Probing ultralight bosons with binary black holes

Baumann, D.; Chia, H.S.; Porto, R.A.

DOI
10.1103/PhysRevD.99.044001

Publication date
2019

Document Version
Final published version

Published in
Physical Review D

Citation for published version (APA):
We study the gravitational-wave (GW) signatures of clouds of ultralight bosons around black holes (BHs) in binary inspirals. These clouds, which are formed via superradiance instabilities for rapidly rotating BHs, produce distinct effects in the population of BH masses and spins, and a continuous monochromatic GW signal. We show that the presence of a binary companion greatly enriches the dynamical evolution of the system, most remarkably through the existence of resonant transitions between the growing and decaying modes of the cloud (analogous to Rabi oscillations in atomic physics). These resonances have rich phenomenological implications for current and future GW detectors. Notably, the amplitude of the GW signal from the clouds may be reduced, and in many cases terminated, much before the binary merger. The presence of a boson cloud can also be revealed in the GW signal from the binary through the imprint of finite-size effects, such as spin-induced multipole moments and tidal Love numbers. The time dependence of the cloud’s energy density during the resonance leads to a sharp feature, or at least attenuation, in the contribution from the finite-size terms to the waveforms. The observation of these effects would constrain the properties of putative ultralight bosons through precision GW data, offering new probes of physics beyond the Standard Model.

DOI: 10.1103/PhysRevD.99.044001

I. INTRODUCTION

The recent detections of gravitational waves (GWs) by the LIGO/Virgo Collaboration [1–5] mark the beginning of multimessenger astronomy [6] and the birth of “precision gravity” [7,8]. Binary systems, of comparable masses or extreme-mass ratios, will become the leading probe to test gravitational dynamics and the physics of compact objects, such as black holes (BHs) and neutron stars (NSs), under unique conditions. While it is indisputable that future GW observations will play a transformative role in astrophysics [9], it is less clear what impact these measurements will have on other branches of physics, and in particular, whether they can shed light on phenomena beyond the Standard Model. In principle, the nonperturbative regime of gravitational dynamics may carry imprints of “new physics”; see e.g., [10]. However, the need for numerical modeling, together with the small number of cycles involved, may hinder our ability to pinpoint different scenarios, if restricted only to the merger phase. In contrast, an accurate analytic reconstruction of the signal during the inspiral, in combination with simulations for the late stages of the coalescence, offers a unique opportunity to study physics beyond the Standard Model through GW precision data.

Using GW observations as probes of new physics is challenging, mainly due to the efficient decoupling of short-distance physics from long-distance observations. Indeed, finite-size effects, characterized by higher-derivative terms in a “worldline” effective field theory (EFT) approach [11–22], encapsulate the physics at scales shorter than the size of the objects sourcing GWs in a binary system. Because Einstein gravity is derivatively coupled, these terms scale with high powers of the ratio between the size and the separation of the bodies. For a nonrotating compact object, the first correction to the structureless point-particle approximation is due to tidal effects and scales with the fifth power of its size. It is often parametrized in terms of the so-called (tidal) “Love numbers,” e.g., [23,24], which are analogous to susceptibilities in electrodynamics. These Love numbers modify the phase of the GW signal for inspiraling binary systems at fifth post-Newtonian (5PN) order [25,26]. For rapidly rotating bodies, finite-size effects become relevant already at 2PN, through intrinsic spin-induced multipole moments [11,27,28]. In either case, the challenge is to extract the...
parameters of the source accurately [29–31], which requires very precise waveforms. As we shall see, both the Love numbers and the spin-dependent multipoles may carry the imprint of new degrees of freedom. The Love numbers in particular offer a unique diagnostic, since they vanish for BHs in Einstein gravity [23,24,49,50], and therefore any nonzero value would point to physics beyond the Standard Model [7,21,22].

New physics may also appear at distances larger than the size of the objects in a binary. This may entail long-distance modifications of general relativity or extra fields as in scalar-tensor theories. In either case, if the scale of new physics is shorter than the separation between the bodies, decoupling still applies, as far as the computation of GW observables is concerned. For example, modifications of general relativity can be evidenced during the inspiral phase through higher-derivative corrections to the Einstein-Hilbert action; see e.g., [51]. On the other hand, additional degrees of freedom can have effects at longer distances if their Compton wavelength is larger than the size of the compact objects. This will lead, as we shall see, to larger natural values for the finite-size coefficients in the effective theory, enhanced by powers of the ratio between the Compton scale and gravitational radius. In this manner, new physics may manifest itself at scales larger than anticipated from the Standard Model plus Einstein gravity alone.

A notable example for physics beyond the Standard Model are “ultralight” particles [52]. Even though the masses for these particles can be many orders of magnitude smaller than those of known elementary particles, they are technically natural if the coupling to ordinary matter is very weak. At the same time, the required weak coupling makes detecting these particles by traditional experimental means extremely challenging. On the other hand, their large Compton wavelengths mean that, if present in nature, they will be efficiently produced by the superradiance instability of rapidly rotating BHs [53–59]. For bosonic fields, superradiance creates a classical condensate which can carry a significant amount of the total mass and angular momentum of the system; see [60] for a review. Rather quickly, on cosmological/astrophysical scales, this leads to the BH carrying an “atmosphere,” with a large “cloud” of the field co-rotating with the BH (see Fig. 1). In the nonrelativistic limit, the eigenfunctions of the system are determined by a Schrödinger-like equation and the whole setup is sometimes referred to as a “gravitational atom.”

FIG. 1. Schematic illustration of the evolution of the boson cloud during the inspiral. The system experiences a resonance when the orbital frequency matches the energy difference between growing and decaying modes. The orbits for which the resonance occurs are indicated by the gray bands, marking the point when the cloud may rapidly deplete leading to novel GW signatures.

The bosons which can be produced in this manner span a vast range of masses, from $10^{-20}$ to $10^{-10}$ eV, corresponding to BHs of a few to billions of solar masses. At the upper end of this range we may find particles which play the role of the QCD axion [61–63], while those at the lower end offer alternative dark matter candidates, both with real [64] and complex [65,66] scalars. In addition, a plethora of ultralight axionlike particles arise also as a natural consequence of most string compactifications [67,68]. The “axiverse” [69], as it is commonly known, includes fields with a wide range of masses, increasing the chances some of them may have formed boson clouds through superradiance instabilities.

In this paper, we investigate the GW signatures from boson clouds surrounding BHs, when the latter are part of a binary. As we will show, the existence of a companion greatly enriches the dynamical evolution of the entire system. One of the most important consequences of the gravitational perturbation occurs when the orbital frequency of the binary matches the energy difference between growing and decaying modes of the cloud, leading to resonant transitions between the two states. This effect is analogous to Rabi transitions in the hydrogen atom under…

...The current state-of-the-art analytic computations are approaching 4PN order [28,32–48], moving forward toward the key 5PN threshold [7,8].
the influence of an oscillating external field. The reso-
nances can deplete (or significantly attenuate) the energy
density stored in the cloud, leading to rich phenomenolo-
gical implications and powerful new probes of physics
beyond the Standard Model.

For real bosonic fields, there are two main sources of
GW radiation: (i) a continuous monochromatic signal from
the cloud [70] and (ii) emission from the two-body system,
that includes the BH-cloud system and its companion. The
former depends on the specific configuration for complex
scalar fields around BHs, and may be absent for axisym-
metric condensates. Both signals, when present, will be
affected by the dynamical evolution of the cloud in the
binary, in particular as the system evolves through the
resonances. Notably, the amplitude of the continuous signal
emitted from the cloud may be reduced, and in many cases
terminated, much before the binary merger. The boson
clouds can also reveal themselves through finite-size effects
in the GW signal from the binary during the inspiral, such
as spin-induced multipole moment(s) and tidal Love
number(s). The time dependence of the cloud during the
resonance leads to a sharp feature, or at least an attenuation,
in the contribution from finite-size effects to the wave-
forms. The observation of these phenomenological effects
would help us elucidate the nature of the BH-cloud system,
and the putative ultralight bosons, through precision
GW data.

Using high-precision GW observations to constrain the
particle content in the Universe resembles the discipline
collider physics. In our case, a scattering process takes
place between the rotating BH and an incident wave of
the bosonic field. The superradiant amplification of the
scattered wave is a type of “resonance” that occurs when
the Compton wavelength of the field exceeds the size of
the BH horizon. This resonance creates a large number
of ultralight bosons in the intermediate state, which, in
the case of real scalars, then decay via self-annihilation
into GWs. These GWs can be thought of as many soft
quanta in the final state, whose frequency is given by
(twice) the mass of the boson. In a binary system, the
initial state features an extra body, producing a final
state that includes the GWs emitted during the inspiral.
The existence of a “three-body” initial state leads to an
additional resonant behavior, opening new decaying
channels for the bosonic field. This new configuration
enables us to extract more information about the inter-
mediate state. In particular, the coefficients of finite-size
terms in the worldline theory carry the imprints of new
particles, which, as we will show, can be constrained by
GW precision data. This is analogous to precision tests
in particle physics, where constraints on the coefficients
of higher-dimension operators probe new physics beyond
the Standard Model [71].

A similar analogy, between collider physics and the
imprints of massive particles in cosmological correlation
functions, was drawn in [72]. Taking inspiration from
cosmology, we will refer to this nascent discipline in GW
science as “gravitational collider physics,” which comple-
ments other searches for new light particles, in the lab [52],
astronomy [84], and cosmology [85–87].

A. Outline

The outline of the paper is as follows: In Sec. II, we
review how superradiance generates boson clouds around
rotating BHs, and compute their energy spectrum. In
Sec. III, we study the dynamics of the BH cloud when it
is part of a binary. We show that the gravitational
perturbation produced by a companion induces resonant
mixing between growing and decaying modes of the cloud.
We determine the efficiency of this level mixing to deplete
the energy density in the cloud, as a function of the
parameters of the system. In Sec. IV, we discuss how
the novel features in the dynamics of the BH cloud affect
GW signals, either directly from the cloud or the two-body
system including the companion. We estimate the relevant
finite-size effects, as well as their time dependence induced
by the evolution of the cloud. We then briefly discuss the
novel type of constraints on ultralight bosons in binary
inspirals, through GW searches both from the cloud and the
binary system. We conclude, in Sec. V, with a summary of
our results and a discussion of open problems. We also
provide an outlook on future applications of the gravi-
tational collider. A number of appendices contain additional
material: In Appendix A, we derive various properties of
the BH cloud, including the higher-order (hyper)fine
structure of its energy spectrum and its quadrupole
moment. In Appendix B, we provide an alternative per-
pective on the absence of a dipole induced by the
companion object. Finally, in Appendix C, we review
aspects of the EFT approach for extended objects and
introduce the relevant finite-size corrections.

B. Notation and conventions

Our metric convention is (−, +, +, +). We will use
natural units, $\hbar = c = G = 1$, throughout. The gravita-
tional radius of an object of mass $M$ is $r_g \equiv GM/c^2 = M$.
Properties of the boson cloud will be denoted by the
subscript $c$, e.g., $M_c$ and $r_c$ are the mass and typical radius
of the cloud. We use the subscript $*$ for quantities related to
the companion, e.g., $M_*$ is its mass and $V_*$ is the induced
gravitational potential. Ignoring the small energy loss in
GW emission, the mass of the BH cloud is equal to the
initial BH mass, and we use $M$ for both of them. We denote

\[^{3}\text{In the case of the “cosmological collider,” particles are produced as “resonances,” when their Compton wavelength exceeds the cosmological horizon during inflation. The properties of the new particles in the intermediate state are reflected in the momentum scaling and angular behavior in the soft limits of correlation functions [72–83].}\]
by $J$ the initial BH angular momentum, which is also equal to that of the BH cloud. We define the BH spin parameters $\tilde{a} = a/M = J/M^2$. The mass and spin of the BH at the moment when superradiance saturates are denoted by the subscript $s$, i.e., $M_s$ and $a_s$. We introduce the dimensionless variables $\kappa$ and $\Lambda$ to describe the spin-induced quadrupole and tidal Love number. For a Kerr BH, we have $\kappa = 1$ and $\Lambda = 0$.

II. BOSON CLOUDS AROUND BLACK HOLES

When a BH rotates faster than the angular phase velocity of an incoming wave, it amplifies the energy and angular momentum of the field in its vicinity.\(^4\) This superradiance effect [53–57] is a natural mechanism to create clouds of ultralight bosons around Kerr BHs (see [60] for a review). In this section, we will review this phenomenon and the properties of the boson clouds that it generates.

