

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences

C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy

Phys. Rev. D **92**, 023002 — Published 6 July 2015

DOI: [10.1103/PhysRevD.92.023002](https://doi.org/10.1103/PhysRevD.92.023002)

# A novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences

C. Pankow,<sup>\*</sup> P. Brady,<sup>†</sup> and E. Ochsner<sup>‡</sup>

*Center for Gravitation, Cosmology, and Astrophysics,  
University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA*

R. O’Shaughnessy<sup>§</sup>

*Center for Computational Relativity and Gravitation,  
Rochester Institute of Technology, Rochester, NY 14623, USA and  
Center for Gravitation, Cosmology, and Astrophysics,  
University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA*

We introduce a highly-parallelizable architecture for estimating parameters of compact binary coalescence using gravitational-wave data and waveform models. Using a spherical harmonic mode decomposition, the waveform is expressed as a sum over modes that depend on the intrinsic parameters (e.g. masses) with coefficients that depend on the observer dependent extrinsic parameters (e.g. distance, sky position). The data is then prefiltered against those modes, at fixed intrinsic parameters, enabling efficient evaluation of the likelihood for generic source positions and orientations, independent of waveform length or generation time. We efficiently parallelize our intrinsic space calculation by integrating over all extrinsic parameters using a Monte Carlo integration strategy. Since the waveform generation and prefiltering happens only once, the cost of integration dominates the procedure. Also, we operate hierarchically, using information from existing gravitational-wave searches to identify the regions of parameter space to emphasize in our sampling. As proof of concept and verification of the result, we have implemented this algorithm using standard time-domain waveforms, processing each event in less than one hour on recent computing hardware. For most events we evaluate the marginalized likelihood (evidence) with statistical errors of  $\lesssim 5\%$ , and even smaller in many cases. With a bounded runtime independent of the waveform model starting frequency, a nearly-unchanged strategy could estimate NS-NS parameters in the 2018 advanced LIGO era. Our algorithm is usable with any noise curve and existing time-domain model at any mass, including some waveforms which are computationally costly to evolve.

## I. INTRODUCTION

The upcoming ground based gravitational-wave detector network (notably including advanced LIGO [1] and advanced Virgo [2]) are sensitive to the gravitational-wave signal from coalescing compact binaries, both the relatively well understood signal from the earlier inspiral phase [3–14] and the less well understood strong-field merger [15–22]. Both of these regimes are important in understanding the underlying physical processes and properties of the binary which likely provide the central engine for phenomena such as short gamma-ray bursts [23–25]. A multimessenger observation of this category of event would be the first of its kind and could answer open questions about these poorly understood events. In the era of multimessenger astronomy, rapid and robust measurements of candidate compact binary gravitational-wave events will be a critical science product for gravitational-wave observatories, as colleagues with other instruments perform followup and coincident observations [26]. The most tantalizing proposed elec-

tromagnetic counterparts are expected to be brief, potentially disappearing within days if not shorter [27–33]. Given limited resources, reliable low-latency parameter estimation of gravitational-wave signals will significantly enhance the science output of multimessenger astronomy.

With the scheduled resumption of data taking in late 2015 [26], the second generation gravitational-wave interferometers in Hanford, Washington, and Livingston, Louisiana, are expected to reach unprecedented sensitivities [34]. The Virgo detector is also expected to resume data taking within a year of this milestone. In preparation for the next run, the LIGO and Virgo Collaborations have implemented and extensively tested a set of low latency gravitational-wave detection pipelines [33, 35–37], capable of compact binary event detection within a few minutes (or less) from the coalescence time. These pipelines trade the ability to accurately determine all but a few of the parameters of the coalescence for speed and breadth of analysis.

Having a more accurate estimation of the binary coalescence’s parameters is valuable not only to gravitational-wave science but also to electromagnetic observatories to guide their pointing. Many astrophysical phenomena which could create transient gravitational waves also have electromagnetic signatures which may decay rapidly. Moreover, gravitational-wave interferometer networks often cannot localize gravitational-

---

<sup>\*</sup>Electronic address: [pankow@gravity.phys.uwm.edu](mailto:pankow@gravity.phys.uwm.edu)

<sup>†</sup>Electronic address: [patrick@gravity.phys.uwm.edu](mailto:patrick@gravity.phys.uwm.edu)

<sup>‡</sup>Electronic address: [evano@gravity.phys.uwm.edu](mailto:evano@gravity.phys.uwm.edu)

<sup>§</sup>Electronic address: [oshaughn@mail.rit.edu](mailto:oshaughn@mail.rit.edu)

wave events to better than a few hundred square degrees on the sky [26, 33, 38], implying follow-up observations should occur expeditiously after the initial identification to promptly localize as accurately as possible. While several Bayesian algorithms for gravitational-wave parameter estimation have been employed in the past, the time scale of a full parameter estimation analysis remains at a much higher latency than the initial detection.

With these goals in mind, and moving hierarchically from the broader search result, fast follow-ups have been developed to rapidly localize the event in the sky, assuming the search pipelines have provided accurate timing and masses [33, 39]. However, the most robust interpretation of gravitational-wave data requires systematically comparing all possible candidate signals to the data, constructing a Bayesian posterior probability distribution for candidate binary parameters [40–47]. These schemes provide measurements of all the physical parameters of the coalescence, albeit at a much higher latency than the search itself. Owing to the complexity and multimodality of the model space, these strategies have adapted variants of Markov Chain Monte Carlo (MCMC) or nested sampling [47, 48] algorithms to estimate the parameters of the coalescence. In physics, similar path-based methods have been enormously successful at a broad range of physical problems, by exploring all possible paths through a configuration space; see, e.g. [49, 50].

Though successful, the nature of these algorithms is functionally serial, with moderate degrees of parallelization requiring intensive communication to coordinate the current state. Without state-of-the-art techniques like ensemble sampling, MCMC and nested sampling techniques therefore do not scale efficiently to beyond a few tens of cores (e.g., for parallel tempering). In contrast, efficient and highly-parallelizable Monte Carlo integration strategies are frequently applied to problems with dimensions comparable or higher than coalescing compact binaries; see, e.g., [51–54]. In this work, we apply such methods to gravitational-wave parameter estimation for the first time.

This work is organized as follows. In Sec. II we provide an executive summary, outlining the principles that enable our algorithm to provide rapid, accurate results for the test cases explored here: time-domain models for non-spinning, circular binaries. Then, in Sec. III and IV, we outline the pertinent features of our waveform decomposition and describe our algorithm in detail. To demonstrate our algorithm provides high performance in an environment mimicking conditions during an observational run, Sec. V presents results drawn from a large sample of events from the “2015 double neutron star mock data challenge,” results from which are described in [33]. In Sec. VI, we discuss the broader significance of our result in the context of other parameter estimation work inside and outside the community.

## II. EXECUTIVE SUMMARY

In this paper, we outline an alternative architecture for gravitational-wave parameter estimation, based on three key ingredients: efficient evaluation of the likelihood as a function of extrinsic parameters, highly-parallelized Monte Carlo integration over a grid of intrinsic parameters, and input derived from gravitational-wave search and sky localization pipelines.

The computational cost to evaluate the likelihood has historically been a limiting step in gravitational-wave parameter estimation. This cost has two factors in each comparison which is required to be made. The first is the cost of waveform generation, as discussed in [55–59] and references therein. Second is the filtering cost related to the required length of the waveform to be compared — this scales in proportion to the lowest frequency accessible by a gravitational-wave interferometer. At fixed sampling rates, the lowest frequency content of the waveform dominates a time-domain likelihood computation. Recently, several methods have been proposed to perform this comparison more efficiently [55–57, 60], by interpolating some combination of the waveform or likelihood or by adopting a sparse representation to reduce the computational cost of data handling. In this work, we introduce a robust and straightforward scheme to reduce the cost per comparison, without resorting to interpolation within or manipulation of the waveform family. We first perform a “precomputation” phase for each intrinsic parameter by decomposing each physically distinct source in the spin-weighted spherical harmonic basis. This allows us to quickly evaluate the likelihood as a function of extrinsic parameters. This method is applicable to any available waveform model, and the computational cost of the precomputation step requires the waveform to be generated and decomposed only once, thereby making this step a small fractional cost of the total computation. To our knowledge, this work is the first time any such method has been implemented for the most accurate but computationally-expensive waveform models like EOB [61, 62] while leveraging several hundred computing cores at once.

Taking advantage of the fast likelihood computation, each set of intrinsic parameters has a straightforward and vectorizable Monte Carlo integration performed to explore and marginalize over the extrinsic parameters. Monte Carlo integration can provide an exhaustive search of the parameter space with a fixed cost in convergence scaling with the number of points drawn. As each Monte Carlo draw is a statistically independent sample, this procedure can engage as many computational resources as are necessary to provide fast convergence of the integral to a satisfactory level of uncertainty. Efficient Monte Carlo integration schemes adjust their sampling strategy while performing the integration by adapting to sample only the meaningful regions of the space. By using detection search results (e.g. those produced by matched-filtered searches [63]) to sample near maxima of the like-

likelihood, Monte Carlo methods can retain generality but still be extremely efficient. We describe a procedure to explore the extrinsic parameter space and evaluate the marginalized likelihood that is embarrassingly parallel, extensively incorporates information provided to accelerate convergence, and provides a well-understood sampling error estimate, allowing us to target a given level of precision.

To complete the parameter estimation process, our algorithm evaluates the likelihood and accumulates intrinsic posterior samples using a well-motivated discrete grid whose position and size is derived from the existing information provided by gravitational-wave detection pipelines. To further accelerate convergence, we also use the information provided by the search to prioritize specific combinations of extrinsic parameters for further investigation. Like the process of marginalizing the likelihood at fixed intrinsic parameters, this grid search is embarrassingly parallel. Typically, parameter estimation strategies make almost no use of the information reported by existing search codes, precisely to avoid self-consistency issues that can arise by, for example, using inferences for the data as priors on the reanalysis of that data. In contrast, by carefully distinguishing between sampling and integration priors, we circumvent these issues.

To summarize, combining these three factors (parallelizable Monte Carlo integration, efficient likelihoods, and search information), we can provide reliable parameter estimates for merging double neutron star binaries within one hour on existing hardware for instruments available in the next five years. This promising and simple first implementation has performance comparable to or better than well-developed state-of-the-art Markov Chain Monte Carlo codes like LALINFERENCE applied to initial detector data with comparable nonprecessing sources. For advanced instruments, with longer signals and more costly waveforms, our method should perform significantly better than the current LALINFERENCE implementation with state-of-the-art nonprecessing binaries. Our method therefore provides a valuable alternative to other promising methods like reduced-order modeling, ensemble sampling, and other tactics to improve scaling and decrease latency [64, 65]. Owing to its superior scaling and transparency, our method provides a well understood and easily implemented alternative to LALINFERENCE for suitable problems, particularly valuable for low-latency and approximate parameter estimation and for investigations into waveform systematics.

### III. BINARY WAVEFORMS

#### A. Intrinsic and extrinsic parameters

On physical grounds, we group waveform parameters  $\vec{\mu} = (\vec{\lambda}, \vec{\theta})$  into two classes: the intrinsic parameters ( $\vec{\lambda}$ ) and extrinsic parameters ( $\vec{\theta}$ ). The intrinsic parameters

are fundamental to the description of the binary: if we change any intrinsic parameters we must recompute the orbital dynamics of the binary (typically through the relatively expensive process of numerically integrating ordinary differential equations). Extrinsic parameters simply describe how the binary is oriented in space and time relative to the detector; changing extrinsic parameters involves a relatively inexpensive rotation, translation or rescaling transformation. As we will show in the next subsection, for the non-spinning case considered here the intrinsic parameters (transformations of the binary component masses  $m_1$  and  $m_2$ ) are

$$\vec{\lambda} = \{\mathcal{M}_c, \eta\}, \quad (1)$$

where  $\mathcal{M}_c = (m_1 m_2)^{3/5} / (m_1 + m_2)^{1/5}$  is the chirp mass and  $\eta = m_1 m_2 / (m_1 + m_2)^2$  is the symmetric mass ratio. The extrinsic parameters are

$$\vec{\theta} = \{t_{\text{geo}}, \alpha, \delta, \iota, D, \psi, \phi_c\}, \quad (2)$$

where  $t_{\text{geo}}$  is the time at which the coalescence point of the waveform arrives at the Earth geocenter,  $\alpha$  and  $\delta$  are the right ascension and declination,  $\iota$  is the inclination angle of the binary's angular momentum vector and the line of sight to Earth,  $D$  is the luminosity distance to the binary<sup>1</sup>,  $\psi$  is the polarization angle, and  $\phi_c$  is the orbital phase of the binary at coalescence.

#### B. Waveform decomposition

The gravitational wave strain measured by the  $k^{\text{th}}$  interferometric detector in a network is given by

$$h_k(t) = F_{+,k}(\delta, \alpha, \psi) h_{+,k}(t) + F_{\times,k}(\delta, \alpha, \psi) h_{\times,k}(t), \quad (3)$$

where  $F_{+,k}$ ,  $F_{\times,k}$  are the antenna patterns of the detector<sup>2</sup> and  $h_{+,k}$ ,  $h_{\times,k}$  are the two components of the gravitational wave strain, evaluated at the  $k^{\text{th}}$  detector. The antenna patterns depend only on the extrinsic sky location and polarization angle, while the polarizations depend on both intrinsic and extrinsic parameters. Meanwhile, at leading order for inspiral-only waveforms the polarizations are described  $\Phi(t)$  the orbital phase, and the post-Newtonian  $v(t)$  “velocity” parameters, these depending only on combinations of the masses of the binary.

<sup>1</sup> For the sources considered in this paper, the redshift correction is assumed to be negligible.

<sup>2</sup> Due to the rotation of the Earth, the antenna patterns change as a function of time. While this has been mostly neglected due to the signal duration versus the Earth's rotational velocity, an accounting of this effect will become necessary as the instruments become sensitive to longer binary coalescence waveforms.

