

Fast waveform generation for gravitational waves using evolutionary algorithms

Quirijn Meijer^{1,2,*} and Sarah Caudill^{3,4}

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

²*Nikhef, Science Park 105, 1098XG Amsterdam, The Netherlands*

³*Department of Physics, University of Massachusetts, Dartmouth, Massachusetts 02747, USA*

⁴*Center for Scientific Computing and Data Science Research, University of Massachusetts, Dartmouth, Massachusetts 02747, USA*



(Received 11 June 2024; accepted 31 July 2024; published 26 August 2024)

Gravitational-wave analyses depend heavily on waveforms that model the evolution of compact binary coalescences as seen by observing detectors. In many cases these waveforms are given by waveform approximants, models that approximate the amplitude and phase of the waveform at a set of frequencies. Because of their omnipresence, improving the speed at which approximants can generate waveforms is crucial to accelerating the overall analysis of gravitational-wave detections. An optimization algorithm is proposed that can select at which frequencies in the spectrum an approximant should compute the power of a waveform, and at which frequencies the power can be safely interpolated at a minor loss in accuracy. The algorithm used is an evolutionary algorithm modeled after the principle of natural selection, iterating frequency arrays that perform better at every iteration. As an application, the candidates proposed by the algorithm are used to reconstruct signal-to-noise ratios. It is shown that the IMRPhenomXPHM approximant can be sped up by at least 30% at a loss of at most 2.87% on the drawn samples, measured by the accuracy of the reconstruction of signal-to-noise ratios. The behavior of the algorithm as well as lower bounds on both speedup and error are explored, leading to a proposed proof of concept candidate that obtains a speedup of 46% with a maximum error of 0.5% on a sample of the parameter space used.

DOI: [10.1103/PhysRevD.110.043035](https://doi.org/10.1103/PhysRevD.110.043035)

I. INTRODUCTION

Gravitational waves were first predicted well over a hundred years ago by Heaviside, Poincaré, and Einstein, with Einstein famously remaining skeptical of their existence throughout his life. Einstein believed that the amplitude of gravitational waves reaching earth would be so small that their detection could never be within human reach. Despite the accurate prediction of their magnitude, the LIGO-Virgo collaboration succeeded in making their first detection of a gravitational wave in 2015 [1]. Since then, the LIGO-Virgo-KAGRA collaboration has confirmed detection of over 90 gravitational waves [2–5]. Gravitational waves are now used to study black holes, neutron stars, and other cosmological sources of gravitational waves.

A prominent routine in the analysis of gravitational wave data is matched filtering [6], a process in which templates that predict the time or frequency evolution of a gravitational wave are convolved with incoming data [6]. In order to cover many possible source systems, templates for

different systems are collected in template banks [7]. In turn, the templates populating this bank are generated by waveform approximants. Waveform approximants [8–12] serve to approximate the waveform of a system either in the time or frequency domain and are faster alternatives to the more precise waveform simulations from numerical relativity [13–17]. Approximants balance the need for both speed and accuracy, and therefore approximate a waveform to a reasonable accuracy [18] within a reasonable time. Other applications for approximants are in parameter estimation [19], where waveforms are generated during run-time, so that there too computational speed is crucial. If waveforms are computed during run-time this is called on-line computation, whereas it is called off-line computation if this task is delegated to the preprocessing stage.

Since the Fourier transform provides a duality between the time domain and frequency domain, the domain of analysis is a choice. Sometimes it is preferential to work in the frequency domain, where a different perspective on data is allowed, or computations are easier to perform. This is the case for the methods presented in this paper, where the frequency contents can be considered the building blocks of a waveform. A waveform approximant generating a

*Contact author: r.h.a.j.meijer@uu.nl

waveform in the frequency domain will do so by solving sets of equations at individual frequencies in the spectrum [20]. Doing so is costly, notably when waveforms are generated on-line during the run-time of a procedure, but also in generally large undertakings such as template bank validation where percentual changes make a large absolute difference [21–24]. One way of reducing the cost of computation would be to evaluate a waveform on as small an array of frequencies as possible, resorting to alternative, faster procedures to determine the behavior at the remaining frequencies. This is the premise of this paper.

With this line of reasoning, a method that can speed up waveform approximants is calling the solver for a subset of the frequencies, and interpolating the values of the frequencies in the complement within the frequency spectrum. Finding where in the spectrum to solve and where to interpolate is then an optimization problem, in particular one with a very large solution space that has size growing exponentially in the number of frequencies. One way to efficiently find a (local) optimum for this problem is through the use of evolutionary algorithms [25].

Evolutionary algorithms mimic natural evolution, where successive generations are expected to be better adapted to their environments. This principle translates to an algorithm that at each iteration tests a set of solutions, collectively called a generation, and then forms a new generation that contains better solutions to the problem. Evolutionary algorithms, sometimes also referred to as genetic algorithms, have been applied to gravitational-wave data analysis in other work [26–28].

Alternative methods with the same motive have been proposed, such as multiband template interpolation [29] and adaptive frequency resolution [30]. However, both methods make assumptions on the frequency evolution of the waveform to reduce the number of frequencies at which the power is computed, and adaptive frequency resolution in particular was shown to be accurate mostly for relatively loud signals. Evolutionary algorithms require no assumptions to be made and are therefore more widely applicable, with no specific boundary on the loudness. Other related methods include the work done on reduced order quadratures [31–33] and meshfree approximations [34].

The application of evolutionary algorithms to the optimization problem of waveform generation is studied in this paper as a proof of concept, proposing an algorithm, and studying both the time speedup and the coverage of the parameter space. Coverage here means the generalizability of the frequency array output by the algorithm to samples taken from the parameter space, or on what subset of the parameter space the interpolation from the output frequency array gives accurate approximations.

The contents of this paper are ordered as follows. In Sec. II approximants are briefly introduced. In Sec. III evolutionary algorithms are treated. Section IV then outlines the used methodology, including the design of the

algorithm and the test cases for the output. Results are reported in Sec. V before the work is reviewed in Sec. VI.

II. APPROXIMANTS

In a compact binary coalescence, as two objects such as black holes or neutron stars orbit each other, energy is lost through gravitational wave emission [35]. A typical process consists of the inspiral, during which the objects inspiral closer together, the merger, at which point the objects coalesce, and the ringdown, where the single resultant object settles down into its nature [35].

In the case of a noneccentric binary black hole system, its evolution is uniquely defined by two mass parameters, six spin parameters, two sky location parameters along with the phase of coalescence, inclination, polarization angle, luminosity distance, and the time of arrival [36]. These parameters are collected in a 15-dimensional parameter space denoted as \mathcal{P} . Since there are no *a priori* restrictions on the spin parameters, a binary black hole can be defined by a set of parameters with arbitrary spin orientations. If this is the case, the total angular momentum and orbital angular momentum may be misaligned, causing precession of the orbital plane [36]. It is then said that this system demonstrates precession, appearing in the waveform through amplitude and phase modulations [36,37].

If instead of a binary black hole a system containing a neutron star is considered, an additional parameter describing the tidal deformability of the star is added to this space [38]. This paper will however restrict itself, and therefore the parameter space, to the case of black holes.

A waveform approximant defined on these parameters can then be interpreted as a function:

$$A_{\mathbb{F}}: \mathcal{P} \rightarrow D(\mathbb{F}, \mathbb{C}) \quad (1)$$

that maps a given set of parameters into the set of discrete complex functions on the frequency spectrum $\mathbb{F} \subset \mathbb{R}$, which should not be confused with an algebraic field. Note that $D(\mathbb{F}, \mathbb{C})$ can alternatively be written as $\mathbb{C}^{\mathbb{F}}$, and that the notation $A_{\mathbb{F}}$ specifically includes the spectrum \mathbb{F} . Such a waveform approximant is typically the best possible option for the modeling of a system's evolution short of an analytical or numerical solution of the Einstein field equations with stress-energy-momentum tensor defined by the parameters of choice. This will be discussed in more detail shortly.