A. Black hole superradiance

Consider a rotating BH of mass $M$ and spin $J \equiv aM$. A bosonic field with mass $\mu$ and angular frequency $\omega$ experiences a superradiant instability if

$$\frac{\omega}{\mu} < \frac{\Omega_H}{2M r_+}, \quad (2.1)$$

where $m$ is the azimuthal angular momentum of the field, $\Omega_H$ is the angular velocity of the black hole, and $r_+\equiv M + \sqrt{M^2 - a^2}$ is the size of the event horizon of the black hole (see Appendix A). Superradiant growth requires $m > 0$, i.e., modes that corotate with the Kerr black hole.

In principle, superradiance occurs for both massive and massless fields. However, the mass of the bosonic field plays the crucial role of a reflecting barrier, so that the superradiantly amplified field is reflected back onto the black hole and continuously extracts angular momentum from it. This provides a natural realization of the “black-hole bomb” scenario proposed by Teukolsky and Press [89]. The amplitude of the field increases while the black hole spin decreases, until the inequality (2.1) is saturated. The black hole spin at saturation is

$$\frac{a_s}{M_s} = \frac{4m(M_s \omega_s)}{m^2 + 4(M_s \omega_s)^2}, \quad (2.2)$$

where $M_s$ and $\omega_s$ are the relevant quantities evaluated after superradiance has ended. The existence of black holes with spins above this critical value would rule out ultralight scalar fields in the corresponding mass range [70].\(^5\)

Although superradiance is a phenomenon that occurs for bosons of any spin [56,57,90], in the rest of the paper we will focus only on the scalar case. The reason is twofold: from a practical point of view, the Klein-Gordon equation for a massive scalar field in a Kerr background is separable in Boyer-Lindquist coordinates [91,92], which makes the problem analytically tractable (see Appendix A). This is not the case for massive vector\(^6\) and tensor fields. From a theoretical point of view, ultralight scalar fields arise naturally as particle candidates in various scenarios for physics beyond the Standard Model. The Kerr-scalar system is therefore a particularly well-motivated system to be studied.

B. Gravitational atom

The equation of motion of a massive scalar field $\Psi$ around a rotating BH is

$$\left(g^{\mu\nu} \nabla_\mu \nabla_\nu - \mu^2\right)\Psi(t, r) = 0, \quad (2.3)$$

where $g_{\mu\nu}$ is the Kerr metric and $\nabla_\mu$ is the associated covariant derivative. For (real) solutions of (2.3) in the nonrelativistic limit, it is convenient to consider the ansatz\(^5\)

$$\Psi(t, r) = \frac{1}{\sqrt{2\mu}} [\psi(t, r) e^{-i\mu t} + \psi^*(t, r) e^{+i\mu t}], \quad (2.4)$$

where $\psi(t, r)$ is a complex scalar field which varies on a timescale that is much longer than $\mu^{-1}$. Substituting (2.4) into (2.3), and keeping only the leading contributions in $r^{-1}$, the field $\psi(t, r)$ satisfies the Schrödinger equation with a Coulomb-like central potential\(^7\)

$$i \frac{\partial}{\partial t} \psi(t, r) = \left[ -\frac{1}{2\mu} \nabla^2 - \frac{\alpha}{r} \right] \psi(t, r), \quad (2.5)$$

where we have defined the “fine-structure constant”

$$\alpha \equiv \frac{GM_\mu}{\hbar c} \simeq 0.02 \left( \frac{M}{3M_\odot} \right) \left( \frac{\mu}{10^{-12} \text{ eV}} \right). \quad (2.6)$$

Notice that $\alpha$ is the ratio of the gravitational radius of the BH, $r_g \equiv GM/c^2$, and the (reduced) Compton wavelength of the scalar field, $\lambda_c \equiv \hbar/\mu c$.

\(^4\)This is the rotational analog of Cherenkov radiation [88], where spontaneous emission of light occurs when a test particle moves with a speed that is faster than the phase velocity of the medium.

\(^5\)See [93] for recent progress in the separability of the Proca field in Kerr backgrounds.

\(^6\)Complex solutions may be described in terms of two real fields without loss of generality.

\(^7\)Although the Schrödinger equation is first order in time derivatives, the Klein-Gordon equation is second order. This means that generic solutions of (2.3) will still depend on the value of the field and its first time derivative as initial conditions.
The individual terms in (2.10) are analogous to the relativistic correction to the kinetic energy, the fine splitting ($\Delta \epsilon \neq 0$) and the hyperfine splitting ($\Delta m \neq 0$) in the hydrogen atom. These corrections break the degeneracies between modes of the same $n$ but different $\ell$ and $m$ (see Fig. 2).\footnote{Similar computations for the energy spectrum of bounded Dirac fields in Kerr spacetime can be found in \cite{94}.} As we shall see in Sec. III, the presence of an external perturbation can induce transitions between these energy levels.

\section{3. Decay width}

The analogy between the gravitational atom and the hydrogen atom is of course not exact. An obvious distinction is that electrons in the hydrogen atom are fermions, while the cloud is bosonic. This means that each level of the gravitational atom can be occupied by an exponentially large number of scalar quanta \cite{70}, forming a boson \textit{condensate}. Moreover, the states are not perfectly stationary, due to the boundary conditions at the horizon. This is encapsulated in the imaginary part of the eigenfrequency,

\begin{equation}
\omega_{n\ell m} \to \omega_{n\ell m} + i\Gamma_{n\ell m},
\end{equation}

which was omitted in (2.8). In general, the instability rate has to be determined numerically \cite{95}. However, in the limit $\alpha \ll 1$, the decay width can be computed analytically. The result, known as “Detweiler’s approximation,” is given by \cite{96}

\begin{equation}
\Gamma_{n\ell m} = \frac{2r+M}{M} C_{n\ell m}(\alpha)(m\Omega_{\text{H}} - \omega)\alpha^{\ell+5},
\end{equation}

where we defined\footnote{See \cite{97,98} which point out a missing factor of 1/2 in Detweiler’s original computation.}

\begin{equation}
C_{n\ell m}(\alpha) \equiv \frac{2^\ell \alpha^{\ell+1} (n + \ell)!}{n^{\ell+4} (n - \ell - 1)!} \left( \frac{\ell!}{(2\ell)!} \right)^2 \bar{\epsilon} \prod_{j=1}^{\ell} \frac{\bar{\epsilon} \times \left[ j^2 (1 - \bar{\alpha}^2) + (\bar{\alpha} \Delta \epsilon - 2\bar{\alpha} \Delta m)^2 \right]}{(\bar{\alpha} \Delta \epsilon - 2\bar{\alpha} \Delta m)^2},
\end{equation}

with $\bar{\alpha} \equiv \alpha/M$ and $\bar{r} \equiv r_+/M$. Although the analytic result (2.12) was obtained for $\alpha \ll 1$, it was found to give reasonable agreements with numerical results for $\alpha < 0.5$ (see Fig. 28 in \cite{60}).

Due to the strong dependence of $\Gamma$ on $\ell$, the dominant growing modes have $\ell_1 = m_1 = 1$. In principle, this includes an infinite tower of overtones $|n_1\rangle$, with $n_1 \geq 3$, but, in practice, the fastest growing mode is $|211\rangle$. Moreover, at the end of the superradiant growth, we have $\omega_{n_1\ell_1} > \omega_{211} = \Omega_{\text{H}}$, so that the higher overtones decay back into the black hole. Hence, in the rest of the paper, we will consider $|211\rangle$ as the
initial condition. The observationally relevant range of $\alpha$ is then given by

$$0.005 \left( \frac{M}{3 M_\odot} \right)^{1/9} \lesssim \alpha < 0.5,$$

(2.14)

where the lower bound is obtained by demanding the cloud to grow significantly within the age of the Universe, while the upper bound is the maximum value of $\alpha$ for superradiant growth of the $|211|$ mode to occur, assuming an initial BH spin of $\tilde{a} = 0.99$. Smaller values of $\tilde{a}$ result in stronger upper bounds on $\alpha$. For instance, $\tilde{a} = 0.8$ implies $\alpha < 0.25$ [95].

C. Continuous emission

While the energy spectra are identical for both real or complex scalar fields, their stability properties may be different. The energy-momentum tensor of a real scalar field is always time dependent and nonaxisymmetric. As a consequence, clouds made out of real fields are a continuous source of GWs [70]. For complex fields, on the other hand, certain configurations may suppress GW emission. This occurs when the real and imaginary components of the field are in the same eigenstate, but their relative phase is such that they produce time-independent and axisymmetric configurations. Whether this is realized in a specific case depends on the initial conditions of the superradiant growth of the cloud. In any case, since the timescale for this GW emission is much longer than the timescale of the superradiance instability, the former does not inhibit the formation of the cloud [99]. However, depending on the values of $M$ and $\alpha$, the clouds may still deplete on cosmological/astrophysical timescales.

The rate of GW emission for a real scalar cloud is given by [10]

$$\frac{dE_{gw}}{dt} \approx 0.01 \left( \frac{M_c(\alpha)}{M} \right)^2 \alpha^{14},$$

(2.15)

where the numerical coefficient was obtained from a fit to the numerical results of [99,100] and $M_c(\alpha)$ is the mass of the cloud as a function of $\alpha$. For $\alpha \ll 1$, and assuming the initial BH spin to be close to maximal, $\tilde{a} \rightarrow 1$, we can approximate $M_c(\alpha) \approx M\alpha$ [100]. In the absence of additional effects such as accretion disks, conservation of energy implies $M_c = -E_{gw}$, and the cloud evolves as

$$M_c(t) = \frac{M_{c,0}}{1 + t/\tau_c},$$

(2.16)

where $M_{c,0}$ is the initial mass of the cloud and $\tau_c$ the lifetime of the cloud, which is given by

$$\tau_c \approx 10^7 \text{ years} \left( \frac{M}{3 M_\odot} \right) \left( \frac{0.07}{\alpha} \right)^{15}$$

$$\approx 10^9 \text{ years} \left( \frac{M}{10^5 M_\odot} \right) \left( \frac{0.1}{\alpha} \right)^{15}.$$  

(2.17)

These estimates are valid for $\alpha \lesssim 0.1$, while for larger values of $\alpha$, (2.17) can underestimate $\tau_c$ up to 2 orders of magnitude (see footnote 10).

We see that the cloud’s lifetime is extremely sensitive to the value of $\alpha$. For stellar-mass BHs, with masses in the range [3,100] $M_\odot$, we can ignore the GW emission from the cloud for typical binary dynamical timescales of order 10 Myr to 1 Gyr (e.g., [102] and references therein), provided that $\alpha \lesssim 0.07$. On the other hand, for supermassive BHs with $M \gtrsim 10^5 M_\odot$, the cloud survives for more than 1 Gyr if $\alpha \lesssim 0.1$. If the BHs continuously accrete matter, long-lived clouds may be allowed with larger values of $\alpha$ [103]. The associated ranges of $\alpha$ will become relevant when we discuss the evolution of the BH cloud in a binary system, and in particular its GW signatures. While GW emission from the cloud may be ignored for evolution timescales that are smaller than $\tau_c$, other depletion channels may become active in the presence of a companion. We will discuss this next.

III. CLOUDS IN BINARY SYSTEMS

So far, we have studied the BH cloud in isolation, where it can be described by quasistationary states. As we shall see, these configurations are altered when the BH carrying the cloud is part of a binary system (with a companion which may or may not carry a cloud itself). We will find that the gravitational perturbations on the BH cloud yield interesting new phenomena. In particular, we will discover the presence of resonant orbital frequencies for the binary, which lead to new instabilities for the BH cloud. We will discuss the observational consequences of these resonances in Sec. IV.

A. Gravitational perturbations

The presence of a companion, of mass $M_c$, introduces a perturbation to the dynamics of the cloud. We will concentrate on the gravitational potential in the free-falling frame of the cloud, but also briefly discuss the possibility of mass/energy transfer between the cloud and the companion. As long as the binary separation, $R_c$, is larger than the size of the cloud, $R_c > r_c$, the gravitational influence of the companion can be treated in a multipole expansion. Furthermore, we will assume that the correction to the
Kerr metric is a small perturbation. This requires the length scale associated to the gravitational (curvature) perturbation $\sqrt{R_s^2/M_*}$ to be larger than $r_c$, or equivalently $R_s > 2^{-2/3} q^{1/3} \alpha^{2/3} r_c$, where $q \equiv M_*/M$. For $\alpha < 1$, the latter condition is somewhat weaker than the requirement of negligible mass transfer (see below), which reads $R_s > 2 q^{1/3} r_c$. Taking this into account, the regime of validity of our perturbative treatment is bounded by $R_s > R_{\text{pt}} \equiv \max\{r_c, 2 q^{1/3} r_c\}$.

