Inspiraling black-hole binary spacetimes: Challenges in transitioning from analytical to numerical techniques
Yosef Zlochower, Hiroyuki Nakano, Bruno C. Mundim, Manuela Campanelli, Scott Noble, and Miguel Zilhão
Phys. Rev. D 93, 124072 — Published 29 June 2016
DOI: 10.1103/PhysRevD.93.124072
I. INTRODUCTION

The field of numerical relativity (NR) has progressed at a remarkable rate since the breakthroughs of 2005 [1–2], when it first became possible to simulate the late-inspiral, plunge, merger, and ringdown of black-hole binaries (BHBs). Recently, Lousto and Healy [3] completed a long-term 50 orbit precessing BHB simulation using the moving punctures approach and Szilagyi et al. [5] completed the longest BHB simulation to date: the last 176 orbits for a non-spinning, intermediate mass ratio \( m_1/m_2 = 1/7 \), BHB using the generalized harmonic approach. This is a remarkable achievement, but the scaling of the inspiral time with the initial separation \( T \sim D^4 \) means that evolving a binary through the long inspiral is prohibitive expensive, even for highly efficient codes. Such a simulation becomes even more expensive when one is interested in performing long-term dynamical evolutions of relativistic magnetohydrodynamics (MHD) circumbinary disks around inspiraling supermassive BHBs (SMBHBs). This is because the circumbinary gas can exhibit significant secular variations on the time scale of hundreds to thousands of binary orbits.

In order to make these long-term simulations possible, our group developed a complementary approach to treat dynamical BHB spacetimes. In a series of papers [3–10] we used an analytic spacetime that is an approximate solution to the Einstein’s field equations in the inspiral regime to describe the evolution of the accretion disks surrounding the binary and each of the individual BHs.

Our initial approach [6], was one in which relativistic effects were present but relatively small. In the situation when gravity is weak \( [r_g/r = GM/(rc^2) \ll 1] \) and motions are slow \( [(v/c)^2 \ll 1] \), the post-Newtonian (PN) approximation gives a very good description of spacetime. One can then simply construct a PN metric which takes energy loss from the binary into account, accurately modeling both the mass loss and inspiral of the binary [11]. Using a spacetime accurate to 2.5PN order (i.e., including terms up to \( \sim (r_g/r)^{5/2} \)), but describing the binary orbital evolution to 3.5PN, we demonstrated that circumbinary disks can track the inspiral of a SMBHB until the binary practically reaches the relativistic merger regime [6]. The shortcoming of this approach was that the PN metric was not valid very close to the BHs, and consequently, we excised any material that fell within 1.5 binary separations. This prevented us from studying the dynamics of the gas all the way down through the horizons of each BH.

In a more recent paper [8], we extended the metric to cover the full BHB spacetime up to the rapid plunge state. We did this by extending the framework established in Refs. [12–14] for constructing a spacetime metric valid for initial data, i.e., a metric accurate for all spatial points but in a very small time interval, to develop a metric valid for arbitrary times. In this approach, the near zone (NZ), i.e., a zone well outside the two BHs, but less than a gravitational wavelength from the binary, is still described using a PN expansion. In the far zone (FZ), i.e., farther than one wavelength from the binary, the metric is described by a post-Minkowskian (PM) expansion. Finally, near each BH, i.e., in the inner zone (IZ), the metric is described using a perturbed Kerr (here Schwarzschild) BH. The metrics covering the different
zones are smoothly stitched together using asymptotic expansions and transition functions in their overlapping regions of validity.

This approach allows us to follow inspiraling SMBHBs over hundreds to thousands of binary orbits, the timescale on which gas accumulates, without having to solve the Einstein’s equations numerically. The numerical advantage here is that the numerical timestep is limited by fluid characteristic speeds, rather than the much faster speed of light (this advantage is diminished if we want to evolve gas right near the horizons).

Here, we explore the possibility of using a hybrid approach, i.e., use the analytic metric for the long inspiral down to separations where our global spacetime metric is still valid, and then transition to a full numerical simulation using the analytic spacetime as initial data (the use of PN techniques to generate initial data for BHs was first developed in Refs. [14–18]). The way to do this is to convert our approximate spacetime prescription into suitable initial data for 3 + 1 NR evolutions, evolve the data forward in time, and compare the orbital evolution, test particle trajectories, and gravitational radiation output with our approximate solution.

Perhaps more well known is the complimentary approach of combining PN and other analytical techniques with numerical waveforms to generate highly accurate hybrid waveforms. Many authors have explored this and we refer the reader to [5, 19–30] and references therein.

The main motivation of this paper is to develop techniques to smoothly transition from an analytically evolved spacetime to a numerically evolved one. By smooth, here we mean that all families of geodesics passing through the transition region will have continuous second derivatives. Here we are considering test particle trajectories as stand-ins for fluid trajectories. In particular, if there is a jump in the second derivative of the fluid, we can expect a quasi-equilibrium fluid configuration to shock and therefore require a re-equilibration that may take longer than the inspiral time. Of course, bulk binary dynamics are important too. Therefore, we want the physics of the inspiral (rate, orbital frequency) to be as unaffected by the transition as possible.

