Reduced Order Modeling of Gravitational Wave Signals from Precessing Compact Binary Neutron Star Systems with Tidal Effects

Amanda Baylor*

Dickinson College, Carlisle, PA 17013, USA

Rory Smith†

OzGrav, School of Physics & Astronomy,

Monash University, Clayton 3800, VIC, Australia

(Dated: July 31, 2019)

Abstract

With the advancement of gravitational wave detectors came a new era of increased gravitational wave detections. There is now a greater need for fast and accurate inference methods to extract astrophysical information about their compact binary sources. Current methods of parameter estimation take excessive computational time. We introduce a reduced order model that will generate a significant speedup on Bayesian inference analyses and produce posterior probability distributions on the parameters of binary neutron star systems.

* baylora98@gmail.com

† rory.smith@monash.edu

I. INTRODUCTION

Gravitational wave (GW) detectors are now more sophisticated than ever before. Advanced LIGO (aLIGO) and Advanced Virgo (AdV) resumed operation for observing run 3 (O3) in April 2019, which is expected to be the longest continuous observing run yet. With improved detector sensitivity, aLIGO and AdV will be able to make a few detections per month [1].

The main sources for aLIGO and AdV are from compact binary coalescence (CBC) events, such as binary black hole (BBH) mergers, binary neutron star (BNS) mergers, or even merging neutron star black hole (NSBH) pairs. One of the key goals of gravitational wave astronomy is to extract astrophysical information from the parameters of these sources, such as the individual masses or spins, using Bayesian inference methods. This process of parameter estimation uses a model for the signal to output the parameters of the source.

Although LIGO already has methods in place to perform parameter estimation, they can be computationally expensive. Without reduced order models (ROMs), PE takes approximately a few weeks for BBH mergers and up to several months for BNS mergers [2]. With aLIGO and AdV making more frequent detections, this can create bottlenecks in the analysis pipeline. Therefore, there is a greater need for fast and accurate inference methods. Reduced order modeling can significantly cut down the computation time required by reducing the number of iterations needed to represent the full model.

Previous reduced order models have been able to estimate most source parameters within just a few hours to days with reasonable accuracy [2]. However, we would like to extract the full set of parameters from BNS signals, which includes precession and tidal effects. We present a ROM which will allow us to significantly speed up parameter estimation on BNS systems. The system, GW190425, is described in detail in Sec. III.

II. PARAMETERS OF GW SIGNALS

A gravitational wave signal, or strain, is a dimensionless quantity:

$$\text{Strain} = \frac{\Delta L}{L} \tag{1}$$

where L is the arm length of the interferometer. The strain can also be represented as:

$$h(\vec{\Lambda}; t) = F_+ h_+ + F_x h_x \tag{2}$$

where h_+ and h_x are the strain in the plus- and cross-polarizations of the GW. These polarizations describe how much the GW is squeezed or stretched in the plus and cross directions [source or image]. F_+ and F_x represent the antenna patterns which project the GW's plus- and cross-polarizations into the detector's frame, and F depends only on the orientation and location of the binary with respect to the detector. Fig. 1 shows the orientation of a binary system in the source's frame.