### 1. Free-falling clouds

The perturbed metrics for BHs in external fields have been studied extensively in the literature; see e.g., [104] and references therein. The companion object induces a time-dependent perturbation to the metric, such that $g_{\mu\nu} = g_{\mu\nu}^{(0)} + \delta g_{\mu\nu}$, where $g_{\mu\nu}^{(0)}$ is the unperturbed Kerr spacetime. In general, the metric perturbation $\delta g_{\mu\nu}$ consists of two separate components: (i) a direct contribution from the gravitational potential due to $M_*$, and (ii) the response of the BH cloud to tidal deformations. We will work in the Newtonian limit, which dominates the gravitational perturbation to the cloud due to the companion, and ignore the subleading tidal contributions, as well as the effect of the BH spin on the metric. (We will discuss tidal and spin-induced effects in Sec. IV B, when we study the GW signal from the binary system.)

In comoving Fermi coordinates $(\vec{t}, \vec{x}^i)$, whose origin is located at the center-of-mass\(^\text{12}\) of the unperturbed BH cloud (see Fig. 3), the $\delta g_{00}$ component reads [104,105]

$$
\delta g_{00}(\vec{t}, \vec{r}) = \frac{2}{\epsilon_*} \mathcal{E}_L(\vec{t}, 0) \vec{x}^i
$$

$$
= \sum_{\ell, \epsilon_\ell} \sum_{m, \leq |\ell|} \frac{2}{\epsilon_\ell(\ell_\ell - 1)} \vec{x}^i \mathcal{E}^{(m_\ell)}(\vec{t}, 0) Y_{\ell, m_\ell}(\vec{0}, \vec{r}),
$$

(3.1)

where we have used the notation $L \equiv (i_1 \cdots i_\ell)$. The tensor $\mathcal{E}_L$ is the symmetric trace-free representation of the tidal tensor, defined as $\mathcal{E}_L(\vec{t}, 0) = -\partial_t U_\ell(\vec{t}, 0)/\ell(\ell_\ell - 2)\!, \text{ with \ } U_\ell(\vec{t}) = -M_*/R_\ell(\vec{t})$ the gravitational potential generated by the companion, evaluated along the geodesic of the BH cloud. In the second line of (3.1) we have decomposed the gravitational field into spherical harmonics in polar coordinates $\vec{r} \equiv (\vec{r}, \vec{\theta}, \vec{\phi})$, with coefficients $\mathcal{E}^{(m_\ell)}(\vec{t}, 0)$. The fact that the expansion (3.1) in the free-falling frame begins with two derivatives of the

\(^\text{12}\)In our case, the center-of-mass of the BH cloud coincides with that of the isolated BH, such that the $(\vec{t}, \vec{x}^i)$ coordinates may be directly related to the Boyer-Lindquist coordinates of the Kerr metric.

\(^\text{13}\)The distinction between time coordinates only plays a role at higher post-Newtonian orders, e.g., [104], when the boost factors begin to contribute.
lead to the appearance of extra terms in the potential, for instance a dipole. However, as we show in Appendix B, this fictitious dipole eventually cancels. Finally, the above reasoning can be generalized to incorporate all types of relativistic corrections.

In summary, the Klein-Gordon equation for the BH-cloud system receives corrections due to the presence of a companion which, in the nonrelativistic limit, lead to a perturbed gravitational potential (3.4). The salient feature is the fact that the multipolar interaction starts at the quadrupole level, as determined by the equivalence principle.

2. Mass transfer

The mutual gravitational attraction between the bodies can also induce transfer of mass/energy between the cloud and the companion. This happens when the characteristic Bohr radius $r_c$ exceeds the Lagrange point, $L1$, located in between the two objects of the binary. The equipotential surface with the same gravitational potential as $L1$ is called the Roche lobe (see Fig. 4).

Mass transfer from the cloud to the companion happens when $r_c \gtrsim R_{\text{Roche}}$. Using Eggleton’s fitting formula [107], this can be converted into a critical orbital separation

$$R_{s,\text{cr}} \equiv \left( \frac{0.49q^{-2/3}}{0.6q^{-2/3} + \ln(1 + q^{-1/3})} \right)^{-1} r_c. \quad (3.5)$$

This phenomenon becomes particularly important when $q \gg 1$, since the Roche lobe surrounding the BH carrying the cloud is then relatively small. In this limit, we find

$$R_{s,\text{cr}} \approx 2q^{1/3} r_c. \quad (3.6)$$

In the rest of the paper, we will impose $R_s > R_{s,\text{cr}}$, such that mass transfer can be ignored.

B. Level mixing

The gravitational perturbation $V_s$ in (3.4) induces an overlap between the modes $|\psi_i\rangle \equiv |n_i,\ell_i, m_i\rangle$ and $|\psi_j\rangle \equiv |n_j,\ell_j, m_j\rangle$, through $\langle \psi_j|V_s|\psi_i\rangle$. Substituting (3.4), we get

$$\langle \psi_j|V_s|\psi_i\rangle = -M_\ast \mu \sum_{\ell_i \geq \ell_j} \sum_{|m_i| \leq \ell_i} \frac{4\pi}{2\ell_i + 1} \frac{Y_{\ell_i m_i}^* (\Theta_\ast, \Phi_\ast)}{R_{c,\ast}^{\ell_i + 1}} I_\ell \times I_{\Omega}, \quad (3.7)$$

where

$$I_\ell \equiv \int_0^\infty d\bar{r} r^{2\ell + 1} \bar{R}_{\ell,m_i} (\bar{r}) \bar{R}_{m_i,\ell} (\bar{r}), \quad (3.8)$$

$$I_{\Omega} \equiv \int d\Omega Y_{\ell m_i}(\bar{\Theta}, \bar{\Phi}) Y_{\ell m_i}(\bar{\Theta}, \bar{\Phi}) Y_{\ell m_i}(\bar{\Theta}, \bar{\Phi}). \quad (3.9)$$

The angular integral $I_{\Omega}$ can be expressed in terms of the Wigner 3-j symbol,

$$I_{\Omega} = \sqrt{\frac{(2\ell_j + 1)(2\ell_i + 1)(2\ell_s + 1)}{4\pi}} \begin{pmatrix} \ell_j & \ell_i & \ell_s \\ 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} \ell_j & \ell_i & \ell_s \\ -m_j & m_i & m_s \end{pmatrix}, \quad (3.10)$$

which implies the following selection rules:

1. $-m_j + m_i + m_s = 0$,
2. $|\ell_j - \ell_i| \leq \ell_s \leq \ell_i + \ell_j$,
3. $\ell_i + \ell_j + \ell_s = 2p$, for $p \in Z$.

Recall that the fastest growing mode [211] has $\ell_g = m_g = 1$, while the fastest decaying mode [100] has $\ell_d = m_d = 0$. In this case, the above selection rules would require $\ell_s = \pm m_s = 1$, namely a dipole coupling, which is absent in (3.4). For the quadrupole, the two fastest decaying modes that can couple to $\ell_g = m_g = 1$ are $\ell_d = 1, m_d = -1$ and $\ell_d = 1, m_d = 0$. Since these rules are obtained purely from the angular dependence of the eigenfunctions, they apply equally to the fundamental mode ($n = 2$) and the overtones ($n \geq 3$).

C. Rabi resonances

We now investigate how level mixings (see Fig. 5), induced by the quadrupole $\ell_s = 2$, affect the dynamical evolution of the cloud. For simplicity, we will restrict ourselves to quasi-circular orbits with orbital frequency.

FIG. 5. Illustration of the frequency spectrum for $\ell = 1$ up to $n = 3$ (there is in principle an infinite tower of overtones), for co-rotating and counter-rotating orbits. Solid arrows represent the allowed resonant transitions, while dashed arrows denote perturbative level mixings.
where $M$ is the total mass of the BH cloud, which is equal to the initial BH mass before superradiance, if we neglect the small mass loss due to GW emission. Since $\Phi_\ast$ denotes the azimuthal angle of $M_\ast$ relative to the BH cloud, the orientation $\Phi_\ast = +\Omega t$, with $\Omega > 0$, corresponds to orbits that corotate with the cloud, called co-rotating orbits, while $\Phi_\ast = -\Omega t$ are counter-rotating orbits. This distinction will play a key role in the survival of the cloud throughout the inspiral stage. For notational convenience, we will write $\Phi_\ast = \pm \Omega t$, where the upper sign corresponds to co-rotating orbits and the lower sign denotes counter-rotating orbits.

### 1. Hyperfine mixing

We first consider the mixing of the dominant growing mode $|211\rangle$ with the nearby decaying modes $|210\rangle$ and $|21\rangle$. The problem then involves the dynamics of a coupled three-state system. As we shall see, for orbital motions along the equatorial plane, the mode $|210\rangle$ decouples and the dynamics reduces to a two-state system.

#### 2. Three-state system

For the three-state system, the expectation value of Hamiltonian for the field $\psi$ takes the form

$$H = H_0 + H_1 + H_2,$$

where $H_0$ represents the diagonal terms that contribute equally to all three modes, $H_1$ contains terms that split the degeneracy, and $H_2$ are the off-diagonal terms induced by the quadrupole perturbation $\epsilon_s = 2$. Using the basis $|211\rangle = (1,0,0)^T$, $|210\rangle = (0,1,0)^T$, and $|21\rangle = (0,0,1)^T$, the explicit matrix representations of the different contributions are

$$H_0 = \begin{pmatrix} E & 0 & 0 \\ 0 & E & 0 \\ 0 & 0 & E \end{pmatrix}, \quad H_1 = \begin{pmatrix} e_h + 3\eta f_0(\Theta_\ast) & 0 & 0 \\ 0 & -6\eta f_0(\Theta_\ast) & 0 \\ 0 & 0 & -e_h + 3\eta f_0(\Theta_\ast) \end{pmatrix},$$

$$H_2 = \begin{pmatrix} 3\eta f_1(\Theta_\ast)e^{\pm i\Omega t} & 0 & 3\eta f_2(\Theta_\ast)e^{\mp 2i\Omega t} \\ 0 & 3\eta f_1(\Theta_\ast)e^{\mp 2i\Omega t} & -3\eta f_1(\Theta_\ast)e^{\pm i\Omega t} \\ 3\eta f_2(\Theta_\ast)e^{\mp 2i\Omega t} & -3\eta f_1(\Theta_\ast)e^{\pm i\Omega t} & 0 \end{pmatrix}.$$ (3.13)

For the special case $\Theta_\ast = \pi/2$ (i.e., equatorial motion), this reduces to $f_0 = -1$, $f_1 = 0$, $f_2 = 3$.

It is convenient to work in the interaction picture, where the evolution of the state $|\psi_\ast\rangle$ is given by

$$|\psi_\ast(t)\rangle = e^{iH_\ast t}|\psi_\ast(0)\rangle,$$ (3.18)

where $H_\ast \equiv H_0 + H_1$ is time independent.\(^{14}\) Since $H_\ast$ is diagonal, it does not couple modes of different angular momenta. The eigenstates of $H_\ast$ are therefore equal to the corresponding eigenstates of the nonrelativistic hydrogen atom, with the same $\ell$ and $m$, up to $O(\alpha^2)$ corrections which will not be relevant in our case.

