Physical template family for gravitational waves from precessing binaries of spinning compact objects: Application to single-spin binaries

Yi Pan,[1] Alessandra Buonanno,[2] Yanbei Chen,[1] and Michele Vallisneri[3]
[1]*Theoretical Astrophysics and Relativity, California Institute of Technology, Pasadena, California 91125, USA*
[2]*Groupe de Gravitation et Cosmologie (GReCO), Institut d'Astrophysique de Paris (CNRS), 98*[bis] *Boulevard Arago, 75014 Paris, France*
[3]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA*
(Received 7 October 2003; published 19 May 2004)

The detection of the gravitational waves (GWs) emitted by precessing binaries of spinning compact objects is complicated by the large number of parameters (such as the magnitudes and initial directions of the spins, and the position and orientation of the binary with respect to the detector) that are required to model accurately the precession-induced modulations of the GW signal. In this paper we describe a fast matched-filtering search scheme for precessing binaries, and we adopt the physical template family proposed by Buonanno, Chen, and Vallisneri [Phys. Rev. D **67**, 104025 (2003)] for ground-based interferometers. This family provides essentially exact waveforms, written directly in terms of the physical parameters, for binaries with a single significant spin, and for which the observed GW signal is emitted during the phase of adiabatic inspiral (for LIGO-I and VIRGO, this corresponds to a total mass $M \lesssim 15 M_{\odot}$). We show how the detection statistic can be maximized automatically over all the parameters (including the position and orientation of the binary with respect to the detector), except four (the two masses, the magnitude of the single spin, and the opening angle between the spin and the orbital angular momentum), so the template bank used in the search is only four-dimensional; this technique is relevant also to the searches for GW from extreme-mass-ratio inspirals and supermassive black hole inspirals to be performed using the space-borne detector LISA. Using the LIGO-I design sensitivity, we compute the detection threshold (~ 10) required for a false-alarm probability of 10^{-3}/yr and the number of templates ($\sim 76\,000$) required for a minimum match of 0.97 for the mass range $(m_1, m_2) = [7,12] M_{\odot} \times [1,3] M_{\odot}$.

DOI: 10.1103/PhysRevD.69.104017 PACS number(s): 04.30.Db, 04.25.Nx, 04.80.Nn, 95.55.Ym

I. INTRODUCTION

Binaries consisting of a black hole (BH) in combination with another BH or with a neutron star (NS) are among the most promising gravitational-wave (GW) sources for first-generation laser-interferometer GW detectors such as the Laser Interferometer Gravitational Wave Observatory (LIGO) [1,2], VIRGO [3], GEO600 [2,4] and TAMA300 [5]. For LIGO-I and VIRGO, and for binaries with a total mass $M \lesssim 15 M_{\odot}$, the observed GW signal is emitted during the adiabatic-inspiral regime, where post-Newtonian (PN) calculations can be used to describe the dynamics of the binary and predict the gravitational waveforms emitted [6–9].

Very little is known about the statistical distribution of BH spin magnitudes in binaries: the spins could very well be large, with a significant impact on both binary dynamics and gravitational waveforms. On the contrary, it is generally believed that NS spins will be small in the NS-BH and NS-NS binaries that are likely to be observed with first-generation GW detectors. For example, the observed NS-NS binary pulsars have rather small spin, $S_{\rm NS}/m_{\rm NS}^2 \sim 10^{-3}$ [6]. One reason the NSs in binaries of interest for GW detectors should carry small spin is that they are old enough to have spun down considerably (even if they once had spins comparable to the theoretical upper limits, $S_{\rm NS}/m_{\rm NS}^2 \simeq 0.6$–$0.7$ [10], where $m_{\rm NS}$ is the NS mass, and where we set $G=c=1$), and because dynamical evolution cannot spin them up significantly (even during the final phase of inspiral when tidal torques become important [11]).

Population-synthesis studies [12,13] suggest that in NS-BH binaries there is a possibility for the BH spin to be substantially misaligned with the orbital angular momentum of the binary. Early investigations [14,15] showed that when this is the case and the BH spin is large, the evolution of the GW phase and amplitude during the adiabatic inspiral is significantly affected by spin-induced modulations. While reliable templates for precessing binaries should include these modulational effects, performing GW searches with template families that include all the *prima facie* relevant parameters (the masses, the spins, the angles that describe the relative orientations of detector and binary, and the direction of propagation of GWs to the detector) is extremely computationally intensive.

Several authors have explored this issue, and they have proposed detection template families (DTFs) that depend on fewer parameters and that can still reproduce well the expected physical signals. An interesting suggestion, built on the results obtained in Ref. [14], came from Apostolatos [15], who introduced a modulational sinusoidal term (the Apostolatos ansatz) in the frequency-domain phase of the templates to capture the effects of precession. This suggestion was tested further by Grandclément, Kalogera, and Vecchio [16]. The resulting template family has significantly fewer parameters, but its computational requirements are still very high, and its signal-fitting performance is not very satisfactory; Grandclément and Kalogera [17] subsequently suggested a modified family of *spiky* templates that fit the signals better.

After investigating the dynamics of precessing binaries, Buonanno, Chen, and Vallisneri [[18], henceforth BCV2] proposed a new convention for quadrupolar GW emission in such binaries, whereby the oscillatory effects of precession are isolated in the evolution of the GW polarization tensors. As a result, the response of the detector to the GWs can be written as the product of a carrier signal and a modulational correction, which can be handled using an extension of the Apostolatos ansatz. On the basis of these observations, BCV2 built a modulated frequency-domain DTF that, for maximal spins, yields average fitting factors ($\overline{\mathrm{FF}}$, see Sec. VI B of Ref. [18]) of $\simeq 0.97$ for $(7+5)M_\odot$ BH-BH binaries, and $\simeq 0.93$ for $(10+1.4)M_\odot$ NS-BH binaries (see also Tables VIII and IX and Fig. 14 of Ref. [18]). Note that the stationary-phase-approximation (SPA) templates developed for nonspinning binaries give much lower $\overline{\mathrm{FF}}$s of $\simeq 0.90$ for $(7+5)M_\odot$ BH-BH binaries and $\simeq 0.78$ for $(10+1.4)M_\odot$ NS-BH binaries, while according to our computations the Apostolatos templates give $\overline{\mathrm{FF}} \simeq 0.81$ for $(10+1.4)M_\odot$ NS-BH binaries [19].

An important feature of the BCV2 templates is that their mathematical structure allows an automatic search over several of the modulational parameters (in strict analogy to the automatic search over initial orbital phase in GW searches for nonspinning binaries), reducing significantly the number of templates in the search banks, and therefore the computational cost. However, since many more signal shapes are effectively (if implicitly) tested against the detector output, the detection threshold for this DTF should be set higher than those for simpler families (for the same false-alarm probability). According to simple false-alarm computations performed with Gaussian, stationary detector noise (see Ref. [18]) for a single template, the gain in FF is larger than the increase in the threshold only for binaries (such as NS-BH binaries) with low symmetric mass ratios $m_1 m_2/(m_1+m_2)^2$; while the opposite is true for high mass ratios. (Ultimately, the issue of FF gain versus threshold increase will be settled only after constructing the mismatch metric for this template family and performing Monte Carlo analyses of false-alarm statistics for the entire template bank under realistic detector noise.) Although the improvement in FF with the BCV2 DTF is relevant, it is still not completely satisfactory, because it translates to a loss of $\sim 20\%$ in detection rate (for the maximal-spin case) with respect to a perfect template bank (the loss will be higher if the higher required threshold is taken into account). Current estimates of binary-inspiral event rates within the distance accessible to first-generation GW interferometers hovers around one event per year, so a reduction of $\sim 20\%$ in the detection rate may not be acceptable.

BCV2 also proposed, but did not test, a new promising family of *physical* templates (i.e., templates that are exact within the approximations made to write the PN equations) for binaries where only one of the two compact bodies carries a significant spin. This family has two remarkable advantages: (i) it consists only of the physical waveforms predicted by the PN equations in the adiabatic limit, so it does not raise the detection threshold unnecessarily by including unphysical templates, as the BCV2 DTF did; (ii) all the template parameters except four are *extrinsic*: that is, they can be searched over semialgebraically without having to compute all of the corresponding waveforms.

In this paper we describe a data-analysis scheme that employs this family, and we estimate the number of templates required for a NS-BH search with LIGO-I: we assume $1M_\odot < m_{\mathrm{NS}} < 3M_\odot$ and $7M_\odot < m_{\mathrm{BH}} < 12M_\odot$ (see Sec. III D). In a companion paper [20], we show how a simple extension of this template family can be used to search for the GWs emitted by binaries when both compact bodies have significant spins (and where of course the adiabatic limit of the PN equations is still valid). The problem of estimating the parameters of the binaries is examined in a forthcoming paper [21].

This paper is organized as follows. In Sec. II we review the formalism of matched-filtering GW detection, and we establish some notation. In Sec. III we review the PN dynamics and GW generation in single-spin binaries, and we discuss the accuracy of the resulting waveforms, indicating the range of masses to which our physical template family can be applied. In Sec. IV we describe the parametrization of the templates, and we discuss the semialgebraic maximization of signal-template correlations with respect to the extrinsic parameters. In Sec. V we describe and test a fast two-stage detection scheme that employs the templates, and we discuss its false-alarm statistics. In Sec. VI we build the template mismatch metric, and we evaluate the number of templates required for an actual GW search. Finally, in Sec. VII we summarize our conclusions.

II. A BRIEF REFRESHER ON MATCHED-FILTERING GW DETECTION

We refer the reader to Ref. [22] (henceforth BCV1), for a self-contained discussion of matched-filtering techniques for GW detection, which includes all relevant bibliographic references. In this section we mainly establish our notation and conventions; the experienced reader may therefore want to jump ahead to Sec. III.

Matched filtering [23–36] is the standard method to detect GW signals of known shape, whereby we compare the detector output with *templates* that approximate closely the signals expected from a given class of sources, for a variety of source parameters. The goodness of fit between the template $h(\lambda^A)$ (where λ^A denotes all the source parameters) and the real GW signal s is quantified by the *overlap*

$$\rho[s,h(\lambda^A)] = \frac{\langle s,h(\lambda^A)\rangle}{\sqrt{\langle h(\lambda^A),h(\lambda^A)\rangle}} \tag{1}$$

[also known as the *signal-to-noise ratio* after filtering s by $h(\lambda^A)$], where the inner product $\langle g(t),\, h(t)\rangle$ of two real signals with Fourier transforms $\tilde{g}(f)$, $\tilde{h}(f)$ is given by [33]

$$\langle g,h\rangle = 2\int_{-\infty}^{+\infty}\frac{\tilde{g}^*(f)\tilde{h}(f)}{S_n(|f|)}df = 4\,\mathrm{Re}\int_0^{+\infty}\frac{\tilde{g}^*(f)\tilde{h}(f)}{S_n(f)}df; \tag{2}$$

throughout this paper we adopt the LIGO-I one-sided noise power spectral density S_n given by Eq. (28) of BCV1. Except where otherwise noted, we shall always consider normalized templates $\hat{h}$ (where the hat denotes normalization), for which $\langle \hat{h}(\lambda^A),\hat{h}(\lambda^A)\rangle = 1$, so we can drop the denominator of Eq. (1).

A large overlap between a given stretch of detector output and a particular template implies that there is a high probability that a GW signal similar to the template is actually present in the output, and is not being merely simulated by noise alone. Therefore the overlap can be used as a *detection statistic*: we may claim a detection if the overlap rises above a *detection threshold* ρ^*, which is set, on the basis of a characterization of the noise, in such a way that false alarms are sufficiently unlikely.

The maximum (*optimal*) overlap that can be achieved for the signal s is $\sqrt{\langle s,s\rangle}$ (the *optimal signal-to-noise ratio*), which is achieved by a perfect (normalized) template $\hat{h} \equiv s/\sqrt{\langle s,s\rangle}$. In practice, however, this value will not be reached, for two distinct reasons. First, the template family $\{\hat{h}(\lambda^A)\}$ might not contain a faithful representation of the physical signal w. The fraction of the theoretical maximum overlap that is recovered by the template family is quantified by the *fitting factor* [35]

$$\mathrm{FF} = \frac{\max_{\lambda^A}\langle w,\hat{h}(\lambda^A)\rangle}{\sqrt{\langle w,w\rangle}}. \tag{3}$$

Second, in practice we will usually not be able to use a *continuous* template family $\{\hat{h}(\lambda^A)\}$, but instead we will have to settle with a discretized template bank $\{\hat{h}(\lambda^A_{(k)})\}$, where (k) indexes a finite lattice in parameter space; so the best template to match a given physical signal will have to be replaced by a nearby template in the bank. (As we shall see in Sec. IV, there is a partial exception to this rule: we can take into account all possible values of certain parameters, known as *extrinsic parameters* [23,27], without actually laying down templates in the bank along that parameter direction.) The fraction of the optimal overlap that is recovered by the template bank, in the worst possible case, is quantified by the *minimum match* [27,31]. Assuming that the physical signal belongs to the continuous template family $\{\hat{h}(\lambda^A)\}$, the minimum match is equal to

$$\mathrm{MM} = \min_{\lambda'^A}\max_{(k)}\langle \hat{h}(\lambda'^A),\hat{h}(\lambda^A_{(k)})\rangle. \tag{4}$$

The required minimum match MM sets the allowable coarseness of the template bank [23,31,32]: the closer to one MM, the closer to one another the templates will need to be laid down. In Sec. VI we shall use a notion of *metric* [27,29,34] in parameter space to characterize the size and the geometry of the template bank corresponding to a given MM.

III. ADIABATIC POST-NEWTONIAN MODEL FOR SINGLE-SPIN BINARY INSPIRALS

In this section we discuss PN adiabatic dynamics and GW generation for NS-BH and BH-BH binaries. Specifically, in Secs. III A–III C we review the relevant PN equations and the GW emission formalism developed by BCV2. In Sec. III D we study the accuracy of the waveforms, and determine the mass range where the waveforms produced by adiabatic models can be considered accurate for the purpose of GW detection. The discussion of this section is continued in Appendix A, where we investigate the effects of quadrupole-monopole interactions (tidal torques) on the waveforms, which have been so far neglected in studies of precessing-binary waveforms, and in fact turn out to be only marginally important for NS-BH binaries. The time-pressed reader may want to skip Secs. III A and III B (which review BCV2 material), move on to Sec. III C (which introduces the template bank examined in this paper), and then jump to the last two paragraphs of Sec. III D (which summarize the comparisons between PN orders).

In this paper, we restrict ourselves to binaries in which only one body has significant spin, leaving a similar study of generic binaries to a companion paper [21]. As a further restriction, we consider only binaries in circular orbits, assuming that they have already been circularized by radiation reaction as they enter the frequency band of ground-based GW detectors. For all binaries, we denote the total mass by $M = m_1 + m_2$ and the symmetric mass ratio by $\eta = m_1 m_2/M^2$; we also assume that the heavier body (with mass $m_1 \geqslant m_2$) carries the spin $S_1 = \chi_1 m_1^2$, with $0 \leqslant \chi_1 \leqslant 1$ (here and throughout this paper we set $G = c = 1$).

A. The PN dynamical evolution

In the adiabatic approach [37,38,6] to the evolution of spinning binaries, one builds a sequence of precessing (due to spin effects) and shrinking (due to radiation reaction) circular orbits. The orbital frequency increases as the orbit shrinks. The time scales of the precession and shrinkage are both long compared to the orbital period (this is the *adiabatic condition*), until the very late stage of binary evolution. (Such orbits are also called *spherical orbits*, since they reside on a sphere with slowly shrinking radius.)

The radiation-reaction-induced evolution of frequency can be calculated by using the energy-balance equation, $\dot{\omega} = -\mathcal{F}/(d\mathcal{E}/d\omega)$, where $\mathcal{E}$ is the orbital-energy function, and $\mathcal{F}$ the GW energy-flux (or luminosity) function. Both have been calculated as functions of the orbital frequency using PN-expansion techniques, and are determined up to 3.5 PN order [7–9]; however, spin effects have been calculated only up to 2 PN order [37]. The resulting evolution equation for ω, obtained by inserting the PN expansions of $\mathcal{E}$ and $\mathcal{F}$ into the balance equation and reexpanding, is

$$\frac{\dot{\omega}}{\omega^2}=\frac{96}{5}\eta(M\omega)^{5/3}\Bigg\{1-\frac{743+924\eta}{336}(M\omega)^{2/3}-\left(\frac{1}{12}\left[\chi_1(\hat{\mathbf{L}}_N\cdot\hat{\mathbf{S}}_1)\left(113\frac{m_1^2}{M^2}+75\eta\right)\right]-4\pi\right)(M\omega)$$
$$+\left(\frac{34103}{18144}+\frac{13661}{2016}\eta+\frac{59}{18}\eta^2\right)(M\omega)^{4/3}-\frac{1}{672}(4159+14532\eta)\pi(M\omega)^{5/3}$$
$$+\Bigg[\left(\frac{16447322263}{139708800}-\frac{1712}{105}\gamma_E+\frac{16}{3}\pi^2\right)+\left(-\frac{273811877}{1088640}+\frac{451}{48}\pi^2-\frac{88}{3}\hat{\theta}\right)\eta+\frac{541}{896}\eta^2-\frac{5605}{2592}\eta^3$$
$$-\frac{856}{105}\log[16(M\omega)^{2/3}]\Bigg](M\omega)^2+\left(-\frac{4415}{4032}+\frac{661775}{12096}\eta+\frac{149789}{3024}\eta^2\right)\pi(M\omega)^{7/3}\Bigg\}, \tag{5}$$

where $\gamma_E=0.577\ldots$ is Euler's constant. We denote by $\hat{\mathbf{L}}_N \propto \mathbf{r}\times\mathbf{v}$ the unit vector along the orbital angular momentum, where $\mathbf{r}$ and $\mathbf{v}$ are the two-body center-of-mass radial separation and relative velocity, respectively. $\hat{\mathbf{L}}_N$ is also the unit normal vector to the orbital plane. (Throughout this paper we shall always add hats to vectors to denote the corresponding unit vectors.) The (scalar) quantity $\hat{\theta}$ is an undetermined regularization parameter that enters the GW flux at 3 PN order [8]. As in Ref. [18], we do not include the (partial) spin contributions to $\dot{\omega}$ at 2.5 PN, 3 PN, and 3.5 PN orders, which arise from known 1.5 PN and 2 PN spin terms of $\mathcal{E}$ and $\mathcal{F}$. (To be fully consistent one should know the spin terms of $\mathcal{E}$ and $\mathcal{F}$ at 2.5 PN, 3 PN, and 3.5 PN orders.) In Sec. III D we shall briefly comment on the effect of these terms. We ignore also the quadrupole-monopole interaction, which we discuss in Sec. A.

The precession equation for the spin is [38,14]

$$\dot{\mathbf{S}}_1=\frac{\eta}{2M}(M\omega)^{5/3}\left(4+3\frac{m_2}{m_1}\right)\hat{\mathbf{L}}_N\times\mathbf{S}_1, \tag{6}$$

where we have replaced $r\equiv\mathbf{r}$ and $|\mathbf{L}_N|$ by their leading-order Newtonian expressions in ω. The precession of the orbital plane (defined by its normal vector $\hat{\mathbf{L}}_N$) can be computed following Eqs. (5)–(8) of Ref. [18], and it reads

$$\dot{\hat{\mathbf{L}}}_N=\frac{\omega^2}{2M}\left(4+3\frac{m_2}{m_1}\right)\mathbf{S}_1\times\hat{\mathbf{L}}_N\equiv\boldsymbol{\Omega}_L\times\hat{\mathbf{L}}_N. \tag{7}$$

Equations (5)–(7) describe the adiabatic evolution of the three variables ω, $\mathbf{S}_1$, and $\hat{\mathbf{L}}_N$. It can be easily deduced that the magnitude of the spin, $S_1=|\mathbf{S}_1|$, and the angle between the spin and the orbital angular momentum, $\kappa_1\equiv\hat{\mathbf{L}}_N\cdot\hat{\mathbf{S}}_1$, are conserved during the evolution.