We note that the use of PN techniques to generate consistent initial data (i.e., data with the correct radiation content) provides the final ingredient proposed in the “Lazarus Approach” to generating waveforms [31, 32]. The proposal there was to transition from PN to NR techniques and then from NR to perturbative techniques. When using the global analytic metric as initial data, the resulting initial data are essentially equivalent (there is only a small difference in the NZ/FZ transition function and the two metric prescriptions coincide at $t = 0$) to the initial data proposed in Johnson-McDaniel et al. [14], and first evolved in Reifdenberger and Tichy [33]. Reifenberger and Tichy compared evolutions of Bowen-York data [34] to several different analytic initial data constructions, including Johnson-McDaniel et al. Our work here extends upon the work of Reifenberger and Tichy in several ways. (i) We use the full 4-dimensional metric of [8] to compare the dynamics of the numerically evolved metric with the analytic one. (ii) We evolve binaries with separations large enough that the PN metric and binary dynamics are expected to be accurate, and (iii) we find techniques to ameliorate the inaccuracies associated with evolving these data that were discovered by Reifenberger and Tichy. These inaccuracies arise from both constraint violations, due to the fact that our global metric solves the Einstein’s field equations only approximately, and inaccuracy in the PN orbital angular momentum and inspiral rate, which can lead to eccentricity in the numerical binary evolution.

For the current work, we start the full numerical simulations when the binary is separated by $D = 20M$ and evolve for six orbits. As shown in Ref. [35], where the authors there explored numerical simulations of Bowen-York data at separations ranging from $D = 100M$ to $D = 20M$, there is good agreement between the predictions of PN theory and numerical simulations at $D = 20M$. Additionally, simulations starting at $D = 20M$ down to merger are possible with our current codes, as demonstrated in Ref. [4] (such a simulation would require approximately 1 million CPU hours on an AMD Opteron machine).

This paper is organized as follows. In Sec. II we review how the analytic BHB inspiral metric is constructed, as well as how it is used to generate 3 + 1 initial data. In Sec. III we describe the techniques we used to numerically evolve the spacetime metric. In Sec. IV we provide details on how the simulations were performed and the key outcomes of these simulations. In Sec. V we compare the results from the numerical simulation at separations $\sim 20M$ with the predictions of PN theory. Finally, in Sec. VI we discuss our results both in terms of the accuracy of the binary dynamics (e.g., inspiral rate and orbital frequency) and in terms of gravitational waveform generation.

Throughout this paper, we use the geometric unit system, where $G = c = 1$, with the useful conversion factor $1M_{\odot} = 1.477 \text{ km} = 4.926 \times 10^{-6} \text{ s}$.

II. ANALYTIC BHB INSPIRAL METRIC

In this paper, we restrict our analysis to non-spinning BHs in quasi-circular orbits. In this context, it is useful to provide a brief review here of our approximate solution to the Einstein’s field equations of a BHB spacetime in the inspiral regime [8]. The inclusion of spins, both aligned [7] and unaligned in this spacetime framework will be the subject of future studies.

This framework was first introduced in Refs. [12, 13] as initial data for BHB evolutions, and was generalized in Ref. [8] to be a full BHB spacetime. In this framework, the spacetime is constructed by asymptotically matching metrics in three different zones characterizing three different spacetime regions of validity for different analytic
metrics: (i) a FZ where the spacetime can be described by a two-body perturbed flat spacetime with outgoing gravitational radiation and where retardation effects are fully accounted for; (ii) a NZ which is less than one GW length from the center of mass of the binary (but not too close to each BH) that is described by a PN metric (this includes retardation effects at a perturbative level and binding interactions between the two BHs); and (iii) IZs that are described by perturbed Schwarzschild (or Kerr) BHs. The full spacetime is then constructed by smoothly transitioning from zone to zone in the so-called buffer zones (BZs). A schematic diagram of these zones, and a table describing where the zone boundaries are located are provided in Fig. 1 and Table I (these figure and table were also presented in Refs. [7, 8]).

In the sections below, we will refer to these initial data as the second-order analytical metric. It is constructed by asymptotically matching a 2.5PN metric in the NZ (the matching is only done for terms up through $O(v/c)^3$) to a Schwarzschild metric with quadrupole (and its time derivatives) and octupole tidal deformations in the IZ. As explained in details in Ref. [3] this matching is approximate in the sense that it does not lead to a formal second-order asymptotic matching in all metric components for all times. However, as demonstrated there, it does lead to a significant improvement against a lower-order analytic metric, the first-order analytical metric, which is constructed by asymptotically matching a 1PN metric in the NZ (only terms of order $O(v/c)^2$ are matched) into a Schwarzschild metric with quadrupole tidal deformations. The matching for this first order metric is exact. The metric in the FZ is constructed from the PM expansion over a flat spacetime with source multipolar decomposition, where the source multipoles are expanded in the PN approximation up to 2.5PN. Note that in the PM formalism the PN metric in the NZ and the multipolar metric in the FZ are formally asymptotically matched up to 2.5PN in the NZ-FZ BZ. The precise orders used for the calculation of the metric pieces composing these analytical metrics are given in Table II.

