	The complex amplitude of the sinusoid at positive QNM frequency $\Omega_+ = \Omega_R + i\Omega_I$
α_-	The complex amplitude of the sinusoid at negative QNM frequency $\Omega_- = -\Omega_R + i\Omega_I$
t_s	The central start time of the Gaussian
β	The frequency width of the Gaussian

Table II: The four free parameters of horizon templates used by Mark et al. [14].

B. Bayesian Analysis and Hypothesis Testing

Bayesian analysis provides us a consistent framework of inductive logic. The central piece of the Bayesian framework is the **Bayes' Theorem**, which allows us to update our degree of belief of a hypothesis H with a prior knowledge I when a new data d is available.

1. Bayes' Theorem

Mathematically, Bayes' theorem (also known as Bayes' rule) states that

$$\Pr(B|A) = \frac{\Pr(A|B)\Pr(B)}{\Pr(A)}, \tag{11}$$

where $\Pr(A)$ and $\Pr(B)$ denote the (discrete) probability of event A and event B occurs, regardless of each other, respectively; $\Pr(A|B)$ is the probability of event A **on the condition that** that event B occurs and likewise for $\Pr(B|A)$.

2. Parameter Estimation

Suppose we have a generative model H that describes the observed data d given some parameters $\vec{\theta} = (\theta_1, \theta_2, ..., \theta_n)$. Before the observation, suppose we already have some knowledge I about the parameters $\vec{\theta}$ assuming that the model H is true, represented by the **prior probability distribution** $p(\vec{\theta}|H, I)$, we can then use Eq.11 to update our knowledge or degree of belief of the value of these parameters to get the **posterior probability distribution** $p(\vec{\theta}|d, H, I)$ as follows:

$$p(\vec{\theta}|d, H, I) = \frac{p(d|\vec{\theta}, H, I)p(\vec{\theta}|H, I)}{p(d|H, I)} \tag{12}$$

$$= \frac{p(d|\vec{\theta}, H, I)p(\vec{\theta}|H, I)}{\int_\theta p(d|\vec{\theta'}, H, I)p(\vec{\theta'}|H, I)d\theta'}. \tag{13}$$

The term $p(d|\vec{\theta}, H, I)$, when viewed as a function of the parameters $\vec{\theta}$, is called the **likelihood** function and denoted as $\mathcal{L}(\vec{\theta}|d, H, I)$. As for the term $p(d|H, I)$, it is called the **evidence**, or marginal likelihood. In this scenario, it can be considered as a normalization constant and often ignored (or, computed numerically so that the LHS is a normalized PDF). Then, Eq.12 can be turned into:

$$p(\vec{\theta}|d, H, I) \propto p(d|\vec{\theta}, H, I)p(\vec{\theta}|H, I). \tag{14}$$

a. Parameter estimation in the context of GW

Suppose that the strain data $d(t)$ from a detector only consists of noise $n(t)$, which we assume to be Gaussian and stationary[15]. The probability that the noise $n(t)$ has a realization $n_0(t)$ (with zero mean) is given by [16]:

$$p(n_0) = \mathcal{N} \exp\left\{-\frac{1}{2}\int_{-\infty}^{+\infty} df \frac{|\tilde{n_0}(f)|^2}{(1/2)S_n(f)}\right\}, \tag{15}$$

where $\mathcal{N}$ is a normalization constant and $S_n(f)$ is the power spectrum density of noise. If we introduce the notion of **noise-weighted inner product**, namely

$$(A|B) = 4\Re \int_0^\infty df \frac{\tilde{A}^*(f)\tilde{B}(f)}{S_n(f)}, \tag{16}$$

then we can rewrite Eq. 15 into

$$p(n_0) = \mathcal{N} \exp\left\{-\frac{1}{2}(n_0|n_0)\right\}. \tag{17}$$

Suppose the strain data $d(t)$ consists of both noise $n_0(t)$ and GW signal $h(t; \vec{\theta})$, namely

$$n_0(t) = d(t) - h(t; \vec{\theta}),$$

then the the likelihood $p(d|\vec{\theta}, H, I)$ for single detector can be obtained from Eq. 15:

$$p(d|\vec{\theta}, H, I) = \mathcal{N} \exp\left\{-\frac{1}{2}(d(t) - h(t; \vec{\theta})|d(t) - h(t; \vec{\theta}))\right\}. \tag{18}$$

For the case of multiple detectors (for example, H1 and L1), if we assume that the noise distributions are all Gaussian and stationary, and more importantly independent of each other, then we have

$$p(d_{\rm H1}, d_{\rm L1}|\vec{\theta}, H, I) = \prod_{i\in\{H1,L1\}} \mathcal{N}_i \exp\left\{-\frac{1}{2}(d_i(t_i) - h(t_i; \vec{\theta})|d_i(t_i) - h(t_i; \vec{\theta}))\right\}. \tag{19}$$

For computational reasons, it is often easier to deal with **log prior**, **log likelihood** and **log posterior**. In particular, Eq. 12 for posterior distribution becomes

$$\ln p(\vec{\theta}|d, H, I) = \ln p(d|\vec{\theta}, H, I) + \ln p(\vec{\theta}|H, I) - \ln p(d|H, I). \tag{20}$$

If we substitute the likelihood function into Eq. 20, we have

$$\ln p(\vec{\theta}|d, H, I) = -\frac{1}{2}(d(t) - h(t; \vec{\theta})|d(t) - h(t; \vec{\theta})) + \ln p(\vec{\theta}|H, I) + \ln K, \tag{21}$$

where K is a constant independent of $\vec{\theta}$.

After the posterior distribution was sampled, for example by means of Markov Chain Monte Carlo discussed

in Section I C, we can calculate various estimators as point estimates, such as maximum likelihood estimator (MLE), which is

$$\hat{\vec{\theta}}_{\rm MLE} = \arg\max \mathcal{L}(\vec{\theta}|d, H, I).$$

Another estimator is maximum a posteriori estimator (MAP), which is

$$\hat{\vec{\theta}}_{\rm MAP} = \arg\max p(\vec{\theta}|d, H, I).$$

Note that these two estimators would change if one marginalizes over some parameters. On the other hand, the Bayes estimator is invariant to marginalization, namely

$$\hat{\vec{\theta}}_{\rm Bayes} = \int \vec{\theta} p(\vec{\theta}|d, H, I) d\vec{\theta}.$$

Note that Bayes estimator is simply the mean of the posterior probability distribution.

b. Bayesian Credible Interval

In addition to obtaining point estimates, we would also like to estimate the precision of these inferred values. This can be done by finding the minimum **Bayesian Credible Interval** of given credibility $(1-\alpha)\times 100\%$. It is simply the *shortest* interval that contains a fraction of $1-\alpha$ of the posterior samples, or equivalently that there is a probability of α that the parameter lies outside the interval. Mathematically, the Bayesian credible interval $[\theta_{\rm lower}, \theta_{\rm upper}]$ is the interval that satisfies

$$1-\alpha = \int_{\theta_{\rm lower}}^{\theta_{\rm upper}} p(\theta|d, H, I) d\theta, \tag{22}$$

where $\theta_{\rm upper} - \theta_{\rm lower}$ is minimized.

3. Hypothesis Testing

Suppose we have two (generative) models H_0 and H_1, and we want to compare which model describes the observed data the best. It is useful to compute the **odds ratio** defined as

$$O_{H_0}^{H_1} = \frac{p(H_1|d, I)}{p(H_0|d, I)}. \tag{23}$$

Applying Eq.11 to Eq.23, we have

$$O_{H_0}^{H_1} = \frac{p(d|H_1, I)p(H_1|I)}{p(d|H_0, I)p(H_0|I)}.$$

The factor $\dfrac{p(H_1|I)}{p(H_0|I)}$ is called the **prior odds** and the factor $\dfrac{p(d|H_1, I)}{p(d|H_0, I)}$ is called the **Bayes factor**, denoted as $B_{H_0}^{H_1}$.

A question that arises naturally is that how can we compute the term $p(d|H_i, I)$, where $i = 0, 1$. Suppose H_0 has no free parameter, whereas H_1 has n free parameters $\vec{\theta} = (\theta_1, \theta_2, ..., \theta_n)$. Then it is straightforward to calculate $p(d|H_0, I)$. But for $p(d|H_1, I)$, by re-arranging Eq.11 we have

$$p(\vec{\theta}|d, H_1, I)p(d|H_1, I) = p(d|\vec{\theta}, H, I)p(\vec{\theta}|H_1, I).$$

Integrating both sides with respect to $\vec{\theta}$ (or *marginalizing over the parameters*), we have

$$\int p(\vec{\theta}|d, H_1, I)p(d|H_1, I)d\vec{\theta} = \int p(d|\vec{\theta}, H_1, I)p(\vec{\theta}|H_1, I)d\vec{\theta}.$$

Note that $p(d|H_1, I)$ is independent of $\vec{\theta}$ and by definition

$$\int p(\vec{\theta}|d, H_1, I)d\vec{\theta} = 1,$$

We finally arrive at

$$p(d|H_1, I) = \int p(d|\vec{\theta}, H_1, I)p(\vec{\theta}|H_1, I)d\vec{\theta}. \tag{24}$$

When the odds ratio $O_{H_0}^{H_1} > 1$, that means the observed data favors (alternative) hypothesis H_1 more than the hypothesis H_0 and vice versa.

