Solving the Teukolsky equation with physics-informed neural networks

Raimon Luna[,1] Juan Calderón Bustillo[,2,3] Juan José Seoane Martínez[,1]
Alejandro Torres-Forné[,1,4] and José A. Font[1,4]

[1]*Departamento de Astronomía y Astrofísica, Universitat de València,
Dr. Moliner 50, 46100, Burjassot (València), Spain*
[2]*Instituto Galego de Física de Altas Enerxías, Universidade de Santiago de Compostela,
15782 Santiago de Compostela, Galicia, Spain*
[3]*Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong*
[4]*Observatori Astronòmic, Universitat de València,
C/ Catedrático José Beltrán 2, 46980, Paterna (València), Spain*

We use physics-informed neural networks (PINNs) to compute the first quasinormal modes of the Kerr geometry via the Teukolsky equation. This technique allows us to extract the complex frequencies and separation constants of the equation without the need for sophisticated numerical techniques, and with an almost immediate implementation under the PyTorch framework. We are able to compute the oscillation frequencies and damping times for arbitrary black hole spins and masses, with accuracy typically below the percentual level as compared to the accepted values in the literature. We find that PINN-computed quasinormal modes are indistinguishable from those obtained through existing methods at signal-to-noise ratios (SNRs) larger than 100, making the former reliable for gravitational-wave data analysis in the mid term, before the arrival of third-generation detectors like LISA or the Einstein Telescope, where SNRs of $\mathcal{O}(1000)$ might be achieved.

I. INTRODUCTION

The observation of gravitational waves from binary black hole (BBH) mergers by the LIGO-Virgo-KAGRA (LVK) detector network [1–4] has opened a window to study gravity in its strong field regime. In particular, some of the current detections have made possible to access the ringdown signal of the highly distorted black holes produced after the merger, opening the prospects to turn black-hole spectroscopy into an observational field (see e.g. [5–12]). This discipline aims to identify and characterize the quasinormal modes (QNMs) of a rotating (Kerr) black hole, as those are the fingerprint of relaxing black holes after a perturbation of their spacetime geometry. In black-hole perturbation theory any type of perturbation (scalar, electromagnetic and gravitational) can be mapped into wave equations with a short range potential. QNMs are solutions of those equations that correspond to damped oscillating modes, i.e. modes which are outgoing at infinity and ingoing at the black hole horizon. The properties of these modes are well understood (see [13–15] for details).

A generic excitation of an isolated black hole leads to a prompt emission of gravitational waves, followed by a distinctive exponentially decaying QNM ringing. Such a perturbation can naturally arise in a variety of astrophysical situations. Examples include the late stages of the gravitational collapse of massive stars or the capture of infalling matter. With the advent of gravitational-wave astronomy and the detection of signals from BBH mergers, the QNM ringdown of the final black hole has become the prime example to test black hole perturbation theory. The ringdown can be searched for in the postmerger section of the gravitational-wave signal, which encodes information on the properties of the remnant object. In particular, the analysis of QNMs can be exploited to test the nature of such object and, specifically, to understand if this is consistent with an astrophysical black hole as predicted by general relativity, ruling out alternative exotic compact objects [16].

One of the most appealing prospects of black-hole spectroscopy is, thus, the testing of general relativity. By comparing the observed QNM spectra of post-merger black holes with the known spectra of the Kerr (or Kerr-Newman) solutions of Einstein(-Maxwell) theories, tests of the no-hair theorem [17–20] may become possible. This conjecture states that for a Kerr-Newman black hole the mode oscillation frequencies and damping rates are determined uniquely by the black hole mass, angular momentum and possibly charge (which is expected to be zero for astrophysical black holes). Consistency tests of general relativity need, at least, the precise identification of three QNM parameters, two to measure the mass and the spin of the black hole and the third one for cross-check [21,22]. Such an identification is expected to be only possible if the

064025-1

signal-to-noise ratio (SNR) in the ringdown stage is $\mathcal{O}(100)$ [5,23,24]. Therefore, the actual realization of strong-confidence tests of GR through black-hole spectroscopy seems out of reach for current second-generation detectors[1] (but see [6] for a more promising estimate when various signal posteriors are *combined*) while it will be possible with the Einstein Telescope [26] and the space detector LISA [27]. At present, searches for violations of the no-hair conjecture are conducted by introducing potential deviations in the QNM parameters as additional degrees of freedom to estimate through Bayesian inference. The most recent study carried out by the LVK collaboration using data form the third gravitational-wave catalog GWTC-3 [4] constrains the frequency deviation parameters for the fundamental and the first overtone of the $\ell = m = 2$ QNM by about 1% at 90% credibility [28] (see [10,25] for the first application of this technique). Black hole spectroscopy is expected to become increasingly important as gravitational-wave detectors become more sensitive to subdominant quasinormal modes. Moreover, the possible deviations from the predictions of general relativity can be parametrized and used to test for exotic compact objects [16] and alternative theories of gravity [29,30].

The QNM frequencies and their decay rates are computed from the linear analysis around the background geometry, and can be extracted from a large variety of spacetime geometries (see [15,31–37] for an incomplete list). In the case of the Schwarzschild geometry, the perturbations were first studied by [20,38,39] and then extended to Kerr black holes by Teukolsky [19], leading to the renowned master equation bearing his name. After separation of variables on the Teukolsky equation, two master equations depending respectively on the radial and zenithal coordinates are obtained, which have to be solved in order to extract the QNMs. There is a surprisingly large diversity of methods to solve such master equations. Those have been devised during the past decades in parallel to the development of computational power. The first method, introduced by Leaver [40] and based on the Fröbenius method, expands the master variables as an infinite series, imposing the convergence of such series in the whole domain through the so-called "quotient method." This is equivalent to finding roots of a function which is expressed as a continued fraction, which is the reason why this approach is commonly known as the "continued fraction" method. Among the many other methods, it is worth mentioning the direct time evolution of the equations [41–43], the computation of bound states in an effective potential [44], the WKB approximation [45], asymptotic iteration methods [46], pseudospectral

basis decomposition [47,48], or even more exotic approaches such as the expansion in the spacetime dimension [49].

In this paper we solve the Teukolsky equation using a machine-learning approach, namely physics-informed neural networks (PINNs) [50] and assess the quality of our results by comparing them with those obtained by the continued fraction method. The main advantage of PINNs as a method for finding QNMs is that it can be easily adapted to different systems with very little changes in the method implementation, bringing us closer to a general-purpose quasinormal mode solver. The networks can in principle solve an arbitrary number of equations with an arbitrary number of eigenvalues and separation constants just by adding them to the loss function of the system with appropriate weighting factors. Moreover, PINNs do not require a particular computational grid or numerical differentiation techniques and the method can be easily implemented with only a few lines of code in a machine-learning environment (as e.g. the PyTorch [51] framework employed here). The equations that are to be solved have no particular requirements for the PINN to process them, as long as their solutions are finite and regular, which makes PINNs also adequate for uses on numerical background solutions.

