Directed searches for gravitational waves from ultralight bosons

Maximiliano Isi,1,2,* Ling Sun,2,3,4,† Richard Brito,5,6,‡ and Andrew Melatos3,4,§

1LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
2LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA
3School of Physics, University of Melbourne, Parkville, Victoria 3010, Australia
4Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav),
   University of Melbourne node, Parkville, Victoria 3010, Australia
5Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
   Am Mühlenberg 1, Potsdam-Golm 14476, Germany
6Dipartimento di Fisica, “Sapienza” Università di Roma & Sezione INFN Roma1,
   Piazzale Aldo Moro 5, 00185 Roma, Italy

(Received 20 October 2018; published 26 April 2019)

Gravitational-wave detectors can be used to search for yet-undiscovered ultralight bosons, including those conjectured to solve problems in particle physics, high-energy theory, and cosmology. In particular, ground-based instruments could probe boson masses between $10^{-15}$ eV and $10^{-11}$ eV, which are largely inaccessible to other experiments. In this paper, we explore the prospect of searching for the continuous gravitational waves generated by boson clouds around known black holes. We carefully study the predicted waveforms and use the latest-available numerical results to model signals for different black-hole and boson parameters. We then demonstrate the suitability of a specific method (hidden Markov model tracking) to efficiently search for such signals, even when the source parameters are not perfectly known as well as allowing for some uncertainty in theoretical predictions. We empirically study this method’s sensitivity and computational cost in the context of boson signals, finding that it will be possible to target remnants from compact-binary mergers localized with at least three instruments. For signals from scalar clouds, we also compute detection horizons for future detectors (Advanced LIGO, LIGO Voyager, Cosmic Explorer, and the Einstein Telescope). Among other results, we find that, after one year of observation, an Advanced LIGO detector at design sensitivity could detect these sources up to over 100 Mpc, while Cosmic Explorer could reach over $10^4$ Mpc. These projections offer a more complete picture than previous estimates based on analytic approximations to the signal power or idealized search strategies. Finally, we discuss specific implications for the follow-up of compact-binary coalescences and black holes in x-ray binaries. Along the way, we review the basic physics of bosons around black holes, in the hope of providing a bridge between the theory and data-analysis literatures.

DOI: 10.1103/PhysRevD.99.084042

I. INTRODUCTION

After decades of dedicated effort, the Advanced Laser Interferometer Gravitational-wave Observatory (aLIGO) [1] and Advanced Virgo [2] detectors have inaugurated gravitational-wave (GW) astronomy with the observation of several compact-binary coalescences (CBCs) [3–9]. Arguably, one of the most exciting prospects in this new era of astronomy is to use gravitational waves to learn about fundamental physics. Common examples of this are attempts to probe the nature of gravity by testing general relativity [10,11], or to probe the nature of nuclear matter through the neutron-star equation of state [9,12,13]. Another exciting possibility is that of using gravitational waves to learn about particle physics. In particular, it may be possible to search for new ultralight bosons with gravitational-wave detectors, a prospect that has recently garnered much attention [14–21]. In this paper, we explore the potential of achieving this using directed searches for continuous gravitational waves with ground-based detectors.

There are strong theoretical reasons to believe in the existence of new weakly interacting, ultralight scalar (spin 0) or vector (spin 1) particles. The prime example of this is the axion, a (pseudo)scalar particle originally proposed to explain the strong constraints on the existence of charge-parity (CP) violating terms in the strong nuclear force sector [22–24]. This quantum-chromodynamics (QCD) axion is among the best-motivated extensions of the standard model, but there are others. For example,
string theory predicts the existence of a variety of axionlike particles (potentially including the QCD axion) with masses populating each decade between $10^{-33}$ eV and $10^{-10}$ eV as a result of the compactification of extra spatial dimensions [25]. Similarly, a hidden sector of light vector particles also naturally arises in compactifications of string theory [26]. Besides motivations from particle and high-energy physics, these bosons are also popular dark matter candidates (see, e.g., [27–29]).

Because of their weak couplings to the standard model and their vanishingly small mass, all these proposed new particles would be extremely hard to detect by conventional means. In particular, all existing constraints on the existence of the QCD axion rely on its expected coupling to Standard Model particles, a property that is heavily model dependent [30]. Such observations loosely constrain the mass of the QCD axion to be $\lesssim 10^{-3}$ eV [31], but values of order $10^{-10}$ eV or lower are favored by theory [14]. For other kinds of conjectured ultralight bosons, whose potential interactions with the standard model are very weak or inexistent, constraints and detection can only be provided through their gravitational coupling.

Given the substantial challenge in detecting ultralight bosons, there has been considerable excitement about the proposal to look for these particles by taking advantage of the universal character of gravitational couplings. The idea hinges on the phenomenon of superradiance [32–36], by which macroscopic clouds of these bosons should form around rapidly spinning black holes (BHs) and, in turn, produce a varied set of observational signatures [14–21, 37–41]. Indeed, constraints based on BH spin measurements of x-ray binaries have already been placed and exclude roughly the mass interval $[10^{-12}, 10^{-11}]$ eV for noninteracting massive scalar fields [18,38] and $[10^{-13}, 10^{-11}]$ eV for noninteracting massive vector fields [21,38]. Constraints derived from observations of x-ray binaries should, however, be interpreted with caution, since there is great uncertainty about the age and history of these systems, as well as caveats about the systematics affecting their spin measurements [42,43].

A much cleaner observational signature is the emission of potentially detectable gravitational waves at a frequency of roughly twice the boson mass [17–21]. As it turns out, this means that clouds formed around stellar mass BHs should emit signals within the most sensitive band of ground-based detectors, probing boson masses in the theoretically interesting mass range of the order of $10^{-15}$ eV through $10^{-11}$ eV [17,20].

In this paper, we explore the prospect for the direct detection of continuous gravitational waves emitted by boson clouds. In particular, we focus on searches in data from present and future ground-based detectors directed at known BHs. As the main observational scenario, we consider the follow-up of remnants from compact-binary coalescences detected through gravitational waves, but we also examine BH candidates known from electromagnetic observations. Besides treating the data analysis, we study in detail the morphology of boson signals and use numerical calculations, combined with the latest analytic results, to estimate their amplitude and other relevant features. This allows us to more accurately predict the potential signals that may be expected from clouds around a given BH. In the case of scalar clouds, we use those estimates to obtain detection horizons for second-generation and proposed third-generation instruments. Along the way, we review the theoretical basics of BH-boson superradiance in a language that we hope facilitates future work by gravitational-wave analysts interested in the topic.

In Sec. II, we review the theory of boson clouds around BHs and the gravitational-wave emission mechanism. In Sec. III we discuss at length the specific morphology of the expected signals, as seen by ground-based detectors. In Sec. IV, we introduce hidden Markov model tracking as an ideal strategy to search for these signals, validate the method, and estimate the sensitivity through Monte Carlo simulations; we also discuss implications for the follow-up of compact-binary mergers and x-ray binaries as potential sources. Finally, we provide a summary and conclusions in Sec. V.

II. THE BOSON CLOUD

The physics of boson fields around BHs has been extensively studied in different limits using both analytic and numerical methods [14,20,25,35,36,44–49]. We summarize those results below, in a way that is best suited for the data-analysis framework introduced later in the paper. We explain how a macroscopic boson cloud spontaneously arises around fast-spinning BHs and proceeds to emit gravitational radiation, providing some essential mathematical detail.

Consider first a Kerr BH of mass $M$ and angular momentum $J$. The characteristic length associated with the BH mass will be $r_g \equiv GM/c^2$, or half the Schwarzschild radius. The other characteristic length, given by the BH spin, is the usual Kerr parameter, $a \equiv J/(Mc)$, from which we can in turn define the dimensionless spin, $\chi \equiv ac^2/(GM)$. In terms of these quantities, the radius of the hole’s outer horizon, in Boyer-Lindquist coordinates, is

$$r_+ = r_g \left( 1 + \sqrt{1 - \chi^2} \right) \equiv r_g \tilde{r}_+,$$

where $\tilde{r}_+$ is defined here to be dimensionless. At this location, the BH will then have a frame-dragging angular speed (with respect to infinity) of

$$\Omega_{\text{BH}} = \frac{1}{2} \frac{c}{r_g} \frac{\chi}{\sqrt{1 - \chi^2}} \equiv \frac{c}{r_g} \tilde{\Omega}_{\text{BH}},$$

where $\tilde{\Omega}_{\text{BH}}$ is defined here to be dimensionless.
where \( \bar{\Omega}_{\text{BH}} \) is defined here to be dimensionless. (See, e.g., [50] for a recent review of the Kerr metric.)

Now, imagine that, beyond the usual particles in the standard model, there exists an ultralight boson of mass

\[
m_b \equiv \mu / c^2,
\]

where \( \mu \) is the boson’s rest energy. The corresponding length and timescales are given by the Compton wavelength, \( \lambda_\mu \equiv 2\pi \hbar / (m_b c) \), and angular frequency, \( \omega_\mu = c / \lambda_\mu = \mu / \hbar \). As soon as the BH is born (perhaps, as a result of stellar collapse or a binary coalescence), the usual quantum fluctuations in the boson field will cause pairs of particles to spontaneously appear in the hole’s vicinity, causing a number of them to unavoidably fall in. What happens next will depend on the properties of the BH and the infalling excitations: under most circumstances, the particle in the boson field will scatter off the BH with a boost in amplitude, effectively increasing the number of particles (occupation number) around the BH [32–34,36,51–53].

From the second law of BH thermodynamics [34] (or more generic kinematic arguments [25,36,53]), we may expect the boson-wave amplification to occur when the following superradiance condition is satisfied:

\[
\omega_\mu / m < \bar{\Omega}_{\text{BH}},
\]

where \( m \) is the (magnetic) quantum number corresponding to the projection of the particle’s total angular momentum along the BH spin direction. This amplification extracts energy from the BH just as in the classical Penrose process [32], which is itself another manifestation of BH superradiance (see [36] for a review). Because of the field’s nonzero mass, a scattered boson will generally tend to be bound to the BH, attracted by its gravitational pull. Consequently, scattered particles may remain confined in that region, facilitating successive scatterings and the associated compounded amplification of the field. This process is similar to the “BH bomb” devised by Press and Teukolsky [33], with the mirror replaced by the boson mass [45–47,54–56].

We may anticipate that the boson amplification will be maximized when the field and BH have comparable characteristic length scales, i.e., \( \lambda_\mu \sim r_g \). If this is the case, then the field amplitude will grow at an exponential rate in a (quasi)bound state around the BH (Sec. II A). As the field grows, it draws energy and angular momentum from the BH until the condition of Eq. (4) is no longer satisfied. Because the field is bosonic, there are \textit{a priori} no limits to the occupation number of any given energy level: the number of particles in a superradiant state will grow exponentially to form a macroscopic “cloud.” However, if formed solely due to the superradiant instability, the cloud will extract at most \( \sim 10\% \) of the BH’s mass [57,58]. This cloud will slowly fade away, as its energy is radiated away in the form of gravitational waves over very long timescales compared to the superradiant rate (Sec. II B).