#### 3. Two-state system

We now consider quasicircular orbits along the equator, where $f_1 = 0$. This reduces the problem to a coupled two-state system, which can be solved exactly. Labeling the

\(^{14}\)Strictly speaking, $H_\ast$ has a time dependence due to the decrease of the orbital radius $R_\ast(t)$ in the binary. However, since $R_\ast(t)/R_\ast \ll \Omega$ during the inspiral stage, this time dependence can be treated adiabatically, and neglected at leading order.
growing and decaying modes as $|\psi_g\rangle \equiv |211\rangle$ and $|\psi_d^{(h)}\rangle \equiv |21 - 1\rangle$, the state in the interaction picture is expressed as a linear combination
\[
|\psi_I(t)\rangle = c_g(t)|\psi_g\rangle + c_d^{(h)}(t)|\psi_d^{(h)}\rangle,
\]
where the time-dependent coefficients satisfy
\[
|c_g(t)|^2 + |c_d^{(h)}(t)|^2 = 1. \tag{3.20}
\]
The Schrödinger equation becomes
\[
\frac{id}{dt} \begin{pmatrix} c_g(t) \\ c_d^{(h)}(t) \end{pmatrix} = \begin{pmatrix} 0 & 9\eta e^{-2i(\pm \Omega - \epsilon_g)t} \\ 9\eta e^{2i(\pm \Omega - \epsilon_g)t} & 0 \end{pmatrix} \begin{pmatrix} c_g(t) \\ c_d^{(h)}(t) \end{pmatrix}. \tag{3.21}
\]
Starting the evolution in the growing mode, $c_g(0) = 1$, $c_d^{(h)}(0) = 0$, we find
\[
c_g(t) = \frac{e^{i(\epsilon_g + \Omega)t}}{2\Delta_R^{(h)}} [(\Delta_R^{(h)} + \epsilon_g + \Omega)e^{-i\Delta_R^{(h)}t} + (\Delta_R^{(h)} - \epsilon_g + \Omega)e^{i\Delta_R^{(h)}t}], \tag{3.22}
\]
\[
c_d^{(h)}(t) = \frac{9\eta e^{-i(\epsilon_g + \Omega)t}}{2\Delta_R^{(h)}} [e^{-i\Delta_R^{(h)}t} - e^{i\Delta_R^{(h)}t}], \tag{3.23}
\]
where we introduced
\[
\Delta_R^{(h)} \equiv \sqrt{(9\eta)^2 + (\epsilon_g + \Omega)^2}. \tag{3.24}
\]
Since $\psi$ is a complex scalar field with a global $U(1)$ symmetry, there is an associated conserved Noether current $J^\mu$ (see Appendix A for details). At leading order, $J^0 \approx \psi^\dagger \psi$, which is interpreted as the occupation density of the system, and is analogous to the probability density in quantum mechanics. The occupation density of the decaying mode is proportional to
\[
|c_d^{(h)}(t)|^2 = \left[ 1 - \frac{(\epsilon_g + \Omega)}{\Delta_R^{(h)}} \right]^2 \sin^2 \left[ \int_0^t dt' \Delta_R^{(h)} (t') \right]. \tag{3.25}
\]
Note that $\Delta_R^{(h)}$ is not constant during the inspiral, but increases as the orbit shrinks, so that the phase of the oscillations in (3.25) has been written as an integral over time. We see that $2\Delta_R^{(h)}$ controls the frequency of oscillations between the growing and decaying modes. In analogy with the quantum mechanical problem, we call this the Rabi frequency.

4. Hyperfine resonance

When the orbital frequency $\Omega$ matches the hyperfine splitting $\epsilon_h$, the system experiences a resonance, and starts to oscillate between the growing and decaying modes. Since $\epsilon_h > 0$, the resonance will only take place for co-rotating orbits. We will refer to this effect as the hyperfine resonance. This happens when the binary separation is
\[
R_{\text{res}}^{(h)} = 144^{1/3} \alpha^{-4}(1 + q)^{1/3} \alpha^{-2/3} r_g,
\]
\[
\approx 9^{1/3} \alpha^{-14/3}(1 + 4\alpha^2)^{2/3}(1 + q)^{1/3} \left( \frac{a_0}{a} \right)^{2/3} r_g, \tag{3.26}
\]
where we have used (2.2) for $a_s$, with $M_s \approx M$ and $M \omega_s \approx \alpha$.

5. Bohr mixing

We now consider the possibility that the dominant growing mode $|211\rangle$ mixes also with the decaying modes $|n0\rangle$ and $|n1 - 1\rangle$, with $n \geq 3$. We will take the coupling to the $n = 3$ modes to be a proxy for the mixing with this infinite tower of overtones. Similar conclusions hold for the higher-order overtones ($n \geq 4$). In particular, since $\omega_{n1-1} > \omega_{31-1}$, the only difference is that these resonant mixings occur at orbital separations that are slightly shorter. As before, we restrict ourselves to motion in the equatorial plane, $\Theta_\star = \pi/2$, so that the $|310\rangle$ mode decouples.

6. Three-level system

Labeling $|\psi_g\rangle \equiv |211\rangle$ and $|\psi_d^{(h)}\rangle \equiv |21 - 1\rangle$ as before, but adding the extra decaying mode $|\psi_d^{(b)}\rangle \equiv |31 - 1\rangle$, the state of the system in the interaction picture is now expressed as the following linear combination
\[
|\psi_I(t)\rangle = c_g(t)|\psi_g\rangle + c_d^{(h)}(t)|\psi_d^{(h)}\rangle + c_d^{(b)}(t)|\psi_d^{(b)}\rangle, \tag{3.27}
\]
with the normalization condition given by
\[
|c_g|^2 + |c_d^{(h)}|^2 + |c_d^{(b)}|^2 = 1. \tag{3.28}
\]
The Schrödinger equation implies
In general, there is no analytic solution for the three-level system (3.29). However, since we have \( \epsilon_b / |\epsilon_b| \approx \alpha^2 \ll 1 \), there is a clear hierarchy of scales in the evolution equation, and the system will probe different decaying modes at different times. The strength of the coupling—away from resonances—is determined by the ratio of the size of the gravitational perturbation to the energy split. During the early stages of the inspiral, when \( \Omega \ll \epsilon_b \), the mixing with the hyperfine state dominates. The solutions for \( c_d(t) \) and \( c_d^{(b)}(t) \) are thus identical to those given in (3.25), with an error in the normalization condition of order \( (\epsilon_b / |\epsilon_b|)^2 \). As we have seen, for co-rotating orbits, the binary experiences a resonance when \( \Omega = \epsilon_b \), while counter-rotating orbits continue smoothly through this region. When the orbital frequency approaches the scale of the Bohr splitting, \( \Omega = |\epsilon_b| \), the overtone \( |31 - 1| \) gets excited; see Fig. 6.

### 7. Bohr resonance

Near \( \Omega \approx |\epsilon_b| \), the phase of the hyperfine mixing term in (3.29) oscillates rapidly over a time of order \( \eta^{-1} \). In this region, we have \( |c_d^{(b)}|^2 \sim (\eta / \Omega)^2 \approx (\eta / \epsilon_b)^2 \ll 1 \), so that we can ignore the small correction due to the \( |21 - 1| \) state. The dynamical evolution, once again, reduces to a two-level system, this time describing the mixing between the states \([211]\) and \([31 - 1]\). Repeating the analysis of §III C 1, we find

\[
|c_d^{(b)}(t)|^2 = \left[ 1 - \frac{(\epsilon_b \mp \Omega)}{\Delta_k^{(b)}} \right]^2 \sin^2 \int_{t_0}^t dt' \Delta_k^{(b)}(t'),
\]

(3.31)

where the modified Rabi frequency is

\[
\Delta_k^{(b)} \equiv \sqrt{(7.6 \eta)^2 + (\epsilon_b \mp \Omega)^2}.
\]

(3.32)

Similar to the hyperfine case, the system undergoes a resonance when the orbital frequency matches with the energy split with the overtones. We refer to this effect as the Bohr resonance. A crucial difference to the case discussed in Sec. III C 1 is that, while for the hyperfine resonance we have \( \omega_{211 - 1} < \omega_{211} \), the Bohr resonance is for an excited state \( \omega_{31 - 1} > \omega_{211} \); see Fig. 5. This implies that the Bohr resonance can only occur for counter-rotating orbits, at a binary separation given by

\[
R_{res}^{(b)} = \left( \frac{144}{5} \right)^{2/3} \alpha^{-2} (1 + q)^{1/3} r_\eta.
\]

(3.33)

Note the \( \alpha^{-2} \) scaling, whereas the corresponding result for the hyperfine resonances scales as \( \alpha^{-4} \). For small \( \alpha \), the Bohr resonance therefore occurs at a much smaller separation than the hyperfine case (see Fig. 6). In terms of the Bohr radius of the cloud, we can write \( R_{res}^{(b)} = (18/5)^{2/3} (1 + q)^{1/3} r_c \). For \( q < 1 \), the Bohr resonance thus occurs at the boundary of the regime of validity of the multipole expansion, \( R_s > r_c \). On the other hand, for \( q \gg 1 \), the multipole expansion remains accurate, while the stronger constraint typically comes from the requirement for the validity of the perturbative treatment of the gravitational interaction with negligible mass transfer, corresponding to \( R_s > R_{pl} = 2q^{1/3} r_c \).

### D. Cloud depletion

So far, we have taken the growing and decaying modes to be stationary, while in reality both have nonzero decay widths; see (2.12). Although \( \Gamma_{211} \) vanishes when the superradiant growth stops, the rates \( \Gamma_{m/m'} \), with \( m < 0 \), remain finite throughout the evolution of the BH cloud. Since the gravitational perturbation of the companion can induce transitions to the decaying modes, this opens up a new depletion channel for the cloud. In this section, we will explore this effect in more detail, both for co-rotating and counter-rotating orbits.

As we shall see, it is convenient to express the time evolution of the system in terms of the orbital frequency, \( \Omega \), of the binary. For quasicircular orbits, the relation between \( t \) and \( \Omega \) due to the leading quadrupolar GW emission is [108]

\[
t(\Omega) = \tau_0 \left[ 1 - \left( \frac{\Omega_0}{\Omega} \right)^{8/3} \right],
\]

(3.34)

where \( \tau_0 \) is the time to merger for an initial orbital frequency \( \Omega_0 \).

\[044001-11\]
\[
\tau_0 = \frac{5}{256} \nu \left( \frac{1}{M_{\text{tot}} \Omega_0} \right)^{8/3},
\]

(3.35)

with \( M_{\text{tot}} \equiv M + M_\star \), and \( \nu \equiv M \mathcal{M}_\star / M_{\text{tot}} \). For simplicity, we will assume that we can neglect the GW emission from the cloud throughout the coalescence of the binary. For the case of real scalar fields, this requires the time to merger, \( \tau_0 \), to be shorter than the lifetime of the cloud, \( \tau_c \), introduced in Sec. II C. This constraint also applies to complex fields, except for the special case where the resulting configuration is time-independent and axisymmetric. Imposing the condition \( \tau_0 < \tau_c \) leads to

\[
\frac{\tau_0}{\tau_c} \approx \frac{10^{-21}}{(M \Omega_0)^{8/3}} \left( \frac{1 + q}{q} \right)^{1/3} \left( \frac{\alpha}{0.07} \right)^{15} < 1,
\]

(3.36)

which can be translated into a bound on the initial orbital frequency

\[
\Omega_0 \gtrsim 0.9 \text{ mHz} \left( \frac{3 M_\odot}{M} \right) \left( \frac{1 + q}{q} \right)^{1/8} \left( \frac{\alpha}{0.07} \right)^{45/8} \equiv \Omega_{\nu}. \tag{3.37}
\]

In the following, we will assume \( \Omega > \Omega_{\nu} \) during the entirety of the binary inspiral. In principle, we can also study binaries which form at smaller initial orbital frequencies, so that GW emission depletes the cloud before merger. We will return to this point in Sec. IV C.

1. Level mixings

We take \( t_0 \) to be the time at which the superradiance has saturated. To estimate the amount of depletion between times \( t_0 \) and \( t \), we introduce the following quantity

\[
\mathcal{A}(t, t_0) = \sum_{n, \ell, m} \sum_{m \geq 0} |\Gamma_{n \ell m}| \int_{t_0}^t \text{d}r |c_{n \ell m}(r')|^2, \tag{3.38}
\]

where \( c_{n \ell m}(r) \) is the overlap of the state \(|\psi\rangle\) with the decaying modes \(|n \ell m\rangle\), with \( m \leq 0 \). The “depletion estimator” (3.38) is thus the ratio of the integrated time spent by the system in the decaying modes to the decay timescale \(|\Gamma_{n \ell m}|^{-1}\), weighted by the occupation density of each state. Physical quantities of the cloud, such as its mass and angular momentum, would decay as \( e^{-2 \mathcal{A}(t, t_0)} \), where the factor of 2 in the exponential arises because the energy-momentum tensor depends quadratically on the field. In the limit where mixing with the decaying states is negligible, \( \mathcal{A} \to 0 \) and the cloud remains stable. On the other hand, significant depletion occurs when \( \mathcal{A} \) becomes of order one.