The integration of Eqs. (5)–(7) should be stopped at the point where the adiabatic approximation breaks down. This point is usually reached (e.g., for 2 PN and 3 PN orders) when the orbital energy $\mathcal{E}_{n\mathrm{PN}}$ reaches a minimum, as found by solving $d\mathcal{E}_{n\mathrm{PN}}/d\omega=0$ (exceptions occur at Newtonian, 1 PN, and 2.5 PN orders, as we shall explain in more detail in Sec. III D). We shall call the corresponding orbit the minimum energy circular orbit, or MECO. Up to 3 PN order, and including spin-orbit effects up to 1.5 PN order, the orbital energy $\mathcal{E}(\omega)$ reads [6,37,40]

$$\mathcal{E}(\omega)=-\frac{\mu}{2}(M\omega)^{2/3}\Bigg\{1-\frac{(9+\eta)}{12}(M\omega)^{2/3}$$
$$+\frac{8}{3}\left(1+\frac{3}{4}\frac{m_2}{m_1}\right)\frac{\hat{\mathbf{L}}_N\cdot\mathbf{S}_1}{M^2}(M\omega)$$
$$-\frac{1}{24}(81-57\eta+\eta^2)(M\omega)^{4/3}$$
$$+\Bigg[-\frac{675}{64}+\left(\frac{34445}{576}-\frac{205}{96}\pi^2+\frac{10}{3}\omega_s\right)\eta$$
$$-\frac{155}{96}\eta^2-\frac{35}{5184}\eta^3\Bigg](M\omega)^2\Bigg\}. \tag{8}$$

Henceforth, we assume the regularization parameter ω_s to be zero, as computed in Refs. [7,9,39].

B. The precessing convention

BCV2 introduced a new convention to express the leading-order mass-quadrupole gravitational waveform generated by binaries of spinning compact objects; here we review it briefly. At this order, the radiative gravitational field emitted by the quasicircular binary motion reads

$$h^{ij}=\frac{2\mu}{D}\left(\frac{M}{r}\right)Q_c^{ij}, \tag{9}$$

where D is the distance between the source and the Earth, and $Q_c^{ij}=2[\hat{\lambda}^i\hat{\lambda}^j-\hat{n}^i\hat{n}^j]$, with $\hat{\mathbf{n}}$ and $\hat{\boldsymbol{\lambda}}$ the unit vectors along the separation vector $\mathbf{r}$ and the relative velocity $\mathbf{v}$ of the binary. In general, $\hat{\mathbf{n}}(t)$ and $\hat{\boldsymbol{\lambda}}(t)$ can be written as

$$\hat{\mathbf{n}}(t)=\mathbf{e}_1(t)\cos\Phi(t)+\mathbf{e}_2(t)\sin\Phi(t), \tag{10}$$

$$\hat{\boldsymbol{\lambda}}(t)=-\mathbf{e}_1(t)\sin\Phi(t)+\mathbf{e}_2(t)\cos\Phi(t), \tag{11}$$

where $\mathbf{e}_1(t)$, $\mathbf{e}_2(t)$, and $\mathbf{e}_3(t)\equiv\hat{\mathbf{L}}_N(t)$ are orthonormal vectors, and $\mathbf{e}_{1,2}(t)$ forms a basis for the instantaneous orbital plane. Now Q_c^{ij} can be written as

$$Q_c^{ij}=-2\{[\mathbf{e}_+]^{ij}\cos 2(\Phi+\Phi_0)+[\mathbf{e}_\times]^{ij}\sin 2(\Phi+\Phi_0)\}, \tag{12}$$

with $\mathbf{e}_+=\mathbf{e}_1\otimes\mathbf{e}_1-\mathbf{e}_2\otimes\mathbf{e}_2$ and $\mathbf{e}_\times=\mathbf{e}_1\otimes\mathbf{e}_2+\mathbf{e}_2\otimes\mathbf{e}_1$, and Φ_0 an arbitrary initial phase. For quasicircular orbits, we have $\dot{\hat{\mathbf{n}}}=\omega\hat{\boldsymbol{\lambda}}$, but in general $\dot{\Phi}\neq\omega$, because of the time dependence of $\mathbf{e}_{1,2}(t)$. BCV2 defined the *precessing convention* by requiring that

$$\dot{\mathbf{e}}_i(t)=\boldsymbol{\Omega}_e(t)\times\mathbf{e}_i(t),\quad i=1,2, \tag{13}$$

$$\boldsymbol{\Omega}_e(t)=\boldsymbol{\Omega}_L-(\boldsymbol{\Omega}_L\cdot\hat{\mathbf{L}}_N)\hat{\mathbf{L}}_N. \tag{14}$$

[See Eq. (7) for the definition of $\boldsymbol{\Omega}_L$.] In this convention, we do have $\dot{\Phi}=\omega$.

C. The detector response

The response of a ground-based interferometric detector to the GW signal of Eq. (9) is given by

$$h=\underbrace{-\frac{2\mu}{D}\frac{M}{r}([\mathbf{e}_+]^{ij}\cos 2(\Phi+\Phi_0)+[\mathbf{e}_\times]^{ij}\sin 2(\Phi+\Phi_0))}_{\text{factor Q: wave generation}}\underbrace{([\mathbf{T}_+]_{ij}F_++[\mathbf{T}_\times]_{ij}F_\times)}_{\text{factor P: detector projection}}; \tag{15}$$

the tensors $[\mathbf{T}_{+,\times}]_{ij}$ are defined by $\mathbf{T}_+\equiv\mathbf{e}_x^R\otimes\mathbf{e}_x^R-\mathbf{e}_y^R\otimes\mathbf{e}_y^R$ and $\mathbf{T}_\times\equiv\mathbf{e}_x^R\otimes\mathbf{e}_y^R+\mathbf{e}_y^R\otimes\mathbf{e}_x^R$, after we introduce the *radiation frame*

$$\mathbf{e}_x^R=-\mathbf{e}_x^S\sin\varphi+\mathbf{e}_y^S\cos\varphi, \tag{16}$$

$$\mathbf{e}_y^R=-\mathbf{e}_x^S\cos\Theta\cos\varphi-\mathbf{e}_y^S\cos\Theta\sin\varphi+\mathbf{e}_z^S\sin\Theta, \tag{17}$$

$$\mathbf{e}_z^R=+\mathbf{e}_x^S\sin\Theta\cos\varphi+\mathbf{e}_y^S\sin\Theta\sin\varphi+\mathbf{e}_z^S\cos\Theta=\hat{\mathbf{N}}, \tag{18}$$

with $\hat{\mathbf{N}}$ the direction of wave propagation and Θ, φ the corresponding angles in an arbitrarily chosen *source frame*, $\{\mathbf{e}_x^S,\mathbf{e}_y^S,\mathbf{e}_z^S\}$ (see Fig. 1 of Ref. [18]). For the antenna patterns $F_{+,\times}$ we have

$$F_{+,\times}=\frac{1}{2}[\bar{\mathbf{e}}_x\otimes\bar{\mathbf{e}}_x-\bar{\mathbf{e}}_y\otimes\bar{\mathbf{e}}_y]^{ij}[\mathbf{T}_{+,\times}]_{ij}, \tag{19}$$

where $\bar{\mathbf{e}}_{x,y}$ are the unit vectors along the orthogonal interferometer arms. More explicitly [26],

$$F_+=\frac{1}{2}(1+\cos^2\theta)\cos 2\phi\cos 2\psi-\cos\theta\sin 2\phi\sin 2\psi, \tag{20}$$

$$F_\times=\frac{1}{2}(1+\cos^2\theta)\cos 2\phi\sin 2\psi+\cos\theta\sin 2\phi\cos 2\psi. \tag{21}$$

Here ϕ, θ, and ψ are the orientation angles of the detector, as defined by Fig. 2 of Ref. [18].

Mathematically, we see that the factor P of Eq. (15), which is independent of time, collects only terms that depend on the position and orientation of the detector and that describe the *reception* of GWs, while factor Q collects only terms that depend on the dynamical evolution of the binary and that describe the *generation* of GWs (at least if the vectors $\mathbf{e}_{1,2,3}$ are defined without reference to the detector, as we will do soon). Using the language of BCV2, in the precessing convention the *directional* parameters Θ, φ, ϕ, θ, and ψ are isolated in factor P, while the *basic* and *local* parameters of the binary are isolated in factor Q.

Physically, we see that factor Q evolves along three different time scales: (i) the orbital period, which sets the GW *carrier* frequency $2\dot{\Phi}=2\omega$; (ii) the precession time scale at which the $\mathbf{e}_{+,\times}$ change their orientation in space, which modulates the GWs; (iii) the radiation-reaction time scale, characterized by $\omega/\dot{\omega}$, which drives the evolution of frequency. In the adiabatic regime, the orbital period is the shortest of the three: so for convenience we shall define the (leading-order) *instantaneous GW frequency* $f_{\rm GW}$ directly from the instantaneous orbital frequency ω, $f_{\rm GW}\equiv(2\omega)/(2\pi)=\omega/\pi$.

Thus, what parameters are needed to specify Q completely? Equation (5) for $\omega(t)$ can be integrated numerically, starting from an arbitrary $\omega(0)$ [41] once we specify the basic parameters M, η, and χ_1 and the local parameter $\kappa_1\equiv\hat{\mathbf{L}}_N\cdot\hat{\mathbf{S}}_1$ (conserved through evolution). With the resulting $\omega(t)$ we can integrate Eqs. (6) and (7), and then Eq. (13). For these we need initial conditions for $\hat{\mathbf{S}}_1$, $\hat{\mathbf{L}}_{\rm N}$, and for the $\mathbf{e}_i$; without loss of generality, we can introduce a (fixed) source frame attached to the configuration of the binary at $t=0$:

$$\mathbf{e}_x^S\propto\mathbf{S}_1(0)-[\mathbf{S}_1(0)\cdot\hat{\mathbf{L}}_{\rm N}(0)]\hat{\mathbf{L}}_{\rm N}(0),$$

$$\mathbf{e}_y^S=\hat{\mathbf{L}}_{\rm N}(0)\times\mathbf{e}_x^S,\quad \mathbf{e}_z^S=\hat{\mathbf{L}}_{\rm N}(0), \tag{22}$$

and then take

$$\mathbf{e}_1(0)=\mathbf{e}_x^S,\quad \mathbf{e}_2(0)=\mathbf{e}_y^S,\quad \mathbf{e}_3(0)=\mathbf{e}_z^S. \tag{23}$$

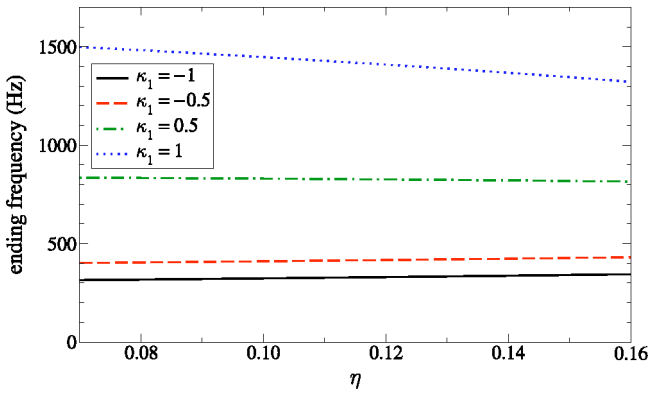

FIG. 1. Ending frequency (instantaneous GW frequency at the MECO) as a function of η, evaluated from Eq. (8) at 2 PN order for $M=15M_\odot$, $\chi_1=1$, and for different values of κ_1.

[If $\mathbf{S}_1(0)$ and $\hat{\mathbf{L}}_{\rm N}(0)$ are parallel, $\mathbf{e}_x^S$ can be chosen to lie in any direction orthogonal to $\hat{\mathbf{L}}_{\rm N}(0)$.] The initial orbital phase Φ_0 that enters the expression of Q is defined by

$$\hat{\mathbf{n}}(0)=\mathbf{e}_1(0)\cos\Phi_0+\mathbf{e}_2(0)\sin\Phi_0, \tag{24}$$

while the initial conditions for $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{L}}_{\rm N}$, as expressed by their components with respect to the source frame, are

$$\hat{\mathbf{L}}_{\rm N}(0)=(0,0,1), \tag{25}$$

$$\hat{\mathbf{S}}_1(0)=(\sqrt{1-\kappa_1^2},0,\kappa_1). \tag{26}$$

BCV2 proposed to use the family of waveforms (detector responses) defined by Eqs. (5)–(7) and (13)–(15) as a family of *physical templates* for compact binaries with a single spin. Depending on the maximum PN order N up to which the terms of Eq. (5) are retained, we shall denote this class of template families ST_N. The ST_N templates deserve to be called *physical* because they are derived from a physical model, namely the adiabatic PN dynamics plus quadrupole GW emission. Each ST_N template family is indexed by *eleven* parameters: M, η, χ_1 (basic), κ_1 (local), Θ, φ, θ, ϕ, ψ (directional), plus the time t_0 at which the instantaneous GW frequency passes through the fiducial value $\omega(0)/\pi$, and the corresponding initial phase Φ_0. Of these, using the distinction between intrinsic and extrinsic parameters introduced in Refs. [27,42] and further discussed by BCV2, the first four are intrinsic parameters: that is, when we search for GWs using ST_N templates, we need to lay down a discrete template bank along the relevant ranges of the intrinsic dimensions. The other seven are extrinsic parameters: that is, their optimal values can be found semialgebraically without generating multiple templates along the extrinsic dimensions (another way of saying this is that the maximization of the overlap over the extrinsic parameters can be incorporated in the detection statistic, which then becomes a function only of the intrinsic parameters). In Sec. IV we shall describe how this maximization over the extrinsic parameters can be achieved in practice.

D. Comparison between different post-Newtonian orders and the choice of mass range

In this section we investigate the range of masses m_1 and m_2 for which the PN-expanded evolution equations (5)–(7) [and therefore the template family (15)] can be considered reliable. As a rule of thumb, we fix the largest acceptable value of the total mass by requiring that the *GW ending frequency* (in our case, the instantaneous GW frequency at the MECO) should not lie in the frequency band of good detector sensitivity for LIGO-I. Considering the results obtained by comparing various nonspinning PN models [24,22], and considering the variation of the ending frequency when spin effects are taken into account [18], we require $M\leqslant 15M_\odot$. In keeping with the focus of this paper on binaries with a single significant spin, we also impose $m_2/m_1\leqslant 0.5$, which constrains the spin of the less massive body to be relatively small (of course, this condition is always satisfied for NS-BH binaries). Population-synthesis calculations [43] suggest that the more massive of the two compact bodies will have the larger spin, since usually it will have been formed first, and it will have been spun up through accretion from the progenitor of its companion. For definiteness, we assume $7M_\odot<m_1<12M_\odot$ and $1M_\odot<m_2<3M_\odot$; the corresponding range of η is 0.07–0.16.

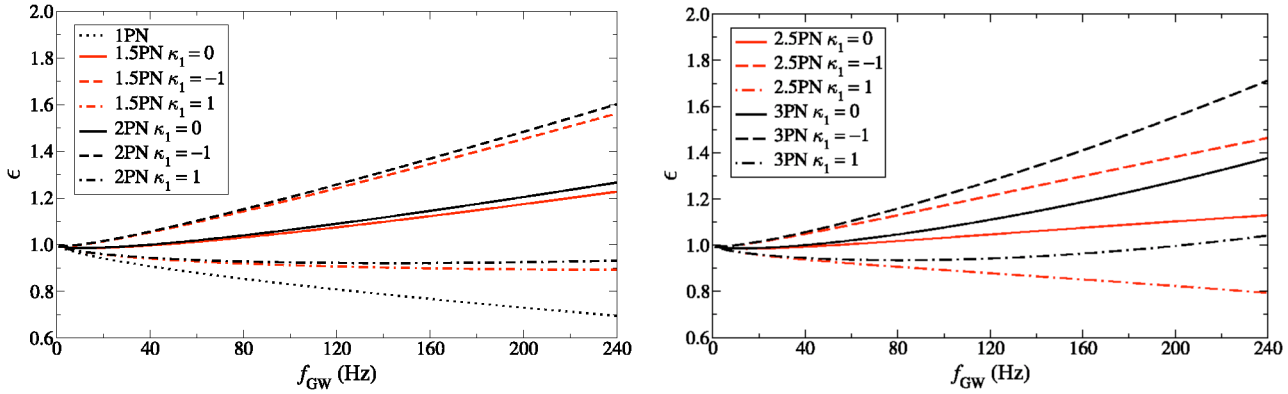

FIG. 2. Plot of $\epsilon\equiv(\dot{\omega}/\omega^2)/[96/5\,\eta(M\omega)^{5/3}]$ as a function of $f_{\rm GW}=\omega/\pi$, evaluated from Eq. (5) at different PN orders for a $(10+1.4)M_\odot$ binary. We do not show the 3.5 PN curves, which are very close to the 3 PN curves.

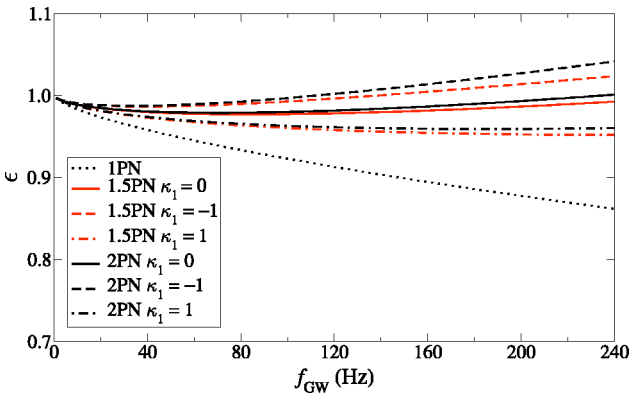

FIG. 3. Plot of $\epsilon \equiv (\dot{\omega}/\omega^2)/[96/5\eta(M\omega)^{5/3}]$ as a function of $f_{GW}=\omega/\pi$, evaluated from Eq. (5) at different PN orders for a $(1.4+1.4)M_\odot$ NS-NS binary. We do not show the 2.5 PN, 3 PN ($\hat{\theta}=0$), and 3.5 PN curves, which are very close to the 2 PN curves. Note the change in scale with respect to Fig. 2.

In Fig. 1 we plot the GW ending frequency as a function of η, evaluated from Eq. (8) at 2 PN order for $M=15M_\odot$ and $\chi_1=1$. The various curves refer to different values of κ_1. The minimum of the GW ending frequency is ~ 300 Hz, and it corresponds to a $(12+1)M_\odot$ binary with spin antialigned with the orbital angular momentum. In Fig. 2 we plot $\dot{\omega}/\omega^2$, normalized to its leading (Newtonian) term $96/5\eta(M\omega)^{5/3}$, as a function of the instantaneous GW frequency; $\dot{\omega}/\omega^2$ is evaluated from Eq. (5) at different PN orders, for a $(10+1.4)M_\odot$ binary with $\chi_1=1$. We see that the effects of the spin-orbit interaction (evident for different κ_1 within the same PN order) are comparable to, or even larger than, the effect of increasing the PN order. We see also that the different PN curves spread out more and more as we increase M and η. For comparison, in Fig. 3 we show the same plot for a $(1.4+1.4)M_\odot$ NS-NS binary; note the different scale on the vertical axis. In this case the various curves remain rather close over the entire frequency band.

Another procedure (often used in the literature) to characterize the effects of spin and PN order on the evolution of the GW frequency is to count the number of GW cycles accumulated within a certain frequency band: $\mathcal{N}_{GW} \equiv (1/\pi)\int_{\omega_{min}}^{\omega_{max}} (\omega/\dot{\omega}) d\omega$. Here we take $\omega_{min}=\pi\times 10$ Hz and $\omega_{max}=\omega_{ISCO}=(6^{3/2}\pi M)^{-1}$, corresponding to the orbital frequency at the innermost stable circular orbit (ISCO) of a Schwarzchild black hole with mass M. In Table I we show $\mathcal{N}_{GW}$ at increasing PN orders for $(10+1.4)M_\odot$, $(12+3)M_\odot$, and $(7+3)M_\odot$ binaries. The contributions in parentheses are partial spin terms present at 2.5 PN, 3 PN, and 3.5 PN orders and due to known 1.5 PN spin terms in the orbital energy and luminosity. These terms were neglected in Eq. (5) to be consistent in including PN terms of the same order, and we list them here only to give their order of magnitude. Unless there are cancellations, the large number of cycles suggests that it is worthwhile to compute spin effects up to the 3.5 PN order.