### A. Black-hole and boson interactions

#### 1. Energy levels

The qualitative picture laid out above is backed up by analytic and semianalytic calculations of boson fields over a Kerr background [14,20,25,35,36,44–47]. In the nonrelativistic \((\alpha / j \ll 1)\) regime implied by Eq. (4), the influence of the BH is effectively reduced to a simple inverse-radius gravitational potential. This potential causes the bosons to present quasibound energy eigenstates essentially identical to those in the hydrogen atom, but with gravity replacing electromagnetism as the relevant interaction. More carefully solving the Schrödinger equation over a Kerr background, one indeed finds that, in this regime, the system has hydrogenic energy levels [44,46],

\[
E_n \approx \mu \left( 1 - \frac{1}{2} \alpha^2 + \cdots \right),
\]

for \( \bar{n} = n + l + 1 \) the principal quantum number, \( n \) the radial quantum number, and \( l \) the \textit{orbital} azimuthal quantum number.\(^2\) As usual, we have \( |j - s| \leq l \leq j + s \) and \(-j \leq m \leq j\), where \( j \) and \( s \) are, respectively, the total and spin angular-momentum quantum numbers. All superradiant levels are hydrogenic with a spectrum well described by Eq. (5) [47]. The quantity \( \alpha \) in Eq. (5) plays exactly the same role as the fine-structure constant in the hydrogen atom and takes the value of the ratio of the two relevant length scales (or, equivalently, timescales),

\[
\alpha \equiv \frac{r_g}{\lambda_\mu} = \frac{GM m_b}{c^2 \hbar} = \frac{GM}{c^3} \omega_\mu,
\]

where \( \lambda_\mu \equiv \lambda_\mu / (2\pi) \). Importantly, Eq. (4) implicitly constrains \( \alpha \) as a function of the BH spin. If we want superradiance to take place, then Eq. (4) and Eq. (2) demand

\[
\alpha < \frac{1}{2} m \sqrt{1 + \sqrt{1 - \chi^2}}^{-1} \leq \frac{m}{2},
\]

where the second inequality is obtained by noting \( 0 \leq \chi < 1 \). Because \( m \leq j \), this condition justifies working in the nonrelativistic, \( \alpha < j \), limit in the first place [14].\(^2\)

\(^2\)Although the spin parameter does not appear at leading order in \( \alpha [47] \), the fact that the BH is spinning does affect the angular part of the boson eigenfunctions; these have to be described using spin-weighted \textit{spheroidal}, rather than spherical, harmonics [20] (see Sec. III). Higher-order corrections to the energy eigenvalues, including corrections due to the BH spin, can be found in [39].
2. Cloud growth

Unlike the hydrogen atom, however, the BH-boson system is non-Hermitian due to the ingoing boundary condition for waves at the horizon. This means that the occupation number of the different energy eigenstates need not be constant—in fact, they will most certainly not be so for the superradiant states we are interested in. For small \( \alpha/m \), the occupation number\(^3\) of a given quantum state will grow exponentially at a rate that may be analytically approximated as [21,44,46]

\[
\Gamma_{jlmn} \approx 2\alpha^{2j+2l+5}\bar{r}_+(m\Omega_{\text{BH}} - \omega_p)C_{jlmn},
\]

with \( C_{jlmn} \) a dimensionless factor, and \( \bar{r}_+ \), the dimensionless radius defined in Eq. (1). In the case of a scalar boson, the orbital angular momentum is necessarily the total angular momentum \( (j = l) \), and it can be shown that [44,46]

\[
C_{jlmn}^{(\text{scalar})} = \frac{2^{l+j+2l+1+n}}{(l+n+1)!} \left[ \frac{l!}{(2l+1)!} \right]^2 \times \prod_{k=1}^{l} \left[ k^2(1 - \chi^2) + \frac{4r_+^2}{\alpha^2}(m\Omega_{\text{BH}} - \omega_p)^2 \right].
\]

Expressions for vector \( C_{jlmn} \) up to leading order in \( \alpha \) can be found in Appendix A of [21]. The validity of this approximation in the regime of interest has been confirmed numerically for scalars [47] and, more recently, vectors [38,49,59–61].

Irrespective of boson spin, there are three key features of the occupation growth rate, \( \Gamma \) in Eq. (8), that can be distilled from the above results:

(i) the sign of \( \Gamma \) depends solely on \( (m\Omega_{\text{BH}} - \omega_p) \), implying that indeed energy levels satisfying Eq. (4) will grow exponentially, while others will be depleted;

(ii) \( \Gamma \) is a high power of \( \alpha \), growing with the sum \( j + l \);

(iii) for a fixed angular momentum \( (j, l) \) and \( \alpha \), \( \Gamma \) decreases mildly with \( n \).

Because we are working in the small-\( \alpha \) limit, these three facts mean that, for a given system (i.e., a given \( \alpha \) and \( \chi \)), the fastest growth will occur for the fundamental radial harmonic of the level with the smallest possible total angular momentum, \( j \), that still supports a magnetic number, \( |m| \leq j \), sufficiently large to satisfy Eq. (4). In other words, if the boson has spin weight \( s = 0, 1 \), the level with the fastest superradiant growth in a given system will have angular quantum numbers \( \{j, l, m\} \) given by

\[
j = l + s = m = \text{cei}(\alpha/\Omega_{\text{BH}}),
\]

where “cei” stands for the operation of rounding up to the closest integer. In particular, the fastest-possible level over all values of \( \alpha \) and \( \chi \) will then be

\[
j = l + s = m = 1, n = 0.
\]

Given this, it follows from Eq. (8) that vector clouds will tend to grow significantly faster than scalar ones \( (\Gamma^{(v)}/\Gamma^{(s)} \propto \alpha^{-2}) \).

3. Final state

As the particle number grows, the energy and angular momentum required to populate the boson energy levels are extracted from the BH.\(^4\) Consequently, the BH quickly loses mass and spin until Eq. (4) is asymptotically saturated and the growth rate, Eq. (8), vanishes. As implied by Eq. (7), the spin of the BH at the end of this process will then be

\[
\chi_f = \frac{4\alpha_fm}{4\alpha_f^2 + m^2},
\]

where \( \alpha_f \) is given by Eq. (6) for the final BH mass. If no other processes (such as accretion) take place in the relevant timescale, the final mass of the cloud \( (M_c) \) will simply be given by the difference between the initial \( (M_i) \) and final \( (M_f) \) BH masses. If only one level is populated, then it may be shown that this will be approximately [20]

\[
M_c = M_i - M_f \approx M_i \frac{\alpha \chi_i}{m},
\]

with the last equality being valid for \( \alpha \lesssim 0.1 \).

A more exact value for this quantity may be obtained by numerically solving a set of difference equations, e.g., Eqs. (17)–(21) in [35], assuming a quasadiabatic evolution. If superradiant growth is the dominant factor, so that we can ignore other processes such as GW emission and accretion, these are just

\[
\dot{M} = -\Gamma_{jlmn} M_c, \quad \dot{M}_c = -\dot{M}, \quad \dot{J} = -m\Gamma_{jlmn} \omega \bar{r}_+^2 c^2 M_c, \quad \dot{J}_c = -\dot{J},
\]

where dots indicate time derivatives and \( (M, J) \) and \( (M_c, J_c) \) are the instantaneous mass and angular momentum for the BH and boson cloud, respectively. In the following sections, we will use the exact value for \( M_c \).

\(^3\)The rate of change of the occupation number is twice that of the field amplitude, which is itself given by the imaginary part of the wave function frequency: \( \Gamma = 2\text{Im}(\omega) \).

\(^4\)As this happens, Eq. (4) guarantees that the BH area increases, satisfying the second law of thermodynamics [34,53].
computed this way to characterize the signal. In any case, it may be shown that the boson cloud may reach a size of at most \( M_c \approx 0.1 \times M_i \) \cite{57,58}.

The time it takes a single-level cloud to grow to its full size is simply the time it takes the BH to reach the spin of Eq. (12). In the absence of significant interaction with the environment (e.g., through accretion or strong gravitational-wave emission), this is inversely linked to the “instability timescale” implied by Eq. (8), namely

\[
\tau_{\text{inst}} \equiv 1/\Gamma_{jlmn}.
\]

This corresponds to an \( e \)-folding in the occupation number of level \((j,l,m,n)\). In the nonrelativistic limit \((\alpha \ll 1)\), \( \tau_{\text{inst}} \) can be approximated for the dominant scalar level \((l = m = 1, n = 0)\) by \cite{20}

\[
\tau_{\text{inst}}^{(s)} \approx 27 \text{days} \left( \frac{M}{10 M_\odot} \right) \left( \frac{1}{\alpha} \right)^9 \frac{1}{\chi_i},
\]

which is generally slower than the timescale for the dominant vector level \((j = m = 1, n = 0)\), approximated by \cite{21}

\[
\tau_{\text{inst}}^{(v)} \approx 2 \text{min} \left( \frac{M}{10 M_\odot} \right) \left( \frac{1}{\alpha} \right)^7 \frac{1}{\chi_i},
\]

in the same small-\( \alpha \) limit.

As it turns out, gravitational interactions between levels might prevent the simultaneous population of more than one state \cite{14}. Whether this is true or not true, the hierarchy of involved timescales also suggests that we need only consider single-level clouds. For instance, for most parameters, the \( l = m = 2 \) scalar level has an instability timescale larger than or comparable to the depletion timescale of the fastest-growing level \( l = m = 1 \), so we should expect the latter to be unoccupied by the time the former reaches any significant size. The same holds for vectors, as discussed in \cite{21}. This means that we need only consider a single level at a time: as soon as the BH is born, the fastest-growing level will be quickly populated; because the field has integer spin, there is no limit to how many particles can occupy it, and the growth will continue until the BH reaches the spin of Eq. (12); the next level, with second-largest \( \Gamma \), will begin to grow only after the first one is depleted (through gravitational wave emission, as discussed below, or any other reason).

All of the results in this section were obtained from perturbative analyses that consider the nonrelativistic behavior of a boson field over a static Kerr background. These do not take into account effects such as the backreaction of the field onto the background metric, gravitational wave emission by the boson condensate, or interaction between energy levels. However, the validity of Eqs. (5)–(8) is confirmed by full numerical relativity simulations for the case of vector fields \cite{48,49,57,62}.

### B. Gravitational-wave emission

Once the boson cloud has reached a macroscopic size, it will emit a significant amount of gravitational radiation. There are three main mechanisms by which this may happen: (i) emission due to annihilation of bosons into gravitons; (ii) boson transitions between energy levels, analogous to electron jumps in the hydrogen atom; and (iii) abrupt collapse of the cloud due to particle self-interactions (“bosenova”).

Due to the high occupation numbers involved, the first two processes can be described purely classically, with GW emission stemming from a time-varying quadrupole (and higher multipoles, to a lesser degree) in the cloud’s stress energy. Transitions become important only if more than one level is occupied with comparable numbers. Therefore, it could take over thousands of years after the birth of the BH for such a signal to become detectable \cite{17,18}, making transitions interesting for very old BHs only. Unfortunately, the typical duration of transition signals would be of order years or shorter \cite{17}, which makes their observation from old potential sources highly unlikely. Meanwhile, bosenovae are only relevant in the presence of large boson self-interactions \cite{63,64}; in particular, they are not expected to occur for the QCD axion \cite{17}. In any case, the typical duration of bosenova signals would be of the order of milliseconds and the numerical simulations required to produce their waveforms are still in their infancy \cite{63,64}—this makes bosenovae a possibility relevant for unmodeled-burst analyses (e.g., \cite{65–67}), rather than the continuous-wave searches we are concerned with.

Given the above considerations, we restrict ourselves to annihilation signals, which are the best understood and most relevant for ground-based detectors. To understand this, consider a BH that has been maximally spun down due to the growth of the boson cloud surrounding it. We should expect this cloud to be composed of a vast number of particles in a coherent state corresponding to the fastest-growing energy level, as determined by Eq. (8). Indeed, we may think of the cloud as a macroscopic object with particle density given by the norm of the boson wave function plus a time-varying component proportional to \( \alpha^2 \), and rotating with angular frequency \( \omega_n = m \Omega_{\text{BH}} \), as implied by the saturation of Eq. (4). Treating this object purely classically, we may then expect the cloud to radiate gravitational waves at twice its rotational frequency \cite{68}. We note, however, that, unlike typical GW sources, the radiation wavelength is always smaller than the cloud’s extent, and therefore the assumptions behind the quadrupole approximation do not apply \cite{14}.

A detailed description of the cloud’s gravitational-wave emission can instead be obtained by using the Teukolsky formalism to solve the linearized Einstein equations for the
cloud’s stress energy, given by the wave function of the relevant quantum state [15,20]. A purely monochromatic boson field has a stress-energy tensor proportional to \(e^{-2\omega_0 t}\), therefore, one indeed finds that the cloud emits gravitational waves with angular frequency

\[
\tilde{\omega} = 2\omega_0
\]  

(18)

and that the emission pattern is described by a set of spin-weighted spheroidal harmonics corresponding to the spin of the final BH, with azimuthal numbers \(l \geq 2j\), and the magnetic quantum number fixed to \(m = 2m\) (see Sec. III A). Both vectors and scalars emit GWs with the same angular pattern (for a given \(j\) and \(m\)), with the fast-

growing level radiating mostly in the \(l = m = 2\) mode.

