


Bayesian inference of multimessenger astrophysical data: Methods and applications to gravitational waves

Matteo Breschi¹, Rossella Gamba¹, and Sebastiano Bernuzzi¹

Theoretisch-Physikalisches Institut, Friedrich-Schiller-Universität Jena, 07743, Jena, Germany

 (Received 30 January 2021; accepted 25 May 2021; published 9 August 2021)

We present BAJES, a parallel and lightweight framework for Bayesian inference of multimessenger transients. BAJES is a PYTHON modular package with minimal dependencies on external libraries adaptable to the majority of the Bayesian models and to various sampling methods. We describe the general workflow and the parameter estimation pipeline for compact-binary-coalescence gravitational-wave transients. The latter is validated against injections of binary black hole and binary neutron star waveforms, including confidence interval tests that demonstrate the inference is well calibrated. Binary neutron star postmerger injections are also studied using a network of five detectors made of LIGO, Virgo, KAGRA, and Einstein Telescope. Postmerger signals will be detectable for sources at $\lesssim 80$ Mpc, with Einstein Telescope contributing over 90% of the total signal-to-noise ratio. As a full scale application, we reanalyze the gravitational-wave transients catalog-1 black hole transients using the effective-one-body TEOBResumS approximant and reproduce selected results with other approximants. BAJES inferences are consistent with previous results; the direct comparison of BAJES and BILBY analyses of GW150914 shows a maximum Jensen-Shannon divergence of 5.2×10^{-4} . GW170817 is reanalyzed using TaylorF2 with 5.5PN point mass and 7.5PN tides, TEOBResumSPA, and IMRPhenomPv2_NRTidal with different cutoff frequencies of 1024 and 2048 Hz. We find that the former choice minimizes systematics on the reduced tidal parameter, while a larger amount of tidal information is gained with the latter choice. BAJES can perform these analyses in about 1 day using 128 CPUs.

DOI: [10.1103/PhysRevD.104.042001](https://doi.org/10.1103/PhysRevD.104.042001)

I. INTRODUCTION

Bayesian inference has become a standard method for the analysis of astrophysical and cosmological events, e.g., [1–5], since it offers a generic statistical framework to rigorously test hypothesis against observational information. Given a set of parametric models (hypothesis) and assumptions on the parameters (prior), Bayesian inference allows one to infer the parameters in terms of probability distributions and also to select the best-fitting model among competing hypotheses. In particular, Bayesian methods are central tools used in gravitational-wave (GW) astronomy to determine the source properties of an observed signal [6–8] and the related applications. Some examples are tests of general relativity [9,10], astrophysical population studies [11], inferences of the neutron star matter properties [12] and cosmology [13,14]. Furthermore, Bayesian inference offers the optimal framework to combine different observational datasets from multimessenger astronomical observations, like GW170817 and the electromagnetic (EM) counterparts [15–22]. Multimessenger inference of astrophysical phenomena such as binary neutron star (BNS) mergers is a fundamental resource to clarify the mechanism at the origin of the radiation, to obtain accurate inferences on the properties of the source, and to improve theoretical models gaining information from observational data (e.g., [23–25]).

In the last years, many efforts have been made by the scientific community in the development of sophisticated parameter inference tools for astronomical observations. In the context of GW data analysis, LALSuite [26] is the official software provided by the LIGO-Virgo Collaboration [27,28] and it offers a variegated framework for generic analysis of GW data. Other mature software for parameter estimation of GW transients are also available; some examples are the semianalytical integrator RIFT [29,30], the user-friendly library BILBY [31], and the inference module of the PYCBC package [32]. Bayesian software for parameter inference of other astrophysical transients have also been developed, such as tools for high-energy photons from compact objects and galaxy clusters [33–35], neutrino radiation [36], supernovae transients [37–39], pulsar arrival timings [40–42], and for cosmological inferences [43–46]. Current pipelines for the analysis of astrophysical and cosmological observations are targeted to specific applications. However, within a multimessenger framework, it is essential to develop a flexible pipeline capable of combining different datasets and physical models. This issue can be tackled allowing the infrastructure to merge different Bayesian models, extending the considered parameter space and generalizing the definition of the likelihood function. This implies the use of large amounts of data and computationally expensive models. It follows that

efficient parallelization techniques and well-calibrated proposal methods are necessary to optimize the performances of such a flexible pipeline.

In this work, we present BAJES [baIes], a PYTHON package for Bayesian inference developed at Friedrich-Schiller-Universität Jena. Our goal is to provide a simple, complete, and reliable implementation capable to robustly perform Bayesian inference on arbitrary sets of data, with specific functionalities for multimessenger astrophysics. The software is designed to be state of art, simple to use, and light weighted with minimal dependencies on external libraries. The paper is structured as follows. In Sec. II, we recall the basic concepts of Bayesian theory of probability. In Sec. III, we describe the design and the workflow of the BAJES software. Section IV describes the tools and the methods implemented in BAJES for the analysis of GW transients, including the available templates. In Sec. V, we describe the GW pipeline and the Bayesian framework of the GW model. In Sec. VI, we present a survey of injection studies and validation tests performed with artificial binary merger signals. Section VII shows the results of the BAJES pipeline on the GW events observed by the LIGO-Virgo interferometers [1,47]. Finally, we conclude in Sec. VIII. The paper concludes with Appendixes on the implemented sampling methods, the proposal methods, and a simple use example.

II. BAYESIAN INFERENCE

The task of a Bayesian inference is the formulation and the computation of conditional probabilities. It is possible to classify this topic in two main problems: parameter estimation (PE) and model selection (MS). With PE, we mean the evaluation of the characteristic distribution for the parameters that define the model of interest. On the other hand, with MS we refer to the discrimination between competing models in light of the data, comparing the suitability of different assumptions directly on the observation. In order to discuss how these tasks are achieved, in the following sections we recall the basic concepts of Bayesian theory of probability. By convention, we label the natural logarithm as log throughout all the paper.

A. Bayes' theorem

Given a set of observed data \mathbf{d} and a set of parameters θ , that characterizes our model within some background hypothesis H , it is possible to estimate the posterior distribution for θ using the Bayes' theorem [48–50],

$$p(\theta|\mathbf{d}, H) = \frac{p(\mathbf{d}|\theta, H)p(\theta|H)}{p(\mathbf{d}|H)}, \quad (1)$$

where $p(\mathbf{d}|\theta, H)$ is the likelihood function, $p(\theta|H)$ is the prior distribution, and $p(\mathbf{d}|H)$ is the evidence. The likelihood function describes the probability of observing the

data \mathbf{d} given θ and assuming that the hypothesis H is true. Therefore, it encodes the observational information and it predicts the agreement between the observed data \mathbf{d} and the expected outcome for every given sample θ of the parameter space. The prior distribution $p(\theta|H)$ depicts the knowledge on the parameters before performing the estimation. Usually, the functional form of this term is chosen in accordance with geometrical and/or physically motivated argumentation. The term $p(\mathbf{d}|H)$ is labeled as evidence and it represents the probability of observing the data \mathbf{d} given the hypothesis H . The evidence is also called marginalized likelihood since, according to the marginalization rule, it can be expressed as

$$p(\mathbf{d}|H) = \int_{\Theta} p(\mathbf{d}|\theta, H)p(\theta|H)d\theta, \quad (2)$$

where the integral is extended over the entire parameter space Θ . Subsequently, the posterior distribution $p(\theta|\mathbf{d}, H)$ represents the probability of the parameters θ in light of the data overhauled by our *a priori* information. The knowledge of $p(\theta|\mathbf{d}, H)$ allows us to compute the expectation of the statistical quantities of interest. For example, the mean values $E[\theta]$ can be estimated as

$$E[\theta] = \int_{\Theta} \theta p(\theta|\mathbf{d}, H)d\theta, \quad (3)$$

and, analogously, it is possible to infer the expectation of a generic function of the parameters θ ,

$$E[f(\theta)] = \int_{\Theta} f(\theta)p(\theta|\mathbf{d}, H)d\theta. \quad (4)$$

From this argumentation it follows that, in order to perform PE, we have to introduce a prior distribution $p(\theta|H)$ and a likelihood function $p(\mathbf{d}|\theta, H)$; then, the properties of the model are encoded in the posterior distribution $p(\theta|\mathbf{d}, H)$ that can be computed imposing Eq. (1).

B. Model selection

In Eq. (1), for a fixed set of assumptions H , the evidence acts like a normalization constant; however, this quantity plays a crucial role in the context of MS. If we are interested in comparing two competing hypotheses, H_A and H_B , quantifying which one is better explaining the data, in the Bayesian framework it is natural to introduce the odds ratio,

$$\mathcal{O}_A^B = \frac{p(H_B|\mathbf{d})}{p(H_A|\mathbf{d})} = \frac{p(H_B)p(\mathbf{d}|H_B)}{p(H_A)p(\mathbf{d}|H_A)}. \quad (5)$$

The term $p(H_i|\mathbf{d})$ represents the posterior probability for the i th hypothesis given the observed data and the ratio \mathcal{O}_A^B encodes the will of the data in favoring one model over another. Assuming that the two hypotheses are

equiprobable $p(H_B) = p(H_A)$, it is possible to reduce the computation to the ratio of the evidences, also known as Bayes' factor,

$$\mathcal{B}_A^B = \frac{p(\mathbf{d}|H_B)}{p(\mathbf{d}|H_A)}. \quad (6)$$

If $\mathcal{B}_A^B < 1$, then the hypothesis A is preferred by the data; otherwise, B is favored if $\mathcal{B}_A^B > 1$. However, this rule is not always straightforward since the estimation of the Bayes' factor might suffer from uncertainties [51,52]. Then, in a realistic scenario, more stringent bounds are required in order to prefer a hypothesis [53].

C. Joint distributions

Let us assume that we performed two independent observations, \mathbf{d}_1 and \mathbf{d}_2 , and each of them can be modeled using two sets of parameters, respectively, θ_1 and θ_2 . In general, it is possible to apply Bayes' theorem, Eq. (1), separately to every set of measurements. However, if the two events are not independent (e.g., the same physical process observed by two different observatories), the joint posterior distribution can be written as

$$p(\theta_1, \theta_2 | \mathbf{d}_1, \mathbf{d}_2, H) = \frac{p(\mathbf{d}_1, \mathbf{d}_2 | \theta_1, \theta_2, H) p(\theta_1, \theta_2 | H)}{p(\mathbf{d}_1, \mathbf{d}_2 | H)}, \quad (7)$$

where $p(\mathbf{d}_1, \mathbf{d}_2 | \theta_1, \theta_2, H)$ is the joint likelihood function and $p(\theta_1, \theta_2 | H)$ is the joint prior distribution. If the employed set of parameters θ_1 and θ_2 are independent, the joint probabilities simply correspond to the product of the single probability terms. However, if θ_1 and θ_2 correlate, the problem could easily become nontrivial; e.g., the intersection between the two parameter spaces $\Theta_1 \cap \Theta_2$ could be not empty, or the value of the two sets of parameters could be related with each other $\theta_i \equiv \theta_i(\theta_j)$, or unexpected correlations might appear introducing a larger parameter space.

D. Samplers

In a realistic scenario, the form of the likelihood function is not always analytically determinable and the parameter space has usually a large number of dimensions. For these reasons, the evaluation of the posterior distribution and the estimation of its integral are performed with stochastic techniques. In particular, sampling methods have proven to be reliable and generic tools for the analysis of non-analytical forms and multidimensional parameter spaces [7,46,54,55], capable to give robust and stable results.

Markov-chain Monte Carlo (MCMC) methods are suitable tools to perform PE, exploring the parameter space through random walks and collecting samples along the evolution. Subsequently, the posterior distribution is estimated using the set of independent samples of the

parameter space. Nevertheless, the nested sampling [56] is an advanced algorithm capable to extract posterior samples and perform accurate estimation of the evidence, which is the key quantity for MS. In order to solve this task, each sample is assumed to be the representative of a isoprobability contour. Then, the evidence is computed as the sum of the likelihood values weighted on the respective prior volume, estimated resorting to Bayesian calculations. The details of the sampling methods are discussed in Appendixes A and B.

III. DESIGN OF THE CODE

BAJES is a pure PYTHON software that aims to provide a versatile and robust framework for generic Bayesian inference within a simple and clear approach. In order to achieve this task, the software relies on a modular and composed architecture and it implements logically specialized objects. Furthermore, we decide to keep a light-weight setup with minimal dependencies on external libraries. These properties make BAJES a simple and general tool with a wide range of applicability. The body of the BAJES software is constituted by two components: the INF module, which represents an implementation of the Bayesian logic, and the OBS module, which contains the tools to manage and process physical data.

The INF module is the Bayesian skeleton of the software and contains the methods required to instantiate the model to be inferred. Section III A describes the tools implemented in BAJES.INF and the general workflow of the module. The Bayesian approach is constituted by three fundamental stages [57]: formulating a model, comparing the model with the data, and inferring the properties of the model. The goal of the INF module is to provide a flexible and general interface capable to adapt itself to a broad variety of problems. This structure promotes usability and applicability, supplying a comprehensive and unique architecture that allows the user to tackle specific problems. In order to enforce these concepts, the software is developed promoting an intuitive manageability and the simplicity of use: providing the necessary basic information, the user can easily set up a full Bayesian analysis. Appendix D shows an example of a practical PE analysis with BAJES tools.

In the context of data analysis, the statistical infrastructure has to be flanked by the physical characterization of the experimental data with the purpose of defining a full Bayesian model. This is necessary in order to connect the statistical properties with the actual observable quantities. Obviously, the specific physical model to be used depends on the nature of the analyzed data, on the assumptions made to build the model, and, in general, different physical events will require tailored treatments and specialized tools. To address this, BAJES provides the OBS module, which is a container of methods needed to characterize and handle specific physical observations. This module is designed aiming to the analysis of GWs and EM counterparts.

Currently, BAJES.OBS includes two submodules: GW and KN. The GW module contains the tools necessary to deal with GW analysis and it is described in Sec. IV. The KN module supplies a framework for the analysis of kilonovae light curves generated by BNS collisions, following our early work [25]. The implementation of the KN module is inspired by the approach presented in Refs. [25,58] and it will be detailed in a follow-up work.

A. Workflow

In order to fulfill the generic tasks of Bayesian inference, BAJES provides the tools necessary to instantiate a Bayesian model [59] and to extract the statistical quantities of interest, such as posterior distribution or evidence.

The BAJES.INF module supplies a PRIOR and a LIKELIHOOD objects, and their combination defines the Bayesian model. The PRIOR provides the implementation of a generic prior distribution: this object can be instantiated with list of PARAMETER, which will define the parameter space of the model. It is also possible to introduce constant and variable quantities to be carried along the sampling. Subsequently, it is possible to introduce a likelihood function customizing the specific LOG_LIKE method of a LIKELIHOOD object. Furthermore, BAJES implements JOINTPRIOR and JOINTLIKELIHOOD objects in order to handle joint posterior distributions. Additionally, planned extensions include hierarchical models (e.g., [38,60]) and Bayesian methods to deal with error propagation.

Once the Bayesian model is defined, it is possible to fit the model to the data estimating its properties. The INF module provides a SAMPLER method that wraps the provided model and initializes the requested sampling algorithm. As mentioned in Sec. IID, the sampler explores the parameter space aiming to reconstruct the posterior distribution. In order to conduct an accurate analysis, the SAMPLER objects rely on auxiliary functions such as the proposal methods described in Appendix C: these are statistical techniques that aim to extract independent samples and to conduct the sampler toward more likely regions of the parameter space (see Appendixes A and B).

When the Sampler completes the analysis, the algorithm returns the conditioned probability of interest and the properties of the model can be inferred. This information allows us to compute the statistical quantities that characterize the model and make it possible to test competing hypotheses and verify different assumptions. Moreover, from these results it is possible to understand the limits of the involved description; then, in general, the model can be improved and the workflow can be iterated with the reviewed model, improving the understanding and the modeling of the observed event. Figure 1 shows a schematic representation of the described workflow.

B. Parallelization

By default, BAJES analyses can be performed taking advantage of the parallel-threading multiprocessing PYTHON library. However, with this method, the number of available processes is strictly limited by the size of the single machine and for nontrivial problems this could be a strong limitation. For this reason, the BAJES software implements a customized method for multinodes communication based on the message passing interface (MPI) protocol.

For ideal scaling, the execution time of a machine computation is inversely proportional to the number of central processing units (CPUs) that leads to a linearly increase of the speed-up. However, in a realistic scenario, the scaling performances of sampling techniques are affected by unavoidable computational steps serially performed (e.g., temperature swapping in parallel chains and live points collection of nested sampling) and by the continuous exchange of information between different processes, required to adapt the evolution of the algorithm.

MCMC and nested sampling algorithms require separate treatments. The performances of MCMC sampling are typically quantified in terms of proposal acceptance and correlation length of the chains [7,61–65], and generally the overall execution time is determined by several contributions, such as the total number of chains, the complexity of the parameter space, and the employed proposal techniques. Estimations of MCMC execution times [66–68]

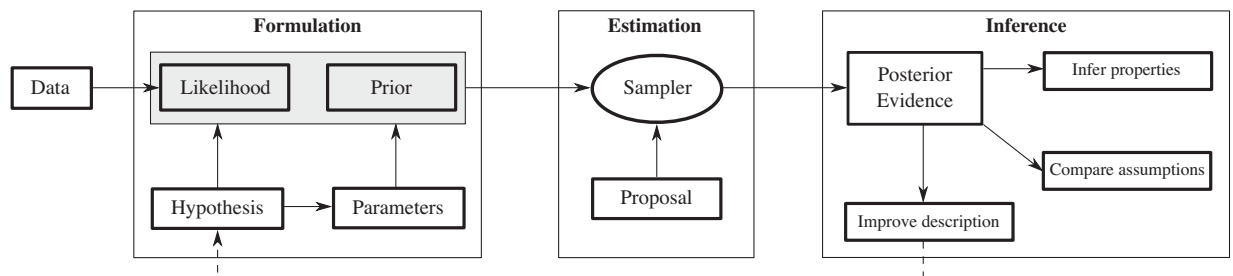


FIG. 1. Schematic representation of the workflow described in Sec. III A. The scheme highlights the three stages of the Bayesian formalism: formulation of the model, fitting the model to the data, and inference of the model properties. The gray box constituted by likelihood function and prior assumptions represents the Bayesian model to be inferred. The dashed back-propagating line refers to the case in which the analysis is iterated with an improved description.

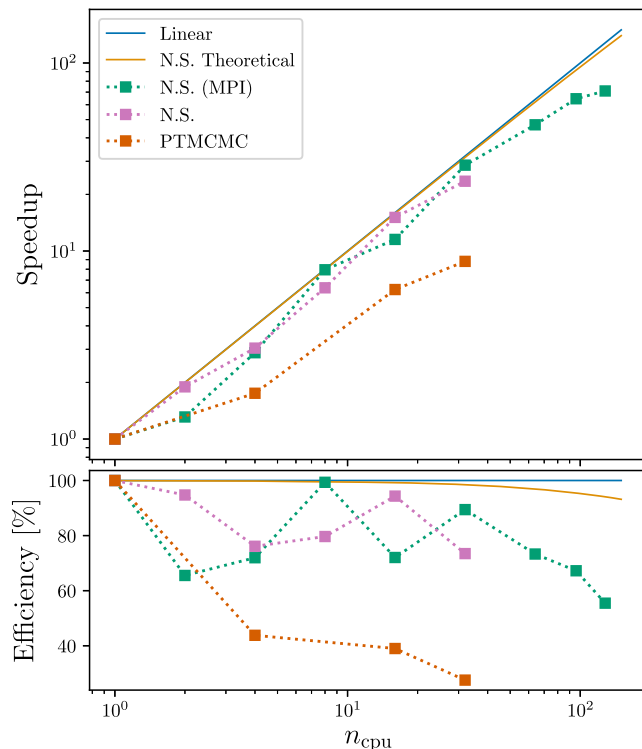


FIG. 2. Scaling plot for the BAJES pipeline (see Sec. V). Top panel: the square markers are the measured speed-up factors with respect to the serial execution time. The execution times for the PTMCMC (orange squares) algorithm are estimated performing 10^3 iterations with 128×4 tempered chains. The nested sampling execution times (pink squares for multithreading and green squares for MPI) are computed employing the dynesty software with 1024 live points and tolerance equal to 0.1. The blue solid line is the ideal linear scaling and the yellow solid line is the theoretical scaling of nested sampling. Bottom panel: same quantities discussed in the previous panel expressed in terms of efficiency.

have shown that the efficiency drastically decreases for an increasing number of parallel chains. On the other hand, the parallelization performances of the nested sampling are well studied [46,69–72] and the theoretical speed-up factor S_{NS} of this algorithm is expressed as a sublinear scaling,

$$S_{\text{NS}}(n_{\text{live}}, n_{\text{cpu}}) = n_{\text{live}} \cdot \log \left(1 + \frac{n_{\text{cpu}}}{n_{\text{live}}} \right). \quad (8)$$

For $n_{\text{live}} \gg n_{\text{cpu}}$, the values predicted by Eq. (8) are comparable with a linear trend.

Figure 2 shows the measured speed-up factors in the execution time as a function of the number of CPUs for different samplers and different parallelization methods. The speed-up factors are computed with respect to the execution time measured from the serial job. The execution times are estimated from the BAJES GW pipeline (see Sec. V) using GW150914 [73] as target signal (see

Sec. VII A). Moreover, Fig. 2 shows the respective efficiency rates that encode the deviation from the ideal linear scaling. The efficiency is computed as the percentage ratio between the measure speed-up factor and the number of employed CPUs. Ideal linear scaling has an efficiency of 100%.

For the parallel-tempered MCMC (PTMCMC) algorithm implemented in BAJES, we estimate the speed-up performing 10^3 iterations with 128×4 tempered chains, while, for nested sampling, we employ the dynesty software with 1024 live points and tolerance equal to 0.1. The PTMCMC is not optimal in terms of execution-time scaling, mainly due to the serial swapping routine. However, it gives acceptable scaling performances with efficiency $\gtrsim 40\%$ up to $n_{\text{cpu}} \simeq 16$ using multithreading methods. The results with MPI are worst compared to multithreading due to the data communication.

Regarding the nested sampling, for a very small number of processes, roughly $n_{\text{cpu}} \lesssim 2$, the multithreading method gives more efficient results, since the MPI protocol requires additional time for data communication. For an increasing number of CPUs, roughly $n_{\text{cpu}} \gtrsim 6$, the two parallelization methods give comparable results. However, the strength of MPI parallelization is the capability of accessing multiple CPUs located in different physical machines: the MPI interface implemented in BAJES gives an efficiency greater than 70% up to $n_{\text{cpu}} \simeq 100$, which is the typical size of standard PE job.

C. Dependencies

The BAJES software is compatible with PYTHON 3.7 or higher versions, and it can be easily installed using the standard PYTHON setup tools. It relies on essential modules, preferring those that can be easily installed via pip. The BAJES software mandatory requires the following libraries: NumPy [74] for algebraic computations and array managing, SciPy [75] for advanced mathematical tools, and ASTROPY [76] is invoked for astrometric computations. However, in order to execute the pipeline supplied with BAJES (see Sec. V), other libraries might be required: GWPY [77,78] is used to access the gravitational-waves open science center (GWOSC) archive [47,79,80], MATPLOTLIB [81] and CORNER [82] are employed for plotting. Moreover, if the MPI parallelization is requested, the software needs the installation of the MPI4PY library [83–85].

In order to perform the sampling, BAJES implements a PTMCMC algorithm based on PTEMCEE [65] or can use additional external packages. In particular, we interface the MCMC EMCEE [63] and the nested sampling algorithms of CPNest [86] and DYNESTY [87].

IV. GRAVITATIONAL WAVES

The BAJES.OBS.GW module contains the functions and the tools needed to deal with GW transients that are

mainly provided by signal processing and matched-filtered techniques [8].

