Numerical tests of evolution systems, gauge conditions, and boundary conditions for 1D colliding gravitational plane waves

J. M. Bardeen
Physics Department, University of Washington, Seattle, Washington 98195

L. T. Buchman
Astronomy Department, University of Washington, Seattle, Washington 98195
(Received 17 October 2001; published 5 March 2002)

We investigate how the accuracy and stability of numerical relativity simulations of 1D colliding plane waves depends on choices of equation formulations, gauge conditions, boundary conditions, and numerical methods, all in the context of a first-order $3+1$ approach to the Einstein equations, with basic variables some combination of first derivatives of the spatial metric and components of the extrinsic curvature tensor. Hyperbolic schemes, specifically variations on schemes proposed by Bona and Massó and Anderson and York, are compared with variations of the Arnowitt-Deser-Misner formulation. Modifications of the three basic schemes include raising one index in the metric derivative and extrinsic curvature variables and adding a multiple of the energy constraint to the extrinsic curvature evolution equations. Redundant variables in the Bona-Massó formulation may be reset frequently or allowed to evolve freely. Gauge conditions which simplify the dynamical structure of the system are imposed during each time step, but the lapse and shift are reset periodically to control the evolution of the spacetime slicing and the longitudinal part of the metric. We show that physically correct boundary conditions, satisfying the energy and momentum constraint equations, generically require the presence of some ingoing eigenmodes of the characteristic matrix. Numerical methods are developed for the hyperbolic systems based on decomposing flux differences into linear combinations of eigenvectors of the characteristic matrix. These methods are shown to be second-order accurate, and in practice second-order convergent, for smooth solutions, even when the eigenvectors and eigenvalues of the characteristic matrix are spatially varying.

DOI: 10.1103/PhysRevD.65.064037 PACS number(s): 04.25.Dm

I. INTRODUCTION

The goal of projects such as the ground-based Laser Interferometric Gravitational Wave Observatory (LIGO) and the space-based Laser Interferometer Space Antenna (LISA) is to detect gravitational waves, and to use them as a new observational window for relativistic astrophysics. A primary source for these gravitational waves is the coalescence of binary black holes [1]. The highly nonlinear and dynamical merger phase of this coalescence process can only be calculated by numerical relativity, and obtaining merger gravitational waveforms, both for theoretical understanding and for detection, is dependent on long-term stable and accurate numerical evolutions. A worldwide collaboration of numerical relativists, physicists, mathematicians, and computer scientists has devoted considerable effort over the last 20 years to develop 3D codes to calculate black hole merger gravitational waveforms, and significant progress has been made, especially in the last few years. However, more groundwork is required before calculations of 3D binary black hole merger templates for a variety of scenarios can be completed. Greater understanding of equation formulations, boundary conditions, and dynamic gauge conditions, and the use of advanced numerical methods, is essential to achieve this goal. We believe that an important foundation for this understanding is extensive testing and analysis in 1D and 2D. Choptuik's discovery of black hole critical phenomena in spherically symmetric gravitational collapse [2] is an example of the potential of careful numerical work in 1D.

This paper reports the methodology, results, and analysis of calculations of 1D nonlinear colliding gravitational plane-wave spacetimes. We have chosen to investigate hyperbolic formulations of the Einstein equations, as they are well-posed, they can be treated with advanced numerical methods, and they can help in the analysis of boundary conditions [3,4]. We call a set of equations *hyperbolic* if the characteristic matrix can be diagonalized with a complete set of eigenvectors and real eigenvalues, following LeVeque [5]. This is called *strongly hyperbolic* [6] in much of the literature. The lapse and the shift are evolved during each time step in a manner which is consistent with a simple hyperbolic scheme. Between time steps, the lapse and the shift are reset according to conditions which are unconstrained by the need to preserve hyperbolicity. In this way, the evolution of the hypersurfaces and spatial coordinates can be controlled to prevent large gradients, coordinate pathologies, and instabilities. Some of the redundant variables of a hyperbolic formulation can also be reset between time steps. This resetting can have positive or negative effects on accuracy and stability, depending on the eigenmode structure of the reset system. Finally, we find that boundary conditions should not be based naively on the eigenmodes of the hyperbolic decomposition for two reasons: (a) satisfying the constraint equations at the boundaries generically requires the presence of incoming eigenmodes, and (b) even whether the ''physical'' eigenmodes are purely outgoing at the boundaries is gauge dependent.

Many ways of formulating evolution equations for the spatial metric in Einstein's theory of general relativity are

possible. The most thoroughly tested formulation in numerical relativity is the Arnowitt-Deser-Misner (ADM) set of equations [7]. The standard ADM equations in vacuum are

$$(\partial_t - \mathcal{L}_\beta) h_{ij} = -2\alpha K_{ij}, \tag{1}$$

$$(\partial_t - \mathcal{L}_\beta) K_{ij} = -\alpha_{|ij} + \alpha[{}^{(3)}R_{ij} + K K_{ij} - 2K_i{}^l K_{lj}], \tag{2}$$

$$(\partial_t - \mathcal{L}_\beta) K^i{}_j = -\alpha^{|i}{}_{|j} + \alpha[{}^{(3)}R^i{}_j + K K^i{}_j]. \tag{3}$$

In these equations, K_{ij} is the extrinsic curvature, $K = K^l{}_l$, β^i is the shift, α is the lapse, h_{ij} is the 3-metric, and ${}^{(3)}R_{ij}$ is the 3D Ricci tensor. The vertical bar represents a covariant derivative taken with respect to the 3-geometry. Equation (3) evolves what we call the ''mixed'' form of the extrinsic curvature tensor. The energy and momentum constraint equations are, respectively,

$$\mathcal{E} = 1/2[{}^{(3)}R - K_i{}^j K^i{}_j + K^2] = 0, \tag{4}$$

$$\mathcal{M}_i = K_i{}^j{}_{|j} - K_{|i} = 0. \tag{5}$$

While successful calculations using the ADM formulation have been done in 2D, 3D calculations generally crash after just a few dynamical times.

Alternative formalisms include many versions of hyperbolic systems, which add redundant variables and/or constraint terms to the equations to allow a complete set of eigenmodes describing evolution along the characteristics. As indicated by Reula [6], there are an infinite number of hyperbolic formulations. We focus on variations of relatively simple schemes proposed by Bona-Massó (BM) [8] and Anderson-York (AY) [9], in which the characteristics propagate either at local light speed or along the hypersurface normals, and in which the variables include first derivatives of the metric.

Initial attempts at using hyperbolic methods in 3D were based on the BM formulation [4], but did not use numerical methods which take advantage of the eigenfields of the system. These codes were not much more successful than ADM. Nonhyperbolic Baumgarte-Shapiro-Shibata-Nakamura (BSSN) schemes [10,11], based on conformal decomposition of the metric, have shown considerable success in improving the stability of 3D calculations for weak and strong gravitational fields and a variety of spacetime slicings [12]. Alcubierre *et al.* [13] report that a BSSN scheme, combined with excision and certain dynamic gauge conditions, allows accurate numerical evolutions of 3D distorted dynamic black holes up to hundreds of dynamical times.

In the context of considering only first derivative variables, a great variety of hyperbolic schemes have been proposed that involve adding constraint terms to the equations [14–19]. Kidder, Scheel, and Teukolsky [19] examine a rather general class of such schemes, which include the AY [9] and Frittelli-Reula [14] formulations as special cases. Among these schemes are some which allow for long-term evolution of a Schwarschild black hole in 3D.

In this paper, we explore ways of using hyperbolic methods that combine superior accuracy with gauge conditions which maintain stability at least for the limited dynamical times we can explore with plane waves. Three basic first order systems are studied: BM, AY, and ADM. Hyperbolicity is obtained in our BM and AY formulations by adding momentum constraint terms to the ADM equations, as in the standard formulations. The BM, AY, and ADM formulations are modified by using ''mixed'' forms (with one index raised) of the extrinsic curvature and metric derivatives as variables. The BM formulations are further modified by resetting redundant variables, which gives an overall ADM-like evolution. Further, all our formulations are varied by adding a multiple of the energy constraint to the evolution equations for the extrinsic curvature. Specifically, $-n\alpha\mathcal{E}h_{ij}/2$ is added to Eq. (2), and $-n\alpha\mathcal{E}\delta^i{}_j/2$ is added to Eq. (3), where n, the energy constraint coefficient, is an arbitrary real number. The ADM formulation is actually hyperbolic as long as the longitudinal-transverse components of the metric and extrinsic curvature can be assumed to vanish identically, and $n<0$ or $0<n<1$. Comparisons of results from our various ADM, BM, and AY calculations allow us to identify and analyze aspects of equation formulation which significantly improve accuracy and/or stability. These are mixed variables, a separation of the constraint error speeds from the other characteristic speeds of the system, and maintaining long-term effective hyperbolicity (taking into account resetting of redundant variables, but ignoring deviation from strict hyperbolicity due to resetting the lapse and the shift).

While the same energy constraint terms as specified above are present in the standard BM formulation, numerical implementations have been carried out, as far as we are aware, only for $n=0$ (the Ricci evolution system) and $n=1$ (the Einstein evolution system). Adding these energy constraint terms to the AY formulation is a special case of the more general Kidder, Scheel, and Teukolsky [19] schemes. Shinkai and Yoneda [15–17] analyzed the stability and accuracy properties of first order hyperbolic systems using Ashtekar's connection variables in plane-symmetric spacetimes, and found that the addition of multiples of the constraints to the dynamical equations improved accuracy and stability. These results have been extended to ADM systems of equations [18].

Gauge choices in most previous implementations of hyperbolic formulations have been limited in order to preserve the hyperbolicity of the system. Since no time derivatives of the lapse and the shift occur in the dynamical equations for the other variables, the lapse and the shift can be reset arbitrarily at any time during the numerical evolution, as pointed out by Balakrishna *et al.* [20]. Our gauge evolution maintains strict hyperbolicity during each time step, but the lapse and shift are reset periodically between time steps in order to control the long-term evolution of the coordinate system. The resetting may be accomplished by imposing algebraic conditions, by solving elliptic equations, or by evolving the lapse and/or shift through dynamical equations implemented independently of the main hyperbolic system.

Poor boundary conditions can result in the introduction of instabilities or inaccuracies into the numerical grid. In numerical relativity, boundary conditions have usually been rather crudely implemented. Some sort of outgoing radiation

conditions are imposed on all components of the metric, or boundary conditions are based on an analytic exterior solution [12,21]. One attraction of hyperbolic methods has been the possibility of basing boundary conditions on the eigenmodes of the characteristic matrix. However, it is clear from our plane-wave calculations that, particularly for the ''non-physical'' eigenmodes involving the non-transverse-traceless parts of the metric, making the amplitudes of the incoming eigenmodes at the boundaries zero can lead to serious violations of the energy and momentum constraints. Furthermore, what constitutes an incoming eigenmode is dependent on the formulation of the equations as well as on gauge conditions. Even imposing purely outgoing boundary conditions on the ''physical'' eigenmodes of the hyperbolic system is not strictly correct, as nonlinear coupling between the ''physical'' and ''non-physical'' eigenmodes in the source terms can generate a gauge-dependent admixture of outgoing and incoming ''physical'' eigenmodes. Our boundary conditions are based on quadratic extrapolation of the variables from inside the grid to the first ghost cells on either side of the grid. The ghost cell values are then corrected to make sure the constraint equations are satisfied on the boundaries. For 1D plane waves, projection of the Weyl tensor onto a null tetrad gives a gauge-independent measure of the left and right-going components of the gravitational radiation. Our numerical solutions for colliding plane waves show that as the wave packets leave the grid, the incoming components of the Weyl tensor are in fact zero even though there are non-zero incoming ''physical'' eigenmodes of the characteristic matrix.

Our focus in applying hyperbolic methods to the Einstein equations is on achieving second order accuracy for smooth solutions, when the eigenvectors and eigenvalues of the system are a function of position. Finite difference methods such as MacCormack, Lax-Wendroff, and staggered leapfrog [22], which are often used in numerical relativity, give good second order accuracy for smooth solutions, but standard wave propagation algorithms for hyperbolic systems as presented by LeVeque [23] are not second order accurate when the eigenvectors and eigenvalues are spatially varying. LeVeque suggested a new wave propagation method [24] for variable coefficient flux problems which we develop and apply to our 1D nonlinear gravitational plane-wave calculations. We show in the Appendix that the new methods are formally second order accurate even with varying eigenvectors and eigenvalues, and verify second order convergence in our numerical results.

II. EVOLUTION EQUATIONS

The most general spatial metric for a nonlinear 1D plane wave traveling in the x-direction is

$$ds^2 = h_{xx}dx^2 + h_{yy}dy^2 + h_{zz}dz^2 + 2h_{yz}dydz, \tag{6}$$

in which h_{xx}, h_{yy}, h_{zz}, and h_{yz} are functions of x alone. We will restrict our discussion to a diagonal metric in this paper.

The standard ADM evolution equations are first order in time and second order in space. Most hyperbolic formalisms are first order in space and time, and incorporate first derivatives of the spatial metric as additional variables. The derivative variables as defined by BM are

$$D_{kij} = \frac{1}{2}\partial_k h_{ij}. \tag{7}$$

In our applications with a diagonal metric, we find that switching to the ''mixed'' variables

$$D_{ki}{}^{j} = \frac{1}{2}(\partial_k h_{il})h^{lj}, \quad K_i{}^{j} = K_{il}h^{lj} \tag{8}$$

from the ''lowered'' variables, D_{kij} and K_{ij}, improves accuracy significantly without other complications. However, with a non-diagonal metric, $D_{ki}{}^{j}$ and $K_i{}^{j}$ are not symmetric in i and j, and the evolution equations for $D_{ki}{}^{j}$ acquire complicated source terms.

Below, we present the first order evolution equations for 1D plane waves travelling along the x-direction and described by a diagonal spatial metric, using our mixed variables. The equations with D_{kij} and K_{ij} as variables are given in the BM papers [8]. A few points need to be made about notation. First, since our 1D problem involves derivatives only in the x-direction, we simplify our notation $D_{xi}{}^{j} \to D_i{}^{j}$. Second, a prime indicates a spatial derivative with respect to x. Third, our symbol for the shift is simply β instead of β^i. We suppress the index on the shift because there is only one non-zero component in this 1D case.

The evolution equations for h_{ij} are obtained from the definition of the extrinsic curvature of the hypersurfaces, Eq. (1), and are

$$\partial_t h_{xx} = 2h_{xx}[\beta D_x{}^{x} + \beta' - \alpha K_x{}^{x}], \tag{9}$$

$$\partial_t h_{yy} = 2h_{yy}[\beta D_y{}^{y} - \alpha K_y{}^{y}],$$

$$\partial_t h_{zz} = 2h_{zz}[\beta D_z{}^{z} - \alpha K_z{}^{z}]. \tag{10}$$

The evolution equations for $D_i{}^{j}$ are obtained by taking the time derivative of $D_{ki}{}^{j}$ in Eq. (8), and interchanging space and time derivatives. The resulting equations are

$$\partial_t D_x{}^{x} + \partial_x[-\beta D_x{}^{x} - \beta' + \alpha K_x{}^{x}] = 0, \tag{11}$$

$$\partial_t D_y{}^{y} + \partial_x[-\beta D_y{}^{y} + \alpha K_y{}^{y}] = 0,$$

$$\partial_t D_z{}^{z} + \partial_x[-\beta D_z{}^{z} + \alpha K_z{}^{z}] = 0. \tag{12}$$

The $K_i{}^{j}$ variables are evolved from the Einstein equations, Eq. (3). We include the addition of an arbitrary multiple, n, of the energy constraint in these equations. After organization into a conservation law form, the $K_i{}^{j}$ evolution equations are

$$\partial_t K_x{}^x + \partial_x\left[-\beta K_x{}^x + \frac{\alpha}{h_{xx}}\left(\frac{\alpha'}{\alpha} + D_y{}^y + D_z{}^z\right)\right]$$
$$= -\beta' K_x{}^x + \alpha\left\{ K_x{}^x K_l{}^l + \frac{1}{h_{xx}}\left[\frac{\alpha'}{\alpha}(D_y{}^y + D_z{}^z) - D_y{}^y D_y{}^y - D_z{}^z D_z{}^z - D_x{}^x\left(\frac{\alpha'}{\alpha} + D_y{}^y + D_z{}^z\right)\right]\right\} - \frac{n}{2}\alpha\mathcal{E}, \tag{13}$$

$$\partial_t K_y{}^y + \partial_x\left[-\beta K_y{}^y + \frac{\alpha}{h_{xx}} D_y{}^y\right]$$
$$= -\beta' K_y{}^y + \alpha\left[K_y{}^y K_l{}^l - \frac{D_y{}^y D_l{}^l}{h_{xx}}\right] - \frac{n}{2}\alpha\mathcal{E}, \tag{14a}$$

$$\partial_t K_z{}^z + \partial_x\left[-\beta K_z{}^z + \frac{\alpha}{h_{xx}} D_z{}^z\right]$$
$$= -\beta' K_z{}^z + \alpha\left[K_z{}^z K_l{}^l - \frac{D_z{}^z D_l{}^l}{h_{xx}}\right] - \frac{n}{2}\alpha\mathcal{E}, \tag{14b}$$

where we write $\alpha\mathcal{E}$ so that the division between the flux terms and the source terms is apparent:

$$\alpha\mathcal{E} = -\partial_x\left[\frac{\alpha}{h_{xx}}(D_y{}^y + D_z{}^z)\right] + \alpha[K_x{}^x(K_y{}^y + K_z{}^z) + K_y{}^y K_z{}^z]$$
$$+ \frac{\alpha}{h_{xx}}\left[\left(\frac{\alpha'}{\alpha} - D_x{}^x\right)(D_y{}^y + D_z{}^z) - (D_y{}^y D_y{}^y + D_y{}^y D_z{}^z + D_z{}^z D_z{}^z)\right]. \tag{15}$$

III. GAUGE EVOLUTION

We let the lapse and the shift evolve during each time step according to a prescription which simplifies the hyperbolic system, and we periodically reset the lapse and the shift between time steps to control the longer term evolution of the coordinates and to keep gauge pathologies from developing. We defer discussion of resetting gauge conditions to Sec. V. Here, we discuss how the gauge evolves between resettings.