The gravitational power radiated in each angular mode \(l\) may be written as

\[
\dot{E}_{GW}(l, m, \tilde{\omega}) \approx \frac{1}{4\pi G} \left(\frac{c}{r_g}\right)^2 (M_\odot/c^2)^2 \mathcal{A}_{l m}^2(\alpha, \chi_i),
\]  

(19)

where \(M_f\) and \(r_g = GM_f/c^2\) are, respectively, the mass and length scale of the final BH and \(\mathcal{A}_{l m}(\alpha, \chi_i)\) is a dimensionless factor. For scalars, we compute the \(\mathcal{A}_{l m}\)'s numerically using BH perturbation theory as in [20], but this is not currently feasible for vectors. Regardless of boson spin, as long as Eq. (7) is satisfied, the emitted power will be a steep function of \(\alpha\) [15,17,20]. In fact, for small \(\alpha\), the power emitted in the dominant angular mode (\(l = 2\)) by the fastest-growing scalar level (\(l = m = 1, n = 0\)) may be roughly approximated as [35]

\[
\dot{E}_{GW}^{(s)} \approx 7 \times 10^{41} \text{ erg/s} \left(\frac{\alpha}{0.1}\right)^{16} \chi_i^2,
\]  

(20)

and for the fastest-growing vector level (\(j = m = 1, l = n = 0\)) as [21]

\[
\dot{E}_{GW}^{(v)} \approx 2 \times 10^{49} \text{ erg/s} \left(\frac{\alpha}{0.1}\right)^{12} \chi_i^2.
\]  

(21)

The difference in the \(\alpha\) dependence in these two expressions arises from the fact that the fastest vector level has no orbital angular momentum (\(l = 0\)), and so lies closer to the BH yielding a more compact cloud.

The energy in the gravitational radiation is drawn from the cloud itself, which slowly fades away as its component particles annihilate into gravitons [14,17]. As a result, the signal will be almost monochromatic, with a slowly decreasing amplitude and slowly increasing frequency. The evolution of both these quantities is tied to the timescale implied by Eq. (19),

\[
\tau_{GW} \equiv M c^2/\dot{E}_{GW}.
\]  

(22)

This “gravitational-wave timescale” is just the time it takes for half of the rest energy of the cloud to be radiated away and can be thought of as the typical duration of the signal. Using the approximations of Eqs. (13) and (20), we get for the dominant scalar level in the nonrelativistic limit, a signal duration of

\[
\tau_{GW}^{(s)} \approx 6.5 \times 10^4 \text{ yr} \left(\frac{M}{10 M_\odot}\right) \left(\frac{0.1}{\alpha}\right)^{15} \frac{1}{\chi_i}.
\]  

(23)

Similarly, using Eq. (21), for the dominant vector level we get

\[
\tau_{GW}^{(v)} \approx 1 \text{ day} \left(\frac{M}{10 M_\odot}\right) \left(\frac{0.1}{\alpha}\right)^{11} \frac{1}{\chi_i}.
\]  

(24)

Clearly, the vector processes tend to take place at a much faster pace than scalar ones, as expected from the higher radiated power. In both cases, however, the duration of the signal is significantly longer than the time it takes for the cloud to grow, as given by Eq. (15). This is an important, general feature that justifies the separate treatment of the early growth and late emission stages in the first place.

### III. THE SIGNAL

Having reviewed the physics of boson clouds around BHs in Sec. II, we will now focus on the properties of the gravitational signal produced by one of these systems, as seen by differential-arm-length detectors on the ground. Given that we only expect one quantum state to be significantly populated at any given time, we will restrict our discussion to signals from single-level clouds. In any case, the signal from a multilevel cloud can be produced trivially by the addition of several single-level waveforms described below, assuming negligible interaction between levels. We describe the signal morphology in Sec. III A and elaborate on the most salient features in Sec. III B.

#### A. Waveform

Consider a cloud made up of bosons in a single quantum state that has just stopped growing, after drawing enough energy and angular momentum from its host BH to saturate Eq. (4). In that case, as anticipated in Sec. II B, we expect the cloud to emit a continuous gravitational signal with a small spin-up and amplitude depending on the properties of both BH and boson. The strain signal, \(h_I\), seen by a given differential-arm-length detector, \(I\), can be written in the usual form as a sum over polarizations,

\[
h_I(t) = F_{+}^{I}(t)a_+ \cos \phi(t) + F_{\times}^{I}(t)a_\times \sin \phi(t),
\]  

(25)

where the \(F_{p}^{I}\)'s are the antenna-response functions of detector \(I\) to signals of plus (+) and cross (×) polarizations.
DIRECTED SEARCHES FOR GRAVITATIONAL WAVES … 

PHYS. REV. D 99, 084042 (2019)

... (see, e.g., Appendix B in [69] for explicit expressions). These depend implicitly on the relative location and orientation of the detector and the source, usually parametrized by its right ascension ($\alpha$), declination ($\delta$), and polarization angle ($\psi$). This last parameter determines how the frame in which the polarizations are defined is oriented in the plane of the sky; for our purposes, this will be the angle between the spin of the BH and the projection of the celestial north onto the plane normal to the line of sight.

The polarization amplitudes, $a_+$ and $a_\times$, are made up of contributions from several angular multipoles, indexed by the wave azimuthal number, $\ell \geq 2j$, and with fixed magnetic number, $\tilde{m} = 2m$,

$$a_{\pm/\times} = -\sum_{\ell \geq 2j} h_0^{(\ell)}[\mp S_{\ell\tilde{m}\tilde{m}} \pm S_{\ell\tilde{m}\tilde{m}}],$$

(26)

with the (plus) minus sign on the right-hand side corresponding to the $+\times$ polarization [20]. Because the boson cloud does not emit GWs isotropically, these amplitudes depend on the orientation of the source relative to the detector.

The angular dependence is encoded in Eq. (26) via the spin-weighted spheroidal harmonics [70], which are analogous to the usual spin-weighted spherical harmonics but account for the nonsphericity of the space around the Kerr BH. As such, these are functions of BH spin, signal frequency, and orientation with respect to the source,

$$sS_{\ell\tilde{m}\tilde{m}} \equiv S_{\ell\tilde{m}}(\alpha\tilde{m}/c, \cos \iota),$$

(27)

where for us the spin weight is always $s = -2$, as needed to describe GWs. The inclination $\iota$ is defined as the angle between the BH spin and the line of sight. We compute these eigenfunctions numerically using Leaver’s method [71,72].

The characteristic amplitude of each mode can be written as

$$h_0^{(\ell)} = \frac{c^4}{G M_{BH}^2} \frac{1}{2\pi^2 f^2 r} A_{\ell\tilde{m}}(\alpha, \chi),$$

(28)

where the dimensionless factor $A_{\ell\tilde{m}}$ encodes the relative amount of energy that the source deposits in each mode, as in Eq. (19). Assuming the BH has been fully spun down by the cloud, this is only a function of the initial BH spin and the fine-structure constant $\alpha$, and it can be computed numerically from BH perturbation theory following [20]. In the nonrelativistic limit ($\alpha \ll 1$), this can be approximated by [15,18,35]

$$h_0^{(\ell)} \approx 8 \times 10^{-28} \left(\frac{M}{10 M_\odot}\right) \left(\frac{\alpha}{0.1}\right)^7 \left(\frac{\text{Mpc}}{r}\right) \left(\frac{\chi - \chi_f}{0.1}\right)$$

(29)

for the dominant scalar mode ($\ell = m = 1$, $n = 0$, $\tilde{\ell} = \tilde{m} = 2$), and by [21]

$$h_0^{(\psi)} \approx 4 \times 10^{-24} \left(\frac{M}{10 M_\odot}\right) \left(\frac{\alpha}{0.1}\right)^5 \left(\frac{\text{Mpc}}{r}\right) \left(\frac{\chi - \chi_f}{0.1}\right)$$

(30)

for the dominant vector mode ($j = m = 1$, $n = 0$, $\tilde{\ell} = \tilde{m} = 2$), corresponding to the approximations in Eqs. (20) and (21), respectively. Equations (29) and (30) include an explicit spin-dependent correction factor to account for the dependence of the mass cloud on $\alpha$, with $\chi_f$ itself a function of $\alpha$ defined in Eq. (12). Although not fully accurate, these expressions will be useful when studying the scalings of the expected signal amplitude—especially in the case of vectors, for which the $A$ factors have yet to be computed numerically.

In both the scalar and the vector cases, the amplitude of the signal will decrease as a function of time, starting from the peak values given by Eq. (28). This is due to the progressive dissipation of the boson cloud sourcing the GW signal. This weakening occurs over a timescale of the order of the signal duration, Eq. (22).

In a frame inertial with respect to the source, the phase evolution of the signal corresponds to a simple monotone that may potentially evolve slowly in frequency. In the frame of the detector, extra timing corrections are needed, so that, in terms of the time $t$ measured at Earth, the phase evolution can be written as

$$\phi(t) = 2\pi \sum_{j=0}^{N} \frac{\partial f_j}{f} [t - t_0 + \delta t(t)]^{(j+1)} + \phi_0,$$

(31)

where $\delta f_j$ is the $j$th time derivative of $f$, the GW frequency measured at fiducial time $t_0$, and $\phi_0$ is a phase offset. The timing corrections $\delta t(t)$ account for delays due to the relative motion of the source and detector, general- and special-relativistic effects, as well as potential corrections due to the presence of a companion if the source is part of a binary (e.g., [73–75]).

As implied by Eq. (18), the source-frame frequency will be given by $f = \omega / (2\pi) = \omega_n / \pi$. Noting that $\omega_n \approx \omega_\mu$ by Eq. (5), this may be approximated as a function of BH mass and fine-structure constant as

$$f \approx \frac{\alpha}{\pi \tau_g} \approx 645 \text{ Hz} \left(\frac{10 M_\odot}{M}\right) \left(\frac{\alpha}{0.1}\right).$$

(32)

This means that stellar-mass BHs should support boson clouds that emit gravitational waves at frequencies within the sensitive band of existing and planned ground-based detectors.

Once the superradiance instability has shut down, the GW signal will expect a (slight) positive change in frequency (a “spin-up”). This is expected on purely classical grounds, as is typical of any gravitationally bound system (as in the characteristic “chirp” of compact-binary coalescences). The value of $\partial f_j$ can be computed...
from the rate of change in the cloud’s binding energy [21]. For scalars, Eq. (20) implies a signal frequency derivative (see Appendix)

\[
j_{s}^{(s)} \approx 3 \times 10^{-14} \, \text{Hz/s} \left( \frac{10 \, M_{\odot}}{M} \right)^{2} \left( \frac{\alpha}{0.1} \right)^{19} \chi_{l}^{2},
\]

(33)

while, for vectors Eq. (21) implies [21]

\[
j_{s}^{(v)} \approx 1 \times 10^{-6} \, \text{Hz/s} \left( \frac{10 \, M_{\odot}}{M} \right)^{2} \left( \frac{\alpha}{0.1} \right)^{18} \chi_{l}^{2}.
\]

(34)

The frequency drift is faster for vectors, as corresponds to quicker cloud dissipation. Besides this spin-up, we will also want to allow for higher-order derivatives of the frequency in Eq. (31) to incorporate potential perturbations caused by the astrophysical environment, presence of a companion, level interactions, or theoretical uncertainty.

**B. Projected properties**

We are interested in making statements about the presence of ultralight bosons based on searches for GW signals from known BHs. This means that we need to know what strain frequencies and amplitudes we may expect from our target BH, without knowing the true mass of the boson (if it exists). In principle, a particular BH could support clouds for a range of boson masses, which we will parametrize implicitly via the fine-structure constant, \( \alpha \) of Eq. (6). Although a particular single-level cloud is expected to emit a GW quasimonotone, an unknown \( \alpha \) means that we could expect signals at a variety of frequencies. This is clear from Eq. (20), in the small-\( \alpha \) limit, or more generally from the fact that the \( A_{l,n} \) factors in Eq. (28) will be nonzero for a range of \( \alpha \)'s. In other words, a given BH could “resonate” with different bosons, allowing us to probe a (narrow) range of particle masses; hence there is a band of signal frequencies to potentially be expected from a given BH.