The number of accumulated GW cycles $\mathcal{N}_{GW}$ can be a useful diagnostic, but taken alone it provides incomplete and sometimes even misleading information. There are three reasons for this. First, $\mathcal{N}_{GW}$ is related only to the number of orbital cycles of the binary *within* the orbital plane, but it does not reflect the precession of the plane, which modulates the detector response in both amplitude and phase. These modulations are very important effects, as witnessed by the fact that neither the standard nonspinning-binary templates (which do not have built-in modulations) nor the original Apostolatos templates (which add only modulations to the phase) can reproduce satisfactorily the detector response to the GWs emitted by precessing binaries. Second, even if two signals have values of $\mathcal{N}_{GW}$ that differ by ~ 1 when ω_{max} equals the GW ending frequency (which apparently represents a total loss of coherence, and hence a significant decrease in overlap), one can always shift their arrival times to obtain higher overlaps. Third, in the context of GW searches the differences in $\mathcal{N}_{GW}$ should be minimized with respect to the search parameters, as done with the fitting factor.

The Cauchy criterion [28] states that the sequence ST_N converges if and only if for every k, $\langle \mathrm{ST}_{N+k}, \mathrm{ST}_N \rangle \to 1$ as $N\to\infty$. In Table II, we test the specific case of $k=0.5$, for maximally spinning and nonspinning $(10+1.4)M_\odot$ and $(12+3)M_\odot$ binaries. The overlaps quoted at the beginning of each column are maximized over the extrinsic parameters t_0 and Φ_0, but not over the five extrinsic directional parameters φ, Θ, θ, ϕ, and ψ or the intrinsic parameters m_1, m_2, χ_1, and κ_1 (we call this the *nonmaximized* overlap). For comparison, we show in parentheses the overlaps maximized over all the parameters of the lower-order family [i.e., the *fitting factors* (FF) for the target family ST_{N+k} as matched by

TABLE I. PN contributions to the number $\mathcal{N}_{GW}$ of GW cycles accumulated from $\omega_{min}=\pi\times 10$ Hz to $\omega_{max}=\omega_{ISCO}=1/(6^{3/2}\pi M)$. The contributions in parentheses, $(\cdots)$, are partial spin terms present at 2.5 PN, 3 PN, and 3.5 PN orders and due to known 1.5 PN spin terms in the orbital energy and luminosity.

	$(10+1.4)M_\odot$	$(12+3)M_\odot$	$(7+3)M_\odot$
Newtonian	3577.0	1522.3	2283.8
1 PN	213.1	114.3	139.0
1.5 PN	$-181.3+114.2\kappa_1\chi_1$	$-99.7+55.7\kappa_1\chi_1$	$-102.3+48.2\kappa_1\chi_1$
2 PN	9.8	6.3	6.4
2.5 PN	$-20.4+(21.1\kappa_1\chi_1)$	$-12.7+(12.1\kappa_1\chi_1)$	$-10.9+(9.0\kappa_1\chi_1)$
3 PN	$2.2+(-17.0\kappa_1\chi_1+2.4\kappa_1^2\chi_1^2)+0.42\hat{\theta}$	$2.2+(-9.7\kappa_1\chi_1+1.2\kappa_1^2\chi_1^2)+0.40\hat{\theta}$	$2.3+(-6.6\kappa_1\chi_1+0.7\kappa_1^2\chi_1^2)+0.43\hat{\theta}$
3.5 PN	$-1.9+(6.4\kappa_1\chi_1)$	$-1.3+(3.8\kappa_1\chi_1)$	$-1.1+(2.6\kappa_1\chi_1)$

TABLE II. Test of Cauchy convergence of the adiabatic templates ST_N at increasing PN orders, for $(10+1.4)M_\odot$ and $(12+3)M_\odot$ binaries, and for maximally spinning BHs ($\chi_1=1$, upper and middle panels) and nonspinning BHs ($\chi_1=0$, lower panel). The overlaps quoted at the beginning of each column are maximized only with respect to the extrinsic parameters t_0 and Φ_0. In parentheses, $(\cdots)$, we give the overlaps maximized over all the parameters of the lower-order family (i.e., the fitting factors FF for the target family ST_{N+k} as matched by the search family ST_N, evaluated at the target masses shown); the fitting factors are obtained by extending the search into the unphysical template region where $\eta>0.25$ and $\chi_1>1$. In brackets, $[\cdots]$, we show the parameters M,η,χ_1,κ_1 (or M,η when $\chi_1=0$) at which the maximum overlap is achieved. The detector is perpendicular to the initial orbital plane, and at 3 PN order we set $\hat{\theta}=0$; in all cases the integration of the equations is started at a GW frequency of 60 Hz. The maximization procedure was stopped whenever an overlap $\geqslant 0.99$ was achieved.

$(N+k,N)$	$\langle ST_{N+k},ST_N\rangle$ for $(10+1.4)M_\odot$ binary, $\chi_1=1$							
	$\kappa_1=0.9$			$\kappa_1=0.5$			$\kappa_1=-0.5$			$\kappa_1=-0.9$	
(1,0)	0.1976	(0.7392)	[24.5,0.02,0.00,0.00]	0.1976	(0.7392)	0.1976	(0.7392)	[24.5,0.02,0.00,−0.00]	0.1976	(0.7392)
(1.5,1)	0.2686	(0.7848)	[4.53,0.54,0.00,0.00]	0.2696	(0.7008)	0.2065	(0.6040)	[6.58,0.36,0.00,−0.00]	0.1800	(0.6255)
(2,1.5)	0.4876	(⩾0.99)	[9.56,0.14,0.83,0.93]	0.5627	(⩾0.99)	0.6623	(⩾0.99)	[11.7,0.10,0.97,−0.50]	0.7728	(0.9760)
(2.5,2)	0.1587	(0.9578)	[10.5,0.13,1.56,0.95]	0.2011	(0.9887)	0.2902	(0.9398)	[10.2,0.13,2.00,−0.19]	0.3460	(⩾0.99)
(3,2)	0.4395	(0.9848)	[11.5,0.10,0.84,0.81]	0.5057	(0.9881)	0.5575	(0.9712)	[12.0,0.10,0.92,−0.48]	0.6606	(⩾0.99)
(3,2.5)	0.1268	(0.9758)	[12.8,0.08,0.05,0.98]	0.1539	(⩾0.99)	0.2520	(0.9744)	[25.6,0.03,0.35,−0.21]	0.2488	(⩾0.99)
(3.5,3)	0.9614	(⩾0.99)	[11.7,0.10,1.00,0.90]	0.9738	(⩾0.99)	0.9907	(⩾0.99)	[11.3,0.11,1.02,−0.49]	0.9939	(⩾0.99)

$(N+k,N)$	$\langle ST_{N+k},ST_N\rangle$ for $(12+3)M_\odot$ binary, $\chi_1=1$							
	$\kappa_1=0.9$			$\kappa_1=0.5$			$\kappa_1=-0.5$			$\kappa_1=-0.9$	
(1,0)	0.2506	(0.7066)	[10.5,0.22,0.00,0.00]	0.2506	(0.7066)	0.2506	(0.7066)	[10.5,0.22,0.00,−0.00]	0.2506	(0.7066)
(1.5,1)	0.3002	(0.7788)	[8.22,0.50,0.00,0.00]	0.2597	(0.7381)	0.2124	(0.6934)	[11.6,0.44,0.00,−0.00]	0.2017	(0.5427)
(2,1.5)	0.6379	(⩾0.99)	[16.0,0.14,1.14,0.92]	0.7089	(⩾0.99)	0.8528	(⩾0.99)	[14.2,0.18,1.14,−0.59]	0.8620	(⩾0.99)
(2.5,2)	0.2039	(0.9397)	[15.4,0.17,1.95,0.87]	0.2800	(0.9863)	0.4696	(0.9756)	[13.5,0.18,1.22,−0.51]	0.4219	(⩾0.99)
(3,2)	0.6679	(0.9851)	[11.0,0.25,0.72,0.84]	0.7267	(⩾0.99)	0.9052	(⩾0.99)	[13.1,0.21,1.50,−0.70]	0.8868	(⩾0.99)
(3,2.5)	0.1603	(⩾0.99)	[18.5,0.10,0.05,0.99]	0.2272	(⩾0.99)	0.3804	(0.9759)	[15.8,0.15,0.94,−0.49]	0.3060	(⩾0.99)
(3.5,3)	0.9517	(⩾0.99)	[15.2,0.15,0.84,0.86]	0.9694	(⩾0.99)	0.9932	(⩾0.99)	[15.3,0.16,1.00,−0.49]	0.9900	(⩾0.99)

$(N+k,N)$	$\langle ST_{N+k},ST_N\rangle$ for $\chi_1=0$					
	$(10+1.4)M_\odot$			$(12+3)M_\odot$		
(1,0)	0.1976	(0.7392)	[24.5,0.02]	0.2506	(0.7066)	[10.5,0.22]
(1.5,1)	0.1721	(0.6427)	[5.22,0.51]	0.2153	(0.6749)	[9.22,0.51]
(2,1.5)	0.7954	(0.9991)	[12.7,0.09]	0.8924	(0.9981)	[16.2,0.14]
(2.5,2)	0.4872	(0.9961)	[6.94,0.25]	0.5921	(0.9977)	[8.05,0.48]
(3,2)	0.7471	(0.9970)	[15.3,0.06]	0.8982	(0.9994)	[19.3,0.10]
(3,2.5)	0.4127	(0.9826)	[26.5,0.02]	0.5282	(0.9783)	[29.0,0.05]
(3.5,3)	0.9931	(⩾0.99)	[11.6,0.11]	0.9924	(⩾0.99)	[15.4,0.15]

the search family ST_N]; we show in brackets the parameters at which the maximum overlaps are achieved. [The overlaps are especially bad when 1 PN and 2.5 PN waveforms are used. These two orders are rather peculiar: the flux function $\mathcal{F}$ can be a decreasing function of ω, and even assume negative values (which is obviously not physical); correspondingly, $\dot{\omega}$ can become negative. Furthermore, the MECO criterion used to set the ending frequency can also fail, because for some configurations the MECO does not exist, or occurs after $\dot{\omega}$ has become negative. To avoid these problems, we stop the numerical integration of the equations of motion when $\dot{\omega}$ decreases to one tenth of its Newtonian value, or at a GW frequency of 1 kHz, whichever comes first. For comparison, in Table II we show also the overlaps between ST_2 and ST_3, which are much higher than those between ST_2 and $ST_{2.5}$ and than those between $ST_{2.5}$ and ST_3.]

While the nonmaximized overlaps can be very low, the FFs are consistently high (note that this requires extending the search into the unphysical template region where $\eta>0.25$ and $\chi_1>1$); however, the best-fit search parameters can be rather different from the target parameters. This suggests that higher-order PN effects can be reabsorbed by a change of parameters, so the ST_N templates can be considered rather reliable for the purpose of detecting GWs from precessing binaries in the mass range examined; however, the estimation of binary parameters can suffer from systematic errors. In the rest of this paper we shall describe and analyze a search scheme that uses the ST_2 template family.

A more thorough analysis of the differences between the various PN orders would be obtained by comparing the PN-expanded adiabatic model used in this paper with PN-resummed adiabatic models (e.g., using the Padé prescription [28]) and nonadiabatic models (e.g., using the effective-one-body model [44]). A similar comparison was carried out for the nonspinning case in Refs. [22,24]. Unfortunately, waveforms that include precessional effects are not yet available for the PN-resummed adiabatic and nonadiabatic models.

IV. A NEW PHYSICAL TEMPLATE FAMILY FOR NS-BH AND BH-BH PRECESSING BINARIES

In this section we discuss the detection of GWs from single-spin precessing binaries using the template family first suggested in Ref. [18], and further discussed in Sec. III. The proposed detection scheme involves the deployment of a discrete template bank along the relevant range of the intrinsic parameters M, η, χ_1, and κ_1 and the use of a detection statistic that incorporates the maximization of the overlap over all the extrinsic parameters: the directional angles Θ, φ, θ, ϕ, and ψ, the time of arrival t_0, and the initial phase Φ_0. In Sec. IV A we describe the reparametrization of the templates used for the formulation of the maximized statistic, which is then discussed in Sec. IV B, where we also present an approximated but computationally cheaper version. The exact and approximated statistics are discussed together in Sec. V in the context of an optimized two-stage detection scheme.

A. Reparametrization of the waveforms

We recall from Eqs. (15)–(21) that the generic functional form of our precessing templates is

$$h[\lambda^A]=Q^{ij}[M,\eta,\chi_1,\kappa_1;\Phi_0,t_0;t]P_{ij}[\Theta,\varphi;\theta,\phi,\psi]. \tag{27}$$

[Please note that for the rest of this paper we shall use coupled raised and lowered indices to denote contraction; however, the implicit metric is always Euclidian, so covariant and contravariant components are equal. This will be true also for the symmetric trace-free (STF) components introduced later, which are denoted by uppercase roman indices.]

The factor $Q^{ij}(t)$ (which describes the time-evolving dynamics of the precessing binary) is given by

$$Q^{ij}=-\frac{2\mu}{D}\frac{M}{r}[[\mathbf{e}_+]^{ij}\cos 2(\Phi+\Phi_0) + [\mathbf{e}_\times]^{ij}\sin 2(\Phi+\Phi_0)], \tag{28}$$

where the GW phase $\Phi(t)$ and the GW polarization tensors $\mathbf{e}_{+,\times}(t)$ evolve according to Eqs. (5), (13), and (14). This factor depends on the intrinsic parameters M, η, χ_1, and κ_1 and on two extrinsic parameters: the initial phase Φ_0 and the time of arrival t_0 of the waveform, referred to a fiducial GW frequency. We can factor out the initial phase Φ_0 by defining

$$Q_0^{ij}\equiv Q^{ij}(\Phi_0=0), \tag{29}$$

$$Q_{\pi/2}^{ij}\equiv Q^{ij}(\Phi_0=\pi/4); \tag{30}$$

we then have

$$Q^{ij}=Q_0^{ij}\cos(2\Phi_0)+Q_{\pi/2}^{ij}\sin(2\Phi_0). \tag{31}$$

The factor P_{ij} (which describes the static relative position and orientation of the detector with respect to the axes initially defined by the binary) is given by

$$P_{ij}=[\mathbf{T}_+]_{ij}F_+ +[\mathbf{T}_\times]_{ij}F_\times, \tag{32}$$

where the detector antenna patterns $F_{+,\times}(\theta,\phi,\psi)$ and the detector polarization tensors $\mathbf{T}_{+,\times}(\Theta,\varphi)$ depend on the orientation angles θ, ϕ, and ψ, and on the position angles Θ and φ, all of them extrinsic parameters. The antenna patterns can be rewritten as

$$\begin{Bmatrix} F_+ \\ F_\times \end{Bmatrix} = \sqrt{F_+^2+F_\times^2}\begin{Bmatrix} \cos\alpha \\ \sin\alpha \end{Bmatrix}; \tag{33}$$

the factor $F\equiv\sqrt{F_+^2+F_\times^2}$ then enters h as an overall multiplicative constant [45]. In what follows we shall be considering normalized signals and templates, where F drops out, so we set $F=1$. We then have

$$P_{ij}=[\mathbf{T}_+]_{ij}\cos\alpha+[\mathbf{T}_\times]_{ij}\sin\alpha. \tag{34}$$

Both $Q^{ij}(t)$ and P_{ij} are three-dimensional STF tensors, with five independent components each. Using an orthonormal STF basis M_{ij}^I, $I=1,\ldots,5$, with $(M^I)_{ij}(M^J)^{ij}=\delta^{IJ}$, we can conveniently express P_{ij} and Q^{ij} in terms of their components on this basis,

$$Q^{ij}=Q^I(M^I)^{ij}, \quad P_{ij}=P^I(M^I)_{ij}, \tag{35}$$

where

$$Q^I=Q^{ij}(M^I)_{ij}, \quad P^I=P_{ij}(M^I)^{ij}. \tag{36}$$

In this paper, we shall adopt a particular orthonormal basis,

$$\begin{aligned}
(M^1)_{ij}&=\sqrt{\frac{4\pi}{15}}(\mathcal{Y}_{ij}^{22}+\mathcal{Y}_{ij}^{2-2}),\\
(M^2)_{ij}&=-i\sqrt{\frac{4\pi}{15}}(\mathcal{Y}_{ij}^{22}-\mathcal{Y}_{ij}^{2-2}),\\
(M^3)_{ij}&=-\sqrt{\frac{4\pi}{15}}(\mathcal{Y}_{ij}^{21}-\mathcal{Y}_{ij}^{2-1}),\\
(M^4)_{ij}&=i\sqrt{\frac{4\pi}{15}}(\mathcal{Y}_{ij}^{21}+\mathcal{Y}_{ij}^{2-1}),\\
(M^5)_{ij}&=-\sqrt{\frac{8\pi}{15}}\mathcal{Y}_{ij}^{20},
\end{aligned} \tag{37}$$

with $\mathcal{Y}_{ij}^{2m}$ defined by

$$\mathcal{Y}_{ij}^{2m}q^iq^j\equiv Y^{2m}(\hat{\mathbf{q}}), \tag{38}$$

where $Y^{2m}(\hat{\mathbf{q}})$, $m=-2,\ldots,2$ are the usual $l=2$ spherical harmonics and $\hat{\mathbf{q}}$ is any unit vector. We bring together this result with Eqs. (31) and (34) to write the final expression

$$h=P_I[Q_0^I\cos(2\Phi_0)+Q_{\pi/2}^I\sin(2\Phi_0)], \tag{39}$$

where

$$P_I(\Theta,\varphi,\alpha)=\{[\mathbf{T}_+(\Theta,\varphi)]_I\cos\alpha+[\mathbf{T}_\times(\Theta,\varphi)]_I\sin\alpha\}. \tag{40}$$

Henceforth, we shall denote the surviving extrinsic parameters collectively as $\Xi^\alpha\equiv(t_0,\Phi_0,\alpha,\Theta,\varphi)$, and the intrinsic parameters as $X^i\equiv(M,\eta,\chi_1,\kappa_1)$.

B. Maximization of the overlap over the extrinsic parameters

As we have anticipated, it is possible to maximize the overlap $\rho=\langle s,\hat{h}\rangle$ *semialgebraically* over the extrinsic directional parameters Θ, φ, θ, ϕ, and ψ, without computing the full representation of $\hat{h}$ for each of their configurations. In addition, it is possible to maximize efficiently also over t_0 and Φ_0, which are routinely treated as extrinsic parameters in nonspinning-binary GW searches.

For a given stretch of detector output s, and for a particular set of template intrinsic parameters $X^i=(M,\eta,\chi_1,\kappa_1)$, we denote the fully maximized overlap as

$$\rho_{\Xi^\alpha}\equiv\max_{\Xi^\alpha}\langle s,\hat{h}(X^i,\Xi^\alpha)\rangle=\max_{t_0,\Phi_0,\Theta,\varphi,\alpha}\left[\underbrace{\frac{P_I[\langle s,Q_0^I\rangle_{t_0}\cos 2\Phi_0+\langle s,Q_{\pi/2}^I\rangle_{t_0}\sin 2\Phi_0]}{\sqrt{P_IP_J\langle Q_0^I\cos 2\Phi_0+Q_{\pi/2}^I\sin 2\Phi_0,Q_0^J\cos 2\Phi_0+Q_{\pi/2}^J\sin 2\Phi_0\rangle}}}_{\rho}\right], \tag{41}$$

where the subscript t_0 denotes the dependence of the signal-template inner products on the time-of-arrival parameter of the templates. In fact, each of these inner products can be computed simultaneously for all t_0 with a single FFT; in this sense, t_0 is an extrinsic parameter [46].

Let us now see how to deal with Φ_0. We start by making an approximation that will be used throughout this paper. We notice that the template components $P_IQ_0^I$ and $P_IQ_{\pi/2}^I$ [Eqs. (29) and (30)] are nearly orthogonal, and have approximately the same signal power,

$$\langle P_IQ_0^I,P_JQ_{\pi/2}^J\rangle\simeq 0, \tag{42}$$

$$\langle P_IQ_0^I,P_JQ_0^J\rangle\simeq\langle P_IQ_{\pi/2}^I,P_JQ_{\pi/2}^J\rangle; \tag{43}$$

this is accurate as long as the time scales for the radiation-reaction-induced evolution of frequency and for the precession-induced evolution of phase and amplitude modulations are both much longer than the orbital period. More precisely, Eqs. (42) and (43) are valid up to the leading-order stationary-phase approximation. Under this hypothesis Eq. (41) simplifies, and its maximum over Φ_0 is found easily:

$$\rho_{\Xi^\alpha}=\max_{t_0,\Phi_0,\Theta,\varphi,\alpha}\frac{P_I[\langle s,Q_0^I\rangle\cos 2\Phi_0+\langle s,Q_{\pi/2}^I\rangle\sin 2\Phi_0]}{\sqrt{P_IP_J\langle Q_0^I,Q_0^J\rangle}}$$

$$=\max_{t_0,\Theta,\varphi,\alpha}\sqrt{\frac{P_IP_JA^{IJ}}{P_IP_JB^{IJ}}}\equiv\max_{t_0,\Theta,\varphi,\alpha}\rho_{\Phi_0}, \tag{44}$$

where we have defined the two matrices

$$A^{IJ}\equiv\langle s,Q_0^I\rangle_{t_0}\langle s,Q_0^J\rangle_{t_0}+\langle s,Q_{\pi/2}^I\rangle_{t_0}\langle s,Q_{\pi/2}^J\rangle_{t_0},$$

$$B^{IJ}\equiv\langle Q_0^I,Q_0^J\rangle, \tag{45}$$

which are functions only of the intrinsic parameters (and, for A^{IJ}, of t_0). We have tested the approximations (42) and (43) by comparing the maximized overlaps obtained from Eq. (44) with the results of full numerical maximization without approximations; both the values and the locations of the maxima agree to one part in a thousand, even for systems with substantial amplitude and phase modulations, where the approximations are expected to be least accurate.

