

Ionization of gravitational atoms

Daniel Baumann,^{1,2,3} Gianfranco Bertone¹, John Stout,⁴ and Giovanni Maria Tomaselli¹

¹*Gravitation Astroparticle Physics Amsterdam (GRAPPA), University of Amsterdam, Amsterdam 1098 XH, Netherlands*

²*Center for Theoretical Physics, National Taiwan University, Taipei 10617, Taiwan*

³*Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan*

⁴*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*



(Received 1 March 2022; accepted 3 June 2022; published 27 June 2022)

Superradiant instabilities may create clouds of ultralight bosons around rotating black holes, forming so-called “gravitational atoms”. It was recently shown that the presence of a binary companion can induce resonant transitions between bound states of these clouds, whose backreaction on the binary’s orbit leads to characteristic signatures in the emitted gravitational waves. In this work, we show that the interaction with the companion can also trigger transitions from bound to unbound states of the cloud—a process that we refer to as “ionization” in analogy with the photoelectric effect in atomic physics. The orbital energy lost in the process overwhelms the losses due to gravitational wave emission and contains sharp features carrying information about the energy spectrum of the cloud. Moreover, we also show that if the companion is a black hole, then the part of the cloud impinging on the event horizon will be absorbed. This “accretion” leads to a significant increase of the companion’s mass, which alters the dynamical evolution and ensuing waveform of the binary. We argue that a combined treatment of resonances, ionization, and accretion is crucial to discover and characterize gravitational atoms with upcoming gravitational-wave detectors.

DOI: [10.1103/PhysRevD.105.115036](https://doi.org/10.1103/PhysRevD.105.115036)

I. INTRODUCTION

Black holes are remarkably simple objects. The spacetime around a black hole is uniquely determined by its mass and spin, and the gravitational waves (GWs) released in the merger of two black holes can be predicted very precisely. This makes black holes exceptionally clean environments to probe the fundamental laws of nature [1–4], with any deviation from the predictions of general relativity being an indication of new physics.

A particularly well-studied example of new physics, accessible with future GW observations, are ultralight bosons. Such bosons can be generated by superradiance [5], forming long-lived condensates (“clouds”) around rotating black holes [6–9]. Due to their strong similarity with the hydrogen atom, such systems have been called “gravitational atoms”. For isolated gravitational atoms, there are essentially two ways of inferring the presence of these boson clouds. First, rotating clouds will emit gravitational waves [7] that can be looked for in continuous-wave searches [10]. Second, the clouds extract spin from their parent black holes and this spin down can be inferred statistically in a population of rotating black holes [7,11–15]. The existence of rapidly-spinning black holes would then rule out ultralight bosons in a certain mass range. Unfortunately, neither of these effects is very distinctive, so it is hard to use them as a way of unambiguously discovering gravitational atoms in the sky.

Recently, a new avenue for detecting gravitational atoms has been explored which exploits their effects in binary systems. When a gravitational atom is part of a binary it gets perturbed by the companion. As was shown in [16], the gravitational interaction between the companion and the boson cloud is resonantly enhanced when the orbital frequency matches the energy difference between two eigenstates of the cloud; see [17–28] for related work. This leads to an analog of the Landau-Zener transition in quantum mechanics [29–31], where the companion forces the cloud to smoothly transition from one state to another. These transitions are a distinctive fingerprint of a boson cloud.

In this paper we extend the treatment of [16,31] to allow for transitions to unbound states of the cloud. When these transitions are effective, the cloud escapes from the parent black hole and the gravitational atom gets “ionized”, like in the photoelectric effect for ordinary atoms. Figure 1 illustrates the main result of this analysis. Shown is the ionization power P_{ion} , the rate of energy lost by the binary due to ionization, as a function of the separation R_* between the parent black hole and companion for typical parameters of the system. We see that the effect of ionization can be orders of magnitude larger than the rate of energy lost due to GW emission, P_{GW} , and therefore dominate the binary’s dynamics. Moreover, this ionization power is not a smooth function of R_* , but contains interesting “discontinuities” at specific separations. These sharp features

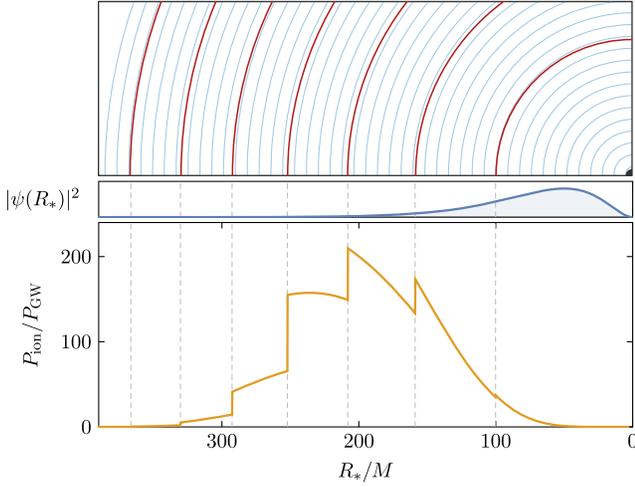


FIG. 1. Schematic illustration of the ionization of the gravitational atom. In the bottom panel, we plot the ratio of the ionization power P_{ion} and the power lost due to GW emission P_{GW} . We see that the energy loss due to ionization can overwhelm that due to GW emission and hence dominate the binary's dynamics. The signal has sharp features when the bound state begins to resonate with the continuum, which occurs at specific separations R_* . Shown also is the density profile of the cloud, $|\psi(R_*)|^2$, for a $|211\rangle$ bound state.

arise when the bound state begins to resonate with the continuum. Like the resonant transitions between bound states [16,31], the ionization signal therefore contains very distinctive information about the microscopic structure of the cloud. We will show how these features are imprinted in the binary's dynamics, and thus in the emitted gravitational waves.

For the first time, we also include the accretion of the cloud onto the companion. Although accretion is suppressed for the wavelike boson cloud (compared to an equal density of particles), it is nevertheless a large effect, since the typical densities of the boson clouds are large. In many cases, the mass of the companion can change by up to an order-one fraction during the inspiral, leading to a significant speedup of the merger compared to the vacuum evolution. Unlike the ionization signal, the effect of accretion is a smooth function of the separation R_* and hence more degenerate with changes in the source parameters.

We conclude that ionization and accretion play a critical role in the phenomenology of gravitational atoms in binaries. Rather remarkably, the effects can be so large that they overwhelm the energy lost due to GW emission and therefore drive the inspiral (rather than just being a small perturbation). A consistent treatment of these systems must therefore take these effects into account, as well as their interplay with the resonances between bound states studied in [16,31].

Outline. The outline of the paper is as follows: In Sec. II we review the energy eigenstates of the gravitational atom and describe the resonant transitions between bound states.

In Sec. III we study transitions to unbound states and describe the ionization of the boson cloud. Special attention is paid to the sharp resonance features in the ionization power. In Sec. IV we compute the accretion of the wavelike boson cloud onto the companion black hole. In Sec. V we show how both ionization and accretion change the dynamics of the binary. We present numerical results for a few representative examples. Finally, we state our conclusions and discuss open problems in Sec. VI.

A number of appendices contain technical details. In Appendix A we describe the various approximations that are used in Sec. III to integrate out the continuum states and derive the effective dynamics of the bound states. Such a description only holds in the Markov approximation, whose validity we discuss in Appendix B, and we argue there that this approximation applies to the systems we consider in the main text. In Appendix C we derive an approximation for the ionization power, which measures how quickly the companion transfers energy from bound to unbound states. In Appendix D, we discuss the low-energy limit of the unbound states and describe under which conditions the discontinuities seen in Fig. 1 appear. Finally, in Appendix E, we describe the exact solutions of the Klein-Gordon equation in the Kerr geometry, and discuss an approximation relevant for our derivation of the accretion rate in Sec. IV.

Notation and conventions. Our metric signature will be $(-, +, +, +)$ and, unless stated otherwise, we will work in natural units with $G = \hbar = c = 1$. Greek letters will stand for spacetime indices. Quantities associated to the boson clouds will be denoted by the subscript c . For example, the initial mass and angular momentum of the cloud are M_c and S_c , respectively. The gravitational fine-structure constant, $\alpha = \mu M$, is the ratio of the gravitational radius of the black hole (which in natural units is simply $r_g = M$) and the (reduced) Compton wavelength of a boson field, $\lambda_c = \mu^{-1}$, where μ is the mass of the field.

The Kerr metric for a black hole of spin J is

$$ds^2 = -\frac{\Delta}{\rho^2}(dt - a \sin^2 \theta d\phi)^2 + \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2 + \frac{\sin^2 \theta}{\rho^2} (adt - (r^2 + a^2)d\phi)^2, \quad (1.1)$$

where $a \equiv J/M$, $\Delta \equiv r^2 - 2Mr + a^2$ and $\rho^2 \equiv r^2 + a^2 \cos^2 \theta$. The roots of Δ determine the inner and outer horizons, located at $r_{\pm} = M \pm \sqrt{M^2 - a^2}$, and the angular velocity at the outer horizon is $\Omega_+ \equiv a/2Mr_+$. Dimensionless quantities, defined with respect to the black hole mass M , are labeled by tildes. For example, the dimensionless spin of the black hole is $\tilde{a} \equiv a/M$. We use an asterisk to denote quantities associated to the black hole companion; for instance, M_* and a_* are the mass and spin of the

companion, while $q = M_*/M$ is the ratio of the black hole masses.

II. GRAVITATIONAL ATOMS IN BINARIES

We begin with a brief review of the structure of the gravitational atom. We start, in Sec. II A, by describing the bound and unbound spectra of the atom in isolation. In Sec. II B, we explain how a binary companion perturbs this atom, mediating transitions between different states. We then describe the case of resonant transitions between bound states in Sec. II C.

A. Scalar field around Kerr

The Klein-Gordon equation for a scalar field of mass μ in a curved spacetime is

$$(g^{\alpha\beta}\nabla_\alpha\nabla_\beta - \mu^2)\Phi(t, \mathbf{r}) = 0. \quad (2.1)$$

As is well known, in the Kerr background (1.1), the Klein-Gordon equation admits bound state solutions that are remarkably similar to those of the hydrogen atom. When the Compton wavelength of the field is much larger than the gravitational radius of the black hole, $\alpha \equiv r_g/\lambda_c \ll 1$, it is useful to consider the following ansatz

$$\Phi(t, \mathbf{r}) = \frac{1}{\sqrt{2\mu}} [\psi(t, \mathbf{r})e^{-i\mu t} + \psi^*(t, \mathbf{r})e^{+i\mu t}], \quad (2.2)$$

where ψ is a complex scalar field which varies on time-scales much longer than μ^{-1} , see e.g., [32]. If Φ is itself a complex scalar field, then we only use the first term in (2.2). We will often refer to ψ as the wave function of the cloud. Far from the black hole and at leading order in α , the Klein-Gordon equation (2.1) is then identical to the Schrödinger equation for the hydrogen atom,

$$i\frac{\partial}{\partial t}\psi(t, \mathbf{r}) = \left(-\frac{1}{2\mu}\nabla^2 - \frac{\alpha}{r}\right)\psi(t, \mathbf{r}). \quad (2.3)$$

In this limit, the scalar field can be studied using standard techniques of nonrelativistic quantum mechanics. This Schrödinger equation permits two qualitatively different sets of eigenstates (see Fig. 2), whose properties we will now review.

1. Bound states

We first consider the familiar bound state solutions, which are labeled by three integers; a principal ‘‘quantum number’’ n , orbital angular momentum ℓ , and azimuthal angular momentum m . At leading order in α , these bound state solutions have the form

$$\psi_{n\ell m}(t, \mathbf{r}) = R_{n\ell}(r)Y_{\ell m}(\theta, \phi)e^{-i(\omega_{n\ell m} - \mu)t}, \quad (2.4)$$

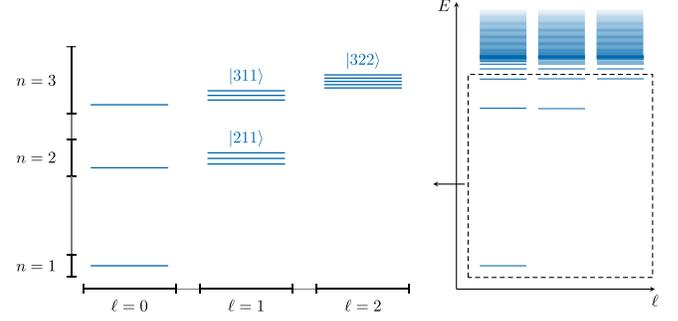


FIG. 2. Illustration of the spectrum of bound and unbound states of the gravitational atom.

where $Y_{\ell m}(\theta, \phi)$ are spherical harmonics and $R_{n\ell}(r)$ are the hydrogenic radial functions. The latter are given by

$$R_{n\ell}(r) = \sqrt{\left(\frac{2\mu\alpha}{n}\right)^3 \frac{(n-\ell-1)!}{2n(n+\ell)!}} \left(\frac{2\alpha\mu r}{n}\right)^\ell \times \exp\left(-\frac{\mu\alpha r}{n}\right) L_{n-\ell-1}^{2\ell+1}\left(\frac{2\mu\alpha r}{n}\right), \quad (2.5)$$

where $L_{n-\ell-1}^{2\ell+1}(x)$ is the associated Laguerre polynomial. For small values of α , the radial profile peaks at a multiple of the ‘‘Bohr radius’’ $r_c \equiv (\mu\alpha)^{-1}$ and decays exponentially as $r \rightarrow \infty$. These bound state solutions are defined for $n \geq \ell + 1$, $\ell \geq 0$, and $\ell \geq |m|$. For notational simplicity, it will be convenient to lean on the quantum mechanical analogy and represent (2.4) using the bra-ket notation $|n\ell m\rangle$. The normalization of the bound states is chosen so that

$$\langle n\ell m | n'\ell' m' \rangle = \int d^3r \psi_{n\ell m}^*(t, \mathbf{r})\psi_{n'\ell' m'}(t, \mathbf{r}) = \delta_{nn'}\delta_{\ell\ell'}\delta_{mm'}. \quad (2.6)$$

The amplitude of (2.4) is determined by the total mass of the cloud and will be restored when necessary.

There is one important difference between the hydrogen atom and the gravitational atom. While the wave functions of the former are regular at $r = 0$, the latter must be purely ingoing at the black hole’s outer horizon since no physical mode can escape from the black hole. This ‘‘dissipative’’ boundary condition forces the bound state eigenfrequencies of the boson cloud to be complex,

$$\omega_{n\ell m} = E_{n\ell m} + i\Gamma_{n\ell m}, \quad (2.7)$$

where $E_{n\ell m}$ and $\Gamma_{n\ell m}$ denote the energies and instability rates, respectively. At leading order in α , these are

$$E_{n\ell m} = \mu \left(1 - \frac{\alpha^2}{2n^2} + \mathcal{O}(\alpha^4)\right), \quad (2.8)$$

$$\Gamma_{n\ell m} = 2\tilde{r}_+ C_{n\ell} g_{\ell m}(\tilde{a}, \alpha, \omega)(m\Omega_+ - \omega_{n\ell m})\alpha^{4\ell+5} + \mathcal{O}(\alpha^{4\ell+7}), \quad (2.9)$$

where the numerical coefficients $C_{n\ell}$ and $g_{\ell m}$ can be found in [32]. As discussed there, these bound states are still labeled by the ‘‘quantum’’ numbers n , ℓ , and m , and the latter two reduce to the orbital and azimuthal angular momenta of the cloud in the $\alpha \rightarrow 0$ limit. Crucially, the nonzero instability rates allow a rapidly spinning black hole to spontaneously shed a sizable fraction of its mass and angular momentum to form the boson cloud. Even though these rates are highly suppressed for $\alpha \ll 1$, the cloud can still grow very quickly on astrophysical timescales. The gravitational atom is the endpoint of this process. Since the state $|211\rangle$ grows fastest, we will take this as the initial configuration of the cloud when the binary inspiral begins.

In the nonrelativistic limit, a cloud in a state $|\psi\rangle$, with wave function $\psi(t, \mathbf{r})$, has mass density

$$\rho(t, \mathbf{r}) = \begin{cases} M_c |\psi(t, \mathbf{r})|^2 & \text{(complex field),} \\ 2M_c |\text{Re}\psi(t, \mathbf{r})|^2 & \text{(real field),} \end{cases} \quad (2.10)$$

where M_c is the initial mass of the cloud. By convention, we require that the cloud’s wave function is initially unit normalized, $\langle \psi | \psi \rangle = 1$. Superradiant growth can be quite efficient and, depending on the initial spin of the parent black hole, the mass of the cloud M_c can be a significant fraction of the total mass of the system (up to $0.1M$, where M is the mass of the central black hole). Since the typical size of the cloud r_c is between 10 and 10^3 times the Schwarzschild radius of the parent black hole for typical values of α , the cloud can be an exceptionally dense region of matter compared to other astrophysical environments. For example, if the cloud sits around a stellar mass black hole with $M = 10 M_\odot$, then the average mass density is between 10^8 kg/m^3 and 10^{12} kg/m^3 . On the other hand, around an intermediate mass black hole $M = 10^5 M_\odot$, the cloud can be much more spread out so that its average mass density is ‘‘only’’ 1 kg/m^3 to 10^4 kg/m^3 . As a point of reference, the density of water is $\rho_{\text{H}_2\text{O}} = 10^3 \text{ kg/m}^3$, so an inspiraling black hole companion moving through the cloud encounters a medium that can be potentially much denser than water. As we will see, the associated large flux of mass through the companion’s horizon can strongly impact the dynamics of the inspiral.

2. Continuum states

The Schrödinger equation (2.3) also permits continuum state solutions. In addition to the orbital and azimuthal angular momentum ℓ and m , these solutions are labeled by a positive, real-valued wave number k ,

$$\psi_{k;\ell m}(t, \mathbf{r}) = R_{k;\ell}(r) Y_{\ell m}(\theta, \phi) e^{-ie_{\ell m}(k)t}. \quad (2.11)$$

We distinguish the continuous index by a trailing semicolon and use the bra-ket notation $|k; \ell m\rangle$. In the hydrogen atom, these continuum states represent those states in which the electron has been unbound from the proton, and can thus be thought of as scattering states. A similar interpretation applies to the gravitational atom; these states represent the situation in which the scalar field is not bound to the black hole. The continuum radial functions are given by

$$R_{k;\ell}(r) = \frac{2ke^{\frac{\pi\mu\alpha}{2k}} |\Gamma(\ell+1 + \frac{i\mu\alpha}{k})|}{(2\ell+1)!} (2kr)^\ell \times e^{-ikr} {}_1F_1\left(\ell+1 + \frac{i\mu\alpha}{k}; 2\ell+2; 2ikr\right), \quad (2.12)$$

where ${}_1F_1(a; b; z)$ is the Kummer confluent hypergeometric function. In contrast to the bound states, these continuum states do not decay exponentially as $r \rightarrow \infty$ and are not unit-normalizable. The normalization is instead chosen so that

$$\begin{aligned} \langle k; \ell m | k'; \ell' m' \rangle &= \int d^3r \psi_{k;\ell m}^*(t, \mathbf{r}) \psi_{k';\ell' m'}(t, \mathbf{r}) \\ &= 2\pi \delta(k - k') \delta_{\ell\ell'} \delta_{mm'}, \end{aligned} \quad (2.13)$$

i.e., these continuum states are δ -function normalized.

Since the boundary conditions for these continuum states are much less restrictive than those for the bound states, the exact eigenfrequencies are known and are purely real $\omega(k) = \sqrt{\mu^2 + k^2}$, with $k \in [0, \infty)$. We will work in the nonrelativistic limit, $k \ll \mu$, where the dispersion relation for the continuum states is

$$\epsilon(k) \equiv \sqrt{\mu^2 + k^2} - \mu \approx \frac{k^2}{2\mu}. \quad (2.14)$$

In Sec. III B, we will find that only the continuum states with $k \sim \mathcal{O}(\mu\alpha^2)$ play an important dynamical role and, since we will always work in the limit $\alpha \ll 1$, we will not need to consider corrections to the nonrelativistic approximation.

According to the normalization condition (2.13), the continuum states are linearly distributed in k ; that is the density of states behaves as $dn \propto dk$. However, in terms of the energy ϵ , this density of states *diverges* as $\epsilon \propto k^2 \rightarrow 0$,

$$dk = \frac{\mu d\epsilon}{k(\epsilon)}. \quad (2.15)$$

This diverging density of states at low energies will play a crucial role in the ionization effects we describe in the next section.

An important related property of the continuum wave functions is that they vanish as \sqrt{k} in the soft limit $k \rightarrow 0$.

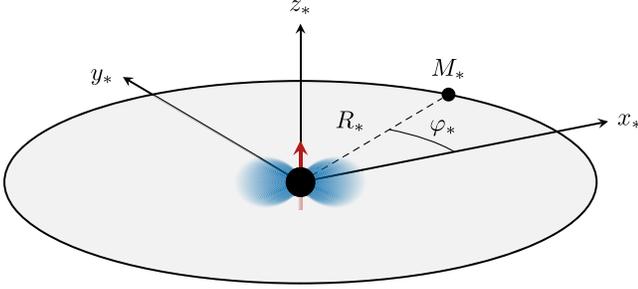


FIG. 3. Schematic diagram of an equatorial binary inspiral. The position of the companion with mass M_* can be described by the distance between the two black holes, R_* , and the true anomaly φ_* , which is the polar angle of the companion in the equatorial plane.

As we discuss in Appendix D, this behavior is ultimately due to the long-range nature of the gravitational potential, and we show there that

$$R_{k;\ell m}(r) \sim \sqrt{\frac{4\pi k}{r}} J_{2\ell+1}(2\sqrt{2\alpha\mu}r), \quad k \rightarrow 0, \quad (2.16)$$

where $J_\nu(z)$ is the Bessel function of the first kind. In contrast to the free particle, the long-range Coulombic potential localizes the zero mode to a Bohr radius-sized region around $r = 0$, instead of spreading out over all of space. As we will discuss in Sec. III, this seemingly innocuous behavior, combined with the divergent density of states (2.15), is responsible for dramatic discontinuities in the ionization power during the inspiral.

B. Perturbation from the companion

Our main goal is to understand the dynamics of the cloud during a binary inspiral. To this end, we must describe the effect that the binary companion has on the cloud through its gravitational field. This can be encoded in an additional potential term in the Schrödinger equation (2.3) and in this section we will describe both the structure of this potential and the transitions it mediates.

For simplicity, we restrict our attention to inspirals that occur in the equatorial plane of the cloud. As illustrated in Fig. 3, the relative motion of the companion is most conveniently described using the distance between the parent black hole and companion, R_* , and the so-called *true anomaly*, φ_* , which is the companion's polar angle in the equatorial plane.

