Template Bank for Compact Binary Coalescence Searches in Gravitational Wave Data: A General Geometric Placement Algorithm

Javier Roulet,† Liang Dai, Tejaswi Venumadhav, Barak Zackay, and Matias Zaldarriaga

1Department of Physics, Princeton University, Princeton, NJ, 08540, USA
2School of Natural Sciences, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA

(Dated: April 4, 2019)

We introduce an algorithm for placing template waveforms for the search of compact binary mergers in gravitational wave interferometer data. We exploit the smooth dependence of the amplitude and unwrapped phase of the frequency-domain waveform on the parameters of the binary. We group waveforms with similar amplitude profiles and perform a singular value decomposition of the phase profiles to obtain an orthonormal basis for the phase functions. The leading basis functions span a lower-dimensional linear space in which the unwrapped phase of any physical waveform is well approximated. The optimal template placement is given by a regular grid in the space of linear coefficients. The algorithm is applicable to any frequency-domain waveform model and detector sensitivity curve. It is computationally efficient and requires little tuning. Applying this method, we construct a set of template banks suitable for the search of aligned-spin binary neutron star, neutron-star–black-hole and binary black hole mergers in LIGO–Virgo data.

PACS numbers:

I. INTRODUCTION

The optimal algorithm to search for known signals in the presence of Gaussian noise is matched-filtering, in which a signal template is cross-correlated with the data and triggers are recorded whenever the correlation exceeds some threshold. In the context of gravitational wave detection, compact binary coalescences are a good example of predictable signals for which we have accurate models, and thus are well suited for matched filtering. Source parameters describing the waveform, however, are not known a priori and therefore one needs a bank of waveform templates that adequately cover the parameter space.

The notion of good coverage is characterized by the requirement that any physical waveform the search aims to detect has a sufficiently large match with at least one waveform in the bank. For example, the LIGO and Virgo Collaborations have aimed at a minimum match of 97% for any aligned-spin binary merger with component masses between 1 and \( \sim 200 M_\odot \), for which they require a template bank consisting of \( \sim 4 \times 10^5 \) waveforms [1]. The large number of templates involved means that matched filter search is a sizable computational task. This means that an efficient bank should not over-cover the parameter space. In other words, the templates should be uniformly spaced with respect to a distance defined in terms of the matched-filtering mismatch between templates (defined in [1]). This incorporates the notion that, from the perspective of signal detection, two waveforms that sufficiently resemble each other are essentially indistinguishable in the presence of noise. Source parameters can be mutually degenerate in the sense that different parameter combinations may describe similar waveforms. The optimal placement of templates in physical parameter space is very non-uniform; for example, an order of magnitude more templates are needed to search for mergers with 1–3\( M_\odot \) components (“neutron stars”) than for mergers with 3–200\( M_\odot \) (“black holes”).

Two broad classes of template-placement algorithms have been developed in the literature. One robust method is “stochastic placement” [2–5]: waveforms are randomly drawn from the desired parameter space, and one gradually builds up the bank by only accepting newly drawn waveforms that differ sufficiently from the ones the bank already has, and rejecting those that are too similar to at least one existing waveform. This method, however, has the shortcoming that a large number of trial waveforms needs to be drawn before convergence is achieved (much more than the required number of templates in the bank). This method also tends to over-cover the parameter space (the average template density is higher than optimal at fixed minimum match [3]).

A different method to construct the bank is “geometric placement”. Here, a metric in the parameter space is defined based on the matched-filtering overlap between waveforms [7, 8]. This metric is then used to define a regular lattice [9–11]. However, it is in general difficult to derive this metric, especially if the parameter space is high dimensional or if the waveform model is not analytic. Approximations to the metric have first been found by using suitably reparameterized analytic, post-Newtonian (PN) non-spinning waveform models [7, 8, 12]; later generalizations include the use of phenomenological waveform models and template parameters [13], the inclusion of aligned-spin PN models [14, 15], or numerical evaluation from arbitrary waveform models [16].

In practice, a combination of the two methods is often a better strategy. For example, one can place templates geometrically at low masses and stochastically at high
masses [15], or one can use many small patches with regularly spaced templates, which are themselves placed in a stochastic fashion to cover the entire parameter space [16].