One thing that we must be cautious about is that a model that fits the data best does not imply the model gives the highest evidence. A more complicated model (i.e. with more free parameters) often subjects to more noise than a simpler model (i.e. with less free parameters). This is similar to over-fitting in regression. When you have N data points for fitting, you can always use a degree N polynomial to fit all points, but very likely the fitted polynomial will not generalize well to new data because it was affected by the noise in the data.

Bayesian analysis embodies the **Occam's razor** and penalizes more complicated models automatically. The following example was taken from [17]. Suppose there are two models H_0 and H_1 with H_0 having no free parameter and H_1 having 1 free parameter θ. Assume that there is no prior knowledge to favor a particular model, i.e. the prior odds is simply 1. We further assume that the prior probability distribution for model H_1 is constant over a range $\theta \in [\theta_{\rm min}, \theta_{\rm max}]$, namely

$$p(\theta|H_1, I) = \begin{cases} \dfrac{1}{\theta_{\rm max} - \theta_{\rm min}}, & \theta \in [\theta_{\rm min}, \theta_{\rm max}] \\ 0, & \text{otherwise} \end{cases}.$$

We also assume that the likelihood to be a Gaussian function, centered around the maximum likelihood when $\theta = \theta_0$, with a standard deviation of σ_θ. Mathematically,

$$p(d|\theta, H_1, I) = p(d|\theta_0, H_1, I)e^{-\frac{(\theta-\theta_0)^2}{2\sigma_\theta^2}}.$$

The odds ratio can be found easily by integration:

$$O_{H_0}^{H_1} = \frac{p(d|\theta_0, H_1, I)}{p(d|H_0, I)} \frac{\sqrt{2\pi}\sigma_\theta}{\theta_{\rm max} - \theta_{\rm min}}.$$

Very often, the first factor $\frac{p(d|\theta_0, H_1, I)}{p(d|H_0, I)} > 1$, since often a more complicated model fits the noisy data better than a simpler model. However, the second factor suggests that if the parameter θ is unnecessary to describe the data in a sense that the width of the likelihood function is much smaller than the width of the prior, i.e. $\sigma_\theta \ll \theta_{\max} - \theta_{\min}$, the second factor will penalize H_1 in a sense that the odds ratio will become smaller.

C. Markov Chain Monte Carlo (MCMC)

In Section I B 2 and I B 3, we outlined the procedures of performing parameter estimation and hypothesis testing respectively. To perform the actual Bayesian analysis, we would need to sample from the posterior distribution $p(\vec{\theta}|d, H, I)$. Markov Chain Monte Carlo (MCMC) is the standard tool to complete this task.

A Markov chain of random variables $\{x_1, x_2, ..., x_N\}$ satisfies the property

$$\Pr(x_{i+1}|x_1, x_2, ..., x_i) = \Pr(x_{i+1}|x_i).$$

That is, the probability of transiting from x_i to x_{i+1} depends on the current state x_i only. Suppose we want to draw samples from the distribution $\pi(x)$. To make sure that the Markov chain will converge to the target distribution $\pi(x)$, the target distribution needs to be stationary, i.e.

$$\pi(x') = \int \pi(x) q(x, x') dx, \tag{25}$$

and detailed balance, i.e.

$$\pi(x) q(x, x') = \pi(x') q(x', x), \tag{26}$$

where $q(x, x')$ is called the transition kernel, and

$$\Pr(x_i|x_1, x_2, ..., x_{i-1}) = \int q(x_{i-1}, x) dx.$$

If the posterior distribution that we desire satisfies the above conditions, then we are in a good shape to utilize MCMC.

1. Metropolis and Metropolis-Hasting algorithm

a. Metropolis algorithm

Metropolis algorithm is the simplest and earliest algorithm for MCMC, proposed by Metropolis and Ulam [18] in 1949. Here we outline the Metropolis algorithm. For the sake of simplicity, we denote the posterior distribution as $\mathcal{P}$.

1. Initialize $\vec{\theta}_{i=0}$
2. Propose a jump $\delta\vec{\theta} = (\theta_{\text{proposed},n} - \theta_{i-1,n})$ according to the proposal distribution, i.e. $\delta\theta_n \sim \Lambda_n$
3. Accept the jump $\vec{\theta}_i = \vec{\theta}_{\text{proposed}}$ with a probability of $\min(\frac{\mathcal{P}_{\text{proposed}}}{\mathcal{P}_{i-1}}, 1)$. Else set $\vec{\theta}_i = \vec{\theta}_{i-1}$
4. Repeat step 2 & 3 for a sufficient amount of steps
5. Discard early steps (also known as burn-in)

Note that there are hyperparameters controlling the proposal distribution. A common choice of the proposal distribution is a multivariate Gaussian distribution, i.e.

$$\Lambda_n = N(0, \sigma_n^2),$$

where σ_n is the standard deviation of the distribution in the n^{th} dimension and it is a hyperparameter of the MCMC algorithm. Albeit its simplicity, the power of Metropolis algorithm is limited in a sense that the proposal distribution Λ_i in this algorithm cannot be asymmetric, which can be problematic when sampling anisotropic distributions. Also the algorithm converges slowly. To solve these problems, we can use the updated version called Metropolis-Hasting algorithm.

b. Metropolis-Hasting algorithm

Metropolis-Hasting algorithm is very similar to Metropolis algorithm, proposed by Hasting [19] in 1970. However, this algorithm does not require a symmetric proposal distribution. Here we outline the Metropolis-Hasting algorithm.

1. Initialize $\vec{\theta}_{i=0}$
2. Propose a jump $\delta\vec{\theta} = (\theta_{\text{proposed},n} - \theta_{i-1,n})$ according to the proposal distribution, i.e. $\delta\theta_n \sim \Lambda_n$
3. Accept the jump $\vec{\theta}_i = \vec{\theta}_{\text{proposed}}$ with a probability of $\min(\frac{\mathcal{P}_{\text{proposed}}}{\mathcal{P}_{i-1}} \frac{\Lambda(\vec{\theta}_{\text{proposed}}|\vec{\theta}_{i-1})}{\Lambda(\vec{\theta}_{i-1}|\vec{\theta}_{\text{proposed}})}, 1)$. Else set $\vec{\theta}_i = \vec{\theta}_{i-1}$
4. Repeat step 2 & 3 for a sufficient amount of steps
5. Discard early steps (also known as burn-in)

A common issue with Metropolis/Metropolis-Hasting algorithm is that it cannot sample multi-modal distributions well. This problem can be overcome by parallel tempering, or replica exchange.

2. Parallel tempering

Parallel tempering, or replica exchange, was originally proposed by Swendsen and Wang [20] in 1986. The idea is that we have N different chains running MCMC at different temperatures T in parallel with slightly modified likelihood function

$$p(d|\vec{\theta}, H, I, \beta) = \Big[p(d|\vec{\theta}, H, I)\Big]^\beta, \tag{27}$$

where $\beta \equiv \frac{1}{T}$. Note that $T = 1$ chain must be present, since this is the chain that contains the desired posterior samples.

During the run, chains of different temperatures swap samples with some criteria [20]. This is desirable because the hotter chain has a flatter likelihood, and hence it can explore the parameter space more easily, whereas the cooler chain explores a local maximum well. This can help us sample multi-modal distributions. By adopting parallel tempering sampler in this project, we can estimate the evidence, which is essential to model selection, using thermodynamic integration.

a. Thermodynamic integration

From Eq. 24 that the evidence is given by

$$p(d|H_1, I) = \int p(d|\vec{\theta}, H_1, I)p(\vec{\theta}|H_1, I)d\vec{\theta}.$$

When performing the thermodynamic integration, derived in [21], we introduce a new quantity which depends on β:

$$p(d|H_1, I, \beta) = \int [p(d|\vec{\theta}, H_1, I)]^\beta p(\vec{\theta}|H_1, I)d\vec{\theta}. \quad (28)$$

Consider

$$\begin{aligned}\frac{d\ln p(d|H_1, I, \beta)}{d\beta} &= \frac{d\ln p(d|H_1, I, \beta)}{dp(d|H_1, I, \beta)}\frac{dp(d|H_1, I, \beta)}{d\beta} \\ &= \frac{1}{p(d|H_1, I, \beta)}\times \\ &\frac{d}{d\beta}\int [p(d|\vec{\theta}, H_1, I)]^\beta p(\vec{\theta}|H_1, I)d\vec{\theta} \\ &= \frac{1}{p(d|H_1, I, \beta)}\times \\ &\int \ln p(d|\vec{\theta}, H_1, I)[p(d|\vec{\theta}, H_1, I)]^\beta \times \\ &p(\vec{\theta}|H_1, I)d\vec{\theta} \\ &= \left\langle \ln p(d|\vec{\theta}, H_1, I)\right\rangle_\beta,\end{aligned}$$

where $\langle\cdots\rangle_\beta$ denotes the average with respect to the posterior distribution at $T = \frac{1}{\beta}$. Then by integration, we have

$$\ln p(d|H_1, I, \beta) = \int \left\langle \ln p(d|\vec{\theta}, H_1, I)\right\rangle_\beta d\beta. \quad (29)$$

Note that ultimately we want $p(d|H_1, I, \beta = 1)$, the evidence, and by definition

$$p(d|H_1, I, \beta = 0) = \int p(\vec{\theta}|H_1, I)d\vec{\theta} = 1.$$

Therefore, the natural choice of the lower and upper bound for evaluating Eq. 29 would be 0 and 1 respectively, also for the LHS of Eq. 29,

$$\ln p(d|H_1, I, \beta = 1) - \ln p(d|H_1, I, \beta = 0) = \ln p(d|H_1, I).$$

Finally, we have

$$\ln p(d|H_1, I) = \int_0^1 \left\langle \ln p(d|\vec{\theta}, H_1, I)\right\rangle_\beta d\beta. \quad (30)$$

Since in parallel tempering we have chains with different values of β, we can estimate the integral in Eq. 30 easily.

