Massive Boson Superradiant Instability of Black Holes: Nonlinear Growth, Saturation, and Gravitational Radiation

William E. East
Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada

We study the superradiant instability of a massive boson around a spinning black hole in full general relativity without assuming spatial symmetries. We focus on the case of a rapidly spinning black hole with a Compton wavelength comparable to the black hole radius, which is the regime where relativistic effects are maximized. We follow the growth of the boson cloud through superradiance into the nonlinear regime as it spins down the black hole, reaches a maximum energy, and begins to dissipate through the emission of gravitational waves. We find that the superradiant instability can efficiently convert a significant fraction of a black hole’s rotational energy into gravitational radiation.

Introduction.—The advent of gravitational wave (GW) astronomy [1] not only provides a window into the dynamics of merging black hole (BH) and neutron star binaries, but also furnishes a tool to look for new fundamental physics. One example of beyond standard model particle physics that could be explored is ultralight bosons that are weakly coupled to ordinary matter, including the QCD axion [2], the string axiverse [3, 4], and dark photons [5, 6]. Massive bosons are unstable in the presence of spinning BHs, and can form clouds that grow exponentially through superradiance [7–9], potentially reaching masses up to \( \sim 10\% \) of the BH, and spinning the BH down in the process [4, 10–12].

Thus, the existence of such bosonic fields can be probed indirectly by measuring the masses and spins of BHs through electromagnetic observations of accretion disks or GW observations of BH mergers [13–16]. One could also hope to directly observe the GWs coming from these oscillating boson clouds, either from resolved sources [13–15], or from a stochastic background [17, 18] with LIGO, LISA, or other GW detectors.

The goal of this work is to study the nonlinear growth and saturation of the superradiant instability of massive bosons in the regime where relativistic effects are maximized, both to understand the dynamics and GW signature in this regime, and to place an upper limit on the importance of nonlinear or “nonadiabatic” effects. To this end, we focus on the case of a massive vector field, which has a significantly faster growth rate and stronger gravitational radiation [19] than the scalar case, and consider its superradiant growth around a rapidly spinning BH with size comparable to the boson Compton wavelength.

There has been much recent progress in understanding the linear regime of the superradiant instability of massive vector fields [15, 19, 20]. There have, however, only been a handful of investigations of the superradiance of rotating BHs beyond the test field limit, including the superradiant scattering of a large amplitude GW [27], and the nonlinear growth of the superradiant instability in anti-de Sitter [28]. In addition, there have been several studies of the analogous process for charged BHs in spherical symmetry [29–31].

In Ref. [12], the growth of the superradiant instability of a massive vector field was studied, and found to be efficient even into the nonlinear regime, with the instability smoothly saturating as the BH horizon frequency decreases to match that of the oscillating cloud (see also Ref. [32] for a nonlinear study around a nonspinning BH). That study focused on a complex vector field with an assumed \( m = 1 \) azimuthal symmetry that gave rise to an axisymmetric spacetime. This assumption reduces the computational expense of a problem that is challenging due to the large separation of timescales between the oscillation period and growth time of the superradiant instability. However, in such a setting the GW radiation and other nonaxisymmetric couplings that would arise through gravity are suppressed. There also exist hairy BH solutions [33, 34] as the plausible end point of the instability fixing this symmetry [35, 36], which is not the case for a real vector field.

Here we tackle for the first time the more challenging—and observationally interesting—problem of a real massive vector field with no assumed spatial symmetries, leveraging accurate initial data for an exponentially growing boson cloud. We show evidence of nonlinear mode coupling through gravity, both in the GWs and in the field. However, we find that even in this regime where the growth rate and gravitational radiation reaction are maximized, and where the saturation of the
instability leads to the formation of a cloud rapidly oscillating close to the BH horizon with \( \approx 6\% \) of its mass, nonlinear effects do not significantly hamper the extraction, and conversion into GWs, of the BH’s rotational energy.

**Methodology.**—To study the superradiant instability, we evolve the Einstein field equations coupled to a (real) massive vector field \( X^a \)—also known as a Proca field—with mass parameter \( \mu \) (boson mass divided by \( \hbar \)), governed by \( \nabla_a F^{ab} = \mu^2 X^b \), with \( F_{ab} = \nabla_a X_b - \nabla_b X_a \). We evolve these equations using the same methods as described in Ref. [12], except that our domain is fully three dimensional and does not assume any spatial symmetries. We use units with \( G = c = 1 \) throughout.

