

DISSERTATION
SUBMITTED TO THE
COMBINED FACULTY OF MATHEMATICS, ENGINEERING AND NATURAL SCIENCES
OF HEIDELBERG UNIVERSITY, GERMANY
FOR THE DEGREE OF
DOCTOR OF NATURAL SCIENCES

PUT FORWARD BY
SHAUN DAVID BROCUS FELL
BORN IN BELLINGHAM, WASHINGTON, USA
ORAL EXAMINATION 11-11-2024

Celestial Beacons: How Black Holes Illuminate the Path to Vector Dark Matter

Referees:

Prof. Dr. Lavinia Heisenberg
Prof. Dr. Luca Amendola

ZUSAMMENFASSUNG

Die Einführung der Gravitationswellenastronomie hat die Astrophysik in eine neue Ära gestürzt. Gravitationswellensignale, die von starken Gravitationsregimen ausgehen, liefern enorme Informationen über die Umgebung aus der sie stammen. Dies ermöglicht die Durchführung bemerkenswert leistungsfähiger Studien zur grundlegenden Natur der Schwerkraft, der dunklen Energie und der dunklen Materie. Diese aktuellen und zukünftigen Studien erfordern eine sehr hohe Genauigkeit und Präzision der physikalischen Modelle, die sie entwickeln, um alle potenziellen Informationen aus den beobachteten Daten zu extrahieren. Insbesondere aufgrund der schwachen Kopplung zwischen Materie und Schwerkraft liefern sehr starke Gravitationsumgebungen die höchste Qualität an Informationen über die umgebenden fundamentalen Felder. Zu diesem Zweck besteht das Hauptziel dieser Arbeit darin, Modelle der Wechselwirkung dunkler Materie mit hochdynamischen Schwarzslochsystemen zu erstellen und die potenzielle Beobachtbarkeit mithilfe zukünftiger Gravitationswellenobservatorien vorherzusagen. Zur Förderung dieses Ziels wurden mehrere Studien und Werkzeuge entwickelt, die in neuen Toolkits der numerischen Relativität, Parameterbeschränkungen und theoretischen Ergebnissen gipfelten. Obwohl noch viel Arbeit nötig ist, um die Leistungsfähigkeit künftiger Gravitationswellenobservatorien voll auszuschöpfen, bietet die Arbeit in dieser Dissertation große Fortschritte und leistungsfähige Werkzeuge für das Verständnis der fundamentalen Bausteine des Universums.

ABSTRACT

The inception of gravitational wave astronomy has plunged astrophysical sciences into a new era. Gravitational wave signals sourced by strong gravity regimes provide enormous information about the surrounding environments from which they came. This allows remarkably powerful studies to be performed regarding the fundamental nature of gravity, dark energy, and dark matter. These current and future studies require very high accuracy and precision in the physical models they develop in order to extract all potential information from observed data. In particular, due to the feeble coupling between matter and gravity, very strong gravitational environments provide the highest quality of information regarding surrounding fundamental fields. To that end, the primary objective of this thesis is to generate models of dark matter interacting with highly dynamical black hole systems and forecast potential observability using upcoming gravitational wave observatories. Several studies and tools were developed to further this goal, culminating in new numerical relativity toolkits, parameter constraints, and theoretical results. While much more work still needs to be done to fully leverage the prowess of upcoming gravitational wave observatories, the work in this thesis provides great advancements and powerful tools in the pursuit of understanding the fundamental building blocks of the universe.

Contents

I	INTRODUCTION	I
I.1	Our Library of Knowledge	2
I.1.1	Quantum Field Theory	3
	Black-body Radiation	3
	The Quantum Revolution	4
	From Particles to Fields and Back	4
	Reigning in the Infinities	5
	Diving into the Nucleus	5
	The Final Pillar	6
I.1.2	The Standard Model	7
	Bosons and Fermions	7
	Building the Subatomic Universe	8
	The Building Blocks of the Cosmos	9
	Language of the Subatomic Universe	10
I.1.3	The General Theory of Relativity	11
	Success in the Face of Scrutiny	12
	A Century of Correct Predictions	13
	The Exotic Realm	18
	A Beautiful Theory	20
I.1.4	The Greater Universe	20
I.2	The Dark Ocean	23
I.2.1	History	24
I.2.2	Lights in the Dark	25
I.2.3	What We Know	26
I.3	The Dark Forest	27
I.3.1	History	28
I.3.2	What We Know	29
I.3.3	Candidates and Experiments	30
	The First Black Holes	31
	Cleaning Up the Standard Model	32
	A Dark Matter Miracle?	33
	The Dark Photon	34
I.4	Heavy Photons and Beyond	37
I.4.1	Going Above and Beyond	39
I.5	Probing the Universe with the Computer	40
I.5.1	History	41
I.5.2	Complexities and Challenges	44
I.5.3	Current Programs	48
I.6	Outline	49

2	PROCA ON KERR	50
2.1	Motivation	52
2.1.1	Previous Studies	52
2.2	Theory	53
2.2.1	Curved Spacetime with Proca	53
	Decomposition in the FKKS Ansatz	55
	Systematics of Solving the Radial and Angular Equations	57
	Computing the Asymptotic Flux	59
2.2.2	Superradiance	59
2.2.3	Modified Gravitational Waves	60
2.3	Results	64
2.3.1	Proca Clouds	64
2.3.2	Identifying Modified GWs	65
	An Estimate Based on SNR	67
	LISA Proca Discovery Potential	72
2.4	Conclusion	72
3	GRBoondi	75
3.1	Motivation	76
3.2	NR Fundamentals	77
3.2.1	3+1 Decomposition of the Field Equations	78
	Slicing Spacetime	79
	Gauge Freedom	82
	Curvature of the Hypersurfaces	83
	Projecting the Riemann Tensor	84
	The Constraints	87
	NR as a Cauchy Problem	88
3.2.2	Well-posedness and Hyperbolicity	89
	Hyperbolicity of the Einstein Equations	92
3.2.3	Evolution and Schemes	93
	The BSSNOK Formulation	93
	CCZ ₄ Formulation: Damping the Constraints	96
3.2.4	Gauge Choices	98
3.3	Evolving Generalized Proca	100
3.3.1	Fundamental Pieces	101
	Structure of Code	101
	Discretization and Time-Stepping	105
	Kreiss-Oliger Dissipation	107
	Boundary Conditions	108
3.3.2	Adaptive Mesh Refinement	111
3.3.3	Parallelization: MPI, OpenMP, and SIMD	115
3.3.4	Extraction and Diagnostics	119
	Volume-integrated Quantities	120
	Surface-integrated Quantities	121
	Diagnostic Quantities	123
3.3.5	Modularity	124
3.4	Convergence Behavior of GRBoondi	127
3.4.1	Theory of Convergence Testing	128
3.4.2	Convergence Checks	130

3.4.3	Consistency Checks	131
3.5	Performance Benchmarking and Analysis	133
3.5.1	Performance Comparisons	134
	CPU and Architecture	135
	Compilation and Optimizations	135
	Assessing Relative Performance	138
3.6	Examples	140
3.6.1	Electromagnetic Fields on Minkowski	140
3.6.2	Base Proca	143
3.6.3	Base Proca in a Spinning, Expanding Universe	145
4	PROCA IN AN EXPANDING UNIVERSE	149
4.1	Theory	150
4.2	Method	154
4.3	Results	156
4.3.1	Convergence Tests	157
	Resolution Test	159
	Analytic Derivatives Test	159
4.4	Conclusion and Future Work	161
5	CONCLUSION	163
APPENDIX A ADDITIONAL DETAILS ON SUPERRADIANT PROCA CLOUDS AROUND EMRIS		165
A.1	Decomposition of Proca Equations	166
A.2	Secular Evolution of Proca Cloud	175
A.3	Comparison to Other Effects	176
APPENDIX B ADDITIONAL CONSIDERATIONS FOR PROCA IN AN EXPANDING UNIVERSE		180
B.1	Coordinate Transformations	181
B.2	Equations of Motions of the Auxiliary Damping Field	182
APPENDIX C GENERALIZED PROCA AS A CONSTRAINED SYSTEM		183
C.1	Hamiltonian Analysis à la Dirac-Bergmann	184
C.2	Well-posedness of the Cauchy Problem	186
PUBLICATIONS		189
REFERENCES		190

Listing of figures

1.1	Hadrons, Fermions, and Bosons	8
1.2	The Standard Model of Particle Physics	9
1.3	The Standard Model Lagrangian	10
1.4	Einstein Rings	13
1.5	Gravitational Lensing of Galaxy	14
1.6	M87 Black Hole	15
1.7	Kerr Metric Singularity	16
1.8	Gravitational Wave Inspiral	17
1.9	Lorentzian Wormhole	19
1.10	Alcubierre Expansion	20
1.11	The Big Bang	21
1.12	Inflation	22
1.13	Cosmological Evolution	23
1.14	Galactic Rotation Curve	28
1.15	Bullet Cluster	29
1.16	Dark Matter Candidates	31
1.17	WIMP Parameter Space	33
1.18	Dark Photon Cloud	36
1.19	Dark Photon Constrained Parameter Space	37
1.20	Proca Lagrangian	38
1.21	Expanded Einstein Equations	40
1.22	Excision Example	42
2.1	EMRI System	51
2.2	Proca Solution Workflow	59
2.3	Proca Cloud Growth Rate	65
2.4	Radial Functions of a Proca Field around a Kerr Black Hole	66
2.5	Radial Functions of a Proca Field around a Kerr Black Hole	66
2.6	Faithfulness Plot for EMRI Scenarios with Proca Clouds	69
3.1	Foliation of Spacetime	79
3.2	Foliation of Spacetime with Time Curve	81
3.3	GRBoondi Class Inheritances	103
3.4	GRBoondi Level Methods	104
3.5	Kreiss Oliger Dissipation	108
3.6	Grid Ghost Cells	109
3.7	Mesh Hierarchy	114
3.8	Subcycling	115
3.9	Distributed-Memory Parallelization	116

3.10	Shared-Memory Parallelization	117
3.11	GRBoondi Performance Metrics	139
3.12	Proca Field Energy Density with AMR Hierarchy	146
4.1	Kerr-de Sitter Discriminant	153
4.2	Kerr-de Sitter Horizons	153
4.3	Kerr-de Sitter Instability Growth	158
4.4	Convergence Tests	160
4.5	Convergence Test Routine	161
A.1	Proca Gravitational Timescale	176
A.2	Orbital Dynamics of EMRIs with Proca	177

DEDICATED TO MY PARENTS, WHOSE UNENDING SUPPORT FORMED THE FOUNDATION OF
MY SUCCESS

Acknowledgments

FIRST AND FOREMOST, I would like to express my sincerest gratitude to my supervisor, Lavinia Heisenberg. Her mentorship has been crucial to the successes I have achieved thus far. Her sincerity, kindness, and mentoring skills have made the final path to my Ph.D. extremely enjoyable. I have no doubt I would not have been able to find another supervisor as excellent as her. She fostered a successful group full of curious and kind minds. My time in her research group has been filled entirely with laughter and pure curiosity. I will always look back to my time in her group fondly.

Second, I would like to thank my parents. Their unending support has made my studies in higher education possible. Without their love and support, I would not have been able to achieve all that I have today. Their continued proudness, willingness to help me travel across the globe, and unwavering universal encouragement has formed the core foundation of my achievements. This thesis is made possible thanks to them.

Third, I have an abundance of gratitude for the Structures YRC who funded many of my travels to conferences and summer schools. Thanks to their support, I was able to experience a vast number of research fields and build friendships with other researchers around Europe.

Finally, I would like to thank my partner, whose understanding and support helped prevent me from working myself into the ground. Her love and compassion pulled me back from pure exhaustion. Without her, I would have overworked myself long ago. I remain healthy in body and mind thanks to her.

Though this thesis has only one author listed, it is through the efforts of many that made it possible, whether by funding my travels, providing thought-provoking discussions, or simply being someone who provided joy during stressful times. They all deserve to be listed because without them, this thesis would be in a very different form, or simply would not exist. It is to them that I offer my most heartfelt thanks.

*The darkest things in the universe simultaneously shine
the brightest.*

1

Introduction

THE STORY OF SCIENCE can be traced back thousands of years to ancient Mesopotamia¹. The beginning of science is fundamentally ontological. What is the nature of our being, the relations we have to the world, and the nature of these relations? The early works during this revolutionary period laid the groundwork for formalized thoughts into the underpinnings of the natural world. These early 'theories' include philosophies such as alchemy and astrology. It wasn't until the scientific revolution in 16th- to 17th-century Europe when natural scientific philosophies dramatically changed, departing from the outdated Greek traditions²⁻⁴.

The new philosophies that emerged were more mechanical, leveraging mathematical methods, concepts, and insights to explain the everyday phenomena. Moreover, the new science was far more reliable and *reproducible*, owing to the emergence of the scientific method^{5,6}. These devel-

opments paved the way for extraordinary breakthroughs and revolutions, some even transforming our very view of the universe overnight. The 18th century saw the chemical revolution, one of the first drastic changes in human scientific thought following the scientific revolution⁶. The 19th century brought new ideas and concepts regarding energy, the age of the Earth, and especially evolution⁶⁻⁸. Finally, the 20th century brought forth transformations in genetic theory and new insights in physics paved the way for new fields such as molecular biology, and particle physics⁶.

It wasn't just the spirit of natural philosophy that influenced scientific progress. The tumultuous political environment of the time often pushed scientific thinking into new paradigms and further scientific progress, the Manhattan project being one of the most well-known examples⁹. These politically-guided advancements even hastened humanity's journey into space, where progress on chemical rocketry and near-Earth space navigation allowed humans, for the first time in history, to step foot on another celestial body¹⁰.

The multi-millennia-spanning story of science has no end in sight. The modern day library of knowledge built by human curiosity of the natural world is so vast and so deep that *thousands* of textbooks have been authored and published concerning an enormous range of disciplines, topics, concepts, and theories. Moreover, this library of knowledge has been an endeavor built upon previous generational insights and continues to be an ever-growing mountain of understanding. Though historically, some have predicted a looming end of physics, today's insights have revealed a seemingly crestless mountain of understanding^{11,12}.

1.1 OUR LIBRARY OF KNOWLEDGE

Our current understanding of the universe covers a vast range of insights into the physical world ranging from subatomic lengths on the scale of femtometers, to the entire universe on the scale of billions of light-years, and everything in between. This body of knowledge has been built both by geniuses of the ages, but also lesser known scientists whose cumulative efforts have greatly advanced human understanding.

Science is very much a collaborative effort, which is especially true in the current age. As little as a century ago, the mountain of knowledge was little more than a hill, taking a curious mind comparatively little effort to summit the top and discover a new phenomena. However, at the turn of

the 20th century, the discovery of quantum theory and relativity sparked a tremendous explosion that radically shifted scientific exploration. Waves behaving like particles, particles appearing like waves, the constancy of the speed of light, the cosmic microwave background — all these discoveries opened new avenues of research, bringing forth even more discoveries that pushed human exploration into both the subatomic and supra-galactic scales.

In the current age, fundamental scientific understanding of the natural world is built upon two large pillars: quantum field theory and general relativity. Moreover, the inception of these two pillars were around the same time, the early 20th century. General relativity was the first radical shift in understanding of gravity for almost three centuries since the age of Isaac Newton^{13,14}. The theory describes the origin of gravity and its deep relationship with matter and energy. Naturally, general relativity describes things at very large scales, from the orbit of the moon to the evolution of the cosmos. However, around the same time, quantum theory was emerging as the sole descriptor of the microscopic universe and arguably represents an even bigger shift in humanity's views of the natural world.

1.1.1 QUANTUM FIELD THEORY

The progenitor to quantum field theory, the theory that describes the microscopic universe, is quantum mechanics. The annals of quantum mechanics is thus a fundamental piece of the history of modern physics. The first emergence of the quantum theory can be dated back to the latter half of the 19th century¹⁵, beginning with attempts to describe individual phenomena such as black-body radiation and solar emission spectra.

BLACK-BODY RADIATION

Black-body radiation is perhaps the greatest beacon of the inadequacy of classical mechanics to describe mesoscopic and microscopic phenomena. Thermal radiation is electromagnetic radiation that has been emitted from an object due to the object's own internal processes. A black-body is a so-called 'perfect emitter' in the sense that all incident electromagnetic radiation is absorbed, leaving none to be reflected. The opposite, called a white body, reflects all incident radiation and emits nothing. A black-body then emits only thermal radiation, which would then be called black-body

radiation¹⁶. Many everyday objects can be modeled as a black-body, such as the planets and the sun.

The prediction for the spectral signature of black-body radiation is what sparked the quantum revolution. Classical physics predicts that the spectral radiance B of electromagnetic radiation from a black-body should scale linearly with the temperature T and quadratically with the frequency f of the emitted radiation, $B \propto T * f^2$. This is a catastrophic result, suggesting that higher frequencies are emitted much more than lower frequencies. This implies that an infinite amount of photons with an infinite frequency should be emitted from everything in the universe that has a non-zero temperature. According to classical physics, the universe should be bathed in infinite energy. This goes by the name *the ultraviolet catastrophe* and the solution to this problem sparked the quantum revolution.

THE QUANTUM REVOLUTION

In the 20th century, physicist Max Planck suggested that the emitted radiation should be 'quantized' into discrete amounts. In other words, the emitted photons can only take discrete energies with the smallest unit called a 'quanta' of energy¹⁶. Yet, Planck only suggested the quanta as a mathematical tool with zero real world correspondence. Finally, in the same year that he proposed special relativity, Einstein suggested that these quanta are real particles, which now go by the name of 'photons'. This sparked an avalanche of physical predictions about the nature of atomic theory, leading Einstein to predict the photoelectric effect, earning him the 1921 Nobel prize^{17,18}.

Quantum theory continued to develop along the route suggested by the quanta. In 1927, Werner Heisenberg penned an early version of his illustrious uncertainty principle¹⁹. In 1926, Erwin Schrödinger formulated his celebrated equation that describes the behavior of quantum waves, becoming a core piece of quantum mechanics²⁰. With his equation, Schrödinger calculated the energy levels of the hydrogen atom which correctly reproduced many different properties of hydrogen, earning Schrödinger and Dirac the 1933 Nobel prize in physics^{21,22}.

FROM PARTICLES TO FIELDS AND BACK

However, around 1927, quantum mechanics stood as an offshoot of classical mechanics, regarding only particles, probability waves, and time as an absolute concept. Relativity, already well-matured

at this point, had yet to be incorporated into the theory. Paul Dirac began to consider the learnings from special relativity in the context of quantum theory, proposing his well-known 'Dirac equation' for the electron²³. This unification of relativity with quantum theory gave birth to quantum field theory, which regards particles as localized excitations of an omnipresent field.

The father of quantum field theory can be attributed to Paul Dirac himself, who was able to compute the spontaneous emission of an atom²⁴. He described the quantization of the electromagnetic field as an ensemble of harmonic oscillators, introducing new mathematical tools along the way. With these new tools, physicists believed any possible computation for any physical process involving photons and charged particles could be performed. However, it was quickly realized that the computations were only reliable to first order in the perturbative expansions²⁵⁻²⁷. It became apparent that at higher orders, infinities in the calculations began to emerge, making the computation physically irrelevant and diminishing the confidence of the theory itself. The very nature of the compatibility between quantum theory and relativity were called into question.

REIGNING IN THE INFINITIES

The resolution to the infinities began to emerge in 1947, when Hans Bethe reabsorbed the infinities into corrections of the mass and charge of the particles, giving birth to the powerful theoretical method of *renormalization*²⁸. This tool allowed finite results to be deduced. Based on these results, several other physicists were finally able to covariantly formulate arbitrarily precise models of quantum electrodynamics (QED), earning them the 1965 Nobel prize in physics²⁹⁻³⁷. Through these works, renormalization became a foundational aspect of quantum field theory. With these techniques, QED became one of the most precisely tested theories in all of physics³⁸. For some of the parameters, the agreement with experimental data is found to be within 1 part in 10^{-9} , or one part in a billion, making it one of the most precisely tested physical theories of all time³⁹⁻⁴¹.

DIVING INTO THE NUCLEUS

The success of QED has made it a model for other quantum field theories, such as quantum chromodynamics (QCD), which describes the strong force between quarks and gluons, the nucleus equivalents of electrons and photons, respectively. These comprise the inner world of the proton

and neutron, making QCD the theory describing the nucleus of the atom. Though equivalent to QED in spirit, QCD is saliently different in practice resulting in very different phenomena.

QCD exhibits several bewildering phenomena relative to QED. The first is color confinement, the phenomenon where quarks and gluons cannot be isolated from one another. In fact, as two color-charged particles such as quarks are continuously separated, the force between them approaches a constant. This is in stark contrast to the inverse-square law of electrodynamics, where the force falls off as the square of the distance between two electrically charged particles^{42,43}.

The second phenomenon is asymptotic freedom, in which the strength of the interaction between quarks and gluons steadily decrease as the energy scale increases and the corresponding length scale decreases^{44,45}. In other words, quarks interact weakly at high energies. At low energies, the interactions are much stronger, leading to confinement of quarks and gluons, and forming composite hadrons such as protons and neutrons.

The third phenomenon is chiral symmetry breaking⁴⁶. This phenomenon is different in nature than the previous two in that the internal symmetries of the theory undergo spontaneous breaking. At the theoretical level, this spontaneous breaking of the symmetry leads to the dynamical generation of mass for otherwise massless particles.

These three strange phenomenon make QCD a very different theory than QED and, at the same time, shows the versatility of the tools allotted by quantum field theories.

THE FINAL PILLAR

The third and final foundational pillar of our understanding of the subatomic world illuminated by the mechinations of quantum field theory is quantum flavordynamics (QFD). The earliest history of QFD can be traced back to Enrico Fermi, the pioneer of the weak interaction^{47,48}. Fermi suggested that the beta decay process can be explained via the weak force which is mediated by the W and Z bosons, analogous to the photon of electromagnetism. In fact, the weak interaction is responsible for all radioactive decay and thus is important in the understanding of many physical processes that occur in stellar matter, fission reactors, and any other radiation-driven processes.

In the 1960s, the pioneering work of Sheldon Glashow, Abdus Salam, and Steven Weinberg unified the electromagnetic and the weak force into a single force termed the *electroweak force*^{49,50}.

They earned the 1979 Nobel prize in physics for their discovery. They realized that above a critical energy scale of around 246 GeV, the electromagnetic and weak interactions would unify into a single force. It was later shown that the electroweak theory of Glashow, Salam, and Weinberg was renormalizable. In other words, the infinities inherent to the quantum field theory could be reabsorbed, leading to well-defined finite predictions. To date, the precision of experimental measurements of the electroweak interaction are on par with that of QED, making it one of the most precisely tested theories⁵¹.

1.1.2 THE STANDARD MODEL

These three revolutionary theories, QED, QCD, and electroweak theory, built using the mathematical tools of modern QFT, form our understanding of the atomic cosmos. These theories classify and describe all known fundamental particles and three out of the four fundamental forces. The collective name of these three theories that form our modern understanding of atomic processes is the *Standard Model*.

The Standard Model of particle physics is the collective model that describes the electromagnetic interaction, the weak interaction, the strong interaction, and all fundamental particles. This collective description can be split broadly into two separate pieces, namely the matter itself and the interactions between the matter particles. The particles representative of the interactions are called 'force carriers' and they include the photon, the W and Z boson for the weak force, and the gluon for the strong force. All fundamental particles can be divided into three large classes, fig. 1.1. They can be bosonic or fermionic, depending on the internal properties of their quantum fields, and they can be hadronic.

BOSONS AND FERMIONS

The bosonic and fermionic descriptors are categorizations that depend on the internal spin of the particle, which is a quantum analog of angular momentum and is a fundamental intrinsic property of the particle. In equivalence to the photon, the spin of particles is also quantized and hence the spin can be used to distinguish various particles. The electron is a famous example of a particle that carries intrinsic spin, which was inferred in 1921, by Otto Stern and Walther Gerlach, by sending an

electron through an inhomogeneous magnetic field. The deflection of the trajectory of the electron led Stern and Gerlach to suggest that all electrons carry a fundamental, irreducible, and quantized spin.

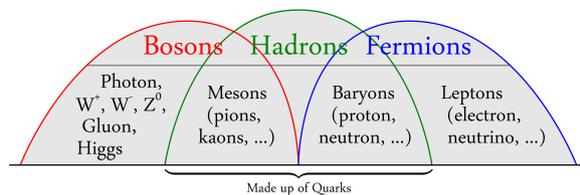


Figure 1.1: ⁵² Division of fundamental, composite, and force-carrying particles into classes. The force-carrying particles are all bosons, while the matter particles can be bosonic or fermionic, depending on the internal properties of the particles and whether they are fundamental or composite.

Bosons are particles that carry a spin which is an integer number, such as 1, 2, etc. Fermionic particles are those particles whose spin is a half-integer, such as $\frac{1}{2}$, $\frac{3}{2}$, etc. ^{53,54} All force-carrying particles, the W, Z bosons and gluons (as well as the gravitational force carrier), are bosonic. Electrons, neutrons, the neutrinos, and protons are all fermionic. However, the proton

and neutron are not fundamental since they are composed of more fundamental particles. Particles that are composed of more fundamental particles are composite particles held together by the strong nuclear force and are called hadrons. They are the subatomic analog of molecules, which are held together by the electromagnetic force.

BUILDING THE SUBATOMIC UNIVERSE

The hadrons and fermions can be further categorized into three more categories. The leptons are fundamental fermionic particles that do not interact with the strong force. Due to the lack of interaction with the strong force-carrying particles, they are termed 'color-free', where color refers to the strong nuclear interactions equivalent of electric charge. The particles that are leptonic are the electron, muon, tauon, and their corresponding neutrino flavors, the electron neutrino, the muon neutrino, and the tau neutrino. The fermionic hadrons are called baryons. They are composite particles that participate in the strong nuclear force. Particles that are baryonic are the proton, neutron, and more exotic flavors such as the pentaquark. The bosonic hadrons are called mesons. They are also composite particles which participate in the strong nuclear force. They differ from the baryons in that they possess an even number of quarks, in contrast to baryons which are composed of an odd number. However, all mesons are unstable, with the longest living one lasting only for tens of nanoseconds ⁵⁵. This implies free mesons only exist for a very short amount of time. On the other

hand, mesons are extremely important as they are responsible for binding together atomic nuclei⁵⁶. Since they are bosonic, they act as the force carrier for the *residual strong force*, or nuclear force. The strong nuclear force can be thought of as certain combinations of quarks 'leaking away' from the internal interactions of the individual protons and neutrons, binding the protons and neutrons together and forming atomic nuclei.

Since mesons are unstable, they also participate in the weak interaction where they can undergo decay processes to lighter mesons. There are hundreds of discovered mesons, all with various properties dependent on their constituent particles⁵⁵.

The force carrier bosons, mesons, baryons, and leptons contain all fundamental particles known to particle physics. There are an enormous number of such particles with most being mesons and baryons. However, all known composite particles are composed of just 12 fundamental matter particles which interact with the 5 force-carrying particles.

THE BUILDING BLOCKS OF THE COSMOS

The 6 quarks compose the inner nuclei, binding together with the gluon to form protons, neutrons, and mesons. The 6 leptons consist of the electron, the muon, the tauon, and the electron neutrino, the muon neutrino, and the tau neutrino. The five force-carrying particles are the gluon, responsible for binding together quarks to form hadrons, the photon, responsible for the interaction between leptons, the W and Z bosons, responsible for radioactive decay of particles, and the Higgs particle, which is responsible for giving mass to all fundamental particles.

These 17 particles constitute the Standard Model, the most accurate model for the fundamental constituents of the universe, fig. 1.2. The Standard Model has been in development since the latter

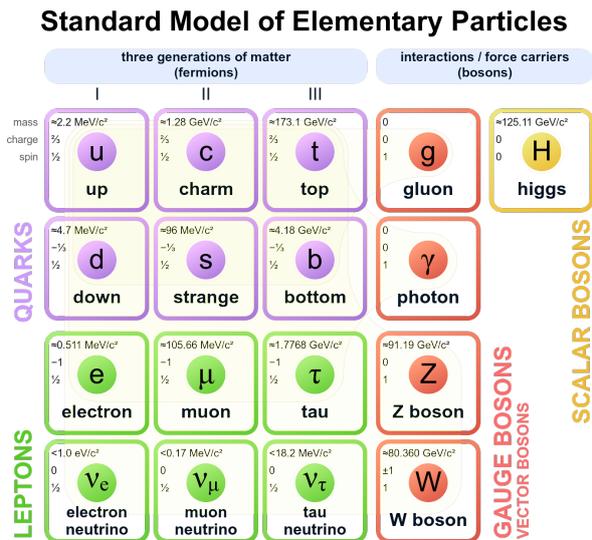


Figure 1.2: ⁵⁷The Standard Model of particle physics consists of twelve fundamental matter particles and five force-carrying particles. The six quarks are the fundamental particles that compose mesons and baryons, yielding atomic nuclei. The six leptons consist of electrons and their heavier cousins, the muon and tauon, along with their associated neutrino flavors. The five force-carrying particles, the gluon, photon, W and Z bosons, and the Higgs particle are responsible for the interactions between the various matter particles or, in the case of the Higgs particle, give mass to them.

half of the 20th century through the enormous work of scientists around the world⁵⁸. The Standard Model was finalized in the 1970s after the existence of the quark was experimentally confirmed⁵⁹⁻⁶¹.

$$\begin{aligned} \mathcal{L} = & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ & + i\bar{\Psi}\mathcal{D}\Psi + \text{h.c} \\ & + \Psi_i y_{ij} \Psi_j \phi + \text{h.c} \\ & + |\mathcal{D}_\mu\phi|^2 - V(\phi) \end{aligned}$$

Figure 1.3: The Standard Model Lagrangian, the fundamental mathematical object that describes the atomic universe. The first line describes the four force-carrying particles, how they exist and how they interact with each other. The second line describes how the force-carrying particles interact with the matter particles. The third line describes how the matter particles interact with the Higgs field ϕ , thereby obtaining mass. The last line describes the Higgs field itself and how the force carriers – in fact, only the weak force carriers – interact with it.

The Standard Model is built using the tools of quantum field theory and is an immensely complicated theory. The Standard Model consists of 19 parameters whose numerical values must be fixed by experimental data. However, in 2010, it was discovered that the neutrino, previously thought to be a massless particle, actually contains a very light mass⁶². This discovery changed the Standard Model and in doing so introduced 7 new parameters, for a total of 26 experimentally determined parameters.

LANGUAGE OF THE SUBATOMIC UNIVERSE

The fundamental mathematical object that defines the theory is the Standard Model Lagrangian, fig. 1.3⁶³. This object contains the 12 fundamental matter particles, the five force-carrying particles, describes how they interact with each other, and how they acquire mass via the Higgs particle. This object encapsulates our best understanding of all particles in the universe, how they interact with each other to form nuclei, atoms, molecules, and eventually stellar matter, planets, and organic material. It describes the propagation of light, how cells form animals and humans, how the atmosphere behaves, how fusion reactors work, how the sun warms the planet, and how the light from stars on the other side of the universe is created. Naturally, using the Standard Model to describe the weather would be analogous to using a hydrogen bomb to pound a nail instead of a hammer. Remarkably good approximations can be built from the model, such as chemistry, condensed matter physics, etc. Nonetheless, it's the fundamental theory that is thought to describe how all fundamental particles in the universe exist and behave.

That isn't to say the Standard Model is perfect. There are many challenges that still need to be

overcome. For instance, the description of the gravitational force is not present anywhere in the model, nor is dark matter or dark energy. But much more fundamentally, the Standard Model is mathematically inconsistent with that of the theory of gravity. This implies one or the other must be wrong. There have been many theories proposed that are collectively denoted as *physics beyond the Standard Model*. This refers to theories that aim to solve the challenges faced by the Standard Model and includes models such as Minimal Supersymmetric Standard Model, string theory, M-theory, loop quantum gravity, and many more. All these theories aim in some way to bring in a quantum description of gravity which is currently described by general relativity. However, due to the nature of the gravitational theory, this has proven to be a remarkably difficult problem and it is currently unclear which direction is correct.

1.1.3 THE GENERAL THEORY OF RELATIVITY

The current best theory of gravity that continues to withstand the scrutiny of experimental data is General Relativity (GR). GR is a vastly different theory from the Standard Model, but its origins can be traced back to around the time of the quantum revolution. Albert Einstein first developed his theory of special relativity in 1905, an offshoot of classical mechanics that hypothesizes the constancy of the speed of light based on Maxwell's equations⁶⁴⁻⁶⁶. He realized later that his 'principle of relativity', the idea that the speed of light is the same for everything no matter their velocity, could be extended to the gravitational field. He later realized that undergoing free fall in a gravitational field is equivalent to non-accelerating motion. Hence, the rules of special relativity must apply to a freely falling observer. This idea is called the *equivalence principle* and is at the very heart of GR⁶⁷. At the same time, Einstein predicted the phenomenon of gravitational time dilation. In 1911, Einstein then predicted the equivalence between accelerated motion and the gravitational force, deducing gravitational light deflection⁶⁸. He later sought to describe the gravitational force as a geometrical manifestation, employing the use of differential geometry. Following several years of thought and development, Einstein finally wrote down field equations he believed to accurately describe the gravitational forces, which took the form

$$R_{ab} = \mathcal{T}_{ab} .$$

The left-hand side describes the geometry of spacetime, the unification of space and time into a single object, and the right-hand side describes the distribution of matter and energy. However, Hermann Weyl realized this equation is only consistent if the entire universe was filled with a uniform distribution of mass⁶⁹. Einstein returned in 1915, with an improved version, which he presented to the Prussian Academy of Sciences⁷⁰, and which took the form

$$R_{ab} - \frac{1}{2}g_{ab}R = \mathfrak{T}_{ab} ,$$

These are nowadays famously known as the Einstein equations.

SUCCESS IN THE FACE OF SCRUTINY

The history of the astronomical verification of Einstein's predictions is turbulent, however they were finally verified in 1919, by Eddington and Dyson and again in 1922, by Campbell⁷¹⁻⁷³. To this day, GR has undergone a litany of tests, making it the most successful gravitational theory of all time. The first tests were proposed by Einstein himself, suggesting that the perihelion precession of Mercury's orbit, the gravitational deflection of light passing by the sun, and the gravitational redshift of light could be used to test the predictions of his theory. The perihelion precession of Mercury's orbit is one of the more powerful tests of GR, which it famously passed.

More modern tests of GR have been carried out in the hopes of finding a breakdown in either GR or the Standard Model. One of the most important modern tests of GR is that of gravitational lensing. The most precise tests are analogous to Eddington's 1919 experiment in which deflections of light from distant sources by the sun are measured. Current day tests confirming the predictions of GR are at the 0.03% level⁷⁴. The Shapiro time delay, a relativistic correction of the time taken by a photon to complete a there-and-back journey to a planet, is another 'classical' test of GR, which agrees with experimental data at the 0.002% level (though there is debate on the experimental data)⁷⁵. Tests of the equivalence principle itself have also been carried out using experimental apparatus called "Eötvös torsion balance" experiments. These test the assertion that the trajectories of falling bodies are independent of their mass and internal structures. They have tested the equivalence principle to a factor of 10^{-15} , making it an extremely precise test⁷⁶.

The direct detection of gravitational waves in 2015 by the Advanced LIGO team was a momen-

tous test of GR. The discovery achieved two things: provide further data in support of GR and heralded in the new age of multimessenger astronomy. GR predicts that orbiting astrophysical objects will emit gravitational waves and the radius of this orbit will slowly decay due to the emissions. The existence of gravitational waves is not unique to GR, but to any theory that predicts the propagation of gravity at some speed. The measurement of the waveform at Earth of these waves is then a powerful test of GR among the swampland of all gravitational theories. So far, these tests have been in excellent agreement with GR⁷⁷⁻⁷⁹. Multimessenger astronomy, the combination of astronomical measurements such as radio, x-ray, and optical observations with gravitational wave detections and measurements, will be a powerful tool in continuing to test the predictions of GR.

A CENTURY OF CORRECT PREDICTIONS

GR makes many new predictions, some that are regarded as 'physical' which could occur in the universe, and even more that may just be exotic mathematical oddities. The first radical prediction of GR that significantly departs from the outdated Newtonian gravity is the existence of gravitational lensing. Since GR associates the gravitational force with the curvature of spacetime, the trajectory of a photon can appear to curve when travelling past a massive object. The most striking example of this is the existence of so-called Einstein rings, fig. 1.4. A massive object in front of a distant luminous object can curve the light from the distant object so drastically, it appears to smear out into a ring with the massive object at its center. This forms what appears to be a ring of light, which are called Einstein rings. Einstein rings are one of the more drastic manifestations of gravitational lensing and are part of a class of gravitational lensing called *strong lensing*. They are distinguished by the clearly visible distortion of a background source that forms Einstein rings, arcs, and multiple images, fig. 1.5.

Weak lensing is a class of gravitational lensing where the light of the luminous source is only

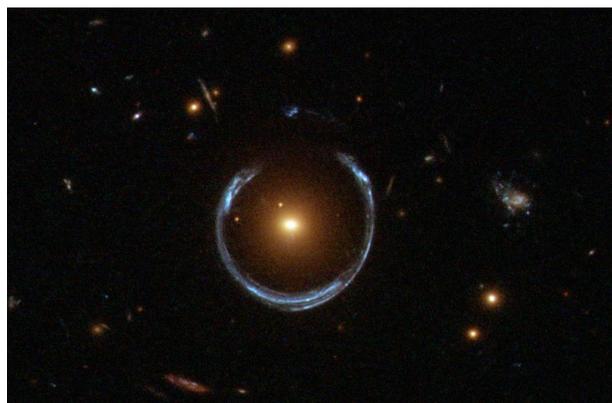


Figure 1.4: Einstein rings are a form of strong gravitational lensing where the light of a distant luminous source is bent into a ring shape by the gravitational effect of a closer massive object. Credit: ESA/Hubble, NASA

slightly modulated and can only be detected using large statistical samples. Usually, a large number of galaxies must be used to reach high enough statistical confidence for the presence of weak lensing due to the fact that galaxies are intrinsically elliptic.

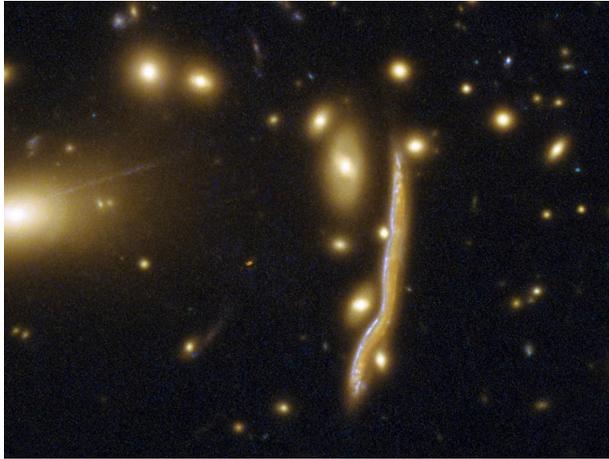


Figure 1.5: Strong gravitational lensing, or strong lensing, is a class of gravitational lensing where the light of a distant luminous source is clearly distorted by the eye, forming Einstein rings, arcs, or multiple images. In this image, the enormous mass of galactic cluster MACSJ1206.2-0847 distorts the image of a more distant galaxy. Credit: ESA/Hubble, NASA

The final class of gravitational lensing effects is that of *microlensing*, where no distortions can be detected yet the amount of light received from a luminous object changes with time. Gravitational lensing can be a powerful tool in astronomical measurements since one can use the degree of lensing to estimate the mass of the lens, the most famous example being that of the Bullet cluster.

The existence of *black holes* is another prediction of GR that has been tested extensively. By definition, a black hole is a region of space-time where the curvature is so strong that not

even electromagnetic waves — the fastest thing in the universe — could escape it, leaving what appears to be a black lightless void. However, the idea of such an object dates back to the 18th century and was suggested by the astronomer John Michell^{80,81}. He suggested that a star massive enough could have an escape velocity that exceeds the speed of light, and that their existence could be inferred by their effect on neighboring bodies.

The modern theory of GR predicts a similar object, the black hole, though the origin of the gravitational force is much different than suggested during the time of Michell. Instead, the curvature of spacetime becomes so strong that light can no longer escape the gravitational 'well' and, if close enough, inevitably plummets to the center. The first solution of the Einstein equations which possessed such properties was suggested by Karl Schwarzschild just a few months after Einstein introduced his field equations^{82,83}.

Schwarzschild developed a solution which results from a spherically symmetric and static point mass. Though simple in its construction, it already possesses many striking properties that portray

just how bewildering black holes are. The first is the existence of a curvature singularity. This is the point at the center of the black hole where the curvature, i.e. the gravitational force, blows up to infinity. All matter that enters the black hole will inevitably reach the singularity, where the density of the matter will diverge.

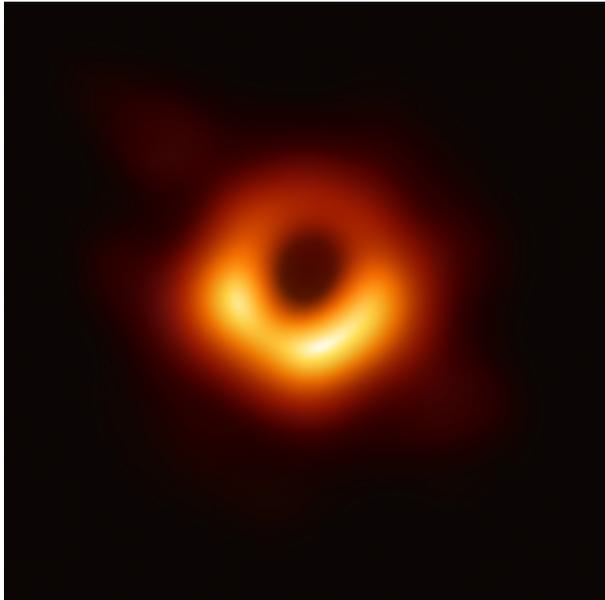


Figure 1.6: An image of the supermassive black hole at the center of the M87 galaxy in the radio band. Credit: EHT Collaboration

Many physicists today believe such intense gravitational forces signal a breakdown in the theory of GR and will require a fully quantum version of gravity to describe. Nonetheless, in the 1960s Roger Penrose and Stephen Hawking proved that singularities are an inevitable consequence of GR, earning Penrose the 2020 Nobel prize in physics, though Hawking unfortunately passed away in 2018⁸⁴⁻⁸⁶. Another striking feature is the existence of an event horizon. Though not a solid surface like a planet, the event horizon is a region defined as a 'coordinate singularity'. Its defining feature is that

it's the radius from the center past which things travelling at the speed of light can no longer escape. This turns the event horizon into a 'causal boundary', as events that occur inside the event horizon can not affect anything on the outside, since the information would have to travel faster than the speed of light to escape. The event horizon is analogous to the surface of Michells 'dark star', where the escape velocity equals the speed of light.

Since the work of Schwarzschild, several new black hole solutions of the Einstein equations have been put forth. The most widely utilized and the one believed to most accurately represent black holes in our universe was formulated by Roy Kerr in 1963⁸⁷. Now called the *Kerr metric*, Roy Kerr wrote down the solution of the Einstein equations that describe an empty spacetime which possesses a spinning black hole, a generalization of the Schwarzschild solution which was static and unmoving. The Kerr solution exhibits many new interesting properties such as the Lense-Thirring precession effect, a Coriolis-like force distinct to solutions of GR which are rotating⁸⁸⁻⁹⁰. The sin-

gularity at the center of the black hole is also no longer point-like, but instead smeared out over a ring, fig. 1.7⁹¹.

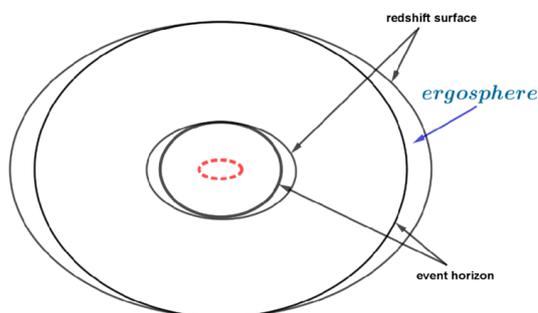


Figure 1.7: ⁹²The various structures of the Kerr black hole. The ergosphere lies outside the outer horizon. The singularity is now a ring instead of a point, represented by the dashed line.

A new feature of the Kerr black hole that differs from the Schwarzschild black hole is the existence of what is called an *ergoregion*, delineated between the event horizon and the *ergo-surface*. The Kerr metric is distinct from the Schwarzschild metric in that it possesses spin. However, black holes are not material objects, they are regions of spacetime itself. Hence, it is spacetime itself that is rotating, dragging along

material particles and other forms of energy with it. The ergosurface is a special surface of the spacetime beyond which material particles can no longer remain at rest and are forced to rotate with the black hole. Matter and energy would have to travel faster than the speed of light to remain rotationally stationary. However, since the ergosphere is outside the event horizon, particles are still free to escape the black hole. Another feature distinct from the Schwarzschild black hole is the existence of *two* event horizons. The outer horizon is analogous to the Schwarzschild horizon and is the surface beyond which nothing can escape. The inner horizon is closer to the singularity and is called a *Cauchy horizon*. This horizon is technically escapable and past this horizon, closed curves in spacetime can exist, providing a possible scenario of time travel, though their existence is questionable at best.

Beyond the simple cases of static and rotating black holes, there is a plethora of other black hole-like solutions, such as the Reissner-Nordström black hole, describing a static black hole that possesses an electromagnetic charge, the Kerr-Newman black hole, describing a rotating and electromagnetically charged black hole, the white hole, the exact opposite of the usual black hole in which nothing can enter its event horizon, and hairy black holes, which possess other charges beside the usual electromagnetic one.

The first astronomical object to be accepted as a black hole is Cygnus X-1, a galactic x-ray source^{96,97}. The very first image of a black hole was captured in 2019 by the Event Horizon Telescope (EHT) collaboration, where radio telescopes scattered around the Earth were used to form a virtual Earth-sized radio telescope to image the center of the M87 galaxy, capturing an image of the surrounding orbital material of the supermassive black hole at its center, fig. 1.6⁹⁸⁻¹⁰⁰.

The EHT collaboration again used the same techniques in 2022 to image the black hole at the center of our Milky Way galaxy, Sagittarius A*¹⁰¹⁻¹⁰³. These images can be used to test the predictions of GR, since the orbiting material around the black hole is in an extremely strong gravitational environment.

Beyond black holes, GR also predicts the stability and evolutions of stars such as main-sequence stars like our sun, white dwarfs and neutron stars, which are collapsed stars composed of degenerate quantum matter, and more exotic stars like strange stars or quark stars which have not been observed. Nonetheless, GR can be used to predict the shape of these stars which in turn, using observational data, place constraints on GR itself.

Another famous prediction of GR that has only recently been discovered is that of gravitational waves. These are waves in spacetime itself that are generated by accelerated masses and propagate at the speed of light. In 1916, Einstein demonstrated that his theory of general relativity predicts the existence of gravitational waves^{104,105}. Nearly a hundred years later, in a groundbreaking experiment, the LIGO-Virgo collaboration made the first direct detection of gravitational waves generated by a pair of stellar mass spinning black holes a distance of about 1.4 billion light years away. The collision of these black holes generated gravitational waves with a power at merger of about 3.6×10^{49} watts which, if this were an electromagnetic signal, would outshine every star in the

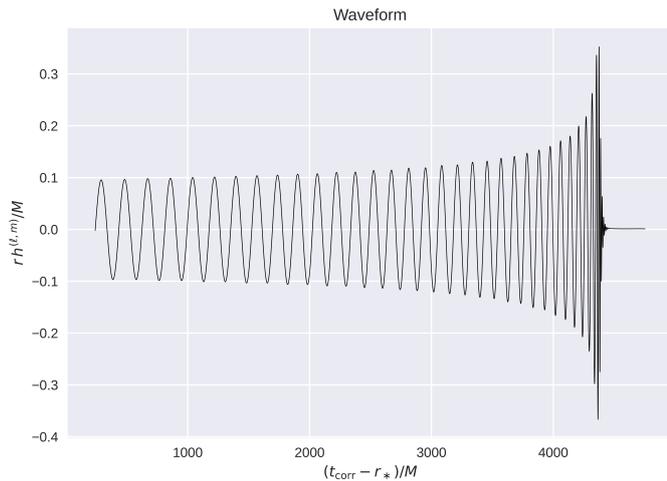


Figure 1.8: ⁹³⁻⁹⁵ An example of a gravitational signal. The last few moments are highly dynamical and the most difficult part of the binary merger to model. This waveform clearly shows the three distinct phases of a merger — the long inspiral phase, the transient merger phase characterized by the large peaks in the waveform, and finally a ringdown phase when the final combined black hole settles down. This waveform was generated using numerical solutions of the Einstein equations with binary black hole initial data.

universe combined^{77,78}. These phenomena are easily one of the most powerful events in the entire universe, showcasing the colossal prowess of black holes.

The direct discovery of gravitational waves confirmed the last remaining directly undetected predictions of GR and provides even more evidence of the geometric nature of gravity. The discovery itself brought forth the new age of gravitational wave astronomy, and multimessenger astronomy in general. Since the famous 2015 discovery, over a hundred more detections have been published, including the mergers of neutron stars and neutron star-black hole binaries¹⁰⁶⁻¹⁰⁸. Multimessenger astronomy has been a powerful tool to study black holes and other compact objects, and provides stringent tests of GR itself, even placing constraints on the quantum nature of black holes¹⁰⁹⁻¹¹⁷.

THE EXOTIC REALM

Beyond the astrophysical predictions that astronomers have observed by testing GR in a huge range of experiments, GR is also a mathematical theory, so the mathematical aspects of GR have also been explored in depth. From these studies, several very interesting results have been published. Some of the more interesting ones include closed timelike curves, wormholes, and warp drives.

A timelike curve is the path of a physical observer, such as a person, bird, plane, planet, or galaxy, through the four-dimensional spacetime. It's the curve that describes how physical matter travels in space and time. A *closed timelike curve* (CTC) is such a physical trajectory that forms a closed loop in spacetime, returning to its starting position. In normal three-dimensional space, a closed trajectory is usually not an issue, since observers are free to travel through space without restrictions, such as planets completing their orbits. However, a closed curve in four-dimensional spacetime possesses the primary issue of possibly allowing time travel, raising paradoxical philosophical issues such as the grandfather paradox. Such CTCs have been shown to exist both in exotic solutions of GR, such as infinite material cylinders, and physical solutions, such as the Kerr black hole. However, Hawking proposed that CTCs are unphysical and proposed the *chronology protection conjecture*, "It seems that there is a Chronology Protection Agency which prevents the appearance of closed timelike curves and so makes the universe safe for historians."¹¹⁸. Many theorists believe the quantum version of the gravitational theory will rule out the existence of CTCs. Nonetheless, they exist as perfectly reasonable trajectories allowed by GR.

Wormholes are another exotic object allowed within the context of GR. They are a purely hypothetical structure that connects two distant, even casually disconnected, regions of spacetime. Some forms of wormholes are even of the interuniversal type. The first type of wormhole was the Schwarzschild wormhole, a wormhole thought to exist at the singularity point of the Schwarzschild black hole.

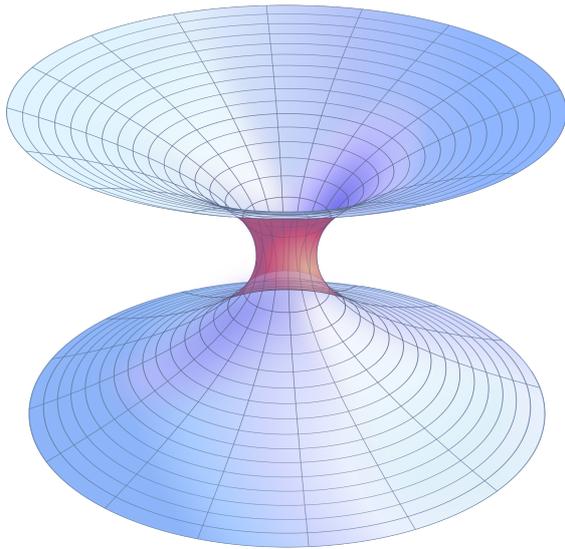


Figure 1.9: ¹¹⁹ An embedding diagram of a Lorentzian wormhole, a type of continuation of the Schwarzschild black hole connecting two separate universes.

However, these types of wormholes have been shown to be unstable, requiring exotic negative matter to stabilize them ¹²⁰. Other types of wormholes are the so-called *traversable wormholes*, wormholes that material particles could travel back and forth through, being held up also by exotic negative mass material. If traversable wormholes exist, they would be a method of faster-than-light travel, since they could connect two separate ends of the universe through a shortcut.

The last exotic solution of GR worth mentioning is the *warp drive*. Warp drives have

a turbulent history since the 1990s when Miguel Alcubierre proposed his now-called Alcubierre warp drive ¹²¹. The main idea of the Alcubierre drive is that the spacetime behind an object is expanded while the spacetime in front is contracted, supposedly accelerating the object to super-light-speed velocities, fig. 1.10. However, already in Alcubierre's proposal, the Alcubierre drive has been shown to violate almost all physicality conditions, requiring enormous amounts of negative energy. Since his proposal, many other proposals have been suggested as reducing or outright eliminating the negative energy requirements ¹²²⁻¹²⁶. The warp drive is still a controversial topic among theorists with some suggesting the warp drive is completely infeasible. Nonetheless, this remains an active area of research in the hopes of aiding humanity's expansion into space and crossing interstellar distances ^{122,127,128}.

Warp drives are a particularly appealing solution to interstellar travel since they are a perfectly valid solution of the Einstein equations, though their physical validity is still an active question in research. It may be possible one day to construct a spacetime distortion using physically realistic material to propel a payload to cosmic distances on human timescales.

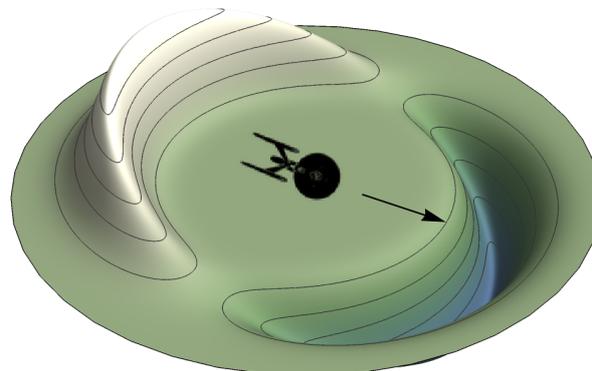


Figure 1.10: A slice of the spacetime expansion in the Alcubierre warp drive. Spacetime behind the object, in this case a spaceship, is expanding and spacetime in front is contracting, in theory propelling the ship to faster-than-light speeds.

A BEAUTIFUL THEORY

GR is easily the most successful gravitational theory of all time, continuing to pass stringent observational tests ranging from terrestrial experiments, such as the Eötvös torsion balance experiment, to strong-gravity astrophysical tests, such as the merger of massive black hole binaries. Repeated successes increase the confidence in the correctness of GR in explaining astrophysical phenomena, allowing theorists to use GR to confidently predict new phenomena and explain currently observed ones. Another of GR's key successes is explaining the evolution of the universe's structure as a whole when cosmologists apply GR to the entire cosmos.

1.1.4 THE GREATER UNIVERSE

The application of GR to the entire universe results in the field of study called *theoretical cosmology*. A cosmological model provides a description of the large-scale structure of the universe and allows questions of its origin, evolution, composition, and eventual fate. The current understanding of cosmology began with the introduction of GR in 1915. This was quickly followed by the first major observational discovery of cosmic expansion in the 1920s by Edwin Hubble, Vesto Slipher, and others^{129,130}. Further efforts revealed the cosmic microwave background, distant supernovae, and the accelerated expansion of the universe, leading to the theory of the big bang and the standard cosmological model — our best understood theory of the composition and history of the universe.

The standard cosmological model, called the Λ -CDM model, is the current best understanding of the contents and evolution of the cosmos. It purports that the universe began from an initial quantum fluctuation during an era called the *Planck epoch* and lasted for 10^{-43} seconds and is believed to be the time when all four fundamental forces were united into a single force from which came the *grand unification epoch*, where gravity is believed to have separated from the other three fundamental forces.

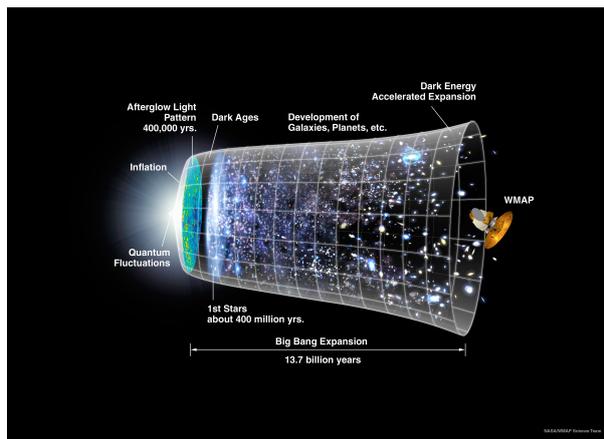


Figure 1.11: An artistic rendition of the evolution of the universe, from the initial quantum fluctuations, through the rapid inflationary epoch, to the present day accelerated expansion phase. Credit: NASA/WMAP Science Team ¹³¹

So far, these epochs are highly theoretical and not much, if anything, is understood about these timescales. It's commonly understood that a full quantum theory of gravity is required to really understand what took place during this time.

Next came the *inflationary epoch* which lasted until the universe was 10^{-32} seconds old. This is where the predictive power of Λ -CDM begins to take place. During this epoch, the universe underwent a drastic and rapid in-

flation from 10^{-36} seconds to 10^{-32} seconds, increasing in size by a factor of 10^{26} , fig. 1.12. This radical idea solves many of the previous issues present in cosmology at the time. During inflation, the initial quantum fluctuations 'froze in', giving birth to the large-scale structure inhomogeneities we see today, such as galaxies and stellar clusters. Following inflation came several dynamical eras characterized by additional particle interactions, the *electroweak epoch*, *quark epoch*, *hadron epoch*, *lepton epoch*, then the *photon epoch*. During the photon epoch, at around 370,000 years after the big bang, the cosmic microwave background (CMB) formed, which is perhaps the greatest evidence we have for the big bang. The CMB formed as a result of electrons recombining with protons to form atoms, allowing photons to stream freely through the universe, some travelling unperturbed until they finally reach our detectors. This period of the universe forms an opaque barrier for cosmological telescopes, since before the CMB, photons could not stream freely as they continuously interacted with free electrons.

Overlapping with the photon epoch and reigning for billions of years was the *matter dominated era*, where the universe cooled enough for matter to clump together to form structures. At this point, the non-relativistic matter dominated the energy content of the universe, hence its name. According to Λ -CDM, the universe was mainly composed of 15.5 % ordinary matter and 84.5 % yet-to-be discovered 'dark matter'.

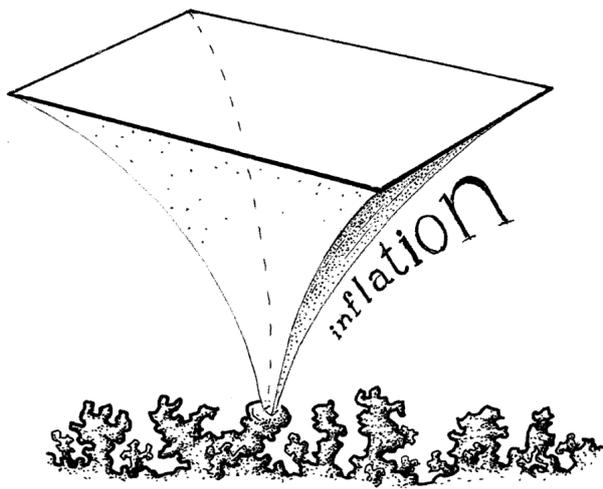


Figure 1.12: ¹³²Artistic rendition of inflation from early quantum fluctuations.

The matter dominated era lasted from about 47,000 years after the big bang until about 9.8 billion years after the big bang. During the era, many physical processes took place, including the formation of the CMB, the beginnings of molecular matter and the formation of stars, galaxies, and clusters.

After the matter dominated era came the *dark energy dominated era*, the current era the universe exists in. The dark energy dominated era is characterized by the dominant content

of the universe being that of a yet-to-be-discovered form of energy that is responsible for the accelerated expansion of the universe. According to Λ -CDM, this is the final epoch of the universe and several scenarios have been suggested to be the ultimate fate of the cosmos:

- *Heat Death*: Cosmic expansion continues to accelerate until thermal equilibrium is reached after 10^{1000} years and no more structure is possible. After this point, the universe has reached a highly entropic state and most matter is in the form of subatomic particles and low-energy photons. ¹³³
- *Big Rip*: Expansion of spacetime becomes so extreme that subatomic particles are ripped apart and eventually spacetime itself, resulting in a kind of singularity event ¹³⁴. Current observations by the Chandra X-ray Observatory and the Planck telescope do not rule out the big rip scenario ^{135,136}.
- *Big Crunch*: Expansion eventually halts and reverses under the influence of all matter and

energy in the universe, accelerating it back towards a hot and dense state. Current estimates suggest this scenario is unlikely.

- *Vacuum Decay*: This suggests the universe is in a false vacuum state, leading to the implication that the universe could later undergo a phase transition to a lower state, replacing all particles and forces with new ones, destroying the old ones in the process.

Currently, Λ -CDM favors the first two scenarios, given the recent cosmological data. However, vacuum decay or some other extremely unlikely quantum phenomena are technically still possible.

At the heart of the standard cosmological model lies GR, the underlying theory that cosmological parameters are fed in to. To date, GR has been a well developed theory that has passed many cosmological tests. However, there remains many open challenges that need to be addressed, such as the nature of dark matter and the energy content that is responsible

for the accelerated expansion of the universe. While it is possible that all these challenges have answers within GR, it could be that GR itself is ultimately the wrong theory. In fact, a large number of theorists believe this to be the case since GR does not tell us anything about the nature of fundamental particles. Should there be a single true theory underpinning the universe, GR is certainly not it. Nonetheless, it provides a remarkably good approximation when describing large-scale physics ranging from the orbits of planets to the evolution of the cosmos.

1.2 THE DARK OCEAN

The true nature of the universe's accelerated expansion is a particularly challenging problem to solve. In reality, it is likely a combination of quantum field theory and GR. After all, material par-

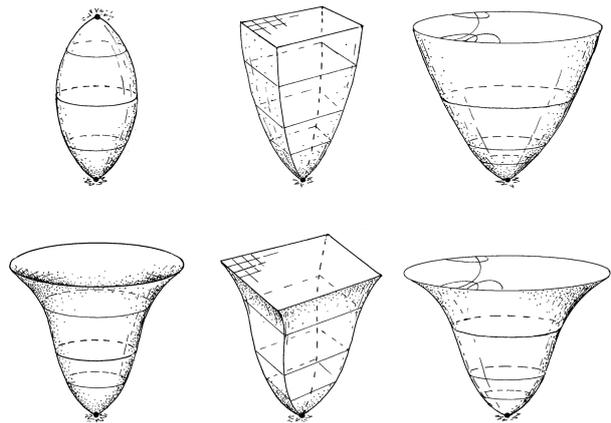


Figure 1.13: ¹³² Different types of evolution scenarios for the cosmos lifetime. The top left is an example of the Big Crunch. The universe is birthed from a singularity and then eventually collapses back into one. The bottom three show different scenarios of accelerated expansion. The top middle shows expansion at a decelerated rate. Finally, the top right shows expansions at a constant rate.

ticles and the force carriers live on spacetime and therefore are affected naturally by the spacetime background. Hence, the description of the underlying energy responsible for the accelerated expansion of the universe should have a natural residency in a fully quantum theory of gravity. As this has yet to be fully developed in a consistent way, the only recourse is to embed a quantum theory in a classical description of gravity, such as that of GR.

1.2.1 HISTORY

The idea of a cosmological constant is traced back to the very early works by Einstein on his new theory of general relativity in 1917¹³⁷. Einstein wished to develop a theory of cosmology in GR, but one in which the universe was static and unchanging, a commonly accepted notion at the time. The idea of a static universe is one in which the universe as a whole is spatially and temporally infinite, i.e. it extends forever in every direction and has always existed, and is neither expanding nor contracting.

Einstein's version of a static universe is slightly different, allowing the universe to remain temporally infinite yet possess some amount of spatial curvature rendering it spatially finite, which is called a *static eternal universe*. This was achieved by adding a positive *cosmological constant* to his equations of GR, counteracting the attractive nature of all the matter in the universe. Einstein's static universe is one which is closed — it contains a uniform distribution of dust and a positive uniform cosmological constant. However, the Einstein static universe possessed a lethal property: it is fundamentally unstable. The cosmological constant must be exactly equal to a specific value. Should the cosmological constant be even a miniscule amount larger, the universe would undergo expansion. If it were even a fraction smaller, the universe would collapse. This implies the Einstein static universe is unstable to small perturbations, rendering it untenable as we know the universe experiences perturbations everywhere, such as the existence of galaxies, clusters, and superclusters. The cosmological constant was then abandoned.

The return of the cosmological constant can be divided into four separate stages¹³⁸. In 1922, Alexander Friedmann showed that the equations of GR remain valid in the presence of a non-zero cosmological constant, provided the universe becomes dynamical. Then in 1927, Georges Lemaître showed that the universe is in fact expanding, combining GR with astronomical observations, in

particular those of Edwin Hubble in 1929. In the late 1960s, Petrosian, Salpeter, and Szekeres introduced the cosmological constant to explain a peculiar expansion history motivated by quasar data (though later data removed this particular motivation). Finally, recent observations of supernovae indicate the cosmological constant is needed after all. The existence of the cosmological constant has been compounded in recent decades by even more high resolution cosmological and astrophysical observations¹³⁶. It appears that the cosmological constant is a peculiar requirement in GR to explain the current view of the cosmos.

1.2.2 LIGHTS IN THE DARK

The current model of cosmology, one based on copious observational data, is that the universe is undergoing cosmic expansion and the rate of expansion is increasing. At the level of the Einstein equations, this can be explained by introducing a universe-spanning energy density. Since the nature of this energy is mysterious and largely unknown, it has been dubbed *dark energy*. The simplest implementation of dark energy in the Einstein equations is the cosmological constant, a homogeneous energy density throughout spacetime. Despite the fact that dark energy is exceedingly sparse, only 10^{-27} kilograms per cubic meter throughout the universe, current data suggests it composes a staggering 68% of the overall energy content of the cosmos. This suggests that this mysterious form of energy could arise from the fundamental properties of empty space itself. In other words, dark energy could be the energy of spacetime, though this is currently only speculative.