In this section, we study in detail the properties of continuous signals that can be expected from clouds around a given BH. Although so far we have kept the discussion general, we now focus on scalars to provide concrete examples. We use numerical techniques to compute the power emitted by different systems, obtaining estimates that should be more reliable than previously published projections. In particular, for each value of \( \alpha \) and initial BH parameters, we numerically solve the differential equations governing the evolution of the cloud to obtain the final BH parameters, as in [35]. We then use the numerical results of [20] for the \( A_{l,n} \) factors to compute the radiated amplitude by means of Eq. (28). Unfortunately, at the moment, it is not possible to do this for vectors, since the corresponding perturbative calculations are significantly more difficult and have yet to be carried out.

To summarize the key points from the discussion below: any given BH can allow us to probe a narrow range of boson masses set by its mass and, to a lesser extent, its spin (Figs. 1–3); heavier BHs will “resonate” with lighter bosons and produce louder signals at lower frequencies (Fig. 4); and signals from heavier BHs (lighter bosons) will both grow and vanish more slowly, resulting in smaller frequency derivatives (Figs. 6–8). We expect similar conclusions to hold for vectors, except that the overall radiated power will be stronger and the timescales shorter; cf. Eqs. (8) and (21).

**1. Amplitude and frequency**

Begin with the example of a BH with initial mass \( M = 60 \, M_{\odot} \) and dimensionless spin \( \chi = 0.70 \), parameters consistent with the remnant from LIGO’s first detection [3]. Assume that the BH is then maximally spun down by the presence of a scalar cloud, and consider the amplitude of GWs emitted immediately after. We will assume the cloud is dominated by the fastest-growing energy level \((l=m=1, n=0)\), and restrict ourselves to the dominant GW mode \((\tilde{l} = \tilde{m} = 2)\). For concreteness, place the source 5 Mpc away and consider the amplitude of the signal as seen from Earth for different values of \( \alpha \). Our BH could potentially support clouds emitting GWs at different frequencies and with different amplitudes, depending on the true value of the boson mass.

The above fact is clearly illustrated in Fig. 1, which shows the characteristic amplitude of waves produced by

![Strain amplitude vs frequency for example black hole (scalar cloud)](image-url)

FIG. 1. Strain amplitude vs frequency for example black hole (scalar cloud). The x axis shows different frequencies at which we might expect a GW signal from a scalar \((l = m = 1, n = 0)\) cloud around a BH with initial mass of \( M_{i} = 60 \, M_{\odot} \) and initial spin \( \chi_{i} = 0.70 \). The colored curve shows the corresponding characteristic amplitude, \( h_{0} \), in Eq. (28) assuming \( r = 5 \) Mpc, parametrized by the fine-structure constant, \( \alpha \) in Eq. (6), as indicated by the color bar. For reference, the other curve shows the small-\( \alpha \) approximation of Eq. (29), including the spin correction responsible for the amplitude turnover. Points to the left of the vertical dotted line have \( \tau_{int} > 1 \) yr.
clouds around our example BH for different initial values of $\alpha$. In the best case scenario for this BH, if there existed a boson with $\mu \approx 4 \times 10^{-13}$ eV such that $\alpha \approx 0.179$, then we would observe a signal with characteristic strain amplitude $h_0 \approx 5.2 \times 10^{-26}(5 \text{ Mpc}/r)$ at $f \approx 191 \text{ Hz}$, corresponding to the peak in Fig. 1. For this value of the boson mass, at the end of the superradiant process the BH will have reached a final spin of $\chi = 0.62$, having lost 1.7% of its mass to the cloud.

Figure 1 also implicitly defines the range of frequencies of interest for searches directed at this BH to be, say, within 150 Hz and 200 Hz. This range could be broader or narrower, depending on the sensitivity of the search and how long one waits from the birth of the BH to make an observation. For instance, points to the left of the dotted gray line correspond to boson masses for which the signal would take longer than 1 year to complete one e-folding in amplitude (more on timescales in Sec. III B 2 below). According to Eq. (7), the maximum value of $\alpha$ for this source is $\sim 0.2$, for which the amplitude vanishes. Note that, because $\mu$ depends linearly on $\alpha$, this means that any given BH will allow us to probe a very narrow range of boson masses.

While the overall shape of the curve in Fig. 1 will be generally the same for all BHs, the location and width of the peak will be a strong function of the initial BH mass and spin. This is represented in Fig. 2, in which we have fixed the BH distance and spin to the values above, but allowed its mass to vary. Color in this figure represents the strain amplitude of GWs emitted at a given frequency (y axis) as a function of initial BH mass (x axis), while the gray dashed line marks the peak amplitude for a given BH mass. Although $\alpha$ is not explicitly shown, it should be clear from Fig. 1 that moving vertically toward higher frequencies and amplitudes corresponds to increasing $\alpha$. In fact, a vertical cut of Fig. 2 at $M_i = 60 M_\odot$ would yield Fig. 1.

Figure 2 shows that heavier BHs can support clouds that emit GWs at lower frequencies but greater amplitudes. This was expected from the discussion in Sec. II B: (i) heavier BHs are also larger, and so must be the boson cloud surrounding it, thus yielding lower GW frequencies; and (ii) heavier BHs result in a heavier cloud, as dictated by Eq. (13), which will in turn radiate more strongly, per Eq. (19). Because the overall radiated power decreases with BH mass, this also means that the band of detectable frequencies is narrower for lighter BHs, which is also visible in Fig. 2. The fact that the peak frequency (dashed line) decreases linearly with BH mass was already anticipated in Eq. (32). This can be understood from the observations that (i) $f \sim \omega/\pi \sim 2/\lambda_\mu$, and that (ii) $\lambda_\mu \sim r_g$ for the boson and BH sizes to match and maximize superradiance. As one moves vertically up the plot (increasing $f$ or, equivalently, $\alpha$), the emitted power vanishes abruptly at a point defined by the saturation of Eq. (7).
this case, because $\chi_i = 0.70$ and $m = 1$, this corresponds to $\alpha = 0.2$. Finally, as in Fig. 1, clouds corresponding to points below the dotted gray line would take longer than 1 year from the birth of the BH to complete one $e$-folding during growth.

The properties of the GW emission will also vary with the initial spin of the BH. This is illustrated in Fig. 3, in which we have fixed the BH distance and mass ($M_i = 60 M_\odot$), but allowed its initial spin to vary. As in Fig. 2, color represents the characteristic strain amplitude emitted at a given frequency (y axis) for different values of the initial spin (x axis); again, the dashed line traces the peak amplitude. It is no surprise to find that BHs with greater initial spins yield louder GWs: the faster the BH is spinning before the superradiant process kicks off, the longer the cloud may grow without saturating Eq. (4) and, consequently, the more mass it will extract before its growth stalls. The fact that higher-spin BHs result in heavier clouds is reflected in Eq. (13), and the corresponding dependence of $h_0$ on $\chi_i$ can be glimpsed from Eq. (20), which is valid for $\alpha \ll 1$. Similarly, for a fixed initial BH mass, a faster initial spin results in a lighter final BH [see Eq. (13) again], and so yields higher GW emission frequencies per Eq. (32). As in Fig. 2, the upper frequency cutoff is given by Eq. (12).

As suggested by Figs. 2 and 3, the characteristic GW amplitude emitted by a boson cloud as a function of frequency may show interesting structure as the initial BH mass and spin are varied. However, in many situations, it suffices to know the expected amplitude of the peak emission from a given system. This information is summarized in Fig. 4, which displays the characteristic amplitude and frequency for the optimal cloud as a function of BH mass, and for different values of the initial spin. The curves in the bottom panel can be understood as constant-spin cuts of the full mass-spin plane shown in Fig. 5. As for the other color maps, the dotted white lines in that plot mark the values of our example BH ($M_i = 60 M_\odot$, $\chi_i = 0.70$), which can at best yield an amplitude of $h_0 = 5.2 \times 10^{-26} (5 \text{ Mpc}/r)$ (peak of Fig. 1). Gray lines mark $\tau_{\text{inst}} = 1$ day and $\tau_{\text{inst}} = 1$ yr for reference.

2. Timescales

The figures discussed so far provide important information about the expected strain as a function of frequency when searching for signals from a given BH, but it is important to also consider the timescales introduced in Sec. II. There are two timescales associated with the gravitational signal: the time it takes to reach its peak amplitude, and its duration thereafter.
For instance, for the BH in Fig. 6 this means that, to constrain the presence of the best-matching boson (vertical dashed line: \( \alpha \approx 0.179 \), i.e., \( \mu \approx 4 \times 10^{-13} \) eV), one must wait at least 1 month from the moment the BH is born before looking for a GW signal in the data. During that first month, the cloud is still growing and the signal has not reached its peak \( (h_0 \approx 5.2 \times 10^{-26}, \text{ according to Fig. 1}) \), meaning it might be too weak and unstable for detection. Thus, the absence of a detectable signal during that initial period would not be evidence against the existence of the boson. The same is true for any other value of \( \alpha \), but the peak strains will be weaker (cf. Fig. 1) and the times required to reach them possibly much longer (if \( \alpha < 0.179 \), for our example with \( \chi_i = 0.70 \)).

The second relevant timescale, the signal duration, is also strongly dependent on \( \alpha \). This is also illustrated in Fig. 6 (purple curve) for our example BH (\( M = 60 M_\odot \), \( \chi = 0.70 \)). For the boson that best matches this BH (\( \alpha \approx 0.179 \)), the characteristic duration of the signal is \( \sim 7.5 \times 10^3 \) yr. Similar to the situation described above, the absence of a detectable signal long after this would not constitute evidence against the existence of such a boson. This is because, if one waits too long, the fastest energy level will have been depleted, and one should instead look for signals corresponding to the next level; cf. Eq. (11). This feature is especially important when targeting old BHs.

Both timescales are a function of BH mass, as reflected in Figs. 7 and 8, and scale inversely with BH spin. The color in Figs. 7 and 8 corresponds to the growth and duration timescales, respectively, both assuming a spin of \( \chi = 0.70 \). The horizontal dashed line marks \( \alpha = 0.179 \), the value of the fine-structure constant that yields peak emission for a BH with that spin. Meanwhile, a vertical cut along the dotted lines (\( M = 60 M_\odot \)) would produce the orange and purple curves in Fig. 6, respectively, for Figs. 7 and 8. Although both timescales vary widely for different \( \alpha \)'s, for any given system \( (M, \chi, \alpha) \), \( \tau_{\text{inst}} \) is always orders of magnitude shorter than \( \tau_{\text{GW}} \) which allows the treatment of the cloud growth and signal emission as two different regimes, as explained in Sec. II.

IV. DIRECTED SEARCHES

There are multiple observational signatures of BH superradiance that could be used to probe the boson-mass space [14–21,38–41,76]. Among all these, we will focus on the prospect for direct detection of the continuous gravitational waves expected from these sources (Sec. III). In particular, we will restrict ourselves to searches directed at specific well-localized targets, rather than searches covering the whole sky (see, e.g., [75] for a review of continuous-wave searches). This means that we are interested in studying known (potential) BHs that could have the right mass, spin, and age to possibly harbor a radiating boson cloud. In order to apply existing search strategies, we would
also like the cloud to be stable enough to make sure the signal lasts sufficiently long and evolves slowly enough to be considered “persistent” (we sharpen these criteria below).

Because the properties of the central BH can be measured \textit{a priori}, directed searches can potentially make unambiguous statements about the existence of ultralight bosons without relying on BH population models, which carry much uncertainty. If a signal were found from a given target, detailed measurements of its morphology (see Sec. III A) would provide invaluable information about the mass and dynamics of the new particle. On the other hand, if a signal were \textit{not} found, knowledge of the BH parameters could allow us to place stringent constraints on the existence of bosons in the corresponding mass range. Furthermore, having a specific sky location allows us to probe deeper in the noise and explore a greater range of parameters to farther distances. This comes at the price, of course, of the restriction to BHs that are already known, which may limit the use of the method in practice if no suitable BHs are discovered to target.

