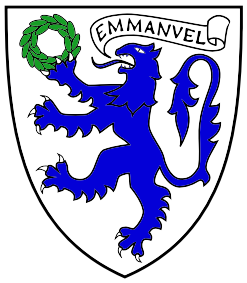
Numerical modelling of gravitational wave sources in general relativity



Miren Raj Radia

Supervisor: Prof. Ulrich Sperhake

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Emmanuel College

January 2023

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This thesis contains work that appears in the following publications:

- T. Andrade, L. A. Salo, J. C. Aurrekoetxea, J. Bamber, K. Clough, R. Croft, E. de Jong, A. Drew, A. Duran, P. G. Ferreira, P. Figueras, H. Finkel, T. França, B.-X. Ge, C. Gu, T. Helfer, J. Jäykkä, C. Joana, M. Kunesch, K. Kornet, E. A. Lim, F. Muia, Z. Nazari, **M. Radia**, J. Ripley, P. Shellard, U. Sperhake, D. Traykova, S. Tunyasuvunakool, Z. Wang, J. Y. Widdicombe, and K. Wong, *GRChombo: An adaptable numerical relativity code for fundamental physics*, Journal of Open Source Software **6**, 3703 (2021)

- **M. Radia**, U. Sperhake, E. Berti, and R. Croft, *Anomalies in the gravitational recoil of eccentric black-hole mergers with unequal mass ratios*, Phys. Rev. D **103**, © 2021 American Physical Society, 104006 (2021), arXiv:2101.11015

- **M. Radia**, U. Sperhake, A. Drew, K. Clough, P. Figueras, E. A. Lim, J. L. Ripley, J. C. Aurrekoetxea, T. França, and T. Helfer, *Lessons for adaptive mesh refinement in numerical relativity*, Class. Quant. Grav. **39**, 135006 (2022), arXiv:2112.10567 [gr-qc]

- T. Helfer, U. Sperhake, R. Croft, **M. Radia**, B.-X. Ge, and E. A. Lim, *Malaise and remedy of binary boson-star initial data*, Class. Quant. Grav. **39**, 074001 (2022), arXiv:2108.11995 [gr-qc]

Miren Raj Radia
January 2023

Numerical modelling of gravitational wave sources in general relativity

Miren Raj Radia

The first direct detection of gravitational waves (GWs) from a black-hole (BH) binary, GW150914, by the advanced Laser Interferometer Gravitational-wave Observatory (aLIGO) detectors in 2015 heralded a new era in GW physics. Since then, over 90 compact binary merger events have been detected by the GW detector network with many more expected in the decades to come. A significant part of the theoretical foundations that underpins this achievement is the modelling of GW sources in General Relativity (GR) using numerical relativity (NR). In this thesis, we discuss the features and capabilities of the NR code GRChombo. Although GRChombo is no longer a new code, its original development and design targeted applications beyond the conventional astrophysical paradigm that other NR codes have focussed on. Here, we describe more recent additions that have allowed GRChombo to model BH binaries and other GW sources with good accuracy. Through direct comparison, we demonstrate that this accuracy is comparable to that of a more mature NR code. One of the key capabilities of GRChombo is its adaptive mesh refinement (AMR). This allows the numerical grid to dynamically adjust itself in order to sufficiently resolve the large range of spatial and temporal scales that characteristically arise in non-trivial solutions of GR as a consequence of the theory's non-linearity. However, this flexibility requires careful control in order to achieve the desired accuracy and we discuss in detail the lessons learned in order to achieve this with GRChombo. We apply GRChombo and these techniques to the investigation of the effect of orbital eccentricity on the GW emission and the gravitational recoil imparted to the BH merger remnant from the inspiral and merger of unequal-mass non-spinning BH binaries. Finally, we explore the modelling of a more exotic type of compact object: boson stars (BSs) which are comprised of complex scalar field matter. In particular, we investigate the construction of suitable initial data describing BS binaries and its effect on the ensuing evolutions.

This thesis is dedicated to my loving parents, Bina and Raj....

Acknowledgements

This thesis would not have been possible without the support of my family, friends, colleagues and collaborators. In this section, I wish to name some of those who I am particularly indebted to but this list is far from exhaustive.

Firstly, I would like to thank my supervisor, Ulrich Sperhake, for his invaluable guidance and support, and for directing me along interesting and fruitful avenues for research. Not only have I benefited greatly from our conversations about physics and his wealth of experience in numerical relativity, but I have also enjoyed our discussions on completely unrelated matters.

I have been fortunate to be a part of the Numerical Relativity group in DAMTP. As a result, I have had many useful and insightful exchanges about physics and numerics with Roxana Roşca-Mead, Robin Croft, Amelia Drew, Michalis Agathos and Justin Ripley amongst others. I am also thankful to Emanuele Berti for his contributions to the work presented in Chapter 5 and Andrea Nerozzi for pointing me in the right direction in order to complete the calculation in Sec. 2.4.3.

A particularly enjoyable aspect of my PhD has been being a part of the GRChombo collaboration which has grown significantly over the last few years. I have had many productive and engaging discussions with a significant number of GRChombo collaborators and wish to thank everyone who is and has been a part of it. I am particularly grateful to Katy Clough, Tiago França, Josu Aurrekoetxea, Thomas Helfer, Pau Figueras, Eugene Lim and Kacper Kornet.

I would like to express my gratitude to my friends in Pavilion B who have made my time as a graduate student a more enjoyable experience: Owain, Fred, Philip, Roland, Wuhyun, Roxana, João, Maeve, Charles, Toby, Jack, Amanda, James, Johnathan and Tobias, in particular, my officemates Fred and Owain for making our office an entertaining albeit probably not the most productive environment. I am also especially indebted to Philip and João for their efforts in encouraging me to finish this thesis over the last few years.

In addition to the above, I also wish to thank my other friends, both in Cambridge and elsewhere who have provided me with much-needed relief from my research: Maria, Matt, Jade, Louisa, Barnaby, Ritu, Rishi, Luke, Alex, Claire and Jamie amongst others.

Finally, I am eternally grateful for the ongoing support and love from my family and in particular, my parents, Bina and Raj, without whom I would not have been able to pursue this path.

The simulations presented in this thesis used the following HPC resources:

- The Cambridge Service for Data Driven Discovery (CSD3), part of which is operated by the University of Cambridge Research Computing Service on behalf of the Science and Technology Facilities Council (STFC) DiRAC HPC Facility. The DiRAC component of CSD3 was funded by BEIS capital funding via STFC capital grants ST/P002307/1 and ST/R002452/1 and STFC operations grant ST/R00689X/1.

- The DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility. The equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1 and ST/R002371/1, Durham University and STFC operations grant ST/R000832/1.

- XSEDE resources under Grant No. PHY-090003 were used on the San Diego Supercomputing Center's clusters Comet and Expanse and the Texas Advanced Supercomputing Center's (TACC) Stampede2.

- The TACC Frontera cluster. Frontera is made possible by National Science Foundation award OAC-1818253.

- PRACE resources under Grant Number 2020225359 were used on the GCS Supercomputer JUWELS at the Jülich Supercomputing Centre (JCS) through the John von Neumann Institute for Computing (NIC), funded by the Gauss Centre for Supercomputing e.V.

- The Fawcett supercomputer in DAMTP funded by STFC Consolidated Grant ST/P000673/1.

DiRAC Resources were used under projects ACSP218, ACTP186 and ACTP238. DiRAC is part of the National e-Infrastructure.

Throughout my PhD, I have been supported by an STFC studentship and have also received invaluable financial support for travel from Emmanuel College, Cambridge. Furthermore, I received support to visit the Yukawa Institute for Theoretical Physics,

Kyoto and Kindai University, Osaka through the European Union Horizon 2020 project H2020-MSCA-RISE-2015 “StronGRHEP” 690904.

Table of contents

List of figures

List of tables

Nomenclature

Roman Symbols

$\boldsymbol{C}$	Weyl tensor
D	Spatial covariant derivative (Levi-Civita connection)
${}^{(4)}\boldsymbol{R}$	Spacetime Riemann tensor
$\boldsymbol{g}$	Spacetime metric
$\boldsymbol{K}$	Trace of the extrinsic curvature
$\mathscr{M}$	Spacetime manifold
$\boldsymbol{n}$	Unit normal to the foliation of spacelike hypersurfaces
$\boldsymbol{R}$	Spatial Riemann tensor
t	Coordinate time
$\boldsymbol{t}$	Timeflow vector equal to ∂_t

Greek Symbols

α	Lapse function
$\boldsymbol{\beta}$	Shift vector
$\boldsymbol{\gamma}$	Spatial metric
Ψ_4	The fourth Weyl/Newman-Penrose scalar
$\psi_{\ell m}$	The (ℓ, m) mode of Ψ_4 with respect to the spin-weight $s = -2$ spherical harmonics ${}_{-2}Y^{\ell m}$

Σ A spatial hypersurface in the foliation

Other Symbols

$\breve{\nabla}$ Flat space derivative

∇ Spacetime covariant derivative (Levi-Civita connection)

$\perp$ Spatial projection operator

Acronyms / Abbreviations

ADM Arnowitt-Deser-Misner

AMR Adaptive Mesh Refinement

BC Boundary Condition

BH Black Hole

BR Berger-Rigoutsos

BS Boson Star

BSSN(OK) Baumgarte-Shapiro-Shibata-Nakamura(-Oohara-Kojima)

BVP Boundary Value Problem

BY Bowen-York

CCZ4 Conformal and Covariant Z4

CFL Courant-Friedrichs-Lewy

CPU Central Processing Unit

DFT Discrete Fourier Transform

ECO Exotic Compact Object

EKG Einstein-Klein-Gordon

EM Energy-Momentum

FFI Fixed Frequency Integration

FFT Fast Fourier Transform

FTCS Forward Time Centred Space

GPU Graphics Processing Units

GR General Relativity

GW Gravitational Wave

IBVP Initial Boundary Value Problem

IVP Initial Value Problem

KAGRA Kamioka Gravitational Wave Detector

KO Kreiss-Oliger

LIGO Laser Interferometer Gravitational-wave Observatory

LISA Laser Interferometer Space Antenna

MOL Method Of Lines

MPI Message Passing Interface

NR Numerical Relativity

NS Neutron Star

ODE Ordinary Differential Equation

PDE Partial Differential Equation

PN Post-Newtonian

QNM Quasinormal Modes

RHS Right-Hand Side

RK4 Fourth order Runge-Kutta

SIMD Single Instruction, Multiple Data

TT Transverse-Traceless

Chapter 1

Introduction

Over a century ago, Einstein revolutionised modern physics by publishing his seminal papers introducing the general theory of relativity [5, 6, 7] or *general relativity* (GR). The foundations of GR rely on several fundamental conceptual leaps from previous theories of gravity. For example, the unification of space and time into spacetime defies the common intuition on their separation. Perhaps the most groundbreaking and beautiful aspect of the theory is the manifestation of gravity not as a force, but as a consequence of the geometry of spacetime. The relationship between the curvature of spacetime and the matter it contains is described by the Einstein equation[1]–arguably one of the most elegant equation of physics. In the notation we shall use in this thesis, and in units where the gravitational constant and the speed of light are both unity, $G = c = 1$, it is

$$ {}^{(4)}R_{\mu\nu} - \frac{1}{2}{}^{(4)}R g_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}, \tag{1.0.1} $$

where Λ is the cosmological constant. As put succinctly by Wheeler [8], this equation says

> *Spacetime tells matter how to move; matter tells spacetime how to curve.*

One of the most striking consequences of GR was first found by Schwarzschild [9] when he published an exact solution of the Einstein equation in 1916. This was later understood to describe a spacetime containing a spherically symmetric *black hole* (BH)–a region of spacetime from which there is no escape [10, 11][2]. This solution was later

[1] Although the Einstein equation can be regarded as a single tensorial equation, it is often thought of as multiple equations for the individual components of the tensors comprising each side, hence we shall use the singular term "Einstein equation" and plural term "Einstein [field] equation*s*" interchangeably.

[2] These references establish the causal structure of the Schwarzschild spacetime but the term black hole was coined slightly later by Wheeler and Dicke [12].

generalized to include rotation by Kerr [13] and electric charge by Newman, Couch, Chinnapared, Exton, Prakash, and Torrence [14]. Furthermore, following the uniqueness theorems of Carter [15] and Robinson [16], the Kerr BH is expected to be the stationary endpoint of gravitational collapse in GR.

Another early corollary of GR that Einstein predicted was the existence of gravitational waves (GWs)–ripples in the fabric of spacetime [17, 18]. He found that in the linearised weak-field limit, the Einstein equation (1.0.1) admits transverse plane-wave solutions that travel at the speed of light. Furthermore, Einstein showed that such waves are generated by the time-variation of the mass quadrupole moment of matter sources. Despite these insights, there was still much debate about the existence and physicality of GWs which was not settled until the late 1950s [19]. Even Einstein changed his mind several times on their existence [20].

The first indirect evidence of the existence of GWs came when Hulse and Taylor [21], in 1974, discovered a binary pulsar system–a two-body system where at least one of the constituent objects is a highly magnetised neutron star (NS) that emits electromagnetic radiation from its poles. The energy loss due to gravitational radiation inferred from observations of the electromagnetic pulses over several years confirmed Einstein's quadrupole formula to remarkable accuracy [22]. This result led to Hulse and Taylor being awarded the Nobel Prize in 1993. However, the astrophysical community still yearned to directly observe GWs in order to provide a window to testing GR in the strong-field regime.

Efforts to directly detect GWs began with resonant mass antennas in the 1960s [23]. Later iterations of these detectors developed in the 1980s, 1990s and early 2000s were cryogenically cooled and operated as part of an international network [24]. Despite these improvements, to the best of our knowledge, no GWs have ever been detected by them. Interferometric detectors were first proposed in the 1960s [25] and 1970s [26] and by the early 2000s several such experiments were operational including TAMA 300 in Japan, GEO600 in Germany, the Laser Interferometer Gravitational-wave Observatory (LIGO) in the United States and Virgo in Italy. Various combinations of these detectors made observation runs between 2002 and 2011. Whilst their data was used to place bounds on GW sources (see, for example, Ref. [27]), as it turned out, the elusive goal of direct detection required greater sensitivity than achieved by these first generation detectors.

In September 2015, almost 100 years after Einstein published his papers on GR, the two advanced[3] LIGO (aLIGO) detectors directly detected a GW signal for the first time. The radiation was emitted from the inspiral, merger and ringdown of a BH binary system approximately 1.3×10^{22} km away [28]. This GW150914 event marked the start of a new and exciting period in GW physics and, like the previous indirect evidence of GWs, also led to three significant LIGO contributors (Weiss, Barish and Thorne) being awarded the Nobel Prize in 2017. Since GW150914, other facilities have joined the worldwide advanced GW detector network, namely the upgaded Virgo interferometer and the Kamioka Gravitational Wave Detector (KAGRA) located in Japan. Together, the network has detected over 90 compact binary merger events [29, 30, 31]. One notable event is that of GW170817, the first observation of GWs emitted from the inspiral of a NS binary system [32], which was corroborated by observations of counterpart electromagnetic radiation. With further observing runs using the existing network coming up and next generation detectors including the Einstein Telescope and the space-based Laser Interferometer Space Antenna (LISA) mission on the horizon, many more events are expected in the years to come including from previously unobserved types of sources.

The key theoretical task that underpins these detections is the computation of predicted waveforms through the modelling of GW sources. Many of the observed events (including GW150914) can be found by searching for transient bursts of power in the detectors without using a waveform model. However, inference of the sources and exploration of the underlying physics relies on precision modelling. GW source modelling is a collaborative effort with several different tools that are best suited to different parts of the waveform. Let us consider the case of a typical BH binary observed by the GW detector network, that is, with a mass ratio that is not extreme [33]. Here, the BHs could complete many thousands of orbits before merging (depending on how the binary originally formed). For the vast majority of such an inspiral, the post-Newtonian (PN) approximation is the best tool to quickly and accurately determine the BH dynamics over these relatively long timescales. Shortly after merger, the remnant BH rapidly settles down to the Kerr solution in the ringdown phase. For this part of the waveform, BH perturbation theory can accurately model the quasinormal modes (QNMs) of the remnant. In the intermediate part of the waveform including the late inspiral, plunge and merger, the aforementioned approximations break down and it is necessary to solve the

[3]The "advanced" specifier is used to denote the upgrade from first to second generation detectors that substantially improved their sensitivity.

full Einstein equations using *numerical relativity* (NR). Our BH binary example should be considered as a proxy for other phenomena of interest in GR; NR is often the only tool that can accurately model in the highly non-linear strong-field regime. We shall focus on modelling using NR in this thesis.

NR as a field can be traced to the pioneering work of Hahn and Lindquist [34] in the mid-1960s but these early simulations only evolved for very short times. More successful simulations involving head-on BH collisions were undertaken by Čadež [35] and Smarr [36] and Eppley [37] in the 1970s. However, these efforts were limited by the modest computational resources available at the time which restricted numerical evolutions to either spherical symmetry or low-resolution axisymmetry. In the late 1980s and 1990s, there was significant progress in the field with many explorations of different problems in GR using NR. However, it took until 2005 to obtain the holy grail – the simulation of orbiting BH binaries through inspiral, merger and ringdown. This was first achieved by Pretorius [38], and soon after by Campanelli, Lousto, Marronetti, and Zlochower [39] and Baker, Centrella, Choi, Koppitz, and Meter [40]. This important accomplishment is often referred to as *the numerical relativity breakthrough*[4]. Since then, NR has continued to flourish as a field. In Chapter 2, we provide an overview of the NR techniques and more general numerical methods that will be relevant to later chapters.

In order to perform simulations using computers, we approximate the continuum spacetime by a discrete computational grid with finite resolution (spacing between the cells/points). It is a generic feature of GR and its character as a highly non-linear theory that its solutions often span a large range of spatial and temporal scales. Combined with the inherent limits of computational resources, it follows that any finite resolution numerical code will require some form of spatial and temporal mesh refinement to fully capture the dynamics of these solutions. Here, by mesh refinement, we mean that in parts of the computational domain (but not the whole domain), the grid resolution is progressively increased by covering these regions with finer meshes. The NR code that we predominantly use in this thesis, GRChombo, uses the technique of *adaptive mesh refinement* (AMR) where the regions covered by the finer meshes are dynamically adjusted during the evolution in order to sufficiently resolve the physics being simulated. We discuss the current state of GRChombo in Chapter 3 and, in particular, recent improvements we have made in order to enhance its capability to model GW sources accurately. Furthermore, we discuss the technique of AMR and the insights we have gained in leveraging its flexibility in the context of NR in Chapter 4. We apply GRChombo and

[4]For a more detailed history of the steps that lead to the breakthrough, see Sec. 5 in Ref. [41].

these lessons to an investigation in the area of BH binary phenomenology in Chapter 5, namely the effect of orbital eccentricity on the gravitational recoil imparted to the remnant BH in the merger of unequal-mass BH binaries. Further motivation on this problem can be found in the introduction to this chapter.

In the conventional astrophysical paradigm, the only compact objects in GR are BHs and NSs. Furthermore, the Tolman-Oppenheimmer-Volkoff limit for NSs (see, for example, Ref. [42]) means that BHs are the only ones with masses greater than $3M_\odot$. Until now, analysis of GW data has largely been consistent with this picture [43, 44, 45, 46, 47]. Nevertheless, with observational data and accuracy only increasing, there are many reasons to consider models beyond this paradigm. For instance, modified gravity theories and/or extensions to the Standard Model of particle physics can give rise to *exotic compact objects* (ECOs) (see Ref. [48] for a recent review on the zoo of ECOs). One of the most well-motivated type of ECOs are *boson stars* (BSs) which we discuss and investigate in Chapter 6.

Notational conventions

Unless otherwise stated, we take the following conventions in this thesis. We use Greek letters $\mu, \nu, \ldots = 0, 1, 2, 3$ for spacetime indices and Latin letters $i, j, \ldots, 1, 2, 3$ for spatial indices. We sometimes denote tensors in index-free notation using bold symbols such as $\boldsymbol{T}$. We use a mostly plus signature $(-+++)$ and geometric units where the the speed of light c and the gravitational constant G are both unity $c = 1 = G$. When there is a potential for ambiguity between spacetime and purely spatial tensors (for example, the Ricci scalar R), we prepend a ${}^{(4)}$ to denote the spacetime quantity. Further conventions can be inferred from Chapter 2 or are explicitly provided in the introduction of the relevant chapter.

Chapter 2

Numerical Relativity Fundamentals

In this chapter, we will provide an overview of the mathematics and techniques that we use for numerical evolutions of the Einstein equation. We start by introducing the 3+1 decomposition in Sec. 2.1. Next in Sec. 2.2, we look at how to construct initial data that can be evolved numerically. Unfortunately, the ADM evolution equations that arise in the 3+1 formalism are unsuitable for most numerical evolutions, so in Sec. 2.3, we look at reformulations that allow for stable and accurate numerical evolutions. In Sec. 2.4 we describe how to measure the gravitational waves in numerical solutions and how this relates to the plane-wave solutions of the linearised Einstein equation. Finally, in Sec. 2.5, we review some of the basic numerical methods that are used for the simulations presented in this thesis.

Though we do not aim to provide a comprehensive overview of these topics, as this should be reserved for a good textbook rather than this chapter, we do hope to provide sufficient detail such that someone with a working knowledge of general relativity can follow the work described in later chapters. The reader is referred to Refs. [49, 50, 51] for a more thorough exploration of these areas and also to Ref. [52] for further information on the numerical aspects.

Whilst this chapter does not present any novel research, the calculation of the electromagnetic decomposition of the Weyl tensor in Sec. 2.4.3, despite being used ubiquitously in NR, is, to the best of our knowledge, difficult to find explicitly in the literature.

2.1 The 3+1 formalism

Before we can solve the Einstein equation numerically, we first need to formulate it as an initial value problem (IVP) that will allow us to specify some initial data and then evolve them in "time". It should be pointed out that this is only possible due to the *hyperbolic* nature of the Einstein equation[1] and it wasn't until over 30 years after Einstein published his general theory of relativity that it was proved that his equation admits a locally well-posed IVP [53].

When learning relativity for the first time, one usually tries [or is forced to] "think covariantly" and treat space and time on an equal footing. Indeed, the general covariance of GR may be viewed as part of its elegance. However, for our practical applications, we will need to devolve to the arguably more intuitive picture of splitting space and time.

Before we proceed, we should point out that the 3+1 formalism is far from the only way to formulate Einstein equation as an IVP. A common alternative is to use a generalized form of harmonic coordinates where the coordinates x^μ satisfy[2]

$$\Box_g x^\mu = H^\mu(x^\nu), \tag{2.1.1}$$

and $\Box_g = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the d'Alembertian. Not only were harmonic coordinates the ones used to prove local well-posedness of the Einstein equation by Fourès-Bruhat [53], but they were also used in the groundbreaking first simulation of an inspiralling black-hole binary by Pretorius [38]. Furthermore, they are used by the SXS collaboration to generate waveforms for GW detector template banks [54]. Other, less common formulations include the Cauchy characteristic approach where the slices of spacetime are null rather than spatial (see Ref. [55] for a review) and the hyperboloidal slices approach where the slices are asymptotically null.

Nevertheless, our discussion will only consider the conventional 3+1 decomposition.

2.1.1 The foliation of spacetime

We start with a spacetime $(\mathscr{M}, \boldsymbol{g})$ which we shall assume is *globally hyperbolic* which means that there exists a Cauchy surface[3] $\Sigma \subset \mathscr{M}$ whose domain of dependence is the whole spacetime $\mathscr{M}$. It follows that there exists a *global* time function $t : \mathscr{M} \to \mathbb{R}$ [56] such that the level sets of t, which we shall denote Σ_t, are Cauchy surfaces that are

[1] More precisely, the Einstein equation is mixed hyperbolic-elliptic, as we will see later in Sec. 2.1.6.
[2] Vanilla harmonic coordinates are obtained if one chooses H^μ to vanish.
[3] A Cauchy surface is one where no two points are causally connected.

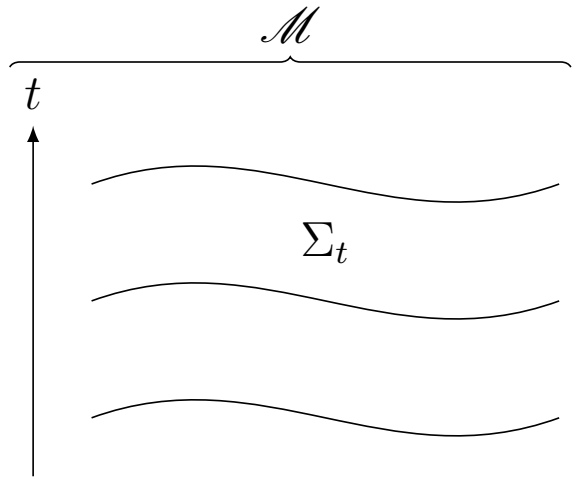


Fig. 2.1 The foliation of spacetime in the 3+1 decomposition

topologically equivalent (hence we shall usually drop the t subscript) and the spacetime can be completely *foliated* by these surfaces. This means we can identify $\mathscr{M}$ with $\mathbb{R} \times \Sigma$ as shown in Fig. 2.1. Furthermore, there exists a "time-flow"[4] vector field $\boldsymbol{t}$ such that

$$t^\alpha \nabla_\alpha t = 1. \tag{2.1.2}$$

The future-directed unit normal to the foliation $\boldsymbol{n}$ is given by

$$\boldsymbol{n} = -\alpha \, \mathrm{d}t, \tag{2.1.3}$$

where α is the *lapse* given by

$$\alpha := ||\mathrm{d}t||^{-1/2} = (-g^{\beta\gamma} \nabla_\beta t \nabla_\gamma t)^{-1/2} > 0, \tag{2.1.4}$$

and the final inequality follows from the foliation being spacelike. The *induced* or *spatial metric* on Σ is

$$\gamma_{\mu\nu} := g_{\mu\nu} + n_\mu n_\nu. \tag{2.1.5}$$

We can project spacetime tensors onto the hypersurfaces Σ using the *projector*

$$\perp^\mu_\nu := \delta^\mu_\nu + n^\mu n_\nu = \gamma^\mu{}_\nu, \tag{2.1.6}$$

[4]Note that t^μ is not necessarily timelike

where the projection $\perp T$ of an arbitrary rank (r,s) tensor $\boldsymbol{T}$ is given by

$$(\perp T)^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} := \perp^{\mu_1}_{\alpha_1}\cdots\perp^{\mu_r}_{\alpha_r}\perp^{\beta_1}_{\nu_1}\cdots\perp^{\beta_s}_{\nu_s} T^{\alpha_1\cdots\alpha_r}{}_{\beta_1\cdots\beta_s}. \tag{2.1.7}$$

From this, we can decompose tensors into their normal and spatial directions using the relation $\delta^\mu_\nu = \perp^\mu_\nu - n^\mu n_\nu$, for example, the time-flow vector field can be decomposed as

$$t^\mu = \delta^\mu_\alpha t^\alpha = (\perp^\mu_\alpha - n^\mu n_\alpha)t^\alpha = \underbrace{-n^\mu n_\alpha t^\alpha}_{\text{normal}} + \underbrace{\perp^\mu_\alpha t^\alpha}_{\text{spatial}}. \tag{2.1.8}$$

We define the *shift vector* $\boldsymbol{\beta}$ as the spatial projection of $\boldsymbol{t}$

$$\beta^\mu := \perp^\mu_\alpha t^\alpha, \tag{2.1.9}$$

and it follows from Eqs. (2.1.2) and (2.1.8) that the time flow vector field can be written as

$$\boldsymbol{t} = \alpha\boldsymbol{n} + \boldsymbol{\beta}. \tag{2.1.10}$$

We now introduce the Levi-Civita connection of $\boldsymbol{\gamma}$ which we denote as D. Remarkably, it turns out that the covariant derivative of spatial tensors defined by this connection is simply the projection of the corresponding spacetime covariant derivative defined by the Levi-Civita connection of $\boldsymbol{g}$, that is, for an arbitrary rank (r,s) spatial tensor $\boldsymbol{S}$,

$$D_\mu S^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_s} = \perp^\alpha_\mu \perp^{\nu_1}_{\beta_1}\cdots\perp^{\nu_r}_{\beta_r}\perp^{\gamma_1}_{\rho_1}\cdots\perp^{\gamma_s}_{\rho_s}\nabla_\alpha S^{\beta_1\cdots\beta_r}{}_{\gamma_1\cdots\gamma_s}. \tag{2.1.11}$$

It should be pointed out that the projections on all of the indices on the right-hand-side are necessary even though $\boldsymbol{S}$ is purely spatial, as, in general,

$$\perp^\alpha_\mu\perp^{\nu_1}_{\beta_1}\cdots\perp^{\nu_r}_{\beta_r}\perp^{\gamma_1}_{\rho_1}\cdots\perp^{\gamma_s}_{\rho_s}\nabla_\alpha S^{\beta_1\cdots\beta_r}{}_{\gamma_1\cdots\gamma_s} \neq \perp^\alpha_\mu \nabla_\alpha S^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_s}. \tag{2.1.12}$$

A common pitfall, that those new to the 3+1 formalism often fall into[5], is omitting these projectors as in the right-hand-side of Eq. (2.1.12).

Until now, we have considered all tensors as being defined on [the tangent spaces of] the whole spacetime $\mathscr{M}$. However, for purely spatial tensors, that is those that are invariant under projection (2.1.7) (or equivalently are orthogonal to $\boldsymbol{n}$), we can identify these with tensors defined on Σ. Similarly for tensors defined on Σ, we can identify these with the unique tensor on $\mathscr{M}$ that is equal on Σ and is invariant under projection.

[5]The author is included here.

2.1.2 Adapted coordinates

Now we define coordinates that are *adapted* to the foliation. We start by choosing Cartesian coordinates x^i on one of the hypersurfaces Σ and parallel transport these coordinates to the rest of the foliation using $\boldsymbol{t}$. Then, $(x^\mu) = (t, x^i)$ are our adapted coordinates.

In these coordinates, the spacetime metric takes the form

$$\boldsymbol{g} = -\alpha^2 \,\mathrm{d}t^2 + \gamma_{ij}\left(\mathrm{d}x^i + \beta^i \,\mathrm{d}t\right)\left(\mathrm{d}x^j + \beta^j \,\mathrm{d}t\right), \tag{2.1.13}$$

or, in components,

$$(g_{\mu\nu}) = \begin{pmatrix} -\alpha^2 + \beta^k\beta_k & \beta_i \\ \beta_j & \gamma_{ij} \end{pmatrix}, \qquad (g^{\mu\nu}) = \begin{pmatrix} -1/\alpha^2 & \beta^i/\alpha^2 \\ \beta^j/\alpha^2 & \gamma^{ij} - \beta^i\beta^j/\alpha^2 \end{pmatrix}, \tag{2.1.14}$$

and the unit vector normal to the foliation (i.e. Eq. (2.1.3) with raised indices) takes the form

$$(n^\mu) = \frac{1}{\alpha}\left(\partial_t - \beta^i\partial_i\right). \tag{2.1.15}$$

Similarly, using Eq. (2.1.10), the time-flow vector field in these coordinates is $\boldsymbol{t} = \partial_t$.

Note that, in these adapted coordinates, purely spatial tensors (for example, those that have been projected with $\perp$ (2.1.7)) have trivially vanishing raised temporal components since they are orthogonal to $\boldsymbol{n}$ and it follows from Eq. (2.1.3) that

$$0 = n_{\mu_p}\,(\perp T)^{\mu_1\cdots\mu_p\cdots\mu_r}{}_{\nu_1\cdots\nu_s} = -\alpha\,(\perp T)^{\mu_1\cdots 0\cdots\mu_r}{}_{\nu_1\cdots\nu_s}, \qquad \alpha > 0. \tag{2.1.16}$$

However, the same is not true for lowered indices since in adapted coordinates, the components n^i do not necessarily vanish by Eq. (2.1.15). Of course, these temporal components are uniquely determined by

$$0 = n^{\nu_p}\,(\perp T)^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_p\cdots\nu_s}, \tag{2.1.17}$$

hence it is sufficient to consider only the spatial components (whether raised or lowered) of purely spatial tensors. We will therefore freely swap between spacetime indices $\mu, \nu, \rho, \sigma, \ldots$ and spatial indices $i, j, k, \ldots$ for purely spatial tensors.

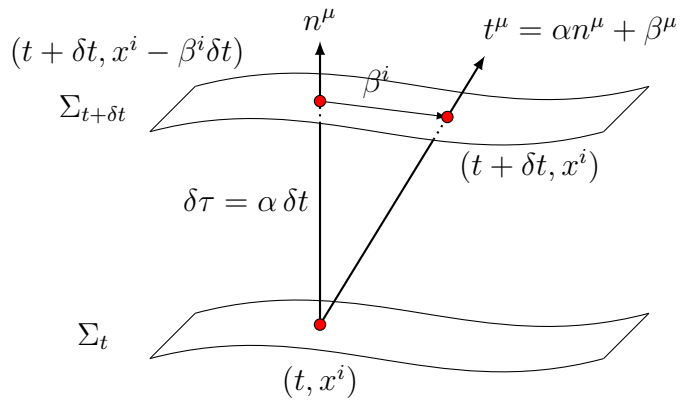


Fig. 2.2 Schematic diagram of how the coordinates on the hypersurface Σ_t are related to the coordinates on the hypersurface $\Sigma_{t+\delta t}$.

2.1.3 Gauge freedom

Unlike in the conventional covariant formulation of GR, the gauge freedom (or diffeomorphism invariance) has been made explicit through the introduction of the lapse α and shift vector $\boldsymbol{\beta}$. Before we proceed any further we should try and understand intuitively what these objects represent.

We call an observer travelling with 4-velocity $\boldsymbol{n}$ (2.1.15) an *Eulerian* or *normal* observer. The "lapse" of proper time measured by such an observer $\delta\tau$ is related to the change in coordinate time by

$$\delta\tau = \alpha\, \delta t. \tag{2.1.18}$$

From this equation, it is not too hard to see that the lapse controls the way that the spacetime is foliated. The shift tells us the velocity of the lines of constant coordinate x^i relative to the Eulerian observer. If the observer is at x^i at coordinate time t, then at coordinate time $t + \delta t$, the observer will be at $x^i - \beta^i \delta t$. We can see that the shift controls the way the spatial coordinates change from one hypersurface to the next. This picture is illustrated schematically in Fig. 2.2.

Since the lapse and shift represent our gauge freedom they are, in principle, freely specifiable. However, in practice, the gauge plays a large part in ensuring the numerical simulation remains stable and accurate. Choosing a bad gauge may lead to an instability or finite-time blowup. For example, if we take the naive choice of $\alpha = 1$ (which is known as *geodesic slicing*) with the Schwarzschild spacetime in isotropic coordinates, it can be shown that any observer which starts within the horizon, reaches the singularity in a proper time $\Delta\tau \leq \pi M$ [57]. Since $\alpha = 1$ means that $\Delta t = \Delta\tau$, if we tried to evolve

this spacetime using geodesic slicing, it is clear that we will run into problems in finite coordinate time. The particular gauge that we use throughout the simulations presented in this work will be described in Sec. 2.3.4.

Finally, it will be useful to consider the *acceleration* of an Eulerian observer which is given by

$$a_\mu := n^\beta \nabla_\beta n_\mu = D_\mu \log \alpha, \tag{2.1.19}$$

where the second equality follows after some algebra.

2.1.4 The extrinsic curvature

Since the Einstein equation is second order in time, we need to introduce an auxiliary variable in order to rewrite the equations in first-order [in time] form. The canonical choice for such an object is known as the *extrinsic curvature* $\boldsymbol{K}$ which is defined by[6]

$$K_{\mu\nu} = - \perp^\alpha_\mu \nabla_\alpha n_\nu. \tag{2.1.20}$$

Note that, despite there being just one projector for two indices, the extrinsic curvature is purely spatial since, for the second index,

$$n^\alpha K_{\mu\alpha} = - \perp^\beta_\mu n^\alpha \nabla_\beta n_\alpha = -\frac{1}{2} \perp^\beta_\mu \nabla_\beta (n^\alpha n_\alpha) = 0, \tag{2.1.21}$$

where the final equality follows because $\boldsymbol{n}$ is a unit vector. Combining this with the acceleration (2.1.19) we can write the gradients of the normal vector as

$$\nabla_\mu n_\nu = -K_{\mu\nu} - n_\mu D_\nu \log \alpha. \tag{2.1.22}$$

It can also be shown that the extrinsic curvature is given by

$$K_{\mu\nu} = -\frac{1}{2} \mathcal{L}_{\boldsymbol{n}} \gamma_{\mu\nu}, \tag{2.1.23}$$

where $\mathcal{L}_{\boldsymbol{n}}$ is the Lie derivative with respect the normal $\boldsymbol{n}$. Indeed, this is sometimes taken as the definition of the extrinsic curvature in some places in the literature. It immediately follows that the extrinsic curvature is symmetric[7].

[6] Whilst this sign convention is the usual one in NR, it differs to that used elsewhere in the literature, for example, Ref. [56].

[7] This can be shown directly from its definition (2.1.20), see, for example Ref. [51]

From Eq. (2.1.23), we can see that, heuristically, $\boldsymbol{K}$ is just the time derivative of the spatial metric and this is typically the main source of intuition one uses in the numerical relativity context.

However, there is also a geometrical picture as to what the extrinsic curvature represents. Looking at the definition (2.1.20), we can see that $\boldsymbol{K}$ encodes information about how the normal $\boldsymbol{n}$ varies over Σ, or in other words, the curvature of the embedding of Σ in $\mathscr{M}$. This geometrical interpretation explains the name "extrinsic curvature". It is important to distinguish between the *extrinsic* curvature which tells us about the embedding of Σ in $\mathscr{M}$ and the *intrinsic* curvature encoded in the spatial Riemann tensor (with respect to the Levi-Civita connection of the spatial metric $\boldsymbol{\gamma}$) which does not know about the embedding.

2.1.5 Projections of the Riemann tensor

In order to obtain evolution equations in the 3+1 formalism, we need to project the spacetime Riemann tensor[8] ${}^{(4)}\boldsymbol{R}$. Before we do so, we set our sign conventions by defining the spacetime Riemann tensor applied to vector fields $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ by

$$ {}^{(4)}\boldsymbol{R}(\boldsymbol{u},\boldsymbol{v})\boldsymbol{w} := \nabla_{\boldsymbol{u}}\nabla_{\boldsymbol{v}}\boldsymbol{w} - \nabla_{\boldsymbol{v}}\nabla_{\boldsymbol{u}}\boldsymbol{w} - \nabla_{[\boldsymbol{u},\boldsymbol{v}]}\boldsymbol{w}, \tag{2.1.24} $$

where, for a vector field $\boldsymbol{u}$, $\nabla_{\boldsymbol{u}} = u^\alpha \nabla_\alpha$ and the commutator $[\boldsymbol{u},\boldsymbol{v}]$ is given by

$$ [\boldsymbol{u},\boldsymbol{v}]^\mu = u^\alpha \nabla_\alpha v^\mu - v^\alpha \nabla_\alpha u^\mu. \tag{2.1.25} $$

The spatial Riemann tensor $\boldsymbol{R}$ is defined similarly with respect to D.

Naively, one might think there are many different projections of the Riemann tensor given that, for each index, one has a choice whether to project in the spatial or normal direction. However, the symmetries of the Riemann tensor mean that there are only three independent ones which we look at below

(i) Projecting all four indices onto Σ gives the *Gauss equation*,

$$ \perp^\mu_\alpha \perp^\beta_\nu \perp^\gamma_\rho \perp^\delta_\sigma {}^{(4)}R^\alpha{}_{\beta\gamma\delta} = R^\mu{}_{\nu\rho\sigma} + 2K^\mu{}_{[\rho}K_{\sigma]\nu}. \tag{2.1.26} $$

[8] This notation for the four dimensional Riemann tensor ${}^{(4)}\boldsymbol{R}$ is conventional in the numerical relativity literature since most of the time we will be using the spatial Ricci tensor.

Contracting on μ and ρ yields the *contracted Gauss equation*

$$\perp^{\beta}_{\nu}\perp^{\delta}_{\sigma} {}^{(4)}R_{\beta\delta} + n^{\alpha}\perp^{\beta}_{\nu} n^{\gamma}\perp^{\delta}_{\sigma} {}^{(4)}R_{\alpha\beta\gamma\delta} = R_{\nu\sigma} + KK_{\nu\sigma} - K^{\alpha}{}_{\sigma}K_{\alpha\nu}, \tag{2.1.27}$$

where $K = \gamma^{\alpha\beta}K_{\alpha\beta}$ is the trace of the extrinsic curvature. If we then contract again with $\gamma^{\nu\sigma}$, we obtain the *scalar Gauss equation*,

$${}^{(4)}R + 2{}^{(4)}R_{\alpha\beta}n^{\alpha}n^{\beta} = R + K^2 - K_{\alpha\beta}K^{\alpha\beta}. \tag{2.1.28}$$

(ii) Projecting three indices onto Σ and one in the normal direction yields the *Codacci equation*,

$$\perp^{\alpha}_{\mu}\perp^{\beta}_{\nu}\perp^{\gamma}_{\rho} n^{\delta}{}^{(4)}R_{\alpha\beta\gamma\delta} = D_{\nu}K_{\mu\rho} - D_{\mu}K_{\nu\rho}. \tag{2.1.29}$$

Contracting this with $\gamma^{\nu\rho}$, one obtains the *contracted Codacci equation*,

$$\perp^{\alpha}_{\mu} n^{\beta}{}^{(4)}R_{\alpha\beta} = D_{\mu}K - D_{\alpha}K^{\alpha}{}_{\mu}. \tag{2.1.30}$$

(iii) Finally, projecting two indices onto Σ and two in the normal direction, one obtains

$$n_{\alpha}\perp^{\beta}_{\mu} n^{\gamma}\perp^{\delta}_{\nu} {}^{(4)}R^{\alpha}{}_{\beta\gamma\delta} = \mathcal{L}_{\boldsymbol{n}}K_{\mu\nu} + \frac{1}{\alpha}D_{\mu}D_{\nu}\alpha + K_{\mu\alpha}K^{\alpha}{}_{\nu}. \tag{2.1.31}$$

This is sometimes referred to as the *Ricci equation*.

Note that both the Gauss (2.1.26) and Codacci (2.1.29) equations (often referred together as the Gauss-Codacci equations), do not feature the gauge variables nor the normal to the foliation on the right-hand-side, hence can be defined for a single hypersurface Σ embedded in $\mathscr{M}$. We have not yet used the Einstein equation (1.0.1) so these equations are purely geometrical. These equations relate the curvature of the spacetime $\mathscr{M}$ to the *intrinsic* curvature of the hypersurfaces Σ encoded in the spatial Ricci tensor and the *extrinsic* curvature of the embedding.

2.1.6 The constraint equations

In order to apply the Einstein equation (1.0.1), we first need to define our projections of the energy-momentum (EM) tensor $\boldsymbol{T}$. Given the symmetry of the tensor, we have three independent contractions as follows.

(i) Projecting both indices in the normal direction gives the *energy density*:

$$\rho = n^\alpha n^\beta T_{\alpha\beta}. \tag{2.1.32}$$

This is the energy density as measured by the Eulerian observer.

(ii) Projecting once in the normal direction and once spatially gives the *momentum density*:

$$j_\mu = -\perp^\alpha_\mu n^\beta T_{\alpha\beta}. \tag{2.1.33}$$

(iii) Finally, projecting both indices spatially gives the *stress tensor*:

$$S_{\mu\nu} = \perp^\alpha_\mu \perp^\beta_\nu T_{\alpha\beta}. \tag{2.1.34}$$

We will also use the shorthand $S = \gamma^{\mu\nu} S_{\mu\nu}$.

We can therefore write the energy-momentum tensor as

$$T_{\mu\nu} = \rho n_\mu n_\nu + n_\mu j_\nu + n_\nu j_\mu + S_{\mu\nu}. \tag{2.1.35}$$

Now, substituting the Einstein equation (1.0.1) and the definition of the energy density (2.1.32) into the scalar Gauss equation (2.1.28), gives the *Hamiltonian constraint equation*,

Hamiltonian constraint equation

$$\mathcal{H} \equiv R + K^2 - K_{ij}K^{ij} - 16\pi\rho = 0. \tag{2.1.36}$$

Similarly, substituting the Einstein equation (1.0.1) and the definition of the momentum density (2.1.33) into the contracted Codacci equation (2.1.30), gives the *momentum constraint equations*,

Momentum constraint equations

$$\mathcal{M}_i \equiv D_l K^l_{\ i} - D_i K - 8\pi j_i = 0. \tag{2.1.37}$$

Both of these equations involve only purely spatial quantities and their derivatives in Σ but do not involve any time derivatives or gauge variables, hence they are not evolution equations but rather constraints that the spatial metric $\boldsymbol{\gamma}$ and extrinsic curvature $\boldsymbol{K}$ must satisfy on each hypersurface Σ. From these equations, we can see that we are

unable to freely specify $\{\boldsymbol{\gamma}, \boldsymbol{K}\}$ on the initial hypersurface Σ_0 but rather must solve these constraints which are of *elliptic* type. We will discuss this in Sec. 2.2.

In the next subsection, we will obtain the evolution equations for the spatial metric and extrinsic curvature. If the constraint equations need to be satisfied on each hypersurface Σ, one might worry that they need to be solved at each timestep of a numerical evolution. Thankfully, it turns out that, due to the Bianchi identity, the constraints are preserved in the continuum limit, during a time evolution (for more details see Ref. [58]). In practice, since we do not work in the continuum limit, the constraints are never satisfied perfectly during a numerical evolution. Sometimes, we will even start with initial data that contain small violations of the constraints (see, for example, Sec. 6.4). If we consider the space of all possible solutions $\{\Sigma, \boldsymbol{\gamma}, \boldsymbol{K}\}$ (including those that violate Eqs. (2.1.36) and (2.1.37)), then we refer to the subspace of *physical* solutions that satisfy Eqs. (2.1.36) and (2.1.37) as the *constraint hypersurface* in solution space. Since we are interested in physical solutions, we will want to stay as close to the constraint hypersurface as possible. However, solving an elliptic equation can be computationally expensive, particularly if it needs to be done at each timestep. Therefore, we will take the *free evolution* approach where we do not enforce the constraints but rather monitor the quantities $\mathcal{H}$ and $\mathcal{M}_i$ to ensure they remain small[9]. Additionally, in Sec. 2.3.3, we will add terms to the evolution equations which have the property of damping constraint violations, bringing our time evolutions closer to the constraint hypersurface.

2.1.7 The ADM evolution equations

We are now ready to obtain the evolution equations. It turns out that we have already essentially met the evolution equation for $\boldsymbol{\gamma}$ in the form of Eq. (2.1.23). However, since we want evolution equations with respect to t in our adapted coordinate system, we will need to massage this equation into the right form. By direct calculation, it can be shown that $\mathcal{L}_{\alpha \boldsymbol{n}} \gamma_{ij} = \alpha \mathcal{L}_{\boldsymbol{n}} \gamma_{ij}$. Then, rearranging Eq. (2.1.10), substituting Eq. (2.1.23) and using the fact that, in adapted coordinates $\mathcal{L}_{\boldsymbol{t}} = \partial_t$, we obtain the evolution equation

$$\partial_t \gamma_{ij} = \mathcal{L}_{\boldsymbol{\beta}} \gamma_{ij} - 2\alpha K_{ij}. \tag{2.1.38}$$

[9] Unfortunately, since these quantities should vanish for physical solutions, it is unclear what we mean by "small". In practice, we might compare $\mathcal{H}$ with a term in Eq. (2.1.36), for example, $16\pi\rho$ and seek to ensure $\mathcal{H} \ll |16\pi\rho|$.

Deriving the evolution equation for $\boldsymbol{K}$ is a little more involved. We start with the Ricci equation (2.1.31) and eliminate the projection of the Riemann tensor using the contracted Gauss equation (2.1.27) to obtain

$$\perp^{\alpha}_{\mu}\perp^{\beta}_{\nu}{}^{(4)}R_{\alpha\beta} = -\mathcal{L}_{\boldsymbol{n}}K_{\mu\nu} - \frac{1}{\alpha}D_\mu D_\nu\alpha + R_{\mu\nu} + KK_{\mu\nu} - 2K_{\mu\alpha}K^{\alpha}{}_{\nu}. \tag{2.1.39}$$

As for the spatial metric, it can also be shown that $\mathcal{L}_{\alpha\boldsymbol{n}}K_{ij} = \alpha\mathcal{L}_{\boldsymbol{n}}K_{ij}$. Then substituting the Einstein equation (1.0.1) and rearranging, we obtain the evolution equation

$$\partial_t K_{ij} = \mathcal{L}_{\boldsymbol{\beta}}K_{ij} - D_iD_j\alpha + \alpha\left(R_{ij} - 2K_{ik}K^{k}{}_{j} + KK_{ij}\right) + 4\pi\alpha\left[\gamma_{ij}(S-\rho) - 2S_{ij}\right]. \tag{2.1.40}$$

The full ADM evolution system is thus

ADM evolution system

$$\partial_t\gamma_{ij} = \mathcal{L}_{\boldsymbol{\beta}}\gamma_{ij} - 2\alpha K_{ij}, \tag{2.1.41a}$$

$$\partial_t K_{ij} = \mathcal{L}_{\boldsymbol{\beta}}K_{ij} - D_iD_j\alpha + \alpha\left(R_{ij} - 2K_{ik}K^{k}{}_{j} + KK_{ij}\right) + 4\pi\alpha\left[\gamma_{ij}(S-\rho) - 2S_{ij}\right], \tag{2.1.41b}$$

$$0 = R + K^2 - K_{ij}K^{ij} - 16\pi\rho, \tag{2.1.41c}$$

$$0 = D_lK^{l}{}_{i} - D_iK - 8\pi j_i. \tag{2.1.41d}$$

Unfortunately, it turns out these equations are usually not stable enough for numerical evolution and this is because they are only *weakly hyperbolic* [49]. In Sec. 2.3, we will look at reformulations of these equations that are suitably stable for numerical evolution and that we use for the simulations presented in this thesis.

2.1.8 The Z4 formulation

In Sec. 2.1.6, we mentioned that, because the constraints are elliptic equations, they are computationally expensive to solve at each timestep. However, there is a way to convert these equations to hyperbolic evolution equations. This involves the addition of a new 4-vector $\boldsymbol{Z}$ [59] which can be added to the Einstein-Hilbert action [60] to obtain the action

$$S_{\text{Z4}} = \frac{1}{16\pi}\int \mathrm{d}^4x\,\sqrt{-g}[{}^{(4)}R + 2g^{\mu\nu}\nabla_\mu Z_\nu] + S_M. \tag{2.1.42}$$

Here and in what follows, we use the colour blue to indicate terms that arise specifically from the Z4 vector $\boldsymbol{Z}$. Varying the action using the Palitini procedure (where the metric $\boldsymbol{g}$ is varied independently of the connection ∇)[10] yields the equations

$$ {}^{(4)}R_{\mu\nu} + 2\nabla_{(\mu}Z_{\nu)} = 8\pi\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right), \tag{2.1.43a} $$

$$ \nabla_\rho g^{\mu\nu} = 0, \tag{2.1.43b} $$

$$ Z_\mu = 0, \tag{2.1.43c} $$

where $T = g^{\mu\nu}T_{\mu\nu}$. These equations are manifestly equivalent to the Einstein equation (1.0.1), so long as we impose $\boldsymbol{Z} = 0$. Since we are taking the "free evolution" approach with respect to the Hamiltonian and momentum constraints (2.1.36)–(2.1.37), one might ask how $\boldsymbol{Z}$ evolves if we do not impose this new constraint. Using the Bianchi identity and the Z4 Einstein equation (2.1.43a), it follows that $\boldsymbol{Z}$ satisfies

$$ \Box_g Z_\mu + R_{\mu\alpha}Z^\alpha = 0, \tag{2.1.44} $$

which is just a [forced] wave equation. Since $\boldsymbol{Z} = 0$ corresponds to physical solutions, $\boldsymbol{Z}$ provides some measure of the deviation from physical solutions of GR, but given that it also satisfies a wave equation, we can see that these deviations propagate away.

We now perform a 3+1 split of Eq. (2.1.43a). First, we introduce the normal and spatial projections of $\boldsymbol{Z}$ as[11]

$$ \Theta = -n_\mu Z^\mu, \qquad \Theta_\mu = \perp^\alpha_\mu Z_\alpha, \qquad \boldsymbol{Z} = \Theta\boldsymbol{n} + \boldsymbol{\Theta}. \tag{2.1.45} $$

Then, the Z4 equations in the 3+1 formalism are

[10] Note that the Levi-Civita connection might need to be assumed in the matter part of the action S_M i.e. not using the Palitini procedure for S_M.

[11] The literature [59, 61] seems to only distinguish between $\boldsymbol{Z}$ and its spatial projection through the type of the index, i.e. Z_μ and Z_i respectively. To avoid confusion, we use $\boldsymbol{\Theta}$ for the spatial projection of $\boldsymbol{Z}$.

3+1 Z4 equations

$$\partial_t \gamma_{ij} = \mathcal{L}_\beta \gamma_{ij} - 2\alpha K_{ij}, \tag{2.1.46a}$$

$$\partial_t K_{ij} = \mathcal{L}_\beta K_{ij} - D_i D_j \alpha + \alpha \Big(R_{ij} + 2D_{(i}\Theta_{j)} - 2K_{ik}K^k{}_j + (K - 2\Theta)K_{ij} \Big) + 4\pi\alpha \left[\gamma_{ij}(S - \rho) - 2S_{ij} \right], \tag{2.1.46b}$$

$$\partial_t \Theta = \mathcal{L}_\beta \Theta + \frac{\alpha}{2} \left[\mathcal{H} + 2D_i \Theta^i - 2\Theta K - 2\partial_k(\log\alpha)\Theta^k \right], \tag{2.1.46c}$$

$$\partial_t \Theta_i = \mathcal{L}_\beta \Theta_i + \alpha \left[\mathcal{M}_i + \partial_i \Theta - \partial_i(\log\alpha)\Theta - 2K_i{}^k \Theta_k \right], \tag{2.1.46d}$$

where $\mathcal{H}$ and $\mathcal{M}_i$ are given by Eq. (2.1.36) and Eq. (2.1.37) respectively. We can see that the elliptic constraint equations $\mathcal{H} = 0 = \mathcal{M}_i$ have become hyperbolic evolution equations for Θ and $\boldsymbol{\Theta}$.

Note that the vanishing of the physical constraints $\mathcal{H}$ and $\mathcal{M}_i$ is equivalent to the vanishing of Θ, $\partial_t\Theta, \Theta_i$ and $\partial_t\Theta_i$ (assuming Θ and $\boldsymbol{\Theta}$ vanish initially). We can also see that the introduction of $\boldsymbol{Z}$ has made more explicit what we meant by the "constraint hypersurface" of solution space mentioned in Sec. 2.1.6. If we consider the solution space of Eq. (2.1.43a) for all arbitrary $\boldsymbol{Z}$, then the constraint hypersurface of *physical* solutions is just the subspace where $\boldsymbol{Z}$ vanishes.

2.2 Initial data

In the previous section, we have seen that, when we perform a 3+1 split, we cannot freely specify the spatial metric $\boldsymbol{\gamma}$ and extrinsic curvature $\boldsymbol{K}$ on the initial hypersurface. Instead, they must satisfy the constraint equations (2.1.36)–(2.1.37) which are elliptic PDEs. In this section, we will briefly outline some methods to solve these equations and the *Bowen-York* solution [62] that can be used to describe multiple spinning boosted BHs and will be used in several simulations presented later in this thesis. We should point out that we will only consider the York-Lichnerowicz conformal decomposition [63, 64, 65, 66, 67] and not the [extended] conformal thin sandwich approach [68, 69] which is popular for neutron star initial data. For a more detailed overview of initial data methods, the reader is referred to Ref. [70].

2.2.1 The York-Lichnerowicz conformal decomposition

The York-Lichnerowicz decomposition starts by conformally rescaling the spatial metric [63, 64, 65] as[12]

$$\gamma_{ij} = \bar{\psi}^4 \bar{\gamma}_{ij}, \tag{2.2.1}$$

and decomposing the extrinsic curvature $\boldsymbol{K}$ into its trace K and trace-free parts $\boldsymbol{A}$ *viz.*

$$K^{ij} = A^{ij} - \frac{1}{3}\gamma^{ij} K. \tag{2.2.2}$$

The Hamiltonian constraint (2.1.36) becomes

$$8\bar{D}^2\bar{\psi} - \bar{R}\bar{\psi} + \bar{\psi}^5\left(A_{kl}A^{kl} - \frac{2}{3}K^2\right) + 16\pi\bar{\psi}^5\rho = 0, \tag{2.2.3}$$

where $\bar{D}$ is the Levi-Civita connection associated to $\bar{\gamma}_{ij}$, $\bar{D}^2 = \bar{D}^l\bar{D}_l$ and $\bar{R}$ is the Ricci scalar associated to $\bar{D}$. The elliptic nature of the Hamiltonian constraint is more explicit since Eq. (2.2.3) is an elliptic equation for $\bar{\psi}$.

Next, we use a result from Refs. [66, 67] that a symmetric traceless tensor S^{ij} can be decomposed into its transverse[13] and longitudinal parts.

$$S^{ij} = \underbrace{S_*^{ij}}_{\text{transverse}} + \underbrace{(\mathbb{L}_{\boldsymbol{\gamma}}\bar{\boldsymbol{W}})^{ij}}_{\text{longitudinal}}, \tag{2.2.4}$$

where S_*^{ij} is symmetric, traceless and transverse, $\bar{\boldsymbol{W}}$ is a vector and $\mathbb{L}_{\boldsymbol{h}}$ is a linear differential operator associated to a metric $\boldsymbol{h}$ and defined by

$$(\mathbb{L}_{\boldsymbol{\gamma}}\bar{\boldsymbol{W}})^{ij} := D^iW^j + D^jW^i - \frac{2}{3}\gamma^{ij}D_lW^l. \tag{2.2.5}$$

Now, there is a choice as to whether we use the differential operator associated to the physical metric $\mathbb{L}_{\boldsymbol{\gamma}}$, or the conformal metric $\mathbb{L}_{\tilde{\boldsymbol{\gamma}}}$ in Eq. (2.2.4). For simplicity, we will only consider the latter, which is referred to in the literature as the *conformal transverse-traceless decomposition*, as that is what is relevant for Bowen-York data in the next section. The reader is referred to Refs. [70, 49] for a discussion of the former

[12] Following Ref. [49] but in contrast to Ref. [70], we use a bar to denote a conformal quantity and reserve tilde for the specific conformal rescaling that makes the conformal metric determinant unity $\tilde{\gamma} = \det\tilde{\boldsymbol{\gamma}} = 1$. However, we also use $\bar{\psi}$ for the conformal factor to avoid confusion with ψ used in other chapters.

[13] A transverse tensor $\boldsymbol{T}$ has zero divergence in every index e.g. $D_{i_p}T^{i_1\cdots i_p\cdots i_r} = 0$.

We start by conformally rescaling the tracefree part of the extrinsic curvature as

$$A^{ij} = \bar{\psi}^{-10}\bar{A}^{ij} \iff A_{ij} = \bar{\psi}^{-2}\bar{A}_{ij}, \tag{2.2.6}$$

where we lower and raise indices of conformal tensors with the conformal metric $\bar{\gamma}_{ij}$ and its inverse.

Applying the decomposition (2.2.4) to the conformal tracefree extrinsic curvature,

$$\bar{A}^{ij} = \bar{A}^{ij}_{*} + (\mathbb{L}_{\bar{\gamma}}\bar{\boldsymbol{W}})^{ij}, \tag{2.2.7}$$

the momentum constraint (2.1.37) reduces to

$$\bar{\Delta}_{\mathbb{L}_{\bar{\gamma}}}\bar{W}^i - \frac{2}{3}\bar{\psi}^6\bar{D}^i K - 8\pi\bar{\psi}^{10}j^i = 0, \tag{2.2.8}$$

where the linear differential operator $\bar{\Delta}_{\mathbb{L}_{\bar{\gamma}}}$ is defined by

$$\bar{\Delta}_{\mathbb{L}_{\bar{\gamma}}}\bar{W}^i := \bar{D}_l(\mathbb{L}_{\bar{\gamma}}\bar{\boldsymbol{W}})^{il} = \bar{D}^2\bar{W}^i + \frac{1}{3}\bar{D}^i\bar{D}_l\bar{W}^l + \bar{R}^i{}_l\bar{W}^l. \tag{2.2.9}$$

Unfortunately, constructing transverse tensors such as $\bar{A}^{ij}_{*}$ is non-trivial, so, if we start with an arbitrary symmetric tracefree tensor $\bar{M}^{ij}$, we decompose it as in Eq. (2.2.4) and find its transverse part to be

$$\bar{M}^{ij}_{*} = \bar{M}^{ij} - (\mathbb{L}_{\bar{\gamma}}\bar{\boldsymbol{U}})^{ij}, \tag{2.2.10}$$

where we find $\bar{\boldsymbol{U}}$ by solving

$$\bar{\Delta}_{\mathbb{L}_{\bar{\gamma}}}\bar{U}^i = \bar{D}_l\bar{M}^{il}. \tag{2.2.11}$$

We can do all of this together by taking $\bar{A}^{ij}_{*} = \bar{M}^{ij}_{*}$, defining $\bar{\boldsymbol{V}} := \bar{\boldsymbol{W}} - \bar{\boldsymbol{U}}$ and using the linearity of $\bar{\Delta}_{\mathbb{L}_{\bar{\gamma}}}$, so the conformal tracefree extrinsic curvature is given by

$$\bar{A}^{ij} = \bar{M}^{ij}_{*} + (\mathbb{L}_{\bar{\gamma}}\bar{\boldsymbol{W}})^{ij} = \bar{M}^{ij} + (\mathbb{L}_{\bar{\gamma}}\bar{\boldsymbol{V}})^{ij}. \tag{2.2.12}$$

Combining all the transformations in Eqs. (2.2.1), (2.2.2), (2.2.6) and (2.2.12), the Hamiltonian and momentum constraints become

York-Lichnerowicz conformal decomposition of the constraints

$$8\bar{D}^2\bar{\psi} - \bar{R}\bar{\psi} + \bar{\psi}^{-7}\bar{A}_{kl}\bar{A}^{kl} - \frac{2}{3}\bar{\psi}^5 K^2 + 16\pi\bar{\psi}^5\rho = 0, \tag{2.2.13a}$$

$$\bar{\Delta}_{\mathbb{L}_{\bar{\gamma}}}\bar{V}^i + \bar{D}_k\bar{M}^{ik} - \frac{2}{3}\bar{\psi}^6\bar{D}^i K - 8\pi\bar{\psi}^{10}j^i = 0. \tag{2.2.13b}$$

In this decomposition, the conformal metric $\bar{\gamma}_{ij}$, the symmetric tracefree tensor $\bar{M}^{ij}$, the trace of the extrinsic curvature K, and the matter energy density ρ and momentum density j^i are specified and the system (2.2.13) is solved with appropriate boundary conditions for $\bar{\psi}$ and $\bar{V}^i$.

Note that these equations are coupled via the conformal factor $\bar{\psi}$, and the decomposed parts of the extrinsic curvature K and A^{ij}. However, there are assumptions one can make in order to decouple them.

One assumption is to choose K constant corresponding to constant mean curvature (so $D^i K$ vanishes in Eq. (2.2.13b)). Then, if j^i vanishes, the momentum constraint (2.2.13b) can be solved first for $\bar{V}^i$, $\bar{A}^{ij}$ can be reconstructed using Eq. (2.2.12) and then the Hamiltonian constraint (2.2.13a) can be solved for $\bar{\psi}$.

Another common assumption, that we will make in the next section, is that the conformal metric is flat, that is $\bar{\gamma}_{ij} = f_{ij}$, where f_{ij} is the flat metric ($f_{ij} = \delta_{ij}$ in Cartesian coordinates). Then $\bar{R} = 0$, the conformal covariant derivative $\bar{D}$ becomes the usual flat-space one which we denote as $\breve{\nabla}$, and the Hamiltonian constraint reduces to

$$8\breve{\nabla}^2\bar{\psi} + \bar{\psi}^{-7}\bar{A}_{kl}\bar{A}^{kl} - \frac{2}{3}\psi^5 K^2 + 16\pi\bar{\psi}^5\rho = 0. \tag{2.2.14}$$

Furthermore, if we take $K_{ij} = 0$ which corresponds to *time-symmetric initial data*[14], Eq. (2.2.14) becomes

$$\breve{\nabla}^2\bar{\psi} + 2\pi\bar{\psi}^5\rho = 0. \tag{2.2.15}$$

2.2.2 Brill-Lindquist initial data

If we seek conformally-flat, vacuum, time-symmetric initial data so that $\gamma_{ij} = f_{ij}$ and $0 = \rho = j^i = K_{ij}$, then the Hamiltonian constraint (2.2.15) becomes the usual flat-space

[14]This term comes from the fact that $K_{ij} = 0$ means that $\mathcal{L}_{\boldsymbol{n}}\gamma$ vanishes. Physically this corresponds to initial data that is instantaneously at rest and looks the same if one evolves forwards or backwards in time. For compact object binaries, this manifests itself as the two objects, having been moving away from each other, being instantaneously stationary as gravity causes them to accelerate back towards each other.

Laplace equation

$$\breve{\nabla}^2 \bar{\psi} = 0. \tag{2.2.16}$$

If we solve this system subject to the boundary condition $\lim_{R\to\infty} \bar{\psi} = 1$ corresponding to an asymptotically flat end, the simplest non-trivial solution is

$$\bar{\psi} = 1 + \frac{k}{R}, \tag{2.2.17}$$

where R is the coordinate radius. The initial spatial metric is thus

$$\boldsymbol{\gamma} = \left(1 + \frac{k}{R}\right) f_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^j, \tag{2.2.18}$$

which is just the spatial part of the Schwarzschild metric [9] in isotropic coordinates. This describes a spherically-symmetric black hole of mass $M = 2k$.

Furthermore, since Eq. (2.2.16) is linear, we can superpose solutions to obtain

$$\bar{\psi}(\mathbf{x}) = 1 + \sum_{i=1}^{N} \frac{m_i}{2|\mathbf{x} - \mathbf{x}_i|}, \tag{2.2.19}$$

which describes initial data for N black holes with *bare masses*[15] m_i and centres $\mathbf{x}_i$ so long as the centres are sufficiently well separated and the horizons are distinct. This *analytic* solution is known as *Brill-Lindquist* initial data [71, 72]. Although the time-symmetry of the solution means that it is not typically the most astrophysically interesting/relevant solution, its simplicity means that it is often used as a test case in numerical relativity codes.

It should be pointed out that since Eq. (2.2.19) is singular at the centres $\mathbf{x} = \mathbf{x}_i$, these points are formally removed from the manifold, that is $\Sigma = \mathbb{R}^3 \setminus \{\mathbf{x}_i\}_{i=1}^N$. We commonly refer to these points as the *punctures*. Furthermore, as is the case for the single Schwarzschild BH (2.2.18), each puncture corresponds to spatial infinity of a different asymptotically flat end. This means that, topologically, the spacetime has $N+1$ asymptotically flat ends. However, since N of these asymptotically flat ends are hidden behind horizons, they are causally disconnected from the rest of the spacetime and this topological complexity should not affect the time evolution. Finally, we note that the

[15] Note the bare mass differs to other mass definitions (e.g. the ADM mass) unless $N = 1$.

ADM mass (see Appendix A) of the ith BH in this solution is [49]

$$M_{\mathrm{ADM},i} = m_i \left(1 + \sum_{j=0, j\neq i}^{N} \frac{m_j}{|\mathbf{x}_j - \mathbf{x}_i|}\right). \tag{2.2.20}$$

2.2.3 Bowen-York initial data

Since astrophysical black holes are expected to have spin and linear momentum, we will need something more generic than Brill-Lindquist initial data. We start by taking the same assumptions of conformal flatness and vacuum, but this time relax the time-symmetric assumption and instead require $K = 0$, which corresponds to maximal slicing[16]. Choosing $\bar{M}^{ij} = 0$ means the momentum constraint (2.2.13b) reduces to

$$0 = \bar{\Delta}_{\mathbb{L}_{\boldsymbol{f}}} \bar{V} \equiv \breve{\nabla}^2 \bar{V}^i + \frac{1}{3} \breve{\nabla}^i \breve{\nabla}_j \bar{V}^j = 0. \tag{2.2.21}$$

This has the analytic solution[17]

$$\bar{V}^i = -\frac{1}{4R}\left[7P^i + \frac{s^i s_j P^j}{R^2}\right] + \frac{1}{R^3} \bar{\epsilon}^{ijk} s_j S_k, \tag{2.2.22}$$

where P^i and S^i are constant vectors, $s^i = (x^i - x^i_0)/|\mathbf{x} - \mathbf{x}_0|$ (in Cartesian coordinates) is the outward pointing unit radial vector centred at $\mathbf{x}_0$ and $\bar{\epsilon}^{ijk}$ is the Levi-Civita tensor in three dimensions. The resulting conformal trace-free extrinsic curvature tensor is

$$\begin{aligned} {}^{(\mathrm{BY})}\bar{A}_{ij}(\mathbf{x}; \mathbf{x}_0, \mathbf{P}, \mathbf{S}) &= (\mathbb{L}_{\boldsymbol{f}} \bar{\boldsymbol{V}})_{ij} && (2.2.23)\\ &= \frac{3}{2R^2}\left[s_i P_j + s_j P_i + s_k P^k (s_i s_j - f_{ij})\right] \\ &\quad - \frac{3}{R^3}\left(\bar{\epsilon}_{ilk} s_j + \bar{\epsilon}_{jlk} s_i\right) s^l S^k. && (2.2.24) \end{aligned}$$

and the physical extrinsic curvature is recovered with $K_{ij} = {}^{(\mathrm{BY})}\bar{A}_{ij}\bar{\psi}^{-2}$. This solution of the momentum constraint is known as the *Bowen-York* solution [73, 62].