In this work, we develop a fast and general method to construct a high-effectualness template bank using geometric placement. We further require that the template bank be built from sub-banks that can be approximated to have a fixed amplitude profile \( A(f) \). This property is useful for implementing the noise amplitude-spectral-density drift correction, a key component for precise matched filtering [17, 18]. Our method relies on the construction of a flat, linear space of phase functions that embeds the space of physical waveforms. The Euclidean distance in this space coincides with the mismatch distance between similar waveforms, making these coordinates naturally suited for geometric placement of templates. These coordinates are similar in essence to the ones introduced in Brown et al. [14], except that we generalize them to arbitrary waveform models and component mass ranges. The paper is organized as follows. In II we define a metric based on the mismatch between templates and show how the desired Euclidean space can be constructed. In III we apply this formalism to the construction of a search-quality template bank that targets stellar-mass compact binary mergers. We summarize our results in IV. The bank presented here was used in the search described in Vennumadhav et al. [17], except for some limited differences that we report in Appendix A.

II. LINEAR METRIC SPACE

In this section, we define the notion of distance between templates and describe the construction of a low-dimensional linear space of phase functions in which the metric is Euclidean. We build this linear space naturally suited for geometric placement of templates. We first introduce the noise-weighted inner product in the frequency domain [20]

\[
(h_i | h_j) := 4 \int_0^\infty df \frac{\tilde{h}_i(f)\tilde{h}_j^*(f)}{S_n(f)}.
\]

Here, \( S_n(f) \) is a fiducial one-sided power spectral density (PSD) of the detector noise and tildes indicate Fourier transforms. The match between \( h_i \) and \( h_j \) is given by \( \text{Re} z_{ij} \), where

\[
z_{ij} := \frac{(h_i | h_j)}{\sqrt{(h_i | h_i)(h_j | h_j)}} = (h_i | h_j).
\]

A general frequency-domain waveform model can be cast to the form

\[
\tilde{h}(f; p) = A(f; p)e^{i\psi(f; p)},
\]

where the frequency dependent functions \( A \) and \( \psi \) vary slowly with the binary parameters \( p \), as illustrated in

![FIG. 1: An example of two waveforms that look very different to each other in the frequency domain (top panel) but have very similar amplitude and phase profiles (middle and bottom panels). The amplitude and phase profiles can be well captured by a low-dimensional linear space spanned by a few basis functions.](https://example.com/fig1.png)
For matched filtering, the phase \( \psi \) is the most important to describe with high accuracy, since loss of phase coherence leads to a rapid degradation of the signal-to-noise ratio (SNR).

Moreover, it is important to analyze templates with different amplitude profiles \( A(f) \) separately as the matched filtering correction from PSD drifts depends on \( A(f) \) [19]. Thus we assume in the following that \( A(f; \mathbf{p}) \approx \overline{A}(f) \) is valid for a suitably chosen subset of parameters. To achieve this, we sort a large number of randomly sampled physical input waveforms into groups of similar amplitude profile. In each group, we require that the match of the amplitudes to a reference \( \overline{A}(f) \) exceeds a minimum

\[
(A_i | \overline{A}) = 4 \int_0^{\infty} df \frac{A_i(f) \overline{A}(f)}{S_n(f)} \geq 0.96
\]

for all input waveforms \( h_i \) in the group. Note that the match of the amplitudes sets an upper bound on the mismatch of the waveforms. Our approach will be to split a template bank into “subbanks”, each subbank describing one group of input waveforms which share the same approximate amplitude profile \( \overline{A}(f) \).

We design the subbanks in order to minimize the average amplitude mismatch as follows. We start with a single subbank that contains all the input waveforms, and define its reference amplitude profile as the root-mean-square

\[
\overline{A}(f) := \sqrt{\langle A^2 \rangle(f)},
\]

where the angled brackets indicate average over the input waveforms in the subbank. This choice inherits the normalization of the input waveforms. We compute the amplitude match Eq. \( \overline{A} \) for all the waveforms; if the worst match satisfies the chosen bound we stop. If it does not, we add a new subbank with a reference amplitude given by the waveform with the worst amplitude match. We then optimize the choice of reference amplitudes using the \( k \)-means algorithm: we reassign waveforms to subbanks by their best amplitude match, redefine the amplitude profile of the subbanks using Eq. \( \overline{A} \), and iterate these two steps a few times to achieve convergence. Finally we recompute the worst match and decide if a new subbank is needed, in which case we repeat the above process.