In the following, we introduce hidden Markov model (HMM) tracking as a well-suited method to carry out directed searches for these signals (Sec. IV A). We evaluate its sensitivity with Monte Carlo simulations and use the results to estimate the scalar-cloud detection horizons for future detectors (Sec. IV B). This discussion is agnostic as to the origin of the target BH, assuming only a known location and reasonably constrained intrinsic parameters. The conclusions are, therefore, generally applicable to any known stellar-mass BH, but we devote special attention to remnants from compact-binary mergers and holes in x-ray binaries (Sec. IV C). As we discuss below, vector signals present unique data-analysis and theoretical challenges, so we focus mainly (though not uniquely) on scalars.

A. Search method

Hidden Markov model tracking is an efficient search strategy for detecting quasimonochromatic gravitational waves \cite{77,78}. It was developed with rapidly spinning neutron stars in mind and has been applied in searches directed at several targets \cite{9,79–81}. This strategy is well suited to searches for gravitational waves from boson clouds because its computational efficiency enables the coverage of a wide range of signal phase parameters and a grid of sky locations. Furthermore, it allows small deviations from restrictive waveform models, unlike other coherent or semicoherent search methods that rely on Taylor-series-based matched filters and are, thus, more model restricted (e.g., \cite{82,83}). This makes it ideal to search for signals over a broad frequency band (cf. Fig. 1), even when the location of the source is only loosely known and when there is potential uncertainty in the signal morphology.

1. Algorithm overview

The goal of HMM tracking is to find the most likely path that a putative signal takes in the time-frequency plane, contingent on the observed noisy data \cite{77,78}. To do so, it
divides the $f - t$ plane into pixels, assuming the signal is monochromatic over a period $T_{\text{drift}}$ and splitting the frequency axis into bins of width $\Delta f = 1/(2T_{\text{drift}})$. The signal power in each bin is then estimated coherently using the $F$-statistic [82,84], a frequency-domain estimator that accounts for the motion of the Earth and is widely used in continuous-wave searches [75]. At each time step $i$, this statistic is computed for each discrete frequency bin $j$ by coherently integrating over the time interval $(t_i, t_i + T_{\text{drift}})$. Henceforth “$f_0$” denotes the central value of the signal frequency, $f_{0j}$ in the $j$th bin. If the total observation time is $T_{\text{obs}}$, then the values of $F(f_0)$ for the $N_F = T_{\text{obs}}/T_{\text{drift}}$ blocks of duration $T_{\text{drift}}$ are combined incoherently as described in [77,78].

Based on this information, the HMM algorithm computes the likelihood of different signal paths, assuming the signal can only transition between adjacent frequency bins from one time step to the next. For application to boson signals, we assume the transition probability $A_{f_0j,f_0k}$ between frequency bins $f_{0j}$ and $f_{0k}$ to be

$$A_{f_0j,1f_0j} = A_{f_0j,f_0i} = \frac{1}{2},$$

and to vanish otherwise (see [77,78] for details). The choice of Eq. (35) amounts to favoring signals with a positive frequency derivative, in agreement with the signal model of Sec. III. We choose a uniform prior $\Pi_{f_0} = N_Q^{-1}$ on $f_0$ over the frequency band being searched, where $N_Q$ is the total number of frequency bins. The result of the HMM tracking algorithm is summarized by a figure of merit representing the significance of the optimal path relative to all others (see [78,79]). This quantity can then be treated as a regular (frequentist) detection statistic, and its background can be computed over several noise-only instantiations data to assign detection significances.

2. Frequency-derivative tolerance

Although, in principle, this method would be able to handle signals with arbitrary frequency evolutions, allowing for large frequency drifts ($\dot{f}$) comes with a significant reduction in sensitivity. In order to allow for a maximum frequency derivative $\max(\dot{f})$ we must choose $T_{\text{drift}} \lesssim \Delta f/\max(\dot{f})$, so as to guarantee that

$$\left| \int_t^{t + T_{\text{drift}}} \text{d}t \dot{f}(t) \right| \lesssim \Delta f,$$

for $0 < t < T_{\text{obs}}$ and where the frequency resolution is set to $\Delta f = (2T_{\text{drift}})^{-1}$, as mentioned above. Therefore, tracking a signal with higher $\dot{f}$ requires reducing the coherent-integration time over which the $F$ statistic is computed, which in turn diminishes the sensitivity of the search [78].

The implementation of the $F$-statistic-based HMM used by the LIGO and Virgo Collaborations can currently track quasimonochromatic signals with derivatives of at most $\dot{f} \sim 10^{-8}$ Hz s$^{-1}$ [78]. This is more than enough to accommodate the majority of scalar signals [Eq. (33)], but puts most (although not all) of the vector parameter space out of reach [Eq. (34)]. Detecting shorter-lived signals may be possible by extending the current method to track not only the signal frequency, but also its first time derivative with a two-dimensional HMM as in [78]. Such a strategy would also naturally handle noticeable decays in signal amplitude over the observation run. The adaptation of the methods in [78] to boson signals will be the subject of future work.\footnote{The HMM tracking based on 1 s short Fourier transforms described in Ref. [81] can be used to search for long-duration transient signals with timescales $\sim$10$^5$–10$^4$ s. The timescale of a vector signal is much longer than that, and hence longer short Fourier transforms need to be used in the tracking. However, the Doppler modulation due to the motion of the Earth with respect to the solar system barycenter is not negligible when the length of short Fourier transforms is longer than $\sim$10–100 s [83].}

3. Computing cost

The computing time for one central processing unit (CPU) over a total observing time $T_{\text{obs}}$ in a frequency band from $f_{\text{min}}$ to $f_{\text{max}}$ is given by [78]

$$T = 2\pi\beta N_{\text{ifo}} T_{\text{drift}} T_{\text{obs}} T_{\text{SFT}}^{-1} N_{\text{sky}} (f_{\text{max}} - f_{\text{min}}),$$

where $T_{\text{SFT}}$ is the length of the short Fourier transforms (SFTs) used to compute the $F$-statistic [84], $N_{\text{ifo}}$ is the number of interferometers, $N_{\text{sky}}$ is the number of sky locations, $\beta$ is the percentage of time that the interferometers collect data (“duty cycle”), and $\kappa$ is the time to compute the $F$-statistic per template per SFT. The value of $\kappa$ depends on $T_{\text{SFT}}$ and the CPU architecture; we adopt the recent estimate that $\kappa = 4 \times 10^{-8}$ s for $T_{\text{SFT}} = 1800$ s.

We normally divide the full frequency band into multiple 1 Hz subbands to allow parallelized computing. For example, if we have 10$^2$ cores running in parallel, a search for $T_{\text{obs}} = 80$ days over the frequency band spanning 100–200 Hz in two detectors and with a fixed sky location ($N_{\text{sky}} = 1$) takes about 7 min to complete. This estimation is consistent with the real cost of our simulations below.

B. Sensitivity estimates

We would like to study the sensitivity of ground-based detectors to continuous GW signals from boson clouds around known BHs. For this purpose, we simulate signals consistent with the morphology described in Sec. III A and
study how well they can be recovered using the HMM tracking (Sec. IV B 1). We then translate expected strain sensitivities into detection horizons for boson signals with current and future ground-based detectors (Sec. IV B 2). Finally, we explore the impact of uncertainties in the source’s sky location (Sec. IV B 3).

1. Strain sensitivity

To study our sensitivity to waves from boson clouds, we inject synthetic signals with parameters consistent with the morphology described in Sec. III A into simulated Gaussian noise corresponding to two aLIGO detectors at design sensitivity [85]. The signal frequency and frequency derivative were chosen to be roughly in agreement with an optimal scalar cloud around the example BH discussed in Sec. IV B 3.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial frequency</td>
<td>( f_{\text{inj}} )</td>
<td>201.2 Hz</td>
</tr>
<tr>
<td>First derivative of ( f_{\text{inj}} )</td>
<td>( \dot{f}_{\text{inj}} )</td>
<td>( 1 \times 10^{-12} ) Hz(^{-1})</td>
</tr>
<tr>
<td>Right ascension</td>
<td>( \alpha_* )</td>
<td>23(^\circ)23(^\circ)26,60(^\circ)</td>
</tr>
<tr>
<td>Declination</td>
<td>( \delta_* )</td>
<td>58(^\circ)48(^\circ)0.0(^\circ)</td>
</tr>
<tr>
<td>Inclination</td>
<td>( \cos i )</td>
<td>([-1, 1])</td>
</tr>
<tr>
<td>Polarization</td>
<td>( \psi )</td>
<td>([0, 2\pi])</td>
</tr>
<tr>
<td>Initial phase</td>
<td>( \Phi_0 )</td>
<td>([0, 2\pi])</td>
</tr>
<tr>
<td>Gaussian noise PSD(^{1/2})</td>
<td>( S_{h}^{1/2}(f) )</td>
<td>( 4 \times 10^{-24} ) Hz(^{-1/2})</td>
</tr>
</tbody>
</table>

To demonstrate that HMM can accurately reconstruct boson signals, Fig. 9 presents two tracking examples for injected signals with \( h_0 = 5 \times 10^{-26} \) and parameters in Table II. The frequency random walks are such that \(|\delta f| \leq 0.1 \Delta f \) and \(|\delta f| \leq 0.5 \Delta f \) for panels (a) and (b), respectively. The optimal HMM paths (dark curves) match the injected path \( f_0(t) \) (light curves) closely: the root-mean-square error between the optimal HMM paths and the actual signals are \( \varepsilon = 1.16 \times 10^{-7} \) Hz = \( 0.16 \Delta f \) for panel (a) and \( 3.65 \times 10^{-7} \) Hz = \( 0.50 \Delta f \) for panel (b). These small discrepancies are mostly due to the frequency discretization carried out by the HMM algorithm.

To next quantify the efficiency of HMM tracking at detecting signals of different amplitudes. For concreteness, we assume a small uncertainty in the signal model by setting \(|\delta f| \leq 0.1 \Delta f \), as in Fig. 9(a). Figure 10 shows the receiver operator characteristic (ROC) curves for injected signals with four values of \( h_0 \), ranging from \( 2 \times 10^{-26} \) to \( 5 \times 10^{-26} \). For each signal amplitude, these curves show the detection probability (1 – \( P_d \), where \( P_d \) is the false-dismissal probability) as a function of required false-alarm probability threshold (\( P_a \)). For instance, if we demand a false alarm probability \( P_a = 1\% \), we can expect to detect a signal with \( h_0 = 4 \times 10^{-26} \) (5 \times 10^{-26}) with 84% (98%) of the time.

The detection threshold in continuous-wave searches is traditionally defined to be 95% false-dismissal rate at 1% false-alarm probability [75,79]. In our case, for an observation time of \( T_{\text{obs}} = 80 \) days with two aLIGO design detectors, this corresponds to a strain amplitude of \( h_0^{95\%} = 4.7 \times 10^{-26} \) for unknown inclination. Based on this, we will consider boson signals “detectable” if they reach an amplitude of \( h_0^{95\%} \) or higher for the observation conditions.

