Cure for unstable numerical evolutions of single black holes: Adjusting the standard ADM equations in the spherically symmetric case

Bernard Kelly, Pablo Laguna, Keith Lockitch, Jorge Pullin, Erik Schnetter, Deirdre Shoemaker, and Manuel Tiglio
Center for Gravitational Physics & Geometry, Penn State University, University Park, Pennsylvania 16802
(Received 28 March 2001; published 24 September 2001)

Numerical codes based on a direct implementation of the standard Arnowitt-Deser-Misner (ADM) formulation of Einstein's equations have generally failed to provide long-term stable and convergent evolutions of black hole space-times when excision is used to remove the singularities. We show that, for the case of a single black hole in spherical symmetry, it is possible to circumvent these problems by adding terms involving the constraints to the evolution equations, thus adjusting the standard ADM system. We investigate the effect that the choice of the lapse and shift has on the stability properties of numerical simulations and thus on the form of the added constraint term. To facilitate this task, we introduce the concept of quasi-well-posedness, a version of well-posedness suitable for ADM-like systems involving second-order spatial derivatives.

DOI: 10.1103/PhysRevD.64.084013 PACS number(s): 04.25.Dm, 04.70.Bw, 04.20.Ex, 04.20.Cv

I. INTRODUCTION

In numerical relativity, Einstein's equations are usually solved as an initial value problem. That is, the space-time is foliated with spacelike hypersurfaces. These hypersurfaces, or slices, are characterized by their intrinsic geometry (spatial metric g_{ij}) and extrinsic curvature K_{ij}. Subsequent slices in a given foliation are connected via the lapse function α and shift vector β^i [1]. Under this framework, the components of Einstein's equation naturally separate into constraint and evolution equations for the dynamical variables g_{ij} and K_{ij}. Thus a typical procedure to construct a space-time consists of first specifying Cauchy data (g_{ij},K_{ij}) that satisfy the constraints on the initial slice, and then applying the evolution equations to update these data into the next slice. This procedure is known as free or unconstrained evolution.

In order to propagate (g_{ij},K_{ij}) using the evolution equations, one must provide, in addition, a prescription for choosing the lapse function and shift vector. One generally regards the lapse-shift prescription as a choice of coordinates. At the continuum level, given a lapse-shift choice and a specific set of Cauchy data, the evolution equations yield not only a unique space-time metric expressed in a given coordinate system but also evolved data (g_{ij},K_{ij}) that continue to satisfy the constraints.

Numerical approximations are likely to complicate the picture described above. For instance, because of truncation errors, the data (g_{ij},K_{ij}) do not satisfy the continuum Einstein constraints but rather their discrete approximations. Even if consistency between the numerical approximations of the evolution and constraint equations is achieved, the evolved data (g_{ij},K_{ij}) will at best satisfy the constraints up to truncation errors. Of course, this is what one should expect. However, the presence of numerical errors could also trigger fast growing modes that render the numerical evolution unstable. Characterizing and controlling these growing modes has been and continues to be one of the most difficult and demanding tasks in numerical relativity.

The origin of these destabilizing modes is multiple. Unstable modes could be purely numerical (e.g., the Courant instability) or they could already be present at the continuum level. An example of continuum instabilities are the so-called constraint-violating modes [2]. These are solutions of the Einstein evolution equations that do not satisfy the Hamiltonian and momentum constraints. The general perception is that this class of solutions consists mostly of rapidly growing solutions, although there is no formal proof that this is indeed the case. In summary, even if one were to start an evolution with ''perfect'' initial data that satisfy the constraints, truncation errors could cause the numerical solution to drift into a constraint-violating, and perhaps unbound, solution.

Various groups accumulated numerical evidence supporting, as a possible cause of instabilities, the particular $3+1$ form used to recast Einstein's equation as an initial value problem. These observations partially motivated the development of hyperbolic formulations of Einstein equations [3–8]. Hyperbolic formulations have the advantage that mathematical tools exist to prove the existence and uniqueness of the solutions as well as well-posedness, thus providing important information for the implementation of stable discretization algorithms. If excision is used to handle the singularities, numerical simulations of black hole space-times require the imposition not only of outgoing, radiative boundary conditions far away from the holes but also of conditions on the excision boundary that respect the propagation of physical modes to flow down into the hole. Because hyperbolic formulations also yield information regarding the characteristics of all the field variables in the system, in principle this knowledge could be extremely useful when applying these boundary conditions.

Unfortunately, so far, none of the numerical efforts based on hyperbolic formulations of Einstein equations have been able to demonstrate their clear superiority. Currently, the formulation originally developed by Shibata and Nakamura [9], and later reintroduced by Baumgarte and Shapiro [10] (BSSN), seems to be the least prone to instabilities. Interestingly enough, the BSSN formulation is not explicitly hyperbolic. Just recently, a numerical implementation of the Einstein-Christoffel hyperbolic system [7] was applied to

single black hole space-time, and produced three-dimensional evolutions with stability properties comparable to those using the BSSN formulation [11]. The common lore these days is, however, that the standard Arnowitt-Deser-Misner (ADM) [12] formulation is the one which most easily suffers from instabilities [13–15]. On the other hand, the ADM formulation has the advantage of containing a minimal number of equations, making it attractive for numerical studies.

The main objective of this paper is to show that, given a choice of lapse function and shift vector, it is possible to obtain long-term stable and convergent simulations using the standard ADM formulation if ''appropriate'' terms involving the constraints are added to the evolution equations. Of course, here the key ingredient is to understand what constitutes an appropriate choice of constraint terms. We show that at least in spherical symmetry, it is possible to obtain a definite prescription for adding these constraint terms that only depends on the choice of gauge or coordinate conditions, namely the lapse and shift. The idea of adding constraint terms to the evolution equations (adjusted ADM formulations) is not new [16]. Just recently, Yoneda and Shinkai [17] studied in full detail the propagation of the constraints in the family of adjusted ADM systems.

In order to help identify what constitutes a suitable choice of constraint terms, we introduce a definition of quasi-well-posedness. The idea of quasi-well-posedness is simply to make choices of lapse and shift such that, when the ADM system is enlarged with the addition of new variables (typically, but not always, quantities related to the first spatial derivatives of the 3 metric), one obtains a strongly hyperbolic system. Once a strongly hyperbolic system is obtained, one can make use of standard results that apply to those systems regarding the existence, uniqueness, and well-posedness of the solutions. The essence of quasi-well-posedness is then that the properties of the strongly hyperbolic system are to some degree inherited by the original ADM system. The advantage of this procedure is that by looking at a strongly hyperbolic system, one is also able to gain insight into the propagation of the constraints and characteristics of the original system, thus improving the chances of obtaining stable numerical evolutions.

In Sec. II, we introduce the concept of quasi-well-posedness. The explicit form of the equations and initial data used in our numerical evolutions are described in Sec. III. The prescription to adjust the standard ADM equations is presented in Sec. IV. In Sec. V we investigate in the implications of the presence of gauge modes. Next, in Sec. VI, we discuss the quasi-well-posedness properties of the adjusted ADM system under three lapse and shift conditions. For each of these conditions, we present numerical results showing the stability properties of the simulations. We end in Sec. VII with concluding remarks.

II. QUASI-WELL-POSEDNESS

To understand the general idea of quasi-well-posedness, let us consider the initial value problem for the wave equation $\partial_{tt}\varphi=\partial_{xx}\varphi$, ignoring boundaries. By adding a new variable $\pi=\partial_t\varphi$, this equation takes the form

$$\partial_t u = P\partial_{xx}u + Qu, \tag{2.1}$$

where

$$u=\begin{pmatrix}\varphi\\ \pi\end{pmatrix}, \quad P=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} \quad \text{and} \quad Q=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.$$

One can view the system of equations (2.1) as an analog of the standard ADM equations. Now, adding one more variable $\xi=\partial_x\varphi$, we rewrite Eq. (2.1) as a first order system in both time and space. That is,

$$\partial_t\tilde{u}=A\partial_x\tilde{u}+B\tilde{u}, \tag{2.2}$$

with

$$\tilde{u}=\begin{pmatrix}\varphi\\ \pi\\ \xi\end{pmatrix}, \quad A=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix} \quad \text{and} \quad B=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Note that one needs to supplement system (2.2) with the the constraint $C=\xi-\partial_x\varphi$. For this simple case, one can find estimates for the solution u from Eq. (2.1) directly, but the idea here is not to do so but rather to use what is known about the first order system (2.2). The matrix A has the following set of eigenvalues and eigenvectors:

$$\begin{aligned}&\lambda_1=1 \quad \text{with eigenvector} \quad \vec{e}_1=[0,1,1],\\ &\lambda_2=0 \quad \text{with eigenvector} \quad \vec{e}_2=[1,0,0],\\ &\lambda_3=-1 \quad \text{with eigenvector} \quad \vec{e}_3=[0,-1,1].\end{aligned} \tag{2.3}$$

Given that matrix A has real eigenvalues and a complete system of eigenvectors, the system of equations (2.2) is strongly hyperbolic[1] and thus well-posed. In this case, the matrix A is also Hermitian (symmetric hyperbolicity). and has real and distinct eigenvalues (strictly hyperbolic).