Under the assumption that $\bar{\psi} \to 1$ as $R \to \infty$ (asymptotic flatness), it turns out that expressions for the ADM linear (A.0.2) and angular momentum (A.0.3) can be calculated

[16] This is because $K = 0$ maximises the action

$$V = \int_\Sigma \sqrt{\gamma}\, \mathrm{d}^3 x.$$

See Sec. 10.2.2 in Ref. [51] for details.

[17] See Appendix B of Ref. [50] for methods on how to solve this equation.

without solving the Hamiltonian constraint for $\bar{\psi}$. They are, respectively,

$$P^i_{\rm ADM} = P^i, \qquad J^i_{\rm ADM} = S^i. \tag{2.2.25}$$

The linearity of Eq. (2.2.22) means that we can superpose multiple solutions,

$$\bar{A}_{ij}(\mathbf{x}) = \sum_{i=1}^{N} {}^{\rm (BY)}\bar{A}_{ij}(\mathbf{x}; \mathbf{x}_i, \mathbf{P}_i, \mathbf{S}_i), \tag{2.2.26}$$

corresponding to multiple sources of linear and angular momentum. As one might expect intuitively, the total linear momentum in the spacetime for such a superposed solution is just the vector sum of the individual linear momenta. Calculating the total angular momentum is more involved due to the contribution of the linear momentum about the centre of mass to the orbital angular momentum. Nevertheless, it is still possible to do without solving the Hamiltonian constraint. One example, that will be particularly relevant to some of the simulations presented later on, is if the two sources are centred at $\mathbf{x} = \mathbf{d}$ and $\mathbf{x} = -\mathbf{d}$ with momenta $\mathbf{P}$ and $-\mathbf{P}$ respectively and $\mathbf{d} \cdot \mathbf{P} = 0$. Then, surprisingly[18], the ADM angular momentum is given by [74]

$$\mathbf{J}_{\rm ADM} = 2\mathbf{d} \times \mathbf{P}. \tag{2.2.27}$$

Unfortunately, solving the Hamiltonian constraint (2.2.14) with $\bar{A}_{ij}$ given by Eq. (2.2.26) cannot be done analytically. Given that we expect a solution with behaviour similar to the Brill-Lindquist solution (2.2.19), starting with the ansatz

$$\bar{\psi} = \bar{\psi}_{\rm BL} + u, \qquad \bar{\psi}_{\rm BL} = \sum_{i=1}^{N} \frac{m_i}{2|\mathbf{x} - \mathbf{x}_i|}, \tag{2.2.28}$$

reduces the Hamiltonian constraint to

$$\breve{\nabla}^2 u + \left(\bar{\psi}_{\rm BL} + u\right)^{-7} \bar{A}_{kl}\bar{A}^{kl} = 0. \tag{2.2.29}$$

Remarkably, it turns out there exists a C^2 solution to this equation in all of $\mathbb{R}^3$ (including at the punctures of $\bar{\psi}_{\rm BL}$) with the outer boundary condition $\partial_R u = (1-u)/R$ (equivalent to $u - 1 \propto 1/R$ corresponding to an asymptotically flat end) [75]. This is known as *puncture initial data*. The ADM mass of each black hole in this solution (i.e. the

[18] Note that $2\mathbf{d}$ is the *coordinate* separation of the two sources and one might naively assume the *proper* distance would be required for this expression.

generalization of Eq. (2.2.20)) is

$$M_{\mathrm{ADM},i} = m_i \left(1 + u(\mathbf{x}_i) + \sum_{j=0, j\neq i}^{N} \frac{m_j}{|\mathbf{x}_j - \mathbf{x}_i|} \right). \tag{2.2.30}$$

Given that the unique *stationary*, axisymmetric black-hole solution to the Einstein equation is the Kerr spacetime [15, 16], one might hope that we have found initial data that "evolves" into this (at least for $N = 1$ and $\mathbf{P} = \mathbf{0}$). This would mean we have stumbled upon a conformally flat slicing of the Kerr spacetime. Unfortunately, under some reasonable assumptions[19], no such slicing exists [76]. We must therefore conclude that the evolution of this initial data cannot be stationary. It turns out that, when this initial data is evolved in time, there is an initial burst of spurious gravitational waves which is commonly referred to as "junk" radiation. Similar observations are made for $\mathbf{P} \neq \mathbf{0}$ (independent of $\mathbf{S}$). An example of this spurious radiation is shown in Fig. 2.3. Despite this issue, Bowen-York puncture data has proved very robust over the past two decades. It is usually straightforward to remove the contribution of the junk radiation to calculations involving gravitational radiation by suitably truncating the signal.

2.2.4 Issues of uniqueness with matter

Consider the Hamiltonian constraint for conformally-flat, time-symmetric initial data (2.2.15). Schematically, this is of the form

$$\breve{\nabla}^2 v + f v^n = 0, \tag{2.2.31}$$

where f is a smooth function on the domain $\mathbb{R}^3$ and we shall assume a boundary condition of the form $v - 1 = \mathcal{O}(1/R)$ as $R \to \infty$. We typically solve nonlinear problems like this by iteratively solving their linearisation. Writing $v = v_0 + \epsilon v_1$, the equation at $\mathcal{O}(\epsilon)$ becomes

$$\breve{\nabla}^2 v_1 + n f v_0^{n-1} v_1 = 0, \tag{2.2.32}$$

and the boundary condition for v_1 is $v_1 = \mathcal{O}(1/R)$ as $R \to \infty$. Suppose we have two solutions for this boundary value problem v_1 and v_1'. Let $w = v_1 - v_1'$. Then since

[19] These include that the foliation is axisymmetric and smoothly reduces to slices of constant Schwarzschild time in the limit of no spin.

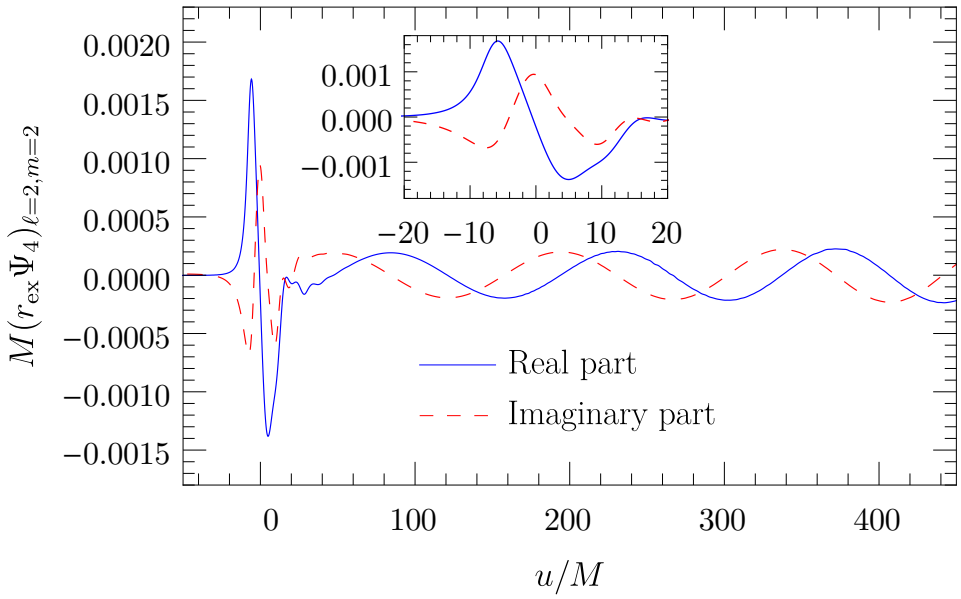


Fig. 2.3 The real and imaginary parts of the dominant $(\ell, m) = (2, 2)$ multipole of the Weyl scalar Ψ_4 (which provides a measure of the gravitational radiation. See Sec. 2.4) extracted at $R = r_{\rm ex} = 80M$. This is from the simulation of an equal-mass, non-spinning BH binary quasicircular inspiral with Bowen-York puncture initial data (2.2.26) and shows the spurious "junk" radiation in a small interval around $u = 0$. $u = t - r^*_{\rm ex}$ is the retarded time so $u = 0$ corresponds to roughly when the centre of the computational domain is in causal contact with the extraction sphere. The inset shows the subset of the signal where the junk radiation dominates.

Eq. (2.2.32) is linear, w also satisfies it with the same boundary condition. Consider

$$I[w] := \int_{\mathbb{R}^3} \breve{\nabla}^k (w \breve{\nabla}_k w)\, \mathrm{d}^3x = \int_{\mathbb{R}^3} |\breve{\boldsymbol{\nabla}} w|^2\, \mathrm{d}^3x - \int_{\mathbb{R}^3} nf v_0^{n-1} w^2\, \mathrm{d}^3x, \tag{2.2.33}$$

where we have applied Eq. (2.2.32) in the final equality. Now, assuming $nf \leq 0$ and $v_0^{n-1} \geq 0$, it immediately follows that $I[w] \geq 0$. However, we can also apply the divergence theorem to see that

$$I[w] = \lim_{R\to\infty} \int_{S^2_R} w \frac{\partial w}{\partial R} \sqrt{h}\, \mathrm{d}^2x, \tag{2.2.34}$$

where S^2_R is the 2-sphere of radius R and h is the determinant of the induced metric on S^2_R. Since $w = \mathcal{O}(1/R)$, it follows that $\partial_R w = \mathcal{O}(1/R^2)$. Furthermore, $\sqrt{h} = \mathcal{O}(R^2)$ so, in the limit, $I[w]$ vanishes. We conclude that w vanishes identically on $\mathbb{R}^3$ and the solutions to Eq. (2.2.32) are unique.

Unfortunately, looking at Eq. (2.2.15) we find $f = 2\pi\rho$ and $n = 5$. Assuming a nonnegative matter energy density $\rho \geq 0$ that is specified a priori[20], we find $nf \geq 0$, hence the uniqueness result does not apply. We can recover uniqueness if we conformally rescale the matter energy density, for example $\rho = \bar{\psi}^{-8}\bar{\rho}$, and consider $\bar{\rho}$ as specified. However this is undesirable for fundamental fields such as the complex scalar fields which comprise boson stars.

Although, this may seem like a dealbreaker, there is evidence to suggest that the failure of uniqueness is not catastrophic [77, 78]. Rather than an infinite number of solutions, there are usually two solution branches and it is possible to obtain a local uniqueness result under certain assumptions [77]. This means that if one starts with an initial guess (i.e. u_0 in the iterative scheme) that is sufficiently close to the final solution, it is still possible for numerical algorithms to converge to the desired physically relevant solution. On the flip-side, if the initial guess is too far away from the desired solution, it is also possible to converge to an undesired solution that lies on the other branch.

2.3 Evolution

Since our ultimate aim is to perform numerical evolutions of the initial data discussed in the previous section, we will need a formulation of the Einstein equation that is numerically stable. Unfortunately, as alluded to at the end of Sec. 2.1.7, the ADM equations do not have the requisite *hyperbolicity* properties for a wide variety of problems of interest so it is necessary to seek alternative formulations. Explorations of modifications to the ADM equations for use in NR simulations had been conducted since at least the 1980s before the hyperbolicity properties of these formulations were fully understood (see Figs. 3 and 4 in Ref. [79]), but it wasn't until the late 1990s/early 2000s that significant progress was made [41].

In this section we start by briefly explaining the notion of *well-posedness* and the relation with the hyperbolicity properties of PDE systems. For a detailed discussion, the reader is referred to Refs. [49, 80]. Next, we introduce the Baumgarte-Shapiro-Shibata-Nakamura(-Oohara-Kojima) [BSSN(OK)][21] formulation [81, 82, 83], by far the most successful formulation of the Einstein equation based on the 3+1 formalism over the past two decades. We will also discuss the conformal and covariant Z4 (CCZ4) formulation

[20] This is not exactly true for the boson stars we will consider in Chapter 6 as ρ also depends on $\bar{\psi}$ but a similar argument applies.

[21] Note that in the literature, it is often referred to as just BSSN hence the use of parentheses around the "OK".

[84, 85] that has become popular in recent years and has been used for the majority of the simulations presented in this work. Finally, we will look at the moving-puncture gauge conditions that, in conjunction with the BSSN(OK) formulation, have proved remarkable robust for a wide class of numerical relativity problems.

2.3.1 Well-posedness and hyperbolicity

A necessary but not sufficient condition for a PDE system that admits stable numerical evolution is that the system is *well-posed*. Consider a general PDE system in first-order form[22],

$$\partial_t \mathbf{u} + \mathsf{M}^i \partial_i \mathbf{u} = \mathbf{f}(\mathbf{u}), \tag{2.3.1}$$

where the components of the m-dimensional vector $\mathbf{u}$ are the m fields we wish to solve for, and, for each i, M^i is an $m \times m$ matrix of coefficients. A well-posed system has solutions that depend continuously on their initial data, that is if we make a small change to the initial data, then the corresponding change in the evolved solution is also small. More precisely, we say the system is well-posed if there exist constants k and p and a norm $\|\cdot\|$, independent of the initial data, such that

$$\|\mathbf{u}(t, \mathbf{x})\| \leq k \mathrm{e}^{pt} \|\mathbf{u}(0, \mathbf{x})\|, \tag{2.3.2}$$

or, in other words, the growth in the solutions are at most exponential.

Given the finite precision of computers, there will always be some numerical noise in initial data. It is therefore reasonably straightforward to see that if solutions of a PDE system fail to satisfy Eq. (2.3.2), then we may run into problems quickly when evolving simulations, or, at the very least, such a system will be unsuitable for obtaining reliable and accurate results.

The well-posedness of a PDE system is related to its *hyperbolicity*. For the purpose of this discussion, we will restrict to the case of Eq. (2.3.1) with each M^i having constant coefficients and $\mathbf{f} \equiv \mathbf{0}$. In the more general case where $\mathsf{M}^i = \mathsf{M}^i(t, \mathbf{x}, \mathbf{u})$, we linearise about a background solution $\mathbf{u}_0$ and consider the local form of the M^i. In the case $\mathbf{f}$ is linear in the u's, it turns out that the well-posedness follows straightforwardly from the well-posedness of the unforced system [49], hence it is sufficient to consider the unforced system. The main difference between the general nonlinear case and the special linear

[22] Do not confuse the vectors denoted in bold-font here, which are just an ordered list of components, with the vectors defined in a differential geometry sense that are relevant to GR.

case we will consider is that the analysis can only show that solutions exist locally in time.

For an arbitrary unit vector $\mathbf{n}$, we define the *principal symbol* of the PDE system (2.3.1) as[23]

$$\mathsf{P}(\mathbf{n}) = \mathsf{M}^i n_i. \tag{2.3.3}$$

We then characterise the hyperbolicity of the system (2.3.1) depending on the type of eigenvalues and the number of linearly independent eigenvectors as follows. We say the system is

(i) *weakly hyperbolic* if all of the eigenvalues of $\mathsf{P}(\mathbf{n})$ are real.

(ii) *strongly hyperbolic* if it is weakly hyperbolic and it has m linearly independent eigenvectors. It follows that there exist positive-definite symmetric matrices $\mathsf{H}(\mathbf{n})$ such that HP is symmetric:

$$\mathsf{HP} - \mathsf{P}^T\mathsf{H}^T = \mathsf{HP} - \mathsf{P}^T\mathsf{H} = 0. \tag{2.3.4}$$

We call H the *symmetriser*.

(iii) *symmetric hyperbolic* if it is strongly hyperbolic and the symmetriser is independent of $\mathbf{n}$.

Note that in the above list, each notion of hyperbolicity implies the previous but the converse is not true.

The existence of the symmetriser H allows us to define an inner product (and subsequently a norm),

$$\langle \mathbf{u}, \mathbf{v} \rangle := \mathbf{u}^\dagger \mathsf{H} \mathbf{v}, \qquad \|\mathbf{u}\| := \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}. \tag{2.3.5}$$

To see how the hyperbolicity of the system is related to its well-posedness, consider a Fourier mode $\mathbf{u}(t, \mathbf{x}) = \tilde{\mathbf{u}}(t) \mathrm{e}^{\mathrm{i} k \mathbf{x} \cdot \mathbf{n}}$. Its norm satisfies

$$\partial_t \|\mathbf{u}\|^2 = \partial_t (\mathbf{u}^\dagger \mathsf{H} \mathbf{u}) = \mathrm{i} k \tilde{\mathbf{u}}^T (\mathsf{P}^T \mathsf{H} - \mathsf{HP}) \tilde{\mathbf{u}} = 0, \tag{2.3.6}$$

where we have applied the equation of motion (2.3.1) with the definition of the principal symbol (2.3.3) in the second equality and used Eq. (2.3.4) in the final equality, hence

[23] We follow here the conventions of Ref. [49] and do not multiply the principal symbol by the imaginary unit i. However, it is commonplace in the literature to do so, for example Ref. [80]

this result only holds for strongly or symmetric hyperbolic systems. We can see that its norm is constant in time and this mode satisfies the definition of well-posedness (2.3.2). In the more general case, where $\mathsf{M}^i = \mathsf{M}^i(t, \mathbf{x}, \mathbf{u})$, the norm remains bounded rather than staying constant. This illustrates why strong hyperbolicity is a necessary condition for well-posedness.

Typically, when analysing the hyperbolicity of a system, rather than considering the principal symbol, we look at the *eigenfields* or *eigenfunctions*. For a strongly hyperbolic system if R is the matrix of eigenvectors of the principal symbol, the *eigenfields* are related to the fundamental evolution variables $\mathbf{u}$ by

$$\mathbf{w} = \mathsf{R}^{-1}\mathbf{u}. \tag{2.3.7}$$

In the case of one spatial dimension (so there is only one choice for $\mathbf{n}$ up to a sign), we can multiply the equation of motion (2.3.1) (with $\mathbf{f} = \mathbf{0}$) by R^{-1} to obtain

$$\partial_t \mathbf{w} + \mathsf{\Lambda}\partial_x \mathbf{w} = \mathbf{0}, \tag{2.3.8}$$

where $\mathsf{\Lambda} = \mathsf{R}\mathsf{P}\mathsf{R}^{-1}$ is the diagonal matrix of eigenvalues; the evolution equations decouple for the eigenfields into m advection equations where the corresponding eigenvalue is the advection speed.[24]. If we can form a set of m linearly independent eigenfields w_a from the fundamental variables u_a that satisfy advection equations, then the system is necessarily strongly hyperbolic.

Using the above method, the hyperbolicity of the ADM evolution system (2.1.41) can be analysed (see Sec. 5.4 of Ref. [49] for details). It turns out that the system can be strongly hyperbolic so long as (i) the momentum constraints are satisfied identically and (ii) the lapse evolves via the Bona-Masso slicing condition (2.3.37) or the densitised lapse

$$\tilde{\alpha} = \alpha/\sqrt{\gamma} \tag{2.3.9}$$

is a specified function of space and time (but importantly not the lapse α itself). Although (ii) is not too onerous, (i) is a bit of a showstopper. Since we always expect constraint violations in numerical evolutions and sometimes even use initial data with small violations, we must inevitably conclude that the ADM evolution system is merely weakly hyperbolic and alternatives must be sought.

[24]Note that this is no longer true in the case of more than 1 spatial dimension as the unit vector $\mathbf{n}$ is no longer trivial.

2.3.2 The BSSNOK formulation

In order to obtain the BSSNOK formulation, we start with a conformal rescaling of the spatial metric[25]

$$\tilde{\gamma}_{ij} = \chi\gamma_{ij}, \qquad \tilde{\gamma}^{ij} = \frac{1}{\chi}\gamma^{ij}, \qquad \chi = (\det(\boldsymbol{\gamma}))^{-1/3} = \gamma^{-1/3}. \tag{2.3.10}$$

This choice of χ ensures the determinant of the conformal metric $\tilde{\gamma}_{ij}$ is unity[26]. This choice for the conformal factor means that when the metric elements blow up, for example, near a black-hole singularity, χ smoothly goes to zero and the conformal metric components $\tilde{\gamma}_{ij}$ remain regular. Note that, since χ factors out the metric determinant, the conformal metric $\tilde{\gamma}_{ij}$ is no longer a tensor but rather a *tensor density* of weight $-2/3$[27]. Similarly, any tensor that is multiplied by χ^p becomes a tensor density of weight $-2p/3$. Care must be taken when evaluating covariant or Lie derivatives of tensor densities as they differ from the formulae used for ordinary tensors. See Appendix A.3 of Ref. [50] for details.

Next, the extrinsic curvature is decomposed into its trace K and trace-free parts A_{ij} as in Eq. (2.2.2). The tracefree part is then conformally rescaled[28]. We have

$$\tilde{A}_{ij} = \chi\left(K_{ij} - \frac{1}{3}K\gamma_{ij}\right). \tag{2.3.13}$$

A particularly important step of the BSSN(OK) construction is the introduction of the *conformal connection functions* [83] (sometimes referred to as the "BSSN Γ") defined

[25]As stated in Sec. 2.1.2 here we assume Cartesian spatial coordinates. The BSSNOK formulation has also been extended to non-Cartesian coordinate systems. See Ref. [86] for details.

[26]This χ choice of the conformal factor is the main choice for BSSNOK codes in the community to the best of my knowledge at the present time. However, the original papers [82, 83] used ϕ where $\chi = \mathrm{e}^{-4\phi}$ and there are also other choices such as $W = \sqrt{\chi}$ [87]. Note that Ref. [88] uses the symbol χ for the W choice.

[27]A tensor density of weight W and rank (r,s) $\mathcal{T}^{a_1\cdots a_r}{}_{b_1\cdots b_s}$ transforms under a coordinate transformation $x^a \mapsto x'^a = x'^a(x^b)$ as

$$\mathcal{T}'^{a_1\cdots a_r}{}_{b_1\cdots b_s} = J^W \frac{\partial x'^{a_1}}{\partial x^{c_1}}\cdots\frac{\partial x'^{a_r}}{\partial x^{c_r}}\frac{\partial x^{d_1}}{\partial x'^{b_1}}\cdots\frac{\partial x^{d_s}}{\partial x'^{b_s}}\mathcal{T}'^{c_1\cdots c_r}{}_{d_1\cdots d_s}, \tag{2.3.11}$$

where J is the Jacobian determinant given by

$$J = \left|\det\left(\frac{\partial x^a}{\partial x'^b}\right)\right|. \tag{2.3.12}$$

A tensor is a tensor density of weight 0.

[28]Note that we take a different conformal rescaling to what we did for the York-Lichnerowicz decomposition (2.2.6).

by

$$\tilde{\Gamma}^i := \tilde{\gamma}^{jk}\tilde{\Gamma}^i_{jk} = -\partial_j\tilde{\gamma}^{ij}, \tag{2.3.14}$$

where $\tilde{\Gamma}^i_{jk}$ are the Christoffel symbols of the conformal metric $\tilde{\gamma}_{ij}$. The conformal connection functions $\tilde{\Gamma}^i$ are promoted to full evolution variables; this allows the conformal Ricci tensor $\tilde{R}_{ij}$ (2.3.19a) to be expressed in a manifestly elliptic form (the only highest derivative term features just the Laplace operator $\tilde{\gamma}^{kl}\partial_k\partial_l$) and is crucial to removing terms that spoil strong hyperbolicity.

Finally, the evolution equations are modified using the constraint equations in order to improve the hyperbolicity properties. In the evolution equation for K, the Hamiltonian constraint is used to remove the Ricci scalar and in the evolution equation for $\tilde{\Gamma}^i$, the momentum constraint is used to remove derivatives of A^{ij}. The latter change has been found empirically to be necessary for numerical stability [83].

The [non-gauge] evolution variables for the BSSNOK formulation are

$$\left\{\chi, \tilde{\gamma}_{ij}, K, \tilde{A}_{ij}, \tilde{\Gamma}^i\right\}, \tag{2.3.15}$$

and the system of equations is

The BSSNOK evolution system

$$\partial_t\chi = \beta^k\partial_k\chi + \frac{2}{3}\chi(\alpha K - \partial_k\beta^k), \tag{2.3.16a}$$

$$\partial_t\tilde{\gamma}_{ij} = \beta^k\partial_k\tilde{\gamma}_{ij} + \tilde{\gamma}_{ki}\partial_j\beta^k + \tilde{\gamma}_{kj}\partial_i\beta^k - 2\alpha\tilde{A}_{ij} - \frac{2}{3}\tilde{\gamma}_{ij}\partial_k\beta^k, \tag{2.3.16b}$$

$$\partial_t K = \beta^k\partial_k K + \alpha\left(\tilde{A}^{kl}\tilde{A}_{kl} + \frac{1}{3}K^2 + 4\pi[S+\rho]\right) - \chi\tilde{\gamma}^{kl}D_kD_l\alpha, \tag{2.3.16c}$$

$$\begin{aligned}\partial_t\tilde{A}_{ij} = {}& \beta^k\partial_k\tilde{A}_{ij} + \chi\left[-D_iD_j\alpha + \alpha(R_{ij} - 8\pi S_{ij})\right]^{\mathrm{TF}} + \tilde{A}_{ij}\left[\alpha K - \frac{2}{3}\partial_k\beta^k\right] \\ & + 2\tilde{A}_{k(i}\partial_{j)}\beta^k - 2\alpha\tilde{\gamma}^{kl}\tilde{A}_{ik}\tilde{A}_{lj},\end{aligned} \tag{2.3.16d}$$

$$\begin{aligned}\partial_t\tilde{\Gamma}^i = {}& \beta^k\partial_k\tilde{\Gamma}^i + \frac{2}{3}\tilde{\Gamma}^i\partial_k\beta^k - \tilde{\Gamma}^k\partial_k\beta^i + \tilde{\gamma}^{kl}\partial_k\partial_l\beta^i + \frac{1}{3}\gamma^{ik}\partial_k\partial_l\beta^l \\ & - \tilde{A}^{ik}\left[3\alpha\frac{\partial_k\chi}{\chi} + 2\partial_k\alpha\right] + 2\alpha\tilde{\Gamma}^i_{kl}\tilde{A}^{kl} - \frac{4}{3}\alpha\tilde{\gamma}^{ik}\partial_k K - 16\pi\alpha\tilde{\gamma}^{ik}j_k,\end{aligned} \tag{2.3.16e}$$

where TF denotes the trace-free part of the expression is taken. The second covariant

derivative of the lapse is given by

$$D_i D_j \alpha = \partial_i \partial_l \alpha - \tilde{\Gamma}^k_{ij} \partial_k \alpha + \frac{1}{2\chi} \left(2\partial_{(i} \alpha \partial_{j)} \chi - \tilde{\gamma}_{ij} \tilde{\gamma}^{kl} \partial_k \alpha \partial_l \chi \right), \tag{2.3.17}$$

and the spatial Ricci tensor is given by

$$R_{ij} = \tilde{R}_{ij} + \frac{R^\chi_{ij}}{\chi}, \tag{2.3.18}$$

with

$$\tilde{R}_{ij} = -\frac{1}{2} \tilde{\gamma}^{kl} \partial_k \partial_l \tilde{\gamma}_{ij} + \tilde{\gamma}_{k(i} \partial_{j)} \tilde{\Gamma}^k + \frac{1}{2} \tilde{\Gamma}^k \partial_k \tilde{\gamma}_{ij} + \tilde{\gamma}^{lm} \left(2\tilde{\Gamma}^k_{l(i} \tilde{\Gamma}_{j)km} + \tilde{\Gamma}^k_{im} \tilde{\Gamma}_{klj} \right), \tag{2.3.19a}$$

$$R^\chi_{ij} = \frac{1}{2} \left[\tilde{D}_i \tilde{D}_j \chi + \tilde{\gamma}_{ij} \tilde{\gamma}^{kl} \tilde{D}_k \tilde{D}_l \chi \right] - \frac{1}{4\chi} \left[\partial_i \chi \partial_j \chi + 3\tilde{\gamma}_{ij} \tilde{\gamma}^{kl} \partial_k \chi \partial_l \chi \right]. \tag{2.3.19b}$$

The promotion of auxiliary variables to full evolution variables means we also have some algebraic constraints:

$$\tilde{\gamma} = 1, \tag{2.3.20}$$

$$\tilde{\gamma}^{kl} \tilde{A}_{kl} = 0, \tag{2.3.21}$$

$$\mathcal{G}^i \equiv \tilde{\Gamma}^i - \tilde{\gamma}^{kl} \tilde{\Gamma}^i_{kl} = 0. \tag{2.3.22}$$

For Eq. (2.3.20), we typically do not enforce this constraint and, in most applications, it turns out to be well satisfied throughout the evolution (assuming it is satisfied by the initial data). On the other hand, the manual enforcement of Eq. (2.3.21)[29] has been found to be necessary for numerical stability[30]. The final algebraic constraint (2.3.22) is enforced in different ways. One approach is to substitute on the right-hand side of Eq. (2.3.16e), all undifferentiated $\tilde{\Gamma}^i$ with their definition in terms of the conformal metric $\tilde{\gamma}_{ij}$ and its derivatives [90]. An alternative is to add a multiple of $\mathcal{G}^i$ to the right-hand side of Eq. (2.3.16e) to act as a damping term [91].

The hyperbolicity of the BSSNOK system was investigated by Sarbach, Calabrese, Pullin, and Tiglio [92] and Gundlach and Martin-Garcia [93] where they show that the formulation is indeed strongly hyperbolic and thus well-posed. It should be pointed out that other strongly hyperbolic formulations of the Einstein equation exist [79]. Nevertheless, for reasons that are still not well-understood, the BSSNOK formulation

[29] We enforce this before every evaluation of the RHS.

[30] For the CCZ4 formulation, which we will discuss in the next section, some groups have experimented with enforcing this constraint using a damping term, for example Ref. [89].

has proved more numerically robust than most others and, at least until the introduction of conformal variants of the Z4 formulation, it dominated 3D numerical relativity codes based on the 3+1 formalism.

2.3.3 The CCZ4 formulation

We now move on to discussing the covariant and conformal Z4 formulation (CCZ4) [84, 85] that will be used for the majority of the simulations presented in this thesis. The CCZ4 formulation modifies the Z4 formulation discussed in Sec. 2.1.8 in two fundamental ways. The first is by adding terms to the equations that damp constraint violations and the second is by performing a conformal decomposition à la BSSNOK.

Let us illustrate the idea behind constraint damping with a very simple example. Consider the [trivial] system

$$\partial_t \lambda = 0, \tag{2.3.23}$$

subject to the "constraint" $\lambda = 0$ and suppose at some time $\lambda = k \neq 0$ (in the more general case, such violations can be caused by numerical truncation error, regridding when using adaptive mesh refinement or just exist in the initial data). Evolving with Eq. (2.3.23), the violation remains throughout the evolution. We can add a constraint damping term to the RHS of Eq. (2.3.23) viz.

$$\partial_t \lambda = -\kappa \lambda, \qquad \kappa > 0. \tag{2.3.24}$$

Evolving with Eq. (2.3.24), we now find that the violations are damped away and the "constraint hypersurface" in solution space (the point $\lambda = 0$ in this simple example) is an *attractor*. This is shown in Fig. 2.4.

In this spirit, Gundlach, Martin-Garcia, Calabrese, and Hinder [61] modified the Z4 Einstein equation (2.1.43a) to include constraint damping terms[31]:

$${}^{(4)}R_{\mu\nu} + 2\nabla_{(\mu}Z_{\nu)} - \kappa_1(2n_{(\mu}Z_{\nu)} - (1+\kappa_2)g_{\mu\nu}n^\alpha Z_\alpha) = 8\pi\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right). \tag{2.3.25}$$

Here, in addition to the colour blue for the Z4-specific terms, we also use the colour red to indicate these covariant damping terms. The forced wave equation for $\boldsymbol{Z}$ (2.1.44) becomes

$$\Box Z_\mu + R_{\mu\alpha}Z^\alpha - \kappa_1\nabla^\alpha\left(2n_{(\alpha}Z_{\mu)} + \kappa_2 g_{\mu\alpha}n^\beta Z_\beta\right) = 0. \tag{2.3.26}$$

[31] Note that κ_1 has dimensions $[\kappa_1] = L^{-1}$ whereas κ_2 is dimensionless.

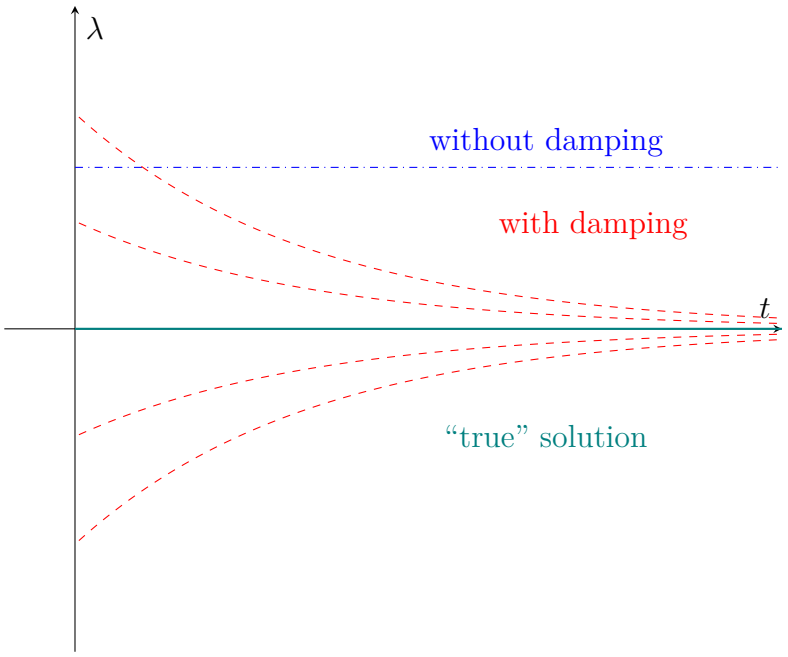


Fig. 2.4 Simple illustration of constraint damping. The constraint or "true" solution is $\lambda = 0$ [teal, solid line]. The undamped solution evolved with Eq. (2.3.23) has constant violations throughout [blue, dot-dashed] and the damped solutions evolved with Eq. (2.3.24) tend towards the true solution [red, dashed].

Linearising about Minkowski space, a mode analysis of Eq. (2.3.26) shows [61] that violations of the constraint $Z_\mu = 0$ are damped so long as

$$\kappa_1 > 0, \qquad \kappa_2 > -1. \tag{2.3.27}$$

Unlike the ADM formulation, the Z4 formulation (2.1.43) *is* strongly hyperbolic (even without the damping terms) [61] so long as the densitised lapse $\tilde{\alpha}$ (2.3.9) is specified as a function of space and time (as opposed to the physical lapse α) or a Bona-Masso slicing condition is used, but the momentum constraint is no longer required. Since the damping terms do not affect the highest order derivatives, they do not therefore affect the principal symbol and therefore the hyperbolicity.

There are several obvious advantages to the addition of constraint damping. Since numerical approximations/mesh refinement (see Sec. 3.1.4) always introduce violations of the constraints, we can mitigate these with damping. It may also help ameliorate the effect of constraint-violating boundary conditions[32]. Furthermore, we can be more confident that the numerical solutions we obtain are closer to the constraint hypersurface

[32]However, Bernuzzi and Hilditch [94] actually found the opposite with their Z4c formulation.

than they would otherwise be and are thus, more physically relevant. The use of constraint damping has even enabled the exploration of new physics [95]. Nevertheless, we should still remain cautious when using constraint damping. Although solutions should move closer to the constraint hypersurface as the solution progresses, it is not possible to control where on the constraint hypersurface they tend toward. It is of course possible that the constraint-damped solution, whilst closer to a true solution of the Einstein equation (1.0.1), may be further away from the physical configuration that was intended to be simulated. Therefore, one should not try to evolve initial data with very large constraint violations and just hope for the best.

We now perform a conformal decomposition of the metric and extrinsic curvature as in Eqs. (2.3.10) and (2.3.13) for BSSNOK. However rather than choosing somewhat arbitrarily to promote $\tilde{\Gamma}^i$ to an evolution variable, we instead combine it with Θ^i and define the new evolution variable

$$\hat{\Gamma}^i = \tilde{\Gamma}^i + 2\tilde{\gamma}^{ik}\Theta_k = \tilde{\Gamma}^i + \frac{2\Theta^i}{\chi}. \tag{2.3.28}$$

At each evaluation of the RHS of the CCZ4 system, the quantity Θ^i/χ is reconstructed using the above equation, where $\tilde{\Gamma}^i$ is calculated from the conformal metric $\tilde{\gamma}_{ij}$ and its derivatives.

There is then a choice as to whether the evolution equations should be modified using the Hamiltonian (2.1.36) and momentum (2.1.37) constraints to bring them closer to the BSSNOK form. In the CCZ4 formulation [84, 85], these modifications are not made (hence the "covariant" in the name) but in the similar Z4c formulation [94, 96], the equations are modified accordingly. There seems to be little difference between the performance of the two formulations [85]. One observation from Ref. [85] is that Z4c exhibited a smaller deviation in the final BH mass compared to CCZ4 for simulations of vacuum BHs and collapsing neutron stars. However, Bernuzzi and Hilditch [94] mention that Z4c seems to be particularly sensitive to constraint violating boundary conditions (hence the development of constraint-preserving BCs by Ruiz, Takahashi, Alcubierre, and Nunez [97]) which has not been reported with CCZ4. From now on, we will only consider CCZ4.

The CCZ4 evolution variables are

$$\{\chi, \tilde{\gamma}_{ij}, K, \tilde{A}_{ij}, \Theta, \hat{\Gamma}^i\} \tag{2.3.29}$$

and the system of equations is

The CCZ4 system

$$\partial_t \chi = \beta^k \partial_k \chi + \frac{2}{3}\chi(\alpha K - \partial_k \beta^k), \tag{2.3.30a}$$

$$\partial_t \tilde{\gamma}_{ij} = \beta^k \partial_k \tilde{\gamma}_{ij} + \tilde{\gamma}_{ki}\partial_j \beta^k + \tilde{\gamma}_{kj}\partial_i \beta^k - 2\alpha \tilde{A}_{ij} - \frac{2}{3}\tilde{\gamma}_{ij}\partial_k \beta^k, \tag{2.3.30b}$$

$$\begin{aligned}\partial_t K = {} & \beta^k \partial_k K + \alpha\left(R^Z + K(K - 2\Theta)\right) - 3\alpha\kappa_1(1+\kappa_2)\Theta - \gamma^{kl} D_k D_l \alpha \\ & + 4\pi\alpha(S - 3\rho),\end{aligned} \tag{2.3.30c}$$

$$\begin{aligned}\partial_t \tilde{A}_{ij} = {} & \beta^k \partial_k \tilde{A}_{ij} + \chi\left[-D_i D_j \alpha + \alpha(R^Z_{ij} - 8\pi S_{ij})\right]^{\mathrm{TF}} \\ & + \tilde{A}_{ij}\left[\alpha(K - 2\Theta) - \frac{2}{3}\partial_k \beta^k\right] + 2\tilde{A}_{k(i}\partial_{j)}\beta^k - 2\alpha\tilde{\gamma}^{kl}\tilde{A}_{ik}\tilde{A}_{lj},\end{aligned} \tag{2.3.30d}$$

$$\begin{aligned}\partial_t \Theta = {} & \beta^k \partial_k \Theta + \frac{1}{2}\alpha\left(R^Z - \tilde{A}_{kl}\tilde{A}^{kl} + \frac{2}{3}K^2 - 2\Theta K\right) - \alpha\kappa_1\Theta(2+\kappa_2) \\ & - \Theta^k \partial_k \alpha - 8\pi\alpha\rho,\end{aligned} \tag{2.3.30e}$$

$$\begin{aligned}\partial_t \hat{\Gamma}^i = {} & \beta^k \partial_k \hat{\Gamma}^i + \frac{2}{3}\left[\partial_k \beta^k\left(\tilde{\Gamma}^i + 2\kappa_3\frac{\Theta^i}{\chi}\right) - 2\alpha K\frac{\Theta^i}{\chi}\right] - 2\alpha\kappa_1\frac{\Theta^i}{\chi} \\ & + 2\tilde{\gamma}^{ik}(\alpha\partial_k\Theta - \Theta\partial_k\alpha) - 2\tilde{A}^{ik}\partial_k\alpha + 2\alpha\tilde{\Gamma}^i_{kl}\tilde{A}^{kl} \\ & - \alpha\left[\frac{4}{3}\tilde{\gamma}^{ik}\partial_k K + 3\tilde{A}^{ik}\frac{\partial_k\chi}{\chi}\right] - \left(\tilde{\Gamma}^k + 2\kappa_3\frac{\Theta^k}{\chi}\right)\partial_k\beta^i \\ & + \tilde{\gamma}^{kl}\partial_k\partial_l\beta^i + \frac{1}{3}\tilde{\gamma}^{ik}\partial_l\partial_k\beta^l - 16\pi\alpha\tilde{\gamma}^{ik}j_k.\end{aligned} \tag{2.3.30f}$$

Here we have introduced the *modified Ricci tensor* R^Z_{ij} which is the only form in which the spatial Ricci tensor appears. It is defined by

$$R^Z_{ij} := R_{ij} + 2D_{(i}\Theta_{j)} = \hat{R}_{ij} + \frac{1}{\chi}\left(R^\chi_{ij} + R^\Theta_{ij}\right), \tag{2.3.31}$$

with the parts given by

$$\hat{R}_{ij} = -\frac{1}{2}\tilde{\gamma}^{kl}\partial_k\partial_l\tilde{\gamma}_{ij} + \tilde{\gamma}_{k(i}\partial_{j)}\hat{\Gamma}^k + \hat{\Gamma}^k\partial_k\tilde{\gamma}_{ij} + \tilde{\gamma}^{lm}\left(\tilde{\Gamma}^k_{li}\tilde{\Gamma}_{jkm} + \tilde{\Gamma}^k_{lj}\tilde{\Gamma}_{ikm} + \tilde{\Gamma}^k_{im}\tilde{\Gamma}_{klj}\right), \tag{2.3.32a}$$

$$R^\chi_{ij} = \frac{1}{2}\left[\tilde{D}_i\tilde{D}_j\chi + \tilde{\gamma}_{ij}\tilde{\gamma}^{kl}\tilde{D}_k\tilde{D}_l\chi\right] - \frac{1}{4\chi}\left[\partial_i\chi\partial_j\chi + 3\tilde{\gamma}_{ij}\tilde{\gamma}^{kl}\partial_k\chi\partial_l\chi\right], \tag{2.3.32b}$$

$$R^\Theta_{ij} = \frac{\Theta^k}{\chi}\left(\tilde{\gamma}_{ik}\partial_j\chi + \tilde{\gamma}_{jk}\partial_i\chi - \tilde{\gamma}_{ij}\partial_k\chi\right), \tag{2.3.32c}$$

This combination means that terms which would spoil strong hyperbolicity cancel out [95]. The second covariant derivative of the lapse $D_iD_j\alpha$ is calculated as in Eq. (2.3.17).

The extra parameter κ_3 (and terms coloured in teal) was introduced by Alic, Bona-Casas, Bona, Rezzolla, and Palenzuela [84] as they were unable to obtain stable BH evolutions without $\kappa_3 \neq 1$. This parameter does not appear in the damped Z4 Einstein equation (2.3.25) and in order to maintain covariance, we must choose $\kappa_3 = 1$. However, in Ref. [85], it is shown that one can retain spatial covariance when evolving BHs by setting $\kappa_3 = 1$ and replacing

$$\alpha\kappa_1 \to \kappa_1. \tag{2.3.33}$$

This is because, for the gauge choice we typically make (and shall discuss next), the lapse becomes very small inside the BH which has the unintended consequence of turning off the constraint damping here. Note that this choice still breaks temporal covariance.

Whilst κ_2 and κ_3 are dimensionless, κ_1 is a decay rate (cf. Eq. (2.3.25)), and thus has units of inverse time. In practice, since we take code units where $M \sim 1$, the default values for these damping parameters is

$$\alpha\kappa_1 = 0.1,\ \kappa_2 = 0 \text{ and } \kappa_3 = 1. \tag{2.3.34}$$

Now, in order to close the system, we need to specify what conditions we will use for the gauge variables.

2.3.4 The moving puncture gauge

As mentioned in Sec. 2.1.3, the choice of gauge has a large influence on the stability and accuracy of a simulation. Arguably, one of the main advantages of the 3+1 formalism over, for example, generalized harmonic coordinates is the intuition it provides over how to control the coordinates through the choice of lapse α and shift β^i (see Fig. 2.2).

We have already met a "bad" gauge condition in the form of geodesic slicing ($\alpha = 1$) in Sec. 2.1.3. Consider the evolution of the volume element $\sqrt{\gamma}$ for fixed coordinate x^i (i.e. along $\boldsymbol{t} = \partial_t$). Using Eq. (2.1.38), we find that this satisfies

$$\partial_t \log\sqrt{\gamma} = -\alpha K + D_k\beta^k. \tag{2.3.35}$$

Now, it follows from Eq. (2.1.40) with $\alpha = 1$ and $\beta^i = 0$ (so the coordinates move along geodesics) that

$$\partial_t K = K_{ij}K^{ij} + 4\pi(\rho + S). \tag{2.3.36}$$

Assuming K_{ij} is non-vanishing, the first term is strictly-positive and the second term is nonnegative if the strong energy condition[33] holds. It follows that K grows indefinitely. Looking back at Eq. (2.3.35), we can see that this implies the volume elements of the coordinates collapse to zero.

One solution to this problem is to enforce $K = 0$ (maximal slicing) [63]. This is achieved by choosing initial data with $K = 0$ and, at each timestep, setting $\partial_t K = 0$ in Eq. (2.3.16c) (or Eq. (2.3.30c) for CCZ4) and solving the resulting elliptic equation for α. A particularly useful feature of maximal slicing is its *singularity avoidance*; it does not allow the hypersurfaces to come arbitrarily close to a physical singularity which is essential if one wishes to simulate black holes. Unfortunately, since maximal slicing involves solving an elliptic equation at each timestep, it is computationally very expensive even with fast elliptic solvers.

This motivated intense efforts in the early 1990s to seek alternatives choices that perform similarly. These efforts culminated in the *Bona-Masso* family of slicing conditions [98] where the lapse satisfies

$$(\partial_t - \mathcal{L}_{\boldsymbol{\beta}})\alpha = -\alpha f(\alpha) K, \tag{2.3.37}$$

where f is an arbitrary positive function. In the case $f(\alpha) = 2/\alpha$ and assuming zero shift, by substituting Eq. (2.3.35) into the RHS of Eq. (2.3.37), we find that $\partial_t \alpha = 2\partial_t \log \sqrt{\gamma}$ which can be integrated up directly to get $\alpha = g(\mathbf{x}) + \log \gamma$ for some arbitrary g. This explains why this choice for f is known as *1+log slicing*. It can be shown that this slicing condition has strong singularity avoidance properties [99] and has proved extremely robust for a wide variety of spacetimes so has become ubiquitous in 3+1 numerical relativity codes. Since 1+log slicing has proved so successful in the BSSNOK formulation, in order to match as closely as possible to it in the CCZ4 formulation, we make the replacement $K \to K - 2\Theta$. This arises because the Hamiltonian constraint (2.1.36) is used to replace the Ricci scalar to obtain the evolution equation for K (2.3.16c). Thus, 1+log slicing condition in the CCZ4 formulation is

CCZ4 1+log slicing

$$\partial_t \alpha = \beta^k \partial_k \alpha - 2\alpha(K - 2\Theta). \tag{2.3.38}$$

In practice, we use a generalised version of the above condition which takes the form

[33] The strong energy condition says that, for any timeline or null vector $\boldsymbol{u}$, we must have $(T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T) u^a u^b \geq 0$.

$$\partial_t \alpha = a_1 \beta^k \partial_k \alpha - a_2 \alpha^{a_3} (K - 2\Theta), \tag{2.3.39}$$

and reduces to Eq. (2.3.38) in the case of our default parameter values,

$$a_1 = 1, \; a_2 = 2 \text{ and } a_3 = 1. \tag{2.3.40}$$

Now that we have nailed down the slicing condition, one might hope that the naive shift condition $\beta^i = 0$ works for our needs [although perhaps the reader is sufficiently wary after seeing how badly the naive geodesic slicing condition failed that they won't fall into this trap again]. As expected, 1+log slicing with zero shift is not sufficient to stably and accurately evolve black holes for long times. This is because as the evolution proceeds, the size of the horizon increases rapidly in coordinate space until it encompasses the entire computational domain. Physically, this can be thought of as the Eulerian observers continuously falling in. Furthermore, since observers closer to the black hole move faster, the physical distance between coordinates increases which results in rapid growth of radial metric components and eventually leads codes to fail. This effect is referred to as *slice stretching*.

In order to prevent the coordinate lines falling into the black hole, looking at Fig. 2.2, it is clear that we need to pick an outward-pointing shift vector. Today, the most popular way to do this is to use a *Gamma driver* condition. We will use such a condition in the form[34]

CCZ4 Gamma driver shift condition

$$\partial_t \beta^i = b_1 \beta^k \partial_k \beta^i + b_2 B^i, \tag{2.3.41a}$$

$$\partial_t B^i = b_1 (\beta^k \partial_k B^i - \beta^k \partial_k \hat{\Gamma}^i) + \partial_t \hat{\Gamma}^i - \eta B^i, \tag{2.3.41b}$$

where b_1, b_2 and η are specifiable parameters. The discovery of this gauge condition led to the breakthrough simulations of binary black-hole systems through inspiral and merger without the need for excision[35] by Campanelli, Lousto, Marronetti, and Zlochower [39] and Baker, Centrella, Choi, Koppitz, and Meter [40][36].

There is still the question of what values for the parameters to choose in (2.3.41). For the advection parameter $b_1 = 0$ turns off the advection terms ("non-shifting shift")

[34] For the BSSNOK version, make the replacement $\hat{\Gamma}^i \to \tilde{\Gamma}^i$.

[35] Note that these were not the first simulations of a binary black-hole system through inspiral and merger as this accolade goes to Pretorius [38] who used generalized harmonic coordinates and excision.

[36] Baker, Centrella, Choi, Koppitz, and Meter [40] actually use a slightly different condition compared with (2.3.41) and this is discussed in Sec. II of [100].

and $b_1 = 1$ includes them ("shifting shift"). A linear mode analysis that compared different variants of the Gamma driver shift condition was conducted by van Meter, Baker, Koppitz, and Choi [100] where it was found that $b_1 = 1$ was the better option. However, in practice, little difference is found between the different variants and we shall typically take $b_1 = 1$. For b_2, the conventional choice is $b_2 = 3/4$. This seems to stems from Ref. [101] as this choice makes the longitudinal speed of the shift be that of the speed of light (when $\alpha = 1$). For the Z4 formulation, it is mentioned in Ref. [84] that $b_2 = 3/4$ can lead to weak hyperbolicity when $\alpha \sim 1$ and hence Ref. [94] suggests $b_2 = 1$. However, $b_2 = 3/4$ seems to work fine in practice for CCZ4 and this is what we shall use. Looking at the form Eq. (2.3.41b), we can see that the final term is acting as a damping term (cf. Eq. (2.3.24)) so η controls the decay timescale. Using dimensional analysis, we find that $[\eta] = L^{-1}$. Typically, for a spacetime of ADM mass M, we choose $\eta \sim 1/M$ but we note that for some scenarios (e.g. BH binaries with small mass ratios), it can help to allow η to vary in space and time [102]. After the gauge has settled, the coordinate size of the black hole horizon can be seen to increase as η increases (see Fig. 4 in [103]). Empirically, for comparable mass black-hole binaries, simulations are found to be reasonably robust against variations in η within an order of magnitude. However, if it is chosen too small $\eta \ll 1/M$, the simulation can become unstable and, on the other hand, if it is made too large $\eta \gg 1/M$, the aforementioned slice stretching effects can come into play.

The position of the punctures can be tracked by integrating their equation of motion (cf. Fig. 2.1)

$$\frac{\mathrm{d}x^i_p}{\mathrm{d}t} = -\beta^i(\mathbf{x}_p). \tag{2.3.42}$$

With the Gamma driver shift condition, the punctures become attractors. It should be stressed that the trajectories one obtains are gauge dependent.

The combination of 1+log slicing (2.3.38) and the Gamma driver shift condition (2.3.41) is known as the *moving puncture gauge* [39, 40] as it allows the BH punctures to move around the computational domain without the need for excision or special care. Although for Brill-Lindquist data (2.2.19) and Bowen-York puncture data (2.2.28), each puncture initially corresponds to its own asymptotically flat end, after a short time of evolution with the moving puncture gauge, they instead correspond to infinitely long "cylinders" at area-radius $r_0 \sim 1.31M < 2M$ [104]. This is known as the *trumpet solution*. A seasoned numerical analyst might worry that evaluating numerical derivatives across the puncture (which is still a coordinate singularity even if it does not correspond to

the physical singularity) might lead to artefacts that cause uncontrollable numerical errors and contaminate the rest of the computational domain destroying any accuracy. Remarkably, this does not turn out to be the case and the numerics are surprisingly regular [105]. This might be attributed to the presence of a horizon which shields the rest of computational domain from the crimes within.

2.4 Gravitational wave extraction

Now that we know how to construct initial data and evolve it, it is time to consider how we extract gravitational waves from numerical simulations. In this section, we will briefly review the plane-wave solution to the linearised Einstein equation that one typically meets in a first course in GR. Next, we look at the Weyl tensor and its electromagnetic decomposition. We then discuss the Newman-Penrose formalism and explain how the Weyl scalar Ψ_4 can be related back to the previously found plane-wave solution. Finally we show how the energy, linear and angular momentum radiated in gravitational waves can be calculated from Ψ_4.

It should be noted that using the Newman-Penrose formalism is not the only way to extract gravitational waves in 3+1 simulations. However, it is arguably one of the more straightforward approaches and has thus become the most popular method in the community. For a review of other methods and a comparison between them, see Ref. [106].

2.4.1 Plane wave solutions to the linearised Einstein equation

For this subsection, we follow the canonical treatment in Ref. [107]. In the weak-field limit of GR, we can consider the spacetime as a perturbation of Minkowski space, so there exist coordinates x^μ such that the metric can be written as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{2.4.1}$$

where $(\eta_{\mu\nu}) = \mathrm{diag}(-1, 1, 1, 1)$ is the Minkowski metric and $|h_{\mu\nu}| = \mathcal{O}(\epsilon) \ll 1$. Here ϵ is merely an expansion parameter that we will use to keep track of orders. Note that we have implicitly assumed that the coordinates x^μ are approximately Cartesian. We regard $h_{\mu\nu}$ as the components of a tensor field on the Minkowski background (so indices of tensors are raised and lowered with the Minkowski metric). To first order in ϵ, the

inverse metric is

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}, \qquad h^{\mu\nu} := \eta^{\mu\alpha}\eta^{\nu\beta}h_{\alpha\beta}. \tag{2.4.2}$$

At linear order in ϵ, the Riemann tensor is given by

$${}^{(4)}R_{\mu\nu\rho\sigma} = \frac{1}{2}\left(\partial_\nu\partial_\rho h_{\mu\sigma} + \partial_\sigma\partial_\mu h_{\nu\rho} - \partial_\nu\partial_\sigma h_{\mu\rho} - \partial_\mu\partial_\rho h_{\nu\sigma}\right). \tag{2.4.3}$$

It turns out that the equations are simplified if we consider the trace reversed metric perturbation

$$\bar{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h^{\alpha}{}_{\alpha}. \tag{2.4.4}$$

Then, at linear order in ϵ, the Einstein equation (1.0.1) becomes

$$\partial^\alpha\partial_{(\mu}\bar{h}_{\nu)\alpha} - \frac{1}{2}\left(\partial_\alpha\partial^\alpha\bar{h}_{\mu\nu} + \eta_{\mu\nu}\partial^\alpha\partial^\beta\bar{h}_{\alpha\beta}\right) = 8\pi T_{\mu\nu}. \tag{2.4.5}$$

We now make a gauge transformation in order to simplify this equation. Under the change of coordinates

$$x^\mu \mapsto \tilde{x}^\mu = x^\mu + \xi^\mu, \tag{2.4.6}$$

where $|\xi^\mu| = \mathcal{O}(\epsilon)$, the spacetime metric becomes

$$\tilde{g}_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} - 2\partial_{(\mu}\xi_{\nu)}, \tag{2.4.7}$$

at linear order in ϵ. We can read off the transformation in the metric perturbation as

$$h_{\mu\nu} \to h_{\mu\nu} - 2\partial_{(\mu}\xi_{\nu)}, \tag{2.4.8}$$

and its trace-reversal as

$$\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} - 2\partial_{(\mu}\xi_{\nu)} + \eta_{\mu\nu}\partial_\alpha\xi^\alpha. \tag{2.4.9}$$

Now, its divergence transforms as

$$\partial^\alpha\bar{h}_{\alpha\mu} \to \partial^\alpha\bar{h}_{\alpha\mu} - \partial^\alpha\partial_\alpha\xi_\mu, \tag{2.4.10}$$

so we can choose ξ^μ such that $\partial^\alpha\partial_\alpha\xi_\mu = \partial^\alpha\bar{h}_{\alpha\mu}$ (this equation is simply a sourced wave equation so there exists a solution). In the new gauge (dropping the tildes), we find

$$\partial^\alpha\bar{h}_{\alpha\mu} = 0, \tag{2.4.11}$$

which is known as Lorenz gauge[37] (in analogy with electromagnetism). Note that, looking at Eq. (2.4.5), we can see that the components of the energy-momentum tensor $|T_{\mu\nu}| = \mathcal{O}(\epsilon)$ so at linear order in ϵ, they do not change under the gauge transformation. In this new gauge, the linearised Einstein equation (2.4.5) becomes

$$\Box_\eta \bar{h}_{\mu\nu} = -16\pi T_{\mu\nu}, \tag{2.4.12}$$

where $\Box_\eta = \partial^\alpha \partial_\alpha$ is the usual flat-space wave operator.

In vacuum $T_{\mu\nu} = 0$, we find that Eq. (2.4.12) admits plane wave solutions of the form

$$\bar{h}_{\mu\nu} = \mathrm{Re}\left[A_{\mu\nu} \exp(\mathrm{i} l_\alpha x^\alpha)\right], \tag{2.4.13}$$

where $A_{\mu\nu}$ is the constant symmetric amplitude tensor and l^μ is the real wavevector which must be null in order for this solution to satisfy Eq. (2.4.12). From now on, we drop the "Re" and take it as understood. The Lorenz gauge condition (2.4.11) implies

$$l^\alpha A_{\alpha\mu} = 0, \tag{2.4.14}$$

which means the waves are transverse (i.e. the amplitudes are orthogonal to the wavevector and thus the direction of propagation), like electromagnetic waves. However, Eq. (2.4.11) does not completely fix the gauge as we can perform another transformation of the form (2.4.6) with ξ'^μ satisfying the wave equation $\Box_\eta \xi'^\mu = 0$ and Eq. (2.4.11) will still be satisfied. In particular, we can choose

$$\xi'^\mu = B^\mu \exp(\mathrm{i} l_\alpha x^\alpha), \tag{2.4.15}$$

where B^μ is an arbitrary constant vector. It can be shown that we can use a gauge transformation of this form to further impose

$$A_{0\mu} = 0, \qquad A^\alpha{}_\alpha = 0. \tag{2.4.16}$$

The conditions (2.4.14) and (2.4.16) are known as *transverse-traceless* (TT) gauge. Note that, in this gauge, $h_{\mu\nu} = \bar{h}_{\mu\nu}$. In the case of a plane wave in the z-direction

[37] This is commonly called Lorentz gauge.

($l^\mu = (1,0,0,1)$), we can write

$$(A_{\mu\nu}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A^+ & A^\times & 0 \\ 0 & A^\times & -A^+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.4.17}$$

The two constants A^+ and $A^\times$ correspond to the two independent polarizations of gravitational waves.

It follows from the equation of geodesic deviation (see, for example Ref. [106]) that the effect of a gravitational wave of the form (2.4.13) (in TT gauge) on the displacement ϵ^i of two nearby freely-falling test particles is governed by

$$\ddot{\epsilon}_i = \frac{1}{2}\ddot{h}^{\rm TT}_{ik}\epsilon^k, \tag{2.4.18}$$

which has solution

$$\epsilon_i(t) = \epsilon^k(0)\left[\delta_{ik} + \frac{1}{2}h^{\rm TT}_{ik}(t)\right], \tag{2.4.19}$$

or, in other words, the relative *strain* $\delta\epsilon/\epsilon$ between these two particles is proportional to the metric perturbation amplitude. This is why $h^{\rm TT}_{ij}$ is commonly referred to as the *gravitational wave strain*.

2.4.2 The Weyl tensor

Given a $\mathcal{D}$ dimensional spacetime with Riemann tensor ${}^{(\mathcal{D})}\boldsymbol{R}$ of the Levi-Civita connection ∇ with respect to the metric $g_{\mu\nu}$, the Weyl tensor is defined by [108]

$$C_{\mu\nu\rho\sigma} := {}^{(\mathcal{D})}R_{\mu\nu\rho\sigma} - \frac{2}{\mathcal{D}-2}\left(g_{\mu[\rho}\,{}^{(\mathcal{D})}R_{\sigma]\nu} - g_{\nu[\rho}\,{}^{(\mathcal{D})}R_{\sigma]\mu}\right) + \frac{2}{(\mathcal{D}-1)(\mathcal{D}-2)}g_{\mu[\rho}g_{\sigma]\nu}\,{}^{(\mathcal{D})}R. \tag{2.4.20}$$

It incorporates the degrees of freedom of the Riemann tensor that are not present in the Ricci tensor and thus shares the symmetries of the Riemann tensor. Furthermore, from the above definition, it can also be verified that the Weyl tensor is traceless in all pairs of indices:

$$C^\alpha{}_{\mu\alpha\nu} = 0. \tag{2.4.21}$$

It is well known that the Riemann tensor has $\mathcal{D}^2(\mathcal{D}^2-1)/12$ independent components and the Ricci tensor, which is symmetric, has $\mathcal{D}(\mathcal{D}+1)/2$ independent components (for

$\mathcal{D} > 2$). Therefore, the number of independent components in the Weyl tensor is

$$\frac{\mathcal{D}^2(\mathcal{D}^2-1)}{12} - \frac{\mathcal{D}(\mathcal{D}+1)}{2} = \frac{\mathcal{D}(\mathcal{D}+1)(\mathcal{D}+2)(\mathcal{D}-3)}{12}, \quad \mathcal{D} > 3 \tag{2.4.22}$$

Note that both in $\mathcal{D} = 2$ and 3, the Ricci and Riemann tensor have the same number of independent components (1 and 6 respectively), so the Weyl tensor vanishes identically. It follows that the first nontrivial dimension is $\mathcal{D} = 4$ which is conveniently the dimension most relevant to us and is what we shall now assume.

An important property of the Weyl tensor is that under a conformal transformation[38] of the form

$$g_{\mu\nu} \to \tilde{g}_{\mu\nu} = \Omega g_{\mu\nu}, \tag{2.4.23}$$

the Weyl tensor is invariant:

$$C^{\mu}{}_{\nu\rho\sigma} = \tilde{C}^{\mu}{}_{\nu\rho\sigma}. \tag{2.4.24}$$

In other words, the Weyl tensor captures the degrees of freedom within a conformal equivalence class of metrics. Note that Eq. (2.4.24) only holds with the indices in these positions as the respective metrics will need to be used to raise and lower indices and these differ by a conformal factor.

2.4.3 The electromagnetic decomposition of the Weyl tensor

Given an arbitrary timelike unit vector n^μ (we will later restrict to the case where n^μ is the unit normal to the foliation (2.1.3) so the ambiguous notation is intentional), we define the electric and magnetic parts of the Weyl tensor respectively by,

$$E_{\mu\nu} := n^\alpha n^\beta C_{\alpha\mu\beta\nu}, \tag{2.4.25a}$$

$$B_{\mu\nu} := n^\alpha n^\beta (*C)_{\alpha\mu\beta\nu}, \tag{2.4.25b}$$

where the dual Weyl tensor, $(*C)_{\mu\nu\rho\sigma}$ is defined by[39]

$$(*C)_{\mu\nu\rho\sigma} := \frac{1}{2}\epsilon^{\alpha\beta}{}_{\rho\sigma} C_{\mu\nu\alpha\beta}, \tag{2.4.26}$$

[38] Some theoretical physicists call such a transformation a *Weyl* transformation. It's not too difficult to see where this comes from.

[39] Note that this is just the Hodge dual if the Weyl tensor is considered to be a 2-form valued 2-form hence the similar notation.

and $\epsilon_{\mu\nu\rho\sigma}$ is the volume form. Note that the dual satisfies

$$(**C)_{\mu\nu\rho\sigma} = -C_{\mu\nu\rho\sigma}, \tag{2.4.27}$$

which is derived by using contractions of the volume form in terms of Kronecker deltas[40]. It follows from the symmetries of the Weyl tensor that the electric and magnetic parts are symmetric, trace-free and spatial in the sense that

$$n^\alpha E_{\alpha\mu} = 0 = n^\alpha B_{\alpha\mu}. \tag{2.4.29}$$

Here, the only non-trivial relation is the symmetry of the magnetic part. One would hope that the dual Weyl tensor shares the same symmetries as the Weyl tensor, in particular,

$$(*C)_{\mu\nu\rho\sigma} \stackrel{?}{=} (*C)_{\rho\sigma\mu\nu} \iff \epsilon^{\alpha\beta}{}_{\rho\sigma} C_{\mu\nu\alpha\beta} \stackrel{?}{=} \epsilon^{\alpha\beta}{}_{\mu\nu} C_{\alpha\beta\rho\sigma}, \tag{2.4.30}$$

where we have already applied the corresponding Weyl tensor symmetry in the final expression. Let $A_{\mu\nu\rho\sigma} = (*C)_{\mu\nu\rho\sigma} - (*C)_{\rho\sigma\mu\nu}$, and consider

$$2\epsilon^{\mu\nu\alpha\beta} A_{\alpha\beta\rho\sigma} = \epsilon^{\mu\nu\alpha\beta}\epsilon_{\gamma\delta\rho\sigma} C_{\alpha\beta}{}^{\gamma\delta} - \epsilon^{\mu\nu\alpha\beta}\epsilon_{\gamma\delta\alpha\beta} C^{\gamma\delta}{}_{\rho\sigma} = 0, \tag{2.4.31}$$

where the final equality follows after expanding products of the epsilon tensor in terms of Kronecker deltas (2.4.28) and simplifying[41]. This equality holds if and only if $A_{[\mu\nu]\rho\sigma} = 0$ whence (2.4.30) holds and then the symmetry of the magnetic part follows.

The electric and magnetic parts each have 16 independent components. However, they are purely spatial so, in practice, they have at most 9 independent components. Their symmetry then restricts this to at most 6 and finally the tracelessness means they only have 5 independent components. Assuming the electric and magnetic parts are independent of one another, this gives 10 total, which is precisely the number of independent components in the Weyl tensor. Therefore, it is not unreasonable to believe that it can be written in terms of the electric and magnetic parts. Unfortunately whilst

[40] In $\mathcal{D}$ dimensions, the contraction of two volume forms over p indices is given by

$$\epsilon^{\alpha_1\cdots\alpha_p\mu_1\cdots\mu_{\mathcal{D}-p}}\epsilon_{\alpha_1\cdots\alpha_p\nu_1\cdots\nu_{\mathcal{D}-p}} = \pm p!(n-p)!\delta^{[\mu_1}_{\beta_1}\cdots\delta^{\alpha_{\mathcal{D}-p}]}_{\beta_{\mathcal{D}-p}}, \tag{2.4.28}$$

where the $+$ is for Riemannian signature and the $-$ is for Lorentzian signature.

[41] The algebra is quite fiddly and the author resorted to a computer algebra system (as well as for many of the other calculations in this subsection) in the form of the *Mathematica* package, *xAct* [109].

the result (2.4.44) is quoted throughout the literature, its derivation is difficult to find so we explain it below in more detail than has been provided in the rest of this chapter.

Before we can directly tackle the Weyl tensor, it is helpful to consider something more simple in the form of self-dual 2-forms.

