A toy model for testing finite element methods to simulate extreme-mass-ratio binary systems

Carlos F Sopuerta\textsuperscript{1,2,3}, Pengtao Sun\textsuperscript{4,5}, Pablo Laguna\textsuperscript{1,2,3,6} and Jinchao Xu\textsuperscript{1,4,5}

\textsuperscript{1} Institute for Gravitational Physics and Geometry, Penn State University, University Park, PA 16802, USA
\textsuperscript{2} Center for Gravitational Wave Physics, Penn State University, University Park, PA 16802, USA
\textsuperscript{3} Department of Astronomy and Astrophysics, Penn State University, University Park, PA 16802, USA
\textsuperscript{4} Department of Mathematics, Penn State University, University Park, PA 16802, USA
\textsuperscript{5} Center for Computational Mathematics and Applications, Penn State University, University Park, PA 16802, USA
\textsuperscript{6} Department of Physics, Penn State University, University Park, PA 16802, USA

Received 25 July 2005, in final form 21 October 2005
Published 14 December 2005
Online at stacks.iop.org/CQG/23/251

Abstract

Extreme-mass-ratio binary systems, binaries involving stellar mass objects orbiting massive black holes, are considered to be a primary source of gravitational radiation to be detected by the space-based interferometer LISA. The numerical modelling of these binary systems is extremely challenging because the scales involved expand over several orders of magnitude. One needs to handle large wavelength scales comparable to the size of the massive black hole and, at the same time, to resolve the scales in the vicinity of the small companion where radiation reaction effects play a crucial role. Adaptive finite element methods, in which quantitative control of errors is achieved automatically by finite element mesh adaptivity based on \textit{a posteriori} error estimation, are a natural choice that has great potential for achieving the high level of adaptivity required in these simulations. To demonstrate this, we present the results of simulations of a toy model, consisting of a point-like source orbiting a black hole under the action of a scalar gravitational field.

PACS numbers: 02.60.Cb, 02.70.Dh, 04.25.Dm, 04.30.Db

(Some figures in this article are in colour only in the electronic version)

1. Introduction

As we enter the era of gravitational wave astronomy, a number of experiments to detect and study gravitational waves have been set up and some others are presently under development.
Among the second group we find one of the experiments that is presently attracting a considerable amount of attention: the Laser Interferometer Space Antenna (LISA) [2–5], a collaboration between ESA and NASA that is scheduled to be launched in the next decade. Extreme-mass-ratio binaries (EMRBs) are considered to be a primary source of gravitational radiation to be detected by LISA [6, 7]. They consist of a ‘small’ object, such as a main sequence star, a stellar mass black hole, or a neutron star, with mass \( m \) ranging from \( 1M_\odot \) to \( 10^2M_\odot \), orbiting a massive black hole (MBH) with mass \( M \) ranging from \( 10^3M_\odot \) (if we consider the case of intermediate mass black holes) to \( 10^9M_\odot \) (the case of big supermassive black holes sitting in the centre of galaxies). This translates to EMRBs with mass ratios, \( \mu = m/M \), in the range \( 10^{-3}–10^{-9} \). In order to exploit these types of systems through LISA, it is crucial to have a good theoretical understanding of the evolution of these systems, good enough to produce accurate waveform templates in support of data analysis efforts.

Because there is no significant coupling between the strong curvature effects between the MBH and its companion, relativistic perturbation theory is a well-suited tool to study EMRBs. Clearly, the accuracy of this approximation depends on the smallness of the ratio \( \mu \). The goal is to study the perturbations generated by the small body in the (background) gravitational field of the MBH, and how these perturbations affect the motion of the small body itself, which follows the geodesics of the perturbed spacetime; that is, studying how the presence of the small body affects its own trajectory. This problem is usually known in the literature as the radiation reaction problem. This is an old problem and several approaches to dealing with it have been proposed (see the recent reviews by Poisson [8, 9], Detweiler [10] and Mino [11–13]). A pragmatic approach is to use energy–momentum balance arguments [14–19]. Under this approach, one estimates the changes in the small body constants of motion by computing the fluxes of energy and angular momentum at infinity and through the MBH horizon. This approach works well in the adiabatic regime, when the time scale of the radiation reaction is much bigger than the orbital time scale. Until now it has dealt with special orbits of the Kerr black hole, but it has not yet produced results for generic orbits because of the difficulty of adjusting the third constant of motion, the Carter constant, present in generic geodesics of the Kerr spacetime. However, there have been some recent advances in this direction [20–23].

An alternative approach consists in trying to describe the radiation reaction effects on the small body as the action of a local self-force that is responsible for the deviations from the geodesic motion. A rigorous formulation of this concept has been given, for the first time, by Mino, Sasaki and Tanaka [24], and later, adopting an axiomatic approach, by Quinn and Wald [25]. These works give a formal prescription to compute the self-force. For the practical implementation of this prescription some techniques have been proposed (see [26] for a recent progress report): the mode-sum scheme [27–30], and a regularization scheme based on zeta-function regularization techniques [31].

In this paper, we advocate an alternative, and at the same time complementary, approach: the direct numerical integration of the (linearized) evolution equations for the gravitational field together with the evolution equations for the small body, which form a system of partial differential equations (PDEs) coupled to a system of ordinary differential equations (ODEs). We can already find in the literature an attempt to use numerical methods for simulating EMRBs [32]. In this work, the perturbations are computed by using the full general relativistic equations in the framework of the characteristic formulation (in contrast to a Cauchy-type formulation), with the small body being described by an energy–momentum distribution that moves ‘rigidly’ along a geodesic of the numerically computed spacetime. The main drawback of this work is the size of the small body, which is bigger than the MBH horizon. Another approach that has been used in the literature [33] is to describe the gravitational field by
using the Teukolsky formalism [34] implemented in the time domain (using a numerical code introduced in [35]) and by modelling the small body by smearing the singularities in the source term by the use of narrow Gaussian distributions. This work also has the problem that the size of the small body is too big (in comparison with the size of the MBH). This shows the main underlying difficulty in the numerical simulation of EMRBs, namely, the problem involves a vast range of physical scales (spatial and temporal) that expand over several orders of magnitude. Specifically, one need not only handle large wavelength scales comparable to the massive black hole, but also resolve the scales in the vicinity of the small object where radiation reaction effects play a crucial role.

An obvious conclusion we can extract from these facts is that, in order to carry out successful numerical simulations of EMRBs, we need a high degree of adaptivity. Our proposal is the use of the finite element method (FEM) as a natural choice to achieve this high level of adaptivity. Finite element methods have only been used occasionally in general relativity (see [36, 37]). To demonstrate that these numerical techniques have great potential in leading to successful simulations of EMRBs, we present results from simulations of a toy problem consisting of a point-like source orbiting a black hole in scalar gravitation in \(2 + 1\) dimensions. Our aim is to test FEM techniques in a simple representative problem that possesses the main ingredients and challenges of the astrophysical EMRBs; that is, at this stage we are basically conducting a feasibility study. In our toy model, the spacetime metric is fixed and aims at describing the gravitational field of a non-rotating black hole. For computational efficiency, we have made a reduction from three spatial dimensions to two. In this reduction, which we explain in detail later, the metric we work with is no longer a solution of the Einstein vacuum equations but it keeps the important property that its geodesics coincide with the equatorial geodesics of the Schwarzschild spacetime. This metric is not dynamical but fixed. The dynamical gravitational field is described by a scalar field on this spacetime, satisfying a wave-like equation with a source term that describes the presence of the small object. An important ingredient of our model is the use of a particle description for the small object. The equations of motion of this particle are the geodesics of the fixed spacetime metric modified by the presence of (spatial) components of the gradient of the gravitational scalar field. In this way, we have that the particle orbits the black hole subject to radiation reaction: indeed, the particle generates the scalar gravitational field which affects its own motion.

In this work, we compare numerical simulations that use the simple classical FEM with simulations that use an adaptive mesh FEM. The essence of the adaptive mesh technique is to produce real-time local mesh coarsening or refinement to achieve the desired level of smoothness in the solution. To that end, a good \textit{a posteriori} error estimator to predict the regions in the computational domain where changes take place rapidly is extremely important. There are several ways to approach this problem. Theoretical research on this subject [38] has found that the Hessian matrix of the numerical solution can accurately predict where the steepest gradients of the solution would take place. We demonstrate that this technique when applied to our toy model easily captures the dynamics of the field in the vicinity of the particle since around the particle location either the field or the source term will change much more than anywhere else. We then refine the local area surrounding the particle and resolve it more accurately so that we are able to achieve our final target.

The plan of this paper is as follows. In section 2 we introduce the particular theoretical model we want to study; namely, the description of the gravitational field and of the point-like source, the computational set-up, and an energy-balance test that can be used to test the numerical computations. In section 3 we describe the computational techniques that we use for the time-domain simulations of our theoretical model: the FEM, the discretization of our equations using finite elements, the numerical method for solving the equations of
motion of the particle, and finally, the adaptive finite element method (AFEM). In section 4 we present and discuss the numerical results of the simulations. Here, we distinguish between the simulations that use the classical FEM and those that use the AFEM. We devote section 5 to discussing the main results of this work and the future perspectives that it opens up.

2. Theoretical aspects of the toy model

In the following, we describe in detail the main aspects of our toy model, namely, (i) the gravitational theory we use, (ii) the derivation of the equations to be solved, (iii) the particle’s description, and (iv) the computational set-up. We use physical units in which \( c = G = 1 \). We use lowercase Greek letters for spacetime indices and lowercase Latin letters for spatial indices. In sections 2.1 and 2.2 indices run as \( \mu, \nu, \ldots = 0–3; i, j, \ldots = 1–3 \).

2.1. Scalar gravity

Scalar gravitation is a theory of gravity that, although it cannot be applied to our physical world (we know it cannot fit all the available experimental data), is a good laboratory for numerical relativity due to its simplicity. Of particular interest for our purposes is the fact that (scalar) gravitational waves exist in scalar gravity. In this theory, the gravitational field is described by a scalar field \( \Phi(x^\mu) \) on a spacetime geometry described by a non-dynamical metric tensor \( g_{\mu\nu} \), which we just prescribe (the version presented in problem 7.1 of the textbook by Misner, Thorne and Wheeler [39] and in [40–42] considers only the case in which the non-dynamical spacetime is the flat spacetime). The scalar field does not affect the spacetime structure defined by \( g_{\mu\nu} \). We are interested in studying the evolution of a particle-like object orbiting a black hole in this theory. To that end, we consider a background spacetime metric \( g_{\mu\nu} \) describing the geometry of a non-rotating black hole, and a particle of mass \( m \) that follows the worldline \( z^\mu(\tau) \), where \( \tau \) denotes the particle proper time. The action of this particle–field system is

\[
S = \int \sqrt{-g} \mathcal{L} \, d^4x = -\int \sqrt{-g} \left( \frac{1}{8\pi} g^{\mu\nu} \nabla_\mu \Phi \nabla_\nu \Phi - \rho \, e^\Phi \right) \, d^4x ,
\]

where the comoving density \( \rho \), describing the particle, is given by

\[
\rho = \int \frac{m}{\sqrt{-g}} g^{44} [x^4 - z^4(\tau)] \, d\tau = \frac{m}{u^t} \sqrt{-g} \delta^4 [x^4 - z^4(t)],
\]

where \( t \) is a time coordinate and \( u^t \) is the corresponding component of the velocity of the particle:

\[
u^\mu = \frac{dz^\mu(\tau)}{d\tau} \implies u^t = \frac{dt}{d\tau} .
\]

Varying the action (1) with respect to the scalar field \( \Phi \), we obtain the gravitational field equation:

\[
g^{\mu\nu} \nabla_\mu \nabla_\nu \Phi = 4\pi e^\Phi \rho .
\]

Varying now the action with respect to \( z^\mu \), we obtain the particle’s equations of motion

\[
u^\nu \nabla_\nu u^\mu + (g^{\mu\nu} + u^\mu u^\nu) \nabla_\nu \Phi = 0.
\]

Equations (4) and (5) form a coupled system of partial and ordinary differential equations respectively. Equation (4) is a hyperbolic and nonlinear equation that describes the dynamical gravitational field whereas (5) is a system of equations of motion for the particle. In the absence of \( \Phi \) the particle follows the geodesics of the background spacetime \( g_{\mu\nu} \), but in the presence of \( \Phi \) it will no longer follow the geodesic of the background. The way in which \( \Phi \) influences
the motion of the particle is through its gradient projected orthogonally to the 4-velocity of the particle. At the same time, the particle influences the gravitational field $\Phi$ through the source term $\rho$. The coupling between the particle and the field is where the radiation reaction is encoded: the particle induces a nonzero $\Phi$, and this field induces a deviation in the motion of the particle from the background geodesic motion, and at the same this motion affects the evolution of the field, and so on. Therefore, the situation is the same as in the general relativistic case where the particle is treated as a perturbation on the background spacetime of the MBH, and where the radiation reaction mechanism works in the same way.