A common explanation of dark energy comes from the realm of quantum field theory. Vacuum energy is the lowest possible energy state of a quantum field. It turns out that the lowest energy state of QED is non-zero, implying that there exists a constant and homogeneous amount of energy associated with the QED vacuum throughout spacetime. Naturally, this provides a seemingly excellent description of dark energy as it portrays many of the same properties required to explain the cosmic expansion. The vacuum energy can be explained using the Heisenberg uncertainty principle, that is, the conjugacy between energy and time. This states that particle-anti particle pairs are continuously generated and mutually annihilated within a very short amount of time, inversely proportional to the energy of the particle-anti particle pair. With regards to the dark energy proposal, this means particle-anti particle pairs and their subsequent annihilation are responsible for

the cosmic expansion. However, there is an error with this proposal that is so large, it has its own name — the *vacuum catastrophe*. When the energy of the vacuum was calculated and subsequently compared to the observational data for the cosmological constant, the two values differed from between 50 to as much as 120 orders of magnitude, making it the worst prediction in scientific history¹³⁹⁻¹⁴³.

Naturally, other explanations for dark energy besides the cosmological constant have been proposed. First and foremost, the observational and theoretical evidence for dark energy are heavily based on GR itself. It's therefore conceivable that the classical theory of GR is wrong and another classical theory would be the more agreeable one, eliminating the need for dark energy. There are many such theories, in principle an infinite number, differing either in the exact mathematical form of the theory or in their interpretation of the gravitational interaction.

Another popular interpretation of dark energy is called *quintessence*. This is popular for two reasons: it introduces a new matter field with very simple behavior and it solves the coincidence problem. If acceleration occurred too early in the universe, then structures such as galaxies would never have had the opportunity to occur. Quintessence solves this issue by allowing the density of the quintessence field to 'track' the radiation density in the early universe until the matter-radiation equality, after which the quintessence field starts behaving as dark energy^{144,145}. Quintessence has neither been suggested by the data nor ruled out. More work, both observationally and theoretically, will need to be performed to confidently make statements on the viability of quintessence.

1.2.3 WHAT WE KNOW

The current knowledge on the true nature of dark energy is, naturally, extremely limited. The so-called 'late time accelerated expansion' of the universe is considered one of the most mysterious aspects of the standard cosmological model. In the base Λ -CDM model, the accelerated expansion is driven by the cosmological constant. This means that in the current standard model of cosmology, the cosmological constant is a purely phenomenological parameter without an underlying theoretical basis. Testing the standard paradigm can be carried out by parameterizing models of dark energy and then constraining those parameters using observational data. For a generic model, this means taking a particular ansatz for the *cosmological equation of state*, which is a rough classifier for model

types. The equation of state is characterized by the ratio of pressure and energy. For the standard cosmological paradigm, this ratio is generically time-dependent and spatially constant, and hence denoted by $\omega(t)$. The most recent cosmological data suggests that $\omega(t) = \omega_0 = -1.028 \pm 0.031$, i.e. a time-independent constant slightly smaller than -1 . The Λ -CDM model fixes $\omega = -1$ exactly. Hence, current observational data taken from the Planck satellite is in quite good agreement with the current cosmological model. Models that exhibit $\omega < -1$ are called *phantom energy* models and they lead to late-time behaviors characteristic of Big Rip scenarios.

Thus, current observational data is in good agreement with the accelerated expansion of the universe being driven by a dark energy density which is constant and unchanging throughout the universe with an equation of state parameter approximately equal to -1 . The nature of this dark energy density is wholly unknown and much more work, both theoretically and observationally, needs to be carried out to determine precisely what lies at the heart of empty space.

1.3 THE DARK FOREST

So far, we have discussed the Standard Model, which describes all known material particles in the universe. Current observations point to the fact that they contribute to about 5 % of all energy in the universe. We have also discussed the cosmological constant, a particular form of dark energy which Λ -CDM suggests is the reason for the accelerated expansion of the universe. Current observations suggest dark energy makes up about 68 % of all energy in the universe. Equally mysterious, the remaining content is colloquially known as *dark matter*. Current observations by the Planck satellite suggest dark matter contributes about 26 % to the total energy content of the universe, five times more than regular matter. It is a very weakly interacting form of matter that does not fit into the Standard Model, making it very difficult to detect using laboratory experiments. However, to our best knowledge, dark matter still interacts via gravitation, the same as all other forms of matter, making its cosmological impact easy to detect. There is a plethora of candidates for dark matter, including abandonments of its particle nature and suggesting it's a new form of gravity. However, a large body of observational evidence suggests the particle nature of dark matter. In the current age of the Standard Model, GR, and Λ -CDM, the mysterious nature of dark matter is one of the greatest challenges in the modern era of physics and cosmology.

1.3.1 HISTORY

The history of dark matter, much like GR, has been one of turbulence¹⁴⁶⁻¹⁵¹. Just a decade after Einstein introduced his new theory of GR, physicists and astronomers attempted to estimate the total amount of luminous material in the galaxy¹⁵²⁻¹⁵⁴. Several scientists observing the data noticed a discrepancy between the velocity dispersions of the rotating material around the galactic center and the amount of visible, luminous material comprising the galactic disk¹⁵⁵⁻¹⁵⁷.

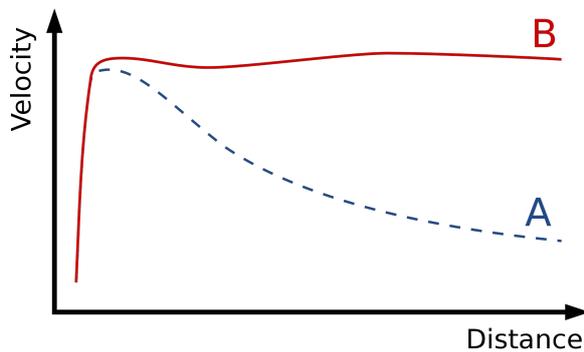


Figure 1.14:¹⁵⁸ Expected orbital speeds of stars in the Milky Way (A) versus the observed orbital speeds (B). This discrepancy represents one of the biggest smoking guns for the existence of dark matter.

These discrepancies were not particularly problematic until much later when astronomers began to observe a discrepancy between the relatively 'flat' rotation curves of galaxies and the predicted 'declining' curve deduced from the visible material comprising that galaxy, fig. 1.14. The modern-day acceptance of the existence of dark matter is owed to the definitive cumulative evidence resulting from two different branches of astronomy — that of the

high velocity dispersions in clusters and the flat galactic rotation curves. It was not until the 1970s when the evidence for a yet undetected, invisible, and very weakly interacting form of matter was sufficient to convince the larger scientific community. Perhaps the most famous observation in support of the existence of dark matter was performed in 1933 by Fritz Zwicky, who studied the velocity dispersion of the Coma cluster^{155,159}. Zwicky estimated the mass of the Coma cluster based on the mass at the edge of the cluster and then again based purely on its brightness. He estimated that the cluster had about 400 times as much mass compared to what could be seen. The velocities of the constituent galaxies was too high to explain the amount of visible matter, suggesting some other matter must be present that is obscured from view, concluding that the bulk of the matter was invisible. Further evidence for dark matter also came in the form of mass-to-light estimations.

Finally, perhaps in the most influential paper regarding dark matter in history, Vera Rubin and Kent Ford used a new spectrograph to measure the orbital velocity curves of spiral galaxies with great accuracy^{160,161}. They studied the orbital velocities of ionized hydrogen clouds in the M31

galaxy, confirming the presence of invisible matter being responsible for increased orbital velocities compared to standard predictions from the luminous material. At the same time, radio astronomers were measuring interstellar atomic ionized hydrogen clouds, since these typically extend much farther than the visible stellar objects. The rotation curves traced using these measurements further compounded the presence of invisible matter. More studies continued to build a mountain of evidence in favor of invisible matter.

1.3.2 WHAT WE KNOW

To date, there is a large amount of observational evidence for the existence of dark matter that can be classified based on the type of data collected. The first are the ones utilized historically, that of galactic rotation curves and velocity dispersions. These simply use frequency data of the measured electromagnetic radiation and known astrophysical stellar/galactic processes to measure velocities of the galactic constituents. These were used by Fritz Zwicky and Vera Rubin in their pioneering work.

The second class of observations regard galactic clusters. These are particularly important for dark matter estimations since the mass distributions of the cluster can be estimated using several methods. The first is the velocity of the constituents along the Earth-cluster line-of-sight. The second is x-ray data emitted by hot intergalactic gas within the cluster. These measurements are typically in agreement with the estimation that dark matter outweighs normal matter in a ratio of 5 to 1, as suggested by the cosmological data of Planck¹⁶². Besides rotation curve and galactic cluster estimations,

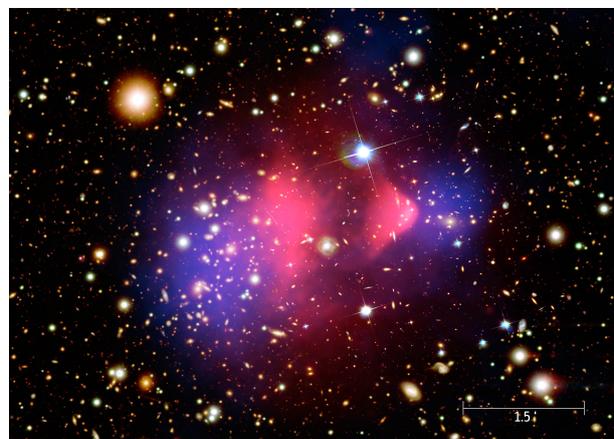


Figure 1.15: X-ray data (pink) collected by the Chandra X-ray Observatory superimposed over optical data, with the matter distribution computed using gravitational lensing (blue). The x-ray data represents the bulk of the baryonic (normal) matter composing the galaxies, while the gravitational lensing data shows all the material composing the galaxies, providing one of the best pieces of evidence for the existence of dark matter.

gravitational lensing is also a powerful tool in estimating the mass of a galaxy or cluster. The primary phenomenon leveraged for these analyses is the fact that dark matter interacts gravitationally

just like any other matter, allowing it to form a gravitational *lens*. This effect was capitalized on in recent decades to provide extremely strong evidence for the existence of galactic dark matter halos^{163,164}.

The bullet cluster provides one of the strongest and widely popular pieces of evidence in favor of the material nature of dark matter, suggesting dark matter is a new form of matter and not a phenomenon associated with an alternative form of gravity, fig. 1.15. The cluster consists of two colliding clusters of galaxies 3.7 billion light years from Earth. The Chandra X-ray Observatory, combined with optical data from the Hubble Space Telescope, can image the gaseous components of the cluster, forming estimations on the total amount of visible material. Gravitational lensing maps then provide a secondary estimation of the total amount of material and its distribution in the colliding clusters. Comparing these two observational datasets provides strong evidence for the existence of material present in the colliding clusters that does not interact with the electromagnetic spectrum.

Cosmological data provides another powerful avenue to probe the dark sector of the universe. Since dark matter behaves differently than normal matter, the CMB will show different imprints depending on the nature of the material. The small-scale differences in the temperature of the CMB can be utilized to measure the density of dark matter¹⁶⁵.

There are several other methods to measuring cosmological and astrophysical distributions of dark matter, including Type Ia supernovae, redshift-space distortions, and the Lyman-alpha forest. All these methods are in agreement with the Λ -CDM cosmological model, providing evidence for the material nature of dark matter.

1.3.3 CANDIDATES AND EXPERIMENTS

Since the true particle nature of dark matter has yet to be confirmed, and with little evidence in the way of any one model, there have been many candidates suggested over the years, and several have had large, expensive detectors built to find them. The types of candidates suggested originate from almost all of fundamental physics, including new, undiscovered particles beyond the Standard Model, extra dimensions, black holes, and even new forms of gravity.

With regards to new particles, these candidates can be split into two broad categories: baryonic

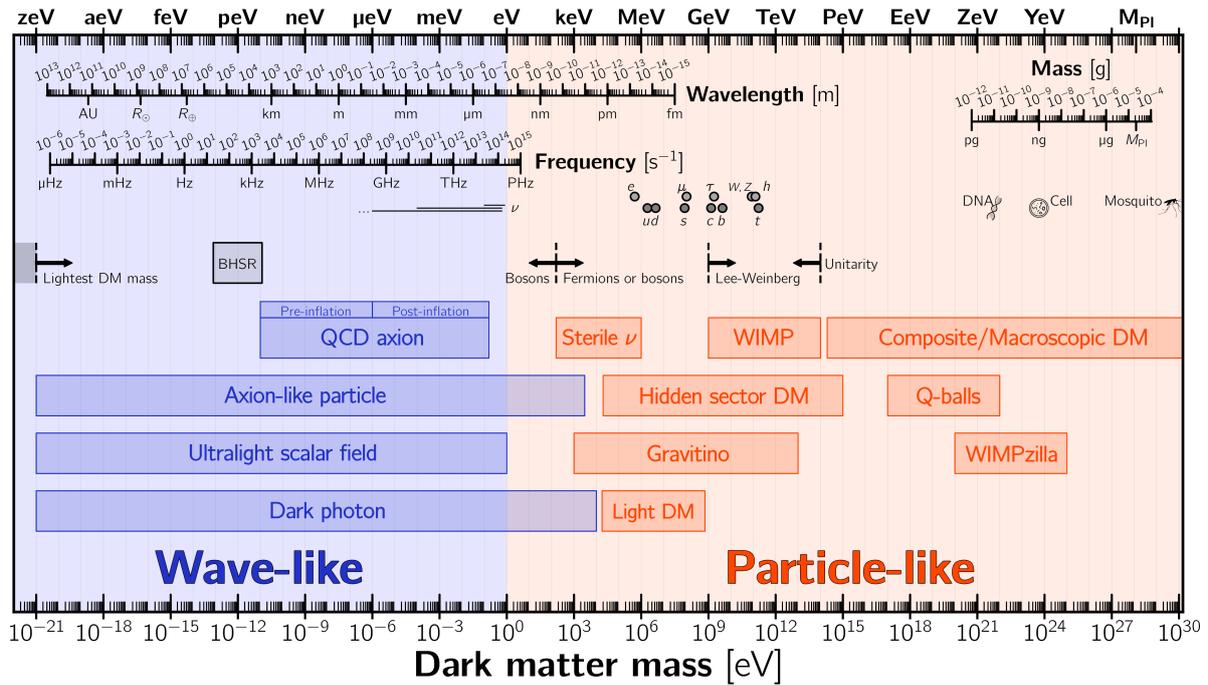


Figure 1.16: ¹⁶⁶ Different types of dark matter candidates as a function of mass. The little knowledge on the true nature of dark matter translates to a huge uncertainty in the mass of the theoretical particle ranging from several times the mass of the sun, in the case of black holes, to ten trillion trillion times lighter than a proton.

and non-baryonic. The baryonic candidates are those which are composed of baryons, such as neutron stars, burnt-out dwarfs and other massive objects. These are collectively called MACHOs, or Massive Compact Halo Objects. These were initially the most popular candidates, since they provide a simple explanation ^{167,168}. However, these types of candidates have largely run out of favor due to lack of evidence for such a high amount of baryonic matter in the early universe.

Non-baryonic matter is currently the most favored candidate for dark matter and can be represented in two large classes: hypothetical new particles and primordial black holes. The most popular hypothetical particles are the Weakly Interacting Massive Particle (WIMP), the axion, sterile neutrinos, the dark photon, and supersymmetric particles. Most of these particles have well-motivated production mechanisms, making them good candidates for dark matter.

THE FIRST BLACK HOLES

Primordial black holes are an interesting candidate since they have a fairly simple production mechanism and are quite distinct from the elementary particle candidates. Primordial black holes (PBH) are black holes that were formed in the very early universe when the causal structure of spacetime was rapidly changing. Their accepted formation scenario is that they were formed when certain re-

gions in the random density fluctuations of the inflationary epoch and early photon era underwent gravitational collapse. Since their formation pre-dated the standard stellar formation era, they could in principle form with a much wider mass range, even in the subatomic mass scales. This means PBHs could be in existence today that have the mass of our moon, the sun, or an asteroid and they could be moving at enormous speeds¹⁶⁹.

There are several possible ways their existence could be inferred using observations. The first is via direct observation of the radiation emitted during their evaporation via Hawking radiation. A second detection mechanism is via temperature fluctuations in the CMB, since PBHs would affect the energy distribution in the CMB. The third avenue for detection relies on the assumption that PBHs only account for a fraction of the total dark matter, and the rest in the form of some sort of fundamental particles. This would imply the particle dark matter could accumulate around the PBHs and, provided the particle dark matter undergoes some sort of annihilation, be detectable as a radiation-emitting halo. PBHs could also form binary systems with themselves, leading to potential gravitational wave observations in the near future when the LISA telescope begins its data collection phase.

Current observations for PBHs come from the NASA Fermi Gamma-ray Space Telescope, Planck space observatory, and LIGO/Virgo. Near future observatories that could conceivably constrain the PBH mass range are the Square Kilometer Array, gravitational wave observatories such as the Laser Interferometer Space Antenna (LISA), pulsar timing arrays, the Vera C. Rubin Observatory, very large arrays, and other electromagnetic observatories.

CLEANING UP THE STANDARD MODEL

Axions are another very common candidate and are part of a larger class of particles called WISPs, or weakly interacting sub-eV particles. This is a generic acronym for particles that interact, as the name suggests, only very weakly with all other matter. The neutrino is a non-dark matter example of a WISP. The axion has quite a long history, being originally postulated independently of the dark matter hypothesis. They were suggested as a possible resolution to the so-called strong CP problem in QCD, which is the question of why QCD preserves a certain kind of internal symmetry, called charge-parity symmetry¹⁷⁰. This is where the name 'axion' derives from, being named

by Frank Wilczek after a cleaning detergent. In 1977, Roberto Peccei and Helen Quinn postulated a mechanism nowadays called the Peccei-Quinn mechanism which naturally solves the strong CP problem by introducing a new particle, now called the axion. QCD effects in the early universe cause a large cosmological population of axions¹⁷¹⁻¹⁷³. Within a certain range of the axion parameter space, these cosmological axions could account for dark matter^{174,175}. This makes the axion a very promising candidate for the dark matter particle, due to the fact it both solves the strong CP problem and is an excellent explanation for cosmological dark matter.

However, as is usual for dark matter candidates, searching for axions via direct detection is extremely challenging. Not only does the axion interact very weakly with the Standard Model, it is also extremely light. Nonetheless, the presence of the axion modifies the Maxwell equations of electromagnetism. This implies the axion can convert to photons and hence be measured as an excess of electromagnetic energy. This is the idea behind the Axion Dark Matter eXperiment (ADMX)¹⁷⁶⁻¹⁷⁹. The experiment uses a resonant microwave cavity that searches for axion-to-photon conversion of axionic dark matter in the local galactic halo. Thus far, the experiment has only produced null results. However, continuous upgrades and searches could reveal an electromagnetic excess. Besides terrestrial searches, helioscopes are another avenue for axion detection since axion conversion could be stimulated by the strong solar magnetic fields. They could also be generated in the extremely high magnetic field in the magnetosphere of neutron stars. Astronomical imaging, such as imaging of the M87 core by the EHT collaboration, is also used to constrain the axion parameter space.

A DARK MATTER MIRACLE?

The Weakly Interacting Massive Particle (WIMP) is one of the most popular dark matter candidates. There is no formal definition of a WIMP, other than the fact that it acts gravitationally and is at least as weak as the weak force. Many different WIMP candidates are expected to be produced in the early universe. For the correct dark matter abundance today, the WIMP must have a mass in the range of $100 \frac{\text{GeV}}{c^2}$ ¹⁸⁰. Certain extensions of the Standard Model, called supersymmetric extensions, naturally predict the existence of a new particle in this mass range.

This apparent coincidence was an exciting prediction and was called the *WIMP miracle*¹⁸². WIMP-like particles are also predicted by models that include universal extra dimensions.

There are many experiments underway hoping to discover a WIMP. Indirect observations aim to discover WIMP annihilation signals or decay products. These usually will occur in overdense regions of dark matter halos, such as those around galactic centers, clusters, and even black holes. Halos that form around binaries would alter the waveform of the gravi-

tational wave radiation, allowing gravitational wave observatories to search for WIMPs alongside other channels. Typical indirect searches look for gamma ray radiation excesses, whose spectral signature depends on the specific model of the WIMP. This allows telescopes such as the Fermi-LAT gamma ray telescope or the VERITAS ground-based gamma ray observatory to place constraints on the WIMP parameter and model space. Future observations from the IceCube Neutrino Observatory may also constrain the WIMP mass. Similar to axion helioscopes, WIMPs interacting with solar photons could also be measured.

Direct detection schemes are also underway. These cover a large range of different detection schemes, such as cryogenic crystal detectors, scintillators, bubble chambers, and time projection chambers. To date, there have been no detection confirmations.

THE DARK PHOTON

The dark photon is another interesting dark matter candidate, owed in part to its curious origination beyond the Standard Model, its various production mechanisms, and the implications of its existence. The dark photon is a new force carrier beyond the usual four that interacts very weakly with electrically charged matter^{183,184}. The existence of such a new force carrier is quite ubiquitous in new physics scenarios. These new forces are said to reside in a *hidden sector*, a portion of the the-

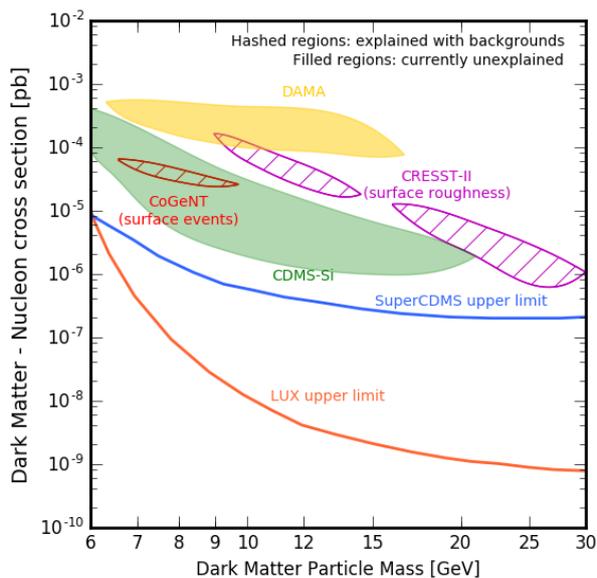


Figure 1.17: ¹⁸¹ Limits on the WIMP parameter space constrained by various direct detection experiments.

ory that very weakly interacts with the rest of the model, within which particle-type dark matter resides. There are many different ways the dark photon could couple to the Standard Model, each giving rise to a new model.

Mass for the dark photon can naturally arise in several scenarios. It could manifest via a Higgs mechanism, similar to the Standard Model matter sector¹⁸⁵⁻¹⁸⁸. This mass production mechanism could produce masses in the MeV-GeV range, often utilizing some form of supersymmetry. However, much smaller masses are possible using a Stückelberg mechanism, which is a theoretical technique to restore some sort of symmetry to a vector theory (such as the dark photon theory) that otherwise is devoid of it. The use of the Stückelberg mechanism has particular relevance in large volume string compactifications with branes, a concept from string theory^{189,190}. The amalgamation of these various models lead the dark photon mass range to span from the meV scale to the TeV scale. Besides coupling to the electromagnetic sector of the Standard Model, the dark photon can also couple to the weak sector, colloquially called the "dark Z", due to its coupling to the Z boson.

Technically, there exist two different kinds of dark photons. The first is massless and cannot interact directly with any of the Standard Model particles. The second is the massive kind that can couple directly to the Standard Model sector. The mass of the massive dark photon covers the enormous range that was just discussed, from meV to TeV scales. However, this mass parameter space can be split into two separate regions, centered around 1 MeV, which is twice the mass of the electron, the lightest Standard Model particle¹⁹¹. The dark photon is said to be *visible* if its mass is greater than 1 MeV, since it can decay into Standard Model charged states which leave a distinct signature in detectors. However, the dark photon could also decay into dark sector states, provided those states are light enough. If the dark photon mass is less than 1 MeV, then it can no longer decay into Standard Model particles and its decay products are hence invisible. The searches for these dark photons rely on searching for energy deficits in astrophysical processes like stars or in direct search schemes.

Collider experiments are also able to search for dark photon productions. These include searching for meson decays, bremsstrahlung, and annihilation signals. Data collection at the LHC and SLAC provide precise experimental data that further constrains the allowable mass range for the

dark photon. For the mass range less than 1 MeV, atomic and nuclear experiments provide further constraints. They aim to detect modifications of the Coulomb force due to the presence of a dark photon. Additionally, corrections to the atomic energy levels of the hydrogen atom translate into bounds on the dark photon parameters. Further, searches for axion and axion-like particles can be translated into constraints on the dark photon parameter space. For example, the experimental data from ADMX has been leveraged to place additional bounds¹⁹².

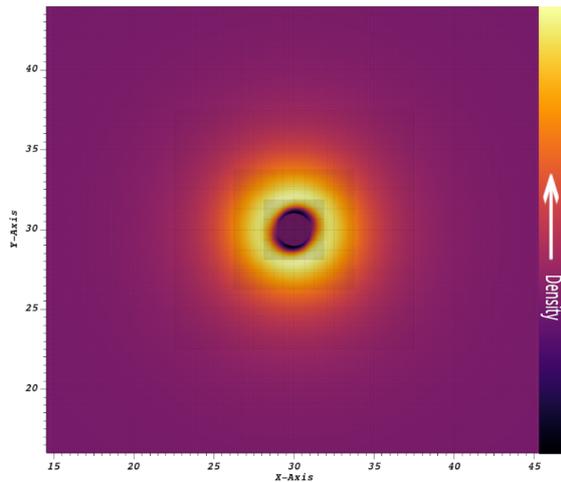


Figure 1.18: Superradiant build-up of a massive dark photon cloud around a rapidly spinning black hole. This image was generated using a numerical evolution of the dark photon equations of motion on a Kerr black hole. The finite grid on which the dark photon was calculated is visible as a hierarchy of boxes.

black hole, fig. 1.18. If these clouds form in binary systems, they can modify the orbital dynamics and resulting gravitational radiation, leading to possible detection scenarios for gravitational observatories. Since the instability develops by channeling energy from the black hole into the matter field, statistical studies can be done on black hole populations, and measure the statistical distributions of mass and spin of the black holes. With a fine enough model for the statistical distributions of 'naked' black holes, the presence of dark photon clouds could be statistically determined. Follow up searches could then be utilized for individual black hole systems to look for the clouds directly.

One of the major studies performed in this thesis pertains to detecting such superradiant clouds using extreme mass-ratio inspirals and detecting the gravitational radiation with the LISA mission. All details on this study are covered in Section 2.

Astrophysical processes also allow further constraints on the dark photon parameter space, opening the door for gravitational wave observatories to behave as cosmic particle detectors¹⁹³⁻¹⁹⁹. A common scenario discussed in the literature in these searches is that of superradiance. This is a process by which a matter field surrounding a spinning black hole, like a Kerr black hole, undergoes a rapid instability, causing the matter field to absorb energy from the black hole and grow exponentially²⁰⁰. The growth of the field will eventually saturate, leading to the formation of a cloud around the

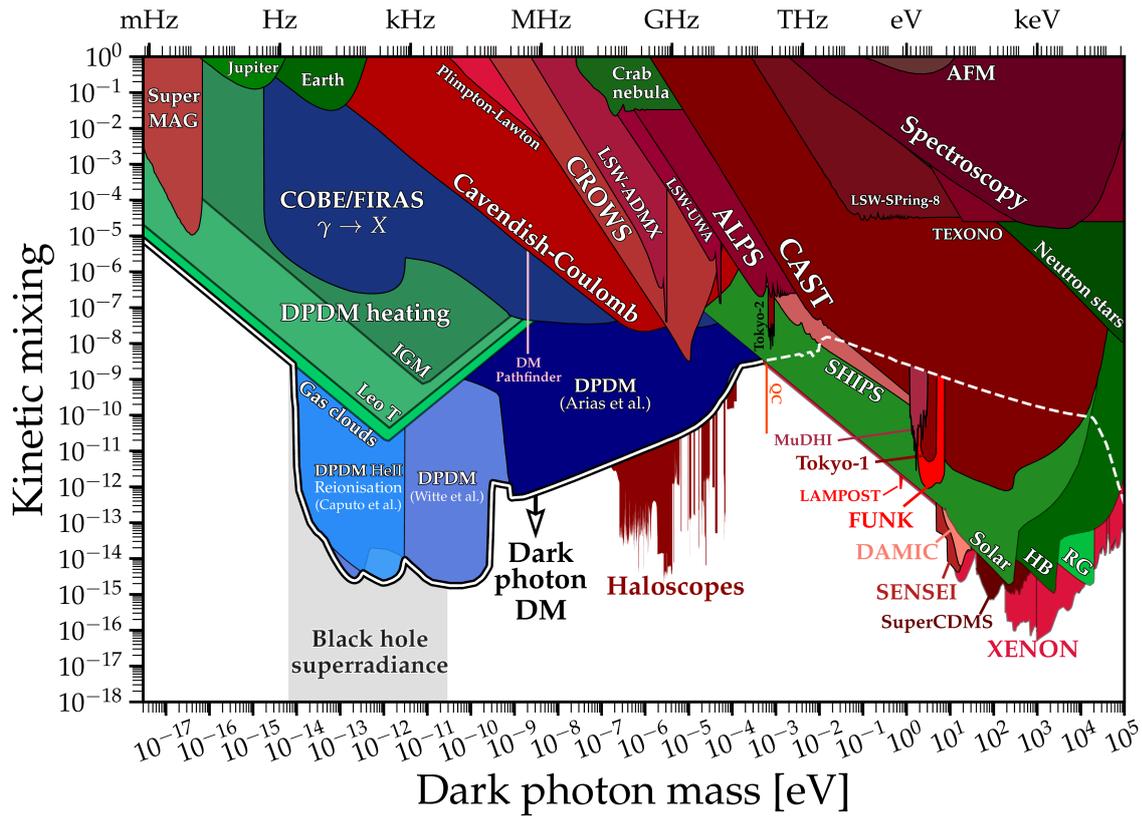


Figure 1.19: ¹⁸¹ Limits on the dark photon parameter space from various channels, including astrophysical and collider types.

The implications of a dark photon discovery would be immense. This would imply not only a possible detection of dark matter, but could also imply an entirely new force carrier in nature that lies beyond the Standard Model ²⁰¹.

1.4 HEAVY PHOTONS AND BEYOND

The Standard Model of particle physics can be mathematically described by a single Lagrangian, fig. 1.3. Adding new fields to the Standard Model, such as the dark photon, amounts to adding new mathematical terms to the Standard Model Lagrangian. The type of terms that are added and their mathematical nature depends on the specific model of dark photon under consideration. In mathematical terms, adding a new dark photon to the Standard Model is accomplished by adding a new $U(1)$ gauge field*. The Standard Model photon, responsible for the electromagnetic interactions, is a type of $U(1)$ gauge field. For the case of a massive dark photon, various mechanisms exist for

* $U(1)$ is a type of very simple symmetry inherent to the dark photon, which can be thought of as a type of rotational symmetry. A gauge field is a type of field that possesses certain types of invariances under localized transformations ²⁰².

generating mass for the gauge field, which explicitly breaks the $U(1)$ symmetry. These could be the Higgs mechanism or a Stückelberg mechanism, depending on the specific model²⁰³.

The most popular basic model for the dark photon can be represented by the so-called *Proca Lagrangian*, fig. 1.20.

The Proca theory can be traced back to Alexandru Proca's 1936 paper where he introduces a massive alternative to the standard wave equation²⁰⁴. Later refinements showed that this new vector equation is the result of a massive spin-one force carrier, i.e. a massive photon²⁰⁵. The two phenomenologies of the photon and this new massive photon hence,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m^2 A_\mu A^\mu$$

Figure 1.20: The Proca Lagrangian describes a massive vector field.

will be very different. For example, since the Proca field is massive, it will now propagate an additional degree of freedom.

If the photon itself were massive, the simplest model for the new electrodynamics would be the Proca theory. The most common theory for the massless photon is, of course, the Maxwell electrodynamics, however it's conceivable that the photon mass is miniscule. If the photon were to carry mass, the speed of electromagnetic waves would develop a frequency dependence. Other effects would also be apparent, such as modifications to the Coulomb law. High precision tests using hollow conductors have placed an upper bound on the mass of the photon of $m \leq 10^{-14} \frac{eV}{c^2}$ ²⁰⁶. Astrophysical processes can also place a limit on the photon mass. For example, experiments designed to detect effects caused by the galactic vector potential have placed extremely small upper bounds on the photon mass. The main idea is that the additional mass term in the Proca Lagrangian, fig. 1.20, would affect the galactic plasma²⁰⁷. The generation of magnetoacoustic waves in the galactic plasma cannot have frequencies lower than a certain critical frequency which depends on the mass of the photon. The lack of observation of a lower frequency bound has placed limits on the photon mass of $m < 3 * 10^{-27} \frac{eV}{c^2}$, which is tremendously smaller than the terrestrial experiments. These sharp upper limits from the non-observation of effects generated by the galactic potential are model-

dependent²⁰⁸. If the photon mass is generated by the Higgs mechanism, the upper limit is $m \leq 10^{-14} \frac{eV}{c^2}$ from the Coulomb law experiments.

1.4.1 GOING ABOVE AND BEYOND

More recent proposals attempt to generalize the Proca theory to include additional interactions that extend the standard Proca terms. Such theories go by the name of *generalized Proca* theories^{209,209,210}. The approach to generalizing the Proca theory is to include additional derivative self-interactions, which are carefully chosen to preserve the three degrees of freedom of the Proca field. The motivation for these new interactions extends beyond previous extensions to electromagnetic theories. Generalizing the standard Einstein equations to include additional degrees of freedom could reconcile the dark energy problem. Introducing these new modes of the field could give rise to a late-time acceleration of the universe, circumventing the need for a cosmological constant²¹¹. It has been shown that a subsector of the full generalized Proca theory can give rise to de Sitter solutions, solutions that possess a positive cosmological constant. It has also been shown that the full generalized Proca landscape generically gives rise to de Sitter solutions. This implies that generalized Proca models can explain the late-time accelerated expansions of the universe without invoking a mysterious energy content.

Generalized Proca theories have also been studied in the context of black hole solutions^{212–215}. These solutions will be important in searching for astrophysical signals of generalized Proca fields, for example in gravitational wave signals, dynamical friction of black holes through dark matter halos, etc.

Recent studies have uncovered an interesting issue that arises when generalized Proca fields are allowed to evolve on spinning black hole backgrounds^{216–220}. Particular cases of generalized Proca theories have been studied on both flat spacetime and Kerr black holes, which have shown that the temporal evolution of the field equations cease to be well-defined after some point in the evolution. This can be traced back to a breakdown in the field equations themselves. This implies that the generalized Proca model is ill-defined and hence would not be a viable candidate for dark matter systems. These studies themselves are limited in their scope, since they don't include the backreaction of the Proca field on spacetime which, by the Einstein equations, would modify the dynamics.

Nonetheless, this suggests that some models within the generalized Proca landscape are intractable. Further details on this problem will be elucidated in appendix C.

The landscape of generalized Proca theories could harbor solutions to some of physics largest problems, including the nature of dark energy and dark matter. However, theoretical issues still exist that must be well understood in order to narrow down the theories that have astrophysical relevance. Much more work is still needed to determine which class of generalized Proca theories, if any, could solve the biggest problems in physics.

1.5 PROBING THE UNIVERSE WITH THE COMPUTER

The fundamental equations of GR are the Einstein equations. In order to make predictions about astrophysical processes, the Einstein equations are the most widely used equations that predict the trajectories of matter and shapes of astrophysical objects. The fundamental object that describes the 'shape' of spacetime is the metric. Given a particular distribution of matter, the Einstein equations will dictate what the shape of spacetime is. Equivalently, given a particular shape of spacetime, the Einstein equations will dictate how the matter will move. The famous aphorism in regard to this is given by John Archibald Wheeler: "Spacetime tells matter how to move and matter tells spacetime how to curve".

$$\begin{aligned} & \frac{1}{2}g^{\alpha\beta}\partial_\alpha\partial_\mu g_{\beta\nu} + \frac{1}{2}g^{\alpha\beta}\partial_\alpha\partial_\nu g_{\mu\beta} - \frac{1}{2}g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu} - \frac{3}{2}g^{\alpha\beta}\partial_\mu\partial_\nu g_{\alpha\beta} \\ & - \frac{1}{2}g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\mu g_{\beta\nu} - \frac{1}{2}g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\nu g_{\mu\beta} + \frac{1}{4}g^{\beta\lambda}g^{\alpha\rho}\partial_\nu g_{\alpha\lambda}\partial_\mu g_{\rho\beta} \\ & - \frac{1}{4|g|}g^{\alpha\beta}\partial_\beta |g| \partial_\alpha g_{\mu\nu} - \frac{1}{4|g|}g^{\alpha\beta}\partial_\beta |g| \partial_\mu g_{\alpha\nu} + \frac{1}{4|g|}g^{\alpha\beta}\partial_\beta |g| \partial_\nu g_{\mu\alpha} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \end{aligned}$$

Figure 1.21: The Einstein equations written out in terms of the metric. The metric $g_{\mu\nu}$ is a four-by-four symmetric matrix, so the Einstein equations are a set of ten partial differential equations, each equation a highly non-linear expression. The symbol $T_{\mu\nu}$ denotes the energy and momentum of the matter field. In a vacuum, $T_{\mu\nu} = 0$.

However, compared to the aged Newtonian theory of gravity, the Einstein equations are vastly complex, fig. 1.21. They consist of ten non-linear partial differential equations that must be solved with appropriate boundary conditions. Further, if matter fields are present in the spacetime, the equations of motion for the matter fields must be solved in tandem with

the Einstein equations, increasing the complexity several times. Only a handful of exact solutions to the Einstein equations are known. The list of astrophysically relevant solutions is even smaller. These solutions usually leverage certain symmetries or algebraic properties in order to simplify the equations. For example, the Kerr black hole uses rotational symmetry in a vacuum while the

Reissner-Nordström black hole assumes complete spherical symmetry in the presence of an electromagnetic field. However, most astrophysical processes are extremely dynamical and, hence, usually do not possess nearly as many symmetries, making it difficult to develop astrophysically relevant exact solutions.

The usual recourse for developing solutions to astrophysical problems is to resort to either approximations or numerical methods. Approximations can be very powerful in their predictive power, such as the Teukolsky master equation²²¹. The most powerful method utilized today to develop highly accurate solutions of the Einstein equations is leveraging computational power to solve them. This is the vast field of numerical relativity. Due to the extreme complexity of both the astrophysical environment and the Einstein equations themselves, supercomputers or high-performance computing clusters are often deployed for solutions.

Numerical relativity covers almost all aspects of astrophysics and cosmology, including black hole perturbations, neutron stars, merging astrophysical objects in binaries, and cosmic evolution models. There is quite some freedom in choosing the exact form of the Einstein equations, thanks in part to the covariant nature. All methods of solving the Einstein equation must solve two broad problems. The first is initial data, which must be carefully chosen so as to be consistent with the Einstein equations. The second is the temporal evolution of the metric field. These two problems require different methods. The first, choosing suitable initial data, requires either an intelligent guess or solving elliptic constraint equations. The second is purely numerical and requires solving hyperbolic equations.

1.5.1 HISTORY

The field of numerical relativity emerged from the desire to study more general solutions of the Einstein equations that are not tractable analytically. A widely enforced precondition for numerical solutions of the Einstein equations is a mathematical decomposition of spacetime back to separated space and time*. The first such formalism can be traced back to the late 1950s to the work of Richard Arnowitt, Stanley Deser, and Charles W. Misner, which is nowadays known as the ADM formalism²²². However, for technical reasons, the exact form in their original paper of the decom-

*There are also techniques that do not perform this split.

position isn't usually used anymore, instead opting for the "3+1 formalism", which decomposes the four-dimensional spacetime into three dimensions of space and one dimension of time.

At the time the ADM decomposition was published, computational power was insufficient to handle the task of performing numerical computations. The first documented case of attempts to numerically solve the ADM equations appears to be in 1964 by Susan Hahn and Richard Lindquist²²³. These early attempts were focused on evolving simple rotationally symmetric data. Around the same period, Tsvi Piran wrote one of the first codes to evolve a system containing gravitational radiation in a cylindrically symmetric system²²⁴. This body of work set the groundwork for many concepts still in use today in solving the ADM equations, such as 'free evolution' versus 'constrained evolution'. These two different approaches deal with the constraint equations present in the ADM formulation in different ways. Additionally, applying symmetry to the problem reduced the computational power and memory requirements, allowing the code to run on available supercomputers at the time.

The earliest realistic results were carried out in the 1980s by Richard Stark and Tsvi Piran²²⁵, when they calculated the gravitational radiation produced from the formation of a rotating black hole. These results stood among the very few numerical relativity results for nearly 20 years, owing to the vast computational power required. Then, in the 1990s, the Binary Black Hole Grand Challenge Alliance successfully simulated a head-on binary black hole collision. In the post-processing stage, they computed the event horizon, although still imposing an axisymmetry of the system²²⁶.

The earliest work to solve the Einstein equations in full three-dimensional space focused on a single spherically symmetric black hole, since this provides an excellent test of the numerical method. First, it's a numerical solution of an already exactly known solution, so the accuracy of the numerical method can be determined by simply comparing it to the known exact solution. Second, it contains one of the most challenging features of GR to numerically handle, the curvature singularity. In the following years, two key advances greatly aided the numerical relativity community — the advancement in computational power and new theoretical tools to improve the simulation's efficiency. In regard to black hole spacetimes, two theoretical concepts were developed to aid the handling of the physical singularity, the idea of 'excision' and the 'puncture' method. Additionally, adaptive mesh refinement techniques were introduced into the numerical relativity field.

The idea of excision is quite simple — during the course of the simulation, continuously cut out a portion of the spacetime within the horizon, fig. 1.22. This resolves the issue of handling the physical singularity at the center of the black hole. This idea was developed in the late 1990s by Miguel Alcubierre and Bernd Brügmann²²⁸. The main reason this works and does not contaminate the rest of the spacetime

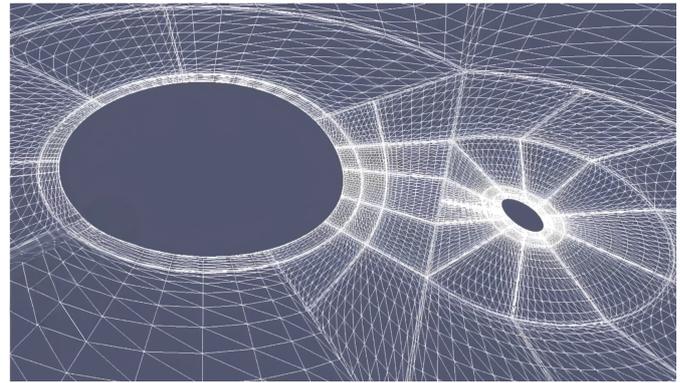


Figure 1.22:²²⁷ Example of excision of a black hole singularity on the computational grid of the simulation.

with noise is because causality prevents information from escaping the event horizon. Thus, errors that could accumulate at the excision boundary will not contaminate the region outside the event horizon, which is the most physically interesting. The main issue with these implementations is twofold. First, while physical information will not escape the event horizon, co-ordinate information could, for example, if the co-ordinate data was elliptic. Second, as the black hole moves, the excision zone will need to move in tandem. Thus, the simulation will need to continuously track the location of not only the singularity, but also determine the location of the horizon, a highly non-trivial task. The first stable, long-term evolution of the orbit and subsequent merger of two black holes using the excision method was in 2005²²⁹.

The puncture method is markedly different from the excision method in that no removal of the singularity throughout the simulation is required. Instead, the analytic part of the solution that contains the singularity is factored out²³⁰. Until 2005, all published usage of the puncture technique required fixing the singularity to a specific co-ordinate location throughout the entire simulation. Since black holes tend to move under the influence of their mutual mass, this caused the co-ordinate grid to become twisted and deformed, resulting in instabilities.

Finally in 2005, a breakthrough occurred, resulting in the year being called the *annus mirabilis of numerical relativity*. A group of researchers demonstrated that a specific choice of co-ordinate conditions along with intelligent choices for the behavior of the singularity, they demonstrated the ability to allow the singularity, or 'puncture', to move throughout the co-ordinate system. This allowed numerical solutions for two black holes orbiting each other to be obtained, as well as accurate

extraction of the gravitational radiation emitted by the binary. This was a remarkable achievement, demonstrating a huge leap forward in the numerical relativity community, allowing accurate models of strong gravitational events to be developed.

Adaptive mesh refinement (AMR) is another concept that greatly increased the efficiency and accuracy of numerical solutions, which was borrowed from computational fluid dynamics. Mesh refinement first appeared in the 1980s, thanks to the work of Choptuik in his studies of critical collapse of scalar fields²³¹. The first studies were performed in one dimension and later extended to two²³². AMR has also been applied to cosmological studies. In today's active field of numerical relativity, AMR is a widely used technique and has been applied to compact object mergers and simulations of the resulting gravitational radiation^{233,234}.

1.5.2 COMPLEXITIES AND CHALLENGES

The field of numerical relativity deals with many challenges that range over two broad disciplines, mathematical relativity and computer science. The challenges faced in mathematical relativity are theoretical challenges that pertain to the causal structure of the problem at hand, co-ordinate difficulties, and well-posedness of the evolution equations themselves. The difficulties faced on the computational side include developing fast code, computer memory handling, and all the challenges brought by parallelization. The solutions to these issues took several decades of intense work to amend, or in the case of computational power, waiting for more advanced processors and computational methods to be built.

Understanding each individual challenge is vital to developing highly efficient and parallelizable code that can fully leverage the latest computing technology.

1. Complexity and Nonlinearity of the Evolution Equations

- **Nonlinearity:** The Einstein field equations are extremely nonlinear, meaning that small perturbations in the initial data for solutions or parameters can lead to large variations in the final state. These nonlinearities make finding stable solutions difficult.
- **Coupling of Equations:** The Einstein equations themselves are a system of coupled partial differential equations, where one solution depends on the solution of another. This system makes finding numerical solutions arduous.

2. Dimensionality and Computational Cost

- **High Dimensions:** The fundamental structure of spacetime is a four-dimensional manifold. This means numerical simulations often need to be performed in a three-dimensional space and evolution through time, requiring significant computational resources. Leveraging symmetries in the problem allows one to reduce the dimensionality of the simulation or reduce the size of the computational grid.
- **Resolution and Scale:** High resolution is vital to capture fine detail, especially near strong gravitational sources where the metric is changing rapidly, like near a black hole. This leads to a large number of grid points and a high computational cost, vastly increasing processor and memory requirements.

3. Stability and Accuracy

- **Numerical Stability:** Ensuring numerical stability over long simulations is a challenging art. Instabilities can arise from the discretization methods or from inherent dynamics of the system. Tracking the origin of these instabilities and whether they arise from the numerical method or from the system itself is vital to ensuring accurate simulations.
- **Accuracy:** Maintaining high accuracy is crucial for long-term stability of simulations. This is especially true for extreme conditions, such as near the event horizon or singularities. Small numerical errors can grow over time, tarnishing the simulation.

4. Constraint Equations

- **Constraint Preservation:** The Einstein equations are a constrained system, meaning there are constraints in the equations that must be satisfied at all times. Violations of constraints imply the simulation is no longer solving the Einstein equations. Numerically ensuring that these constraints remain satisfied throughout the entire simulation is formidable and several different schemes have been developed to handle this challenge.

5. Boundary Conditions

- **Artificial Boundaries:** Simulating an entire spacetime is an extremely complex undertaking and only recently have efforts been made in this direction. Almost all simulation software actively utilized in research employs a finite computational domain, requiring artificial boundary conditions at the outer edges. Choosing the appropriate boundary conditions, which is a highly system-dependent choice, is vital to ensure unphysical reflections or inaccuracies are not present at the boundaries.
- **Asymptotic Boundaries:** For systems such as black hole mergers, an ideal boundary is at infinity, however this is impractical since this would either require an infinite number of grid points or an entirely new computational method. This approach is still in its infancy.

6. Singularities and Black Holes

- **Spacetime Singularities:** Singularities, such as those within a black hole, pose a tremendous numerical challenge. Many computational quantities tend to diverge near this point, hence methods to numerically handle them are vital for stability. Techniques such as excision or puncture methods are used, but these come with their own complexities.
- **Event Horizons:** Accurately locating and evolving event horizons is crucial for black hole simulations. Typically, a high resolution near the event horizon is required for long-term stability and accuracy. However, locating event horizons is difficult due to their dynamical nature, especially in binary systems. Searching for event horizons greatly increases computational complexity and cost.

7. Gravitational Waves

- **Extraction of Gravitational Waves:** Extracting and interpreting gravitational wave data from numerical simulations requires precise techniques and is sensitive to numerical errors. Overcoming these difficulties is crucial for developing accurate templates for use in gravitational wave detection, such as LIGO, VIRGO, and the future LISA mission.

-
- **Wave Propagation:** Accurately modeling the propagation of gravitational waves over large distances is computationally demanding, especially in the presence of matter. Typically, integrations of quantities over discretized surfaces are required to compute waveforms, further reducing precision. Maintaining a high precision for waveform modeling comes at a high computational cost.

8. Initial Conditions and Physical Modeling

- **Realistic Initial Conditions:** Setting up realistic and physically accurate initial conditions for simulations, such as those representing inspiraling neutron stars or black holes, is an artform due to its high difficulty. Entire software toolkits have been developed just to handle the task of developing initial data. Highly accurate initial data is vital for long-term stability and accuracy of simulations, especially for methods that do not employ constraint violation damping.
- **Matter and Fields:** Including additional physical elements such as matter, electromagnetic fields, and other fields introduces further complications and increases computational complexity. These additional elements typically come with their own evolution equations, stability requirements, and initial data. Adding such elements to a simulation vastly increases computational complexity and cost.

9. Computational Resources

- **High Performance Computing:** The need for significant computational resources often requires access to high-performance computing clusters, typically requiring academic standing. Managing and optimizing these resources is a challenge and is typically specialized to each facility. Moreover, each computing cluster runs their own type of processors, schedulers, and login procedures, making optimizing simulations for these architectures arduous.
- **Parallelization and Efficiency:** Efficient development of software to run on modern multi-core and distributed computing systems is essential but technically demanding. To generate efficient code typically requires additional knowledge about modern processor architecture, parallelization libraries, and resource management.

These challenges make numerical relativity a slowly growing area of gravitational research, sometimes taking several decades to satisfactorily overcome them. Nonetheless, advancements in computer science, manufacturing and processing, and theoretical methods has led to groundbreaking discoveries in gravitational research. This makes the task of overcoming these challenges deeply rewarding.

1.5.3 CURRENT PROGRAMS

To date, there are many different numerical relativity toolkits available. Each one tackles different challenges and physical systems.

- *Einstein Toolkit*²³⁵: The Einstein toolkit is a comprehensive collection of software components and tools for simulating and analyzing relativistic astrophysical systems. It has been used to simulate binary black hole systems, neutron star mergers, gravitational wave generation, and other relativistic astrophysical phenomena.
- *Spectral Einstein Code (SpEC)*²³⁶: The SpEC code is a flexible, multi-domain spectral solver for partial differential equations. It has been used to simulate gravitational radiation emission, horizon topologies, and binary evolution.
- *Bi-functional Adaptive Mesh (BAM)*²³⁷: The BAM code is a closed-source code that specializes in binary black hole simulations, neutron star processes, and binary neutron star mergers.
- *Numerical Relativity in Python (NRPy+)*²³⁸: NRPy+ is a python package that aims to accelerate the encoding of differential equations into high performance software. Alongside the "Simple, Efficient Numerical Relativity Code" (SENR), they form a fully-functional, highly parallelized numerical relativity code written in the C programming language. It has been used for binary black hole simulations, as well as gravitational waveform modeling.
- *GRChombo*^{239,240}: GRChombo is a new open-source NR code written in the C++ programming language and uses hybrid MPI and OpenMP parallelism, alongside vector intrinsics. It uses an adaptive mesh refinement library that allows it to dynamically adjust the numerical grid resolution based on pre-defined metrics. It has been used in many areas, including early universe cosmology, astrophysics, and mathematical GR.

-
- *Langage Objet pour la RElativité Numérique (Lorene)*²⁴¹: LORENE is a code written in the C++ programming language that uses spectral methods to solve partial differential equations. It has been used in the simulation of modified gravity theories as well as neutron stars.
 - *Adaptive Mesh Spectral Scheme (AMSS-NCKU)*^{242,243}: AMSS-NCKU uses characteristic evolution to solve the Einstein equations and has been used to study binary black holes, multiple black hole systems, and black holes in modified gravity theories.

There are many more NR libraries that tackle different physical systems, such as cosmological systems and hydrodynamical evolutions, with too many to list. The large number of NR libraries is tantamount to the complexity of numerical solutions of the Einstein equations. There are many challenges one is faced with when pursuing numerical solutions.

The GRChombo library will be further explained in Section 3, where it was used to build a new library dedicated to numerical solutions of generalized Proca theories.

1.6 OUTLINE

The rest of the thesis is divided into three pieces. Section 2 will detail a study of a standard Proca field on a spinning black hole. The solutions found in this study will be applied to extreme mass ratio inspirals to determine the viability of the LISA mission to detect such Proca clouds using the gravitational wave information. Section 3 will go into details about a new numerical relativity library built during this thesis that is dedicated to studying generalized Proca fields on arbitrary fixed backgrounds. Finally, Section 4 applies the new library to the case of a standard Proca field on a Kerr-de Sitter background.

The important thing is not to stop questioning. Curiosity has its own reason for existence. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery each day.

Albert Einstein

2

Proca on Kerr

THE ADVENT OF GRAVITATIONAL WAVE ASTRONOMY has brought forth a plethora of avenues to study a wide range of physical phenomena, including fundamental physics. The coupling of gravity to all fundamental particles allows for the study of fundamental fields in the proximity of a strong gravitational field, such as coalescing compact objects. These studies have already placed stringent limits on the allowed parameter space of some physical theories^{244–249}. Current GW observatories such as LIGO²⁵⁰, Virgo^{251,252}, KAGRA^{253,254}; future missions such as LISA^{255,256}, Einstein Telescope²⁵⁷, Cosmic Explorer^{258,259}, DECIGO^{260,261}; and indirect detectors such as pulsar timing arrays, for example EPTA^{262,263}, NANOGrav^{264,265}, and the PPTA²⁶⁶ are thus invaluable tools to study fundamental physics. They shed light on the accuracy of theoretical models of strong gravity, but they can also potentially ameliorate current big problems in physics such as the nature of dark

matter and dark energy. In the scope of astrophysical and fundamental research, they are extremely powerful tools.

The LISA mission in particular, operating in the millihertz frequency range, offers a unique capability to answer fundamental questions. These range from probing the dynamics of extreme mass ratio inspirals (EMRIs) to studying the expansion of the universe²⁶⁷. The EMRI scenario has particular interest due to the long timescale for which the system remains in the strong gravity regime. Moreover, the wide difference in their respective masses translates to a wide difference in the curvature of spacetime they generate, fig. 2.1.

The substantial gap in masses implies the central super massive black hole (SMBH) can be treated as generating a Kerr geometry on which the secondary compact object acts as a perturbing particle. This approximation drastically simplifies estimation techniques for the gravitational waveform, without requiring an appeal to full numerical relativity²⁶⁹⁻²⁷¹.

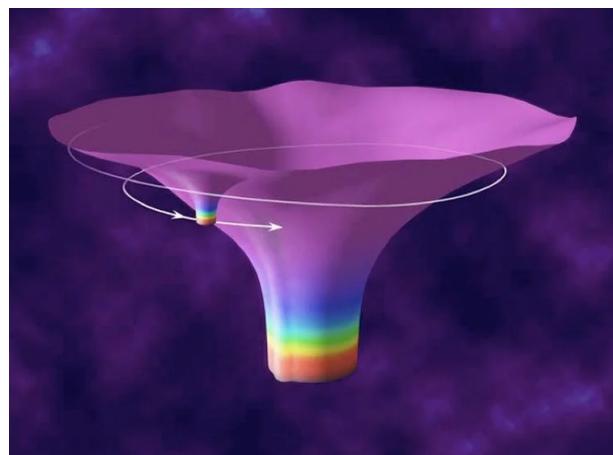


Figure 2.1: ²⁶⁸ An EMRI system represented graphically by the curvatures they generate.

The most common formation mechanism of EMRIs is thought to occur by gravitational capture of a stellar mass compact object onto a highly eccentric orbit²⁷²⁻²⁷⁵. These captures are thought to happen as a result of two body relaxations of the objects where an object is deflected to an orbit around the SMBH with a small pericenter distance. Eventually, emission of gravitational waves (GWs) will reduce the eccentricity to more circular values. Only compact objects such as stellar mass black holes (BHs), neutron stars, white dwarfs or helium cores of giant stars can produce detectable extragalactic EMRI signals. Less compact or lighter objects such as main sequence stars either cannot withstand the tidal forces of the SMBH or are not massive enough to produce sufficiently strong GWs. Other less common formation scenarios exist such as tidal separation of binaries, Bondi-like capture of passing objects, separation of a massive star's core from its envelope, and compact object formation in accretion disks around the SMBH²⁷⁶.

2.1 MOTIVATION

The timescale with which the secondary BH experiences relativistic orbital velocities suggests EMRIs will be a prime target for performing a litany of tests of GR, study environmental effects around SMBHs, and enabling high precision estimation of the physical parameters of the system, such as redshifted masses and black hole spins^{267,277,278}. The high precision measurements of massive black holes (MBHs) and EMRI properties allows a remarkably precise test of many theories, including the predictions of new fundamental fields outside the Standard Model.

Such fundamental fields include the predicted dark matter particle. Dark matter candidates cover a wide range of predicted mass values, depending on the underlying theory. Some models predict the dark matter particle to have a mass as low as 10^{-22} eV²⁷⁹, so-called fuzzy dark matter, and as high as (sub-) solar masses in the form of primordial black holes²⁸⁰. One model for dark matter includes the dark photon, which can be extremely light, $m \gtrsim 10^{-22}$ eV²⁸¹, and behaves as non-relativistic matter.

Near future constraints on the couplings of the dark photon to the Standard Model will come from GW measurements, for example from black hole superradiance^{197,282-284}*. Current constraints in the literature coming from superradiance suggest LISA would be able to constrain the mass of the vector field in the range 1×10^{-16} eV to 6×10^{-16} eV²⁸⁷. The superradiance phenomenon, and more generally GW measurements, will be a powerful tool in probing the potential vector nature of dark matter.

2.1.1 PREVIOUS STUDIES

The study of the instability of matter fields on black hole spacetimes has seen much attention in recent years, owing to the detection prospects using gravitational observatories^{197,198,282,283,288-299}. Most of these studies pertain to either the superradiant process itself and how it affects the black holes spin and mass, which will have an astrophysical statistical effect on the black hole populations, or on the gravitational radiation emitted from the resulting cloud itself. For example, full numerical simulations have been performed to study the resulting gravitational radiation emitted from a

*See ref. 285,286 for the case of a scalar-charged secondary BH in an EMRI and the associated effects on the GW waveform.

superradiantly-excited Proca cloud around a spinning black hole, including evolving through the initial instability phase. Other studies have performed analytic derivations using specialized ansatzes for the solution.

More in line with the details of this study, literature exists for research into the effect of the superradiant clouds, so-called *gravitational atoms*, on the internal motion of a binary system. One study looked at the effect of a scalar and massive vector cloud around one of the black holes and energy transitions induced by a companion black hole³⁰⁰. They showed that this effect could be detectable as a dephasing in the gravitational wave signal.

The study performed here investigates the ability for the LISA mission to detect the modifications to the gravitational wave signal from an EMRI system due to the presence of a superradiant cloud around the supermassive black hole. To that end, numerical solutions of the Proca cloud are computed on a Kerr background, giving the resulting energy distribution of the superradiant cloud. Then the modifications of the gravitational wave signal resulting from the presence of this energy distribution is computed. A LISA forecast is then carried out to determine which parameter values of the Proca field allow LISA to detect such modifications.

2.2 THEORY

2.2.1 CURVED SPACETIME WITH PROCA

The starting point is the specification of the relevant fields, via the action functional

$$S[g, A, \psi] = S_0[g, A] + S_m[g, A, \psi] , \quad (2.1)$$

where g is the metric tensor, A is the Proca field, $S_0[g, A]$ describes the background, and $S_m[g, A, \psi]$ is the action for the matter field ψ . Here, the "skeletonized" approach is adopted. The matter action for the generic matter field ψ is replaced by the action for the point particle. This is a phenomenological reduction of the description of the secondary black hole to that of a "probe" particle following the geodesics set by the background. This is achieved via the replacement

$$S_m[g, A, \psi] \rightarrow S_p[g, A, \{x\}] , \quad (2.2)$$

where $\{x\}$ are the co-ordinates of the secondary black hole.

Specifying the background to be that of a Proca field minimally coupled to a supermassive Kerr black hole, the background action becomes

$$S_0[g, A] = \int d^4x \sqrt{-g} \left[\frac{1}{\kappa} \mathfrak{R} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} \mu^2 A_\mu A^\mu \right], \quad (2.3)$$

where $\kappa = \frac{16\pi G}{c^4}$. Due to the lack of an obvious separability of the point particle current in the chosen ansatz, we consider no coupling between the secondary BH and the Proca field*

$$S_p[g, A, \psi] \equiv - \int m_p d\tau + q \int A_\mu J^\mu \quad (2.4)$$

$$\rightarrow -m_p \int \sqrt{-g_{\mu\nu} \frac{dx^\mu}{dt\tau} \frac{dx^\mu}{d\tau}} d\tau. \quad (2.5)$$

The equations of motion (EOM) associated to the action functional are

$$G^{\rho\sigma} = 8\pi \left(\frac{-1}{4} F_{\mu\nu} F^{\mu\nu} g^{\rho\sigma} + F^{\rho\nu} F_\nu^\sigma - \frac{1}{2} \mu^2 g^{\rho\sigma} A_\mu A^\mu + \mu^2 A^\rho A^\sigma \right) + 8\pi \mathfrak{T}_p^{\rho\sigma} \quad (2.6)$$

$$0 = \nabla_\rho F^{\rho\sigma} - \mu^2 A^\sigma, \quad (2.7)$$

where $\mathfrak{T}_p^{\rho\sigma}$ is the energy-momentum tensor of the point particle. Note that, due to the source terms on the RHS of the Einstein equations, the Proca equations cannot be written in a Klein-Gordon-type form. Instead, using the Lorentz constraint from the conserved current, one finds

$$\nabla^2 A^\nu - R_\sigma^\nu A^\sigma - \mu^2 A^\nu = 0. \quad (2.8)$$

However, this is a generic statement for the full non-linear system. If we work in a perturbative regime in which the Proca amplitude is small and the mass ratio between the secondary and the SMBH is large, then the right-hand side of the Einstein equations vanish, so we can write

$$\nabla^2 A^\nu - \mu^2 A^\nu = 0. \quad (2.9)$$

*More generally, we don't consider any direct coupling between the Proca field and the Standard Model fields.

So, the Proca equations become, in the linearized regime,

$$\nabla^2 A^\nu - \mu^2 A^\nu = 0 \quad (2.10)$$

$$\nabla_\sigma A^\sigma = 0 . \quad (2.11)$$

DECOMPOSITION IN THE FKKS ANSATZ

Analytic solutions of the Proca EOM in the Kerr background have historically been extremely difficult to achieve. This is due to the fact that there is no obvious separable ansatz for the Proca field. The usual approach is to take perturbative approximations. However, a recent analysis has shown that there is such an ansatz that allows the equations to separate, the so-called Frolov-Krtouš-Kubizňák-Santos (FKKS) ansatz³⁰¹. Using this ansatz, we can progress in the semi-analytic computation of the Proca spectra.

The first step in solving the Proca equations is to decompose the Proca EOM in the FKKS ansatz. We define the ansatz for the Proca field as

$$A^\mu = B^{\mu\nu} \nabla_\nu Z , \quad (2.12)$$

where $B^{\mu\nu}$ is implicitly defined through the complex-valued algebraic equation

$$B^{\mu\nu} (g_{\nu\gamma} + \frac{i}{\lambda} h_{\nu\gamma}) = \delta_\gamma^\mu , \quad (2.13)$$

and $h_{\mu\nu}$ is the so-called principle tensor of the Kerr spacetime. In Boyer-Lindquist co-ordinates, it is defined as

$$h_{\mu\nu} = \begin{bmatrix} 0 & r & a^2 \cos(\theta) \sin(\theta) & 0 \\ -r & 0 & 0 & ar \sin(\theta)^2 \\ -a^2 \cos(\theta) \sin(\theta) & 0 & 0 & a \cos(\theta)(a^2 + r^2) \sin(\theta) \\ 0 & -ar \sin(\theta)^2 & -a \cos(\theta) \sin(\theta)(a^2 + r^2) & 0 \end{bmatrix} . \quad (2.14)$$

We then perform a separation of variables in the FKKS ansatz via

$$Z = R(r)S(\theta)e^{-i\omega t}e^{im\phi} . \quad (2.15)$$

This is a multiplicative separation into two arbitrary single-coordinate-dependent functions and two eigenfunctions of the spacetime Killing vectors *

$$\mathfrak{L}_T Z = -i\omega Z \quad (2.16)$$

$$\mathfrak{L}_\Phi Z = imZ , \quad (2.17)$$

where \mathfrak{L} is the Lie derivative, and T, Φ are the temporal and azimuthal Killing vectors of the Kerr spacetime, respectively. After inserting this separated form into the Proca EOM, one finds the following coupled second-order system of differential equations

$$\frac{d}{dr} \left(\frac{\Delta}{q_r} \frac{dR}{dr} \right) + \left(\frac{K_r^2}{q_r \Delta} + \frac{2 - q_r \sigma}{q_r^2 \nu} - \frac{\mu^2}{\nu^2} \right) R = 0 \quad (2.18)$$

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\frac{\sin \theta}{q_\theta} \frac{dS}{d\theta} \right) - \left(\frac{K_\theta^2}{q_\theta \sin^2 \theta} + \frac{2 - q_\theta \sigma}{q_\theta^2 \nu} - \frac{\mu^2}{\nu^2} \right) S = 0 , \quad (2.19)$$

where

$$K_r = am - (a^2 + r^2)\omega \quad q_r = 1 + \nu^2 r^2 \quad (2.20)$$

$$\sigma = a\nu^2(\mathbf{m} - a\omega) + \omega \quad K_\theta = \mathbf{m} - a\omega \sin^2 \theta \quad (2.21)$$

$$q_\theta = 1 - \nu^2 a^2 \cos^2 \theta \quad \Delta = r^2 + a^2 - 2Mr . \quad (2.22)$$

Henceforth, we focus only on a single mode specified by the tuple $(\nu, \omega, \mathbf{m})$. A generic solution to the Proca EOM will be a linear combination of the single mode solutions. See appendix A.1 for a more in-depth discussion of the method to solving the coupled eigenvalue problem 2.18-2.19. Some additional quantities that are important for the later analysis are the total energy and the normalization of the Proca field. Eqs. 2.18-2.19 determine the Proca field only up to an overall nor-

*We write the mode number as \mathbf{m} in order to distinguish it from the mass of the secondary black hole, which is denoted as m .

malization constant. This constant must be determined as part of a full description of the state of the cloud. We choose to normalize the field by the requirement that the total energy of the Proca field matches the reduction in energy of the black hole (energy conservation), utilizing the clear separation of timescales between the superradiant instability and the gravitational radiation from the cloud. Were this separation of timescales not present, a non-negligible amount of energy from the Proca cloud could be radiated away in gravitational energy during the instability phase, preventing this method from determining the Proca normalization. Fortunately, the cloud generally builds up more rapidly than it depletes via gravitational radiation, hence we can safely neglect the depletion during the instability³⁰⁰. The total energy of the cloud at a particular instance is defined by

$$E_c = - \int \mathfrak{T}_t^t \sqrt{-g} dr d\theta d\phi , \quad (2.23)$$

where g is the metric determinant and \mathfrak{T} is the stress-energy tensor of the Proca field, eq. 2.6. Normalization of the Proca field then follows from the requirement $E_c = M_{0,bh} - M_{f,bh}$, where $M_{0,bh}$ and $M_{f,bh}$ are the masses of the black hole before and immediately after the superradiant instability, respectively.