The main idea behind PINNs is to use neural networks as universal approximants of mathematical functions [52] in order to solve differential equations. In the simplest case, one can approximate a function $f(x)$ with a fully connected feed-forward neural network with a single neuron in the input layer, which receives values of x. Then, a single neuron (or two, if f is complex) in the output layer will generate the value of $f(x)$. An important advantage with respect to other numerical techniques is the ability to evaluate the derivatives of $f(x)$ by standard automatic differentiation on every point, without the need for a numerical differentiation scheme on a fixed grid.

Automatic differentiation in general, and backpropagation in particular, are the commonly used techniques to take derivatives of machine learning models with respect to their parameters (weights and biases), which are then used by an optimization algorithm to train the model. In our case, we use the same infrastructure of automatic differentiation to compute the derivatives of the model outputs with respect to the input coordinates of the differential equation (in the present case, the radial and angular coordinates). Given a differential equation of the form $F[x; f(x), f'(x), f''(x), \dots] = 0$, then it is straightforward to define a loss function $\mathcal{L}$ as the norm of the values of F on some set of points in the domain. Minimizing $\mathcal{L}$ is then equivalent to solving the differential equation for $f(x)$.

When the differential equation has one or multiple parameters to be computed, these can be fitted in the same manner as the weights and biases of the neural network, making the practical implementation remarkably

---

[1]As an example, the ringdown SNR for the first signal detected, GW150914, was around 7 when the analysis is started 3 ms after the signal peak. Moreover, this reaches $\simeq 15$ if the analysis is started at the signal peak introducing overtones [10,11,25].

straightforward. General boundary conditions and solution normalizations can be included in the loss function (weak enforcement) or imposed on the neural network output (hard enforcement), see for instance [53–55]. PINNs have been used in the past for the computation of black-hole QNMs [56–58] in the case of the Schwarzschild metric. In this paper we extend this approach to the more astrophysically relevant case of the Kerr metric.

The rest of the paper is structured as follows. In Sec. II we lay out our analytic treatment of the Teukolsky equation in order to make it suitable for its implementation in a PINN, which is described in Sec. III. We present our results in Sec. IV and close with our conclusions in Sec. V.

II. THE TEUKOLSKY EQUATION

In the Teukolsky equation [19] and its solution [40], linear perturbations around the Kerr metric are encoded in a master variable ψ. These are classified by their spin-weight under the Lorentz group as scalar ($s = 0$), vector ($s = -1$), and tensor ($s = -2$) perturbations. In this paper we will always take $s = -2$. After performing a standard separation of variables, we obtain

$$\psi = \frac{1}{2\pi} \int e^{-i\omega t} \sum_{l=|s|}^{\infty} \sum_{m=-l}^{l} e^{im\phi} S_{lm}(\theta) R_{lm}(r) d\omega, \quad (1)$$

where θ and ϕ are the zenithal and azimuthal angular coordinates, and ω is the (complex) frequency of the mode. We are left with two ordinary differential equations (ODEs) for the radial and angular parts of the solution. When solving for the modes, we will take $G = c = 2M = 1$, with c denoting the speed of light in vacuum, G Newton's constant, and M the black hole mass. We reinstate these constants when generating waveforms for particular values of the mass M and spin a. The radial part, for a solution with intrinsic angular momentum $J = 2Ma$, has the form

$$\Delta(r)R''(r) + (s+1)(2r-1)R'(r) + V(r)R(r) = 0. \quad (2)$$

Above, we have defined

$$\Delta(r) = r^2 - r + a^2,$$
$$V(r) = \alpha(r)\omega^2 + \beta(r)\omega + \gamma(r),$$
$$\alpha(r) = \frac{(r^2 + a^2)^2}{\Delta(r)} - a^2,$$
$$\beta(r) = \frac{is(a^2 - r^2) - 2amr}{\Delta(r)} + 2isr,$$
$$\gamma(r) = \frac{a^2m^2 + isam(2r-1)}{\Delta(r)} - A, \quad (3)$$

where A is a separation constant that will depend on the angular momentum numbers l, m. Furthermore, in the Schwarzschild limit we know that [40]

$$\lim_{a \to 0} A = l(l+1) - s(s+1). \quad (4)$$

Correspondingly, the angular equation reads

$$(1 - u^2)S''(u) - 2uS'(u)$$
$$+ \left[a^2\omega^2u^2 - 2a\omega su + s + A - \frac{(m + su)^2}{1 - u^2} \right] S(u) = 0, \quad (5)$$

where $u = \cos\theta$. In addition to these ODEs, any perturbation must obey ingoing boundary conditions at the horizon and outgoing at spatial infinity, as well as regularity at the spherical poles $\theta = 0, \pi$. A set of ansätze obeying such conditions, derived by Leaver [40], are

$$R(r) = e^{i\omega r}(r - r_-)^{p_-}(r - r_+)^{p_+}f(x),$$
$$p_- = -1 - s + i\omega + i\sigma_+,$$
$$p_+ = -s - i\sigma_+,$$
$$\sigma_+ = \frac{\omega r_+ - am}{\sqrt{1 - 4a^2}}, \quad (6)$$

with $x = r_+/r$ being used to compactify the radial domain, and

$$S(u) = e^{a\omega u}(1 + u)^{|m-s|/2}(1 - u)^{|m+s|/2}g(u), \quad (7)$$

where $r_\pm = (1 \pm \sqrt{1 - 4a^2})/2$ are respectively the inner (r_-) and outer (r_+) horizon radii of the Kerr metric. The functions $f(x)$ and $g(u)$ have to be regular in the ranges $[0, 1]$ and $[-1, 1]$ respectively. By substitution of (6) and (7) into (2) and (5), we obtain a set of equations for the functions $f(x)$ and $g(u)$ of the form

$$\mathcal{L}_F[f(x)] = F_2f''(x) + F_1f'(x) + F_0f(x) = 0,$$
$$\mathcal{L}_G[g(u)] = G_2g''(u) + G_1g'(u) + G_0g(u) = 0, \quad (8)$$

where the explicit form of the $F_i(x)$ and $G_i(u)$ coefficients is given in the Appendix. With this, the problem reduces to finding the set of $\{\omega_n, A_n\}$ that admit nonzero solutions for (8). Following Leaver, we choose the normalization conditions

$$f(1) = 1, \qquad g(-1) = 1. \quad (9)$$

III. THE NEURAL NETWORK

In order for the PINN to solve the equations, we must define two sets of values for the variables x and u where the networks will be evaluated. For the sake of simplicity, we pick uniformly spaced grids, as