This system has a well-defined total energy–momentum tensor $T_{\mu\nu}$, which is given by

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta (\sqrt{-g} \mathcal{L})}{\delta g_{\mu\nu}}. \quad (6)$$

It has two differentiated components, one associated with the scalar gravitational field, $(\Phi)T_{\mu\nu}$, and the other with the particle, $(\rho)T_{\mu\nu}$: $T_{\mu\nu} = (\Phi)T_{\mu\nu} + (\rho)T_{\mu\nu}$. Their expressions are, respectively,

$$(\Phi)T_{\mu\nu} = \frac{1}{4\pi} \left( \nabla_\mu \Phi \nabla_\nu \Phi - \frac{1}{2} g_{\mu\nu} \nabla^\sigma \Phi \nabla_\sigma \Phi \right), \quad (\rho)T_{\mu\nu} = \rho \epsilon^\Phi u_\mu u_\nu. \quad (7)$$

One can show that the energy–momentum conservation equation

$$\nabla_\mu T^{\mu\nu} = 0, \quad (9)$$

follows provided matter conservation holds, that is, $\nabla_\mu (\rho u^\mu) = 0$.

### 2.2. 3+1 decomposition of the equations

In order to numerically solve equations (4), (5) it is convenient to rewrite them in a 3 + 1 language, which allows for an initial-value problem formulation. To that end, we follow the usual 3 + 1 decomposition used in numerical relativity; that is, we write our spacetime line-element as follows,

$$ds^2 = -\alpha^2 dt^2 + h_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt), \quad (10)$$

where $h_{ij}$ is the spatial metric of the $\{t = \text{const}\}$ hypersurfaces, $\alpha$ denotes the lapse function and $\beta^i$ is the shift vector. The normal to the hypersurfaces is

$$n_\mu = (-\alpha, 0), \quad n^\mu = \frac{1}{\alpha} (1, -\beta^i). \quad (11)$$

The covariant and contravariant components of the spacetime metric are

$$(g_{\mu\nu}) = \begin{pmatrix} -\alpha^2 + \beta^i \beta^j \alpha h_{ij} / \alpha & \beta^i \beta^j \alpha h_{ij} / \alpha \\ \beta^i h_{ij} / \alpha & h_{ij} / \alpha \end{pmatrix}, \quad (g^{\mu\nu}) = \begin{pmatrix} -\alpha^{-2} & \alpha^{-2} \beta^i / \alpha \\ \alpha^{-2} \beta^i / \alpha & h^{ij} - \alpha^{-2} \beta^i \beta^j / \alpha \end{pmatrix}. \quad (12)$$

where $h^{ij}$ is the inverse of $h_{ij}$. Then, $h^{\mu\nu} = g^{\mu\nu} + n^\mu n_\nu$ is the projector orthogonal to the hypersurfaces $\{t = \text{const}\}$, which also contains the induced metric that these hypersurfaces inherit: its spatial components coincide with $h_{ij}$.

Equation (4) for $\Phi$ is second order in space and in time. We will split it into two equations that are first order in time. To that end, we introduce the following definition:

$$\Pi = n^\mu \partial_\mu \Phi = \frac{1}{\alpha} (\partial_t - \beta^i \partial_i) \Phi. \quad (13)$$
Then, equation (4) can be split into the following two equations:
\[
(\partial_t - \beta^i \partial_i) \Phi = \alpha \Pi, \\
(\partial_t - \beta^i \partial_i) \Pi = \alpha (K \Pi + a^i D_i \Phi + \Delta \Phi - 4\pi e^\Phi \rho).
\]
In the second equation, \(K\) denotes the trace of the extrinsic curvature \(K_{\mu\nu}\) of the hypersurfaces \([t = \text{const}]\):
\[
K_{\mu\nu} = - \frac{1}{2} \mathcal{G}_\mu h_{\mu\nu} \implies K = -\nabla_\mu n^\mu.
\]
The symbols \(D_i\) and \(\Delta\) denote the covariant derivative and Laplacian associated with the induced metric \(h_{ij}\). Then we can write
\[
\Delta \Phi = h^{ij} D_i D_j \Phi = \frac{1}{\sqrt{h}} \partial_i (\sqrt{h} h^{ij} \partial_j \Phi).
\]
Moreover, \(a^i\) denotes the spatial components of \(h^{\mu\nu} n_\sigma \nabla_\sigma n^\nu\). After some calculations we find the following expression for \(a^i\):
\[
a^i = h^{ij} D_j \ln \alpha = h^{ij} \partial_j \ln \alpha.
\]
Now, let us perform the 3 + 1 decomposition of the equations for the trajectory of the particle. Our coordinate system \((t, x^i)\) is, in principle, not adapted to the particle trajectory in the sense that \(t\) is not the proper time and \(x^i\) are not comoving coordinates. Then, in this coordinate system the trajectory of the particle is given by \((t, x^i = z^i(t))\), and the unit tangent velocity vector is
\[
u^\mu \partial_\mu = u^t \partial_t + u^i \partial_i.
\]
But also
\[
u^\mu \partial_\mu = \partial_t \implies u^t = \frac{dz^i(t)}{dt} = v^i \quad \text{where} \quad v^i = \frac{dz^i}{dt}.
\]
Then, on the trajectory of the particle the following holds:
\[
u^i = \frac{dz^i(t)}{dt} \frac{dt}{dr} = v^i u^t, \quad \text{where} \quad v^i = \frac{dz^i}{dt}.
\]
Due to the fact that \(u^\mu\) is a unit timelike vector field \((g_{\mu\nu} u^\mu u^\nu = -1)\) we do not need to solve for all the components. One possibility is to solve the particle equations for the quantities \((z^i(t), v^i(t))\). Then, the equations that we get are
\[
d\frac{d}{dr} z^i = v^i,
\]
\[
d\frac{d}{dr} v^i = f^i_g + f^i_\Phi,
\]
where
\[
f^i_g = \frac{1}{(u^t)^2} \left( v^i \Gamma^\rho_{\mu\nu} \Gamma_{\rho\mu} - \Gamma_{\rho\mu} \right) u^\rho u^\nu,
\]
\[
f^i_\Phi = \frac{1}{(u^t)^2} \left[ v^i (g^{ij} u^j + u^i u^j) - (g^{ij} + u^i u^j) |\partial_\mu \Phi, \right.
\]
and the velocity component \(u^i\) can be written in terms of \(v^i\) and the spacetime metric as follows:
\[
(u^t)^2 = -(g_{tt} + 2g_{ti} v^i + g_{ij} v^i v^j)^{-1}.
\]
The term \(f^i_g\) gives the contribution to the geodesic motion in the background spacetime \(g_{\mu\nu}\), whereas the term \(f^i_\Phi\) describes the deviation from the geodesic motion due to the action of the scalar gravitational field \(\Phi\).
2.3. Pseudo-Schwarzschild black hole background

We want the background metric $g_{\mu\nu}$ to represent the geometry of a non-rotating black hole. In 3 + 1 we can use the Schwarzschild metric. However, to reduce the computational cost of our simulations, we reduce the space of our toy model to two spatial dimensions. There are several ways in which this reduction can be performed. For instance, one can start directly with the 2 + 1 metric obtained from the Schwarzschild metric on the hypersurface $z = 0$. Another possibility would be to start from the equations in the 3 + 1 set-up, expanding all the different terms and then, to neglect all the derivatives with respect to coordinate $z$ and the components of the different objects in that direction, and finally to set $z = 0$. In performing these dimensional reductions, the metrics we obtain are no longer solutions of Einstein’s equations in the dimensionally reduced spacetime, but this is not an issue in our toy model since the spacetime metric is not a dynamical object. The important thing is that the metric we obtain from these reductions keeps most of the important properties of the Schwarzschild metric, in particular, the fact that from the two different reductions mentioned above the equations for the geodesics are the same, and they coincide with the geodesics of Schwarzschild in the plane $z = 0$. The equations for $\Phi$, when we introduce the expression of the dimensionally reduced metrics, would be in general different. In this work, we use the first possibility for the dimensional reduction. As a consequence of this reduction the indices we use in this subsection run as follows: $\mu, \nu, \ldots = 0–2; i, j, \ldots = 1, 2$.

In what follows, we specify the form of the 2 + 1 background. We start from the Schwarzschild metric in Cartesian Kerr–Schild coordinates but reducing a space dimension. That is, our background metric is described by the following line element,

$$\eta_{\mu\nu} \, dx^\mu \, dx^\nu = -dt^2 + \delta_{ij} \, dx^i \, dx^j + 2H \ell_\mu \ell_\nu \, dx^\mu \, dx^\nu,$$

(27)

where $\eta_{\mu\nu} = \text{diag}(-1, 1, 1)$ is the 2 + 1 Minkowski metric, $\ell^\mu$ is a future-directed light-like vector field (both with respect to the Minkowski and Schwarzschild metric), and $H$ is a scalar. In Cartesian Kerr–Schild coordinates these three objects are given by

$$\eta_{\mu\nu} \, dx^\mu \, dx^\nu = -dt^2 + \delta_{ij} \, dx^i \, dx^j, \quad \ell_\mu \, dx^\mu = -dt + \frac{x^i}{r} \, dx^i, \quad H = \frac{M}{r},$$

(28)

where $M$ is the black hole mass and $r = \sqrt{b_{ij} x^i x^j}$.

We choose the $\{t = \text{const}\}$ foliation of the 2 + 1 spacetime. Then, the values of the relevant 2 + 1 quantities that result from this choice are

$$\alpha^2 = \frac{1}{1 + 2H} = \frac{1}{1 + \frac{2M}{r}},$$

(29)

$$\beta^i = -\frac{2H}{1 + 2H} \xi^i, \quad \xi^i = \frac{x^i}{r},$$

(30)

$$h_{ij} = \delta_{ij} + 2H \ell_i \ell_j, \quad h^{ij} = \delta^{ij} - \frac{2H}{1 + 2H} \xi^i \xi^j,$$

(31)

$$\sqrt{h} = \sqrt{\det(h_{ij})} = \sqrt{1 + 2H} = \sqrt{1 + \frac{2M}{r}},$$

(32)

$$K = -\frac{2M}{r^3} \left(1 + \frac{2M}{r}\right)^{-3/2},$$

(33)

$$a^i = -h^{ij} \frac{\partial_j H}{1 + 2H} = \frac{H}{r} \alpha^2 \xi^i = \frac{1}{(1 + \frac{2M}{r})^2} \frac{M x^i}{r^3}.$$  

(34)
As happens in any Kerr–Schild metric, the determinant of the spacetime metric is equal to $-1$, and hence we have the following relation between the lapse and the determinant of the spatial metric, $\sqrt{h} = 1/\alpha$; this agrees with equations (29), (32).

The Dirac delta function that appears in the source $\rho$ is regularized by using a Gaussian distribution,

$$\delta^2[x^i - \zeta^i(t)] \approx \frac{1}{(\sqrt{2\pi}\sigma)^2} e^{-\frac{R^2}{2\sigma^2}},$$

where

$$R^2 = \sum_{i=1}^2 [x^i - \zeta^i(t)]^2.$$  \hfill (35)

The source function $\rho$ takes the form

$$\rho = \frac{m}{(\sqrt{2\pi}\sigma)^2 u^i u^j \sqrt{h}} e^{-\frac{R^2}{2\sigma^2}} = \frac{m}{(\sqrt{2\pi}\sigma)^2 u^i} e^{-\frac{R^2}{2\sigma^2}}.$$  \hfill (37)

As we have already mentioned, $m$ is the total rest mass of the particle. We can recover this quantity, which is a constant of motion, from the following expression,

$$m = \int \gamma \rho \sqrt{h} \, d^2x.$$  \hfill (38)

where $\gamma = \sqrt{1 + h^{ij} u_i u_j} = au^i$. Note that by making the approximation (35), the exact character of this relation does not change due to the fact that the integral of the Dirac delta and a given Gaussian over the whole two-dimensional (2D) domain yields the same result. On the other hand, the matter conservation relation $\nabla_\mu (\rho u^\mu) = 0$ ensures that $m$, as defined by equation (38), is a conserved quantity.