SYSTEMATICS OF SOLVING THE RADIAL AND ANGULAR EQUATIONS

We can now proceed to solve the angular and radial equations, eqs. 2.18-2.19. First, define the parameters and eigenvalues of the problem as $(\mathbf{m}, S, n, a, \mu, M) = \mathcal{P}$ and $(\nu, \omega) = \mathcal{E}$, respectively. The system of equations can then be represented schematically as

$$\begin{aligned} L[r; \mathcal{P}, \mathcal{E}]R(r) &= 0 \\ O[\theta; \mathcal{P}, \mathcal{E}]S(\theta) &= 0 , \end{aligned} \quad (2.24)$$

where L and O are linear operators defined in eq. 2.18 and 2.19 and which are coupled only through the set of eigenvalues \mathcal{E} . In the non-relativistic limit $\mu M \ll 1$, the real and imaginary parts of the frequency read^{294,302,303}

$$\frac{\omega_R}{\mu} \approx 1 - \frac{\mu^2 M^2}{2(|\mathbf{m}| + n + S + 1)^2} + O((\mu M)^4) \quad (2.25)$$

$$M\omega_I \approx 2r_+ C_{lmSn} g_{jm}(a, \mu M, \omega) (\mathbf{m}\Omega_H - \omega_R) (M\mu)^{4|\mathbf{m}|+5+2S}. \quad (2.26)$$

These, together with the non-relativistic limits for the eigenvalue ν in appendix A.1, provide good starting guesses for iteratively solving the system eq. 2.24. The algorithm we employ to numerically solve the Proca field in the FKKS ansatz follows similarly to 293 and goes as follows:

- After specifying the initial parameters to be considered, i.e. choosing a particular $p \in \mathcal{P}$, an initial guess for the ω and ν eigenvalues are formed by the non-relativistic limit.
- Solve the determinant of the angular equation matrix and pick the eigenvalue that is nearest to either the non-relativistic limit or the previous result for a different mass.
- Solve the Proca radial equation using the Frobenius method and find the initial conditions from evaluating the Frobenius solution at a starting radius, close to the outer horizon.
- Numerically solve the radial equation with the previously obtained ω and ν eigenvalues and boundary conditions.
- Find the logarithmic minimum of the radial equation at the outer boundary of the radial integration over ω -space. Minimization is carried out in ω -space, recalculating ν for each ω -value, using a native Nelder-Mead algorithm in the software system Mathematica³⁰⁴.
- With the found value of ν and ω , the angular matrix can be solved for the expansion of the angular function in terms of the spherical harmonics, completing the solution of the EOM.

This process can be repeated for varying choices of the Proca mass parameter, overtone number, and mode number. Initial guesses for ω and ν switch from using the non-relativistic limit to using a 4th-order polynomial fit to previous results in $\mu - \omega$ space. These fits perform much better than the non-relativistic limit for higher mass parameters, typically $\mu \gtrsim 0.3$.

The flow of the algorithm proceeds graphically as in fig. 2.2.

COMPUTING THE ASYMPTOTIC FLUX

We used the package SUPERRAD²⁸⁷ for the calculation of the asymptotic fluxes from the Proca cloud. It uses a combination of analytic and numerical results to compute the asymptotic energy flux from a Proca cloud, assuming all the energy of the cloud resides in a single mode. The asymptotic angular momentum flux can then be computed from the Teukolsky formalism³⁰⁵ as

$$\left\langle \frac{dJ}{dt} \right\rangle = \frac{\mathbf{m}}{\omega} \left\langle \frac{dE}{dt} \right\rangle. \quad (2.27)$$

2.2.2 SUPERRADIANCE

Black hole superradiance is a dissipative phenomenon which involves the unstable growth of field amplitudes due to the collection of negative energy states by the ergoregion*. This superradiant instability of matter fields around spinning black holes can lead to, under certain conditions, a quasibound state. In fact, a quite general argument for the existence of superradiance can be shown to follow from the black hole area theorem, which states

$$\delta M = \frac{T_H}{4} \delta A + \Omega_H \delta J \quad (2.28)$$

for an uncharged black hole, where T_H is the Hawking temperature, A_H is the area of the horizon, $\Omega_H = \frac{a}{r_+^2 + a^2}$, a is the spin of the black hole, and r_+ is the radius of the outer event horizon. For a matter wave of frequency ω and azimuthal number \mathbf{m} , the ratio of angular momentum to energy is

$$\frac{L}{E} = \frac{\mathbf{m}}{\omega}. \quad (2.29)$$

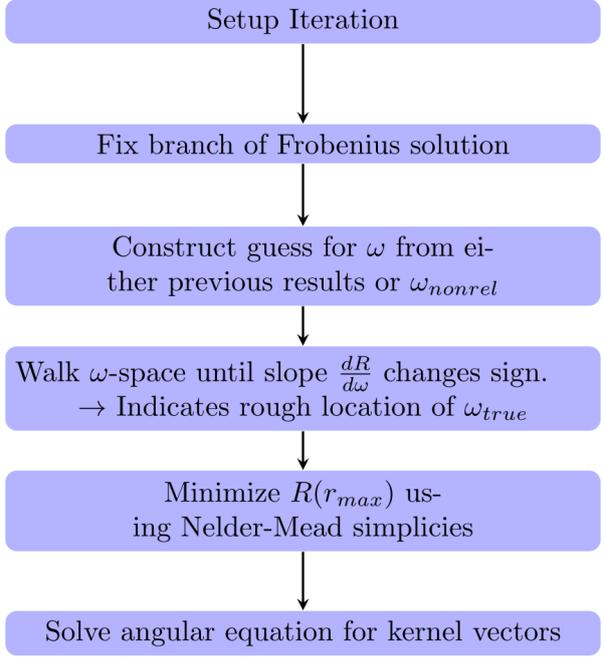


Figure 2.2: Graphical representation of routine to solve the Proca equations in the Kerr background.

*For a review of superradiance, see 200.

Hence, an interaction of the matter wave with the black hole causes the latter to change its angular momentum by

$$\frac{\delta J}{\delta M} = \frac{\mathfrak{m}}{\omega}. \quad (2.30)$$

The area theorem then tells us

$$\delta M = \frac{\omega T_H}{4} \frac{\delta A_H}{\omega - \mathfrak{m}\Omega_H}. \quad (2.31)$$

The second law of black hole thermodynamics, $\delta A_H > 0$, implies waves impinging on the event horizon with frequency

$$\omega < \mathfrak{m}\Omega_H \quad (2.32)$$

causes the black hole to lose mass and hence energy is extracted by the wave, increasing its own energy. This wave can become trapped by the potential well of the black hole, causing the wave to again impinge on the black hole.

This is the mechanism of superradiance. A small amplitude wave initially impinging from past infinity will be continuously excited in a runaway process until the black hole loses enough angular momentum and mass to turn off the superradiant condition, eq. 2.32. This is a purely classical description. A quantum description, in which vacuum states at past and future infinity contain different particles numbers, has also been formulated^{306–308}. The process is reminiscent of the well-known Penrose process, though they are distinct phenomena^{309*}.

2.2.3 MODIFIED GRAVITATIONAL WAVES

After the superradiant instability has turned off, saturating the superradiant threshold eq. 2.32, the system exists in a quasi-stationary state consisting of a black hole surrounded by a quasibound Proca condensate. It is not in an eternal bound state due to gravitational emission from the cloud itself, which manifests as a long-duration depletion of the condensate³⁰⁰. This long-timescale deple-

*It has been shown that the Penrose process requires the existence of an ergoregion, while superradiance requires an event horizon^{310,311}. Additionally, a stationary axi-symmetric black hole necessarily has an ergoregion. Thus, superradiance is a sufficient condition for the Penrose process, but the converse is not necessarily true.

tion of the cloud provides a secular change in the mass and angular momentum of the background, in addition to any other emissions from the system. Hence, the presence of a Proca environment surrounding the primary black hole in an EMRI system modifies the inspiraling dynamics of the secondary black hole, resulting in a modification to the measured waveform at the detector *. In an EMRI system, the inspiraling dynamics is well approximated by assuming the trajectory follows a sequential evolution of geodesics of the Kerr spacetime. In the Kerr spacetime, geodesics are determined by three constants of motion: the energy E , the projection of the angular momentum along the spin axis L , and the Carter constant C . Assuming an equatorial orbit, the Carter constant vanishes and receives no evolution. Hence, the geodesic motion is determined by only two constants, E and L . Sequential evolution along a series of geodesics corresponds to an adiabatic evolution of the orbital constants. This adiabatic change in the integrals of motion arises due to the asymptotic flux of energy and angular momentum from the system, sourced by either environmental effects or gravitational emission.

In particular, for the case of an EMRI system immersed in a superradiantly-generated Proca cloud, the evolution of the integrals of motion is given by

$$\frac{dE_{geo}}{dt} = - \left(\frac{dE_{GW}}{dt} + \left(\frac{dE_{geo}}{dt} \right)_{Proca} \right) \quad (2.33)$$

$$\frac{dL_{geo}}{dt} = - \left(\frac{dL_{GW}}{dt} + \left(\frac{dL_{geo}}{dt} \right)_{Proca} \right), \quad (2.34)$$

where I_{geo} represents the integral of motion for the geodesic and $\left(\frac{dI_{geo}}{dt} \right)_{Proca}$ represents the change in the orbital constants due to the flux of energy and momentum from the quasibound Proca cloud. Since the secondary black hole is minimally coupled to the energy-momentum of the Proca field, via the Einstein equations, the change in the orbital constants will not be the same as the change in the energy and angular momentum of the Proca cloud. Instead, at the linear level, the presence of the Proca cloud modifies the energy and angular momentum of the background spacetime, which enters as an additional change in the orbital constants. In particular, the change in the

*Here, we neglect higher order effects such as resonant depletion of the Proca cloud (see 302,312 for resonant depletion of a superradiant scalar cloud in an EMRI system). We also neglect dynamical friction effects on the secondary black hole. These effects have been shown to have dramatic effects in the scalar field case and conceivably will also have large effects in the vector case^{302,313}, though a study of this nature has yet to be performed. However, for the preliminary and simplified analysis considered here, we relegate these effects to studies that will be performed in the future.

integrals of motion, due to the presence of an uncoupled Proca cloud, arises due to the change in the energy and angular momentum of the Kerr background:

$$dE_{geo} = \frac{\partial E_{geo}}{\partial L_{Kerr}} dL_{Kerr} + \frac{\partial E_{geo}}{\partial E_{Kerr}} dE_{Kerr} \quad (2.35)$$

$$dL_{geo} = \frac{\partial L_{geo}}{\partial L_{Kerr}} dL_{Kerr} + \frac{\partial L_{geo}}{\partial E_{Kerr}} dE_{Kerr} \quad (2.36)$$

such that

$$\left(\frac{dE_{geo}}{dt} \right)_{Proca} = \frac{dE_{Proca}}{dt} \Gamma(r) \left[\frac{dE_{geo}}{dE_{Kerr}} + \frac{\mathfrak{m}}{\omega} \frac{dE_{geo}}{dL_{Kerr}} \right] \quad (2.37)$$

$$\left(\frac{dL_{geo}}{dt} \right)_{Proca} = \frac{dE_{Proca}}{dt} \Gamma(r) \left[\frac{dL_{geo}}{dE_{Kerr}} + \frac{\mathfrak{m}}{\omega} \frac{dL_{geo}}{dL_{Kerr}} \right], \quad (2.38)$$

where we introduced a radially dependent prefactor that accounts for the fraction of the Proca cloud within the orbital radius, and E_{Kerr} and L_{Kerr} are the total mass and angular momentum of the Kerr spacetime, respectively. E_{Kerr} and L_{Kerr} are calculated prior to the superradiant instability and hence represent the total mass and angular momentum of the black hole-cloud system after the instability has turned off. It follows from the Teukolsky equation and our choice of normalization of the Proca field that $\frac{dE_{Proca}}{dt} \propto E_{Proca}^2$, and hence $\Gamma(r) = \left(\frac{E(r < r_{orbit})}{E_{total}} \right)^2$ (see A.2). This prefactor accounts for the portion of the radiating Proca cloud that modifies the orbital trajectory. $\Gamma(r)$ asymptotes to unity at asymptotic infinity, meaning all of the Proca cloud is within the orbital radius and contributes to the trajectory modification. At the other extreme, near the horizon, $\Gamma(r)$ approaches zero since all of the cloud is external to the orbital radius. This prefactor hence represents the fraction of the Proca cloud the inspiraling black hole "sees". At infinity, the mass the secondary black hole "sees" is the total mass of the black hole-cloud system; At the horizon, it's the mass of the central black hole which is "seen" by the secondary black hole and hence none of the radiating Proca cloud modifies the trajectory at this point.

This is an approximation in several respects. First, the angular structure of the cloud is integrated out to produce a purely radial function. Secondly, the energy integral eq. 2.23 is calculated using the stress-energy tensor from the perturbative calculation of the Proca field and not from the full Einstein-Proca system. Thirdly, an additional averaging of the radial distance over an orbital period

is performed when calculating $\Gamma(r)$, due to limitations of the waveform generator.

For the gravitational terms in the flux functions, 5PN accurate analytic expressions for the energy and momentum fluxes (hence, semi-latus rectum and eccentricity evolution. See below.) are employed.

For our purposes, it's more convenient to express adiabatic evolution of the integrals of motion in terms of the eccentricity and semi-latus rectum, from which the integrals of motion can be expressed. The asymptotic fluxes can be transformed into rates of change of the orbital parameters by inverting $\dot{I}_{geo} = \frac{dI_{geo}}{dp}\dot{p} + \frac{dI_{geo}}{de}\dot{e}$. This then gives us the rate of change of the geometry of the trajectory which the secondary black hole follows.

The full trajectory is calculated by integrating the flux equations, after choosing suitable initial conditions, using an 8th-order explicit Runge-Kutta integrator. The trajectory is integrated to within 0.2 gravitational radii of the separatrix, calculated using the previous iteration loop of the integration. Initial conditions for the integration, namely $(p_0, e_0, \Phi_{\theta,0}, \Phi_{\phi,0}, \Phi_{r,0})$, are chosen such that coalescence occurs approximately after five years. This gives the greatest possible chance a Proca cloud will be detected during the mission lifetime of the LISA observatory.

The trajectory, once computed, is then fed into a waveform model. The model currently employed is the Fast EMRI Waveforms (FEW) Augmented Analytic Kludge (AAK) model^{271,314}. The AAK model is built using Keplerian ellipses for the orbital trajectory, and evolves the inspiral, periastron precession, and Lense-Thirring precession using PN fluxes. The difference to the original Analytic Kludge model is that the orbital frequencies and two precession rates are enforced to be the original Kerr values, which is achieved by solving an algebraic expression for some unphysical values of the mass, spins, and semi-latus rectum. This defines a map $(M, a, p) \rightarrow (\tilde{M}, \tilde{a}, \tilde{p})$ which maps the frequencies of the Keplerian orbit onto the frequencies for the Kerr geodesic motion. This greatly improves the accuracy of the original AK model and agrees remarkably well with Teukolsky-based waveforms. The FEW version, the version employed in this study, removes this mapping and instead directly calculates the fundamental frequencies and converts them into the basis for the AAK model, eqs. 2.39-2.41.

$$\dot{\Phi} = \Omega_r \quad (2.39)$$

$$\dot{\gamma} = \Omega_\theta - \Omega_r \quad (2.40)$$

$$\dot{\alpha} = \Omega_\phi - \Omega_\theta, \quad (2.41)$$

where $\dot{\Phi}$ is the variation of the quasi-Keplerian mean anomaly, $\dot{\alpha}$ is the Lense-Thirring precession, and $\dot{\gamma} + \dot{\alpha}$ is the periapsis precession. These phase evolutions are then fed into the Peters-Matthew formula for the gravitational strain amplitudes³¹⁵ *.

The states of the cloud in this analysis are restricted to the $m = 1$ mode and $n = 0$ overtone. This is for several reasons. First, the asymptotic flux values from the numerical solver are numerically unstable for larger mode and overtone values. Secondly, gravitational emission from higher modes is α -suppressed^{287,292}, with higher modes being suppressed by powers of α^4 . Hence, the secular variation in the cloud largely comes from the $m = 1$ mode. Moreover, the total mass contained in the higher modes is less than that in the $m = 1$ mode. Reducing the analysis to the single choice of these parameters is thus reasonable within the approximations already employed.

2.3 RESULTS

2.3.1 PROCA CLOUDS

Following the procedure laid out in Section 2.2.1, the Proca field equations were solved for mode numbers $\mathbf{m} = \{1, 2, 3, 4\}$, overtone numbers $n = \{0, 1, 2, 3, 4\}$, SMBH dimensionless spin $\chi \in [0.6, 0.9]$ and Proca spin $S = -1$ [†]. For our purposes, we restrict to $S = -1$ as this is the most unstable²⁹⁴. Examples of our generated data are shown in fig. 2.3,2.4,2.5. Figure 2.3 shows the evolution of the Proca field frequency as a function of the gravitational coupling for the $\mathbf{m} = 1$ mode and SMBH dimensionless spin $\chi = 0.9$, for various overtone numbers. The imaginary part of the frequency gives the instability rate of the cloud, while the real part yields the oscillation frequency. As can be seen, the $n = 0$ overtone number is the most unstable. The maximum instability occurs

*The custom code used to generate these trajectories and waveforms is available at <https://github.com/Shawn-F/GWGenerator.git>.

[†]The code used to generate this dataset is available at <https://github.com/Shawn-F/KerrDressedWithProca.git>.

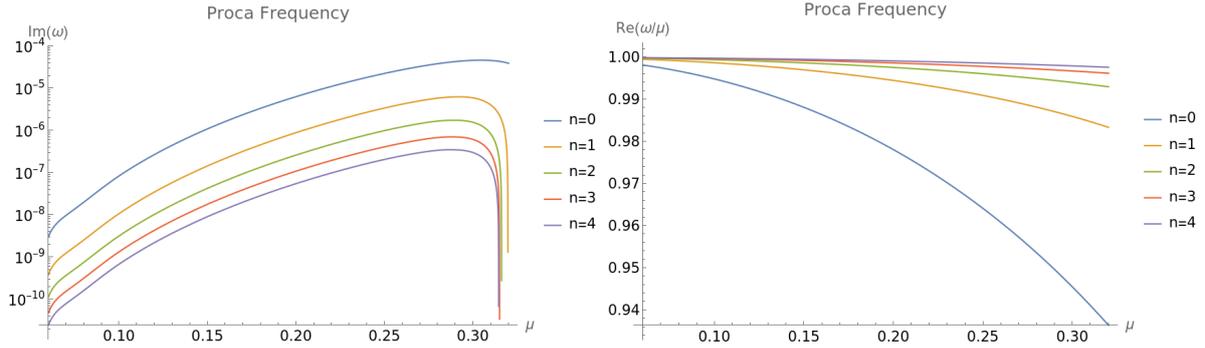


Figure 2.3: Superradiant Proca data for $\mathbf{m} = 1$ mode, dimensionless spin $\chi = 0.9$, and Proca spin $S = -1$. The left plot displays the imaginary part of the Proca field frequency, which directly yields the instability rate of the cloud. The right plot displays the real part of the frequency, which yields the oscillatory part of the field. As expected, the $\mathbf{m} = 1, n = 0$ mode is the most unstable. Here we take $G = c = \hbar = 1$, so that $\alpha = \mu M \frac{G}{c\hbar} \rightarrow \mu M$.

at, for $(\mathbf{m}, n, \chi) = (1, 0, 0.9)$, $\alpha = 0.304$ with an instability rate of $\tau = 2.1 \times 10^4 \frac{GM}{c^3} = 0.105$ s. Compared to the maximum instability of the corresponding scalar superradiant cloud, this is ~ 2500 times faster²⁹⁴.

Fig. 2.4, 2.5 show example radial functions for the $\mathbf{m} = 1$ mode and dimensionless spin $\chi = 0.9$. The overtone structure of the Proca field is clearly displayed. The number of roots of the radial function is given by the overtone number, which also specifies the number of maxima and minima. The compactness of the cloud is also apparent, being directly given by the gravitational coupling, as expected. A higher gravitational coupling translates to a more compact Proca cloud. Lower values of the gravitational coupling yield a Proca cloud that can span thousands of gravitational radii, as expected from the rough scaling of the radial function as $\sim \frac{1}{\alpha}$. Higher values of the gravitational coupling yield Proca clouds that span tens of gravitational radii. Hence, higher values of the gravitational coupling are expected to have the greatest effect on an EMRI system*.

2.3.2 IDENTIFYING MODIFIED GWs

To assess the detectability and measurability of a superradiantly-generated Proca cloud around an EMRI system from the emitted GWs, we leverage a simple figure of merit. We perform a simple analysis based on generated SNR and traditionally used faithfulness criteria to have an idea about the distinguishable mass region of the Proca particles. This merit relies on comparing two waveforms, and so can be easily and quickly computed.

*For a more in-depth analysis of superradiant Proca fields on Kerr backgrounds, see 283,293,294,297,301,316.

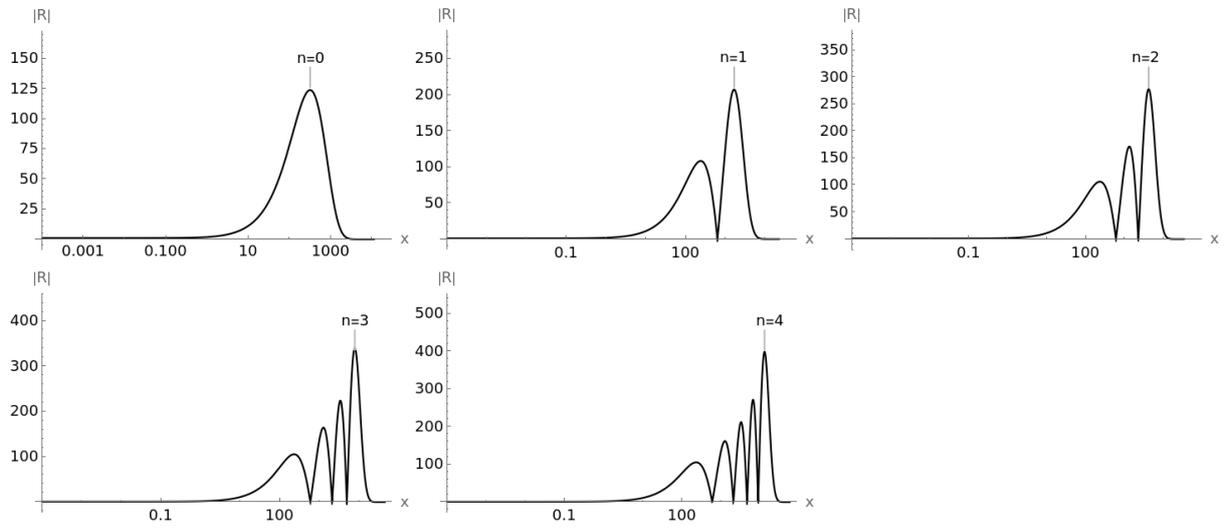


Figure 2.4: Example radial functions for $m = 1$ mode with dimensionless spin $\chi = 0.9$ and gravitational coupling $\alpha = \frac{11}{100}$ for various overtone numbers.

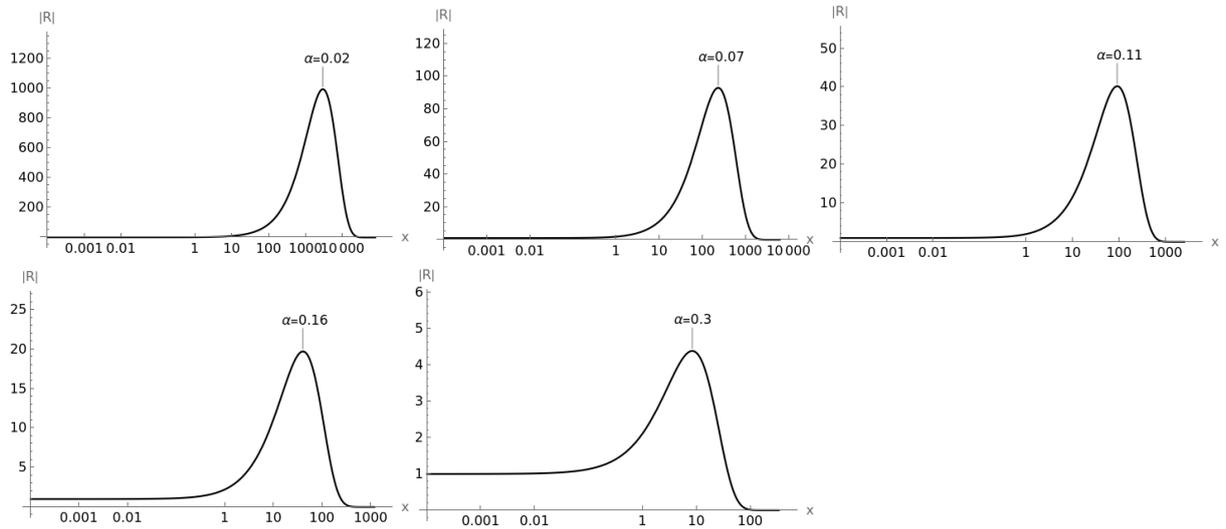


Figure 2.5: Example radial functions for $m = 1$ mode with dimensionless spin $\chi = 0.9$ and overtone number $n = 0$ for various gravitational couplings. As expected, higher gravitational couplings translate to a denser cloud, with most of the mass being concentrated closer to the horizon.

AN ESTIMATE BASED ON SNR

First, we define a noise-weighted inner product between two waveforms as

$$\langle h_1, h_2 \rangle = 4 \times \Re \sum_{\alpha=I,II} \int_{f_{min}}^{f_{max}} \frac{\tilde{h}_{1,\alpha}(f) \tilde{h}_{2,\alpha}^*(f)}{S_n(f)} df, \quad (2.42)$$

where $\tilde{h}_{1/2,\alpha}$ are the Fourier-transforms of the detector response signals, $\tilde{h}_{2,\alpha}^*$ the latter's complex conjugate, and $S_n(f)$ the one-sided noise power spectral density (PSD) of LISA^{270,317}. The LISA PSD model receives contributions from three noise sources. The analytic expression for the PSD is

$$S_n(f) = \frac{10}{3L^2} \left(P_{OMS}(f) + \frac{4P_{acc}(f)}{(2\pi f)^4} \right) \left(1 + \frac{6}{10} \left(\frac{f}{f_*} \right)^2 \right) + S_c(f), \quad (2.43)$$

where P_{OMS} is the single-link optical metrology noise, P_{acc} is the single test mass acceleration noise, and S_c is the galactic confusion noise, accounting for unresolved galactic sources that manifest in the noise. The analytic expressions for the three noise contributions are

$$P_{OMS} = (1.5 \times 10^{-11} \text{ m})^2 \left(1 + \left(\frac{2 \text{ mHz}}{f} \right)^4 \right) \text{ Hz}^{-1} \quad (2.44)$$

$$P_{acc} = \left(3 \times 10^{-15} \frac{\text{m}}{\text{s}^2} \right) \left(1 + \left(\frac{0.4 \text{ mHz}}{f} \right)^2 \right) \left(1 + \left(\frac{f}{8 \text{ mHz}} \right)^4 \right) \text{ Hz}^{-1} \quad (2.45)$$

$$S_c = A f^{-\frac{7}{3}} e^{-f^\alpha + \beta f \sin(\kappa f)} (1 + \tanh(\gamma(f_k - f))) \text{ Hz}^{-1}, \quad (2.46)$$

where the parameters of the galactic confusion noise are fit to simulations for a four-year data run. These fit values are $(A, \alpha, \beta, \kappa, \gamma, f_k) = (1.8 \times 10^{-44}, 0.138, -221, 521, 1680, 0.00113)$.

Using the noise-weighted inner product, the SNR of a particular signal is defined as $\text{SNR}^2 = \langle h|h \rangle$. Further, we also define the faithfulness between two signals as^{318,319}

$$F \equiv \max_{t_c, \phi_c} \frac{\langle h_1|h_2 \rangle}{\sqrt{\langle h_1|h_1 \rangle \langle h_2|h_2 \rangle}}, \quad (2.47)$$

which is maximized over time and phase offsets of the two signals and takes values between -1 and 1, where the latter indicates perfect agreement between the waveforms. The maximization translates

to maximizing over the variables T in

$$\langle h_{1,T}|h_2 \rangle = \Re \sum_{\alpha=I,II} \int \frac{\tilde{h}_{1,\alpha}(f)\tilde{h}_{2,\alpha}^*(f)}{S_{n,2}(f)} e^{-2\pi fT} df, \quad (2.48)$$

where $h_{1,T}$ is the time-offset version of the original waveform h_1 by time T , $S_{n,2}$ is now the two-sided noise power spectral density, and we have extended the integration domain to the entire reals using the fact that the noise PSD forces the integrand to vanishing values outside the range $[-f_{max}, -f_{min}] \cup [f_{min}, f_{max}]$ and that it's an even function of f . Using the convolution theorem, this translates into the convolution

$$\langle h_{1,T}|h_2 \rangle = \Re \sum_{\alpha=I,II} (H_{1,\alpha} * H_{2,\alpha}), \quad (2.49)$$

where $H_{1,\alpha} = \mathfrak{F}^{-1}\left(\frac{\tilde{h}_{1,\alpha}(f)}{S_{n,2}(f)}\right)$, $H_{2,\alpha} = \mathfrak{F}^{-1}(\tilde{h}_{2,\alpha}^*(f))$, and \mathfrak{F} denotes the Fourier transform. The faithfulness then becomes *

$$F \equiv \max \left\{ \left(\frac{\Re \sum_{\alpha=I,II} (H_{1,\alpha} * H_{2,\alpha})}{\sqrt{\langle h_1|h_1 \rangle \langle h_2|h_2 \rangle}} \right) (T, \phi_c) \right\}. \quad (2.50)$$

The detectability requirement places a threshold on the faithfulness statistic. Under Gaussian likelihoods for the parameters, this threshold arises from the requirement that a systematic mismodeling error, i.e. the error between the true waveform and the model waveform, should be smaller than the statistical measurement error. If the mismodeling error were larger than the measurement error, then the signals would be measurably different in the LISA data. Thus, if the faithfulness between the bare EMRI and the dressed EMRI waveforms are below this critical threshold, LISA should likely be able to distinguish between the two EMRI systems. This threshold assumes a Gaussian distribution for the model parameters around the true values. This prerequisite does not hold for the Proca mass, since the distribution is asymmetric and the waveforms are being compared against the vacuum case (c.f. fig. 2.6). However, it provides a rough estimate for the threshold of distinguish-

*Maximization over the coalescence phase ϕ_c is achieved by iteration over phase offsets of one of the signals. See 320 for further discussion of the maximization procedure.

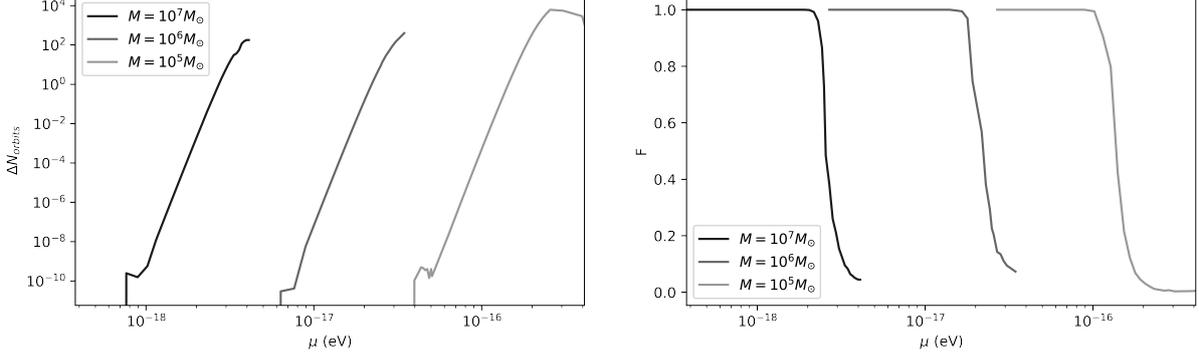


Figure 2.6: The difference in number of orbits and faithfulness as a function of Proca mass, respectively. The total spacetime dimensionless spin is $\chi=0.9$ and the initial eccentricity was chosen to be 0.2. The difference in number of orbits is the absolute difference between the number of orbits completed by the dressed and undressed waveform at separatrix. This data suggests LISA should be able to distinguish GR-in-vacua waveforms and Proca-modified waveforms, for the given initial data, for Proca masses above $\sim 2 \times 10^{-18}$ eV. The upper limit on the mass is enforced by the superradiant threshold cutoff. Above this threshold, no bound state exists. For the $M = 10^5 M_\odot$ data, this upper limit is $\mu_{max} = 4.47 \times 10^{-16}$ eV.

bility. The expression for this threshold is

$$F_c = 1 - \frac{D - 1}{2 \times \text{SNR}^2}, \quad (2.51)$$

where D is the size of the parameter space³²¹. The inclusion of a Proca mass increases the number of parameters by one, hence the total parameter space is specified by

$$(M, m, \mu, a, p_0, e_0, x_0, d_L, \theta_s, \phi_s, \theta_K, \phi_K, \Phi_{\theta,0}, \Phi_{\phi,0}, \Phi_{r,0}), \quad (2.52)$$

where M is the mass of the SMBH (the primary), m is the mass of the smaller black hole (the secondary), μ is the mass of the Proca field, a is the spin value of the SMBH, p_0 , e_0 , and x_0 are the initial semi-latus rectum, eccentricity and inclination, respectively, d_L is the luminosity distance to the system, θ_S and ϕ_S are the barycentric sky location of the system, θ_K and ϕ_K describe the orientation of the EMRI angular momentum vector in the barycentric co-ordinate system, and $\Phi_{\theta/\phi/r,0}$ are the initial phases for the polar, azimuthal, and radial motion, respectively. Orientation of the spin vector of the SMBH is currently ignored in waveform generation due to limitations of the waveform generator package employed during this study. For signals with SNR on the order of 20, the faithfulness threshold for detectability is ~ 0.98 . For any pair of signals that produces a faithfulness below this threshold, they will likely be distinguishable with LISA.

The second measure of detectability we utilize is the number of orbits completed by the binary,

as this is easily inferred from the waveform directly measured by LISA. The number of orbits completed by the binary at separatrix is directly related to the orbital frequency by

$$N_{orbits} = \frac{1}{2\pi} \int_0^{T_{sep}} \Omega_\phi dt, \quad (2.53)$$

where Ω_ϕ is the azimuthal orbital frequency, which is related to the phase of the trajectory Φ_ϕ by $\frac{d\Phi_\phi}{dt} = \Omega_\phi$. The resolvable deviations in the phase of the trajectory of an EMRI system by LISA can be estimated by a Fisher-matrix analysis^{321,322}. The result is that LISA should be able to distinguish a phase shift of $\Delta\Phi_\phi \sim 0.05$ radians, or in terms of the GW phase $\Delta\Psi \sim 0.1$ radians. This roughly translates to a resolution in the number of completed orbits at the time the trajectory reaches the separatrix as $\Delta N_{orbits} \sim 0.001$.

Figure 2.6 shows three example EMRI systems with SMBH masses $M = (10^5, 10^6, 10^7)M_\odot$, SMBH dimensionless spin $\chi = 0.9$, and initial eccentricity of the orbit $e = 0.2$. The initial semi-latus rectum is chosen so that the lifetime of the system approximately equals the LISA mission lifetime. As expected, for higher Proca masses, the dressed waveform increasingly deviates from the GR-in-vacua (undressed) waveform. This is due to the increased coupling between the Proca cloud and the background spacetime. Since the size of the Proca cloud roughly scales as $\frac{1}{\mu}$ (c.f. eq. A.22), higher mass translates to a more compact Proca cloud. Thus, as the secondary black hole inspirals, Proca clouds with greater Proca masses have a larger effect on the trajectory due to the compactness (c.f. eq. 2.37, in particular $\Gamma(r)$). In addition, the energy and angular momentum fluxes from the Proca cloud are monotonically increasing up until just before the superradiant condition fails. Hence, higher compactness of the cloud together with larger values of the asymptotic fluxes explains the greater deviation from the GR-in-vacua scenario.

It was found that the initial value of the eccentricity has little effect on the faithfulness, i.e. the difference between the Proca waveform and the GR-in-vacua waveform doesn't change much with eccentricity. This is likely due to the averaging procedure over eccentricity of the effect the Proca flux has on the secondary's trajectory. In other words, the function $\Gamma(r)$ is only a function of the radial distance and not the eccentricity. Though factors such as $\left[\frac{dE_{orbit}}{dE_{Kerr}} + \frac{m}{\omega} \frac{dE_{orbit}}{dL_{Kerr}} \right]$ are expected to change with eccentricity, this effect is evidently small.

The spin of the SMBH, on the other hand, plays a larger role. While the spin doesn't affect the

'knee' of the graph in fig. 2.6, it does change the minimum value of the faithfulness. In other words, lower spins of black holes are less able to constrain lower Proca masses.

It's also important to determine the astrophysical relevance of this type of modification to the background. This is achieved by comparing the timescale for gravitational emission from the Proca cloud to the timescale for inspiral of the secondary object. The latter is fixed to be ~ 5 years, the approximate LISA mission lifetime, and the former depends on the gravitational coupling α . In other words, the astrophysical relevance for the study performed in this paper depends on the gravitational coupling. The 'knee' of the faithfulness statistic as a function of α in fig. 2.6 occurs approximately at $\alpha \sim 0.07$. The corresponding timescale for gravitational emission from the Proca condensate ranges from $10^3 - 10^5$ years (cf. appendix A.2 and fig. A.1). This means the modification to the gravitational wave signal due to the time-dependent background as discussed here is of astrophysical relevance since the two relevant timescales are separated by three or more orders of magnitude. For higher gravitational couplings, the timescale for gravitational emission shrinks (see fig. A.1). This translates to less observational relevance since the inspiral phase and gravitational emission must occur coincidentally for LISA to be able to probe the Proca cloud properties using the results of this study. It would be astrophysically remarkable to observe a high gravitational coupling using EMRIs as probes with LISA. Nonetheless, for the purposes of this study, perfect coincidence is assumed. This is a reasonable approximation for lower α values, but becomes unreasonable for $\alpha \gtrsim 0.2$, where the cloud decay timescales roughly equals the inspiral timescale which are both extremely short. However, the faithfulness statistic is much less than unity at this point, so the Proca condensate will already be observable with LISA for the astrophysically-relevant timescales. Hence, there is a range of small gravitational coupling, translating to large cloud decay timescales, where the effect on the inspiral will be observable with LISA and where the cloud is sufficiently long lived. To quantitatively take into account the likeliness of coincidence for these two timescales, a full population synthesis study, analysis, and potentially an N-body simulation would be required, which is beyond the scope of this study.

LISA PROCA DISCOVERY POTENTIAL

This study suggests LISA should be able to detect superradiantly-generated Proca fields in the mass range $\mu \in [1.8 \times 10^{-17}, 4.47 \times 10^{-16}]$ eV. The lower bound comes from both the $\frac{1}{\mu}$ scaling of the Proca cloud radius and the reduced SNR for higher SMBH masses. At lower mass values, the Proca cloud extends over several thousands of gravitational radii and so not enough mass is within the orbital radius to significantly modify the secondary BH's trajectory. The resulting waveforms are thus not 'different' enough to be distinguishable with LISA (as determined through the faithfulness and number of orbits statistics). The upper bound comes from the saturation of the superradiant condition. Above a critical mass, the superradiant condition is no longer satisfied and hence a superradiant bound state is not formed. Since the energy flux from the Proca cloud depends on the combination $M \cdot \mu$, lower mass values of the SMBH allow for increased detection probability for higher mass values of the Proca field (and vice versa).

2.4 CONCLUSION

EMRI systems provide a unique arena to study fundamental fields beyond the Standard Model, most notably dark matter candidates. Due to the feeble interaction between dark matter candidates and the Standard Model fields, the mass range of such candidates covers tens of orders of magnitudes. The dark photon, in particular, is a well-motivated candidate for dark matter. It has several production mechanisms including the misalignment mechanism, quantum fluctuations during inflation, tachyonic instabilities arising from couplings to a misaligned axion, and topological defect decays. Various couplings have been proposed for the dark photon. Direct detection searches assume different couplings in the hope of detecting a dark photon-involved process.

Gravitational interactions, on the other hand, need only assume a minimal coupling between gravity and the dark photon. The study performed here investigated the gravitational instability arising from a perturbation in the Proca field in the vicinity of a rotating uncharged black hole. The superradiant phenomenon is responsible for a buildup of a Proca cloud around a Kerr black hole. The resulting cloud modifies the dynamical behavior of an EMRI system during the inspiraling phase, which in turn causes a modification of the measured GW waveform. Depending on

the regime of the parameter space, such a modification can be measured by the future space-borne gravitational observatory LISA.

After solving the Proca equations on a Kerr background and determining the fluxes of energy and angular momentum from the resulting quasibound state, the modification of the waveform was calculated along with derived statistical quantities that allows one to quantify the difference between modified and unmodified waveforms. Limits on the detectable mass range of the Proca field using LISA were obtained and found that LISA should be able to detect Proca fields in the mass range $[1.8 \times 10^{-17}, 4.47 \times 10^{-16}]$ eV.

Throughout the study, several approximations were leveraged. Primarily, the Einstein-Proca system is linearized with respect to the Kerr background. This separates the Proca field from the gravitational perturbations, greatly aiding analytical developments. Secondly, the secondary black hole is approximated to be travelling adiabatically on a sequence of geodesics and is only described as a point particle, the so-called skeletonized approach. Higher order corrections to the motion due to self-force interactions are neglected. Thirdly, the coupling between the secondary black hole and Proca cloud is minimal. The only coupling between them is through the modification to the background the Proca cloud induces via the integrals of motion. This neglects resonant effects between the Proca cloud and secondary black hole. It should be pointed out that dynamical friction effects, accretion, and resonant transitions have been shown to have drastic effects on the orbital phase for the case of a superradiant scalar field. It's reasonable to suspect such dramatic effects will also be present in the Proca field scenario. For example, resonant transitions can produce floating/sinking orbits which have a drastic effect on the resulting gravitational wave signal. Dynamical friction and accretion produce an additional torque on the binary, gradually dephasing it with respect to the vacuum scenario. The results in this study are thus a conservative treatment since incorporating these effects will provide additional ways in which the signal can be modified, widening the observational potential of LISA (see appendix A.3). Finally, the variation of the local gravitational potential is approximated via the function $\Gamma(r)$ in eq. 2.37.

Future studies planned by the authors involve a more accurate calculation of the Einstein-Proca equations of motion, involving numerical relativity calculations to accurately determine the geodesics, as well as dynamical friction effects and accretion effects on the secondary BH which would pro-

vide further modifications to the signal. Additionally, transitions between bound-bound states and bound-unbound states induced by the secondary black hole will likely also have a significant effect. These resonant transitions are planned for future studies as well. Hence, future plans involve more accurate predictions for the interplay between a Proca cloud around an EMRI system with the plan to generate data analysis-ready templates for the future LISA mission. These templates will form the foundation for probing the existence of macroscopic Bose-Einstein condensates around EMRIs. We also plan to extend our analysis to the generalized Proca theories^{210,323}, since the presence of derivative interactions will have large implications both for the background dynamics as well as for the perturbations.

*The task is not to see what has never been seen before, but
to think what has never been thought before about what
you see every day.*

Erwin Schrödinger

3

GRBoondi

GENERALIZED PROCA IS A COMPLICATED EXTENSION of the base theory of Proca. Searching for new numerical solutions of even a single model within the generalized Proca landscape requires building entirely new solvers, making searching the landscape a taxing endeavor. Optimizing these searches will be vital in looking for cosmologically and astrophysically relevant models. To that end, GRBoondi was built to facilitate finding numerical solutions of generalized Proca theories*. GRBoondi is an open-source numerical relativity code written in the C++14 programming language[†]. It uses a hybrid set of Message Passing Interface (MPI) and Open Multi-Processing (OpenMP) parallelization libraries to achieve good performance on the latest architectures, allowing GRBoondi to have excellent scaling on both high-performance computing clusters and supercomputers.

*Boondi is a Koori word for a multi-purpose tool used for hunting and digging.

[†]GRBoondi is freely available for download at <https://github.com/ShawnFell/GRBoondi.git>.

3.1 MOTIVATION

The theory of generalized Proca is an extension of the standard Proca model that aims to extend Proca theories in a consistent way. A natural question that often arises is whether it is possible to extend the interactions present in the Proca field in a way that does not introduce extra propagating degrees of freedom and still maintain second order equations of motion. Previous results have proven a no-go theorem for similar theories, making such extensions potentially problematic. Fortunately, there is a path forward and the solution is given by the generalized Proca models. The main objectives of such an extension are:

- enforce the equations of motion to be of second order
- restrict the temporal component of the Proca field to be non-dynamical.

The second objective is vital to ensure no instabilities arise via negative energy modes, or *ghosts*. This requirement comes from the massive representation of the Lorentz group, which only carries 3 dynamical fields and hence the inclusion of additional interactions should not alter this.

The derivation of the interaction terms resulting from these two requirements is rather lengthy. However, there are a finite number of such terms³²³. The resulting generalized Proca Lagrangian takes the form

$$\mathcal{L}_{\text{g.p.}} = \sqrt{-g} \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \sum_{n=2}^6 \alpha_n \mathcal{L}_n \right), \quad (3.1)$$

where the self-interaction terms are

$$\mathcal{L}_2 = G_2(X, F, Y) \quad (3.2)$$

$$\mathcal{L}_3 = G_3(X) \nabla_\mu A^\mu \quad (3.3)$$

$$\mathcal{L}_4 = G_4(X) \mathcal{R} + \frac{\partial G_4}{\partial X} \left((\nabla_\mu A^\mu)^2 - \nabla_\rho A_\sigma \nabla^\sigma A^\rho \right) \quad (3.4)$$

$$\begin{aligned} \mathcal{L}_5 = G_5(X) G_{\mu\nu} \nabla^\mu A^\nu - \frac{1}{6} \frac{\partial G_5}{\partial X} \left((\nabla_\mu A^\mu)^3 + 2 \nabla_\rho A_\sigma \nabla^\gamma A^\rho \nabla^\sigma A_\gamma \right. \\ \left. - 3 (\nabla_\mu A^\mu) \nabla_\rho A_\sigma \nabla^\sigma A^\rho \right) - g_5(X) \tilde{F}^{\alpha\mu} \tilde{F}_\mu^\beta \nabla_\alpha A_\beta \end{aligned} \quad (3.5)$$

$$\mathcal{L}_6 = G_6(x) \mathcal{L}^{\mu\nu\alpha\beta} \nabla_\mu A_\nu \nabla_\alpha A_\beta + \frac{1}{2} \frac{\partial G_6}{\partial X} \tilde{F}^{\alpha\beta} \tilde{F}^{\mu\nu} \nabla_\alpha A_\mu \nabla_\beta A_\nu, \quad (3.6)$$

where G_i are arbitrary functions of $X = -\frac{1}{2} A_\mu A^\mu$, $Y = A^\mu A^\nu F_\mu^\alpha F_{\nu\alpha}$, $F = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$,

$\mathcal{L}^{\mu\nu\alpha\beta} = \frac{1}{4}E^{\mu\nu\rho\sigma}E^{\alpha\beta\gamma\delta}R_{\rho\sigma\gamma\delta}$ is the dual Riemann tensor, $E^{\alpha\beta\mu\nu}$ is the four-dimensional Levi-Civita tensor, $R_{\alpha\beta\mu\nu}$ is the Riemann tensor, $F_{\mu\nu}$ is the field strength tensor of the Proca field A_μ , $G_{\mu\nu}$ is the Einstein tensor, ∇_μ is the covariant derivative operator with respect to the metric $g_{\mu\nu}$, g is the metric determinant, and $\tilde{F}_{\mu\nu}$ is the dual field strength tensor^{210,324}.

These are the most general interactions giving rise to second order equations of motion and that propagate only three dynamical degrees of freedom. Due to the appearance of the arbitrary functions G_i , the generalized Proca Lagrangian is actually a *landscape* of theories, with each one determined by a specific choice of the arbitrary functions. Studying certain physical systems within the framework of generalized Proca theories requires specifications of the arbitrary functions. This leads to difficulties in making general statements about the landscape of all generalized Proca theories, since there are in principle an infinite number of different choices for the arbitrary functions.

For this reason, to facilitate the numerical study of generalized Proca theories in arbitrary spacetimes, GRBoondi was built. This new NR code allows for arbitrary modifications to the base theory, expediting research into all generalized Proca models.

3.2 NR FUNDAMENTALS

GR is a highly complicated theory and numerical GR is even more so. There are many different conceptual pieces, mathematical tools, and numerical methods that enter in the field of numerical computation. Section 3.2.1 will discuss the basic theoretical method of writing the Einstein equations in a form conducive to numerical computation. This section expands on this by writing the Einstein equations in a form pertinent to temporal evolution, the so-called Cauchy problem. Finally, section 3.2.3 will discuss various different schemes developed throughout the years along with their advantages and disadvantages. These following sections will lay the theoretical groundwork necessary to begin discussing the numerical code.

This section does not aim to provide a comprehensive pedagogical sermon for numerical relativity, instead laying a broad foundation for the rest of this chapter. There are several well-established lecture notes and textbooks that give excellent introductions to the field of numerical relativity. These references include refs. 325 and 326 which present comprehensive introductions on the topic.

3.2.1 3+1 DECOMPOSITION OF THE FIELD EQUATIONS

Before the Einstein equations can be input into a computer, they need to be written in the form of an Initial Value Problem (IVP). This is a form of the equations that, given data for the fields at a specific time, can be uniquely solved for at later times using the field equations. It should be pointed out that this is only possible because the Einstein equations, which are a set of partial differential equations, are *hyperbolic* in nature. This is the statement that the Cauchy problem, which an IVP is, can be uniquely solvable everywhere. A hyperbolic partial differential equation (PDE) is a type of PDE characterized by properties that resemble the wave equation. In fact, the Einstein equations can be written in the form of a non-linear wave equation, called the *relaxed Einstein equations*

$$\square h^{\alpha\beta} = -16\pi\tau^{\alpha\beta} \quad (3.7)$$

$$\partial_\mu h^{\mu\alpha} = 0, \quad (3.8)$$

where $h^{\alpha\beta}$ is the so-called gravitational field amplitude, $\tau^{\alpha\beta}$ is a stress-energy *pseudo*-tensor of the matter fields *and* the gravitational field, and \square is the standard covariant d'Alembertian operator. It is thus clear that the Einstein equations possess a wave-like form, allowing the IVP of the Einstein equations to be solved uniquely.

To write the Einstein problem as an IVP, a common approach is to write the equations in the so-called *3+1 form*. That is to say, the 'covariant-ness' of the equations is undone. Recall that one of Einstein's greatest insights was to combine the three-dimensional space and one dimension of time into a single object called *spacetime*. The 3 + 1 formalism unwinds this process and separates the temporal direction from the spatial one. This is obvious if one wants to do explicit temporal evolution of initial spatial data and is the approach adopted here.

However, there are other formalisms such as the generalized Harmonic approach, where the co-ordinates satisfy $\square x^\mu = H^\mu(x^\mu)$, and H^μ is some four-dimensional function of the co-ordinates. This choice of formalism was used to prove the well-posedness of the Einstein equations and the one chosen in the first groundbreaking simulation of an inspiraling black hole^{229,327}. It is also commonly used today by the SXS collaboration to generate gravitational waveform templates for the

data analysis pipelines in current gravitational wave observatories⁹⁴. Another, albeit less used, formalism is the Cauchy characteristic approach which chooses to slice spacetime using null surfaces instead of temporal surfaces³²⁸.

SLICING SPACETIME

The starting point is to consider a *globally hyperbolic* spacetime (\mathcal{M}, g) , which is to say a spacetime manifold \mathcal{M} equipped with a metric g that admits a Cauchy surface Σ . A Cauchy surface Σ is a surface whose intersecting causal curves, which have no end points, intersect Σ once and only once. This implies that there exists a global time co-ordinate $t : \mathcal{M} \rightarrow \mathbb{R}$ whose level surfaces are Cauchy surfaces. This means the topology of \mathcal{M} is necessarily $\Sigma \times \mathbb{R}$. In other words, spacetime can be consistently sliced according to a constant time co-ordinate, fig. 3.1.

The existence of this temporal co-ordinate allows one to define concepts such as 'evolution', which is a key ingredient in solving the Einstein equations. Moreover, the temporal co-ordinate defines a normal vector for each Cauchy surface, or leaf/hypersurface, of the slicing, or 'foliation', which is defined by

$$n^\mu = -\alpha \nabla^\mu t, \quad (3.9)$$

where α is called the *lapse*, defined by

$$\alpha = (-\nabla^\mu t \nabla_\mu t)^{-\frac{1}{2}}. \quad (3.10)$$

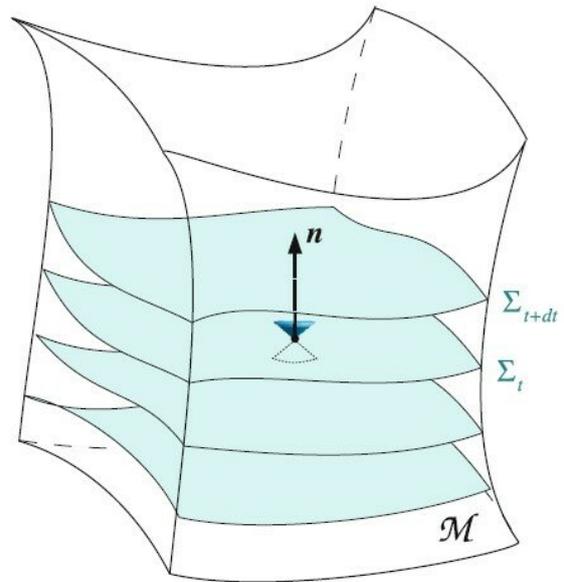


Figure 3.1: ³²⁵ Slicing of spacetime according to Cauchy surfaces, which are defined as the level surfaces of some 'time' co-ordinate.

Another important quantity is the *normal evolution vector* $m^\mu = \alpha n^\mu$. This will be important in defining the evolution equations for various quantities later.

With a four-dimensional metric g and normal vectors to the hypersurface, one can define the projection operator, which takes tensors defined in the four-dimensional spacetime and projects

them down onto the hypersurface

$$\gamma_{\beta}^{\mu} = g_{\beta}^{\mu} + n^{\mu}n_{\beta} . \quad (3.11)$$

For example, a four-dimensional vector $V^{\mu} \in \mathcal{M}$ can be projected to the hypersurface Σ via $\gamma_{\alpha}^i V^{\alpha} = V^i \in \Sigma$. For an arbitrary (r,s)-rank tensor, the projection onto the hypersurface Σ is accomplished via

$$\gamma_{\mu}^i \gamma_{\nu}^j \cdots \gamma_k^{\alpha} \gamma_l^{\beta} \cdots T_{\alpha\beta \cdots}^{\mu\nu \cdots} = T_{kl \cdots}^{ij \cdots} \in \Sigma . \quad (3.12)$$

This projection operator also defines a metric on the hypersurfaces, defined by

$$\gamma_{\mu\nu} = g_{\mu\nu} + n_{\mu}n_{\nu} . \quad (3.13)$$

We have thus split the spacetime and its fundamental object, the metric, into three spatial dimensions defined as the level surfaces of a time co-ordinate and equipped with a three-dimensional spatial metric γ and a single temporal dimension which defines the normal vector for the hypersurfaces. For any tensorial quantity defined on the manifold \mathcal{M} , it can be split into spatial and temporal contributions via the projection operator and the normal vector. For example, given a vector $V^{\mu} \in \mathbb{R}$, the spatial and temporal contributions can be factored out by defining

$$\gamma_{\mu}^i V^{\mu} = X^i \in \Sigma \quad (3.14)$$

$$n_{\mu} V^{\mu} = \phi , \quad (3.15)$$

where ϕ is the scalar quantity that signifies 'how much' V^{μ} points out of the hypersurface Σ . The full four-dimensional vector can then be reconstructed via $V^{\mu} = X^{\mu} + n^{\mu}\phi$. This is a general procedure that will be vital when decomposing additional matter fields living on the spacetime manifold.

Towards a decomposition of the full four-dimensional metric g in terms of quantities orthogonal and transverse to the hypersurface, the remaining quantity needed is the *shift vector*, which is defined by*

$$\beta^i = \gamma_{\mu}^i t^{\mu} , \quad (3.16)$$

*Actually, choosing this definition is choosing a specific set of co-ordinates that are said to be *adapted co-ordinates*, which are co-ordinates of the form $x^{\mu} = (t, x^i)$, where the x^i are co-ordinates on the hypersurface.

where t^μ is the tangent vector to curves of constant spatial co-ordinate, i.e. temporal curves, fig. 3.2.

With this decomposition, the tangent vector to curves of constant spatial co-ordinates can be decomposed as

$$\begin{aligned} t^\mu &= (\gamma_\nu^\mu - n^\mu n_\nu) t^\nu \\ &= -n^\mu n_\nu t^\nu + \gamma_\nu^\mu t^\nu = \alpha n^\mu + \beta^\mu . \end{aligned} \quad (3.17)$$

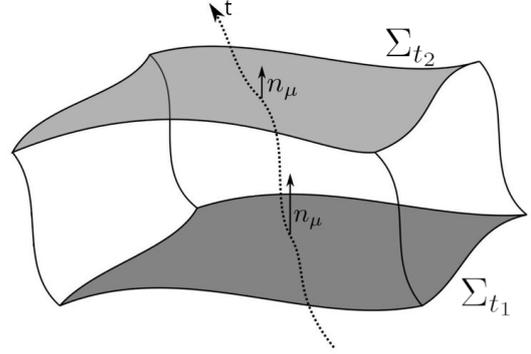


Figure 3.2: Slicing of spacetime according to Cauchy surfaces, which are defined as the level surfaces of some 'time' co-ordinate whose curve is graphically represented as the dashed line.

With these ingredients, the full four-dimensional metric can be decomposed entirely in terms of quantities defined on the hypersurface and scalar functions and it takes the form

$$g_{\mu\nu} = \begin{bmatrix} -\alpha^2 + \beta_i \beta^i & \beta_j \\ \beta_i & \gamma_{ij} \end{bmatrix} , \quad (3.18)$$

or, in terms of the infinitesimal line element

$$ds^2 = -\alpha^2 dt^2 + \gamma_{ij} (dx^i + \beta^i dt) (dx^j + \beta^j dt) . \quad (3.19)$$

Using these definitions, one can deduce a simple relation between the determinant of the spatial and four-dimensional metric. Let $g = \det(g_{\mu\nu})$ and $\gamma = \det(\gamma_{ij})$. It follows from a simple computation that

$$\sqrt{-g} = \alpha \sqrt{\gamma} . \quad (3.20)$$

This relation will find much importance in computing various integrals over hypersurfaces, especially in the context of GRBoondi.

Derivations on the hypersurface and the four-dimensional manifold can, remarkably, be written in terms of one another. Let D be the covariant derivative with respect to the spatial metric $\gamma_{\mu\nu}$ and, as before, ∇ the covariant derivative with respect to the full four-dimensional metric g^* . For

*These are also called *connection coefficients* and relate to how quantities defined in two different tangent spaces can be compared.

a tensor $T \in \mathcal{T}(\mathcal{M})$, the covariant derivative of T is also a tensor, $\nabla(T) \in \mathcal{T}(\mathcal{M})$. The same holds true for a tensor $P \in \mathcal{T}(\Sigma)$ and D . In fact, one can find a relation between the two covariant derivative operations acting on the four-dimensional tensor and its projected component, which takes the form*

$$D_w T_{kl\dots}^{ij\dots} = \gamma_\mu^i \gamma_\nu^j \dots \gamma_k^\alpha \gamma_l^\beta \dots \gamma_w^\sigma \nabla_\sigma T_{\alpha\beta\dots}^{\mu\nu\dots} . \quad (3.21)$$

Thus, tensors and their derivations have been decomposed in terms of scalar and three-dimensional quantities defined on the hypersurfaces of the foliations. This allows any four-dimensional equation to be written in terms of spatial and temporal quantities.

GAUGE FREEDOM

The Einstein equations constitute a type of gauge theory. This means there is a certain kind of freedom in defining the metric field that does not change the physical states of the theory. The decomposition in the previous section has made this gauge freedom manifest in terms of the lapse and shift vector. This is rather apparent by inspection of eqs. 3.10 and 3.16. The temporal co-ordinate can be freely scaled, which should not change any of the dynamical equations governing the physical states. Since the shift and lapse are defined in terms of this temporal co-ordinate, they are also freely specifiable without changing the physical states, hence encoding the gauge freedom apparent in the theory.

There is a particular choice of observers that become special in a foliation by spatial hypersurfaces, which are called the *Eulerian observers*. Since the normal vector is a unit timelike vector, it can be regarded as the four-dimensional velocity of some physical observer, called the Eulerian observer. The proper time τ for this observer is related to the global time co-ordinate t via

$$\delta\tau = \alpha\delta t . \quad (3.22)$$

This justifies the use of the word 'lapse function'. The lapse encodes the relation between the 'co-ordinate time' t and the physical time τ measured by an Eulerian observer. Additionally, the shift

*It should be noted that the T on the left-hand side is a three-dimensional projected version of the four-dimensional tensor T on the right-hand side. It's written here using the same T in order to avoid useless cluttering of the equations.

encodes the 'velocity' of the lines of constant co-ordinate x^i relative to the Eulerian observer. If the observer is at position x^i at time t , then at a later time of $t + \delta t$, the observer will be at position $x^i - \beta^i \delta t$.

Since the lapse and shift are completely freely specifiable, we could in principle choose whatever functions we want. However, practically, the gauge choice plays a huge role in ensuring the numerical simulations remain stable and accurate, especially in the context of simulations that contain curvature singularities. For example, taking the choice $\alpha = 1$, called *geodesic slicing*, and the Schwarzschild spacetime in isotropic co-ordinates, it can be shown that observers within the horizon reach the singularity in finite proper time which scales proportionately to the mass of the black hole. Since $\alpha = 1$, this means the observer will reach the singularity in finite *co-ordinate* time as well, posing a huge problem for numerical simulations. When trying to evolve the spacetime using this gauge choice, observers will reach the singularity within the simulation timeframe, causing the simulation to break down as the curvature measured by these observers becomes infinite.

CURVATURE OF THE HYPERSURFACES

The curvature of the hypersurface, relative to the four-dimensional spacetime, can be encoded in a quantity called the *extrinsic curvature*, which is nothing but the projected part of the covariant derivative of the normal vector

$$K_{\mu\nu} = -\gamma_{\mu}^{\alpha} \nabla_{\alpha} n_{\nu} . \quad (3.23)$$

The extrinsic curvature is sometimes called the *second fundamental form*, the spatial metric being the *first fundamental form*. The extrinsic curvature is related to the principal curvatures of the hypersurface at a given point, via the Weingarten map, or *shape operator*. The shape operator is nothing but the directional covariant derivative of the normal vector, hence encoding information about the curvature of the hypersurface in that direction.

Remarkably, the second fundamental form can be related to the first fundamental form in a succinct way,

$$K_{ij} = -\frac{1}{2} \mathcal{L}_n \gamma_{ij} , \quad (3.24)$$

where \mathcal{L}_n is the Lie derivative in the direction of the normal vector. It can also be written directly in

terms of the covariant derivative of the normal vector as

$$K_{\mu\nu} = -\nabla_\nu n_\mu - n_\nu D_\mu \ln(\alpha) . \quad (3.25)$$

PROJECTING THE RIEMANN TENSOR

The Einstein equations* are written in terms of various projections of the Riemann curvature tensor, namely the Ricci tensor $R_{\mu\nu}$ and Ricci scalar \mathcal{R}

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = \frac{8\pi G}{c^4}\mathcal{T}_{\mu\nu} , \quad (3.26)$$

where \mathcal{T} is the stress-energy tensor of the matter fields present in the spacetime. Towards the objective of decomposing the evolution equations into spatial and temporal components, the Ricci tensor and Ricci scalar must be decomposed as well. This will inevitably come from a full decomposition of the Riemann tensor itself. These projected forms of the Riemann tensor will form a foundation of the 3+1 formalism for GR.

In order to avoid further cluttering of equations, we introduce a new notation for spatial and full four-dimensional quantities. Unless otherwise stated, full four-dimensional tensors will be denoted with a ⁴ behind the symbol, such as the full four-dimensional Ricci scalar ⁴ \mathcal{R} . Three dimensional tensors will be denoted with no superscript, such as the three-dimensional Ricci scalar \mathcal{R} .

The Riemann tensor is a rank-4 tensor, meaning it is a $4 \times 4 \times 4 \times 4$ matrix of components. This means it will have various projections, depending on the index that is projected. The first projection will be a complete projection of all four indices to the hypersurface, $\gamma_\mu^\alpha \gamma_\nu^\beta \gamma_\rho^\sigma \gamma_\lambda^\gamma ({}^4R)_{\alpha\beta\gamma}^\rho$. This can be written using the three-dimensional Riemann tensor and the extrinsic curvature tensor as

$$\gamma_\mu^\alpha \gamma_\nu^\beta \gamma_\rho^\sigma \gamma_\lambda^\gamma ({}^4R)_{\alpha\beta\gamma}^\rho = R_{\mu\nu\lambda}^\sigma + K_\nu^\sigma K_{\mu\lambda} - K_\lambda^\sigma K_{\nu\mu} . \quad (3.27)$$

This is the so-called *Gauss relation*. We can contract the σ and ν indices to obtain the *contracted*

*Here, we ignore the cosmological constant, which would contribute a term $\Lambda g_{\mu\nu}$ to the left-hand side.