Denoting the spatial coordinates of the cloud in its Fermi frame with $\mathbf{r} = \{r, \theta, \phi\}^1$ and working at leading order in α , the Schrödinger equation (2.3) is modified by the addition of the companion's gravitational potential

¹These coordinates coincide with the familiar Boyer-Lindquist coordinates at leading order in the post-Newtonian expansion. See [16,31] for more details.

$$V_*(t) = -q\alpha \sum_{\substack{\ell_* \geq 2 \\ |m_*| \leq \ell_*}} \varepsilon_{\ell_* m_*} e^{-im_* \varphi_*} Y_{\ell_* m_*}(\theta, \phi) \left(\frac{r^{\ell_*}}{R_*^{\ell_*+1}} \Theta(R_* - r) + \frac{R_*^{\ell_*}}{r^{\ell_*+1}} \Theta(r - R_*) \right), \quad (2.17)$$

where $q \equiv M_*/M$ is the mass ratio between the companion and the parent black hole, Θ is the Heaviside step function, and $\varepsilon_{\ell_* m_*} \equiv \frac{4\pi}{2\ell_*+1} Y_{\ell_* m_*}^*(\frac{\pi}{2}, 0)$. Importantly, we explicitly exclude the fictitious $\ell_* = 1$ dipole contribution, as it vanishes in the freely-falling frame and always eventually cancels in others [16].

This perturbation acts like a periodic driving force whose frequency slowly increases with time. In terms of the instantaneous frequency $\Omega(t) \equiv |\dot{\varphi}_*(t)|$, the true anomaly evolves according to

$$\varphi_*(t) = \pm \int_0^t dt' \Omega(t'), \quad (2.18)$$

where $t = 0$ is an initial reference time, and the upper (lower) sign denotes an orbit in which the companion corotates (counter-rotates) with the cloud. For the quasi-circular equatorial orbits we consider in this paper, the power emitted by gravitational waves is

$$P_{\text{GW}} \equiv \frac{dE_{\text{GW}}}{dt} = -\frac{32}{5} \frac{q^2}{(1+q)^2} M^2 R_*^4 \Omega^6, \quad (2.19)$$

and the orbital frequency evolves according to [33]

$$\frac{d\Omega}{dt} = \gamma \left(\frac{\Omega}{\Omega_0} \right)^{11/3}, \quad \text{with} \quad \gamma \equiv \frac{96}{5} \frac{q}{(1+q)^{1/3}} M^{5/3} \Omega_0^{11/3}, \quad (2.20)$$

where Ω_0 is a reference orbital frequency and γ is the “chirp rate”.

While the equation of motion (2.20) can be solved exactly, $\Omega(t) = \Omega_0 (1 - 8\gamma t / (3\Omega_0))^{-3/8}$, it will be convenient to work on timescales shorter than Ω_0/γ and linearize this solution to

$$\Omega(t) \approx \Omega_0 + \gamma t, \quad (2.21)$$

so that $\varphi_*(t) \approx \pm(\Omega_0 + \frac{1}{2}\gamma t)t$. Note that the frequency “chirps,” and thus the two black holes merge at $t = \frac{3}{8}\Omega_0/\gamma$, so that this linear approximation is useful as long as the inspiral has not reached the merger phase (see Fig. 4).

The chirp rate γ is defined in (2.20) with respect to a reference frequency Ω_0 . Our primary interest in Sec. III and beyond, is in understanding how the cloud responds to the companion's gravitational perturbation when Ω_0 , or an integer multiple of it, matches the energy difference ΔE between an occupied bound state and one of the continuum

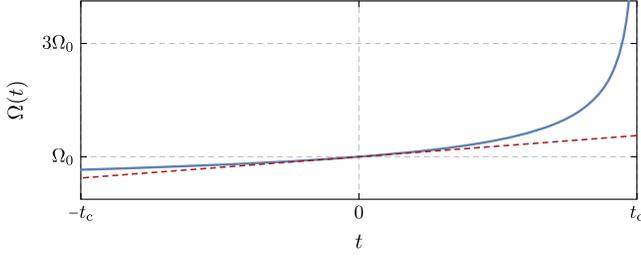


FIG. 4. Time dependence of the orbital frequency and its linear approximation.

bands. Throughout this paper, we will use γ to denote the chirp rate for the specific transition under consideration, with reference frequency $\Omega_0 = \Delta E$. This should be contrasted with the instantaneous chirp rate $\dot{\varphi}_*(t)$, which is equal to γ up to small corrections since the inspiral evolves very slowly. We justify this definition of the chirp rate γ in Appendix A 5.

C. Resonant transitions

In [16,31], it was shown that the companion's gravitational perturbation can force the cloud to transition from one bound state to another. We will briefly review these resonant transitions and establish a convenient notation.

Throughout this work, we will denote a generic bound state with a lower-case multi-index, e.g., $|a\rangle \equiv |n\ell m\rangle$. The matrix elements $\eta_{ab}(t) = \langle a|V_*(t)|b\rangle$ enable resonant transitions between different bound states when the orbital frequency satisfies a resonance condition. Because of the quasi-periodic nature of φ_* , we can decompose each of the matrix elements into their Fourier coefficients,

$$\eta_{ab}(t) = \sum_{m_\varphi \in \mathbb{Z}} \eta_{ab}^{(m_\varphi)}(t) e^{-im_\varphi \varphi_*(t)}, \quad (2.22)$$

where the functions $\eta_{ab}^{(m_\varphi)}(t)$ are slowly varying in time. Since both $|a\rangle$ and $|b\rangle$ have definite angular momentum—say m_a and m_b , respectively—the coupling oscillates with a definite frequency

$$\eta_{ab}(t) = e^{-i(m_a - m_b)\varphi_*(t)} \eta_{ab}^{(m_a - m_b)}(t). \quad (2.23)$$

When the oscillation frequency matches the energy difference between the two states,

$$(m_a - m_b)\Omega(t) = E_a - E_b, \quad (2.24)$$

the binary can resonate with the cloud and we expect that transitions between the two states will be enhanced [16]. Indeed, as the companion slowly moves through the resonance, the cloud is forced to transfer its population from one state to the other [31].

This process is the analog of the Landau-Zener transition in quantum mechanics [29,30]. The fraction of the cloud that is transferred from the initial state $|b\rangle$ to the final state $|a\rangle$ is controlled by the dimensionless Landau-Zener parameter $z_{ab} \equiv \eta_{ab}^2/\gamma$. Long after the transition, the total fraction of the cloud populating the state $|a\rangle$ is

$$|\langle a|\psi(\infty)\rangle|^2 = 1 - e^{-2\pi z_{ab}}. \quad (2.25)$$

There are two limiting behaviors of these transitions. For $z_{ab} \gg 1$, the transition is *adiabatic* and the cloud is transferred almost entirely from $|b\rangle$ to $|a\rangle$. On the other hand, for $z_{ab} \ll 1$, the transition is *nonadiabatic*, in which case the system is driven too quickly for it to respond and almost none of the cloud is transferred from $|b\rangle$ to $|a\rangle$.

During these resonant transitions, the cloud's angular momentum changes macroscopically, which must be compensated for by the binary's orbital angular momentum, i.e., the cloud *backreacts* significantly on the orbital dynamics. If the orbit *gains* angular momentum during this process, it can almost completely balance the angular momentum lost due to GW emission and cause the companion to *float*, temporarily slowing down the inspiral until the transition is completed. On the other hand, if the orbit *loses* angular momentum, then the orbit will *sink*, speeding up the inspiral temporarily. Both types of transitions impart a characteristic signature on the GW signal coming from the inspiral that can be used to detect the presence of a cloud.

In [31], it was shown that multiple of these transitions occur during the inspiral, leading to a characteristic fingerprint for the cloud that can be used to unambiguously determine the mass and spin of the ultralight boson. This tree of transitions ends when the orbital frequency Ω becomes too large and the resonance condition (2.24) between bound states can no longer be satisfied. However, the orbital frequency can then be high enough to ionize the cloud, unbinding it from its parent black hole. Indeed, this process occurs throughout the inspiral and so we will need to understand it, and its backreaction on the orbit, in order to fully characterize the phenomenology of these cloud-binary systems. This ionization process is the subject of the next section.

III. IONIZATION: EXCITING UNBOUND STATES

We will now study transitions between bound and unbound states of the gravitational atom, induced by the gravitational perturbation of the companion (see Fig. 5).² Since the analysis is somewhat technical, we will start with

²In principle, the companion can also mediate transitions from one continuum state to another. In this paper, we will ignore these, as we will only be concerned with the leading-order evolution of the cloud that remains bound to the parent black hole. We will justify this approximation in Sec. A 2.

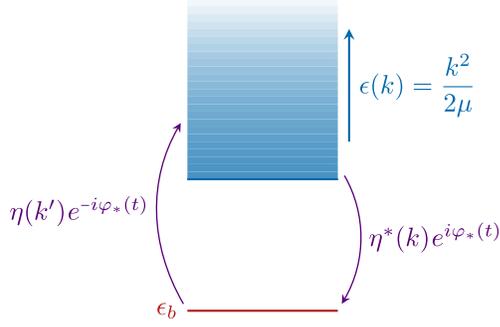


FIG. 5. Schematic illustration of the transitions between a bound state and the continuum.

a simple toy model involving a single bound state interacting with the continuum. We ignore the interactions between the semi-infinite number of continuum states and also neglect the angular momentum of the continuum states. This simplified model will capture the main dynamical features of the system without too many technical distractions. After we have gained intuition from this toy model, we will extend it to the real system of interest.

A. A toy model

Consider a single bound state $|b\rangle$, with energy $\epsilon_b < 0$, interacting with a semi-infinite continuum of states $|k\rangle$. For simplicity, we will assume that the continuum states depend only on the wave number k , with dispersion relation $\epsilon(k) = k^2/2\mu$, and that they do not interact with one another. We will also assume that the interaction between the bound state and the continuum oscillates at a frequency $\dot{\varphi}_*(t)$ that grows slowly and linearly in time, $\ddot{\varphi}_*(t) = \gamma$. This is the simplest generalization of the familiar two-state Landau-Zener system to include the coupling to the continuum. Despite its simplicity, this toy model will illustrate many of the phenomena we will encounter in the more realistic scenario.

The Hamiltonian of our toy model is³

$$\mathcal{H} = \epsilon_b |b\rangle\langle b| + \frac{1}{2\pi} \int_0^\infty dk [\eta(k) e^{-i\varphi_*(t)} |k\rangle\langle b| + \eta^*(k) e^{i\varphi_*(t)} |b\rangle\langle k| + \epsilon(k) |k\rangle\langle k|]. \quad (3.1)$$

As in Sec. II, the continuum states are normalized such that $\langle k|k'\rangle = 2\pi\delta(k-k')$, while the phase is $\varphi_*(t) = \varphi_0 + \Omega_0 t + \gamma t^2/2$. A general state in the Hilbert space can be written as

³This is an extension of the Demkov-Osherov model [34] to a single bound state interacting with a semi-infinite continuum. A similar model was studied in [35], but with a different focus and using different techniques.

$$|\psi\rangle = c_b(t) e^{-i\epsilon_b t} |b\rangle + \frac{1}{2\pi} \int_0^\infty dk c_k(t) e^{-i\epsilon(k)t} |k\rangle, \quad (3.2)$$

where we have peeled off the standard oscillatory behavior caused by the nonzero energies of each state—this will help us isolate the effect of the interactions $\eta(k)$. The Schrödinger equation associated to the Hamiltonian (3.1) leads to the equations of motion

$$i \frac{dc_b}{dt} = \frac{1}{2\pi} \int_0^\infty dk \eta^*(k) e^{i\varphi_*(t) + i(\epsilon_b - \epsilon(k))t} c_k(t), \quad (3.3)$$

$$i \frac{dc_k}{dt} = \eta(k) e^{-i\varphi_*(t) + i(\epsilon(k) - \epsilon_b)t} c_b(t). \quad (3.4)$$

Our goal is to “integrate out” the continuum to find an approximate description of the system entirely in terms of the bound state’s dynamics. We do so using the so-called Weisskopf-Wigner method; see e.g., [36–38].

Assuming that the system begins its life in the bound state, $c_k(t) \rightarrow 0$ as $t \rightarrow -\infty$, for all k , we can solve (3.4) exactly,

$$c_k(t) = -i \int_{-\infty}^t dt' \eta(k) e^{i(\epsilon(k) - \epsilon_b)t' - i\varphi_*(t')} c_b(t'). \quad (3.5)$$

Substituting this into (3.3), we find an (integro-differential) equation for the dynamics of the entire system purely in terms of the bound state amplitude,

$$i \frac{dc_b}{dt} = \int_{-\infty}^t dt' \Sigma_b(t, t') c_b(t'), \quad (3.6)$$

where we have defined the *self-energy*

$$\Sigma_b(t, t') \equiv \frac{1}{2\pi i} \int_0^\infty dk |\eta(k)|^2 e^{i(\varphi_*(t) - \varphi_*(t')) - i(\epsilon(k) - \epsilon_b)(t - t')}. \quad (3.7)$$

This equation of motion is still quite complicated, but we can make significant progress via the *Markov approximation* [38], wherein we integrate by parts and drop the remainder term. The bound state Schrödinger equation then simplifies to

$$i \frac{dc_b}{dt} = \mathcal{E}_b(t) c_b(t), \quad (3.8)$$

where we have introduced the *induced energy*

$$\begin{aligned} \mathcal{E}_b(t) &= \int_{-\infty}^t dt' \Sigma_b(t, t') \\ &= \frac{1}{2\pi i} \int_{-\infty}^t dt' \int_0^\infty dk |\eta(k)|^2 e^{i(\varphi_*(t) - \varphi_*(t')) - i(\epsilon(k) - \epsilon_b)(t - t')}. \end{aligned} \quad (3.9)$$

As we discuss in Appendix B, this approximation consists of dropping terms that are higher order in $\mathcal{E}_b(t)$ and its time integrals. The imaginary part of the induced energy completely determines the behavior of the bound state occupation probability, which may be approximated as

$$\frac{d \log |c_b(t)|^2}{dt} = 2 \text{Im} \mathcal{E}_b(t) \approx -\frac{\mu |\eta(k_*(t))|^2}{k_*(t)} \Theta(k_*^2(t)), \quad (3.10)$$

where $k_*(t) = \sqrt{2\mu(\dot{\varphi}_*(t) + \epsilon_b)}$ and $\Theta(x)$ is the Heaviside function, with $\Theta(k_*^2(t)) = \Theta(\dot{\varphi}_*(t) + \epsilon_b)$. We will devote the rest of this section to understanding the time dependence of $\text{Im} \mathcal{E}_b(t)$ and qualitatively justifying the approximation in (3.10).

As we might expect, the bound state only starts to significantly interact with the continuum when the frequency of the perturbation $\dot{\varphi}_*(t)$ is high enough to compensate for the bound state's binding energy, $-\epsilon_b$. This is when the bound state starts to “resonate” with the continuum and we can choose our time coordinate so that this resonance occurs at $t = 0$. This is not a resonance in the classic sense, but we find it useful to continue using this language. As illustrated in Fig. 6, the system (3.8) evolves on a time scale set by $\gamma^{-1/2}$ and its behavior can be divided into three distinct stages.

Far before the resonance, in the left-shaded region, where $\sqrt{\gamma}t \ll -1$, the perturbation cannot provide enough energy for the bound and continuum states to interact and so the population of the bound state is, to good approximation, completely unaffected by the presence of the continuum. This changes when $|\sqrt{\gamma}t| \lesssim 1$, in the unshaded transient region, where the system goes on resonance and develops a relatively complicated time dependence. We do

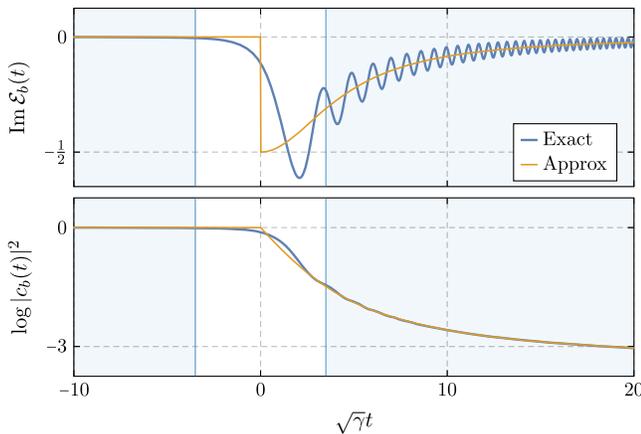


FIG. 6. The imaginary part of the induced energy $\mathcal{E}_b(t)$ (top) and the log occupation of the bound state $\log |c_b(t)|^2$ (bottom) as functions of dimensionless time $\sqrt{\gamma}t$, using both the exact expression (3.9) [blue] and our approximation (3.10) [orange]. Here, we assume that the bound-to-continuum couplings take the form $|\eta(k)|^2 = (k/\mu)/[1 + k^4/(81\mu^2\gamma)]$.

not need to fully understand this complicated stage, other than to note that this region serves to smoothly interpolate between the $\sqrt{\gamma}t \ll -1$ stage and the final $\sqrt{\gamma}t \gg 1$ stage.

In the right-shaded region, where $\sqrt{\gamma}t \gg 1$, the system approaches a type of steady state where the imaginary part of the induced energy $\mathcal{E}_b(t)$ is well approximated by two distinct behaviors. The first is a remaining transient oscillation whose amplitude decays in time and whose properties depend only on the behavior of the coupling $|\eta(k)|^2$ as $k \rightarrow 0$. As described in Appendix A, when $|\eta(k)|^2$ goes to zero linearly in k at the edge of the continuum, these oscillations decay as $(\sqrt{\gamma}t)^{-1}$, and thus their effect on the solution $\log |c_b(t)|^2$ decays as $(\sqrt{\gamma}t)^{-2}$. As illustrated in Fig. 6, these oscillations provide a subleading correction to the dominant behavior, which is a steady and smooth deoccupation of the cloud, whose instantaneous rate depends only on the properties of the continuum state that the system is currently “resonating” with, i.e., the continuum state whose energy is $\frac{1}{2\mu}k_*^2(t) = \dot{\varphi}_*(t) + \epsilon_b$. This dominant contribution (3.10) is the only one we will consider in the text.⁴

We will mostly be interested in the total energy that has been ionized by the perturbation, as a function of time. Assuming that the system only occupies the bound state in the far past, this ionized energy can be defined as the total energy contained within the continuum,

$$E_{\text{ion}}(t) \equiv \frac{1}{2\pi} \frac{M_c}{\mu} \int_0^\infty dk (\epsilon(k) - \epsilon_b) |c_k(t)|^2. \quad (3.11)$$

As we describe in Appendix C, the rate at which energy is ionized dE_{ion}/dt , which we will call the *ionization power* P_{ion} , can be expressed in a particularly simple form by again working in the Markov approximation and ignoring subleading transient contributions,

$$P_{\text{ion}}(t) \approx \frac{M_c}{\mu} \left[\dot{\varphi}_*(t) \frac{\mu |\eta(k_*(t))|^2}{k_*(t)} \right] \Theta(k_*^2(t)) |c_b(t)|^2. \quad (3.12)$$

This is clearly evocative of (3.10) and enjoys a simple interpretation; how quickly the ionized energy grows is equal to the rate at which the bound state “resonates” into the state at $k_*(t)$, namely $\mu |\eta(k_*(t))|^2/k_*(t)$, weighted both by the energy difference $\epsilon(k_*(t)) - \epsilon_b = \dot{\varphi}_*(t)$ between these states and by how much is still left in the bound state at that time, $(M_c/\mu) |c_b(t)|^2$.

Perhaps the most striking phenomenon we will encounter is the appearance of seemingly discontinuous jumps in the ionization power. We will find that these jumps occur when the perturbation begins to resonate with the

⁴As we explain in Appendix A, we can also derive the deoccupation rate (3.10) using stationary perturbation theory, $\gamma = 0$, and assuming that the obtained answer holds for $\gamma \neq 0$, if the frequency evolves slow enough.

continuum—that is, when the perturbation’s frequency is just enough to excite the bound state into the very edge of the continuum. These discontinuities are apparent in our approximation of the time evolution (3.10), shown in Fig. 6, and are ultimately due to the behavior of the continuum wave functions as $k \rightarrow 0$. As we explain in Appendix D, the long-range nature of the r^{-1} potential localizes this zero mode to a Bohr radius-sized region around $r = 0$ and, by a matching argument, this implies that the wave function’s normalization scales like \sqrt{k} as $k \rightarrow 0$, as do all matrix elements between the bound and continuum states. The combination $\mu|\eta(k_*(t))|^2/k_*(t)$ thus approaches a *finite* limit for $k_*(t) \rightarrow 0$, when the bound state begins to resonate with the continuum, leading to an apparent discontinuity in our approximation (3.10). Said differently, the coupling per unit energy $|\eta(\epsilon)|^2 = dk(\epsilon)/d\epsilon|\eta(k(\epsilon))|^2$ is finite in the zero-energy limit because the zero-energy modes are still localized about the origin. Of course, this approximation does not capture the transient region shown in Fig. 6, which smooths out these apparent discontinuities on a timescale $\gamma^{-1/2}$.

It is instructive to compare the timescale of the transition, $\gamma^{-1/2}$, to the characteristic timescale of the inspiral, Ω_0/γ , which measures how long it takes for the separation between the two black holes to change by a $\mathcal{O}(1)$ fraction. Using the definition of γ in (2.20), the ratio of the two timescales is⁵

$$\frac{\gamma^{-1/2}}{\Omega_0/\gamma} = \sqrt{\frac{96}{5}} \frac{q^{1/2}}{(1+q)^{1/6}} \left(\frac{\alpha\Omega_0}{\mu}\right)^{5/6} \propto \sqrt{q\alpha^3}, \quad (3.13)$$

where we used that the transitions occur for $\Omega_0 \sim \mu\alpha^2$ to get the scaling in the final equality. For small q and α , the transitions therefore are very fast on the timescale of the inspiral.

B. The realistic case

Conceptually, extending our analysis to the realistic case of the gravitational atom requires very little extra work beyond what we have already done, the main complication being that there are simply many more states to keep track of. Our goal is again to integrate out the continuum states and encode their effects on the bound states in terms of a set of induced energies and couplings, analogous to (3.9). These effective interactions will be relatively complicated functions of time, but will contain a simple “steady-state” behavior similar to (3.10).

⁵Here, we have ignored the backreaction of ionization on the binary’s dynamics, which can increase the effective chirp rate $\dot{\varphi}_*(t) \approx \gamma$ by a factor of $\mathcal{O}(100)$. This changes the estimate (3.13), which scales as $\gamma^{1/2}$, by an $\mathcal{O}(10)$ factor. However, for the values of q and α we consider, this does not change the fact that these transitions are fast.