The next thing we need is the explicit form of the equations of motions for the particle, which means that we need to compute the right-hand side of equation (23), or equivalently, the terms (24), (25), for our spacetime metric (27). The expressions that we obtain are

$$f_i^g = \frac{2M}{r^2} u^i \left\{ - \frac{M}{r} + \left( 1 + \frac{2M}{r} \right) \frac{x_j u^j}{r} + \delta_{jk} v^j v^k - \left( 2 + \frac{M}{r} \right) \frac{(x_j v^j)^2}{r^2} \right\} - \frac{M}{r^3} \left\{ 1 - \frac{2M}{r^2} + \frac{4M}{r^2} x_j u^j + 2\delta_{jk} v^j v^k - \left( 3 + \frac{2M}{r} \right) \frac{(x_j v^j)^2}{r^2} \right\},$$  \hfill (39)

$$f_i^\Phi = \left[ 1 - \delta_{jk} v^j v^k - \frac{2M}{r} \left( 1 - \frac{x_j v^j}{r} \right)^2 \right] \times \left\{ \frac{2M}{r^3} x^i - \left( 1 + \frac{2M}{r} \right) u^i \right\} \frac{\Pi}{\sqrt{1 + \frac{2M}{r}}} + \frac{2M}{1 + \frac{2M}{r}} \frac{x^i x^j}{r} \partial_j \Phi - \partial_i \Phi \right\},$$  \hfill (40)

where we have used that

$$(u')^2 = \left[ 1 - v^i v_i - \frac{2M}{r} \left( 1 - \frac{x_i v^i}{r} \right)^2 \right]^{-1}.$$  \hfill (41)

It is important to remark that expression (39) has the same form as in the 3+1 case, confirming the fact that the geodesics of our 2+1 spacetime coincide with the equatorial geodesics ($z = 0$) of the Schwarzschild spacetime.
2.4. Toy model set-up

To summarize, we have to solve a set of six equations. Two of them are PDEs, equations (14), (15), and contain a nonlinear term, the one that introduces the coupling between the gravitational scalar field and the matter sources. The other four equations are ODEs, equations (22), (23), describing the trajectory followed by the particle of mass $m$. The two sets of equations are coupled.

To integrate these equations we will consider spacetime domains of the type, $[t_0, t_f] \times \Omega$, where $t_0$ and $t_f$ are the initial and final integration times, and $\Omega$ is the spatial domain (see figure 1), the domain at every $\{t = \text{const}\}$ slice, consisting of two circular boundaries: an outer boundary $\partial\Omega_{\text{out}}$ at $r = r_{\text{out}} > 50M$ and an inner boundary $\partial\Omega_{\text{in}}$ at $r = r_{\text{in}} < r_h$ where $r_h = 2M$ is the horizon radius. The aim is to locate $r_{\text{out}}$ much close to the radiation zone so we can impose standard outgoing radiation boundary conditions. With regard to the inner boundary, the idea is to excise the black hole singularity from the computational domain, as is done in many fully numerical relativity calculations of black hole dynamics [43, 44], without affecting the computation of the field. This can be done as long as we have $r_{\text{in}} < r_h$ since the characteristics of the PDEs (14), (15), for $r < r_h$, all point in the direction of $r = 0$, which entitles us to perform the singularity excision. Moreover, because of this property of the characteristics, we do not need to impose any kind of boundary conditions at the inner boundary. On the outer boundary ($r = r_{\text{out}}$) we use the two-dimensional Sommerfeld boundary condition (see, e.g., [45, 46]):

$$\left. \left( \frac{\partial_t + \frac{x^i}{r} \partial_i + \frac{1}{2r} \right) \Phi \right|_{r=r_{\text{out}}} = 0. \quad (42)$$

We can rewrite this condition, using polar coordinates, as $(\partial_t + \partial_r)(\sqrt{r} \Phi)|_{r=r_{\text{out}}} = 0$. However, in contrast to what happens in the three-dimensional (3D) case, where $F(t - r)/r$ is an exact solution coinciding with the radiative behaviour of the field (at large $r$), in the two-dimensional case $F(t - r)/\sqrt{r}$ is not a solution of the equations, and therefore the boundary condition (42) does not capture correctly the radiative behaviour of the model. In this sense, this boundary condition could be improved along the lines shown in [45, 46], by
considering higher order derivative boundary conditions, or along the works [47–50] where exact radiative boundary conditions are explored.

The initial data consist of two functions on \( \mathbb{R}^2 \), \((\Phi_0, \partial_t \Phi_0)\) or equivalently \((\Phi_0, \Pi_0)\), the initial position of the particle \( z'_0 = z'(t_0) \) and its initial velocity \( v'_0 = v'(t_0) \). For the scalar gravitational field we will use the simplest initial data, which consist in setting the initial value of \( \Phi \) and its time derivative equal to zero, \( \Phi_0(x, y) \equiv \Phi(t = t_0, x, y), \Phi_0(x, y) \equiv \partial_t \Phi(t, x, y)|_{t=t_0} \):

\[
\Phi_0(x, y) = 0, \quad \text{and} \quad \Phi_0(x, y) = 0.
\]

In terms of \((\Phi, \Pi)\) this translates to

\[
\Phi_0(x, y) = 0, \quad \text{and} \quad \Pi_0(x, y) = -\frac{\beta^i}{\alpha} \partial_i \Phi_0(x, y).
\]

From a physical point of view, these data correspond to a situation in which the particle comes into existence at the initial time \( t_0 \) and through the source term in equation (15) induces a non-zero scalar gravitational field. A consequence of this initial set-up is the triggering of an spurious burst of radiation.

The particle’s initial data, \((z'_0, v'_0)\), are chosen in such a way that they coincide with initial data that, in the absence of a scalar gravitational field, would correspond to circular geodesics of our background spacetime (27). Without loss of generality, we assume that the particle is located initially at

\[
z'_0 = (x_0, 0), \quad \text{and} \quad v'_0 = \left(0, \sqrt{\frac{M}{x_0}}\right).
\]

By using our knowledge of the geodesics (we recall that the geodesics of (27) coincide with the equatorial geodesics of the 3 + 1 Schwarzschild spacetime), we can have an estimation of the time that the particle takes to complete a circular orbit, which we call \( T \). This time can be obtained by using Kepler’s third law,

\[
T^2 = \frac{4\pi^2}{M} R^3,
\]

where \( R \) is the radius of the circular orbit. It is important to remark that this expression is exact when we neglect the effect of the scalar gravitational field \( \Phi \), otherwise it just gives an estimation for the orbital period. For an orbital radius \( r = 10M \) we have \( T \approx 198.7M \).

2.5. Energy-balance test

An important feature of the toy model is the existence of a local conservation law (9) that involves both the energy–momentum of the particle and of the dynamical scalar gravitational field. This local conservation law together with the symmetries of the spacetime leads to global conservation laws. We can use these global conservation laws as a test for the numerical simulations of our toy model.

Let us derive the conservation law associated with the timelike Killing vector field that makes our spacetime static, \( \xi = \partial_t \). Contracting equation (9) with \( \xi^\mu \) and using the Killing equations,

\[
\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0, \quad \text{we obtain}
\]

\[
\xi_\mu \nabla_\nu T^{\mu \nu} = 0 \Rightarrow \nabla_\nu (T^{\mu \nu} \xi_\mu) - T^{\mu \nu} \nabla_\mu \xi_\mu = 0 \Rightarrow \nabla_\nu (T^{\mu \nu} \xi_\mu) = 0.
\]

(48)
We will now integrate this relation over a compact region of the spacetime, $\mathcal{V}$, and use the Gauss theorem to convert volume integrals into surface integrals,

$$\int_{\mathcal{V}} \nabla_\nu (\xi_\mu T^\mu_\nu) \, d\mathcal{V} = 0 \quad \Rightarrow \quad \int_{\partial \mathcal{V}} T^\mu_\nu \xi_\mu \, d\Sigma_\nu = 0,$$

where $\partial \mathcal{V}$ is the boundary of $\mathcal{V}$, which is a closed hypersurface, and $d\Sigma_\mu$ the volume 1-form associated with $\partial \mathcal{V}$. In this way, we have obtained an energy–momentum conservation law which tells us that the flux of the vector $T^\mu_\nu \xi_\nu$ across the boundary of the region $\mathcal{V}$ must vanish.

The region $\mathcal{V}$ is chosen as follows (see figure 2). It consists of two cylinders, $C_1$ and $C_2$, that extend along the timelike Killing direction, and two slices, $S_1$ and $S_2$, orthogonal to the timelike Killing, such that they cut the cylinders forming a closed $2 + 1$ spacetime region. Then, the spacetime region $\mathcal{V}$ can be described as

$$\mathcal{V} \equiv \{(x^\mu) = (t, x, y) \mid t_f > t > t_i, r_2 > r > r_1, 2\pi > \theta > 0\},$$

where the angle $\theta$ is defined by $\tan \theta = y/x$; $t_i$ and $t_f$ can be chosen as the initial and final integration times respectively; $r_2$ can be taken to coincide with the outer boundary and $r_1$ can be chosen so that $r_1 > 2M$, and hence the black hole region is excluded and $\xi$ is timelike everywhere in $\mathcal{V}$. The boundary is then given by

$$\partial \mathcal{V} = C_1 \cup C_2 \cup S_1 \cup S_2,$$

where

$$C_1 = \{(x^\mu) \mid t_f > t > t_i, r = r_1, 2\pi > \theta > 0\},$$

$$C_2 = \{(x^\mu) \mid t_f > t > t_i, r = r_2, 2\pi > \theta > 0\},$$

and $S_1$ and $S_2$ are the slicing surfaces.

Figure 2. Spacetime region $\mathcal{V}$ (see equation (50)) where the conservation law (49) is tested.
\[ S_1 = \{ (x^\mu) \mid t = t_i, r_2 > r > r_1, 2\pi > \theta > 0 \}, \]
\[ S_2 = \{ (x^\mu) \mid t = t_f, r_2 > r > r_1, 2\pi > \theta > 0 \}. \]

(53)

When we apply the conservation law (49) to our domain, we obtain the following relation between boundary integrals:
\[
\int_{C_1} T^{\mu \nu} \xi_\mu \, d\Sigma_\nu + \int_{C_2} T^{\mu \nu} \xi_\mu \, d\Sigma_\nu + \int_{S_1} T^{\mu \nu} \xi_\mu \, d\Sigma_\nu + \int_{S_2} T^{\mu \nu} \xi_\mu \, d\Sigma_\nu = 0. \]

(54)

This relation can be interpreted by stating that the difference in energy between two given instances of time, \( t_i \) and \( t_f \) (which correspond to the integrals on the spacelike slices \( S_1 \) and \( S_2 \)), is due to the loss of energy through the cylinders \( C_1 \) (gravitational waves being absorbed by the horizon) and \( C_2 \) (gravitational waves escaping to infinity).