For a hyperbolic formulation of the equations, the natural choice for the lapse between resettings is the Choquet-Bruhat algebraic gauge condition [25,26], because it simplifies the fluxes and source terms in the hyperbolic system of equations considerably. This gauge condition is

$$\alpha = Q\sqrt{\det(h_{ij})}, \tag{16}$$

where Q is a specified function of x,t.

We vary the Choquet-Bruhat algebraic gauge condition by making Q and Q' (which equals $\partial_x Q$) variables in the hyperbolic system, rather than specified functions of x,t. We choose Q and Q' as variables so that Q' can be included in the flux of $K_x{}^x$ as part of the hyperbolic system. Otherwise, Q'' would have to be considered part of the source of $K_x{}^x$, and evaluating Q'' from the lapse involves second derivatives of the metric. By advecting Q and Q' along the hypersurface normals, we incorporate them into the hyperbolic system in a consistent way. Our advection equations are

$$\partial_t Q - \beta Q' = 0, \quad \partial_t D_Q - \partial_x[\beta D_Q] = 0, \tag{17}$$

where $D_Q = Q'/Q$. Our advection of Q corresponds to harmonic slicing [25].

There is a danger with resetting the lapse and the shift, in that fluctuations in β' and D_Q can feed back on one another through the evolution equations for $D_x{}^x$ and $K_x{}^x$. The resetting gauge conditions of Sec. V imply that a fluctuation in $K_x{}^x$ is balanced by a fluctuation in β', and a fluctuation in $D_x{}^x$ is balanced by a fluctuation in D_Q. For certain time intervals between resetting, if these fluctuations propagate at different speeds, they may drift in such a way that they reinforce rather than cancel over much of the time interval. Although the standard procedure is to keep the shift constant in hyperbolic schemes, we find that if we advect D_Q with β' constant, such a positive feedback can occur, resulting in a runaway instability. However, if we advect both D_Q and β' along hypersurface normals, the evolution is stable. Our advection equations for β and β' are

$$\partial_t \beta - \beta\beta' = 0, \quad \partial_t \beta' - \partial_x[\beta\beta'] = 0. \tag{18}$$

IV. CONSTRAINT EQUATIONS

The energy and momentum constraints must be satisfied by the initial conditions and throughout the evolution. We use these constraints to obtain the initial conditions. We do not impose the constraints during the evolution of the dynamical equations. However, we do insure that the boundary conditions are consistent with the constraint equations, and we use the constraints to check for accuracy and convergence as the numerical evolution proceeds. The energy and momentum constraint equations are, respectively,

$$\mathcal{E} = -\partial_x\left[\frac{1}{h_{xx}}(D_y{}^y + D_z{}^z)\right]$$
$$- \frac{1}{h_{xx}}[D_y{}^y D_y{}^y + D_y{}^y D_z{}^z + D_z{}^z D_z{}^z + D_x{}^x(D_y{}^y + D_z{}^z)]$$
$$+ K_x{}^x(K_y{}^y + K_z{}^z) + K_y{}^y K_z{}^z$$
$$= 0, \tag{19}$$

$$\mathcal{M}_x = -\partial_x(K_y{}^y + K_z{}^z) - D_y{}^y K_y{}^y - D_z{}^z K_z{}^z + (D_y{}^y + D_z{}^z)K_x{}^x$$
$$= 0. \tag{20}$$

V. RESETTING GAUGE CONDITIONS

The lapse and shift are periodically reset between time steps in order to implement a dynamic spacetime slicing which is unconstrained by the need to maintain a hyperbolic system. Our resetting gauge conditions are chosen to prevent pathologies and/or strong gradients from developing in the hypersurfaces and spatial coordinates, and to help stability properties at the boundaries of the grid. Resetting introduces discontinuities into the time evolution of the lapse and the shift at a given spatial location. However, in the $3+1$ for-

malism, no time derivatives of the lapse and the shift appear in the equations. Resetting has no effect on the dynamical state of the system or on the gauge at a given time; it just affects how the coordinates evolve after the resetting.

Changes in spacetime slicing which maintain the explicit planar symmetry only directly impact $K_x{}^x$. Nonlinear source terms in the evolution equation for $K_x{}^x$ have the potential to generate runaway growth of $K_x{}^x$ when $K_x{}^x$ is positive. Our lapse resetting condition drives $K_x{}^x$ toward a small negative value to insure against this. In addition, a negative $K_x{}^x$ implies that the proper distance between hypersurface normals displaced in the x-direction increases with time. Together with our shift resetting condition, which keeps h_{xx} roughly constant, this results in hypersurface normals which point outward at the boundaries of the computational domain. Some features in $K_x{}^x$ and $D_x{}^x$ potentially associated with instability advect along hypersurface normals (see Sec. IX), and are then advected out of the grid before they can do much harm.

The equation for the lapse is derived by imposing the condition at the time of resetting

$$\partial_t K_x{}^x - \beta \partial_x K_x{}^x = -\Gamma [K_x{}^x - (K_x{}^x)_T], \tag{21}$$

where $(K_x{}^x)_T$ is a specified ''target'' value, and Γ is a damping constant, which is chosen to be comparable to the characteristic frequency of the waves we are propagating. Substituting this condition into the evolution equation for $K_x{}^x$ [Eq. (13)] and simplifying using the energy constraint, we obtain our lapse resetting condition,

$$\partial_x \left(\frac{\alpha'}{h_{xx}} \right) = \alpha \left[\Gamma [K_x{}^x - (K_x{}^x)_T] + K_x{}^x K_x{}^x - K_y{}^y K_z{}^z + \frac{D_y{}^y D_z{}^z}{h_{xx}} \right] - D_x{}^x \left(\frac{\alpha'}{h_{xx}} \right). \tag{22}$$

To limit initial transients in the lapse, given our initial condition $K_x{}^x = 0$, the target value is made proportional to $(1 - e^{-\Gamma t/4})$.

In our colliding wave calculations, Eq. (22) as it stands can cause the lapse to become negative at the edges of the grid, if the second derivative of the lapse becomes too negative. To prevent this, we replace $\mathcal{S}$, the expression in square brackets in Eq. (22), by

$$\mathcal{S} \to \frac{\mathcal{S}}{\sqrt{1 + (\mathcal{S}/\mathcal{S}_{lim})^2}} \tag{23}$$

when $\mathcal{S}$ is negative, so $\mathcal{S} > -|\mathcal{S}_{lim}|$. A side effect of the limiter is to allow $K_x{}^x$ to become more negative than its target value.

We choose an equation for the shift so that at the time of resetting, h_{xx} is advected along hypersurface normals:

$$\partial_t h_{xx} - \beta \partial_x h_{xx} = 0. \tag{24}$$

Substituting this requirement into the evolution equation for h_{xx} [Eq. (9)], we obtain the shift resetting condition,

$$\partial_x \beta = \alpha K_x{}^x. \tag{25}$$

VI. HYPERBOLIC SYSTEMS

The evolution equations presented in Sec. II have been cast in a first order, flux-conservative form, represented by the following set of l equations

$$\partial_t \mathbf{q} + \partial_x [\mathbf{F}(\mathbf{q})] = \mathbf{S}(\mathbf{q}), \tag{26}$$

where $\mathbf{q}$ is a vector of l variables. The flux vector is given by

$$\mathbf{F}(\mathbf{q}) = \mathbf{A}(x)\mathbf{q}, \tag{27}$$

where the $l \times l$ characteristic matrix $\mathbf{A}(x)$ is the flux Jacobian, $\partial_{\mathbf{q}}[\mathbf{F}(\mathbf{q})]$. The system is hyperbolic if $\mathbf{A}(x)$ has a complete set of eigenvectors and real eigenvalues.

A. Modified Bona-Massó formulation

The standard BM formulation [8] creates a hyperbolic scheme by introducing the redundant variables V_i, which are defined as

$$V_i = D_{ik}{}^k - D^k{}_{ki} \Rightarrow V_x = D_y{}^y + D_z{}^z. \tag{28}$$

The momentum constraint is used to evolve V_x:

$$\partial_t V_x + \partial_x [-\beta V_x] = \alpha [D_y{}^y K_y{}^y + D_z{}^z K_z{}^z - (D_y{}^y + D_z{}^z) K_x{}^x - (\alpha'/\alpha)(K_y{}^y + K_z{}^z)]. \tag{29}$$

We densitize the lapse according to the Choquet-Bruhat algebraic condition (see Sec. III), which simplifies the standard BM system of equations considerably. With $\alpha \to Q\sqrt{\det(h_{ij})}$ and $(A_x = \partial_x \ln \alpha) \to D_Q + D_x{}^x + D_y{}^y + D_z{}^z$, the fluxes for $K_i{}^j$ reduce to

$$F(K_x{}^x) = -\beta K_x{}^x + \frac{\alpha}{h_{xx}} \left[D_Q + D_x{}^x + \left(2 - \frac{n}{2} \right) V_x \right], \tag{30}$$

$$F(K_y{}^y) = -\beta K_y{}^y + \frac{\alpha}{h_{xx}} \left(D_y{}^y - \frac{n}{2} V_x \right),$$

$$F(K_z{}^z) = -\beta K_z{}^z + \frac{\alpha}{h_{xx}} \left(D_z{}^z - \frac{n}{2} V_x \right). \tag{31}$$

Our advection of Q and D_Q, as described in Sec. III, corresponds to harmonic slicing, a special case of the standard BM lapse evolution equation. We also advect β and β', whereas the standard BM formulation specifies the shift as a known function of x and t.

B. Modified Anderson-York formulation

The AY formulation differs from the BM scheme in how the momentum constraint is used to make the system hyperbolic. The AY scheme eliminates the need for the BM redun-

dant variables V_i by incorporating the momentum constraint into the evolution equation for the f_{kij} variables, which are defined as

$$f_{kij}=D_{kij}+h_{ki}V_j+h_{kj}V_i. \tag{32}$$

The V_i variables in this equation are not separate variables, but rather denote the combinations of the D's given in Eq. (28).

The AY formulation replaces the BM D_{kij} with f_{kij}, which are simply the spatial metric derivative terms in the K_{ij} fluxes of the BM formulation. Using this as a guide, we generalize the AY scheme (whose original form is restricted to Ricci evolution, $n=0$) to allow for non-zero energy constraint contributions. This leads to

$$f_{kij}=D_{kij}+h_{ki}V_j+h_{kj}V_i-\frac{n}{2}V_k h_{ij}. \tag{33}$$

The generalization in Eq. (33) works as long as the inverse transformation from f_{kij} to D_{kij} exists, which is the case for $n\neq 1$. An evolution equation is obtained for f_{kij} from Eq. (33) by using the momentum constraint to eliminate the time derivative of the V_i variables. For our modified AY scheme, we then raise one index so that $f_{ki}{}^{j}=f_{kil}h^{lj}$ are our basic variables. A hyperbolic system results without the need for the BM redundant *variables* V_i.

For the diagonal metric plane-wave case under consideration, Eq. (33) reduces to

$$f_x{}^x=D_x{}^x+\left(2-\frac{n}{2}\right)V_x, \tag{34}$$

$$f_y{}^y=D_y{}^y-\frac{n}{2}V_x, \quad f_z{}^z=D_z{}^z-\frac{n}{2}V_x. \tag{35}$$

We have simplified our notation in that $f_{ki}{}^{j}\to f_i{}^j$ for this 1D problem. Notice that $f_{yx}{}^{y}=V_x$ and $f_{zx}{}^{z}=V_x$, which contribute to fluxes in the y and z directions, are not zero. However, with planar symmetry the divergence of these flux components vanishes identically.

The evolution equations for $f_i{}^j$ are

$$\partial_t f_x{}^x+\partial_x[-\beta f_x{}^x-\beta'+\alpha K_x{}^x]=\left(2-\frac{n}{2}\right)\alpha\mathcal{C}, \tag{36}$$

$$\partial_t f_y{}^y+\partial_x[-\beta f_y{}^y+\alpha K_y{}^y]=-\frac{n}{2}\alpha\mathcal{C},$$

$$\partial_t f_z{}^z+\partial_x[-\beta f_z{}^z+\alpha K_z{}^z]=-\frac{n}{2}\alpha\mathcal{C}, \tag{37}$$

where

$$\mathcal{C}=[D_y{}^y K_y{}^y+D_z{}^z K_z{}^z-(D_y{}^y+D_z{}^z)K_x{}^x -(\alpha'/\alpha)(K_y{}^y+K_z{}^z)]. \tag{38}$$

The D's in Eq. (38) are not separate variables, but denote

$$D_x{}^x=f_x{}^x-\left(\frac{2-\frac{n}{2}}{1-n}\right)[f_y{}^y+f_z{}^z], \tag{39}$$

$$D_y{}^y=\frac{1}{2(1-n)}[(2-n)f_y{}^y+nf_z{}^z],$$

$$D_z{}^z=\frac{1}{2(1-n)}[(2-n)f_z{}^z+nf_y{}^y]. \tag{40}$$

These relations are the inverse transformation of the system of Eqs. (34) and (35). One can see that $n=1$ is not allowed.

The $K_i{}^j$ evolution equations are the same as in our modified BM scheme, with the understanding again that the D's in the source terms are not separate variables, but the above linear combinations of f's [Eqs. (39) and (40)]. The fluxes are defined so the $f_i{}^j$ variables can replace the expressions involving $D_i{}^j$ in the fluxes of our modified BM scheme. The following fluxes result:

$$F(K_x{}^x)=-\beta K_x{}^x+\frac{\alpha}{h_{xx}}(D_Q+f_x{}^x), \tag{41}$$

$$F(K_y{}^y)=-\beta K_y{}^y+\frac{\alpha}{h_{xx}}f_y{}^y,$$

$$F(K_z{}^z)=-\beta K_z{}^z+\frac{\alpha}{h_{xx}}f_z{}^z. \tag{42}$$

The AY formalism imposes the Choquet-Bruhat algebraic condition on the lapse, as we did in our modified BM scheme. The evolution of the lapse and the shift between gauge resettings is treated in exactly the same way as in our modified BM formalism.

C. Modified Arnowitt-Deser-Misner formulation

The simplest of the hyperbolic schemes we present is our modified ADM formulation, which consists of Eqs. (9) to (14), (17), and (18), with $\alpha=Q\sqrt{\det(h_{ij})}$ and $\alpha'/\alpha=D_Q+D_x{}^x+D_y{}^y+D_z{}^z$. This system is hyperbolic when the metric is diagonal if $n<0$ or $0<n<1$. The fluxes for $K_i{}^j$ are

$$F(K_x{}^x)=-\beta K_x{}^x+\frac{\alpha}{h_{xx}}\left[D_Q+D_x{}^x+\left(2-\frac{n}{2}\right)(D_y{}^y+D_z{}^z)\right], \tag{43}$$

$$F(K_y{}^y)=-\beta K_y{}^y+\frac{\alpha}{h_{xx}}\left[\left(1-\frac{n}{2}\right)D_y{}^y-\frac{n}{2}D_z{}^z\right],$$

$$F(K_z{}^z)=-\beta K_z{}^z+\frac{\alpha}{h_{xx}}\left[\left(1-\frac{n}{2}\right)D_z{}^z-\frac{n}{2}D_y{}^y\right]. \tag{44}$$

The hyperbolicity of our modified ADM system of equations breaks down for $n=0$ and $n\geqslant 1$. Although our ADM formulation at $n=0$ is non-hyperbolic, it is stable. At $n=1$, however, the system is both non-hyperbolic and on the verge

of being unstable. For $n>1$, the equations have complex eigenvalues, giving unstable exponential growth of errors.