Waveform approximants can include what are known as higher order modes [35]. In the time domain, a waveform $h(t)$ with polarizations $h_+(t)$ and $h_{\times}(t)$ can be written in multipole expansion as a linear combination of the spherical harmonics $Y_{l,m}^{-2}$ of weight 2 [35,39–45]:

$$h_+(t) + ih_{\times}(t) = \sum_{l \leq 2} \sum_{m=-l}^{m=l} Y_{l,m}^{-2}(t, \varphi) h_{l,m}(t), \quad (2)$$

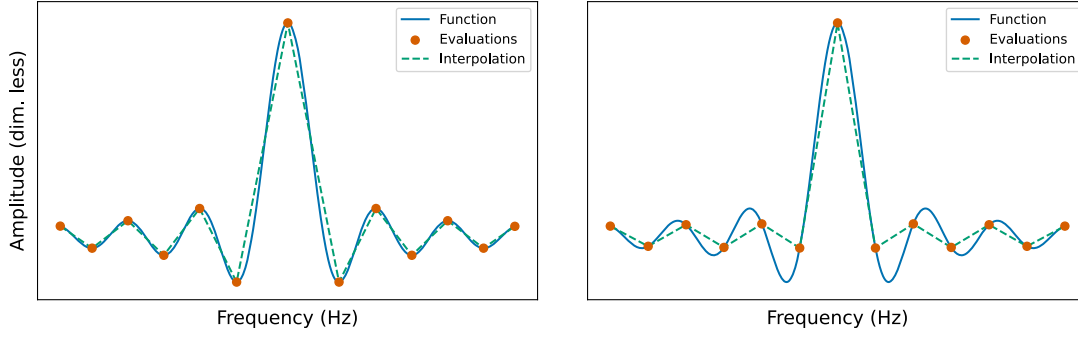


FIG. 1. A damped sinusoid interpolated on two different frequency arrays. The left frequency array consists of the turning points of the function, whereas the right frequency array is uniformly spaced. This figure shows both how frequency arrays can be realized differently, as well as the importance of the choice of array. Although both contain the same number of frequencies, corresponding to the evaluations in the plots, the left interpolation is clearly a better approximation to the true function.

where the $h_{l,m}(t)$ are the modes of gravitational-wave emission. Note that the spherical harmonics depend on the polar angle (ι) and the azimuthal angle (φ), which partially define a frame of reference for the binary black hole with origin in the center of mass of the system. In typical searches it is assumed the quadrupole $(l, m) = (2, \pm 2)$ mode dominates the expansion [41], and moreover, the maxima of $Y_{2,\pm 2}^{-2}(\iota, \varphi)$ at $\iota = 0, \pi$ coincide with face-on sources for which the dominant direction of gravitational-wave emission is aligned with the normal vector to the detector plane. For these reasons the quadrupole mode may serve as a good approximation for many waveforms, akin to a lower-order Taylor expansion. However, for other systems the higher order modes may contribute significantly to the expansion [41], for instance for edge-on systems where $\iota = \pi/2$. This value of ι corresponds to a minimum of $Y_{2,\pm 2}^{-2}(\iota, \varphi)$, but a maximum of many other spherical harmonics. Approximants that take into account modes for $l, |m| > 2$ are said to include higher order modes.

The most accurate way of obtaining waveforms is through numerical relativity [13–17, 46, 47], where the Einstein field equations are solved numerically. Although precise, these methods are very computationally expensive. Approximants arise through the need to obtain waveforms that approximate these solutions at a reasonable accuracy, but with far greater speed. As a calibration benchmark therefore, numerical relativity remains an important part of the study of approximants. In practice, different families of waveform approximants exist, such as the Taylor family [48, 49] for post-Newtonian approximations, the effective-one-body (EOB) family [50, 51] that reduces two-body dynamics to the case of a corresponding single body, and the phenomenological family [20, 44, 52], with the latter being the fastest as they are designed specifically to be faster than the EOB family while retaining similar accuracy. A sophisticated example from the phenomenological family is the frequency-domain approximant IMRPhenomXPHM [20], which can compute the waveforms of the precessing system with the inclusion

of higher order modes. Internally, this approximant fits behavior to individual frequencies before constructing a precessing waveform through modulation by the twisting up procedure [53]. In this paper IMRPhenomXPHM is used as the approximant of choice, as it is fast, commonly used online, and incorporates both precession and higher order modes. Note, however, that the methods presented in this paper are agnostic to this choice and are applicable to all approximants.

As will be demonstrated in this paper using IMRPhenomXPHM for $A_{\mathbb{F}}$, obtaining A_F for a subset $F \subset \mathbb{F}$ and interpolating A_F to A_F^* on \mathbb{F} , where the asterisk indicates interpolation, is faster on a per-frequency basis than solving for all frequencies in \mathbb{F} , yet still accurate. The subset F will be referred to as the frequency array, and two coarse examples of such frequency arrays for a damped sinusoid are shown in Fig. 1. The convention of asterisks indicating linear interpolations will be upheld. Notation will however be shortened, with the exact waveform $A_{\mathbb{F}}(\theta)$ being denoted by $h(\theta)$, and the approximate waveform $A_F^*(\theta)$ by $h_F^*(\theta)$. If the choices of θ and F are implicit, the notation can be further reduced to h and h^* , respectively.

Since this speedup is obtained through interpolation and not algorithmically, h and h^* cannot be expected to be equal, and a loss in accuracy will be incurred. Two metrics will be used to measure this loss. The first is the residual sum of squares (RSS) for the real and imaginary parts of h and h^* . This metric is used by the evolutionary algorithm internally. The second metric is based on a widely used quantity, the signal-to-noise ratio (SNR).

The concept of SNR comes from signal processing, and is defined for a given power spectral density (PSD) as the normalization with respect to u of the inner product:

$$\langle u, v \rangle = \frac{4}{T} \sum_{f \in \mathbb{F}} \frac{u(f) \overline{v(f)}}{\text{PSD}(f)}, \quad (3)$$

where T is the duration of the waveform, the bar denotes the complex conjugate, and for the applications in this paper,

the PSD describes the sensitivity of the detector at different frequencies. The value of $\langle u, v \rangle$ can be interpreted as a quantization of the presence of u in v , as a ratio of signal to noise.

III. EVOLUTIONARY ALGORITHMS

Evolutionary algorithms [25] are inspired by biological evolution, where natural selection leads populations to evolve to adapt to their environment. As such, this class of algorithms is used as approximation schemes for difficult optimization problems. As in genetics, data is represented internally by genotypes that contain the hereditary information of an organism, whereas the phenotype refers to the externally observable characteristics. In practice this means a solution is encoded in a genotype, with the full solution phenotype being constructed from the genotype. Besides an injective mapping from genotypes and phenotypes, an evolutionary algorithm is governed by the following set of functions.

- (1) an objective function or fitness evaluation function $O: G \rightarrow \mathbb{R}$ that assigns a genotype a fitness score measuring performance;
- (2) a (stochastic) parent selection function P that draws two parents from the set of genotypes; and
- (3) a recombination function $R: G \times G \rightarrow G$ that combines two genotypes to form a new genotype.

Note that the recombination function can include stochastic mutations of the genotype, useful to add small perturbations that might nudge a solution from a non-optimal potential well in the objective field. The steps in an evolutionary algorithm are then summarized as follows.

- (1) A first generation of solutions is instantiated.
- (2) Using O , each genotype is assigned a fitness.
- (3) A new generation is populated through repeated calls to the composition function $R \circ P$.
- (4) The above steps are repeated until a set of stopping conditions has been met.

Examples of stopping conditions include reaching a preset number of generations, or the stalling of variation and improvement within the generations.

In general, evolutionary algorithms offer a method of exploring a solution space where the solutions can be complicated objects. Under the right assumptions, each subsequent generation of solutions is expected to improve on the previous generation, leading to well-performing solutions in practice, even if the determination of a global optimum is not guaranteed. A famous example is the evolved antenna by NASA, designed in automation for the detection of unusual radiation patterns [54].

Few theoretical results that explain the effectivity and convergence of evolutionary algorithms are known. However, a notable theorem often cited is the Schema theorem [55] that states better adapted solutions are likely to drive up the average fitness of successive populations.

IV. METHODOLOGY

With the optimization problem and solution method defined, this section discusses the methodology, split into three subsections. The preprocessing section discusses all considerations that precede the introduction of the evolutionary algorithm to the problem. Following this, the hyperparameters that define the evolutionary algorithm for the established problem of finding the optimal frequency arrays are discussed in the second subsection, with the third subsection offering the performance measures used to evaluate the proposed solutions. Three such solutions will be considered.

In typical applications the frequency spectrum \mathbb{F} is a uniformly spaced set, and in the tests presented here the set is defined by upper and lower limits of 0 and 2048 Hz, respectively, with a sampling frequency of 4 kHz and a duration of 16 seconds.

A. Preprocessing

In order to obtain a faster rate of convergence for the evolutionary algorithm, reference events are used to populate initial frequency arrays during preprocessing. An event is defined here as a fixed set of parameters, and in the case of a cataloged candidate, the parameters used are retrieved from the parameter estimation data releases [2–5]. Three such reference events are tested. First, GW150914 [1], as it was a landmark detection without additional complexities. Second and third are GW190412 [56] and GW190814 [57] since there is evidence for the contribution of higher order modes in these events, as well as precession [58,59]. For this proof of concept, only the plus-polarized components are used.

Starting from the parameters of a reference event, the parameters other than the chirp mass are fixed, and the chirp mass is varied by ± 5 solar masses to generate a number of different waveforms on \mathbb{F} . The choice to vary the chirp mass was made as it induces a large variation in the phase of the resultant waveforms [60]. Each of the waveforms are normalized in amplitude to $[-1, 1]$, and the turning points in the frequency domain are added to a set B called the background or prior. Every frequency array for this event will include this ever-present background of frequencies. This is done under the assumption that the evolutionary algorithm finding and including turning points is inevitable, since a piecewise linear interpolation between the turning points of a function generally gives a very well performing approximation to the full function. The inclusion of this prior information in preprocessing greatly reduces the work the evolutionary algorithm would have to perform in identifying these frequencies during a random walk. Choosing the right number of waveforms, or equivalently the permitted cardinality of B , is an example of bias-variance tradeoff [61]. For the tests described here, five waveforms were used, including the reference event.

TABLE I. The collection of hyperparameters that defines the evolutionary algorithm. The value column contains the values for these parameters used for the realization of the algorithm.