2.4.3.1 Self-dual 2-forms

Let $\boldsymbol{\omega}$ be a complex-valued 2-form that satisfies

$$\boldsymbol{\omega} = \mathrm{i} \star \boldsymbol{\omega} \iff \omega_{\mu\nu} = \frac{1}{2}\mathrm{i}\epsilon_{\mu\nu}{}^{\alpha\beta}\omega_{\alpha\beta}, \tag{2.4.32}$$

where $\star$ denotes the usual Hodge star[42] and define the associated 1-form $\boldsymbol{\eta}$ by

$$\eta_\mu := n^\alpha \omega_{\alpha\mu}. \tag{2.4.34}$$

Then, we can recover $\boldsymbol{\omega}$ from $\boldsymbol{\eta}$ using the identity tensor

$$\mathcal{I}_{\mu\nu\rho\sigma} = -\left(2g_{\mu[\rho}g_{\sigma]\nu} + \mathrm{i}\epsilon_{\mu\nu\rho\sigma}\right) \tag{2.4.35}$$

We have

$$\begin{aligned}
\mathcal{I}_{\mu\nu\alpha}{}^{\beta} n^\alpha \eta_\beta &= -\left(g_{\mu\alpha}\delta^\beta_\nu - g_{\nu\alpha}\delta^\beta_\mu + \mathrm{i}\epsilon_{\mu\nu\alpha}{}^{\beta}\right) n^\alpha n^\gamma \omega_{\gamma\beta} \\
&= -n_\mu n^\gamma \omega_{\gamma\nu} + n_\nu n^\gamma \omega_{\gamma\mu} + \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}\epsilon^{\gamma\beta\delta\varepsilon} n^\alpha n_\gamma \omega_{\delta\varepsilon} \\
&= -2n^\gamma n_{[\nu}\omega_{\mu]\gamma} - \frac{1}{2}(3!)\delta^{[\delta}_\mu \delta^\varepsilon_\nu \delta^{\gamma]}_\alpha n^\alpha n_\gamma \omega_{\delta\varepsilon} \\
&= -2n^\gamma n_{[\nu}\omega_{\mu]\gamma} - \frac{1}{2}\left(-2\omega_{[\mu\nu]} - 4n^\delta n_{[\nu}\omega_{\mu]\delta}\right) \\
&= \omega_{\mu\nu}.
\end{aligned} \tag{2.4.36}$$

Of course, we could change (2.4.32) to

$$\boldsymbol{\omega} = -\mathrm{i} \star \boldsymbol{\omega}, \tag{2.4.37}$$

[42] For a p-form $\boldsymbol{\xi}$ on a $\mathcal{D}$ dimensional manifold, the Hodge dual $\star\boldsymbol{\xi}$ is a $(\mathcal{D}-p)$-form defined by

$$(\star\xi)_{\alpha_1\ldots\alpha_{\mathcal{D}-p}} = \frac{1}{p!}\epsilon_{\alpha_1\ldots\alpha_{\mathcal{D}-p}\beta_1\ldots\beta_p}\xi^{\beta_1\ldots\beta_p}. \tag{2.4.33}$$

which can be recovered with the complex conjugate identity tensor

$$\bar{\mathcal{I}}_{\mu\nu\rho\sigma} = -2g_{\mu[\rho}g_{\sigma]\nu} + \mathrm{i}\epsilon_{\mu\nu\rho\sigma}. \tag{2.4.38}$$

2.4.3.2 Recovering the Weyl tensor from its electric and magnetic parts

Now, in order to apply the result in Eq. (2.4.36), we want to write the Weyl tensor in a form that satisfies something similar to (2.4.32). We do this by defining a new tensor

$$\mathcal{C}_{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma} + \mathrm{i}(*C)_{\mu\nu\rho\sigma}. \tag{2.4.39}$$

This satisfies a property similar to (2.4.32) since

$$\mathrm{i}(*\mathcal{C})_{\mu\nu\rho\sigma} = -(**C)_{\mu\nu\rho\sigma} + \mathrm{i}(*C)_{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma} + \mathrm{i}(*C)_{\mu\nu\rho\sigma} = \mathcal{C}_{\mu\nu\rho\sigma}, \tag{2.4.40}$$

where we have used Eq. (2.4.27) in the second equality. The object analogous to the "associated 1-form" (2.4.34) is

$$\bar{Q}_{\mu\nu} := n^\alpha n^\beta \mathcal{C}_{\alpha\mu\beta\nu} = E_{\mu\nu} + \mathrm{i}B_{\mu\nu}, \tag{2.4.41}$$

where we have written $\bar{Q}$ for consistency with the quantity Q which is often defined in the literature (e.g. Eq. (8.6.18) in Ref. [49]). Then, we apply the recovery identity (2.4.36) on both pairs of indices to obtain

$$\begin{aligned} C_{\mu\nu\rho\sigma} &= \mathrm{Re}\left(\mathcal{C}_{\mu\nu\rho\sigma}\right) \\ &= \mathrm{Re}\left(\mathcal{I}_{\mu\nu\alpha}{}^{\beta}\mathcal{I}_{\rho\sigma\gamma}{}^{\delta}n^\alpha n^\gamma \bar{Q}_{\beta\delta}\right) \\ &= \mathrm{Re}\left\{\left(g_{\mu\alpha}\delta^\beta_\nu - g_{\nu\alpha}\delta^\beta_\mu + \mathrm{i}\epsilon_{\mu\nu\alpha}{}^{\beta}\right)\left(g_{\rho\gamma}\delta^\delta_\sigma - g_{\sigma\gamma}\delta^\delta_\rho + \mathrm{i}\epsilon_{\rho\sigma\gamma}{}^{\delta}\right)n^\alpha n^\gamma (E_{\beta\delta} + \mathrm{i}B_{\beta\delta})\right\} \\ &= 2n_\mu n_{[\rho}E_{\sigma]\nu} - 2n_\nu n_{[\rho}E_{\sigma]\mu} - \epsilon_{\mu\nu}{}^{\beta}\epsilon_{\rho\sigma}{}^{\delta}E_{\beta\delta} - 2\epsilon_{\mu\nu}{}^{\beta}n_{[\rho}B_{\sigma]\beta} - 2\epsilon_{\rho\sigma}{}^{\delta}n_{[\mu}B_{\nu]\delta}, \end{aligned} \tag{2.4.42}$$

where we have chosen the convention $\epsilon_{\mu\nu\rho} = n^\alpha \epsilon_{\alpha\mu\nu\rho}$ for the 3-dimensional volume element and implicitly used the symmetry (2.4.30) in the second equality. Using Eq. (2.4.28), we find that

$$\epsilon_{\mu\nu}{}^{\beta}\epsilon_{\rho\sigma}{}^{\delta}E_{\beta\delta} = 2\gamma_{\nu[\rho}E_{\sigma]\mu} - 2\gamma_{\mu[\rho}E_{\sigma]\nu}, \tag{2.4.43}$$

where $\gamma_{\mu\nu}$ is given by Eq. (2.1.5) (except we still do not require that n^μ is necessarily the normal to the foliation). Substituting back into the final line of Eq. (2.4.42), we obtain

$$C_{\mu\nu\rho\sigma} = 2\left(l_{\mu[\rho}E_{\sigma]\nu} - l_{\nu[\rho}E_{\sigma]\mu} - n_{[\rho}B_{\sigma]\beta}\epsilon^{\beta}{}_{\mu\nu} - n_{[\mu}B_{\nu]\delta}\epsilon^{\delta}{}_{\rho\sigma}\right), \tag{2.4.44}$$

where

$$l_{\mu\nu} = \gamma_{\mu\nu} + n_\mu n_\nu = g_{\mu\nu} + 2n_\mu n_\nu. \tag{2.4.45}$$

This result allows us to calculate the Weyl tensor in terms of purely spatial quantities that we can compute in a 3+1 simulation.

2.4.3.3 3+1 expressions for the electric and magnetic parts

We now show how to derive expressions for the electric and magnetic parts in the 3+1 formalism. In order to do this, we have to assume that n_μ is the unit normal to the foliation (2.1.3). Since we mainly use the CCZ4 formulation, we will also include the terms that depend on the Z4 vector (or rather its projections).

Starting from the contracted Gauss equation (2.1.27), we use the definition of the Weyl tensor (2.4.20) to substitute for the Riemann tensor in terms of the Weyl tensor, the Ricci tensor and the Ricci scalar. Then we use the electromagnetic decomposition of the Weyl tensor (2.4.44) to substitute for the electric and magnetic parts. Rearranging for the electric part, we get

$$E_{\mu\nu} = R_{\mu\nu} - K_\mu{}^\alpha K_{\nu\alpha} + KK_{\mu\nu} + \frac{1}{3}{}^{(4)}R\gamma_{\mu\nu} - {}^{(4)}R^{\alpha\beta}\gamma_{\mu(\alpha}\gamma_{\nu)\beta}. \tag{2.4.46}$$

We then use the damped Z4 equation (2.3.25) to remove the spacetime Ricci tensor and scalar. Furthermore we replace all spacetime quantities (e.g. Z_μ via Eq. (2.1.45) and $T_{\mu\nu}$ via Eq. (2.1.35)) with their respective projections and apply Eqs. (2.1.19) and (2.1.22) as necessary to remove gradients of the normal. This yields

$$\begin{aligned} E_{ij} = {} & R_{ij} - K_i{}^m K_{jm} + KK_{ij} - \Theta\left(K_{ij} + \frac{1}{3}K\gamma_{ij}\right) - \frac{2}{3}\Lambda\gamma_{ij} - \frac{4}{3}G\pi(4\rho\gamma_{ij} + 3S_{ij} - S\gamma_{ij}) \\ & - \frac{2}{3}\kappa_1\Theta(2+\kappa_2)\gamma_{ij} + D_{(i}\Theta_{j)} + \frac{1}{3}\gamma_{ij}D_m\Theta^m - \frac{2}{3}\gamma_{ij}\left(\Theta^m D_m \log\alpha + \mathcal{L}_{\boldsymbol{n}}\Theta\right), \end{aligned} \tag{2.4.47}$$

where we have switched to spatial indices since all the quantities are purely spatial. Finally, we use the evolution equation for Θ (2.1.46d) to replace $\mathcal{L}_{\boldsymbol{n}}\Theta$. This gives an explicitly-trace-free form which we can write more succinctly as

$$E_{ij} = \left[R_{ij} - K_i{}^m K_{jm} + K_{ij}(K-\Theta) - 4G\pi S_{ij} + D_{(i}\Theta_{j)}\right]^{\mathrm{TF}}. \tag{2.4.48}$$

where the $^{\mathrm{TF}}$ indicates the tracefree part is taken. Unfortunately, the Ricci tensor does not appear in the usual modified Ricci tensor form (2.3.31) due to the missing factor of 2 in front of the $D_{(i}\Theta_{j)}$ term which means existing code to calculate the modified Ricci tensor needs to be modified in order to calculate Eq. (2.4.48). Note that this expression differs from the usual [non-Z4] 3+1 expression which is

$$E_{ij} = R_{ij} + KK_{ij} - K_{im}K^{m}{}_{j} - 4\pi\left[S_{ij} + \frac{1}{3}\gamma_{ij}(4\rho - S)\right]. \tag{2.4.49}$$

This expression relies on the Hamiltonian constraint (2.1.36) in order to be tracefree whereas the tracelessness appears manifestly in the Z4 expression.

For the magnetic part, we start with the Codacci equation (2.1.29) and play the same game as we did for the electric part, using Eq. (2.4.20) to substitute for the Riemann tensor and then applying the electromagnetic decomposition of the Weyl tensor (2.4.44). Rearranging for the magnetic part, we find

$$B_{\mu\nu} = \frac{1}{2}\epsilon_{\nu\mu\alpha}n^{\beta}\,{}^{(4)}R_{\beta}{}^{\alpha} + \epsilon_{\nu}{}^{\alpha\beta}D_{\alpha}K_{\mu\beta}, \tag{2.4.50}$$

Now we could use the damped Z4 equation (2.3.25) to substitute for the Ricci tensor, but instead we just use the symmetry of $B_{\mu\nu}$ and symmetrize both sides to obtain

$$B_{ij} = \epsilon_{mn(i}D^{m}K_{j)}{}^{n}, \tag{2.4.51}$$

where we have switched to spatial indices as all the quantities are purely spatial. Note that the symmetry of K_{ij} and the antisymmetry of ϵ_{ijk} ensures that this expression is tracefree.

2.4.4 The Newman-Penrose formalism and null tetrads

The Newman-Penrose formalism [110] starts by introducing a tetrad of *null* vector fields. Normally, we can only construct two linearly independent null vectors at a point but if we allow complex numbers, we can extend this to four. To construct these null vectors we can start with an orthonormal tetrad $\{e^{\mu}_{(a)} : a = 0\ldots3\}$ which satisfy $g_{\alpha\beta}e^{\alpha}_{(a)}e^{\beta}_{(b)} = \eta_{(a)(b)}$,

where $(\eta_{(a)(b)}) = \mathrm{diag}(-1,1,1,1)$. Then, we define our null tetrad as[43]

$$l^\mu := \frac{1}{\sqrt{2}}\left(e^\mu_{(0)} + e^\mu_{(1)}\right), \tag{2.4.52a}$$

$$k^\mu := \frac{1}{\sqrt{2}}\left(e^\mu_{(0)} - e^\mu_{(1)}\right), \tag{2.4.52b}$$

$$m^\mu := \frac{1}{\sqrt{2}}\left(e^\mu_{(2)} + \mathrm{i}e^\mu_{(3)}\right), \tag{2.4.52c}$$

$$\bar{m}^\mu := \frac{1}{\sqrt{2}}\left(e^\mu_{(2)} - \mathrm{i}e^\mu_{(3)}\right), \tag{2.4.52d}$$

Writing $l^\mu_{(0)} = l^\mu$, $l^\mu_{(1)} = k^\mu$, $l^\mu_{(2)} = m^\mu$ and $l^\mu_{(3)} = \bar{m}^\mu$, these satisfy

$$g_{\alpha\beta} l^\alpha_{(a)} l^\beta_{(b)} = \zeta_{(a)(b)}, \qquad (\zeta_{(a)(b)}) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.4.53}$$

We can now represent the 10 independent components of the Weyl tensor in terms of its projection onto the null tetrad (2.4.52). The five complex scalars Ψ_a, $a = 0,\ldots,4$ are defined as such and are known as the *Weyl scalars* (also commonly referred to as the *Newman-Penrose scalars*). In particular, the outgoing gravitational radiation is encapsulated in[44]

$$\Psi_4 := C_{\alpha\beta\gamma\delta} k^\alpha \bar{m}^\beta k^\gamma \bar{m}^\delta. \tag{2.4.54}$$

Using Eq. (2.4.44), it can be shown that

$$\Psi_4 = Q_{ij} \bar{m}^i \bar{m}^j, \tag{2.4.55}$$

where $Q_{ij} = E_{ij} - \mathrm{i}B_{ij}$ is the complex conjugate of Eq. (2.4.41); this is the formula we use in our numerical code to calculate Ψ_4.

One of the main conceptual issues with the Newman-Penrose formalism is that Ψ_4 depends on the choice of the null tetrad. For the Kerr spacetime, the "natural" null tetrad which has the required asymptotic behaviour and principal null directions is known as the *Kinnersley tetrad* [111]. However, it can be shown that under a small perturbation of the Kinnersley tetrad, Ψ_4 is invariant at first order [112]. Thus, the ambiguity in the choice

[43] Following Ref. [49], we use k^μ to denote the second null vector rather than the more conventional n^μ to avoid confusion with the unit normal to the foliation (2.1.3).

[44] Note that we follow the sign convention of Ref. [49] which differs to that of Refs. [50, 106]

of tetrad is a second-order effect and should not be too important. Of course in practice, we are not typically evolving Kerr but one would expect that far from the sources of gravitational radiation, the spacetime is close to Kerr. We therefore wish to construct what is known as a *quasi-Kinnersley tetrad* which tends to the Kinnersley tetrad as the spacetime tends to Kerr. We will avoid discussing the details of this ambiguity further, but the reader is referred to Ref. [106] for a brief overview particularly with respect to the intended numerical application.

In practice, during a numerical simulation, we follow the construction of a null tetrad using step (a) from Sec. V A in Ref. [113] which we now outline. We start by constructing an orthonormal basis. First, we pick $e^{\mu}_{(0)} = n^{\mu}$. Then, we define the purely spatial vectors

$$\tilde{e}^{i}_{\phi} = -y(\partial_x)^i + y(\partial_x)^i, \tag{2.4.56}$$

$$\tilde{e}^{i}_{r} = x(\partial_x)^i + y(\partial_y)^i + z(\partial_z)^i, \tag{2.4.57}$$

$$\tilde{e}^{i}_{\theta} = \sqrt{\gamma}\gamma^{im}\tilde{\epsilon}_{mnp}\tilde{e}^{n}_{\phi}\tilde{e}^{p}_{r}, \tag{2.4.58}$$

where $\tilde{\epsilon}_{ijk}$ is the totally-antisymmetric alternating symbol. This is just the flat-space tetrad in spherical coordinates. Note that these vectors are neither normalised nor orthogonal so we orthonormalise them using a Gram-Schmidt procedure:

$$\hat{e}^{i}_{\phi} = \frac{\tilde{e}^{i}_{\phi}}{\sqrt{\gamma_{kl}\tilde{e}^{k}_{\phi}\tilde{e}^{l}_{\phi}}}, \tag{2.4.59}$$

$$\hat{e}^{i}_{r} = \frac{\tilde{e}^{i}_{r} - \hat{e}^{i}_{\phi}(\gamma_{kl}\hat{e}^{k}_{\phi}\tilde{e}^{l}_{r})}{\sqrt{\gamma_{mn}[\tilde{e}^{m}_{r} - \hat{e}^{m}_{\phi}(\gamma_{kl}\hat{e}^{k}_{\phi}\tilde{e}^{l}_{r})][\tilde{e}^{n}_{r} - \hat{e}^{n}_{\phi}(\gamma_{kl}\hat{e}^{k}_{\phi}\tilde{e}^{l}_{r})]}}, \tag{2.4.60}$$

$$\hat{e}^{i}_{\theta} = \frac{\tilde{e}^{i}_{\theta} - \hat{e}^{i}_{\phi}(\gamma_{kl}\hat{e}^{k}_{\phi}\tilde{e}^{l}_{\theta}) - \hat{e}^{i}_{r}(\gamma_{kl}\hat{e}^{k}_{r}\tilde{e}^{l}_{\theta})}{\sqrt{\gamma_{mn}[\tilde{e}^{m}_{\theta} - \hat{e}^{m}_{\phi}(\gamma_{kl}\hat{e}^{k}_{\phi}\tilde{e}^{l}_{\theta}) - \hat{e}^{m}_{r}(\gamma_{kl}\hat{e}^{k}_{r}\tilde{e}^{l}_{\theta})][\tilde{e}^{n}_{\theta} - \hat{e}^{n}_{\phi}(\gamma_{kl}\hat{e}^{k}_{\phi}\tilde{e}^{l}_{\theta}) - \hat{e}^{n}_{r}(\gamma_{kl}\hat{e}^{k}_{r}\tilde{e}^{l}_{\theta})]}}. \tag{2.4.61}$$

The strange order in which this is applied is for consistency with Ref. [113]. There, they claim it is important to start the Gram-Schmidt procedure with the azimuthal direction vector in order to avoid frame-dragging effects although these are attributed to the use of a vanishing shift vector whereas we use the Gamma driver shift condition (2.3.41). Finally we set $e^{\mu}_{(1)} = \hat{e}^{\mu}_{r}$, $e^{\mu}_{(2)} = \hat{e}^{\mu}_{\theta}$ and $e^{\mu}_{(3)} = \hat{e}^{\mu}_{\phi}$ and construct the null tetrad using (2.4.52).

Now, Ref. [113] proceeds to apply a set of null rotations and boosts (that preserve Eq. (2.4.53)) in order to bring the tetrad into a quasi-Kinnersley form. However, following Refs. [114, 103], we dispense with this step.

2.4.5 The Weyl scalar Ψ_4 and gravitational radiation

In order to relate Ψ_4 to outgoing gravitational waves, we assume the null tetrad is constructed using (2.4.52) from the standard flat-space orthonormal tetrad in spherical coordinates $\{e^\mu_t, e^\mu_r, e^\mu_\theta, e^\mu_\phi\}$. This means that our expressions will only be valid asymptotically.

For the plane-wave solution derived in Sec. 2.4.1, we can write the metric perturbation more generally in the TT gauge as

$$h_{\mu\nu} = h^+ A^+_{\mu\nu} + h^\times A^\times_{\mu\nu}, \tag{2.4.62}$$

where $A^{+,\times}_{\mu\nu}$ are constant symmetric polarization tensors. For an outgoing wave in the radial direction, we choose

$$A^+_{(\theta)(\theta)} = -A^+_{(\phi)(\phi)} = 1, \tag{2.4.63}$$

$$A^\times_{(\theta)(\phi)} = A^\times_{(\phi)(\theta)} = 1, \tag{2.4.64}$$

with all other components vanishing (these are components with respect to the orthonormal basis rather than the coordinate basis hence the parentheses). Note the similarity with the expression for the amplitude tensor of a wave moving in the z-direction (2.4.17).

Now, note that when substituting the definition of the Weyl tensor (2.4.20) into the definition of Ψ_4 (2.4.54), the only non-vanishing term is the first one containing the Riemann tensor (since most contractions of two null tetrad vectors vanish (2.4.53)). We can then use the linearised expression for the Riemann tensor (2.4.3) with our expression for the metric perturbation above (2.4.62) to calculate Ψ_4. Putting this all together, one can show that

$$\Psi_4 = \frac{1}{4}[\partial_t^2 - 2\partial_t\partial_r + \partial_r^2]h, \qquad h = -h^+ + \mathrm{i}h^\times, \tag{2.4.65}$$

where we refer to h as the *strain*. For an outgoing wave, we expect $h = h(t-r)$, so $\partial_r h = -\partial_t h$ and we find that

$$\Psi_4 = -\ddot{h}^+ + \mathrm{i}\ddot{h}^\times. \tag{2.4.66}$$

2.4.6 Spherical harmonic decomposition

It is often convenient to decompose Ψ_4 and h into spherical harmonic modes. Not only is numerical noise filtered by the decomposition process since it tends to be of

higher frequency than the physical signal, but knowledge of the symmetries or even the expected dominant modes of the configuration being simulated enables sanity checks of the numerical solution. Furthermore, some characteristics of gravitational radiation, such as quasinormal modes, are best understood in terms of these multipoles. We compute these modes with respect to the spin weight[45] $s = -2$ spherical harmonics ${}_{-2}Y^{\ell m}$ [115] which are defined and computed using the formula

$$ {}_{s}Y^{\ell m}(\theta,\phi) = (-1)^{s}\sqrt{\frac{2\ell+1}{4\pi}}d^{\ell}_{m(-s)}(\theta)\mathrm{e}^{\mathrm{i}m\phi}, \tag{2.4.67} $$

where the Wigner d functions are given by [103]

$$ d^{\ell}_{ms}(\theta) = \sum_{t=\max(0,m-s)}^{\min(\ell+m,\ell-s)} \frac{(-1)^{t}\sqrt{(\ell+m)!(\ell-m)!(\ell+s)!(\ell-s)!}}{(\ell+m-t)!(\ell-s-t)!t!(t+s-m)!} \times \left(\frac{\cos\theta}{2}\right)^{2\ell+m-s-2t}\left(\frac{\sin\theta}{2}\right)^{2t+s-m}. \tag{2.4.68} $$

Using the orthogonality relation over the unit sphere,

$$ \oint_{S^2} {}_{s}Y^{\ell m}\,{}_{s'}\bar{Y}^{\ell' m'}\,\mathrm{d}\Omega = \delta_{ss'}\delta_{\ell\ell'}\delta_{mm'}, \tag{2.4.69} $$

we can decompose Ψ_4 into a superposition of modes:

$$ \Psi_4(t,r,\theta,\phi) = \sum_{\ell=2}^{\infty}\sum_{m=-\ell}^{m=\ell}\psi_{\ell m}(t,r)\,{}_{-2}Y^{\ell m}(\theta,\phi), \tag{2.4.70} $$

where the multipolar amplitudes are given by

$$ \psi_{\ell m}(t,r) = \oint_{S^2}\Psi_4(t,r,\theta,\phi)\,{}_{-2}\bar{Y}^{\ell m}(\theta,\phi)\,\mathrm{d}\Omega. \tag{2.4.71} $$

In a similar fashion, we decompose the strain as

$$ h = -h^{+} + \mathrm{i}h^{\times} = \sum_{\ell=2}^{\infty}\sum_{m=-\ell}^{\ell}h_{\ell m}(t,r)\,{}_{-2}Y^{\ell m}(\theta,\phi), \tag{2.4.72} $$

[45] A function f has spin weight s if under a rotation by angle ϑ of the angular basis $\{e_\theta, e_\phi\}$, it transforms as $f \to \mathrm{e}^{-\mathrm{i}s\vartheta}f$. It can be shown that Ψ_4 (and thus also h) has spin weight $s = -2$ (see appendix D of Ref. [49])

where we also write $h_{\ell m} = -h^{+}_{\ell m} + \mathrm{i} h^{\times}_{\ell m}$. It then follows from Eqs. (2.4.66) and (2.4.70) that

$$\ddot{h}^{+}_{\ell m} = -\operatorname{Re}(\psi_{\ell m}), \qquad \ddot{h}^{\times}_{\ell m} = \operatorname{Im}(\psi_{\ell m}). \tag{2.4.73}$$

Typically, we compute $\psi_{\ell m}$ according to Eq. (2.4.71) during a numerical simulation and then compute the strain amplitudes by integrating Eq. (2.4.73).

By the peeling theorem [110], we expect $\Psi_4 = \mathcal{O}(1/r)$ as $r \to \infty$, hence we typically consider $r\Psi_4$ (or $r\psi_{\ell m}$, $rh_{\ell m}$, etc.) rather than just Ψ_4. Because we typically extract at finite radius $r = r_{\mathrm{ex}}$, this makes the extrapolation to null infinity more straightforward.

For most astrophysical problems, the $\ell = 2$ modes dominate and for some applications it is reasonable to ignore contributions from $\ell > 2$. However, we will see an explicit example in Chapter 5 where it is important to include them. In the case of axisymmetry, so long as the polar axis for the spherical harmonics is aligned with the axis of symmetry, the $m \neq 0$ modes necessarily vanish (Eq. (2.4.67) depends on m through $\mathrm{e}^{\mathrm{i}m\phi}$).

2.4.7 Radiated energy and momentum

We can calculate an object that can be interpreted as the energy-momentum tensor for GWs by adding a second order correction to the linearised metric (2.4.1) viz.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} + h^{(2)}_{\mu\nu}, \tag{2.4.74}$$

where $h^{(2)}_{\mu\nu} = \mathcal{O}(\epsilon^2)$. Then, at quadratic order in ϵ, the Einstein equation (in vacuum) can be written as [116]

$$ {}^{(4)}R^{(1)}_{\mu\nu}\left[\boldsymbol{h}^{(2)}\right] - \frac{1}{2}\,{}^{(4)}R^{(1)}\left[\boldsymbol{h}^{(2)}\right]\eta_{\mu\nu} = 8\pi t_{\mu\nu}[\boldsymbol{h}], \tag{2.4.75}$$

where ${}^{(4)}R^{(1)}_{\mu\nu}\left[\boldsymbol{h}^{(2)}\right]$ and ${}^{(4)}R^{(1)}\left[\boldsymbol{h}^{(2)}\right]$ are the contributions to the full Ricci tensor and scalar respectively that contain parts linear in $h^{(2)}_{\mu\nu}$ and

$$t_{\mu\nu} = -\frac{1}{8\pi}\left\{ {}^{(4)}R^{(2)}_{\mu\nu}[\boldsymbol{h}] - \frac{1}{2}\,{}^{(4)}R^{(2)}[\boldsymbol{h}]\eta_{\mu\nu}\right\}. \tag{2.4.76}$$

Here ${}^{(4)}R^{(2)}_{\mu\nu}[\boldsymbol{h}]$ and ${}^{(4)}R^{(2)}[\boldsymbol{h}]$ are the contributions to the full Ricci tensor and scalar respectively that are quadratic in $h_{\mu\nu}$[46]. It can be shown that $t_{\mu\nu}$ is symmetric and transverse, hence it can almost be interpreted as an energy-momentum tensor for GWs.

[46]The observant reader may have noticed that we have omitted the term $-\frac{1}{2}\,{}^{(4)}R^{(1)}[\boldsymbol{h}]h_{\mu\nu}$ in this equation. However, this vanishes by the Einstein equation at linear order in ϵ.

Unfortunately it is not invariant under a gauge transformation of the form (2.4.6). This problem can be eliminated if we consider the average over distances much greater than the wavelength of the radiation, which we denote with $\langle\cdot\rangle$. This is known as the *Isaacson stress-energy tensor* [116] and, in the TT gauge, it is given by

$$T_{\mu\nu} := \langle t_{\mu\nu} \rangle = \frac{1}{32\pi} \left\langle \partial_\mu h^{\rm TT}_{\alpha\beta} \partial_\nu h^{{\rm TT},\alpha\beta} \right\rangle . \tag{2.4.77}$$

For an outgoing gravitational wave in the radial direction as discussed in Sec. 2.4.5, in locally Cartesian coordinates, the outgoing energy flux is[47],

$$\frac{{\rm d}E}{{\rm d}t{\rm d}\Omega} = T^{0r} = \frac{1}{16\pi} \left\langle \partial^0 h \partial^r \bar{h} \right\rangle = \frac{1}{16\pi} \left\langle |\dot{h}|^2 \right\rangle , \tag{2.4.78}$$

where we have used $\partial_t h = -\partial_r h$ in the final equality. Integrating this expression in time and over the unit sphere S^2, and using Eq. (2.4.66), we find that the accumulated energy radiated in GWs is given by [117]

$$E^{\rm rad}(t) = \lim_{r\to\infty} \frac{r^2}{16\pi} \int_{t_0}^{t} {\rm d}t' \oint_{S^2} {\rm d}\Omega \left| \int_{-\infty}^{t'} {\rm d}t'' \, \Psi_4(t'', r, \theta, \phi) \right|^2 . \tag{2.4.79}$$

The accumulated linear momentum radiated in GWs $\mathbf{P}^{\rm rad}$ can be calculated similarly:

$$\mathbf{P}^{\rm rad}(t) = \lim_{r\to\infty} \frac{r^2}{16\pi} \int_{t_0}^{t} {\rm d}t' \oint_{S^2} {\rm d}\Omega \, \mathbf{e}_r \left| \int_{-\infty}^{t'} {\rm d}t'' \, \Psi_4(t'', r, \theta, \phi) \right|^2 , \tag{2.4.80}$$

where $\mathbf{e}_r$ is the unit radial vector with Cartesian components

$$\mathbf{e}_r = (\sin\theta\cos\phi,\, \sin\theta\sin\phi,\, \cos\theta). \tag{2.4.81}$$

The formula for the accumulated radiated angular momentum $\mathbf{J}^{\rm rad}$ is a little more complicated to derive (see e.g. Ref. [49] for further details). It takes the form [117]

[47] Here, $\bar{h}$ denotes the complex conjugate of the strain h rather anything to do with trace-reversal.

$$\mathbf{J}^{\rm rad}(t) = -\lim_{r\to\infty}\frac{r^2}{16\pi}\,{\rm Re}\int_{t_0}^{t}{\rm d}t'\left\{\oint_{S^2}\left(\int_{-\infty}^{t'}{\rm d}t''\,\bar{\Psi}_4(t'',r,\theta,\phi)\right)\right.$$
$$\left.\times\hat{\mathbf{J}}\left(\int_{-\infty}^{t'}{\rm d}t''\int_{-\infty}^{t''}{\rm d}t'''\,\Psi_4(t''',r,\theta,\phi)\right)\,{\rm d}\Omega\right\}, \quad (2.4.82)$$

where the angular momentum operator $\hat{\mathbf{J}}$ for spin weight $s=-2$ is given by

$$\hat{\mathbf{J}} = \left({\rm Re}\,\hat{\mathbf{J}}_+, {\rm Im}\,\hat{\mathbf{J}}_+, \frac{\partial}{\partial\phi}\right), \quad (2.4.83)$$

and

$$\hat{\mathbf{J}}_+ = {\rm e}^{{\rm i}\phi}\left({\rm i}\frac{\partial}{\partial\theta} - \cot\theta\frac{\partial}{\partial\phi} + 2{\rm i}\,{\rm cosec}\,\theta\right). \quad (2.4.84)$$

Typically, we calculate Eqs. (2.4.79), (2.4.80) and (2.4.82) at finite extraction radius $r = r_{\rm ex}$ using the extracted values of Ψ_4 there and then extrapolate these to infinity (or at least use the extrapolated value as an estimate of the error from the result at finite extraction radius). The value of the lower integration limit in these formulae should be $t_0 = -\infty$. However, in practice, we choose t_0 in order to exclude the contribution of the spurious "junk" radiation coming from the initial data (see, for example the discussion at the end of Sec. 2.2.3).

2.5 Numerical methods

Now that we have introduced the differential equations we wish to solve, it is time to explore how we translate these equations into a form that can be solved using a computer program. In this section, we focus mainly on techniques applicable to the evolution of hyperbolic equations rather than finding solutions to elliptic problems as is necessary for solving the constraint equations for initial data. However, some of the ideas are applicable to both kinds of problem.

Here we will focus on the method of *finite differences* which is often the first numerical method to solve a differential equation that one typically meets, for example, at an undergraduate level. Whilst the method of finite differences is arguably one of the simpler methods, it is definitely not the only method. Some popular alternatives within the numerical relativity community are *spectral* methods where the solution is expanded in terms of a basis of functions. For example, these are used by the SXS collaboration in the Spectral Einstein Code (SpEC) [118] and by the TwoPunctures code for

constructing binary Bowen-York puncture data [119] (see Sec. 3.2.3). Another approach that is somewhat less common within the numerical relativity community is the method of *finite elements*.

Although the simulations presented later on in this thesis are in three spatial dimensions and use adaptive mesh refinement, for the purposes of this section we will focus on a much simpler example. This will allow us to explain the basic concepts of numerical analysis without getting bogged down in technical details.

2.5.1 The one-dimensional linear wave equation

For many examples in this section, we will consider the linear wave equation in one dimension which is probably the archetypal example of a linear hyperbolic partial differential equation (PDE). In units where the wavespeed $c = 1$, it takes the form

$$\partial_t^2 u(x,t) = \partial_x^2 u(x,t), \qquad (x,t) \in \mathbb{R}^2 \tag{2.5.1}$$

Given suitable initial data, for example,

$$u(x,0) = f(x), \qquad \partial_t u(x,0) = g(x), \qquad x \in \mathbb{R}, \tag{2.5.2}$$

the problem is well posed. In particular, we have d'Alembert's solution

$$u(x,t) = \frac{1}{2}\left[f(x+t) + f(x-t)\right] + \frac{1}{2}\int_{x-t}^{x+t} g(y)\,\mathrm{d}y. \tag{2.5.3}$$

Although the linear wave equation appears to be a relatively simple example that can be solved analytically, we will explore how to solve it numerically as some of the features it exhibits and methods we use will be applicable to the more complicated equations we have already met. Furthermore, as we have seen in Sec. 2.4.1, the linearised Einstein equation can be written in such a way that it is essentially just a simple wave equation (albeit typically in three spatial dimensions). Indeed, even the full nonlinear Einstein equation can be written as a [rather complicated] nonlinear wave equation in some coordinate systems (for example, the harmonic coordinates mentioned at the start of Sec. 2.1) and this is often how mathematical properties of the equation are understood.

2.5.2 Finite differences and evaluating numerical derivatives

In the method of finite differences, the continuous spacetime coordinates x and t are replaced with a grid of a countable number of points with spacings Δx and Δt. For this section, we will restrict to the case of uniform and fixed Δx and Δt. Furthermore, we will consider a finite computational domain, $(x,t) \in \mathscr{D} = [x_{\min}, x_{\max}] \times [0, t_{\max}] \subset \mathbb{R}^2$, rather than considering compactifications. Then, we label our points in the grid as

$$x_j = x_{\min} + j\Delta x, \qquad j = 0, \ldots, J, \tag{2.5.4}$$

$$t_n = n\Delta t, \qquad n = 0, \ldots, N, \tag{2.5.5}$$

where Δx, J, Δt and N are chosen such that $J\Delta x = x_{\max} - x_{\min}$ and $N\Delta t = t_{\max}$. It is conventional to use the following shorthand notation to label the value of u at points in the grid:

$$u^n_j = u(x_j, t_n). \tag{2.5.6}$$

We can use Taylor expansions to approximately calculate derivatives of u given multiple neighbouring grid values u^n_j. For example, using

$$u^n_{j+1} = u^n_j + \Delta x(\partial_x u)^n_j + \frac{(\Delta x)^2}{2}(\partial^2_x u)^n_j + \frac{(\Delta x)^3}{6}(\partial^3_x u)^n_j + \mathcal{O}\left(\Delta x^4\right), \tag{2.5.7}$$

$$u^n_{j-1} = u^n_j - \Delta x(\partial_x u)^n_j + \frac{(\Delta x)^2}{2}(\partial^2_x u)^n_j - \frac{(\Delta x)^3}{6}(\partial^3_x u)^n_j + \mathcal{O}\left(\Delta x^4\right), \tag{2.5.8}$$

we find that

$$(\partial^2_x u)^n_j = \frac{u^n_{j+1} - 2u^n_j + u^n_{j-1}}{(\Delta x)^2} + \mathcal{O}\left(\Delta x^2\right). \tag{2.5.9}$$

It is commonplace to refer to finite difference formulae that calculate approximate derivatives as *stencils*. This example is *centred* because it is symmetric about the point of evaluation (x_t, t_n). In general, we will prefer centred stencils as these minimise the largest number of points needed on one side of the evaluation point and tend to have smaller coefficients in their leading order corrections. In parallelised codes such as GRChombo (see Chapter 3), the points are distributed across multiple processors which means that *ghost* points are needed on either side of the subset of the grid an individual processor is working on in order to evaluate derivatives. At each step, these ghost points are exchanged between processors. Therefore, minimising the number of needed ghost points reduces the amount of communication between processors and memory required to store them thereby improving computational efficiency. Unfortunately, it turns out

that lopsided stencils are necessary for stability when evaluating the *advection* terms (i.e. terms of the form $\beta^k \partial_k u$ in (2.3.16), (2.3.30), (2.3.38) and (2.3.41)) where the "sign" of the lopsidedness depends on the sign of the shift vector component. These stencils tend to be the limiting factor in terms of minimum number of ghost points required.

2.5.3 Consistency, convergence and stability

Before we look at some examples of numerical schemes for the wave equation (2.5.1), it is important to understand the difference between *consistency*, *convergence* and *stability* as these concepts describe how "good" a particular numerical scheme is.

Suppose we have a PDE

$$\mathscr{L} u = 0, \tag{2.5.10}$$

where $\mathscr{L}$ is a differential operator which we approximate using finite differences as $\mathscr{L}_{\Delta x, \Delta t}$. Let $\tilde{u}$ be the true solution to Eq. (2.5.10) and $u_{\Delta x, \Delta t}$ be the finite-difference solution.

2.5.3.1 Consistency

We say the numerical scheme is *consistent* if

$$\lim_{\Delta x, \Delta t \to 0} \mathscr{L}_{\Delta x, \Delta t} \tilde{u} = 0. \tag{2.5.11}$$

In other words, a consistent finite difference scheme correctly approximates the original differential equation and not some other equation in the continuum limit. Typically if one finds the numerical scheme is inconsistent, then one has made a mistake when trying to derive it.

The *truncation error* is

$$\tau_{\Delta x, \Delta t} = \mathscr{L}_{\Delta x, \Delta t} \tilde{u}, \tag{2.5.12}$$

and, if $\tau_{\Delta x, \Delta t} = \mathcal{O}(\Delta x^m, \Delta t^n)$, as $\Delta x, \Delta t \to 0$, we say the method is mth order in space and nth order in time or just kth order where $k = \min(m, n)$.

2.5.3.2 Convergence

We say the numerical scheme is *convergent* if, in the limit as $\Delta x, \Delta t \to 0$ with $x = j\Delta x$ and $t = n\Delta t$ fixed (so $j, n \to \infty$),

$$(u_{\Delta x, \Delta t})^n_j \to \tilde{u}(x, t), \tag{2.5.13}$$

that is the finite-difference solution approaches the true solution in the continuum limit. Note that the difference between convergence and consistency is subtle; consistency is about the *equations* (or scheme) and convergence is about the *solution*.

The *solution error* is

$$\varepsilon_{\Delta x \Delta t} = u_{\Delta x, \Delta t} - \tilde{u}. \tag{2.5.14}$$

Given that we are usually trying to numerically solve a problem that cannot be solved analytically, it is often impossible to determine the convergence of a numerical solution to its continuum limit. However, there may exist a quantity which is a function of the evolution variables for which we *do* know the analytic solution (for example the Hamiltonian and momentum constraints should vanish in the continuum limit) and we can use these as a verification of the convergence. Of course, whilst it is necessary for this derived quantity to converge if the solution is converging, it is usually not sufficient to show convergence of the solution

2.5.3.3 Stability

Firstly, we need to define the L^p norm $\|\cdot\|_{\Delta x,p}$ of a grid function $w_{\Delta x,\Delta t}$ which is defined by

$$\|(w_{\Delta x,\Delta t})^n\|_p := \left(\Delta x \sum_{j=0}^{J} \left|(w_{\Delta x,\Delta t})^n_j\right|^p\right)^{1/p} \approx \left(\int_{x_{\min}}^{x_{\max}} |w|^p \, \mathrm{d}x\right)^{1/p}. \tag{2.5.15}$$

Note that the generalization to higher dimensional spaces is straightforward. For example in three dimensions, we have

$$\|(w_{\Delta \mathbf{x},\Delta t})^n\|_p := \left(\Delta x \Delta y \Delta z \sum_{j_x=0}^{J_x} \sum_{j_y=0}^{J_y} \sum_{j_z=0}^{J_z} \left|(w_{\Delta \mathbf{x},\Delta t})^n_{j_x,j_y,j_z}\right|^p\right)^{1/p}. \tag{2.5.16}$$

We say the numerical scheme is *stable* in some region $\Lambda \subset \mathbb{R}^2$, if, for any $T > 0$, there exists a constant C_T such that

$$\|(u_{\Delta x,\Delta t})^n\|_2 \leq C_T \left\|(u_{\Delta x,\Delta t})^0\right\|_2, \tag{2.5.17}$$

where $0 \leq n\Delta t \leq T$ and $(\Delta x, \Delta t) \in \Lambda$. Intuitively, a stable numerical scheme does not amplify some component of the initial data arbitrarily. Comparing with Eq. (2.3.2), we can see that this is essentially the discrete analogue of well-posedness.

2.5.3.4 The relation between them

An important result in the study of finite difference approximations is the *Lax equivalence theorem* [120] which states that for a consistent finite difference approximation to a well-posed linear initial value problem, stability is a necessary and sufficient condition for convergence. Whilst this applies to our one dimensional wave equation (2.5.1), it does not apply to the nonlinear equations we typically want to solve in NR. Nevertheless, it provides intuition that the convergence of a finite difference approximation is related to its stability.

2.5.4 Von Neumann stability analysis

Proving stability using the definition (2.5.17) is non-trivial, even for linear problems, and we will often rely instead on *von Neumann stability analysis* [121] (also known as Fourier mode analysis). Below we go through the general theory behind the analysis and specific examples can be found in the next section.

Suppose we have a finite difference approximation to a linear IVP of the form

$$\sum_{k=0}^{K}\sum_{r=0}^{R}\mathsf{A}_{rk}\mathbf{u}_{j+k}^{n+r}=0, \tag{2.5.18}$$

where $\mathbf{u}$ is a vector of evolution variables of length m and A_{rk} is an $m\times m$ matrix of constant coefficients for each k and r. In the case of non-constant coefficients, we assume they are slowly varying enough to be considered constants and the analysis applies only locally. We proceed by considering a Fourier mode of the form

$$\mathbf{u}_j^n=\xi^n\mathrm{e}^{\mathrm{i}kj\Delta x}\mathbf{u}^0, \tag{2.5.19}$$

where the superscript ^{n} on the right-hand side is an exponent as opposed to an index. Substituting into Eq. (2.5.18) yields a matrix equation of the form $B\mathbf{u}^0=\mathbf{0}$ for some $m\times m$ matrix B with non-trivial solution if and only if $\det B=0$. This is a polynomial in ξ of degree mR so let the roots be $\xi_i(k)$ for $i=1,\ldots,mR$. We call these the *amplification factors*. The von Neumann stability condition is

$$|\xi_i(k)|\leq 1,\qquad \text{for all } i=1,\ldots,mR,\ k\in\mathbb{R}. \tag{2.5.20}$$

In other words, a mode should not grow in magnitude. This is somewhat similar to our heuristic argument that strong hyperbolicity implied well-posedness for an IVP (2.3.6).

For nonlinear equations, we can linearise and then apply the analysis. Although von Neumann stability analysis is far from rigorous, empirically it is found to be quite reliable in providing the right intuition.

2.5.5 Numerical schemes

We are now ready to look at a few numerical schemes. Often, it is convenient to rewrite the equations so they are first order in time by introducing the auxiliary variable $s = \partial_t u$. Then the wave equation (2.5.1) becomes

$$\partial_t u = s, \tag{2.5.21a}$$

$$\partial_t s = \partial_x^2 u. \tag{2.5.21b}$$

2.5.5.1 Forward time centred space

In the forward time centred space scheme (FTCS), the system (2.5.21) becomes

$$u_j^{n+1} = u_j^n + \Delta t\, s_j^n \tag{2.5.22a}$$

$$s_j^{n+1} = s_j^n + \frac{\Delta t}{4\Delta x^2}\left(u_{j+2}^n - 2u_j^n + u_{j-2}^n\right). \tag{2.5.22b}$$

Note that the non-centred temporal differencing and centred spatial differencing means that this scheme is first-order accurate in time and second-order accurate in space. This is an example of an *explicit* scheme because the value of variables at the latest timestep can be written explicitly in terms of the variables at earlier timesteps. On the other hand *implicit* schemes require a [sparse] matrix equation to be solved at each timestep. This can be computationally expensive, particularly in higher than one dimension which is why we will only consider explicit methods.

Now, consider a von Neumann stability analysis. Rewriting the scheme just in terms of u, we have

$$\frac{u_j^{n+2} - 2u_j^{n+1} + u_j^n}{\Delta t^2} = \frac{u_{j+2}^n - 2u_j^n + u_{j-2}^n}{4\Delta x^2}. \tag{2.5.23}$$

Substituting a Fourier mode of the form (2.5.19) (with $m = 1$) gives the quadratic equation

$$\xi^2 - 2\xi + \left(1 + \alpha_C^2 \sin^2(k\Delta x)\right) = 0, \tag{2.5.24}$$

where $\alpha_C = \Delta t/\Delta x$ with solutions

$$\xi_\pm = 1 \pm \mathrm{i}\alpha_C \sin(k\Delta x). \tag{2.5.25}$$

We can see that $|\xi_\pm| > 1$ in general and therefore, this scheme is unconditionally unstable. We would expect the numerical solution to quickly blow up and this is indeed what happens for generic initial data.

2.5.5.2 Staggered Leapfrog

In the staggered leapfrog scheme we think of $s = \partial_t u$ as living halfway between two timesteps, so the stencil

$$s_j^{n+1/2} = \frac{u_j^{n+1} - u_j^n}{\Delta t} = (\partial_t u)(x_j, t_n + \Delta t/2) + \mathcal{O}(\Delta t^2) \tag{2.5.26}$$

is second-order accurate. The full system is

$$u_j^{n+1} = u_j^n + \Delta t\, s_j^{n+1/2}, \tag{2.5.27}$$

$$s_j^{n+1/2} = s_j^{n-1/2} + \frac{\Delta t}{\Delta x^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n). \tag{2.5.28}$$

Performing a von Neumann analysis yields the amplification factors

$$\xi_\pm = 1 - 2\alpha_C^2 \sin^2(k\Delta x/2) \pm 2\alpha_C \sin(k\Delta x/2)\sqrt{\alpha_C^2 \sin^2(k\Delta x/2) - 1}. \tag{2.5.29}$$

If we have

$$\alpha_C = \Delta t/\Delta x \leq 1, \tag{2.5.30}$$

then $|\xi_\pm| = 1$ for all k but if $\alpha > 1$, then one of $\xi_\pm$ has magnitude greater than 1 and the scheme is unstable. This is known as the *Courant-Friedrichs-Lewy* (CFL or just Courant) stability condition [122] (see Ref. [123] for an English translation) and α_C is called the *Courant factor*. It can be interpreted geometrically as requiring the numerical domain of dependence to contain the physical domain of dependence. For the nonlinear system we typically encounter in numerical relativity (for example (2.3.30)), it is often found empirically that $\alpha_C < \alpha_{C,\max}$ is required for stability, where $\alpha_{C,\max}$ depends on the system of equations and numerical scheme but is often approximately 1/2.

2.5.5.3 Method of lines

It is often convenient to separate the differencing in time and in space. This is known as the Method Of Lines (MOL). More precisely, consider our first-order-in-time wave equation system (2.5.21) and suppose $\Delta_x^2 u$ is some finite difference approximation to $\partial_x^2 u$. Then, consider the system

$$\partial_t u = s, \tag{2.5.31}$$

$$\partial_t s = \Delta_x^2 u, \tag{2.5.32}$$

which is now simply an ODE system in time and hence we can apply standard methods for numerical ODE integration.

In particular, we could use fourth order Runge-Kutta (RK4). Consider a general first-order ODE of the form

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \mathbf{F}(\mathbf{u}), \tag{2.5.33}$$

where $\mathbf{u}$ is a vector of variables. We then calculate the four *substeps*

$$\mathbf{k}_1 = \mathbf{F}(\mathbf{u}), \tag{2.5.34a}$$

$$\mathbf{k}_2 = \mathbf{F}(\mathbf{u} + \mathbf{k}_1 \Delta t/2), \tag{2.5.34b}$$

$$\mathbf{k}_3 = \mathbf{F}(\mathbf{u} + \mathbf{k}_2 \Delta t/2), \tag{2.5.34c}$$

$$\mathbf{k}_4 = \mathbf{F}(\mathbf{u} + \mathbf{k}_3 \Delta_t), \tag{2.5.34d}$$

and the full step is performed using

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \frac{\Delta t}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4). \tag{2.5.35}$$

So long as the IVP (Eq. (2.5.31) with suitable initial data) is well-posed, there exists a nontrivial stability region for the method of lines with RK4 [124]. Because of the simplicity in adapting this method to nonlinear equations, it has become ubiquitous amongst finite difference NR codes and it is what we shall mainly use.

2.5.6 Boundary conditions

Since we work on a finite [non-compactified] computational domain, we need to impose some boundary conditions. One particularly simple example consists of prescribing periodic boundary conditions; for the first-order-in-time wave equation system (2.5.21)

this can be imposed as

$$u_J^n = u_0^n, \tag{2.5.36a}$$
$$s_J^n = s_0^n. \tag{2.5.36b}$$

Unfortunately periodic boundary conditions are often not physically relevant to astrophysical scenarios and instead we prefer a *radiative* boundary condition that does not allow incoming waves from $x < x_{\min}$ or $x > x_{\max}$. This can be written in differential form as

$$\partial_t u = \begin{cases} \partial_x u, & x = x_{\min}, \\ -\partial_x u, & x = x_{\max}. \end{cases} \tag{2.5.37}$$

If we use the method of lines, we can perform the time integration using the same scheme as in the main part of the problem domain. For the spatial discretisation, if our main domain uses stencils that require m points either side of the evaluation point, then we can add m ghost points just outside the main computational domain and use one-sided stencils for the spatial derivatives in Eq. (2.5.37) as necessary. It is usually sufficient to use at most second order stencils for these derivatives even if using higher stencils elsewhere. In the nonlinear case, using lower-order stencils can even be helpful as they tend to be more dissipative than higher order ones.

In the three dimensional case, we typically assume an outgoing spherical radiation condition so that an evolution variable u behaves as

$$u \sim u_\infty + \frac{f(t-r)}{r}, \tag{2.5.38}$$

as $r \to \infty$, where f is an arbitrary function and u_∞ is a prespecified constant that we have added to take into account the fact that for some variables such as the conformal factor χ in the BSSNOK and CCZ4 systems, we require $\chi \to 1$ as $r \to \infty$. In differential terms, this can be written as

$$\partial_t u = -\frac{x^i}{r}\partial_i u - \frac{u - u_\infty}{r}. \tag{2.5.39}$$

This is often called a *Sommerfeld* radiation condition [125].

Although Eq. (2.5.37) is exact for our linear wave equation (and Eq. (2.5.39) for the linear wave equation in three dimensions), in the nonlinear case, this will no longer be true. We generally find that so long as these conditions are applied sufficiently far away

from highly nonlinear physics they work reasonably well in practice, although inevitably there is a small amount of reflection at the boundaries.

2.5.7 Convergence testing

When performing numerical simulations, it is important to test their convergence so that we can obtain a quantitative error estimate, otherwise the results obtained are essentially worthless. Since the CFL condition (2.5.30) means that we always choose $\Delta t \propto \Delta x$, we will now only refer to a single discretisation parameter Δ.

For a stable convergent finite-difference approximation u_Δ to a solution of a PDE $\mathscr{L}u = 0$, if we interpret u_Δ as continuous function of t and x (for example using sufficiently high order interpolation), we expect

$$u_\Delta(x,t) = \tilde{u}(x,t) + \Delta\xi_1(x,t) + \Delta^2\xi_2(x,t) + \ldots + \Delta^i\xi_i(x,t) + \ldots, \tag{2.5.40}$$

where the ξ_i are independent of Δ. This idea is due to an observation by Richardson [126]. For a kth order scheme, we expect that $\xi_i = 0$ for $i = 1, \ldots, k-1$ and $\xi_i \neq 0$ for $i \geq k$.

If we know the true solution $\tilde{u}$ to our problem, we can verify the convergence of our kth order scheme by computing the ratio of the solution errors ε_Δ for two different values of Δ. For example, if $\Delta_1 > \Delta_2$, then

$$\frac{\varepsilon_{\Delta_1}(x,t)}{\varepsilon_{\Delta_2}(x,t)} = \frac{u_{\Delta_1}(x,t) - \tilde{u}(x,t)}{u_{\Delta_2}(x,t) - \tilde{u}(x,t)} = \left(\frac{\Delta_1}{\Delta_2}\right)^k + \mathcal{O}(\Delta_1). \tag{2.5.41}$$

Note that if interpolation is used, it should ideally be at least kth order. However, in practice, these interpolation errors tend to be subdominant at the typical resolutions used.

More often, if we are using numerical methods, we do not know the true solution to our problem. In this case, we can only verify the *Cauchy convergence* of our solution. In other words, we can only check that our method is converging to something at the rate we would expect from the scheme we are employing. This time, we take our approximate solution at three different resolutions $\Delta_1 > \Delta_2 > \Delta_3$ and compute the ratio

$$\frac{u_{\Delta_1}(x,t) - u_{\Delta_2}(x,t)}{u_{\Delta_2}(x,t) - u_{\Delta_3}(x,t)} = Q_k(\Delta_1, \Delta_2, \Delta_3) + \mathcal{O}(\Delta_1) \tag{2.5.42}$$

where Q_k is the *convergence factor* which is given by

$$Q_k(\Delta_1, \Delta_2, \Delta_3) := \frac{\Delta_1^k - \Delta_2^k}{\Delta_2^k - \Delta_3^k}. \tag{2.5.43}$$

In practice, in order to avoid dividing by zero, one usually plots the numerator $u_{\Delta_1} - u_{\Delta_2}$ and the *rescaled* denominator $Q_k(u_{\Delta_2} - u_{\Delta_3})$ and then checks that they roughly agree in order to verify convergence. For more complicated codes, particularly those that use mesh refinement, it is not always obvious what order we expect the solution to converge at. These codes employ a variety of ingredients at different orders (in order to handle the added complications e.g. interpolation at the boundaries of refined regions) and it is not known a priori which of these will dominate at the resolutions used. Therefore, we usually plot the rescaled denominator for several different values of k and assume the convergence order is the one that gives the best agreement with the numerator.

Verification of the convergence order allows us to compute an error estimate via *Richardson extrapolation*. For example suppose we have two numerical solutions at resolutions Δ_1 and Δ_2, where $\Delta_1 > \Delta_2$, and we have verified that these are converging at kth order. Then using Eq. (2.5.40), we find

$$u_k^{\text{Rich}} := \frac{\Delta_1^k u_{\Delta_2} - \Delta_2^k u_{\Delta_1}}{\Delta_1^k - \Delta_2^k} = \tilde{u} + \mathcal{O}(\Delta_1^{k+1}), \tag{2.5.44}$$

that is the kth order *Richardson extrapolant* u_k^{Rich} converges at least one order higher than the two individual solutions. We can then estimate the solution error of a numerical solution (not necessarily one of the two used to compute the extrapolant) by computing

$$\varepsilon_\Delta \simeq u_\Delta - u_k^{\text{Rich}}. \tag{2.5.45}$$

So far we have only considered *local* convergence tests (i.e. the convergence of a solution at a point). It is often more practical, particularly in higher dimensions, to consider *global* convergence tests where we take a norm $\|\cdot\|_p$ of the difference between two solutions. However, rather than analysing the convergence of the evolution variables we will usually analyse the convergence of a diagnostic computed from them, for example a gravitational wave multipole $\psi_{\ell m}$ (2.4.71). Like the global convergence test, this has the advantage of only being a single number (assuming we only look at one extraction radius and a single mode for the case of $\psi_{\ell m}$) for each timestep, so we can easily plot the

differences as a function of time and assess how the convergence changes. Furthermore, this allows us to directly obtain an error estimate for particular quantities of interest.

Chapter 3

GRChombo

In this chapter, we look at the current state of the GRChombo numerical relativity code [1] that has been used to conduct many of the simulations presented in this thesis. Much of the code infrastructure was in place when I was first introduced to it, and thus I am not responsible for large parts of its implementation. Nevertheless, over the last few years I have been one of the main developers of the code and its largest contributor[1] since its public release in January 2018.

The GRChombo code was originally developed by Clough, Figueras, Finkel, Kunesch, Lim, and Tunyasuvunakool [88], a group of collaborators from the University of Cambridge, King's College London and Queen Mary, University of London, in the mid 2010s. Its intention was to provide greater flexibility and enable the exploration of new physics through the use of fully adaptive mesh refinement (AMR) and a modular design. Originally, large parts of the code were written in Fortran. However, in order to enable the use of vector intrinsics for better performance on the latest CPU architectures, the code was ported fully to C++ in 2016. Since then, the GRChombo collaboration has grown in size to over twenty researchers from over eight institutions covering a wider variety of different numerical relativity applications

In Sec. 3.1 we describe the fundamental features of the code and their implementation. Next in Sec. 3.2, we discuss in more detail some specific features that have been contributed to by the author, although this is not comprehensive, and it may be more natural to discuss some features in other sections. Finally, we explain some postprocessing tools in Sec. 3.3 that have been developed in order to analyse GRChombo simulations.

This chapter contains sections from the coauthored publication, Ref. [3] for which I was the sole or main contributor.

[1] This is based on the number of lines added at the time of writing according to GitHub.

3.1 Fundamental code features

The GRChombo code is based on the Chombo library [127] for solving PDEs with adaptive mesh refinement. The tools from Chombo that GRChombo inherits, include the timestepping algorithm, the distributed memory load balancing and the Berger-Rigoutsos style adaptive mesh refinement. The rest of the code has mostly been implemented by the GRChombo collaboration. We describe both parts below.

3.1.1 Structure of the code

In order to improve maintainability and reduce code duplication, we have attempted to consistently organise the main parts of the code according to the following hierarchy:

(i) **Chombo**: These are mostly features related to the underlying Chombo library tools and most of the code in this part of the hierarchy is provided by Chombo itself.

(ii) **GRChombo generic**: These are generic features that are implemented in the GRChombo code that are applicable to all problems that one might use GRChombo for. In particular, this does not include anything about the underlying physics or GR as GRChombo can be used to evolve any suitable hyperbolic equation.

(iii) **Problem-specific**: This is for code that will likely be different for each problem.

We call the collection of code for a specific problem a GRChombo *example*. For the rest of this section, we will discuss the structure of the `BinaryBH` example in GRChombo that is used to evolve binary BH spacetimes.

The fundamental class structure for the `BinaryBH` example is shown in Fig. 3.1. The central class in a GRChombo example is the `AMR` class or rather its derived classes `GRAMR` and `BHAMR`. This stores a vector of pointers of `AMRLevel` objects, each corresponding to a single level in the mesh refinement hierarchy. At the start of a simulation the `AMR` object sets up the initial grid and then orchestrates the evolution. Furthermore, it manages the outputting of checkpoint files and regridding (this is the reshaping of levels finer than the coarsest level in order to increase resolution where it is needed) at user-specified intervals. Although the `AMR` object manages most of the simulation, it does not contain or know about the implementation of how to perform many of the necessary tasks and in most cases this involves just calling member functions of the `AMRLevel` objects and their derived objects. For example, at each timestep on a particular level, in order to "advance"

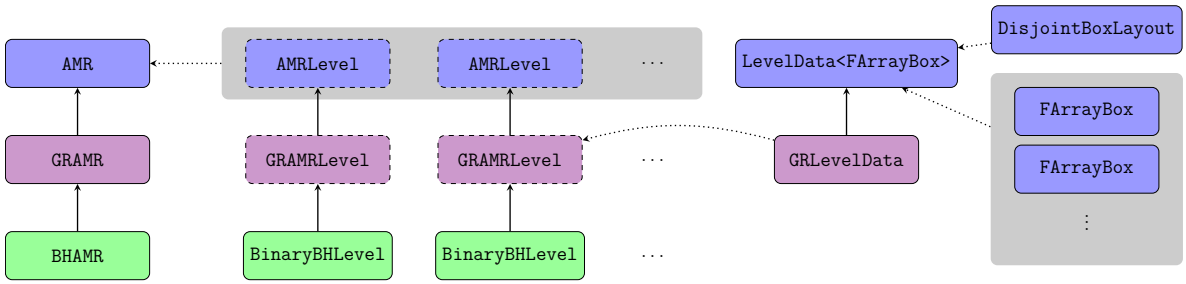


Fig. 3.1 Diagram showing the class structure of the `BinaryBH` example in GRChombo. The blue rounded rectangles are Chombo classes, the purple rounded rectangle are GRChombo generic classes and the green rectangles are problem-specific classes. A solid arrow denotes class inheritance in the direction of child to parent and a dotted arrow indicates an object is a member of another object in the direction of member to aggregator. The grey rectangles indicate that the enclosed objects are stored as a vector of pointers. Finally, a dashed border indicates the class is abstract meaning that part of its implementation must be defined by the child

the solution forward one step, the `AMR` object calls the `GRAMRLevel::advance()` function; this function in turn calls other functions to implement the particular timestepping algorithm (see Sec. 3.1.2).

The `GRAMR` child of the `AMR` class includes generic GRChombo additions to the `AMR` class that require access to the full mesh refinement hierarchy such as a pointer to an `AMRInterpolator` object that allows interpolation of data on the grid using the best resolution available (see Sec. 3.1.6). The `BHAMR` class includes black-hole-specific additions that also require access to the full level hierarchy, for example an `AHFinder` object for finding apparent horizons[2] and a `PunctureTracker` object that tracks the location of the punctures.

Each refinement level is represented by a `BinaryBHLevel` object. The `BinaryBHLevel` class is a child of the `GRAMRLevel` class. The latter implements problem-agnostic features such as the particular timestepping algorithm and how to perform a regrid on the actual data on the level. The former defines all the problem-specific operations on a single level, such as the calculation of the initial data, the evaluation of the right-hand side (RHS) for the timestepping algorithm and any analysis tasks to perform at the end of timesteps (e.g. extracting Ψ_4 or finding apparent horizons).

[2]At the time of writing, the `AHFinder` is still in development and has not been added to the main branch of the public code

Each `GRAMRLevel` object contains a `GRLevelData` object[3] which [via its parent `LevelData<FArrayBox>` class] stores a collection of pointers to `FArrayBox`es (which each store the values of the data in a box) and a `DisjointBoxLayout` object which knows about the layout of these `FArrayBoxes` on the grid and how many ghost cells they contain. Since these boxes are distributed across many processors with MPI, the complexity that arises from this is abstracted away from the user by these classes.

3.1.2 Discretisation and timestepping

GRChombo evolves in time using the method of lines (see the end of Sec. 2.5.5). The time stepping is performed using the standard fourth order Runge-Kutta method (RK4) (2.5.34)-(2.5.35). Until recently, for the spatial discretisation, GRChombo only supported fourth order stencils. However, in order to achieve the necessary accuracy for the simulations presented in Chapters 4 and 5, it was necessary to add support for sixth order stencils. The improvement of sixth order over fourth order is shown in terms of the strain from a BH binary inspiral in Fig. 3.2. For both orders, we use centred stencils except for the advection terms for which we switch to lopsided stencils of the same order depending on the sign of the shift vector component. For completeness, the expressions for the stencils are provided in Appendix 3.A.

Finite difference methods can often introduce spurious high-frequency modes, particularly when using adaptive mesh refinement and regridding. To ameliorate this, GRChombo uses $N = 3$ *Kreiss-Oliger* (KO) dissipation [128]. At every evaluation of the RHS for an evolution variable F, we add the term

$$\frac{\sigma}{64\Delta x}\left(F_{i-3} - 6F_{i-2} + 15F_{i-1} - 20F_i + 15F_{i+1} - 6F_{i+2} + F_{i+3}\right) \tag{3.1.1}$$

to the RHS. Note that the inclusion of this term does not ruin the consistency of our scheme since it vanishes in the continuum limit (it is merely a finite-difference approximation to $(\Delta x^5)\partial_x^6 F = \mathcal{O}(\Delta x^5)$). A von Neumann stability analysis shows that this scheme, when applied to the trivial PDE,

$$\partial_t F = 0, \tag{3.1.2}$$

[3]It actually has three of these objects corresponding to the evolution variables at the current step, the values of the evolution variables at the previous step and the values of any diagnostic variables.

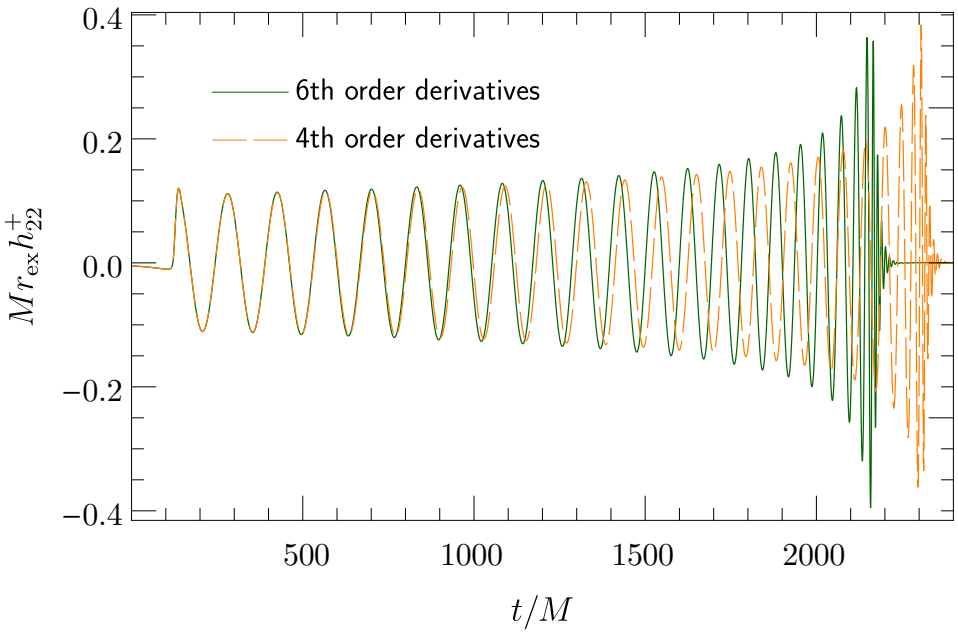


Fig. 3.2 An example of the improved accuracy one obtains with sixth order spatial derivatives compared to fourth order ones. The + polarization of the $(2,2)$ mode of the strain h^+_{22} extracted at $r_{\rm ex} = 120M$ from the simulation of an equal-mass, non-spinning BH binary (configuration `q1-d12` from Table 4.1) is shown. Both cases use the same grid setup and resolution (with finest grid spacing $h = M/80$). The superior accuracy of the sixth order stencils in this case can be inferred from this plot combined with the convergence analysis of this configuration shown in Fig. 4.2.

is linearly stable only if

$$0 \leq \sigma \leq \frac{2}{\alpha_C}, \tag{3.1.3}$$

where $\alpha_C = \Delta t/\Delta x$ is the CFL factor which we typically set to $1/4$[4]. Note that we always use $N = 3$ KO dissipation, independent of the order of spatial discretisation. Naively, one might question this choice for sixth order spatial derivative stencils as the conventional wisdom is to pick N such that $2N - 1 > m$, where m is the order of finite difference scheme (see, for example, Ref. [49]). However, in this case, what matters is the order of the time stepping which is still fourth order[5], hence the dissipation operator does not "spoil" the convergence properties of the scheme. This is consistent with the approach discussed in Sec. 3.2 of Ref. [129].

[4]Note that the stability analysis makes a number of assumptions and problems can begin to appear towards the upper end of this range. A typical symptom of the instability from too high σ is the appearance of a checkerboard-like pattern in otherwise spatially homogeneous regions of the spacetime.

[5]Also note that Theorem 9.1 in Ref. [128] only refers to the order of the time stepping

3.1.3 Parallelisation

Like other numerical relativity codes and, more generally, scientific computing codes, GRChombo exploits parallelisation at several different levels in order to achieve good performance and scaling on modern supercomputers.