Thus, it is convenient to introduce  $h(t|\mathcal{M}_c, \eta, \iota, \phi_c)$  given by:

$$h_{+,k}(t) - i h_{\times,k}(t) = h(t - t_k | \mathcal{M}_c, \eta, \iota, \phi_c) . \quad (4)$$

Due to the propagation of the wave, each member of the network will record a different time of arrival at the site depending only on sky location. Time offsets are then measured relative to the geocenter arrival time. The expression,  $t_k$  denotes the time of arrival of the coalescence at the  $k^{th}$  detector,

$$t_k = t_{\text{geo}} + \frac{\vec{x}_k \cdot \hat{N}(\alpha, \delta)}{c} , \quad (5)$$

where  $\vec{x}_k$  is a vector pointing from the geocenter to the  $k^{th}$  detector and  $\hat{N}(\alpha, \delta)$  is the direction of gravitational-wave propagation.

At this stage, we make no assumptions about the functional form of  $h(t|\mathcal{M}_c, \eta, \iota, \phi_c)$ . We can use a  $-2$  spin-weighted spherical harmonic mode decomposition (denoted  $h_{lm}$ ) to further separate intrinsic and extrinsic parameters appearing in the polarizations as

$$h_{+,k}(t) - i h_{\times,k}(t) = \frac{D_{\text{ref}}}{D} \sum_{lm} \hat{h}_{lm}(\mathcal{M}_c, \eta, t_k; t) Y^{(-2)}_{lm}(\iota, -\phi_c), \quad (6)$$

evaluated at some fixed distance  $D_{\text{ref}}$  (in this work we choose  $D_{\text{ref}} = 100$  Mpc).

If we define a complex-valued antenna pattern for each detector as

$$F_k = F_{+,k} + i F_{\times,k} , \quad (7)$$

then we can re-express the measured strain in the  $k^{th}$  detector as

$$h_k(\vec{\lambda}, \vec{\theta}; t) = \text{Re} \frac{D_{\text{ref}}}{D} F_k(\alpha, \delta, \psi) \sum_{lm} \hat{h}_{lm}(\mathcal{M}_c, \eta, t_k; t) Y^{(-2)}_{lm}(\iota, -\phi_c) . \quad (8)$$

Aside from  $t_k$ , we have now completely separated the intrinsic parameters (which enter only the  $\hat{h}_{lm}$ ) from the extrinsic parameters (which enter only the  $F_k$  and  $Y^{(-2)}_{lm}$ ).

Given a time-domain representation of the gravitational wave strain, we can define a frequency-domain version of this strain via a Fourier transform

$$\tilde{h}(f) = \int_{-\infty}^{\infty} h(t) e^{-2\pi i f t} dt . \quad (9)$$

Time translation of frequency-domain waveforms is trivial. If  $\tilde{h}(t_k; f)$  is the Fourier-domain representation of a strain for some arrival time  $t_k$ , then the same strain arriving at another time  $t'_k$  can be simply related by

$$\tilde{h}(t'_k; f) = \tilde{h}(t_k; f) e^{-2\pi i f (t'_k - t_k)} . \quad (10)$$

Thus, if we work with frequency-domain waveforms, the arrival time  $t_k$  can be factored out as  $\exp(-2\pi i f t_k)$  and we can complete the separation of intrinsic and extrinsic parameters.

For the explicit decomposition above, we have focused on non-spinning waveforms and found that they can be separated into two intrinsic parameters and seven extrinsic parameters. In the general case, which could include precession, tidal effects and any other physics, this intrinsic and extrinsic separation is still possible. In fact, there will always be seven extrinsic parameters which enter only through inexpensive geometric factors, while any additional parameters will always be encoded in the expensive  $h_{lm}$  modes.

To see that this is true, first note that Eq. (10) holds for an arbitrary strain and so can always be used to factor out the dependence on time of arrival. Similarly, a gravitational wave far from its source will always fall off

as  $1/D$ . Furthermore, the antenna patterns  $F_+$  and  $F_\times$  depend on the detector geometry, not the source, and so are unchanged. This gives a total of five extrinsic parameters that in general can be factored out exactly as in our non-spinning waveform. Lastly, we have two angles which enter into the  $Y^{(-2)}_{lm}$ . The physical interpretation of these angles depends on the choice of the frame relative to which the harmonic mode decomposition is performed. One should choose a frame which is convenient for expressing and computing the  $h_{lm}$ . For the non-spinning case, this means aligning the frame with the orbital angular momentum,  $\hat{L}$ , in which case one can show that the zenith angle is  $\iota$  and the azimuth is  $-\phi_c$ . In the precessing case, one would most likely use a frame aligned with the total angular momentum,  $\hat{J}$ , e.g. using the parameterization described in [66]. In the notation of that paper, the extrinsic spherical harmonic dependence is  $Y^{(-2)}_{lm}(\theta_{JN}, -\phi_{JL})$ , where  $\theta_{JN}$  is the inclination of the *total* angular momentum to line of sight, and  $\phi_{JL}$  marks the azimuthal position of  $\hat{L}$  on its precession cone about  $\hat{J}$ .

#### IV. METHODS

A set of data  $d_k(t)$  collected from a gravitational-wave detector is typically decomposed into the inherent noise of the detector and a putative gravitational-wave signal:

$$d_k(t) = h_k(t) + n_k(t). \quad (11)$$

In the absence of a signal and under the assumption that each detector produces stationary Gaussian noise, the noise  $\tilde{n}_k(f)$  in the  $k^{\text{th}}$  detector is characterized by its one-sided power spectral density  $S_k(f)$ :

$$\langle \tilde{n}_k(f)^* \tilde{n}_k(f') \rangle = \frac{1}{2} S_k(|f|) \delta(f - f'). \quad (12)$$

In the following discussion, we define a weighted inner-product of two complex Fourier-domain functions  $\tilde{a}(f), \tilde{b}(f)$  with a weighting function  $1/S(f)$ :

$$\langle a|b \rangle \equiv 2 \int_{-\infty}^{\infty} df \frac{\tilde{a}^*(f) \tilde{b}(f)}{S(f)}. \quad (13)$$

The mode decomposed waveforms  $\hat{h}_{lm}$ , can be shifted in time in reference to a detector relative to the geocentric time via Eq. (10). For this reason, the overlap between a single detector data time series  $d(t)$  and a time-shifted complex function  $h(t - t_k)$  is

$$\langle h(t_k)|d \rangle = 2 \int_{-\infty}^{\infty} \frac{\tilde{d}(f) \tilde{h}^*(f)}{S(f)} e^{2\pi i f t_k} df. \quad (14)$$

Note that this is simply the Fourier transform of the integrand in Eq. (13), which means we can compute the overlap for all possible time shifts with a single Fourier transform.

#### A. Bayesian evidence and posteriors

If the detector noise is Gaussian, it follows that the probability of some set of noise realizations in each of our detectors is

$$p(\{d\}|\mathcal{H}_0) \propto \prod_k \exp\left(-\frac{\langle d|d \rangle_k}{2}\right) \quad (15)$$

in the absence of a gravitational-wave signal. Here,  $\mathcal{H}_0$  indicates the hypothesis that the data is only Gaussian noise.

With a gravitational-wave signal present, the data from each detector is noise plus the response of an interferometer to the gravitational-wave  $h(\vec{\mu})$ . In this case, the probability of some measured  $d$  given the presence of the signal with parameter values  $\vec{\mu}$  is

$$p(\{d\}|\vec{\mu}, \mathcal{H}_1) \propto \prod_k \exp\left(-\frac{\langle d - h(\vec{\mu})|d - h(\vec{\mu}) \rangle_k}{2}\right) \quad (16)$$

where  $\mathcal{H}_1$  indicates the hypothesis that the data consists of Gaussian noise plus a gravitational-wave signal.

By Bayes' theorem, the posterior probability of the parameters  $\vec{\mu}$  under the signal present hypothesis  $\mathcal{H}_1$  is

$$p(\vec{\mu}, \mathcal{H}_1|\{d\}) = \frac{p(\{d\}|\vec{\mu}, \mathcal{H}_1)p(\vec{\mu})}{p(\{d\}|\mathcal{H}_1)} \quad (17)$$

where

$$p(\{d\}|\mathcal{H}_1) = \int p(\{d\}|\vec{\mu}, \mathcal{H}_1)p(\vec{\mu})d\vec{\mu}. \quad (18)$$

It is convenient, at this stage, to introduce the likelihood ratio

$$\mathcal{L}(\{d\}|\vec{\mu}) = \prod_k \frac{\exp\{-\langle d - h(\vec{\mu})|d - h(\vec{\mu}) \rangle_k/2\}}{\exp\{-\langle d|d \rangle_k/2\}} \quad (19)$$

and to rewrite Eq. (17) in the form

$$p(\vec{\mu}, \mathcal{H}_1|\{d\}) = \frac{\mathcal{L}(\{d\}|\vec{\mu})p(\vec{\mu})}{Z}, \quad (20)$$

where  $Z$  is defined to be

$$Z = \int \mathcal{L}(\{d\}|\vec{\mu})p(\vec{\mu})d\vec{\mu} \equiv \frac{p(\{d\}|\mathcal{H}_1)}{p(\{d\}|\mathcal{H}_0)}. \quad (21)$$

We can also compute the posterior probability density (which we will abbreviate as simply the posterior) for one or multiple parameters (marginalized over all other parameters). Let  $x$  be one or more parameters in  $\vec{\mu}$  and  $y$  be all other parameters, such that  $\vec{\mu} = x \cup y$ . Then, the posterior for  $x$  is

$$p(\mathcal{H}_1|x) \equiv \frac{p(x|\mathcal{H}_1)}{Z(\{d\}|\mathcal{H}_1)} \int dy p(y|\mathcal{H}_1)\mathcal{L}(x, y). \quad (22)$$

### B. Efficiently evaluating the likelihood

We now show how the waveform decomposition derived above can be exploited to speed up evaluations of the likelihood ratio. In Eq. (8) note that we have taken the observed strain in a detector  $H_k(t)$  and rewritten it as a linear combination of harmonic mode time series  $\hat{h}_{lm}(t)$ . The harmonic mode time series depend on the intrinsic parameters, but the extrinsic parameters (apart from  $t_k$ ) are entirely encoded in the coefficients of the linear combination. Because computing the likelihood is just an inner product, which is a linear operator, we can pull these coefficients outside the inner product integral. Thus, we need only compute inner products involving the  $\hat{h}_{lm}$  and data  $\{d\}$ . If we store these, we can compute the likelihood for any extrinsic parameters by simply re-computing the coefficients and reconstructing the linear combination.

To that end, we define the following quantities:

$$\begin{aligned} Q_{k,lm}(\vec{\lambda}, t_k) &\equiv \langle h_{lm}(\vec{\lambda}, t_k) | d \rangle_k \\ &= 2 \int_{-\infty}^{\infty} \frac{df}{S_k(|f|)} e^{2\pi i f t_k} \tilde{h}_{lm}^*(\vec{\lambda}; f) \tilde{d}(f) , \end{aligned} \quad (23a)$$

$$U_{k,lm,l'm'}(\vec{\lambda}) = \langle h_{lm} | h_{l'm'} \rangle_k , \quad (23b)$$

$$V_{k,lm,l'm'}(\vec{\lambda}) = \langle h_{lm}^* | h_{l'm'} \rangle_k . \quad (23c)$$

Note that each  $Q_{k,lm}(\vec{\lambda}, t_k)$  is computed for all  $t_k$  with a single inverse Fourier transform, as in Eq. (14). A signal will produce a peak in the filtered outputs localized to a short millisecond time window around the coalescence time, so the  $Q_{k,lm}(\vec{\lambda}, t_k)$  will be sharply peaked as functions of  $t_k$  and we need only retain the values for a narrow range of  $t_k$ . To allow for detector arrival times that differ from the geocenter time, the range of  $t_k$  for which we must store the  $Q_{k,lm}$  is set by the light travel time across Earth ( $2R_{\oplus}/c \simeq 42\text{ms}$ ). Conservatively, work we store the  $Q_{k,lm}$  for a much longer 300 ms range of  $t_k$ . The  $U_{k,lm,l'm'}(\vec{\lambda})$  and  $V_{k,lm,l'm'}(\vec{\lambda})$  are independent of  $t_k$  and are computed once by a straightforward use of Eq. (13).

By plugging Eq. (8) into Eq. (19), taking a log and collecting terms, we obtain

---


$$\begin{aligned} \ln \mathcal{L}(\vec{\lambda}; \vec{\theta}) &= (D_{\text{ref}}/D) \text{Re} \sum_k \sum_{lm} (F_k Y^{(-2)}_{lm})^* Q_{k,lm}(\vec{\lambda}, t_k) \\ &\quad - \frac{(D_{\text{ref}}/D)^2}{4} \sum_k \sum_{lm,l'm'} \left[ |F_k|^2 [Y^{(-2)}_{lm}]^* Y^{(-2)}_{l'm'} U_{k,lm,l'm'}(\vec{\lambda}) + \text{Re} \left( F_k^2 Y^{(-2)}_{lm} Y^{(-2)*}_{l'm'} V_{k,lm,l'm'}(\vec{\lambda}) \right) \right] . \end{aligned} \quad (24)$$


---

Importantly, the intrinsic parameters  $\vec{\lambda}$  enter only through the  $Q_{k,lm}$ ,  $U_{k,lm,l'm'}$  and  $V_{k,lm,l'm'}$ . These are the dominant cost, as they require computing the orbital dynamics, the  $h_{lm}$ , inner product integrals and inverse Fourier transforms. By contrast, the extrinsic parameters enter the  $F_k$  and  $Y^{(-2)}_{lm}$ , which are much cheaper to compute.