Parameter	Type	Description	Value
δ	Float	Confidence	0
w	Float	Penalty weight	8
g	Integer	Population size	6000
d	Boolean	Algorithm direction	True
s	Integer	Population sample size	6
r	Integer	Number of generations	30

Finally, all frequencies under 20 Hz were removed as candidates for inclusion in a genotype. The result of these preprocessing steps is one set of frequencies (the background) per event that will be present in all frequency arrays learned on the corresponding event, as well as one permissible set of frequencies to include in the genotypes for each.

B. Hyperparameters

Given the preprocessed data, the hyperparameters uniquely define the evolutionary algorithm up to stochasticity. In this subsection the different hyperparameters are treated, guided by the context wherein they arise. The reader is recommended to look ahead at Table I before reading this subsection in order to get an overview of how this subsection is structured.

Recall that for a set S and a subset $T \subset S$, an indicator vector I_T is a binary vector indexed by the elements of S such that $I_T(s) = 1$ if and only if $s \in T$. Genotypes, as subsets of the frequency spectrum, are encoded in indicator vectors over \mathbb{F} , and a phenotype is constructed from such a genotype by adding the indicated frequencies to the frequencies in the background B . This map therefore realizes a frequency array as a phenotype. Additionally, a parameter $\delta \in \mathbb{R}_{\geq 0}$ representing the confidence of the algorithm can be set: if a frequency f is included in the genotype, then the elements of the set,

$$\{f' \in \mathbb{F} : |f - f'| \leq \delta\}, \quad (4)$$

will be included in the phenotype. Note that this means that with $\delta = 0$ only the frequency f itself is included in the phenotype, therefore reducing to the special case described above. The phenotype (or frequency array) construction process for different values of δ , with B set to the empty set for simplicity, is shown in Fig. 2.

For a genotype G (with corresponding phenotype F) and a reference event waveform h , the objective function is defined as

$$O(G) = \sum_{f \in \mathbb{F}-F} |h(f) - h_F^*(f)| + w \cdot \#F, \quad (5)$$

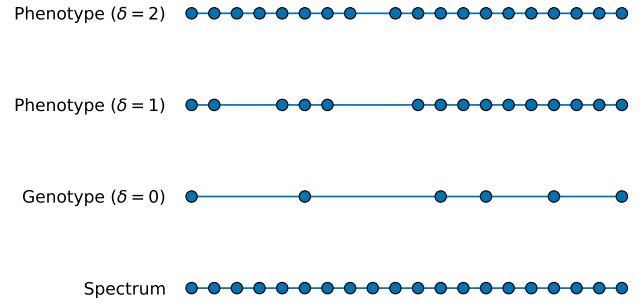


FIG. 2. An example genotype and corresponding phenotypes for a spectrum of subsequent natural numbers with empty background. For $\delta = 0$ the genotype and phenotype coincide. With $\delta = 1$ nearby frequencies are added to the genotype to form the phenotype, and with $\delta = 2$ additional nearby frequencies are added. This figure shows the hyperparameter δ controls the confidence with which the algorithm includes frequencies in the solution phenotypes.

where $|\cdot|$ denotes the modulus and the hash denotes set cardinality. The sum of moduli amounts to the residual sum of squares of the exact and approximate waveforms for both the real and complex parts, as can be seen from expanding the modulus for a single frequency. This sum can therefore also be expressed as

$$\text{RSS}(\text{Re}[h], \text{Re}[h_F^*]) + \text{RSS}(\text{Im}[h], \text{Im}[h_F^*]). \quad (6)$$

An additive penalty, weighed by a variable w , is added for the phenotype size. One should take note of the range for the summation in the objective function defined in Eq. (5): for all frequencies in F the exact waveform coincides with the approximate waveform, meaning these terms reduce to zero, and therefore do not have to be computed.

The complexity of the solution space, and therefore the need for optimization methods like evolutionary algorithms, is demonstrated through the objective function. A first approach to optimization could be by way of a greedy algorithm [62], where at every iteration the frequency f' corresponding to the largest residual, or equivalently the largest term in the sum, is added to the phenotype. However, due to the addition of this frequency, the approximate waveform is changed to

$$h_{F \cup \{f'\}}^*, \quad (7)$$

and the terms of the sum in the objective function will have to be recomputed. At the same time the cross section of the (very high-dimensional) solution space containing solutions without the frequency f' is pruned. There is no guarantee that the frequency f' that is now included is part of the optimal solution as this objective function has no optimal substructure [62], and no choice of frequency is ever reconsidered. It follows that the greedy

algorithm would now be barred from finding a global optimum, whereas an evolutionary algorithm could rectify this misstep.

If a phenotype frequency array is used for the computation of the SNR, the error and rate of convergence are quantifiable, and minimizing the sum of moduli in the objective function is equivalent to minimization of the error term. Assume D is an arbitrary but fixed strain of data, and h is a waveform. It then follows from the linearity of the inner product that the error $E(h)$ for this waveform is given by

$$E(h) = \langle h, D \rangle - \langle h^*, D \rangle = \langle h - h^*, D \rangle. \quad (8)$$

Since taking the limit of an inner product is interchangeable with taking the limit of an argument, and the continuous inner product will tend to 0 if either argument tends to 0, it is given that $E(h) \rightarrow 0$ if $h - h^* \rightarrow 0$. The latter is exactly the case if $|h - h_F^*| \rightarrow 0$ in the L^1 norm.

Once the objective function has been used to evaluate the fitness of the genotypes in a generation, a new generation of offspring can be formed. The parent selection function P is implemented according to fitness proportionate selection with elitism [63,64]. Elitism refers to the carrying over of the top percentage of genotypes (by fitness ranking) to the next generation, and fitness proportionate selection means that the higher a genotype is ranked, the higher the probability will be that it is drawn as a parent by P . The probability measure so constructed on the genotypes sorted by ascending fitness is given by $\mathbb{P}(i) = i \cdot p$ with

$$p = \frac{1}{\sum_{j=1}^g j} = \frac{2}{g(g+1)} \quad (9)$$

for the candidate ranked at position i , and g the fixed population size of a generation. The use of this measure ensures that the lowest ranked genotype (at $i = 1$) has the lowest probability of being chosen as a parent, with every higher ranked genotype having a proportionally higher probability of being selected. The second equality in Eq. (9) follows by substituting in Gauss' summation formula. Note too the importance of the generation size, as for larger values the probabilities will be more evenly distributed over the candidates.

Finally, the recombination function R is defined as the sum of the two genotypes, assuming the genotypes to be a Boolean algebra. This means that a frequency will be present in the offspring if and only if that frequency is present in at least one of the parent genotypes.

Because of a symmetry argument applied to the phenotype construction process, the algorithm can be ran both forwards and backwards. In the forwards case, the phenotype grows in size towards the full frequency spectrum as frequencies are added in each iteration of the algorithm. In the backwards case, the frequency spectrum is reduced in

size towards the background by removing frequencies in each iteration. As a rule of thumb, the forwards case is better suited to finding smaller frequency arrays, whereas the backwards case is better suited to finding frequency arrays that result in a low error. This is a result of the stopping conditions typically being met before the frequency array reaches half the size of the frequency spectrum. Note that the parameter that dictates the direction stands in direct relation to the penalty w included in Eq. (5). In both cases the algorithm needs to be incentivized to move through the solution space by the penalty. In, for instance, the backwards direction, if w were to be set to 0, there would be no reason to remove any frequency from the phenotype, as doing so would only lead to a loss in accuracy. Setting $w > 0$, the removal of this same frequency might result in an increased fitness, as long as the loss in accuracy is less than the penalty w for including the frequency.

The evolutionary algorithm that has now been established can only be run once an initial generation is given. Initialization of this first generation is done at random through a uniform distribution on the set of allowed frequencies. A generation of genotypes is formed by drawing a sample of frequencies of size s from this distribution, repeating this until g such genotypes have been drawn to form the full first generation. This marks the frequency sample size s as a hyperparameter. Note that the number of possible genotypes that can be drawn is given by a binomial coefficient, and due to the factorial in this coefficient this number will be very large. In order to explore a large section of the solution space, the sample size s needs to be chosen proportional to both g and the maximum run-time r of the algorithm expressed in the number of generations. In the k th generation, a maximum in the order of magnitude of $s \cdot 2^k$ frequencies is included in the genotype. Using this information an appropriate combination of values for s , g , and r can be set.

Recapitulating this subsection, besides the initial choices of background size during preprocessing and the probability measure for the parent selection function, the algorithm is defined by the hyperparameters (δ, w, g, d, s, r) given in Table I. The values in this table represent the hyperparameters used in this proof of concept, with True for the direction d meaning the algorithm was ran backwards. This choice was made to prioritize accuracy, which also inspired the choice of δ . The run-time r was set to well above the observed limit of stagnation in trial runs, with the population size g and frequency sample size s empirically determined following the previous discussion. The penalty weight w is set to 8, but any value to counterbalance the increasing modulus in the objective function as the evolutionary algorithm progresses and frequencies are removed would have sufficed. These choices of hyperparameters consistently drive the evolutionary algorithm to converge to similar and well-performing frequency array phenotypes on

repeated runs, suggesting the optimization does not get stuck in easily avoidable local minima. The run-time for the algorithm (including preprocessing) on an Intel Gold 6154 CPU is in the order of magnitudes of minutes, bound by the longest recorded run.