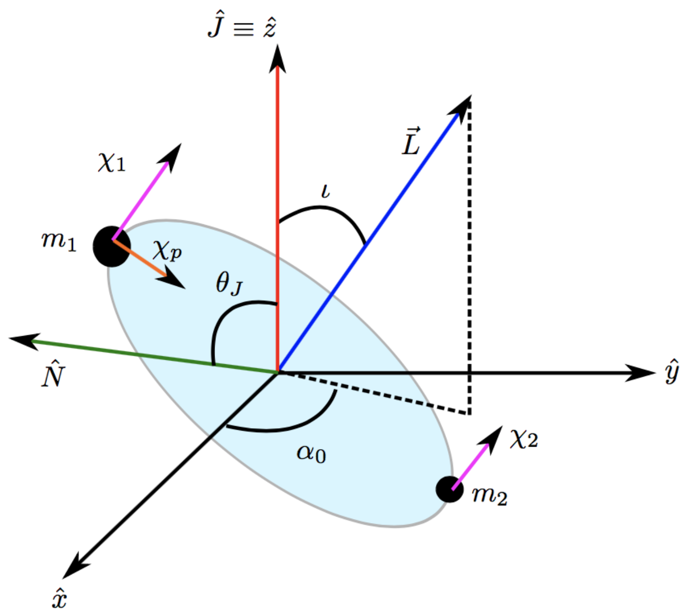

FIG. 1. The orientation and characteristic parameters of a binary system in the source's frame. $\hat{N}$ points in the direction of Earth. The angular momentum $\hat{J} = \hat{L} + S_1 + S_2$ is defined to be aligned with the z-axis since it stays approximately fixed as $\hat{L}$ evolves with time for a precessing binary. We always label the heavier component as m_1. [2]

BBH signals are characterized by fifteen parameters, and BNS by seventeen. The two additional parameters account for the tidal deformability of each neutron star component (λ_1, λ_2), which describe the amount that a neutron star is stretched or squeezed due to its companion [3]. All parameters and their descriptions are listed in Table 1.

Certain parameters, ra, dec, r, ψ, t_c, ϕ_c are considered trivial because they have little effect on the shape of the waveform. Trivial parameters can be set to a fixed constant, as they usually enter the strain as an overall scaling factor. See Sec. VI for how we do this in practice.

Parameters	**Description**
ra	Right Ascension
dec	Declination
r	Distance to source
ψ	Polarization angle
t_c	Time of coalescence
ϕ_c	Orbital phase at coalescence
$\mathcal{M}_c$	Chirp mass $= \frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}}$
q	Mass ratio $= \frac{m_1}{m_2}$
χ_1, χ_2	Component spins in the $\hat{L}$ direction
χ_{p_1}, χ_{p_2}	Component spins in the orbital plane, orthogonal to $\hat{L}$
θ_j	Amount binary is tilted with respect to the coordinate plane
α	Amount L has rotated around J
i	Inclination angle
λ_1, λ_2	Tidal deformability (BNS only)

TABLE I. The list of parameters that characterize a binary system. BBH systems are characterized by the first fifteen and BNS by all seventeen.

All other parameters are non-trivial and must be estimated carefully, especially for precessing models. The spin parameter and proxy for precession, χ_p, modulates the waveform. The component masses affect the frequency and the length of the signal. For example, in Fig. 2, we see that the BNS merger GW170817 is significantly longer in duration than the BBH merger GW150914. Lighter objects like neutron stars will take longer to merge from a given orbital frequency.

F is a function of r, ra, dec, and ψ only, but h depends on the full set of parameters $\vec{\Lambda}$. Consequently, it is difficult to create a model for h that accurately estimates all of them simultaneously. Moreover, h also depends on the duration of the signal. The longer a binary signal, the harder it is to perform parameter estimation. It is for this reason that signals from BNS mergers require much more computational time than BBH mergers. See Sec. VII for the calculation of the speedup on our analyses. The next section will introduce the BNS system we have performed parameter estimation on.