Since system (2.2) is well-posed, we can establish the existence and uniqueness of the solution to system (2.2), together with the usual condition

$$\|\tilde{u}(t,\ )\|^2 \leqslant De^{at}\|\tilde{u}(0,\ )\|^2, \tag{2.4}$$

with a and D independent of the initial data and $\|\ \|$ the L_2 norm. Furthermore, because the vector $\tilde{u}$ contains u as part of its components,

$$\|u(t,\ )\|^2 \leqslant \|\tilde{u}(t,\ )\|^2.$$

[1]Following Ref. [18], a system is called symmetric hyperbolic if the matrix A is Hermitian. It is called strictly hyperbolic if the eigenvalues are real and distinct; it is called strongly hyperbolic if the eigenvalues are real and there exists a complete system of eigenvectors; finally, it is called weakly hyperbolic if the eigenvalues are real.

It then follows from Eq. (2.4) that the solution of system (2.1) satisfies the condition

$$\|u(t,\ )\|^2 \leq De^{at}(\|u(0,\ )\|^2+\|\partial_x u(0,\ )\|^2). \quad (2.5)$$

We should remark that, although Eq. (2.5) was derived using L_2 norms, the solution could in general be bounded by the initial data in some other Sobolev norm. The importance of condition (2.5) and similar bounds is that it guarantees, among other things, that the solutions to system (2.1) will not have unbounded high frequency modes. In numerical calculations however, this is not enough. For practical purposes, even large power-law growths such as t^4 are likely to be extremely problematic [2].

The idea is then to investigate quasi-well-posedness in the standard ADM formulation. That is, given a choice of lapse and shift, one enlarges the standard ADM system introducing new variables that render the system at least strongly hyperbolic. Existence, uniqueness and well-posedness for the ADM equations will then follow from the corresponding usual properties of the strongly hyperbolic system. In some instances, to obtain such hyperbolic formulations it is not enough just to add the spatial derivatives of the metric as new variables; one also has to add the constraints to the evolution equations, or even take combinations of these evolution equations. But the general idea is the same: add the constraints and/or take the same combination in the ADM equations as one does in the strongly hyperbolic system. It will be in this sense that we will refer, in what follows, to the quasi-well-posedness of some ADM systems.

It is important to mention that for this analysis we have neglected the effects from boundary conditions. For the case under consideration (single black hole evolutions via excision of the singularity) boundary conditions at the excised region do not affect quasi-well-posedness since, as we shall see, all the characteristics of the system are outgoing into the hole. At the outer edge of the computational domain, we use the exact analytic solution as boundary condition. Our numerical results show that this does not influence the properties drawn from quasi-well-posedness for the class of gauge choices and added constraint-terms we considered.

III. $\dot{g}$-$\dot{K}$ EQUATIONS

Although the idea of quasi-well-posedness can be applied to more general systems, throughout this paper we will concentrate on the ADM evolution equations ($\dot{g}$-$\dot{K}$ equations) obtained from the vanishing of the Ricci tensor, i.e.,

$$\partial_t g_{ij}-\mathcal{L}_\beta g_{ij}=-2\alpha K_{ij}, \quad (3.1)$$

$$\partial_t K_{ij}-\mathcal{L}_\beta K_{ij}=-\nabla_i\nabla_j\alpha+\alpha(R_{ij}+KK_{ij}-2K_{ik}K^k_j). \quad (3.2)$$

As customary in numerical relativity, we will focus on free evolutions; that is, we will not enforce the constraints

$$2H\equiv R+K^2-K_{ij}K^{ij}=0, \quad (3.3)$$

$$M^i\equiv\nabla_j(K^{ij}-g^{ij}K)=0, \quad (3.4)$$

except at the initial data. Finally, to facilitate our analysis, we will only consider a single black hole in spherical coordinates. The most general time-dependent spherically symmetric metric in a $3+1$ form is

$$ds^2=(-\alpha^2+a^2\beta^2)dt^2+2a^2\beta dtdr+a^2dr^2+b^2d\Omega^2, \quad (3.5)$$

where all functions are assumed to depend only on r and t, $d\Omega^2\equiv d\theta^2+\sin^2\theta d\phi^2$, and $\beta\equiv\beta^r$. In these spherical coordinates, the spatial metric, extrinsic curvature and Ricci tensor are diagonal:

$$g_{ij}=\mathrm{diag}(a^2,b^2,b^2\sin^2\theta), \quad (3.6)$$

$$K^i_{\ j}=\mathrm{diag}(K_a,K_b,K_b), \quad (3.7)$$

$$R^i_{\ j}=\mathrm{diag}(R_a,R_b,R_b). \quad (3.8)$$

The $\dot{g}$-$\dot{K}$ equations take the forms

$$(\partial_t-\beta\partial_r)a=-\alpha aK_a+a\partial_r\beta, \quad (3.9)$$

$$(\partial_t-\beta\partial_r)b=-\alpha bK_b, \quad (3.10)$$

$$(\partial_t-\beta\partial_r)K_a=-\frac{1}{a^2}\left(\partial_r^2\alpha-\frac{1}{a}\partial_r a\partial_r\alpha\right)+\alpha[R_a+(K_a+2K_b)K_a], \quad (3.11)$$

$$(\partial_t-\beta\partial_r)K_b=-\frac{1}{ba^2}\partial_r b\partial_r\alpha+\alpha[R_b+(K_a+2K_b)K_b] \quad (3.12)$$

and the constraints

$$H\equiv\frac{R_a}{2}+R_b+2K_aK_b+K_b^2=0, \quad (3.13)$$

$$M\equiv\partial_r K_b+(K_b-K_a)\frac{\partial_r b}{b}=0, \quad (3.14)$$

where the components of the Ricci tensor are given by

$$R_a=\frac{2}{a^3b}(-a\partial_r^2 b+\partial_r a\partial_r b), \quad (3.15)$$

$$R_b=\frac{1}{a^3b^2}[-ba\partial_r^2 b+b\partial_r a\partial_r b+a^3-a(\partial_r b)^2]. \quad (3.16)$$

Our numerical results consist of free evolutions of analytic, single black hole initial data (a,b,K_a,K_b) in a computational domain $r_e\leq r\leq r_o$, where r_o denotes the location of the outer boundary far from the hole, and r_e the excision boundary inside the black hole horizon. The numerical code used solves the $\dot{g}$-$\dot{K}$ equations in the interior of the computational domain by the method of lines. That is, the $\dot{g}$-$\dot{K}$ system has the generic form

$$\partial_t u = \beta \partial_r u + \rho, \tag{3.17}$$

with $u = u(a,b,K_a,K_b)$ and ρ given by the right hand side of Eqs. (3.9)–(3.12). The starting point is then to approximate the derivatives appearing in ρ by second order accurate, centered finite differences. On the other hand, the the spatial derivative in the ''advection'' term $\beta \partial_r u$ is approximated by an upwind discretization [19] of the form

$$\partial_r u_i \approx \Delta_r u_i \equiv \frac{(u_{i+1} - u_{i-1})}{2\,\Delta r} \pm q \frac{(u_{i\mp 1} - 3u_i + 3u_{i\pm 1} - u_{i\pm 2})}{3\Delta r}, \tag{3.18}$$

where $q \geqslant 0$ and the choice of sign is given by $\pm = \mathrm{sgn}(\beta)$. As we shall see below, for the black hole space-time metrics under consideration, the shift vector is always non-negative. Thus, the discretization [Eq. (3.18)] is only needed with the upper sign. The above upwind discretization has a truncation error

$$\tau_i \equiv \Delta_r u_i - \partial_r u_i = \frac{1}{6}(1-2q)\Delta r^2 \partial_r^3 u_i \mp \frac{q}{6}\Delta r^3 \partial_r^4 u_i\,. \tag{3.19}$$

Thus the discretization is second order accurate for any choice of the adjustable parameter q, except for $q = 0.5$ when the truncation errors are of $O(\Delta r^3)$. In our simulations, we set $q = 0.5$. This value yields both stable evolutions and minimizes the amount of dissipation introduced by the discretization. After the discretization of the spatial differential operators, Eq. (3.17) becomes a set of ordinary differential equations for the interior grid points $\{r_i\}_{i=1,\ldots,N-1}$. The grid values r_0 and r_N denote the locations of the inner and outer boundaries, respectively. The temporal updating of these equations is carried out via an iterated Crank-Nicholson method [20]. Note that at the outermost grid point r_{N-1}, the upwind scheme [Eq. (3.18)] requires ''ghost'' values at r_{N+1}. These values are computed using second or third order accurate extrapolations.

Next is the implementation of boundary conditions. As mentioned above, there are two boundaries in the problem, one far from the black hole at $r_N = r_o$ and a second at the point of excision, $r_0 = r_e$. The boundary condition used at r_o is that all field variables take the values provided by the exact analytic solution. It is important to note that in previous efforts in numerical evolutions of spherically symmetric black hole space-times, the stability of the evolutions depended not only in the location of the outer boundary but also on the imposition of ''outgoing'' boundary conditions. One of these outgoing conditions is constructed by assuming that the solution error $u - \bar{u}$ is of the form $\epsilon(t-r)/r^n$, with $\bar{u}$ the analytic exact solution. This condition yields

$$\partial_t u = -\partial_r u + \partial_r \bar{u} + n\frac{(u-\bar{u})}{r}, \tag{3.20}$$

to be applied at the nodal grid point $r_N = r_o$ The outgoing condition [Eq. (3.20)] is then added to system (3.17) and handled with the same method of lines used for the interior points.[2] Obviously, there is no reason to believe that the solution error should behave as $\epsilon(t-r)/r^n$ for general choices of lapse and shift. Another approach to construct outer boundary conditions that has been used is to blend [21] the numerical solution to the analytic one beyond a certain radius, r.