3. Burn-in and thinning

Since the MCMC chain starts at a random position in the parameter space, it is not guaranteed that samples in the initial stages are drawn from the desired posterior distribution. To mitigate this problem, we usually discard the some fraction of the initial chain, such that the memory of the chain to the initial position is lost. Although we wish that the posterior samples drawn using MCMC are independent of each other, often this is not the case. To remove the undesired correlation between samples, we take a sample from the MCMC chain every L sample, where L is the **autocorrelation length** (ACL) or **integrated autocorrelation time** (ACT), forming effective samples. This process is known as *thinning*. The autocorrelation length is defined as

$$L = \int_{-\infty}^{\infty} \rho(t)dt, \quad (31)$$

where $\rho(t)$ is the Pearson correlation coefficient between the chain and itself shifted by t samples. However, since we only have finite number of samples, we can estimate ACL/ACT by

$$L \approx 1 + 2\sum_t \rho(t). \quad (32)$$

II. OBJECTIVES

In this project, we aim to search for echoes of gravitational waves in the three events LIGO detected as of this writing, and verify the results of [8] using their phenomenological model with Bayesian analysis instead. We perform a Bayesian parameter estimation on the parameters of the echo waveform templates described in section I A 1 and Bayesian model selection of presence of echoes versus their absence, to incorporate the Occam's razor described in section I B 3, and to provide stronger evidence for the presence or absence of echoes in these detections. The analysis technique developed in this project could be repeated with different models to provide even more robust evidence of the existence of echoes from ECOs.

III. IMPLEMENTING TEMPLATES OF ECHOES

A. Phenomenological model by Mark et al.

As part of the work, the template model described in [14] by Mark et al. was implemented in Python. Table III summarizes the seven main functions in the implementation `phenom_echoes_waveform_tapir.py`. Note that the Fourier transform is defined as

$$\psi(t) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} Z(\omega)e^{-i\omega t}. \quad (33)$$

1. *Computing the reflection amplitude $\tilde{\mathcal{R}}_{BH}(\omega)$ and transmission amplitude $\tilde{\mathcal{T}}_{BH}(\omega)$*

In both Nakano et al.'s and Mark et al.'s models, there are two quantities that are required: reflection amplitude $\tilde{\mathcal{R}}_{\rm BH}(\omega)$ in Mark et al.'s notation or $\sqrt{R}$ in Nakano's notation, and transmission amplitude $\tilde{\mathcal{T}}_{\rm BH}(\omega)$ in Mark et al.'s notation or $\sqrt{1-R}$ in Nakano's notation. From now on, we will use Mark et al.'s notations for the sake of clarity and consistency.

In Nakano et al.'s paper, there are two fitting formulas for $\tilde{\mathcal{R}}_{\rm BH}$, one for positive and one for negative frequencies f. In practice, we would need to compute the reflection amplitude (and hence transmission amplitude by the normalization condition) for each of the systems LIGO had detected. For now, we will be using Nakano et al.'s fit for testing purposes to build the infrastructure of the search pipeline first. Figure 3 shows a plot of the reflection and transmission amplitude fit described in [13] for positive frequency f, parameterized by $x = 2\pi M f$ and q, where M is the mass of the black hole and q is the dimensionless spin parameter.

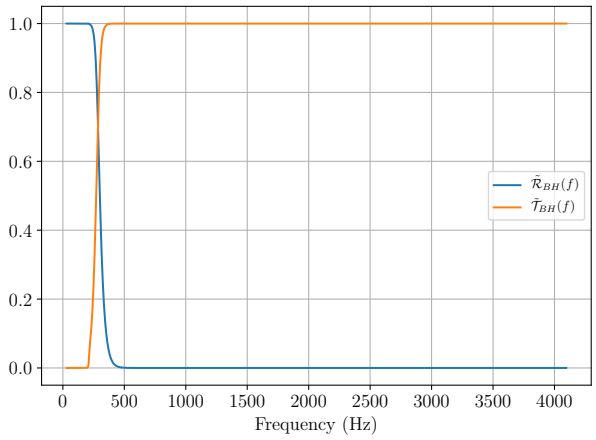


Figure 3: Nakano et al.'s fit for reflection and transmission amplitude as a function of frequency f, using the final mass and spin of the black hole given in [2] for the GW150914 event.

2. *Using Mark et al.'s model in the context of GW*

In Section I A, we mentioned that Nakano et al. argued that their templates work in the context of gravitational waves. However, attention must be paid when using the template proposed by Mark et al., since they derived the results in the context of scalar waves, whereas gravitational waves are metric perturbations from Minkowski spacetime $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$. Also, the template was derived using a Schwarzschild spacetime, but the sources of LIGO's three discoveries are all rotating instead of static. Therefore, the template becomes a phenomenological model when we try to extend it to the case of gravitational waves from rotating ECOs.

For an observer far away ($r \to \infty$), the Weyl scalar Ψ_4 relates to the second time derivative of the gravitational waves as

$$\Psi_4 = \ddot{h}_+ - i\ddot{h}_\times. \tag{34}$$

We **claim** (or rather *hope*) that template for modeling ψ proposed by Mark et al. also models Ψ_4. That means we can generate echo templates for gravitational waves by performing double integration with respect to time from the TAPIR templates for scalar fields.

From Eq. 6 and Eq. 33, we have

$$\psi_{\rm echo}(t) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} Z_{\rm echo}(\omega) e^{-i\omega t}.$$

Note that the only time dependence is in the term $e^{-i\omega t}$, that means double integration in time is equivalent to dividing the frequency waveform $Z_{\rm echo}$ by $(-i\omega)^2 = -\omega^2$. Hence, we have

$$h_+(t) - ih_\times(t) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{Z_{\rm echo}(\omega)}{-\omega^2} e^{-i\omega t}. \tag{35}$$

Since we are doing analysis on a *finite segment* of data, we must truncate the echo waveform, up to $N^{\rm th}$ echo, as follows

$$\begin{aligned} h_+(t) - ih_\times(t) &= \sum_{n=1}^{N} \int_{-\infty}^{+\infty} -\frac{d\omega}{2\pi\omega^2} \tilde{K}^{(n)} Z^{\rm H}_{\rm T}(\omega) e^{-i\omega t} \\ &= \sum_{n=1}^{N} \int_{-\infty}^{+\infty} -\frac{d\omega}{2\pi\omega^2} (\tilde{\mathcal{T}}_{\rm BH}\tilde{\mathcal{R}})(\tilde{\mathcal{R}}_{\rm BH}\tilde{\mathcal{R}})^{n-1} \times \\ & \quad e^{-i\omega(t-t_s+2nx_0)} e^{-\omega^2/(2\beta^2)} \left(\frac{\alpha_+}{\omega - \Omega_+} + \frac{\alpha_-}{\omega - \Omega_-} \right). \end{aligned}$$

We examine the term $e^{-i\omega(t-t_s+2nx_0)}$, it implies that the first echo starts roughly at $t = t_s + 2|x_0|$, and each successive echo is $2|x_0|$ apart. Note that we can infer the **compactness** of the ECO from the value of $|x_0|$.

3. *Echo template for GW150914 using Mark et al.'s model*

To demonstrate that we could use Mark et al.'s model to search for echoes of gravitational waves, we have generated a template with echoes for GW150914, the loudest event among the three we had. The QNM frequencies $\Omega_\pm$ for the horizon waveform used for generating this echo template was constructed using the least-damped QNM ($l = 2, m = 2$ mode) of the event, taken from [22], namely

$$\Omega_\pm = \pm 502\pi + i\frac{-1}{4.0 \times 10^{-3}}.$$

Figure 4a and 4b show the frequency-domain and time-domain horizon waveform for generating the echoes in GW150914 respectively.

Function	Description
`get_fd_horizon_waveform`	Generate the horizon waveform in frequency-domain $Z_{\rm T}^{\rm H}$
`get_fd_total_transfer`	Compute the total transfer function $\tilde{K}$
`get_fd_nth_transfer`	Compute the $n^{\rm th}$ transfer function $\tilde{K}^{(n)}$
`get_fd_total_echoes`	Generate the total echo waveform in frequency-domain $Z_{\rm T}$
`get_td_total_echoes`	Generate the total echo waveform in time-domain $\psi(t)$ by performing Fourier Transform on the frequency domain waveform
`get_fd_up_to_n_echoes`	Generate the echo waveform up to $n^{\rm th}$ echo in frequency domain $Z_{\rm T} = \sum_{k=1}^{n} Z_{\rm T}^{(k)}$
`get_td_up_to_n_echoes`	Generate the echo waveform up to $n^{\rm th}$ echo in time-domain by performing Fourier Transform to the frequency domain waveform

Table III: The seven main functions in the Python implementation of template model described in Mark et al. [14].