In order to evaluate these surface integrals we need first to find the normal vector everywhere on the boundary \( \partial V \), that is, on each of the disjoint pieces of (51). The pieces \( S_1 \) and \( S_2 \) (see equation (53)) have constant time \( t \), therefore the (timelike) normals there are given by
\[
n_{|S_1} = \left. -\frac{dt}{\sqrt{1+2M/r}} \right|_{t=t_i}, \quad n_{|S_2} = \left. \frac{dt}{\sqrt{1+2M/r}} \right|_{t=t_f},
\]
(55)

whereas the pieces \( C_1 \) and \( C_2 \) (see equation (52)) are cylinders of constant radius \( r \). Here, it is important to remark that both of the cylinders are assumed to be located at \( r > r_h = 2M \), so that they are timelike hypersurfaces. In practice, we will take \( r_1 \) very close to \( r_h \). Taking this into account we can write the (spacelike) normals as follows:
\[
n_{|C_1} = -\frac{dx^i}{r_1 \sqrt{1-2M/r_1}} = -\frac{dr}{r_1 \sqrt{1-2M/r_1}}, \quad \quad n_{|C_2} = \frac{dx^i}{r_2 \sqrt{1-2M/r_2}} = \frac{dr}{r_2 \sqrt{1-2M/r_2}}.
\]
(56)

Then, the contributions of the gravitational scalar field energy–momentum to the surface integrals are
\[
\int_{S_1} (\Phi) T^{\mu \nu} \xi_\mu \, d\Sigma_\nu = \frac{1}{8\pi} \int \int_{S_1} \, dx \, dy \left\{ \Pi^2 + (\partial_\xi \Phi)^2 + (\partial_r \Phi)^2 \right\} - \frac{2M/r}{1+2M/r} \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right)^2 - \frac{4M/r}{\sqrt{1+2M/r}} \Pi \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right),
\]
(57)

\[
\int_{S_2} (\Phi) T^{\mu \nu} \xi_\mu \, d\Sigma_\nu = -\frac{1}{8\pi} \int \int_{S_2} \, dx \, dy \left\{ \Pi^2 + (\partial_\xi \Phi)^2 + (\partial_r \Phi)^2 \right\} - \frac{2M/r}{1+2M/r} \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right)^2 - \frac{4M/r}{\sqrt{1+2M/r}} \Pi \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right),
\]
(58)

\[
\int_{C_1} (\Phi) T^{\mu \nu} \xi_\mu \, d\Sigma_\nu = -\frac{1}{4\pi} \frac{r_1}{1+2M/r_1} \int t_f \int_0^{2\pi} \, d\theta \left[ \Pi - \frac{2M/r_1}{\sqrt{1+2M/r_1}} (\cos \theta \partial_\theta \Phi + \sin \theta \partial_\phi \Phi) \right] \times \left[ -\frac{2M}{r_1} \Pi + \frac{1}{\sqrt{1+2M/r_1}} (\cos \theta \partial_\theta \Phi + \sin \theta \partial_\phi \Phi) \right],
\]
(59)
\[ \int_{C_2} (\Phi \mathcal{T}^{\mu \nu} \xi_\mu \xi_\nu) d\Sigma_v = \frac{1}{4\pi} \frac{r_2}{1 + 2M/r_2} \times \int_{t_f}^{t_i} dr \int_0^{2\pi} d\theta \left[ \Pi - \frac{2M/r_2}{\sqrt{1 + 2M/r_2}} (\cos \theta \partial_x \Phi + \sin \theta \partial_y \Phi) \right] \]

\[ \times \left[ -\frac{2M}{r_2} \Pi + \frac{1}{\sqrt{1 + 2M/r_2}} (\cos \theta \partial_x \Phi + \sin \theta \partial_y \Phi) \right]. \quad (60) \]

The contributions of the particle's energy–momentum to the surface integrals are

\[ \int_{S_1} (\rho \mathcal{T}^{\mu \nu} \xi_\mu \xi_\nu) d\Sigma_v = \int_{S_1} dx \ dy \rho e^\Phi \left[ 1 - \frac{2M}{r} + \frac{2M}{r^2}(x_i u'_i) \right] (u')^2, \quad (61) \]

\[ \int_{S_1} (\rho \mathcal{T}^{\mu \nu} \xi_\mu \xi_\nu) d\Sigma_v = -\int_{S_1} dx \ dy \rho e^\Phi \left[ 1 - \frac{2M}{r} + \frac{2M}{r^2}(x_i u'_i) \right] (u')^2. \quad (62) \]

Since the particle density is very small outside a small neighbourhood around the particle location, the contribution due to the particle to the integrals on the cylinders \( C_1 \) and \( C_2 \) can be neglected,

\[ \int_{C_1} (\rho \mathcal{T}^{\mu \nu} \xi_\mu \xi_\nu) d\Sigma_v \approx 0, \quad \int_{C_2} (\rho \mathcal{T}^{\mu \nu} \xi_\mu \xi_\nu) d\Sigma_v \approx 0, \quad (63) \]

where in (57) and (61), \( \Phi \) and \( \Pi \) mean \( \Phi(t_i, x, y) \) and \( \Pi(t_i, x, y) \) respectively. In (58) and (62), \( \Phi \) and \( \Pi \) mean \( \Phi(t_f, x, y) \) and \( \Pi(t_f, x, y) \) respectively. Both in (59) and in (60) we have used polar coordinates instead of Cartesian ones \((x = r \cos \theta, y = r \sin \theta, dx \ dy = r \ dr \ d\theta)\). Therefore, in (59), \( \Phi \) and \( \Pi \) mean \( \Phi(t, r_1 \cos \theta, r_1 \sin \theta) \) and \( \Pi(t, r_1 \cos \theta, r_1 \sin \theta) \) respectively; and in (60), \( \Phi \) and \( \Pi \) mean \( \Phi(t, r_2 \cos \theta, r_2 \sin \theta) \) and \( \Pi(t, r_2 \cos \theta, r_2 \sin \theta) \) respectively. In (61) and (62), the objects \( \rho \) and \( u' \) must be substituted by their expressions (37) and (41) respectively.

Then, the conservation law that we have to test numerically is

\[ \frac{1}{8\pi} \int_{S_1} dx \ dy \left\{ \Pi^2 + (\partial_x \Phi)^2 + (\partial_y \Phi)^2 - \frac{2M}{r} \frac{\Pi}{1 + 2M/r} \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right)^2 \right. \]

\[ - \left. \frac{4M}{\sqrt{1 + 2M/r}} \Pi \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right) \right\} - \frac{1}{8\pi} \int_{S_1} dx \ dy \]

\[ \times \left\{ \Pi^2 + (\partial_x \Phi)^2 + (\partial_y \Phi)^2 - \frac{2M}{r} \frac{\Pi}{1 + 2M/r} \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right)^2 \right. \]

\[ - \left. \frac{4M}{\sqrt{1 + 2M/r}} \Pi \left( \frac{x}{r} \partial_x \Phi + \frac{y}{r} \partial_y \Phi \right) \right\} - \frac{1}{4\pi} \frac{r_1}{1 + 2M/r_1} \int_0^{r_f} dr \int_0^{2\pi} d\theta \]

\[ \times \left[ \Pi - \frac{2M}{r_1} \frac{\Pi}{\sqrt{1 + 2M/r_1}} (\cos \theta \partial_x \Phi + \sin \theta \partial_y \Phi) \right] \]

\[ \times \left[ -\frac{2M}{r_1} \frac{\Pi}{\sqrt{1 + 2M/r_1}} + \frac{\cos \theta \partial_x \Phi + \sin \theta \partial_y \Phi}{\sqrt{1 + 2M/r_1}} \right] \]

\[ + \frac{1}{4\pi} \frac{r_2}{1 + 2M/r_2} \int_0^{r_f} dr \int_0^{2\pi} d\theta \left[ \Pi - \frac{2M}{r_2} \frac{\Pi}{\sqrt{1 + 2M/r_2}} (\cos \theta \partial_x \Phi + \sin \theta \partial_y \Phi) \right] \]

\[ \times \left[ -\frac{2M}{r_2} \frac{\Pi}{\sqrt{1 + 2M/r_2}} + \frac{\cos \theta \partial_x \Phi + \sin \theta \partial_y \Phi}{\sqrt{1 + 2M/r_2}} \right]. \]
\begin{align}
+ \iint_{S_1} dx \, dy \, e^\Phi \left[ 1 - \frac{2M}{r} + \frac{2M}{r^2} (x_i v_i') \right] (u')^2 \\
- \iint_{S_2} dx \, dy \, e^\Phi \left[ 1 - \frac{2M}{r} + \frac{2M}{r^2} (x_i v_i') \right] (u')^2 = 0. \tag{64}
\end{align}

In order to decide whether the result from the numerical computation is satisfactory, we can normalize the previous expression with respect to, for instance, the initial energy of the particle, i.e., the last but one line in this conservation law.

3. The numerical framework

The goal of this section is twofold. First, we give a brief introduction to the finite element method, a numerical technique that has rarely been used in numerical relativistic calculations. A detailed basic exposition of the FEM can be found in classical textbooks such as [51–55]. Second, we want to describe the adaptive finite element method, where we present a new local mesh refinement technique based on an interpolation error estimate introduced recently by Chen, Sun and Xu [38]. This is the technique that we investigate in this paper as a possible tool for achieving the adaptivity that the simulations of EMRBs require.

3.1. Introduction to the finite element method

The FEM is a numerical technique for solving problems described by PDEs or those that can be formulated as functional minimization problems. In what follows, we briefly introduce the basic ideas and procedures of the FEM by using a simple example involving the following wave equation:

\[(\partial^2_t - \partial^2_x - \partial^2_y) \psi_1(t, x, y) = 2, \quad t \in [0, T], \quad x \in \Omega_1 = \{(x, y) | r = \sqrt{x^2 + y^2} \leq R\}, \tag{65}\]

\[n^i \partial_{n^i} \psi_1 |_{\partial \Omega_1} = 2R, \tag{66}\]

\[\psi_1(t = 0, x, y) = x^2 + y^2, \quad (\partial_t \psi_1)(t = 0, x, y) = 0. \tag{67}\]

The domain \(\Omega_1\) is a disc of radius \(R\). We prescribe the von Neumann boundary condition (66) where \(n^i\) are the components of the normal to the boundary \(\partial \Omega_1\). The initial data are described by (67). This problem has the simple analytic solution:

\[\psi_1(t, x, y) = t^2 + x^2 + y^2. \tag{68}\]

We start from the discretization of the computational domain \(\Omega_1\) into an assembly of disjoint element domains \(\Omega_\alpha\), that is,

\[\Omega = \bigcup_{\alpha} \Omega_\alpha, \quad \Omega_\beta \cap \Omega_\gamma = \emptyset \quad \text{for} \quad \beta \neq \gamma. \tag{69}\]

In two dimensions, the element domains are typically triangles and quadrilaterals. In this work we use only triangles. In practice, mesh generation is carried out by using the software included in the general-purpose software package FEPG [56]. A typical mesh (domain discretization) for our example is given in figure 3. Every element is equipped with a finite-dimensional functional space \(\mathcal{F}_\alpha\), so that we approximate our physical solution locally, at every element, as a linear combination of functions of \(\mathcal{F}_\alpha\) (usually, a special treatment of the boundary elements is required). It is very common that the functional spaces \(\mathcal{F}_\alpha\) are formed by piecewise polynomials. In this paper, we consider linear elements, where the element functions are first-order polynomials, i.e., \(a + bx + cy\). This choice of element functions implies second-order convergence to the solution in the \(L^2\) norm.
Figure 3. Mesh corresponding to the domain $\Omega$ in equation (65).

A very important ingredient of the FEM is that it works with an integral form of the equation we want to solve, which is called the weak form of the problem. To obtain it, we multiply equation (65) by a test function $\phi$, integrate over the domain $\Omega$, and use the Gauss or divergence theorem which introduces the boundary conditions (66). The result can be written in the form

$$L[\phi, \psi] \equiv (\phi, \partial^2_t \psi) + (\nabla \phi, \nabla \psi) - 2 \int_{\partial \Omega} r \phi \, ds + 2(\phi, 1) = 0,$$

where $(\cdot, \cdot)$ denotes the inner product $(u, v) = \int_{\Omega} uv \, dx \, dy$, $\nabla$ denotes the gradient operator, and $s$ is a coordinate on $\partial \Omega$. It is important to note that the third term is the result of introducing into the weak formulation the boundary conditions (66).

Using the element functional spaces we can expand our solution in terms of nodal basis functions, $n_A(x, y)$ ($A = 1, \ldots, N$, with $N$ being the number of nodes), which are associated with the nodes or grid points of the mesh. Nodal functions take the value unity at the node with which they are associated and zero at all the other nodes ($n_A(x_B) = \delta_{AB}$ for any node $x_B$). Since we will produce a FEM discretization only in space, the expansion of our solution in terms of the nodal functions can be written in the form

$$\psi_h \in H^2(\Omega), \quad \psi_h(t, x, y) = \sum_B \psi_B(t)n_B(x, y),$$

where $H^2(\Omega)$ is the Sobolev space of functions on $\Omega$ that are, together with their first and second generalized spatial derivatives, square integrable, that is, they belong to $L^2(\Omega)$. We assume that $\psi_h$ belongs to $H^2(\Omega)$ for any time $t \in [0, T]$. The subscript $h$ denotes a scale associated with the domain discretization, for instance, it may be proportional to the square root of the average area of the elements that compose the mesh. In our example $h$ refers to the maximum mesh diameter in the whole domain. An obvious property of $h$ is that the size of the elements goes to zero as $h$ goes to zero.

