

Leveraging rapid parameter estimates for efficient gravitational-wave Bayesian inference via posterior repartitioning

Metha Prathaban*

*Kavli Institute for Cosmology, Madingley Road, Cambridge CB3 0HA, United Kingdom
and Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom*

Charlie Hoy¹ and Michael J. Williams²

Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, United Kingdom



(Received 9 February 2026; revised 6 August 2026; accepted 13 August 2026; published 17 September 2026)

Gravitational-wave astronomy typically relies on rigorous, computationally expensive Bayesian analyses. Several methods have also been developed to perform rapid, approximate Bayesian inference. We present a novel approach to leverage the results of these low-latency analyses to accelerate the final inference, while ensuring that the Bayesian prior remains independent of the data. By combining the fast constraints from the `simple-pe` algorithm with the nested sampling acceleration technique of posterior repartitioning, we demonstrate that our method can guide the nested sampler toward the most probable regions of parameter space more efficiently for signal-to-noise ratios (SNR) greater than 20, while mathematically guaranteeing that the final inference is identical to that of a standard, uninformed analysis. We validate the method through an injection study on signals with $\text{SNR} < 150$, demonstrating that it produces statistically robust and unbiased results while providing speedups of up to 210%, with a mean speedup of 34% for $\text{SNRs} > 20$. Importantly, we show that the performance gain provided by our method scales with SNR, establishing it as a powerful technique to mitigate the cost of analyzing signals from current and future gravitational-wave observatories.

DOI: [10.1103/pnrs-9zd5](https://doi.org/10.1103/pnrs-9zd5)

I. INTRODUCTION

A crucial part of gravitational-wave data analysis is parameter estimation (PE), the process of inferring physical properties of a source, such as its component masses or sky location, from the observed signal, given a model [1]. The results of PE are fundamental to our understanding of stellar evolution, general relativity, and cosmology (most recently, see e.g. [2–10]). As the global network of gravitational-wave (GW) observatories improves in sensitivity and detects an ever-increasing number of events, the demand for efficient analysis methods is growing [11–16], and will become more pressing as more GW detectors come online [17]. This has led to several recent efforts in this area [18–26], including using machine learning algorithms [27–32]. This new era of high signal-to-noise ratio (SNR) signals is exemplified by the recent loud event GW250114_082203, which, with an estimated network

SNR of ~ 80 , is significantly louder than previous detections and promises a new frontier for precision science [33,34].

The standard approach to perform GW parameter estimation is through Bayesian inference [35,36]. This is typically accomplished using stochastic sampling algorithms like nested sampling (NS) [37,38] (although see e.g. Refs. [18,22] for alternative techniques). Nested sampling is highly effective and reliable, providing not only estimates for the posterior distribution but also the Bayesian evidence required for model comparison [38]. However, the algorithm's computational cost presents a significant bottleneck. A standard PE analysis for a single event can require millions of likelihood evaluations, often translating to hours or days of runtime on high-performance computing clusters [39]. This computational demand poses an impending challenge for the higher event rates of future observing runs, and already makes the routine use of the most physically comprehensive models, such as those including orbital eccentricity, expensive.

To address the growing computational challenge associated with nested sampling, we introduce a method that combines two existing techniques: posterior repartitioning (PR) [40,41] and the `simple-pe` analysis [42]. PR is a technique that can accelerate NS by using an approximate posterior, learned from a preliminary NS analysis, to define

*Contact author: myp23@cam.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

a more efficient sampling prior [43]. `simple-pe` leverages intuitive physical arguments to produce fast, probabilistic constraints on the source parameters [42], including general relativistic effects such as spin-induced orbital precession [44] and higher order multipole moments [45,46]. Designed to produce initial parameter estimates in minutes on a single CPU, `simple-pe` can be employed for real-time alerts, as well as for informing more detailed parameter estimation offline.

In this work, we propose a new framework that replaces the preliminary NS run previously used for PR with the rapid, physically motivated constraints from `simple-pe`. This integration has the potential to be faster and more robust than previous implementations [43], as it avoids a potentially slow and unreliable preliminary stochastic search. The result is a parameter estimation framework that significantly reduces the number of likelihood evaluations required for convergence, enabling more efficient analysis of both current and future data. Combining our method with techniques that accelerate the single likelihood evaluation time would further improve the performance gains [47–50].

This paper presents and validates this new PE framework. First, we conduct a large-scale injection study on a population of signals representative of current detections to validate the statistical integrity of the method. We show that it produces unbiased posteriors that are consistent with those from standard NS analyses. Second, we perform a study to characterize how the method’s performance scales with network SNR. This establishes the efficiency gains that can be expected, particularly for the loud events that will become more common in future observing runs.

The remainder of this paper is organized as follows. After a review of Bayesian inference and nested sampling in Sec. II, we detail our PE framework in Sec. III. We then present two distinct studies to assess the method. In Sec. IV, we validate its statistical integrity and in Sec. V, we characterize its performance scaling with SNR. Both sections detail the study design and present the corresponding results. Finally, we present conclusions and discuss future work in Sec. VI.

II. BACKGROUND

A. Bayesian parameter estimation in gravitational wave astronomy

The Bayesian analysis of GW signals involves determining the properties of the astrophysical source, described by a set of parameters θ , given the observed strain data $\mathbf{d}$. Within the Bayesian inference framework, this is achieved by computing the posterior probability distribution, $\mathcal{P}(\theta|\mathbf{d}, \mathfrak{M})$, which represents probability of the parameters given the data and a particular underlying model, $\mathfrak{M}$ [35]. According to Bayes’ theorem, the posterior is given by

$$\mathcal{P}(\theta|\mathbf{d}, \mathfrak{M}) = \frac{\pi(\theta|\mathfrak{M})\mathcal{L}(\mathbf{d}|\theta, \mathfrak{M})}{\mathcal{Z}(\mathbf{d}|\mathfrak{M})}. \quad (1)$$

The model remains unchanged throughout this study, so we will omit the explicit dependence on the model but we refer interested readers to e.g. [51–53] for multimodel inference. Here, $\pi(\theta)$ is the prior probability distribution, encoding our preexisting knowledge about the parameters. The term $\mathcal{L}(\mathbf{d}|\theta)$ is the likelihood, which quantifies the probability of observing the data $\mathbf{d}$ given a specific set of parameters. The denominator, $\mathcal{Z}(\mathbf{d})$, is the Bayesian evidence, obtained by integrating the product of the likelihood and prior over the entire parameter space:

$$\mathcal{Z}(\mathbf{d}) = \int \pi(\theta)\mathcal{L}(\mathbf{d}|\theta)d\theta. \quad (2)$$

While essential for comparing competing models within Bayesian methodology [38] (although this does not account model accuracy, see Refs. [52,53]), the evidence is computationally expensive to calculate.

B. Nested sampling

There exist several stochastic sampling methods to estimate the posterior distributions of GW data, such as MCMC [54,55] and nested sampling [36,38,56–58]. NS is a Monte Carlo algorithm specifically designed to calculate the Bayesian evidence. NS works by drawing a set of “live points” from the prior distribution. At each iteration, the point with the lowest likelihood is removed from the live set, being made a dead point, and is replaced with a new point drawn from the prior, subject to the constraint that its likelihood must be higher than that of the point it is replacing. This process generates a sequence of live points that iteratively move from the broad prior volume into progressively smaller regions of higher likelihood (see Fig. 1). The collection of dead points, when appropriately weighted, provides samples from the posterior, while the sum of their likelihood values multiplied by their enclosed prior volume fractions yields a robust estimate of the evidence.

Despite its effectiveness, NS is computationally intensive. A single analysis often requires hours to days of processing time on high-performance computing clusters (although it has been shown that nested sampling can be parallelized over many CPUs [59], reducing the overall wall time, and techniques have been developed to optimize the GW likelihood [21,24,47,48,60–65], reducing the overall processing time). This cost is an obstacle to achieving rapid scientific turnaround for the large event catalogs of current and future observatories, assuming nested sampling continues to be used [13,66]. For example, it has been shown that a full nested sampling algorithm with the future Laser Interferometer Space Antenna (LISA) observatory can take more than 10^5 CPU hours [67].

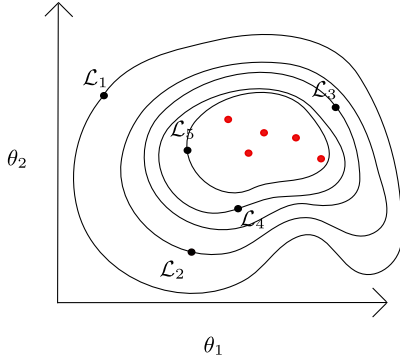


FIG. 1. Cartoon of nested sampling algorithm in two dimensions. Dead points define nested likelihood contours, indicated by $\mathcal{L}_1, \mathcal{L}_2, \dots$, in the parameter space, and the live points compress exponentially toward the peak of the likelihood as the algorithm proceeds. The fractional prior volumes enclosed between successive contours determine the weights of each dead point, used for calculating the Bayesian evidence and drawing posterior samples.

III. METHODS

A. Posterior repartitioning

One method to reduce the computational cost of nested sampling is posterior repartitioning, which improves sampling efficiency by exploiting a fundamental property of Bayesian inference [40,41,43,68]. The posterior distribution and Bayesian evidence depend only on the product of the likelihood and the prior, $\mathcal{L}(\theta)\pi(\theta)$. We are therefore free to redefine these two components for computational purposes, so long as their product remains invariant.

Nested sampling is uniquely suited to this manipulation because its algorithm explicitly separates the prior and the likelihood, by “sampling from the prior subject to a hard likelihood constraint.” PR leverages this by redefining the prior that the NS algorithm sees. The procedure is as follows (also see Fig. 2):

- (1) An approximation to the target posterior, $g(\theta)$, is generated. In the past this has been achieved by training a normalizing flow on samples from a preliminary, low-resolution nested sampling run [43]. Normalizing flows learn a series of invertible mappings from a base distribution to a target [69,70].
- (2) This learned distribution $g(\theta)$ is assigned as the new, effective prior for the NS algorithm. This is the repartitioned prior, $\pi'(\theta) \equiv g(\theta)$.
- (3) To ensure the likelihood-prior product is unchanged, the likelihood must be redefined with a corrective term:

$$\mathcal{L}'(\theta) = \mathcal{L}(\theta) \left[\frac{\pi(\theta)}{\pi'(\theta)} \right] \quad (3)$$

The term in the brackets is the repartitioning factor. The nested sampler is then executed using $\pi'(\theta)$ as

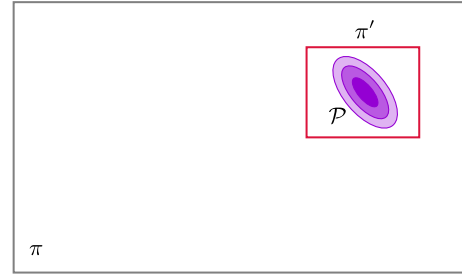


FIG. 2. Most of the computational time of nested sampling is spent in locating the bulk of the posterior, $\mathcal{P}$, within the much larger prior volume, π . In posterior repartitioning, this region of interest is first identified by an approximate method. The nested sampling run is then initialized within this smaller volume, π' , using a repartitioned likelihood [Eq. (3)] to ensure the resulting posterior and evidence are identical to a standard analysis.

its sampling prior and $\mathcal{L}'(\theta)$ for its likelihood evaluations.

It is important to note that this is purely a computational technique; the true Bayesian prior $\pi(\theta)$ of the inference problem is not altered. The result is an NS run that converges faster because its sampling is concentrated within the smaller volume of the repartitioned prior, which by design already resembles the posterior. The acceleration from PR fundamentally depends on the quality of the initial approximation, $\pi'(\theta)$. The approach of using a low-resolution NS run for initialization, as in [43], presents two key drawbacks. First, the preliminary run itself can be computationally expensive, offsetting some of the efficiency gains. Second, and more critically, a stochastic sampler run in a low resolution configuration may fail to identify complex posterior structures, such as sharp peaks or secondary modes. If the initial approximation is poor, posterior repartitioning can be ineffective or, in the worst case, can cut off at the algorithmic level regions of the parameter space with non-negligible posterior probability, leading to biased results. Signals with lower signal-to-noise ratios (SNR) can be particularly susceptible to these issues.

We note that similar techniques that leverage preliminary or intermediate results have been developed for Markov chain Monte Carlo and sequential Monte Carlo [71,72]. We focus on PR, as nested sampling is the current standard for GW analyses. We leave a comparison to other approaches for further work. We also note that acceleration techniques which optimize the GW likelihood are complementary to PR and, while not considered, could also be applied.

B. simple-pe

The simple-pe algorithm provides a rapid analysis framework that generates estimates for the posterior probability distribution within minutes on a single CPU [42]. Instead of performing a full Bayesian analysis with stochastic sampling, simple-pe uses simple, intuitive, and physical insights to estimate the source properties.

simple-pe takes advantage of the well-known physical decomposition of GW signals, whereby a GW signal can be decomposed into an infinite sum of spin-weighted spherical harmonics [45,46]. However, simple-pe relies on the fact that only a small number of these terms significantly affect the observed signal; it can be shown that a GW signal consistent with general relativity (assuming quasicircular orbits) can be described to high fidelity as a sum of three orthogonal components [42,73]. In essence, the dominant term describes the base chirplike signal, the second term describes the contribution due to spin-induced precession and the third term describes the contribution from higher order multipole moments. The central principle of simple-pe is that a measurement of the dominant term provides the primary information for the intrinsic parameters: chirp mass, mass ratio, and aligned spins. A measurement of the precessing and/or higher order multipole terms then further constrains these intrinsic measurements and provides information on the extrinsic properties. While simple-pe is capable of analysing precessing systems, with the possibility of incorporating eccentric binaries [74], this is not explored in this work.