From the empirical result that \( h_0^{95\%} = 4.7 \times 10^{-26} \) obtained for the simulations above, it is straightforward to estimate how the sensitivity of the search would scale for different detector networks and observation times. The sensitivity scaling will be given by [78]

\[
h_0^{95\%}(f) \propto N_{\text{ifo}}^{-1/2} S_h(f)^{1/2}(T_{\text{drift}} T_{\text{obs}})^{-1/4},
\]

assuming a network of \( N_{\text{ifo}} \) detectors with power-spectral density (PSD) \( S_h(f) \) at the signal frequency.\(^8\)

\(^8\)This requires that all detectors have comparable sensitivities given by \( S_h(f) \); were this not the case, the \( N_{\text{ifo}}^{-1/2} S_h(f)^{1/2} \) factor would have to be replaced by the square root of the effective PSD \( S_{\text{eff}}(f) \), which is itself given by the harmonic mean of the PSDs for each detector.
Using this, Fig. 11 presents projected 95%-confidence strain upper limits, $h_{95\%}$, for different detectors as a function of GW frequency (assuming there is no detection). We show results for aLIGO design sensitivity\(^9\) (gray line), as well as proposed third-generation detectors: LIGO Voyager (yellow line), Cosmic Explorer (purple line), and Einstein Telescope (red line)\(^8\). All curves in Fig. 11 were produced assuming $N_{ifo} = 1$ and $T_{obs} = 1$ yr, but it is straightforward to rescale them for different configurations using Eq. (38). In particular for the Einstein Telescope, we show the sensitivity of a 10 km, $90^\circ$ interferometer in the “D” configuration—the sensitivity of the full triangular layout is up to 50% better for a circularly polarized wave\(^9\). Finally, note that these curves were obtained by effectively marginalizing over source orientation: they represent the value of $h_{95\%}$ marginalized over the distributions of $\cos \psi$ and $\psi$ in the ranges shown in Table II. To obtain the values of $h_{95\%}$ corresponding to optimal source orientation, one should divide the curves of Fig. 11 by a factor of $\sim 2.8$\(^9\).

Using this, Fig. 11 presents projected 95%-confidence strain upper limits, $h_{95\%}$, for different detectors as a function of GW frequency (assuming there is no detection). We show results for aLIGO design sensitivity\(^9\) (gray line), as well as proposed third-generation detectors: LIGO Voyager (yellow line), Cosmic Explorer (purple line), and Einstein Telescope (red line)\(^8\). All curves in Fig. 11 were produced assuming $N_{ifo} = 1$ and $T_{obs} = 1$ yr, but it is straightforward to rescale them for different configurations using Eq. (38). In particular for the Einstein Telescope, we show the sensitivity of a 10 km, $90^\circ$ interferometer in the “D” configuration—the sensitivity of the full triangular layout is up to 50% better for a circularly polarized wave\(^9\). Finally, note that these curves were obtained by effectively marginalizing over source orientation: they represent the value of $h_{95\%}$ marginalized over the distributions of $\cos \psi$ and $\psi$ in the ranges shown in Table II. To obtain the values of $h_{95\%}$ corresponding to optimal source orientation, one should divide the curves of Fig. 11 by a factor of $\sim 2.8$\(^9\).

The sky position of the source with respect to the detector does not impact the search sensitivity significantly because the variation due to the antenna pattern is averaged.

\(^9\)For aLIGO, we use the latest-available sensitivity projections\(^8\), which correspond to those in [1] but with a reflectivity of 32.5% in the signal recycling mirror (SRM).
It is useful to translate the projected strain sensitivities of Fig. 11 into detection horizons for boson signals from BHs with different parameters. The detection horizon is the farthest distance up to which we should expect to be able to detect an optimal boson signal—namely, a signal from a boson cloud that perfectly matches its host BH (to maximize intrinsic strain) and is optimally oriented with respect to the detector (to maximize measured strain).

Consequently, horizons are a measure of how well we can do in the best-case scenario and are, thus, not generally representative of most detections (see, e.g., [93] for an overview of distance measures in GW astronomy)—yet, they are a straightforward proxy for the reach of our instruments to this type of source. We compute this quantity for scalar clouds based on the results from Sec. III B; we defer computation of horizons for vector clouds until better numerical estimates of their intrinsic amplitudes become available and analysis methods suitable for higher frequency derivatives are developed.

Figure 12 shows the horizon luminosity distance (color) for scalar signals as a function of initial BH mass $M_i$ and spin $\chi_i$, for (a) design Advanced LIGO, (b) LIGO Voyager, (c) Cosmic Explorer, and (d) the Einstein Telescope. White contours indicate the values of the boson rest-energy $\mu$ (eV) that we would be able to probe with a BH of that mass and spin. The shaded region marks values for which we expect the signal to evolve too rapidly for current data-analysis techniques to handle, based on Eqs. (33) and (36). Note that this varies slightly among plots due to minor differences in redshift. In all cases we assume one year of uninterrupted observation by a single detector, as in Fig. 11.

The horizon plots were obtained by finding the luminosity distance at which an optimal cloud for the given BH parameters would become barely detectable, i.e., $h_{0}(d_L) = h_{0}^{95\%}(f_{\text{det}})$ for detector-frame signal frequency $f_{\text{det}}$. In order to obtain the relevant value of $h_{0}^{95\%}$, we rescale the curves of Fig. 11 using Eq. (38) to account for variations in $T_{\text{drift}}$. This is needed because the expected $f_{\text{det}}$ varies widely over the parameter space, affecting the maximum-allowed coherence time; cf. Eq. (36). We also take into account the fact that both frequencies and frequency derivatives get redshifted as the signal makes its way to Earth, i.e., $f_{\text{det}} = \frac{f_{\text{src}}}{(1+z)^{-1}}$ and $\dot{f}_{\text{det}} = \frac{f_{\text{src}}}{(1+z)^{-2}}$ for a BH at redshift $z$ and source-frame frequency $f_{\text{src}}$. Finally, we rescale the curves in Fig. 11 to obtain values corresponding to optimal source orientation, as explained above. All these different factors modulate the intrinsic strain inferred from Fig. 5 to yield Fig. 12.

For all detectors, the horizon generally increases with initial BH mass and spin, as expected from Fig. 5. Furthermore, higher masses and spins are expected to yield smaller $\dot{f}$’s, which enables longer coherent times (longer $T_{\text{drift}}$’s) and, thus, slightly higher sensitivity; cf. Eq. (38). Yet, this tendency is offset by the fact that heavier systems yield lower frequencies (Fig. 4), causing the horizon to quickly drop as signals reach the lower end of the detector’s sensitive band (cf. Fig. 11). Moreover,

10In an actual analysis, we might want to set a $T_{\text{drift}}$ shorter than that implied by Eq. (33) in order to allow for theoretical uncertainty in the predicted value of the frequency derivative.

11Assuming standard $\Lambda$CDM cosmology with present parameters: $\Omega_m = 0.308$, $\Omega_\Lambda = 0.692$, $\Omega_\kappa = 0.0$, $h = 0.678$ [94].

TABLE IV. Detection efficiency vs sky location ($h_0 = 5 \times 10^{-26}$).

<table>
<thead>
<tr>
<th>Right ascension</th>
<th>Declination</th>
<th>Detection efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>23h 23m 26.0s</td>
<td>58°48'0.0&quot;</td>
<td>0.98</td>
</tr>
<tr>
<td>23h 23m 26.0s</td>
<td>-59°35'0.0&quot;</td>
<td>0.97</td>
</tr>
<tr>
<td>23h 23m 26.0s</td>
<td>00°02'0.0&quot;</td>
<td>0.92</td>
</tr>
<tr>
<td>23h 23m 26.0s</td>
<td>88°48'0.0&quot;</td>
<td>0.98</td>
</tr>
<tr>
<td>23h 23m 26.0s</td>
<td>-89°18'0.0&quot;</td>
<td>0.98</td>
</tr>
<tr>
<td>05h 23m 26.0s</td>
<td>58°48'0.0&quot;</td>
<td>0.98</td>
</tr>
<tr>
<td>11h 23m 26.0s</td>
<td>58°48'0.0&quot;</td>
<td>0.98</td>
</tr>
<tr>
<td>17h 23m 26.0s</td>
<td>58°48'0.0&quot;</td>
<td>0.99</td>
</tr>
</tbody>
</table>
Signals from clouds around heavier BHs can more easily get redshifted out of the band. At the other end of the spectrum, the instruments we consider tend to be more sensitive at higher frequencies, but these correspond to lower masses and, thus, lower radiated power (for a given $\chi_i$). On the other hand, increasing the BH spin yields both higher GW amplitudes and, to an extent, frequencies. Unfortunately, however, lower masses and higher spins also result in high $\dot{f}$’s that make much of that part of parameter space inaccessible to current methods (shaded regions).

All this means, roughly, that the farthest horizons will be obtained for BHs with masses in the range $10^2 < M_i / M_\odot < 10^3$ and spin as high as possible, corresponding to boson masses within $10^{-14} < \mu / \text{eV} < 10^{-12}$ (depending on $\chi_i$). Even outside this range, these horizons are significantly more distant than the sources at which these searches are generally directed, which tend to lie within the Milky Way (see, e.g., [75]).

As a concrete example, consider again a GW150914-like remnant with $M_i = 60 M_\odot$ and $\chi_i = 0.70$. As we saw back in Sec. III B 1, this BH would be best matched by a scalar...
bozon with $\mu = 4 \times 10^{-13}$ eV. A scalar of that mass would yield an optimal cloud ($l = m = 1$, $n = 0$) that radiates gravitational waves ($\tilde{l} = \tilde{m} = 2$) at a source-frame frequency of 191 Hz with characteristic amplitude $h_0 = 5.2 \times 10^{-26} (5 \text{ Mpc}/d_L)$, corresponding to the peak in Fig. 1. From the intersection of the dotted lines in Fig. 12(a), we see that such a signal would be detectable, at most, up to 12 Mpc away with one aLIGO detector at design sensitivity observing continuously for 1 yr—or, equivalently, $\sim 20$ Mpc for three such detectors. This agrees with previous estimates in [18].

Prospects are even better for third-generation detectors, with farther horizons over most of the parameter space. Third-generation detectors would offer significant improvements for almost any target with $M_i \lesssim 10^3 \, M_\odot$. In particular, we find that Cosmic Explorer could reach ranges of over $10^8$ Mpc ($\zeta \gtrsim 1.4$) for fast-spinning BHs over a range of masses [Fig. 12(c)]. The Einstein Telescope could also reach such distances, but for a more limited choice of parameters, and would have a shorter reach for most of the sources we consider; on the other hand, this instrument would outperform all others at higher masses (lower frequencies). Some representative values are presented in Table V to ease comparison between instruments.

We underscore that the horizons computed in this section (in both Fig. 12 and Table V) correspond to single detectors and not to a network. This is to facilitate rescaling for any specific configuration. For instance, without taking into account detector orientation and signal polarization, the horizon for a GW150914-like remnant seen by a network of two Cosmic Explorers would be roughly $\sim \sqrt{2} \times 300$ Mpc $\approx 424$ Mpc. For a network composed of one Cosmic Explorer and one Einstein Telescope in the full triangular “D” configuration, this would be instead $\sim \sqrt{(1.06 \times 100)^2 + 300^2}$ Mpc $\approx 320$ Mpc, where the factor of 1.06 accounts for the 6% improvement in sensitivity to a linearly polarized wave when using the full triangular Einstein Telescope instead of the 90° configuration assumed in Fig. 12 and Table V [91,92].

As mentioned above, for each BH mass and spin, our horizon computation assumes the coherent time window $T_{\text{drift}}$ is set to the largest value that can accommodate the expected frequency derivative [cf. Eq. (36)]. This choice is tailored to optimize sensitivity over the full parameter space [cf. Eq. (38)]. Instead, we could lengthen $T_{\text{drift}}$ to slightly increase horizon distances, at the expense of losing all sensitivity to signals with $|\dot{f}| > 1/(2T_{\text{drift}}^2)$. This would not be advisable except for targets somehow known not to follow Eq. (33). Figure 13 shows how the horizon scales with $T_{\text{drift}}$ for fixed $T_{\text{obs}} = 1$ yr, as implied by Eq. (38), together with the maximum $|\dot{f}|$ allowable for any given $T_{\text{drift}}$ (top axis). The different curves correspond to the representative systems of Table I, with vertical dotted lines marking the $T_{\text{drift}}$ assumed in Table V and Fig. 12, i.e., the highest $T_{\text{drift}}$ compatible with Eq. (33).

Note that the horizons for $T_{\text{drift}} = T_{\text{obs}} = 1$ yr shown in Fig. 13 are the same as would be obtained in a fully coherent search of that duration, if we had enough computing resources to explicitly search over $\dot{f}$ as well as $f$. This is the optimal sensitivity we could ever hope to achieve with 1 yr of data (scaling as $\sqrt{T_{\text{obs}}}$. Even in this idealized case, the aLIGO range to a GW150914-like remnant would fall short of 100 Mpc.