Using Nakano et al.'s fitting of $\tilde{\mathcal{R}}_{\rm BH}(f)$, $\tilde{\mathcal{T}}_{\rm BH}(f)$, and assuming that the reflectivity $\tilde{R}$ of the ECO surface is frequency-independent, we were able to generate a template for echoes in GW150914. Figure 5a and 5b show the best-fit IMR waveform of GW150914 with three echoes appended and the three echoes respectively in time domain.

B. Phenomenological model by Abedi et al.

As part of the work, the template model described in [8] by Abedi et al. was implemented in Python. Table IV summarizes the five main functions in the implementation `phenom_echoes_waveform_abedi.py`. In our implementation, we have slightly changed the meaning of two time-related parameters $t_{\rm echo}$ and t_0, in a sense that they are now measured with respect to the time of merger $t_{\rm merger}$, for the sake of computational efficiency and clarity.

To demonstrate that we could use Abedi et al.'s model to search for echoes of gravitational waves, we have generated a template with echoes for GW150914. Figure 6a and 6b shows the frequency evolution $\omega_I(t)$ of the best-fit IMR template of GW150914 and smooth cut-off function $\Theta(t, t_0)$ to be applied to the IMR waveform to produce the echo waveform respectively. Figure 7a and 7b show the best-fit IMR time-domain waveform of GW150914 with three echoes appended and the truncated IMR time-domain waveform $\mathcal{M}_{\rm T,I}(t)$ used to generate the echo template respectively in the source frame.

IV. BAYESIAN ANALYSIS ON FAKE STRAIN DATA

Before performing the Bayesian analysis on actual strain data, it is educational to perform the same analysis on fake strain data first, namely strain data with Gaussian noise and IMR signal with echoes (which we will refer as *IMRE*) of known parameters. By recovering the injected signal and inferring the parameters correctly, we can learn that the analysis method proposed would be able to find signals in real strain data, and that the method has been properly implemented.

A. Generating fake strain data

In order to generate fake strain data, a program `generate_fake_strain.py` was written in Python. The power spectrum density used to produce the noise here is the analytical PSD `aLIGOZeroDetHighPower`. To generate the IMRE strain $h_{\rm H1/L1}(t)$ in detector frame from the plus and cross polarizations $h_+(t)$, $h_\times(t)$ in source frame, note that

$$\begin{aligned} h_{\rm H1/L1}(t) = {} & F_{+,\rm H1/L1}(t;\alpha,\delta,\psi)h_+(t) \\ & + F_{\times,\rm H1/L1}(t;\alpha,\delta,\psi)h_\times(t), \end{aligned} \tag{36}$$

where F_+, $F_\times$ are the detector responses for plus and cross polarization respectively, and α, δ and ψ are the right ascension, declination of the source, polarization of the gravitational wave respectively.

Figure 8a shows the 10-second long of strain of pure Gaussian noise in H1, with lower frequency cut-off $f_{\rm low} =$ 30 Hz. Figure 8b shows the 10-second long strain of Gaussian noise with an IMRE signal injected in H1 at geocentric GPS time $t_{\rm merger} = 1126259462.430$ s, with the same lower frequency cut-off. We can observe from Figure 8b that only the merger part of the signal stands out from the noise floor.

B. Parameter estimation on fake strain data

Recall from Sec. I B 2 a that we can infer the parameters of the source binary system from the gravitational wave strain data by the virtue of Bayes' Theorem. We implemented a parameter estimation pipeline from start to end in Python, capable of digesting `gwf` frame file (`mcmc_preprocessing.py`), signal injection (`generate_fake_strain.py`), parameter estimation (`mcmc_parameter_estimation.py`) and post-processing (`mcmc_postprocessing.py`) such as thinning, point estimation calculation and generating summary page.

Function	Description
`get_td_ligo_template`	Generate the IMR template in time domain. Currently it is a wrapper of `pycbc.waveform.get_td_waveform` in `PyCBC`
`get_frequency_evolution`	Compute the frequency evolution of the IMR template $\omega_I(t)$. Currently it is a wrapper of `waveform.utils.frequency_from_polarizations` in `PyCBC`
`get_td_cutoff`	Compute the cut-off function $\Theta(t, t_0)$
`get_td_echo_up_to_n`	Generate the echo waveform up to n^{th} echo in time-domain
`get_fd_echo_up_to_n`	Generate the echo waveform up to n^{th} echo in frequency-domain by performing Fourier transform on the time-domain template

Table IV: The five main functions in the Python implementation of template model described in Abedi et al. [8].

Table V summarizes the technical details, such as the IMR approximant used and number of iterations when performing MCMC, of a parameter estimation run with IMRE injected to colored Gaussian noise. The characteristic SNR [23] of the signal in H1 and L1 are 50.0 and 38.3 respectively. The IMR parameters of the injected signal, such as masses and spins, were chosen to be the inferred values of GW150914.

The program `mcmc_parameter_estimation.py` is parallelizable, with the aid of OpenMPI implementation for Python `mpi4py`. As the ensemble sampler updates the position of half of the walkers at a time, we used 5 CPU cores (which is one-half of the number of walkers) in order to minimize the latency of the calculation.

We have chosen the prior distribution of the echo parameters to be uniform over a range, and the prior range of the parameters used are listed in Table VI. The summary statistics of the parameter estimation run, such as MLE and MAP, are listed in Table VII. A very useful visualization of the sampled posterior distribution, corner or triangle plot, is shown in Figure 11. We can see that the inferred values of the echo parameters are both accurate (close to injected value) and precise (narrow posterior distribution), especially for those time-related parameters. For example, we can see from the 1D histogram of Δt_{echo} in Figure 9a that the MAP is very close the injected value (represented by the vertical dashed line), and the 90% Bayesian credible interval ([0.1498, 0.1503]) is much narrower than the prior range ([0.1, 0.2]), that means the range is shrunk by about 99.5%. Figure 9b shows that the chain converges near the true value and hovers around it. Note that the posterior samples are obtained from the thinned chain, where the estimated ACL/ACT is 153.

As for amplitude-related parameters A and γ, the parameter estimation is not as accurate and precise as those time-related parameters. For instance, Figure 10a shows the 1D histogram of A, and Figure 10b shows the posterior thinned samples of A. We can see that the range is not shrunk as much as compared with Δt_{echo}, only by about 90.8%. The posterior samples for A also did not converge well as compared with Δt_{echo}. This is not too surprising because those time-related parameters can be inferred nicely using coherence of the strain with the template generated on-the-fly, while those amplitude-related parameters can only be inferred using noisy strain data.

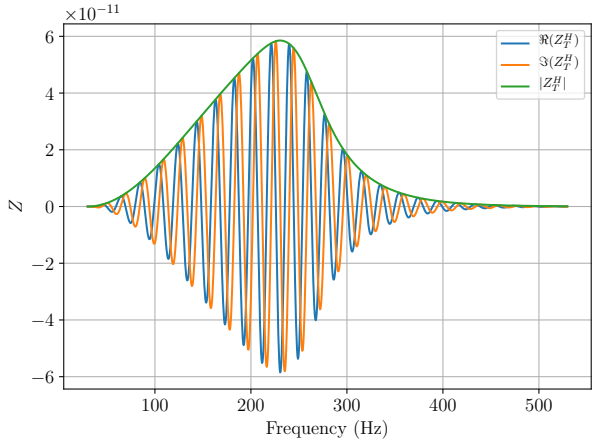
Similar plots for other echo parameters can be found in the summary page `https://ldas-jobs.ligo.caltech.edu/~ka-lok.lo/projects/echoes/production/parameter_estimation_170814/summary_page.html`.

Generation of fake strain	
IMR approximant	SEOBNRv4_ROM
Echo template model	Abedi et al's
Analytical PSD	aLIGOZeroDetHighPower
Computational resources used	
Run time	17:18:19
Number of CPU cores used	5
Total CPU time	3 days 08:44:35
Memory usage	4.58 GB
Hyperparameters in MCMC	
Number of iterations	20000
Number of walkers	10
Number of temperature chains	8
Fraction of burn-in	0.1

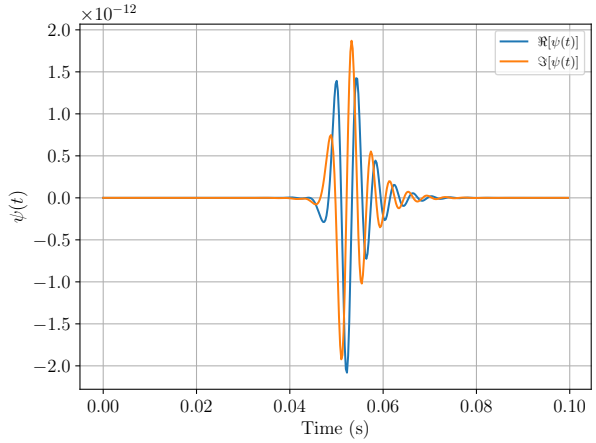
Table V: Technical details of the parameter estimation run

Parameter	Prior range
A	[0.01, 0.99]
γ	[0.01, 0.99]
t_0 (s)	[-0.05, 0.0]
t_{echo} (s)	[0, 0.1]
Δt_{echo} (s)	[0.1, 0.2]