$$x_i = \frac{i-1}{N_x - 1}, \quad i = 1, \ldots, N_x,$$

$$u_j = \frac{2(j-1)}{N_u - 1} - 1, \quad j = 1, \ldots, N_u, \qquad (10)$$

with $N_x = N_u = 100$, although it is important to keep in mind that a uniform grid is not a requirement. We can, therefore, naturally define a loss function as the weighted sum of the averaged moduli of the right-hand side of Eq. (8), as

$$\mathcal{L}[\{x_i, u_j\}] = \frac{W}{N_x} \sum_{i=0}^{N_x} |\mathcal{L}_F[f(x_i)]| + \frac{1}{N_u} \sum_{j=0}^{N_u} |\mathcal{L}_G[g(u_j)]|,$$

$$(11)$$

where the constant W can be used to give different relative weights to the radial and angular equations. The specific value of W is not very important as long as it is larger than 1, which gives a higher importance to the (more complicated) radial part. For the case at hand we chose $W = 10$. Under this setup, we will define two fully-connected neural networks, $\mathcal{N}_f$ and $\mathcal{N}_g$, to act as approximants to the solutions $f(x)$ and $g(u)$. A schematic representation of the structure of the algorithm is shown in Fig. 1. The input layers of both networks will contain a single neuron, which receives the value of the dummy variable (x and u respectively). The output layers consist of two neurons that output the real and imaginary parts of the approximated functions. The computations across both networks are performed on real numbers exclusively. The trainable parameters of the setup will then be the weights and biases of the networks and, very importantly, the parameters ω and A. Furthermore, we adopt a "hard enforcement" scheme to impose the normalization conditions (9), by defining

$$f(x) = (e^{x-1} - 1)\mathcal{N}_f(x) + 1,$$
$$g(u) = (e^{u+1} - 1)\mathcal{N}_g(u) + 1. \qquad (12)$$

This automatically ensures that (9) are satisfied without the need for any specially devoted term in the loss (11), which significantly accelerates the convergence of the method.

The system is implemented under the PyTorch framework, which provides automatic differentiation capabilities, making it very easy to implement the neural networks as well as the optimizing tools. Optimization is performed by the Adam algorithm [59], with an initial learning rate of 0.005 which exponentially decays with a factor 0.999 per epoch. Each value of the spin parameter a is trained for a total of 2000 epochs. The gradients of the loss with respect to the parameters of both networks are computed by automatic differentiation, so both networks are trained together as a single model.
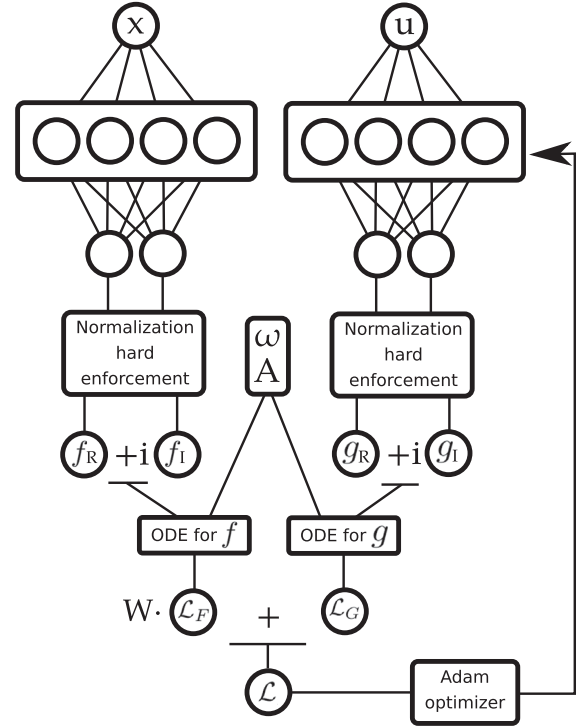

FIG. 1. Schematic flowchart of the structure of the neural networks. The functions $f(x)$ and $g(u)$ are evaluated independently and then introduced into the ODEs (which share the eigenvalues ω and A) to produce the two contributions to the loss. A weighted sum of both contributions is defined as the total loss function of the system.

In order to compute the QNMs for a whole range of spins in $a \in [0, 1/2)$ of the Kerr geometry, which can be then interpolated, we can proceed in a sequential manner. We start by training the Schwarzschild case ($a = 0$) with the weights and biases of the network randomly initialized. Initial random weights are normally distributed with zero mean and standard deviations of 0.05 (f network) and 0.01 (g network), while all biases are set to zero. The initial value of ω can be set to a value close to the expected mode (in our case we choose $\omega_0 = 0.7 - 0.1i$), and A is used to fix l, by initializing it to its analytic value at zero spin: $A_0 = l(l+1) - s(s+1)$.

After the solution for $a = 0$ has been trained, we start progressively increasing the value of a in small steps, and retrain both networks to extract the QNM at the new value of a, until we reach the maximum value of $a = 0.4999$. At each training step, we take as initial weights and biases the ones resulting from the training at the previous step. In this way, we start the convergence process for each value of a with a configuration that is closer to the optimum than a randomly initialized network. We also guarantee that the network will always converge to the mode with the same value of l, which would otherwise be hard to impose for $a > 0$.

We find that, when gradually increasing the value of the spin parameter a, it is convenient to cluster the training at

higher spins. In order to achieve this, we choose increasing values of a as

$$a_k = \frac{1}{2}\left(1 - 10^{-\frac{3(k-1)}{N_a-1}}\right), \quad k = 1, ..., N_a, \quad (13)$$

with $N_a = 20$.

With this method, the initialization of the network at each value of a is closer to the solution, which improves the accuracy. When the QNM frequencies have to be evaluated for a large number of different spin parameters a, it is convenient to perform an interpolation over the PINN results on a sparser set of values of a.

A. Hyperparameter study

We have performed a hyperparameter study on the possible neural network architectures we could use. It revealed that moderately simple neural networks are enough for this problem. Figures 2 and 3 show the relative deviation of the real and imaginary parts of the mode $(2, 0, 0)$ when compared to [40] as a function of the number of neurons per layer and the number of hidden layers, respectively. As an example we plot the results for the spinless case $(a = 0)$ and the near extremal case $(a = 0.4999)$. A measure of the accuracy on the actual differential equations for the radial and angular parts is the value of the loss function. For our computations, the final loss is typically of the order of 10^{-1}.