### III. TECHNIQUES

We evolved the BHB initial data using the LazEv [36] implementation of the moving puncture approach [2, 3] with the conformal function $W = \sqrt{\chi} = \exp(-2\phi)$ suggested by Ref. [37] and the Z4 [38, 40] and BSSN [41, 43] in the global metric. The virial theorem implies that $m/r$ is taken to be $O(v^2/c^3)$ and the $O$ symbol denotes the highest order term included in the expansion. Here $g_{tt}$ and $g_{rr}$ refer to the diagonal components of the metric, the rest are the off-diagonal components. Note that the version of the first-order metric used in Refs. [8, 9] differed from this table in that $g_{tt}$ was only $O(v/c)^2$ there and here we use resumination techniques on both the first and second order metrics (in Refs. [8, 9] resummation was only used for the second-order metric).

### TABLE I. Location of the inner, near, and far zones, as well as the buffer zones joining them. Here $r_{in}$ and $r_{out}$ are the approximate inner and outer boundaries of a given zone, $m_i$ is the mass of BH $i$, $r_i$ is the coordinate distance to BH $i$, $r_{12}$ is the binary separation, and $\lambda$ is the wavelength of gravitational radiation emitted by the binary.

<table>
<thead>
<tr>
<th>Zone</th>
<th>$r_{in}$</th>
<th>$r_{out}$</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>IZ BH1 ($r_1$)</td>
<td>0</td>
<td>$\ll r_{12}$</td>
<td>-</td>
</tr>
<tr>
<td>IZ BH2 ($r_2$)</td>
<td>0</td>
<td>$\ll r_{12}$</td>
<td>-</td>
</tr>
<tr>
<td>NZ ($r_A$)</td>
<td>$\gg m_A$</td>
<td>$\ll \lambda$</td>
<td>-</td>
</tr>
<tr>
<td>FZ ($r$)</td>
<td>$\gg r_{12}$</td>
<td>$\infty$</td>
<td>-</td>
</tr>
<tr>
<td>NZ-NZ BZ</td>
<td>-</td>
<td>-</td>
<td>$m_A \ll r_A \ll r_{12}$</td>
</tr>
<tr>
<td>NZ-FZ BZ</td>
<td>-</td>
<td>-</td>
<td>$r_{12} \ll r \ll \lambda$</td>
</tr>
</tbody>
</table>
evolution systems. Here we use the conformal covariant Z4 (CCZ4) implementation of Ref. [40]. Note that the same technique has been recently applied to the evolution of binary neutron stars [43, 44]. For the CCZ4 system, we again used the conformal factor $W$. We used centered eighth-order finite differencing for all spatial derivatives, a fourth-order Runge Kutta time integrator, and both fifth and seventh-order Kreiss-Oliger dissipation [46].

Our code uses the EINSTEIN TOOLKIT [47–49] / CACTUS [50] / CARPET [51, 52] infrastructure. The CARPET mesh refinement driver provides a “moving boxes” style of mesh refinement. In this approach, refined grids of fixed size are arranged about the coordinate centers of both holes. The CARPET code then moves these fine grids about the computational domain by following the trajectories of the two BHs.

We use AHFINDERDIRECT [53] to locate apparent horizons. We also use the Antenna code [32] to calculate the Weyl scalar $\psi_4$.

We measure the distance between the two BHs using the simple proper distance or SPD. The SPD is the proper distance, on a given spatial slice, between the two BH apparent horizons as measured along the coordinate line joining the two centers. As such, it is gauge dependent, but still gives reasonable results (see Ref. [35] for more details).

To obtain initial data, we use eighth-order finite-differencing of the analytic global metric to obtain the 4-metric and all its first derivatives at every point on our simulation grid. The finite-differencing of the global metric to obtain the finite-differencing of the analytic global metric to obtain the details).

![FIG. 2. The metric function $W$ at $t = 0$ plotted versus $x$ showing the effects of stuffing the BH interior and the same function at $t = 165M$ (but plotted versus $y$). The more regular shape of $W$ near the center of the BH at $t = 165M$ is typical of moving punctures simulations (note that the puncture is offset from the $y$-axis by 0.01M). For reference, the function $W$ for a Bowen-York puncture simulation (solid curve) when the puncture cross the $x$ axis for the second time is included (the plot of the Bowen-York data has been shifted). Note how at the stuffed $W$ appears to evolve to be very similar to the standard trumpet $W$ typical of puncture initial data.]

The initial data parameters for our BHB simulations are given in Table III. To evolve the second-order analytical data we used the following grid structure. The coarsest grid spanned $0 \leq x \leq 3200M$, $-3200M \leq y \leq 3200M$, and $0 \leq z \leq 3200M$ (we used $\pi$-rotation symmetry and $z$-reflection symmetry). The refinement levels were centered on the two BHs with half-widths of 1600, 800, 440, 220, 110, 55, 25, 10, 5, 2, and 0.75. In the figures below we denote the resolution of the coarsest grid by $h$. Our lowest resolution runs had a coarsest resolution of $h = h_0 = 32M$. We increased the resolution by successive factors of 1.2 for the higher resolution runs. Our
TABLE III. Initial data parameters. $m_1$ and $m_2$ are the masses of the two BHs, $D$ is the orbital separation, $\delta \Omega_{orb}$ and $\delta \tau$ are the modifications to the 3.5PN orbital frequency and inspiral required to reduce the eccentricity, $\Xi$ is a scale factor (see text), and $r_H$ is the measured horizon radius.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1/M$</td>
<td>0.5</td>
</tr>
<tr>
<td>$m_2/M$</td>
<td>0.5</td>
</tr>
<tr>
<td>$D/M$</td>
<td>20.00</td>
</tr>
<tr>
<td>$\delta \Omega_{orb}$</td>
<td>$7.88515 \times 10^{-6}$</td>
</tr>
<tr>
<td>$\delta \tau$</td>
<td>$-1.54103 \times 10^{-4}$</td>
</tr>
<tr>
<td>$r_{\text{max}}/M$</td>
<td>0.25</td>
</tr>
<tr>
<td>$r_H/M$</td>
<td>0.484</td>
</tr>
<tr>
<td>$r_{\text{min}}/M$</td>
<td>0.05</td>
</tr>
<tr>
<td>$\Xi$</td>
<td>800</td>
</tr>
</tbody>
</table>