Having decided on the division of subbanks, we wish to find an efficient representation of the set of phases \( \psi(f) \) as a linear combination of a small number of basis functions,

\[
\psi(f; \mathbf{p}) = \overline{\psi}(f) + \sum_\alpha c_\alpha(\mathbf{p}) \psi_\alpha(f),
\]

where \( \alpha \) is an integer index that enumerates the basis functions and \( \overline{\psi}(f) \) is an average phase which we are free to define. From now on, we abandon the physical parameters \( \mathbf{p} \) and describe the waveforms in terms of their \( c_\alpha \) components:

\[
h(f; \mathbf{c}) = \overline{A}(f) \exp \left[ i \left( \overline{\psi}(f) + \sum_\alpha c_\alpha \psi_\alpha(f) \right) \right].
\]

We now express the match between two waveforms using the above decomposition. As mentioned earlier, template waveforms are defined up to arbitrary time and phase offsets, namely an additive piece to the phase that is a linear function of the frequency \( \Delta \psi(f) = \phi_0 + 2\pi f \tau_0 \).

To achieve this, we first choose two basis functions \( \psi_0(f) \) and \( \psi_1(f) \) to span the subspace of linear phases so that \( \phi_0 \) and \( \phi_1 \) capture phase and time offsets, respectively, and in particular \( \psi_0(f) \equiv 1 \). If two waveforms are similar, their inner product Eq. \( \overline{A} \) to second order in \( \delta c_\alpha \) is approximately

\[
\langle h, h + \delta \mathbf{c} \rangle = 4 \int_0^{\infty} df \frac{\overline{A}^2(f)}{S_n(f)} e^{i \sum_\alpha \delta c_\alpha \psi_\alpha(f)}
\]

\[
\approx 4 e^{i \delta \phi_0} \int_0^{\infty} df \frac{\overline{A}^2(f)}{S_n(f)} \left[ 1 + i \sum_\alpha \delta c_\alpha \psi_\alpha(f) \right] - \frac{1}{2} \sum_{\alpha, \beta} \delta c_\alpha \delta c_\beta \psi_\alpha(f) \psi_\beta(f) + \mathcal{O}(\delta c^3).
\]

This motivates a new inner product, with respect to which we will orthonormalize the basis functions:

\[
\langle \psi_\alpha, \psi_\beta \rangle := 4 \int_0^{\infty} df \frac{\overline{A}^2(f)}{S_n(f)} \psi_\alpha(f) \psi_\beta(f)
\]

\[
= \delta_{\alpha \beta},
\]

which we enforce by a suitable choice of the basis functions \( \psi_\alpha(f) \) (described below). In particular, the first condition \( \langle \psi_0, \psi_0 \rangle = 1 \) is the normalization Eq. \( \overline{A} \), and the two first basis functions are

\[
\psi_0(f) = 1,
\]

\[
\psi_1(f) = \frac{f - \overline{\tau}}{\sqrt{\overline{\tau}^2 - \overline{\tau}^2}},
\]

where we define \( \overline{\tau} := 4 \int_0^{\infty} df f^\infty A^2(f)/S_n(f) \).

Using orthonormality, Eq. \( \overline{A} \) becomes

\[
\langle h, h + \delta \mathbf{c} \rangle \approx e^{i \delta \phi_0} \left( 1 - \frac{1}{2} \sum_{\alpha \neq 1} \delta c_\alpha^2 \right) + \mathcal{O}(\delta c^3).
\]

Thus, for nearby templates the distance Eq. \( \overline{A} \) is

\[
d_{\mathbf{c}, \mathbf{c} + \delta \mathbf{c}} \approx \frac{1}{2} \sum_{\alpha \neq 1} \delta c_\alpha^2 + \mathcal{O}(\delta c^3),
\]