Even though the accuracy of the results does not vary dramatically once a sufficient network complexity is achieved, for all calculations in this paper we will use 2 hidden layers of 200 neurons, with the hyperbolic tangent
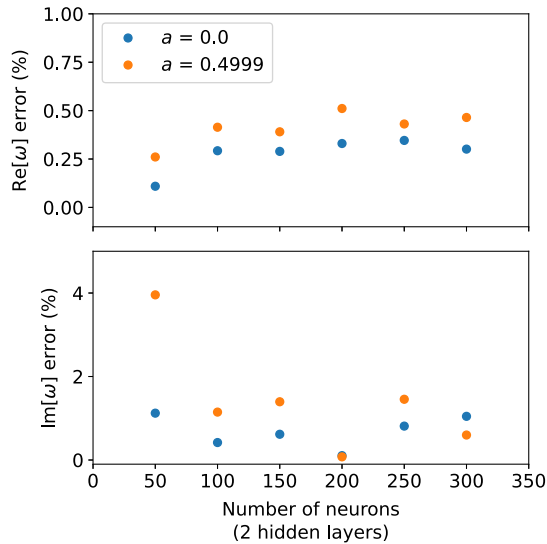

FIG. 2. Error of the $(l, m, n) = (2, 0, 0)$ QNM frequency for a zero spin black hole (blue circles) and a near extremal Kerr black hole (orange circles), when compared to [40], as a function of the number of neurons in each of the two hidden layers.
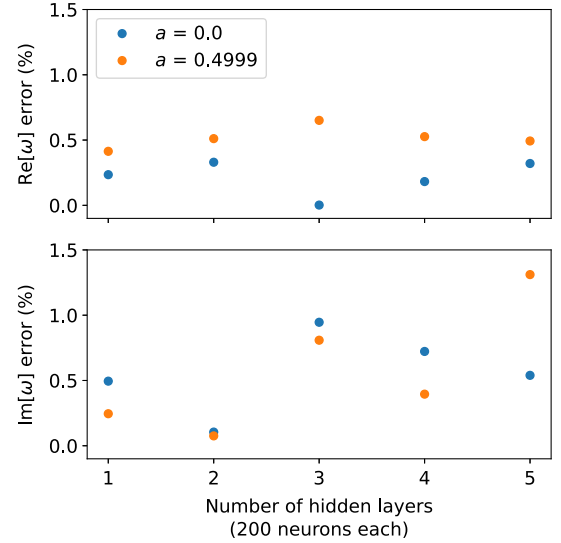

FIG. 3. As Fig. 2 but showing the error of the $(l, m, n) = (2, 0, 0)$ QNM frequency as a function of the number of hidden layers, with 200 neurons each.

as the activation function. This configuration is enough to correctly obtain the values of the modes with errors below 1%, while making the training of each mode reasonably fast: the full training of 2000 iterations takes about 2 minutes on a single Intel Core i7-10750H at 2.60 GHz CPU.

IV. RESULTS

A. QNM computation

We use the neural networks to compute the values of the QNMs reported in [40], particularly those with $(l, m, n) = (2, 2, 0)$ and $(2, 0, 0)$, with spin parameters ranging from $a = 0$ (Schwarzschild black hole) to $a = 0.4999$. While our method is applicable to any ringdown mode, we choose to showcase these two. The first one is the fundamental emission mode from circular black-hole merger remnants. The second is codominant in the case of highly eccentric mergers, a type of source that has received much attention recently (e.g. [60–64]), and therefore may serve as a smoking gun for such scenario. In Fig. 4 we show the final results for the near-extremality case, i.e., $a = 0.4999$ as the values of the learned radial and angular functions $f(x)$ and $g(u)$ and the learning curve for this value of the spin. The learning curve is a plot of the loss function from Eq. (11) as a function of the number of iterations. As expected for a converging system, it starts by dropping very fast, and then the convergence rate becomes smaller as we approach the minimum.

The real and imaginary parts of the QNMs are explicitly reported in Tables I and II, respectively, together with their relative deviation from the values in [40]. In this particular case, we randomly initialize the networks for
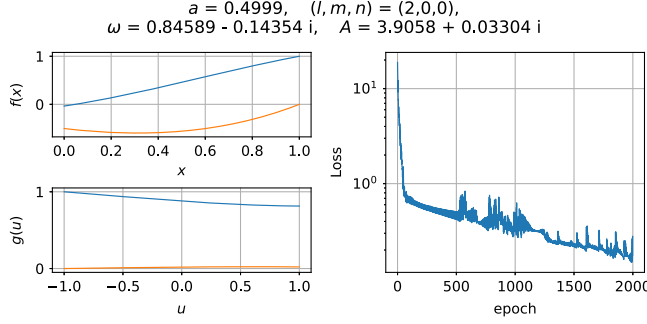
FIG. 4.　$f(x)$ and $g(u)$ obtained from the PINN for the dominant QNM in the case $(l, m, n) = (2, 0, 0)$, together with their learning curve, for $a = 0.4999$. The blue and orange curves correspond to the real and imaginary parts, respectively.

each value of a. It can be observed that all of the modes in the tables are less than 1% away.

Subdominant QNMs can also be computed with the PINN, although higher overtones ($n > 0$) become increasingly difficult to obtain and usually require larger training times. The simplest approach is to initialize the frequency parameter ω to a value closer to the desired mode prior to

TABLE I.　Numerical results obtained by the PINN for the oscillatory (real) part of the QNM ω for $(l, m, n) = (2, 0, 0)$, compared to Leaver's results. The error lies below the percentual level in all cases.

	Real part of ω		
a	PINN	Leaver	% error
0	0.74981	0.74734	0.330
0.1	0.75137	0.75025	0.149
0.2	0.76151	0.75936	0.282
0.3	0.77758	0.77611	0.190
0.4	0.80571	0.80384	0.233
0.45	0.82380	0.82401	0.025
0.49	0.83974	0.84451	0.564
0.4999	0.84589	0.85023	0.511

TABLE II.　Numerical results obtained by the PINN for the damping (imaginary) part of the QNM ω for $(l, m, n) = (2, 0, 0)$, compared to Leaver's results. The error lies below the percentual level in all cases.

	Imaginary part of ω		
a	PINN	Leaver	% error
0	−0.17812	−0.17793	0.108
0.1	−0.17719	−0.17740	0.122
0.2	−0.17610	−0.17565	0.255
0.3	−0.17332	−0.17199	0.775
0.4	−0.16572	−0.16431	0.853
0.45	−0.15785	−0.15697	0.566
0.49	−0.14657	−0.14707	0.337
0.4999	−0.14354	−0.14365	0.073

TABLE III.　Subdominant modes obtained by the PINN for the Schwarzschild black hole, together with their errors with respect to the results in [40].

	Subdominant mode examples		
(l, m, n)	PINN ω	δ Re (%)	δ Im (%)
(3, 3, 0)	$1.2016 - 0.1852i$	0.23	0.11
(2, 2, 1)	$0.6943 - 0.5490i$	0.13	0.21

the training, which can make the training process converge to it. Slowly lowering the initial imaginary part of ω can be used as a way to scan the first overtones. More sophisticated methods use driving terms in the loss function to favor higher modes, as in [54]. As an illustration, in Table III we show the subdominant modes (3, 3, 0) and (2, 2, 1) for the spinless case ($a = 0$), together with their errors.[2] These two QNMs are considered the best candidates for black hole spectroscopy [10,12,65,66], as the (3, 3, 0) is the loudest and longest-lived subdominant mode in BBH mergers and the (2, 2, 1) mode is the strongest overtone when only the spherical mode (2,2) is observable in the signal during the merger stage, as it is the case for all GW detections to date (see however [12]). For the sake of simplicity, for the rest this paper we will restrict to the fundamental $n = 0$ modes.