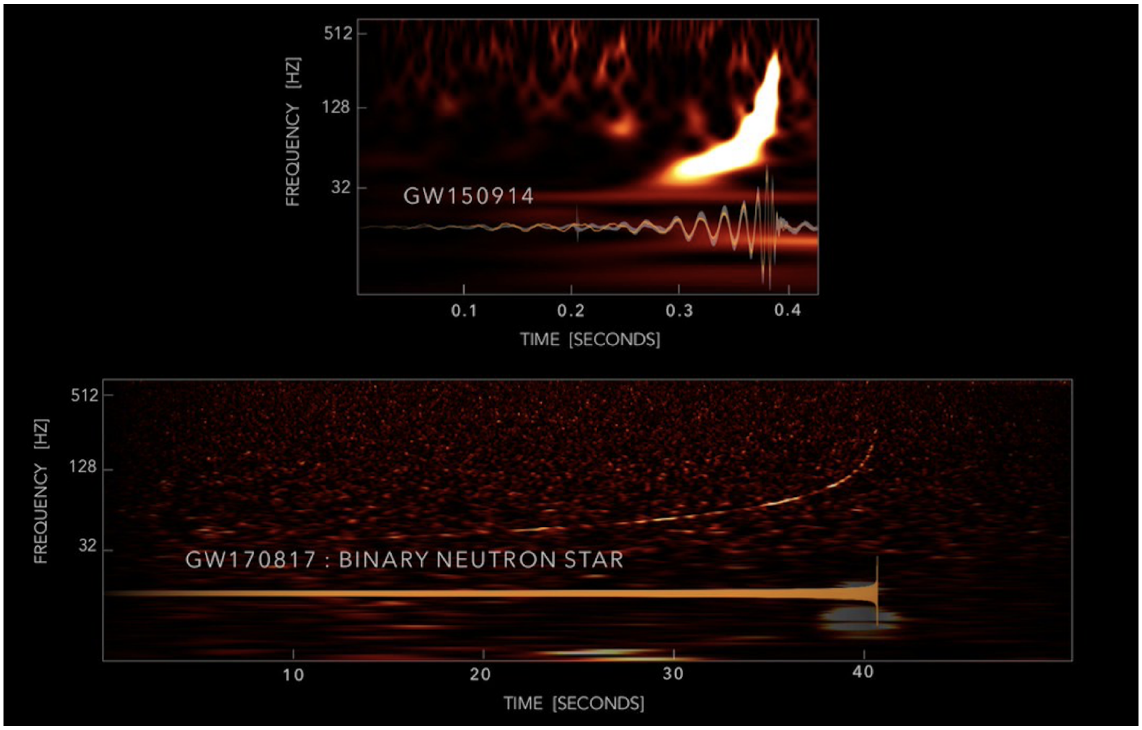

FIG. 2. The spectograms and waveforms for the first detected BBH merger GW150914 and BNS merger GW170817. The duration of the BNS merger is significantly longer than that of the BBH [4]

.

III. GW190425

GW190425 was recently detected at the beginning of O3 by aLIGO and AdV [Source?]. All public information about this event can be accessed through the online Gravitational-Wave Candidate Event Database (GraceDB) [5]. GW190425 has been identified as a BNS merger with greater than 99 percent confidence and a false alarm rate of 1 per 69834 years. This makes it extremely desirable for astrophysical analysis.

GW190425 has been our motivation for this project. As the only detected BNS merger since GW170817, GW190425 can reveal substantial astrophysical information through gravitational waves. By performing parameter estimation on the full set of parameters for GW190425 using our ROM, we were able to access this information even faster.

Estimating the mass distributions on GW190425 was important to contribute to the evidence of a constraint on neutron star masses [6]. Moreover, we performed parameter estimation for two sets of chirp mass priors. The first run was a "regular" distribution of $1.42M_{\odot} \leqslant \mathcal{M}_c \leqslant 2.6M_{\odot}$. After obtaining the probability distribution on $\mathcal{M}_c$ for the regular chirp mass, we observed that we could further constrain the chirp mass to a narrower region for a more accurate result. For more on priors, see section VI.

The tidal deformability of BNS systems allows us to reconstruct the components' equations of state [3]. We determined $\lambda_{1,2}$ for GW190425, which depend on the radius of the star and the Love number. These quantities can then be used to solve the Tolman-Oppenheimer-Volkoff (TOV) equation. By recovering the equation of state for a neutron star, we gain information about the behavior of matter in the cores of the densest objects in the universe.

The spin measurements on GW190425 are also able to tell us about BNS formation history. If the spin alignment is found to be close to zero, then we can conclude that the two neutron star components formed close together, causing their spins to become aligned over time. A non-zero spin alignment would indicate that the components formed separately and merged before their spins aligned.

The following sections describe the background behind parameter estimation and how we do this in practice.