Gauss relation,

$$\gamma_{\alpha}^{\mu}\gamma_{\beta}^{\nu}({}^4\mathcal{R})_{\mu\nu} + \gamma_{\alpha\mu}n^{\nu}\gamma_{\beta}^{\rho}n^{\sigma}({}^4R)_{\nu\rho\sigma}^{\mu} = \mathcal{R}_{\alpha\beta} + \mathcal{K}K_{\alpha\beta} - K_{\alpha\mu}K_{\beta}^{\mu}, \quad (3.28)$$

where $\mathcal{K} = g^{\mu\nu}K_{\mu\nu}$ is the trace of the extrinsic curvature. This can be contracted again over the indices α and β to obtain

$${}^4\mathcal{R} + 2({}^4\mathcal{R})_{\mu\nu}n^{\mu}n^{\nu} = \mathcal{R} + \mathcal{K}^2 - K_{ij}K^{ij}. \quad (3.29)$$

This famous equation is called the *scalar Gauss relation* and is a generalization of Gauss' famous *Theorema Egregium*, which relates the intrinsic curvature of Σ , encoded by the Ricci scalar \mathcal{R} , to its extrinsic curvature, encoded by $\mathcal{K}^2 - K_{ij}K^{ij}$.

We can derive another relation using the *Ricci identity*, which is a statement about the non-commutativity of the covariant derivative operator,

$$[\nabla_{\alpha}, \nabla_{\beta}]n^{\gamma} = {}^4R_{\mu\alpha\beta}^{\gamma}n^{\mu}. \quad (3.30)$$

Projecting this relation using the tools just developed, one finds

$$\gamma_{\rho}^{\gamma}n^{\sigma}\gamma_{\alpha}^{\mu}\gamma_{\beta}^{\nu}({}^4R)_{\sigma\mu\nu}^{\rho} = D_{\beta}K_{\alpha}^{\gamma} - D_{\alpha}K_{\beta}^{\gamma}, \quad (3.31)$$

which is called the *Codazzi relation*. Contracting the Codazzi relation once along the α and γ indices yields the *contracted Codazzi relation*

$$\gamma_{\alpha}^{\mu}n^{\nu}({}^4\mathcal{R}_{\mu\nu}) = D_{\alpha}\mathcal{K} - D_{\mu}K_{\alpha}^{\mu}. \quad (3.32)$$

Taken together, eqs. 3.27 and 3.31 form the *Gauss-Codazzi relations*, which underpin the 3+1 formalism of GR. They will be explicitly used in projecting the constraints hidden within the GR field equations onto the hypersurfaces.

Instead of projecting the entire Riemann tensor onto the hypersurface in eq. 3.31, we can project

it twice onto the hypersurface and twice along the normal vector

$$\gamma_{\alpha\mu}n^\sigma\gamma_\beta^\nu n^\rho({}^4R)_{\rho\nu\sigma}^\mu = \gamma_{\alpha\mu}n^\sigma\gamma_\beta^\nu[\nabla_\nu, \nabla_\sigma]n^\mu. \quad (3.33)$$

These leads to

$$\gamma_{\alpha\mu}n^\sigma\gamma_\beta^\nu n^\rho({}^4R)_{\rho\nu\sigma}^\mu = \frac{1}{\alpha}\mathcal{L}_m K_{\alpha\beta} + \frac{1}{\alpha}D_\alpha D_\beta \alpha + K_{\alpha\mu}K_\beta^\mu, \quad (3.34)$$

where $m^\mu = \alpha n^\mu$ is the *evolution normal vector*. This is sometimes called the *Ricci equation*.

Together with the Gauss-Codazzi relations, these three relations complete the 3+1 decomposition of the Riemann tensor. One can take the trace of the Ricci equation, combine it with the scalar Gauss equation, and find an expression for the four-dimensional Ricci scalar in terms of three-dimensional quantities

$${}^4\mathcal{R} = \mathcal{R} + \mathcal{K}^2 + K_{ij}K^{ij} - \frac{2}{\alpha}\mathcal{L}_m \mathcal{K} - \frac{2}{\alpha}D_i D^i \alpha. \quad (3.35)$$

So far, the projection of the left-hand side of the Einstein equations have been formulated in terms of the 3+1 variables. What remains is the decomposition of the right-hand side, the stress energy tensor of the matter fields $\mathcal{T}_{\mu\nu}$. It is a rank-2 symmetric tensor, so there are three possible projections

- The *matter energy density* measured by the Eulerian observer is given by the double projection along the normal direction

$$\rho = \mathcal{T}_{\mu\nu}n^\mu n^\nu, \quad (3.36)$$

which follows from the fact that the four velocity of the Eulerian observer is the unit normal vector itself.

- Similarly, the *matter momentum density* as measured by the Eulerian observer is the linear form

$$p_\alpha = -\mathcal{T}_{\mu\nu}\gamma_\alpha^\mu n^\nu \quad (3.37)$$

whose sign is chosen so that it is future-oriented.

- Finally, the *matter stress tensor* as measured by the Eulerian observer is the bilinear form

$$S_{\mu\nu} = \mathcal{T}_{\alpha\beta} \gamma_{\mu}^{\alpha} \gamma_{\nu}^{\beta} . \quad (3.38)$$

Under this decomposition, the full four-dimensional stress energy tensor can be reconstructed as

$$\mathcal{T}_{\mu\nu} = S_{\mu\nu} + n_{\mu} p_{\nu} + n_{\nu} p_{\mu} + \rho n_{\mu} n_{\nu} . \quad (3.39)$$

THE CONSTRAINTS

Behind the scenes, GR is a constrained system. It's not obvious what those constraints are from first inspection of the equations themselves, however they are revealed by the 3+1 decomposition.

Projecting eq. 3.26 entirely along the normal direction and applying the scalar Gauss equation eq. 3.29 results in the equation

$$\begin{aligned} {}^4\mathcal{R}_{\mu\nu} n^{\mu} n^{\nu} + \frac{1}{2}({}^4\mathcal{R}) &= 8\pi \mathcal{T}_{\mu\nu} n^{\mu} n^{\nu} \\ \Downarrow \\ \mathcal{R} + \mathcal{K}^2 - K_{ij} K^{ij} &= 16\pi E . \end{aligned} \quad (3.40)$$

This is the *Hamiltonian constraint* and, as its name suggests, follows also from a Hamiltonian treatment of GR.

Projecting eq. 3.26 once along the hypersurface and once along the normal vector, then applying the contracted Codazzi equation eq. 3.32, yields

$$\begin{aligned} {}^4\mathcal{R}_{\mu\nu} \gamma_i^{\mu} n^{\nu} - \frac{1}{2}({}^4\mathcal{R}) g_{\mu\nu} \gamma_i^{\mu} n^{\nu} &= 8\pi \mathcal{T}_{\mu\nu} \gamma_i^{\mu} n^{\nu} \\ \Downarrow \\ D_j K_i^j - D_i \mathcal{K} &= 8\pi p_i . \end{aligned} \quad (3.41)$$

This is the *momentum constraint*.

The remaining projection of the Einstein equations is the complete projection along the hypersurface. Combining the contracted Gauss equation with the Ricci equation, eqs. 3.28 and 3.34, yields

$$\mathcal{L}_m K_{ij} = -D_i D_j \alpha + \alpha (\mathcal{R}_{ij} + \mathcal{K} K_{ij} - 2K_{il} K_j^l + 4\pi [(S - \rho) \gamma_{ij} - 2S_{ij}]) , \quad (3.42)$$

where $S = \gamma^{ij} S_{ij}$ is the trace of the matter stress tensor.

eqs. 3.40-3.42 are completely equivalent to the original Einstein equations, eq. 3.26. Eq. 3.42 is an equation for a rank-2 bilinear form in the leaves of the foliation, involving only symmetric tensors. Hence it therefore has six independent components. Eq. 3.40 is a scalar equation, yielding an additional component. Eq. 3.41 is a vectorial equation, yielding three more components. In total, eqs. 3.40-3.42 are a system of equations for ten independent components, the same as the original Einstein equations. Including eq. 3.24, all variables can be evolved using spatial and temporal information.

NR AS A CAUCHY PROBLEM

As mentioned earlier, the Einstein equations can be placed in a form that makes the problem of solving them a Cauchy IVP. The evolution normal vector can be written as $m^\mu = t^\mu - \beta^\mu$. Using properties of the Lie derivative, this implies

$$\mathcal{L}_m = \mathcal{L}_t - \mathcal{L}_\beta . \quad (3.43)$$

Using so-called adapted co-ordinates, i.e. co-ordinates whose temporal component is simply t , one can write

$$\mathcal{L}_t = \frac{\partial}{\partial t} , \quad (3.44)$$

which implies

$$\mathcal{L}_m T_{j\dots}^{i\dots} = \left(\frac{\partial}{\partial t} - \mathcal{L}_\beta \right) T_{j\dots}^{i\dots} . \quad (3.45)$$

Hence, eq. 3.24 and 3.42 can be rewritten as

$$\left(\frac{\partial}{\partial t} - \mathcal{L}_\beta\right) K_{ij} = -D_i D_j \alpha + \alpha (\mathcal{R}_{ij} + \mathcal{K} K_{ij} - 2K_{il} K_j^l + 4\pi [(S - \rho) \gamma_{ij} - 2S_{ij}]) \quad (3.46)$$

$$\left(\frac{\partial}{\partial t} - \mathcal{L}_\beta\right) \gamma_{ij} = -2\alpha K_{ij} . \quad (3.47)$$

Thus, the full Einstein equations may be written as

3+1 Einstein Equations

$$\left(\frac{\partial}{\partial t} - \mathcal{L}_\beta\right) \gamma_{ij} = -2\alpha K_{ij}$$

$$\left(\frac{\partial}{\partial t} - \mathcal{L}_\beta\right) K_{ij} = -D_i D_j \alpha + \alpha (\mathcal{R}_{ij} + \mathcal{K} K_{ij} - 2K_{il} K_j^l + 4\pi [(S - \rho) \gamma_{ij} - 2S_{ij}])$$

$$D_j K_i^j - D_i \mathcal{K} = 8\pi p_i$$

$$\mathcal{R} + \mathcal{K}^2 - K_{ij} K^{ij} = 16\pi E .$$

Hence, given constraint-satisfying initial data at some particular time slice, the 3+1 Einstein equations will determine the physical state at later times. However, this is not enough to ensure the solutions obtained from the 3+1 Einstein equations are unique and valid solutions. One also has to ensure these equations are *well-posed*, which is closely related to the concept of hyperbolicity.

3.2.2 WELL-POSEDNESS AND HYPERBOLICITY

Stability and accuracy of the IVP of the 3+1 Einstein equations relies on the concept of well-posedness, which comes from the mathematical theory of PDEs. Since stability of the numerical solutions of GRBoondi is vital for simulating a wide range of physical systems, this concept will be touched on here, although a pedagogical discussion is deferred to the literature^{329,330}.

The starting point for a discussion of well-posedness begins with a general discussion of PDEs. Consider a system of PDEs of the generic form

$$\partial_t u = P(D)u \quad (3.48)$$

for some arbitrary n -dimensional vector-valued function of space and time, and $P(D)$ an $n \times n$ matrix whose components are smooth functions of the spatial derivation operators. The IVP for such a system corresponds to finding a solution of the form $u(x, t)$ given some initial data $u(x, t = t_0)$. A critical property for such a system of PDEs is that of *well-posedness*. This is the idea that the solution $u(x, t)$ depends continuously on the initial data $u(x, t = t_0)$. In other words, small perturbations in the initial data $u(x, t = t_0)$ produces only small changes in the final states $u(x, t)$. This can be encapsulated in the following definition ³³⁰:

Definition 3.2.1

A system of partial differential equations is called well-posed if there exists a norm (usually a Sobolev norm) and two constants k, α , such that for all initial data and all positive times,

$$\|u(\cdot, t)\| \leq ke^{\alpha t} \|u(\cdot, 0)\| .$$

Most systems of evolution in mathematical physics are well-posed, however simple examples can already showcase systems that are ill-posed. Consider for example the simple inverse heat equation

$$\partial_t u = -\partial_x^2 u . \tag{3.49}$$

Assume the initial data is single Fourier mode $u(x, 0) = e^{ikx}$. The solution is easily found to be

$$u(x, t) = e^{k^2 t + ikx} , \tag{3.50}$$

which grows exponentially in time with an exponent that depends on the frequency of the Fourier mode k . By increasing k , the rate of growth can be increased arbitrarily, so the general solution cannot be bounded by an exponential that is independent of the initial data. Additionally, given any generic initial data, we can always add a small perturbation of the form ϵe^{ikx} for $\epsilon \ll 1$ and $k \gg 1$, such that after a finite amount of time has elapsed, the solution can be very different than the original unperturbed system. Hence, there is no continuity of the solution with respect to the initial data. Therefore, this system is ill-posed.

Hyperbolicity is a concept closely related to well-posedness. In fact, it can be shown that a hyper-

bolic system of PDEs is well-posed under very general conditions. Consider a system of first order evolution equations

$$\partial_t u + M^i \partial_i u = s(u) , \quad (3.51)$$

where M^i are $n \times n$ constant-valued matrices and $s(u)$ is a source vector-valued function that may depend on the u s but not on their derivatives. It can be shown quite generally that if the system of equations with $s(u) = 0$ is well-posed, then the system of equations with $s(u) \neq 0$ is also well-posed. Intuitively, hyperbolicity is attributed to systems of evolution equations that behave in some way as wave equations. Such systems are well-posed and the speed of propagation of information should be finite. The notion of hyperbolicity can be related to properties of the matrices M^i , which are also called the characteristic matrices. To that end, consider an arbitrary unit vector n_i and construct the matrix $P(n_i) = M^i n_i$, which is known as the *principal symbol* of the system of equations. Then, using the principal symbol, one can define several different versions of hyperbolicity:

- The system of equations eqs. 3.51 is called *strongly hyperbolic* if the principal symbol has real eigenvalues and a complete set of eigenvectors for all n_i .
- The system of equations is called *weakly hyperbolic* if the principal symbol has real eigenvalues for all n_i but does not have a complete set of eigenvectors.
- If all the characteristic matrices are symmetric, then the system is called *symmetric hyperbolic*.
- The system of equations is called *strictly hyperbolic* if all the eigenvalues of the principal symbol are not only real but also distinct for all n_i .

It can be shown that strongly and symmetric hyperbolic systems are well-posed. Moreover, strictly hyperbolic systems are automatically strongly hyperbolic.

The system eq. 3.51 can be recast into the form

$$\partial_t u + \partial_i F^i(u) = s(u) , \quad (3.52)$$

where F^i are vector-valued functions of the u s and spacetime co-ordinates, but not of derivatives of

the u s. The vectors F^i are called the *flux vectors*, and the characteristic matrices are then given by

$$(M^i)_{ab} = \frac{\partial F_a^i}{\partial u_b}. \quad (3.53)$$

A system of the form eq. 3.52 is called a *balance law* since the change of the solution u in a volume is given by a balance between the fluxes entering or leaving the volume element and sources.

If $s(u) = 0$, then eq. 3.52 is called a *conservation law*.

A simple example of a hyperbolic equation is the one-dimensional advection equation

$$\partial_t u + v \partial_x u = 0 \quad (3.54)$$

with $v \in \mathbb{R}$. This solution propagates the initial data with a speed v without changing the initial data's profile. In other words, for initial data $u(x, 0) = f(x)$, the solution to eq. 3.54 is simply $u(x, t) = f(x - vt)$. This is an example of a *strongly hyperbolic* system.

HYPERBOLICITY OF THE EINSTEIN EQUATIONS

Now the following question arises: Under what conditions are the $3+1$ form of the Einstein equations well-posed? It turns out the $3+1$ Einstein equations are well-posed under two mutually independent conditions:

- The momentum constraints eq. 3.41 can be guaranteed to be identically satisfied.
- Either the 'densitized lapse' $\tilde{\alpha} = \alpha/\sqrt{\gamma}$ is assumed to be a known function of spacetime (but not the lapse itself), or we use the slicing condition of the so-called *Bona-Masso* family:

$$\frac{d}{dt} \alpha = -\alpha^2 f(\alpha) \mathcal{K}.$$

Under these conditions, the Einstein evolution equations would be strongly hyperbolic, implying well-posedness. While choosing a type of gauge is easy enough, ensuring the momentum constraints are continuously satisfied is an entirely different challenge. Firstly, using finite methods for numerical integrations will inevitably lead to violations of the constraints, due to the inability for the solver to resolve details finer than the highest resolution of the computation. Secondly, even in the continuum limit, when the resolution is taken to infinity, the strong hyperbolicity of the system

would only be guaranteed for very specific choices of the initial data. Hence, even in the absence of matter, when $S = \rho = 0$, the evolution equations in the form of the $3+1$ Einstein equations are only weakly hyperbolic. Thus, alternative versions of the evolution equations must be sought to ensure stability and accuracy of the numerical scheme.

3.2.3 EVOLUTION AND SCHEMES

A large number of alternative formulations of the Einstein equations have been proposed in the literature, more than there are groups that can code the solvers. These formulations all search for evolution equations that result in stable and accurate evolutions.

THE BSSNOK FORMULATION

In 1987, Nakamura, Oohara, and Kojima presented a new formulation of the Einstein equations based on a conformal transformation and which showed improved stability compared to the $3+1$ Einstein equations³³¹. Largely unnoticed for over a decade, it wasn't until Baumgarte and Shapiro rigorously compared it to the $3+1$ Einstein equations and showed that the new formulation had excellent stability properties in all considered cases. Today, this new formulation is widely accepted and the most commonly used in full three-dimensional solvers. The most common version of this formulation is called the *BSSN* (Baumgarte, Shapiro, Shibata, Nakamura), or *BSSNOK* (Baumgarte, Shapiro, Shibata, Nakamura, Oohara, Kojima), formulation.

A central point in the BSSN formulation is the application of a conformal transformation of the spatial metric

$$\tilde{\gamma}_{ij} = \psi^{-4} \gamma_{ij} , \quad (3.55)$$

where ψ is a conformal factor that can in principle be freely specified. For example, in black hole spacetimes, the conformal factor can be chosen to be the initial singular conformal factor and then enforce this factor to remain temporally fixed. This allows the *non-singular* part of the metric to be evolved and is known as the *puncture* method. In the BSSN formulation, the conformal factor is chosen such that the conformal spatial metric $\tilde{\gamma}_{ij}$ has unit determinant

$$\psi^4 = \gamma^{1/3} \Rightarrow \psi = \gamma^{1/12} . \quad (3.56)$$

Additionally, the extrinsic curvature is separated into its trace and its trace-free part,

$$A_{ij} = K_{ij} - \frac{1}{3}\gamma_{ij}\mathcal{K} , \quad (3.57)$$

which is then conformally rescaled

$$\tilde{A}_{ij} = \psi^{-4} A_{ij} . \quad (3.58)$$

An important additional piece for the BSSN formulation is the introduction of three new auxiliary functions called *conformal connection functions*

$$\tilde{\Gamma}^i = \tilde{\gamma}^{jk}\tilde{\Gamma}_{jk}^i = -\partial_j\tilde{\gamma}^{ij} , \quad (3.59)$$

where $\tilde{\Gamma}$ are the Christoffel symbols of the conformally rescaled metric. These additional variables are promoted to full dynamical variables, meaning they will possess evolution equations. The introduction of these variables is to allow the conformal Ricci tensor $\tilde{\mathcal{R}}_{ij}$ to be expressed in a manifestly elliptic form, which is crucial for ensuring strong hyperbolicity and hence well-posedness. In order to improve further the hyperbolicity properties of the evolution equations, multiples of the constraints are added in. For example, in the evolution equation for the extrinsic curvature, the Hamiltonian constraint eq. 3.40 is used to remove the Ricci scalar. In the evolution equation for $\tilde{\Gamma}^i$, the momentum constraint is used to remove derivatives of A^{ij} , which is important for stability. Additionally, instead of evolving the conformal factor explicitly, a common approach for black hole spacetimes is to evolve $\chi = \gamma^{-1/3}$.

The evolution variables for the BSSN formulation of the Einstein equations are then a set of 16 variables $\{\chi, \tilde{\gamma}_{ij}, \mathcal{K}, \tilde{A}_{ij}, \tilde{\Gamma}^i\}$ and whose evolution is governed by eqs. 3.60-3.64.

The second covariant derivative of the lapse can be calculated using the BSSN variables as

$$D_i D_j \alpha = \partial_i \partial_j \alpha - \tilde{\Gamma}_{ij}^k \partial_k \alpha + \frac{1}{2\chi} (\partial_i \alpha \partial_j \chi + \partial_j \alpha \partial_i \chi - \tilde{\gamma}_{ij} \tilde{\gamma}^{kl} \partial_k \alpha \partial_l \chi) \quad (3.65)$$

and the full spatial Ricci tensor is computed via

$$\mathcal{R}_{ij} = \tilde{\mathcal{R}}_{ij} + \frac{1}{\chi} \mathcal{R}_{ij}^\chi \quad (3.66)$$

$$\partial_t \tilde{\gamma}_{ij} = -2\alpha \tilde{A}_{ij} + \tilde{\gamma}_{ik} \partial_j \beta^k + \tilde{\gamma}_{jk} \partial_i \beta^k - \frac{2}{3} \tilde{\gamma}_{ij} \partial_k \beta^k + \beta^k \partial_k \tilde{\gamma}_{ij} \quad (3.60)$$

$$\begin{aligned} \partial_t \tilde{A}_{ij} = & \chi [-D_i D_j \alpha + \alpha (R_{ij} - 8\pi S_{ij})]^{TF} + \alpha (K \tilde{A}_{ij} - 2 \tilde{A}_{ik} \tilde{A}_j^k) + \beta^k \partial_k \tilde{A}_{ij} + \\ & \tilde{A}_{ik} \partial_j \beta^k + \tilde{A}_{jk} \partial_i \beta^k - \frac{2}{3} \tilde{A}_{ij} \partial_k \beta^k \end{aligned} \quad (3.61)$$

$$\partial_t \chi = \frac{2}{3} \chi (\alpha K - \partial_i \beta^i) + \beta^i \partial_i \chi \quad (3.62)$$

$$\partial_t K = -\chi \tilde{\gamma}^{kl} D_k D_l \alpha + \alpha \left(\tilde{A}^{ij} \tilde{A}_{ij} + \frac{1}{3} K^2 \right) + 4\pi \alpha (\rho + S) + \beta^i \partial_i K \quad (3.63)$$

$$\begin{aligned} \partial_t \tilde{\Gamma}^i = & \tilde{\gamma}^{jk} \partial_j \partial_k \beta^i + \frac{1}{3} \tilde{\gamma}^{ij} \partial_j \partial_k \beta^k - \tilde{\Gamma}^k \partial_k \beta^i + \frac{2}{3} \tilde{\Gamma}^i \partial_k \beta^k + \beta^k \partial_k \tilde{\Gamma}^i - \\ & \tilde{A}^{ik} \left(\frac{3\alpha}{\chi} \partial_k \chi + 2 \partial_k \alpha \right) + 2\alpha \tilde{\Gamma}_{kl}^i \tilde{A}^{kl} - \frac{4}{3} \alpha \tilde{\gamma}^{ik} \partial_k \mathcal{K} - 16\pi \alpha \tilde{\gamma}^{ik} p_k \end{aligned} \quad (3.64)$$

$$\tilde{\mathcal{R}}_{ij} = -\frac{1}{2} \tilde{\gamma}^{kl} \partial_k \partial_l \tilde{\gamma}_{ij} + \tilde{\gamma}_{k(i} \partial_{j)} \tilde{\Gamma}^k + \frac{1}{2} \tilde{\Gamma}^k \partial_k \tilde{\gamma}_{ij} + \tilde{\gamma}^{lm} \left(2 \tilde{\Gamma}_{l(i} \tilde{\Gamma}_{j)km} + \tilde{\Gamma}_{im}^k \tilde{\Gamma}_{klj} \right) \quad (3.67)$$

$$\mathcal{R}_{ij}^\chi = \frac{1}{2} \left(\tilde{D}_i \tilde{D}_j \chi + \tilde{\gamma}_{ij} \tilde{\gamma}^{kl} \tilde{D}_k \tilde{D}_l \chi \right) - \frac{1}{4\chi} \left(\partial_i \chi \partial_j \chi + 3 \tilde{\gamma}_{ij} \tilde{\gamma}^{kl} \partial_k \chi \partial_l \chi \right), \quad (3.68)$$

where \tilde{D} is the covariant derivative with respect to the conformal spatial metric $\tilde{\gamma}_{ij}$. The BSSN evolution equations are supplemented with algebraic constraints due to the definition of the new variables

$$\tilde{\gamma} = 1 \quad (3.69)$$

$$\tilde{\gamma}^{ij} \tilde{A}_{ij} = 1 \quad (3.70)$$

$$\tilde{\Gamma}^i - \tilde{\gamma}^{kl} \tilde{\Gamma}_{kl}^i = 0. \quad (3.71)$$

The first constraint typically is not enforced since it is fairly stable for simulations that have well-behaved initial data. The second constraint is required to be manually enforced throughout the simulation in order to maintain stability. The third also needs to be enforced.

It turns out, after performing a detailed analysis, that the BSSN evolution equations are strongly hyperbolic, implying that they are well-posed. The BSSN system has been proven time and again to be more numerically stable than other formulations.

CCZ₄ FORMULATION: DAMPING THE CONSTRAINTS

An important point to be made about the BSSN formulation is that the constraints need to be carefully monitored throughout the entire simulation, since errors can accumulate over time due to finite differencing, boundary effects, and possible other origins. The accumulation of these errors drives the simulation data away from solutions of the Einstein equations. To counteract these challenges, a new formulation was developed called the *CCZ₄* formalism. The *CCZ₄* formalism adds additional auxiliary fields that act as damping fields which suppress the Hamiltonian and momentum constraint equations.

The idea of a constraint suppression mechanism can be largely understood using a simple example. Consider the trivial system

$$\partial_t \lambda = 0 \tag{3.72}$$

which is subject to the constraint $\lambda = 0$ for all time. Suppose at a later time, $\lambda = k \neq 0$, possibly due to numerical errors, mesh refinement effects, or other origins. Numerically evolving eq. 3.72 past this point means the violation of the constraint will remain constant throughout the rest of the simulation. Instead, we could add a constraint damping term to the right-hand side of eq. 3.72

$$\partial_t \lambda = -\kappa \lambda \tag{3.73}$$

for $\kappa > 0$. If λ becomes non-zero at some point in the evolution, the constraint damping term will push the solution back to the constraint 'surface' $\lambda = 0$. The constraint surface becomes an *attractor* in the solution space.

In the spirit of this, the *CCZ₄* formulation of the Einstein equations was introduced which includes a new auxiliary field that is responsible for damping away violations of the constraints. The modified Einstein equations take the form

$$\mathcal{R}_{\mu\nu} + 2\nabla_{(\mu} Z_{\nu)} - \kappa_1 (2n_{(\mu} Z_{\nu)} - (1 + \kappa_2) g_{\mu\nu} n^\alpha Z_\alpha) = 8\pi \left(\mathcal{T}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{T} \right). \tag{3.74}$$

This new vector field is assumed to vanish on physical solutions, which amounts to adding a new

constraint $Z_\mu = 0$. This form of the Einstein equations implies an evolution equation for Z ,

$$\nabla^\nu \nabla_\nu Z_\mu + \mathcal{R}_{\mu\nu} Z^\nu - \kappa_1 \nabla^\alpha (2n_{(\alpha} Z_{\mu)}) + \kappa_2 g_{\mu\alpha} n^\beta Z_\beta = 0, \quad (3.75)$$

which takes the form of a generalized telegraph equation, implying the Z field propagates as an exponentially damped wave at the speed of light³³². A careful analysis shows that for the Z field to be properly damped, $\kappa_1 > 0$ and $\kappa_2 > -1$. The requirement that this system be strongly hyperbolic is simply that one uses the densitized lapse $\alpha/\sqrt{\gamma}$ or a Bona-Masso slicing gauge condition. Notably, the momentum constraint is no longer required.

This approach has several advantages. The first is obviously that the constraints are damped and violations propagate at the speed of light outside the computational domain. This is a highly desirable trait since numerical schemes inevitably introduce some amount of error, whether it's finite differencing related, boundary related, or from some other source. The damping of these violations increases the confidence that the simulation data is closer to the constraint surface than they would have otherwise been. While in principle starting with initial data that possesses some amount of constraint violation is permissible, relying on the damping terms to reduce the violations during the evolution is discouraged. This is due to the fact that the flow of the solution towards the constraint surface may converge to a solution not intended during the initial data construction.

The CCZ4 formalism continues towards decomposing the covariant equations of motion eq. 3.74. The decomposition follows closely to the BSSN approach, however, different auxiliary functions are chosen

$$\hat{\Gamma} = \tilde{\Gamma}^i + \frac{2}{\chi} Z^i, \quad (3.76)$$

where the Z^i are the spatially projected version of the four vector Z^μ , $Z^i = \gamma^i_\mu Z^\mu$. The modification of these auxiliary connection functions changes the dynamical variables compared to BSSN. Now, the dynamical variables are $\{\chi, \tilde{\gamma}_{ij}, \mathcal{K}, \tilde{A}_{ij}, \Theta, \hat{\Gamma}^i\}$, where $\Theta = -n_\mu Z^\mu$, the component of the auxiliary damping field normal to the hypersurfaces. This formulation results in the evolution equations 3.77-3.82. In order for this system to be strongly hyperbolic, the combination $\mathcal{R} + 2D_i Z^i$

$$\partial_t \tilde{\gamma}_{ij} = -2\alpha \tilde{A}_{ij} + 2\tilde{\gamma}_{k(i}\partial_{j)} \beta^k - \frac{2}{3}\tilde{\gamma}_{ij}\partial_k \beta^k + \beta^k \partial_k \tilde{\gamma}_{ij} \quad (3.77)$$

$$\begin{aligned} \partial_t \tilde{A}_{ij} = & \chi [-D_i D_j \alpha + \alpha (\mathcal{R}_{ij} + D_i Z_j + D_j Z_i - 8\pi\alpha S_{ij})]^{\text{TF}} + \alpha \tilde{A}_{ij} (\mathcal{K} - 2\Theta) \\ & - 2\alpha \tilde{A}_{il} \tilde{A}_j^l + 2\tilde{A}_{k(i}\partial_{j)} \beta^k - \frac{2}{3}\tilde{A}_{ij}\partial_k \beta^k + \beta^k \partial_k \tilde{A}_{ij} \end{aligned} \quad (3.78)$$

$$\partial_t \chi = \frac{2}{3}\alpha \chi \mathcal{K} - \frac{2}{3}\chi \partial_k \beta^k + \beta^k \partial_k \chi \quad (3.79)$$

$$\begin{aligned} \partial_t \mathcal{K} = & -D^i D_i \alpha + \alpha (\mathcal{R} + 2D_i Z^i + \mathcal{K}^2 - 2\Theta \mathcal{K}) + \beta^j D_j \mathcal{K} - 3\alpha \kappa_1 (1 + \kappa_2) \Theta \\ & + 4\pi\alpha (S - 3\rho) \end{aligned} \quad (3.80)$$

$$\begin{aligned} \partial_t \Theta = & \frac{1}{2}\alpha \left(\mathcal{R} + 2D_i Z^i - \tilde{A}_{ij} \tilde{A}^{ij} + \frac{2}{3}\mathcal{K}^2 - 2\Theta \mathcal{K} \right) - Z^i \partial_i \alpha + \beta^k D_k \Theta \\ & - \alpha \kappa_1 (2 + \kappa_2) \Theta - 8\pi\alpha \rho \end{aligned} \quad (3.81)$$

$$\begin{aligned} \partial_t \hat{\Gamma}^i = & 2\alpha \left(\tilde{\Gamma}_{jk}^i \tilde{A}^{jk} - \frac{3}{2} \tilde{A}^{ij} \frac{\partial_j \chi}{\chi} - \frac{2}{3} \tilde{\gamma}^{ij} \partial_j \mathcal{K} \right) + 2\tilde{\gamma}^{ki} \left(\alpha \partial_k \Theta - \Theta \partial_k \alpha - \frac{2}{3} \alpha \mathcal{K} Z_k \right) - \\ & 2\tilde{A}^{ij} \partial_j \alpha + \beta^k \partial_k \hat{\Gamma}^i + \tilde{\gamma}^{kl} \partial_k \partial_l \beta^i + \frac{1}{3} \tilde{\gamma}^{ik} \partial_k \partial_l \beta^l + \frac{2}{3} \tilde{\Gamma}^i \partial_k \beta^k - \tilde{\Gamma}^k \partial_k \beta^i + \\ & 2\kappa_3 \left(\frac{2}{3} \tilde{\gamma}^{ij} Z_j \partial_k \beta^k - \tilde{\gamma}^{jk} Z_j \partial_k \beta^i \right) - 2\alpha \kappa_1 \tilde{\gamma}^{ij} Z_j - 16\pi\alpha \tilde{\gamma}^{ij} p_k \end{aligned} \quad (3.82)$$

must be computed as

$$\begin{aligned} \mathcal{R}_{ij} + 2D_{(i} Z_{j)} = & -\frac{1}{2} \tilde{\gamma}^{mn} \partial_m \partial_n \tilde{\gamma}_{ij} + \tilde{\gamma}^{mn} \left(2\tilde{\Gamma}_{m(i}^k \tilde{\Gamma}_{j)kn} + \tilde{\Gamma}_{im}^k \tilde{\Gamma}_{kjn} \right) \\ & + \tilde{\gamma}_{m(i}\partial_{j)} \hat{\Gamma}^m + \tilde{\Gamma}^m \tilde{\Gamma}_{(ij)m} + \frac{1}{2\chi} \left(\tilde{D}_i \tilde{D}_j \chi - \frac{1}{2\chi} \partial_i \chi \partial_j \chi \right) \\ & + \frac{1}{2\chi} \tilde{\gamma}_{ij} \left(\tilde{\gamma}^{mn} \tilde{D}_m \tilde{D}_n \chi - \frac{3}{2\chi} \tilde{\gamma}^{mn} \partial_m \chi \partial_n \chi \right). \end{aligned}$$

3.2.4 GAUGE CHOICES

With the evolution equations in hand, there is one remaining piece before turning to numerical solvers and that is the choice of gauge. This choice is vital to ensure stability of simulations, *especially* in the presence of singularities. Consider the evolution of the spatial metric, given by eq. 3.24. Fixing the spatial co-ordinates to x^i , then

$$\partial_t \ln(\sqrt{\gamma}) = -\alpha \mathcal{K} + D_k \beta^k. \quad (3.83)$$

Using the choice $\alpha = 1$ and $\beta^k = 0$, then the evolution of the $3+1$ Einstein equations yields

$$\partial_t \mathcal{K} = K_{ij} K^{ij} + 4\pi(\rho + S) . \quad (3.84)$$

The first term is positive-definite and the second is at least non-negative, implying that the volume element shrinks to vanishing values.

The solution to this issue is by finding an alternative choice of gauge. Intense effort was put into this search, finally arriving at the *Bona-Masso* family of gauge conditions³³³

$$(\partial_t - \mathfrak{L}_\beta) \alpha = -\alpha f(\alpha) \mathcal{K} \quad (3.85)$$

with $f(\alpha)$ an arbitrary function of the lapse. Substituting the spatial metric evolution eq. 3.24 into the above equation and choosing $f(\alpha) = \frac{2}{\alpha}$ yields

$$\partial_t \alpha = 2\partial_t \log(\sqrt{\gamma}) , \quad (3.86)$$

which can be directly integrated to yield

$$\alpha = g(x) + \log(\gamma) . \quad (3.87)$$

This choice of $f(\alpha)$ is commonly known as $1+\log$ slicing. It turns out this choice of slicing has very strong singularity avoidance properties, making it a robust choice for evolving spacetimes containing singularities³³⁴. In the CCZ4 formulation, the $1+\log$ slicing becomes

$$\partial_t \alpha = \beta^k \partial_k \alpha - 2\alpha(\mathcal{K} - 2\Theta) . \quad (3.88)$$

This is not enough for stable evolution, since the choice of the shift vectors needs to be gauge fixed as well. It turns out the choice $\beta^i = 0$ is insufficient as well. Instead, the most common choice is

the so-called *Gamma-driver* condition, which amounts to the choice

$$\begin{aligned}\partial_t \beta^i &= b_1 \beta^k \partial_k \beta^i + b_2 B^i \\ \partial_t B^i &= b_1 \left(\beta^k \partial_k B^i - \beta^k \partial_k \hat{\Gamma}^i \right) + \partial_t \hat{\Gamma}^i - \eta B^i ,\end{aligned}$$

where b_1 , b_2 , and η are specifiable parameters. The choice implemented in codes such as GRChombo is $b_1 = 1$, $b_2 = \frac{3}{4}$, and $\eta \sim 1/M$ for simulations of black holes of mass M .

The combination of Γ +log slicing and the Gamma driver condition is known as the *moving puncture gauge* as it permits black hole singularities to traverse the computational grid without the need for excision or extra care. For this very reason, this is a common gauge choice when dealing with black holes and is the most common choice of gauge for codes such as GRChombo.

With these gauge choices and evolution formalisms in hand, one can finally turn to generating computer code and running simulations.

3.3 EVOLVING GENERALIZED PROCA

This section discusses the intricate details of a new numerical relativity code developed during this thesis titled GRBoondi. The code is entirely dedicated towards numerical evolutions of systems containing generalized Proca fields. In other words, GRBoondi is used to solve eqs. 3.1-3.6.

The inception of GRBoondi came about due to several difficulties arising from computing the evolution of the Proca field using full numerical relativity in GRChombo. Primarily, a catastrophic result prevented long-term evolution of the Proca field in a dynamical spinning black hole background: The black hole background was unstable to long-term evolution. Even with a proper choice of gauge and using the best initial data that's actively used in black hole evolutions, the horizon of the black hole had trouble evolving past several hundred simulation clock cycles. For a superradiant Proca cloud, this is extremely far from sufficient. Typical timescales are on the order of thousands, or even tens of thousands, of clock cycles. Hence, one of the central projects of the thesis had a catastrophic problem. Therefore the attention was turned to evolving the Proca field on a fixed background. At first, this was heavily avoided, since the initial goal was to derive results for generalized Proca fields in the full Einstein-generalized Proca theory. However, after facing the

show-stopping problem in the full simulation, closer inspection of the fixed background problem revealed that the fixed background approximation was quite acceptable. This is actually a permissible approximation for these types of studies since the Proca field is largely a perturbation on the background. Hence the errors in the Einstein equation, whose magnitude is roughly proportional to the *square* of the Proca field, are otherwise small, even for large superradiant Proca clouds. Moreover, it was quickly realized that studying several different generalized Proca theories all resulted in roughly the same computer code. Hence, GRBoondi was developed, allowing for rapid reusability of huge parts of the computer code, even for more complex generalized Proca theories.

This section will be split into several parts. Subsection 3.3.1 spells out the fundamental pieces of GRBoondi. In particular, the structure of the code, the discretization procedure, dissipation algorithms, and available boundary conditions will be elucidated. Subsection 3.3.2 will be dedicated to discussing the complexities of adaptive mesh refinement. Subsection 3.3.3 will discuss the parallelization procedures utilized by GRBoondi, an important piece in developing highly scalable simulations. Subsection 3.3.4 discusses diagnostic quantities and their extraction from the computational grid. This is crucial in extracting useful information from the simulations and especially in studying the superradiant process, where knowing the total energy of the Proca cloud is vital. Subsection 3.3.5 finally discusses how GRBoondi can incorporate arbitrary modifications to the base Proca theory, allowing any generalized Proca model to be studied.

3.3.1 FUNDAMENTAL PIECES

There are many pieces and active parts in numerical relativity codes, making development of an accurate and precise simulator a challenging task. Fortunately, many libraries have been developed over the years that provide comprehensive functions, allowing developers to relegate challenging methods to more established software libraries. These libraries are heavily utilized by GRBoondi. We describe these fundamental methods, concepts, and libraries here.

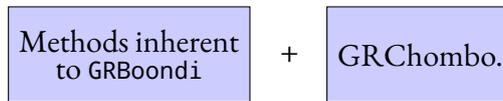
STRUCTURE OF CODE

From the outset, GRBoondi aims to be two things: 1) A highly modular codebase and 2) easy to use. This means many low-level function methods and algorithms are entrusted to existing libraries.

Two of the most widely used ones are:

- *Chombo*³³⁵: a set of tools for implementing finite differencing and finite volume methods for solving generic PDEs on block-structured adaptively refined grids.
- *GRChombo*^{239,240}: a numerical relativity code for solving Einstein’s equations, which allows matter fields to be mutually evolved.

Since GRChombo is based off the Chombo library itself, the Chombo methods are not directly utilized within GRBoondi, instead only indirectly through its use of GRChombo. Because of this inheritance, the structure of GRBoondi is basically of the form



Many of the grid generation, refinement methods, checkpointing, and other methods come directly from GRChombo, while the methods for evolving the Proca evolution equations, setting up appropriate level classes, and ability to incorporate additional terms in the generalized Proca Lagrangian are in GRBoondi.

It’s instructive to use an example in order to explain the basic mechanics of running a simulation using GRBoondi. To this end, the problem of simulating a standard Proca field on a Kerr background is used, whose code is incorporated into the GRBoondi Github repository and henceforth will be called the *Proca example*. The fundamental class structure of the Proca example is shown in fig. 3.3. The central orchestrating class is the AMR class, which is a parent class to GRAMR. The AMR class stores pointers to AMRLevel classes, which contain methods and information about each refinement level of the grid. Each AMRLevel class corresponds directly to a single refinement level. At runtime, the AMR class instantiates the grid, generating each level class, and then finally executing the evolution. The AMR class is thus the orchestrator of the entire simulation. It generates checkpoint files and, as the grid orchestrator, is responsible for the regridding procedure. While the AMR class manages the overall mechinations of the simulation, it doesn’t know how to perform finer-grained tasks, instead calling methods from each level class and their children. For example, after each timestep, the AMR class calls the `GRAMRLevel1::postTimeStep()` method, which executes functions after each timestep on the level that that GRAMRLevel1 instance points to.

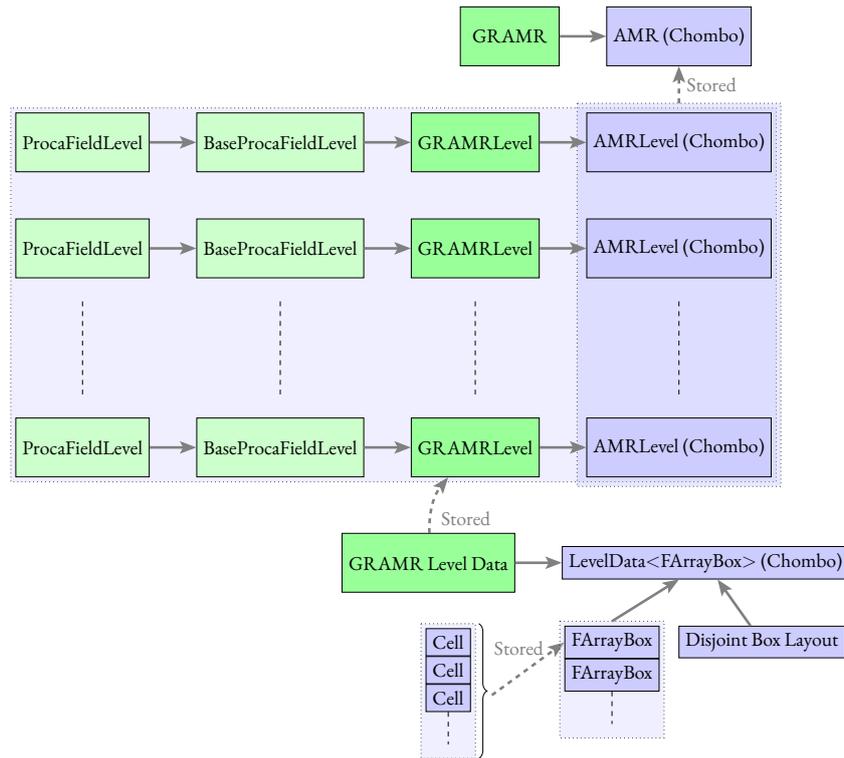


Figure 3.3: Diagrammatic representation of the class structure and hierarchy for the Proca example. The lilac rectangles are classes from the Chombo library, the green rectangles are GRChombo classes, and the light green boxes are GRBoondi classes.

The GRAMR class is a child of the full AMR class and handles generic methods from GRChombo that need access to the full refinement hierarchy, for example, a pointer to the interpolator object for interpolating data on the grid. It can also handle filling in ghost cells across multiple levels (see sec. 3.3.2) or accessing the GRAMRLevel children of the AMRLevel instances. Users could add another level to the hierarchy by inheriting from GRAMR and defining their own AMR orchestrator. For example, in GRChombo, there is the BHAMR which handles apparent horizon locating and tracking the black hole puncture. Users of GRBoondi could easily incorporate similar inheritances.

Each refinement level in the grid is represented programmatically by the ProcaFieldLevel class, itself a great-grandchild of the AMRLevel class, grandchild of the GRAMRLevel class, and child of the BaseProcaFieldLevel class. The AMRLevel class contains information about the levels above and below it, and virtualizes methods such as setting up initial data, performing the regridding on its level, and advancing the state of the simulation on its level. These methods are then defined by the child of the class, in this case GRAMRLevel. The GRAMRLevel class does the brute force work on its level. It executes code that advances the levels state, carries out the process of regridding, and executes the functions that define the initial data. The BaseProcaFieldLevel class inherits from

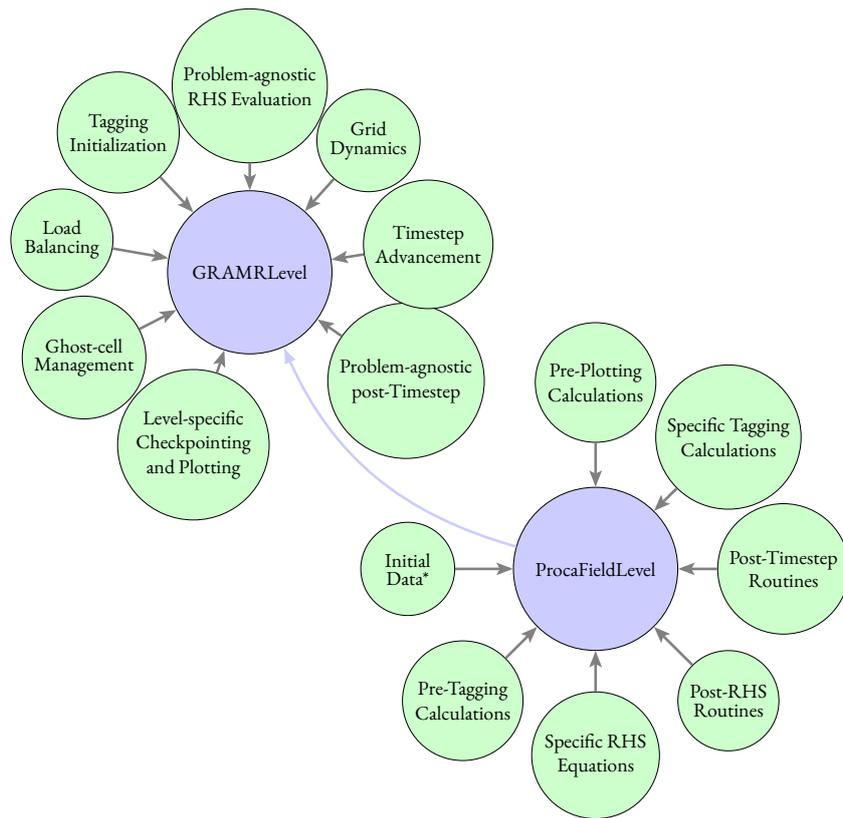


Figure 3.4: Diagrammatic representation of the level methods and level class inheritance of GRBoondi. Starred methods in the ProcaFieldLevel planet are required to be set by the user. All other methods are virtual and have default definitions which can be freely modified by the user.

GRAMRLevel. It contains almost all of the methods required for evolving the Proca field on its level, including computing diagnostic quantities for plotting, executing the time-stepping code, and integrating quantities across the level. The methods of BaseProcaFieldLevel are catered towards generalized Proca systems. They use the variables and data specifically designed for evolving such fields. The ProcaFieldLevel class is entirely problem specific and must be defined by the user. This is the final level of the level class hierarchy. It must have at least one method defined, which is the initial data method — it must define exactly how to compute the initial data. Beyond that, additional methods are available for use and are inherited from BaseProcaFieldLevel. These include additional steps to be performed in the BaseProcaFieldLevel::postTimeStep or in the BaseProcaField::PrePlotLevel execution. Users can even overwrite methods inherited from BaseProcaField, such as the specific time-stepping code, since the methods have been virtualized using c++14 specifiers.

DISCRETIZATION AND TIME-STEPPING

GRBoondi evolves the equations using the *Method of Lines*, which separates the differencing into time and space. As an example, consider the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} . \quad (3.89)$$

Then let $\Delta^2 u$ represent a discretized version of the second partial derivative $\Delta^2 u \approx \frac{\partial^2 u}{\partial x^2}$. In particular, $\Delta^2 u$ is determined by neighboring points and is no longer a derivation. This transforms the heat equation into

$$\frac{\partial u}{\partial t} = \Delta^2 u , \quad (3.90)$$

which is now an ordinary differential equation that approximates the PDE eq. 3.89 and can be solved using standard ordinary differential equation techniques.

GRBoondi uses the fourth order Runge-Kutta (RK4) method for computing the temporal integrals. To illustrate the RK4 procedure, consider an arbitrary IVP of the form

$$\frac{dy}{dt} = f(t, y) \quad (3.91)$$

with initial data $y(t_0) = y_0$, $y(t)$ is the yet-to-be-determined function, and $f(t, y)$ is an arbitrary function of time and the variable y itself. Then, with a step size of $\Delta t = h > 0$, four substeps are computed

$$\begin{aligned} k_1 &= f(t_n, y_n) \\ k_2 &= f\left(t_n + \frac{h}{2}, y_n + h\frac{k_1}{2}\right) \\ k_3 &= f\left(t_n + \frac{h}{2}, y_n + h\frac{k_2}{2}\right) \\ k_4 &= f(t_n + h, y_n + hk_3) . \end{aligned}$$

For a given n , the solution one timestep later is computed as

$$y_{n+1} = y_n + \frac{h}{6} (k_1 + 2k_2 + 2k_3 + k_4) . \quad (3.92)$$

So long as the original IVP is well-posed, then so is the RK4 method³³⁶. The simplicity and stability of the RK4 method makes it an extremely popular choice for solving finite difference problems. Additionally, the accuracy of the integration can be dramatically increased simply by adding more substeps, resulting in the so-called explicit Runge-Kutta method³³⁷. Moreover, the total accumulated error from the approximation scales as $\mathcal{O}(h^4)$, making it easy to substantially increase accuracy by simply reducing the time step.

For the discretization itself, the continuous spatial and temporal directions are replaced by finite versions up to some boundary. This means the spacetime is truncated to finite spatial and temporal extent and the continuous surface is replaced by a mesh. For simplicity, this mesh is here assumed to be uniform in all spatial and temporal directions, however the real mesh setup will be discussed in section 3.3.2, where an adaptive mesh routine is discussed.

The computational domain then becomes $\mathcal{D} = [x_{min}, x_{max}] \cup [0, t_{max}] \in \mathbb{R}^2$ and the grid is defined by

$$x_j = x_{min} + j\Delta x \quad (3.93)$$

$$t_n = n\Delta t \quad (3.94)$$

for $j = 0, \dots, J$ and $n = 0, \dots, N$, with $J = \frac{x_{max}-x_{min}}{\Delta x}$ and $N = \frac{t_{max}}{\Delta t}$. The evolution variables then become functions over this mesh $Q(x, t) \rightarrow Q_m^n = Q(x_m, t_n)$. Spatial derivatives are computed by using *stencils*. GRBoondi uses fourth-order centered stencils, which means derivatives are computed using the following formulas³³⁸

$$\begin{aligned} \partial_x Q &= \frac{1}{12h} (Q_{i-2} - 8Q_{i-1} + 8Q_{i+1} - Q_{i+2}) \\ \partial_x^2 Q &= \frac{1}{12h^2} (-Q_{i-2} + 16Q_{i-1} - 30Q_i + 16Q_{i+1} - Q_{i+2}) \\ \partial_x \partial_y Q &= \frac{1}{144h^2} (Q_{i-2,j-2} - 8Q_{i-2,j-1} + 8Q_{i-2,j+1} - Q_{i-2,j+2} - 8Q_{i-1,j-2} + 64Q_{i-1,j-1} \\ &\quad - 64Q_{i-1,j+1} + 8Q_{i-1,j+2} + 8Q_{i+1,j-2} - 64Q_{i+1,j-1} + 64Q_{i+1,j+1} \\ &\quad - 8Q_{i+1,j+2} - Q_{i+2,j-2} + 8Q_{i+2,j-1} - 8Q_{i+2,j+1} + Q_{i+2,j+2}) , \end{aligned}$$

where $Q_{i,j}$ is the value of the evolution variable at grid point (i, j) . The x and y symbols are place-

holders for any of the three spatial co-ordinates. For computing so-called advection terms, which are of the form $\beta^i \partial_i Q$, different stencils are required in order to maintain stability

$$\partial_x Q = \begin{cases} \frac{1}{12h} (-3Q_{i-1} - 10Q_i + 18Q_{i+1} - 6Q_{i+2} + Q_{i+3}), & \text{if } \beta^x > 0 \\ \frac{1}{12h} (-Q_{i-3} + 6Q_{i-2} - 18Q_{i-1} + 10Q_i + 3Q_{i+1}), & \text{if } \beta^x < 0 \end{cases}. \quad (3.95)$$

Sixth order stencils are available as well, though these typically aren't necessary.

KREISS-OLIGER DISSIPATION

An issue that arises from using finite methods to compute systems of PDEs is the appearance of high frequency noise, which is especially true for multi-level refinement hierarchies and the use of regridding. These can arise from reflections off level boundaries or simply from the spacing between grid points. Even worse, these high frequencies can develop amplitudes which grow very fast. Hence, it's important to impose a scheme that can deal with these high frequency (unphysical) modes and preserve the low frequency (physical) ones. This is achieved by implementing an $N = 3$ Kreiss-Oliger (KO) dissipation term³³⁹. The basic idea of the Kreiss-Oliger dissipation term is to add an additional term to the evolution equations that damps these unphysical high frequency modes. Consider a finite difference scheme that schematically is of the form

$$u_m^{n+1} = u_m^n + \Delta t S(u_m^n), \quad (3.96)$$

where $S(u^n)$ is some spatial finite differencing operator. To this evolution scheme, an additional term is added

$$u_m^{n+1} = u_m^n + \Delta t S(u_m^n) - \epsilon \frac{\Delta t}{\Delta x} (-1)^N \Delta_x^{2N}(u_m^n) \quad (3.97)$$

with $\epsilon > 0$, $N \geq 1$ an integer, and where Δ_x^{2N} is the $2N$ centered difference operator. In GRBoondi, the $N = 3$ KO dissipation term is used, adding the term

$$\frac{\sigma}{64\Delta x} (Q_{i-3} - 6Q_{i-2} + 15Q_{i-1} - 20Q_i + 15Q_{i+1} - 6Q_{i+2} + Q_{i+3}) \quad (3.98)$$

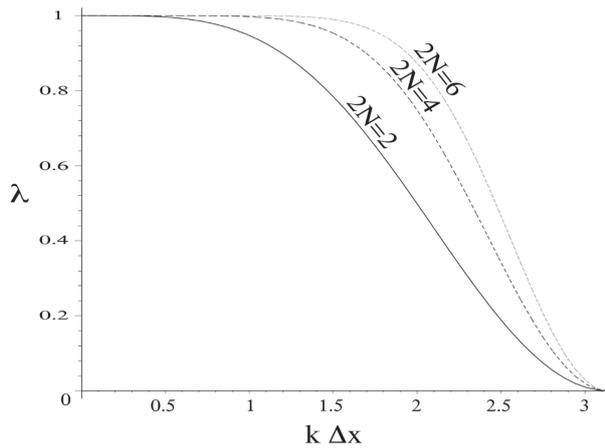


Figure 3.5: ³²⁶ Dissipation of modes as a function of wavenumber $k\Delta x$. Higher frequencies correspond to lower values of $k\Delta x$, while lower frequencies correspond to higher values of $k\Delta x$. This shows that higher orders of the KO dissipation approach a step function, meaning the dissipation strongly dampens frequencies higher than the inverse of the grid spacing and leaves lower frequencies, the physical ones, virtually untouched.

to the evolution of each variable Q , where σ is the so-called Kreiss-Oliger coefficient. In the continuum limit, this term vanishes. This term does exactly what is desired, it damps high frequency modes corresponding to the grid spacing, and leaves the lower frequencies minimally affected, fig. 3.5. A stability analysis shows that this scheme is stable as long as

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

where $\alpha_C = \frac{\Delta t}{\Delta x}$ is the so-called *Courant factor*

and has a typical value of $\alpha = 0.25$. The dissipation term should be tuned based on the problem at hand. A typical value of $\sigma = 0.3$ is the default, however increasing it can improve long term stability. Contrarily, increasing it too much runs into the upper bound set by the stability condition. A value that is too high can result in a checkboard type of pattern developing in the simulation, spoiling it. This problem was encountered periodically during some simulations. The remedy is simply to slightly reduce the coefficient.

BOUNDARY CONDITIONS

An important piece in choosing the conditions of a simulation is picking the right boundary conditions at the edge of the computational space. Broadly speaking, there are three possible ways to choose a boundary condition:

- Modify the evolution equations close to the boundary;
- Add additional cells beyond the boundary;
- Use different derivative stencils close to the boundary.

The appropriate approach depends on the physical system under consideration, for example, if the system has asymptotic waves, a uniform but time varying value at the boundary, or is asymptotically vanishing.

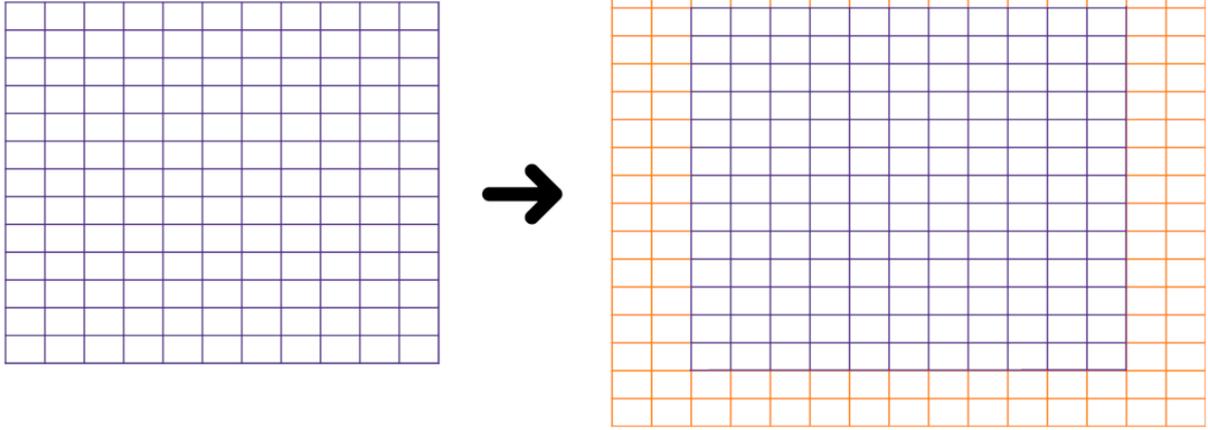


Figure 3.6: Example of adding ghost cells to the outer boundary of the numerical grid. These allow the usual derivative stencils to be used at the outer boundary of the computational box.

A particularly simple example is that of *periodic boundary conditions*. Periodic boundary conditions are enforced by adding in so-called *ghost cells* outside the domain whose values are set to the value of the fields on the opposite side of the box. In other words, for a two-dimensional system in a centered square of side length a , the values of the derivatives are calculated using the standard stencils via

$$\begin{aligned} \frac{\partial^p}{\partial x_1^p} \phi(a/2, x_2) &= \frac{\partial^p}{\partial x_1^p} \phi(-a/2, x_2) \\ \frac{\partial^p}{\partial x_2^p} \phi(x_1, a/2) &= \frac{\partial^p}{\partial x_2^p} \phi(x_1, -a/2) , \end{aligned} \quad (3.99)$$

where the stencils use ghost cells for cells that lie on the exterior of the boundary. Periodic boundary conditions are not typically astrophysically relevant, although some use cases include large cosmological simulations (motivated by the cosmological principle) and other large computational scenarios. More physically relevant to the cases in GRBoondi are *radiative conditions*, or *Sommerfeld radiation conditions*. This is a very common condition in numerical relativity, since it ensures that outgoing waves do not get reflected back into the computational space at the boundary. The conditions on the evolution variables assume an outgoing spherical wave, so the evolution variables behave as

$$u \sim u_{r \rightarrow \infty} + \frac{f(t-r)}{r} \quad (3.100)$$

in the limit that $r \rightarrow \infty$, where f is an arbitrary function and $u_{r \rightarrow \infty}$ is a constant that is added for

variables that asymptote to non-zero values. This can be enforced at the derivative level by rewriting eq. 3.100 as

$$\frac{\partial Q}{\partial t} = -\frac{x^i}{r} \partial_i Q - \frac{Q - Q_{r \rightarrow \infty}}{r} \quad (3.101)$$

for the evolution variable Q . This works well for linear wave equations, but in the vast majority of cases relevant to GRBoondi, this will no longer be true. Nonetheless, this is an adequate condition so long as it is applied far enough away from the highly dynamical and nonlinear parts of the simulation.

Extrapolation is another boundary condition that sees some use. At the outer boundary of the computational box, the evolution variables are assumed to satisfy

$$Q = a + b * R , \quad (3.102)$$

where R is the radius from the user-defined center, and a and b are constants. For zeroth order extraction, $b = 0$ and a is the value of the field in the cell closest to the boundary. For first order extraction, the two coefficients are computed using the two outermost non-boundary cells. This boundary condition is particularly useful when the field is spatially uniform, or approximately so, at the boundary but has some time variance.

Reflection symmetry is a particularly useful boundary condition since it can be used to dramatically reduce the size of the computational box, vastly speeding up simulation time. Across a reflective boundary, the evolution variable is assumed to be either odd or even. For example, if the reflective boundary is the $x = 0$ plane, then the evolution variable is assumed to satisfy

$$Q(x, y, z) = \pm Q(-x, y, z) , \quad (3.103)$$

where the choice of sign depends on the nature of the variable and is user-defined. This condition sees use in highly symmetric systems, such as Schwarzschild or Kerr spacetimes. For example, the example in GRBoondi of a Proca field on a Kerr background uses the symmetry in the $z = 0$ plane to reduce the size of the computational box by $\frac{1}{2}$.

The last boundary condition is a static boundary condition, which fixes the value of the evolu-

tion variable to its initial data by imposing

$$\partial_t Q = 0 . \quad (3.104)$$

GRBoondi also supports mixing various boundary conditions. For example, in the Proca example, the symmetry across the $z = 0$ plane is enforced while the Sommerfeld boundary condition is enforced at the outer edges of the $z > 0$ computational domain. This allows the outgoing radiation condition to be enforced while also leveraging the z-symmetry to reduce the computational domain size, substantially reducing computation time.

3.3.2 ADAPTIVE MESH REFINEMENT

The computational grid is a core part of setting up a simulation and choosing the right type of grid is highly problem dependent, but also very important for extracting accurate data. GRBoondi uses the Berger-Rigoutsos (BR) adaptive mesh refinement (AMR) implementation from Chombo. This grid is built from a hierarchy of increasingly fine levels, called *refinement levels*, which are labeled $l = 0, \dots, l_{max}$ and are of the form

$$\Delta x_l = \Delta x_0 / 2^l , \quad (3.105)$$

where Δx_0 is the spatial size of the coarsest cells, i.e. the size of the cells on the coarsest level of the hierarchy. The BRAMR algorithm uses a block-structured approach, meaning each level of the hierarchy is split into variably sized boxes, which are then distributed between CPUs, see sec. 3.3.3.

The allowable boxes are constrained by two conditions:

- *Proper Nesting*: a box on the l^{th} refinement level must not touch level $(l - 2)$ directly. Instead, there must be at least one intermediate $(l - 1)$ -level grid cell in between.
- *Proper Refinement*: a box on level l must not refine parts of a level $(l - 1)$ grid cell. It must refine it completely or not at all.

At regridding or initial generation of the grid, on a single refinement level l , cells are flagged for refinement according to a *tagging criterion* $\tau = \tau(I)$. For a given cell with indices $I = (i, j, k)$

and Cartesian co-ordinates $X_I = (x, y, z)$, if the tagging criterion computed for that cell is above a user-defined threshold, $\tau(I) > \tau_U$, then that cell is flagged for refinement and its resolution increased by a factor of 2 once the refinement stage is reached*.

In block-structured AMR, a challenge commonly faced is how to properly and efficiently partition cells that need refinement into boxes. GRBoondi, being a derivative of GRChombo, uses the BR algorithm for this purpose. A *block factor* is user-defined and enforces the minimum number of cells that can partition the length of a box. Additionally, there is a user-defined maximum box size. Typically, these two parameters are set equal to each other so that all boxes are roughly the same size. In order to enforce the block factor on a newly refined level, Chombo generates a temporary *coarser* level built from a set of tagged cells whose size corresponds to the block factor on the newly refined level. Chombo then applies the BR partitioning algorithm on the new virtual coarse level to construct boxes of grids which obey the user-specified block factor and maximum box size. The size of the box is typically chosen with the number of CPU threads in mind, since the cells of each box are iterated through using the available CPU threads (OpenMP; see sec. 3.3.3).

The BRAMR algorithm is reviewed here, since it is a central component of the GRBoondi foundation³⁴⁰. The algorithm itself follows four broad tenets:

- There should be as little unnecessarily refined area as possible.
- There should be as few boxes as possible.
- The boxes should 'fit' the data.
- The algorithm should be fast.

The idea is to find the minimum box size that encloses all tagged cells on each level. Define the *tagging indicator function* as

$$T(I) = \begin{cases} 1 & \tau(I) > \tau_U \\ 0 & \textit{else} \end{cases} . \quad (3.106)$$

*The factor of 2 is known as the *ref ratio* and is the resolution difference between levels. In other words, it's the ratio of the size length of the cells on level $l - 1$ to that of cells on level l .