The simple-pe algorithm treats parameters hierarchically. It first identifies the most likely chirp mass, mass ratio and aligned-spin based on the best fitting dominant only term that maximizes the matched filter SNR. Although a starting point is needed, which can come from a preliminary search, simple-pe has routines to determine a best-fitting template through an SNR optimization step, similar to what is done in some online searches by the LVK [75]. It then approximates the likelihood surface around this 3 dimensional peak as a multivariate Gaussian, where the width is based on the expected accuracy at which the mass and spin can be measured, accounting for known degeneracies throughout the parameter space [76]. This simple distribution allows for the rapid generation of initial mass and spin samples. Extrinsic parameters, such as distance and orientation, are handled separately by sampling from constrained distributions that depend on the network SNR and response. Finally, SNRs for features beyond the dominant term are obtained by matched filtering the data again, and these initial samples are refined through a reweighting step; each sample is assigned a weight based on the probability of obtaining each of the observed SNRs.

While this approach sacrifices the full rigour of an NS analysis, its physically informed logic makes it a more robust tool for generating initial parameter estimates. Because its constraints are derived from specific physical features rather than stochastic exploration, it may be less susceptible to the same failure modes as a low-resolution sampler. For example, it is potentially less likely to miss secondary modes of the posterior, such as in the inclination angle, which is known to be often bimodal. This makes it an ideal candidate for improving the initialization stage of posterior repartitioning.

However, due to the approximations employed, simple-pe in its current form may not guarantee the most reliable estimates. For example, for binaries with significant spin-precession or higher order multipoles, more than three terms are needed in the waveform expansion to fully describe the observed signal [77,78]. By neglecting some higher order multipole power, biases may be expected particularly in high mass systems (see e.g. [79]). Similarly, by using only the dominant term to initially provide samples for the mass and spin of the binary, biases may be expected for high SNR signals with significant contributions from subdominant terms [73] (although the initial samples are reweighted to account for the observed power in subdominant terms; if the initial distribution lies outside of the region consistent with the observed power, the reweighting procedure fails). Finally, simple-pe may struggle for low SNR signals since the multivariate Gaussian approximation for the initial mass and spin samples begins to break down.

C. The simple-pe-PR algorithm

Our method integrates the constraints from simple-pe into the posterior repartitioning framework through a multistep algorithm. The full procedure, shown schematically in Fig. 3, is as follows:

- (1) Initial constraints: The method begins by running simple-pe on the gravitational-wave strain data to generate a set of samples. These samples provide a rapid, physically motivated initial approximation of the posterior.
- (2) Learning a continuous distribution: The discrete samples from simple-pe must be converted into a continuous probability distribution from which we can both sample and evaluate the log-probability. We achieve this by training a normalizing flow, using a modified version of the MARGARINE software [80–82], on the simple-pe samples (see Appendix B for details). This flow serves as a learned, continuous representation of the initial posterior approximation. However, other methods besides normalizing flows are possible for obtaining this, such as generation of a multidimensional kernel density estimate [83], or Gaussian processes [84].
- (3) Widening for robustness: Due to the approximations employed, the distribution produced by simple-pe can be narrower than, or offset from, the true posterior [85]. To prevent this from biasing the final inference, the learned distribution is widened. The specific widening factors applied are an important choice, and their empirical calibration based on our injection study is detailed in Sec. IV A. The widening itself is achieved by modifying the base distribution of the trained normalizing flow. Specifically, we increase the standard deviation of the base multivariate Gaussian distribution. The covariance matrix

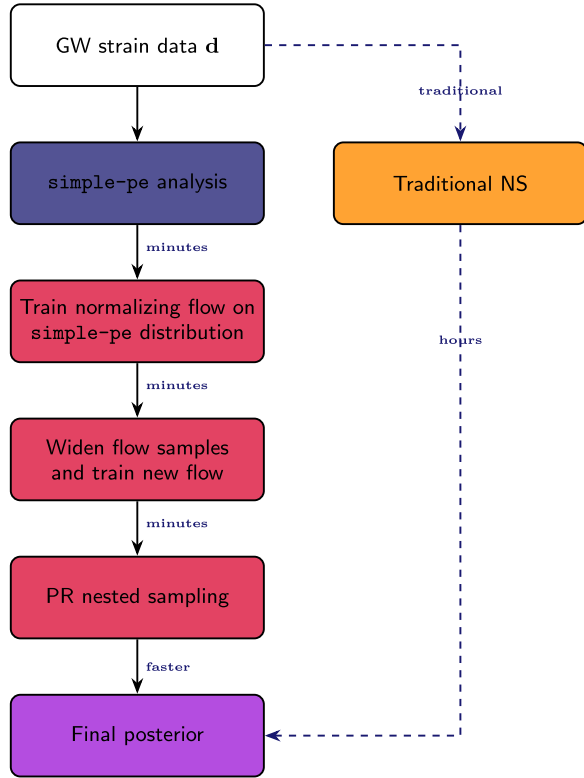


FIG. 3. A flowchart of our proposed algorithm. Our method (left) introduces several fast preprocessing steps before running a posterior repartitioned (PR) nested sampling analysis. This reduces the runtime compared to a traditional nested sampling (NS) analysis (right). The primary computational bottleneck in both methods is the nested sampling step itself. Provided the initial `simple-pe` distribution is sufficiently widened, our method produces the same posteriors as the traditional approach but with significantly fewer likelihood evaluations.

is diagonal, meaning that we can change the standard deviation for each dimension independently. 2000 samples are then drawn from this “widened” flow. A second normalizing flow is then trained on these widened samples to produce the final, robust repartitioned prior, $\pi'(\theta)$. During training of both flows, we do not apply bounds and allow the flows to have support outside the original prior. This will not affect the final inference as samples outside the original prior will not be accepted by the sampler anyway, and we discuss this further in Appendix B.

- (4) Final parameter estimation: The final PE analysis is performed using `Bilby-pr`, a custom `Bilby` plugin package developed for this work and made publicly available on GitHub [86]. This package extends the standard `Dynesty` sampler [57] as implemented in `Bilby` [56], with arguments to accept a pretrained MARGARINE normalizing flow. It then implements custom prior and likelihood wrappers that employ the flow as the new repartitioned prior, $\pi'(\theta)$, and adjust the likelihood according to Eq. (3).

The repartitioned prior, $\pi'(\theta)$, does not incorporate information for all parameters. The constraints from `simple-pe` on the sky location parameters, right ascension (α) and declination (δ), exhibited a variance in accuracy that made them unsuitable for reliable repartitioning [87]. Similarly, `simple-pe` does not provide distributions for the phase, polarization angle, and time of coalescence. For these parameters, we retain the original priors. Finally, since we analytically marginalize over luminosity distance in the likelihood in order to improve convergence [88,89], we do not use the `simple-pe` distance estimates. In summary, our repartitioned prior only incorporates information from `simple-pe` for the intrinsic binary parameters and the inclination angle, where its performance is most robust.

IV. INJECTIONS STUDY

We conducted an injection study to validate the statistical robustness of `simple-pe-PR` on a population of simulated binary black hole merger signals, each 8 seconds in duration. The parameters for these injections were drawn from the prior distributions detailed in Table III in Appendix A. This choice of priors is representative of those used in LVK analyses, including the choice of mass priors (see Appendix A for further discussion). It therefore allows us to validate our method’s statistical robustness in a similar parameter range.

The signals were generated using the aligned-spin waveform model `IMRPhenomXHM`, which includes higher order multipoles [90]. They were injected into colored Gaussian noise representative of the design sensitivity of the Advanced LIGO-Hanford (H1), LIGO-Livingston (L1), and Virgo (V1) detectors [91].

A. Calibration of widening procedure

A key component of our method is the widening of the initial `simple-pe` distribution, by applying a multiplicative increase to the standard deviation of the flow’s base distribution. This ensures the resulting posteriors are unbiased. The primary goal of this first study is to determine the requisite widening.

To perform an empirical calibration of the necessary widening factors, we drew a set of 100 injections from the prior. For each of these calibration signals, we compared the initial `simple-pe` distribution to the full posterior from a standard NS analysis. For each repartitioned parameter, we then determined the widening factor required to ensure the 1st and 99th percentiles of the reference one-dimensional marginal posterior were contained within the 1st and 99th percentiles of the one-dimensional marginal widened distribution.

This calibration confirmed that the performance of `simple-pe` depends on the network SNR [42]. A notable change in behavior was observed around an SNR of 20:

TABLE I. Widening factors (w) applied to the standard deviation of the normalizing flow’s base distribution for signals with network SNR < 20 (SNR ≥ 20).

Parameter	Widening factor (w)
Mass Ratio (q)	2.5
Chirp Mass ($\mathcal{M}_c$)	2.0 (1.5)
Spin of lighter BH (χ_1)	2.0 (1.5)
Spin of heavier BH (χ_2)	1.0
Inclination angle (θ_{IN})	2.5

simple-pe generally performed better on signals louder than this, and hence needed less widening. Consequently, we adopted two distinct sets of widening factors: one for signals in the lower-SNR regime (network SNR < 20) and a less conservative set for high-SNR signals (network SNR ≥ 20), shown in Table I.

The widening factors we use are intentionally conservative to ensure robustness for the majority of signals with SNRs above 10. For many signals, the initial distributions require little to no widening; 25% of the injections in the calibration set needed widening factors less than 1.25 in all parameters. As discussed in Sec. V, this has an impact on the observed performance gains. For signals below SNR 10, the required widening becomes large enough to negate any computational benefit from posterior repartitioning. While we apply the lower-SNR factors to these signals for the purpose of our proceeding injection study, this fixed widening is often insufficient to guarantee an unbiased posterior and we do not recommend application of our method in this regime (see Sec. IV B for further discussion).

Finally, we note that these widening factors were calibrated on aligned-spin injections in a three-detector network. They may differ for analyses involving two-detector networks or precessing binary systems. We also expect these widening factors to vary with signal duration and prior choices (Appendix A), but do not expect this to be the primary determinant in our method’s performance (see Sec. V). The widening factors were observed to be largely independent of mass over the parameter range considered (where the primary mass spanned from 11.5 to $152.3M_\odot$); other structures in the likelihood, such as those related to SNR, have a more significant impact. While this widening process could be automated in future work, using techniques similar to those introduced in [43], we do not explore this here.

B. Statistical validation and performance

We now conduct a large-scale injection study to demonstrate that our method achieves correct statistical coverage and is consistent with the standard version of Bilby and nested sampling [92]. We also assess the method’s performance. To perform the injection study, we drew a new set of

100 injections from the prior. For each of the signals, we performed two analyses: one using a standard nested sampling algorithm and one using the simple-pe-PR algorithm described in Sec. III C. For the simple-pe-PR analyses, we applied the appropriate SNR-dependent widening factors from Table I.

For both analyses, the nested sampling step was performed with the Dynesty sampler [57] as implemented in Bilby. The right ascension (α), polarization angle (ψ), and phase at coalescence (ϕ) were treated with periodic boundary conditions. Identical sampler settings were used for both the standard and PR analyses to ensure a direct comparison; we used 1000 live points and the “acceptance walk” sampler, with its default settings [1].

1. Percentile-percentile comparison with standard nested sampling

The statistical integrity of our method is validated by a population-level comparison with the standard NS analysis, using the statistics of the probability-probability (P-P) plot [93]. A P-P plot shows the fraction of signals for which the injected parameter values lie within a given credible interval, as a function of the credible interval. In the ideal case where the parameter inference procedure leads to the correct coverage, this plot should be a diagonal line. However, there is a statistical uncertainty that arises from the finite number of signals analyzed.

The recovered distributions for all parameters from our method are consistent with those from the normal NS method. We note that the individual P-P plots for both methods revealed small biases in the declination and time of coalescence at the geocenter parameters. Importantly, though, these biases do not arise from changes related to our method; our results are consistent with standard NS. We attribute these biases to statistical variance within this specific realization of 100 injections, as they were not present in the separate realization used for calibration. Sky location and time of coalescence at the geocenter have known degeneracies [94], which may cause statistical fluctuations in one parameter to manifest as apparent biases in both. The P-P plots obtained from this set and the calibration set are included in Appendix F for reference. It should also be noted that summary statistics, such as the P-P plot, have been previously shown to be subject to an additional source of uncertainty that is not typically quantified in P-P plot error bars [95], which could be relevant to these apparent biases; we did not include this source of uncertainty in our analysis, in line with standard practice in the literature.

To more directly compare the population level agreement between the methods, the differences between the two P-P plots for some representative parameters are shown in Fig. 4 (the remaining parameters are plotted in Fig. 19). The 2σ error bars estimated from the injections are also plotted,

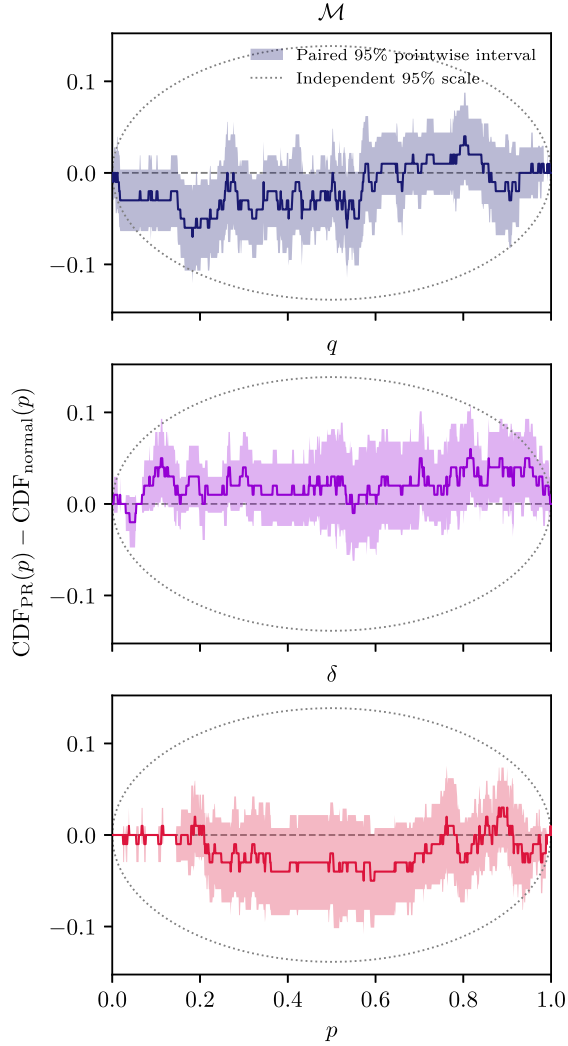


FIG. 4. The solid lines show for each parameter the difference between the two P-P plots obtained from `simple-pe-PR` and standard NS, $\text{CDF}_{\text{PR}}(p) - \text{CDF}_{\text{normal}}(p)$, as a function of the credible interval, p . The ideal case for two methods which perfectly agree is a horizontal line at zero, but statistical uncertainties play a role. The dotted lines show the 2σ error bars if the two CDFs were statistically independent. This is shown for reference only, since in our case both methods run on the same set of injections so their P-P curves are positively correlated. The shaded bands therefore show the 2σ pointwise uncertainties estimated from the standard deviation of the differences. In particular, even the biased declination (bottom panel) shows consistency between the two methods, confirming that the biases are not due to changes related to `simple-pe-PR`.

which demonstrate again that the two methods are consistent with each other to within statistical fluctuations.