Ground-based GW detectors measure the local perturbations of the spacetime as time series. Then, if we can believe that a physical GW transient is recorded in the data, the detector output $d(t)$ is assumed to be the sum of the noise contribution $n(t)$ and the GW strain $s(t)$, such as

$$d(t) = n(t) + s(t). \quad (9)$$

If $n(t)$ respects the conditions of Gaussianity and stationarity and if we dispose of a template $h(t)$ able to reproduce the *real signal* $s(t)$, then it is possible to filter out the noise contribution revealing the presence of a signal in the observed data. It is useful to introduce the inner product between two time-domain series, $a(t)$ and $b(t)$, as

$$(a|b) = 4\Re \int_0^\infty \frac{a^*(f)b(f)}{S_n(f)} df, \quad (10)$$

where $a(f)$ is the Fourier transform of the time series $a(t)$,

$$a(f) = \int_{-\infty}^{+\infty} a(t)e^{-2\pi ift} dt, \quad (11)$$

and analogously for $b(f)$, while $S_n(f)$ is the power spectral density (PSD) of the noise $n(t)$.

In order to perform inference on GW data, it is necessary to provide an infrastructure capable to process data segments, characterize the noise contamination, localize sources and detectors, and to generate waveform templates. In the following sections, we discuss how these tasks are achieved by the OBS.GW module.

A. Time and frequency series

A realistic portion of data measured by an interferometer is a time series with constant sampling rate F_s and finite duration T . The restriction to a finite domain leads to approximate Eqs. (10) and (11) numerically, taking the advantage of the fast Fourier transform (FFT) algorithm [88]. Within this framework, the time step $\Delta t = 1/F_s$ and the duration T of the time-domain series are reflected in the spacing $\Delta f = 1/T$ of the frequency bins and in the maximum frequency accessible from the data $f_{\text{Nyq}} = F_s/2$, known as Nyquist's frequency. Then, we can approximate Eq. (10) as

$$(a|b) \approx \frac{4}{T} \Re \sum_i \frac{a^*(f_i)b(f_i)}{S_n(f_i)}, \quad (12)$$

where $f_i = i \cdot \Delta f$ and the index i takes integer values from 0 to $F_s T/2$. Generally, this summation is performed on a restricted frequency band, identified by a lower and an upper cutoff frequencies, respectively, f_{min} and f_{max} , in order to neglect irrelevant portion of the spectrum. From Eq. (10), or

its approximation Eq. (12), it is possible to introduce the signal-to-noise ratio (SNR) as

$$\rho = \frac{(d|h)}{\sqrt{(h|h)}}, \quad (13)$$

which estimates the power of the signal $h(t)$ enfolded in the data $d(t)$ weighted on the variance of the background noise.

BAJES implements a SERIES object designed to manage time-domain and frequency-domain series. This instance stores the input series and it computes the FFT (or the inverse FFT) of the given data in order to provide both representations of the series. The SERIES object supplies also an interface capable to perform tapering, filtering, and padding of the input series: we make use of the Tukey window for the tapering, while the filtering is performed using a Butterworth filter. Furthermore, the SERIES object implements a summation and a product between objects of the same type, defined in the frequency domain, and contains methods to compute inner products and SNRs.

B. Noise characterization

As shown in Eq. (9), the measured data $d(t)$ are intrinsically related with the noise fluctuations $n(t)$. The noise of a GW detector is represented by stochastic fluctuations [8] that propagate to the output. The primary noise sources in a ground-based interferometer are gravity gradients and seismic motions [89–91], thermal agitation [92,93], quantum perturbations [94,95], and internal optomechanical resonances [96,97]. Moreover, the time series recorded by a GW detector are also affected by external nongravitational signals [98], such as the ac power grid transmission responsible for the 60 Hz peak of LIGOs and the 50 Hz one of Virgo. The noise fluctuations are assumed to be Gaussian distributed and stationary on a relatively large timescale [8]. Then, the PSD shows the distribution of the noise power for every frequency component and it can be computed as

$$E[|n(f)|^2] = \frac{T}{2} S_n(f), \quad (14)$$

where the expectation is computed with the temporal average over a period T . In other words, the PSD characterizes the uncertainty of the observed data in frequency domain as the variance of the associated noise spectrum. The amplitude spectral density (ASD) is usually defined as the square root of the PSD, $\sqrt{S_n(f)}$.

The PSD is a key quantity in order to estimate the product Eq. (10), since it describes the weight of each frequency component. It follows that a full characterization of the noise sources and a proper estimation of its contribution are essential in order to perform accurate measurements of GW transients. A general tool for estimating PSD is the Welch's method [99], consisting

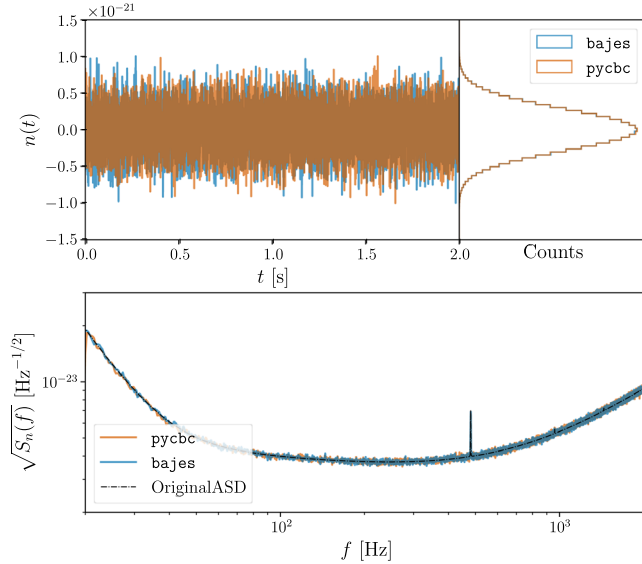


FIG. 3. Top left panel: comparison of artificial noise segment produced using BAJES (blue) and PyCBC (orange). The segments are computed using LIGO design sensitivity P1200087 [105] with a lower cutoff frequency at 20 Hz and a sampling rate of 4096 Hz. The panel shows a 2 s segment extracted from a segment with total duration of $T = 1024$ s. Top right panel: histograms of the time-domain samples computed using the whole artificial segment of length $T = 1024$ s. Bottom panel: ASDs reconstructed from the artificial noise segments using the Welch’s method. The spectra are computed using the whole segments with a chunk length of 4 s and an overlap fraction of 99%. The black dashed line represents the original ASD.

in the average of the spectral densities computed on chunks of the full noise segment. However, this is not the only technique suitable for this task [100,101].

Aiming to data analysis purposes, BAJES implements a NOISE object. This component is capable to estimate the PSD of a given noise time-domain series using the Welch’s method, generate an artificial segment of stationary and Gaussian noise from a given PSD, and it disposes of methods for PSD interpolation. Figure 3 shows a comparison of artificial noise segments generated with BAJES and with the PyCBC routines [32,102–104], where the total length of the artificial segment is 1024 s. The histograms and the PSDs show that the generated noise fluctuations respect the conditions of Gaussianity and stationarity with a frequency spectrum described by the requested PSD.

C. Ground-based interferometers

The current ground-based observatories for GW detection are L-shaped Michelson interferometers with Fabry-Perot cavities [106,107]. Each arm has a length of $L \approx 3\text{--}4$ km, depending on the actual detector [27,28], and it is constituted of two mirrors acting like test masses. The detector performs measurements of the gravitational strain as a difference in length of the orthogonal arms,

$$d(t) = \frac{\Delta L(t)}{L}, \quad (15)$$

where $\Delta L = \Delta L_x - \Delta L_y$ is the difference between the displacements along the two orthogonal arms. The projection of the signal on the detector can be computed from the GW polarization components $h_{+,\times}$ as

$$h(t) = D_{ij}h_{ij}(t) = F_+h_+(t) + F_\times h_\times(t), \quad (16)$$

where D_{ij} is labeled as detector tensor and it depends on the geometry of the interferometer, while $F_{+,\times}$ are the *antenna pattern* functions for each polarization. The antenna patterns $F_{+,\times}$ characterize the detector sensitivity in the different regions of the sky for every given time [108].

Generally, in standard observing conditions, the observations of GW transients are performed simultaneously by a worldwide network of ground-based interferometers. Thanks to this, it is possible to correlate strains observed independently in different locations, improving the estimation of the astrophysical origin of the transients [109,110].

The necessity to localize a GW observatory in a fixed frame arises in order to project the expected signal on the detector and to estimate the light travel time from each detector in the network. For these tasks, BAJES disposes of a DETECTOR class able to characterize a ground-based interferometer. This object is identified with the coordinates of the site of the interferometer (latitude, longitude, and altitude) and the angles of the two arms (azimuth and tilt). It is also possible to initialize the DETECTOR object to precomputed detector configurations using the two-digits string identifying a ground-based interferometer, e.g., H1 for LIGO-Hanford [27], L1 for LIGO-Livingston [27], V1 for Virgo [28], G1 for GEO600 [106,111], K1 for KAGRA [112,113], and ET for Einstein Telescope (configuration D) [114,115].

D. Waveform templates

The last ingredient necessary to complete the framework is a waveform template, i.e., a model able to compute the expected representation of the signal $h(t, \theta)$ (or $h(f, \theta)$) for every given set of parameters θ . The WAVEFORM object is a class that accesses the methods disposable in the BAJES.OBS.GW.APPROX submodule and computes the expected GW polarization components for every given set of parameters. The frequency-domain waveform in Eq. (16) can be written in terms of the amplitude $A(f)$ and the phase $\phi(f)$,

$$h(f) = A(f)e^{-i[\phi_0 + 2\pi t_0 f + \phi(f)]}, \quad (17)$$

where ϕ_0 and t_0 are, respectively, phase and time references.

BAJES directly implements and interfaces with all the most advanced GW templates for quasicircular compact

binary mergers, and includes state-of-art templates for eccentric and hyperbolic binary black hole (BBH) mergers [116,117] as well as for BNS postmerger [118]. In particular:

- (i) TaylorF2: The post-Newtonian (PN) framework [119–124] represents a milestone for the description of the two-body problem. This approach solves the dynamical evolution of a compact binary with a perturbative method assuming low velocities and weak field approximations, which are reflected in the condition $v = (\pi G M f)^{1/3} \ll c$, where v is the characteristic velocity in the binary, $M = m_1 + m_2$ is the total mass, and f is the GW frequency. The exact analytic solution of the gravitational radiation emitted by a point-particle compact binary is known up to the 3.5PN order. BAJES also implements a pseudo-5.5PN accurate description of the point mass baseline, as derived in Ref. [125]. Pure tidal corrections are implemented up to 7.5PN [126,127] and include the recently computed tail terms (6.5, 7.5 PN) of [128], whereas spin-quadrupole terms are included up to 3.5PN [129].
- (ii) TEOBResumS: The effective-one-body (EOB) formalism [130–138] is a Hamiltonian approach that reduces the two-body problem to a single orbiting object in an effective potential. TEOBResumS [139] is an EOB approximant for spinning coalescing compact binaries [140,141] (black holes or neutron stars) inspiralling along generic orbits. It includes tidal effects [142–144] and subdominant modes [145,146] up to $\ell = 5$, as well as a description of precession-induced modulations up to merger [147]. When considering systems evolving along quasicircular orbits, the computational performance of the model is enhanced by means of a postadiabatic description of the early inspiral [148]. Moreover, the model includes the implementations of hyperbolic encounters [117], eccentric mergers [116], and a frequency-domain approximation valid for quasicircular BNS coalescences, labeled as TEOBResumSPA [149]. The model is publicly available at [150] and used in BAJES from the provided PYTHON interface.
- (iii) NRPM: BNS merger remnants are expected to be loud sources of GWs [151–153] that convey unique information on the equation of state of hot matter at extreme densities [12,154,155]. NRPM [118] is an analytical model based on numerical relativity BNS simulations. The model is tuned on a set of simulations covering the range $q \leq 1.5$ and $\Lambda_{1,2} \gtrsim 400$. For smaller values of the tidal parameters, the model is identically zero since it is not expected to have a postmerger signal in these cases, due to prompt black-hole formation [156–158]. NRPM is directly implemented in BAJES and it can be attached to the TEOBResumS premerger template, obtaining a

complete model for the gravitational radiation expected from BNS coalescences.

- (iv) MLGW: Machine learning tools can be employed to construct accurate representations of GW signals. The MLGW package [159,160] takes advantage of these methods to generate fast and reliable GW templates for BBH coalescences. The model is composed by contributions extracted with principal component analysis and a linear combination of regression models, which maps the orbital parameters of the black holes to the reduced representation of the wave. A complete model includes two principal component models, for both phase and amplitude of the wave, and a mixture of regression models for each of the principal components considered. The algorithm is trained on time-domain models and tested only for aligned-spin BBHs. Currently, the released software provides the representations of EOB templates, TEOBResumS [139] and SEOBNRv4 [161].
- (v) GWSURROGATE: The templates provided by the GWSURROGATE package [162] implements fast waveforms based on reduced-order models [163] trained on numerical relativity simulations. The NRSur7DQ4 model [164] is a precessing extension of the model presented in Ref. [165] trained on a set of simulations with $q \leq 4$ and $\chi_1, \chi_2 \leq 0.8$ that contains all higher-order modes with $\ell \leq 4$. On the other hand, NRHYBSUR3DQ8 [166] and its tidal version [167] are calibrated using hybridized waveforms in order to increase the number of orbits of the training templates, improving the quality of the approximation. This model is tuned in a wider range in the mass ratio, $q \leq 8$, but it does not include precession contributions.
- (vi) LALSsimulation: The LIGO Algorithm Library LALSuite [26] is the official LIGO-Virgo Collaboration software and it provides the largest variety of waveform template models. BAJES implements the waveform generator of LALSsimulation, a module of LALSuite. For the results of this paper, we make use of this implementation in order to employ IMRPhenomPv2 approximant [168,169] and its tidal extension, IMRPhenomPv2_NRTidal [170]. A list of all the approximants available through LALSsimulation can be found at [171].

V. PIPELINE

BAJES provides a customized and automatized pipeline for the analysis of GW transients and EM counterparts. In this section, we discuss the model implemented to perform PE analysis on GW transients with BAJES.

In the context of GW data analysis, we introduce the working hypotheses that are going to define the employed Bayesian model. We call the assumption that the data

contain a nonvanishing GW transient *signal hypothesis* H_S , i.e., $d(t) = n(t) + s(t)$ with $s(t) \neq 0$. On the other hand, the *noise hypothesis* H_N is the condition for which the recorded data correspond only to pure noise realization, i.e., $d(t) = n(t)$. Then, in the signal hypothesis condition, a GW signal emitted by a quasicircular compact binary coalescence (CBC) can be fully characterized with a set of 17 parameters according to general relativity. Precisely,

$$\theta_{\text{cbc}} = \{m_1, m_2, \chi_1, \chi_2, \Lambda_1, \Lambda_2, D_L, \iota, \alpha, \delta, \psi, t_0, \phi_0\}, \quad (18)$$

where

- (i) $m_{1,2}$ refer to the detector-frame masses of the two objects, taken with the convention $m_1 \geq m_2$.
- (ii) $\chi_{1,2}$ are the dimensionless spin vectors,

$$\chi_i = \frac{c\mathbf{S}_i}{Gm_i^2}, \quad i = 1, 2, \quad (19)$$

where $\mathbf{S}_{1,2}$ are the spins of the two objects, c is the speed of light, and G is the gravitational constant.

- (iii) $\Lambda_{1,2}$ are the dimensionless tidal polarizability parameters that encode the quadrupolar deformability of the i th object under the effect of an external force,

$$\Lambda_i = \frac{2}{3}k_{2,i} \left(\frac{c^2 R_i}{Gm_i} \right)^5, \quad i = 1, 2, \quad (20)$$

where $k_{2,i}$ and R_i are, respectively, the second Love number and the radius of the i th object (k_2 is identically zero for black holes).

- (iv) D_L is the luminosity distance of the source from the observer.
- (v) ι is the angle between the line of sight and the total angular momentum of the system and it takes value in the range $[0, \pi]$.
- (vi) $\{\alpha, \delta\}$ are, respectively, right ascension and declination angles that identify the sky position of the source.
- (vii) ψ is the polarization angle and it takes value in the range $[0, \pi]$.
- (viii) $\{t_0, \phi_0\}$ are, respectively, reference time and reference phase.

The sampling is performed promoting the chirp mass \mathcal{M} and the mass ratio q ,

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \quad q = \frac{m_1}{m_2} \geq 1, \quad (21)$$

over the single mass components, since they appear to be less correlated for this type of signals [6,7]. For spinning binary mergers, we define the effective spin parameter χ_{eff} as

$$\chi_{\text{eff}} = \frac{m_1 \chi_{1,z} + m_2 \chi_{2,z}}{m_1 + m_2} \quad (22)$$

that encodes the aligned-spin contribution and it can lead to narrower uncertainties than the single spin components [172]. Furthermore, in the context of BNS mergers, it is useful to introduce the reduced tidal parameter $\tilde{\Lambda}$,

$$\tilde{\Lambda} = \frac{16}{13} \left[\frac{(m_1 + 12m_2)m_1^4 \Lambda_1}{M^5} + (1 \leftrightarrow 2) \right], \quad (23)$$

and the asymmetric tidal parameter $\delta\tilde{\Lambda}$,

$$\delta\tilde{\Lambda} = \left[1 - \frac{7996 m_2}{1319 m_1} - \frac{11005}{1319} \left(\frac{m_2}{m_1} \right)^2 \right] \frac{m_1^6 \Lambda_1}{M^6} - (1 \leftrightarrow 2), \quad (24)$$

where $M = m_1 + m_2$. The tidal parameters $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ are, respectively, proportional to the leading order and the next-to-leading-order tidal contributions according to PN expansion.

Generally, concerning GW analysis, it is common to label $\theta_{\text{int}} = \{m_1, m_2, \chi_1, \chi_2, \Lambda_1, \Lambda_2\}$ as intrinsic parameters, since they affect the physical dynamics of the system; while, the extrinsic parameters $\theta_{\text{ext}} = \{D_L, \iota, \alpha, \delta, \psi, t_0, \phi_0\}$ are related with the observed signal by trivial geometrical argumentation. Moreover, it is possible to include in the pipeline further parameters in order to take into account the calibration of the input quantities, such as calibration envelopes and PSD uncertainties. For a detailed discussion about these topics, see Sec. V D.

In the following sections, we present the implementations available in the BAJES GW pipeline.

A. Data segments

The default GW routine implemented in BAJES is designed for binary mergers analyses. In general, the pipeline is able to access to the open-source database of GWOSC [47,79,80] to all the data released with gravitational-wave transients catalog-1 (GWTC-1) [1] and to the expected PSD curves for current and next-generation detectors [114,115].

The input data to be analyzed by the pipeline can be gathered in different ways. The GW pipeline provides a customized *injection* generator capable to produce artificial data given a prescribed set of parameters and the detector sensitivity curves. Moreover, the BAJES pipeline allows to access the observational data recorded by the LIGO-Virgo interferometers [27,28,105] from the GWOSC, specifying the central value of the GPS time and the duration of the segment.

When the data information is gathered, the pipeline initializes the LIKELIHOOD function and the PRIOR with the requested parameters belonging to the set θ_{cbc} , and it passes

these arguments to the requested sampler which performs the Bayesian inference.

B. Prior distributions

The prior distribution for the masses is chosen flat in the components $\{m_1, m_2\}$ that can be written in terms of $\{\mathcal{M}, q\}$ as

$$p(\mathcal{M}, q|H_S) = \frac{\mathcal{M}}{\Pi_{\mathcal{M}}\Pi_q} \left(\frac{1+q}{q^3} \right)^{2/5}, \quad (25)$$

where $\Pi_{\mathcal{M}}$ and Π_q are the prior volumes limited by the bounds $[\mathcal{M}_{\min}, \mathcal{M}_{\max}]$ in chirp mass and $[1, q_{\max}]$ in mass ratio,

$$\Pi_{\mathcal{M}} = \frac{1}{2} (\mathcal{M}_{\max}^2 - \mathcal{M}_{\min}^2),$$

$$\Pi_q = 5 \left[\frac{2^{2/5} \sqrt{\pi} \Gamma(\frac{4}{5})}{\Gamma(\frac{3}{10})} - \frac{{}_2F_1(-\frac{2}{5}, -\frac{1}{5}, \frac{4}{5}, -q_{\max})}{q_{\max}^{1/5}} \right], \quad (26)$$

where ${}_2F_1(a, b, c; z)$ is the hypergeometric function and $\Gamma(x)$ is the Euler function.

The spin vectors can be written in the polar frame of the binary as $\chi_i = \{\chi_i, \vartheta_i, \varphi_i\}$, where χ_i is the spin magnitude, ϑ_i is the tilt angle, and φ_i is the complimentary azimuthal angle between χ_i and the orbital angular momentum \mathbf{L} of the binary. The prior distribution for these quantities is specified by the maximum value of spin magnitude $\chi_{\max} \geq 0$ and it can be chosen between the following:

- (i) *Isotropic prior with precessing spins*: The prior on the angular components $\{\vartheta_i, \varphi_i\}$ is isotropic over the solid angle, while the spin magnitude is uniformly distributed in the range $[0, \chi_{\max}]$,

$$p(\chi_i, \vartheta_i, \varphi_i|H_S) = \frac{\sin \vartheta_i}{4\pi\chi_{\max}}. \quad (27)$$

- (ii) *Isotropic prior with aligned spins*: This case is identical to the isotropic one except for the assumption of aligned spins, $\vartheta_i = 0, \pi$. The xy components of the spin vectors are marginalized, obtaining the form [30]

$$p(\chi_{i,z}|H_S) = \frac{1}{2\chi_{\max}} \log \left| \frac{\chi_{\max}}{\chi_{i,z}} \right|. \quad (28)$$

- (iii) *Volumetric prior with precessing spins*: The distribution is taken uniform in all Cartesian components, i.e., flat over the sphere with radius χ_{\max} . This prior can be written as

$$p(\chi_i, \vartheta_i, \varphi_i|H_S) = \frac{3}{4\pi} \frac{\chi_i^2 \sin \vartheta_i}{\chi_{\max}^3}. \quad (29)$$

- (iv) *Volumetric prior with aligned spins*: The same of volumetric case with aligned components; the marginalization over the xy components leads to the form

$$p(\chi_{i,z}|H_S) = \frac{9}{16\pi} \frac{\chi_{\max}^2 - \chi_{i,z}^2}{\chi_{\max}^3}. \quad (30)$$

The prior distribution for the sky position parameters $\{\alpha, \delta\}$ is taken isotropic over the entire solid angle, such that $\alpha \in [0, 2\pi]$ and $\delta \in [-\pi/2, +\pi/2]$,

$$p(\alpha, \delta|H_S) = \frac{\cos \delta}{4\pi}, \quad (31)$$

and analogously for the inclination ι in the range $[0, \pi]$,

$$p(\iota|H_S) = \frac{\sin \iota}{2}. \quad (32)$$

Regarding the luminosity distance, the bounds are specified by the lower and the upper bounds $[D_{\min}, D_{\max}]$, and the analytic form of the prior can be chosen between the following:

- (i) *Volumetric prior*: General analysis assumes that the source is uniformly distributed over the sphere centered around the detectors, then

$$p(D_L|H_S) = \frac{3D_L^2}{D_{\max}^3 - D_{\min}^3}. \quad (33)$$

- (ii) *Comoving-volumetric prior*: In order to take into account the cosmological expansion of the Universe, a prior uniformly distributed over the comoving volume V_C is a more suitable physically motivated choice. Within this assumption, the prior on the luminosity distance can be written as

$$p(D_L|H_S) \propto \frac{dV_C}{dD_L}. \quad (34)$$

The luminosity distance D_L and the comoving volume V_C are related through the redshift z for a fixed cosmological model; by default, BAJES acquires the values of the cosmological constants from Ref. [173].

- (iii) *Source-frame prior*: As shown in Refs. [4,14], Eq. (34) does not take into account contributions due to time dilatation. Then, we can introduce a prior distribution uniformly distributed in the source-frame volume as

$$p(D_L|H_S) \propto \frac{1}{1+z} \frac{dV_C}{dD_L}, \quad (35)$$

where the factor $(1+z)^{-1}$ converts the merger rate from the source frame to the detector frame.