Finally, at the inner or excision boundary, the working assumption is that all of the fields have outgoing (into the hole) characteristics. We will show that for the gauge or coordinate conditions under consideration, this is indeed the situation. Therefore, there is no need to impose any boundary condition, and at r_e one can just apply the same system of equations used in the interior of the computational domain. However, a finite difference discretization would require ''ghost'' values at r_{-1}. We construct these values by extrapolation. Another possibility, which is becoming popular in three-dimensional simulations [22], is instead to extrapolate the right hand side of Eq. (3.17) to $r_0 = r_e$. Either approach was stable for the cases considered in the present work.

An important requirement when performing numerical simulations of black holes in which the singularities are excised from the computational domain is to use coordinates that are regular and penetrate the horizon. Foliations that penetrate the horizon facilitate the task of removing (i.e., excising) a region containing the black hole singularity while preserving the causal structure of the space-time exterior to the event horizon. Thus the numerical results we present here consist of numerical evolutions that for infinite resolution correspond to the solution of a single black hole expressed in ingoing-Eddington-Finkelstein (IEF) and also Painlevé-Gullstrand (PG) [23,24]. coordinates. The IEF coordinates coincide with Kerr-Schild coordinates in the case of zero angular momentum [25]. Recently, Martel and Poisson [26] have showed that the PG and IEF coordinates are members of the same one-parameter family of coordinate systems.

In IEF coordinates, the line element [Eq. (3.5)] takes the explicit form

$$ds^2 = -\left(1 - \frac{2m}{r}\right)dt^2 + \frac{4m}{r}dt\,dr + \left(1 + \frac{2\,m}{r}\right)dr^2 + r^2 d\Omega^2. \tag{3.21}$$

The ADM variables in these coordinates are given by

$$\alpha = \left(1 + \frac{2m}{r}\right)^{-1/2}, \tag{3.22}$$

$$\beta = \frac{2m}{r}\left(1 + \frac{2m}{r}\right)^{-1}, \tag{3.23}$$

$$a = \left(1 + \frac{2m}{r}\right)^{1/2}, \tag{3.24}$$

$$b = r, \tag{3.25}$$

[2]It is important to emphasize that this outgoing condition applies only to the assumed behavior of the solution error since in spherically symmetric space-times there is no gravitational radiation.

$$K_a = -\frac{2m}{r^3}(r+m)\left(1+\frac{2m}{r}\right)^{-3/2}, \tag{3.26}$$

$$K_b = \frac{2m}{r^2}\left(1+\frac{2m}{r}\right)^{-1/2}. \tag{3.27}$$

Similarly, for PG coordinates,

$$ds^2 = -\left(1-\frac{2m}{r}\right)dt^2 + 2\sqrt{\frac{2m}{r}}dtdr + dr^2 + r^2 d\Omega^2 \tag{3.28}$$

and

$$\alpha = 1, \tag{3.29}$$

$$\beta = \sqrt{\frac{2m}{r}}, \tag{3.30}$$

$$a = 1, \tag{3.31}$$

$$b = r, \tag{3.32}$$

$$K_a = -\frac{\beta}{2r}, \tag{3.33}$$

$$K_b = \frac{\beta}{r}. \tag{3.34}$$

The geometrical interpretation of the IEF coordinate system is that in addition to having a timelike killing vector, the combination of timelike and radial tangent vectors $\vec{\partial}_t$-$\vec{\partial}_r$ remains null. In terms of the space-time metric, this condition is stated as $g_{tt}-2g_{tr}+g_{rr}=0$, or similarly in terms of 3+1 metric functions as $\alpha = a(1-\beta)$. On the other hand, the PG coordinate system can be viewed as that anchored to a family of freely moving observers (timelike) starting at infinity with vanishing velocity [26].

Finally, in numerical evolutions of spherically symmetric space-times, it is useful to monitor not only the Hamiltonian and momentum constraints but to pay attention also to the mass function [30]

$$\mathcal{M}(r,t) = \frac{b}{2}(1-\nabla_\mu b \nabla^\mu b). \tag{3.35}$$

In vacuum, this $\mathcal{M}$ is the gauge invariant definition of mass.

IV. ADJUSTED ADM SYSTEMS

Each of the equations in the $\dot{g}$-$\dot{K}$ or standard ADM system (3.9)–(3.12) has the form

$$\partial_t u_{(n)} - \beta \partial_r u_{(n)} = B_{(n)} u_{(n)} + C_{(n)} \tag{4.1}$$

(no summation over n), with

$$u_{(n)} = (a, b, K_a, K_b), \tag{4.2}$$

$$B_{(n)} = \mathrm{diag}(-\alpha K_a + \partial_r \beta, -\alpha K_b, 2\alpha K_b, \alpha K_a). \tag{4.3}$$

$C_{(n)}$ contain the remaining terms that cannot be written in the form $B_{(n)}u_{(n)}$, with $B_{(n)}$ independent of $u_{(n)}$. If one views each equation as independent, that is β, $B_{(n)}$, and $C_{(n)}$ as given and independent of $u_{(n)}$, Eqs. (4.1) rapidly admit growing solutions if $B_{(n)}>0$, exponentially growing if $B_{(n)}$ and β are non-negative constants. Obviously, when Eqs. (4.1) are considered as a coupled system of equations, one cannot guarantee the above conclusion. However, we have been able to track the problems observed in our numerical evolutions to those terms for which $B_{(n)}>0$.

In IEF and PG coordinates, the second and fourth components of $B_{(n)}$ are nonpositive. The first component of $B_{(n)}$ vanishes for PG coordinates and is positive for IEF coordinats. Finally, the third component of $B_{(n)}$ is always nonnegative for both IEF and PG coordinates. Our numerical experiments indicate that the origin of the instabilities is due to the term involving $B_{(n=3)}=2\alpha K_b$. The term in IEF coordinates involving $B_{(n=1)}$ with the ''wrong'' sign did not seem to affect the evolutions (see results below).

The objective is then to find a way to ''change the sign'' of $B_{(n=3)}$. Fortunately, the combination $K_b K_a$ also appears in the Hamiltonian constraint [Eq. (3.13)]. One can then add a term of the form $-\mu\alpha H$ to the evolution equation for K_a. As a consequence, $B_{(n=3)}=2(1-\mu)\alpha K_b$. Therefore, in principle, any value of $\mu \geqslant 1$ should produce stable evolutions. Our numerical simulations show that $\mu=2$ is an optimal value to reach quickly the time independent solution. However, the same experiments indicate also that even values of $\mu \approx 0.5$ yield stable evolutions. The reason for this is likely because the simple analysis above does not take into consideration the non-linear coupling in the equations.

V. LINEAR GAUGE MODES

A potential source of instabilities in numerical simulations is the presence of gauge modes. Gauge modes arise because a prescription of the lapse and shift is not sufficient for complete gauge fixing. If the lapse and shift functions are, for instance, determined from algebraic expressions or differential equations, there still remains freedom to perform a coordinate transformation that takes us from (g_{ij},K_{ij}) to $(\tilde{g}_{ij},\tilde{K}_{ij})$, leaving the lapse-shift prescription invariant. In other words, the data (g_{ij},K_{ij}) will be unique up to coordinate transformations that leave the lapse-shift prescription invariant. One can then encounter a situation in which the transformed pair $(\tilde{g}_{ij},\tilde{K}_{ij})$ continues to satisfy the constraints, but possesses unbounded growth. In principle, because $(\tilde{g}_{ij},\tilde{K}_{ij})$ are allowed solutions to the evolution equations, through numerical convergence and monitoring of the constraints one could single out these modes. In practice the situation is not that simple. Gauge conditions that allow the rapid growth of grid functions are likely to trigger numerical instabilities.

