Bayesian inference of overlapping gravitational wave signals

Elia Pizzati,1,2 Surabhi Sachdev,1 Anuradha Gupta,3 and B. S. Sathyaprakash1,4,5
1Institute for Gravitation and the Cosmos, Department of Physics, Penn State University, University Park PA 16802, USA
2Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
3Department of Physics and Astronomy, The University of Mississippi, Oxford MS 38677, USA
4Department of Astronomy and Astrophysics, Penn State University, University Park PA 16802, USA
5School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, United Kingdom

(Dated: February 16, 2021)

The observation of gravitational waves from LIGO and Virgo detectors inferred the merger rates to be $23.9^{+14.9}_{-10.1}$ $\text{Gpc}^{-3} \text{yr}^{-1}$ for binary black holes and $320^{+500}_{-240}$ $\text{Gpc}^{-3} \text{yr}^{-1}$ for binary neutron stars. These rates suggest that there is a significant chance that two or more of these signals will overlap with each other during their lifetime in the sensitivity-band of future gravitational-wave detectors such as the Cosmic Explorer and Einstein Telescope. The detection pipelines provide the coalescence time of each signal with an accuracy $\mathcal{O}(10 \text{ ms})$. We show that using the information of the coalescence time, it is possible to correctly infer the properties of these overlapping signals with the current data-analysis infrastructure. Studying different configurations of the signals, we conclude that the inference is robust provided that the two signals are not coalescing within less than $\sim 1$–$2\text{ s}$. Signals whose coalescence epochs lie within $\sim 0.5 \text{ s}$ of each other suffer from significant biases in parameter inference, and new strategies and algorithms are required to overcome such biases.

I. INTRODUCTION

The advent of the third generation (3G) gravitational-wave (GW) observatories, such as the Cosmic Explorer (CE) [1, 2] and the Einstein Telescope (ET) [3], will offer the possibility to reach cosmological distances, thanks to an order of magnitude improved sensitivity compared to the current generation of detectors such as Advanced LIGO [4], Advanced Virgo [5], and KAGRA [6]. Indeed, 3G observatories will have unprecedented sensitivity to detect coalescence events from an epoch when the Universe was still in its infancy assembling its first stars and will routinely detect mergers with stupendously large signal-to-noise ratios of several thousands [7, 8]. An order of magnitude greater redshift reach and access to extremely high-fidelity signals compared to current interferometers promises many new discoveries, while allowing completely independent, precision tests of cosmological models, alternative gravity theories, and astrophysical scenarios of compact binary formation and evolution [9]. With an expected rate of hundreds of thousands of binary coalescence signals each year on top of weak but persistent radiation from isolated neutron stars, rare bursts from supernovae and other transient sources and stochastic backgrounds, 3G observatories demand novel algorithms for signal detection and characterization. Therefore, a proper understanding of systematics arising from overlapping loud and quiet signals alike will answer a range of scientific questions that are at the forefront of fundamental physics and astronomy, as well as a realistic estimation of the computational cost.

According to current estimates, 3G observatories are expected to detect tens of thousands of binary black hole (BBH) and millions of binary neutron star (BNS) mergers each year [10, 11]. If we take account of the fact that signals will last longer due to a lower starting frequency ($3 \text{ Hz for ET}$ and $7 \text{ Hz for CE}$), then it is clear that 3G data will be dominated by many overlapping signals [12–15]. This poses two challenges: first, the detection of individual signals could, in principle, be affected by the presence of multiple signals. Second, the current Bayesian inference methods [16, 17] may not guarantee unbiased estimation of source parameters which is crucial to deliver the science promises of 3G observatories.

The Laser Interferometer Space Antenna (LISA) is expected to produce a data set containing many overlapping astrophysical signals. Indeed, galactic white dwarf binaries are persistent sources of gravitational waves and they produce a “foreground” noise [18] that could maskerade the detection and parameter estimation of other astrophysical signals. Several authors have studied the problem of both detection [19, 20] and Bayesian inference [21, 22] in this context, while others have focused on the global solution to the full family of potential signals [23, 24] as well as implementation of novel algorithms that are specifically suited to characterize multiple overlapping signals [25].

The problem of overlapping signals producing a confusion background in future terrestrial detectors was identified more than a decade ago [26]. Detecting overlapping GW signals has been shown to be possible by two ET mock data challenges [12, 13]. These studies were able to correctly identify and recover signals even when they were overlapping with multiple other signals. However, no effort to study the problem of inference in the case of terrestrial detectors has so far been made.

In this work, we explore Bayesian parameter inference of overlapping signals in terrestrial detectors. Even though the signal detection may provide unbiased results, there is no guarantee that the parameter inference in the case of overlapping signals is possible within the current framework. This is because current methods heavily rely on the efficiency of sampling algorithms, which are used to explore the posterior distribution of parameters. If we analyze overlapping signals with the current parameter estimation (PE) procedures (i.e., the assumption that the parameter space for multiple signals is the same as in the case of data containing only one signal at a time), we expect Markov Chains and the posterior distribution to exhibit a non-trivial behavior such as slowly or non-convergence of chains, multi-modal and biased posterior distributions, etc. To this end we will deploy the Fisher information matrix formalism, to gauge the limit between the region where overlapping signals could lead to biases in parameter
inference and the region where they don’t. The Fisher study tells us that as long as the difference in merger time $\Delta t_C$ of two overlapping signals is larger than the accuracy $\delta t$ with which their merger times can be measured (i.e., $\Delta t_C \gg \delta t$), irrespective of how long the individual signals are, parameter inference will not be a problem. We exploit this result in Bayesian inference by choosing the prior on the merger epoch $\delta t$ of overlapping signals there are at any one time but if two or more signals have their merger times lie within a duration $\Delta t$. This is what we will set out to compute next.