B. Detectability prospects

In order to assess the quality of the QNMs generated from the PINN for parameter estimation (mass, spin, amplitude and phase) of a ringing astrophysical black hole detection, we can produce ringdown waveforms as damped sinusoid strain signals of the form

$$h = h_0 e^{-t/\tau} \cos(2\pi f t + \phi), \tag{14}$$

for every set of (l, m, n). Above, h_0 and ϕ are an initial amplitude and phase determined by the particular perturbation, which in Fig. 5 we choose to be 10^{-21} and 0 respectively. The frequency f and damping rate τ for a black hole of mass M are computed respectively from the real and imaginary parts of ω as

$$f = \frac{c^3}{4\pi GM} \mathrm{Re}[\omega],$$

$$\tau = -\frac{2GM}{c^3} \frac{1}{\mathrm{Im}[\omega]}. \tag{15}$$

For each mass and spin configuration, we produce two waveforms using the ω values respectively computed by the

---

[2]Due to a slower convergence, the training of the (2, 2, 1) mode needed 5000 iterations instead of 2000, and required the initial guess for ω to be $0.7 - 0.5i$, quite close to the correct value.
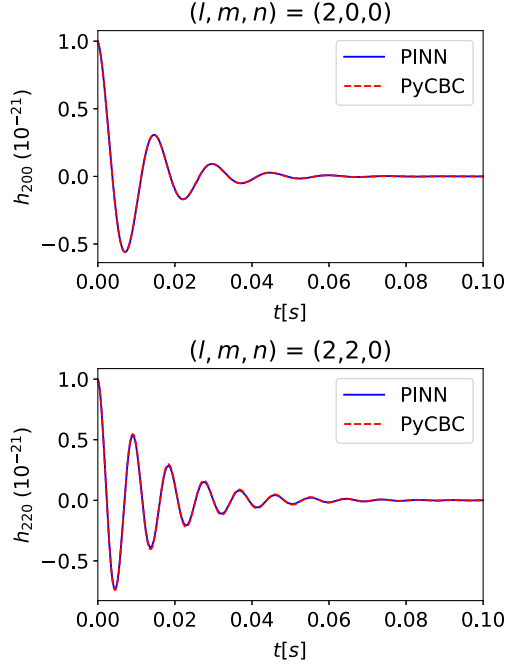
FIG. 5. Time-domain waveform comparison between the h_{200} (top) and h_{220} (bottom) strain predictions of the PINN and the reference by PYKERR, for a black hole of mass $M = 200M_\odot$ and angular momentum $J = 0.9M$.

PINN and the Python library PYKERR, which uses the tabulated values from [5] available at the GRIT Ringdown website [67].

The top and bottom panels of Fig. 5 respectively show the (2, 0, 0) and (2, 2, 0) waveforms obtained for the case of a black hole with parameters $M = 200M_\odot$ and $J = 0.9M$ in the time domain, while the corresponding Fourier transformed waveforms are shown in Fig. 6.

We quantify the discrepancy between the two waveforms by means of their match [68–70], through the library PYCBC [71]. The match can be used to estimate the SNR needed to distinguish the corresponding waveforms or, in other words, the signal loudness required for the difference between the two waveforms to drive parameter biases in a real detection scenario. This is given by [72–74]

$$F = 1 - \frac{D}{2\,\mathrm{SNR}^2_{\min}}, \tag{16}$$

where F is the match between the signals and D is the number of intrinsic parameters of the model. In our case these are $D = 4$: mass, spin, amplitude and phase. The match of two waveforms a and b is defined as their overlap maximized over time and phase shifts as

$$F(a, b) = \max_{t_c, \phi} \frac{(a|b)}{\sqrt{(a|a)(b|b)}}, \tag{17}$$

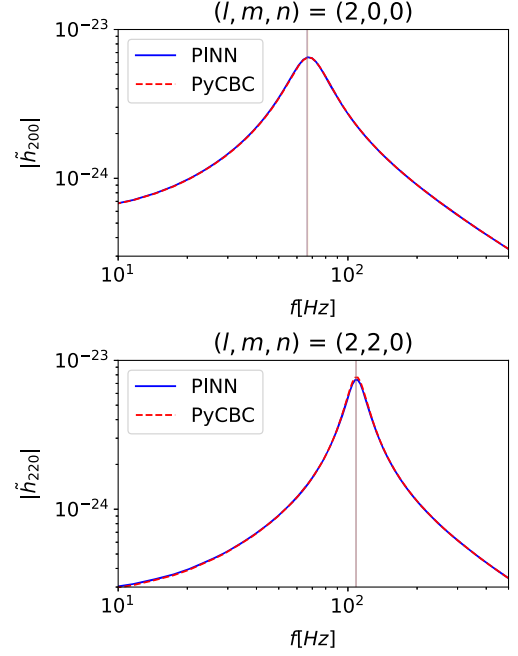where (a, b) denotes the inner product defined as

FIG. 6. Fourier transform of the h_{200} (top) and h_{220} (bottom) strain predictions, for a black hole of mass $M = 200M_\odot$ and angular momentum $J = 0.9M$. The vertical lines indicate the predictions by the PINN and by PYCBC. However, they cannot be told apart in the figure, since their frequencies only differ by 0.17 Hz and 0.07 Hz, respectively.

$$(a|b) = 4 \times \mathrm{Re} \int_{f_0}^{f_{\max}} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} df, \tag{18}$$

with $\tilde{a}(f)$ denoting the Fourier transform of $a(t)$, and $\tilde{a}^*(f)$ its complex conjugate. In addition, $S_n(f)$ is the one-sided noise power spectral density (PSD). In our case we use the aLIGOZeroDetHighPower model from the LALSimulation package [75], representative of the design sensitivity of Advanced LIGO. In the calculation of the match F, we numerically maximize the match over possible values of the phase difference in Eq. (14). Therefore, we can argue that the QNM extracted by the PINNs could be used for experimental parameter estimation as long as the SNR produced by the ringdown signal in the detector is lower than our estimation. Figure 7 shows the $\mathrm{SNR}_{\min}$ as a function of the black hole mass and spin. We assume a detection window of 300 ms with a sampling rate of 4096 Hz. We observe that $\mathrm{SNR}_{\min}$ well above 100 are obtained for spins lower than $\simeq 0.9$, typical of LIGO-Virgo observations. This indicates that our QNMs are suitable for parameter estimation in the near future, before the arrival of third-generation detectors like LISA or Einstein Telescope, when SNRs of $\mathcal{O}(1000)$ can be expected [66]. We note that the quality of the PINN modes significantly decays for higher values of the spin, as the black hole approaches extremality, especially for the case of the $(l, m, n) = (2, 2, 0)$ mode.
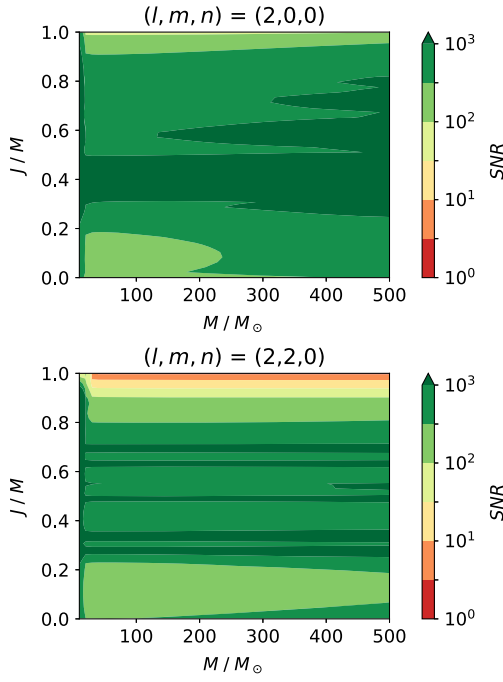
FIG. 7. Signal-to-noise ratio estimates to resolve the waveforms generated by using modes from the PINN and PYKERR.