We can write the Hamiltonian of the gravitational atom as

$$\mathcal{H} = \sum_b \epsilon_b(t) |b\rangle\langle b| + \sum_{a \neq b} \eta_{ab}(t) |a\rangle\langle b| + \sum_K \epsilon_K(t) |K\rangle\langle K| + \sum_{K,b} [\eta_{Kb}(t) |K\rangle\langle b| + \text{H.c.}], \quad (3.14)$$

where we use a, b, \dots as a bound state multi-index,⁶ $|a\rangle \equiv |n_a \ell_a m_a\rangle$ and $|b\rangle = |n_b \ell_b m_b\rangle$, while K, L, \dots is a continuum state multi-index, $|K\rangle \equiv |k; \ell m\rangle$. We take $\epsilon_b(t)$, $\epsilon_K(t)$, $\eta_{ab}(t)$, and $\eta_{Kb}(t)$ as shorthands for the energies and couplings $\epsilon_{n_b \ell_b m_b}(t)$, $\epsilon_{\ell m}(k; t)$, $\eta_{n_a \ell_a m_a | n_b \ell_b m_b}(t)$, and $\eta_{k; \ell m | n_b \ell_b m_b}(t)$, respectively. Sums over multi-indices should be understood to include a sum over *all* states of a given type. For instance, the analog of (3.2) for a generic state is⁷

$$|\psi\rangle = \sum_b c_b(t) e^{-i\epsilon_b t} |b\rangle + \sum_K c_K(t) e^{-i\epsilon_K t} |K\rangle = \sum_{n, \ell, m} c_{n\ell m}(t) e^{-iE_{n\ell m} t} |n\ell m\rangle + \frac{1}{2\pi} \sum_{\ell, m} \int_0^\infty dk c_{k; \ell m}(t) e^{-i\epsilon(k)t} |k; \ell m\rangle, \quad (3.15)$$

where $E_{n\ell m}$ and $\epsilon(k)$ are defined in (2.8) and (2.14), respectively.

In this abbreviated notation, the coefficients obey the following equations of motion

$$i \frac{dc_b}{dt} = \sum_{a \neq b} \eta_{ba}(t) c_a(t) e^{i(\epsilon_b - \epsilon_a)t} + \sum_K \eta_{bK}(t) c_K(t) e^{i(\epsilon_b - \epsilon_K)t}, \quad (3.16)$$

$$i \frac{dc_K}{dt} = \sum_a \eta_{Ka}(t) c_a(t) e^{i(\epsilon_K - \epsilon_a)t}. \quad (3.17)$$

Assuming that the continuum states are completely deoccupied in the far past, $t \rightarrow -\infty$, we can solve (3.17) exactly,

⁶In the previous subsection we used the subscript b to denote “bound state” whereas now we use it as a bound state *index*, slightly abusing notation.

⁷Since the energies $\epsilon_b(t)$ and $\epsilon_K(t)$ depend on time, the appropriate “integrating factor” in this ansatz should be $\exp(-i \int dt' \epsilon_b(t'))$ instead of $\exp(-i\epsilon_b t)$, etc. However, the time dependence of these energies is extremely suppressed, $\dot{\epsilon}_b \sim \mathcal{O}(\gamma(q\alpha)^2)$, since it only arises from the radial dynamics of the companion. Such time-dependent terms are not critical to the resonant effects we discuss in this section, and only provide very small corrections to details like the time at which the resonance begins. We will thus ignore them.

$$c_K(t) = -i \int_{-\infty}^t dt' \left[\sum_a \eta_{Ka}(t') c_a(t') e^{i(\epsilon_K - \epsilon_a)t'} \right]. \quad (3.18)$$

Substituting this into (3.16) yields an integro-differential equation purely in terms of the bound states

$$i \frac{dc_b}{dt} = \sum_{a \neq b} \eta_{ba}(t) c_a(t) e^{i(\epsilon_b - \epsilon_a)t} + \sum_a \int_{-\infty}^t dt' \Sigma_{ba}(t, t') c_a(t'), \quad (3.19)$$

where we have defined the self-energies

$$\Sigma_{ba}(t, t') = -i \sum_K \eta_{bK}(t) \eta_{Ka}(t') e^{i(\epsilon_b - \epsilon_K)t + i(\epsilon_K - \epsilon_a)t'}, \quad (3.20)$$

which generalize (3.6) to include multiple bound states. The main complication, compared to the toy model presented in Sec. III A, is that the continuum can mediate transitions between different bound states, and will thus induce off-diagonal couplings.

Again working in the Markov approximation, we can rewrite (3.19) as an effective Schrödinger equation for the bound states

$$i \frac{dc_b}{dt} = \mathcal{E}_b(t) c_b(t) + \sum_{a \neq b} [\eta_{ba}(t) e^{i(\epsilon_b - \epsilon_a)t} + \mathcal{E}_{ba}(t)] c_a(t), \quad (3.21)$$

where we have defined both the *induced couplings*

$$\mathcal{E}_{ba}(t) = -i \int_{-\infty}^t dt' \sum_K \eta_{bK}(t) \eta_{Ka}(t') e^{i(\epsilon_b - \epsilon_K)t + i(\epsilon_K - \epsilon_a)t'} \quad (3.22)$$

and the *induced energies* $\mathcal{E}_b(t) = \mathcal{E}_{bb}(t)$, the realistic analog of (3.9). As before, we have reduced the complicated problem of bound states interacting with a continuum to the analysis of a set of (complicated) time-dependent functions $\mathcal{E}_{ba}(t)$.

These induced couplings take a much simpler form when we remember that both the bound and continuum states have definite azimuthal angular momentum, which we will denote as m for the continuum state K and m_a or m_b for the bound states $|a\rangle$ or $|b\rangle$, respectively. Since the couplings between the bound and continuum states $\eta_{Ka}(t)$ reduce to a single Floquet component (2.23), we can write the induced couplings appearing in (3.21) as⁸

⁸The Floquet components $\eta_{Ka}^{(m-m_a)}$ inherit their time dependence purely from the radial motion of the companion. Though this slow radial motion is extremely important when it forces the frequency of the perturbation to slowly increase in time and cannot be ignored there, taking the adiabatic approximation $\eta_{Ka}^{(m-m_a)}(t') \approx \eta_{Ka}^{(m-m_b)}(t)$ only requires dropping subleading terms of $\mathcal{O}(\gamma)$, and so we will use this approximation throughout.

$$\begin{aligned} \mathcal{E}_{ba}(t) = & -i \int_{-\infty}^t dt' \sum_K \eta_{Kb}^{*(m-m_b)}(t') \eta_{Ka}^{(m-m_a)}(t) \\ & \times e^{i(m-m_b)\varphi_*(t) - i(m-m_a)\varphi_*(t') + i(\epsilon_b - \epsilon_K)t + i(\epsilon_K - \epsilon_a)t'}. \end{aligned} \quad (3.23)$$

As we argue in Appendix A, the off-diagonal terms oscillate as $\mathcal{E}_{ba}(t) \propto e^{i(\epsilon_b - \epsilon_a)t - i(m_b - m_a)\varphi_*(t)}$, just like the directly mediated transitions between the bound states

$$\eta_{ba}(t) e^{i(\epsilon_b - \epsilon_a)t} = \eta_{ba}^{(m_b - m_a)}(t) e^{i(\epsilon_b - \epsilon_a)t - i(m_b - m_a)\varphi_*(t)}. \quad (3.24)$$

The total coupling between the bound states $|a\rangle$ and $|b\rangle$, $\eta_{ba}(t) e^{i(\epsilon_b - \epsilon_a)t} + \mathcal{E}_{ba}(t)$, thus oscillates extremely rapidly unless the argument of this exponential becomes stationary, which occurs when

$$(m_b - m_a) \dot{\varphi}_*(t) = \epsilon_b - \epsilon_a. \quad (3.25)$$

This is exactly the resonance condition (2.24) and so, even including the effects of the continuum, we can ignore transitions between bound states as long the system is not actively in resonance, cf. [31]. That is, away from resonances the coupling between $|a\rangle$ and $|b\rangle$ oscillates rapidly enough so as to effectively average out to zero. Ignoring these resonances, we can thus dramatically simplify the effective Schrödinger equation (3.21) to

$$i \frac{dc_b}{dt} = \mathcal{E}_b(t) c_b(t), \quad (3.26)$$

where the induced energies,

$$\begin{aligned} \mathcal{E}_b(t) = & -i \sum_K \int_{-\infty}^t dt' |\eta_{Kb}^{(m-m_b)}(t')|^2 \\ & \times e^{i(m-m_b)(\varphi_*(t) - \varphi_*(t')) - i(\epsilon_K - \epsilon_b)(t - t')}, \end{aligned} \quad (3.27)$$

are simply the generalization of (3.9) to include continuum states with different angular momenta.

The dynamics of this effective Schrödinger equation are very similar to those of the toy model. Assuming that the system occupies a single bound state and ignoring the transient oscillations as we discussed in Sec. III A, we may write the analog of (3.10) as

$$\frac{d \log |c_b(t)|^2}{dt} = 2 \text{Im} \mathcal{E}_b(t) \approx - \sum_{\ell, g} \left[\frac{\mu |\eta_{K_* b}^{(g)}(t)|^2}{k_*^{(g)}(t)} \Theta(k_*^{(g)}(t)^2) \right], \quad (3.28)$$

with $K_* = \{k_*^{(g)}(t), \ell, m = g + m_b\}$ and $k_*^{(g)}(t) = \sqrt{2\mu(g\dot{\varphi}_*(t) + \epsilon_b)}$, where the sum ranges from $\ell = 0, 1, \dots, \infty$ and over all g such that $|g + m_b| \leq \ell$. As before, the instantaneous rate of deoccupation only relies

on the properties of the state that the system currently “resonates” with. However, in contrast to our toy model, there are two main complications. First, the perturbation oscillates at every overtone $g \in \mathbb{Z}$ of the base frequency $\dot{\varphi}_*(t)$. Second, the continuum state with energy $\frac{1}{2\mu}k_*^2(t) = g\dot{\varphi}_* + \epsilon_b$ is infinitely degenerate. The selection rule (2.23) kills the sum over overtones, but we still need to account for this infinite degeneracy, leading to the sum over total and azimuthal orbital angular momentum.

The same simplifications apply to the ionization power, which we may write as

$$P_{\text{ion}} \equiv \frac{dE_{\text{ion}}}{dt} \approx \sum_{\ell, g} \frac{M_c}{\mu} \left[g\dot{\varphi}_*(t) \frac{\mu |\eta_{K_* b}^{(g)}(t)|^2}{k_*^{(g)}(t)} \right] \Theta(k_*^{(g)}(t)^2) |c_b(t)|^2, \quad (3.29)$$

assuming that the system initially only occupies one bound state $|b\rangle$, where the sum is again over all states that can participate in the resonance. If the system occupies multiple bound states, we can approximate the ionization power by summing (3.29) over each occupied state.

We plot this ionization power as a function of the binary separation R_* in Fig. 7. We show this ionization power normalized arbitrarily (*top*) and by the energy lost due to gravitational wave emission $P_{\text{GW}} \equiv dE_{\text{GW}}/dt$ (*bottom*), ignoring cloud depletion $|c_b(t)|^2 = 1$, for both corotating and counter-rotating orbits. As we explained in the previous subsection, the discontinuous jumps that appear in both panels are due to the bound state beginning to resonate with the continuum and the fact that all couplings $|\eta_{Kb}|^2$ are $\propto k$ as $k \rightarrow 0$. The fact that the perturbation now has multiple overtones means that this resonance can occur at multiple points in the orbit. Specifically, for a cloud whose initial state is $|n_b \ell_b m_b\rangle$, these discontinuities will appear at the orbital separations

$$\frac{R_*^{(g)}}{M} = \alpha^{-2} [4g^2(1+q)n_b^4]^{1/3}, \quad g = 1, 2, \dots, \quad (3.30)$$

though they become progressively weaker for higher overtones g . From the bottom panel, we also see that ionization is a *large* effect compared to the energy loss due to GW emission; for $M_c = 0.01M$, P_{ion} can be two orders of magnitude larger than P_{GW} . To understand this more intuitively, we note that the cloud’s binding energy per unit mass, $\alpha^2/(2n^2)$, is comparable to the same quantity for the binary, $M/(2R_*)$, when $R_* \sim r_c$. If ionization reduces the cloud’s mass by an amount of order the companion’s mass M_* , this will therefore cause a large backreaction on the orbit. We confirm this intuitive expectation numerically in Sec. V.

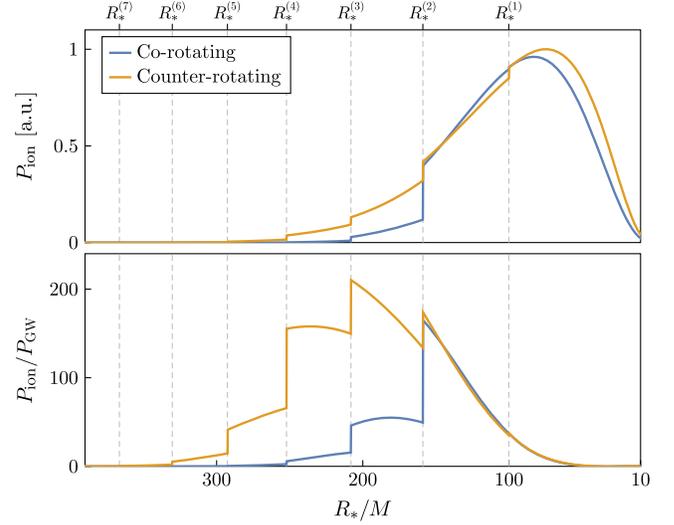


FIG. 7. The ionization power (3.29) as a function of the binary separation R_* , for $\alpha = 0.2$, $q = 10^{-3}$, $M_c = 0.01M$, and a cloud in the $|211\rangle$ state. We ignore both cloud depletion and the backreaction on the orbital dynamics (see Sec. V). In the top panel, we normalize both curves by the peak ionization power of the counter-rotating orbit; so-called arbitrary units. In the bottom panel, we have normalized each curve by P_{GW} , the energy lost due to GW emission (2.19). We see that the energy lost due to ionization, whose overall amplitude is controlled by the mass of the cloud $M_c |c_b(t)|^2$, can dominate over GW emission.

It is worth noting that, for small q , the curves shown in Fig. 7 exhibit a universal scaling behavior. The radial wave functions $R_{n\ell}(r)$ and $R_{k;\ell}(r)$, given in (2.5) and (2.12), only depend on the dimensionless variables $r/r_c = \alpha^2 r/M$ and kr , respectively. The wavelength k_* appearing in (3.28) and (3.29) is also a function of r/r_c that scales as α^2 and is independent of q , when $q \ll 1$. Because the matrix elements $|\eta_{K_* b}^{(g)}(t)|^2$ are evaluated at k_* , every radial variable in the overlap integrals will therefore appear in the combination $\alpha^2 r/M$. The overlaps themselves thus also inherit a homogeneous α -scaling, which can be found by power counting. For the ionization power and the deoccupation rate, this leads to

$$P_{\text{ion}} = \alpha^5 q^2 \frac{M_c}{M} \mathcal{P}(\alpha^2 R_*/M), \quad (3.31)$$

$$\frac{d \log |c_b(t)|^2}{dt} = \frac{\alpha^3 q^2}{M} \mathcal{R}(\alpha^2 R_*/M), \quad (3.32)$$

where \mathcal{P} and \mathcal{R} are universal functions for each bound state $|n_b \ell_b m_b\rangle$ that have to be found numerically. These relations are particularly useful when results are needed for many points in parameter space, as we now only need to compute the relatively complicated functions \mathcal{P} and \mathcal{R} once for a fiducial set of parameters.

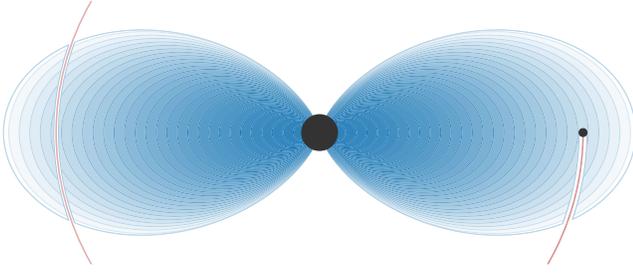


FIG. 8. Cartoon illustrating the accretion of the boson cloud by the companion black hole. As explained in the text, the cloud will respond rapidly and replenish the local density behind the companion.

IV. ACCRETION: ABSORBING THE CLOUD

So far, we have treated the perturbing object as pointlike and studied only its gravitational influence on the cloud. In this section, we will take the finite size of the companion into account and compute its absorption of the cloud (see Fig. 8).⁹ If the secondary object is a black hole of mass M_* and spin a_* , then this absorption will play an important role in the binary's dynamics.

A. Motion in a uniform medium

We start by solving the problem in the idealized case of a black hole moving with a constant velocity in a medium with a uniform density ρ . If the medium were made of small particles at rest at infinity, the problem would be relatively straightforward to solve via geodesic motion in the rest frame of the black hole. In the Schwarzschild case, the energy flux takes the form [43]

$$\frac{dM_*}{dt} = \frac{\pi\rho M_*^2 (4v^2 + \sqrt{8v^2 + 1} - 1)^3}{2v^3 (\sqrt{8v^2 + 1} - 1)^2} \sim \frac{4\pi\rho(2M_*)^2}{v}, \quad (4.1)$$

$v \rightarrow 0,$

where v is the asymptotic value of the relative velocity between the particles and the black hole. The divergence at $v \rightarrow 0$ signals the nonexistence of a stationary configuration with $v = 0$ where the density of the medium approaches a finite nonzero value at infinity.

In the case of interest, the Compton wavelength of the medium is much larger than the gravitational radius, $r_{g,*} = M_*$, and therefore (4.1) does not hold. We expect the true answer to be smaller because the quantum pressure of the field suppresses small-scale overdensities. Because of the relative motion, the black hole will see the scalar field

⁹The absorption cross section of a scalar field by a black hole has been studied extensively; in the massless case for rotating black holes in [39–42], in the massive case for Schwarzschild black holes in [43], and more recently, in the massive case for charged and/or rotating black holes in [44,45]. Our analysis will be similar to that in [43,45].

as having a wave number $k \sim \mu v$. Besides the (reduced) Compton wavelength, $\lambda_C = \mu^{-1}$, the other relevant scale in the problem is then the (reduced) de Broglie wavelength, $\lambda_{dB} = k^{-1}$. It will also be useful to define the dimensionless ratios $r_{g,*}/\lambda_C = \mu M_*$ and $\lambda_C/\lambda_{dB} = k/\mu$. We are interested in the limit where both of these ratios are small,

$$\begin{aligned} \mu M_* &\ll 1 \quad (\text{“fuzzy”}), \\ k/\mu &\ll 1 \quad (\text{“nonrelativistic”}). \end{aligned} \quad (4.2)$$

We will see, in Sec. IV B, why these are the relevant limits in the realistic setting.

Our goal is to compute the radial energy flux at the outer horizon $r = r_+$,

$$\frac{dM_*}{dt} = \int d\theta d\phi \sqrt{g_{\theta\theta}g_{\phi\phi}} T^r{}_0(r_+), \quad (4.3)$$

where the energy-momentum tensor $T_{\mu\nu}$ is that of the field profile $\Phi(t, \mathbf{r})$. Expanding this profile in modes with definite frequency $\omega^2 = \mu^2 + k^2$, we have (cf. Appendix E)

$$\Phi(t, \mathbf{r}) = \sum_{\ell_*, m_*} R_{k; \ell_*, m_*}(r) S_{\ell_*, m_*}(ka_*; \cos\theta) e^{-i\omega t + im_*\phi}, \quad (4.4)$$

where $S_{\ell_*, m_*}(c; \cos\theta)$ are spheroidal harmonics with spheroidicity c , we can write the radial energy flux associated to this profile as

$$\begin{aligned} T^r{}_0 &= \frac{2\omega(r - r_+)(r - r_-)}{r^2 + a^2 \cos^2\theta} \\ &\times \sum_{\ell_*, m_*} \text{Im}(\partial_r R_{\ell_*, m_*}^* R_{\ell_*, m_*}) |S_{\ell_*, m_*}|^2 + \dots, \end{aligned} \quad (4.5)$$

where the ellipses represent terms that mix different angular momenta and will vanish when integrated in (4.3) to compute the radial energy flux. We denote the angular momentum quantum numbers measured with respect to the companion's position as ℓ_* and m_* , to distinguish them from those measured with respect to the parent black hole.

The presence of the black hole deforms the field profile and determines its shape at the horizon, and thus the flux, as function of the boundary conditions at large distances. We work in the rest frame of the black hole and consider an incident monochromatic plane wave from infinity with wave vector \mathbf{k} . In Minkowski spacetime, the asymptotic field profile would be

$$\begin{aligned} \Phi(t, \mathbf{r}) &\sim \sqrt{\frac{\rho}{2\omega^2}} e^{i\mathbf{k}\cdot\mathbf{r}} e^{-i\omega t} \\ &= \sqrt{\frac{\rho}{2\omega^2}} \sum_{\ell_*=0}^{\infty} (2\ell_* + 1) i^{\ell_*} j_{\ell_*}(kr) P_{\ell_*}(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) e^{-i\omega t}, \\ r/M_* &\rightarrow \infty, \end{aligned} \quad (4.6)$$

where $\omega = \sqrt{\mu^2 + k^2}$, with $k = \mu v / \sqrt{1 - v^2}$. In this expression, $j_{\ell_*}(kr)$ is the spherical Bessel function, $P_{\ell_*}(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}})$ is the Legendre polynomial and the normalization has been chosen so that $\rho \approx T_{00} = 2\omega^2 \Phi^* \Phi$. The long-range nature of the gravitational field, however, deforms the field; in a spherically symmetric spacetime, we have [44]

$$\Phi(t, \mathbf{r}) \sim \sqrt{\frac{\rho}{2\omega^2}} \sum_{\ell_*=0}^{\infty} (2\ell_* + 1) i^{\ell_*} j_{\ell_*}(kr + \delta(r)) \times P_{\ell_*}(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) e^{-i\omega t}, \quad r/M_* \rightarrow \infty, \quad (4.7)$$

where $\delta(r) = kM_*(1 + \omega^2/k^2) \log(2kr) + \delta_{\ell_*}$, and δ_{ℓ_*} is a constant phase shift. Although our case is not quite spherically symmetric, deviations from (4.7) are controlled by the spheroidicity parameter, which is $ka_* \ll 1$ in the nonrelativistic limit we are considering.