standard choice, which we used for all long-term runs shown below used a medium resolution of $h = h_0/1.2^3$. The highest resolution run had $h = h_0/1.2^3$.

To demonstrate the smoothness of the transition from analytical to numerical evolution we evolve a set of test particles using both the second-order analytic metric and the numerical evolution of the second-order analytic data. Note that at $t = 0$ the 4-metrics associated with the two evolution schemes are geometrically identical (i.e., only differ by a coordinate transformation implicit in using different choices for the lapse and shift at $t = 0$). However, because the evolution schemes are different, the two spacetimes will have different effective stress energies (i.e., $G_{ab}$ will differ for the two spacetimes) even at $t = 0$. Thus, even if the two spacetimes were initially in the same coordinate system, higher-order time derivatives (third and higher) of the geodesics will not agree. Thus we only expect continuity of the force acting on the geodesics as we transition from analytical to numerical evolutions of the metric. As shown in Fig. 3 we do see a relatively smooth transition at early times with the two sets of geodesics initially agreeing quite well and then deviating more significantly at later times. Importantly, this latter deviation is due to two effects, differences in the later time coordinate systems and differences in the curvature. Plots of the same geodesics from the CCZ4 and BSSN evolutions are nearly identical.

Figure 3 provides evidence that the transition from analytical to numerical evolutions is sufficiently smooth that no sudden impulses are imparted to timelike geodesics. However, we still need to demonstrate that the subsequent dynamics are accurate. This will require that the constraint violations do not significantly affect the dynamics of the binary and that the binary remains quasicircular.

A. Initial Hamiltonian and momentum constraint violations

Evolutions of second-order analytical data with BSSN were first performed in Tichy [33], where, like we do here, they looked at mass conservation, inspiral time, anomalous eccentricity, and constraint violations (see also Ref. 17). The conclusion there, as well as here, is that residual constraint violations lead to relatively large errors in the subsequent dynamics.

Our expectation is that inaccuracies in the second-order analytic metric will decrease as the binary separation is increased. To test this assumption, we need a measure of the constraint violation that is related to the dynamics of the binary. Since we can interpret violations of the Hamiltonian constraint as an unphysical matter field on our spacetime, a natural measure of the degree of violation is the total amount of unphysical matter compared to the total amount of physical mass (in this case the total amount of physical mass is $\approx 1M$). One subtlety we have to contend with is that both positive and negative mass densities are dynamically important and there is no reason to expect their respective contributions to the total error will cancel. Thus we consider two measures of the unphysical mass given by,

\begin{align*}
    m_{\text{unphys}} &= \frac{1}{16\pi} \int \mathcal{H} \sqrt{\gamma} dv, \quad (5) \\
    m_{\text{abs}} &= \frac{1}{16\pi} \int |\mathcal{H}| \sqrt{\gamma} dv, \quad (6)
\end{align*}

where $\mathcal{H}$ is the Hamiltonian constraint violation and the
FIG. 4. The unphysical mass of the binary versus binary separation $D$ for the second-order analytical data (i.e., initial data) as measured using the integrals $m_{\text{unphys}}$ and $m_{\text{abs}}$. The horizontal axis is in units of $D/M$, where $D$ is the separation of the binary. The larger mass is $m_{\text{abs}}$. Note that while a significant amount of unphysical matter is present, it is spread out such that only a fraction of it is absorbed by the BHs (see Fig. 5). Note also that there is a near equal amount of positive and negative mass (which is why $m_{\text{unphys}}$ is about a factor of 20 smaller than $m_{\text{abs}}$). Rather than plotting $m_{\text{unphys}}$, we plot $-m_{\text{unphys}}$ since the net unphysical mass is actually negative.

The integral is over the cube centered on the origin with side length 400$M$ excluding the interiors of the horizons. The former gives the net unphysical mass in the spacetime, while the absolute value in the latter ensures that positive and negative mass densities do not cancel.

We show both of these measures of unphysical mass versus separation in Fig. 4. Note that there is a near equal amount of positive and negative mass (which is why $m_{\text{unphys}}$ is about a factor of 20 smaller than $m_{\text{abs}}$). The magnitude of $m_{\text{unphys}}$ decreases with binary separation as roughly $(D/M)^{-1.8}$, while $m_{\text{abs}}$ decreases at rate of $(D/M)^{-1.6}$. The masses were calculated at $t = 0$ for three resolutions. In all cases, the truncation errors for the highest resolution corresponded to an uncertainty in second or higher significant digit in the mass.