In a time-dependent problem like the one we are considering, the unknowns are the functions $\psi_A(t)$. The equations for these functions are obtained from the spatial discretization.
In a Galerkin-type formulation of the FEM, the discretized equations come from the imposition of the vanishing of the residuals,

\[ E_A = L[n_A, \psi_B] = 0, \]

(72)

which consists in taking \( \phi = n_A \) in (70). Introducing also the expansion (71) yields the following system of equations for the functions \( \psi_A(t) \),

\[ \sum_B M_{AB} \ddot{\psi}_B + \sum_B K_{AB} \psi_B = F_A, \]

(73)

where the matrix \( M_{AB} = (n_A, n_B) \), the so-called mass matrix, and the matrix \( K_{AB} = (\nabla n_A, \nabla n_B) \), the so-called stiffness matrix, are symmetric and positive-definite matrices. The vector \( F_A = -2(n_A, 1) + \int_{\Omega} n_A \, ds \) is sometimes called the force vector. Equation (73) is the outcome of the FEM spatial discretization, which is sometimes called the semi-discrete form because it consist of a linear system of second-order ordinary differential equations in time. They are usually solved by using finite difference methods. One of the most popular methods for second order in time equations is the Newmark method, which is second-order accurate in time (see, e.g., [54], for details).

Before we discuss the numerical implementation it is worth mentioning two important features of the FEM: (i) the piecewise approximations (piecewise linear in our example) of physical fields on finite elements provide good precision even with simple approximating functions, increasing the number of elements we can achieve the desired precision; (ii) the local character of the approximation leads to sparse systems of equations once the problem is discretized. This helps considerably in solving problems with a very large number of nodal unknowns.

The numerical implementation of the equations of this paper has been carried out by using the general-purpose software package FEPG [56], which can automatically generate finite element Fortran source code based on component programming. It can handle many types of problems, including time-dependent nonlinear ones, like the one we are interested in. As a test, we have implemented the example described in this section in FEPG (see a snapshot of the evolution in figure 4). We have also studied the convergence properties of the solution. Here, it is important to point out that in an unstructured mesh, like the one we are using in this example, to perform a convergence test is not as straightforward as it is for structured meshes. We have to define the scale \( h \) properly such that by changing it we obtain the correct
convergence. Starting from the initial mesh (see figure 3), we call the solution we obtain \( \psi_h \). Then, we globally refine this initial mesh by transforming every initial triangular element into four smaller triangular elements by connecting the three mid points of each edge. Solving our example equation on this mesh leads to a more accurate solution that we call \( \psi_{h/2} \), and whose associated scale is \( h/2 \). By repeating this refinement process we get finer meshes, and by solving our equation on them we obtain more accurate solutions, \( \psi_{h/2^k} \), with associated scale \( h/2^k \). We have checked that the solution we obtain converges quadratically in the scale \( h \) to the exact solution (68) by studying the norms \( \| \psi_{h/2^k} - \psi_{\text{exact}} \|_{L^2} \) (see the left of figure 5).

We have also checked that it converges quadratically in the usual way, without making use of the exact solution, just by comparing the norms \( \| \psi_h - \psi_{h/2} \|_{L^2} \) and \( \| \psi_{h/2} - \psi_{h/4} \|_{L^2} \) (see the right of figure 5).

### 3.2. Finite element discretizations of the toy model

In this section, we develop a FEM formulation for our toy model in the spirit of the ideas presented above. The ingredients of this problem are the following. (i) The computational domain was described in section 2.4 and shown in figure 1. (ii) The equations that describe our model are the PDEs given by equations (14), (15). (iii) The boundary conditions and initial data are described in section 2.4.

We start from the discretization of the computational domain (see figure 1). We use linear triangular elements. The aspect of the resulting triangularization is shown in figure 6. It is worth pointing out that the FEM is specially well suited for complex domains. In our case, this allows us to use circular boundaries, which adapt better to the characteristics of our problem. This is specially important in the case of the inner boundary, which corresponds to the fact that we have excised the black hole singularity from the computational domain.

The next step in our development is the finite element discretization of equations (14), (15). We need to construct a weak formulation of these equations. To that end, it is very convenient to rewrite the equations in the following equivalent form,

\[
\alpha^{-1} \partial_t \Phi = \alpha^{-1} \beta^i \partial_i \Phi + \Pi, \tag{74}
\]

\[
\alpha^{-2} \partial_t \Pi = \alpha^{-2} \beta^i \partial_i \Pi + \partial_i (\sqrt{h} h^{ij} \partial_j \Phi) + \alpha^{-1} \alpha^i \partial_i \Phi + \alpha^{-1} K \Pi - 4\pi \alpha^{-1} e^\Phi \rho, \tag{75}
\]
where we have used the fact that $\alpha \sqrt{h} = 1$. The main reason for casting the equations in this form is that it brings the second-derivative terms into the form of the divergence of a spatial vector, without any additional factors. As we did with the example of the previous subsection, we will discretize the system of PDEs (74), (75) for the unknowns $\Phi$ and $\Pi$ by using a FEM discretization (a Galerkin-type formulation) for the spatial dimensions and by using finite difference methods in time. In this sense, it is important to take into account that our system of equations is of first order with respect to time derivatives and of second order with respect to spatial derivatives. In order to discretize these equations, particular attention has to be paid to the convection terms in order to keep them under control. We deal with this issue by including additional artificial viscosity into the finite element equations that we obtain.

For the FEM spatial discretization the finite element space that we use is $\mathcal{S}_h \subset H^2(\Omega)$, which consists of piecewise triangular linear interpolation functions. Then, taking all this into account and operating in a similar way as we did in the example of section 3.1, the discretized problem that we obtain can be introduced in the following way. We need to find $(\Phi^{n+1}, \Pi^{n+1}) \in \mathcal{S}_h \subset H^2(\Omega)$, such that the equations

\[
(\alpha^{-1} \Phi^{n+1}, \phi) = (\alpha^{-1} \Phi^n, \phi) + \left( \frac{\Delta t}{\alpha} \frac{\partial^i}{\partial x} \Phi^n, \phi \right) - \delta(h) \left( \frac{\Delta t}{\alpha} \frac{\partial^i}{\partial x} \Phi^n, \frac{\partial^j}{\partial x} \Phi^n \right) + \frac{\Delta t}{\alpha} \left( \alpha \Pi^n, \phi \right),
\]

(\alpha^{-2} \Pi^{n+1}, \phi) = (\alpha^{-2} \Pi^n, \phi) + \left( \frac{\Delta t}{\alpha} \frac{\partial^i}{\partial x} \Pi^n, \phi \right) - \delta(h) \left( \frac{\Delta t}{\alpha} \frac{\partial^i}{\partial x} \Pi^n, \frac{\partial^j}{\partial x} \Pi^n \right) - \frac{\Delta t}{\alpha} \left( \sqrt{h} \frac{\partial^i}{\partial x} \Phi^{n+1}, \partial_j \phi \right) + \frac{\Delta t}{\alpha} \left( \alpha \Pi^n, \phi \right)
\]

(76)
\[ -\Delta t \left( 4\pi \alpha^{-1} e^{\Phi_1^\alpha n} \rho, \psi \right) - \int_{\partial\Omega_{\text{out}}} \frac{\alpha^{-1}}{1 + 2M r} \left( \Phi_1^{n+1} - \Phi_1^n + \frac{\Phi_1^{n+1} - \Phi_1^n}{2r} \Delta t \right) \psi \, ds \\
- \Delta t \int_{\partial\Omega_{\text{in}}} \frac{\alpha^{-1}}{1 + 2M r} \frac{\partial_i \Phi_1^{n+1}}{r} \psi \, ds, \]  
(77)

hold for any \((\phi, \psi) \in S_h\) and for \(n = 0, 1, \ldots, N, N \Delta t = T\), where \(T = t_f - t_0\) is the total computational time. \(\Delta t\) is the time-step size we use to evolve \(\Phi_1^n\) and \(\Pi_1^n\). We use the notation \((\cdot, \cdot)\) to denote the inner product in \(S_h\), which is defined as \((u, v) = \int_{\Omega} uv \, d\Omega\) \((\forall u, v \in S_h)\).

Finally, to deal with the convection terms, arising from the derivatives of the fields along the shift vector, we use a streamline diffusion scheme that includes artificial viscosity and which is specially adapted to the FEM. The factor \(\delta(h)\) is the penalty ratio of artificial viscosity, which of course depends on the mesh size \(h\). It has to be tuned properly in order to obtain a stable computation and at the same time an accurate solution.

The last terms of (77) consist of integrals on the boundaries \(\partial\Omega_{\text{in}}\) and \(\partial\Omega_{\text{out}}\). The way they come into play is as follows. In the case of the outer boundary \(\partial\Omega_{\text{out}}\), we obtain the last but one integral in (77) after integration by parts of the term with second-order spatial derivatives in (75) and imposition of the outgoing boundary condition (42). In the case of the inner boundary \(\partial\Omega_{\text{in}}\), we obtain the last integral in (77) from the same integration by parts. However, the resulting integral does not correspond to a proper boundary condition. As we have already mentioned before, we do not need any boundary condition at the inner boundary due to the particular structure of the characteristics there (remember that the inner boundary is inside the horizon), which all point inwards. Therefore, the resulting integral is just an integration of a term proportional to \(\partial_\nu \Phi_1^{n+1}\) on the inner boundary.

Equations (76), (77) are the basis of the computational procedure that we follow to obtain the solution of our problem. The basic algorithm consists in computing first \(\Phi_1^{n+1}\) from (76) in terms of \(\Phi_1^n\) and \(\Pi_1^n\). The next step is to introduce \(\Pi_1^{n+1}\) into the right-hand side of (77) to obtain \(\Pi_n^{n+1}\) in terms of \(\Phi_1^n\) and \(\Pi_n^n\) together. Then, \(\Pi_n^{n+1}\) has to be used in order to compute the value of \(\Phi_1\) in the next time step, and so on. It is not difficult to show that this algorithm is second-order accurate both in space and in time.

### 3.3. Computing the motion of the particle

In the procedure we have just described we have omitted the role of the motion of the particle. It is clear that in order to evaluate the right-hand side of (77) we need to introduce the position of the particle. And for that, we need to solve the set of equations (22), (23). These ODEs have to be integrated simultaneously with the PDEs, which means that every time we evaluate the right-hand side of (77) we need to evolve (22), (23) a time step \(\Delta t\), and use the result (new position and velocity of the particle) in order to evaluate the source of equation (15) (equation (77) in the discretized system).

The type of numerical algorithm we use to solve the ODEs (22), (23) has to take into account the particular structure of equations (23), which are the non-trivial ones. These equations contain two differentiated terms, \(f_\nu^g\) and \(f_\nu^\Phi\). The first term would give us the geodesic motion around a Schwarzschild black hole, and therefore the time scale of the changes induced by the term \(f_\nu^g\) is the orbital period of the geodesic that the particle would follow by ignoring the radiation reaction effects. The second term contains the gradients of the scalar field \(\Phi\) in the neighbourhood of the particle position. These terms are responsible, in our toy model, for the radiation reaction effects, and therefore, the time scale of the changes they induce will be, in general, much smaller than the orbital time scale.
Taking into account that these ODEs are nonlinear both in $z_i$ and in $v_i$, we will just use an explicit scheme with a time step $\Delta t_1$ much smaller than the PDE time step $\Delta t$. That is, we split each PDE time step $\Delta t$ in many ODE time steps $\Delta t_1$, so that we solve equations (22), (23) in each time step $\Delta t_1$ in order to guarantee the accuracy of the solution. This scheme is able to approximate the accuracy that an implicit scheme would have provided. The specific discretization algorithm we use in our numerical computations to evolve the particle’s position, $z$, and velocity, $v$, from a PDE time step $t_n$, with values $(z^n, v^n)$, to the next time step $t_{n+1}$, with values $(z^{n+1}, v^{n+1})$, is given by

$$
\begin{align*}
  v^{n+1}_0 &= v^n, \\
  z^{n+1}_0 &= z^n, \\
  \text{From } m = 0 \text{ to } m = M - 1: & \quad \begin{align*}
    v^{n+1}_m &= v^n + \Delta t_1 \left[ f_g \left( v^{n+1}_m, z^{n+1}_m \right) + f_\Phi \left( v^{n+1}_m, z^{n+1}_m, \Pi^n, \nabla \Phi^v \right) \right], \\
    z^{n+1}_m &= z^n + \Delta t_1 v^{n+1}_m, \\
    v^{n+1} &\equiv v^{n+1}_M, \\
    z^{n+1} &\equiv z^{n+1}_M, 
  \end{align*}
\end{align*}
$$

(78)

where $\Delta t_1$ denotes the ODE time step, which is related to the PDE time step by the relation $\Delta t_1 = \Delta t / M$. We choose the integer $M$ in order to achieve the accuracy we want. The way in which $\nabla \Phi$ is computed is by using the corresponding FEM piecewise polynomial expansion that follows from the expansion in nodal functions of $\Phi$.