- (iv) *Log-uniform prior*: It could be useful to introduce a uniform prior in $\log D_L$, then

$$p(D_L|H_S) = \frac{D_L^{-1}}{\log(D_{\max}/D_{\min})}. \quad (36)$$

For the remaining parameters, i.e., $\{\psi, t_0, \phi_0\}$ and $\{\Lambda_1, \Lambda_2\}$ (if required), their prior distributions are taken uniform within the provided bounds. Then, the overall prior is the product of the priors of the single parameters.

C. Likelihood function

The key ingredient of the inference is the likelihood function that encodes the capability of a given model $h(t, \boldsymbol{\theta}_{\text{cbc}})$ to match the observed data $d(t)$. For Gaussian and stationary noise $n(t)$, we expect the mean of the noise fluctuations to be centered around zero with a variance described by the PSD in the frequency domain, i.e.,

$$p(n|H_N) \propto e^{-\frac{1}{2}(n|n)}. \quad (37)$$

It follows that within the signal hypothesis H_S and supposing that we dispose of a template $h(t, \boldsymbol{\theta}_{\text{cbc}})$ capable to reproduce the *real signal* $s(t)$ for a given set of $\boldsymbol{\theta}_{\text{cbc}}$, the log-likelihood function can be written as the frequency-domain residuals between the recorded data and the template with the product defined in Eq. (10),

$$p(d|\boldsymbol{\theta}_{\text{cbc}}, H_S) = \frac{1}{\mathcal{N}} e^{-\frac{1}{2}(d-h|d-h)}, \quad (38)$$

where \mathcal{N} is the normalization constant that can be expressed in terms of the PSD using the numerical approximation Eq. (12),

$$\mathcal{N} \approx \prod_i \sqrt{\frac{\pi T S_n(f_i)}{2}}. \quad (39)$$

Then, the Bayes' factor of the signal hypothesis against the noise assumption can be computed as

$$\mathcal{B}_N^S = \frac{p(d|H_S)}{p(d|H_N)} = \int_{\boldsymbol{\theta}} \exp \left[(d|h(\boldsymbol{\theta}_{\text{cbc}})) - \frac{1}{2}(h(\boldsymbol{\theta}_{\text{cbc}})|h(\boldsymbol{\theta}_{\text{cbc}})) \right] \times p(\boldsymbol{\theta}_{\text{cbc}}|H_S) d\boldsymbol{\theta}_{\text{cbc}}. \quad (40)$$

Combining Eq. (38) with Eq. (17), it is possible to write the explicit dependency of the likelihood with respect to the reference parameters $\{\phi_0, t_0\}$; since these values have no physical relevance, we marginalize the posterior distribution with respect to $\{\phi_0, t_0\}$ in order to increase the efficiency of the sampling. The marginalization over ϕ_0

can be computed analytically and the solution can be written in terms of the modified Bessel function of the first kind [174]. For the time-shift t_0 , the computation is semianalytical since the values of the likelihood are evaluated on an equally spaced grid resorting on the FFT computation [175].

D. Additional implementations

In order to perform accurate and reliable inferences of GW transients, the pipeline requires refinements and auxiliary control systems. In this section, we discuss some of the additional tools implemented in the GW pipeline supplied with BAJES.

1. Calibration envelopes

The necessity of calibration envelopes [176] arises due to imperfect knowledge of the interferometer response to differential arm length changes [177–179], which affects the transfer functions of the detector components introducing systematic errors that propagate to the recorded data. These uncertainties are estimated by inspecting the detector control systems and propagating the measurements into a frequency-dependent probability distribution. Subsequently, the information on calibration errors must be taken into account when inferring the astrophysical parameters of GW signals. In order to achieve this task, it is useful to introduce two auxiliary functions $\delta A(f)$ and $\delta\phi(f)$ that characterize, respectively, the amplitude and the phase uncertainties of the measured data segments. Then, the calibration envelopes $\{\delta A(f), \delta\phi(f)\}$ can be taken into account in the Bayesian model as

$$h(f) \rightarrow [1 + \delta A(f)] e^{i\delta\phi(f)} h(f). \quad (41)$$

This procedure is accomplished specifying the values of the calibration envelopes at predefined logarithmic-spaced frequency nodes f_j and linearly interpolated over the interested frequency axis. The calibration parameters $\{\delta A(f_j), \delta\phi(f_j)\}$ are introduced in the sampling and estimated along with the signal parameters. The prior for calibration envelopes $\{\delta A(f_j), \delta\phi(f_j)\}$ is a multivariate normal distribution with variance specified by the measured calibration errors.

2. PSD uncertainties

The usage of a fixed estimated PSD might generate biases due to nonstationary effects and unaccounted slow variations in the noise spectrum. Then, it arises the necessity to take into account the uncertainty of the PSD estimate during the inference of the properties of a GW signal. For this reason, the pipeline allows the possibility to include PSD uncertainty weights η_j [7]: the Fourier domain is divided in predefined logarithmic-spaced bins $[f_j, f_{j+1}]$ and the weights are included such that

$$S_n(f) \rightarrow \eta_j S_n(f), \quad \text{for } f_j \leq f < f_{j+1}, \quad (42)$$

where η_j is taken constant in respective frequency bin. The full set of $\{\eta_j\}$ parameters, one for every frequency bin, is introduced in the sampling and they are estimated during the exploration with the signal parameters. The prior distribution for the PSD uncertainty η_j is taken normal with mean zero and variance $1/N_j$, where N_j is the number of data sample enclosed in the bin $[f_j, f_{j+1}]$. This scheme has shown to improve the robustness of the GW inference [180], and it offers a flexible model capable to quantify the differences between the estimated PSD and the spectrum of the analyzed data.

VI. INJECTION STUDIES

In this section, we show the results coming from a set of *injection* studies performed with the BAJES pipeline in order to test the sampling routines with the GW infrastructure. An injection is a simulated GW signal that has been added into a time-domain segment. For our studies, we generate artificial noise segments according with a prescribed PSD assuming Gaussian and stationary fluctuations, as discussed in Sec. IV B. Subsequently, the artificial signal $h(t)$ is simulated according with the input parameters θ_{cbc} , projected on the requested detectors and included in the data segment. Finally, the data are analyzed by the BAJES pipeline resorting to the framework described above. The sensitivity curves employed for these studies correspond to the noise spectra expected at design sensitivity for current ground-based detector [27,28,105,107,112,181] and for next-generation interferometers [114,115]. The properties of the injected signals are described in the following paragraphs depending on the particular kind of source.

A. Binary black holes

The first example corresponds to an aligned-spinning BBH coalescence with intrinsic parameters $\{\mathcal{M} = 30 M_\odot, q = 2, \chi_{1,z} = 0.3, \chi_{2,z} = 0.\}$ located at a luminosity distance $D_L = 3$ Gpc with inclination angle $i = 0$. The signal is injected such that the merger occurs at GPS time 1126259462.0 with a sampling rate of 4096 Hz and a signal length of 16 s. The data are analyzed using two detectors, H1 and L1, with LIGO design sensitivity P1200087 [27,105,181]. The sky location of the source corresponds to the position of maximum sensitivity for the detector H1, $\{\alpha = 0.372, \delta = 0.811\}$. The injected signal is generated with TEOBResumS waveform model (employing only the dominant mode) with a network SNR of 14, corresponding to 11 in H1 and 9 in L1.

The recovery of the posterior distribution is performed with PTMCMC sampling with 8 tempered ensembles and 128 chains per ensemble. Moreover, we requested 8×10^3 burn-in iterations. The injected strain is analyzed in the frequency domain from 20 Hz to 1 kHz employing three

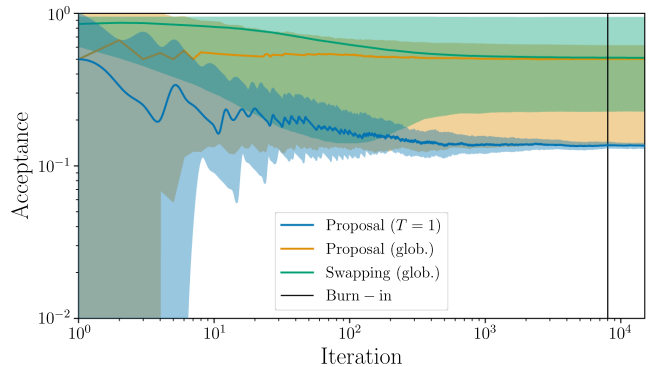


FIG. 4. Sampler acceptances during the BBH injection study performed with TEOBResumS described in Sec. VI A. The blue line is the proposal acceptance of the untempered ensemble averaged over the chains, the yellow and the green lines are, respectively, the proposal and the swapping acceptances of whole sampler averaged over all the tempered ensembles. The solid lines represent the median values and the shadowed areas are the 90% credible regions. The vertical black line is the requested last burn-in iteration.

different templates: TEOBResumS, IMRPhenomPv2, and NRSur7DQ4. The likelihood is marginalized over reference time and phase. We set the chirp mass prior in $[23, 37] M_\odot$ and the mass ratio in $[1, 8]$. The spins are kept aligned with an isotropic prior in the range $[-0.9, +0.9]$ for every component. We employ volumetric prior for the luminosity distance in the support $[100 \text{ Mpc}, 5 \text{ Gpc}]$ and the prior distributions for the remaining parameters are chosen according with Sec. V B.

Figure 4 shows the acceptance fractions for the analysis performed with TEOBResumS waveform model. The inferences performed with other approximants show similar behaviors. The untempered ensemble required less than 10^4 iterations to converge to the maximum-posterior value. After the requested burn-in, the untempered ensemble shown an average acceptance of 15% and, averaging over all the tempered ensembles, the sampler advanced with a global proposal acceptance of $\sim 45\%$ and with a global swapping acceptance of $\sim 50\%$. The final autocorrelation length (ACL) of the untempered ensemble corresponds to a lag of 70 iterations and the sampler collected a final amount of 1.5×10^4 independent posterior samples.

Table I shows the recovered mean values, and Fig. 5 shows the recovered marginalized posterior distribution for some exemplary parameters. The marginalized posterior distributions enclose the injected values within 90% credible intervals for all the waveform approximants. The estimated evidences slightly prefer TEOBResumS waveform, accordingly with the injected template. However, these values lie in the same range for all the analyses, leading to a not fully resolved model selection. This is due to the large uncertainties associated to the evidence estimation of the PTMCMC and with relatively low SNR of the injected signal. For the latter reason, it is also not possible to reveal

TABLE I. Recovered parameters during the BBH injections studies. The signal has been injected in H1 + L1 using design sensitivity curves, with an overall network SNR of 14. The data have been analyzed with the BAJES PTMCMC sampling. The reported values correspond to the medians with the 90% credible regions. The last column reports the estimated logarithmic Bayes' factor and the associated standard deviation.

| Approximant | \mathcal{M} [M_{\odot}] | q | $\chi_{1,z}$ | $\chi_{2,z}$ | χ_{eff} | D_L [Gpc] | ι [rad] | α [rad] | δ [rad] | $\log \mathcal{B}_N^S$ |
|--------------|-------------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|-------------------------|
| Injected | 30.0 | 2.0 | 0.3 | 0.0 | 0.2 | 3.0 | 0.0 | 0.372 | 0.811 | ... |
| TEOBResumS | $30.63^{+1.84}_{-1.64}$ | $1.56^{+0.78}_{-0.49}$ | $0.24^{+0.36}_{-0.41}$ | $0.17^{+0.59}_{-0.63}$ | $0.23^{+0.14}_{-0.15}$ | $2.49^{+1.29}_{-1.17}$ | $0.82^{+2.05}_{-0.63}$ | $2.17^{+4.05}_{-2.13}$ | $0.21^{+1.10}_{-1.09}$ | $73.29^{+2.64}_{-2.64}$ |
| IMRPhenomPv2 | $30.34^{+1.77}_{-1.69}$ | $1.50^{+0.83}_{-0.46}$ | $0.17^{+0.43}_{-0.43}$ | $0.17^{+0.57}_{-0.60}$ | $0.19^{+0.14}_{-0.16}$ | $2.47^{+1.30}_{-1.22}$ | $0.80^{+2.06}_{-0.62}$ | $2.34^{+3.88}_{-2.31}$ | $0.17^{+1.13}_{-1.05}$ | $73.16^{+2.66}_{-2.66}$ |
| NRSur7DQ4 | $30.35^{+1.75}_{-1.60}$ | $1.56^{+0.82}_{-0.53}$ | $0.18^{+0.37}_{-0.39}$ | $0.17^{+0.59}_{-0.63}$ | $0.19^{+0.56}_{-0.63}$ | $2.49^{+1.25}_{-1.23}$ | $0.84^{+2.02}_{-0.64}$ | $2.11^{+4.12}_{-2.07}$ | $0.21^{+1.10}_{-1.09}$ | $73.11^{+2.72}_{-2.72}$ |

systematic differences between the different approximants [182], and the results of the employed templates are largely consistent between each other.

B. Binary neutron stars inspiral

In this section, we analyze an inspiralling nonspinning BNS merger with intrinsic parameters are $\{\mathcal{M} = 1.188 M_{\odot}, q = 1, \Lambda_1 = \Lambda_2 = 600\}$. The source is located at a luminosity distance $D_L = 120$ Mpc with inclination angle $\iota = 0$. The signal is injected such that the merger occurs at GPS time 1126259462.0 with a sampling rate of 4096 Hz and a signal length of 128 s. The data are analyzed using two detectors, H1 and L1, with LIGO design sensitivity P1200087 [27,105,181]. The sky location of the source corresponds

to the position of maximum sensitivity for the detector H1, $\{\alpha = 0.372, \delta = 0.811\}$. The injected signal is generated with TEOBResumS waveform model with a network SNR of 20, corresponding to 15 in H1 and 13 in L1.

The recovery of the posterior distribution is performed employing nested sampling algorithm with 1024 live point and tolerance equal to 0.1. Furthermore, we set, respectively, the minimum and the maximum number of iterations for every MCMC subchain to 32 and 4×10^3 . The injected strain is analyzed in the frequency domain from 20 Hz to 1 kHz employing three different templates: TEOBResumSPA, IMRPhenomPv2_NRTidal, and TaylorF2 including 5.5PN point-mass corrections and 6PN tidal contributions. We set the chirp mass prior in $[0.8, 2] M_{\odot}$ and the mass ratio in $[1, 4]$. The spins are kept aligned with an isotropic prior in the

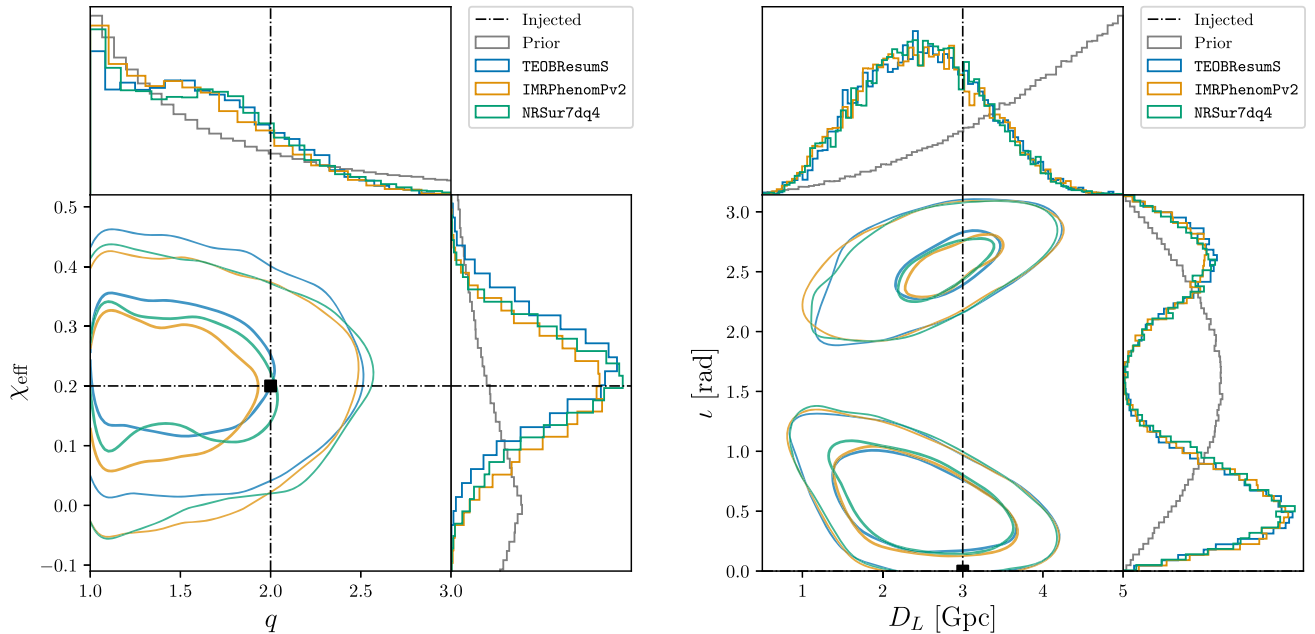


FIG. 5. Posterior distributions for $\{q, \chi_{\text{eff}}\}$ and $\{D_L, \iota\}$ recovered from the injection studies performed on a BBH signal with two interferometers (H1 + L1) at design sensitivities with network SNR of 14. The artificial signal has been generated with TEOBResumS model and the injected parameters are marked with black lines and squares. The contours represent the 50% (thick) and the 90% (thin) credible regions. The recovery has been performed with four different approximants analyzing the frequency range from 20 Hz to 1 kHz. The estimation of the luminosity distance is affected by the degeneracy with the inclination angle [183], due to the correlations in the strain amplitude for aligned-spin sources.

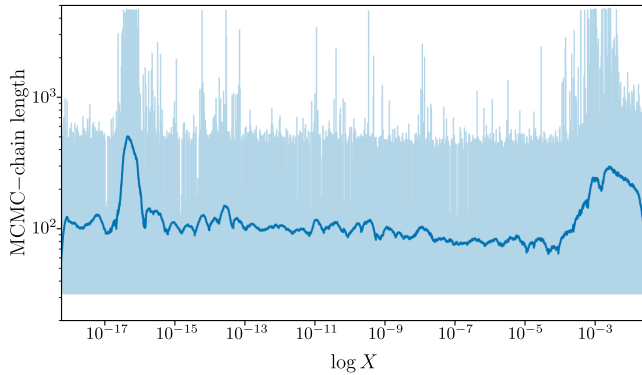


FIG. 6. Length of the MCMC subchains during the nested sampling performed with TEOBResumSPA described in Sec. VI B. The actual values extracted from the sampler have been smoothed with a Savitzky-Golay filter for visualization purposes. The shadowed region shows the difference between the actual lengths and the filtered estimation. The values are lower bounded by the requested minimum value, set equal to 32.

range $[-0.9, +0.9]$ for every component. The tidal parameters are extracted in the range $[0, 5000]$. We employ volumetric prior for the luminosity distance in the support $[10, 400]$ Mpc and the prior distributions for the remaining parameters are chosen according with Sec. V B.

Figure 6 shows the number of iterations of the MCMC subchains employed to propose a new point as a function of the prior mass X (see Appendix B for the definition). The actual values extracted from the sampler have been smoothed with a Savitzky-Golay filter for visualization purposes. The shadowed region shows the difference between the actual MCMC-chain lengths and the filtered estimation. Large values of MCMC-chain length (say $\gtrsim 250$) characterize the more expensive steps, where the proposal method requires more iterations in order to propose new samples and estimate the boundaries of the current contour. During the initial stages ($X \approx 1$), the boundaries defined by the current live points were comparable with the prior and the sampler required many iterations in order to propose new samples. Subsequently, for $\log X \lesssim 10^4$, the sampler identified the region that encloses the majority of the posterior volume and the algorithm advanced spending ~ 100 iterations to propose a new sample. The length of the MCMC-chain slightly increases again during the latest stages, since the sampler has to reach the bulk of the posterior distribution restricting the boundaries to a neighborhood of the maximum-likelihood values.

Table II shows the recovered median values, and Fig. 7 shows the marginalized posterior distribution for some exemplary parameters. The recovered values are in agreement with the properties of the injected signal: the posterior distribution encloses the injected sample for all the parameter in the 50% credible region, with a small bias in the in the maximum-posterior value for the reduced tidal

parameter $\tilde{\Lambda}$, corresponding to roughly ~ 150 . However, this behavior is expected [184–188] considering that we employed an upper cutoff frequency of 1 kHz¹ combined with the large aligned-spin prior.² The estimated evidences slightly prefer TEOBResumSPA approximant, accordingly with the injected model; however, due to the low SNR, they do not show any strong preference.

C. Binary neutron stars postmerger

We perform a set of BNS postmerger injections using a five-detector network which includes H1 and L1 with LIGO design sensitivity P1200087 [27,105,181], V1 with Virgo design sensitivity P1200087 [28], K1 with KAGRA design sensitivity T1600593 [107,112], and the third-generation interferometer ET (configuration D) P1600143 [114,115]. The injected signals are generated with NRPM and correspond to the postmerger radiations of nonspinning BNSs with intrinsic parameters $\{\mathcal{M} = 1.188 M_{\odot}, q = 1, \Lambda_1 = \Lambda_2 = 600\}$. The signal is injected such that the merger occurs at GPS time 1126259462.0 with a sampling rate of 8192 Hz and a signal length of 4 s. The sky location of the source corresponds to the position of maximum sensitivity for the detector ET, $\{\alpha = 2.640, \delta = 0.762\}$. The signals are injected at different luminosity distances in order to simulate different SNRs, spanning the range from 8 to 32, which corresponds to $D_L \in [20, 80]$ Mpc. We observe that, for these kinds of signal, ET is the most relevant detector and it concurs in the determination of the SNR with a weight larger than 90% for all the analyzed cases.

For these studies, we employ the PTMCMC sampler using eight tempered ensembles with 128 chains and 2×10^4 burn-in iterations. The injected strain is analyzed in the frequency domain from 1 kHz to 4 kHz employing NRPM waveforms.³ We set the chirp mass prior in $[0.8, 2] M_{\odot}$ and the mass ratio in $[1, 1.5]$. The spins are kept fixed to zeros for every component. The tidal parameters are extracted in the range $[0, 5000]$. We employ volumetric prior for the luminosity distance in the support $[5 \text{ Mpc}, 400 \text{ Mpc}]$ and the prior distributions for the remaining parameters are chosen according to Sec. V B.

As shown in Ref. [190], the postmerger GW radiation of a long-lived BNS merger is characterized by a main frequency peak in the Fourier domain (f_2) that can be parametrized with quasiuniversal (EOS-insensitive) relations involving the tidal polarizability parameters. The NRPM model is constructed using these relations that allow one to constrain $\tilde{\Lambda}$ from postmerger observations and at the

¹For typical BNS, the information on the tidal parameters is gathered in frequency range above 800 Hz [127,188].

²Large spin effects can mitigate the tidal contributions, leading to an underestimate of the tidal parameters [187].

³The same template model is used for both injection and recovery in order to avoid noise contributions different from the simulated detector noise.

TABLE II. Recovered parameters during the inspiralling BNS injections studies. The signal has been injected in H1 + L1 using design sensitivity curves, with an overall network SNR of 20. The data have been analyzed with the nested sampling provided by DYNESTY. The reported values correspond to the medians with the 90% credible regions. The last column reports the estimated logarithmic Bayes' factor and the associated standard deviation.

| Approximant | \mathcal{M} [M_{\odot}] | q | χ_{eff} | $\tilde{\Lambda}$ | D_L [Mpc] | ι [rad] | α [rad] | δ [rad] | $\log \mathcal{B}_N^S$ |
|-----------------|-------------------------------|------------------------|------------------------|---------------------|-------------------|------------------------|------------------------|------------------------|------------------------|
| Injected | 1.188 | 1.00 | 0.00 | 600 | 120 | 0.00 | 0.372 | 0.811 | ... |
| TEOBResumSPA | $1.1880^{+0.0002}_{-0.0002}$ | $1.20^{+0.42}_{-0.17}$ | $0.01^{+0.05}_{-0.01}$ | 435^{+305}_{-248} | 113^{+19}_{-38} | $0.60^{+1.98}_{-0.42}$ | $0.48^{+0.26}_{-0.12}$ | $0.86^{+0.20}_{-0.21}$ | $564.6^{+0.3}_{-0.3}$ |
| TaylorF2 | $1.1880^{+0.0005}_{-0.0001}$ | $1.28^{+1.16}_{-0.25}$ | $0.01^{+0.10}_{-0.01}$ | 392^{+415}_{-260} | 106^{+26}_{-38} | $0.82^{+1.91}_{-0.56}$ | $0.75^{+3.51}_{-0.33}$ | $0.81^{+0.28}_{-1.56}$ | $564.3^{+0.3}_{-0.3}$ |
| IMRPhenomPV2NRT | $1.1880^{+0.0002}_{-0.0001}$ | $1.25^{+0.36}_{-0.21}$ | $0.01^{+0.02}_{-0.01}$ | 316^{+304}_{-215} | 111^{+20}_{-50} | $0.76^{+1.98}_{-0.55}$ | $0.63^{+3.54}_{-0.21}$ | $0.79^{+0.24}_{-1.49}$ | $563.9^{+0.3}_{-0.3}$ |

same time map the properties of a postmerger signal into the inspiral parameters of the binary [118]. Furthermore, numerical relativity simulations [118,191] have shown that the postmerger frequency peak f_2 can be related to the radius R_{max} of the maximum mass configuration of a nonrotating neutron star. For the injected sources, we get $f_2 = 2.94$ kHz and $R_{\text{max}} = 10.8$ km.