Although Eq. (44) looks innocent enough, the maximization of ρ_{Φ_0} is not a trivial operation. The five components of P_I in Eq. (44) are not all independent, but they are specific functions of only three parameters, Θ, φ, and α [see the discussion leading to Eqs. (34) and (40)]. We can therefore think of ρ_{Ξ^α} as the result of maximizing ρ_{Φ_0} with respect to the five-dimensional vector P_I, *constrained to the three-dimensional physical submanifold* $P_I(\Theta,\varphi,\alpha)$. We shall then refer to ρ_{Ξ^α} as the *constrained* maximized overlap.

What is the nature of the constraint surface? We can easily find the two constraint equations that define it. First, we notice from Eqs. (41) and (44) that the magnitude of the vector P_I does not affect the overlap: so we may rescale P_I and set one of the constraints as $P_IP^I=1$; even better, we may require that the denominator of Eq. (44) be unity, $P_IP_JB^{IJ}=1$. Second, we remember that P_{ij} [Eq. (32)] is the polarization tensor for a plane GW propagating along the direction vector

$$\hat{N}^i=(\sin\Theta\cos\varphi,\sin\Theta\sin\varphi,\cos\Theta). \tag{46}$$

Because GWs are transverse, P_{ij} must admit $\hat{N}^i$ as an eigenvector with null eigenvalue; it follows that

$$\det P_{ij}=0. \tag{47}$$

This equation can be turned into the second constraint for the P_I [see Eq. (B6) of Appendix B].

Armed with the two constraint equations, we can reformulate our maximization problem using the method of Lagrangian multipliers [Eq. (B7) in Appendix B]. However, the resulting system of cubic algebraic equations does not appear to have closed-form analytic solutions. In Appendix B we develop an iterative algebraic procedure to solve the system, obtaining the constrained maximum and the corresponding P_I. In practice, we have found it operationally more robust to use a closed-form expression for the partial maximum over Φ_0 and α (which seems to be the farthest we can go analytically), and then feed it into a numerical maximum-finding routine (such as the well-known AMOEBA [47]) to explore the (Θ,φ) sphere, repeating this procedure for all t_0 to obtain the full maximum.

To maximize ρ_{Φ_0} over α, we use Eq. (40) to factor out the dependence of the P_I on α, and write

$$\sqrt{\frac{P_I P_J A^{IJ}}{P_I P_J B^{IJ}}} = \sqrt{\frac{\mathbf{u}\mathbf{A}_\alpha \mathbf{u}^T}{\mathbf{u}\mathbf{B}_\alpha \mathbf{u}^T}}, \tag{48}$$

where $\mathbf{u}$ is the two-dimensional row vector $(\cos\alpha,\sin\alpha)$, and where $\mathbf{A}_\alpha$ and $\mathbf{B}_\alpha$ are the 2×2 matrices

$$\mathbf{A}_\alpha = A^{IJ}\begin{pmatrix} [\mathbf{T}_+]_I[\mathbf{T}_+]_J & [\mathbf{T}_+]_I[\mathbf{T}_\times]_J \\ [\mathbf{T}_+]_I[\mathbf{T}_\times]_J & [\mathbf{T}_\times]_I[\mathbf{T}_\times]_J \end{pmatrix}, \tag{49}$$

$$\mathbf{B}_\alpha = B^{IJ}\begin{pmatrix} [\mathbf{T}_+]_I[\mathbf{T}_+]_J & [\mathbf{T}_+]_I[\mathbf{T}_\times]_J \\ [\mathbf{T}_+]_I[\mathbf{T}_\times]_J & [\mathbf{T}_\times]_I[\mathbf{T}_\times]_J \end{pmatrix}; \tag{50}$$

in these definitions we sum over the indices I and J. The matrices $\mathbf{A}_\alpha$ and $\mathbf{B}_\alpha$ are implicitly functions of the angles Θ and φ through the polarization tensors $\mathbf{T}_+$ and $\mathbf{T}_\times$. It is straightforward to maximize Eq. (48) over α, yielding [48]

$$\rho_{\Xi^\alpha} = \max_{t_0,\Theta,\varphi} \sqrt{\max \mathrm{eigv}[\mathbf{A}_\alpha \mathbf{B}_\alpha^{-1}]} \equiv \max_{t_0,\Theta,\varphi} \rho_{\Phi_0,\alpha}. \tag{51}$$

The overlap $\rho_{\Phi_0,\alpha}$ is essentially equivalent to the $\mathcal{F}$ statistic used in the search of GWs from pulsars [49].

The last step in obtaining ρ_{Ξ^α} is to maximize $\rho_{\Phi_0,\alpha}$ numerically over the (Θ,φ) sphere, repeating this procedure for all t_0 to obtain the full maximum. Now, t_0 enters Eq. (51) only through the ten signal-template inner products $\langle s, Q^I_{0,\pi/2}\rangle$ contained in $\mathbf{A}_\alpha$, and each such product can be computed for all t_0 with a single FFT. Even then, the semi-algebraic maximization procedure outlined above can still be very computationally expensive if the search over Θ and φ has to be performed for each individual t_0. We have been able to reduce computational costs further by identifying a rapidly computed, fully algebraic statistic ρ'_{Ξ^α} that approximates ρ_{Ξ^α} from above. We then economize by performing the semialgebraic maximization procedure only for the values of t_0 for which ρ'_{Ξ^α} rises above a certain threshold; furthermore, the location of the approximated maximum provides good initial guesses for Θ and φ, needed to kickstart their numerical maximization.

Quite simply, our fast approximation consists in neglecting the functional dependence of the P_I on the directional parameters, computing instead the maximum of ρ_{Φ_0} [Eq. (44)] as if the five P_I were free parameters. Using the method of Lagrangian multipliers outlined in the beginning of Appendix B [Eqs. (B3)–(B5)], we get

$$\rho'_{\Xi^\alpha} = \max_{P_I} \sqrt{\frac{P_I P_J A^{IJ}}{P_I P_J B^{IJ}}} = \sqrt{\max \mathrm{eigv}[\mathbf{A}\mathbf{B}^{-1}]}, \tag{52}$$

with

$$(A^{IJ} - \lambda B^{IJ}) P_J = 0, \quad \lambda = \max \mathrm{eigv}[\mathbf{A}\mathbf{B}^{-1}]. \tag{53}$$

Here the prime stands for *unconstrained* maximization over P_I. We shall henceforth refer to ρ'_{Ξ^α} as the *unconstrained maximum*.

Note that the value of the P_I at the unconstrained maximum will not in general correspond to a physical set of directional parameters, so P_{ij} will not admit any direction vector $\hat{N}^i$ [Eq. (46)] as a null eigenvector. However, we can still get approximate values of Θ and φ by using instead the eigenvector of P_{ij} with the smallest eigenvalue (in absolute value).

V. DESCRIPTION AND TEST OF A TWO-STAGE SEARCH SCHEME

In Sec. IV we have described a robust computational procedure to find the maximum overlap ρ_{Ξ^α} (which is maximized over the extrinsic parameters Φ_0, t_0, and P_I, where the allowed values of the P_I are constrained by their functional dependence on the directional angles). We have also established a convenient analytic approximation for ρ_{Ξ^α}, the unconstrained maximized overlap ρ'_{Ξ^α} (which is maximized over the extrinsic parameters Φ_0, t_0, and P_I, but where the P_I are treated as five independent and unconstrained coefficients). Because the unconstrained maximization has access to a larger set of P_I, it is clear that $\rho'_{\Xi^\alpha} > \rho_{\Xi^\alpha}$. Still, at least when the target signal s is very close to the template $h(X_i)$, we expect ρ'_{Ξ^α} to be a very good approximation for ρ_{Ξ^α}.

A quick look at the procedures outlined in Sec. IV shows that, for the filtering of experimental data against a discrete bank of templates $\{h(X^i_{(k)})\}$, the computation of ρ'_{Ξ^α} is going to be much faster than the computation of ρ_{Ξ^α}. Under these conditions, it makes sense to implement a two-stage search scheme where the discrete bank $\{h(X^i_{(k)})\}$ is first reduced by selecting the templates that have high ρ'_{Ξ^α} against the experimental data; at this stage we identify also the promising times of arrival t_0. The exact ρ_{Ξ^α} is computed only for these first-stage triggers, and compared with the detection threshold ρ^* to identify detection candidates (one would use the same threshold ρ^* in the first stage to guar-

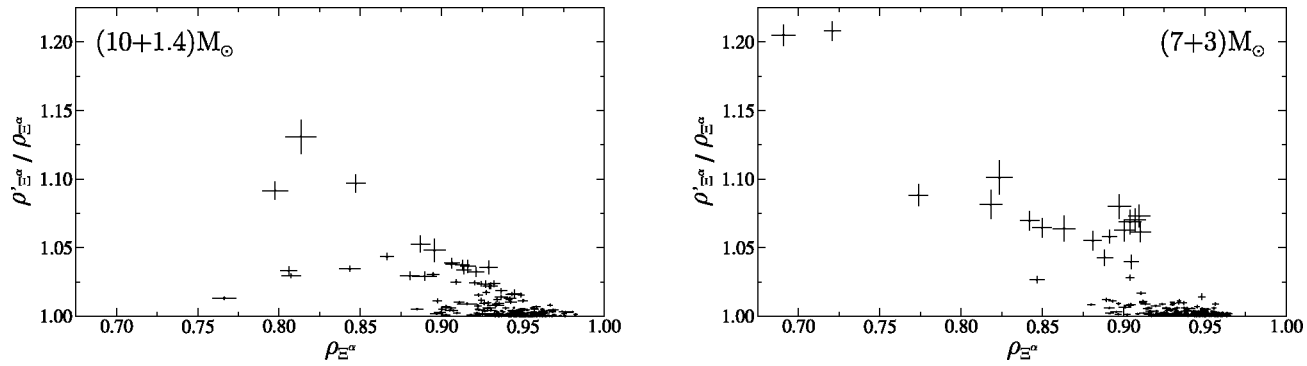

FIG. 4. Ratio between the unconstrained (ρ'_{Ξ^α}) and constrained (ρ_{Ξ^α}) maximized overlaps, as a function of ρ_{Ξ^α}. Each point corresponds to one out of 20×50 sets of intrinsic parameters for target signal and template, and is averaged over 100 sets of extrinsic parameters for the target signal. The error bars show the standard deviations of the sample means (the standard deviations of the samples themselves will be 10 times larger, since we sample 100 sets of extrinsic parameters). The two panels show results separately for $(10+1.4)M_\odot$ (left) and $(7+3)M_\odot$ target systems (right). The few points scattered toward higher ratios and lower ρ_{Ξ^α} are obtained when the first set of extrinsic parameters happens to yield a high ρ'_{Ξ^α} that is not representative of most other values of the extrinsic parameters; then the magnitude of the intrinsic-parameter deviation is set too high, and the comparison between ρ'_{Ξ^α} and ρ_{Ξ^α} is done at low ρ_{Ξ^α}, where the unconstrained maximized overlap is a poor approximation for its constrained version.

antee that all the detection candidates will make it into the second stage) [50].

To prove the viability of such a search scheme, we shall first establish that ρ'_{Ξ^α} is a good approximation for ρ_{Ξ^α} for target signals and templates computed using the adiabatic model of Sec. III. We will take slightly displaced intrinsic parameters for target signals and templates, to reproduce the experimental situation where we are trying to detect a signal of arbitrary physical parameters with the closest template belonging to a discrete bank. This first test is described in Sec. V A. We shall then study the false-alarm statistics of ρ_{Ξ^α} and ρ'_{Ξ^α}, and we shall show that, for a given detection threshold, the number of first-stage triggers caused by pure noise is only a few times larger than the number of *bona fide* second-stage false alarms. Such a condition is necessary because the two-stage detection scheme is computationally efficient only if few templates need ever be examined in the expensive second stage. The false-alarm statistics (in Gaussian stationary noise) are obtained in Sec. V B, and the second test is described in Sec. V C.

A. Numerical comparison of constrained and unconstrained maximized overlaps

In this section we describe a set of Monte Carlo runs designed to test how well ρ'_{Ξ^α} can approximate ρ_{Ξ^α}, for the target signals and templates computed using the adiabatic model of Sec. III, for typical signal parameters, and for signal-template parameter displacements characteristic of an actual search. We choose target signals with 20 different sets of intrinsic parameters given by

$$(m_1,m_2,\chi_1,\kappa_1)=\left\{\begin{matrix}(10,1.4)M_\odot\\(7,3)M_\odot\end{matrix}\right\}\times\left\{\begin{matrix}0.5\\1\end{matrix}\right\}\times\left\{\begin{matrix}-0.9\\-0.5\\0.0\\0.5\\0.9\end{matrix}\right\}. \tag{54}$$

For each set of target-signal intrinsic parameters, we choose 100 random sets of extrinsic parameters $(\Theta,\varphi,\alpha,\Phi_0)$, where the combination (Θ,φ) is distributed uniformly on the solid angle, and where α and Φ_0 are distributed uniformly in the $[0,2\pi]$ interval. The target signals are normalized, so the allowed range for ρ_{Ξ^α} and ρ'_{Ξ^α} is $[0,1]$.

For each target signal, we test 50 (normalized) templates displaced in the intrinsic-parameter space (M,η,χ_1,κ_1) (the optimal extrinsic parameters will be determined by the optimization of ρ_{Ξ^α} and ρ'_{Ξ^α}, so we do not need to set them). The direction of the displacement is chosen randomly in the (M,η,χ_1,κ_1) space. For simplicity, the magnitude of the displacement is chosen so that, for each set of target-signal intrinsic parameters and for the *first set* of target-signal extrinsic parameters, the overlap ρ'_{Ξ^α} is about 0.95; the magnitude is then kept fixed for the other 99 extrinsic-parameter sets, so ρ'_{Ξ^α} can be very different in those cases.

Figure 4 shows the ratio $\rho'_{\Xi^\alpha}/\rho_{\Xi^\alpha}$, for each pair ($20\times 50$ in total) of target and template intrinsic-parameter points, averaged over the 100 target extrinsic-parameter points, as a function of the averaged ρ_{Ξ^α}. The ρ'_{Ξ^α} get closer to the ρ_{Ξ^α} as the latter get higher; most important, the difference is within $\sim$ 2% when $\rho_{\Xi^\alpha}>0.95$, which one would almost certainly want to achieve in an actual search for signals. We conclude that ρ'_{Ξ^α} can indeed be used as an approximation for ρ_{Ξ^α} in the first stage of a two-stage search. The second stage is still necessary, because the false-alarm statistics are worse for the unconstrained maximized overlap (where more degrees of freedom are available) than for its constrained version. We will come back to this in the next two sections.

It is also interesting to compare the set of extrinsic parameters of the target signal with the set of extrinsic parameters that maximize ρ_{Ξ^α}, as characterized by the corresponding source direction vectors, $\hat{\mathbf{N}}_{\rm true}$ and $\hat{\mathbf{N}}_{\rm max}$, respectively. Figure 5 shows the inner product $\hat{\mathbf{N}}_{\rm true}\cdot\hat{\mathbf{N}}_{\rm max}$, averaged over the 100

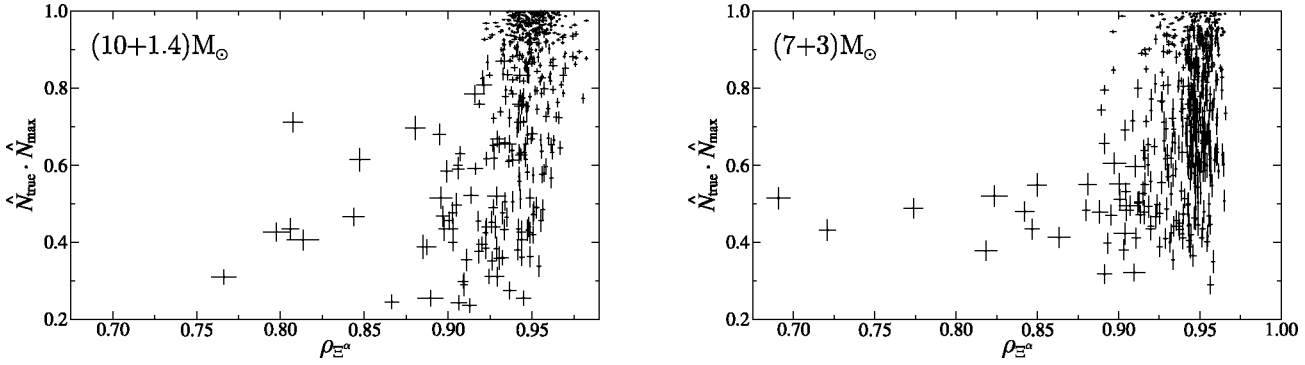

FIG. 5. Inner product between target-signal source direction $\hat{\mathbf{N}}_{\rm true}$ and ρ_{Ξ^α}-maximizing source direction $\hat{\mathbf{N}}_{\rm max}$, as a function of ρ_{Ξ^α}. Each point corresponds to one out of 20×50 sets of intrinsic parameters for target signal and template, and is averaged over 100 sets of extrinsic parameters for the target signal. Standard deviations of the sample means are shown as error bars, as in Fig. 4. The two panels show separately $(10+1.4)M_\odot$ target systems (left) and $(7+3)M_\odot$ target systems (right).

target extrinsic-parameter points, as a function of the averaged ρ_{Ξ^α}. The difference between the vectors can be very large, even when $\rho_{\Xi^\alpha}>0.95$: this happens because the intrinsic-parameter displacement between target signal and template can be compensated by a change in the extrinsic parameters of template (in other words, the effects of the intrinsic and extrinsic parameters on the waveforms are highly correlated).

B. False-alarm statistics for the constrained and unconstrained maximized overlaps

In this section we derive and compare the false-alarm statistics of ρ_{Ξ^α} and ρ'_{Ξ^α}. Our purpose is to estimate the number of additional triggers that are caused by replacing the detection statistic ρ_{Ξ^α} by the first-stage statistic ρ'_{Ξ^α}. Our two-stage detection scheme, which employs the rapidly computed ρ'_{Ξ^α} to choose candidates for the more computationally expensive ρ_{Ξ^α}, will be viable only if the number of those candidates is small enough.

By definition, a false alarm happens when, with interferometer output consisting of pure noise, the detection statistic computed for a given template happens to rise above the detection threshold. Although the detection statistics ρ_{Ξ^α} and ρ'_{Ξ^α} include maximization over the time of arrival t_0, we find it convenient to exclude t_0 from this computation, and to include it later when we evaluate the total false-alarm probability for all the templates in the bank.