Therefore, if we fix a point in the intrinsic parameter space we can compute the  $Q_{k,lm}(\vec{\lambda}, t_k)$ ,  $U_{k,lm,l'm'}(\vec{\lambda})$  and  $V_{k,lm,l'm'}(\vec{\lambda})$  only once, vary the extrinsic parameters and compute the likelihood for only the cost of the  $F_k$  and  $Y^{(-2)}_{lm}$ . This allows us to efficiently integrate over the extrinsic parameters and obtain a marginalized posterior for the intrinsic parameters  $p(\vec{\lambda})$  as in Eq. (22). If we do this for a collection of points in the intrinsic parameter

space, we can integrate over  $\vec{\lambda}$  as well and obtain  $Z$  as in Eq. (21). Note that the computation for each point in the intrinsic space is completely independent of the others. This makes the algorithm embarrassingly parallel and given enough CPU cores the entire analysis can be run in the time it takes to integrate over the extrinsic parameters (modulo some brief startup and post-processing steps).

The remainder of this section provides more details on the various steps of our algorithm.

### C. Placement over intrinsic parameters

In the current work, we restrict ourselves to consider non-spinning binaries in which we also neglect tidal effects.

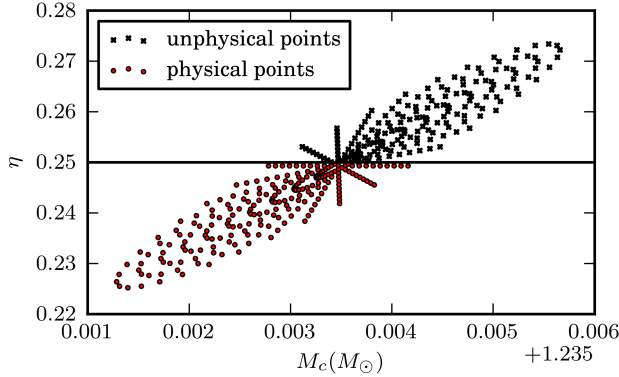


FIG. 1: We use an effective Fisher matrix to compute an approximate ellipsoidal region of overlap  $\geq 90\%$  with the masses reported by a detection pipeline. We then fill this ellipsoid with discrete points and cut any with unphysical values of symmetric mass ratio  $\eta$ . In this case  $m_1 = 1.5M_\odot$  and  $m_2 = 1.35M_\odot$ , so a little under half of the points placed would have been unphysical. At each physical grid point, we marginalize the likelihood over all extrinsic parameters as described in Sec. IV D.

Therefore, we need only two mass parameters to describe the intrinsic parameter space, for which we use the symmetric mass ratio  $\eta = m_1 m_2 / M^2$  and the chirp mass  $\mathcal{M}_c = M\eta^{3/5}$  (where  $M = m_1 + m_2$  is the total mass). Since detection searches will report masses for the candidate event, we use these to guide which region of the intrinsic parameter space to explore.

Let  $(\mathcal{M}^*, \eta^*)$  be the masses reported by a detection pipeline. We then perform an effective Fisher matrix calculation as described in [67, 68] centered about this point. This involves evaluating the overlap

$$\frac{|\langle h(\vec{\lambda}^*) | h(\vec{\lambda}) \rangle|}{\sqrt{\langle h(\vec{\lambda}^*) | h(\vec{\lambda}^*) \rangle \langle h(\vec{\lambda}) | h(\vec{\lambda}) \rangle}} \quad (25)$$

between our waveform with masses  $(\mathcal{M}^*, \eta^*)$  and approximately tens of nearby waveforms with different intrinsic parameters (while extrinsic parameters are held constant). The measured overlap values are then fit with a multi-dimensional quadratic. The coefficients of this quadratic fit are called the effective Fisher matrix. Like the standard Fisher matrix, the effective Fisher matrix serves as a quick, crude estimate of expected parameter estimation performance and can be used to predict surfaces of constant overlap, which will in general be ellipsoids. In this work, we use the effective Fisher matrix to approximate the region of intrinsic parameter space which will have overlap  $\geq 90\%$  with the masses  $(\mathcal{M}^*, \eta^*)$ .

Once we have defined this 90% overlap ellipsoid, we must fill it with a set of discrete points at which we will compute the likelihood. To do this, we first specify the total number of intrinsic parameter points we wish to place, here 200 points. Then, these points are arranged within

a unit sphere. In this work, we placed points along 20 radial lines, with 10 points per spoke. Along each spoke, the points are placed uniformly in radial distance. Now, the eigenvalues and eigenvectors of the effective Fisher matrix tell us the lengths and orientations of the axes of the 90% overlap ellipsoid. We use these to deform and rotate our set of points in the sphere to a set of points in the 90% overlap ellipsoid. Once we have filled our ellipsoid, we remove any points that have unphysical  $\eta$  ( $> 1/4$ ). We ensure that we always place spokes in our ellipsoid along the direction of constant  $\eta$ , so that we always have many points along this boundary of the parameter space. One reason for the choice of spoked placement is so that we can always ensure near-equal mass binaries will have many points near the  $\eta = 0.25$  boundary of the parameter space, the region where most astrophysical binary neutron stars are expected. Fig. 1 illustrates this placement of intrinsic parameter points.

#### D. Integrating over extrinsic parameters

Because precomputed quantities allow us to efficiently evaluate the likelihood  $\mathcal{L}(\vec{\lambda}, \vec{\theta})$  as a function of  $\vec{\theta}$ , we first integrate the likelihood  $\mathcal{L}(\vec{\lambda}, \vec{\theta})$  over all extrinsic parameters  $\vec{\theta}$  to get

$$\mathcal{L}_{\text{red}}(\vec{\lambda}) = \int \mathcal{L}(\vec{\lambda}, \vec{\theta}) p(\vec{\theta}) d\vec{\theta}, \quad (26)$$

where  $p(\vec{\theta})$  is our prior over the extrinsic parameters. We assume the sources analyzed are randomly-oriented and randomly distributed in the universe out to a fiducial radius<sup>3</sup>. With the potential exception of the sky position, our priors are independent, and thus separable. We then evaluate the reduced likelihood  $\mathcal{L}_{\text{red}}$  by integrating over the geocentric time of arrival  $t$ , distance  $D$ , sky position  $\Omega_{\text{sky}}$  represented as right ascension  $\alpha$  and declination  $\delta$ , angular momentum orientation as measured by inclination  $\iota$ , polarization angle  $\psi$ , and coalescence phase  $\phi_c$ , using the separable prior to get

$$\mathcal{L}_{\text{red}}(\vec{\lambda}) = \int \frac{dt}{T_{\text{window}}} \frac{D^2 dD d\Omega_{\text{sky}}}{V_{\text{max}}} \frac{d(\cos \iota) d\phi_c}{4\pi} \frac{d\psi}{\pi} p(\vec{\theta}) \mathcal{L}(\vec{\lambda}, \vec{\theta}). \quad (27)$$

In this expression and our calculations, we adopt a maximum distance  $D_{\text{max}}$  (here 300 Mpc, beyond which we do not expect appreciable sensitivity from a 2015 era interferometer) and a time window  $T_{\text{window}}$  (here, 300 ms) surrounding the event.

<sup>3</sup> Beyond which we do not expect our instruments to be appreciably sensitive in the next two years.



With the exception of time (described below), we evaluate these integrals and reconstruct the posterior distribution using Monte Carlo integration [69, 70]. If  $p_s(\vec{\theta})$  is a distribution which is never zero, then

$$\mathcal{L}_{\text{red}}(\vec{\lambda}) = \int \frac{\mathcal{L}(\vec{\lambda}, \vec{\theta}) p(\vec{\theta})}{p_s(\vec{\theta})} [p_s(\vec{\theta}) d\vec{\theta}] . \quad (28)$$

If we draw  $N$  random samples  $\vec{\theta}_q$  from  $p_s$ , we can estimate the value of the integral  $\hat{\mathcal{L}}_{\text{red}}$  and its error  $\sigma_{\mathcal{L}_{\text{red}}}$  as follows. Let

$$w_q \equiv \frac{\mathcal{L}(\vec{\lambda}, \vec{\theta}_q) p(\vec{\theta}_q)}{p_s(\vec{\theta}_q)} , \quad (29)$$

for a given draw  $q$  of  $\vec{\theta}$ , then

$$\hat{\mathcal{L}}_{\text{red}}(\vec{\lambda}) \equiv \frac{1}{N} \sum_q w_q = \langle w \rangle , \quad (30)$$

$$\sigma_{\mathcal{L}_{\text{red}}}^2 = \langle w^2 \rangle - \langle w \rangle^2 . \quad (31)$$

The weighted samples also provide an estimate of the marginalized one-dimensional cumulative distributions  $\hat{P}(< x)$  at fixed  $\vec{\lambda}$ :

$$\hat{P}(< x) \equiv \frac{1}{\sum_q w_q} \sum_q w_q \Theta(x - x_q) , \quad (32)$$

where  $x$  is one of the extrinsic variables in  $\vec{\theta}$ , and  $\Theta$  is the Heaviside step function. This formula follows by performing Monte Carlo integration on the parameter volume  $< x$ , keeping track of overall normalization. In the limit of many samples, this discontinuous estimate should converge to a smooth, marginalized posterior distribution. For any set of samples and any  $x$ , the error in  $\hat{P}$  follows from the (correlated) statistics of the Monte Carlo integrals in its numerator and denominator. In the typical case that all samples  $x_q$  are distinct, the unique sample with the largest weight corresponds to the largest discontinuity in  $\hat{P}$ . The magnitude of this discontinuity, or equivalently its inverse  $n_{\text{eff}}$ , provides a practical measure of how reliable we expect this one-dimensional posterior to be:

$$n_{\text{eff}} \equiv \frac{\sum_q w_q}{\max\{w_q\}} . \quad (33)$$

Equivalently, the “effective number of samples”  $n_{\text{eff}}$  measures how many independent samples produce similar weights near the largest observed weight.

Following the discussion of priors earlier in this section, unless otherwise indicated, we draw samples using

a separable sampling distribution  $p_s(\vec{\theta}) = \prod_{\alpha} p_{s,\alpha}(\theta_{\alpha})$ . Each factor  $p_{s,\alpha}$  is equal to the corresponding prior in that dimension.

Furthermore, being a pure Monte Carlo integral, all the draws are independent. We can simply average the results of multiple instances to achieve a smaller  $\sigma_{\mathcal{L}_{\text{red}}}$ , even if these evaluations adopted to a different sampling distribution. Currently, we perform  $n_{\text{trials}} (= 10)$  evaluations of the integral at each fixed intrinsic point, terminating when either  $N$  iterations crosses a threshold ( $= 10^6$ ) or to some fixed  $n_{\text{eff}}$  threshold ( $= 1000$ ), whichever comes first. This approach is thus highly parallelizable: simply instantiate instances (with different seeds) of the desired integral across however many computing resources are required for the target execution time and precision, effectively dividing the time required to converge by the number of processes available. Combined with gridding the intrinsic parameters, this allows for a degree of parallelization which has not been achieved by other sampling methods in this parameter space. There is a practical limit to this strategy since there is a fixed start up time, but the degree of parallelization realized is at least an order of magnitude or more over current schemes.

### 1. Adaptive Monte Carlo integration

To better concentrate our samples in regions of high significance, we implemented an importance sampling algorithm using a simple adaptive Monte Carlo procedure. Every  $n_{\text{adapt}}$  samples, we updated the sampling distribution based on measured weights  $w_k^{\beta}$ . We choose to temper the distribution by raising the likelihood to an exponent which is chosen heuristically,

$$\beta = \min \left( 0.8, 4 \frac{\ln(n_{\text{adapt}})}{\rho_{\text{net}}^2} \right) . \quad (34)$$

where  $n_{\text{adapt}}$  is the number of samples in the histogram and  $\rho_{\text{net}}$  is the network SNR. This choice attempts to mitigate the large dynamic range of  $\mathcal{L}$  in Eq. 30 to a scale comparable to the number of samples used in each histogram. At the end of every adaptation period, we reconstruct the one-dimensional sampling distributions in each adapting dimension, using the last  $n_{\text{adapt}}$  samples. In each dimension ( $\theta$ ), we subdivide the full range into  $n_{\text{bins}}$  equal-length bins then evaluate a tempered, weighted histogram. Since the dimensionality of the space is high, but sparse, there is a danger that early adaptations will have one-dimensional sampling distributions which suffer from fluctuations in other parameters after marginalization. To avoid this effect and better ensure the full prior space is covered adequately, we then average the histogram  $W$  with a uniform distribution with weight  $s$ :

$$\hat{W}_{\alpha} = s W_{\alpha} + (1 - s) / n_{\text{bins}} . \quad (35)$$

Finally, we transformed from this discrete, bin-by-bin representation to a continuous integral by constructing a one-dimensional sampling distribution  $p_{s,\alpha}(\theta_\alpha)$  by interpolating between bin centers, then constructing the one-dimensional inverse  $P_{s,\alpha}(< \theta_\alpha)^{-1}$  by integrating  $P'_{s,\alpha}(\theta_\alpha) = p_{s,\alpha}(\theta_\alpha)$ . The latter process ensures that the sampling distribution and inverse cumulative distribution used to generate random samples are normalized and self-consistent.

As configured for this paper, we used  $n_{\text{bins}} = 100$ ,  $n_{\text{adapt}} = 1000$ , and  $s = 0.1$ . The adapted sampling distributions are frozen in after  $10^5$  samples are drawn.

The adaptive sampler was not used for all parameters. The parameters that are useful to adapt often have a significant dynamic range (i.e., distance) or could be represented with more detailed prior information (i.e. sky location). Conversely, parameters whose marginalized posterior is expected to resemble the prior, like orbital phase and polarization, were not adapted as our one-dimensional method would provide no benefit.

## 2. Using search results to target specific areas of the sky

Gravitational-wave search pipelines have usually already identified candidate event times in two or more interferometers' data, typically to much less than the light crossing time of the earth.