For each AMR level, GRChombo splits the domain into boxes and these boxes are shared between processes running on multiple distributed-memory nodes using the Message Passing Interface (MPI). In practice, even though the memory is shared within a node, we typically still use multiple MPI processes per node in order to achieve optimal performance. For example, if a node has n cores, we might choose to use between $n/4$-$n/2$ MPI processes per node. At every *regrid*, we use a load balancing routine in Chombo in order to evenly distribute the boxes. We sort the boxes using a Morton ordering [130], as this minimises communication by increasing the chance that neighbouring boxes are on the same or nearby MPI processes.

One of the most common operations in an NR code is looping through all the cells/points on the grid, calculating some expression and then storing its value in a grid variable. An example is the calculation of the RHS at every RK4 substep which is often where a code spends a large proportion of its time. Within an MPI process, GRChombo uses OpenMP to thread these loop over the z and y coordinates of the boxes. For the x direction, GRChombo relies on SIMD/vector intrinsics[6] in order to utilize the full vector-width of the targeted architecture. We use intrinsics because the complexity of the CCZ4 equations (2.3.30) means that compilers will usually fail to auto-vectorize these loops. The main disadvantage of using SIMD intrinsics is that they are complex and difficult to implement properly. In order to hide the technical implementation from users, many NR codes rely on code-generation scripts to convert more familiar Mathematica/Python expressions to optimized and vectorized Fortran/C/C++ code, for example, Kranc [131] and NRPy+ [132]. GRChombo takes a different approach, keeping the programming at the lower level but relying on C++14 templates to provide a somewhat more user-friendly interface for writing optimized code. Vectorized expressions can be enforced by replacing the C++ type `double` with a template type `data_t`, which represents a vector of values of the variables on the grid of arbitrary length (e.g. the value of χ at the points with x index $i_x = 0, 1, 2, 3$, and constant y and z). In order to make this functionality work, the user is required to write their code in a *compute class*

[6] For the `x86_64` architecture, GRChombo currently supports SSE2, AVX and AVX-512 instructions. We have also tested intrinsics for the ARM `AArch64` architecture using Neon and SVE instructions and plan to add these to the public code soon.

with a `compute` member function which can then be instantiated as an object and then passed to a loop function which calls the `compute` member function in each vector of cells. Multiple compute objects can be combined into a *compute pack* which can then be called by the loop function for added efficiency. For a more detailed description and examples, see Sec 2.5 in Ref. [95].

3.1.4 Berger-Rigoutsos adaptive mesh refinement

In GRChombo, the grid comprises a hierarchy of cell-centred Cartesian meshes consisting of up to $l_{\max}+1$ refinement levels labelled from[7] $l=0,\ldots,l_{\max}$ each with grid spacing

$$\Delta x_l = 2^{l_{\max}-l}\Delta x_{l_{\max}} = \Delta x_0/2^l. \tag{3.1.4}$$

GRChombo uses block-structured AMR, so each refinement level is split into boxes which are distributed between the CPUs as described in 3.1.3.

At regridding or initial grid creation, on a given refinement level l, cells are tagged for refinement according to a *tagging criterion* $C=C(\mathbf{i})$. In a given cell with indices $\mathbf{i}=(i,j,k)$ and corresponding Cartesian coordinates[8] $\mathbf{x}=(x_i,y_j,z_k)$, if $C(\mathbf{i})>\tau_R$, where τ_R is a prespecified threshold value (which may vary with l), then the cell is tagged for refinement. We discuss techniques to design a suitable tagging criterion and aspects to consider in Sec. 4.2.

In block-structured AMR, the main challenge after tagging cells is finding an efficient algorithm to *partition* the cells that require refinement into blocks or boxes. GRChombo uses Chombo's implementation of the Berger-Rigoutsos grid generation algorithm [133] in order to do this.

For this purpose, we define the *block factor* as the number of cells that must divide the side lengths of all blocks; it is a specifiable parameter. Furthermore, these side lengths must not exceed the specified maximum box size. In order to enforce the block factor on level $l+1$, starting with the tagged cells on level l, Chombo generates a temporary new set of tagged cells on a virtual coarser level $l-n$ where n is chosen such that the length of one cell on level $l-n$ corresponds to the block factor on level $l+1$[9]. The new set of coarser tags are derived using a global OR operation, i.e. as long as any of the l level cells corresponding to the coarser level cell is tagged, the virtual coarser level cell will

[7] Note that the finest level that exists may be less than $l_{\max}$.

[8] Note that the indices here are of the discrete cells on the grid as opposed to spacetime components.

[9] For example if the block factor is 4, then $n=1$ since the refinement ratio is 2 and $2^{(l+1)-(l-1)}=4$. Note that this means the block factor must be a power of 2.

be tagged. Chombo then applies the Berger-Rigoutsos partitioning algorithm on this coarser level to construct boxes of grids which obey both the desired block factor and maximum box size. We typically choose both to be a multiple of the processor vector width for optimal performance.

For completeness, we next review the Berger-Rigoustos algorithm (see also Fig. 3.3). We find the minimum box that encloses all of the tagged cells on this level. Let $T(\mathbf{i})$ be the tagging indicator function defined by

$$T(\mathbf{i}) = \begin{cases} 1, & \text{if } C(\mathbf{i}) > \tau_R, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1.5}$$

and define the *signatures* or traces of the tagging by

$$\mathcal{S}_x(i) := \sum_{j,k} T(\mathbf{i}) = \int T(\mathbf{i})\, \mathrm{d}y\, \mathrm{d}z, \tag{3.1.6}$$

$$\mathcal{S}_y(j) := \sum_{i,k} T(\mathbf{i}) = \int T(\mathbf{i})\, \mathrm{d}x\, \mathrm{d}z, \tag{3.1.7}$$

$$\mathcal{S}_z(k) := \sum_{i,j} T(\mathbf{i}) = \int T(\mathbf{i})\, \mathrm{d}x\, \mathrm{d}y. \tag{3.1.8}$$

First, we look for "holes" in the signatures i.e. if there exist i, j or k for which $\mathcal{S}_x(i)$, $\mathcal{S}_y(j)$ or $\mathcal{S}_z(k)$ vanish which corresponds to there being no tagged cells along the plane orthogonal to the signature direction. If there are holes, we choose the one with the largest index over all the dimensions (since it is more efficient to have fewer big boxes than more small boxes) as the plane of partition. If there are no holes, we next look for inflections (see Ref. [133] and their discussion of Fig. 10 for details) in the signatures by computing their discrete Laplacian, for example,

$$\Delta\mathcal{S}_x(i) = \mathcal{S}_x(i-1) - 2\mathcal{S}_x(i) + \mathcal{S}_x(i+1), \tag{3.1.9}$$

and searching for zero-crossings in $\Delta\mathcal{S}_x(i)$. Heuristically, this corresponds to a rough boundary between tagged and untagged regions. (cf. the partitioning in step (2) of Fig. 3.3) If there are inflections, then, in each direction, we pick the inflection with the greatest difference, for example,

$$|\delta(\Delta\mathcal{S}_x(i))| = |\Delta\mathcal{S}_x(i-1) - \Delta\mathcal{S}_x(i)|, \tag{3.1.10}$$

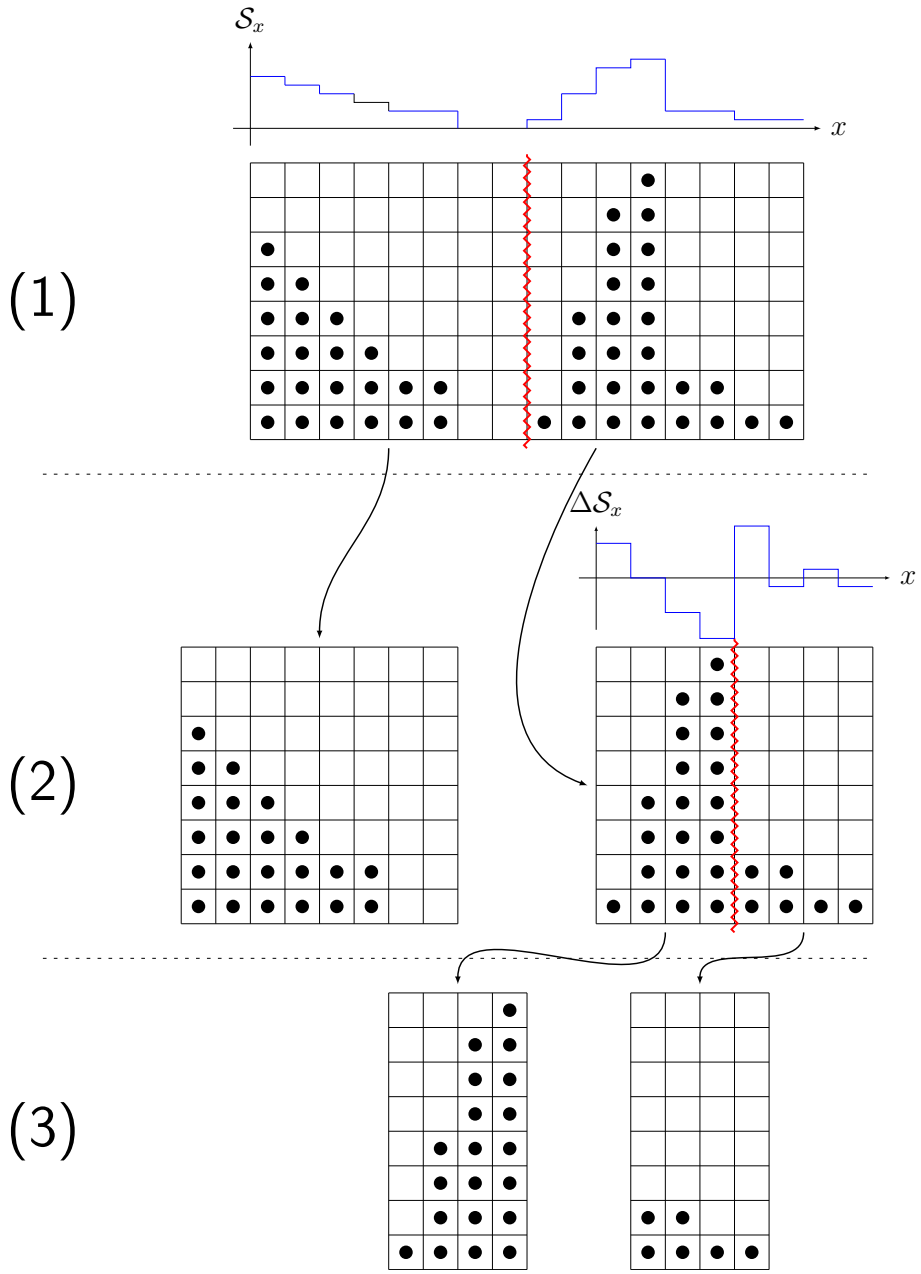


Fig. 3.3 Schematic illustration of the partitioning algorithm. For simplicity, we show a 2D grid and only consider partitioning in the x direction. The cells tagged for refinement are indicated with •. In (1), the signature $\mathcal{S}_x$ is computed and two "holes" are found where the signature vanishes. The line (plane in 3D) of partition is then at the hole with the largest index (rightmost). The result of the partitioning is shown in (2). To partition the right box in (2), the signature is computed, but this time there are no holes, so the algorithm looks for zero crossings of the discrete Laplacian of the signature $\Delta\mathcal{S}_x$. The zero-crossing with the largest change is then selected. This algorithm terminates once all boxes have reached the required fill ratio ϵ_{FR}.

As for the holes, we then pick the greatest inflection index over all the dimensions as our plane of partition. If there are no holes or inflections in the signatures, we simply split the box along the midpoint of the direction with the longest side.

After partitioning, we check whether the partition is sufficiently *efficient*, specifically whether the proportion of tagged cells to all cells in the partition exceeds a user-specified *fill ratio* threshold, $\epsilon_{\text{FR}} < 1$ and that the lengths of the boxes are at most the prespecified maximum box size (which we choose in order to ensure sufficient load balancing). If these tests are passed then we accept the partition and, if not, we continue to partition recursively discarding any boxes that do not contain tagged cells.

Note that, whilst a higher value of ϵ_{FR} will result in a more efficient partition in the sense that there will be a greater ratio of tagged to untagged cells, this is not always the most computationally efficient choice as there are greater overheads with smaller boxes (for example, there will be more ghost cells). There could also be more fluctuation in the structure of the grids between consecutive regrids which may result in greater noise. Although the optimal fill ratio depends on the particular physical problem and the computational resources, we typically use $\epsilon_{\text{FR}} = 0.7$.

Finally, the boxes in the partition are refined, that is, they are defined on the next finer level $(l+1)$ with twice the resolution (3.1.4). For newly created regions on this finer level, we interpolate the data from the coarser level using fourth-order interpolation.

The regridding process starts on the finest level that currently exist (or at most level $l_{\text{max}} - 1$) and works up the hierarchy on increasingly coarse levels until the base level, from whose timestep the regrid was called, is reached (which need not be $l = 0$). It is therefore only possible to add a single extra level (up to l_{max}) at each regrid. After the regrid on level l, the union of the set of cells in the new boxes on this level (plus an additional prespecified *buffer region*) with the set of cells tagged on level $l - 1$ is used as the final set of tagged cells on level $l - 1$ in order to ensure *proper nesting*[10] [134]. This also ensures that cells on coarser levels will be tagged if any of their corresponding finer level cells are tagged.

[10]By proper nesting we mean that

(i) The physical region corresponding to a level $l - 1$ cell must be fully refined or not refined at all, that is it must be completely covered by level l cells and not partially refined.

(ii) There must be at least one level l cell between the boundary of $l + 1$ and the boundary of level l except at the boundary of the entire computational domain. In practice, we even need two such buffer cells (corresponding to 4 cells on level $l + 1$) for fourth and sixth-order spatial stencils.

The frequency of regridding is user-specifiable on each refinement level $l < l_{\max}$, though, since a regrid on level $l = l'$ enforces a regrid on all levels $l' \leq l < l_{\max}$, for problems without rapidly varying (in time) length scales, it is usually sufficient to regrid every few timesteps on one of the coarser levels (e.g. for compact object binaries). Not only does reducing the frequency of regridding reduce the computational cost, but since regridding introduces errors/noise due to interpolation, we have also found that this can improve the accuracy of the simulation.

The Courant condition (2.5.30) limits the size of the maximum time steps one can take on the finer levels. Rather than evolving all refinement levels with the same timestep we use *subcycling* by following the Berger-Colella evolution algorithm [134], which we now review. As the algorithm is recursive, we can consider evolving a set of coarser and finer grids at level l and $l+1$ respectively in the AMR grid hierarchy. First, one time step is taken on the coarser grids (i.e. those at level l). One then evolves the finer (level $l+1$) grids for as many time steps until they have advanced to the same time as the coarse grid. As we have hard-coded the time steps on each level, Δt_l, so that $\Delta t_l = \Delta t_{l-1}/2$, the grids on level $l+1$ will then take two time steps after the grids on level l that takes one time step. After level $l+1$ has "caught up" with level l, the mean of the data in the $[2^3 = 8]$ cells covering a single level l cell is calculated and this value is copied back onto level l. Note that this particular procedure is only second-order accurate in contrast to the restriction operation in a vertex-centred code which requires no approximation. This may partially explain some of the difference in convergence orders we observed between GRChombo and the vertex-centred code Lean in Sec. 4.3.

The ghost cells at the interface between the finer and coarser grids are set by interpolating the values of the coarser grid in both space (due to the cell centered grid) and time (due to the requirement for intermediate values in the RK4 timestepping). The time interpolation is achieved by fitting the coefficients of a 3rd order polynomial in t using the values obtained at each substep of the RK4 timestepping on the relevant cells of the coarser level (see Ref. [135] for more detail).[11]

[11] An alternative approach would be the use of larger ghost zones in the finer level, with the outer ones discarded at each RK step (for example, see Sec. 2.3 of Ref. [136]). One disadvantage here is the extra memory use, especially beyond the fixed-box-hierarchy case.

3.1.5 Boundary conditions

GRChombo implements several types of boundary condition which can be applied independently to the "high" and "low" boundaries in each Cartesian direction. These are:

- Periodic: evolution variables φ obey

$$\varphi(\mathbf{x} + L_i \mathbf{e}_i) = \varphi(\mathbf{x}), \qquad \text{(no sum)} \tag{3.1.11}$$

  where L_i is the length of the computational domain in the ith direction.

- Static: the boundary values of the evolution variables φ are fixed at their initial values by imposing

$$\partial_t \varphi = 0. \tag{3.1.12}$$

- Reflective: across each reflective boundary, an evolution variable φ is assumed to satisfy either odd or even parity reflective symmetry. For example, if imposed on the boundary $x = 0$, we have

$$\varphi(x, y, z) = \pm\varphi(-x, y, z), \tag{3.1.13}$$

  where the parity depends on the specific variable (and the direction of the boundary). One can use these boundary conditions and the symmetry of the spacetime to reduce the volume evolved. For instance, in the case of a simple equal-mass head-on binary BH merger, one needs only 1/8 of the domain[12]; the rest can be inferred from the evolved values [137].

- Extrapolating: the evolution variables φ are assumed to satisfy

$$\varphi = a + bR \tag{3.1.14}$$

  in the boundary cells, where R is the coordinate radius from a prespecified centre (typically the centre of the computational domain), and a and b are coefficients. In the case of zeroth order extrapolation, $b = 0$ and the value of a is simply the value of the φ in the cell closest to the boundary. In the case of first order extrapolation, the coefficients a and b are calculated using the two outermost non-boundary cells that

[12] Of course this particular example is axisymmetric, so one could use an explicitly axisymmetric code instead.

lie on the line perpendicular to the boundary that passes through the boundary cell. These are especially useful for variables which asymptote to a spatially uniform but time varying value (see Refs. [138, 139, 140]).

- Sommerfeld/radiative: see Sec. 2.5.6. Note that, unlike the other conditions, since spatial derivatives are taken parallel to the boundaries, it is important that ghost cells are exchanged between neighbouring boxes that contain boundary cells[13]

Furthermore, for an individual boundary GRChombo supports having mixed boundary conditions with some variables satisfying Sommerfeld conditions and others using extrapolating conditions.

3.1.6 Interpolation of AMR data

Often, when analysing simulations, one requires data to be defined at points that do not coincide with the numerical grid. In order to circumvent this problem, we use interpolation to calculate an approximation to the values of the variables at these points.

Consider the problem of interpolating a function f on a one-dimensional grid with points x_j, $j = 0, \dots J$ and uniform spacing $x_j - x_{j-1} = \Delta x$. Let $f_j = f(x_j)$ be given. If we want to interpolate f at x, where $x \in [x_0, x_J]$, then there must be some $s \in [0, 1)$ and j^* such that $x = x_{j^*} + s\,\Delta x$ (i.e. x_{j^*} is the largest grid point less than or equal to x). Then, we seek an *interpolant* $\tilde{f}(x)$ of the form

$$f(x) \simeq \tilde{f}(x_{j^*} + s\,\Delta x) = \sum_{j=-k+1}^{k} f_{j^*+j} w_j(s), \tag{3.1.15}$$

where $w_j(s)$ are called the *weights*. We use *Lagrange polynomials* where the weights are given by

$$w_j(s) = \prod_{i=-k+1, i\neq j}^{k} \frac{x_{j^*} + s\,\Delta x - x_{j^*+i}}{x_{j^*+j} - x_{j^*+i}} = \prod_{i=-k+1, i\neq j}^{k} \frac{s-i}{j-i}. \tag{3.1.16}$$

These polynomials are of degree $2k-1$ and the error converges as $|f(x) - \tilde{f}(x)| = \mathcal{O}(\Delta x^{2k})$ as $\Delta x \to 0$. We always use $k = 2$ (fourth order) and use the algorithm in Ref. [141] to generate the weights. It should be noted that GRChombo also supports other interpolation algorithms (see, for example, Sec. 7.4 in Ref. [142] for a discussion on the advantages of polynomial convolution methods which are implemented in the code) but we will not discuss these here.

[13] Indeed, without this exchange of ghost cells, it is found empirically that an instability develops.

In order to extend this problem to three spatial dimensions, we recursively iterate on the dimension and apply the algorithm above to interpolate one dimension at a time. We start along the z direction by interpolating $(2k)^2$ points that are at the desired z coordinate but aligned with the grid in the x and y directions. Next we interpolate $2k$ points that are at the desired z and y coordinates but aligned with the grid in the x direction. We then perform the last interpolation along the x axis to obtain the final answer.

Finally, we could ignore the mesh refinement and just interpolate points on the coarsest level that covers the whole computational domain but these results would be far too inaccurate. In GRChombo, the `AMRInterpolator` class starts searching for points on the finest available level and then progresses down the hierarchy to increasingly coarse levels until all requested points are found. This way, the best resolution available for a particular point is always used. Furthermore, the `AMRInterpolator` can be made aware of any reflective boundary conditions corresponding to bitant, quadrant or octant symmetry and infer the value of a point outside the main computational domain but within the unsymmetrised domain from the corresponding point within the domain, applying parity changes as necessary.

Since GRChombo splits the domain across multiple processors using MPI, the `AMRInterpolator` allows each process to request an arbitrary and not necessarily identical set of points. Each point can be anywhere in the computational domain (a process need not restrict to the subset of the domain which it covers). For an individual point, the finest level covering it may, in general, lie on any process, hence, in principle, a process requesting points may need to communicate with an arbitrary number of other processes. In order to do this, an expensive MPI collective [`Alltoallv`] operation is used. This operation is sufficiently flexible such that redundant communication should be minimised. Often, the processing of interpolated data is sufficiently simple so as to not warrant parallelisation (see, for example, Sec. 3.2.1). In these cases, it is important to only request the points from a *single* process rather than *every* process, otherwise the complexity of the MPI collective operations will scale as $\mathcal{O}(N^2)$ rather than $\mathcal{O}(N)$ where N is the number of MPI processes leading to poor scaling of the whole code.

3.2 New code features

3.2.1 General extraction of data on a 2D surface

Although we will primarily be interested in the extraction of Ψ_4 on spheres of fixed coordinate radius in this thesis, in order to reduce code duplication, we have built a general `SurfaceExtraction` class that is *templated*[14] over a `SurfaceGeometry` class. The latter specifies a two-dimensional surface S parametrised by coordinates (u, v) and the former allows the user to straightforwardly extract values of grid variables and their derivatives on S (which uses the `AMRInterpolator` class discussed in Sec. 3.1.6) and then write these values to a file with a uniform format. Furthermore, it then allows the user to perform surface integrals of the form

$$\iint_S f(u, v, \varphi^a)\epsilon_S \,\mathrm{d}u\,\mathrm{d}v, \tag{3.2.1}$$

where f is a user-specified function, φ^a denotes a collection of grid variables that have been interpolated onto the surface and ϵ_S is the *coordinate*[15] area element of S which must be provided in the `SurfaceGeometry` class and write these results to a file in a similarly uniform format. For the numerical integration in each dimension, the code allows the user to choose the quadrature such as the trapezium rule (second-order accurate), Simpson's rule (fourth-order accurate) and Boole's rule (sixth-order accurate).

We now specialise to the case of a sphere at coordinate radius R, $S = S_R^2$. Typically, we use $u = \theta$ and $v = \phi$, the usual spherical polar coordinates and constant grid spacings: $\Delta\theta = \pi/(n_\theta - 1)$ and $\Delta\phi = /(n_\phi - 1)$. This gives the usual $\epsilon_S = R^2 \sin\theta$. Unfortunately, since the aforementioned quadratures for the numerical integration require inclusion of the endpoints $(\theta = 0, \pi)$ as they are *closed*, this means we end up extracting the same value on each pole n_ϕ times despite its contribution to the surface integral vanishing. Furthermore, due to the uniform spacing in θ, the parts of the sphere nearest the pole are better resolved than the parts near the equator $(\theta = \pi/2)$. In order to rectify the latter problems we are currently investigating a more uniform coordinate system on the sphere, where the polar coordinate θ is replaced by $u = z/R$ and the surface area element becomes $\epsilon_S = R^2$. Nevertheless, for the simulations in this thesis, we will use the more conventional

[14] In C++, a class template allows a class to be written for a generic template parameter type `T` that can then be reused with different *specialisations* of `T` (e.g. `T` = `int`, `double`, etc.) without having to rewrite the class.

[15] The metric in the context of GR is not taken into account here and this must be handled separately in f.

parametrisation. We have written a child class `SphericalExtraction` that inherits from `SurfaceExtraction<SphericalGeometry>` (this is the `SurfaceExtraction` template class specialised with `SphericalGeometry`) and additionally provides functions to extract the multipoles of functions on the spheres with respect to the spin-weighted spherical harmonics, ${}_sY^{\ell m}$. In this case, the function f in Eq. (3.2.1) is

$$f(\theta, \phi, \varphi^a) = \frac{1}{R^2} g(R, \theta, \phi, \varphi^a)\, {}_s\bar{Y}^{\ell m}(\theta, \phi), \tag{3.2.2}$$

where g is a user-specified function. Since ϕ is a periodic coordinate on the sphere, the convergence in the ϕ integration, as $\Delta\phi \to 0$, is exponentially fast independent of the type of quadrature used [143]. For this reason, we default to using Simpson's rule for the integration in θ but just use the trapezium rule for the ϕ integration with $n_\theta \approx 49 > n_\phi \approx 32$. It should be noted that the spherical geometry is not the only `SurfaceGeometry` in GRChombo as we have also provided an implementation for a cylinder that was used for some of the simulations in Ref. [144].

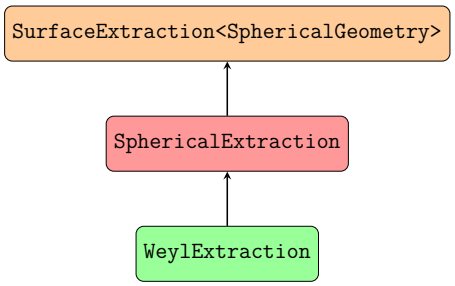


Fig. 3.4 Class inheritance structure of the code used to extract values of $R\Psi_4$ on coordinate spheres.

Finally, we define a child class `WeylExtraction` of `SphericalGeometry` that extracts the real and imaginary parts of Ψ_4 and then calculates the multipolar amplitudes with Eqs. (3.2.1) and (3.2.2) and

$$g(R, \theta, \phi, \Psi_4) = R\Psi_4. \tag{3.2.3}$$

Here the right-hand side is automatically multiplied by R, since we know $\Psi_4 = \mathcal{O}(1/R)$ as $R \to \infty$ (see Sec. 2.4.6). The final class inheritance structure is shown in Fig. 3.4.

3.2.2 Apparent horizon finding

In addition to the extraction of gravitational waves via the extraction of Ψ_4 described in the previous subsection, we will also typically want to be able to locate black hole horizons during a numerical simulation. Typically, the main type of horizon that one thinks of when it comes to black holes are *event horizons* which correspond to the boundary of the region of the spacetime that is causally disconnected from future null infinity $\mathscr{I}^+$. Indeed, the existence of such a region is the defining characteristic of a black hole. Unfortunately, since finding the event horizon requires the maximal development of the spacetime (or at

least a significant part of it) up to $\mathscr{I}^+$, this is generally impractical during a numerical simulation. Instead, we will seek *apparent horizons* which are defined locally as the *outermost marginally trapped surface* in Σ. More precisely, in $3+1$ dimensions, a *trapped surface* is a closed two-dimensional surface $S \subset \Sigma$ such that the *expansion* of outgoing null geodesics orthogonal to S is negative and it is *marginally trapped* if the expansion is non-positive. We then define a *trapped region* as the subset of Σ that contains trapped surfaces and the apparent horizon $\mathscr{H}$ as the boundary of this region. Thus, the expansion must vanish on $\mathscr{H}$. The reason why we can consider apparent horizons is because, if the weak cosmic censorship conjecture and the null energy condition holds, then the apparent horizon must lie within the black hole (Proposition 9.2.8 in Ref. [145]) and in certain cases can be expected to approximate the intersection of the event horizon with Σ.

Let $\boldsymbol{s}$ be the outward-pointing unit normal to S lying in Σ (so $\boldsymbol{s}$ is spacelike) and let $\boldsymbol{n}$ be our usual future-directed unit normal to the foliation (2.1.3). The outgoing null vector on S is given by $\boldsymbol{l} = \boldsymbol{n} + \boldsymbol{s}$ and the induced metric on S is

$$h_{\mu\nu} := \gamma_{\mu\nu} - s_\mu s_\nu. \tag{3.2.4}$$

We next define the expansion Θ of outgoing null geodesics as

$$\Theta := h^{\alpha\beta}\nabla_\alpha l_\beta. \tag{3.2.5}$$

Then, it follows from the formula for the gradients of the unit normal (2.1.22) and $s^\alpha s_\alpha = 1$ that

$$\Theta = h^{\alpha\beta}\nabla_\alpha(n_\beta + s_\beta) = h^{mn}(D_m s_n - K_{mn}) = (\gamma^{mn} - s^m s^n)(D_m s_n - K_{mn}). \tag{3.2.6}$$

The condition for finding the apparent horizon is therefore

$$0 = \Theta = (\gamma^{mn} - s^m s^n)(D_m s_n - K_{mn}). \tag{3.2.7}$$

One of the most common approaches to finding apparent horizons is to assume that $\mathscr{H}$ coincides with the zero contour of some scalar function:

$$F = 0. \tag{3.2.8}$$

The unit normal is then given by $s_i = D_i F/|DF|$, where $|DF| = \sqrt{\gamma^{mn} D_m F D_n F}$ is the normalisation factor. Substituting into Eq. (3.2.7) yields

$$\left(\gamma^{mn} - \frac{D^m F D^n F}{|DF|^2}\right)\left(\frac{D_m D_n F}{|DF|} - K_{mn}\right) = 0, \tag{3.2.9}$$

which is an elliptic PDE for F. There are a variety of different methods to solve this equation and for a detailed review, the reader is referred to Ref. [146].

It is convenient to choose coordinates (u, v, w) that are adapted to the shape of the horizon such that the F can be written as

$$F(u, v, w) = u - f(v, w), \tag{3.2.10}$$

and Eq. (3.2.9) becomes a 2D PDE on v and w. We will assume the horizon is *star-shaped* and choose $(u, v, w) = (R, \theta, \phi)$, the usual spherical polar coordinates. In GRChombo we solve this equation using the SNES (Scalable Nonlinear Equation Solver) library of the PETSc software package [147, 148, 149]. Internally, this uses a line search based on Newton's method to iteratively find the solution.

Once we have found the location of the horizon at $R = f(\theta, \phi)$, we can determine the metric on it in terms of θ and ϕ by substituting $\mathrm{d}R = \partial_\theta f \,\mathrm{d}\theta + \partial_\phi f \,\mathrm{d}\phi$ into the expression for the spatial metric $\boldsymbol{\gamma}$ in spherical polar coordinates to obtain

$$\begin{aligned}\boldsymbol{h} = (\gamma_{\theta\theta} + \gamma_{RR}\partial_\theta f \partial_\theta f + 2\gamma_{R\theta}\partial_\theta f)\,\mathrm{d}\theta^2 + 2(\gamma_{\theta\phi} + \gamma_{RR}\partial_\theta\partial_\phi f + \gamma_{R\theta}\partial_\phi f + \gamma_{R\phi}\partial_\theta f)\,\mathrm{d}\theta\,\mathrm{d}\phi \\ + (\gamma_{\phi\phi} + \gamma_{RR}\partial_\phi\partial_\phi f + 2\gamma_{R\phi}\partial_\phi f)\,\mathrm{d}\phi^2.\end{aligned} \tag{3.2.11}$$

The area of the horizon is then given by

$$A_{\mathscr{H}} = \iint_{\mathscr{H}} \sqrt{h}\,\mathrm{d}\theta\,\mathrm{d}\phi, \tag{3.2.12}$$

where h is the determinant of (3.2.11).

To determine the angular momentum of the horizon, we use the formula from the isolated horizon formalism [150]:

$$J^a_{\mathscr{H}} = \frac{1}{8\pi}\iint_{\mathscr{H}} \varphi^{la} s^m K_{lm}\sqrt{h}\,\mathrm{d}\theta\,\mathrm{d}\phi, \tag{3.2.13}$$

where, for each $a = 1, 2, 3$, the components of the array (i.e. not a tensor) φ^{ia} are the flat-space Killing vectors associated to angular momentum:

$$\varphi^{ia} = \bar{\epsilon}^{iak} x_k. \tag{3.2.14}$$

Here $\bar{\epsilon}^{ijk}$ is the totally antisymmetric alternating symbol ($\bar{\epsilon}^{123} = 1$) and x_k are the Cartesian coordinates. Note that unlike Ref. [150], we do not determine an approximate Killing vector but rather use the so-called "coordinate angular momenta" as in Ref. [151]. Therefore, there is, a priori, no reason to believe that Eq. (3.2.13) accurately calculates the angular momentum. However, it has been found empirically to work very well when the angular momentum is known (e.g. for Bowen-York initial data) and with the conservation of angular momentum calculated from the radiated angular momentum (2.4.82) (assuming the initial angular momentum is known).

Finally, we calculate the mass of the horizon by assuming it is sufficiently close to Kerr, so we can apply Christodoulou's formula [152] with Eqs. (3.2.12) and (3.2.13):

$$M_{\mathscr{H}} = \sqrt{\frac{A_{\mathscr{H}}}{16\pi} + \frac{4\pi |\mathbf{J}_{\mathscr{H}}|^2}{A_{\mathscr{H}}}}. \tag{3.2.15}$$

We have used this quantity to assess the ability of different tagging criteria to suitably resolve the horizons in simulations of black-hole spacetimes. When the horizon is poorly resolved, we have found that this leads to unacceptably large drifts in the horizon mass. This is discussed further in Sec. 4.2.2.

3.2.3 TwoPunctures initial data

Recall from Sec. 2.2 that constructing suitable initial data in GR is non-trivial since one needs to solve the elliptic constraint equations (2.1.36)-(2.1.37) in addition to ensuring the solution physically represents the problem at hand.

As part of the preparation for the simulations presented in Chapter 5, we have integrated the TwoPunctures code [119] into GRChombo. This code constructs binary Bowen-York puncture initial data (see Sec. 2.2.3) using a pseudospectral method. However, since the exponential convergence of spectral methods only applies when the solution is C^∞, and we know that the solution to the Hamiltonian constraint (2.2.29) for the auxiliary function $u = \psi - \psi_{\rm BL}$ is only C^2, the TwoPunctures code switches to a

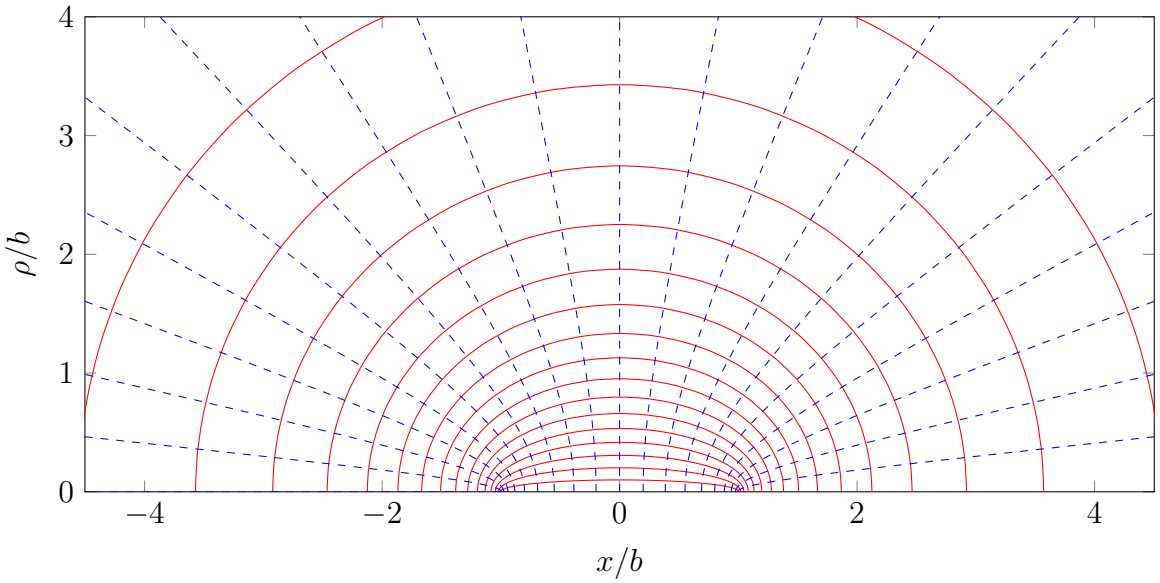


Fig. 3.5 The lines of constant A [red, solid] and B [blue, dashed] from Eq. (3.2.16). Since the grid is axisymmetric about the x-axis, the radius on the vertical axis is $\rho = \sqrt{y^2 + z^2}$. In TwoPunctures, the punctures are located at $x = \pm b$.

new coordinate system (A, B, φ). These are related to the Cartesian coordinates via

$$x = b\frac{A^2+1}{A^2-1}\frac{2B}{1+B^2}, \tag{3.2.16a}$$

$$y = b\frac{2A}{1-A^2}\frac{1-B^2}{1+B^2}\cos\varphi, \tag{3.2.16b}$$

$$z = b\frac{2A}{1-A^2}\frac{1-B^2}{1+B^2}\sin\varphi, \tag{3.2.16c}$$

with $A \in [0,1]$, $B \in [-1,1]$ and $\varphi \in [0, 2\pi)$ and b is half of the coordinate separation between the punctures. The lines of constant A and B are shown in Fig. 3.5. Since the punctures are placed at $x = \pm b$ (note that in the case of unequal masses, this means the origin is not at the centre of mass), it can be seen from the figure that these new coordinates are well adapted to the symmetries of the problem. Furthermore, it can be shown that u is now C^∞ in these new coordinates. [119].

3.2.4 Optimisation of GRChombo

Here we briefly discuss some recent optimisation of the GRChombo code. Since GRChombo is already a relatively highly optimised code (see Ref. [95] for scaling tests performed in 2018), additional optimisations tend to lead to diminishing returns.

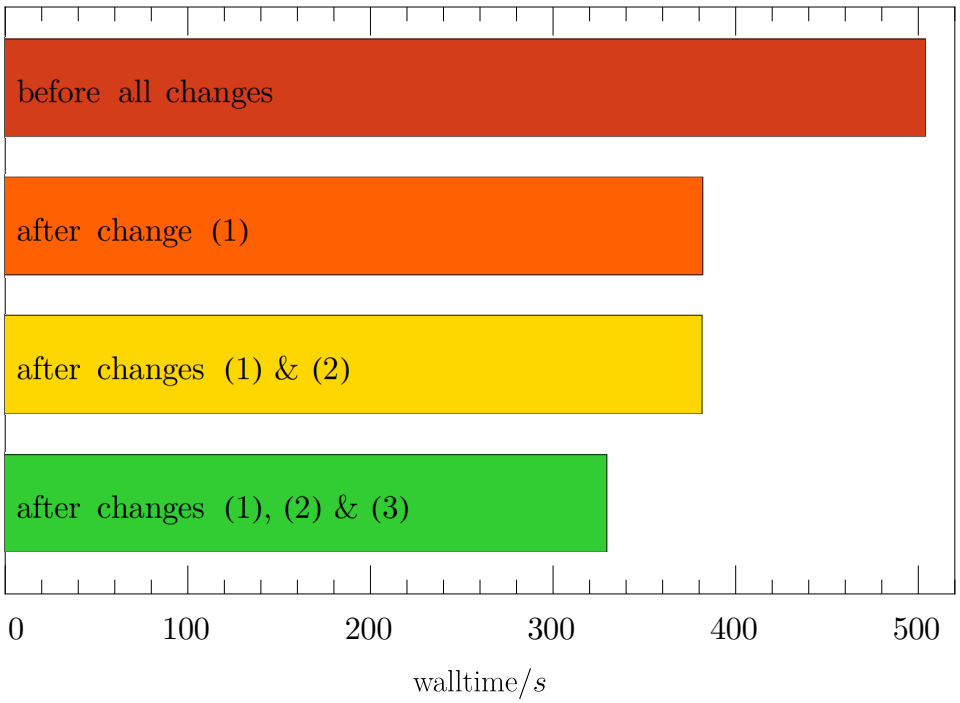


Fig. 3.6 Time taken to evolve 8 timesteps (on the coarsest level) of a typical equal-mass non-spinning BH binary before the optimisations described in Sec. 3.2.4 and after each one. For these simulations we have used 8 dual socket Intel Xeon Platinum 8168 (Skylake) nodes, each with 48 cores. Each simulation used 96 MPI processes and 4 OpenMP threads per process. The size of the whole computational domain is $(384M)^3$, there are $l_{\rm max}+1=8$ refinement levels and the resolution on the finest level is $h=M/64$, where M is the sum of the two BH masses.

Nevertheless, below we describe some recent changes that have led to $\sim 30\%$ speedups in typical BH binary simulations with extraction of Ψ_4:

(1) **Optimisation of SurfaceExtraction:** The first optimisation we have performed on the code is directly related to the comment at the end of Sec. 3.1.6. Before this change, every MPI process requested all of the points on the extraction spheres to be interpolated despite the calculation of modes taking place in serial and the output of data only occurring on a single process, resulting in a large amount of unnecessary communication between processes. After this change, only process 0 requests the points for interpolation and calculates the modes. We found that this reduced the time spent in this part of the code by over 50%.

(2) **Optimisation of the addition of the RHS to the solution:** Next, we considered the function which implements the operation schematically represented by the

equation

$$\texttt{new\_solution} = \texttt{old\_solution} + \Delta t * \texttt{RHS}. \tag{3.2.17}$$

Considering we use RK4 (2.5.34)-(2.5.35), this operation is performed at every substep after the first and also the final completion step (2.5.35), that is 4 times per timestep on every level. Previously, this operation used a function from Chombo but we decided to reimplement it in order to add OpenMP parallelisation for the loops in the y and z directions (as for the calculation of the RHS described in Sec. 3.1.3). However, we typically only use 2-4 OpenMP threads and thus the speedup gained from this optimisation is relatively minor.

(3) **Reduction in the unnecessary exchanging of ghost cells:** The final optimisation concerns the exchanging of ghost cells. Since the `AMRInterpolator` uses ghost cells to interpolate points that are near the sides of the boxes (cf. Eq. (3.1.15)), it is necessary to exchange ghost cells on the levels that will be used for interpolation. Previously, we would exchange ghost cells for all evolution and diagnostic variables on all levels. However, this is computationally expensive due to the large amounts of communication required between the MPI processes, so we added the ability to exchange only a subset of the ghosts. For example, the tracking of the location of puncture (2.3.42) requires interpolation of the shift vector β^i. For any reasonable tagging criterion, we can expect the puncture to lie on one of the finest few levels. Therefore, a significant reduction in the number of ghost cells that need to be exchanged can be made if only the ghost cells for the shift vector components are exchanged and only on the finest few levels. We found this change sped up the tracking of the punctures by over 80%.

The effect of these optimisations is shown in Fig. 3.6 where we show the time taken to evolve a typical BH binary configuration for several timesteps before and after each change.

3.3 Analysis tools

In order to analyse the results coming from GRChombo simulations, we have developed several tools to postprocess the generated data which we briefly describe here.

3.3.1 Calculation of the radiated energy and momentum from Ψ_4 output

In order to calculate the expressions for the radiated energy (2.4.79), linear momentum (2.4.80) and angular momentum (2.4.82) in GWs (see Sec. 2.4.7), we first read in the values of Ψ_4 extracted on the spheres of fixed coordinate radii outputted by the `WeylExtraction` class in GRChombo at each step. Recall that, on each sphere, these lie on a uniform grid in the angular coordinates (θ, ϕ) with uniform spacing in each direction. Consider the expression for the radiated angular momentum

$$\mathbf{J}^{\mathrm{rad}}(t) = -\lim_{r\to\infty} \frac{r^2}{16\pi} \mathrm{Re} \int_{t_0}^{t} \mathrm{d}t' \left\{ \oint_{S^2} \left(\int_{-\infty}^{t'} \mathrm{d}t'' \, \bar{\Psi}_4(t'', r, \theta, \phi) \right) \times \hat{\mathbf{J}} \left(\int_{-\infty}^{t'} \mathrm{d}t'' \int_{-\infty}^{t''} \mathrm{d}t''' \, \Psi_4(t''', r, \theta, \phi) \right) \mathrm{d}\Omega \right\}, \quad (3.3.1)$$

which we use as an example for any of the radiated quantities. We perform all time integrations using Simpson's rule. Note that the lower integration limit for all but the outermost integrals ($-\infty$ in the above formula) is always set to $t = 0$ in practice (i.e. when the simulation starts), whereas we adjust t_0 to exclude the contribution from the junk radiation. For the calculation of the angular derivatives in $\hat{\mathbf{J}}$ (2.4.84), we use second-order centred stencils everywhere except at $\theta = 0, \pi$ where we use second order one-sided stencils. The computation of the angular integration is performed in the same way as for the `SphericalExtraction` class in GRChombo: we use the trapezium rule for the integration over ϕ (due to the previously mentioned exponential convergence of periodic integrals) and Simpson's rule for the integration over θ. We do not evaluate the limit $r \to \infty$ at this step but rather compute a separate result for each finite extraction radius extracted in GRChombo. We then use a separate tool to extrapolate the results to infinity (see Sec. 3.3.3).

3.3.2 Integration of Ψ_4 to obtain the strain

We will often want to translate the Weyl scalar Ψ_4 into the strain h which requires two integrations in time:

$$h = -h^+ + \mathrm{i}h^\times = \int_{t_0}^{t} \mathrm{d}t' \int_{t_0'}^{t'} \mathrm{d}t'' \Psi_4. \quad (3.3.2)$$

The "correct" choice for the lower integration limits is $t_0 = t_0' = -\infty$ but given that our numerical simulations start at some finite time (without loss of generality $t = 0$), we

have two free integration constants which we typically choose to minimise the linear drift. Unfortunately, it is found in practice that even when the linear drift is removed there can still be a residual nonlinear drift when the integration is performed with standard methods such as Simpson's rule. In order to minimise this drift, we switch to the Fourier domain and use the method of *fixed frequency integration* (FFI) first introduced in Ref. [153] which we briefly describe below.

First we set our conventions by defining the Fourier transform of $g(t)$ by[16]

$$\tilde{g}(f) = \mathcal{F}_t[g(t)](f) := \int_{-\infty}^{\infty} \mathrm{e}^{2\pi \mathrm{i} f t} g(t)\, \mathrm{d}t, \tag{3.3.3}$$

and its inverse by

$$g(t) = \mathcal{F}_f^{-1}[\tilde{g}(f)](t) = \int_{-\infty}^{\infty} \mathrm{e}^{-2\pi \mathrm{i} f t} g(f)\, \mathrm{d}f. \tag{3.3.4}$$

Then, integrations in the time domain correspond to frequency divisions in the frequency domain:

$$\int_{-\infty}^{t} g(t')\, \mathrm{d}t' = \mathcal{F}_f^{-1}\left[\mathrm{i}\frac{\tilde{g}(f)}{2\pi f}\right](t) = \mathrm{i}\int_{-\infty}^{\infty} \mathrm{e}^{-2\pi \mathrm{i} f t}\frac{g(f)}{2\pi f}\, \mathrm{d}f \tag{3.3.5}$$

Note that the division by f means that spurious low-frequency modes are amplified by the integration leading to the observed nonlinear drifts that we observe from naive integration in the time domain. In order to eliminate these we pick a *cut-off frequency* f_0 and replace any fs in the denominators in Eq. (3.3.5) with $\max(f, f_0)$. Typically, we will integrate individual modes using this method so, in particular, we set

$$h^{+}_{\ell m} = \mathcal{F}_f^{-1}\left[\frac{\mathcal{F}[\mathrm{Re}\,\psi_{\ell m}](f)}{(2\pi \max(f, f_0))^2}\right], \tag{3.3.6a}$$

$$h^{\times}_{\ell m} = -\mathcal{F}_f^{-1}\left[\frac{\mathcal{F}[\mathrm{Im}\,\psi_{\ell m}](f)}{(2\pi \max(f, f_0))^2}\right]. \tag{3.3.6b}$$

Note that it is important that the real and imaginary parts are computed separately. Given that we have discretely sampled data, in practice we calculate the Fourier transform and its inverse via the discrete Fourier transform (DFT) computed using fast Fourier transforms (FFT) provided by the FFTW3 library [154].

Of course we still have to choose a value for the parameter f_0. In Ref. [153], it is stipulated that this should be set "according to the lowest expected physical frequency for the given wave mode", however this is often not known a priori. In practice, we typically experiment and fine-tune the value. If f_0 is too low, then the nonlinear drifts are not

[16] We follow the conventions of Ref. [52] as it makes the translation to the discrete Fourier transform for discretely sampled data easier.

suppressed and if it is too high, then *physical* low-frequency modes are suppressed (for example, during the early part of the inspiral phase for a waveform coming from a BH binary inspiral).

3.3.3 Extrapolation of radiation to null infinity

For a given *radiated* quantity f (for example, Ψ_4, $E^{\rm rad}$, h, etc.) computed at several finite extraction radii, we extrapolate to infinity by fitting a polynomial in $1/R$ of the form

$$f(u,R) \simeq \hat{f}_N(u,R) = \sum_{n=0}^{N} \frac{\hat{f}_{n,N}(u)}{R^n}, \tag{3.3.7}$$

where R is the coordinate radius and $u = t - r^*$ denotes the retarded time evaluated with the tortoise coordinate

$$r^* = R + 2M \log\left|\frac{R}{2M} - 1\right|. \tag{3.3.8}$$

We compute the fit by resampling f from each extraction radius $R = R_{\rm ex}$ onto a common grid in u and then, at each discrete value of u, performing a linear least squares regression for the coefficients of the $1/R^n$ using the GNU Scientific Library [155].

If we take $\hat{f}_{0,N}(u)$ as our estimate of the extrapolated quantity, we then estimate the error by computing the difference between it and $\hat{f}_{0,N+1}(u)$. Alternatively, we estimate the error of our result from finite radius $R = R_{\rm ex}$ by computing the difference between it and $\hat{f}_{0,1}(u)$. Since we sometimes find that fitting for $N > 1$ results in spurious oscillations or drift, we tend to either take $N = 1$ or use the result from finite extraction radius.

3.3.4 Converting between complex number representations

We will often want to convert a complex oscillating function of time $f(t)$ from its Cartesian representation $f(t) = a(t) + \mathrm{i}b(t)$ to modulus-argument (also known as amplitude-phase) form $f(t) = A(t)\exp(\mathrm{i}\varphi(t))$. Although this conversion is something one typically learns at school level and indeed there is nothing further to add about the calculation of the amplitude ($A = \sqrt{a^2 + b^2}$), we will typically want to choose $\varphi(t)$ such that it is monotonic. Given that we have a discrete number of time points, we assume that Δt is relatively small compared to the inverse of the dominant frequency of $f(t)$. Then, using the standard finite-difference notation (2.5.6) (e.g. $\varphi^n = \varphi(t_n)$), we choose $\varphi^0 = \arg(a(t_0) + \mathrm{i}b(t_0)) \in (-\pi, \pi]$

and

$$\varphi^n = \arg(a(t_n) + \mathrm{i}b(t_n)) + 2k\pi, \tag{3.3.9}$$

where $k \in \mathbb{Z}$ is chosen to minimise $|\varphi^n - \varphi^{n-1}|$. This procedure works well when the phase is a smooth function of time (e.g. during the inspiral phase of a gravitational waveform) but may lead to undesirable jumps when it is not (e.g. for the "junk radiation" part of a waveform cf. Fig. 2.3).

Appendix 3.A Spatial Derivative Stencils

We use the formulae in Ref. [156] for the fourth order stencils. Using the conventional notation for finite differences where,

$$F_i = F|_{x=x_i}, \qquad F_{i,j} = F|_{x=x_i,y=y_j}, \tag{3.A.1}$$

and x_i and y_i are coordinates of the discrete points on a uniform grid, the centred stencils are

$$\partial_x F = \frac{1}{12h}\left(F_{i-2} - 8F_{i-1} + 8F_{i+1} - F_{i+2}\right), \tag{3.A.2}$$

$$\partial_x^2 F = \frac{1}{12h^2}\left(-F_{i-2} + 16F_{i-1} - 30F_i + 16F_{i+1} - F_{i+2}\right), \tag{3.A.3}$$

$$\begin{aligned}\partial_{xy}^2 F = \frac{1}{144h^2}\big(&F_{i-2,j-2} - 8F_{i-2,j-1} + 8F_{i-2,j+1} - F_{i-2,j+2} - 8F_{i-1,j-2} + 64F_{i-1,j-1}\\ &- 64F_{i-1,j+1} + 8F_{i-1,j+2} + 8F_{i+1,j-2} - 64F_{i+1,j-1} + 64F_{i+1,j+1} - 8F_{i+1,j+2}\\ &- F_{i+2,j-2} + 8F_{i+2,j-1} - 8F_{i+1,j+1} + F_{i+1,j+2}\big),\end{aligned} \tag{3.A.4}$$

and, for the advection term, the lopsided stencils are

$$\partial_x F = \begin{cases} \dfrac{1}{12h}\left(-3F_{i-1} - 10F_i + 18F_{i+1} - 6F_{i+2} + F_{i+3}\right), & \text{if } \beta^x > 0, \\ \dfrac{1}{12h}\left(-F_{i-3} + 6F_{i-2} - 18F_{i-1} + 10F_i + 3F_{i+1}\right), & \text{if } \beta^x < 0. \end{cases} \tag{3.A.5}$$

We follow Ref. [129] for the sixth order stencils. The centred stencils are

$$\partial_x F = \frac{1}{60h}\left(-F_{i-3} + 9F_{i-2} - 45F_{i-1} + 45F_{i+1} - 9F_{i+2} + F_{i+3}\right), \tag{3.A.6}$$

$$\partial_x^2 F = \frac{1}{180h^2}\left(2F_{i-3} - 27F_{i-2} + 270F_{i-1} - 490F_i + 270F_{i+1} - 27F_{i+2} + 2F_{i+3}\right), \tag{3.A.7}$$

$$\begin{aligned}\partial_{xy}^2 F = \frac{1}{3600h}\big(&F_{i-3.j-3} - 9F_{i-3,j-2} + 45F_{i-3,j-1} - 45F_{i-3.j+1} + 9F_{i-3,j+2} - F_{i-3,j+3}\\ &- 9F_{i-2,j-3} + 81F_{i-2,j-2} - 405F_{i-2,j-1} + 405F_{i-2,j+1} - 81F_{i-2,j+2} + 9F_{i-2,j+3}\\ &+ 45F_{i-1,j-3} - 405F_{i-1,j-2} + 2025F_{i-1,j-1} - 2025F_{i-1,j+1} + 405F_{i-1,j+2} - 45F_{i-1,j+3}\\ &- 45F_{i+1,j-3} + 405F_{i+1,j-2} - 2025F_{i+1,j-1} + 2025F_{i+1,j+1} - 405F_{i+1,j+2} + 45F_{i-1,j+3}\\ &+ 9F_{i+2,j-3} - 81F_{i+2,j-2} + 405F_{i+2,j-1} - 405F_{i+2,j+1} + 81F_{i+2,j+2} - 9F_{i+2,j+3}\\ &- F_{i+3,j-3} + 9F_{i+3,j-2} - 45F_{i+3,j-1} + 45F_{i+3,j+1} - 9F_{i+3,j+2} + F_{i+3,j+3}\big),\end{aligned} \tag{3.A.8}$$

and, for the advection terms, the lopsided stencils are

$$\partial_x F = \begin{cases} \frac{1}{60h}\left(2F_{i-2} - 24F_{i-1} - 35F_i + 80F_{i+1}\right. & \\ \quad \left. -30F_{i+2} + 8F_{i+3} - F_{i+4}\right), & \text{if } \beta^x > 0, \\ \frac{1}{60h}\left(F_{i-4} - 8F_{i-3} + 30F_{i-2} - 80F_{i-1}\right. & \\ \quad \left. +35F_i + 24F_{i+1} - 2F_{i+2}\right), & \text{if } \beta^x < 0. \end{cases} \tag{3.A.9}$$

Chapter 4

Adaptive mesh refinement in numerical relativity

In this chapter, we demonstrate the flexibility and utility of the Berger-Rigoutsos AMR algorithm described in Sec. 3.1.4 for generating gravitational waveforms from binary BH inspirals, and for studying other problems involving non-trivial matter configurations. We show that GRChombo can produce high quality black-hole waveforms though a code comparison with the established numerical relativity code Lean. We also discuss some of the technical challenges involved in making use of full AMR (as opposed to, e.g. moving box mesh refinement), including the numerical effects caused by using various refinement criteria when regridding. We suggest several "rules of thumb" for when to use different tagging criteria for simulating a variety of physical phenomena. We demonstrate the use of these different criteria through example evolutions of a scalar field theory.

This chapter contains parts of the coauthored publication, Ref. [3]. I performed and postprocessed the GRChombo BH binary inspiral simulations. Furthermore, I implemented the modifications in the GRChombo code that were required in order to test truncation error tagging for the axion star simulations. Finally, I conducted all convergence analyses.

4.1 Introduction

In the numerical relativity (NR) community, many codes rely on the technique of so-called "moving boxes" for mesh refinement, where a hierarchy of nested boxes with increasingly fine meshes centred around specified points (also sometimes referred to as the "box-in-box" approach). Within this framework, boxes move around either dynamically or along

predetermined paths, in order to track objects' trajectories[1]. This technique has proved remarkably successful, particularly in the case of generating gravitational waveforms from binaries of compact objects for the template banks for gravitational wave (GW) detectors, such as LIGO-Virgo-KAGRA [167, 168, 169, 54, 170]. Moving box codes have matured to allow exploration of a wide variety of physics with a plethora of diagnostic tools [49, 50].

However, there are classes of problems for which the moving boxes technique becomes impractical due to the topology of the system. Here, the use of "fully adaptive" mesh refinement (AMR) is required where the mesh dynamically adjusts itself in response to the underlying physical system being simulated, following user-specified mesh refinement criteria. In general, there are two broad classes of AMR, depending on whether newly refined meshes are added to the grid on a cell-by-cell "tree-structured" basis [171, 172, 173] or on a box-by-box "block-structured" basis. In this chapter, we will exclusively discuss the latter.

In block-structured AMR, first described and implemented by Berger and Oliger [174], the computational domain is built from a hierarchy of increasingly fine levels, with each one containing a set of (not necessarily contiguous) boxes of meshes, with the only condition being that a finer mesh must lie on top of one *or possibly more* meshes from the next coarsest level. It is important to stress that this means it is allowed for a fine mesh to straddle more than one coarse mesh—in other words the grid structure is *level*-centric rather than box-centric. In contrast to the moving boxes approach, this approach allows for highly flexible "many-boxes-in-many-boxes" mesh topologies, enabling the study of dynamical systems where the spacetime dynamics are not driven by localized compact systems e.g. in studying nonspherical collapse scenarios [175, 176], higher dimensional black holes/black string evolution [177, 178, 179, 180, 181, 182], cosmic string evolution [183, 184, 185], and the behaviour of strongly inhomogeneous cosmological spacetimes [186, 187, 188, 189, 190, 191, 192, 193, 194].

Despite its advantages, AMR is a double-edged sword and its flexibility comes with a cost—each coarse-fine transition may introduce unwanted interpolation and prolongation errors whose magnitude depends on the order of the coarse-fine boundary operators, in addition to introducing a "hard surface" which can generate spurious unphysical reflections. We emphasize that AMR should not be treated as a "black box", but requires

[1] In addition, boxes may be allowed to merge if they come close enough together. These codes include those built upon the popular Cactus computational framework [157, 158, 136] such as the McLachlan [159], LazEv [156], Maya [160], Lean [161, 162] and Canuda [163] codes, and the BAM code [164, 103, 165, 166].

careful control and fine-tuning of refinement criteria, that often depend on the physics being simulated, in order to work effectively. In particular, the creation/destruction of a finer grid is determined by the tagging of cells for refinement, which in turn is controlled by a *tagging criterion*. Although this ability to refine regions can be incredibly powerful, in practice it can be difficult to manage the exact placement of refined grids. Furthermore, we find that the management of coarse-fine boundaries in dynamically sensitive regions of spacetime, such as near apparent horizons, is essential for producing accurate results.

In this chapter we explain some of the tagging criteria and numerical techniques we have used to obtain convergent, reliable results when using block-structured AMR. We will discuss these issues in the context of the AMR NR code GRChombo [1, 88][2], which was first introduced in 2015 and uses the Chombo [127] library. While our methods apply directly to GRChombo, we believe many of the lessons we have learned are general and may be useful to researchers who work with other numerical relativity codes that make use of block-structured AMR, in particular those which rely on the Berger-Rigoutsos [133] style grid generation methods.

We demonstrate the utility of the techniques we have employed through a direct comparison of gravitational waveforms generated by binary black-hole inspiral and merger calculated by GRChombo and the more established Lean code which uses the aforementioned "moving boxes" style mesh refinement, and show that GRChombo is capable of achieving comparable production-level accuracy. We secondly apply AMR techniques to the evolution of several scalar field models which exhibit dynamics on a wide range of spatial and temporal scales, to demonstrate the relative advantages of several tagging criteria implemented in GRChombo.

This chapter is organised as follows:

- In Sec. 4.2 we discuss considerations for tagging criteria in AMR grid generation.

[2]Adaptive mesh refinement is now being used in several other NR codes. For example, those based on the PAMR/AMRD mesh refinement libraries [195, 196, 197], the Hahndol code [40], which uses PARAMESH [198], the HAD code [199], and the pseudospectral codes, SpEC [200], bamps [201] and SFINGE [202] in which the AMR implementation is somewhat different to finite difference codes like GRChombo. More recently, CosmoGraPH [203], and Simflowny [204] both based on the SAMRAI library [205, 206, 207], GRAthena++ [208], Gmunu [209] and Dendro-GR [210], all based on oct-tree AMR, and GRAMSES [173] have been introduced. Alternatives are problem-adapted coordinate systems, e.g. NRPy+ [132] or discontinuous Galerkin methods as in SpECTRE [211, 212] (see also [213, 214]). Furthermore, it should be noted that some code frameworks, such as those based on the Carpet mesh refinement driver [136], are technically capable of performing block-structured AMR. However, it can be cumbersome to use, and these codes typically rely on moving-box type methods (e.g. the codes referenced in footnote 1). A brief overview of the history of NR codes can be found in [41].

- In Secs. 4.3 and 4.4, we illustrate how these techniques are applied in practice to simulations of BH binaries and spacetimes with a (self-interacting) scalar field.

Throughout this chapter we use geometric units $c = G = 1$. In the sections on black holes we set the mass scales with respect to the ADM mass of the spacetime, whereas for the section on scalar fields we set $\mu = mc/\hbar = 1$, which then describes lengths relative to the scalar Compton wavelength[3].

4.2 Considerations for tagging criteria used for grid generation

We have found that the choice of tagging criteria can greatly impact the stability and accuracy of a given simulation. Here we mention several factors to consider when designing tagging criteria for use in GRChombo and other codes with similar AMR algorithms. We also provide some explicit examples of tagging criteria and discuss their relative merits.

4.2.1 Buffer Regions

One of the problems of many tagging criteria we have tried is that they can often introduce several refinement levels over a relatively small distance in space. This leads to the boundaries of these refinement levels being particularly close to one another. Due to the errors introduced by interpolation at these boundaries, they can add spurious reflections or noise. This is exacerbated when other refinement boundaries are nearby, allowing for this noise to be repeatedly reflected and even amplified before it has time to dissipate (e.g. via Kreiss-Oliger dissipation – see Sec. 3.1.2). A particularly simple way to mitigate this problem is to increase the *buffer regions*, i.e. the number of cells n_{B} between refinement levels. Since the regridding algorithm starts at the finest level and works up the hierarchy to coarser levels (see Sec. 3.1.4), increasing this parameter actually increases the size of the coarser levels rather than shrinking the finer levels in order to enforce this buffer region restriction.

[3]If we are interpreting the results in terms of a physical particle mass, this is equivalent to setting the value of $\hbar$ in the code. Note that in general $\hbar \neq 1$ in NR simulations (as this would imply that one unit in the length scale is equal to the Planck length l_{Pl}).

4.2.2 Considerations for black-hole spacetimes

Here we describe several techniques that we have used when creating tagging criteria to evolve black-hole spacetimes. The major complication with evolving black holes is that they have an event horizon. In practice, it is often challenging to find the true event horizon, which would require tracing geodesics through the full evolution of the spacetime. Therefore NR simulations typically consider the location of the apparent horizon instead, which always lies inside the event horizon [145]. Mathematically, the region within an apparent horizon is causally disconnected from its exterior. For a given numerical approximation, however, artefacts from the discretization can propagate from behind the horizon and contaminate the rest of the computational domain.

As a consequence of this superluminal propagation of numerical noise, we often find that GRChombo simulations of BHs are particularly sensitive to the presence of refinement boundaries. One should avoid adding refinement within the horizon (which in any case is unobservable and not usually of interest), but problems are particularly severe where a refinement boundary intersects the apparent horizon. In such cases we have observed significant phase inaccuracies and drifts in the horizon area (some even violating the second law of black hole mechanics). Similar problems may occur if a refinement boundary is close to but does not intersect the horizon. In order to avoid these issues, we typically enforce the tagging of all cells within the horizon plus a buffer radius up to a maximum level $l_{\rm BH}^{\rm max}$ (which need not necessarily be $l_{\rm max}$ and may differ for each BH in the simulation). If r_p is the coordinate distance from the puncture of a BH of mass $M_{\rm BH}$ in a spacetime with total mass $M \sim 2M_{\rm BH}$, then, for $\eta \sim 1/M$ in the moving puncture gauge (2.3.38-2.3.41), after the initial gauge adjustment the apparent horizon is at approximately $r_p = M_{\rm BH}$ (see Fig. 4 in Ref. [103]). Guided by this approximation, we can tag all cells with $r_p < M_{\rm BH} + b$, where b is a prespecified parameter. Although one might think choosing $b \propto M_{\rm BH}$ for each BH might be the most sensible choice for unequal mass configurations, we have found larger BHs less sensitive to the presence of refinement boundaries. Thus, choosing $b \propto M$ the same for each BH in a binary usually works sufficiently well.

Increasing the size of the buffer regions between refinement boundaries by adjusting $n_{\rm B}$ as discussed in Sec. 4.2.1 can help to keep refinement boundaries sufficiently spaced apart. However, we have also separately enforced the spacing out of refinement boundaries by doubling the radius of the second and third-finest levels covering a BH. This leads to

tagging cells on level l (to be refined on level $l+1$) with

$$r_p < (M_{\rm BH} + b) 2^{\min(l_{\rm BH}^{\rm max} - l - 1, 2)} \tag{4.2.1}$$

In spacetimes where BH horizons are dynamical (often the target of AMR simulations), one can in principle use the locations of apparent horizons to define tagged regions. However, rather than incorporating the output of a horizon finder into the tagging criterion, a simpler and in most cases equally effective method can be obtained from using contours of the conformal factor χ and tagging regions with $\chi < \chi_0$, where χ_0 is a prespecified threshold value which may vary on each refinement level. This gives a reasonably robust and general method of identifying the approximate locations of horizons. Further details on precise values and their dependence on the BH spin are given in Appendix 4.A.

4.2.3 Asymmetric grids

The grid-generation algorithm (Sec. 3.1.4) is inherently asymmetric, for example, it picks the "hole" with the *largest* index as the partition plane. This means that even if the tagging has symmetries, the grids themselves may not obey the same or any symmetries. For example, whilst one might expect that, for tagging cells with Eq. (4.2.1), the grids would have reflective symmetry in all three coordinate directions, this is often not the case, particularly for larger $\epsilon_{\rm FR}$. This asymmetry can lead to undesirable behaviour. For example, when simulating the head-on collision of two BHs with no symmetry assumptions (as described in Sec. 3.1.5) with the tagging of Eq. (4.2.1), the punctures can deviate slightly from the collision axis. We can "fix" this asymmetry by replacing

$$r_p \to \varrho = \max(|x - x_p|, |y - y_p|, |z - z_p|) \tag{4.2.2}$$

in Eq. (4.2.1) so that the tagged regions are boxes rather than spheres (this tagging is similar to what is done in some moving-box style mesh refinement codes). Whilst there is inevitably a loss of efficiency from this choice, this is typically outweighed by the reductions in error achieved. Clearly, this approach pushes the AMR method in the direction of a moving boxes approach; in practice, we therefore apply it predominantly to BH simulations but not for more complex matter structures that require the full flexibility of AMR.

4.2.4 Using truncation error for tagging cells

Truncation error tagging was introduced by Berger and Oliger [174]. We have implemented truncation error tagging in GRChombo by using a *shadow hierarchy* (e.g. Ref. [215]). In this scheme, we estimate the truncation error on a grid at level l by comparing the solution of a specially chosen variable f on that level to the coarser level directly "beneath" it on the grid:

$$\tau_{l,f}(\mathbf{i}) = |f_l(\mathbf{i}) - f_{l-1}(\mathbf{i})| \,. \tag{4.2.3}$$

We note that the error (4.2.3) clearly must be computed before we average the finer grid values onto the coarser grid. As Chombo uses a cell-centred scheme, in order to compare the values of f on the two levels, we interpolate f from the coarser level onto the finer level using fourth order interpolation. If we compute the truncation error of multiple grid variables, we combine the error estimates for each variable at each point:

$$\tau_l(\mathbf{i}) = \sqrt{\sum_f \frac{\left(\tau_{l,f}(\mathbf{i})\right)^2}{L_f}}, \tag{4.2.4}$$

where L_f is a normalizing factor for each variable f. We then set this as our tagging criterion in Eq. (3.1.5): $C(\mathbf{i}) = \tau_l(\mathbf{i})$. The free parameters in this scheme of tagging are the choice of grid variables that one computes truncation error estimates for and the normalization factors for each variable.

