

Reversible jump MCMC for model selection in the analysis of gravitational-wave signals from the coalescence of compact objects

Anna Puecher^{1,*}, Tim Dietrich^{1,2}, Hauke Koehn¹, Jonathan Gair²,
Gregory Ashton³, Luca Negri^{4,5} and Anuradha Samajdar^{4,5}

¹*Institut für Physik und Astronomie, Universität Potsdam, Haus 28, Karl-Liebknecht-Str. 24/25, 14476, Potsdam, Germany*

²*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, Potsdam 14476, Germany*

³*School of Mathematical Sciences, University of Southampton*

⁴*Institute for Gravitational and Subatomic Physics (GRASP), Utrecht University, Princetonplein 1, 3584 CC Utrecht, The Netherlands*

⁵*Nikhef, Science Park 105, 1098 XG Amsterdam, The Netherlands*



(Received 24 July 2026; accepted 11 September 2026; published 6 October 2026)

Bayesian inference is commonly employed in the analysis of gravitational-wave signals not only to estimate the source parameters, but also for model selection. The latter provides insight into the physics of the source and has the potential to inform the direction for future model development. Although model comparison is usually performed by analyzing the data separately with different models and comparing the obtained Bayesian evidences, an alternative approach consists in sampling directly over the model itself. Here, we present `t-roo`, a reversible jump Markov chain Monte Carlo sampler capable of performing transdimensional inference on gravitational-wave signals from compact binary coalescences. Employing `t-roo`, a single analysis provides simultaneously the model odds ratio and the parameter posteriors for the favored models, hence yielding a potentially substantial computational advantage, particularly when comparing many models or analyzing highly informative data. `t-roo` is built on the sampler `eryn` and is specifically designed to compare models describing different kinds of sources, i.e., binary black hole, binary neutron star, or neutron star-black hole systems, as well as multiple models for the same source class. We validate the sampler on a set of injections, finding agreement with the results obtained with the nested sampler `dynesty`. We then use `t-roo` to analyze the real events GW190425 and GW230529, for which the system's parameters alone do not provide conclusive evidence of the presence of a neutron star component. `t-roo` can be adapted to any model-comparison scenario, thus providing a valuable tool in particular for next-generation detectors, where analyzing data separately with competing models becomes computationally even more demanding.

DOI: [10.1103/k5ts-8cpr](https://doi.org/10.1103/k5ts-8cpr)

I. INTRODUCTION

In the past ten years, a growing number of gravitational wave (GW) detections from different sources has enabled us to study the population and properties of binary systems of compact objects [1,2], to place constraints on the nuclear equation of state and cosmological parameters [3–7], and to test the theory of general relativity [8–11]. The properties of the emitting system are commonly inferred from GW

signals through Bayesian inference, not only to measure the source's parameters, but also to test hypotheses and compare different models in order to determine which one provides the best description of the data. Bayesian analyses require estimating parameter posteriors or model evidences, a computationally challenging task for the highly-dimensional parameter spaces involved in GW data-analysis. Therefore, stochastic sampling methods, such as nested or Markov chain Monte Carlo (MCMC) samplers [12–14], are usually employed to estimate these quantities. Currently, various software exist for this purpose and are regularly employed by the LIGO-Virgo-KAGRA (LVK) [15–17] collaboration to analyze GW data, e.g., Refs. [18–20].

Adopting different approximants to analyze GW detections from a given type of source can hint at which features of the various models describe the data best, thus providing

*Contact author: anna.puecher@uni-potsdam.de

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

useful information for the development of future models, as shown in Refs. [21,22] for binary neutron star (BNS) and binary black hole (BBH) detections, respectively.

However, in the context of GW analyses from compact binary coalescences (CBC), the relevant question often becomes what was the nature of the objects in the system from which the signal originated. Unless a clear electromagnetic counterpart is present, the classification of a GW signal source is usually based on the estimated component masses [23,24], or, if the signal-to-noise ratio (SNR) is large enough, on the presence of tidal information in the phase. Nevertheless, for events such as GW190814 [25], GW230529 [26], and GW190425 [27], the nature of the components remains uncertain [see, e.g., Refs. [28–41]]. Even for GW170817 [3], an exact source classification based only on the GW signal would not be possible.

In these situations, analyzing GW data with different models can provide information beyond the mere comparison of component masses, thus helping classify the signal's source. Gravitational-waveform models incorporate effects specific to the source they aim at describing: BNS models include tidal phase information, alongside its interplay with spin, while the amplitude corrections in neutron star-black hole (NSBH) models describe the different behaviors near merger and the possibility of tidal disruption.

Bayesian statistics provides a natural way to compare different models or hypotheses describing a certain signal. This is conventionally achieved by performing separate analyses on the same data with the various models investigated, and then comparing evidences to compute the so-called odds ratio. Alternatively, the model can be considered as one of the parameters to sample over. The main advantage of this second approach is that the more informative the data, the more time the sampler will spend in the correct model, reducing the iterations in the disfavored ones and the overall computational cost.

References [21,22] employed a specific case of this second method, introducing a categorical parameter to describe the model to be used at each sampler iteration. However, this approach cannot be applied when comparing models with different parameter-space dimensionality. This issue arises, for example, when comparing aligned-spin and precessing models, or models describing different sources, since in the case of NS components also tidal effects should be accounted for.

In these cases, when we want to compare models with different dimensionalities, we must resort to reversible jump MCMC (RJMCMC) [42,43]. RJMCMC samplers are commonly employed for LISA (Laser Interferometer Space Antenna) data analysis [44–46], for X-ray binaries analyses [47], and for GW ground-based detectors in the context of modelling a burst (or noise) signal with a flexible number of given basis functions, for example in relation to glitch mitigation [48–52].

In this work, we develop, to the best of our knowledge, the first RJMCMC sampler to analyze full GW signals from different CBC sources with state-of-the-art waveform approximants. In particular, our sampler, called *t-roo* (Topsy kangaROO),¹ can analyze GW data employing multiple BBH, NSBH, or BNS models at the same time. In the end, the sampler provides parameter posteriors and odds ratios between the different models.

The main advantages of this approach, compared to performing separate runs and then comparing evidences, are as follows:

- (i) A significant reduction of the “human overhead” during the analysis, since only one run, with no model-specific settings, is needed.
- (ii) A reduction of the computational cost, especially for informative events. In particular, the computational advantage increases the more models are compared at the same time. This is in general expected for RJMCMC samplers, but in the specific case of the sampler presented here, we remark that improvements, especially in the code parallelization, are needed in order to benefit from the expected computational gain.
- (iii) Although the odds ratio convergence requires some time, *t-roo* provides the source classification rather quickly, and could therefore be employed in the future also for low-latency purposes.

Our sampler is based on elements from BILBY [20], where it relies especially on its well-tested utilities for evaluating the likelihood function for CBC signal analyses, and the *eryn* [44] sampling package, from which it inherits the structure of the RJMCMC sampler, including parallel tempering.

We test *t-roo* performance on simulated signals for both current and future-generation detectors, such as the Einstein Telescope (ET), and we analyze the real GW events GW190425 and GW230529.

The paper is structured as follows: in Sec. II, we provide an overview of the elements inherited from *eryn* and BILBY, while in Sec. III we illustrate the specific proposals used in *t-roo*. In Sec. IV, we describe the analysis methods, and we show the results of a set of analyses on simulated data and on the real events in Sec. V. We summarize our conclusions in Sec. VI. In Appendix A, we provide a more detailed and technical description of the working mechanisms of RJMCMC samplers, focusing on the specific application to our case in Appendix B. We illustrate some possible empirical criteria for convergence in Appendix C, and in Appendix D we discuss the

¹Regarding the name, *kangaroo* was chosen to recall the *jump* in “reversible jump”, and for consistency with Australian names for samplers such as BILBY and DINGO. The kangaroo is *tipsy* because it not only jumps, but it can also perform *reverse* jumps.

robustness of the evidence computation by comparing the estimates obtained with different samplers.

`t-roo` is publicly available on GitHub at <https://github.com/AnnaPuecher/t-roo>.

II. T-ROO BASICS

MCMC sampling algorithms differ from nested ones in that they directly sample the target distribution. The main condition that MCMC samplers have to satisfy is the so-called *detailed balance*, which essentially means that the probability of a chain being in a state x and transitioning to a state x' must be the same as being in the state x' and transitioning to x . In the specific implementation of the Metropolis Hastings algorithm [53], at each step, we propose a new point in the parameter space and decide whether to accept it or not, where the exact form of the acceptance probability is what ensures detailed balance. RJMCMC is a generalized version of this algorithm, which allows the chain to move also across parameter spaces with different dimensionalities. A detailed overview of MCMC and RJMCMC samplers is provided in Appendix A.

For CBC analyses, MCMC algorithms are conventionally employed to sample the source posterior $p(\vec{\theta}|d, M)$, where $\vec{\theta}$ denotes the parameters of interest, d the detector data, and M the model employed to describe the data [19,54]. If, instead, we have a set of models $\{M_k\}$ and want to establish which one describes the data best, we can either analyze the data separately with the various models and then compare evidences, usually computed with nested samplers, or, with MCMC samplers, we can directly sample the joint probability

$$p(\vec{\theta}_k, k|d) = p(\vec{\theta}_k|k, d)p(k|d), \quad (1)$$

where k is the model label and $\vec{\theta}_k$ the parameters of model k . Notably, reversible jump MCMC allows us to do this even if the models considered have different dimensions. The better a model manages to describe the data, the more time the sampler will spend in that model. Therefore, once the sampler is converged, the final probability of a model will be given by

$$p_{\text{MOD}_k} := p(k|d) = \frac{n_k}{\sum_k n_k}, \quad (2)$$

where n_k is the number of samples specifically in the model k .

`t-roo` is based on `eryn` [44], a sampler purposefully designed to tackle Bayesian analyses of complex and multidimensional problems, such as the LISA global fit, which is in turn based on the affine invariant sampler `emcee` [54]. In MCMC ensemble samplers, a collection of N_w *walkers* explore the parameter space in parallel, with proposals based on the position of other walkers. The key property of affine invariant samplers is that they have the

same efficiency in sampling different distributions related by affine transformations. This implies that they can rather efficiently sample highly correlated distributions, which are commonly encountered in GW analysis from CBC sources.

In the following, we provide a brief explanation of the features that `t-roo` inherited from `eryn` and `Bilby`.

A. Stretch move

In Ref. [55], the authors suggest the so-called *stretch move* as the simplest proposal to update the walkers' positions that satisfies both affine invariance and detailed balance. In a stretch move, the new position of a walker X_i at time t is proposed by moving by a fraction of the distance with respect to another random walker X_j as

$$X_{i,\text{new}} = X_j(t) - z(X_i(t) - X_j(t)), \quad (3)$$

with z being a random variable drawn from a distribution $g(z)$. This probability distribution is

$$g(z) = \begin{cases} \frac{1}{\sqrt{z}}, & \text{if } z \in [1/a, a] \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

with a arbitrarily set to 2 [55]. The new proposed point is then accepted with probability

$$\alpha(X_{i,\text{new}}, X_i(t)) = \min \left\{ 1, z^{d-1} \frac{p(X_{i,\text{new}})}{p(X_i(t))} \right\}, \quad (5)$$

where $p(X)$ denotes the target probability density, evaluated at the point X , d is the dimension of the parameter space involved in the stretch move, and z^{d-1} a factor needed to preserve the symmetry condition; see Appendix B 3 for details.

When choosing the other random walker X_j needed for the stretch move, one must ensure that the position of this walker is not updated at the same time as the move we are currently attempting. In the group stretch move in `eryn`, the walkers are split into two (or more) separate groups and are updated sequentially. When updating the positions of walkers in a given group, the random walkers X_j are then chosen from the other one(s), therefore called *complementary*.

As explained in detail later, we will employ stretch moves in `t-roo` for different purposes: to update the single models' parameters and for the proposal of tidal deformability parameters when considering jumps between models (cf. Sec. III).

B. Parallel tempering

MCMC samplers can struggle with complex and multimodal distributions since there is a possibility that walkers get stuck in one of the modes. One way to mitigate this issue is through *parallel tempering* [56–58]. In parallel

tempering, for each walker, n_T chains are run in parallel, each one with a different temperature T_j , which transforms the likelihood as

$$\mathcal{L}(d|\vec{\theta}, M) \rightarrow \mathcal{L}(d|\vec{\theta}, M)^{1/T_j}, \quad (6)$$

meaning that the j th chain is effectively exploring the posterior

$$p_{T_j}(\vec{\theta}, d) \propto \mathcal{L}(d|\vec{\theta}, M)^{1/T_j} \pi(\vec{\theta}). \quad (7)$$

Basically, the temperature effect “smoothens” the likelihood profile, allowing the walkers to explore low-likelihood regions of the parameter space more easily. The chain with $T_j = 1$ explores the “true” posterior distribution, i.e., the target distribution, while the $T_j = \infty$ chain explores the prior. Different chains with different temperatures run in parallel and periodically exchange information. Chains with higher temperatures sweep across the parameter space, discovering (new) regions with potentially high likelihood, and then pass this information to the low-temperature chains, which can better explore the likelihood peaks. In order to preserve detailed balance, the swaps between chains with different temperatures are accepted with a probability:

$$\alpha(T_i, T_j) = \min \left\{ 1, \left(\frac{\mathcal{L}(d|\vec{\theta}_i)}{\mathcal{L}(d|\vec{\theta}_j)} \right)^{\frac{1}{T_j} - \frac{1}{T_i}} \right\}. \quad (8)$$

The number of required temperature chains depends on the problem at hand, but usually at least 10-20 are needed. The choice of the specific temperature values is non-trivial, since ideally it should maximize the information exchange among chains. To find the optimal temperature ladder, *eryn*, and consequently *t-roo*, employ the dynamic adaptation algorithm developed in Ref. [59]. It is good practice to consider the samples accumulated during the temperature optimization as part of the burn-in. In *t-roo*, temperature adaptation happens only during the preliminary sampling stage (see Sec. III A 2) and is automatically switched off during the RJMCMC phase. Therefore, when extracting the odds ratio or posterior parameters from the RJMCMC iterations, these samples are automatically not taken into account. As we will see later in Appendix D, parallel tempering also allows one to compute evidence in MCMC samplers.

C. Bilby features

As mentioned in the introduction, *t-roo* employs the sampler infrastructure of *eryn*, combining it with *BILBY* features in order to analyze CBC signals. In particular, the detector classes, the waveform generators, and the likelihood are directly called from *BILBY* with specific wrappers, which allow us to import advanced features

implemented there, e.g., likelihood marginalization [60], or injections in different synthetic noise realizations.

For the runs presented in this work, we employ the *multibanding* approximation of the likelihood [61], which restricts the waveform evaluation to a reduced set of frequency points, adapted according to the signal duration, hence significantly reducing the computational cost. However, we note that in principle all the likelihood methods—e.g., the “standard” *GravitationalWaveTransient* likelihood, the relative binning [62], or the reduced-order-quadrature (ROQ) [63] ones—can be employed in *t-roo*.² Similarly, the prior infrastructure is heavily based on *BILBY*, and therefore, also somewhat specific priors such as the aligned spin one for the spin magnitudes or the uniform in comoving volume for the luminosity distance are available in *t-roo*.

III. T-ROO PROPOSALS

As discussed in detail in Appendix A, an RJMCMC sampler performs two different moves during each sampling step.³ For each walker we have the following:

- (1) *Within-model move*: propose new parameter values for the model the walker is currently in and decide whether to accept the move to this new point in this parameter space or not.
- (2) *Between-models move*: propose a jump to a different model, and simultaneously update the parameters in the new model; also in this case, there is a criterion to decide whether to accept the proposed jump or not.

In the following, we describe the specific moves that we implemented in *t-roo*. We highlight that one strength of our approach is that we are able not only to compare different kinds of compact binaries, but also to simultaneously rank multiple waveform models within a certain binary category. Moreover, we can directly jump between BNS and BBH models, without the need to go through an NSBH one, which reduces the number of iterations and, consequently, the computational cost. Figure 1 shows a schematic overview of the *t-roo* algorithm and the different moves involved.

The proposals are written assuming sampling in chirp mass $\mathcal{M}_c$, mass ratio $q = m_2/m_1$, and component tidal deformabilities Λ_1 and Λ_2 . Although, in principle, it would be possible to sample over different combinations of these parameters and convert them before the proposal, this would yield an (unnecessary) computational cost. The posterior probabilities corresponding to different choices

²In the case of ROQ, this would require the ROQ data to be available in the needed mass range for all the models investigated.

³One could also decide to have multiple within-model moves in a sampling step. This possibility is implemented in *t-roo* but not used by default, since the preliminary sampling already helps identify the parameter regions with the largest support for each model.

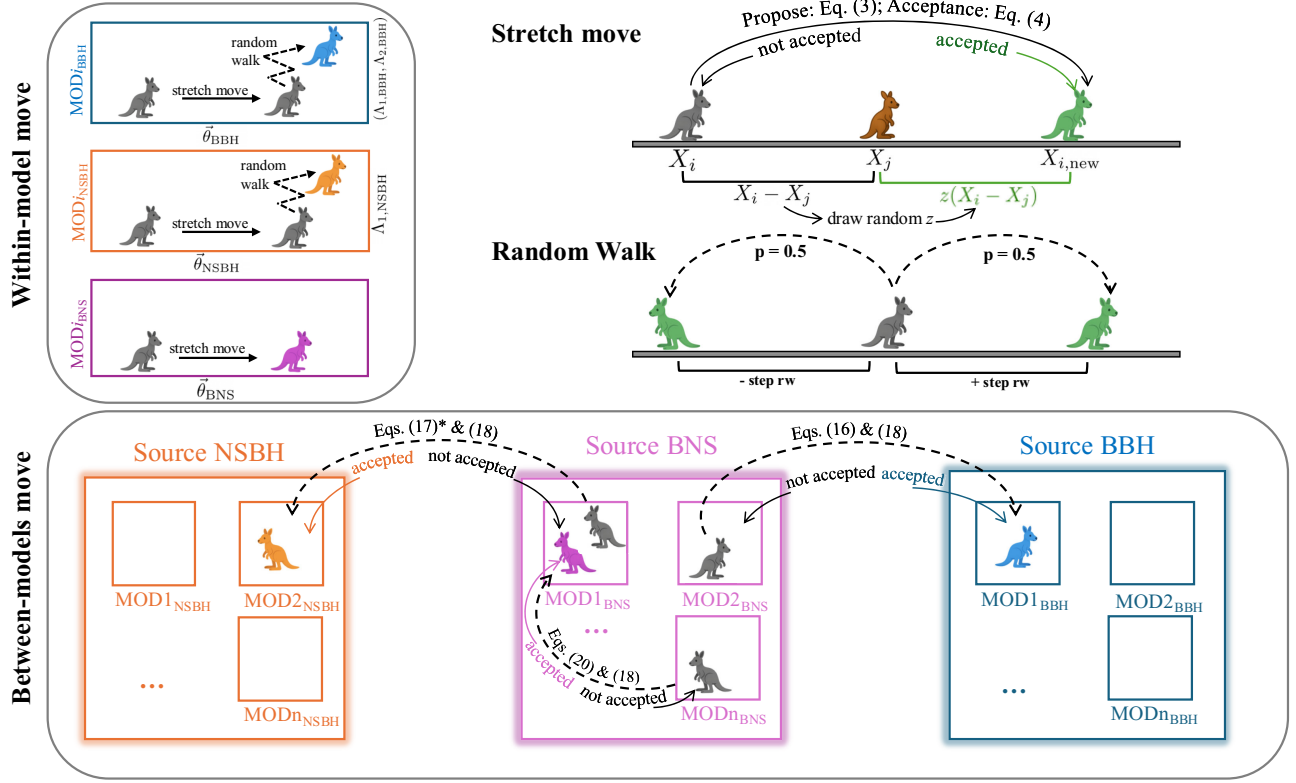


FIG. 1. Schematic representation of the moves implemented in τ -roo, with reference to the respective equations in the text; the asterisk (*) denotes the inverse transformation. For the between-model move, the walkers can jump to any model, of a different or the same source class, with the same probability. For the within-model move, we update the *pseudo*-parameters with a random walk with step “step rw” and all the other parameters with a stretch move. See Sec. III for a complete discussion of the proposals.

of sampling parameters can be obtained anyway by reweighting the final posteriors.