## A. Overlapping signals of the same family

Let $r$ denote the Poisson detection rate of a given signal family (BBH or BNS). In an interval $\Delta t$, the expected Poisson rate is $\nu = r\Delta t$ and the probability of observing exactly $k$ mergers during $\Delta t$ is given by

$$P_k(\nu) = \frac{\nu^k e^{-\nu}}{k!}.$$  

Thus, the probability of observing two or more mergers during $\Delta t$ is

$$P_{k \geq 2} = \sum_{k=2}^{\infty} P_k(\nu) = \sum_{k=2}^{\infty} \frac{\nu^k e^{-\nu}}{k!} = 1 - e^{-\nu}(1 + \nu).$$  

We have made use of the fact that the Poisson distribution is normalized, namely $\sum_{k=0}^{\infty} P_k(\nu) = 1$. To compute the number of chunks $N_{k \geq 2}$ in which two or more mergers will be observed we must multiply the probability $P_{k \geq 2}$ by the number of chunks

$$n_M \equiv T/\Delta t$$

in an observational period $T$.

$$N_{k \geq 2} \equiv P_{k \geq 2} n_M = \left[1 - e^{-\nu}(1 + \nu)\right] \frac{T}{\Delta t}.$$  

Substituting $\Delta t = \nu/r$ and noting that $N_T \equiv rT$ is the total number of signals detected during the period $T$, we get

$$N_{k \geq 2} = \left[1 - e^{-\nu(1 + \nu)}\right] \frac{N_T}{\nu}.$$  

It is easy to see that in the limit $\Delta t \rightarrow 0$ (equivalently, $\nu \rightarrow 0$), $N_{k \geq 2} \approx \nu N_T / 2$. The factor of $1/2$ assures that the number of instances when two or more signals are found in a chunk is never greater than half of the total number of observed signals but it is also weighed down by the Poisson rate $\nu$. In the other limit, when $\Delta t \rightarrow T$ (and $\nu \gg 1$), $N_{k \geq 2} \approx 1$ but less than 1.

Figure 1 plots the number of chunks $N_{k \geq 2}$ in which we can expect to find two or more mergers in a year’s worth of data (i.e., using $T = 1$ yr and $\nu = r \Delta t$). Also indicated in the plot are the detection rate of BBH (BNS) which is expected to be in the range $r_{\text{BBH}} \in [1.6, 4.8] \times 10^{-3} \text{ s}^{-1}$ ($r_{\text{BNS}} \in [3.5, 35] \times 10^{-3} \text{ s}^{-1}$, respectively) [15] in a 3G detector network comprising of one ET and two CE (one in north America and the other in Australia). As we shall see in Sec. III, parameter inference should not be a problem if the difference in coalescence times of a pair of signals is larger than ~1 s; this is indicated in Fig. 1 by the horizontal line drawn at $\Delta t = 1$ s. Thus, in Sec. IV we will focus on Bayesian inference of signals whose merger times differ by about one second. We see that at the higher end of the BNS rate, we expect ~15,000 one-second long chunks with two or more mergers while at the lower end of the BNS rate this number is ~200. Likewise, ~300 chunks will contain two or more BBH mergers at the higher end of the BBH detection rate while this number is ~40 at the lower end of the BBH rate.

The detection rate of BBH signals in the current detector network of LIGO, Virgo and KAGRA at their design sensitivity

## II. NUMBER OF OVERLAPPING SIGNALS

The number of overlapping signals depends on (a) the typical duration of signals and (b) the rate at which they arrive at the detector. At the leading order, the length $\xi$ of a coalescing compact binary signal starting from a gravitational-wave frequency $f_s$ merger is given by

$$\xi = \frac{5}{256} \left(GMf_s^3\right)^{-5/3}(c_s^3f_s^{-8/3}),$$  

(1)

where $G$ is Newton’s constant, $c$ is the speed of light and the chirp mass $M$ is related to the component masses $m_1$ and $m_2$ via $M \equiv (m_1 m_2)^{1/3}/(m_1 + m_2)^{1/5}$. A BNS system consisting of a pair of $1.4M_\odot$ would last for $\xi \approx 10^{0.3}$ s starting from a frequency of $f_s = 10$ Hz (relevant for Advanced LIGO and Advanced Virgo), 1.8 hr for $f_s = 5$ Hz (CE) and almost 7 hr for $f_s = 3$ Hz (ET). A source of intrinsic chirp mass $M$ at a cosmological redshift of $z$ would appear in the detector to have a chirp mass of $(1+z)M$, and hence lives for a shorter duration in a detector’s sensitivity band. Thus, BNSs ($1M_\odot \leq m_1, m_2 \leq 3M_\odot$) could last for tens of minutes to several hours in band while BBH signals ($3M_\odot \leq m_1, m_2 \leq 50M_\odot$) could last for tens of seconds to thousands of seconds.

The cosmic merger rate of compact binary coalescences determined by the first three observing runs of LIGO and Virgo [27, 28] implies that in a network of 3G observatories the detection rate $r$, defined as the number of signals whose matched filter signal-to-noise ratio is larger than 12, lies in the range $r_{\text{BBH}} \in [5 \times 10^4, 1.5 \times 10^5]$ yr$^{-1}$ for BBHs and $r_{\text{BNS}} \in [10^5, 10^6]$ yr$^{-1}$ for BNSs [15]. Thus, given that signals last for several hours, 3G data would contain several loud overlapping signals. We shall see below that for the purpose of parameter inference the relevant quantity is not how many overlapping signals there are at any one time but if two or more
Thus, the probability of observing multiple mergers in a chunk of size 1 s is at best \( r \sim 2.3 \times 10^{-5} \, s^{-1} \) (or 730 yr\(^{-1}\)) \([28]\). Thus, the probability of observing multiple mergers in a chunk of size 1 s or less is negligibly small in the Advanced detector era. This will also be the case in the \( \Lambda \)+ era \([29]\) where the detection rates are expected to be 3 times larger.