D. Wave modes

1. BM

The hyperbolic system of equations obtained from the modified BM formulation described in Sec. VI A is

$$\partial_t \mathbf{q} + \partial_x [\mathbf{A}(x)\mathbf{q}] = \mathbf{S}(\mathbf{q}), \tag{45}$$

where

$$\mathbf{q} = \begin{pmatrix} D_x{}^x \\ D_y{}^y \\ D_z{}^z \\ K_x{}^x \\ K_y{}^y \\ K_z{}^z \\ V_x \\ D_Q \\ \beta' \end{pmatrix}, \tag{46}$$

and

$$\mathbf{A}(x) = \begin{pmatrix} -\beta & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & -1 \\ 0 & -\beta & 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & -\beta & 0 & 0 & \alpha & 0 & 0 & 0 \\ \frac{\alpha}{h_{xx}} & 0 & 0 & -\beta & 0 & 0 & \frac{\alpha}{h_{xx}}\left(2-\frac{n}{2}\right) & \frac{\alpha}{h_{xx}} & 0 \\ 0 & \frac{\alpha}{h_{xx}} & 0 & 0 & -\beta & 0 & -\frac{\alpha}{h_{xx}}\left(\frac{n}{2}\right) & 0 & 0 \\ 0 & 0 & \frac{\alpha}{h_{xx}} & 0 & 0 & -\beta & -\frac{\alpha}{h_{xx}}\left(\frac{n}{2}\right) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta \end{pmatrix}. \tag{47}$$

The nine eigenmodes of the homogeneous system are obtained from the characteristic matrix, $\mathbf{A}(x)$. Six of the eigenmodes travel along the light cones. They are

$$\left.\begin{array}{l} \frac{1}{\sqrt{h_{xx}}}\left[D_x{}^x + D_Q + \left(2-\frac{n}{2}\right)V_x\right] \pm \left[K_x{}^x - \frac{\beta'}{\alpha}\right], \\ \frac{1}{\sqrt{h_{xx}}}\left[D_y{}^y - \frac{n}{2}V_x\right] \pm K_y{}^y, \quad \frac{1}{\sqrt{h_{xx}}}\left[D_z{}^z - \frac{n}{2}V_x\right] \pm K_z{}^z \end{array}\right\} \text{speeds} = -\beta \pm \frac{\alpha}{\sqrt{h_{xx}}}. \tag{48}$$

The remaining three eigenmodes are simply the variables V_x, D_Q, and β', which travel along the hypersurface normals, with speeds $-\beta$.

The eigenmodes of the characteristic matrix, however, do not necessarily describe how solutions of the full nonlinear system of equations propagate. It is a special property of plane-wave systems that eigenmodes of the full nonlinear system of equations exist which consist of purely right-going waves with $K_y{}^y \pm K_z{}^z = (D_y{}^y \pm D_z{}^z)/\sqrt{h_{xx}}$, purely left-going waves with $K_y{}^y \pm K_z{}^z = -(D_y{}^y \pm D_z{}^z)/\sqrt{h_{xx}}$, and $D_x{}^x = K_x{}^x = 0$. These are solutions of the Einstein equations in a gauge with $\alpha = 1$ and $\beta' = 0$. In our nonlinear colliding plane wave calculations, our initial conditions are such that the waves have this form. The right-going wave is in the left half of the grid, the left-going wave is in the right half of the grid, and they are just at the point of colliding. When discussing solutions of the full nonlinear system of equations, we refer to the transverse-traceless quantities $(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y - K_z{}^z)$, the constraint quantities $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y + K_z{}^z)$, and the longitudinal variables $D_x{}^x/\sqrt{h_{xx}}$ and $K_x{}^x$. After the waves pass through each other, it is only approximately true that the transverse-traceless quantities have the form of purely right-going and purely left-going waves as described above and it is not at all true that the constraint quantities have this form.

The characteristic speeds apply to small amplitude, short wavelength perturbations in the variables, so that the principal terms (which are first derivative terms) dominate over the source terms. The disturbances in the constraint quantities which propagate along the characteristics will generally be constraint-violating because the constraints explicitly tie the principal terms to the nonlinear source terms, and require that they cancel. The longitudinal variables, $D_x{}^x/\sqrt{h_{xx}}$ and $K_x{}^x$, are not eigenmodes of the homogeneous system. In the full nonlinear system, $D_x{}^x/\sqrt{h_{xx}}$ and $K_x{}^x$ have some features which propagate along the light cones, and some features which propagate along the hypersurface normals. The propagation of the longitudinal variables is strongly dependent on the choice of gauge.

2. AY

There is a complete set of eight eigenmodes of the modified AY homogeneous system of equations. The six eigenmodes which travel along the light cones are

$$\left.\begin{array}{l} \dfrac{1}{\sqrt{h_{xx}}}(f_x{}^x+D_Q)\pm\left(K_x{}^x-\dfrac{\beta'}{\alpha}\right), \\ \dfrac{f_y{}^y}{\sqrt{h_{xx}}}\pm K_y{}^y, \quad \dfrac{f_z{}^z}{\sqrt{h_{xx}}}\pm K_z{}^z \end{array}\right\} \text{speeds}=-\beta\pm\frac{\alpha}{\sqrt{h_{xx}}}. \tag{49}$$

The remaining two eigenmodes are the variables D_Q and β', which travel along the hypersurface normals, with speeds $-\beta$.

3. ADM

For the modified ADM homogeneous system, the eigenmodes, which form a complete hyperbolic system for $n<0$ or $0<n<1$, consist of ''longitudinal'' and ''physical'' eigenmodes propagating along the light cone,

$$\left.\begin{array}{l} \dfrac{1}{\sqrt{h_{xx}}}\left[nD_x{}^x+nD_Q+\left(2-\dfrac{n}{2}\right)(D_y{}^y+D_z{}^z)\right]\pm\left[nK_x{}^x-\dfrac{n\beta'}{\alpha}+\left(2-\dfrac{n}{2}\right)(K_y{}^y+K_z{}^z)\right], \\ \dfrac{1}{\sqrt{h_{xx}}}[D_y{}^y-D_z{}^z]\pm[K_y{}^y-K_z{}^z] \end{array}\right\} \text{speeds}=-\beta\pm\frac{\alpha}{\sqrt{h_{xx}}}, \tag{50a}$$

''constraint'' eigenmodes propagating inside the light cone for $0<n<1$,

$$\left.\sqrt{\frac{1-n}{h_{xx}}}[D_y{}^y+D_z{}^z]\pm[K_y{}^y+K_z{}^z]\right\}\text{speeds} = -\beta\pm\frac{\alpha}{\sqrt{h_{xx}}}\sqrt{1-n}, \tag{50b}$$

and the D_Q and β' eigenmodes with speeds $-\beta$. Hyperbolicity fails for $n=0$ because the ''longitudinal'' eigenmodes are not independent of the ''constraint'' eigenmodes, for $n=1$ because the two ''constraint'' eigenmodes are not independent of each other, and for $n>1$ because the ''constraint'' eigenvalues are complex.

VII. BOUNDARY CONDITIONS

Since in numerical relativity, computations are usually performed on a limited grid within a much larger space, the boundary conditions should be designed to be consistent with how waves propagate while they are still inside the grid. Even more important, since the evolution equations admit constraint-violating solutions, constraint violations will propagate into the grid unless boundary conditions are carefully designed to suppress them.

Consider the ''constraint'' eigenmodes. They are $[D_y{}^y+D_z{}^z-nV_x]/\sqrt{h_{xx}}\pm[K_y{}^y+K_z{}^z]$ in BM and AY (though expressed in terms of different variables), and $\sqrt{1-n}[D_y{}^y+D_z{}^z]/\sqrt{h_{xx}}\pm[K_y{}^y+K_z{}^z]$ in ADM. Even for the same value of the energy constraint coefficient n, what is outgoing in BM and AY is different from what is outgoing in ADM. Furthermore, for a given solution, the amplitudes of the BM, AY, and ADM modes depend on n. Whatever the correct boundary condition, its effect on the solution should be independent of the equation formulation. The relative amount of right and left-going ''constraint'' modes is also gauge dependent, in the sense that the choice of boundary conditions in solving the constraint equations in the initial conditions is a gauge choice, and this affects the relative values of $(D_y{}^y+D_z{}^z)$ and $(K_y{}^y+K_z{}^z)$ at all later times. The initial conditions symmetric about the midpoint of the grid at $x=10$ give purely *incoming* ''constraint'' modes for $n=0$ ($[D_y{}^y+D_z{}^z]/\sqrt{h_{xx}}=\pm[K_y{}^y+K_z{}^z]$ on the left/right edges of the grid) initially and at all times until the effects of the wave collision reach the boundaries.

The ''longitudinal'' eigenmodes involving $D_x{}^x$ and $K_x{}^x$ are also formulation dependent, since they are different in ADM from what they are in BM and AY, and they depend on n in all three formulations. There is gauge freedom to pose any boundary conditions one likes on these modes, but a poor choice might give rise to singularities in $D_x{}^x$ or $K_x{}^x$ inside the grid.

The ''physical'' eigenmodes $[D_y{}^y-D_z{}^z]/\sqrt{h_{xx}}\pm[K_y{}^y-K_z{}^z]$ are the same in all three formulations, and are inde-

pendent of n. However, their time evolution is gauge-dependent because the nonlinear source terms in their evolution equations involve the gauge-dependent constraint quantities. With our choice of initial gauge, the amplitudes of $(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y - K_z{}^z)$ differ by about 3% after the wave collision, so there is typically a 3% admixture of the incoming ''physical'' eigenmode as the outgoing waves approach the boundaries.

A gauge-independent measure of the amplitudes of left and right-going gravitational waves can be obtained by projecting the Weyl tensor onto a complex null tetrad, as in the Newman-Penrose spin coefficient formalism [27],

$$\Psi^0_{\pm} = R_{(t)(y)(t)(y)} - R_{(t)(z)(t)(z)} + R_{(x)(y)(x)(y)} - R_{(x)(z)(x)(z)} \mp 2[R_{(t)(y)(x)(y)} - R_{(t)(z)(x)(z)}] \tag{51}$$

for right/left propagation. A purely right-going wave would have $\Psi^0_- = 0$. Our numerical results indicate that plane waves after a collision are indeed purely outgoing by this standard. As an outgoing wave boundary condition, using the evolution equations to evaluate the time derivatives of the extrinsic curvature in the Riemann tensor, this becomes

$$\left(\partial_x \left[\frac{(D_y{}^y - D_z{}^z)}{\sqrt{h_{xx}}} \mp (K_y{}^y - K_z{}^z) \right] \right) \frac{1}{\sqrt{h_{xx}}} + \frac{1}{2} \frac{(D_y{}^y - D_z{}^z)}{\sqrt{h_{xx}}} \left[\frac{(D_y{}^y + D_z{}^z)}{\sqrt{h_{xx}}} \mp (K_y{}^y + K_z{}^z) \right] + \frac{1}{2} \frac{(D_y{}^y + D_z{}^z)}{\sqrt{h_{xx}}} \left[\frac{(D_y{}^y - D_z{}^z)}{\sqrt{h_{xx}}} \mp (K_y{}^y - K_z{}^z) \right] = 0 \tag{52}$$

at the right/left boundaries. This expression is consistent with $[D_y{}^y - D_z{}^z]/\sqrt{h_{xx}} = \pm [K_y{}^y - K_z{}^z]$ if and only if $[D_y{}^y + D_z{}^z]/\sqrt{h_{xx}} = \pm [K_y{}^y + K_z{}^z]$.

Since conventional outgoing wave boundary conditions are not appropriate, our boundary conditions are based on a smooth second order extrapolation of the variables, which is corrected to make sure the energy and momentum constraint equations are satisfied on the boundaries. Equation (52) could also be imposed at the boundaries to further improve the extrapolation, but we have not tried this. Our procedure is detailed further in Sec. VIII B 3.

In addition to the eigenmodes discussed above, there are eigenmodes propagating along the hypersurface normals, which can be incoming or outgoing, depending on the sign of the shift on the boundaries. It seems to be important for stability that the hypersurface normals do not point into the grid (see Secs. IX C and IX D).

Our results show that quadratic extrapolation without correction for the energy and momentum constraints produces a significant but not dominant error (see Sec. IX D). However, errors from imposing outgoing boundary conditions on the ''constraint'' eigenmodes, or from using standard constant extrapolation, would swamp all other errors as they propagate into the grid. Standard constant extrapolation, which gives the same values for the variables, and therefore the fluxes, in the ghost cell and adjoining physical cell, also eliminates the incoming ''constraint'' eigenmodes.

VIII. NUMERICAL METHODS

A. Strang splitting

As described in Sec. VI, all of the formulations we tested, both hyperbolic and non-hyperbolic, are in first order, flux conservative form. We solve all these systems of equations using a Strang-split method [22]. In this method, the homogeneous transport part of Eq. (26) and the contributions from the source terms are treated separately. In particular, the following straightforward system of ordinary differential equations is first solved over half a time step

$$\partial_t \mathbf{q} = \mathbf{S}(\mathbf{q}). \tag{53}$$

Then, the transport part of Eq. (26), which contains the flux terms, is solved over a full time step

$$\partial_t \mathbf{q} + \partial_x [\mathbf{F}(\mathbf{q})] = 0. \tag{54}$$

Our methods for solving the transport step are discussed in Sec. VIII B below. The calculation is completed by again solving Eq. (53) over half a time step.

We choose to use the Strang-split method because it is simpler in the context of how we are handling boundary conditions. An iterative scheme such as the MacCormack method [28] requires repeated implementation of the boundary conditions each time step. However, in the Strang-split scheme, the boundary conditions are imposed only once during each time step. The fewer applications of the boundary conditions in the Strang-split method is advantageous because we are using quadratic extrapolation to obtain ghost cell values. Quadratic extrapolation amplifies any jitter at the boundaries, and the frequent application of quadratic extrapolation in iterative schemes such as MacCormack could easily lead to an instability.

B. Transport step

In the transport step, we solve Eq. (54) both with a finite difference method and with a wave propagation approach, which takes advantage of the eigenfields of a diagonalizable hyperbolic system. Advanced numerical methods for diagonalizable hyperbolic systems introduce limiter functions to resolve sharp discontinuities that typically arise in hydrodynamics problems. A smooth problem can be solved just as accurately and more efficiently with a finite difference method. In vacuum general relativity, discontinuities may or may not arise, depending on the gauge conditions. Commonly used gauge conditions lead to steep gradients near black hole horizons. One can deal with these gradients by using high resolution methods requiring diagonalizable hyperbolic formulations; or, one can dynamically adjust the gauge conditions so as to avoid the steep gradients altogether [29].

Whether one uses a finite difference method or a sophisticated hyperbolic technique, it is important to have a nu-

merical scheme which is fully second order accurate for smooth solutions and generalizable to black hole spacetimes and higher dimensions. It is straightforward to devise a finite difference scheme based on a Taylor series expansion which is formally second order accurate. High resolution Riemann-based wave propagation algorithms introduced by LeVeque [23], which decompose $\Delta\mathbf{q}$ across a grid cell interface into a linear combination of eigenvectors of the $\mathbf{A}(x)$ matrix, are applicable to a wide variety of diagonalizable hyperbolic problems. Flux differences are calculated from the $\Delta\mathbf{q}$ decomposition. We refer to these algorithms as ''standard wave decomposition'' methods. However, the standard wave decomposition methods are not second order accurate for smooth solutions when the characteristic matrix $\mathbf{A}(x)$ is a function of position, because the changes in $\mathbf{A}(x)$ across cell boundaries as well as $\Delta\mathbf{q}$'s must be accounted for in flux differences. In numerical relativity problems, $\mathbf{A}(x)$ depends on the lapse, the shift, and the spatial metric, and can have gradients comparable with the gradients of $\mathbf{q}$.

LeVeque has suggested a wave propagation approach for solving variable coefficient flux problems based on splitting up the jump in $\mathbf{F}(\mathbf{q})$ rather than the jump in $\mathbf{q}$ [24]. We refer to this approach as ''flux-based wave decomposition.'' We develop and apply this method to solve the Einstein equations for 1D nonlinear plane waves as described below in Sec. VIII B 1. We show in the Appendix that flux-based wave decomposition methods are formally second order accurate for sufficiently smooth solutions for arbitrary smooth variations of the eigenvalues and eigenvectors (see also Bale *et al.* [30]). For further discussion and analysis of flux-based wave decomposition methods, in the context of more general approximate Riemann solvers, see [31]. While it is difficult to formally prove second order convergence for numerical methods since this also requires proving stability, our numerical tests of these methods, and those of Ref. [30], typically exhibit second order convergence.

1. Flux-based wave decomposition

Using Eq. (54) to update average grid cell values of the variables $\mathbf{q}$ requires knowing flux values at grid cell interfaces. The interface flux values are found by solving the following equation, obtained by multiplying Eq. (54) by $\mathbf{A}(x)$ on the left hand side:

$$\partial_t[\mathbf{F}(\mathbf{q})]+\mathbf{A}(x)\partial_x[\mathbf{F}(\mathbf{q})]=0. \tag{55}$$

The time derivative of $\mathbf{A}(x)$ vanishes because the variables on which $\mathbf{A}(x)$ depends (the lapse, the shift, and the longitudinal part of the spatial metric) have no fluxes, and are not updated during the transport step. Using Eq. (55) to compute the interface fluxes was originally described by Bona *et al.* [8]. However, it is not clear from [8] how they handled problems in which $\mathbf{A}(x)$ varies from cell to cell.