A. Between-models move

The choice for a specific form of the between-model moves is the hardest challenge in RJMCMC algorithms, because their acceptance ratio strongly affects the sampler’s overall performance.

The proposed values for the parameters in the new model depend on the parameters in the current one: to avoid the risk of updating the two states simultaneously, in between-model jumps, we update all walkers in a given model at the time, proceeding sequentially through the models. This approach is similar to the idea of a *group* stretch move, since the walkers are effectively divided into different groups that are updated in sequence. The difference is that the groups are not created with a random partition, but based on the model the walkers are currently in.

When more than two models are compared, τ -roo randomly chooses which model a jump is proposed to, assuming that all the competing models have an equal probability. One could take into account prior information about the preferred models or source type by assigning

model priors, that could be alternatively assigned (or further divided) based on the source classification. This would enable also the incorporation of astrophysical information, e.g., the expected probability to detect systems with specific features, such as eccentricity or precession. This is left for future work, and currently the models’ priors can be incorporated a-posteriori by multiplying the final odds ratios.

1. Tidal deformability

Since our primary goal is to compare models that describe BBH, BNS, and NSBH systems, the core of the between-model proposal concerns the tidal deformability parameters, Λ_1 and Λ_2 . In principle, one could simply add and remove Λ_1 and Λ_2 based on the model jump (for example, when going from NSBH to BBH, we could just remove Λ_2 , while when going from BBH to BNS we could add Λ_1 and Λ_2). However, this would mean losing the tidal information gained while sampling with one model when going to a model with fewer tidal parameters. To avoid this, we add the *pseudo*-tidal parameters $\Lambda_{1,BBH}$, $\Lambda_{2,BBH}$, and $\Lambda_{1,NSBH}$ for BBH and NSBH models, which essentially act as “storage” variables. Crucially, these *pseudo*-parameters are not included in the likelihood computation for the

model they belong to, but serve just to cache the tidal information until the next jump to a model where that tidal parameter is actually included.

The proposal between tidal parameters in different models is a *stretch* move, the primary move employed in `emcee` and explained in Sec. II A. In the implementation proposed by Goodman and Weare in Ref. [55], the random variable z defining the stretch move is sampled in $[1/a, a]$, with a arbitrarily set to $a = 2$. However, in our between-model moves, the *pseudo*-parameters are used to store the tidal information; thus, ideally, we do not want this stretch move to propose too large jumps. Therefore, we choose

$$g(z) = \begin{cases} \frac{1}{\sqrt{z}}, & \text{if } z \in [1 - \epsilon, 1 + \epsilon], \\ 0, & \text{otherwise,} \end{cases} \quad (9)$$

with ϵ set by default to 0.01, but adjustable by the user. Proposing very small jumps, of the order of a few percent, effectively means that we collect most of the information about the tidal parameters during the within-model moves in NSBH or BNS models. X_j is picked from the walkers in the other models, i.e., the ones that are currently not updated. In order to avoid unphysical values of tidal deformability, if the proposed value becomes negative, we simply retain the old one.

2. Proposals for mass parameters

The presence of tidal effects in a signal can alter which values for $\mathcal{M}_c$ and q maximize the likelihood, because of the accelerated inspiral (due to the additional dissipated energy) and the degeneracy between tidal parameters and mass ratio. This circumstance might cause the sampler to get stuck in one of the models: if we propose the jump to a different model assuming the same values for the mass parameters, the new model might not have significant likelihood support at those masses, especially for high-SNR signals. Therefore, we implement specific data-informed proposals for chirp mass and mass ratio, which take into account the effect of different tidal contents.⁴

A `t-foo` run starts with a *preliminary sampling* phase, during which the sampler runs on each model separately for some time without proposing between-model moves.⁵

⁴We only look empirically at the effect of including or not tidal phase contributions in the model, but a more complete proposal would need to take into account other contributions of matter effects, for example, the differences in the spin-spin term in the phase description due to the different quadrupole moment values in NSs and BHs.

⁵Note that parallel tempering is enabled during this stage; therefore, swaps between models due to tempering are still possible.

Figure 2 shows an example of the distribution of $\mathcal{M}_c$ and $\tilde{\Lambda}$ samples after the preliminary sampling, illustrating how the NSBH and BBH models find disjoint chirp mass intervals. To identify a mapping between these different intervals, and similarly for the mass ratio, the sampler considers the $\mathcal{M}_c$, q , and $\tilde{\Lambda}$ samples from each model after the preliminary sampling phase has finished, with $\tilde{\Lambda}$ being the mass-weighted (or “reduced”) tidal deformability parameter [64]. It determines the likelihood “center-of-mass” of the $\mathcal{M}_c$, q , and $\tilde{\Lambda}$ samples collected during the last 200 steps:

$$\mathcal{M}_{c,\text{CM}} = \frac{\sum_i \mathcal{L}(\vec{\theta}_i) \mathcal{M}_{c,i}}{\sum_i \mathcal{L}(\vec{\theta}_i)}, \quad (10)$$

$$q_{\text{CM}} = \frac{\sum_i \mathcal{L}(\vec{\theta}_i) q_i}{\sum_i \mathcal{L}(\vec{\theta}_i)}, \quad (11)$$

$$\bar{\tilde{\Lambda}}_{\text{CM}} = \frac{\sum_i \mathcal{L}(\vec{\theta}_i) \tilde{\Lambda}_i}{\sum_i \mathcal{L}(\vec{\theta}_i)}. \quad (12)$$

Here, the index i runs over the model’s samples, and $\vec{\theta}$ refers to the parameters of that specific model. Furthermore, the bar on $\tilde{\Lambda}$ denotes that, for the computation of $\bar{\tilde{\Lambda}}_{\text{CM}}$ in the BBH and NSBH models, we assume the physical value of zero tidal deformability for the BH object, i.e., $\Lambda_{1,\text{BBH}} = \Lambda_{2,\text{BBH}} = \Lambda_{1,\text{NSBH}} = 0$ (yielding, for BBH models, $\tilde{\Lambda}_{\text{CM}} = 0$), because these are the values employed in the likelihood calculation. We consider only the last 200 iterations to ensure no samples from the burn-in phase are used. However, how many final iterations should be used can be adjusted by the user. Using the equivalent of the center of the mass of the likelihood instead of, for example, the sample corresponding to its maximum value, ensures that we take into account the overall trend of the distribution, making the point choice more robust against potential biases or outliers.

We consider the line connecting the different likelihood-center-of-mass points in the $\mathcal{M}_c$ - q - $\tilde{\Lambda}$ space. For two different models MOD_1 and MOD_2 , its slope is given by

$$s_{\mathcal{M}_c(\text{MOD}_1, \text{MOD}_2)} = \frac{\mathcal{M}_{c,\text{CM},\text{MOD}_1} - \mathcal{M}_{c,\text{CM},\text{MOD}_2}}{\bar{\tilde{\Lambda}}_{\text{CM},\text{MOD}_1} - \bar{\tilde{\Lambda}}_{\text{CM},\text{MOD}_2}} \quad (13)$$

$$s_{q(\text{MOD}_1, \text{MOD}_2)} = \frac{q_{\text{CM},\text{MOD}_1} - q_{\text{CM},\text{MOD}_2}}{\bar{\tilde{\Lambda}}_{\text{CM},\text{MOD}_1} - \bar{\tilde{\Lambda}}_{\text{CM},\text{MOD}_2}}. \quad (14)$$

For each between-models move, the proposed $\mathcal{M}_c$ and q for the new model are computed following the lines with these slopes, starting from the point defined by the current sample in the current model, and assuming the physical values $\Lambda_{1,\text{BBH}} = 0$, $\Lambda_{2,\text{BBH}} = 0$, and $\Lambda_{1,\text{NSBH}} = 0$.

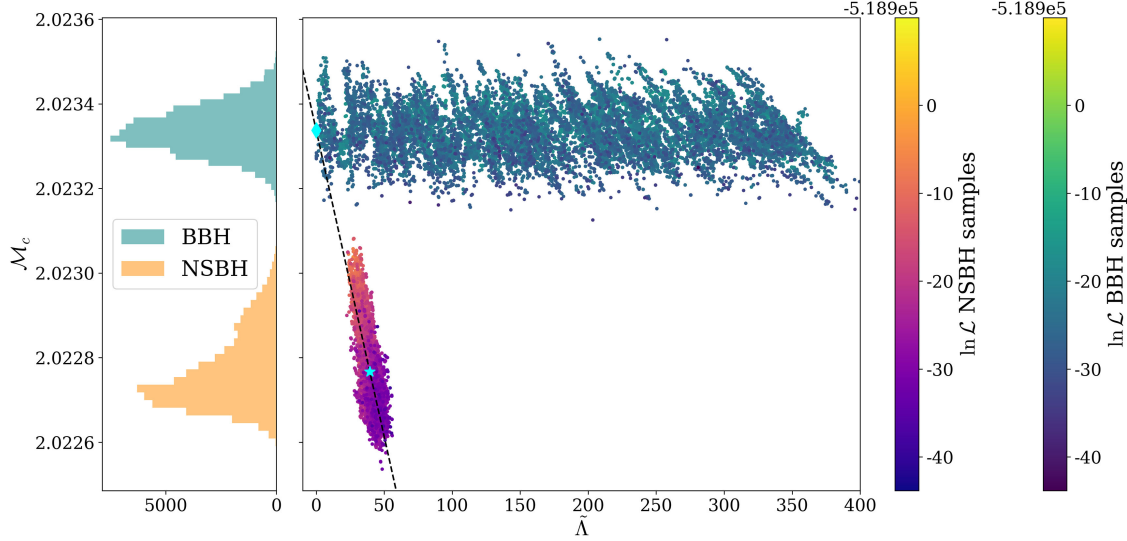


FIG. 2. Example of $\mathcal{M}_c$ and $\tilde{\Lambda}$ distribution of samples collected by all the walkers in the last 200 steps of the preliminary sampling phase in the case of an NSBH injection with $\Lambda_2 = 300$ as recovered by ET (NSBH_LAM300_ET, see Sec. IV A for details), color-coded by the likelihood value of each sample; the orange-red colors correspond to samples recovered with the NSBH approximant IMRPhenomNSBH, and the blue-green with the BBH approximant IMRPhenomD. The left panel displays the one-dimensional $\mathcal{M}_c$ distribution for the BBH and NSBH case, showing how the recovered support lies in disjoint ranges for the two models. We note that in the BBH model $\Lambda_{1,\text{BBH}}$ and $\Lambda_{2,\text{BBH}}$, from which the $\tilde{\Lambda}$ parameter shown here is computed, are *pseudo*-parameters used merely as storage variables, and therefore do not affect the likelihood. The cyan diamond and star indicate the likelihood center-of-mass values for the BBH and NSBH model, respectively. The proposal slope in Eq. (13) is computed from the dashed line connecting them.

However, if the ranges for the $\mathcal{M}_c$ support in different models overlap sufficiently, this specific proposal is not needed. In this case, the chirp mass value is not updated in the between-model move.⁶ Specifically, when going from MOD_1 to MOD_2 , the chirp-mass proposal is *not* employed when the likelihood center-of-mass value of $\mathcal{M}_c$ for MOD_2 lies within the 68% interval of the $\mathcal{M}_c$ samples from MOD_1 . We do not apply this condition to the mass ratio, because the posteriors for this parameter are usually much wider and this distinguishability criterion will rarely be satisfied.

Besides identifying the right map between the different model parameters, the preliminary sampling phase also improves the overall convergence of the sampler. Before the reversible jump even starts, the walkers will already have reduced the region of parameter space to explore for each model. The number of iterations in the preliminary sampling phase should be enough to pass the burn-in phase and provide reliable, although not definitive, posteriors for the parameters of interest, but not too long to affect the computational performance of the algorithm.

⁶We skip the chirp mass proposal mostly to save computational time, although we note that the main computational bottleneck is in any case the waveform evaluation.

3. The complete proposal

Taken all together, the between-model proposals read (see Appendix B 1 for the derivation)

BBH to NSBH

$$\begin{cases} \vec{\Theta}_{\text{NSBH},i+1} = \vec{\Theta}_{\text{BBH},i} \\ \Lambda_{1,\text{NSBH},i+1} = \Lambda_{1,\text{BBH},i} \\ \Lambda_{2,\text{NSBH},i+1} = \Lambda_{2,j,i} + u_2(\Lambda_{2,\text{BBH},i} - \Lambda_{2,j,i}) \\ \mathcal{M}_{c,\text{NSBH},i+1} = \mathcal{M}_{c,\text{BBH},i} + s_{\mathcal{M}_c(\text{NSBH},\text{BBH})} \cdot \tilde{\Lambda}_{\text{NSBH},i+1} \\ q_{\text{NSBH},i+1} = q_{\text{BBH},i} + s_{q(\text{NSBH},\text{BBH})} \cdot \tilde{\Lambda}_{\text{NSBH},i+1} \\ v_2 = 1/u_2 \end{cases} \quad (15)$$

BNS to BBH

$$\begin{cases} \vec{\Theta}_{\text{BBH},i+1} = \vec{\Theta}_{\text{BNS},i} \\ \Lambda_{1,\text{BBH},i+1} = \Lambda_{1,j,i} + w_1(\Lambda_{1,\text{BNS},i} - \Lambda_{1,j,i}) \\ \Lambda_{2,\text{BBH},i+1} = \Lambda_{2,j,i} + w_2(\Lambda_{1,\text{BNS},i} - \Lambda_{2,j,i}) \\ \mathcal{M}_{c,\text{BBH},i+1} = \mathcal{M}_{c,\text{BNS},i} - s_{\mathcal{M}_c(\text{BBH},\text{BNS})} \tilde{\Lambda}_{\text{BNS},i} \\ q_{\text{BBH},i+1} = q_{\text{BNS},i} - s_{q(\text{BBH},\text{BNS})} \tilde{\Lambda}_{\text{BNS},i} \\ u_1 = 1/w_1 \\ u_2 = 1/w_2 \end{cases} \quad (16)$$

NSBH to BNS

$$\left\{ \begin{array}{l} \vec{\Theta}_{\text{BNS},i+1} = \vec{\Theta}_{\text{NSBH},i} \\ \Lambda_{1,\text{BNS},i+1} = \Lambda_{1,j,i} + v_1(\Lambda_{1,\text{NSBH},i} - \Lambda_{1,j,i}) \\ \Lambda_{2,\text{BNS},i+1} = \Lambda_{2,j,i} + v_2(\Lambda_{1,\text{NSBH},i} - \Lambda_{2,j,i}) \\ \mathcal{M}_{c,\text{BNS},i+1} = \mathcal{M}_{c,\text{NSBH},i} + \\ \quad -s_{\mathcal{M}_c(\text{BNS,NSBH})}(\tilde{\Lambda}_{\text{NSBH},i} - \tilde{\Lambda}_{\text{BNS},i+1}) \\ q_{\text{BNS},i+1} = q_{\text{NSBH},i} - s_{q(\text{BNS,NSBH})}(\tilde{\Lambda}_{\text{NSBH},i} - \tilde{\Lambda}_{\text{BNS},i+1}) \\ w_1 = 1/v_1 \\ w_2 = 1/v_2 \end{array} \right. \quad (17)$$

for the $\text{BBH} \rightarrow \text{NSBH}$, $\text{BNS} \rightarrow \text{BBH}$, and $\text{NSBH} \rightarrow \text{BNS}$ moves, respectively. The subscript i denotes current samples, while $i+1$ denotes the proposed ones, and j refers to the complementary set of samples used for the stretch move. The parameter vector $\vec{\Theta}$ stands for the parameters for which no specific proposal is implemented, i.e., $\vec{\Theta} = \vec{\theta} \setminus \{\mathcal{M}_c, q, \Lambda_1, \Lambda_2\}$. In the chirp mass and mass ratio proposal, the barred $\tilde{\Lambda}$ indicates that, in the case of NSBH or BBH systems, we do not use the stored values of the *pseudo*-parameters for the computation of $\tilde{\Lambda}$, but we assume $\Lambda_{1,\text{BBH}} = \Lambda_{2,\text{BBH}} = \Lambda_{1,\text{NSBH}} = 0$; i.e., we assume physical values of zero tidal deformability for BHs, as for the computation of the slopes in Eqs. (13) and (14). The proposal for the auxiliary variables $u_1, u_2, v_1, v_2, w_1, w_2$ is determined by the requirement that the transformations must be reversible (see Appendixes A 5 and B 1 for details). The auxiliary variables w_1 and v_1 are not needed when swapping between NSBH and BBH models, because $\Lambda_{1,\text{BBH}}$ and $\Lambda_{1,\text{NSBH}}$ are both *pseudo*-parameters that can simply be set equal to each other, skipping the stretch-move. The opposite transformations, i.e., $\text{NSBH} \rightarrow \text{BBH}$, $\text{BBH} \rightarrow \text{BNS}$ and $\text{BNS} \rightarrow \text{NSBH}$, are simply the inverse ones, consistent with the requirements that, in RJMCMC algorithms, each proposal should be reversible (the inverse mapping is provided in Appendix B 1).

Once the new model and parameters values have been proposed, they are accepted according to the acceptance probability

$$\alpha(x, x') = \min \left\{ 1, \frac{\mathcal{L}(\vec{\theta}_{x'}) p(\vec{\theta}_{x'}) g'_{x'}}{\mathcal{L}(\vec{\theta}_x) p(\vec{\theta}_x) g_x} \mathbb{J}_{xx'} \right\}, \quad (18)$$

where x denotes the “old” state model and parameters at step i , while x' the new model and parameters of the proposed state at the following step $i+1$. g_x and $g_{x'}$ represent the probability distributions of the auxiliary variables for the model in the states x and x' , respectively (see Appendix A for details). $\mathcal{L}(\vec{\theta}_y)$ and $p(\vec{\theta}_y)$ denote the likelihood and prior for the parameters and model in state y . As explained in

detail in Appendix B 4, in order to preserve the detailed balance, the priors $p(\vec{\theta}_x)$, or $p(\vec{\theta}_{x'})$, must incorporate all the parameters we are sampling over, including the *pseudo*-parameters $\Lambda_{1,\text{BBH}}$, $\Lambda_{1,\text{NSBH}}$, and $\Lambda_{2,\text{BBH}}$, even if they are not used in the likelihood calculation.

For the proposals in Eqs. (15)–(17) and their inverses, the Jacobians $\mathbb{J}_{xx'}$ depend exclusively on the auxiliary variables, and, thanks to the specific mapping between auxiliary variables in the different models,

$$\frac{g_{x'}}{g_x} \mathbb{J}_{xx'} = 1 \quad (19)$$

in all cases; see Appendix B 2 for the computation.

4. Proposal between models for the same type of source and parameter space

`t-foo` provides the possibility to compare not only models that describe different kinds of binary systems, but also multiple models for a given source type. In this case, we do not need to take into account the effects of different tidal contents, and therefore, the parameter transformation between one model and the other is simply an identity transformation:

$$\vec{\theta}_{\text{MOD}_2} = \mathbf{I} \vec{\theta}_{\text{MOD}_1}, \quad (20)$$

with, consequently, a Jacobian $\mathbb{J}_{xx'} = 1$.

We note that this is possible because, for the time being, we are assuming only models with the same physics, for example, aligned-spin only. Comparing models describing not only different sources, but also different features for the same source, such as the presence of precession, or eccentricity, would require dedicated proposals for the between-model moves, and is left for future work.

B. Within-model moves

We update the parameters within a model with a group stretch move as implemented in `eryn` and described in Sec. II A, with the exception of the *pseudo*-parameters $\Lambda_{1,\text{BBH}}$, $\Lambda_{1,\text{NSBH}}$, and $\Lambda_{2,\text{BBH}}$. Given that they are included in the between-model moves, the *pseudo*-parameters are effectively part of the sampled parameter spaces, and as such have to be updated also in the within-model moves. However, since they are not relevant for the likelihood computation, we do not expect to gain information about them. Thus, their posterior should ideally recover the prior, which, for simplicity, we assume to be the same as the actual tidal parameters in the BNS and NSBH models, i.e., uniform in the component tidal deformabilities. To facilitate the recovery of uniform distributions, we update the *pseudo*-parameters in the within-model move through a random walk with a default step size of 10 (in units of dimensionless tidal deformability). If the walkers spend a sufficient amount of time in a model without tidal information, the random walk ensures that the posterior samples

TABLE I. Injections analyzed to assess the sampler’s performance. Together with the labels, we provide the component masses, spins, and tidal deformabilities. The second to last column reports the models used to create the mock signal for the injection, while the last one lists the models employed for recovery.