IV. PARAMETER ESTIMATION

A. Bayesian Inference

We base our method of parameter estimation off of Bayes' Theorem and the use of conditional probability: $P(A|B)$ = the probability that A happens given that B happens. Our objective is to compute the posterior probability density function (PDF):

$$P(\vec{\Lambda}|d) = \frac{\pi(\vec{\Lambda})\mathcal{L}(d|\vec{\Lambda})}{z(d)} \tag{3}$$

such that we obtain the probability of our source's parameters, given the data. An example posterior PDF is shown in Fig. 3, where we created a fake signal and used our ROM to extract the real parameter value.

Our data is the sum of the signal and the total noise: $d = h(\vec{\Lambda}) + n$. $\pi(\vec{\Lambda})$ represents the prior probability, or what we think we know about the parameters before analyzing. $z(d)$

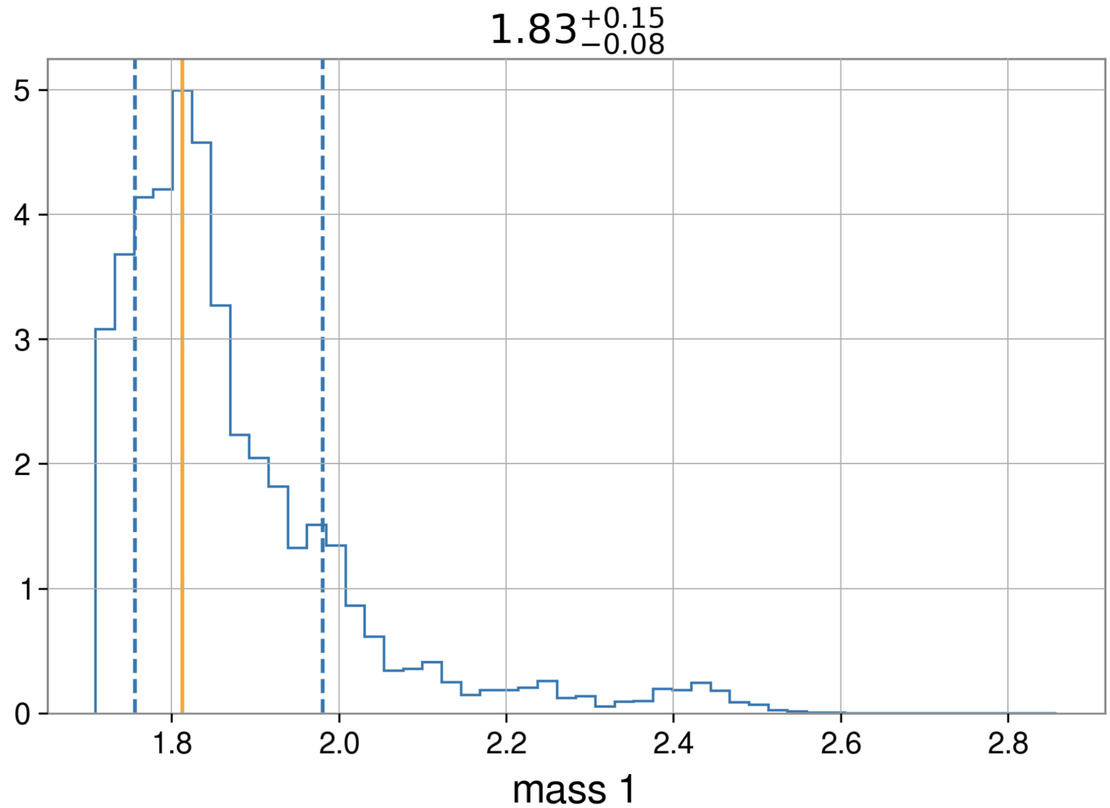

FIG. 3. An example posterior probability density function from a test of our ROM on fake data. The orange line represents the true parameter value for mass 1, and the dashed lines mark the percent credible region with 60 percent chance of containing the true value.

is the Bayesian evidence, a normalization factor that gives us the probability of our data, assuming it contains a signal.