V. DISCUSSION

We have used physics-informed neural networks [50] to solve the Teukolsky master equation and compute the first quasinormal modes of the Kerr geometry for arbitrary black-hole masses and spins. Our results show an accuracy typically below the percentual level as compared to the accepted values in the literature [40]. This validates the method as an acceptable approach to use for black-hole spectroscopy in the foreseeable future. We have focused here on the cases of the fundamental $(\ell, m, n) = (2, 2, 0)$ ringdown mode, dominant for quasicircular BBH merger remnants and the $(2, 0, 0)$ mode, which can serve as a smoking gun for highly eccentric mergers [64]. In addition, we have displayed the calculation of the overtone $(2, 2, 1)$ and the strongest higher angular mode in quasicircular BBH mergers $(3, 3, 0)$, as these are the most promising prospects to realize black-hole spectroscopy in the near future. While in this paper we have showcased the utility of PINNs in solving the astrophysically relevant case of the Teukolsky equation, this work can be regarded as a first step toward the computation of QNMs under more generic conditions (e.g. [76]).

PINNs are very versatile, general-purpose tools that, as the particular physics case discussed in this work has shown, can solve in a remarkably straightforward way quasinormal mode (eigenvalue) problems that would otherwise require a more involved numerical setup. Their implementation is very simple when compared to other techniques, requiring only a few lines of code in a suitable machine learning framework providing automatic differentiation capabilities. In addition, the method has also the advantage of being able to solve several equations with several eigenvalues at once, in our case the radial and angular equations with the eigenvalue ω and the separation constant A. As opposed to other methods, this generalization is trivially done by performing a weighted sum of each equation's contribution to the total loss function.

Moreover, the requirements on the type of equations to be used are very low, and PINNs are expected to solve very generic equations. In particular, they can be applied to any system of the form (8) with arbitrary coefficients $F_i(x), G_i(u)$. Remarkably, these coefficients do not even need to have an analytical expression, which makes this approach especially interesting for the computation of QNMs for complicated numerical background solutions, such as those involving exotic compact objects or in modified gravity. Therefore, we regard this paper as a preliminary step toward the use of PINNs for the extraction of QNMs in a variety of more complex scenarios in strong-field gravity.

The major drawbacks of a solution procedure based on PINNs are the computation time, which is usually larger than in other methods, and the accuracy of the results. However, for the physics case studied in this work these two drawbacks do not seem to be much of a concern. On the one hand, the very small size of the required network architectures makes the training affordably fast, with approximately 2 minutes per mode. On the other hand, the precision that we obtain, albeit not as high as that achieved by some other numerical techniques (e.g. spectral methods), is sufficient to be used on experimental detection data for black-hole spectroscopy and parameter estimation for current detector sensitivities, before the arrival of third-generation detectors like LISA or Einstein Telescope.

APPENDIX: PINN EQUATIONS

The coefficients for Eq. (8) that are actually solved by the PINN are

$$
\begin{aligned}
F_0(x) = &-a^4x^2\omega^2 - 2a^3mx^2\omega + a^2(-Ax^2 + x^2(4(r_+ + 1)\omega^2 + 2i(r_+ + 2)\omega + 2is(\omega + i) - 2) + x\omega^2 - \omega^2) \\
&+ 2am(r_+x^2(2\omega + i) - x(\omega + i) - \omega) + A(x - 1) \\
&- ir_+(2\omega + i)(x^2(s - 2i\omega + 1) - 2(s + 1)x + 2i\omega) + (s + 1)(x - 2i\omega), \\
F_1(x) = &\,2a^4x^4(x - i\omega) - 2ia^3mx^4 + a^2x^2(2r_+x^2(-1 + 2i\omega) - (s + 3)x^2 + 2x(s + i\omega + 2) - 4i\omega) \\
&+ 2iam(x - 1)x^2 + (x - 1)(2r_+x^2(1 - 2i\omega) + (s + 1)x^2 - 2(s + 1)x + 2i\omega), \\
F_2(x) = &\,a^4x^6 - 2a^2(x - 1)x^4 + (x - 1)^2x^2,
\end{aligned}
\tag{A1}
$$

$$
\begin{aligned}
G_0(u) = &\,4a^2(u^2 - 1)\omega^2 - 4a(u^2 - 1)\omega((u - 1)|m - s| + (u + 1)|m + s| + 2(s + 1)u) \\
&+ 4(A(u^2 - 1) + m^2 + 2msu + s((s + 1)u^2 - 1)) - 2(u^2 - 1)|m + s| \\
&- 2(u^2 - 1)|m - s|(|m + s| + 1) - (u - 1)^2|m - s|^2 - (u + 1)^2|m + s|^2, \\
G_1(u) = &-8a(u^2 - 1)^2\omega - 4(u^2 - 1)((u - 1)|m - s| + (u + 1)|m + s| + 2u), \\
G_2(u) = &-4(u^2 - 1)^2.
\end{aligned}
\tag{A2}
$$

[1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett. **116,** 061102 (2016).

[2] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, Phys. Rev. X **9,** 031040 (2019).

[3] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run, Phys. Rev. X **11,** 021053 (2021).

[4] The LIGO Scientific, the Virgo, and the KAGRA Collaborations, GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, arXiv:2111.03606.

[5] E. Berti, V. Cardoso, and C. M. Will, On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA, Phys. Rev. D **73,** 064030 (2006).

[6] G. Carullo, L. van der Schaaf, L. London, P. T. H. Pang, K. W. Tsang, O. A. Hannuksela, J. Meidam, M. Agathos, A. Samajdar, A. Ghosh, T. G. F. Li, W. Del Pozzo, and C. Van Den Broeck, Empirical tests of the black hole no-hair conjecture using gravitational-wave observations, Phys. Rev. D **98,** 104020 (2018).