To investigate these modes, let us consider an infinitesimal coordinate transformation

$$x^\mu \to x^\mu + \xi^\mu. \tag{5.1}$$

This transformation induces a linear perturbation of the space-time metric,

$$g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}, \tag{5.2}$$

of the form

$$\delta g_{\mu\nu} = -\mathcal{L}_\xi g_{\mu\nu}. \tag{5.3}$$

In spherical symmetry, the most general coordinate transformations are

$$t \to t + \delta t(r,t), \tag{5.4}$$

$$r \to r + \delta r(r,t), \tag{5.5}$$

or equivalently

$$\xi^\mu = (\delta t, \delta r, 0, 0). \tag{5.6}$$

Given Eq. (5.6), the gauge induced perturbations of the metric take the forms

$$\begin{aligned}\delta g_{tt} = &-2(-\alpha^2 + a^2\beta^2)\partial_t \delta t - 2a^2\beta\partial_t \delta r \\ &- \partial_t(-\alpha^2 + a^2\beta^2)\delta t - \partial_r(-\alpha^2 + a^2\beta^2)\delta r,\end{aligned} \tag{5.7}$$

$$\begin{aligned}\delta g_{tr} = &-a^2\beta\partial_t \delta t - a^2\partial_t \delta r - (-\alpha^2 + a^2\beta^2)\partial_r \delta t \\ &- a^2\beta\partial_r \delta r - \partial_t(a^2\beta)\delta t - \partial_r(a^2\beta)\delta r,\end{aligned} \tag{5.8}$$

$$\delta g_{rr} = -2a^2\beta\partial_r \delta t - 2a^2\partial_r \delta r - \partial_t(a^2)\delta t - \partial_r(a^2)\delta r, \tag{5.9}$$

$$\delta g_{\theta\theta} = -\partial_t(b^2)\delta t - \partial_r(b^2)\delta r, \tag{5.10}$$

$$\delta g_{\phi\phi} = \delta g_{\theta\theta}\sin^2\theta, \tag{5.11}$$

with all remaining components vanishing. On the other hand, from Eq. (3.5), the components of the perturbed metric $\delta g_{\mu\nu}$ above are also given by

$$\delta g_{tt} = -2\alpha\delta\alpha + 2a\,\delta a\beta^2 + 2a^2\beta\delta\beta, \tag{5.12}$$

$$\delta g_{tr} = 2a\,\delta a\beta + a^2\delta\beta, \tag{5.13}$$

$$\delta g_{rr} = 2a\,\delta a, \tag{5.14}$$

$$\delta g_{\theta\theta} = 2b\,\delta b. \tag{5.15}$$

In studying these gauge perturbations, we consider a number of different prescriptions for choosing the lapse and shift. These impose restrictions on the components of $\delta g_{\mu\nu}$ to linear order, but are not sufficient for complete gauge fixing.

One piece of evidence usually given to argue for the highly unstable properties of the standard ADM formulation is its inability to simulate single black hole space-times. Specifically, the problem consists of using a known analytic black hole solution of Einstein's equation to construct initial data, set boundary conditions, and specify the lapse and shift. Given this input, the output is the numerical evolution of the spatial metric g_{ij} and extrinsic curvature K_{ij}. This simple setup does not yield stable evolutions.

Choosing the lapse and shift to be given by exact analytic solutions implies that in the coordinate transformations above, $\delta\alpha = \delta\beta = 0$. We shall call this the "exact lapse plus exact shift" (EL+ES) condition. It is clear from Eqs. (5.7)–(5.11) that the EL+ES choice does not fix the coordinates uniquely, but rather results in a nontrivial system of partial differential equations (PDE's) to be satisfied by δt and δr. The solutions to this PDE system form an equivalence class of gauges all satisfying the same condition on the lapse and shift. At the analytic level, one selects a unique member of this equivalence class by imposing boundary and initial conditions. Numerically, however, because of truncation errors, this is only possible if the evolution is stable.

Substitution of $\delta\alpha = \delta\beta = 0$ in Eqs. (5.12)–(5.14) yields $\delta g_{tr} - \beta\delta g_{rr} = \delta g_{tt} - \beta^2\,\delta g_{rr} = 0$. Using Eqs. (5.7)–(5.9), these conditions can be rewritten in terms of the gauge perturbation ξ^μ as

$$\partial_t \xi = A\partial_r \xi + B\xi, \tag{5.16}$$

where

$$\xi = \begin{pmatrix} \delta t \\ \delta r \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} \beta & 0 \\ \dfrac{\alpha^2}{a^2} & \beta \end{pmatrix}.$$

The matrix A has a degenerate eigenvalue $\lambda = \beta$ and a corresponding eigenvector $\vec{e} = [0,1]$. Therefore, the system is only weakly hyperbolic and thus ill-posed. It is then not possible to guarantee the absence of rapidly growing gauge modes. Gauge instabilities, by themselves, do not violate the constraints; however, they are likely to trigger numerical instabilities and thereby couple to constraint-violating modes. Note that this conclusion does not depend on the use of the standard ADM formulation, but applies to any initial value formulation of Einstein's equation. This is likely the reason why it has not been possible to produce hyperbolic formulations of Einstein equations using the EL+ES lapse-shift condition. Other groups reported numerical instabilities associated with an EL+ES prescription. (See, for example, Refs. [27,28].) Our result, which provides some analytic insight into the source of these instabilities, extends the analysis of Ref. [27] by including perturbations of a general spherically symmetric space-time [Eq. (3.5)] and by making no assumptions about the form of the coupling between gauge and constraint-violating perturbations.

Finally, the inability to guarantee the absence of unbound gauge modes does not necessarily imply that it is impossible to design an evolution scheme that is long term stable and convergent. As we shall see, by adjusting the standard ADM system with constraint terms, stable evolutions are possible even in the presence of these gauge modes.

VI. STABLE ADJUSTED ADM SYSTEMS IN ONE DIMENSION

We consider next a series of lapse-shift choices. For each choice, we investigate (1) the quasi-well-posedness of the resulting system of adjusted ADM evolution equations, (2) the propagation of the constraints, and (3) the convergence and stability of numerical simulations.

A. Exact lapse plus area locking

We can take advantage of the assumed spherical symmetry of the problem and ''lock'' the area of constant-r surfaces. That is, we exploit the lapse-shift freedom and set $\partial_t g_{\theta\theta}=0$, or equivalently $\partial_t b=0 \ \ \forall \ \ t$. From Eq. (3.10), this yields

$$0=-\beta\partial_r b+\alpha b K_b, \tag{6.1}$$

which can be seen as an algebraic equation to solve for β or α. We will use Eq. (6.1) as

$$\beta=\frac{\alpha\, b\, K_b}{\partial_r b}, \tag{6.2}$$

with b determined by the initial data. In our case, for both IEF and PG coordinates, $b=r$. In addition, we choose an exact lapse, i.e., an arbitrary but *a priori* specified function of space-time. Here again, since the goal is to reproduce numerically the analytic solution, we set the lapse to that given by the IEF or PG solutions.

This exact lapse, area locking (EL+AL) gauge condition was previously investigated in Ref. [27]. However, the implementation of locking the area was done at the numerical level. That is, condition (6.2) was not explicitly used. Instead, during the temporal updating of grid functions, a correction to the shift was introduced to keep the area locked. With this numerical area locking and with a blending of outer boundary conditions, stable simulations were reported in Ref.. [27] for computational domains with $r_o \leqslant 11m$.

Interestingly enough, this EL+AL choice of lapse and shift yields an ADM system of equations already first order in space and quasi linear: namely,

$$\partial_t u=A\partial_r u+Bu, \tag{6.3}$$

with

$$u=\begin{pmatrix} a \\ K_a \\ K_b \end{pmatrix} \quad \text{and} \quad A=\begin{pmatrix} \beta & 0 & \alpha a r \\ \dfrac{r\partial_r\alpha+2(1-\mu)\alpha}{a^3 r} & \beta & 0 \\ \dfrac{\alpha}{a^3 r} & 0 & \beta \end{pmatrix},$$

where β is given by Eq. (6.2) and we have set $b=r$ to simplify notation. In the numerical evolutions, however, we do not set $b=r$. The numerical code includes the evolution equation for b. We do not explicitly write the matrix B since it is not necessary for the analysis below. However, we must emphasize that, in order to achieve stability, a value of $\mu \geqslant 1$ is in principle needed (see Sec. IV).

Although this EL+AL case is not representative of the general ADM equations, where second spatial derivatives do appear, we will use it to introduce the main techniques and ideas regarding quasi-well-posedness. We note first that the matrix A has eigenvalues

$$\lambda_1=\beta, \tag{6.4}$$

$$\lambda_2=\beta+\frac{\alpha}{a}, \tag{6.5}$$

$$\lambda_3=\beta-\frac{\alpha}{a}, \tag{6.6}$$

and corresponding eigenvectors

$$\vec{e}_1=[0,1,0] \tag{6.7}$$

$$\vec{e}_2=\left[a^2 r,\frac{r\partial_r\alpha+2(1-\mu)\alpha}{\alpha},1\right], \tag{6.8}$$

$$\vec{e}_3=\left[-a^2 r,\frac{r\partial_r\alpha+2\,(1-\mu)\alpha}{\alpha},1\right]. \tag{6.9}$$

Because all of the eigenvalues are real and distinct, system (6.3) is strictly hyperbolic independent of the addition of the constraint term. The EL+AL system of equations is then an example of a hyperbolic system that, unless suitable constraint-terms are added, is subject to developing rapidly growing solutions (see Ref. [2] for another example). It is also important to note that λ_1 represents a characteristic speed corresponding to propagation along the timelike normal to the foliation. Similarly, λ_2 and λ_3 are characteristic speeds along the light cone.