The likelihood $\mathcal{L}$ tells us the probability of having obtained the data if it contained a signal with parameters $\vec{\Lambda}$ and noise n. $\mathcal{L}$ is the computationally expensive piece of parameter estimation:

$$log[\mathcal{L}(d|\vec{\Lambda})] = -\frac{1}{2}\langle d - h(\vec{\Lambda}), d - h(\vec{\Lambda})\rangle \tag{4}$$

Since the noise n is Gaussian, the likelihood $\mathcal{L}$ is an exponential, and we take the $log(\mathcal{L})$ for simplicity. We can break the likelihood up into separate pieces which can be approximated individually:

$$log[\mathcal{L}(d|\vec{\Lambda})] = \langle d, h\rangle - \frac{1}{2}(\langle h, h\rangle - \langle d, d\rangle) \tag{5}$$

where $\langle d, d\rangle$ is given by our data. We are left with the linear term $\langle d, h\rangle$ and the quadratic term $\langle h, h\rangle$. The linear term is an overlap integral, or a weighted inner product for discretely sampled noisy data:

$$\langle d, h\rangle = 4\Re\,\Delta f \sum_{k=1}^{L} \frac{\tilde{d}^*(f_k)\tilde{h}(\vec{\Lambda}; f_k)}{S_n(f_k)} \tag{6}$$

where $S_n(f_k)$ is the noise power spectral density (PSD) of the detector. $\tilde{d}$ and $\tilde{h}$ are discrete Fourier transforms of the data and the strain, which can be evaluated by:

$$\tilde{h}(\vec{\Lambda}; f) = \int_{-\infty}^{\infty} h(\vec{\Lambda}; t) e^{2\pi i f t} dt \tag{7}$$

and similarly for $\tilde{d}$. The properties of the detector's noise are well-understood in the frequency domain, and it is more efficient to take the Fourier transform prior to evaluating the likelihood rather than after. It is also easier to work directly with PSD of the detector, which describes the average power at each frequency bin in units of [Strain/Hz].

The likelihood is proportional to the number of data points, and evaluating over all frequencies $\{f_k\}_{k=1}^{L}$ takes significant computational time. Moreover, we want to estimate the full set of parameters $\vec{\Lambda}$, which the likelihood also depends on. Computing the sum (6) and repeating our evaluation creates the bottlenecks in our analysis pipeline. Therefore, we use a reduced order quadrature (ROQ) rule to reduce the number of iterations needed to evaluate the likelihood.

B. ROMs

The ROM is a computationally efficient representation of our waveform. $\tilde{h}$ is usually expressed as:

$$\tilde{h}^{ROM}(\vec{\Lambda}; f) = \sum_{i=1}^{N} \tilde{h}^{Full}(\vec{\Lambda}; F_i) B_i(f) \tag{8}$$

where $\{F_i\}_{i=1}^{N}$ is a set of judiciously chosen frequency nodes (F-nodes) which are selected from the full set $\{f_k\}_{k=1}^{L}$. $\{B_i(f)\}_{i=1}^{N}$ is a set of basis functions which each represent a unique waveform. We can then re-express (8) as:

$$\tilde{h}^{ROM}(\vec{\Lambda}; f) \approx \tilde{h}^{Full}(\vec{\Lambda}; f) \tag{9}$$

We utilize a reduced basis method to build our ROM and approximate the waveform, such that we find a better representation of the *Full* model that will mitigate its complexity. We can reconstruct the waveform from the F-nodes without performing as many mathematical operations. Our goal is to have $N \lll L$, where the theoretical speedup of our analysis is $\approx \frac{L}{N}$.