Recall that ρ_{Ξ^α} [Eq. (44)] and ρ'_{Ξ^α} [Eq. (52)] are functions of the matrices **A** and **B**, which contain the inner products $\langle s,Q^I_{0,\pi/2}\rangle$ and $\langle Q^I_{0,\pi/2},Q^J_{0,\pi/2}\rangle$, respectively. In this case the signal s is a realization of the noise, n. We combine the vectors Q^I_0 and $Q^I_{\pi/2}$ together as $Q^{\mathcal{I}}$ with $\mathcal{I}=1,\ldots,10$; under the assumption of Gaussian stationary noise, $Y^{\mathcal{I}}\equiv\langle n,Q^{\mathcal{I}}\rangle$ is a ten-dimensional Gaussian random vector with zero mean and covariance matrix [36]

$$C^{\mathcal{I}\mathcal{J}}=\overline{\langle n,Q^{\mathcal{I}}\rangle\langle n,Q^{\mathcal{J}}\rangle}=\langle Q^{\mathcal{I}},Q^{\mathcal{J}}\rangle. \tag{55}$$

The covariance matrix $C^{\mathcal{I}\mathcal{J}}$ specifies completely the statistical properties of the random vector $Y^{\mathcal{I}}$, and it is a function only of **B**, and therefore only of the intrinsic parameters of the template. We can also combine $P_I\cos 2\Phi_0$ and $P_I\sin 2\Phi_0$ together as $P_{\mathcal{I}}$, and then write the maximized overlaps ρ_{Ξ^α} and ρ'_{Ξ^α} as

$$\max_{P_{\mathcal{I}}}\frac{P_{\mathcal{I}}\langle n,Q^{\mathcal{I}}\rangle}{\sqrt{P_{\mathcal{I}}P_{\mathcal{J}}\langle Q^{\mathcal{I}},Q^{\mathcal{J}}\rangle}}=\max_{P_{\mathcal{I}}}\frac{P_{\mathcal{I}}Y^{\mathcal{I}}}{\sqrt{P_{\mathcal{I}}P_{\mathcal{J}}C^{\mathcal{I}\mathcal{J}}}}, \tag{56}$$

where maximization is performed over the appropriate range of the $P_{\mathcal{I}}$. In the rest of this section we shall use the shorthand ρ to denote both ρ_{Ξ^α} and ρ'_{Ξ^α}.

Equation (56) is very general: it describes ρ_{Ξ^α} and ρ'_{Ξ^α}, but it can also incorporate other maximization ranges over the $P_{\mathcal{I}}$, and it can even treat different template families. In fact, the maximized detection statistic for the $(\psi_0\psi_{3/2}\mathcal{B})_6$ DTF of Ref. [18] can be put into the same form, with $P_{\mathcal{I}}\equiv\alpha_{\mathcal{I}}$, for $\mathcal{I}=1,\ldots,6$, and with completely unconstrained maximization.

We can now generate a distribution of the detection statistic ρ for a given set of intrinsic parameters by generating a distribution of the Gaussian random vector $Y^{\mathcal{I}}$, and then computing ρ from Eq. (56). The first step is performed easily by starting from ten independent Gaussian random variables $Z^{\mathcal{I}}$ of zero mean and unit variance, and then setting $Y^{\mathcal{I}}=[\sqrt{C}]^{\mathcal{I}\mathcal{J}}Z_{\mathcal{J}}$ [51]. Thus, there is no need to generate actual realizations of the noise as time series, and no need to compute the inner products $\langle n,Q^{\mathcal{I}}\rangle$ explicitly.

The statistics ρ [Eq. (56)] are homogeneous with respect to the vector $Z^{\mathcal{I}}$: that is, if we define $Z^{\mathcal{I}}=r\hat{Z}^{\mathcal{I}}$ (where $r\equiv\sqrt{Z^{\mathcal{I}}Z_{\mathcal{I}}}$ and $\hat{Z}^{\mathcal{I}}\hat{Z}_{\mathcal{I}}=1$) we have

$$\rho[Y^{\mathcal{I}}(Z^{\mathcal{I}})]=r\rho[Y^{\mathcal{I}}(\hat{Z}^{\mathcal{I}})]\equiv r\rho_1(\Omega); \tag{57}$$

here Ω represents the direction of $\hat{Z}^{\mathcal{I}}$ in its ten-dimensional Euclidian space. The random variable r has the marginal probability density

$$p_r(r)=\frac{r^{\nu-1}\exp(-r^2/2)}{2^{\nu/2-1}\Gamma(\nu/2)}, \tag{58}$$

where the direction Ω is distributed uniformly over a tensphere. (For the rest of this section we shall write equations in the general ν-dimensional case; the special case of our template family is recovered by setting $\nu=10$.) The random variables r and Ω [and therefore $\rho_1(\Omega)$] are statistically independent, so the cumulative distribution function for the statistic ρ is given by the integral

$$P(\rho<\rho^*)=\int d\Omega \int_0^{\rho^*/\rho_1(\Omega)} p_r(r)dr \Big/ \int d\Omega$$
$$=1-\int \frac{\Gamma[\nu/2\,,[\rho^*/\rho_1(\Omega)]^2/2]}{\Gamma[\nu/2]}d\Omega \Big/ \int d\Omega, \tag{59}$$

where $\Gamma[k,z]=\int_z^{+\infty} t^{k-1}e^{-t}dt$ is the *incomplete gamma function*.

The false-alarm probability for a single set of intrinsic parameters and for a single time of arrival is then $1-P(\rho<\rho^*)$. The final integral over the ν-dimensional solid angle can be performed by Monte Carlo integration, averaging the integrand over randomly chosen directions Ω. Each sample of the integrand is obtained by generating a normalized $\hat{Z}^{\mathcal{I}}$ (that is, a direction Ω), obtaining the corresponding $Y^{\mathcal{I}}$, computing $\rho_1(\Omega)$ from Eq. (56), and finally plugging $\rho_1(\Omega)$ into the Γ function.

Equation (59) shows that if we set $\rho_1(\Omega)=1$, the random variable ρ follows the $\chi_{(\nu)}$ distribution; this is obvious because in that case $\rho=r=\sqrt{Z^{\mathcal{I}}Z_{\mathcal{I}}}$ [see Eq. (57)], where the $Z^{\mathcal{I}}$ are ν-independent Gaussian random variables. In fact, $\rho_1(\Omega)$ can be written as

$$\rho_1(\Omega)=\max_{P^{\mathcal{I}}}\frac{R_{\mathcal{I}}\hat{Z}^{\mathcal{I}}}{\sqrt{R_{\mathcal{I}}R_{\mathcal{J}}\delta^{\mathcal{I}\mathcal{J}}}}, \quad R_{\mathcal{I}}=[\sqrt{C}]_{\mathcal{I}\mathcal{J}}P^{\mathcal{J}} \tag{60}$$

which shows that $\rho_1(\Omega)=1$ uniformly for every Ω if and only if the range of maximization for $P_{\mathcal{I}}$ is the *entire* ν-dimensional linear space generated by the basis $\{Q^{\mathcal{I}}\}$; however, once we start using the entire linear space, the particular basis used to generate it ceases to be important, so the covariance matrix $C^{\mathcal{I}\mathcal{J}}$ drops out of the equations for the false-alarm probabilities. That is the case, for instance, for the $(\psi_0\psi_{3/2}\mathcal{B})_6$ DTF (see Sec. V B of Ref. [18]), whose false-alarm probability is described by the $\chi_{(\nu=6)}$ distribution. For our template family $\nu=10$, but both ρ_{Ξ^α} and ρ'_{Ξ^α} have very restrictive maximization ranges for $P_{\mathcal{I}}$ (because $P_{\mathcal{I}=1,\ldots,5}$ and $P_{\mathcal{I}=6,\ldots,10}$ are strongly connected): so both ρ_{Ξ^α} and ρ'_{Ξ^α} will have much lower false-alarm probability, for the same threshold ρ^*, than suggested by the $\chi_{(\nu=10)}$ distribution. In fact, in the next section we shall see that the effective ν for the detection statistic ρ'_{Ξ^α} is about 6, while the effective ν for ρ_{Ξ^α} is even lower.

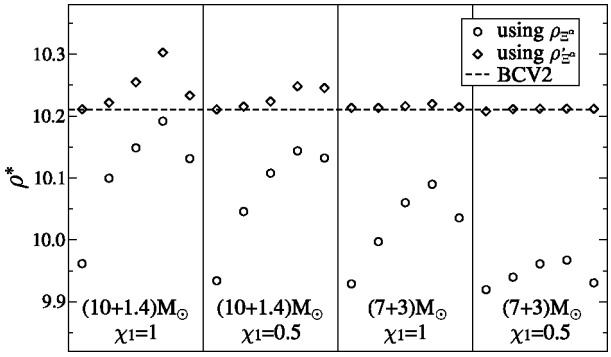

FIG. 6. Detection thresholds ρ^* for a false-alarm rate of 10^{-3}/yr, using the constrained statistic ρ_{Ξ^α} (circles), the approximated (unconstrained) statistic ρ'_{Ξ^α} (diamonds), and the detection statistic for the $(\psi_0\psi_{3/2}\mathcal{B})_6$ DTF from BCV2 (dashed line). The four windows correspond to the masses and χ_1 shown; the points in each window correspond to $\kappa_1=0.9,\ 0.5,\ 0,\ -0.5,\ -0.9$.

C. Numerical investigation of false-alarm statistics

The total false-alarm probability for the filtering of experimental data by a template bank over a time T is

$$P_{\rm tot}(\rho>\rho^*)=1-[P(\rho<\rho^*)]^{\mathcal{N}_{\rm shapes}\mathcal{N}_{\rm times}} \tag{61}$$

(see, for instance, Ref. [22]), where the exponent $\mathcal{N}_{\rm shapes}\mathcal{N}_{\rm times}$ is an estimate of the number of effective independent statistical tests. The number of independent signal shapes $\mathcal{N}_{\rm shapes}$ is related to (and smaller than) the number of templates in the bank [52]; the number of independent times of arrival $\mathcal{N}_{\rm times}$ is roughly $T/\delta t_0$, where δt_0 is the mismatch in the time of arrival needed for two nearby templates to have, on average, very small overlap. In our tests we set $\mathcal{N}_{\rm shapes}=10^6$ and $\mathcal{N}_{\rm times}=3\times10^{10}$ (or equivalently $\delta t_0 \simeq 1$ ms), as suggested by the results of Sec. VI for template counts and for the full mismatch metric; in fact, both numbers represent rather conservative choices.

We compute single-test false-alarm probabilities from Eq. (59), averaging the integrand over 10^5 randomly chosen values of Ω to perform the integration over Ω. Our convergence tests indicate that this many samples are enough to obtain the required precision [53]. In Fig. 6 we show the thresholds ρ^* required to achieve a total false-alarm rate of 10^{-3}/yr; the figure suggests that a threshold close to 10 is adequate. The thresholds are only marginally higher for the unconstrained statistic, so the number of first-stage false alarms that are dismissed in the second stage is limited. We show also the threshold required to achieve the same false-alarm rate with the $(\psi_0\psi_{3/2}\mathcal{B})_6$ DTF of BCV2: this threshold is very close to the values found for ρ'_{Ξ^α}, indicating that ρ'_{Ξ^α} has roughly six effective degrees of freedom (as it seems reasonable from counting the five P^I plus Φ_0). The BCV2 threshold is consistently higher than the ρ_{Ξ^α} threshold for the same single-test false-alarm rate; this suggests that the detection scheme discussed in this paper is less wasteful (with respect to the available signal power) than the BCV2 scheme, assuming of course that the number of templates used in the two banks is similar.

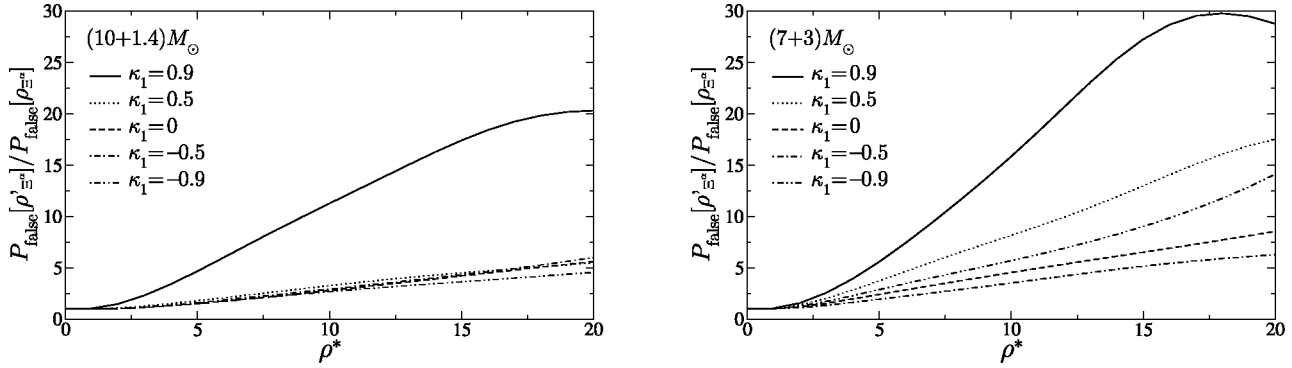

FIG. 7. Ratio $(1-P[\rho'_{\Xi^\alpha}<\rho^*])/(1-P[\rho_{\Xi^\alpha}<\rho^*])$ between single-test false-alarm probabilities for the unconstrained and constrained detection statistics, as a function of threshold ρ^*. The two panels represent systems with masses equal to $(10+1.4)M_\odot$ (left) and to $(7+3)M_\odot$ (right). The five curves in each plot correspond to different κ_1.

In Fig. 7 we show the ratio between the single-test false-alarm probabilities for ρ_{Ξ^α} and ρ'_{Ξ^α}: for a common threshold around 10, we can expect about five times more false alarms using ρ'_{Ξ^α} than using ρ_{Ξ^α}, for most values of the intrinsic parameters (for some of them, this number could be as high as ~ 15). These results corroborate our suggestion of using ρ'_{Ξ^α} in the first-stage of a two-stage detection scheme, to weed out most of the detection candidates before computing the more computationally expensive ρ_{Ξ^α}.

We need to remember that our assumption of Gaussian stationary noise is usually not realized in practice. For non-Gaussian noise, the probability density function [the counterpart of $p_r(r)$ defined in Eq. (58), now a function of both r and Ω] drops more slowly for larger r, so more high-ρ events will appear than are expected for the Gaussian case. These events will cause false alarms for both the unconstrained and constrained statistics, but we expect the ratio of the ρ'_{Ξ^α} and ρ_{Ξ^α} false-alarm probabilities at a common threshold (shown in Fig. 7 for Gaussian noise) *to be reduced*, at least if the threshold is high enough; so our proposed two-stage scheme should still be relevant in reducing computational costs. This expectation is justified by the following argument. For a common threshold ρ^* and a given direction $\hat{Z}^{\mathcal{I}}$, constrained-statistic triggers require $r=|Z^{\mathcal{I}}|$ larger by a certain factor $N_r>1$ than unconstrained-statistic triggers; this factor N_r is independent of ρ^* and of the probability distribution for r. Now, the ratio $P[r>N_r r_0]/P[r>r_0]$ drops much more quickly for a Gaussian distribution (or rather, for a χ^2 distribution with several degrees of freedom) than for a distribution with larger tails; thus, our Gaussian-noise estimates undercount the ρ_{Ξ^α} triggers with respect to the ρ'_{Ξ^α} triggers. As for nonstationarity, real-world data-analysis protocols try to cope by introducing vetoing schemes and by requesting coincident triggers between two or more detectors [54]. It is hard to predict how these additional elements might play into our proposed search scheme; a full statistical analysis including non-Gaussianity and nonstationarity is beyond the scope of this paper, and it will probably need to be quite empirical in nature.

VI. TEMPLATE COUNTING AND PLACEMENT

The last aspect to examine before we can recommend the template family of Sec. IV for actual use with the two-stage search scheme of Sec. V is the total number of templates that are needed in practice. As mentioned in Sec. II, the template-bank size and geometry required to achieve a certain minimum match can be studied using the *mismatch metric* [27,29,34], which describes, to quadratic order, the degrading overlap between nearby elements in a template bank:

$$1-\langle\hat{h}(\lambda^A),\hat{h}(\lambda^A+\Delta\lambda^A)\rangle\equiv\delta[\lambda^A,\lambda^A+\Delta\lambda^A]=g_{BC}\Delta\lambda^B\Delta\lambda^C, \tag{62}$$

where δ denotes the mismatch, and where

$$g_{BC}=-\frac{1}{2}\frac{\partial^2\langle\hat{h}(\lambda^A),\hat{h}(\lambda^A+\Delta\lambda^A)\rangle}{\partial(\Delta\lambda^B)\partial(\Delta\lambda^C)}. \tag{63}$$

No zeroth- or first-order terms are needed in the expansion (62), because the overlap has a maximum of 1 (for normalized templates) at $\Delta\lambda^A=0$. The metric is positive definite, because $\delta>0$. Note that, according to this definition, the mismatch δ is the *square* of the *metric distance* between λ^A and $\lambda^A+\Delta\lambda^A$. It is also *half* the square of the *inner-product distance* $\sqrt{\langle\Delta\hat{h},\Delta\hat{h}\rangle}$, where $\Delta\hat{h}\equiv\hat{h}(\lambda^A)-\hat{h}(\lambda^A+\Delta\lambda^A)$ [55].

Ideally, for a given continuous template family, one could find a reparametrization in which the metric is a Kronecker delta, and then lay down a template bank as a uniform hypercubic lattice in these coordinates, with the appropriate density to yield the required MM. For a hypercubic lattice in n dimensions [56], the (metric) side δl of the lattice cell is given by the relation $1-\mathrm{MM}=n(\delta l/2)^2$ [22,27]; we then get the total number of templates in the bank by dividing the total (metric) volume of parameter space by the volume of each cell:

$$\mathcal{N}_{\rm templates}=\int\sqrt{|\det g_{BC}|}d^n\lambda^A/[2\sqrt{(1-\mathrm{MM})/n}]^n. \tag{64}$$

In practice, this expression will usually underestimate the total number of templates, for two reasons: first, for more than two dimensions it is usually impossible to find coordinates where the metric is diagonalized everywhere at once;

second, the fact that the actual parameter space is bounded will also introduce corrections to Eq. (64). (The presence of null parameter directions, discussed in Sec. VI B, can also be seen as an extreme case of boundary effects.)

As we showed in Secs. IV and V, the overlap of the detector output with one of the ST_N templates can be maximized automatically over all the extrinsic parameters Ξ^α; it follows that a discrete template bank will need to extend only along the four intrinsic parameters X^i. So the estimate (64) for the number of templates should be computed on the *projected* metric g_{ij}^{proj} that satisfies

$$\begin{aligned} 1-\rho_{\Xi^\alpha}[\hat{h}(X^i,\Xi^\alpha),\hat{h}(X^i+\Delta X^i)] &\equiv 1-\max_{\Xi'^\alpha}\langle \hat{h}(X^i,\Xi^\alpha),\hat{h}(X^i+\Delta X^i,\Xi'^\alpha)\rangle \\ &= g_{ij}^{\mathrm{proj}}\Delta X^i \Delta X^j. \end{aligned} \tag{65}$$

Note that g_{ij}^{proj} is still a function of *all* the parameters. In Sec. VI A we compute g_{ij}^{proj} from the full metric g_{BC}; we then proceed to construct an *average* metric, $\overline{g_{ij}^{\mathrm{proj}}}$, which is connected closely to detection rates and does not depend on the extrinsic parameters.

In fact, it turns out that *not all four intrinsic parameters are needed to set up a template bank that achieves a reasonable* MM: we can do almost as well by replacing a 4D bank with a 3D bank where (for instance) we set $\kappa_1=0$. As a geometrical counterpart to this fact, the projected metric must allow a quasinull direction: that is, it must be possible to move along a certain direction in parameter space while accumulating almost no mismatch. The correct template counting for the 3D bank is then described by a *reduced* metric, which we discuss in Sec. VI B. Finally, we give our results for the total number of templates in Sec. VI C.