To compute the energy flux at the horizon, we must understand the dependence of the near-field solution on the boundary condition (4.7). This will be achieved by a *matched asymptotic expansion*; the far-field and near-field solutions will be studied separately and matched in the overlap region, where both expansions hold. The boundary condition will then fix the overall amplitude of the solution. This procedure is schematically illustrated in Fig. 9.

Near-field solution—With the ansatz (4.4), the Klein-Gordon equation is separable. The exact solution of the equation for $R_{k;\ell_*,m_*}(r)$ can be expressed in terms of the confluent Heun function (see Appendix E and [46]). We expect the contributions from modes with $\ell_* \geq 1$ to be suppressed at radii smaller than about $\ell_*^2/(\mu^2 M_*)$ (due to the angular momentum barrier), so that the $\ell_* = m_* = 0$ mode dominates near the horizon. Expanding the confluent Heun function around $r = r_+$, one can show that

$$R_k(r) = C_k e^{-i\omega(\tilde{r}-r) - im_*\tilde{\phi}} (1 + \mathcal{O}(\mu M_*, kM_*)), \quad \text{for } r_+ \leq r < r_{\max}, \quad (4.8)$$

where we use $R_k(r) = R_{k;0,0}(r)$ as a shorthand, the coefficient $C_k = C_{k;0,0}$ defines the near-horizon amplitude of the $\ell_* = m_* = 0$ mode, \tilde{r} and $\tilde{\phi}$ are the radial and angular tortoise coordinates (defined in Appendix E), and the breakdown of the expansion is at

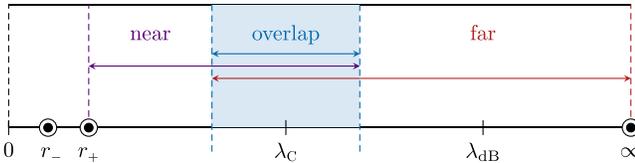


FIG. 9. Schematic illustration of the near-field and far-field expansions, where r_{\pm} are the inner and outer horizons of the black hole. The two asymptotic solutions are matched in the overlap region, $M_* \ll r \ll 1/k$.

$$\frac{r_{\max}}{M_*} \sim \min \left\{ \frac{1}{(\mu M_*)^2}, \frac{1}{kM_*} \right\} \gg 1. \quad (4.9)$$

Using the explicit expressions of the tortoise coordinates, and plugging (4.8) into (4.3), we get

$$\frac{dM_*}{dt} = 4M_* r_+ \omega^2 |C_k|^2. \quad (4.10)$$

We will now determine C_k by matching (4.8) to the far-field solution.

Far-field solution—Far from the companion, $r \gg M_*$, the equation for $R_k(r)$ becomes

$$\frac{d^2 R_k}{dr^2} + \left(\frac{2}{r} + \dots \right) \frac{dR_k}{dr} + \left(k^2 + \frac{2M_*(\omega^2 + k^2)}{r} + \dots \right) R_k = 0. \quad (4.11)$$

This equation is solved by a linear combination of confluent hypergeometric functions,

$$e^{ikr} R_k = C_{F1} F_1(1 + ikM_*(1 + \omega^2/k^2); 2; 2ikr) + C_U U(1 + ikM_*(1 + \omega^2/k^2); 2; 2ikr). \quad (4.12)$$

For $kr \ll 1$, this solution overlaps with the near-field solution (4.8). Expanding (4.12) in this limit and matching to (4.8) then gives $C_F = C_k$ and $C_U \leq \mathcal{O}((\mu M_*)^2)$. To determine the overall amplitude of the solution, we then expand (4.12) for $kr \gg 1$, where it reduces to a spherical Bessel function, $R_k(r) \propto j_0(kr + \delta(r))$, and compare it to the $\ell_* = 0$ mode of the boundary condition (4.7). This gives

$$C_F = C_k = \frac{\sqrt{2\pi\rho}}{\omega} \left| \Gamma(1 + ikM_*(1 + \omega^2/k^2)) e^{\frac{1}{2}\pi kM_*(1 + \omega^2/k^2)} \right|. \quad (4.13)$$

Plugging this back into (4.10), we get

$$\frac{dM_*}{dt} = \mathcal{A}_* \rho |\Gamma(1 + ikM_*(1 + \omega^2/k^2))|^2 e^{\pi kM_*(1 + \omega^2/k^2)}, \quad (4.14)$$

where $\mathcal{A}_* \equiv 8\pi M_* r_{+,*}$ is the area of the outer horizon of the Kerr black hole. This is our final answer for the mass accretion rate.

The result is shown in Fig. 10 for $2\mu M_* = 10^{-4}$. As anticipated, the flux is smaller than for particles, but still divergent for $v \rightarrow 0$. For nonrelativistic momenta, $k < \mu$, we can identify two different regimes

$$\frac{dM_*}{dt} = \mathcal{A}_* \rho \begin{cases} 1 & \text{for } k \gg 2\pi\mu^2 M_*, \\ \frac{2\pi\mu^2 M_*}{k} & \text{for } k \ll 2\pi\mu^2 M_*. \end{cases} \quad (4.15)$$

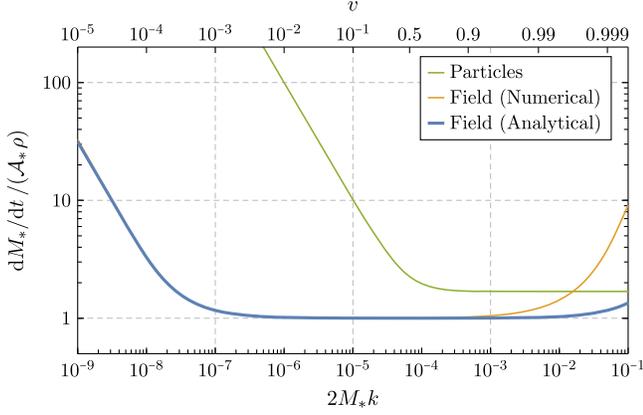


FIG. 10. Mass accretion rate of a Schwarzschild black hole computed analytically—from (4.14)—and numerically for a scalar field with $2\mu M_* = 10^{-4}$. Shown for comparison is also the accretion rate for particles given by (4.1).

It is worth noting that, at the cross-over point $k = 2\pi\mu^2 M_*$, the de Broglie wavelength of the scalar field equals the Bohr radius of the gravitational atom, $2\pi/k = r_c$.¹⁰ For $k \ll 2\pi\mu^2 M_*$, the energy flux diverges as $1/v$, just like in the particle case, but with a smaller normalization. For $k \gg 2\pi\mu^2 M_*$, instead, the energy flux is independent of v and takes the very natural form $\mathcal{A}_* \rho c$, if we restore a factor of c . This indeed matches the result for the low-energy cross section for a massless field [39–42]. The regime holds until relativistic corrections kick in at $k \sim \mu$, and our derivation breaks down.

Numerical solution—Figure 10 also shows the result of a numerical approach to the problem. In the Schwarzschild case, we numerically integrated the confluent Heun function for different values of k and ℓ_* , with the main goal of confirming that the $\ell_* = m_* = 0$ mode indeed dominates in the fuzzy limit. This allowed us to determine the near-horizon amplitudes $C_{k;\ell_* m_*}$ of modes with $\ell_* \geq 1$ as a function of the asymptotic density ρ by comparing the asymptotic limit of the confluent Heun function with the partial wave expansion of the boundary conditions (4.7). The results are in remarkable agreement with the analytical estimate for all $\mu M_* \ll 1$ and $k \ll \mu$, and explicitly show the suppression of $C_{k;\ell_* m_*}$ for $\ell_* \geq 1$.

B. Application to the realistic case

So far, we have studied an idealized model of a black hole moving through a uniform scalar field mass density. However, we would like to apply these results to the case we are actually interested in: a companion black hole of

¹⁰To give an interpretation of this result, recall that a particle with impact parameter b and velocity v is scattered by an angle $\sim M/(v^2 b)$ by the Coulomb interaction. Taking $b \sim \lambda_{\text{dB}}$, we get an order-one deflection angle for $\lambda_{\text{dB}} \sim r_c$. Scattering of waves with more (less) energy will be less (more) effective.

mass $M_* = qM$ moving through a nonuniform cloud that is bound to its parent black hole. This more realistic scenario has a few major complications over its idealized counterpart and in this section we confront them.

First and foremost, the scalar field mass density can have nontrivial azimuthal structure and so the companion can experience different densities along a single orbit. For instance, if the cloud is composed of a *real* scalar field occupying the $|211\rangle$ state, its mass density (2.10) behaves as $\rho(\mathbf{r}) \propto \cos^2 \phi$. In contrast, if it is a *complex* scalar field occupying the same state (or any other pure eigenstate), its mass density does not vary along the orbit, $\rho(\mathbf{r}) = \rho(r, \theta)$. When the mass density has nontrivial ϕ dependence, we will assume that we can replace it with its azimuthal average, $\rho(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} d\phi \rho(r, \theta, \phi)$. In this case, both real and complex scalar fields are treated equally and give identical predictions. We do not expect this to be a bad approximation, as it is roughly akin to only tracking quantities that have been averaged over an orbit, like those we work with in Sec. V.

Even assuming that we can azimuthally average the scalar field density, it is still nonuniform in the radial direction and the relative asymptotic velocity between the companion and scalar field is ill defined. We will assume that accretion occurs dynamically in a region that is much smaller than the size of the cloud, so that we can define this velocity “locally”. We will later justify this assumption. This dynamical region is mesoscopic, in the sense that the dynamics is only sensitive to the local properties of the cloud (like its density and velocity), but the region is still much larger than the size of the companion object. In place of the asymptotic fluid density, we can then use the local density $\rho(\mathbf{R}_*)$ of the cloud at the position of the companion. Similarly, we define the local velocity to be the ratio of the probability current to the probability density,

$$\mathbf{v}_c(\mathbf{R}_*) = \frac{i}{2\mu|\psi|^2} (\psi \nabla \psi^* - \psi^* \nabla \psi) = \frac{m}{\mu R_*^\perp} \hat{\phi}, \quad (4.16)$$

where m is the azimuthal angular momentum of the cloud and R_*^\perp is the length of the projection of \mathbf{R}_* on the equatorial plane, so that the difference between (4.16) and the orbital velocity of the companion, $\mathbf{v}_* \sim \pm \sqrt{M/R_*} \hat{\phi}$, is the relative fluid-black hole velocity. For equatorial circular orbits, with $R_*^\perp = R_*$, this relative velocity is

$$v = \left| \sqrt{\frac{M}{R_*}} \mp \frac{m}{\mu R_*} \right| = \frac{\alpha}{\sqrt{R_*/r_c}} \left| 1 \mp \frac{m}{\sqrt{R_*/r_c}} \right|, \quad (4.17)$$

where the $- (+)$ sign refers to corotating (counter-rotating) orbits and $r_c = (\mu\alpha)^{-1}$ is the typical radius of the cloud. We stress that the quantities $\rho(\mathbf{R}_*)$ and $\mathbf{v}_c(\mathbf{R}_*)$ are computed

without taking the backreaction of the companion into account. For small q , this is a good approximation.

Under these assumptions, and for the systems we study, the mass accretion flux is approximately independent of velocity,

$$\frac{dM_*}{dt} \approx \mathcal{A}_* \rho(\mathbf{R}_*), \quad (4.18)$$

where $\mathcal{A}_* \approx 4\pi(2qM)^2$ is the area of the companion's horizon. From the discussion of the previous section (see the ‘‘plateau’’ in Fig. 10), this approximation is valid as long as the relative fluid velocity is neither too slow nor too fast,

$$2\pi q\alpha \ll v \ll 1. \quad (4.19)$$

From (4.17), we see that this condition can be violated when either the orbital separation is very small, $R_* \sim \alpha^2 r_c$, in which case the fluid is moving too quickly, $v \sim 1$, or when the orbital separation is very large, $R_* \sim r_c/q^2$, in which case the fluid is moving too slowly, $v \ll 2\pi\alpha q^2$. Both of these cases occur during a typical inspiral. However, for small q and α , the cloud is extremely dilute whenever (4.19) is violated, because the companion is either too close¹¹ or too far away from the parent black hole to see an appreciable density, and so accretion is negligible whenever (4.18) does not apply.¹²

Finally, let us now check that the accretion process actually happens in a mesoscopic region where we can assume that the companion sees a uniform medium. The mass absorption formula (4.18) can be written as

$$\frac{dM_*}{dt} = (\pi b_{\max}^2) v \rho, \quad (4.20)$$

where $b_{\max} \equiv 4qM/\sqrt{v}$ is the radius of the absorption cross section, or the maximum impact parameter for absorption in a particle analogy. To apply the idealized derivation, we need to satisfy two conditions: (1) the density and velocity of the cloud are approximately constant over a region of size b_{\max} , and (2) the region of size b_{\max} is gravitationally dominated by the companion,

¹¹We have assumed that the cloud has nontrivial angular momentum, which pushes the density of the cloud away from the parent black hole. This is a fair assumption, as these are the types of states prepared by superradiance. Moreover, we do not expect accretion to be significant for $\ell = 0$ states anyway, since the time spent by the companion in the region $R_* \lesssim \alpha^2 r_c$ is very short.

¹²This reasoning can fail when the relative velocity (4.17) vanishes and the companion orbits the parent black hole at the same local speed as the cloud, which occurs for corotating orbits at $R_* = m^2 r_c$. In an orbital band of width $\Delta R_* \sim \pi q m^3 r_c$ around this special orbit, the constraint $2\pi q\alpha \ll v$ is violated and (4.18) cannot be applied. Rather, the low-velocity limit of (4.15) must be used instead and accretion is enhanced.

i.e., it is smaller than the radius of the Hill sphere $r_{\text{Hill}} = R_*(q/3)^{1/3}$. These two conditions then require that

$$(1) \quad b_{\max} \ll r_c \Rightarrow \frac{R_*}{M} \ll (4q\alpha^2)^{-4}, \quad (4.21)$$

$$(2) \quad b_{\max} \ll r_{\text{Hill}} \Rightarrow \frac{R_*}{M} \gg (8q/\sqrt{3})^{8/9}. \quad (4.22)$$

Both of these conditions are easily satisfied for the typical values of α , q , and R_* that we are interested in.

There are two ways the companion can fail to see such a uniform medium. The first is simply if the azimuthally-averaged density $\rho(\mathbf{R}_*)$ vanishes, or changes dramatically, at a particular orbital separation. This can occur when the cloud occupies a state $|n\ell m\rangle$, with $\ell \neq n-1$, for which the radial wave function has zeros away from the origin. In this case, we can think of the density that the companion sees as simply being the averaged density within a Hill sphere about the companion. Similarly, as illustrated in Fig. 8, the companion itself changes the local density—it vacuums up the scalar field as it passes through the cloud and leaves an empty ‘‘tube’’ of diameter $\mathcal{O}(M_*)$. However, the cloud will respond and replenish this local density on a relatively short timescale. This perturbation excites modes with typical wavelength of $\mathcal{O}(M_*)$, whose frequencies $\omega^2 = \mu^2 + k^2$ scale as $\mathcal{O}(\mu/(aq))$. These modes respond extremely quickly, and we expect that this empty ‘‘tube’’ is rapidly filled in before companion can complete an orbit and encounter this locally depleted region again. So, the companion should see a relatively uniform medium throughout the inspiral, and we will thus use the approximation (4.18) throughout Sec. V to capture the effect accretion has on the binary's dynamics.

V. BACKREACTION ON THE ORBIT

We will now study the effect that both ionization and accretion have on a binary inspiral. We are mostly interested in intermediate or extreme mass ratio inspirals, where the light companion moves inside the cloud of the much heavier parent black hole. In Sec. VA, we describe the system and its evolution equations, while in Sec. VB we show numerical solutions to these equations for a few representative examples.

A. Evolution equations

Chronologically, the first resonant transitions in the inspiral are those with the lowest frequency. These typically happen before the separation becomes comparable to r_c . During those resonances, the state of the cloud can be transformed to decaying states. For example, for an initial $|211\rangle$ state, the first resonances mediated by the quadrupolar perturbation ($\ell_* = 2$) connect it to the $|21-1\rangle$ and $|31-1\rangle$ states in the corotating and counter-rotating cases, respectively. It is nontrivial to understand whether or not

the cloud survives after these transitions, though it has recently been shown that it can [23] in some cases.

Our main goal in this section is to understand the physics of the subsequent inspiral, away from resonances, and under the hypothesis that the cloud is still present when ionization and accretion kick in. Our results should thus not be read as a fully realistic solution of the dynamics of the system, as that would require including the resonances (and their impact on the evolution of the cloud). Rather, we present an example of the impact of ionization and accretion only, and their interplay. We will restrict to quasi-circular, equatorial orbits, and study separately orbits that are corotating and counter-rotating with respect to the cloud. The gravitational field of the cloud will also be neglected, as it gives a correction of order M_c/M to the orbital quantities, which, as we will see, is subdominant with respect to the impact of ionization and accretion. We will numerically solve the time evolution of three quantities: the companion's mass M_* , the cloud's mass M_c , and the separation R_* .

The evolution of M_* and M_c is determined by mass conservation. As we discussed in Sec. IV, the mass of the companion increases by accretion, while the mass of the cloud decreases by the corresponding amount. In addition, the cloud loses mass through ionization. We therefore have

$$\frac{dM_*}{dt} = 4\pi(2M_*)^2\rho(\mathbf{R}_*), \quad (5.1)$$

$$\frac{dM_c}{dt} = -\frac{dM_*}{dt} - M_c \sum_{\ell,g} \left[\frac{\mu |\eta_{K,b}^{(g)}(t)|^2}{k_*^{(g)}(t)} \Theta(k_*^{(g)}(t)^2) \right], \quad (5.2)$$

where $\rho(\mathbf{R}_*) = M_c |R_{n\ell}(R_*) Y_{\ell m}(\theta_*, \phi_*)|^2$ is the local density of the cloud at the position of the companion.¹³ The accretion formula (5.1) holds for a nonrotating black hole, while for a rotating black hole it has to be rescaled to account for the reduced area of the horizon. The last term of (5.2) is the ionization rate, defined in (3.28).

To determine the backreaction on the inspiral, we use the conservation of angular momentum. The system carries angular momentum in the form of the orbital angular momentum of the binary and the spin of the cloud, which are given by

$$L \equiv \frac{M_* \Omega R_*^2}{1+q} \quad \text{and} \quad S_c \equiv \frac{m M_c}{\mu}, \quad (5.3)$$

¹³Here, we have ignored the possibility that accretion is enhanced, as it is for corotating orbits at $R_* = m^2 r_c$. This enhancement occurs in a region that is too narrow to be resolved for the values of q we consider. We also assumed that either the scalar field is complex, or that we have azimuthally averaged the mass density of the real scalar field. Both provide the same result.

where $\Omega^2 R_*^3 = (1+q)M$ for quasi-circular Keplerian orbits. Gravitational waves carry angular momentum to infinity at a rate $dL_{\text{GW}}/dt = P_{\text{GW}}/\Omega$, where P_{GW} is given in (2.19). In the vacuum solution, these gravitational waves are the reason for the shrinking orbit. Ionization leads to an additional loss of angular momentum through the emission of scalar waves. These waves carry angular momentum to infinity at a rate given by an expression analogous to (3.29),

$$\frac{dL_{\text{out}}}{dt} = \sum_{\ell,g} \left[(m+g) \frac{\mu |\eta_{K,b}^{(g)}(t)|^2}{k_*^{(g)}(t)} \right] \Theta(k_*^{(g)}(t)^2) |c_b(t)|^2. \quad (5.4)$$

The conservation of the total angular momentum then implies

$$\frac{dL}{dt} + \frac{dS_c}{dt} = -\left(\frac{P_{\text{GW}}}{\Omega} + \frac{dL_{\text{out}}}{dt} \right). \quad (5.5)$$

Using (5.2) for the evolution of M_c in dS_c/dt , we can express the difference between its last term and dL_{out}/dt in terms of the ionization power, P_{ion} , defined in (3.29). This leads to an equation for the evolution of the binary's separation,¹⁴

$$\begin{aligned} \frac{qM^2}{2R_*^2} \frac{dR_*}{dt} = & -P_{\text{GW}} - P_{\text{ion}} - \left[\frac{2+q}{2(1+q)^{3/2}} \sqrt{MR_*} \mp \frac{mM}{\alpha} \right] \\ & \times M |\Omega| \frac{dq}{dt}, \end{aligned} \quad (5.6)$$

where the minus (plus) sign refers to corotating (counter-rotating) orbits. We see that the inspiral dynamics is determined by three different ‘‘forces’’. The first two have the obvious interpretation of the drag induced by the energy lost in gravitational waves and scalar waves, respectively. The third term, instead, is the accretion of momentum that comes along with the accretion of mass. The sign of this force depends on whether the cloud is locally rotating faster or slower than the companion. Not surprisingly, the two behaviors are separated by $R_* = m^2 r_c$ (in the small- q limit), corresponding to the special corotating orbit identified in Sec. IV B where the relative velocity vanishes.

As a final note, we observe that the backreaction of the gravitational interaction between an object and the medium

¹⁴Note that this expression neglects the transient oscillations associated with the discontinuities. As we showed in Sec. III, these oscillations decay over a very narrow region of R_* in the small backreaction limit. This region remains small even in the cases studied here, as even though the backreaction is strong and the instantaneous chirp rate $\dot{\varphi}_*(t)$, cf. Appendix A 5, is enhanced roughly by a factor of $P_{\text{ion}}/P_{\text{GW}} \sim \mathcal{O}(100)$, this narrow region scales as $\gamma^{1/2}$ and, especially for the parameters we are interested in, this region is still small enough to ignore the effect that the transient oscillations and varying chirp rate $\dot{\varphi}_*(t)$ can have on (5.6).

it is moving through is known in the literature as “dynamical friction”. For uniform density media, the interpretation of the effect is simple; the wake of the overdensity behind the moving object exerts a gravitational pull on it, creating a drag force. The effect has been computed for a light field in [47], and there have been some recent attempts to apply it to the case of the gravitational atom in [17,48]. The length of the wake and the intensity of the drag force depend on the history of the system, with divergent results found for stationary configurations in asymptotically uniform media. The gravitational atom, however, is special in two ways; first, it is localized in space, providing a natural regulation for the divergence mentioned previously, and second, its spectrum is composed of bound and unbound states, but only the latter can carry (angular) momentum to infinity [49]. Despite these complications, the physical origin of the drag force is the same. It is therefore a question of semantics whether one calls the drag induced by the backreaction of ionization “dynamical friction”. In any case, because bound and unbound states together form a complete set, the description of the evolution of their occupations, and the associated backreaction, either in the form of resonances or drag, provides a full description of the interaction between the cloud and the moving object.

B. Numerical results

The system of equations (5.1), (5.2), and (5.6) determines the evolution of M_* , M_c , and R_* . In this section we solve these equations numerically for some benchmark parameters.