### B. Overlapping signals from two different families

If the detection rate of signal families A and B are \( r_A \) and \( r_B \), then probability that one or more mergers of each of these signal families would occur during an interval \( \Delta t \) is

\[
P_{A,k\geq 1} = 1 - e^{-\Delta t r_A}, \quad P_{B,k\geq 1} = 1 - e^{-\Delta t r_B}. \tag{6}
\]

Thus, the probability \( P_{AB} \) that an interval \( \Delta t \) contains one or more from each of the two signal families is simply the product \( P_{AB} = P_{A,k\geq 1} P_{B,k\geq 1} \). If the rates are small, this reduces to \( P_{AB} = (\Delta t)^2 r_A r_B \) and the number of such chunks over a period \( T \) is \( N_{AB} = (\Delta t)^2 r_A r_B T = N_A N_B \Delta t A \), where \( N_A \) and \( N_B \) are the total number of mergers during the period \( T \) of families A and B, respectively, and \( n_{\Delta t} = T/\Delta t \) is the number of chunks of width \( \Delta t \) during \( T \). Using the range of BNS and BBH rates quoted before, we find that \( N_{AB} \) would lie in the range 170–5100 for \( T = 1 \) yr and \( \Delta t = 1 \) s.

From the foregoing discussions it is clear that a small but significant fraction of signals would have their coalescence time within an interval of 1 s. As we shall see in the next Section, due to their long duration, overlapping BNS signals are far less correlated with each other than overlapping BBH signals. For the same reason, a pair of overlapping BNS and BBH signals are poorly correlated. Hence, in the Bayesian inference problem (Sec. IV) we will only consider overlapping BBH signals.

### III. COVARIANCE AMONG OVERLAPPING SIGNALS

If two signals are well separated then the covariance between their parameters is zero and we do not expect one signal to affect the parameter inference of the other. As we bring the two signals closer together in time, at some point the presence of one of the signals will begin to bias the estimation of parameters of the other. In this Section we estimate the covariance between the parameters of a pair of overlapping signals using the Fisher matrix formalism. Although Fisher matrix is valid in the limit of large signal-to-noise ratios, any inferences we can draw from the correlation will guide us in choosing the parameter space of compact binaries where systematic biases could be large.

To this end, we assume that the data contains a pair of signals \( s_A \) and \( s_B \) buried in stationary, Gaussian noise \( n \). The detector output is a sum of the overlapping signals buried in detector noise:

\[
x(t) = n(t) + s_A(t, \lambda^{(A)}_a) + s_B(t, \lambda^{(B)}_a). \tag{7}
\]

where \( \lambda^{(A)}_a, \lambda^{(B)}_a \), for \( a = 1, \ldots, p \), are the set of parameters...
corresponding to signals \(s_A\) and \(s_B\), respectively. Note that since both \(s_A\) and \(s_B\) are assumed to belong to the same signal family they are specified by the same number of parameters. Furthermore, we shall only consider a single detector for this exercise. The relevant parameters for a binary with non-spinning companions are the chirp mass \(M_c\), symmetric mass ratio \(\eta\), the epoch \(t_C\) when the signal amplitude reaches its peak and the phase \(\phi_C\) of the signal at that epoch and so: 

\[
\lambda_c^{(A)} = (M_c^{(A)}, \eta^{(A)}, t_C^{(A)}, \phi_C^{(A)}) \text{ and similarly for signal } s_B. 
\]

In this Section, we assume the waveform model to be the TaylorF2 approximation at 3.5 post-Newtonian order \([30–32]\). The epoch of merger is taken to be the time when the instantaneous gravitational-wave frequency reaches the value \(f_{\text{ISCO}} = (6^{3/2}\pi M)^{-1}\), which corresponds to be the inner most stable circular orbit for Schwarzschild black holes.

For the computation of the covariance matrix it is more convenient to consider that the data contains only one signal, i.e., the sum of the two signals \(s = s_A + s_B\), and it is characterized by a double number of parameters: \(\theta_a = \lambda_a^{(A)}\) for \(a = 1, \ldots, p\) and \(\theta_a = \lambda_{a-p}^{(B)}\) for \(a = p + 1, \ldots, 2p\). For a noise background that is stationary and Gaussian the covariance matrix \(C\) which is inverse of the Fisher matrix \(\Gamma\), is given by:

\[
C_{ab} = \Gamma_{ab}^{-1}, \quad \Gamma_{ab} = \left(\frac{\partial s}{\partial \theta_a}, \frac{\partial s}{\partial \theta_b}\right). \tag{8}
\]

Here the scalar product of two waveforms (or any pair of functions of time for that matter) \(h\) and \(g\) is defined as

\[
\langle h, g \rangle \equiv 4\Re \int_{f_{\text{low}}}^{f_{\text{high}}} \tilde{h}(f) \tilde{g}^*(f) \frac{S_b(f)}{S_h(f)} \, df, \tag{9}
\]

where \(\Re\) stands for the real part of the integral, \(\tilde{h}\) and \(\tilde{g}\) are the Fourier transforms of the signals \(h\) and \(g\), respectively, \(g^*\) denotes the complex conjugate of \(g\) and \(S_b(f)\) is the one-sided noise spectral density of the detector. In our study we will use either the noise spectral density of Advanced LIGO \([4]\) or that of the Cosmic Explorer \([9]\). The lower frequency cutoff \(f_{\text{low}}\) is chosen to be 20 Hz for Advanced LIGO and 5 Hz for Cosmic Explorer. The upper frequency cutoff \(f_{\text{high}}\) is assumed to be the larger of the inner-most stable circular orbit frequency of the two overlapping signals, i.e., \(f_{\text{high}} = \max[(6^{3/2}\pi M_1)^{-1}, (6^{3/2}\pi M_2)^{-1}]\), where \(M_1\) and \(M_2\) are the total mass of the two overlapping signals.