A. Computation of the full, projected, and average metric

According to Eq. (63), the full metric g_{BC} can be computed numerically by fitting the quadratic decrease of the overlap $\langle \hat{h}(\lambda^A),\hat{h}(\lambda^A+\Delta\lambda^A)\rangle$ around $\Delta\lambda^A=0$. It is also possible to rewrite g_{BC} in terms of first-order derivatives of the waveforms: since $\langle \hat{h}(\lambda^A),\hat{h}(\lambda^A)\rangle=1$ for all λ^A,

$$\frac{\partial}{\partial\lambda^B}\langle \hat{h},\hat{h}\rangle = 2\left\langle \hat{h},\frac{\partial\hat{h}}{\partial\lambda^B}\right\rangle = 0 \tag{66}$$

[in this equation and in the following, we omit the parametric dependence $\hat{h}(\lambda^A)$ for ease of notation]; taking one more derivative with respect to λ^C, we get

$$\left\langle \frac{\partial\hat{h}}{\partial\lambda^C},\frac{\partial\hat{h}}{\partial\lambda^B}\right\rangle + \left\langle \hat{h},\frac{\partial^2\hat{h}}{\partial\lambda^C\partial\lambda^B}\right\rangle = 0, \tag{67}$$

which implies [by Eq. (63)]

$$g_{BC}=\frac{1}{2}\left\langle \frac{\partial\hat{h}}{\partial\lambda^B},\frac{\partial\hat{h}}{\partial\lambda^C}\right\rangle. \tag{68}$$

The inner product in the right-hand side of Eq. (68) expresses the *Fisher information matrix* for the normalized waveforms $\hat{h}(\lambda^A)$ (see for instance Ref. [30]); for nonnormalized waveforms $h(\lambda^A)$ we can write

$$g_{BC}=\frac{1}{2\langle h,h\rangle}\left\langle \frac{\partial h}{\partial\lambda^B},\frac{\partial h}{\partial\lambda^C}\right\rangle - \frac{1}{2\langle h,h\rangle^2}\left\langle \frac{\partial h}{\partial\lambda^B},h\right\rangle\left\langle h,\frac{\partial h}{\partial\lambda^C}\right\rangle. \tag{69}$$

It is much easier to compute the mismatch metric from Eq. (68) rather than from Eq. (63), for two reasons. First, we know the analytic dependence of the templates on all the extrinsic parameters (except t_0), so we can compute the derivatives $\partial\hat{h}/\partial\Xi^\alpha$ analytically (the derivative with respect to t_0 can be handled by means of the Fourier-transform time-shift property $\mathcal{F}[h(t+t_0)]=\mathcal{F}[h(t)]\exp[2\pi i f t_0]$). Second, although the derivatives $\partial\hat{h}/\partial X^i$ have to be computed numerically with finite-difference expressions such as $[\hat{h}(X^i+\Delta X^i,\Xi^a)-\hat{h}(X^i,\Xi^a)]/\Delta X^i$, this is still easier than fitting the second-order derivatives of the mismatch numerically [57].

To obtain the projected metric g_{ij}^{proj}, we rewrite the mismatch $\delta(\lambda^A,\lambda^A+\Delta\lambda^A)$ by separating intrinsic and extrinsic parameters,

$$\begin{aligned} &\delta(X^i,\Xi^\alpha;X^i+\Delta X^i,\Xi^\alpha+\Delta\Xi^\alpha) \\ &\quad=(\Delta X^i\Delta\Xi^\alpha)\begin{pmatrix} G_{ij} & C_{i\beta}\\ C_{\alpha j} & \gamma_{\alpha\beta}\end{pmatrix}\begin{pmatrix}\Delta X^j\\ \Delta\Xi^\beta\end{pmatrix}; \end{aligned} \tag{70}$$

here we have split the full metric g_{BC} into four sections corresponding to intrinsic-intrinsic (G_{ij}), extrinsic-extrinsic ($\gamma_{\alpha\beta}$), and mixed ($C_{\alpha j}=C_{j\alpha}$) components. Maximizing the overlap over the extrinsic parameters is then equivalent to minimizing Eq. (70) over the $\Delta\Xi^\alpha$ for a given ΔX^i, which is achieved when

$$\gamma_{\alpha\beta}\Delta\Xi^\beta + C_{\alpha j}\Delta X^j = 0, \tag{71}$$

while the resulting mismatch is

$$\begin{aligned} \min_{\Delta\Xi^\alpha}\delta(X^i,\Xi^\alpha;X^i+\Delta X^i,\Xi^\alpha+\Delta\Xi^\alpha) &= [G_{ij}-C_{i\alpha}(\gamma^{-1})^{\alpha\beta}C_{\beta j}]\Delta X^i\Delta X^j \\ &\equiv g_{ij}^{\mathrm{proj}}\Delta X^i\Delta X^j. \end{aligned} \tag{72}$$

Here $(\gamma^{-1})^{\alpha\beta}$ is the matrix inverse of $\gamma_{\alpha\beta}$. For each point (X^i,Ξ^α) in the *full* parameter space, the *projected metric* g_{ij}^{proj} describes a set of concentric ellipsoids of constant ρ_{Ξ^α} in the intrinsic-parameter subspace. We emphasize that the projected metric has tensor indices corresponding to the intrinsic parameters, but it is a function of both the intrinsic and the extrinsic parameters, and so are the constant-ρ_{Ξ^α} ellipsoids.

Therefore, to build a template bank that covers all the signals (for all X^i and Ξ^α) with a guaranteed MM, we must

use the projected metric at each X^i to construct the constant-mismatch ellipsoids for all possible Ξ^α, and then take the intersection of these ellipsoids to determine the size of the unit template-bank cell. This is a *minimax* prescription [28], because we are maximizing the overlap over the extrinsic parameters of the templates, and then setting the template-bank spacing according to the least favorable extrinsic parameters of the signal. In general, the intersection of constant-mismatch ellipsoids is not an ellipsoid, even in the limit $\delta\to 0$, so it is impossible to find a single intrinsic-parameter metric that can be used to enforce the minimax prescription. There is an exception: the projected metric is not a function of t_0 or Φ_0 [58], so it can be used directly to lay down banks of nonspinning-binary templates [27,29] for which t_0 and Φ_0 are the only extrinsic parameters.

Returning to the generic case, we can still use the projected metric to guide the placement of a template bank if we relax the minimax prescription and request that the minimum match be guaranteed *on the average* for a distribution of signal extrinsic parameters [59]. It turns out that this *average-mismatch* prescription is closely related to the expected detection rates. Let us see how. The matched-filtering detection rate for a signal $s\equiv \mathrm{SA}\times\hat{h}(X^i,\Xi^\alpha)$, where $\mathrm{SA}=\langle s,s\rangle^{1/2}$ is the *signal amplitude* at a fiducial luminosity distance, is proportional to $\mathrm{SA}^3\rho^3_{\Xi^\alpha}[\hat{s},\hat{h}_{\mathrm{near}}]$, where $\hat{h}_{\mathrm{near}}\equiv\hat{h}(X^i+\Delta X^i,\Xi'^\alpha)$ is the closest template in the bank, and where we assume that sources are uniformly distributed throughout the volume accessible to the detector (see, for instance, Ref. [22]). The minimax prescription is given by

$$\rho_{\Xi^\alpha}[\hat{s},\hat{h}_{\mathrm{near}}]\simeq 1-g^{\mathrm{proj}}_{ij}(X^i,\Xi^\alpha)\Delta X^i\Delta X^j\geqslant \mathrm{MM} \qquad (73)$$

for all Ξ^α, which ensures that the detection rate is reduced at most by a factor MM^3 for every combination of signal extrinsic and intrinsic parameters.

Averaging over a uniform distribution of signal extrinsic parameters [60], we get a detection rate proportional to

$$\begin{aligned}\int d\Xi^\alpha \mathrm{SA}^3\rho^3_{\Xi^\alpha}&\simeq\int d\Xi^\alpha \mathrm{SA}^3[1-g^{\mathrm{proj}}_{ij}\Delta X^i\Delta X^j]^3\\ &\simeq\overline{\mathrm{SA}^3}-3\left[\int d\Xi^\alpha \mathrm{SA}^3 g^{\mathrm{proj}}_{ij}\right]\Delta X^i\Delta X^j\\ &\simeq\overline{\mathrm{SA}^3}[1-\overline{g^{\mathrm{proj}}_{ij}}\Delta X^i\Delta X^j]^3, \qquad (74)\end{aligned}$$

where $\overline{\mathrm{SA}^3}=\int d\Xi^\alpha \mathrm{SA}^3$, and where the *average metric* $\overline{g^{\mathrm{proj}}_{ij}}$, now a function only of X^i, is defined as

$$\overline{g^{\mathrm{proj}}_{ij}}=\int d\Xi^\alpha \mathrm{SA}^3 g^{\mathrm{proj}}_{ij}/\overline{\mathrm{SA}^3}. \qquad (75)$$

[To derive Eq. (74) we assume that $1-\rho_{\Xi^\alpha}[\hat{s},\hat{h}_{\mathrm{near}}]\ll 1$ for all Ξ^α.] We can now state the new average-mismatch prescription as

$$1-\overline{g^{\mathrm{proj}}_{ij}}(X^i)\Delta X^i\Delta X^j\geqslant\overline{\mathrm{MM}}, \qquad (76)$$

which ensures that the detection rate, *averaged over the extrinsic parameters of the signal*, is reduced at most by the factor $\overline{\mathrm{MM}}^3$. We shall call $\overline{\mathrm{MM}}$ the *average minimum match*.

B. Null parameter directions and reduced metric

As discussed by Sathyaprakash and Schutz [61] and by Cutler [59], an extreme example of boundary effect occurs when one of the eigenvalues of g_{BC} at λ^A (say, $\Lambda_{(1)}$) becomes so small that it is possible to move away in parameter space along the corresponding eigendirection (say, $e^A_{(1)}$) and *reach the boundary of the allowed parameter region while keeping the mismatch* $\delta(\lambda^A,\lambda^A+\tau e^A_{(1)})$ *well below the required value* $\delta_{\mathrm{MM}}=1-\mathrm{MM}$. In other words, the ellipsoid of constant mismatch δ_{MM} extends far beyond the allowed parameter region in the quasinull-eigenvalue direction. In such a situation, Eq. (64) will underestimate the total number of templates, because the denominator should now express the volume of the intersection of each lattice cell with the allowed parameter region [62]. A simple-minded fix to Eq. (64) is the following: write $\det g_{BC}=\Pi_{(k)}\Lambda_{(k)}$, where the $\Lambda_{(k)}$ are the n eigenvalues of g_{BC}; identify all the small eigenvalues, where *small* can be defined by $\Lambda_{(i)}\ll(1-\mathrm{MM})/l^2_{(i)}$, with $l_{(i)}$ the coordinate diameter of the allowed parameter range along the eigenvector $e^A_{(i)}$; replace the small eigenvalues by the corresponding value of the expression $(1-\mathrm{MM})/l^2_{(i)}$; use this modified determinant in Eq. (64).

Physically, the presence of k small eigenvectors suggests that the variety of waveform shapes spanned by an n-dimensional template family can be approximated with very high overlap by an $(n-k)$-dimensional *reduced* family. A lower-dimensional template bank is certainly desirable for practical purposes, but it is necessary to exercise caution: because the metric g_{BC} is not homogeneous, the quasinull eigendirections rotate as we move in parameter space [63], so we need to show explicitly that any signal in the n-dimensional family can be reached from a given $(n-k)$-dimensional submanifold along a quasinull trajectory. For this to happen, the small eigenvalues must exist throughout the entire n-dimensional parameter space, and the flow of the quasinull eigenvectors must map the submanifold into the entire space. To see that under these conditions the mismatch between the points on the submanifold and the points outside is indeed small, consider the following argument, due to Cutler [59]. The triangle inequality for the inner-product distance guarantees that

$$\delta^{1/2}[\lambda^A(0),\lambda^A(1)]\leqslant\int_0^1\sqrt{g_{BC}\frac{d\lambda^B}{d\nu}\frac{d\lambda^C}{d\nu}}d\nu \qquad (77)$$

along *any* path $\lambda^A(\nu)$; for a path that follows the flow of the quasinull eigenvector $e^A_{(i)}$ (a *reduction curve*), the total mismatch is then bounded by the average of $\Lambda_{(i)}$ along the curve, times an integrated squared parameter length of order $l_{(i)}$ [64].

For the ST_N template bank and for the two-stage search scheme of Sec. V, we find that the projected metric g^{proj}_{ij} admits a small eigenvalue for all values of the intrinsic and

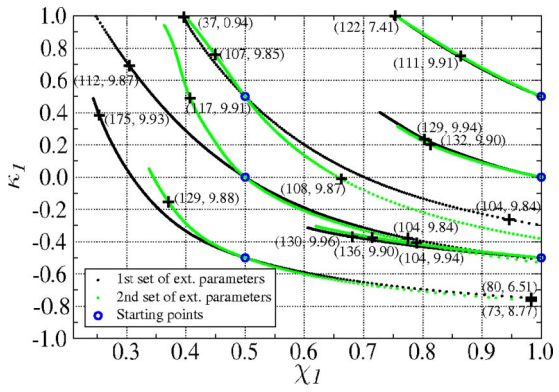

FIG. 8. Plot of (χ_1,κ_1) reduction curves in the (χ_1,κ_1) plane. We show curves for two sets of starting extrinsic parameters, corresponding to detector directions perpendicular (dark dots) and parallel (light dots) to the initial orbital plane. The curves start at the points marked with circles, and proceed in steps of 10^{-6} for the nominal mismatch (i.e., the mismatch computed using the projected metric). For starting points at $\chi_1=0.5$, we follow the quasinull eigenvector for both positive and negative increments. The curves end at the (χ_1,κ_1) boundary, or (roughly) where the true mismatch (i.e., the exact mismatch between the local and the starting template) becomes greater than 0.01. The ending points are marked with crosses, and they are annotated with the number of steps taken since the starting point, and with the true mismatch in units of 10^{-3}.

extrinsic parameters. Figures 8 and 9 show several examples of reduction curves that follow the quasinull eigendirections (the subtleties related to projected-metric reduction curves are discussed in Appendix C). The curves shown [65] begin at the points marked with circles, where $(m_1+m_2)=(10+1.4)M_\odot$ and

$$(\chi_1,\kappa_1)=\begin{Bmatrix}0.5\\1.0\end{Bmatrix}\times\begin{Bmatrix}-0.5\\0.0\\0.5\end{Bmatrix}; \tag{78}$$

the curves then proceed in steps of 10^{-6} for the *nominal* mismatch (i.e., the mismatch computed using the local projected metric) until they reach the (χ_1,κ_1) boundary, or (roughly) until the *true* mismatch (i.e., the exact mismatch between the local and the starting template) is greater than 0.01. We show curves for two sets of starting extrinsic parameters, corresponding to detector directions perpendicular (dark dots) and parallel (light dots) to the initial orbital plane. Figure 8 shows the projection of the reduction curves in the (χ_1,κ_1) plane; the ending points are marked with crosses, and they are annotated with the number of steps taken since the starting point, and with the true mismatch in units of 10^{-3}. Comparing the two numbers at each cross, we see that the triangle inequality is always respected: the true mismatch δ_N is always less than the accumulated nominal mismatch $10^{-6}N^2$ (where N is the number of steps); in fact, we see that the latter is a good approximation for the former. Figure 9 shows the projection of the same reduction curves in the $(M,\mathcal{M})$ plane. The chirp mass $\mathcal{M}\equiv M\eta^{3/5}$ varies by less than 2% along the curves: this is natural, since $\mathcal{M}$ dominates the evolution of the GW phase [see Eq. (5)].

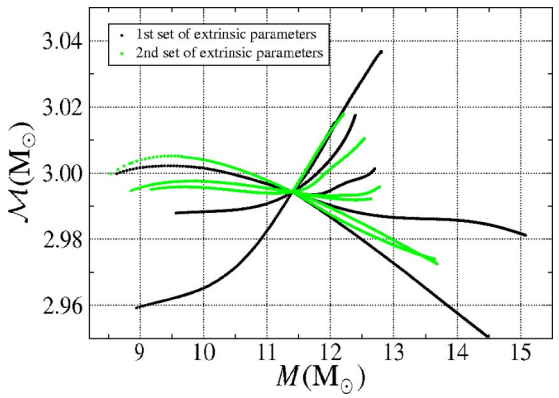

FIG. 9. Plot of (χ_1,κ_1) reduction curves in the $(M,\mathcal{M})$ plane. The curves are the same as shown in Fig. 8, but we omit all markings.

Figure 8 suggests that we can reduce the dimensionality of our template bank by collapsing the (χ_1,κ_1) plane into $\sim$ three curves, while retaining the full (M,η) plane. Templates laid down on these 3D submanifolds with a required minimum match MM will then cover every signal in the full 4D family with mismatch no larger than $(1-\mathrm{MM})+\delta_{\mathrm{red}}$, where $\delta_{\mathrm{red}}\simeq 0.01$ is the *reduction mismatch* introduced by the reduction procedure. Further investigations will be needed to find the optimal choice of reduction curves in the (χ_1,κ_1) plane, and to investigate the reduction curves of the average metric $\overline{g_{ij}^{\mathrm{proj}}}$.

C. Template counting

While three or more reduction curves will probably be necessary to limit $\delta_{\mathrm{red}}\simeq 0.01$, for the sake of definiteness we select a 3D reduced template space corresponding to $(m_1,m_2)\in[7,12]\times[1,3]$, $\kappa_1=0$, and $\chi_1\in(0,1]$ [66]. We compute the total number of templates in this 3D template bank according to

$$\mathcal{N}_{\mathrm{templates}}=\frac{\int\sqrt{|\det\overline{g_{i'j'}^{\mathrm{proj}}}|}\,dM\,d\eta\,d\chi}{[2\sqrt{(1-\mathrm{MM})/3}]^3}, \tag{79}$$

where the primed indices i',j' run through M, η, and χ, and we set $X^4\equiv\kappa_1=0$; furthermore, $\overline{g_{ij}^{\mathrm{proj}}}$ denotes the metric averaged over the extrinsic parameters Θ, φ, and α, as given by Eq. (75). The integral is carried out by evaluating the projected metric at the parameter sets

$$(m_1,m_2,\chi_1)=\begin{Bmatrix}7M_\odot\\12M_\odot\end{Bmatrix}\times\begin{Bmatrix}1M_\odot\\2M_\odot\\3M_\odot\end{Bmatrix}\times\begin{Bmatrix}0.1\\0.3\\0.5\\0.7\\1.0\end{Bmatrix}; \tag{80}$$

TABLE III. Effects of quadrupole-monopole terms, for $(10+1.4)M_\odot$ binaries with maximally spinning BH. At the beginning of each column we quote the overlaps between ST_2 templates and ST_2^{QM} templates that include quadrupole-monopole effects. Just as in Table II, these overlaps are maximizing only over the extrinsic parameters t_0 and Φ_0. In parentheses, $(\cdots)$, we show the fitting factors for the ST_2^{QM} family as matched by the ST_2 family; in brackets, $[\cdots]$, we show the intrinsic parameters at which the fitting factors are achieved. The View column describes the position of the detector with respect to the initial orbital plane. In all cases the integration of the equations is started at a GW frequency of 60 Hz. The maximization procedure was stopped whenever an overlap $\geqslant 0.99$ was achieved.

View	$(10+1.4)M_\odot$ with $\chi_1=1$							
	$\kappa_1=0.9$		$\kappa_1=0.5$		$\kappa_1=-0.5$		$\kappa_1=-0.9$	
Top	0.4796	($\geqslant$0.99) [10.3,0.13,1.21,0.89]	0.9890	($\geqslant$0.99)	0.1873	($\geqslant$0.99) [11.3,0.11,1.08,−0.48]	0.7245	(0.9877)
Side	0.3503	($\geqslant$0.99) [10.0,0.13,0.77,0.94]	0.8033	($\geqslant$0.99)	0.8754	($\geqslant$0.99) [11.4,0.11,1.03,−0.39]	0.7598	($\geqslant$0.99)
Diagonal	0.3292	($\geqslant$0.99) [11.2,0.11,0.80,0.94]	0.6669	($\geqslant$0.99)	0.4546	($\geqslant$0.99) [11.3,0.11,1.08,−0.49]	0.8437	(0.9887)

at each of the points the metric is averaged on 100 pseudo-random sets of extrinsic parameters. The integration then proceeds by interpolating across the parameter sets (80) The final result is $\mathcal{N}_{\text{templates}} \simeq 76\,000$ for MM$=0.98$ (not including the reduction mismatch). Given the uncertainties implicit in the numerical computation of the metric, in the interpolation, in the choice of the reduction curves, and in the actual placement of the templates in the bank, this number should be understood as an order-of-magnitude estimate. Most of the templates, by a factor of about 10 to 1, come from the parameter region near $m_2=1$ (that is, from the small-η region).

VII. SUMMARY

Buonanno, Chen, and Vallisneri (BCV2) recently proposed [18] a family of physical templates that can be used to detect the GWs emitted by single-spin precessing binaries. The attribute *physical* refers to the fact that the templates are exact within the approximations used to write the PN equations that rule the adiabatic evolution of the binary. In this paper, after reviewing the definition of this template family (here denoted as ST_N), we discuss the range of binary masses for which the templates can be considered accurate, and examine the effects of higher-order PN corrections, including quadrupole-monopole interactions. We then describe an optimized two-stage detection scheme that employs the ST_N family, and investigate its false-alarm statistics. Finally, we estimate the number of templates needed in a GW search with LIGO-I. Our results can be summarized as follows.

We determine the range of binary masses where the ST_N templates can be considered accurate by imposing two conditions: first, for the orbital separations that correspond to GWs in the frequency band of good interferometer sensitivity, the dynamics of the binary must be described faithfully by an adiabatic sequence of quasispherical orbits; second, the nonspinning body must be light enough that its spin will be negligible for purely dimensional reasons. The selected mass range is $(m_1,m_2) \simeq [7,12]M_\odot \times [1,3]M_\odot$.