Table VI: The prior range of the echo parameters

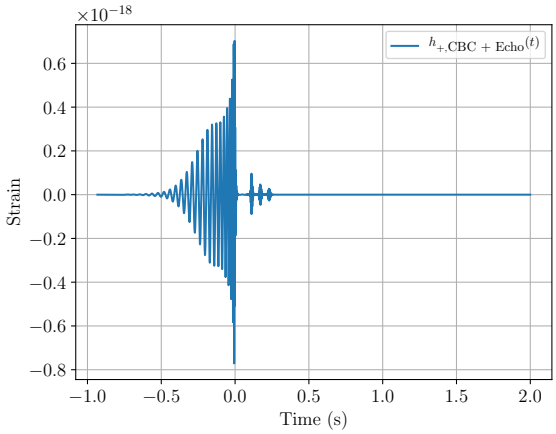


(a) Frequency domain

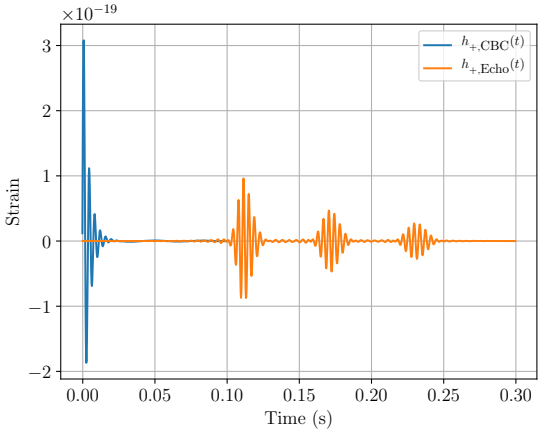


(b) Time domain

Figure 4: The horizon waveform for generating the echoes in GW150914.



(a) GW150914 best-fit IMR waveform with three echoes appended

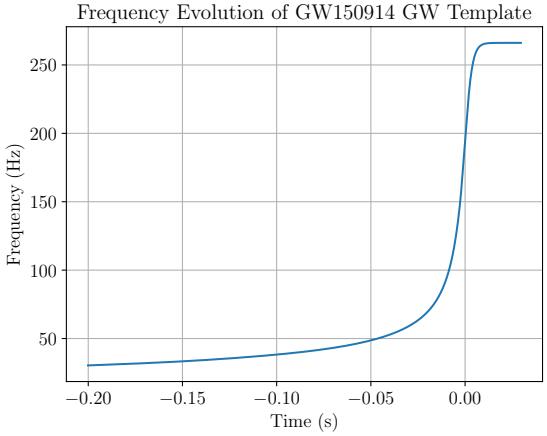


(b) The train of echoes, with $n = 3$

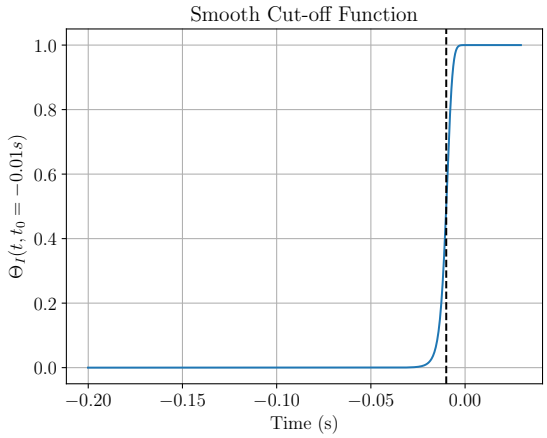
Figure 5: The best-fit IMR time-domain template GW150914 with three echoes in the source frame.

Parameter	MLE	MAP	Mean	Median	SD	90% CI	Injected Value
A	0.184	0.184	0.180	0.178	0.0272	[0.133, 0.224]	0.23
γ	0.725	0.725	0.708	0.706	0.105	[0.548, 0.887]	0.64
t_0 (s)	-0.0135	-0.0135	-0.0158	-0.0149	0.00427	[-0.0220, -0.0104]	-0.01
$t_{\rm echo}$ (s)	0.0500	0.0500	0.0499	0.0499	0.000137	[0.0496, 0.0500]	0.05
$\Delta t_{\rm echo}$ (s)	0.150	0.150	0.150	0.150	0.000159	[0.150, 0.150]	0.15

Table VII: The summary statistics of a parameter estimation run. 90% CI refers to the 90% minimum Bayesian credible interval, introduced in Sec I B 2 b

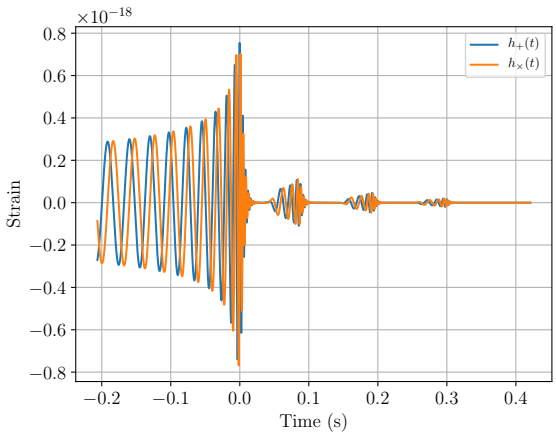


(a) The frequency evolution $\omega_I(t)$ of the IMR template

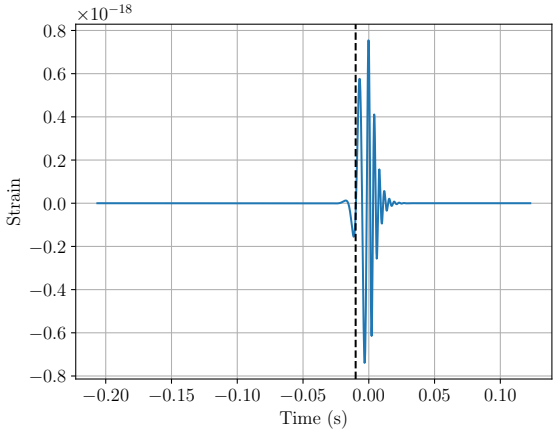


(b) The smooth cut-off function $\Theta(t, t_0)$. The dashed vertical line is $t = t_0$.

Figure 6: (6a) The frequency evolution of the best-fit IMR template for GW150914. (6b) The smooth cut-off function to be applied to GW150914 IMR template to generate the echo template.

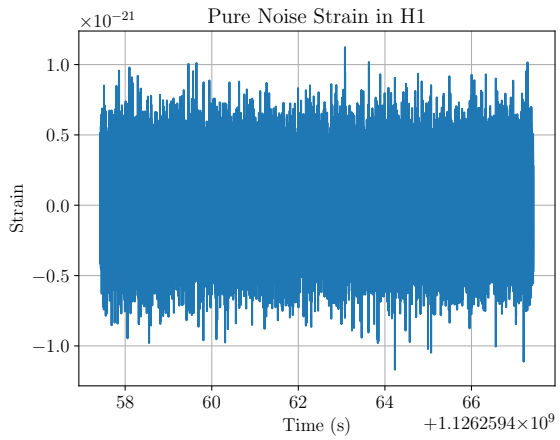


(a) GW150914 best-fit IMR waveform with three echoes appended

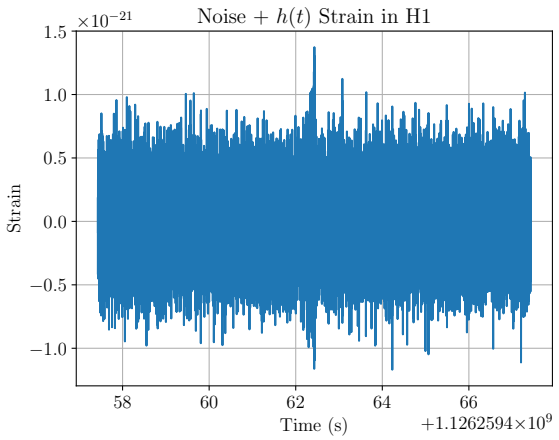


(b) The truncated IMR waveform $\mathcal{M}_{\mathrm{T,I}}(t)$ used to produced the echo template

Figure 7: (7a) The best-fit IMR template for GW150914 with three echoes appended in the source frame. (7b) The truncated IMR waveform $\mathcal{M}_{\mathrm{T,I}}(t)$ used to generate the echo template.

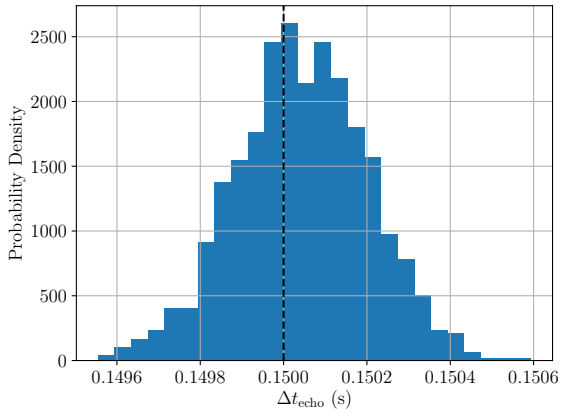


(a) The strain of pure Gaussian noise in H1

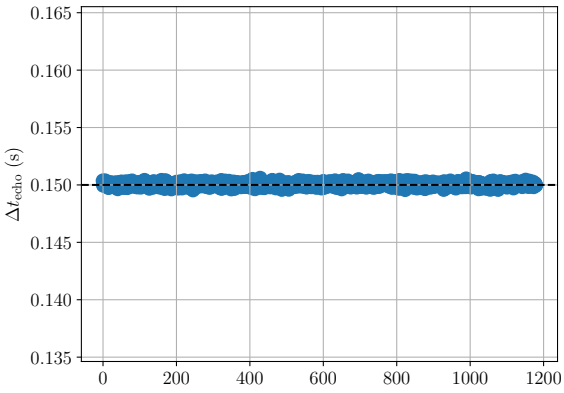


(b) The strain of Gaussian noise with IMRE template injected in H1

Figure 8: The fake strain in H1 generated for the analysis. An IMRE signal was injected at geocentric GPS time $t_{\text{merger}} = 1126259462.430$ s. We can see from Fig 8b that the signal stands out from the noise floor.