Table III shows the recovered mean values and Fig. 8 shows the marginalized posterior distributions for f_2 and for $\tilde{\Lambda}$ recovered during the survey described above, and Fig. 9 presents the recovered postmerger signal recovered in the ET detector for the case with SNR 16. The Bayes'

factors show evidence of signal from SNR 9; however, in order to estimate f_2 with an accuracy of $O(0.1$ kHz), the method requires a SNR $\gtrsim 12$. The mean values estimated from the marginalized posteriors agree with the injected properties of the signal within the 90% credible intervals; however, it is interesting to observe that, due to the correlations induced by the EOS-insensitive relations, the sampler explores nontrivial degeneracy between the intrinsic parameters. The same behavior has been shown in Ref. [118]. These correlations strongly affect the estimation of $\tilde{\Lambda}$ and PE of a postmerger signal is only capable of imposing an upper bound for this parameter. For example,

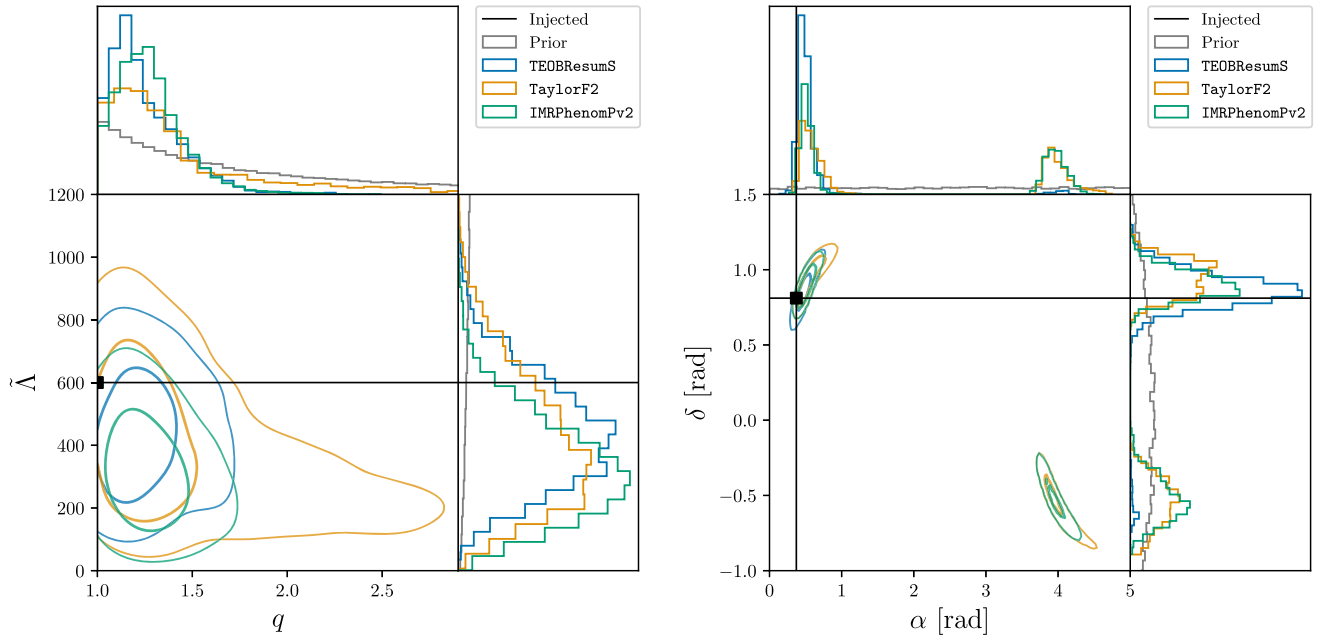


FIG. 7. Posterior distributions for $\{q, \tilde{\Lambda}\}$ and $\{\alpha, \delta\}$ recovered from the injection studies performed on an inspiralling BNS signal with two interferometers (H1 + L1) at design sensitivities with network SNR of 20. The artificial signal has been generated with TEOBResumS model and the injected parameters are marked with black lines and squares. The recovery has been performed with three different approximants analyzing frequency range from 20 Hz to 1 kHz. The degeneracy in the sky location can be removed introducing a third detector [189], and it is due to the correlations between longitudinal and latitudinal angles that concur in the estimation of times of arrival in the different interferometers.

TABLE III. Recovered parameters during the BNS postmerger injection survey. The reported values correspond to the medians with the 90% credible regions.

| SNR | \mathcal{M} [M_{\odot}] | $\tilde{\Lambda}$ | f_2 [kHz] | R_{\max} [km] | $\log \mathcal{B}_{\text{N}}^{\text{S}}$ |
|----------|-------------------------------|------------------------|------------------------|---------------------|--|
| Injected | 1.188 | 600 | 2.94 | 10.8 | ... |
| 8 | $2.2^{+0.8}_{-1.5}$ | 2700^{+1800}_{-1800} | $0.8^{+2.1}_{-0.3}$ | 27^{+10}_{-19} | $0.2^{+0.2}_{-0.2}$ |
| 9 | $1.3^{+1.6}_{-0.7}$ | 2350^{+1900}_{-1700} | $1.7^{+1.2}_{-1.1}$ | 14^{+22}_{-6} | $2.6^{+1.3}_{-1.3}$ |
| 10 | $0.8^{+2.2}_{-0.2}$ | 2830^{+1630}_{-1910} | $2.7^{+0.3}_{-2.0}$ | $8.9^{+2.8}_{-0.9}$ | $11.0^{+4.8}_{-4.8}$ |
| 12 | $0.8^{+2.1}_{-0.2}$ | 1860^{+2570}_{-1410} | $2.9^{+0.7}_{-2.3}$ | $9.0^{+2.7}_{-0.8}$ | $13.9^{+6.3}_{-6.3}$ |
| 16 | $0.78^{+0.43}_{-0.18}$ | 1780^{+2640}_{-1220} | $2.93^{+0.02}_{-0.39}$ | $8.9^{+2.0}_{-0.8}$ | $45.2^{+4.2}_{-4.2}$ |
| 32 | $0.80^{+0.45}_{-0.12}$ | 1730^{+730}_{-1220} | $2.93^{+0.02}_{-0.03}$ | $9.0^{+2.1}_{-0.6}$ | 271^{+29}_{-29} |

with SNR 32, it would be possible to constrain the value of $\tilde{\Lambda}$ with an uncertainty of $\sim 10^4$ that is a much larger value compared with the uncertainties coming from the analysis of the inspiral data (see Table II). Nevertheless, the observation of a postmerger signal would extraordinarily extend our knowledge regarding neutron star matter [154,157,192], allowing us to verify our current models and to constrain the extreme-density properties of the EOS, such as the radius of the maximum mass star R_{\max} and the inference of softening effects at high densities [118,193]. These constraints can be further improved with the inclusion of the inspiral information within a full inspiral-merger-postmerger analysis of the observed BNS signal.

D. Confidence interval test

The confidence interval (CI) test has become a standard control check to verify the reliability of a GW pipeline (e.g., [2,7,194]), since it ensures that the recovered probability distributions are truly representative of the inferred

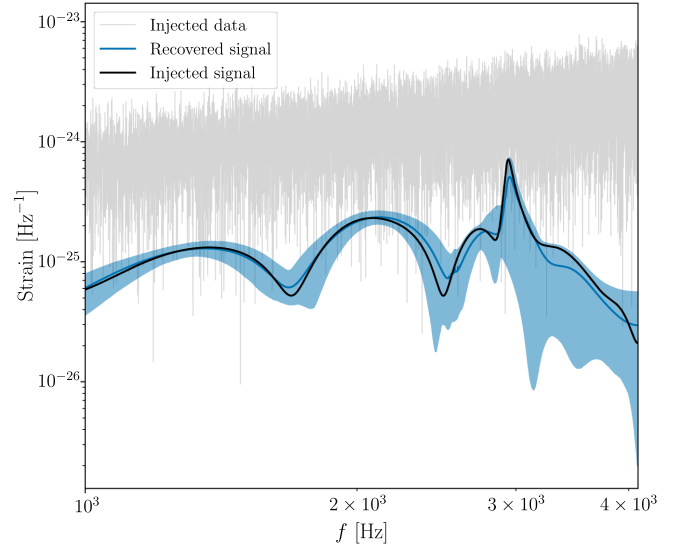


FIG. 9. Frequency-domain amplitude strains in the ET detector for the analysis at SNR 16 corresponding to a luminosity distance of 40 Mpc. The plotted quantities correspond to the injected signal (black), the full artificial data strain (gray), and the recovered template (blue), where the solid line is the mean value and the shadowed area is the 90% credible region.

confidence levels. For each parameter employed in the analyses, the CI test measures the fraction of true values that follow below a given credible level and, if the algorithm is well calibrated, we expect these two quantities to be proportional. The test is performed using a large set of injected signals with parameters extracted from a population prior $\mathbf{p}(\theta_{\text{cbc}})$. Then, we conduct PE analyses in order to recover the posterior distributions for every injection and the CIs can be estimated from the posterior distributions by determining the quantiles under which the true parameters

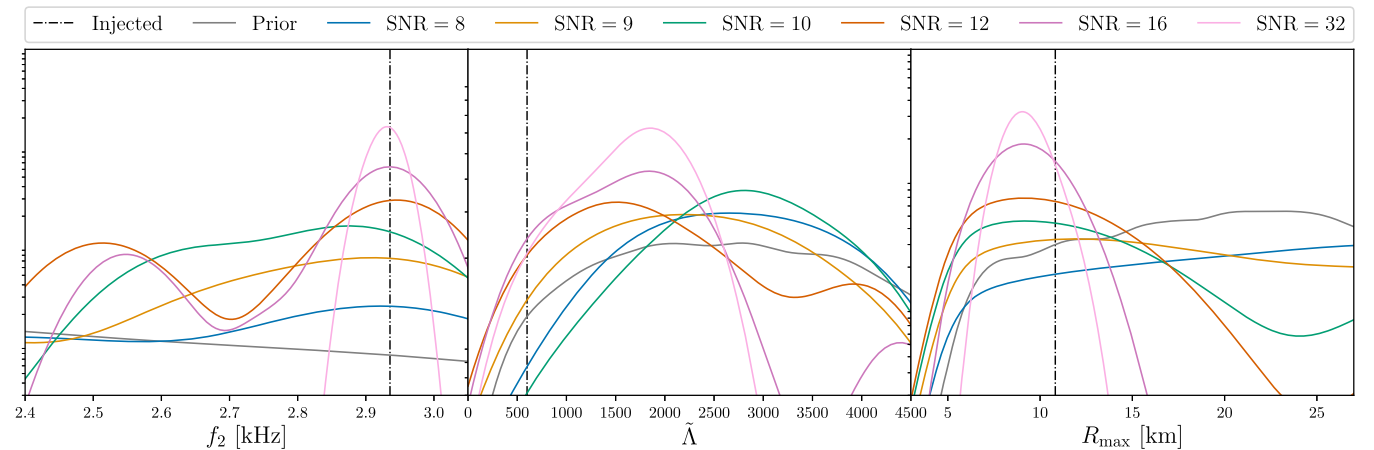


FIG. 8. Marginalized posterior distributions for $\{f_2, \tilde{\Lambda}, R_{\max}\}$ recovered in the BNS postmerger injection survey using a five-detector-network H1 + L1 + V1 + K1 + ET at design sensitivities varying the luminosity distance and locating the source in the position of maximum sensitivity for ET. The injected signals have been generated with NRPM template and recovered with the same model.

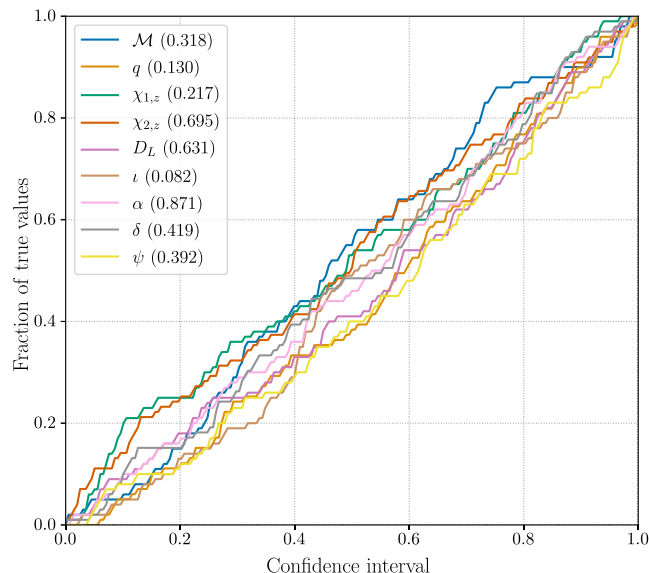


FIG. 10. Fraction of events found within a confidence interval for test described in Sec. VID, for which a set of 100 BBH injections was used. For every parameter, the label shows the p -value of the KS test. The recovered p -values are uniformly distributed with a p -value of 58%.

lie. Then, the population prior $\mathbf{p}(\theta_{\text{cbc}})$ is used as input prior distribution for the analysis of the injected signal.⁴

For our CI test, we inject 100 BBH signals employing the prior distribution used for the parameters of GW150914 (see Sec. VII A) that includes nine parameters. The signals are generated using MLGW waveform template and injected in the two LIGO detectors, H1 and L1, at design sensitivity using segments of duration 8 s. The analyzed frequency range goes from 20 Hz to 1 kHz. We use the nested sampling provided by DYNESTY with 1024 live points and tolerance equal to 0.1. The likelihood function is marginalized over reference phase and time shift. Figure 10 shows the recovered fractions of events found within an increasing confidence level. The fraction of event is expected to be uniformly distributed if the prior distributions are a good approximation of the underlying injected population distributions.

For each parameter, we compute the p -value of Kolmogorov-Smirnov (KS) test, quantifying the consistency with uniformly distributed events. The results are shown between round brackets in Fig. 10. From these results, we estimate the combined p -values quantifying the probability that the individual p -values are extracted from a uniform distribution. We estimate an overall p -value of 58%, according with analogous computations performed with a set of nine random uniformly distributed samples.

⁴We observe that in order to perform an accurate test the employed prior has to be a good representation of the population distribution. In our case, this is ensured by definition since the injected signals are extracted from the employed prior $\mathbf{p}(\theta_{\text{cbc}})$.

Furthermore, we compute the Jensen-Shannon (JS) divergence between the distribution of fraction of events with respect to a uniform distribution. The results lie around 2×10^{-3} for all the parameters, in agreement with analogous estimations performed on a set of 100 uniformly distributed random samples. These results confirm that the pipeline is well calibrated.

VII. LIGO-VIRGO TRANSIENTS

In this section, we apply the BAJES pipeline to the GW transients [1] observed by the LIGO-Virgo interferometers [27,28,105]. For all the analyses, the data are extracted from the GWOSC archive [47,79,80] with a sampling rate of 4096 Hz resorting to the GWPY facilities. The analyzed strains are centered around the nominal GPS time. Subsequently, the strains are windowed and transformed in the Fourier space, using the tools described in Sec. IV. PSD curves and calibration uncertainties are taken from the official LIGO-Virgo data release of GWTC-1 [1,195,196]. We use the nested sampling implemented in DYNESTY, employing 2048 live points with a tolerance equal to 0.1. With these settings, we collect $\sim 5 \times 10^4$ samples from each

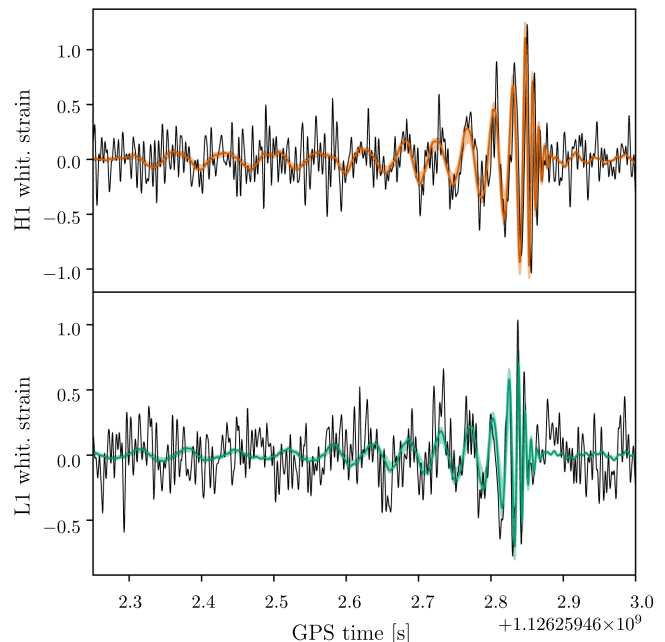


FIG. 11. Waveform template reconstructed from the analysis of GW150914 with TEOBResumS compared with the LIGO-Hanford (top panel) and LIGO-Livingston (bottom panel) data. The black lines are the whitened strains recorded by the LIGO interferometers, where we applied a band-pass filter in the frequency window [20 Hz, 480 Hz] only for visualization purposes. The colored lines are the median recovered template projected on the respective detector (orange for LIGO-Hanford and green for LIGO-Livingston), and the shadowed areas represent the 90% credible regions. The estimated network SNR of the signal corresponds to 22.

PE analysis. The measured quantities reported in the text correspond to the median values and to the 90% credible regions, except when explicitly mentioned.

We note that the prior assumptions employed in BAJES slightly differ from the ones of the official LIGO-Virgo analysis. In the latter, the sampling is performed imposing additional bounds in the mass components [12,15,73,197]; while, in BAJES, the samples are extracted from the entire square defined by the $\{\mathcal{M}, q\}$ bounds. Moreover, the strains analyzed by the BAJES pipeline are slightly shifted in time with respect to the official LIGO-Virgo segments due to different reference conventions.

A. GW150914

In this section, we discuss the results obtained from the analysis of the first GW transient observed by the LIGO interferometers, GW150914 [73,183,198,199]. For all the discussed cases, the analyzed strains correspond to the GWTC-1 release [1] of LIGO-Hanford and LIGO-Livingston data centered around GPS time 1126259462 with a sampling rate of 4096 Hz and a duration of 8 s. We set the lower cutoff frequency to 20 Hz and the highest frequency to 1 kHz. The employed prior is isotropic in spin components and volumetric in luminosity distance, and it spans the ranges $\mathcal{M} \in [12, 45] M_\odot$, $q \in [1, 8]$, $\chi_{1,2} \in [0, 0.99]$, and $D_L \in [100, 5000]$ Mpc. We include eight spectral nodes for the calibrations of the analyzed strains.

We discuss first the PE analysis employing TEOBResumS model with aligned spins, including all high-order modes up to $\ell = 5$ with $m = \pm\ell$, i.e., $(2, \pm 2)$, $(3, \pm 3)$, $(4, \pm 4)$, $(5, \pm 5)$. Figure 11 shows the recovered waveform template compared with the whitened strains recorded by the LIGO interferometers, and Fig. 12 shows the recovered posterior distribution. We estimated a network SNR of 23. The results are consistent with similar studies performed with the same approximant [139,146], recovering the signal of a nonspinning equal-mass BBH merger with $\mathcal{M} = 31.57^{+0.82}_{-0.70} M_\odot$. The inference of the extrinsic parameters shows a source located at a luminosity distance of ~ 490 Mpc, and the area covered by the 90% isoprobability level of the sky location posterior distributions corresponds to ~ 670 deg². The estimated Bayes' factor corresponds to $\log \mathcal{B}_N^S = 267.8 \pm 0.2$, where the uncertainty is given by the standard deviation.

We repeat the PE analysis with IMRPhenomPv2 template employing precessing spin components. In Fig. 13, the marginalized posterior distribution of the recovered masses is compared with the official LIGO-Virgo posterior samples release with GWTC-1 [1] performed the LALInference routines [7,26] using the same waveform approximant. The two analyses are consistent with each other, recovering a BBH signal with chirp mass $\mathcal{M} = 31.00^{+1.52}_{-1.49} M_\odot$ and mass ratio well constrained around the equal mass case, $q = 1.18^{+0.36}_{-0.17}$. The inference of the effective spin parameter is consistent with zero and the posterior distribution of the

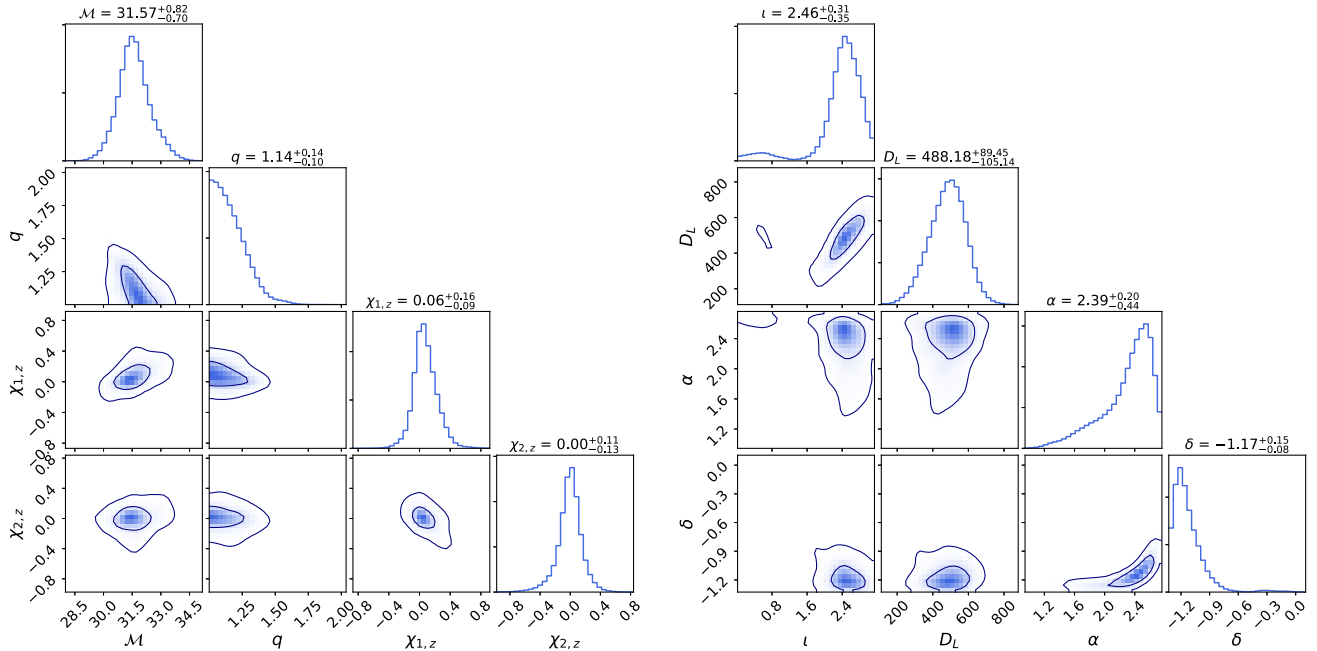


FIG. 12. Posterior distributions for the intrinsic (left) and extrinsic (right) parameters of GW150914 estimated with BAJES pipeline employing TEOBResumS waveform model with aligned-spin components. All higher-order modes up to $\ell = 5$ (with $m = \pm\ell$) have been used to generate the waveform models. The chirp mass is expressed in solar masses M_\odot , the luminosity distance is expressed in megaparsec Mpc, while the angles $\{i, \alpha, \delta\}$ are in radians. We report the median value and the 90% credible regions for each parameter and the contours represent the 50% and the 90% credible regions.