The Fisher matrix contains interference terms of the following type:

\[
\Gamma_{\alpha, \beta'} = \begin{pmatrix} \frac{\partial s_A}{\partial \lambda_{\alpha}} \frac{\partial s_B}{\partial \lambda_{\beta'}} & \frac{\partial s_A}{\partial \lambda_{\alpha}} \frac{\partial s_B}{\partial \phi_{\beta'}} & \frac{\partial s_A}{\partial \phi_{\alpha}} \frac{\partial s_B}{\partial \lambda_{\beta'}} & \frac{\partial s_A}{\partial \phi_{\alpha}} \frac{\partial s_B}{\partial \phi_{\beta'}} \end{pmatrix}. \tag{10}
\]

Covariances are of primary interest in this Section as they can tell us the degree to which the presence of one signal affects the
parameter inference of the other. In order to measure the extent of covariance we consider two sets of overlapping signals:

1. overlapping BBHs with masses:

\[
(m_1^{(A)}, m_2^{(A)}) = (21 M_\odot, 15 M_\odot) \quad (11)
\]
\[
(m_1^{(B)}, m_2^{(B)}) = (33 M_\odot, 29 M_\odot). \quad (12)
\]

2. overlapping BNSs with companion masses:

\[
(m_1^{(A)}, m_2^{(A)}) = (1.45 M_\odot, 1.35 M_\odot) \quad (13)
\]
\[
(m_1^{(B)}, m_2^{(B)}) = (1.50 M_\odot, 1.40 M_\odot). \quad (14)
\]

Furthermore, in all cases we choose

\[
(i_c^{(A)}, \phi_c^{(A)}) = (0, 0), \quad (i_c^{(B)}, \phi_c^{(B)}) = (\tau, \pi/3), \quad (15)
\]

and vary \(\tau\) over the range \([-2, 2]\) s.

The covariances between the chirpmass, symmetric mass ratio and epoch of coalescence are plotted in Fig. 2 as a function of the parameter \(\tau\) for overlapping BBHs (top panels) and BNSs (bottom panels) for noise spectral densities of Advanced LIGO (left panels) and Cosmic Explorer (right panels). Other cross-covariances are negligibly small and not shown. What we plot are the normalized covariances, i.e., a combination of the correlation coefficients defined as:

\[
\sigma_{ab} \equiv \frac{C_{ab}}{\sqrt{C_{aa}C_{bb}}}, \quad a \neq b. \quad (16)
\]

This quantity is strictly bounded between \(-1\) and \(+1\). A correlation coefficient of \(+1\) implies that the parameters are perfectly correlated, \(-1\) implies they are perfectly anti-correlated, and a value of \(0\) would imply they are uncorrelated. We will take \(\sigma_{ab} \sim 0.1\) to be small enough to indicate that the presence of the second signal does not significantly bias parameter inference of the other signal. For the purpose of parameter inference correlations are only important when two overlapping signals are of comparable strengths. If the amplitude one of them is far smaller than the other then the weaker signal has negligible influence on the parameter inference of the stronger signal.

In all cases, the correlation coefficients are negligibly small beyond \(|\tau| \sim 0.5\) s except in the case of Advanced LIGO when the correlation continues to be significant up to \(|\tau| \sim 1\) s. The correlation remains insignificant in the case of BNSs both in Advanced LIGO and Cosmic Explorer except when \(\tau \sim 0\).

From this analysis, we conclude that the parameter inference of overlapping BNS signals is likely to be less severe than that of overlapping BBH signals. We will therefore consider only the latter class of signals in the remainder of this paper.

IV. BAYESIAN INFERENCE OF OVERLAPPING SIGNALS

In this Section, we support the results we have derived using the Fisher information matrix formalism (Sec. III) with a full Bayesian inference procedure. With this parameter estimation (PE) process, we are able to fully explore the posterior distribution of the parameters that generated the signals. This is important, because it allows us to confirm the presence (expected from the Fisher study) of distinct maxima in the posterior, one for each signal coalescing within the time chunk considered. Moreover, thanks to this numerical approach, we can explore more carefully the region where biases are expected (i.e., within \(\Delta_\theta < 1\) s), assessing their significance and gauging the conditions for which they seem to happen.

Within the Bayesian framework, given a set of parameters \(\lambda\) describing a compact binary coalescence (CBC) waveform \(h(\lambda, t)\), we can write the posterior distribution for \(\lambda\) as:

\[
P(\lambda | x, h) = \frac{\pi(\lambda) \mathcal{L}(x | \lambda, h)}{\mathcal{Z}(x)}, \quad (17)
\]

where \(x\) is the detector output. This posterior can be explored by using a sampling algorithm (e.g., MCMC, nested sampling). As in Sec. III, assuming that the data \(x\) contains two overlapping signals \(s_A\) (signal \(A\)) and \(s_B\) (signal \(B\)), then it can be written as:

\[
x = n + s_A + s_B, \quad (18)
\]

where \(n\) is the noise of the interferometer. Note that, in principle, to perform a Bayesian analysis of two or more overlapping signals we should broaden the parameter space, e.g., \(\theta = [\lambda^A, \lambda^B]\), in order to account for the presence of multiple overlapping signals. However, since running a sampling algorithm requires significant amount of computational resources, in most cases this is not required. In fact, as argued in Sec. III, if the signals’ coalescence times are wide apart we do not expect the presence of one signals to influence posterior distribution of parameters of the other. For this reason, in what follows we consider the parameter space of a single CBC signal. We will return on this point later on when discussing possible biases arising because of this choice.

A. Choice of signal families

As already mentioned, in this analysis we focus only on BBH signals. This choice is motivated by the fact that covariances among overlapping BNS signals as demonstrated in Sec. III is negligibly small when their merger times differ by more than a few milliseconds—a duration during which we do not expect more than one merger to occur even in 3G detectors. Moreover, BNS signals last for several hours in 3G detectors and tens of minutes in Advanced LIGO and Virgo implying Bayesian inference takes a formidable amount of computational resources (although new algorithms are already showing the promise of greatly reducing the computational requirement [32]). Furthermore, we also restrict our analysis using Advanced LIGO sensitivity. As argued before, LIGO is not affected by the problem of overlapping signals, because the rate and the duration of the signals are far too small to create any overlap. Here we are not really interested in reproducing a realistic set of overlapping data, though; instead, we want to focus on the parameter estimation process. To do so, there is no substantial advantage in using 3G mock data: we expect that our conclusions will be valid even if they are based on the analysis Advanced LIGO mock data.