2. Event level agreement with standard nested sampling

For every injection, we also compute the individual one-dimensional Jensen-Shannon divergence (JSD) between each pair of posterior distributions (see Appendix D for

JSD methodology and full results, as well as [96] [97]). For the majority of injections, the posteriors from `simple-pe-PR` are in good agreement with those from standard NS, confirming the general robustness of the widening procedure. The JSD agreement also supports the hypothesis that the biases observed in the P-P plot are due to a statistical fluctuation in the properties of this specific set of injections. Nevertheless, we observed two injections above SNR 20 that exhibited significant differences in the posterior estimates. In both cases, this was due to `simple-pe` identifying the opposite sky mode, resulting in an offset inclination angle estimate. Our widening procedure is not designed to account for such cases where the initial estimate has significantly different support to the posterior. This type of failure occurs more frequently at lower signal strengths. Between SNR 10 and 20, six injections, representing 12.5% of signals in this range, show significant posterior differences. Below SNR 10, this number increases to 10 of the 12 injections, an expected result as our widening factors were not calibrated for this regime.

Although the inclination angle failure mode can produce biased posteriors in rare instances, we still include it in our repartitioned prior to leverage the noticeable speedups it provides for the majority of signals. Further improvements to the `simple-pe` algorithm will alleviate this failure mode, both ensuring more robust results and reducing the need for our widening procedure. Moreover, the widening procedure itself could be improved by varying the degree of widening on the fly during the analysis, by sampling over an additional widening factor parameter, similarly to [43]. This would permit larger widening factors to be incorporated for more challenging signals without sacrificing performance on signals which require little to no widening.

3. Computational performance

We quantify the performance improvement, or speedup, of our method as the ratio of a given metric (e.g., number of likelihood evaluations) for a standard NS analysis to that of our `simple-pe-PR` analysis. The total runtime of our method is dominated by the nested sampling step. On average, `simple-pe` took 5 CPU minutes, and training both the original and widened normalizing flows together took less than one CPU minute. By comparison, the nested sampling step took on average 7 hours, parallelized over 38 CPU cores. Therefore, when comparing runtimes, we only consider the respective nested sampling times of each method. We could consider the speedup in terms of total likelihood evaluations, but one characteristic of PR is that the analysis typically spends a larger fraction of its runtime sampling the high-likelihood region of the parameter space, yielding a larger effective sample size than the equivalent standard analysis (see Appendix E). Accordingly, our method can be performed with fewer live points than a standard NS analysis for the same final sample size. As such, rather than the total speedup, the per-sample speedup

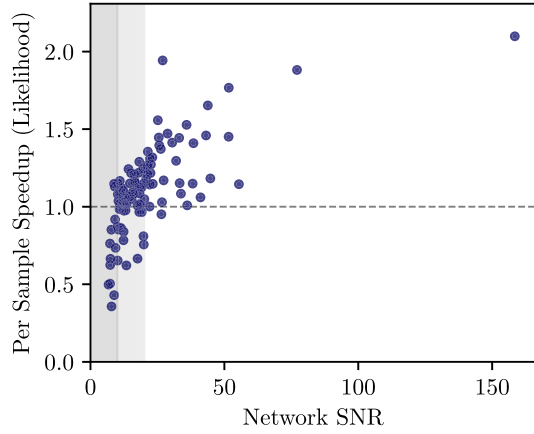


FIG. 5. Per-sample likelihood evaluation speedup factor for each of the 100 injections, as a function of network SNR. Each point represents a single analysis. The lighter shaded region ($\text{SNR} < 20$) highlights the regime where the speedup is unreliable. It also marks the boundary of the two sets of SNR-dependent widening factors. The darker shaded region ($\text{SNR} < 10$) indicates where the calibrated widening factors may not be large enough to obtain unbiased results. The dashed line at 1.0 indicates no change in performance relative to a standard analysis.

(the speedup factor normalized by the number of effective posterior samples from each analysis, calculated as in [98]) is the more informative performance metric, as it compares like-for-like inference products. Figure 5 shows this per-sample speedup in terms of likelihood evaluations for each injection, as a function of network SNR. As anticipated from the expected performance of `simple-pe` [42] and the SNR-dependence of the widening factors, there is a clear positive trend in performance with SNR, albeit with significant scatter. The maximum speedup observed is 2.1, for an injection with SNR 158. The performance distribution reflects the characteristics of our injection set, which was drawn from the prior distributions used in LVK analyses. Although the prior on the luminosity distance was narrowed to ensure that most signals have detectable SNRs (Table III), the majority of signals still have network SNRs below 20 (shaded region). The performance of our method in this regime is expected to be modest, or in some cases worse than standard NS. We find that 71% of injections in our population show speedups greater than 1, and for signals with $\text{SNR} > 10$, this figure increases to 78%. The mean speedup for injections with $\text{SNR} > 20$ was 1.34.

As discussed in Sec. IV A, our chosen widening factors are intentionally conservative for the majority of signals above SNR 10; on the other hand, below SNR 10, the applied widening factors are generally expected to be insufficient. These results therefore highlight an important distinction between the statistical robustness and the computational efficiency of the method at different signal strengths. For signals below SNR 10, the method is

unlikely to produce reliable results and we do not recommend its use. For signals with SNR between 10 and 20, the empirically calibrated widening factors are robust for most cases, and the final inference will be statistically consistent with the standard analysis, excepting certain failure modes. However, as shown in Fig. 5, the performance in this regime is unreliable, with some analyses resulting in a slowdown. Therefore, for both statistical validity and computational efficiency, we recommend the application of our method for the time being to signals with a network $\text{SNR} > 20$. While this represents a minority ($\lesssim 10\%$) of current detections in GWTC-4.0, these high-SNR events are often the most information-rich signals and warrant repeated analyses to study carefully [99–101]. As such, they account for a disproportionate fraction of the total computational budget for parameter estimation. For future observing runs, and observatories like the Einstein Telescope, signals with $\text{SNR} > 20$ are predicted to constitute a larger fraction of the observed population; for Einstein Telescope approximately 50% of all detections will have $\text{SNR} > 20$ [102].

There is only one injection with $\text{SNR} > 20$ that resulted in a marginal per-sample slowdown (5% slower than standard NS); the final posterior distributions for this analysis were in excellent agreement with the standard NS analysis. The slowdown resulted from an over-broad repartitioned prior compared to the highly constrained target posterior, due to the conservative widening factors. Consequently, the internal random walk of the nested sampler frequently proposed points in regions of high repartitioned prior probability but low likelihood. This reduced the internal acceptance rate, thereby increasing the number of likelihood evaluations required to find a valid new live point at each nested sampling iteration. This event illustrates a key trade-off of our current design. The fixed widening procedure prioritizes statistically unbiased inference, but is therefore occasionally too conservative, leading to localized sampling inefficiencies. Automating the widening procedure in future would lead to better performance improvements.

4. Multimodality

A key advantage of `simple-pe` is that it is in general more robust to multimodal posteriors, a common feature in gravitational-wave analysis that presents a significant challenge for other PR initialization methods. A low-resolution NS run or a Fisher matrix approximation can both fail to identify secondary posterior modes, particularly when they are well separated from the primary mode. Such a failure leads to biased inference, as the repartitioned prior can inadvertently exclude regions of significant posterior probability. Figure 6 provides a practical demonstration of our method’s advantage over other PR initializations in this regard. The posterior for this injection is bimodal in the inclination angle, which the `simple-pe` analysis

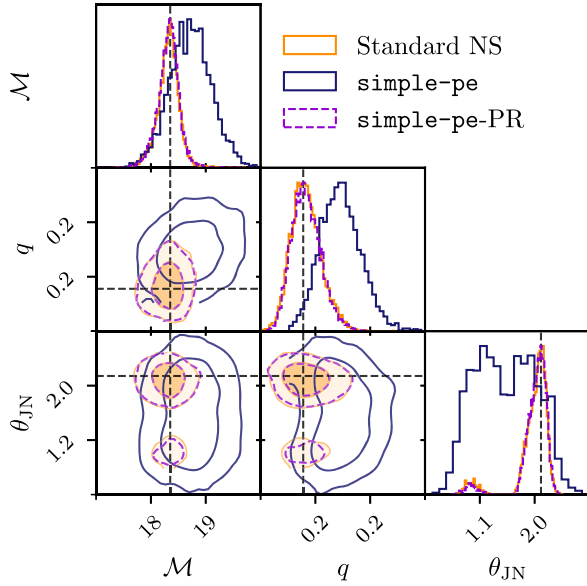


FIG. 6. A representative injection from the study with a bimodal marginal posterior for the inclination angle (θ_{JN}). The `simple-pe` approximation (unwidened) correctly identifies both modes of the posterior. This ensures the repartitioned prior encompasses the full posterior structure, including the secondary mode, leading to an unbiased final inference.

successfully captures. This information is then passed to the repartitioned prior, which ensures the subsequent NS analysis explores both modes and converges to the correct, unbiased posterior distribution. This case highlights how the physically motivated model of `simple-pe` can provide a more reliable starting point for PR than a simple stochastic search, preserving the statistical integrity of the analysis for complex posterior geometries.

In summary, the injection study successfully validates that our method produces statistically consistent posteriors across the considered parameter range, and demonstrates the potential performance improvements.

V. SNR SCALING STUDY

The efficiency of PR is determined by the quality of the initial posterior approximation. Since both the precision of `simple-pe` and the concentration of the true posterior are expected to improve at higher SNR, we anticipate that the computational speedup of the `simple-pe-PR` algorithm will grow with signal strength. This was empirically supported by the results in Sec. IV, but now we conduct a targeted study to investigate the performance of our method for signals with varying SNR.

We focus on a single fiducial binary black hole merger. The parameters for this injection are listed in Table II, and were chosen randomly from the prior specified in Table III. We then generated a series of signals, each 8 seconds in duration. For each signal, all source parameters were kept constant except for the luminosity distance, which was

TABLE II. Fiducial detector frame parameters for the injection used in the SNR scaling study. The luminosity distance is not listed as it was varied to achieve the target SNRs.

Parameter	Value
Chirp mass ($\mathcal{M}_c$)	$22.5 M_\odot$
Mass ratio (q)	0.11
Right ascension (α)	1.11 rad
Declination (δ)	1.33 rad
Inclination angle (θ_{JN})	2.59 rad
Polarization angle (ψ)	1.01 rad
Phase at coalescence (ϕ)	3.82 rad
Spin of lighter BH (χ_1)	0.28
Spin of heavier BH (χ_2)	0.49
Time of coalescence (t_c)	1367638517.66 s

scaled to produce signals with target network SNRs of 10, 20, 30, 40, 50, 75, 100, and 150.

The injection and recovery of each signal followed the procedure detailed in Sec. IV. We used the priors detailed in Table III in Appendix A, but adopted a wider luminosity distance prior, from 100 to 5000 Mpc, which was more suitable for the low-SNR injections in this study.

A. Statistical agreement

The posterior distributions produced by `simple-pe-PR` and standard NS showed excellent agreement for the signals with $\text{SNR} > 20$ (Fig. 11). For the SNR 10 signal, our `simple-pe-PR` method produced a narrower posterior on the mass ratio, q , than standard NS, due to insufficient widening of the initial `simple-pe` distribution. For the SNR 20 signal, the standard NS analysis revealed a very small secondary mode in the inclination angle (containing 0.6% of the samples) that was absent from our method's result. This feature was not present in the initial `simple-pe` estimate and was sufficiently

TABLE III. Prior distributions for the 100 injections used in the injection study. The choice of a source-frame uniform luminosity distance prior and uniform component (comp) mass priors are chosen to be consistent with standard LVK analyses.