It is first useful to comment on our choice of fiducial parameters and their astrophysical plausibility. To make a strong observational case, we choose parameters for which ionization and accretion occur mostly in-band for a future space-based detector like LISA. At the same time, we must require that $q \ll 1$ in order for our perturbative treatment to be applicable (see Appendix B). We thus consider intermediate mass ratio inspirals, with $M = 10^4 M_\odot$ and $q = 10^{-3}$, as we want the companion to be a reasonably-sized black hole. In order for the discontinuities in the ionization power P_{ion} to appear in the LISA band, we take $\alpha = 0.2$. This allows very fast superradiant growth of the cloud, but also makes it decay relatively rapidly to gravitational waves when the scalar field is real. The exact depletion rate depends on the initial mass of the cloud, but for these parameters M_c/M is expected to fall to 0.01 after 10^5 years and to 0.001 after 10^6 years, with an extremely strong dependence on α . It is thus not unreasonable to take $M_c/M = 0.01$ as a reference point for its initial value when ionization and accretion kick in; however, we will also show that even for $M_c/M = 0.001$ the impact of the cloud is still very large.

It is possible to adjust the values of M , q , and α . For example, we could reduce the value of α to make the cloud longer lived. If we want the ionization features of the signal

to stay in the LISA band, then we would have to simultaneously reduce the value of M (which would increase q if we keep M_* fixed). However, in this work we only want to illustrate that ionization has a *large* and *sharp* effect on the inspiral, and we therefore do not attempt to find the region of parameter space with the most observational relevance. In the same vein, we fix the initial state of the cloud to $|211\rangle$ for simplicity. As previously mentioned, there is an uncertainty in the initial bound state due to the previous history of the system, both from past resonant transitions and the superradiant growth of other modes like the $|322\rangle$ state which becomes relevant for larger values of α . More concretely, for counter-rotating orbits, the state $|211\rangle$ cannot undergo any hyperfine transitions and the first Bohr transition (to the state $|31-1\rangle$) occurs around $R_*/M \sim 200$, when ionization is already a large effect (see Fig. 7). For corotating orbits, the hyperfine transition to the state $|21-1\rangle$ can be significant and would have to be included in the analysis. We do not expect that choosing a different initial state would qualitatively affect our conclusions, but leave a more detailed analysis for future work.

Let us now describe the numerical results. To understand the magnitude of the different effects, we show in Fig. 11 the evolution of the parameters separately under the effects of ionization and accretion and then both combined, starting from a separation of $R_* = 400M$. In all cases, we observe a very significant shortening of the time to merger, with the orbits suddenly *sinking* as soon as the ionization energy losses overcome those in gravitational radiation. The dynamical evolution of the system is thus *driven*, and not just *perturbed*, by the interaction of the binary with the cloud. The binary merges faster for counter-rotating orbits, since the ionization power is larger at

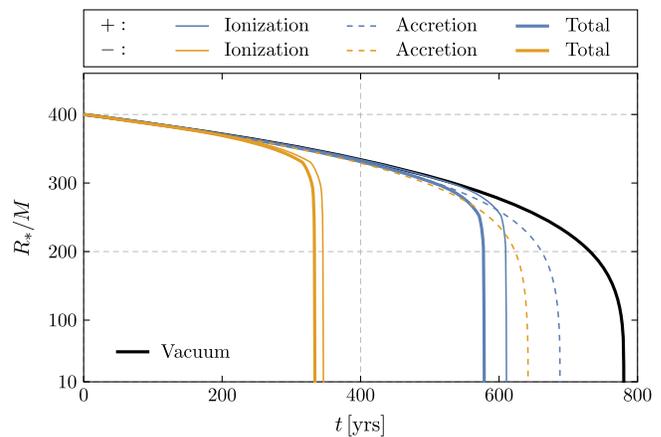


FIG. 11. Evolution of the separation R_* , for $M = 10^4 M_\odot$ and $\alpha = 0.2$, with initial values of $R_* = 400M$, $q = 10^{-3}$ and $M_c/M = 0.01$ in a $|211\rangle$ state. Shown are the results for both corotating (+) and counter-rotating (-) orbits. The vacuum system, where no cloud is present, is shown for comparison. We see that accretion and ionization significantly reduce the merger time.

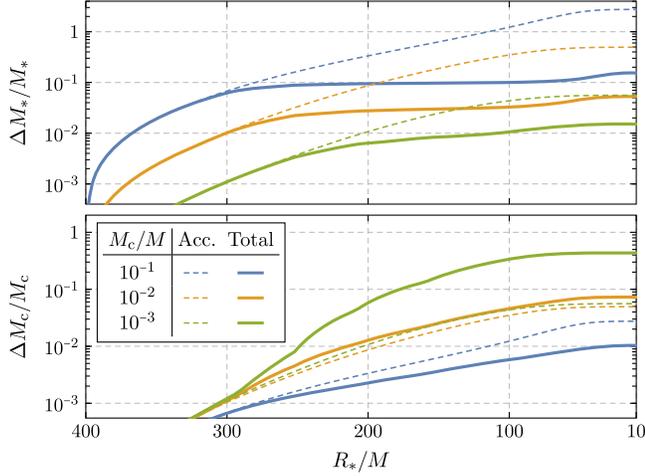


FIG. 12. Fractional changes of the mass of the companion M_* and the mass of the cloud M_c , for $M = 10^4 M_\odot$ and $\alpha = 0.2$, with initial values of $R_* = 400M$, $M_* = 10^{-3}M$. Shown are the results for three different initial values of M_c . All curves refer to corotating orbits and a $|211\rangle$ bound state.

large R_* and the accretion force is opposite to the motion, cf. Fig. 7 and (5.6).

In the top panel of Fig. 12, we show the fractional change of the mass of the companion M_* as function of the separation R_* , for corotating orbits and three different initial values of M_c . We observe that accretion takes place throughout the entire inspiral, without a clear hierarchy between the timescales of accretion and merger. Not surprisingly, we see that the accreted mass is very sensitive to M_c/M , with the total $\Delta M_*/M_*$ being roughly proportional to it, at least in the early stages of the inspiral. In fact, in the case with only accretion, the final value of $\Delta M_*/M_*$ can be predicted from a simple order-of-magnitude estimate; multiplying the average accretion flux, (4.18), by the time-to-merger in vacuum, we get $\Delta M_*/M_* \sim (M_c/M)(r_c/M)$, which is in good agreement with the numerical results shown in Fig. 12. What is maybe more surprising is that the inclusion of ionization strongly limits the accretion of mass. This phenomenon can be explained by noting that ionization does not have a big effect on the accretion rate (5.1), which only depends on ρ and q , but significantly reduces the time spent inside the cloud, and therefore the total accreted mass.

In the bottom panel of Fig. 12, we show the fractional change of the mass of the cloud M_c . We see that the cloud is partially depleted during the inspiral, due to both ionization and accretion. The hierarchy between the two effects depends on the initial value of M_c . For more massive clouds, the primary mechanism of mass loss is accretion, which is limited by the inclusion of ionization due to the reduced time spent inside the cloud. Instead, for lighter clouds, ionization is the primary mechanism of mass loss. We see that the total mass loss does not seem to depend

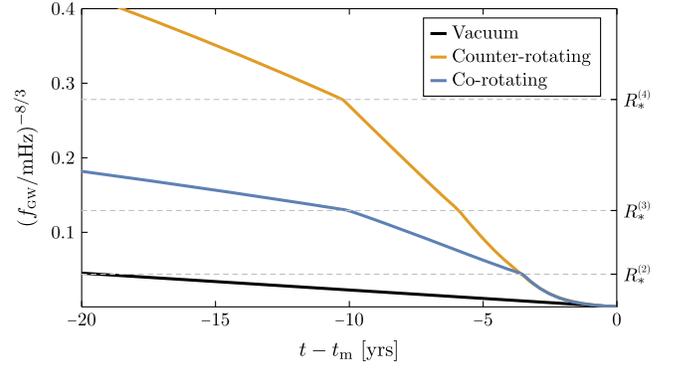


FIG. 13. Evolution of the GW frequency as a function of the remaining time to merger, $t - t_m$, for $M = 10^4 M_\odot$ and $\alpha = 0.2$, with initial values of $R_* = 400M$, $q = 10^{-3}$, and $M_c/M = 10^{-3}$ in a $|211\rangle$ state. The central region of the range shown on the y axis corresponds to a few millihertz, falling inside the LISA sensitivity band. The “kinks” at separations $R_*^{(g)}$ correspond to the discontinuities in the ionization power, see Fig. 7.

sensitively on the initial value of M_c , so that the fractional mass loss is larger for smaller clouds. In our example, with $M_c/M = 0.1$, only about 1% of the initial mass is lost at the end of the inspiral; instead, more than 50% would be depleted for an initial $M_c/M = 10^{-3}$.

It is natural to wonder how degenerate the observables are with the expected signal from a binary in vacuum with different parameters. Although we postpone a systematic study of this issue to future work, it is useful to compare the evolution of the GW frequency f_{GW} as a function of the time to merger. This is done in Fig. 13 for the very conservative case of initial $M_c/M = 10^{-3}$, demonstrating that even a tiny cloud can have a strong impact on the inspiral. In the plot, the scale of the frequency axis has been chosen such that the nonrelativistic vacuum evolution, $f_{\text{GW}} \propto (t_m - t)^{-3/8}$, where t_m is the merger time, becomes a straight line. It is apparent that the shape of $f_{\text{GW}}(t)$ deviates significantly from a straight line: a decisive role is played by the “kinks” appearing at the frequencies where the ionization power P_{ion} is discontinuous, cf. Fig. 7. From (3.30), kinks appear at the frequencies

$$f_{\text{GW}}^{(g)} = \frac{6.45 \text{ mHz}}{g} \left(\frac{10^4 M_\odot}{M} \right) \left(\frac{\alpha}{0.2} \right)^3 \left(\frac{2}{n_b} \right)^2 = \frac{33.5 \text{ mHz}}{g} \left(\frac{M}{10^4 M_\odot} \right)^2 \left(\frac{\mu}{10^{-14} \text{ eV}} \right)^3 \left(\frac{2}{n_b} \right)^2, \quad (5.7)$$

where the overtone number g ranges over positive integers and n_b is the principal number of the cloud’s initial state. These kinks thus constitute a sharp observational signature of ionization caught in the act. If only a region between two kinks is observed, then the evolution is likely to be more degenerate with a signal from a vacuum system, whose

parameters would however differ from the true parameters of the binary.

C. Open problems

We now address a number of unresolved questions regarding the phenomenology of the system, discussing the limitations of our analysis and some future prospects.

Gravitational field of the cloud—By using the simple Keplerian relation $\Omega^2 R_*^3 = (1 + q)M$, we have neglected the backreaction due to the gravitational field of the cloud. This backreaction would manifest itself as $\mathcal{O}(M_c/M)$ corrections to the orbital dynamics. At the Newtonian level, the effect of the cloud is twofold; the enclosed mass “seen” by the companion varies with R_* , due to the diffuse nature of the cloud, and the nontrivial angular structure of the cloud generates higher-mass multipoles. The first effect is only relevant when the companion orbits inside the cloud, $R_* \sim r_c$, while the second can also provide corrections at large distances.

Angular structure of the cloud—Similarly, we have ignored the angular structure of the cloud in our treatment of accretion, where we azimuthally averaged the mass density and assumed that the accretion process was accurately captured by averaging over each orbit. This implicitly assumes that the orbit remains quasicircular even after we include accretion effects. However, we expect that this assumption can break down at certain points during the inspiral, like when the relative velocity between the cloud and companion vanishes for corotating orbits and the companion has enough time to develop nontrivial eccentricity.

Inclination and eccentricity—For simplicity, we have only studied equatorial quasicircular orbits. The phenomenology of inclined orbits is potentially much richer, as the transfer of angular momentum between the cloud and orbit can cause the orbital plane to precess. The companion would also explore regions of the cloud with different densities, resulting in an uneven distribution of the “forces” appearing in (5.6) over the course of an orbit, potentially causing the orbit to become more eccentric. Taking into account eccentricity is necessary for a more complete analysis even in the simple case of equatorial orbits, especially in situations where the forces in (5.6) have a nontrivial R_* dependence. For example, for corotating orbits the accretion force does not always act as a drag, changing sign with R_* .

Resonances—We have not studied the interplay of the resonances between bound states with the ionization and accretion processes. The effect of resonances on the dynamics is twofold. First, they introduce periods of either accelerated (“sinking” orbits), or decelerated (“floating” orbits) inspiral; these would appear as distinctive features in the evolution of the separation and frequency. Second, the resonances can change the state of the cloud. Both of these effects can interact nontrivially with ionization and

accretion, as the total mass accreted or ionized depends on the time spent at a given orbital separation, and on the state of the cloud. For instance, the effects of both ionization and accretion will be enhanced during a floating orbit, while a rapidly sinking orbit can break many the various approximations we have relied on in our analysis. Furthermore, when the cloud transitions to an excited state it becomes easier to ionize, so this dependence on the evolution of the state has to be included in a self-consistent analysis of the ionization. It would be interesting to study the state dependence of the ionization signal in more detail.

Equal mass ratios—We have only studied the case of a large mass ratio $q \ll 1$, where the gravitational influence of the companion could be treated perturbatively. The parameter q is one of the main order parameters in our perturbative analysis and many of our approximations do not hold when $q \sim 1$. It would be interesting to develop a formalism that is able to treat the case of equal mass ratios,¹⁵ where ionization can be efficient enough to completely evaporate the cloud before the merger.

Transient oscillations—Our analysis relied on replacing the dynamics of the ionization process with its “steady state” behavior (3.28). However, as we described in Sec. III A, there is interesting transient behavior that occurs when a bound state just begins to resonate with a continuum band. How are these transient oscillations modified when we include the cloud’s backreaction on the orbit? Do these oscillations also affect the orbital dynamics, and can we observe them in the resulting gravitational wave signal? These are interesting questions for the future that require a different formalism to answer.

Relativistic corrections—Our treatment was nonrelativistic, both in the derivation of the mass accretion and ionization, and in the orbital evolution. Hence, our results do not apply in the final phase of the inspiral—closer to the merger—when the velocity approaches the speed of light and the post-Newtonian expansion breaks down. This period of the inspiral is notoriously difficult to model even for vacuum systems, especially for large mass ratios. However, close to the merger, we expect the effects of the cloud to fade in comparison to the increasingly strong nonlinearities of the vacuum evolution (see Figs. 7 and 13). The region where resonances, ionization and accretion are most relevant is thus within the applicability of the non-relativistic approximation.

VI. CONCLUSIONS

Gravitational waves provide an interesting new window into the weak-coupling frontier of particle physics, giving us access to physics that is invisible to traditional collider experiments [31]. Such weakly-coupled sectors arise in the string landscape as ultralight axions [6,51–53] and are also

¹⁵For a recent attempt to describe this regime see [50].

interesting dark matter candidates [47]. In this paper, we have studied the dynamical effects of clouds of ultralight scalars around black holes when they are part of binary systems. We have calculated two novel types of cloud-binary interactions; the ionization of the cloud due to the gravitational perturbation from the binary companion and the accretion of mass onto the secondary object, in the case it is a black hole.

When unbound states are excited by the gravitational perturbation due to the companion, the cloud loses mass to outgoing scalar waves. This ionization induces a backreaction on the orbit of the binary, which loses energy and angular momentum to the scalar field. These losses are notable for two reasons: (1) they can significantly exceed the energy loss due to GW emission, thus dominating the dynamics of the system, and (2) they contain sharp features (see Fig. 7), which carry detailed information about the microscopic structure of the cloud.

During the inspiral, the companion will move inside the scalar cloud. If the companion is a black hole, then its event horizon will absorb parts of the cloud. Due to the high densities reachable by superradiantly-generated clouds, the mass of the secondary object can significantly increase during the inspiral, which impacts the dynamics of the system. The momentum accreted by the object is also non-negligible.

Both ionization and accretion affect the orbital dynamics of the binary. We studied this backreaction numerically, finding that the deviation from the expectations for a vacuum system can be rather dramatic. The inspiral happens much faster than in the absence of the cloud, and both the mass of the companion and of the cloud evolve significantly in time. Even with conservative choices of parameters, the frequency evolution is quantitatively and qualitatively modified, especially due to the discontinuities in the ionization power producing “kinks” in the frequency evolution of the gravitational waves (see Fig. 13). These features are a new and distinctive signature of gravitational atoms in black hole binaries.

Our analysis made a number of simplifying assumptions. First, we restricted ourselves to extreme mass ratio inspirals on quasicircular orbits in the equatorial plane. We expect that qualitatively new behavior appears for equal mass ratios, and that both inclination and eccentricity can lead to a rich phenomenology in the presence of the cloud. These are both interesting directions for future work. Similarly, we did not explore the interplay between bound state resonances and both ionization and accretion, nor did we account for the interesting transient phenomena that occur when the ionization process begins. A more complete analysis should take both of these into account.

A combined treatment of the resonances studied in [16,31], together with the ionization and accretion discussed in this work, is required to achieve a complete understanding of the phenomenology of gravitational

atoms in binaries. This in turn will serve as a starting point to devise suitable strategies to discover and characterize these systems with upcoming gravitational-wave detectors. Current data analysis techniques mostly rely on matched filtering, where waveform templates are compared to observations. Waveforms based on vacuum systems may thus produce a very low signal-to-noise ratio when applied to our case, because of the drastically different evolution of the observables. In a template-based approach, dedicated searches are thus needed to not miss inspirals involving gravitational atoms and to distinguish them from other kinds of environmental effects, like dark matter overdensities [54–57]. We postpone a systematic study of these phenomenological issues to future work.

ACKNOWLEDGMENTS

D. B. and J. S. are grateful to Horng Sheng Chia and Rafael Porto for previous collaborations on this topic. We thank Thomas Spieksma for verifying some of the numerical results in this paper. D. B. receives funding from a VIDI grant of the Netherlands Organisation for Scientific Research (NWO) and is part of the Delta-ITP consortium. D. B. is also supported by a Yushan Professorship at National Taiwan University funded by the Ministry of Science and Technology (Taiwan). J. S. is supported by NASA Grant No. 80NSSC20K0506.

APPENDIX A: INTEGRATING OUT THE CONTINUUM

As explained in the main text, the dynamics of the gravitational atom in a binary, including both bound and continuum states, can be captured by integrating out the continuum and incorporating its effects in terms of a set of induced couplings and energies for the bound states alone. This process yields an effective Schrödinger equation for the bound states that describes the behavior of the entire system. In this appendix we justify the approximations we used to derive these continuum-induced couplings. First, we explain how our approximation for the fractional deoccupation rate (3.10) in the toy model arises from the large time asymptotics of the induced energy. This derivation relies on ignoring the transitions between continuum states, so we then justify this assumption. Next, we discuss the complications that arise in the more realistic case, which includes many more bound and continuum states. We then describe an alternative, albeit uncontrolled, derivation of (3.10) using stationary perturbation theory. Finally, we conclude with a discussion of the effects a nonlinearly ramping frequency $\dot{\varphi}_*(t)$ has on our approximations.

1. Saddle point approximation

We are interested in the asymptotic behavior of the induced energy

$$\begin{aligned}\mathcal{E}_b(t) &= \int_{-\infty}^t dt' \Sigma_b(t, t') \\ &= \frac{1}{2\pi i} \int_{-\infty}^t dt' \int_0^{\infty} dk |\eta(k)|^2 e^{-i(\epsilon(k) - \epsilon_b)(t-t') + i(\varphi_*(t) - \varphi_*(t'))},\end{aligned}\quad (\text{A1})$$

where $\epsilon(k) = k^2/(2\mu)$. Without loss of generality, we can absorb the bound-state energy into our reference frequency, $\varphi_*(t) = -\epsilon_b t + \gamma t^2/2$, and assume that $\gamma > 0$. The bound state then begins to “resonate” with the continuum for $t \gtrsim 0$, and we would like to determine the asymptotic behavior of this function before and after this time, $|\sqrt{\gamma}t| \gg 1$, as a way of approximating its behavior away from the complicated transient region around $t = 0$.

There are two representations of this function that will be useful. We can either first perform the integral over t' to find

$$\begin{aligned}\mathcal{E}_b(t) &= \frac{1}{\sqrt{8\pi\gamma}} \int_0^{\infty} d\epsilon |\eta(\epsilon)|^2 \exp\left[\frac{i(\epsilon - \gamma t)^2}{2\gamma} - \frac{3\pi i}{4}\right] \\ &\quad \times \operatorname{erfc}\left[\frac{e^{\frac{i\pi}{4}}(\epsilon - \gamma t)}{\sqrt{2\gamma}}\right],\end{aligned}\quad (\text{A2})$$

or we can define $z = \sqrt{\gamma}(t - t')$ and write

$$\mathcal{E}_b(t) = \frac{1}{2\pi i \sqrt{\gamma}} \int_0^{\infty} dz e^{itz} \mathcal{K}(z), \quad (\text{A3})$$

where we introduced the dimensionless time $\tau \equiv \sqrt{\gamma}t$ and the kernel

$$\mathcal{K}(z) \equiv e^{-\frac{1}{2}iz^2} \int_0^{\infty} d\epsilon e^{-i\epsilon z/\sqrt{\gamma}} |\eta(\epsilon)|^2. \quad (\text{A4})$$

The former has the benefit of making the “resonance” behavior much clearer, while the latter is useful for understanding the large time $|\tau| \gg 1$ asymptotics since it has the form of a standard Laplace-like integral. In both representations, we have transformed the integral over momenta k into an integral over the energy ϵ and defined $|\eta(\epsilon)|^2 = dk(\epsilon)/d\epsilon |\eta(k(\epsilon))|^2 = \mu |\eta(k)|^2/k$. In the cases of interest, $|\eta(\epsilon)|^2$ approaches a constant as $\epsilon \rightarrow 0$ and decays algebraically as $\epsilon \rightarrow \infty$, so that the “total coupling” of the bound state to the continuum $\int_0^{\infty} d\epsilon |\eta(\epsilon)|^2$ is finite.