However, the initial search pipeline is not designed to provide estimates of the sky location, so the BAYESTAR pipeline [71] rapidly processes the results of a gravitational-wave search to identify candidate sky locations consistent with a gravitational-wave event, using only the information provided by the top-level search. This code produces a posterior distribution  $p_{BS}(\alpha, \delta)$  for the gravitational-wave signal sky location in a discretized, interlocking, equal-area grid of pixels corresponding to sky regions, with probabilities for each pixel (a HEALPix<sup>4</sup> skymap). The refinement of the grid is variable and scales with the resolution required to resolve the features of the posterior in  $\alpha$  and  $\delta$ .

While we could allow the adaptive sampling to reduce the sky area required to be searched, we can also use the two-dimensional  $(\alpha, \delta)$  posterior from BAYESTAR as a sampling distribution to immediately target our Monte Carlo at a region of high support. This is more effective than the adaptive sampler because the one-dimensional sampling distributions, even if well converged to the true marginalized one-dimensional posteriors, over cover the two-dimensional space, thus limiting their effectiveness in sampling only where support truly exists.

For simplicity, when using a BAYESTAR skymap, we adopt a purely discrete sky: the samples are selected from the set of sky region centers as calculated by the healpix

library. In practice, the pixels are at a sufficient resolution such that detrimental effects from the discretization are rarely noticeable.

## 3. Time marginalization

Having already computed the  $Q_{k,lm}(\vec{\lambda}, t_k)$ , we can evaluate the likelihood versus time  $\mathcal{L}(t)$  cheaply, by array addition operations. Hence, rather than performing Monte Carlo integration, we can likewise efficiently *integrate* over time.

Specifically, for every sky location drawn by the adaptive Monte Carlo routine, we compute the corresponding time shift between geocenter and each detector with Eq. (5). To evaluate the value of  $\ln \mathcal{L}$  for some geocenter time  $t_{\text{geo}}$ , we simply look up the corresponding value of the  $Q_{k,lm}(\vec{\lambda}, t_k)$  from our precomputed values and plug them into Eq. (24). The resulting time series  $\mathcal{L}(t)$  is then numerically integrated using Simpson's rule in a window of 300 ms, centered on  $t_{\text{geo}}$ . The integrated time series is required to be at a high enough sampling rate to accurately capture the peak feature. We integrate it at the same sample rate (16 kHz) of the native  $h(t)$  data.

## E. Intrinsic priors and posterior construction

At this point, we have evaluated  $\mathcal{L}_{\text{red}}(\vec{\lambda})$  over a structured grid of intrinsic parameters (here indexed by  $r$ )  $\lambda_r$ . By construction, the values  $\mathcal{L}_{\text{red}}(\vec{\lambda})$  have small statistical errors (e.g., typically less than a few percent). We can then interpolate  $\mathcal{L}_{\text{red}}(\vec{\lambda})$  throughout the sampled grid. Combined with the prior  $p(\vec{\lambda})$  over intrinsic parameters, we evaluate the overall evidence  $Z$  and posterior distribution  $p(\vec{\lambda})$  over intrinsic parameters via

$$Z = \int p(\vec{\lambda}) \mathcal{L}_{\text{red}}(\vec{\lambda}) d\vec{\lambda}, \quad (36)$$

$$p(\vec{\lambda}) = \frac{1}{Z} p(\vec{\lambda}) \mathcal{L}_{\text{red}}(\vec{\lambda}). \quad (37)$$

To construct posteriors for the intrinsic parameters  $\mathcal{M}_c, \eta$ , we adopt a uniform prior in  $m_1, m_2$  with  $m_1, m_2 \in [1, 2]M_\odot$ : inside the specified region, the (uniform) prior density is  $p(m_1, m_2) dm_1 dm_2 = dm_1 dm_2 / (M_\odot)^2$ . Changing coordinates using the Jacobian  $d(m_1, m_2) / d(\mathcal{M}_c, \eta) = \delta \mathcal{M}_c / M^2 = \delta \eta^{6/5} \mathcal{M}_c^{-1}$  where  $\delta = (m_1 - m_2) / M = \sqrt{1 - 4\eta}$ , we find the prior density in  $\mathcal{M}_c, \eta$  coordinates is

$$p(\mathcal{M}_c, \eta) d\mathcal{M}_c d\eta = \frac{1}{M_\odot^2} \frac{\mathcal{M}_c d\mathcal{M}_c d\eta}{\eta^{6/5} \sqrt{1 - 4\eta}}. \quad (38)$$

In other words, due to a coordinate singularity, the prior in  $\eta$  diverges near the equal-mass line. This coordinate

<sup>4</sup> <http://healpix.sourceforge.net/>

singularity has a disproportionate impact on comparable-mass binary posteriors.

However, unlike MCMC and nested-sampling codes, we can reweight our results, allowing the user to reprocess the posterior using any user-specified intrinsic-parameter prior. As a result, our approach also enables several other calculations of practical interest: reanalysis given alternative astrophysical priors, simultaneous analysis of multiple events, adapting the “prior” intrinsic (mass, tide [72, 73], cosmological parameter [74–76]) distribution to reproduce multiple observations, and simultaneous independent constraints from multimessenger observations.

To a good approximation, the intrinsic and extrinsic parameter distributions often separate after marginalizing in time. In other words, after marginalizing in time, the extrinsic parameter distributions are nearly independent of  $\vec{\lambda}$ . In that case, each individual extrinsic parameter distribution provides a reliable estimate of the posterior. As a crude approximation to the one-dimensional cumulative posterior distribution  $P(\theta)$ , we simply average all weighted samples  $(x_{r,q}, w_{r,q})$  from all mass points:

$$\hat{P}(< x) = \frac{\sum_{r,q} w_{r,q} \Theta(x - x_{r,q})}{\sum_{r,q} w_{r,q}}, \quad (39)$$

where  $\Theta$  is the Heaviside step function.

## V. RESULTS: RUNTIME DETAILS AND LARGE SCALE TESTS

We present a proof-of-principle study of our parameter estimation pipeline, employed in an environment which should resemble the steps taken after the identification of search triggers. To this purpose, a mock data challenge was constructed to exercise both the low latency gravitational wave searches as well as the parameter estimation follow-up processes expected to be applied to gravitational-wave candidates identified by the search [33]. This challenge proceeded in several stages, each designed to emulate the anticipated workflow in a low latency binary neutron-star (BNS) detection environment.

From the reported search pipeline results, we use the  $\mathcal{M}_c$  and  $\eta$  coordinates from which the intrinsic search space is extrapolated, and the window for time marginalization (see Sec. IV D 3) is formed around the reported coalescence time  $t_r$ . The search pipeline also provides reference power spectral densities  $S(f)$  utilized in evaluating the likelihood.

Unless otherwise stated, the likelihoods were evaluated using nonspinning TaylorT4 templates at 3.5 post-Newtonian (PN) order, including only the  $(2, \pm 2)$  and  $(2, 0)$  modes. As we will discuss at length below, this signal model does *not* include all degrees of freedom permitted to the binary in the data. When evaluating the likelihood, waveforms started at  $f_{\text{low}} = 40\text{Hz}$ .

The inner product integration Eq. (13) uses an inverse power spectrum filter targeting the frequency range  $[f_{\text{min}}, f_{\text{max}}] = [f_{\text{min}}, 2000\text{Hz}]$ , constructed from the measured power spectrum.

Only distance used an adaptive sampling function, starting initially with a constant function. Our distance prior was uniform in volume out to  $D = 300\text{Mpc}$ . The declination and right ascension were sampled from the posterior provided by BAYESTAR. As used here, the skymap had  $12 \times 64^2$  pixels, roughly one per square degree.

### A. Ensemble of events

We first present the agglomerated results from 450 events drawn randomly from a set of recovered events in [33, 77]. These events were also processed using BAYESTAR and a subset were processed by the LALINFERENCE Monte Carlo sampler codes. All events have false alarm rate — as determined by the GSTLAL search pipeline — smaller than one per century. This threshold is motivated by the selection criteria outlined in [26]. As in [33], this data set adopts a noise PSD set at the median sensitivity from fiducial predictions for 2015 LIGO observing [26, 78]. It is expected that the Virgo interferometer will not be functional during this time period, leaving a two detector LIGO site network.

A population of gravitational-wave signals from BNSs was added into the data, as described in [33] and reviewed here. Events in this set (the 2015 MDC data) were distributed isotropically on the sky, and uniformly in volume out to 219 Mpc. The BNS injections had uniform random component masses in  $1.2M_\odot - 1.6M_\odot$  and randomly oriented spins with the dimensionless magnitude not exceeding 0.05. These (precessing) binary signals were generated via precessing SpinTaylorT4 templates at 3.5 PN order [6], including all known post-Newtonian modes [13].

First, the GSTLAL BNS search was performed over the MDC set, identifying events for further follow up. For each event identified by the gravitational-wave search pipeline, we distributed the intrinsic points according to the procedure described in Sec. IV C and evaluated the Monte Carlo integral 10 times for each mass point, combining each of the runs together after. We outline the form of the priors and sampling functions for each parameter in Table I.

In addition to the information provided by the search pipeline, we also make use of the posterior probability of  $\alpha$  and  $\delta$ , provided by BAYESTAR. It is expected that such information will be available within a few minutes of trigger identification. The gains in time to convergence in the Monte Carlo integral are expected to outweigh the loss of time waiting for BAYESTAR to finish processing if the BAYESTAR posterior is used as a sampling function for the sky location.

In order to validate the results of the pipeline, we present a comparison of the ensemble distribution of re-

parameter	physical prior	initial sampling distribution
distance ( $D$ )	uniform in volume	uniform in distance
inclination ( $\iota$ )	uniform in $\cos \iota$ , $\iota$ between $(-\pi, \pi)$	same
polarization ( $\psi$ )	uniform in $(0, 2\pi)$	same
orbital phase ( $\phi$ )	uniform in $(0, 2\pi)$	same
declination ( $\delta$ )	uniform in $\cos \delta$ , $\delta$ between $(-\pi, \pi)$	drawn from BAYESTAR skymap
right ascension ( $\alpha$ )	uniform in $(0, 2\pi)$	drawn from BAYESTAR skymap

TABLE I: Prior distributions used for the extrinsic parameters in the 2015 MDC study.

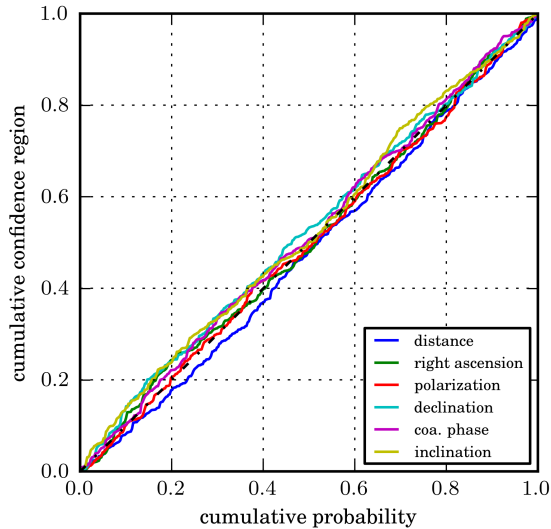


FIG. 2: For 450 randomly-selected NS-NS binaries, a plot of the cumulative distribution of  $P_\theta(\theta_k)$  for each extrinsic variable  $\theta = D, \alpha, \delta, \iota, \psi, \phi$ . The curves represent the normalized fraction of events with estimated cumulative probability less than the injected parameter value. The grey shaded regions are the  $1\sigma$  uncertainty regions.

covered parameters versus their intrinsic distribution as determined by the prior — a so called PP plot — in Fig. 2. If our estimates for the one-dimensional cumulative distributions  $P(< x)$  are unbiased and if  $x_*$  is a random variable consistent with the prior, then  $P(x_*)$  should be a uniformly-distributed random variable. To test this hypothesis, we use the one-dimensional posteriors provided by the MDC.

For each parameter  $x$ , each colored curve in Fig. 2 is the fraction of events with estimated cumulative probability  $P(< x_*)$  at the injected parameter value  $x_*$ . Specifically, if  $P(x_{*q})$  are the sorted cumulative probabilities for the  $q = 1 \dots n$  events with  $P(x_{*1}) < P(x_{*2})$ , then the points on the plot are  $\{P(x_{*q}), q/n\}$ .

There are a few possible sources of bias in the recovery of the posterior distribution. While the spins are nom-

inally mild, it is possible that the use of non-spinning templates in the signal analysis for both the GSTLAL and our parameter estimation codes could cause biases. Additionally, the event selection process outlined in [33] introduces a small selection (Malmquist) bias which slightly disfavors edge-on binaries relative to our uniform prior. Our parameter estimation strategy does not account for selection biases; for a sufficiently large ensemble of events, small deviations between the statistical properties of our posteriors and the ensemble are expected.

## B. Marginalization details

Two events were selected for demonstration. One event represents a marginal detection by the GSTLAL pipeline. It is suggestive of what might be seen in the 2015-2016 era, and is not a confident detection by itself. The second event is a “gold-plated” event, one which is an assured single-event detection by the GSTLAL pipeline. They are qualitatively comparable in most regards, however, the stronger event (#21091), being more strongly peaked in likelihood did not accumulate as many effective samples, and hence the error on the reduced likelihood ( $\mathcal{L}_{\text{red}}(\lambda)$ ) tends to be larger by an order of magnitude, yet still small in the relative sense. We provide the simulation and GSTLAL event IDs so that the interested reader can cross-correlate results here with those obtained by other samplers in [33]. We will refer to them by their GSTLAL IDs #21091 and #14631. Their properties are enumerated along with the values obtained by the GSTLAL pipeline and the recovered parameters in Table V B.