The main advantages of truncation error tagging are that it allows for a conceptually straightforward way to implement convergence tests in AMR codes: as one increases the base grid resolution, one should scale the truncation error tagging threshold for grid generation with the expected convergence of the code. Additionally, truncation error tagging is a "natural" tagging criterion as it refines regions that are most likely to be under resolved.

4.2.5 Tagging criteria based on grid variables and derived quantities

Some physical problems lend themselves to other tagging criteria, and GRChombo permits the user to easily specify refinement criteria based on any properties of the local grid variables or derived expressions of them, for example, derivatives or curvature scalars. We caution though that the tagging criteria we discuss below are not functions of geometric scalars, so the performance of a given criterion will depend on the formulation and gauge

conditions used. Nevertheless, for the Bona-Masso-type slicing (2.3.37) and gamma-driver (2.3.41) conditions we have the most experience with, these gauge conditions have proven to be reliable and robust.

First we discuss tagging criteria based on the conformal factor of the spatial metric χ. Contours of χ can provide a good choice in dynamical BH cases as detailed in Sec. 4.2.2 and Appendix 4.A, to ensure that horizons are covered. Taking differences of χ across a cell using locally evaluated derivatives, i.e. using $C = \sqrt{\delta^{ij}\partial_i\chi\partial_j\chi}\Delta x$, also provides an efficient measure to refine key areas[4]. In particular, using the second derivative of χ, i.e. $C = \sqrt{\delta^{ij}\partial_i\partial_j\chi}\Delta x$, is efficient because usually it is the regions in which gradients are changing most rapidly that require greater resolution, rather than steep linear gradients. However, in practice any derivative can be used provided the thresholds are tuned appropriately for the problem at hand.

Alternatively, we find empirically that the sum of the absolute value of the different components of the Hamiltonian constraint proves to be an efficient tagging criteria in dynamical matter spacetimes. The condition is

$$C = \mathcal{H}_{\rm abs} = |R| + |\tilde{A}^{kl}\tilde{A}_{kl}| + \frac{2}{3}K^2 + 16\pi|\rho| + 2|\Lambda|, \tag{4.2.5}$$

where $R = \gamma^{ij}R_{ij}$ is the Ricci scalar. As we will see in Sec. 4.4.2, this quantity generally remains constant in regions of spacetime where the individual metric and matter components oscillate in a stable, time-invariant manner (as in the case of the stable axion star we present later). Thus, using this measure reduces the amount of spurious regridding that occurs, which in turn reduces errors introduced by that process. Where it starts to grow in some region, this generally reflects a decrease in the local dynamical timescales and thus physically motivates regridding.

A disadvantage of using these more arbitrary criteria over error tagging is that convergence testing is more challenging - one must ensure that similar regions are refined at the appropriate resolutions in the convergence runs, which necessitates tuning of the threshold at each different base resolution. Depending on the regridding condition, halving the threshold τ_R, for example, may not result in double the resolution being applied. Nevertheless, if one ensures that the regions of most physical significance have an appropriate increase in refinement, convergence can usually be demonstrated, as we show below.

[4]Imposing simply that $\partial_i\chi$ (without the factor of Δx) is higher that some threshold results in unlimited regridding, since one does not reduce the local gradient in a variable by refinement, only the difference across the cell.

In order to obtain convergent results and use resolution most efficiently in a physical problem, it is often helpful to implement rules to enforce that given regions are refined for a given amount of time at least to a given level. Whilst this may seem to go against the spirit of AMR, it is easy to implement within that formalism as a secondary condition, and is often required to avoid excessive or insufficient tagging in very dynamical cases. For example, when one is not interested in resolving outgoing scalar radiation, one may choose to suppress regridding above a particular level outside a particular radius. In the opposite sense, we often enforce extra regridding over the extraction surfaces for the Weyl scalars, to ensure that they have sufficient resolution and that noise is not introduced from grid boundaries crossing the spheres.

Several examples of the application of these criteria to black-hole binary inspirals and matter field evolutions are presented in the following, sections 4.3 and 4.4. These two examples cover the main considerations when using AMR in NR codes. Spacetimes with singularities have particular requirements related to the resolution of the horizons. Furthermore, in dynamical matter spacetimes achieving an optimum frequency of regridding can be crucial in obtaining convergence. In the matter case we focus on an isolated real scalar (axion) star, which provides a very good test of AMR capabilities. In particular, it tests the ability to resolve stably oscillating matter configurations without excessive gridding and ungridding, and to follow the dynamical timescales of gravitational collapse, which are the key challenges in many simulations of fundamental fields in NR, including the modelling of cosmic strings, inflationary spacetimes and exotic compact objects. Fully AMR techniques are also likely to create significant challenges for high-resolution shock capturing, but we leave this topic for future investigations.

4.3 Binary black-hole simulations with adaptive mesh refinement

In this section, we demonstrate the efficacy of some of the techniques discussed in Sec. 4.2 in the context of BH binaries. To do this, we select a representative sample of BH binary configurations, analyse the accuracy of the resulting gravitational waveforms and compare the results obtained with GRChombo to that obtained with a more conventional moving boxes style mesh refinement code, Lean [161].

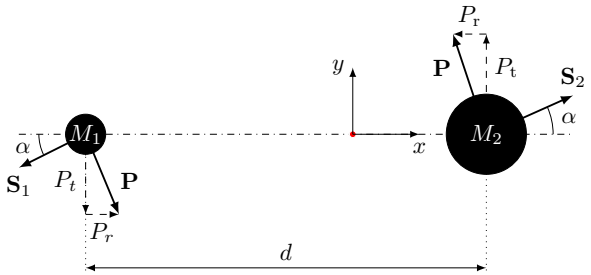


Fig. 4.1 Schematic illustration of the parameters characterizing the BH binary configurations under consideration: the mass ratio $q = M_2/M_1 > 1$, the initial separation d, the initial tangential linear momentum $P_{\rm t}$, the initial inward radial linear momentum $P_{\rm r}$, the dimensionless spin of each black hole $\chi_i = |\mathbf{S}_i|/M_i^2$ and the angle of the spin in the orbital plane α relative to the outward radial direction of the initial BH positions.

Table 4.1 A list of the parameter values (cf. Fig. 4.1) for the BH binary configurations simulated in this chapter. $M = M_1 + M_2$ denotes the total black hole mass of the spacetime.

Label	q	d/M	$P_{\rm t}/M$	$P_{\rm r}/M$	χ_i	α	Reference
`q1-d12`	1	12.21358	0.08417	5.10846×10^{-4}	0	-	[216]
`q2-d10`	2	10	0.08566	0	0	-	[2]
`q1-s09`	1	11.01768	0.075	0	0.9	30°	[217]

Before we present our results, we first provide details of the explicit tagging criteria used in our GRChombo simulations and the methods we use to analyse and compare our results.

4.3.1 Methods

We consider three different BH binary configurations with the parameters provided in Table 4.1 and illustrated schematically in Fig. 4.1. All simulations include an inspiral, merger and ringdown.

The first configuration consists of two equal-mass non-spinning BHs with a quasi-circular inspiral lasting about 10 orbits; this configuration is labelled `q1-d12` (for mass ratio $q = 1$ and distance $d \approx 12$). The parameters were computed in order to minimize the initial eccentricity of the simulation using standard techniques [216].

The second configuration involves two BHs with mass ratio 2 : 1. The inspiral is about 6 orbits and is approximately quasicircular. This is one of the configurations

simulated in the `lq1:2` sequence of Chapter 5. Here we label this configuration `q2-d10` (for mass ratio $q = 2$ and distance $d = 10$).

The final configuration consists of a mildly eccentric inspiral of two equal-mass highly-spinning BHs. The spins lie in the orbital plane as shown in Fig. 4.1, which is the "superkick" configuration [218, 219, 220]. Here, the quantity we analyse is the gravitational recoil of the remnant BH. This configuration is taken from the sequence simulated in Ref. [217] and we label it `q1-s09` (for mass ratio $q = 1$ and spin $\chi = 0.9$).

4.3.1.1 GRChombo setup and tagging criteria

For the GRChombo simulations of the BH binary configurations in Table 4.1, we use the CCZ4 equations (2.3.30) with the default damping parameters (2.3.34) (note that in code units, $M = 1$). We use the moving puncture gauge (2.3.39, 2.3.41) with the default lapse parameters (2.3.40) and the shift parameters $b_1 = 1$, $b_2 = 3/4$, $M\eta = 1$ for `q1-d12` and `q1-s09` and $b_1 = 1$, $b_2 = 3/4$, $M\eta = 3/4$ for `q2-d10`. For `q1-d12` and `q2-d10`, we use reflective BCs along one boundary to impose bitant symmetry (i.e. symmetry across the equatorial plane) and Sommerfeld BCs for all other boundaries. Following sections Secs. 4.2.2, 4.2.3 and 4.2.5, we use a tagging criterion of the following form

$$C = \max\left(C_\chi, C_{\rm punc}, C_{\rm ex}\right), \tag{4.3.1}$$

where the quantities on the right-hand side are defined below. Note that we use the value $+\infty$ to denote a large value that always exceeds the threshold τ_R.

(i) C_χ tags regions in which the derivatives of the conformal factor χ become steep. It is the dominant criterion for the intermediate levels $l_{\rm ex}^{\rm max} \leq l < l_{\rm max} - 3$, where $l_{\rm ex}^{\rm max}$ is the maximum extraction level (see item (iii) below). It is given by

$$C_\chi = \Delta x_l \sqrt{\sum_{i,j} \left(\partial_i \partial_j \chi\right)^2}, \tag{4.3.2}$$

where Δx_l is the grid spacing on refinement level l.

(ii) $C_{\rm punc}$ includes parts of the tagging criterion that use the location of the punctures. It is the dominant criterion on the finest three levels and is comprised of two parts, $C_{\rm insp}$ and $C_{\rm merg}$ that are used depending on the coordinate distance between the

punctures $s_p = |\mathbf{x}_{\mathrm{p},1} - \mathbf{x}_{\mathrm{p},2}|$ as follows:

$$C_{\mathrm{punc}} = \begin{cases} C_{\mathrm{insp}}, & s_p \geq M + b \\ \max(C_{\mathrm{insp}}, C_{\mathrm{merg}}), & 10^{-3} \leq s_p < M + b \\ C_{\mathrm{merg}}, & s_p < 10^{-3}, \end{cases} \tag{4.3.3}$$

where $M = M_1 + M_2$ is the sum of the individual BH masses, b is a buffer parameter (cf. Sec. 4.2.2), and 10^{-3} is a choice in the cutoff for the distance between the punctures s_p which determines when the merger has completed. The inspiral criterion is given by

$$C_{\mathrm{insp}} = \begin{cases} +\infty, & \text{if } \varrho_1 < (M_1 + b)2^{\min(l_{\max} - l - 1, 2)}, \\ & \text{or } \varrho_2 < (M_2 + b)2^{\min(l_{\max} - l - 1, 2)}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.3.4}$$

where ϱ_i is the "max" or "infinity" norm (4.2.2) of the coordinate position vector relative to puncture i. Similarly, the merger criterion is given by

$$C_{\mathrm{merg}} = \begin{cases} +\infty, & \text{if } \varrho_{\bullet} < (M + b)2^{\min(l_{\max} - l - 1, 2)}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.3.5}$$

where $\varrho_{\bullet}$ is the max-norm (4.2.2) of the coordinate position vector relative to the centre of mass $\mathbf{x}_{\bullet} = (M_1 \mathbf{x}_{p,1} + M_2 \mathbf{x}_{p,2})/M$.

(iii) C_{ex} ensures the Ψ_4 extraction spheres are suitably well resolved. It is the dominant tagging criterion for $0 \leq l < l_{\mathrm{ex}}^{\max}$ and is given by $C_{\mathrm{ex}} = \max_i\{C_{\mathrm{ex},i}\}$, where i labels the extraction spheres and

$$C_{\mathrm{ex},i} = \begin{cases} +\infty, & \text{if } r < 1.2 r_{\mathrm{ex},i} \text{ and } l < l_{\mathrm{ex},i}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.3.6}$$

where $r_{\mathrm{ex},i}$ and $l_{\mathrm{ex},i}$ are the radius and level of the ith extraction sphere and $l_{\mathrm{ex}}^{\max} = \max_i l_{\mathrm{ex},i}$. The factor of 1.2 is present to add a 20 % buffer radius around the extraction spheres in order to reduce the effect of spurious reflections off the refinement level boundaries.

A summary of the grid configuration parameters is given in Table 4.2.

Table 4.2 GRChombo grid parameters for the simulation of configurations in Table 4.1 There are $(l_{\max}+1)$ refinement levels and the coarsest level has length (without symmetries applied) L. The grid spacing on the finest level is $\Delta x_{l_{\max}}$ and the minimum number of cells in the buffer regions between consecutive refinement level boundaries is n_B. The regrid threshold for the tagging criterion (4.3.1) is τ_R and the BH buffer parameter is b.

Configuration	$l_{\max}$	L/M	$\Delta x_{l_{\max}}/M$	n_B	τ_R	b/M
`q1-d12` low	9	1024	1/80	20	0.016	0.7
`q1-d12` medium	9	1024	1/96	24	0.0133	0.7
`q1-d12` high	9	1024	1/128	32	0.01	0.7
`q2-d10` low	7	512	1/88	48	0.01	0.467
`q2-d10` medium	7	512	1/104	52	0.00923	0.467
`q2-d10` high	7	512	1/112	56	0.00857	0.467
`q1-s09` low	7	512	1/64	16	0.02	0.7
`q1-s09` medium	7	512	1/96	24	0.0133	0.7
`q1-s09` high	7	512	1/112	28	0.0114	0.7

4.3.1.2 Comparison code: Lean

The Lean code [161] is based on the Cactus computational toolkit [157] and uses the method of lines with fourth-order Runge-Kutta time stepping and sixth-order spatial stencils. The Einstein equations are implemented in the form of the Baumgarte-Shapiro-Shibata-Nakamura-Oohara-Kojima (BSSNOK) formulation [81, 82, 83] with the moving-puncture gauge [39, 40] (cf. Eqs. (2.3.37) and (2.3.41)). The Carpet driver [136] provides mesh refinement using the method of "moving boxes." For the non-spinning binary configurations `q1-d12` and `q2-d10`, we use bitant symmetry to reduce computational expense, whereas configuration `q1-s09` is evolved without symmetries. The computational domains used for these simulations are characterized by the parameters listed in Table 4.3. The domain comprises a hierarchy of $l_{\max}+1$ refinement levels labelled from $l = 0, \dots l_F, \dots, l_{\max}$, with grid spacing given by Eq. (3.1.4). Before applying the symmetry, for $l \leq l_F$ each level consists of a single fixed cubic grid of half-length $R_l = R_0/2^l$, and for $l_F < l \leq l_{\max}$, each level consists of two cubic components of half-length $R_l = 2^{l_{\max}-l} R_{l_{\max}}$ centred around each BH puncture. We adopt this notation for consistency with that used to describe GRChombo. This translates into the more conventional Lean grid setup notation (cf. Ref. [161]) as

$$\left\{(R_0, \dots, 2^{-l_F} R_0) \times (2^{l_{\max}-l_F-1} R_{l_{\max}}, \dots, R_{l_{\max}}), \Delta x_{l_{\max}}\right\}. \tag{4.3.7}$$

Table 4.3 LEAN grid parameters for the simulation of configurations in Table 4.1. There are $(l_{\max}+1)$ levels of which the first (l_F+1) comprises a single box that covers both BHs with the remaining levels consisting of two separate box components that cover each BH separately. The half-width of the coarsest level is R_0 and the half-width of a single component on the finest level is $R_{l_{\max}}$. The grid spacing on the finest level for the three resolutions used in the convergence analysis is $\Delta x_{l_{\max}}$.

Configuration	$l_{\max}$	l_F	R_0/M	$R_{l_{\max}}/M$	$\Delta x_{l_{\max}}/M$
`q1-d12`	9	5	512	1/2	1/64, 1/96, 1/128
`q2-d10`	8	4	256	1/3	1/84, 1/96, 1/108
`q1-s09`	8	3	256	1	1/80, 1/88, 1/96

A CFL factor of 1/2 is used in all simulations, and apparent horizons are computed with AHFINDERDIRECT [221, 222].

For all our BH evolutions, with LEAN and GRCHOMBO, the initial data are constructed with the TWOPUNCTURES spectral solver [119].

4.3.1.3 Gravitational wave analysis

One of the most important diagnostics from our simulations is the GW signal which we compute from the Weyl scalar Ψ_4. For GRCHOMBO, the calculation of Ψ_4 is explained in Sec. 2.4 and technical details of the extraction procedure can be found in Sec. 3.2.1. For LEAN, details can be found in Ref. [161]. Below, we describe further analysis we have performed in order to compare the gravitational wave output from each code.

We start with the multipolar decomposition of the Weyl scalar,

$$\Psi_4 = \sum_{l=2}^{\infty} \sum_{m=-\ell}^{\ell} {}_{-2}Y^{\ell m} \psi_{\ell m}. \tag{4.3.8}$$

Next, we translate to the gravitational-wave strain h according to

$$\Psi_4 = \ddot{h} = -\ddot{h}^{+} + \mathrm{i}\ddot{h}^{\times} \tag{4.3.9}$$

which gives us the strain multipoles as $\ddot{h}^{+}_{\ell m} = -\mathsf{Re}(\psi_{\ell m})$ and $\ddot{h}^{\times}_{\ell m} = \mathsf{Im}(\psi_{\ell m})$. To avoid spurious drift resulting from numerical inaccuracies, we perform the necessary integrations in time in the Fourier domain [153]. We then rewrite the strain modes in terms of their amplitude and phase

$$-h^{+}_{\ell m} + \mathrm{i}h^{\times}_{\ell m} = h^{A}_{\ell m} \exp\left(\mathrm{i}h^{\phi}_{\ell m}\right), \tag{4.3.10}$$

where multiples of 2π are added to $h^{\phi}_{\ell m}$ appropriately in order to minimize the difference between consecutive data points.

The *radiated* quantities derived from Ψ_4 are affected by two main error sources; the discretization error due to finite resolution and an uncertainty arising from the extraction at finite radii instead of null infinity. We determine the former by conducting a convergence analysis of the quantities extracted at finite radius. In order to determine the second error contribution, we compute a given radiated quantity f at several finite extraction radii and extrapolate to infinity by fitting a polynomial in $1/r$ of the form

$$f_N(u,r) = \sum_{n=0}^{N} \frac{f_{n,N}(u)}{r^n}. \tag{4.3.11}$$

Here, r is the coordinate radius and $u = t - r^*$ denotes the retarded time evaluated with the tortoise coordinate

$$r^* = r + 2M \ln \left| \frac{r}{2M} - 1 \right|. \tag{4.3.12}$$

We uniformly observe that time shifts in terms of r^* result in slightly better alignment of wave signals extracted at different coordinate radii r. If we take $f_{0,N}(u)$ as our estimate of the extrapolated quantity, we then estimate the error ϵ in our result from $r = r_{\rm ex}$ by computing

$$\epsilon_{f,r_{\rm ex},N} = \left| f(u, r_{\rm ex}) - f_{0,N}(u) \right|. \tag{4.3.13}$$

Typically, and unless stated otherwise, we set $N = 1$ and drop the N subscripts. Our total error budget is then given by the sum of the discretization and extraction uncertainties.

We quantify the agreement between the two codes' results in the context of GW analysis by computing the overlap following the procedure of Refs. [223, 224]. In the following, we restrict our analysis to the dominant (2,2) quadrupole part of the signal and drop the subscript "$\ell = 2, m = 2$". Before computing the overlap, we extrapolate the strain to infinity using the procedure explained above.

Given the power spectral density $S_n(f)$ of a detector's strain noise as a function of frequency f, the inner product of two signals g, h on the space of waveforms is given by[5]

$$\langle g|h\rangle := 4\,\mathrm{Re}\left\{ \int_0^\infty \frac{\tilde{g}^*(f)\tilde{h}(f)}{S_n(f)}\,\mathrm{d}f \right\}, \tag{4.3.14}$$

[5]We use in our calculation one-sided, as opposed to two-sided, spectral power densities, i.e. we only consider non-negative frequencies, hence the factor 4 in Eq. (4.3.14).

where the Fourier transform is defined by

$$\tilde{g}(f) := \int_{-\infty}^{\infty} g(t) \mathrm{e}^{-2\pi \mathrm{i} f t} \, \mathrm{d}t. \tag{4.3.15}$$

We next define the overlap of the two signals as the normalized inner product maximized over shifts Δt, Δh^{ϕ} in time and phase,

$$\rho(g,h) := \max_{\Delta h^{\phi}, \Delta t} \frac{\langle g|h\rangle}{\sqrt{\langle g|g\rangle\langle h|h\rangle}} \,. \tag{4.3.16}$$

The quantity $1-\rho(g,h)$ then provides a measure for the *discrepancy* between the two waveforms, analogous to the *mismatch* introduced as a measure for signal-to-noise reduction due to model imperfections in GW data analysis [225, 226].

For `q1-s09`, we instead analyse the convergence of the linear momentum radiated in GWs in the form of the BH recoil velocity or kick. First we compute the radiated momentum $\mathbf{P}^{\mathrm{rad}}$ using Eq. (2.4.80) and then compute the recoil velocity–which must lie in the z-direction by symmetry–using $v = -P_z^{\mathrm{rad}}/M_{\mathrm{fin}}$. Since the radiated momentum can be written in terms of a sum, with each term involving several multipolar amplitudes $\psi_{\ell m}$ (Eq. (40) in [97]), analysing this quantity has the benefit of additionally indirectly comparing the agreement of higher order multipoles (i.e. $\ell > 2$) between the codes.

4.3.2 Results

For each configuration in Table 4.1, we have performed three simulations at different resolutions with both GRChombo and Lean in order to calibrate their accuracy which we discuss below. The respective grid configurations are given in Tables 4.2 and 4.3.

For the first configuration `q1-d12` of an equal-mass binary, we show the convergence analysis in Fig. 4.2 with the analysis for GRChombo on the left and for Lean on the right. For Lean, we observe convergence of about fourth order in the amplitude and between fifth and sixth order in the phase of the quadrupole mode h_{22} of the strain (4.3.10). For GRChombo we observe convergence of about second order in the amplitude and about fourth order in the phase of the same mode. We note that, as mentioned in Sec. 5.A, higher resolutions were required with GRChombo in order to enter the convergent regime.

By comparison with a Richardson extrapolation, we estimate the discretization errors in the amplitude and phase of the finest resolution simulations from both codes as follows. Excluding the early parts of the signal dominated by "junk" radiation and the

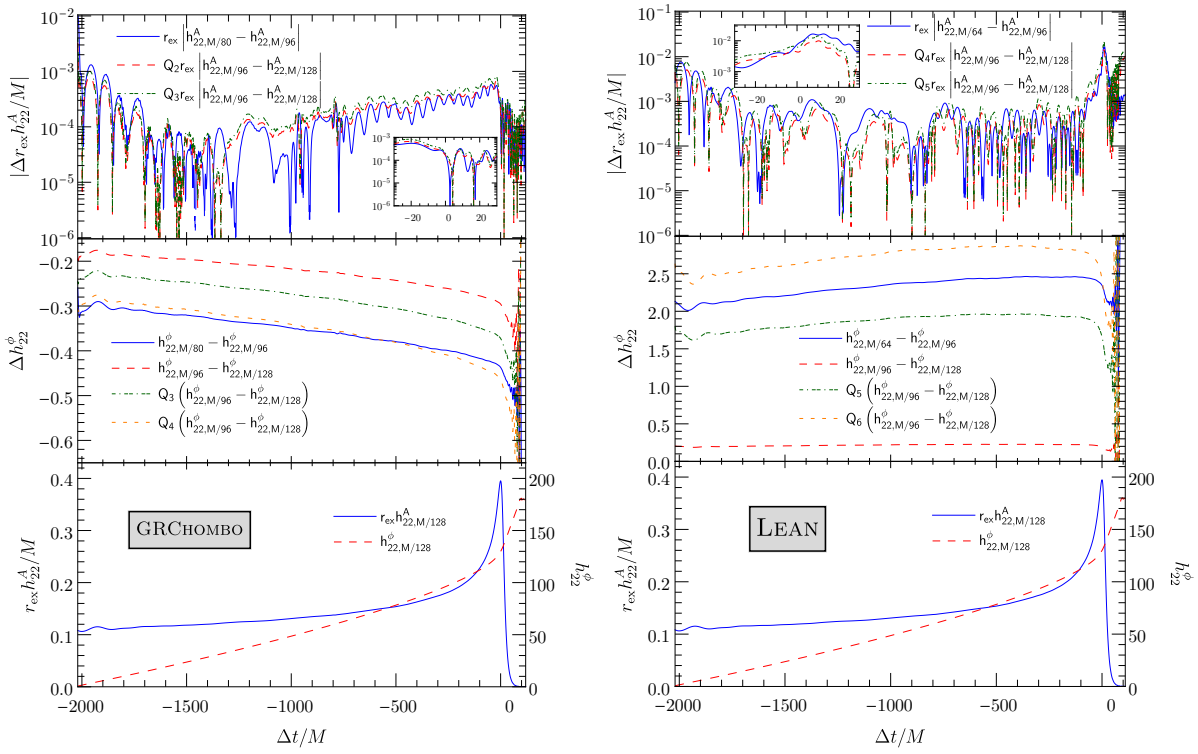


Fig. 4.2 Convergence of the quadrupole mode of the strain $h_{22} = -h^{+}_{22} + \mathrm{i}h^{\times}_{22}$ calculated from the values of Ψ_4 extracted for configuration `q1-d12` at $r_{\mathrm{ex}} = 120\,M$ for both GRChombo with finest grid resolutions $\Delta x_{l_{\max}} = M/80$, $M/96$ and $M/128$ (a) and Lean with finest grid resolutions $\Delta x_{l_{\max}} = M/64$, $M/96$ and $M/128$ (b). *Top panels*: Convergence of the amplitude $h^{A}_{22} = |h_{22}|$. The difference between the higher resolution results is rescaled according to fourth and fifth order convergence for Lean and according to second and third-order convergence for GRChombo. In each case, the inset shows an interval around the peak amplitude. *Middle panels*: Convergence of the phase $h^{\phi}_{22} = \mathrm{Arg}(h_{22})$. The difference between the higher resolution results is rescaled according to fifth and sixth order convergence for Lean and third and fourth order convergence for GRChombo. *Bottom panels*: For reference we plot the amplitude h^{A}_{22} and the phase h^{ϕ}_{22} of the highest resolution waveform on the same time axis. For the two lower resolution waveforms from each code, we have time-shifted each of them in order to maximize the overlap (cf. Eq. (4.3.16)) with the highest resolution waveform. $\Delta t = 0$ corresponds to the maximum in h^{A}_{22} for the highest resolution waveform.

late part of the ringdown which is dominated by noise, we obtain a discretization error of $\Delta h^{A}_{22}/h^{A}_{22} \lesssim 1\,\%$ in the amplitude assuming fourth order convergence for Lean and second order convergence for GRChombo. Up to the late ringdown where the phase becomes inaccurate, we estimate the phase error is $\Delta h^{\phi}_{22} \leq 0.15$ assuming sixth order convergence for Lean and fourth order convergence for GRChombo.

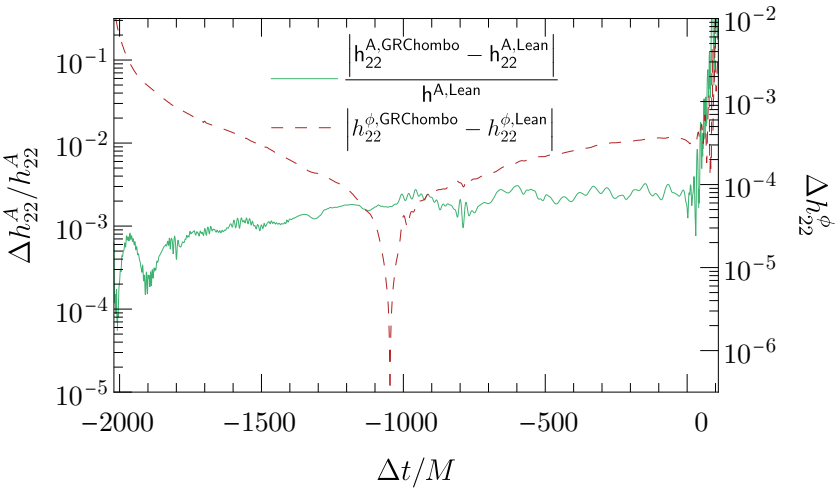


Fig. 4.3 The relative and absolute difference between the GRChombo and Lean outputs for the amplitude and phase of h_{22} respectively from the simulation of configuration `q1-d12`. In both cases, the data comes from the simulations with finest grid spacing $\Delta x_{l_{\max}} = M/128$ with Ψ_4 extracted at $r_{\rm ex} = 120M$. As for the convergence plots in Fig. 4.2, the time has been shifted in order to maximize the overlap (cf. Eq. (4.3.16)) between the two waveforms and $\Delta t = 0$ at the peak in $h_{22}^{A,\rm Lean}$.

Following the procedure in Sec. 4.3.1.3, we estimate the error in the phase, due to finite-radius extraction, is $\epsilon_{h_{22}^{\phi},120M} \lesssim 0.4$, and, in the amplitude is $\epsilon_{h_{22}^{A},120M}/h_{22}^{A} \lesssim 8\,\%$ (although this steadily decreases towards $\lesssim 2\%$ near merger) for both codes. Here we have ignored the early part of the signal where the amplitude is dominated by the "junk" radiation up to $u = 300M$.

We next directly compare the results of the two codes by computing the relative difference in the amplitude h_{22}^{A} and the absolute difference in the phase h_{22}^{ϕ} which is shown in Fig. 4.3. Again, ignoring the early part of the signal and the late ringdown, the relative difference in the amplitudes is $\lesssim 1\,\%$, consistent with the individual error estimates from the two codes. The discrepancy in phase remains $\mathcal{O}(10^{-3})$ or smaller throughout the inspiral, merger and early ringdown–well within the error estimates of each code.

For the first asymmetric BH binary configuration, `q2-d10`, we proceed in the same way. We study the convergence in analogy to figure 4.2. Ignoring again the contamination at early times, we obtain third-order convergence in the amplitude and fifth-order convergence in the phase for Lean. For GRChombo, we obtain fourth order convergence in the amplitude and mild overconvergence of about eighth order in the phase[6]. This

[6]We assume fourth order convergence for our GRChombo phase error.

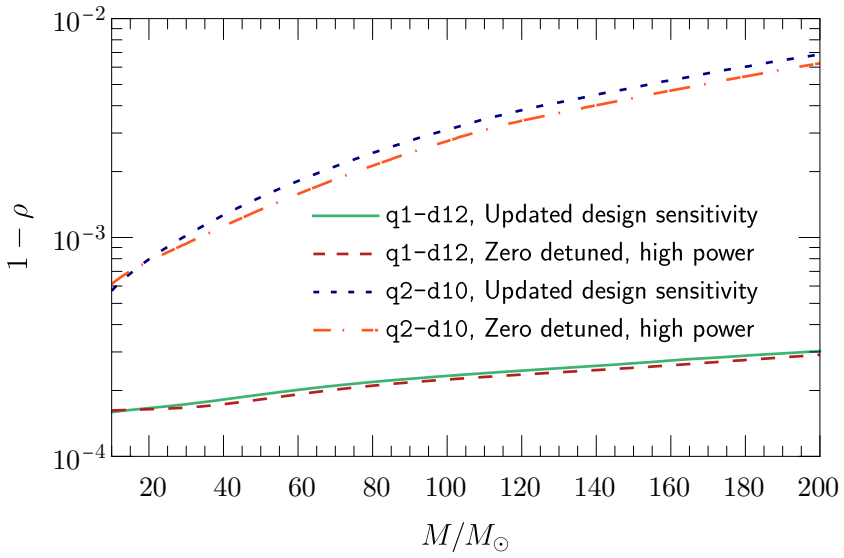
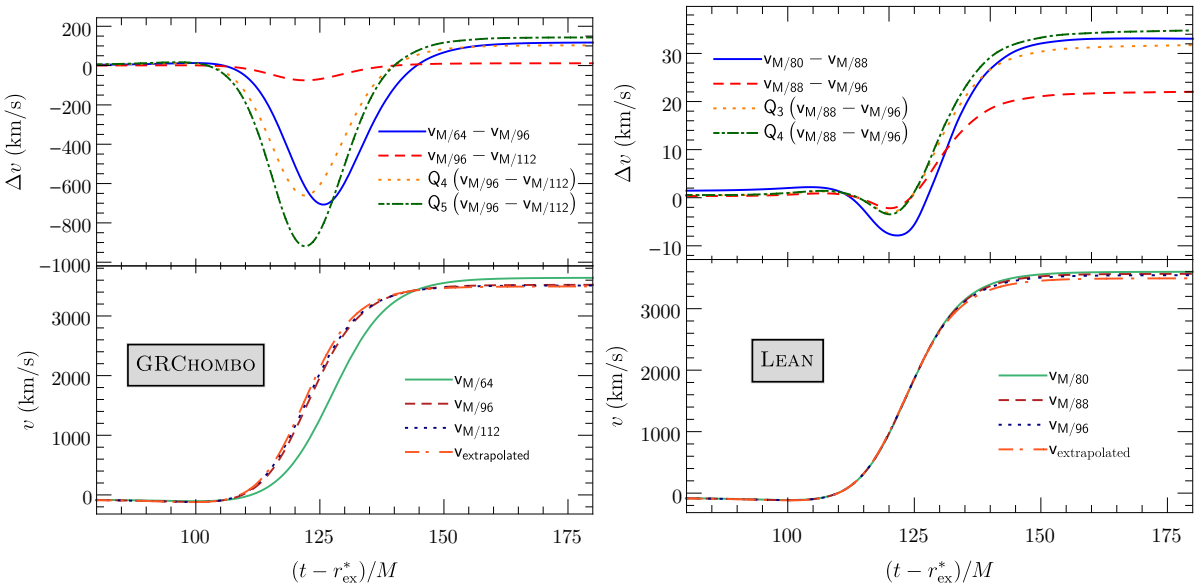


Fig. 4.4 The discrepancy $1-\rho$ between the $(\ell, m) = (2, 2)$ mode of the gravitational wave signal from the `q1-d12` (10 orbits, non-spinning, equal mass) and `q2-d10` (6 orbits, non-spinning, 2 : 1 mass ratio) BH binary configurations simulated with Lean and GRChombo. For `q1-d12`, we use the simulation with resolution $\Delta x_{l_{\max}} = M/128$ for both codes and for `q2-d10`, we use the simulation with resolution $\Delta x_{l_{\max}} = M/96$ for both codes. For each configuration, we show the difference computed with the updated Advanced LIGO sensitivity design curve (`aLIGODesign.txt` in Ref. [227]) and the zero detuned, high power noise curve from the Advanced LIGO anticipated sensitivity curves (`ZERO_DET_high_P.txt` in Ref. [228]).

leads to uncertainty estimates of $\Delta h^A_{22}/h^A_{22} \lesssim 2.5\,\%$ in the amplitude and $\Delta h^\phi_{22} \lesssim 0.25$ in the phase for both codes.

The error due to finite-radius extraction in the amplitudes is $\epsilon_{h^A_{22},86.7M}/h^A_{22} \lesssim 10\,\%$ in the early inspiral decreasing down to $\lesssim 2\,\%$ in the late inspiral, and in the phase is $\epsilon_{h^\phi_{22},86.7M} \sim 0.5$ for both codes.

In Fig. 4.4, we display as a function of the total mass M the discrepancy $1-\rho$ (where ρ is the overlap given by Eq. (4.3.16)) between the GRChombo and Lean waveforms for both `q1-d12` and `q2-d10`, with the spectral noise density $S_n(f)$ given by (i) the updated Advanced LIGO sensitivity design curve (`aLIGODesign.txt` in Ref. [227]) and (ii) the zero detuned, high power noise curve from the Advanced LIGO anticipated sensitivity curves (`ZERO_DET_high_P.txt` in Ref. [228]). For `q1-d12`, the figure demonstrates excellent agreement of the two waveforms for the entire range $M = 10 \ldots 200\,M_\odot$ with a discrepancy $1-\rho \approx 0.03\,\%$ or less, whereas for `q2-d10`, the agreement is not quite as strong but nevertheless demonstrates very good consistency with a discrepancy $1-\rho \approx 0.7\,\%$ or less. The larger difference for `q2-d10` compared to `q1-d12` may be attributed to the slightly



(a) (b)

Fig. 4.5 Convergence plots for the accumulated linear momentum radiated from configuration `q1-s09` for GRChombo with finest grid resolutions $\Delta x_{l_{\max}} = M/64$, $M/96$ and $M/112$ (a) and for Lean with finest grid resolutions $\Delta x_{l_{\max}} = M/80$, $M/88$ and $M/96$ (b). This is shown in the form of the BH recoil velocity in the bottom panels. For both codes, the radiated linear momentum is calculated from the extracted Ψ_4 values at $r_{ex} = 90M$, and the extrapolated curve corresponds to a Richardson extrapolation assuming fourth order convergence. In the top panels, we show the difference between the results from different resolutions with rescalings according to third and fourth order convergence for Lean and according to fourth and fifth order convergence for GRChombo.

lower resolutions employed for this configuration, especially near the smaller BH. To put these numbers into context, Lindblom, Owen, and Brown [226] estimate that a mismatch of 3.5 % would result in a reduction in the GW event detection rate by about 10 %.

Our final BH binary features asymmetry in the form of non-zero spins. This time, we focus on the BH recoil velocity v calculated from the linear momentum radiated in GWs, and the analysis is shown in Fig. 4.5. From the plots, we can see that Lean exhibits convergence between third and fourth order, whilst GRChombo exhibits convergence between fourth and fifth order. We illustrate our estimate of the total error for each code in Fig. 4.6. Here, the error bands—around the curve from the highest resolution simulation in each case—correspond to the difference with the Richardson extrapolated curve assuming fourth-order convergence plus the estimated error due to finite-radius extraction (about 1.5 % + 2 % for Lean and 0.5 % + 3 % for GRChombo). This total error is about 3.5 % for both codes.

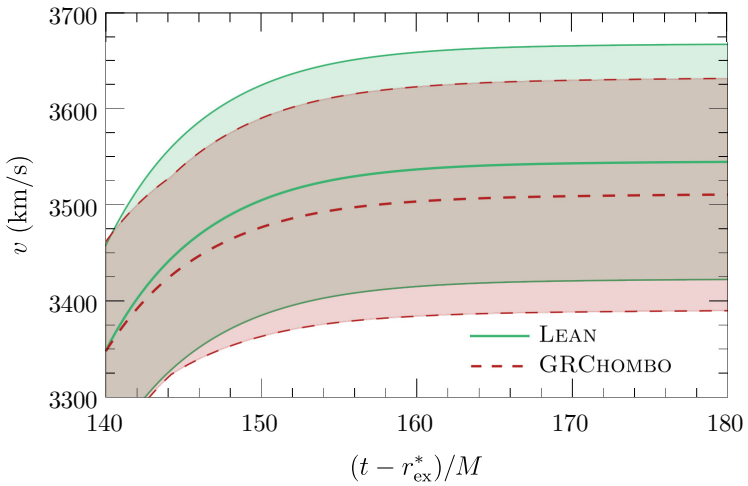


Fig. 4.6 The accumulated radiated linear momentum at the end of the highest resolution simulations of configuration `q1-s09` from each code. The linear momentum is shown in the form of the BH recoil velocity and the error bands show our estimate of the total error coming from both discretization and finite-radius effects.

4.4 Comparing tagging criteria using axion stars

In order to demonstrate the application of our techniques to problems with matter fields and dynamically varying length scales, we consider the evolution of a single axion star—a compact object composed of a real scalar bosonic field. We analyse the evolution of a star that is stable on the timescale of the simulation, as well as one in which the self-interaction is increased in order to trigger gravitational collapse to a BH. As discussed previously, this simple example tests many of the key requirements in using AMR to evolve fundamental fields coupled to gravity, in particular, the ability to follow stable oscillations and to adapt to changing dynamical timescales. Similar considerations apply, for example, to cosmological spacetimes, cosmic strings and collisions of exotic compact objects.

We demonstrate the use of two effective tagging criteria; first, tagging by the magnitude of terms in the Hamiltonian constraint (4.2.5), and second, by the numerical truncation error between refinement levels (4.2.4).

4.4.1 Methods

4.4.1.1 Setup

We consider the evolution of two different axion star configurations. Axion stars are quasi-equilibrium configurations of a self-gravitating real scalar field ϕ [229] that is subject

Table 4.4 GRChombo grid parameters for axion star configurations using different tagging criteria. There are $(l_{\max}+1)$ refinement levels and the coarsest level has length (without symmetries applied) L. The grid spacing on the coarsest level is Δx_0 and the minimum number of cells in the buffer regions between consecutive refinement level boundaries is n_B. The regrid thresholds for the different tagging criteria are given by τ_R. We consider two cases; a stable axion star ($f_a = 1$) and an unstable collapse to a BH ($f_a = 0.05$), with $\mu = m_a c/\hbar = 1$ in code units.

f_a	Tagging	$l_{\max}$	μL	$\mu\Delta x_0$	n_B	τ_R
1.0	Ham	3	512	4	8	0.1
1.0	Ham	3	512	2	16	0.05
1.0	Ham	3	512	1	32	0.025
1.0	Trunc	3	512	4	8	0.0625
1.0	Trunc	3	512	2	16	$3.91\,\mathrm{x}\,10^{-3}$
1.0	Trunc	3	512	1	32	$2.44\,\mathrm{x}\,10^{-4}$
0.05	Ham	8	1024	4	8	0.1
0.05	Ham	8	1024	2.67	12	0.067
0.05	Trunc	8	1024	4	8	0.001
0.05	Trunc	8	1024	2.67	12	$2.96\,\mathrm{x}\,10^{-4}$

to a periodic self-interaction potential $V(\phi)$. A canonical potential is

$$V(\phi) = \mu^2 f_a^2 [1 - \cos(\phi/f_a)]\,, \tag{4.4.1}$$

where this form arises as a result of the spontaneously broken $U(1)$ Peccei-Quinn symmetry and subsequent "tilting" of the potential due to instanton effects [230, 231]. The decay constant f_a quantifies the symmetry breaking scale and determines the strength of the scalar field self-interactions (their strength for a given central amplitude is inversely related to f_a) and $\mu = m_a c/\hbar$ is an inverse length scale related to the scalar mass[7] m_a. Axion stars on the main stability branch are characterized by their central amplitude ϕ_0 or equivalently their ADM mass $M_{\rm ADM} \sim \mu^{-1}$. They have a physical size R (defined as the radius containing 99 % of the total mass) that is approximately inversely related to their ADM mass, and thus a useful descriptor is their compactness $\mathcal{C} = M_{\rm ADM}/R$. Axion stars with $\mathcal{C} \sim 1/2$ are highly relativistic and may form BHs if they collapse or collide. For $m_a \sim 10^{-14}$ eV, they are of a mass and size comparable to solar mass BHs, and thus potentially of astrophysical interest. Further details related to the setup used here can be found in [232, 140], and a useful general review of axion physics is provided in [233].

[7] In Planck units one can write $\mu = m_a/M_{\rm Pl}^2$, where $M_{\rm Pl}$ is the Planck mass.

The equation of motion for the scalar field ϕ is given by the Klein-Gordon equation for a real scalar field minimally coupled to gravity

$$\nabla_\mu \nabla^\mu \phi - \frac{dV}{d\phi} = 0\,, \tag{4.4.2}$$

and the system is completed with the Z4 equations (2.1.43a) for the metric components. To construct localised, quasi-equilibrium oscillatory (axion star) solutions, we solve the Einstein-Klein-Gordon (EKG) system of equations with a harmonic field ansatz and appropriate boundary conditions [234, 229].

Unlike the case of complex scalar boson stars, for axion star solutions the metric components $g_{\mu\nu}$ also oscillate in time, with energy being transferred between the matter and curvature terms in the Hamiltonian constraint [140]. This makes them challenging targets for dynamical refinement; simple criteria based solely on matter field gradients will fail to achieve a stable grid structure, as the gradients change over time even in the quasi-stable case. If the gradients are close in value to a tagging threshold, frequent regridding will occur, which introduces errors.

The stability of the axion star solution comes from the balance of its tendency to disperse due to gradient pressure from spatial field derivatives, with the tendency to collapse due to its energy density and attractive self interactions. The relative strengths of these effects determines whether the axion star remains stable, disperses through scalar radiation or collapses to a black hole when perturbed. In particular, if the self-interaction scale f_a is too low, this can cause the axion star to collapse to a BH [235]. In the last case, a key AMR challenge is determining tagging criteria that progressively track the axion star collapse without triggering too frequent regridding from the more rapid field oscillations.

We consider two cases, both with central amplitude $\phi_0 = 0.020$ and $\mu M_{\rm ADM} = 0.4131$:

(i) An axion star with weak self-interactions ($f_a = 1$)[8], where the scalar field and metric oscillate over time in a localized configuration that is stable over time periods much longer than that of the simulation. We would ideally like the refinement to remain constant, despite the oscillations of the fields.

(ii) An unstable configuration where we increase the attractive self interaction by reducing the self interaction scale to $f_a = 0.05$, such that the axion star is destabilized

[8]The decay constant f_a is dimensionless in geometric units as used here; to obtain its value in Planck units one simply multiplies by the Planck mass $M_{\rm Pl}$.

and undergoes collapse to a BH. We need the mesh refinement to follow this process sufficiently rapidly, but without excessive regridding.

To evolve this system in GRChombo, the EKG equation (4.4.2) is decomposed into two first order equations in the 3+1 formulation,

$$\partial_t \phi = \beta^i \partial_i \phi + \alpha \Pi, \tag{4.4.3}$$

$$\partial_t \Pi = \beta^i \partial_i \Pi + \alpha \gamma^{ij} (\partial_i \partial_j \phi + \partial_i \phi \partial_j \alpha) + \alpha \left(K\Pi - \gamma^{ij} \Gamma^k_{ij} \partial_k \phi - \frac{dV}{d\phi} \right), \tag{4.4.4}$$

and added to the CCZ4 evolution scheme (2.3.30). The initial data are set up as in the previous study [235] using the numerically obtained oscillaton profile for an $m^2\phi^2$ potential [229, 234, 236]. We choose the initial hypersurface such that $\phi = 0$ and hence $V(\phi) = 0$ everywhere. The Hamiltonian constraint is thereby satisfied for both the $V = m^2\phi^2$ and the axion potential (4.4.1) cases. Furthermore, if we impose the extrinsic curvature $K_{ij} = 0$, the momentum constraint is trivially satisfied and all the dynamical information is encoded in the kinetic term of the field Π. The system is evolved in the moving puncture gauge using the default 1+log parameters (2.3.40) for the lapse evolution equation (2.3.39) with the exception of $a_1 = 0$; and $b_1 = 0$, $b_2 = 3/4$ and $\eta = \mu$ for the shift evolution (2.3.41).

4.4.1.2 Tagging criteria

We demonstrate the suitability of two different tagging methods for tracking the axion star evolution in both the stable and unstable cases. In the case of collapse to a BH, for both tagging methods, the threshold τ_R must be chosen such that the apparent horizon is covered entirely by the finest refinement level, as discussed in Sec. 4.2.2. In our case, we choose a maximum refinement level $l_{\max} = 8$, which covers up to $\mu r \geq 0.5 > \mu M_{\rm ADM}$.

The first tagging criterion we consider is based on physical quantities in the simulation, as outlined in Sec. 4.2.5. We choose the absolute sum of the terms in the Hamiltonian constraint $\mathcal{H}_{\rm abs}$ (4.2.5), setting the criterion

$$C(\mathbf{i}) \equiv \mathcal{H}_{\rm abs} \tag{4.4.5}$$

in the tagging indicator function (3.1.5).

We also show the efficacy of truncation error tagging, outlined in Sec. 4.2.4. We choose the variables f as defined in Eq. (4.2.3) to be χ, K, ϕ and π, as these capture

the information in the Hamiltonian constraint $\mathcal{H}$. We use the tagging criterion (4.2.4), explicitly

$$C(\mathbf{i}) = \sqrt{(\tau_{l,\chi}(\mathbf{i}))^2 + (\tau_{l,K}(\mathbf{i}))^2 + (\tau_{l,\phi}(\mathbf{i}))^2 + (\tau_{l,\pi}(\mathbf{i}))^2}\,, \tag{4.4.6}$$

where $\tau_{l,f}(\mathbf{i})$ is defined by Eq. (4.2.3) and we have set the normalizing factor $L_f = 1$ for all f.

4.4.1.3 Diagnostics and convergence testing

We perform convergence tests using several key physical quantities from the evolution. The first quantity is the L^2 norm of the Hamiltonian constraint violations (2.1.36) $||\mathcal{H}||_2$ over a coordinate volume $\mathcal{V}$

$$||\mathcal{H}||_2 = \sqrt{\int_{\mathcal{V}} \mathrm{d}^3x\, \mathcal{H}^2}\,. \tag{4.4.7}$$

For the case of a stable axion star, in order to exclude the constraint violation at the outer boundaries, we choose $\mathcal{V}$ to be the volume enclosed by a sphere of fixed coordinate radius $r_{\rm out}$, $B_{r_{\rm out}}$, with a centre that coincides with that of the star:

$$\mathcal{V} = B_{r_{\rm out}}. \tag{4.4.8}$$

In the case of an unstable axion star that collapses to a BH, we furthermore excise the volume enclosed by a smaller sphere of fixed coordinate radius $r_{\rm in} < r_{\rm out}$ with the same centre in order to exclude the constraint violations near the puncture that arises after the collapse. The radius $r_{\rm in}$ is chosen such that the sphere will lie within the apparent horizon once it is formed. This means that

$$\mathcal{V} = B_{r_{\rm out}} \setminus B_{r_{\rm in}}. \tag{4.4.9}$$

For the stable axion star, we also test convergence using the total mass of the matter content $M_{\rm mat}$ within the same volume

$$M_{\rm mat} = -\int_{\mathcal{V}} \mathrm{d}^3x\, \sqrt{\gamma}\, T^0_0 = \int_{\mathcal{V}} \mathrm{d}^3x\, \sqrt{\gamma}\left(\alpha\rho - \beta_k j^k\right), \tag{4.4.10}$$

where ρ and j^i are defined by Eqs. (2.1.32) and (2.1.33), $\gamma = \det(\gamma_{ij})$ and α and β_i are the lapse and shift as defined by Eqs. (2.1.4) and (2.1.9). Further details can be found in Ref. [140].

4.4.2 Results

We have performed simulations of axion stars at different resolutions for each configuration in Table 4.4. We evolve with GRChombo and investigate two different tagging criteria: refinement using (i) the Hamiltonian constraint $\mathcal{H}_{\rm abs}$ (4.4.5) and (ii) the truncation error of the variables χ, K, ϕ and π as defined by (4.4.6).

As outlined in Sec. 4.2.5, for the stable axion star configuration ($f_a = 1.0$), we expect Hamiltonian constraint tagging to generate stable refinement levels. We set τ_R such that we obtain an appropriate initial grid structure, which we choose to have a maximum refinement level $l_{\rm max} = 3$ with refinement concentrated on the axion star.

The middle panel of Fig. 4.7 shows $M_{\rm mat}$ for the finest grid configuration for a stable axion star in table 4.4, where $M_{\rm mat}$ is calculated as defined in Eq. (4.4.10) for a coordinate sphere with radius $\mu r_{\rm out} = 25$. We observe some initial gauge evolution of $M_{\rm mat}$ due to the transition from the initial polar-areal gauge to puncture gauge, with subsequent regular oscillations over time, which are physical. Given that $M_{\rm mat}$ includes only matter contributions and the ADM mass of an axion star is approximately constant (i.e. the total flux out of the outer spherical surface is zero), these oscillations indicate the transfer of energy between curvature and matter terms as discussed in Sec. 4.4.1.1.

The upper panel of Fig. 4.7 shows the difference in mass $\Delta M_{\rm mat}$ for the simulations in table 4.4. This demonstrates convergence between third and fourth order in $M_{\rm mat}$ at late times, but second order near the beginning of the simulation. This agrees with expectations; the initial data used is accurate to second order, and this error dominates at early times, with the fourth order convergence related to the evolution scheme only being recovered at later times. Some error is also introduced by the interpolation of the initial conditions onto the grid, which is first order (but with a high spatial resolution in the numerical solution so this is subdominant). By comparing the highest resolution simulation with a Richardson extrapolation, we obtain a discretization error estimate of $\Delta M_{\rm mat}/M_{\rm mat} \lesssim 4 \times 10^{-5}$ at late times (using third order extrapolation).

The lower panel of Fig. 4.7 shows the L^2 norm of the Hamiltonian constraint violations $||\mathcal{H}||_2$ for the same simulations. Again, we measure between third and fourth order convergence at late times, and between first and second order initially. We obtain an error measure at late times of $\sqrt{M_{\rm ADM}}\,\Delta||\mathcal{H}||_2/16\pi \lesssim 8\times10^{-8}$, where we have normalized with the ADM mass to create a dimensionless measure of the spurious energy density.

We perform the same analysis of $M_{\rm mat}$ and $||\mathcal{H}||_2$ for the stable axion star using truncation error tagging with the parameters in table Table 4.4. We obtain a very similar grid structure and evolution behaviour to Hamiltonian tagging, with the same convergence

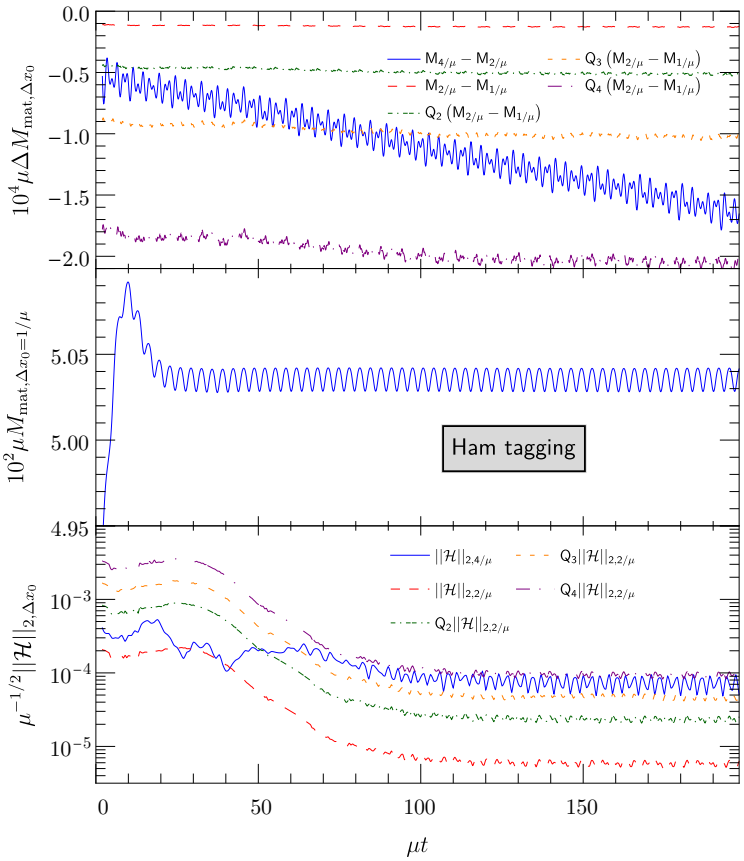


Fig. 4.7 Convergence plots for the stable axion star configuration ($f_a = 1.0$) with Hamiltonian constraint tagging (4.4.5) and grid configurations given in table 4.4. The top panel shows the difference in the calculated matter mass M_{mat} (4.4.10) within a sphere of radius $\mu r_{\mathrm{out}} = 25$ between the resolutions with rescalings according to second, third and fourth order convergence. For reference, in the middle panel, we show M_{mat} for the highest resolution simulation (with $\mu\Delta x_0 = 1$). In the bottom panel, we plot the L^2 norm of the Hamiltonian constraint $||\mathcal{H}||_2$ (4.4.7) for the two lower resolution simulations ($\mu\Delta x_0 = 4, 2$) in addition to rescalings according to second, third and fourth order convergence. We omit the corresponding plots for the simulations with truncation error tagging (4.4.6) as they are qualitatively very similar.

and error estimates, demonstrating that both methods can achieve equivalent, accurate results.

For the unstable axion star configuration ($f_a = 0.05$), we perform convergence testing on $||\mathcal{H}||_2$ within the spatial volume with $0.5 < \mu r < 25$, excising the region $\mu r < 0.5$ where a BH is formed. We use the parameters given in Table 4.4. We do not perform a convergence test of M_{mat}, as the quantity oscillates with a high frequency about a

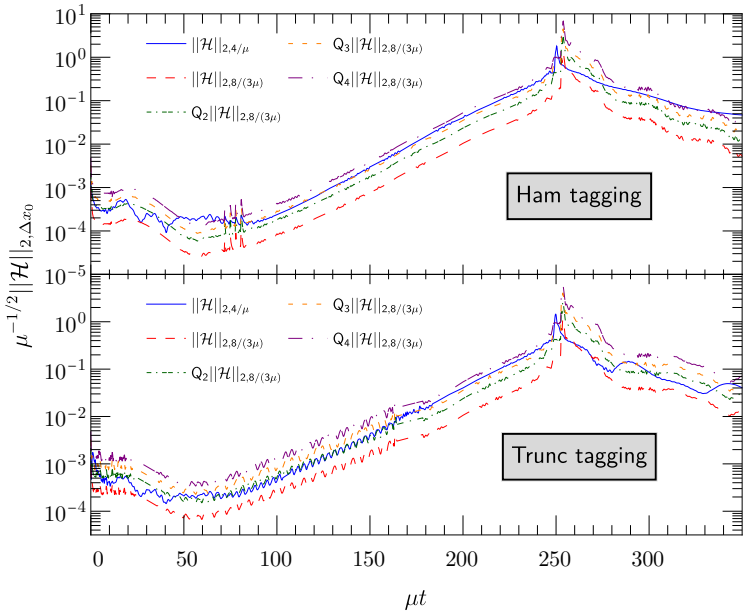


Fig. 4.8 Convergence plots of the L^2 norm of the Hamiltonian constraint (4.4.7) contained between spheres of radius $\mu r_{\rm in} = 0.5$ and $\mu r_{\rm out} = 25$ for the unstable axion star configuration ($f_a = 0.05$) with both Hamiltonian constraint tagging (4.4.5) (top panel) and truncation error tagging (4.4.6) (bottom panel). The grid configurations are provided in Table 4.4 and we additionally plot rescalings according to second, third and fourth order convergence.

mean value that rapidly decreases to zero around the collapse, making such an analysis impractical.

The top panel of Fig. 4.8 shows the convergence analysis of $||\mathcal{H}||_2$ with Hamiltonian constraint tagging. We observe approximately third order convergence prior to collapse to a BH, then approximately fourth order convergence at late time. We obtain a maximum error measure on the finer grid of $\sqrt{M_{\rm ADM}}\Delta||\mathcal{H}||_2/16\pi \lesssim 0.021$, with this value occurring approximately at the collapse.

For the same configuration with truncation error tagging in the lower panel of Fig. 4.8, we obtain similar convergence prior to the collapse with a maximum error measure $\sqrt{M_{\rm ADM}}\Delta||\mathcal{H}||_2/16\pi \lesssim 0.015$. The convergence after the collapse is lower: between first and third order. In general, we see that the rapid regridding that occurs during a collapse (often triggered at different times at the different resolutions) introduces errors which can reduce the convergence order. This illustrates one of the main challenges of AMR, which is to obtain good convergence in highly dynamical regimes.

4.5 Conclusion

In this chapter, we have presented a detailed discussion of the use of fully adaptive mesh refinement (AMR) in numerical relativity simulations with the GRChombo code. To avoid confusion, we first summarize how the term "fully adaptive" is meant to distinguish AMR from the common (and often highly successful) box-in-a-box approach. This distinction consists in two main features. First, we use the term AMR in the sense that it allows for refined regions of essentially arbitrary shape. Second, it identifies regions for refinement based on a point-by-point interpretation of one or more user-specifiable functions of grid variables. Of course, a region of arbitrary shape will inevitably be approximated by a large number of boxes on Cartesian grids; the distinction from a box-in-a-box approach therefore consists in the *large number* of boxes used in AMR. Likewise, every box-in-a-box approach will ultimately base its dynamic regridding on some function of the evolved grid variables, as the apparent horizon. The key feature of AMR is the *pointwise* evaluation of grid variables or their derived quantities.

The advantages of AMR based simulations over the simpler box-in-a-box structure evidently arise from its capability to flexibly adapt to essentially any changes in the shape or structure of the physical system under consideration. These advantages, however, do not come without new challenges; the identification of these challenges and the development of tools to overcome them are the main result of this chapter.

The first and most elementary result of our study is the (hardly surprising) observation that there exist no "one size fits all" criterion for refinement that automatically handles all possible physical systems. Many of the challenges, however, can be effectively addressed with a combination of a small number of criteria for tagging and refining regions of the domain. We summarize these challenges and techniques as follows.

(i) With AMR, it is more difficult to test for (and obtain) convergence, because of the loss of direct control over the resolution in a given region of spacetime. While the refinement in AMR is every bit as deterministic as it is in a box-in-a-box approach, the complexity of the underlying algorithm makes it practically impossible for a user to predict when, if and where refinement will take place. Consider for example the convergence analysis of a simulation using the truncation-error based tagging criterion of (4.2.3); in some regions of the spacetime a low resolution run may encounter a sufficiently large truncation error to trigger refinement whereas a higher-resolution run will not. To counteract this effect, one may adjust the tagging threshold in anticipation of the reduction in the truncation error, but some

experimentation is often necessary because different ingredients of the code have different orders of accuracy. Additionally, one may enforce refinement using a priori knowledge, as for example, through enforced tagging around the spheres of wave extraction. An alternative approach would be to record the grid structure over time for one simulation (e.g. the lowest resolution run) and then "replay" this grid structure (or as close as possible to it) for simulations at different resolutions as is done for the Had code [237].

(ii) A further challenge arises from the use of too many refinement regions/boundaries over a small volume in spacetime. The interpolation at refinement boundaries is prone to generating small levels of high-frequency numerical noise that may bounce off neighbouring boundaries if these are too close in space (or time). An effective way to handle this problem is the use of buffer zones in space and to avoid unnecessarily frequent regridding.

(iii) In the case of BH simulations, we often observe a degradation of numerical accuracy when refinement levels cross or even exist close to the apparent horizon. This typically manifests itself as an unphysical drift in the horizon area and, in the case of binaries, a loss of phase accuracy and/or a drift in the BH trajectory. These problems can be cured by enforced tagging of all grid points inside the apparent horizon. In practice, we add an additional buffer zone to ensure all refinement boundaries are sufficiently far away from the apparent horizon(s).

(iv) The Berger-Rigoutsos algorithm detailed in Sec. 3.1.4 does not treat the x, y and z direction on exactly equal footing; the partitioning algorithm (cf. Fig. 3.3) inevitably handles the coordinate directions in a specific order. This can lead to asymmetries in the refined grids even when the underlying spacetime region is symmetric. In some simulations of BHs, we noticed this to cause a loss in accuracy. A simple way to overcome this problem is to enforce a boxlike structure around BHs.

(v) A single tagging variable (or one of its spatial derivatives) may not always be suitable to achieve appropriate refinement throughout the course of an entire simulation; for example, this may be due to gauge dependence or dramatic changes in the dynamics of the physical evolution. GRChombo allows for tagging regions based on arbitrary functions of *multiple* variables and their derivatives to overcome problems of this kind.

In order to avoid the difficulties listed here, we often combine two or more tagging criteria. The efficacy of this approach is demonstrated in Secs. 4.3 and 4.4 where we present in detail AMR simulations of inspiralling BH binaries and stable as well as collapsing oscillatons. By comparing the BH simulations with those from the box-in-a-box based Lean code, we demonstrate that with an appropriate choice of tagging criteria, AMR simulations reach the same accuracy and convergence as state-of-the-art BH binary codes using Cartesian grids. While AMR does not directly bestow major benefits on the modelling of vacuum BH binaries, it offers greater flexibility in generalizing these to BH spacetimes with scalar or vector fields, other forms of matter, or BHs of nearly fractal shape that can form in higher-dimensional collisions [179]. The simulations of rapidly oscillating or gravitationally collapsing scalar fields demonstrate GRChombo's capacity to evolve highly compact and dynamic matter configurations of this type.

Finally, we note the potential of AMR for hydrodynamic simulations. However, high-resolution shock capturing methods present qualitatively new challenges for AMR, and we leave the investigation of this topic for future work.

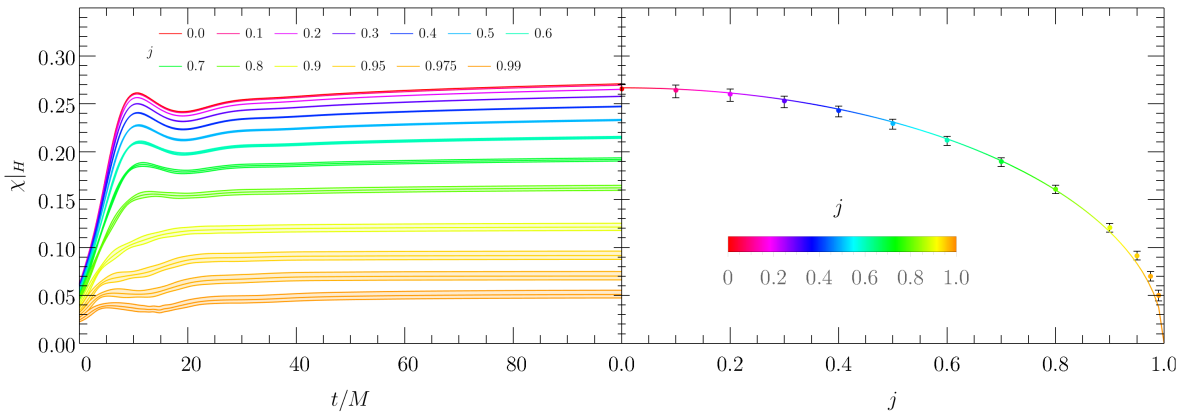


Fig. 4.9 Plots illustrating the dependence of the value of the conformal factor χ on the apparent horizon surface H in the moving puncture gauge (2.3.39, 2.3.41) for different values of the dimensionless spin j. For all plots, we use the quasi-isotropic Kerr initial data [238] and the default values of the lapse parameters (2.3.40) along with the shift parameters $b_1 = 0$, $b_2 = 3/4$ and $M\eta = 1$. Although we would expect the plots to vary for different gauge parameters (in particular, as η is varied), these plots provide a rough rule-of-thumb. The left panel shows the mean value of χ as a function of time with the error bands around each curve corresponding to the maximum and minimum on H. The right panel shows the mean value of χ over the interval $t/M \in [40, 100]$ for each j with the error bars corresponding to the minimum and maximum values of χ over the same interval. Furthermore, we show a fit of the mean value of χ against j which takes the form $\langle\chi\rangle|_H \simeq 0.2666\sqrt{1-j^2}$.

Appendix 4.A Approximate horizon locations as a tagging criteria

As noted in Sec. 4.2, the use of the horizon location can be an essential part of an adaptive mesh scheme for refinement. In particular, one does not usually want to put additional refinement within a horizon (where effects are unobservable anyway), and should take care to avoid grid boundaries overlapping the horizon, since this can lead to instabilities that strongly affect the physical results. Whilst using an apparent horizon finder for this is a possibility, often a more "quick and dirty" scheme using contours of the conformal factor can be just as effective, and significantly easier to implement. Whilst in principle there is a dependence on simulation and gauge parameters (in particular η [103]), in general the approximate values are quite robust.

The key dependence is on the [dimensionless] spin of the black hole j, as illustrated in Fig. 4.9, with a good fit obtained from the relation

$$\langle\chi\rangle|_H = 0.2666\sqrt{1-j^2}. \tag{4.A.1}$$

One key advantage is that one does not need to know a priori the mass of the BH spacetime which forms, and it can be seen that simply using the $j = 0$ values will give a conservative coverage of the horizon. These types of criteria were used extensively in the higher dimensional black ring spacetimes studies in Refs. [179, 180, 181, 182] and for the investigation into gravitational collapse in a modified gravity theory in Ref. [178].

Chapter 5

Gravitational recoil from eccentric binary black-hole mergers

The radiation of linear momentum imparts a recoil (or "kick") to the centre of mass of a merging black-hole binary system. In this chapter we present an investigation of the impact of nonzero eccentricity on the kick magnitude and gravitational-wave emission of nonspinning, unequal-mass black hole binaries. This work also demonstrates application of the methods discussed in Chapter 4.

This chapter is based on the coauthored publication Ref. [2] in collaboration with U. Sperhake, E. Berti and R. Croft. I performed the vast majority of the GRChombo simulations presented in this chapter as well as their analysis. I also implemented the calculation of the radiated momentum from the sums of overlaps of multipolar amplitude for all simulations.