In each partition, the *signatures* of the tagging function of any given box are defined as

$$X(x) = \int T(I)dydz \quad (3.107)$$

$$Y(y) = \int T(I)dx dz \quad (3.108)$$

$$Z(z) = \int T(I)dx dy . \quad (3.109)$$

Given these signatures, the Laplacian of the signatures are also computed, $\partial_x^2 X(x)$, $\partial_y^2 Y(y)$, and $\partial_z^2 Z(z)$. Given these results, the BR algorithm searches for all (if any) inflection points individually for each direction. Finally, the one whose difference $\delta(\partial_i^2 X_i)$ is the greatest becomes the line of partitioning for this particular dimension. This roughly corresponds to the line between tagged and untagged cells in the orthogonal directions of the signature. If there exists a point with zero signature, then this 'hole' is chosen to be the line of partition instead. If there are no holes or inflections, then the box is split along the midpoint.

After partitioning, the new partition is checked for *efficiency*, by checking whether the proportion of tagged cells to all cells in the partition exceeds a user-defined `fill ratio` ϵ and that the boxes are within the block factor and maximum box size requirements. If the tests are passed, then the partition is accepted. If not, then the boxes are recursively partitioned, discarding any boxes that don't meet the requirements. The value of ϵ naturally plays a role in the computational efficiency. A higher value of ϵ leads to a greater ratio of tagged to untagged cells, making for a more efficient partitioning. However, this is not always *computationally* efficient since there will be greater overhead for the increased number of boxes, especially since each box needs to be allocated to an MPI instance (more on this in sec. 3.3.3). There could also be greater fluctuation in the overall structure of the grid, leading to more noise generation. Likewise, a smaller value of ϵ is more computationally efficient but produces a less efficient partitioning. The default value is `fill ratio` = 0.7, though the optimal value is case and processor dependent.

The partitioned box is finally refined, its constituent cells split into a finer mesh using the refinement ratio $\frac{\delta x^{l+1}}{\delta x^l}$. This process is recursively performed until there are no more tagged cells, or the top of the refinement hierarchy is reached, l_{max} . On each new level, the data is interpolated from the coarser level using fourth-order interpolation.

One now needs to specify a prescription for tagging cells that need to be refined. GRBoondi tags cells based on expressions involving user-selected fields, $F(Q)$. When any one of the expressions passes a user-defined global threshold $\sigma(Q)$, then the cell is tagged for regridding

$$\tau(I) = \begin{cases} 1 & F(Q(I)) > \sigma(Q) \\ 0 & \text{else} \end{cases} . \quad (3.110)$$

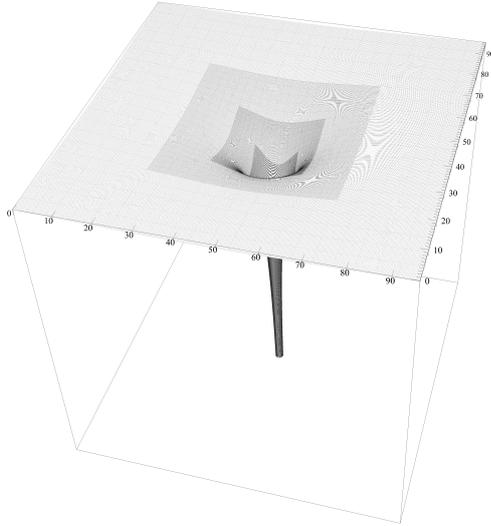


Figure 3.7: An example of a mesh generated using the BRAMR algorithm. The mesh is overlaid on a two-dimensional extruded plot of the conformal factor for a Kerr black hole spacetime. The various refinement levels are clearly visible, showing areas of strong gravity are sufficiently resolved.

The rate at which regridding is performed is user-defined. In an ideal world, regridding would occur at every timestep. However, this is usually not the best approach. One should consider that it's important to let numerical errors dissipate before remeshing. Moreover, it's more computationally efficient to not regrid very often. In some cases, it may be best to turn off regridding completely and evolve the system on a static hierarchy of levels. This is the case in the Proca on a Kerr black hole example. The highest levels of the hierarchy capture almost all of the physics, which do not deviate from the center very much, thus making a fixed grid the best choice.

Fig. 3.7 shows an example of a computational grid generated using the BRAMR algorithm. Since the finer levels have a smaller Courant number, each mesh levels timestep is appropriately reduced according to

$$\Delta t^{l+1} = \Delta t^l \frac{\Delta x^{l+1}}{\Delta x^l} . \quad (3.111)$$

This means the grids on level $l + 1$ will take two timesteps for each timestep taken on level l . The entire grid is then evolved according to the Berger-Colella evolution algorithm, which uses the concept of *subcycling*, fig. 3.8. The first step is to evolve grid level l one timestep. Then the next level in the hierarchy $l + 1$ is evolved until it reaches the same time as grid level l . For a refinement ratio of

2, level $l + 1$ takes 2 timesteps to reach the same time at level l . After the level has caught up to the coarser level, the data on the coarser level is generated using the data from the finer level using interpolation. Ghost cells on the boundary between levels are populated with data from the coarser level in both space and time. The temporal interpolation is generated using 3rd order polynomial fits in time and the data from the substeps of the RK4 time-stepping.

3.3.3 PARALLELIZATION: MPI, OPENMP, AND SIMD

Modern computational power has reached the exascale, with modern supercomputers being able to execute $\sim 10^{18}$ floating point operations per second (FLOPS)³⁴¹. With modern graphics processing units (GPUs) and central processing units (CPUs) reaching such enormous throughput, modern numerical relativity codes have access to tremendous resources. However, being able to access the computational potential for modern day processing units requires fundamental design changes for the code.

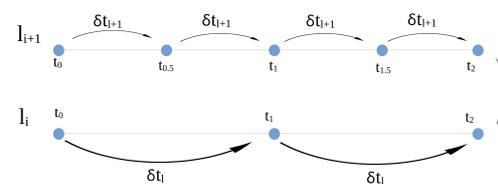


Figure 3.8: Pictorial representation of subcycling. Finer levels undergo evolution at a finer timestep until they reach the time of the coarser level, which then updates the data on the coarser level via interpolation of the data on the finer level. This algorithm is recursive, so levels higher in the hierarchy undergo frequent time-stepping compared to the coarser levels.

In the modern day of multi-node, multi-core, and multi-threaded systems, parallelization can take several forms. There are four main types:

- *Distributed-memory parallelization:* Multiple instances of the same program, called *processes*, run simultaneously, with each process being run on a different CPU core. Each process has its own copy of the data and information required for the simulation. This is a form of inter-node parallelization.
- *Shared-memory parallelization:* Multiple 'threads' are run using a single, shared memory allocation in the CPU's random access memory (RAM). This is typically more memory efficient since the program has access to the same memory locations of the data. However, inter-node parallelization cannot be used because of this. This is instead of form of intra-node parallelization.

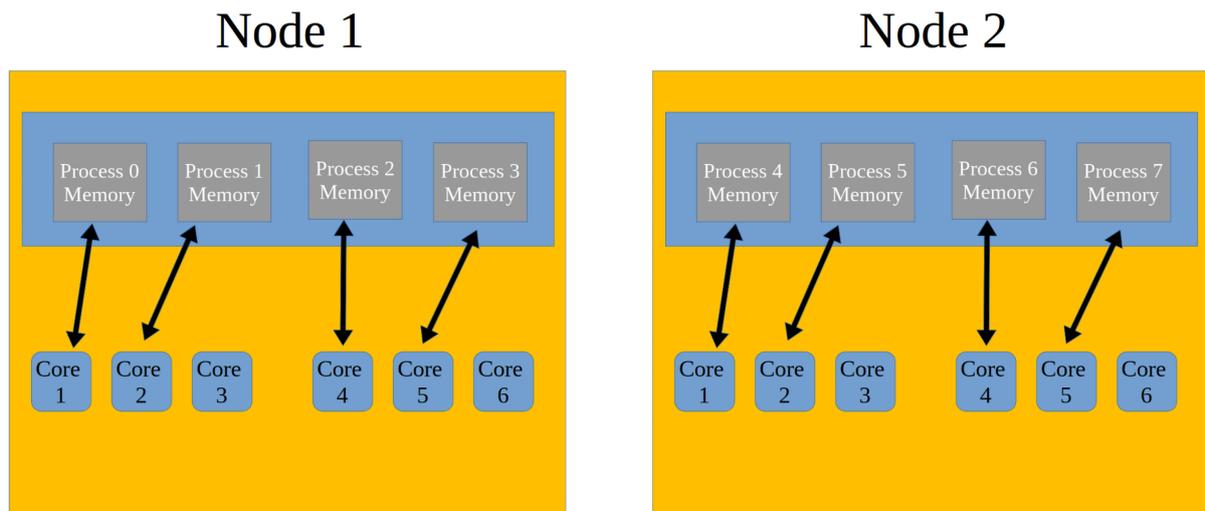


Figure 3.9: Distributed-memory parallelization allows a program to run on several nodes simultaneously, each node supplying its own pool of memory. While this allows for enormous parallelization, it introduces many complexities in program design.

- *Vectorization:* Modern CPUs support so-called ‘single instruction, multiple data’ (SIMD), which means they can perform the same operation on multiple operands simultaneously. For example, multiplying each element of a vector by a scalar can be handled simultaneously on a single core. This type of parallelization is handled by the program compiler, though depending on the program, the programmer may need to alter the code to better leverage vectorization.
- *Heterogeneous computing:* This leverages the massive computational ability of modern day GPUs to perform many simple tasks over a huge number of compute cores.

Distributed-memory allows for programs to run on multiple computers (called *nodes*) simultaneously. For example, a high performance computer cluster (HPC), such as the Baden-Württemberg UniCluster*, is a collection of hundreds of computers connected via high-throughput linkages. Distributed-memory parallelization allows a program to run on several of these nodes at once and access each node’s memory pool. The program is split into several processes, called *ranks*. Each process is an exact duplicate of all the others. Thus, the programmer has to design their program around the fact that it will be split into several copies. Each rank is assigned a unique identifier, thus each copy can perform different tasks depending on the identifier of the rank. The message passing interface (MPI) is a standardized and portable library that enables distributed-memory parallelization. MPI is what enables GRBoondi to be massively parallel.

*All the simulations in this thesis were performed on this cluster.

Shared-memory parallelization involves running a computation using a single process that spawns multiple *threads*, each running on a different core of the CPU, fig. 3.10. These threads all run on a single node, thus accessing the same memory pool. However, shared-memory programs tend to be far more challenging to develop since many factors can cause a huge heap of problems, including race-conditions, different CPU architectures, even the weather can cause variations in results using shared-memory parallelization. Nonetheless, shared-memory parallelization is typically faster than distributed-memory parallelization, since the computations and communications all take place usually on a single CPU, and requires considerably less memory.

Vectorization is the remaining type of parallelization supported by GRBoondi. In particular, the type of vectorization employed is the 'single instruction, multiple data' (SIMD) type. It is a type of parallel processing in which a single operation acts on multiple data points simultaneously. The data is grouped into vectors that the hardware can process in parallel. The size of these vectors depends on the architecture of the CPU. For instance, a 256-bit wide SIMD operator can hold four 64-bit floating-point numbers. This allows significant performance improvement, particularly in tasks with high data parallelism, such as the ones simulated with GRBoondi. The most common SIMD implementations relevant for GRBoondi are the Intel Streaming SIMD Extensions (SSE) and Advanced Vector Extensions (AVX). The AVX SIMD extensions are compatible with x86 instruction sets for processors from both Intel and AMD, the most common CPU manufacturers. AVX-512 extends the standard AVX extension to include 512-bit support, first supported by Intel's Knights Landing processors. The AVX-512 extension can process 8 floating-point numbers simultaneously, a tremendous speed improvement that is hugely beneficial for numerical relativity simulations. Activating SIMD vectorization in the compiler for the simulations has improved processing speeds close to 100 times.

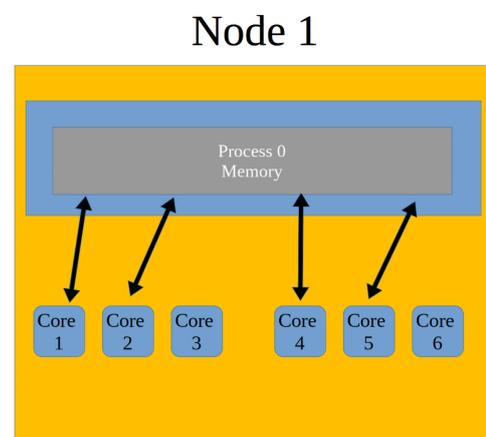


Figure 3.10: Shared-memory parallelization lets a program spawn many different threads on a single CPU, who can all access the same memory pool. While this can provide a considerable reduction in communication latency and has a much smaller memory footprint than distributed-memory, it can also introduce many complexities and bizarre problems since the physical CPU chip can have defects that alter results.

```

1  class DampingFieldDiagnostic
2  {
3  ...
4  template <class data_t> void compute(Cell<data_t> current_cell) const
5  {
6      MatterVars<data_t> matter_vars{current_cell.template load_vars<MatterVars>()}; // load the matter variables from the
7      Chombo grid
8      current_cell.store_vars(matter_vars.Z, c_Z_out);
9  };

```

Listing 3.1: Example of a class that computes a diagnostic quantity using either SIMD vectorization or the built-in double datatype.

GRBoondi utilizes distributed-memory, shared-memory, and vectorization parallelization. Utilization of heterogenous computing is underway by the GRChombo collaborators³⁴². GRBoondi splits the computational domain into boxes, which are then shared between processes via the MPI implementation. The typical CPU used during this thesis is the Intel Xeon Platinum 8358 processor, which features 64 cores and a clock speed of 2.6 GHz. Since intra-node communication is typically much faster than inter-node communication (which takes place via Infiniband HDR200 linkages), multiple MPI processes are spawned on a single node, typically between 4 to 8 processes. The individual cells within a box are then parallelized using the shared-memory distribution library OpenMP. A large portion of the simulation time is spent computing the evolution equations at each RK4 timestep. Hence, parallelizing these steps is vital for performance. GRChombo abstractifies away the SIMD vectorization utilities by defining new C++ datatypes which automatically handle both vectorized and serialized operations, depending on whether the user compiles with SIMD or not. All of the box loops and cell loops are handled by a single `BoxLoops::loop` C++ method. This further abstractifies away the complex task of designing MPI and OpenMP operations. The only requirement from the user is to write a class that takes a (in the terminology used in the C++ language) template datatype, which replaces the standard `double` or `float` datatype. The classes which compute various quantities written by the user must have a `compute` method which executes the computation. The instantiated class is then passed to a `BoxLoops::loop` member function which executes the `compute` method over all cells and boxes for each level of the grid.

As an example of the multiple levels of parallelism in the GRBoondi code, we look at the Proca example. A simple usage of the box loop class is the computation of diagnostic variables. The Proca example has a diagnostic variable called the Damping Field diagnostic, which computes the value of an evolution variable and stores it for plotting. The template datatype `data_t` is what enables the

```

1 DampingFieldDiagnostic z_field_diagnostic{};
2 // compute diagnostics on each cell of current level
3 BoxLoops::loop(
4     make_compute_pack(Asquared, EM, z_field_diagnostic, background_init),
5     m_state_new, m_state_diagnostics, EXCLUDE_GHOST_CELLS
6 );

```

Listing 3.2: An example of looping over the boxes and cells using MPI and OpenMP implementations. The parallelism is abstracted away in a single method called `BoxLoops::loop`.

method `compute` to take either double or SIMD datatypes. In the `BaseProcaFieldLevel` class, the damping field is computed (or in this case extracted) from each cell using a call to the `BoxLoops::loop` method. The damping field class is initialized and then passed to the `BoxLoops::loop` method. Since many other diagnostic quantities are typically computed after each timestep, they are packaged together using the `make_compute_pack` function. Each call to the `compute` method in the `BoxLoops::loop` method automatically threads that call to each class in the compute pack. Additionally, the ghost cells on the grid can be ignored by passing the `EXCLUDE_GHOST_CELLS` flag, which tells the box loops to ignore any ghost cells and only execute the computation on the physical grid cells. Should the user write a class that cannot utilize the SIMD extension, such as when the code contains conditional statements, the vectorization can be disabled at the call to `BoxLoops::loop` by passing another flag via `BoxLoops::loop(..., disable_simd())`. This is a common occurrence when excising certain regions of the grid, where a conditional is employed to only excise cells within a certain zone of the computational domain. Conditionals are quite a challenge to implement alongside SIMD, so sometimes it's easier to just disable SIMD in favor of simpler code. Fortunately, computations such as excision are executed much less frequently than the RK4 steps, so disabling SIMD for these computations has little impact on the computation speed.

3.3.4 EXTRACTION AND DIAGNOSTICS

`GRBoondi` offers various methods to extract quantities from the grid, either by saving directly each value in every cell, integrating quantities over the entire mesh, or integrating quantities over pre-defined surfaces. Primarily, the total energy and fluxes are computed by default in `GRBoondi`, though the user can very easily add additional quantities. This is a particularly vital feature since derived data is almost always required. Hence, this section will review the main pieces of this procedure. Since `GRBoondi` inherits from `GRChombo`, most of the class structure comes from this inheritance,

```

1  AMRReductions<VariableType::diagnostic> amr_reductions(m_gr_amr);
2  std::vector<double> integrals;
3  integrals.push_back(m_p.SymmetryFactor * amr_reductions.sum(c_rho));
4  SmallDataIO constraint_file(m_p.data_path + m_p.integrals_filename, m_dt, m_time, m_restart_time, SmallDataIO::APPEND,
5  first_step);
6  constraint_file.write_time_data_line(integrals);

```

Listing 3.3: An example of computing a volume-integrated quantity. Using the methods of Chombo and the interface classes of GRChombo, such a procedure is simple.

though GRBoondi adds additional functionality specific for generalized Proca theories, including dedicated post-processing routines.

VOLUME-INTEGRATED QUANTITIES

Volume-integrated quantities are those quantities that are formed by integrating variables over the grid

$$\int_{\mathcal{G}} f(x(i), y(j), z(k), Q) dx dy dz , \quad (3.112)$$

where \mathcal{G} is the entire three-dimensional grid, $x(i), y(j), z(k)$ are the three spatial co-ordinates as functions of the grid indices, Q is a grid variable (usually one of the evolution variables), and f is some arbitrary function of position and Q . The procedure to compute this sum is fairly straightforward. Chombo has built-in tools, called `AMRTools`, that allow for methods to be run that need access to the entire level hierarchy. In this case, the `computeSum` method is used. This method computes the sum of a specified variable over all cells in a level and over all levels.

Since the sum computed in this way is blind to the metric, the function f needs to already have the volume element taken into account. For example, in computing the total energy in the Proca example simulations, the energy density ρ is multiplied by the determinant of the spatial metric before being saved to the grid. Hence, $f(x(i), y(j), z(k), Q) = \rho(x(i), y(j), z(k)) * \sqrt{\gamma}$.

At the level of the code, it's very easy to compute volume-integrated quantities, `lst. 3.3`. The `AMRReductions` object is templated over the type of value defined on the grid. In this instance, the type of `c_rho` is a diagnostic variable, in contrast to an evolution variable. This means `c_rho` is understood to be a variable that is not relevant in any way to the evolution equations, instead only as a value that should be computed a posteriori. The `AMRReductions` class is then initialized using the `AMR` object, which holds all the information about the grid. The summation is then computed using the `sum` method of `AMRReductions`, which is then stored in a vector, accounting for any

```

1  for (auto var_enum : vars_to_integrate)
2  {
3    integrals.push_back(m_p.SymmetryFactor * amr_reductions.sum(var_enum));
4  }

```

Listing 3.4: An example of integrating several quantities and storing their output. The vector storage can then be saved to file exactly the same as lst.3.3.

symmetries implied by the boundary conditions. The vector is then stored to disk as a time series datapoint. Multiple quantities can be computed this way with each integral being appended to the vector storage. For example, by iterating over an enum which labels the various diagnostic quantities, one can compute all their integrals as in lst. 3.4.

SURFACE-INTEGRATED QUANTITIES

Extraction of quantities over more general surfaces requires a more involved process, since the surface itself needs to be explicitly defined. This section will use the example of extracting fluxes across a two-dimensional surface, though in principle any quantity can be computed over a general surface using similar methods discussed here.

Integrals of the form

$$\int_{\mathcal{S}} f(u, v, Q) \epsilon_{\mathcal{S}} du dv \quad (3.113)$$

are the primary quantities computed here. \mathcal{S} is the surface over which the integral is evaluated, $f(u, v, Q)$ is an arbitrary function of the co-ordinates of the surface, u and v , and a variable Q , which is typically an evolution variable, and finally $\epsilon_{\mathcal{S}}$ is the co-ordinate area element. Note again that f needs to have the spacetime volume element already taken into account. At each point on the surface, the variable Q is interpolated using the values on the grid cells.

For the example of extracting fluxes, the `FluxExtraction` class takes care of all the necessary computations. It is a child class which inherits from the `SphericalExtraction` class, which itself inherits from the `SurfaceExtraction<SphericalGeometry>` class. The `SurfaceExtraction` class is templated over a `SurfaceGeometry` class, which defines several properties about the surface being integrated, such as the embedding relations of the surface in the computational space, the volume elements, and co-ordinate differentials. The `SurfaceExtraction` class handles the low-level methods for computing the actual integrals using user-specified integration methods*. The

*At the time of writing, the trapezium, Simpson's, and Boole's rule have been implemented as possible integration

```

1  for (auto var_enum : vars_to_extract)
2  {
3      m_gr_amr.fill_multilevel_ghosts(VariableType::diagnostic, Interval(var_enum, var_enum), min_level);
4  };
5  FluxExtraction my_extraction(
6  m_flux_container, m_p.extraction_params, vars_to_extract, m_dt, m_time, first_step, m_restart_time, m_p.SymmetryFactor);
7  my_extraction.execute_query(m_gr_amr.m_interpolator);

```

Listing 3.5: An example of extracting fluxes through a spherical surface.

SphericalExtraction class is a specification of the templated SurfaceExtraction class to extract quantities over spherical surfaces of radius R , $\mathcal{S} = S^2$. In this case, the surface co-ordinates are chosen to be the standard spherical co-ordinates, $(u, v) = (\theta, \phi)$ and the discretization is a homogeneous grid with spacings $\Delta\theta = \frac{\pi}{n_\theta-1}$ and $\Delta\phi = \frac{2\pi}{n_\phi-1}$. This results in the usual co-ordinate volume element $\epsilon_{\mathcal{S}} = R^2 \sin \theta$. The default integration rule is the Simpson rule in both spherical directions with cell counts of $n_\theta = 49$ and $n_\phi = 32$, though these numbers are freely specifiable in the simulation's parameter file.

In the code, implementation of these methods is straightforward. However, since the interpolator needs values defined around a queried cell, the ghost cells need to be filled. Hence, the procedure for computing fluxes is a little more involved than just querying the AMRReductions class, see lst. 3.5 The first step is to fill the ghost cells for variables that are going to be integrated over the spherical surface, to ensure the interpolator has accurate data. Then the FluxExtraction code is initialized, including all variables that are going to be extracted using the vars_to_extract variable. Finally, the query is executed and all variables are integrated over the spherical surface and saved to disk.

The FluxExtraction class has several different pieces but it's conceptually fairly straightforward. The first piece happens at class initialization and involves specifying which grid variables should be extracted, lst. 3.6. The add_var method is a member of the SurfaceExtraction class. It simply appends the variable to be extracted to a vector of names. This is populated by iterating over the m_vars_to_extract member variables of the FluxExtraction class, which finds its origin in the simulation's parameter file, i.e. specified by the user at runtime. The next step is to execute the query, which orchestrates several different steps. Step one is to compute the interpolated values of the grid variables onto the surface itself, lst. 3.7. Step two is to compute the integrands, which sets

methods.

```

1 FluxExtraction(Container &a_flux_container, spherical_extraction_params_t &a_params, std::vector<int> a_vars_to_extract,
   double a_dt, double a_time, bool a_first_step, double a_restart_time = 0.0, double a_symmetry_mult = 1.0)
2 : SphericalExtraction(a_params, a_dt, a_time, a_first_step, a_restart_time), m_vars_to_extract(a_vars_to_extract),
   m_flux_container(a_flux_container), m_symmetry_mult(a_symmetry_mult)
3 {
4 // iterate over variables to extract and add them to the extractor
5 for (auto var : m_vars_to_extract)
6 {
7     add_var(var, VariableType::diagnostic);
8 }
9 }

```

Listing 3.6: Initialization of the FluxExtraction class specifies which variables should be extracted.

```

1 extract(a_interpolator);

```

Listing 3.7: In the FluxExtraction class the first step in executing a surface integration query is to compute the interpolated grid variables on the surface itself.

up the integration method and performs several checks of the discretization. This is performed for each variable that is to be extracted. This step also sets up the output vector, where each integral is stored, see lst. 3.8. Finally, the integration is executed for all variables, see lst. 3.9. The results are then stored in the `flux_integrals` vector, which is a two-dimensional matrix that holds the results of the integrals for each radii of the spherical surface and for each variable.

DIAGNOSTIC QUANTITIES

GRBoondi offers the ability for users to select in-built diagnostic quantities that should be computed throughout the simulation. Moreover, users can define their own classes and quantities that are computed from the evolutionary variables. These are specified in the parameter file of the simulations and hence at runtime. The user can specify which variables should be used as diagnostic variables and hence which ones are saved to disk. These are separated into three distinct categories. The first is plot variables. These are variables whose value at each cell is saved to an HDF5 file format and read in by specialized visualization toolkits (or GRBoondi's own post-processing routines). The second type is the `integration_vars`, or variables that will be integrated across the entire grid. For

```

1 std::vector<std::vector<double>> flux_integrals(m_vars_to_extract.size());
2 for (int var{0}; var < m_vars_to_extract.size(); var++)
3 {
4     add_var_integrand(var, flux_integrals[var], IntegrationMethod::simpson, IntegrationMethod::simpson);
5 }

```

Listing 3.8: In the FluxExtraction class the second step in executing a surface integration query is to set up the integration methods and storage container for the results.

```
1 integrate();
```

Listing 3.9: In the `FLuxExtraction` class the final step in executing a surface integration query is to execute the integration.

```
1 plot_vars = Z Asquared rho rhoE EMtrace EMSquared
2 extraction_vars = Edot Jdot
3 integration_vars = rho rhoE rhoJ EMSquared
```

Listing 3.10: User can specify which variables are used as diagnostic variables.

example, the energy density is integrated to find the total energy across the entire computational domain. The third and final type is the `extraction_vars`, or the extraction variables. These are variables that will be integrated across spherical surfaces of user-specified radii. For example, the energy flux is integrated over a sphere of large radius to find the total flux of a quantity from a spherical region. All three variable types are specified in the parameter file, `lst. 3.10`. Moreover, users can turn off either integration, extraction, or both using distinct flags, `lst. 3.11`. By default, `GRBoondi` offers several diagnostic quantities that users can choose from. These include the Proca energy density, angular momentum scalar density, energy flux, angular momentum flux, Eulerian energy density, the trace of the stress-energy tensor, the square of the stress-energy tensor $\mathcal{T}^{\mu\nu}\mathcal{T}_{\mu\nu}$, the auxiliary constraint-damping scalar, the square of the Proca field $A^\mu A_\mu$, approximately conserved momentum $\alpha\mathcal{T}_i^0$, flux of the conserved momentum, and another quantity associated with the flux of linear momentum³⁴³ *.

3.3.5 MODULARITY

Modularity of `GRBoondi` is one of its core features. It's one of the ways in which `GRBoondi` allows arbitrary modifications coming from generalized Proca theories. The main class that specifies various steps in the simulation, including computing the evolution equations, generating plot files, and postprocessing after each timestep, is the `BaseProcaFieldLevel` class. In the language of C++, the `BaseProcaFieldLevel` class is an abstract class. It cannot be instantiated directly. Instead, it serves

```
1 activate_extraction = 1
2 activate_integration = 1
```

Listing 3.11: User can specify which variables are used as diagnostic variables.

*If the spacetime does not admit the killing vectors associated to these conserved quantities, the user can simply turn off any computation involving them by not specifying them in the parameter file.

```

1 template <class background_t, class proca_t>
2 void BaseProcaFieldLevel<background_t, proca_t>::specificAdvance()
3 {
4     // check for nans
5     if (m_p.nan_check)
6     {
7         BoxLoops::loop(NanCheck(), m_state_new, m_state_new,
8             EXCLUDE_GHOST_CELLS, disable_simd());
9     }
10 };

```

Listing 3.12: Default code for the method `specificAdvance`.

```

1 virtual void specificAdvance() override
2 {};

```

Listing 3.13: Default code for the method `specificAdvance`.

to be purely a parent class, being inherited by other classes. This primarily comes from the need to specify initial data, which is unique for each simulation. Every method in `BaseProcaFieldLevel` is a virtual method, meaning the child class can define their own version of the method, allowing users to change specific functionality in each step of the simulation. For example, one method defines what to do after each RK4 calculation, [lst. 3.12](#). Instead of running this code, users can define their own version of this method in their child class by simply overriding the method itself. For example, if the user defines the child class as `ProcaFieldLevel1`, which inherits from `BaseProcaFieldLevel1`, and they wish for no code to be run in this step, they could simply define the method to be trivial, [lst. 3.13](#). They could do this with any method in `BaseProcaFieldLevel1`. Moreover, some methods in `BaseProcaFieldLevel1` have additional steps built in which, instead of overriding the method, simply add additional functionality after the default code has run. For example, a virtual method called `additionalPostTimeStep` exists which is called after all the code in `BaseProcaFieldLevel1::specificPostTimeStep` is executed, [lst. 3.14](#). There is also the `additionalPrePlotLevel` which adds functionality to the `BaseProcaFieldLevel1::prePlotLevel` method.

The `BaseProcaField` class is the base class which holds the actual evolution equations and equa-

```

1 template <class background_t, class proca_t>
2 void BaseProcaFieldLevel<background_t, proca_t>::specificPostTimeStep()
3 {
4     ...
5     // add any other computations from the user, via virtual function
6     additionalPostTimeStep();
7 };

```

Listing 3.14: Adding additional functionality on top of the default behavior in `specificPostTimeStep`.

tions for the stress energy tensor. This class is then instantiated in the `BaseProcaFieldLevel::specificEvalRHS` method, which computes the RHS and evolves the fields. As the name suggests, `BaseProcaField` only holds the basic evolution equations for the generalized Proca fields. More specifically, it holds the equations for the theory $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ and sets the time derivative of the temporal part of the Proca field and the auxiliary damping field to zero, $\dot{\phi} = \dot{Z} = 0$. In total, the base evolution equations are

$$\dot{X}_i = -\alpha\partial_i\phi - \phi\partial_i\alpha + \beta^j\partial_jX_i - \alpha\gamma_{ij}E^j + X_j\partial_i\beta^j \quad (3.114)$$

$$\begin{aligned} \dot{E}^i = & \alpha\mathcal{K}E^i + \beta^j\partial_jE^i - E^j\partial_j\beta^i + \gamma^{jk}\gamma^{il}(\partial_j\alpha * [\partial_lX_k - \partial_kX_l] \\ & + \alpha[\partial_j\partial_lX_k - \partial_j\partial_kX_l]) - \alpha\gamma^{jk}\gamma^{il}[\Gamma_{jl}^m(\partial_mX_k - \partial_kX_m) + \Gamma_{jk}^m(\partial_lX_m - \partial_mX_l)] \end{aligned} \quad (3.115)$$

$$\dot{\phi} = 0 \quad (3.116)$$

$$\dot{Z} = 0, \quad (3.117)$$

where the $3+1$ variables of the Proca field are $\phi = -n^\mu A_\mu$, $E^i = \gamma_\mu^i n_\nu F^{\mu\nu}$, $X_i = \gamma_{i\mu} A^\mu$, and Z is an auxiliary field introduced to damp violations of the Proca constraint, similar to the `CCZ4` formalism, eq. 3.74. Higher order terms in the generalized Proca theory will modify these evolution equations. To deal with these new modifications, the `BaseProcaField` class must somehow allow the evolution equations to be modified. Due to subtleties in templated and virtualized functions in C++, it's not possible to create a virtualized function that is templated over typename. However, there does exist a type of coding idiom called *curiously recurring template pattern*, or `CRTP`, which allows parent classes to access methods of the child class by passing the child class as a template argument to the parent class itself. If this wasn't the case, then the templated method in `BaseProcaField` would have to access a virtualized method, which would be defined by the user, which is not possible in C++. The specific implementation of `CRTP` by `GRBoondi` follows:

- `BaseProcaField` is templated over two class types, one for the background spacetime and another for the evolution modifications.

```
1 template <class background_t, class modification_t> class BaseProcaField
```

- After the evolution equations are computed, a method of the templated class is called, which adds the modifications to the evolution equations.

```

1 void BaseProcaField<background_t, modification_t>::matter_rhs(...)
2 {
3     ...
4     // add modifications ala CRTP
5     static_cast<const modification_t *>(this)->matter_rhs_modification(
6     total_rhs, matter_vars, metric_vars, d1, d2, advec);
7 }

```

This static casts the BaseProcaField object to the templated class, which is simultaneously a child class. Finally, the method which modifies the evolution variables is called from the casted object.

- Users then define a ProcaField class which inherits from BaseProcaField as

```

1 class ProcaField : public BaseProcaField<background_t, ProcaField>

```

and defines a method called matter_rhs_modification which modifies the evolution equations to account for higher order terms in the generalized Proca Lagrangian

```

1 void matter_rhs_modification(...)
2 {
3     FOR1(i)
4     {
5         total_rhs.Evec[i] += ...
6         total_rhs.Avec[i] += ...
7     }
8     total_rhs.Z += ...
9     total_rhs.phi += ...
10 }

```

where FOR1(i) is a preprocessor definition for for(int{i} = 0; i<3; ++i).

With these procedures, users can easily add modifications to the base $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ Lagrangian. These modifications are then automatically taken into account in the BaseProcaField evolution computation.

3.4 CONVERGENCE BEHAVIOR OF GRBoondi

When performing numerical computations of any kind, relativity included, it's vital that convergence and consistency checks are performed alongside the main simulations. Without these quan-

titative error estimations, the resulting data from the main simulations are completely worthless*. Numerical simulations unavoidably introduce errors in the final solution, originating from multiple places. Firstly, the finite size of the grid can, and does, introduce high frequency noise with frequencies comparable to the grid spacing. This is the reason GRBoondi uses Kreiss-Oliger dissipation. Secondly, the interpolation errors from, for example, computing derivatives using ghost cells near box boundaries or level boundaries introduce additional errors. Thirdly, highly dynamical spacetimes such as black hole spacetimes can exacerbate existing errors, compounding them and causing them to dramatically increase. Finally, in the context of finite numerical simulations, boundary conditions are unphysical and will contribute to the noise and errors of the simulation. There are potentially many more sources of errors that can arise in numerical relativity. Hence, it's absolutely critical that users of the code approximately quantify the total amount of errors in the simulation, otherwise there is no confidence in the accuracy and *precision* of the resulting data. To that end, GRBoondi performs several convergence tests, correctness tests, and consistency tests for various methods and background spacetimes.

3.4.1 THEORY OF CONVERGENCE TESTING

The key idea of convergence testing is the observation that the solution of a stable finite differencing scheme can be expanded as a continuous function in a power series of the discretization parameter ϵ ³⁴⁵

$$q_\epsilon(t, x) = q(t, x) + \epsilon e_1(t, x) + \dots + \epsilon^n e_n(t, x) + \dots, \quad (3.118)$$

where $q(t, x)$ is the (continuous) solution of the original differential equation and the $e_i(t, x)$ are so-called *error functions* at different orders in ϵ . For an "n'th order accurate approximation", we expect $e_{i < n} = 0$ and $e_n \neq 0$.

Assume for the moment that we know the exact solution to the differential equation and we want to compare the numerically computed solution against the known solution. To test the convergence of the numerical solution to the true solution, we perform the computation at two resolutions, say Δ_1 and Δ_2 , with $r \equiv \frac{\Delta_1}{\Delta_2} > 1$.

*Boyd defines an 'idiot' as someone who publishes a numerical calculation without checking it against an identical calculation with a different resolution³⁴⁴.

In each case, we compute the error between the numerical and true solution

$$E_1 = q - q_{\Delta_1} \quad (3.119)$$

$$E_2 = q - q_{\Delta_2} . \quad (3.120)$$

Notice that this difference can only be computed at the grid points of the simulation, since that is the only place where q_{Δ_1} and q_{Δ_2} are respectively defined. Moreover, these grid points will be different between the two simulations. We then compute the r.m.s. norm for each solution error and calculate the ratio

$$c(t) = \frac{\|E_1(t)\|}{\|E_2(t)\|} . \quad (3.121)$$

This ratio, being a sort of spatial integration of the solution errors, is purely a function of time and is called the *convergence factor*. If we have a numerical scheme that is of n 'th accuracy, the solution expansion eq. 3.118 will yield the convergence factor in the continuum limit

$$\lim_{\Delta \rightarrow 0} c(t) = \left(\frac{\Delta_1}{\Delta_2} \right)^n = r^n . \quad (3.122)$$

The convergence tests are typically performed with $r = 2$, i.e. the resolution of the simulation is doubled. However, in principle it can be performed with any ratio. The simulation is performed with several higher resolutions and, if the behavior is close to the expected value, we are said to be in the *convergence regime*.

However, the issue with this approach is that it assumes we know the exact solution of the differential equation to begin with. Most of the time, including all cases relevant to GRBoondi*, this is not the case. The best we can do is prove that the simulations converge to *something* and hope that thing is the true solution†. To that end, the simulation is repeated with *three* different resolutions $\Delta_1 > \Delta_2 > \Delta_3$. The *relative* errors are then computed and the convergence factor is defined as

$$c(t) = \frac{\|q_{\Delta_1} - q_{\Delta_2}\|}{\|q_{\Delta_2} - q_{\Delta_3}\|} . \quad (3.123)$$

*Except for some testing cases, see below.

†This is where good initial data is vital. If the initial data is too far off of the expected result, the simulation can convergence to something that is *not* the desired solution, but is a solution to an entirely different problem.

In the continuum limit, we expect

$$\lim_{\Delta \rightarrow 0} c(t) = \frac{\Delta_1^n - \Delta_2^n}{\Delta_2^n - \Delta_3^n} . \quad (3.124)$$

For equal resolution ratios, $\Delta_1/\Delta_2 = \Delta_2/\Delta_3 = r$, we have

$$\lim_{\Delta \rightarrow 0} c(t) = r^n . \quad (3.125)$$

Convergence testing not only allows us to ensure the errors are decreasing at the expected rate as resolution is increased, but also estimate the error in the final solution itself. Assume we have an n 'th order accurate scheme and we've performed the simulations for two different resolutions, Δ_1 and Δ_2 . The series expansion eq. 3.118 tells us

$$q_{\Delta_1} - q_{\Delta_2} = e_n (\Delta_1^n - \Delta_2^n) + \mathcal{O}(\Delta^{n+1}) \sim E_{\Delta_2} (r^n - 1) . \quad (3.126)$$

The solution error on the highest resolution of the grid is then

$$E_{\Delta_2} \sim \frac{1}{r^n - 1} (q_{\Delta_1} - q_{\Delta_2}) . \quad (3.127)$$

This error estimate allows the creation of error bars for the simulations.

3.4.2 CONVERGENCE CHECKS

GRBoondi comes built in with several different background spacetimes and so several convergence tests are performed, one convergence check for each grid variable. Since the convergence check is functionally identical for each background spacetime, the basic layout will be discussed once and then the results for each check will follow.

The first step is to define the result containers and number of resolutions to run at, lst. 3.15. Next, the iteration over the resolutions begins and the grid is initialized based on the current resolution, lst. 3.16. Whichever background spacetime is being tested is then initialized, lst. 3.17. The GR constraints are then computed to verify that the background spacetime is indeed a solution of the Einstein equations. The Proca field class is then initialized and the evolution equations are com-

```

1  const std::vector<int> resolutions{96, 192, 384}; // resolutions to run at
2  const int num_resolutions = 2; // how many of the resolutions to actually use
3  // vector of norms for convergence checking
4  std::array<std::array<double, NUM_VARS>, num_resolutions> error_norms;

```

Listing 3.15: Setup phase for convergence testing.

```

1  for (int ires{0}; ires < num_resolutions; ++ires)
2  {
3      // setup the array boxes for various inputs and outputs
4      const int N_GRID{resolutions[ires]}; // number of cells on each side of box
5      // setup boxes
6      Box box(IntVect(0, 0, 0), IntVect(N_GRID - 1, N_GRID - 1, N_GRID - 1)); // The computational box

```

Listing 3.16: Resolution iteration begins and grid is initialized.

puted using *numerically* computed derivatives, including the background ones, lst. 3.18. where `fixedbg_FAB` and `rhs_FAB` are the boxes of cells on the grid. Then the evolution equations are computed again, however this time using the analytic derivatives for the background variables, lst. 3.19. Then the difference between the two solutions is taken, lst. 3.20.

A basic check is run to ensure constraints aren't violated and that the difference between the two solutions isn't too large. Then the convergence factor is computed via eq. 3.123. The results of the convergence checks for each background built in to `GRBoondi` are displayed in table 3.1.

It's not clear what convergence factor we should expect, since there are many levels of refinement, and various ingredients at different orders. However, since `GRBoondi` uses fourth-order stencils, we can expect $c(t)$ to be somewhere in the range of $c(t) = 16$. For the Minkowski background, we obtain exactly $c(t) = 16$, however for the Boosted Schwarzschild background, the convergence factor is less, closer to what we would expect for a third-order accurate approximation. Nonetheless, a high convergence factor such as 12.7084 is sufficient.

3.4.3 CONSISTENCY CHECKS

Besides convergence tests, `GRBoondi` also performs consistency checks to ensure the code behaves as it was designed to. The first is checking that the modifications introduced via a `ProcaField` class reproduces the expected evolution equations. To verify this, the equations of motion derived from

```

1  background_t background_init(bg_params, dx);

```

Listing 3.17: Initialization of the background class.

```

1 ProcaFieldTest matter(background_init, proca_params);
2 // setup matterccz4 rhs with matter class
3 MatterCCZ4RHS<ProcaFieldTest> matter_ccz4_rhs(
4 matter, ccz4_params, dx, sigma, CCZ4RHS<>::USE_BSSN, G_Newt);
5 BoxLoops::loop(matter_ccz4_rhs, fixedbg_FAB, rhs_FAB);

```

Listing 3.18: Evolution equations are computed using numerically computed derivatives for the Proca *and* background variables.

```

1 ProcaField::params_t proca_params = {1, 1, 1};
2 ProcaField analytic_matter(background_init, proca_params);
3 // compute RHS using analytic expressions
4 MatterEvolution<ProcaField, background_t> my_an_evolution(
5 analytic_matter, background_init, sigma, dx, center_vector);
6 BoxLoops::loop(my_an_evolution, fixedbg_FAB, fixedbg_rhs_FAB);

```

Listing 3.19: Evolution equations are recomputed, however this time using the known analytic derivatives for the background variables.

the Lagrangian $\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\mu^2 A^\mu A_\mu$ are computed, yielding

$$\frac{1}{\alpha}\mathcal{L}_m E^i = E^i K - D^i Z + \mu^2 X^i - \frac{2}{\alpha}D_j(\alpha D^{[j} X^{i]}) \quad (3.128)$$

$$\frac{1}{\alpha}\mathcal{L}_m Z = -\mu^2 \phi - D_i E^i - \kappa Z \quad (3.129)$$

$$\frac{1}{\alpha}\mathcal{L}_m X_i = -E_i - D_i \phi - \phi D_i \ln(\alpha) \quad (3.130)$$

$$\frac{1}{\alpha}\mathcal{L}_m \phi = \frac{Z}{\mu^2} + \phi K - D_i X^i - x^i D_i \ln(\alpha), \quad (3.131)$$

where $\mathcal{L}_m Q = (\partial_t - \mathcal{L}_\beta)Q$ and the auxiliary field Z is introduced to damp constraint violations.

These equations are then modifications of the base equations used in GRBoondi, eqs. 3.114. This check will test the ability for GRBoondi to correctly incorporate modifications to the field equations.

The test proceeds similarly to the convergence checks in the previous section, but this time only choosing a single resolution. The grid is set up in the same manner and a Kerr black hole background is initialized. Two separate ProcaField classes are instantiated. The first contains all the evolution equations from eqs. 3.128-3.131 and the second contains only the modifications from the base evolution equations eqs. 3.114-3.117. Finally, the two evolution equations are computed and their results subtracted, lst. 3.21. The results of each evolution are stored in my_rhs_FAB and ref_rhs_FAB, respectively, and their difference taken in the last line. The error is then computed

```

1 rhs_FAB -= fixedbg_rhs_FAB;

```

Listing 3.20: Difference between the two solutions using numerically and analytically computed derivatives.

Convergence Tests	
Background	Convergence Factor
Boosted Schwarzschild	12.7084
Minkowski	16
Kerr-de Sitter	14.3488
Kerr	15.5201

Table 3.1: Convergence test results for background spacetimes shipped with GRBoondi.

```

1 MatterEvolution<ProcaField, KerrSchild> my_matter(my_proca_field, kerr_init, 0.0, dx, center_vector); // GRBoondi evolution.
   set Kreiss-Oliger to zero
2 MatterEvolution<FixedBGProcaField<Potential>, KerrSchild> ref_matter(test_proca_field, kerr_init, 0.0, dx, center_vector); //
   Reference evolution. set Kreiss-Oliger to zero
3 // Now loop over the box and compute the RHS
4 BoxLoops::loop(my_matter, my_FAB, my_rhs_FAB);
5 BoxLoops::loop(ref_matter, ref_FAB, ref_rhs_FAB);
6 ref_rhs_FAB -= my_rhs_FAB; // now subtract the two

```

Listing 3.21: Checking the ability for GRBoondi to correctly incorporate modifications to the Proca equations of motion.

using the L^∞ -norm, lst. 3.22. The current version of GRBoondi passes this check perfectly, with all errors falling below the critical threshold.

GRBoondi also contains other smaller checks to test the correctness of various helper functions, though these will not be detailed here to avoid pedantism.

3.5 PERFORMANCE BENCHMARKING AND ANALYSIS

Benchmarking the performance of any numerical relativity code is a vital part of ensuring the software runs as it was designed to. There are a huge number of places where performance potential could be lost. For example, by improper memory management, poorly optimized coding choices, compiler options, processor architecture, etc. Running performance checks allows one to verify the code runs as intended and to find potential performance bottlenecks.

The performance of GRChombo has been thoroughly tested up to 285,600 cores on the Stam-

```

1 for (int i{c_phi}; i <= c_Z; i++) //iterate over each evolution variable
2 {
3   double max_err = ref_rhs_FAB.norm(max_norm, i, num_comps); //compute the norm of the difference between the two evolution
   computations for current variable
4   if (max_err > error_limit)
5   {
6     failed = -1; //if the error is larger than a predefined limit, then test fails
7   }
8 }

```

Listing 3.22: iterating over the errors and checking they fall below a predefined threshold.

pede2 supercomputer, which uses Intel Xeon Phi 7250 processors. These tests showed that GRChombo scales both strongly and weakly with increased resources. *Strong scaling* refers to the speed increase for simulations when the available resources is increased. For an ideal runtime, doubling the number of resources should cut the simulation walltime in half. GRChombo also shows *weak scaling*. This refers to the ability for more expensive simulations to run in the same amount of time for a proportional increase in the number of computational resources. Since GRBoondi inherits most of the mesh generation and time-stepping code, it automatically inherits these traits as well.

These scaling tests refer to how well the program itself handles increased resources. Another place programs can gain efficiency, even without increasing the resource pool, is by researching and applying compiler options when building the program. Sometimes this can feel like an artistic movement rather than a scientific approach, due to the huge number of possible compiler options and number of compilers. In the development of GRBoondi, it was found that carefully picking compiler options can increase the speed of simulations up to hundreds of times, so it is worth running optimization benchmarks to understand the effect of each option and search for improvements.

Another benchmark will be comparing GRBoondi to GRChombo itself. Since GRChombo is a versatile toolset for evolving matter and spacetime in tandem, for some problems it can be too large of a tool than is necessary*. This versatility can even cause greater computational overhead, slowing simulations down just from these added features existing, especially when multiple unnecessary code files are incorporated at compile time[†]. The advantage of using GRBoondi against GRChombo for simulating superradiant Proca fields in black hole spacetimes will be analyzed. The effect of using a fixed background compared to a fully evolving one will be tested, showing that the fixed background approximation in GRBoondi is both precise enough for useful data and a huge optimization.

3.5.1 PERFORMANCE COMPARISONS

The two main benchmarks performed for GRBoondi are then

- Compilation optimizations using the Intel OneAPI DPC++/C++ Compiler. Since the simulations performed in this test are ran on Intel processors, the Intel C++ compiler typically

*Analogously, using GRChombo for simpler problems is like using a sledgehammer to crack a nut.

[†]Although, advanced compilers should be able to optimize these redundant files away.

offers the best performance. With sufficient knowledge of compiler options, similar performance can also be obtained with the GNU compilers.

- Performance comparison between evolving the full Einstein+Proca system using GRChombo and evolving just the Proca system on a fixed background using GRBoondi.

CPU AND ARCHITECTURE

Since performance metrics, especially compiler optimizations, are largely architecture dependent, it's vital that the testing computers are standardized across all tests. The tests in this section are performed on an Intel Xeon Platinum 8358 processor with a base frequency of 2.6 GHz and turbo frequency of 3.4 GHz. The turbo frequency is the maximum single-core frequency at which the processor is capable of operating*. The CPU cache is 48 megabytes, which is memory storage located physically on the processor itself, in contrast to system memory, which is typically stored on the main motherboard of the computer and accessed by the processor via PCIe channels. Each processor contains 32 cores, or 32 independent central computing units, and 64 threads, or 64 'logical' cores.

These tests run on a various number of compute nodes in the Baden-Württemberg high performance computing cluster (bwHPC), with each compute node containing two Intel Xeon processors and 256 gigabytes of system memory. Each node is connected via an Infiniband HDR200 (High Data Rate 200) interconnect, which supports data transfer speeds up to 200 gigabytes per second (Gbps) per port. These interconnects have extremely low latency, allowing code highly parallelized using tools such as MPI to achieve extremely fast speeds, even when several processes are scattered across several nodes. The operating system of the nodes is Red Hat Enterprise Linux release 8.8.

COMPILATION AND OPTIMIZATIONS

The first benchmark is evaluating the effects of different compilation options for the Intel OneAPI C++ compiler, henceforth denoted by `icpc`. There are many different types of compiler options,

*Though, consumer processors can be, and are, overclocked to allow higher clock speeds, though this comes at the cost of increased thermal energy and decreased stability.

ranging from optimizations, code generation, floating point operations, and linker settings. By far and large, the optimization options are naturally the most relevant, though other options can subtly impact performance. Reviewed here are some of the most important options that should be enabled.

- `-O[n]`: This flag tells the compiler the broad level of optimization to use for the compilation. The maximum of `-O3` tells `icc` to be aggressive in its optimization. The compiler will use many fancy techniques to generate a highly optimized executable. This option is especially useful for loops that have many floating-point operations, which is highly relevant to `GRBoondi`.
- `-parallel`: This tells the compiler to enable the auto-parallelizer, which generates multi-threaded code for loops.
- `-xCORE-AVX512`: This is an extremely important compiler option, which tells the compiler to generate AVX instruction code, a type of vectorization (see sec. 3.3.3). There are many different Intel vectorization features enabled with this option. The code compiled with this option can see significant speed up. It should be noted the exact vectorization flag is highly architecture dependent. The option `-xCORE-AVX512` works on the Intel Xeon processors that the simulations were executed on. Users of `GRBoondi` need to be aware of the architecture the code is being compiled on and adjust this option accordingly.
- `-qopt-zmm-usage=high`: This option pertains to the vectorization of the program and should be used in accordance with the `-xCORE-AVX512` option. This option tells the vectorizer to maximize the usage of particular SIMD registers in the processor, providing further speed-up.
- `fp-model fast=2`: This option controls the semantics of floating-point calculations. The floating-point domain of the processor is a collection of registers that control the floating point behavior of the instruction sets. Setting the model to `fast=2` tells the compiler to use more aggressive optimizations, increasing speed at the slight detriment in accuracy and reproducibility. Analysis of `GRBoondi` shows these detriments are unnoticeable.

-
- `-ipo`: This option enables interprocedural optimization between files, or *Whole Program Optimization* (WPO). This option inlines functions that are defined in separate files (replaces the function call with the function code itself).

The options that are absolutely critical for highly optimized code are the `-O` and `-xAVX-512` option set. There are also various other options set for the `icpc` compiler, though they are usually warning-suppression or smaller floating point optimizations and are largely irrelevant compared to the vectorization and optimization features.

Since GRBoondi inherits from Chombo, there is a very high-level compilation option offered by the Chombo library called `OPT`. It has three possible choices: `FALSE`, `TRUE`, and `HIGH`. Setting it to `TRUE` turns on the optimizations (though it's not entirely clear the effect of this. Setting `OPT=TRUE` at the very least has the effect of including `-O3` in all executable compilations.). Setting the option to `HIGH` turns off all assert statements and initializes memory locations to zero during allocation. Naturally, this option also has a substantial effect on performance.

The performance improvement garnered by setting these options can be quantified by turning on each option incrementally. This benchmark was executed by selectively turning on various compilation options and executing a standardized simulation. The simulation consists of a Kerr black hole of mass 1 and dimensionful spin $a = 0.99$. Plot files and checkpoint files are disabled, to ensure no unnecessary performance penalties from writing to files. The grid consists of four static refinement levels, with grid co-ordinate size 64^3 and each level has 64^3 cells. The box sizes are fixed to be built from 16^3 cells. Sommerfeld outgoing radiation boundary conditions are chosen along with reflective symmetry about the z -plane to take advantage of the z -symmetry. The CFL factor is 0.2 and nan checking is disabled. The spatial resolution on the finest grid is $dx_{fine} = 0.0625M$, corresponding to a temporal resolution of $dt_{fine} = 0.0125M$. The simulation rate is sampled exactly when the lowest level in the AMR hierarchy reaches $t = 20M$. The simulation is repeated three times and the average of the simulation rate at $t = 20M$ is taken to obtain the result for that set of compilation options. This usually gives the simulation enough time to settle down into a steady state after all the necessary boilerplate components have been initialized and set up.

The simulations were ran on two compute nodes of the bwHPC, each containing two Intel Xeon Platinum 8358 processors using 86GB of system memory per node. Eight MPI processes are

spawned on each node and eight cores are allocated to each process. The results of the benchmark are displayed in table 3.2. Let R_t denote the rate of the simulation measured in geometrized seconds per hour, $\frac{M}{hr}$, measured at simulation time t . This is a conversion between how many co-ordinate seconds can be simulated in one hour of walltime.

Compilation Benchmarking	
Compiler Options	Simulation rate
DEBUG=TRUE, OPT=FALSE, no other options	D.N.F.
DEBUG=FALSE, OPT=FALSE, no other options	$R_{20M} = 97.86 \frac{M}{hr}$
-O3	$R_{20M} = 372.96 \frac{M}{hr}$
-parallel, -O3	$R_{20M} = 376.44 \frac{M}{hr}$
-xCORE-AVX512, -parallel, -O3	$R_{20M} = 491.87 \frac{M}{hr}$
-xCORE-AVX512, -qopt-zmm-usage=high, -parallel, -O3	$R_{20M} = 516.92 \frac{M}{hr}$
-fp-model fast=2, -xCORE-AVX512, -qopt-zmm-usage=high, -parallel, -O3	$R_{20M} = 518.09 \frac{M}{hr}$
-no-prec-div, -fp-model fast=2, -xCORE-AVX512, -qopt-zmm-usage=high, -parallel, -O3	$R_{20M} = 516.95 \frac{M}{hr}$
-qoverride-limits, -no-prec-div, -fp-model fast=2, -xCORE-AVX512, -qopt-zmm-usage=high, -parallel, -O3	$R_{20M} = 504.95 \frac{M}{hr}$
-ipo, -qoverride-limits, -no-prec-div, -fp-model fast=2, -xCORE-AVX512, -qopt-zmm-usage=high, -parallel, -O3	$R_{20M} = 517.07 \frac{M}{hr}$

Table 3.2: Performance benchmarking of compilation options for the Intel OneAPI C++ compiler, averaged over three separate runs.

These results show how important vectorization is for generating efficient simulation code, as well as using the high-level optimization flag -O3. Turning off all debugging features and enabling all optimization features increases performance several times.

ASSESSING RELATIVE PERFORMANCE

Comparing the simulation rates for GRBoondi and GRChombo highlights the efficiency boost garnered by simply neglecting backreaction. Although one typically has to be careful with this approximation to ensure it's physically realistic, the vast majority of cases relevant for GRBoondi permit such an approximation. Utilizing this approximation allows one to turn off the evolution of the background spacetime and simply evolve the Proca field on a fixed background. This can permit

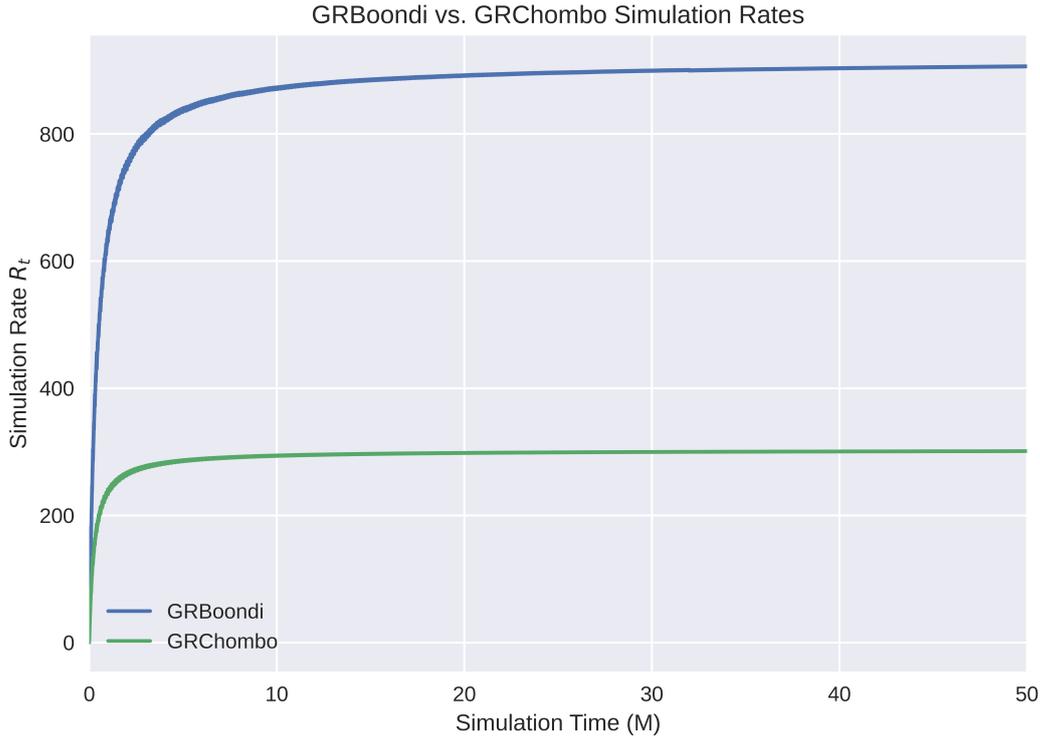


Figure 3.11: Performance metric of GRBoondi compared to GRChombo. The reduction in evolution equations from choosing the fixed background approximations results in a dramatic speed-up in simulation rates.

a tremendous boost in simulation performance since there are less equations to solve using the Runge-Kutta method. In this section, we estimate the performance boost by comparing the simulation rates of GRBoondi and an equivalent physical system using GRChombo.

The simulation setups are equivalent to the previous section. A $\mu = 0.4$ Proca field begins as a perturbation around a rapidly spinning $M = 1$ and $a = 0.99$ Kerr black hole. 4 levels of mesh refinement are chosen on a fixed grid with co-ordinate volume $V = 64^3$ and the side lengths of the computational box are partitioned into $N = 64$ cells. The box dimensions are constrained to consist of 16^3 cells. The z -axis symmetry is exploited by imposing symmetric boundary conditions at the $z=0$ plane and only evolving the $z>0$ domain. This choice of refinement levels and cell partitions implies the finest level has a spatial resolution of $dx_{fine} = 0.0625M$. Choosing a CFL factor of 0.2 implies the temporal resolution is $dt_{fine} = 0.0125M$. Generation of plot and checkpoint files was disabled.

The simulations were ran on four compute nodes of the bwHPC, each containing two Intel Xeon Platinum 8358 processors using 86GB of system memory per node. 8 MPI processes are

```

1 template <class data_t> void compute_emtensor_modification(...) const {};
2 template <class data_t, ...> void matter_rhs_modification(...) const {};

```

Listing 3.23: The electromagnetic example in GRBoondi highlights the most basic usage of the software adding no additional terms to the base equations.

spawned on each node and 8 cores are allocated to each process. The results of the benchmark are displayed in table 3.3. Let R_t denote the same metric as the previous section. The simulation rates are extracted on the same level at a simulation time of $t = 50M$. The tremendous performance

Performance Benchmarking	
Simulation Toolkit	Simulation rate
GRChombo	$R_{50M} = 301.18 \frac{M}{hr}$
GRBoondi	$R_{50M} = 909.24 \frac{M}{hr}$

Table 3.3: Performance improvement of GRBoondi relative to GRChombo. This test highlights the efficiency boost gained from exploiting a perturbative backreaction in many physical systems. This test uses a superradiant Proca cloud around a rapidly spinning black hole as a proxy.

boost from using GRBoondi is evident. The simulation runs at three times the rate compared to GRChombo, thanks to the dramatic reduction in evolution equations, fig. 3.11.

3.6 EXAMPLES

GRBoondi comes equipped with several example systems to highlight the various features of the code. Each example highlights a different aspect of the code and shows users a good starting point for building their simulations.

3.6.1 ELECTROMAGNETIC FIELDS ON MINKOWSKI

The first and simplest example is that of electromagnetic fields on a Minkowski background. This highlights the most basic usage of GRBoondi, without any additional terms in the Lagrangian. The theory under study in this example is the standard electromagnetic field, $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$. In this system, the modifications to the evolution equations are trivial and the ProcaField class adds no additional modification to the evolution equations where the dots represent additional arguments, either additional template parameters or function arguments. They are omitted here for clarity.

The only equations that are computed are then the base evolution equations, eqs. 3.114-3.117. The level class definition is the most irreducible version that any level class can be, only specifying

```

1 class EMFieldLevel : public BaseProcaFieldLevel<DefaultBackground, ProcaField>
2 {
3     public:
4         // inherit constructor from base class
5         using BaseProcaFieldLevel::BaseProcaFieldLevel;
6
7         // override method to calculate initial data
8         virtual void initialData() override
9         {
10            // Initialize the initial conditions class
11            Initial_EM_Conditions initial_conditions(m_dx, m_p.initial_conditions_params, m_p.background_params);
12
13            // Loop over box cells and assign initial EM field
14            BoxLoops::loop(initial_conditions, m_state_new, m_state_new, INCLUDE_GHOST_CELLS);
15        }
16 };

```

Listing 3.24: Level class of the electromagnetic example, representing the most irreducible version that any level class can be. The necessary definition is specification of the initial data.

the initial data for the electromagnetic field. This highlights a very basic setup for initial data. The class which computes the initial data at each grid point, as a function of co-ordinates, is initialized using parameters specified by the user in the parameter file. Then the class is looped over the boxes and cells, populating the grid with initial values for the variables, including the ghost cells.

The initial data itself is also simple, specifying the initial data of the electromagnetic field to be that of an electrically charged point particle, lst. 3.25.

This also highlights the simple procedure of storing data on the grid. The co-ordinates of the current grid cell are computed using the `Coordinates<>` class, which computes the co-ordinate position of the cell. The co-ordinate radius, $r = \sqrt{x^2 + y^2 + z^2}$, can simply be computed using the built-in `.get_radius()` method of the `Coordinates<>` class. The initial amplitude is then extracted from the user-defined parameters. The matter evolution variables on the grid are then flushed with zeroes to ensure no memory initialization values tarnish further computations. Finally, the initial data is computed using analytic expressions and saved to the current cell using the `.store_vars()` method.

Another important fundamental piece that users need to specify is the parameter class, called `SimulationParameters`, which inherits from the base `ProcaSimulationParameters`. This class loads the parameters defined in a parameter field and stores them as members of the class, lst. 3.26.

The members of the class can be simple datatypes such as `doubles`, or they can be `structs` defined in other classes, which is the case here. For example, in lst. 3.25 line 4, the initial data struct is defined and in lst. 3.26 line 16, it's initialized as a member of the `SimulationParameters` class.

```

1 class Initial_EM_Conditions
2 {
3 public:
4     struct params_t
5     {
6         double init_amplitude;
7     };
8
9 protected:
10    ...
11 public:
12    Initial_EM_Conditions(const double a_dx, params_t a_params, const DefaultBackground::params_t a_Kerr_params): m_params{
13        a_params}, m_dx{a_dx}, m_Kerr_params{a_Kerr_params} {};
14
15    template <class data_t> void compute(Cell<data_t> current_cell) const
16    {
17        Coordinates<data_t> coords(current_cell, m_dx, m_Kerr_params.center); // location of cell
18
19        data_t radius{coords.get_radius()}; //radius of the current cell
20        double amp = m_params.init_amplitude; // initial amplitude of the EM field
21
22        MatterVars<data_t> matter_vars; //initialize the matter variables
23        VarsTools::assign(matter_vars, 0.); // assign all matter variables in this cell to zero
24
25        matter_vars.phi = -amp / radius; // assign scalar part to be that of electrically charged point particle
26        matter_vars.Evec[1] = amp / radius / radius; //assign corresponding electric part
27
28        current_cell.store_vars(matter_vars); // push matter vars to cell
29    }
30 };

```

Listing 3.25: Initial data for the electromagnetic field takes the simple form of an electrically charged point particle.

```

1 class SimulationParameters : public ProcaSimulationParameters
2 {
3 public:
4     SimulationParameters(GRParmParse &pp) : ProcaSimulationParameters(pp)
5     {
6         read_params(pp);
7         check_params();
8     }
9
10    void read_params(GRParmParse &pp)
11    {
12        pp.load("center", background_params.center, center);
13        pp.load("initial_em_field", initial_conditions_params.init_amplitude); // Initial EM field params
14    }
15    DefaultBackground::params_t background_params; // parameters of kerr bh
16    Initial_EM_Conditions::params_t initial_conditions_params; // initial conditions parameters
17    ProcaField::params_t matter_params; // Proca field params. In this class, an empty struct
18 };

```

Listing 3.26: Simple example of setting up the SimulationParameters class, which reads in the parameter file and saves the values to members of the class.

The initial data parameters can then be stored in this struct and used throughout classes that contain references to the class.