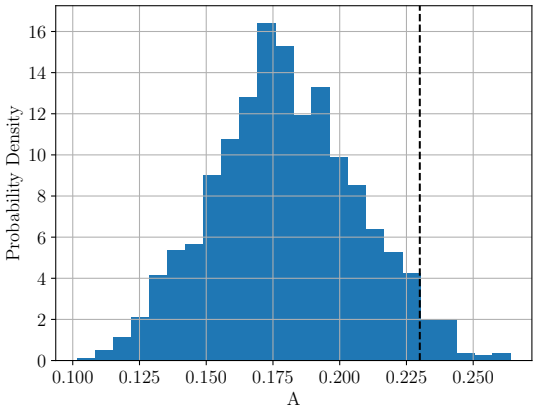


(a) The estimate of 1D marginal posterior probability density for $\Delta t_{\rm echo}$ by histogram. Note that the distribution is very narrow compared to its prior range, and the MAP is very close to the injected value (indicated by the dashed line)

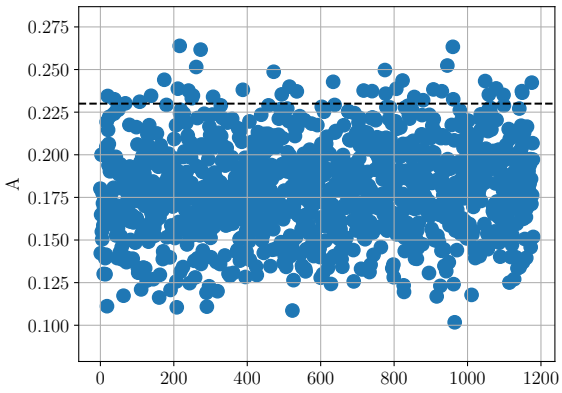


(b) The posterior samples of $\Delta t_{\rm echo}$. We can see that the chain converges to the true value and hovers around the injected value

Figure 9: The plots for 1D marginal posterior probability density function for the parameter $\Delta t_{\rm echo}$



(a) The estimate of 1D marginal posterior probability density for A by histogram. Note that the distribution is narrow compared to its prior range, and the MAP is reasonably close to the injected value (indicated by the dashed line)



(b) The posterior samples of A. We can see that the chain does not converge very well as compared with the case of $\Delta t_{\rm echo}$

Figure 10: The plots for 1D marginal posterior probability density function for the parameter A

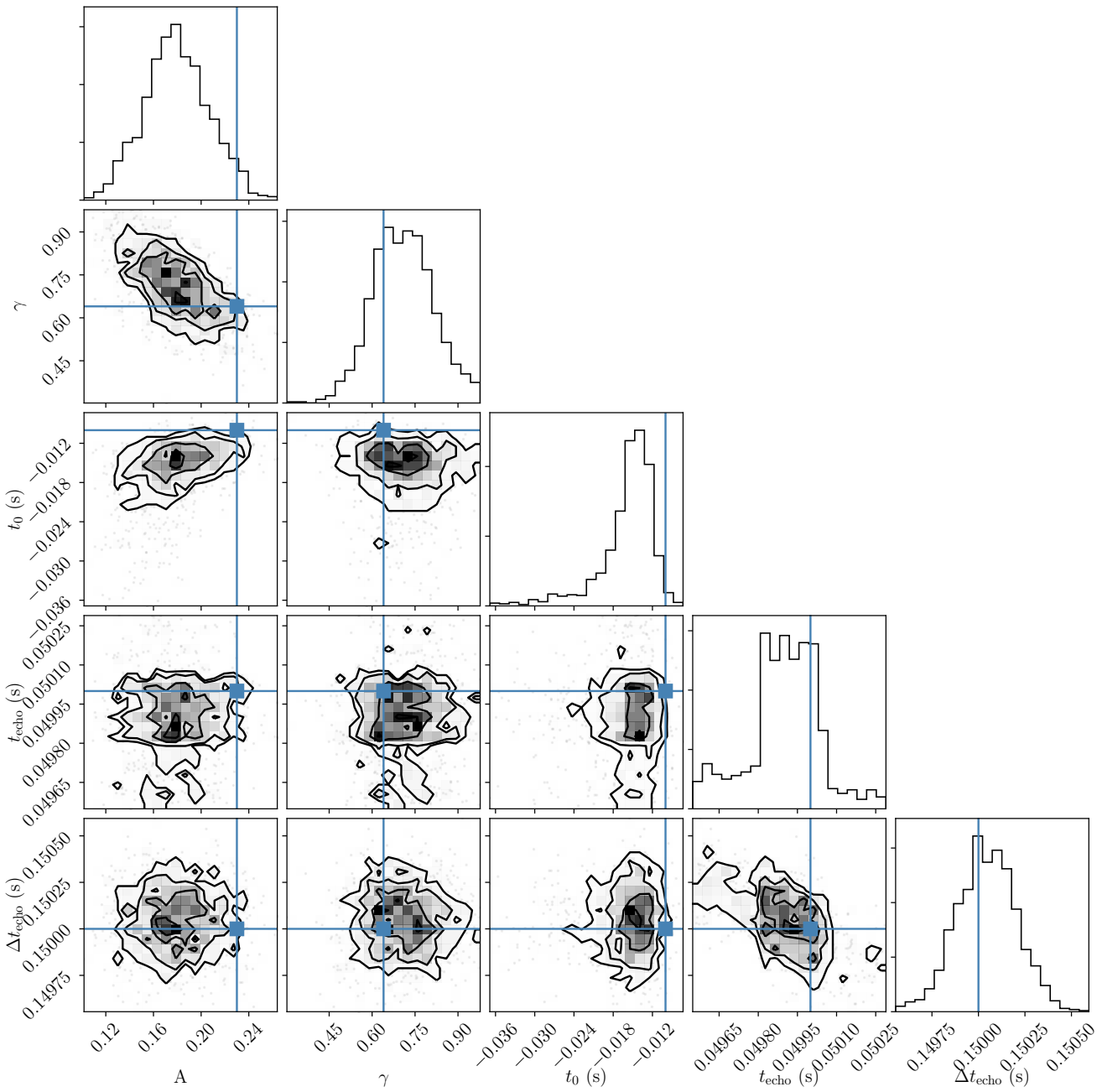


Figure 11: The corner/triangle plot for this parameter estimation run. The blue solid lines represent the values of parameters being injected

C. Hypothesis testing on fake strain data

Ultimately, we want to perform hypothesis testing on a given strain data, and parameter estimation is just an intermediate step. In the language of Bayesian inference, we have two competing models H_0 and H_1, which are just null hypothesis and alternative hypothesis in frequentist language, and they are

$$H_0 : \neg\text{Echo} \Rightarrow d = n + h_{\rm IMR},$$
$$H_1 : \text{Echo} \Rightarrow d = n + h_{\rm IMRE}.$$

Note that we *assumed that there is a gravitational wave signal in the data*, and we are only interested in knowing whether there are echoes in the data or not. When the null hypothesis is true, that means the data only contain a GW signal with inspiral, merger and ringdown. When the alternative hypothesis is true instead, that means the data contain not only IMR, but also echoes.

Recall from Sec I B 3 that we can compute the odds ratio. In this case, the odds ratio is given by

$$O^{\rm Echo}_{\neg\rm Echo} = \frac{p(d\,|\text{Echo}, \text{GW})}{p(d\,|\neg\text{Echo}, \text{GW})} \frac{p(\text{Echo}\,|\text{GW})}{p(\neg\text{Echo}\,|\text{GW})}. \tag{37}$$

If we assume that the prior odds is simply 1, namely

$$p(\text{Echo}\,|\text{GW}) = p(\neg\text{Echo}\,|\text{GW}) = \frac{1}{2},$$

then the odds ratio is simply the Bayes factor, equivalently the marginal likelihood ratio

$$O^{\rm Echo}_{\neg\rm Echo} = B^{\rm Echo}_{\neg\rm Echo} = \frac{p(d\,|\text{Echo}, \text{GW})}{p(d\,|\neg\text{Echo}, \text{GW})}. \tag{38}$$

Since we work with log posterior, log likelihood and log prior, we can take natural logarithm on both sides of Eq. 38, we have

$$\ln O^{\rm Echo}_{\neg\rm Echo} = \ln B^{\rm Echo}_{\neg\rm Echo} = \ln p(d\,|\text{Echo}, \text{GW}) - \ln p(d\,|\neg\text{Echo}, \text{GW}) = \ln Z_{\rm IMRE} - \ln Z_{\rm IMR}. \tag{39}$$

The term $\ln p(d\,|\text{Echo}, \text{GW})$ can be obtained from parameter estimation result using thermodynamic integration as described in Sec I C 2 a. As for the term $\ln p(d\,|\neg\text{Echo}, \text{GW})$, this can be obtained easily since we are fixing the IMR parameters during parameter estimation. The likelihood function is actually a Dirac-delta function peaked at $\vec{\theta} = \vec{\theta}_{\rm IMR}$ as we are not varying the IMR parameters, hence the marginal likelihood is simply

$$p(d\,|\neg\text{Echo}, \text{GW}) = \int p(d\,|\vec{\theta}_{\rm IMR}, \neg\text{Echo}, \text{GW})\delta(\vec{\theta} - \vec{\theta}_{\rm IMR})d\vec{\theta} = p(d\,|\vec{\theta}_{\rm IMR}, \neg\text{Echo}, \text{GW}), \tag{40}$$

which is just the likelihood evaluated at $\vec{\theta} = \vec{\theta}_{\rm IMR}$, given that there are no echoes.