which means that the mismatch distance is given by an Euclidean metric in \( \mathbf{c} \) space at small displacements. We construct the bank on a regular grid in \( \mathbf{c} \) space with spacings \( \Delta c_\alpha \lesssim 1 \), chosen sufficiently small so as to guarantee a minimal loss of match.
We note in passing that we can also compute the distance in the opposite limit of large separation, which is
useful for estimating the long-range correlations between
triggers from different templates during a search. Assuming
now that the templates are separated by \( \delta c = D \hat{n} \),
with \( \sum_{\alpha} n_{\alpha}^2 = 1 \) and \( D \gg 1 \), we can perform a stationary
phase approximation around the frequencies \( f_j \) at which
\( \sum_{\alpha} n_{\alpha} \psi_{\alpha}^\prime(f_j) = 0 \). This yields
\[
(h(c) \mid h(c + D \hat{n})) = 4 \int_{0}^{\infty} df \frac{\overline{A}(f)}{S_n(f)} e^{i D \sum_{\alpha} n_{\alpha} \psi_\alpha(f)} \\
\approx \frac{4}{\sqrt{D}} \sum_{j} \frac{\overline{A}(f_j)}{S_n(f_j)} \sqrt{2\pi} \exp(-i \frac{\pi}{2} + i D \sum_{\alpha} n_{\alpha} \psi_\alpha(f_j))
\]
(16)
Thus, the long-range correlation between two templates
separated by \( D \) decays as \( 1/\sqrt{D} \) (this holds for the match
without maximization over time).
In practice we choose the set of basis functions \( \psi_\alpha(f) \) as follows:
1. Define a discrete frequency grid \( \{f_k\} \) (our choice is
described in [11]). The integrals over frequency will be
approximated by quadratures \( \sum_{k} \Delta f_k \ldots \);
2. Compute a moderately large number of waveforms
for random parameter choices (we use \( 5 \times 10^4 \)), and extract the unwrapped phases, \( \{\psi_{\alpha}(f_k)\} \), as illustrated in the top panel of Fig. 2;
3. Subtract the average phase \( \overline{\psi}(f) \);
4. Subtract the projection onto the first two dimen-
sions so that the phase residuals
\[
\Delta \psi_{\alpha}(f) = \psi_{\alpha}(f) - \overline{\psi}(f) - \sum_{\alpha=0}^{1} \langle \psi_{\alpha}-\overline{\psi}, \psi_{\alpha} \rangle \psi_{\alpha}(f)
\]
(17)
are orthogonal to \( \psi_0 \) and \( \psi_1 \) with respect to the inner product Eq. (12) (middle panel of Fig. 2);
5. Construct a matrix of weighted phase residuals
\[
M_{ik} = w_k \Delta \psi_{\alpha}(f_k),
\]
\[
w_k = 2 \overline{A}(f_k) \sqrt{\Delta f_k/S_n(f_k)}
\]
(18)
and find its singular-value decomposition (SVD)
\[
M_{ik} = \sum_{\alpha} U_{i\alpha} D_{\alpha} V_{\alpha k}.
\]
(19)
\( U, V \) are orthogonal matrices and we sort the axes
so that the eigenvalues \( D_{\alpha} > 0 \) are in decreasing
order. From the orthogonality of \( V \), i.e.
\( \sum_{k} V_{\alpha k} V_{\beta k} = \delta_{\alpha \beta} \), we can identify
\[
V_{\alpha k} = w_k \psi_{\alpha}(f_k)
\]
(20)
which satisfies the orthonormality Eq. (12) and defines the basis functions, with the convention that the \( \alpha \) start at 2 (bottom panel of Fig. 2).

III. Constructing a Search Quality Template Bank

In this Section, we apply the method developed in [11]
for the construction of a template bank suitable to the
search of gravitational wave strain signals from binary
neutron stars, neutron-star–black-hole and binary black
hole mergers.
We choose lower and upper frequency cutoffs of \( f_{\text{min}} = 24 \) Hz, and \( f_{\text{max}} = 512 \) Hz, respectively. These cutoffs
are chosen such that the resulting loss in SNR is lower than
2% for binary neutron star templates (the amplitude pro-
files of these whitened waveforms, i.e., \( A(f)/\sqrt{S_n(f)} \), are

\[
\psi_\alpha = U_{i\alpha} D_{\alpha}.
\]
(21)
Since \( U \) is an orthogonal matrix, \( |U_{i\alpha}| \leq 1 \) and \( |D_{\alpha}| \leq 1 \), that is, the extent spanned by the input samples
along each dimension in component space is bounded by \( D_{\alpha} \). This means that the information in the templates
is captured by the first few components along the larger
dimensions, and we can reduce the dimensionality of our
description by dropping the dimensions that have \( D_{\alpha} \ll 1 \).
essentially independent of parameters since the cut-off scale is outside the LIGO sensitivity band). Formally, the accumulated SNR\(^2\) = \(4 \int_0^{24 \text{Hz}} df A^2(f)/S_n(f) \approx 4 \int_{512 \text{Hz}}^\infty df A^2(f)/S_n(f) \approx 10^{-2}\) outside our frequency range. It is advisable to restrict the frequency range because the linear-free phase, and thus the basis functions, grow rapidly at both ends (see Fig. 3), and our Taylor expansion Eq. (11) would become inaccurate. As we noted above, it is exactly at these frequencies where the contribution to the matched-filtering SNR vanishes. It is better to discard these frequencies rather than to try and capture the negligible information content within by adding extra dimensions to the template bank. Furthermore, this has the additional benefit that the strain data can be down-sampled during analysis, which reduces the computational cost of the search.