Parameter	Shape	Range
Chirp mass ($\mathcal{M}_c$)	Uniform (comp)	$[10, 25] M_\odot$
Mass ratio (q)	Uniform (comp)	$[0.05, 1.0]$
Luminosity distance (d_L)	Uniform (source)	$[100, 1000]$ Mpc
Right ascension (α)	Uniform	$[0, 2\pi]$ rad
Declination (δ)	Cosine	$[-\pi/2, \pi/2]$ rad
Inclination angle (θ_{JN})	Sine	$[0, \pi]$ rad
Polarization angle (ψ)	Uniform	$[0, \pi]$ rad
Phase at coalescence (ϕ)	Uniform	$[0, 2\pi]$ rad
Spin of lighter BH (χ_1)	Uniform (aligned)	$[-0.99, 0.99]$
Spin of heavier BH (χ_2)	Uniform (aligned)	$[-0.99, 0.99]$
Time of coalescence (t_c)	Uniform	trigger time + $[-0.1, +0.1]$ s

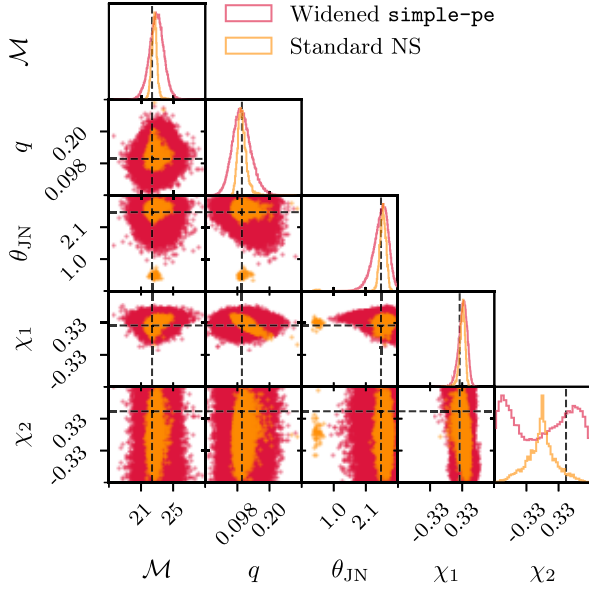


FIG. 7. For the SNR 20 signal, `simple-pe` did not produce sufficient samples from the secondary mode of the true posterior in order for the flow to learn this feature. The widened flow is able to generate samples which lie in this secondary mode, but with very low probability. Since the sampler recognizes that this region of the parameter space has a high repartitioned likelihood, it attempts to explore it, but inefficiently and unsuccessfully, leading a marginally slower analysis than standard NS and a slightly different final posterior. This issue can be mitigated by drawing more initial samples from `simple-pe` to make sure the secondary mode is sufficiently covered.

distant from the primary mode that our widening procedure did not fully encompass it (see Fig. 7). This exemplifies another limitation of the method: although `simple-pe` can handle multimodal posteriors, as demonstrated in Sec. IV, nevertheless it can still miss secondary features, particularly when their relative posterior mass is very small. This risk is most pronounced at low SNRs where posteriors can be complex, and is significantly mitigated for high SNR signals. Moreover, `simple-pe` produces samples drawn from a weighted distribution. The risk can be further mitigated, therefore, even at lower SNRs by drawing more samples to provide to the flow, which takes negligible additional time. Alternatively, the flow could also be trained directly on the weighted samples themselves.

B. Performance scaling with SNR

Figure 8 shows the speedup factors as a function of optimal network SNR. Two metrics are considered when quantifying the performance improvement: the number of likelihood evaluations and the sampling time. Note that the speedup difference between these two metrics at a given SNR arises from the computational overhead of the repartitioned likelihood [Eq. (3)]. This calculation requires evaluating not only the original likelihood, but also the

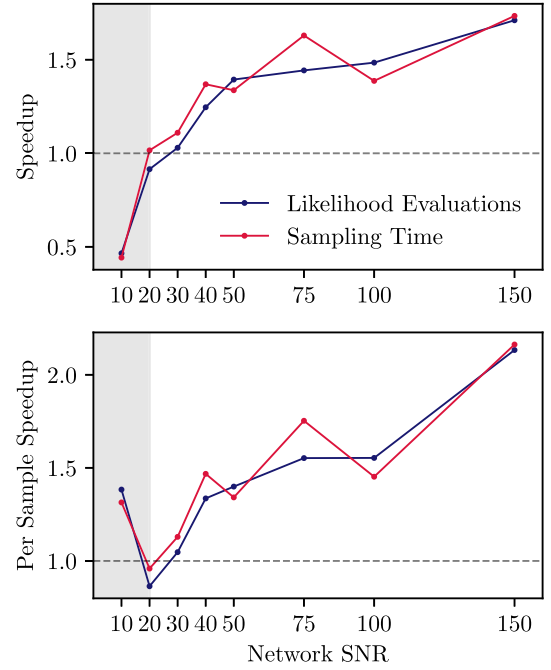


FIG. 8. Performance of the `simple-pe-PR` method relative to a standard nested sampling analysis as a function of network SNR. The speedup factor is the ratio of a metric for the standard analysis to that of the PR analysis. Top: total speedup factors, showing the ratio of total likelihood evaluations (blue) and total sampling time (crimson). Bottom: speedup factors normalized by the number of effective posterior samples. The dashed line at 1.0 indicates no change in performance. The shaded and nonshaded regions separate the two different regimes used for the widening factors.

log-probabilities of both the normalizing flow and the original prior. This additional cost would be mitigated if a vectorized sampler were used, which could evaluate the flow probabilities of multiple points at once, reducing the average single particle evaluation time [22,58,103–106]. The top panel of Fig. 8 plots the overall speedup, while the bottom panel shows the per-sample speedup.

In terms of overall speedup (top panel), our method is slower than the standard approach at SNR 10 and 20. As the SNR increases, however, the method’s performance improves substantially. Above SNR 20, `simple-pe-PR` becomes more efficient than standard NS and by SNR 150, the number of likelihood evaluations and the total sampling time are reduced by a factor of 1.7, corresponding to a 41% reduction. While the overall trend shows a clear monotonic increase in performance with SNR, minor deviations in the sampling time are consistent with temporal variations in CPU node performance on high performance computing systems.

Considering instead the per-sample speedup (bottom panel), the performance at low SNR is now more favorable; only the SNR 20 analysis is marginally slower. At SNR 150, the number of likelihood evaluations and total

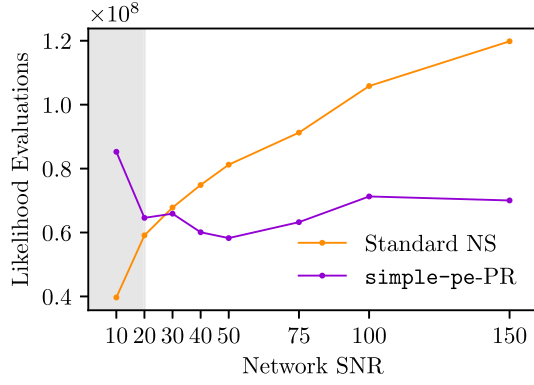


FIG. 9. Total number of likelihood evaluations as a function of network SNR for a standard analysis (orange) and our `simple-pe-PR` analysis (violet). The computational cost of the standard analysis grows with SNR, while the cost of the PR analysis remains approximately constant for $\text{SNR} > 20$.

sampling time are now reduced by a factor of 2.1 (a 53% reduction). Furthermore, the per-sample speedup at high SNR shows no sign of saturating, suggesting that further efficiency gains may be possible for even louder signals. Although SNR 150 signals are unlikely to be routinely observed until the advent of third-generation detectors or LISA [107–111], our method still provides an improvement on standard NS above SNR 20. We already observe speedups of above 1.5 at $\text{SNR} > 50$, a regime where signals are starting to be observed in O4 [33,34] and will become more frequent in O5.

At SNR 10, the poor performance of our method can be explained by the posterior distribution itself being broader at low SNR. At SNR 20, the missed secondary mode in the widened flow leads to sampling inefficiencies. The sampler attempts to explore the part of the parameter space containing the secondary mode, recognizing it as a region of high likelihood for the repartitioned analysis [Eq. (3)], but since this lies deep in the tails of the widened flow, the analysis requires more likelihood evaluations (Fig. 7).

The origin of our method’s performance improvement beyond SNR 20 is illustrated in Fig. 9. As expected, the number of likelihood evaluations required by the standard NS analysis grows with SNR. This is because the posterior volume becomes increasingly small relative to the prior, increasing the number of iterations of the algorithm required for convergence. In contrast, the number of likelihood evaluations for the PR analysis shows no strong scaling with SNR for signals above SNR 20, remaining approximately constant within the stochastic variations expected from the normalizing flow training and the sampling process. Our method effectively removes the expensive initial “search phase” of the nested sampling algorithm (see Fig. 10).

This efficiency gain has significant practical implications. As well as reducing runtimes by hours to days, this reduction in cost lowers the barrier for performing analyses

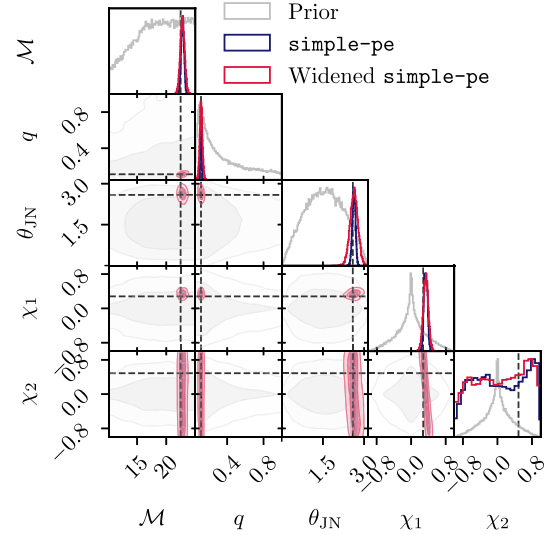


FIG. 10. The prior, π , `simple-pe` distribution and widened distribution used as the repartitioned prior, π' , are shown for the SNR 40 injection. The repartitioned prior occupies a much smaller volume than the original prior, reducing the number of likelihood evaluations required for convergence. In all parameters but the spin of the heavier BH (χ_2), the repartitioned prior is much more constrained than the original prior and effectively removes the initial “search phase” of NS.

with more physically complete but expensive waveform models, such as those including orbital eccentricity [112–115]. Although many of the signals in version 4.0 of the Gravitational-Wave Transient Catalog (GWTC-4.0) have SNRs < 20 , 6% have SNRs > 20 [116], and this percentage is likely to increase as detector sensitivity improves [13].

Overall, the performance of our method is primarily driven by the Kullback-Leibler (KL) divergence between the prior and the posterior, which can be thought of as the “amount of compression” required for the sampler to converge [38,43,68]. This depends most strongly on the SNR of the signal and the initial prior volume. For example, for a low-SNR signal this divergence is modest, and the standard NS analysis already converges relatively quickly. In this regime, applying a conservative widening factor to the already broad `simple-pe` approximation can increase the effective KL divergence, leading to a longer runtime (see Appendix C for further discussion). Additional sampling inefficiencies can be introduced if the widened flow proposes samples outside the support of the original prior (see Appendix B for further discussion). We do not expect the speedups to vary significantly with the binary’s component masses independently; any observed correlation with mass is an indirect consequence of these two dominant factors. This conclusion relies on the assumption that the quality of the initial `simple-pe` approximation is reasonably consistent across the parameter space. While a detailed study of `simple-pe`’s performance is beyond the

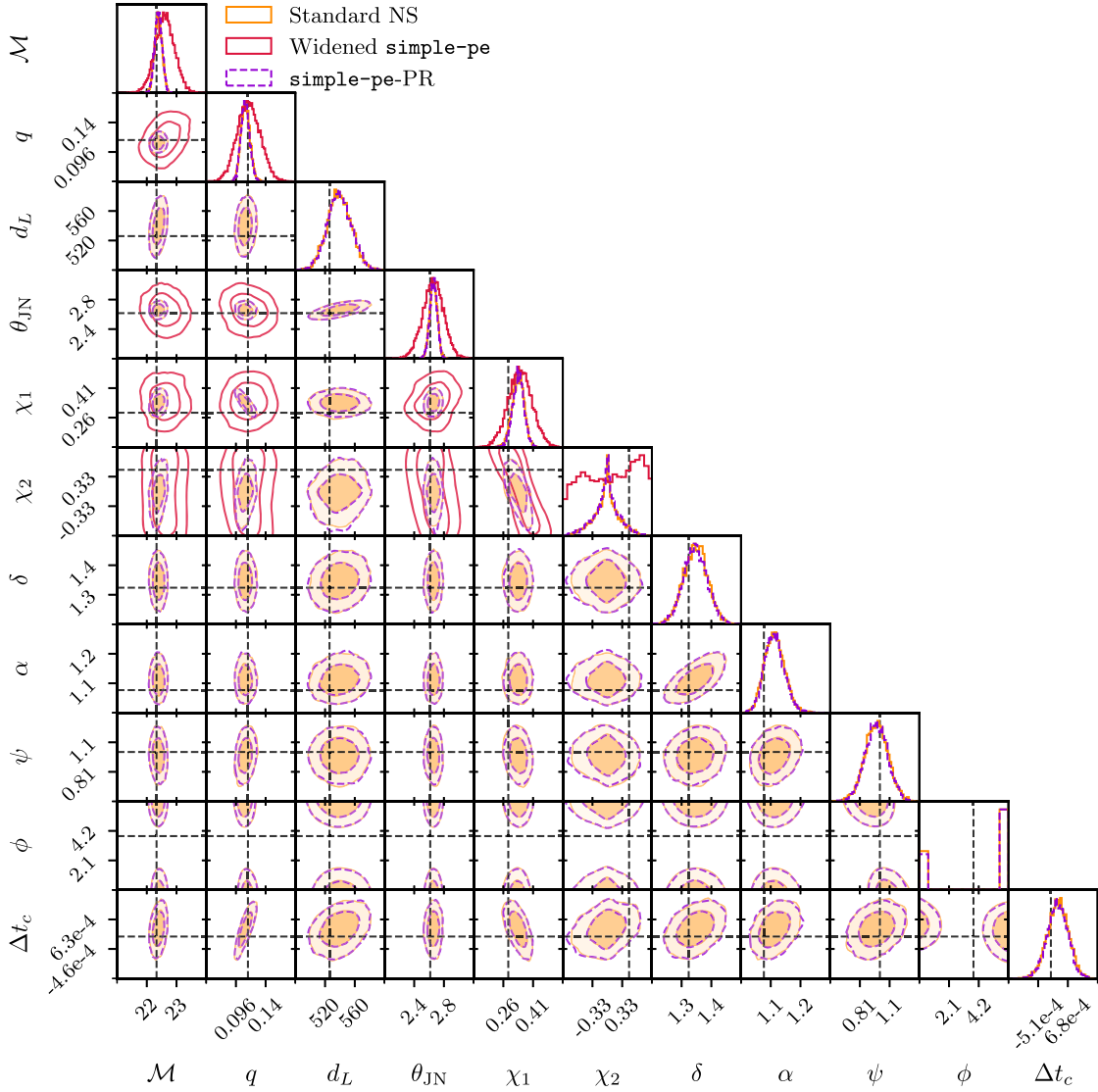


FIG. 11. Validation of the `simple-pe-PR` method on the SNR 40 injection. The one- and two-dimensional marginal posterior distributions from our method (violet) are plotted alongside those from a standard NS analysis (orange), showing excellent agreement. The repartitioned prior, π' , used to accelerate the analysis is also shown for the five parameters to which it was applied (crimson).

scope of this paper, our findings indicate that it is likely a subdominant effect regardless. For the injections in this section, the `simple-pe` approximation became more accurate at higher SNRs, yet we achieved significant speedup improvements beyond SNR 20 without changing the multiplicative widening factors. This demonstrates that the dominant driver of the acceleration is the overall reduction in the prior-to-posterior volume, rather than marginal improvements in the precision of the initial guess.