	$m_1, m_2 [M_\odot]$	χ_1, χ_2	Λ_1, Λ_2	Injected model	Recovery models
NSBH_LAM300	3.75, 1.5	0.6, 0.02	0, 300	IMRPhenomNSBH	IMRPhenomNSBH IMRPhenomD
NSBH_LAM600	3.75, 1.5	0.6, 0.02	0, 600	IMRPhenomNSBH	IMRPhenomNSBH IMRPhenomD
NSBH_LAM1000	3.75, 1.5	0.6, 0.02	0, 1000	IMRPhenomNSBH	IMRPhenomNSBH IMRPhenomD
BBH_NSBHMASSES	3.75, 1.5	0.6, 0.02	0, 0	IMRPhenomD	IMRPhenomNSBH IMRPhenomD
BNS_LAM600	1.4, 1.4	0.01, 0.02	600, 600	IMRPhenomD_NRTidalv2	IMRPhenomD_NRTidalv2 IMRPhenomNSBH IMRPhenomD
BBH_BNSMASSES	1.4, 1.4	0.01, 0.02	0, 0	IMRPhenomD	IMRPhenomD_NRTidalv2 IMRPhenomNSBH IMRPhenomD
BNS_LAM600_MM	1.4, 1.4	0.01, 0.02	600, 600	IMRPhenomXAS_NRTidalv3	IMRPhenomD_NRTidalv2 IMRPhenomXAS_NRTidalv3 SEOBNRv5_ROM_NRTidalv3 IMRPhenomD IMRPhenomXAS SEOBNRv5_ROM

of the *pseudo*-parameters will be uniformly distributed in the prior range.

We include the possibility of *block-sampling* in the within-model moves, meaning that we can update only specific subsets of the parameters at each step. We employ the same structure as in the BILBY-MCMC proposal cycle [19], but in our case only the parameters involved in the proposal change, while the move remains the same, i.e., a stretch move, or a random walk for $\Lambda_{1,\text{BBH}}$, $\Lambda_{2,\text{BBH}}$, or $\Lambda_{1,\text{NSBH}}$. Block sampling is particularly useful in case of multi-modal distributions, since it allows the walkers to better explore the full range of each parameter. However, one of the strengths of affine samplers, such as `emcee`, on which `eryn`, and consequently `t-roc`, are based, is their efficiency in dealing with highly-correlated parameters, which is the case for CBC GW signals. Therefore, in this case, employing block sampling can actually slow down convergence, and it is by default not used in `t-roc`.

IV. ANALYSIS METHODS

A. Injections setup

To assess the robustness of the sampler, we test it on simulated data obtained by injecting different mock GW signals on top of synthetic detector noise. We assume a detector network composed of the two Advanced LIGO detectors, Hanford and Livingston [15], and the Advanced Virgo detector [16] at design sensitivity [65,66]. We consider NSBH signals with different tidal contributions and recover them with an NSBH and a BBH model, to investigate when `t-roc` can identify the presence of a NS

component. We also analyze BNS injections and recover them using different models describing BNS, NSBH, and BBH systems. In addition, we look at signals from BBH sources, with masses compatible with either BNS or NSBH systems, to evaluate how the sampler behaves in the absence of tidal information. All injections are summarized with their labels and source properties in Table I.

For NSBH sources, we use the IMRPhenomNSBH approximant [67], which is based on the BBH amplitude prescription in Ref. [68], in turn based on IMRPhenomC, and the tidal phase description in IMRPhenomD [69] and NRTidalv2 [70]. Therefore, the comparison is performed with the IMRPhenomD approximant for the BBH case and IMRPhenomD_NRTidalv2 for the BNS one, in order to reduce potential modelling systematics. On the other hand, when comparing multiple models for each source, we consider also other state-of-the-art approximants [71–73].

For all injections, we place the source at a luminosity distance of $d_L = 30$ Mpc in order to have a large signal-to-noise ratio, which enables us to identify tidal signatures if present. SNRs range roughly between 50 and 100 depending on the simulated system. For the injection NSBH_LAM600, we perform an additional run at $d_L = 150$ Mpc (NSBH_LAM600_LOWSNR), to test the performance at lower SNR (specifically, this signal has an SNR of 14). We also perform the NSBH_LAM300 and BNS_LAM600 injections for an ET setup (NSBH_LAM300_ET and BNS_LAM600_ET, respectively), assuming two L-shaped detectors with a 15 km arm length and misaligned orientation, placed in the proposed Sardinia and Lusatia sites [74–76].

We employ the multibanding likelihood, as implemented in BILBY, together with phase and distance marginalization in order to reduce the computational cost. Signals are analyzed starting from 20 Hz, and in all the tests we use $N_w = 500$ walkers, each with 16 temperature chains. These are implemented as the default values in `t-roo`, but can be changed depending on the analysis requirements. In particular, when sampling with a large number of models, it is advisable to increase the number of walkers, to ensure an adequate exploration of all parameter spaces.

B. Priors

The parameter priors are kept the same across different models, in order to avoid penalizing or favoring models with different prior support. For chirp mass, we choose a uniform prior in $[1.15, 1.30]M_\odot$ and $[1.95, 2.10]M_\odot$ for the injections of sources with masses compatible with BNS systems or not, respectively. In the analysis of real events, chirp mass is sampled uniformly in the ranges $[1.45, 1.52]M_\odot$ and $[2.02, 2.04]M_\odot$ in the cases of GW190425 and GW230529, respectively. We use BNS models in the recovery only if the injected masses are fully compatible with BNS systems, meaning that we demand the full prior range to lie in the region of validity of the model.⁷ For signals with masses close to the maximum value allowed for neutron star systems, the wider mass prior would cause the waveform evaluation for the BNS models to fail, thus artificially disfavoring the model. Although intuitively correct, since for quite large masses we expect the BNS hypothesis to be disfavored, a consistent treatment of this scenario requires an appropriate rescaling of the mass priors, or of the prior of the models themselves, which are left for future work.

The mass ratio and the tidal deformabilities are sampled uniformly in $q \in [0.16, 1]$ and $\Lambda_1, \Lambda_2 \in [0, 5000]$, respectively. For the component spins, we take an aligned spin prior with the spin magnitudes in the range $[0, 0.90]$, while the luminosity distance is sampled uniformly in comoving volume in the range $[0, 500]$ Mpc.⁸

C. Convergence

The convergence of MCMC analyses is usually based on the accumulated number of effective samples, i.e., the samples thinned by the autocorrelation time, or on metrics such as the Gelman-Rubin coefficient [77].

⁷In the case of real events, one would refer to low-latency point estimates, or to preliminary parameter estimation results.

⁸Spin magnitudes are usually assumed to be smaller for NS components, given that the Kepler limit implies $\chi \lesssim 0.7$. However, in order to analyze the data without assumptions about the source, one has to take into account that each component could also be a BH, thus with spin up to $\chi = 1$. The specific limit $\chi_{\max} = 0.90$ is due to waveform validity constraints. This large range of spin magnitudes is then employed for all the models, even BNS approximants, in order to ensure a consistent prior volume.

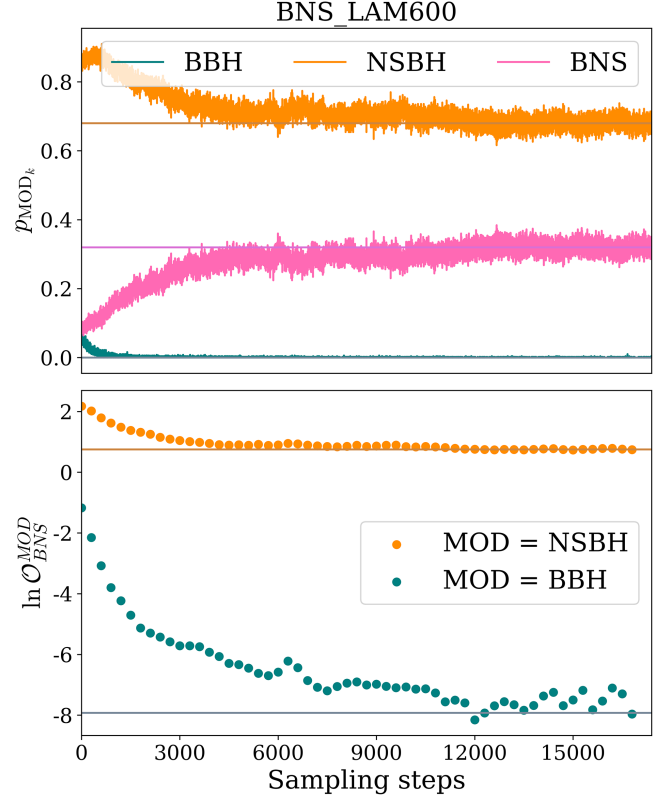


FIG. 3. Top panel: model probabilities, computed as the average number of samples over walkers in each model as a function of the sampling step for the BNS_LAM600 analysis (see Sec. IV A). Bottom panel: for the same analysis, odds ratios with respect to the injected model computed every 300 steps over the last 500 iterations. In both panels, the horizontal lines show the final values obtained in the last 500 iterations and discussed in Sec. V.

In the case of RJMCMC analysis, quantities such as the autocorrelation time and the convergence statistics would need to be defined in a way that takes into account that each chain (walker) jumps between models and dimensions.

Trace plots are usually the best indicator of convergence for the parameter posteriors, and in the case of RJMCMC one can look at whether the time spent in each model has become stable. Figure 3 shows how the probabilities for each model and the odds ratios evolve over the sampler iterations, for the example of the BNS_LAM600 injection (see Sec. IV A for details).

However, it is nontrivial to establish a definitive convergence criterion for RJMCMC samplers, and there are no clear prescriptions in literature. We investigated a few empirical criteria with which the sampler could automatically estimate convergence (as outlined in Appendix C), but we did not deem them sufficiently robust to be employed as the default convergence checks. Therefore, in this initial implementation, `t-roo` by default runs for a fixed number of steps and, if needed, the analysis can be resumed and run for more iterations, thanks to the implemented checkpoints. In the future, we aim at developing more reliable and robust

TABLE II. Posterior model probabilities and odds ratios recovered by `t-roo` compared to the corresponding odds ratios obtained with `dynesty`, including their conservative uncertainty estimates.

	P_{MOD}	odds ratios	
	t-roo	t-roo	dynesty
NSBH_LAM300	$p_{\text{NSBH}} = 0.12$ $p_{\text{BBH}} = 0.88$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = 2.0$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = 2.1 \pm 0.35$
NSBH_LAM600	$p_{\text{NSBH}} = 0.92$ $p_{\text{BBH}} = 0.083$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = -2.4$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = -1.9 \pm 0.35$
NSBH_LAM1000	$p_{\text{NSBH}} = 0.97$ $p_{\text{BBH}} = 0.033$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = -3.4$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = -3.6 \pm 0.35$
BBH_NSBHMASSES	$p_{\text{NSBH}} = 0.24$ $p_{\text{BBH}} = 0.76$	$\ln \mathcal{O}_{\text{BBH}}^{\text{NSBH}} = -1.1$	$\ln \mathcal{O}_{\text{BBH}}^{\text{NSBH}} = -1.8 \pm 0.35$
BNS_LAM600	$p_{\text{BNS}} = 0.32$ $p_{\text{NSBH}} = 0.68$ $p_{\text{BBH}} = 1.2 \times 10^{-4}$	$\ln \mathcal{O}_{\text{BNS}}^{\text{NSBH}} = 0.75$ $\ln \mathcal{O}_{\text{BNS}}^{\text{BBH}} = -7.9$	$\ln \mathcal{O}_{\text{BNS}}^{\text{NSBH}} = 2.0 \pm 0.34$ $\ln \mathcal{O}_{\text{BNS}}^{\text{BBH}} = -6.8 \pm 0.34$
BBH_BNSMASSES	$p_{\text{BNS}} = 0.019$ $p_{\text{NSBH}} = 0.083$ $p_{\text{BBH}} = 0.90$	$\ln \mathcal{O}_{\text{BBH}}^{\text{BNS}} = -3.8$ $\ln \mathcal{O}_{\text{BBH}}^{\text{NSBH}} = -2.4$	$\ln \mathcal{O}_{\text{BBH}}^{\text{BNS}} = -9.3 \pm 0.34$ $\ln \mathcal{O}_{\text{BBH}}^{\text{NSBH}} = -4.0 \pm 0.33$
BNS_LAM600_MM	$p_{\text{PHEND}} = 1.2 \times 10^{-5}$ $p_{\text{XAS}} = 4.8 \times 10^{-5}$ $p_{\text{SEOB}} = 3.2 \times 10^{-5}$ $p_{\text{PHENDV2}} = 0.32$ $p_{\text{XASV3}} = 0.30$ $p_{\text{SEOBV3}} = 0.37$	$\ln \mathcal{O}_{\text{XASV3}}^{\text{PHEND}} = -10$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{XAS}} = -8.8$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{SEOB}} = -9.2$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{PHENDV2}} = 0.065$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{SEOBV3}} = 0.20$	$\ln \mathcal{O}_{\text{XASV3}}^{\text{PHEND}} = -11 \pm 0.32$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{XAS}} = -11 \pm 0.32$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{SEOB}} = -11 \pm 0.32$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{PHENDV2}} = -0.22 \pm 0.33$ $\ln \mathcal{O}_{\text{XASV3}}^{\text{SEOBV3}} = 0.43 \pm 0.32$
NSBH_LAM600_LOWSNR	$p_{\text{NSBH}} = 0.70$ $p_{\text{BBH}} = 0.30$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = -0.86$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = 0.40 \pm 0.27$
NSBH_LAM300_ET	$p_{\text{NSBH}} = 0.99$ $p_{\text{BBH}} = 0.011$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = -4.5$	$\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}} = -7.9 \pm 0.44$
BNS_LAM600_ET	$p_{\text{BNS}} = 0.64$ $p_{\text{NSBH}} = 0.36$ $p_{\text{BBH}} = 0.0$	$\ln \mathcal{O}_{\text{BNS}}^{\text{NSBH}} = -0.57$ $\ln \mathcal{O}_{\text{BNS}}^{\text{BBH}} = -\infty$	$\ln \mathcal{O}_{\text{BNS}}^{\text{NSBH}} = -1.9 \pm 0.38$ $\ln \mathcal{O}_{\text{BNS}}^{\text{BBH}} = -217 \pm 0.37$

criteria that can be automatically implemented in the sampler, and we may also investigate alternative stopping criteria, more similar to standard MCMC analyses.

$$\mathcal{O}_A^B = \frac{n_B}{n_{\text{tot}}} \frac{n_{\text{tot}}}{n_A} = \frac{n_B}{n_A}. \quad (22)$$

V. RESULTS: VALIDATION ON SIMULATED DATA AND ANALYSIS OF REAL EVENTS

To assess the performance of `t-roo`, we compare the results from the injection runs to an equivalent set of analyses performed with `BILBY` and the `dynesty` sampler [18], both for parameter posteriors and model selection.

Usually, the odds ratio between two hypotheses, or models, $\mathcal{H}_B$ and $\mathcal{H}_A$ is computed as

$$\mathcal{O}_A^B = \frac{\mathcal{Z}_{\mathcal{H}_B}}{\mathcal{Z}_{\mathcal{H}_A}} = \frac{p(\mathcal{H}_B|d)}{p(\mathcal{H}_A|d)} = \frac{p(d|\mathcal{H}_B) p(\mathcal{H}_B)}{p(d|\mathcal{H}_A) p(\mathcal{H}_A)}. \quad (21)$$

If we assume no prior knowledge about the model preferences, this reduces to the *Bayes factor* $\mathcal{B}_A^B$, i.e., to the ratio of evidence $\mathcal{Z}_{\mathcal{H}_B}/\mathcal{Z}_{\mathcal{H}_A}$. As mentioned earlier, for RJMCM samplers, the odds ratio is given by the percentage of samples in a given model, i.e.,

Table II lists the model probabilities and odds ratios, computed using Eq. (22), found with `t-roo` and the corresponding odds ratios from the `dynesty` runs, evaluated with Eq. (21), for comparison. Model probabilities in `t-roo` are computed by taking the total number of samples over walkers in each model during the last 500 sampling steps, and odds ratios are always shown with respect to the injected model.

Overall, in most cases, we find good consistency between odds ratios obtained with `t-roo` and `dynesty`. `dynesty` provides an uncertainty estimate for the computed evidence $\sigma(\ln \mathcal{Z})$, quantified from both statistical and sampling error contributions [18]. In Table II, next to the `dynesty` estimates of the odds ratio between two models MOD_1 and MOD_2 , we report the uncertainty⁹

⁹This is a conservative estimate that effectively ignores potential correlations between evidences on the two models.

$$\sigma(\ln \mathcal{O}_{\text{MOD}_1}^{\text{MOD}_2}) \sim \sigma(\ln \mathcal{Z}_{\text{MOD}_1}) + \sigma(\ln \mathcal{Z}_{\text{MOD}_2}). \quad (23)$$

For the odds ratios recovered with RJMCMC samplers, we cannot estimate uncertainty in the same way, since the models' probabilities derive as a consequence of the sampler visiting more often the more favored models.¹⁰

However, once the sampler is converged, we expect a good agreement with the odds ratio recovered by *dynesty*. Nevertheless, we do not expect to recover the exact same value obtained with *dynesty*, given the differences between the samplers and that also the evidence recovered with nested samplers can fluctuate; see Appendix D for a more extended discussion.

As an important validation, in all the tests, the model preferred by *t-roo* and *dynesty* is the same. We find larger differences in the recovered odds ratios in the cases where we compare more than two models. This is likely due to the fact that the evidence, as computed in *dynesty* and shown in Eq. (A2), involves a marginalization over the entire parameter space. In *t-roo*, the less favored models are visited less often, and consequently also the parameter space is less explored. To verify that, instead, this is not caused by some issue within the sampler, we reanalyze the injection BNS_LAM600 with *t-roo* using for recovery only the NSBH and BNS models. If the sampler works properly, we expect to recover the same model probabilities, and consequently odds ratio $\ln \mathcal{O}_{\text{BNS}}^{\text{NSBH}}$, as in the analysis including also the BBH model, because, once converged, a RJMCMC sampler should return the “correct” model probabilities, independently of the number of models investigated. With this new analysis, we obtain $p_{\text{NSBH}} = 0.68$ and $p_{\text{BNS}} = 0.32$, yielding $\ln \mathcal{O}_{\text{BNS}}^{\text{NSBH}} = 0.78$, which correspond to the values reported in Table II for the analysis including also the BBH model (small differences in the odds ratio value are due to fluctuations in the next significant digits in the model probabilities). Therefore, we confirm that the differences with the *dynesty* results are not due to issues in the sampler, but to the fact that for the least favored models the collected samples do not cover the full parameter space. In order to recover the same evidences, and consequently odds ratio, one could run the analysis for longer and potentially tweak the priors to ensure a full exploration of all the parameter spaces. However, this would undermine the computational advantage of *t-roo* and would be in contrast to the wish of sampling more preferred models more often, while avoiding unnecessary evaluations of the less favored ones.

¹⁰To provide some uncertainty, one could compute the probability of each model as the mean number of samples in that model over the walkers. This value would be characterized by the statistical uncertainty $\sigma = 1/\sqrt{N_w}$, with N_w being the number of walkers considered. However, given the large number of walkers employed (specifically, $N_w = 500$ in all the analyses presented here), this uncertainty would still be negligible compared to the *dynesty* one.