We define the fiducial PSD empirically from the PSDs of 200 LIGO Handford and LIGO Livingston data files chosen randomly from the Second Advanced LIGO Observing Run (O2) release \([21,22]\). Each individual PSD was computed as described in \([17]\). The fiducial PSD is constructed using the 10th percentile of all the sample PSDs in each frequency bin. This choice is robust to large fluctuations in the sample PSDs, and is representative of optimal detector conditions.

We choose a target parameter space of compact binary mergers satisfying the following bounds:

\[
1 \, \text{M}_\odot < m_1, m_2 < 100 \, \text{M}_\odot, \quad (22)
\]

\[
q > \begin{cases} 
1/50 & \text{if } m_2 < 3 \, \text{M}_\odot \\
1/18 & \text{otherwise,}
\end{cases} \quad (23)
\]

\[
|\chi_1|, |\chi_2| < 0.99, \quad (24)
\]

where \(m_1\) and \(m_2\) are the primary and secondary masses, respectively, \(q = m_2/m_1 \lesssim 1\) is the mass ratio, and \(\chi_1\) and \(\chi_2\) are the individual dimensionless spin projections in the direction of the orbital angular momentum.

As mentioned before, the number of templates required to describe waveforms from low-mass mergers is significantly larger than that for high-mass mergers, due to the larger number of wave cycles in band. Searches with larger template banks suffer a penalty in sensitivity because of the increased look-elsewhere effect. To prevent the high penalty inherent to the lower-mass region of parameter space from affecting the higher-mass regions, we propose to divide the search space into a number of regions and perform an independent search in each. Each search then only pays an additional look-elsewhere penalty that a few other searches are performed, but is unaffected by the potentially huge size of the other banks. This division can be interpreted as implementing a prior about which templates are more likely to produce an astrophysical trigger: if we expect comparable numbers of high- and low-mass signals but have vastly more templates at low-mass, any particular low-mass template is much less likely to produce an astrophysical trigger. In addition, templates in different regions of parameter space are sensitive to different types of noise transients in the strain data. Dividing the search into several regions enables us to recognize the different types of noise background that a search using each class of templates is subject to.

Under the above motivations, we divide the search space into regions based on the component masses, and construct a separate template bank for each of them. The division is illustrated in Fig. 3 and is defined as follows. We refer to binary components with masses between 1 and 3 M\(_\odot\) as neutron stars, and to components with masses between 3 and 100 M\(_\odot\) as black holes. We make three binary neutron star (BNS) template banks, three neutron-star–black-hole (NSBH) banks, and five binary black hole (BBH) banks. The banks within each of these categories are defined by bins in the chirp mass \(M := (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}\). We put the bounds between the three BNS banks at \(M = \{1.1, 1.3\} \, \text{M}_\odot\). This choice is motivated by the observation that the chirp masses of the known Galactic binary neutron stars expected to merge within a Hubble time lie in a narrow range \([23]\), and therefore we might expect more astrophysical signals from this chirp mass range (which we further expand to account for the redshift of the detector-frame masses up to \(z \sim 0.05\), or a luminosity distance \(d_L \sim 200 \, \text{Mpc}\)). In this way, we minimize the number of templates in the most astrophysically probable BNS bank, BNS 1, enhancing our sensitivity to those sys-

![FIG. 3: Component masses of the input waveforms used for constructing our template banks. We divide the parameter space according to the component masses into three banks of binary neutron star waveforms (BNS 0-2), three banks of neutron-star–black-hole waveforms (NSBH 0-2), and five banks of binary black hole waveforms (BBH 0-4). We further divide each bank into subbanks (color coded) under the criterion that the match of the amplitude profiles with a reference profile specific to each subbank exceeds 0.96. The dots show the 5 × 10\(^4\) input waveforms we use to build each bank. We do not scatter plot here individual template waveforms in each bank, as those lack an explicit representation in terms of the source physical parameters.](image-url)
s1. A similar strategy was adopted in Refs. 24, 25. For other banks, we use logarithmic chirp-mass bins: we place the bounds between the three NSBH banks at $M = \{3, 6\} M_\odot$, and those between the five BBH banks at $M = \{5, 10, 20, 40\} M_\odot$. We generate $5 \times 10^4$ input waveforms in each bank using the IMRPhenomD approximant [26]. Based on the amplitude profiles $A_i(f)$ of the input waveforms, we further divide each bank into sub-banks as explained in [11]. We find that a single subbank is sufficient for waveforms with $m_{1,2} \lesssim 15 M_\odot$, but multiple amplitude subbanks are needed for heavier mergers as the frequency at which $A(f)$ is cut-off falls within the LIGO sensitive band. Table II summarizes the parameters of all template banks. The banks differ greatly in size, which justifies the division of the search space into multiple banks.