C. ROQ Rule

Built from the ROM, the ROQ is a compressed inner product that will reduce the size of the inner products of the likelihood function. The ROQ builds off of (8) by reducing the number of frequency bins in some optimal way. Beginning with the linear term:

$$\langle d, h \rangle_{ROQ} \approx \sum_{i=1}^{N} \tilde{h}(\vec{\Lambda}; F_i)\omega_i \tag{10}$$

where ω_i are the data-dependent ROQ weights:

$$\omega_i = 4\Re\,\Delta f \sum_{j=1}^{M} \frac{\tilde{d}^*(f_j)B_i(f_j)}{S_n(f_k)} e^{-2\pi i f_j t_c} \tag{11}$$

where $M \gg N$. For the quadratic term, we need a separate ROQ:

$$\langle h, h \rangle_{ROQ} \approx \sum_{k=1}^{N} \tilde{h}^*(\vec{\Lambda}; \mathcal{F}_k)\,\tilde{h}(\vec{\Lambda}; \mathcal{F}_k)\psi_k \tag{12}$$

for a unique set of F-nodes $\{\mathcal{F}_k\}_{k=1}^{N}$ with a different integration weight that is not data-dependent:

$$\psi_k = 4\Re\,\Delta f \sum_{j=1}^{M} \frac{C_k(f_j)}{S_n(f_j)} \tag{13}$$

which is built from a different basis set $\{C_k\}_{k=1}^{N}$. Since the bases are not orthonormal, we need to handle them individually to avoid complications.

V. BUILDING THE ROM

The model off which we are basing our ROM is IMRPhenomPv2_NRTidal, version two of the inspiral-merger-ringdown, phenomenological model for precessing BNS systems with tidal effects, estimated using numerical relativity. Previous models have been used to estimate a subset of parameters for BBH and BNS systems [2, 7]. However, this will be the first time we use a ROM on IMRPhenomPv2_NRTidal.

1. *Construct the Training Set*

The first step is to construct a training set, which is a list of numbers that each represent a unique waveform with parameters $\vec{\Lambda}$. Our representation of the model will be built from elements of the training set.

We define the minimum and maximum values for each parameter and then construct grids of parameter spaces. To achieve a dense set of waveforms, we use a mixture of grid-based sampling as well as random sampling. Then, we grab only unique sets of values so there are no repeated waveforms.

Our model is designed for BNS signals up to 128 seconds in duration. We constructed two different training sets for each of the two chirp mass priors. The number of training points for each set and the frequency range is listed in Table II.

Training Set Size		f (Hz)	
Regular M_c	Narrow M_c	Min	Max
438,309	1,100,000	20	2048

TABLE II. Sizes of the two training sets constructed for two separate chirp mass priors. Both training sets span the same range of frequencies.

2. *Implement the Greedy Algorithm*

Next we apply a greedy algorithm to our training set to build our reduced basis. We utilize the code greedycpp, a parallelized algorithm written in C++ [8]. The algorithm will select parameter points from the training set with their corresponding waveforms and add them to the basis set, $\{B_i(f)\}_{i=1}^{N}$, such that $N \lll L$. We want the smallest, most accurate basis possible.

We check how accurate the basis is by calculating the overlap:

$$O = \int_{-\infty}^{\infty} \hat{h}^{ROM} \hat{h}^{True} df \tag{14}$$

which tells us how well our ROM matches the $True$ waveform. If $\hat{h}^{ROM}$ was perfectly overlapping $\hat{h}^{True}$, then O would equal 1 and the overlap error $\epsilon = 1 - O$ would equal zero.

The algorithm will find points with the worst error by calculating ϵ and add them back into the basis. The greedy will stop after it has reached our defined tolerance value of $\epsilon \approx 10^{-8}$.

3. Validation

Validation must be performed on our basis set before we can be sure we have the most accurate basis for building our ROM. This step involves random sampling of points outside of our training set. Again using greedycpp, we compute the error of random parameter values and their associated waveforms with errors larger than the tolerance value. These points are then added back into our original training set to create an enriched set.

We then reapply the greedy algorithm to the enriched set to produce a more accurate basis. We can repeat steps 2 and 3 as much as we like until we are satisfied with the accuracy and size of our basis, but sufficient accuracy can be reached with just one or two validations.