The parameters of the overlapping BBH signals used in Bayesian inferences is the same as what we used in Sec. III:
nonspinning BBHs with masses as given in Eq. (12) and coalescence times and phases as given in Eq. (15). We ignore the position of the sources in the sky and their orientation relative to the detectors (setting all angles to zero). We do, however, include in our analysis the luminosity distance \( d_L \) of the source. The parameter space we use in our analysis is thus:

\[
\lambda = \{m_1, m_2, \phi_C, t_C, d_L\}
\]

Note that our choice of sky position is the worst case scenario, because we are considering the two sources to have the same exact location in the celestial sphere. In reality, if overlapping signals arrive from different directions in the sky, they will have different phase coherence amongst a network of detectors and thus easier to discriminate. Thus, since our choice of sky position is the worst case scenario, the parameter estimation problem can only be better when sky position and orientation are taken into account.

To explore different configurations of the parameters, we vary the two luminosity distances of the sources \( d_L^{(A)} \) and \( d_L^{(B)} \), keeping the distance of one of the sources fixed to 1 Gpc and setting the other at either 500 Mpc, 1 Gpc, or 2 Gpc. We also vary the time shift \( \tau \) defined in Eq. (15) as the epoch coalescence of signal B. The resulting variations in the parameter sets are:

\[
\tau = [-1.5 \text{ s}, -1.0 \text{ s}, -0.5 \text{ s}, 0.0 \text{ s}, 0.5 \text{ s}]
\]

\[
d_{L}^{(B)} = 500 \text{ Mpc}, d_{L}^{(A)} = 1 \text{ Gpc}, d_{L}^{(B)} = 2 \text{ Gpc}
\]

With these choices, there are 25 different possible configurations, each of which is analyzed for Bayesian parameter inference.

In the inference problem we use a signal model that accurately represents the BBH waveforms. To this end, we use the IMRPhenomV2 approximant to create waveforms in the frequency domain, fixing the low frequency cutoff to be 20 Hz, which is consistent with the minimum frequency used in the LIGO/Virgo PE. In Fig. 3, we plot the two waveforms in the time domain, for the different configurations of the parameters. We vary only the time shift since changing the distances of the signals results in a simple re-scaling of the amplitudes as we have neglected the effect of redshift on masses. The resulting overlapping waveform is plotted as well. In Table I, we compute the expected matched filter SNR for the different possible configurations of the parameters, here we focus on the distances, since the coalescence time does not affect the SNR value.

### Table I. SNRs for the two signals we have chosen to focus on in our analysis, created with different values of the luminosity distances \( d_L \).

| SNR  | \( d_L = 0.5 \text{ Gpc} \) | \( d_L = 1 \text{ Gpc} \) | \( d_L = 2 \text{ Gpc} \) |
|------|-----------------|-----------------|-----------------|
| signal A | 49.8            | 24.9            | 12.4            |
| signal B | 76.2            | 38.1            | 19.0            |

Note that applying a time shift to the signals do not change the value of the SNR.

### B. Setting up Bayesian inference runs

Having created the mock data with overlapping signals we next focus on parameter inference. Our analysis uses two LIGO interferometers, but our conclusions are not significantly affected by this choice; considering a different detector network would simply result in different SNRs for the signals as we
FIG. 4. Summary of the results for the set of 25 runs, each one with a different configuration of the parameters $\tau$, $d^\text{A}_L$, and $d^\text{B}_L$. A runs are shown on the left panel, and B runs are on the right panel. The recovered values for the masses $m_1$ and $m_2$ are plotted, together with their uncertainties. The true values of the masses for signal A and signal B are highlighted with the dashed horizontal lines. Note that the points in the plots, referring to the same time shift $\tau$, are slightly shifted with respect to their exact value of $\tau$ so that they do not overlap with each other. The $\tau = 0.0$ s runs are highlighted with a grey shadowed band (see discussion in Sec. IV D). The inset in the right panel shows the distributions for the coalescence time in the form of a ‘violin’ plot: these are just samples to show that the coalescence time is recovered very well (within the ms range) in every configuration.

are not focusing on the sky position of the source. Although this could in principle change the heights of the peaks in the posterior distribution, we do expect it to influence their relative ratios significantly, and hence the PE process we consider is expected to hold for any network.

The data set consists of 4 s of mock data from the two LIGO interferometers with a stationary, Gaussian noise background. 4 s is large enough to span the full length of the longer signal. We use the package \texttt{bilby} [17, 34] to facilitate injection of signals into mock data that mimics the power spectral density of Advanced LIGO.

For Bayesian inference of the parameters of the injected signals we use the \texttt{bilby} software package to perform parameter inference, running the \texttt{dynesty} sampler [35]. \texttt{Dynesty} is a dynamic, nested sampling algorithm [36, 37], which is well suited for our purposes because it quickly achieves convergence, but at the same time it is able to handle non-trivial, multi-modal distributions better than MCMC-based algorithms. We allow the sampler to explore the likelihood surface with respect to all the parameters except $\phi_C$, over which the likelihood is analytically marginalized, and $d_L$, over which the likelihood is numerically maximized. Marginalization over $\phi_C$ and $d_L$ correctly accounts for the effects of the parameters $\phi_C$ and $d_L$ on the resulting 3-d posterior [38, 39]. Consequently, our final posterior distribution is over the parameters $m_1$, $m_2$, and $t_C$.