C. Performance measures

In the remainder of this section, two metrics to measure performance will be introduced. The first is a fractional measure of the speedup obtained by using a frequency array phenotype output by the evolutionary algorithm. The second measures the loss in accuracy for a frequency array, where the exact waveform was injected into LIGO Livingston detector noise before the SNR was computed by matched filtering using both the exact and approximate waveforms as templates. The percentual difference in SNR is then used as a measure of loss.

For the definition of the speedup metric, let t_a be the average time in seconds elapsed by the approximant per frequency, and t_i the average time required to interpolate at a frequency in seconds. With $\rho := \#F/\#F$ the relative size of the frequency array, the speedup of F can be computed as

$$S[F] = \frac{\#F \cdot t_a}{\#F \cdot [\rho t_a + (1 - \rho)t_i]}, \quad (10)$$

since the numerator specifies the time if the waveform is computed at every frequency, and the denominator specifies the time if the waveform is computed at a fraction ρ of the frequencies, with interpolation at a fraction $(1 - \rho)$ of the frequencies. This expression reduces to

$$S[F] = \frac{1}{(1 - \rho) \cdot t_i/t_a + \rho}. \quad (11)$$

Note that this quantity is dimensionless, measuring the speedup as a ratio in an isolated section of a procedure where the waveform is generated, without depending on the machine. In order to obtain accurate values for t_a and t_i , 100 waveforms were generated on the full spectrum using Bilby [65], and the same 100 waveforms were interpolated on the frequency array phenotype. For every frequency array the parameters of the corresponding reference event were used without loss of generality since only time is measured, with the primary mass perturbed by ± 1 at every iteration to ensure no values were cached. The times per frequency were then averaged out for each event and type, repeating and averaging this process 100 times to remove possible inconsistencies.

Finding the neighborhood in the parameter space \mathcal{P} where the frequency arrays are effective is important, as the reference event input to the algorithm during preprocessing adds a level of specificity. Nearby the reference events the frequency arrays are likely to allow a more accurate

reconstruction of the SNR. In order to analyze the generalizability, 10,000 points are sampled from the default preprocessing binary black hole prior in Bilby version 2.2.2. At every sampled set of parameters, the exact plus polarization of the waveform is generated and injected into simulated noise generated using the LIGO Livingston PSD included in Bilby version 2.2.2. The SNR of both the exact plus polarization, hereafter also called the exact SNR, and the approximate plus polarization, hereafter called the approximate SNR or SNR^* , is computed against this strain. The percentual error for the frequency array F is then calculated as

$$E_{\%}[F] = \frac{|\text{SNR} - \text{SNR}^*|}{\text{SNR}} \times 100. \quad (12)$$

This metric is computed as a percentage to account for the different values of the exact SNR. A given absolute error might be negligible for large SNR values, whereas it might be detrimental to a low value SNR reconstruction. Computing the error as a percentage removes this consideration. A threshold of 0.01 is set for $E_{\%}[F]$, and a parameter sample is said to be rejected if the percentual error either attains or exceeds this threshold. Considering the uncertainties in SNR values in the GWTC-3 catalog [4] this threshold is particularly strict. This threshold is used as an aid to quantize results and is inconsequential to their objective interpretation.

V. RESULTS

In this section the results from running the evolutionary algorithm with the three reference events GW150914, GW190412, and GW190814 are presented. The three frequency arrays output by the evolutionary algorithm are analyzed in terms of speedup, accuracy, and their generalizability before selecting one of the three based on how well these three qualities are balanced. This frequency array is used for additional tests.

The results for the reference events are shown in Figs. 3–5, respectively, where both constituents of the objective function are shown in a visually separable way. The waveform error in these figures refers to the residual sum of squares, normalized to $[0, 100]$ so it is more easily compared to the percentual array size at this same scale. Normalization here means the minimum achieved waveform error is mapped to 0, and the maximum achieved waveform error is mapped to 100. This means that the visualized waveform errors give little qualitative information and cannot be compared across figures, because the original scales are not shown. The benefit, however, is that this allows for a visual rule to select the stopping time, and one that is in fact automatable: halting at the first sign of a significant jump in error. At this time the evolutionary algorithm is stopped and the frequency array with the highest fitness score is chosen as the solution.

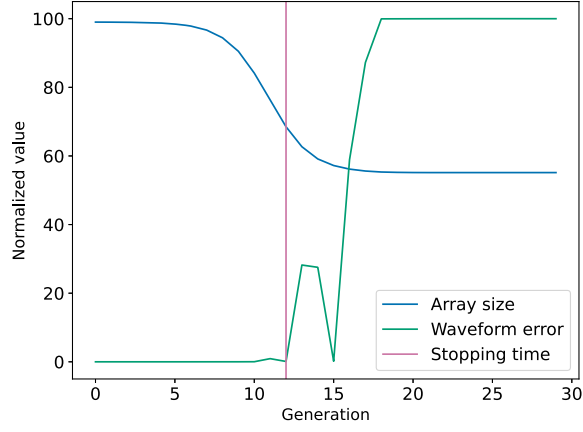


FIG. 3. A visualization of the two parts that make up the objective function for GW150914, along with the stopping time for this event (generation 12). The waveform error is normalized to $[0, 100]$ so that the graphs can be compared at the same scale.

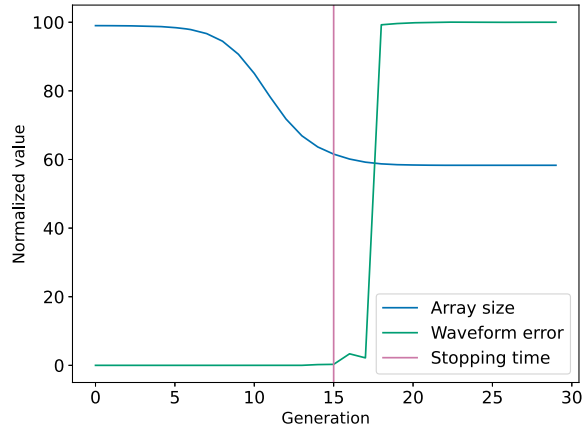


FIG. 4. A visualization of the two parts that make up the objective function for GW190412, along with the stopping time for this event (generation 15). The waveform error is normalized to $[0, 100]$ so that the graphs can be compared at the same scale.

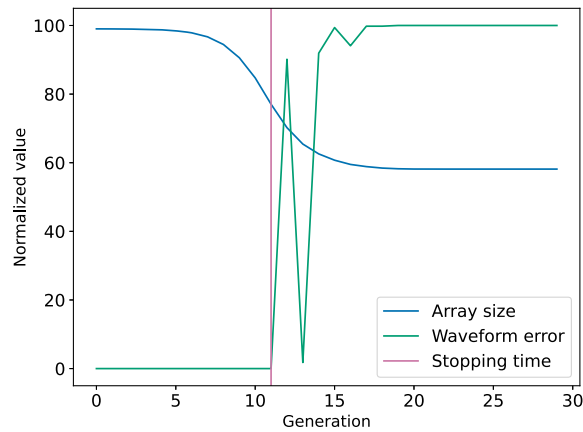


FIG. 5. A visualization of the two parts that make up the objective function for GW190814, along with the stopping time for this event (generation 11). The waveform error is normalized to $[0, 100]$ so that the graphs can be compared at the same scale.

At first glance, the nonmonotonicity of the waveform error might seem conflicting with the assumption that every subsequent generation contains better adapted solutions. However, one should keep in mind that this statement is made with respect to the objective function, where the waveform error is combined with a size penalty, and that the objective function will show monotonicity. Overall, the fact that the waveform error is only relative is underlined by the difference between the exact and approximate waveform visualized in Fig. 6. Despite the dramatic jump in waveform error seen in Fig. 4, the difference $h - h^*$ for the frequency array proposed in the last generation (where the error is maximized) barely registers at the scale of the waveform, being at least an order of magnitude smaller. These magnitudes can be considered as upper bounds for the waveform errors over the run-time of the evolutionary algorithm.

In the current implementation, the error is more pronounced in the phase than it is in amplitude. Depending on the application, the objective function can be modified to either balance this error or favor one of the two.

The typical development of the proposed frequency array for GW150914 over the generations is visualized in Fig. 7. The generations shown were selected to equally space the process before stagnation occurs at generation 22, and includes the ultimately selected candidate from generation 12. Both the stagnation and jump in waveform error that underlies this choice of candidate can be seen in Fig. 3. At every generation, a distribution function over the frequency spectrum is shown. This distribution is computed by dividing the spectrum into bins, counting the frequencies present in each bin, and normalizing the counts to the maximum count over all bins. Mirroring the development of the array size in Fig. 3, few frequencies are removed from the early candidates, with only a minor visible move downwards from the dotted line seen in generation 8. Change accelerates between generations 8 and 12. Note that from the peaks and valleys that persist from this generation onwards, the algorithm has decided at an early stage which frequency ranges are most important. The lower end of the frequencies are kept consistently, although the dropoff towards 250 Hz becomes steeper during the later generations. Interpreting this figure gives an indication of how the evolutionary algorithm has traversed the search space, favoring the lower range of the spectrum, but always locally to the extent that above 250 Hz, the densities look like a diffused horizontal line. This means the evolutionary algorithm does not favor any band of the spectrum as strongly as it does $[0, 250]$ Hz.