A. Injection-recoveries with NSBH and BBH models

Figure 4 shows the recovered posteriors for the intrinsic parameters $\mathcal{M}_c$, q , χ_1 , χ_2 , and Λ_2 for the NSBH_LAM600 analysis, comparing the posteriors obtained both for *t-roo* and *dynesty* with the NSBH model IMRPhenomNSBH and the BBH one IMRPhenomD. The injected parameters are recovered correctly for the NSBH approximant and slightly biased for the recovery with the BBH one, as expected given that the former was used to inject the signal. In particular, the separated peaks of the $\mathcal{M}_c$ and q posteriors illustrate why we need specific proposals in the between-model move, as explained in Sec. III A 2. The posteriors recovered by *t-roo* are consistent with the ones obtained from *Bilby* using *dynesty* for both models. The *t-roo* posterior for the BBH model appears somewhat rugged, as a consequence of the BBH model being disfavored and therefore rarely visited by the walkers. In fact, this shows the advantage of RJMCMC samplers and how they can reduce the computational cost of these analyses when one of the models is strongly preferred, as the sampler will then focus automatically on the model with the higher preference.

The marginalized distribution for Λ_2 shows how *t-roo* accurately recovers the right value in the NSBH model, whereas in the BBH model Λ_2 is distributed across the entire prior range thanks to the random walk proposal. In the absence of tidal information, the BBH model would be preferred and the sampler would spend more time there; consequently, the large number of iterations of the random walk would consistently yield a final posterior that just recovers the prior on Λ_2 .

For the NSBH injections NSBH_LAM600 and NSBH_LAM1000, *t-roo* correctly recovers IMRPhenomNSBH as the favored model, while for NSBH_LAM300, the sampler favors the BBH approximant IMRPhenomD. This is expected, because the tidal information recovered by the data in the latter case is not sufficient to discriminate between the two scenarios, and the sampler prefers the simpler model. Notably, this is consistent with the evidence values recovered by *dynesty*, which also indicate a preference for the BBH model for this particular injection.

On the other hand, for NSBH_LAM600_LOWSNR, i.e., the injection with $\Lambda_2 = 600$ but placed at a larger distance, we still find that the NSBH model is favored, but only with $p_{\text{NSBH}} = 70\%$ compared to $p_{\text{NSBH}} = 92\%$ for the injection at $d_L = 30$ Mpc. Also these results are consistent with the ones from *dynesty*.

The situation changes when we analyze the signal from the same source but detected by ET: thanks to the larger SNR (~ 519), also a smaller tidal deformability is enough to clearly recover the tidal content, and therefore discriminate between the two scenarios. For the NSBH_LAM300_ET test, both *t-roo* and *dynesty* correctly identify IMRPhenomNSBH as the preferred model.

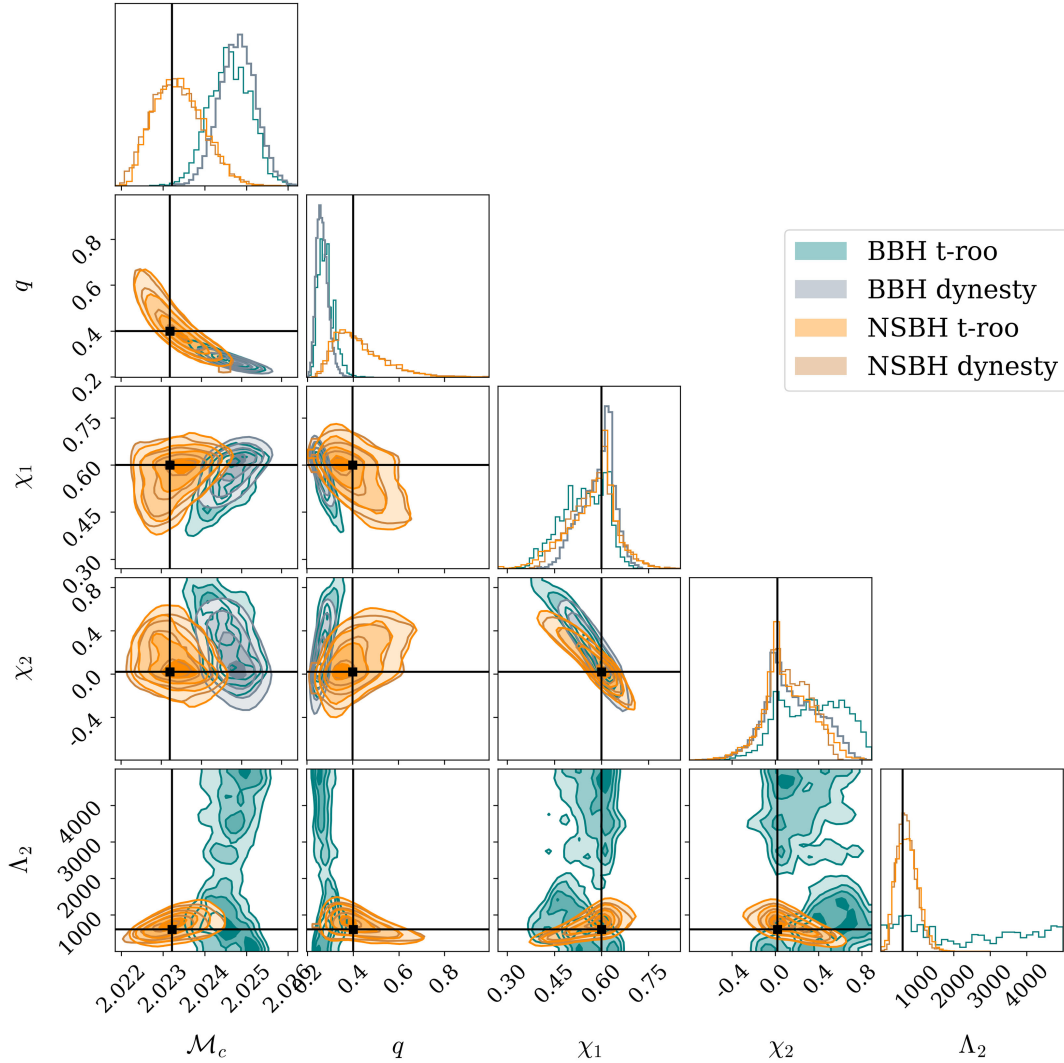


FIG. 4. Parameter posterior distributions recovered by `t-roo` for both the NSBH and BBH models for the NSBH_LAM600 injection, compared to the posteriors obtained for the same analysis with `dynesty` and `Bilby`. For `t-roo`, we show only the samples of the last 1000 sampling steps to exclude samples where the sampler is still exploring the preferred model. The black lines indicate the injected values.

Moreover, we note how, for such high-SNR signals, the data are very informative and therefore `t-roo` manages to converge very quickly, as shown in Fig. 5. This highlights the potential of RJMCMC algorithms in these scenarios: large SNRs usually make parameter estimation analyses more computationally expensive, and in these cases running separate analyses with all the models could become prohibitive. On the other hand, with `t-roo`, large SNRs make it easier for the sampler to understand what the favored model is, yielding a significant computational advantage.

If, instead, we simulate a BBH signal with the same parameters (i.e., the analysis BBH_NSBHMASSES in Table I), we see from Table II that `t-roo` correctly recovers the BBH model as the preferred one, with an odds ratio again consistent with the `dynesty` one. Also in this case, the inferred posteriors are in excellent agreement with the ones

from `Bilby` and the Λ_2 parameter for the NSBH model rails against zero, indicating that no tidal information is recovered in the data.

B. Injection-recoveries with BBH, NSBH, and BNS models

In Fig. 6, we show the posteriors recovered for the BNS injection BNS_LAM600. We used BBH, NSBH, and BNS models for the recovery, as listed in Table I. We find again very good agreement between the distributions inferred with `t-roo` and `dynesty`. The injected values are correctly recovered by the BNS and NSBH models, while we find biases in the BBH ones, especially in chirp mass and mass ratio, where the latter peaks at $q = 0.44$ although the injected signal corresponds to an equal-mass system. This is expected given the lack of tidal content in the BBH

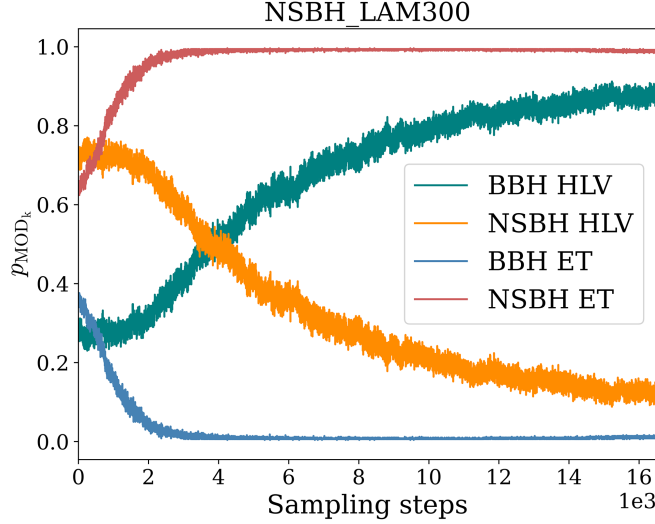


FIG. 5. Model probabilities for the NSBH_LAM300 injection as recovered by a LIGO-VIRGO and an ET network (labeled “HLV” and “ET”, respectively), computed as the average number of samples per walker accumulated in each model over the RJMCMC sampling steps. The model probabilities do not start exactly at 0.5 because of the temperature swaps in the preliminary sampling.

model and the correlation between the tidal parameters and mass ratio.

The t - roo posterior for the BBH model is rather scattered, because it encompasses less than 0.1% of the total number of samples, due to its low preference. We highlight that this is the expected outcome of RJMCMC samplers: the walkers will systematically explore mostly the favored model(s), therefore yielding a good estimate of their posteriors. Posteriors for the disfavored models could be obtained by running the sampler for longer, until the accumulated number of samples in the disfavored model is enough to reconstruct the parameter posterior probability. However, this undermines the purpose of RJMCMC, i.e., understanding which models are more likely and focusing on them. The underlying idea is that a RJMCMC sampler will provide us with an estimate of the preferred model and the “preferred” posterior probabilities, i.e., the posteriors conditional to the model deemed “correct”.

From the models’ probabilities in Table II, the favored model is IMRPhenomNSBH, even though the injected signal was simulated with the BNS approximant IMRPhenomD_NRTidalv2. This is again in agreement with the odds ratios estimated by dynesty. Looking at the marginalized posterior for Λ_2 in Fig. 6, we see that, for the NSBH model, both t - roo and dynesty recover a value larger than the injected one, $\Lambda_{2, \text{inj}} = 600$, peaking roughly at $\Lambda_2 \sim 1800$. At leading order, the tidal information enters waveform approximants as the mass-weighted combination $\tilde{\Lambda}$ [78,79], which, in the case of the BNS_LAM600 injection, is $\tilde{\Lambda} = 600$. When using the NSBH model for this specific

system, we would need a NS tidal deformability of $\Lambda_2 = 1200$ to yield the injected value of $\tilde{\Lambda}$, since the tidal deformability of the BH is set to zero. Therefore, the NSBH model is trying to recover the tidal information, but assuming it all belongs to just one object. This also explains why the NSBH model is preferred: since the tidal content at this point is basically equivalent in the BNS and NSBH models, in the absence of other discriminant information, the sampler prefers the simpler model. This means that in principle we cannot distinguish signals from a BNS or NSBH source just based on the leading-order tidal information in different approximants, but this is not caused by issues in the specific sampling methods used. Employing more advanced models, such as NRTidalv3 [73], which can distinguish the individual component’s tidal deformabilities, will help to solve this degeneracy.

Because of the large number of samples accumulated for the NSBH model, its posterior for Λ_1 shows how the sampler recovers the uniform prior for this *pseudo*-parameter that, as explained in Sec. III, basically serves as a storage variable.

As for Fig. 4, we see a noticeable difference in the posterior peaks of $\mathcal{M}_c$ and q between the BBH model and the NSBH or BNS one. Considering that, as discussed above, the NSBH and BNS models recover the same tidal information, this difference arises likely from the lack of tidal content in the BBH approximant, and supports the need of a tidal-dependent proposal for these two parameters.

Finally, for the BBH_BNSMASSES injection, where we analyze a signal coming from a source with the same parameters, but being a BBH system instead of a BNS one, t - roo correctly recovers IMRPhenomD as the preferred model. In this case, the mass ratio is recovered correctly by all models, since the injection is performed with a BBH approximant (i.e., the data do not contain tidal information), and both the NSBH and BNS models recover correctly Λ_1 and/or Λ_2 being zero.

C. Injection comparing multiple models

Finally, we discuss the case in which we compare not only one model per binary type in the recovery, but multiple models describing the same source. For BNS_LAM600_MM, we use IMRPhenomXAS_NRTidalv3 to inject a signal from an equal-mass BNS system with $\Lambda_1 = \Lambda_2 = 600$, and analyze it with three BNS and three BBH approximants; cf. Table I.

Figure 7 shows the posteriors recovered with the different models. Although not shown here for simplicity, we find in all cases a very good agreement with the posteriors recovered by BILBY runs with dynesty. We only see slight differences in the q posteriors recovered by BNS models. This is likely due to the interplay with the BBH model, which systematically prefers more unequal-mass systems and slows down the convergence for this

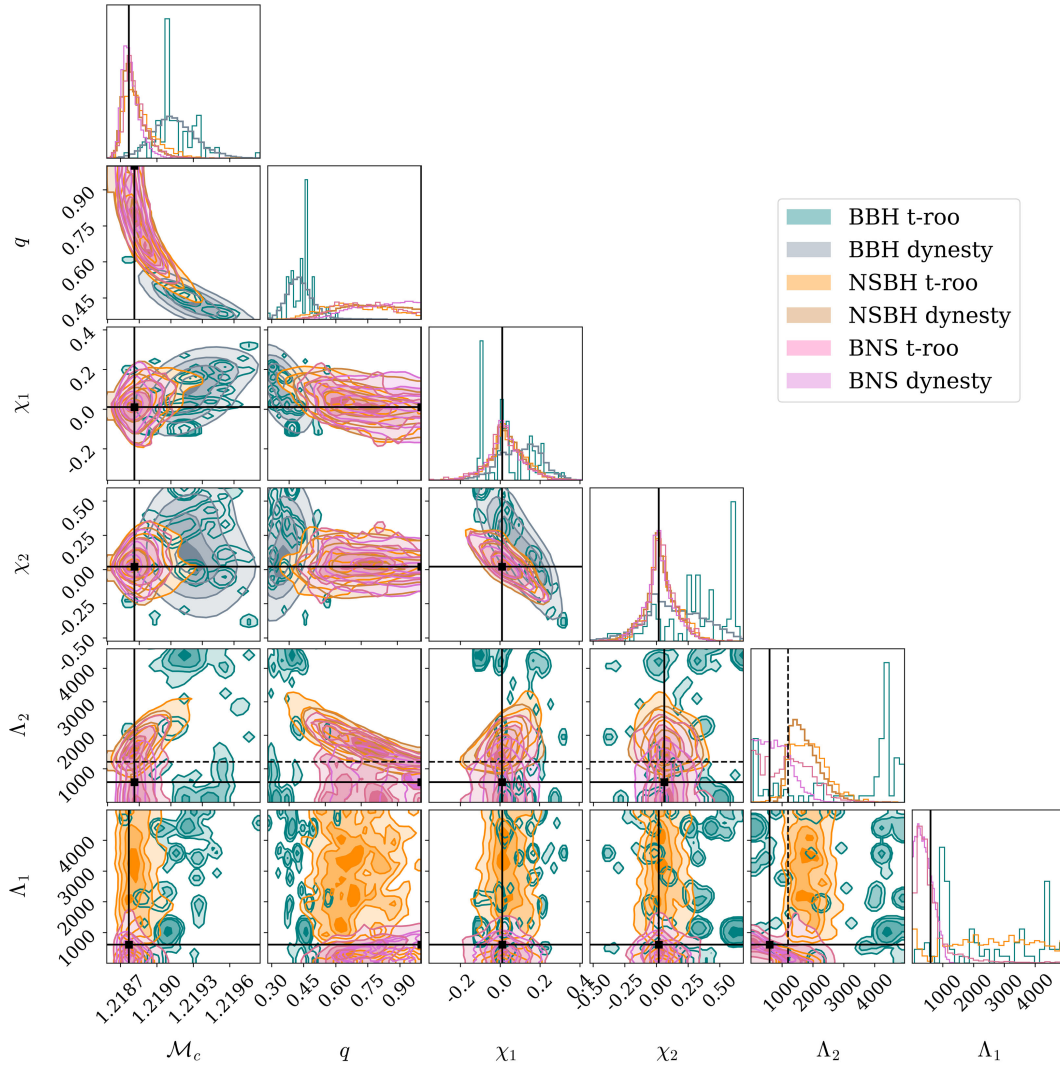


FIG. 6. Parameter posterior distributions recovered by t -roo with BNS, NSBH, and BBH models for the BNS_LAM600 injection, compared to the posteriors obtained for the same analysis with dynesty. For t -roo, we show only the samples of the last 1000 sampling steps, to avoid including samples of a phase when the sampler is still exploring the preferred model. The black solid lines indicate the injected values, while the dashed line in the panel for the Λ_2 one-dimensional distribution denotes the Λ_2 value needed by the NSBH model to yield the same $\tilde{\Lambda}$ value as the injected one.

parameter: a more sophisticated proposal for the mass ratio, potentially also including spin information, would help mitigate this effect.

We see that all BBH approximants recover very similar posteriors, and likewise all posteriors from BNS models are essentially identical. However, there are noticeable differences between BBH and BNS posteriors, especially for $\mathcal{M}_c$ and q , in analogy to the previous results. The BNS models recover Λ_1 and Λ_2 correctly, while the BBH ones tend to return the prior. t -roo correctly favors the BNS models, and in particular SEOBNRv5_ROM_NRTidalv3. Moreover, although the odds ratios are slightly different, the order of the model preference, even within the same kind of source, is the same as dynesty. The only difference is that t -roo slightly prefers IMRPhenomD_

NRTidalv2 over IMRPhenomXAS_NRTidalv3. However, the difference with respect to the dynesty odds ratio lies within the uncertainty of the latter.

We remark that the t -roo results are obtained with only one run, in which we can compare an arbitrary number of models for different kinds of sources. Even though a larger set of models might require a larger number of iterations for the sampler to reach convergence, this is still computationally and practically more convenient than running separate analyses for each model, especially if some of the models, like the BBH ones here, are quite strongly disfavored.