VI. CONCLUSIONS

The increasing detection rate and SNR of gravitational-wave events from the global detector network are enabling an era of high-precision science, but the computational cost of the required parameter estimation presents a significant

challenge [11–13]. Standard nested sampling, while robust, is computationally expensive, particularly for the loudest signals. In this work, we have presented and validated a method to accelerate this process by combining the rapid, physically motivated constraints from `simple-pe` with the nested sampling acceleration technique of posterior repartitioning.

An injection study, performed on a population drawn from standard analysis priors for 8 second signals, rigorously validated the statistical integrity of our method. The successful probability-probability test and the detailed analysis of challenging cases confirm that our method produces unbiased posteriors consistent with standard techniques, correctly prioritizing statistical robustness over speed when necessary.

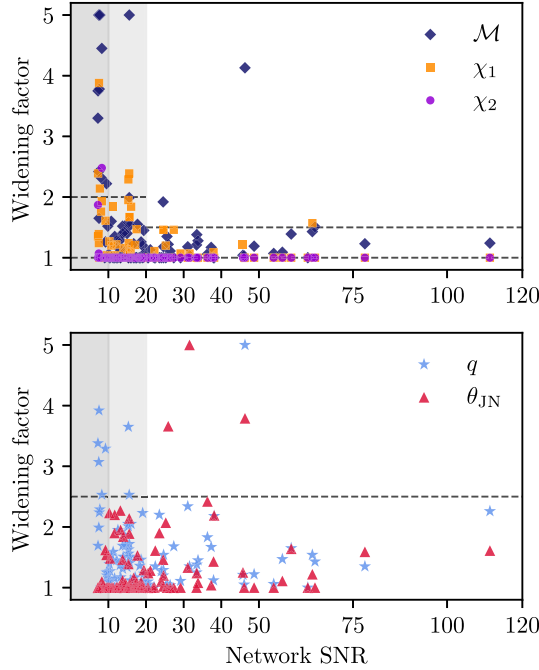


FIG. 12. The minimum widening factor required for the repartitioned prior to encompass the 1st and 99th percentiles of the standard nested sampling posterior for each of the 100 injections. The widening factor represents the multiplicative value applied to the standard deviation of the normalizing flow’s base distribution for a given parameter. These results were used to calibrate the two sets of fixed, conservative factors used in this work (dashed lines).

We have established a clear relationship between SNR and performance, showing that the computational speedup increases with SNR. For a fiducial signal at an SNR of 150, we achieve a per-sample speedup factor of approximately 2.1 in the number of likelihood evaluations. By comparison, a signal produced by the same parameters but a smaller luminosity distance, with SNR 30, achieves a per-sample speedup of 1.2, showing that the efficiency grows significantly with signal strength. Furthermore, the per-sample speedup does not show signs of saturating at SNR 150, indicating that further gains are possible for yet louder signals. This positions our method as a powerful algorithm for next-generation signals from observatories like the Einstein Telescope and LISA [107–111].

Beyond the numerical speedup, our method provides a way to capitalize on the standard GW analysis workflow. Low-latency analyses, like `simple-pe`, are essential for rapid event alerts, but the valuable information they provide about the likely region of parameter space is typically set aside when the final, high-fidelity parameter estimation is performed. This separation is necessary to preserve the integrity of the Bayesian prior, which must not be informed by the data being analyzed. Our method provides an innovative strategy for this purpose. As a purely computational technique, it uses the low-latency result to focus the

sampler on the region of interest, while the repartitioning factor [Eq. (3)] mathematically ensures that the final posterior is identical to one produced from the original, unaltered prior. The method thus gains the efficiency of an informed analysis while rigorously preserving the statistical purity of the final inference.

The repartitioning algorithm presented here is not specific to `simple-pe`. It is possible to use our technique with machine learning algorithms, which also rapidly produce initial posterior estimates [27–32]. However, care should be taken when using initialization methods which are not robust to multimodal posteriors, such as a Fisher matrix approximation, as these could cut off secondary modes in the subsequent analysis and bias the inference. Our method is also complementary to likelihood acceleration techniques; `simple-pe-PR` reduces the number of required likelihood evaluations, and thus can be combined with optimized likelihoods that reduce the cost of each individual evaluation for even further speedups [47–50]. `simple-pe` could also be used to seed alternative sampling methods besides NS, such as sequential Monte Carlo, to accelerate GW inference [72,117]. An important future direction for this work is in automating the widening procedure, in order to improve both the robustness and the performance gains. Moreover, our method could be extended to explore precessing systems.

In summary, we have developed a robust algorithm for accelerating gravitational-wave parameter estimation in the high-SNR regime. This method focuses on speeding up the core nested sampling algorithm itself, meaning it can be used in conjunction with other established acceleration techniques for further gains. This work represents a substantial step toward preparing for the data analysis challenges of the next generation of gravitational-wave observatories and will help to realize the full scientific potential of future detections.

This work made use of NumPy [118], SciPy [83], `SIMPLE-PE` [42], `PESUMMARY` [119], `Bilby=2.3.0` [56,120], `ANESTHETIC` [121] and `MARGARINE` [80,81]. Plots were prepared with Matplotlib [122] and `ANESTHETIC` [121].

ACKNOWLEDGMENTS

We thank Sama Al-Shammari for helpful comments during the LIGO-Virgo-KAGRA internal review. We additionally thank Stephen Fairhurst, Laura Nuttall and Mukesh Singh for valuable discussions, and Stephen Fairhurst and Mukesh Singh for their continued development of `simple-pe` (alongside C.H). M. P was supported by the research environment and infrastructure of the Handley Lab at the University of Cambridge. C. H thanks the UKRI Future Leaders Fellowship for support through the Grant No. MR/T01881X/1, and the University of Portsmouth for support through the Dennis Sciamia Fellowship. M.J. W acknowledges support from Grants No. ST/X002225/1, No. ST/Y004876/1 and the University of Portsmouth.

This work was performed using the Cambridge Service for Data Driven Discovery (CSD3), part of which is operated by the University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility [[123]]. The DiRAC component of CSD3 was funded by BEIS capital funding via STFC Capital Grants No. ST/P002307/1 and No. ST/R002452/1 and STFC Operations Grants No. ST/R00689X/1. DiRAC is part of the National e-Infrastructure.

DATA AVAILABILITY

The data that support the findings of this article are openly available [96].

APPENDIX A: PRIORS

The prior distributions used for the main nested sampling analyses, detailed in Table III, are chosen to reflect the common priors used by the LVK [116] and such that the gravitational-wave signal fits within the specified 8s data duration. This is commonly done in the literature (see e.g. [124]). This includes priors that are uniform in the component masses, though the sampler itself explores the parameter space in terms of chirp mass and mass ratio. In contrast, the `simple-pe` analysis that initializes our method uses priors uniform in chirp mass and mass ratio directly. Similarly, the sampling priors on the spins are distributed uniformly over the sphere, whereas `simple-pe` uses a uniform prior on the effective aligned spin of the binary.

While this mismatch does not affect the statistical validity of the final posterior, our investigations show that it may have an impact on the computational efficiency of the method. The initial posterior approximation from `simple-pe` is conditioned on its own prior. Therefore, when the final analysis uses a different prior, the `simple-pe` guess can be more offset from the target posterior. This may necessitate more significant widening of the repartitioned prior and results in a larger effective KL divergence for the sampler to traverse, reducing the overall speedup.

Improved performance could be achieved in future work by using consistent priors throughout the method. This could be accomplished by adapting `simple-pe` to use component-mass priors or by implementing a prior reweighting scheme for the initial samples before they are used to train the normalizing flow.

APPENDIX B: EMPIRICAL CALIBRATION OF THE WIDENING FACTORS

To ensure the repartitioned prior robustly encompasses the true posterior, the initial distribution from `simple-pe` must be widened. We detail here the empirical procedure used to calibrate the widening factors applied throughout this work.

The calibration was performed using the 100-injection set described in Sec. IV. For each injection, we determined

the minimum widening required for the repartitioned prior to be statistically consistent with the posterior from the standard NS analysis. The procedure was as follows. First, a normalizing flow was trained on the initial `simple-pe` samples. Then, the standard deviation of each diagonal element of the flow’s base covariance matrix was systematically increased. This process was repeated until the 1st and 99th percentiles of the standard NS posterior were contained within the 1st and 99th percentiles of the distribution generated by the widened flow. This calibration was performed independently for each repartitioned parameter.

The minimum required widening factor for each parameter is shown as a function of network SNR in Fig. 12. The results show a clear dependence on SNR for several parameters, particularly the chirp mass and spins, with a notable change in behavior around an SNR of 20. For many signals, especially those above SNR 10, little to no widening was required. Based on these results, we adopted the two sets of fixed, conservative widening factors detailed in Table I.

These widening factors were chosen to be robust against the small, expected variations between the `simple-pe` approximation and the true posterior. We note that these factors are not tuned to accommodate rare outlier cases where the initial guess is significantly offset (as in the examples discussed in Sec. IV). A more advanced implementation could employ an automated widening procedure, where the widening factors themselves are sampled during the analysis. This would be robust to outliers, applying large widening only when needed, while retaining higher efficiency for typical cases where minimal widening is sufficient. We leave the development of such a widening procedure for future work.

APPENDIX C: KL DIVERGENCES AND PR SPEEDUPS

The performance scaling of our method, presented in Sec. V, can be understood better through considering the Kullback-Leibler (KL) divergence. In the context of nested sampling, the KL divergence between the sampling prior and the posterior, $D_{\text{KL}}(\mathcal{P}||\pi)$, quantifies the information gain, or the “amount of compression,” required for the sampler to converge from the prior to the posterior [38,43,68]. A larger KL divergence generally corresponds to a longer and more computationally expensive nested sampling run. By design, our method aims to reduce this divergence by replacing the broad original prior with a more compact repartitioned prior that is already a closer approximation of the posterior.

Figure 13 plots the change in the KL divergence between our method and the standard NS analysis for the SNR scaling study. The results provide further insight into the observed performance. At SNR 10, the substantial widening applied to the already broad `simple-pe` guess results

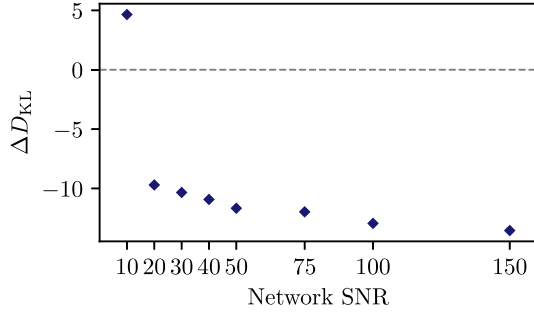


FIG. 13. The change in the KL divergence between the sampling prior and posterior for our method relative to the standard NS analysis. A negative value indicates a reduction in the KL divergence for our method and an expected speedup. The plot shows that the method increases the required ‘compression’ at SNR 10 but successfully reduces it for all higher SNRs, providing an explanation for the performance trend observed in Sec. V.

in a repartitioned prior that is a poorer representation of the posterior than the original prior. This causes the KL divergence to increase, leading to the performance reduction reported in Sec. V. For $\text{SNR} \geq 20$, our method

successfully reduces the KL divergence. However, at SNR 20 specifically, this gain is offset by the sampling inefficiency introduced by the missed secondary posterior mode, which explains the slight performance reduction in this case. For $\text{SNR} \geq 30$, the performance gains are driven directly by the reduction in KL divergence. The growing difference in KL divergence between our method and the standard analysis as SNR increases is the underlying reason for the trend of increasing speedups.

APPENDIX D: JENSEN-SHANNON DIVERGENCES

To quantitatively assess the statistical consistency between the posteriors from standard NS and our `simple-pe-PR` method, we compute the Jensen-Shannon divergence (JSD) for the one-dimensional marginal distributions. The JSD is a measure of the distance between two probability distributions, taking values between 0 and $\ln 2$ (1) nats (bits); 0 represents exactly identical distributions and a larger JSD indicates greater disagreement.

To estimate the mean and uncertainty of the JSD for each comparison, we generate 100 posterior realizations per

TABLE IV. Jensen-Shannon divergences for marginal posterior distributions between standard NS and our method, from the SNR scaling study. The quoted values are in units of 1×10^{-3} nats. 100 different realizations of 5000 posterior samples were generated for each analysis by resampling the nested sampling weights and performing rejecting sampling again. The mean and 1σ quantiles are shown, where for injections above SNR 10 the JSD values are within the threshold of 0.05 bits, or 35 millinats [126]. There are significant disagreements for the SNR 10 signal for the mass ratio parameter, but we do not recommend the application of our method in this SNR regime.