### 3.4. The adaptive finite element method

The AFEM is the application of the classical FEM to a series of local adaptive meshes in order to get more accurate numerical results with less computational cost. Starting with a given initial coarse mesh, the adaptive mesh on each level is generated locally and adaptively in terms of a posteriori error estimate of the finite element solution. Usually, local refinement takes place where the a posteriori errors are much bigger than elsewhere, or in other words, where the finite element solution changes steeply. On the other hand, for those places in which the a posteriori errors are sufficiently small, the local derefinement will operate in order to eliminate extra grids because in those places the solution changes slowly. Thus, during the finite element computation we can adaptively adjust the mesh density without losing the numerical accuracy but reducing considerably the computational cost.

It is obvious that the key part of the AFEM is the a posteriori error estimate. To have such a good estimate means that one can precisely find out in which places the mesh should be refined or derefined without introducing much numerical pollution. After this is done, the rest of the procedure is standard; for instance, the mesh bisection and the finite element approximation. There are presently a number of works in the literature on these issues (see, e.g., [57–60]). In what follows, we introduce our own a posteriori error estimate, which is an interpolation error estimate. Then, we elaborate on our local mesh improvement techniques such as refinement, coarsening and the smoothing strategy, which aim to minimize the interpolation error.

The interpolation error estimate comes from recent work by Chen, Sun and Xu [38]. This estimate can be seen as the theoretical foundation of our adaptive mesh techniques; that is, our algorithms are aimed at minimizing (or at least reducing) the interpolation error by iteratively modifying the grids. We introduce the estimate through the following definition.
Definition. Let $\Omega$ be a bounded domain in $\mathbb{R}^n$. Given a function $u \in C^2(\Omega)$, we say that a symmetric positive-definite matrix $H \in \mathbb{R}^{n \times n}$ is a majoring Hessian of $u$ if
\[
|\xi^t (\nabla^2 u)(x) \xi| \leq c_0 |\xi^t H(x) \xi|,
\]
for some positive constant $c_0$.

Here, $\xi^t$ denotes the transpose of the vector $\xi$. We then use the majoring Hessian to define a new metric
\[
H_p = (\det H)^{-\frac{1}{2p}} H, \quad (p \geq 1).
\]

There are two conditions for a triangulation $T_N$, where $N$ is the number of simplices (generic elements), to be a nearly optimal mesh in the sense of minimizing the interpolation error in the $L^p$ norm. The first condition consists in asking the mesh to capture the high oscillations of the Hessian metric; that is, $H$ should not change very much on each element. This condition can be expressed in a more precise way by means of the following statement.

There exist two positive constants $\alpha_0$ and $\alpha_1$ such that
\[
\alpha_0 |\xi^t H_t \xi| \leq |\xi^t H(x) \xi| \leq \alpha_1 |\xi^t H_t \xi|, \quad (\xi \in \mathbb{R}^n, x \in \Omega),
\]
where $H_t$ denotes the average of $H$ over $\tau \in T_N$.

The second condition demands the triangulation $T_N$ to be quasi-uniform under the new metric induced by $H_p$. This condition can also be expressed in a more precise way through the following statement.

There exist two positive constants, $\beta_0$ and $\beta_1$, such that
\[
\sum_J \tilde{d}_{\tau,J}^2 \leq \beta_0 (\forall \tau \in T_N) \quad \text{and} \quad \max_{\tau \in T} |\tilde{\tau}| \leq \beta_1,
\]
where $|\tilde{\tau}|$ denotes the volume of $\tau$, and $\tilde{d}_{\tau,J}$ the length of the $J$th edge of $\tau$ under the new metric $H_p$.

The first inequality in (82) means that each $\tau$ is shape-regular under the metric $H_p$. The second inequality means that all elements $\tau$ are of comparable size (also under the new metric), which is a global condition. This means that the mesh is more dense in the regions where $\det H_p(x)$ is larger. In [38], an important result has been proven that a triangulation which satisfies both the local condition (81) and the global one (82) yields a good approximation. This result can be expressed in a precise way in the form of a theorem:

Theorem. Let $u$ be a function belonging to $C^2(\Omega)$, $T_N$ a triangulation satisfying the conditions (81) and (82), and $u_T$ the linear finite element interpolation of $u$ based on the triangulation $T_N$. Then, the following error estimate holds,
\[
\|u - u_T\|_{L^p(\Omega)} \leq C N^{-\frac{1}{2}} \|\sqrt{\det(H)}\|_{L^{\frac{2p}{p-1}}(\Omega)},
\]
for some constant $C = C(n, p, c_0, \alpha_0, \alpha_1, \beta_0, \beta_1)$. This error estimate is optimal in the sense that for a strictly convex (or concave) function, the above inequality holds in a reversed direction.

The result expressed by this theorem is the basis of the grid adaptation algorithms that we use in this work. Roughly speaking, for a given function $u$, we will adapt our grids in such
a way that the conditions given in (81) and (82) are satisfied in the best possible way. One important remark we need to make is that the validity of the theorem stated above allows for a few exceptions of the condition (82) for \( p < \infty \) (see [38] for details). This is of particular importance since in practice it is very difficult to guarantee that the condition (82) is satisfied everywhere.

The next point is to see how the Hessian matrix of the solution can be obtained when the linear finite element approximation is used for the discretization of the PDEs of our problem. Since taking piecewise second derivatives of piecewise linear functions will give no approximation to the Hessian matrix, special post-processing techniques need to be used in order to obtain a reasonable Hessian matrix approximation from linear finite elements. One of the most popular techniques is the patch recovery technique proposed by Zienkiewicz and Zhu (ZZ) [61, 62], which is based on the least-squares fitting on local patches. Results from the application of this technique demonstrate that it is robust and efficient. The theoretical reason why the ZZ method works is largely understood to be related to the superconvergence phenomenon for second-order elliptic boundary-value problems discretized on a finite element grid that has certain local symmetry (see the works of Walhbin [63] and Babuška and Strouboulis [55]). These classic superconvergence results can be used in order to justify the effectiveness of the ZZ method. A significant improvement of this type of analysis was introduced recently by Bank and Xu [64, 65]. In [64] they give superconvergence estimates for piecewise linear finite element approximations on quasi-uniform triangular meshes where most pairs of triangles that share a common edge form approximate parallelograms. In [64] they also analyse a post-processing gradient recovery scheme, showing that \( Q_h \nabla u_h \), where \( Q_h \) is the global \( L^2 \) projection, is a superconvergent approximation to \( \nabla u \). This result leads to a theoretical justification of the ZZ method for such types of grids (see [66] for details).

The gradient recovery algorithm that we use in the numerical examples of this paper is based on a new approach due to Bank and Xu [65], where they use the smoothing iteration of the multigrid method to develop a post-processing gradient recovery scheme. This scheme proves to be very efficient for recovering Hessian matrices. All the methods mentioned above can be extended to anisotropic grids with some appropriate modifications, but a theoretical justification of such extensions is still lacking. Nevertheless, numerical experiments have given satisfactory results.

Let us now discuss techniques to improve the mesh quality. We define the mesh quality of a given triangulation \( \mathcal{T} \) in terms of the interpolation error:

\[
Q(\mathcal{T}, u, p) = \|u - u_{\mathcal{T}}\|_{L^p(\Omega)}, \quad (1 \leq p < \infty).
\]  

There are three main ways of improving a mesh: (i) refinement or coarsening through splitting or merging of edges [67–69]; (ii) edge swapping by replacing sets of elements by other such sets while preserving the position of the points (nodes) [70]; and (iii) mesh smoothing, which moves the vertices of the mesh while keeping the connectivity [71–74]. We derive those techniques by minimizing the interpolation error in the \( L^p \) norm.

For the first method, this can be done by equidistributing the edge lengths with respect to the new metric. Thus, we compute the edge lengths with respect to the new metric \( H_p \) and mark edges whose length is greater than \( r_1 d \), where \( r_1 \geq 1 \) is a parameter and \( d \) is a fixed edge length, the global average edge length. Then, we connect marked edges element-wise according to the different situations that can be given. This is illustrated in figure 7. The coarsening operates like an inverse procedure to the refinement process. It marks the edges whose length is smaller than \( r_2 d \), where \( r_2 \leq 1 \) is another parameter. We then shrink this edge to a point and connect to the vertices of the patch of the edge.
We consider now the case of edge swapping involving four points \( \{ x_\alpha \} (\alpha = 1–4) \) constituting two adjacent triangles and a convex quadrilateral. Let \( T_1 = \triangle_{123} \cup \triangle_{134} \) and \( T_2 = \triangle_{124} \cup \triangle_{234} \) be two triangulations, where \( \triangle_{\alpha \beta \gamma} \) stands for the triangle made up of the points \( x_\alpha, x_\beta, \) and \( x_\gamma \). Then, we choose the triangulation \( T_1 \) if and only if \( Q(T_1, u, p) \leq Q(T_2, u, p) \), for some \( 1 \leq p \leq \infty \). In [75], we show that this criterion is equivalent to the empty circle criterion when \( u(x) = \|x\|^2 \). Thus it is an appropriate generation of the edge swapping used in the isotropic case to the anisotropic case.

Finally, local smoothing of the mesh adjusts the location of a vertex in its patch \( \Omega_{\alpha} \), which consists of all simplices containing the vertex \( x_\alpha \), without changing the connectivity. The moving of a vertex to its new location is expected to improve the quality of the mesh. Several sweeps throughout the whole mesh can be performed to improve the overall mesh quality. By minimizing the interpolation error in \( \Omega_{\alpha} \), we move the vertex \( x_\alpha \) to the position \( x^* \) in such a way that

\[
\nabla u(x^*) = -\frac{1}{|\Omega_{\alpha}|} \sum_{\tau_\beta \in \Omega_{\alpha}} \left( \nabla|\tau_\beta| \sum_{x_\gamma \in \tau_\beta, x_\gamma \neq x_\alpha} u(x_\gamma) \right).
\]

The derivation of this formula can be found in [75].

For the application to our numerical computations we use \( Q_h \nabla u_h \) and \( u_h \) in (85).

Moreover, we need to make choices for the different parameters of this adaptive mesh technique. For the order of the \( p \) of the \( L^p \) norm we use \( p = 1 \), the \( L^1 \) norm. According to [75], the \( L^1 \) norm can catch singularities more efficiently than other norms. For the multiple \( r \) in the text of the global average edge length \( d \) under the new metric \( H_p \) (we use \( rd \) as the threshold to determine which edge lengths under the new metric are bigger than it; we will then bisect those edges in the local refinement process) we take \( r = 3 \) initially. Our local refinement procedure is a nested iteration process. There is an outer iteration for which, on each step, we reduce the multiple \( r \) down to \( r_{\text{min}} \) (we take \( r_{\text{min}} = 1 \)). The main purpose of this iteration is to resolve the singularity as precisely as possible and to control the refinement pollution at the lowest level. On each outer iteration step, we perform an inner iteration, where we have another parameter \( r \) that controls this iteration. Usually, we do not need this number to establish a criterion to stop the inner iteration of local refinement. For the sake of getting the optimal mesh, we let this iteration go until all the edge lengths are smaller than \( rd \) as before. In practice, there are occasions when this may take too much time, and in those cases we set up a criterion as before in order to control the number of inner iterations.

4. Numerical results

In this section, we report the results of the simulations of the toy model. We have distinguished two types of simulations. Those that were performed by just using the classical FEM, without
adding any extra adaptivity. The results are discussed in section 4.1; those that were performed by using the AFEM are described in the previous section. The results of these simulations are discussed in section 4.2.