The remaining piece is to write a main function for the program to execute. The main function has many responsibilities. It must:

- set up and coordinate the MPI processes before anything else;
- initialize the `SimulationParameters` class and pass the parameter file to the class in order for the parameters to be parsed;
- execute the `ProcaLevelFactory`, a class which is responsible for setting up each level of the AMR hierarchy;
- set up the AMR interpolator, for interpolating data across grid cells;
- set up any timing functionality for analyzing the simulation performance metrics;
- execute the simulation;
- perform cleanup of the AMR object;
- finalize the MPI processes and clean up memory.

Fortunately, almost all of these processes are constant boilerplate procedures for every simulation. `GRBoondi` relegates almost all of these setup procedures to a single function called `runGRBoondi`.

The example can then be compiled using the standard procedure of running `make all`. This generates an executable which can then be run by either calling the executable directly or using a job scheduler such as `slurm`. This example highlights the irreducible pieces of a simulation. More complicated parts can be added by the user for more complex systems.

3.6.2 BASE PROCA

The next example represents a step up in complexity by introducing a mass term to the base Lagrangian, $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}\mu^2 A^\mu A_\mu$. This is actually a simplified case of the most simple generalized Proca theory. This theory corresponds to setting $G_2 = \frac{1}{2}A^\mu A_\mu$, $\alpha_2 = -1$, and $\alpha_{i>2} = 0$. `GRBoondi` contains built-in functions to handle the simple case when $\mathcal{L}_2 = \alpha G_2(A^\mu A_\mu)$, for

```

1 template <class G2 = DefaultG> class L2
2 {
3 public:
4     struct params_t
5     {
6         double alpha2;
7     };
8     ...
9 public:
10    L2(){}; // Default constructor for default initialization in matter class
11    L2(G2 a_G2_function, params_t a_params): m_g2_function(a_G2_function), m_params{a_params} {};
12
13    template < ... > void compute_emptensor_modification( ... ) const;
14
15    template < ... > void matter_rhs_modification( ... ) const;
16 };

```

Listing 3.27: GRBoondi contains built-in functions to evolve generic definitions of $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \alpha_2 G_2(A^\mu A_\mu)$.

```

1 ProcaField(KerrSchild a_background, params_t a_params)
2 : BaseProcaField<KerrSchild, ProcaField>(a_background),
3   m_background(a_background), m_params(a_params)
4 { // set up the L2 Lagrangian
5
6     DefaultG::params_t G2_params{m_params.mass}; // Initialize G2 function parameters
7     L2_t::params_t L2_params{m_params.alpha2}; // Initialize L2 Lagrangian parameters
8
9     DefaultG a_G2(G2_params);
10    this->m_L2 = L2_t(a_G2, L2_params);
11    this->m_G2 = a_G2;
12 };

```

Listing 3.28: Using the built-in \mathcal{L}_2 functions is very straightforward, amounting to initializing the functions in the constructor of the ProcaField class.

generic choices of G_2 . This is achieved by templating the evolution equations and stress-energy equations over a generic definition of G_2 , allowing users to define any choice of mass term. GRBoondi contains a class called L2 which encapsulates all modifications to the equations of motion coming from such terms. Users then simply have to define their own G2 function or use the default built-in definition of $G_2 = A^\mu A_\mu$.

The base Proca example uses the default definition of the G_2 function and the ProcaField class is initialized and sets these functions. Then, the modification to the evolution equations is achieved by calling the built-in modification methods of the L2 class. Users can also add the evolution equa-

```

1 template < ... >
2 void matter_rhs_modification( ... ) const
3 {
4     // add modifications coming from L2 Lagrangian
5     m_L2.matter_rhs_modification(total_rhs, matter_vars, metric_vars, d1,d2, advec);
6     ...
7 }

```

Listing 3.29: Modifying the evolution equations from the \mathcal{L}_2 Lagrangian is as simple as calling the built-in function of the L2 class.

```

1 class Initial_Proca_Conditions
2 {
3     ...
4     template <class data_t> void compute(Cell<data_t> current_cell) const
5     {
6         ...
7         mattervars.Avec[0] = m_params.init_amplitude * exp(-radius / r0_BL) / gamma_det;
8         current_cell.store_vars(mattervars); // export to grid
9     }
10 };

```

Listing 3.30: Initial data for the Proca field is taken from analytic approximate computations of the Proca bound states.

tion for the damping term, which isn't automatically added by the L2 class, which is done in the base Proca example.

For the base Proca example, the initial data is taken from analytic approximate computations of the superradiant bound state for a Proca field, lst. 3.30, where $r0_BL$ is the approximate radius of the bound state. The energy density at a late time is displayed in fig. 3.12.

GRBoondi also contains another example called the NonLinearProcaKerBH example, which evolves the system originating from the Lagrangian $\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}\mu^2 A^\mu A_\mu - \frac{1}{4}\lambda\mu^2 (A^\mu A_\mu)^2$. Incorporating these new terms amounts to modifying the $G2$ function. In this example, a new $G2$ function is defined, called NonLinearG2, which defines the function itself and its first and second derivatives with respect to its arguments. The rest of the code follows almost identically to the base Proca example previously. This example highlights the great modularity of GRBoondi, making it extremely easy to modify the evolution equations coming from additional terms in the Lagrangian.

3.6.3 BASE PROCA IN A SPINNING, EXPANDING UNIVERSE

The final example is a slight modification of the base Proca example and highlights the modularity of GRBoondi with respect to the background spacetime*. For this example, a new background spacetime was built, called the Kerr-de Sitter spacetime, and the new class called KerrdeSitter. The key pieces of building a new background spacetime will be elucidated here†.

The first step in defining a new background class is specifying the parameters of the background. In this case, there are four new parameters, lst. 3.31. The struct also contains additional parame-

*At the time of writing, the background spacetime defined in this example has become a standard class in the source code of GRBoondi. Nonetheless, it highlights how users can easily define their own background spacetime and very easily incorporate the new class into existing code, taking advantage of the classes templated over the background class.

†For a discussion of the Kerr-de Sitter spacetime itself, see Chapter 4.

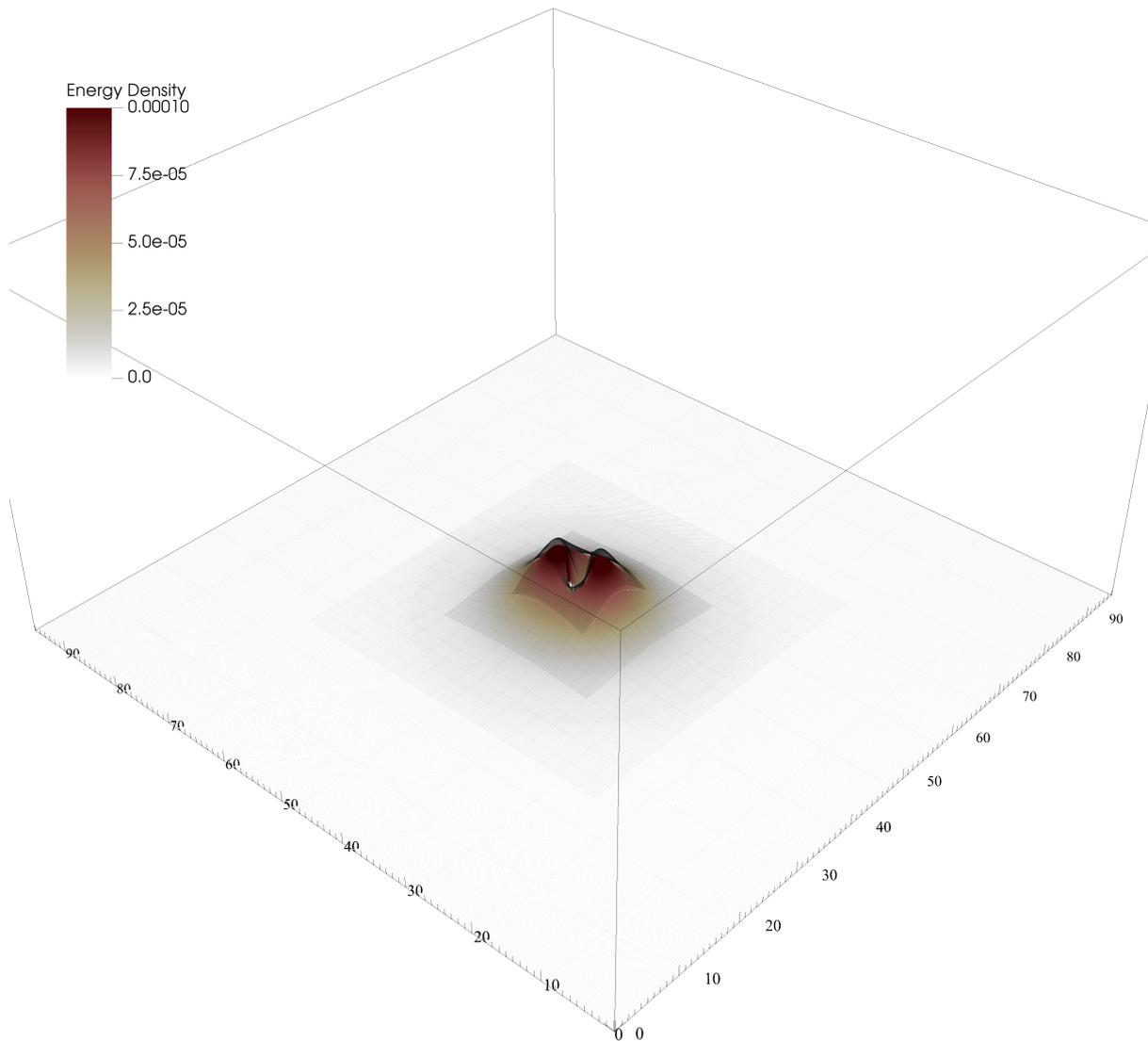


Figure 3.12: A snapshot of the energy density of a standard Proca field at late times in the $z = 0$ plane. The value of the energy density is represented both by the colored portion and a vertical extrusion. The AMR hierarchy is overlaid on this extrusion, highlighting the fact that the finest levels sufficiently cover the most dynamical parts of the simulation.

```

1 class KerrdeSitter
2 {
3   public:
4   struct params_t
5   {
6     double mass = 1.0;           //!< The mass of the BH in solar masses
7     double cosmo_constant = 0.0; //!< The cosmological constant
8     double spin = 0.0;           //!< The spin param a = J / M
9     std::array<double, CH_SPACEDIM> center; //!< The center of the BH
10    double r_plus = 0.0;          //!< The outer horizon. Precomputed
11    double r_minus = 0.0;         //!< The inner horizon. Precomputed
12  };
13  ...
14 };

```

Listing 3.31: The first step in defining a new background is specifying the relevant parameters.

```

1 template <class data_t>
2 void compute_metric_background(MetricVars<data_t> &metric_vars, const Coordinates<data_t> &coords) const
3 {
4     metric_vars.lapse = ...
5     FOR1(i)
6     {
7         metric_vars.shift[i] = ...
8         FOR1(j)
9         {
10            metric_vars.gamma[i][j] = ...
11        }
12
13        metric_vars.d1_lapse[i] = ...
14        ...
15    }

```

Listing 3.32: The new background class is required to have a method called `compute_metric_background` which computes the metric variables and its derivatives.

ters, the outer and inner horizon of the black hole, which are precomputed using GRBoondi’s utility functions and external computations. This is to prevent loss of performance due to expensive computations of the outer and inner horizon, which is quite involved.

The next step is to define the initializer. Typically, a default constructor can be used. However, in this example, a constructor was built that automatically verifies the user-defined parameters satisfy certain consistency checks, such as ensuring a naked singularity isn’t formed. Since the constructor is called at initialization time of the class, which occurs only once for each level method, it’s inexpensive to check the parameters at initialization time.

Finally, the background class should have a function called `compute_metric_background`, which computes the metric variables and their derivatives for a given co-ordinate point, lst. 3.32. The function signature of the method is required to be of this form, taking in a reference to the `MetricVars` class and `Coordinates` class.

The background class should also have a method called `check_if_excised`, which returns a boolean value telling the excision code if a cell should be excised or not depending on which region of the computational domain the cell is in. This is extremely important for black hole space times since the curvature at the center diverges and can cause significant numerical errors, spoiling the simulation. For singularity free backgrounds, such as the Minkowski background, the method can return `false` to never excise, lst. 3.33.

The rest of the code follows almost identically to the base Proca example, except every instance of `KerrSchild` is replaced with `KerrdeSitter`. For example, the definition of the `ProcaFieldLevel` class is shown in lst. 3.34. These are all the necessary ingredients to implement a new background

```
1 virtual bool check_if_excised(const Coordinates<double> &coords,  
2 const double buffer = 1.0) const  
3 {  
4     return false; // Dont ever excise  
5 }
```

Listing 3.33: New background classes are required to have a `check_if_excised` method, which computes where cells are within a domain where their values should be excised, such as within the black hole horizon. The example here shows a trivial implementation where excision never takes place.

```
1 class ProcaFieldLevel : public BaseProcaFieldLevel<KerrdeSitter, ProcaField>  
2 {  
3     ...  
4 }
```

Listing 3.34: Initialization of the `ProcaFieldLevel` class using a custom background class.

spacetime. All the features showcased in the examples highlight the power of GRBoondi to simulate generalized Proca theories in arbitrary spacetime backgrounds, greatly accelerating the study of the landscape of generalized Proca.

My goal is simple. It is a complete understanding of the universe, why it is as it is and why it exists at all.

Stephen Hawking

4

Proca in an Expanding Universe

IT IS CRUCIAL THAT ACCURATE MODELS ARE DEVELOPED for the evolution of superradiant Proca clouds around spinning black holes. Near future gravitational observatories are going to offer unprecedented tests for massive vector dark matter, hence accurate models are necessary to probe experimental data. Several theoretical analyses have already been carried out to understand the evolutionary behavior of massive vector fields surrounding black holes and their subsequent secular decay ^{199,282,283,289,292–295,297,301,303,346,347}. All of these studies explicitly neglect the cosmological constant, an overwhelming component of our universe. Normally, this is a justifiable assumption, since recent observational data has shown the cosmological constant to be quite small ^{136,348,349}. In terms of SI units, the observed cosmological constant is $\Lambda \sim 10^{-52} m^{-2}$, a tremendously small value. It is hence a reasonable approximation to take $\Lambda = 0$ in the studies of superradiant Proca

fields. Nonetheless, the effects of a cosmological constant on the growth rate of the Proca cloud is still required to complete the understanding of Proca superradiance in our universe.

The purpose of this chapter is to detail the study of Proca fields in spinning and expanding universes by specifying the background spacetime to be that of a Kerr black hole in a de Sitter background, the so-called Kerr-de Sitter solution. At first, the general theory of (standard) Proca fields in a Kerr-de Sitter background is discussed, elucidating the interesting causal properties of the background and general dynamical evolution of the Proca field. Next, as a first application of GRBoondi, the Proca evolution equations are solved in the Kerr-de Sitter background spacetime and the results discussed. Finally, as is standard for all numerical relativity computations, resolution and convergence tests are performed to verify the data converges to some continuous solution.

4.1 THEORY

The starting point is the Einstein-Hilbert-Proca theory, where the Einstein-Hilbert action is augmented by the Proca action

$$S[g_{\mu\nu}, A^\sigma] = \int d^4x \sqrt{-g} \left(\frac{1}{16\pi G} (\mathfrak{R} - 2\Lambda) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{2} \mu^2 A^\mu A_\mu \right), \quad (4.1)$$

where $g_{\mu\nu}$ is the metric, g its determinant, \mathfrak{R} is the Ricci scalar, Λ is the cosmological constant, and A^μ is the Proca four-vector. Variation of $S[g, A]$ with respect to these two fields yields the equations of motion

$$G^{\rho\sigma} + \Lambda g^{\rho\sigma} = 8\pi \left(\frac{-1}{4} F_{\mu\nu} F^{\mu\nu} g^{\rho\sigma} + F^{\rho\nu} F_\nu^\sigma - \frac{1}{2} \mu^2 g^{\rho\sigma} A_\mu A^\mu + \mu^2 A^\rho A^\sigma \right) = 8\pi \mathfrak{T}_{\mu\nu} \quad (4.2)$$

$$0 = \nabla_\rho F^{\rho\sigma} - \mu^2 A^\sigma, \quad (4.3)$$

where $G^{\rho\sigma}$ is the Einstein tensor and $\mathfrak{T}_{\mu\nu}$ is the stress-energy tensor of the Proca field. We assume the backreaction of the Proca field on the spacetime is negligible, as is usually the case, meaning we can set $\mathfrak{T}_{\mu\nu} = 0$ without significant loss of accuracy. This implies that the field equations reduce to

$$G^{\rho\sigma} + \Lambda g^{\rho\sigma} = 0 \quad (4.4)$$

$$\nabla_\rho F^{\rho\sigma} - \mu^2 A^\sigma = 0 . \quad (4.5)$$

The solution of the gravity sector describing a spinning black hole is called the Kerr-de Sitter (KdS) solution. This solution is in fact a special case of the more general Plebański-Demiański family of metrics^{350,351}. We choose to work in the Kerr-Schild form of the solution, which takes the form

$$g_{\mu\nu} = g_{0,\mu\nu} + 2HK_\mu K_\nu , \quad (4.6)$$

where $g_{0,\mu\nu}$ is the background de Sitter metric and K_μ is a null vector (with respect to both g and g_0). In Kerr-Schild co-ordinates (t, r, θ, ϕ) , the de Sitter background metric takes the form

$$g_{0,\mu\nu} = \begin{bmatrix} \frac{-\Delta_\theta}{\Theta} \Lambda_r & 0 & 0 & 0 \\ 0 & \frac{\rho^2}{(r^2+a^2)} \Lambda_r & 0 & 0 \\ 0 & 0 & \frac{\rho^2}{\Delta_\theta} & 0 \\ 0 & 0 & 0 & \frac{(r^2+a^2)}{\Theta} \sin^2 \theta \end{bmatrix} , \quad (4.7)$$

where we've defined

$$\begin{aligned} \Delta_\theta &= 1 + \frac{\Lambda}{3} a^2 \cos^2 \theta & \Delta_r &= r^2 - 2Mr + a^2 - \frac{\Lambda}{3} r^2 (r^2 + a^2) \\ \Theta &= 1 + \frac{\Lambda}{3} a^2 & \rho^2 &= r^2 + a^2 \cos^2 \theta \end{aligned} \quad (4.8)$$

and $H = \frac{2Mr}{\rho^2}$. The Δ_r definition becomes important for the analysis of the black hole horizons. The poles of Δ_r correspond to the poles of the metric in Boyer-Lindquist co-ordinates³⁵². In a similar fashion, the null vector is defined via

$$K_\mu = \left(\frac{\Delta_\theta}{\Theta}, \frac{\rho^2}{(r^2+a^2)\Lambda_r}, 0, \frac{-a}{\Theta} \sin^2(\theta) \right) . \quad (4.9)$$

Prior to plugging this metric into the numerical solver, we need to understand the basic causal structure of the background spacetime on which the Proca cloud will evolve. This amounts to determining the location of the horizons, which follows by solving the quartic polynomial $\Delta_r = 0$. The *existence* of the horizons can be determined by analyzing the discriminant of the quartic polynomial, which will be denoted by Q . The positivity of Q guarantees that all four roots are simulta-

neously either real or complex. This means we only have to verify one of the roots is real to ensure the other three are as well. The discriminant is easily solved for and takes the form

$$Q = -\frac{16\Lambda}{243} \left[12a^8\Lambda^8 + a^{10}\Lambda^4 + a^2(81 - 891M^2\Lambda) + 3a^6\Lambda^2(18 + M^2\Lambda) \right] \quad (4.10)$$

$$+ 81M^2(-1 + 9M^2\Lambda) + 27a^4\Lambda(4 + 11M^2\Lambda) \quad (4.11)$$

Fig. 4.1 shows a contour plot of Q , with the spin and cosmological constant rescaled by the black hole mass. The region $Q > 0$ shows the allowed parameters for the background. The shaded region denotes the disallowed region. Beyond the allowed region, one or more horizons will disappear, yielding a naked singularity. Fig. 4.1 hence tells us the allowed values of the black hole parameters that can be used in the simulations. An absolute maximum value of the cosmological constant is $\Lambda_{max} = \frac{16}{45+26\sqrt{3}} \frac{1}{M^2}$ and corresponding absolute maximum value of the black hole spin of $a_{max} = \sqrt{\frac{9}{16} + \frac{3\sqrt{3}}{8}} M$. This point corresponds to the cusp of the non-shaded region on the upper-right quadrant of fig. 4.1.

Another interesting feature is the existence of a minimum spin for a certain range of values for the cosmological constant. For spacetimes satisfying $\Lambda \geq 1/9$, the black hole is required to possess spin in order for the Kerr-de Sitter black hole to exist. That is to say, for cosmological constant values greater than $1/9$, static black holes in the form of eq. 4.6 do not exist.

We label the three positive roots as r_- , r_+ , and r_Λ , which denote the inner, outer, and cosmological horizons, respectively. The fourth root is negative and corresponds to a 'horizon' inside the singularity at $r = 0$. In Kerr-Schild co-ordinates, besides the singularity at $r = 0$, the poles in the metric disappear, including r_Λ . However, a new pole appears at $\tilde{r}_\Lambda = \sqrt{\frac{3}{\Lambda}}$. We find that $r_\Lambda \leq \tilde{r}_\Lambda$ throughout the allowed parameter space. In the limit of small cosmological constant,

$$r_\Lambda = \tilde{r}_\Lambda - 1 + \mathcal{O}(\sqrt{\Lambda}) \quad (4.12)$$

A peculiar property of the causal structure is the existence of a maximum cosmological constant which permits the existence of a black hole. At the maximum cosmological constant value Λ_{max} , the outer and cosmological horizons merge, see for example fig. 4.2. Additionally, dimensionless spin values greater than unity are allowed without producing a naked singularity, as long as the cos-

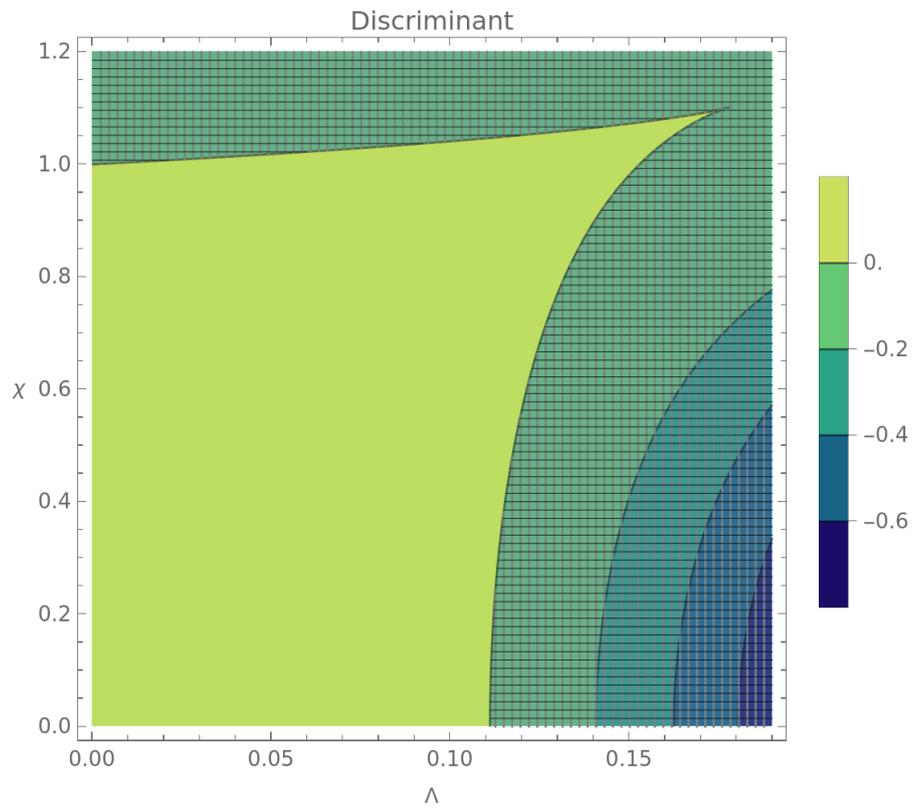


Figure 4.1: Plot of the values of the discriminant Q . The shaded region shows the unallowed parameters for the existence of a Kerr-de Sitter black hole. The unshaded region are the allowable parameters.

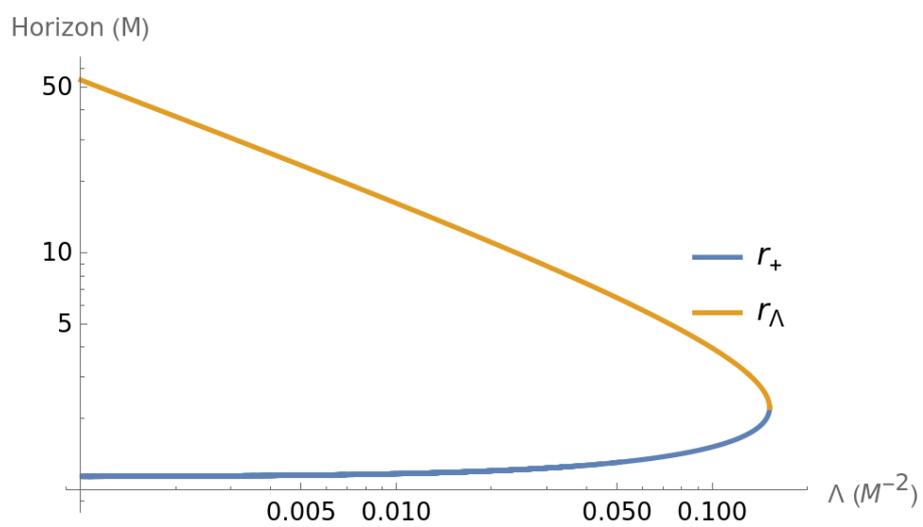


Figure 4.2: Evolution of the cosmological and outer horizons, eventually merging. A value of $a=0.5$ is chosen as a representative value. Past the value of Λ where the horizons merge, no Kerr-de Sitter black hole is possible.

mological constant is greater than zero, in stark contrast to spinning black hole solutions in non-expanding universes.

On the other hand, the solution in the Proca sector of eqs. 4.4-4.5 is unclear, since analytic studies have yet to be performed. However, based on knowledge from the Proca on Kerr solutions, we can expect the cloud to undergo a superradiant instability due to the existence of the black hole horizon. This implies that the energy of the cloud will be of the form

$$E(t) = E_0 e^{2i\omega t}, \quad (4.13)$$

where ω is the complex frequency of the Proca field. The factor of 2 comes from the fact that the stress-energy tensor of the Proca field is quadratic in the Proca field itself. Since the frequency is of the form $\omega = \omega_r + i\omega_i$, the imaginary part implies an exponential evolution superimposed over an oscillatory one. The imaginary frequency is what is solved for in this study.

4.2 METHOD

Towards a numerical solution of the Proca system, the next step is to decompose the field equations into a form pertinent for numerical computations. We follow the standard procedure, which is to decompose the spacetime via a foliation into a series of three-dimensional time-like hypersurfaces. Since the spacetime admits a time-like Killing vector, we can choose the Kerr-Schild time coordinate as the function that defines the foliation leaves. Hence, the (r, θ, ϕ) co-ordinates become co-ordinates on the hypersurfaces. We thus define our 3 + 1 decomposition as³²⁵

$$\alpha = -\frac{g}{\gamma} \quad (4.14)$$

$$\beta^i = -\frac{g^{0i}}{g^{00}} \quad (4.15)$$

$$\gamma_{ij} = g_{ij}, \quad (4.16)$$

where g and γ are the determinants of the full and spatial metrics, respectively, α is the lapse function, β^i is the shift vector, γ_{ij} is the spatial metric, and Latin indices range from one to three. In the

co-ordinate system of eq. 4.7, these take the form

$$\alpha = \sqrt{\frac{\Delta_\theta \Lambda_r^2 \rho^2}{\Gamma}} \quad (4.17)$$

$$\beta^i = \left[\begin{array}{ccc} \frac{2Mr(a^2+r^2)\Delta_\theta\Lambda_r}{2a^2Mr\sin^2(\theta)\Lambda_r+\Theta(2Mr+(a^2+r^2)\Lambda_r)\rho^2} & 0 & -\frac{2Mr a \Delta_\theta \Lambda_r}{2a^2Mr\sin^2(\theta)\Lambda_r+\Theta(2Mr+(a^2+r^2)\Lambda_r)\rho^2} \end{array} \right] \quad (4.18)$$

$$\gamma_{ij} = \left[\begin{array}{ccc} \frac{\rho^2(2Mr+(a^2+r^2)\Lambda_r)}{\Lambda_r^2(a^2+r^2)^2} & 0 & -\frac{2aMr\sin^2(\theta)}{a^2\Theta\Lambda_r+r^2\Theta\Lambda_r} \\ 0 & \frac{\rho^2}{\Delta_\theta} & 0 \\ -\frac{2aMr\sin^2(\theta)}{a^2\Theta\Lambda_r+r^2\Theta\Lambda_r} & 0 & \frac{\sin^2(\theta)}{\Theta^2} \left(\Theta(r^2+a^2) + \frac{2a^2Mr\sin^2(\theta)}{\rho^2} \right) \end{array} \right], \quad (4.19)$$

where we've defined $\Gamma = 2Mr\Delta_\theta + \rho^2\Theta\Lambda_r$. As our numerical solver computes the time evolution using cubic cells, we transform the (r, θ, ϕ) co-ordinates to a Cartesian-like co-ordinate system defined by

$$x = r \sin(\theta) \cos(\phi) \quad (4.20)$$

$$y = r \sin(\theta) \sin(\phi) \quad (4.21)$$

$$z = r \cos(\theta) . \quad (4.22)$$

The shift, spatial metric, and all derivatives are then transformed using the resulting Jacobian matrix. See App. B.1 for more details. The last step is to decompose the Proca equations eq. 4.5 under the foliation. A standard calculation yields

$$\frac{1}{\alpha} \mathcal{L}_m E^i = E^i K - D^i Z + \mu^2 X^i - \frac{2}{\alpha} D_j (\alpha D^{[j} X^{i]}) \quad (4.23)$$

$$\frac{1}{\alpha} \mathcal{L}_m Z = -\mu^2 \phi - D_i E^i - \kappa Z \quad (4.24)$$

$$\frac{1}{\alpha} \mathcal{L}_m X_i = -E_i - D_i \phi - \phi D_i \ln(\alpha) \quad (4.25)$$

$$\frac{1}{\alpha} \mathcal{L}_m \phi = \frac{Z}{\mu^2} + \phi K - D_i X^i - x^i D_i \ln(\alpha) , \quad (4.26)$$

where $E^i = \gamma_\mu^i F^{\mu\nu}$, $X_i = \gamma_{i\mu} A^\mu$, $\phi = -n_\mu A^\mu$, Z is an auxiliary field introduced to damp violations of the Proca constraint with a tuning parameter κ ^{216,283,353}, and n^μ is the time-like normal to the spatial hypersurfaces.

4.3 RESULTS

For the numerical evolution of eqs. 4.23-4.26, we use GRBoondi. We use a box width of $L = 60M$ with $N = 192$ grid points across each edge of the computational box. We use 4 refinements levels at a 2 : 1 refinement ratio, resulting in a resolution of the finest level of $dx_{fine} = 0.01953M$. To prevent boundary effects from contaminating the simulation, we use Sommerfeld-outgoing radiation boundary conditions, which allows oscillations to exit the simulation region with minimal reflections due to finite-size effects. This is especially important since we introduced an auxiliary field which dampens violations of the constraint equation and the evolution equation for Z is a generalized telegraph equation. This implies that not only are the values of Z damped, but also propagate at the speed of light. Hence, the outgoing radiation boundary conditions are vital for ensuring violations of the constraint equation propagate outside the computational domain.

To understand the effect of a cosmological constant on the dynamical evolution of the superradiant Proca cloud, we perform a various number of simulations with parameters that yield the highest growth rates. We choose three different values of the cosmological constant, $\Lambda = (5 \cdot 10^{-6}, 10^{-4}, 10^{-3})$. Higher values of the cosmological constant are more difficult to simulate numerically as the cosmological horizon quickly becomes small. We reserve probing this region of the parameter space to future studies, which will likely entail a new co-ordinate system. Additionally, we fix the black hole spin to $\chi = 0.99$. We sample the Proca mass at six different values, $\mu = (0.35, 0.4, 0.45, 0.5, 0.6, 0.7)$. In addition to the main simulations, we also perform a convergence study to ensure our choice of resolution produces accurate data, which we discuss in sec. 4.3.1.

For initial data, we take a Gaussian profile with width determined by analytic approximation studies¹⁹⁹, $r_0 = \frac{1}{M^2\mu}$. The initial data is then $A_x = \frac{A}{\gamma} e^{-\frac{r}{r_0}}$, where A is some pre-determined amplitude which we take to be $A = 0.1$, and all other variables are chosen to be zero.

The data from our simulations is available in table 4.1. The cosmological constant has been rescaled to $\Lambda = \frac{\bar{\Lambda}}{M^2}$, where $\bar{\Lambda}$ is the unscaled parameter. Plots of the normalized total energy as a function of time are shown in fig. 4.3. A curious new feature seems to arise in the case of decaying modes, namely a secondary scale. For example, in the case of $\mu = 0.6$ and $\Lambda = 5 * 10^{-6}$, the decay rate of the total energy slows at around $t = 4000M$. Similar features can be seen in the

$\mu = 0.7/\Lambda = 10^{-3}$ plot as well as $\mu = 0.6/\Lambda = 10^{-4}$. Whether this is a numerical artifact or a real emerging scale is uncertain. Analytic studies will need to be performed to determine the true nature of this emerging scale. Due to the variations in the $\mu = 0.35/\Lambda = 10^{-3}$ simulation data, such a feature cannot be determined. It should be noted that the rapid oscillations and cessation of decay in the $\mu = 0.35/\Lambda = 10^{-3}$ plot is likely purely numerical errors. As the energy of the Proca cloud becomes incredibly small, it likely runs in to a precision floor and the simulation becomes numerically meaningless past $t \sim 2000M$.

Simulation Data		
μ	Λ	ω_i
0.35	0.000005	0.0000378479679286
0.40	0.000005	0.0004368389589392
0.45	0.000005	0.0006612210382883
0.50	0.000005	0.0007077262886612
0.60	0.000005	-0.0025030874787604
0.70	0.000005	-0.0003803920672232
0.35	0.0001	-0.0000509955398002
0.40	0.0001	0.0004173920772338
0.45	0.0001	0.0006586744541765
0.50	0.0001	0.0007037383784603
0.60	0.0001	-0.0027137433850122
0.70	0.0001	-0.0007733224329400
0.35	0.001	-0.0086145055245226
0.40	0.001	-0.0017417580794997
0.45	0.001	0.0002659457748499
0.50	0.001	0.0006576997786869
0.60	0.001	-0.0027797013519735
0.70	0.001	-0.0025289981313056

Table 4.1: All available simulation data.

4.3.1 CONVERGENCE TESTS

To ensure the reliability and accuracy of our simulations, we conduct two types of convergence tests. These tests are crucial for validating the fidelity of the simulated data and confirming that it accurately reflects the underlying physical system. The first type of test we perform is the grid resolution test. By running our simulations at multiple grid resolutions, we can verify that the results converge to a stable solution as the grid is refined. This process helps us identify and minimize numerical artifacts that might arise from discretization errors. Specifically, we monitor key physi-

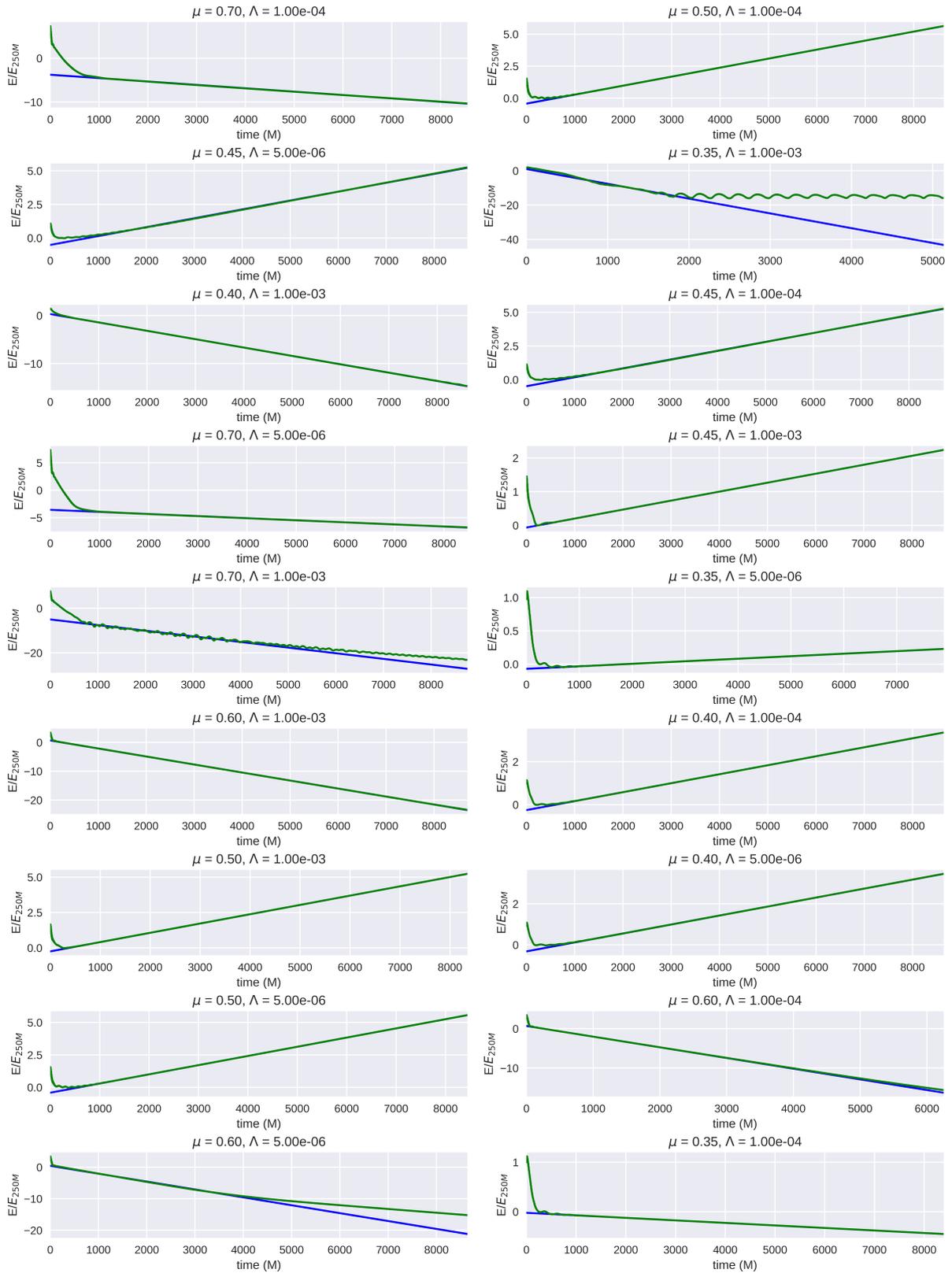


Figure 4.3: Growth of the energy of the cloud over simulation time. Energy values are normalized to their value at $t = 250M$. The energy data is fitted in logarithmic space to a linear function, which captures the exponential characteristics of the instability. The fitting of the data starts at $t = 250M$. The green line is the numerical data and the blue straight line is the fit function.

cal quantities and ensure that their values stabilize with increasing grid resolution, indicating that our results are not dependent on the grid size. The second type of test is the time step refinement test. Here, we run our simulations with progressively smaller time steps to ensure that the temporal evolution of the system is accurately captured. By comparing results obtained with different time steps, we can check for stability and convergence in the time integration scheme. This test is particularly important for dynamic simulations where rapid changes in the system need to be accurately resolved. In order to maintain the Courant–Friedrichs–Lewy condition, the temporal timestep decreases proportionally with the spatial timestep. Thus, the two checks are unified in our resolution tests. Below, we provide detailed insights into the methodologies employed for conducting these two tests and assess the reliability of our simulations based on their outcomes.

RESOLUTION TEST

The first convergence test is a resolution test. This ensures that finite-differencing effects have a diminishing effect on the data as the resolution is turned up. In other words, the resulting data should converge to stable values as the resolution is increased. After some point, increasing the resolution further should have little effect on the data. This ensures that the resolution we have chosen for our simulations is satisfactory enough to produce good data.

The results of this convergence test are as follow: we repeat two of the simulations with parameters $(\mu, \Lambda) = (0.4, 5 * 10^{-6})$ and $(0.4, 10^{-3})$ with a resolution on the finest level of $dx_{fine} = 0.0167M$. Since the Courant-Friedrichs-Lewy condition is automatically adjusted for in GRBoondi, the corresponding time resolution is $dt_{fine} = 0.0033M$, an increase in resolution from the main data, which utilized $dt_{fine} = 0.0039M$. It was found that the instability rate for both the growing and decay mode change by less than a percent, signifying good precision for the main data. The two simulations and their differences are shown in fig. 4.4.

ANALYTIC DERIVATIVES TEST

The second convergence test is a self-check test on the metric derivatives. Since the procedure in transforming the metric variables and their derivatives from Kerr-Schild co-ordinates to a Cartesian-like co-ordinate system is involved, we run a check to make sure the numerically computed deriva-

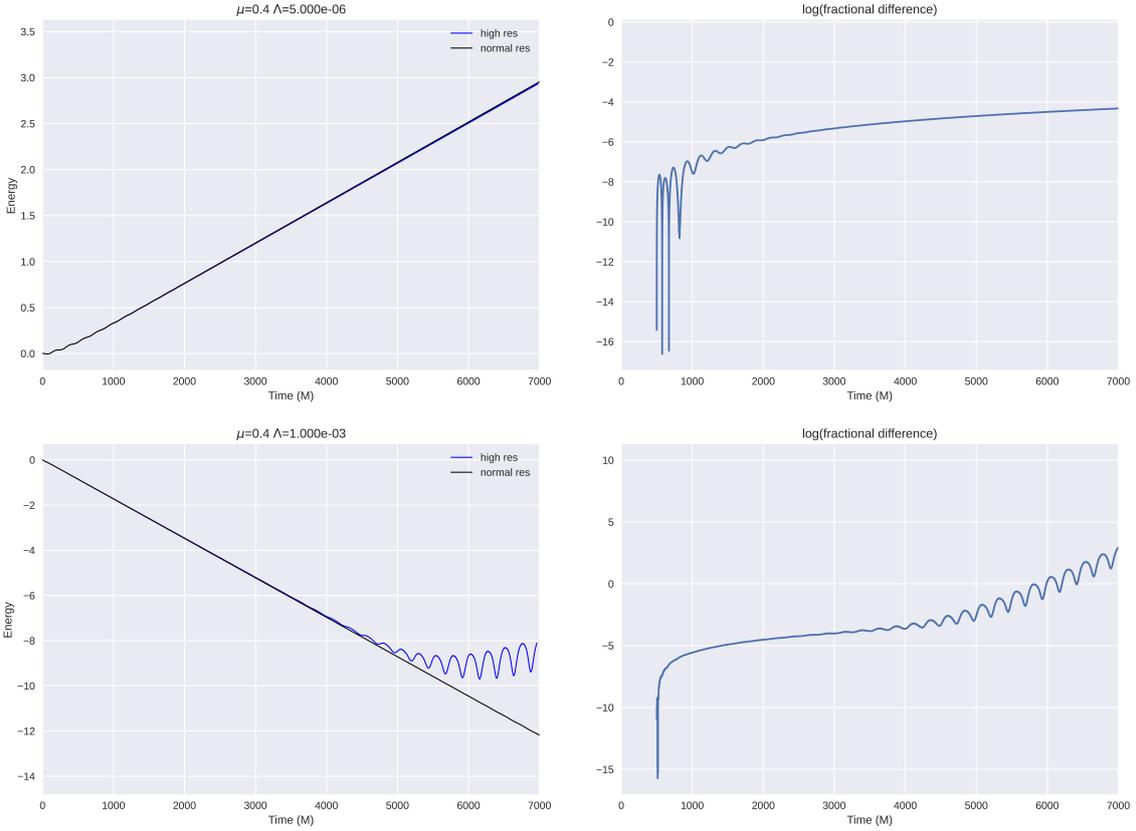


Figure 4.4: Test of the convergence of the simulations as resolution is increased. The normal resolution is the resolution used for the main data, which uses $N = 192$ grid points. The high-resolution simulations use $N = 224$ grid points, resulting in a resolution on the finest level of $dx_{fine} = 0.0167M$. The fractional difference is computed as $|E_{norm.res.} - E_{highres.}|/E_{norm.res.}$. The two energy datasets were interpolated over a common time dataset using cubic splines.

tives converge to the analytically computed ones from eqs. B.3-B.4. We follow the procedure of 354 for carrying out this convergence test.

The test procedure is graphically represented in fig. 4.5. The convergence factor $c(t)$ is defined as

$$c_i(t) \equiv \frac{\|\epsilon_{\Delta_i}\|}{\|\epsilon_{\Delta_{i+1}}\|} \quad (4.27)$$

and Δ_i denotes the i 'th resolution, $\epsilon_{\Delta_i} = U(t, x) - U_{\Delta_i}(t, x)$, $U(t, x)$ denotes the variable computed using the exact analytic derivatives, and $U_{\Delta_i}(t, x)$ denotes the variable computed using the numerically computed derivatives at a resolution of Δ_i . The resolutions are typically chosen to be twice the previous resolution. In other words, $\Delta_i = 2^i \Delta_1$. This implies that for an n th-order finite differencing scheme,

$$c(t) = 2^n. \quad (4.28)$$

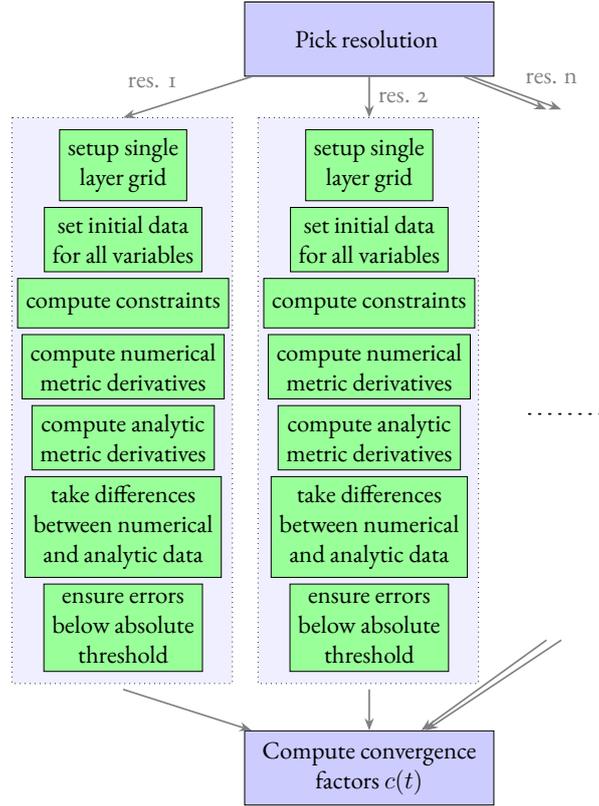


Figure 4.5: Graphical representation of the convergence test routine. Multiple resolutions are iterated over to compute the convergence factors $c_i(t)$.

Since GRBoondi uses fourth-order finite differencing stencils for the spatial derivatives, we expect $c(t) = 16$ for each variable. We perform the convergence test for two different resolutions, which differ by a multiple of two per the previous discussion. We find that the minimum convergence factor across all grid variables was 14.35, in fairly good agreement with the expected convergence factor of 16 for a fourth-order finite differencing routine.

4.4 CONCLUSION AND FUTURE WORK

This study investigated the dynamics of a Proca field surrounding a spinning black hole within an expanding universe, a novel approach that advances our understanding of superradiant vector fields interacting with the cosmological constant—a crucial yet overlooked aspect in prior research. Leveraging the advanced capabilities of the recently developed GRBoondi software, we conducted a series of simulations aimed at quantifying the growth rates under diverse scenarios involving different Proca masses and cosmological constants. Furthermore, we conducted a rigorous resolution analysis to validate the accuracy and reliability of our simulation data, revealing robust agreement with

expected theoretical predictions.

There are a few limitations of the current study. Firstly, the effects of backreaction were ignored, however this approximation is suitable for this analysis as the energy densities were small and gravitational radiation was not desired. Should gravitational emission from the resulting Proca cloud be desired, full numerical computations will need to be performed. Due to GRBoondi outputting checkpoint files in the same format as GRChombo, using GRChombo to perform full computations with initial data from the simulations here will be straightforward. Additionally, many parts of the parameter space were left unsampled, due to the numerical complexities. Primarily, simulating larger cosmological constants is difficult due to the cosmological horizon rapidly approaching the outer horizon of the black hole. Additionally, simulating lower spins increases the simulation time considerably, hence we only focused on a single value of the black hole spin, $\chi = 0.99$. A more complete analysis would likely require a different co-ordinate choice that penetrates the cosmological horizon.

Future studies to be performed include theoretical analyses of various co-ordinate systems in the search of ones that can penetrate the outer and cosmological horizons. Additionally, future studies will turn to full evolution of the Einstein+Proca system to compute the emitted gravitational radiation and apply the results to gravitational observatory forecasting. An intriguing progression would involve extending beyond the Proca fields mass term to incorporate derivative self-interactions inherent in generalized Proca theories. This expansion would enhance our exploration of the field's dynamics, encompassing interactions that go beyond simple mass considerations and delve into the complexities introduced by derivative couplings within these theories.

5

Conclusion

THE ONSET OF MULTIMESSENGER ASTRONOMY has opened up new channels for humans to peer into the universe. Next generation gravitational observatories will provide excellent data to study the nature of dark matter and dark energy, perhaps providing indirect detections of such mysterious phenomena. Strong gravity regimes will hence provide an excellent laboratory to study fundamental physics. In particular, the superradiant phenomenon is an invaluable tool to study dark matter and dark energy models, especially the landscape of generalized Proca theories. To that end, highly accurate models will need to be developed to facilitate data analysis pipelines. Without these accurate models, searching the data for possible detections is analogous to roaming around a dark room hoping to find something you know nothing about.

The work in this thesis took steps towards developing such models by placing constraints on the

possible mass range of Proca fields detectable by LISA, developing powerful tools for studying generalized Proca models and, for the first time, analyzing the effect of a cosmological constant on the superradiant growth rate of Proca fields. The detection prospects of LISA utilized the dephasing effect of a gravitational wave induced by the presence of a massive vector field around a supermassive spinning black hole. That study placed constraints on the mass parameter space of the Proca field observable by LISA.

In the pursuit of models for generalized Proca theories, no unified tool existed in the literature for studying generalized Proca theories in general. This inspired the creation of GRBoondi, a versatile and powerful tool for simulating the evolution of any generalized Proca theory in an arbitrary fixed background. The broad field of numerical relativity and the finer details of GRBoondi were discussed, illustrating its modularity and prowess. The example systems that ship with GRBoondi were also discussed, elucidating the key properties that make GRBoondi a powerful tool for studying generalized Proca theories.

The first results using GRBoondi were then detailed by studying a standard Proca field in a spinning and expanding universe. This was a first of its kind study in two respects. Primarily, it was the first study to utilize GRBoondi, demonstrating its modularity with respect to the background spacetime. It was a simple task to adjust the built-in examples to use a new background spacetime class. Secondly, it is the first study to incorporate the cosmological constant into the evolution of a superradiant Proca field around a spinning black hole. It was shown that the introduction of a non-zero cosmological constant can have tremendous impacts on the evolution. States that originally are exponentially growing in the $\Lambda = 0$ case transform to exponentially decaying states in the $\Lambda > 0$ case. This has huge impacts on the phenomenology, providing another avenue to search for massive vectorial dark matter, especially in more general dark energy scenarios that vary both spatially and temporally.

While the mysteries of the cosmos remain vast and enigmatic, the insights gained in this thesis illuminates new pathways for exploration, promising a deeper understanding of the universe and our place within it.

Equipped with his five senses, man explores the universe around him and calls the adventure Science.

-- Edwin Powell Hubble

A

Additional Details on Superradiant Proca Clouds around EMRIs

This appendix contains further details pertaining to Section 2 that are not relevant for the main text but aid in further understanding of the finer details of the study. Section A.1 provides details on the decomposition of the Proca equations of motion in the FKKS ansatz. Section A.2 discusses the long-term evolution of the Proca cloud due to the emission of gravitational waves sourced by the cloud itself. Section A.3 discusses other physical effects that could alter the results of Section 2.

Angular Equation

Multiply the angular equation eq. 2.19 by q_θ and define $\Lambda = \frac{\mu^2}{\nu^2} - \frac{\sigma}{\nu} + 2a\omega m - a^2\omega^2$ and $\gamma^2 = \omega^2 - \mu^2$, where we've denoted the mode number by m and no longer by \mathbf{m} as in the main text. Rearranging the terms, one finds

$$q_\theta \left(\partial_\theta^2 + \cot \theta \partial_\theta - \frac{m^2}{\sin^2 \theta} + \Lambda \right) S + \left((\gamma^2 - 2\sigma\nu) a^2 \cos \theta^2 - \gamma^2 \nu^2 a^4 \cos \theta^4 - 2a^2 \nu^2 \cos \theta \sin \theta \partial_\theta \right) S = 0. \quad (\text{A.1})$$

Now we expand the angular variable S in terms of functions proportional to the associated Legendre polynomials as

$$S = \sum_{l'=|m|} b_{l'} Y_{2l'-|m|+\eta}^m(\theta), \quad (\text{A.2})$$

where $Y_{l'}^m(\theta) = Y_{l'}^m(\theta, 0)$, and we expand the function in a basis with definite parity as the angular equation respects parity. We then insert this into the angular equation above.

Using various relations of the associated Legendre polynomials in the first term of eq. A.1, one can easily show

$$q_\theta \left(\partial_\theta^2 + \cot \theta \partial_\theta - \frac{m^2}{\sin^2 \theta} + \Lambda \right) S \rightarrow q_\theta (-l'(l'+1) + \Lambda) Y_{l'}^m, \quad (\text{A.3})$$

where we dropped the $b_{l'}$ coefficient. We operate on eq. A.1 with $\int d\Omega \bar{Y}_l^m$ and define the quantity $\langle lm|X|l'm \rangle \equiv \int d\Omega \bar{Y}_l^m X Y_{l'}^m$. Then we have the expression

$$\begin{aligned} 0 = & (-l'(l'+1) + \Lambda) \langle lm|l'm \rangle + \\ & ((\gamma^2 - 2\sigma\nu)a^2 - a^2\nu^2(-l'(l'+1) + \Lambda)) \langle lm|\cos \theta^2|l'm \rangle - \\ & 2a^2\nu^2 \langle lm|\cos \theta \sin \theta \partial_\theta|l'm \rangle - \\ & \gamma^2\nu^2 a^4 \langle lm|\cos \theta^4|l'm \rangle. \end{aligned}$$

Representing the trigonometric functions in a spherical harmonic basis, we can easily calculate

$$\begin{aligned}
\langle lm|l'm\rangle &= \delta_{ll'} \\
a_{ll'} &\equiv \langle lm|\cos\theta^2|l'm\rangle = \frac{1}{3}\sqrt{\frac{16\pi}{5}}\langle l|2|l'\rangle + \frac{\sqrt{4\pi}}{3}\langle l|0|l'\rangle \\
b_{ll'} &\equiv \langle lm|\cos\theta^4|l'm\rangle = \frac{16\sqrt{\pi}}{105}\langle l|4|l'\rangle + \frac{7\sqrt{4\pi}}{35}\langle l|0|l'\rangle + \frac{10}{35}\sqrt{\frac{16\pi}{5}}\langle l|2|l'\rangle \\
d_{ll'} &\equiv \langle lm|\cos\theta\sin\theta\partial_\theta|l'm\rangle = \sqrt{\frac{4\pi}{3}}\frac{l'\sqrt{(l'+1)^2-m^2}}{\sqrt{(2l'+1)(2l+3)}}\langle l|1|l'+1\rangle - \\
&\quad \sqrt{\frac{4\pi}{3}}\frac{(l+m)(\sqrt{l'^2-m^2})}{\sqrt{(2l'+1)(2l-1)}}\langle l|1|l'-1\rangle,
\end{aligned}$$

where $\langle l1, l2, l3\rangle$ denotes the triple product integral $\int d\Omega Y_{l1}^m Y_{l2}^m Y_{l3}^m$ and can be represented in terms of the 3J-symbols. Thus, the angular equation, after reinserting the $b_{l'}$ coefficients, becomes

$$\mathcal{M}_{ll'} b_{l'} = 0, \quad (\text{A.4})$$

where

$$\mathcal{M}_{ll'} = (\Lambda - l'(l'+1))\delta_{ll'} + (\nu^2(l'(l'+1) - \Lambda) - 2\sigma\nu + \gamma^2) a^2 a_{ll'} - \gamma^2 \nu^2 a^4 b_{ll'} - 2a^2 \nu^2 d_{ll'}. \quad (\text{A.5})$$

This implies that $b_{l'}$ lives in the kernel of the map \mathcal{M} . For there to be nontrivial solutions to this eigenvalue equation, we require $\text{Det}(\mathcal{M}) = 0$, which places restrictions on the complex-valued eigenvalue ν . In general, the complex eigenvalue ν depends on μ, ω, l , and a . A generic solution can be obtained numerically, though insight can be found by taking suitable limits.

STATIC LIMIT: In the limit of staticity ($a \rightarrow 0$), the map takes the diagonal form

$$\mathcal{M} = (\Lambda - l'(l'+1))\delta_{ll'}. \quad (\text{A.6})$$

The nontrivial solution constraint then enforces

$$\Lambda = l'(l'+1), \quad (\text{A.7})$$

which, in terms of the eigenvalue ν , yields

$$\mu^2 - \omega\nu - l'(l' + 1)\nu^2 = 0 . \quad (\text{A.8})$$

We thus find the solutions

$$\nu = \begin{cases} \frac{\mu^2}{\omega} & l' = 0 \\ -\frac{\omega \pm \sqrt{\omega^2 + 4\mu^2 l'(l'+1)}}{2l'(l'+1)} & l' > 0 \end{cases} . \quad (\text{A.9})$$

We have a pair of modes for $l' > 0$ and a single (even-parity) mode for $l' = 0$. For the latter, the associated eigenvector is a single spherical harmonic and is thus even-parity under a parity transformation, while for the former we have a pair of modes. As the angular equation respects parity, this pair have the same parity. Thus, we've found the even-parity solutions.

MARGINALLY-BOUND CASE: $\gamma^2 = 0$ Consider now the case $\omega^2 = \mu^2$. This is the threshold between a quasibound mode and an unbound state. The matrix \mathcal{M} is tridiagonal, as can be seen using properties of the 3J-symbols. Consider the truncated series

$$S = Y_l^m + b_1 Y_{l+2}^m + 0 * Y_{l+4}^m , \quad (\text{A.10})$$

where $l = |m| + \eta$ and $\eta = 0, 1$ denotes the parity.

In general, we have three equations coming from

$$\mathcal{M} \cdot b \quad (\text{A.11})$$

with $b = (1, b_1, 0)^T$.

In the $\eta = 0$ case, we have $m = \pm l$. The solution $S = Y_l^{\pm l}$ with $\nu = \frac{\pm\omega}{m-a\omega}$ is an exact solution since $\Lambda = l(l+1)$, $\sigma = \pm m\nu$. It can be shown this solution together with this eigenvalue solves the angular equation by direct insertion. It can also be shown this eigenvalue corresponds to the $S = -1$ polarization state of the Proca field.

In the case $\eta = 1$ and $b_1 = 0$, we now have two nontrivial equations, with $m = \pm(l-1)$. It can

be shown by direct computation that the corresponding eigenvalues are

$$\nu = \frac{1}{2a} \left(\pm l - a * \omega + \epsilon \sqrt{(\mp l + a\omega)^2 + 4a\omega} \right) . \quad (\text{A.12})$$

These correspond to both parity-odd and parity-even polarization states, as shown in²⁹⁴.

For $\eta = 0$ and $m = 1$, the last polarization state can be recovered by finding the middle root of the following cubic expression²⁹⁴

$$a\nu^3(m - a\omega) - \nu^2((m + 1)(m + 2) - a\omega(2m - a\omega)) + \omega\nu + \omega^2 = 0 . \quad (\text{A.13})$$

Radial equation

The radial equation eq. 2.18 at asymptotic infinity can be shown to reduce to

$$\left(\left(1 - \frac{r_s}{r} \right) \partial_r \partial_r + \frac{\omega^2}{1 - \frac{r_s}{r}} - \mu^2 \right) R(r) = 0 . \quad (\text{A.14})$$

Remarkably, there is an exact solution in terms of Whittaker M and W functions. The exact solution is

$$R(r) = c_1 M_{\xi, \chi} [2(r_s - r)Q] + c_2 W_{\xi, \chi} [2(r_s - r)Q] , \quad (\text{A.15})$$

where $\xi = \frac{r_s(\mu^2 - 2\omega^2)}{2Q}$, $\chi = \frac{-i\sqrt{-1 + 4r_s^2\omega^2}}{2}$, $c_{1,2}$ are constants, $Q \equiv \sqrt{\mu^2 - \omega^2}$, and $W_{\xi, \chi}[x]$, $M_{\xi, \chi}[x]$ are the Whittaker W and M functions. Asymptotically, the solution takes the form

$$R(r \rightarrow \infty) = r^{\frac{(2\omega^2 - \mu^2)M}{Q}} e^{-Qr} , \quad (\text{A.16})$$

assuming vanishing boundary conditions at infinity, as employed later.

Now, we rewrite the derivatives of the radial equation in terms of the tortoise co-ordinate, $\frac{\partial r^*}{dr} = \frac{r^2 + a^2}{\Delta}$. One finds

$$\left(\frac{(r^2 + a^2)^2}{q_r \Delta} \partial_{r^*}^2 + \frac{K_r^2}{q_r \Delta} \right) R(r) + \left(\frac{2 - q_r}{q_r^2} \frac{\sigma}{\nu} - \frac{\mu^2}{\nu^2} + \frac{2r(2 - q_r)}{q_r^2} \partial_{r^*} \right) R(r) = 0 . \quad (\text{A.17})$$

Multiplying by $q_r \Delta$,

$$((r^2 + a^2)^2 \partial_{r^*}^2 + K_r^2) R(r) + q_r \Delta \left(\frac{2 - q_r}{q_r^2} \frac{\sigma}{\nu} - \frac{\mu^2}{\nu^2} + \frac{2r(2 - q_r)}{q_r^2} \partial_{r^*} \right) R(r) = 0. \quad (\text{A.18})$$

In the limit $r \rightarrow r_+$, $\Delta = (r - r_+)(r - r_-) \rightarrow 0$, and so the above differential equation reduces to

$$((r_+^2 + a^2)^2 \partial_{r^*}^2 + K_{r_+}^2) R(r) = 0, \quad (\text{A.19})$$

assuming $\partial_{r^*} R(r)$ is finite at the outer horizon. Taking $m = 0$, this reduces K_{r_+} to $-(r_+^2 + a^2)\omega$.

The ODE is readily solved, yielding

$$R(r \rightarrow r_+) = e^{-i\omega r^*}. \quad (\text{A.20})$$

If we take a generic m , then the asymptotic form is

$$R(r \rightarrow r_+) \sim e^{-r^* \sqrt{\frac{m(a\omega - m)}{r_s r_+} + \frac{1}{r_s^2} (m - r_s \omega)(m + r_s \omega)}} = e^{-ir^* \frac{K_{r_+}}{(r_+^2 + a^2)}}. \quad (\text{A.21})$$

ASYMPTOTIC FORM The two asymptotic forms of the radial equation are then

$$R(r) = \begin{cases} e^{-i\omega r^*} & r \rightarrow r_+ \\ r^{\frac{(2\omega^2 - \mu^2)M}{Q}} e^{-Qr} & r \rightarrow \infty \end{cases} \quad (\text{A.22})$$

with $Q \equiv \sqrt{\mu^2 - \omega^2}$.