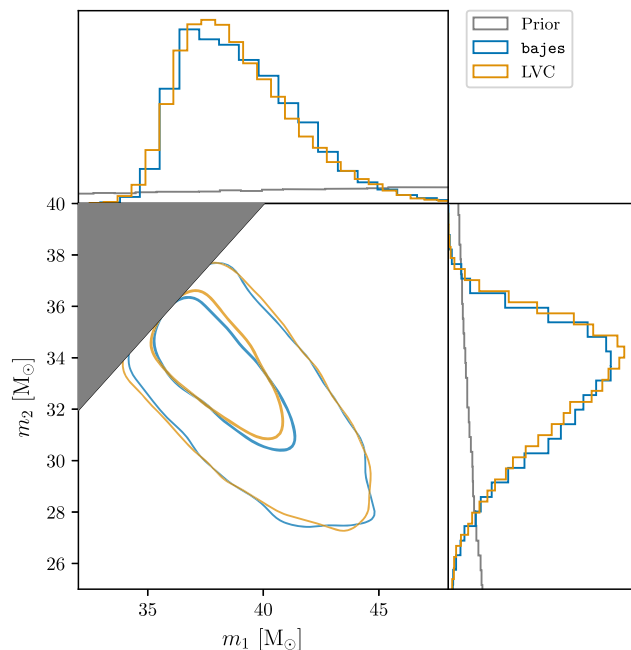


FIG. 13. Posterior distributions of the detector-frame mass components $\{m_1, m_2\}$ recovered in analysis of GW150914 with IMRPhenomPv2 (blue line). The results are compared with the official LIGO-Virgo posterior samples released with GWTC-1 [1] (yellow line) computed using the LALInference routines [7,26]. The central panel shows the 50% and the 90% credible regions.

spin components does not show evidence of precession, according to Refs. [183,197,198]. Also, the extrinsic parameters show an overall good agreement with previous estimations performed with the same approximant, locating the source at a luminosity distance of $D_L = 458_{-169}^{+123}$ Mpc with a posterior sky-location area of ~ 640 deg² at the 90% credible level.

The main difference between the TEOBResumS posterior and IMRPhenomPv2 one is the uncertainty on the mass ratio parameter, for which TEOBResumS admits a largest value of 1.28 at the 90% credible region, while the IMRPhenomPv2 posterior reaches 1.94 with the same confidence. However, this disagreement is mainly due to the different spin assumptions employed for the two analyses [183,197]. Moreover, the posterior distribution recovered with IMRPhenomPv2 shows slightly smaller \mathcal{M} and larger D_L compared with the TEOBResumS inference, as shown also in Ref. [139].

Finally, we verify the compatibility of the recovered posterior distribution against of existing GW pipelines. In particular, we employ the BILBY pipeline [31,72,194] in order to estimate the posterior distributions of GW150914, using the same prior assumptions and settings discussed above. We observe that GW150914 is a suitable candidate to test the statistical significance of the results and the agreement between the pipelines: due to the loudness of this event (corresponding to a SNR > 20), the overall

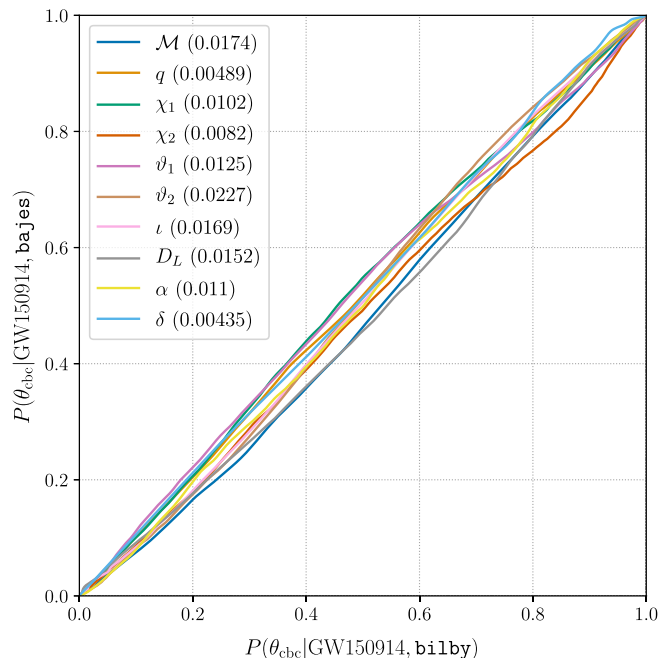


FIG. 14. PP plot for the marginalized posterior distributions of GW150914 parameters. On the x axis, the cumulative posterior probabilities estimated with the BILBY pipeline and on the y axis the same quantities computed with the BAJES pipeline. Different colors refer to different parameters and the legend shows also the square root of the JensenShannon divergence.

impact of statistical noise fluctuations on the recovered posterior distribution is expected to be less determinant compared with the other BBH mergers presented in GWTC-1 [1]. Figure 14 shows the probability-probability (PP) plot of the marginalized posterior distributions recovered for every parameter. A PP plot compares the cumulative distributions estimated with two methods, plotting one against the other. Then, if two probability distributions are identical, the associated PP plot is represented by a bisector line. In our case, the results coming from the two pipelines are largely consistent between each other, with observed deviation fully consistent with statistical fluctuations. This fact is confirmed by the p -values computed between the marginalized distributions of each parameter: the values are comparable to or larger than 0.4 except for ϑ_2 and D_L parameters, for which we, respectively, get p -values of 0.15 and 0.11. Furthermore, the legend in Fig. 14 shows the square root of the JS divergence of the marginalized posterior distributions for the employed parameters: the largest recovered value corresponds to 5.2×10^{-4} bit for the tilt angle ϑ_2 .

B. GW170817

We analyze the LIGO-Virgo data corresponding to GW170817 [12,15,185,200], the first GW signal observed from a BNS merger. The employed data correspond to the GWTC-1 release [1] of LIGO-Virgo data centered around

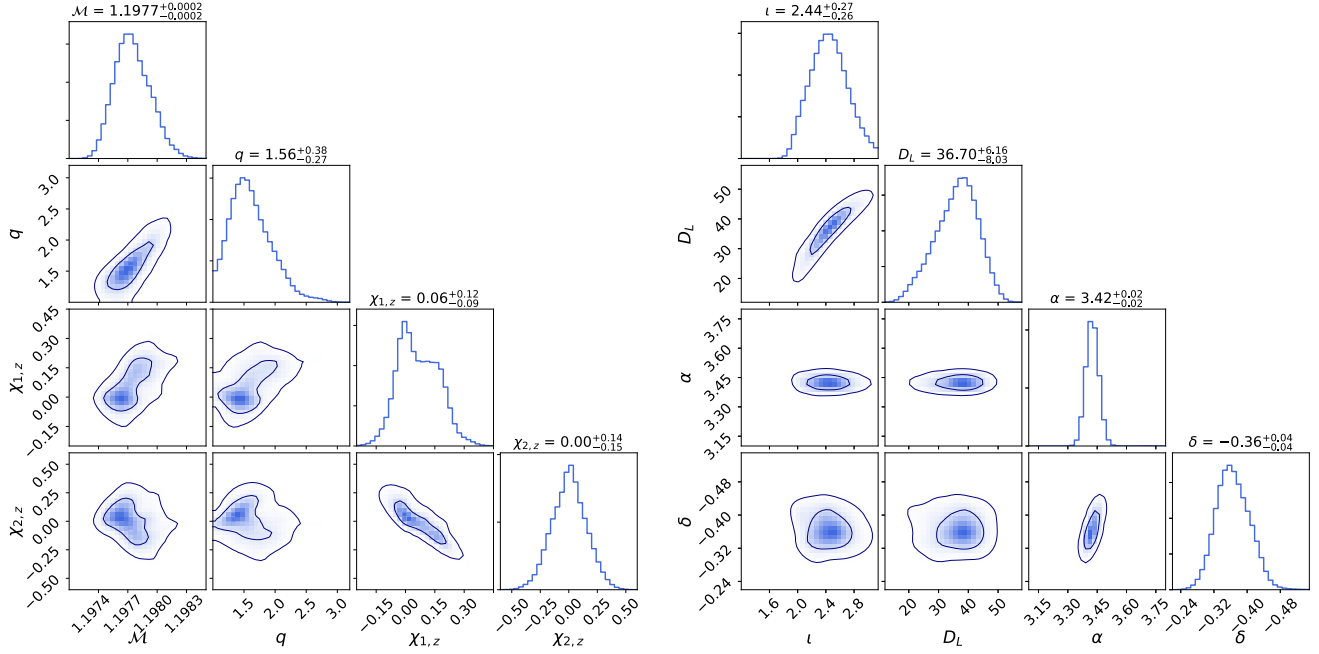


FIG. 15. Posterior distributions for the intrinsic (left) and extrinsic (right) parameters of GW170817 estimated with BAJES pipeline employing TEOBResumSPA waveform model with aligned-spin components. The chirp mass is expressed in solar masses M_\odot , the luminosity distance is expressed in megaparsec Mpc, while the angles $\{\iota, \alpha, \delta\}$ are in radians. We report the median value and the 90% credible levels for each parameter and the contours represent the 50% and the 90% credible regions.

GPS time 1187008857 with a sampling rate of 4096 Hz and a duration of 128 s analyzing the frequency range from 20 Hz to 2 kHz. The employed prior is isotropic in spin components and volumetric in luminosity distance, and it spans the ranges $\mathcal{M} \in [1.18, 1.21] M_\odot$, $q \in [1, 8]$, $\chi_{1,2} \in [0, 0.89]$, $\Lambda_{1,2} \in [0, 5000]$, and $D_L \in [5, 75]$ Mpc. We include four spectral nodes for the calibrations of the analyzed segments.

Figure 15 shows the posterior distributions for the parameters recovered employing TEOBResumSPA approximant (with $\ell = 2$ and $|m| = \ell$) with aligned spins. The recovered detector-frame chirp mass corresponds to $\mathcal{M} = 1.1977^{+0.0002}_{-0.0002} M_\odot$ and the mass ratio lies around the equal mass case, $q = 1.56^{+0.38}_{-0.27}$. The spin components do not show evidence of spin contributions, consistently with Ref. [185], with an estimated effective spin of $\chi_{\text{eff}} = 0.04^{+0.06}_{-0.03}$. The primary tidal parameter Λ_1 is constrained to be $\lesssim 950$ at the 90% confidence level, while the secondary component Λ_2 is more broadly distributed over the prior. The recovered tidal parameter posterior estimates a value of $\tilde{\Lambda} = 607^{+477}_{-356}$, in rough agreement with previous estimations obtained from EOB models [1,149]. The asymmetric tidal parameter $\delta\tilde{\Lambda}$ shows a posterior distribution centered around nonzero values, $\delta\tilde{\Lambda} = 92^{+200}_{-258}$; however, the hypothesis $\delta\tilde{\Lambda} = 0$ is confidently included in the posterior support, corresponding to the 27th percentile. Moreover, the measured $\tilde{\Lambda}$ is overall consistent with independent estimations coming from the analysis of the

EM counterpart AT2017gfo [25,155]. Regarding the extrinsic parameters, we recovered a luminosity distance of $D_L = 36.7^{+6.2}_{-8.0}$ Mpc and a sky location at $\{\alpha = 3.42^{+0.02}_{-0.02}$ rad, $\delta = -0.36^{+0.04}_{-0.04}$ rad $\}$. The estimation of the extrinsic parameters is generally consistent with previous estimations [1,149,185].

Furthermore, the analysis of GW170817 is repeated with a TaylorF2 waveform template that includes for the first time 5.5PN point-mass corrections [125] and 7.5PN tidal contributions [126,127]. This analysis is performed using the same prior assumptions described above and the posterior distribution for the intrinsic parameters is shown in Fig. 16. The mass parameters recover the values $\mathcal{M} = 1.1976^{+0.0003}_{-0.0002} M_\odot$ and $q = 1.42^{+0.79}_{-0.36}$, while the effective spin $\chi_{\text{eff}} = 0.02^{+0.07}_{-0.04}$. The estimated luminosity distance of $D_L = 37.1^{+10.3}_{-12.1}$ Mpc. The inference of the tidal components roughly coincides with the estimations coming from the analogous analysis with PN templates [1,185]. The primary tidal component is constrained below $\Lambda_1 < 730$ at 90% credible region, while the secondary is more broadly distributed over the prior. The reduced tidal parameter is measured to be $\tilde{\Lambda} = 404^{+701}_{-246}$, and the asymmetric tidal term $\delta\tilde{\Lambda}$ is well constrained around zero with an uncertainty of ~ 150 at the 90% confidence level. The median values of the sky location angles coincide with the estimation performed with TEOBResumSPA.

The GW170817 PE studies are repeated with IMRPhenomPv2_NRTidal template employing precessing spin

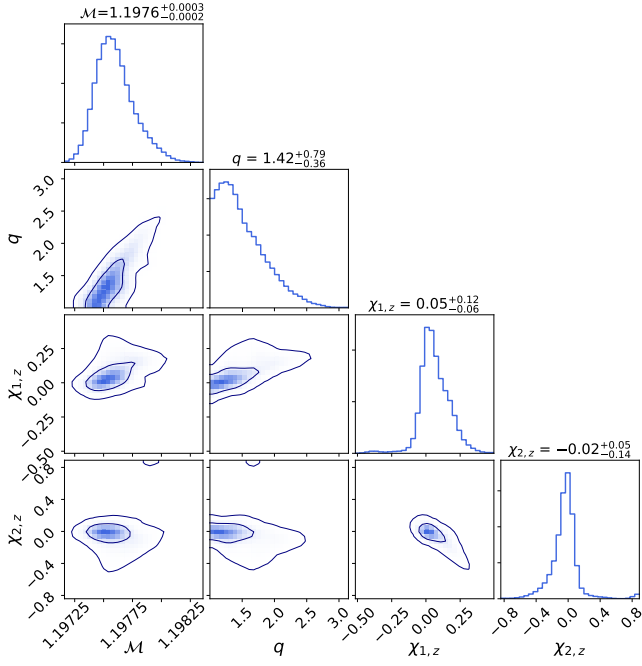


FIG. 16. Posterior distributions for the intrinsic parameters of GW170817 estimated with BAJES pipeline employing TaylorF2 (5.5PN + 7.5PN) waveform model with aligned-spin components. The chirp mass is expressed in solar masses M_{\odot} . We report the median value and the 90% credible levels for each parameter and the contours represent the 50% and the 90% credible regions.

components with high-spins prior ($\chi_{\max} = 0.89$) and low-spins prior ($\chi_{\max} = 0.05$) in order to compare our results with the official LIGO-Virgo posterior samples presented in Ref. [185]. The recovered posterior distributions for the mass parameters for the low-spin case show a chirp mass of $1.1975^{+0.0002}_{-0.0002} M_{\odot}$, and the mass ratio is constrained below 1.46 at the 90% credible region. Regarding the high-spin prior analysis, we recovered a chirp mass of $1.1976^{+0.0002}_{-0.0002} M_{\odot}$ and the mass ratio favors more asymmetric values, $q = 1.49^{+0.35}_{-0.32}$. Focusing on the tidal parameters, Fig. 18 shows the comparison of the posterior distribution in the $\{\tilde{\Lambda}, \delta\tilde{\Lambda}\}$ plane: the marginalized distributions and the 90% credible regions coming from BAJES are largely consistent with the official LIGO-Virgo samples, with estimated JS divergences below 10^{-4} .

Finally, Fig. 17 shows the recovered reduced tidal parameters $\tilde{\Lambda}$, where the posterior distributions have been reweighted to uniform prior distribution. The figure includes the results computed employing the same waveform models and using a smaller upper cutoff frequency $f_{\max} = 1$ kHz. The main differences between the analyses with different f_{\max} lie in the results of the tidal sector. Overall, the recovered tidal parameters with $f_{\max} = 2$ kHz appear more constrained with respect to the cases with $f_{\max} = 1$ kHz. This behavior is expected considering that the tidal information is gathered in high-frequency

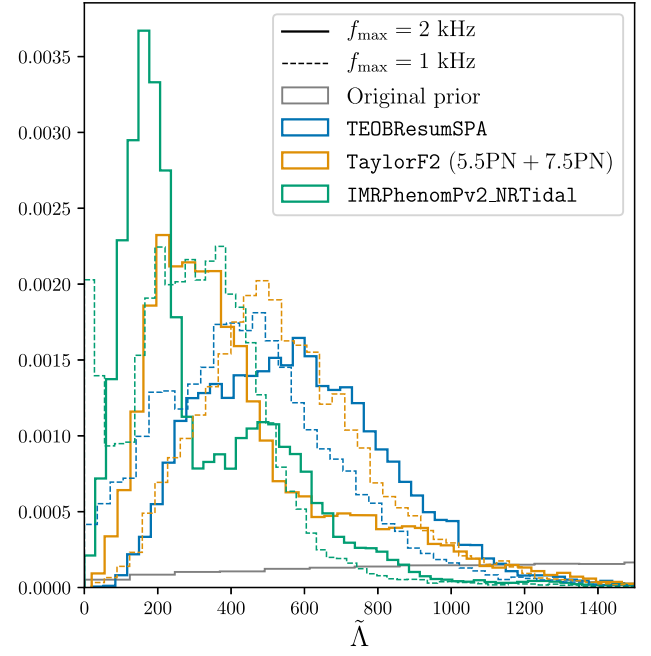


FIG. 17. Posterior distributions for the reduced tidal parameter $\tilde{\Lambda}$ of GW170817 estimated using the BAJES pipeline with different waveform approximants and different upper cutoff frequencies. The posterior samples have been reweighted to uniform prior distribution and the plot shows the original employed prior (gray line). Solid lines refer to the results with $f_{\max} = 2$ kHz, while dashed lines are estimated with $f_{\max} = 1$ kHz. TEOBResumSPA (blue line) and TaylorF2 (yellow line) samples are computed employing aligned spins, while precessing spin components were included for IMRPhenomPv2_NRTidal (green line).

regimes [127,188]. On the other hand, the choice of $f_{\max} = 2$ kHz enlarges multimodal and asymmetric behaviors in the posterior distribution of the reduced tidal parameter and systematic effects appear to be more relevant between different template families, as well known from previous studies [188,201,202]. The differences in the $\tilde{\Lambda}$ parameters can be led back to the modeling choices of the employed approximants [1,125,185][see][]. The results estimated with TEOBResumSPA at 1 kHz show a posterior distribution slightly shifted toward lower values with respect to the analysis at 2 kHz, consistently with what has been observed in the BNS injection study (Sec. VI B). The posterior distributions for EOB and PN approximants show a good agreement with $f_{\max} = 1$ kHz.

C. EOB catalog

As full scale application, we reproduce the analyses of the BBH mergers published in GWTC-1 [1] employing the BAJES pipeline and the time-domain EOB waveform model TEOBResumS, including only the dominant (2,2) mode. Table IV shows the priors used for each event, where the nominal GPS time refers to the central value of the

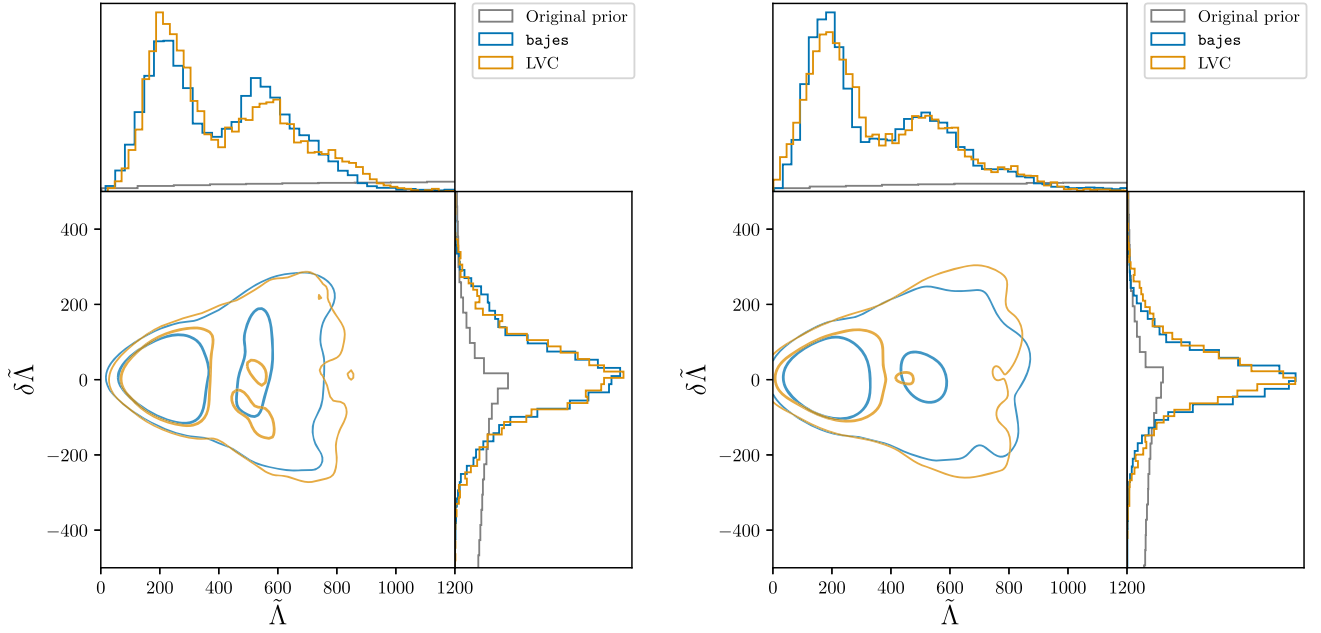


FIG. 18. Comparison between the posterior distributions for the tidal parameters $\{\tilde{\Lambda}, \delta\tilde{\Lambda}\}$ estimated with the BAJES pipeline (blue lines) with the official LIGO-Virgo released samples [185] (yellow lines) computed with the LALInference routines [7,26]. The analyses are performed with IMRPhenomPv2_NRTidal waveform approximant. The left panel shows the results with low-spin prior ($\chi_{\max} = 0.05$), while the right panel presents the high-spin prior results ($\chi_{\max} = 0.89$)s. For both panels, the central plots show the 50% and 90% credible regions. We estimate the JS divergences for the marginalized distributions, finding the values of 8.9×10^{-5} bit and 5.3×10^{-5} bit, respectively, for $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ for the high-spin prior case. Regarding the low-spin prior studies, the JS divergences correspond to 8.3×10^{-5} and 4.7×10^{-5} bit, respectively, for $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$.

analyzed time axis. For all the studies, the analyzed frequency range goes from 20 Hz to 1 kHz and we assume aligned-spin components with isotropic prior distribution and volumetric prior for the luminosity distance. The prior distributions for the other parameters are chosen

accordingly with Sec. V B. For these studies, we employ eight calibration envelope nodes for each detector, the phase ϕ_0 is marginalized during the likelihood evaluations, and the time-shift parameter t_0 is sampled from a 2 s window centered around the nominal GPS time.