The speedup of the frequency arrays is shown in Table II, showing the input variables and values for $S[F]$ as defined in Eq. (11). A detail that should be noted is that there is a natural inverse relation between the relative array size ρ and speedup S . The more frequencies are omitted from the frequency array, the higher the

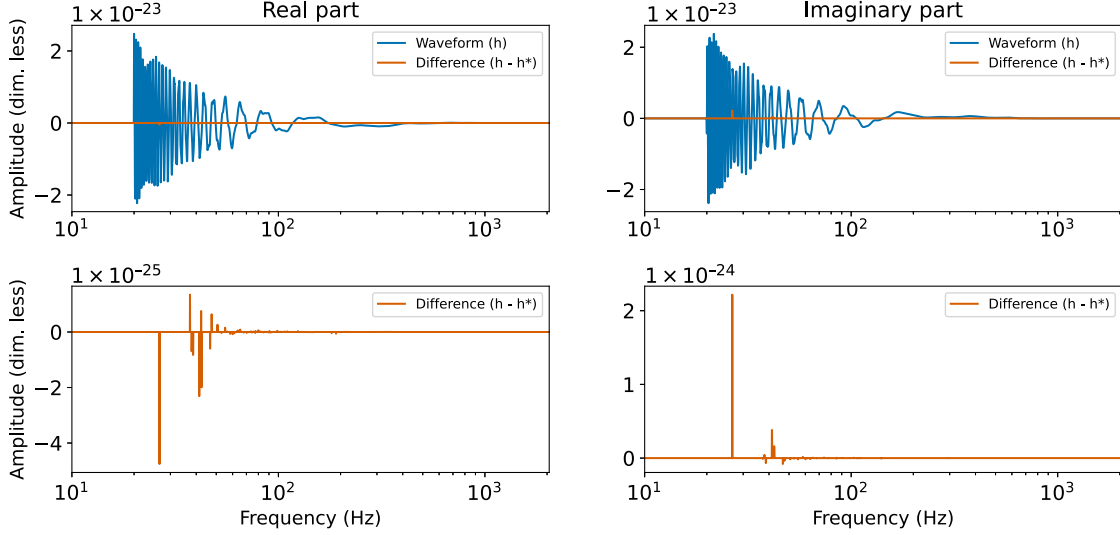


FIG. 6. The real and imaginary parts of the plus-polarized component of GW190412 (h) and the difference with the approximation ($h - h^*$) computed on the frequency array proposed in the last generation (29) of the algorithm. The bottom two panels show the differences on a smaller scale, and the horizontal axes are in log scale.

speedup. However, the values in the time columns differ to an extent. The differences in the t_i column can be explained by the density of the frequencies included in the array. The time measurements given in this column correspond to the average time elapsed interpolating at a frequency that is not included in the array. If many subsequent frequencies are omitted, the value at these frequencies can be interpolated at once from a single linear segment, whereas the same number of values for nonsubsequent frequencies need to be interpolated using several such segments. The computation times scale

linearly with the number of segments that are computed and will therefore be higher in the latter scenario. For the t_a column, the time difference between frequency arrays is likely to stem from the additional effects added to the waveforms, such as twisting up.

As a note on data compression and memory usage, before interpolating to the full frequency spectrum, one can opt to write the waveform to disk, indexed by the frequency array. By choosing to interpolate to the full spectrum later, waveforms can be stored in a compressed format, and the compression factor will equal ρ .

The results of the generalizability tests to determine the regions of the parameter space \mathcal{P} where the frequency arrays are effective are shown in Table III. For each of the 15 parameters (minus the geocentric time) the sampled ranges are shown, as well as the ranges that were recovered for the frequency arrays. Recovered here means that the parameter samples from these ranges were accepted by their error being below the given threshold, with

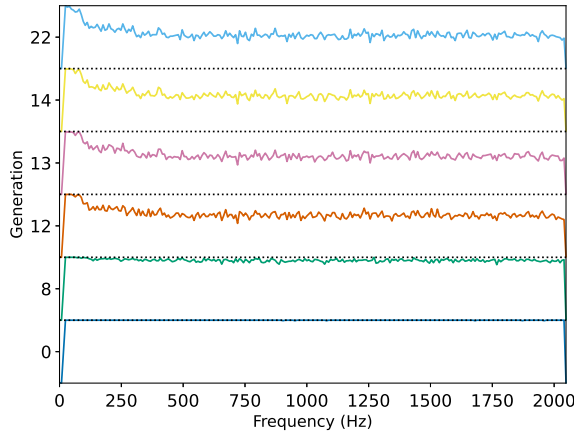


FIG. 7. The evolution of the proposed frequency arrays for GW150914 visualized for a selection of generations. The densities are computed by binning the frequency spectrum, counting frequencies, and normalizing to the maximum count. The evolution can be seen to follow the trend in the array size seen in Fig. 3, and shows what frequencies are removed during the run-time of the evolutionary algorithm. Notably, the frequencies between 0 and 250 Hz persist the longest.

TABLE II. Measurements of the values required for the computation of the speedup $S[F]$ given in Eq. (11), along with the computed speedup. These values are ρ , the size of the frequency array relative to the spectrum, t_a , the average time in seconds spent computing the waveform at a frequency by the approximant, and t_i , the average time in seconds spent to interpolate the waveform at a frequency. All values are rounded to two decimals for readability.

	ρ	t_a	t_i	S
GW150914	0.69	5.52×10^{-7}	6.95×10^{-12}	1.46
GW190412	0.62	8.48×10^{-7}	7.66×10^{-12}	1.62
GW190814	0.77	1.71×10^{-6}	5.90×10^{-12}	1.30

TABLE III. The ranges for each of the variables in the parameter space and the recovered ranges per frequency array, in rounded numerical format. If no unit is included, the quantity is dimensionless. Note that the geocentric time is not included, as it is inconsequential to this test. All parameter ranges save for the luminosity distance and chirp mass were recovered, with both of these not being fully covered by the drawn sample of parameters.

	Range	GW150914	GW190412	GW190814
Mass ratio	[0.125, 1]	[0.125, 1]	[0.125, 1]	[0.126, 1]
Chirp mass (M_{\odot})	[25, 100]	[25.005, 86.922]	[25.007, 86.855]	[25, 86.509]
Luminosity distance (Mpc)	[100, 5000]	[101.091, 4999.683]	[170.862, 4999.658]	[128.847, 4999.707]
Declination (Rad)	$[-1.571, 1.571]$	$[-1.552, 1.557]$	$[-1.553, 1.553]$	$[-1.545, 1.559]$
Right ascension (Rad)	[0, 6.283]	[0, 6.282]	[0, 6.281]	[0, 6.283]
Orbital inclination (Rad)	[0, 3.142]	[0.003, 3.114]	[0.005, 3.138]	[0.026, 3.121]
Polarization angle (Rad)	[0, 3.142]	[0, 3.141]	[0, 3.141]	[0.001, 3.141]
Coalescence phase (Rad)	[0, 6.283]	[0.001, 6.283]	[0, 6.283]	[0, 6.282]
Spin of primary mass	[0, 0.99]	[0, 0.99]	[0, 0.99]	[0, 0.99]
Spin of secondary mass	[0, 0.99]	[0, 0.99]	[0, 0.99]	[0, 0.99]
Tilt of primary mass (Rad)	[0, 3.142]	[0.014, 3.136]	[0.023, 3.126]	[0.018, 3.1]
Tilt of secondary mass (Rad)	[0, 3.142]	[0.004, 3.13]	[0.026, 3.127]	[0.017, 3.125]
Precession angle (Rad)	[0, 6.283]	[0, 6.283]	[0.003, 6.283]	[0, 6.283]
Azimuthal angle (Rad)	[0, 6.283]	[0, 6.282]	[0, 6.282]	[0, 6.283]

$E_{\%}[F] < 0.01$. For most parameters, the full ranges were recovered, save for the luminosity distance and chirp mass. In the case of the luminosity distance, further inspection reveals that the lower ranges were simply not reached in the parameter sample drawn. The same holds for the chirp masses due to the way Bilby samples the chirp mass through uniform sampling in the component masses. If this is manually changed the results might vary slightly.

Additional statistics on the generalizability tests are shown in Table IV. A number of parameter sample points were rejected. However, the errors for these samples were evaluated at a strict error threshold. Keeping in mind the generalizability, these results can be considered remarkable, as the error is bounded above by 2.87%. The distribution of the errors is further visualized in Fig. 8. This plot shows that for all three frequency arrays the percentual errors in SNR are mostly concentrated below the error threshold of $10^{-2} = 0.01$.