We emphasize again that t -roo and RJMCMC samplers focus on model selection, returning a good posterior estimation only if a model is sufficiently favored. Basically,

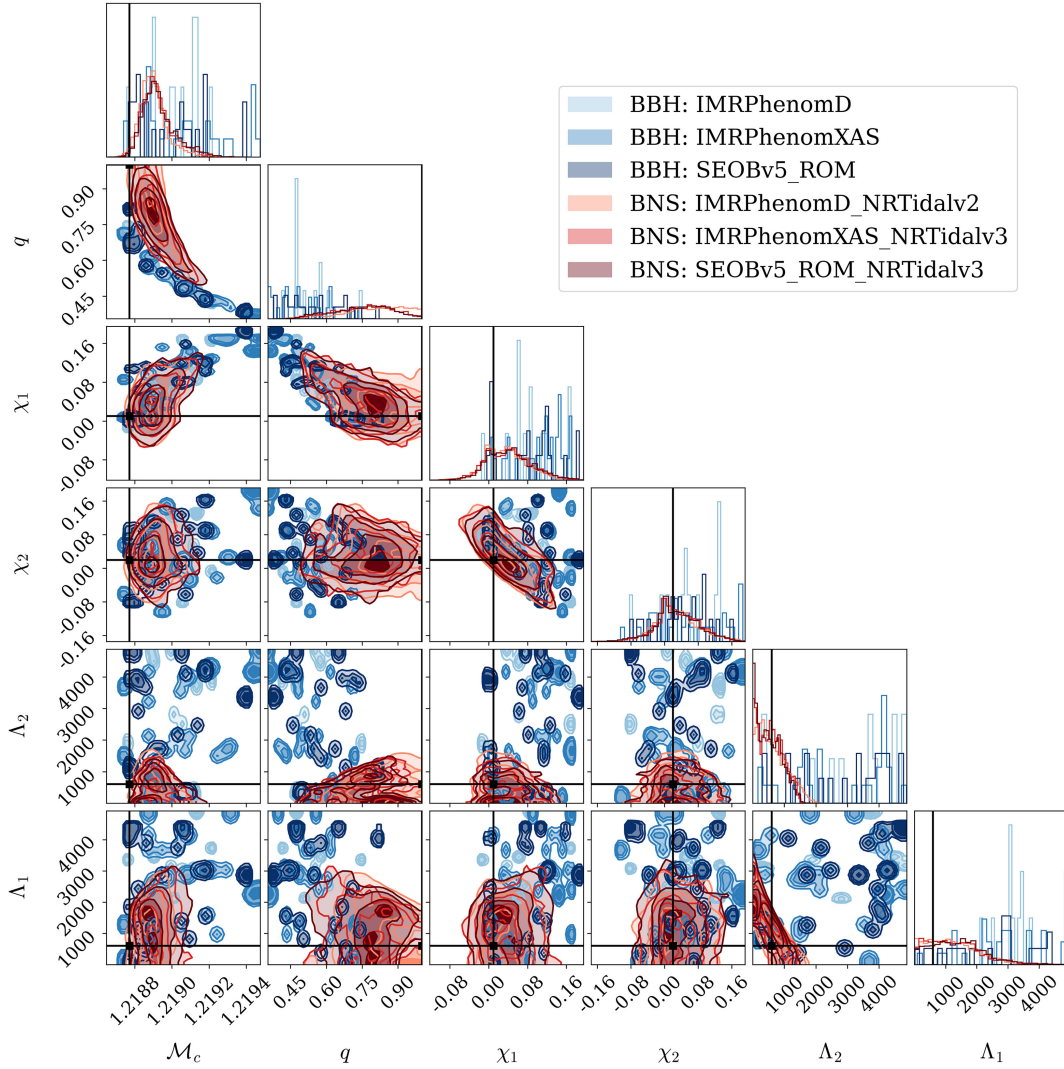


FIG. 7. Parameter posterior distributions recovered by `t-roo` with three different BNS and three different BBH models for the BNS_LAM600_MM injection. To better show the comparison of the posteriors for the BBH models, in this case, we include in the corner plot the last 1500 iterations. The black solid lines correspond to the injected values.

the sampler does not “waste” time exploring the posteriors for models that it has inferred do not describe the data sufficiently well, which yields a significant computational gain.

D. Real events

We employ `t-roo` to analyze the real events GW190425 and GW230529,¹¹ for which the absence of electromagnetic counterparts and of clear tidal signatures did not enable a certain classification of the components as NSs. We do not analyze GW190814 because, given the

large mass ratio, higher-order-modes would need to be included in all models used to analyze this signal.¹²

We analyze both events with `IMRPhenomXAS` for the BBH case, `IMRPhenomNSBH` for the NSBH case, and `IMRPhenomXAS_NRTidalv3` for the BNS one. The recovered probabilities for each source are shown in Table III. For GW190425, `t-roo` prefers the BBH model, with a probability $p_{\text{BBH}} = 0.66$, consistent with the fact that no significant tidal information was recovered for this event. Among the NSBH and BNS models, the NSBH one is preferred, again consistently with the fact that it has fewer free parameters.

¹¹Data were downloaded from the Gravitational Wave Open Science Center [80,81], with the estimated power spectral densities provided in Refs. [82,83], respectively.

¹²Recently, an NSBH model including higher-order modes has been developed in Ref. [84]. However, for this work, we employ only publicly available models reviewed by the LVK collaboration.

TABLE III. `t-roo` recovered probabilities for models corresponding to different source types for the real events GW190425 and GW230529.

Event	p_{BBH}	p_{NSBH}	p_{BNS}
GW190425	0.66	0.31	0.033
GW230529	0.38	0.55	0.067

For GW230529, `t-roo` finds that the NSBH model IMRPhenomNSBH is the one that best describes the data, with a probability $p_{\text{NSBH}} = 0.55$.

E. Computational cost

One of the main advantages of RJMCMC algorithms for Bayesian model selection is, especially when the data are informative, the expected reduction in computational cost with respect to performing separate analysis for each model and then comparing the evidence.

A consistent comparison between the samplers' efficiency is beyond the scope of this paper because it would require specific prescriptions to describe computational costs that take into account not only the number of accumulated samples and the likelihood evaluation time, but also the parallelization inefficiency, the burn-in time, and the effect of parallel tempering, as shown, for example, in Ref. [19] for BILBY-MCMC.¹³

To provide an indication of the computational gain, the `t-roo` analysis for NSBH_LAM600 performed the preliminary sampling stage and accumulated 10^4 RJMCMC samples in 14 hours, while for `dynesty` the BBH run took a bit more than 24 hours, accumulating 10159 samples, and the NSBH one took ~ 31 hours, accumulating 10889 samples. All these times refer to runs performed with 8 cores on the same machine. Even though `eryn`, and consequently `t-roo`, is parallelized, the current implementation is suboptimal, as only the likelihood evaluations are performed on different physical CPUs, while every other operation is executed on one main device, leading to a significant overhead cost. Consequently, `eryn` does not scale properly when running on more than 8 cores. `dynesty`, on the other hand, can run on many CPUs in parallel, making it faster in practice.

We stress that, in order to fully exploit the potential of `t-roo`, and especially to make it competitive with samplers like `dynesty` that can run in parallel on many more cores, the code parallelization has to be optimized.

VI. CONCLUSIONS AND FUTURE APPLICATIONS

We have developed `t-roo`, the first RJMCMC sampler to analyze GW signals from CBCs with different state-of-the-art

waveform models, even with a different number of parameters, at the same time. Our implementation is focused toward applying this method to source classification, i.e., comparing models describing different types of binary systems such as BNSs, BBHs, and NSBHs. We implement specific between- and within-model proposals that handle the different tidal content of the models. Moreover, we encompass the possibility of comparing multiple different models for each of various source types at the same time.

We assessed the performance of `t-roo` in different scenarios, analyzing simulated GW detections from different sources (see Sec. V for details). In all the cases, the inferred posterior distributions and odds ratios are in agreement with the ones obtained from BILBY analyses with the `dynesty` sampler.

For NSBH signals, `t-roo` correctly identifies the binary type when the tidal information in the data is sufficient to discriminate the two options, otherwise it prefers the simpler BBH model. This means that signals from systems with a lower tidal deformability require larger SNR to be correctly identified, as expected. However, a NSBH system with $\Lambda_2 = 600$ is correctly identified already at an SNR of 14. On the other hand, when injecting BNS signals, the NSBH model is favored, due to the fact that it can mimic the tidal content of the injection, while requiring one fewer free parameter. However, in all cases, the models preferred by `t-roo` are consistent with the `dynesty` results. We also investigated BBH signals from sources where masses are within the BNS or NSBH range, finding that the BBH model is correctly preferred in both cases. Finally, we analyzed a BNS injection with three BNS and three BBH models, finding again that the inferred odds ratios and posteriors agree with the ones recovered by `dynesty`.

We emphasize that the main advantage of an RJMCMC approach is that it only requires one run to provide model probabilities and posteriors for the models sufficiently favored, while otherwise one run for each model is needed. It can significantly reduce the computational cost of the analysis, because, the less favored a model is, the less time the sampler will spend in that model. However, as explained in Sec. VE, `t-roo` specifically needs to be better parallelized in order to fully exploit its computational-gain potential and to provide performances competitive with other commonly used samplers.

This reduction of the computational cost becomes more pronounced when the data is particularly informative, and when we are comparing many models. The advantage of performing only one run is not only computational, but also practical, since it reduces the human overhead, the setup time, and the potential risk of mistakes or mismatches in settings between runs.

Further developments require, first of all, an improvement of the code parallelization, together with investigating robust convergence criteria and potential methods to estimate uncertainties. Moreover, future work will not just

¹³As `dynesty` is a nested sampler, a different model for the computational cost estimation would be needed.

focus on the improvements mentioned throughout the work, but also on the application to different case studies. Given the flexibility of the framework, different use-cases simply require the design of specific proposals, in order to make the RJMCMC jumps efficient. For example, it will be interesting to compare models with precession and eccentricity, or models with or without deviations from GR described by extra parameters.

The potential computational gain of this approach is going to be particularly important with the observation of long and loud signals. While the analysis of high-SNR signals is usually more computationally expensive, for RJMCMC samplers it becomes easier to distinguish the information, and consequently it requires fewer iterations to converge. Therefore, `t-roo` is a very powerful starting tool for a large number of applications in GW data analysis not only for current detectors, but also for future ones such as the third-generation ground-based detectors Einstein Telescope and Cosmic Explorer, and the space-based LISA.

ACKNOWLEDGMENTS

We thank Alexandre Toubiana and Alessandro Santini for the help with `eryn`, and Nihar Gupte for the useful comments. A. P., T. D., and H. K. acknowledge funding from the EU Horizon under ERC Starting Grant, No. SMArt-101076369. J. G. acknowledges support from the Simons Foundation International via Grant No. SFI-MPS- BH-00012593-06. Computations were performed on the DFG-funded research cluster Jarvis at the University of Potsdam (INST336/173-1; Project No. 502227537). This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation. Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

DATA AVAILABILITY

`t-roo` is publicly available on GitHub at <https://github.com/AnnaPuecher/t-roo>.

APPENDIX A: REVERSIBLE JUMP MCMC THEORY

1. Bayesian model selection

According to Bayes' theorem, we can infer the parameters $\vec{\theta}$ of an observed system from the observed data d as

$$p(\vec{\theta}|d, \mathcal{H}) = \frac{p(d|\vec{\theta}, \mathcal{H})p(\vec{\theta}|\mathcal{H})}{p(d|\mathcal{H})}, \quad (\text{A1})$$

where $p(\vec{\theta}|d, \mathcal{H})$ is the posterior probability of the parameters $\vec{\theta}$ given the data d and a model or hypothesis $\mathcal{H}$ to

describe the data, $p(d|\vec{\theta}, \mathcal{H})$ is the *likelihood*, and $p(\vec{\theta}|\mathcal{H})$ is the prior probability of the parameters under the assumption of the model or hypothesis $\mathcal{H}$, i.e., the prior describes our knowledge about the parameters before the observation.

The term at the denominator is called *evidence* and corresponds to the marginalized likelihood

$$p(d|\mathcal{H}) = \int d\vec{\theta} p(d|\mathcal{H}, \vec{\theta}) p(\vec{\theta}|\mathcal{H}). \quad (\text{A2})$$

When interested in the parameters posterior probabilities $p(\vec{\theta}|d, \mathcal{H})$, the evidence serves just as a normalization factor. However, it plays a crucial role when the goal is model selection instead.

Let us suppose that we have two models $\mathcal{H}_A$ and $\mathcal{H}_B$, and we want to know which one provides a better description of some observed data d . Using Bayes' theorem, the probability of each model given the data is

$$p(\mathcal{H}_A|d) = \frac{p(\mathcal{H}_A)p(d|\mathcal{H}_A)}{p(d)}, \quad (\text{A3})$$

and similarly for $\mathcal{H}_B$. To compare the models' probabilities, one usually refers to the odds ratio

$$\mathcal{O}_A^B = \frac{p(\mathcal{H}_B|d)}{p(\mathcal{H}_A|d)} = \frac{p(\mathcal{H}_B)}{p(\mathcal{H}_A)} \underbrace{\frac{p(d|\mathcal{H}_B)}{p(d|\mathcal{H}_A)}}_{\mathcal{B}_A^B} \underbrace{\frac{p(d)}{p(d)}}_{1}, \quad (\text{A4})$$

where the ratio of the models' prior probabilities π_A^B is usually set to 1, meaning that we do not have *a priori* preference between them. Therefore, usually the odds ratio reduces to the *Bayes factor* $\mathcal{B}_A^B$, defined as the ratio between the evidences obtained with the different models, computed as in Eq. (A2).

An alternative approach consists in sampling directly over the model, i.e., the target of the Bayesian inference becomes the joint posterior probability

$$p(\vec{\theta}_k, k|d) = \frac{p(d|k, \vec{\theta}_k)p(k, \vec{\theta}_k)}{\sum_{k' \in K} \int p(k', \vec{\theta}_{k'}) p(d|k', \vec{\theta}_{k'}) d\vec{\theta}_{k'}}, \quad (\text{A5})$$

with k being the model's label for a finite set K of models, $\vec{\theta}_k$ the parameters of the specific model k , and $p(k, \vec{\theta}_k) = p(k)p(\vec{\theta}_k|k)$ the prior for each model and its parameters. This joint posterior probabilities can always be factorized in

$$p(k, \vec{\theta}_k|d) = p(k|d)p(\vec{\theta}_k|k, d), \quad (\text{A6})$$

hence providing the model's probability given the data and the parameters probability for a given model, given the data and a specific choice of model.

Reversible Jump MCMC algorithms sample directly the joint posterior $p(\theta_k, k|d)$. The key idea is that, the more (less) favored a model is, the more (less) time the sampler will spend in it, and consequently the more (less) samples it will accumulate for that model.

Therefore, at the end of a RJMCMC algorithm, the odds ratio between two models $\mathcal{H}_A$ and $\mathcal{H}_B$ will be given by

$$\mathcal{O}_A^B = \frac{n_B}{n_A}, \quad (\text{A7})$$

with n_A and n_B being the number of samples in model $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively.

2. MCMC review

Markov chains are stochastic processes that define a sequence of states of a system, in which each state depends on the previous one, and only on that. A Markov chain is characterized by a transition kernel $\mathcal{T}$, which describes the transition probability from one state to the other—i.e., $\mathcal{T}(x, y)$ denotes the conditional probability of the next state in the chain to be y given that the present state is x —and a limiting distribution π , to which the states probabilities stabilize after an infinite time, or an infinite number of steps. The limiting distribution of a Markov chain is stationary; i.e., it does not change over time, or $\pi\mathcal{T} = \pi$.

If a Markov chain is *irreducible*, i.e., each state can be reached from any other state in a finite, however large, number of steps, and *aperiodic*, i.e., we can return to the same state in an unknown number of steps, and not with a fixed periodicity; it will have a unique limiting distribution (see Meyn and Tweedie [85] for proof and explanation).¹⁴

A Markov chain is said to be (*time*) *reversible* if the transition kernel satisfies the so-called *detailed balance* condition:

$$\pi(x)\mathcal{T}(x, y) = \pi(y)\mathcal{T}(y, x), \quad (\text{A8})$$

which basically means that the probability of being in a state x and going to a state y , proposed with the transition kernel $\mathcal{T}(x, y)$, must be the same as being in a state y and going back to the state x . Any Markov chain with a transition kernel $\mathcal{T}$ that satisfies Eq. (A8) will have a stationary distribution π . This can be shown by reading the stationary distribution condition simply as “if we are in a state x described by the limiting distribution π , and we go to a new state y through the transition kernel $\mathcal{T}(x, y)$, also the new state y will be described by the limiting distribution π .”

The probability of being at a state y can then be computed as the probability of being in x and going to y , for all possible states x , i.e.,

$$\begin{aligned} p(y) &= \int dx \pi(x) \mathcal{T}(x, y) = \int dx \pi(y) \mathcal{T}(y, x) \\ &= \pi(y) \int dx \mathcal{T}(y, x) = \pi(y), \end{aligned} \quad (\text{A9})$$

where, in the last step, we used the fact that the transition kernel is normalized $\int \mathcal{T}(y, x) dx = 1$. Therefore, the detailed balance condition ensures that the probability of the new state y is also described by the limiting distribution π .

In MCMC samplers, the goal is to build Markov chains whose stationary distribution corresponds to the distribution we want to reconstruct. In Bayesian inference, the desired stationary distribution is the posterior.

There exist different MCMC algorithms; in the following, we will discuss the Metropolis-Hastings one [53,86], which can be then easily generalized to the reversible jump case.

3. Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm for MCMC sampling can be summarized as follows. Suppose we are in a state x_i in the Markov Chain, to move to the next state x_{i+1} one needs to

- (1) Propose the new state x_{i+1} , generating it from the proposal distribution $q(x_{i+1}|x_i)$, which describes the probability of x_{i+1} given x_i (and only x_i , since Markov chains depend only on the previous point).
- (2) Decide whether to accept the new state or not, based on the *acceptance probability* α . To decide whether the proposed point is accepted, a random value r between $[0, 1]$ is drawn: if $\alpha \geq r$, the proposed x_{i+1} is accepted and saved as the next state in the chain, otherwise it is rejected and x_i is saved again as the chain state.

The acceptance ratio is computed as

$$\alpha = \min \left\{ 1, \frac{\pi(\theta_{i+1})q(\theta_{i+1}|\theta_i)}{\pi(\theta_i)q(\theta_i|\theta_{i+1})} \right\}. \quad (\text{A10})$$

The expression for the acceptance probability is not unique, but Ref. [87] showed how this form optimizes efficiency.

The acceptance probability in Eq. (A10) was derived by Hastings in Ref. [14] in the following way: assume that the chain is in a state x and wants to go to a state y in the same space. Based on the Metropolis-Hastings algorithm, the transition kernel can be written as $\mathcal{T}(x, y) = q(y|x)\alpha(x, y)$, where we explicitly wrote the two probabilities of proposing a new point, $q(x|y)$, and accepting it, $\alpha(x, y)$. The *detailed balance* condition [Eq. (A8)] now reads

$$\pi(x)q(y|x)\alpha(x, y) = \pi(y)q(x|y)\alpha(y, x). \quad (\text{A11})$$

Hastings proposed to write the acceptance probability as

$$\alpha(x, y) = \frac{s(x, y)}{1 + \frac{\pi(x)q(y|x)}{\pi(y)q(x|y)}}, \quad (\text{A12})$$

¹⁴If the space of states is continuous, also the so-called “positive recurrence” property is needed, meaning that, starting from any state, we can always go back in a finite number of steps.

with the only condition that s must be symmetric, i.e., $s(x, y) = s(y, x)$. With this expression for the acceptance ratio, Eq. (A11) can be rewritten as

$$\pi(x)q(y|x) \frac{s(x, y)}{\frac{\pi(y)q(x|y) + \pi(x)q(y|x)}{\pi(y)q(x|y)}} = \pi(y)q(x|y) \frac{s(y, x)}{\frac{\pi(x)q(y|x) + \pi(y)q(x|y)}{\pi(x)q(y|x)}} \quad (\text{A13})$$

$$\begin{aligned} & \frac{\pi(x)q(y|x)s(x, y)\pi(y)q(x|y)}{\pi(y)q(x|y) + \pi(x)q(y|x)} \\ &= \frac{\pi(y)q(x|y)s(y, x)\pi(x)q(y|x)}{\pi(x)q(y|x) + \pi(y)q(x|y)}. \end{aligned} \quad (\text{A14})$$

Therefore, provided that $s(x, y)$ is symmetric, this form of the acceptance probability ensures that detailed balance is satisfied.

An easy choice for $s(x, y)$ is

$$s(x, y) = \begin{cases} 1 + \frac{\pi(x)q(y|x)}{\pi(y)q(x|y)} & \text{if } \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)} \geq 1 \\ 1 + \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)} & \text{if } \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)} < 1. \end{cases} \quad (\text{A15})$$

Therefore, the acceptance ratio becomes

$$\alpha(x, y) = \begin{cases} 1 & \text{if } \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)} \geq 1 \\ \frac{1 + \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)}}{1 + \frac{\pi(x)q(y|x)}{\pi(y)q(x|y)}} & \text{if } \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)} < 1, \end{cases} \quad (\text{A16})$$

or

$$\alpha(x, y) = \min \left\{ 1, \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)} \right\}. \quad (\text{A17})$$

This expression can be further simplified if the chosen proposal function is symmetric, i.e., if $q(x|y) = q(y|x)$. In this case, the acceptance probability simply becomes

$$\alpha(x, y) = \min \left\{ 1, \frac{\pi(y)}{\pi(x)} \right\}. \quad (\text{A18})$$

4. Generalizing Metropolis-Hastings

To facilitate the extension to the transdimensional case, we first consider the Metropolis-Hastings MCMC in a more general way. We now consider two states x and x' , belonging to the sets A and A' , respectively, which for now we assume are subsets (more precisely Borel sets) of a larger parameter space, e.g., $A, A' \subseteq \mathbb{R}^n$.

