The gravity of particle physics:
dark matter, black holes, and axions

Zachary S. C. Picker

This thesis is presented as part of the requirements for the conferral of the degree:

Doctor of Philosophy (PhD)

Supervisors:
A/Prof Archil Kobakhidze
Prof Celine Boehm

The University of Sydney
School of Physics

2022

Declaration

I, *Zachary S. C. Picker*, declare that this thesis is submitted in partial fulfilment of the requirements for the conferral of the degree *Doctor of Philosophy (PhD)*, from the University of Sydney, is wholly my own work unless otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.

Zachary S. C. Picker

July 6, 2022

Abstract

The greatest successes of fundamental physics occur at the two extremes of scale. At the smallest length, nature is understood with the tools of quantum field theory and the framework of the Standard Model of particle physics. Meanwhile, on astronomical and cosmological scales, the dominant physics is gravity, described by the formalism of General Relativity. Despite the incredible triumphs of both these theories, there is still much to explore in the intersection of these two regimes of size. Perhaps chief among these mysteries is the nature of dark matter. In this thesis, I will explore the physics of two dark matter candidates—primordial black holes, and axions. A central theme emerged from my investigations, related to the fundamental tensions between gravity and particle physics, between the largest and smallest regimes of physics. Namely, that the isolated treatment of each of these candidates misses novel phenomenology that only comes to light when *both* gravity and particle physics are accounted for.

Primordial black holes form in the early universe and so could comprise a fraction (or all) of the cold dark matter. Indeed, black holes are probably the only dark matter candidate that we have certainly observed. However, the Schwarzschild metric describes black holes in a flat, empty background. In the very early universe, a phenomenologically important period for primordial black holes, this is certainly not applicable. The black holes would be surrounded by the hot and dense cosmic fluid, and so we require a black hole solution which is properly cosmologically embedded. In this thesis, I explore the physics of these cosmological black holes generally, before choosing a specific metric—the Thakurta metric—for particular study. This metric was the only one we found in the literature which was pathology-free and valid in radiation dominated eras, although it is not without its criticisms, which I discuss also at length. The Thakurta metric has a time-dependent Misner-Sharp mass roughly proportional to the cosmological scale factor. We found that this greatly affects the landscape of black hole dark matter constraints. For one, binary formation in the early universe is significantly suppressed, removing entirely

the gravitational wave bounds on primordial black hole abundance. Secondly, these black holes evaporate significantly quicker, greatly increasing the mass of the smallest black hole which can survive to today. As a result, we close entirely the previously-unconstrained asteroid mass range of primordial black hole dark matter.

In the second half of my thesis, I examine a different dark matter candidate—the axion. This hypothetical light scalar particle was proposed to solve the strong-CP problem, a theoretical issue with the strong force related to its nontrivial background structure. However, it was shown recently that the inclusion of coloured gravitational instantons spoils the axion solution to the strong-CP problem. The most natural solution to this new *strong-gravity-CP problem* is the introduction of a second, coupled axion, which we called the 'companion' axion. The companion axion solution has qualitatively different phenomenology when compared to the single axion scenario, since one of the coupled axions can be much lighter than the other. First, we recomputed the axion-photon constraints in the new context of the companion axion. Then, we recomputed a number of cosmological considerations related to axions—perhaps most importantly, the misalignment mechanism for dark matter production. Notably, the 'favored' dark matter regime (where haloscopes would preferentially guide their searches), can occur at much lighter masses than are currently being probed. In addition, we found that the domain-wall problem which plagues the standard axion scenario is automatically avoided.

Publication list and attribution statement

The following papers form the basis for this thesis:

[1] A. Kobakhidze and Z. S. C. Picker, *Apparent horizons of the Thakurta spacetime and the description of cosmological black holes*, *Eur. Phys. J. C* **82** (2022) 347 [2112.13921]

[2] Z. Chen, A. Kobakhidze, C. A. J. O'Hare, Z. S. C. Picker and G. Pierobon, *Phenomenology of the companion-axion model: photon couplings*, 2109.12920

[3] Z. Chen, A. Kobakhidze, C. A. J. O'Hare, Z. S. C. Picker and G. Pierobon, *Cosmology of the companion-axion model: dark matter, gravitational waves, and primordial black holes*, 2110.11014

[4] C. Boehm, A. Kobakhidze, C. A. J. O'Hare, Z. S. C. Picker and M. Sakellariadou, *Comment on: Cosmological black holes are not described by the Thakurta metric*, 2105.14908

[5] Z. S. C. Picker, *Navigating the asteroid field: New evaporation constraints for primordial black holes as dark matter*, 2103.02815

[6] C. Boehm, A. Kobakhidze, C. A. J. O'hare, Z. S. C. Picker and M. Sakellariadou, *Eliminating the LIGO bounds on primordial black hole dark matter*, *JCAP* **03** (2021) 078 [2008.10743]

Authors lists are alphabetical in all cases. In the spirit of theoretical physics, all of these papers were as a result of a true collaboration between all authors, so it is not always easy to delineate which calculations I was specifically responsible for. I can say only that I was involved as much as possible in all aspects—across the brainstorming of ideas,

performing calculations, and writing and editing the final papers. I hope that this thesis might suffice as evidence of my engagement and contribution throughout these projects.

I have not reproduced the papers exactly in this thesis, but have rewritten and shuffled around the content in the pursuit of logic and flow. As a result, they are not exactly mapped to specific chapters. However, the contents of Refs. [1, 4–6] are primarily summarized in Chapters 2 and 3, while Refs. [2, 3] are almost entirely found in Chapter 5.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Zachary S. C. Picker, July 6, 2022

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Archil Kobakhidze, July 6, 2022

Acknowledgements

Just like dark matter, 80% of my gratitude must go to my supervisors. Firstly, to Archil Kobakhidze—it is hard to overstate Archil's influence on my growth as a physicist (for better or worse). For all the five-minute-questions which became two-hour-collaborations on his whiteboard, referee reports which are 'just sociology', surprisingly-intense soccer games, forgotten historical insights and genuine consideration of his students as real collaborators, I am truly thankful. If I ever rise to the ranks of distinguished dotty professor, it will be at least partially in his image.

Secondly, to Celine Bœhm—for putting up with my philosophical tangents and silly ideas, building from scratch an amazing astroparticle collaboration at Usyd, somehow finding time for all of her diverse students, and her sometimes-harsh but always-genuine care for everyone. I am looking forward to a beer after her head-of-school tenure when she can give me all the juicy gossip.

Then, there is the baryonic matter that is everyone else (for without them, structures would not form correctly). To the rest of my physics collaborators and friends: Cirian Oh"aire [sic], whose guidance and advice in our astroparticle group has absolutely made me a better researcher and communicator; Mairi Sakellariadou, for sticking with us on some crazy ideas; Giovanni Pierobon, Zhe Chen, and Yunho Kim, for sticking with us on some crazy ideas (a recurring theme); Markus Mosbech, Joseph Allingham, Natasa Lazarevic, Daniel Collison and everyone else in the group at Sydney or UNSW, for all the beers and crazy ideas. And to the rest of my friends in physics, whose fields I do not understand—thank you for doing the boring stuff I do not care about (but which is probably important).

Finally, to my lay-friends, lay-family and interested laypeople everywhere: Nell, Elle, Donny, Jimmy, Harry, Eve, Joel, Gareth, Shirley, Miro, Prof. Picker, Ms. Cohen, Sophia, Eli, Nefi, the RCT community, and everyone else who was a part of my life in the last few

years. Thanks for everything, and I apologize for all the rants (not really). The next page is for you:

The layperson box: an introduction

Unfortunately for all you laypeople—you are not exactly the intended audience of a thesis. Still, I enjoy babbling about physics, and sometimes it is hard to explain why this thing that I complain about all the time is making me move overseas in a few months.

So if you do find yourself flipping through this thesis, I have placed these boxes around at the end of a number of sections. The intention is to explain, in plain English, more or less what the section is about, so you at least know what I've been doing the last three years. The full list of layperson boxes is:

Contents

Abbreviations

ΛCDM	Λ–cold dark matter
ABRACADABRA	A Broadband/Resonant Approach to Cosmic Axion Detection with an Amplifying B-field Ring Apparatus
ALP	Axion-like particle
ALPS	Any Light Particle Search
CASPEr	Cosmic Axion Spin Precession Experiment
CAST	CERN Axion Solar Telescope
CERN	Conseil Européen pour la Recherche Nucléaire
COBE	Cosmic Background Explorer
DESI	Dark Energy Spectroscopic Instrument
EROS	Expérience pour la Recherche d'Objets Sombres
HAYSTAC	The Haloscope at Yale Sensitive to Axion CDM
IAXO	International Axion Observatory
LIGO	Laser Interferometer Gravitational-Wave Observatory
LISA	Laser Interferometer Space Antenna
LUX	Large Underground Xenon Experiment
MACHO	Massive compact halo object

MADMAX	Magnetized Disc and Mirror Axion experiment
MOND	Modified Newtonian dynamics
NANOGrav	North American Nanohertz Observatory for Gravitational Waves
OGLE	Optical Gravitational Lensing Experiment
ORGAN	Oscillating Resonant Group Axion experiment
OSQAR	Optical Search for QED Vacuum Bifringence, Axions and Photon Regeneration
TeVeS	Tensor–vector–scalar gravity
WIMP	Weakly interacting massive particle
WMAP	Wilkinson Microwave Anisotropy Probe
ADM	Arnowitt–Deser–Misner
ADMX	Axion Dark Matter eXperiment
BBN	Big bang nucleosynthesis
BPST	Belavin–Polyakov–Schwarz–Tyupkin
BRST	Becchi–Rouet–Stora–Tyutin
BSM	Beyond the Standard Model
CEH	Colored Eguchi–Hanson
CKM	Cabibbo–Kobayashi–Maskawa
CMB	Cosmic microwave background
CP	Charge-parity
DFSZ	Dine–Fischler–Srednicki–Zhitnitsky
DMRadio	Dark Matter Radio experiment
EH	Eguchi–Hanson

EPTA	European Pulsar Timing Array
FLRW	Friedmann–Lemaître–Robertson–Walker
GWyymmdd	Gravitational wave event, at given date
HSC	Subaru Hyper Suprime-Cam
KSVZ	Kim–Shifman–Vainshtein–Zakharov
LSST	Legacy Survey of Space and Time
M87	Messier 87
MS	Misner–Sharp
PandaX	Particle and Astrophysical Xenon Detector
PBH	Primordial black hole
PG	Painlevé–Gullstrand
PhD	Doctor of Philosophy
PPTA	Parkes Pulsar Timing Array
PQ	Peccei–Quinn
PQWW	Peccei–Quinn–Weinberg–Wilczek
QCD	Quantum chromodynamics
QED	Quantum electrodynamics
QFT	Quantum field theory
RCT	RollerCoaster Tycoon
SC	Shadrach Cohen
SDSS	Sloan Digital Sky Survey
Sgr A*	Sagittarius A*
SKA	Square Kilometre Array

SO	Special orthogonal (group)
SU	Special unitary (group)
UV	Ultraviolet
WKB	Wentzel–Kramers–Brillouin

Acronyms written in SMALLCAPS are intended to be read phonetically, while acronyms in Full Caps (FC) are to be read out letter-by-letter.

1
Introduction

All the attraction, the tension—can't you see, baby, this is perfection?

—Shakira (on the Standard Model), *Hips Don't Lie*

Hello, and welcome to my thesis—please, take a seat. During a rainy April in 2022, I find myself wrapping up my PhD after three-or-so years of research, making for a cumulative total of seven-ish years of physics learning. This is an unfortunate state of affairs, since the rest of the world has regrettably been studying physics for many more than seven years. Actually, physicists have been hard at work for some millenia now, and as a result, there has been quite a lot to catch up on.

On the one hand, humans have spent significant time asking the fractal-like question of, 'and what is that made out of?' This innocent persistence leads one down the helter-skelter of earth to atoms to nuclei to particles to quantum field theory, and other related mischief.

Similarly, many bathtimes have been spent wondering in the opposite direction, 'and what is that inside of?' Once we leave the Earth, we must pass through the solar system, the galaxy, the local cluster, through the cosmos and beyond the particle horizon.

Still, even though I have generously given everyone else a few-thousand-year head start, they have not gotten to the bottom of everything yet. The elementary particles and forces are extremely well-described by the beautiful Standard Model of particle physics, while the largest scales of the universe fall into the domain of gravity via General Relativity. However, the Standard Model is imperfect, with both theoretical and experimental anomalies remaining. And General Relativity has its own issues—on top of its incompatibility with quantum mechanics, many exotic predictions, such as black holes, still have much to be grappled with.

I do not mean to imply that physicists have been slacking off on these topics until I arrived. Rather the opposite—there is so much written on these issues, both old and new,

that it might take a long and full career before a truly informed literature review could be written.

However, of chief interest to this thesis is the areas of intersection *between* the two regimes of the large and small, where many unusual mysteries remain. The universe seems to be filled with some unknown elementary materials, such as the *dark energy* which appears to injecting energy into the universe to drive its expansion, and the *dark matter* which makes up about 80% of the mass content of the universe. Increasingly, it does not seem like particle physics alone, or astronomy alone, will be able to address these questions—precisely, perhaps, because they exist at the intersection of the two paradigms at opposite length scales. In recent years, the field of *astroparticle* physics has emerged to occupy this niche. The ultimate intention of astroparticle physicists is to use the tools of particle physics to study the broader universe while simultaneously using astronomical probes to study particle physics.

I have been privileged throughout my PhD to research a relatively broad range of astroparticle topics, mainly centering around the nature of dark matter. Maybe the main philosophical direction guiding these projects, however, has been the necessity of integrating gravity and particle physics—and the consequences of taking these lessons seriously.

Roughly one half of my thesis regards the difficulty of properly embedding a black hole in the dense, hot fluid of the early universe. We found that the solutions to this problem are qualitatively very different to the usual black hole models, which treat the black holes as separate from their environment. As a result, the landscape of black hole dark matter constraints is drastically modified.

The second half focuses on a hypothetical particle, the axion, which both solves some theoretical issues in the Standard Model while being a handy dark matter candidate. When gravitational interactions are properly included, however, we found that the axion solution breaks down. Again, the solution to this problem—funnily enough, adding a second axion—has rather drastic phenomenological consequences, especially for the dark matter question.