For each subbank, we apply the procedure outlined in [11] to define a set of basis phase functions that generate a linear space and obtain the projections of the input waveforms onto that space. These are shown in Fig. 4 for the example case of the BBH 0 bank, with the points color coded by their chirp mass. The first three dimensions capture practically all the diversity of the input waveforms. Also note the large differences in size from the leading dimension to the sub-leading ones. The chirp mass, which on physical grounds we expect to be the variable that carries the most information about the waveform, indeed corresponds very well to the coefficient of the leading dimension $c_2$.

Next, we choose a grid spacing $\Delta c_\alpha$ common to all dimensions and define a rectangular grid in component space as follows. We force the point $c = 0$ to be a grid point, because the SVD typically aligns the highest density regions (where the input physical waveforms tend to be) with the axes. Along each dimension, we add uniformly-spaced points until the whole range spanned by the input waveforms is covered. We allow the spacing to slightly decrease so that the most extreme input component is half the grid spacing away from the most extreme grid point. We do this for each dimension and in the positive and negative directions separately. Finally, not all the points of the rectangular grid describe physically viable waveforms. We only keep the templates that are close to at least one input waveform, with the following criterion. For every input waveform set of components, we keep the closest grid point and a patch of the grid around it, with size equal to the corresponding dimension times a tunable fudge factor $\zeta \sim 0.1$.

Indeed, as Fig. 4 shows, the input physical waveforms do not fill the entire rectangular volume but are distributed within some irregularly shaped region. Furthermore, the density of input waveforms is low in the low-$M$ region, where the waveforms have more wave cycles in band and hence are mutually more distinguishable. Holes can be produced in the physically viable region if the fudge factor $\zeta$ is too small, and there is an excess of unphysical templates if $\zeta$ is too large. We choose the $\Delta c_\alpha$ and $\zeta$ parameters such that we achieve a good balance between economic template bank size and high bank effectualness. The values chosen for each bank are reported in Table II.

In Table II we observe a general trend with the mass: the banks for lighter mergers tend to have fewer subbanks and the first dimension spans a wider range. By comparison, the banks for heavier mergers have more subbanks, with smaller dimensions. The increase in the number of subbanks for heavier mergers is caused by the cutoff frequency falling in the band, which increases the variety of amplitude profiles.

There are interesting implications of the number of dimensions and their size for parameter estimation. Given an astrophysical signal, in the limit of high SNR $\rho$, the parameter likelihood is approximately given by $P(d \mid p) \propto \exp(-\rho^2|z|^2/2)$, where $z = (h(p_\star) - h(p))$ is the complex match of $p$ to the best-fit parameters $p_\star$ [27]. By virtue of Eq. (15), this means that the likelihood is approximately an isotropic Gaussian in terms of the $c_\alpha$ coordinates, with a width $\sim 1/\rho$. The number of dimensions can therefore be interpreted as the number of independent parameters that can be measured, and the size of each dimension as the relative precision that can be obtained for a fixed SNR (with the caveat that we have restricted the frequency range; for example, information about the tidal deformability comes from frequencies higher than our 512 Hz cutoff).

For example, for BNS (and effectively for light BBH) the banks have two dimensions, with a large first dimension well correlated with the chirp mass (Fig. 4). The two measurable parameters are the chirp mass, which indeed can be measured to much higher precision than for

---

1 The authors thank Thomas Dent for suggesting this approach.
heavy systems, and a combination of the mass ratio and effective spin which can be measured with a lower precision. These are the leading contributions to the phase evolution as can be understood from the post-Newtonian expansion.

An important advantage of our geometric coordinates is that they are well suited for a two-step search that effectively achieves a smaller grid spacing at reduced computational cost. We realize this by refining the template grid on demand around all triggers that exceed an appropriately lowered SNR threshold \cite{17,28}. During the search, we first use a coarse grid, and refine every trigger using neighboring templates from a denser grid that has half the spacing along each dimension. The fact that the distance between $c_a$ components translates directly to mismatch (Eq. 15) makes this method straightforward to implement.