Investigating the cause for the rejection of parameter samples, it becomes apparent that nearly all rejected samples have a low exact SNR value for h . This is visualized in Fig. 9, showing the exact SNR distributions

for the rejected parameter samples versus the exact SNR distributions for the full samples. It can be seen from this figure that for the rejected parameter samples, the SNR is on average an order of magnitude smaller than that of the full parameter sample, and moreover that the distributions are left-tail heavy. In order to verify the strength of the SNR as a predictor for rejection, logistic regression [61] of the binary acceptance class onto the waveform parameters was employed. The sampled parameters for the three candidates were stacked, with the luminosity distance replaced by the SNR. The removal of the luminosity distance is a requirement so that the multicollinearity assumption of logistic regression is not violated, as luminosity distance correlates strongly with SNR. Despite the SNR scaling with several

TABLE IV. The extrema for the percentual errors $E_{\%}[F]$ on the parameter sample, along with the number of parameter samples rejected, as for these rejected samples $E_{\%}[F] \geq 0.01$.

	Minimum	Maximum	Rejected
GW150914	2.69×10^{-10}	0.50	909
GW190412	1.12×10^{-7}	2.87	902
GW190814	1.47×10^{-9}	0.21	203

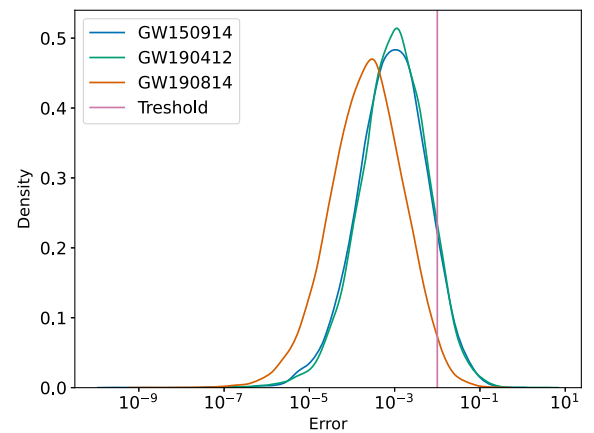


FIG. 8. Densities of the percentual error in SNR $E_{\%}[F]$ for the full parameter samples, along with the threshold for acceptance. The horizontal axis is in log scale.

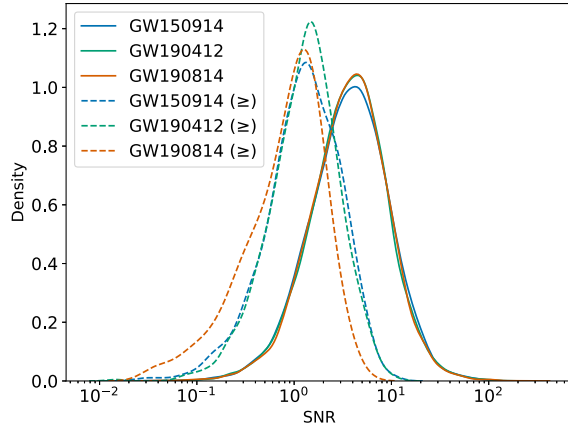


FIG. 9. The exact SNR distributions for the full parameter samples (solid) and the rejected parameter sample points that exceed (\geq) the threshold of 0.01 set on $E_{\%}[F]$ (dashed). The horizontal axis is in log scale.

other parameters, the removal of the luminosity distance should sufficiently reduce the correlation. A binary labeling was created for this dataset by assigning a dependent variable 1 to the accepted parameter samples and 0 to the rejected samples. Fitting a logistic regression model, the SNR was found to be the predictor with the highest coefficient by far, with a value of $\beta_{\text{SNR}} = 16.5$. This coefficient was followed by the chirp mass with a value of $\beta_{\text{Chirp}} = 3.74$, meaning the SNR is a significantly stronger predictor for the rejection of a parameter sample according to the regression model. As with accuracy, the generalizability too is bounded from below by the results presented.

Based on the analysis in this section, the frequency array obtained by using GW150914 as the reference event is proposed as a proof of concept, as it obtains a speedup in computational time of 46% at a maximum error of 0.5% in SNR. A heatmap for the generalizability of this frequency array is shown in Fig. 10. This figure was created by drawing ten partial parameter samples, meaning every parameter except for the mass ratio and chirp mass were sampled, and then varying the mass ratio and chirp mass before computing the percentual error in SNR $E_{\%}$ for every such combination. A point in the plane then corresponds to the average error of the ten parameter samples for that combination of the chirp mass and mass ratio, and the full figure shows 100,000 such parameter samples in total. It can be seen from the figure that, especially in the band of the chirp mass where the event belongs to, moving away from the mass ratio of the event will increase the error.

As an additional test for the applicability of the proof of concept, the maxima of the log likelihoods [66] that quantify the likelihood of waveforms being present in detector strains were approximated using the frequency array. Using the inner product defined in Eq. (3), the log likelihood can be defined as

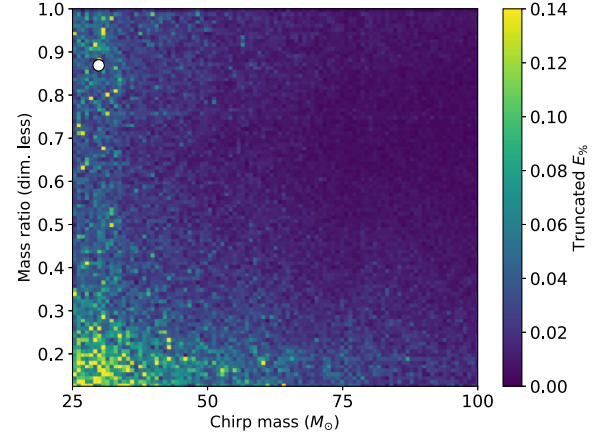


FIG. 10. A heatmap for the error in SNR $E_{\%}$ based on partial parameter samples when using the proposed frequency array obtained using GW150914 as the reference event. A partial parameter sample means that a sample was drawn from the parameter space without sampling the chirp mass or mass ratio. For every such sample, the chirp mass and mass ratio were varied within the ranges of [25, 100] and [0.125, 1], respectively. The average errors for ten such samples are shown in this figure, truncated at 0.14% for increased visibility. The mass ratio and chirp mass of GW150914 are indicated by a white marker.

$$\mathcal{L}(h, D) = -\frac{1}{2} \langle D - h, D - h \rangle \quad (13)$$

for h the waveform projected onto the detector frame and D the strain. Setting $D := N + h'$ for N noise, the maximum of the log likelihood should be attained at the injected waveform h' . Analogous to the SNR reconstruction, the exact log likelihood maximum $\mathcal{L}(h, N + h)$, or \mathcal{L}_m for shorthand, and approximate log likelihood maximum $\mathcal{L}(h^*, N + h)$, or \mathcal{L}_m^* , were computed for a parameter sample of size 10,000. Here, however, both polarizations are included. Note that these errors do not need to be converted to percentages as was done for the errors in SNR, since the variance in the values will be significantly smaller. The absolute error distribution is shown in Fig. 11 and shows promising results. The reader should be aware, however, that the evolutionary algorithm was not fine-tuned to this computation, suggesting that a higher accuracy can likely be obtained. Specifically, the objective function (that defines the optimization problem) should be adapted to the log likelihood function, and the hyper-parameters should be rebalanced.

Further study in general might reveal marginally better performing frequency arrays, or frequency arrays better suited to specific sections of the parameter space. It is, however, expected that the improvement from frequency subsampling and linear interpolation represented by the values in this section is near optimal. Implementing the naive frequency subsampling scheme of interpolating every second frequency (setting the relative array size at $\rho = 0.5$) for GW150914, the speedup S evaluates to 1.99, but the

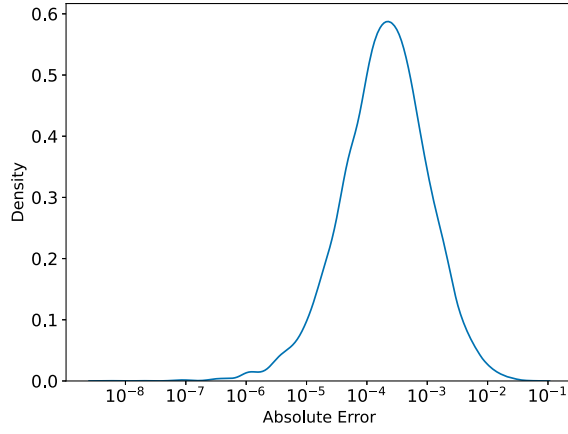


FIG. 11. The distribution of the absolute error $|\mathcal{L}_m - \mathcal{L}_m^*|$ between the exact log likelihood maximum \mathcal{L}_m computed using the exact waveform and the approximate log likelihood maximum \mathcal{L}_m^* computed using the approximate waveform. This distribution was constructed on a sample of size 10,000. The horizontal axis is in log scale.

minimum error of $E_\%$ blows up by a factor with order of magnitude 10^4 , in the process causing a rejection for 5972 parameter samples.