Since my studies have ranged somewhat extensively across particle physics, gravity, and astrophysics, I have included a relatively lengthy introduction to these topics in this first chapter. Chapter 2 comprises a dedicated review of black holes, while Chapter 3 contains most of my original research on cosmological black holes. Chapter 4 is a moderately pedagogical review of axions, and in Chapter 5 most of our novel work on companion

axions can be found. Finally, I conclude with some rather chatty remarks in Chapter 6.

Writing philosophy

Before I begin, I would like to briefly describe my writing philosophy throughout this thesis, and particularly regarding literature review. For one, this thesis is not a pedagogical textbook, and the intended audience is of relatively informed physicists. As a result, I will not be giving comprehensive and introductory explanations—there are uncountable excellent sources already existing for topics such as the mathematics of general relativity, or the spectrum of dark matter candidates. A small subset of this uncountable set might include Refs. [7–15], which were particularly impactful on the writing and structure of this thesis[1].

Instead, I have tried to emphasize and elaborate on topics and perspectives that either do not get nearly as much pedagogical attention, or that I found myself particularly excited to synthesize my thoughts on. For example, I spent multiple pages on the physics between Special and General Relativity, and significantly less time on the mathematical formalism of General Relativity itself. And I spent significant time discussing the strong-CP problem on a more introductory level, since I found that I often struggled with so-called 'introductory' texts on axion physics—whereas, I summarize much more briefly the status of axion constraints. In all, I just wanted to write a thesis whose literature review was not merely a retreading of familiar territory, and more accurately summarized the questions and confusions I struggled with throughout my studies. I hope that you find these elaborations and tangents as interesting as I did writing them.

Where possible (at the very least, in this introduction), I have tried to cite the original texts of a concept. This posed a particular challenge for older works or broader ideas (e.g. the ΛCDM framework), so in some places modern reviews have had to suffice. I have also attempted to minimize the use of acronyms where possible, in attempt to curb the ever-increasing mess of jargon which appears to be yelled at you in academic writing these days.

Generally, spacetime indices are labelled with Greek letters, and spatial indices with Latin letters, but there are some noted exceptions (e.g., sometimes Latin indices run over the gluon fields). I mostly use natural units where $\hbar = c = 1$, and occasionally geometric

[1]Perhaps even more impactful on the structure itself was Ref. [16], a thesis template which I forked before modifying substantially.

units where $G = 1$, unless I am including units for particular emphasis.

Finally, a quick note on narrative voice—I swap back between the personal 'I' and the scientific 'we' throughout the thesis. In general, I have tried to use my personal voice when editorializing or making a more opinionated point, and 'we' in more impersonal contexts, such as a derivation or presenting research findings.

1.1 Gravity and general relativity

Of all the fundamental forces, gravity has arguably been around the longest in human thought and learning. Aristotelian ideas of earth which longs to return to its 'natural' place held for almost two millenia [17–20]. The revelation of heliocentrism by Copernicus and others [21–24] eventually paved the way for Newton's theories of gravity and force [25], often considered the birthplace of modern physics (at least, in the west—the pioneering work of astronomers and mathematicians throughout the Middle East, India, China and the rest of the world cannot be understated [26–34]). Newton's gravitational theory, which lasted for centuries as the preeminent description of gravity, posited an attractive force between objects with mass given by,

$$F = G\frac{m_1 m_2}{r} \,, \tag{1.1}$$

in terms of the masses m_i of the two objects, their separation r, and Newton's gravitational constant G. Newton's theory can also be written in field theoretic fashion, using the gravitational potential ϕ:

$$\Delta\phi = 4\pi\rho \,, \tag{1.2}$$

where the spatially-defined scalar field ρ is the density of matter. Then the dynamics of a test particle are given by,

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\nabla\phi \,. \tag{1.3}$$

From this perspective, it will be slightly easier to see the conflicts between Newtonian gravity and Special Relativity.

1.1.1 Special Relativity

It was the meddling of physicists with another of the fundamental forces, electromagnetism, that eventually led to thin cracks in Newtonian gravity. Maxwell's equations [35–38] are inherently (and at least historically, accidentally) Lorentz invariant [39–42], leading to the prediction that the speed of light is constant in all reference frames. Although regarded as somewhat of a curiosity at the time, this was at odds with the Galilean principles of relativity that underpinned Newtonian physics [43]. Famously, Albert Einstein's resolution of

this discrepancy led to his theory of Special Relativity [44], which gave 'new' prescriptions for coordinate changes in order to respect this different kind of relativity, along with the famous revelation that energy, mass and momentum are related. Of particular note was that the coordinate changes when *boosting*—moving to a system with a different velocity at the same point in space—naturally rotated the time and spatial coordinates (specifically, with a hyperbolic rotation).

I use 'new' in quotation marks above because Lorentz transformations were already understood at the time, primarily by mathematicians studying geometry [39–42]. It was quickly realized that a geometrical perspective was indeed valuable for the interpretation of special relativity, especially when grappling with the somewhat confusing situation where time and spatial coordinates can be mixed. Vector-like quantities, such as position and momentum, are more appropriately written as four-vectors—position is promoted to a spacetime four-vector, and momentum to the energy-momentum four-vector. These four-vectors are defined on a four-dimensional background geometry, known as Minkowski space [42, 45], and points on the manifold are spacetime *events*.

This spacetime background is understood mathematically as a pseudo-Riemannian manifold, meaning that it comes equipped with a natural inner product for vectors called the metric (it is 'pseudo-' because the manifold is not positive-definite—the temporal and spatial coordinates have a relative minus sign). Compared to the old Euclidean geometry of three-vectors with Galilean relativity, the spatial length of a position vector or the magnitude of a momentum vector are not conserved under Lorentz transformations. Rather, the metric inner product of a *four-vector* with itself is conserved. For the energy-momentum four-vector, the invariant is actually mass, which is the origin for the popular energy-mass equivalence relation. For the spacetime four-vector, meanwhile, the 'spacetime interval' is conserved:

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \tag{1.4}$$

$$= -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \,, \tag{1.5}$$

where $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2$ in polar coordinates. The specific spacetime interval above, denoting 'flat' spacetime, is the Minkowski spacetime [45]. Actually, the spacetime interval is commonly used interchangeably in physics as a way to define the metric $g_{\mu\nu}$. Technically, the metric is an inner product, but this can be written succinctly as a second-rank tensor. The spacetime metric can then be used to define spacetime backgrounds more

complicated than the Minkowski space above:

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \ , \tag{1.6}$$

where the so-called *Einstein summation notation* implies that we sum over the four space-time coordinates for both contracted μ and ν indices, giving sixteen (possibly zero or identical) terms. For reference, the Minkowski metric specifically is commonly written as $\eta_{\mu\nu}$. This understanding of background spacetime, as we will see, is central to General Relativity, where the motion of particles under gravity is reunderstood as the motion of particles moving along geodesics of a curved spacetime background.

The layperson box: Special Relativity Previous — Next

You may have heard the rule, 'nothing can go faster than light.' This was a particular piece of wisdom that first appeared somewhat accidentally in the late nineteenth century, when physicists were studying the equations of electricity and magnetism. We will take here it as an axiom, an elementary law which should be assumed (see, maybe, the end of appendix A for some philosophical musings on physical assumptions).

One curious implication of this rule is that not even light can go faster than light. What I mean is, if someone were on a moving train and shot a beam of light alongside a stationary observer, the stationary observer would *not* see that the light is going at light-speed + train-speed. They would just see it at regular light-speed. No matter what *frame of reference* the observer is in, light always goes at light-speed.

This curious fact has rather severe consequences, which were spelled out carefully by Albert Einstein. How can the scenario above be consistent at all? To oversimplify, we should remember that velocity is given by 'distance divided by time'. In order for both of the observers to see the same light velocity then, one of the observers must be *measuring distance or time differently*.

Actually, both are happening. To the stationary observer, the moving observer appears to have both a slower clock and a shorter train compared to what the moving person observes. This is called 'Special Relativity', where time and length are mixed up with each other. In fact, Einstein realized, this also applied to energy, mass and momentum—they all depended on the frame of reference of the observer. For the stationary observer, even, this implied a particular equivalence had to hold: $E = mc^2$.

1.1.2 Between Special and General Relativity

Introductions often move quickly from Special to General Relativity. However, I find the decade-or-so period between these theories quite remarkable, and also instructive for understanding General Relativity [46]. There are a number of immediate problems for the reconciliation of Newtonian gravity with Special Relativity. If energy and mass are equivalent, how does energy gravitate? What reference frame should be used when considering distances between masses? And, Newtonian gravity has instantaneous action-at-a-distance, whereas Special Relativity has causal structure given by the finite speed of light.

While it was obvious to physicists in the early twentieth century that a relativistic theory of gravitation was necessary, it was not at all obvious how it should look. It is perhaps comforting (and relatable) that physicists navigating the fog of relativistic gravity spent much of the decade at times arguing vehemently and other times working together collegially [46].

Sometimes, the discrepancies between Special Relativity and Newtonian gravity are used to justify the move to General Relativity. However, at least as best as I can tell, General Relativity does not follow uniquely as the solution to these problems. Indeed, physicists spent the better part of a decade developing a wide range of fascinating and instructional gravity theories. As we will see, it actually requires the insistence of Einstein and others on additional assumptions, related to the equivalence of inertial and gravitational mass, and Einstein's freefall equivalence principle, to intuit our way to General Relativity. Based on Ref. [46], I will give an account of this intuition, partly following history and partly following my own understanding of the logical process which would in principle have guided physicists at the time.

Einstein's dismissed his initial forays into Lorentz-invariant gravity very quickly. Today, we might call these first theories 'scalar' gravity. At the heart of his issues was the recurring result that the internal energy of systems affected their vertical acceleration under gravity. This violated an observation close to Einstein's heart—his equivalence principle of acceleration and gravity, wherein an observer cannot distinguish between a gravitational field and an accelerating reference frame. Or, in other language, a freefalling observer feels at rest. Einstein appears to have seen this equivalence as a fundamental flaw of Special Relativity itself, where inertial reference frames are naturally 'preferred' [47, 48]. He insisted on a system of mechanics with *no* preferred reference frames. Notably, this equiv-

alence principle implies that light must be bent by gravity—an important observational test that would eventually validate Einstein's gravity theory.

These objections are reasonable to a modern observer, but we should remember that at the time, there was less solid experimental evidence to guide these ideas. The equivalence principle prediction of curved light from gravity was not yet observed, and the Eötvös experiments which measured precisely the difference between inertial and gravitational mass were only just being presented publicly [49]. As a result, relativistic gravity theories were developed which attempted simpler modifications of Newtonian gravity.

One idea I find particularly fascinating regards the similarity between Coulomb's law and Newton's law [50]. Actually, neither are Lorentz invariant, but the former is saved by the Lorentz invariance of Maxwell's full equations, via Ampere's law and others. It was hypothesized that a similar 'gravitomagnetic' field could exist, being sufficiently weak to avoid detection but rescuing gravity from the Lorentz transformations. However, this idea is ruined by the nature of the opposite sign of the gravitational force compared to Coulomb's Law. It was pointed out by Max Abraham [51, 52] (a recurring rival/colleague of Einstein's) that this sign would imply that a moving mass would disastrously absorb the analogous gravity-gravitomagnetic waves, rather than the situation where a moving charge *emits* electromagnetic waves, losing energy instead of gaining it.

Meanwhile, a simple and natural modification to Newtonian gravity was developed by Gunnar Nordström [53] (in close correspondence with Einstein, who had tried similar ideas years earlier), who proposed that the Laplacian of Eq. 1.3 should be replaced with the Special Relativity-flavored d'Alembertian $\partial_\mu \partial^\mu$. Still, it was not obvious how to modify the actual force law of Eq. 1.3, and Einstein maintained his objection that the internal energy (from rotations, stresses, etc.) was altering the acceleration. In addition, Einstein had crafted a thought experiment involving radiation in a box under gravity, in order to show that shear stresses must also be accounted for, or else one could gain unlimited energy by moving the box up and down [54].

The stress-energy tensor

In retrospect, we can see that these pains arise from the transition from the simple mass density of Newtonian gravity to the stress-energy tensor we are more familiar with, developed also concurrently for spacetime by Max von Laue and Hermann Minkowski, amongst others [45, 46, 55, 56]. Because all forms of energy need to gravitate, any attempt to simply

modify Newton's laws without accounting for all forms of energy were doomed to fail.

The stress-energy tensor $T^{\mu\nu}$ is a second rank tensor with the following interpretation: the element $T^{\mu\nu}$ gives the flux of the μ-th component of the energy-momentum four-vector over the surface of constant x^ν. In simpler English, T^{00} gives the energy density, T^{0i} gives the momentum density (with $i \in \{1, 2, 3\}$), T^{i0} gives the energy flux, and T^{ii} gives the momentum flux. The diagonal of this last block is indeed pressure, while the off-diagonal elements represent shear stress. Realizing this, Nordström modified his scalar equation [57, 58] so that the mass density could be replaced by the only reasonable choice of the stress-energy tensor, the trace T.

Gravity and geometry

Meanwhile, Einstein (with mathematician Marcel Grossman) had become increasingly interested in geometrical interpretations of spacetime, where objects move along geodesics in a curved background described by some metric tensor $g_{\mu\nu}$. With Adrian Fokker, Einstein showed that Nordstrm̈s new gravity theory was equivalent to motion along a curved spacetime with metric tensor $g_{\mu\nu} = \phi^2 \eta_{\mu\nu}$ [59]. Although we now know that such a conformally flat metric has no light bending, this was not a theoretical problem before such observations had been made. As a result, Nordström's gravity actually greatly resembled General Relativity, in that it coupled explicitly geometry to an energy source, via,

$$R = 24\pi T \,, \tag{1.7}$$

where R is the Ricci scalar. Interestingly, Nordström was more interested in embedding his scalar theory into a five dimensional spacetime which would unify gravity and electromagnetism [60–62]. This often-forgotten idea would be revived in the famous Kaluza-Klein theory [63–65], an important step in the long theoretical road of unifying the fundamental forces.

Despite the surprising success of this scalar gravitational theory, at the end of the day, Einstein's insistence on *general* covariance (inspired by his 'happiest thought' of the equivalence principle), would turn out to be empirically correct—gravitational bending of light was observed [66–68], as well as an accurate prediction for Mercury's perihelion shift [69, 70]. After the fog of covariant gravity theories settled, it was General Relativity that remained.