In order to make it numerically tractable to follow the growth of the superradiant instability through saturation, we need to start with consistent initial data describing a non-negligible Proca cloud arising from the superradiant instability. To do that, we leverage the results of Ref. [12] where a complex massive vector field was evolved on an axisymmetric spacetime. We focus on a case from there with \( \mu = 0.4/M_0 \) around a BH with initial mass \( M_0 \) and dimensionless spin \( a = 0.9 \) and use the metric and Proca field configurations after the Proca cloud has grown from \( \sim 10^{-3} M_0 \) to \( 10^{-2} M_0 \), which is \( \sim 1/6 \) its saturation value. We then take just the real part of the Proca field rescaled to give approximately the same total mass, i.e. \( X_a \rightarrow \sqrt{2} \text{Re}(X_a) \), and use this for constructing initial data. There will be a small correction to the metric since we have a transformed an axisymmetric stress energy into a nonaxisymmetric one. We therefore obtain consistent initial data by solving the Einstein constraint equations using the methods of Ref. [27], and using the above configurations as the free data. This slightly decreases the BH mass and angular momentum, by \( \sim 0.3\% \), compared to the axisymmetric spacetime. We give more details on this procedure in the appendix.

In order to evolve these systems for timescales of \( > 10^4 M_0 \) at modest resolution, without accumulating large numerical errors just from the isolated BH spacetime, we use the background error subtraction technique [37]. Unless otherwise noted, the results presented were obtained using a numerical grid that has seven levels of mesh refinement, with a 2:1 refinement ratio, centered on the BH. The finest level has a grid spacing of \( dz \approx 0.02 M_0 \). To estimate truncation error, we also perform the same calculation using lower resolution. Details are given in the appendix.

To characterize our solutions, we extract several relevant quantities. Though our solutions are neither stationary nor axisymmetric, it will be useful to compute several mildly gauge dependent quantities with respect to the unit normal to slices of constant coordinate time \( n^a \), and the coordinate axial vector \( \hat{\phi}^a \) which would be the axisymmetric Killing vector in the unperturbed BH spacetime. From the stress-energy tensor of the massive vector field \( T^{ab} \), we can calculate the total energy \( E \) and angular momentum \( J \) from the volume integrals of \( n_a T^a_t \) and \( -n_a \hat{\phi}_b T^{ab} \), respectively. We also use the time component of the vector field \( \chi = -n_a X^a \) as a convenient scalar quantity with respect to our spacetime splitting. During the evolution, we keep track of the BH apparent horizon, and measure its area and its (approximate) angular momentum with respect to \( \hat{\phi}^a \). We combine these two quantities using the Christodoulou formula to construct a measure of the total BH mass \( M_{BH} \). To measure the GW emission, we extract the Newman-Penrose scalar \( \Psi_4 \).

**Results.**—We follow the boson cloud as it grows from \( E/M_0 \approx 0.01 \) to a maximum value of 0.06 over a timescale of \( \sim 9 \times 10^3 M_0 \), and then subsequently decreases due to the emission of gravitational radiation. This evolution of the cloud energy and angular momentum closely matches the decrease in the BH mass and angular momentum, as shown in Fig. 1. It is this spin-down of the BH that causes the instability to shut off—by \( t = 8 \times 10^3 M_0 \), the BH horizon frequency is within \( \sim 2\% \) of the cloud oscillation frequency, and the condition for superradiance is nearly saturated. For comparison, we also show the evolution of the Proca field quantities from the corresponding axisymmetric case. There is a small difference in the saturation values of these quantities, primarily due to the shift in BH parameters when constructing the initial data as described above. Also, unlike the axisymmetric case, at late times \( E \) and \( J \) begin to decrease due to GW emission, as described below.

The oscillations of the boson cloud produce GWs, as shown in Fig. 2. The GW signal is predominately quadrupolar, i.e., \( \ell = |m| = 2 \) spin weight \( -2 \) spherical harmonics, and has approximately twice the frequency of the dominant unstable mode, \( \omega_0 \approx 0.359/M_0 \) [the \((j, m, n) = (0, 1, 0)\) test field mode for \( \mu = 0.4/M_0 \) and \( a = 0.99 \)] [19]. In the language of
FIG. 1. The energy (top) and angular momentum (bottom) in the Proca field as a function of time (solid blue lines). In addition, we show the corresponding decrease in the mass and angular momentum measured from the BH horizon, with initial offset as given by the initial Proca field values (dotted black lines). For comparison, we also show the energy and angular momentum in the Proca field from the axisymmetric evolutions described in Ref. [12]. The difference in the saturation values between the two cases is predominantly due to the small change in the BH parameters caused by solving the constraint equations when constructing the nonaxisymmetric initial data.