5.1 Introduction

Gravitational waves (GWs) carry energy, angular momentum and linear momentum away from the source with potentially observable consequences. The radiated energy corresponds to an often enormous mass deficit in the source; for example the first ever detected black-hole (BH) binary merger, GW150914 [28], radiated $\Delta M \approx 3\,M_\odot$, or about 4.6 % of the total mass of the source. A tiny fraction of this energy is deposited into GW interferometers, thus enabling us to detect and characterise the signal [19]. The angular momentum radiated in GWs reduces the rotation rate of possible merger remnants and—at least in four spacetime dimensions—plays a critical role in avoiding the formation of naked singularities in the form of BHs spinning above the Kerr limit;

see e.g. Refs. [239, 240]. Therefore, GW emission is a necessary ingredient of the theory of general relativity, in the sense that it avoids the formation of spacetime singularities and preserves its predictive power.

In this chapter, we focus on the radiated linear momentum, which imparts a recoil (commonly referred to as a *kick*) on the centre of mass of the emitting system [241, 242, 243].

Whereas GWs inevitably carry energy and angular momentum—provided their sources do—the radiation of linear momentum requires some degree of asymmetry, as realised in nonspherical supernova explosions or unequal-mass ratios and/or spin misalignments in binary BH mergers. The inspiral of two equal-mass, nonspinning BHs, for example, radiates energy and angular momentum, whereas the emitted linear momentum is zero by symmetry. By turning these considerations around, we may also regard the study of recoiling GW emitters as a guided search for characteristic (in some loose sense "asymmetric") features in their orbital dynamics which, in turn, might help us to better understand astrophysical sources through GW observations. A recoiling postmerger BH, for example, can induce a blue (or red) shift in parts of its GW signal that may be exploited in future GW observations to directly measure BH kicks [244, 245, 246], and the effect of kicks should be taken into account in future ringdown tests of general relativity with third-generation GW detectors to avoid systematic biases [247]. The asymmetric emission of GWs is not the only mechanism that can contribute to recoils; if there is an accretion disk or some other astrophysical background, this can also impart a kick on the remnant BH that can be $\mathcal{O}(100)$ km/s [248].

For binary BH mergers, early estimates of the recoil speeds of the remnant BH relied on a variety of approximations, including post-Newtonian (PN) theory [249, 250], BH perturbation theory [251], the effective-one-body formalism [252], the close-limit approximation [253, 254], and combinations thereof [255]. Not long afterwards, during the numerical relativity (NR) gold rush, several groups obtained more accurate results for the kick velocity from the merger of nonspinning BHs along quasicircular orbits [256, 257, 160]. These calculations were followed by the discovery that the merger of spinning BHs can lead to kick velocities of ~ 3000 km/s when the spins lie in the orbital plane and point in opposite directions ("superkick" configurations [218, 219, 220]), and to even larger kicks of order ~ 5000 km/s when the spins are partially aligned with the orbital angular momentum ("hang-up kick" configurations [258]). The probability of such large recoils occurring in nature depends therefore on spin alignment, and this has been studied by several authors (see, e.g., Refs. [259, 260, 261, 262, 263, 262]).

The possible occurrence of superkicks has important consequences for astrophysical BHs and their environments [264, 265, 266, 267]. It is pertinent to compare the recoil velocities obtained from NR simulations with the escape velocities of various astrophysical environments [268]. For example, stellar-mass BH binaries are believed to form dynamically in globular clusters [269]. In this case the escape velocities are generally $\mathcal{O}(10)$ km/s, smaller than the $\mathcal{O}(100)$ km/s kicks predicted for quasicircular, nonspinning binaries [257]. Then relativistic recoils can affect the proportion of BH merger remnants that are retained by globular clusters even if the BHs are nonspinning [270]. At the other end of the scale, the recoil velocities of supermassive BHs can be used to constrain theories of their growth at the centre of dark matter halos [271]. Kicked remnants in the accretion disk of an active galactic nucleus may also lead to detectable electromagnetic counterparts for stellar-origin BH mergers [272, 273].

As mentioned above, a net gravitational recoil requires some asymmetry in the system, so that the GW emission is anisotropic. A natural way to accentuate the asymmetry is through the addition of orbital eccentricity. Early calculations in the close-limit approximation [254] predicted a kick proportional to $1+e$ for small eccentricities, $e \lesssim 0.1$. More recently, numerical relativity calculations led to the conclusion that eccentricity can lead to an approximate 25% increase in recoil velocities for superkick configurations with moderate eccentricities [217].

The main goal of this study is to investigate the impact of nonzero eccentricity on the kick magnitude and the corresponding GW emission of nonspinning, unequal-mass BH binaries. As we shall see, the eccentricity has a subtle but significant effect on the kick magnitude, which manifests itself in corresponding patterns in the GW signal, especially in subdominant multipoles.

For isolated binary systems with large initial separations, the emission of GWs acts to circularise the orbit by the time the signal enters the frequency band of ground-based detectors. However, viable dynamical formation channels of stellar-origin BH binaries could result in a non-negligible population of merging BHs that still retain moderate eccentricities at frequencies relevant for ground-based GW detection (see, for example, Refs. [274, 275, 276, 277, 278, 279]). Furthermore, the presence of astrophysical media such as accretion disks may increase the eccentricity during the inspiral [280]. Most of the events observed by the LIGO/Virgo Collaboration show no evidence of significant eccentricities [281] but the extraordinary GW190521 event [282] is potentially consistent with an eccentricity as high as $e \approx 0.7$ [283, 284].

Orbital eccentricity is expected to be a distinguishing feature of stellar-origin BH binaries that form dynamically, but a nonzero eccentricity is more likely at the low frequencies accessible by LISA, where gravitational radiation reaction has less time to circularise the binary [285, 286, 287]. If confirmed, a nonzero eccentricity would hint at a possible dynamical origin for this event [283].

Eccentricity is expected to play an even more prominent role for massive BH binaries: the dynamics of these binaries in stellar and gaseous environments is expected to lead to distinct (but generically nonzero) orbital eccentricities by the time the binaries enter the LISA sensitivity window (see Ref. [288] and references therein). Even larger eccentricities are possible if BH binary coalescence occurs through the interaction with a third BH [289].

Our work is an exploration of the effect of large eccentricities near merger, and it differs in several ways from the catalogue of eccentric, unequal-mass simulations presented in Ref. [290]. While their study considered a larger range of mass ratios (in our notation, $1/10 \leq q \leq 1$), they carried out fewer simulations for each value of q. The binaries in their simulations have initial eccentricities smaller than $e_0 = 0.18$ 15 cycles before merger, and since they start at larger orbital separations, their eccentricity will have further decreased by the time of merger. As we will see below, the larger initial eccentricities in our simulations allow us to highlight interesting periodicities in the emission of gravitational radiation and the behaviour of the recoil velocity.

The remainder of this chapter is organised as follows. In Sec. 5.2 we discuss our two numerical codes (GRChombo and Lean), the computational framework, and the catalogue of simulations we produced for this study. In Sec. 5.3 we present the main results of our simulations. In Sec. 5.4 we summarise these results and point out possible directions for future work. In Appendix 5.A we detail our tests for numerical accuracy and verify that our two codes give comparable results. Finally, in Appendix 5.B we discuss the tagging criterion used in GRChombo. Throughout this chapter we use geometrical units ($G = c = 1$).

5.2 Computational framework and set of simulations

5.2.1 Numerical methods

The simulations reported in this work have been performed with the GRChombo (see Chapter 3) and Lean [161] codes. We estimate the error budget of our simulations from both codes to be up to 3.5 %. Details of our convergence analyses are provided in

Appendix 5.A. Though different codes were used for each sequence of configurations, we undertook comparison tests in order to ensure consistent results, and these can also be found in Appendix 5.A.

5.2.1.1 GRChombo setup

For all GRChombo simulations we use sixth-order spatial stencils in order to improve phase accuracy (see Fig. 3.2). We evolve using the CCZ4 formulation (see Sec. 2.3.3) and the moving-puncture gauge (see Sec. 2.3.4) with the replacement $\kappa_1 \to \kappa_1/\alpha$ (2.3.33), in order to stably evolve BHs and maintain spatial covariance. After this replacement, we use the constraint damping parameters $\kappa_1 = 0.1$, $\kappa_2 = 0$ and $\kappa_3 = 1$ in all simulations. The regridding is controlled by the tagging of cells for refinement in the Berger-Rigoutsos algorithm (see Sec. 3.1.4), with cells being tagged if the tagging criterion C exceeds a specified threshold value τ_R. Details of the tagging criterion used in this work are provided in Appendix 5.B. We take a CFL factor of 1/4 in all simulations. Due to the inherent symmetry of the configurations considered, we employ bitant symmetry in order to reduce the computational expense.

5.2.1.2 Lean setup

The Lean code [161] is based on the Cactus computational toolkit [157] and uses the method of lines with fourth-order Runge-Kutta time stepping and sixth-order spatial stencils for improved phase accuracy [129]. The Einstein equations are implemented in the form of the BSSNOK formulation (see Sec. 2.3.2) with the moving-puncture gauge (see Sec. 2.3.4). The Carpet driver [136] provides AMR using the technique of "moving boxes." We use bitant symmetry to exploit the symmetry of the simulations and reduce computational expense. The computational domain comprises a hierarchy of $l_{\max}+1$ refinement levels labelled from $l = 0, \dots l_F, \dots, l_{\max}$, each with grid spacing $h_l = h_0/2^l$. Before applying the symmetry, for $l \leq l_F$ each level consists of a single fixed cubic grid of half-length[1] $R_l = R_0/2^l$, and for $l_F < l \leq l_{\max}$, each level consists of two cubic components of half-length $R_l = 2^{l_{\max}-l} R_{l_{\max}}$ centred around each BH. We adopt this notation for consistency with that used to describe GRChombo. This translates into the more conventional Lean grid setup notation (cf. Ref. [161]) as

$$\left\{(R_0, \dots, 2^{-l_F} R_0) \times (2^{l_{\max}-l_F-1} R_{l_{\max}}, \dots, R_{l_{\max}}), h_{l_{\max}}\right\}. \tag{5.2.1}$$

[1] In one departure from this rule, we enhance R_2 by a factor of 4/3 for the simulations of sequence `lq1:2` of Table 5.1.

A CFL factor of 1/2 is used in all simulations, and apparent horizons are computed with AHFinderDirect [221, 222].

5.2.1.3 Initial data

For both codes, we use puncture data of Bowen-York type (see Sec. 2.2.3) provided by the spectral solver of Ref. [119] in the form of the Cactus thorn TwoPunctures for Lean, and a standalone version integrated into GRChombo (see Sec. 3.2.3). In the latter case, we take advantage of the improvements made in Ref. [291] to use spectral interpolation.

5.2.2 Black-hole binary configurations

We follow the construction of sequences of BH binary configurations and the notation of Ref. [292]. In particular, we denote by M_1 and M_2 the initial BH masses. Without loss of generality, since we are only considering unequal masses ($M_1 \neq M_2$), we take $M_2 > M_1$ and denote their sum by $M = M_1 + M_2$. The reduced mass is $\mu = M_1 M_2/M$ and to quantify the mass ratio, we use either

$$q = \frac{M_1}{M_2} \tag{5.2.2}$$

or the symmetric mass ratio $\eta = \mu/M$. Finally, the total Arnowitt-Deser-Misner (ADM) mass is denoted by $M_{\rm ADM}$ (see Appendix A).

In order to construct a sequence for a fixed mass ratio, we first determine an initial quasicircular configuration. We specify the initial coordinate separation D/M along the x axis, and the scale in the codes is fixed by choosing $M_1 = 0.5$. Next, Eq. (65) in Ref. [103] is used to calculate the initial tangential momentum of each BH, $\mathbf{p} = (0, \pm p, 0)$ (as shown in Fig. 5.1). We use a Newton-Raphson method to iteratively solve for the Bowen-York bare mass parameters that give the desired BH masses. The binding energy of this quasicircular configuration is then computed using

$$E_{\rm b} = M_{\rm ADM} - M. \tag{5.2.3}$$

The rest of the sequence with increasing orbital eccentricity is constructed by fixing the binding energy and gradually reducing the initial linear momentum parameter p. We decide to reduce the linear momentum rather than, for example, alter its direction, so that the x axis is fixed as the initial apoapsis for all configurations. For a given

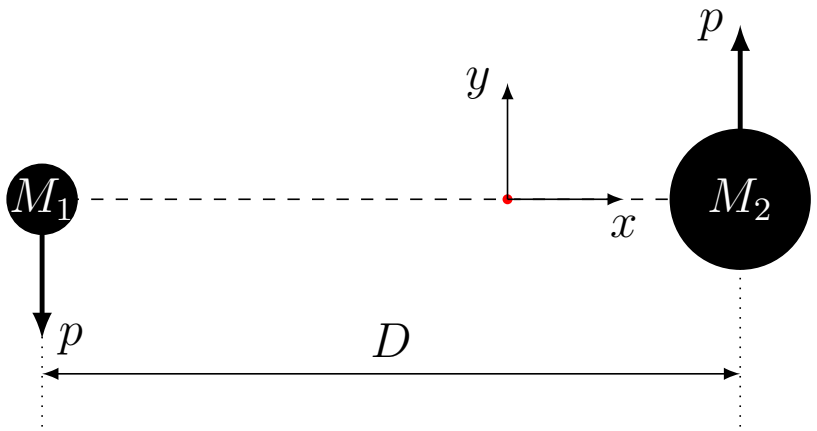


Fig. 5.1 Schematic diagram of the initial BH binary setup for an arbitrary configuration in one of the sequences.

configuration with fixed p, we iteratively solve for the separation D and bare masses that give the required binding energy and BH masses. The choice to keep the binding energy constant as the momentum parameter (and thus the initial kinetic energy) is reduced means that the initial separation increases along the sequence. This ensures an inspiral phase of comparable duration as the eccentricity increases. The initial orbital angular momentum of the system is given by $L = Dp$ (2.2.27). Even though D increases as p decreases, the initial angular momentum of the system monotonically decreases as p decreases for all but the least one or two eccentric configurations in a sequence.

We have parametrised the configurations within a sequence by their initial tangential momentum p, but we would like to measure the eccentricity of these configurations. Unfortunately, there is no gauge-invariant measure of eccentricity [293] and the ambiguity in any definition is particularly pronounced in the late stages of inspiral from which our simulations start. Following Ref. [292], we use the formalism in Ref. [294] to obtain a PN estimate for the eccentricity. Note that this formalism has three eccentricity parameters—e_t, e_r and e_ϕ—and employs two different types of coordinates: ADM-like and harmonic. The choice of which parameter and coordinate type to use is somewhat arbitrary. We mostly focus on the eccentricity parameter e_t in harmonic coordinates[2] as in Ref. [217]. This estimate should be taken with a pinch of salt due to the relatively small initial binary separations D in our simulations. Furthermore, e_t has an infinite gradient as a function of the initial orbital angular momentum in the quasicircular limit (see Fig. 1 in Ref. [292]), such that values of $e_t \lesssim 0.1$ are difficult to realise in practice,

[2]The ADM-like estimate of Ref. [294] differs by only a few percent for $e_t \lesssim 0.8$, and would not significantly alter our results.

Table 5.1 Sequences of binary BH configurations studied in this chapter with their mass ratio, binding energy E_b/M, and the GW extraction radius r_{ex}. For reference, we also list for each sequence the kick velocities v_c in the quasicircular limit. These values agree, within the numerical uncertainties, with the results of Ref. [257].

Sequence	Code	q	E_b/M	r_{ex}/M	v_c (km/s)
`sq2:3`	GRChombo	2/3	−0.0113386	88	102
`sq1:2`	Lean	1/2	−0.0106964	80	149
`lq1:2`	Lean	1/2	−0.0090858	80	150
`sq1:3`	GRChombo	1/3	−0.0093684	65	178

unless the BHs start from large initial distance. In the head-on limit e_t diverges, and a Keplerian/Newtonian interpretation ceases to be valid. Despite these shortcomings, this estimate provides us with a helpful approximation of the eccentricity and a criterion to quantify deviations away from quasicircularity.

The sequences considered in this work are given in Table 5.1. Note that there are two sequences corresponding to the mass ratio $q = 1/2$. The sequence `lq1:2` has a longer inspiral phase compared to the other sequences. For the nearly quasicircular configurations, the binary completes about six orbits before merger in the `lq1:2` sequence, and about three orbits in all other sequences. The longer sequence of simulations was conducted in order to identify any possible artefacts in the shorter sequences due to the exclusion of the earlier inspiral phase. In addition to the labelling of sequences in Table 5.1, we refer to individual simulations within a sequence by appending "`-p`" to the sequence label followed by a four digit integer which is given by $10^3 p/M$ truncated appropriately; for example, `sq1:2-p0100` denotes the simulation in sequence `sq1:2` with initial tangential momentum $p = 0.1M$.

5.2.3 Diagnostics

For all simulations, we have extracted values of the Weyl scalar Ψ_4 on spheres of finite coordinate radius given in Table 5.1 for each sequence. We also computed the dominant terms in the multipolar decomposition (2.4.70).

Our main diagnostics are the energy, linear momentum and angular momentum radiated in GWs, which are computed directly using standard methods (see Secs. 2.4.7 and 3.3.1).

Additionally, we compute the radiated linear momentum from the multipolar amplitudes $\psi_{\ell,m}$ in Eq. (2.4.70) using the formulae of Ref. [97]. From the symmetry of our

configurations, the z component vanishes identically: $P_z^{\rm rad} = 0$. For the components in the orbital plane, we write $P_+^{\rm rad} = P_x^{\rm rad} + \mathrm{i} P_y^{\rm rad}$. Then,

$$P_+^{\rm rad}(t) = \sum_{\tilde{\ell}=2}^{\infty} \sum_{\tilde{m}=-\tilde{\ell}}^{\tilde{\ell}} P_+^{\tilde{\ell},\tilde{m}}, \tag{5.2.4}$$

where

$$P_+^{\tilde{\ell},\tilde{m}}(t) = \lim_{r\to\infty} \frac{r^2}{8\pi} \int_{t_0}^{t} \mathrm{d}t' \left\{ \left(\int_{-\infty}^{t'} \mathrm{d}t'' \, \psi_{\tilde{\ell},\tilde{m}} \right) \left(\int_{-\infty}^{t'} \left[a_{\tilde{\ell},\tilde{m}} \bar{\psi}_{\tilde{\ell},\tilde{m}+1} + b_{\tilde{\ell},-\tilde{m}} \bar{\psi}_{\tilde{\ell}-1,\tilde{m}+1} - b_{\tilde{\ell}+1,\tilde{m}+1} \bar{\psi}_{\tilde{\ell}+1,\tilde{m}+1} \right] \mathrm{d}t'' \right) \right\}, \tag{5.2.5}$$

and the coefficients $a_{\ell,m}$ and $b_{\ell,m}$ are given by

$$a_{\ell,m} = \frac{\sqrt{(\ell-m)\,(\ell+m+1)}}{\ell\,(\ell+1)}, \tag{5.2.6a}$$

$$b_{\ell,m} = \frac{1}{2\ell} \sqrt{\frac{(\ell-2)\,(\ell+2)\,(\ell+m)\,(\ell+m-1)}{(2\ell-1)\,(\ell+1)}}. \tag{5.2.6b}$$

We will find it helpful to define the partial sums,

$$P_+^{\tilde{\ell}} = \sum_{\tilde{m}=-\tilde{\ell}}^{\tilde{\ell}} P_+^{\tilde{\ell},\tilde{m}}, \tag{5.2.7a}$$

$$P_+^{\leq\tilde{\ell}} = \sum_{\tilde{\ell}'=2}^{\tilde{\ell}} P_+^{\tilde{\ell}'}. \tag{5.2.7b}$$

In practice, we do not evaluate the limit in Eqs. (2.4.79), (2.4.80), (2.4.82) and (5.2.5), but rather just evaluate them at the finite extraction radius $r = r_{\rm ex}$, as given in Table 5.1. A discussion of the error this introduces is given in Appendix 5.B.

In order to exclude the spurious radiation inherent in Bowen-York initial data, we start the integration in Eqs. (2.4.79), (2.4.80), (2.4.82), and (5.2.5) at $t_0 = 50M + r_{\rm ex}$. The recoil velocity is computed from the radiated momentum according to

$$\mathbf{v} = -\frac{\mathbf{P}^{\rm rad}}{M_{\rm fin}}, \tag{5.2.8}$$

where $M_{\rm fin}$ is the mass of the BH merger remnant. The quantity $M_{\rm fin}$ can be computed using energy balance:

$$M_{\rm fin} = M_{\rm ADM} - \tilde{E}^{\rm rad}\,, \tag{5.2.9}$$

where $\tilde{E}^{\rm rad}$ denotes the radiated energy *including* the spurious radiation. We similarly compute the spin of the final BH $\chi_{\rm fin}$ (which, by symmetry, must be in the z direction) using the radiated angular momentum:

$$\chi_{\rm fin} = \frac{L - J_z^{\rm rad}}{M_{\rm fin}^2}, \tag{5.2.10}$$

where the initial angular momentum is $L = pD$. For LEAN simulations, we have compared $M_{\rm fin}$ and $\chi_{\rm fin}$ with the corresponding values derived from the apparent horizon properties, and find agreement to within $\leq 0.1\,\%$.

5.3 Results

Using the framework summarised in the previous section, we have simulated four sequences of nonspinning BH binaries, characterised by their mass ratio (5.2.2) and binding energy (5.2.3). The parameters of these sequences are listed in Table 5.1. We have selected our mass ratios such that they cover the regime of maximum recoil, realised for $\eta = 0.195$ or $q = 1/2.77$ (cf. Fig. 5.3). Recall that sequences `sq2:3`, `sq1:2` and `sq1:3` complete about three orbits and sequence `lq1:2` completes about six orbits, respectively, in the quasicircular limit.

Our main results are displayed in Fig. 5.2, where we plot for all sequences the total recoil speed $v_{\rm tot}$, various truncations of the multipolar contributions to the total recoil according to Eqs. (5.2.4)–(5.2.7b), the total radiated GW energy $E^{\rm rad}$ and the dimensionless spin $\chi_{\rm fin}$ of the BH resulting from the merger.

Let us first focus on the total recoil $v_{\rm tot}$, displayed in each of the figure's top panels as the blue solid line. For each mass ratio, the global maximum of the kick velocity is realised for moderate eccentricities $e_t \approx 0.5$. We also illustrate this kick variation in Fig. 5.3, where the solid blue curve shows the quasicircular kick as a function of the symmetric mass ratio η according to Fit 3 in Table V of Ref. [295]. The velocity ranges obtained for our eccentric binaries are overlaid as the vertical bars for each of our sequences. The bar for each constant-η sequence is obtained by starting at the quasicircular limit on the right of each panel in Fig. 5.2 and identifying the minimum and maximum of $v(p)$, excluding the plunge regime to the left of the global maximum.

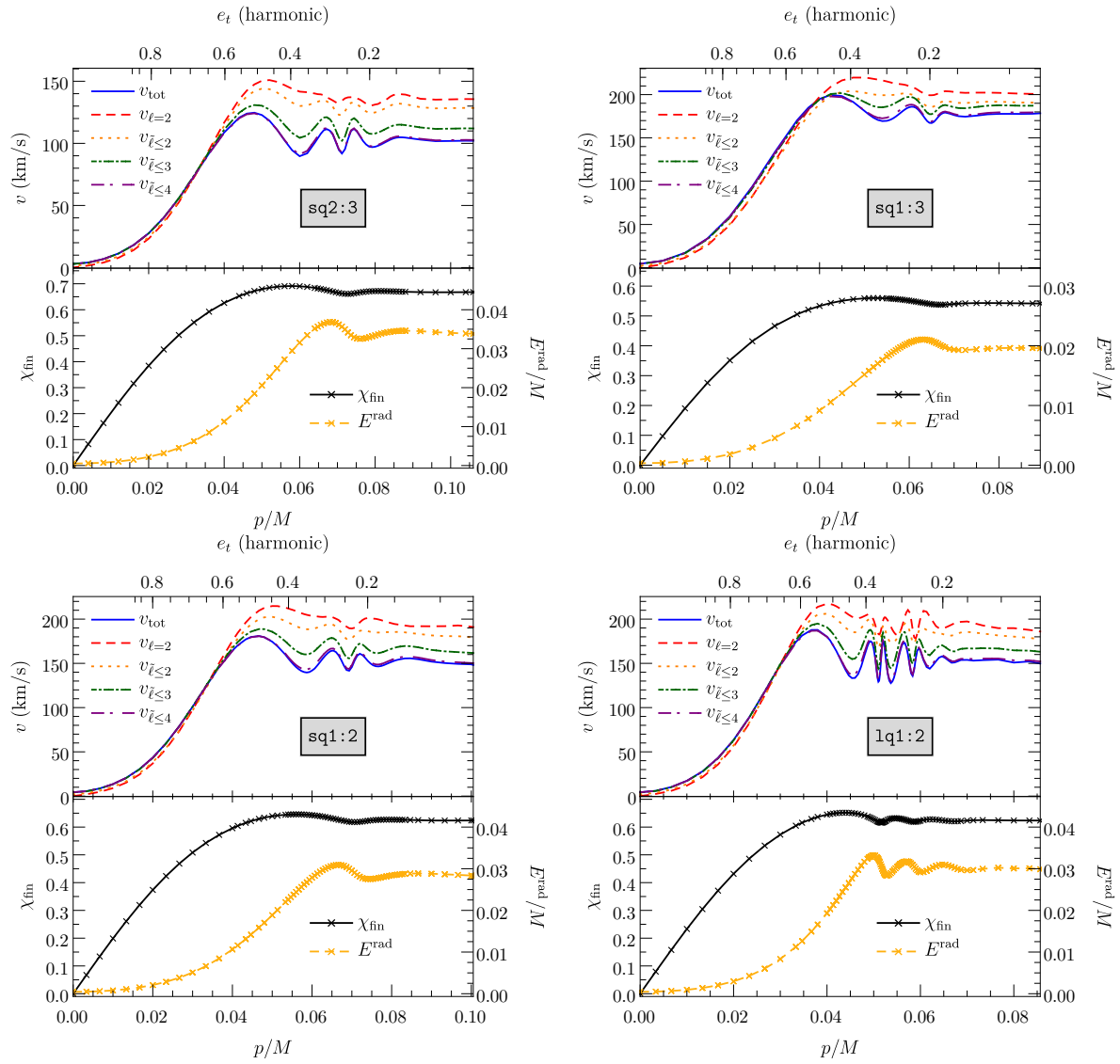


Fig. 5.2 For each sequence of simulations in Table 5.1: Top panel: the recoil velocity v is plotted as a function of the initial tangential momentum p/M. The individual curves represent the total kick $v_{\rm tot}$ (blue, solid), the contribution to the kick from $\ell = 2$ modes of Ψ_4, $\psi_{2,m}$, only in Eqs. (5.2.4)–(5.2.5) $v_{\ell=2}$ (red, dashed), and the contributions to the kick from $P_+^{\leq\tilde{\ell}'}$ defined in Eq. (5.2.7b) $v_{\tilde{\ell}\leq\tilde{\ell}'}$ for $\tilde{\ell}' = 2$ (orange, dotted), $\tilde{\ell}' = 3$ (green, dot-dashed) and $\tilde{\ell}' = 4$ (purple, long dot-dashed). Our estimate of the eccentricity (see Sec. 5.2.2) is provided on the upper horizontal axis. Bottom panel: The final BH spin $\chi_{\rm fin}$ (black, solid) and the energy radiated in GWs $E^{\rm rad}$ (gold, dashed) are also plotted as functions of p/M. For both curves, the individual simulations performed for this analysis are shown by $\times$ symbols.

For our sequences `sq2:3`, `sq1:2` and `lq1:2`, the magnification of the kick through moderate values of the orbital eccentricity is similar to the enhancement by up to 25 %

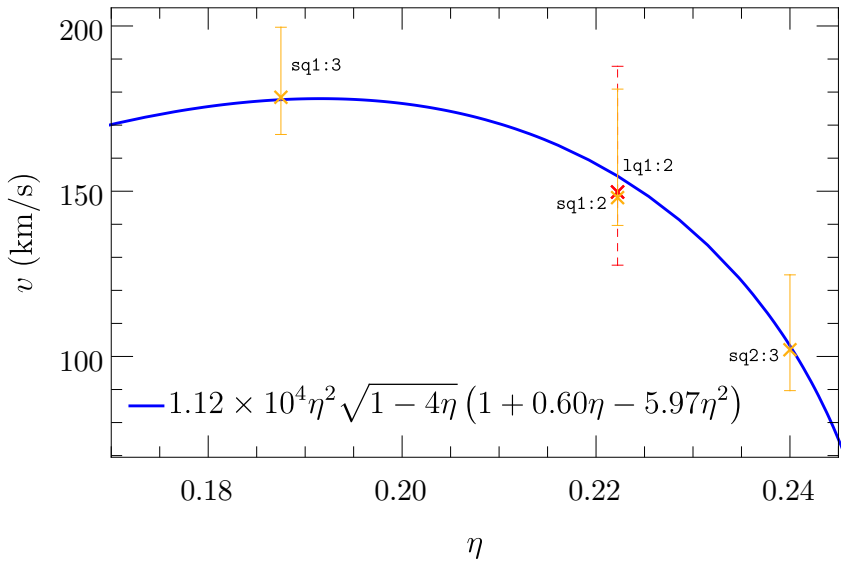


Fig. 5.3 The range of recoil velocities obtained for each sequence is plotted against the symmetric mass ratio η. Note that for each sequence we exclude the configurations with $p < p_{\max}$ (i.e. the head-on limit), where $p = p_{\max}$ is the tangential momentum that maximises the kick. The three short sequences are marked in gold and the long sequence is marked in red (dashed). A fitted formula for the quasicircular kick as a function of η from Ref. [295] is also shown in blue for comparison.

reported in Ref. [217] for the so-called superkick configurations [218, 219]. For `sq1:3` the effect is milder, with a $\sim 12\,\%$ amplification, but still well above the uncertainty estimates of our simulations. On the other hand, as evidenced by the oscillatory pattern of the function $v(p)$ in Fig. 5.2, appropriate nonzero values of the eccentricity can also lead to a *reduction* of the maximum kick at a given mass ratio by $\sim 10\,\%$. This overall modification of the gravitational recoil in the merger of eccentric, nonspinning BH binaries is the first main result of our study.

Besides the global maximum, we also note a number of local minima and maxima in the kick velocity as we vary the eccentricity in Fig. 5.2. For all mass ratios ($q = 2/3,\ 1/2,\ 1/3$) we see about five local extrema in $v(p)$ in our three short sequences, corresponding to the two upper panels and the bottom-left panel. We notice a similar, albeit less pronounced, oscillatory pattern in the functions $E_{\rm rad}(p)$ and $\chi_{\rm fin}(p)$ for the radiated energy and final spin in the lower subpanels in Fig. 5.2. Our results display no systematic correlation, however, between the extrema of the respective quantities; neither global nor local extrema in v, $E_{\rm rad}$ or $\chi_{\rm fin}$ coincide in magnitude or their eccentricity values. We believe this diversity is due to the qualitatively different dependence of the radiated quantities on the GW multipoles: overlaps of *different* multipoles for the kick, a sum of terms $\propto \psi^2_{lm}$

for the energy, and the interaction of first and second time integrals for the angular momentum in Eq. (2.4.82).

We added to our study the $q = 1/2$ sequence of longer BH binary inspirals to investigate whether these anomalies in $v = v(p)$ might merely result from ignoring in our simulations the earlier inspiral phase. The remarkable outcome of this test, however, is that the oscillatory behaviour in the kick as a function of eccentricity is *more* pronounced in the long sequence. The solid blue curve in the bottom-right panel of Fig. 5.2 displays significantly more rapid oscillations in the eccentricity regime $0.2 \lesssim e_t \lesssim 0.4$ as compared to the shorter inspiral sequences. This oscillatory behaviour, and the apparent increase in the number of oscillations as we increase the initial separation of the BHs, is the second of our results.

We next attempt to gain insight into the origin of this behaviour. For this purpose, we have computed the multipolar contributions to the total kick according to Eqs. (5.2.5)–(5.2.7b). The resulting velocities are displayed in Fig. 5.2 by the additional dashed, dotted and dash-dotted curves. Here, the curves labelled $v_{\ell=2}$ have been computed from the $\ell = 2$ modes of Ψ_4 ($\psi_{2,m}$ only) in Eqs. (5.2.4)-(5.2.5). We computed this additional contribution (red dashed curves in the figure) to determine whether the oscillatory behaviour is also present in the pure quadrupole signal. The answer is yes: the oscillations are clearly perceptible in $v_{\ell=2}$, even though they are a bit milder than in the total kick $v_{\rm tot}$. Considering all (cumulative) multipolar contributions shown in Fig. 5.2, we notice the following behaviour:

(1) The oscillatory dependence of the kick on eccentricity is present at any level of truncating the multipolar contributions in the cumulative sum (5.2.7b).

(2) The partial sum of the kick up to $\tilde{\ell} = 4$ barely differs from the total kick, indicating that higher-order overlap terms do not significantly contribute to the kick.

(3) The higher-order contributions $\tilde{\ell} > 2$ to the cumulative kick (5.2.7b) systematically decrease the kick, counteracting the pure quadrupole contribution $v_{\ell=2}$.

In short, we have not identified any specific multipoles dominating the variation in the kick function $v = v_{\rm tot}(p)$.

In our search for an explanation, we turn next to the infall direction of the BH binary just before merger. A well-known feature of the superkicks generated in the inspiral of BHs with opposite spins $\boldsymbol{S}_1 = -\boldsymbol{S}_2$ pointing in the orbital plane is the sinusoidal variation with the initial azimuthal angle of the spin vectors; cf. Fig. 4 in Ref. [296].

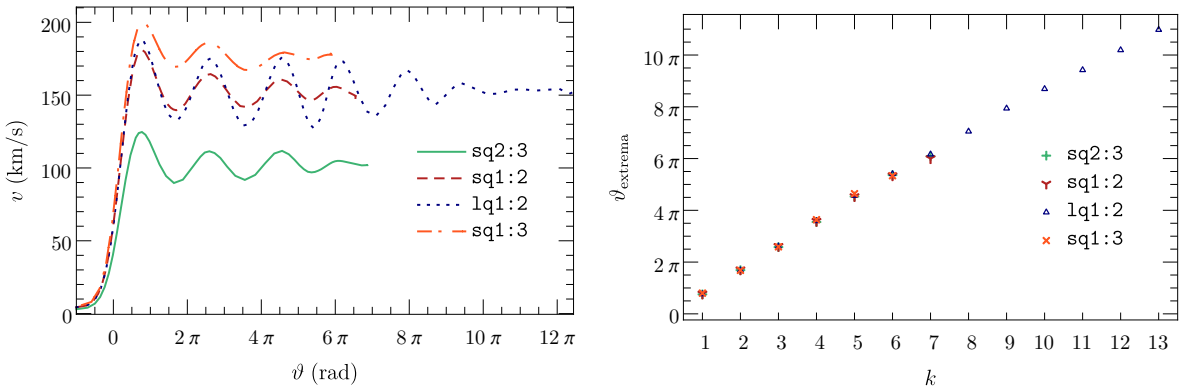


Fig. 5.4 Plots involving the angle of the kick ϑ for all sequences. In the left panel we plot the BH recoil velocity v against ϑ. In the right panel we plot the location of the local extrema $\vartheta_{\text{extrema}}$ of the left panel against the index of the extrema k counting rightwards from the global maximum on the left.

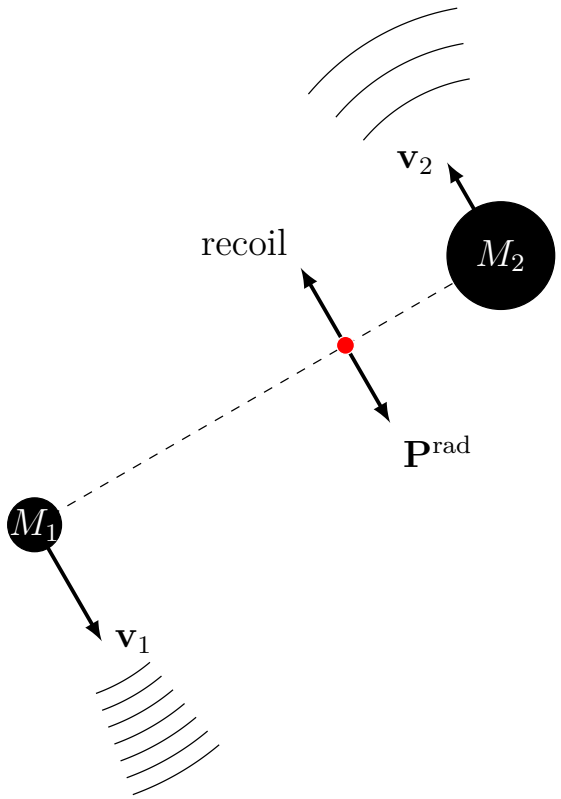


Fig. 5.5 Reproduced from Fig. 3 in Ref. [297]. Schematic diagram showing the excess beaming of the GWs in the direction of the smaller and faster BH in unequal-mass binaries.

The initial orientation of the spins can, alternatively, be interpreted as a measure for the angle between the in-plane spin components and the BH binary's infall direction at merger [298]. The superkick is therefore commonly determined by simulating otherwise identical BH binary configurations for different values of this angle and fitting the resulting data with a cosine function; see, e.g., Sec. III A in Ref. [217]. For the eccentric, nonspinning BH binaries considered in this chapter, it is the initial apsis (either a periapsis or an apoapsis) that defines a reference direction. Unfortunately, neither the apsis nor a "binary infall direction" are rigorously defined quantities in the strong-field regime of general relativity, and we consider instead the orientation of the final kick relative to the x axis, defined by

$$\tilde{\vartheta} = \arg(v_x + \mathrm{i}v_y)\,. \tag{5.3.1}$$

For convenience, we define

$$\vartheta = \tilde{\vartheta} + 2n\pi, \tag{5.3.2}$$

where $n \geq 0$ is chosen minimally for each configuration in order to obtain ϑ as a monotonic function of the initial tangential momentum p for each sequence. We will interchangeably refer to ϑ and $\tilde{\vartheta}$ as the angle of the kick. Since all of our simulations start with the BHs located on the x axis with purely tangential initial momentum $\mathbf{p} = (0, \pm p, 0)$ (Fig. 5.1), the x direction can be regarded as the initial direction of the apoapsis. If we furthermore interpret the gravitational recoil to be predominantly generated by the excess beaming of the GWs in the direction of the smaller and faster BH (see Fig. 5.5) during the short merger phase, the kick direction can serve as an approximate measure for the infall direction of the binary.

We can test this prediction by computing the kick magnitude as a function of the angle ϑ; if correct, we would expect a periodic variation with a period close to 2π. We do not expect an exact 2π periodicity because the relevant periapsis (or apoapsis) direction should be the last one before merger, and will shift away from the x axis during the inspiral due to apsidal precession—the BH analogue of Mercury's perihelion precession around the Sun. More specifically, we would expect deviations from a 2π periodicity to be more pronounced for longer inspirals, i.e., lower eccentricity and/or larger initial separations, but only mildly dependent on the mass ratio q. Quite remarkably, all of these features are borne out by the functions $v = v(\vartheta)$ displayed for our four sequences in the left panel of Fig. 5.4 and the location of the extrema in this plot shown in the right panel of Fig. 5.4. For all sequences we observe the same approximate 2π periodicity, with deviations from this value increasing at larger ϑ, i.e. for longer inspirals. Note also that $\vartheta = -\pi$ in the head-on limit, as expected for our initial configurations, that start with the heavier BH located on the positive x axis.

While short of a rigorous proof, this result provides considerable evidence in favour of interpreting the oscillatory dependence of the kick on the eccentricity as a consequence of the corresponding variation in the infall direction as measured relative to the last apoapsis (or periapsis) of the eccentric binary. This interpretation also explains why the longer sequence `lq1:2` exhibits more oscillations than the shorter sequences `sq1:3`, `sq1:2` and `sq2:3`. Let us consider for this purpose two binary configurations that only differ by a tiny amount of eccentricity δe. The longer the inspiral phase, the more time these two binaries have to build up a considerable phase difference and, hence, a different kick and merger GW signal. Note the potentially dramatic consequences of this behaviour for the GW emission from eccentric binaries over astrophysical time scales. For long

astrophysical inspirals retaining some eccentricity near merger, the kick and GW merger signal should exhibit critical dependence on the eccentricity. In terms of our Fig. 5.2, the function $v = v(e_t)$ would display a huge number of oscillations rather than the handful observed in our case, and the resulting curve would look like a "band" rather than a single line. Within the band, a very small change δe_t in eccentricity can produce a finite change in the kick and merger waveform.

As indicated by our analysis of the multipolar contributions to the total recoil, the variations in the GW signal are of a complex nature. We defer a more comprehensive analysis of the GW pattern to future work, but merely illustrate with an example the type of variations that are encountered. For this purpose, we show in Fig. 5.6 the $(\ell, m) = (2, 2)$ and $(3, 3)$ multipoles of the GW signal around merger for the configurations `lq1:2-p0537` and `lq1:2-p0567`, corresponding to a local minimum and maximum in the kick, respectively; cf. the bottom-right panel of Fig. 5.2. In Fig. 5.6, the time has been shifted such that $\Delta t = 0$ corresponds to the first occurrence of a common apparent horizon. The main difference perceptible in the figure is the relative phase shift of the (3,3) mode relative to the dominant quadrupole (2,2). For the case $p = 0.567M$ with maximal kick, the global peaks of both multipoles are aligned, whereas for $p = 0.537M$ with minimal kick, the global peak of the $(2, 2)$ mode coincides with a minimum in $(\ell, m) = (3, 3)$. We have made similar observations for other pairs of modes such as $(2, 2)$ and $(2, 1)$, and find these pairs to dominate the oscillatory variation in the multipolar series expansion (5.2.7b).

5.4 Conclusion

In this chapter we have studied the gravitational recoil and GW emission of sequences of nonspinning BH binaries with mass ratios $q = 2/3$, $1/2$ and $1/3$, and eccentricity varying from the quasicircular to the head-on limit. For this purpose we have evolved 274 configurations with the GRChombo and Lean codes. Both codes yield convergent results for the recoil with a total error budget of 3-4 % and exhibit excellent agreement, well within this uncertainty estimate, for a verification configuration simulated with both codes. In order to estimate the impact of variations in the overall length of the inspirals, we have evolved two sequences for the case $q = 1/2$ which complete about three and six orbits, respectively, in the quasicircular limit.

The findings of our study are summarised as follows.

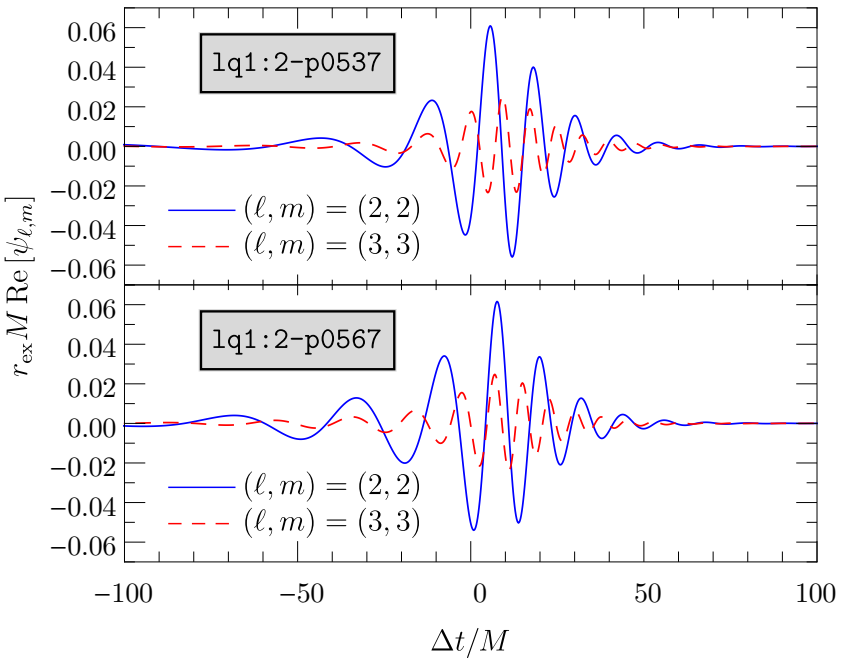


Fig. 5.6 The real parts of the $(\ell, m) = (2, 2)$ and $(3, 3)$ modes of Ψ_4 are shown as functions of time for the two binaries of sequence `lq1:2` with $p/M = 0.537$ and $p/M = 0.567$, resulting in kick velocities of $v = 128$ and $173\,\mathrm{km/s}$, respectively.

(i) For all sequences, the total recoil reaches a global maximum for moderate eccentricities $e \sim 0.5$. As in the case of the enhancement of superkicks studied in Ref. [217], the maximum kick is enhanced by up to about 25 % relative to the value obtained for quasicircular configurations.

(ii) Besides this global maximum, we observe an oscillatory dependence of the kick v as a function of eccentricity, with several local minima and maxima in the function $v = v(e)$. Appropriate nonzero values of the eccentricity can lead to a *reduction* of the kick by $\sim 10\,\%$ relative to the quasicircular value instead of an increase. By splitting the kick into separate multipolar contributions, we notice that this oscillatory dependence is already present, albeit in a slightly weaker form, when we consider only quadrupole terms in the series expansion (5.2.4). Further contributions involving $\ell \geq 2$ multipoles tend to decrease the overall kick and mildly enhance the oscillatory variation; see Fig. 5.2.

(iii) We interpret this oscillatory variation in the kick as a consequence of changes in the angle between the infall direction at merger and the apoapsis (or periapsis) direction. In the absence of rigorous definitions for either of these directions, we

approximate this angular variation by considering the direction of the final kick and the x axis, assuming that the former is related via relativistic GW beaming to the infall direction and by taking into account that our BHs start on the x axis with zero radial momentum. Displayed as a function of this angle, the kick displays the expected periodic behaviour with a period close to but mildly deviating from 2π, presumably due to periapsis precession.

(iv) We have explored the dependence of this oscillatory behaviour of the recoil by simulating an additional sequence of eccentric binaries with mass ratio $q = 1/2$, but less negative binding energy, corresponding to about six orbits in the quasicircular limit. We find the oscillations in $v = v(e)$ to be more pronounced and numerous than in the shorter sequence. We attribute this feature to the longer available time window during which otherwise identical binaries with tiny differences in the initial eccentricity build up a phase difference prior to merger. This observation raises the intriguing possibility that the total recoil depends highly sensitively on the initial eccentricity.

(v) The variations in the kick velocity are accompanied by relative time shifts in the peak amplitudes of subdominant multipoles relative to the peaks of the (2,2) mode; cf. Fig. 5.6. For configurations with a large (small) kick, the peak amplitude of subdominant multipoles tends to be aligned (misaligned) with the quadrupole peak.

Our findings point to a variety of future investigations. While our simulations indicate an increased sensitivity of the GW merger signal to the initial eccentricity for larger initial separations (i.e. longer inspirals), it is not clear how this will be affected by the circularizing nature of GW emission. In this context, it will also be important to analyse in more quantitative terms the differences in the GW signals and possible implications for parameter inference in GW observations. A thorough investigation of long eccentric inspirals on astrophysical time scales will likely require PN methods and may benefit greatly from a multi-time-scale analysis in phase space, as applied to spin-precessing BH binaries in Refs. [299, 300] or to the dynamics of binary systems in external gravitational background potentials in Refs. [301, 302]. If there is a single conclusion to draw from the results of this work, it is the surprisingly rich phenomenology of the GW signals of eccentric compact binaries—even in the absence of spins—which merits as much as it requires further investigation.

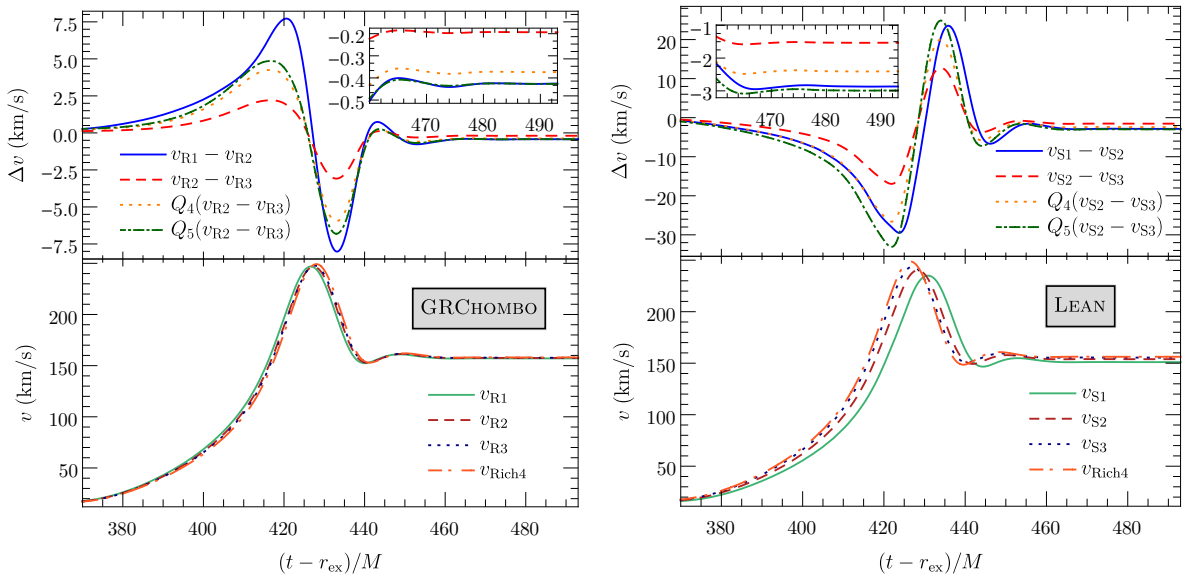


Fig. 5.7 For each code, we show convergence plots for the accumulated linear momentum radiated from sq1:2-p0100 by plotting the BH recoil velocity in the bottom panels. The Richardson extrapolated curve, v_{Rich4}, assuming fourth-order convergence, is also shown in the bottom panel. The grid configurations are given in Table 5.2 for GRChombo and in Table 5.3 for Lean. The top panel shows the difference between the configurations along with rescalings corresponding to fourth- and fifth-order convergence. The inset shows a magnification of the right side of the plot: the final value of the recoil velocity is what we show in Fig. 5.2.

Appendix 5.A Numerical accuracy

As in Ref. [217], the uncertainty in our numerical results for the recoil velocities has two predominant contributions: the discretization error and the finite extraction radii for the Weyl scalar Ψ_4.

To estimate the uncertainty arising from the latter, we have selected a representative sample of the simulations from each sequence and extrapolated the cumulative radiated momentum to infinity from about six extraction radii in the range $r_{\mathrm{ex}}/2 \leq r_{\mathrm{ex}}$ using the procedure explained in Sec. 3.3.3. We report the results from the finite extraction radii given in Table 5.1 and estimate the error by comparing with the linear-order extrapolation. For both codes, we estimate that the contribution from this error is about 2% for all sequences.

In order to estimate the error contribution from finite differencing and verify that our codes give consistent results, we have performed simulations of sq1:2-p0100 (the

Table 5.2 Grid configurations used for GRChombo simulations. As explained in Sec. 5.2.1.1 and Appendix 5.B, the total number of refinement levels is $l_{\max}+1$, the number of cells along each dimension on the coarsest level is N, τ_R is the regridding threshold value, b is the BH tagging buffer parameter that we set proportional to the mass M_i ($i=1,\,2$) of the nearest BH for all configurations except R4, and h denotes the grid spacing on the finest level.

Label	$l_{\max}$	N	τ_R	b	h/M_1	tagging
R1	7	320	0.012	$0.5M_i$	3/80	Spherical
R2	7	368	0.01043	$0.5M_i$	3/92	Spherical
R3	7	416	0.00923	$0.5M_i$	3/104	Spherical
R4	7	352	0.01091	0.7	3/88	Box

binary in sequence `sq1:2` with $p/M=0.1$) with both codes. We discuss the analyses of the convergence of each code separately before comparing.

5.A.1 GRChombo convergence

For GRChombo, we have performed the simulations of `sq1:2-p0100` with finest resolutions $h=3M_1/80$, $3M_1/92$ and $3M_1/104$, and we refer to the configurations corresponding to these resolutions as R1, R2 and R3, respectively. The full grid configurations are given in Table 5.2 and the results of this analysis are shown in the left panel of Fig. 5.7. Around merger, at $(t-r_{\rm ex})/M\sim 420$, our results exhibit mild overconvergence in the top-left panel of Fig. 5.7. The important results for our analysis in Fig. 5.2, however, are the final kick values after the merged BH has settled down. As can be seen from the inset, the convergence here is close to fifth order. From our convergence analysis, the difference between the result obtained from the R1 simulation and the more conservative fourth-order Richardson-extrapolated result leads to an estimate of the discretization error of about 1%. A similar error estimate is also obtained for the radiated energy, $E^{\rm rad}$. From experience, we have found smaller values for the mass ratio $q<1$ more challenging to accurately simulate than larger values, and we therefore feel justified in using this error estimate (for a $q=1/2$ configuration) as a conservative estimate for the error in the `sq2:3` sequence simulations ($q=2/3$). We therefore used the R1 grid configuration for this sequence with $l_1^{\max}=l_2^{\max}=l_{\max}=7$ (both BHs are covered by the finest level; see Appendix 5.B for details).

For the `sq1:3` simulations, we used the R4 grid configuration (see Table 5.2) with $l_1^{\max}=l_{\max}=7$ and $l_2^{\max}=l_{\max}-1=6$ (the larger BH is not covered by the finest level:

Table 5.3 Grid configurations used for LEAN simulations. As explained in Sec. 5.2.1.2, the total number of refinement levels is $l_{\max}+1$, the number of fixed refinement levels is l_F+1, R_0 is the half-length of the outer grid, $R_{l_{\max}}$ is the half-length of one cubic component of the innermost grid, and h is the grid spacing on the finest level.

Label	$l_{\max}$	l_F	R_0	$R_{l_{\max}}$	h/M_1
S1	7	4	384	1	1/20
S2	7	4	384	1	1/24
S3	7	4	384	1	1/32
S4	7	4	384	1	1/28

see Appendix 5.B for details). This corresponds to a resolution of $h = 3M_1/88$. We performed a separate convergence analysis of `sq1:3-p0089`, which led to an estimated 1% discretization error.

Combining both the finite extraction radius and discretization errors, our estimate for the total error budget of the GRCHOMBO simulations is about 3%.

5.A.2 Lean convergence

With LEAN, we have simulated `sq1:2-p0100` with finest resolutions $h = M_1/20$, $M_1/24$ and $M_1/32$. We refer to these grid configurations as S1, S2 and S3, respectively (cf. Table 5.3). The right panel of Fig. 5.7 shows convergence between fourth and fifth order. For simulations in `sq1:2`, we used the S2 grid configuration. From the convergence analysis, the difference between the result obtained from the S2 simulation and the fourth-order Richardson extrapolation leads to an estimate of the discretization error of about 1.5%.

For the `lq1:2` simulations, we have undertaken a separate convergence analysis of `lq1:2-p0086` using the same grid setup as in Table 5.3, but using higher resolutions $h/M_1 = 1/24$, 1/28 and 1/32. We observe convergence close to fourth order and obtain an error estimate of 1 % from the Richardson-extrapolated kick for the medium resolution $h/M_1 = 1/28$.

In summary, the LEAN simulations of sequence `sq1:2` are performed with resolution grid S2 of Table 5.3 and an error budget of 3.5 %, and those of sequence `lq1:2` with grid S4 of Table 5.3 and an error budget of 3 %.

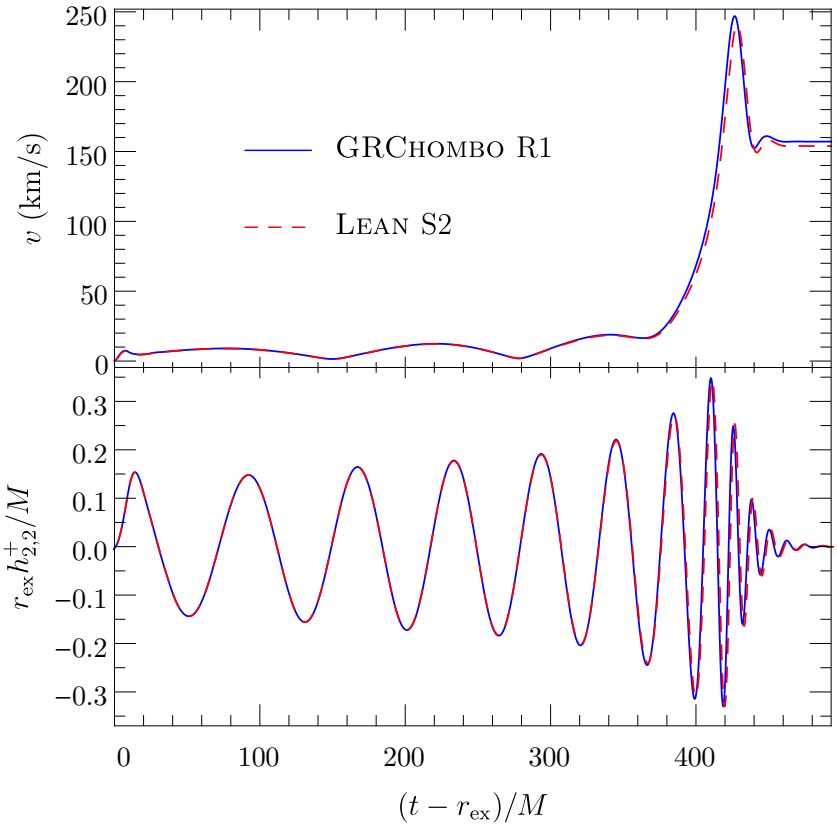


Fig. 5.8 Comparison between GRChombo and Lean for the accumulated linear momentum radiated in GWs in simulations of `sq1:2-p0100` with $e_t = 0.10$. We compare the BH recoil velocity (top panel) and the corresponding plus-polarized $\ell = m = 2$ strain amplitude (bottom panel).

5.A.3 Comparison between GRChombo and Lean

A comparison of the recoil velocity computed from GRChombo and Lean simulations of `sq1:2-p0100` with the grid configurations R1 and S2 (used for the `sq2:3` and `sq1:2` runs) respectively, is shown in the top panel of Fig. 5.8. The eccentricity estimate for this system is $e_t = 0.10$. We have chosen this configuration for two reasons. First, to determine appropriate resolutions, we had to calibrate our codes' accuracy at the start of our exploration, which we began in the regime of mild eccentricities to acquire an intuitive understanding of their behaviour. Second, configurations with mild eccentricity have a longer inspiral phase than highly eccentric ones, and therefore impose a stronger requirement on phase accuracy. A mildly eccentric binary is therefore ideally suited to obtain a conservative estimate of the numerical accuracy, which is representative across the targeted parameter space.

The final recoil velocities obtained for this configuration with our two codes differ by about 2%, which is well within the error budget of each code. We also show the quadrupole contribution h^{+}_{22} to the '+' polarization strain defined by Eq. (2.4.73) in the bottom panel of the figure, to better illustrate the agreement between the codes for these grid configurations.

In Fig. 5.7 the differences between the results of different resolutions with Lean are greater than that of GRChombo. However, we found that Lean entered the convergent regime at lower resolutions than GRChombo. This is compatible with the observations of Ref. [84] that higher resolutions were required for convergence with CCZ4 compared to BSSNOK.

Appendix 5.B GRChombo tagging criterion

As explained in Sec. 3.1.4, the regridding is controlled by the tagging of cells for refinement in the Berger-Rigoutsos algorithm [133], with cells being tagged if the tagging criterion C exceeds the specified threshold value τ_R as given in Table 5.2. For this work, we use the tagging criterion

$$C = \begin{cases} 0, & \text{if } l \geq l_{\rm BH}^{\rm max} \text{ and } r_{\rm BH} < (M_{\rm BH} + b), \\ \max(C_\chi, C_{\rm punc}, C_{\rm ex}), & \text{otherwise,} \end{cases} \tag{5.B.1}$$

where $l_{\rm BH}^{\rm max}$ is a specifiable maximum level parameter for each BH (so that it is not unnecessarily over resolved), $r_{\rm BH}$ is the coordinate distance to the puncture, $M_{\rm BH}$ is the mass of the corresponding BH, b is a buffer parameter, and C_χ, $C_{\rm punc}$, and $C_{\rm ex}$ are given as follows:

(i) C_χ tags regions in which the gradients of the conformal factor χ become steep. It is given by

$$C_\chi = h_l \sqrt{\sum_{i,j} \left(\partial_i \partial_j \chi\right)^2}, \tag{5.B.2}$$

where h_l is the grid spacing on refinement level l.

(ii) $C_{\rm punc}$ tags within spheres around each puncture in order to ensure the horizon is suitably well resolved. It is given by

$$C_{\rm punc} = \begin{cases} 100, & \text{if } r_{\rm BH} < (M_{\rm BH} + b) 2^{\max(l_{\rm BH}^{\rm max} - l - 1, 2)}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.B.3}$$

(iii) $C_{\rm ex}$ ensures each sphere on which we extract the Weyl scalar Ψ_4 is suitably well resolved. It is given by

$$C_{\rm ex} = \begin{cases} 100, & \text{if } r < 1.2 r_{\rm ex} \text{ and } l < l_{\rm ex}, \\ 0, & \text{otherwise,} \end{cases} \tag{5.B.4}$$

where r is the coordinate distance to the centre of mass, $r = r_{\rm ex}$ gives the location of the extraction sphere, and $l_{\rm ex}$ is a specifiable extraction level parameter for each sphere.

We also used this tagging criterion with the replacement $r_{\rm BH} \to \max(x_{\rm BH}, y_{\rm BH}, z_{\rm BH})$, where, e.g. $x_{\rm BH}$ is the distance to the puncture in the x direction. We refer to this as "box" tagging and the original as "spherical" tagging. Naively, one might hope that C_χ is sufficient to ensure suitable refinement around the BHs, since the gradients of χ become increasingly steep close to the punctures. However, we found empirically that, without $C_{\rm punc}$, the horizons are perturbed significantly by the refinement boundaries, leading to lower accuracy.

Chapter 6

Numerical investigations of boson stars

The content presented in this chapter contains material that appears in the co-authored publication Ref. [4] with T. Helfer, U. Sperhake, R. Croft, B. Ge and E. Lim. In particular, the simulations presented in Sec. 6.4.4 were conducted and analysed by co-authors and are presented here with interpretation in my own words and in an abbreviated form for coherence. The rest of the work in this chapter is either entirely my own or that fully reproduced by me and given in my own words.

6.1 Introduction

The origins of boson stars (BS) can be traced back to the 1950s, when Wheeler [303] sought to find spatially localised configurations ("gravitational atoms") of electromagnetic fields in GR, which he termed *geons*. By switching from real to complex-valued fundamental fields, it is even possible to obtain configurations that are genuinely stationary solutions to the Einstein equation. First established for spin-0 (scalar) fields [304, 305, 306], this idea has more recently been extended to spin-1 (vector) fields [307] (commonly referred to as *Proca* stars) in addition to systems with multiple scalar fields [308]. In comparison to the relatively simple nature of black holes, the existence of these equilibrium configurations relies on an elaborate balance between the intrinsically dispersive character of these fundamental fields and their self-gravitation.

The first BS models computed in the 1960s comprised a massive free (i.e. non-interacting) complex scalar field φ. The resulting equilibrium states of interest, commonly referred to as *mini boson stars*, form a one parameter family of ground-state solutions

characterised by the central scalar-field amplitude φ_0. These solutions exhibit a stability structure analogous to that of Tolman-Oppenheimer-Volkoff [309, 310] stars with a stable and unstable branch separated by the configuration with maximal mass [311, 312, 313, 314]. For each value of φ_0, there also exists a countable hierarchy of excited states with $n > 0$ nodes in the scalar profile [315, 316], but numerical evolutions show that these states are unstable [317]. Intriguingly, their decay to either ground state BSs or BHs occurs through a cascade of intermediate states similar to atomic transitions–in keeping with the "gravitational atom" description.

The addition of self-interaction terms for the scalar field can provide further repulsion against gravity which enables the existence of objects far more compact than mini boson stars and even neutron stars (NS) [318, 319, 320, 321]. For plausible scalar field masses, these models can have more astrophysically relevant masses. This raises the question whether such stars may be able to act as black hole mimickers in binary coalescences or if their GW emissions might be distinguishable from that of BHs or NSs [322]. Recent studies conclude that this may be feasible with next-generation GW detectors and, in certain situations, even with advanced LIGO [323, 324, 325]. Beyond the potential GW phenomenology interest in BSs, they have also been suggested as a potential model of *dark matter halos* that could explain the flat rotation curve observed for most galaxies [326].

For BS binaries, the simplest case is a head-on collision and this is investigated in Refs. [327, 328, 89] with the effect of shifts in the phase and differences in the sign of the frequency (BS-anti-BS) considered. The final state of these mergers depends significantly on these differences with the possibilities being either a BH, a BS, two BSs that fail to merge (in the case of a phase offset close to π) or near annihilation of the scalar field (in the case of opposite frequencies). The inclusion of orbital angular momentum resulting in inspirals is considered in Refs. [329, 330]. The GW signals from such binaries are qualitatively similar to those of BH binaries but display more complex structure around merger.

Even if boson stars are not realised in nature, their relative simplicity means they can act as a proxy for other types of horizonless compact objects in General Relativity. Unlike fluid matter, there is no need to deal with the complications of shocks and flux conservation. Furthermore, whereas neutron stars have sharp boundaries which often require workarounds such as artificial atmospheres in order to perform stable numerical evolutions (see, for example, Refs. [331, 332]), the smooth profile of boson stars dispenses with the need for such measures.

This chapter is organised as follows. In Sec. 6.2, we introduce the Einstein-Klein-Gordon system and the 3+1 evolution formalism. Next, in Sec. 6.3, we detail how to construct some single boson star models and show some of their properties. Then in Sec. 6.4, we consider the construction of binary initial data, describe some of the issues that arise with existing methods and explain how these issues can be ameliorated.

In this chapter, we set $\hbar = 1$ but restore factors of Newton's constant G.

6.2 The Einstein-Klein-Gordon system

6.2.1 Action and covariant equations

The action for a complex scalar field φ minimally coupled to gravity is

$$S = S_{\rm EH} + S_{\rm M}, \quad \text{where} \tag{6.2.1a}$$

$$S_{\rm EH} = \frac{1}{16\pi G}\int_{\mathscr{M}} \mathrm{d}^4x\,\sqrt{-g}^{(4)}R, \tag{6.2.1b}$$

$$S_{\rm M} = -\frac{1}{2}\int_{\mathscr{M}} \mathrm{d}^4x\,\sqrt{-g}\left[g^{\alpha\beta}\nabla_\alpha\bar{\varphi}\nabla_\beta\varphi + V\left(|\varphi|^2\right)\right], \tag{6.2.1c}$$

where V is the scalar potential[1]. The choice of this potential determines the character of the resulting BS models, and we shall discuss the relevant choices we make later on.

Variation of the action (6.2.1) with respect to the spacetime metric and scalar field yields the *Einstein-Klein-Gordon* (EKG) system[2],

$${}^{(4)}R_{\mu\nu} - \frac{1}{2}{}^{(4)}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{6.2.2a}$$

$$g^{\alpha\beta}\nabla_\alpha\nabla_\beta\varphi = V'\left(|\varphi|^2\right)\varphi, \tag{6.2.2b}$$

where the energy-momentum (EM) tensor is given by

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_{\rm M}}{\delta g^{\mu\nu}} = \nabla_{(\mu}\bar{\varphi}\nabla_{\nu)}\varphi - \frac{1}{2}g_{\mu\nu}\left[g^{\rho\sigma}\nabla_\rho\bar{\varphi}\nabla_\sigma\varphi + V\left(|\varphi|^2\right)\right]. \tag{6.2.3}$$

[1]Note the slightly unusual convention of V being multiplied by $-1/2$.

[2]Note that since we specified the potential V as a function of $|\varphi|^2$, we use $V'(|\varphi|^2)$ to denote differentiation with respect to $|\varphi|^2$ which leads to an additional factor of φ from the variation in Eq. (6.2.2b).

An important quantity arises from the U(1) invariance of the action (6.2.1). Noether's theorem [333] gives the conserved current

$$J^\mu = \frac{\mathrm{i}}{2} g^{\mu\alpha} \left(\bar{\varphi} \nabla_\alpha \varphi - \varphi \nabla_\alpha \bar{\varphi} \right), \tag{6.2.4}$$

such that $\nabla_\mu J^\mu = 0$. The corresponding *Noether charge* is

$$N = -\int_\Sigma \mathrm{d}^3x \, \sqrt{\gamma} n_\alpha J^\alpha, \tag{6.2.5}$$

for a spatial slice Σ with unit normal n_μ and induced metric determinant γ. This quantity can be interpreted as the total number of bosonic particles [306]. Given that we work on a finite computational domain, of course we may not be able to fully realise the conservation of N if scalar matter is radiated outwards. Nevertheless, it can be shown that a local conservation law holds over a closed volume $V \subset \Sigma$ which remains fixed in *coordinate* space [140]. Writing the current (6.2.4) in terms of its normal and spatial projections (cf. Eq. (2.1.8)),

$$\mathcal{N} = -n_\alpha J^\alpha, \qquad \mathcal{J}^\mu = -\perp^\mu_\alpha J^\alpha, \tag{6.2.6}$$

as $J^\mu = \mathcal{N} n^\mu + \mathcal{J}^\mu$, we find

$$\frac{\partial}{\partial t} \int_V \mathrm{d}^3x \, \sqrt{\gamma} \mathcal{N} = \int_{\partial V} \mathrm{d}^2x \, \sqrt{h} s_m (\beta^m \mathcal{N} - \alpha \mathcal{J}^m), \tag{6.2.7}$$

where h is the determinant of the metric induced on ∂V, s^i is the outward pointing unit normal to ∂V in Σ, β^i is the shift vector and α is the lapse function.