At a separation of $D = 20M$, we find that $m_{\text{unphys}} = 0.001M$, while $m_{\text{abs}} = 0.033M$ (Bowen-York data for a $D = 20M$ binary solved using the TwoPunctures [5N] code with $40^3$ collocation points gives $|m_{\text{abs}}| < 7 \times 10^{-7}M$). Note that $m_{\text{abs}}$ increases much more rapidly than the power-law prediction with decreasing separation for $D < 25M$, but $m_{\text{unphys}}$ is only 30% larger than the power-law prediction at $D = 20M$. It is important to note that the BHs will not absorb this much mass as their cross sections are quite small.

An important result from Fig. 4 is that while the unphysical mass tends to zero at infinite binary separation, for practical purposes, it is never small in the regime where we would use NR evolutions. Thus we need a way of removing the unphysical mass from the system.

We also examine how the quantity of unphysical matter ($m_{\text{abs}}$) depends on the locations of the inner/near buffer zones. For our runs, we used the transition parameters of Ref. [14], which were optimized for a separation of $D = 10M$. Using these parameters, we find that $m_{\text{abs}}$ is 0.033$M$. By optimizing the parameters to reduce $m_{\text{abs}}$, we reduce this by only 4% to 0.032$M$. Interestingly, while $m_{\text{abs}}$ is reduced, the constraint violations are more concentrated near the two BHs and thus allowing for more absorption of constraint violating matter by the BHs. Importantly, the constraint violations cannot be significantly reduced by moving the locations of the zone boundaries, since they are fixed analytic functions of the masses and separation delimiting overlapping regions of validity for different metric approximations.

It is important to not only determine how much unphysical matter is present, but also where it is located. To this end, we plot the Hamiltonian constraint violations on the equatorial plane for BHBs at separations of $D = 20M$ and $D = 100M$ using both the standard second-order analytical metric described in Sec. [11] and the first-order version. The main difference between the two is described in Table [11] The Hamiltonian constraint violations show a clear improvement as we switch from the first order to the second order analytical metric, indicating that it is the low PN order which dominates the error. Perhaps unexpectedly, even at a separation of $D = 100M$, the first order metric has $m_{\text{abs}} = 0.02M$, while the second order metric has $m_{\text{abs}} = 0.001M$. Examining Fig. 5 we see that the constraint violations are concentrated in extended clouds well outside the horizons in the buffer zone between the inner and near zones. Most importantly, the first order metric shows a strong shell of high constraint violation surrounding the two BHs. The second-order metric, on the other hand, has a lower amplitude, more diffuse cloud of constraint violation that is less likely to be absorbed.

B. Effects of the Hamiltonian and momentum constraint violations

In this section we examine how the numerical evolution scheme can compound or mitigate issues associated with initial constraint violations. To this end, we evolved the second-order analytical data using both the BSSN formulation [11][49] and the constraint damping CCZ4 approach.

Our initial explorations of the dynamics of the second-order analytical data were based on the BSSN and CCZ4 systems. As shown in this section, we find that the CCZ4 is uniformly better than BSSN in evolving data with non-trivial constraint violations.

One of the most important differences between the BSSN and CCZ4 evolutions is in the horizon mass conservation. As shown in Fig. 3 the mass conservation of the BHs was relatively poor for BSSN and substantially better for CCZ4. In the figure, we see that the BSSN run showed an initial increase in mass of $10^{-3}M$, followed by a mass loss of similar magnitude. While a change of 2

- mass
FIG. 5. The Hamiltonian constraint violation in the vicinity of the two BHs in the binary for binary separations of $D = 20M$ (top) and $D = 100M$ (bottom). First order metrics are shown to the left and second order metrics to the right. The very small white circles at the centers of the BHs are the horizons.

parts in 1000 may seem small, the effect of this mass change on the orbital trajectory is quite large.

To determine the cause of the lack of conservation of the (apparent) horizon mass, we compare the time derivative of the horizon mass $dM_H/dt = dM_{1,2}/dt$ ($M_1 = M_2$ by symmetry) with the average value of the constraint on the horizons $\mathcal{H}_H$ and the flux of constraint violation into the horizon $\mathcal{C}_H$ (since the spacetime around the two horizons is identical by symmetry, we only plot the constraint violation averaged over one of the BHs). We define $\mathcal{H}_H$ and $\mathcal{C}_H$ as

$$\mathcal{H}_H = \frac{\oint \mathcal{H} \sqrt{\sigma} dA dB}{\oint \sqrt{\sigma} dA dB}, \quad (7)$$

$$\mathcal{C}_H = -\oint \mathcal{C}^i n_i \sqrt{\sigma} dA dB, \quad (8)$$

where $\mathcal{H}$ is the Hamiltonian constraint violation, $\mathcal{C}^i$ is the momentum constraint violation, $n_i$ is the unit (outward) normal to the horizon, $\sqrt{\sigma} dA dB$ is the proper area element on the horizon, and the integrals are performed over the surface of the horizon.