It follows from the results in this section that as expected, for a given threshold, smaller frequency arrays are effective on smaller neighborhoods of the parameter space \mathcal{P} . Under the assumption of continuity, parameter instances nearby the reference event parameters will give better performance because of their similarity. If evolutionary algorithms are to be used to speed up computations on \mathcal{P} with a given maximum allowable error, a single frequency array might not suffice. One option would then be to assign different frequency arrays to different regions of \mathcal{P} . In order to do this efficiently, the effective neighborhood of an array will have to be computed. A first proposition could be akin to the generalizability test used in this paper. After setting a threshold, recover the ranges in which all parameter samples were accepted, and take the convex hull of these ranges in \mathcal{P} . There exist efficient algorithms for the computation of convex hulls in higher dimensions under the assumption of flat space [67] that enable this procedure. Finding an optimal coverage of \mathcal{P} by the resulting polytopes, however, reduces to an instance of the sphere packing problem through a topological argument, which depending on specifics is known to be NP-complete (non-deterministic polynomial-time complete) [62,68]. This obstacle can be overcome at the cost of efficiency if overlap of the polytopes is allowed and a global optimum is not required. Furthermore, for sufficiently large polytopes, the largest embeddable hypercube can be determined and used as an approximation to the polytope. These cubes in turn can be used to tile \mathcal{P} more easily due to their shape. If such a tiling of \mathcal{P} is generated, the speedups demonstrated in this section can serve as lower bounds for the speedups achievable on the entirety of \mathcal{P} .

VI. CONCLUSIONS

Evolutionary algorithms were introduced as a method for the optimization problem of subsampling frequency spectra to speedup waveform approximants, and per extension, the computation of SNR. It was shown that even with strict bounds on the percentual errors, speedups of at least 30% in computational time are achievable on nearly the entire sampled parameter ranges. If the error bound is heightened to at most 2.87%, the frequency arrays generalize to the full parameter ranges (specified in Table III). The margins of the parameter ranges where the frequency arrays were not effective were explored, as well as the possibility of tiling the parameter space to force a set upper bound on the error. It was demonstrated that the ranges where the frequency arrays were ineffective are mostly determined through correlation with the SNR, and it could therefore be argued that the frequency arrays generalize to the full ranges depending on the error metric in use. The proposed proof of concept frequency array, obtained by using GW150914 as a reference event, shows a speedup of 46% with a maximum error of 0.5% on the full parameter ranges. Additionally, it was shown that this frequency array performs exceptionally well for the reconstruction of the maxima of log likelihood functions.

Possible future lines of work include improving the methods presented in this paper specifically for parameter estimation [69] or time domain problems such as the fast computation of Euler angles that describe the evolution of precessing binary black hole systems [20,37,51,70]. Alternatively, the tiling of the parameter space, and simultaneously the optimum in trade-off between accuracy and generalizability, can be studied further. The evolutionary algorithm can in theory also be calibrated to perform on targeted sections of the parameter space, such as those corresponding to intermediate-mass black hole binaries or extreme mass ratio inspirals. Within these sections a single frequency array might suffice.

Further improvements can be made to the evolutionary algorithm, for instance by incorporating more prior information. The algorithm can be discouraged from traversing sections of the solution space, for example by using a supplied power spectral density, or by more hyperparameter tuning. Currently the algorithm is hardcoded to run for a total of 30 generations. Despite a stopping condition being in place, the algorithm was run for the full duration to study its behavior. By tuning the hyperparameters and designing more advanced stopping conditions, the evolutionary algorithm can be modified to propose frequency arrays meeting a set of required conditions without additional user input. A future milestone is to release the adjustable code as a package.

For the evolutionary algorithm itself, the investigation of its applications to data compression for gravitational-wave data analysis could prove useful. This can be coupled to the exploration of alternative interpolation schemes.

With the number of gravitational-wave detections increasing through the arrival of detector upgrades and

new generations of observatories, the need for faster yet accurate analyses is becoming more apparent. The use of optimization schemes such as evolutionary algorithms to realize these targets is indispensable, and will enable the scientific community to learn more about the nature of gravitational waves in less time.

ACKNOWLEDGMENTS

We thank Harsh Narola, Stefano Schmidt, Melissa Lopez, Justin Janquart, Bhooshan Gadre, Marc van der Sluys, Chris

van den Broeck, Justin Perez, and the anonymous referee. Q.M. is supported by the research program of the Netherlands Organization for Scientific Research (NWO). S.C. is supported by the National Science Foundation under Grant No. PHY-2309332. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by the National Science Foundation Grants No. PHY-0757058 and No. PHY-0823459. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

-
- [1] LIGO Scientific and Virgo Collaborations, Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* **116**, 061102 (2016).
 - [2] LIGO Scientific and Virgo Collaborations, GWTC-1: A gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs, *Phys. Rev. X* **9**, 031040 (2019).
 - [3] LIGO Scientific and Virgo Collaborations, GWTC-2: Compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, *Phys. Rev. X* **11**, 021053 (2021).
 - [4] LIGO Scientific and Virgo Collaborations, GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, *Phys. Rev. X* **13**, 041039 (2023).
 - [5] LIGO Scientific and Virgo Collaborations, GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, *Phys. Rev. D* **109**, 022001 (2024).
 - [6] C. W. Helstrom, D. W. Fry, L. Costrell, and K. Kandiah., *Statistical Theory of Signal Detection*, International Series of Monographs on Electronics and Instrumentation (Elsevier Science, 2013).
 - [7] B. S. Sathyaprakash and S. V. Dhurandhar, Choice of filters for the detection of gravitational waves from coalescing binaries, *Phys. Rev. D* **44**, 3819 (1991).
 - [8] A. Buonanno and T. Damour, Effective one-body approach to general relativistic two-body dynamics, *Phys. Rev. D* **59**, 084006 (1999).
 - [9] Yi Pan, Alessandra Buonanno, Andrea Taracchini, Lawrence E. Kidder, Abdul H. Mroué, Harald P. Pfeiffer, Mark A. Scheel, and Béla Szilágyi, Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism, *Phys. Rev. D* **89**, 084006 (2014).
 - [10] Héctor Estellés, Marta Colleoni, Cecilio García-Quirós, Sascha Husa, David Keitel, Maite Mateu-Lucena, Maria de Lluc Planas, and Antoni Ramos-Buades, New twists in compact binary waveform modeling: A fast time-domain model for precession, *Phys. Rev. D* **105**, 084040 (2022).
 - [11] Sascha Husa, Sebastian Khan, Mark Hannam, Michael Pürrer, Frank Ohme, Xisco Jiménez Forteza, and Alejandro Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal, *Phys. Rev. D* **93**, 044006 (2016).
 - [12] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Pürrer, Xisco Jiménez Forteza, and Alejandro Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era, *Phys. Rev. D* **93**, 044007 (2016).
 - [13] Emmanuel A. Tassone, Paula A. Mandrilli, Carlos N. Kozameh, Gonzalo D. Quiroga, and José I. Nieva, Numerical evolution of the center of mass and angular momentum for binary black holes, *Phys. Rev. D* **104**, 084038 (2021).
 - [14] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, T. Accadia, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and C. Affeldt, The NINJA-2 project: Detecting and characterizing gravitational waveforms modelled using numerical binary black hole simulations, *Classical Quantum Gravity* **31**, 115004 (2014).
 - [15] Karan Jani, James Healy, James A. Clark, Lionel London, Pablo Laguna, and Deirdre Shoemaker, Georgia Tech catalog of gravitational waveforms, *Classical Quantum Gravity* **33**, 204001 (2016).
 - [16] James Healy, Carlos O. Lousto, Yosef Zlochower, and Manuela Campanelli, The RIT binary black hole simulations catalog, *Classical Quantum Gravity* **34**, 224001 (2017).
 - [17] James Healy and Carlos O. Lousto, Fourth RIT binary black hole simulations catalog: Extension to eccentric orbits, *Phys. Rev. D* **105**, 124010 (2022).
 - [18] Stanislav Babak, Andrea Taracchini, and Alessandra Buonanno, Validating the effective-one-body model of spinning, precessing binary black holes against numerical relativity, *Phys. Rev. D* **95**, 024010 (2017).
 - [19] Nelson Christensen and Renate Meyer, Parameter estimation with gravitational waves, *Rev. Mod. Phys.* **94**, 025001 (2022).
 - [20] Geraint Pratten, Cecilio García-Quirós, Marta Colleoni, Antoni Ramos-Buades, Héctor Estellés, Maite Mateu-Lucena, Rafel Jaume, Maria Haney, David Keitel, Jonathan E. Thompson, and Sascha Husa, Computationally