6.2.2 Evolution equations

We use the CCZ4 system (2.3.30) to evolve the spacetime variables but we still need an evolution system for the scalar field. In order to reduce the Klein-Gordon equation (6.2.2b) to a first-order[-in-time] system of evolution equations, we introduce the auxiliary variable

$$\Pi = -\mathcal{L}_{\boldsymbol{n}} \varphi = -n^\alpha \nabla_\alpha \varphi = -\frac{1}{\alpha} (\partial_t - \beta^k \partial_k) \varphi. \tag{6.2.8}$$

The choice of the temporal first-order reduction variable Π is arbitrary so long as the resultant formulation is sufficiently well-posed. We choose Π as above for consistency with the other variables in the CCZ4 formulation and with the rest of the literature (e.g.

Refs. [89, 334]). Since we are reducing to first-order derivatives in time, we could also reduce to first-order in space. However, we choose to not do so in order to minimise the memory required during numerical simulations and also for consistency with elsewhere in the literature. Rewriting the Klein-Gordon equation in terms of the CCZ4 variables (2.3.29) yields the following evolution system for our matter variables $\{\varphi, \Pi\}$,

$$\partial_t \varphi = \beta^k \partial_k \varphi - \alpha \Pi, \tag{6.2.9a}$$

$$\begin{aligned}\partial_t \Pi = \beta^k \partial_k \Pi - \chi \tilde{\gamma}^{kl} \partial_k \alpha \partial_l \varphi + \alpha \Big[\chi \tilde{\Gamma}^k \partial_k \varphi + \frac{1}{2} \tilde{\gamma}^{kl} \partial_k \chi \partial_l \varphi \\ - \chi \tilde{\gamma}^{kl} \partial_k \partial_l \varphi + \Pi K + V' \left(|\varphi|^2 \right) \varphi \Big] .\end{aligned} \tag{6.2.9b}$$

The projections of the EM tensor (see Sec. 2.1.6) in terms of the CCZ4 and matter variables are

$$\rho = \frac{1}{2} \left[\Pi_R^2 + \Pi_I^2 + \chi \tilde{\gamma}^{ij} \left(\partial_i \varphi_R \partial_j \varphi_R + \partial_i \varphi_I \partial_j \varphi_I \right) + V \left(\varphi_R^2 + \varphi_I^2 \right) \right] , \tag{6.2.10a}$$

$$j_i = \frac{1}{2} [\Pi \partial_i \bar{\varphi} + \bar{\Pi} \partial_i \varphi], \tag{6.2.10b}$$

$$S_{ij} = \partial_{(i} \bar{\varphi} \partial_{j)} \varphi - \frac{1}{2\chi} \tilde{\gamma}_{ij} \left[\chi \tilde{\gamma}^{kl} \partial_k \bar{\varphi} \partial_l \varphi - |\Pi|^2 + V \left(|\varphi|^2 \right) \right] . \tag{6.2.10c}$$

6.3 Single boson star models

6.3.1 Static, spherically symmetric models

We now restrict to the case of spherical symmetry and seek a solution with a static metric in *polar-areal* coordinates[3] of the form

$$ds^2 = -\mathrm{e}^{2a(r)} \, \mathrm{d}t^2 + \mathrm{e}^{2b(r)} \, \mathrm{d}r^2 + r^2 \left(\mathrm{d}\theta^2 + \sin^2 \theta \, \mathrm{d}\phi^2 \right) . \tag{6.3.1}$$

It turns out that, in order to minimise the energy, the scalar field must have a harmonic dependence on time [335],

$$\varphi(r, t) = \tilde{\varphi}(r) \mathrm{e}^{\mathrm{i}\omega t} . \tag{6.3.2}$$

Fortunately, this time dependence does not break the staticity of the spacetime, since the energy momentum tensor (6.2.3) only depends on φ through $|\varphi|^2$ and $\nabla_\mu \bar{\varphi} \nabla_\nu \varphi$.

[3] This is, of course, an arbitrary choice but, since the degrees of freedom only appear in two metric components, it helps to simplify the equations.

Introducing the auxiliary variable

$$\Phi(r) = \tilde{\varphi}'(r), \tag{6.3.3}$$

the EKG system (6.2.2) reduces to a set of four coupled first-order ordinary differential equations for the variables

$$\{\tilde{\varphi}, \Phi, a, b\} \tag{6.3.4}$$

with the unknown "eigenvalue" ω. To close the system, we need to specify some boundary conditions. Since it is an eigenvalue problem, we require five such conditions which are

$$\tilde{\varphi}(0) = \varphi_0, \tag{6.3.5a}$$
$$\Phi(0) = 0, \tag{6.3.5b}$$
$$b(0) = 0, \tag{6.3.5c}$$
$$\lim_{r\to\infty} \tilde{\varphi}(r) = 0, \tag{6.3.5d}$$
$$\lim_{r\to\infty} [a(r) + b(r)] = 0. \tag{6.3.5e}$$

The first condition (6.3.5a) is just the amplitude of the scalar field at the centre of the star, φ_0 and, for a fixed potential, we will use it to parametrise the space of solutions. The second (6.3.5b) ensures φ is regular at $r = 0$, the third (6.3.5c) prevents a conical singularity at the origin, the fourth (6.3.5d) imposes spatial locality of the scalar field and the final condition (6.3.5e) ensures that we recover the Minkowski metric at spatial infinity.

6.3.2 Choice of potential

We next need to choose the specific potential that, as alluded to in the previous section, determines the character of our boson star models.

The first potential we consider, and arguably one of the simplest examples with a self-interaction term, is a quartic potential of the form

$$V\left(|\varphi|^2\right) = V_{\text{quar}}\left(|\varphi|^2\right) := m^2|\varphi|^2 + \frac{\lambda}{2}|\varphi|^4, \tag{6.3.6}$$

where m is the bosonic mass and λ is a mass-dimensionless coupling constant first considered in Ref. [318] for boson star models. We also consider the potential

$$V\left(|\varphi|^2\right) = V_{\rm sol}\left(|\varphi|^2\right) := m^2|\varphi|^2\left(1 - 2\frac{|\varphi|^2}{\sigma_0^2}\right)^2, \tag{6.3.7}$$

where σ_0 is a dimensionful coupling constant. This form of potential arises in particle physics as an example with false vacua (corresponding to $|\varphi| = \sigma_0/\sqrt{2}$) that can lead to spontaneous symmetry breaking. More importantly for us, this potential can give rise to highly compact boson star configurations [319, 315].

Following Ref. [326], we rescale the variables in order to remove constants as

$$\hat{r} = mr, \quad \hat{t} = \omega t, \quad f(\hat{r}) = a(r) - \log(\omega/m), \quad g(\hat{r}) = b(r), \tag{6.3.8a}$$

$$\psi(\hat{r}) = \sqrt{4\pi G}\tilde{\varphi}(r), \quad \Psi(\hat{r}) = \frac{\sqrt{4\pi G}}{m}\Phi(r). \tag{6.3.8b}$$

The boundary conditions for the rescaled variables are

$$\psi(0) = \psi_0, \tag{6.3.9a}$$

$$\Psi(0) = 0, \tag{6.3.9b}$$

$$g(0) = 0, \tag{6.3.9c}$$

$$\lim_{\hat{r}\to\infty} \psi(\hat{r}) = 0, \tag{6.3.9d}$$

$$\lim_{\hat{r}\to\infty} [f(\hat{r}) + g(\hat{r})] = \log(m/\omega), \tag{6.3.9e}$$

where ψ_0 is the rescaled central scalar field amplitude. Note that the rescaling has recast the eigenvalue problem as a boundary value problem with the oscillation frequency ω recoverable from Eq. (6.3.9e). For the quartic potential, we additionally rescale the coupling constant as

$$\Lambda = \frac{\lambda}{4\pi G m^2}, \tag{6.3.10}$$

so the system of equations to solve becomes

$$f'(\hat{r}) = \frac{1}{2}\left\{\hat{r}\left[\mathrm{e}^{2g(\hat{r})}\psi(\hat{r})^2\left(\mathrm{e}^{-2f(\hat{r})} - 1 - \frac{\Lambda}{2}\psi(\hat{r})^2\right) + \Psi(\hat{r})^2\right] + \frac{\mathrm{e}^{2g(\hat{r})}-1}{\hat{r}}\right\}, \tag{6.3.11a}$$

$$g'(\hat{r}) = \frac{1}{2}\left\{\hat{r}\left[\mathrm{e}^{2g(\hat{r})}\psi(\hat{r})^2\left(\mathrm{e}^{-2f(\hat{r})} + 1 + \frac{\Lambda}{2}\psi(\hat{r})^2\right) + \Psi(\hat{r})^2\right] - \frac{\mathrm{e}^{2g(\hat{r})}-1}{\hat{r}}\right\}, \tag{6.3.11b}$$

$$\psi'(\hat{r}) = \Psi(\hat{r}), \tag{6.3.11c}$$

$$\begin{aligned}\Psi'(\hat{r}) = \left[\mathrm{e}^{2g(\hat{r})}\hat{r}\psi(\hat{r})^2\left(1 + \frac{\Lambda}{2}\psi(\hat{r})^2\right) - \frac{\mathrm{e}^{2g(\hat{r})}+1}{\hat{r}}\right]\Psi(\hat{r}) + \\ \mathrm{e}^{2g(\hat{r})}\left(1 + \Lambda\psi(\hat{r})^2 - \mathrm{e}^{-2f(\hat{r})}\right)\psi(\hat{r}).\end{aligned} \tag{6.3.11d}$$

Similarly, for the solitonic potential, we rescale the coupling constant as

$$\Sigma = \sqrt{4\pi G}\sigma_0, \tag{6.3.12}$$

so the system of equations to solve becomes

$$f'(\hat{r}) = \frac{1}{2}\left\{\hat{r}\left[\mathrm{e}^{2g(\hat{r})}\psi(\hat{r})^2\left(\mathrm{e}^{-2f(\hat{r})} - \left(1 - 2\frac{\psi(\hat{r})^2}{\Sigma^2}\right)^2\right) + \Psi(\hat{r})^2\right] + \frac{\mathrm{e}^{2g(\hat{r})}-1}{\hat{r}}\right\}, \tag{6.3.13a}$$

$$g'(\hat{r}) = \frac{1}{2}\left\{\hat{r}\left[\mathrm{e}^{2g(\hat{r})}\psi(\hat{r})^2\left(\mathrm{e}^{-2f(\hat{r})} + \left(1 - 2\frac{\psi(\hat{r})^2}{\Sigma^2}\right)^2\right) + \Psi(\hat{r})^2\right] - \frac{\mathrm{e}^{2g(\hat{r})}-1}{\hat{r}}\right\}, \tag{6.3.13b}$$

$$\psi'(\hat{r}) = \Psi(\hat{r}), \tag{6.3.13c}$$

$$\begin{aligned}\Psi'(\hat{r}) = \left[\mathrm{e}^{2g(\hat{r})}\hat{r}\psi(\hat{r})^2\left(1 - 2\frac{\psi(\hat{r})^2}{\Sigma^2}\right)^2 - \frac{\mathrm{e}^{2g(\hat{r})}+1}{\hat{r}}\right]\Psi(\hat{r}) + \\ \mathrm{e}^{2g(\hat{r})}\left[1 - 8\frac{\psi(\hat{r})^2}{\Sigma^2} + 12\frac{\psi(\hat{r})^4}{\Sigma^4} - \mathrm{e}^{-2f(\hat{r})}\right]\psi(\hat{r}).\end{aligned} \tag{6.3.13d}$$

6.3.3 Shooting for the ground state

For a fixed central scalar field amplitude ψ_0, there is a sequence of eigenvalues

$$\omega = \omega^{(n)}, \qquad \omega^{(n+1)} > \omega^{(n)}, \qquad n = 0, 1, 2, \ldots \tag{6.3.14}$$

It turns out that n corresponds to the number of nodes (roots) in ψ [315, 316] with $\omega = \omega^{(0)}$ being the *ground state* and $\omega = \omega^{(n)}$ being the *nth excited state*.

Since excited states turn out to be unstable [317], hereafter we shall focus only on the ground state and drop the $^{(0)}$ superscript.

To find the ground state, a standard shooting algorithm with a binary search is used which we describe below.

1. Fix the values of $\psi(0)$, $\Psi(0)$ and $g(0)$ using Eq. (6.3.9) and choose a value of $f(0) < 0$ (corresponding to the lapse $\alpha < 1$ at the centre of the star) as an initial guess.

2. Integrate outwards. As $\hat{r} \to \infty$, it can be shown that for both Eqs. (6.3.11) and (6.3.13), the solutions asymptotically satisfy a linear equation with two solutions of the form
$$\psi \sim \frac{1}{\hat{r}} \exp\left[\pm\hat{r}\sqrt{1 - \frac{\omega^2}{m^2}}\right]. \tag{6.3.15}$$
We therefore expect any solutions of Eqs. (6.3.11) and (6.3.13) to asymptote to a linear superposition of the $\pm$ solutions in Eq. (6.3.15). Since the $+$ solutions are exponentially growing and finite precision numerics imply that there will always be some contribution from this solution, at some radius we expect this growing solution to dominate and ψ to stop decaying.

3. If there are no roots, take this initial $f(0)$ as an upper bound for the ground state $f(0)$. Otherwise, keep increasing $f(0) < 0$ until there are no roots and ψ blows up to $+\infty$ (in practice the integration is stopped once $|\psi| > \psi_0$ or Ψ is larger than some threshold).

4. Obtain a lower bound for the ground state $f(0)$ by gradually decreasing $f(0)$ until there is exactly 1 root and ψ blows up to $-\infty$. The change in the number of roots from 0 to many might occur suddenly (particularly for solitonic stars) so it may be necessary to perform a binary search (interval bisection) until there is exactly 1 root.

5. Perform a binary search to find the value of $f(0)$ that makes ψ decay to the largest radius. Because of finite precision, inevitably the solution will blow up at some finite radius but it is often the case that the precision is high enough such that the "bad" growing part of the solution can simply be replaced with the expected

decaying asymptotic behaviour (6.3.15) at some large radius (in practice we choose $\hat{r}^*$ where $\psi'(\hat{r}^*) = \Psi(\hat{r}^*) = 0$).

Because we can not fully eliminate the growing mode (6.3.15) with finite-precision numerics, the ADM mass which is given by the limit of the mass aspect function

$$\hat{M}\left(\hat{r}\right) = \frac{\hat{r}}{2}\left(1 - \mathrm{e}^{-2g(\hat{r})}\right), \tag{6.3.16}$$

as $\hat{r} \to \infty$, that is

$$\hat{M}_{\mathrm{ADM}} = \lim_{\hat{r}\to\infty} \hat{M}(\hat{r}), \tag{6.3.17}$$

is typically just calculated by evaluating $\hat{M}_{\mathrm{ADM}} \simeq \hat{M}(\hat{r}^*)$. Similarly, the frequency ω is computed using

$$\frac{\omega}{m} = \lim_{\hat{r}\to\infty} \exp\left[-f(\hat{r})\right] \simeq \exp\left[-f(\hat{r}^*)\right], \tag{6.3.18}$$

and the Noether charge (6.2.5) is computed using

$$N = \frac{1}{Gm^2}\int_0^\infty \hat{r}^2 \mathrm{e}^{g(\hat{r})-f(\hat{r})}\left[\psi(\hat{r})\right]^2 \mathrm{d}\hat{r} \simeq \frac{1}{Gm^2}\int_0^{\hat{r}^*} \hat{r}^2 \mathrm{e}^{g(\hat{r})-f(\hat{r})}\left[\psi(\hat{r})\right]^2 \mathrm{d}\hat{r} \tag{6.3.19}$$

In addition to the mass and oscillation frequency, we would also like to determine some notion of "radius" for our stars. Unlike fluid stars, boson stars do not have a sharp boundary; the matter fields do not vanish outside a bounded region but rather just decay exponentially (6.3.15). This means that there does not exist a well-defined "radius" of the star. There are several ways as to how one might get around this. One conventional option is to define the radius r_p (or rescaled radius $\hat{r}_p = m r_p$) which contains a fraction p of the mass according to the mass aspect function, where p is just smaller than unity (typically 0.99 or 0.95), that is

$$\hat{M}(\hat{r}_p) = p\hat{M}_{\mathrm{ADM}}. \tag{6.3.20}$$

Another option is to to choose the radius where ψ is $(1-p)\psi_0$. For the compactness parameter, $C = M_{\mathrm{ADM}}/r < 1/2$, if using r_p, a convenient choice is

$$C_p = \frac{pM_{\mathrm{ADM}}}{r_p}, \tag{6.3.21}$$

but an alternative is simply

$$C_{\max} = \max_{\hat{r}} \frac{\hat{M}(\hat{r})}{\hat{r}}. \tag{6.3.22}$$

Both options are empirically found to give similar numbers.

6.3.4 Properties of boson star solutions

For a fixed potential, the ground-state solutions form a one-parameter family characterised by the central scalar field amplitude Ψ_0. The mass-radius plots for two such families (one for mini boson stars and another for solitonic stars) (6.3.7) are shown in Fig. 6.1. For

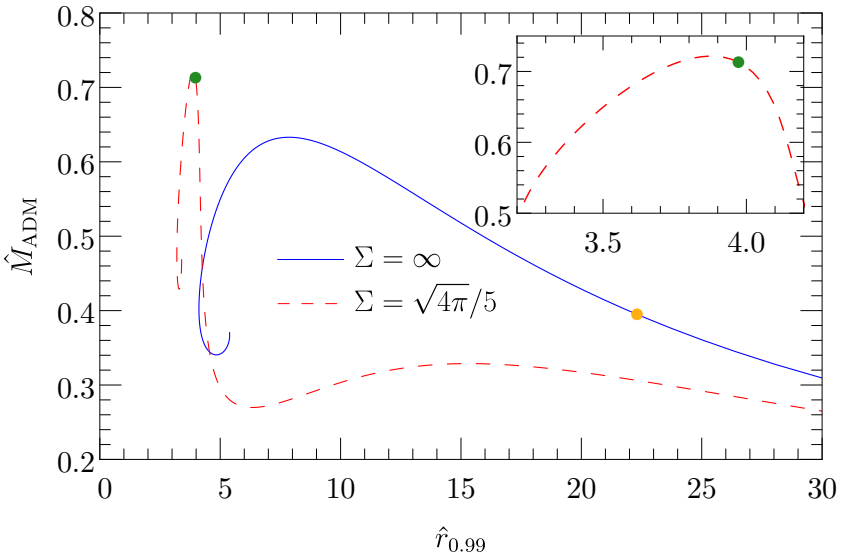


Fig. 6.1 Mass-radius plots for the one-parameter family of mini boson stars (blue solid) and a family of solitonic stars (red dashed). Both families arise from the solitonic potential (6.3.7) with the rescaled coupling constant (6.3.12) $\Sigma = \infty$ and $\sqrt{4\pi}/5$ respectively. The properties of the two models highlighted by dots are given in Table 6.1.

Table 6.1 Properties of the two boson star models used for the simulations presented in Sec. 6.4.4. Both stars arise from the solitonic potential (6.3.7). The rescaled central amplitude is ψ_0 (6.3.8), the rescaled coupling constant is Σ (6.3.12), the Noether charge is N (6.3.19), the rescaled ADM mass is $\hat{M}_{\rm ADM}$, the oscillation frequency is ω, the rescaled 99 % stellar radius (i.e. the radius containing 99 % of the mass) is $\hat{r}_{0.99}$ (6.3.20), the 99 % measure of the compactness is $C_{0.99}$ (6.3.21) and the max measure of the compactness is $C_{\rm max}$ (6.3.22).

Model	ψ_0	Σ	NGm^2	$\hat{M}_{\rm ADM}$	ω/m	$\hat{r}_{0.99}$	$C_{0.99}$	$C_{\rm max}$
mini	0.0440	∞	0.399	0.395	0.971	22.31	0.0175	0.0249
solitonic	0.603	$\sqrt{4\pi}/5$	1.05	0.713	0.439	3.97	0.180	0.222

each family, a specific model is highlighted by a point that we will discuss in Sec. 6.4.4.

These models have been selected as representative examples of a relatively compact star and a more squishy one. Note that both models lie to the right of the maximum mass in the plot and are thus stable [311, 312, 313, 314]. Their properties are shown in Table 6.1.

6.3.5 Conversion to isotropic coordinates

Before we can evolve the boson star models constructed by solving Eqs. (6.3.9) and (6.3.13), in order to minimise gauge adjustment during the evolution, we first transform the solution to *isotropic* coordinates $\{t, R, \theta, \phi\}$ where the metric takes the form

$$ds^2 = -\alpha^2\,\mathrm{d}t^2 + \chi^{-1}\left[\mathrm{d}R^2 + R^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right]. \tag{6.3.23}$$

Comparison of the polar-areal metric (6.3.1) with the above yields

$$\frac{\mathrm{d}R}{\mathrm{d}r} = \mathrm{e}^{b(r)}\frac{R}{r} = \mathrm{e}^{g(mr)}\frac{R}{r}, \qquad \chi = \frac{R^2}{r^2}, \tag{6.3.24}$$

where we have implicitly chosen $R \to +\infty$ as $r \to +\infty$. Assuming that both metrics are asymptotically Schwarzschild as $R, r \to \infty$, that is

$$\chi \sim \left(1 + \frac{M_{\mathrm{ADM}}}{2R}\right)^{-4}, \qquad \mathrm{e}^{-2b(r)} \sim 1 - \frac{2M_{\mathrm{ADM}}}{r}, \tag{6.3.25}$$

then we expect, as $r \to \infty$,

$$R \sim \frac{r}{4}\left(1 + \mathrm{e}^{-g(mr)}\right)^2. \tag{6.3.26}$$

In practice, since we are working on a finite [non-compactified] grid, we solve Eq. (6.3.23) by applying Eq. (6.3.26) at a large finite radius $r = r_{\mathrm{max}}$ and integrating inwards. The solution in this case is

$$R(r) = \frac{r_{\mathrm{max}}}{4}\left(\mathrm{e}^{-g(mr_{\mathrm{max}})} + 1\right)^2 \exp\left[-\int_r^{r_{\mathrm{max}}} \frac{\mathrm{e}^{g(mr')}}{r'}\,\mathrm{d}r'\right]. \tag{6.3.27}$$

6.4 Binary initial data

6.4.1 Boosting spacetimes

Before we superpose the initial data for two single boson stars in order to construct binary initial data, we first consider how to *Lorentz boost* an individual star. Applying

appropriate boosts enables us to perform head-on collisions in a reasonable simulation time and further allows the possibility of adding orbital angular momentum for inspirals.

6.4.1.1 Lorentz boosts

Consider an observer $\tilde{\mathcal{O}}$ with Cartesian coordinates $\tilde{x}^{\tilde{\mu}}$ moving at velocity[4] $-v^i$ with respect to another observer $\mathcal{O}$ with coordinates x^μ. Then, their coordinates are related by (assuming, without loss of generality, that their origins coincide)

$$\tilde{x}^{\tilde{\mu}} = \Lambda^{\tilde{\mu}}{}_{\alpha} x^\alpha, \tag{6.4.1a}$$

$$x^\mu = \left(\Lambda^{-1}\right)^{\mu}{}_{\tilde{\alpha}} \tilde{x}^{\tilde{\alpha}}, \tag{6.4.1b}$$

where,

$$\left(\Lambda^{\tilde{\mu}}{}_{\nu}\right) = \left(\begin{array}{c|c} \gamma & \gamma v_j \\ \hline \gamma v^{\tilde{i}} & \delta^{\tilde{i}}_j + (\gamma - 1)\frac{v^{\tilde{i}} v_j}{|\mathbf{v}|^2} \end{array}\right), \tag{6.4.2}$$

$v_i = \delta_{i\tilde{j}} v^{\tilde{j}}$, and, as is conventional[5], $\gamma = (1 - |\mathbf{v}|^2)^{-1/2}$. Similarly, the expression for $(\Lambda^{-1})^{\mu}{}_{\tilde{\nu}}$ is given by replacing $\mathbf{v} \to -\mathbf{v}$ in the right-hand side of Eq. (6.4.2) (the inverse of a boost is simply a boost in the opposite direction with the same magnitude). Hereafter, we shall refer to $\mathcal{O}$'s frame as the *rest frame* and $\tilde{\mathcal{O}}$'s frame as the *boosted frame*.

6.4.1.2 Boosting the spacetime variables

Suppose that we have expressions for the components of the spacetime metric, $g_{\mu\nu}$ and its first partial derivatives, $\partial_\rho g_{\mu\nu}$, in the rest frame as functions of the rest frame coordinates x^μ. Since we will want to evolve in the boosted frame, the coordinates on the numerical grid will be the boosted ones, $\tilde{x}^{\tilde{\mu}}$ rather than the rest frame ones x^μ, so the first step is to convert back to the rest frame coordinates using Eq. (6.4.1b).

Next, we transform the components of the metric and first derivatives using

$$\tilde{g}_{\tilde{\mu}\tilde{\nu}} = \left(\Lambda^{-1}\right)^{\alpha}{}_{\tilde{\mu}} \left(\Lambda^{-1}\right)^{\beta}{}_{\tilde{\nu}} g_{\alpha\beta} \tag{6.4.3a}$$

$$\tilde{\partial}_{\tilde{\rho}} \tilde{g}_{\tilde{\mu}\tilde{\nu}} = \left(\Lambda^{-1}\right)^{\delta}{}_{\tilde{\rho}} \left(\Lambda^{-1}\right)^{\alpha}{}_{\tilde{\mu}} \left(\Lambda^{-1}\right)^{\beta}{}_{\tilde{\nu}} \partial_\delta g_{\alpha\beta}. \tag{6.4.3b}$$

[4] We take a minus sign here so that in $\tilde{\mathcal{O}}$'s frame, objects that are stationary in $\mathcal{O}$'s frame appear to be moving with velocity v^i

[5] There is potential for confusion between the Lorentz factor and the spatial metric (2.1.5) which share the same symbol, but it should be possible to disambiguate from context.

Comparing with the 3+1 form of the metric in adapted coordinates (2.1.14) (which also applies in the boosted frame with $\{\alpha, \beta^i, \gamma_{ij}\}$ replaced by $\{\tilde{\alpha}, \tilde{\beta}^{\tilde{i}}, \tilde{\gamma}_{\tilde{i}\tilde{j}}\}$), we find

$$\tilde{\alpha} = \sqrt{-\tilde{g}_{\tilde{t}\tilde{t}} + \tilde{\beta}_{\tilde{k}}\tilde{\beta}^{\tilde{k}}}, \tag{6.4.4a}$$

$$\tilde{\beta}_{\tilde{i}} = \tilde{g}_{\tilde{t}\tilde{i}}, \tag{6.4.4b}$$

$$\tilde{\gamma}_{\tilde{i}\tilde{j}} = \tilde{g}_{\tilde{i}\tilde{j}}, \tag{6.4.4c}$$

where the shift vector, $\tilde{\beta}^{\tilde{i}}$, is obtained by first inverting $\tilde{\gamma}_{\tilde{i}\tilde{j}}$ to obtain $\tilde{\gamma}^{\tilde{i}\tilde{j}}$ and then computing

$$\tilde{\beta}^{\tilde{i}} = \tilde{\gamma}^{\tilde{i}\tilde{j}}\tilde{\beta}_{\tilde{j}}.. \tag{6.4.5}$$

We can then use Eqs. (6.4.4b) and (6.4.5) to compute the right-hand side of Eq. (6.4.4a).

Finally, we wish to compute the extrinsic curvature in the boosted frame $\tilde{K}_{\tilde{i}\tilde{j}}$. Often, in the rest frame of a stationary spacetime, one chooses a foliation (typically constant t slices) which results in a trivially vanishing extrinsic curvature. Naively, one might think the extrinsic curvature transforms similar to Eq. (6.4.3a) which would mean it would also vanish in the boosted frame. Unfortunately, this is not true as the extrinsic curvature is defined by the foliation (cf. Sec. 2.1.4). Since we will be considering slices of constant $\tilde{t} = \gamma(t + v_j x^j)$ instead of slices of constant t, the foliation is manifestly different. Instead, we use the expression for the extrinsic curvature that arises from rearranging the γ_{ij} evolution equation (2.1.38) (in the boosted frame),

$$\tilde{K}_{\tilde{i}\tilde{j}} = -\frac{1}{2\tilde{\alpha}}\left(\tilde{\partial}_{\tilde{t}}\tilde{\gamma}_{\tilde{i}\tilde{j}} - 2\tilde{D}_{(\tilde{i}}\tilde{\beta}_{\tilde{j})}\right), \tag{6.4.6}$$

where $\tilde{D}$ is the covariant derivative associated to $\tilde{\gamma}_{\tilde{i}\tilde{j}}$ which can be computed in the usual way.

6.4.1.3 Boosting the scalar field

A scalar field φ transforms trivially under a Lorentz transformation so the only thing to do is evaluate it at the rest frame coordinates obtained from the boosted ones using Eq. (6.4.1a). Since the boson stars we consider have a harmonic time dependence $\propto \mathrm{e}^{\mathrm{i}\omega t}$ (6.3.2), an easy trap to fall into[6] is substituting $t = 0$ when one actually needs to

[6]The author is guilty of having fallen into this trap.

substitute t in terms of $\tilde{x}^{\tilde{\mu}}$ at $\tilde{t}=0$ (the initial slice), which from Eq. (6.4.1a) is given by

$$t|_{\tilde{t}=0} = -\gamma v_j \tilde{x}^j. \tag{6.4.7}$$

In order to compute the auxiliary variable Π in the boosted frame, we first note that the first partial derivative of the scalar field transform as

$$\tilde{\partial}_{\tilde{\mu}}\tilde{\varphi} = \left(\Lambda^{-1}\right)^{\alpha}{}_{\tilde{\mu}}\,\partial_\alpha\varphi, \tag{6.4.8}$$

and using this, we can evaluate

$$\tilde{\Pi} = -\frac{1}{\tilde{\alpha}}\left(\tilde{\partial}_{\tilde{t}}\tilde{\varphi} - \tilde{\beta}^{\tilde{i}}\tilde{\partial}_{\tilde{i}}\tilde{\varphi}\right). \tag{6.4.9}$$

6.4.2 Superposition of two boson stars

Suppose we have two boson star solutions with centres at $\mathbf{x}=\mathbf{x}^{(A)}$, $A=1,2$ described by the quantities

$$\left\{\gamma_{ij}^{(A)}\left(x;\mathbf{x}^{(A)}\right), K_{ij}^{(A)}\left(x;\mathbf{x}^{(A)}\right),\ \varphi^{(A)}\left(x;\mathbf{x}^{(A)}\right), \Pi^{(A)}\left(x;\mathbf{x}^{(A)}\right)\right\},\quad A=1,2. \tag{6.4.10}$$

Then, the *superposition* of these two stars is given by the spacetime specified by the variables $\{\gamma_{ij}, K_{ij}, \varphi, \Pi\}$, where, pointwise,

$$\gamma_{ij} = \gamma_{ij}^{(1)} + \gamma_{ij}^{(2)} - \delta_{ij} + \delta\gamma_{ij}, \tag{6.4.11a}$$
$$K_{ij} = \gamma_{k(i}\left[K_{j)l}^{(1)}\gamma^{(1)lk} + K_{j)l}^{(2)}\gamma^{(2)lk}\right] + \delta K_{ij}, \tag{6.4.11b}$$
$$\varphi = \varphi^{(1)} + \varphi^{(2)}, \tag{6.4.11c}$$
$$\Pi = \Pi^{(1)} + \Pi^{(2)}, \tag{6.4.11d}$$

and the specified centres of the two constituent solutions, as given explicitly in Eq. (6.4.10), are understood. The correction term, $\delta\gamma_{ij}$ and δK_{ij}, are included in order to solve the constraint equations (2.1.36-2.1.37). One could similarly construct a lapse and shift by superposing the lapses and shifts that arise from the boosted single-star solutions (6.4.4a) but since they do not affect the physical dynamics of the resultant spacetime, we instead take $\alpha=\sqrt{\chi}$ and $\beta^i=0$. The choice to raise an index on the constituent extrinsic curvature tensors before adding (6.4.11b) is somewhat arbitrary but follows what is done for superposed Kerr-Schild (BH) data in Refs. [336, 337, 161].

6.4.2.1 Simple superposition

For *simple* or *naive* superposition, we make the approximation

$$\delta\gamma_{ij} = 0 = \delta K_{ij}. \tag{6.4.12}$$

Of course, this is not generally constraint-satisfying, so there is no reason, a priori, for this approach to work. Nevertheless, such an approximation has been used in numerous studies involving BH binaries [338, 339, 340] and also BS binaries [327, 329, 89, 204]. For BH spacetimes and higher dimensional ones in particular, this approach has proved remarkably successful. In some cases such as infinite initial separation, Brill-Lindquist data[7] (see Sec. 2.2.2) or Aichelburg-Sexl shockwaves (head-on BH collisions at the speed of light) [341], simple superposition is even exact. One hopes that, if the initial separation is sufficiently large, then the correction terms will be small enough so as to make the simple superposition approximation valid. Unfortunately, it turns out that there can be unintended consequences. Helfer, Lim, Garcia, and Amin [342] find that (see appendix A therein), in the case of binary oscillatons (real-scalar-field counterparts of our complex BSs), simple superposition results in spurious low-frequency amplitude modulations during the evolution (see Fig. 7 in Ref. [342]). As we will see in Sec. 6.4.4, these problems also occur for BSs. We believe that these horizonless compact objects may be more vulnerable to these superposition artefacts compared to the BHs due to the lack of a horizon with its accompanying causal barrier.

These problems arise because, for simple superposition, the metric at the centre of star 1 differs from that of the isolated star case by

$$\Delta\gamma_{ij}(0, \mathbf{x}^{(1)}) = \gamma^{(2)}_{ij}(0, \mathbf{x}^{(1)}; \mathbf{x}^{(2)}) - \delta_{ij}, \tag{6.4.13}$$

(and similarly for star 2) which only vanishes in the limit of infinite separation. This perturbation can be interpreted as a distortion of the volume element $\sqrt{\gamma}$ at the centre of each star. For realistic separations of $\mathcal{O}(100M)$[8] we have found deviations in the volume element at the centre of each star by $\mathcal{O}(1)\%$ though this should be caveated by noting that it strongly depends on the individual BS models comprising the binary. The corresponding deviation in the energy density ρ is significantly smaller due to the

[7] Note that in this case the superposition is in the conformal factor ψ rather than in the ψ^4 that appears in the metric

[8] Unfortunately, since the effect scales with $\sqrt{|\mathbf{x}^{(1)} - \mathbf{x}^{(2)}|}$, it is impractical to mitigate it by just increasing the initial separation.

exponential fall-off in the scalar field (6.3.15). Since the distortion in the volume element seems to be the dominant cause of our issues, in what follows, we will focus only on modifying $\delta\gamma_{ij}$ and continue to make the approximation $\delta K_{ij} = 0$. In principle, one might be able to consider modifications to δK_{ij} that improve the superposition further, but we believe that any benefit is likely to be very minor for the configurations of interest (with relatively small boosts). These efforts are arguably better spent in attempting to solve the constraints fully.

Finally, it is worth reiterating the point made in Sec. 2.3.3 in the context of CCZ4 constraint damping that minimising the constraints is necessary but not sufficient to find the "desired" physical solution. It is entirely possible to have well-controlled constraint violations (i.e. close to the constraint hypersurface in solution space) but not be close to the intended solution (i.e. far away from the intended physical configuration along the constraint hypersurface). In the current context, we have found that the significant undesired excitations we observe are retained even after applying CCZ4 constraint damping.

6.4.2.2 Improved superposition for identical binaries

In the case of identical binaries with individual boosts satisfying

$$v_i^{(1)} v_j^{(1)} = v_i^{(2)} v_j^{(2)} \tag{6.4.14}$$

(cf. Eq. (6.4.2)), Eq. (6.4.13) provides us with a method to resolve the problems identified above, namely choosing

$$\delta\gamma_{ij}(0, \mathbf{x}) = \delta_{ij} - \gamma_{ij}^{(2)}(0, \mathbf{x}^{(1)}; \mathbf{x}^{(2)}) = \delta_{ij} - \gamma_{ij}^{(1)}(0, \mathbf{x}^{(2)}; \mathbf{x}^{(1)}). \tag{6.4.15}$$

Note that this correction term is constant on the initial hypersurface and the second equality follows from the symmetry in our binary (and our condition on the boosts (6.4.14)). This choice means that the deviation of the metric at each star centre from the isolated star values now vanishes.

A consequence of the above choice is that the value of the superposed solution at spatial infinity is no longer δ_{ij} but rather

$$2\delta_{ij} - \gamma_{ij}^{(2)}(0, \mathbf{x}^{(1)}; \mathbf{x}^{(2)}). \tag{6.4.16}$$

This entails a slight modification to the asymptotic value in the Sommerfeld boundary condition (2.5.39).

6.4.2.3 Alternative improved superposition for stationary binaries

In the case of no boosts (but not necessarily identical stars), the extrinsic curvature vanishes identically, and the metric is conformally flat, so the only non-trivial non-matter variable to superpose is the conformal factor χ, that is, simple superposition reduces to

$$\chi = \left[\frac{1}{\chi^{(1)}(R^{(1)})} + \frac{1}{\chi^{(2)}(R^{(2)})} - 1\right]^{-1}, \tag{6.4.17}$$

where $R^{(A)} = |\mathbf{x} - \mathbf{x}^{(A)}|$.

Recall, from Sec. 6.3.5 that it is only the boundary condition of matching with standard isotropic Schwarzschild that fixes the scaling of the isotropic radius R. Let's drop this requirement and apply it later to the superposed binary.

For simplicity, assume we start with R computed as in Sec. 6.3.5 and write $\check{R} = \kappa R$, where $\check{R}$ is the *rescaled* isotropic radius and κ is the rescaling factor. In what follows, quantities with a $\check{}$, are rescaled and those without are not. In the new coordinates, the conformal factor becomes

$$\check{\chi} = \kappa^2 \chi. \tag{6.4.18}$$

By choosing

$$\kappa = \frac{\check{R}_0}{R(\check{R}_0/\sqrt{\check{\chi}_0})}, \tag{6.4.19}$$

where $R = R(r)$ is the isotropic coordinate determined as in (6.3.27) as a function of area-radius, we can set $\check{\chi} = \check{\chi}_0$ at $\check{R} = \check{R}_0$ (which corresponds to area-radius $r = \check{R}_0/\sqrt{\check{\chi}_0}$).

We now set our binary conformal factor to

$$\chi = \left[\frac{1}{\check{\chi}^{(1)}(\check{R}_1)} + \frac{1}{\check{\chi}^{(2)}(\check{R}_2)} - s\right]^{-1} = \left[\frac{1}{\kappa_1^2\chi^{(1)}(\kappa_1\check{R}_1)} + \frac{1}{\kappa_2^2\chi^{(2)}(\kappa_2\check{R}_2)} - s\right]^{-1}, \tag{6.4.20}$$

where κ_1, κ_2 and s are chosen such that

$$\frac{1}{\kappa_1^2} + \frac{1}{\kappa_2^2} - s = 1, \qquad s = \frac{1}{\check{\chi}^{(1)}(\check{R}^{(1,2)})} = \frac{1}{\check{\chi}^{(2)}(\check{R}^{(1,2)})}, \tag{6.4.21}$$

for fixed $\check{R}^{(1,2)} = |\mathbf{x}^{(1)} - \mathbf{x}^{(2)}|$. These conditions are chosen so that the conformal factor at the centre of each star is the same as in the single isolated star case and $\chi = 1$ at

spatial infinity. They can be rewritten as

$$\left(\frac{R^{(1)}(\sqrt{s}\check{R}^{(1,2)})}{\check{R}^{(1,2)}}\right)^2+\left(\frac{R^{(2)}(\sqrt{s}\check{R}^{(1,2)})}{\check{R}^{(1,2)}}\right)^2-s=1, \tag{6.4.22a}$$

$$\kappa_1=\frac{\check{R}^{(1,2)}}{R^{(1)}(\sqrt{s}\check{R}^{(1,2)})}, \tag{6.4.22b}$$

$$\kappa_2=\frac{\check{R}^{(1,2)}}{R^{(2)}(\sqrt{s}\check{R}^{(1,2)})}. \tag{6.4.22c}$$

Equation (6.4.22a) can be solved numerically to find the required value of s and then this can be substituted back into Eq. (6.4.22b) and Eq. (6.4.22c) to find the values of κ_1 and κ_2.

6.4.2.4 More general improved superposition

The previous two improved superposition methods both required some restriction on the configuration (identical stars with specific boosts and stationarity respectively) which begs the question as to how one can extend these methods to the more general case of non-identical stars and arbitrary boosts. One way to generalise is to introduce *weight functions* $w^{(A)}(\mathbf{x}),\ A=1,2$ such that the metric correction term takes the form

$$\delta\gamma_{ij}=w^{(1)}(\mathbf{x})\left(\delta_{ij}-\gamma^{(2)}_{ij}(0,\mathbf{x}^{(1)};\mathbf{x}^{(2)})\right)+w^{(2)}(\mathbf{x})\left(\delta_{ij}-\gamma^{(1)}_{ij}(0,\mathbf{x}^{(2)};\mathbf{x}^{(1)})\right). \tag{6.4.23}$$

These weight functions are included in order to 'correct' the value of the metric at the centre of each star as in Eq. (6.4.15). Furthermore, we require the weight functions to vanish at spatial infinity so that the metric asymptotes to δ_{ij} (this fixes the problem mentioned in the final paragraph of Sec. 6.4.2.2). These requirements mean that the weight functions must satisfy

1. $w^{(1)}(\mathbf{x}^{(1)})=1=w^{(2)}(\mathbf{x}^{(2)})$.
2. $w^{(1)}(\mathbf{x}^{(2)})=0=w^{(2)}(\mathbf{x}^{(1)})$.
3. $w^{(A)}(\mathbf{x})\to\infty,\ A=1,2$ as $r\to\infty$.

Such weight functions can be constructed from *bump functions* (which are smooth and compactly supported real-valued functions on $\mathbb{R}^n$). One example is given by $\zeta:\mathbb{R}^3\to\mathbb{R}$

where

$$\zeta(r) = \begin{cases} \exp(-1/(1-r^2)), & r < 1, \\ 0, & r \geq 1. \end{cases} \tag{6.4.24}$$

6.4.3 Mixed boson-star black-hole binary initial data

For simple superposition of an isotropic Schwarzschild BH (2.2.18) with a BS solution in isotropic coordinates, the BS is found to suffer from similar problems as in the BS-BS case (see Sec. 6.4.4). Even in the case of using this construction as an initial guess for solving the York-Lichnerowicz Hamiltonian constraint equation (2.2.14) in an elliptic solver, these issues can persist after relaxation. Furthermore, the uniqueness issues with matter mentioned in Sec. 2.2.4 mean that it is vital that initial guesses are close to the desired solution in order for the elliptic solver to converge to it.

Here we briefly explain the procedure to solve for mixed Bowen-York boson-star initial data including how the alternative improved superposition described in Sec. 6.4.2.3 can be adapted for the choice of initial guess in the elliptic solver alleviating some of the concerns mentioned in the previous paragraph.

We start with the York-Lichnerowicz decomposition of the constraints (2.2.13) but with the additional assumptions that the trace of the extrinsic curvature K vanishes and that the conformal metric is flat. The constraint equations become[9]

$$8\breve{\nabla}^2\bar{\psi} + \bar{\psi}^{-7}\bar{A}_{kl}\bar{A}^{kl} + 16\pi G\bar{\psi}^5\rho = 0, \tag{6.4.25a}$$

$$\breve{\nabla}_j\bar{A}^{ij} - 8\pi G\bar{\psi}^{10}j^i = 0, \tag{6.4.25b}$$

where $\breve{\nabla}$ is the flat-space derivative. We set the conformal, trace-free extrinsic curvature to the Bowen-York solution (2.2.24) that describes a BH with initial momentum $\mathbf{P}$ and spin $\mathbf{S}$ centred at $\mathbf{x} = \mathbf{x}_{\rm BH}$,

$$\bar{A}_{ij} = {}^{(\rm BY)}\bar{A}_{ij}(\mathbf{x}; \mathbf{x}_{\rm BH}, \mathbf{P}, \mathbf{S}). \tag{6.4.26}$$

Furthermore, we set the EM tensor to that of a single stationary (i.e. unboosted) BS centred at $\mathbf{x} = \mathbf{x}_{\rm BS}$. Since the BS is stationary, there is no contribution to the extrinsic curvature from it and the two terms in Eq. (6.4.25b) decouple. The first term vanishes

[9] Note that here the overbars are denoting the York-Lichenerowicz conformal decomposition as in Sec. 2.2.1 and *not* complex conjugation.

since the Bowen-York solution satisfies

$$\breve{\nabla}_k{}^{(\mathrm{BY})}\bar{A}^{ik} = 0, \tag{6.4.27}$$

and the second term vanishes since a single stationary BS has vanishing momentum density. It remains only to solve the Hamiltonian constraint (6.4.25a) as for Bowen-York data (but with some extra terms coming from the energy density).

For the Hamiltonian constraint, we use the puncture method described in Sec. 2.2.3, and write

$$\bar{\psi} = u + \bar{\psi}_{\mathrm{BL}}, \tag{6.4.28}$$

where the Brill-Lindquist part is

$$\bar{\psi}_{\mathrm{BL}} = \frac{M_{\mathrm{BH}}}{2|\mathbf{x} - \mathbf{x}_{\mathrm{BH}}|}. \tag{6.4.29}$$

Since we're using the conformal factor $\bar{\psi}$ rather than $\chi = \bar{\psi}^{-4}$, we choose as our initial guess for u,

$$u_0 = \left[\check{\chi}^{(\mathrm{BS})}(\check{R}^{(\mathrm{BS})})\right]^{-1/4} + s, \tag{6.4.30}$$

where the constant s and rescaling factor κ (6.4.19) are to be determined. If $R_{\mathrm{sep}} = |\mathbf{x}_{\mathrm{BS}} - \mathbf{x}_{\mathrm{BH}}|$ is the coordinate separation between the centres of the boson star and black hole, then, in order for the value of $\bar{\psi}_0 = u_0 + \bar{\psi}_{\mathrm{BL}}$ to be the same at the centre of the star as for the isolated star case, we must choose

$$s = -\frac{M_{\mathrm{BH}}}{2R_{\mathrm{sep}}}. \tag{6.4.31}$$

Then, in order for $\bar{\psi} \to 1$ as $R \to \infty$ (asymptotic flatness), it follows that the rescaling constant must be

$$\kappa = \left(1 + \frac{M_{\mathrm{BH}}}{2R_{\mathrm{sep}}}\right)^{-2}. \tag{6.4.32}$$

One should note that even though we have assumed the boson star is stationary in order to solve the constraints, once we have solved the constraints, we can then apply an arbitrary boost to the whole spacetime using the formulae in Sec. 6.4.1. By choosing the Bowen-York parameter $\mathbf{P}$ and this final boost appropriately, it is then possible in principle to construct a spacetime where both the BH and the BS have arbitrary boosts.

6.4.4 Comparison between simple and improved superposition for binary boson-star initial data

We now show some results which illustrate the differences between simple superposition and improved superposition described in Sec. 6.4.2.2 for boson-star binaries. For this purpose, we consider the head-on collision of two identical boson stars. We use the two models highlighted in Fig. 6.1 and Table 6.1 (one relatively squishy mini boson star and one more compact solitonic boson star). These simulations were performed with the Lean code (cf. Sec. 4.3.1.2).

For all simulations, we fix the initial boost of each star to $v^{(1)} = -v^{(2)} = v = 0.1$ along the collision axis. The simulations are then characterised by the boson star model, the type of superposition and the initial separation of the two stars d. We label the boson star model as `mini` or `soli` and prepend a `+` to denote improved superposition with its absence indicating simple superposition. The configurations of our simulations are provided in Table 6.2. We vary the initial separation in order to investigate the dependence of our results on d.

Table 6.2 The simulated binary BS configurations. The constituent BSs are either the mini or solitonic models described in Table 6.1, and they are boosted towards each other at initial velocity v. The initial data is constructed using superposition (6.4.11) of the simple (6.4.12) or improved (6.4.15) kind. The initial separations of the simulated configurations is d and the extraction radius for GWs is $r_{\rm ex}$ which we both provide in units of the total ADM mass M.

Label	Star 1	Star 2	v	Superposition	d/M	$r_{\rm ex}/M$
`mini`	mini	mini	0.1	simple	75.5, 101, 126, 151, 176	300
`+mini`	mini	mini	0.1	improved	75.5, 101, 126, 151, 176	300
`soli`	solitonic	solitonic	0.1	simple	16.7, 22.3, 27.9, 33.5, 39.1	84
`+soli`	solitonic	solitonic	0.1	improved	16.7, 22.3, 27.9, 33.5, 39.1	84

6.4.4.1 Radiated gravitational-wave energy

We start by comparing the energy radiated in GWs $E^{\rm rad}$ (2.4.79) and illustrate its dependence on the initial separation d of the binary in Fig. 6.2. As in, Sec. 5.2.3, we evaluate Eq. (2.4.79) at finite extraction radius $r = r_{\rm ex}$, where $r_{\rm ex} = 300M$ for the mini BS binaries and $r_{\rm ex} = 84M$ for the solitonic binaries. Similarly, in order to exclude the contribution from any spurious radiation in the initial data, we start the integration at $t_0 = r_{\rm ex} + 40M$. A full convergence analysis for a simulation in each sequence has been

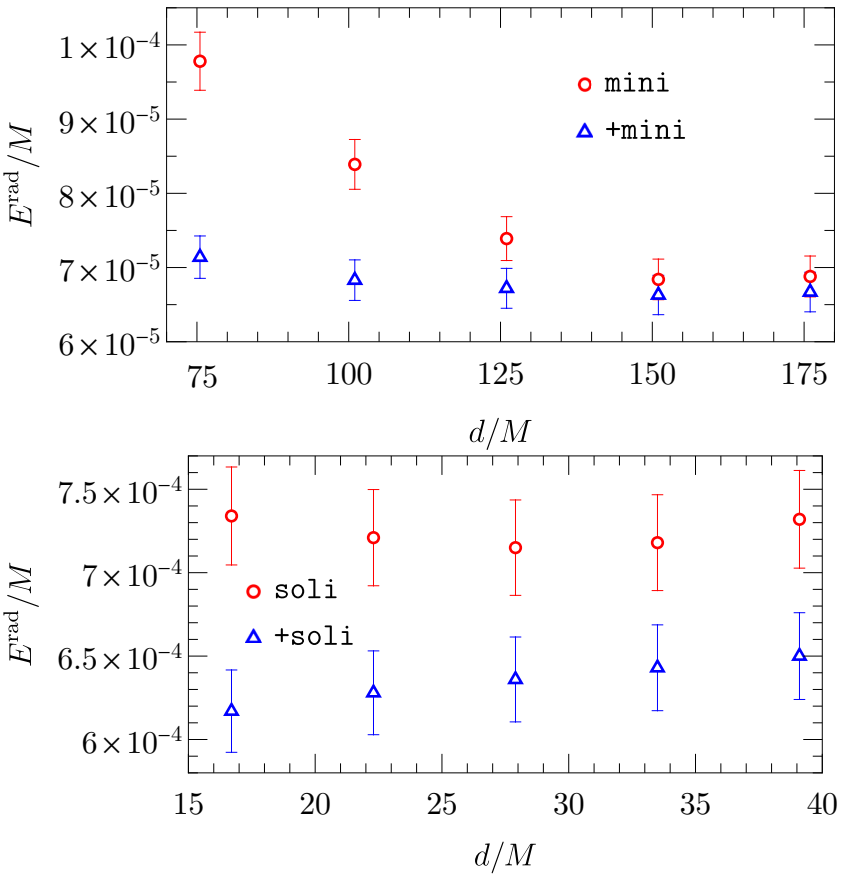


Fig. 6.2 The energy radiated in GWs $E^{\rm rad}$ (2.4.79) in units of the ADM mass M from the head-on collision of two boson stars for the configurations provided in Table 6.2.

performed and is reported more comprehensively in Ref. [4] but we briefly summarise the estimated uncertainties here. As in Appendix 5.A, the total uncertainty is comprised of the discretization error and the error due to finite-radius extraction. For the former we conservatively estimate an error of about 1 %, and, for the latter, we find the error is up to 3 % for all of our simulations. We therefore estimate the uncertainty of our simulations at around 4 %.

As the initial separation is increased, the [negative] binding energy of the binary becomes smaller in magnitude with a corresponding increase of the collision velocity around merger. However, since the binding energy vanishes in the large d limit, this effect becomes negligible. We would therefore expect $E^{\rm rad}$ to increase with d and then level off. Given the relatively large initial separations we use, we would expect any increase to be small.

The results from the `mini` configurations shown as ∘ in the top panel of Fig. 6.2 exhibit a rather different behaviour: the radiated energy rapidly decreases with d and

levels off only for $d \gtrsim 150M$. We have verified that this excess energy for smaller d is not due to extra spurious radiation. Even if the contribution from "junk" radiation were included in Fig. 6.2 (by starting the time-integration earlier), we find that this leads to differences in the radiated energy of well below 0.1 % of the total. In contrast to this behaviour, the radiated energy from the `+mini` configurations (△ in the top panel of Fig. 6.2) is approximately constant at the level of the numerical uncertainties. For $d \gtrsim 150M$, both types of initial data yield compatible results. The main improvement afforded by the adjusted superposition is that reliable results are obtained at significantly smaller initial separations which are more suitable for starting BS inspirals.

In the case of solitonic configurations (bottom panel of Fig. 6.2), the difference in the behaviour of the radiated energy as d varies between the two sequences is less pronounced with both yielding approximately constant $E^{\rm rad}$. However, the discrepancy between the calculated amounts is significant when compared to the numerical uncertainties. This discrepancy is accompanied by drastic differences in the dynamics of the collision which we explain further below.

6.4.4.2 Scalar field dynamics and gravitational collapse

The improved superposition modification in Sec. 6.4.2.2 was originally developed in Ref. [342] in order to ameliorate spurious long-wavelength oscillations in the [real] scalar field. In our case, with a complex scalar field, this effect can be readily observed in the scalar-field amplitude (or modulus) $|\varphi|$ and its consequences most dramatically affect the collision of our solitonic BSs. Recall, from Fig. 6.1 that our constituent solitonic BS model, whilst stable, is close to the instability threshold (the local maxima on the curve). It is therefore not surprising that such a star might be rather sensitive to spurious disturbances such as Eq. (6.4.13). Indeed, this is precisely what we observe with our `soli` simulations of simply-superposed initial data. In the top panel of Fig. 6.3, we show the evolution of $|\varphi|$ at the centre[10] of one of the constituent stars for the simulations of the `soli` and `+soli` configurations with initial separation $d = 22.3M$. We also show the trajectories of the stellar centres in both cases along the x-axis in the bottom panel. Consider the `soli` configuration displayed with the solid blue curves. The scalar-field amplitude initially increases up to a maximum around $t \simeq 30M$ and then rapidly decays. At around the maximum, an apparent horizon is first found with horizon mass (3.2.15)

[10]Note that the centre of each star is tracked by locally fitting an inverted Gaussian to the conformal factor χ and approximating the centre as the minimum of this fitted curve. We choose χ instead of $|\varphi|$ so that the tracking is robust to stellar collapse. Around merger, this procedure becomes inaccurate so should be interpreted as a qualitative guide only here.

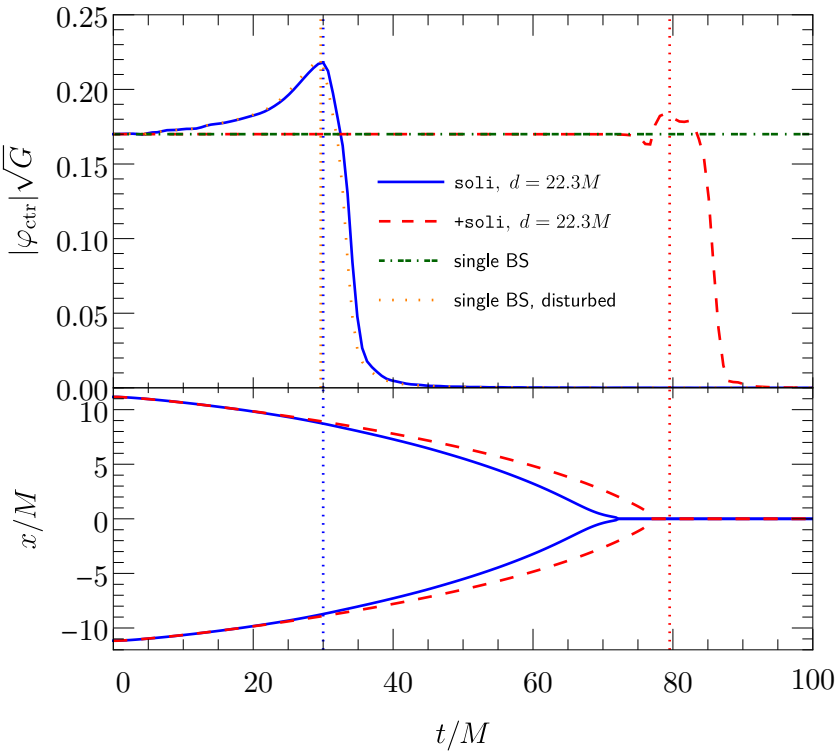


Fig. 6.3 The central scalar-field amplitude $|\varphi_{\rm ctr}|$ is plotted as a function of time for one of the constituent stars in the head-on collisions of solitonic boson stars from Table 6.2 with initial separation $d = 22.3M$. The same quantity is also shown for a single isolated solitonic BS with the same boost and another single isolated solitonic BS that is disturbed according to Eq. (6.4.13). Each dotted vertical line denotes the time an apparent horizon is first found for the simulation with the corresponding colour (note that no horizon forms for the undisturbed single BS). The bottom panel shows the coordinate trajectories of the BS centres[10] in the binary cases.

$M_{\mathscr{H}} \simeq 0.5M$; this time is marked by the vertical dotted blue line. Looking at the solid blue curve in the bottom panel, we see that the BSs are still too far away to have merged ($\hat{r}_{0.99} = 2.78M$) which means that the individual BSs have collapsed to BHs. We interpret this early gravitational collapse as an artificial consequence of the use of simple superposition.

For comparison, we plot the evolution of the central scalar-field amplitude for a single isolated solitonic BS boosted at the same velocity $v = 0.1$ displayed as the dark green dot-dashed curve in Fig. 6.3. As expected, it remains constant to relatively high precision—within $\mathcal{O}(10^{-5}/\sqrt{G})$.

In contrast to the behaviour for simply superposed initial data, the evolution of the improved initial data (`+soli`) exhibits significantly different behaviour shown by the

dashed red curves in the figure. For most of the infall, $|\varphi_{\rm ctr}|$ remains constant at the same value as the single isolated BS, and only increases mildly just after the BSs merge (when the trajectories meet at $x = 0$) before collapsing to zero. An apparent horizon of mass $M_{\mathscr{H}} = 0.99M$ is first found at $t \simeq 79.5M$ (dotted vertical red line); this approximately coincides with the maxima in the amplitude.

To further test our interpretation, we consider again the evolution of a single isolated boosted solitonic BS but, this time, disturbed by exactly the same term as for simple superposition (6.4.13). The resulting central scalar-field amplitude is shown in Fig. 6.3 as the dotted orange curve. Remarkably, this curve almost overlaps that of the `soli` binary and the time of first apparent horizon formation is almost coincident (shown by the dotted vertical orange line). This behaviour is clearly unphysical and strongly suggests that it is this disturbance (6.4.13) that is responsible for the same unphysical behaviour in the binary case. This analysis has been repeated for the other configurations in the `soli` sequence (i.e. with different d in Table 6.2) with very similar results: the constituent BSs always collapse to distinct BHs approximately $\Delta t \approx 50M$ before merger.

Finally, we note that the trajectories shown in the bottom panel of Fig. 6.3 show that the `soli` binary merges a little earlier than the `+soli` case. This is indeed a systematic effect that we observe for all initial separations. Whilst, we do not have a definitive explanation for this effect, note that the two trajectories start diverging at approximately the time of spurious collapse in the `soli` case. One explanation is that some energy in BS collisions is converted into deformation energy rather than just the kinetic energy of the stars' centres of mass resulting in a slower infall compared to the BH case. An alternative may come as a result of the repulsive nature of the scalar fields—a property which supports it against gravitational collapse and allows BSs to exist—which may slow down the infall for BSs when compared to BHs. Whatever the true explanation is, the key takeaway is that even mild disturbances in the initial data can dramatically change the ensuing dynamics.

6.5 Conclusion

In this chapter, we have presented an introduction to the numerical study of boson stars including the construction of single star models and binary initial data. Our main results at the end of the previous section concern the construction and validity of BS binary initial data. In particular, for the case of head-on collisions of identical non-spinning BSs, we have compared the commonly-used method of simple superposition (6.4.12) with one

of our improved methods, namely Eq. (6.4.15), that has first been used in Ref. [342] for oscillatons.

Our results demonstrate that the improved procedure yields significant benefits in the time evolutions of these initial data. Conversely, we find that the use of simple superposition may not only lead to incorrect quantitative results but also qualitatively different behaviour such as premature gravitational collapse. Although this improved method is remarkably effective at ameliorating some of the most severe problems with simple superposition, given its simplicity, it does have its own shortcomings. In particular, it only applies to a restricted class of configurations[11] and there still exist some residual constraint violations. We therefore believe that these adjustments should be thought of as a first step rather than the whole solution. These weaknesses immediately point towards the need to generalise the adjustments so that the symmetry restrictions can be relaxed using, for example, the procedure described in Sec. 6.4.2.4. The ultimate goal is to fully solve the constraints where these methods can be used to provide an improved initial guess.

In addition to the many possible avenues for exploration involving BS binaries using accurate initial data, there are yet more opportunities for investigations with mixed BS-BH binaries. However, we believe similar caveats with respect to the quality of rudimentary initial data may apply here, hence the brief discussion in Sec. 6.4.3. We defer the investigation of this mixed initial data to future work.

[11]The alternative improved method proposed in Sec. 6.4.2.3 yields similar results (not presented here) but is also only applicable to a highly restricted set of configurations, albeit a different one.

Chapter 7

Conclusions and outlook

In this thesis, we have looked at various aspects of gravitational-wave (GW) source modelling using numerical relativity (NR). Firstly, we have presented the current state of the GRChombo NR code and described recent enhancements including some specifically relevant to the modelling of GW sources. Next, we have discussed the use of adaptive mesh refinement (AMR) in GRChombo and the experiences we have gained from our investigations. These improvements to GRChombo and techniques for AMR have then been used to study the effect of black-hole (BH) binary eccentricity on the remnant BH recoil in unequal-mass mergers. Finally, we have considered a more exotic type of GW source in the form of boson stars (BSs) with a particular focus on the construction of binary initial data.

Over the past several years, GRChombo has matured into a capable code that achieves good performance on a wide variety of CPU architectures and can tackle a wide variety of NR problem. However, HPC systems are becoming more heterogeneous with accelerators such as Graphics Processing Units (GPUs) increasingly being relied on to provide most of the floating point performance. Until a few years ago, the NR community had largely not made the transition to GPU-accelerated codes, in part because of the complexity of the equations that we solve (2.3.30); it is challenging to obtain good performance with these on GPUs. However, if we are to continue pushing the boundary on the types of problems we can solve, there is a growing need to exploit these computational resources. The next generation of GW detectors will be able to observe currently inaccessible frequency bands that will allow observations of GWs from sources that have not yet been modelled accurately with NR (for example, compact binary mergers with more unequal mass ratios). To that end, we are currently in the process of porting GRChombo from the Chombo libraries to the AMReX library [343]

which supports GPU acceleration on all major vendor platforms. Given the large number of diagnostic tools and analysis pipelines for GRChombo that have been developed for by users, it is important that we make this transition carefully so as to make porting these tools as simple as possible. Thankfully, AMReX shares a common origin with Chombo which should make this goal more straightforward.

In addition to the transition towards GPU acceleration within the NR community, the prevalence of fully AMR codes has also grown since the origins of GRChombo. As we have remarked several times, taming the complexity of AMR is crucial to obtaining efficient and accurate simulations. Though moving-box style refinement has proved very robust for conventional astrophysical compact binaries, AMR can be particularly useful for the modelling of GW sources that reside outside this paradigm (for example, cosmic strings [183, 185]). Again, the arrival of next generation GW detectors in the decades to come stands to only increase the need to model these kinds of sources accurately.

Eccentricity is expected to play a greater role when it comes to analysis of GWs from compact binary sources that will be observed by next generation detectors. In particular, if we are to further our [currently rather limited] understanding on how black hole binaries are formed, inference of the eccentricity can provide vital insight. As a result, in the last few years, there has been renewed interest in modelling eccentric binaries [344, 345] as well as on methods of inferring eccentricity from gravitational waveforms [346, 347]. Our work in Chapter 5 provides a small taste of the richness that eccentricity adds to BH binary phenomenology. However, there is without a doubt, much more yet to uncover in this area.

Finally, BSs have a long history of providing manifestations of new physics such as extensions to the standard model of particle physics as well as possible resolutions to unexplained phenomena such as dark matter. The modelling of GWs from BSs is still in its relative infancy, especially when compared to BHs and NSs. Even when restricting to modelling of BSs using NR, there are many unexplored avenues left to pursue; these may be necessary to deduce constraints from future GW observations. Furthermore, other experiments, from those that study the very small such as the Large Hadron Collider (LHC) to the very big such as the James Webb Space Telescope (JWST) are expected to place further bounds on the role of new scalar fields in our universe. Even if, as we expect, these ground-breaking experiments place further bounds on the existence of BSs, there will always be regimes that are not reached by them where the idea of boson stars will continue to thrive.

Although the work contained in this thesis may only provide a small contribution to the long journey of understanding our universe, it is nevertheless one step further along that path.

Bibliography

[1] T. Andrade et al., *GRChombo: An adaptable numerical relativity code for fundamental physics*, Journal of Open Source Software **6**, 3703 (2021).

[2] **M. Radia**, U. Sperhake, E. Berti, and R. Croft, *Anomalies in the gravitational recoil of eccentric black-hole mergers with unequal mass ratios*, Phys. Rev. D **103**, © 2021 American Physical Society, 104006 (2021), arXiv:2101.11015.

[3] **M. Radia** et al., *Lessons for adaptive mesh refinement in numerical relativity*, Class. Quant. Grav. **39**, 135006 (2022), arXiv:2112.10567 [gr-qc].

[4] T. Helfer et al., *Malaise and remedy of binary boson-star initial data*, Class. Quant. Grav. **39**, 074001 (2022), arXiv:2108.11995 [gr-qc].

[5] A. Einstein, *Zur Allgemeinen Relativitätstheorie*, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1915**, 778–786 (1915).

[6] A. Einstein, *Die Feldgleichungen der Gravitation*, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1915**, 844–847 (1915).

[7] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, Annalen Phys. **49**, edited by J.-P. Hsu and D. Fine, 769–822 (1916).

[8] J. A. Wheeler, *Geons, black holes, and quantum foam : a life in physics* (Norton, New York ; London, 1998).

[9] K. Schwarzschild, *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie*, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), 189–196 (1916).

[10] D. Finkelstein, *Past-Future Asymmetry of the Gravitational Field of a Point Particle*, Phys. Rev. **110**, 965–967 (1958).

[11] M. D. Kruskal, *Maximal extension of Schwarzschild metric*, Phys. Rev. **119**, 1743–1745 (1960).

[12] C. A. R. Herdeiro and J. P. S. Lemos, *The black hole fifty years after: Genesis of the name*, Nov. 2018, arXiv:1811.06587 [physics.hist-ph].

[13] R. P. Kerr, *Gravitational Field of a Spinning Mass as an Example of Algebraically Special Metrics*, Phys. Rev. Lett. **11**, 237–238 (1963).

[14] E. T. Newman et al., *Metric of a Rotating, Charged Mass*, J. Math. Phys. **6**, 918–919 (1965).

[15] B. Carter, *Axisymmetric Black Hole Has Only Two Degrees of Freedom*, Phys. Rev. Lett. **26**, 331–333 (1971).

[16] D. C. Robinson, *Uniqueness of the Kerr Black Hole*, Phys. Rev. Lett. **34**, 905–906 (1975).

[17] A. Einstein, *Näherungsweise Integration der Feldgleichungen der Gravitation*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 688–696 (1916).

[18] A. Einstein, *Über Gravitationswellen*, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1918**, 154–167 (1918).