[7] G. Carullo, W. Del Pozzo, and J. Veitch, Observational black hole spectroscopy: A time-domain multimode analysis of GW150914, Phys. Rev. D **99,** 123029 (2019).

[8] G. Carullo, Enhancing modified gravity detection from gravitational-wave observations using the parametrized ringdown spin expansion coefficients formalism, Phys. Rev. D **103,** 124043 (2021).

[9] G. Carullo, D. Laghi, N. K. Johnson-McDaniel, W. Del Pozzo, Ó. J. C. Dias, M. Godazgar, and J. E. Santos, Constraints on Kerr-Newman black holes from merger-ringdown gravitational-wave observations, Phys. Rev. D **105,** 062009 (2022).

[10] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, Testing the No-Hair Theorem with GW150914, Phys. Rev. Lett. **123,** 111102 (2019).

[11] J. C. Bustillo, P. D. Lasky, and E. Thrane, Black-hole spectroscopy, the no-hair theorem, and GW150914: Kerr versus Occam, Phys. Rev. D **103,** 024041 (2021).

[12] C. D. Capano, M. Cabero, J. Westerweck, J. Abedi, S. Kastha, A. H. Nitz, Y.-F. Wang, A. B. Nielsen, and B. Krishnan, Observation of a multimode quasi-normal spectrum from a perturbed black hole, arXiv:2105.05238.

[13] S. Chandrasekhar, *The Mathematical Theory of Black Holes,* Oxford Classic Texts in the Physical Sciences (Clarendon Press, 1998).

[14] K. D. Kokkotas and B. G. Schmidt, Quasi-normal modes of stars and black holes, Living Rev. Relativity **2**, 2 (1999).

[15] E. Berti, V. Cardoso, and A. O. Starinets, Quasinormal modes of black holes and black branes, Classical Quantum Gravity **26**, 163001 (2009).

[16] V. Cardoso and P. Pani, Testing the nature of dark compact objects: A status report, Living Rev. Relativity **22**, 4 (2019).

[17] W. H. Press, Long wave trains of gravitational waves from a vibrating black hole, Astrophys. J. **170**, L105 (1971).

[18] C. V. Vishveshwara, Stability of the Schwarzschild metric, Phys. Rev. D **1**, 2870 (1970).

[19] S. A. Teukolsky, Perturbations of a rotating black hole. I. Fundamental equations for gravitational, electromagnetic, and neutrino-field perturbations, Astrophys. J. **185**, 635 (1973).

[20] S. Chandrasekhar and S. L. Detweiler, The quasi-normal modes of the Schwarzschild black hole, Proc. R. Soc. A **344**, 441 (1975).

[21] O. Dreyer, B. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, Black-hole spectroscopy: Testing general relativity through gravitational-wave observations, Classical Quantum Gravity **21**, 787 (2004).

[22] E. Berti *et al.*, Testing general relativity with present and future astrophysical observations, Classical Quantum Gravity **32**, 243001 (2015).

[23] V. Baibhav and E. Berti, Multimode black hole spectroscopy, Phys. Rev. D **99**, 024005 (2019).

[24] E. Berti, J. Cardoso, V. Cardoso, and M. Cavaglià, Matched filtering and parameter estimation of ringdown waveforms, Phys. Rev. D **76**, 104044 (2007).

[25] M. Giesler, M. Isi, M. A. Scheel, and S. A. Teukolsky, Black Hole Ringdown: The Importance of Overtones, Phys. Rev. X **9**, 041060 (2019).

[26] M. Maggiore, C. Van Den Broeck, N. Bartolo, E. Belgacem, D. Bertacca, M. A. Bizouard, M. Branchesi, S. Clesse, S. Foffa, J. García-Bellido, S. Grimm, J. Harms, T. Hinderer, S. Matarrese, C. Palomba, M. Peloso, A. Ricciardone, and M. Sakellariadou, Science case for the Einstein Telescope, J. Cosmol. Astropart. Phys. 03 (2020) 050.

[27] P. Amaro-Seoane *et al.*, Laser interferometer space antenna, arXiv:1702.00786.

[28] The LIGO Scientific, the Virgo, and the KAGRA Collaborations, Tests of general relativity with GWTC-3, arXiv:2112.06861.

[29] G. W. Horndeski, Second-order scalar-tensor field equations in a four-dimensional space, Int. J. Theor. Phys. **10**, 363 (1974).

[30] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Modified gravity and cosmology, Phys. Rep. **513**, 1 (2012).

[31] V. Cardoso and J. P. S. Lemos, Quasinormal modes of Schwarzschild anti-de Sitter black holes: Electromagnetic and gravitational perturbations, Phys. Rev. D **64**, 084017 (2001).

[32] V. Cardoso and J. P. S. Lemos, Quasinormal modes of the near extremal Schwarzschild-de Sitter black hole, Phys. Rev. D **67**, 084020 (2003).

[33] V. Cardoso, A. S. Miranda, E. Berti, H. Witek, and V. T. Zanchin, Geodesic stability, Lyapunov exponents and quasinormal modes, Phys. Rev. D **79**, 064016 (2009).

[34] R. A. Konoplya and A. Zhidenko, Quasinormal modes of black holes: From astrophysics to string theory, Rev. Mod. Phys. **83**, 793 (2011).

[35] H. Yang, D. A. Nichols, F. Zhang, A. Zimmerman, Z. Zhang, and Y. Chen, Quasinormal-mode spectrum of Kerr black holes and its geometric interpretation, Phys. Rev. D **86**, 104006 (2012).

[36] P. Pani, E. Berti, and L. Gualtieri, Gravitoelectromagnetic Perturbations of Kerr-Newman Black Holes: Stability and Isospectrality in the Slow-Rotation Limit, Phys. Rev. Lett. **110**, 241103 (2013).

[37] O. J. C. Dias, M. Godazgar, and J. E. Santos, Linear Mode Stability of the Kerr-Newman Black Hole and Its Quasinormal Modes, Phys. Rev. Lett. **114**, 151101 (2015).

[38] T. Regge and J. A. Wheeler, Stability of a Schwarzschild singularity, Phys. Rev. **108**, 1063 (1957).

[39] F. J. Zerilli, Effective Potential for Even Parity Regge-Wheeler Gravitational Perturbation Equations, Phys. Rev. Lett. **24**, 737 (1970).

[40] E. W. Leaver, An analytic representation for the quasi-normal modes of Kerr black holes, Proc. R. Soc. A **402**, 285 (1985).

[41] W. Krivan, P. Laguna, P. Papadopoulos, and N. Andersson, Dynamics of perturbations of rotating black holes, Phys. Rev. D **56**, 3395 (1997).