Equation (55) is a linear advection equation for the flux vector, $\mathbf{F}(\mathbf{q})$. As such, flux values at cell interfaces can be updated by solving Riemann problems based on decomposing flux differences between adjacent grid cells into eigenvector expansions (see [5] for a discussion of solving Riemann problems for the advection equation), and including correction terms to give second order accuracy. We develop two wave propagation methods based on this idea which we call methods **I** and **II**. A wave in this approach is defined as a discontinuity in the *flux* associated with a certain eigenmode across the characteristic corresponding to that eigenmode.

We explicitly deal with the fact that the eigenvalues and eigenvectors of the characteristic matrix are varying across the grid. The magnitudes of the eigenvalues give the wave speeds and the signs of the eigenvalues give the wave directions. In the flux decomposition for method **I**, we need to decide if a wave is left-going or right-going at a given cell interface. This is determined by the sign of the average of the eigenvalues obtained from the characteristic matrices on either side of the interface. If the average eigenvalue for a particular eigenmode is negative, then the corresponding eigenvector is evaluated in the cell to the left of the interface. If the average eigenvalue is positive, then the eigenvector is evaluated in the cell to the right of the interface. In method **II**, the eigenvalues and eigenvectors at a cell interface are obtained from the characteristic matrix at the interface, calculated as an average from the adjacent cells. For both methods, waves with zero interface speed still contribute to the flux difference. We can include these contributions in either the left- or right-going waves of method **I**, as long as we do so consistently.

In method **I**, the flux difference decomposition takes the following form at the interface between cells i and $i-1$:

$$\begin{aligned}\mathbf{F}(\mathbf{q}_i)-\mathbf{F}(\mathbf{q}_{i-1})&=\mathbf{A}_i\mathbf{q}_i-\mathbf{A}_{i-1}\mathbf{q}_{i-1}\\&=\sum_{L=1}^{m}\gamma^L_{i-(1/2)}\mathbf{r}^L_{i-1}+\sum_{R=m+1}^{M}\gamma^R_{i-(1/2)}\mathbf{r}^R_i,\end{aligned} \tag{56}$$

where $\mathbf{r}$ are right eigenvectors of the characteristic matrix, and M is the total number of eigenmodes. We denote the left-going waves at this interface as $\mathbf{W}^L_{i-(1/2)}=\gamma^L_{i-(1/2)}\mathbf{r}^L_{i-1}$, where $1\leqslant L\leqslant m$. The right-going waves are given by $\mathbf{W}^R_{i-(1/2)}=\gamma^R_{i-(1/2)}\mathbf{r}^R_i$, where $m+1\leqslant R\leqslant M$. The number of left-going waves, m, can vary from interface to interface since the sign of the average eigenvalue can change from cell to cell. The eigenvectors $\mathbf{r}_{i-1}$ are evaluated in cell $i-1$. Likewise, $\mathbf{r}_i$ are evaluated in cell i. The coefficients $\gamma_{i-(1/2)}$ are obtained by solving Eq. (56); the subscripts $i-\frac{1}{2}$ indicate interface values. In method **II**, the flux difference decomposition at a given interface between cells i and $i-1$ is the same as Eq. (56), except the eigenvectors $\mathbf{r}_{i-(1/2)}$ of the averaged characteristic matrix $\mathbf{A}_{i-(1/2)}=(\mathbf{A}_{i-1}+\mathbf{A}_i)/2$ replace both $\mathbf{r}_{i-1}$ and $\mathbf{r}_i$.

Method **I** is implemented in the context of the CLAWPACK software package [32]. The first order wave propagation and second order corrections in both methods **I** and **II** are analogous to Eqs. (18) and (19) of LeVeque's paper on standard wave decomposition methods [23]. The updated value of $\mathbf{q}_i$ is given by

$$\overline{\mathbf{q}}_i = \mathbf{q}_i - \frac{\Delta t}{\Delta x}\left(\sum_R \mathbf{W}^R_{i-(1/2)} + \sum_L \mathbf{W}^L_{i+(1/2)}\right) - \frac{\Delta t}{\Delta x}(\tilde{\mathbf{F}}_{i+(1/2)} - \tilde{\mathbf{F}}_{i-(1/2)}). \quad (57)$$

$\tilde{\mathbf{F}}_{i\pm(1/2)}$ are flux correction terms which can be reduced near discontinuities by introducing limiter functions. Limiters prevent the oscillatory behavior around discontinuities seen with finite difference methods. In the absence of limiting, the flux corrections are

$$\tilde{\mathbf{F}}_{i\pm(1/2)} = \frac{1}{2}\left(\sum_R \mathbf{W}^R_{i\pm(1/2)} - \sum_L \mathbf{W}^L_{i\pm(1/2)}\right) - \frac{1}{2}\frac{\Delta t}{\Delta x}\sum_{p=1}^{M} \lambda^p_{i\pm(1/2)}\mathbf{W}^p_{i\pm(1/2)}, \quad (58)$$

where $\lambda^p_{i\pm(1/2)}$ denote cell-interface speeds.

Both flux-based wave decomposition methods **I** and **II** are successful in giving second order convergent results in our numerical calculations.

2. Finite difference method

To solve Eq. (54) using a Lax-Wendroff finite difference method, we first perform a second order Taylor expansion of $\mathbf{q}$ around t:

$$\mathbf{q}(x,t+\Delta t) = \mathbf{q}(x,t) + \Delta t\,\partial_t\mathbf{q}(x,t) + \frac{1}{2}\Delta t^2\partial_t^2\mathbf{q}(x,t). \quad (59)$$

Observe that

$$\partial_t\mathbf{q} = -\partial_x[\mathbf{A}(x)\mathbf{q}], \quad (60)$$

and, taking another time derivative,

$$\partial_t^2\mathbf{q} = -\partial_x[\mathbf{A}(x)(\partial_t\mathbf{q})] = \partial_x[\mathbf{A}(x)\partial_x(\mathbf{A}(x)\mathbf{q})]. \quad (61)$$

Note that the time derivative of $\mathbf{A}(x)$ vanishes as in Eq. (55). Plugging these expressions for $\partial_t\mathbf{q}$ and $\partial_t^2\mathbf{q}$ into Eq. (59) gives

$$\mathbf{q}(x,t+\Delta t) = \mathbf{q}(x,t) - \Delta t\,\partial_x[\mathbf{A}(x)\mathbf{q}(x,t)] + \frac{1}{2}\Delta t^2\partial_x[\mathbf{A}(x)\partial_x(\mathbf{A}(x)\mathbf{q}(x,t))]. \quad (62)$$

Making the centered finite difference approximation to the derivatives in Eq. (62), the updated value of $\mathbf{q}_i$ is given by

$$\overline{\mathbf{q}}_i = \mathbf{q}_i - \frac{\Delta t}{\Delta x}(\mathbf{A}_{i+1}\mathbf{q}_{i+1} - \mathbf{A}_i\mathbf{q}_i) + \frac{1}{4}\left(\frac{\Delta t}{\Delta x}\right)^2 \times[(\mathbf{A}_i + \mathbf{A}_{i+1})(\mathbf{A}_{i+1}\mathbf{q}_{i+1} - \mathbf{A}_i\mathbf{q}_i) - (\mathbf{A}_{i-1} + \mathbf{A}_i) \times(\mathbf{A}_i\mathbf{q}_i - \mathbf{A}_{i-1}\mathbf{q}_{i-1})]. \quad (63)$$

3. Boundary conditions

Our numerical methods only require values in one ghost cell at each boundary. We obtain values for all variables in the ghost cell by quadratic extrapolation from the three adjacent physical cells. Numerical integration of the constraint equations from the last physical cell to the ghost cell by the trapezoidal rule is used to correct the constraint quantities $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y + K_z{}^z)$ in the ghost cell, with iteration to convergence.

IX. RESULTS

A. Initial conditions

The initial conditions must satisfy the constraint equations, Eqs. (19) and (20). Since the constraint equations are differential equations, they require boundary conditions for their solutions. Different choices of boundary conditions correspond to different gauge conditions. We choose symmetric boundary conditions which give flat space between two waves. This means that we choose $(D_y{}^y + D_z{}^z)$ and $(K_y{}^y + K_z{}^z)$ to vanish initially between the waves.

The variables h_{xx}, $K_x{}^x$, and the combinations $(D_y{}^y - D_z{}^z)$ and $(K_y{}^y - K_z{}^z)$ are freely specifiable. We normally take $h_{xx} = 1$, $K_x{}^x = 0$, and

$$0.25\ln\left(\frac{h_{yy}}{h_{zz}}\right) = \sum_{i=1}^{2} A_i \cos^2\left[\frac{2}{\pi}\frac{(x - x_{0i})}{w_i}\right]\sin[k_i(x - x_{0i}) + \delta_i], \quad (64)$$

for $-w_i < (x - x_{0i}) < w_i$, and zero outside that range. The x derivative of Eq. (64) gives $(D_y{}^y - D_z{}^z)/2$. In our standard initial conditions for colliding plane waves, one wave is initially on the left and moving to the right, with $[K_y{}^y - K_z{}^z]_1 = [(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}]_1$. The other wave is initially on the right and moving to the left, with $[K_y{}^y - K_z{}^z]_2 = -[(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}]_2$. The parameters for $0 \leqslant x \leqslant 20$ are $w_i = 4.0$, $k_i = 1.6$, $A_i = 0.08$, $x_{01} = 6.0$, $x_{02} = 14.0$, $\delta_1 = 0$, and $\delta_2 = \pi$. These initial conditions, depending on the variables, are symmetric (or antisymmetric) about $x = 10$, and symmetry (or antisymmetry) is preserved throughout the evolution. Hence, our figures only show the range $0 \leqslant x \leqslant 10$. Since the initial plane waves do not overlap, and $(D_y{}^y + D_z{}^z)$ and $(K_y{}^y + K_z{}^z)$ vanish at $x = 0$, the initial conditions are two analytic single plane waves of the type described by Misner, Thorne, and Wheeler [33].

Our initial conditions produce large amplitude, nonlinear colliding gravitational plane waves. Our measure of ''large amplitude'' is that h_{yy} and h_{zz} are substantially different from 1 by the time the waves have traversed the grid. It is known that nonlinear plane-wave spacetimes develop a singularity behind the wave [34,35]. For a single plane wave, this is only a coordinate singularity, while for colliding plane waves, a physical singularity also develops. The values we take for our wave amplitudes are about as large as possible without allowing a singularity to develop during the crossing time of the waves. One can get a feel for this value by asking at what amplitude does a singularity develop at the left edge of the grid for a single plane wave exiting the right edge? For

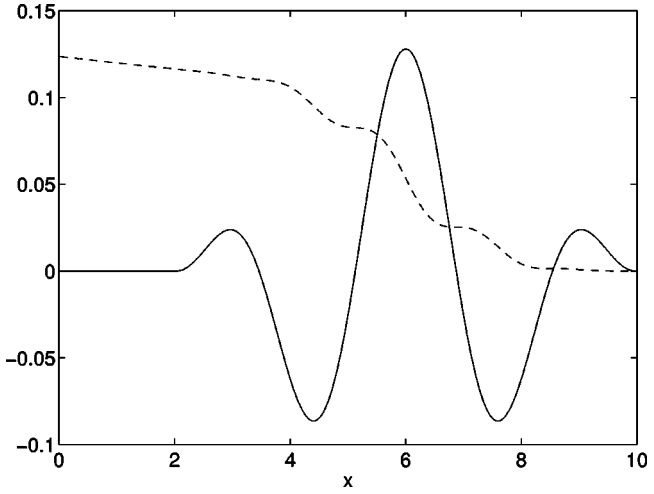

FIG. 1. Initial conditions for derivatives of the transverse metric. The solid line is $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$ and the dashed line is $2(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$. Note that $x=10$ is the center of the grid.

a single wave as given by Eq. (64) with the shape specified by our values for w_i and k_i, and a flat metric ahead of the wave, the answer is approximately 0.11. This is an upper limit, however, because the effects of colliding waves add together in a way which is hard to estimate. The initial conditions for $(D_y{}^y \pm D_z{}^z)$ are shown in Fig. 1.

B. Comparing evolution systems

1. Testing for the optimal system

We experiment with several different formulations of the Einstein equations to determine the factors involved in improving the global accuracy of 1D colliding gravitational plane wave calculations. The basic formalisms we test are the modified BM, AY, and ADM schemes of Sec. VI. In all of these schemes, using mixed variables rather than lowered variables improves accuracy significantly. We also compare alternative ways of handling the redundant variable V_x in the BM schemes. V_x can be left to evolve independently (no-reset BM), or it can be reset periodically to enforce the constraint that $V_x = D_y{}^y + D_z{}^z$ (reset BM). Results have been calculated for a range of values of the coefficient n of the energy constraint term in the extrinsic curvature evolution equations, from about -0.4 to 1.0, and in some cases for values of $n>1$. For the ADM and reset BM schemes, the results near 0 and 1 reflect the breakdown of hyperbolicity at these values of n. Results are primarily shown for $t=12$ since this is the latest time at which the physical waves are largely within the grid.

Figure 2 shows the evolution of linear combinations of metric derivatives appearing in the eigenmodes from $t=8$ to $t=12$, after the physical waves have finished colliding. In these high resolution (4000 cell) calculations, the numerical errors are negligible on the scale of the graph, and we have verified that all the different formulations seem to be converging to the same solution. The quantity $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$ is shown in Fig. 2(a). The coordinate speed of propagation can be read off the graph: it is roughly two

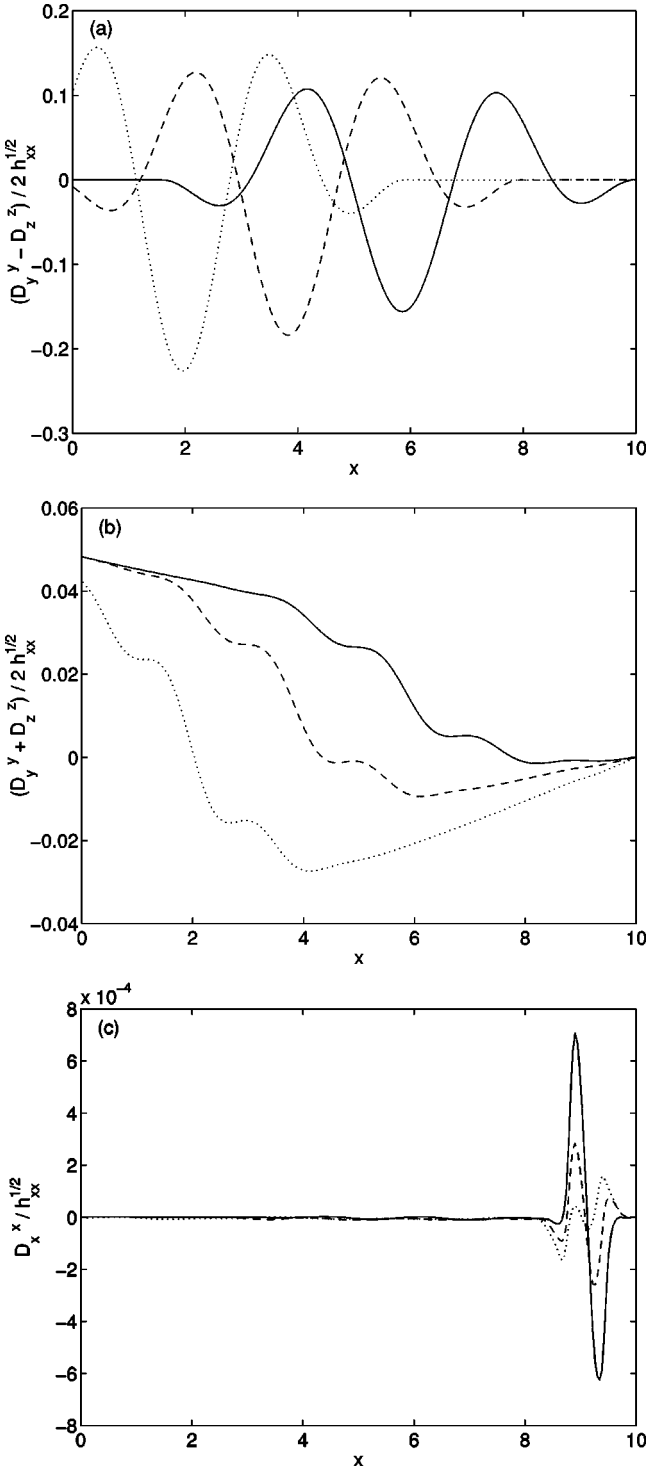

FIG. 2. Evolution of the metric derivatives. The solid line is at $t=8$, the dashed line at $t=10$, and the dotted line at $t=12$. (a) $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$, (b) $(D_y{}^y + D_z{}^z)/(2\sqrt{h_{xx}})$, and (c) $D_x{}^x/\sqrt{h_{xx}}$.

units in x for every two units of time until around $t=12$, when the coordinate speed of light starts to differ significantly from 1. Over the same range of times, the quantity $(K_y{}^y - K_z{}^z)/2$ is within about 3% of $-(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$, close to but not identical to what is expected from the left-propagating ''physical'' eigenmode. In Fig. 2(b), we see that

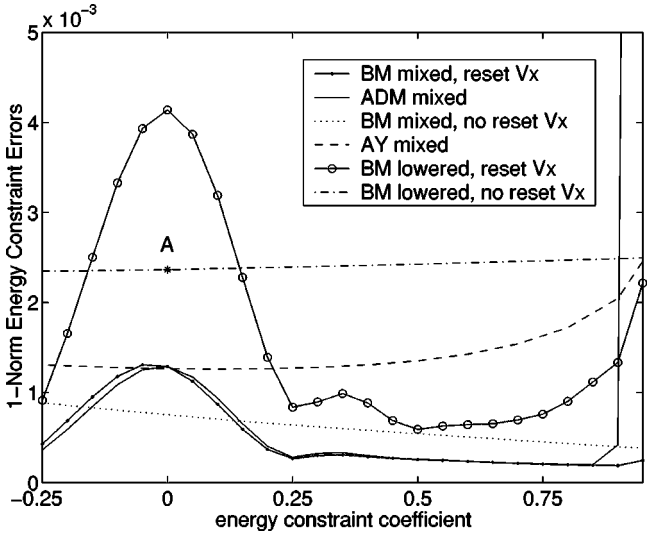

FIG. 3. One-norm errors of the energy constraint plotted against the energy constraint coefficient, n, for several different formulations of the Einstein equations. Evaluated at $t=12$ with a grid resolution of 500 cells. Note that the largest value of n plotted is 0.95. Point A is the formulation closest to the standard BM scheme [4].