Let us now consider the particular case of initial data and lapse function constructed from the single black hole solution in IEF or PG coordinates. The eigenvalues are given in these coordinates by

$$\lambda_1=\frac{2m}{r+2m}=\sqrt{\frac{2m}{r}}, \tag{6.10}$$

$$\lambda_2=1=\sqrt{\frac{2m}{r}}+1, \tag{6.11}$$

$$\lambda_3=\frac{2m-r}{2m+r}=\sqrt{\frac{2m}{r}}-1, \tag{6.12}$$

where the first and second equalities are for IEF and PG coordinates, respectively. Since the excision boundary r_e is by construction inside the black hole horizon (i.e., $r_e \leqslant r_h \equiv 2m$), we have that all the eigenvalues are positive there. By looking at the principal part of Eq. (6.3), it is easy to see that non-negative eigenvalues imply a propagation of field

variables in the direction of decreasing r-coordinate. Therefore, at r_e all the fields propagate out of the computational domain into the hole singularity; thus boundary conditions are not required. On the other hand, at the outer boundary r_o, one has $\lambda_3<0$, with the remaining eigenvalues still nonnegative. At r_o, one has then two ingoing modes (λ_1 and λ_2) and only one outgoing mode (λ_3). Given this information, it is perfectly possible to impose a condition suppressing modes entering the computational domain. However, as we mentioned above, we choose not to do so and impose at r_o the analytic, exact solutions. The direct consequence will be the appearance of a pulse at the outer boundary of the computational domain due to discontinuities in the truncation errors between the outermost evolved point and the boundary point r_o. This pulse propagates in the direction of the black hole and leaves the computational domain through the excision boundary.

The next step is to analyze the effect that the EL+AL choice has on the propagation of the constraints. At the continuum level, for arbitrary choice of lapse, shift and initial data (g_{ij},K_{ij}) satisfying constraints (3.3) and (3.4) and the evolution equations (3.1) and (3.2) guarantee, ignoring for the moment boundary conditions, that the evolved data will continue to satisfy the constraints. If one now takes boundary conditions, into consideration it is important to keep in mind that boundary data (g_{ij},K_{ij}) must satisfy the constraints. By looking at the way constraints propagate, i.e., their characteristics, one gains insight into the allowed boundary conditions consistent with the constraints. Another important aspect of well-posedness in the propagation of the constraints is that it guarantees that there will be no unbounded high frequency growth appearing in the constraints if they are not exactly satisfied at the initial slice (for example, due to numerical errors). This well-posedness for the constraint propagation is a nontrivial property, not possible to prove for a generic formulation of Einstein's equations [29].

What we look for are evolution equations for the constraints, equations that would hold if system (6.3) is satisfied. They can be found by taking time derivatives in both sides of Eqs. (3.13) and (3.14), replacing the time derivatives of the metric by the right hand sides of Eqs. (3.9)–(3.12), and finally expressing the metric and its spatial derivatives in terms of the constraints and their spatial derivatives. Following this procedure, it is not too difficult to show that

$$\partial_t v = P\,\partial_r v + Q v, \tag{6.13}$$

where now

$$v=\begin{pmatrix} H \\ M \end{pmatrix} \quad \text{and} \quad P=\begin{pmatrix} \beta & \dfrac{4\alpha}{a^2} \\ \dfrac{\alpha}{4} & \beta \end{pmatrix}. \tag{6.14}$$

The matrix P has eigenvalues $\overline{\lambda}_1=\lambda_2$ and $\overline{\lambda}_2=\lambda_3$, with λ_2 and λ_3 given by Eqs. (6.5) and (6.6), respectively. This implies that system (6.13) is also strictly hyperbolic with characteristic speeds along the light cone. Note that the constraints at r_e propagate out of the computational domain into the hole singularity, consistent with the outgoing propagation of the field variables at r_e, namely the tilting, into the black hole, of the light cone. At the outer boundary r_o, there is an ingoing mode ($\overline{\lambda}_2>0$). Therefore, one has, as expected, to be careful to provide data at r_o consistent with this entering mode. Since we are imposing at the outer boundary the analytic IEF and PG solutions, the data at r_o already satisfy the constraints. However, as mentioned above, choosing the lapse and shift does not completely fix the gauge freedom; thus one still has to be careful handling the gauge modes described in Sec. V.

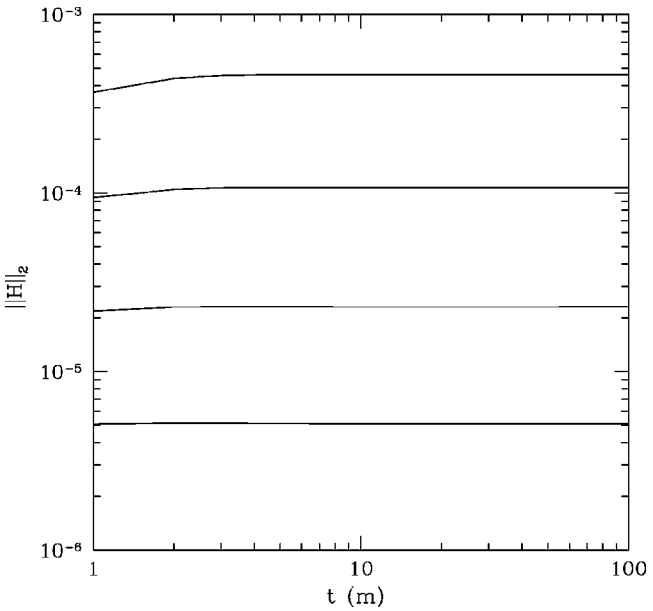

FIG. 1. L_2 norm of the Hamiltonian constraint as a function of time for IEF initial data and the EL+AL lapse shift. The computational domain extends from $r_e=1m$ to $r_o=40m$. Lines from top to bottom correspond to resolutions of $\Delta r=m/5, m/10, m/20, m/40$, respectively.

Figure 1 shows the L_2 norm of the Hamiltonian constraint as a function of time. The initial data are given by the IEF analytic solution [Eq. (3.21)], and the lapse and shift are chosen from the EL+AL conditions. Similar results were obtained with PG coordinates. The computational domain extends from $r_e=1m$ to $r_o=40m$. We tried larger and smaller values for r_o. However, the stability of the simulations was not affected by the location of the outer boundary. We use an upwind parameter $q=0.5$ and a constraint-term parameter $\mu=2$. We show runs for resolutions of $\Delta r=m/5$, $m/10$, $m/20$, and $m/40$, with $\Delta t=0.25\,\Delta r$. The run with $\Delta r=m/5$ has a resolution similar to those used in three-dimensional simulations of black hole collisions.

Figure 2 shows the L_2 norm of the Hamiltonian constraint (top) and the L_2 norm of the mass [Eq. (3.35)] error (bottom) taken at time $t=200m$ as a function of resolution. The convergence rate implied by the Hamiltonian constraint is 2.18, and that imphed by the mass error is 1.7. The convergence rate from the Hamiltonian constraint is larger than second order because we used third order accurate discretizations of

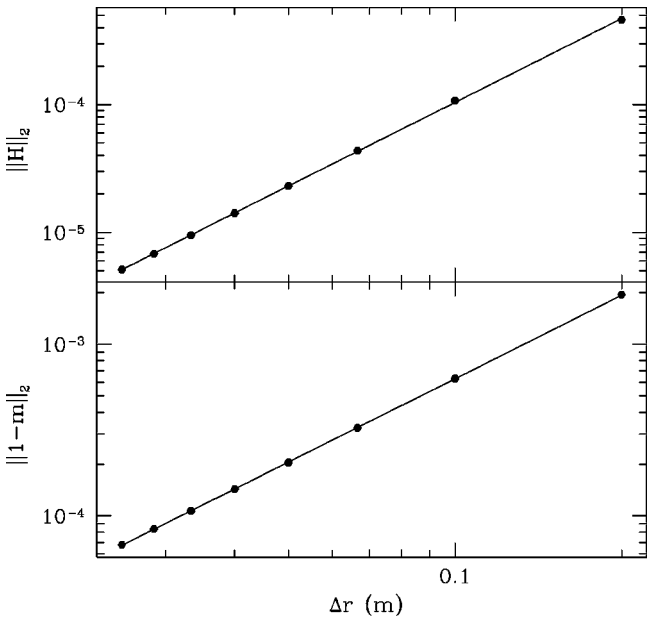

FIG. 2. L_2 norm of the Hamiltonian constraint (top) and L_2 norm of the mass error (bottom) as functions of resolution. The errors plotted were obtained at $t=200m$. The convergence rate implied by the Hamiltonian errors is 2.18, and that implied by the mass error 1.7.

the advection term as well as third order accurate extrapolations at the excision. On the other hand, the convergence rate obtained from the mass error is less than second order because the mass function is proportional to g^{rr}, a quantity difficult to handle numerically near the singularity. The reason for using the Hamiltonian constraint and mass to monitor accuracies and convergence is because of the gauge invariant nature of these quantities. Finally for reference, in Fig. 3 we show the L_2 norm of the solution Hamiltonian constraint as a function of time for runs with resolution $\Delta r=m/10$, varying the parameter μ. Different lines correspond to values of $\mu=0.0$, 0.025, 0.5, and 2 in order of stability improvement. It is clear from this figure the dramatic effect that the added constraint term has on the stability of the simulations.