4.1. Numerical simulations of the toy model using the FEM

The mass ratio we have considered is $\mu = 0.01$, that is, the particle’s mass is $m = 0.01M$. The inner boundary is located at $r_{in} = 1M$, inside the horizon $r_h = 2M$. The outer boundary has been located at $r_{out} = 50M$. Regarding the initial data, the only parameters that need to be given, in addition to those already given, in order to completely specify it (see section 2.4) are the width of the Gaussian that we use to regularize the Dirac delta distribution, which we take to be $\sigma = 1M$, and the initial position of the particle, which we take to be $(x_0, y_0) = (10M, 0)$. Finally, we give the resolution that we have used for the simulations. To give a measure of the spatial resolution we describe the structure of the mesh. It is composed of $N = 9266$ triangles, quasi-uniformly distributed, more concentrated near the centre and coarsening gradually as we approach the outer boundary. Regarding the resolution in the time direction, the step size we use for the time evolution is $\Delta t = 0.01M$. The algorithm we use to study the evolution of the gravitational scalar field $\Phi$ and the motion of the particle is completely described by the explicit schemes of equations (76), (77) and (78).

What we have observed is that the orbit of the particle (remember that the initial data correspond to a circular orbit in the absence of the scalar gravitational field) shrinks gradually until the particle reaches the horizon at $r_h = 2M$. The trajectory followed by the particle is drawn in figure 8. Snapshots characterized by the time $t$ and time-step number $n$ of the evolution of the scalar gravitational field are shown by means of the contour plots given in figures 9–13. The position of the particle is evident in these graphs.

We have checked the energy-balance law (64) for the parameters given above. The result can be found in figure 14, where the horizontal axis denotes the time and the vertical axis indicates the value of the left-hand side of (64) in units of $m$. As we can see there, the energy law (64) is satisfied up to a certain level along the evolution. However, from $t = 40M$ the error in the energy-balance test grows and stabilizes around a different but bigger value. This growth is due to outer boundary effects: the particle started from a position $r = 10M$ and the outer
A toy model for testing finite element methods to simulate extreme-mass-ratio binary systems

Figure 9. Scalar gravitational field $\Phi$ at time $t = 10M$ ($n = 1000$) (left) and at time $t = 20M$ ($n = 2000$) (right).

Figure 10. Scalar gravitational field $\Phi$ at time $t = 30M$ ($n = 3000$) (left) and at time $t = 40M$ ($n = 4000$) (right).

Figure 11. Scalar gravitational field $\Phi$ at time $t = 50M$ ($n = 5000$) (left) and at time $t = 100M$ ($n = 10000$) (right).

boundary is located at $r = 50M$. As we have mentioned in the discussion of the boundary conditions in section 2.4, the outer boundary condition (42) is not optimal, and induces some
error in the solution. This could be improved by either moving the outer boundary further out or by using an improved outer boundary condition.

The key point in these computations is the ability of our numerical scheme to resolve the source term describing the particle. It enters the equations for the scalar gravitational field $\Phi$ as a very singular source term. This source term depends on the position of the particle, but at the same time, the position of the particle depends on the gradients of $\Phi$. Therefore, it is very important first to compute properly the effect of the source $\rho$ on the field $\Phi$, and then to compute accurately the gradient of the field itself. We have already mentioned that in the computations shown in this subsection the width of the Gaussian is taken to be $\sigma = 1M$ in order to make the source sufficiently smooth for the resolution we have used. In this sense, if we try to use a smaller width, for instance $\sigma = 0.1M$, the classical FEM will fail to provide the expected accuracy, although it still provides a numerical solution. This is where one realizes the potential of the AFEM as a better choice to carry numerical computation for a model presenting features similar to those of our model. In the next subsection, we describe how we have implemented the AFEM in the case of the toy model and show that it provides a reasonable solution for the case in which $\sigma = 0.1M$. 

Figure 12. Scalar gravitational field $\Phi$ at time $t = 200M$ ($n = 20000$) (left) and at time $t = 300M$ ($n = 30000$) (right).

Figure 13. Scalar gravitational field $\Phi$ at time $t = 400M$ ($n = 40000$) (left) and at time $t = 500M$ ($n = 50000$) (right).
4.2. Numerical simulations of the toy model using the AFEM

We have just concluded that the central part of the numerical simulation of our toy model, and more in general of EMRBs, is the proper resolution of the matter source, $\rho$, and of the gravitational field, $\Phi_1$, in the surroundings of the particle position, which is a key issue in order to compute accurately the motion of the particle, and hence the waves that are emitted as a consequence. To that end, within the framework we have established above, we are able to compute accurately for small values of the particle’s width $\sigma$. In the limit $\sigma \to 0$ we recover the distributional description of the particle that appears in the continuum description. Then, it is obvious that better accuracy in the terms discussed above means an increase in the resolution around the particle’s position, and in this sense the AFEM is a natural choice since it provides the resolution required in the different regions of the computational domain, which maintains the computational cost at realistic levels.

According to the theoretical foundations of the adaptive mesh technique presented above, in order to apply mesh adaptivity successfully in our numerical computations, first of all, we have to study which quantity in the problem we are dealing with can be used in order to determine where and how the mesh should be refined. In technical terms, this means that we have to be able to determine the places where refinement (or derefinement) has to take place in terms of the Hessian matrix of the selected quantity. We have to look for places where the majoring Hessian matrix is much bigger than anywhere else, which are the places where the quantity we have chosen changes very fast. In the case of our toy model there is not too much choice. We can either take the field $\Phi$ or the source $\rho$. In the case of $\Phi$, it is interesting in the sense that it is one of the goals of our computations. However, since it is the solution of a wave-type equation, it evolves like a wave in the sense that the profile of the solution will have maxima and minima which will be captured by the adaptive mesh procedure, and hence most parts of the domain will be eventually refined, and this may not be the most efficient way.
to proceed. The other choice, the particle’s source $\rho$, is a better choice for the simple reason that we are solving for the particle’s position at the same time as we are solving the PDEs for the field. Therefore, the recognition of the regions that need refinement will be simpler in our problem. That is, the most efficient way of adapting the mesh for our computations is to concentrate on the region where the particle is present and follow it from there.

If we look at the behaviour found in the previous subsections (see figures 9–10), it happens that the bigger values of $\Phi$ occur around the particle’s position. Therefore, refining there is very convenient at that stage of the computation. A possible drawback of using the particle’s source as the refinement criterion is that the situation changes at later times when the field $\Phi$ is present all over the computational domain and there may be a need for refinement in other places such as close to the horizon. However, in the numerical simulations we have performed we find that using the source term $\rho$ for refinement is much better than using the field $\Phi$. By calculating the Hessian matrix of the particle’s source we can know exactly where and how to get local mesh refinement on the particle.

For the numerical simulations of the toy model with the AFEM we use the same parameters as in section 4.1, except for the width of the Gaussian, which we have been able to reduce down to $\sigma = 0.1M$. And this has been done using a number of triangular elements in the interval 11 000–15 000 (the number of elements changes with the evolution, depending on how the adaptivity is implemented), a number comparable to that used in the calculations without adaptivity. That is, working with the AFEM we have been able to use a width $\sigma$ an order of magnitude smaller than with the classical FEM, using a comparable number of elements. The time step is now $\Delta t = 0.005M$, half that used in the previous simulations. The trajectory followed by the particle is drawn in figure 15. We show the evolution of the scalar gravitational field in figures 16–21. These are contour plots such as those used previously. In order to show how we refine the mesh in the evolution we have superposed the mesh onto the contour plots. In this way, we can see how the high-resolution part of the mesh moves with the particle. One can clearly see the difference between this case and the previous case: the particle inspirals much faster in this case, and it hardly completes more than one orbit. To understand this difference, it is important to remark the fact that the particle sources are different since the Gaussian profiles have a different width $\sigma$ and the fact that the sources couple nonlinearly with the scalar gravitational field. As a consequence, the difference in $\sigma$ leads to a different evolution. We have checked that when we take $\sigma = 1M$ with the AFEM we recover the type of trajectories we get in the case without adaptivity.
Figure 16. Scalar gravitational field $\Phi$ at time $t = 5M$ ($n = 1000$) (left) and at time $t = 20M$ ($n = 4000$) (right).

Figure 17. Scalar gravitational field $\Phi$ at time $t = 35M$ ($n = 7000$) (left) and at time $t = 50M$ ($n = 10000$) (right).

Figure 18. Scalar gravitational field $\Phi$ at time $t = 65M$ ($n = 13000$) (left) and at time $t = 80M$ ($n = 160004$) (right).

A more detailed view of the structure of the refined region around the particle is shown in figure 22, where one can see the different levels of refinement, and how the resolution...
Figure 19. Scalar gravitational field $\Phi$ at time $t = 95M$ ($n = 19\,000$) (left) and at time $t = 110M$ ($n = 22\,000$) (right).

Figure 20. Scalar gravitational field $\Phi$ at time $t = 125M$ ($n = 25\,000$) (left) and at time $t = 140M$ ($n = 30\,000$) (right).

Figure 21. Scalar gravitational field $\Phi$ at time $t = 150M$ ($n = 25\,000$) (left) and at time $t = 160M$ ($n = 32\,000$) (right).

increases as one approaches the particle. A global perspective of the structure of the mesh, containing the excised region around the singularity of the black hole, and the refined region around the particle can be seen in the three-dimensional graph shown in figure 23.

We have also checked the energy-balance law (64) for the simulations with adaptivity. We have plotted the result in figure 24, where again the horizontal axis denotes time and
the vertical axis the value of the left-hand side of (64) normalized with the mass $m$ of the particle. The behaviour of the error in the conservation law is very similar to that in the case without adaptivity, where the error in the energy-balance test grows after the boundary affects the particle. In this case, the error does not seem to stabilize as clearly as in the previous case, and on average the error in the conservation law is bigger. However, one has to take into account that the evolution in this case is much faster since the particle has plunged very quickly. Moreover, since the number of elements in both cases is comparable, in the case of the AFEM a considerable number of them have been used to increase the resolution around the particle; this means that the density of elements near the boundaries is less and therefore if errors are propagated from there, they can affect the result more than in the case without adaptivity. This can be avoided by using more computational power to push the boundaries far enough so that we can neglect their effect during the significant part of the evolution.
5. Remarks and conclusions

In this work, we have studied the application of finite element techniques to the time-domain numerical simulations of EMRBs. More specifically, we have shown how finite elements can help us in achieving the degree of adaptivity that the computation of radiation reaction effects around the small body requires. However, adaptivity is not the only advantage in using finite elements for the study of EMRBs, some of which are obvious from the study of the toy model we have considered in this work. In short, the main additional advantages that finite elements provide are as follows. (i) Versatility. The FEM can be applied to a wide range of problems: static, quasi-static, transient, highly dynamical, linear and nonlinear, etc. Moreover, the practical implementation of the FEM can be easily designed in a modular way, as is shown by the existence of a number of multi-purpose finite element packages. This is a good property for the design of complex numerical codes, such as those needed for the description of astrophysical EMRBs. (ii) Many of the procedures that one uses in the framework of the FEM have a solid theoretical foundation, in the sense that behind them one often finds rigorous mathematical analysis (examples of this can be found in the classical textbooks [51–55]). (iii) The ability of the FEM to manage problems with complex geometries. For instance, we can easily consider spherical boundaries which is helpful in imposing boundary conditions and also in implementing excision techniques for black hole singularities. This is something that may be worth exporting to fully numerical relativistic calculations of spacetimes containing black holes. (iv) In relation to the previous issue, the imposition of boundary conditions can be done in a somewhat natural way in a FEM framework. The idea is that by adapting the mesh to the geometry of the problem under study, boundary conditions which are also adapted to that geometry will be incorporated in a natural way into the FEM discretization through the weak formulation of the problem. The outgoing boundary conditions that we have used in our toy
model are a good example; its implementation is simpler than it would have been in a similar finite differences scheme. (v) Another advantage of the FEM, which we have not explored in this paper, is that it is to some extent natural to handle distributions, like, for instance, the typical Dirac delta distributions that appear in physical systems containing particles (e.g., EMRBs). The idea is that we can use the exact properties of the distributions at the level of the weak formulation of the equations, which in the case of Dirac delta distributions leads automatically to a regularization of the singularity associated with these kinds of distributions. We can apply this idea, for instance, to the calculation of the general relativistic perturbations of a non-rotating black hole, produced by the presence of a point-like object, via the Regge–Wheeler–Zerilli formalism. Then, the type of FEM discretization we would get is similar to that obtained by Price and Lousto [76–78], where they also used an integral form of the equations in order to obtain a discretization. In addition, the FEM procedure can be generalized to higher dimensional problems in a straightforward way. These ideas will be the subject of future investigations.