	SNR								
	10	20	30	40	50	75	100	150	
$\mathcal{M}$	1.49 ^{0.61} _{0.53}	1.04 ^{0.24} _{0.25}	1.22 ^{0.38} _{0.30}	0.84 ^{0.26} _{0.21}	0.84 ^{0.21} _{0.20}	1.03 ^{0.37} _{0.24}	0.95 ^{0.32} _{0.24}	0.95 ^{0.33} _{0.26}	
q	51.85 ^{3.72} _{2.95}	3.79 ^{0.65} _{0.65}	0.75 ^{0.24} _{0.20}	1.03 ^{0.43} _{0.29}	0.68 ^{0.23} _{0.23}	0.66 ^{0.22} _{0.15}	0.93 ^{0.37} _{0.17}	2.00 ^{0.44} _{0.46}	
α	2.86 ^{0.81} _{0.60}	1.41 ^{0.40} _{0.37}	0.73 ^{0.24} _{0.19}	0.69 ^{0.22} _{0.18}	0.81 ^{0.27} _{0.26}	0.65 ^{0.18} _{0.17}	0.67 ^{0.23} _{0.21}	0.62 ^{0.22} _{0.13}	
δ	4.62 ^{0.77} _{0.74}	2.58 ^{0.49} _{0.40}	0.68 ^{0.22} _{0.23}	0.63 ^{0.25} _{0.21}	0.59 ^{0.26} _{0.15}	0.71 ^{0.22} _{0.20}	0.68 ^{0.22} _{0.19}	0.75 ^{0.28} _{0.25}	
θ_{JN}	6.27 ^{0.92} _{0.96}	3.39 ^{0.58} _{0.62}	0.79 ^{0.34} _{0.23}	0.64 ^{0.23} _{0.22}	0.65 ^{0.23} _{0.17}	0.86 ^{0.27} _{0.24}	0.62 ^{0.21} _{0.19}	0.67 ^{0.23} _{0.15}	
ψ	1.80 ^{0.66} _{0.57}	1.05 ^{0.19} _{0.17}	0.75 ^{0.25} _{0.21}	0.78 ^{0.27} _{0.14}	0.59 ^{0.21} _{0.16}	0.75 ^{0.25} _{0.23}	0.75 ^{0.24} _{0.11}	0.73 ^{0.23} _{0.18}	
ϕ	2.40 ^{0.71} _{0.45}	0.50 ^{0.30} _{0.18}	0.24 ^{0.12} _{0.12}	0.24 ^{0.18} _{0.10}	0.15 ^{0.09} _{0.07}	0.12 ^{0.14} _{0.06}	0.09 ^{0.11} _{0.05}	0.05 ^{0.05} _{0.03}	
χ_1	24.40 ^{2.92} _{2.68}	4.02 ^{0.70} _{0.56}	0.69 ^{0.27} _{0.19}	0.85 ^{0.26} _{0.23}	0.81 ^{0.29} _{0.29}	0.98 ^{0.38} _{0.27}	1.03 ^{0.23} _{0.23}	2.64 ^{0.55} _{0.51}	
χ_2	1.24 ^{0.28} _{0.29}	0.76 ^{0.26} _{0.21}	1.20 ^{0.34} _{0.31}	1.62 ^{0.36} _{0.40}	1.11 ^{0.31} _{0.21}	0.94 ^{0.31} _{0.22}	1.69 ^{0.43} _{0.37}	1.75 ^{0.40} _{0.32}	
t_c	16.56 ^{2.01} _{1.98}	0.93 ^{0.35} _{0.22}	0.78 ^{0.26} _{0.18}	0.78 ^{0.24} _{0.17}	0.88 ^{0.26} _{0.23}	0.75 ^{0.18} _{0.13}	1.04 ^{0.38} _{0.25}	1.12 ^{0.31} _{0.25}	

JSD divergence [$\times 10^{-3}$ nat]

analysis. This is achieved by resampling the nested sampling weights associated with the full set of dead points, using the functionality in ANESTHETIC [121]. From each set of newly weighted points, a final realization of 5000 posterior samples is produced via rejection sampling. The JSD is then calculated between kernel density estimates of these marginal distributions, constructed using PESUMMARY [125]. JSD values below 0.05 bits, or 35 millinats, indicate that the distributions are statistically consistent [126].

The JSD values for the SNR scaling study are presented in Table IV. For all signals with $\text{SNR} > 10$, the JSDs are below this threshold, confirming excellent agreement between the methods. As expected from the discussion in Sec. V, the SNR 10 signal exhibits a JSD value above the

threshold, for the mass ratio parameter, providing a quantitative measure of the observed disagreement.

The JSD values for the full injection study are presented in Fig. 14. The results show a clear trend with SNR and provide quantitative support for the conclusions drawn in the main text. Below SNR 10, several injections exhibit JSDs above the 35 millinat threshold for consistency. This is expected, as our empirical widening factors were not calibrated for this low-SNR regime and can be insufficient to fully encompass the true posterior. In the $10 \leq \text{SNR} < 20$ regime, the agreement improves significantly, though a few analyses still produce posteriors that are statistically distinguishable from the standard result. For signals with $\text{SNR} > 20$, only two injections fail the JSD test. These correspond to the specific outlier cases identified in Section IV, where a significant offset in the initial `simple-pe` guess was not fully corrected by the widening procedure. An automated widening scheme would likely

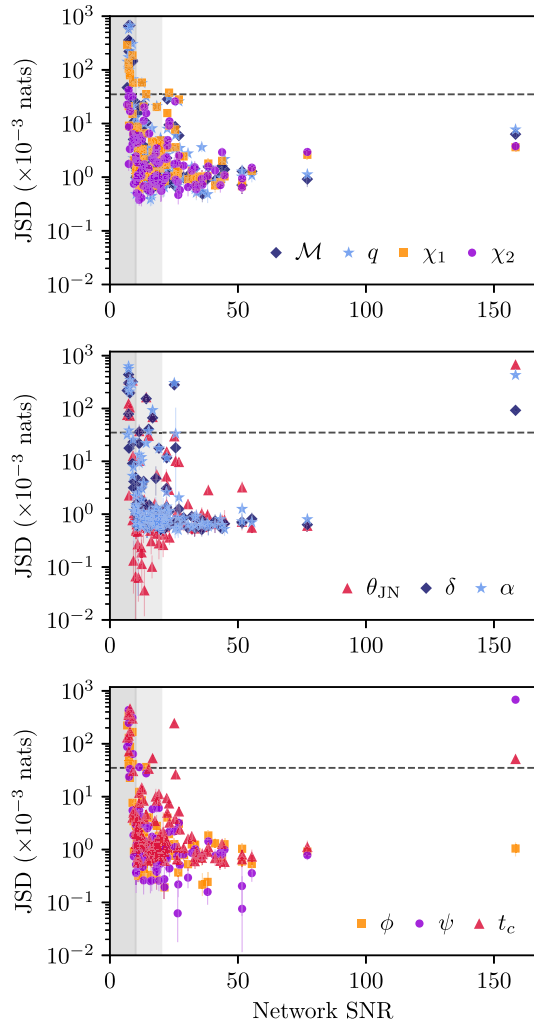


FIG. 14. The Jensen-Shannon divergence (JSD) between the one-dimensional marginal posteriors from our method and the standard NS analysis for each of the 100 injections. Each point represents the JSD for a single parameter and for a single event. The horizontal dashed line at 35 millinats represents the threshold for statistical consistency [126]. The plot shows a clear trend of improved agreement at higher SNR.

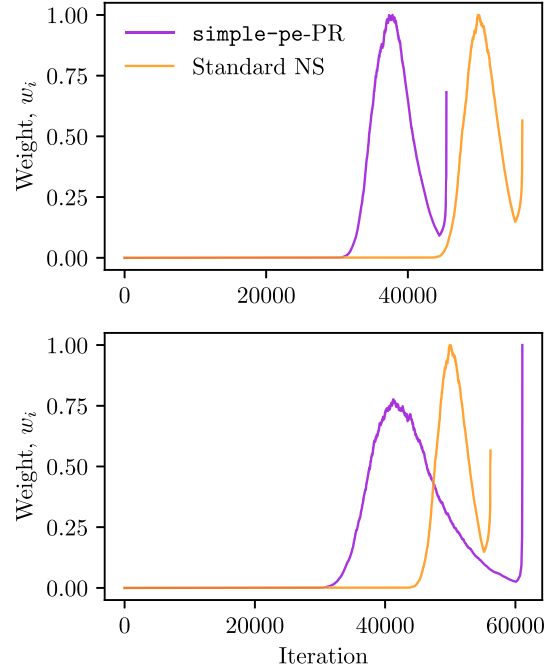


FIG. 15. Comparison of the posterior weight distribution of the dead points for a standard NS analysis (orange) and our `simple-pe-PR` method (violet). The x-axis represents the nested sampling iteration, and the y-axis is the posterior weight. The top panel shows a typical case where the PR analysis enters the high-likelihood region (where weights become non-negligible) earlier but the phase of significant weight extends over slightly more iterations, resulting in a larger effective sample size. The bottom panel shows a more extreme case where the compression through the posterior is much slower for the PR analysis, leading to a substantially larger effective sample size at the cost of a longer runtime. Both cases demonstrate that a direct comparison of runtimes at a fixed number of live points is misleading, and the per-sample speedup is a more informative metric.

resolve these remaining discrepancies. These results provide quantitative evidence for our recommendation to apply the method to signals with $\text{SNR} > 20$, the regime where it is both computationally efficient and statistically robust.

APPENDIX E: EFFECTIVE SAMPLE SIZE

The effective sample size from a nested sampling analysis can be obtained using the set of nested sampling weights, w_i , for each dead point, i . This can be calculated as [98]

$$n_{\text{eff}} = \frac{(\sum_i w_i)^2}{\sum_i w_i^2}. \quad (\text{E1})$$

Posterior repartitioned NS typically spends a longer time exploring the bulk of the posterior than standard NS for a given number of live points (see Fig. 15). This can be explained by PR having a greater density of points in this region, due to the smaller effective prior volume. Thus, the rate of prior volume compression occurs more slowly in this region, generating a larger number of dead points with

non-negligible weights, and therefore resulting in a larger effective sample size. As a result, our method should be performed with fewer live points than a standard NS analysis in order to obtain a given final sample size.

APPENDIX F: P-P PLOTS

The P-P plots obtained from standard NS and `simple-pe-PR` on both the injection study set and the calibration set used to determine the necessary widening factors are included here for reference (Figs. 16–18). We note that in the case of the calibration set, neither P-P plot shows the same biases in declination and time of coalescence at the geocentre as seen for the injection set. This supports our hypothesis that the biases seen for the injection set can be attributed to statistical fluctuations in drawing the injections. In both cases, the P-P plot from our method is consistent with the P-P plot from standard NS, indicating agreement between the two methods. Finally, the difference plots for the remaining parameters not shown in the main text are also included here (Fig. 19).

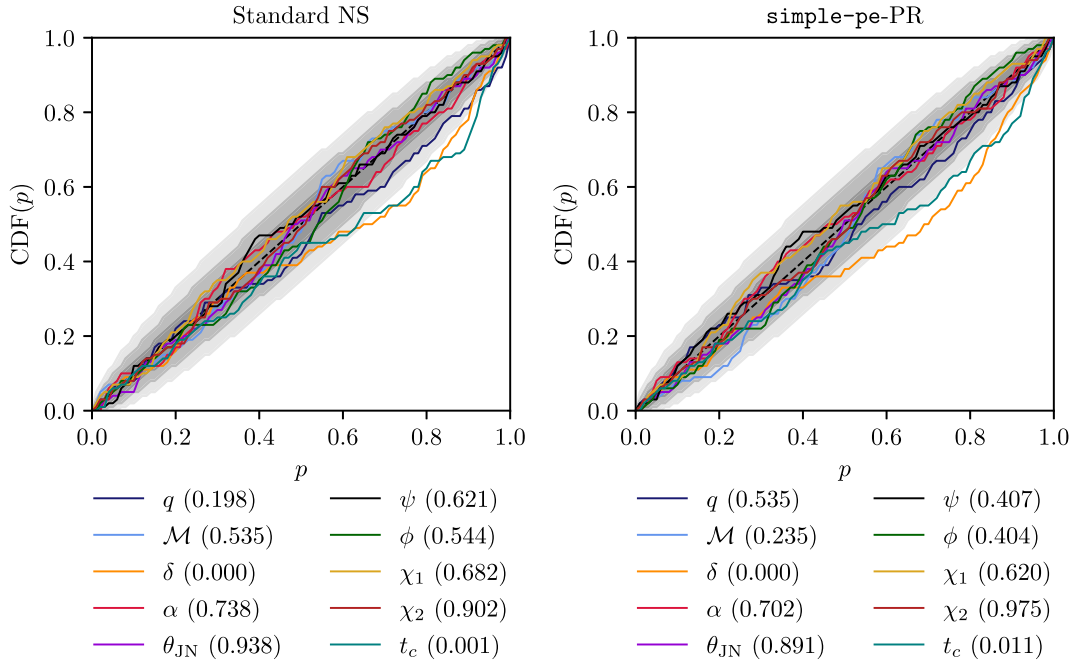


FIG. 16. Probability-probability (P-P) plot for the 100 injections in our validation study, analyzed using both standard NS and `simple-pe-PR`. Each colored line represents the cumulative distribution function (CDF) of the posterior quantiles of the true injected values for a specific parameter. The diagonal dashed line represents a perfect uniform distribution, while the shaded gray regions indicate the 1σ , 2σ and 3σ confidence intervals. The p-values from a Kolmogorov-Smirnov test for each parameter are shown in the legend. We observe biases in the declination and time of coalescence at the geocentre for both methods, which we attribute to statistical fluctuations in drawing our injections, since this bias was not observed in the calibration set. Importantly, our algorithm produces percentiles consistent with those from standard NS, demonstrating agreement between the two methods on a population level.

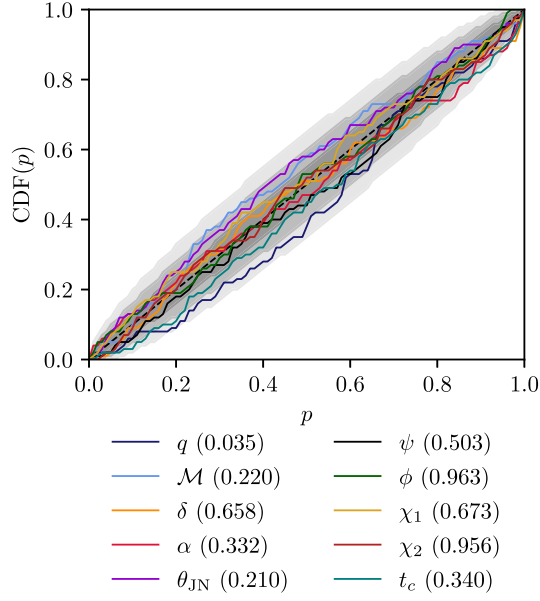


FIG. 17. P-P plot obtained from the standard nested sampling analyses performed on our calibration set from Section IV A using the Dynesty sampler.