To characterize the effectualness of the bank at recovering the target physical signals, we generate a set of $10^4$ random “test waveforms” within the parameter range of each bank, using the same approximant with which the input waveforms were generated. We choose the parameters from a distribution that is uniform in the component masses $m_1, m_2$ and aligned spins $\chi_1, \chi_2$. In principle, we would have to match each test waveform against every waveform in the bank to look for the best match. To save computational effort, we select a candidate best-match based on the approximate metric Eq. (15) by extracting the phase of the test waveform $\psi^{(i)}(f)$, projecting it onto the linear space, $c_{a(i)} = \left(\psi^{(i)} - \bar{\psi}, \psi_{a(i)}\right)$, and finding the closest grid point with respect to the Euclidean metric (15). Since \textit{a priori} we do not know which subbank best describes the test waveform, we pick the best candidate from each subbank and compute the match with all. The best match with our reduced set of candidates is a lower bound on the best match over all the waveforms in the bank. Rather than using Eq. (4) directly, we compute the match by following the detection strategy described in Vennumadhav \textit{et al.} \cite{17}: we account for the finite time resolution of the Fourier transform by downsampling the waveforms to 512 Hz and sinc-interpolating the matched-filter output twice. We show the result of this test in Fig. 5 in terms of the cumulative fraction of the matches with each bank before and after applying the grid refinement, which we use to assess the collection threshold on the coarse grid and the effectualness achieved for each bank, respectively. We find that depending on the bank 99% of the templates have a match higher than 0.95–0.98.

![FIG. 5: Effectualness of our template banks, tested on random waveforms drawn from a uniform distribution in individual masses and aligned spins. The vertical axis shows the fraction of the random trials that do not achieve a given match in the bank.](image-url)
IV. CONCLUSIONS

We have developed a general and computationally efficient geometric placement algorithm to construct high-effectualness template banks for detecting gravitational waves from compact binary mergers. We have constructed a basis of functions that generate a linear space of phase profiles on which the mismatch metric is Euclidean. For the purpose of signal detection, we shift the focus away from physical parameters to the linear coefficients for the basis phase profiles. We identified which components carry the largest amount of information about physical waveforms and what is the minimal set required to guarantee a desired match. The basis functions can be determined from a set of input waveforms whose size is small compared to that of the bank. The basis functions can be generated with any frequency-domain waveform model. The resolution of the bank can be decided independently after the basis functions have been found; in particular, it can be increased arbitrarily at negligible computational cost since no further evaluations of the physical waveform approximants need to be done. Our algorithm guarantees that within each of the few subbanks that make up one template bank, all templates share the same amplitude profile, a property that is critical for the correction of the power-spectral-density drift in signal processing.

We have applied our algorithm to the construction of a collection of eleven template banks that together cover the parameter space associated to stellar-mass compact binary mergers with aligned spins. We find the effectualness and total number of templates to be comparable to the ones obtained by other algorithms in the literature [1,16]; detailed comparisons are difficult due to the different parameter spaces targeted in various works. We note that our template bank includes rapidly spinning neutron stars, which to date have not been searched for in the gravitational wave data. We implement a two-step search with a coarse grid that we refine around triggers at the time of search, a task for which our new formalism is ideally suited. This is an important step to reduce the number of templates while preserving a high effectualness.

Looking forward, an accurate and fast interpolation from physical parameters to the $c_{\alpha}$ component space would be extremely useful for rapid parameter estimation. First, because waveforms can be generated at negligible computational cost once the components are known. At least in cases where analytical waveform models are not valid, waveform generation dominates the computational cost of parameter estimation. Moreover, the likelihood would look close to an isotropic Gaussian in terms of the $c_{\alpha}$ coordinates due to orthonormality, making them a natural choice from the data analysis perspective.

The template bank described here is publicly available at [https://github.com/jroulet/template_bank](https://github.com/jroulet/template_bank).

Acknowledgements

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center [https://www.gw-openscience.org](https://www.gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. LIGO is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

LD acknowledges the support from the Raymond and Beverly Sackler Foundation Fund. TV acknowledges support by the Friends of the Institute for Advanced Study. BZ acknowledges the support of The Peter Svennilson Membership fund. MZ is supported by NSF grants AST-1409709, PHY-1521097 and PHY-1820775 the Canadian Institute for Advanced Research (CIFAR) program on Gravity and the Extreme Universe and the Simons Foundation Modern Inflationary Cosmology initiative.

Appendix A: Differences with the bank used in Venumadhav et al. [17]

In this appendix we report the differences between the template bank presented in this work and the one that was actually used in the binary black hole search reported by Venumadhav et al. [17].