We impose detailed balance as the condition that the probability of a chain to be in a state in set A and going to a state in the set A' must be the same as being in a state in A' and going to a state in the set A . This can be written as

$$\int_A \pi(x)P(x, A')dx = \int_{A'} \pi(x')P(x', A)dx'. \quad (\text{A19})$$

The transition kernel is now $P(x, A')$ (and similarly $P(x', A)$), which describes the probability of being in a state x and ending up in the set A' . This can happen in two ways:

- (1) The chain is in a state $x \in A$, it proposes a move to $x' \in A'$, and the move is accepted.
- (2) The chain is in a state $x \in A'$, it proposes a move anywhere, but the move is not accepted, and therefore, the chain remains in the set A' .

This can be written as

$$\begin{aligned} P(x, A') &= \underbrace{\int_{A'} q(x'|x)\alpha(x, x')dx'}_{\text{case 1}} \\ &+ \underbrace{\left[\int q(x'|x)[1 - \alpha(x, x')]dx' \right] \cdot I_{\{x \in A'\}}}_{\text{case 2}}, \end{aligned} \quad (\text{A20})$$

where $[1 - \alpha(x, x')]$ denotes the probability of *not* accepting the move to x' , and $I_{\{x \in A'\}}$ is the indicator function, i.e.,

$$I_{\{x \in A'\}} = \begin{cases} 1 & \text{if } x \in A' \\ 0 & \text{if } x \notin A'. \end{cases} \quad (\text{A21})$$

Explicitly writing these two contributions in the detailed balance Eq. (A19), we obtain

$$\begin{aligned} & \int_A dx \pi(x) \left[\int_{A'} q(x'|x)\alpha(x, x')dx' + r(x) \cdot I_{\{x \in A'\}} \right] \\ &= \int_{A'} dx' \pi(x') \left[\int_A q(x|x')\alpha(x', x)dx + r(x') \cdot I_{\{x' \in A\}} \right], \end{aligned} \quad (\text{A22})$$

where we defined

$$r(x) = \int q(x'|x)[1 - \alpha(x, x')]dx' \quad (\text{A23})$$

and

$$r(x') = \int q(x|x')[1 - \alpha(x', x)]dx. \quad (\text{A24})$$

Using

$$\int_A dx \cdot I_{\{x \in A'\}} = \int_{A \cap A'} dx, \quad (\text{A25})$$

the equation above becomes

$$\begin{aligned}
& \int_A \int_{A'} dx dx' \pi(x) q(x'|x) \alpha(x, x') + \int_{A \cap A'} dx \pi(x) r(x) \\
&= \int_{A'} \int_A dx' dx \pi(x') q(x|x') \alpha(x', x) + \int_{A \cap A'} dx' \pi(x') r(x').
\end{aligned} \tag{A26}$$

The second terms on the right- and left-hand side are the same; therefore, it reduces to

$$\begin{aligned}
& \int_A \int_{A'} dx dx' \pi(x) q(x'|x) \alpha(x, x') \\
&= \int_{A'} \int_A dx' dx \pi(x') q(x|x') \alpha(x', x),
\end{aligned} \tag{A27}$$

which is called the *integrated detailed balance* condition. One can show that if Eq. (A27) holds for all subsets $A, A' \subseteq \mathbb{R}^n$, then the “regular” detailed balance [e.g., Eq. (A8)] also holds.

5. The transdimensional case

So far we have assumed that the Markov chains always move between spaces with the same dimensionality. However, especially in the case where we want to compare different models, the states involved might belong to spaces with different dimensions. A “solution” to this issue, proposed by Green in Ref. [42], consists in introducing so-called *auxiliary variables* to ensure that, in the end, we always move between spaces with the same dimensionality.

Suppose we want to move between a set with dimension n_k and another one with dimensions $n_{k'}$, i.e., $A \subseteq \mathbb{R}^{n_k}$ and $A' \subseteq \mathbb{R}^{n_{k'}}$. When the chain is in a state $x \in A$, we generate a random variable $\vec{u}$ with dimension r from a distribution $g(\vec{u})$, and, similarly, a random variable $\vec{u}'$ with dimension r' from a distribution $g'(\vec{u}')$ for the chain that is at the state x' in A' . We define a mapping h between the states and auxiliary variables of the two sets, such that

$$(x', \vec{u}') = h(x, \vec{u}) \quad (x, \vec{u}) = h^{-1}(x', \vec{u}'), \tag{A28}$$

where h^{-1} is the inverse function of h .

To ensure that the detailed balance is respected, the only condition upon the deterministic mapping h is to be a diffeomorphism, which in turn requires *dimension-matching*, i.e.,

$$n_k + r = n_{k'} + r'. \tag{A29}$$

Now let us look at the detailed balance condition in Eq. (A27). When working with auxiliary variables, the probability of proposing a new state reduces to the probability of drawing the specific random variable involved in the proposal, since the jump is then dictated by the deterministic transformation [Eq. (A28)]. If the chain is in the current state x , the new point x' is proposed through the function $h(x, \vec{u})$. Since h is a deterministic mapping, it in turn depends on the drawn value for the

random auxiliary variable $\vec{u}$, thus the probability of proposing the new point x' corresponds to the probability of drawing a specific value $\vec{u}$, i.e., it is given by $g(\vec{u})$. Therefore, when working with maps between spaces, Eq. (A27) can be rewritten as

$$\begin{aligned}
& \int_A \int_{A'} \pi(x) g(\vec{u}) \alpha(x, x') dx d\vec{u} \\
&= \int_{A'} \int_A \pi(x') g'(\vec{u}') \alpha(x', x) dx' d\vec{u}',
\end{aligned} \tag{A30}$$

where we remind the reader that $\pi(x)$ is the probability of being in state x , $g(\vec{u})$ is the probability of drawing the random variable $\vec{u}$, and $\alpha(x, x')$ represents the probability to accept the jump from state x to state x' , and similarly for the right-hand side of the equation.

Since we demand that the mapping h is a diffeomorphism, h and its inverse h^{-1} are both differentiable, we can perform a change of variables in the right-hand side integral in Eq. (A30) as $dx' d\vec{u}' = \frac{\partial h(x, \vec{u})}{\partial(x, \vec{u})} dx d\vec{u}$, obtaining

$$\begin{aligned}
& \int_A \int_{A'} \pi(x) g(\vec{u}) \alpha(x, x') dx d\vec{u} \\
&= \int_{A'} \int_A \pi(x') g'(\vec{u}') \alpha(x', x) \left| \frac{\partial h(x, \vec{u})}{\partial(x, \vec{u})} \right| dx d\vec{u},
\end{aligned} \tag{A31}$$

where the term $\left| \frac{\partial h(x, \vec{u})}{\partial(x, \vec{u})} \right|$ is the Jacobian arising from the change of variables between $(x, \vec{u})$ and $(x', \vec{u}')$. This implies

$$\pi(x) g(\vec{u}) \alpha(x, x') = \pi(x') g'(\vec{u}') \alpha(x', x) \left| \frac{\partial h(x, \vec{u})}{\partial(x, \vec{u})} \right|, \tag{A32}$$

and, repeating the derivation in Sec. A 3, we find that the acceptance probability now reads

$$\alpha(x, x') = \min \left\{ 1, \frac{\pi(x') g'(\vec{u}')}{\pi(x) g(\vec{u})} \left| \frac{\partial h(x, \vec{u})}{\partial(x, \vec{u})} \right| \right\}. \tag{A33}$$

6. Reversible jump MCMC

As discussed earlier, when using RJMCMC algorithms in Bayesian inference to compare models or hypotheses, the goal is to sample the joint posterior probability $p(\vec{\theta}_k, k | d)$ of the parameters *and* the model describing the data. Therefore, the chains move between states given by $x = (\vec{\theta}_k, k)$, with k being the label denoting the model and $\vec{\theta}_k$ the parameter vector for that specific model. At each step the sampler has to perform two kinds of moves:

- (1) *Within-model move*: within a fixed model k , update the model’s parameters $\vec{\theta}_k$ with a standard MCMC procedure (e.g., the Metropolis Hastings algorithm described above).

(2) *Between-models move*: propose a jump to a different model, thus updating k and $\vec{\theta}_k$ simultaneously.

For the *within-model moves*, a “standard” Metropolis-Hastings algorithm can be employed. In this move, we want to go from a state $x = (\vec{\theta}_k, k)$ to a state $(\vec{\theta}'_k, k)$. Since the model k is fixed, we can drop its identifying label in the states. Writing explicitly the stationary distribution π that we want to reconstruct as the joint posterior $p(\vec{\theta}_k, k|d)$, we find that the acceptance ratio in Eq. (A10) becomes

$$\begin{aligned} \alpha(\vec{\theta}_k, \vec{\theta}'_k) &= \min \left\{ 1, \frac{p(\vec{\theta}'_k|d)q(\vec{\theta}_k|\vec{\theta}'_k)}{p(\vec{\theta}_k|d)q(\vec{\theta}'_k|\vec{\theta}_k)} \right\} \\ &= \min \left\{ 1, \frac{\mathcal{L}(d|\vec{\theta}'_k, k)p(\vec{\theta}'_k|k)q(\vec{\theta}_k|\vec{\theta}'_k)}{\mathcal{L}(d|\vec{\theta}_k, k)p(\vec{\theta}_k|k)q(\vec{\theta}'_k|\vec{\theta}_k)} \right\}, \end{aligned} \quad (\text{A34})$$

$$\alpha(x, x') = \min \left\{ 1, \frac{\mathcal{L}(d|\vec{\theta}'_k, k')p(k'|\vec{\theta}'_k)g'(\vec{u}')}{\mathcal{L}(d|\vec{\theta}_k, k)p(k|\vec{\theta}_k)g(\vec{u})} \left| \frac{\partial h(\vec{\theta}_k, \vec{u})}{\partial(\vec{\theta}_k, \vec{u})} \right| \right\} = \min \left\{ 1, \frac{\mathcal{L}(d|\vec{\theta}'_k, k')p(k')p(\vec{\theta}'_k|k')g'(\vec{u}')}{\mathcal{L}(d|\vec{\theta}_k, k)p(k)p(\vec{\theta}_k|k)g(\vec{u})} \left| \frac{\partial h(\vec{\theta}_k, \vec{u})}{\partial(\vec{\theta}_k, \vec{u})} \right| \right\}, \quad (\text{A35})$$

where the evidence cancels out. We also used the fact that the state prior can be written as $p(k, \vec{\theta}_k) = p(k)p(\vec{\theta}_k|k)$; however, this time the term describing the model prior does not cancel out since different models can have different prior probabilities.

APPENDIX B: RJMCMC DETAILS FOR T-ROO

1. Full proposal derivation and inverse moves

In the following, we show how the move proposals in Eqs. (15)–(17) in Sec. III are derived, and how we obtain the proposals for the inverse move.

Let us start from the proposal for the tidal parameters and look at the case BNS $\rightarrow$ BBH as an example. As discussed in the main text, we update the tidal parameter through a stretch move according to Eq. (16). If now we want to go back to the BNS case, we write the proposal for the new $\Lambda'_{1,j}$ in the BNS model as

$$\Lambda'_{1,\text{BNS}} = \Lambda_{1,j} + u_1(\Lambda_{1,\text{BBH}} - \Lambda_{1,j}). \quad (\text{B1})$$

For the transformation to be reversible, we want that this proposal allows us to go back to the “original” point in the BNS model, i.e., to the $\Lambda_{1,\text{BNS}}$ from which we proposed the move to $\Lambda_{1,\text{BBH}}$ in Eq. (15). Substituting the $\Lambda_{1,\text{BBH}}$ proposal in Eq. (16) into Eq. (B1) leads to

$$\Lambda'_{1,\text{BNS}} = \Lambda_{1,j} + u_1(\Lambda_{1,j} + w_1(\Lambda_{1,\text{BNS}} - \Lambda_{1,j}) - \Lambda_{1,j}) \quad (\text{B2})$$

$$= \Lambda_{1,j} - u_1 w_1 \Lambda_{1,j} + u_1 w_1 \Lambda_{1,\text{BNS}} \quad (\text{B3})$$

$$= (1 - u_1 w_1) \Lambda_{1,j} + u_1 w_1 \Lambda_{1,\text{BNS}}. \quad (\text{B4})$$

where in the second step we applied Bayes’ theorem [see Eq. (A5)], and $\mathcal{L}$ denotes the likelihood term. For the prior, we used the fact that the joint model-parameters prior can be factorized in $p(k, \vec{\theta}_k) = p(k)p(\vec{\theta}_k|k)$, and the $p(k)$ terms cancel out because the chain is moving within the same model. Since the evidence is a normalization constant in the posterior probabilities, it is the same across all states; therefore, it cancels out in the acceptance probability, and we need to compute only the likelihood and the prior of each state.

In the *between-model move*, instead, the acceptance ratio derived for the general and transdimensional case in Eq. (A33) must be used. This time, the chain is in a state $x = (\vec{\theta}_k, k)$ and wants to move to $x' = (\vec{\theta}'_{k'}, k')$. Again writing explicitly the stationary distribution as the joint posterior $p(\vec{\theta}_k, k|d)$, we find

This shows that the condition for this move to be reversible is

$$u_1 w_1 = 1. \quad (\text{B5})$$

Similarly, one can show that the reversibility condition for the Λ_2 proposal reads

$$u_2 w_2 = 1. \quad (\text{B6})$$

Therefore, the move between auxiliary variables is

$$\begin{aligned} \text{BNS} \rightarrow \text{BBH}: u_1 &= \frac{1}{w_1}, & u_2 &= \frac{1}{w_2} \quad \text{and} \\ \text{BBH} \rightarrow \text{BNS}: w_1 &= \frac{1}{u_1}, & w_2 &= \frac{1}{u_2}. \end{aligned} \quad (\text{B7})$$

In the case of moves between an NSBH and a BBH model, both $\Lambda_{1,\text{BBH}}$ and $\Lambda_{1,\text{NSBH}}$ are *pseudo*-parameters. Hence, we do not propose a stretch move in this case and instead just keep the same values when jumping to the other model. This also implies that the auxiliary variables u_1, v_1 are not needed in this proposal.¹⁵

For the chirp mass and mass ratio proposals, the idea, outlined in Sec. III A, is to move along a line with slope determined by the likelihood center-of-mass samples in the preliminary sampling, translating this line in the $\mathcal{M}_c - \tilde{\Lambda}$ or $q - \tilde{\Lambda}$ space to make it pass through the current sample

¹⁵We remind the reader that we have auxiliary variables (u_1, u_2) , (v_1, v_2) , and (w_1, w_2) in the BBH, NSBH, and BNS spaces, respectively.

point. Following the slope definition in Eq. (13), the line equation, in the BBH $\rightarrow$ BNS case, reads

$$(\mathcal{M}_{c,\text{BBH}} - \mathcal{M}_{c,\text{BNS}}) = s_{\mathcal{M}_c(\text{BBH,BNS})} \cdot (\tilde{\Lambda}_{\text{BBH}} - \tilde{\Lambda}_{\text{BNS}}), \quad (\text{B8})$$

where the bar on $\tilde{\Lambda}$ denotes that $\tilde{\Lambda}$ is computed assuming, for the *pseudo*-parameters, the physical values $\Lambda_{1,\text{BBH}} = \Lambda_{2,\text{BBH}} = \Lambda_{1,\text{NSBH}} = 0$, which therefore imply $\tilde{\Lambda}_{\text{BBH}} = 0$. Therefore, the chirp-mass proposal in Eq. (16) becomes

$$\begin{aligned} \mathcal{M}_{c,\text{BBH}} &= \mathcal{M}_{c,\text{BNS}} + s_{\mathcal{M}_c(\text{BBH,BNS})}(\tilde{\Lambda}_{\text{BBH}} - \tilde{\Lambda}_{\text{BNS}}) \\ &= \mathcal{M}_{c,\text{BNS}} - s_{\mathcal{M}_c(\text{BBH,BNS})}\tilde{\Lambda}_{\text{BNS}}, \end{aligned} \quad (\text{B9})$$

and similarly we can derive the proposal for the mass ratio.

Since the mapping in RJMCMC must be invertible, the proposals for the moves in the opposite direction, i.e., BBH $\rightarrow$ BNS, can be found by simply inverting this mapping.

In the specific case of the chirp mass proposal when going from the BBH to the BNS model, we have to use

$$\mathcal{M}'_{c,\text{BNS}} = \mathcal{M}_{c,\text{BBH}} + s_{\mathcal{M}_c(\text{BBH,BNS})}\tilde{\Lambda}'_{\text{BNS}}, \quad (\text{B10})$$

where the prime denotes the new point proposed in the BNS space. For the transformation to be reversible, we need to ensure that we can go back to the starting point in the BNS space, i.e., that we can have $\mathcal{M}'_{c,\text{BNS}} = \mathcal{M}_{c,\text{BNS}}$. Plugging Eq. (B10) into Eq. (B10),

$$\begin{aligned} \mathcal{M}'_{c,\text{BNS}} &= \mathcal{M}_{c,\text{BNS}} - s_{\mathcal{M}_c(\text{BBH,BNS})}\tilde{\Lambda}_{\text{BNS}} + s_{\mathcal{M}_c(\text{BBH,BNS})}\tilde{\Lambda}'_{\text{BNS}} \\ &= \mathcal{M}_{c,\text{BNS}} - s_{\mathcal{M}_c(\text{BBH,BNS})}(\tilde{\Lambda}_{\text{BNS}} - \tilde{\Lambda}'_{\text{BNS}}). \end{aligned} \quad (\text{B11})$$

This shows that the transformation is reversible when $\tilde{\Lambda}_{\text{BNS}} = \tilde{\Lambda}'_{\text{BNS}}$, which in turn is ensured by the conditions $u_1 w_1 = 1$ and $u_2 w_2 = 1$ derived above.

In the same way, one can derive the full proposals for the moves between BBH $\rightarrow$ NSBH and NSBH $\rightarrow$ BNS [see Eqs. (15) and (17), respectively], and their inverse. For completeness, we list the inverse mappings here:

$$\text{NSBH to BBH} \left\{ \begin{array}{l} \tilde{\Theta}_{\text{BBH},i+1} = \tilde{\Theta}_{\text{NSBH},i} \\ \Lambda_{1,\text{BBH},i+1} = \Lambda_{1,\text{NSBH},i} \\ \Lambda_{2,\text{BBH},i+1} = \Lambda_{2,j,i} + v_2(\Lambda_{2,\text{NSBH},i} - \Lambda_{2,j,i}) \\ \mathcal{M}_{c,\text{BBH},i+1} = \mathcal{M}_{c,\text{NSBH},i} - s_{\mathcal{M}_c(\text{NSBH,BBH})} \cdot \tilde{\Lambda}_{\text{NSBH},i} \\ q_{\text{BBH},i+1} = q_{\text{BBH},i} - s_{q(\text{NSBH,BBH})} \cdot \tilde{\Lambda}_{\text{NSBH},i} \\ u_2 = 1/v_2 \end{array} \right. \quad (\text{B12})$$

$$\text{BBH to BNS} \left\{ \begin{array}{l} \tilde{\Theta}_{\text{BNS},i+1} = \tilde{\Theta}_{\text{BBH},i} \\ \Lambda_{1,\text{BNS},i+1} = \Lambda_{1,j,i} + u_1(\Lambda_{1,\text{BBH},i} - \Lambda_{1,j,i}) \\ \Lambda_{2,\text{BNS},i+1} = \Lambda_{2,j,i} + u_2(\Lambda_{1,\text{BBH},i} - \Lambda_{2,j,i}) \\ \mathcal{M}_{c,\text{BNS},i+1} = \mathcal{M}_{c,\text{BBH},i} + s_{\mathcal{M}_c(\text{BBH,BNS})}\tilde{\Lambda}_{\text{BNS},i+1} \\ q_{\text{BNS},i+1} = q_{\text{BBH},i} + s_{q(\text{BBH,BNS})}\tilde{\Lambda}_{\text{BNS},i+1} \\ w_1 = 1/u_1 \\ w_2 = 1/u_2 \end{array} \right. \quad (\text{B13})$$

$$\text{BNS to NSBH} \left\{ \begin{array}{l} \tilde{\Theta}_{\text{NSBH},i+1} = \tilde{\Theta}_{\text{BNS},i} \\ \Lambda_{1,\text{NSBH},i+1} = \Lambda_{1,j,i} + w_1(\Lambda_{1,\text{BNS},i} - \Lambda_{1,j,i}) \\ \Lambda_{2,\text{NSBH},i+1} = \Lambda_{2,j,i} + w_2(\Lambda_{1,\text{BNS},i} - \Lambda_{2,j,i}) \\ \mathcal{M}_{c,\text{NSBH},i+1} = \mathcal{M}_{c,\text{BNS},i} + s_{\mathcal{M}_c(\text{BNS,NSBH})}(\tilde{\Lambda}_{\text{NSBH},i+1} - \tilde{\Lambda}_{\text{BNS},i}) \\ q_{\text{BNS},i+1} = q_{\text{NSBH},i} + s_{q(\text{BNS,NSBH})}(\tilde{\Lambda}_{\text{NSBH},i+1} - \tilde{\Lambda}_{\text{BNS},i}) \\ v_1 = 1/w_1 \\ v_2 = 1/w_2 \end{array} \right. \quad (\text{B14})$$

2. Jacobians

As discussed in Sec. A 5, it is essential to include the Jacobian in the acceptance ratio computation for the between-model move. In the following, we compute the Jacobian specifically for the move BNS $\rightarrow$ BBH, i.e., for the mapping in Eq. (16). The Jacobians for all the other transformations can be computed with the same procedure.