To evaluate the effect of higher-order PN corrections for binaries in this mass range, we compute the overlaps between templates computed at successive PN orders. When computed between templates with the same parameters, such overlaps can be rather low; however, they become very high when maximized over the parameters (both intrinsic and extrinsic) of the lower-order PN template (see Table II). This means that the ST_2 template family should be considered acceptable for the purpose of GW detection, but this means also that the estimation of certain combinations of binary parameters can be affected by large systematic errors [20]. (When precessing-binary gravitational waveforms computed within PN-resummed and nonadiabatic approaches [28,44] become available, it will be interesting to compare them with the PN-expanded, adiabatic ST_N templates to see if the maximized overlaps remain high. We do expect this to be the case, because the spin and directional parameters of the ST_N templates provide much leeway to compensate for nontrivial variations in the PN phasing.) Again by considering maximized overlaps, we establish that quadrupole-monopole effects [67,68] can be safely neglected for the range of masses investigated (Table III).

We describe a two-stage GW detection scheme that employs a discrete bank of ST_2 templates laid down along the intrinsic parameters (M,η,χ_1,κ_1) [although the (χ_1,κ_1) may be collapsed to one or few 1D curves, in light of the discussion of dimensional reduction of Sec. VI]. The detection statistic $\rho_{\Xi\alpha}(M,\eta,\chi_1,\kappa_1)$ is the overlap between the template and the detector output, maximized over template extrinsic parameters: $(t_0,\Phi_0,P_I) \equiv (t_0,\Phi_0,\theta,\phi,\psi,\Theta,\varphi)$. This maximization is performed semialgebraically, in two stages. First, for all possible times of arrival t_0, we maximize the overlap over Φ_0 and over P_I without accounting for the constraints that express the functional dependence of the P_I on $(\theta,\phi,\psi,\Theta,\varphi)$: this step yields the approximated (unconstrained) maximum $\rho'_{\Xi\alpha}$, which can be computed very rapidly, and which sets an upper bound for $\rho_{\Xi\alpha}$. Second, only for the times of arrival t_0 at which $\rho'_{\Xi\alpha}$ passes the detection threshold, we compute the fully constrained maximum $\rho_{\Xi\alpha}$, which is more expensive to compute. (Note that this scheme differs from traditional hierarchical schemes because we use the *same* threshold in the first and second stages.) We find that $\rho'_{\Xi\alpha}$ is a good approximation to $\rho_{\Xi\alpha}$, so the number of first-stage triggers passed to the second stage is small.

For a total false-alarm probability of 10^{-3}/yr, and for a conservative estimate for the number of independent statistical tests, the detection threshold is around 10. For this value, between 5 and 15 first-stage triggers are passed to the second stage for each eventual detection. For the same threshold, the single-test false-alarm probability is lower for ST_2 templates than for the $(\psi_0\psi_{3/2}\mathcal{B})_6$ DTF of [18] [the total false-alarm

probability depends on the number of independent statistical tests, which is not available at this time for the $(\psi_0\psi_{3/2}\mathcal{B})_6$ DTF].

The procedure of maximization over the extrinsic parameters outlined in this paper can also be adapted for the task of detecting GWs from extreme-mass-ratio inspirals (i.e., the inspiral of solar-mass compact objects into the supermassive BHs at the center of galaxies [69]) and inspirals of two supermassive black holes with LISA [70]. This is possible under the simplifying assumptions of coherent matched filtering over times short enough that the LISA antenna patterns can be considered constant, and of GW emission described by the quadrupole formula. Furthermore, the formalism of projected and reduced mismatch metrics developed in Sec. VI can treat GW sources, such as extreme-mass-ratio inspirals, where many physical parameters are present, but only few of their combinations have significant effects on the emitted waveforms [59,61]. In fact, this formalism is closely related to the procedures and approximations used in the ongoing effort (motivated by mission-design considerations) to count the templates needed to detect extreme-mass-ratio inspirals with LISA [71].

It should be possible to generalize the formalism beyond quadrupole GW emission, at least to some extent. When higher-multipole contributions are included, the detector response becomes much more complicated than Eq. (15) [see, e.g., Eqs. (3.22b)–(3.22h) of Ref. [38]]. In particular, the response cannot be factorized into a factor that depends only on the dynamical evolution of the binary, and a factor that depends only on the position and orientation of the detector; it is instead a sum over a number of such terms, each containing different harmonics of the orbital and modulation frequencies. Despite these complications, it should still be possible to maximize the overlap over the extrinsic parameters, using a relatively small number of signal-template and template-template inner products. The constrained-maximization procedure would however be very complicated, and although the (fully algebraic) unconstrained maximum would still be easy to compute, the dimensionality of the unconstrained template space would now be so large that it may increase the false alarm probability too dramatically to make the two-stage scheme useful.

The last result of this paper is an estimate of the number of ST_2 templates needed for a GW search in the mass range $[7,12]M_\odot\times[1,3]M_\odot$. To obtain this estimate, we first compute the full mismatch metric, which describes the mismatch for small displacements in the intrinsic and extrinsic parameters; we then obtain the projected metric, which reproduces the effect of maximizing the overlap over the extrinsic parameters. At this point we observe that the projected metric has an eigenvector corresponding to a very small eigenvalue; this indicates that we can choose one of the four intrinsic parameters to be a function of the other three, so the dimensionality of the ST_2 template bank can be reduced to three. For simplicity, we perform this reduction by setting $\kappa_1=0$. We then compute the reduced mismatch metric, and obtain a rough estimate of $\sim 76\,000$ as the number of templates required for an average MM of 0.98, or 0.97 including an estimated reduction mismatch of 0.01.

ACKNOWLEDGMENTS

We wish to thank Leor Barack, Teviet Creighton, Curt Cutler, and Jonathan Gair for useful discussions. A.B. thanks the LIGO Caltech Laboratory under NSF cooperative agreement PHY92-10038 for support and the Theoretical Astrophysics group for hospitality during her visit at Caltech. The research of Y.C. and Y.P. was supported by NSF grant PHY-0099568 and NASA grant NAG5-12834. Y.C. is also supported by the David and Barbara Groce Fund at the San Diego Foundation. M.V.'s research was supported by the LISA Mission Science office at the Jet Propulsion Laboratory, California Institute of Technology, where it was performed under contract with the National Aeronautics and Space Administration.

APPENDIX A: THE QUADRUPOLE-MONOPOLE INTERACTION

In this appendix we investigate the effect of the quadrupole-monopole interaction, which we have so far neglected in describing the dynamics of precessing binaries. It is well known [67] that the quadrupole moment of a compact body in a binary creates a distortion in its gravitational field, which affects orbital motion (both in the evolution of ω and in the precession of $\hat{\mathbf{L}}_N$), and therefore GW emission; the orbital motion, on the other hand, exerts a torque on the compact body, changing its angular momentum (i.e., it induces a *torqued precession*). Although the lowest-order quadrupole-monopole effect is Newtonian, it is smaller than spin-orbit effects and of the same order as spin-spin effects.

When the spinning body is a black hole, the equations for the orbital evolution and GW templates are modified as follows to include quadrupole-monopole effects. Equation (5) gets the additional term [68]

$$\left(\frac{\dot{\omega}}{\omega^2}\right)_{\mathrm{Q-M}}=\frac{96}{5}\eta(M\omega)^{5/3}\left[\frac{5}{2}\chi_1^2\frac{m_1^2}{M^2}(3\kappa_1^2-1)(M\omega)^{4/3}\right], \tag{A1}$$

while the precession equations (6) and (7) become [68]

$$\dot{\mathbf{S}}_1=\frac{\eta}{2M}(M\omega)^{5/3}\left[\left(4+3\frac{m_2}{m_1}\right)-3\chi_1\kappa_1(M\omega)^{1/3}\right](\hat{\mathbf{L}}_N\times\mathbf{S}_1) \tag{A2}$$

and

$$\begin{aligned}\dot{\hat{\mathbf{L}}}_N&=\frac{\omega^2}{2M}\left[\left(4+3\frac{m_2}{m_1}\right)-3\chi_1\kappa_1(M\omega)^{1/3}\right](\mathbf{S}_1\times\hat{\mathbf{L}}_N)\\&\equiv\boldsymbol{\Omega}_L'\times\hat{\mathbf{L}}_N;\end{aligned} \tag{A3}$$

furthermore, the orbital energy (8) gets the additional term

$$\mathcal{E}_{\mathrm{Q-M}}(\omega)=-\frac{\eta}{2}(M\omega)^{2/3}\left[-\frac{1}{2}\chi_1^2\frac{m_1^2}{M^2}(3\kappa_1^2-1)(M\omega)^{4/3}\right]; \tag{A4}$$

finally, $\mathbf{\Omega}_e$ is again obtained from Eq. (14), using the modified $\mathbf{\Omega}'_L$ in Eq. (A3). (Note that $\kappa_1 \equiv \hat{\mathbf{L}}_N \cdot \hat{\mathbf{S}}_1$.)

The quadrupole-monopole interaction changes the number of GW cycles listed in Table I at 2 PN order. The additional contributions are $5.2\chi_1^2 - 15.5\kappa_1^2\chi_1^2$ for a $(10+1.4)M_\odot$ binary, $2.5\chi_1^2 - 7.6\kappa_1^2\chi_1^2$ for a $(12+3)M_\odot$ binary, and $1.8\chi_1^2 - 5.4\kappa_1^2\chi_1^2$ for a $(7+3)M_\odot$ binary. To estimate more quantitatively the effect of the quadrupole-monopole terms, we evaluate the nonmaximized overlaps (in the sense of Sec. III D) between 2 PN templates, computed with and without the new terms. The results for $(10+1.4)M_\odot$ binaries are summarized in Table III. In parentheses we show the fitting factors, which are all very high; in brackets we show the intrinsic parameters at which the maximum overlaps are obtained. We conclude that for the purpose of GW searches, we can indeed neglect the effects of the quadrupole-monopole interaction on the dynamical evolution of the binary.

APPENDIX B: ALGEBRAIC MAXIMIZATION OF THE OVERLAP OVER THE P_I

In this section, we explore the algebraic maximization of ρ_{Φ_0} [see Eq. (44)], given by

$$\rho_{\Phi_0} = \sqrt{\frac{A^{IJ}P_I P_J}{B^{IJ}P_I P_J}}, \tag{B1}$$

over the P_I. We recall that the five P_I are combinations of trigonometric functions of three angles, and therefore must satisfy two constraints: luckily, both of these can be formulated algebraically. In light of the discussion of Sec. IV B, the overall normalization of the P_I does not affect the value of the overlap (44). As a consequence, we can rescale the P_I and replace the first constraint by

$$B^{IJ}P_I P_J = 1, \tag{B2}$$

which enforces the normalization of the templates. This constraint is chosen only for convenience: the maximum, subject to this constraint, is exactly the same as the unconstrained maximum found by searching over the entire five-dimensional space. Let us work out its value, which will be useful later. Introducing the first Lagrangian multiplier λ, we impose

$$\frac{\partial}{\partial P_I}[A^{IJ}P_I P_J - \lambda(B^{IJ}P_I P_J - 1)] = (A^{IJ} - \lambda B^{IJ})P_J = 0, \tag{B3}$$

which has solutions only for λ corresponding to the eigenvalues of $\mathbf{AB}^{-1}$. For those solutions, we multiply Eq. (B3) by P_I to obtain

$$\lambda = A^{IJ}P_I P_J; \tag{B4}$$

using Eqs. (B1) and (B2), we then see that λ is the square of the overlap, so it should be chosen as the largest eigenvalue of $\mathbf{AB}^{-1}$. We then write the *unconstrained maximum* as

$$\rho'_{\Xi^\alpha} = \max_{t_0} \sqrt{\max \text{ eigv } \mathbf{AB}^{-1}}. \tag{B5}$$

By construction, ρ'_{Ξ^α} will always be larger than or equal to the constrained maximum, ρ_{Ξ^α}.

The second constraint comes from Eq. (47). Writing out the STF components, we get

$$\det P_{ij} \equiv \det \frac{1}{\sqrt{2}} \begin{pmatrix} P_1 + P_5/\sqrt{3} & P_2 & P_3 \\ P_2 & -P_1 + P_5/\sqrt{3} & P_4 \\ P_3 & P_4 & -2P_5/\sqrt{3} \end{pmatrix} \equiv D^{IJK}P_I P_J P_K = 0. \tag{B6}$$

(The tensor D^{IJK} can be chosen to be symmetric since $D^{IJK}P_I P_J P_K = D^{(IJK)}P_I P_J P_K$.) The constrained maximum of ρ_{Φ_0} over the P_I, subject to the two constraints, can be obtained as the maximum of the function

$$A^{IJ}P_I P_J - \lambda(B^{IJ}P_I P_J - 1) - \mu(D^{IJK}P_I P_J P_K) \tag{B7}$$

over P_I and over the two Lagrange multipliers λ and μ. After taking partial derivatives, we get a system of seven equations,

$$A^{IJ}P_J - \lambda B^{IJ}P_J - \frac{3}{2}\mu D^{IJK}P_J P_K = 0, \tag{B8}$$

$$B^{IJ}P_I P_J - 1 = 0, \tag{B9}$$

$$D^{IJK}P_I P_J P_K = 0, \tag{B10}$$

where the last two equations come from the constraints (B2) and (B6). Multiplying the first equation by P_I and using the two constraints, we obtain Eq. (B4) again. So the first Lagrange multiplier λ is still the square of the overlap. The second Lagrange multiplier μ is zero when the signal s belongs to ST_N template family, and has exactly the same intrinsic parameters as the template. In this case, the extrinsic parameters of the signal correspond to a vector P_I that satisfies Eq. (B8) with $\mu = 0$ (the multiplier λ is still needed to enforce normalization of the template). When the intrinsic parameters are not exactly equal, but close, μ becomes finite, but small. Equations (B8)–(B10) can then be solved iteratively by expanding P_I in terms of μ,

$$P_I = \sum_{n=0}^{\infty} P_I^{(n)} \mu^n. \tag{B11}$$

Inserting this expansion into Eqs. (B8) and (B10), we get the zeroth-order equation

$$A^{IJ}P_J^{(0)} - (A^{LM}P_L^{(0)}P_M^{(0)})B^{IJ}P_J^{(0)} = 0, \tag{B12}$$

where we have already used the zeroth-order version of Eq. (B4) to eliminate λ.

Multiplying by $(B^{-1})^{KI}$, we see that the zeroth-order solution $P_J^{(0)}$ must lie along an eigenvector of $(B^{-1})^{KI}A^{IJ}$, and that the corresponding eigenvalue must be equal to $A^{LM}P_L^{(0)}P_M^{(0)}$, and therefore also to the square of the zeroth-order extremized overlap. To get the maximum overlap, we must therefore choose $P_I^{(0)}$ as the eigenvector corresponding to the largest eigenvalue. So the zeroth-order constrained maximum is exactly the unconstrained maximum obtained above [Eqs. (B3)–(B5)].

We can then proceed to nth-order equations:

$$[A^{IJ}-2(A^{JM}P_M^{(0)}B^{IL}P_L^{(0)})-(A^{LM}P_L^{(0)}P_M^{(0)})B^{IJ}]P_J^{(n)} = \sum_{m_1=0}^{n-1}\sum_{m_2=0}^{n-1} A^{LM}P_L^{(m_1)}P_M^{(m_2)}B^{IJ}P_J^{(n-m_1-m_2)} + \sum_{m=0}^{n-1}\frac{3}{2}D^{IJK}P_J^{(m)}P_K^{(n-m-1)}. \quad \text{(B13)}$$

At each order, we insert the nth-order expansion of P_I into Eq. (B10), and select the real solution closest to zero as the nth-order approximation to μ (such a solution is guaranteed to exist for all odd n). We then obtain the nth-order approximation to λ (and therefore to ρ_{Ξ^α}) using Eq. (B4). We proceed in this way, until λ and μ converge to our satisfaction.

This iterative procedure succeeds when the intrinsic parameters of signal and template are close; as their distance increases, the procedure becomes more and more unstable and eventually fails to converge. The iteration fails often also when the overlap is optimized against pure noise. For these reasons, a practical implementation of the detection statistic ρ_{Ξ^α} must eventually rely on the semialgebraic maximization procedure discussed in Sec. IV B. Indeed, we have used the semialgebraic procedure for all the tests discussed in Sec. V.

APPENDIX C: DIMENSIONAL REDUCTION WITH A NONUNIFORM PROJECTED METRIC

In this appendix we extend the reasoning of Sec. VI B to study dimensional reduction under the projected metric $g_{ij}^{\rm proj}(\lambda^A)$, which lives in the intrinsic parameter space, but is a function of all parameters. For each point $\lambda^A=(X^i,\Xi^\alpha)$ in parameter space, we denote $\Lambda_{(1)}(\lambda^A)$ the smallest eigenvalue of $g_{ij}^{\rm proj}(\lambda^A)$, and $e^i_{(1)}(\lambda^A)$ the corresponding eigenvector in the intrinsic parameter space. Suppose we have

$$\Lambda_{(1)}(\lambda^A) \ll \frac{1-\mathrm{MM}}{l_{(1)}^2}, \quad \text{(C1)}$$

for all values of λ^A in the allowed parameter region, where $l_{(1)}$ is the coordinate diameter of the allowed parameter range along the eigenvector $e^i_{(1)}$.

Now let us start from a generic point $\lambda_0^A=(X_0^i,\Xi_0^\alpha)$ in parameter space and follow the normal eigenvector $e^i_{(1)}$ for a tiny parameter length ϵ, reaching $\lambda_1^A=(X_1^i,\Xi_1^\alpha)$, according to

$$X_1^i = X_0^i + \epsilon e^i_{(1)}(\lambda_0^A), \quad \text{(C2)}$$

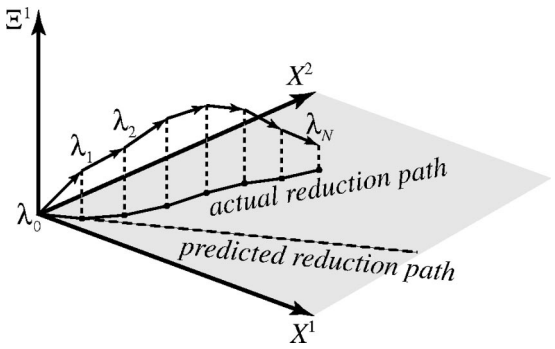

FIG. 10. Illustration of dimensional reduction. Here we show a signal space with one extrinsic parameter (Ξ^1) and two intrinsic parameters ($X^{1,2}$), and we assume that the projected metric has one small eigenvalue all through parameter space. Starting from a generic point λ_0^A, we follow the flow of the quasinull eigenvector of $g_{ij}^{\rm proj}$ for an infinitesimal parameter distance to reach λ_1^A; we then repeat this process, each time adjusting the direction of the eigenvector according to the metric (hence the difference between the reduction path *predicted* at λ_0^A and the *actual* reduction path). In the end we reach λ_N^A after having accumulated a parameter length l in the *intrinsic* parameter space. The mismatch between λ_0 and λ_N will be smaller than $\delta_{\rm MM}=1-\mathrm{MM}$, if l is not much larger than $l_{(1)}$, the coordinate diameter of the intrinsic parameter space in the approximate direction of the quasinull eigenvector.