TABLE IV. Prior and posterior information for the analyses of the BBH events of GWTC-1 with TEOBResumS. The GPS time refers to the central value of the time axis. For all studies, we assume aligned spins with isotropic prior. The inferred values refer to the medians of the marginalized posterior distributions and the uncertainties are 90% credible regions, except for the log Bayes' factors $\log\mathcal{B}_{\mathcal{N}}^{\mathcal{S}}$, for which we report the standard deviations.

| Data information | | Prior bounds | | | | Inferred values | | | | | | | |
|------------------|--------------|--------------|-------------------------------|-------|--------------|-----------------|-------------------------------|------------------------|-------------------------|-------------------------|-------------------------|------------------------|---|
| Event | GPS time [s] | Duration [s] | \mathcal{M} [M_{\odot}] | q | $\chi_{1,2}$ | D_L [Mpc] | \mathcal{M} [M_{\odot}] | q | $\chi_{1,z}$ | $\chi_{2,z}$ | χ_{eff} | D_L [Mpc] | $\log\mathcal{B}_{\mathcal{N}}^{\mathcal{S}}$ |
| GW150914 | 1126259462 | 8 | [12,45] | [1,8] | [0,0.99] | [100,5000] | $31.9^{+1.1}_{-1.5}$ | $1.20^{+0.29}_{-0.17}$ | $0.07^{+0.39}_{-0.28}$ | $0.00^{+0.40}_{-0.42}$ | $0.05^{+0.10}_{-0.13}$ | 471^{+167}_{-231} | $267.1^{+0.2}_{-0.2}$ |
| GW151012 | 1128678900 | 16 | [12,45] | [1,8] | [0,0.99] | [100,5000] | $18.3^{+1.8}_{-1.0}$ | $1.86^{+2.86}_{-0.76}$ | $0.05^{+0.33}_{-0.28}$ | $0.11^{+0.53}_{-0.48}$ | $0.09^{+0.22}_{-0.17}$ | 1039^{+627}_{-626} | $16.0^{+0.2}_{-0.2}$ |
| GW151226 | 1135136350 | 16 | [6.5,15] | [1,8] | [0,0.99] | [50,5000] | $9.71^{+0.07}_{-0.07}$ | $2.04^{+1.59}_{-0.93}$ | $0.38^{+0.23}_{-0.24}$ | $0.16^{+0.58}_{-0.47}$ | $-0.05^{+0.45}_{-0.61}$ | 490^{+222}_{-240} | $36.4^{+0.2}_{-0.2}$ |
| GW170104 | 1167559936 | 16 | [12,45] | [1,8] | [0,0.99] | [100,5000] | $25.6^{+1.8}_{-2.1}$ | $1.56^{+0.86}_{-0.46}$ | $0.00^{+0.29}_{-0.32}$ | $-0.04^{+0.40}_{-0.40}$ | $-0.03^{+0.21}_{-0.25}$ | 1069^{+423}_{-446} | $58.4^{+0.2}_{-0.2}$ |
| GW170608 | 1180922494 | 16 | [5,10] | [1,8] | [0,0.99] | [50,5000] | $8.49^{+0.05}_{-0.04}$ | $1.48^{+1.22}_{-0.42}$ | $0.06^{+0.30}_{-0.28}$ | $0.03^{+0.58}_{-0.39}$ | $0.06^{+0.27}_{-0.09}$ | 298^{+146}_{-128} | $80.3^{+0.2}_{-0.2}$ |
| GW170729 | 1187529256 | 4 | [25,175] | [1,8] | [0,0.99] | [100,7000] | $51.4^{+9.1}_{-9.6}$ | $1.84^{+0.95}_{-0.77}$ | $0.47^{+0.39}_{-0.48}$ | $-0.05^{+0.83}_{-0.36}$ | $0.30^{+0.28}_{-0.21}$ | 2495^{+1600}_{-1300} | $27.1^{+0.2}_{-0.2}$ |
| GW170809 | 1185389807 | 16 | [12,45] | [1,8] | [0,0.99] | [100,5000] | $30.3^{+2.3}_{-2.0}$ | $1.45^{+0.72}_{-0.39}$ | $0.07^{+0.33}_{-0.27}$ | $0.17^{+0.58}_{-0.21}$ | $0.17^{+0.24}_{-0.21}$ | 999^{+473}_{-483} | $41.8^{+0.2}_{-0.2}$ |
| GW170814 | 1186302519 | 16 | [12,45] | [1,8] | [0,0.99] | [100,5000] | $26.8^{+1.3}_{-1.0}$ | $1.29^{+0.52}_{-0.26}$ | $0.07^{+0.39}_{-0.28}$ | $0.02^{+0.49}_{-0.38}$ | $0.08^{+0.16}_{-0.12}$ | 540^{+224}_{-189} | $99.6^{+0.2}_{-0.2}$ |
| GW170818 | 1186741861 | 16 | [12,45] | [1,8] | [0,0.99] | [100,5000] | $31.8^{+3.4}_{-2.9}$ | $1.48^{+0.96}_{-0.43}$ | $-0.08^{+0.27}_{-0.35}$ | $0.00^{+0.45}_{-0.36}$ | $-0.06^{+0.33}_{-0.27}$ | 1190^{+594}_{-438} | $29.7^{+0.2}_{-0.2}$ |
| GW170823 | 1187058327 | 4 | [25,175] | [1,8] | [0,0.99] | [100,7000] | $37.4^{+5.5}_{-5.1}$ | $1.57^{+0.94}_{-0.51}$ | $-0.01^{+0.42}_{-0.33}$ | $0.06^{+0.56}_{-0.55}$ | $0.03^{+0.30}_{-0.28}$ | 1690^{+1081}_{-880} | $39.5^{+0.2}_{-0.2}$ |

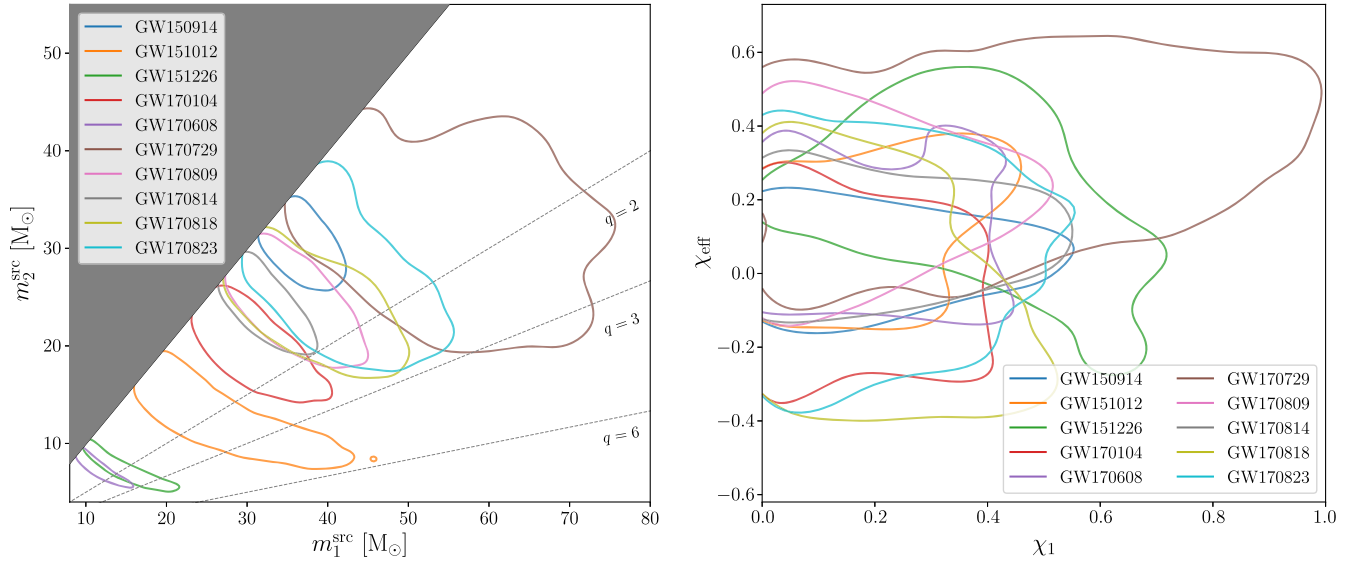


FIG. 19. Marginalized posterior distribution for the source-frame mass components $\{m_1, m_2\}$ (left) and the spins $\{\chi_1, \chi_{\text{eff}}\}$ (right) of the BBH mergers presented in GWTC-1 [1]. The PE studies have been performed with TEOBResums model. The masses are expressed in the binary source frame employing the cosmological model proposed in Ref. [173]. The contours refer to the 90% credible regions.

Figure 19 shows the posterior distributions marginalized in the mass components and spin magnitudes planes for all the analyzed BBH events. The mass components are expressed in the source frame of the binaries assuming the cosmological model presented in Ref. [173]. The detector-frame mass components m_i can be estimated in the source frame of the binary as

$$m_i^{\text{src}} = \frac{m_i}{1+z}, \quad i = 1, 2, \quad (43)$$

where z is the cosmological redshift of the source. In general, the recovered mass parameters show a predominance of equal mass binaries with mass ratio well constrained below $q \lesssim 3$, except for the low-mass binary GW151012, that admits values of $q \simeq 5$ at 90% credible level. The recovered mass components are distributed between $7 M_\odot$ and $70 M_\odot$, with an abundance in the range $[20 M_\odot, 50 M_\odot]$. In terms of spin contributions, the most interesting events are GW151226, whose posterior distribution excludes the nonspinning case at 90% confidence level, consistently with Ref. [1,160], and GW170729, which recovers an effective spin of $\chi_{\text{eff}} \approx 0.3$ and admits values up to $\chi_1 \gtrsim 0.9$. The other GW transients show mitigated spin contributions, with $\chi_1 \lesssim 0.5$ at the 90% credible level. Generally, the posterior distributions for the secondary spin component χ_2 are more broadly distributed and less informative than those for the primary component χ_1 .

Furthermore, Fig. 20 shows the posterior distributions of final masses M_f^{src} (estimated in the source frame) and final spin magnitudes χ_f expected for the merger remnants. The values are computed resorting to NR formulas presented in

Ref. [203], calibrated with aligned-spin BBH simulations. The majority of the recovered final spins χ_f lie around ~ 0.67 due to the moderated spin contributions of the observed mergers. Regarding the extrinsic parameters, Fig. 21 shows the posterior distributions of the sky location.

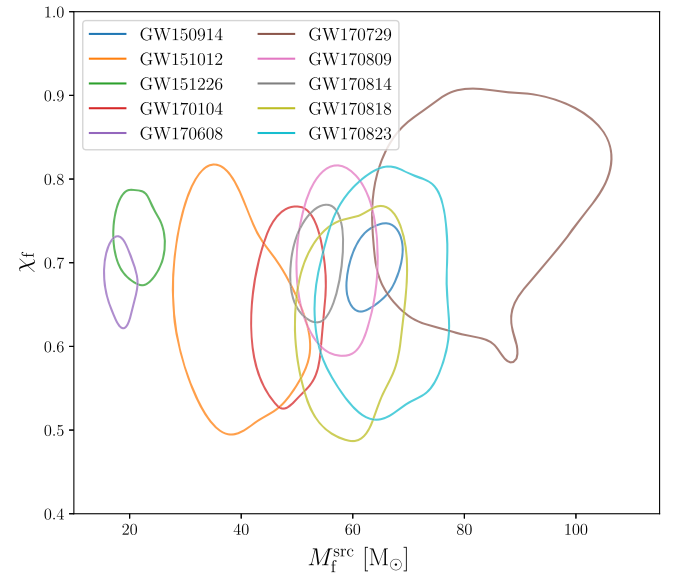


FIG. 20. Marginalized posterior distribution for the final masses and spins $\{M_f^{\text{src}}, \chi_f\}$ for the remnants of the BBH mergers presented in GWTC-1 [1]. The analyses are performed using the BAJES pipeline and TEOBResums waveform approximant. The values have been computed from the posterior samples employing NR formulas presented in Ref. [203]. The final masses are expressed in the source frame of the binary. The contours refer to the 90% credible regions.

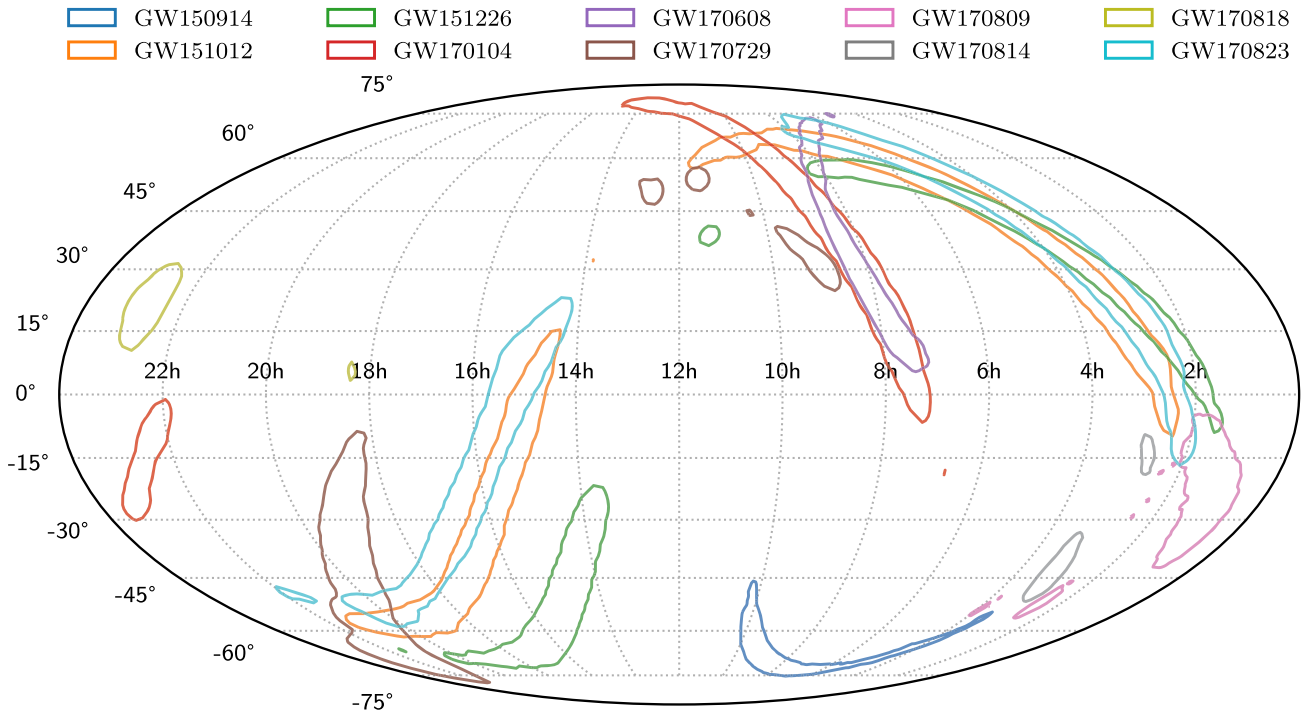


FIG. 21. Marginalized posterior distribution for right ascension and declination angles $\{\alpha, \delta\}$ (represented using a Mollweide projection) for the BBH mergers presented in GWTC-1 [1]. The analyses are performed using the BAJES pipeline and TEOBResumS waveform approximant. The right ascension α is expressed in hours, while the declination δ is reported in degrees. The contours refer to the 90% credible regions.

The sky maps for GW170814 and GW170818 show slightly larger bimodal behavior compared with the results presented in GWTC-1 [1]. On the other hand, the 90% credible region for GW170104 appears to be more constrained. Note that the aligned-spin assumption affects the

overall SNR and, then, the recovered posterior distributions [204,205]. The measurements for the sky locations of the other events do not show considerable deviations from the GWTC-1 estimations [1]. Figure 22 shows the correlations of the luminosity distance D_L with the inclination angle i

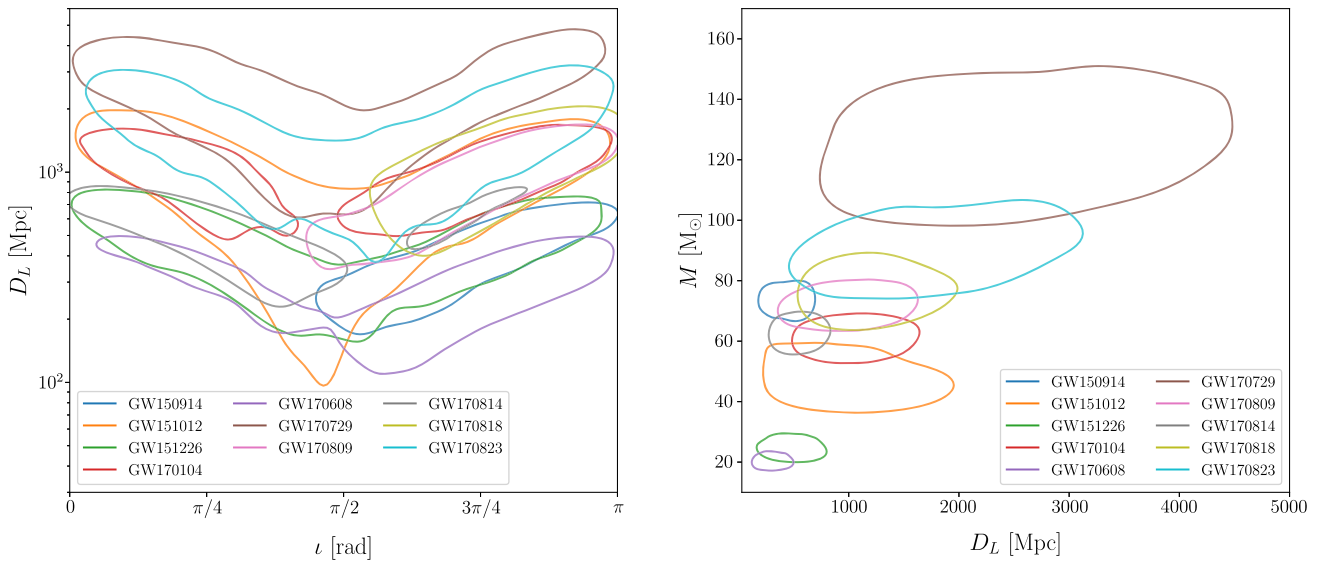


FIG. 22. Marginalized posterior distribution for inclination angle and luminosity distance $\{i, D_L\}$ (left) and for luminosity distance and total mass $\{D_L, M\}$ (right) for the BBH mergers presented in GWTC-1 [1]. The analyses are performed using the BAJES pipeline and TEOBResumS waveform approximant. The contours refer to the 90% credible regions.

and with the total mass $M = m_1 + m_2$ (estimated in the detector frame). The luminosity distances are in agreement with the GWTC-1 estimations, while the inclination angles show slightly wider posterior supports due to the degeneracy introduced with the aligned-spins assumption [206].

For the GW150914 case, we can compare the analysis with higher-order modes (presented in Sec. VII A) with the results estimated using only dominant mode. First of all, the results with $\ell \geq 5$ show narrower uncertainties, consistently with the inclusion of a larger amount of physical information [10,207,208]. On the other hand, the estimated Bayes' factors do not show strong evidence in favor or against the inclusion of higher-order modes, as expected for this kind of source; higher-order modes are expected to be more relevant for large mass ratios and edge-off binaries [209,210]. Overall, the median values of the recovered parameters are consistent between each other except for the mass ratio that appears to be more constrained around $q = 1$ including higher modes.

In conclusion, we shown that TEOBResumS can be effectively applied to BBH signals, obtaining robust and consistent results [139]. The main limitation of the presented results comes from the aligned-spins assumption that introduces degeneracy with other parameters and induces biases in population inferences [211–213]. We are planning to extend the presented catalog including precessing spin terms [147] and verifying the importance of eccentric contributions [116]. In terms of computational cost, TEOBResumS shows an overall good behavior, performing the analysis of a typical BBH (with length of 8 s) in ~ 14 hours on 32 CPUs.

VIII. CONCLUSION

In this paper, we presented BAJES, a parallel, lightweight infrastructure for Bayesian inference, whose main application is the data analysis of gravitational-wave and multimessenger transients. BAJES is implemented in PYTHON and comes with a versatile framework for Bayesian inference and different state-of-art samplers. Furthermore, it provides methods for the analysis of GW and EM transients emitted by compact binary coalescences. We benchmarked BAJES by means of injection-recovery experiments with BBH merger, BNS inspiral-merger and postmerger signals. The injection studies and statistical tests show that the BAJES pipeline is well calibrated and it provides robust results, within the expected statistical fluctuations.

The injections of BNS postmerger signals also offered the first detectability study with a five-detector network including LIGO, Virgo, KAGRA, and third-generation ET [114,115]. We find BNS postmerger signal will be detectable for optimally oriented sources located at $\lesssim 80$ Mpc. This result is largely merit of the ET sensitivity [114,115] that contributes to 90% of the SNR. According with recent population studies [11] and using the distance threshold estimated from our survey with the third-generation

network (~ 80 Mpc), the detection rate of these sources is expected to be 0.5–2 events per year. As discussed in Sec. VI C, the detection of such a transient, combined with the knowledge of EOS-insensitive relations, can reveal essential properties of the nuclear matter at high densities, improving significantly the EOS constraints [118].

We demonstrated the reliability of BAJES in analyzing the observational data recorded by the LIGO-Virgo interferometers [1,47]. The posterior distributions for the parameters of GW transients computed with the BAJES are in agreement with the results from other GW pipelines [7,30]. The direct comparison of the BAJES results on GW150914 with the ones obtained with the BILBY [31,72,194] pipeline shows a maximum JS divergence of 5.2×10^{-4} bit for the tilt angle ϑ_2 and the marginalized posterior distributions are largely consistent between each other. Furthermore, the analyses of GW150914 with TEOBResumS approximant slightly emphasize the relevance of higher-order modes for improving the accuracy of the binary properties' estimations.

We performed PE studies on GW170817 using TEOBResumSPA, IMRPhenomPv2_NRTidal and, for the first time, TaylorF2 including 5.5PN point-mass contributions and 7.5PN tidal terms. The novel analysis with the extended PN model shows a good agreement with previous estimations performed with the same template family [1,185]. Using IMRPhenomPv2_NRTidal template, we found full consistency with previous results and the official LIGO-Virgo posterior samples [1,185]. Posterior distributions for the reduced tidal parameter $\tilde{\Lambda}$ recovered with upper cutoff frequency $f_{\max} = 2$ kHz show larger systematic biases between different waveform templates compared with the 1 kHz analyses. However, using a larger f_{\max} , it is possible to take into account a larger amount of tidal information that leads to more constrained measurements. The tidal parameter $\tilde{\Lambda}$ estimated with TEOBResumSPA has slightly larger values compared with the measure of TaylorF2 and IMRPhenomPv2_NRTidal with $f_{\max} = 2$ kHz. The results of EOB and PN models are in overall good agreement if $f_{\max} = 1$ kHz is employed and also with independent analysis of AT2017gfo [25]. We note that parallelization methods are key for the PE of BNS signals associated to long data segments $\gtrsim 100$ s. The BAJES runs discussed in Sec. VII B were efficiently performed on 128 CPUs with total execution time of ~ 1 day.

Future work will present the validation and the application of BAJES to multimessenger analyses, including EM counterparts like kilonovae and γ -ray burst (e.g., [25,214]). Moreover, we are implementing reduced-order quadrature [215–217] and the relative binning [201,218] in order to speed-up the likelihood evaluations in the GW studies. Inferences on the properties of neutron star matter will be supported with the inclusion of a parametrized EOS sampling method [219,220]. Moreover, future BAJES releases will include an extended set of nested samplers,

in particular algorithms based on machine learning (e.g., [221]) and efficiently parallelizable routines (e.g., [222]).

BAJES is publicly available at [223] and contributions from the community are welcome. The posterior samples presented in Sec. VII C and the configuration files to reproduce the runs are available on Zenodo [224].

ACKNOWLEDGMENTS

The authors would like to thank Walter Del Pozzo for useful discussions. M. B. and S. B. acknowledge support by the EU H2020 under ERC Starting Grant No. BinGraSp-714626. R. G. and M. B. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG) under Grant No. 406116891 within the Research Training Group RTG 2522/1. The computational experiments were performed on the ARA cluster at the Friedrich Schiller University Jena supported in part by DFG Grants No. INST 275/334-1 FUGG and No. INST 275/363-1 FUGG and ERC Starting Grant, No. BinGraSp-714626. This research made use of data obtained from the Gravitational Wave Open Science Center [80], a service of LIGO Laboratory, the LIGO Scientific Collaboration, and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique, the Italian Istituto Nazionale della Fisica Nucleare, and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

APPENDIX A: MCMC AND PTMCMC

A generic MCMC algorithm explores the parameter space moving forward region with increasing value of the probability and returns a set of independent samples representative of the target probability density. The MCMC samplers implemented in BAJES is based on EMCEE [63]: this routine represents a simple and complete implementation of a Metropolis-Hastings sampling that takes advantage of parallel chains.