To get a sense for the behavior of this function, it is useful to first rescale the integral in (A2) by taking $\epsilon \rightarrow \sqrt{\gamma}|\tau|\tilde{\epsilon}$,

$$\mathcal{E}_b(\tau) = \frac{|\tau|}{\sqrt{2\pi}} \int_0^{\infty} d\tilde{\epsilon} |\eta(\sqrt{\gamma}|\tau|\tilde{\epsilon})|^2 \mathcal{I}(\tilde{\epsilon}, \tau), \quad (\text{A5})$$

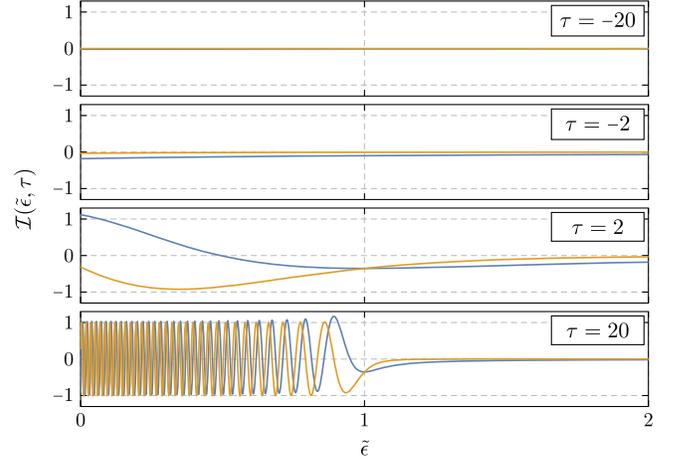


FIG. 14. The real [blue] and imaginary [orange] parts of the modulating function $\mathcal{I}(\tilde{\epsilon}, \tau)$, for several values of the dimensionless time τ . For large negative values of τ , the integrand of (A5) is highly suppressed for $\tilde{\epsilon} \in [0, \infty)$. For large positive times $\tau \gg 1$, the integrand oscillates rapidly when $\tilde{\epsilon} \in [0, 1]$, slowing down when $\tilde{\epsilon} \sim 1$, and is then again highly suppressed for $\tilde{\epsilon} \gg 1$.

where we defined the kernel

$$\mathcal{I}(\tilde{\epsilon}, \tau) \equiv \frac{1}{2} e^{\frac{i^2}{2}(\tilde{\epsilon} - \operatorname{sgn}\tau)^2 - \frac{3\pi i}{4}} \operatorname{erfc}\left[\frac{|\tau|}{\sqrt{2}} e^{\frac{i\pi}{4}}(\tilde{\epsilon} - \operatorname{sgn}\tau)\right]. \quad (\text{A6})$$

We plot this kernel for several values of τ in Fig. 14. We see that, for $\tau \rightarrow -\infty$, the integrand of (A5) is strongly suppressed throughout the entire integration region, and so both the real and imaginary parts of the induced energy will be small. In the opposite limit, $\tau \rightarrow +\infty$, the integrand oscillates rapidly in the interval $\tilde{\epsilon} \in (0, 1)$, so we expect only the end point $\tilde{\epsilon} = 0$ and the region around $\tilde{\epsilon} = 1$ to contribute to the integral. For $\tilde{\epsilon} \in (1, \infty)$, the integrand no longer oscillates, but instead decays algebraically. The integrand—and especially the real part in [blue]—has a very heavy tail which the saddle point approximation is not able to fully capture. Instead, we will need to use the Laplace-like form (A3) to compute these additional contributions.

Keeping in mind that the saddle point approximation does not capture the full behavior of the induced energy as $\tau \rightarrow \infty$, we will apply it anyway. As stated before, there are two relevant contributions—from the endpoint at $\tilde{\epsilon} = 0$ and from the “saddle point” at $\tilde{\epsilon} = 1$. From Fig. 14, we expect that the contribution at $\tilde{\epsilon} = 0$ produces an oscillatory *ringing* that is left over from when the bound state first hits the edge of the continuum, and how quickly these oscillations decay depends on how the bound state couples to the lowest-energy continuum modes, i.e., how $|\eta(\epsilon)|^2$ scales as $\epsilon \rightarrow 0$. In contrast, the saddle point at $\tilde{\epsilon} = 1$ gives a nonoscillatory decay which only depends on the coupling between the bound state and the particular continuum state

it is “resonating with” $|\eta(\epsilon = \gamma t)|^2$. Assuming that $|\eta(\epsilon)|^2$ approaches a constant $|\eta|^2$ as $\epsilon \rightarrow 0$, we find that

$$\mathcal{E}_b(t) \sim -\frac{i\mu|\eta(k_*(t))|^2}{2k_*(t)} - \frac{|\eta|^2 e^{\frac{1}{2}i\gamma t^2 - \frac{i\pi}{4}}}{2\sqrt{2\pi\gamma t}} \left[1 + \operatorname{erf}\left(\frac{e^{\frac{i\pi}{4}}\sqrt{\gamma t}}{\sqrt{2}}\right) \right],$$

$$\sqrt{\gamma t} \rightarrow +\infty, \quad (\text{A7})$$

where we have switched back to parametrizing the system in terms of the momentum and introduced $k_*(t) = \sqrt{2\mu\gamma t}$, the momentum of the state at the saddle point.

To find the dominant behavior of $\operatorname{Re}\mathcal{E}_b(t)$ as $\tau \rightarrow \pm\infty$, we can use (A3) and repeatedly integrate by parts in z to generate an expansion in powers of τ^{-1} . However, the aforementioned heavy tail can hinder this iterative process. Each integration by parts generates higher derivatives of the kernel evaluated at $z = 0$, but these derivatives are not necessarily finite. From (A4), we see that $\partial_z^k \mathcal{K}(z)|_{z=0}$ contains a term proportional to $\int_0^\infty d\epsilon \epsilon^k |\eta(\epsilon)|^2$, and since $|\eta(\epsilon)|^2$ decays only algebraically, sufficiently high derivatives will diverge. This signals that $\mathcal{K}(z)$ has terms of the form $z^k \log^n z$, which produce asymptotic behavior of the form $\log^n \tau / \tau^{k+1}$, i.e., logarithmic behavior that is not captured in the standard saddle point approximation.

For our purposes, we will only concentrate on the leading order $|\tau| \rightarrow \infty$ behavior. This is governed by the total coupling $\mathcal{K}(0) = \int_0^\infty d\epsilon |\eta(\epsilon)|^2 = \int_0^\infty dk |\eta(k)|^2$, and direct integration yields

$$\mathcal{E}_b(t) \sim \frac{1}{2\pi\gamma t} \left[\int_0^\infty dk |\eta(k)|^2 \right] + \dots \quad (\text{A8})$$

As $\tau \rightarrow -\infty$, this is the dominant contribution and gives an accurate approximation—as the effective energy gap between the bound and continuum states shrinks, the coupling to the continuum induces a correction to the bound state’s energy. There is, however, no appreciable deoccupation of the bound state until after the transition at $\tau = 0$. As $\tau \rightarrow +\infty$, the integral picks up an additional saddle point and the induced energy is well approximated by

$$\mathcal{E}_b(t) \sim -\frac{i\mu|\eta(k_*(t))|^2}{2k_*(t)} - \frac{|\eta|^2 e^{\frac{1}{2}i\gamma t^2 - \frac{i\pi}{4}}}{2\sqrt{2\pi\gamma t}} \left[1 + \operatorname{erf}\left(\frac{e^{\frac{i\pi}{4}}\sqrt{\gamma t}}{\sqrt{2}}\right) \right]$$

$$+ \frac{1}{2\pi\gamma t} \left[\int_0^\infty dk |\eta(k)|^2 \right] + \dots \quad (\text{A9})$$

Since we are mainly concerned with the imaginary part of this expression, we use the first term in (A9) throughout the main text.

2. Unbound-unbound transitions

It will be helpful to address our assumption that we can ignore the transitions between the continuum states in our analysis of the ionization process. We will do so in the toy model studied above and in Sec. III A. Numerical experiments show that the bound state’s dynamics are relatively unaffected if we include these transitions and is still well described by the first term in (A9). We can understand better why they may be ignored, and justify our assumption, by including these couplings in the toy Hamiltonian (3.1) and arguing that they should, at least at weak coupling, provide a subleading correction to the effective Schrödinger equation (3.8).

A nontrivial coupling between continuum states $\eta(k, k') = \langle k | \mathcal{H} | k' \rangle$, for $k \neq k'$, changes the solution (3.5) for the continuum amplitudes to

$$c_b(t) = -i \int_{-\infty}^t dt' \eta(k) e^{-i\varphi_*(t') + i(\epsilon(k) - \epsilon_b)t'} c_b(t')$$

$$+ \frac{1}{2\pi i} \int_{-\infty}^t dt' \int_0^\infty dk' \eta(k, k') e^{i(\epsilon(k) - \epsilon(k'))t'} c_{k'}(t'). \quad (\text{A10})$$

Importantly, both the bound-to-unbound couplings $\eta(k)$ and unbound-to-unbound couplings $\eta(k, k'; t)$ are $\mathcal{O}(q\alpha)$ and we work exclusively in the $q\alpha \ll 1$ regime. By plugging this solution back into itself, we can generate a solution purely in terms of the bound state amplitude, with the first correction to the $\eta(k, k') \rightarrow 0$ limit of (A10) being

$$c_k(t) \supset -\frac{1}{2\pi} \int_{-\infty}^t dt_1 \int_{-\infty}^{t_1} dt_2 \int_0^\infty dk' \eta(k, k') \eta(k') e^{i(\epsilon(k) - \epsilon(k'))t_1 - i\varphi_*(t_2) + i(\epsilon(k') - \epsilon_b)t_2} c_b(t_2), \quad (\text{A11})$$

which is $\mathcal{O}(q^2\alpha^2)$, while other corrections are higher order.

In the bound state Schrödinger equation (3.8), this correction contributes a term involving the chain of matrix elements $\langle b | \mathcal{H} | k \rangle \langle k | \mathcal{H} | k' \rangle \langle k' | \mathcal{H} | b \rangle$, while the leading-order solution only involves the chain of elements $\langle b | \mathcal{H} | k \rangle \langle k | \mathcal{H} | b \rangle$. Clearly, the leading-order contribution only accounts for the system transitioning into the

continuum and then back to the bound state, while higher-order corrections involve the system going into the continuum and then bouncing around between different continuum states before returning to the bound state. Each of these transitions is thus penalized by an additional factor of $q\alpha$ and so we expect that they provide a subleading effect, especially at weak coupling $q\alpha \ll 1$.

We might worry that, over long times, a substantial enough continuum population can be built up so that the second term in (A10) can overcome its $\mathcal{O}(q^2\alpha^2)$ -suppression and compete with the first. However, this sort of coherent effect is extremely unlikely in light of the oscillatory factors in (A10), which serve to randomize the “direction” of this perturbation and suppress its effects on long time scales. These arguments can be trivially extended to the more realistic case discussed in the next section, so we will ignore continuum-to-continuum transitions throughout our analysis

and focus only on how the bound states interact with the continuum.

3. Extension to the realistic case

The main complication in going to the more realistic case is that there are many more bound and continuum states, and the continuum now mediates transitions between different bound states. These effects appear in the form of off-diagonal induced couplings,

$$\mathcal{E}_{ba}(t) = -i \sum_K \eta_{Kb}^{*(\Delta m_b)}(t) \eta_{Ka}^{(\Delta m_a)}(t) \int_{-\infty}^t dt' e^{i\Delta m_b \varphi_*(t) - i\Delta m_a \varphi_*(t') + i(\epsilon_b - \epsilon_K)t + i(\epsilon_K - \epsilon_a)t'}, \quad (\text{A12})$$

where we have introduced the shorthand $\Delta m_a \equiv m - m_a$ and $\Delta m_b \equiv m - m_b$. We would like to understand the general behavior of these off-diagonal terms and argue that they can be ignored whenever the resonance condition between the states $|a\rangle$ and $|b\rangle$ is not satisfied. On resonance, they provide a small correction compared to the direct coupling between these states and so they can be neglected.

Assuming that the frequency $\dot{\varphi}_*(t)$ is linear, we can again define the variable $z \equiv t - t'$ and write (A12) as

$$\mathcal{E}_{ba}(t) = e^{i(\epsilon_b - \epsilon_a)t - i(m_b - m_a)\varphi_*(t)} \left[-i \sum_K \int_0^\infty dz e^{-\frac{1}{2}i\Delta m_a \gamma z^2 + i(\Delta m_a \dot{\varphi}_*(t) - (\epsilon_K - \epsilon_a))z} \eta_{Kb}^{*(\Delta m_b)}(t) \eta_{Ka}^{(\Delta m_a)}(t) \right]. \quad (\text{A13})$$

The term in braces is of a similar form to the induced energy (A3), whose behavior we have already analyzed in (A9). It contains both oscillating and smoothly decaying terms. Ignoring these oscillating terms for now, we see that the induced couplings oscillate rapidly with phase $\exp[i(\epsilon_b - \epsilon_a)t - i(m_b - m_a)\varphi_*(t)]$. As we argue in Sec. III B, the direct couplings between $|a\rangle$ and $|b\rangle$ also oscillate with this phase, and if these oscillations are too rapid the contribution to the bound state solution will quickly average out. Of course, this oscillation slows down when the resonance condition $(m_b - m_a)\dot{\varphi}_*(t) = (\epsilon_b - \epsilon_a)$ is satisfied, but again these induced couplings, which are $\mathcal{O}(q^2\alpha^2)$, must compete with the $\mathcal{O}(q\alpha)$ direct couplings η_{ba} , and so even then they have a small effect on the behavior of the resonance for $q\alpha \ll 1$.

We might worry about the oscillations that arise in (A9) as transients when the state $|a\rangle$ begins to resonate with the continuum might spoil this story, and that these induced couplings might become relevant. Fortunately, this is not the case. These transient oscillations “start” when the

companion can excite $|a\rangle$ into the continuum, $\Delta m_a \dot{\varphi}_*(t) = -\epsilon_a$, and if they are present they modify the overall exponential in (A13) to

$$\exp[-i(m_b - m_a)\varphi_*(t) + i(\epsilon_b - \epsilon_a)t + i(\Delta m_a \dot{\varphi}_*(t) + \epsilon_a)^2 / (2\Delta m_a \gamma)]. \quad (\text{A14})$$

This term can contribute appreciably when the argument of the exponential slows down, that is when $\Delta m_b \dot{\varphi}_*(t) = -\epsilon_b$. The two conditions $\Delta m_i \dot{\varphi}_*(t) = -\epsilon_i$, for $i = a, b$, can only simultaneously be satisfied when $(m_b - m_a)\dot{\varphi}_*(t) = \epsilon_b - \epsilon_a$, i.e., exactly on resonance. So, the transient oscillatory terms in (A9) may “smear out” the resonance slightly, but again since they are $\mathcal{O}(q^2\alpha^2)$ and must compete with the $\mathcal{O}(q\alpha)$ direct couplings $\eta_{ba}(t)$, we do not expect that they provide a qualitative change in behavior in the dynamics, and away from resonance we can ignore the induced couplings entirely.

With this out of the way, we can focus entirely on the diagonal terms, $\mathcal{E}_b(t) \equiv \mathcal{E}_{bb}(t)$, which are much simpler,

$$\begin{aligned} \mathcal{E}_b(t) &= -i \int_{-\infty}^t dt' \sum_K |\eta_{Kb}^{(\Delta m_b)}(t)|^2 e^{i\Delta m_b(\varphi_*(t) - \varphi_*(t')) - i(\epsilon_K - \epsilon_b)(t - t')} \\ &= \frac{1}{2\pi i} \sum_{\ell, m} \int_{-\infty}^t dt' \int_0^\infty dk |\eta_{Kb}^{(\Delta m_b)}(t)|^2 e^{i\Delta m_b(\varphi_*(t) - \varphi_*(t')) - i(\epsilon(k) - \epsilon_b)(t - t')}. \end{aligned} \quad (\text{A15})$$

This is nothing more than a sum over integrals of the form we have already analyzed, and we can use the same techniques as before to attack this. In particular, the integral over t' yields

$$\begin{aligned} \mathcal{E}_b(t) = & \frac{1}{2\pi} \sum_{\ell, m} \sqrt{\frac{\pi}{2\Delta m_b \gamma}} \int_0^\infty dk |\eta_{Kb}^{(\Delta m_b)}(t)|^2 \exp\left(\frac{i(\Delta m_b \dot{\varphi}_*(t) - (\epsilon(k) - \epsilon_b))^2}{2\Delta m_b \gamma} - \frac{3\pi i}{4}\right) \\ & \times \left[\text{sgn} \Delta m_b \gamma + \text{erf}\left(\frac{e^{\frac{i\pi}{4}}(\Delta m_b \dot{\varphi}_*(t) - (\epsilon(k) - \epsilon_b))}{\sqrt{2\Delta m_b \gamma}}\right) \right]. \end{aligned} \quad (\text{A16})$$

As discussed previously, we can think of the imaginary part as getting a saddle point contribution at $k_*^{(g)}(t) = \sqrt{2\mu(g\dot{\varphi}_*(t) + \epsilon_b)}$, which again only contributes if $k_*^{(g)}(t)^2 > 0$. For this to ever happen (since $\epsilon_b < 0$), we must have that $\Delta m_b \gamma = (m - m_b)\gamma > 0$. Thus, ignoring the oscillatory terms and other transients, we have

$$\mathcal{E}_b(t) \approx - \sum_{\ell, g} \left[\frac{i\mu |\eta_{K_*b}^{(g)}(t)|^2}{2k_*^{(g)}(t)} \Theta(k_*^{(g)}(t)^2) \right], \quad (\text{A17})$$

with $K_* = \{k_*^{(g)}(t), \ell, m = g + m_b\}$ and $k_*^{(g)}(t) = \sqrt{2\mu(g\dot{\varphi}_*(t) + \epsilon_b)}$, where the sum ranges from $\ell = 0, 1, \dots, \infty$ and over all g such that $|g + m_b| \leq \ell$. This is the extension of the first term in (A9) to include other sectors of continuum states, with different angular momenta, connected to the bound state by perturbations that oscillate at different frequencies.

4. Stationary perturbation theory

We can get a better sense for the origin of the first term in (A9) by deriving it via stationary perturbation theory. We start with the toy Hamiltonian (3.1), with $\varphi_*(t) = \Omega_0 t$, so that

$$\begin{aligned} \mathcal{H} = & \epsilon_b |b\rangle\langle b| + \frac{1}{2\pi} \int_0^\infty dk [\eta(k) e^{-i\Omega_0 t} |k\rangle\langle b| \\ & + \eta^*(k) e^{i\Omega_0 t} |b\rangle\langle k| + \epsilon(k) |k\rangle\langle k|]. \end{aligned} \quad (\text{A18})$$

The transition rate from the bound to the unbound states is then computed with Fermi's Golden Rule, which states that the transition probability per unit time per unit phase space volume is

$$d\Gamma = 2\pi |\eta(k)|^2 \delta(\epsilon(k) - \epsilon_b - \Omega_0) \frac{dk}{2\pi}. \quad (\text{A19})$$

Using $\epsilon(k) = k^2/(2\mu)$, the fractional change in the bound state population is

$$\frac{d \log |c_b(t)|^2}{dt} = - \int d\Gamma = - \frac{\mu |\eta(k_*)|^2}{k_*} \Theta(k_*^2), \quad (\text{A20})$$

where $k_* = \sqrt{2\mu(\Omega_0 + \epsilon_b)}$ and the Θ function ensures that this is only nonzero when k_* is real. This is the same as (3.10), with $\gamma = 0$, and is equivalent to the quantum mechanical derivation of the cross section in the photoelectric effect.

We see that the first term in (A9) has a simple interpretation—it represents the “steady state” deoccupation of the bound state into the continuum that is captured by assuming the perturbation's frequency does not change in time. We can extend this to the case of interest by adiabatically increasing the frequency $\dot{\varphi}_*(t) = \Omega_0 + \gamma t$ in (A20). It is not clear from Fermi's Golden Rule how slowly this frequency change needs to be in order for (A20) to be valid, but we see from (A9) that this stationary picture accurately captures the most important aspect of the true dynamics we use throughout the main text.

In the realistic case, the companion connects the states $|b\rangle$ and $|K\rangle$, each with definite azimuthal angular momentum m_b and m , respectively, with a perturbation that oscillates with definite frequency, $\eta_{Kb} \propto \exp[-i(m - m_b)\varphi_*(t)]$. It is trivial to extend the above discussion to the case where there are many such decay channels for the bound state, in which case we sum (A20) over all of them. Once we adiabatically restore the frequency's time dependence, we find that this stationary perturbation theory approach recovers (A17).

5. Nonlinear chirp frequency

Throughout this work, we have assumed that we can linearize the frequency and write the phase as $\varphi_*(t) = -\epsilon_b t + \gamma t^2/2$. It will be useful to justify this approximation.

Let us return to (A1) and try to understand the behavior of the t' integral,

$$\int_{-\infty}^t dt' e^{i(\epsilon - \epsilon_b)t' - i\varphi_*(t')}, \quad (\text{A21})$$

for a phase $\varphi_*(t)$ with general time dependence. This integral has essentially two contributions. One comes from the end point, which we can isolate through integration by parts,

$$\int_{-\infty}^t dt' e^{i(\epsilon - \epsilon_b)t' - i\varphi_*(t')} \supset \frac{i e^{i(\epsilon - \epsilon_b)t - i\varphi_*(t)}}{\dot{\varphi}_*(t) - (\epsilon - \epsilon_b)} + \dots, \quad (\text{A22})$$

while another can arise if $\dot{\varphi}(t_*) = \epsilon - \epsilon_b$ for some $t' = t_*$ in the integration interval. When such a time exists, the integral receives an additional contribution

$$\int_{-\infty}^t dt' e^{i(\epsilon - \epsilon_b)t' - i\varphi_*(t')} \supset \sqrt{\frac{2\pi}{\dot{\varphi}_*(t_*)}} e^{i(\epsilon - \epsilon_b)t_* - i\varphi_*(t_*) - \frac{i\pi}{4}}, \quad (\text{A23})$$

which we should divide in half when $t = t_*$. We obtain a rough approximation for the t' integral,

$$\int_{-\infty}^t dt' e^{i(\epsilon - \epsilon_b)t' - i\varphi_*(t')} \approx \begin{cases} \frac{i e^{i(\epsilon - \epsilon_b)t - i\varphi_*(t)}}{\dot{\varphi}_*(t) - (\epsilon - \epsilon_b)}, & t < t_* \\ \sqrt{\frac{\pi}{2\dot{\varphi}_*(t_*)}} e^{i(\epsilon - \epsilon_b)t_* - i\varphi_*(t_*) - \frac{i\pi}{4}}, & t = t_* \\ \frac{i e^{i(\epsilon - \epsilon_b)t - i\varphi_*(t)}}{\dot{\varphi}_*(t) - (\epsilon - \epsilon_b)} + \sqrt{\frac{2\pi}{\dot{\varphi}_*(t_*)}} e^{i(\epsilon - \epsilon_b)t_* - i\varphi_*(t_*) - \frac{i\pi}{4}}, & t > t_* \end{cases} \quad (\text{A24})$$

by adding these different contributions.