C. Bayesian priors

At the beginning of the analysis, we have to set the priors on the various parameters. We consider a uniform prior on the phase $\phi_C$, with periodic boundary conditions, a power-law prior on the luminosity distance, $p(d_L) \propto d_L^\alpha$ with $\alpha = 2$, and a uniform prior on the two masses $m_1$ and $m_2$ over the range $[10 M_\odot, 50 M_\odot]$. As for the coalescence time, selecting the best possible prior turns out to be a game-changing strategy. In fact, running a simulation with a wide prior on the time $t_C$ that spans the merger times of the two overlapping signals leads to significant problems: while one of the two signals is always recovered correctly, the other is completely ignored by the sampling algorithm. A wide prior on $t_C$, therefore, would only allow us to infer the parameters of the louder signal, without access to the weaker one.

However, as already pointed out, previous work suggests that signal can always be detected, even if they are overlapping, and their merger time correctly identified [12, 13]. The detection of a signal allow us to know its epoch of coalescence with very low uncertainty (at the order of 10 ms). For this reason, we can assume to know the time of coalescence of the two overlapping signals with a good degree of accuracy, and constrain our parameter space choosing a prior on the coalescence time which is centered on the (fiducial) true value of the time $t_C$, with a width of 100 ms. In this way, for each of the signals we can isolate the region of the parameter space where we expect
to find the correct parameters. This choice allows us to recover the correct parameters for both signal A and signal B.

Therefore, for each of the 25 injections, we run the Bayesian inference procedure two times: the first one (we refer to it as run A) aims to recover the true values of the parameters of signal A; to this end, since \( t_{\text{c}}^{\text{A}} = 0.0 \) s, we set the prior on the coalescence time centered around zero. Run B, on the other hand, focuses on the signal B peak in the parameter space; thus, the prior is chosen to be centered in \( t_{\text{c}} = \tau \).

D. Results

In Fig. 4, we show the results for the recovered values of the parameters: results from run A are shown on the left panel, while run B ones are shown on the right panel. Note that in the two panels only the values for the masses are shown: that is simply because, since we are analytically marginalizing over \( \phi \) and \( d_{\text{L}} \), the only parameters for which we can create a posterior distribution are the two masses \( m_1 \) and \( m_2 \), and the coalescence time \( t_{\text{c}} \). However, the latter is always recovered correctly and with a very small error on its value (around 1 ms), as shown in the inset the right-hand panel of Fig. 4. The ‘violin’ plot was created from the distributions of \( t_{\text{c}} \) for a single time shift configuration, namely run B with \( \tau = -1.5 \) s, which clearly demonstrates that \( t_{\text{c}} \) is determined to millisecond accuracy.

Furthermore, we show the corner plots in Fig. 5 for two randomly chosen runs (run A on the right side and run B on the left), created with the following choice of the parameters: \( \tau = -1.0 \) s, \( d_{\text{L}}^{\text{A}} = d_{\text{L}}^{\text{B}} = 1 \) Gpc. The corner plots show that the results resemble the ones obtained when the data contains only one signal: the posterior on the two masses presents a degeneracy along the line of equal chirp mass, and the true values of the masses (and of the coalescence time) lie inside the 2\( \sigma \) credible region. Therefore, we confirm that – as expected from discussion in Sec. III – using current PE algorithms it is possible to treat the case of overlapping signals. We find that that setting the appropriate prior on the coalescence time \( t_{\text{c}} \), as determined by the search pipeline, is critical in determining the parameters of the signal without any bias.

Looking at Fig. 4, we note that promising results are found for all the possible parameter configurations, except for the case in which the time shift \( \tau \) is equal to zero (i.e., the two signals are coalescing at the exact same time). In fact, for \( \tau \neq 0.0 \) s, the majority of the A runs (left panel), give results for the parameters that are within one (sometimes two) sigma from the true values \( \mathcal{L}^{\text{A}} = \{ m_1 = 21 \text{ M}_\odot, m_2 = 15 \text{ M}_\odot \} \). Similarly, in most of the considered cases, the B runs shown on the right panel give results for the parameters that are close to the true values \( \mathcal{L}^{\text{B}} = \{ m_1 = 33 \text{ M}_\odot, m_2 = 29 \text{ M}_\odot \} \).

If the time shift is zero (the corresponding runs are highlighted with a grey shadowed box in Fig. 4), then both runs recover only same signal. This means the the two-peaked posterior is not well sampled by the algorithm. The chains are able to find only one peak, and they are not able to sample efficiently the other one: a possible explanation for this behaviour is that the two peaks have very different values (on average, the two peaks differ for 100 – 500 orders of magnitude), thus sampling the minor peak turns out to be impossible for the algorithm.

This hypothesis is supported by an examination of Table I: the highest peak is the one that’s associated with the highest SNR. As expected, then, in the runs in which signal B is louder than signal A, we recover a peak close to \( \mathcal{L}^{\text{B}} \), and vice-versa.

We note here that the fact that the algorithm was able to focus on both peaks at the same time was already clear from the results in the case of a wide prior on the coalescence time: if we do not isolate the correct value of \( t_{\text{c}} \) by using a narrow \( t_{\text{c}} \) prior, again we recover only the loudest peak and its corresponding time shift, with no sign of the other peak in the posterior. We could also try to isolate one peak at a time by imposing a narrower prior on the two masses \( m_1 \) and \( m_2 \). This is more difficult though, because of the relatively large uncertainties of the recovered values of the masses (see e.g., Fig. 5). For this reason, it is not easy to isolate one peak by preventing any influence of the other one. Moreover, imposing a narrower prior on the component masses does not have a clear justification as signal detection pipelines return fiducial values for the two masses only within a large credibility region, and this could prevent in principle a reliable narrowing of the mass priors.