1.1.3 General Relativity

It is time to leave the history behind and look at General Relativity from the top down. General covariance makes it necessary to write the gravitational field equations with second rank tensors, coupling on one side geometrical objects such as the Ricci tensor $R_{\mu\nu}$, and on the other the *full* stress-energy tensor $T_{\mu\nu}$. Actually, it was almost entirely from intuition alone that Einstein suggested the correct form of the gravitational field equations:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = T_{\mu\nu} \ . \tag{1.8}$$

Freefalling test particles move according to the geodesic equation,

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2} + \Gamma^\mu_{\alpha\beta} \frac{\mathrm{d}x^\alpha}{\mathrm{d}s} \frac{\mathrm{d}x^\beta}{\mathrm{d}s} = 0 \ , \tag{1.9}$$

in terms of the Christoffel symbols Γ (otherwise known as the Levi-Civita connection), which are arrays of numbers used to compare objects defined at separate spacetime points. Because the curvature of spacetime itself is nontrivial, comparing for example two vectors is also nontrivial—the partial derivative ∂_μ is not useful in that it cannot distinguish between internal changes in the vector field and changes due to the curvature of spacetime. Instead, the covariant derivative is required:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\rho} V^\rho = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} V^\nu \right) \ , \tag{1.10}$$

for arbitrary vector field V, and where g is here the determinant of the metric tensor.

There is, however, a more fundamental approach to the Einstein equations. It was shown by David Hilbert [71] (rapidly, or even, in advance of Einstein) that the field equations could be derived via the Euler-Lagrange equations from a simple action:

$$S_{\mathrm{EH}} \equiv \frac{1}{8\pi G c^{-4}} \int R \sqrt{-g} \, d^4 x \ , \tag{1.11}$$

where the determinant of the metric tensor is required to keep the action itself generally covariant. Fundamental physicists love a good Lagrangian—as we will later see, the rest of the fundamental forces, as part of the Standard Model of particle physics, are also defined in this way.

This formulation leads us to two important points. First, the action of General Relativity can be seen as the *simplest* possible theory consisting of the metric tensor and its second derivatives (first derivatives can be made to vanish in suitable coordinate systems). This is because it can be shown that any tensor comprised of first and second derivatives of the metric can be written in terms of the Riemann tensor and the metric tensor itself—the Ricci scalar being the only independent scalar that can be made from the Riemann tensor. This realization gives special importance to General Relativity, along with giving an easy avenue to extend and modify the theory—we can add extra curvature invariants, or even particle fields, to the action. Extending or modifying General Relativity is a common practice, for example when attempting to explain phenomena like dark matter and dark energy [72–77], or exploring the interaction of General Relativity with quantum systems [78, 79].

The second advantage of this formulation is that we can see that the Einstein field equations can also be given an integration constant Λ, so that they become,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = T_{\mu\nu} \ . \tag{1.12}$$

It is popular folklore that Einstein introduced this factor originally and erroneously to preserve the steady-state universe from gravitational collapse. However, observations of the expansion of the universe today indeed support the existence of such a 'cosmological constant', as we will see in Sec. 1.1.4.

Tests of General Relativity

General Relativity is incredibly well-tested across a wide range of scales, leading to a large number of novel predictions. Famously, the precession in the perihelion of Mercury [69, 70] and the bending of light near the sun [66–68] conspired to send Einstein's reputation from a big name within the scientific community, to international and lasting fame with the general public [80, 81]. Torsion-balance laboratory tests have probed General Relativity even on small scales [49, 82, 83]. Gravitational fields also warp time, which must be accounted for by GPS satellites [84]. Exotic objects like black holes solve Einstein's equations, which I will discuss in Ch. 2. Quadrupole moments of masses source gravitational waves [42, 85, 86], ripples in spacetime analogous to electromagnetic radiation—incredibly, gravitational waves from merging black hole binary systems have been observed by LIGO/Virgo [87–90].

There are also endless astrophysical and cosmological observations, like strong lensing in galaxy clusters [91] or solar system tests [92] such as lunar laser ranging [93, 94]. And of course, there is cosmology—the study of the large-scale universe with General Relativity is one of its most powerful applications. As such, it deserves its own subsection.

The layperson box: General Relativity Previous — Next

If you have ever wondered why Albert Einstein is so famous when there have been other highly talented scientists throughout history, General Relativity is the reason why.

Isaac Newton had two good points about physics and gravity. Firstly, objects that are moving should keep moving in straight lines, unless something forces them to accelerate off their straight line. Secondly, objects with mass attract each other.

General Relativity could be considered a way of combining those two points into one single idea. Although it seemed to model nature somewhat well, Einstein did not like the second point. Firstly, nothing can go faster than light—but according to Newton's theory, nudging a mass would instantaneously move a distant mass, which would let you communicate instantly across arbitrary distances. Secondly, Einstein had previously realized that energy and mass were 'mixed up', so gravity had to account for the gravitation of energy itself, and not just mass.

Einstein's solution is quite ingenious. First, he dispelled the Newtonian notion that objects attracted each other. Then all we are left with is the first idea, that objects move in straight lines. But—what is a straight line? On Earth, straight lines actually are most like equatorial circles, and not straight at all. Really, what a 'straight' line is, depends on the shape of the surface that you are moving along. This is the heart of General Relativity—all objects are moving on a big background, which is warped and bent by the energy and mass of objects on it. That means that the 'straight lines' they follow actually lead objects naturally towards each other, as they follow the curving of the background down into the dips that each object makes.

Part of what made Einstein so famous was the testing of his idea. Light too, follows 'straight' lines—meaning that a massive object, like the sun, would appear to bend the light rays. Not long after Einstein suggested this theory, astronomers realized that there was going to soon be an opportunity to see this. In both Brazil and São Tomé and Príncipe, a solar eclipse in 1919 would allow astronomers to observe stars near the sun. Einstein's prediction was widely publicized in newspapers, so when the moment finally came and the bending was observed, Einstein was propelled into lasting fame [80, 81].

1.1.4 Cosmology

One of the most important philosophical adjustments in science concerns the place of Earth in the cosmos. With the advent of the Copernican revolution [22], the tenets of Aristotelian and classical geocentrism were abandoned in favor of a heliocentric solar system. In fact, this principle can be extended even more generally, into what is known now as the Copernican principle [95]. Not only is the Earth not the center of the solar system, but the sun is not in the center of the galaxy, the galaxy is not the center of the universe, and none of these particular structures we inhabit is more special than any other one. Although there is a philosophical aura to this principle, we should not forget that it is in fact an observation. The more we examine the distant (and wider) universe, the more we are inclined to assume this cosmological principle: on the largest scales, the universe is *isotropic* and *homogeneous*.

It turns out that Einstein's equations have actually a rather simple form under these two extremely constraining assumptions. Actually, we are left with only one free scalar function after applying this principle. The metric is known as the Friedmann–Lemaître–Robertson–Walker (FLRW) spacetime [96–103]:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right) , \tag{1.13}$$

where the time-dependent function $a(t)$ is known as the scale factor and $k \in \{-1, 0, +1\}$ is a constant curvature parameter which determines whether the spacetime topology is hyperbolic, flat or spherical, respectively. Throughout, I will set $a \equiv 1$ today. On the largest scales, then, the universe can simply be described by its global topological structure, and a scaling function which determines at what rate it is growing or shrinking. The source of this metric is a perfect fluid,

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu - P g_{\mu\nu} , \tag{1.14}$$

in terms of the energy density ρ, pressure P, and fluid four-velocity u. Of course, we actually do know something of the energy-matter content of the universe, and we can parametrize this content by an equation of state $P = \omega\rho$. It is common to split the energy density into its components, each with their own value of ω:

$$\rho = \rho_{\text{radiation}} + \rho_{\text{matter}} + \rho_{\text{cosm. constant}} . \tag{1.15}$$

Radiation (light and matter at relativistic speeds) has an equation of state with $\omega = 1/3$, matter with $\omega = 0$, and the cosmological constant with $\omega = -1$. It is not hard to derive the time dependence of these components, but it is easily understood from physical principles. Matter is diluted proportional to the change in volume, whereas radiation is additionally redshifted as the universe grows, and the cosmological constant is, unsurprisingly, constant. So the time dependent energy density can then be written as,

$$\rho(t) = \rho_{\text{radiation}} a^{-4}(t) + \rho_{\text{matter}} a^{-3}(t) + \rho_{\text{cosm. constant}} \ . \tag{1.16}$$

Because redshift is an important physical observable in astronomy, it is also conventional to use redshift z as both a time and distance measurement instead of the scale factor:

$$1 + z = \frac{a(t_0)}{a(t)} \tag{1.17}$$

The Einstein equations can be used to derive what is known as the Friedmann equations [96, 97]. The first of these is written using the important Hubble parameter, $H \equiv \dot{a}/a$:

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \ . \tag{1.18}$$

For convenience, we define a critical density $\rho_{\text{c},0} \equiv 3H_0^2/8\pi G$, which would be the density of a flat $k = 0$ universe today. Then we can define the dimensionless densities $\Omega_i = \rho_i/\rho_{\text{c}}$, where the subscript 0 is traditionally dropped since it is understood they are defined today. Then we can rewrite the Friedmann equation in the very useful form,

$$H^2(a) = H_0^2 \left(\Omega_{\text{r}} a^{-4} + \Omega_{\text{m}} a^{-3} + \Omega_{\text{k}} a^{-2} + \Omega_\Lambda \right) \ , \tag{1.19}$$

in terms of the *Hubble constant*, $H_0 \equiv H(a = 1)$ [104] and the somewhat-contrived curvature parameter $\Omega_{\text{k}} \equiv -k/(a_0 H_0)^2$ (and I have shorted the energy density components to ‘r’, ‘m’, and ‘Λ’).

The history of the universe

This form of the Friedmann equation makes it very easy to quickly see the evolution of the universe, since there are distinct eras where each of these components dominates—at smallest a, radiation dominates, then matter, etc. It was observed early in the twentieth century that distant galaxies are moving away from us, implying that the universe is

expanding [96, 98, 104, 105]. This means that a smaller scale factor does indeed correlate to earlier times (although the scale factor could shrink again, if the fate of the universe was to collapse). In this picture of the universe then, the universe began at a single moment in time with a Big Bang [106–108], expanding (and decelerating) through radiation- and matter-domination, before accelerating again during Λ-domination (as long as $\Lambda \neq 0$).

The detailed story of the early universe is even more interesting, but I will give just a qualitative picture here. As we go back in time, the universe becomes denser and hotter, and so we require high energy particle physics (see the upcoming Sec. 1.2) to describe the state of this thermal bath. Of course, our knowledge of particle physics is limited to theory and laboratory experiments, so the hottest periods of the universe are actually a probe of unknown physics, which is simultaneously a powerful use of cosmology and a source of uncertainty.

Starting from the earliest times, as the universe cools, a number of phase transitions occur (not dissimilar to the cooling of a magnet). First, we might pass through the hypothetical Grand Unification scale [109, 110], where the strong and electroweak forces would cease to be unified. At yet cooler temperatures, the electroweak [111, 112] phase transition occurs, separating electromagnetism from the weak force (which is spontaneously broken). Cooler yet, hadrons are able to form—composite particles such as protons and neutrons—during the quantum chromodynamics (QCD) [113–117] phase transition.

Around one second after the big bang, nucleosynthesis (BBN) would begin [107, 118–120]. The proportions of light nuclei such as Helium, Deuterium and Lithium which are produced here are a very powerful cosmological observable. At a redshift of $z \sim 5000$, matter finally begins to dominate. Eventually, the universe is sufficiently cool that electrons couple to ions to form atoms—at this point, photons can suddenly freely stream and the universe becomes transparent. This forms a surface of last scattering known as the Cosmic Microwave Background (CMB) [107, 108, 121–123], because the light from this surface is redshifted into the microwave band by the time we observe it.

The surface of the CMB is easily the most important physical observable in cosmology, and the amount of information prised out of the CMB by crafty physicists is truly staggering. Perhaps chief amongst these observables is the power spectrum of the small observed temperature anisotropies [124–126]. Although we do assume that the universe is homogenous, small perturbations either in the thermal bath or along the photon's path after the CMB [127] contain a significant amount of physical information. Oscillations in the

photon-baryon fluid tell us about the pressure, gravity, and even the particle content and earlier history of the fluid [128]. As a result, we can learn a surprisingly great deal about the composition of the early universe just from the CMB temperature map [124–126].

After the CMB, the universe enters a dark age before the first stars are formed and while large scale structures are still forming. The formation of the first stars and galaxies, however, produces sufficient ionizing radiation to reionize neutral hydrogen, starting from around $z \sim 11$ and ending at around $z \sim 6$ [129, 130] when all of the hydrogen has been reionized. After this, the universe enters the long age of galaxy evolution that takes us to today. Interestingly, around only 4 billion years ago, observations show that the cosmological constant began dominating in our universe [126]. As a result, we are currently in a period of accelerating expansion—this last revelation is quite recent, having been observed only recently using distant Type 1A supernovae as standard candles [131, 132].

The nature of the cosmological constant is not well known, and is often referred to interchangeably as dark energy. More carefully, dark energy is a hypothetical component of the energy content of the universe with equation of state specified by $\omega < -1/3$. The cosmological constant corresponds to $\omega = -1$, and observations seem to indicate that the actual value is indeed around -1 [133]. Because we do not fully understand dark energy, it is hard to make predictions for the future of the universe. Maybe it will approach absolute zero temperature, forever. Possibly, dark energy will eventually be sufficiently strong that all structures (even atoms) will be ripped apart. Perhaps the universe will eventually collapse, and maybe even bounce again after. I do not believe I will be around to find out.

The horizon problem and inflation

However, there is still one interesting paradox remaining, whose resolution is required to complete our picture of the evolution of the universe. Based on observations of the parameters in Eq. 1.19, it is possible to estimate the age of the universe to be roughly 14 billion years old. However, consider two regions at the antipodes of our night sky—the distance between them is necessarily more than 14 billion light years, so they could never have been in causal contact. In fact, when we observe the CMB, it is possible to show that patches of only ~ 1 degree on the sky were ever actually in causal contact. And yet, the universe appears to be incredibly homogeneous in all directions—not only do galaxies and astrophysical structures appear the same, but the CMB temperature is incredibly uniform.

This is a paradox in its truest form. We began by assuming homogeneity and isotropy, and the logical conclusion of this argument was an FLRW universe which is not sufficiently old that it should be homogeneous and isotropic in the first place—our conclusions contradict our assumptions. This is known as the horizon problem [134].