Write out the tortoise co-ordinate in terms of the Boyer-Lindquist radial co-ordinate to find

$$r^* = r + \frac{r_s r_+}{r_+ - r_-} \ln \frac{r - r_+}{r_s} - \frac{r_s r_-}{r_+ - r_-} \ln \frac{r - r_-}{r_s} \quad (\text{A.23})$$

with $r_s = 2 * M$ and r_{\pm} the inner and outer horizon radii in Boyer-Lindquist co-ords. Then we rewrite the asymptotic form near the horizon as

$$e^{-i\omega r^*} = e^{-i\omega r} \left(\frac{r - r_+}{r_s} \right)^{-i\omega \frac{r_s r_+}{r_+ - r_-}} \left(\frac{r - r_-}{r_s} \right)^{i\omega \frac{r_s r_-}{r_+ - r_-}}. \quad (\text{A.24})$$

We see there is a pole at $r = r_+$. We can then expand this function in terms of a generalized power series as

$$R(r) = x^{-i\kappa}(r_0 + r_1x + r_2x^2 + \dots) \quad (\text{A.25})$$

with $x \equiv \frac{r-r_+}{r_+-r_-}$ and $\kappa = \frac{\omega r_s r_+}{r_+-r_-}$, or more concisely,

$$R(r) = \sum_{n=0}^{\infty} r_n x^{n-i\kappa} . \quad (\text{A.26})$$

The coefficients r_n can be determined by inserting the above into the radial equation. We first recast the radial equation into the form

$$\partial_r^2 R(r) + \left(\frac{q_r}{\Delta} \left(\frac{2r - r_s}{q_r} - \frac{2r\nu^2\Delta}{q_r^2} \right) \right) \partial_r R(r) + \left(\frac{q_r}{\Delta} \left(-\frac{\mu^2}{\nu^2} + \frac{K_r^2}{\Delta q_r} + \frac{(2 - q_r)\sigma}{q_r^2\nu} \right) \right) R(r) = 0 , \quad (\text{A.27})$$

which is of the form

$$\partial_r^2 R(r) + P(r; \lambda) \partial_r R(r) + Q(r; \lambda) R(r) = 0 , \quad (\text{A.28})$$

where λ are all the other parameters, i.e. ν, μ , etc, and

$$P(r; \lambda) = \frac{1}{(r - r_+)} \left(\frac{q_r}{(r - r_-)} \left(\frac{2r - r_s}{q_r} - \frac{2r\nu^2\Delta}{q_r^2} \right) \right) = \frac{1}{(r - r_+)} * \tilde{P}(r; \lambda)$$

$$Q(r; \lambda) = \frac{1}{(r - r_+)^2} \left(\frac{q_r}{(r - r_-)^2} \left(-\frac{\mu^2}{\nu^2}\Delta + \frac{K_r^2}{q_r} + \Delta \frac{(2 - q_r)\sigma}{q_r^2\nu} \right) \right) = \frac{1}{(r - r_+)^2} \tilde{Q}(r; \lambda) .$$

It's clear then that, as $\Delta = (r - r_+)(r - r_-)$, then the $P(r; \lambda)$ diverges like $\frac{1}{r-r_+}$ and $Q(r; \lambda)$ diverges as $\frac{1}{(r-r_+)^2}$. Hence, $r - r_+$ is a pole of order 1 for $P(r; \lambda)$ and a pole of order 2 for $Q(r; \lambda)$ and both $\tilde{P}(r; \lambda)$ and $\tilde{Q}(r; \lambda)$ are regular at $r = r_+$. Thus, by Fuch's theorem, we can apply the method of Frobenius to find a power series solution of the form eq. A.26 for the radial function near the outer horizon. Define the following functions

$$\tilde{P}_1(r; \lambda) = \frac{2r - r_s}{(r - r_-)}$$

$$\tilde{P}_2(r; \lambda) = -\frac{2r\nu^2}{q_r}$$

$$\tilde{Q}_1(r; \lambda) = \frac{K_r^2}{(r - r_-)^2}$$

$$\tilde{Q}_2(r; \lambda) = \frac{q_r}{(r - r_-)} \left(\frac{2 - q_r \sigma}{q_r^2 \nu} - \frac{\mu^2}{\nu^2} \right)$$

so that

$$\tilde{P} = \tilde{P}_1 + (r - r_+) \tilde{P}_2$$

$$\tilde{Q} = \tilde{Q}_1 + (r - r_+) \tilde{Q}_2 .$$

The differential equation then becomes

$$\partial_r^2 R(r) + \frac{\tilde{P}_1}{r - r_+} \partial_r R(r) + \tilde{P}_2 \partial_r R(r) + \frac{\tilde{Q}_1}{(r - r_+)^2} R(r) + \frac{\tilde{Q}_2}{(r - r_+)} R(r) = 0 . \quad (\text{A.29})$$

Inserting the expansion eq. A.26, one finds

$$0 = \sum_{n=0} \left(r_n (n - \kappa) (n - \kappa - 1) x^{n-2-\kappa} + \tilde{P}_1 r_n (n - \kappa) x^{n-2-\kappa} + \right.$$

$$\left. \tilde{Q}_1 r_n x^{n-\kappa-2} + \tilde{P}_2 r_n (n - \kappa) x^{n-\kappa-1} + \tilde{Q}_2 r_n x^{n-\kappa-1} \right)$$

$$0 = \sum_{n=0} \left(r_n (n - \kappa) (n - \kappa - 1) x^{n-2-\kappa} + \tilde{P}_1 r_n (n - \kappa) x^{n-2-\kappa} + \tilde{Q}_1 r_n x^{n-\kappa-2} \right) +$$

$$\sum_{n=0} \left(\tilde{P}_2 r_n (n - \kappa) x^{n-\kappa-1} + \tilde{Q}_2 r_n x^{n-\kappa-1} \right)$$

$$0 = \sum_{n=0} \left((n - \kappa) (n - \kappa - 1) + \tilde{P}_1 (n - \kappa) + \tilde{Q}_1 \right) r_n x^{n-2-\kappa} +$$

$$\sum_{n=1} \left(\tilde{P}_2 (n - 1 - \kappa) + \tilde{Q}_2 \right) r_{n-1} x^{n-\kappa-2}$$

$$0 = \left((-\kappa) (-\kappa - 1) + \tilde{P}_1 (-\kappa) + \tilde{Q}_1 \right) r_0 x^{-2-\kappa} + \sum_{n=1} \left[\left((n - \kappa) (n - \kappa - 1) + \tilde{P}_1 (n - \kappa) + \tilde{Q}_1 \right) r_n \right.$$

$$\left. + \left(\tilde{P}_2 (n - 1 - \kappa) + \tilde{Q}_2 \right) r_{n-1} \right] x^{n-\kappa-2} .$$

The indicial equation can be read off as

$$\left((-\kappa)(-\kappa - 1) + \tilde{P}_1(r_+; \lambda)(-\kappa) + \tilde{Q}_1(r_+; \lambda) \right) = 0. \quad (\text{A.30})$$

Solving for κ ,

$$\kappa = \frac{\tilde{P}_1(r_+) - 1 \pm \sqrt{(1 - \tilde{P}_1(r_+))^2 - 4\tilde{Q}_1(r_+)}}{2}. \quad (\text{A.31})$$

Evaluating the tilde functions, one finds

$$\begin{aligned} \tilde{P}_1(r_+) &= 1 \\ \tilde{Q}_1(r_+) &= \left(\frac{\omega r_+ r_s - am}{r_+ - r_-} \right)^2. \end{aligned}$$

Hence, we find κ to be

$$\kappa = \pm i \left(\frac{\omega r_+ r_s - am}{r_+ - r_-} \right). \quad (\text{A.32})$$

To find the recursion relation between the coefficients, we reindex the sum to

$$\sum_{n=0} \left(\left((n-k)(n-k-1) + \tilde{P}_1(r)(n-k) + \tilde{Q}_1(r) \right) r_n + \left(\tilde{P}_2(r)(n-1-k) + \tilde{Q}_2(r) \right) r_{n-1} \right) x^{n-k-2}. \quad (\text{A.33})$$

Now expand the tilde functions in a Taylor series around the outer horizon radius. After an application of the Cauchy product formula, one finds

$$\begin{aligned} \sum_{n=0} \left[(n-\kappa)(n-\kappa-1)r_n + \sum_{j=0}^n \left(p_{1,j}(n-j-\kappa)r_{n-j} + q_{1,j}r_{n-j} + p_{2,j}(n-j-1-\kappa)r_{n-j-1} + \right. \right. \\ \left. \left. q_{2,j}r_{n-1-j} \right) \right] x^{n-\kappa-2}, \end{aligned} \quad (\text{A.34})$$

where $p_{1/2,j}$ and $q_{1/2,j}$ denote the j 'th coefficient in the Taylor series expansion of the four tilde

functions. Vanishing of the coefficients enforces

$$r_n = -\frac{1}{(n-\kappa)(n-\kappa-1)} \sum_{j=0}^n \left(p_{1,j}(n-j-\kappa)r_{n-j} + q_{1,j}r_{n-j} + p_{2,j}(n-j-1-\kappa)r_{n-j-1} + q_{2,j}r_{n-1-j} \right). \quad (\text{A.35})$$

Solving for r_n , we find the recursion relation for the Frobenius coefficients

$$r_n = \frac{-1}{n(n-2*\kappa)} \left[(p_{2,0}(n-\kappa-1) + q_{2,0})r_{n-1} + \sum_{j=1}^n ((p_{1,j}(n-j-\kappa) + q_{1,j})r_{n-j} + (p_{2,j}(n-j-1-\kappa) + q_{2,j})r_{n-1-j}) \right]. \quad (\text{A.36})$$

Taking $r_0 = 1$ and calculating the Taylor series expansion of the four tilde functions, together with the value for κ derived from the indicial equation, we have completed our asymptotic expansion of the radial function. By Fuchs theorem, we have a fundamental set of solutions given by eq. A.26, together with eq. A.32 and eq. A.36. This solution will be used to calculate the boundary condition at the outer horizon as required for numerically solving eq. A.27.

Now the boundary conditions at infinity can be enforced by minimizing the quantity $\ln R(r_{max})^2$ over the complex-frequency space for a specified large value of r_{max} . This not only enforces the asymptotic boundary conditions at infinity, namely $R(r \rightarrow \infty) \sim e^{-Qr}$, but also will yield the complex-frequency parameter ω .

With these two boundary conditions enforced, we can numerically solve the radial equation by integrating from $r = r_+ + \epsilon$ to r_{max} , with initial conditions determined by the Frobenius expansion at the outer horizon and the asymptotic boundary condition enforced by the minimization of $\ln R(r_{max})^2$, which yields the complex-frequency ω .

After solving the radial and angular equation for the complex eigenvalues, we've determined the Proca field in terms of the parameters (m, a, M, n, S, μ) , where m is the total angular momentum projection, a is the dimensionless spin of the BH, M is the mass of the BH, n is the overtone number specifying the number of zero crossings of the radial function and comes from imposing boundary conditions on a Schrödinger-like equation in the non-relativistic regime, S is the spin of the Proca field, taking values $S = -1, 0, +1$, and $\mu = \frac{m_A}{h}$ is the mass parameter of the Proca field.

A.2 SECULAR EVOLUTION OF PROCA CLOUD

Here, we display various figures for certain orbital and system parameters. These aid in visualizing the relation between a dressed EMRI system and the vacuum case. The variation of the mass of the cloud over time can be determined from $\frac{dE_c}{dt} \propto E_c^2$, which follows from the Teukolsky formalism for the Proca field on a Kerr background. In particular, our choice of normalization is the statement of energy conservation, i.e. $M_c = M_{0,bh} - M_{f,bh}$, where M_c is the mass of the cloud at saturation, $M_{0,bh}$ and $M_{f,bh}$ are the initial and final masses of the black hole at saturation, respectively. Using eq. 2.23, and the fact $\mathfrak{T}_{\mu\nu} \sim (A^\mu)^2$, it then follows that the normalization coefficient of the Proca field is

$$C = \sqrt{\frac{M_{0,bh} - M_{f,bh}}{\bar{E}_c}}, \quad (\text{A.37})$$

where \bar{E}_c is the unnormalized energy calculated directly from eq. 2.23. The final mass of the black hole can be determined directly from eq. 2.30 and the requirement the saturation condition it met, $\omega = m\Omega_H$.

The true normalized energy of the cloud is then $E_c = C^2 \bar{E}_c$. This implies the amplitude of the Proca field scales likes $\sim \sqrt{M_c}$. From the Teukolsky formalism, the asymptotic energy flux due to perturbations of the Kerr spacetime obey the scaling relations $\dot{E}_c \propto |Z|^2 \propto C^4 = E_c^2$. In other words, $\frac{dE_c}{dt} \propto E_c^2$. This relation yields

$$M_c(t) = \frac{M_0}{1 + t/\tau}, \quad (\text{A.38})$$

where τ is the gravitational emission timescale and is determined from the relation

$$\tau = \frac{M_0}{\frac{dM}{dt}(0)}. \quad (\text{A.39})$$

The initial time is taken to be the time when the superradiant instability saturates and the cloud reaches a quasibound state. The gravitational emission during the instability is neglected due to the clear separation of timescales. The timescale for depletion via gravitational emission and evolution of the Proca cloud mass is shown in figure A.1. The evolution of the orbital parameters for an example system of $M = 10^6 M_\odot$ and $\chi = 0.9$ due to the secular change in the mass and angu-

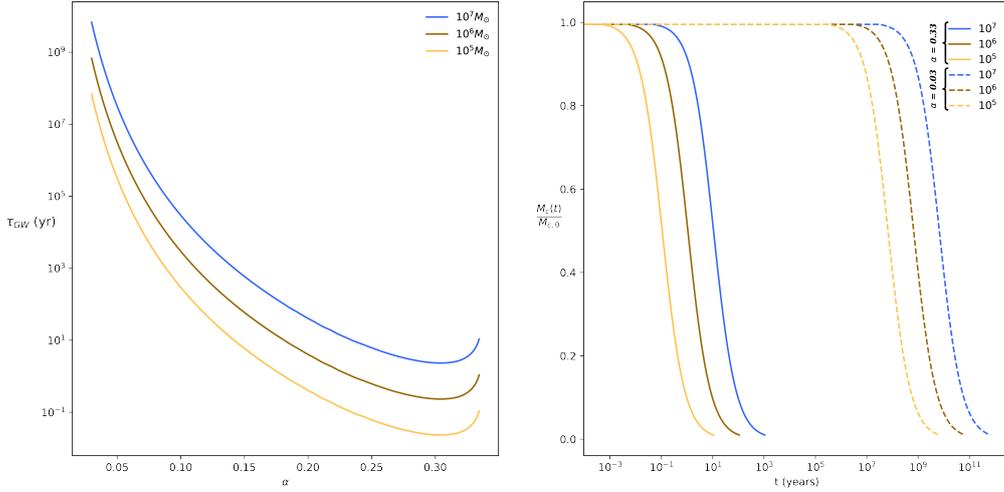


Figure A.1: Gravitational emission timescale (left) and relative depletion of the mass of the radiating Proca cloud (right) versus gravitational coupling for various choices of the SMBH mass. The emission timescale tends to shrink for higher couplings, up until the superradiant condition is violated. It's also apparent that lower gravitational couplings permit longer lived clouds.

lar momentum of the background is shown in figure A.2. The strength of the deviation from the vacuum scenario (labeled $\alpha = 0$ in the figure) as a function of the gravitational coupling is clearly visible. Stronger values of α correspond to greater deviations from the vacuum scenario, producing waveforms that differ by a larger amount from the undressed counterpart.

It should be noted that lower values of α use analytic expressions for the energy flux, while higher values of α use fits to numerical relativity calculations²⁸⁷. Moreover, for higher values of α , τ diverges near the superradiant threshold as the Proca cloud mass drops off in this region and $\tau \sim \frac{1}{M_0}$.

A.3 COMPARISON TO OTHER EFFECTS

As stated in the main text, several effects have been neglected when computing the potential Proca mass range observable with the LISA mission. These include dynamical friction^{279,313,355-361}, accretion of the Proca cloud onto the secondary black hole^{362,363}, and resonant transitions between Proca states^{300,302,358,364}. Within a limited scope, each effect is studied here to understand the role it plays in the potential Proca mass range observable with LISA.

First, it has been shown that dynamical friction has the same order of magnitude effect as that arising from transitions between states of the cloud. Indeed, the transition from bound to unbound states has been suggested to be interpreted as dynamical friction³⁶¹ of the Proca cloud on the

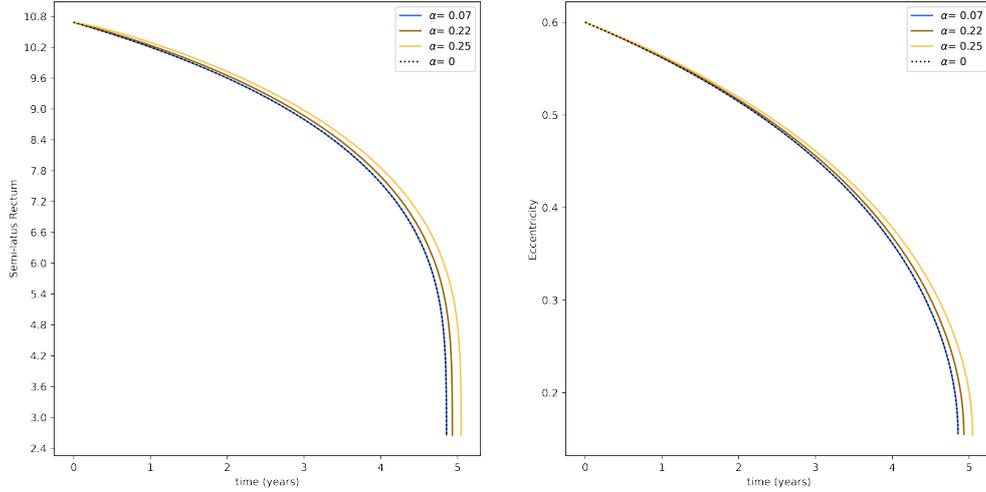


Figure A.2: Variation of the semi-latus rectum (left) and eccentricity (right) over the course of the inspiral, for various choices of the gravitational coupling α . Here, the SMBH mass is taken to be $M=10^6 M_\odot$. The semi-latus rectum is expressed in gravitational units, $\frac{p}{M}$, where p is the semi-latus rectum in geometrized units.

secondary black hole. However, the full phenomenon of state transitions has one of three effects on the orbital trajectory. The orbit either floats, sinks, or is kicked, depending on the initial and final states of the transition. Either of these three effects individually yield a reduction in the faithfulness, with respect to the vacuum case. When including the modification to the background, a sinking orbit may counteract the effect of the modification to the background discussed in the main text. However, a floating orbit has the opposite effect, it enhances the deviation from the vacuum scenario. This can be seen from the expression for the waveform inner product eq. 2.42. Consider a signal in the detector, whose functional form is $h(f) = A(f)e^{i\phi(f)}$, where $A(f)$ and $\phi(f)$ are the amplitude and phase as a function of frequency, respectively. Let $A_0(f)$ and $\phi_0(f)$ be the amplitude and phase, respectively, for the vacuum inspiral. Let $\delta\phi_1(f)$ denote the deviation from the vacuum scenario due to the modification of the background, as discussed in the main text, and let $\delta\phi_2(f)$ denote the additional phase deviation due to the floating or sinking orbits. At lowest order, the amplitude remains unchanged, so the waveform of the perturbed spacetime is

$$h_{\text{Proca}}(f) = A_0(f)e^{i(\phi_0(f)+\delta\phi_1(f)+\delta\phi_2(f))} = h_{\text{vacuum}}(f)e^{i(\delta\phi_1(f)+\delta\phi_2(f))} , \quad (\text{A.40})$$

where we've defined $h_{\text{vacuum}}(f) = A_0(f)e^{i\phi_0(f)}$. The waveform inner product then takes the form

$$\langle h_{\text{vacuum}} | h_{\text{Proca}} \rangle = 4 * Re \int \frac{h_{\text{vacuum}} h_{\text{Proca}}^*}{S_n} df = 4 * Re \int \frac{|A_0|^2}{S_n} e^{i(\delta\phi_1(f) + \delta\phi_2(f))} df . \quad (\text{A.41})$$

Assuming a small deviation of the phase*, then the inner products, at lowest order, becomes

$$\langle h_{\text{vacuum}} | h_{\text{Proca}} \rangle = 4 * Re \int \frac{|A_0|^2}{S_n} df - 2 * Re \int \frac{|A_0|^2}{S_n} (\delta\phi_1(f) + \delta\phi_2(f))^2 df , \quad (\text{A.42})$$

where we've dropped purely imaginary terms. For floating orbits, $\delta\phi_1$ and $\delta\phi_2$ have the same sign, so the second term is purely positive, reducing the value of the faithfulness. This implies that the inclusion of state transitions in the cloud will increase the observable mass range of the Proca cloud with LISA. To further elucidate this point, consider for example a $10^7 M_\odot$ primary black hole surrounded by a Proca cloud with mass $\mu = 9.35 * 10^{-19} eV$. The results of this study suggest this Proca mass would be unobservable with LISA. However, inclusion of floating reduces the faithfulness even further, potentially pushing it below the critical faithfulness threshold for observability. Over the entire parameter, the observable mass range of the Proca field with LISA is then further extended beyond the range suggested in the main text. This is owed to the fact that floating orbits contribute scenarios where the Proca mass is observable, while sinking orbits, that partially cancel the effect of this study, do not. These estimates were confirmed with explicit numerical computation for an example scenario. It was found that for $\mu = 9.35 * 10^{-19} eV$, $M = 10^7 M_\odot$, $e0 = 0.2$, and $\chi = 0.9$, inclusion of either a sinking or floating orbit reduced the faithfulness from values above the critical threshold for observability (see figure 2.6) to values below it. Despite the fact there are regions of the parameter space where the effects can partially cancel, over the entire range of the parameter space, the observable Proca mass range will increase due to those regions which enhance the background modification and thus push the faithfulness statistic below the critical threshold for observability.

The second additional effect neglected is accretion of the Proca cloud onto the secondary black hole. It has been suggested in³⁶⁵ that accretion is strongly α -suppressed relative to dynamical friction (and hence state transitions discussed above) and the change in the cloud mass is even further

*For state transitions, this is typically a good approximation since the deviation in frequency scales as q for $q \ll 1$ ³⁰⁰.

suppressed*. Hence, it's reasonable to neglect the effect of accretion at this order in the perturbative expansion. Nonetheless, accretion acts as an additional force that can either enhance or impair the radial inspiral^{362,363}. In a similar argument as above, over the entire parameter space, the region that enhances the modification to the background will widen the observable Proca mass range with LISA by pushing the faithfulness below the critical threshold. Thus, including accretion will also further expand the potentially observable Proca mass range of LISA.

The inclusion of these two effects are thus shown to further widen the mass range observable with LISA, further suggesting the results of this study are a conservative estimate. It should also be noted that self-gravity is another effect commonly studied in the literature^{366,367}. Since this study uses the SUPERRAD²⁸⁷ package to compute the asymptotic energy and angular momentum fluxes from the Proca cloud, which uses numerical relativity-fitted formulas for large α -values and a Newtonian treatment for $\alpha \ll 1$, the self-gravity of the cloud is automatically included in this study.

*The estimations in 365 were calculated assuming a scalar cloud, but then suggested to be within the same order of magnitude for the Proca cloud as well.

B

Additional Considerations for Proca in an Expanding Universe

This appendix contains elucidations on the theoretical study in Chapter 4. Section B.1 elucidates the co-ordinate transformations performed in the numerical analysis. In particular, the transformation of the metric variables and their derivatives from the transformation $(r, \theta, \phi) \rightarrow (x, y, z)$ is detailed. Section B.2 goes into detail about the evolution equations for the auxiliary damping field introduced in eqs. 4.23-4.26.

B.1 COORDINATE TRANSFORMATIONS

In transforming metric variables from Kerr-Schild to the Cartesian-like co-ordinate system eqs. 4.20-4.22, we follow the standard procedure for co-ordinate transformations. That is

$$\beta_{cart}^i(x, y, z) = \Lambda_j^i \beta_{KS}^j(r, \theta, \phi) \quad (\text{B.1})$$

$$\gamma_{ij, cart} = (\Lambda^{-1})_i^k (\Lambda^{-1})_j^l \gamma_{kl, KS} , \quad (\text{B.2})$$

where $\Lambda_j^i = \frac{dX^i}{dR^j}$, $X = (x, y, z)$, and $R = (r, \theta, \phi)$. The derivatives of the metric variables are then computed straight forwardly using the following rule for a rank-r contravariant tensor

$$\partial_p T^{ij\dots} = (\Lambda^{-1})_p^r \left[\tilde{\partial}_r (\Lambda_m^i \Lambda_n^j \dots) \tilde{T}^{mn\dots} + (\Lambda_m^i \Lambda_n^j \dots) \tilde{\partial}_r \tilde{T}^{mn\dots} \right] , \quad (\text{B.3})$$

where a tilde represents the quantity in the base co-ordinate system and non-tilde in the new co-ordinate system. For a covariant rank-r tensor, we find a similar rule

$$\partial_r T_{ij\dots} = (\Lambda^{-1})_r^p \left((\Lambda^{-1})_i^m (\Lambda^{-1})_j^n \dots \right) \tilde{\partial}_p \tilde{T}_{mn\dots} + \tilde{T}_{mn\dots} \partial_r \left((\Lambda^{-1})_i^m (\Lambda^{-1})_j^n \dots \right) . \quad (\text{B.4})$$

For example, the spatial metric in the new Cartesian-like co-ordinate system can be computed as

$$\partial_r \gamma_{ij} = \frac{\partial \tilde{\gamma}_{kl}}{\partial R^m} \frac{\partial R^m}{\partial X^r} \frac{\partial R^k}{\partial X^i} \frac{\partial R^l}{\partial X^j} + \tilde{\gamma}_{kl} \frac{\partial}{\partial X^r} \left(\frac{\partial R^k}{\partial X^i} \frac{\partial R^l}{\partial X^j} \right) . \quad (\text{B.5})$$

The rest of the derivatives of the metric variables follow similarly. The Jacobian matrix of the transformation eqs. 4.20-4.22 is simple to compute. Since at each grid point, we are given the (x, y, z) co-ordinates, we represent the Jacobian matrix in terms of the (x, y, z) variables. Hence, it is given by

$$\Lambda_j^i = \begin{bmatrix} \frac{x}{r} & \frac{y}{r} & \frac{z}{r} \\ \frac{xz}{r^2 \rho} & \frac{yz}{r^2 \rho} & \frac{-\rho}{r^2} \\ \frac{-y}{\rho^2} & \frac{x}{\rho^2} & 0 \end{bmatrix} , \quad (\text{B.6})$$

where $r^2 = x^2 + y^2 + z^2$ and $\rho^2 = x^2 + y^2$. Due to the division by the two radii, GRBoondi sets a minimum value for the radii of 10^{-12} . For more details on the numerical implementation, see 368.

B.2 EQUATIONS OF MOTIONS OF THE AUXILIARY DAMPING FIELD

The details about the auxiliary damping field introduced in eqs. 4.23-4.26 are discussed here. The covariant equations of motion that lead to eqs. 3.128-3.131 are

$$\nabla_\alpha(F^{\alpha\beta} + g^{\alpha\beta}Z) = \mu^2 A^\beta + \kappa Z n^\beta . \quad (\text{B.7})$$

In the absence of the auxiliary field, eq. 4.5 is recovered and taking the divergence would yield the Lorenz constraint $\nabla_\mu A^\mu = 0$. However, in the case of the added auxiliary field, a new equation is derived

$$\square Z - \kappa n^\mu \nabla_\mu Z + \kappa Z \mathcal{K} = \mu^2 \nabla_\nu A^\nu , \quad (\text{B.8})$$

which is manifestly an evolution equation for the auxiliary field. In fact, due to the presence of the \square operator, this is the form of a modified wave equation. Since the constraint $\nabla_\nu A^\nu$ is the function that should be minimized by the Z field, and with a little foresight, we make the intelligent choice $Z = \mu^2 \nabla_\nu A^\nu$. Eq. B.8 is now a modified wave equation for the *Lorenz constraint* and takes the form

$$\square Z - \kappa n^\mu \nabla_\mu Z + Z(\kappa \mathcal{K} - 1) = 0 \quad (\text{B.9})$$

called the generalized telegraph equation³³² in curved space. This equation describes exponentially suppressed wave motion. This can be seen by restricting to flat space and setting $\kappa = 1$. In this case, $Z = f(t - x)e^{-t}$ is a solution, which describes waves propagating at the speed of light and experiencing exponential temporal suppression. This is the desired behavior here since Z represents violations of the Proca constraint equation. In the context of numerical schemes, violations of the Proca constraint propagate at the speed of light outside of the finite computational domain and are simultaneously suppressed. This explains the terminology used in the main text of the thesis which describes Z as an 'auxiliary damping field'. It's auxiliary since it was added by hand and it damps due to the nature of the generalized telegraph equation.



Generalized Proca as a Constrained System

Broadly speaking, there are two possible avenues for studying generalized Proca theories. The first was the one adopted throughout the main text of the thesis — the Lagrangian formalism. This is largely the angle adopted throughout most field theories. However, the other approach can offer new insights, which is the Hamiltonian formalism. This is the approach usually taken in the context of quantum *mechanics*.

One of the fathers of quantum mechanics, Paul Dirac, developed a theory of *constrained systems* using the Hamiltonian formalism. This theory offers insights into the gauge nature of many fields. In the modern age, Dirac's theory forms one of the key foundations of modern theoretical physics and is widely utilized by theories such as modern QFT and GR. It provides a rigorous treatment of all theories that contain constraints and systematically describes gauge invariance and even quantization. We can thus use Dirac's theory of constrained Hamiltonian systems to acquire further

insights into the nature of generalized Proca systems. Though this analysis is largely incomplete, we use one of the most basic generalized Proca theories to illustrate possible issues that can easily sneak up unsuspectingly when evolving seemingly innocuous evolution equations.

C.1 HAMILTONIAN ANALYSIS À LA DIRAC-BERGMANN

We consider the \mathcal{L}_2 Lagrangian, restricted to be only a function of the Proca norm

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - V(A^\mu A_\mu) . \quad (\text{C.1})$$

Converting to the Hamiltonian via the standard Legendre transformation yields

$$\mathcal{H}_{can} = \frac{\alpha}{2}\Pi_i\Pi^i + D_j\Pi^j (\alpha\phi - \beta^i X_i) + D_i X_j (\Pi^j \beta^i - \Pi^i \beta^j) + \quad (\text{C.2})$$

$$\frac{\alpha}{2}D_i X_j (D^i X^j - D^j X^i) + \alpha V(A^2) , \quad (\text{C.3})$$

where $\Pi^i \equiv \frac{\delta\mathcal{L}}{\delta X_i}$ and $\pi \equiv \frac{\delta\mathcal{L}}{\delta\phi}$. Its immediately clear from the canonical Hamiltonian that we have a single primary constraint $C_1 \equiv \pi = 0$. The primary Hamiltonian is thus

$$\mathcal{H}' = \mathcal{H}_{can} + \lambda_1 C_1 , \quad (\text{C.4})$$

where λ_1 is a Lagrange multiplier. It's important to note here that there are cases where the Dirac-Bergmann algorithm can breakdown³⁶⁹. In particular, the algorithm will fail when the primary constraints contain spatial derivatives of the configuration space variables. Fortunately, this is not the case here, so the algorithm can be followed as usual. We can continue forward by following the Dirac consistency conditions for the constraints.

As usual, we define the Poisson bracket in field theory by

$$[A, B] = \int d^3w \left[\frac{\delta A}{\delta\phi(w)} \frac{\delta B}{\delta\pi(w)} - \frac{\delta A}{\delta\pi(w)} \frac{\delta B}{\delta\phi(w)} + \quad (\text{C.5})$$

$$\frac{\delta A}{\delta X_i(w)} \frac{\delta B}{\delta\Pi^i(w)} - \frac{\delta A}{\delta\Pi^i(w)} \frac{\delta B}{\delta X_i(w)} \right] . \quad (\text{C.6})$$

The Dirac consistency conditions then yield an additional constraint

$$\dot{C}_1 = [C_1, H'] \approx 0 \rightarrow C_2 \equiv -D_i \Pi^i + 2\phi V' , \quad (\text{C.7})$$

where the symbol \approx denotes equality on the constraint surface. This constraint is reminiscent of the Gauss constraint from Maxwell electrodynamics. In fact, one can show $\Pi^i = -E^i$, thus recovering the previous constraint from the corresponding Lagrangian theory.

Continuing with the consistency relations, one would calculate $\dot{C}_2 = [C_2, \mathcal{H}']$. However, this yields relations for the Lagrange multiplier, signaling the end of the recursive algorithm. Hence, for this theory there is only one primary and one secondary constraint, $C_1 = \pi$ and $C_2 = -D_i \Pi^i + 2\phi V'$.

It immediately follows from these constraints that this system is purely second class, just as regular Proca, since

$$[C_1, C_2] = V' - 2\phi^2 V'' . \quad (\text{C.8})$$

Since this system is purely second class for generic field configurations, the first class, total, and extended Hamiltonians are all equivalent. These are obtained by plugging in the solution from $[C_2, H']$ in terms of the Lagrange multiplier into the primary Hamiltonian.

With the second-class constraints in hand, we can continue on and define the Dirac bracket. Since we're dealing with a field theory, we need to modify the usual finite-dimensional version. Here, we have an infinite number of second-class constraints — two at each hypersurface point. Thus, the matrix product in the finite-dimensional Dirac bracket gets promoted to an integral over the hypersurfaces

$$[A, B]^* \equiv [A, B] - \int d^3z d^3z' [A, C_a(z)] M_{ab}^{-1}(z, z') [C_b(z'), \mathcal{H}_{pr}] , \quad (\text{C.9})$$

where \mathcal{H}_{pr} is the partially reduced Hamiltonian, obtained from the extended Hamiltonian by ap-

plying the second-class constraints. For our system

$$\begin{aligned} \mathcal{H}_{pr} = & \frac{\alpha}{2} \Pi_i \Pi^i + D_i X_j (\Pi^j \beta^i - \Pi^i \beta^j) + \\ & \frac{\alpha}{2} D_i X_j (D^i X^j - D^j X^i) + \alpha V(A^2) + 2\phi V'(\alpha\phi - \beta^i X_i) . \end{aligned} \quad (\text{C.10})$$

This definition of the Dirac bracket can be straight-forwardly shown to satisfy the usual properties of the finite-dimensional Dirac bracket, namely vanishing on the second-class constraints, anti-symmetry, linearity, etc.

The evolution of the canonical variables is then determined via the Dirac bracket. We hence find the equations of motion

$$\dot{\pi} \approx 0 \quad (\text{C.11})$$

$$\dot{\phi} \approx \frac{1}{g_{nn}} [C_2, \mathcal{H}_{pr}] \quad (\text{C.12})$$

$$\mathcal{L}_m X_i \approx \alpha \Pi_i - D_i(\alpha\phi) \quad (\text{C.13})$$

$$\mathcal{L}_m \Pi^i \approx 2\alpha X^i V' - \Pi^i \alpha K - D_j(\alpha D^j X^i) + \alpha D_j D^i X^j . \quad (\text{C.14})$$

We leave the expression for $\dot{\phi}$ unevaluated here since the resulting expression is quite lengthy due to the presence of the term $D_i \frac{\delta \mathcal{H}_{pr}}{\delta X_i}$. This results in third derivatives of the potential.

These evolution equations automatically take into account the second-class constraints and so are free evolution equations. Though this assumes we have a purely second class system. A peculiarity of the second-class constraints is that they're spatially varying quantities, hence their Poisson bracket are as well. This implies that there can be certain field configurations where $[C_1, C_2] \approx 0$ in local patches of spacetime, while remaining non-zero everywhere else. This has immense implications for the well-posedness of the evolution equations, which we now investigate.

C.2 WELL-POSEDNESS OF THE CAUCHY PROBLEM

We wish to analyze the well-posedness of the Cauchy problem from the point of view of the Hamiltonian theory. The primary issue will occur in the region of phase space where the second-class

constraints become first class,

$$[C_1, C_2](x) = 2g_{nn}(x) \rightarrow 0 . \quad (\text{C.15})$$

Since $g_{nn}(x)$ is an explicitly local quantity, the region where $g_{nn}(x) = 0$ is a patch of the spacetime (which may or may not be an open region and it may be point-like). The important consequence of $g_{nn} \rightarrow 0$ is that the constraints become first class.

It follows directly from the Dirac-Bergmann algorithm and the Dirac conjecture that the first class constraints generate infinitesimal gauge transformations, while second-class constraints don't affect the evolution and only the initial data. Explicitly, consider the canonical variable $\phi(x)$. Once $\phi(x)$ enters the region of phase space where $g_{nn}(x) = 0$, it acquires an additional gauge freedom, via

$$\delta\phi(x) = \epsilon^a [\phi(x), C_a(x)] . \quad (\text{C.16})$$

This is precisely where the issue arises in the well-posedness of the Cauchy problem, which we now show.

Consider the canonical variable ϕ and set up initial data for ϕ far away from any region where $g_{nn} = 0$. Since the constraints are all second class on the initial data surface, the evolution of ϕ is uniquely determined by the total Hamiltonian. Let ϕ evolve off the initial data surface according to this evolution and suppose its trajectory takes it into a region where $g_{nn} = 0$. Let the time this occurs be called t_1 .

At t_1 , ϕ acquires gauge freedom due to the reduction of the second-class constraints to first class. So ϕ is subject to specifiable gauge conditions, in order to have a well-defined evolution. Hence, the solutions behavior is subject to additional freedom that is independent of the initial data. As one cannot impose these gauge conditions at the initial data surface, the solutions behavior changes discontinuously with the initial data. Hence, the system is ill-posed, with respect to Hadamard. If at any point in the evolution, $g_{nn} = 0$, then the system becomes ill-posed and well-defined evolution is no longer possible.

The ill-posedness of the equations of motion is also evident from eq. C.11. When $g_{nn} \rightarrow 0$, the solution of the evolution equations is no longer possible since $\dot{\phi} \rightarrow \infty$. This reflects the fact that we

lose a degree of freedom in this patch of spacetime. Due to the fact the constraints become locally first class, the number of dynamical degrees of freedom reduces from three to two.

Publications

The following is a list of publications by the author. From this list, the starred items were used directly in this thesis.

- * Shaun Fell, Lavinia Heisenberg, and Doğa Veske. Detecting fundamental vector fields with lisa. *Physical Review D*, 108(8):083010, October 2023. ISSN 2470-0029. doi: 10.1103/PhysRevD.108.083010. URL <http://dx.doi.org/10.1103/PhysRevD.108.083010>
- * Shaun David Brocus Fell and Lavinia Heisenberg. Grboondi: A code for evolving generalized proca theories on arbitrary backgrounds. 2024. doi: <https://doi.org/10.48550/arXiv.2405.01348>. URL <https://arxiv.org/abs/2405.01348>
- * Shaun David Brocus Fell and Lavinia Heisenberg. Proca in an expanding universe. *Fortschritte der Physik*, 72(7-8):2400110, 2024. doi: <https://doi.org/10.1002/prop.202400110>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/prop.202400110>
- Fabio D'Ambrosio, Shaun D. B. Fell, Lavinia Heisenberg, et al. Black holes in $f(\mathbb{Q})$ gravity. *Phys. Rev. D*, 105:024042, January 2022. doi: 10.1103/PhysRevD.105.024042. URL <https://link.aps.org/doi/10.1103/PhysRevD.105.024042>
- Fabio D'Ambrosio, Shaun D. B. Fell, Lavinia Heisenberg, et al. Gravitational waves in full, non-linear general relativity, 2022. URL <https://arxiv.org/abs/2201.11634>
- Shaun D B Fell and Lavinia Heisenberg. Positive energy warp drive from hidden geometric structures. *Classical and Quantum Gravity*, 38(15):155020, July 2021. ISSN 1361-6382. doi: 10.1088/1361-6382/ac0e47. URL <http://dx.doi.org/10.1088/1361-6382/ac0e47>

References

- [1] David C Lindberg. *The Beginnings of Western Science: The European scientific tradition in philosophical, religious, and institutional context, prehistory to A.d. 1450, second edition*. University of Chicago Press, Chicago, IL, February 2010.
- [2] David C. Lindberg and Robert S. Westman, editors. *Reappraisals of the Scientific Revolution, Ed. By and (Cambridge:)*. Cambridge University Press, 1990.
- [3] David C. Lindberg. *The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, Prehistory to A.D. 1450, Second Edition*. University of Chicago Press, 2008. ISBN 9780226482040. doi: 10.7208/chicago/9780226482040.001.0001. URL <http://dx.doi.org/10.7208/chicago/9780226482040.001.0001>.
- [4] Edward Grant. *A history of natural philosophy*. Cambridge University Press, Cambridge, England, January 2007.
- [5] Ofer Gal. *The origins of modern science*. Cambridge University Press, Cambridge, England, February 2021.
- [6] Peter J Bowler and Iwan Rhys Morus. *Making modern science, second edition*. University of Chicago Press, Chicago, IL, 2 edition, August 2020.
- [7] David Cahan. *From natural philosophy to the sciences*. University of Chicago Press, Chicago, IL, October 2003.
- [8] Bernard Lightman. *CHAPTER 13. Science and the Public*, pages 337–376. University of Chicago Press, Chicago, 2011. ISBN 9780226318035. doi: doi:10.7208/9780226318035-015. URL <https://doi.org/10.7208/9780226318035-015>.
- [9] Adolf A. Berle. Scientists in national and world affairs: The atomic age . edited by morton grodzins and eugene rabinowitch. basic books, new york, 1963. 634 pp. \$10. *Science*, 143 (3604):341–344, January 1964. ISSN 1095-9203. doi: 10.1126/science.143.3604.341. URL <http://dx.doi.org/10.1126/science.143.3604.341>.
- [10] Ted Spitzmiller. *The history of human space flight*. University Press of Florida, Gainesville, FL, January 2017.
- [11] Sean Carroll. In 1894, physics seemed complete. then kelvin spotted 2 looming "clouds" — bigthink.com. <https://bigthink.com/hard-science/19th-century-physics-kelvin/>, 2024. [Accessed 28-05-2024].
- [12] Sean Carroll. *Quanta and fields: The Biggest Ideas in the universe*. Penguin, New York, NY, May 2024.

-
- [13] Albert Einstein. *Relativity*. Wings Books, New York, NY, May 2001.
- [14] Don Howard and John Stachel, editors. *Einstein and the history of general relativity*. Einstein Studies. Birkhauser Boston, Secaucus, NJ, 1989 edition, January 1989.
- [15] Edmund Whittaker. *History of the theories of aether and electricity, vol. I*. Dover Publications, Mineola, NY, May 2017.
- [16] Max Planck. *The theory of heat radiation*. Dover Books on Physics. Dover Publications, Mineola, NY, March 2003.
- [17] A Douglas Stone. *Einstein and the quantum*. Princeton University Press, Princeton, NJ, October 2013.
- [18] Einstein. The nobel prize in physics 1921. <https://www.nobelprize.org/prizes/physics/1921/summary/>, 1921. [Accessed 29-05-2024].
- [19] W. Heisenberg. über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, March 1927. doi: 10.1007/BF01397280.
- [20] Rita G Lerner. *Encyclopedia of physics*. Titles Supplied by John Wiley & Sons, Milton, QLD, Australia, 2 edition, December 1990.
- [21] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.*, 28:1049–1070, December 1926. doi: 10.1103/PhysRev.28.1049. URL <https://link.aps.org/doi/10.1103/PhysRev.28.1049>.
- [22] Schrödinger and Dirac. The nobel prize in physics 1933. <https://www.nobelprize.org/prizes/physics/1933/summary/>, 1933. [Accessed 29-05-2024].
- [23] Paul Dirac. The quantum theory of the electron. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 117(778):610–624, February 1928. ISSN 2053-9150. doi: 10.1098/rspa.1928.0023. URL <http://dx.doi.org/10.1098/rspa.1928.0023>.
- [24] Paul Adrien Maurice Dirac and Niels Henrik David Bohr. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 114(767):243–265, 1927. doi: 10.1098/rspa.1927.0039. URL <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1927.0039>.
- [25] J. R. Oppenheimer. Note on the theory of the interaction of field and matter. *Phys. Rev.*, 35:461–477, March 1930. doi: 10.1103/PhysRev.35.461. URL <https://link.aps.org/doi/10.1103/PhysRev.35.461>.
- [26] F. Bloch and A. Nordsieck. Note on the radiation field of the electron. *Phys. Rev.*, 52:54–59, July 1937. doi: 10.1103/PhysRev.52.54. URL <https://link.aps.org/doi/10.1103/PhysRev.52.54>.
- [27] V. F. Weisskopf. On the self-energy and the electromagnetic field of the electron. *Phys. Rev.*, 56:72–85, July 1939. doi: 10.1103/PhysRev.56.72. URL <https://link.aps.org/doi/10.1103/PhysRev.56.72>.

-
- [28] H. A. Bethe. The electromagnetic shift of energy levels. *Phys. Rev.*, 72:339–341, August 1947. doi: 10.1103/PhysRev.72.339. URL <https://link.aps.org/doi/10.1103/PhysRev.72.339>.
- [29] S. Tomonaga. On a relativistically invariant formulation of the quantum theory of wave fields. *Progress of Theoretical Physics*, 1(2):27–42, August 1946. doi: 10.1143/PTP.1.27.
- [30] Julian Schwinger. On quantum-electrodynamics and the magnetic moment of the electron. *Physical Review*, 73(4):416–417, February 1948. doi: 10.1103/PhysRev.73.416.
- [31] Julian Schwinger. Quantum electrodynamics. i. a covariant formulation. *Phys. Rev.*, 74:1439–1461, November 1948. doi: 10.1103/PhysRev.74.1439. URL <https://link.aps.org/doi/10.1103/PhysRev.74.1439>.
- [32] R. P. Feynman. Space-time approach to quantum electrodynamics. *Phys. Rev.*, 76:769–789, September 1949. doi: 10.1103/PhysRev.76.769. URL <https://link.aps.org/doi/10.1103/PhysRev.76.769>.
- [33] R. P. Feynman. The theory of positrons. *Phys. Rev.*, 76:749–759, September 1949. doi: 10.1103/PhysRev.76.749. URL <https://link.aps.org/doi/10.1103/PhysRev.76.749>.
- [34] R. P. Feynman. Mathematical formulation of the quantum theory of electromagnetic interaction. *Physical Review*, 80(3):440–457, November 1950. doi: 10.1103/PhysRev.80.440.
- [35] F. J. Dyson. The radiation theories of tomonaga, schwinger, and feynman. *Physical Review*, 75(3):486–502, February 1949. doi: 10.1103/PhysRev.75.486.
- [36] F. J. Dyson. The s matrix in quantum electrodynamics. *Physical Review*, 75(11):1736–1755, June 1949. doi: 10.1103/PhysRev.75.1736.
- [37] Feynman Tomonaga, Schwinger. The nobel prize in physics 1965 — nobelprize.org. <https://www.nobelprize.org/prizes/physics/1965/summary/>, 1965. [Accessed 29-05-2024].
- [38] Michael E Peskin and Daniel V Schroeder. *An introduction to quantum field theory*. CRC Press, London, England, September 2019.
- [39] G. W. Bennett, B. Bousquet, H. N. Brown, et al. Measurement of the positive muon anomalous magnetic moment to 0.7 ppm. *Phys. Rev. Lett.*, 89:101804, August 2002. doi: 10.1103/PhysRevLett.89.101804. URL <https://link.aps.org/doi/10.1103/PhysRevLett.89.101804>.
- [40] G. W. Bennett, B. Bousquet, H. N. Brown, et al. Measurement of the negative muon anomalous magnetic moment to 0.7 ppm. *Physical Review Letters*, 92(16), April 2004. ISSN 1079-7114. doi: 10.1103/physrevlett.92.161802. URL <http://dx.doi.org/10.1103/PhysRevLett.92.161802>.
- [41] X. Fan, T. G. Myers, B. A. D. Sukra, et al. Measurement of the electron magnetic moment. *prl*, 130(7):071801, February 2023. doi: 10.1103/PhysRevLett.130.071801.
- [42] Taizo Muta. *Foundations of quantum chromodynamics: An introduction to perturbative methods in gauge theories (3rd edition)*. World Scientific Lecture Notes In Physics. World Scientific Publishing, Singapore, Singapore, 3 edition, May 2008.

-
- [43] Andrei Smilga. *Lectures On Quantum Chromodynamics*. World Scientific Publishing, Singapore, Singapore, November 2001.
- [44] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *prl*, 30(26):1343–1346, June 1973. doi: 10.1103/PhysRevLett.30.1343.
- [45] H. David Politzer. Reliable perturbative results for strong interactions? *prl*, 30(26):1346–1349, June 1973. doi: 10.1103/PhysRevLett.30.1346.
- [46] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. i. *Phys. Rev.*, 122:345–358, April 1961. doi: 10.1103/PhysRev.122.345. URL <https://link.aps.org/doi/10.1103/PhysRev.122.345>.
- [47] E. Fermi. Versuch einer Theorie der β -Strahlen. I. *Zeitschrift fur Physik*, 88(3-4):161–177, March 1934. doi: 10.1007/BF01351864.
- [48] Fred L. Wilson. Fermi’s theory of beta decay. *American Journal of Physics*, 36(12):1150–1160, December 1968. doi: 10.1119/1.1974382.
- [49] Steven Weinberg. Conceptual foundations of the unified theory of weak and electromagnetic interactions. *Rev. Mod. Phys.*, 52:515–523, July 1980. doi: 10.1103/RevModPhys.52.515. URL <https://link.aps.org/doi/10.1103/RevModPhys.52.515>.
- [50] Sheldon Glashow, Abdus Salam, and Steven Weinberg. The Nobel Prize in Physics 1979 — nobelprize.org. <https://www.nobelprize.org/prizes/physics/1979/summary/>, 1979. [Accessed 29-05-2024].
- [51] Chris Hays. [www-pnp.physics.ox.ac.uk. http://www-pnp.physics.ox.ac.uk/~hays/ewk2016.pdf](http://www-pnp.physics.ox.ac.uk/~hays/ewk2016.pdf), 2016. [Accessed 29-05-2024].
- [52] Hugo Spinelli. File:Bosons-Hadrons-Fermions-RGB-png2.png - Wikipedia — w.wiki. <https://w.wiki/ADz6>, 2013. [Accessed 29-05-2024].
- [53] A Das and T Ferbel. *Introduction to Nuclear and Particle Physics*. WORLD SCIENTIFIC, December 2003. ISBN 9789812564351. doi: 10.1142/5460. URL <http://dx.doi.org/10.1142/5460>.
- [54] Robert M Eisberg and Robert Resnick. *Quantum physics of atoms, molecules, solids, nuclei, and particles*. John Wiley & Sons, Nashville, TN, 2 edition, January 1985.
- [55] R. L. Workman et al. Review of particle physics. *PTEP*, 2022:083Co1, 2022. doi: 10.1093/ptep/ptac097.
- [56] David Griffiths. *Introduction to elementary particles*. Wiley-VCH Verlag, Weinheim, Germany, 2 edition, August 2008.
- [57] Cush. File:standard model of elementary particles.svg - wikipedia — en.wikipedia.org. https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg, 2019. [Accessed 30-05-2024].
- [58] Robert Oerter. *The theory of almost everything*. Pi Press, July 2005.
- [59] Bill Carithers and P. Grannis. Discovery of the top quark. *SLAC Beam Line*, 25N3:4–16, 1995.

-
- [60] E. D. Bloom, D. H. Coward, H. Destaebler, et al. High-Energy Inelastic e-p Scattering at 6° and 10°. *prl*, 23(16):930–934, October 1969. doi: 10.1103/PhysRevLett.23.930.
- [61] M. Breidenbach, J. I. Friedman, H. W. Kendall, et al. Observed behavior of highly inelastic electron-proton scattering. *prl*, 23(16):935–939, October 1969. doi: 10.1103/PhysRevLett.23.935.
- [62] Cern. Particle chameleon caught in the act of changing — home.cern. <https://home.cern/news/press-release/cern/particle-chameleon-caught-act-changing>, 2010. [Accessed 30-05-2024].
- [63] Julia Woithe, Gerfried J Wiener, and Frederik F Van der Veken. Let’s have a coffee with the standard model of particle physics! *Physics Education*, 52(3):034001, March 2017. ISSN 1361-6552. doi: 10.1088/1361-6552/aa5b25. URL <http://dx.doi.org/10.1088/1361-6552/aa5b25>.
- [64] A. Einstein. Zur elektrodynamik bewegter körper [adp 17, 891 (1905)]. *Annalen der Physik*, 517(S1):194–224, February 2005. ISSN 1521-3889. doi: 10.1002/andp.2005517S113. URL <http://dx.doi.org/10.1002/andp.2005517S113>.
- [65] David J Griffiths. *Introduction to Electrodynamics*. Pearson, Upper Saddle River, NJ, 4 edition, September 2012.
- [66] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, Nashville, TN, 3 edition, July 1998.
- [67] Karl von Meyenn. *Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen*, pages 160–214. Vieweg+Teubner Verlag, Wiesbaden, 1990. ISBN 978-3-322-83770-7. doi: 10.1007/978-3-322-83770-7_6. URL https://doi.org/10.1007/978-3-322-83770-7_6.
- [68] A. Einstein. Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes. *Annalen der Physik*, 340(10):898–908, January 1911. doi: 10.1002/andp.19113401005.
- [69] Hermann Weyl. Zur gravitationstheorie. *Annalen der Physik*, 359(18):117–145, January 1917. doi: 10.1002/andp.19173591804.
- [70] Abraham Pais. *Subtle is the lord : The science and the life of Albert Einstein: The science and the life of Albert Einstein*. Oxford University Press, London, England, September 1982.
- [71] F. W. Dyson, A. S. Eddington, and C. Davidson. A determination of the deflection of light by the sun’s gravitational field, from observations made at the total eclipse of may 29, 1919. *Philosophical Transactions of the Royal Society of London Series A*, 220:291–333, January 1920. doi: 10.1098/rsta.1920.0009.
- [72] W. W. Campbell and R. Trumpler. Observations on the deflection of light in passing through the sun’s gravitational field, made during the total solar eclipse of september 21, 1923. *Publications of the Astronomical Society of the Pacific*, 35(205):158, June 1923. doi: 10.1088/123292a. URL <https://dx.doi.org/10.1088/123292a>.
- [73] Don Howard and John Stachel, editors. *Einstein and the history of general relativity*. Einstein Studies. Birkhauser Boston, Secaucus, NJ, 1989 edition, January 1989.

-
- [74] E. Fomalont, S. Kopeikin, G. Lanyi, et al. Progress in measurements of the gravitational bending of radio waves using the vlba. *The Astrophysical Journal*, 699(2):1395, June 2009. doi: 10.1088/0004-637X/699/2/1395. URL <https://dx.doi.org/10.1088/0004-637X/699/2/1395>.
- [75] B. Bertotti, L. Iess, and P. Tortora. A test of general relativity using radio links with the cassini spacecraft. *nat*, 425(6956):374–376, September 2003. doi: 10.1038/nature01997.
- [76] S. Schlamminger, K. Y. Choi, T. A. Wagner, et al. Test of the equivalence principle using a rotating torsion balance. *prl*, 100(4):041101, February 2008. doi: 10.1103/PhysRevLett.100.041101. URL <https://doi.org/10.1103/PhysRevLett.100.041101>.
- [77] Davide Castelvecchi and Alexandra Witze. Einstein’s gravitational waves found at last. *Nature*, February 2016. ISSN 1476-4687. doi: 10.1038/nature.2016.19361. URL <http://dx.doi.org/10.1038/nature.2016.19361>.
- [78] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Observation of gravitational waves from a binary black hole merger. *prl*, 116(6):061102, February 2016. doi: 10.1103/PhysRevLett.116.061102. URL <https://link.aps.org/doi/10.1103/PhysRevLett.116.061102>.
- [79] NSF. Gravitational waves detected 100 years after Einstein’s prediction — nsf.gov. https://www.nsf.gov/news/news_summ.jsp?cntn_id=137628, 2016. [Accessed 30-05-2024].
- [80] Colin Montgomery, Wayne Orchiston, and Ian Whittingham. Michell, laplace and the origin of the black hole concept. *Journal of Astronomical History and Heritage*, 12(2):90–96, July 2009.
- [81] John Michell. On the means of discovering the distance, magnitude, &c. of the fixed stars, in consequence of the diminution of the velocity of their light, in case such a diminution should be found to take place in any of them, and such other data should be procured from observations, as would be farther necessary for that purpose. by the rev. john michell, b. d. f. r. s. in a letter to henry cavendish, esq. f. r. s. and a. s. *Philosophical Transactions of the Royal Society of London Series I*, 74:35–57, January 1784.
- [82] Adam Levy. How black holes morphed from theory to reality. *Knowable Magazine*, January 2021. ISSN 2575-4459. doi: 10.1146/knowable-010921-1. URL <http://dx.doi.org/10.1146/knowable-010921-1>.
- [83] Karl Schwarzschild. Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pages 189–196, January 1916.
- [84] Roger Penrose. Gravitational collapse and space-time singularities. *prl*, 14(3):57–59, January 1965. doi: 10.1103/PhysRevLett.14.57.
- [85] L. H. Ford. The classical singularity theorems and their quantum loopholes. *arXiv e-prints*, art. gr-qc/0301045, January 2003. doi: 10.48550/arXiv.gr-qc/0301045.
- [86] Roger Penrose, Reinhard Genzel, and Andrea Ghez. The Nobel Prize in Physics 2020 — nobelprize.org. <https://www.nobelprize.org/prizes/physics/2020/summary/>, 2020. [Accessed 31-05-2024].

-
- [87] Roy P. Kerr. Gravitational field of a spinning mass as an example of algebraically special metrics. *prl*, 11(5):237–238, September 1963. doi: 10.1103/PhysRevLett.11.237.
- [88] Herbert Pfister. On the history of the so-called lense-thirring effect. *General Relativity and Gravitation*, 39(11):1735–1748, September 2007. ISSN 1572-9532. doi: 10.1007/s10714-007-0521-4. URL <http://dx.doi.org/10.1007/s10714-007-0521-4>.
- [89] Ronald J Adler and etc. *Introduction to general relativity*. Pure & Applied Physics S. McGraw Hill Higher Education, Maidenhead, England, 2 edition, June 1975.
- [90] Charles W Misner and etc. *Gravitation*. W.H. Freeman, New York, NY, December 1973.
- [91] Matt Visser. The kerr spacetime: A brief introduction, 2008.
- [92] John Kogut. Supplemental lecture 6 rotating sources, gravito-magnetism and the kerr black hole, 01 2019.
- [93] Michael Boyle and Mark Scheel. The sxs package, May 2024.
- [94] Michael Boyle, Daniel Hemberger, Dante A B Iozzo, et al. The sxs collaboration catalog of binary black hole simulations. *Classical and Quantum Gravity*, 36(19):195006, September 2019. ISSN 1361-6382. doi: 10.1088/1361-6382/ab34e2. URL <http://dx.doi.org/10.1088/1361-6382/ab34e2>.
- [95] Abdul H. Mroué, Mark A. Scheel, Béla Szilágyi, et al. Catalog of 174 binary black hole simulations for gravitational wave astronomy. *Physical Review Letters*, 111(24), December 2013. ISSN 1079-7114. doi: 10.1103/physrevlett.111.241104. URL <http://dx.doi.org/10.1103/PhysRevLett.111.241104>.
- [96] Bruce Rolston. U. of T. The Bulletin, November 10, 1997, Articles — web.archive.org. https://web.archive.org/web/20080307181205/http://www.news.utoronto.ca/bin/bulletin/nov10_97/art4.htm, 1997. [Accessed 31-05-2024].
- [97] H. L. Shipman. The implausible history of triple star models for cygnus x-1: Evidence for a black hole. *aplett*, 16:9, February 1975.
- [98] The Associated Press. Video: Astronomers Reveal the First Picture of a Black Hole — nytimes.com. <https://www.nytimes.com/video/science/space/10000006453594/first-image-black-hole.html?smid=url-share>, 2019. [Accessed 31-05-2024].
- [99] Kazunori Akiyama, Antxon Alberdi, Walter Alef, et al. First m87 event horizon telescope results. i. the shadow of the supermassive black hole. *The Astrophysical Journal Letters*, 875(1):L1, April 2019. ISSN 2041-8213. doi: 10.3847/2041-8213/ab0ec7. URL <http://dx.doi.org/10.3847/2041-8213/ab0ec7>.
- [100] Kazunori Akiyama, Antxon Alberdi, Walter Alef, et al. First m87 event horizon telescope results. iv. imaging the central supermassive black hole. *The Astrophysical Journal Letters*, 875(1):L4, April 2019. ISSN 2041-8213. doi: 10.3847/2041-8213/ab0e85. URL <http://dx.doi.org/10.3847/2041-8213/ab0e85>.

-
- [101] Event Horizon Telescope Collaboration, Kazunori Akiyama, Antxon Alberdi, et al. First sagittarius a*event horizon telescope results. i. the shadow of the supermassive black hole in the center of the milky way. *The Astrophysical Journal Letters*, 930(2):L12, May 2022. ISSN 2041-8213. doi: 10.3847/2041-8213/ac6674. URL <http://dx.doi.org/10.3847/2041-8213/ac6674>.
- [102] Event Horizon Telescope Collaboration, Kazunori Akiyama, Antxon Alberdi, et al. First sagittarius a*event horizon telescope results. iii. imaging of the galactic center supermassive black hole. *The Astrophysical Journal Letters*, 930(2):L14, May 2022. ISSN 2041-8213. doi: 10.3847/2041-8213/ac6429. URL <http://dx.doi.org/10.3847/2041-8213/ac6429>.
- [103] Event Horizon Telescope Collaboration, Kazunori Akiyama, Antxon Alberdi, et al. First sagittarius a*event horizon telescope results. vi. testing the black hole metric. *The Astrophysical Journal Letters*, 930(2):L17, May 2022. ISSN 2041-8213. doi: 10.3847/2041-8213/ac6756. URL <http://dx.doi.org/10.3847/2041-8213/ac6756>.
- [104] Albert Einstein. Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pages 688–696, January 1916.
- [105] Albert Einstein. Über Gravitationswellen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pages 154–167, January 1918.
- [106] Alexander H. Nitz, Collin Capano, Alex B. Nielsen, et al. 1-ogc: The first open gravitational-wave catalog of binary mergers from analysis of public advanced ligo data. *The Astrophysical Journal*, 872(2):195, February 2019. ISSN 1538-4357. doi: 10.3847/1538-4357/ab0108. URL <http://dx.doi.org/10.3847/1538-4357/ab0108>.
- [107] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs. *Physical Review X*, 9(3):031040, July 2019. doi: 10.1103/PhysRevX.9.031040.
- [108] Alexander H. Nitz, Thomas Dent, Gareth S. Davies, et al. 2-ogc: Open gravitational-wave catalog of binary mergers from analysis of public advanced ligo and virgo data. *apj*, 891(2):123, March 2020. doi: 10.3847/1538-4357/ab733f.
- [109] Fabio D’Ambrosio, Francesco Gozzini, Lavinia Heisenberg, et al. Testing gravitational waveforms in full general relativity, 2024.
- [110] Vitor Cardoso and Paolo Pani. Testing the nature of dark compact objects: a status report. *Living Reviews in Relativity*, 22(1), July 2019. ISSN 1433-8351. doi: 10.1007/s41114-019-0020-4. URL <http://dx.doi.org/10.1007/s41114-019-0020-4>.
- [111] N. V. Krishnendu and Sumanta Chakraborty. Probing black hole ‘charge’ from the binary black hole inspiral, 2024.
- [112] Tousif Islam, Ajit Kumar Mehta, Abhirup Ghosh, et al. Testing the no-hair nature of binary black holes using the consistency of multipolar gravitational radiation. *Phys. Rev. D*, 101:024032, January 2020. doi: 10.1103/PhysRevD.101.024032. URL <https://link.aps.org/doi/10.1103/PhysRevD.101.024032>.

-
- [113] Valentino F Foit and Matthew Kleban. Testing quantum black holes with gravitational waves. *Classical and Quantum Gravity*, 36(3):035006, January 2019. ISSN 1361-6382. doi: 10.1088/1361-6382/aafcba. URL <http://dx.doi.org/10.1088/1361-6382/aafcba>.
- [114] Andrea Maselli, Paolo Pani, Vitor Cardoso, et al. Probing planckian corrections at the horizon scale with lisa binaries. *Phys. Rev. Lett.*, 120:081101, February 2018. doi: 10.1103/PhysRevLett.120.081101. URL <https://link.aps.org/doi/10.1103/PhysRevLett.120.081101>.
- [115] Vitor Cardoso, Valentino F. Foit, and Matthew Kleban. Gravitational wave echoes from black hole area quantization. *Journal of Cosmology and Astroparticle Physics*, 2019(08):006–006, August 2019. ISSN 1475-7516. doi: 10.1088/1475-7516/2019/08/006. URL <http://dx.doi.org/10.1088/1475-7516/2019/08/006>.
- [116] Ivan Agullo, Vitor Cardoso, Adrián del Río, et al. Potential gravitational wave signatures of quantum gravity. *Phys. Rev. Lett.*, 126:041302, January 2021. doi: 10.1103/PhysRevLett.126.041302. URL <https://link.aps.org/doi/10.1103/PhysRevLett.126.041302>.
- [117] Andrew Coates, Sebastian H. Völkel, and Kostas D. Kokkotas. Spectral lines of quantized, spinning black holes and their astrophysical relevance. *Phys. Rev. Lett.*, 123:171104, October 2019. doi: 10.1103/PhysRevLett.123.171104. URL <https://link.aps.org/doi/10.1103/PhysRevLett.123.171104>.
- [118] S. W. Hawking. Chronology protection conjecture. *prd*, 46(2):603–611, July 1992. doi: 10.1103/PhysRevD.46.603.
- [119] AllenMcC. File:Lorentzian Wormhole.svg - Wikipedia — en.wikipedia.org. https://en.wikipedia.org/wiki/File:Lorentzian_Wormhole.svg, 2013. [Accessed 31-05-2024].
- [120] Enrico Rodrigo. *The physics of stargates*. Eridanus Press, September 2010.
- [121] Miguel Alcubierre. The warp drive: hyper-fast travel within general relativity. *Classical and Quantum Gravity*, 11(5):L73–L77, May 1994. ISSN 1361-6382. doi: 10.1088/0264-9381/11/5/001. URL <http://dx.doi.org/10.1088/0264-9381/11/5/001>.
- [122] Alexey Bobrick and Gianni Martire. Introducing physical warp drives. *Classical and Quantum Gravity*, 38(10):105009, April 2021. ISSN 1361-6382. doi: 10.1088/1361-6382/abdf6e. URL <http://dx.doi.org/10.1088/1361-6382/abdf6e>.
- [123] José Natário. Warp drive with zero expansion. *Classical and Quantum Gravity*, 19(6):1157–1165, March 2002. ISSN 1361-6382. doi: 10.1088/0264-9381/19/6/308. URL <http://dx.doi.org/10.1088/0264-9381/19/6/308>.
- [124] Chris Van Den Broeck. A ‘warp drive’ with more reasonable total energy requirements. *Classical and Quantum Gravity*, 16(12):3973–3979, November 1999. ISSN 1361-6382. doi: 10.1088/0264-9381/16/12/314. URL <http://dx.doi.org/10.1088/0264-9381/16/12/314>.
- [125] Erik W. Lentz. Breaking the warp barrier: hyper-fast solitons in einstein-maxwell-plasma theory. *Classical and Quantum Gravity*, 38(7):075015, April 2021. doi: 10.1088/1361-6382/abe692.

-
- [126] Shaun D B Fell and Lavinia Heisenberg. Positive energy warp drive from hidden geometric structures. *Classical and Quantum Gravity*, 38(15):155020, July 2021. ISSN 1361-6382. doi: 10.1088/1361-6382/ac0e47. URL <http://dx.doi.org/10.1088/1361-6382/ac0e47>.
- [127] Christopher Helmerich, Jared Fuchs, Alexey Bobrick, et al. Analyzing warp drive spacetimes with warp factory. *Classical and Quantum Gravity*, 41(9):095009, April 2024. ISSN 1361-6382. doi: 10.1088/1361-6382/ad2e42. URL <http://dx.doi.org/10.1088/1361-6382/ad2e42>.
- [128] Jared Fuchs, Christopher Helmerich, Alexey Bobrick, et al. Constant velocity physical warp drive solution. *Classical and Quantum Gravity*, 41(9):095013, April 2024. ISSN 1361-6382. doi: 10.1088/1361-6382/ad26aa. URL <http://dx.doi.org/10.1088/1361-6382/ad26aa>.
- [129] E. P. Hubble. Extragalactic nebulae. *apj*, 64:321–369, December 1926. doi: 10.1086/143018.
- [130] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, March 1929. ISSN 1091-6490. doi: 10.1073/pnas.15.3.168. URL <http://dx.doi.org/10.1073/pnas.15.3.168>.
- [131] NASA/WMAP. Timeline of the Universe Image — map.gsfc.nasa.gov. <https://map.gsfc.nasa.gov/media/060915/index.html>, 2006. [Accessed 31-05-2024].
- [132] Roger Penrose. *The road to reality*. Jonathan Cape, London, England, July 2004.
- [133] Fred C. Adams and Gregory Laughlin. A dying universe: the long-term fate and evolution of astrophysical objects. *Reviews of Modern Physics*, 69(2):337–372, April 1997. doi: 10.1103/RevModPhys.69.337.
- [134] George Ellis, Roy Maartens, and Malcolm A H MacCallum. *Relativistic Cosmology*. Cambridge University Press, Cambridge, England, March 2012.
- [135] A. Vikhlinin, A. V. Kravtsov, R. A. Burenin, et al. Chandra cluster cosmology project iii: Cosmological parameter constraints. *The Astrophysical Journal*, 692(2):1060–1074, February 2009. ISSN 1538-4357. doi: 10.1088/0004-637x/692/2/1060. URL <http://dx.doi.org/10.1088/0004-637x/692/2/1060>.
- [136] N. Aghanim, Y. Akrami, M. Ashdown, et al. Planck2018 results: Vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, September 2020. ISSN 1432-0746. doi: 10.1051/0004-6361/201833910. URL <http://dx.doi.org/10.1051/0004-6361/201833910>.
- [137] Albert Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pages 142–152, January 1917.
- [138] S.E. Rugh and H. Zinkernagel. The quantum vacuum and the cosmological constant problem. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 33(4):663–705, 2002. ISSN 1355-2198. doi: [https://doi.org/10.1016/S1355-2198\(02\)00033-3](https://doi.org/10.1016/S1355-2198(02)00033-3). URL <https://www.sciencedirect.com/science/article/pii/S1355219802000333>.