For \( \alpha < 1 \), the decay rate (2.12) is suppressed for higher \( \ell \)'s, so that the sum in (3.38) will be dominated by the modes with \( \ell = 1 \) and \( m = -1 \) (the \( m = 0 \) mode decouples for orbits along the equatorial plane). Restricting further to the dominant decaying channels, with \( n_h = 2 \) and \( n_b = 3\),

\[3.15\] and using (3.34) to convert the integral over time in (3.38) into an integral over orbital frequency, we have

\[
\mathcal{A}(\Omega, \Omega_0) \approx \sum_{i = (h, b)} |\Gamma_{i}^{0}| \frac{5}{96 \nu} \left( \frac{1}{M \Omega_0^{2/3}} \right) \int_{\Omega_0}^{\Omega} \text{d}\Omega \Omega^{-11/3} |\langle \psi_i^{0} | \Omega \rangle|^2.
\]

(3.39)

Using Detweiler’s approximation (2.12), the relevant decay rates are

\[
|\Gamma_{i}^{0}| = \frac{B_i}{24 \nu^{3/4} \left( \frac{1 - 4\alpha^2}{1 + 4\alpha^2} \right)^2} \left( \frac{2}{1 + 4\alpha^2} \right)^{7/4} \right), \tag{3.40}
\]

where \( B_{(h)} \equiv 1 \) and \( B_{(b)} \equiv 256/729 \), and we used \( \tilde{a} = \tilde{a}_s \), at saturation. In the limit of small \( \alpha \), the corresponding decay time is

\[
\tau_{d}^{(i)} \approx 1 \mathcal{B}_{(i)} \text{ years} \left( \frac{M}{3 M_\odot} \right) \left( \frac{0.07}{\alpha} \right)^{10} \]

\[
\approx 0.9 \mathcal{B}_{(i)}^{-1} \text{ years} \left( \frac{M}{10^{-5} M_\odot} \right) \left( \frac{0.2}{\alpha} \right)^{10}. \tag{3.41}
\]

This timescale is very sensitive to \( \alpha \), but can be shorter than the duration of GW observations, so that the decay can, in principle, occur in observational bands. In the following, we will use the estimator (3.39) to analyze the stability of the BH cloud against level mixing, during the inspiral phase of the binary dynamics.

2. Co-rotating orbits

We first consider the evolution of the cloud for co-rotating orbits, which is dominated by the existence of the hyperfine resonance. For each combination of \( \alpha \) and \( q \), we chose the initial condition \( \Omega_0 = \Omega_{\nu} \), see (3.37). Figure 7 illustrates the total amount of depletion due to level mixing between \( \Omega_0 \) and the breakdown of perturbation theory at \( \Omega_{\nu} \). As shown in the top panel of Fig. 7, the parameter space divides into three distinct regions:

(i) Region I (red). The cloud depletes dramatically during the resonance.

(ii) Region II (blue). The cloud depletes only slightly during the resonance.

(iii) Region III (gray). The cloud remains stable.

15Ignoring the \( n \geq 4 \) overtones amounts to underestimating the total contribution from \( n \geq 3 \) by approximately a factor of 2, mostly from the evolution of the cloud prior to the resonances. This arises from the suppression of \( \Gamma_{n \ell m} \) for increasing \( n \); see (2.13).

16In the numerical integration of (3.39), we ignore the oscillatory terms in (3.25) and (3.31), which overestimates \( \mathcal{A} \) by roughly a factor of 2.

044001-12
only spends a short amount of time within the resonance epoch, relative to the decay time \( t_d^{(h)} \). \(^{17}\)

The results in the other regions can be understood qualitatively as follows. The larger the value of \( q \), for fixed \( M \) and \( \alpha \), the faster the binary transits through the inspiral stage, due to the efficiency of GW emission from the binary. At the same time, in the regime of perturbative level mixing, \( |c_d|^2 \) is proportional to \( q \). Since \( \mathcal{A} \) is roughly proportional to both the time spent in the inspiral and the strength of the coupling to the decaying modes, the two effects compete with each other. By continuity, \( \mathcal{A} \) peaks at intermediate values of \( q \) in region II. As shown for the case \( \{ \alpha = 0.1, q = 1 \} \) (red curve), the depletion can be significant, yet it extends over a larger time than in the case of the resonant decay. In regions III\(_a\) and III\(_b\), the rapid decay of the orbit due to GW emission (\( q \gg 1 \)) and the suppression of the perturbative mixing (\( q \ll 1 \)) dominate, respectively. As a consequence, the cloud hardly depletes during the inspiral, as shown for \( \{ \alpha = 0.03, q = 10^4 \} \) (yellow curve) and \( \{ \alpha = 0.1, q = 10^{-2} \} \) (blue curve). However, this does not necessarily imply that the cloud is stable beyond \( \Omega > \Omega_{pt} \), since our perturbative treatment of the problem starts breaking down. Only a full numerical simulation could then inform us about the fate of the cloud towards the merger stage.

### 3. Counter-rotating orbits

Next, we consider the case of counter-rotating orbits, for which the Bohr resonance can be excited. As before, we chose the initial condition \( \Omega_0 = \Omega_{cr} \). Figure 8 illustrates the total amount of depletion between \( \Omega_0 \) and \( \Omega_{pt} \). The labeling of the different regions is in accordance to Fig. 7. Once again, the basic features can be understood as follows.

The ratio of the initial orbital frequency and the frequency of the Bohr resonance is given by

\[
\frac{\Omega_0}{\epsilon_b} \simeq 0.003 \left( \frac{\alpha}{0.1} \right)^{21/8} \frac{(1 + q)^{1/8}}{q^{7/8}}. \tag{3.43}
\]

Except for very large values of \( \alpha \) and very small \( q \), this implies \( \Omega_0 < \epsilon_b \), so that essentially all the orbits shown in Fig. 8 experience the Bohr resonance at some point during the inspiral. However, since the Bohr resonance occurs at much smaller orbital separations than the hyperfine resonance, the binary moves through it much faster. As a consequence, the amount of depletion is significantly less dramatic than for co-rotating orbits experiencing the

\(^{17}\)For fixed mass of the BH cloud, \( M \), large values of \( q \) increase the rate of GW emission and hence the shrinking of the orbit, while larger values of \( \alpha \) push the resonance towards shorter distances. Notice that, while the GW emission is invariant under \( q \rightarrow 1/q \), the timescale associated with the decaying mode, as well as the separation at resonance, is controlled by \( M \), which breaks this symmetry.
particles forming the cloud, and whose GW signatures are within reach of current and planned GW observatories. Our treatment will be rather schematic and a more detailed investigation will be the subject of future work.

A. Signal from the cloud

As we have seen in Sec. II C, a real scalar cloud emits continuous, monochromatic GWs [70,99]. The frequency of these GWs is determined by the mass of the scalar field,

$$f_c \approx \frac{\mu}{\pi} = 484 \text{ Hz} \left( \frac{\mu}{10^{-12} \text{ eV}} \right),$$

and their rms strain amplitude is [99,103]

$$h_c \approx 2 \times 10^{-26} \left( \frac{M}{3 M_\odot} \right) \left( \frac{M_c(\alpha)/M}{0.1} \right) \left( \frac{\alpha}{0.07} \right) \left( \frac{9 \text{ kpc}}{d} \right),$$

where $d$ is the (Euclidean) distance of the source. Since the signal is emitted continuously, its detectability depends on the total observation time $T_{\text{obs}}$. The signal-to-noise ratio (SNR) of the cloud signal is

$$\text{SNR} = h_c \langle F_+^2 \rangle^{1/2} \frac{T_{\text{obs}}^{1/2}}{\sqrt{S_n(f_c)}},$$

where $S_n(f_c)$ is the (one-sided) noise spectral density at the frequency $f_c$ and $\langle F_+^2 \rangle = \langle F_\times^2 \rangle$ is the angular average of the square of the detector pattern functions $F_+,\times$ for each GW polarization. Using (4.2), we get

$$\text{SNR} \approx 13 \langle F_+^2 \rangle^{1/2} \left( \frac{T_{\text{obs}}}{1 \text{ yr}} \right)^{1/2} \left( \frac{10^{-23} \text{ Hz}^{-1}}{\sqrt{S_n(f_c)}} \right) \left( \frac{M}{3 M_\odot} \right) \times \left( \frac{M_c/M}{0.1} \right) \left( \frac{\alpha}{0.07} \right) \left( \frac{10 \text{ kpc}}{d} \right),$$

which is strongly dependent on $\alpha$. We see that BH-cloud systems with $\alpha < 0.07$ and $M < 10^6 M_\odot$ may be detectable only within our own galaxy, while those with larger values of $\alpha$, and bigger masses, may be observed at extragalactic distances. Further discussion and detailed forecasts can be found in [100,109–111].

I. Resonant extinction/attenuation

When the BH cloud is part of a binary system, the cloud may deplete over time, according to

$\Delta$ The ratio (4.5) need not be a smooth exponential decay, but may contain periodic features; cf. (3.25) and (3.31). The depletion process switches on and off as the binary transits through the resonance. This effect manifests itself most prominently during the Bohr resonance.
where \( \mathcal{A} \) is the depletion estimator introduced in (3.38).

This decay is most prominent for co-rotating orbits, since the hyperfine resonance occurs at large separations, so that the binary spends a significant amount of time near the resonance orbit. For counter-rotating orbits, the Bohr resonance attenuates the signal, but does not completely terminate it. In either case, the extinction/attenuation of the GW signal, \( \Delta f_c \), is experimentally (and computationally) challenging, this Doppler effect would be very convincing evidence that the BH cloud is a part of a binary system. At the same time, it will open a new avenue to study these quantities would require incorporating the backreaction of the cloud to the spacetime geometry, which is beyond the scope of this work.\(^{20}\) Instead, we will assume that the cloud gives rise to quasistationary spacetimes, with typical relative velocity of the binary and \( \kappa \equiv J/M^2 \) the dimensionless spin parameter.

As long as the mass of the boson cloud constitutes a sizable fraction of the initial BH mass, the metric of the (isolated) BH cloud would depart from that of the Kerr background. These departures include spin-induced multipole moments, which are not uniquely determined by the mass and spin of the cloud. A rigorous computation of these quantities would require incorporating the backreaction on the spacetime geometry, which is beyond the scope of this work.\(^{20}\) Instead, we will assume that the cloud dominates the contribution to the multipole moments of the BH cloud, and estimate \( \kappa \) by comparing the cloud’s mass quadrupole, \( Q_c \), to \( J^2/M \), where \( M \) and \( J \) are the total mass and angular momentum of BH-cloud system. We choose to normalize \( \kappa \) with \( (M,J)\) of the BH cloud, instead of \( (M_c,J_c) \) of the cloud itself, because (i) the quantities \( (M,J) \) can be directly measured through GW observations, and (ii) they are conserved throughout the evolution (up to small losses due to GW emission). The parameters \( (M,J) \) therefore also coincide with the mass and angular momentum of the initial BH prior to the formation of the cloud.

From the stress-energy tensor of the scalar field, we find for the \([211]\) state (see Appendix A)

\[
\kappa(\alpha) \gtrsim Q_c(\alpha)/M^3 \sim 10^3 \left(\frac{M_c(\alpha)/M}{0.1}\right)^4 \left(\frac{\alpha}{0.1}\right)^4 ,
\]

where we have imposed the weak cosmic censorship condition, \( J \leq M^2 \).\(^{22}\) The effect on the GW phase then scales as

\[
\frac{h_c(t)}{h_c(t_0)} = \frac{M_c(t)}{M_c(t_0)} \sim e^{-2A(t-t_0)},
\]

where \( A \) is the depletion estimator introduced in (3.38).
where we used $\chi \approx 1$, and assumed that the initial BH is rapidly rotating (which is required for the cloud to form in the first place). Notice that, in the regime of validity of the perturbative expansion, the relative velocity satisfies $v \lesssim \alpha/2$.\footnote{For $q \lesssim 1$, the virial theorem implies $v \approx (\alpha/2) \sqrt{r_c/R}$, and hence $v \lesssim \alpha/2$ in the regime of validity of the multipole expansion, $r_c \ll R$.} For $v > \alpha/2$, the companion experiences a smaller amount of the cloud according to Gauss’s law. Even though larger relative velocities may seem favorable, the reduction in the effective mass of the cloud will dominate, leading to negligible finite-size effects once the companion enters the BH-cloud system.

The existence of resonances in the orbital dynamics can lead to an abrupt drop, or significant change, in the mass of the cloud $M_c$, with a corresponding frequency-dependent variation of the GW signal. As the binary scans through the Rabi frequency, we have

$$\kappa(t) - 1 \sim M_c(t) \sim e^{-2A(t,t_0)},$$

(4.10)

This time dependence constitutes a distinctive signature of the existence of a BH cloud in a binary system and may leave a measurable imprint in the waveforms. We will return to this in Sec. IV C.