Particle physics, this is the “annihilation” GW signal. There is also a small modulation in the amplitude that occurs with lower frequency which can be attributed to a beating of the frequency of the dominant unstable mode \( n = 0 \) with a smaller amplitude overtones \( n = 1 \) with slightly higher frequency, \( \Delta \omega_{10} = \omega_1 - \omega_0 = 0.029/M_0 \) \(^2\), also referred to as the “transition” GW signal. The amplitude of this secondary mode is somewhat an artifact of our initial data, and the fact that we only track the cloud through a small number of e-folds. The growth rate of this mode is approximately \( 1/4 \times \) smaller than the dominant mode, and so would be exponentially suppressed if the cloud grows for many e-folds (and will also begin to decay at late times when \( \omega_1 \) is above the BH horizon frequency). However, such a lower frequency transition GW signal might be relevant, in particular, if there is part of the parameter space where an overtone mode has a faster growth rate than the fundamental mode \([15]\), as can occur for the scalar case \([38]\).

Figure 2 also shows that the GW signal has higher \( \ell \) spherical harmonic components, which is expected since the cloud is localized close to the nonspherical, spinning BH. In the bottom panel we can see that there is also an \( m = 4 \) component to the GWs at twice the frequency of the \( m = 2 \) component. The \( m = 4 \) component has a smaller amplitude that grows quadratically with the cloud energy, consistent with being due to mode doubling effects.

The GW radiation is subdominant during the exponential growth phase of the cloud. This is illustrated in the top panel of Fig. 3, where we show the rate of increase of the boson cloud energy and the GW luminosity \( P_{GW} \) as a function of time. The value of \( P_{GW}/E^2 \) found here when the cloud size is at maximum mass matches the value found in the test field regime [19] to within the estimated numerical error (which is \( \sim 17\% \)). This peak value of the GW luminosity is \( \sim 20 \times \) smaller than the maximum growth rate of the Proca cloud. However, eventually as the BH is spun down and the superradiant instability saturates, the boson cloud begins to lose energy with a rate that roughly matches the GW emission, as shown in the bottom panel. (At earlier times in the bottom panel of Fig. 3, the energy and angular momentum does briefly plateau, perhaps due to the presence of the overtone mode.) The cloud will dissipate on a characteristic timescale of \( t_{GW} := E/P_{GW} \sim 10^5 M_0 \).

If we look at the vector field itself—in particular, \( \chi \) as shown in Fig. 4—in addition to the dominant

\(^2\)There are also other polarization modes in addition to the dominant \( S = -1 \). However, they grow at a much slower rate than the first overtone \([20]\).
FIG. 2. Spherical harmonics of the real part of $\Psi_4$, which encodes the GW signal. The top panel shows the $2 \leq \ell \leq 4$, $m = 2$ components, while the the bottom panel shows the $4 \leq \ell \leq 6$, $m = 4$ components. The left-hand inset demonstrates that the $m = 2$ ($m = 4$) signal is predominantly at $2$ ($4$) times the frequency of the most unstable mode $f_0 = \omega_0/(2\pi)$, while the right-hand inset demonstrates that there is a smaller variation in the amplitude due to beating of the first and second most unstable modes at frequency $\Delta f_{10} := f_1 - f_0$.

FIG. 3. Top: The time derivative of the (smoothed) Proca energy (black line) and the GW power (red line). Bottom: The decrease in the Proca cloud energy (black lines) and angular momentum (blue lines) from its peak value. The dashed lines show the results from the lower resolution run, which is continued for longer times. For comparison, we also show the rate of decrease expected from GW dissipation (dotted red and magenta lines, respectively) for both resolutions.

$m = 1$, there are higher $m$ components (predominantly $\ell = m$) at a much smaller amplitude (though as noted above, since there is no axisymmetry, such a decomposition is not free of gauge dependence). The increase of these appear consistent with nonlinear mode coupling through gravity, with the $m = 3$ and $m = 2$ mode increasing in proportion to the third and fourth power of the $m = 1$ mode, respectively. We note that though there are superradiantly unstable modes with higher $m$, which can continue to spin down the BH after the saturation of the $m = 1$ mode and subsequent dissipation of the cloud due to GWs, their growth time is much longer than considered here ($> 10^7 M_0$). However, nonlinear mode coupling could play a role in seeding such superradiant modes at a higher amplitude.