In this paper, we have illustrated the capabilities of the FEM and its suitability to simulate EMRBs, by applying it to a simplified model, a toy model, that retains the basic ingredients and difficulties of general relativistic EMRBs. We have shown simulations of this system based on the classical FEM and simulations based on the AFEM. One outcome of this work is the realization that in order to increase the resolution around the small body, treated as a particle, the introduction of adaptivity in the region of the particle is necessary and in this sense the AFEM is a natural tool to use. The primary benefit of the AFEM is that it can provide an efficient, accurate and reliable computational analysis of very large continuum problems, for only a relatively small fraction of the cost associated with the non-adaptive FEM. The accuracy of a finite element solution is directly dependent on the number of free parameters used to mathematically represent the problem, and how effectively those parameters, or mathematical degrees of freedom, are distributed over the problem space. Furthermore, the full computational cost associated with obtaining a finite element solution is related to both the number and the interconnectivity of the degrees of freedom used in the problem discretization. Consequently, the most efficient distribution of degrees of freedom for a problem is the one that provides a sufficiently accurate solution for the lowest number of free parameters. Currently, the only practical way to achieve this objective is by using adaptive solution strategies which are capable of cleverly evolving and improving an efficient distribution of degrees of freedom over the problem domain by establishing solution error distributions, and then adjusting or adding degrees of freedom to the discretization in order to correct them. By increasing the numbers of degrees of freedom only in the regions of higher error in the solution, it is possible to make the most significant improvement in the global accuracy of the finite element solution for the minimum additional computational cost. In this paper, we employed the Hessian matrix of the solution to describe the solution error distribution, which can perfectly guide us to where and how we have to locally refine the mesh without any unnecessary pollution. In this sense, it is important to emphasize that the rest of the adaptive strategies available do not have this advantageous property. Moreover, this comes without paying an extra price since the main advantages of the other adaptivity techniques are present in the AFEM that we have used.

The natural continuation of this work is the transfer of the technology used here to the general relativistic problem. One possible way is to consider the general framework of metric perturbations in the set-up of the $3+1$ decomposition, that is, to use 3D perturbative numerical relativity, trying at the same time to profit from the experience gained in 3D full numerical relativity. However, 3D calculations using the AFEM may be at present computationally too demanding, and therefore other avenues should also be explored. Among them we
can consider one-dimensional calculations restricted to the case of a non-rotating MBH, by using the well-known Regge–Wheeler and Zerilli–Moncrief formalisms, or just by using a harmonic gauge, where the computation of self-forces seems more natural. This would be an interesting benchmark to test further the FEM techniques in a general relativistic context. From here, one can study problems of more interest for gravitational wave physics related to LISA (involving spinning MBHs) by considering 2D calculations using the curvature based formalism of Teukolsky for linear perturbations around Kerr black holes. One can go beyond by considering also perturbations of Kerr black holes using metric perturbations in a harmonic gauge, where the computation of self-forces can be carried out by using techniques already present in the literature.

Acknowledgments

CFS and PL acknowledge the support of the Center for Gravitational Wave Physics funded by the National Science Foundation under Cooperative Agreement PHY-0114375. This work was partially supported by NSF grant no PHY-0244788 to Penn State University. PS and JX acknowledge partial support from NSF grant no DMS-0209497 and from the Center for Computational Mathematics and Applications of Penn State University.

References

[1] Hough J, Rowan S and Sathyaprakash B 2005 Preprint gr-qc/0501007 (J. Phys. B: At. Mol. Opt. Phys. at press)
[2] Vitale S et al 2002 Nucl. Phys. Proc. Suppl. 110 209
[3] Danzmann K 2003 Adv. Space Res. 32 1233
[4] Danzmann K and Rudiger A 2003 Class. Quantum Grav. 20 S1
[5] Prince T 2003 Am. Astron. Soc. Meeting 202 3701
[6] Barack L and Cutler C 2004 Phys. Rev. D 69 084005 (Preprint gr-qc/0310125)
[7] Gair J R, Barack L, Creighton T, Cutler C, Larson S L, Phinney E S and Vallisneri M 2004 Class. Quantum Grav. 21 S1 S59
[8] Poisson E 2004 Living Rev. Rel. 7 6 (Preprint gr-qc/0306052) (http://www.livingreviews.org/lrr-2004-6)
[9] Poisson E 2005 Class. Quantum Grav. 21 R153
[10] Detweiler S 2005 Class. Quantum Grav. 22 S681 (Preprint gr-qc/0501004)
[11] Mino Y 2005 Class. Quantum Grav. 22 S717 (Preprint gr-qc/0506002)
[12] Mino Y 2005 Prog. Theor. Phys. 113 733 (Preprint gr-qc/0506003)
[13] Mino Y 2005 Class. Quantum Grav. 22 S375 (Preprint gr-qc/0506008)
[14] Tanaka T, Shibata M, Sasaki M, Tagoshi H and Nakamura T 1993 Prog. Theor. Phys. 90 65
[15] Cutler C, Kennefick D and Poisson E 1994 Phys. Rev. D 50 3816
[16] Hughes S A 2000 Phys. Rev. D 61 045004 (Preprint gr-qc/9910091)
[17] Hughes S A 2001 Class. Quantum Grav. 18 4067 (Preprint gr-qc/0008058)
[18] Hughes S A 2001 Phys. Rev. D 64 044004 (Preprint gr-qc/0104041)
[19] Glampedakis K, Hughes S A and Kennefick D 2002 Phys. Rev. D 66 064005 (Preprint gr-qc/0205053)
[20] Mino Y 2003 Phys. Rev. D 67 084027 (Preprint gr-qc/0302075)
[21] Hughes S A, Drasco S, Flanagan E E and Franklin J 2005 Phys. Rev. Lett. 94 221101 (Preprint gr-qc/0504015)
[22] Drasco S, Flanagan E E and Hughes S A 2005 Class. Quantum Grav. 22 S801 (Preprint gr-qc/0505075)
[23] Sago N, Tanaka T, Hikida W and Nakano H 2005 Prog. Theor. Phys. 114 509 (Preprint gr-qc/0506092)
[24] Mino Y, Sasaki M and Tanaka T 1997 Phys. Rev. D 55 3457 (Preprint gr-qc/9606018)
[25] Quinn T C and Wald R M 1997 Phys. Rev. D 56 3381 (Preprint gr-qc/9610053)
[26] Lousto C O 2005 Class. Quantum Grav. 22 S369 (Preprint astro-ph/0501238)
[27] Barack L and Ori A 2000 Phys. Rev. D 61 061502 (Preprint gr-qc/9912010)
[28] Barack L 2000 Phys. Rev. D 62 084027 (Preprint gr-qc/0005042)
[29] Barack L 2001 Phys. Rev. D 64 084021 (Preprint gr-qc/0105040)
[30] Barack L, Mino Y, Nakano H, Ori A and Sasaki M 2002 Phys. Rev. Lett. 88 091101 (Preprint gr-qc/0111001)
[31] Lousto C O 2000 Phys. Rev. Lett. 84 5251 (Preprint gr-qc/9912017)
[32] Bishop N T, Gomez R, Husa S, Lehner L and Winicour J 2003 Phys. Rev. D 68 084015 (Preprint gr-qc/0301060)
A toy model for testing finite element methods to simulate extreme-mass-ratio binary systems 285

[33] Lopez-Aleman R, Khanna G and Pullin J 2003 Class. Quantum Grav. 20 3259 (Preprint gr-qc/0303054)
[34] Teukolsky S A 1973 Astrophys. J. 185 635
[35] Krivan W, Laguna P, Papadopoulos P and Andersson N 1997 Phys. Rev. D 56 3395 (Preprint gr-qc/9702048)
[36] Arnold D N, Mukherjee A and Pouly L 1998 Numerical Analysis 1997 ed D F Griffiths, D J Higham and G A Watson (Essex: Addison-Wesley) pp 1–15
[37] Metzger J 2004 Class. Quantum Grav. 21 4625 (Preprint gr-qc/0408059)
[38] Chen L, Sun P and Xu J 2003 Math. Comput. at press
[39] Misner C W, Thorne K S and Wheeler J A 1973 Gravitation (San Francisco, CA: Freeman)
[40] Shapiro S L and Teukolsky S A 1993 Phys. Rev. D 47 1529
[41] Shapiro S L and Teukolsky S A 1994 Phys. Rev. D 49 1886
[42] Scheel M A, Shapiro S L and Teukolsky S A 1994 Phys. Rev. D 49 1894
[43] Shoemaker D et al. 2003 Class. Quantum Grav. 20 3729 (Preprint gr-qc/0301111)
[44] Sperhake U, Smith K L, Kelly B, Laguna P and Shoemaker D 2004 Phys. Rev. D 69 024012 (Preprint gr-qc/0307015)
[45] Bayliss A and Turkel E 1980 Commun. Pure Appl. Math. 33 707
[46] Bayliss A, Gunzburger M and Turkel E 1982 SIAM J. Appl. Math. 42 430
[47] Alpert B, Greengard L and Hagstrom T 2002 J. Comput. Phys. 180 270
[48] Lau S R 2004 J. Comput. Phys. 199 376 (Preprint gr-qc/0401001)
[49] Lau S R 2004 Class. Quantum Grav. 21 4147
[50] Strang G and Fix J G 1973 An Analysis of the Finite Element Method (Englewood Cliffs, NJ: Prentice-Hall)
[51] Zienkiewicz O C 1977 The Finite Element Method (New York: McGraw-Hill)
[52] Ciarlet P G 1978 The Finite Element Method for Elliptic Problems (Amsterdam: North-Holland)
[53] Hughes T J R 1993 The Finite Element Method (Englewood Cliffs, NJ: Prentice-Hall)
[54] Babuška I and Strouboulis T 2001 The finite element method and its reliability Numerical Mathematics and Scientific Computation (Oxford Science Publications) (Oxford: Oxford University Press)
[55] Beijing FEGEN Software Co. Ltd, http://www.femonline.net/
[56] Morin P, Nochetto R H and Siebert K G 2002 SIAM Rev. 44 631
[57] Rivara M C and Venere M 1996 Eng. Comput. 12 224
[58] Rivara M C and Iribarren G 1996 Math. Comput. 65 1485
[59] Dolejsi V 1998 Comput. Vis. Sci. 1 165
[60] Zienkiewicz O C and Zhou J Z, 1992 Int. J. Numer. Methods Eng. 33 1331
[61] Zienkiewicz O C and Zhou J Z, 1992 Int. J. Numer. Methods Eng. 33 1365
[62] Wahlbin L B 1995 Superconvergence in Galerkin Finite Element Methods (Berlin: Springer)
[63] Bank R and Xu J SIAM J. Numer. Anal. at press
[64] Bank R and Xu J SIAM J. Numer. Anal. at press
[65] Xu J and Zhang Z 2003 Math. Comput. (electronically published on August 19)
[66] Bank R E, Sherman A H and Weiser A 1983 Scientific Computing ed R S et al (Amsterdam: North-Holland) pp 3–17
[67] Rivara M C 1984 SIAM J. Numer. Anal. 21 604
[68] Komluheber R and Roitzsch R 1990 IMPACT Comput. Sci. Eng. 2 40
[69] Field D A 1988 Commun. Appl. Numer. Methods 4 709
[70] Bank R E and Smith R K 1997 SIAM J. Numer. Anal. 34 979
[71] Shephard M and Georges M 1991 Int. J. Numer. Methods Eng. 32 709
[72] Freitag L A, Jones M T and Plassmann P E 1995 4th Int. Meshing Roundtables (Sandia Labs) pp 47–58
[73] Chen L and Xu J 2004 Preprint
[74] Lousto C O and Price R H 1997 Phys. Rev. D 56 6439 (Preprint gr-qc/9705071)
[75] Martel K and Poisson E 2002 Phys. Rev. D 66 084001 (Preprint gr-qc/0107104)
[76] Martel K 2004 Phys. Rev. D 69 044025 (Preprint gr-qc/0311017)