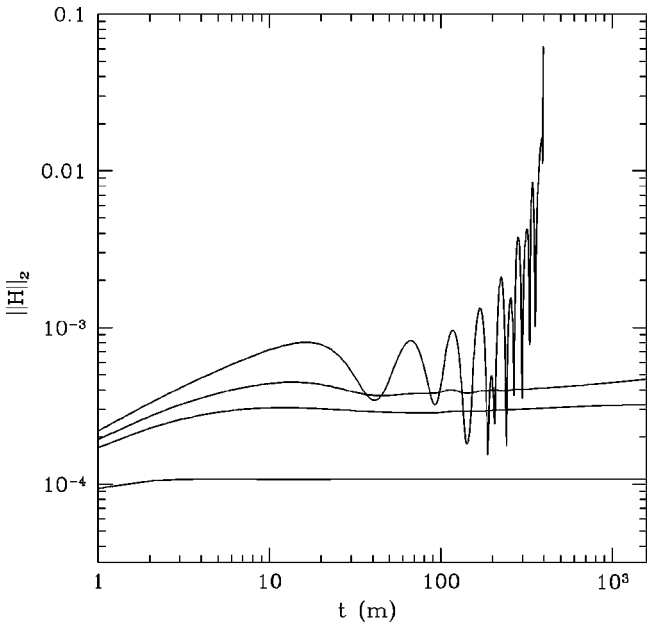

FIG. 3. L_2 norm of the Hamiltonian constraint as a function of time for $\Delta=m/10$ and $r_o=40m$. Each line corresponds to different values of the constraint-term parameter μ. The values of μ are 0.0, 0.025, 0.5, and 2 in order of stability improvement.

B. Ingoing null plus area locking

The ingoing null + area locking (IN+AL) recipe to specify the lapse and shift consists, in addition to locking the areal coordinate, of imposing the condition that the vector $\partial_t-\partial_r$ remains null throughout the evolution. This null condition is only compatible with the IEF case, since, by construction, the IEF solution is based on ingoing null observers. An analogous (ingoing timelike) condition can be obtained for the PG case. In terms of the space-time metric, the ingoing null condition is stated as $g_{tt}-2\,g_{tr}+g_{rr}=0$, or similarly in terms of 3+1 metric functions in Eq. (3.5) as

$$\alpha=a(1-\beta). \tag{6.15}$$

Conditions (6.15) and (6.1) yield

$$\alpha=\frac{a}{arK_b+1} \quad \text{and} \quad \beta=\frac{arK_b}{arK_b+1}, \tag{6.16}$$

where we have set $b=r$, since by construction b remains locked to r. This prescription for the lapse and shift was successfully applied in the past [30–32].

In order to make the ADM equations in the IN+AL gauge a first order in space system, we need to introduce two new variables $w\equiv\partial_r a$ and $y\equiv\partial_r K_b$. Even after this choice is made, there is no unique way of writing the resulting equations as a quasilinear system. The reason for this is the ambiguity one encounters when dealing with terms involving $\partial_r a$. One has the choice to either keep it as $\partial_r a$, substitute it with w, or a combination of both. Either choice changes the principal part of the equation. However, in our case, we need only find a choice that yields a strongly hyperbolic system. It turns out that the simplest choice of replacing $\partial_r a$ by w everywhere yields a system that is well-posed. That is, the resulting ADM equations have the form

$$\partial_t u=A\partial_r u+Bu, \tag{6.17}$$

with

$$u=\begin{pmatrix} a \\ K_a \\ K_b \\ y \\ w \end{pmatrix} \tag{6.18}$$

and

$$A=\frac{1}{z^2}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & z\,a\,r\,K_b & 0 & r & -a^{-2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & z\,a\,K_b & 0 & z+a^2r^2K_b^2 & a^{-1}K_b \\ 0 & -z\,a^2 & 0 & a^2\,r & a\,r\,K_b(1+z) \end{pmatrix}, \tag{6.19}$$

where we have introduced $z\equiv 1+a\,r\,K_b$ to simplify the notation. The corresponding eigenvalues and eigenvectors are

$$\lambda_1=0 \quad \text{with eigenvectors} \quad \vec{e}_1=[1,0,0,0,0],$$

$$\vec{e}_2=[0,0,1,0,0], \tag{6.20}$$

$$\lambda_2=1 \quad \text{with eigenvectors} \quad \vec{e}_3=[0,1,0,0,-z\,a^2],$$

$$\vec{e}_4=[0,0,0,1,a^2r], \tag{6.21}$$

$$\lambda_3=\frac{arK_b-1}{z} \quad \text{with eigenvector}$$

$$\vec{e}_5=[0,1,0,-aK_b,a^2]. \tag{6.22}$$

Note that all the eigenvectors are independent, and thus the system is strongly hyperbolic. Also note that the eigenvalues λ_2 and λ_3 are again characteristic speeds $(\beta\pm\alpha/a)$ along the light cone. Furthermore, in IEF coordinates, $\lambda_3=(2m-r)/(2m+r)$. Therefore, one encounters a situation similar to that of the EL+AL case; namely, at r_e all the eigenvalues are non-negative, and at r_0 one has that only $\lambda_3<0$. The existence, uniqueness, and well-posedness for IN+AL then follows as with the EL+AL case.

Regarding the constraints, their evolution is also described by a strongly hyperbolic system with characteristic speeds along the light cone. It is important to stress that our ADM equations already imply this, i.e., the evolution does not have any relation to making the evolution equations first order in space. In fact, the principal part of the evolution equations for the constraints is exactly the same as in the EL+AL gauge, but now the lapse and shift are given by Eq. (6.16). Thus the analysis and conclusions also follow as in that case.

An interesting aspect of the IN+AL choice is that it is possible to find a general solution to Einstein's equations. We start by defining $f\equiv r\,a\,K_b$. We then use momentum constraint (3.14) to eliminate K_a from the other equations. The outcome is that we need only to solve three of the four equations (3.9) and (3.11)–(3.13). We choose to work with the Hamiltonian constraint, [Eq. (3.13)] and Eqs. (3.9) and (3.12). The resulting system of equations reads

$$0=a(f^2+a^2-1)+2raf\partial_r f+2r(1-f^2)\partial_r a, \tag{6.23}$$

$$(\partial_t-\partial_r)a=\frac{a(f^2+a^2-1)}{2r(1+f)^2}, \tag{6.24}$$

$$(\partial_t-\partial_r)f=\frac{(f^2+a^2-1)}{r(1+f)}. \tag{6.25}$$

The general solutions of these equations are

$$f=\frac{2m}{r}-C\left(1-\frac{2\,m}{r}\right), \tag{6.26}$$

$$a^2=(1+C)(1+f), \tag{6.27}$$

where $C=C(t+r)$, and m is a constant. In terms of the 3+1 variables, the solutions read

$$a^2=(1+C)(1+f), \quad b^2=r^2, \quad K_a=\partial_r\left(\frac{f}{a}\right),$$

$$K_b=\frac{f}{ar}, \quad \alpha=\frac{a}{1+f}, \quad \beta=\frac{f}{1+f}, \tag{6.28}$$

with the line element [Eq. (3.5)] given explicitly by

$$ds^2=-(1+C)^2\left(1-\frac{2\,m}{r}\right)dt^2+2(1+C)\times\left[\frac{2m}{r}-\left(1-\frac{2m}{r}\right)C\right]dt\,dr+(1+C)\left[1+\frac{2m}{r}-\left(1-\frac{2m}{r}\right)C\right]dr^2+r^2d\Omega^2. \tag{6.29}$$

By setting $C=0$ one recovers the IEF solution [Eq. (3.21)]. Also, it is not difficult to show from the gauge invariant definition of mass [Eq. (3.35)] that the parameter m is indeed the mass of the black hole. An important property of general solution (6.29) is that it explicitly shows the residual gauge freedom associated with the IN+AL choice of lapse and shift. We have explicitly found the equivalence class of solutions gauge related to the IEF solutions that satisfy the IN+AL lapse-shift condition.

We repeat the same type of numerical experiment as with the EL+AL case; the same parameter values, resolutions, boundary conditions, and initial data. However, lapse and shift are constructed in this case from Eq. (6.16). Figure 4 shows the L_2 norm of the Hamiltonian constraint as a function of time for different resolutions. Figure 5 shows the L_2 norms of the Hamiltonian constraint (top) and mass error (bottom) for different resolutions. The convergence rates are similar to those in EL+AL, namely 2.19 from the Hamiltonian constraint and 1.7 from the mass error. Finally, Fig. 6 shows the L_2 norm of the Hamiltonian constraint as a function of time for a resolution of $\Delta r=m/10$. Different lines correspond to values of $\mu=0.0$, 0.025, 0.5, and 2 in order of stability improvement. It is clear from these results that the stability behavior of the system of equations under the IN+AL gauge choice closely follows that of EL+AL gauge choice.