2. Tidal deformations

In addition to the permanent multipole moments, a compact object may also acquire induced multipoles in the presence of a gravitational perturbation. It is an interesting (and somewhat puzzling) fact that for BHs the tidal response to an external field vanishes in four-dimensional Einstein gravity.\footnote{So far, this has been shown to hold for nonrotating BHs. However, it is also expected to be the case for Kerr BHs. See e.g., \cite{124} and references therein for progress along this direction.} This is often parameterized in terms of the tidal Love numbers,\footnote{These Love numbers are the analog of susceptibilities in electrodynamics, which describe the response of a distribution of charge to an applied electric (or magnetic) field.} which are zero for black holes \cite{7,23,24,49,50}, but may have sizable values for neutron stars \cite{23,24} and other more exotic objects (like boson stars) \cite{30,31,125}. While the Love numbers first enter in the GW phase at 5PN order \cite{11,25,26,126}, the lack of standard model background in Einstein’s theory offers a powerful opportunity to probe the dynamics of vacuum spacetimes in general relativity, through precision GW data. See \cite{7} for more on this issue.

Since the cloud is much less compact than an isolated BH, but carries a significant fraction of the mass and angular momentum of the system, we can, in principle, have large Love numbers for the BH-cloud system. Although a detailed computation is beyond the scope of this work, on dimensional grounds we expect the “tidal deformability” parameter, $\Lambda$, to scale as (see Appendix C)

$$\Lambda(\alpha) \sim \left( \frac{M_c(\alpha)/M}{\alpha} \right)^4 \sim 10^{\frac{1}{2}} \left( \frac{M_c(\alpha)/M}{0.1} \right)^8 \left( \frac{\alpha}{\alpha} \right)^8.$$  

(4.11)

The parameter $\Lambda$ enters in the phase of the waveform at 5PN order, and its imprint on the signal from the binary therefore scales as

$$\Lambda(\alpha) v^{10} \sim 10^{-6} \left( \frac{M_c(\alpha)/M}{0.1} \right)^2 \left( \frac{v}{\alpha} \right)^{10},$$

(4.12)

where $v \lesssim \alpha/2$ in the regime of validity of the multipole expansion.

As for the spin-induced quadrupole, the presence of the binary companion can induce a characteristic time dependence of the Love numbers

$$\Lambda(t) \sim M_c(t) \sim e^{-2A(t,t_0)},$$

(4.13)

most notably for Bohr resonances, which occur at shorter binary separations.

As the companion progresses into the cloud and approaches the merger with the central BH, both the effects of the spin-induced quadrupole and Love numbers will be dominated by the Kerr solution. In particular, the contribution from the Love number will become negligible. By comparing the values obtained before and after the resonance, the variation of the spin-induced quadrupole and Love number(s) can turn into a key signature of the presence of a cloud in a binary system.

C. Probing ultralight scalars

In this section, we study the new ways in which ultralight scalars can be probed with binary inspirals. We will concentrate on paradigmatic GW sources for which both the hyperfine and Bohr resonances can play a major role, accessible to present and planned GW detectors.

To be able to probe the resonances, the lifetime of the cloud has to be sufficiently long, and the resonance sharp enough, to be detected within the operational time window of GW observatories:

1. Cloud lifetime. Survival of the cloud until the resonance typically requires $t_c \gtrsim 10$ Myr for stellar-mass binaries, using generic lower bounds on their merger times \cite{102}. For supermassive BHs, we conservatively take $t_c \gtrsim 1$ Gyr to account for the large uncertainties in the formation channels.
and compared to the frequency bands of ground-based Fig. 9, for physically motivated values of $M \lesssim \alpha$ out of complex fields can be observed for larger values on $M$ for larger $M$. Accretion may alleviate the stability issue [103].

In general, we can relax the condition (3.36) for the cloud’s stability by requiring the lifetime to be sufficiently long to reach the resonance after formation, e.g., $\tau_c \gtrsim 1$ Myr, but not necessarily to last until merger.

The frequency of the signal from the cloud is shown in Fig. 9, for physically motivated ranges of $\alpha$ and $M$. The vertical dotted lines illustrate the typical observational bands of various GW observatories.

1. Continuous-wave searches

The frequency of the signal from the cloud is shown in Fig. 9, for physically motivated values of $\alpha$ and $M$; cf. (4.1), and compared to the frequency bands of ground-based [127–131] and space-based [132–134] detectors, as well as pulsar timing arrays (PTAs) [135–140]. We see that ground-based experiments probe $M \lesssim 3 \times 10^3 M_\odot$, while LISA would give us access to $M \gtrsim 3 \times 10^5 M_\odot$. For LIGO, the majority of resolvable events with significant SNR are produced by galactic sources; cf. (4.4). The present binary event rates inferred from LIGO [141] indicate that the most promising sources in the galaxy are far from merging. Indeed, population synthesis codes [142–144] suggest that there are presently of the order of $10^5$ to $10^6$ galactic BH-BH and BH-NS binaries of comparable masses with orbital frequencies distributed in the mHz region. Provided a fraction of these binary systems transit through a resonant epoch, a sharp decay in the GW signal from the cloud may be observed. This is particularly suitable for the hyperfine decay, for which the orbital resonances for stellar-mass BH binaries fall naturally in the mHz region; see (4.14) below. However, due to the finite lifetime of the clouds, detecting this effect in the galaxy requires the clouds to be formed relatively recently, while the majority of these binaries may have formed much earlier [144]. Despite these inherent uncertainties in the binary formation mechanisms, there are fortunately a large number of galactic binaries, each sampling a distribution of initial conditions. This constitutes an ensemble which allows for a “scanning” of the orbital frequencies for the resonance, and the corresponding depletion of the monochromatic GW signal.

The scanning for both the hyperfine and Bohr resonances is more promising with LISA, which would also be sensitive to extra-galactic sources. In particular, a sharp decay may be observed for clouds surrounding supermassive BHs with a small companion, in what are known as extreme-mass-ratio inspirals (EMRIs). Similarly to the tuning required for the binary signals from EMRIs to fall into the LISA band [145], a companion may excite orbital resonances, inducing a sharp depletion on the GW signal from a supermassive BH cloud. The formation mechanism and population of intermediate-mass and supermassive BHs, as well as EMRIs, is less well understood. However, since LISA will be able to probe sources of extra-galactic origin, the search volume increases considerably, allowing for the detection of GW signals at higher redshifts.

As a concrete example, consider a real scalar field with $\mu \simeq 10^{-17}$ eV, forming a cloud surrounding a supermassive BH with mass $M \simeq 5 \times 10^6 M_\odot$. For simplicity, let us assume that the BH accretes matter, so that the issue of the lifetime of the cloud may be ignored. In this scenario we have $\alpha \simeq 0.37$, and a resonant decay time of about a few months. The continuous monochromatic signal from the cloud occurs around $f_c \simeq 5$ mHz, the most sensitive part of the LISA band. For binaries that transit through the hyperfine resonance, the complete termination of the signal occurs for $q \lesssim 10^3$. The decay can arise either from significant perturbative depletion in the early stages or, provided that the initial orbital frequency $\Omega_0$ lies close to...
\(\epsilon_h\), the cloud may survive the perturbative decay and experience a sharp depletion within a few months.\(^{25}\) On the other hand, significant attenuation of the signal due to Bohr resonances requires \(q < 1\); cf. Fig. 8. These type of binary systems, scanning over the space of orbital resonances, are expected to be abundant in nature, which motivates comprehensive continuous-wave searches with space-based interferometers.

### 2. Binary searches

Precision finite-size measurements in binary searches are a new probe of boson clouds around BHs. If the resonance falls within the frequency window of future observatories, it becomes a distinctive feature of the GW signal. To see when that may be the case, it is instructive to translate the orbital frequency at which the resonance transitions occur, \(\Omega_{\text{res}}\), to the GW frequency, \(f_{\text{res}}\), emitted from the binary during the inspiral:

\[
f^{(i)}_{\text{res}} = \frac{\Omega^{(i)}_{\text{res}}}{\pi} = \left| \epsilon_i \right| \begin{cases} \frac{7.2 \text{ mHz}}{1 + 4\alpha^2} \left( \frac{\alpha}{0.1} \right)^7 \left( \frac{3M_\odot}{M} \right) \text{ hyperfine,} \\ 0.75 \text{ Hz} \left( \frac{\alpha}{0.1} \right)^3 \left( \frac{3M_\odot}{M} \right) \text{ Bohr.} \end{cases}
\]

(4.14)

This frequency is shown in Fig. 10, for physically motivated ranges of \(\alpha\) and \(M\). We see that, for most cases, both the hyperfine and Bohr frequencies are too low to be observed by ground-based detectors. The best prospects for detecting the resonance feature in future binary searches is therefore through space-based detectors, such as LISA.

The targeted values of \(\mu\) (and \(\alpha\)) for real scalars is restricted mostly by the stability of the cloud. For axisymmetric configurations with complex fields, GW emission from the cloud is not present and therefore binary searches are perhaps the only way in which these bosons could be potentially observable.\(^{26}\) As we see in Fig. 10, observing the resonant decay is most promising in the deci-Hertz and LISA frequency bands. For simplicity, let us focus on sources that are observable by LISA, which is sensitive to compact binaries with large masses \(M \in [10^5, 10^7] M_\odot\), although in principle it can also detect stellar-mass BHs.

\[\text{FIG. 10. The resonance frequency } f_{\text{res}} \text{ for physically motivated ranges of } \alpha \text{ and } M. \text{ The vertical dotted lines illustrate the typical observational bands of various GW observatories.}\]

\[\text{[146]. Binary searches can then broadly be divided into two classes: (i) binaries of comparable masses, } q \in [10^{-2}, 10^{2}] \text{, and (ii) EMRIs, with } q \in [10^{-6}, 10^{-2}] \cup [10^{2}, 10^{6}] \text{.}\]

In the case of comparable-mass binaries, finite-size effects manifest themselves at high PN orders, and therefore become accessible when the relative velocity of the binary becomes sizable. Once the companion enters the cloud, the finite-size effects will be dominated by the BH solution, for instance the spin-induced quadrupole will be given by \(\kappa \approx 1\) and the Love numbers will vanish. To detect the time-dependent effects in (4.10) and (4.13), the imprints in the GW phase have to be observable near the resonance. For co-rotating orbits, where the hyperfine resonance occurs at larger distances, measuring the depletion of the finite-size effects is challenging. Nevertheless, it is also possible that the cloud forms in a binary systems with \(\Omega_0 > \epsilon_h\), therefore missing the hyperfine resonance. While level mixings would still deplete the cloud, it is possible that it survives until later times; e.g., see Fig. 7 for \(\alpha \approx 0.1\) and \(q \approx 10^{-2}\). In this scenario, resonant depletion would not occur and finite-size effects may still be observed at later stages of the inspiral.

Detecting the resonant decay is more promising for counter-rotating orbits, where the resonance generically occurs at shorter separations. In that case, the relative velocity near the resonance is given by \(v_{(b)} \approx 0.65(\alpha/2)\), which can be significant. For example, for \(\alpha \approx 0.1\) and \(M_\odot \approx 0.1 M_\odot\), the finite-size effects near the Bohr resonance are of order,

\[
\kappa(\alpha) v_{(b)}^4 \chi^2 \sim 10^{-3},
\]

(4.15)

\[
\Lambda(\alpha) v_{(b)}^10^4 \sim 10^{-8}.
\]

(4.16)

The resonant depletion of these quantities can be as large as \(\approx 20\% - 30\%\) (see Fig. 8 for \(\alpha \approx 0.1\) and \(q \approx 10^{-2}\)). Notice that, while small, both types of finite-size effects are larger than the usual 2PN and 5PN terms, respectively.

\[\text{[31]. However, unlike the BH cloud, boson stars do not experience resonant depletion, which therefore is a unique feature of boson condensates surrounding spinning BHs.}\]
In particular, near the resonance, tidal effects are comparable to typical 3PN contributions and the corrections to the GW phase appear to be within reach of future GW observatories such as LISA [30].