The most famous of the solutions to the horizon problem is the introduction of a period of exponential inflation at the beginning of the universe [135–139]. In this paradigm, the universe began in causal contact, before a period of rapid expansion occurred such that two nearby points would be expanded beyond each other's cosmological horizons. As a result, two points which might not appear today to have ever been in causal contact could have been in contact at the beginning of inflation. Inflation also handily solves two other problems—the flatness problem (we observe an incredibly finely-tuned flat universe), and the problem of the rarity of exotic objects such as magnetic monopoles which might be expected to be formed in the very early universe.

We require new physics to include inflation into our cosmological understanding. There should be some new particle field which dominates the energy density before radiation, commonly named the *inflaton*, in a manner similar perhaps to dark energy. Inflation must also end in a period of reheating, since inflation cools the universe down significantly. Then the inflaton must be coupled to Standard Model particles in such a way that the thermal bath of radiation domination is formed at the end of inflation [140]. Finally, inflation serves as a mechanism to seed inhomogeneities in the early universe, since quantum fluctuations are blown up, leaving a spectrum of large-scale fluctuations [141–147].

Cosmological observations

I will conclude this whirlwind, qualitative tour of cosmology with a brief note on the observations themselves. There are a large range of astronomical observations which give cosmological information, from dedicated microwave surveys like COBE [125], WMAP [148] and Planck [126], to large scale structure surveys such as by SDSS [149], DESI [150], 2dF [151] and the future Euclid [152] and LSST [153], to more local measurements such as supernovae [131, 132, 154] or elemental abundance [107, 155, 156]. That is only a small sample of cosmological observations—the last couple decades has already been dubbed a 'golden age' for cosmology, in both the precision and breadth of measurement.

It is important to note, however, that we often do not *directly* measure an observable like Ω_{m} or even H_0. Rather, we measure slightly more esoteric parameters, such as

the angular scale of CMB anisotropies or metalicity of stars. Then, the 'fundamental' cosmological parameters, like the energy densities, must be *derived* following specific models—the most popular of which is the ΛCDM model [157] which assumes a cosmological constant with a dark matter population, although there is still plenty of wiggle-room for alternative cosmological models. This is on top of whatever additional modelling may be necessary to connect the observation to cosmology, such as galaxy formation or supernovae physics. Of course, this situation is not unusual in astronomy and our knowledge of the cosmological parameters is increasingly precise. Still, it is important to remember, from a theory-oriented mindset, that there is a nontrivial amount of model-dependency in the cosmological parameters.

The layperson box: cosmology Previous — Next

What happened at the beginning of the universe? And what's going to happen at the end? These are fun questions that physicists get asked often, and the answer to them is called 'cosmology'. Actually, cosmology might be more carefully described as the study of the largest scales. When you zoom out enough, galaxies just look like specks of dust, and suddenly it is much easier to simulate the universe—all you have to do is add gravity to this pile of sand (and radiation, like light) and see what happens.

A century ago, astronomers like Edwin Hubble (as in, the space telescope) realized that galaxies were moving away from us. This makes the zoomed-out sand picture somewhat simple. The universe today is just a big, expanding clump of sand. Then it is easy to wind the clock backwards (theoretically) step by step and estimate what it used to look like.

Earlier in time, the dust and radiation and everything had to be closer together and so under higher pressure—and therefore, higher temperatures. Eventually, it would have been so hot that not even atom could stay together anymore, and the universe would stop being a kind of lumpy object-filled-space, and become more like a hot soup of particles whizzing around each other. If you look through a telescope 'between' the galaxies (really, it must be a telescope which sees microwaves instead of optical light), you can actually see exactly the point when atoms first formed, thus ending the opaque-soup-times. Everything past this point is totally obscured from telescopes, forming a kind of sphere all around us. The inside surface of this sphere that we observe is known as the cosmic microwave background.

This surface is very handy because, 1; the maths of the soup is way simpler, and 2; all particle physicists have ever wanted is high-energy environments like this to create exotic particles (see: page 32). That means that if we can observe the earliest universe with telescopes, we could potentially learn about particle physics for free. And, because experiments on Earth have helped us learn about particle physics up to really quite high energies, we can actually predict what the hot soup of the early universe was doing, even in extremely early times.

If you keep turning the handle of the clock enough, eventually you enter an era where we don't know the particle physics anymore. If Einstein's General Relativity is to be trusted, it seems that eventually the universe shrinks back to one tiny point that encompasses all of space simultaneously. The explosion from this point is, of course, known as the Big Bang. I think my favorite thing about all of this, however, is how recent these revelations are. The Big Bang is now entrenched in popular culture, but it wasn't until the 1960s that it was even the preferred theory among physicists. The amount we know about the universe compared to fifty or even twenty years ago is astounding—for example, it was only at the beginning of the millennium that distant supernovae (exploding stars) were carefully measured, showing that the universe was actually *accelerating* in its expansion—a mystery whose answer is still very much unknown.

1.2 Particle physics and the Standard Model

While gravity might be considered the physics of the largest scales, the second of the two fundamental regimes of physics concerns the very smallest scales—*particle physics*. Or, if you would believe the sign above my office door, *high energy physics*. These two names are easily related by the famous Heisenberg uncertainty relation [158] for position and momentum, $\Delta x \Delta p \geq \hbar/2$, where $\hbar$ is the reduced Planck constant [159–162]. Indeed, much of the history of particle physics is a story of colliding particles together at increasingly higher speeds so that we can probe the so-called *ultraviolet* (higher energy) regions of particle physics.

The chief theory of particle physics is known as the Standard Model of particle physics, or Standard Model for short [163, 164]. This theory describes the properties and interactions of the known fundamental particles (quarks, leptons, gauge bosons and the Higgs boson), and the various ways in which they combine to form the 'particle zoo' of composite particles (hadrons, mesons, etc.). The basic aim of this theory is to make quantitative predictions for questions such as, 'If I collide two "X" particles together, what are the chances I get two "Y" particles out again? What are the energies of those "Y" particles?' The Standard Model, which describes the elementary particles and their interactions, is one of the crowning achievements of modern science. Not only is it aesthetically beautiful, but it is has incredible predictive power.

Of course, the Standard Model is not a complete theory of nature—if it was, I would be out of a job. It does not include dark energy, gravity, or dark matter (see: Sec. 1.3). In its minimal formulation, the Standard Model has the rather-large-quantity of nineteen free parameters, a curious problem of hierarchies and naturalness [109, 165–168], and no explanation for the 'generations' of particles [169, 170]. Then there is the ongoing puzzle of nonzero neutrino masses [171–174], the strong-CP problem [175–181], the primordial matter-antimatter symmetry [182–184], and the flavor anomalies [185, 186] which are slowly gaining more statistical significance. We do not yet know how or if the strong force [113–117] should be unified with the electroweak force [111, 112] in a grand unified theory [109, 110], or if the spacetime symmetries should be extended with some kind of supersymmetry [187–189], or something yet more exotic like string theory [190–197].

Despite all of its flaws, however, the Standard Model remains one of the jewels of human achievement. It is elegantly simple but simultaneously deep and complex, and almost no other scientific model comes close to it for pure predictive power. While I gave

a somewhat gentle and historical account of General Relativity, in this section I will take a more top-down approach and introduce the tools of particle physics at a more technical level. Much of the great depth of the Standard Model is not required background for this thesis, but I think it may be instructive to briefly recall the mathematical formalism of this theory.

1.2.1 Quantum field theory

The Standard Model is a quantum field theory (QFT) [198–207]. The *field theory* component means that the Standard Model is specified with a Lagrangian density, with terms built out of mathematical fields of various kinds defined at spacetime points. Because we are dealing with small scales, it is generally appropriate to define these fields on a Minkowski spacetime, which is always tangential to the true spacetime background of General Relativity.

The *quantum* part of QFT, in the *canonical* quantization, means that these fields are promoted to operators, in a similar way to how observables like momentum and position are promoted to operators in classical quantum mechanics ('first quantization', whereas QFT is sometimes called 'second quantization'). The field operators of QFT act on the so-called Fock space [208], a Hilbert space composed of *particle number* states, each defined with respect to both spacetime location and energy. In this sense particle states are considered field excitations—the classical concept of field value corresponds to particle number.

Alternatively, the field theory can be quantized with the famous path integral formalism [204, 209, 210]. In this description, the action principle is extended so that the probability of some process includes contributions from *every* possible intermediate path. Incredibly, canonical quantization and the path integral formalism are formally equivalent. I am not sure there is a more mysterious (or more aesthetically pleasing) wonder in all of physics.

In both descriptions, one usually computes *correlators* (correlation functions)—vacuum expectation values of products of field operators. These are often expressible as perturbation series and expressed visually via Feynman diagrams [211]. These diagrams are a way of organizing the vertices and internal or external lines of an interaction which are mapped to expressions in the perturbation series via Feynman rules. These diagrams are so ubiquitous that they have become the *de facto* representation of particle

physics processes.

1.2.2 Symmetries of the Standard Model

Terms in the Standard Model Lagrangian density (or just 'Lagrangian') are generally composed of multiple fields, as well as dimensionless or dimensionful parameters. While the beauty of General Relativity is found in its geometric arguments and relativistic thought experiments, the beauty of the Standard Model is quite different—but no less astounding. Instead, the construction of the Standard Model Lagrangian follows from quite remarkable symmetry arguments.

Here, symmetries are expressed in the language of *transformations*. If the action is invariant under some transformation, it is called symmetric. It is useful to use mathematical groups to describe these symmetries, but determining how to apply the action of a group to a specific particle field in the Lagrangian is not always trivial. Generally, we must use a *representation* of the group—a group homomorphism from the symmetry group into (as far as particle physicists are concerned) a particular matrix space of specified dimension. For a particle field, we are able to choose this representation, including making it the trivial representation, but generally we are interested only in the irreducible representations, since we are attempting to describe elementary particles and forces. In this sense, it is common to label particle fields by the way we choose them to transform under some representation of a symmetry. I will elaborate with some examples to clarify this concept shortly.

Noether's Theorem

One important class of symmetries are continuous global transformations. The remarkable Nöther's theorem [212] tells us that there is a conserved quantity for every generator of these symmetry groups. Maybe the most famous global symmetries are the spacetime symmetries of time translation, spatial translation, and rotations. The first corresponds to conservation of energy, the second to conservation of momentum, and the third to conservation of angular momentum. The final Lorentz transformation, boosts, correspond to conservation of center-of-mass momentum.

Spacetime symmetries

The representation theory of the Lorentz group is fundamental to particle physics. Actually, there are four disconnected components of the Lorentz group $\mathrm{SO}(1,3)$—the discrete actions of time reversal T and parity reversal P move between them. The connected component preserving orientation and time direction is called the proper orthochronous Lorentz group $\mathrm{SO}^+(1,3)$. Usually every term in the Lagrangian is required to be Lorentz invariant, so that the action itself remains Lorentz invariant.

However, the individual particle fields have non-trivial Lorentz structure, and in fact each kind of field is labelled by how it transforms under a particular representation of the Lorentz group. These representations are labelled by a number referred to a the *intrinsic spin* of the field, so particle fields are named spin-0 (*scalar*), spin-1/2 (*spinor*), spin-1 (*vector*) and so on. For spinor fields, these representations also determine whether the field is *left*- or *right*-chiral, and we often combine the left- and right-chiral spinors into what is known as a *bispinor*.

Particles with spin in integers are known as bosons, and follow Bose-Einstein statistics—the creation and annihilation operators of the fields commute and multiple particles can occupy the same quantum state. Particles with spin in half-integers are known as fermions, and follow Fermi-Dirac statistics—their creation and annihilation operators anticommute and so the Pauli exclusion principle [213] means only one particle can occupy a particular quantum state.

In relatively non-rigorous language it is often said that the Lorentz group is equivalent to two copies of the group $\mathrm{SU}(2)$, which is responsible for the existence of the left- and right-chiral particle fields, and the labelling of 'spin' for the particle representations. More technically, when forming representations of the Lorentz group we use the Lie algebras. The complexification of the Lorentz algebra satisfies $\mathfrak{so}(1,3)_{\mathbb{C}} \simeq \mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C})$, where $\mathfrak{sl}_2$ is the special linear Lie algebra of order 2.

Gauge symmetries

The theory of gauge invariance [214–217] which defines the Standard Model might be the most beautiful aesthetic achievement in all of physics, as far as I am concerned. While the spacetime symmetries are used to define the fields themselves, the gauge symmetries define the interactions of the fields. The principle states that Standard Model fields are invariant under *local internal symmetries*. These symmetries are spacetime-dependent and

act on the fields only (in contrast with the Lorentz symmetries, which act both on the field and on the spacetime point the field is defined on). Gauge symmetries are philosophically quite different to the global symmetries which are used in Nöether's theorem—while the latter takes one physical state to a different physical state (and asserts there is conserved quantities across the two systems), two states related by a gauge symmetry are genuinely the *same* state. In other words, there is a redundancy in our description of the system. The gauge group of the standard model is,

$$\mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3) \ . \tag{1.20}$$

These groups roughly correspond to electromagnetism, the weak force, and the strong force respectively, although I will explain a small wrinkle in this description below. Every particle in the standard model is transformed under a representation of this group, although it is usually easier to look at the representations under each of the component groups separately (the particle field will therefore transform under the tensor product of each of the representations). Remarkably, almost the entire dynamics of the Standard Model is contained in this gauge group.

For example, let us introduce a bispinor field ψ, which might represent a particle such as an electron. Let's define this field so that it transforms under the fundamental representation of $\mathrm{U}(1)$, the smallest-dimensional faithful representation (meaning, the group structure is fully preserved). This transformation is given by,

$$\psi \to e^{-iq\alpha(x)}\psi \ , \tag{1.21}$$

where $\alpha(x)$ is an arbitrary continuous function. However, the electron Lagrangian on its own (the Dirac Lagrangian) is *not* invariant under this transformation, as can easily be seen by inspecting the derivative terms:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - mc^2\bar{\psi}\psi \ . \tag{1.22}$$

So is it not possible to have such a particle field? The answer is no—this problem can be saved by defining a vector field which transforms under the adjoint representation of $\mathrm{U}(1)$, the distinguished representation of a Lie group which is more-or-less just its own

Lie group:

$$A_\mu \to A_\mu + \partial_\mu \alpha(x) \ . \tag{1.23}$$

Incredibly, when one introduces a suitable photon-electron interaction term in the Lagrangian, the photon transformation terms can be used to exactly cancel out the non-invariant part of the electron transformation, so that the total $\mathrm{U}(1)$ invariance is preserved. Specifically, the remedy is to promote the partial derivative in the Dirac Lagrangian to a covariant derivative, defined like,

$$\partial_\mu \to D_\mu \equiv \partial_\mu - iqA_\mu \ . \tag{1.24}$$

The constant q then gains physical interpretation as the *charge* of the bispinor field under this symmetry. Incredibly, the predictions of the photon-electron system defined this way seem to describe nature correctly. This is the beginning of quantum electrodynamics (QED) [199, 202–207, 218, 219], one of the most precise physical theories in all of science. Measurements of the fine-structure constant (related to the photon-electron coupling constant) agree to within one part in a billion [220]. The moral of the story is—if we allow other particle fields to transform nontrivially under the $\mathrm{U}(1)$ symmetry in the Lagrangian, we are *forced* to implement interactions between these particles and photons.