[42] E. Pazos-Ávalos and C. O. Lousto, Numerical integration of the Teukolsky equation in the time domain, Phys. Rev. D **72**, 084022 (2005).

[43] E. Harms, S. Bernuzzi, and B. Brügmann, Numerical solution of the 2 + 1 Teukolsky equation on a hyperboloidal and horizon penetrating foliation of Kerr and application to late-time decays, Classical Quantum Gravity **30**, 115013 (2013).

[44] V. Ferrari and B. Mashhoon, New approach to the quasi-normal modes of a black hole, Phys. Rev. D **30**, 295 (1984).

[45] S. Iyer and C. M. Will, Black hole normal modes: A WKB approach. 1. Foundations and application of a higher order WKB analysis of potential barrier scattering, Phys. Rev. D **35**, 3621 (1987).

[46] H. T. Cho, A. S. Cornell, J. Doukas, T. R. Huang, and W. Naylor, A new approach to black hole quasinormal modes: A review of the asymptotic iteration method, Adv. Theor. Math. Phys. **2012**, 281705 (2012).

[47] G. B. Cook and M. Zalutskiy, Gravitational perturbations of the Kerr geometry: High-accuracy study, Phys. Rev. D **90**, 124021 (2014).

[48] A. Jansen, Overdamped modes in Schwarzschild-de Sitter and a *Mathematica* package for the numerical computation of quasinormal modes, Eur. Phys. J. Plus **132**, 546 (2017).

[49] R. Emparan, R. Suzuki, and K. Tanabe, Decoupling and non-decoupling dynamics of large D black holes, J. High Energy Phys. 07 (2014) 113.

[50] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. **378**, 686 (2019).

[51] A. Paszke *et al.*, PyTorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems 32* (Curran Associates, Inc., 2019), pp. 8024–8035.

[52] K. Hornik, M. Stinchcombe, and H. White, Multilayer feedforward networks are universal approximators, Neural Netw. **2**, 359 (1989).

[53] C. Rao, H. Sun, and Y. Liu, Physics-informed deep learning for computational elastodynamics without labeled data, J. Eng. Mech. **147**, 04021043 (2021).

[54] H. Jin, M. Mattheakis, and P. Protopapas, Unsupervised neural networks for quantum eigenvalue problems, in *2020 NeurIPS Workshop on Machine Learning and the Physical Sciences* (NeurIPS, 2020), 10.48550/arxiv.2010.05075.

[55] M. Mattheakis, D. Sondak, A. S. Dogra, and P. Protopapas, Hamiltonian neural networks for solving equations of motion, Phys. Rev. E **105**, 065305 (2022).

[56] A. Övgün, I. Sakallı, and H. Mutuk, Quasinormal modes of dS and AdS black holes: Feedforward neural network method, Int. J. Geom. Method Mod. Phys. **18**, 2150154 (2021).

[57] A. M. Ncube, G. E. Harmsen, and A. S. Cornell, Investigating a new approach to quasinormal modes: Physics-informed neural networks, in 65th Annual Conference of the South African Institute of Physics (2021), arXiv:2108.05867.

[58] A. S. Cornell, A. Ncube, and G. Harmsen, Using physics-informed neural networks to compute quasinormal modes, Phys. Rev. D **106**, 124047 (2022).

[59] D. Kingma and J. Ba, Adam: A method for stochastic optimization, in *International Conference on Learning Representations* (2014), 10.48550/arxiv.1412.6980.

[60] J. C. Bustillo, N. Sanchis-Gual, A. Torres-Forné, and J. A. Font, Confusing Head-On Collisions with Precessing Intermediate-Mass Binary Black Hole Mergers, Phys. Rev. Lett. **126**, 201101 (2021).

[61] I. Romero-Shaw, P. D. Lasky, E. Thrane, and J. C. Bustillo, GW190521: Orbital eccentricity and signatures of dynamical formation in a binary black hole merger signal, Astrophys. J. **903**, L5 (2020).

[62] V. Gayathri, J. Healy, J. Lange, B. O'Brien, M. Szczepańczyk, I. Bartos, M. Campanelli, S. Klimenko, C. O. Lousto, and R. O'Shaughnessy, Eccentricity estimate for black hole mergers with numerical relativity simulations, Nat. Astron. **6**, 344 (2022).

[63] R. Gamba, M. Breschi, G. Carullo, S. Albanesi, P. Rettegno, S. Bernuzzi, and A. Nagar, GW190521 as a dynamical capture of two nonspinning black holes, Nat. Astron. **7**, 11 (2023).

[64] J. C. Bustillo, N. Sanchis-Gual, A. Torres-Forné, J. A. Font, A. Vajpeyi, R. Smith, C. Herdeiro, E. Radu, and S. H. W. Leong, GW190521 as a Merger of Proca Stars: A Potential New Vector Boson of 8.7×10^{-13} eV, Phys. Rev. Lett. **126**, 081101 (2021).

[65] M. Cabero, J. Westerweck, C. D. Capano, S. Kumar, A. B. Nielsen, and B. Krishnan, Black hole spectroscopy in the next decade, Phys. Rev. D **101**, 064044 (2020).

[66] S. Bhagwat, C. Pacilio, E. Barausse, and P. Pani, Landscape of massive black-hole spectroscopy with LISA and the Einstein Telescope, Phys. Rev. D **105**, 124063 (2022).

[67] GRIT, Ringdown data and routines (2022), https://centra.tecnico.ulisboa.pt/network/grit/files/ringdown/.

[68] C. Cutler and E. E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form?, Phys. Rev. D **49**, 2658 (1994).

[69] T. A. Apostolatos, Search templates for gravitational waves from precessing, inspiraling binaries, Phys. Rev. D **52**, 605 (1995).

[70] L. S. Finn, Detection, measurement and gravitational radiation, Phys. Rev. D **46**, 5236 (1992).

[71] A. Nitz *et al.*, gwastro/pycbc: v2.0.5 release of PyCBC (2022).

[72] L. Lindblom, B. J. Owen, and D. A. Brown, Model waveform accuracy standards for gravitational wave data analysis, Phys. Rev. D **78**, 124020 (2008).

[73] K. Chatziioannou, A. Klein, N. Yunes, and N. Cornish, Constructing gravitational waves from generic spin-precessing compact binary inspirals, Phys. Rev. D **95**, 104004 (2017).

[74] K. Chatziioannou, A. Klein, N. Cornish, and N. Yunes, Analytic Gravitational Waveforms for Generic Precessing Binary Inspirals, Phys. Rev. Lett. **118**, 051101 (2017).

[75] LIGO Scientific Collaboration, LIGO Algorithm Library— LALSuite, free software (GPL) (2018).

[76] M. H.-Y. Cheung, L. W.-H. Poon, A. K.-W. Chung, and T. G. F. Li, Ringdown spectroscopy of rotating black holes pierced by cosmic strings, J. Cosmol. Astropart. Phys. 02 (2021) 040.