In the case of EMRIs, the evolution of the companion in the background of a supermassive BH can be obtained through perturbative self-force computations [147,148]. The tidal deformability of the object is typically not taken into account, since it is highly suppressed when $M_{\text{c}} \ll M$. Nonetheless, systems with $q \ll 1$ still offer an excellent probe of deviations from $\kappa = 1$ for the background spacetime. It has been estimated that a variation in the spin-induced quadrupole of order $\Delta \kappa/\kappa \approx 10^{-2}$ may be detectable by LISA [149]. For example, counter-rotating binaries with $\alpha \approx 0.1–0.4$ and $q \lesssim 10^{-2}$ would behave in a similar way as the blue curves in Fig. 8. This strongly motivates searches for large spin-induced quadrupole moments in EMRIs and their resonant depletion.

V. CONCLUSIONS AND OUTLOOK

Astrophysical BHs exist with a wide range of masses, from a few to billions of solar masses. When these BHs are rapidly rotating, they can produce condensates of ultralight particles in the range $[10^{-20}, 10^{-10}]$ eV, which includes well-motivated candidates for physics beyond the Standard Model, such as the QCD axion [61–63], axionlike particles in string theory [67–69], and interesting new possibilities for dark matter [64]. Extremely ambitious observational programs are under way to search for these light particles in the lab [52], in astrophysics [84] and in cosmology [85–87]. As impressive as these efforts are, they are blind to particles that couple to ordinary matter only through gravitational interactions. Such extremely weakly coupled particles would nevertheless be produced by BH superradiance. The resulting boson clouds can be long-lived (on cosmological/astrophysical timescales) creating temporary BH hair that can be searched for in future GW observations, either in isolation or as a member of a binary system.

In this paper, we studied the dynamics of scalar condensates in binary inspirals, both with real and complex fields. We showed that the presence of a companion greatly enriches the dynamics of the system, by inducing mixing between growing and decaying modes of the cloud. At certain critical orbital frequencies the evolution becomes nonperturbative, leading to enhanced mixing through resonances with the orbital motion of the system. By restricting to simplified two- and three-state models of the cloud, we have shown that the resulting time dependence can significantly affect physical observables through a sharp depletion (or at least an attenuation) of the energy density in the bosonic field. This has important phenomenological consequences, both for the monochromatic signal emitted by the cloud and the finite-size effects imprinted in the waveforms of the binary signal. The characteristic time dependence of the signals thus become a very distinctive feature of the dynamics of boson clouds in binary systems, which would not be present in other scenarios, e.g., boson stars.

The results presented here should be considered a first step towards more accurate descriptions of the dynamics of the cloud. For instance, a proper calculation of the finite-size effects associated with the BH cloud will require numerical simulations, to incorporate the backreaction on the background geometry. We have also ignored the backreaction on the BH mass and spin due to the evolution of the surrounding cloud. Moreover, our analytical treatment breaks down when higher modes are excited during the resonances, as well as during the late stages of the inspiral, where the dynamical system becomes strongly time dependent. In these cases, only numerical approaches are able to properly inform us about the exact details of the dynamics. Other aspects of our estimations can also be improved upon. For example, we have restricted our analysis to quasi-circular orbits on the equator. This was only for simplicity, and it would be relatively straightforward to extend our analysis to the case of elliptic orbits. Finally, we did not perform a detailed forecast of the expected event rates, and signal strengths, with future GW observatories. This requires a careful treatment of formation mechanisms, as well as the associated generation of boson clouds.

In spite of our simplifications, we have identified robust GW signatures for clouds in binary inspirals, which open up new ways of probing physics beyond the Standard Model through current and future GW observations. In particular, we hope that future measurements could not only detect these ultralight particles, but also help us elucidate their putative properties. Following an analogy with collider physics, we expect to extract not only the particle’s mass, but also its spins and self-couplings [150], ultimately realizing the idea of the “gravitational collider.”

ACKNOWLEDGMENTS

We thank Mina Arvanitaki, Masha Baryakhtar, Alessandra Buonanno, Sam Dolan, William East, Lotte ter Haar, Tanja Hinderer, Badri Krishnan, Luis Lehner, Ilya Mandel, Mehrdad Mirbabayi, David Nichols, Samaya Nissanka, Paolo Pani, Frans Pretorius, John Stout, Ka Wa Tsang, Benjamin Wallisch, and Matias Zaldarriaga for helpful comments and discussions. We thank Benjamin Wallisch for assistance in producing some of the figures. D. B. is grateful to the Leung Center for Cosmology and Particle Astrophysics at the National Taiwan University and the Institute for Advanced Study for hospitality while some of this work was being performed. The work of D. B. is supported by a Vidi grant of the Netherlands Organisation for Scientific Research (NWO) that is funded by the Dutch Ministry of Education, Culture and Science (OCW). The work of H. S. C. is supported by NWO. The work of R. A. P. was supported by the Simons Foundation.
and São Paulo Research Foundation Young Investigator Awards, Grants No. 2014/25212-3 and No. 2014/10748-5, and by the German Science Foundation (DFG) within the Collaborative Research Center (SFB) 676 “Particles, Strings and the Early Universe.” R. A. P. would like to thank the International Center for Theoretical Physics (ICTP) in Trieste for hospitality while this work was being completed. R. A. P. is grateful to Leonidas K. Porto, Emilianio A. Porto and Imme Roewer for the unconditional support.

APPENDIX A: GRAVITATION ATOM

In this appendix, we derive some of the properties of the BH cloud presented in the main text. In particular, we compute the corrections to the frequency eigenvalues of the Hamiltonian in the nonrelativistic limit. We also determine the mass quadrupole moment of the BH cloud in the Newtonian approximation. Unless stated otherwise, these results apply to both real and complex scalar fields.

1. Kerr geometry

The Kerr metric in Boyer-Lindquist (BL) coordinates is

\[
ds^2 = -\frac{\Delta}{\rho^2} (dt - a\sin^2\theta d\phi)^2 + \frac{\rho^2}{\Delta} dr^2 + r^2 d\theta^2 + \frac{\sin^2\theta}{\rho^2} (a dt - (r^2 + a^2) d\phi)^2,
\]

where we have defined

\[
\Delta = r^2 - 2Mr + a^2 \quad \text{and} \quad \rho^2 = r^2 + a^2 \cos^2 \theta. \tag{A2}
\]

Since the metric components have no explicit dependence on \( t \) and \( \phi \), the Kerr spacetime is stationary and axisymmetric. The roots of \( \Delta \) determine the event horizon at \( r_+ = M + \sqrt{M^2 - a^2} \) and the Cauchy horizon at \( r_- = M - \sqrt{M^2 - a^2} \). The surface \( \mathcal{H} \) is an ergosphere. The angular velocity of the spinning BH at its event horizon is

\[
\Omega_H = \frac{a}{2Mr_+}. \tag{A3}
\]

The BH angular velocity \( \Omega_H \) plays a crucial role in the superradiance condition (2.1).

2. Test scalar field

The quadratic Lagrangian of a real scalar field is\(^{27}\)

\[
\mathcal{L} = -\frac{1}{2} g^{ab} \nabla_a \Psi \nabla_b \Psi - \frac{1}{2} \mu^2 \Psi^2. \tag{A4}
\]

Ignoring the backreaction of the field’s stress-energy on the metric, we can use the Kerr solution (A1) for the metric \( g^{ab} \). Expressed in BL coordinates, the Klein-Gordon equation is separable into a set of ordinary differential equations. To find the stationary solutions of the KG equation, we consider the following ansatz

\[
\Psi(t, r, \theta, \phi) = \sum_{\ell m} \frac{1}{2\pi} \int \frac{d\omega}{2\pi} \Re[e^{-i\omega t} e^{i\ell \phi} R_{\ell m}(r) S_{\ell m}(\theta, c)]. \tag{A5}
\]

The plane wave solutions in \( t \) and \( \phi \) are a direct consequence of the isometries of the Kerr solution. The angular functions \( S_{\ell m}(\theta, c) \) are called spherical harmonics with spheroidicity parameter \( c^2 \equiv a^2(\sin^2 \theta - \mu^2) \). In the limit \( c^2 \to 0 \), we have \( S_{\ell m}(\theta, c) e^{i\ell \phi} \to Y_{\ell m}(\theta, \phi) \), where \( Y_{\ell m} \) are the ordinary spherical harmonics. The radial functions \( R_{\ell m}(r) \) do not admit analytic solutions and have to be determined numerically. However, as we shall see, at leading order in an expansion in \( r^{-1} \), they take the form of the radial functions of the hydrogen atom.

3. Nonrelativistic limit

The ansatz (A5) corresponds to stationary solutions of a massive scalar field around the BH. To study the dynamical evolution of \( \Psi \), it proves useful to make another ansatz

\[
\Psi(t, r) = \frac{1}{\sqrt{2\mu}} [\psi(t, r)e^{-i\mu t} + \psi^*(t, r)e^{i\mu t}], \tag{A6}
\]

where \( \psi \) is a complex scalar field which varies on a timescale that is much longer than \( \mu^{-1} \). The action for \( \psi \) reads

\[
S = \int d^4x \sqrt{-g} \left( -\frac{1}{2\mu} \left( \nabla a \nabla^a \psi \nabla^a \psi + i\mu g^{0a} (\nabla^a \psi \nabla^a \psi - \nabla^a \psi \nabla a \psi^*) + \mu^2 (g^{00} + 1) \psi \psi^* \right) \right), \tag{A7}
\]

which spontaneously breaks time reparametrization. Expanding in powers of \( r^{-1} \), we obtain the following effective action

\[
S = \int dt dv r d\phi d^2 \psi \sin \theta [\mathcal{L}_2 + \mathcal{L}_4 + \mathcal{L}_5 + \cdots], \tag{A8}
\]

where \( \mathcal{L}_n \) denotes terms of order \( O(\alpha^n) \),

\(^{27}\)Since a complex scalar field may be described in terms of two real fields, we will restrict the following discussion to the real case only.
This is analogous to the nonrelativistic limit of the hydrogen atom. By expanding $\psi$ in terms of the stationary eigenstates

$$\psi(t, r) = \int \frac{d\omega}{2\pi} \sum_{\ell m} \psi_{n\ell m}(t, r),$$

and using (A5), we can identify

$$\psi_{n\ell m}(t, r, \theta, \phi) \simeq e^{-i(\omega-\mu)t} R_{n\ell m}(r) Y_{\ell m}(\theta, \phi),$$

where we have used the small spheroidicity approximation $S_{\ell m}(\theta)e^{in\phi} \simeq Y_{\ell m}(\theta, \phi) + \mathcal{O}(\alpha^2)$, and defined $R_{n\ell m}(r) \equiv \sqrt{\mu/2} R_{n\ell m}(r)$. Since we have expanded the action in (A8) at large distances, the boundary condition at the BH event horizon is replaced by a regular boundary condition at the origin. The radial function $R_{n\ell m}(r)$ therefore take the form of the radial functions of the hydrogen atom.

The Lagrangians $L_4$ and $L_5$ are higher-order self-interaction terms. Using the nonrelativistic equation of motion (A10), we can substitute $i\partial_t \psi$ and $i\partial_r \psi = im\psi$ back into $L_4$ and $L_5$. This gives

$$L_4 + L_5 = \frac{1}{2\mu} \partial_\phi \psi^* \partial_\phi \psi + \frac{\alpha}{\mu^2 r^3} \psi^*$$

$$\times |\ell(\ell + 1) - 2r\partial_r(r\partial_r) - 2\tilde{a}\alpha| \psi,$$

where we have used the eigenvalue equation of the spherical harmonics

$$\left( \frac{1}{\sin \theta} \partial_\theta \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) Y_{\ell m}(\theta, \phi)$$

$$= -\ell(\ell + 1) Y_{\ell m}(\theta, \phi).$$

Finally, it is convenient to write the dynamics of $\psi$ in terms of the Hamiltonian

$$H_c \equiv T_c + V_c,$$

where we have separated $H_c$ into its kinetic and potential components,

$$T_c \equiv -\frac{1}{2\mu} \mathbf{\nabla}^2 + \frac{1}{\mu^2 r^3},$$

$$V_c \equiv -\frac{\alpha}{\mu^2 r^3} [\ell(\ell + 1) - 2r\partial_r(r\partial_r)] + \frac{2\tilde{a}\alpha^2}{\mu^2 r^3}.$$
occupation density of a particular state. This is analogous to the probability density in quantum mechanics, after suitable normalization.