The logic is similar for the $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$ gauge groups, although the maths is slightly more complicated—partly because these groups are not Abelian, and partly because the representations are more complicated. The fundamental representations have dimension 2 and 3 respectively, while the adjoint have dimensions 3 and 8. As a result, fermions which feel the weak force are found in doublets, while those that feel the strong force are in triplets (commonly labelled by their *color* of red, green, and blue). Meanwhile, there are three weak bosons and eight so-called gluons for the strong force.

I will also briefly comment on the use of the word 'covariant derivative', which is obviously striking in comparison to General Relativity. When considering geometry, the covariant derivative provides instructions for *parallel transporting* vector fields across an arbitrary manifold. The use of the word in the gauge context here is not accidental. Instead of measuring the the 'parallel-ness' of vectors, however, we are interested in what it means to compare the phases of fields at different spacetime locations. In this sense, the gauge covariant derivative provides instruction for transporting the phases of fields. Generally, gauge theory can be understood purely geometrically, as a study of connections on bundles

of various mathematical types—something of which I know relatively little, however.

The direction of my summary here is sometimes cause for confusion (including for myself). Often, gauge theory is introduced by 'promoting' a global symmetry to a local one, and exploring the consequences of this assumption. This is more-or-less the direction of the argument I made above. However, this leads sometimes to a confusion of the $\mathrm{U}(1)$ symmetry with the phase invariance of quantum mechanical states. These are *not* the same symmetries, and phase invariance is not the motivation for gauge theory. Rather, phase invariance is responsible for conservation of particle number. This shouldn't be unexpected, because the wavefunction corresponding to electromagnetically neutral states must also be invariant under an arbitrary phase change. Meanwhile, in our interacting QFT, we have transitions between particle states. As a result, electromagnetically charged particles can violate particle number conservation, so long as electric charge is conserved (for example, a muon decaying to an electron and neutrino-antineutrino pair).

Instead, as noted earlier, the motivation for gauge theories is primarily related to the fact that relativistic descriptions of vector fields (and other higher-spin fields) contain redundant degrees of freedom. This is most familiar from the gauge potential of Maxwell's equations. When quantizing the gauge bosons, this redundancy leads to ambiguity in the path integral, and the gauge must be fixed in such a way that physical observables aren't affected. However, for the non-Abelian symmetries, the path integral contains infinities which can't be cured by Abelian gauge-fixing procedures, forcing the addition of Fadeev-Popov ghosts [221]. Curiously, these quantized Yang-Mills theories actually possess a symmetry known as the Becchi–Rouet–Stora–Tyutin (BRST) symmetry [222, 223], which might be considered a kind of 'quantum version' of the gauge symmetry, rendering the Fadeev-Popov ghosts harmless.

I should probably note that my summary above follows what is known as the 'gauge argument'. Philosophers or philosophically-inclined physicists [224–226] occasionally concern themselves with ontological or epistemological aspects of this argument, since in some ways the construction is a bit *ad hoc*, even from purely geometric perspectives. Still, as a physicist (and only a somewhat-interested amateur philosopher), I should be content to push these issues aside, call this construction a 'mathematical formalism', and be content that it seems to get things correct in the real world—the philosophers can figure out why that is so. It may be that there is not (and will never be) an 'answer' to why the mathematical language of gauge symmetries and irreducible representations so adequately describes the interactions of particles.

The Weinberg-Salam-Glashow model

I mentioned a wrinkle in the association of the Standard Model gauge group to the fundamental forces. That is because the group $\mathrm{U}(1)\times\mathrm{SU}(2)$ actually refers to the unified *electroweak* force [111, 112], in a somewhat analogous fashion to how electricity and magnetism were unified by Maxwell's equations into electromagnetism. This gauge group, unsurprisingly, corresponds to four massless gauge bosons. However, somewhat unusually, experiments found that the weak bosons were in fact massive. This is somewhat of a problem, since a mass term for the vector bosons is clearly not invariant under the transformation Eq. 1.23. How does one construct a gauge theory with massive gauge bosons?

The solution to this problem is now known as the Higgs mechanism [227–233] and involves adding a new particle to the spectrum—specifically, a complex doublet, which would have four degrees of freedom. This particle is in a nontrivial representation of the electroweak gauge group, and given a potential known as the 'sombrero' potential. This potential has a continuous set of minima occurring at nonzero vacuum expectation value, of which one is particularly chosen by nature.

This process is known (somewhat misleadingly) as spontaneous symmetry breaking—misleading, in the sense that the gauge symmetry is never broken at any time. What we actually do is fix the gauge of the electroweak theory. There is a particularly convenient gauge called the *unitary gauge* in which three of the components of the complex doublet are set to zero, with one remaining in the particle spectrum—the Higgs boson. The other three degrees of freedom are transferred to the transverse polarization of three of the vector bosons—in other words, they are given mass. Specifically, we need to rotate the basis of the electroweak bosons by an angle known as the weak (or Weinberg) angle, so that we end up with a massless photon and three massive weak bosons. In other words, we rotate from a basis consisting of weak hypercharge and weak isospin (the basis with which $\mathrm{U}(1)\times\mathrm{SU}(2)$ is really defined) to the basis of electromagnetism and weak interactions that we observe in nature.

Notably, we can also introduce interaction terms between the complex doublet and fermionic fields, known as Yukawa interactions [112, 234]. The spontaneous symmetry breaking of the Higgs doublet means that these terms actually become mass terms, with mass proportional to the Higgs vacuum expectation value and the coupling constant of the interaction. In this way, we also say that the Higgs field gives fermions their mass.

Other symmetries

The Standard Model has yet more symmetries, not all of which were 'intentional'. There is strong isospin and baryon number, for example—but I will review those more in detail in Ch. 4, when introducing the axion.

There are also the discrete C, P, and T symmetries, standing for charge, parity, and time. The charge symmetry changes particles to antiparticles. The parity symmetry is a reflection about the origin (and so of left- to right-handed systems), and the time symmetry refers to time-reversal. Originally, it was thought that each of these were separately symmetries of the Standard Model Lagrangian. However, Wu's experiment in 1956 showed that parity was not conserved by weak interactions [235]. In fact, it is maximally violated—there do not appear to be any right-handed neutrinos or left-handed antineutrinos [236]. To account for this, it was then assumed that the combined CP-symmetry should be conserved (meaning that C on its own must also be violated) [237]. However, experimental evidence from kaon decays showed that this too was violated by the weak force [238]. We know, however, that CPT should be conserved for any local QFT with hermitian Hamiltonians [239, 240]. Then, we must face the rather unusual consequence that time symmetry is actually not conserved.

1.2.3 Gravity and particle physics

So far, I have introduced gravity in terms of spacetime geometry, and particle physics in terms of quantum field theory. However, it is also useful to understand gravity in the language of field theory, since it will also highlight more clearly the discrepencies between gravity and quantum mechanics.

First, it is sometimes useful to rewrite the metric tensor in the language of *tetrads* (or *vierbeins*) e^a_μ, defined in the following way [215, 241]:

$$g_{\mu\nu} \equiv e^a_\mu e^b_\nu \eta_{ab} \ . \tag{1.25}$$

This allows us to rewrite the Christoffel connection as [9],

$$\Gamma^\nu_{\mu\lambda} = e^\nu_a \partial_\mu e^a_\lambda + e^\nu_a e^b_\lambda \omega^a_{\mu b} \ , \tag{1.26}$$

where ω is known as the spin-connection field. In this way, we have decomposed

the connection into a part related to spatial translations, and a part related to Lorentz transformations—the spin connection fields can be considered the gauge field generated by Lorentz transformations, while the tetrads are the gauge field generated by translations. The freedom to set torsion to zero (as in General Relativity) provides a relation between the spin connection fields and the tetrads, allowing us to safely use a single dynamical metric tensor. It appears that there is a strong connection between General Relativity and Poincaré gauge invariance (with zero torsion), although the technical details of this relation seem to be an active area of research even today [242, 243].

Secondly, it is always possible to decompose the metric tensor into a background component and a dynamical component:

$$g_{\mu\nu} \equiv \eta_{\mu\nu} + h_{\mu\nu} \ . \tag{1.27}$$

Then it is possible to treat the dynamical component as a perturbation around the background, in the same way we treat any other particle field. In fact, if we examine the Einstein-Hilbert action, including other particle fields,

$$S = \frac{1}{8\pi G} \int d^4x \, \sqrt{-g} R + \mathcal{L}_{\text{matter}} \ , \tag{1.28}$$

then the stress-energy tensor of General Relativity can be defined variationally by,

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta \left(\mathcal{L}_{\text{matter}} \sqrt{-g} \right)}{\delta g_{\mu\nu}} \ . \tag{1.29}$$

The perturbation $h_{\mu\nu}$ can be interpreted as a spin-2 particle field called the graviton [244–246], sourced by the other particle fields.

The issue with gravity arises from an important aspect of particle physics that I have so far neglected—renormalizability [202–205, 247–252]. This is the process of curing infinities that routinely appear in QFT calculations by adding a finite number of counterterms to the Lagrangian which modify the 'bare' parameters (such as mass or charge). Then, the physical, measured parameters are rendered finite after calculation. Such a procedure is not actually so *ad hoc* and it is understood now, via the group renormalization equations, that these infinities are related to the dependence of physical parameters on the scale of observation, and the difficulties which arise when different length scales must be compared simultaneously.

Unfortunately, General Relativity turns out to be a non-renormalizable field theory within standard perturbation theory. This is because the gravitational coupling constant is Newton's constant G, which has inverse mass-squared dimensions. If we were to expand the action around Eq. 1.27, we would find that every term has an extra power of $1/G$. This means that the perturbative corrections at any subsequent order of perturbation theory contain terms which were not present in the lower-order approximation, and so you would require an infinite number of counterterms to renormalize such a theory.

However, such an argument only applies to perturbation theory, which does not capture the entire behavior of a field theory. The high-energy spectrum of General Relativity is dominated by black holes (among other nonperturbative objects). These objects are required to contribute to the full theory at high energies, corresponding to short distance scales, but cannot currently be fully evaluated. All up, General Relativity is incompatible with quantum mechanics both perturbatively *and* non-perturbatively, ultimately meaning that we require knowledge of a full quantum gravity theory. [253]

It should be noted, however, that non-renormalizability is not a death sentence for a theory. It is often the case that the low-energy limits of a non-renormalizable theory still have reliable predictive power. This is in fact the case for General Relativity, below the scale of the Planck mass, and for other effective field theories, for example the QCD effective field theory.

The layperson box: particle physics Previous — Next

Particle physics is the study of the very smallest objects in the universe. In the earliest days, they studied atoms. Eventually, they realized that atoms were made of yet smaller particles—electrons orbiting nuclei. Then, they realized again that nuclei are also made of smaller particles—protons and neutrons. And yet deeper, protons and neutrons are made out of quarks and gluons.

Actually, they found that there was a gigantic 'zoo' of both elementary and composite particles to catalogue. Eventually, a model for all these particles coalesced, called the Standard Model, developed in the mid-twentieth century by a number of brilliant scientists. Perhaps the most famous of these would be Richard Feynman, but the work was truly a collective effort over many decades by both theorists and experimentalists.

The main kind of question that particle physicists are faced with is, 'if I put some collection of particles together (and let them interact with each other), what kind of particles do I get out'? Or more carefully—'what is the probability I get some specific particles out?'

There is a special connection between short lengths and high energies, and particle physicists tend to use those terms interchangeably. One of the defining rules of quantum mechanics (the strange physics of the very small) is the Heisenberg uncertainty principle, $\Delta x \Delta p \geq \hbar/2$. In somewhat approximate English, this rule states that a measurement of a very small scale must be done with a measurement at very high momenta (and vice-versa). Meaning, if we want to study the smallest scales, we need very large energies. To find a new particles, we must accelerate other particles we know (like protons or electrons) to very high speeds and smash them into each other. Then all that energy can go into creating more exotic and short-lived particles to study. This is also why the sign on my office door reads 'high energy physics', and not just 'particle physics'.

The Standard Model is not merely a catalogue of the elementary (and composite) particles, and the rules for how they interact. Incredibly, the Standard Model also comes with reasons for *why* the particles interact with each other the way they do. The word that is always tossed around is 'symmetry'. In fact, when you ask the rules of the elementary particles to follow certain fundamental symmetries, you recover all of the interaction rules as a consequence. In my opinion, this symmetry-based construction is possibly the most aesthetically beautiful revelation in all of physics—see Appendix A for my attempt at a longer explanation.

Specifically, the rules for particle interactions are known as the three fundamental forces: electromagnetism, which rules electrons and atoms and chemicals and pretty much everything big; the weak force, which rules radioactive decay; and the strong force, which rules the nuclei of atoms and glues quarks into composite particles like protons and neutrons.

1.3 Dark matter

Dark matter refers to the hypothetical invisible content of the universe, comprising roughly eighty percent of the matter density of the universe. Whether dark matter is a larger problem for gravity, or for particle physics, is a question whose answer depends somewhat on personal opinion—certainly, long and full careers can be made in either direction. Either way, the nature of dark matter is, in my opinion, perhaps the largest remaining open question in physics. This problem has been troubling physicists for almost a century now, since Fritz Zwicky first coined the term in 1933 after observing the velocities of galaxies in the Coma cluster [254, 255]. It should be mentioned, however, that a number of people both before and after Zwicky made similar observations and proposed similar concepts [256–259]. Without getting into too much of a history lesson, by around the 1970s, evidence for dark matter had snowballed into a problem too large to ignore [260–265].

1.3.1 Evidence for dark matter

I will briefly summarize the mountain of evidence for dark matter, going from the smallest scales to the largest. It should be noted, of course, that the smallest scales here are still rather large—dark matter has not yet been detected by any laboratory experiments, and so we primarily must look towards astronomical observations. Even the solar system appears to be too small—we don't yet have the precision to detect the effects of dark matter on planetary dynamics [266], but perhaps future interstellar probes could measure a tidal force from dark matter [267].