Each event was gridded in the elliptical region of 10% mismatch in the same manner described in Sec. IV C. After unphysical points were thrown out, there were 128 points for event #14631 and 124 points for event #21091. Each point had ten independent instances of the integrator applied, both to accumulate samples in parallel fashion as well as to cross-validate the results. The approximate wall clock time for any given instance was about 45 minutes. With approximately 100 intrinsic points and 10 copies of the integrator instance for a given event, the total time of order 750 CPU hours.

The evolution of the Monte Carlo integral evaluation is shown in Figs. 3(a) (#21091) and 3(b) (#14631). This is for a single mass point corresponding to the center of

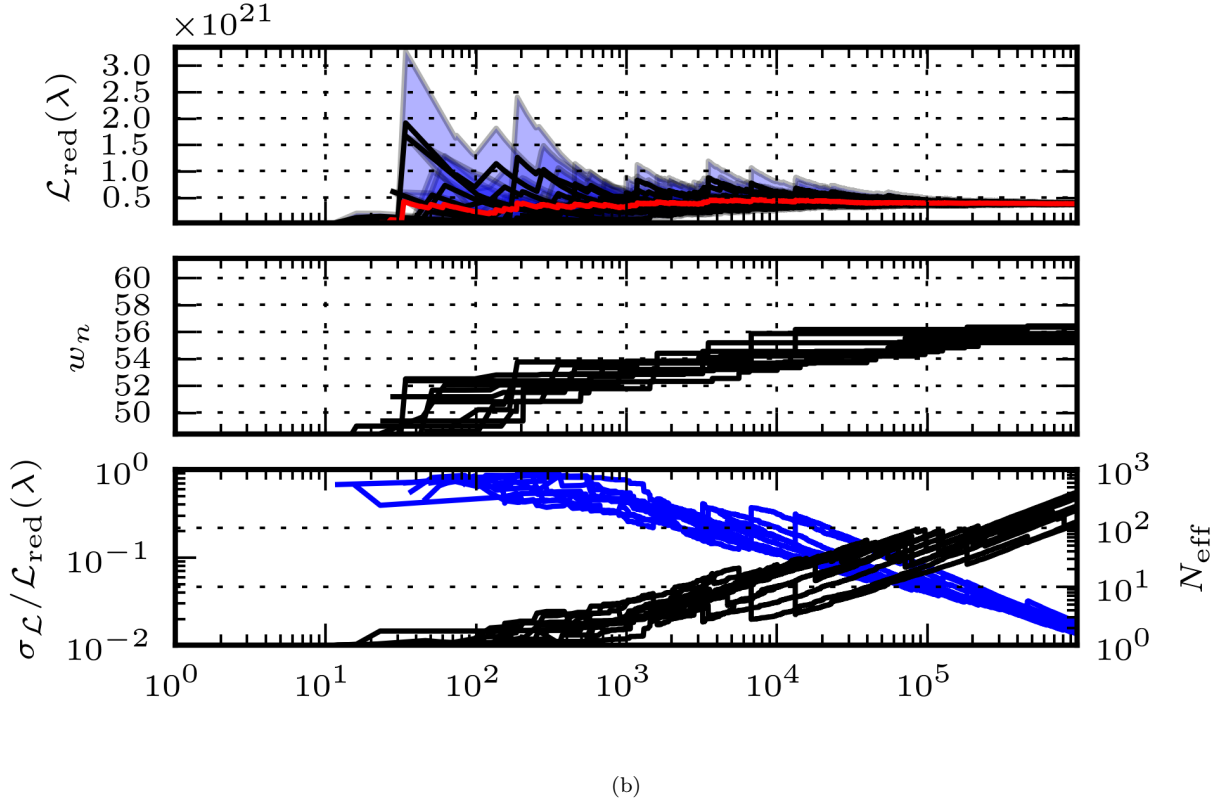
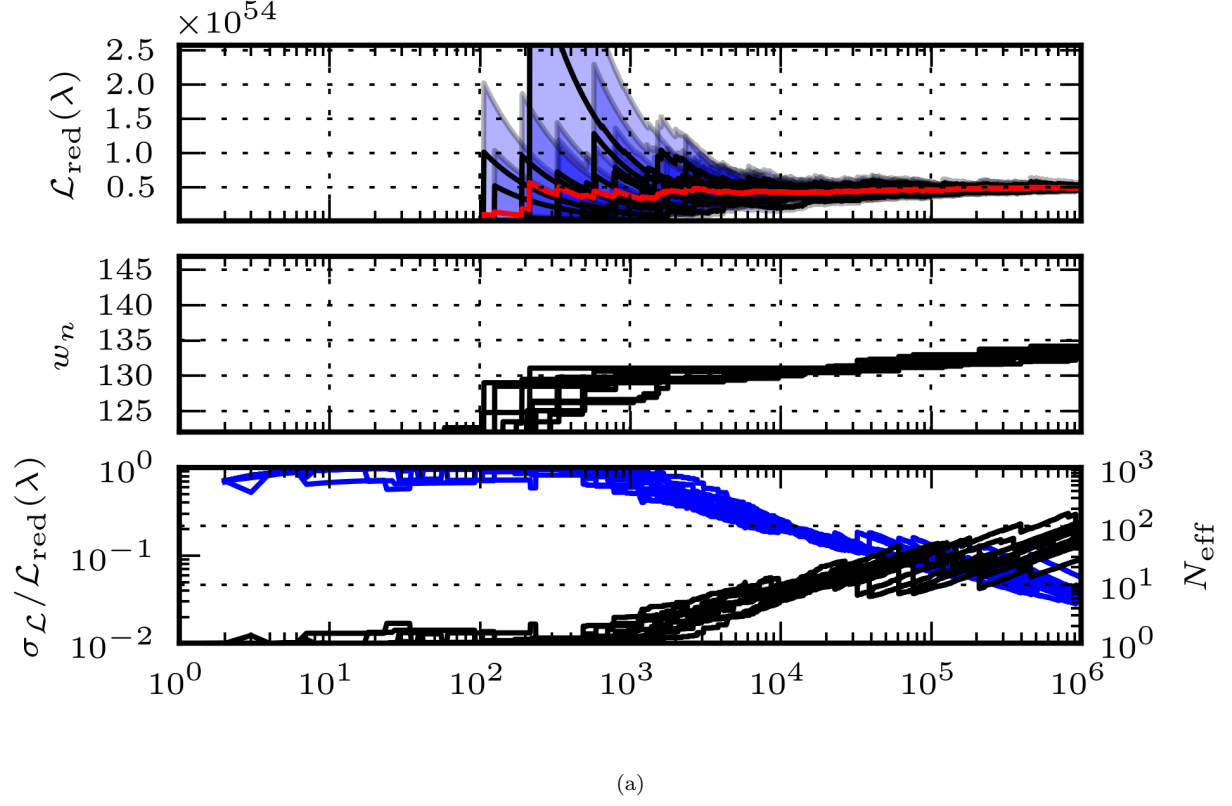


FIG. 3: Figs. (a) and (b) are correspond to events #21091 and #14631, respectively. The top panel in each figure shows the estimated value (black lines) of the integral, for each of the ten integrators, through a given sample count.  $1\text{-}\sigma$  error regions are shaded in blue around the curve. Additionally, the value of the integral if all ten integrators are combined is shown in red. The middle panel shows the maximum value found by the integrator instance throughout the iterations. The last panel plots the number of effective samples (Eq. (33)) in black, with corresponding ordinate axis the right. In the same panel, the relative error (Eq. (31)) is plotted in blue, with corresponding ordinate axis on the left.

GSTLAL Search ID #21091, Injection ID #26465				GSTLAL Search ID #14631, Injection ID #30639			
Parameter	Recovered (parameter estimation)			Parameter	Recovered (parameter estimation)		
	Injected (GSTLAL search)	Recovered	median / mode		Injected (GSTLAL search)	Recovered	median / mode
$m_1$ ( $M_\odot$ )	1.60	1.68	1.55 / 1.55	$m_1$ ( $M_\odot$ )	1.32	1.49	1.49 / 1.49
$m_2$ ( $M_\odot$ )	1.40	1.33	1.54 / 1.54	$m_2$ ( $M_\odot$ )	1.28	1.14	1.13 / 1.13
$D$ (Mpc)	68	-	59 / 67	$D$ (Mpc)	102	-	85 / 101
$\iota$	3.0	-	1.3 / 0.47	$\iota$	3.09	-	1.53 / 2.76
$\delta$	-0.12	-	-0.06 / -1.02	$\delta$	0.28	-	0.28 / 0.91
$\alpha$	4.00	-	3.9 / 0.82	$\alpha$	1.89	-	2.39 / 1.83
$\psi$	3.12	-	3.08 / 1.01	$\psi$	2.32	-	3.14 / 3.33
$\phi$	4.03	-	3.08 / 0.94	$\phi$	5.73	-	3.08 / 2.26

Other parameters				Other parameters			
Parameter	Recovered (parameter estimation)			Parameter	Recovered (parameter estimation)		
	Injected (GSTLAL search)	Recovered	Recovered		Injected (GSTLAL search)	Recovered	Recovered
$ \chi_1 $	0.026	-	-	$ \chi_1 $	0.032	-	-
$ \chi_2 $	0.008	-	-	$ \chi_2 $	0.036	-	-
$\rho$	16.28	17.24 <sup>a</sup>	16.85 <sup>b</sup>	$\rho$	11.06	12.04 <sup>a</sup>	11.07 <sup>b</sup>

<sup>a</sup>with corresponding false alarm rate  $3.8 \times 10^{-14}$

<sup>b</sup>as measured by  $2\sqrt{\ln(\max\{L_i\})}$

TABLE II: Injected and recovered intrinsic and extrinsic parameters for an injected signal. The GSTLAL search only reports time of arrival, signal-to-noise ratio, and mass information. The parameters from our algorithm are quoted at a weighted median and mode of the marginalized posteriors. Some other parameters not considered in this study (e.g. the dimensionless component spin  $\chi_{1,2}$ ) are listed separately.

the ellipsoid constructed for sampling the  $M_c, \eta$  plane, and so also the mass reported by the GSTLAL search. As explained in section IV C each mass point is independently evaluated 10 times. Each integral evaluation is shown in the top panel as a function of the number of samples drawn to that point. The blue regions surrounding each line are the error estimate at that point in the sampling. It is readily apparent that the integrators have sampled near the maximum point within a few thousand samples (middle panel, Figs. 3(a) and 3(b)), and at this point the integrators all begin to converge towards the same evidence value. Within  $10^5$  samples, most of the integrators have highly overlapping error regions, and only moderately small changes in the maximum likelihood point searched. The relative error for each integrator is plotted in blue in the bottom panel, with the typical  $1/\sqrt{N}$  behavior exhibited. The “jumps” in the curves correspond to a new maximum being found in the integrand. The final relative error for any of the integrators is typically of order 5% or less for all cases (with only one 10% relative error for #21091), and typically less than 1% for #14631. The red line in the top panel of the figures shows the evidence value if all ten integrators were joined as one integrator.

Figs. 4(a) (#21091) and 4(b) (#14631) display the results of each of the ten integrator instances scattered in the  $N_{\text{eff}}$ , evidence plane. The dashed line indicates the value of the evidence as calculated across all samples from all ten points. As can be noted, the error on the evidence values are small and most of the points are clustered near the total evidence dashed line, further indicating that all ten of the integrators have reached an internally consistent value, with very little spread around

the average. These evidences are then averaged together to obtain the reduced likelihood for that mass point. A contour plot of the evidence for this event is shown in Figs. 5(a) (#21091) and 5(b) (#14631), with the actual value of  $M_c$  and  $\eta$  of the injected signal marked by the crosshairs and the value obtained by the GSTLAL search pipeline is at the center of the ellipse. The evidence follows roughly the expected quadratic-shape of a Fisher-matrix manifold for these two parameters, but the maximum evidence point is slightly offset from the true value. We expect that this difference arises from the mild spin of the system biasing the result, as the waveform family used to generate templates and calculate  $\mathcal{L}_{\text{red}}$  does not include the effects of spin.