Conclusion.—We have studied the development of the superradiant instability of a spinning BH in the presence of a massive vector field, following the spin-down of the BH and the growth of an oscillating boson cloud with significant mass and attendant GW radia-
As a consequence of the fact that, at saturation, the BH horizon has a frequency that matches the oscillated spinning BH’s mass, and convert it into GWs. This amount of GW energy and frequency is comparable to that of a BH merger—though the peak luminosity is \( \sim 10^{-3} \times \) smaller—and such a signal would be visible by LIGO at distances on the order of a Gpc. This exact value of the vector boson mass is in tension with reported BH spins measurements, though observations of higher BH masses can probe new parts of the parameter space [15]. We have focused on a maximally relativistic case, while for larger vector bosonic Compton wavelengths (compared to the BH radius), as well as for the case of a scalar boson, the growth rates and GW luminosity are smaller, and we expect the adiabatic approximation to be even more accurate. These sources will thus be promising targets for continuous GW searches [10][11].

Besides the spin-down of the BH leading to saturation of the instability, and the dissipation of the boson cloud through GWs, we see subdominant nonlinear effects, including evidence of mode coupling both in the vector field and in the GWs, similar to that reported in Refs. [27][28]. It would interesting for future work to investigate what role these could play in speeding up the onset of higher azimuthal number superradiant instabilities, which could liberate further rotational energy from the BH at late times. The large energy densities and high frequency oscillations of the boson clouds exhibited here also suggest studying possible observational signatures of coupling to an accretion disk or other matter.

**Acknowledgements.**—I thank Frans Pretorius, Helvi Witek, Huan Yang, and Miguel Zilhão for stimulating discussions and comments on this work. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science. Computational resources were provided by the Perseus cluster at Princeton University and by XSEDE under Grant No. TG-PHY100053.

---

1. B. P. Abbott *et al.* (Virgo, LIGO Scientific), Phys. Rev. Lett. 116, 061102 (2016), arXiv:1602.03837 [gr-qc].
2. S. Weinberg, Phys. Rev. Lett. 40, 223 (1978).
3. A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, Phys. Rev. D81, 123530 (2010), arXiv:0905.4720 [hep-th].
4. A. Arvanitaki and S. Dubovsky, Phys. Rev. D83, 044026 (2011), arXiv:1004.3558 [hep-th].
5. B. Holdom, Phys. Lett. B166, 196 (1986).
6. M. Cicolli, M. Goodsell, J. Jaeckel, and A. Ringwald, JHEP 07, 114 (2011) [arXiv:1103.3705 [hep-th]].
7. T. Damour, N. Deruelle, and R. Ruffini, Lettere Al Nuovo Cimento Series 2 15, 257 (1976).
8. S. L. Detweiler, Phys. Rev. D22, 2323 (1980).
9. T. Zouros and D. Eardley, Annals Phys. 118, 139 (1979).
10. S. R. Dolan, Phys. Rev. D87, 124026 (2013), arXiv:1212.1477 [gr-qc].
11. R. Brito, V. Cardoso, and P. Pani, Classical and Quantum Gravity 32, 134001 (2015) arXiv:1411.0686 [gr-qc].
12. W. E. East and F. Pretorius, Phys. Rev. Lett. 119, 041101 (2017), arXiv:1704.04791 [gr-qc].
Constructing initial data

We construct initial data that satisfy the constraint part of the Einstein equations by solving the conformal thin-sandwich equations using the methods described in \[27\]. In this formulation, one freely chooses the conformal three-metric $\tilde{\gamma}_{ij}$, the time derivative $\dot{\tilde{\gamma}}_{ij}$, the trace of the extrinsic curvature $K$, and the conformal lapse $\tilde{\alpha}$. Once the Hamiltonian and momentum constraint are solved for the conformal factor $\tilde{\alpha}$ and the shift vector $\tilde{\beta}^i$, we can construct the full metric by conformally rescaling these quantities as $\gamma_{ij} = \tilde{\gamma}^{\alpha} \tilde{\alpha}^6 \delta_{ij}$, $\psi = \tilde{\psi} \tilde{\alpha}^2$, etc.