If we use $\varphi_*(t) = -\epsilon_b t + \gamma t^2/2$ and consider the exact answer, we find that

$$\sqrt{\frac{\pi}{2\gamma}} e^{\frac{i\epsilon^2}{2\gamma} - \frac{i\pi}{4}} \text{erfc} \left[\frac{e^{\frac{i\pi}{4}}(\epsilon - \gamma t)}{\sqrt{2\gamma}} \right] \approx \begin{cases} \frac{i e^{-\frac{1}{2}\gamma t^2 + i\epsilon t}}{\gamma t}, & t \ll t_* \\ \sqrt{\frac{\pi}{2\gamma}} e^{\frac{i\epsilon^2}{2\gamma} - \frac{i\pi}{4}}, & t = t_* \\ \frac{i e^{-\frac{1}{2}\gamma t^2 + i\epsilon t}}{\gamma t} + \sqrt{\frac{2\pi}{\gamma}} e^{\frac{i\epsilon^2}{2\gamma} - \frac{i\pi}{4}}, & t \gg t_* \end{cases} \quad (\text{A25})$$

where $t_* = \epsilon(k)/\gamma$. We see that (A24) accurately captures the large $|t|$ asymptotics of the integral, and that the complicated error function is merely present to interpolate between these three regimes. Furthermore, the relevant chirp rate for the induced energy (A1) is just the instantaneous chirp rate $\dot{\varphi}_*(t)$ which we can, to excellent approximation, replace with the chirp rate defined in (2.20) associated to the frequency $\Omega_0 = -\epsilon_b$ of the energy gap between the bound state and the continuum.

We see then that the linearization of $\dot{\varphi}_*(t)$ is not such a dramatic approximation. The integrand in (A2) will still have a similar form as to the one considered there, and we would still be able to do a saddle point computation isolating the large $|z|$ asymptotics and get effectively the same results we have found in the main text, up to corrections in the (small) nonlinearities we have ignored.

APPENDIX B: MARKOV APPROXIMATION

In the main text, we studied how the cloud is ionized by first constructing an effective Schrödinger equation (3.21)

for the bound states, fully integrating out the dynamics of the continuum states and incorporating their effects in the induced couplings (3.22). This was valid in the so-called ‘‘Markov approximation,’’ which we justify in this appendix.

Let us review how the Markov approximation comes about for a single bound state interacting with the continuum. We argued in Sec. III B that we can ignore the continuum-induced interactions between the bound states off-resonance, and so this truncation to a single bound state still accurately captures the true dynamics of the system, especially when the orbital frequency is too high for any resonance to occur. By solving (3.17) for the continuum state amplitudes and plugging the result into (3.16), we arrive at a single equation for the bound state amplitude

$$i \frac{dc_b}{dt} = \int_{-\infty}^t dt' \Sigma_b(t, t') c_b(t'), \quad (\text{B1})$$

in terms of the self-energy

$$\Sigma_b(t, t') \equiv -i \sum_K \eta_{bK}(t) \eta_{Kb}(t') e^{-i(\epsilon_K - \epsilon_b)(t - t')}. \quad (\text{B2})$$

Assuming that the couplings between the continuum states vanish and ignoring the transitions into other bound states, this equation of motion is exact. We then implement the Markov approximation by first integrating by parts,

$$i \frac{dc_b}{dt} = \mathcal{E}_b(t) c_b(t) - \int_{-\infty}^t dt_1 \mathcal{E}_b(t, t_1) \frac{dc_b(t_1)}{dt_1}, \quad (\text{B3})$$

and dropping the second term, which we will argue can be neglected. Here, we have defined

$$\mathcal{E}_b(t, t') = \int_{-\infty}^{t'} dt_1 \Sigma_b(t, t_1), \quad (\text{B4})$$

and the induced energy $\mathcal{E}_b(t) \equiv \mathcal{E}_b(t, t)$.

Our goal now is to estimate the effect of the second term in (B3). To do this, we first strip off the first-order behavior by defining $\tilde{c}_b(t) = e^{i\varphi_b(t)} c_b(t)$, where $\varphi_b(t) = \int_{-\infty}^t dt_1 \mathcal{E}_b(t)$ is the time-dependent phase induced at first order by the continuum. Plugging this into (B3) yields

$$i \frac{d\tilde{c}_b(t)}{dt} = i \int_{-\infty}^t dt_1 e^{i\varphi_b(t) - i\varphi_b(t_1)} [\mathcal{E}_b(t, t_1) \mathcal{E}_b(t_1, t_1) \times \tilde{c}_b(t_1) + i \mathcal{E}_b(t, t_1) \dot{\tilde{c}}_b(t_1)]. \quad (\text{B5})$$

Defining the second-order induced energy

$$\mathcal{E}_b^{(2)}(t, t') = i \int_{-\infty}^{t'} dt_1 e^{i\varphi_b(t) - i\varphi_b(t_1)} \mathcal{E}_b(t, t_1) \mathcal{E}_b(t_1, t_1), \quad (\text{B6})$$

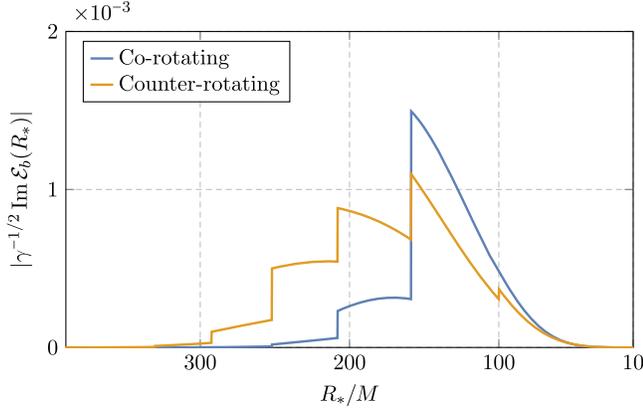


FIG. 15. The dimensionless ratio $|\gamma^{-1/2} \text{Im} \mathcal{E}_b(R_*)|$ as a function of the orbital separation R_* , using our approximation (3.28) as an estimate, for an inspiral with $q = 10^{-3}$ and $\alpha = 0.2$, where γ is the instantaneous chirp rate $\gamma = \dot{\varphi}_*(t)$, defined in (2.20) with $\Omega_0^2 R_*^3 = (1+q)M$.

with $\mathcal{E}_b^{(2)}(t) \equiv \mathcal{E}_b^{(2)}(t, t)$, integrating the first term in (B5) by parts, and dropping terms that contain factors of $d\tilde{c}_b/dt$, (B5) reduces to

$$i \frac{d\tilde{c}_b}{dt} = \mathcal{E}_b^{(2)}(t) \tilde{c}_b(t), \quad (\text{B7})$$

As long as we can argue that this contribution is small compared to the first-order motion, this step of dropping terms containing $d\tilde{c}_b/dt$ is consistent. In principle, we could also iterate this process to find ever more accurate approximations to the true dynamics.

It will be helpful to write the second-order induced energy as

$$\begin{aligned} \mathcal{E}_b^{(2)}(t) &= i \int_{-\infty}^t dt_1 e^{-\text{Im}[\varphi_b(t) - \varphi_b(t_1)] + i \text{Re}[\varphi_b(t) - \varphi_b(t_1)]} \\ &\quad \times \mathcal{E}_b(t, t_1) \mathcal{E}_b(t_1, t_1). \end{aligned} \quad (\text{B8})$$

Of particular importance is the oscillating phase factor, which depends on the real part of the induced phase

difference $\text{Re}[\varphi_b(t) - \varphi_b(t_1)]$. Contributions to this integral will cancel unless t_1 is close to t . Since the relevant time scale of the transition is

$$\gamma^{-1/2} = \sqrt{\frac{5}{96}} \frac{\alpha}{\mu} \frac{q^{-1/2}}{(1+q)^{3/4}} \left(\frac{\mu R_*}{\alpha} \right)^{1/4}, \quad (\text{B9})$$

we can think of $\mathcal{E}_b^{(2)}(t)$ as being on the same order as $\gamma^{-1/2} \mathcal{E}_b(t, t)^2$. These second-order corrections are thus small as long as $|\gamma^{-1/2} \mathcal{E}_b(t, t)^2| \ll |\mathcal{E}_b(t, t)|$. Since there is typically not a hierarchy between the real and imaginary parts of $\mathcal{E}_b(t)$, we can instead write this condition as $|\gamma^{-1/2} \text{Im} \mathcal{E}_b(t)| \ll 1$. We plot this quantity in Fig. 15 for the parameter values we consider in the main text and we see that it is comfortably small, so the Markov approximation is justified.

APPENDIX C: IONIZATION POWER

In this appendix, we justify our approximation of the ionization power $P_{\text{ion}} \equiv dE_{\text{ion}}/dt$ in the toy model of Sec. III A. The extension to the realistic case is conceptually trivial.

The total ionized energy is defined as

$$E_{\text{ion}}(t) = \frac{1}{2\pi} \frac{M_c}{\mu} \int_0^\infty dk (\epsilon(k) - \epsilon_b) |c_k(t)|^2, \quad (\text{C1})$$

where M_c/μ represents the total occupation number of the cloud. We will set this to one and restore it at the end of the calculation. By taking a single time derivative we can express the ionization power as

$$P_{\text{ion}} = \frac{1}{2\pi} \int_0^\infty dk (\epsilon(k) - \epsilon_b) [\dot{c}_k^*(t) c_k(t) + c_k^*(t) \dot{c}_k(t)], \quad (\text{C2})$$

and inserting both the Schrödinger equation (3.4) and the solution (3.5), we can find an equation of motion for the ionized energy purely in terms of the bound state

$$P_{\text{ion}} = \frac{1}{2\pi} \int_0^\infty dk \int_{-\infty}^t dt' [(\epsilon(k) - \epsilon_b) |\eta(k)|^2 e^{i(\varphi_*(t) - \varphi_*(t')) - i(\epsilon(k) - \epsilon_b)(t - t')} c_b^*(t) c_b(t') + \text{c.c.}]. \quad (\text{C3})$$

This has a very similar flavor to the effective bound state equation of motion (3.6), and we can implement the Markov approximation by integrating by parts and dropping the remainder,

$$P_{\text{ion}} = 2 \text{Re} \left[\frac{1}{2\pi} \int_0^\infty dk \int_{-\infty}^t dt' (\epsilon(k) - \epsilon_b) |\eta(k)|^2 e^{i(\varphi_*(t) - \varphi_*(t')) - i(\epsilon(k) - \epsilon_b)(t - t')} |c_b(t)|^2 \right]. \quad (\text{C4})$$

This equation of motion is very similar to (3.10), though now the term analogous to the induced energy $\mathcal{E}_b(t)$ is weighted with the energy difference $\epsilon(k) - \epsilon_b$.

This expression for the ionization power can be analyzed with the same techniques as used in Appendix A—ignoring the transient region around $\dot{\varphi}_*(t) + \epsilon_b = 0$ and the sub-leading oscillatory terms, we can approximate (C4) with its steady-state growth

$$P_{\text{ion}} \approx \frac{M_c}{\mu} \left[\frac{\mu \dot{\varphi}_*(t) |\eta(k_*(t))|^2}{k_*(t)} \right] |c_b(t)|^2 \Theta(k_*(t)), \quad (\text{C5})$$

where we have replaced $\epsilon(k_*(t)) - \epsilon_b = \dot{\varphi}_*(t)$.

One of the main benefits of the derivation of the deoccupation rate using stationary perturbation theory, presented in Sec. A 4, is that it makes inferring rates like the ionization power trivial. For instance, the amount of energy it takes to ionize the bound state $|b\rangle$ into the continuum state $|k; \ell m\rangle$ is $\epsilon(k) - \epsilon_b$. However, this only can happen if the perturbation's frequency matches this difference, $\epsilon(k) - \epsilon_b = g\dot{\varphi}_*$ with g an integer. The rate at which energy is ionized is then determined by the rate at which the bound state is ionized into the continuum state (A20), weighted by this energy difference and the total occupation of the bound state $|c_b(t)|^2$, and then summed over all different decay channels. This is the content of the final expression (3.29).

APPENDIX D: ZERO MODE

As we explained in the main text, the dramatic “discontinuous” behavior of the ionization power P_{ion} is due to the fact that the coupling function $|\eta(k)|^2$ goes to zero linearly in k as $k \rightarrow 0$. We mentioned there that this is because the long-range Coulombic potential keeps the zero mode relatively well localized about the origin, as illustrated in Fig. 16, such that the couplings in energy $|\eta(\epsilon)|^2 \equiv dk(\epsilon)/d\epsilon |\eta(k(\epsilon))|^2$ are finite as $\epsilon \rightarrow 0$. In this

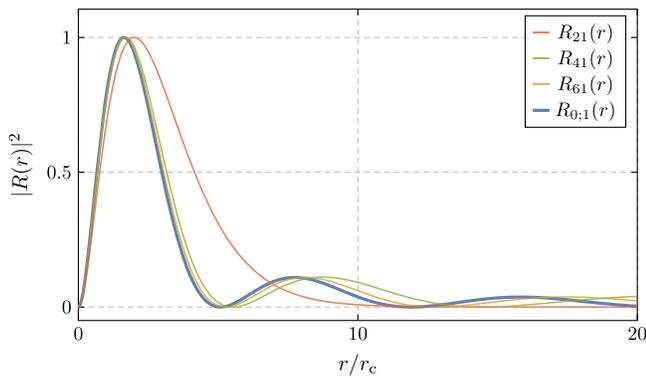


FIG. 16. The radial zero mode density $\lim_{k \rightarrow 0} |k^{-1/2} R_{k;1}(r)|^2$ compared to several bound state densities, all with orbital angular momentum $\ell = 1$. Here, $r_c = (\mu\alpha)^{-1}$ is the typical radius of the cloud, and we have normalized each density so that it has unit maximum. Ignoring the overall normalization, the zero-mode wave function can also be thought of as the limit of the bound state wave functions as $n \rightarrow \infty$.

appendix we discuss the zero mode of the hydrogen atom, its normalization, and the role the long-ranged $1/r$ potential plays in its radial behavior.

In order to determine the overall normalization of the zero mode, we begin by writing the normalized continuum radial wave functions (2.12) as

$$R_{k;\ell}(r) = \frac{2ki^\ell e^{\frac{\pi\mu\alpha}{2k}} |\Gamma(\ell + 1 + \frac{i\mu\alpha}{k})|}{(-2ikr)^{\frac{1}{2}} \Gamma(\ell + 1 + \frac{i\mu\alpha}{k})} \times e^{-ikr} \int_0^\infty d\zeta e^{-\zeta + \frac{i\mu\alpha}{k} \log \zeta} \zeta^{-\frac{1}{2}} J_{2\ell+1}(2\sqrt{-2ikr\zeta}), \quad (\text{D1})$$

where we have used a standard integral representation of the confluent hypergeometric function in terms of the Bessel function of the first kind $J_\nu(z)$. As $k \rightarrow 0$, the integral is localized around its saddle point $\zeta = i\mu\alpha/k$ and asymptotes to

$$R_{k;\ell}(r) \sim \sqrt{\frac{4\pi k}{r}} J_{2\ell+1}(2\sqrt{2\mu\alpha r}), \quad k \rightarrow 0. \quad (\text{D2})$$

It is then clear that any matrix element between a continuum state and a bound state will also scale as \sqrt{k} for $k \rightarrow 0$, so that $|\eta(k)|^2/k$ approaches a finite, nonzero limit as $k \rightarrow 0$.

We can understand this scaling in a less opaque way by considering the Schrödinger equation with a potential that asymptotes to a generic power law, $V(r) \sim 1/r^\Delta$ as $r \rightarrow \infty$, with $\Delta > 0$. Defining $\rho = 1/r$, the radial Schrödinger equation for a state with energy $\epsilon(k) = k^2/2\mu$ can then be written as

$$\left(-\frac{d^2}{d\rho^2} + \frac{\ell(\ell+1)}{\rho^2} - \frac{2\alpha\mu^2 \rho^\Delta}{\rho^4 \mu^\Delta} - \frac{k^2}{\rho^4} \right) R_{k;\ell}(\rho) = 0, \quad (\text{D3})$$

where we have introduced additional factors of μ to keep α dimensionless. We will only be concerned with the behavior of the solutions as $\rho \rightarrow 0$ or, analogously, as $r \rightarrow \infty$, so we have replaced the potential with its dominant long-distance behavior. If $\Delta > 2$, then the potential term is subleading to the centrifugal $\ell(\ell+1)/r^2$ term and the asymptotics of $R_{k;\ell}(\rho)$ are identical to that of a free particle.

For long-ranged potentials, $0 < \Delta < 2$, we can determine the overall normalization of the continuum wave functions as $k \rightarrow 0$ via a matching procedure. The basic idea is that the potential singularity $2\alpha\mu^{2-\Delta}/\rho^{4-\Delta}$ in (D3) dominates over the energy singularity k^2/ρ^4 in the region $\rho \gtrsim \mu[(k/\mu)^2/\alpha]^{1/\Delta}$. When ρ is smaller than this, the energy singularity dominates, so we can construct asymptotic approximations to $R_{k;\ell}(\rho)$ that are valid in these two different regions. When k is very small, the region $\rho \gtrsim \mu[(k/\mu)^2/\alpha]^{1/\Delta}$ comprises most of space, and so this is the relevant solution in the $k \rightarrow 0$ limit. However, the overall normalization of the continuum wave functions is set for

$\rho \lesssim \mu[(k/\mu)^2/\alpha]^{1/\Delta}$, and so we must deduce the overall normalization in the $k \rightarrow 0$ limit by matching. Our goal then is to first determine the asymptotic behavior of $R_{k;\ell}(\rho)$ around each of these singularities and then match them.

Depending on the value of Δ , the asymptotic behavior of $R_{k;\ell}(\rho)$ in the region near the energy singularity can be relatively complicated,

$$R_{k;\ell}(\rho) \sim A\rho \sin\left(\frac{k}{\rho} + \sum_{n=1}^{n\Delta \leq 1} \frac{(-\frac{1}{2}\alpha)^n (2n)!}{(2n-1)(n!)^2} \frac{(k/\mu)^{1-2n} \rho^{n\Delta-1}}{n\Delta-1} + \delta\right),$$

$$\rho \lesssim \mu \left[\frac{(k/\mu)^2}{\alpha} \right]^{1/\Delta}, \quad (\text{D4})$$

where the sum is over all n such that $n\Delta \leq 1$, and a $n\Delta = 1$ term should be understood to give a logarithmic correction. Here, A and δ are the overall normalization and phase, respectively. For example, the asymptotic behavior of wave functions for the Coulombic potential, with $\Delta = 1$, is

$$R_{k;\ell}(\rho) \sim A\rho \sin\left(\frac{k}{\rho} + \frac{\mu\alpha}{k} \log \frac{k}{\rho} + \delta\right), \quad (\text{D5})$$

and demanding these wave functions are appropriately normalized, $\langle k; \ell m | k'; \ell m \rangle = 2\pi\delta(k-k')$, sets the overall amplitude in this region to $A = 2$. In contrast, the asymptotic behavior of $R_{k;\ell}(\rho)$ in the region where the potential singularity dominates is relatively simple,

$$R_{k;\ell}(r) \sim A'\rho^{1-\Delta/4} \sin\left(\frac{2\sqrt{2\alpha}(\rho/\mu)^{\Delta-2}}{2-\Delta} + \delta'\right),$$

$$\rho \gtrsim \mu \left[\frac{(k/\mu)^2}{\alpha} \right]^{1/\Delta}, \quad (\text{D6})$$

where again A' and δ' are an undetermined amplitude and phase.

In the limit $k \rightarrow 0$, the region of (D4)'s validity, $\rho \lesssim \mu[(k/\mu)^2/\alpha]^{1/\Delta}$, shrinks to a point, and the continuum wave functions are well approximated by (D6) as $\rho \rightarrow 0$. However, we do not yet know its amplitude A' or, specifically, the k -scaling of its amplitude. We can determine this scaling by matching the amplitudes of (D4) and (D6) in the region where both expansions apply, $\rho \sim \mu[(k/\mu)^2/\alpha]^{1/\Delta}$. We find that the continuum wave functions then behave as

$$R_{k;\ell}(r) \propto \frac{\sqrt{k}}{r^{\frac{1}{4}(4-\Delta)}} \sin\left(\frac{2\sqrt{2\alpha}(\mu r)^{2-\Delta}}{2-\Delta} + \tilde{\delta}\right), \quad \begin{array}{l} k \rightarrow 0 \\ r \rightarrow \infty \end{array}, \quad (\text{D7})$$

for arbitrary $0 < \Delta < 2$, with $\tilde{\delta}$ an undetermined phase. As long as the potential is sufficiently long-ranged, $\Delta < 2$, the continuum wave functions therefore asymptote to a fixed radial function multiplied by an overall factor of \sqrt{k}

as $k \rightarrow 0$. This implies that, for $\Delta < 2$, the potential is sufficiently long-ranged enough to localize the zero mode. We can compare this general result with the asymptotic expansion of (D2), in which case $\Delta = 1$ and

$$R_{k;\ell}(r) \sim \frac{2\sqrt{k}}{(2\mu\alpha)^{1/4} r^{3/4}} \sin\left(2\sqrt{2\mu\alpha}r - \pi\ell - \frac{\pi}{4}\right), \quad \begin{array}{l} k \rightarrow 0 \\ r \rightarrow \infty \end{array}, \quad (\text{D8})$$

in agreement with our predicted scaling.

This scaling can be contrasted with that of a free particle. In this case, the effective potential due to angular momentum $\ell(\ell+1)/\rho^2$ dominates the $\rho \rightarrow 0$ limit, and, for $k \rightarrow 0$, the radial wave function behaves as

$$R_{k;\ell}(\rho) \sim C_1\rho^{\ell+1} + C_2\rho^{-\ell}, \quad \begin{array}{l} k \rightarrow 0 \\ \rho \rightarrow 0 \end{array}. \quad (\text{D9})$$

The appropriate $k \neq 0$ continuum wave functions are, instead, just the spherical Bessel functions,

$$R_{k;\ell}(\rho) = 2kj_\ell(k/\rho), \quad (\text{D10})$$

which obey the asymptotic scaling

$$R_{k;\ell}(\rho) \sim \frac{2^\ell \ell! k}{(\ell + \frac{1}{2})(2\ell)!} \left(\frac{k}{\rho}\right)^\ell, \quad k \rightarrow 0. \quad (\text{D11})$$

Unlike for potentials with $0 < \Delta < 2$, these continuum wave functions do not have a normalization that scales as \sqrt{k} as $k \rightarrow 0$, and indeed are *not* localized near the origin. We see that $\Delta = 2$ represents a qualitative dividing line in the behavior of the continuum modes in the $k \rightarrow 0$ limit. The matrix elements between a bound state and the zero mode of a potential with $\Delta \geq 2$ obeys $|\eta(k)|^2/k \rightarrow 0$, while this approaches a finite limit for potentials with $0 < \Delta < 2$.