Table VIII shows the log evidence of strain data containing IMRE and IMR signal respectively, and the log Bayes factor for IMRE versus IMR computed from the injection run. The log Bayes factor is the **detection statistic** we are going to use to decide whether we claim there is an IMRE signal or IMR signal in the data. Since the log Bayes factor $\ln B^{\rm Echo}_{\neg\rm Echo}$, and equivalently log odds ratio $\ln O^{\rm Echo}_{\neg\rm Echo}$, is greater than 1, we can conclude, from the Bayesian point of view, that the data favors the existence of echoes, which is indeed the case. Readers should be reminded that the Bayes factor and hence odds ratio incorporate the Occam factor that biases simpler model, in this case the absence of echoes. For typical results with IMR-only injections, see section IV D 2, below.

	Value
Log evidence for IMRE in the data, $\ln Z_{\rm IMRE}$	-40490.980
Estimated statistical error of $\ln Z_{\rm IMRE}$	3.305
Log evidence for IMR in the data, $\ln Z_{\rm IMR}$	-40529.718
Log Bayes factor for IMRE vs. IMR, $\ln B^{\rm Echo}_{\neg\rm Echo}$	38.738

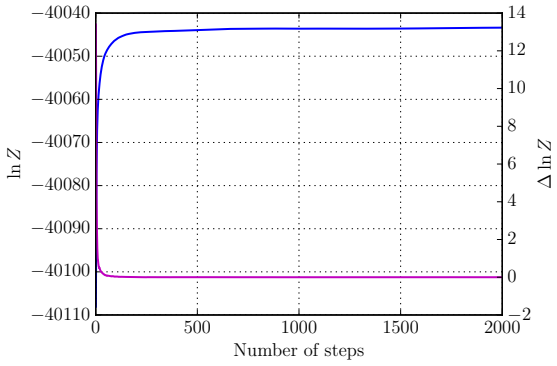
Table VIII: A summary of log evidence and log Bayes factor calculated for hypothesis testing

1. *Test of convergence of evidence calculation using thermodynamic integration*

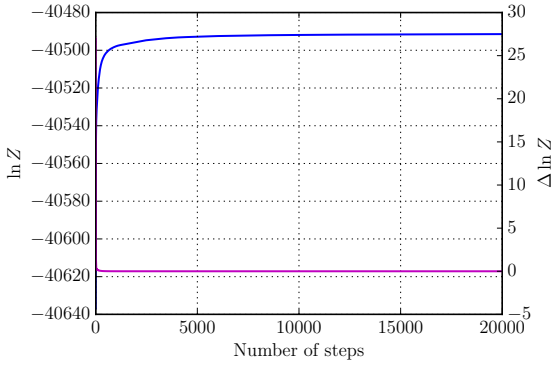
As the detection statistic log Bayes factor is calculated from the log evidence of IMRE $\ln Z_{\rm IMRE}$ and that of IMR $\ln Z_{\rm IMR}$, which are in turn calculated from the posterior samples obtained using a stochastic sampling method, it is imperative to check that the evidence calculated using thermodynamic integration has converged after some finite number of steps.

For the log evidence of IMR, since the IMR parameters were fixed during an injection run for now, the log evidence calculation is exact, and the value can be obtained easily using Eq. 40.

As for the log evidence of IMRE, it was calculated using thermodynamic integration, described in Sec I C 2 a. Figure 12a and 12b show the plot of the estimated value of $\ln Z$, and the difference in the estimated value between each evaluation $\Delta \ln Z$, as a function of number of steps, for a background run (IMR injection) and the IMRE injection run (described in previous section) respectively. For background run, we see that $\ln Z$ quickly converges after about 500 steps, and this can be confirmed by the fact that $\Delta \ln Z$ drops to nearly zero at about 500 steps. This result implies that we actually do not need 2000 iterations but 500 iterations in each run would be sufficient, bringing a 4X improvement in speed. The quick convergence, compared with that of IMRE injection run, exhibited in the background run is likely due to the fact that the posterior distribution is flatter for IMR injection run than IMRE injection run, therefore it takes less



(a) IMR injection (background)



(b) IMRE injection

Figure 12: The estimated value of $\ln Z$ using thermodynamic integration as a function of number of steps is plotted in blue. The difference in value between each evaluation $\Delta \ln Z$ as a function of number of steps is plotted in magenta

iterations for the evidence calculation of background run to converge.

D. Estimation of statistical significance of detection in colored Gaussian noise

After we obtained the detection statistic, a natural question to ask is that how significant statistically is the detection statistic. Simply put, how strong is the evidence[24] of the detection? In the Bayesian school, there are different empirical scales, such as Jeffreys' scale, to interpret the strength of the evidence. However, they are subjective and not universally applicable. Therefore, we are not going to use them.

1. *Posterior probability of hypotheses*

In general, we cannot compute the posterior probability of a model being true from doing Bayesian model comparison/selection, because the models in consideration are not necessarily exhaustive, i.e. model 1 not being true does not imply that model 2 is true, and vice versa. However, in our case, the models, i.e. existence of echoes and absence of echoes, are exhaustive. This allows us to compute the posterior probability of a hypothesis from odds ratio.

$$\Pr(\neg\text{Echo}\,|d,\text{GW}) = \frac{1}{1+e^{\ln O^{\text{Echo}}_{\neg\text{Echo}}}}, \tag{41}$$

$$\Pr(\text{Echo}\,|d,\text{GW}) = 1-\Pr(\neg\text{Echo}\,|d,\text{GW}) = \frac{e^{\ln O^{\text{Echo}}_{\neg\text{Echo}}}}{1+e^{\ln O^{\text{Echo}}_{\neg\text{Echo}}}}. \tag{42}$$

Using Eq. 42 and value from Table VIII, the posterior probability that the data contains echoes, given that there is a gravitational wave signal is

$$\Pr(\text{Echo}\,|d,\text{GW}) = 1-1.50\times10^{-17},$$

which tells you that, in the *Bayesian framework*, the probability of the data containing no echoes is as small as 1.50×10^{-17}!

2. *The frequentist approach: p-value*

Calculating the posterior probability of a hypothesis is certainly much better than using a subjective scale to determine the strength of evidence. The approach described in previous section is probably optimal given that there is only one set of data d available [25]. However, the Bayesian posterior probability fails to tell us that how likely the so-called evidence is simply due to random background noise, since we only consider merely one set of data! The frequentist approach can answer the following question: Given the null hypothesis H_0 is true, how likely (i.e. the probability) are the data going to be as *extreme or more extreme than* the observed data? The probability that we are looking for is exactly the frequentist **p-value**. We can also interpret this p-value as the **false-alarm probability**.

To compute, or estimate, the p-value, we need a **test statistic** T, a natural choice is our detection statistic $\ln B^{\text{Echo}}_{\neg\text{Echo}}$, as it satisfies two important criteria of a test statistic:

1. Continuous,
2. Larger value of T corresponds to better agreement with the alternative hypothesis H_1.

The p-value, where we denote it as simply p, is related to the **null distribution** of test statistic by

$$p = \Pr(T \geq T_{\text{detected}}|H_0) = \int_{-\infty}^{T_{\text{detected}}} p(T|H_0)dT, \tag{43}$$

where T_{detected} is the test/detection statistic we obtained in an experiment, and $p(T|H_0)$ is called the null distribution of T, i.e. the distribution of T given that H_0 is

true. To sample the null distribution of T in the case of colored Gaussian noise, we can perform a lot of runs with **IMR** injection (i.e. assuming that H_0 is true) and different realization of the noise. In this way, we can see how likely the noise will cause the detection statistic as high as, or higher than, the detected value.

We performed 8500 such background runs. The histogram of the sampled null distribution is shown in Figure 13. The vertical dashed line corresponds to the test statistic obtained in the parameter estimation run described in IV B. We can see that the detected value stands out from the null distribution, i.e. the background. This is *consistent* with the posterior probability approach that the the detected value has a very high statistical significance. However, this also means that it is very difficult to estimate the p-value accurately. In fact, the highest sampled value of T under null hypothesis is 9.925. Naively, this problem can be solved by obtaining even more samples, or we can place an upper bound on the p-value. A rough estimation, which will be discussed in the following section, reveals that we need about 10^9 samples in order to obtain a sample as high as the detected value that is caused by the background, which is impractical.