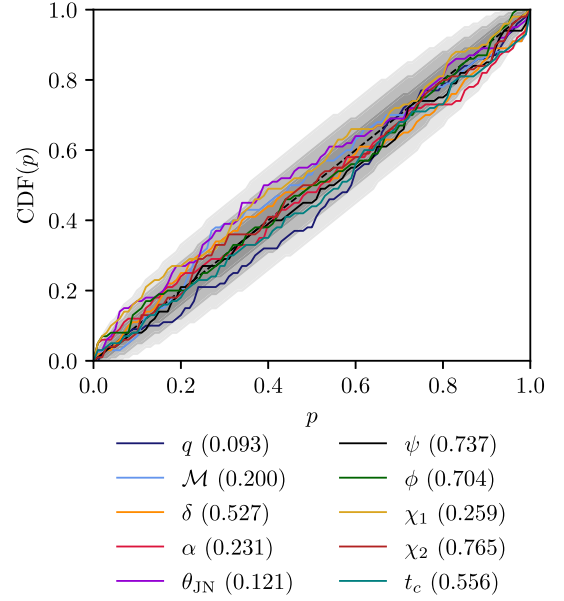


FIG. 18. P-P plot obtained from the simple-pe-PR analyses performed on our calibration set from Section IV A using the Dynesty sampler, modified to implement PR.

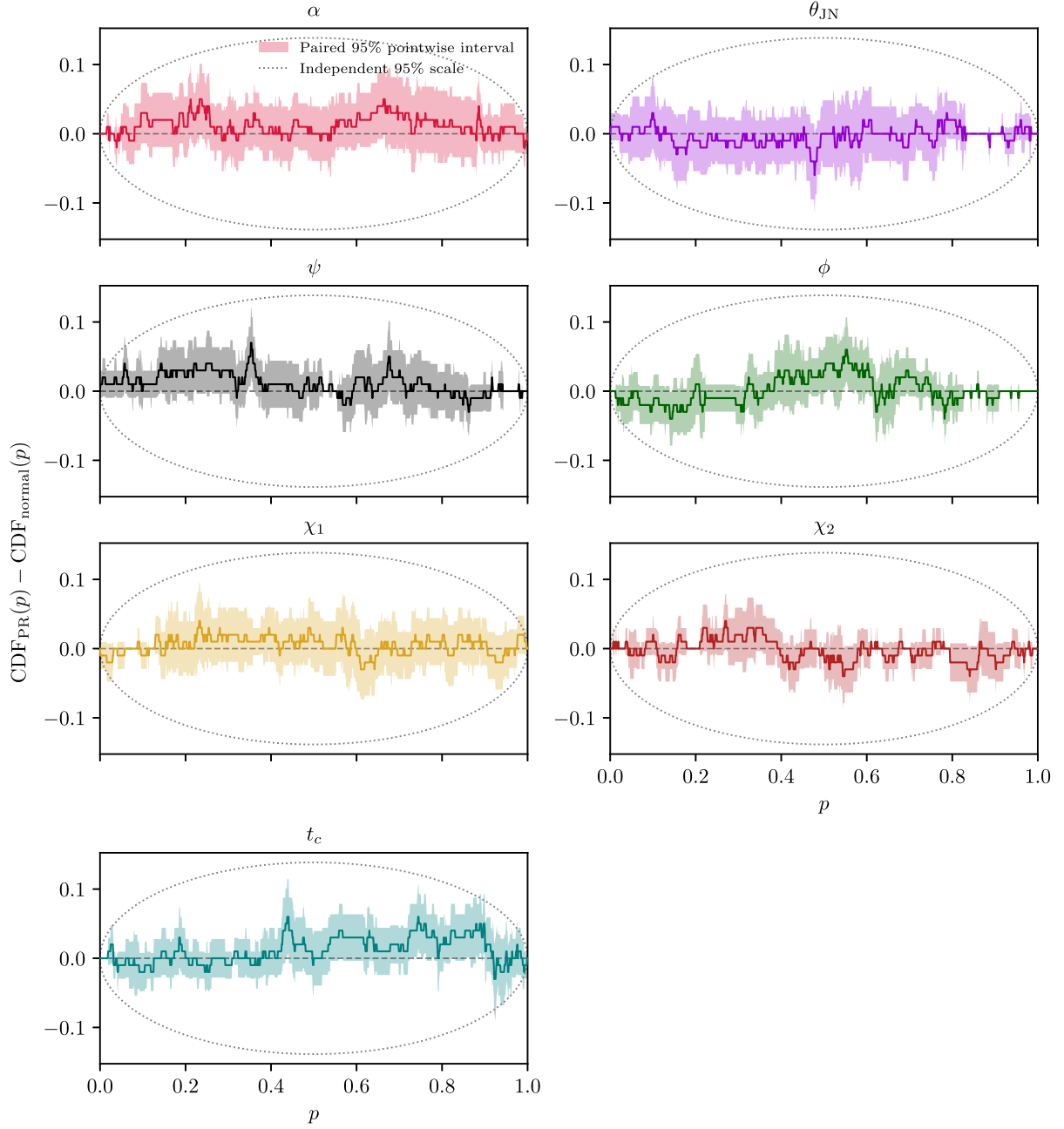


FIG. 19. The differences between the P-P plots obtained from simple-pe-PR and normal NS are shown for the remaining parameters (solid). Perfect agreement corresponds to a horizontal line at zero, but there is statistical uncertainty; within the error bars (shaded), both methods agree at the population level, including the time of coalescence at the geocentre parameter, which exhibited small biases in the P-P plots from both methods.

- [1] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), GWTC-4.0: Methods for identifying and characterizing gravitational-wave transients, [arXiv:2508.18081](#).
- [2] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), GWTC-4.0: Population properties of merging compact binaries, [arXiv:2508.18083](#).
- [3] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), GWTC-4.0: Constraints on the cosmic expansion rate and modified gravitational-wave propagation, [arXiv:2509.04348](#).
- [4] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), Tests of general relativity with GWTC-3, [Phys. Rev. D **112**, 084080 \(2025\)](#).
- [5] A. Ray and V. Kalogera, Reexamining evidence of a pair-instability mass gap in the binary black hole population, [arXiv:2510.18867](#).
- [6] A. Hussain, M. Isi, and A. Zimmerman, Living on the edge: Testing for compact population features at the edges of parameter space, [arXiv:2510.20010](#).
- [7] H. Tong, T. A. Callister, M. Fishbach, E. Thrane, F. Antonini, S. Stevenson, I. M. Romero-Shaw, and F. Dosopoulou, A subpopulation of low-mass, spinning black holes: Signatures of dynamical assembly, [arXiv:2511.05316](#).
- [8] N. V. Krishnendu, F. Ohme, and K. G. Arun, Testing the nature of compact objects in the lower mass gap using gravitational wave observations, [arXiv:2509.10420](#).
- [9] C. Adamcewicz, N. Guttman, P. D. Lasky, and E. Thrane, Do both black holes spin in merging binaries? Evidence from GWTC-4 and astrophysical implications, [arXiv:2509.04706](#).
- [10] J. Stegmann and J. Klencki, Orbital eccentricity and spin-orbit misalignment are evidence that neutron star–black hole mergers form through triple star evolution, [Astrophys. J. Lett. **991**, L54 \(2025\)](#).
- [11] The LIGO Scientific Collaboration, Advanced LIGO, [Classical Quantum Gravity **32**, 074001 \(2015\)](#).
- [12] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin *et al.*, Advanced Virgo: A second-generation interferometric gravitational wave detector, [Classical Quantum Gravity **32**, 024001 \(2014\)](#).
- [13] B. P. Abbott *et al.* (KAGRA, LIGO Scientific, and Virgo Collaborations), Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA, [Living Rev. Relativity **23**, 3 \(2020\)](#).
- [14] J. Aasi *et al.* (LIGO Scientific Collaboration), Advanced LIGO, [Classical Quantum Gravity **32**, 074001 \(2015\)](#).
- [15] F. Acernese *et al.* (Virgo Collaboration), Advanced Virgo: A second-generation interferometric gravitational wave detector, [Classical Quantum Gravity **32**, 024001 \(2015\)](#).
- [16] T. Akutsu *et al.* (KAGRA Collaboration), Overview of KAGRA: Detector design and construction history, [Prog. Theor. Exp. Phys. **2021**, 05A101 \(2021\)](#).
- [17] B. Iyer *et al.*, IndIGO and LIGO-India: Scope and plans for gravitational wave research and precision metrology in India, [arXiv:1510.06059](#).
- [18] J. Lange, R. O’Shaughnessy, and M. Rizzo, Rapid and accurate parameter inference for coalescing, precessing compact binaries, [arXiv:1805.10457](#).
- [19] J. Roulet, J. Mushkin, D. Wadekar, T. Venumadhav, B. Zackay, and M. Zaldarriaga, Fast marginalization algorithm for optimizing gravitational wave detection, parameter estimation, and sky localization, [Phys. Rev. D **110**, 044010 \(2024\)](#).
- [20] C. Pankow, P. Brady, E. Ochsner, and R. O’Shaughnessy, Novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences, [Phys. Rev. D **92**, 023002 \(2015\)](#).
- [21] L. Pathak, A. Reza, and A. S. Sengupta, Fast likelihood evaluation using meshfree approximations for reconstructing compact binary sources, [Phys. Rev. D **108**, 064055 \(2023\)](#).
- [22] V. Tiwari, C. Hoy, S. Fairhurst, and D. MacLeod, Fast non-Markovian sampler for estimating gravitational-wave posteriors, [Phys. Rev. D **108**, 023001 \(2023\)](#).
- [23] A. H. Nitz, Robust, rapid, and simple gravitational-wave parameter estimation, [Phys. Rev. D **112**, 023032 \(2025\)](#).
- [24] A. Sharma, L. Pathak, S. Roy, and A. S. Sengupta, Rapid parameter estimation with the full symphony of compact binary mergers using meshfree approximation, [arXiv:2508.04172](#).
- [25] J. Mushkin, J. Roulet, B. Zackay, T. Venumadhav, O. Ivashtenko, D. Wadekar, A. K. Mehta, and M. Zaldarriaga, Sampler-free gravitational wave inference using matrix multiplication, [Phys. Rev. D **112**, 104025 \(2025\)](#).
- [26] H.-T. Wang, A robust and efficient F-statistic-based framework for consistent bayesian inference of compact binary coalescences, [arXiv:2509.14849](#).
- [27] H. Gabbard, C. Messenger, I. S. Heng, F. Tonolini, and R. Murray-Smith, Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy, [Nat. Phys. **18**, 112 \(2022\)](#).
- [28] S. R. Green and J. Gair, Complete parameter inference for GW150914 using deep learning, [Mach. Learn. Sci. Tech. **2**, 03LT01 \(2021\)](#).
- [29] S. R. Green, C. Simpson, and J. Gair, Gravitational-wave parameter estimation with autoregressive neural network flows, [Phys. Rev. D **102**, 104057 \(2020\)](#).
- [30] M. Dax, S. R. Green, J. Gair, J. H. Macke, A. Buonanno, and B. Schölkopf, Real-time gravitational wave science with neural posterior estimation, [Phys. Rev. Lett. **127**, 241103 \(2021\)](#).
- [31] M. Dax, S. R. Green, J. Gair, N. Gupte, M. Pürrer, V. Raymond, J. Wildberger, J. H. Macke, A. Buonanno, and B. Schölkopf, Real-time inference for binary neutron star mergers using machine learning, [Nature \(London\) **639**, 49 \(2025\)](#).
- [32] V. Raymond, S. Al-Shammari, and A. Göttel, Simulation-based inference for gravitational-waves from intermediate-mass binary black holes in real noise, [Mon. Not. R. Astron. Soc. **542**, 1108 \(2025\)](#).
- [33] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), GW250114: Testing Hawking’s area law and the Kerr nature of black holes, [Phys. Rev. Lett. **135**, 111403 \(2025\)](#).