4. Empirical Interpolation

Finally, we form the empirical interpolant which creates the ROM (9) so that we can replace the weighted inner product (6) with the ROQ approximate (10). We would like to be able to express any waveform evaluated at any time for an arbitrary set of parameters as a sum of reduced basis elements with very high accuracy. We use the empirical interpolation method, also executed by greedycpp, to judiciously select our F-nodes $\{F_i\}_{i=1}^{N}$. We want our selections to be good, not randomly or equally spaced. A similar series of steps is taken to form the quadratic term (12) from building the basis $\{C_k\}_{k=1}^{N}$ to determining the set of F-nodes $\{\mathcal{F}_k\}_{k=1}^{N}$.

VI. ASTROPHYSICAL ANALYSIS

With the reduced bases and empirical interpolation nodes in hand, we can now analyze GW signals. Recall that the objective is to extract the full set of parameters $\vec{\Lambda}$ by computing the posterior PDF (3). We have estimated the likelihood (4) using a ROQ rule, which can now be substituted into (3).

Next, we define our priors $\pi(\vec{\Lambda})$ according to the GW source we wish to analyze. Trivial parameters can be determined by uniform distributions on their priors, such that there is no assumption on a maximum or minimum value. We do have reason to constrain certain parameters, such as chirp mass or individual spins. The upper limit on spin priors for BNS systems is usually 0.8. Additionally, tidal deformability is a dimensionaless number that runs from 0 to 5000. For non-trivial parameters, we define maxima and minima based on prior knowledge from previous analyses, with a uniform distribution on the priors.

Parameter	**Min**	**Max**
Regular $\mathcal{M}_c$	1.42	2.60
q	0.125	1
χ_1, χ_2 (High)	0	0.8
χ_1, χ_2 (Low)	0	0.4
χ_{p1}, χ_{p2}	0	0.89
λ_1, λ_2	0	5000

TABLE III. Ranges for priors of select non-trivial parameters. We built separate training sets for the two different chirp mass priors, with all other prior ranges remaining the same, and we performed separate analyses for two different spin priors.

We use nested sampling to obtain the posterior PDF. We input the data and our likelihood, and in return we receive the probability distributions on our model parameters and the evidence factor $z(d)$. Recall that the evidence gives us the probability of the data given the model. Nested sampling estimates this factor for us, by taking the integral:

$$z(d) = \int \mathcal{L}(d|\vec{\Lambda})\pi(\vec{\Lambda})d\vec{\Lambda} \tag{15}$$

It also computes the Bayes factor, which is the ratio of two evidences, one for the data containing signal plus noise, and one for the data being just noise:

$$z_{ratio} = \frac{z(d|h+n)}{z(d|n)} \tag{16}$$

Since the z_{ratio} is usually a significantly large exponential, we use the $log(z_{ratio})$ to tell us how much more likely it is for the data to contain a signal plus noise versus noise only. For example, a $log(z_{ratio})$ of 44 means that our z_{ratio} is e^{44} and it is e^{44} times more likely that our data contains a signal. The Bayes factor can be used for model selection by comparing the strength of one model to another. The higher the Bayes factor, the better the model can represent the data.

A. LALInference

We begin by running our analysis on GW190425 through LIGO's parameter estimation library, LALInference [9]. LALInference is written in C, with some post-processing tools written in python. The nested sampling algorithm implemented by LALInference will compute the evidence (16) and stop once the difference in the estimate of $log(z)$ from the last iteration to the current one is less than 0.1. It will also compute the Bayes factor (17).

We decided to perform four separate analyses in LALInference. First, we performed nested sampling with the basis built for the regular chirp mass prior for a high spin prior of $0 \leqslant \chi_{1,2} \leqslant 0.8$. We then decided to do the same for a low spin prior of $0 \leqslant \chi_{1,2} \leqslant 0.4$. This process for both low and high spins was then performed with the basis built for the narrow chirp mass prior. An outline of all runs performed on GW190425 can be found in Table IV.

Once LALInference completes the algorithm, we receive posterior PDFs on each one of our model parameters. As described in Sec. IV, the posterior PDF will tell us the most likely values for all seventeen parameters of GW190425.