All possible sets of one and two-dimensional posteriors over the parameters are shown in Figs. 6(a) (#21091) and 6(b) (#14631). We also display here the posteriors for the intrinsic parameters ( $M_c$  and  $\eta$ ). The discreteness of the grid makes determination difficult for the component masses, however, the  $M_c$  and  $\eta$  posterior (using the prior in IV E) is consistent with the result shown in Figs. 5(a) and 5(b). In both cases, the distinct degeneracy between distance and inclination is clearly shown. In the case of #21091, a southerly inclination is weakly favored; this being an example of a bimodal distribution where the wrong mode is selected. The better sampled #14631 does choose the correct mode, but the degeneracy between the two is still quite strong, and the true inclination is nearly exactly face on. Also apparent in both cases is the favoring of moderately off-axis binary configurations which tend to bias the distance measurement towards systematically closer values. The  $p(\alpha, \delta)$  posterior, in effect the sky position, resembles the BAYESTAR

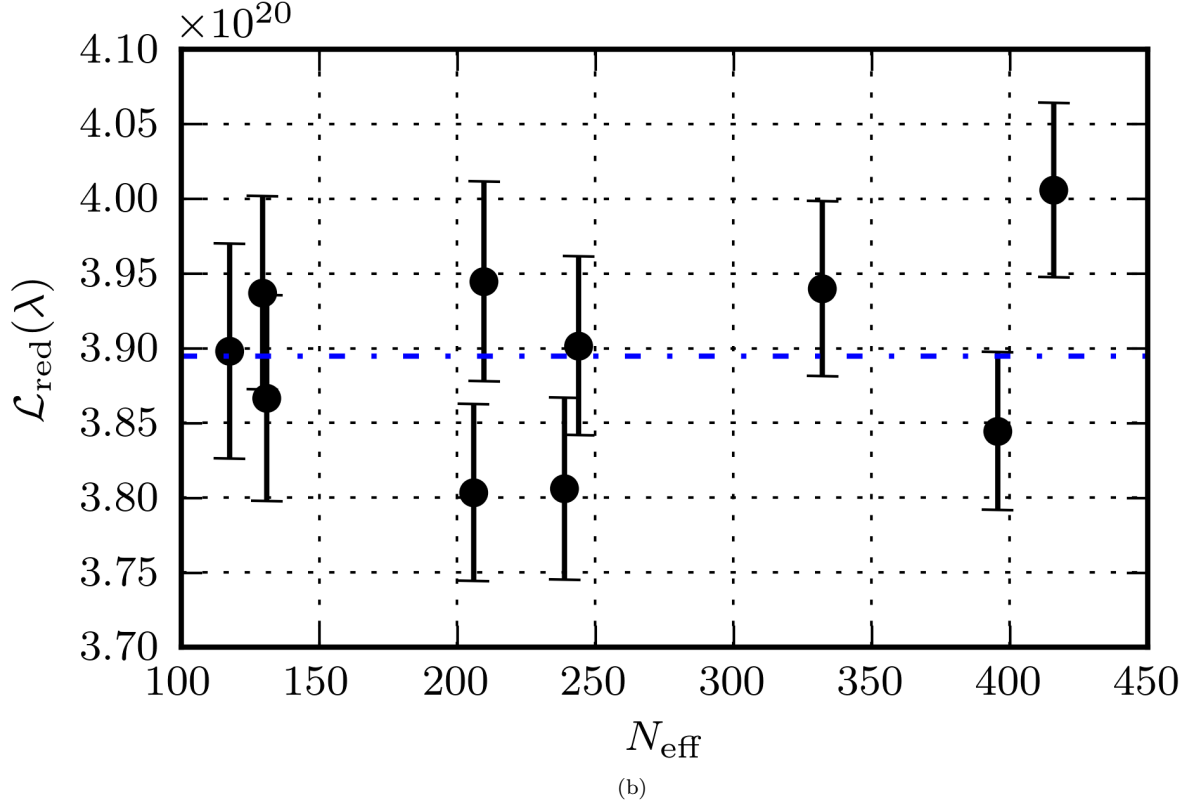
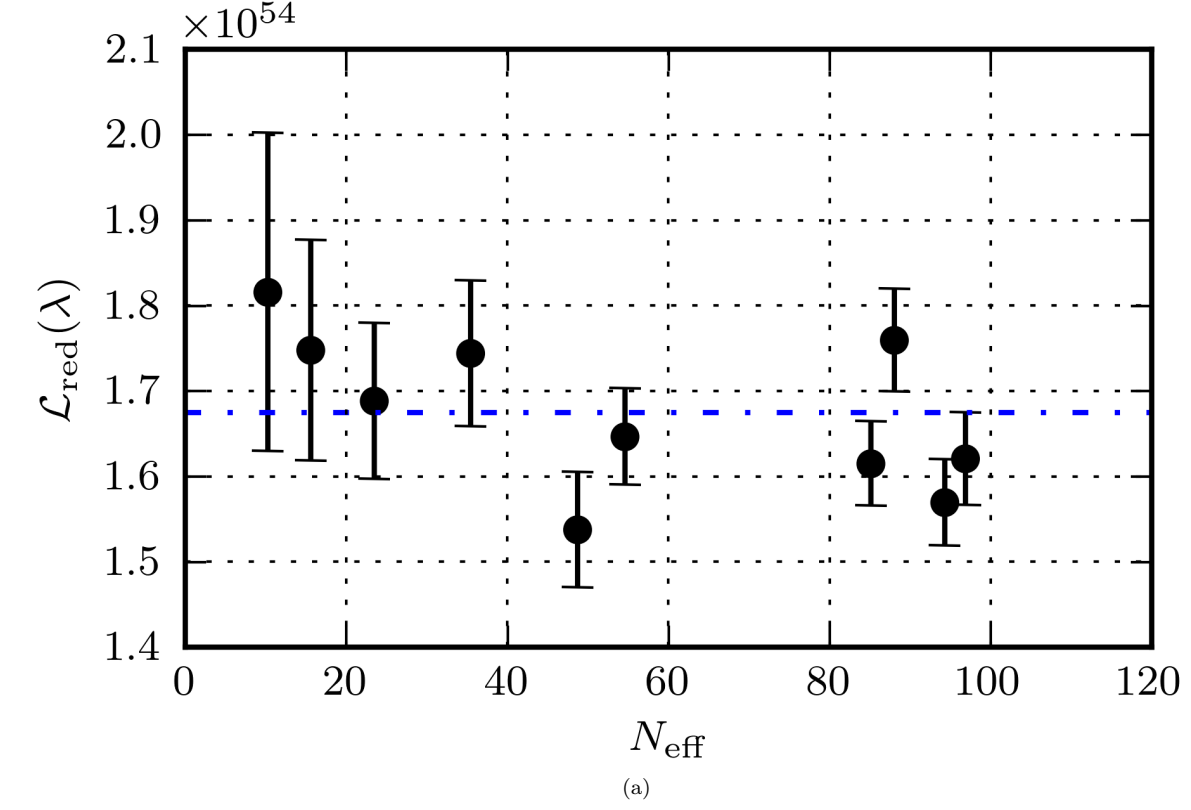


FIG. 4: Figs. (a) and (b) correspond to events #21091 and #14631, respectively. Each of the 10 integrator results is plotted against the reduced likelihood and number of effective samples collected ( $N_{\text{eff}}$ ). The error bars represent the uncertainty in the integral value from statistical error. The horizontal blue dashed line represents the value of the evidence averaged over all 10 integrator estimates.

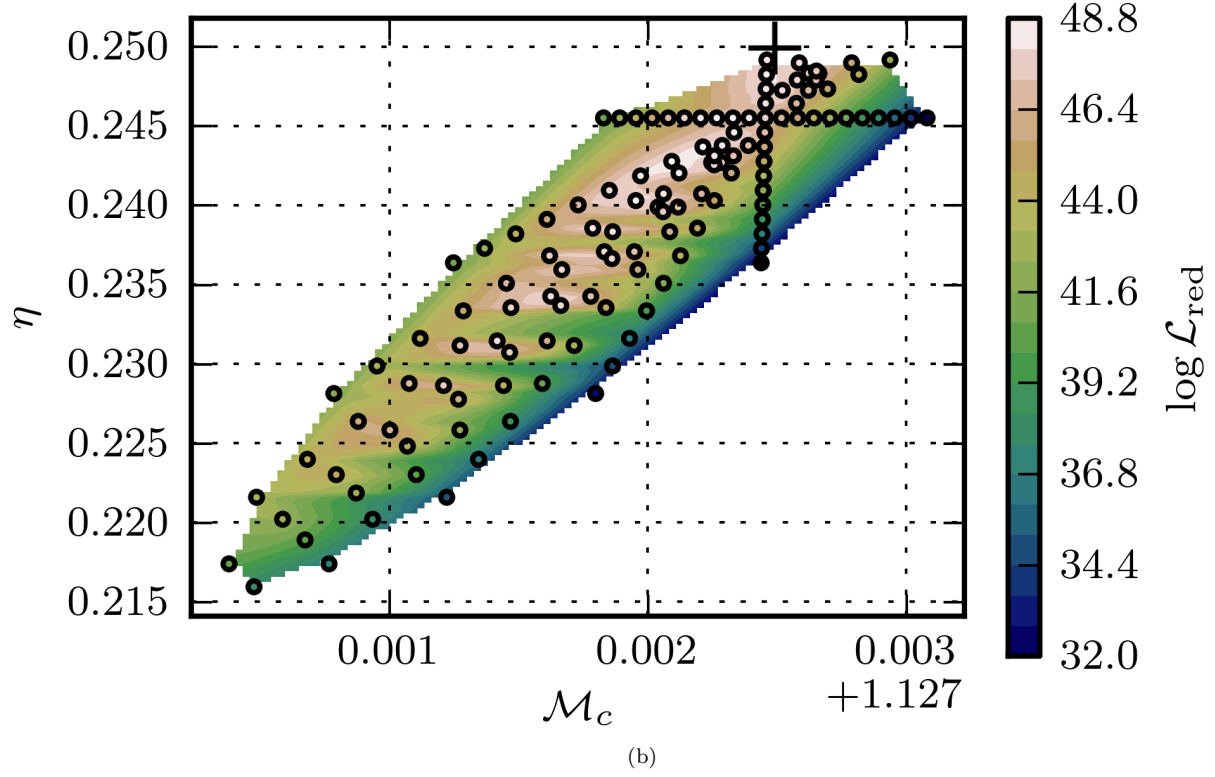
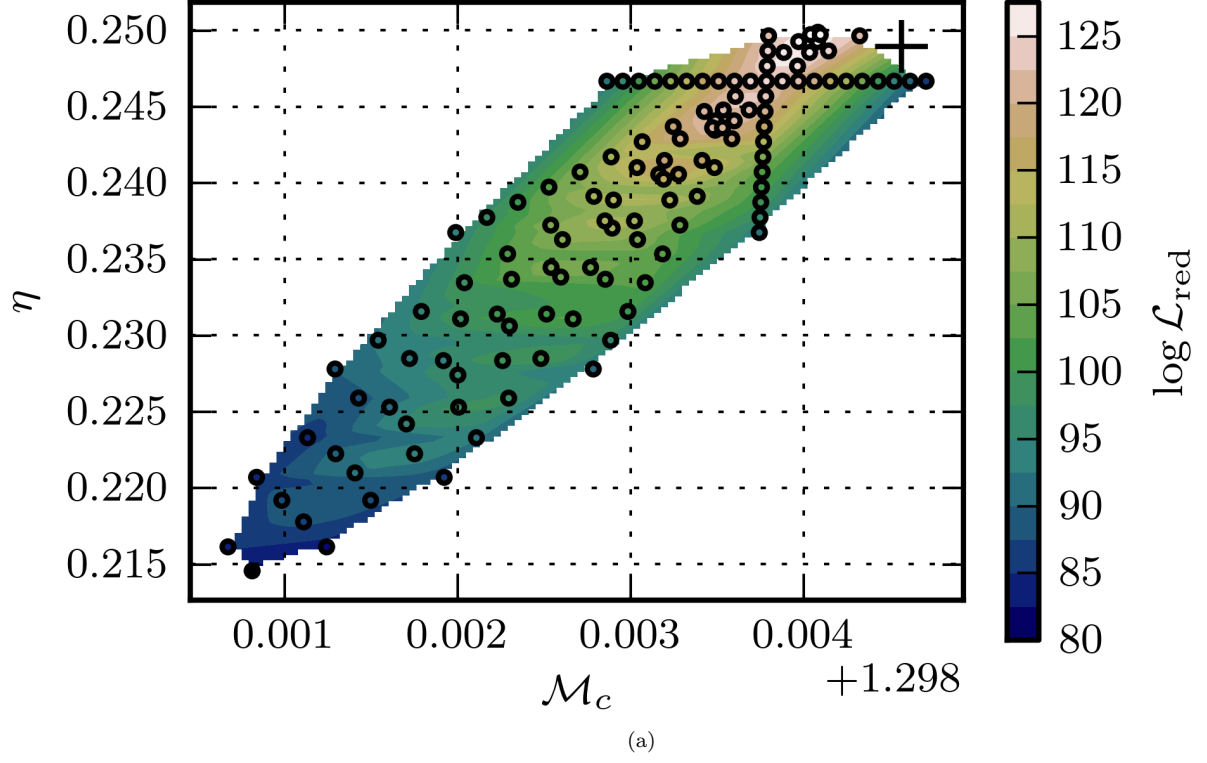


FIG. 5: Figs. (a) and (b) correspond to events #21091 and #14631, respectively. Level contours of the interpolated evidence surfaces in  $\mathcal{M}_c$  and  $\eta$  space. The points are the values of  $\mathcal{M}_c, \eta$ , which were used to do the integral over the extrinsic parameters, and all 10 integrator instances have been averaged to get the final value. The injected value is marked with a large cross.