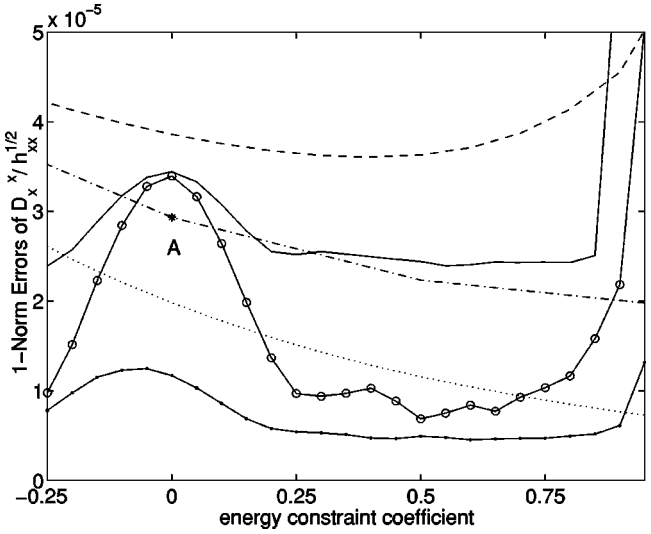

FIG. 4. One-norm errors of $D_x{}^x/\sqrt{h_{xx}}$ plotted against the energy constraint coefficient, n, for several different formulations of the Einstein equations. Evaluated at $t=12$ with a grid resolution of 500 cells. Errors are estimated from comparisons with 1000 cell calculations, assuming quadratic convergence. Note that the largest value of n plotted is 0.95. The legend is the same as in Fig. 3 *except* that the AY mixed values are multiplied by 0.045.

the steps in $(D_y{}^y+D_z{}^z)/(2\sqrt{h_{xx}})$ are associated with extrema of $(D_y{}^y-D_z{}^z)/(2\sqrt{h_{xx}})$. At these times, $(K_y{}^y+K_z{}^z)/2=(D_y{}^y+D_z{}^z)/(2\sqrt{h_{xx}})$ to the left of the physical wave. In the vicinity of the physical wave, $(K_y{}^y+K_z{}^z)/2$ has step-like features associated with steps in $(D_y{}^y+D_z{}^z)/(2\sqrt{h_{xx}})$, but ascending to the right. In the region between the waves, $(K_y{}^y+K_z{}^z)/2$ is much larger than $(D_y{}^y+D_z{}^z)/(2\sqrt{h_{xx}})$ and increases with time. Figure 2(c) shows the evolution of $D_x{}^x/\sqrt{h_{xx}}$. The prominent feature in this figure is a small residual effect (note that the scale of the graph is 10^{-4}) of the prominent feature in $K_x{}^x$ shown in Fig. 9 which survives the near cancellation of $K_x{}^x$ in the evolution equation of $D_x{}^x$ from our shift resetting condition [Eq. (25)]. Since $\alpha K_x{}^x-\beta'=0$ each time the shift is reset, this feature in $D_x{}^x/\sqrt{h_{xx}}$ tends to advect along the hypersurface normals. The generation and modification of features in $D_x{}^x/\sqrt{h_{xx}}$ is due to the different evolutions of $\alpha K_x{}^x$ and β' between gauge resettings. The feature in $K_x{}^x$ results from our lapse resetting condition, Eqs. (22) and (23), when a strong imbalance between the transverse D's and K's occurs near the center of the grid as the waves collide, creating negative values for $\mathcal{S}$. This in turn causes the limiter to take effect, which allows $K_x{}^x$ to dip in the negative direction.

To compare the overall accuracies of different formulations, we present 1-norms of the energy constraint errors in Fig. 3 and 1-norms of errors in $D_x{}^x/\sqrt{h_{xx}}$ in Fig. 4 at $t=12$ for 500 cell grids. The constraint errors are predominantly errors in the derivatives of the constraint quantities, and are insensitive to errors in the longitudinal variables. For each scheme, the 1-norm errors are plotted for a number of values of the energy constraint coefficient, ranging from -0.25 to 0.95 at 0.05 increments. Since our ADM scheme is not hyperbolic for $n=0$, the transport steps of the ADM calculations are solved with the finite difference numerical method, whereas the transport steps of the BM and AY calculations use our flux-based wave decomposition methods. The choice of numerical method makes little difference to the results.

Figure 3 identifies factors which affect accuracy as measured by 1-norm energy constraint errors. It is apparent that using mixed variables improves accuracy significantly in the BM formulation. Similar improvements occur in the AY and ADM formulations. The 1-norm energy constraint errors for ADM and the ADM-like reset BM schemes are almost identical, differing by only 1–2 % over the range $-0.25\leqslant n\leqslant 0.85$, and are minimized for $0.25\leqslant n\leqslant 0.80$. Point "A" in Fig. 3 is the formulation closest to the BM scheme as implemented in Ref. [4]. The identical formulation using mixed instead of lowered variables decreases the 1-norm energy constraint error 3.1 times. If the mixed BM system of equations is transformed into an ADM-like scheme by frequently resetting V_x, and an energy constraint coefficient of 0.5 is used, a 9.3-fold decrease in the 1-norm energy constraint error compared to point "A" is obtained. Both the ADM and the reset BM error curves peak at $n=0$, and increase rapidly as $n\rightarrow 1$, though the increase as $n\rightarrow 1$ for mixed reset BM occurs too close to $n=1$ to be apparent in Fig. 3. The rise in energy constraint errors at $n=0$ and $n=1$ reflects in part the breakdown of hyperbolicity in ADM at these values of n. The effects of this breakdown are more severe at $n=1$ than at $n=0$ because ADM is unstable for $n>1$. Momentum constraint errors are similar to or smaller than the energy constraint errors.

The 1-norm energy constraint errors in the no-reset BM schemes vary slowly for all n. These schemes are well-behaved for $n\geqslant 1$, and the errors for the mixed version continue to decrease. Despite the breakdown in the AY scheme at $n=1$, the constraint errors do not increase strongly until n

gets close to 1. The AY formulation is well-behaved for $n>1$.

Since the true value of $D_x{}^x/\sqrt{h_{xx}}$ is not known exactly, we must extrapolate to estimate the true value and calculate the 1-norm errors shown in Fig. 4. Assuming quadratic convergence, the error estimate at each grid cell of a 500 cell calculation is $\frac{4}{3}$ times the difference between the 500 and 1000 cell results. For $D_x{}^x/\sqrt{h_{xx}}$ at time $t=12$, the 500 cell errors deviate from quadratic scaling in the region $8<x<10$, where the feature in $D_x{}^x/\sqrt{h_{xx}}$ associated with the spike in $K_x{}^x$ is located. Here, our standard extrapolation underestimates the errors for the ADM and no-reset BM formulations, and overestimates the errors for the reset BM formulations, but the effects on Fig. 4 are not very significant.

Figure 4 shows that the same factors which decrease the 1-norm energy constraint errors also decrease the 1-norm errors in $D_x{}^x/\sqrt{h_{xx}}$. Specifically, using mixed variables instead of lowered variables in the no-reset BM formulation at $n=0$ decreases the 1-norm error in $D_x{}^x/\sqrt{h_{xx}}$ 1.5 times. The shapes of the curves fall into the same two classes as described above for Fig. 3. If one frequently resets V_x and sets $n=0.5$ in mixed BM, the 1-norm error decreases 6-fold from point ''A.'' The errors in $D_x{}^x/\sqrt{h_{xx}}$ in the ADM and reset BM schemes peak at $n=0$ and increase rapidly as $n\rightarrow 1$, again reflecting, in part, the failure of hyperbolicity at these values of n.

The fact that Fig. 4 is at all similar to Fig. 3 is because the $D_x{}^x/\sqrt{h_{xx}}$ errors and the constraint errors behave similarly in the region of the physical wave (this is described in detail in Sec. IX B 2). However, there are several differences between these two figures. First, the ADM curve in Fig. 4 is well above the reset BM curve. Second, the reset BM curves have smaller drop-offs from $n=0$ in Fig. 4. Third, the decreasing slopes of the no-reset BM curves are bigger in Fig. 4 than in Fig. 3. These differences are due to relatively large formulation-dependent numerical errors in $D_x{}^x/\sqrt{h_{xx}}$ in the region $8<x<10$, which contribute to the 1-norms. Recall that the evolution of the feature in $D_x{}^x/\sqrt{h_{xx}}$ in this region depends mainly on the evolutions of $K_x{}^x$ and β' between gauge resettings. We expect the numerical errors for $K_x{}^x$ to differ from those for β', since these two variables evolve by quite different equations. Further, we expect these numerical errors to be larger where $K_x{}^x$ (and β') vary rapidly (see Fig. 9). Small differences in these numerical errors among the different formulations result in large differences in numerical errors for $D_x{}^x/\sqrt{h_{xx}}$ in this region. For example, the errors in $D_x{}^x/\sqrt{h_{xx}}$ for $8<x<10$ are larger at $n=0$ than at $n=0.5$ for the mixed no-reset BM scheme, explaining the decreasing slopes in Fig. 4, and are larger for ADM than for reset BM, explaining the displacement between the mixed ADM and mixed reset BM curves.

Another difference between Figs. 3 and 4 is that the 1-norm errors of $D_x{}^x/\sqrt{h_{xx}}$ are one to two orders of magnitude higher for AY than for the other formulations, whereas the 1-norm energy constraint errors for AY and the other formulations are comparable. For example, the 1-norm error in $D_x{}^x/\sqrt{h_{xx}}$ at $n=0.5$ is about 70 times higher for the mixed AY formulation than for the mixed no-reset BM formulation, whereas the 1-norm AY energy constraint error is about 2.5 times higher. The subtraction of two numbers with large errors in Eq. (39) gives a large error for $D_x{}^x/\sqrt{h_{xx}}$ in the AY formulation. When we introduce substantial variations in h_{xx} in the initial conditions, so that $D_x{}^x$ is much larger than $(D_y{}^y+D_z{}^z)$, and h_{xx} is substantially different from 1, then equally large errors are introduced into all the formulations. In our particular calculations, the errors already present in the AY scheme fortuitously cancel the introduced errors, resulting in $D_x{}^x/\sqrt{h_{xx}}$ errors for mixed AY and mixed no-reset BM which are within a factor of 2.

These results demonstrate our ability to significantly increase the accuracy of 1D highly nonlinear colliding gravitational plane wave calculations through equation formulation. In particular, the choice of a mixed set of indices improves the accuracy of the formulations for all values of the energy constraint coefficient tested. Resetting V_x in the BM formulations creates ADM-like schemes. This is an advantage for $0.25\leqslant n\leqslant 0.80$, where ADM is hyperbolic, and a disadvantage for $n=0$ or $n=1$, where ADM is not hyperbolic.

2. Error propagation

In order to understand the variations in accuracy among the formulations, it is instructive to look at how errors vary with position, and how they propagate over time. We focus on the energy constraint errors and the errors in $D_x{}^x/\sqrt{h_{xx}}$, since they are representative of errors in the transverse and longitudinal parts of the metric, respectively. Momentum constraint errors are comparable to or less than the energy constraint errors.

Since the principal parts of the energy and momentum constraints are the derivatives of the constraint quantities, the constraint errors propagate with the same speeds as the ''constraint'' eigenmodes (which are given in Sec. VI D for BM, AY, and ADM). The constraint error propagation can also be obtained from the evolution equations for the energy and momentum constraints, which form their own closed hyperbolic system. The constraint errors propagate along the light cones for the no-reset BM and AY formulations. For ADM, the constraint errors propagate at

$$v_{ADM}=-\beta\pm\frac{\alpha}{\sqrt{h_{xx}}}\sqrt{1-n}. \tag{65}$$

We expect the constraint errors to also propagate at v_{ADM} for the reset BM scheme. Note that for $n<0$ or $0<n<1$, v_{ADM} is different from any of the other characteristic speeds of the system. We find that a separation of the constraint error speeds from the other characteristic speeds improves accuracy.

Figure 5(a) shows the energy constraint error propagation for reset BM and $n=0$. First, notice that the constraint errors are large where the physical wave is present, and that at a given location, the error decreases almost to zero when the physical wave has passed. However, there is a rapid increase in the energy constraint errors propagating with the physical wave. Second, from the graph we see that the waveform of

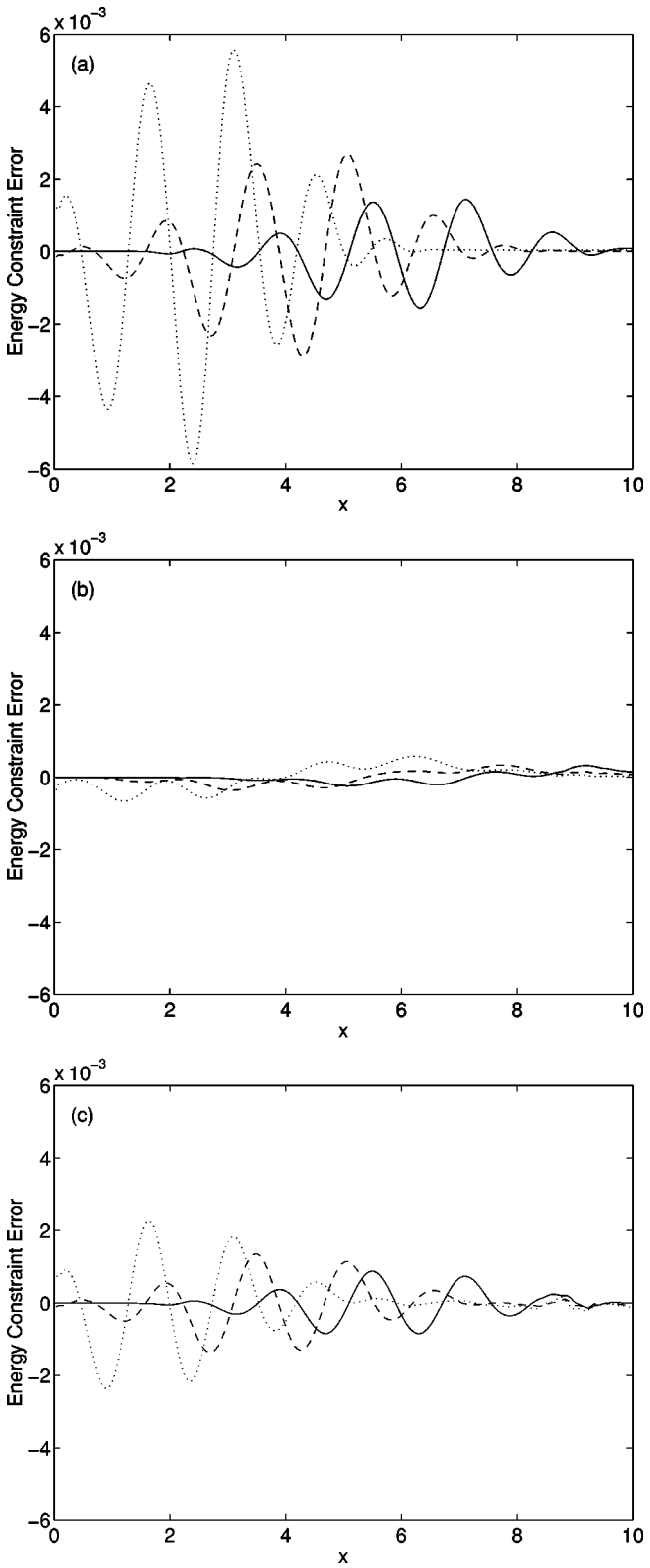

FIG. 5. Energy constraint error propagation using our mixed BM formulation. The solid line is at $t=8$, the dashed line at $t=10$, and the dotted line at $t=12$. Evaluated with a grid resolution of 500 cells, for (a) $n=0$, with V_x resetting, (b) $n=0.5$, with V_x resetting, (c) $n=0.5$, with *no* V_x resetting.

the energy constraint errors propagates at roughly unit coordinate speed, which, over the times we are considering, is approximately coordinate light speed. The energy constraint errors are predominantly errors in the derivatives of $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$. Errors in the propagation of $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ are largest where its second derivative is largest, at the corners of the steps visible in Fig. 2(b). From the energy constraint equation, the steps in $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ are associated with large values of the physical quantities $(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y-K_z{}^z)$. These physical quantities propagate at light speed, and constraint errors, once generated, propagate with the velocity of the ''constraint'' eigenmodes, which is also light speed for reset BM with $n=0$. Since new errors remain in phase with the propagating old errors, the constraint errors are continuously reinforced. Careful comparison of Fig. 2(b) with Fig. 5(a) shows that the constraint error peaks are coincident with the corners of the steps in $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ at all three times shown.