-
- [139] Ronald J. Adler, Brendan Casey, and Ovid C. Jacob. Vacuum catastrophe: An elementary exposition of the cosmological constant problem. *American Journal of Physics*, 63(7):620–626, July 1995. doi: 10.1119/1.17850.
- [140] Gabriel R. Bengochea, Gabriel León, Elias Okon, et al. Can the quantum vacuum fluctuations really solve the cosmological constant problem? *The European Physical Journal C*, 80(1), January 2020. ISSN 1434-6052. doi: 10.1140/epjc/s10052-019-7554-1. URL <http://dx.doi.org/10.1140/epjc/s10052-019-7554-1>.
- [141] Jérôme Martin. Everything you always wanted to know about the cosmological constant problem (but were afraid to ask). *Comptes Rendus. Physique*, 13(6–7):566–665, July 2012. ISSN 1878-1535. doi: 10.1016/j.crhy.2012.04.008. URL <http://dx.doi.org/10.1016/j.crhy.2012.04.008>.
- [142] Norbert Straumann. The history of the cosmological constant problem, 2002.
- [143] Steven Weinberg. The cosmological constant problem. *Rev. Mod. Phys.*, 61:1–23, January 1989. doi: 10.1103/RevModPhys.61.1. URL <https://link.aps.org/doi/10.1103/RevModPhys.61.1>.
- [144] Bharat Ratra and P. J. E. Peebles. Cosmological consequences of a rolling homogeneous scalar field. *Phys. Rev. D*, 37:3406–3427, June 1988. doi: 10.1103/PhysRevD.37.3406. URL <https://link.aps.org/doi/10.1103/PhysRevD.37.3406>.
- [145] Paul J. Steinhardt, Limin Wang, and Ivaylo Zlatev. Cosmological tracking solutions. *Phys. Rev. D*, 59:123504, May 1999. doi: 10.1103/PhysRevD.59.123504. URL <https://link.aps.org/doi/10.1103/PhysRevD.59.123504>.
- [146] Virginia Trimble. Existence and nature of dark matter in the universe. *Annual Review of Astronomy and Astrophysics*, 25(Volume 25, 1987):425–472, 1987. ISSN 1545-4282. doi: <https://doi.org/10.1146/annurev.aa.25.090187.002233>. URL <https://www.annualreviews.org/content/journals/10.1146/annurev.aa.25.090187.002233>.
- [147] Virginia Trimble. *History of Dark Matter in Galaxies*, page 1091–1118. Springer Netherlands, 2013. ISBN 9789400756120. doi: 10.1007/978-94-007-5612-0_21. URL <http://dx.doi.org/10.1007/978-94-007-5612-0%5F21>.
- [148] Robert H Sanders. *The dark matter problem*. Cambridge University Press, Cambridge, England, February 2014.
- [149] Jaan Einasto. *Dark matter and cosmic web story*. Advanced Series in Astrophysics and Cosmology. World Scientific Publishing Company, November 2013.
- [150] Gianfranco Bertone and Dan Hooper. History of dark matter. *Reviews of Modern Physics*, 90(4), October 2018. ISSN 1539-0756. doi: 10.1103/revmodphys.90.045002. URL <http://dx.doi.org/10.1103/RevModPhys.90.045002>.
- [151] J. G. de Swart, G. Bertone, and J. van Dongen. How dark matter came to matter. *Nature Astronomy*, 1(3), March 2017. ISSN 2397-3366. doi: 10.1038/s41550-017-0059. URL <http://dx.doi.org/10.1038/s41550-017-0059>.
- [152] J. C. Kapteyn. First attempt at a theory of the arrangement and motion of the sidereal system. *apj*, 55:302, May 1922. doi: 10.1086/142670.

-
- [153] J. H. Jeans. The motions of stars in a kapteyn-universe. *Monthly Notices of the Royal Astronomical Society*, 82(3):122–132, 01 1922. ISSN 0035-8711. doi: 10.1093/mnras/82.3.122. URL <https://doi.org/10.1093/mnras/82.3.122>.
- [154] J. H. Oort. Observational evidence confirming lindblad’s hypothesis of a rotation of the galactic system. *bain*, 3:275, April 1927.
- [155] F. Zwicky. Die rotverschiebung von extragalaktischen nebeln. *Helvetica Physica Acta*, 6: 110–127, January 1933.
- [156] Sinclair Smith. The mass of the virgo cluster. *apj*, 83:23, January 1936. doi: 10.1086/143697.
- [157] Erik Holmberg. A study of double and multiple galaxies together with inquiries into some general metagalactic problems. *Annals of the Observatory of Lund*, 6:1–173, January 1937.
- [158] KeighleyAstro. Dark matter and dark energy and the formation of the Universe. <http://keighleyastronomicalsociety.co.uk/dark-matter-and-dark-energy-and-the-formation-of-the-universe/>, 2019. [Accessed 03-06-2024].
- [159] F. Zwicky. On the masses of nebulae and of clusters of nebulae. *apj*, 86:217, October 1937. doi: 10.1086/143864.
- [160] Vera C. Rubin and Jr. Ford, W. Kent. Rotation of the andromeda nebula from a spectroscopic survey of emission regions. *apj*, 159:379, February 1970. doi: 10.1086/150317.
- [161] V. C. Rubin, Jr. Ford, W. K., and N. Thonnard. Rotational properties of 21 sc galaxies with a large range of luminosities and radii, from ngc 4605 ($r=4\text{kpc}$) to ugc 2885 ($r=122\text{kpc}$). *apj*, 238:471–487, June 1980. doi: 10.1086/158003.
- [162] Steven W. Allen, August E. Evrard, and Adam B. Mantz. Cosmological parameters from observations of galaxy clusters. *araa*, 49(1):409–470, September 2011. doi: 10.1146/annurev-astro-081710-102514.
- [163] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. Weak-lensing mass reconstruction of the interacting cluster 1e 0657-558: Direct evidence for the existence of dark matter. *apj*, 604(2):596–603, April 2004. doi: 10.1086/381970.
- [164] M. Markevitch, A. H. Gonzalez, D. Clowe, et al. Direct constraints on the dark matter self-interaction cross section from the merging galaxy cluster 1e 0657-56. *The Astrophysical Journal*, 606(2):819–824, May 2004. ISSN 1538-4357. doi: 10.1086/383178. URL <http://dx.doi.org/10.1086/383178>.
- [165] Wayne Hu. CMB Intermediate — background.uchicago.edu. <http://background.uchicago.edu/~whu/intermediate/intermediate.html>, 2001. [Accessed 03-06-2024].
- [166] Cohare9. File:Dark matter candidates.pdf - Wikipedia — en.wikipedia.org. https://en.wikipedia.org/wiki/File:Dark_matter_candidates.pdf, 2024. [Accessed 03-06-2024].
- [167] Lisa Randall. *Dark matter and the dinosaurs*. Ecco Press, Hopewell, VA, October 2015.

-
- [168] Allen Zeyher. MACHOs may be out of the running as a dark matter candidate — astronomy.com. <https://www.astronomy.com/science/machos-may-be-out-of-the-running-as-a-dark-matter-candidate/>, 2016. [Accessed 03-06-2024].
- [169] Earl P. Bellinger, Matt E. Caplan, Taeho Ryu, et al. Solar evolution models with a central black hole. *apj*, 959(2):113, December 2023. doi: 10.3847/1538-4357/ado4de.
- [170] G. 't Hooft. Computation of the quantum effects due to a four-dimensional pseudoparticle. *Phys. Rev. D*, 14:3432–3450, December 1976. doi: 10.1103/PhysRevD.14.3432. URL <https://link.aps.org/doi/10.1103/PhysRevD.14.3432>.
- [171] L.F. Abbott and P. Sikivie. A cosmological bound on the invisible axion. *Physics Letters B*, 120(1–3):133–136, January 1983. ISSN 0370-2693. doi: 10.1016/0370-2693(83)90638-x. URL [http://dx.doi.org/10.1016/0370-2693\(83\)90638-X](http://dx.doi.org/10.1016/0370-2693(83)90638-X).
- [172] John Preskill, Mark B. Wise, and Frank Wilczek. Cosmology of the invisible axion. *Physics Letters B*, 120(1–3):127–132, January 1983. ISSN 0370-2693. doi: 10.1016/0370-2693(83)90637-8. URL [http://dx.doi.org/10.1016/0370-2693\(83\)90637-8](http://dx.doi.org/10.1016/0370-2693(83)90637-8).
- [173] Michael Dine and Willy Fischler. The not-so-harmless axion. *Physics Letters B*, 120(1–3): 137–141, January 1983. ISSN 0370-2693. doi: 10.1016/0370-2693(83)90639-1. URL [http://dx.doi.org/10.1016/0370-2693\(83\)90639-1](http://dx.doi.org/10.1016/0370-2693(83)90639-1).
- [174] Luca Di Luzio, Maurizio Giannotti, Enrico Nardi, et al. The landscape of qcd axion models. *Physics Reports*, 870:1–117, July 2020. ISSN 0370-1573. doi: 10.1016/j.physrep.2020.06.002. URL <http://dx.doi.org/10.1016/j.physrep.2020.06.002>.
- [175] Fuminobu Takahashi, Wen Yin, and Alan H. Guth. Qcd axion window and low-scale inflation. *Phys. Rev. D*, 98:015042, July 2018. doi: 10.1103/PhysRevD.98.015042. URL <https://link.aps.org/doi/10.1103/PhysRevD.98.015042>.
- [176] C. Bartram, T. Braine, E. Burns, et al. Search for invisible axion dark matter in the 3.3 – 4.2 μeV mass range. *Phys. Rev. Lett.*, 127(26):261803, December 2021. ISSN 1079-7114. doi: 10.1103/PhysRevLett.127.261803. URL <https://link.aps.org/doi/10.1103/PhysRevLett.127.261803>.
- [177] R. Khatiwada, D. Bowring, A. S. Chou, et al. Axion dark matter experiment: Detailed design and operations. *Review of Scientific Instruments*, 92(12):124502, December 2021. ISSN 0034-6748. doi: 10.1063/5.0037857. URL <https://doi.org/10.1063/5.0037857>.
- [178] C. Bartram, T. Braine, R. Cervantes, et al. Axion dark matter experiment: Run 1b analysis details. *Phys. Rev. D*, 103:032002, February 2021. doi: 10.1103/PhysRevD.103.032002. URL <https://link.aps.org/doi/10.1103/PhysRevD.103.032002>.
- [179] T. Braine, R. Cervantes, N. Crisosto, et al. Extended search for the invisible axion with the axion dark matter experiment. *Phys. Rev. Lett.*, 124:101303, March 2020. doi: 10.1103/PhysRevLett.124.101303. URL <https://link.aps.org/doi/10.1103/PhysRevLett.124.101303>.
- [180] Katherine Garrett and Gintaras Duda. Dark matter: A primer. *Advances in Astronomy*, 2011:1–22, 2011. ISSN 1687-7977. doi: 10.1155/2011/968283. URL <http://dx.doi.org/10.1155/2011/968283>.

-
- [181] JonHMDavis. File:Direct Detection Constraints.png - Wikipedia — en.wikipedia.org. https://en.wikipedia.org/wiki/File:Direct_Detection_Constraints.png, 2017. [Accessed 04-06-2024].
- [182] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Physics Reports*, 267(5–6):195–373, March 1996. ISSN 0370-1573. doi: 10.1016/0370-1573(95)00058-5. URL [http://dx.doi.org/10.1016/0370-1573\(95\)00058-5](http://dx.doi.org/10.1016/0370-1573(95)00058-5).
- [183] Bob Holdom. Two $u(1)$'s and \square charge shifts. *Physics Letters B*, 166(2):196–198, January 1986. ISSN 0370-2693. doi: 10.1016/0370-2693(86)91377-8. URL [http://dx.doi.org/10.1016/0370-2693\(86\)91377-8](http://dx.doi.org/10.1016/0370-2693(86)91377-8).
- [184] Peter Galison and Aneesh Manohar. Two z 's or not two z 's? *Physics Letters B*, 136(4):279–283, March 1984. ISSN 0370-2693. doi: 10.1016/0370-2693(84)91161-4. URL [http://dx.doi.org/10.1016/0370-2693\(84\)91161-4](http://dx.doi.org/10.1016/0370-2693(84)91161-4).
- [185] Nima Arkani-Hamed and Neal Weiner. Lhc signals for a superunified theory of dark matter. *Journal of High Energy Physics*, 2008(12):104–104, December 2008. ISSN 1029-8479. doi: 10.1088/1126-6708/2008/12/104. URL <http://dx.doi.org/10.1088/1126-6708/2008/12/104>.
- [186] Pierre Fayet. U-boson production in $e+e-$ annihilations, ψ and ψ decays, and light dark matter. *Physical Review D*, 75(11), June 2007. ISSN 1550-2368. doi: 10.1103/physrevd.75.115017. URL <http://dx.doi.org/10.1103/PhysRevD.75.115017>.
- [187] Clifford Cheung, Joshua T. Ruderman, Lian-Tao Wang, et al. Kinetic mixing as the origin of a light dark-gauge-group scale. *Physical Review D*, 80(3), August 2009. ISSN 1550-2368. doi: 10.1103/physrevd.80.035008. URL <http://dx.doi.org/10.1103/PhysRevD.80.035008>.
- [188] David E Morrissey, David Poland, and Kathryn M Zurek. Abelian hidden sectors at a gev. *Journal of High Energy Physics*, 2009(07):050–050, July 2009. ISSN 1029-8479. doi: 10.1088/1126-6708/2009/07/050. URL <http://dx.doi.org/10.1088/1126-6708/2009/07/050>.
- [189] Mark Goodsell, Joerg Jaeckel, Javier Redondo, et al. Naturally light hidden photons in large volume string compactifications. *Journal of High Energy Physics*, 2009(11):027–027, November 2009. ISSN 1029-8479. doi: 10.1088/1126-6708/2009/11/027. URL <http://dx.doi.org/10.1088/1126-6708/2009/11/027>. arXiv:0909.0515 [hep-ph, physics:hep-th].
- [190] Michele Cicoli, Mark Goodsell, Joerg Jaeckel, et al. Testing string vacua in the lab: from a hidden cmb to dark forces in flux compactifications. *Journal of High Energy Physics*, 2011(7), July 2011. ISSN 1029-8479. doi: 10.1007/jhep07(2011)114. URL [http://dx.doi.org/10.1007/JHEP07\(2011\)114](http://dx.doi.org/10.1007/JHEP07(2011)114).
- [191] Marco Fabbrichesi, Emidio Gabrielli, and Gaia Lanfranchi. *The Physics of the Dark Photon: A Primer*. Springer International Publishing, 2021. ISBN 9783030625191. doi: 10.1007/978-3-030-62519-1. URL <http://dx.doi.org/10.1007/978-3-030-62519-1>.

-
- [192] R. Cervantes, G. Carosi, S. Kimes, et al. Admx-orpheus first search for dark photon dark matter: Detailed design, operations, and analysis. *Physical Review D*, 106(10), November 2022. ISSN 2470-0029. doi: 10.1103/physrevd.106.102002. URL <http://dx.doi.org/10.1103/PhysRevD.106.102002>.
- [193] Nils Siemonsen, Cristina Mondino, Daniel Egaña Ugrinovic, et al. Dark photon superradiance: Electrodynamics and multimessenger signals. *Phys. Rev. D*, 107:075025, April 2023. doi: 10.1103/PhysRevD.107.075025. URL <https://link.aps.org/doi/10.1103/PhysRevD.107.075025>.
- [194] Andrea Caputo, Samuel J. Witte, Diego Blas, et al. Electromagnetic signatures of dark photon superradiance. *Physical Review D*, 104(4), August 2021. ISSN 2470-0029. doi: 10.1103/physrevd.104.043006. URL <http://dx.doi.org/10.1103/PhysRevD.104.043006>.
- [195] Enrico Cannizzaro, Laura Sberna, Andrea Caputo, et al. Dark photon superradiance quenched by dark matter. *Physical Review D*, 106(8), October 2022. ISSN 2470-0029. doi: 10.1103/physrevd.106.083019. URL <http://dx.doi.org/10.1103/PhysRevD.106.083019>.
- [196] Amit Bhoonah, Joseph Bramante, and Ningqiang Song. Superradiant searches for dark photons in two stage atomic transitions. *Phys. Rev. D*, 101:055040, March 2020. doi: 10.1103/PhysRevD.101.055040. URL <https://link.aps.org/doi/10.1103/PhysRevD.101.055040>.
- [197] William E. East and Frans Pretorius. Superradiant instability and backreaction of massive vector fields around kerr black holes. *Phys. Rev. Lett.*, 119(4):041101, July 2017. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.119.041101. URL <https://link.aps.org/doi/10.1103/PhysRevLett.119.041101>.
- [198] Vitor Cardoso, Óscar J.C. Dias, Gavin S. Hartnett, et al. Constraining the mass of dark photons and axion-like particles through black-hole superradiance. *Journal of Cosmology and Astroparticle Physics*, 2018(03):043–043, March 2018. ISSN 1475-7516. doi: 10.1088/1475-7516/2018/03/043. URL <http://dx.doi.org/10.1088/1475-7516/2018/03/043>.
- [199] Shaun Fell, Lavinia Heisenberg, and Doğa Veske. Detecting fundamental vector fields with lisa. *Physical Review D*, 108(8):083010, October 2023. ISSN 2470-0029. doi: 10.1103/physrevd.108.083010. URL <http://dx.doi.org/10.1103/PhysRevD.108.083010>.
- [200] Richard Brito, Vitor Cardoso, and Paolo Pani. *Superradiance*, volume 906 of *Lecture Notes in Physics*. Springer International Publishing, Cham, 2015. ISBN 978-3-319-18999-4. doi: 10.1007/978-3-319-19000-6. URL <http://link.springer.com/10.1007/978-3-319-19000-6>.
- [201] R. Essig, J. A. Jaros, W. Wester, et al. Dark sectors and new, light, weakly-coupled particles, 2013.
- [202] Valery Rubakov. *Classical theory of gauge fields*. Princeton University Press, Princeton, NJ, May 2002.

-
- [203] HENRI RUEGG and MARTÍ RUIZ-ALTABA. The stueckelberg field. *International Journal of Modern Physics A*, 19(20):3265–3347, August 2004. ISSN 1793-656X. doi: 10.1142/S0217751X04019755. URL <http://dx.doi.org/10.1142/S0217751X04019755>.
- [204] Al. Proca. Sur la théorie ondulatoire des électrons positifs et négatifs. *Journal de Physique et le Radium*, 7(8):347–353, 1936. ISSN 0368-3842. doi: 10.1051/jphysrad:0193600708034700. URL <http://dx.doi.org/10.1051/jphysrad:0193600708034700>.
- [205] Al. Proca. Théorie non relativiste des particules à spin entier. *Journal de Physique et le Radium*, 9(2):61–66, 1938. ISSN 0368-3842. doi: 10.1051/jphysrad:019380090206100. URL <http://dx.doi.org/10.1051/jphysrad:019380090206100>.
- [206] E. R. Williams, J. E. Faller, and H. A. Hill. New experimental test of coulomb’s law: A laboratory upper limit on the photon rest mass. *Phys. Rev. Lett.*, 26:721–724, March 1971. doi: 10.1103/PhysRevLett.26.721. URL <https://link.aps.org/doi/10.1103/PhysRevLett.26.721>.
- [207] G V Chibisov. Astrophysical upper limits on the photon rest mass. *Soviet Physics Uspekhi*, 19(7):624–626, July 1976. ISSN 0038-5670. doi: 10.1070/PU1976v019n07ABEH005277. URL <http://dx.doi.org/10.1070/PU1976v019n07ABEH005277>.
- [208] Eric Adelberger, Gia Dvali, and Andrei Gruzinov. Photon-mass bound destroyed by vortices. *Phys. Rev. Lett.*, 98:010402, January 2007. doi: 10.1103/PhysRevLett.98.010402. URL <https://link.aps.org/doi/10.1103/PhysRevLett.98.010402>.
- [209] Gianmassimo Tasinato. Cosmic acceleration from abelian symmetry breaking. *Journal of High Energy Physics*, 2014(4):225004, April 2014. ISSN 1029-8479. doi: 10.1007/JHEP04(2014)067. URL [http://dx.doi.org/10.1007/JHEP04\(2014\)067](http://dx.doi.org/10.1007/JHEP04(2014)067).
- [210] Lavinia Heisenberg. Generalization of the proca action. *Journal of Cosmology and Astroparticle Physics*, 2014(05):015–015, May 2014. ISSN 1475-7516. doi: 10.1088/1475-7516/2014/05/015. URL <http://dx.doi.org/10.1088/1475-7516/2014/05/015>. arXiv: 1402.7026.
- [211] Antonio De Felice, Lavinia Heisenberg, Ryotaro Kase, et al. Cosmology in generalized proca theories. *Journal of Cosmology and Astroparticle Physics*, 2016(06):048–048, June 2016. ISSN 1475-7516. doi: 10.1088/1475-7516/2016/06/048. URL <http://dx.doi.org/10.1088/1475-7516/2016/06/048>.
- [212] Sebastian Garcia-Saenz, Aaron Held, and Jun Zhang. Destabilization of black holes and stars by generalized proca fields. *Physical Review Letters*, 127(13), September 2021. ISSN 1079-7114. doi: 10.1103/physrevlett.127.131104. URL <http://dx.doi.org/10.1103/PhysRevLett.127.131104>.
- [213] Masato Minamitsuji. Black holes in the generalized proca theory. *General Relativity and Gravitation*, 49(7), June 2017. ISSN 1572-9532. doi: 10.1007/s10714-017-2250-7. URL <http://dx.doi.org/10.1007/s10714-017-2250-7>.
- [214] Mostafizur Rahman and Anjan A. Sen. Astrophysical signatures of black holes in generalized proca theories. *Physical Review D*, 99(2), January 2019. ISSN 2470-0029. doi: 10.1103/physrevd.99.024052. URL <http://dx.doi.org/10.1103/PhysRevD.99.024052>.

-
- [215] Lavinia Heisenberg, Ryotaro Kase, Masato Minamitsuji, et al. Hairy black-hole solutions in generalized proca theories. *Phys. Rev. D*, 96(8):084049, 2017. doi: 10.1103/PhysRevD.96.084049.
- [216] Katy Clough, Thomas Helfer, Helvi Witek, et al. The problem with proca: ghost instabilities in self-interacting vector fields. *Physical Review Letters*, 129(15):151102, October 2022. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.129.151102. URL <http://dx.doi.org/10.1103/PhysRevLett.129.151102>. arXiv:2204.10868 [astro-ph, physics:gr-qc, physics:hep-ph].
- [217] Andrew Coates and Fethi M. Ramazanoğlu. Intrinsic pathology of self-interacting vector fields. *Physical Review Letters*, 129(15), October 2022. ISSN 1079-7114. doi: 10.1103/physrevlett.129.151103. URL <http://dx.doi.org/10.1103/PhysRevLett.129.151103>.
- [218] Andrew Coates and Fethi M. Ramazanoğlu. Pervasiveness of the breakdown of self-interacting vector field theories. *Phys. Rev. D*, 107:104036, May 2023. doi: 10.1103/PhysRevD.107.104036. URL <https://link.aps.org/doi/10.1103/PhysRevD.107.104036>.
- [219] Kivanç İ. Ünlütürk, Andrew Coates, and Fethi M. Ramazanoğlu. Loss of hyperbolicity and tachyons in generalized proca theories. *Physical Review D*, 108(4), August 2023. ISSN 2470-0029. doi: 10.1103/physrevd.108.044022. URL <http://dx.doi.org/10.1103/PhysRevD.108.044022>.
- [220] Enrico Barausse, Miguel Bezares, Marco Crisostomi, et al. The well-posedness of the cauchy problem for self-interacting vector fields. *Journal of Cosmology and Astroparticle Physics*, 2022(11):050, November 2022. ISSN 1475-7516. doi: 10.1088/1475-7516/2022/11/050. URL <http://dx.doi.org/10.1088/1475-7516/2022/11/050>.
- [221] Saul A. Teukolsky. Perturbations of a rotating black hole. i. fundamental equations for gravitational, electromagnetic, and neutrino-field perturbations. *The Astrophysical Journal*, 185:635, October 1973. ISSN 0004-637X, 1538-4357. doi: 10.1086/152444.
- [222] Charles W Misner, etc., Kip S Thorne, et al. *Gravitation*. W.H. Freeman, New York, NY, January 1973.
- [223] Susan G Hahn and Richard W Lindquist. The two-body problem in geometrodynamics. *Annals of Physics*, 29(2):304–331, September 1964. ISSN 0003-4916. doi: 10.1016/0003-4916(64)90223-4. URL [http://dx.doi.org/10.1016/0003-4916\(64\)90223-4](http://dx.doi.org/10.1016/0003-4916(64)90223-4).
- [224] Tsvi Piran. Cylindrical general relativistic collapse. *Phys. Rev. Lett.*, 41:1085–1088, October 1978. doi: 10.1103/PhysRevLett.41.1085. URL <https://link.aps.org/doi/10.1103/PhysRevLett.41.1085>.
- [225] Richard F. Stark and Tsvi Piran. Gravitational-wave emission from rotating gravitational collapse. *Phys. Rev. Lett.*, 55:891–894, August 1985. doi: 10.1103/PhysRevLett.55.891. URL <https://link.aps.org/doi/10.1103/PhysRevLett.55.891>.
- [226] Richard A. Matzner, H. E. Seidel, Stuart L. Shapiro, et al. Geometry of a black hole collision. *Science*, 270(5238):941–947, November 1995. ISSN 1095-9203. doi: 10.1126/science.270.5238.941. URL <http://dx.doi.org/10.1126/science.270.5238.941>.
- [227] Mark A. Scheel. [tapir.caltech.edu. http://www.tapir.caltech.edu/~rom-gr/slides/Mark_Scheel.pdf](http://www.tapir.caltech.edu/~rom-gr/slides/Mark_Scheel.pdf), 2013. [Accessed 05-06-2024].

-
- [228] Miguel Alcubierre and Bernd Brügmann. Simple excision of a black hole in 3 + 1 numerical relativity. *Phys. Rev. D*, 63:104006, April 2001. doi: 10.1103/PhysRevD.63.104006. URL <https://link.aps.org/doi/10.1103/PhysRevD.63.104006>.
- [229] Frans Pretorius. Evolution of binary black-hole spacetimes. *Phys. Rev. Lett.*, 95:121101, September 2005. doi: 10.1103/PhysRevLett.95.121101. URL <https://link.aps.org/doi/10.1103/PhysRevLett.95.121101>.
- [230] Steven Brandt and Bernd Brügmann. A simple construction of initial data for multiple black holes. *prl*, 78(19):3606–3609, May 1997. doi: 10.1103/PhysRevLett.78.3606.
- [231] Charles R Evans, Lee S Finn, and David W Hobill, editors. *Frontiers in numerical relativity*. Cambridge University Press, Cambridge, England, April 1989.
- [232] Matthew W. Choptuik, Eric W. Hirschmann, Steven L. Liebling, et al. Critical collapse of the massless scalar field in axisymmetry. *Phys. Rev. D*, 68:044007, August 2003. doi: 10.1103/PhysRevD.68.044007. URL <https://link.aps.org/doi/10.1103/PhysRevD.68.044007>.
- [233] Breno Imbiriba, John Baker, Dae-Il Choi, et al. Evolving a puncture black hole with fixed mesh refinement. *prd*, 70(12):124025, December 2004. doi: 10.1103/PhysRevD.70.124025.
- [234] David R. Fiske, John G. Baker, James R. van Meter, et al. Wave zone extraction of gravitational radiation in three-dimensional numerical relativity. *prd*, 71(10):104036, May 2005. doi: 10.1103/PhysRevD.71.104036.
- [235] Samuel Cupp et al. The einstein toolkit, December 2023. URL <https://doi.org/10.5281/zenodo.10380404>.
- [236] Harald P. Pfeiffer, Lawrence E. Kidder, Mark A. Scheel, et al. A multidomain spectral method for solving elliptic equations. *Comput. Phys. Commun.*, 152:253–273, 2003. doi: 10.1016/S0010-4655(02)00847-0.
- [237] Bernd Bruegmann, Jose A. Gonzalez, Mark Hannam, et al. Calibration of moving puncture simulations. *Phys. Rev. D*, 77:024027, 2008. doi: 10.1103/PhysRevD.77.024027.
- [238] Ian Ruchlin, Zachariah B. Etienne, and Thomas W. Baumgarte. Senr/nrpy+: Numerical relativity in singular curvilinear coordinate systems. *Phys. Rev. D*, 97(6):064036, 2018. doi: 10.1103/PhysRevD.97.064036.
- [239] Tomas Andrade, Llibert Areste Salo, Josu C. Aurrekoetxea, et al. Grchombo: An adaptable numerical relativity code for fundamental physics. *Journal of Open Source Software*, 6(68): 3703, 2021. doi: 10.21105/joss.03703. URL <https://doi.org/10.21105/joss.03703>.
- [240] Katy Clough, Pau Figueras, Hal Finkel, et al. Grchombo: Numerical relativity with adaptive mesh refinement. *Classical and Quantum Gravity*, 32(24):245011, December 2015. doi: 10.1088/0264-9381/32/24/245011. URL <https://dx.doi.org/10.1088/0264-9381/32/24/245011>.
- [241] Jean-Alain Marck and Ericourgoulhon. LORENE — lorene.obspm.fr. <https://lorene.obspm.fr/>, 1997. [Accessed 06-06-2024].

-
- [242] Zhoujian Cao, Hwei-Jang Yo, and Jui-Ping Yu. Reinvestigation of moving punctured black holes with a new code. *Phys. Rev. D*, 78:124011, December 2008. doi: 10.1103/PhysRevD.78.124011. URL <https://link.aps.org/doi/10.1103/PhysRevD.78.124011>.
- [243] ZHOUIJIAN CAO. New method for characteristic evolutions in numerical relativity. *International Journal of Modern Physics D*, 22(08):1350042, 2013. doi: 10.1142/S0218271813500429. URL <https://doi.org/10.1142/S0218271813500429>.
- [244] The LIGO Scientific Collaboration and the Virgo Collaboration and the KAGRA Collaboration et al. Tests of general relativity with gwtc-3, 2021.
- [245] The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, et al. Constraints on dark photon dark matter using data from ligo’s and virgo’s third observing run. *Physical Review D*, 105(6):063030, March 2022. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.105.063030. URL <http://dx.doi.org/10.1103/PhysRevD.105.063030>. arXiv:2105.13085 [astro-ph, physics:gr-qc, physics:hep-ph].
- [246] Haipeng An, Maxim Pospelov, Josef Pradler, et al. New limits on dark photons from solar emission and kev scale dark matter. *Physical Review D*, 102(11):115022, December 2020. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.102.115022.
- [247] Francesco D’Eramo, Eleonora Di Valentino, William Giarè, et al. Cosmological bound on the QCD axion mass, redux. *Journal of Cosmology and Astroparticle Physics*, 2022(09):022, September 2022. doi: 10.1088/1475-7516/2022/09/022. URL <https://doi.org/10.1088/1475-7516/2022/09/022>.
- [248] N. P. Pitjev and E. V. Pitjeva. Constraints on dark matter in the solar system. *Astronomy Letters*, 39(3):141–149, March 2013. doi: 10.1134/S1063773713020060. URL <https://doi.org/10.1134/S1063773713020060>.
- [249] Emil Polisensky and Massimo Ricotti. Constraints on the dark matter particle mass from the number of milky way satellites. *Physical Review D*, 83(4), February 2011. doi: 10.1103/physrevd.83.043506. URL <https://doi.org/10.1103/physrevd.83.043506>.
- [250] J Aasi, B P Abbott, R Abbott, et al. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, March 2015. doi: 10.1088/0264-9381/32/7/074001. URL <https://doi.org/10.1088/0264-9381/32/7/074001>.
- [251] T Accadia, F Acernese, M Alshourbagy, et al. Virgo: a laser interferometer to detect gravitational waves. *Journal of Instrumentation*, 7(03):P03012–P03012, March 2012. ISSN 1748-0221. doi: 10.1088/1748-0221/7/03/P03012.
- [252] F. Acernese, M. Agathos, K. Agatsuma, et al. Advanced virgo: a 2nd generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, January 2015. ISSN 0264-9381, 1361-6382. doi: 10.1088/0264-9381/32/2/024001. arXiv:1408.3978 [gr-qc, physics:physics].
- [253] T Akutsu, M Ando, K Arai, et al. Overview of kagra: Detector design and construction history. *Progress of Theoretical and Experimental Physics*, 2021(5):05A101, May 2021. ISSN 2050-3911. doi: 10.1093/ptep/ptaa125.

-
- [254] KAGRA collaboration et al. Kagra: 2.5 generation interferometric gravitational wave detector. *Nature Astronomy*, 3(1):35–40, January 2019. ISSN 2397-3366. doi: 10.1038/s41550-018-0658-y.
- [255] K. Danzmann. Lisa mission overview. *Advances in Space Research*, 25(6):1129–1136, January 2000. ISSN 02731177. doi: 10.1016/S0273-1177(99)00973-4.
- [256] Pau Amaro-Seoane, Heather Audley, Stanislav Babak, et al. Laser interferometer space antenna, 2017. URL <https://arxiv.org/abs/1702.00786>.
- [257] et-gw.eu. <https://www.et-gw.eu/>, 2023. [Accessed 05-Apr-2023].
- [258] David Reitze, Rana X Adhikari, Stefan Ballmer, et al. Cosmic explorer: The u.s. contribution to gravitational-wave astronomy beyond ligo, 2019.
- [259] Sheila Dwyer, Daniel Sigg, Stefan W. Ballmer, et al. Gravitational wave detector with cosmological reach. *Phys. Rev. D*, 91:082001, April 2015. doi: 10.1103/PhysRevD.91.082001. URL <https://link.aps.org/doi/10.1103/PhysRevD.91.082001>.
- [260] Naoki Seto, Seiji Kawamura, and Takashi Nakamura. Possibility of direct measurement of the acceleration of the universe using 0.1 hz band laser interferometer gravitational wave antenna in space. *Physical Review Letters*, 87(22):221103, November 2001. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.87.221103.
- [261] Seiji Kawamura, Takashi Nakamura, Masaki Ando, et al. Space gravitational-wave antennas decigo and b-decigo. *International Journal of Modern Physics D*, 28(12):1845001, 2019. doi: 10.1142/S0218271818450013.
- [262] J Antoniadis, Z Arzoumanian, S Babak, et al. The international pulsar timing array second data release: Search for an isotropic gravitational wave background. *Monthly Notices of the Royal Astronomical Society*, 510(4):4873–4887, January 2022. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stab3418.
- [263] B. B. P. Perera, M. E. DeCesar, P. B. Demorest, et al. The international pulsar timing array: Second data release. *Monthly Notices of the Royal Astronomical Society*, 490(4):4666–4687, December 2019. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stz2857. arXiv:1909.04534 [astro-ph].
- [264] Xing-Jiang Zhu, Linqing Wen, George Hobbs, et al. Detection and localization of single-source gravitational waves with pulsar timing arrays. *Monthly Notices of the Royal Astronomical Society*, 449(2):1650–1663, May 2015. ISSN 1365-2966, 0035-8711. doi: 10.1093/mnras/stv381. arXiv:1502.06001 [astro-ph, physics:gr-qc].
- [265] M. T. Lam, M. A. McLaughlin, Z. Arzoumanian, et al. The nanograv 12.5-year data set: The frequency dependence of pulse jitter in precision millisecond pulsars. *The Astrophysical Journal*, 872(2):193, February 2019. ISSN 1538-4357. doi: 10.3847/1538-4357/ab01cd. arXiv:1809.03058 [astro-ph].
- [266] R. N. Manchester. The parkes pulsar timing array. *Chinese Journal of Astronomy and Astrophysics*, 6(S2):139, October 2006. doi: 10.1088/1009-9271/6/S2/27. URL <https://dx.doi.org/10.1088/1009-9271/6/S2/27>.

-
- [267] Emanuele Berti, Enrico Barausse, Ilias Cholis, et al. Tests of general relativity and fundamental physics with space-based gravitational wave detectors, 2019.
- [268] Girish Mahajan. Extreme mass ratio inspiral - Alchetron, the free social encyclopedia — alchetron.com. <https://alchetron.com/Extreme-mass-ratio-inspiral#extreme-mass-ratio-inspiral-f7170df3-8a37-4335-b5d2-83b7984da22-resize-750.jpg>, 2023. [Accessed 07-06-2024].
- [269] Stanislav Babak, Hua Fang, Jonathan R. Gair, et al. “kludge” gravitational waveforms for a test-body orbiting a kerr black hole. *arXiv:gr-qc/0607007*, January 2008. doi: 10.1103/PhysRevD.77.04990. URL <http://arxiv.org/abs/gr-qc/0607007>. arXiv: gr-qc/0607007.
- [270] Leor Barack and Curt Cutler. Lisa capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy. *Physical Review D*, 69(8):082005, April 2004. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.69.082005.
- [271] Alvin J. K. Chua, Christopher J. Moore, and Jonathan R. Gair. Augmented kludge waveforms for detecting extreme-mass-ratio inspirals. *Phys. Rev. D*, 96:044005, August 2017. doi: 10.1103/PhysRevD.96.044005. URL <https://link.aps.org/doi/10.1103/PhysRevD.96.044005>.
- [272] S. Sigurdsson and M. J. Rees. Capture of stellar mass compact objects by massive black holes in galactic cusps. *Monthly Notices of the Royal Astronomical Society*, 284(2):318–326, January 1997. doi: 10.1093/mnras/284.2.318. URL <https://doi.org/10.1093%2Fmnras%2F284.2.318>.
- [273] Tal Alexander. Stellar processes near the massive black hole in the galactic center. *Physics Reports*, 419(2-3):65–142, November 2005. doi: 10.1016/j.physrep.2005.08.002. URL <https://doi.org/10.1016%2Fj.physrep.2005.08.002>.
- [274] David Merritt. Dynamics of galaxy cores and supermassive black holes. *Reports on Progress in Physics*, 69(9):R01, September 2006. doi: 10.1088/0034-4885/69/9/r01. URL <https://doi.org/10.1088%2F0034-4885%2F69%2F9%2Fr01>.
- [275] Elisa Bortolas and Michela Mapelli. Can supernova kicks trigger EMRIs in the galactic centre? *Monthly Notices of the Royal Astronomical Society*, 485(2):2125–2138, February 2019. doi: 10.1093/mnras/stz440. URL <https://doi.org/10.1093%2Fmnras%2Fstz440>.
- [276] Pau Amaro-Seoane, Jonathan R Gair, Marc Freitag, et al. Intermediate and extreme mass-ratio inspirals—astrophysics, science applications and detection using LISA. *Classical and Quantum Gravity*, 24(17):R113–R169, August 2007. doi: 10.1088/0264-9381/24/17/r01. URL <https://doi.org/10.1088%2F0264-9381%2F24%2F17%2Fr01>.
- [277] Jonathan R. Gair, Michele Vallisneri, Shane L. Larson, et al. Testing general relativity with low-frequency, space-based gravitational-wave detectors. *Living Reviews in Relativity*, 16(1):7, December 2013. ISSN 2367-3613, 1433-8351. doi: 10.12942/lrr-2013-7. arXiv:1212.5575 [gr-qc].
- [278] Jonathan R Gair, Stanislav Babak, Alberto Sesana, et al. Prospects for observing extreme-mass-ratio inspirals with lisa. *Journal of Physics: Conference Series*, 840(1):012021, May 2017. doi: 10.1088/1742-6596/840/1/012021. URL <https://dx.doi.org/10.1088/1742-6596/840/1/012021>.

-
- [279] Lam Hui, Jeremiah P. Ostriker, Scott Tremaine, et al. Ultralight scalars as cosmological dark matter. *Physical Review D*, 95(4):043541, February 2017. ISSN 2470-0010, 2470-0029. doi: 10.1103/physrevd.95.043541. URL <https://doi.org/10.1103/PhysRevD.95.043541>. arXiv:1610.08297 [astro-ph, physics:hep-ph, physics:hep-th].
- [280] Andrew L. Miller, Sébastien Clesse, Federico De Lillo, et al. Probing planetary-mass primordial black holes with continuous gravitational waves. *Physics of the Dark Universe*, 32: 100836, May 2021. ISSN 22126864. doi: 10.1016/j.dark.2021.100836. arXiv:2012.12983 [astro-ph, physics:gr-qc].
- [281] Shu-Rong Chen, Hsi-Yu Schive, and Tzihong Chiueh. Jeans analysis for dwarf spheroidal galaxies in wave dark matter. *Monthly Notices of the Royal Astronomical Society*, 468(2): 1338–1348, June 2017. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stx449. arXiv:1606.09030 [astro-ph].
- [282] Masha Baryakhtar, Robert Lasenby, and Mae Teo. Black hole superradiance signatures of ultralight vectors. *Physical Review D*, 96(3):035019, August 2017. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.96.035019. URL <http://dx.doi.org/10.1103/PhysRevD.96.035019>. arXiv: 1704.05081.
- [283] William E. East. Superradiant instability of massive vector fields around spinning black holes in the relativistic regime. *Physical Review D*, 96(2):024004, July 2017. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.96.024004. URL <http://dx.doi.org/10.1103/PhysRevD.96.024004>. arXiv:1705.01544 [astro-ph, physics:gr-qc, physics:hep-ph].
- [284] Aaron Pierce, Keith Riles, and Yue Zhao. Searching for dark photon dark matter with gravitational wave detectors. *Physical Review Letters*, 121(6):061102, August 2018. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.121.061102. arXiv:1801.10161 [astro-ph, physics:gr-qc, physics:hep-ph].
- [285] Andrea Maselli, Nicola Franchini, Leonardo Gualtieri, et al. Detecting scalar fields with extreme mass ratio inspirals. *Physical Review Letters*, 125(14):141101, September 2020. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.125.141101. arXiv:2004.11895 [astro-ph, physics:gr-qc, physics:hep-ph].
- [286] Andrea Maselli, Nicola Franchini, Leonardo Gualtieri, et al. Detecting fundamental fields with lisa observations of gravitational waves from extreme mass-ratio inspirals. *Nature Astronomy*, 6(4):464–470, February 2022. ISSN 2397-3366. doi: 10.1038/s41550-021-01589-5. arXiv:2106.11325 [astro-ph, physics:gr-qc, physics:hep-th].
- [287] Nils Siemonsen, Taillte May, and William E. East. Superrad: A black hole superradiance gravitational waveform model, 2022. URL <https://arxiv.org/abs/2211.03845>.
- [288] Richard Brito, Vitor Cardoso, and Paolo Pani. Black holes as particle detectors: evolution of superradiant instabilities. *Classical and Quantum Gravity*, 32(13):134001, June 2015. ISSN 1361-6382. doi: 10.1088/0264-9381/32/13/134001. URL <http://dx.doi.org/10.1088/0264-9381/32/13/134001>.
- [289] Nuno M. Santos, Carolina L. Benone, Luís C.B. Crispino, et al. Black holes with synchronised procia hair: linear clouds and fundamental non-linear solutions. *Journal of High Energy Physics*, 2020(7):10, July 2020. ISSN 1029-8479. doi: 10.1007/jhep07(2020)010. URL [http://dx.doi.org/10.1007/JHEP07\(2020\)010](http://dx.doi.org/10.1007/JHEP07(2020)010). arXiv: 2004.09536.

-
- [290] Paolo Pani, Vitor Cardoso, Leonardo Gualtieri, et al. Black-hole bombs and photon-mass bounds. *Physical Review Letters*, 109(13), September 2012. ISSN 1079-7114. doi: 10.1103/physrevlett.109.131102. URL <http://dx.doi.org/10.1103/PhysRevLett.109.131102>.
- [291] WILLIAM H. PRESS and SAUL A. TEUKOLSKY. Floating orbits, superradiant scattering and the black-hole bomb. *Nature*, 238(5361):211–212, July 1972. ISSN 1476-4687. doi: 10.1038/238211a0. URL <http://dx.doi.org/10.1038/238211a0>.
- [292] Hirotaka Yoshino and Hideo Kodama. Gravitational radiation from an axion cloud around a black hole: Superradiant phase. *Progress of Theoretical and Experimental Physics*, 2014(4), April 2014. ISSN 2050-3911. doi: 10.1093/ptep/ptu029. URL <http://arxiv.org/abs/1312.2326>. arXiv:1312.2326 [gr-qc, physics:hep-ph, physics:hep-th].
- [293] Nils Siemonsen and William E. East. Gravitational wave signatures of ultralight vector bosons from black hole superradiance. *Physical Review D*, 101(2):024019, January 2020. ISSN 2470-0029. doi: 10.1103/physrevd.101.024019. URL <https://doi.org/10.1103/2Fphysrevd.101.024019>. arXiv: 1910.09476.
- [294] Sam R. Dolan. Instability of the proca field on kerr spacetime. *Physical Review D*, 98(10): 104006, November 2018. ISSN 2470-0029. doi: 10.1103/physrevd.98.104006. URL <https://doi.org/10.1103/2Fphysrevd.98.104006>. arXiv: 1806.01604.
- [295] Carlos Herdeiro, Eugen Radu, and Helgi Rúnarsson. Kerr black holes with proca hair. *Classical and Quantum Gravity*, 33(15):154001, June 2016. ISSN 1361-6382. doi: 10.1088/0264-9381/33/15/154001. URL <http://dx.doi.org/10.1088/0264-9381/33/15/154001>. arXiv:1603.02687 [astro-ph, physics:gr-qc, physics:hep-th].
- [296] Steven Detweiler. Klein-gordon equation and rotating black holes. *Phys. Rev. D*, 22:2323–2326, November 1980. doi: 10.1103/PhysRevD.22.2323. URL <https://link.aps.org/doi/10.1103/PhysRevD.22.2323>.
- [297] William E. East. Massive boson superradiant instability of black holes: Nonlinear growth, saturation, and gravitational radiation. *Physical Review Letters*, 121(13):131104, September 2018. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.121.131104. URL <http://dx.doi.org/10.1103/PhysRevLett.121.131104>. arXiv:1807.00043 [astro-ph, physics:gr-qc, physics:hep-ph].
- [298] Nils Siemonsen, Tailte May, and William E. East. Modeling the black hole superradiance gravitational waveform. *Physical Review D*, 107(10), May 2023. ISSN 2470-0029. doi: 10.1103/physrevd.107.104003. URL <http://dx.doi.org/10.1103/PhysRevD.107.104003>.
- [299] Helvi Witek, Vitor Cardoso, Akihiro Ishibashi, et al. Superradiant instabilities in astrophysical systems. *Physical Review D*, 87(4), February 2013. ISSN 1550-2368. doi: 10.1103/physrevd.87.043513. URL <http://dx.doi.org/10.1103/PhysRevD.87.043513>.
- [300] Daniel Baumann, Horng Sheng Chia, Rafael A. Porto, et al. Gravitational collider physics. *Physical Review D*, 101(8):083019, April 2020. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.101.083019. arXiv:1912.04932 [gr-qc, physics:hep-ph, physics:hep-th].
- [301] Valeri P. Frolov, Pavel Krtoš, David Kubizňák, et al. Massive vector fields in rotating black-hole spacetimes: Separability and quasinormal modes. *Physical Review Letters*, 120(23): 231103, June 2018. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.120.231103. arXiv: 1804.00030.

-
- [302] Daniel Baumann, Horng Sheng Chia, and Rafael A. Porto. Probing ultralight bosons with binary black holes. *Physical Review D*, 99(4):044001, February 2019. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.99.044001. arXiv:1804.03208 [astro-ph, physics:gr-qc, physics:hep-ph, physics:hep-th].
- [303] Daniel Baumann, Horng Sheng Chia, John Stout, et al. The spectra of gravitational atoms. *Journal of Cosmology and Astroparticle Physics*, 2019(12):006–006, December 2019. ISSN 1475-7516. doi: 10.1088/1475-7516/2019/12/006. arXiv: 1908.10370.
- [304] Wolfram Research, Inc. Mathematica, Version 13.2, 2023. URL <https://www.wolfram.com/mathematica>. Champaign, IL, 2022.
- [305] S. A. Teukolsky and W. H. Press. Perturbations of a rotating black hole. iii. interaction of the hole with gravitational and electromagnetic radiation. *apj*, 193:443–461, October 1974. doi: 10.1086/153180.
- [306] W. G. Unruh. Second quantization in the kerr metric. *Phys. Rev. D*, 10:3194–3205, November 1974. doi: 10.1103/PhysRevD.10.3194. URL <https://link.aps.org/doi/10.1103/PhysRevD.10.3194>.
- [307] A A Starobinskii. Amplification of waves reflected from a rotating black hole. *Zh. Eksp. Teor. Fiz. 64: No. 1, 48-57(Jan 1973)*, January 1973. URL <https://www.osti.gov/biblio/4580095>.
- [308] Visakan Balakumar, Elizabeth Winstanley, Rafael P. Bernar, et al. Quantum superradiance on static black hole space-times. *Physics Letters B*, 811:135904, 2020. ISSN 0370-2693. doi: <https://doi.org/10.1016/j.physletb.2020.135904>. URL <https://www.sciencedirect.com/science/article/pii/S0370269320307073>.
- [309] Mauricio Richartz, Silke Weinfurtner, A. J. Penner, et al. Generalised superradiant scattering, 2009.
- [310] Mauricio Richartz, Silke Weinfurtner, A. J. Penner, et al. Generalized superradiant scattering. *Physical Review D*, 80(12), December 2009. ISSN 1550-2368. doi: 10.1103/physrevd.80.124016. URL <http://dx.doi.org/10.1103/PhysRevD.80.124016>.
- [311] Vitor Cardoso and Paolo Pani. Tidal acceleration of black holes and superradiance. *Classical and Quantum Gravity*, 30(4):045011, January 2013. ISSN 1361-6382. doi: 10.1088/0264-9381/30/4/045011. URL <http://dx.doi.org/10.1088/0264-9381/30/4/045011>.
- [312] Emanuele Berti, Richard Brito, Caio F. B. Macedo, et al. Ultralight boson cloud depletion in binary systems. *Physical Review D*, 99(10):104039, May 2019. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.99.104039. arXiv:1904.03131 [gr-qc].
- [313] Dina Traykova, Katy Clough, Thomas Helfer, et al. Dynamical friction from scalar dark matter in the relativistic regime. *Physical Review D*, 104(10), November 2021. doi: 10.1103/physrevd.104.103014. URL <https://doi.org/10.1103/PhysRevD.104.103014>. arXiv: 2106.08280.
- [314] Michael L. Katz, Alvin J. K. Chua, Lorenzo Speri, et al. Fast extreme-mass-ratio-inspiral waveforms: New tools for millihertz gravitational-wave data analysis. *Physical Review D*, 104(6), September 2021. doi: 10.1103/physrevd.104.064047. URL <https://doi.org/10.1103/PhysRevD.104.064047>.

-
- [315] P. C. Peters and J. Mathews. Gravitational radiation from point masses in a keplerian orbit. *Phys. Rev.*, 131:435–440, July 1963. doi: 10.1103/PhysRev.131.435. URL <https://link.aps.org/doi/10.1103/PhysRev.131.435>.
- [316] Paolo Pani, Vitor Cardoso, Leonardo Gualtieri, et al. Perturbations of slowly rotating black holes: massive vector fields in the kerr metric. *Physical Review D*, 86(10):104017, November 2012. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.86.104017. arXiv: 1209.0773.
- [317] Travis Robson, Neil J Cornish, and Chang Liu. The construction and use of LISA sensitivity curves. *Classical and Quantum Gravity*, 36(10):105011, April 2019. doi: 10.1088/1361-6382/ab1101. URL <https://doi.org/10.1088/1361-6382/ab1101>.
- [318] Lee Samuel Finn. Detection, measurement and gravitational radiation. *Physical Review D*, 46(12):5236–5249, December 1992. ISSN 0556-2821. doi: 10.1103/PhysRevD.46.5236. arXiv:gr-qc/9209010.
- [319] Thibault Damour, Bala R. Iyer, and B. S. Sathyaprakash. Improved filters for gravitational waves from inspiralling compact binaries. *Physical Review D*, 57(2):885–907, January 1998. ISSN 0556-2821, 1089-4918. doi: 10.1103/PhysRevD.57.885. arXiv:gr-qc/9708034.
- [320] Ian Harry, Stephen Privitera, Alejandro Bohé, et al. Searching for gravitational waves from compact binaries with precessing spins. *Physical Review D*, 94(2):024012, July 2016. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.94.024012. arXiv:1603.02444 [astro-ph, physics:gr-qc].
- [321] Katerina Chatziioannou, Antoine Klein, Nicolas Yunes, et al. Constructing gravitational waves from generic spin-precessing compact binary inspirals. *Physical Review D*, 95(10):104004, May 2017. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.95.104004. arXiv:1703.03967 [astro-ph, physics:gr-qc].
- [322] Lee Lindblom, Benjamin J. Owen, and Duncan A. Brown. Model waveform accuracy standards for gravitational wave data analysis. *Physical Review D*, 78(12), December 2008. doi: 10.1103/physrevd.78.124020. URL <https://doi.org/10.1103/physrevd.78.124020>.
- [323] Lavinia Heisenberg. A systematic approach to generalisations of general relativity and their cosmological implications. *Phys. Rept.*, 796:1–113, March 2019. ISSN 03701573. doi: 10.1016/j.physrep.2018.11.006. arXiv: 1807.01725.
- [324] Jose Beltrán Jiménez and Lavinia Heisenberg. Derivative self-interactions for a massive vector field. *Physics Letters B*, 757:405–411, June 2016. ISSN 0370-2693. doi: 10.1016/j.physletb.2016.04.017. URL <http://dx.doi.org/10.1016/j.physletb.2016.04.017>.
- [325] Ericourgoulhon. *3+1 Formalism in General Relativity: Bases of Numerical Relativity*. Lecture notes in physics. Springer Berlin Heidelberg, New York, NY, January 2012. ISBN 9783642245251. doi: 10.1007/978-3-642-24525-1. URL <http://dx.doi.org/10.1007/978-3-642-24525-1>.
- [326] Miguel Alcubierre. *Introduction to 3+1 numerical relativity*. International Series of Monographs on Physics. Oxford University Press, London, England, May 2008.
- [327] Y. Fourès-Bruhat. Théorème d’existence pour certains systèmes d’équations aux dérivées partielles non linéaires. *Acta Mathematica*, 88(0):141–225, 1952. ISSN 0001-5962. doi: 10.1007/bf02392131. URL <http://dx.doi.org/10.1007/BF02392131>.

-
- [328] Jeffrey Winicour. Characteristic evolution and matching. *Living Reviews in Relativity*, 15(1), January 2012. ISSN 1433-8351. doi: 10.12942/lrr-2012-2. URL <http://dx.doi.org/10.12942/lrr-2012-2>.
- [329] H (heinz-Otto) Kreiss. *Initial-boundary value problems and the Navier-Stokes equations*. Classics in applied mathematics. Society for Industrial and Applied Mathematics, Great Britain, 2004.
- [330] Oscar A. Reula. Hyperbolic methods for einstein’s equations. *Living Reviews in Relativity*, 1(1), January 1998. ISSN 1433-8351. doi: 10.12942/lrr-1998-3. URL <http://dx.doi.org/10.12942/lrr-1998-3>.
- [331] T. Nakamura, K. Oohara, and Y. Kojima. General relativistic collapse to black holes and gravitational waves from black holes. *Progress of Theoretical Physics Supplement*, 90:1–218, January 1987. doi: 10.1143/PTPS.90.1.
- [332] V H Weston and S He. Wave splitting of the telegraph equation in r_3 and its application to inverse scattering. *Inverse Problems*, 9(6):789–812, December 1993. ISSN 1361-6420. doi: 10.1088/0266-5611/9/6/013. URL <http://dx.doi.org/10.1088/0266-5611/9/6/013>.
- [333] Carles Bona, Joan Massó, Edward Seidel, et al. New formalism for numerical relativity. *Physical Review Letters*, 75(4):600–603, July 1995. ISSN 1079-7114. doi: 10.1103/physrevlett.75.600. URL <http://dx.doi.org/10.1103/PhysRevLett.75.600>.
- [334] Miguel Alcubierre. Hyperbolic slicings of spacetime: singularity avoidance and gauge shocks. *Classical and Quantum Gravity*, 20(4):607–623, January 2003. ISSN 0264-9381. doi: 10.1088/0264-9381/20/4/304. URL <http://dx.doi.org/10.1088/0264-9381/20/4/304>.
- [335] M. Adams, P. Colella, D. T. Graves, et al. Chombo software package for amr applications - design document, December 2015. URL <https://commons.lbl.gov/display/chombo>.
- [336] Heinz-Otto Kreiss and Godela Scherer. Method of lines for hyperbolic differential equations. *SIAM Journal on Numerical Analysis*, 29(3):640–646, June 1992. ISSN 0036-1429, 1095-7170. doi: 10.1137/0729041.
- [337] J D Lambert. *Numerical methods for ordinary differential systems*. John Wiley & Sons, Chichester, England, July 1991.
- [338] Y. Zlochower, J. G. Baker, M. Campanelli, et al. Accurate black hole evolutions by fourth-order numerical relativity. *Physical Review D*, 72(2), July 2005. ISSN 1550-2368. doi: 10.1103/physrevd.72.024021. URL <http://dx.doi.org/10.1103/PhysRevD.72.024021>.
- [339] Heinz-Otto Kreiss. Methods for the approximate solution of time dependent problems, 1973. URL <https://api.semanticscholar.org/CorpusID:118627871>.
- [340] M. Berger and I. Rigoutsos. An algorithm for point clustering and grid generation. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(5):1278–1286, October 1991. ISSN 00189472. doi: 10.1109/21.120081.
- [341] Katy Bethea. The olcf’s 2021 exascale computer. <https://www.olcf.ornl.gov/2018/02/13/frontier-olcfs-exascale-future/>, 2018. [Accessed 14-06-2024].

-
- [342] GRTLCollaboration. GitHub - GRTLCollaboration/GRTeclyn: Port of GRChombo to AMReX - under development! — github.com. <https://github.com/GRTLCollaboration/GRTeclyn>, 2024. [Accessed 14-06-2024].
- [343] Katy Clough. Continuity equations for general matter: applications in numerical relativity. *Classical and Quantum Gravity*, 38(16):167001, July 2021. ISSN 1361-6382. doi: 10.1088/1361-6382/ac10ee. URL <http://dx.doi.org/10.1088/1361-6382/ac10ee>.
- [344] Boyd. *Chebyshev and Fourier Spectral Meth.* Dover Books on Mathematics. Dover Publications, Mineola, NY, 2 edition, December 2001.
- [345] Lewis Fry 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 A*, 210:307–357, 1910. URL <https://api.semanticscholar.org/CorpusID:119559073>.
- [346] Joao G. Rosa and Sam R. Dolan. Massive vector fields on the schwarzschild spacetime: quasinormal modes and bound states. *Physical Review D*, 85(4):044043, February 2012. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.85.044043. arXiv: 1110.4494.
- [347] Curt Cutler, Daniel Kennefick, and Eric Poisson. Gravitational radiation reaction for bound motion around a schwarzschild black hole. *Physical Review D*, 50(6):3816–3835, September 1994. ISSN 0556-2821. doi: 10.1103/PhysRevD.50.3816.
- [348] Adam G. Riess, Gagandeep S. Anand, Wenlong Yuan, et al. Crowded no more: The accuracy of the hubble constant tested with high resolution observations of cepheids by jwst, 2023.
- [349] V. Ghirardini, E. Bulbul, E. Artis, et al. The srg/erosita all-sky survey: Cosmology constraints from cluster abundances in the western galactic hemisphere, 2024.
- [350] Jerry B. Griffiths and Jiří Podolský. *Plebański–Demiański solutions*, page 304–322. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2009.
- [351] J.F Plebanski and M Demianski. Rotating, charged, and uniformly accelerating mass in general relativity. *Annals of Physics*, 98(1):98–127, 1976. ISSN 0003-4916. doi: [https://doi.org/10.1016/0003-4916\(76\)90240-2](https://doi.org/10.1016/0003-4916(76)90240-2). URL <https://www.sciencedirect.com/science/article/pii/0003491676902402>.
- [352] Sarp Akcay and Richard Matzner. Kerr-de sitter universe. *Classical and Quantum Gravity*, 28(8):085012, April 2011. ISSN 0264-9381, 1361-6382. doi: 10.1088/0264-9381/28/8/085012. arXiv:1011.0479 [gr-qc].
- [353] Miguel Zilhão, Helvi Witek, and Vitor Cardoso. Nonlinear interactions between black holes and proca fields. *Classical and Quantum Gravity*, 32(23):234003, November 2015. ISSN 1361-6382. doi: 10.1088/0264-9381/32/23/234003. URL <http://dx.doi.org/10.1088/0264-9381/32/23/234003>. arXiv:1505.00797 [gr-qc, physics:hep-th].
- [354] Miguel Alcubierre. *Introduction to 3+1 Numerical Relativity*. Oxford University Press, 04 2008. ISBN 9780199205677. doi: 10.1093/acprof:oso/9780199205677.001.0001. URL <https://doi.org/10.1093/acprof:oso/9780199205677.001.0001>.

-
- [355] Rodrigo Vicente and Vitor Cardoso. Dynamical friction of black holes in ultralight dark matter. *Physical Review D*, 105(8), April 2022. doi: 10.1103/physrevd.105.083008. URL <https://doi.org/10.1103%2Fphysrevd.105.083008>.
- [356] Dina Traykova, Rodrigo Vicente, Katy Clough, et al. Relativistic drag forces on black holes from scalar dark matter clouds of all sizes, 2023.
- [357] Jun Zhang and Huan Yang. Dynamic signatures of black hole binaries with superradiant clouds. *Physical Review D*, 101(4), February 2020. doi: 10.1103/physrevd.101.043020. URL <https://doi.org/10.1103%2Fphysrevd.101.043020>.
- [358] Daniel Baumann, Gianfranco Bertone, John Stout, et al. Sharp signals of boson clouds in black hole binary inspirals. *Physical Review Letters*, 128(22), June 2022. doi: 10.1103/physrevlett.128.221102. URL <https://doi.org/10.1103%2Fphysrevlett.128.221102>.
- [359] Robin Buehler and Vincent Desjacques. Dynamical friction in fuzzy dark matter: Circular orbits. *Physical Review D*, 107(2), January 2023. doi: 10.1103/physrevd.107.023516. URL <https://doi.org/10.1103%2Fphysrevd.107.023516>.
- [360] Philippa S. Cole, Gianfranco Bertone, Adam Coogan, et al. Disks, spikes, and clouds: distinguishing environmental effects on bbh gravitational waveforms, 2022.
- [361] Giovanni Maria Tomaselli, Thomas F. M. Spieksma, and Gianfranco Bertone. Dynamical friction in gravitational atoms, 2023.
- [362] Daniel Baumann, Gianfranco Bertone, John Stout, et al. Ionization of gravitational atoms. *Physical Review D*, 105(11), June 2022. doi: 10.1103/physrevd.105.115036. URL <https://doi.org/10.1103%2Fphysrevd.105.115036>.
- [363] W. G. Unruh. Absorption cross section of small black holes. *Phys. Rev. D*, 14:3251–3259, December 1976. doi: 10.1103/PhysRevD.14.3251. URL <https://link.aps.org/doi/10.1103/PhysRevD.14.3251>.
- [364] Jun Zhang and Huan Yang. Gravitational floating orbits around hairy black holes. *Physical Review D*, 99(6), March 2019. doi: 10.1103/physrevd.99.064018. URL <https://doi.org/10.1103%2Fphysrevd.99.064018>.
- [365] Yan Cao and Yong Tang. Signatures of ultralight bosons in compact binary inspiral and outspiral, 2023.
- [366] Miguel C. Ferreira, Caio F. B. Macedo, and Vitor Cardoso. Orbital fingerprints of ultralight scalar fields around black holes. *Physical Review D*, 96(8), October 2017. doi: 10.1103/physrevd.96.083017. URL <https://doi.org/10.1103%2Fphysrevd.96.083017>.
- [367] Otto A. Hannuksela, Kaze W. K. Wong, Richard Brito, et al. Probing the existence of ultralight bosons with a single gravitational-wave measurement. *Nature Astronomy*, 3(5): 447–451, March 2019. doi: 10.1038/s41550-019-0712-4. URL <https://doi.org/10.1038%2Fs41550-019-0712-4>.
- [368] Shaun David Brocus Fell. Grboondi, 2024. URL <https://github.com/ShawnFell/GRBoondi>.

-
- [369] Fabio D'Ambrosio, Lavinia Heisenberg, and Stefan Zentarra. Hamiltonian analysis of $f(q)$ gravity and the failure of the dirac-bergmann algorithm for teleparallel theories of gravity, 2023.
- [370] Shaun David Brocus Fell and Lavinia Heisenberg. Grboondi: A code for evolving generalized proca theories on arbitrary backgrounds. 2024. doi: <https://doi.org/10.48550/arXiv.2405.01348>. URL <https://arxiv.org/abs/2405.01348>.
- [371] Shaun David Brocus Fell and Lavinia Heisenberg. Proca in an expanding universe. *Fortschritte der Physik*, 72(7-8):2400110, 2024. doi: <https://doi.org/10.1002/prop.202400110>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/prop.202400110>.
- [372] Fabio D'Ambrosio, Shaun D. B. Fell, Lavinia Heisenberg, et al. Black holes in $f(\mathbb{Q})$ gravity. *Phys. Rev. D*, 105:024042, January 2022. doi: [10.1103/PhysRevD.105.024042](https://doi.org/10.1103/PhysRevD.105.024042). URL <https://link.aps.org/doi/10.1103/PhysRevD.105.024042>.
- [373] Fabio D'Ambrosio, Shaun D. B. Fell, Lavinia Heisenberg, et al. Gravitational waves in full, non-linear general relativity, 2022. URL <https://arxiv.org/abs/2201.11634>.