Galactic dynamics

The smallest scales in which the local dark matter density has been measured today sits in the ~ 1 kpc range, primarily from surveys of stellar motions [268, 269] in our neighborhood of the Milky Way. Beyond our neighborhood, the velocity dispersion of galaxies and globular clusters additionally provides evidence for dark matter [270].

The related measurements of galaxy rotation curves perhaps comprises the most famous evidence for dark matter [259, 262, 263, 271–275]. The rotational velocity of stars has been measured to have a flat radial distribution, all the way out to the edges of spiral galaxies. If the majority of matter was in stars and other 'baryonic' matter (conventionally referring to all hadronic and leptonic matter in astrophysics), this curve should drop off.

Instead it supports the existence of a large *halo* of invisible matter, four times more massive than the stellar mass of the galaxy and extending to a radius a few times that of the visible galaxy.

Galaxy clusters

Moving to the next hierarchical tier of structure, galaxy clusters provide a wealth of evidence for dark matter—both in general, and in some famous specific cases. Of course, there is again the velocity of the virialized galaxies to support the dark matter [254, 255, 276, 277]. Measurement of X-ray temperature from the intergalactic electron plasma also gives an independent measure of the mass of clusters [278–280]. Intriguing also as a test of General Relativity are the strong lensing images observed in clusters [91, 281, 282]. These striking images of magnified background sources provide yet another fully independent way to measure the cluster mass. It is perhaps the incredible agreement of all these varied measurements that gives galaxy clusters so much power as evidence for dark matter—anyone wishing to suggest a dark matter-less hypothesis must contend with an extremely difficult-to-reproduce set of observations.

On top of this, collisions of galaxy clusters provide additional persuasive evidence for dark matter [283]. The most famous of these, the Bullet Cluster [284–286], has a center of mass displaced from the baryonic center of mass. Such a displacement is relatively easy to explain with dark matter but extremely difficult to reproduce otherwise.

Structure and simulation

The next tier to examine is the universe-wide distribution of structure. It is possible to straightforwardly estimate that baryonic matter alone could not lead to the collapse and formation of structure within the timescales required for the universe today [287]. More recently this has been studied with detailed many-body simulations of universe evolution [288–292]. Although there are necessary simplifications required to make these simulations feasible, this field is incredibly active today and constantly improving—for example, by including the effects of baryonic physics such as supernovae, stellar evolution, and realistic gas physics. The strong resemblances between the simulations and real-life observations [149–151, 153] provides consistent evidence for dark matter. Still, there were some interesting issues relating to these simulations, such as discrepencies with the cuspiness of the cores of galaxy dark matter profiles [293–295], as well as the number

and composition of dwarf galaxies [296, 297]. It may be that these indicate some new physical problems to explore, or it could be that the simulations were missing some level of sophistication needed to accurately reproduce the observed profiles and abundances of dwarf galaxies [298–301]. It does appear that the inclusion of more complicated physics (and better observations) have alleviated many of these issues [302, 303], although new discrepancies indicate the need for ongoing study.

Cosmology

On the largest scales, the cosmological evidence for dark matter is incredibly strong, and consistently the most difficult to reproduce with any other hypothesis. The power spectrum of CMB anisotropies [124] is only correctly reproduced when there is a large cold dark matter component in the early universe, and similar results hold for all of the various CMB observables measured by missions such Planck [126]. Coupled to a wide range of information from large scale structure surveys [149–151, 153], cluster measurements from the Sunyaev-Zel'dovich effect [304, 305], measurements of baryon acoustic oscillations [128, 306], and Lyman-α forest measurements [307], the $\Lambda\mathrm{CDM}$ cosmological model consistently emerges as the most heavily-favored cosmological model [126, 157]. Even farther back in time, calculations of BBN [107, 118–120] determine the proportion of light nuclei like Hydrogen, Helium and Lithium, limiting the quantity of baryonic matter and so requiring the existence of dark matter.

1.3.2 Dark matter candidates

All these observations give us some information about the necessary properties of dark matter. Most obvious is that the dark matter should be sufficiently dark—it should interact weakly enough with photons that we do not (yet) observe it. It certainly should interact gravitationally, and it may even be the case that it *only* interacts gravitiationally (as unfortunate as that would be). Across all the length scales, from globular clusters to the CMB, dark matter should have energy density around $\Omega_{\mathrm{DM}} = 0.26$, with the baryon density being $\Omega_{\mathrm{B}} = 0.05$ [126, 269]. Cosmological models generally prefer cold dark matter, so that its free-streaming length is sufficiently small that structures can be formed correctly and explain the CMB anisotropies [308–311]. However, it is not too difficult to construct warm or hot dark matter candidates that are still compatible with observation [312–314].

Historically, dark matter was often considered to be collisionless, in that it interacts

only weakly with other Standard Model particles [315], and that the dark matter could not form structures equivalent to planets or stars, since it could not lose kinetic energy efficiently enough [316]. However, in recent years, more careful analyses of the 'particle physics' of dark matter has become increasingly pertinent, challenging many of these assumptions. Indeed, warm, interacting, and even self-interacting dark matter candidates are now well-studied, and may even alleviate some cosmological tensions [317–322]. In fact, after so many decades of observations without detection, theorists have had ample time to construct hundreds and hundreds of dark matter models—for almost any general rule dark matter is expected to follow, you will certainly find a friendly professor in the back of a conference talk with an interesting model defying it (or, even, sitting at the front and supervising you [323]).

Dark matter candidates might be broadly classed into three groups—particles, non-particles, and modified gravities. The first is a huge category, containing WIMPs, axions, fuzzy dark matter, sterile neutrinos, and everything else under the sun. The second, non-particles, contains compact objects like black holes, as well as other kind of composite or macroscopic objects. Finally, modified gravities attempt to tackle gravity instead of adding new particle fields.

In the interest of brevity and relevance, I will not give a maximally in-depth overview of dark matter candidates and their properties. Two candidates—primordial black holes and axions—are most relevant to my thesis, and I will review them in more detail in Chapters 2 and 4 respectively. If you are reading this and are particularly fond of a dark matter candidate which I have neglected in this introduction, I can only offer my sincerest apologies.

Particle dark matter

Particle dark matter is probably the most 'popular' dark mattery category, spanning a wide range of models—from single particles with differing spins and interactions [324–326], to entire dark sectors [327] and 'mirror' Standard Models [328]. In light of this, I will summarize just a couple of the benchmark candidates here.

One of the earliest suggestions for dark matter was neutrinos [312–314, 329, 330], which have very small masses, are thermally produced in the early universe, and are indeed quite dark. However, Standard Model neutrinos are a hot dark matter candidate, and it was found that they free-stream far too readily for correct structure formation. While cosmo-

logical measurements also rule out a large neutrino fraction, sterile neutrinos [331–334] are another popular candidate—massive extra neutrinos which are often introduced in beyond-the-Standard-Model theories such as the the 'seesaw' mechanism [335] which gives the Standard Model neutrinos their small masses. These neutrinos are a specific example of the more general warm dark matter, with properties in between that of cold and hot dark matter.

Weakly interacting massive particles (WIMPs) [336, 337], are a somewhat general class of particles with properties similar to the weak bosons in terms of mass and Standard Model couplings. These candidates are most famous for the so-called 'WIMP miracle', the name for the observation that the thermal production of a Standard Model-weak-scale dark matter particle gives the correct dark matter abundance [338–342]. Originally, WIMPs were motivated from a theoretical perspective as the lightest stable supersymmetric particle [187–189, 343], but the minimal supersymmetric Standard Model [344] is no longer favored experimentally [345].

Direct detection searches have been ongoing for decades now, and the constraints are fast approaching the 'neutrino floor', where a signal would not be distinguishable from neutrinos [346]. Although novel detection paths like directional detection experiments can alleviate this problem [347–349], the 'classic' WIMP has seen a decline in popularity in recent years, as other alternatives, such as the axion, gain prominence [337]. Still, the term 'WIMP' is somewhat vague from a more modern perspective regarding the particle physics of dark matter. Indeed, WIMP-like dark matter at both lower and higher masses remains an interesting dark matter candidate [350–352], and we could even consider non-thermal WIMPs, as in Ref. [353].

I explore axions and axion dark matter [354–360] in great detail in Ch. 4, so I will defer their introduction until then. In contrast to WIMP dark matter, axions are extremely light. The more general case of *fuzzy dark matter* refers to general scalar particles which is sufficiently light that the particle wavelengths are already in the order of parsecs, leading to wavelike effects such as interference patterns in the dark matter distribution [361].

Macroscopic dark matter

Macroscopic dark matter generally refers to dark matter candidates which can be measured in units such as grams and centimeters, and so have true geometrical cross sections and elastic scattering with regular matter. These could, for example, be composed of Standard

Model particles—an example of which are hypothetical *quark nuggets* [362, 363]. Perhaps the most prominent candidate in this category is primordial black holes (PBHs) [364–367], which I will introduce in detail in Ch. 2.

Apart from the extensive bounds on PBH dark matter, there are a small number of unique and sometimes funny constraints on macroscopic dark matter. Some of the strongest constraints in the $10^{-10} - 10^{2}$ g range are from the lack of tracks in ancient mica samples [368], but perhaps my favorite constraint is placed from the non-observation of sudden, unexplained human deaths with approximately bullet-like wounds [369].

Modified gravity

Finally, modified gravity theories [370, 371] could be a natural and elegant solution to the dark matter problem, which is, after all, primarily observed via gravitational effects. Besides, it is well understood that General Relativity should be modified at quantum scales. Unfortunately, none of the simplest natural extensions to General Relativity, such as quadratic gravity [78] or massive bimetric gravity [372–374], provide a dark matter explanation. However, theories which modify gravity at the level of the Einstein-Hilbert action can be useful for both inflation [135, 375, 376] and dark energy [377]).

To explain the dark matter, somewhat more *ad hoc* modifications of General Relativity must be undertaken, such as Modified Newtonian dynamics (MOND) [76], where Newton's laws are modified in the galaxy-scale regime to correct for observations such as galaxy rotation curves. Various relativistic extensions of MOND exist, most notably tensor-vector-scalar gravity (TeVeS) [77, 378]. While these theories do indeed explain galaxy-scale dark matter phenomena, it is more difficult to explain evidence from galaxy clusters and the CMB power spectrum [379–381]. In particular, cluster mergers such as the Bullet Cluster [382] can be a challenge to explain, although there are counter-examples to these claims [383–385]. It is perhaps unfortunate that these modifications of gravity must be so complicated in order to attempt to explain the dark matter—the complicated field spectrum of TeVeS does not strike me as particularly more palatable than adding additional dark matter particle fields. Still, it is important to continue exploring these possibilities.

The layperson box: dark matter Previous — Next

Dark matter has always been the central theme of my PhD work. It refers to the observation that there seems to be roughly four-to-five times more matter in the universe than we can see. Evidence for this invisible mass can be found throughout the universe—in the motions of stars throughout our galaxy and others, in huge clusters of multiple galaxies, and even on the scale of the entire universe. When we run simulations or make theoretical predictions about the history of the universe, we do not end up at the 'correct' conditions today, unless dark matter is included in very early stages.

The fact that dark matter has to have been around since not long after the big bang rules out a number of otherwise-fun scenarios, like alien Dyson spheres or dead stars and planets. Physicists' best guesses for dark matter fall into three categories—some kind of yet-unknown particle, some kind of exotic dark object, or some kind of 'we got gravity wrong'.

In this thesis, I look at some specific examples of those first two options. The third, while interesting, is probably the most 'unexplored' option—it is very difficult to make a new gravity theory which gets everything correct that General Relativity gets correct, while still accounting for dark matter on every scale of the universe. For the first two categories, thousands of ideas have been proposed in the last century. Some of them are ruled out, but there are still many viable options. It could even end up that in reality, there are multiple dark matter components.

I could understand why someone might not care what the answer to this question is. So we find out its some kind of particle—who cares? Besides plain curiosity, I can only answer with some long-sighted context. There have been many times in history where the problems left unsolved by physics seemed esoteric and unimportant. At the end of the nineteenth century, there were only a handful of nagging issues with the laws of electromagnetism, and one could be excused for imagining their eventual solutions to be trivial or dull. One of these issues had to do with the constant speed of light, and another with the electric signal found when light struck a material (i.e. the photoelectric effect). With the perfect glasses of hindsight, however, we know that these issues snowballed into General Relativity and quantum mechanics, respectively—two of the most important theoretical paradigm shifts in science. It would be hard to overstate the technological and sociological impacts of these discoveries over the last century.

I do not claim that the discovery of dark matter will change the future of the world. But, you never know—and my salary is relatively cheap. Might as well make a long term investment...

2
Black holes

Make some sort of celestial sphere, too, and poke holes in it, and make the lights move in uncertain Copernican ways
and in the night when the hogs are resting they softly sway and sometimes in retrograde

—Shadrach Cohen, *Seedhead*

Black holes are, in my opinion at least, the most fascinating objects in the universe. Perhaps the chief reason for this is that they are a perfect confluence of gravity and particle physics. For one, black holes spew out a constant stream of all particles in a process known as Hawking radiation [386, 387]. Secondly, black holes in binaries or in galactic centers can accrete material to produce extremely energetic signals, which have proved invaluable in our understanding of galactic and even cosmological history [14, 366, 388–391]. Thirdly, gravitational waves [42, 85, 86] from black hole binary systems measured by LIGO/Virgo [87–89] provide us with a unique new astronomical 'eye', entirely complimentary and orthogonal to electromagnetic observation.

In addition, the possibility that black holes form in the very early universe [364–367] gives us two leverage points to probe new physics. We can use cosmological observations to constrain the possible physics of these black holes, and better our understanding of black holes and gravity. However, we can also turn this idea upside down, and invoke these black holes to explain a number of outstanding cosmological problems—one of which is the nature of dark matter.

2.1 Black hole basics

Although black holes are thoroughly entrenched in popular thought today, it was not always the case that these objects were taken so seriously. It is interesting and perhaps instructive to briefly examine the history of black hole research which led us to today. With this context in mind, I will then introduce some of the basic physics of black holes—chief of which, as far as the work in this thesis is concerned, regard the nature of the black hole horizon, and the thermodynamics of Hawking radiation.