The MCMC algorithm can be summarized as follows. An arbitrary number of chains, say n_{chain} , are initialized with as many random prior samples. For each chain, the last stored sample is evolved and a new sample θ^* is proposed according to predefined proposal methods (see Appendix C). The new sample θ^* is accepted with probability

$$\min \left[1, \frac{p(\theta^*|\mathbf{d}, H) q(\theta_i|\theta^*)}{p(\theta_i|\mathbf{d}, H) q(\theta^*|\theta_i)} \right], \quad (\text{A1})$$

where θ_i is the last sample of the chain and $q(\theta_i|\theta_j)$ is the proposal density function computed between θ_i and θ_j . This procedure is iterated for every chain of the ensemble and samples are collected during the exploration. Note that, according with prescription Eq. (A1), the probability of the proposed sample is not required to be strictly greater than that of the current sample. The initial exploration is called *burn-*

in, in which the chains randomly explore the surrounding prior volume. The algorithm spends these iterations in order to localize the maximum-probability region. After a certain amount of iterations, depending on the complexity of the parameter space, the chains can converge and the actual collection of posterior samples starts. Subsequently, when the algorithm reaches the stopping condition, the burn-in portion is removed, the samples from different chains are joined together, the ACL is computed in order to estimate the effective number of independent posterior samples, and the final set of samples is extracted from the joined set according with the value of ACL. The stopping criterion implemented in BAJES for the MCMC algorithms is defined by the requested number of output samples n_{out} ,

$$\frac{(i - n_{\text{burn}})n_{\text{chain}}}{\text{ACL}} \geq n_{\text{out}}, \quad (\text{A2})$$

where i is the current iteration, n_{chain} is the total number of chains, n_{burn} is the number of iterations required for burn-in, and ACL is computed on the set of post-burn-in samples.⁵

The MCMC algorithm disposes of a light-weighted settings and it is a fast and versatile algorithm. However, when the parameter space becomes large or the distributions have multimodalities, this method could have many issues, such as insufficient and inaccurate exploration of the parameter space; some chains could get trapped in local minima or the ensemble might not be able to reach convergence. These issues can be mitigated resorting to a large number of parallel chains or to specific proposal methods (see Appendix C). Moreover, given a set of posterior samples, it is possible to estimate the evidence using the approximation

$$p(\mathbf{d}|H) \approx \frac{1}{n_{\text{samples}}} \sum_{i=1}^{n_{\text{samples}}} p(\theta_i|\mathbf{d}, H), \quad (\text{A3})$$

where the index $i = 1, \dots, n_{\text{samples}}$ runs over the posterior samples. However, in general, Eq. (A3) is unable to perform accurate estimations of the evidence, since the MCMC algorithm is not designed to minutely inspect all the parameter space.

On the other hand, the parallel tempering MCMC (PTMCMC) [227–229] performs improved exploration of the parameter space and it provides a more accurate estimations of the evidence integral compared to standard MCMC techniques. The PTMCMC sampler implemented in BAJES is inspired by PTEMCEE [65]. The PTMCMC introduces an inverse temperature coefficient $\beta = 1/T \in [0, 1]$ in the computation of posterior distribution, such that

⁵We are planning to modify the current MCMC stopping condition implementing the Gelman-Rubin diagnostic test [225,226].

$$p_\beta(\boldsymbol{\theta}|\mathbf{d}, H) \propto [p(\mathbf{d}|\boldsymbol{\theta}, H)]^\beta p(\boldsymbol{\theta}|H). \quad (\text{A4})$$

The set of all chains is grouped in equally populated subensembles and a different value of β is associated to each tempered ensembles. The default β ladder is geometrically spaced in the range $[0, 1]$. The algorithm proceeds as the usual MCMC for every chain using the tempered posterior distribution Eq. (A4). For $T = 1$, the tempered posterior is identical to the original one and low-temperature chains will move toward regions with large likelihood values focusing on the estimation of the volume of the bulge. However, the contribution of the likelihood function is mitigated by increasing values of T , up to the limit $T \rightarrow \infty$, where the posterior is identically equal to the prior. Then, high-temperature chains will be able to freely explore the majority the prior support, inspecting the tails of the targeted posterior distribution and providing a good coverage of the entire prior volume. Furthermore, the algorithm proposes swaps between consecutive pairs of chains, received with acceptance

$$\min \left[1, \left(\frac{p(\mathbf{d}|\boldsymbol{\theta}_i, H)}{p(\mathbf{d}|\boldsymbol{\theta}_j, H)} \right)^{\beta_j - \beta_i} \right], \quad (\text{A5})$$

where $\boldsymbol{\theta}_i$ and β_i are, respectively, the last sample and the inverse temperature of the i th chain, and analogously for j . If the swap is accepted, the position of the two samples is exchanged in the different selected chains. This procedure allows the information of the high-temperature chains to propagate to the low-temperature ones and vice versa, improving the correlation between the samples. Another key feature of parallel tempering is that it satisfies the detailed balance condition [230], required for convergence of the MCMC chains.

Finally, the stopping criterion Eq. (A2) is estimated for the untempered chains; when it is satisfied, the sampler stops and reproduces the posterior samples using only the chains of the $T = 1$ subensemble. Furthermore, using the auxiliary coefficient β and thermodynamic integration [231,232], it is possible to write the evidence as

$$\log p(\mathbf{d}|H) = \int_0^1 E_\beta[\log p(\mathbf{d}|\boldsymbol{\theta}, H)] d\beta, \quad (\text{A6})$$

where the expectation value is estimated using the tempered posterior, i.e.,

$$E_\beta[f(\boldsymbol{\theta})] = \int_{\Theta} f(\boldsymbol{\theta}) [p(\mathbf{d}|\boldsymbol{\theta}, H)]^\beta p(\boldsymbol{\theta}|H) d\boldsymbol{\theta}. \quad (\text{A7})$$

Equation (A6) can be estimated through numerical integration. The terms $E_\beta[\log p(\mathbf{d}|\boldsymbol{\theta}, H)]$ are estimated over the initial β ladder applying Eq. (A3) to the tempered posterior samples and the integral is approximated using the trapezoidal rule. The PTMCMC represents an improved

version of a standard MCMC technique that aims to provide much accurate estimations of the evidence. However, the accuracy of the estimation Eq. (A6) strongly depends on the number of employed temperatures: in complex situations, the total number of chains needed to accurately estimate the evidence could overcome the number of available processes, affecting the efficiency of the sampler [233].

APPENDIX B: NESTED SAMPLING

The nested sampling [50,56] is a finely designed Bayesian technique designed to accurately estimate the evidence integral and, nevertheless, it provides a set of posterior samples as final product of the algorithm. The strength of this technique is the capability to succeed even in cases of high-dimensional parameter space or multimodal distributions. Nested sampling computes the evidence identifying nested isoprobability contours and estimating the amount of prior volume enclosed by each level. The main advantage of this technique is the reduction of the multidimensional integral Eq. (2) to a one-dimensional problem [234,235], introducing the variable

$$X(\lambda) = \int_{p(\mathbf{d}|\boldsymbol{\theta}, H) > \lambda} p(\boldsymbol{\theta}|H) d\boldsymbol{\theta}. \quad (\text{B1})$$

The quantity $X(\lambda)$ is usually labeled as *prior mass* and it is the cumulative prior volume covering all likelihood values greater than λ . The prior mass takes values in the range $[0, 1]$, where $X = 1$ corresponds to the entire prior volume. Then, we can rewrite the likelihood as function of the prior mass, i.e., $p(\mathbf{d}|X(\lambda), H) = \lambda$, from which follows

$$p(\mathbf{d}|H) = \int_0^1 p(\mathbf{d}|X, H) dX. \quad (\text{B2})$$

Equation (B2) has a further advantage: by definition, the likelihood $p(\mathbf{d}|X, H)$ is a monotonic decreasing function of X . Then, for $X \rightarrow 0$, the likelihood tends to its maximum value. Accomplishing the transformation $\boldsymbol{\theta} \rightarrow X$ involves dividing the unit prior mass range into small bins and sorting them by likelihood.

A standard nested sampling routine requires an input number of live points n_{live} and a real positive number ζ representing the final tolerance of the computation. The live points are samples of the parameter space that are evolved during the routine: starting from a set of n_{live} initial samples (usually extracted from the prior distribution), the live point with lowest likelihood value, say $\boldsymbol{\theta}_i$, is discarded and replaced with a new point $\boldsymbol{\theta}^*$ extracted from the prior distribution that satisfies the relation $p(\mathbf{d}|\boldsymbol{\theta}^*, H) > p(\mathbf{d}|\boldsymbol{\theta}_i, H)$. The new point $\boldsymbol{\theta}^*$ is usually proposed using internal MCMC routines (see Appendix C). The procedure is repeated taking the lowest-likelihood live point at every iteration, such that the algorithm starts inspecting the entire prior volume ($X_0 = 1$), and it moves toward lower value of the prior mass,

$$0 < X_i < \dots < X_2 < X_1 < X_0 = 1, \quad (\text{B3})$$

up to the most likely value(s), where the majority of the likelihood volume is located. At the n th iteration, the evidence is approximated from Eq. (B2) using trapezoidal rule,

$$p_n(\mathbf{d}|H) \approx \frac{1}{2} \sum_{i=1}^n (X_{i-1} - X_{i+1}) p(\mathbf{d}|\boldsymbol{\theta}_i, H), \quad (\text{B4})$$

where X_i is estimated with the expectation value [235]

$$E[X_i] = e^{-i/n_{\text{live}}}. \quad (\text{B5})$$

From Eq. (B5), we can deduce that the average volume occupied by a live point corresponds to the n_{live} th part of the current prior mass. Then, increasing n_{live} , the sampling will perform a finer grained inspection of the prior volume. In the nested sampling context, the i th extracted sample is taken as representative element of the respective isoprobability level of likelihood and, since the algorithm accepts strictly increasing likelihood values, it ensures that each level is nested in the previous one. Then, the overall evidence is computed summing all the likelihood contributions from each nested level weighted on the expected difference in prior mass, according with Eq. (B5). The specific stopping condition depends on the requested sampler. In general, if the algorithm converged to the global maximum-likelihood value, at the n th iteration, the evidence is expected to vary at most of

$$\Delta_n = \max_{i \leq n} [p(\mathbf{d}|\boldsymbol{\theta}_i, H)] \cdot X_n, \quad (\text{B6})$$

where $\max_{i \leq n} [p(\mathbf{d}|\boldsymbol{\theta}_i, H)]$ is the maximum likelihood discovered up to the n th iteration. Then, the general stopping criterion requires that the estimated evidence is not expected to change more than a factor e^ζ , i.e.,

$$\zeta \geq \log \left(1 + \frac{\Delta_n}{p_n(\mathbf{d}|H)} \right). \quad (\text{B7})$$

When the stopping condition is satisfied, the sampler stops and it includes the contributions of the remaining live points to the overall evidence. Then, the posterior distribution can be reconstructed by the chain of collected samples, weighting each point according with its probability distribution

$$p(\boldsymbol{\theta}_k|\mathbf{d}, H) \approx \frac{(X_{k-1} - X_{k+1}) p(\mathbf{d}|\boldsymbol{\theta}_k, H)}{2 p(\mathbf{d}|H)}, \quad (\text{B8})$$

where the index k runs over the extracted samples [236].

The nested sampling routine offers a much better architecture for evidence estimation than MCMC techniques. In general, the estimated log evidence carries a statistical uncertainty inversely proportional to n_{live} due to

the marginalization over the prior mass; while, numerical errors are dominated by the use of point estimates and by the length of the MCMC subchains n_{MCMC} used to propose new samples, as shown in Ref. [6]. This inefficiency can be suppressed estimating the ACL of the MCMC subchains and proposing a new sample independent of the previous one. Note that also the estimation of the posterior samples Eq. (B8) is affected by statistical and numerical uncertainty [71].

The CPNest [86] software represents an exemplary implementation of a standard nested sampling: the code is designed to be nicely interfaced with user-defined models and its sampling methods can be easily customized. On the other hand, DYNESTY [87] takes advantage of flexible bounding methods [237,238] that aim to define isoprobability contours in order to exclude least likely regions of the parameter space improving the robustness of the algorithm. Moreover, DYNESTY provides an implementation of dynamic nested sampling [239]: this technique allocates an adaptive number of live points at each iteration i , i.e., $n_{\text{live}} \equiv n_{\text{live}}(i)$. Since the change in prior volume at a given iteration depends on the number of live points, as shown in Eq. (B5), the possibility of varying n_{live} gives the algorithm the freedom to control the effective resolution of the sampling as a function of prior mass, adapting it to the shape of the posterior in real time and improving the evaluation of the posterior density. Since the architecture of the dynamic nested sampling differs from the standard, it requires modified methods in order to compute the evidence and estimate the posterior. By default, the importance of each sample for the evidence computation is proportional to the amount of the posterior density enclosed in the prior mass probed by that point.

APPENDIX C: PROPOSAL METHODS

The exploration of the parameter space is defined by proposal methods which aim to move a sample toward a more likely position in the parameter space independent from the previous. The efficiency of the proposals determines the rate of acceptance of new samples and it affects the final ACL and subsequently the efficiency of the whole sampler. It follows that these tools are fundamental for the chains progress and a generic, broad, and varied combination of proposal methods is needed to accurately inspect the entire parameter space. For this reason, BAJES implements an assorted combination of proposal methods.

Before discussing the specific proposals implemented in BAJES, we observe that a generic proposal method requires the introduction of a proposal distribution q in order to satisfy the detailed balance. A proposal distribution $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_i)$ quantifies the probability of proposing $\boldsymbol{\theta}^*$ given $\boldsymbol{\theta}_i$. A symmetric proposal is such that the proposed point is fully independent from the initial sample, i.e., $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_i) \propto 1$ for every $(\boldsymbol{\theta}^*, \boldsymbol{\theta}_i)$. The interested reader might look at Refs. [240,241] for details on Markovian process.

- (i) *Prior proposal*: A new point is extracted from the prior distribution generating a sample uniformly distributed over a hypercube and projecting it in the current parameter space according with the prescribed prior. This method could show low acceptance on long timescales, especially for complex posterior distribution. However, it can improve the exploration of the parameter space and it does not require expensive computations.
- (ii) *Stretch proposal*: This method is introduced in Ref. [242] and it make use of an auxiliary random sample θ_a extracted from the history of the chains. A new sample θ^* is proposed from the initial position θ_i according with

$$\theta^* = \theta_a + \xi(\theta_i - \theta_a), \quad (\text{C1})$$

where ξ is a scale factor. As pointed out in Ref. [243], this proposal is symmetric if the probability density of scaling factor $g(\xi)$ satisfies the condition $g(1/\xi) = \xi g(\xi)$. The BAJES implementation adopts the settings presented in Ref. [63]. In order to satisfy the detailed balance, this proposal method requires an acceptance factor,

$$\frac{q(\theta_i|\theta^*)}{q(\theta^*|\theta_i)} = \xi^{n_{\text{dim}}-1}, \quad (\text{C2})$$

where n_{dim} is the number of dimensions of the parameter space. Equation (C2) is computed conditioning of the target distribution on the trajectory described by Eq. (C1). The method shows a good adaptation to arbitrary shapes of the distributions [242] but it might become inefficient for multimodal cases.

- (iii) *Random walk proposal*: This method is introduced in Ref. [242]. The new sample θ^* is extracted from a multivariate normal distribution, centered in the initial sample θ_i and with covariance defined using a subset of N auxiliary points randomly extracted from the history of the chains. By default, BAJES walk proposal uses $N = 25$. This symmetric method is efficient with unimodal distributions and it can improve the correlation between the samples. However, it becomes inefficient in case of complex posterior distribution since it is not capable to arbitrarily adapt its shape.
- (iv) *Replacement proposal*: This method is introduced in Ref. [242]. A subset of N auxiliary points $\bar{\Theta} \equiv \{\bar{\theta}_k\}$ is randomly chosen from the history of the chains and it is used to identify a probability distribution $p(\theta|\bar{\Theta})$ from which the new samples will be extracted. The idea is to estimate $p(\theta|\bar{\Theta})$ such that is it capable to approximate the target distribution, increasing the acceptance. Moreover, the estimation

can be refined and adapted during the exploration of the parameter space. In order to estimate $p(\theta|\bar{\Theta})$, the replacement proposal implemented in BAJES employs a Gaussian kernel density estimation with $N = 25$. However, this method does not access to unexplored regions of the parameter space, with the possibility of leading to highly correlated chains. Furthermore, this method is not symmetric and the proposal distribution is described by $p(\theta|\bar{\Theta})$.

- (v) *Differential evolution proposal*: This method is introduced in Ref. [244] and it aims to solve problems due to multimodal distributions using a differential move based on the information on the explored samples. Two auxiliary random samples $\{\theta_{a,b}\}$ are extracted from the from the history of the chains and a new sample θ^* is proposed from θ_i as

$$\theta^* = \theta_i + \gamma(\theta_a - \theta_b), \quad (\text{C3})$$

where γ is a scale factor whose value is randomly extracted when a new sample is proposed. The differential evolution proposal of BAJES assign $\gamma = 1$ with 50% of probability in order to improve the mixing between different modes. The remaining 50% of the time the scale factor is extracted from a normal distribution such that $\gamma \sim N(0, 2.38/\sqrt{2n_{\text{dim}}})$, where n_{dim} is the dimension of the parameter space. This choice has been proved to increase the acceptance of the algorithm [245,245]. In general, differential evolution is capable to capture linear correlations and improve mixing between different modes; however, it can perform poorly in more complicated scenarios.

- (vi) *Eigenvector proposal*: This method computes the covariance from history of the chains and estimates the relative eigenvectors. Then, the new point is proposed moving the initial sample along a random eigenvector of the covariance with a scale prescribed by the respective eigenvalue. As shown in Ref. [7], this method can improve the efficiency of the sampler and decrease the correlation of the chains.
- (vii) *Ensemble slice proposal*: This method has been introduced in Ref. [246] and it represents an ensemble-based improvement of the standard slice proposal [247]. Let us call θ_i and p_i , respectively, the initial sample and its probability. The method extracts a value $y \sim U(0, p_i)$ and estimates a direction in the parameter space η , resorting to the information of the ensemble samples and using differential and Gaussian estimations. Then, the initial sample is moved along the slide defined by the direction η and a new point θ^* is proposed when the associated probability value p^* is greater than y . With respect to standard slice sampling [247], this method takes advantage of adaptive scale factors refined during the evolution of the sampler that

increase the efficiency of the proposal and drastically reduce the correlation between the collected samples. On the other hand, the ensemble slice proposal requires multiple likelihood evaluations; then, when the likelihood is computationally expensive, it might affect the computational cost of the whole algorithm.

- (viii) *GW-targeted proposal*: As discussed in Refs. [6,7], generic posterior distributions of GW signal parameters show many multimodalities and large correlations between different parameters. Moreover, its shape is usually elongated and nonregular. Then, in order to properly and efficiently explore the parameter space, it is useful to inform the proposal method with known structures expected in the posterior distributions and it is possible to take advantage of the determined analytical dependencies on the extrinsic parameters. The GW-targeted proposal implemented in BAJES explores most of the GW specific methods introduced in Refs. [6,7] such as sky reflection, Gibbs sampling for the luminosity distance, and specific methods to explore the phase polarization and the distance-inclination correlations.

APPENDIX D: SIMPLE EXAMPLE

In this section, we show a simple example of Bayesian inference performed with the samplers implemented in BAJES. We use a three-dimensional parameter space $\theta = \{\theta_1, \theta_2, \theta_3\}$ bounded to $\theta_i \in [-8, +8]$ for $i = 1, 2, 3$ with a uniformly distributed prior distribution. For the sake of simplicity, we employ a fully analytical likelihood function; however, we include multimodalities and non-regular shapes in order to test the behavior of the samplers. Introducing the auxiliary variables

$$r_{\pm} = \sqrt{\theta_1^2 + (\theta_2 \mp 2)^2}, \quad \zeta = \theta_3 - \sqrt{1 + \theta_2^2}, \quad (\text{D1})$$

we define the likelihood as

$$p(d|\theta, H) = [e^{-5(r_+ - 2)^2} + e^{-5(r_- - 2)^2}]e^{-5\zeta^2}. \quad (\text{D2})$$

To give an idea, the isoprobability contours described by this function are roughly similar to the union of two toroidal surfaces where the second is identical to the first except for a rotation of $\pi/2$. The likelihood function in Eq. (D2) can be numerically integrated using the quadrature rule, resulting in the evidence $\log p(d|H) \approx -5.5583$ with an error of the order of $O(10^{-5})$.

In order to infer the described model, BAJES provides a user-friendly and simple-to-use interface for generic Bayesian inference. In order to define the prior distributions, it is sufficient to write a prior configuration file specifying the name of the parameters and the bounds. For

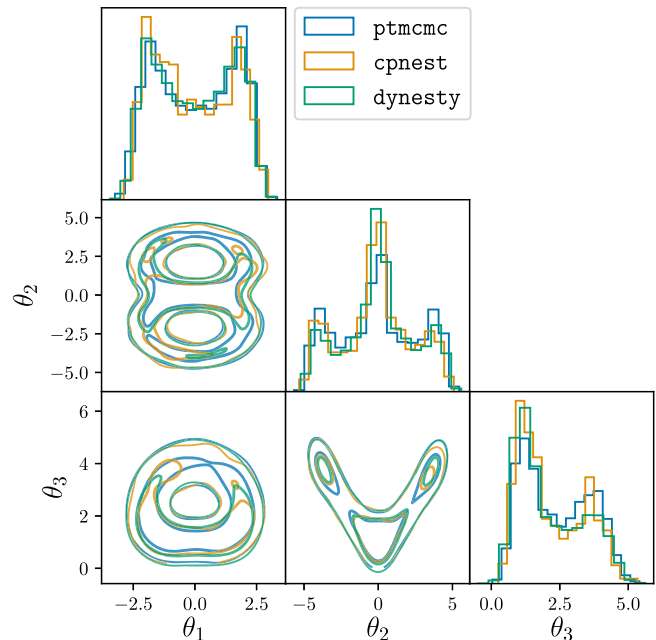


FIG. 23. Posterior distribution for the parameters $\{\theta_1, \theta_2, \theta_3\}$ discussed in Appendix D. The blue lines correspond to the results obtained with PTMCMC sampler, while the yellow and green lines refer, respectively, to CPNest and DYNESTY results. The marginalized contours are the 50% and 90% credible regions.

the case discussed above, we can write the following PRIOR.INI file.

Then, the likelihood function can be written in an auxiliary PYTHON file defining a LOG_LIKE method. This method will be imported by the BAJES routine and used to determine the log-likelihood function for each sample. We observe that the only argument of the customized LOG_LIKE method has to be a dictionary whose keywords are specified by the prior file. This procedure easily allows the user to make use of the BAJES inference introducing arbitrary external data or packages. Following our example, we write the following pseudo code.

Once the model is defined, the PE job can be submitted with the command `PYTHON -M BAJES -P PRIOR.INI -L LIKE.PY`.⁶ For our exercise, we employ three samplers: the PTMCMC, the nested sampling with CPNest, and the dynamic nested sampling with DYNESTY. The PTMCMC algorithm estimated a log-evidence equal to -6.4 ± 5.0 , where the reported uncertainty is the standard deviation. The estimation agrees with the numerical result; however, its uncertainty is of the same order of the measurement. This reflects the inability of MCMC methods to meticulously integrate the features of the targeted parameter space. On the other hand, CPNest estimated a log-evidence equal to

⁶The full list of input arguments can be visualized with the command `PYTHON -M BAJES-HELP`.

Algorithm 1. Prior.ini.

```

1: [x1]
2: MIN= -8
3: MAX= +8
4:
5: [x2]
6: MIN= -8
7: MAX= +8
8:
9: [x3]
10: MIN= -8
11: MAX= +8

```

Algorithm 2. LIKE.PY.

```

1: FROM MATH IMPORT SQRT, EXP, LOG
2:
3: DEF LOG_LIKE(P):
4:   RP = SQRT(P['x1']**2 + (P['x2'] - 2)**2)
5:   RM = SQRT(P['x1']**2 + (P['x2'] + 2)**2)
6:   ZT = P['x3'] - SQRT(1 + P['x2']**2)
7:   G1 = EXP(- 5*(RP - 2)**2)
8:   G2 = EXP(-5*(RM - 2)**2)
9:   RETURN LOG(G1 + G2) - 5*ZT**2

```

-5.50 ± 0.09 and the dynamic nested sampler of DYNESTY returned the value of -5.58 ± 0.14 . These results highlight the strength of the nested sampling with respect to MCMC

techniques in the evidence evaluation. Figure 23 shows the marginalized posterior distributions extracted from the posterior samples.