### 3. Effect of sky-location uncertainty

We would like to understand the effect of uncertainty in the source sky location on the HMM tracking, mainly

![Horizon scaling with $T_{\text{drift}}$.](image)

---

**TABLE V.** Scalar-cloud horizons (Mpc) for representative BHs (boson and signal parameters shown in Table I). The bold row corresponds to the intersection of the dotted lines in Fig. 12.

<table>
<thead>
<tr>
<th>$M_i$ ($M_\odot$)</th>
<th>$\chi_i$</th>
<th>aLIGO</th>
<th>Voy</th>
<th>CE</th>
<th>ET</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.90</td>
<td>0.2</td>
<td>0.4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>0.90</td>
<td>3</td>
<td>6</td>
<td>35</td>
<td>24</td>
</tr>
<tr>
<td><strong>60</strong></td>
<td><strong>0.70</strong></td>
<td><strong>12</strong></td>
<td><strong>49</strong></td>
<td>$3 \times 10^2$</td>
<td>$1 \times 10^2$</td>
</tr>
<tr>
<td>60</td>
<td>0.90</td>
<td>60</td>
<td>$2 \times 10^2$</td>
<td>$1 \times 10^3$</td>
<td>$7 \times 10^2$</td>
</tr>
<tr>
<td>200</td>
<td>0.85</td>
<td>$2 \times 10^2$</td>
<td>$6 \times 10^2$</td>
<td>$5 \times 10^3$</td>
<td>$1 \times 10^3$</td>
</tr>
<tr>
<td>300</td>
<td>0.95</td>
<td>$5 \times 10^2$</td>
<td>$2 \times 10^3$</td>
<td>$2 \times 10^4$</td>
<td>$4 \times 10^3$</td>
</tr>
</tbody>
</table>

---

084042-18
motivated by the prospect of following up compact-binary mergers. In order to find a continuous signal coming from some area in the sky, we would have to analyze the gravitational-wave data with the HMM multiple times assuming slightly different sky locations to tile the patch where the source is thought to lie. The number of iterations (the number of “templates”) needed, \( N_{\text{sky}} \), is determined both by the size of the target area and by the sky resolution of the analysis, which is in turn tied to the frequency resolution of the search and the amplitude of the signal.

To estimate \( N_{\text{sky}} \), we run two sets of simulations by injecting signals with \( h_0 = 5 \times 10^{-26} \) and \( h_0 = 2 \times 10^{-25} \) into simulated Gaussian noise for two aLIGO detectors at design sensitivity. All other parameters, including the sky location, are as listed in Table II. HMM searches are then conducted using the settings shown in Table III, but for a grid of sky locations in the neighborhood of the injected signal. In other words, for each injection amplitude (\( h_0 = 5 \times 10^{-26} \) or \( h_0 = 2 \times 10^{-25} \)) and for each sky location assumed by the search, we inject a signal with random orientation and phase, but a location fixed to the value in Table II; we repeat this 200 times to obtain detection probabilities for each of the search locations.

The results of this study are summarized in Fig. 14 for both the soft (left panel) and loud (right panel) injections. Color in these figures encodes the detection efficiency \( P_d \) for searches assuming a sky location indicated by their offset in right ascension (\( \Delta \alpha \)) and declination (\( \Delta \delta \)) with respect to the true location (\( \Delta \alpha = \Delta \delta = 0 \)). The white (black) contours mark points at which signals were detected 90% (50%) of the time at 1% false-alarm probability. Notice that, for the weaker signal, the 90%-contour encloses an area of \( \sim 0.001 \text{deg}^2 \) around the true location, while for the stronger signal this is roughly 4 orders of magnitude larger. We may take the size of the 90% contours as indicative of the spacing of the sky grid needed to capture a signal.

In an actual search, we need \( N_{\text{sky}} \approx 10^3 \) sky templates per deg\(^2\) to detect a weak signal near the detection limit. The sky resolution generally agrees with other coherent or semicoherent continuous-wave search methods [95,96]. Here we discuss the search feasibility given the required number of sky templates. As a representative example, consider that the existing three-detector network (Advanced LIGO and Advanced Virgo) was able to localize the binary-neutron start merger GW170817 to a sky region spanning \( \sim 30 \text{ deg}^2 \) with 90% credibility [9]. Based on Fig. 14, we would then need \( N_{\text{sky}} \approx 10^4 \) to obtain 90% detection efficiency of a signal with \( h_0 = 5 \times 10^{-26} \) lying somewhere inside the GW170817 90% credible region; by contrast, \( N_{\text{sky}} \approx 3 \) would suffice for a signal with \( h_0 = 2 \times 10^{-25} \). Based on estimates for standard computing architectures and algorithm settings, finding a signal at 200 Hz with \( h_0 = 5 \times 10^{-26} \) in a \( \sim 30 \text{ deg}^2 \) region (\( T_{\text{obs}} = 80 \text{ d} \), \( N_{\text{ifo}} = 2 \)) would take 5 days of computing on 1 k CPUs, which is feasible but not cheap. Note that \( N_{\text{sky}} \) scales as \( f^2 \) [95–98], so more templates will be required at higher frequencies. Because computing cost scales directly with \( N_{\text{sky}} \), the burden will be vastly reduced once more gravitational detectors join the network and the sky locations reach the projected \( O(1 \text{ deg}^2) \) [99].

### C. Potential sources

The discussion thus far has been largely unconcerned with the kind of BH being targeted. In this section, we flesh out the implications of the above conclusions for two types of promising sources: remnants from compact-binary mergers (Sec. IV C 1) and BHs in x-ray binaries (Sec. IV C 2).

![FIG. 14. Sky resolution. Color shows detection efficiency \( (1 - P_d) \), for \( P_d \) the false-dismissal probability) as a function of offsets in right ascension (x axis) and declination (y axis) with respect to the true location for injections with \( h_0 = 5 \times 10^{-26} \) (left) and \( h_0 = 2 \times 10^{-25} \) (right). All other injection parameters are as in Table II and search settings are shown in Table III. The left (right) plot was interpolated from a square grid with 5 (7) sky locations on each side.](image-url)
1. Merger remnants

As pointed out before [18,21], nearby CBC remnants would be ideal targets for searches for gravitational signals from ultralight bosons. Because we witness their birth firsthand, the age of remnant BHs is perfectly known and their mass and spin well constrained. This would enable accurate estimation of the continuous-signal amplitude that should be expected for any given \( \alpha \) (Sec. III B 1), allowing us to potentially place interesting constraints on the existence of matching bosons. Furthermore, the location and orientation inferred from the initial chirp would allow us to take advantage of existing infrastructure for directed searches for continuous waves in LIGO and Virgo data (Secs. IV A and IV B).

Ideally, we would follow up any and all mergers, as soon as a reasonable time has passed for the cloud to form (Sec. III B 2). In practice, however, we may be limited by the uncertainty in the sky location. Signals detected with only two instruments will be too loosely localized to allow for follow-up (e.g., LIGO’s first detection was localized to the sky region of \( \sim 260 \text{ deg}^2 \) [3]). Fortunately, as we saw in Sec. IV B 3, the localization provided by a three-detector network would already be manageable with existing computational resources (e.g., \( \sim 28 \text{ deg}^2 \) for the binary neutron star [9]). This will be further improved when more, and more sensitive, detectors join the effort: a network including LIGO India [100,101] and KAGRA [102,103], on top of the three existing detectors operating at design sensitivity, is expected to routinely locate events to an area of order \( \sim 1 \text{ deg}^2 \) [99]. Regardless of the number of instruments, events with an electromagnetic counterpart (e.g., mergers involving a neutron star) will always be sufficiently well localized. The presence of an electromagnetic counterpart does not directly lower the detection threshold, but does significantly reduce the number of required sky templates and makes it practical to achieve the desired sensitivity with a reasonable computing cost (Sec. IV B).

Extracting information about bosons from one of these observations would also require good knowledge of the remnant distance and orientation. This is required to translate strain (\( h_0 \)) constraints into limits on radiated power (\( \dot{E}_{GW} \)), which can then be turned into statements about the existence of a boson with a given mass [cf. Eqs. (19) and (28)]. In particular, if the source is too far away, the signal from the hypothetical cloud would be undetectable at Earth, rendering constraints on its amplitude moot. Thus, if the distance is not determined by other means (e.g., association with a host galaxy), the implications for bosons will be contingent on the uncertainty in the luminosity distance inferred from the CBC observation.

In the case of a scalar cloud, for most remnant masses and spins, the source would have to be relatively close for the signal to be detectable by ground-based detectors (Fig. 12). For a second-generation network at design sensitivity, the horizon would lie below 100 Mpc \( \times \sqrt{N_{ifo}} \) for most signal parameters [Fig. 12(a)]. Given that we have yet to observe a BH merger that close [3–9], this projection is not too auspicious. Yet, note that the horizon can reach close to \( 10^3 \text{ Mpc} \times \sqrt{N_{ifo}} \) in some regions of parameter space—although taking advantage of this with existing algorithms would require a population CBCs yielding remnants with \( M \gtrsim 100 M_\odot \) (\( f < 10^{-8} \text{ Hz/s} \)).

Estimates of rates from BH population models were provided in [18] based on the nonrelativistic approximation to the amplitude Eq. (29).

As we saw in Sec. IV B 2, prospects are better for next-generation detectors, especially Cosmic Explorer [Fig. 12(c)]. Even then, a good reach to remnants with \( M \sim \mathcal{O}(10 M_\odot) \) would require spins roughly \( \gtrsim 0.85 \), possibly less depending on the mass. Although we have not yet observed any such events [3–9], numerical-relativity simulations routinely produce remnants with such spins [104–107]. Note that the horizons for boson signals are always significantly closer than those for compact-binary coalescences [87].

The vector case is slightly different. Detection horizons are in principle considerably farther for vector clouds due to the intrinsically higher radiated power (see Sec. II B), making most remnant masses and spins accessible. However, more radiated power also means shorter cloud lifetimes and, consequently, faster rates of change for the signal amplitude and frequency [cf. Eq. (34)]. For much of the parameter space, the expected signal would then evolve too rapidly for existing continuous-wave algorithms to handle (see discussion in Sec. IV B). Therefore, the more powerful and quickly evolving vector signals would currently not be detectable, effectively reducing our horizon to such sources. Detection rates for vectors taking this into account were estimated in [21] by using the nonrelativistic approximation of Eq. (30).

To get a sense of our potential reach to vector signals, we may use Eq. (38) to obtain vector horizons starting from the scalar ones in Fig. 12. To do this, note that \( h_0^{(v)}/h_0^{(s)} \approx 5 \times 10^3 (0.1/\alpha)^2 \) by Eqs. (29) and (30), while \( j^{(v)}/j^{(s)} \approx 3 \times 10^5 (0.1/\alpha)^4 \) by Eqs. (33) and (34). Then, if we had the analysis infrastructure to handle the quickly varying vector signals and could coherently search over \( f \) and \( \dot{f} \) (cf. Fig. 13), our horizon for a vector signal from a GW150914-like remnant would be \( \sim 1575 \text{ Mpc} \), instead of the \( \sim 12 \text{ Mpc} \) for scalar signals (Table V). This assumes \( \alpha = 0.1 \), \( T_{\text{obs}} = 1 \text{ day} \), and \( T_{\text{drift}} = 23 \text{ s} \), which is the highest value consistent with \( j^{(v)} \) for that system.

Even for scalar signals, the restriction to small frequency derivatives is quite detrimental, preventing us from accessing higher boson masses (lower BH masses). Because the estimates of Eqs. (33) and (34) are only approximate, there is still sense in searching for signals with \( f > 10^{-8} \text{ Hz s}^{-1} \) in the shaded regions of Fig. 12—although a negative result
would be harder to interpret as evidence against the existence of a boson in that mass range. As suggested above, this is a strong motivation to adapt analysis techniques that can handle quickly evolving continuous signals to make them suitable for boson searches—this is work in progress.