The Jacobian for the transformation between the BNS and BBH space, i.e., from $(\vec{\Theta}_{\text{BNS}}, \Lambda_{1,\text{BNS}}, \Lambda_{2,\text{BNS}}, \mathcal{M}_{c,\text{BNS}}, q_{\text{BNS}}, w_1, w_2)$ to $(\vec{\Theta}_{\text{BBH}}, \Lambda_{1,\text{BBH}}, \Lambda_{2,\text{BBH}}, \mathcal{M}_{c,\text{BBH}}, q_{\text{BBH}}, u_1, u_2)$, is given by

$$\mathbb{J}_{\text{BNS, BBH}} = \begin{vmatrix} \frac{\partial \vec{\Theta}_{\text{BBH}}}{\partial \vec{\Theta}_{\text{BNS}}} & \frac{\partial \vec{\Theta}_{\text{BBH}}}{\partial \Lambda_{1,\text{BNS}}} & \frac{\partial \vec{\Theta}_{\text{BBH}}}{\partial \Lambda_{2,\text{BNS}}} & \frac{\partial \vec{\Theta}_{\text{BBH}}}{\partial \mathcal{M}_{c,\text{BNS}}} & \frac{\partial \vec{\Theta}_{\text{BBH}}}{\partial q_{\text{BNS}}} & \frac{\partial \vec{\Theta}_{\text{BBH}}}{\partial w_1} & \frac{\partial \vec{\Theta}_{\text{BBH}}}{\partial w_2} \\ \frac{\partial \Lambda_{1,\text{BBH}}}{\partial \vec{\Theta}_{\text{BNS}}} & \frac{\partial \Lambda_{1,\text{BBH}}}{\partial \Lambda_{1,\text{BNS}}} & \frac{\partial \Lambda_{1,\text{BBH}}}{\partial \Lambda_{2,\text{BNS}}} & \frac{\partial \Lambda_{1,\text{BBH}}}{\partial \mathcal{M}_{c,\text{BNS}}} & \frac{\partial \Lambda_{1,\text{BBH}}}{\partial q_{\text{BNS}}} & \frac{\partial \Lambda_{1,\text{BBH}}}{\partial w_1} & \frac{\partial \Lambda_{1,\text{BBH}}}{\partial w_2} \\ \frac{\partial \Lambda_{2,\text{BBH}}}{\partial \vec{\Theta}_{\text{BNS}}} & \frac{\partial \Lambda_{2,\text{BBH}}}{\partial \Lambda_{1,\text{BNS}}} & \frac{\partial \Lambda_{2,\text{BBH}}}{\partial \Lambda_{2,\text{BNS}}} & \frac{\partial \Lambda_{2,\text{BBH}}}{\partial \mathcal{M}_{c,\text{BNS}}} & \frac{\partial \Lambda_{2,\text{BBH}}}{\partial q_{\text{BNS}}} & \frac{\partial \Lambda_{2,\text{BBH}}}{\partial w_1} & \frac{\partial \Lambda_{2,\text{BBH}}}{\partial w_2} \\ \frac{\partial \mathcal{M}_{c,\text{BBH}}}{\partial \vec{\Theta}_{\text{BNS}}} & \frac{\partial \mathcal{M}_{c,\text{BBH}}}{\partial \Lambda_{1,\text{BNS}}} & \frac{\partial \mathcal{M}_{c,\text{BBH}}}{\partial \Lambda_{2,\text{BNS}}} & \frac{\partial \mathcal{M}_{c,\text{BBH}}}{\partial \mathcal{M}_{c,\text{BNS}}} & \frac{\partial \mathcal{M}_{c,\text{BBH}}}{\partial q_{\text{BNS}}} & \frac{\partial \mathcal{M}_{c,\text{BBH}}}{\partial w_1} & \frac{\partial \mathcal{M}_{c,\text{BBH}}}{\partial w_2} \\ \frac{\partial q_{\text{BBH}}}{\partial \vec{\Theta}_{\text{BNS}}} & \frac{\partial q_{\text{BBH}}}{\partial \Lambda_{1,\text{BNS}}} & \frac{\partial q_{\text{BBH}}}{\partial \Lambda_{2,\text{BNS}}} & \frac{\partial q_{\text{BBH}}}{\partial \mathcal{M}_{c,\text{BNS}}} & \frac{\partial q_{\text{BBH}}}{\partial q_{\text{BNS}}} & \frac{\partial q_{\text{BBH}}}{\partial w_1} & \frac{\partial q_{\text{BBH}}}{\partial w_2} \\ \frac{\partial u_1}{\partial \vec{\Theta}_{\text{BNS}}} & \frac{\partial u_1}{\partial \Lambda_{1,\text{BNS}}} & \frac{\partial u_1}{\partial \Lambda_{2,\text{BNS}}} & \frac{\partial u_1}{\partial \mathcal{M}_{c,\text{BNS}}} & \frac{\partial u_1}{\partial q_{\text{BNS}}} & \frac{\partial u_1}{\partial w_1} & \frac{\partial u_1}{\partial w_2} \\ \frac{\partial u_2}{\partial \vec{\Theta}_{\text{BNS}}} & \frac{\partial u_2}{\partial \Lambda_{1,\text{BNS}}} & \frac{\partial u_2}{\partial \Lambda_{2,\text{BNS}}} & \frac{\partial u_2}{\partial \mathcal{M}_{c,\text{BNS}}} & \frac{\partial u_2}{\partial q_{\text{BNS}}} & \frac{\partial u_2}{\partial w_1} & \frac{\partial u_2}{\partial w_2} \end{vmatrix} \quad (\text{B15})$$

Looking at the mapping in Eq. (16) and computing all the derivatives, we find

$$\mathbb{J}_{\text{BNS, BBH}} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & w_1 & 0 & 0 & 0 & (\Lambda_{1,\text{BNS}} - \Lambda_{1,j}) & 0 \\ 0 & 0 & w_2 & 0 & 0 & 0 & (\Lambda_{2,\text{BNS}} - \Lambda_{2,j}) \\ 0 & -s_{\mathcal{M}_{c,(\text{BBH,BNS})}} \left| \frac{\partial \vec{\Lambda}_{\text{BNS}}}{\partial \Lambda_{1,\text{BNS}}} \right| & -s_{\mathcal{M}_{c,(\text{BBH,BNS})}} \left| \frac{\partial \vec{\Lambda}_{\text{BNS}}}{\partial \Lambda_{2,\text{BNS}}} \right| & 1 & 0 & 0 & 0 \\ 0 & -s_{q_{(\text{BBH,BNS})}} \left| \frac{\partial \vec{\Lambda}_{\text{BNS}}}{\partial \Lambda_{1,\text{BNS}}} \right| & -s_{q_{(\text{BBH,BNS})}} \left| \frac{\partial \vec{\Lambda}_{\text{BNS}}}{\partial \Lambda_{2,\text{BNS}}} \right| & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{w_1^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{w_2^2} \end{vmatrix}$$

$$= w_1 w_2 \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{w_1^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{w_2^2} \end{vmatrix} = \frac{1}{w_1 w_2}. \quad (\text{B16})$$

When we compute the acceptance ratio for this move, i.e., Eq. (18), we need to include this Jacobian and the probability of drawing the auxiliary variables, which are given by $g(z) = 1/\sqrt{z}$ (see Sec. III). For the auxiliary variables in the BNS, we have $g_{\text{BNS}}(w_1) = 1/\sqrt{w_1}$ and $g_{\text{BNS}}(w_2) = 1/\sqrt{w_2}$. In the BBH space, instead, $g_{\text{BBH}}(u_1) = 1/\sqrt{u_1} = \sqrt{w_1}$ and $g_{\text{BBH}}(u_2) = 1/\sqrt{u_2} = \sqrt{w_2}$, where we have used the reversibility condition $u_1 w_1 = 1$ and $u_2 w_2 = 1$. Together with the Jacobian, the factor in the acceptance probability becomes

$$\frac{g_{\text{BBH}}(u_1, u_2)}{g_{\text{BNS}}(w_1, w_2)} \mathbb{J}_{\text{BNS, BBH}} = \frac{\sqrt{w_1 w_2}}{\frac{1}{\sqrt{w_1 w_2}}} \cdot \frac{1}{w_1 w_2} = 1. \quad (\text{B17})$$

Similarly one can compute the Jacobian for all the other moves, finding in all cases

$$\frac{g'(x')}{g(x)} \mathbb{J}_{xx'} = 1, \quad (\text{B18})$$

as shown in Eq. (19).

3. Symmetry conditions for stretch moves

The stretch move employed in `t-roc` proposals is described in Sec. II A. The acceptance ratio for the stretch move shown in Eq. (5) includes a term to preserve the symmetry condition in the proposal. Note that this is different from the detailed balance condition, which includes in its definition the acceptance ratio itself. Here, we are interested in enforcing that the kernel to propose a new point given the current one is symmetric. This simplifies the calculation of the acceptance probability in the algorithm, but is not a necessary condition such as the detailed balance one, which is instead needed to ensure that the limiting distribution of the chain is stationary.

Assume the chain is in X_k and we use a stretch move to propose a new point Y with

$$Y = X_j + z(X_k - X_j), \quad (\text{B19})$$

where X_j is the current position of another chain chosen randomly and z is a random variable drawn from a distribution $g(z)$. Then, from Y , we propose a new point Y' with

$$Y' = X_j + z'(Y - X_j). \quad (\text{B20})$$

To ensure that the proposal is symmetric, i.e., that the probability of going in one direction or the other is the same, we first then need to check whether it is possible that $Y' = X_k$. Combining Eqs. (B19) and (B20),

$$Y' = X_j + z'(Y - X_j) \quad (\text{B21})$$

$$= X_j + z'(X_j + z(X_k - X_j) - X_j) \quad (\text{B22})$$

$$= X_j + z'zX_k - zz'X_j. \quad (\text{B23})$$

Therefore, in order to have $Y' = X_k$, we need $zz' = 1$.

To find the general symmetry condition for the proposal, we now consider the probability of proposing one point from another one, integrated over all the possible starting points. For the proposals in Eqs. (B19) and (B20),

$$\int P(Y|X_k)dY = \int g(z)dz, \quad (\text{B24})$$

$$\int P(Y'|Y)dY' = \int g(z')dz', \quad (\text{B25})$$

since the probability of a specific point being proposed with a stretch move corresponds to the probability of a certain value for z being chosen.

Stretch moves can be pictured as moving on a sphere with the radius given by the distance to the position of the other randomly selected chain X_j , and centered in the origin X_j . Switching to radial coordinates, in a n -dimensional space,

$$dY = d\Omega_{n-1}/, dr/, r^{n-1} \quad \text{with radius } r = z|X_k - X_j|^n \quad (\text{B26})$$

$$dY' = d\Omega_{n-1}/, dr'/, r'^{n-1} \quad \text{with radius } r' = z'|Y - X_j|^n. \quad (\text{B27})$$

Changing variables with $dr = |X_k - X_j|dz$ and $dr' = |Y - X_j|dz'$, from Eqs. (B25), we find

$$P(Y|X_k) = \frac{g(z)dz}{d\Omega_{n-1}dz|X_k - X_j|z^{n-1}|X_k - X_j|^{n-1}} \quad (\text{B28})$$

and

$$P(Y'|Y) = \frac{g(z')dz'}{d\Omega_{n-1}dz'|Y - X_j|z'^{n-1}|Y - X_j|^{n-1}}. \quad (\text{B29})$$

In a symmetric proposal, we want to go back to the starting point, i.e., $P(Y|X_k) = P(Y'|Y)$. Therefore, we can find the symmetry condition as

$$1 = \frac{P(Y|X_k)}{P(Y'|Y)} \quad (\text{B30})$$

$$= \frac{g(z)}{g(z')} \frac{z^{n-1}|Y - X_j|^n}{z'^{n-1}|X_k - X_j|^n} \quad (\text{B31})$$

$$= \frac{g(z)}{g(1/z)} \left(\frac{1}{z}\right)^{2n-2} \frac{z^n|X_k - X_j|^n}{|X_k - X_j|^n} \quad (\text{B32})$$

$$= \frac{g(z)}{g(1/z)} z^{2-n}, \quad (\text{B33})$$

where in the third line we imposed $Y' = X_k$, i.e., $z' = 1/z$, and in the end we used $|Y - X_j|^n = z^n|X_k - X_j|^n$, from Eq. (B19).

Therefore, the most general condition in n dimensions is

$$\frac{g(z)}{g(1/z)} z^{2-n} = 1. \quad (\text{B34})$$

If we consider the one-dimensional case, $n = 1$ and

$$g(z)z = g(1/z). \quad (\text{B35})$$

Adding the generic symmetric factor to the acceptance probability in Eq. (A10) allows us to cancel out the specific proposal distributions, i.e.,

$$\alpha = \min \left\{ 1, \frac{\pi(\theta_{i+1})q(\theta_{i+1}|\theta_i)}{\pi(\theta_i)q(\theta_i|\theta_{i+1})} \frac{g(z)}{g(1/z)} z^{2-n} \frac{g(1/z)}{g(z)} z^{n-2} \right\} \quad (\text{B36})$$

$$= \min \left\{ 1, \frac{\pi(\theta_{i+1})}{\pi(\theta_i)} \frac{g(1/z)}{g(z)} z^{n-2} \right\}. \quad (\text{B37})$$

With the specific random variable distribution $g(z)$ in Eq. (4),

$$\frac{g(1/z)}{g(z)} z^{n-2} = \frac{1}{\sqrt{1/z}} \frac{1}{1/\sqrt{z}} z^{n-2} = z^{n-1}, \quad (\text{B38})$$

and therefore, we recover the expression for the acceptance ratio in Eq. (A10).

4. Priors on *pseudo*-parameters

As explained in Sec. III, in `t-roo`, we add the *pseudo*-parameters $\Lambda_{1,\text{BBH}}$, $\Lambda_{2,\text{BBH}}$, and $\Lambda_{1,\text{NSBH}}$ to store the information about tidal parameters in models that would not include such parameters. In this section, we show that the prior on these *pseudo*-parameters has to (crucially) be included in the acceptance ratio in order to preserve detailed balance.

For simplicity, let us look at the case where we attempt a move between a BBH and NSBH model, with the proposal in Eq. (15). The move happens between the spaces $(\vec{\theta}_{\text{BBH}}, \Lambda_{1,\text{BBH}}, \Lambda_{2,\text{BBH}}, u_2)$ for the BBH model and $(\vec{\theta}_{\text{NSBH}}, \Lambda_{1,\text{NSBH}}, v_2)$ for the NSBH one, with u_2 and v_2 being the auxiliary variables. In the detailed balance condition in Eq. (A32),

$$\pi(x)g(\vec{u})\alpha(x, x') = \pi(x')g'(\vec{u}')\alpha(x', x) \left| \frac{\partial h(x, \vec{u})}{\partial(x, \vec{u})} \right|, \quad (\text{B39})$$

in this specific case, the state x corresponds to $(k = \text{BBH}, \vec{\theta}_k = \vec{\theta}_{\text{BBH}})$, while the state x' corresponds to $(k = \text{NSBH}, \vec{\theta}_k = \vec{\theta}_{\text{NSBH}})$. The stationary distribution π corresponds to the posterior probability densities $\pi(x) = p(k = \text{BBH}, \vec{\theta}_{\text{BBH}}|d) := \text{pos}_{\text{BBH}}(\vec{\theta}_{\text{BBH}})$ and $\pi(x') = p(k = \text{NSBH}, \vec{\theta}_{\text{NSBH}}|d) := \text{pos}_{\text{NSBH}}(\vec{\theta}_{\text{NSBH}})$, and the deterministic mapping $(\vec{\theta}_{\text{BBH}}, \vec{u}) = h(\vec{\theta}_{\text{NSBH}}, \vec{v})$ and $(\vec{\theta}_{\text{NSBH}}, \vec{v}) = h^{-1}(\vec{\theta}_{\text{BBH}}, \vec{u})$ are given by Eqs. (14) and (15), respectively. In the following, we explicitly write the model label as pos_k , instead of including it in the state as in $\pi(k, \vec{\theta}_k)$. Therefore,

$$\text{pos}_{\text{BBH}}g(\vec{u})\alpha_{\text{BBH} \rightarrow \text{NSBH}} = \text{pos}_{\text{NSBH}}g'(\vec{v})\alpha_{\text{NSBH} \rightarrow \text{BBH}} \mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}, \quad (\text{B40})$$

where $\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}$ denotes the Jacobian relative to the mapping between the NSBH and BBH space, and for now we dropped the parameters' dependence in the posterior for simplicity.

Following the derivation of the expression for the acceptance ratio discussed in Sec. A 3, we can define

$$\alpha_{\text{BBH} \rightarrow \text{NSBH}} = \frac{s_{\text{BBH} \rightarrow \text{NSBH}}}{1 + \frac{\text{pos}_{\text{BBH}}g(u)}{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}} \quad (\text{B41})$$

and similarly for $\alpha_{\text{NSBH} \rightarrow \text{BBH}}$. As before, the only constraint is that s is symmetric, i.e., $s_{\text{BBH} \rightarrow \text{NSBH}} = s_{\text{NSBH} \rightarrow \text{BBH}}$. Hence, we can define

$$s_{\text{BBH} \rightarrow \text{NSBH}} = \begin{cases} 1 + \frac{\text{pos}_{\text{BBH}}g(u)}{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}} & \text{if } \frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} \geq 1 \\ 1 + \frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} & \text{if } \frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} < 1, \end{cases} \quad (\text{B42})$$

which results in

$$\alpha_{\text{BBH} \rightarrow \text{NSBH}} = \begin{cases} 1 & \text{if } \frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} \geq 1 \\ \frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} & \text{if } \frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} < 1. \end{cases} \quad (\text{B43})$$

Focusing on the case $\frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} < 1$, since in the other one simply $\alpha_{\text{BBH} \rightarrow \text{NSBH}} = 1$, we can use Bayes' theorem to write the acceptance probability as

$$\alpha_{\text{BBH} \rightarrow \text{NSBH}} = \frac{\text{pos}_{\text{NSBH}}g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\text{pos}_{\text{BBH}}g(u)} = \frac{\mathcal{L}_{\text{NSBH}}(\vec{\theta}_{\text{NSBH}}, \Lambda_{1,\text{NSBH}})p_{\text{NSBH}}(\vec{\theta}_{\text{NSBH}}, \Lambda_{1,\text{NSBH}})g'(v)\mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\mathcal{L}_{\text{BBH}}(\vec{\theta}_{\text{BBH}}, \Lambda_{1,\text{BBH}}, \Lambda_{2,\text{BBH}})p_{\text{BBH}}(\vec{\theta}_{\text{BBH}}, \Lambda_{1,\text{BBH}}, \Lambda_{2,\text{BBH}})g(u)}, \quad (\text{B44})$$

where $\mathcal{L}$ denotes the likelihood and p the prior. Factorizing the prior, we obtain

$$\alpha_{\text{BBH} \rightarrow \text{NSBH}} = \frac{\mathcal{L}_{\text{NSBH}}(\vec{\theta}_{\text{NSBH}}, \Lambda_{1,\text{NSBH}}) p_{\text{NSBH}}(\vec{\theta}_{\text{NSBH}}) p_{\text{NSBH}}(\Lambda_{1,\text{NSBH}}) g'(v) \mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\mathcal{L}_{\text{BBH}}(\vec{\theta}_{\text{BBH}}, \Lambda_{1,\text{BBH}}, \Lambda_{2,\text{BBH}}) p_{\text{BBH}}(\vec{\theta}_{\text{BBH}}) p_{\text{BBH}}(\Lambda_{1,\text{BBH}}) p_{\text{BBH}}(\Lambda_{2,\text{BBH}}) g(u)} \quad (\text{B45})$$

$$= \frac{\mathcal{L}_{\text{NSBH}}(\vec{\theta}_{\text{NSBH}}) p_{\text{NSBH}}(\vec{\theta}_{\text{NSBH}}) p_{\text{NSBH}}(\Lambda_{1,\text{NSBH}}) g'(v) \mathbb{J}_{\text{BBH} \rightarrow \text{NSBH}}}{\mathcal{L}_{\text{BBH}}(\vec{\theta}_{\text{BBH}}) p_{\text{BBH}}(\vec{\theta}_{\text{BBH}}) p_{\text{BBH}}(\Lambda_{1,\text{BBH}}) p_{\text{BBH}}(\Lambda_{2,\text{BBH}}) g(u)}, \quad (\text{B46})$$

where in the last step we used the fact that the likelihood does *not* depend on the *pseudo*-parameters.