6. Quadrupole moment

The conserved energy associated with the energy-momentum tensor of the scalar field is

$$E \equiv - \int_{\Sigma} T^0_\alpha k^\alpha, \quad (A20)$$

where \( k^\alpha = \delta^\alpha_0 \) is the time-like Killing vector field of the Kerr metric and \( \Sigma \) is a spacelike hypersurface of constant \( t \). At leading order in \( r^{-1} \) and \( \alpha \), the energy density of the [211] mode of the cloud is

$$\rho_E(t, \mathbf{r}) = -T^0_0 = \frac{A^2 \mu^6 \alpha^5}{64 \pi} r^2 e^{-\alpha r} \sin^2 \theta, \quad (A21)$$

where \( A \) is the amplitude of the scalar field, which is determined by the efficiency of the supersymmetric instability. At leading order in \( \alpha \), the energy density \( \rho_E \) is equal to the mass density \( \rho_m \) of the cloud and the mass quadrupole moment can be approximated by

$$Q_c = \int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\phi r^2 \sin \theta |\rho_m|^2 P_2(\cos \theta), \quad (A22)$$

where \( P_2(\cos \theta) \) is the Legendre polynomial. Normalizing the amplitude \( A \) in \( (A22) \) by the mass of the cloud

$$M_c = \int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\phi r^2 \sin \theta \rho_m, \quad (A23)$$

we obtain

$$Q_c \approx -\frac{6M_c \alpha^2}{\mu^2 \alpha^2} \approx \frac{3}{8} M_c r_c^2. \quad (A24)$$

The negative sign arises as a direct consequence of the spinning motion of the cloud.

APPENDIX B: FREE-FALLING CLOUDS

In Sec. III A, we derived the effect of the companion on the BH cloud by considering the perturbed metric in Fermi coordinates. In this appendix, we present an alternative perspective of this derivation by considering a three-body analogy. Our main goal is to show how a fictitious dipole in the gravitational potential, generated by a change of coordinates, cancels out.

1. Three-body analogy

The eigenfunctions of the cloud are determined by the Hamiltonian \( H_c \) of a test particle of mass \( \mu \), cf. (A15). This is analogous to the Hamiltonian of the electron in a hydrogen atom: while the electron wavefunction, as a solution of the Schrödinger equation, has a wavelike distribution around its nucleus, it is treated as a point particle at the level of the Hamiltonian. In the rest of this appendix, we will adopt this particle picture of the Hamiltonian.

In the presence of a companion, of mass \( M_\ast \), the BH cloud will accelerate due to the external gravitational field. The Hamiltonian (A15) must hence be modified to include both the kinetic term of \( M \) as well as the contributions from \( M_\ast \). It is convenient to introduce the center-of-mass coordinate, \( \mathbf{r} \equiv (M\mathbf{r}_1 + \mu \mathbf{r}_2)/(M + \mu) \), so that the total Hamiltonian \( H_{\text{tot}} \) of the three-body system can be written as

$$H_{\text{tot}} = \frac{p_\mathbf{r}^2}{2(M + \mu)} + \frac{p_\mu^2}{2\mu} + V_c(|\mathbf{r}|) + \left[ \frac{p_r^2}{2M_\ast} + V_s(\mathbf{r}, \mathbf{R}_s) \right], \quad (B1)$$

where \( \mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 \) is the relative spatial separation between \( \mu \) and \( M \), and we introduced the reduced mass \( \tilde{\mu} \equiv M\mu/(M + \mu) \), the momenta \( \mathbf{p}_\mu \equiv (M + \mu)\dot{\mathbf{r}} \) and \( \mathbf{p}_\mathbf{r} \equiv \mu \dot{\mathbf{r}} \), and the coordinates \( \mathbf{R}_s \equiv \{R_s, \Theta_s, \Phi_s\} \) and \( \mathbf{r} \equiv \{\tilde{r}, \tilde{\theta}, \tilde{\phi}\} \) relative to the center-of-mass (cf. Fig. 11). Working in the Newtonian limit, the external potential \( V_s \) is given by

$$V_s(\mathbf{r}, \mathbf{R}_s) = -\frac{M_\ast M}{|\mathbf{R}_s + \mu \mathbf{r}/M|} - \frac{M_\ast \mu}{|\mathbf{R}_s - \mathbf{r}|}$$

$$= -M_\ast \sum_{\ell, m} \frac{4\pi}{2\ell + 1} \left( \frac{M(-\mu/M)^{\ell+\mu}}{R_s} \right) \times \left( \frac{\tilde{r}}{R_s} \right)^{\ell} Y_{\ell, m}^\ast(\Theta_s, \Phi_s)Y_{\ell', m}(\tilde{\theta}, \tilde{\phi}). \quad (B2)$$

Substituting \( \mathbf{r} = (1 + \mu/M)^{-1} \mathbf{r} \), and expanding in \( \mu/M \ll 1 \), we find
The leading monopole term \( V_{s,0} \) determines the acceleration of the center-of-mass of the BH-cloud system. It does not depend on the relative separation \( r \), so it does not lead to a shift in the energy levels or the mode functions of the system. The remaining terms are a sum over harmonics, starting with the quadrupole \( \ell_s = 2 \). Importantly, the dipole contribution \( \ell_s = 1 \) vanishes in the center-of-mass frame.

2. Fictitious dipole

We have shown that the contribution from the dipole vanishes in the centre-of-mass frame. By virtue of the equivalence principle, this property must hold for all coordinate systems. More generally, this is a manifestation of the fact that a constant gravitational gradient—in this case the dipole term induced by the external gravitational field produced by the companion \( M_c \)—is physically unobservable. We now show explicitly that this is indeed the case for the coordinate system centered at \( M_c \).

Consider expressing \( H_{tot} \) in terms of \((r_1, r)\) instead of \((\rho, r)\),

\[
H_{tot} = \left[ \frac{1}{M} + \frac{p_1^2}{2M} + \frac{p^2}{2M} \right] + \frac{\mu}{M} p_1 \cdot p + V_s(r) \tag{B4}
\]

where the gravitational potential in this choice of coordinates reads (cf. Fig. 11)

\[
V_s(r, R) = -\frac{M_s M}{|R - r|} - \frac{M_s \mu}{|R - r|}, \tag{B5}
\]

Expanding \( V_s \) for \( r \ll R \) produces a dipole term

\[
-\frac{M_s \mu}{|R - r|} \approx -\left( \frac{M_s \mu}{R} \right) \left( \hat{n} \cdot \frac{r}{R} \right), \tag{B6}
\]

where \( \hat{n} \) is the unit vector along \( R \), and we have used \( |R - r| = R - \hat{n} \cdot r + \mathcal{O}(r^2) \). We will now show that this dipole is canceled by the kinetic mixing term \( \mu p_1 \cdot p / M \hat{\mu} \).

This is manifested most transparently in the \( M_c / R \gg V_s \) limit, in which the gravitational attraction between \( M \) and \( \mu \) is negligible, compared to the force exerted by \( M_c \). In this limit, \( M \) and \( \mu \) free fall separately under the gravitational influence of \( M_c \):

\[
r_1^2 = \frac{2M_s}{R}, \quad r_2^2 = \frac{2M_s}{|R - r|}, \tag{B7}
\]

and the angle between \( \hat{r}_1 \) and \( \hat{r}_2 \), denoted by \( \gamma \), is given by

\[
\cos \gamma = \frac{R^2 + |R - r|^2 - r^2}{2R|\text{R} - \text{r}|}. \tag{B8}
\]

The dipole arising from the kinetic mixing becomes,

\[
\frac{\mu}{M} p_1 \cdot p_r = \mu (\hat{r}_1 \cdot \hat{r}_2 - \hat{r}_1^2) \tag{B6}
\]

\[
= \mu \left( \frac{M_s (R^2 + |R - r|^2 - r^2)}{(R|R - r|)^{3/2}} - \frac{2M_s}{R} \right)
\]

\[
= \mu \left( \frac{M_s}{R} \left( 2 + \frac{n \cdot r}{R} \right) - \frac{2M_s}{R} + \mathcal{O} \left( \frac{r}{R} \right)^2 \right)
\]

\[
= + \left( \frac{M_s \mu}{R} \right) \left( \hat{n} \cdot \frac{r}{R} \right) + \mathcal{O} \left( \frac{r}{R} \right)^2. \tag{B9}
\]

As advertised, this precisely cancels the dipole contribution in (B6).

APPENDIX C: EFT OF EXTENDED OBJECTS

In the language of effective field theory, spinning extended objects are described in terms of an effective worldline action [11–22,28,33–38]. One of the virtues of the formalism is the inclusion of finite-size effects as a series of higher-dimension terms beyond the “minimally coupled” point-particle worldline action, \( S_{pp} = M \int dt \), where \( M \) and \( \tau \) are the mass and proper time, respectively. One such term is the electric-type quadrupole coupling,

\[ S_Q = -\frac{1}{2} \int d\tau Q^{ij}(\tau)E_{ij}(\tau), \tag{C1} \]

where \( E_{ij} \) is the electric component of the Weyl tensor projected onto the spatial hypersurface of a free-falling basis,\(^{28}\) and \( Q^{ij} \) is the mass quadrupole moment of the object. A similar coupling appears for the magnetic component. Higher-order multipole moments can also be easily incorporated.

The quadrupole moment can be further split into two parts: i) a background term, which is independent of any external perturbations, and ii) the response induced by an external field, for instance the gravitational field induced by a companion in a binary system. The effects of a permanent quadrupole moment depend on the scaling of \( Q^{ij} \) with the size and spin of the body. For example, the symmetric and trace-free spin-induced quadrupole moment of a rotating body can be parameterized as

\[ (Q^{ij})_S = -\frac{C_{ES}}{M} \left( S'^i S'^j - \delta^{ij} S'^k S^k S^k \right). \tag{C2} \]

\(^{28}\)The electric part of the Weyl tensor \( W_{abcd} \) is defined as \( E_{\alpha\beta} = W_{abcd} u^a u^b \), with \( u^a \) being the four-velocity. The tetrad is chosen such that \( e^0_0 = u^0 \). Since \( E_{ij} u^0 = 0 \), only the spatial indices matter in the projection to the free-falling frame \( E_{ij} = e^0_i e^0_j E_{\alpha\beta} \).
where \( S^i \) is the spin vector and \( C_{ES} \) is a dimensionless Wilson coefficient that incorporates the intrinsic properties of the object. The factor of \( 1/M \) is chosen for convenience, so that \( C_{ES} = 1 \) for Kerr BHs. This parameter is equivalent to \( \kappa \), introduced in Sec. IV B, and measures the spin-induced quadrupolar deformability of the object, relative to the value for a rotating BH with the same mass and spin. At leading order, the quadrupole term leads to the following additional term in the worldline effective action \([12,28,33,34]\)

\[
S_{ES} = \frac{C_{ES}}{2M} \int \mathrm{d}t S^{ij} \frac{S^{jk} E_{ij}}{R},
\]

where \( S^{ij} \) is the spin tensor, so that \( S^i = \frac{1}{2} E^{ijk} S_{jk} \). Using the power counting rules of the EFT, it is straightforward to show that this term contributes at 2PN order for binary systems with a rapidly spinning compact object (see \([22]\) for more details).

In the presence of an external perturbation, the object’s multipole moments also receive induced corrections proportional to the external field. For example, in the static limit, the response part of the quadrupole moment is \(^{29}\)

\[
(Q_{ij})_R = -C_E E_{ij},
\]

where we have introduced the Wilson coefficient \( C_E \), often called the “Love number.” In the worldline theory, this leads to the following higher-dimension term in the effective action,

\[
S_{E^2} = \frac{C_E}{2} \int \mathrm{d}t E_{ij} E^{ij},
\]

which is quadratic in the external field. Notice that the parameter \( C_E \) scales as [mass] \( \times \) [size] \(^4\) for general bodies, and hence as [mass] \(^5\) for compact objects. It is thus conventional to introduce the dimensionless Love number, or “deformability parameter,” as (see e.g., \([5]\)\)

\[
\Lambda \equiv \frac{C_E}{M^5}.
\]

Because of the many derivatives involved, this term enters at 5PN order for compact objects in a binary system, e.g., \([11]\). Perhaps somewhat surprisingly, the Love numbers vanish for BHs in four-dimensional general relativity. This unexpected “fine-tuning” offers a unique opportunity to probe the nature of spacetime, and the existence of exotic objects, through precision GW data \([7]\).

\(^{29}\)In general, the response function depends on the frequency of the external field, with \( C_E \) being the real part in the low-frequency limit. The imaginary part of the response is responsible for the absorptive properties of the body. See \([13,14]\) for more details.

---


[150] D. Baumann et al. (to be published).