As indicated in the main text, we use the solutions of \[12\] describing a Proca cloud around a spinning BH that has grown through superradiance to a non-negligible mass, but is still far from saturation, as a starting point in constructing initial data. These solutions have an axisymmetric spacetime, but the complex Proca field has an azimuthal symmetry, so one cannot use them directly. According to the Proca field from the above mentioned complex solution, we set the conformal metric functions to be equal to the values from the axisymmetric solution. In addition, we must also specify the energy density $U$ and three momentum density $p^i$. With the Proca field equations, there is an additional complication since the Proca field equations also have constraints. In order to guarantee that these will be satisfied, we make the following choices for conformally rescaling the Proca fields. We let $\tilde{E}^i = \psi^6 E^i$, $\tilde{\chi} = \psi^5 \chi$, and $\tilde{\chi}_i = \dot{\chi}_i$. It follows then that $B^i = \psi^6 \tilde{B}^i$. We set these quantities according to the Proca field from the above mentioned complex solution, but using only the real part of the field, and rescaling to get the same amount of energy: $X_a \rightarrow \sqrt{2} \tilde{R}(X_a)$. Because of this, the Hamiltonian and momentum constraints must be solved.

Since the conformal Proca fields satisfy the Proca field constraint equations with respect to the conformal metric $\tilde{D}_i \tilde{E}^i = -\mu^2 \tilde{\chi}$, it follows then that the
rescaled Proca fields will satisfy these constraints with respect to the rescaled metric:

\[ D_i E^i = \frac{1}{\sqrt{\gamma}} \partial_i (\sqrt{\gamma} E^i) = \psi^{-6} \tilde{D}_i \tilde{E}^i = -\mu^2 \chi . \] (1)

The three momentum density is then given by

\[ p_i = \epsilon_{ijk} E^j B^k + \mu^2 \chi \chi = \psi^{-6} \tilde{p}_i \] (2)

where \( \tilde{p}_i \) is the equivalent expression in terms of the conformal metric and Proca fields. The energy density is given by

\[ U = \frac{1}{2} \psi^{-8} \left( \tilde{B}^i \tilde{B}_i + \tilde{E}^i \tilde{E}_i \right) + \frac{1}{2} \mu^2 \left( \psi^{-12} \tilde{\chi}^2 + \psi^{-4} \tilde{\chi}^i \tilde{\chi}_i \right) . \] (3)

For the massless case \( \mu = 0 \), i.e. electromagnetism, then \( U = \psi^{-8} \tilde{U} \), and hence the conformal momentum and energy have the usual dependence taken for the conformal thin-sandwich equations. For the massive case, there is an additional contribution to the energy which does not scale with this power, but for the cases considered here we are still able to relax to a solution.

We also note in passing that a similar procedure can be used for constructing superposed initial data, for example describing a charged binary BH, without solving the vector field constraints. If we begin with two solutions that individually satisfy the vector field constraints, then we can merely superpose the fields weighted by their metric determinants, e.g.

\[ \sqrt{\gamma} \tilde{E}^i = \sqrt{\gamma^{(1)}} E^{(1)i} + \sqrt{\gamma^{(2)}} E^{(2)i} \] (4)

to get a conformal field that when rescaled will still satisfy the constraints.

**Numerical convergence**

We demonstrate that the constraints are indeed converging to zero during the evolution in Fig. 5, where we show a norm of the generalized harmonic constraint violation \( C_a = H_a - \Box x_a \) [12], for the standard resolution and low resolution, the latter of which has \( 3/4 \times \) the grid spacing of the former. The decrease in this quantity with increased resolution is consistent with the expected fourth-order convergence.

Comparing the two different resolutions, and assuming fourth-order convergence, we can estimate the truncation error in various quantities. We estimate the error in the saturation energy for the Proca cloud to be \( \sim 3\% \). For the GWs we estimate the error in the amplitude of the dominant \( \ell = m = 2 \) mode to be \( \sim 10\% \) while the error in the smaller amplitude and higher frequency \( \ell = m = 4 \) mode is \( \sim 30\% \). For all these quantities, the trend with resolution suggests that these quantities are underestimated due to truncation error.

FIG. 5. The norm of the generalized harmonic constraint violation \( C_a = H_a - \Box x_a \) (average value in the \( x = 0, (y,z) \in [-50M_0, 50M_0]^2 \) plane) at two resolutions. The low resolution case is also shown rescaled by the factor expected for fourth-order convergence.