B. Bilby

We then run the same analysis through Bilby, a Bayesian inference library for gravitational wave astronomy [10]. Bilby is written in python and will be used to replace LALInference during O3, so it is useful to have results from both algorithms. Bilby will perform nested sampling in the same way as LALInference, but is usually faster by a few days. The four runs which we performed in LALInference were repeated with Bilby. Again, we receive posterior PDFs on each one of the parameters of GW190425.

	LALInference		Bilby	
$\mathcal{M}_c$ **Prior**	Regular M_c	Narrow M_c	Regular M_c	Narrow M_c
High Spin Prior	$\chi_{1,2} < 0.8$	$\chi_{1,2} < 0.8$	$\chi_{1,2} < 0.8$	$\chi_{1,2} < 0.8$
Low Spin Prior	$\chi_{1,2} < 0.4$	$\chi_{1,2} < 0.4$	$\chi_{1,2} < 0.4$	$\chi_{1,2} < 0.4$

TABLE IV. We performed a total of eight analyses on GW190425 through the nested sampling algorithms implemented by Bilby and LALInference, one for each a low and high spin prior for both the regular and the narrow chirp mass priors.

VII. MODEL EFFICIENCY

To determine the efficiency of our ROM, we calculate the approximate speedup factor using the sizes of our reduced bases:

$$\text{Speedup} \approx \frac{\#\text{ of total frequency points}}{\#\text{ of linear basis elements} + \#\text{ of quadratic basis elements}} \tag{17}$$

We built separate training sets for the two different parameter estimation runs on GW190425, one for the regular chirp mass and one for the narrow. In each case, we just changed the minimum and maximum values for grid sampling of the chirp mass.

$\mathcal{M}_c$	**Basis Size**		**Speedup**
	Linear	Quadratic	
Regular	1415	402	143
Narrow	406	366	336

TABLE V. The calculated speedup factor and sizes of bases constructed from two different training sets used for parameter estimation on GW190425. The regular chirp mass prior is defined in Sec. III. The narrow chirp mass prior cannot be revealed at this time.

The speedup is determined largely by the likelihood function, which dominates the computational cost of parameter estimation. If we are able to successfully perform parameter estimation on a BNS signal in one day, the theoretical speedup of our analysis means that it would normally take 143 days to analyze the signal with our regular chirp mass prior, and 336 days for the narrow. However, parameter estimation with our ROM typically took a few days, so the computational time that we saved is significant. Such ROMs can be used for future analyses of detected BNS signals.

VIII. CONCLUSIONS

In a short amount of time we were able to construct a fast and efficient ROM for the IMRPhenomPv2_NRTidal waveform that successfully estimated all seventeen parameters of the BNS system GW190425. The theoretical speedup achieved with this ROM can significantly reduce the computational time for future analyses on BNS signals.

IX. ACKNOWLEDGEMENTS

We thank Monash for the creation and development of Bilby. We also thank the National Science Foundation and the University of Florida for supporting our work through a research grant.

[1] https://wiki.gw-astronomy.org/pub/OpenLVEM/TownHallMeetings2018/O3_rates_amsterdam.pdf.

[2] K. B. C.-J. H. M. P. V. R. Rory Smith, Scott E. Field and P. Schmidt, Phys. Rev. **D94** (2016).

[3] T. Hinderer, B. D. Lackey, R. N. Lang, and J. S. Read, Phys. Rev. **D81** (2010).

[4] https://twitter.com/hashtag/gwtc1.

[5] https://gracedb.ligo.org/superevents/S190425z/view/.

[6] J. Alsing, H. O. Silva, and E. Berti, Mon. Not. Roy. Astron. Soc. **478**, 1377 (2018).

[7] P. Schmidt, F. Ohme, and M. Hannam, Phys. Rev. **D91** (2015).

[8] S. E. Field, "greedycpp," https://bitbucket.org/sfield83/greedycpp/src/master/ (2019).

[9] J. Veitch *et al.*, Phys. Rev. **D91** (2015).

[10] G. Ashton *et al.*, Astrophys. J. Suppl. **241** (2019).