- The bank used in Ref. [17] was restricted to BBHs, with a narrower range for the aligned spins $|\chi_{1,2}| < 0.85$ instead of 0.99.

- The reference PSD used to build the bank was $aLIGO_{MID-LOW}$ [29] instead of an empirically measured one.

- As a consequence, the bank used a low-frequency cutoff at 20Hz instead of 24Hz. We found that with the empirical PSD, the relative contribution to squared SNR from frequencies below 24Hz is < 1%.

- The optimization of the subbank amplitude profiles with the $k$-means clustering algorithm was not done; instead the division into subbanks was computed in a way analogous to the “stochastic placement” approach to building a template bank described in [1] with a required amplitude match $(A_i \mid \mathcal{A}) > 0.95$. The reference amplitude of the subbank was given by the first waveform generated in the subbank. Using the $k$-means clustering improves the best match by $\sim 0.01$ on average.
[1] T. Dal Canton and I. W. Harry, (2017), arXiv:1705.01845v1 [gr-qc].
[2] I. W. Harry, B. Allen, and B. S. Sathyaprakash, Phys. Rev. D 80, 104014 (2009).
[3] P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, N. Mazumder, and A. J. Weinstein, Phys. Rev. D 89, 084041 (2014).
[4] S. Privitera, S. R. P. Mohapatra, P. Ajith, K. Cannon, et al., Phys. Rev. D 89, 024003 (2014).
[5] C. Capano, I. Harry, S. Privitera, and A. Buonanno, Phys. Rev. D93, 124007 (2016).
[6] S. Roy, A. S. Sengupta, and N. Thakor, Phys. Rev. D 95, 104045 (2017).
[7] B. J. Owen, Phys. Rev. D 53, 6749 (1996).
[8] B. J. Owen and B. S. Sathyaprakash, Phys. Rev. D 60, 022002 (1999).
[9] S. Babak, R. Balasubramanian, D. Churches, T. Cokelaer, and B. S. Sathyaprakash, Classical and Quantum Gravity 23, 5477 (2006).
[10] T. Cokelaer, Phys. Rev. D 76, 102004 (2007).
[11] S. Babak, R. Biswas, P. Brady, D. A. Brown, et al., Physical Review D 87, 024033 (2013).
[12] T. Tanaka and H. Tagoshi, Phys. Rev. D 62, 082001 (2000).
[13] P. Ajith, S. Babak, Y. Chen, M. Hewitson, et al., Phys. Rev. D 77, 104017 (2008).
[14] D. A. Brown, I. Harry, A. Lundgren, and A. H. Nitz, Phys. Rev. D 86, 084017 (2012).
[15] I. W. Harry, A. H. Nitz, D. A. Brown, A. P. Lundgren, E. Ochsner, and D. Keppel, Phys. Rev. D 89, 024010 (2014).
[16] S. Roy, A. S. Sengupta, and P. Ajith, Phys. Rev. D 99, 024048 (2019).
[17] T. Vennumadhu, B. Zachay, J. Roulet, L. Dai, and M. Zaldarriaga, (2019), arXiv:1902.10341 [astro-ph.IM].
[18] B. Zachay et al., in preparation.
[19] B. Zachay, L. Dai, and T. Vennumadhu, arXiv:1806.08792v2 [astro-ph.IM].
[20] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, Phys. Rev. D 85, 122006 (2012).
[21] “Gravitational Wave Open Science Center (GWOSC),” www.gw-openscience.org/02/, accessed 2019-02-27.
[22] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, Journal of Physics: Conference Series 610, 012021 (2015).
[23] N. Farrow, X.-J. Zhu, and E. Thrane, (2019), arXiv:1902.03300v1 [astro-ph.HE].
[24] R. Magee, H. Fong, S. Caudill, C. Messick, et al., arXiv:1901.09884v2 [gr-qc].
[25] A. H. Nitz, A. B. Nielsen, and C. D. Capano, arXiv:1902.09496v1 [astro-ph.HE].
[26] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Purrrer, X. J. Forteza, and A. Bohé, Phys. Rev. D 93, 044007 (2016).
[27] J. Roulet and M. Zaldarriaga, Monthly Notices of the Royal Astronomical Society 484, 4216 (2019).
[28] B. Gadre, S. Mitra, and S. Dhurandhar, (2018), arXiv:1807.06803v1 [astro-ph.IM].
[29] LIGO Scientific Collaboration, “LIGO Algorithm Library - LALSuite,” free software (GPL) (2018).