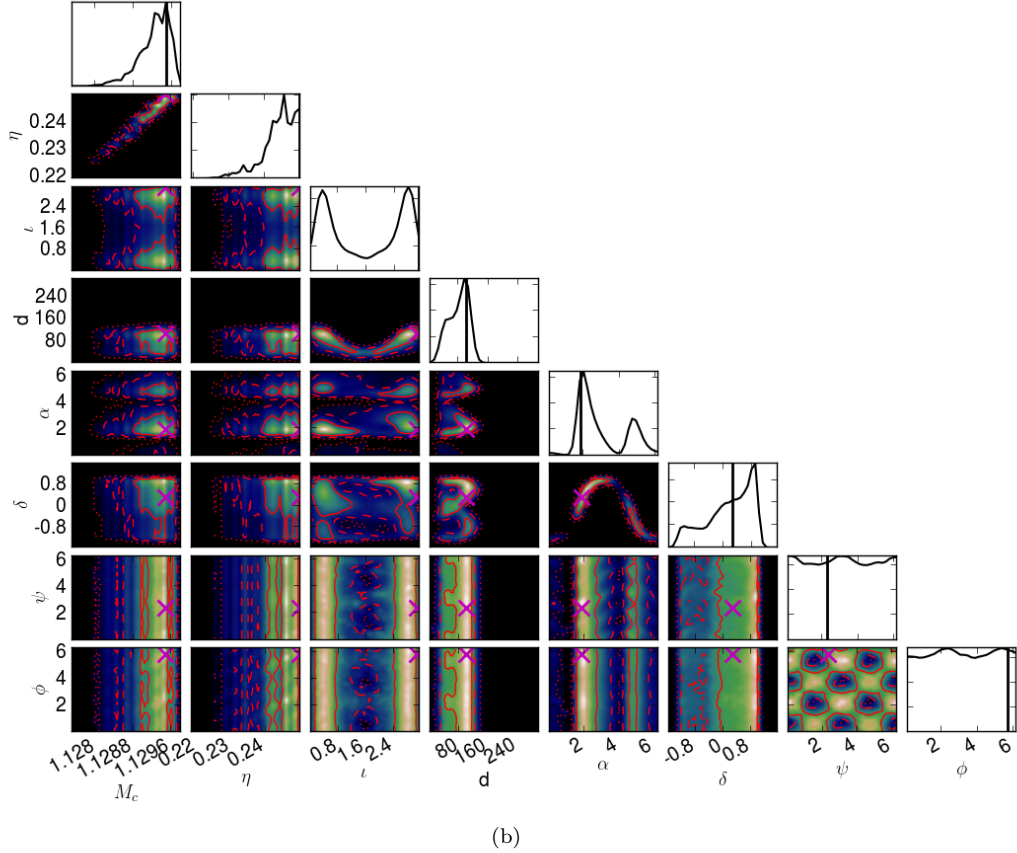
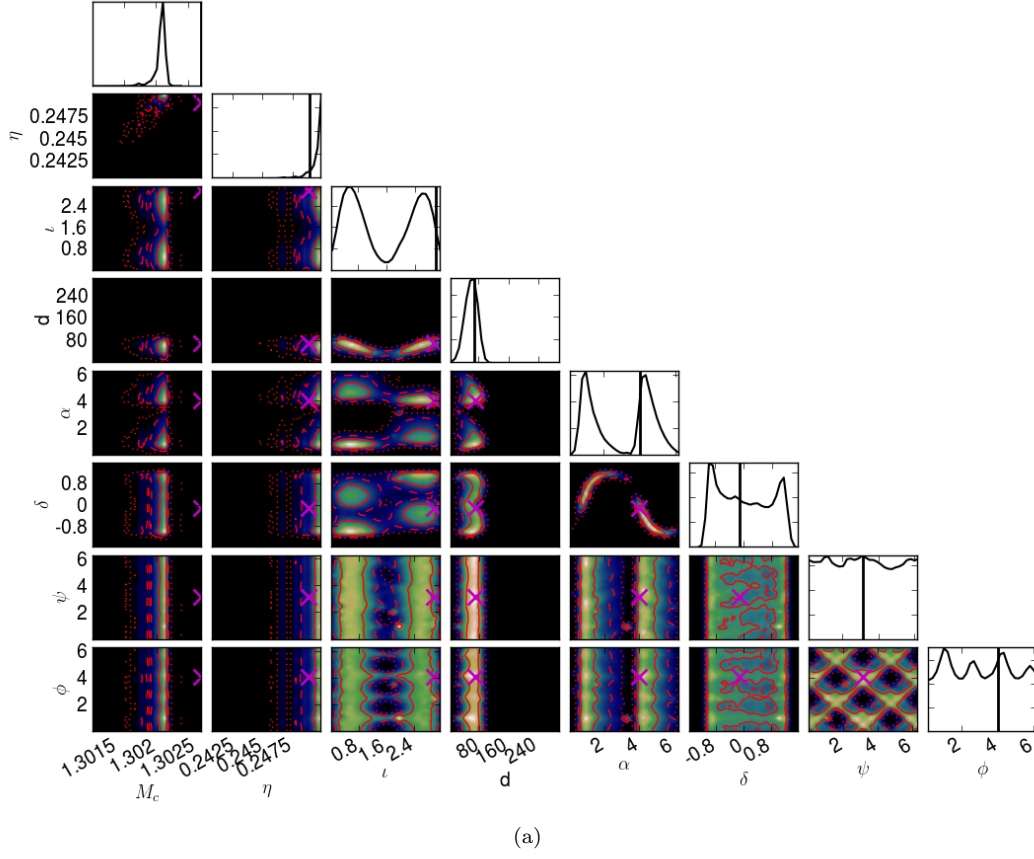


FIG. 6: Posteriors for event #21091, left and event #14631, right. All two-dimensional posteriors are represented as off-diagonal elements, and the diagonal elements are the one-dimensional fully marginalized posteriors. The true values of each parameter are marked with X's in the two-dimensional plots and vertical lines in the one-dimensional plots. From left to right, the parameters are  $M_c$ ,  $\eta$ ,  $l$ , distance, right ascension, declination, polarization, and coalescence phase. Note that the geocentric arrival time is omitted because it is marginalized numerically.

skymap which is used by our algorithm to select sample points for these parameters. For #21091, there are two modes along the triangulation ring, corresponding to a maxima and its mirror image. The sampler was not able to resolve the degeneracy between the points for #21091<sup>5</sup>, but the true location of the event lies near one of the maxima on the ring. In the case of #14631, one mode is suppressed relative to the other, and the true location lies very near the maximum. Also notable is the degeneracy between orbital phase ( $\phi$ ) and polarization angle ( $\psi$ ). The recovered posterior's strong degeneracy both qualitatively agrees with first principles and suggests further performance improvements (e.g., via direct phase marginalization).

### C. Scaling

For a quasicircular compact binary, it is well-known that the time to coalescence from a given gravitational-wave frequency scales as  $t(f) \propto f^{-8/3}$ . As the sensitivity of detectors improves at low frequencies, this requires the use of considerably longer waveforms for detection and parameter estimation. For example, the initial LIGO detectors were sensitive down to 40 Hz, while the advanced LIGO detectors could be sensitive down to 10 Hz. To cover the extra low frequency portion would require waveforms that are  $\approx 40$  times longer.

Traditional Bayesian parameter estimation is computationally limited by waveform generation and likelihood evaluations. Both of these are linearly proportional to waveform length. Note that the likelihood evaluations involve computing an inner product as in Eq. 13, which is approximated as a finite sum. The number of points in the sum is determined by the length of the waveform and data being analyzed, which is why the cost of likelihood evaluations scales with waveform length. Therefore, one would expect the cost of Bayesian parameter estimation using a seismic cutoff of  $f_{\min} = 10$  Hz to be roughly 40 times more expensive than the same analysis using  $f_{\min} = 40$  Hz.

The method proposed here is not computationally limited by waveform generation. Recall that for each point in the intrinsic parameter space we compute the waveform and the inner products between the various modes and the data (the assorted  $Q_{k,lm}$ ,  $U_{k,lm,l'm'}$ ,  $V_{k,lm,l'm'}$ ) only once. We then integrate over the extrinsic parameters, which involves evaluating  $F_+ + iF_\times$  and the  $Y^{(-2)}_{lm}$ 's for different values of the extrinsic parameters. While generating the waveform and computing inner products does scale with waveform duration, this cost is insignificant (even for  $f_{\min} = 15$  Hz) compared to the integration over extrinsic parameters, which is wholly independent of waveform duration. Therefore, as illus-

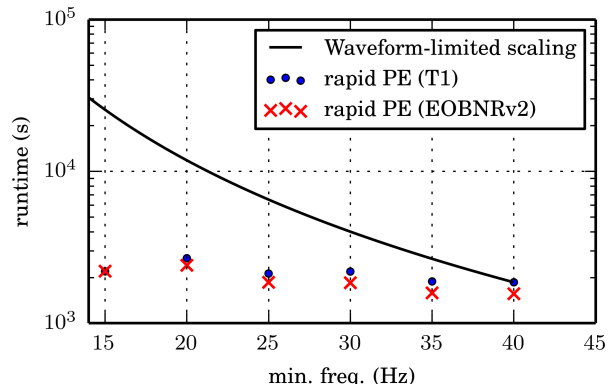


FIG. 7: Points show the runtime of our parameter estimation strategy as a function of the minimum frequency  $f_{\min}$ . For comparison, the solid curve shows the scaling  $\propto f_{\min}^{-8/3}$  expected if our runtime was proportional to the waveform duration (e.g., runtime proportional to the number of time samples). Waveforms were generated using the standard *TaylorT1* time-domain code, with  $m_1 = 1.55M_\odot$  and  $m_2 = 1.23M_\odot$ .

trated by Fig. 7. The cost of our method increases only a little as  $f_{\min}$  is decreased, in contrast to the sharp increase that occurs for waveform-limited techniques such as traditional Bayesian parameter estimation.

## VI. CONCLUSIONS

We have demonstrated a viable strategy for low-latency parameter estimation for long-duration binary neutron star signals, using an environment which resembles running conditions during the first advanced LIGO and Virgo observing runs.

In the era of multimessenger astronomy, rapid and robust inference about candidate compact binary gravitational-wave events will be a critical science product for LIGO, as colleagues with other instruments perform followup and coincident observations [26]. Low-latency sky localization [26, 38], already provides reasonable accuracy by approximate methods like BAYESTAR, so low-latency parameter estimation will further enable rapid electromagnetic followup and interpretation of candidate gravitational-wave events.

Motivated by the need for speed, we have introduced an alternative, highly-parallelizable architecture for compact binary parameter estimation. First, by using a mode decomposition ( $h_{lm}$ ) to represent each physically distinct source and by prefiltering the data against those modes, we can efficiently evaluate the likelihood for generic source positions and orientations, independent of waveform length or generation time. Second, by integrating over all observer-dependent (extrinsic) parameters and by using a purely Monte Carlo integration strategy, we can efficiently parallelize our calculation over the

<sup>5</sup> Though neither did the samplers in [33].

intrinsic and extrinsic space. Third, to target specific intrinsic (and extrinsic) parameters for further investigation, we ingest information provided by the searches and BAYESTAR: the trigger masses and estimated sky position. Using standard time-domain waveforms in a production environment, we can already fully process one event in less than 1 hour, using roughly 1000 cores in parallel, producing posteriors and evidence with reproducibly small statistical errors. By dramatically decreasing the turnaround time for each analysis and by scaling to harness all available resources efficiently, our strategy may significantly increase size and scope of parameter estimation investigations.

Our implementation has bounded runtime – one hour is the *worst* case – so we know what resources will be needed to analyze a given NS-NS binary. Moreover, the parallel algorithm can exploit all available computing resources, without need for communication or coordination between jobs, allowing it to operate in computing environments with tightly constrained wallclock time.

Our algorithm can also be immediately applied to any noise curve and existing time-domain model that provides  $h_{lm}$ , at any mass, including EOBNRv2HM [62] and SEOB [61]. Finally, by construction our dominant operation count cost is independent of the waveform’s length (or number of basis vectors). Hence, unlike reduced-order methods, our code will run in nearly same amount of time now and with full aLIGO-scale instruments with  $f_{\text{low}} \simeq 10\text{Hz}$ .

### A. Comparison with reduced-order quadrature

Reduced-order quadrature methods provide an efficient representation of a waveform family and any inner products against it. Other authors have recently proposed prefiltering the data against the reduced-order basis [60], achieving significant speedup. For example, using TaylorF2 templates, [60] claim runtimes of order 1 hour, comparable to our end-to-end time in the high-precision configuration described above.

Our strategy and reduced-order modeling achieve a similar speedup for qualitatively similar reasons: both strategies prefilter the data. In our algorithm, at each mass point, the data is prefiltered against a set of  $h_{lm}$ , then efficiently reconstruct the likelihood for generic source orientations and distances. By integrating the likelihood at each mass point over all extrinsic parameters, we are dominated by extrinsic-parameter sampling and hence not limited by waveform generation.

### B. Future work

We first recognize that our sampling strategy, especially the adaptation steps, have some deficiencies, especially in light of earlier literature [69, 70]. We intend to fortify our current strategy by expanding it be-

yond the independent product of one-dimensional sampling distributions. High-dimensional sampling distributions can be approximated by kernel density estimators. This will allow us to better capture  $n$ -dimensional correlations with no appreciable loss of computational speed. This approach will also better allow us to incorporate more detailed information from other stochastic samplers to use as sampling distributions. Improvements to the BAYESTAR algorithm will also likely allow its direct inclusion as a sampler, instead of just using the data products. Furthermore, the type of parallelization employed here is highly exploitable by GPU accelerated architectures.

While the alternative architecture proposed here is efficient and highly parallelizable over extrinsic parameters, all other parameters are currently suboptimally explored. For example, the algorithm described and implemented here adopts a fixed, low-resolution grid to sample two mass dimensions. Recent work [79] suggests that the simple grid placement scheme employed here could be dramatically improved upon. All the compact binary specific searches expected to run in the next two years are also capable of providing the specific portion(s) of the template bank that was triggered. We will soon be in a position to develop a more tightly hierarchical approach, and construct our mass grid to resemble those portions. While the method described here should generalize to a few additional dimensions, it is not yet clear what additional computational resources or architectural changes would be needed to apply our technique to all the intrinsic dimensions (e.g. to include tides and component spins). These higher-dimensional problems are being addressed with Markov Chain or Nested Sampling codes, including tests of GR, models which include non-astronomical environmental effects [80], and self-consistent electromagnetic and gravitational-wave parameter estimation. In the short-term future, we plan to explore straightforward extension of the intrinsic grid with additional tidal parameters. Also under exploration is using stochastic banks: these types of banks have been used successfully with aligned-spin but otherwise generic binary coalescence searches. That said, several methods have been proposed for rapid waveform interpolation, including SVD and reduced-order methods. In the long run, we anticipate being able to perform Monte Carlo integration over intrinsic dimensions as well, without being forced to adopt the relatively ad-hoc intrinsic/extrinsic split presented here.

To provide a complete proof-of-principle illustration of our algorithm, we developed an independent production-ready code. That said, the standard LALINFERENCE parameter estimation library in general and existing parameter estimation codes (LALINFERENCE\_MCMC and LALINFERENCE\_NEST) could implement some or all of the low-level and algorithmic changes we describe. For example, MCMC codes could implement our  $h_{lm}$ -based likelihood, then de-facto marginalize over all extrinsic parameters by strongly favoring jumps at fixed intrinsic parameters ( $\lambda$ ). Any implementation which provides accurate

marginalized probabilities (e.g.,  $\mathcal{L}_{\text{red}}$ ) can be parallelized across parameter space. We hope that by combining paradigms, future parameter estimation strategies can reach extremely low latencies, ideally of order a few minutes, when advanced detectors reach design sensitivity.