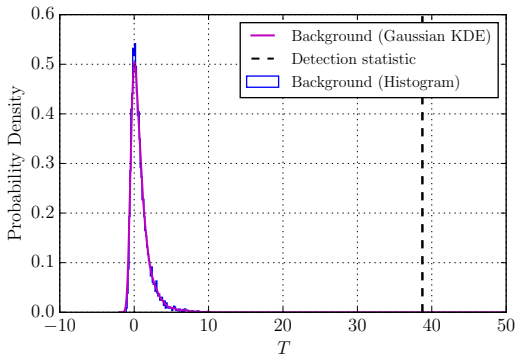


Figure 13: The histogram of null distribution of test statistic T is plotted in blue. The vertical dashed line corresponds to the detected value from the simulated signal described in section IV C and Table VIII. The estimated probability density function using Gaussian KDE is plotted in magenta

To solve this problem, we can use **Gaussian Kernel Density Estimation**, or simply Gaussian KDE, to estimate the probability density function using samples of the null distribution. The kernel density estimate is a non-parametric method to estimate the probability density function (pdf), in this case the null distribution of T, using some samples drawn from the distribution. The basic idea of KDE is that the estimated pdf $\hat{p}(x)$ is proportional to the sum of all samples, weighted by some function called *kernel* K. Mathematically,

$$\hat{p}(x) = \frac{1}{Nb}\sum_{i=0}^{N-1} K(\frac{x - X_i}{b}), \tag{44}$$

where b is the bandwidth and X_i is the $(i+1)^{\text{th}}$sample. Readers can refer to [26] for detailed discussions of Gaussian KDE. The bandwidth parameter b used in this analysis was chosen using Scott's Rule [27].

Figure 13 shows the Gaussian KDE (in magenta) plotted on top of the histogram. We see that the Gaussian KDE fits the histogram well. Now, we can compute the p-value using the estimated pdf of null distribution. The p-value obtained is

$$p = 9.52 \times 10^{-10},$$

which means that the detected value has a statistical significance of 6.01σ. Note that the injection is relatively loud, and we used colored Gaussian noise, which is the most optimal case. We do not expect such a high statistical significance when we use real LIGO power spectrum density, where the noise is non-Gaussian.

V. CONCLUSION AND FUTURE WORK

In this project, we have successfully recovered the IMRE injection buried in colored Gaussian noise. The recovered echo parameters were both accurate and precise. More importantly, we can perform Bayesian hypothesis testing to test the existence of echoes in the data, and reporting the statistical significance of the detection by estimating the background using many IMR injection runs. This implies that the analysis method we proposed here can find an IMRE signal and report the significance of the detection, and infer its parameters if there is one in the real LIGO strain data, and that the pipeline was implemented properly.

In short term, the foreground distribution of the detection statistic will be sampled using many IMRE injection runs with various echo parameters, to make sure that our detected value is consistent with the foreground. The same analysis will be performed on GW150914, LVT151012, GW151226 and GW170104 to search for echoes in the real gravitational wave data. The same analysis will also be repeated using different approximants such as that by Mark et al. and Nakano et al., to provide more robust evidence of the detection.

In long term, in order to speed up the evidence calculation and to have more control over the uncertainty in the estimation of evidence, the pipeline should use (multi-)nested sampling instead of parallel tempered MCMC with thermodynamic integration. This is because previous work has shown that nested sampling requires about a factor of one hundred [28] less evaluation of posterior distribution than thermodynamic integration. Although nested sampling was designed to evaluate the evidence, posterior samples can also be generated as a by-product with little computational effort, which are important when we want to infer the properties of the object using echoes.

Also, in the distant future, when we understand the physics of exotic compact objects better in a sense that

we can come up with physical approximants of the echoes from different types of ECOs, the methodology proposed in this project can be readily modified to test the nature of ECOs, using sub-hypotheses of H_1 such as $H_{\text{Gravastar}}$ and H_{Fuzzball}, that is

$$H_1 = H_{\text{Gravastar}} \vee H_{\text{Fuzzball}} \vee \dots,$$

so that we can learn even more about the properties and structure of ECOs. Earlier work done by Cardoso et al. showed that the time delay between each echo Δt_{echo} can be used to infer the nature of the ECO [6], namely

$$\Delta t_{\text{echo}} \sim -nM \log\left(\frac{l}{M}\right),$$

where M is the mass of the ECO, $l \ll M$ is the microscopic correction of the location of the ECO surface from the Schwarzschild radius, and n is an integer of the order of 1 which depends on the nature of the ECO.

VI. ACKNOWLEDGEMENT

The author would like to express his gratitude to Zachary Mark, the first author of [14], for his help and the friutful discussions of the project, and Hiroyuki Nakano, the first author of [13] for his swift and constructive replies to the author's inquiries. The author would also like to thank Rory Smith for his help with Bayesian analysis. The project has made extensive use of `PyCBC` [29, 30] as the backbone and `emcee` [31] for MCMC. Figure 11 was generated using `corner.py`[32]. The source code of all the programs developed for this project are available at GitLab: `http://137.189.40.44/ka-lok.lo/echoes`.

This summer research would not be possible without Tjonnie Li, the principal investigator of the CUHK GW Group, and National Science Foundation (NSF), Charles K. Kao Research Exchange Scholarship, CN YANG Research Exchange Scholarship and Global Learning Award Scheme from Wu Yee Sun College for the financial support.

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 241103 (2016).
[2] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).
[3] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaboration), Phys. Rev. Lett. **118**, 221101 (2017).
[4] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Physical Review X **6**, 041015 (2016), arXiv:1606.04856 [gr-qc].
[5] V. Cardoso, E. Franzin, and P. Pani, Physical Review Letters **116**, 171101 (2016), arXiv:1602.07309 [gr-qc].
[6] V. Cardoso, S. Hopper, C. F. B. Macedo, C. Palenzuela, and P. Pani, Physical Review D **94**, 084031 (2016), arXiv:1608.08637 [gr-qc].
[7] Gravastar stands for Gravitational Vacuum Star.
[8] J. Abedi, H. Dykaar, and N. Afshordi, ArXiv e-prints (2016), arXiv:1612.00266 [gr-qc].
[9] Or equivalently time of coalescence, denoted by t_c.
[10] IMR stands for **I**nspiral **M**erger **R**ingdown.
[11] G. Ashton, O. Birnholtz, M. Cabero, C. Capano, T. Dent, B. Krishnan, G. D. Meadors, A. B. Nielsen, A. Nitz, and J. Westerweck, ArXiv e-prints (2016), arXiv:1612.05625 [gr-qc].
[12] J. Abedi, H. Dykaar, and N. Afshordi, ArXiv e-prints (2017), arXiv:1701.03485 [gr-qc].
[13] H. Nakano, N. Sago, H. Tagoshi, and T. Tanaka, ArXiv e-prints (2017), arXiv:1704.07175 [gr-qc].
[14] Z. Mark, A. Zimmerman, S. M. Du, and Y. Chen, ArXiv e-prints (2017), arXiv:1706.06155 [gr-qc].
[15] A stationary random process means that the statistical properties of the process do not change with time [33].
[16] M. Maggiore, , 351 (2008).
[17] T. G. F. Li, *Extracting Physics from Gravitational Waves Testing the Strong-field Dynamics of General Relativity and Inferring the Large-scale Structure of the Universe*, 1st ed. (Springer International Publishing, 2015) p. 235.
[18] N. Metropolis and S. Ulam, Journal of the American Statistical Association **44**, 335 (1949).
[19] W. K. Hastings, Biometrika **57**, 97 (1970).
[20] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).
[21] P. M. Goggans and Y. Chi, AIP Conference Proceedings **707**, 59 (2004), http://aip.scitation.org/doi/pdf/10.1063/1.1751356.
[22] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Physical Review Letters **116**, 221101 (2016), arXiv:1602.03841 [gr-qc].
[23] The characteristic SNR is defined as $\sqrt{(h|h)}$.
[24] Here by evidence we refer to the detection statistic, not the marginal likelihood.
[25] The author have no formal proof that calculating the posterior probability of a hypothesis **is the most optimal**, but considering that the frequentist approach is basically useless here since we only have **one** observation. Given one set of data d, a value computed using probability theory seems much more reliable than an empirical scale.
[26] B. W. Silverman, *Density estimation for statistics and*

data analysis / B.W. Silverman (Chapman and Hall London ; New York, 1986) p. 175 pages :.
[27] D. W. Scott, "Multivariate density estimation: theory, practice, and visualization," (2015).
[28] P. Mukherjee, D. Parkinson, and A. R. Liddle, The Astrophysical Journal Letters **638**, L51 (2006).
[29] T. Dal Canton *et al.*, Phys. Rev. **D90**, 082004 (2014), arXiv:1405.6731 [gr-qc].
[30] S. A. Usman *et al.*, Class. Quant. Grav. **33**, 215004 (2016), arXiv:1508.02357 [gr-qc].
[31] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, Publications of the Astronomical Society of the Pacific **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].
[32] D. Foreman-Mackey, The Journal of Open Source Software **24** (2016), 10.21105/joss.00024.
[33] J. D. E. Creighton and W. G. Anderson, , 269 (2011).