Finally, note that we expect to infer the remnant mass and spin from the CBC signal with enough precision to obtain a reasonably accurate prediction of the cloud GW amplitude for any given boson mass. For instance, the mass and dimensionless spin of the GW150914 remnant were each measured with one-sided relative errors of under 10% at 90% credibility, which is sufficiently narrow to make a follow-up search possible (Sec. IV B). The characteristic magnitude of such errors is expected to be significantly reduced for detections at a higher signal-to-noise ratio, which should be commonplace once the current network achieves design sensitivity and for next-generation detectors (see, e.g., [99,108,109]).

In any case, a simplistic way to deal with parameter uncertainty would be to compute the optimal strain for a cloud around a BH corresponding to the upper bounds of the mass and spin credible intervals. A value computed that way would itself be an upper limit on the boson strain, because this quantity scales directly with mass and spin (Fig. 5). Alternatively, a rigorous statistical analysis would take in the full-dimensional posterior probability density on the BH parameters (intrinsic and extrinsic) and marginalize over all parameters to obtain a posterior on the expected boson strain as a function of $\alpha$. The development of this more sophisticated strategy is a work in progress.

2. X-ray binaries

Another type of potentially interesting targets are known BHs in x-ray binaries (see, e.g., [110,111] for reviews treating such systems). The relevance of x-ray binaries to this research program has been pointed out since the outset (e.g., [14,16,17,21]). They have the advantage of being much closer and better located in the sky compared to the CBC remnants, with good measurements of their mass and, in some cases, spin [112]. In fact, some limits on the boson-mass space have already been placed contingent on these measurements, roughly excluding the mass interval $10^{-12} \leq \mu / \text{eV} \leq 10^{-11}$ for scalars [18,38] and $10^{-13} \leq \mu / \text{eV} \leq 10^{-11}$ for vectors [21,38]. Unfortunately, there is large uncertainty about the age and history of these systems, as well as important caveats about the systematics affecting their spin measurements [42,43]. Furthermore, the effect of the active astrophysical environments surrounding these BHs is only understood at the order-of-magnitude level [17,21,39]. For all these reasons, boson constraints derived from existing observations of x-ray binaries should be interpreted with caution.

There are also data-analysis challenges intrinsic to signals coming from sources in a binary system: the Doppler modulation due to the motion of the source within the binary causes the signal power to spread over multiple frequencies. The signal must then be collected from “orbital sidebands” that span a frequency band

$$B \approx 4\pi f_0 a_0 / c P,$$

where $a_0$ is the BH’s projected semimajor axis, $P$ is the orbital period, and $c$ is the speed of light [77]. Frequency-domain matched filters, such as those presented in [77,113], can be applied to sum up the distributed signal power using (imperfect) knowledge of the orbital parameters. Those methods would generally demand $B \lesssim 0.5$ Hz in order to achieve the required sensitivity (see Sec. IV B 1). It becomes prohibitively expensive to detect a weak signal from a binary if the orbital parameters (e.g., $P$, $a_0$, and time of passage through the orbit’s ascending node $T_p$) are poorly measured. More details can be found in Sec. III B of Ref. [113].

As an example, consider the nearby Cygnus X-1 binary, which has been proposed as an interesting target for boson searches [16]. If we take the source parameters in Table VI and assume $f_0 \sim 500$ Hz based on Eq. (32), then the power of a signal from Cygnus X-1 would span a frequency band $B \approx 0.3$ Hz, which is acceptable using existing methods. The strain amplitude estimated given the source distance is $\sim 10^{-22} - 10^{-21}$, possibly detectable with aLIGO even with imperfect knowledge of the binary orbit [77]. However, the non-negligible uncertainty in $a_0$ and the limited knowledge of $T_p$ would require searching a large number of templates.

Unfortunately, for most of the interesting x-ray systems, the orbital parameters are not well measured electromagnetically, and the sidebands in most of the high-mass x-ray binaries would be broader than $\sim 1$ Hz. BHs in low-mass x-ray binaries (i.e., the companion star is less massive and hence $B$ is generally narrower) with well-measured parameters are likely better candidates. Due to these considerations, only a handful of potential sources will be of interest and targets in x-ray binaries will need to be chosen carefully. Future improved electromagnetic measurements of x-ray systems would benefit the analysis by improving overall sensitivity and result interpretability.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass ($M_\odot$)</td>
<td>$M$</td>
<td>$14.8 \pm 1.0$</td>
<td>[111,114]</td>
</tr>
<tr>
<td>Spin</td>
<td>$\chi$</td>
<td>$\geq 0.95$</td>
<td>[111]</td>
</tr>
<tr>
<td>Right ascension</td>
<td>$\alpha_*$</td>
<td>$19^h 58^m 22^s$</td>
<td>[115]</td>
</tr>
<tr>
<td>Declination</td>
<td>$\delta_*$</td>
<td>$35^\circ 12' 6''$</td>
<td>[115]</td>
</tr>
<tr>
<td>Inclination (deg)</td>
<td>$i$</td>
<td>$27.1 \pm 0.8$</td>
<td>[114]</td>
</tr>
<tr>
<td>Distance (kpc)</td>
<td>$r$</td>
<td>$1.86_{-0.11}^{+0.12}$</td>
<td>[111]</td>
</tr>
<tr>
<td>Orbital period (days)</td>
<td>$P$</td>
<td>$5.6$</td>
<td>[114,116]</td>
</tr>
<tr>
<td>Proj. semimajor axis (l-s)</td>
<td>$a_0$</td>
<td>$25.56_{-5.11}^{+3.15}$</td>
<td>[114]</td>
</tr>
</tbody>
</table>
V. CONCLUSION

Black-hole superradiance could be the key that allows gravitational-wave detectors to uncover evidence of ultralight bosons, thus bringing particle physics within the reach of gravitational-wave science. In this paper, we explored the prospect for achieving this goal by looking for the continuous gravitational signals expected from scalar and vector clouds, using searches directed at known BHs.

We began by reviewing the physics of boson clouds (Sec. II) and examined in detail the properties of continuous signals from clouds around a known BH (Sec. III). In doing so, we hoped to provide a bridge between the theory and data-analysis literatures. We then used numerical techniques, combined with the latest analytic results, to compute the features of gravitational waves emitted by scalar clouds around BHs with different initial parameters (Figs. 1–8).

We put forward the use of HMM tracking [9,77–79] to carry out directed searches for boson signals (Sec. IV A). This strategy is well suited to searches for gravitational waves from boson clouds because its computational efficiency enables the coverage of a wide range of signal parameters, and because it does not rely on restrictive waveform models. This makes it ideal to search for signals over a broad frequency band (cf. Fig. 1), even when the location of the source is only loosely known and when there is potential uncertainty in the signal morphology. We demonstrated this through a series of Monte Carlo simulations (Sec. IV B).

From our simulations, we obtained an empirical estimate of the sensitivity of directed searches to boson signals in data from future ground-based detectors: aLIGO design, LIGO Voyager, Cosmic Explorer, and the Einstein telescope (Fig. 11). For scalar clouds, we translated the expected strain sensitivities into detection horizons for those four detectors (Fig. 12), assuming one year of observation by a single detector. We found that, for a second-generation network at design sensitivity, the horizon would lie below $100 \, \text{Mpc} \times \sqrt{N_{\text{ifo}}}$ for most signal parameters; prospects are better for next-generation detectors, especially Cosmic Explorer, for which horizons could reach up to $\sim 10^5 \, \text{Mpc}$. Generally speaking, these horizons lie much farther than the sources at which continuous-wave searches are generally directed [75], but significantly closer than horizons for compact-binary coalescences [87]. Some representative values are shown in Table I, and their scaling with the drift time, one of the primary algorithm settings, is shown in Fig. 13.

In computing signal amplitudes from scalar clouds, we numerically solved the evolution equations governing cloud growth and made use of numerical estimates from BH perturbation theory to obtain the radiated power [20]. Furthermore, to estimate horizons, we incorporated the effect of redshifts on the signal frequency and frequency derivative. We also took into account that the settings of the search algorithm should be varied across parameter space for optimal performance. This allowed us to obtain sensitivity estimates that should be more reliable than previously published projections.

Finally, we discussed implications for the follow-up of remnants from compact-binary coalescences (Sec. IV C 1), as well as BHs in x-ray binaries (Sec. IV C 2). We explored the impact of uncertainties in the source’s sky location and showed that HMM tracking will be able to efficiently cover the localization credible regions obtained from CBC signals with a network of at least three detectors. We also discussed the challenges intrinsic to vector signals that make their analysis difficult in spite of their higher radiated power. We emphasized the strong motivation to extend existing search techniques to handle signals with higher frequency derivatives, so as to bring a significant portion of the scalar and vector signal space into reach. The implementation of such techniques, as well as development of statistical strategies to rigorously handle uncertainty in BH parameters, is a work in progress.

ACKNOWLEDGMENTS

The authors thank Evan Hall for useful input regarding the sensitivities of present and future detectors. M. I. and R. B. thank Asimina Arvanitaki, Masha Baryakhtar, William East, and Robert Lasenby for organizing the meeting “Searching for New Particles with Black Hole Superradiance” held at the Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development & Innovation. M. I. and L. S. are members of the LIGO Laboratory. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY–0757058. Support for this work was provided by NASA through the NASA Hubble Fellowship Grant No. HST-HF2-51410.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under Contract No. NAS5-26555. L. S. was supported by an Australian Research Training Program Stipend Scholarship and the Albert Shimmins Fund at earlier stages of this project. The research was also supported by Australian Research Council (ARC) Discovery Project No. DP170103625 and the ARC Centre of Excellence for Gravitational Wave Discovery CE170100004. R. B. acknowledges financial support from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 792862. This paper carries LIGO Document No. LIGO–P1800270.
APPENDIX: FREQUENCY DRIFT

The gravitational self-energy of the cloud affects the boson’s eigenfrequencies and, consequently, the gravitational-wave frequency. As the cloud dissipates due to gravitational emission, this causes an increase in the emitted signal frequency, similar to what happens in a compact-binary coalescence. Therefore, we may treat the system adiabatically to obtain the frequency drift from the radiated power. This computation was presented in [21] for vectors, and we reproduce it here for scalars.

The gravitational self-energy of a bound state per particle is given by

\[ U_c = -G \frac{m_p}{M_c} \int \frac{\rho(r, \theta, \phi) m(r)}{r} d^3x, \quad (A1) \]

where \( m_p \) is the mass of the boson field, \( M_c \) is the overall mass in the cloud, \( \rho(r, \theta, \phi) \) is its density, and \( m(r) \) is the mass of the cloud enclosed in the radius \( r \), namely

\[ m(r) = \int_0^r \rho(r, \theta, \phi) d^3x. \quad (A2) \]

The rate of change of the GW frequency can then be estimated by [21]

\[ \dot{f} \approx \frac{1}{2\pi \hbar} \times 2U_c. \quad (A3) \]

As the cloud dissipates, the total mass of the system decreases, causing the binding energy to increase (\( U_c < 0 \)) and the GW frequency to increase.

The dominant scalar field mode can be approximated by the \( c = m = 1 \) hydrogen wave function, with a density \( \rho \) given by Eq. (11) in Ref. [35]. After some algebra one finds that, at leading order in \( \alpha \),

\[ \dot{f} \approx \frac{93}{1024\pi R} \frac{c^3}{G} \frac{\dot{E}_G}{M} = \frac{93 c^3 \alpha^3 M_c}{1024 \pi G M^2 \tau_G}, \quad (A4) \]

where we have used \( M_c = -\dot{E}_G c^{-2} \) and \( \dot{E}_G = M_c c^2 / \tau_G \). Using Eq. (13) to approximate \( M_c \sim \alpha M \), (valid in the limit \( \alpha \ll 1 \) for \( m = 1 \)) and Eq. (23) to write \( \tau_G \sim GM^2 / (0.025 M \alpha^4 c^3) \), we get

\[ \dot{f} \approx 3 \times 10^{-14} \text{ Hz/s} \left( \frac{10 M_\odot}{M} \right)^2 \left( \frac{\alpha}{0.1} \right)^{19} \chi_i. \quad (A5) \]