Equation (B46) shows how, even if the *pseudo*-parameters do not enter the likelihood calculation, we have to include their prior in the acceptance ratio calculation in order to respect the detailed balance.

APPENDIX C: CONVERGENCE CRITERIA

As discussed in Sec. IV C, it is not trivial to establish a criterion to automatically check the convergence of RJMCMC analyses [see, e.g., Refs. [88,89] for some examples and discussion]. By default, `t-roo` runs for a set number of steps and, if needed, the run can be resumed and run for longer. However, we also implemented some “empirical” convergence criteria that can be used to decide when to stop the sampler, both in the preliminary and main RJMCMC sampling phases. All the different stopping criteria available in `t-roo` are described below.

For the preliminary sampling stage, we implemented two stopping criteria:

- (i) *Number of steps* (default): the preliminary sampling is stopped after a given number of steps, defined by the user. Although different analyses might require a different number of steps, we find that in general $\mathcal{O}(10^3)$ is sufficient and not too computationally expensive.
- (ii) *Likelihood convergence*: the preliminary sampling is stopped when 80% of the walkers have converged to the maximum likelihood among them within a given tolerance, where the default tolerance is set to 0.1. The idea behind this criterion is that when the sampler finishes exploring the parameter space and finds the region(s) where the support for each model lives, the likelihood of the samples around that region will stabilize around the maximum value and yield similar values among the various walkers.

For the main RJMCMC sampling, we explored three different stopping criteria. In this case, we do not look at the likelihood but at the model probabilities, because in RJMCMC the number of samples in each model should converge to the respective probabilities of that model. The underlying idea is that, once the sampler is converged, each walker will ideally spend an amount of time (i.e., samples) in each model corresponding to that model probability:

- (i) *Number of steps* (default): the run is stopped after a specified number of iterations.
- (ii) *Walkers mean*: we compute the average probability among walkers of each model in the last n steps, and we stop the run when this average does not change for a given number of steps.
- (iii) *Multinomial*: the run is stopped when the standard deviation between each model probability as computed from different walkers is less than 1.3 times the standard deviation expected from a multinomial distribution for N_w walkers.

In order to save computational time, users can decide after how many iterations to start checking whether stopping criteria are satisfied, and how often to check after. In any case, the run can always be resumed and run for more steps.

APPENDIX D: ROBUSTNESS OF EVIDENCE ESTIMATION

Throughout this work, we have compared the `t-roo` model probabilities to the ones obtained by the default nested sampler in `bilby`, i.e., `dynesty`. In order to check the robustness of the evidence calculation, we here compare its results to evidences obtained by different methods, focusing on the injection `NSBH_LAM600`; results are shown in Fig. 8.

We repeat the `dynesty` analysis for the `NSBH_LAM600` injection with different sampling seeds, using 2000 live points in all cases.¹⁶ The odds ratio estimate within `dynesty` is robust, with the fluctuations due to different sampling seeds lying within the uncertainty band; therefore, it constitutes an appropriate reference to assess the reliability of the sampler we developed.

Similarly, we repeat the `t-roo` analysis with different sampling seeds: also in this case, we see fluctuations in the recovered odds ratio value, but the estimate is overall consistent across the different sampling seeds. The blue and red Gaussians in Fig. 8 represent the distribution of odds ratio values from the `dynesty` and `t-roo` analyses of this injection, respectively, with mean and standard deviation derived from the distribution of 10 runs with different sampling seeds.

¹⁶Reference [90] showed in Sec. A2 that the true evidence consistently lies within the errors provided by `dynesty` 68% of the times when more than 1000 live points are used, even though the error on `dynesty` evidence is not technically Gaussian.

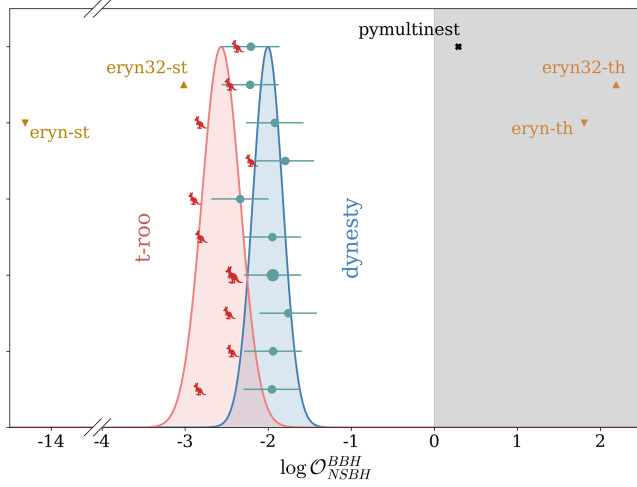


FIG. 8. Comparison between odds ratios for the NSBH_LAM600 run with different samplers. The gray shaded area marks the region in which the BBH model is incorrectly favored. The blue and red Gaussians show the expected distributions of odds ratio from *dynesty* and *t-roo*, respectively, as inferred from the results of 10 analyses with different sampling seeds. The markers show the $\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}}$ recovered by each run, with the larger marker highlighting the run whose results are reported in Table II; in the case of *dynesty*, the error bar corresponds to the conservative estimate of the uncertainty as described in Sec. V. For *eryn* runs, “th” and “st”, refer to evidences computed via the thermodynamic integration and stepping stone methods, respectively, while “eryn-32” refers to the *eryn* runs with 32 temperatures.

Furthermore, we also analyze this injection with a different nested sampler, i.e., *pymultinest* [91,92]. In this case, the recovered $\ln \mathcal{O}_{\text{NSBH}}^{\text{BBH}}$ deviates substantially from the ones recovered by *dynesty*, even favoring the BBH model instead of the NSBH one.

Since *t-roo* is based on *eryn*, we finally check the evidences obtained by single-model *eryn* runs, without the specific within- and between-model moves described above. MCMC algorithms sample directly the target posterior distribution, and consequently do not automatically provide Bayesian evidence. However, one can exploit the parallel tempering chains to compute them. In particular, two different methods are implemented in *eryn*: *thermodynamic integration* [93] and *stepping stone* [94]. Given that the accuracy in evidence estimates increases with the number of parallel temperature chains employed, we additionally repeated these runs with 32 temperatures instead of 16. The odds ratios obtained with *eryn* show large differences not only with respect to the *dynesty* runs, lying well outside the uncertainty region and, in the case of thermodynamic integration, even favoring the wrong model. However, it is known that accurate thermodynamic evidence requires a sufficiently large number of temperatures, and that evidence computed with the stepping stone algorithm require fewer temperatures than standard thermodynamic integration. Therefore, we expect the *eryn32-st* evidence to be most accurate, and this is indeed the run that agrees with the *t-roo* evidence. Most likely the inaccuracies in the other odds ratio estimations are due to insufficient temperatures. This shows another advantage of *t-roo*, meaning that it can obtain good estimates of odds ratios with less temperatures (i.e., further reducing the computational cost of the analysis).

Overall, Fig. 8 shows how different methods can provide quite different estimates of evidence values, and consequently odds ratios, calling for general caution when basing strong claims on them. Nevertheless, the odds ratio estimation in *t-roo* appears to be quite robust, with minimum fluctuations across the results for different sampling seeds.

-
- [1] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), [arXiv:2605.27225](#).
 - [2] N. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), [arXiv:2605.27226](#).
 - [3] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **119**, 161101 (2017).
 - [4] B. P. Abbott *et al.* (SKA South Africa/MeerKAT Collaborations), *Astrophys. Lett.* **848**, L12 (2017).
 - [5] B. P. Abbott *et al.* (MASTER Collaborations), *Nature (London)* **551**, 85 (2017).
 - [6] T. Dietrich, M. W. Coughlin, P. T. H. Pang, M. Bulla, J. Heinzl, L. Issa, I. Tews, and S. Antier, *Science* **370**, 1450 (2020).
 - [7] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), [arXiv:2605.27227](#).
 - [8] R. Abbott *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), *Phys. Rev. D* **112**, 084080 (2025).
 - [9] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), [arXiv:2603.19019](#).
 - [10] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), [arXiv:2603.19020](#).
 - [11] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), [arXiv:2603.19021](#).

- [12] J. Skilling, *Bayesian Anal.* **1**, 833 (2006).
- [13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [14] W. K. Hastings, *Biometrika* **57**, 97 (1970).
- [15] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
- [16] F. Acernese *et al.* (VIRGO Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [17] T. Akutsu *et al.*, *Prog. Theor. Exp. Phys.* **2021**, 05A101 (2021).
- [18] J. S. Speagle, *Mon. Not. R. Astron. Soc.* **493**, 3132 (2020).
- [19] G. Ashton and C. Talbot, *Mon. Not. R. Astron. Soc.* **507**, 2037 (2021).
- [20] G. Ashton *et al.*, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [21] G. Ashton and T. Dietrich, *Nat. Astron.* **6**, 961 (2022).
- [22] A. Puecher, A. Samajdar, G. Ashton, C. Van Den Broeck, and T. Dietrich, *Phys. Rev. D* **109**, 023019 (2024).
- [23] A. G. Abac *et al.* (LIGO Scientific Collaboration, VIRGO Collaboration, and KAGRA Collaboration), *arXiv:2508.18083*.
- [24] <https://emfollow.docs.ligo.org/userguide/> (2026), LIGO/Virgo Public Alerts User Guide.
- [25] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Astrophys. J. Lett.* **896**, L44 (2020).
- [26] A. G. Abac *et al.* (LIGO Scientific Collaboration, KAGRA Collaboration, and VIRGO Collaboration), *Astrophys. J. Lett.* **970**, L34 (2024).
- [27] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Astrophys. J. Lett.* **892**, L3 (2020).
- [28] C. Barbieri, O. S. Salafia, M. Colpi, G. Ghirlanda, and A. Perego, *Astron. Astrophys.* **654**, A12 (2021).
- [29] R. Dudi, A. Adhikari, B. Brügmann, T. Dietrich, K. Hayashi, K. Kawaguchi, K. Kiuchi, K. Kyutoku, M. Shibata, and W. Tichy, *Phys. Rev. D* **106**, 084039 (2022).
- [30] R. J. Foley, D. A. Coulter, C. D. Kilpatrick, A. L. Piro, E. Ramirez-Ruiz, and J. Schwab, *Mon. Not. R. Astron. Soc.* **494**, 190 (2020).
- [31] K. Kyutoku, S. Fujibayashi, K. Hayashi, K. Kawaguchi, K. Kiuchi, M. Shibata, and M. Tanaka, *Astrophys. J. Lett.* **890**, L4 (2020).
- [32] M.-Z. Han, S.-P. Tang, Y.-M. Hu, Y.-J. Li, J.-L. Jiang, Z.-P. Jin, Y.-Z. Fan, and D.-M. Wei, *Astrophys. J. Lett.* **891**, L5 (2020).
- [33] S. Khadkikar and D. Singh, *Phys. Rev. D* **113**, 023029 (2026).
- [34] I. Tews, P. T. H. Pang, T. Dietrich, M. W. Coughlin, S. Antier, M. Bulla, J. Heinzl, and L. Issa, *Astrophys. J. Lett.* **908**, L1 (2021).
- [35] L. L. Lopes and D. P. Menezes, *Astrophys. J.* **936**, 41 (2022).
- [36] B. K. K. Lee, M.-c. Chu, and L.-M. Lin, *Astrophys. J.* **922**, 242 (2021).
- [37] B. Biswas, R. Nandi, P. Char, S. Bose, and N. Stergioulas, *Mon. Not. R. Astron. Soc.* **505**, 1600 (2021).
- [38] I. Bombaci, A. Drago, D. Logoteta, G. Pagliara, and I. Vidaña, *Phys. Rev. Lett.* **126**, 162702 (2021).
- [39] H. Koehn, T. Wouters, H. Rose, P. T. H. Pang, R. Somasundaram, I. Tews, and T. Dietrich, *Phys. Rev. D* **110**, 103015 (2024).
- [40] J. Janquart *et al.*, *Mon. Not. R. Astron. Soc.* **537**, 1001 (2025).
- [41] I. Markin, A. Puecher, M. Bulla, and T. Dietrich, *Phys. Rev. D* **113**, 024031 (2026).
- [42] P. J. Green, *Biometrika* **82**, 711 (1995).
- [43] D. Hastie and P. Green, *Stat. Neerl.* **66**, 309 (2012).
- [44] N. Karnesis, M. L. Katz, N. Korsakova, J. R. Gair, and N. Stergioulas, *Mon. Not. R. Astron. Soc.* **526**, 4814 (2023).
- [45] R. Umstätter, N. Christensen, M. Hendry, R. Meyer, V. Simha, J. Veitch, S. Vigeland, and G. Woan, *Classical Quantum Gravity* **22**, S901 (2005).
- [46] R. Umstätter, N. Christensen, M. Hendry, R. Meyer, V. Simha, J. Veitch, S. Vigeland, and G. Woan, *Phys. Rev. D* **72**, 022001 (2005).
- [47] W. M. Farr, I. Mandel, and D. Stevens, *R. Soc. Open Sci.* **2**, 150030 (2015).
- [48] N. J. Cornish and T. B. Littenberg, *Classical Quantum Gravity* **32**, 135012 (2015).
- [49] T. B. Littenberg and N. J. Cornish, *Phys. Rev. D* **91**, 084034 (2015).
- [50] H. Tong *et al.*, *Astrophys. J. Suppl. Ser.* **276**, 50 (2025).
- [51] K. Chatziioannou, N. Cornish, M. Wijngaarden, and T. B. Littenberg, *Phys. Rev. D* **103**, 044013 (2021).
- [52] S. Y. Cheung, R. Udall, D. Davis, P. D. Lasky, and E. Thrane, *arXiv:2607.09111*.
- [53] W. K. Hastings, *Biometrika* **57**, 97 (1970).
- [54] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [55] J. Goodman and J. Weare, *Commun. Appl. Math. Comput. Sci.* **5**, 65 (2010).
- [56] J. Wang and R. H. Swendsen, *Prog. Theor. Phys. Suppl.* **157**, 317 (2005).
- [57] W. R. Gilks, G. O. Roberts, and S. K. Sahu, *J. Am. Stat. Assoc.* **93**, 1045 (1998).
- [58] D. J. Earl and M. W. Deem, *Phys. Chem. Chem. Phys.* **7**, 3910 (2005).
- [59] W. D. Vousden, W. M. Farr, and I. Mandel, *Mon. Not. R. Astron. Soc.* **455**, 1919 (2016).
- [60] E. Thrane and C. Talbot, *Publ. Astron. Soc. Pac.* **36**, e010 (2019).
- [61] S. Morisaki, *Phys. Rev. D* **104**, 044062 (2021).
- [62] B. Zackay, L. Dai, and T. Venumadhav, *arXiv:1806.08792*.
- [63] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt, *Phys. Rev. D* **94**, 044031 (2016).
- [64] M. Favata, *Phys. Rev. Lett.* **112**, 101101 (2014).
- [65] LIGO Scientific Collaboration, Noise Curves Used for Simulations in the Update of the Observing Scenarios Paper, Technical Report No. LIGO-T2000012-v2 (LIGO Scientific Collaboration, 2022).
- [66] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration, *Living Rev. Relativity* **23**, 3 (2020).
- [67] J. E. Thompson, E. Fauchon-Jones, S. Khan, E. Nitoglia, F. Pannarale, T. Dietrich, and M. Hannam, *Phys. Rev. D* **101**, 124059 (2020).
- [68] F. Pannarale, E. Berti, K. Kyutoku, B. D. Lackey, and M. Shibata, *Phys. Rev. D* **92**, 084050 (2015).
- [69] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044007 (2016).
- [70] T. Dietrich, A. Samajdar, S. Khan, N. K. Johnson-McDaniel, R. Dudi, and W. Tichy, *Phys. Rev. D* **100**, 044003 (2019).

- [71] G. Pratten, S. Husa, C. Garcia-Quiros, M. Colleoni, A. Ramos-Buades, H. Estelles, and R. Jaume, *Phys. Rev. D* **102**, 064001 (2020).
- [72] L. Pompili *et al.*, *Phys. Rev. D* **108**, 124035 (2023).
- [73] A. Abac, T. Dietrich, A. Buonanno, J. Steinhoff, and M. Ujevic, *Phys. Rev. D* **109**, 024062 (2024).
- [74] M. Branchesi *et al.*, *J. Cosmol. Astropart. Phys.* **07** (2023) 068.
- [75] A. Abac *et al.* (ET Collaboration), *J. Cosmol. Astropart. Phys.* **03** (2026) 081.
- [76] S. Danilishin and T. Zhang, Einstein Telescope sensitivity curves used for CoBA Science study, Technical Report No. ET-0304A-22 (Einstein Telescope Collaboration, 2022).
- [77] A. Gelman and D. B. Rubin, *Stat. Sci.* **7**, 457 (1992).
- [78] E. E. Flanagan and T. Hinderer, *Phys. Rev. D* **77**, 021502 (2008).
- [79] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, *Phys. Rev. D* **89**, 103012 (2014).
- [80] <https://doi.org/10.7935/ggb8-1v94>, GWOSC data release GW190425.
- [81] <https://doi.org/10.7935/6xqf-ba54>, GWOSC data release GW230529.
- [82] <https://dcc.ligo.org/LIGO-P2000026/public> (2020), Parameter estimation sample release for GW190425.
- [83] <https://doi.org/10.5281/zenodo.10845779> (2020), Parameter estimation sample release for GW230529.
- [84] F. A. Ramis Vidal, A. Abac, M. Colleoni, T. Dietrich, P. Mourier, A. Gonzalez, I. Markin, and A. Puecher, [arXiv:2606.04810](https://arxiv.org/abs/2606.04810).
- [85] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability* (Taylor & Francis, 1993).
- [86] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [87] P. H. Peskun, *Biometrika* **60**, 607 (1973).
- [88] Y. Fan, S. A. Sisson, and L. Davies, [arXiv:1001.2055](https://arxiv.org/abs/1001.2055).
- [89] S. P. Brooks and P. Giudici, *J. Comput. Graph. Stat.* **9**, 266 (2000).
- [90] I. M. Romero-Shaw *et al.*, *Mon. Not. R. Astron. Soc.* **499**, 3295 (2020).
- [91] F. Feroz, M. P. Hobson, and M. Bridges, *Mon. Not. R. Astron. Soc.* **398**, 1601 (2009).
- [92] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski, *Astron. Astrophys.* **564**, A125 (2014).
- [93] N. Lartillot and H. Philippe, *Syst. Biol.* **55**, 195 (2006).
- [94] W. Xie, P. O. Lewis, Y. Fan, L. Kuo, and M.-H. Chen, *Syst. Biol.* **60**, 150 (2011).