### C. Acknowledgments

This work was supported by National Science Foundation awards PHY 1104371, and PHY 1307429. We

are grateful for computational resources provided by the Leonard E Parker Center for Gravitation, Cosmology and Astrophysics at University of Wisconsin-Milwaukee. We would like to thank Jolien Creighton for useful discussions regarding this paper. We would also like to thank Saeed Mirshekari for a detailed reading of this manuscript. Some of the results in the paper, particularly those derived from BAYESTAR skymaps make use of the HEALPix[81] library. This document has been assigned the internal report number LIGO P1500012.

- 
- [1] G. M. Harry and the LIGO Scientific Collaboration, *Classical and Quantum Gravity* **27**, 084006 (2010), URL <http://stacks.iop.org/0264-9381/27/i=8/a=084006>.
  - [2] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin, et al., *Classical and Quantum Gravity* **32**, 024001 (2015), URL <http://stacks.iop.org/0264-9381/32/i=2/a=024001>.
  - [3] A. Buonanno, Y. Chen, and M. Vallisneri, *Phys. Rev. D* **67**, 104025 (2003), gr-qc/0211087.
  - [4] A. Buonanno, Y. Chen, Y. Pan, and M. Vallisneri, *Phys. Rev. D* **70**, 104003 (2004), gr-qc/0405090.
  - [5] T. Damour, A. Gopakumar, and B. R. Iyer, *Phys. Rev. D* **70**, 064028 (2004), gr-qc/0404128.
  - [6] Y. Pan, A. Buonanno, Y. Chen, and M. Vallisneri, *Phys. Rev. D* **69**, 104017 (2004), URL <http://xxx.lanl.gov/abs/gr-qc/0310034>.
  - [7] C. Königsdörffer and A. Gopakumar, *Phys. Rev. D* **71**, 024039 (2005), gr-qc/0501011.
  - [8] A. Buonanno, Y. Chen, Y. Pan, H. Tagoshi, and M. Vallisneri, *Phys. Rev. D* **72**, 084027 (2005), gr-qc/0508064.
  - [9] C. Königsdörffer and A. Gopakumar, *Phys. Rev. D* **73**, 124012 (2006), gr-qc/0603056.
  - [10] M. Tessmer and A. Gopakumar, *MNRAS* **374**, 721 (2007), gr-qc/0610139.
  - [11] M. Hannam, S. Husa, B. Brügmann, and A. Gopakumar, *Phys. Rev. D* **78**, 104007 (2008), 0712.3787.
  - [12] I. Hinder, F. Herrmann, P. Laguna, and D. Shoemaker, *Phys. Rev. D* **82**, 024033 (2010), 0806.1037.
  - [13] K. G. Arun, A. Buonanno, G. Faye, and E. Ochsner, *Phys. Rev. D* **79**, 104023 (2009).
  - [14] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, *Phys. Rev. D* **80**, 084043 (2009).
  - [15] J. Abadie, B. P. Abbott, R. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, R. Adhikari, P. Ajith, B. Allen, et al., *Phys. Rev. D* **83**, 122005 (2011), 1102.3781.
  - [16] B. Aylott, J. G. Baker, W. D. Boggs, M. Boyle, P. R. Brady, D. A. Brown, B. Brügmann, L. T. Buchman, A. Buonanno, L. Cadonati, et al., *Classical and Quantum Gravity* **26**, 165008 (2009), 0901.4399.
  - [17] P. Kumar, I. MacDonald, D. A. Brown, H. P. Pfeiffer, K. Cannon, M. Boyle, L. E. Kidder, A. H. Mroué, M. A. Scheel, B. Szilágyi, et al., *Phys. Rev. D* **89**, 042002 (2014), 1310.7949.
  - [18] A. Buonanno, Y. Pan, H. P. Pfeiffer, M. A. Scheel, L. T. Buchman, and L. E. Kidder, *Phys. Rev. D* **79**, 124028 (2009), 0902.0790.
  - [19] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, et al., *Phys. Rev. D* **82**, 064016 (2010), 1005.3306.
  - [20] I. MacDonald, S. Nissanke, and H. P. Pfeiffer, *Classical and Quantum Gravity* **28**, 134002 (2011), 1102.5128.
  - [21] I. MacDonald, A. H. Mroué, H. P. Pfeiffer, M. Boyle, L. E. Kidder, M. A. Scheel, B. Szilágyi, and N. W. Taylor, *Phys. Rev. D* **87**, 024009 (2013), 1210.3007.
  - [22] I. Hinder, A. Buonanno, M. Boyle, Z. B. Etienne, J. Healy, N. K. Johnson-McDaniel, A. Nagar, H. Nakano, Y. Pan, H. P. Pfeiffer, et al., *Classical and Quantum Gravity* **31**, 025012 (2013), 1307.5307.
  - [23] N. Gehrels, E. Ramirez-Ruiz, and D. B. Fox, *ARAA* **47**, 567 (2009), 0909.1531.
  - [24] N. Gehrels, J. K. Cannizzo, and J. P. Norris, *New Journal of Physics* **9**, 37 (2007).
  - [25] P. Meszaros, *Reports of Progress in Physics* **69**, 2259 (2006), arXiv:astro-ph/0605208.
  - [26] LIGO Scientific Collaboration, Virgo Collaboration, J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, T. Accadia, F. Acernese, et al., *ArXiv e-prints* (2013), 1304.0670.
  - [27] B. D. Metzger and E. Berger, *Astrophys. J.* **746**, 48 (2012), 1108.6056.
  - [28] D. Grossman, O. Korobkin, S. Rosswog, and T. Piran, *MNRAS* **439**, 757 (2014), 1307.2943.
  - [29] K. Kyutoku, K. Ioka, and M. Shibata, *MNRAS* **437**, L6 (2014), 1209.5747.
  - [30] B. D. Metzger and G. C. Bower, *MNRAS* **437**, 1821 (2014), 1310.4506.
  - [31] M. Tanaka, K. Hotokezaka, K. Kyutoku, S. Wanajo, K. Kiuchi, Y. Sekiguchi, and M. Shibata, *Astrophys. J.* **780**, 31 (2014), 1310.2774.
  - [32] J. Barnes and D. Kasen, *Astrophys. J.* **775**, 18 (2013), 1303.5787.
  - [33] L. P. Singer, L. R. Price, B. Farr, A. L. Urban, C. Pankow, S. Vitale, J. Veitch, W. M. Farr, C. Hanna, K. Cannon, et al., *The Astrophysical Journal* **795**, 105 (2014), URL <http://stacks.iop.org/0004-637X/795/i=2/a=105>.
  - [34] J. Abadie and et al, *Tech. Rep. T1100338* (2013), URL <https://dcc.ligo.org/cgi-bin/DocDB/ShowDocument?docid=63432>.
  - [35] Abadie, J. et al. (The LIGO Scientific Collaboration), *A&A* **541**, A155 (2012), URL <http://dx.doi.org/10.1051/0004-637X/201219505>.

- 1051/0004-6361/201218860.
- [36] J. Aasi and T. L. S. Collaboration), The Astrophysical Journal Supplement Series **211**, 7 (2014), URL <http://stacks.iop.org/0067-0049/211/i=1/a=7>.
  - [37] K. Cannon, R. Cariou, A. Chapman, M. Crispin-Ortuzar, N. Fotopoulos, M. Frei, C. Hanna, E. Kara, D. Keppel, L. Liao, et al., The Astrophysical Journal **748**, 136 (2012), URL <http://stacks.iop.org/0004-637X/748/i=2/a=136>.
  - [38] T. Sidery, B. Aylott, N. Christensen, B. Farr, W. Farr, F. Feroz, J. Gair, K. Grover, P. Graff, C. Hanna, et al., arxiv:1312.6013 (2013), 1312.6013.
  - [39] S. Fairhurst, Classical and Quantum Gravity **28**, 105021 (2011), 1010.6192.
  - [40] J. Aasi et al (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. D **88**, 062001 (2013), URL <http://xxx.lanl.gov/abs/arXiv:1304.1775>.
  - [41] W. Del Pozzo, J. Veitch, and A. Vecchio, Phys. Rev. D **83**, 082002 (2011), 1101.1391.
  - [42] N. Cornish, L. Sampson, N. Yunes, and F. Pretorius, Phys. Rev. D **84**, 062003 (2011), 1105.2088.
  - [43] T. G. F. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio, Phys. Rev. D **85**, 082003 (2012), 1110.0530.
  - [44] J. Veitch, I. Mandel, B. Aylott, B. Farr, V. Raymond, C. Rodriguez, M. van der Sluys, V. Kalogera, and A. Vecchio, Phys. Rev. D **85**, 104045 (2012), 1201.1195.
  - [45] S. Nissanke, J. Sievers, N. Dalal, and D. Holz, Astrophys. J. **739**, 99 (2011), 1105.3184.
  - [46] V. Raymond, Ph.D. thesis, Northwestern University (2012), URL <https://gwic.ligo.org/thesisprize/2012/raymond-thesis.pdf>.
  - [47] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, et al., Phys. Rev. D **91**, 042003 (2015), URL <http://link.aps.org/doi/10.1103/PhysRevD.91.042003>.
  - [48] U. von Toussaint, Reviews of Modern Physics **83**, 943 (2011).
  - [49] W. M. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Reviews of Modern Physics **73**, 33 (2001).
  - [50] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Physics Letters B **195**, 216 (1987).
  - [51] G. P. Lepage, Tech. Rep. (1980).
  - [52] M. Creutz, Phys. Rev. D **21**, 2308 (1980).
  - [53] P. Jeackel, *Monte Carlo methods in finance*. (John Wiley and Sons, 2002), ISBN 0-471-49741-X.
  - [54] A. Papageorgiou, Math. Computation **70**, 297 (2001), URL <http://www.ams.org/journals/mcom/2001-70-233/S0025-5718-00-01231-X/home.html>.
  - [55] P. Canizares, S. E. Field, J. R. Gair, and M. Tiglio, Phys. Rev. D **87**, 124005 (2013), 1304.0462, URL <http://xxx.lanl.gov/abs/arXiv:1304.0462>.
  - [56] R. J. E. Smith, K. Cannon, C. Hanna, D. Keppel, and I. Mandel, Phys. Rev. D **87**, 122002 (2013), 1211.1254.
  - [57] K. Cannon, J. D. Emberson, C. Hanna, D. Keppel, and H. P. Pfeiffer, Phys. Rev. D **87**, 044008 (2013), 1211.7095.
  - [58] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, ArXiv e-prints (2013), 1308.3271.
  - [59] A. Lundgren and R. O'Shaughnessy, Phys. Rev. D **89**, 044021 (2014), URL <http://link.aps.org/doi/10.1103/PhysRevD.89.044021>.
  - [60] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, et al. (2014), 1404.6284.
  - [61] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel, eprint arXiv:1202.0790 (2012), 1202.0790.
  - [62] Y. Pan, A. Buonanno, L. T. Buchman, T. Chu, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, Phys. Rev. D **81**, 084041 (2010), 0912.3466.
  - [63] S. Babak, R. Biswas, P. R. Brady, D. A. Brown, K. Cannon, C. D. Capano, J. H. Clayton, T. Cokelaer, J. D. E. Creighton, T. Dent, et al., Phys. Rev. D **87**, 024033 (2013), URL <http://link.aps.org/doi/10.1103/PhysRevD.87.024033>.
  - [64] B. Farr and V. Kalogera, in *APS Meeting Abstracts* (2013), p. 10003.
  - [65] B. Farr, V. Kalogera, and E. Luijten, Phys. Rev. D **90**, 024014 (2014), 1309.7709.
  - [66] B. Farr, E. Ochsner, W. M. Farr, and R. O'Shaughnessy (2014), 1404.7070.
  - [67] H. Cho, E. Ochsner, R. O'Shaughnessy, C. Kim, and C. Lee, Phys. Rev. D **87**, 02400 (2013), URL <http://xxx.lanl.gov/abs/arXiv:1209.4494>.
  - [68] R. O'Shaughnessy, B. Farr, E. Ochsner, C. Cho, H.S. Kim, and C. Lee, Submitted to PRD (arXiv:1308.4704) (2013), URL <http://arxiv.org/abs/1308.4704>.
  - [69] Press, Teukolsky, Flannery, and Vetterling, *Numerical recipes* (1992), URL <http://www.nr.com>.
  - [70] G. Peter Lepage, Journal of Computational Physics **27**, 192 (1978).
  - [71] L. P. Singer, Ph.D. thesis, California Institute of Technology (2015).
  - [72] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, ArXiv e-prints (2014), 1402.5156.
  - [73] J. S. Read, L. Baiotti, J. D. E. Creighton, J. L. Friedman, B. Giacomazzo, K. Kyutoku, C. Markakis, L. Rezzolla, M. Shibata, and K. Taniguchi, Phys. Rev. D **88**, 044042 (2013), 1306.4065.
  - [74] S. Nissanke, D. E. Holz, S. A. Hughes, N. Dalal, and J. L. Sievers, Astrophys. J. **725**, 496 (2010), 0904.1017.
  - [75] S. R. Taylor, J. R. Gair, and I. Mandel, Phys. Rev. D **85**, 023535 (2012), 1108.5161.
  - [76] C. Messenger and J. Read, Physical Review Letters **108**, 091101 (2012), 1107.5725.
  - [77] *The first two years of electromagnetic follow-up with advanced ligo and virgo* (2015), URL <http://www.ligo.org/scientists/first2years/>.
  - [78] J. Abadie et al (The LIGO-Virgo Scientific Collaboration), CQG **27**, 173001 (2010), 1003.2480, URL <http://arxiv.org/abs/1003.2480>.
  - [79] C.-J. Haster, I. Mandel, and W. Farr (2015), arXiv:1502.05407, URL <http://arxiv.org/abs/1502.05407>.
  - [80] N. J. Cornish and T. B. Littenberg (2015), URL <http://arxiv.org/abs/1410.3835>.
  - [81] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. **622**, 759 (2005), astro-ph/0409513.