- efficient models for the dominant and subdominant harmonic modes of precessing binary black holes, *Phys. Rev. D* **103**, 104056 (2021).
- [21] I. W. Harry, B. Allen, and B. S. Sathyaprakash, Stochastic template placement algorithm for gravitational wave data analysis, *Phys. Rev. D* **80**, 104014 (2009).
- [22] P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, N. Mazumder, and A. J. Weinstein, Effectual template bank for the detection of gravitational waves from inspiralling compact binaries with generic spins, *Phys. Rev. D* **89**, 084041 (2014).
- [23] Ian Harry, Stephen Privitera, Alejandro Bohé, and Alessandra Buonanno, Searching for gravitational waves from compact binaries with precessing spins, *Phys. Rev. D* **94**, 024012 (2016).
- [24] Connor McIsaac, Charlie Hoy, and Ian Harry, A search technique to observe precessing compact binary mergers in the advanced detector era, *Phys. Rev. D* **108**, 123016 (2023).
- [25] Pradnya A. Vikhar, *Evolutionary Algorithms: A Critical Review and its Future Prospects* (IEEE, 2016), pp. 261–265.
- [26] M. Cavaglià, S. Gaudio, T. Hansen, K. Staats, M. Szczepańczyk, and M. Zanolin, Improving the background of gravitational-wave searches for core collapse supernovae: A machine learning approach, *Mach. Learn.* **1**, 015005 (2020).
- [27] Teerth Gill Marco Cavaglia and Kai Staats, Finding the origin of noise transients in LIGO data with machine learning, *Commun. Comput. Phys.* **25**, 963 (2018).
- [28] Vasileios Skliris, Michael R. K. Norman, and Patrick J. Sutton, Real-time detection of unmodelled gravitational-wave transients using convolutional neural networks, *arXiv:2009.14611*.
- [29] Serena Vinciguerra, John Veitch, and Ilya Mandel, Accelerating gravitational wave parameter estimation with multi-band template interpolation, *Classical Quantum Gravity* **34**, 115006 (2017).
- [30] Soichiro Morisaki, Accelerating parameter estimation of gravitational waves from compact binary coalescence using adaptive frequency resolutions, *Phys. Rev. D* **104**, 044062 (2021).
- [31] Rory Smith, Scott E. Field, Kent Blackburn, Carl-Johan Haster, Michael Pürrer, Vivien Raymond, and Patricia Schmidt, Fast and accurate inference on gravitational waves from precessing compact binaries, *Phys. Rev. D* **94**, 044031 (2016).
- [32] Priscilla Canizares, Scott E. Field, Jonathan Gair, Vivien Raymond, Rory Smith, and Manuel Tiglio, Accelerated gravitational wave parameter estimation with reduced order modeling, *Phys. Rev. Lett.* **114**, 071104 (2015).
- [33] Soichiro Morisaki, Rory Smith, Leo Tsukada, Surabhi Sachdev, Simon Stevenson, Colm Talbot, and Aaron Zimmerman, Rapid localization and inference on compact binary coalescences with the Advanced LIGO-Virgo-KAGRA gravitational-wave detector network, *Phys. Rev. D* **108**, 123040 (2023).
- [34] Lalit Pathak, Amit Reza, and Anand S. Sengupta, Fast and faithful interpolation of numerical relativity surrogate waveforms using meshfree approximation, *arXiv:2403.19162*.
- [35] Michele Maggiore, *Gravitational Waves: Volume 1: Theory and Experiments* (OUP Oxford, 2007).
- [36] Theoharis A. Apostolatos, Curt Cutler, Gerald J. Sussman, and Kip S. Thorne, Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries, *Phys. Rev. D* **49**, 6274 (1994).
- [37] Patricia Schmidt, Mark Hannam, and Sascha Husa, Towards models of gravitational waveforms from generic binaries: A simple approximate mapping between precessing and non-precessing inspiral signals, *Phys. Rev. D* **86**, 104063 (2012).
- [38] Tanja Hinderer, Benjamin D. Lackey, Ryan N. Lang, and Jocelyn S. Read, Tidal deformability of neutron stars with realistic equations of state and their gravitational wave signatures in binary inspiral, *Phys. Rev. D* **81**, 123016 (2010).
- [39] K. Atkinson and W. Han, *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*, Lecture Notes in Mathematics (Springer, New York, 2012).
- [40] Y. Kosmann-Schwarzbach and S. F. Singer, *Groups and Symmetries: From Finite Groups to Lie Groups*, Universitext (Springer, New York, 2009).
- [41] Ian Harry, Juan Calderón Bustillo, and Alex Nitz, Searching for the full symphony of black hole binary mergers, *Phys. Rev. D* **97**, 023004 (2018).
- [42] Geraint Pratten *et al.*, Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes, *Phys. Rev. D* **103**, 104056 (2021).
- [43] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leïla Haegel, Sascha Husa, Frank Ohme, Geraint Pratten, and Michael Pürrer, Simple model of complete precessing black-hole-binary gravitational waveforms, *Phys. Rev. Lett.* **113**, 151101 (2014).
- [44] Cecilio García-Quirós, Marta Colleoni, Sascha Husa, Héctor Estellés, Geraint Pratten, Antoni Ramos-Buades, Maite Mateu-Lucena, and Rafel Jaume, Multimode frequency-domain model for the gravitational wave signal from nonprecessing black-hole binaries, *Phys. Rev. D* **102**, 064002 (2020).
- [45] K. G. Arun, Alessandra Buonanno, Guillaume Faye, and Evan Ochsner, Higher-order spin effects in the amplitude and phase of gravitational waveforms emitted by inspiralling compact binaries: Ready-to-use gravitational waveforms, *Phys. Rev. D* **79**, 104023 (2009).
- [46] T. W. Baumgarte and S. L. Shapiro, *Numerical Relativity: Solving Einstein's Equations on the Computer* (Cambridge University Press, Cambridge, England, 2010).
- [47] Frans Pretorius, Evolution of binary black-hole spacetimes, *Phys. Rev. Lett.* **95**, 121101 (2005).
- [48] Alessandra Buonanno, Bala R. Iyer, Evan Ochsner, Yi Pan, and B. S. Sathyaprakash, Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors, *Phys. Rev. D* **80**, 084043 (2009).
- [49] Francesco Messina, Reetika Dudi, Alessandro Nagar, and Sebastiano Bernuzzi, Quasi-5.5PN TaylorF2 approximant for compact binaries: Point-mass phasing and impact on the tidal polarizability inference, *Phys. Rev. D* **99**, 124051 (2019).
- [50] A. Buonanno and T. Damour, Effective one-body approach to general relativistic two-body dynamics, *Phys. Rev. D* **59**, 084006 (1999).

- [51] Rossella Gamba, Sarp Akçay, Sebastiano Bernuzzi, and Jake Williams, Effective-one-body waveforms for precessing coalescing compact binaries with post-Newtonian twist, *Phys. Rev. D* **106**, 024020 (2022).
- [52] Geraint Pratten, Sascha Husa, Cecilio García-Quirós, Marta Colleoni, Antoni Ramos-Buades, Héctor Estellés, and Rafel Jaume, Setting the cornerstone for a family of models for gravitational waves from compact binaries: The dominant harmonic for nonprecessing quasicircular black holes, *Phys. Rev. D* **102**, 064001 (2020).
- [53] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leïla Haegel, Sascha Husa, Frank Ohme, Geraint Pratten, and Michael Pürrer, Simple model of complete precessing black-hole-binary gravitational waveforms, *Phys. Rev. Lett.* **113**, 151101 (2014).
- [54] Greg Hornby, Al Globus, Derek Linden, and Jason Lohn, Automated antenna design with evolutionary algorithms, in *Collection of Technical Papers—Space 2006 Conference* (2006), 10.2514/6.2006-7242.
- [55] Lee Altenberg, *The Schema Theorem and Price's Theorem* (Elsevier, New York, 1995), pp. 23–49.
- [56] LIGO Scientific and Virgo Collaborations, GW190412: Observation of a binary-black-hole coalescence with asymmetric masses, *Phys. Rev. D* **102**, 043015 (2020).
- [57] LIGO Scientific and Virgo Collaborations, GW190814: Gravitational waves from the coalescence of a 23 solar mass black hole with a 2.6 solar mass compact object, *Astrophys. J. Lett.* **896**, L44 (2020).
- [58] Geraint Pratten, Patricia Schmidt, Riccardo Buscicchio, and Lucy M. Thomas, Measuring precession in asymmetric compact binaries, *Phys. Rev. Res.* **2**, 043096 (2020).
- [59] Charlie Hoy, Cameron Mills, and Stephen Fairhurst, Evidence for subdominant multipole moments and precession in merging black-hole-binaries from GWTC-2.1, *Phys. Rev. D* **106**, 023019 (2022).
- [60] Curt Cutler and Éanna E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform?, *Phys. Rev. D* **49**, 2658 (1994).
- [61] Christopher M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)* (Springer-Verlag, Berlin, Heidelberg, 2006).
- [62] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, *Introduction to Algorithms* (MIT Press, 2001).
- [63] Tobias Blickle and Lothar Thiele, A comparison of selection schemes used in evolutionary algorithms, *Evol. Comput.* **4**, 361 (1996).
- [64] Thomas Bäck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms* (Oxford University Press, 1996).
- [65] Gregory Ashton *et al.*, Bilby: A user-friendly Bayesian inference library for gravitational-wave astronomy, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [66] J. Veitch and A. Vecchio, Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network, *Phys. Rev. D* **81**, 062003 (2010).
- [67] F. P. Preparata and M. I. Shamos, *Computational Geometry: An Introduction*, Monographs in Computer Science (Springer New York, 1993).
- [68] Robert J. Fowler, Michael S. Paterson, and Steven L. Tanimoto, Optimal packing and covering in the plane are NP-complete, *Inf. Proc. Lett.* **12**, 133 (1981).
- [69] Eric Thrane and Colm Talbot, An introduction to Bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models, *Publ. Astron. Soc. Aust.* **36**, e010 (2019).
- [70] Patricia Schmidt, Frank Ohme, and Mark Hannam, Towards models of gravitational waveforms from generic binaries: II. Modeling precession effects with a single effective precession parameter, *Phys. Rev. D* **91**, 024043 (2015).