In some particular cases, the recovered masses are more than two-sigma away from the expected values of \( m_1 \) and \( m_2 \). This behavior is more common for \( \tau \neq 0.0 \) s (note that for \( \tau = 0.0 \) s, the expected values are the ones belonging to the loudest signal), but it happens also for other values of \( \tau \). In order to assess the reasons for this wrong behavior, we run the biased configurations without the presence of the interferometer noise. This allows us to discern between the influence of the noise from the one linked to the presence of an overlap in the data. We find that, in the case of the runs with \( \tau \neq 0.0 \) s, the results are substantially better in the absence of noise, with the true values re-entering inside the 90% credibility region. In some cases with \( \tau = 0.0 \) s, however, we still find the same kind of biases even with zero noise. In these latter cases, we conclude that the overlap may play a role in preventing to recover the correct values for the parameters.

Therefore, our Bayesian inference analysis confirms the results we found in Sec. III: if the two signals do not coalesce too close in the time domain, then two distinct peaks are present in the posterior, and they can be well-sampled if a suitable prior on the coalescence time is chosen. However, if the time shift between the two signals is small enough (i.e., smaller than the resolution we have adopted for our analysis, \( \tau = 0.5 \) s), then biases could arise, and a more careful analysis of the Bayesian inference methods is needed.

V. DISCUSSION AND OUTLOOK

We presented a Bayesian inference analysis in the case of overlapping gravitational waves signals. Our goal was to assess the capabilities of current Bayesian inference infrastructure to handle the non-trivial case of one or multiple overlaps happening within a data segment. This problem is destined to play a major role in 3G detector planning, since the dramatic increase in sensitivity will result in a great number of signals coalescing within a few seconds.

We started from a study based on the Fisher matrix formalism, in which we analyzed the correlation between two
overlapping signals. In this way, we were able to determine whether in some regions of parameter space the two signals are strongly correlated between each other, thus preventing a distinct inference procedure for one signal at a time. We found that BNS signals are not expected to be significantly correlated, and therefore their inference could be a problem only for coalescence times really close between each other (at the 10 − 100 ms level). BBHs, instead, suffer from the presence of a correlation starting from a much greater time shift $\tau$ (i.e., the difference between the two coalescence times). However, also for BBHs we find that for time shifts wider than $\sim 1$ s, we do not expect the Bayesian inference to be a significant problem.

We confirmed these findings with a full Bayesian analysis of the two overlapping BBHs. The analysis used the dynesty sampling algorithm to describe the posterior distribution for the parameters considered. We showed that, in order to sample a single peak without worrying for the presence of the other one, a possible solution is to impose a narrow prior around the fiducial value (provided by the signal detection pipeline) of the coalescence time of the signal of interest. This procedure allows to isolate one single peak at a time, and worked well in the configurations we explored. However, as the time shift approaches zero, we witnessed the emergence of biases that prevented the recovery of the true values of the parameters. These biases happened only in the zero time-shift case (although our time resolution was at the 0.5 s level so they could also happen for $0.0 < |\tau| < 0.5$ s, i.e., when the correlation between the two signals is expected to peak.

Dealing with these biases needs a different approach that we did not attempt in this work. One possible solution is to simply broaden the parameter space searching for multiple signals in the same Bayesian inference run. This significantly increases the computational costs of the Bayesian algorithms; however, since such an algorithm would be needed only in the worst case scenario of $|\tau| \sim 0$, then the number of overlaps in such a small time region will be rarely greater than two. This means that, in the light of our work, heavy Bayesian inference methods with a very large number of signals (and parameters) are not really necessary in the case of overlapping signals. Another possible solution to the biases would be to create an iterative procedure where one hierarchically determines the parameters of louder signals (as inferred from search algorithms) and subtracts them from the data before analysing weaker ones [11, 40]. Although it is not possible to use narrower priors on component masses to isolate multiple peaks, it should be possible to impose constraints on the chirp mass as found by detection pipelines. It is, however, important to ascertain the extent to which such constraints can be imposed by carrying out the detection problem on a large sample of injections and the accuracy with which detection pipelines are able to measure chirp mass. Our study did not include other source parameters such as companion spins, the position of the source in the sky and the orientation of binary relative to the detector frame. In particular, when overlapping signals arrive from different positions in the sky then they would, in general, have different coalescence times in different detectors, which might help to isolate one of the peaks better [41] but we need to study this problem more carefully. These and related problems will be explored in a future study.

ACKNOWLEDGMENTS

We thank Anuradha Samajdar, Justin Janquart, Chris Van Den Broeck and Tim Dietrich for sharing and discussing their results on a similar study [15]. We are indebted to useful comments by Rory Smith and Salvatore Vitale. EP is grateful to
Walter Del Pozzo for useful suggestions. EP was supported by INFN in the framework of the 2019 NSF-INFN Summer exchange program. SS is supported by the Eberly Postdoctoral Fellowship of Penn State. BSS is supported in part by NSF Grant No. PHY-1836779. The authors are grateful for computational resources provided by the LIGO-Caltech Computing Cluster. This paper has the LIGO document number P2100044.