APPENDIX E: MORE ON SCALARS AROUND KERR

The aim of this appendix is to present self-contained overview of the exact solutions for the definite frequency modes of a massive scalar field around a Kerr black hole.

1. Definite frequency solutions

The Kerr geometry has two relevant isometries; time translations and azimuthal rotations. This suggests that we choose an ansatz for the scalar field profile, with a definite frequency, ω , and azimuthal angular momentum, $m \in \mathbb{Z}$,

$$\Phi(t, \mathbf{r}) = e^{-i\omega t + im\phi} R(r)S(\theta). \quad (\text{E1})$$

It is a special property of the Kerr background that this ansatz separates the Klein-Gordon equation (2.1) into the angular spheroidal equation

$$\left(-\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d}{d\theta}\right)-k^2a^2\cos^2\theta+\frac{m^2}{\sin^2\theta}\right)S(\theta)=\lambda S(\theta), \quad (\text{E2})$$

and the radial equation

$$0 = \frac{1}{\Delta R} \frac{d}{dr} \left(\Delta \frac{dR}{dr} \right) + k^2 + \frac{P_+^2}{(r-r_+)^2} + \frac{P_-^2}{(r-r_-)^2} - \frac{A_+}{(r_+ - r_-)(r-r_+)} + \frac{A_-}{(r_+ - r_-)(r-r_-)}, \quad (\text{E3})$$

where we have introduced $k^2 = \omega^2 - \mu^2$, the eigenvalue of the spheroidal equation λ , and the parameter combinations

$$P_{\pm} = \frac{ma - 2M\omega r_{\pm}}{r_+ - r_-},$$

$$A_{\pm} = P_+^2 + P_-^2 + \gamma_{\pm}^2 + \lambda, \quad (\text{E4})$$

with $\gamma_{\pm}^2 = \mu^2 r_{\pm}^2 - \omega^2(4M^2 + 2Mr_{\pm} + r_{\pm}^2)$.

Requiring the solution to be regular at $\theta = 0$ and π , forces the spheroidal eigenvalue $\lambda = \lambda_{\ell m}(c)$ to take a set of discrete values, depending on the spheroidicity parameter $c = ka$ and labeled by $\ell = 0, 1, \dots$ and $|m| \leq \ell$. The corresponding angular functions $S(\theta) = S_{\ell m}(c; \cos\theta)$ are the ‘‘spheroidal harmonics,’’ which reduce to the ordinary spherical harmonics for $c = 0$.

The radial equation (E3) has three singularities; one at the outer horizon $r = r_+$ controlled by the parameter P_+^2 , one at the inner horizon $r = r_-$ controlled by P_-^2 , and an irregular singularity at $r = \infty$ controlled by k^2 , which can be understood as the confluence of two regular singularities. This uniquely identifies the radial equation as a form of the ‘‘confluent Heun equation’’, and we expect the radial solutions $R(r)$ to be proportional to the confluent Heun function, which we will define now.¹⁶

Our goal is to find solutions on $r \in [r_+, \infty)$ that are purely ingoing at the outer horizon $r = r_+$, since no physical mode can escape from the black hole. Near the

¹⁶It is useful to compare this to the Schrödinger equation of the hydrogen atom, which has both a regular singularity at $r = 0$ and an irregular singularity at $r = \infty$ that can also be understood as the confluence of two regular singularities. Any linear differential equation with three regular singular points can be mapped to the hypergeometric equation with singularities at $z = 0, 1$ and ∞ . The solution to this equation that is regular about $z = 0$ is the familiar hypergeometric function ${}_2F_1(a, b; c; z)$. Upon the confluence of the singularities at $z = 1$ and $z = \infty$, this turns into the confluent hypergeometric equation, and the regular solution ${}_2F_1(a, b; c; z)$ turns into the confluent hypergeometric function ${}_1F_1(a; c; z)$. An analogous story applies to the radial equation in the Kerr background, except it has an additional regular singularity at the inner horizon $r = r_-$. Any linear differential equation with four regular singular points can be mapped to the Heun equation, and upon a confluence of two singularities this reduced to the confluent Heun equation.

outer horizon, the singularity forces solutions to behave as $R(r) \sim (r - r_+)^{\pm iP_+}$, where the plus sign in the exponent corresponds to purely ingoing modes. Similarly, the singularity at $r = \infty$ forces the modes to behave as $R(r) \sim e^{\pm ikr}$. It will be convenient to define $z \equiv -(r - r_+)/(r_+ - r_-)$ and peel these asymptotic behaviors from the solution,

$$R(r) = e^{-ik(r-r_+)} z^{iP_+} (z-1)^{-iP_-} H(z). \quad (\text{E5})$$

The function $H(z)$ then satisfies the confluent Heun equation [58,59],

$$\frac{d^2 H}{dz^2} + \left(\alpha + \frac{1+\beta}{z} + \frac{1+\gamma}{z-1} \right) \frac{dH}{dz} + \left(\frac{\mu}{z} + \frac{\nu}{z-1} \right) H = 0, \quad (\text{E6})$$

where

$$\mu = \frac{1}{2}(\alpha - \beta - \gamma + \alpha\beta - \beta\gamma) - \eta,$$

$$\nu = \frac{1}{2}(\alpha + \beta + \gamma + \alpha\gamma + \beta\gamma) + \delta + \eta, \quad (\text{E7})$$

with $\alpha = 2ik(r_+ - r_-)$, $\beta = 2iP_+$, $\gamma = -2iP_-$, $\delta = A_+ - A_-$, and $\eta = -A_+$. Equation (E6) has a solution that is regular at the origin, $H(0) = 1$, called the confluent Heun function, $H(z) = \text{HeunC}(\alpha, \beta, \gamma, \delta, \eta; z)$, and one which behaves as z^{-2iP_-} as $z \rightarrow 0$. Since we impose purely ingoing boundary conditions, we discard the latter and find that

$$\begin{aligned} \Phi(t, \mathbf{r}) &= R_{k;\ell m}(r) S_{\ell m}(ka; \cos\theta) e^{-i\omega t + im\phi} \\ &= \mathcal{C} e^{-i\omega t - ik(r-r_+) + im\phi} z^{iP_+} (z-1)^{-iP_-} \\ &\quad \times \text{HeunC}(\alpha, \beta, \gamma, \delta, \eta; z) S_{\ell m}(ka; \cos\theta), \end{aligned} \quad (\text{E8})$$

where \mathcal{C} is a normalization constant.

Using the tortoise coordinates,

$$\begin{aligned} \tilde{r} &= \frac{2M}{r_+ - r_-} \left[r_+ \log\left(\frac{r-r_+}{r_+ - r_-}\right) - r_- \log\left(\frac{r-r_-}{r_+ - r_-}\right) \right] + r, \\ \tilde{\phi} &= \frac{a}{r_+ - r_-} \left[\log\left(\frac{r-r_+}{r_+ - r_-}\right) - \log\left(\frac{r-r_-}{r_+ - r_-}\right) \right], \end{aligned} \quad (\text{E9})$$

the solution can be written as

$$\begin{aligned} \Phi(t, \mathbf{r}) &= \mathcal{C} e^{-ik(r-r_+) - i\omega(t+\tilde{r}-r) + im(\phi+\tilde{\phi})} \\ &\quad \times \text{HeunC}(\alpha, \beta, \gamma, \delta, \eta; z) S_{\ell m}(ka; \cos\theta). \end{aligned} \quad (\text{E10})$$

Since the combination $\tilde{r} - r$ increases as we move away from the outer horizon, this mode indeed represents a purely ingoing wave.

There are two classes of solutions that we use throughout the main text. The first are the quasi-bound states, which

are purely ingoing at the outer horizon and exponentially decaying as $r \rightarrow \infty$. These two boundary conditions can only be satisfied for a discrete set of frequencies $\omega_{n\ell m} = E_{n\ell m} + i\Gamma_{n\ell m}$, cf. (2.8), and so these mode only come in a discrete set. The second are the unbound continuum states, which are purely ingoing at the outer horizon, but oscillate as $r \rightarrow \infty$. Since we impose only one boundary condition, these unbound modes comprise a continuous set with frequencies $\omega^2 = \mu^2 + k^2$.

2. Nonrelativistic limit

The first four parameters of the confluent Heun function are either first order (α, β, γ) or second order (δ) in the dimensionless combinations μM and kM . The fifth parameter, on the other hand, is generally $\eta = \mathcal{O}(1)$, because

$$\lambda_{\ell m}(c) = \ell(\ell+1) - \frac{1}{2} \left[1 - \frac{(2m-1)(2m+1)}{(2\ell-1)(2\ell+3)} \right] c^2 + \mathcal{O}(c^4). \quad (\text{E11})$$

The only exception is when $\ell = 0$, where η is second order in both μM and kM . Modes with nonzero angular momentum see a centrifugal barrier which forces the field away from the black hole, suppressing its amplitude at radii below $\sim \ell^2/(\mu^2 M)$. This is not the case for the $\ell = 0$ mode, whose amplitude is not suppressed near the horizon.

In the main text, we need the profile of the $\ell = 0$ mode in the nonrelativistic ($kM \ll 1$) and fuzzy ($\mu M \ll 1$) limits. In this case, the confluent Heun function can be expanded to second order in μM and kM , but at fixed z , as^{17,18}

$$\begin{aligned} \text{HeunC}(\alpha, \beta, \gamma, \delta, \eta; z) &= 1 - \frac{1}{2}\alpha z + \frac{1}{6}\alpha^2 z^2 - \frac{1}{24}(\alpha^2 + 12\delta)z + \frac{1}{4}(\alpha\beta + \alpha\gamma)z \log(1-z) \\ &\quad - \frac{1}{2}(\beta + \gamma) \log(1-z) + \frac{1}{4}(\gamma^2 - \beta^2) \text{dilog}(1-z) + \frac{1}{4}(\beta\gamma + \gamma^2) \log^2(1-z) \\ &\quad - \frac{1}{24}(\alpha^2 - 6\beta^2 - 6\gamma^2 + 24\eta + 12\delta) \log(1-z) + \dots \end{aligned} \quad (\text{E12})$$

In the first line, we have grouped terms that are dominant as $z \rightarrow -\infty$, while the next two lines contain terms that are subdominant and can be ignored. Given that $\alpha \sim \mathcal{O}(kM)$ and $\delta \sim \beta^2 \sim \gamma^2 \sim \mathcal{O}(\mu^2 M^2)$, we see that the confluent Heun function is approximately constant

$$\text{HeunC}(\alpha, \beta, \gamma, \delta, \eta; z) \sim 1 + \mathcal{O}(\mu M, kM), \quad r_+ \leq r < r_{\max}, \quad (\text{E13})$$

until the linear or quadratic terms in the first line of (E12) become $\mathcal{O}(1)$. This occurs at the radius

$$\frac{r_{\max}}{M} \sim \min \left\{ \frac{1}{(\mu M)^2}, \frac{1}{kM} \right\} \gg 1. \quad (\text{E14})$$

We use this approximation to derive the accretion rate in Sec. IV.

¹⁷Here, $\text{dilog}(1-z) = \text{Li}_2(z) = \sum_{n=1}^{\infty} z^n/n^2$.

¹⁸The procedure consists in finding a recurrence relation among the coefficients of the power series $\text{HeunC}(\alpha, \beta, \gamma, \delta, \eta; z) = \sum_{n=0}^{\infty} a_n z^n$, of the form $P_n a_n = Q_n a_{n-1} + R_n a_{n-2}$, see e.g., [58,60]. After solving it to second order in α, β, γ , and first order in δ, η , the series can be resummed to give (E12).

[1] L. Barack *et al.*, Black holes, gravitational waves and fundamental physics: A roadmap, *Classical Quantum Gravity* **36**, 143001 (2019).
 [2] E. Barausse *et al.*, Prospects for fundamental physics with LISA, *Gen. Relativ. Gravit.* **52**, 81 (2020).

[3] G. Bertone *et al.*, Gravitational wave probes of dark matter: Challenges and opportunities, *SciPost Phys. Core* **3**, 007 (2020).
 [4] H.S. Chia, Probing particle physics with gravitational waves. Ph.D. thesis, Amsterdam University, 2020, [arXiv: 2012.09167](https://arxiv.org/abs/2012.09167).

- [5] R. Brito, V. Cardoso, and P. Pani, Superradiance: New frontiers in black hole physics, *Lect. Notes Phys.* **906**, 1 (2015).
- [6] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, String axiverse, *Phys. Rev. D* **81**, 123530 (2010).
- [7] A. Arvanitaki and S. Dubovsky, Exploring the string axiverse with precision black hole physics, *Phys. Rev. D* **83**, 044026 (2011).
- [8] W. East and F. Pretorius, Superradiant Instability and Backreaction of Massive Vector Fields around Kerr Black Holes, *Phys. Rev. Lett.* **119**, 041101 (2017).
- [9] W. East, Massive Boson Superradiant Instability of Black Holes: Nonlinear Growth, Saturation, and Gravitational Radiation, *Phys. Rev. Lett.* **121**, 131104 (2018).
- [10] R. Abbott *et al.* (LIGO/VIRGO/KAGRA Collaborations), All-sky search for gravitational wave emission from scalar boson clouds around spinning black holes in LIGO O3 data, *Phys. Rev. D* **105**, 102001 (2022).
- [11] R. Brito, V. Cardoso, and P. Pani, Black holes as particle detectors: Evolution of superradiant instabilities, *Classical Quantum Gravity* **32**, 134001 (2015).
- [12] R. Brito, S. Ghosh, E. Barausse, E. Berti, V. Cardoso, I. Dvorkin, A. Klein, and P. Pani, Gravitational wave searches for ultralight bosons with LIGO and LISA, *Phys. Rev. D* **96**, 064050 (2017).
- [13] K. Ng, O. Hannuksela, S. Vitale, and T. Li, Searching for ultralight bosons within spin measurements of a population of binary black hole mergers, *Phys. Rev. D* **103**, 063010 (2021).
- [14] K. Ng, S. Vitale, O. Hannuksela, and T. Li, Constraints on Ultralight Scalar Bosons within Black Hole Spin Measurements from the LIGO-Virgo GWTC-2, *Phys. Rev. Lett.* **126**, 151102 (2021).
- [15] N. Fernandez, A. Ghalsasi, and S. Profumo, Superradiance and the spins of black holes from LIGO and x-ray binaries, [arXiv:1911.07862](https://arxiv.org/abs/1911.07862).
- [16] D. Baumann, H. S. Chia, and R. Porto, Probing ultralight bosons with binary black holes, *Phys. Rev. D* **99**, 044001 (2019).
- [17] J. Zhang and H. Yang, Dynamic signatures of black hole binaries with superradiant clouds, *Phys. Rev. D* **101**, 043020 (2020).
- [18] E. Berti, R. Brito, C. Macedo, G. Raposo, and J. Rosa, Ultralight boson cloud depletion in binary systems, *Phys. Rev. D* **99**, 104039 (2019).
- [19] J. Zhang and H. Yang, Gravitational floating orbits around hairy black holes, *Phys. Rev. D* **99**, 064018 (2019).
- [20] B. Su, Z.-Z. Xianyu, and X. Zhang, Probing ultralight bosons with compact eccentric binaries, *Astrophys. J.* **923**, 114 (2021).
- [21] X. Tong, Y. Wang, and H.-Y. Zhu, Gravitational collider physics via pulsar-black hole binaries II: Fine and hyperfine structures are favored, *Astrophys. J.* **924**, 99 (2022).
- [22] Q. Ding, X. Tong, and Y. Wang, Gravitational collider physics via pulsar-black hole binaries, *Astrophys. J.* **908**, 78 (2021).
- [23] T. Takahashi and T. Tanaka, Axion clouds may survive the perturbative tidal interaction over the early inspiral phase of black hole binaries, *J. Cosmol. Astropart. Phys.* **10** (2021) 031.
- [24] S. Choudhary, N. Sanchis-Gual, A. Gupta, J. Degollado, S. Bose, and J. Font, Gravitational waves from binary black hole mergers surrounded by scalar field clouds: Numerical simulations and observational implications, *Phys. Rev. D* **103**, 044032 (2021).
- [25] Y. Asali, P. Pang, A. Samajdar, and C. Van Den Broeck, Probing resonant excitations in exotic compact objects via gravitational waves, *Phys. Rev. D* **102**, 024016 (2020).
- [26] T. Ikeda, L. Bernard, V. Cardoso, and M. Zilhão, Black hole binaries and light fields: Gravitational molecules, *Phys. Rev. D* **103**, 024020 (2021).
- [27] L. Wong, Evolution of diffuse scalar clouds around binary black holes, *Phys. Rev. D* **101**, 124049 (2020).
- [28] V. Cardoso, F. Duque, and T. Ikeda, Tidal effects and disruption in superradiant clouds: A numerical investigation, *Phys. Rev. D* **101**, 064054 (2020).
- [29] L. Landau, Zur Theorie der Energieübertragung. II, *Phys. Z. Sowjetunion* **2**, 46 (1932).
- [30] C. Zener, Non-adiabatic crossing of energy levels, *Proc. R. Soc. A* **137**, 696 (1932).
- [31] D. Baumann, H. S. Chia, R. Porto, and J. Stout, Gravitational collider physics, *Phys. Rev. D* **101**, 083019 (2020).
- [32] D. Baumann, H. S. Chia, J. Stout, and L. ter Haar, The spectra of gravitational atoms, *J. Cosmol. Astropart. Phys.* **12** (2019) 006.
- [33] P. Peters and J. Mathews, Gravitational radiation from point masses in a Keplerian orbit, *Phys. Rev.* **131**, 435 (1963).
- [34] Y. Demkov and V. Osherov, Stationary and nonstationary problems in quantum mechanics that can be solved by means of contour integration, *Sov. J. Exp. Theor. Phys.* **26**, 916 (1968).
- [35] D. Basko, Landau-Zener-Stückelberg Physics with a Singular Continuum of States, *Phys. Rev. Lett.* **118**, 016805 (2017).
- [36] V. Weisskopf and E. Wigner, Berechnung der Natürlichen Linienbreite auf Grund der Diracschen Lichttheorie, *Z. Phys.* **63**, 54 (1930).
- [37] M. Scully and M. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, England, 1997).
- [38] N. Herring, B. Pardo, D. Boyanovsky, and A. Zentner, Particle decay in post-inflationary cosmology, *Phys. Rev. D* **98**, 083503 (2018).
- [39] S. Das, G. Gibbons, and S. Mathur, Universality of Low-Energy Absorption Cross-Sections for Black Holes, *Phys. Rev. Lett.* **78**, 417 (1997).
- [40] A. Higuchi, Low frequency scalar absorption cross-sections for stationary black holes, *Classical Quantum Gravity* **18**, L139 (2001); **19**, 599(A) (2002).
- [41] C. Macedo, L. Leite, E. Oliveira, S. Dolan, and L. Crispino, Absorption of planar massless scalar waves by Kerr black holes, *Phys. Rev. D* **88**, 064033 (2013).
- [42] V. Cardoso and R. Vicente, Moving black holes: Energy extraction, absorption cross-section and the ring of fire, *Phys. Rev. D* **100**, 084001 (2019).
- [43] W. Unruh, Absorption cross-section of small black holes, *Phys. Rev. D* **14**, 3251 (1976).
- [44] C. Benone, E. de Oliveira, S. Dolan, and L. Crispino, Absorption of a massive scalar field by a charged black hole, *Phys. Rev. D* **89**, 104053 (2014).

- [45] C. Benone and L. Crispino, Massive and charged scalar field in Kerr-Newman spacetime: Absorption and superradiance, *Phys. Rev. D* **99**, 044009 (2019).
- [46] V. Bezerra, H. Vieira, and A. Costa, The Klein-Gordon equation in the spacetime of a charged and rotating black hole, *Classical Quantum Gravity* **31**, 045003 (2014).
- [47] L. Hui, J. Ostriker, S. Tremaine, and E. Witten, Ultralight scalars as cosmological dark matter, *Phys. Rev. D* **95**, 043541 (2017).
- [48] D. Traykova, K. Clough, T. Helfer, P. Ferreira, E. Berti, and L. Hui, Dynamical friction from scalar dark matter in the relativistic regime, *Phys. Rev. D* **104**, 103014 (2021).
- [49] L. Annulli, V. Cardoso, and R. Vicente, Stirred and shaken: Dynamical behavior of boson stars and dark matter cores, *Phys. Lett. B* **811**, 135944 (2020).
- [50] T. Takahashi, H. Omiya, and T. Tanaka, Axion cloud evaporation during inspiral of black hole binaries—The effects of backreaction and radiation, *Prog. Theor. Exp. Phys.* **2022**, 043E01 (2022).
- [51] V. Mehta, M. Demirtas, C. Long, D.J.E. Marsh, L. McAllister, and M. Stott, Superradiance in string theory, *J. Cosmol. Astropart. Phys.* **07** (2021) 033.
- [52] V. Mehta, M. Demirtas, C. Long, D.J.E. Marsh, L. McAllister, and M. Stott, Superradiance exclusions in the landscape of type IIB string theory, [arXiv:2011.08693](https://arxiv.org/abs/2011.08693).
- [53] M. Demirtas, C. Long, L. McAllister, and M. Stillman, The Kreuzer-Skarke axiverse, *J. High Energy Phys.* **04** (2020) 138.
- [54] K. Eda, Y. Itoh, S. Kuroyanagi, and J. Silk, New Probe of Dark-Matter Properties: Gravitational Waves from an Intermediate-Mass Black Hole Embedded in a Dark-Matter Minispike, *Phys. Rev. Lett.* **110**, 221101 (2013).
- [55] K. Eda, Y. Itoh, S. Kuroyanagi, and J. Silk, Gravitational waves as a probe of dark matter minispikes, *Phys. Rev. D* **91**, 044045 (2015).
- [56] B. Kavanagh, D. Nichols, G. Bertone, and D. Gaggero, Detecting dark matter around black holes with gravitational waves: Effects of dark matter dynamics on the gravitational waveform, *Phys. Rev. D* **102**, 083006 (2020).
- [57] A. Coogan, G. Bertone, D. Gaggero, B. Kavanagh, and D. Nichols, Measuring the dark matter environments of black hole binaries with gravitational waves, *Phys. Rev. D* **105**, 043009 (2022).
- [58] P. Ronveaux *et al.*, *Heun's Differential Equations* (Oxford University Press, New York, 1995).
- [59] P. Fiziev, Classes of exact solutions to Regge-Wheeler and Teukolsky equations, [arXiv:0902.1277](https://arxiv.org/abs/0902.1277).
- [60] L. Hui, D. Kabat, X. Li, L. Santoni, and S.S.C. Wong, Black hole hair from scalar dark matter, *J. Cosmol. Astropart. Phys.* **06** (2019) 038.