In Fig. 7, we show the constraint violation averaged over the individual horizons for both BSSN and CCZ4.

FIG. 6. The individual apparent horizon masses for a BHB with initial separation of 20M for both CCZ4 (black) and BSSN (red) evolutions of the second-order analytical data. The BSSN curves show very large oscillations, while the CCZ4 curves show a much smaller linear growth. The large mass oscillations in the BSSN run are due to absorption of constraint violations. Note how the effects changing the evolution system from BSSN to CCZ4 are much larger than truncation error effects.
The Hamiltonian constraint averaged over the horizon as a function of time. This linear log plot shows the absolute value of the constraint violation for the BSSN and CCZ4 evolutions of the second-order analytic metric, as well as the evolution of Bowen-York data (BY) using CCZ4. The BSSN, CCZ4, and BY data points are colored according to the sign of the constraint violation. The lower plot shows the early time behavior.

A large positive violation is observed at early times for BSSN, which is followed by a negative constraint violation. This links well with the initial increase in horizon mass for BSSN, which is followed by a later time decrease. The CCZ4 constraints are a factor of 100 smaller and do not appear to be correlated with the CCZ4 horizon mass. Because of these results, all of our long-term simulations used CCZ4.

Aside from an overall positive (constant) numerical factor, plots of $\mathcal{H}_H$, $\mathcal{C}_H$, and $dM_H/dt$ are nearly identical for BSSN (see Fig. 8). This means that all three are strongly correlating (i.e., $dM_H/dt \propto \mathcal{C}_H \propto \mathcal{H}_H$). This provides a compelling argument that it is the constraint violations that cause the horizon masses to fluctuate. On the other hand, for CCZ4, there is no compelling correlation between $dM_H/dt$ and the (much smaller) constraint violations (see Fig. 9).

Finally, we examine how the constraint violations in the bulk of the simulation domain behave with time. As shown in Fig. 10, the $L^2$ norms of the constraint violations for CCZ4 and BSSN evolutions behave quite differently (here we restrict the $L^2$ norm to the volume inside a ball of radius $30M$ and outside the two horizons so that the norm is dominated by constraint violations relatively close to the binary). The CCZ4 constraints fall to a much lower level (about a factor of 1000 smaller for the Hamiltonian and a factor of 100 for the momentum) than BSSN. For comparison purposes, we also performed equivalent evolutions with standard Bowen-York initial data [34].

As shown in Fig. 11 we see clear convergence of the $L^2$ norm of the momentum constraint violation to zero for CCZ4, while the Hamiltonian constraint violation, though small, seems to bottom out at about $10^{-8}$. The
Hamiltonian constraint for an equivalent Bowen-York run bottoms out at roughly half this value. For this convergence check, we only ran the highest resolution run for 350M due to computational costs.

The amount of unphysical mass that the BHs can absorb depends not only on the amount of unphysical matter, but also on the dynamics of the unphysical matter. For BSSN, the constraint violations largely stay in place (and can therefore be accreted) due the presence of a zero-speed constraint mode in BSSN [39], while for CCZ4 the constraints quickly leave the vicinity of the BHs. The very different behavior of BSSN and CCZ4 is shown in Fig. 12.

The overall efficacy of using CCZ4 to drive the constraint violations to zero can be measured by examining in detail how well the horizon masses are conserved. As shown in Figs. 13 and 14, there is a relatively strong linear trend in the mass that, while converging to a small value, is substantially larger than the Bowen-York result. Here we also see a significant advantage to using higher-order dissipation. Note that even with the highest resolution runs, the horizon mass increase is an order of magnitude larger than for Bowen-York data evolved with CCZ4. Since the Bowen-York data were evolved with the same evolution system and grid structure, it appears that there are peculiarities associated with the analytic initial data driving the mass increase (note, as seen in Fig. 9 absorption of constraint violation seems not to be the cause of this mass increase). One possibility which we have not explored in detail here is that the methods used to staff the BHs may be affecting the mass conservation. The other candidate would be residual constraint violations. Even for our highest resolution run, the average constraint violation on the horizon surface at late time was 50% larger than for Bowen-York (the global constraint violations were, on average, an order of magnitude larger, as shown in Fig. 10). As observed in Ref. [9], differences in accuracy of the spacetime at this level will likely not be important for MHD simulations.

C. Eccentricity

As shown in Fig. 8 the transition from analytical to numerical evolutions is relatively smooth, and, as shown above, evolutions with CCZ4 drive the constraint violations down to acceptable levels (i.e., within a factor of 10 of the levels obtained by evolving the constraint satisfying Bowen-York data). The last step required for a successful continuation of the evolution is to ensure that the binary remains quasicircular (the PN inspiral used to generate the data is quasicircular). To accomplish this, we applied the eccentricity reduction procedure of Ref. [59] to our data (we found that we needed to set $M\Omega_{\text{orb}} = 7.88515 \times 10^{-6}$ and $\delta t = 1.54103 \times 10^{-4}$). After three iterations, we were left with a residual eccentricity of $e = 0.002$, which was small enough for this test (see, Fig. 15). In Fig. 15 we show the SPD versus time for both CCZ4 and BSSN evolutions of the eccentricity-reduced data, we also show a CCZ4 evolution of the original data. From the figure, the non-physical dynamics (overall increase in radius) of the BSSN evolution is ap-
FIG. 12. The Hamiltonian constraint on the \(xy\) plane near the two BHs at \(t = 200M\) for the BSSN (left) and CCZ4 (right) evolutions of the second-order analytical data starting at a separation of \(D = 20M\). The scale is logarithmic and goes from \(10^{-7}\) to \(10^{-3}\). The absolute value of the constraint violations for the BSSN simulation match closely the violations on the initial slice, while the CCZ4 violations are three orders of magnitude smaller. In each plot the interiors of the BHs have been masked out. The constraint violations at \(t = 0\) are given in the top right plot in Fig. 5.