[1] LIGO Scientific collaboration, Exploring the Sensitivity of Next Generation Gravitational Wave Detectors, Class. Quant. Grav. 34 (2017) 044001 [1667.68697].
[2] D. Reitze et al., Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO, [1997.04833].
[3] M. Punturo et al., The Einstein Telescope: A third-generation gravitational wave observatory, Class. Quant. Grav. 27 (2010) 194002.
[4] LIGO Scientific collaboration, Advanced LIGO, Class. Quant. Grav. 32 (2015) 074001 [1411.4547].
[5] VIRGO collaboration, Advanced Virgo: a second-generation interferometric gravitational-wave detector, Class. Quant. Grav. 32 (2015) 024001 [1406.2978].
[6] KAGRA collaboration, KAGRA: 2.5 Generation Interferometric Gravitational Wave Detector, Nat. Astron. 3 (2019) 35 [1811.08679].
[7] B. Sathyaprakash et al., Scientific Objectives of an Einstein Telescope, Class. Quant. Grav. 29 (2012) 124013 [1206.3631].
[8] S. Vitale and M. Evans, Parameter estimation for binary black holes with networks of third generation gravitational-wave detectors, Phys. Rev. D95 (2017) 064052 [1610.06917].
[9] D. Reitze et al., The US Program in Ground-Based Gravitational Wave Science: Contribution from the LIGO Laboratory, Bull. Am. Astron. Soc. 51 (2019) 141 [1903.04615].
[10] V. Baibhav, E. Berti, D. Gerosa, M. Mapelli, N. Giacobbo, Y. Bouffanais et al., Gravitational-wave detection rates for compact binaries formed in isolation: LIGO/Virgo O3 and beyond, Phys. Rev. D100 (2019) 064060 [1906.04197].
[11] S. Sachdev, T. Regimbau and B. Sathyaprakash, Subtracting compact binary foreground sources to reveal primordial gravitational-wave backgrounds, Physical Review D 102 (2020) .
[12] T. Regimbau, T. Dent, W. Del Pozzo, S. Giampapa, T. G. Li, C. Robinson et al., Mock data challenge for the einstein gravitational-wave telescope, Physical Review D 86 (2012) 122001.
[13] D. Meacher, K. Cannon, C. Hanna, T. Regimbau and B. Sathyaprakash, Second einstein telescope mock data and science challenge: Low frequency binary neutron star data analysis, Physical Review D 93 (2016) .
[14] T. Regimbau, M. Evans, N. Christensen, E. Katsavounidis, B. Sathyaprakash and S. Vitale, Digging deeper: Observing primordial gravitational waves below the binary black hole produced stochastic background, Phys. Rev. Lett. 118 (2017) 151105 [1611.08943].
[15] A. Samajdar, J. Janquart, C. V. D. Broeck and T. Dietrich, Biases in parameter estimation from overlapping gravitational wave signals in the third generation detector era, [2021.2021].
[16] J. Veitch et al., Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library, Phys. Rev. D91 (2015) 042003 [1409.7215].
[17] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu et al., Bilby: A user-friendly bayesian inference library for gravitational-wave astronomy, The Astrophysical Journal Supplement Series 241 (2019) 27.
[18] J. Crowder and N. J. Cornish, LISA source confusion, Phys. Rev. D 70 (2004) 082004 [gr-qc/0404129].
[19] N. J. Cornish and E. K. Porter, The Search for supermassive black hole binaries with LISA, Class. Quant. Grav. 24 (2007) 5729 [gr-qc/0612091].
[20] T. B. Littenberg, A detection pipeline for galactic binaries in LISA data, Phys. Rev. D 84 (2011) 063009 [1106.6355].
[21] N. J. Cornish and J. Crowder, LISA data analysis using MCMC methods, Phys. Rev. D 72 (2005) 043005 [gr-qc/0506059].
[22] J. Crowder and N. Cornish, A Solution to the Galactic Foreground Problem for LISA, Phys. Rev. D 75 (2007) 043008 [astro-ph/0611546].
[23] T. Littenberg, N. Cornish, K. Lackeoes and T. Robson, Global Analysis of the Gravitational Wave Signal from Galactic Binaries, Phys. Rev. D 101 (2020) 123021 [2004.08464].
[24] T. Robson and N. Cornish, Impact of galactic foreground characterization on a global analysis for the LISA gravitational wave observatory, Class. Quant. Grav. 34 (2017) 244002 [1705.09421].
[25] A. Pettitay, S. Babak, A. Sesana and M. de Araújo, Resolving multiple supermassive black hole binaries with pulsar timing arrays II: genetic algorithm implementation, Phys. Rev. D 87 (2013) 064036 [1210.2396].
[26] T. Regimbau and S. A. Hughes, Gravitational-wave confusion background from cosmological compact binaries: Implications for future terrestrial detectors, Phys. Rev. D 79 (2009) 062002 [0901.2958].
[27] LIGO Scientific, Virgo collaboration, Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo, Astrophys. J. 882 (2019) L24 [1811.12940].
[28] LIGO Scientific, Virgo collaboration, Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog, [2019.14533].
[29] KAGRA Scientific collaboration, KAGRA Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA, Living Rev. Rel. 21 (2018) 3 [1304.0678].
[30] B. S. Sathyaprakash and S. V. Dhurandhar, Choice of filters for the detection of gravitational waves from coalescing binaries, Phys. Rev. D 44 (1991) 3819.
[31] T. Damour, B. R. Iyer and B. S. Sathyaprakash, A Comparison of search templates for gravitational waves from binary inspiral, Phys. Rev. D 63 (2001) 044023 [gr-qc/0010099].
[32] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan and B. Sathyaprakash, Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors, Phys. Rev. D 80 (2009) 084043 [0907.0709].
[33] D. Finstad and D. A. Brown, Fast Parameter Estimation of Binary Mergers for Multimessenger Follow-up, Astrophys. J. Lett. 905 (2020) L9 [2009.13759].
[34] I. Romero-Shaw et al., Bayesian inference for compact binary coalescences with BILBY: Validation and application to the first LIGO–Virgo gravitational-wave transient catalogue, [2006.06714].
[35] J. S. Speagle, dynesty: a dynamic nested sampling package for estimating bayesian posteriors and evidences, Monthly Notices
of the Royal Astronomical Society 493 (2020) 3132–3158.

[36] J. Skilling, Nested Sampling.

[37] E. Higson, W. Handley, M. Hobson and A. Lasenby, Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation, Statistics and Computing 29 (2018) 891–913.

[38] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale et al., Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library, Phys. Rev. D 91 (2015) 042003.

[39] L. P. Singer and L. R. Price, Rapid Bayesian position reconstruction for gravitational-wave transients, Phys. Rev. D 93 (2016) 024013 [1508.03634].

[40] C. Cutler and J. Harms, BBO and the neutron-star-binary subtraction problem, Phys. Rev. D 73 (2006) 042001 [gr-qc/0511092].

[41] P. Christian, S. Vitale and A. Loeb, Detecting Stellar Lensing of Gravitational Waves with Ground-Based Observatories, Phys. Rev. D 98 (2018) 103022 [1802.02586].