-
- [1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
 - [2] W. Del Pozzo, C. Berry, A. Ghosh, T. Haines, L. Singer, and A. Vecchio, *Mon. Not. R. Astron. Soc.* **479**, 601 (2018).
 - [3] D. Wang, *Int. J. Mod. Phys. D* **28**, 1950106 (2019).
 - [4] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), [arXiv:2010.14527](https://arxiv.org/abs/2010.14527) [*Phys. Rev. X* (to be published)].
 - [5] H. J. Horta, R. Volpi, D. Marinelli, and L. Malag, *Phys. Rev. D* **102**, 103509 (2020).
 - [6] J. Veitch and A. Vecchio, *Phys. Rev. D* **81**, 062003 (2010).
 - [7] J. Veitch *et al.*, *Phys. Rev. D* **91**, 042003 (2015).
 - [8] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Classical Quantum Gravity* **37**, 055002 (2020).
 - [9] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **100**, 104036 (2019).
 - [10] M. Breschi, R. O’Shaughnessy, J. Lange, and O. Birnholtz, *Classical Quantum Gravity* **36**, 245019 (2019).
 - [11] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), [arXiv:2010.14533](https://arxiv.org/abs/2010.14533).
 - [12] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **121**, 161101 (2018).
 - [13] B. P. Abbott *et al.* (LIGO Scientific, VINROUGE, Las Cumbres Observatory, DLT40, Virgo, 1M2H, and MASTER Collaborations), *Nature (London)* (2017), 10.1038/nature24471, [arXiv:1710.05835](https://arxiv.org/abs/1710.05835).
 - [14] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J.* **909**, 218 (2021).
 - [15] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), *Phys. Rev. Lett.* **119**, 161101 (2017).
 - [16] B. P. Abbott *et al.* (Virgo, Fermi-GBM, INTEGRAL, and LIGO Scientific Collaborations), *Astrophys. J.* **848**, L13 (2017).
 - [17] V. Savchenko *et al.*, *Astrophys. J.* **848**, L15 (2017).
 - [18] E. Pian *et al.*, *Nature (London)* **551**, 67 (2017).
 - [19] S. J. Smartt *et al.*, *Nature (London)* **551**, 75 (2017).
 - [20] N. R. Tanvir *et al.*, *Astrophys. J.* **848**, L27 (2017).
 - [21] M. Tanaka *et al.*, *Publ. Astron. Soc. Jpn.* 10.1093/pasj/psx121, [arXiv:1710.05850](https://arxiv.org/abs/1710.05850).
 - [22] V. A. Villar *et al.*, *Astrophys. J.* **851**, L21 (2017).
 - [23] D. Radice and L. Dai, *Eur. Phys. J. A* **55**, 50 (2019).
 - [24] T. Dietrich, M. W. Coughlin, P. T. Pang, M. Bulla, J. Heinzel, L. Issa, I. Tews, and S. Antier, *Science* **370**, 1450 (2020).
 - [25] M. Breschi, A. Perego, S. Bernuzzi, W. Del Pozzo, V. Nedora, D. Radice, and D. Vescovi, [arXiv:2101.01201](https://arxiv.org/abs/2101.01201).
 - [26] LIGO Scientific Collaboration, LIGO Algorithm Library —LALSuite, free software (GPL) (2018).
 - [27] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 115012 (2015).
 - [28] F. Acernese *et al.* (VIRGO Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
 - [29] C. Pankow, P. Brady, E. Ochsner, and R. O’Shaughnessy, *Phys. Rev. D* **92**, 023002 (2015).
 - [30] J. Lange, R. O’Shaughnessy, and M. Rizzo, [arXiv:1805.10457](https://arxiv.org/abs/1805.10457).
 - [31] G. Ashton *et al.*, *Astrophys. J. Suppl.* **241**, 27 (2019).
 - [32] C. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz, and V. Raymond, *Publ. Astron. Soc. Pac.* **131**, 024503 (2019).
 - [33] NASA’s High Energy Astrophysics Science Archive Research Center (HEASARC), HEASoft: Unified release of FTOOLS and XANADU (2014), [ascl:1408.004](https://arxiv.org/abs/1408.004).
 - [34] NASA’s HEASARC, HEASARC software, <https://heasarc.gsfc.nasa.gov/docs/software.html>.
 - [35] M. Olamaie, F. Feroz, K. J. B. Grainge, M. P. Hobson, J. S. Sanders, and R. D. E. Saunders, *Mon. Not. R. Astron. Soc.* **446**, 1799 (2015).
 - [36] M. G. Aartsen *et al.* (IceCube Collaboration), *Astrophys. J.* **893**, 102 (2020).

- [97] C. Bond, D. Brown, A. Freise, and K. A. Strain, *Living Rev. Relativity* **19** (2017).
- [98] M. W. Coughlin *et al.*, *Phys. Rev. D* **97**, 102007 (2018).
- [99] P. Welch, *IEEE Trans. Audio Electroacoust.* **15**, 70 (1967).
- [100] J. A. R. Blais, *J. Geodes.* **70**, 349 (1996).
- [101] N. J. Cornish and T. B. Littenberg, *Classical Quantum Gravity* **32**, 135012 (2015).
- [102] S. A. Usman *et al.*, *Classical Quantum Gravity* **33**, 215004 (2016).
- [103] A. H. Nitz, T. Dal Canton, D. Davis, and S. Reyes, *Phys. Rev. D* **98**, 024050 (2018).
- [104] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, C. Capano, L. Pekowsky, T. Dent, A. R. Williamson, G. S. Davies, S. De, M. Cabero, B. Machenschalk, P. Kumar, S. Reyes, D. Macleod, dfinstad, F. Pannarale, T. Massinger, S. Kumar, M. Tpai, L. Singer, S. Khan, S. Fairhurst, A. Nielsen, S. Singh, shasvath, and B. U. V. Gadre, gwastro/pycbc: 1.18.0 release of pycbc (2021).
- [105] B. Abbott *et al.* (KAGRA, LIGO Scientific, and VIRGO Collaborations), *Living Rev. Relativity* **21**, 3 (2018).
- [106] K. L. Dooley (LIGO Scientific Collaboration), *J. Phys. Conf. Ser.* **610**, 012015 (2015).
- [107] T. Akutsu *et al.* (KAGRA Collaboration), *Nat. Astron.* **3**, 35 (2019).
- [108] V. Raymond and W. Farr, [arXiv:1402.0053](https://arxiv.org/abs/1402.0053).
- [109] A. H. Nitz, C. Capano, A. B. Nielsen, S. Reyes, R. White, D. A. Brown, and B. Krishnan, *Astrophys. J.* **872**, 195 (2019).
- [110] C. Pankow, E. A. Chase, S. Coughlin, M. Zevin, and V. Kalogera, *Astrophys. J. Lett.* **854**, L25 (2018).
- [111] H. Luck *et al.*, *J. Phys. Conf. Ser.* **228**, 012012 (2010).
- [112] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto (KAGRA Collaboration), *Phys. Rev. D* **88**, 043007 (2013).
- [113] T. Akutsu *et al.* (KAGRA Collaboration), [arXiv:2005.05574](https://arxiv.org/abs/2005.05574).
- [114] M. Punturo, M. Abernathy, F. Acernese, B. Allen, N. Andersson *et al.*, *Classical Quantum Gravity* **27**, 194002 (2010).
- [115] S. Hild *et al.*, *Classical Quantum Gravity* **28**, 094013 (2011).
- [116] D. Chiaramello and A. Nagar, *Phys. Rev. D* **101**, 101501 (2020).
- [117] A. Nagar, P. Rettegno, R. Gamba, and S. Bernuzzi, *Phys. Rev. D* **103**, 064013 (2021).
- [118] M. Breschi, S. Bernuzzi, F. Zappa, M. Agathos, A. Perego, D. Radice, and A. Nagar, *Phys. Rev. D* **100**, 104029 (2019).
- [119] L. Blanchet, T. Damour, and G. Schafer, *Mon. Not. R. Astron. Soc.* **242**, 289 (1990).
- [120] G. Faye, S. Marsat, L. Blanchet, and B. R. Iyer, *Classical Quantum Gravity* **29**, 175004 (2012).
- [121] M. Levi and J. Steinhoff, *J. Cosmol. Astropart. Phys.* **01** (2016) 008.
- [122] M. Levi and J. Steinhoff, *J. Cosmol. Astropart. Phys.* **01** (2016) 011.
- [123] C. K. Mishra, K. G. Arun, and B. R. Iyer, *Phys. Rev. D* **91**, 084040 (2015).
- [124] M. Levi and J. Steinhoff, [arXiv:1607.04252](https://arxiv.org/abs/1607.04252).
- [125] F. Messina, R. Dudi, A. Nagar, and S. Bernuzzi, *Phys. Rev. D* **99**, 124051 (2019).
- [126] J. Vines, E. E. Flanagan, and T. Hinderer, *Phys. Rev. D* **83**, 084051 (2011).
- [127] T. Damour, A. Nagar, and L. Villain, *Phys. Rev. D* **85**, 123007 (2012).
- [128] Q. Henry, G. Faye, and L. Blanchet, *Phys. Rev. D* **102**, 044033 (2020).
- [129] A. Nagar, F. Messina, P. Rettegno, D. Bini, T. Damour, A. Geralico, S. Akcay, and S. Bernuzzi, *Phys. Rev. D* **99**, 044007 (2019).
- [130] A. Buonanno and T. Damour, *Phys. Rev. D* **59**, 084006 (1999).
- [131] A. Buonanno and T. Damour, *Phys. Rev. D* **62**, 064015 (2000).
- [132] T. Damour, P. Jaranowski, and G. Schaefer, *Phys. Rev. D* **62**, 084011 (2000).
- [133] T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
- [134] T. Damour, P. Jaranowski, and G. Schäfer, *Phys. Rev. D* **78**, 024009 (2008).
- [135] T. Damour, P. Jaranowski, and G. Schfer, *Phys. Rev. D* **91**, 084024 (2015).
- [136] D. Bini, T. Damour, and A. Geralico, *Phys. Rev. Lett.* **123**, 231104 (2019).
- [137] D. Bini, T. Damour, and A. Geralico, *Phys. Rev. D* **102**, 024062 (2020).
- [138] D. Bini, T. Damour, and A. Geralico, *Phys. Rev. D* **102**, 024061 (2020).
- [139] A. Nagar *et al.*, *Phys. Rev. D* **98**, 104052 (2018).
- [140] T. Damour and A. Nagar, *Phys. Rev. D* **90**, 044018 (2014).
- [141] A. Nagar, G. Riemenschneider, and G. Pratten, *Phys. Rev. D* **96**, 084045 (2017).
- [142] D. Bini and T. Damour, *Phys. Rev. D* **90**, 124037 (2014).
- [143] S. Bernuzzi, A. Nagar, T. Dietrich, and T. Damour, *Phys. Rev. Lett.* **114**, 161103 (2015).
- [144] S. Akcay, S. Bernuzzi, F. Messina, A. Nagar, N. Ortiz, and P. Rettegno, *Phys. Rev. D* **99**, 044051 (2019).
- [145] A. Nagar, G. Pratten, G. Riemenschneider, and R. Gamba, *Phys. Rev. D* **101**, 024041 (2020).
- [146] A. Nagar, G. Riemenschneider, G. Pratten, P. Rettegno, and F. Messina, *Phys. Rev. D* **102**, 024077 (2020).
- [147] S. Akcay, R. Gamba, and S. Bernuzzi, *Phys. Rev. D* **103**, 024014 (2021).
- [148] A. Nagar and P. Rettegno, *Phys. Rev. D* **99**, 021501 (2019).
- [149] R. Gamba, S. Bernuzzi, and A. Nagar, [arXiv:2012.00027](https://arxiv.org/abs/2012.00027).
- [150] TEOBResumS code, https://bitbucket.org/eob_ihes/teobresums/src/master/.
- [151] K. Takami, L. Rezzolla, and L. Baiotti, *Phys. Rev. D* **91**, 064001 (2015).
- [152] J. A. Clark, A. Bauswein, N. Stergioulas, and D. Shoemaker, *Classical Quantum Gravity* **33**, 085003 (2016).
- [153] D. Radice, A. Perego, S. Bernuzzi, and B. Zhang, *Mon. Not. R. Astron. Soc.* **481**, 3670 (2018).
- [154] D. Radice, S. Bernuzzi, W. Del Pozzo, L. F. Roberts, and C. D. Ott, *Astrophys. J.* **842**, L10 (2017).
- [155] D. Radice, A. Perego, F. Zappa, and S. Bernuzzi, *Astrophys. J.* **852**, L29 (2018).
- [156] F. Zappa, S. Bernuzzi, D. Radice, A. Perego, and T. Dietrich, *Phys. Rev. Lett.* **120**, 111101 (2018).

- [157] M. Agathos, F. Zappa, S. Bernuzzi, A. Perego, M. Breschi, and D. Radice, *Phys. Rev. D* **101**, 044006 (2020).
- [158] S. Bernuzzi *et al.*, *Mon. Not. R. Astron. Soc.* **497**, 1488 (2020).
- [159] Stefano Schmidt and Walter Del Pozzo, mlgw, <https://pypi.org/project/mlgw/> (2020).
- [160] S. Schmidt, M. Breschi, R. Gamba, G. Pagano, P. Rettengo, G. Riemenschneider, S. Bernuzzi, A. Nagar, and W. Del Pozzo, *Phys. Rev. D* **103**, 043020 (2021).
- [161] A. Bohé *et al.*, *Phys. Rev. D* **95**, 044028 (2017).
- [162] J. Blackman, S. Field, C. Galley, V. Varma, and K. Barkett, gwsurrogate, <https://pypi.org/project/gwsurrogate/>.
- [163] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio, *Phys. Rev. X* **4**, 031006 (2014).
- [164] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev. Research* **1**, 033015 (2019).
- [165] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, *Phys. Rev. D* **96**, 024058 (2017).
- [166] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev. D* **99**, 064045 (2019).
- [167] K. Barkett, Y. Chen, M. A. Scheel, and V. Varma, *Phys. Rev. D* **102**, 024031 (2020).
- [168] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044006 (2016).
- [169] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044007 (2016).
- [170] T. Dietrich, S. Bernuzzi, and W. Tichy, *Phys. Rev. D* **96**, 121501 (2017).
- [171] LIGO Scientific Collaboration, LIGO Algorithm Library—LALSuite, free software (GPL) (2018).
- [172] K. K. Ng, S. Vitale, A. Zimmerman, K. Chatziioannou, D. Gerosa, and C.-J. Haster, *Phys. Rev. D* **98**, 083007 (2018).
- [173] N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A6 (2020).
- [174] J. Veitch and W. Del Pozzo, Analytic marginalisation of phase parameter, <https://dcc.ligo.org/LIGO-T1300326>.
- [175] W. M. Farr, Marginalisation of the time parameter in gravitational wave parameter estimation, <https://dcc.ligo.org/LIGO-T1400460>.
- [176] S. Vitale, W. Del Pozzo, T. G. Li, C. Van Den Broeck, I. Mandel, B. Aylott, and J. Veitch, *Phys. Rev. D* **85**, 064034 (2012).
- [177] B. Abbott *et al.* (LIGO Scientific Collaboration), *Phys. Rev. D* **95**, 062003 (2017).
- [178] C. Cahillane, J. Betzwieser, D. A. Brown, E. Goetz, E. D. Hall, K. Izumi, S. Kandhasamy, S. Karki, J. S. Kissel, G. Mendell, R. L. Savage, D. Tuyenbayev, A. Urban, A. Viets, M. Wade, and A. J. Weinstein, *Phys. Rev. D* **96**, 102001 (2017).
- [179] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **35**, 205004 (2018).
- [180] T. B. Littenberg, M. Coughlin, B. Farr, and W. M. Farr, *Phys. Rev. D* **88**, 084044 (2013).
- [181] G. M. Harry (LIGO Scientific Collaboration), *Classical Quantum Gravity* **27**, 084006 (2010).
- [182] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Classical Quantum Gravity* **34**, 104002 (2017).
- [183] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), *Phys. Rev. Lett.* **116**, 241102 (2016).
- [184] L. Wade, J. D. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, *Phys. Rev. D* **89**, 103012 (2014).
- [185] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 011001 (2019).
- [186] R. Dudi, F. Pannarale, T. Dietrich, M. Hannam, S. Bernuzzi, F. Ohme, and B. Bruegmann, *Phys. Rev. D* **98**, 084061 (2018).
- [187] A. Samajdar and T. Dietrich, *Phys. Rev. D* **100**, 024046 (2019).
- [188] R. Gamba, M. Breschi, S. Bernuzzi, M. Agathos, and A. Nagar, [arXiv:2009.08467](https://arxiv.org/abs/2009.08467) [*Phys. Rev. D* (to be published)].
- [189] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), *Phys. Rev. Lett.* **119**, 141101 (2017).
- [190] S. Bernuzzi, T. Dietrich, and A. Nagar, *Phys. Rev. Lett.* **115**, 091101 (2015).
- [191] A. Bauswein, H. Janka, K. Hebeler, and A. Schwenk, *Phys. Rev. D* **86**, 063001 (2012).
- [192] A. Bauswein, N. Stergioulas, and H.-T. Janka, *Phys. Rev. D* **90**, 023002 (2014).
- [193] A. Bauswein, N.-U.F. Bastian, D. B. Blaschke, K. Chatziioannou, J. A. Clark, T. Fischer, and M. Oertel, *Phys. Rev. Lett.* **122**, 061102 (2019).
- [194] I. Romero-Shaw *et al.*, *Mon. Not. R. Astron. Soc.* **499**, 3295 (2020).
- [195] LIGO Scientific and Virgo Collaborations, Power spectral densities (PSD) release for GWTC-1, <https://dcc.ligo.org/LIGO-P1900040/public>.
- [196] LIGO Scientific and Virgo Collaborations, Calibration uncertainty envelope release for GWTC-1, <https://dcc.ligo.org/LIGO-P1900011/public>.
- [197] T. D. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), *Phys. Rev. X* **6**, 041014 (2016).
- [198] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), *Astrophys. J.* **818**, L22 (2016).
- [199] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 221101 (2016); **121**, 129902 (E) (2018).
- [200] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), [arXiv:1810.02581](https://arxiv.org/abs/1810.02581).
- [201] L. Dai, T. Venumadhav, and B. Zackay, [arXiv:1806.08793](https://arxiv.org/abs/1806.08793).
- [202] T. Narikawa, N. Uchikata, K. Kawaguchi, K. Kiuchi, K. Kyutoku, M. Shibata, and H. Tagoshi, *Phys. Rev. Research* **1**, 033055 (2019).
- [203] X. Jiménez-Forteza, D. Keitel, S. Husa, M. Hannam, S. Khan, and M. Pürrer, *Phys. Rev. D* **95**, 064024 (2017).
- [204] T. Dal Canton *et al.*, *Phys. Rev. D* **90**, 082004 (2014).
- [205] T. Dal Canton, A. P. Lundgren, and A. B. Nielsen, *Phys. Rev. D* **91**, 062010 (2015).
- [206] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).
- [207] J. C. Bustillo, P. Laguna, and D. Shoemaker, *Phys. Rev. D* **95**, 104038 (2017).
- [208] C. Mills and S. Fairhurst, *Phys. Rev. D* **103**, 024042 (2021).

- [209] V. Varma, P. Ajith, S. Husa, J. C. Bustillo, M. Hannam, and M. Prerer, *Phys. Rev. D* **90**, 124004 (2014).
- [210] J. C. Bustillo, A. Bohé, S. Husa, A. M. Sintes, M. Hannam, and M. Pürrer, [arXiv:1501.00918](https://arxiv.org/abs/1501.00918).
- [211] S. Stevenson, C. P. L. Berry, and I. Mandel, *Mon. Not. R. Astron. Soc.* **471**, 2801 (2017).
- [212] W. M. Farr, S. Stevenson, M. Coleman Miller, I. Mandel, B. Farr, and A. Vecchio, *Nature (London)* **548**, 426 (2017).
- [213] V. Tiwari, S. Fairhurst, and M. Hannam, *Astrophys. J.* **868**, 140 (2018).
- [214] F. Hayes, I. S. Heng, J. Veitch, and D. Williams, *Astrophys. J.* **891**, 124 (2020).
- [215] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, and M. Tiglio, *Phys. Rev. Lett.* **114**, 071104 (2015).
- [216] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Prerer, V. Raymond, and P. Schmidt, *Phys. Rev. D* **94**, 044031 (2016).
- [217] S. Morisaki and V. Raymond, *Phys. Rev. D* **102**, 104020 (2020).
- [218] B. Zackay, L. Dai, and T. Venumadhav, [arXiv:1806.08792](https://arxiv.org/abs/1806.08792).
- [219] J. S. Read, B. D. Lackey, B. J. Owen, and J. L. Friedman, *Phys. Rev. D* **79**, 124032 (2009).
- [220] C. A. Raiethel, F. Ozel, and D. Psaltis, *Astrophys. J.* **831**, 44 (2016).
- [221] J. G. Albert, [arXiv:2012.15286](https://arxiv.org/abs/2012.15286).
- [222] J. Buchner, [arXiv:2101.09604](https://arxiv.org/abs/2101.09604).
- [223] <https://github.com/matteobreschi/bajes>
- [224] M. Breschi, R. Gamba, and S. Bernuzzi, *bajes: Bayesian inference of multimessenger astrophysical data, methods and application to gravitational-waves* (2021).
- [225] A. Gelman and D. B. Rubin, *Stat. Sci.* **7**, 457 (1992).
- [226] S. P. Brooks and A. Gelman, *J. Comput. Graph. Stat.* **7**, 434 (1998).
- [227] D. J. Earl and M. W. Deem, *Phys. Chem. Chem. Phys.* **7**, 3910 (2005).
- [228] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **57**, 2607 (1986).
- [229] R. M. Neal, *Stat. Comput.* **6**, 353 (1996).
- [230] M. Sambridge, *Geophys. J. Int.* **196**, 357 (2014).
- [231] P. M. Goggans and Y. Chi, *AIP Conf. Proc.* **707**, 59 (2004).
- [232] N. Lartillot and H. A. Philippe, *Syst. Biol.* **55**, 195 (2006).
- [233] N. Chopin and C. P. Robert, *Biometrika* **97**, 741 (2010).
- [234] N. Chopin and C. P. Robert, *Biometrika* **97**, 741 (2010).
- [235] M. Betancourt, A. Mohammad-Djafari, J.-F. Bercher, and P. Bessiere, 10.1063/1.3573613.
- [236] J. Buchner, *Stat. Comput.* **26**, 383 (2016).
- [237] J. Buchner, [arXiv:1407.5459](https://arxiv.org/abs/1407.5459).
- [238] J. Buchner, [arXiv:1707.04476](https://arxiv.org/abs/1707.04476).
- [239] E. Higson, W. Handley, M. Hobson, and A. Lasenby, *Stat. Comput.* **29**, 891 (2019).
- [240] W. K. Hastings, *Biometrika* **57**, 97 (1970).
- [241] G. O. Roberts and J. S. Rosenthal, *J. Appl. Probab.* **44**, 458 (2007).
- [242] J. Goodman and J. Weare, *Commun. Appl. Math. Comput. Sci.* **5** (2010).
- [243] J. A. Christen and C. Fox, *Bayesian Anal.* **5**, 263 (2010).
- [244] B. Nelson, E. B. Ford, and M. J. Payne, *Astrophys. J. Suppl. Ser.* **210**, 11 (2013).
- [245] G. O. Roberts and J. S. Rosenthal, *Stat. Sci.* **16**, 351 (2001).
- [246] M. Karamanis and F. Beutler, [arXiv:2002.06212](https://arxiv.org/abs/2002.06212).
- [247] R. M. Neal, *Ann. Stat.* **31**, 705 (2003).