FIG. 13. The mass of the individual horizons versus time for the CCZ4 simulations using the standard fifth-order dissipation, seventh-order dissipation, and fifth-order dissipation of Bowen-York data. The linear trend in the mass, while converging to a small value, is substantially larger than that for Bowen-York.

FIG. 14. The rate of horizon mass increase versus time. Here D5 indicates that fifth-order dissipation was used, while D7 indicates that seventh-order was used.

D. Waveform

The waveforms presented below are relatively short (due to the expense of running the simulation to merger, which would take about 6 months on 80 Opteron cores). We will be comparing the numerical waveforms to PN waveforms. For these “short” runs, the dominant error is due to finite extraction radius. In Fig. 16 we show a “late” segment of the waveform extracted at \(r = 800M\), \(r = 1600M\), a linear extrapolation (in \(l = 1/r\)) to \(r = \infty\) from these two waveforms, as well as the extrapolation to \(r = \infty\) using the perturbative approach of Refs. [60, 61]. As expected, the dominant errors due to finite radius are phase errors. Finally, in Fig. 17 we show the waveform...
To have a more gauge invariant measure of the accuracy of the evolution we compare the waveform (as extracted at 1600M) with the 3.5PN prediction for quasi-circular orbits [62] (similar to what was done in Refs. [35] and [24]). All waveforms are shown in Fig. [19].

When extracting at $r = 1600M$, we get very good agreement between the raw $(\ell = 2, m = 2)$ mode of $\psi_4$ and the extrapolation to infinity using the techniques of Ref. [60]. Note that the numerical waveform prior to the burst of radiation is purely a function of the initial data. The initial data used the 3.5PN equations of motion, thus the agreement in the frequency at early time with 3.5PN is expected. On the other hand the PM metric in the FZ contains terms up to 2.5PN order, which naturally leads to a lower order approximation for the wave amplitude since it not only depends on the orbital parameters but also on the metric perturbation order. After the initial data burst, the waveform becomes noisier but the dynamics of the binary with the predictions of PN.

In Fig. [18] we show the SPD versus time and PN separation versus time. Since the SPD at $t = 0$ is larger than 20M we translate the SPD vertically. Note that the SPD is not expected to be equal to the PN separation. The SPD includes effects due to the non-flatness of the spatial metric and measures how distant the two horizons are to each other, while the PN separation extends from the center of one BH to the other and the proper separation corresponding to this would not be finite. While it is interesting that the numerical SPD matched the PN separation reasonably well, these are not gauge invariant quantities. We also show the SPD for an equivalent Bowen-York simulation (evolved with BSSN) first reported in Ref. [35]. The Bowen-York run also had the eccentricity reduction procedure applied.

V. COMPARISON TO PN

To gauge the accuracy of our transition from analytical to numerical evolutions, we compare the subsequent

(post-initial burst) extracted at $r = 800M$ for the resolutions $h = h_0/1.2^2$ and $h = h_0/1.2^3$. We also show the extrapolation of these to $r = \infty$, and an extrapolation to $r = \infty$ using perturbative techniques. Both the standard extrapolation in $r$ and the perturbative approach give very similar waveforms.

FIG. 15. The SPD versus time as calculated using a CCZ4 evolution of the new data both before and after the eccentricity reduction procedure. For reference, a BSSN evolution of the low-eccentricity data is also given. Constraint violation drive the eccentricity for this latter case.

FIG. 16. The $(\ell = 2, m = 2)$ mode of the waveform ($r\psi_4$) extracted at $r = 800M$ and $r = 1600M$, a linear extrapolation of these to $r = \infty$, and an extrapolation to $r = \infty$ using perturbative techniques. Both the standard extrapolation in $r$ and the perturbative approach give very similar waveforms.

FIG. 17. The $(\ell = 2, m = 2)$ mode of the waveform ($r\psi_4$) extracted at $r = 800M$ at two resolutions. Both the raw waveform and the extrapolation to infinity are shown. The dominant error is the extrapolation error which manifests itself predominately as a phase error.
FIG. 18. The SPD and PN separation versus time as calculated using a CCZ4 evolution of the new data and an older BSSN evolution of Bowen-York data. The SPD is always larger than the coordinate (and hence PN) separations. We shift the SPDs downward by 3.25M so that they agree with initial PN separation at $t = 0$.

agreement with 3.5PN is still quite good. The numerical waveform amplitude, however, seems to be closer to the average of 1.5PN and 3.5PN.