2.1.1 History

The idea of an object with such strong gravity that even light could not escape was first suggested as long ago as the eighteenth century [392, 393], although in many ways the original *dark star* concept differs significantly from the modern understanding of black holes. The first step in the modern black hole story occurred only a few months after Einstein's 1915 works on general relativity [394, 395], when Karl Schwarzschild solved the Einstein equations assuming spherical symmetry [396].

The generality and simplicity of the Einstein equations mean that there are endless solutions which can be constructed. However, it is easy to find 'unphysical' solutions, corresponding to situations which could not exist in the real universe. It was debated for some time whether Schwarzschild's spacetime was real or mere curiosity. First, Arthur Eddington [397] and Lemaître [99] showed that the singularity at the Schwarzschild horizon was a coordinate singularity, meaning it was non-physical. Around the same time, Subrahmanyan Chandrasekhar derived an upper mass limit for white dwarfs [398]—objects held up by electron degeneracy pressure—lending credence to the idea of stellar collapse to black holes. The prediction of neutron stars, held up by the even stronger neutron degeneracy pressure and nuclear forces, came shortly after. Robert Oppenheimer, George Volkoff and Richard Tolman [399, 400] similarly showed that neutron stars also had an upper mass limit, not much higher than the white dwarf limit. Without a clear force to resist gravitational collapse above this limit, the existence of black holes started to look more likely.

Despite these advancements, it took until the late '50s and '60s for physics to enter the 'golden age' of black holes, beginning with the identification of the event horizon by David Finkelstein [401], and the complete extension of a more useful black hole coordinate system

by Martin Kruskal [402]. Meanwhile, more generic black hole solutions were found, including the Kerr solution for rotating black holes [403]. Theoretical work on black holes was bolstered by topological insights from Roger Penrose [404] and Stephen Hawking, continuing through the 1970s (along with James Bardeen and Jacob Bekenstein) with the quite radical development of black hole thermodynamics [386, 387, 405, 406]. Throughout this golden age, astronomical observations were also flowing in—the discovery by Jocelyn Bell Burnell of pulsars [407] provided evidence for the theoretical predictions of relativity, and the the galactic X-ray source Cygnus X–1 [408] was eventually accepted to indeed be a black hole itself [409].

Of course, the history of black hole research did not end in the '70s. Although the golden age may be in the past, the era of *primordial black holes* [364–367] is still in full swing, particularly with the exciting discovery by LIGO [90] of gravitational waves from black hole binary coalescences. And as we will see, there is still much which is not fully understood yet about black holes (and gravity in general).

2.1.2 The Schwarzschild solution

With that in mind, I will leave the history lesson behind and more carefully explore the physics of black holes. The Schwarzschild metric [396] is a spherically symmetric solution to the Einstein equations in vacuum. In fact, Birkhoff's theorem [410, 411] states that the Schwarzschild solution is in fact the *unique* spherically symmetric vacuum solution. Fast-forwarding through the pleasantries of its derivation, the Schwarzschild spacetime interval can be written,

$$\mathrm{d}s^2 = -\left(1 - \frac{2Gm}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2Gm}{r}\right)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \,. \tag{2.1}$$

It is interesting, especially in the context of later work in this thesis, to dwell momentarily on the fact that this is a *vacuum* solution to Einstein's equations, where the stress energy source $T^{\mu\nu}$ is identically zero. There are two important lessons to be gleamed from this. The first is that it is easy to naively imagine matter falling into a black hole (e.g., via stellar collapse) and 'staying there', in some sense, at the center of the black hole. This is not strictly correct—the Schwarzschild black hole does not contain any matter. In that sense our intuitions are challenged from physics such as electrostatics, where a point charge resides at the center of the electric field. Here, it is gravity itself which is gravitating (or more carefully, gravitational energy is itself a source of gravity). From

another perspective, the Schwarzschild solution is pure 'geometry' of spacetime.

2.1.3 Defining the black hole mass

The second lesson from the vacuum solution, continuing the thoughts above, regards the notion of mass in General Relativity. In theories which posses a time-reparametrization invariance (and not merely a time-translation invariance), it can be shown that the Hamiltonian is identically zero. This is certainly the case in General Relativity, where we have full coordinate-reparametrization invariance. This means that notions like energy and mass need to be carefully defined in General Relativity, and they may not be conserved—the classic example being the cosmological redshifting of photons from the expansion of the universe. I have included a longer discussion of this point in appendix B.

In the Schwarzschild metric, there is a parameter which we tellingly label as 'm', but strictly is just a parameter of the metric. How do we know it refers to mass then? One way is to examine the Newtonian limit of the theory. If we compare geodesics in General Relativity with geodesics in Newtonian gravity, one can find that the Schwarzschild 'm' parameter must correspond to the Newtonian gravitational mass.

There is perhaps a second way to identify 'm' with the mass. One could construct a (different) time-dependent metric, which corresponds to the collapse of a non-compact mass, such as a star, into a black hole. In this case, we would begin with an object of well-defined mass, and the distant-future limit of this nonstationary metric would presumably be the Schwarzschild metric. In this limit, we should be able to identify the initial mass of the object with the final 'm' parameter. The conclusion of this thought, then, is that identifying mass is not trivial in General Relativity—although it is straightforward for the Schwarzschild black hole, we had to look at either Newtonian limits or more complicated collapse metrics to draw conclusions about the mass.

This hairiness needs resolved with some more careful definitions of black hole mass. Deep dives into the technical details of General Relativity often invokes the dybbuk of unintelligible maths, so I will give only a brief review here. The Arnowitt–Deser–Misner (ADM) formalism is a Hamiltonian formulation of General Relativity which foliates spacetime into spatial surfaces, parametrized by time [412]. As well as being useful for numerical General Relativity, this system allows a mass to be defined essentially by examining the energy of the spacetime at a well-defined infinity. This is an important point to emphasize—the Schwarzschild metric is asymptotically Minkowski space at spatial infinity, so the ADM

mass is well-defined. As we will see, the situation is more complicated when we examine black holes in a true cosmological background.

The Misner–Sharp mass

It is therefore useful to have a truly local definition of mass in General Relativity, recalling from above that the mass of the black hole is actually just the gravitational energy. Using this idea, we can define an alternative notion of mass known as the Misner–Sharp mass, m_{MS} [413, 414]. This mass is the quasi-local measurement of the curvature-producing energy of a spacetime configuration. As any good notion of mass should be, the Misner–Sharp mass is an invariant quantity, defined geometrically as,

$$1 - \frac{2m_{\mathrm{MS}}}{R} \equiv \nabla^c R \nabla_c R \,, \tag{2.2}$$

where R is the 'areal' radius $R = \sqrt{\mathcal{A}/(4\pi)}$ for 2-sphere area $\mathcal{A}$, and ∇_c is defined with respect to the 2-metric on the submanifold orthogonal to the time and radial directions [415]. For the Schwarzschild metric in the coordinate system of Eq. 2.1, the coordinate r is the areal radius and the Misner–Sharp mass properly reduces to m. This quasi-local mass notion will come in useful when examining black holes with more complicated metrics than the Schwarzschild metric—particularly dynamic black holes, which change size over time, as can be the case in cosmological settings.

2.1.4 Black hole horizons

With a more nuanced understanding of black hole mass behind us, let us move on to the exciting notion of the black hole horizon. It is again the case here that while the Schwarzschild metric permits a relatively straightforward definition of its event horizon, we require somewhat more powerful notions in order to carefully describe other black hole metrics.

When examining the Schwarzschild metric Eq. 2.1, we can see that there is a kind of qualitative transition that occurs at the coordinate singularity $r = 2Gm$, known as the Schwarzschild radius r_{schw}. At smaller radii, the the signs of the $\mathrm{d}t^2$ and $\mathrm{d}r^2$ terms swap. This has a profound effect on the geodesics of light and matter within the Schwarzschild radius—all paths are forced into the central 'true' singularity. The surface at $r_{\mathrm{schw.}}$ is then an *event horizon*, a boundary beyond which no information can be extracted. More

rigorously, we say that an event horizon is generated by null (light-like) geodesics which do not reach spatial infinity.

Similarly, an important radius is the 'photon sphere', defined as the radius in which photons have a circular orbit. This occurs at $r = 3Gm$, so even photons travelling outside the event horizon may not escape the singularity in the end.

Apparent horizons

The event horizon has a rather stringent definition, since it requires knowledge of the entire future history of the spacetime. The situation is more complicated for dynamical black holes, for similar reasons to the difficulty in defining mass above. In this case, it is preferable then to have a local definition of horizons, and a way to categorize their properties [416].

Building on the the more rigorous comment above regarding null geodesics, we define ℓ^μ and n^μ as the tangent vectors to outgoing and incoming null radial geodesics, respectively. Since they are null, they satisfy,

$$\begin{aligned} \ell^\mu \ell_\mu = n^\mu n_\mu &= 0 \ , \\ \ell^\mu n_\mu &= 1 \ . \end{aligned} \tag{2.3}$$

For the cross-normalization above, there is a difference in minus sign depending on metric signature, and sometimes the normalization 2 is used instead of 1.

These null geodesics can be used to define vector fields over a given surface. In the spherically symmetric case, we are particularly interested in surfaces defined with some fixed radius. To examine the characteristics of this surface, we can take the Lie derivative with respect to these null radial geodesics. Then the extrinsic curvature of the surface is encoded in a number of standard scalar fields, known as expansions, shears, inaffinities, and twists. For our purposes, the expansions are the most relevant field—by comparing the expansion scalars for both the incoming and outgoing radial geodesics, we can find the surfaces where there are the analogous qualitative shifts in geodesic behavior as we observed for the Schwarzschild metric. Without getting too bogged down in Lie

derivatives, these scalars can be calculated to be [415, 417],

$$\theta_\ell = \nabla_\mu \ell^\mu + \ell^\mu \ell_\nu \nabla_\mu n^\nu \ , \tag{2.4}$$

$$\theta_n = \nabla_\mu n^\mu + n^\mu n_\nu \nabla_\mu \ell^\nu \ . \tag{2.5}$$

The signs of these expansions give us the behavior of geodesics on the surface. A normal surface, (i.e., far away from any pathological spacetime), corresponds to the 'usual' geodesic behavior, $\theta_\ell > 0$ and $\theta_n < 0$. In contrast, a surface is *trapped* if $\theta_\ell < 0$ and $\theta_n < 0$, implying that the outgoing rays converge on this surface, instead of propagating outward. There is a special case, then, when one of these two expansion scalars is zero, since this will be the boundary where the behavior of geodesics qualitatively changes. This is known as the *apparent horizon*. To ensure that this behavior does indeed swap over this boundary, we also require the condition on the Lie derivative $\mathcal{L}_n \theta_\ell < 0$, to ensure that the apparent horizon is also *trapping* (sometimes referred to as *outer* or *inner*).

It is certainly important which of the two expansions is zero at the horizon. The case where $\theta_\ell = 0$ and $\theta_n < 0$ desribes a *future* apparent horizon, since light rays are being dragged back down beyond this surface. It can be checked relatively easily that the Schwarzschild horizon is indeed a future, outer trapping horizon. In contrast, the case $\theta_n = 0$ and $\theta_\ell > 0$ corresponds to a *past* horizon, where light rays propagate outwards without returning back. This is clearly the 'opposite' of expected black hole behavior, and so this kind of horizon is associated with either a white hole, or a cosmological horizon. In this case, $\mathcal{L}_\ell \theta_n > 0$ is the trapping condition.

Such rigorous definitions of apparent horizons are certainly a mouthful, and it is not helped by the fact that the describing phrases 'trapping', 'inner/outer', and 'past/future' are used in very slightly different ways throughout the literature. However, as we will see when studying dynamical black holes, these local definitions of horizons will prove invaluable.

2.1.5 Thermodynamics and Hawking radiation

Hawking radiation refers to the spontaneous emissions of particles from the surface of the black hole [386, 387, 405, 406]. A reasonably accurate derivation of Hawking radiation can be performed with essentially classical physics—I will reproduce this calculation here briefly, while following a somewhat historical story based on the interesting review by

Don Page [418].

The study of black hole thermodynamics began with Bekenstein, who suggested that black holes should have some entropy S proportional to the area A of the horizon [405]. This idea was based on various thought experiments involving the lowering of quantum systems slowly into black holes, and attempting to reconcile the results with the standard laws of thermodynamics. Separately, Hawking, Bardeen and Carter [406] noticed the connection between black hole area and entropy, based on the idea that the black hole horizon can never decrease (e.g., when merging with other black holes or matter). They generalized this connection to thermodynamics into the four laws of *black hole thermodynamics*:

0. The surface gravity κ, angular velocity Ω and electrostatic potential Φ are constant across the horizon. This is compared to regular thermodynamics, which requires constant temperature T for systems in thermal equilibrium.
1. The black hole mass is related to area A, angular momentum J, and electric charge Q via the compellingly-familiar relation,

$$\delta M = \frac{1}{8\pi}\kappa\delta A + \Omega\delta J + \Phi\delta Q \tag{2.6}$$

2. The area of the black hole horizon cannot decrease, exactly analogously to entropy in standard thermodynamics.
3. The surface gravity cannot be reduced to zero by a finite number of 'operations'. This is related to the weaker form of the third law in thermodynamics, regarding reducing the temperature of a system to zero.

Although these coincidences are certainly entertaining, it was not realized at first that they were physically significant. However, using QFT in black hole spacetimes, Hawking found that black holes did indeed emit radiation in a perfect blackbody spectrum with temperature given by $T_{\rm BH} = \kappa/2\pi$ [386, 387]. The famous relation $S_{\rm BH} = A/4$ then follows from the first law of black hole thermodynamics (with the 'BH' standing either for black hole, or Bekenstein-Hawking).