- [34] The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, Black hole spectroscopy and tests of general relativity with GW250114, [arXiv:2509.08099](#).
- [35] E. Thrane and C. Talbot, An introduction to bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models, *Pub. Astron. Soc. Aust.* **36**, e010 (2019).
- [36] J. Veitch *et al.*, Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library, *Phys. Rev. D* **91**, 042003 (2015).
- [37] J. Skilling, Nested sampling, *AIP Conf. Proc.* **735**, 395 (2004).
- [38] J. Skilling, Nested sampling for general Bayesian computation, *Bayesian Anal.* **1**, 833 (2006).
- [39] B. P. Abbott *et al.*, A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals, *Classical Quantum Gravity* **37**, 055002 (2020).
- [40] X. Chen, M. Hobson, S. Das, and P. Gelderblom, Improving the efficiency and robustness of nested sampling using posterior repartitioning, [arXiv:1803.06387](#).
- [41] X. Chen, F. Feroz, and M. Hobson, Bayesian posterior repartitioning for nested sampling, [arXiv:1908.04655](#).
- [42] S. Fairhurst, C. Hoy, R. Green, C. Mills, and S. A. Usman, Simple parameter estimation using observable features of gravitational-wave signals, [arXiv:2304.03731](#).
- [43] M. Prathaban, H. Bevens, and W. Handley, Accelerated nested sampling with posterior repartitioning and β -flows for gravitational waves, *Mon. Not. R. Astron. Soc.* **541**, 200 (2025).
- [44] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Spin induced orbital precession and its modulation of the gravitational wave forms from merging binaries, *Phys. Rev. D* **49**, 6274 (1994).
- [45] J. N. Goldberg, A. J. MacFarlane, E. T. Newman, F. Rohrlich, and E. C. G. Sudarshan, Spin- s spherical harmonics and δ , *J. Math. Phys. (N.Y.)* **8**, 2155 (1967).
- [46] K. S. Thorne, Multipole expansions of gravitational radiation, *Rev. Mod. Phys.* **52**, 299 (1980).
- [47] N. J. Cornish, Heterodyned likelihood for rapid gravitational wave parameter inference, *Phys. Rev. D* **104**, 104054 (2021).
- [48] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, and M. Tiglio, Accelerated gravitational-wave parameter estimation with reduced order modeling, *Phys. Rev. Lett.* **114**, 071104 (2015).
- [49] S. Vinciguerra, J. Veitch, and I. Mandel, Accelerating gravitational wave parameter estimation with multi-band template interpolation, *Classical Quantum Gravity* **34**, 115006 (2017).
- [50] S. Morisaki, Accelerating parameter estimation of gravitational waves from compact binary coalescence using adaptive frequency resolutions, *Phys. Rev. D* **104**, 044062 (2021).
- [51] G. Ashton and T. Dietrich, The use of hypermodels to understand binary neutron star collisions, *Nat. Astron.* **6**, 961 (2022).
- [52] C. Hoy, Accelerating multimodel Bayesian inference, model selection, and systematic studies for gravitational wave astronomy, *Phys. Rev. D* **106**, 083003 (2022).
- [53] C. Hoy, S. Akcay, J. Mac Uilliam, and J. E. Thompson, Incorporation of model accuracy in gravitational wave Bayesian inference, *Nat. Astron.* **9**, 1256 (2025).
- [54] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57**, 97 (1970).
- [55] G. Ashton and C. Talbot, bilby-MCMC: An MCMC sampler for gravitational-wave inference, *Mon. Not. R. Astron. Soc.* **507**, 2037 (2021).
- [56] G. Ashton *et al.*, bilby: A user-friendly Bayesian inference library for gravitational-wave astronomy, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [57] J. S. Speagle, Dynesty: A dynamic nested sampling package for estimating Bayesian posteriors and evidences, *Mon. Not. R. Astron. Soc.* **493**, 3132 (2020).
- [58] M. J. Williams, J. Veitch, and C. Messenger, Nested sampling with normalizing flows for gravitational-wave inference, *Phys. Rev. D* **103**, 103006 (2021).
- [59] R. J. E. Smith, G. Ashton, A. Vajpeyi, and C. Talbot, Massively parallel Bayesian inference for transient gravitational-wave astronomy, *Mon. Not. R. Astron. Soc.* **498**, 4492 (2020).
- [60] N. J. Cornish, Fast Fisher matrices and lazy likelihoods, [arXiv:1007.4820](#).
- [61] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt, Fast and accurate inference on gravitational waves from precessing compact binaries, *Phys. Rev. D* **94**, 044031 (2016).
- [62] S. Vinciguerra, J. Veitch, and I. Mandel, Accelerating gravitational wave parameter estimation with multi-band template interpolation, *Classical Quantum Gravity* **34**, 115006 (2017).
- [63] B. Zackay, L. Dai, and T. Venumadhav, Relative binning and fast likelihood evaluation for gravitational wave parameter estimation, [arXiv:1806.08792](#).
- [64] S. Morisaki, Accelerating parameter estimation of gravitational waves from compact binary coalescence using adaptive frequency resolutions, *Phys. Rev. D* **104**, 044062 (2021).
- [65] S. Morisaki, R. Smith, L. Tsukada, S. Sachdev, S. Stevenson, C. Talbot, and A. Zimmerman, Rapid localization and inference on compact binary coalescences with the Advanced LIGO–Virgo–KAGRA gravitational-wave detector network, *Phys. Rev. D* **108**, 123040 (2023).
- [66] Q. Hu and J. Veitch, Costs of Bayesian parameter estimation in third-generation gravitational wave detectors: A review of acceleration methods, [arXiv:2412.02651](#).
- [67] C. Hoy and L. K. Nuttall, bilby in space: Bayesian inference for transient gravitational-wave signals observed with LISA, *Mon. Not. R. Astron. Soc.* **529**, 3052 (2024).
- [68] A. Petrosyan and W. Handley, Supernest: Accelerated nested sampling applied to astrophysics and cosmology *Phys. Sci. Forum* **5**, 51 (2022).
- [69] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan, Normalizing flows for probabilistic modeling and inference, [arXiv:1912.02762](#).

- [70] I. Kobyzev, S. J. Prince, and M. A. Brubaker, Normalizing flows: An introduction and review of current methods, *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 3964 (2021).
- [71] N. E. Wolfe, C. Talbot, and J. Golomb, Accelerating tests of general relativity with gravitational-wave signals using hybrid sampling, *Phys. Rev. D* **107**, 104056 (2023).
- [72] M. J. Williams, Accelerated sequential posterior inference via reuse for gravitational-wave analyses, [arXiv:2511.04218](https://arxiv.org/abs/2511.04218).
- [73] C. Hoy, C. R. Weaving, L. K. Nuttall, and I. Harry, A rapid multi-modal parameter estimation technique for LISA, *Classical Quantum Gravity* **41**, 245012 (2024).
- [74] B. G. Patterson, S. M. Tomson, and S. Fairhurst, Identifying eccentricity in binary black hole mergers using a harmonic decomposition of the gravitational waveform, *Phys. Rev. D* **111**, 044073 (2025).
- [75] A. H. Nitz, T. Dal Canton, D. Davis, and S. Reyes, Rapid detection of gravitational waves from compact binary mergers with PyCBC live, *Phys. Rev. D* **98**, 024050 (2018).
- [76] B. J. Owen, Search templates for gravitational waves from inspiraling binaries: Choice of template spacing, *Phys. Rev. D* **53**, 6749 (1996).
- [77] S. Fairhurst, R. Green, C. Hoy, M. Hannam, and A. Muir, Two-harmonic approximation for gravitational waveforms from precessing binaries, *Phys. Rev. D* **102**, 024055 (2020).
- [78] C. McIsaac, C. Hoy, and I. Harry, Search technique to observe precessing compact binary mergers in the advanced detector era, *Phys. Rev. D* **108**, 123016 (2023).
- [79] R. Ursell, C. Hoy, I. Harry, and L. K. Nuttall, The missing multipole problem: Investigating biases from model starting frequency in gravitational-wave analyses, [arXiv:2510.15048](https://arxiv.org/abs/2510.15048).
- [80] H. Bevens, W. Handley, P. Lemos, P. Sims, E. de Lera Acedo, and A. Fialkov, Marginal Bayesian statistics using masked autoregressive flows and kernel density estimators with examples in cosmology, [arXiv:2207.11457](https://arxiv.org/abs/2207.11457).
- [81] H. T. J. Bevens, W. J. Handley, P. Lemos, P. H. Sims, E. de Lera Acedo, A. Fialkov, and J. Alsing, Marginal post-processing of Bayesian inference products with normalizing flows and kernel density estimators, *Mon. Not. R. Astron. Soc.* **526**, 4613 (2023).
- [82] M. Prathaban and H. Bevens, *margarine_unbounded*: A fork of margarine with support for unbounded distributions, https://github.com/mrosep/margarine_unbounded (2026).
- [83] P. Virtanen *et al.* (SciPy 1.0 Contributors), SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nat. Methods* **17**, 261 (2020).
- [84] V. D’Emilio, R. Green, and V. Raymond, Density estimation with Gaussian processes for gravitational wave posteriors, *Mon. Not. R. Astron. Soc.* **508**, 2090 (2021).
- [85] We note that `simple-pe` is still in active development and it is possible that in the future this widening procedure will not be required to obtain robust results.
- [86] M. Prathaban, C. Hoy, and M. J. Williams, `bilby-pr`: v0.1.0 (2026).
- [87] Although we note that `simple-pe` is still in active development and this will likely be improved in future versions.
- [88] L. P. Singer and L. R. Price, Rapid Bayesian position reconstruction for gravitational-wave transients, *Phys. Rev. D* **93**, 024013 (2016).
- [89] L. P. Singer *et al.*, Going the distance: Mapping host galaxies of LIGO and Virgo sources in three dimensions using local cosmography and targeted Follow-up, *Astrophys. J. Lett.* **829**, L15 (2016).
- [90] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, Multimode frequency-domain model for the gravitational wave signal from nonprecessing black-hole binaries, *Phys. Rev. D* **102**, 064002 (2020).
- [91] LIGO Scientific Collaboration and Virgo Collaboration, Noise curves for use in simulations pre-o4 (2022), <https://dcc.ligo.org/T2200043-v3/public>.
- [92] I. M. Romero-Shaw *et al.*, Bayesian inference for compact binary coalescences with bilby: Validation and application to the first LIGO–Virgo gravitational-wave transient catalogue, *Mon. Not. R. Astron. Soc.* **499**, 3295 (2020).
- [93] A. G. Samantha R Cook and D. B. Rubin, Validation of software for Bayesian models using posterior quantiles, *J. Comput. Graph. Stat.* **15**, 675 (2006).
- [94] S. Fairhurst, Triangulation of gravitational wave sources with a network of detectors, *New J. Phys.* **11**, 123006 (2009).
- [95] M. Prathaban and W. Handley, Costless correction of chain based nested sampling parameter estimation in gravitational wave data and beyond, *Mon. Not. R. Astron. Soc.* **533**, 1839 (2024).
- [96] M. Prathaban, C. Hoy, and M. J. Williams, Leveraging simple-pe for efficient gravitational-wave parameter estimation via posterior repartitioning (2025), available at <https://zenodo.org/record/17776522>.
- [97] J. Lin, Divergence measures based on the shannon entropy, *IEEE Trans. Inf. Theory* **37**, 145 (1991).
- [98] L. Kish, *Survey Sampling*, 3rd ed. (Wiley-Interscience, Oxford, England, 1995).
- [99] R. Abbott *et al.*, GW190412: Observation of a binary-black-hole coalescence with asymmetric masses, *Phys. Rev. D* **102**, 043015 (2020).
- [100] R. Abbott *et al.*, GW190521: A binary black hole merger with a total mass of $150m_{\odot}$, *Phys. Rev. Lett.* **125**, 101102 (2020).
- [101] A. G. Abac *et al.*, GW231123: A binary black hole merger with total mass $190\text{--}265 m_{\odot}$, *Astrophys. J. Lett.* **993**, L25 (2025).
- [102] A. Begnani, S. Anselmi, M. Pieroni, A. Renzi, and A. Ricciardone, Detectability and parameter estimation for Einstein telescope configurations with GWJulia, [arXiv:2506.21530](https://arxiv.org/abs/2506.21530).
- [103] A. Cabezas, A. Corenflos, J. Lao, and R. Louf, Blackjax: Composable Bayesian inference in JAX, [arXiv:2402.10797](https://arxiv.org/abs/2402.10797).
- [104] K. W. k. Wong, M. Gabrié, and D. Foreman-Mackey, flowMC: Normalizing flow enhanced sampling package for probabilistic inference in JAX, *J. Open Source Software* **8**, 5021 (2023).
- [105] J. G. Albert, Jaxns: A high-performance nested sampling package based on JAX, [arXiv:2012.15286](https://arxiv.org/abs/2012.15286).

- [106] D. Yallup, N. Kroupa, and W. Handley, Nested slice sampling, in *Frontiers in Probabilistic Inference: Learning Meets Sampling* (2025).
- [107] M. Punturo *et al.*, The third generation of gravitational wave observatories and their science reach, *Classical Quantum Gravity* **27**, 084007 (2010).
- [108] S. Hild *et al.*, Sensitivity studies for third-generation gravitational wave observatories, *Classical Quantum Gravity* **28**, 094013 (2011).
- [109] M. Maggiore *et al.* (ET Collaboration), Science case for the Einstein telescope, *J. Cosmol. Astropart. Phys.* **03** (2019) 050.
- [110] S. Babak, A. Petiteau, and M. Hewitson, LISA sensitivity and SNR calculations, [arXiv:2108.01167](https://arxiv.org/abs/2108.01167).
- [111] M. Colpi *et al.* (LISA Collaboration), LISA definition study report, [arXiv:2402.07571](https://arxiv.org/abs/2402.07571).
- [112] A. Ramos-Buades, A. Buonanno, M. Khalil, and S. Ossokine, Effective-one-body multipolar waveforms for eccentric binary black holes with nonprecessing spins, *Phys. Rev. D* **105**, 044035 (2022).
- [113] A. Gamboa *et al.*, Accurate waveforms for eccentric, aligned-spin binary black holes: The multipolar effective-one-body model seobnr5ehm, *Phys. Rev. D* **112**, 044038 (2025).
- [114] A. Nagar, R. Gamba, P. Rettengo, V. Fantini, and S. Bernuzzi, Effective-one-body waveform model for non-circularized, planar, coalescing black hole binaries: The importance of radiation reaction, *Phys. Rev. D* **110**, 084001 (2024).
- [115] M. d. L. Planas, A. Ramos-Buades, C. García-Quirós, H. Estellés, S. Husa, and M. Haney, Time-domain phenomenological multipolar waveforms for aligned-spin binary black holes in elliptical orbits, [arXiv:2503.13062](https://arxiv.org/abs/2503.13062).
- [116] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), GWTC-4.0: Updating the gravitational-wave transient catalog with observations from the first part of the fourth LIGO-Virgo-KAGRA observing run, [arXiv:2508.18082](https://arxiv.org/abs/2508.18082).
- [117] M. J. Williams, M. Karamanis, Y. Luo, and U. Seljak, Validating sequential Monte Carlo for gravitational-wave inference, *Mon. Not. R. Astron. Soc.* **543**, 1479 (2025).
- [118] C. R. Harris *et al.*, Array programming with NumPy, *Nature (London)* **585**, 357 (2020).
- [119] C. Hoy and V. Raymond, PESummary: The code agnostic parameter estimation summary page builder, *SoftwareX* **15**, 100765 (2021).
- [120] G. Ashton *et al.*, bilby: A user-friendly Bayesian inference library, <https://zenodo.org/records/14171800> (2024), version 2.4.0.
- [121] W. Handley, anesthetic: Nested sampling visualisation, *J. Open Source Software* **4**, 1414 (2019).
- [122] J. D. Hunter, Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).
- [123] <http://www.dirac.ac.uk>
- [124] E. Hamilton, M. Colleoni, J. E. Thompson, C. Hoy, A. Heffernan, M. Kinnear, J. Valencia, F. A. R. Vidal, C. García-Quirós, S. Ghosh, L. London, M. Hannam, and S. Husa, Phenomxpnr: An improved gravitational wave model linking precessing inspirals and nr-calibrated merger-ringdown, [arXiv:2507.02604](https://arxiv.org/abs/2507.02604).
- [125] C. Hoy and V. Raymond, Pesummary: The code agnostic parameter estimation summary page builder, *SoftwareX* **15**, 100765 (2021).
- [126] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-1: A gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs, *Phys. Rev. X* **9**, 031040 (2019).