The same argument applies to AY and no-reset BM, since these formulations also have ''constraint'' mode errors propagating at light speed, but Fig. 3 shows larger errors for ADM and reset BM at $n=0$ than for AY and no-reset BM. We attribute the larger ADM and reset BM errors to the breakdown of hyperbolicity in ADM at $n=0$, so that the ''constraint'' eigenmodes, which propagate at light speed with constant amplitude, interact with the ''longitudinal'' eigenmodes through the constraint quantities. As $n\rightarrow 0$, any errors in the constraint quantities tend to produce amplified errors in the longitudinal variables, because $n(D_x{}^x+D_Q)$ and $n(K_x{}^x-\beta'/\alpha)$ are the same order of magnitude as the constraint quantities in the ADM ''longitudinal'' eigenmodes. The errors in $D_x{}^x$ and $K_x{}^x$ then feed back to the constraint quantities through the source terms of the evolution equation for $(K_y{}^y+K_z{}^z)$.

Figure 5(b) shows a dramatic decrease in both the energy constraint errors and the growth rate of the errors in the mixed reset BM formulation when $n=0.5$. Furthermore, the full waveform of the energy constraint errors does not maintain its shape as it propagates, as does the waveform in Fig. 5(a). By tracking the rightmost bump in this figure, one can determine the speed of the errors to be approximately 0.7 times light speed, which agrees with the value predicted by Eq. (65). Because the energy constraint errors lag behind the source of the errors, namely, the steps in $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$, there is not constant reinforcement and rapid growth of the errors in the region of the physical wave. This results in greater overall accuracy for reset BM than for no-reset BM at $n=0.5$, as seen in Fig. 3.

Figure 5(c) shows the energy constraint error propagation for mixed no-reset BM at $n=0.5$. In contrast to Fig. 5(b), the waveform is maintained reasonably well, and the errors grow more rapidly in time. This is because the constraint errors travel at light speed; so, as in Fig. 5(a), there is a continuous reinforcement of the errors. For no-reset BM at $n=0$, the curve is practically the same as what we show here at $n=0.5$. The error waveform has a smaller growth rate than that for reset BM at $n=0$, presumably because of the hyperbolicity of the no-reset formulation, as discussed earlier. We have also looked at the AY constraint error propagation and

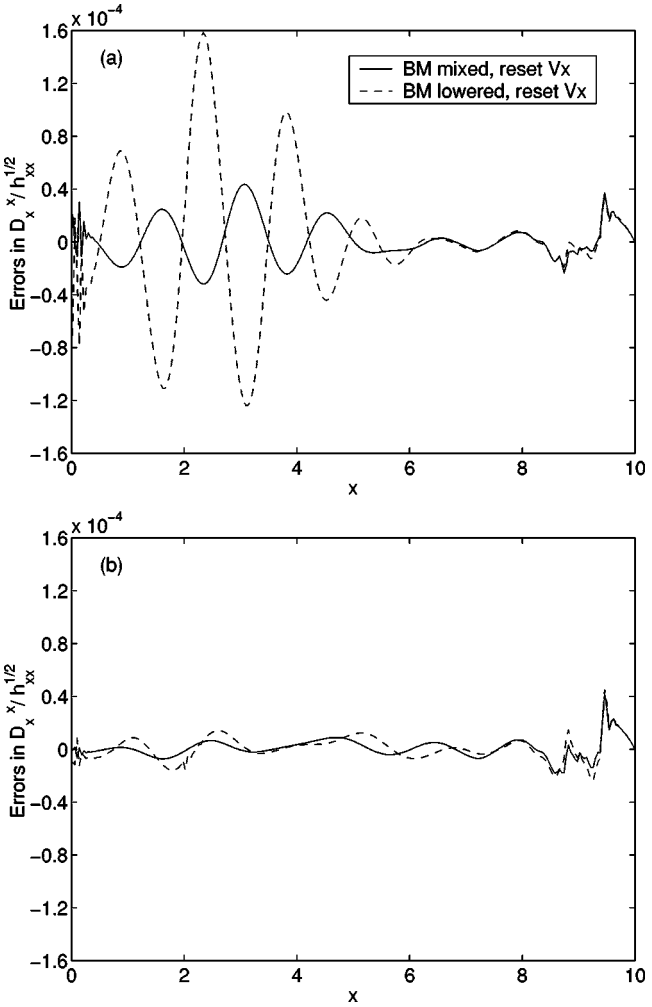

FIG. 6. Errors in $D_x{}^x/\sqrt{h_{xx}}$ versus x for our BM mixed and lowered formulations, with V_x resetting. Evaluated at $t=12$ with a grid resolution of 500 cells. Errors are estimated from comparisons with 1000 cell calculations, assuming quadratic convergence, for (a) $n=0$, and (b) $n=0.5$.

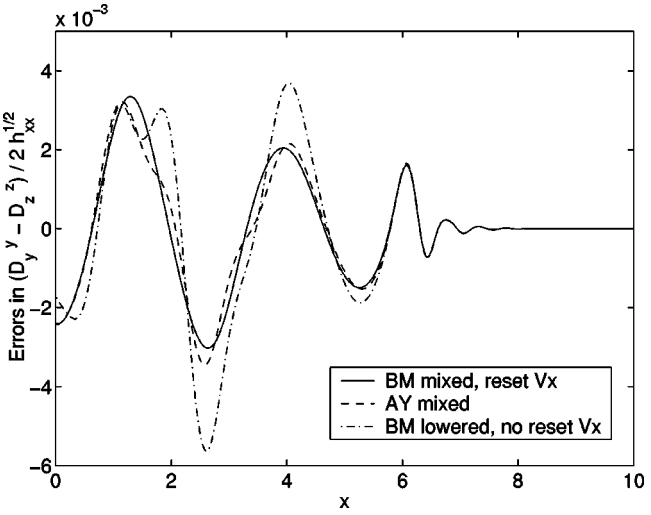

FIG. 7. Errors in $(D_y{}^y-D_z{}^z)/(2\sqrt{h_{xx}})$ versus x for different formulations. Evaluated at $t=12$ with a grid resolution of 500 cells. Errors are estimated from comparisons with 1000 cell calculations, assuming quadratic convergence, for $n=0.5$.

confirm that errors propagate at light speed for all values of n. Because the constraint errors travel with the physical waves for all n in these formulations, their 1-norm errors vary slowly with n in Fig. 3.

The errors in $D_x{}^x/\sqrt{h_{xx}}$ versus x at $t=12$ for the reset BM formulations are shown in Fig. 6, for energy constraint coefficients of (a) 0 and (b) 0.5. Figure 6(a) shows a localization of the errors in the region of the physical wave ($0\leqslant x\leqslant 6$). Figure 6(b) shows the dramatic decrease in the errors for $0\leqslant x\leqslant 6$ when $n=0.5$. The mixed no-reset errors in this region at $n=0.5$ are larger by a factor of about 2 than the mixed reset errors at $n=0.5$, and only slightly smaller than the mixed no-reset errors at $n=0$. The fact that there is a similar reduction in constraint errors and $D_x{}^x/\sqrt{h_{xx}}$ errors when going from no-reset BM to reset BM at $n=0.5$ suggests that the constraint errors are a major source of errors for $D_x{}^x/\sqrt{h_{xx}}$. From the failure of hyperbolicity in ADM at $n=0$, one expects the errors in $D_x{}^x/\sqrt{h_{xx}}$ to increase more rapidly for reset BM than the constraint errors, but we do not see clear evidence for this. The increase in errors is about the same going from $t=8$ to $t=12$, perhaps because the $D_x{}^x/\sqrt{h_{xx}}$ errors have not yet reached their asymptotic limit. The spiky errors at approximately $8.7\leqslant x\leqslant 10$ in both Figs. 6(a) and 6(b) are discussed in Sec. IX B 1.

The ADM and ADM-like reset BM systems show rapid increases in errors as $n\rightarrow 1$ in Figs. 3 and 4. The breakdown of hyperbolicity at $n=1$ results in errors in $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ becoming large compared to errors in $(K_y{}^y+K_z{}^z)$. The contribution to the energy constraint errors from errors in the derivatives of $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ increases correspondingly. Since the ''constraint'' eigenmodes propagate along the hypersurface normals at $n=1$, the errors in $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ also reinforce errors in variables which propagate along the hypersurface normals, namely, $D_x{}^x/\sqrt{h_{xx}}$, β', and D_Q.

Figure 7 shows the errors in $(D_y{}^y-D_z{}^z)/(2\sqrt{h_{xx}})$ as a function of x for three formulations at an n of 0.5. The amplitudes and shapes of the errors for the different formulations are roughly the same, because the evolution equations for $(D_y{}^y-D_z{}^z)/(2\sqrt{h_{xx}})$ are similar for the different formalisms. The curve shapes do not change significantly when one uses $n=0$ instead of $n=0.5$ because n does not enter into the evolution equations for $(D_y{}^y-D_z{}^z)/(2\sqrt{h_{xx}})$. The entire graph shown in Fig. 7 converges quadratically except for the bump around $x=6$, which converges linearly. This bump is near the trailing edge of the physical wave, where the initial conditions are not smooth enough to give second order accuracy. The errors for the formulations which are not shown are similar.

Calculations of the distribution and propagation of energy constraint errors and errors in $D_x{}^x/\sqrt{h_{xx}}$ have illuminated how resetting V_x in the BM formulations to create overall ADM-like evolutions increases or decreases accuracy and stability, depending on the value of n. Intermediate values of n separate constraint error speeds from the other characteristic speeds of the system, resulting in significant improve-

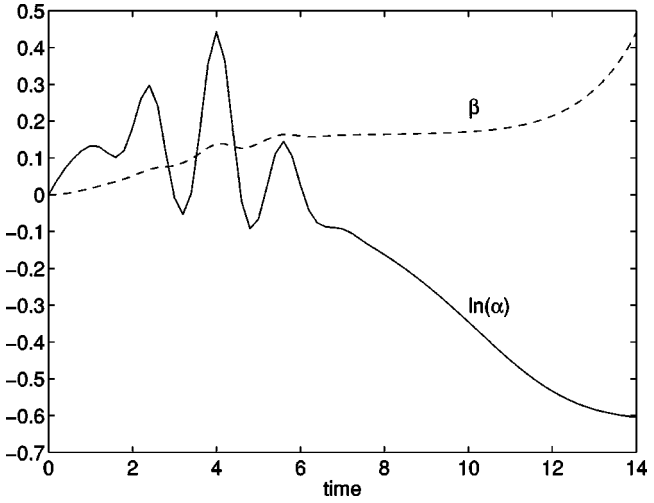

FIG. 8. The shift, β, and the logarithm of the lapse, $\ln(\alpha)$, at the left grid edge plotted against time at intervals $\Delta t=0.2$. The values of β and $\ln(\alpha)$ are fixed at zero at the grid center.

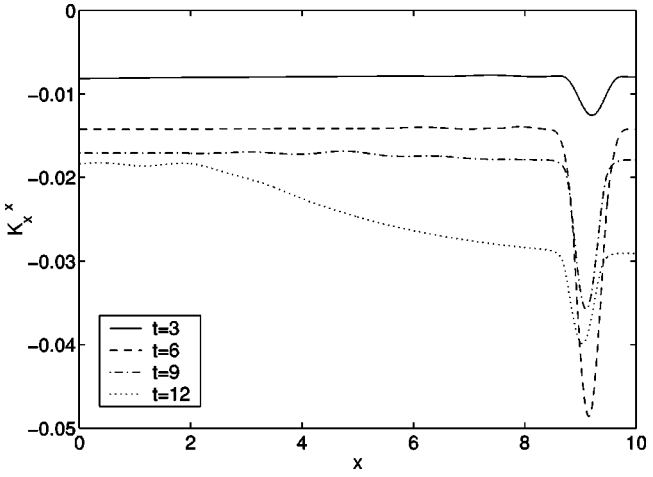

FIG. 9. $K_x{}^x$ versus x at the times indicated in the legend. $K_x{}^x$ is driven to a target value of -0.02 by our lapse resetting condition. The spike at $x\approx 9$ develops as the waves collide, due to the limiter in the equation for the lapse. The effects of the limiter are felt over a wider range of x at $t=12$.

ments in accuracy. Values of n for which ADM hyperbolicity fails result in rapid increases in errors.

The higher accuracy of reset BM compared to the other formulations when $0.25\leqslant n\leqslant 0.80$ is seen when $h_{xx}=1$ and $D_x{}^x=0$ in our initial conditions. When we introduce substantial variations in h_{xx} in the initial conditions, so that $D_x{}^x$ is at least as large as $(D_y{}^y+D_z{}^z)$, and $\sqrt{h_{xx}}$ varies by a factor of 2 to 3, the accuracy results are dominated by the errors directly associated with the variations in h_{xx}. These errors depend primarily on the amplitude of $D_x{}^x/\sqrt{h_{xx}}$, and are largest where the derivative of $D_x{}^x/\sqrt{h_{xx}}$ is largest. They are similar in all the formulations and are independent of n; thus, they tend to equalize the results from the different formulations.

C. Gauge conditions

Our gauge conditions involve both a periodic resetting of the lapse and the shift, and an evolution of the lapse and the shift between resettings. The lapse and shift are reset according to Eqs. (22) and (25), with $\Gamma=1$, and $(K_x{}^x)_T=-0.02$. They are reset at constant time intervals rather than after a specific number of time steps, to give resolution-independent results. Between resettings, β, β', Q, and D_Q are advected along hypersurface normals.

Figure 8 shows how the lapse and the shift at the left boundary of the grid vary with time. The resetting interval is $\Delta t=0.02$, while the graph is drawn from values calculated at intervals of 0.2 in t; hence, the graph contains no information about the evolution of the lapse and the shift between resettings. The lapse and the shift are fixed at zero at the grid center, $x=10$, and are more or less monotonic between the center and the edge, so the variations at intermediate values of x have a similar form, but smaller amplitude.

In Fig. 8, the lapse oscillates at the grid edge until about $t=6$. From $t=6$ to $t=12$, $\ln(\alpha)$ becomes increasingly negative until it starts leveling off around $t=12$. The lapse is determined by Eq. (22). Since we set the lapse to one and its first derivative to zero at the center of the grid, whether the lapse is greater than or less than one at the edge is determined by the sign of the second derivative of the lapse with respect to proper distance, the quantity $\mathcal{S}$ in Eq. (22). The large terms in $\mathcal{S}$ are typically $-K_y{}^yK_z{}^z$ and $D_y{}^yD_z{}^z/h_{xx}$. These nearly cancel when $D_y{}^y$, $D_z{}^z$, $K_y{}^y$, and $K_z{}^z$ are dominated by a single propagating ''physical'' wave pulse. As two such pulses collide, constructive interference in $D_y{}^y$ and $D_z{}^z$ implies destructive interference in $K_y{}^y$ and $K_z{}^z$, and vice versa, which causes rather large oscillations in the value of the lapse at the edge of the grid. Note that at $t=4$, there is maximum constructive interference of $D_y{}^y$ and $D_z{}^z$. Once the waves have largely passed through each other ($t>6$), $K_y{}^y$ and $K_z{}^z$ are large, and $D_y{}^y$ and $D_z{}^z$ are small in the region between the separating waves, so the second derivative of the lapse is strongly negative, and becomes more strongly negative as $K_y{}^y$ and $K_z{}^z$ become steadily larger. By $t=12$, the limiter in Eq. (22), which is necessary to prevent the lapse from becoming negative at the edge of the grid, largely stops further decrease of the lapse at the edge.