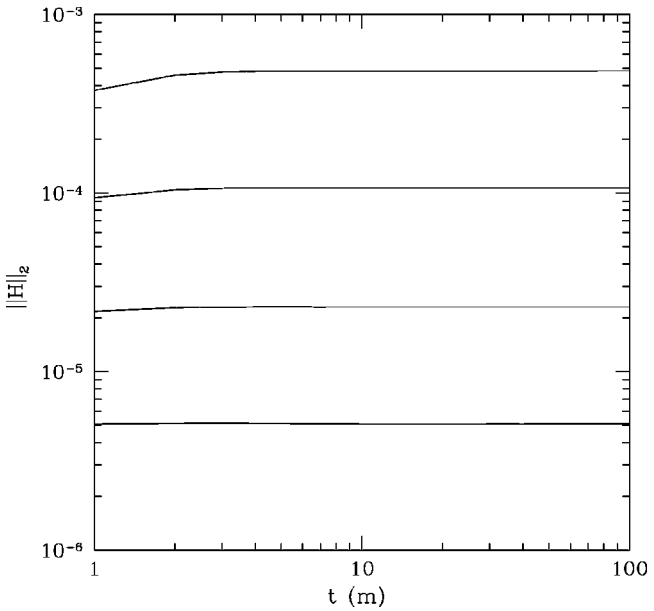

FIG. 4. Same as in Fig. 1 but for the IN+AL case.

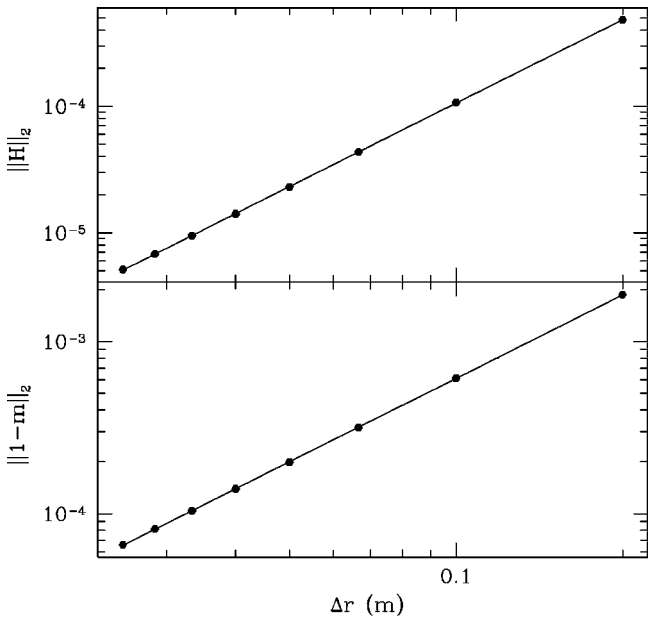

FIG. 5. Same as in Fig. 2 but for the IN+AL case.

C. Exact lapse plus exact shift

Finally, we consider the case in which the lapse function and shift vector are prescribed by exact analytic solutions. To date, with the standard ADM system, it has not been possible to obtain long-term stable and convergent numerical evolutions under the EL+ES choice. Before we present results of EL+ES evolutions using the adjusted ADM system of equations, let us investigate its quasi-well-posedness properties. To make the evolution equations first order in space, we add a new variable $y=\partial_r b$. Once more, the resulting set of equations has the form

$$\partial_t u = A\partial_r u + Bu, \tag{6.30}$$

with

$$u=\begin{pmatrix} a \\ b \\ K_a \\ K_b \\ y \end{pmatrix} \tag{6.31}$$

and

$$A=\begin{pmatrix} \beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \dfrac{b\partial_r\alpha+2(1-\mu)\alpha y}{a^3 b} & 0 & \beta & 0 & -\dfrac{2(1-\mu)\alpha}{a^2 b} \\ \dfrac{\alpha y}{b a^3} & 0 & 0 & \beta & -\dfrac{\alpha}{a^2 b} \\ 0 & 0 & 0 & -\alpha b & \beta \end{pmatrix}. \tag{6.32}$$

The matrix A has eigenvalues

$$\lambda_1=\beta, \tag{6.33}$$

$$\lambda_2=\beta, \tag{6.34}$$

$$\lambda_3=0, \tag{6.35}$$

$$\lambda_4=\beta+\alpha/a, \tag{6.36}$$

$$\lambda_5=\beta-\alpha/a. \tag{6.37}$$

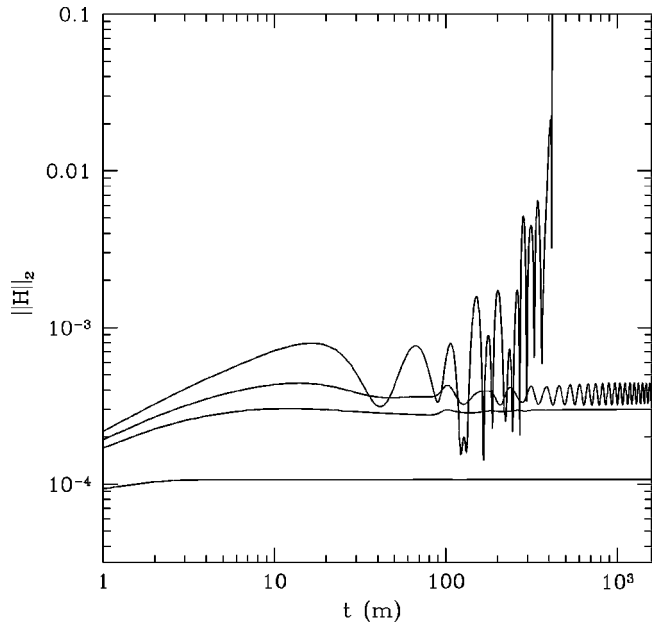

FIG. 6. Same as in Fig. 3 but for the IN+AL case.

There are two fields that propagate with a characteristic speed (β) along the normal to the hypersurfaces in the foliation, one field with zero speed, and two other fields with characteristic speeds ($\beta \pm \alpha/a$) along the light cone. It is not difficult to show that eigenvectors corresponding to the eigenvalue $\lambda_1 = \lambda_2 = \beta$ are not distinct, whatever the value of μ. Therefore, the system of equations is only weakly hyperbolic and thus not quasi-well-posed, as we have already seen by considering the EL+ES gauge perturbations by themselves in Sec. V.

As in the previous two cases, the constraints propagate according to

$$\partial_t v = P \partial_r v + Q v, \tag{6.38}$$

where now

$$v = \begin{pmatrix} H \\ M \end{pmatrix} \tag{6.39}$$

and

$$P = \begin{pmatrix} \beta & \dfrac{4\alpha}{a^2} \\ \dfrac{\alpha}{4} & \beta \end{pmatrix}. \tag{6.40}$$

This is, we have the same situation as before, namely propagation of the constraints along with characteristic speeds along the light cone.

We have carried out the same numerical experiments as for the previous two cases. However, for the IEF solution, the simulations with computational domains for $r_o \geqslant 40m$ crashed. We have been able to identify two possible sources behind this problem. One of them is the pulse originated at the outer boundary due to discontinuities of the truncation error. This pulse propagates in the direction of the black hole (i.e., the decreasing r coordinate) with a characteristic speed β. As the pulse travels its amplitude grows in time. This effect is shown in Fig. 7. Here we plot the solution error for the metric function a as a function of space and time for a resolution $\Delta r = m/10$ and outer boundary located at $r_o = 40m$. The simulation stops because this pulse error increases to the point that the metric function a becomes negative. For small computational domains ($r_o < 40m$), the crossing time of this pulse is short enough and does not allow a catastrophic growth of the pulse. The pulse is able to leave the computational domain through the excision boundary without crashing the simulation. Since the initial amplitude of this pulse is $O(\Delta r^2)$, i.e., the accuracy of the discretization, in principle one could find a resolution small enough such that the growth of the pulse would not affect the life of the simulation. However, accessing those fine resolutions in three-dimensional simulations is likely to be impractical. The second, and perhaps more severe, source of the problem is the presence of the zero velocity mode. Zero velocity modes were in principle also allowed in the case of IN+AL; see Eq. (6.20). However, the adjusted IN+AL system yields only numerical solutions of the general form of Eq. (6.29), which clearly does not contain a zero velocity mode. If $\mu = 0$, these

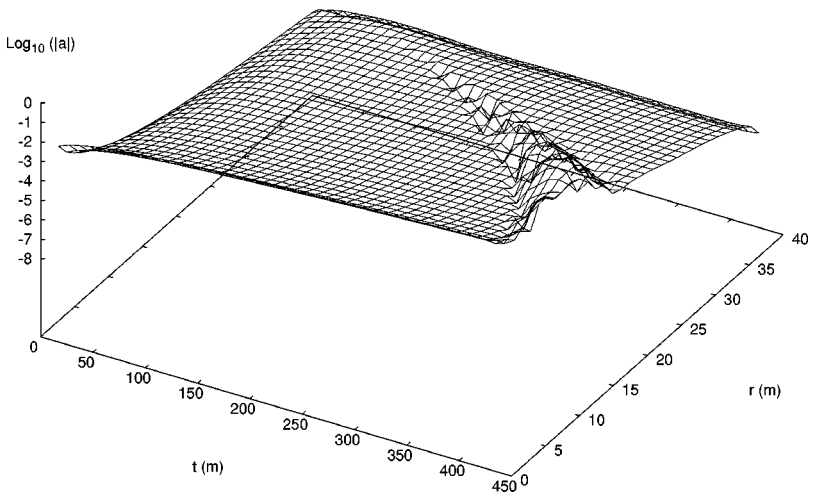

FIG. 7. Solution error for the metric function a in IEF coordinates for the EL+ES case. The resolution is $\Delta r = m/10$, and the outer boundary is located at $r_o = 40\,m$. A pulse originated at the location of the outer boundary, due to discontinuities in the truncation error, propagates in the direction of the black hole (decreasing r coordinate). The simulation stops because this pulse-error increases to the point that the metric function a becomes negative.