[19] P. R. Saulson, *Josh Goldberg and the physical reality of gravitational waves*, Gen. Rel. Grav. **43**, 3289–3299 (2011).

[20] D. Kennefick, *Traveling at the Speed of Thought* (Princeton University Press, 2007).

[21] R. A. Hulse and J. H. Taylor, *Discovery of a pulsar in a binary system.*, Astrophys. J. **195**, L51–L53 (1975).

[22] J. H. Taylor and J. M. Weisberg, *A new test of general relativity: Gravitational radiation and the binary pulsar PS R 1913+16*, Astrophys. J. **253**, 908–920 (1982).

[23] J. Weber, *Detection and Generation of Gravitational Waves*, Phys. Rev. **117**, 306–313 (1960).

[24] P. Astone et al., *IGEC2: A 17-month search for gravitational wave bursts in 2005-2007*, Phys. Rev. D **82**, 022003 (2010), arXiv:1002.3515 [gr-qc].

[25] M. E. Gertsenshtein and V. I. Pustovoit, *On the Detection of Low Frequency Gravitational Waves*, Sov. Phys. JETP **16**, 433 (1962).

[26] G. E. Moss, L. R. Miller, and R. L. Forward, *Photon-noise-limited laser transducer for gravitational antenna*, Appl. Opt. **10**, 2495–2498 (1971).

[27] J. Abadie et al. (LIGO Scientific, VIRGO), *Search for Gravitational Waves from Compact Binary Coalescence in LIGO and Virgo Data from S5 and VSR1*, Phys. Rev. D **82**, 102001 (2010), arXiv:1005.4655 [gr-qc].

[28] B. P. Abbott et al. (LIGO Scientific, Virgo), *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].

[29] B. P. Abbott et al. (LIGO Scientific, Virgo), *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, Phys. Rev. X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].

[30] R. Abbott et al. (LIGO Scientific, Virgo), *GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run*, Phys. Rev. X **11**, 021053 (2021), arXiv:2010.14527 [gr-qc].

[31] R. Abbott et al. (LIGO Scientific, VIRGO, KAGRA), *GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run*, tech. rep. LIGO-P2000318 (The LIGO Scientific Collaboration, the Virgo Collaboration and the KAGRA Collaboration, Nov. 2021), arXiv:2111.03606 [gr-qc].

[32] B. P. Abbott et al. (LIGO Scientific, Virgo), *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, Phys. Rev. Lett. **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].

[33] R. Abbott et al. (LIGO Scientific, VIRGO, KAGRA), *The population of merging compact binaries inferred using gravitational waves through GWTC-3*, tech. rep. LIGO-P2100239 (Nov. 2021), arXiv:2111.03634 [astro-ph.HE].

[34] S. G. Hahn and R. W. Lindquist, *The two-body problem in geometrodynamics*, Ann. Phys. **29**, 304–331 (1964).

[35] A. Čadež, *Colliding black holes*, PhD thesis (University of North Carolina at Chapel Hill, 1971).

[36] L. L. Smarr, *The Structure of General Relativity with a Numerical Illustration: The Collision of Two Black Holes.*, PhD thesis (The University of Texas at Austin, United States – Texas, 1975).

[37] K. R. Eppley, *The Numerical Evolution of the Collision of Two Black Holes.*, PhD thesis (Princeton University, United States – New Jersey, 1976).

[38] F. Pretorius, *Evolution of binary black hole spacetimes*, Phys. Rev. Lett. **95**, 121101 (2005), arXiv:gr-qc/0507014.

[39] M. Campanelli, C. Lousto, P. Marronetti, and Y. Zlochower, *Accurate evolutions of orbiting black-hole binaries without excision*, Phys. Rev. Lett. **96**, 111101 (2006), arXiv:gr-qc/0511048.

[40] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, *Gravitational wave extraction from an inspiraling configuration of merging black holes*, Phys. Rev. Lett. **96**, 111102 (2006), arXiv:gr-qc/0511103.

[41] U. Sperhake, *The numerical relativity breakthrough for binary black holes*, Class. Quant. Grav. **32**, 124011 (2015), arXiv:1411.3997.

[42] V. Kalogera and G. Baym, *The maximum mass of a neutron star*, Astrophys. J. Lett. **470**, L61–L64 (1996), arXiv:astro-ph/9608059.

[43] B. P. Abbott et al. (LIGO Scientific, Virgo), *Tests of general relativity with GW150914*, Phys. Rev. Lett. **116**, 221101 (2016), arXiv:1602.03841 [gr-qc].

[44] B. P. Abbott et al. (LIGO Scientific, Virgo), *Tests of General Relativity with GW170817*, Phys. Rev. Lett. **123**, 011102 (2019), arXiv:1811.00364 [gr-qc].

[45] B. P. Abbott et al. (LIGO Scientific, Virgo), *Tests of General Relativity with the Binary Black Hole Signals from the LIGO-Virgo Catalog GWTC-1*, Phys. Rev. D **100**, 104036 (2019), arXiv:1903.04467 [gr-qc].

[46] R. Abbott et al. (LIGO Scientific, Virgo), *Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog*, Phys. Rev. D **103**, 122002 (2021), arXiv:2010.14529 [gr-qc].

[47] R. Abbott et al. (LIGO Scientific, VIRGO, KAGRA), *Tests of General Relativity with GWTC-3*, tech. rep. LIGO-P2100275 (Dec. 2021), arXiv:2112.06861 [gr-qc].

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

[49] M. Alcubierre, *Introduction to 3+1 Numerical Relativity*, International Series of Monographs on Physics (Oxford University Press, Oxford, 2008).

[50] T. W. Baumgarte and S. L. Shapiro, *Numerical relativity: solving einstein's equations on the computer* (Cambridge University Press, 2010).

[51] É. Gourgoulhon, *3+1 Formalism in General Relativity: Bases of Numerical Relativity*, Lecture Notes in Physics (Springer-Verlag, Berlin Heidelberg, 2012).

[52] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Third (Cambridge University Press, USA, 2007).

[53] Y. Fourès-Bruhat, *Théorème d'existence pour certains systèmes d'équations aux dérivées partielles non linéaires*, Acta Math. **88**, 141–225 (1952).

[54] M. Boyle et al., *The SXS Collaboration catalog of binary black hole simulations*, Class. Quant. Grav. **36**, 195006 (2019), arXiv:1904.04831.

[55] J. Winicour, *Characteristic Evolution and Matching*, Living Rev. Relativ. **15**, 2 (2012).

[56] R. M. Wald, *General Relativity* (Chicago University Press, Jan. 1984).

[57] L. Smarr and J. W. York, *Kinematical conditions in the construction of spacetime*, Phys. Rev. D **17**, 2529–2551 (1978).

[58] S. Frittelli, *Note on the propagation of the constraints in standard 3+1 general relativity*, Phys. Rev. D **55**, 5992–5996 (1997).

[59] C. Bona, T. Ledvinka, C. Palenzuela, and M. Zacek, *General covariant evolution formalism for numerical relativity*, Phys. Rev. D **67**, 104005 (2003), arXiv:gr-qc/0302083.

[60] C. Bona, C. Bona-Casas, and C. Palenzuela, *Action principle for Numerical Relativity evolution systems*, Phys. Rev. D **82**, 124010 (2010), arXiv:1008.0747.

[61] C. Gundlach, J. M. Martin-Garcia, G. Calabrese, and I. Hinder, *Constraint damping in the Z4 formulation and harmonic gauge*, Class. Quant. Grav. **22**, 3767–3774 (2005), arXiv:gr-qc/0504114.

[62] J. M. Bowen and J. W. York, *Time asymmetric initial data for black holes and black hole collisions*, Phys. Rev. D **21**, 2047–2056 (1980).

[63] A. Lichnerowicz, *L'integration des équations de la gravitation relativiste et le problème des n corps*, Journal de mathématiques pures et appliquées **23**, 37–63 (1944).

[64] J. W. York, *Gravitational Degrees of Freedom and the Initial-Value Problem*, Phys. Rev. Lett. **26**, 1656–1658 (1971).

[65] J. W. York, *Role of Conformal Three-Geometry in the Dynamics of Gravitation*, Phys. Rev. Lett. **28**, 1082–1085 (1972).

[66] J. W. York, *Conformally invariant orthogonal decomposition of symmetric tensors on Riemannian manifolds and the initial-value problem of general relativity*, J. Math. Phys. **14**, 456–464 (1973).

[67] J. W. J. York, *Covariant decompositions of symmetric tensors in the theory of gravitation*, Annales de l'I.H.P. Physique théorique **21**, 319–332 (1974).

[68] J. W. York, *Conformal 'thin sandwich' data for the initial-value problem*, Phys. Rev. Lett. **82**, 1350–1353 (1999), arXiv:gr-qc/9810051.

[69] H. P. Pfeiffer and J. W. York, *Extrinsic curvature and the Einstein constraints*, Phys. Rev. D **67**, 044022 (2003), arXiv:gr-qc/0207095.

[70] G. B. Cook, *Initial data for numerical relativity*, Living Rev. Rel. **3**, 5 (2000), arXiv:gr-qc/0007085.

[71] D. R. Brill and R. W. Lindquist, *Interaction Energy in Geometrostatics*, Phys. Rev. **131**, 471–476 (1963).

[72] R. W. Lindquist, *Initial-Value Problem on Einstein-Rosen Manifolds*, J. Math. Phys. **4**, 938–950 (1963).

[73] J. M. Bowen, *General form for the longitudinal momentum of a spherically symmetric source*, Gen. Relativity Gravitation **11**, 227–231 (1979).

[74] J. W. York Jr., *Initial data for collisions of black holes and other gravitational miscellany*, in (Cambridge University Press, Cambridge, England, Jan. 1989), p. 89.

[75] S. Brandt and B. Bruegmann, *A Simple construction of initial data for multiple black holes*, Phys. Rev. Lett. **78**, 3606–3609 (1997), arXiv:gr-qc/9703066.

[76] A. Garat and R. H. Price, *Nonexistence of conformally flat slices of the Kerr space-time*, Phys. Rev. D **61**, 124011 (2000), arXiv:gr-qc/0002013.

[77] T. W. Baumgarte, N. O. Murchadha, and H. P. Pfeiffer, *The Einstein constraints: Uniqueness and non-uniqueness in the conformal thin sandwich approach*, Phys. Rev. D **75**, 044009 (2007), arXiv:gr-qc/0610120.

[78] H. P. Pfeiffer and J. W. York, *Uniqueness and non-uniqueness in the Einstein constraints*, Phys. Rev. Lett. **95**, 091101 (2005), arXiv:gr-qc/0504142.

[79] H.-a. Shinkai, *Formulations of the Einstein equations for numerical simulations*, J. Korean Phys. Soc. **54**, 2513–2528 (2009).

[80] O. Sarbach and M. Tiglio, *Continuum and Discrete Initial-Boundary Value Problems and Einstein's Field Equations*, Living Rev. Relativ. **15**, 9 (2012).

[81] T. Nakamura, K. Oohara, and Y. Kojima, *General Relativistic Collapse to Black Holes and Gravitational Waves from Black Holes*, Prog. Theor. Phys. Suppl. **90**, 1–218 (1987).

[82] M. Shibata and T. Nakamura, *Evolution of three-dimensional gravitational waves: Harmonic slicing case*, Phys. Rev. D **52**, 5428–5444 (1995).

[83] T. W. Baumgarte and S. L. Shapiro, *On the numerical integration of Einstein's field equations*, Phys. Rev. D **59**, 024007 (1998), arXiv:gr-qc/9810065.

[84] D. Alic, C. Bona-Casas, C. Bona, L. Rezzolla, and C. Palenzuela, *Conformal and covariant formulation of the Z4 system with constraint-violation damping*, Phys. Rev. D **85**, 064040 (2012), arXiv:1106.2254 [gr-qc].

[85] D. Alic, W. Kastaun, and L. Rezzolla, *Constraint damping of the conformal and covariant formulation of the Z4 system in simulations of binary neutron stars*, Phys. Rev. D **88**, 064049 (2013), arXiv:1307.7391 [gr-qc].

[86] T. W. Baumgarte, P. J. Montero, I. Cordero-Carrion, and E. Muller, *Numerical Relativity in Spherical Polar Coordinates: Evolution Calculations with the BSSN Formulation*, Phys. Rev. D **87**, 044026 (2013), arXiv:1211.6632.

[87] P. Marronetti, W. Tichy, B. Bruegmann, J. Gonzalez, and U. Sperhake, *High-spin binary black hole mergers*, Phys. Rev. D **77**, 064010 (2008), arXiv:0709.2160.

[88] K. Clough et al., *GRChombo : Numerical Relativity with Adaptive Mesh Refinement*, Class. Quant. Grav. **32**, 245011 (2015), arXiv:1503.03436 [gr-qc].

[89] M. Bezares, C. Palenzuela, and C. Bona, *Final fate of compact boson star mergers*, Phys. Rev. D **95**, 124005 (2017), arXiv:1705.01071.

[90] M. Alcubierre and B. Bruegmann, *Simple excision of a black hole in (3+1)-numerical relativity*, Phys. Rev. D **63**, 104006 (2001), arXiv:gr-qc/0008067.

[91] H.-J. Yo, T. W. Baumgarte, and S. L. Shapiro, *Improved numerical stability of stationary black hole evolution calculations*, Phys. Rev. D **66**, 084026 (2002), arXiv:gr-qc/0209066.

[92] O. Sarbach, G. Calabrese, J. Pullin, and M. Tiglio, *Hyperbolicity of the BSSN system of Einstein evolution equations*, Phys. Rev. D **66**, 064002 (2002), arXiv:gr-qc/0205064.

[93] C. Gundlach and J. M. Martin-Garcia, *Well-posedness of formulations of the Einstein equations with dynamical lapse and shift conditions*, Phys. Rev. D **74**, 024016 (2006).

[94] S. Bernuzzi and D. Hilditch, *Constraint violation in free evolution schemes: Comparing BSSNOK with a conformal decomposition of Z4*, Phys. Rev. D **81**, 084003 (2010), arXiv:0912.2920 [gr-qc].

[95] M. Kunesch, *Numerical simulations of instabilities in general relativity*, PhD thesis (University of Cambridge, Oct. 2018).

[96] M. Ruiz, D. Hilditch, and S. Bernuzzi, *Constraint preserving boundary conditions for the Z4c formulation of general relativity*, Phys. Rev. D **83**, 024025 (2011), arXiv:1010.0523.

[97] M. Ruiz, R. Takahashi, M. Alcubierre, and D. Nunez, *Multipole expansions for energy and momenta carried by gravitational waves*, Gen. Rel. Grav. **40**, 2467 (2008), arXiv:0707.4654 [gr-qc].

[98] C. Bona, J. Masso, E. Seidel, and J. Stela, *A New formalism for numerical relativity*, Phys. Rev. Lett. **75**, 600–603 (1995), arXiv:gr-qc/9412071.

[99] M. Alcubierre, *Hyperbolic slicings of space-time: Singularity avoidance and gauge shocks*, Class. Quant. Grav. **20**, 607–624 (2003), arXiv:gr-qc/0210050.

[100] J. R. van Meter, J. G. Baker, M. Koppitz, and D.-I. Choi, *How to move a black hole without excision: Gauge conditions for the numerical evolution of a moving puncture*, Phys. Rev. D **73**, 124011 (2006), arXiv:gr-qc/0605030.

[101] M. Alcubierre et al., *Gauge conditions for long term numerical black hole evolutions without excision*, Phys. Rev. D **67**, 084023 (2003), arXiv:gr-qc/0206072.

[102] N. Rosato, J. Healy, and C. O. Lousto, *Adapted gauge to small mass ratio binary black hole evolutions*, Phys. Rev. D **103**, 104068 (2021), arXiv:2103.09326.

[103] B. Bruegmann et al., *Calibration of Moving Puncture Simulations*, Phys. Rev. D **77**, 024027 (2008), arXiv:gr-qc/0610128.

[104] M. Hannam, S. Husa, F. Ohme, B. Bruegmann, and N. O'Murchadha, *Wormholes and trumpets: The Schwarzschild spacetime for the moving-puncture generation*, Phys. Rev. D **78**, 064020 (2008), arXiv:0804.0628 [gr-qc].

[105] M. Hannam, S. Husa, D. Pollney, B. Bruegmann, and N. O'Murchadha, *Geometry and regularity of moving punctures*, Phys. Rev. Lett. **99**, 241102 (2007), arXiv:gr-qc/0606099.

[106] N. T. Bishop and L. Rezzolla, *Extraction of Gravitational Waves in Numerical Relativity*, Living Rev. Rel. **19**, 2 (2016), arXiv:1606.02532 [gr-qc].

[107] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman, Sept. 1973).

[108] H. Weyl, *Reine Infinitesimalgeometrie*, Math. Z. **2**, 384–411 (1918).

[109] J. M. Martín-García, *xAct: Efficient tensor computer algebra for the Wolfram Language*, 2002.

[110] E. Newman and R. Penrose, *An Approach to Gravitational Radiation by a Method of Spin Coefficients*, J. Math. Phys. **3**, 566–578 (1962).

[111] W. Kinnersley, *Type D Vacuum Metrics*, J. Math. Phys. **10**, 1195–1203 (1969).

[112] S. A. Teukolsky, *Perturbations of a Rotating Black Hole. I. Fundamental Equations for Gravitational, Electromagnetic, and Neutrino-Field Perturbations*, Astrophys. J. **185**, 635–648 (1973).

[113] J. G. Baker, M. Campanelli, and C. O. Lousto, *The Lazarus project: A Pragmatic approach to binary black hole evolutions*, Phys. Rev. D **65**, 044001 (2002), arXiv:gr-qc/0104063.

[114] D. R. Fiske, J. G. Baker, J. R. van Meter, D.-I. Choi, and J. M. Centrella, *Wave zone extraction of gravitational radiation in three-dimensional numerical relativity*, Phys. Rev. D **71**, 104036 (2005), arXiv:gr-qc/0503100.

[115] E. T. Newman and R. Penrose, *Note on the Bondi-Metzner-Sachs Group*, J. Math. Phys. **7**, 863–870 (1966).

[116] R. A. Isaacson, *Gravitational Radiation in the Limit of High Frequency. II. Nonlinear Terms and the Effective Stress Tensor*, Phys. Rev. **166**, 1272–1280 (1968).

[117] M. Campanelli and C. O. Lousto, *Second order gauge invariant gravitational perturbations of a Kerr black hole*, Phys. Rev. D **59**, 124022 (1999), arXiv:gr-qc/9811019.

[118] M. Boyle, L. Lindblom, H. Pfeiffer, M. Scheel, and L. E. Kidder, *Testing the accuracy and stability of spectral methods in numerical relativity*, Phys. Rev. D **75**, 024006 (2007), arXiv:gr-qc/0609047.

[119] M. Ansorg, B. Bruegmann, and W. Tichy, *A Single-domain spectral method for black hole puncture data*, Phys. Rev. D **70**, 064011 (2004), arXiv:gr-qc/0404056.

[120] P. D. Lax and R. D. Richtmyer, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math. **9**, 267–293 (1956).

[121] J. G. Charney, R. FjÖrtoft, and J. V. Neumann, *Numerical Integration of the Barotropic Vorticity Equation*, Tellus **2**, 237–254 (1950).

[122] R. Courant, H. Lewy, and K. Friedrichs, *Über die partiellen Differenzengleichungen der mathematischen Physik*, Math. Ann. **100**, 32–74 (1928).

[123] R. Courant, K. Friedrichs, and H. Lewy, *On the Partial Difference Equations of Mathematical Physics*, IBM J. Res. Dev. **11**, 215–234 (1967).

[124] H.-O. Kreiss and G. Scherer, *Method of Lines for Hyperbolic Differential Equations*, SIAM J. Numer. Anal. **29**, 640–646 (1992).

[125] A. Sommerfeld, *Die Greensche funktion der schwingungsgleichung*, J.-Ber. Deutsch Math.-Verein **21**, 309–353 (1912).

[126] L. F. Richardson, *The Approximate Arithmetical Solution by Finite Differences of Physical Problems Involving Differential Equations, with an Application to the Stresses in a Masonry Dam*, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character **210**, 307–357 (1911).

[127] M. Adams et al., *Chombo Software Package for AMR Applications - Design Document*, tech. rep. LBNL-6616E (Lawrence Berkeley National Laboratory, 2019).

[128] H. Kreiss and J. Oliger, *Methods for the approximate solution of time dependant problems*, GARP Publication Series 10 (WMO, Geneva, 1973).

[129] S. Husa, J. A. Gonzalez, M. Hannam, B. Bruegmann, and U. Sperhake, *Reducing phase error in long numerical binary black hole evolutions with sixth order finite differencing*, Class. Quant. Grav. **25**, 105006 (2008), arXiv:0706.0740 [gr-qc].

[130] G. M. Morton, *A computer oriented geodetic data base and a new technique in file sequencing*, tech. rep. (1966).

[131] S. Husa, I. Hinder, and C. Lechner, *Kranc: A Mathematica application to generate numerical codes for tensorial evolution equations*, Comput. Phys. Commun. **174**, 983–1004 (2006), arXiv:gr-qc/0404023.

[132] I. Ruchlin, Z. B. Etienne, and T. W. Baumgarte, *SENR/NRPy+: Numerical Relativity in Singular Curvilinear Coordinate Systems*, Phys. Rev. D **97**, 064036 (2018), arXiv:1712.07658.

[133] M. Berger and I. Rigoutsos, *An algorithm for point clustering and grid generation*, IEEE Trans. Sys. Man & Cybernet. **21**, 1278–1286 (1991).

[134] M. J. Berger and P. Colella, *Local adaptive mesh refinement for shock hydrodynamics*, J. Comput. Phys. **82**, 64–84 (1989).

[135] P. McCorquodale and P. Colella, *A high-order finite-volume method for conservation laws on locally refined grids*, Commun. Appl. Math. Comput. Sci. **6**, 1–25 (2011).

[136] E. Schnetter, S. H. Hawley, and I. Hawke, *Evolutions in 3-D numerical relativity using fixed mesh refinement*, Class. Quant. Grav. **21**, 1465–1488 (2004), arXiv:gr-qc/0310042.

[137] M. Bondarescu, *Orbifolds in numerical relativity*, Mar. 2006, arXiv:gr-qc/0603047.

[138] J. Bamber, K. Clough, P. G. Ferreira, L. Hui, and M. Lagos, *Growth of accretion driven scalar hair around Kerr black holes*, Phys. Rev. D **103**, 044059 (2021), arXiv:2011.07870 [gr-qc].

[139] D. Traykova et al., *Dynamical friction from scalar dark matter in the relativistic regime*, Phys. Rev. D **104**, 103014 (2021), arXiv:2106.08280 [gr-qc].

[140] K. Clough, *Continuity equations for general matter: applications in numerical relativity*, Class. Quant. Grav. **38**, 10.1088/1361-6382/ac10ee (2021), arXiv:2104.13420 [gr-qc].

[141] B. Fornberg, *Classroom Note:Calculation of Weights in Finite Difference Formulas*, SIAM Rev. **40**, 685–691 (1998).

[142] S. Tunyasuvunakool, *Applications of Numerical Relativity Beyond Astrophysics*, Thesis (Department of Applied Mathematics and Theoretical Physics, University of Cambridge, 2017).

[143] S.-D. Poisson, *Mémoire sur le calcul numérique des intégrales défines*, Mémoires de l'Académie Royale des sciences de l'Institut de France. **6**, 571–602 (1827).

[144] A. Drew, *Cosmic String Radiation with Adaptive Mesh Refinement*, Thesis (University of Cambridge, May 2021).

[145] S. Hawking and G. Ellis, *The large scale structure of space-time*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 1973).

[146] J. Thornburg, *Event and Apparent Horizon Finders for 3 + 1 Numerical Relativity*, Living Rev. Relativ. **10**, 3 (2007), arXiv:gr-qc/0512169.

[147] S. Balay et al., *PETSc Web page*, https://www.mcs.anl.gov/petsc, 2021.

[148] S. Balay et al., *PETSc users manual*, tech. rep. ANL-95/11 - Revision 3.15 (Argonne National Laboratory, 2021).

[149] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith, *Efficient management of parallelism in object oriented numerical software libraries*, in Modern software tools in scientific computing, edited by E. Arge, A. M. Bruaset, and H. P. Langtangen (1997), pp. 163–202.

[150] O. Dreyer, B. Krishnan, E. Schnetter, and D. Shoemaker, *Introduction to Isolated Horizons in Numerical Relativity*, Phys. Rev. D **67**, 024018 (2003), arXiv:gr-qc/0206008.

[151] M. Campanelli, C. O. Lousto, Y. Zlochower, B. Krishnan, and D. Merritt, *Spin Flips and Precession in Black-Hole-Binary Mergers*, Phys. Rev. D **75**, 064030 (2007), arXiv:gr-qc/0612076.

[152] D. Christodoulou, *Reversible and Irreversible Transformations in Black-Hole Physics*, Phys. Rev. Lett. **25**, 1596–1597 (1970).

[153] C. Reisswig and D. Pollney, *Notes on the integration of numerical relativity waveforms*, Class. Quant. Grav. **28**, 195015 (2011), arXiv:1006.1632 [gr-qc].

[154] M. Frigo and S. G. Johnson, *The design and implementation of FFTW3*, Proc. IEEE **93**, Special issue on "Program Generation, Optimization, and Platform Adaptation", 216–231 (2005).

[155] B. Gough, *GNU Scientific Library Reference Manual - Third Edition*, 3rd (Network Theory Ltd., 2009).

[156] Y. Zlochower, J. Baker, M. Campanelli, and C. Lousto, *Accurate black hole evolutions by fourth-order numerical relativity*, Phys. Rev. D **72**, 024021 (2005), arXiv:gr-qc/0505055.

[157] T. Goodale et al., *The Cactus framework and toolkit: design and applications*, in Vector and parallel processing – vecpar'2002, 5th international conference, lecture notes in computer science (2003).

[158] F. Loffler et al., *The Einstein Toolkit: A Community Computational Infrastructure for Relativistic Astrophysics*, Class. Quant. Grav. **29**, 115001 (2012), arXiv:1111.3344 [gr-qc].

[159] J. Brown, P. Diener, O. Sarbach, E. Schnetter, and M. Tiglio, *Turduckening black holes: An Analytical and computational study*, Phys. Rev. D **79**, 044023 (2009), arXiv:0809.3533 [gr-qc].

[160] F. Herrmann, I. Hinder, D. Shoemaker, and P. Laguna, *Unequal mass binary black hole plunges and gravitational recoil*, Classical Quantum Gravity. **24**, S33–S42 (2007).

[161] U. Sperhake, *Binary black-hole evolutions of excision and puncture data*, Phys. Rev. D **76**, 104015 (2007), arXiv:gr-qc/0606079.

[162] M. Zilhao et al., *Numerical relativity for D dimensional axially symmetric spacetimes: formalism and code tests*, Phys. Rev. D **81**, 084052 (2010), arXiv:1001.2302 [gr-qc].

[163] H. Witek et al., *Canuda: a public numerical relativity library to probe fundamental physics*, Zenodo, Oct. 2021.

[164] B. Brügmann, *Adaptive mesh and geodesically sliced Schwarzschild spacetime in 3+1 dimensions*, Phys. Rev. D **54**, gr-qc/9608050, 7361–7372 (1996).

[165] M. Thierfelder, S. Bernuzzi, and B. Bruegmann, *Numerical relativity simulations of binary neutron stars*, Phys. Rev. D **84**, 044012 (2011), arXiv:1104.4751 [gr-qc].

[166] P. Marronetti et al., *Binary black holes on a budget: Simulations using workstations*, Class. Quant. Grav. **24**, edited by M. Campanelli and L. Rezzolla, S43–S58 (2007), arXiv:gr-qc/0701123.

[167] B. P. Abbott et al. (KAGRA, LIGO Scientific, Virgo, VIRGO), *Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA*, Living Rev. Rel. **21**, 3 (2018), arXiv:1304.0670 [gr-qc].

[168] B. Szilágyi, *Key Elements of Robustness in Binary Black Hole Evolutions using Spectral Methods*, Int. J. Mod. Phys. D **23**, 1430014 (2014), arXiv:1405.3693 [gr-qc].

[169] P. Kumar et al., *Template Banks for Binary black hole searches with Numerical Relativity waveforms*, Phys. Rev. D **89**, 042002 (2014), arXiv:1310.7949 [gr-qc].

[170] B. Szilágyi et al., *Approaching the Post-Newtonian Regime with Numerical Relativity: A Compact-Object Binary Simulation Spanning 350 Gravitational-Wave Cycles*, Phys. Rev. Lett. **115**, 031102 (2015), arXiv:1502.04953 [gr-qc].

[171] A. V. Kravtsov, A. A. Klypin, and A. M. Khokhlov, *Adaptive refinement tree: A New high resolution N body code for cosmological simulations*, Astrophys. J. Suppl. **111**, 73 (1997), arXiv:astro-ph/9701195.

[172] R. Teyssier, *Cosmological hydrodynamics with adaptive mesh refinement: a new high resolution code called ramses*, Astron. Astrophys. **385**, 337–364 (2002), arXiv:astro-ph/0111367.

[173] C. Barrera-Hinojosa and B. Li, *GRAMSES: a new route to general relativistic N-body simulations in cosmology. Part I. Methodology and code description*, JCAP **01**, 007 (2020), arXiv:1905.08890 [astro-ph.CO].

[174] M. J. Berger and J. Oliger, *Adaptive mesh refinement for hyperbolic partial differential equations*, J. Comp. Phys. **53**, 484–512 (1984).

[175] W. E. East, *Cosmic Censorship Upheld in Spheroidal Collapse of Collisionless Matter*, Phys. Rev. Lett. **122**, 231103 (2019), arXiv:1901.04498 [gr-qc].

[176] F. Pretorius and W. E. East, *Black Hole Formation from the Collision of Plane-Fronted Gravitational Waves*, Phys. Rev. D **98**, 084053 (2018), arXiv:1807.11562 [gr-qc].

[177] L. Lehner and F. Pretorius, *Black Strings, Low Viscosity Fluids, and Violation of Cosmic Censorship*, Phys. Rev. Lett. **105**, 101102 (2010), arXiv:1006.5960 [hep-th].

[178] P. Figueras and T. França, *Gravitational Collapse in Cubic Horndeski Theories*, Class.Quant.Grav. **37**, 225009 (2020), arXiv:2006.09414.

[179] T. Andrade, P. Figueras, and U. Sperhake, *Evidence for violations of Weak Cosmic Censorship in black hole collisions in higher dimensions*, JHEP **03**, 111 (2022), arXiv:2011.03049 [hep-th].

[180] H. Bantilan, P. Figueras, M. Kunesch, and R. Panosso Macedo, *End point of nonaxisymmetric black hole instabilities in higher dimensions*, Phys. Rev. D **100**, 086014 (2019), arXiv:1906.10696 [hep-th].

[181] P. Figueras, M. Kunesch, L. Lehner, and S. Tunyasuvunakool, *End Point of the Ultraspinning Instability and Violation of Cosmic Censorship*, Phys. Rev. Lett. **118**, 151103 (2017), arXiv:1702.01755 [hep-th].

[182] P. Figueras, M. Kunesch, and S. Tunyasuvunakool, *End Point of Black Ring Instabilities and the Weak Cosmic Censorship Conjecture*, Phys. Rev. Lett. **116**, 071102 (2016), arXiv:1512.04532 [hep-th].

[183] A. Drew and E. P. S. Shellard, *Radiation from global topological strings using adaptive mesh refinement: Methodology and massless modes*, Phys. Rev. D **105**, 063517 (2022), arXiv:1910.01718 [gr-qc].

[184] T. Helfer, J. C. Aurrekoetxea, and E. A. Lim, *Cosmic String Loop Collapse in Full General Relativity*, Phys. Rev. D **99**, 104028 (2019), arXiv:1808.06678 [gr-qc].

[185] J. C. Aurrekoetxea, T. Helfer, and E. A. Lim, *Coherent Gravitational Waveforms and Memory from Cosmic String Loops*, Class. Quant. Grav. **37**, 204001 (2020), arXiv:2002.05177 [gr-qc].

[186] W. E. East, M. Kleban, A. Linde, and L. Senatore, *Beginning inflation in an inhomogeneous universe*, JCAP **09**, 010 (2016), arXiv:1511.05143 [hep-th].

[187] W. G. Cook, I. A. Glushchenko, A. Ijjas, F. Pretorius, and P. J. Steinhardt, *Supersmoothing through Slow Contraction*, Phys. Lett. B **808**, 135690 (2020), arXiv:2006.01172 [gr-qc].

[188] A. Ijjas, W. G. Cook, F. Pretorius, P. J. Steinhardt, and E. Y. Davies, *Robustness of slow contraction to cosmic initial conditions*, JCAP **08**, 030 (2020), arXiv:2006.04999 [gr-qc].

[189] A. Ijjas, A. P. Sullivan, F. Pretorius, P. J. Steinhardt, and W. G. Cook, *Ultralocality and slow contraction*, JCAP **06**, 013 (2021), arXiv:2103.00584 [gr-qc].

[190] C. Joana and S. Clesse, *Inhomogeneous preinflation across Hubble scales in full general relativity*, Phys. Rev. D **103**, 083501 (2021), arXiv:2011.12190 [astro-ph.CO].

[191] J. C. Aurrekoetxea, K. Clough, R. Flauger, and E. A. Lim, *The Effects of Potential Shape on Inhomogeneous Inflation*, JCAP **05**, 030 (2020), arXiv:1910.12547 [astro-ph.CO].

[192] K. Clough, R. Flauger, and E. A. Lim, *Robustness of Inflation to Large Tensor Perturbations*, JCAP **05**, 065 (2018), arXiv:1712.07352 [hep-th].

[193] K. Clough et al., *Robustness of Inflation to Inhomogeneous Initial Conditions*, JCAP **09**, 025 (2017), arXiv:1608.04408 [hep-th].

[194] E. de Jong, J. C. Aurrekoetxea, and E. A. Lim, *Primordial black hole formation with full numerical relativity*, Sept. 2021, arXiv:2109.04896 [astro-ph.CO].

[195] F. Pretorius, *Numerical relativity using a generalized harmonic decomposition*, Class. Quant. Grav. **22**, 425–452 (2005), arXiv:gr-qc/0407110.

[196] F. Pretorius and M. W. Choptuik, *Adaptive mesh refinement for coupled elliptic-hyperbolic systems*, J. Comput. Phys. **218**, 246–274 (2006), arXiv:gr-qc/0508110.

[197] W. E. East, F. Pretorius, and B. C. Stephens, *Hydrodynamics in full general relativity with conservative AMR*, Phys. Rev. D **85**, 124010 (2012), arXiv:1112.3094 [gr-qc].

[198] P. MacNeice, K. Olson, C. Mobarry, R. de Fainchtein, and C. Packer, *PARAMESH: A Parallel Adaptive Mesh Refinement Community Toolkit*, Computer Physics Comm. **136**, 330 (2000).

[199] M. Anderson, E. Hirschmann, S. L. Liebling, and D. Neilsen, *Relativistic MHD with Adaptive Mesh Refinement*, Class. Quant. Grav. **23**, 6503–6524 (2006), arXiv:gr-qc/0605102.

[200] H. P. Pfeiffer, L. E. Kidder, M. A. Scheel, and S. A. Teukolsky, *A Multidomain spectral method for solving elliptic equations*, Comput. Phys. Commun. **152**, 253–273 (2003), arXiv:gr-qc/0202096.

[201] D. Hilditch, A. Weyhausen, and B. Brügmann, *Pseudospectral method for gravitational wave collapse*, Phys. Rev. D **93**, 063006 (2016), arXiv:1504.04732 [gr-qc].

[202] C. Meringolo, S. Servidio, and P. Veltri, *A spectral method algorithm for numerical simulations of gravitational fields*, Class. Quant. Grav. **38**, 075027 (2021), arXiv:2011.10354 [gr-qc].

[203] J. B. Mertens, J. T. Giblin, and G. D. Starkman, *Integration of inhomogeneous cosmological spacetimes in the BSSN formalism*, Phys. Rev. D **93**, 124059 (2016), arXiv:1511.01106 [gr-qc].

[204] C. Palenzuela et al., *A Simflowny-based finite-difference code for high-performance computing in numerical relativity*, Class. Quant. Grav. **35**, 185007 (2018), arXiv:1806.04182 [physics.comp-ph].

[205] R. D. Hornung and S. R. Kohn, *Managing application complexity in the samrai object-oriented framework*, Concurrency and Computation: Practice and Experience **14**, 347–368 (2002).

[206] B. T. Gunney and R. W. Anderson, *Advances in patch-based adaptive mesh refinement scalability*, J. Parallel Distr. Com. **89**, 65–84 (2016).

[207] *SAMRAI project website*, 2021.

[208] B. Daszuta et al., *GR-Athena++: Puncture Evolutions on Vertex-centered Oct-tree Adaptive Mesh Refinement*, Astrophys. J. Supp. **257**, 25 (2021), arXiv:2101.08289 [gr-qc].

[209] P. C.-K. Cheong, A. T.-L. Lam, H. H.-Y. Ng, and T. G. F. Li, *Gmunu: paralleled, grid-adaptive, general-relativistic magnetohydrodynamics in curvilinear geometries in dynamical space–times*, Mon.Not.Roy.Astron.Soc. **508**, 2279–2301 (2021), arXiv:2012.07322 [astro-ph.IM].

[210] M. Fernando, D. Neilsen, H. Lim, E. Hirschmann, and H. Sundar, *Massively Parallel Simulations of Binary Black Hole Intermediate-Mass-Ratio Inspirals*, SIAM J. Sci. Comput. **41**, C97–C138 (2019), arXiv:1807.06128 [gr-qc].

[211] N. Deppe et al., *SpECTRE*, version 2021.05.03, May 2021.

[212] L. E. Kidder et al., *SpECTRE: A Task-based Discontinuous Galerkin Code for Relativistic Astrophysics*, J. Comput. Phys. **335**, 84–114 (2017), arXiv:1609.00098 [astro-ph.HE].

[213] Z. Cao, P. Fu, L.-W. Ji, and Y. Xia, *Application of local discontinuous Galerkin method to Einstein equations*, Int. J. Mod. Phys. D **28**, 1950014 (2018).

[214] L.-W. Ji, R.-G. Cai, and Z. Cao, *Binary black hole simulation with an adaptive finite element method +III: Evolving a single black hole*, May 2018, arXiv:1805.10642 [gr-qc].

[215] F. Pretorius, *Numerical simulations of gravitational collapse*, PhD thesis (University of British Columbia, Vancouver, Canada, 2002).

[216] S. Khan, *(private communications)*, 2019.

[217] U. Sperhake, R. Roşca-Mead, D. Gerosa, and E. Berti, *Amplification of superkicks in black-hole binaries through orbital eccentricity*, Phys. Rev. D **101**, 024044 (2020), arXiv:1910.01598 [gr-qc].

[218] J. A. González, M. D. Hannam, U. Sperhake, B. Brügmann, and S. Husa, *Supermassive recoil velocities for binary black-hole mergers with antialigned spins*, Phys. Rev. Lett. **98**, 231101 (2007), arXiv:gr-qc/0702052.

[219] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt, *Maximum gravitational recoil*, Phys. Rev. Lett. **98**, 231102 (2007), arXiv:gr-qc/0702133.

[220] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt, *Large merger recoils and spin flips from generic black-hole binaries*, Astrophys. J. Lett. **659**, L5–L8 (2007), arXiv:gr-qc/0701164 [gr-qc].

[221] J. Thornburg, *Finding apparent horizons in numerical relativity*, Phys. Rev. D **54**, 4899–4918 (1996), eprint: gr-qc/9508014.

[222] J. Thornburg, *A Fast apparent horizon finder for three-dimensional Cartesian grids in numerical relativity*, Classical Quantum Gravity. **21**, 743–766 (2004), arXiv:gr-qc/0306056.

[223] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, *Improved filters for gravitational waves form inspiraling compact binaries*, Phys. Rev. D **57**, 885–907 (1998).

[224] M. A. Miller, *Accuracy requirements for the calculation of gravitational waveforms from coalescing compact binaries in numerical relativity*, Phys. Rev. D **71**, 104016 (2005), arXiv:gr-qc/0502087.

[225] B. J. Owen, *Search templates for gravitational waves from inspiraling binaries: Choice of template spacing*, Phys. Rev. D **53**, 6749–6761 (1996), arXiv:gr-qc/9511032.

[226] L. Lindblom, B. J. Owen, and D. A. Brown, *Model Waveform Accuracy Standards for Gravitational Wave Data Analysis*, Phys. Rev. D **78**, 124020 (2008), arXiv:0809.3844 [gr-qc].

[227] L. Barsotti, P. Fritschel, M. Evans, and S. Gras, *Updated advanced ligo sensitivity design curve*, tech. rep. T1800044-v5 (The LIGO Scientific Collaboration, 2018).

[228] D. Shoemaker, *Advanced ligo anticipated sensitivity curves*, tech. rep. T0900288-v3 (The LIGO Scientific Collaboration, 2010).

[229] M. Alcubierre et al., *Numerical studies of Phi**2 oscillatons*, Class. Quant. Grav. **20**, 2883–2904 (2003), arXiv:gr-qc/0301105.

[230] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, Phys. Rev. Lett. **38**, 1440–1443 (1977).

[231] S. Weinberg, *A New Light Boson?*, Phys. Rev. Lett. **40**, 223–226 (1978).

[232] K. Clough, T. Dietrich, and J. C. Niemeyer, *Axion star collisions with black holes and neutron stars in full 3D numerical relativity*, Phys. Rev. D **98**, 083020 (2018), arXiv:1808.04668 [gr-qc].

[233] D. J. E. Marsh, *Axion Cosmology*, Phys. Rept. **643**, 1–79 (2016), arXiv:1510.07633 [astro-ph.CO].

[234] L. A. Urena-Lopez, T. Matos, and R. Becerril, *Inside oscillatons*, Class. Quant. Grav. **19**, 6259–6277 (2002).

[235] T. Helfer et al., *Black hole formation from axion stars*, JCAP **03**, 055 (2017), arXiv:1609.04724 [astro-ph.CO].

[236] L. A. Urena-Lopez, *Oscillatons revisited*, Class. Quant. Grav. **19**, 2617–2632 (2002), arXiv:gr-qc/0104093.

[237] S. L. Liebling, *The Nonlinear sigma model with distributed adaptive mesh refinement*, Class. Quant. Grav. **21**, 3995 (2004), arXiv:gr-qc/0403076.

[238] Y. T. Liu, Z. B. Etienne, and S. L. Shapiro, *Evolution of near-extremal-spin black holes using the moving puncture technique*, Phys. Rev. D **80**, 121503 (2009), arXiv:1001.4077.

[239] M. Campanelli, C. O. Lousto, and Y. Zlochower, *Spinning-black-hole binaries: The orbital hang up*, Phys. Rev. D **74**, 041501(R) (2006), arXiv:gr-qc/0604012.

[240] U. Sperhake et al., *Cross section, final spin and zoom-whirl behavior in high-energy black hole collisions*, Phys. Rev. Lett. **103**, 131102 (2009), arXiv:0907.1252 [gr-qc].

[241] W. B. Bonnor and M. A. Rotenberg, *Transport of momentum by gravitational waves: the linear approximation*, Proc. R. Soc. A **265**, 109–116 (1961).

[242] A. Peres, *Classical Radiation Recoil*, Phys. Rev. **128**, 2471–2475 (1962).

[243] J. D. Bekenstein, *Gravitational-Radiation Recoil and Runaway Black Holes*, Astrophys. J. **183**, 657–664 (1973).

[244] D. Gerosa and C. J. Moore, *Black hole kicks as new gravitational wave observables*, Phys. Rev. Lett. **117**, 011101 (2016), arXiv:1606.04226 [gr-qc].

[245] J. Calderón Bustillo, J. A. Clark, P. Laguna, and D. Shoemaker, *Tracking black hole kicks from gravitational wave observations*, Phys. Rev. Lett. **121**, 191102 (2018), arXiv:1806.11160 [gr-qc].

[246] C. O. Lousto and J. Healy, *Kicking gravitational wave detectors with recoiling black holes*, Phys. Rev. D **100**, 104039 (2019), eprint: arXiv:1908.04382[gr-qc].

[247] V. Varma, M. Isi, and S. Biscoveanu, *Extracting the Gravitational Recoil from Black Hole Merger Signals*, Phys. Rev. Lett. **124**, 101104 (2020), arXiv:2002.00296 [gr-qc].

[248] V. Cardoso and C. F. B. Macedo, *Drifting through the medium: kicks and self-propulsion of binaries within accretion discs and other environments*, Mon. Not. R. Astron. Soc. **498**, 1963–1972 (2020), arXiv:2008.01091.

[249] M. J. Fitchett, *The influence of gravitational wave momentum losses on the centre of mass motion of a Newtonian binary system*, Mon. Not. R. Astron. Soc. **203**, 1049–1062 (1983).

[250] L. Blanchet, M. S. Qusailah, and C. M. Will, *Gravitational recoil of inspiralling black-hole binaries to second post-Newtonian order*, Astrophys. J. **635**, 508 (2005), arXiv:astro-ph/0507692.

[251] S. A. Hughes, M. Favata, and D. E. Holz, *How black holes get their kicks: Radiation recoil in binary black hole mergers*, in Conference on Growing Black Holes: Accretion in a Cosmological Context (Aug. 2004), arXiv:astro-ph/0408492.

[252] T. Damour and A. Gopakumar, *Gravitational recoil during binary black hole coalescence using the effective one body approach*, Phys. Rev. D **73**, 124006 (2006), arXiv:gr-qc/0602117.

[253] C. F. Sopuerta, N. Yunes, and P. Laguna, *Gravitational Recoil from Binary Black Hole Mergers: The Close-Limit Approximation*, Phys. Rev. D **74**, [Erratum: Phys.Rev.D 75, 069903(E) (2007), Erratum: Phys.Rev.D 78, 049901 (2008)], 124010 (2006), arXiv:astro-ph/0608600.

[254] C. F. Sopuerta, N. Yunes, and P. Laguna, *Gravitational recoil velocities from eccentric binary black hole mergers*, Astrophys. J. Lett. **656**, L9–L12 (2007), arXiv:astro-ph/0611110.

[255] A. Le Tiec, L. Blanchet, and C. M. Will, *Gravitational-Wave Recoil from the Ringdown Phase of Coalescing Black Hole Binaries*, Classical Quantum Gravity. **27**, 012001 (2010), arXiv:0910.4594 [gr-qc].

[256] J. G. Baker et al., *Getting a kick out of numerical relativity*, Astrophys. J. Lett. **653**, L93–L96 (2006), arXiv:astro-ph/0603204.

[257] J. A. Gonzalez, U. Sperhake, B. Bruegmann, M. Hannam, and S. Husa, *Total recoil: The Maximum kick from nonspinning black-hole binary inspiral*, Phys. Rev. Lett. **98**, 091101 (2007), arXiv:gr-qc/0610154.

[258] C. O. Lousto and Y. Zlochower, *Hangup Kicks: Still Larger Recoils by Partial Spin/Orbit Alignment of Black-Hole Binaries*, Phys. Rev. Lett. **107**, 231102 (2011), arXiv:1108.2009 [gr-qc].

[259] J. D. Schnittman and A. Buonanno, *The Distribution of Recoil Velocities from Merging Black Holes*, Astrophys. J. Lett. **662**, L63 (2007), arXiv:astro-ph/0702641 [ASTRO-PH].

[260] M. Dotti et al., *Dual black holes in merger remnants. II: spin evolution and gravitational recoil*, Mon. Not. R. Astron. Soc. **402**, 682 (2010), arXiv:0910.5729 [astro-ph.HE].

[261] M. Kesden, U. Sperhake, and E. Berti, *Relativistic Suppression of Black Hole Recoils*, Astrophys. J. **715**, 1006–1011 (2010), arXiv:1003.4993 [astro-ph.CO].

[262] C. O. Lousto, Y. Zlochower, M. Dotti, and M. Volonteri, *Gravitational Recoil From Accretion-Aligned Black-Hole Binaries*, Phys. Rev. D **85**, 084015 (2012), arXiv:1201.1923 [gr-qc].

[263] E. Berti, M. Kesden, and U. Sperhake, *Effects of post-Newtonian Spin Alignment on the Distribution of Black-Hole Recoils*, Phys. Rev. D **85**, 124049 (2012), arXiv:1203.2920 [astro-ph.HE].

[264] S. Komossa, *Recoiling black holes: electromagnetic signatures, candidates, and astrophysical implications*, Adv. Astron. **2012**, 364973 (2012), arXiv:1202.1977 [astro-ph].

[265] M. Colpi, *Massive binary black holes in galactic nuclei and their path to coalescence*, Space Sci. Rev. **183**, 189–221 (2014), arXiv:1407.3102 [astro-ph.GA].

[266] L. Blecha et al., *Recoiling black holes: prospects for detection and implications of spin alignment*, Mon. Not. R. Astron. Soc. **456**, 961–989 (2016), arXiv:1508.01524 [astro-ph.GA].

[267] L. Barack et al., *Black holes, gravitational waves and fundamental physics: a roadmap*, Classical Quantum Gravity. **36**, 143001 (2019), arXiv:1806.05195 [gr-qc].

[268] D. Merritt, M. Milosavljevic, M. Favata, S. A. Hughes, and D. E. Holz, *Consequences of gravitational radiation recoil*, Astrophys. J. Lett. **607**, L9–L12 (2004), arXiv:astro-ph/0402057.

[269] M. J. Benacquista and J. M. Downing, *Relativistic Binaries in Globular Clusters*, Living Rev. Relaltivity **16**, 4 (2013), arXiv:1110.4423 [astro-ph.SR].

[270] J. Morawski, M. Giersz, A. Askar, and K. Belczynski, *MOCCA-SURVEY Database I: Assessing GW kick retention fractions for BH–BH mergers in globular clusters*, Mon. Not. R. Astron. Soc. **481**, 2168–2179 (2018), arXiv:1802.01192 [astro-ph.GA].

[271] Z. Haiman, *Constraints from gravitational recoil on the growth of supermassive black holes at high redshift*, Astrophys. J. **613**, 36–40 (2004), arXiv:astro-ph/0404196.

[272] M. J. Graham et al., *Candidate Electromagnetic Counterpart to the Binary Black Hole Merger Gravitational Wave Event S190521g*, Phys. Rev. Lett. **124**, 251102 (2020), arXiv:2006.14122 [astro-ph.HE].

[273] H.-Y. Chen, C.-J. Haster, S. Vitale, W. M. Farr, and M. Isi, *A standard siren cosmological measurement from the potential GW190521 electromagnetic counterpart ZTF19abanrhr*, Mon. Not. Roy. Astron. Soc. **513**, 2152–2157 (2022), arXiv:2009.14057 [astro-ph.CO].

[274] J. Samsing and E. Ramirez-Ruiz, *On the Assembly Rate of Highly Eccentric Binary Black Hole Mergers*, Astrophys. J. Lett. **840**, L14 (2017), arXiv:1703.09703 [astro-ph.HE].

[275] J. Samsing, *Eccentric Black Hole Mergers Forming in Globular Clusters*, Phys. Rev. D **97**, 103014 (2018), arXiv:1711.07452 [astro-ph.HE].

[276] J. Samsing, A. Askar, and M. Giersz, *MOCCA-SURVEY Database. I. Eccentric Black Hole Mergers during Binary–Single Interactions in Globular Clusters*, Astrophys. J. **855**, 124 (2018), arXiv:1712.06186 [astro-ph.HE].

[277] C. L. Rodriguez et al., *Post-Newtonian Dynamics in Dense Star Clusters: Formation, Masses, and Merger Rates of Highly-Eccentric Black Hole Binaries*, Phys. Rev. D **98**, 10.1103/PhysRevD.98.123005 (2018), arXiv:1811.04926.

[278] J. Samsing et al., *Active Galactic Nuclei as Factories for Eccentric Black Hole Mergers*, Oct. 2020, arXiv:2010.09765 [astro-ph.HE].

[279] H. Tagawa et al., *Eccentric Black Hole Mergers in Active Galactic Nuclei*, Astrophys.J.Lett. **907**, L20 (2021).

[280] V. Cardoso, C. F. B. Macedo, and R. Vicente, *Eccentricity evolution of compact binaries and applications to gravitational-wave physics*, Phys. Rev. D **103**, 10.1103/PhysRevD.103.023015 (2021), arXiv:2010.15151.

[281] B. Abbott et al. (LIGO Scientific and Virgo Collaborations), *Search for Eccentric Binary Black Hole Mergers with Advanced LIGO and Advanced Virgo during their First and Second Observing Runs*, Astrophys. J. **883**, 149 (2019), arXiv:1907.09384 [astro-ph.HE].

[282] R. Abbott et al. (LIGO Scientific and Virgo Collaborations), *GW190521: A Binary Black Hole Merger with a Total Mass of 150 $M_\odot$*, Phys. Rev. Lett. **125**, 101102 (2020), arXiv:2009.01075 [gr-qc].

[283] I. M. Romero-Shaw, P. D. Lasky, E. Thrane, and J. C. Bustillo, *GW190521: orbital eccentricity and signatures of dynamical formation in a binary black hole merger signal*, Astrophys.J.Lett. **903**, L5 (2020), arXiv:2009.04771.

[284] V. Gayathri et al., *GW190521 as a Highly Eccentric Black Hole Merger*, Sept. 2020, arXiv:2009.05461 [astro-ph.HE].

[285] A. Nishizawa, E. Berti, A. Klein, and A. Sesana, *eLISA eccentricity measurements as tracers of binary black hole formation*, Phys. Rev. D **94**, 064020 (2016), arXiv:1605.01341 [gr-qc].

[286] K. Breivik, C. L. Rodriguez, S. L. Larson, V. Kalogera, and F. A. Rasio, *Distinguishing Between Formation Channels for Binary Black Holes with LISA*, Astrophys. J. Lett. **830**, L18 (2016), arXiv:1606.09558 [astro-ph.GA].

[287] A. Nishizawa, A. Sesana, E. Berti, and A. Klein, *Constraining stellar binary black hole formation scenarios with eLISA eccentricity measurements*, Mon. Not. R. Astron. Soc. **465**, 4375–4380 (2017), arXiv:1606.09295 [astro-ph.HE].

[288] C. Roedig and A. Sesana, *Origin and Implications of high eccentricities in massive black hole binaries at sub-pc scales*, J. Phys. Conf. Ser. **363**, 012035 (2012), arXiv:1111.3742 [astro-ph.CO].

[289] M. Bonetti, A. Sesana, F. Haardt, E. Barausse, and M. Colpi, *Post-Newtonian evolution of massive black hole triplets in galactic nuclei – IV. Implications for LISA*, Mon. Not. Roy. Astron. Soc. **486**, 4044–4060 (2019), arXiv:1812.01011 [astro-ph.GA].

[290] E. Huerta et al., *Physics of eccentric binary black hole mergers: A numerical relativity perspective*, Phys. Rev. D **100**, 064003 (2019), arXiv:1901.07038 [gr-qc].

[291] V. Paschalidis, Z. B. Etienne, R. Gold, and S. L. Shapiro, *An efficient spectral interpolation routine for the TwoPunctures code*, Apr. 2013, arXiv:1304.0457 [astro-ph, physics:gr-qc].

[292] U. Sperhake et al., *Eccentric binary black-hole mergers: The Transition from inspiral to plunge in general relativity*, Phys. Rev. D **78**, 064069 (2008), arXiv:0710.3823 [gr-qc].

[293] N. Loutrel, S. Liebersbach, N. Yunes, and N. Cornish, *The eccentric behavior of inspiralling compact binaries*, Classical Quantum Gravity. **36**, 025004 (2019), arXiv:1810.03521 [gr-qc].

[294] R.-M. Memmesheimer, A. Gopakumar, and G. Schaefer, *Third post-Newtonian accurate generalized quasi-Keplerian parametrization for compact binaries in eccentric orbits*, Phys. Rev. D **70**, 104011 (2004), arXiv:gr-qc/0407049.

[295] J. Healy, C. O. Lousto, and Y. Zlochower, *Nonspinning binary black hole merger scenario revisited*, Phys. Rev. D **96**, 024031 (2017), arXiv:1705.07034 [gr-qc].

[296] B. Brügmann, J. A. González, M. D. Hannam, S. Husa, and U. Sperhake, *Exploring black hole superkicks*, Phys. Rev. D **77**, 124047 (2008), arXiv:0707.0135 [gr-qc].

[297] A. G. Wiseman, *Coalescing binary systems of compact objects to (post)5/2 Newtonian order. 2. Higher order wave forms and radiation recoil*, Phys. Rev. D **46**, 1517–1539 (1992).

[298] C. O. Lousto, H. Nakano, Y. Zlochower, and M. Campanelli, *Statistical studies of Spinning Black-Hole Binaries*, Phys. Rev. D **81**, [Erratum: Phys. Rev. D 82, 129902 (2010)], 084023 (2010), arXiv:0910.3197 [gr-qc].

[299] M. Kesden, D. Gerosa, R. O'Shaughnessy, E. Berti, and U. Sperhake, *Effective potentials and morphological transitions for binary black-hole spin precession*, Phys. Rev. Lett. **114**, 081103 (2015), arXiv:1411.0674 [gr-qc].

[300] D. Gerosa, M. Kesden, U. Sperhake, E. Berti, and R. O'Shaughnessy, *Multi-timescale analysis of phase transitions in precessing black-hole binaries*, Phys. Rev. D **92**, 064016 (2015), arXiv:1506.03492 [gr-qc].

[301] C. Hamilton and R. R. Rafikov, *Secular dynamics of binaries in stellar clusters - I. General formulation and dependence on cluster potential*, Mon. Not. R. Astron. Soc. **488**, 5489–5511 (2019), arXiv:1902.01344.

[302] C. Hamilton and R. R. Rafikov, *Secular dynamics of binaries in stellar clusters - II. Dynamical evolution*, Mon. Not. R. Astron. Soc. **488**, 5512–5535 (2019), arXiv:1902.01345.

[303] J. A. Wheeler, *Geons*, Phys. Rev. **97**, 511–536 (1955).

[304] D. A. Feinblum and W. A. McKinley, *Stable States of a Scalar Particle in Its Own Gravational Field*, Phys. Rev. **168**, 1445–1450 (1968).

[305] D. J. Kaup, *Klein-Gordon Geon*, Phys. Rev. **172**, 1331–1342 (1968).

[306] R. Ruffini and S. Bonazzola, *Systems of Self-Gravitating Particles in General Relativity and the Concept of an Equation of State*, Phys. Rev. **187**, 1767–1783 (1969).

[307] R. Brito, V. Cardoso, C. A. R. Herdeiro, and E. Radu, *Proca stars: Gravitating Bose–Einstein condensates of massive spin 1 particles*, Phys. Lett. B **752**, 291–295 (2016), arXiv:1508.05395 [gr-qc].

[308] M. Alcubierre et al., *ℓ-Boson stars*, Class. Quant. Grav. **35**, 10.1088/1361-6382/aadcb6 (2018), arXiv:1805.11488 [gr-qc].

[309] R. C. Tolman, *Static Solutions of Einstein's Field Equations for Spheres of Fluid*, Phys. Rev. **55**, 364–373 (1939).

[310] J. R. Oppenheimer and G. M. Volkoff, *On Massive Neutron Cores*, Phys. Rev. **55**, 374–381 (1939).

[311] J. D. Breit, S. Gupta, and A. Zaks, *Cold Bose stars*, Phys. Lett. B **140**, 329–332 (1984).

[312] M. Gleiser and R. Watkins, *Gravitational Stability of Scalar Matter*, Nucl. Phys. B **319**, 733–746 (1989).

[313] T. D. Lee and Y. Pang, *Stability of mini-boson stars*, Nucl. Phys. B **315**, 477–516 (1989).

[314] E. Seidel and W.-M. Suen, *Dynamical Evolution of Boson Stars. 1. Perturbing the Ground State*, Phys. Rev. D **42**, 384–403 (1990).

[315] T. D. Lee and Y. Pang, *Nontopological solitons*, Phys. Rept. **221**, edited by H.-C. Ren and Y. Pang, 251–350 (1992).

[316] P. Jetzer, *Boson stars*, Phys. Rep. **220**, 163–227 (1992).

[317] J. Balakrishna, E. Seidel, and W.-M. Suen, *Dynamical evolution of boson stars. 2. Excited states and selfinteracting fields*, Phys. Rev. D **58**, 104004 (1998), arXiv:gr-qc/9712064.

[318] M. Colpi, S. L. Shapiro, and I. Wasserman, *Boson Stars: Gravitational Equilibria of Self-Interacting Scalar Fields*, Phys. Rev. Lett. **57**, 2485–2488 (1986).

[319] T. D. Lee, *Soliton stars and the critical masses of black holes*, Phys. Rev. D **35**, 3637–3639 (1987).

[320] F. E. Schunck and D. F. Torres, *Boson stars with generic selfinteractions*, Int. J. Mod. Phys. D **9**, 601–618 (2000), arXiv:gr-qc/9911038.

[321] B. Hartmann, B. Kleihaus, J. Kunz, and I. Schaffer, *Compact Boson Stars*, Phys.Lett.B **714**, 120–126 (2012), arXiv:1205.0899 [gr-qc].

[322] J. C. Bustillo et al., *GW190521 as a Merger of Proca Stars: A Potential New Vector Boson of -13 eV*, Phys. Rev. Lett. **126**, 081101 (2021), arXiv:2009.05376 [gr-qc].

[323] N. Sennett, T. Hinderer, J. Steinhoff, A. Buonanno, and S. Ossokine, *Distinguishing Boson Stars from Black Holes and Neutron Stars from Tidal Interactions in Inspiraling Binary Systems*, Phys. Rev. D **96**, 024002 (2017).

[324] F. Di Giovanni et al., *Dynamical bar-mode instability in spinning bosonic stars*, Phys. Rev. D **102**, 124009 (2020), arXiv:2010.05845 [gr-qc].

[325] A. Toubiana, S. Babak, E. Barausse, and L. Lehner, *Modeling gravitational waves from exotic compact objects*, Phys. Rev. D **103**, 10.1103/PhysRevD.103.064042 (2021), arXiv:2011.12122 [gr-qc].

[326] S. L. Liebling and C. Palenzuela, *Dynamical Boson Stars*, Living Rev. Rel. **15**, 6 (2012), arXiv:1202.5809 [gr-qc].

[327] C. Palenzuela, I. Olabarrieta, L. Lehner, and S. L. Liebling, *Head-on collisions of boson stars*, Phys.Rev.D **75**, 10.1103/PhysRevD.75.064005 (2007), arXiv:gr-qc/0612067.

[328] M. W. Choptuik and F. Pretorius, *Ultra Relativistic Particle Collisions*, Phys. Rev. Lett. **104**, 111101 (2010), arXiv:0908.1780 [gr-qc].

[329] C. Palenzuela, L. Lehner, and S. L. Liebling, *Orbital Dynamics of Binary Boson Star Systems*, Phys.Rev.D **77**, 10.1103/PhysRevD.77.044036 (2008), arXiv:0706.2435.

[330] C. Palenzuela et al., *Gravitational Wave Signatures of Highly Compact Boson Star Binaries*, Phys. Rev. D **96**, 10.1103/PhysRevD.96.104058 (2017), arXiv:1710.09432 [gr-qc].

[331] J. A. Font, M. A. Miller, W.-M. Suen, and M. Tobias, *Three-dimensional numerical general relativistic hydrodynamics. 1. Formulations, methods, and code tests*, Phys. Rev. D **61**, 044011 (2000), arXiv:gr-qc/9811015.

[332] F. D. Swesty, E. Y. M. Wang, and A. C. Calder, *Numerical models of binary neutron star system mergers. 1. Numerical methods and equilibrium data for Newtonian models*, Astrophys. J. **541**, 937–958 (2000), arXiv:astro-ph/9911192.

[333] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918**, 235–257 (1918), arXiv:physics/0503066.

[334] T. Dietrich, S. Ossokine, and K. Clough, *Full 3D numerical relativity simulations of neutron star–boson star collisions with BAM*, Class. Quant. Grav. **36**, 025002 (2019), arXiv:1807.06959 [gr-qc].

[335] R. Friedberg, T. D. Lee, and Y. Pang, *Mini-soliton stars*, Phys. Rev. D **35**, 3640–3657 (1987).

[336] R. A. Matzner, M. F. Huq, and D. Shoemaker, *Initial data and coordinates for multiple black hole systems*, Phys. Rev. D **59**, 024015 (1999), arXiv:gr-qc/9805023.

[337] U. Sperhake, B. J. Kelly, P. Laguna, K. L. Smith, and E. Schnetter, *Black hole head-on collisions and gravitational waves with fixed mesh-refinement and dynamic singularity excision*, Phys.Rev.D **71**, 124042 (2005), arXiv:gr-qc/0503071.

[338] M. Shibata, H. Okawa, and T. Yamamoto, *High-velocity collision of two black holes*, Phys. Rev. D **78**, 101501 (2008), arXiv:0810.4735 [gr-qc].

[339] H. Okawa, K.-i. Nakao, and M. Shibata, *Is super-Planckian physics visible? – Scattering of black holes in 5 dimensions*, Phys. Rev. D **83**, 121501 (2011), arXiv:1105.3331 [gr-qc].

[340] U. Sperhake, W. Cook, and D. Wang, *High-energy collision of black holes in higher dimensions*, Phys. Rev. D **100**, 104046 (2019), arXiv:1909.02997 [gr-qc].

[341] P. C. Aichelburg and R. U. Sexl, *On the gravitational field of a massless particle*, Gen. Relativ. Gravitation **2**, 303–312 (1971).

[342] T. Helfer, E. A. Lim, M. A. G. Garcia, and M. A. Amin, *Gravitational Wave Emission from Collisions of Compact Scalar Solitons*, Phys. Rev. D **99**, 044046 (2019), arXiv:1802.06733 [gr-qc].

[343] W. Zhang et al., *AMReX: a framework for block-structured adaptive mesh refinement*, Journal of Open Source Software **4**, 1370 (2019).

[344] A. Ramos-Buades et al., *Eccentric binary black holes: Comparing numerical relativity and small mass-ratio perturbation theory*, Phys. Rev. D **106**, 124040 (2022), arXiv:2209.03390 [gr-qc].

[345] A. V. Joshi, S. G. Rosofsky, R. Haas, and E. A. Huerta, *Numerical relativity higher order gravitational waveforms of eccentric, spinning, non-precessing binary black hole mergers*, Oct. 2022, arXiv:2210.01852 [gr-qc].

[346] Z. Xuan, S. Naoz, and X. Chen, *Eccentricity to the rescue! Detecting Accelerating Eccentric Binaries in the LISA Band*, Oct. 2022, arXiv:2210.03129 [astro-ph.HE].

[347] A. M. Knee, I. M. Romero-Shaw, P. D. Lasky, J. McIver, and E. Thrane, *A Rosetta Stone for Eccentric Gravitational Waveform Models*, Astrophys. J. **936**, 172 (2022), arXiv:2207.14346 [gr-qc].

[348] R. Arnowitt, S. Deser, and C. W. Misner, *The dynamics of general relativity*, in Gravitation an introduction to current research, edited by L. Witten (1962), pp. 227–265, arXiv:gr-qc/0405109.

Appendix A

ADM mass and momentum

In numerical simulations, it is often useful to measure the total mass and momentum contained in the spacetime. Unfortunately, it is not sufficient to integrate the matter energy density ρ and momentum density j^i over a volume since these do not include any contributions from the curvature of the spacetime (a vacuum black-hole spacetime can contain mass but no matter). Instead, for spacetimes with asymptotically flat ends, we typically use the Arnowitt-Deser-Misner (ADM) integrals [348] which we provide expressions for below.

The *ADM mass* is given by

$$M_{\mathrm{ADM}} = \frac{1}{16\pi} \lim_{r\to\infty} \oint_{S_r^2} \delta^{kl} s^m \left[\partial_k \gamma_{lm} - \partial_m \gamma_{kl}\right] \mathrm{d}A, \tag{A.0.1}$$

where s^m is the outward pointing unit normal to S_r^2, the 2-sphere of radius r, γ_{ij} is the spatial metric and $\mathrm{d}A$ is area element of S_r^2. The *ADM linear momentum* is similarly

$$P^i_{\mathrm{ADM}} = \frac{1}{8\pi} \lim_{r\to\infty} \oint_{S_r^2} s_m \left[K^{im} - \delta^{im} K\right] \mathrm{d}A. \tag{A.0.2}$$

Finally, the *ADM angular momentum* is defined by

$$J^i_{\mathrm{ADM}} = \frac{1}{8\pi} \lim_{r\to\infty} \oint_{S_r^2} \epsilon^{klm} x_l s^n \left[K_{mn} - \delta_{mn} K\right] \mathrm{d}A. \tag{A.0.3}$$

It is important to note that these expressions are only valid in Cartesian-like coordinates. Furthermore, the convergence of the ADM angular momentum in the limit requires faster decay than what is typically required for an asymptotically flat end.

Since the ADM integrals are defined at spatial infinity i^0, they are unaffected by gravitational radiation as these carry energy and momentum to future null infinity $\mathscr{I}^+$. Therefore, one might expect the ADM mass, linear momentum and angular momentum to remain constant during a numerical simulation. However, in practice we use a finite non-compactified computational domain, and calculate the limits in Eqs. (A.0.1) to (A.0.3) by extrapolating the integrals evaluated at finite radii so this conservation does not appear to hold.