The behavior of the shift in Fig. 8 reflects the behavior of $K_x{}^x$ in Fig. 9, since β' is proportional to $K_x{}^x$ by Eq. (25). The resetting of the lapse keeps $K_x{}^x$ near its target value of -0.02 until $t>10$, when the limiter in Eq. (22) takes hold, allowing $K_x{}^x$ to become steadily more negative. Since the derivative of the shift is proportional to $K_x{}^x$, and the shift is clamped to zero at the center of the grid, the shift at the left edge of the grid is positive and starts getting steadily larger for $t>10$.

Recall that a negative $K_x{}^x$, in combination with our shift resetting condition, causes the hypersurface normals to point away from the computational grid at the boundaries (see Sec. V). These resetting gauge conditions are designed so that Q and β are advected out of the grid, in order to suppress the development of instabilities associated with extrapolation at the grid edge. As long as $(K_x{}^x)_T\leqslant 0$, the shift is positive at the left edge of the grid and the advection velocity is negative. However, Fig. 10 shows that when $(K_x{}^x)_T=+0.02$, so

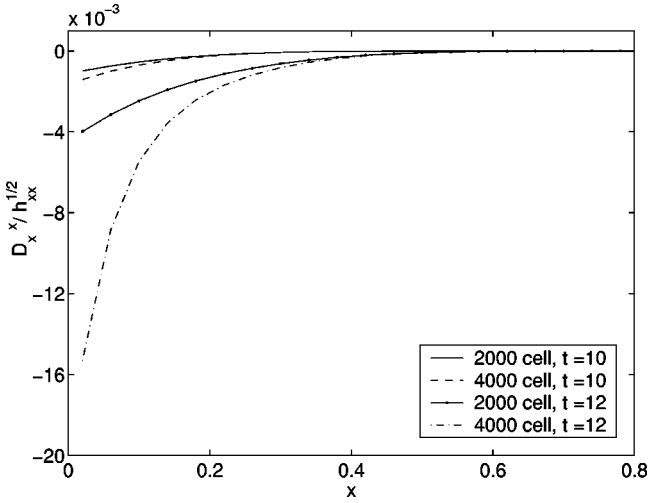

FIG. 10. Development of an instability at the left boundary in the evolution of $D_x^x/\sqrt{h_{xx}}$ when the $K_x{}^x$ target value is $+0.02$ in our lapse resetting condition. The instability is substantially stronger with the higher resolution grid. Note that the physical wave reaches the left edge at $t=10$.

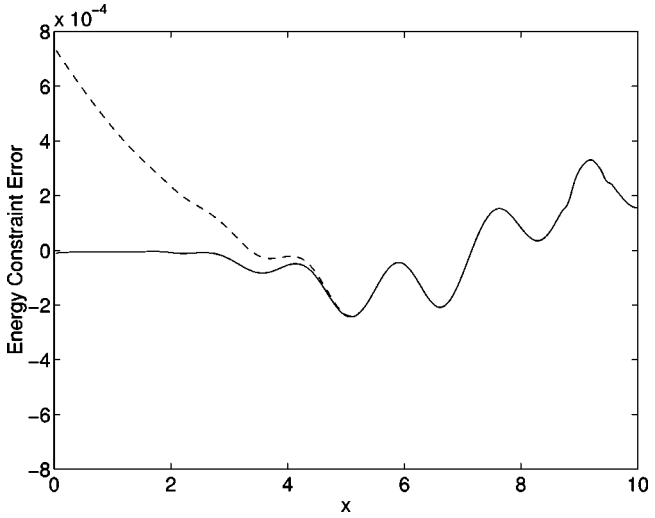

FIG. 11. Energy constraint error for our standard boundary condition, which corrects quadratic extrapolation of the variables at the boundaries using the constraint equations (solid line), versus quadratic extrapolation only (dashed line). Evaluated at $t=8$ with a grid resolution of 500 cells using our mixed BM formulation with V_x resetting and $n=0.5$.

that the shift is negative at the left edge, the solution becomes unstable at the grid edge once the physical wave reaches the edge.

D. Boundary conditions

Figure 11 compares two methods for implementing boundary conditions: quadratic extrapolation of all the variables with and without correcting the extrapolated values of $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y+K_z{}^z)$ to insure that the energy and momentum constraint equations are satisfied at the boundaries. Failure to strictly enforce the constraints at the edges of the grid results in significant constraint errors propagating over much of the grid by $t=8$.

The jitter at the left edge of Fig. 6, which graphs the errors in $D_x{}^x/\sqrt{h_{xx}}$ versus x, is due to numerical errors and the use of quadratic extrapolation in our boundary conditions. It is most prominent when the physical wave is crossing the boundary, but it is not unstable as long as the hypersurface normal at the boundary does not point into the grid (see Sec. V). We will explore ways of reducing this jitter either through the use of limiters near the boundary and/or improved extrapolation methods.

E. Numerical methods

We compare convergence results from our two flux-based wave decomposition schemes with those from the traditional finite difference approach. Tables I and II show quadratic convergence for errors in the metric derivative variables and in the momentum and energy constraints at $t=12$ using our optimal evolution scheme and flux-based wave decomposition method **I**. Quadratic convergence for the metric derivative variables is tested by comparing 1-norms of differences between results evaluated at 500 and 1000 cell resolutions, 1000 and 2000 cell resolutions, and 2000 and 4000 cell resolutions. Convergence is calculated in this way because true values are unavailable for the metric derivative variables. Constraint errors can be calculated directly so quadratic convergence is determined simply by taking the ratios of results from grid resolutions that differ by a factor of 2. There is very little difference in the results of the three numerical methods tested. Further, the 1-norm results from the three numerical methods all converge to the same answer.

Tables I and II show that our flux-based wave decomposition methods are second order convergent when the char-

TABLE I. Convergence results for errors in the metric derivative variables. Evaluated at $t=12$ using our mixed BM formalism, resetting V_x, with $n=0.5$, and flux-based wave decomposition method **I**.

	$\frac{1}{2}\left[\frac{(D_y{}^y-D_z{}^z)}{\sqrt{h_{xx}}}\right]$	$\frac{1}{2}\left[\frac{(D_y{}^y+D_z{}^z)}{\sqrt{h_{xx}}}\right]$	$\frac{D_x{}^x}{\sqrt{h_{xx}}}$
$(500-1000)/(1000-2000)$ [a]	3.98	3.99	5.09
$(1000-2000)/(2000-4000)$ [b]	3.97	3.99	4.38

[a]The 1-norm of the difference between the 500 and 1000 cell calculations is divided by the 1-norm of the difference between the 1000 and 2000 cell calculations.

[b]The 1-norm of the difference between the 1000 and 2000 cell calculations is divided by the 1-norm of the difference between the 2000 and 4000 cell calculations.

TABLE II. Convergence results for the energy and momentum constraint errors. Evaluated at $t=12$ using our mixed BM formalism, resetting V_x, with $n=0.5$, and flux-based wave decomposition method **I**.

	Energy constraint	Momemtum constraint
500/1000 [a]	4.02	3.97
1000/2000 [b]	4.00	3.98
2000/4000 [c]	3.99	3.99

[a]The 1-norm of the 500 cell calculation is divided by the 1-norm of the 1000 cell calculation.

[b]The 1-norm of the 1000 cell calculation is divided by the 1-norm of the 2000 cell calculation.

[c]The 1-norm of the 2000 cell calculation is divided by the 1-norm of the 4000 cell calculation.

acteristic matrix $\mathbf{A}(x)$ has significant spatial variation, because of the spatial dependence of the lapse and the shift [see Eq. (47) and Fig. 8]. However, in these calculations the eigenvectors of $\mathbf{A}(x)$, which depend only on $\sqrt{h_{xx}}$, are roughly constant. We have tested convergence of both our flux-based wave decomposition methods when the eigenvectors of $\mathbf{A}(x)$ vary rapidly, by introducing a 3-fold variation in $\sqrt{h_{xx}}$. Our results are still second order convergent.

X. DISCUSSION

A. Summary

We have identified ways of improving the accuracy and stability of 1D nonlinear colliding gravitational plane wave calculations through an in-depth study of equation formulations, dynamic gauge conditions, boundary conditions, and numerical methods. Three issues stand out in our study of equation formulations. The first issue which improves the accuracy of all the formulations tested is raising an index in the metric derivative and extrinsic curvature variables. The separation of constraint and physical behavior in our calculations is much simpler with mixed variables than with lowered variables. Since the dominant feature of plane wave solutions is the physical wave, and since the mixed variables have a simple linear relation to the ''physical'' eigenmodes, the mixed form gives an advantage in accuracy. In addition, the variables in the source terms of the evolution equations are in mixed form, so a cleaner system of equations results.

Second, it is advantageous if the ''constraint'' eigenmodes, which are constraint violating, propagate at different speeds from other features in the solution, so that nothing in the actual solution is constantly in step with the constraint errors. Our gauge conditions result in features which propagate on the hypersurface normals as well as on the light cones. With a multiple between 0.25 and 0.80 of the energy constraint equation added to the evolution equations for the extrinsic curvature in the ADM and reset BM schemes, the ''constraint'' eigenmodes propagate neither on the light cones, nor on the hypersurface normals, and the growth rates of constraint errors and errors in $D_x{}^x/\sqrt{h_{xx}}$ significantly decrease. However, when the amplitude of $D_x{}^x$ is large compared to $(D_y{}^y+D_z{}^z)$, formulation-independent errors associated with the derivatives of $D_x{}^x/\sqrt{h_{xx}}$ dominate the overall errors for $0.25\leqslant n\leqslant 0.80$.

Third, we find that when hyperbolicity fails in ADM, for values of the energy constraint coefficient equal to 0 or 1, the constraint errors and errors in $D_x{}^x/\sqrt{h_{xx}}$ increase much more quickly in both ADM and reset BM than they do in no-reset BM. At $n=0$, the ''constraint'' and ''longitudinal'' eigenmodes are no longer independent, causing a rapid growth of errors which travel with the physical waves. These errors decrease and stabilize, however, when the waves exit the grid. At $n=1$, the two ''constraint'' eigenmodes are not independent, causing errors in $(D_y{}^y+D_z{}^z)/\sqrt{h_{xx}}$ to increase rapidly. This results in large errors in the energy constraint and in variables which propagate along the hypersurface normals.

The key to our approach to dynamic gauge conditions for hyperbolic calculations is to maintain a simple diagonalizable hyperbolic evolution during each time step, while allowing periodic, flexible resetting of the lapse and the shift between time steps. By resetting the lapse and the shift, we can control the coordinate system in a dynamic and (in principle) arbitrary way, unconstrained by the need to maintain hyperbolicity. Our gauge resetting conditions control the longitudinal components of the extrinsic curvature and the spatial metric so as to prevent pathologies and strong gradients from developing in the hypersurfaces and spatial coordinates. Further, our gauge resetting conditions cause the hypersurface normals to point away from the grid at the edges. This helps to suppress the development of instabilities at the boundaries which are associated with features advecting along the hypersurface normals.

A careful study of boundary conditions has led us to the conclusion that it is incorrect to impose outgoing wave boundary conditions based on the eigenmodes of the hyperbolic decomposition. A substantial contribution from the incoming ''constraint'' eigenmodes is necessary to satisfy the constraint equations at the boundaries. In the presence of nonlinearities, the interaction of the ''physical'' eigenmodes with the incoming ''constraint'' eigenmodes generates an admixture of incoming and outgoing ''physical'' eigenmodes. The right and left-going gravitational wave amplitudes can be determined by projecting the Weyl tensor onto a null tetrad, and our numerical results indicate that the gravitational waves after the collision are purely outgoing according to this definition. Our boundary condition procedure consists of an accurate calculation of the incoming eigenmodes of the characteristic matrix at the boundaries by quadratic extrapolation of all the variables, and correction of the ghost cell values using the energy and momentum constraint equations. This procedure, in combination with our lapse and shift resetting conditions which insure that the hypersurface normals at the boundaries do not point into the grid, has given stable, accurate, and second order convergent results.

Finally, we have developed flux splitting numerical methods for solving hyperbolic formulations of the Einstein equations which are second order accurate for smooth solutions, even when the eigenvalues and eigenvectors of the characteristic matrix are spatially varying. These methods are based on decomposing flux differences between adjacent grid cells

into linear combinations of the eigenvectors of the characteristic matrix. We show that these methods are formally second order accurate and, in practice, second order convergent.

B. Relevance of results and future directions

Our results suggest that hyperbolicity can be a useful guide to picking equation formulations for numerical integration of the Einstein equations. When eigenvectors are incomplete, or some eigenvectors are nearly linearly dependent, some solutions to the equations will tend to grow without bound or by large factors. In 2D and 3D, the ADM and ADM-like equation formulations for non-diagonal metrics fail to have a complete set of characteristic eigenvectors for *any* value of the energy constraint coefficient. This may be at the root of some of the instabilities seen in 3D ADM codes.

However, hyperbolicity, which involves only the principal terms in the equations, is not the whole story. The constraint equations relate derivatives of some quantities (the constraint quantities) to nonlinear terms involving additional quantities, so the actual evolution of the constraint quantities for constraint-satisfying solutions may be very different from the evolution implied by the ''constraint'' eigenmodes, which satisfy equations without source terms. It seems to be advantageous if the speeds of the short-wavelength errors in the constraint quantities, which are the ''constraint'' mode eigenvalues, differ substantially from the propagation speeds of the major features in the constraint quantities. In 1D, both the ADM and the no-reset BM formulations are safely hyperbolic for $n \approx 0.5$, but the former is substantially more accurate because no-reset BM has the same speeds for the ''constraint'' eigenmodes and the features in the constraint quantities, whereas ADM has different speeds. How much of an advantage it is to have different speeds depends on the dominant source of errors. In a gauge where h_{xx} is close to 1, the dominant numerical errors in the constraint quantities are associated with features in the same quantities, but in a gauge where the dominant numerical errors are associated with features in h_{xx}, it does not make much difference whether the ''constraint'' eigenmodes and features in the constraint quantities have the same velocities. In generic black hole spacetimes, it is probably not possible to find a gauge where the physical waves dominate the metric; therefore, the separation of speeds may not make much difference in accuracy for 2D and 3D calculations of greatest physical interest.

The eigenmode decomposition in higher dimensions depends on a chosen direction of propagation, as do the combinations of the actual variables which can be identified as ''longitudinal,'' ''constraint,'' and ''physical'' quantities. In addition, putting the equations into first-order form requires a choice of derivative ordering, which has important effects on the hyperbolicity of the system. A ''directional splitting'' approach to solving the transport steps, in which $\partial_t \mathbf{q} + \partial_k \mathbf{F}^k = 0$ is solved separately for each coordinate direction k, does, at least for some choices of derivative ordering, allow the identification of ''longitudinal,'' ''constraint,'' and ''physical'' eigenmodes for one coordinate direction at a time. The decomposition involves projecting the D_{kij}'s and K_{ij}'s perpendicular to and into the constant-x^k surfaces. The eigenvectors can be constructed explicitly.

Using the mixed coordinate components $D_{ki}{}^{j}$ and $K_i{}^{j}$ as variables when the metric is not diagonal gives complicated source terms in the $D_{ki}{}^{j}$ evolution equations, because the $D_{ki}{}^{j}$ are no longer pure derivatives. Another possibility is to take as the variables the components of an orthonormal triad and the projection of the extrinsic curvature onto the triad (or the Ashtekar variables [15–17]). The triad formalism, because the extrinsic curvature tensor projected on the triad is symmetric, has fewer variables than the mixed coordinate component formalism. In neither case would the variables be simply related to the eigenvectors of the characteristic matrix when the metric is non-diagonal or the triad vectors are not along the coordinate directions; thus, there is no obvious advantage to trying to generalize our mixed variables in 1D to generic 2D and 3D calculations. However, we plan to explore these questions further.

Our approach to dynamic gauge conditions, namely, implementing a simple hyperbolic evolution during each time step, then resetting the lapse and the shift between time steps, is general. A hyperbolic evolution during each time step allows the use of flux-based wave decomposition numerical methods. A long-term strictly hyperbolic evolution, however, is destroyed in all our hyperbolic formulations when we reset the lapse and shift. Despite this, resetting the gauge can be used to increase the accuracy and stability of our 1D calculations. Achieving the same goal in 2D or 3D calculations will be more difficult. There is no one longitudinal direction, and there is not enough gauge freedom to control the longitudinal components of the metric and extrinsic curvature for all directions. What gauge resetting conditions are most effective in 2D and 3D remains an open question.

Our boundary condition results indicate that simple outgoing wave boundary conditions based on the eigenmodes of the characteristic matrix are not valid in general and are inconsistent with the constraint equations. On the other hand, extrapolation (in particular, quadratic extrapolation) is a numerically dangerous procedure. Although one can use constraint extrapolation to constrain certain linear combinations of the variables, this technique is likely to be less effective in higher dimensions than in 1D because the number of variables increases more rapidly than the number of constraints. Using an outgoing wave boundary condition based on the Weyl tensor to further constrain the variables is problematic in the general case. While the peeling theorem [36] in an asymptotically flat spacetime does show that the incoming wave projection of the Weyl tensor falls off much more rapidly than the outgoing wave projection in the wave zone, one cannot assume that the outgoing wave dominates in 3D numerical relativity calculations, which typically have to be truncated at best in the inner part of the wave zone. The ingoing part of the Weyl tensor contains non-radiative quasistatic quadrupole moment contributions which cannot be assumed to vanish.