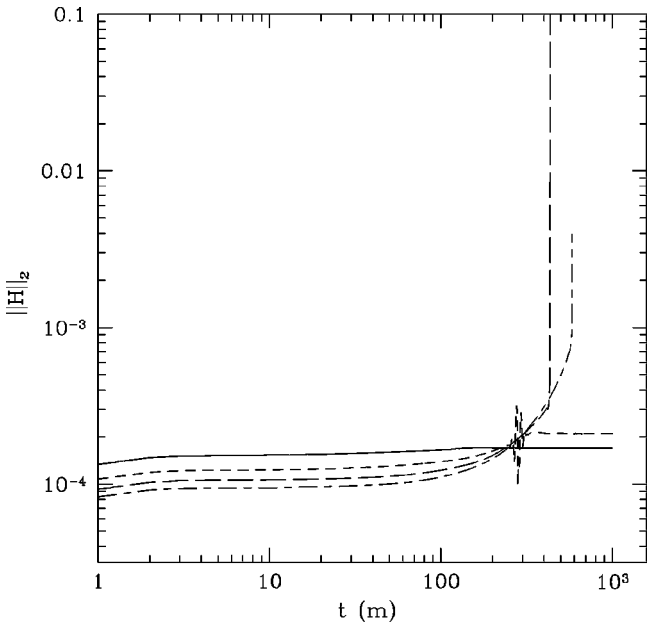

FIG. 8. L_2 norm of the Hamiltonian constraint in IEF coordinates for the EL+ES case as a function of time. The resolutions are $\Delta r = m/10$ and $\mu = 2$. Each line corresponds to different location of the outer boundary: $r_o = 20m$ (solid line), $30m$ (short dashed line), $40m$ (long dashed line), $50m$ (short dashed–long dashed line).

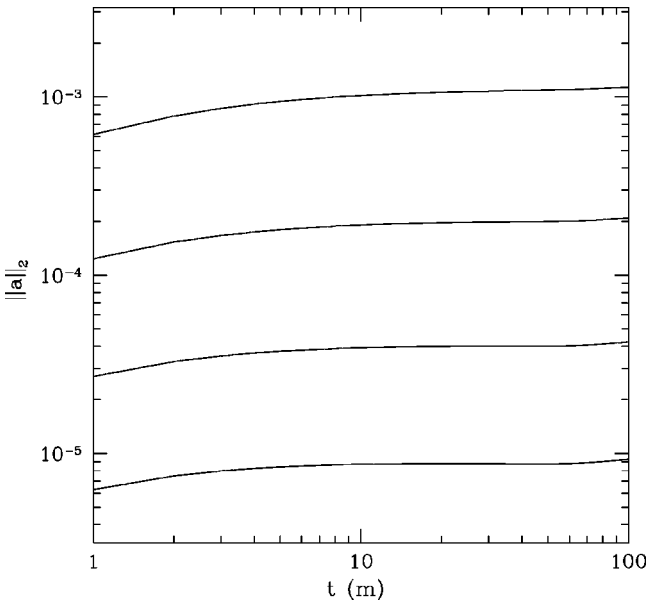

FIG. 9. L_2 norm of the solution error in PG coordinates of the metric function a as a function of time. Each line correspond to different resolutions: $\Delta r = m/5$, $m/10$, $m/20$, and $m/40$, respectively, from top to bottom.

zero velocity modes are not suppressed and eventually terminate the simulation. This catastrophic effect induced by zero velocity modes was previously noted by Alcubierre *et al.* [13].

Figure 8 shows the L_2 norm of the Hamiltonian constraint as a function of time with resolution $\Delta r = m/10$ for different locations of the outer boundary. For computational domains with approximately $r_o \leqslant 40m$, the zero velocity mode is still present, but is eventually damped. The reasons why this mode stops growing remain unclear. Nonetheless, there is a strong indication that this behavior is connected to the particular choice of coordinates used to set the exact lapse and shift. If, instead of the IEF solution, one sets the lapse and shift from the PG solution, the outcome of the simulations is completely different. EL+ES simulations with a PG lapse and shift are long term stable and convergent for arbitrary sizes of the computational domain as long as the system of equations was adjusted with $\mu \geqslant 1$. Figure 9 shows the L_2 norm of the solution error for the metric function a as a function of time for different resolutions with $r_o = 40m$ for EL+ES in PG coordinates.

VII. CONCLUSIONS

We have demonstrated that, at least for the case of single black hole space-times in spherical symmetry, it is possible to obtain long-term stable and convergent numerical simulations using the standard ADM system of equations if the equations are adjusted by introducing terms involving the constraints. Results were presented for three choices of lapse and shift. In addition, we introduced the concept of quasi-well-posedness, which appears to be useful in characterizing the properties of the system of evolution equations. We are currently investigating the extension of this approach to three-dimensional evolutions.

ACKNOWLEDGMENTS

This work was supported by NSF PHY9800973, NSF PHY0090091, and by Eberly research funds of the Pennsylvania State University. We wish to thank D. Arnold, A. Ashtekar, L. Lehner, R. Matzner, R. Price, O. Reula, H. Shinkai, and B. Whiting for comments and helpful discussions.

[1] J. York, in *Sources of Gravitational Radiation*, edited by L. Smarr (Cambridge University Press, Cambridge, England, 1979).
[2] M.A. Scheel, T.W. Baumgarte, G.B. Cook, S.L. Shapiro, and S.A. Teukolsky, Phys. Rev. D **58**, 044020 (1998).
[3] H. Friedrich, Proc. R. Soc. London, Ser. A **375**, 169 (1981); **378**, 401 (1981).
[4] Y. Choquet-Bruhat and J.W. York, C. R. Acad. Sci., Ser. I: Math. **321**, 1089 (1995).
[5] S. Frittelli and O. Reula, Phys. Rev. Lett. **76**, 4667 (1996).
[6] C. Bona, J. Masso, E. Seidel, and J. Stela, Phys. Rev. D **56**, 3405 (1997).
[7] A. Anderson and J. York, Phys. Rev. Lett. **82**, 4384 (1999).
[8] For reviews on hyperbolic methods for Einstein's equations,

see H. Friedrich and A. Rendall, in *Einstein's Field Equations and their Physical Interpretation*, edited by B.G. Schmidt (Springer, Berlin, 2000), gr-qc/000207. Also O. Reula, ''Living Reviews in Relativity,'' http://www.livingreviews.org.
[9] M. Shibata and T. Nakamura, Phys. Rev. D **52**, 5428 (1995).
[10] T. Baumgarte and S. Shapiro, Phys. Rev. D **59**, 024007 (1999).
[11] M. Scheel (private communication).
[12] R. Arnowitt, S. Deser, and C.W. Misner, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962), pp. 227–265.
[13] M. Alcubierre, G. Allen, B. Bruegman, E. Seidel, and W.M. Suen, Phys. Rev. D **62**, 124011 (2000).
[14] M. Alcubierre, G. Allen, B. Bruegman, T. Dramlitsch, J.A. Font, P. Papadopoulos, E. Seidel, N. Stergioulas, W.M. Suen, and R. Takahashi, Phys. Rev. D **62**, 044034 (2000).
[15] M. Miller, gr-qc/0008017.
[16] O. Brodbeck, S. Frittelli, P. Huebner, and O.A. Reula, J. Math. Phys. **40**, 909 (1999).
[17] G. Yoneda and H. Shinkai, Phys. Rev. D **63**, 124019 (2001).
[18] B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods* (Wiley, New York, 1995)
[19] C.A.J. Fletcher, *Computational Techniques for Fluid Dynamics* (Springer-Verlag, Berlin 1991), Vol. I, p. 335.
[20] S.A. Teukolsky, Phys. Rev. D **61**, 087501 (2000).
[21] L. Rezzolla, A. Abrahams, R. Matzner, M. Rupright, and S. Shapiro, Phys. Rev. D **59**, 064001 (1999).
[22] M. Alcubierre and B. Bruegman, Phys. Rev. D **63**, 104006 (2001).
[23] P. Painlevé, C. R. Hebd. Seances Acad. Sci. **173**, 667 (1921).
[24] A. Gullstrand, Ark. Mat., Astron. Fys. **15(6)**, 1 (1922).
[25] C. Misner, K.S. Thorne, and J. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[26] K. Martel and E. Poisson, Am. J. Phys. **69**, 476 (2001).
[27] L. Lehner, M. Huq, M. Anderson, E. Bonning, D. Schaefer, and R. Matzner, Phys. Rev. D **62**, 044037 (2000).
[28] T.W. Baumgarte, S.A. Hughes, and S.L. Shapiro, Phys. Rev. D **60**, 087501 (1999).
[29] S. Frittelli, Phys. Rev. D **55**, 5992 (1997).
[30] R.L. Marsa and M.W. Choptuik, Phys. Rev. D **54**, 4929 (1996).
[31] J. Thornburgh, gr-qc/9906022.
[32] M. Iriondo and O. Reula, gr-qc/0102027.