One important note is that the PN waveform given in the initial data is slightly out of phase with the resulting numerical waveform, as shown in Fig. [20]. That is to say, after translating the waveform in time by $r^*$ (the tortoise coordinate of the extraction observer), the PN and NR waveforms agree quite well for the part of the waveform after the initial burst has hit the extraction sphere (in the plot this would be from $t = 0$ to $t = 2400M$). However, prior to this burst arriving at the observer (in the plot, prior to $t = 0$), the PN and NR waveforms are out of phase by 0.255 radians. This initial part of the NR waveform is produced by the far-zone metric in the initial data, while the latter part of the waveform is produced by subsequent fully nonlinear binary dynamics. This will have repercussions if one wants to smoothly attach a PN waveform to the numerical waveform. It is important to note that other than a translation by the tortoise coordinate $r^*$ corresponding to the extraction radius, the NR and PN waveforms have not been translated.

The phase error itself can be explained by how we construct the metric in the far zone. In the far zone, the metric at some point at a (coordinate) distance $r$ from the origin depends on the dynamics of the binary at a retarded time given by the light propagation time from the binary to that point. We use the expression $t_{\text{ret}} = t - \left[ (r + M) + 2M \log \left( \frac{r + M}{2M} - 1 \right) \right]$, where $M$ is mass of the spacetime. We thus find that for a given waveform frequency $\omega$, using the flat space retarded time will introduce a phase error of approximately

$$\delta \phi = \omega \left[ M + 2M \log \left( \frac{r + M}{2M} - 1 \right) \right].$$

Since the binary’s orbital period here is $M\Omega_{\text{orb}} \approx 0.01$, and the ($\ell = 2, m = 2$) mode of the waveform has twice this frequency, we expect a phase error introduced by the flat space retarded time of $\approx 0.287$ rad, which is reasonably close to our measured phase error of 0.255 rad.

One final note concerns the amplitude of the initial data pulse in the waveform. As seen in Fig. [21] the ini-
potential pulse of radiation is suppressed relative to equivalent Bowen-York data. At $r = 1600M$, the suppression roughly a factor of 2, while at the $r = 400M$ extraction radius, the suppression is closer to a factor of 3. This is mostly due to the fact we have an initial data which models the astrophysical BHB system better and therefore possess less spurious radiation content when compared to the conformally flat BY initial data. Here, because the resolution of the grid at $r = 1600M$, the high-frequency gauge pulse (near $t = -500M$) is completely dissipated away. The high-frequency components of the initial data pulse are similarly suppressed.

VI. DISCUSSION

In order to perform accurate GRMHD simulations of gas accreting onto a BHB, including the minidisks around each BH, we need a spacetime that is accurate for the entire lifetime of the binary, i.e., from the slow inspiral at extremely large separation all the way to merger and relaxation of the remnant BH. Our goal, therefore, is to produce a 4-dimensional metric that is accurate outside the horizons at all times, of sufficient smoothness that timelike geodesics vary smoothly (i.e., are $C^\infty$) except at a single transition time where they are $C^2$, and the subsequent binary dynamics should match PN predictions to a high degree in the vicinity of this transition time. To do this, we extended an analytic metric that is accurate when the binary’s separation is $D \gg M$ by continuing the evolution using fully nonlinear numerical techniques for closer separations.

The main questions we addressed here concerned how we can accurately transition from using an analytically evolved spacetime metric to a fully nonlinear numerically evolved metric that describes the binary during the rapid plunge and merger. At the transition, we used the analytical spacetime metric to construct initial data for a subsequent numerical evolution (as was previously done in Reifenberger and Tichy [33]). Our work builds upon Reifenberger and Tichy in two main ways. We start from an analytic spacetime that can be extended arbitrarily far into the past, and we can thus compare dynamics of particles pre and post transition. We also perform the transition at a binary separation of $D \sim 20M$, where the binary’s dynamics are still well described by PN theory and errors introduce in the gas dynamics by the approximate metric are washed out by MHD turbulence (see Ref. [9] for an analysis of MHD evolutions on this analytical background for various separations).

In order for the transition from an analytical evolution to a numerical one to be smooth enough, the binary’s orbital dynamics could not change significantly as a result of the transition. The binary’s dynamics in the fully nonlinear numerical simulation had two main sources of error. First, constraint violations led to rapid unphysical oscillations in the orbital separation. Second, small errors in the PN expressions for the orbital angular momentum and inspiral rate led to eccentricity in the binary. We were able to ameliorate the first source of error by evolving with the constraint-damping CCZ4 [40] formulation of Einstein’s equations, which causes constraint violations to rapidly propagate away from the BHs, sig-
nificantly reducing unphysical binary dynamics. In addition, by adding small changes to the initial inspiral rate and orbital frequency, we significantly reduced the eccentricity of the numerical binary using the techniques of [13].

We subsequently found that the NR evolution leads to both the expected gravitational waveform, orbital frequency, and binary inspiral rate (to within the truncation error of the simulation). The remaining error we found is a phase error in the early part of the waveform. This phase error is about 0.255 rad. We ascribe this error to our use of the flat-space retarded time in the far zone. By not including effects due to the mass of the spacetime far zone metric, which is something we will explore in an upcoming paper.

ACKNOWLEDGMENTS

We thank Carlos Lousto for a careful reading of this manuscript. We thank Carlos Lousto, Zachariah Etie