The flux-based wave decomposition numerical methods we have presented for hyperbolic formulations of the Einstein equations are completely generalizable, and should

prove useful for calculations in black hole and Brill wave spacetimes where the eigenvectors and eigenvalues have significant spatial variation. We hope to extend our exploration of hyperbolic methods to include the use of limiters and upwind differencing. Limiters will be used to suppress short wavelength numerical instabilities. With upwind differencing, one does not need ghost cells at the apparent horizon boundary of an excised black hole, where the eigenmodes are purely ingoing.

In conclusion, we have developed basic methodologies for hyperbolic formulations of the Einstein equations, which improve accuracy and stability in 1D, and which we think merit further exploration in the context of numerical relativity calculations of more substantial physical interest.

ACKNOWLEDGMENTS

This paper presents the results of one phase of research supported by the NASA Graduate Student Researchers Program under Grant No. NGT5-50298 in collaboration with the Jet Propulsion Laboratory. LTB also gratefully acknowledges partial support from NSF grant DMS-9803442. We thank R. LeVeque and F. Estabrook for discussions and for careful reading of the manuscript. Gratitude is also extended to D. Bale, J. Rossmanith, M. Alcubierre, E. Seidel, and K. Thorne for insightful discussions, and to the Numerical Relativity group in Potsdam, Germany for their warm generosity during an extended visit. Finally, J.M.B. would like to thank the Institute for Theoretical Physics at UCSB for their support during the initial stages of this research.

APPENDIX: SECOND ORDER ACCURACY OF FLUX-BASED WAVE DECOMPOSITION

In the wave propagation stage of solving for the time evolution of a hyperbolic system, the equations being solved are exact conservation laws [Eq. (54)]. The integral form of the conservation law can be used to update the average $\mathbf{q}$'s in the ith cell,

$$\begin{aligned}\mathbf{q}_i^{n+1} &\equiv \frac{1}{\Delta x}\int_{cell\ i} \mathbf{q}^{n+1} dx \\ &= \mathbf{q}_i^n - \frac{1}{\Delta x} \\ &\quad \times \left[\int_{t_n}^{t_{n+1}} \mathbf{F}_{i+(1/2)} dt - \int_{t_n}^{t_{n+1}} \mathbf{F}_{i-(1/2)} dt \right], \end{aligned} \tag{A1}$$

where $i \pm \frac{1}{2}$ denotes values on the cell boundaries. We base our discussions of accuracy on Taylor series expansions in both t and x, assuming that the cell size Δx and the time step $\Delta t = t_{n+1} - t_n$ are the same order. Time and spatial derivatives are related by the wave propagation equation, Eq. (54), and $\mathbf{F}$ is assumed to depend on $\mathbf{q}$ and x. Second order accuracy means that the error in one time step in $\mathbf{q}_i^{n+1}$ decreases faster than $(\Delta t)^2$ as Δt and Δx go to zero.

Expanding $\mathbf{F}$ in a Taylor series in time,

$$\begin{aligned}\int_{t_n}^{t_{n+1}} \mathbf{F}_{i\pm(1/2)} dt &\approx \Delta t \ \ \mathbf{F}^n_{i\pm(1/2)} + \frac{1}{2}\Delta t^2 \partial_t \mathbf{F}^n_{i\pm(1/2)} \\ &\quad + \frac{1}{6}\Delta t^3 \partial_t^2 \mathbf{F}^n_{i\pm(1/2)}. \end{aligned} \tag{A2}$$

In order that $\mathbf{q}_i^{n+1}$ be accurate to second order while neglecting the last term in Eq. (A2), $\partial_t^2 \mathbf{F}$ must be continuous in x. From the evolution Eq. (55) for $\mathbf{F}$, assuming the dependence of $\mathbf{F}$ on $\mathbf{q}$ is analytic, second order continuity in $\partial_t^2 \mathbf{F}$ is guaranteed by second order continuity in $\partial_x^2 \mathbf{F}$. The resulting equation is

$$\int_{t_n}^{t_{n+1}} \mathbf{F}_{i\pm(1/2)} dt \approx \Delta t \mathbf{F}^n_{i\pm(1/2)} - \frac{1}{2}\Delta t^2 \mathbf{A}_{i\pm(1/2)} (\partial_x \mathbf{F})^n_{i\pm(1/2)}. \tag{A3}$$

Let $\mathbf{F}_i$ be the cell-centered values of $\mathbf{F}$, which differ from $\mathbf{F}(\mathbf{q}_i, x_i)$ by a term of order $\Delta x^2 \partial_x^2 \mathbf{F}_i$. Then expanding $\mathbf{F}$ in a Taylor series in x about the cell-center value,

$$\begin{aligned}\mathbf{F}^n_{i\pm(1/2)} &\approx \frac{1}{2}(\mathbf{F}^n_i + \mathbf{F}^n_{i\pm1}) - \frac{1}{8}\Delta x^2 (\partial_x^2 \mathbf{F})^n_i \\ &\equiv \mathbf{F}^n_i \pm \frac{1}{2}\Delta \mathbf{F}^n_{i\pm(1/2)} - \frac{1}{8}\Delta x^2 (\partial_x^2 \mathbf{F})^n_i. \end{aligned} \tag{A4}$$

Since the second-order error terms are the *same* for $\mathbf{F}_{i+(1/2)}$ and $\mathbf{F}_{i-(1/2)}$, and change the same way when $\mathbf{F}_i$ and $\mathbf{F}_{i+1}$ are replaced by $\mathbf{F}(\mathbf{q}_i, x_i)$ and $\mathbf{F}(\mathbf{q}_{i+1}, x_{i+1})$, they cancel in the difference of the flux integrals in Eq. (A1) and can be omitted, *if* $\partial_x^2 \mathbf{F}$ is continuous.

To obtain a second-order accurate contribution to $\mathbf{q}_i^{n+1}$ from the second term in Eq. (A3), it is sufficient that $\mathbf{A}_{i\pm(1/2)}$ and $(\partial_x \mathbf{F})^n_{i\pm(1/2)}$ be accurate to first order. In method **II**, $\mathbf{A}_{i\pm(1/2)}$ is approximated to first order by $\frac{1}{2}(\mathbf{A}_i + \mathbf{A}_{i\pm1})$. A Taylor expansion of $(\partial_x \mathbf{F})^n_{i\pm(1/2)}$ gives, to first order,

$$(\partial_x \mathbf{F})^n_{i\pm(1/2)} \approx \frac{\Delta \mathbf{F}^n_{i\pm(1/2)}}{\Delta x}. \tag{A5}$$

Flux-based wave decomposition consists of decomposing the flux differences $\Delta \mathbf{F}^n_{i+(1/2)}$ into a sum over eigenvectors of the characteristic matrix, and similarly for $\Delta \mathbf{F}^n_{i-(1/2)}$. Different versions evaluate these eigenvectors at different locations. From the point of view of smooth solutions, the preferred location is to base the decomposition of $\mathbf{A}_{i\pm(1/2)}$ at the respective cell boundaries. The decomposition of $\Delta \mathbf{F}^n_{i+(1/2)}$ into a sum of right eigenvectors of the matrix $\mathbf{A}_{i+(1/2)}$ can be written

$$\Delta \mathbf{F}^n_{i+(1/2)} = \mathbf{R}_{i+(1/2)} \mathbf{\Gamma}_{i+(1/2)}, \tag{A6}$$

where $\mathbf{R}_{i+(1/2)}$ is a matrix whose columns are the right eigenvectors $\mathbf{r}^p_{i+(1/2)}$, and the column vector $\mathbf{\Gamma}_{i+(1/2)}$ consists of the coefficients of the decomposition, $\gamma^p_{i+(1/2)}$. Then

$$\mathbf{A}_{i+(1/2)} \Delta \mathbf{F}^n_{i+(1/2)} = \mathbf{R}_{i+(1/2)} \mathbf{\Lambda}_{i+(1/2)} \mathbf{\Gamma}_{i+(1/2)}, \tag{A7}$$

where $\mathbf{\Lambda}_{i+(1/2)}$ is the diagonal matrix of eigenvalues of $\mathbf{A}_{i+(1/2)}$. Let $\lambda^p_{i+(1/2)}$ be the pth eigenvalue, or wavespeed, and let $\mathbf{W}^p_{i+(1/2)} \equiv \gamma^p_{i+(1/2)} \mathbf{r}^p_{i+(1/2)}$, where $\mathbf{r}^p_{i+(1/2)}$ is the pth eigenvector, i.e., the pth column of $\mathbf{R}_{i+(1/2)}$. Equation (A7) can be written

$$\mathbf{A}_{i+(1/2)} \Delta \mathbf{F}^n_{i+(1/2)} = \sum_p \lambda^p_{i+(1/2)} \mathbf{W}^p_{i+(1/2)}. \tag{A8}$$

Combining Eqs. (A3)–(A5) and (A8), and noting that $\Delta \mathbf{F}^n_{i+(1/2)} = \Sigma_p \mathbf{W}^p_{i+(1/2)}$, gives

$$\int_{t_n}^{t_{n+1}} \mathbf{F}_{i+(1/2)} dt \approx \Delta t \left[\mathbf{F}^n_i + \frac{1}{2} \sum_p \mathbf{W}^p_{i+(1/2)} - \frac{1}{2} \frac{\Delta t}{\Delta x} \sum_p \lambda^p_{i+(1/2)} \mathbf{W}^p_{i+(1/2)} \right], \tag{A9}$$

and a similar expression for the flux integral at $i - \frac{1}{2}$. In Eq. (A1), these give for $\mathbf{q}^{n+1}_i$ an expression which is easily shown to be equivalent to the substitution of Eqs. (58) into Eq. (57). Thus, we have shown that method **II** of Sec. VIII B 1 is second-order accurate for $\partial_x^2 \mathbf{F}$ continuous.

Method **I** is based on expressions similar to Eq. (A9) for the flux integrals, but the eigenvectors in $\mathbf{W}_{i+(1/2)}$, for instance, are derived from $\mathbf{A}_i$ for the left-going eigenmodes at the $i + \frac{1}{2}$ interface and derived from $\mathbf{A}_{i+1}$ for the right-going eigenmodes. In the last term in Eq. (A9), the eigenvalues $\lambda^p_{i+(1/2)}$ are approximated to first-order accuracy as the average of the eigenvalues associated with the adjacent cells, but the $\mathbf{W}_{i+(1/2)}$ are only zeroth-order accurate. The first-order corrections to the eigenvectors are the same at the $i \pm \frac{1}{2}$ interfaces, if the same eigenmodes are left and right-going at each interface and the first derivatives of the eigenvectors are continuous. (Recall that in method **I**, the sign of the averaged eigenvalue at a given interface determines the direction of the eigenmode at that interface.) The second condition normally follows from continuity of $\partial_x \mathbf{A}$, but may fail if some of the eigenmodes are nearly degenerate. With this qualification, the first-order corrections in this term will cancel between the two interfaces, as long as the same eigenmodes are left- and right-going at each interface. When the averaged eigenvalues at the cell interfaces do change sign, continuity of the first derivatives of the eigenvalues implied by continuity of $\partial_x \mathbf{A}$ means $\lambda^p_{i+(1/2)}$ and $\lambda^p_{i-(1/2)}$ are both of order Δx. The last term in Eq. (A9), which becomes the last term in Eq. (58), is then second-order accurate by itself. The terms first-order in Δt in Eqs. (57) and (58) will differ in second order from their values in method **II**, but combine by construction to give

$$\sum_L \mathbf{W}^L_{i+(1/2)} + \frac{1}{2} \left[\sum_R \mathbf{W}^R_{i+(1/2)} - \sum_L \mathbf{W}^L_{i+(1/2)} \right] = \frac{1}{2} \left[\sum_R \mathbf{W}^R_{i+(1/2)} + \sum_L \mathbf{W}^L_{i+(1/2)} \right] \equiv \frac{1}{2} \Delta \mathbf{F}^n_{i+(1/2)} \tag{A10}$$

and

$$\sum_R \mathbf{W}^R_{i-(1/2)} - \frac{1}{2} \left[\sum_R \mathbf{W}^R_{i-(1/2)} - \sum_L \mathbf{W}^L_{i-(1/2)} \right] = \frac{1}{2} \Delta \mathbf{F}^n_{i-(1/2)}. \tag{A11}$$

The result for $\mathbf{q}^{n+1}_i$ is the same as method **II** through second order, as long as the eigenvectors are smooth as discussed above. While eigenmodes are degenerate for many hyperbolic formulations of the Einstein equations, the eigenvectors can be chosen with the required smoothness.

Both methods, when the smoothness conditions are satisfied, are equivalent through second order to the Lax-Wendroff finite difference scheme presented in Sec. VIII B 2.

[1] É. É. Flanagan and S. A. Hughes, Phys. Rev. D **57**, 4535 (1998).
[2] M. W. Choptuik, Phys. Rev. Lett. **70**, 9 (1993).
[3] H. Friedrich and A. D. Rendall, Lect. Notes Phys. **540**, 127 (2000).
[4] C. Bona, J. Massó, E. Seidel, and P. Walker, gr-qc/9804052.
[5] R. J. LeVeque, *Numerical Methods for Conservation Laws* (Birkhäuser Verlag, Basel, 1992).
[6] O. A. Reula, Living Rev. Relativ. **1998-3** at http://www.livingreviews.org/
[7] 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.
[8] C. Bona, J. Massó, E. Seidel, and J. Stela, Phys. Rev. D **56**, 3405 (1997); Phys. Rev. Lett. **75**, 600 (1995); C. Bona and J. Massó, *ibid.* **68**, 1097 (1992).
[9] A. Anderson and J. W. York, Jr., Phys. Rev. Lett. **82**, 4384 (1999).
[10] T. W. Baumgarte and S. L. Shapiro, Phys. Rev. D **59**, 024007 (1999).
[11] M. Shibata and T. Nakamura, Phys. Rev. D **52**, 5428 (1995).
[12] M. Alcubierre *et al.*, Phys. Rev. D **62**, 044034 (2000).
[13] M. Alcubierre, B. Brügmann, D. Pollney, E. Seidel, and R. Takahashi, Phys. Rev. D **64**, 061501 (2001).
[14] S. Frittelli and O. A. Reula, Phys. Rev. Lett. **76**, 4667 (1996).
[15] H. Shinkai and G. Yoneda, gr-qc/0103031.
[16] H. Shinkai and G. Yoneda, Class. Quantum Grav. **17**, 4799 (2000).
[17] G. Yoneda and H. Shinkai, Class. Quantum Grav. **18**, 441 (2001).
[18] G. Yoneda and H. A. Shinkai, Phys. Rev. D **63**, 124019 (2001).
[19] L. E. Kidder, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. D **64**, 064017 (2001).
[20] J. Balakrishna *et al.*, Class. Quantum Grav. **13**, L135 (1996).
[21] L. Rezzolla, A. M. Abrahams, R. A. Matzner, M. E. Rupright, and S. L. Shapiro, Phys. Rev. D **59**, 064001 (1999).

[22] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, England, 1986).
[23] R. J. LeVeque, J. Comput. Phys. **131**, 327 (1997).
[24] R. J. LeVeque (private discussion).
[25] Y. Choquet-Bruhat and T. Ruggeri, Commun. Math. Phys. **89**, 269 (1983).
[26] Y. Choquet-Bruhat and J. W. York, C. R. Acad. Sci., Ser. I: Math. **321**, 1089 (1995).
[27] E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962); **4**, 998(E) (1963).
[28] R. W. MacCormack, in *Proceedings of the Second International Conference on Numerical Methods in Fluid Dynamics*, edited by M. Holt, Lecture Notes in Physics Vol. 8 (Springer-Verlag, Berlin, 1971).
[29] P. Anninos, K. Camarda, J. Massó, E. Seidel, W.-M. Suen, and J. Towns, Phys. Rev. D **52**, 2059 (1995).
[30] D. S. Bale, R. J. LeVeque, S. Mitran, and J. A. Rossmanith, ''A wave propagation method for conservation laws and balance laws with spatially varying flux functions,'' report (2001).
[31] R. J. LeVeque and M. Pelanti, J. Comput. Phys. **172**, 572 (2001).
[32] R. J. LeVeque, CLAWPACK software. http://www.amath.washington.edu/ claw
[33] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman and Co., New York, 1973), pp. 943–973.
[34] U. Yurtsever, Phys. Rev. D **37**, 2790 (1988).
[35] U. Yurtsever, Phys. Rev. D **38**, 1731 (1988).
[36] R. Penrose, Proc. R. Soc. London, Ser. A **284**, 159 (1965).