$$\Xi_1^i = \Xi_0^i + \epsilon[\gamma^{-1}(\lambda_0^A)]^{\alpha\beta}[C(\lambda_0^A)]_{\beta j}e^j_{(1)}(\lambda_0^A); \quad \text{(C3)}$$

this choice of $\Delta\Xi^\alpha$ makes Ξ_1^α the extrinsic parameter that minimizes $\delta(X_0^i,\Xi_0^\alpha;X_1^i,\Xi_1^\alpha)$. Denoting the inner-product distance as $\mathrm{dist}(\lambda_0^A,\lambda_1^A)\equiv\sqrt{2\,\delta(\lambda_0^A,\lambda_1^A)}$, we can write

$$\mathrm{dist}(\lambda_0^A,\lambda_1^A) = \epsilon\sqrt{2\Lambda_{(1)}(\lambda_0^A)}+O(\epsilon^2); \quad \text{(C4)}$$

from λ_1^A, we follow the eigenvector $e^i_{(1)}(\lambda_1^A)$ for another parameter length ϵ, and reach λ_2^A; then from λ_2^A we reach λ_3^A, and so on. Up to the Nth step, we have traveled a cumulative parameter distance $l=N\epsilon$ in the intrinsic parameter space, and an inner-product distance

$$\begin{aligned}\mathrm{dist}(\lambda_0^A,\lambda_N^A) &\leqslant \sum_{n=1}^{N}\mathrm{dist}(\lambda_{n-1}^A,\lambda_n^A)\\ &= \sum_{n=1}^{N}[\epsilon\sqrt{2\Lambda_{(1)}(\lambda_{n-1}^A)}+O(\epsilon^2)]\\ &\leqslant l\sqrt{2\max_{\lambda^A}\Lambda_{(1)}(\lambda^A)}+O(N\epsilon^2), \end{aligned} \quad \text{(C5)}$$

where in the first line we have used the triangle inequality for the inner-product distance. The term $O(N\epsilon^2)$ vanishes in the limit $\epsilon\to 0$, $N\to\infty$, keeping $l=N\epsilon$ finite (see Fig. 10). So we can take the continuous limit of Eqs. (C2) and (C3) and arrive at two differential equations that define the resulting trajectory:

$$\dot{X}^i(l) = e^i_{(1)}, \quad \dot{\Xi}^\alpha(l) = [\gamma^{-1}]^{\alpha\beta} C_{\beta j} e^j_{(1)}, \tag{C6}$$

where X^i and Ξ^α are parametrized by the cumulative parameter length l, with

$$X^i(l=0) = X^i_0, \quad \Xi^\alpha(l=0) = \Xi^\alpha_0. \tag{C7}$$

We can allow l to be either positive or negative, in order to describe the two trajectories that initially propagate along $\pm e^i_{(1)}(\lambda^A_0)$. Equation (C5) then becomes

$$\begin{aligned} \mathrm{dist}[\lambda^A_0, \lambda^A(l)] &\leqslant \int_0^l dl' \sqrt{2\Lambda_{(1)}[\lambda^A(l')]} \\ &\leqslant |l| \sqrt{2 \max_{\lambda^A} \Lambda_{(1)}(\lambda^A)}. \end{aligned} \tag{C8}$$

In terms of mismatch,

$$\begin{aligned} \min_{\Xi^\alpha} \delta[\lambda^A_0; X^i(l), \Xi^\alpha] &= \frac{1}{2} \{ \min_{\Xi^\alpha} \mathrm{dist}[\lambda^A_0; X^i(l), \Xi^\alpha] \}^2 \\ &\leqslant \frac{1}{2} \{ \mathrm{dist}[\lambda^A_0; \lambda^A(l)] \}^2 \\ &\leqslant \frac{1}{2} \left[\int_0^l dl' \sqrt{2\Lambda_{(1)}[\lambda^A(l')]} \right]^2 \\ &\leqslant l^2 \max_{\lambda^A} \Lambda_{(1)}(\lambda^A) \\ &\leqslant \left(\frac{l}{l_{(1)}} \right)^2 \delta_{\mathrm{MM}}, \end{aligned} \tag{C9}$$

where the hybrid notation of the first line indicates the mismatch along the solution of Eq. (C6), and where of course $\delta_{\mathrm{MM}} = 1 - \mathrm{MM}$. Here, although we evolve X^i and Ξ^α simultaneously; it is the trajectory $X^i(l)$ in the *intrinsic* parameter space that we are ultimately interested in. In the context of dimensional reduction for the projected metric, we shall call $X^i(l)$ the reduction curve.

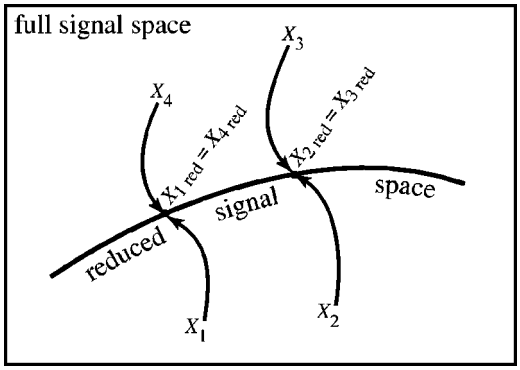

FIG. 11. Illustration of reduced signal space as a hypersurface inside full signal space. Here we show only the directions along the intrinsic parameters. Starting from the points $(X^i_1, \Xi^\alpha_1), \ldots, (X^i_4, \Xi^\alpha_4)$, we follow the trajectory (C6) and reach the hypersurface at $(X^i_{1\,\mathrm{red}}, \Xi^\alpha_{1\,\mathrm{red}}), \ldots, (X^i_{4\,\mathrm{red}}, \Xi^\alpha_{4\,\mathrm{red}})$. For these particular initial points, $X_{1\,\mathrm{red}}$ happens to coincide with $X_{4\,\mathrm{red}}$, and $X_{2\,\mathrm{red}}$ with $X_{3\,\mathrm{red}}$. We can see that λ^A_1 and λ^A_4 (and indeed all points that reduce to $X_{1\,\mathrm{red}}$, including the points along the reduction curve) will be indistinguishable upon detection with the reduced template bank. The same is true for λ^A_2, λ^A_3, and for all the points that reduce to $X_{1\,\mathrm{red}}$.

If the reduction curves are reasonably straight, it should be easy to find a (dimensionally reduced) hypersurface with the property that any given point (X^i_0, Ξ^α_0) in the full parameter space admits a reduction curve that reaches the hypersurface at a parameter l_* not much larger than the coordinate diameter of parameter space (see Fig. 11). From Eq. (C9), we then have $\min_{\Xi^\alpha} \delta [X^i_0, \Xi^\alpha_0; X^i(l_*), \Xi^\alpha] < \delta_{\mathrm{MM}}$. So any point in the full parameter space can be fitted with a mismatch smaller than δ_{MM} by a point on the hypersurface.

[1] A. Abramovici *et al.*, Science **256**, 325 (1992); http://www.ligo.caltech.edu.

[2] The LIGO Scientific Collaboration, Nucl. Instrum. Methods Phys. Res. A (to be published).

[3] B. Caron *et al.*, Class. Quantum Grav. **14**, 1461 (1997); http://www.virgo.infn.it

[4] H. Lück *et al.*, Class. Quantum Grav. **14**, 1471 (1997); http://www.geo600.uni-hannover.de

[5] M. Ando *et al.*, Phys. Rev. Lett. **86**, 3950 (2001); http://tamago.mtk.nao.ac.jp.

[6] L. Blanchet, T. Damour, B.R. Iyer, C.M. Will, and A.G. Wiseman, Phys. Rev. Lett. **74**, 3515 (1995); L. Blanchet, T. Damour, and B.R. Iyer, Phys. Rev. D **51**, 5360 (1995); C.M. Will and A.G. Wiseman, *ibid.* **54**, 4813 (1996).

[7] P. Jaranowski and G. Schäfer, Phys. Rev. D **57**, 7274 (1998); **60**, 124003 (1999); T. Damour, P. Jaranowski, and G. Schäfer, *ibid.* **62**, 044024 (2000); **62**, 021501(R) (2000); **63**, 044021 (2001).

[8] L. Blanchet, G. Faye, B.R. Iyer, and B. Joguet, Phys. Rev. D **65**, 061501(R) (2002).

[9] T. Damour, P. Jaranowski, and G. Schäfer, Phys. Lett. B **513**, 147 (2001).

[10] M. Salgado, S. Bonazzola, E. Gourgoulhon, and P. Haensel, Astrophys. J. **291**, 155 (1994); G.B. Cook, S.L. Shapiro, and S.A. Teukolsky, *ibid.* **424**, 823 (1994); P. Haensel, M. Salgado, and S. Bonazzola, Astron. Astrophys. **296**, 745 (1995); W.G. Laarakkers and E. Poisson, Astrophys. J. Lett. **512**, 282L (1999).

[11] C.S. Kochanek, Astrophys. J. **398**, 234 (1992); L. Bildsten and C. Cutler, *ibid.* **400**, 175 (1992).

[12] V. Kalogera, Astrophys. J. **541**, 319 (2000).

[13] P. Grandclément, M. Ihm, V. Kalogera, and K. Belczynski, gr-qc/0312084.

[14] T.A. Apostolatos, C. Cutler, G.J. Sussman, and K.S. Thorne, Phys. Rev. D **49**, 6274 (1994).

[15] T.A. Apostolatos, Phys. Rev. D **54**, 2438 (1996).

[16] P. Grandclément, V. Kalogera, and A. Vecchio, Phys. Rev. D **67**, 042003 (2003).

[17] P. Grandclément and V. Kalogera, Phys. Rev. D **67**, 082002 (2003).

[18] A. Buonanno, Y. Chen, and M. Vallisneri (BCV2), Phys. Rev.

D **67**, 104025 (2003).
[19] The authors of Refs. [16,17] did not include a Thomas precession term in the physical model used to test the templates; for this reason, the fitting factors quoted in Refs. [16,17] are substantially lower than our result. We note that those authors have recently investigated the effect of that term and found agreement with our results [13].
[20] A. Buonanno, Y. Chen, Y. Pan, and M. Vallisneri (unpublished).
[21] A. Buonanno, Y. Chen, Y. Pan, and M. Vallisneri (unpublished).
[22] A. Buonanno, Y. Chen, and M. Vallisneri (BCV1), Phys. Rev. D **67**, 024016 (2003).
[23] B.S. Sathyaprakash, Phys. Rev. D **50**, R7111 (1994).
[24] T. Damour, B.R. Iyer, and B.S. Sathyaprakash, Phys. Rev. D **63**, 044023 (2001).
[25] É.É. Flanagan and S.A. Hughes, Phys. Rev. D **57**, 4535 (1998); **57**, 4566 (1998).
[26] L.S. Finn and D.F. Chernoff, Phys. Rev. D **47**, 2198 (1993).
[27] B.J. Owen, Phys. Rev. D **53**, 6749 (1996).
[28] T. Damour, B.R. Iyer, and B.S. Sathyaprakash, Phys. Rev. D **57**, 885 (1998).
[29] R. Balasubramanian, B.S. Sathyaprakash, and S.V. Dhurandhar, Phys. Rev. D **53**, 3033 (1996).
[30] L.S. Finn, Phys. Rev. D **46**, 5236 (1992).
[31] B.S. Sathyaprakash and S.V. Dhurandhar, Phys. Rev. D **44**, 3819 (1991).
[32] S.V. Dhurandhar and B.S. Sathyaprakash, Phys. Rev. D **49**, 1707 (1994).
[33] C. Cutler and É.É. Flanagan, Phys. Rev. D **49**, 2658 (1994).
[34] B.J. Owen and B. Sathyaprakash, Phys. Rev. D **60**, 022002 (1999).
[35] T.A. Apostolatos, Phys. Rev. D **52**, 605 (1995).
[36] L. A. Wainstein and L. D. Zubakov, *Extraction of Signals from Noise* (Prentice-Hall, Englewood Cliffs, NJ, 1962).
[37] L.E. Kidder, C.M. Will, and A.G. Wiseman, Phys. Rev. D **47**, 4183(R) (1993).
[38] L.E. Kidder, Phys. Rev. D **52**, 821 (1995).
[39] L. Blanchet, T. Damour, and G. Esposito-Farese, gr-qc/0311052; Y. Itoh and T. Futamese, Phys. Rev. D **68**, 121501(R) (2003); Y. Itoh, *ibid.* **69**, 064018 (2004).
[40] Equation (11) of Ref. [18] suffers from two misprints: the spin-orbit term should be divided by M^2, and the spin-spin term by M^4.
[41] When templates are used in actual GW searches, the initial orbital frequency $\omega(0)$ must be chosen so that most of the signal power (i.e., the square of the optimal signal to noise) is accumulated at GW frequencies higher than the corresponding $f_{\rm GW}(0)=\omega(0)/\pi$,

$$\int_{f_{\rm GW}(0)}^{+\infty}\frac{\tilde{h}^*(f)\tilde{h}(f)}{S_n(|f|)}df\simeq\int_0^{+\infty}\frac{\tilde{h}^*(f)\tilde{h}(f)}{S_n(|f|)}df.$$

For the range of binary masses considered in this paper, and for the LIGO-I noise curve, such a $f_{\rm GW}(0)$ should be about 40 Hz. Most of the calculations performed in this paper (for instance, the convergence tests and the calculation of the mismatch metric) set instead $f_{\rm GW}(0)=60$ Hz to save on computational time; experience has proved that the results are quite stable with respect to this change.
[42] Note that the concept of extrinsic and intrinsic parameters had been present in the data-analysis literature for a long time (see, e.g., Ref. [36]). Sathyaprakash [23] draws the same distinction between *kinematical* and *dynamical* parameters.
[43] K. Belczynski, V. Kalogera, and T. Bulik, Astrophys. J. **572**, 407 (2001).
[44] A. Buonanno and T. Damour, Phys. Rev. D **59**, 084006 (1999); **62**, 064015 (2000); T. Damour, P. Jaranowski, and G. Schäfer, *ibid.* **62**, 084011 (2000); T. Damour, *ibid.* **64**, 124013 (2001); T. Damour, B.R. Iyer, P. Jaranowski, and B.S. Sathyaprakash, *ibid.* **67**, 064028 (2003).
[45] In fact, multiple GW detectors are needed to disentangle this factor from the distance D to the source.
[46] B. F. Schutz, in *The Detection of Gravitational Radiation*, edited by D. Blair (Cambridge University Press, Cambridge, England, 1989).
[47] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1992).
[48] Just as it happens for the P_I, the magnitude of $\mathbf{u}$ does not affect the value of Eq. (48), so the maximization can be carried out equivalently over all the vectors $\mathbf{u}$ that satisfy $\mathbf{u}\mathbf{B}_\alpha\mathbf{u}^T=1$. We can then use a Lagrangian-multiplier method to find the maximum, Eq. (51), and the corresponding $\mathbf{u}$, in a manner similar to the procedure used in Appendix B.
[49] P. Jaranowski, A. Królak, and B.F. Schutz, Phys. Rev. D **58**, 063001 (1998); LIGO Scientific Collaboration, B. Abbott *et al.*, Phys. Rev. D **69**, 082004 (2004).
[50] This is not a conventional hierarchical scheme, at least not in the sense that there is a tradeoff between performance and accuracy.
[51] The square root of the matrix $[\sqrt{C}]^{\mathcal{IJ}}$ can be defined, for instance, by $\sqrt{\mathbf{C}}\sqrt{\mathbf{C}}^T=\mathbf{C}$, and it can always be found because the covariance matrix $C^{\mathcal{IJ}}$ is positive definite. It follows that $\overline{Y^{\mathcal{I}}Y^{\mathcal{J}}}=[\sqrt{C}]^{\mathcal{IL}}[\sqrt{C}]^{\mathcal{JM}}\overline{X_{\mathcal{L}}X_{\mathcal{M}}}=[\sqrt{C}]^{\mathcal{IL}}[\sqrt{C}]^{\mathcal{JM}}\delta_{\mathcal{LM}}=C^{\mathcal{IJ}}$, as required.
[52] It is given, very roughly, by the number of templates when the minimum match is set to 0. See, for instance, Ref. [22].
[53] In fact, the average is dominated by the samples that yield the larger values of $\rho_1(\Omega)$, since the Γ function amplifies small changes in its argument. So the number of samples used in the Monte Carlo integration needs to be such that enough large $\rho_1(\Omega)$ do come up.
[54] LIGO Scientific Collaboration, gr-qc/0308069.
[55] So g_{BC} is truly a metric, except for a factor of 1/2.
[56] For specific dimensionalities, other regular packings might be more efficient: for instance, in two dimensions a lattice of equilateral triangles requires fewer templates than a lattice of squares.
[57] We have found that we can obtain a satisfactory precision for our metrics by taking several cautions: (i) reducing the parameter displacement ΔX^i along a sequence ${}^{(k)}\Delta X^i$ until the norm $\langle {}^{(k)}\Delta\hat{h}-{}^{(k-1)}\Delta\hat{h},{}^{(k)}\Delta\hat{h}-{}^{(k-1)}\Delta\hat{h}\rangle$ of the kth correction becomes smaller than a certain tolerance, where ${}^{(k)}\Delta\hat{h}=[\hat{h}(X^i+{}^{(k)}\Delta X^i,\Xi^a)-\hat{h}(X^i,\Xi^a)]/{}^{(k)}\Delta X^i$ is the kth approximation to the numerical derivative; (ii) employing higher-order finite-difference expressions; (iii) aligning both the starting and ending times of the waveforms $\hat{h}(X^i,\Xi^a)$ and $\hat{h}(X^i+\Delta X^i,\Xi^a)$ by

suitably modifying their lengths [by shifting the two waveforms in time, and by truncating or extending $\hat{h}(X^i + \Delta X^i, \Xi^a)$ at its starting point].

[58] The overlap $\langle \hat{h}(X^i, \Xi^\alpha), \hat{h}(X'^i, \Xi'^\alpha) \rangle$ depends only on $t_0 - t_0'$ and $\Phi_0 - \Phi_0'$: we have $\hat{h}(f) \sim \exp\{2\pi i f t_0 + i\Phi_0\}$ and $\hat{h}'(f) \sim \exp\{2\pi i f t_0' + i\Phi_0'\}$, so $\hat{h}'(f)\tilde{h}^*(f) \sim \exp\{2\pi i f(t_0' - t_0) + i(\Phi_0' - \Phi_0)\}$.

[59] C. Cutler (personal communication).

[60] All the expressions to follow can be adapted to the case of *a priori* known probability distribution for the extrinsic parameters. However, in our case it seems quite right to assume that the orientation angles Θ and φ are distributed uniformly over a sphere, and that α is distributed uniformly in the interval $[0, 2\pi]$.

[61] B.S. Sathyaprakash and B.F. Schutz, Class. Quantum Grav. **20**, S209 (2003).

[62] Pictorially, the error that we make with Eq. (64) is to let the template bank be thinner than a single template in the direction $e^A_{(1)}$.

[63] In fact, in the context of our templates this rotation is such that Eq. (62) ceases to be true in the quasinull eigendirections for $\delta \gtrsim 0.01$. As soon as we move away from the point λ^A where the metric is computed, any rotation of the eigenvectors means that the original quasinull direction is no longer the path along which the mismatch grows most slowly. If the larger eigenvalues are several orders of magnitude larger than the smaller ones, as is true in our case, a tiny rotation is enough to mask the contribution from the smallest eigenvalue.

[64] At least if the geometry of the reduction curve is not very convoluted.

[65] The curves of Figs. 8 and 9 are in fact obtained by following the quasinull eigenvectors of the *fully projected* (χ_1, κ_1) metric, which is a 2D metric on χ_1 and κ_1 obtained by projecting $g^{\rm proj}_{ij}$ again over M and η [using Eq. (72)], *as if they were extrinsic parameters.* The two projection steps are equivalent to projecting the 9D full metric into the 2D (χ_1, κ_1) plane in a single step. This procedure estimates correctly the *reduction mismatch* introduced by adopting the reduced template family created by first collapsing the 2D (χ_1, κ_1) plane into several 1D curves, and then including all the values of (M, η) for each point on the curves. We choose to collapse the (χ_1, κ_1) plane for empirical reasons: the χ_1 and κ_1 parameter bounds are simple, $(\chi_1, \kappa_1) \in [0,1] \times [-1,1]$, and the reduction curves have large parameter lengths.

[66] In fact, a second small eigenvector appears as we get close to $\chi_1 = 0$; this is because spin effects vanish in that limit, so a 2D family of nonspinning waveforms should be sufficient to fit all signals with small χ_1.

[67] See, e.g., H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1980), Chap. 5.

[68] E. Poisson, Phys. Rev. D **57**, 5287 (1998).

[69] D. Hils and P.L. Bender, Astrophys. J. Lett. **445**, L7 (1995); S. Sigurdsson and M.J. Rees, Mon. Not. R. Astron. Soc. **284**, 318 (1996); L.S. Finn and K.S. Thorne, Phys. Rev. D **62**, 124021 (2000); A. Ori and K.S. Thorne, *ibid.* **62**, 124022 (2000). For the scientific payoffs of observing extreme-mass-ratio inspirals, see F.D. Ryan, Phys. Rev. D **52**, 5707 (1995); **56**, 1845 (1997).

[70] P. L. Bender, K. Danzmann, and the LISA Study Team, Laser Interferometer Space Antenna for the Detection of Gravitational Waves, Pre-Phase A Report, Document MPQ233 (Max-Planck-Institüt für Quantenoptik, Garching, 1998).

[71] L. Barack and C. Cutler, Phys. Rev. D **69** 082005 (2004); T. Creighton and J. Gair (private communication).