Actually, armed only with the above, it is possible to straightforwardly derive the black hole mass loss equation, under the assumption that the particle spectrum is given by the Stefan-Boltzmann law for blackbody radiation, $\dot{U} = -\sigma T^4 A$, where U is the internal energy of some system (i.e., we are only counting photons—in reality, all particle species are emitted, and we need to introduce *greybody* or Page factors to account for this

properly [419]). Then, we only need to use the thermodynamic identity,

$$\mathrm{d}U = T\mathrm{d}S \ , \tag{2.7}$$

where we will assume the black hole is at least approximately stationary and ignore the $P\mathrm{d}V$ term. Then we only need to substitute the Stefan-Boltzmann law for $\mathrm{d}U$, and use the above relations for T and S. Since S ultimately depends on the black hole mass, there will be an $\dot{M}$ term on the right hand side. After rearranging we then find,

$$\dot{M} = \frac{\hbar c^6 \alpha}{G^2 M^2} \ , \tag{2.8}$$

where $\alpha = 1/15360\pi$. When considering the full particle spectrum, α is modified nontrivially, since the exact spectra produced depends on the temperature of the black hole (typically, $\alpha \sim 10^{-4}$). This differential equation is easily solved, allowing us to find the lifetime of a black hole, in terms of its initial mass, assuming α is constant:

$$t_{\mathrm{BH}} = \frac{5120\pi G^2 M^3}{\hbar c^4} \sim 10^{67} \text{ yrs} \left(\frac{M}{M_\odot}\right)^3 \ . \tag{2.9}$$

More accurately, we would like to account for all particle species which can be emitted. In that case, with a more worked calculation one can find the emission rate for each of these species:

$$\frac{d^2 N_i}{dtdE} = \frac{1}{2\pi} \sum_{\mathrm{dof}} \frac{\Gamma_i(E, M, a^*)}{e^{E'/T} \pm 1} \ , \tag{2.10}$$

where N_i is the number of particles emitted for species i, Γ_i is the 'greybody factor', E' is the energy of the particle (including the BH spin), a^* is the reduced spin parameter, the sum is over the degrees of freedom of the particle (including color and helicity), and the $\pm$ sign accounts for fermions and bosons respectively. Of course, this is not analytically tractable, so full details of black hole evolution must be performed numerically, e.g. with the BlackHawk software. [420, 421]. As a result, the analytic estimate is actually quite handy, although it will predict slightly too-long lifetimes, as it only accounts for photons (however, it is possible to correct for this with some numerically-based semi-analytic estimates [422]).

Hawking radiation is one of the most exotic features of black holes, explicitly connecting gravity with particle physics. Indeed, evaporating black holes are often employed as a tool within the folklore of quantum gravity, as a way to violate sacred rules of quan-

tum mechanics like unitarity. Many full theses could be written on such problems, and surrounding thought experiments, but we must move on.

2.1.6 Naked singularities

The central singularity of black holes has always been a point of interest, both to physicists and to a lay audience. Of course, describing such a singularity is out of the realm of validity of General Relativity (if not many classical gravity theories). It is perhaps the case that the central object of a black hole may be describable by some quantum gravity theory, but without one, observing a singularity makes it difficult to form well-posed Cauchy problems for spacetimes. As such, it was suggested by Penrose [404, 423] that there are no naked singularities—or rather, all singularities must be hidden behind a horizon. This hypothesis is known as the cosmic censorship hypothesis, and is supposed to restore determinism to General Relativity.

It is not hard to dream up spacetimes which possess naked singularities, and so the formulation of this hypothesis is somewhat ill-defined, and subject to much technical disagreement [424, 425]. It may be considered a kind of additional assumption accompanying General Relativity that spacetimes which possess naked singularities should be considered unphysical. Or, perhaps, the cosmic censorship hypothesis is a actually a consequence of a deeper theory that is not yet well understood. Either way, it is not considered particularly controversial to suggest that good black hole metrics should not contain a naked singularity, and we will use this principle later when searching for useful cosmological metrics.

2.1.7 Stellar black holes

A basic introduction to black holes would not be complete without a brief examination of the production of black holes from stellar collapse. Of particular importance here is the possible mass spectrum which can be produced by these processes—only a relatively narrow range of black hole masses can be produced, and there are multiple 'mass gaps' where it may be difficult to find stellar black holes [426]. These gaps are particularly important in the context of this thesis, since black holes detected in these gaps (e.g., by LIGO/Virgo [87–89]) could possibly constitute evidence for primordial black holes.

Stellar black holes are formed by the gravitational collapse of large stars. The simplest

model of stellar collapse occurs when nuclear fusion is no longer able to provide enough energy to support a star (e.g., at the onset of iron fusion, which sucks energy from the core). Stars may collapse in a large explosion known as a supernova, or hypernova if the explosion is sufficiently energetic and conditions are right. Collapsing stars can leave behind a variety of remnants, depending mainly on the mass of the star. If the remnant is too large to be a white dwarf or neutron star, then our best understanding is that a black hole should be left.

Stellar black holes are expected to be roughly more massive than neutron stars. The lack of observational evidence for black holes a few solar masses above the neutron star cutoff has led to the concept of a 'lower' mass gap for black holes at around $\sim 5\ M_\odot$ [427–429]. It is not well understood why this is the case—it could be a lack of understanding of stellar evolution, but it also could merely be that our observational methods are biased against such black hole masses.

There is also an 'upper' mass gap, from roughly $50 - 150\ M_\odot$ (although the exact size of this range is in dispute). Here, sufficiently large stars suffer from pair-instability, where the core is hot enough to produce electron-positron pairs [426, 430, 431]. This production either results in the total destruction of the star, or significant mass loss leading to usual supernovae after. This mass gap is really more of an upper limit on stellar black holes. The largest star is only $\sim 250\ M_\odot$ [432, 433] and so it is hard to to imagine stellar black holes being produced above $\sim 150\ M_\odot$.

I have introduced stellar black holes in this thesis to highlight the interesting gravitational wave observations by LIGO/Virgo [87–89], which appear to have found black holes in these mass gaps. Most notably, GW190521 [434] measured black holes of masses $\sim 85\ M_\odot$ and $\sim 66\ M_\odot$, with an uncertainty of around ± 20 for both, putting them arguably in the upper-mass gap. There has been a fervent campaign to explain these observations, including scenarios with hierarchical mergers of smaller stellar mass black holes [435, 436], or reexploring the possibility that extremely large stars could in fact collapse to these black holes [437]. Of course, it could be the case that these black holes are *primordial black holes* (PBHs), but there is not yet definitive evidence for that hypothesis [438–440]—I will discuss PBHs in more detail again in Sec. 2.3. With that intriguing clue in mind, let us now look more closely at the current evidence and observations for black holes.

2.2 Observations of black holes

As I alluded to earlier, part of what makes black holes such an exciting dark matter candidate is that there are a large number of independent observations which collectively make it very hard to doubt the existence of black holes. Even more thrilling is that many of the most persuasive observations have only come in relatively recent years.

Gravitational waves

Perhaps the most famous of these recent observations is the detection of gravitational waves by LIGO in 2015 [90]. The waveform was consistent with the coalescence of two $\sim 30M_\odot$ compact objects, since the separation between the two masses could only have been a few hundred kilometers before the merger. In addition, the shape of the ringdown signal gave extra weight to the binary black hole interpretation. Since 2015, LIGO/Virgo [87–89] has detected almost one hundred more events.

One particularly prominent event was the merger of a binary system containing neutron stars, GW170817 [441]. This event featured an accompanying electromagnetic signal, leading to an absolutely enormous wealth of new physics results and arguably the beginning of gravitational wave multi-messenger astronomy [442]. We seem to have gotten somewhat lucky, as GW170817 was observed relatively early into the observational runs of LIGO/Virgo, and we have not yet seen another accompanying electromagnetic signal.

Radio interferometry

Another recent observation was made by the Event Horizon Telescope [443], a large array of radio telescopes allowing for very-long-baseline interferometry. The telescope was able to 'photograph' the hot accreting gas around the event horizon of the supermassive black hole at the centre of the Messier 87 (M87) galaxy, producing what might be considered the 'truest' image of a black hole to date. In addition, during the time between submission and corrections of this thesis, the Event Horizon Telescope reported a similar image of the central supermassive black hole of the Milky Way, Sagittarius A* (Sgr. A*) [444]. Despite being significantly closer than the M87 black hole, the image of Sgr. A* was made more difficult by the rapidity of its oscillations, meaning that radio images from different telescopes could not be straightforwardly combined as in a standard interferometer. Instead, complex numerical simulations of the expected signal at each telescope had to be

made and compared to the true images.

Stellar motions

Also very recently, a rogue stellar mass black hole was detected through astrometric microlensing [445] with the Hubble Space Telescope. Observation of a source star over several years revealed a clear deflection in the star's position, which could be correlated to a lens of about $\sim 7M_\odot$ around one kiloparsec away. Since the lens did not emit any light, and is above the neutron star mass limit, a black hole is again the most likely explanation for this event.

Supermassive black holes in the centers of galaxies also provide a wealth of observations. In our own galaxy, the motions of stars around the radio source Sagittarius A* [446–448] at the center of our galaxy have been shown to be orbiting a $\sim 10^6 M_\odot$ object [449, 450].

Acretion and active galactic nuclei

Meanwhile, in distant galaxies, quasars are believed to be powered by the accretion of gas onto the supermassive black holes in active galactic nuclei [388–390, 451–454]. The massive energies required to power quasars can be explained by the heating of infalling gas around these black holes, forming relativistic jets at the poles of the accretion disks. In some sense, the discovery of the quasar might be considered the final straw which cemented black holes as genuine astronomical objects, and not merely a theoretical curiosity—see the entertaining Ref. [455] for an account of the First Texas Symposium on Relativistic Astrophysics, based around this discovery (and also, where Roy Kerr first introduced his rotating black hole solution [403]).

Accretion also explains observations from X-ray binary systems [456, 457], where a compact object accretes matter from its accompanying star. If the compact object appears to be invisible, and above the neutron star mass limit, we have yet another piece of evidence for black holes.

The nature of black holes makes direct observation difficult, but as we can see, major strides have been made even in the last decade and black holes now have a solid place in the pantheon of astrophysical phenomena. Anyone wishing to propose that black holes do not exist must explain quite a large number of independent observations, ranging from astronomical observations like quasars, to the exact shape of the LIGO gravitational

wave signal. That is not to say that we know the black holes are exactly given by the Schwarzschild or Kerr metrics—there may be more to understand about the exact modelling of black holes, which may or may not affect the above observations.

In addition, the existence of black holes is a separate question to whether or not these black holes could be the dark matter. Since stars presumably formed after the dark matter already existed, we will need to examine black holes which are produced in the early universe—primordial black holes—if we we wish to explain dark matter.

The layperson box: black holes Previous — Next

If gravity is pulling objects together, then why are you not falling to the center of the Earth right now, along with everything else? Think about it for a second, because the answer is not, in my opinion, totally trivial. You are stopped from falling by the ground. Why does the ground resist you? Atoms are almost entirely empty space—can they not just move through each other? Here, the answer is that the electric charge from the atoms (specifically, their electrons) repels each other, like pushing similar magnets together. But what about at the Earth's core, which is a possibly-solid hot metal ball—why does it not collapse further? I would guess that the collapse of the core is resisted this time by the repulsion of the positively-charged *nuclei* of the atoms from each other, but probably the electrons, which act like a 'soup' in the metal, play a role as well (I am not sure why, but it was particularly hard to find an exact, physical answer to this question).

But now, what about stars, which are hot balls of gas—how are they stable? When forming stars, dust clouds in space collapse, getting hotter and hotter until suddenly they are hot enough to ignite nuclear fusion in their cores. The incredible heat produced stops the collapse, as hot particles push outwards from the core. The star then settles into an equilibrium, where the heat and particles from the core push outwards against the gravitational pressure.

In each case, some kind of fundamental force prevents gravitational collapse. As you look at more and more massive objects, eventually some fundamental force is overcome and replaced with a yet-stronger one. But there are a finite amount of fundamental forces. Once fusion stops inside the star, it collapse into a 'white dwarf'. The pressure here is quite exotic—the electrons in a white dwarf are supported by the fact that the electrons are as close together as they could possibly be. If a white dwarf gets too big, however, eventually the gravitational pressure is too much, and the electrons are forced into the nuclei of atoms, combining with protons to make more neutrons and neutrinos. Then, an object known as a 'neutron star' is born—now, it is held up by the fact that *neutrons* can only get so close together. However, if the mass gets above around two solar masses, even this pressure is not sufficient. There is nowhere for the particles to go, and no more fundamental forces to stop the collapse. All the matter falls into an infinitely dense point, unless there is some as-yet unknown physics to prevent it. A *black hole* is born.

In that sense, a black hole really is 'nothing'. All the stuff that was there falls right to an infinitely dense point and is *converted* into pure gravity—curvature of spacetime—at least as far as General Relativity is concerned (a quantum gravity theory might say differently, but none exist yet). The reason we call it a black hole is because there is a massive amount of stuff crammed into a tiny space. If you go too close, then, not even light is fast enough to resist gravity, and so there is a kind of dark sphere all around the central point where you will never* see light escape. There is nothing otherwise 'special' about black holes—they are not big malicious vacuum cleaners in space, or sci-fi villains. In fact, they are technically *pure gravity*, and that is not hyperbole.

Finally, I should qualify the 'never*' above. The work that made Stephen Hawking a household name was his theoretical prediction that black holes actually softly glow, like heated up pieces of metal. The smaller the black hole, the 'hotter' it is, and the more particles and light it emits. Eventually, (very, very slowly) all black holes will evaporate away.

2.3 Primordial black holes

Stellar collapse is not the only mechanism which can produce black holes. It was realized in 1967 by Zel'dovich and Novikov [391], and then Hawking [365] that sufficiently large perturbations in the early universe could collapse to form black holes as well. These were dubbed 'primordial black holes' (PBHs) [366, 367]. In fact, it was not clear at the time whether PBHs were physically viable—but I will return to this interesting historical point when we discuss cosmologically-embedded PBH metrics in the following chapter.

Although there remains no conclusive evidence for the existence of primordial black holes, it was quickly realized that PBHs had very interesting phenomenology which distinguished them from stellar black holes. For one, PBHs can nominally form over an extremely large range of masses—from tiny black holes of mass $m < 10^{15}$ g which would evaporate before today, to supermassive black holes with $m > 10^5 M_\odot$.

As a result of this, PBHs have been invoked to solve a huge variety of problems in cosmology and astrophysics [14, 458]. Perhaps most significantly, they are a relatively attractive dark matter candidate, since they form before BBN (which otherwise constrains the fraction of baryons in the energy density [459]). In addition, the evaporation of small PBHs has been used to explain γ-ray backgrounds [419, 460], cosmic ray distributions [461], radiation from the Galactic center [462], reionization [463], and even gamma-ray bursts [464]. Non-evaporating PBHs are also interesting in their own right, however, leading to explanations of gravitational lensing, heating of stars in galaxies and globular clusters, the seeding of supermassive black holes [465–467], generation of large-scale structure [468], and maybe even the binary black holes observed by LIGO/Virgo [90, 434, 438–440].

2.3.1 Formation and predictions

As noted above, one of the strengths of PBHs—and perhaps the reason for their long-standing popularity amongst physicists—is that theorists and modellers are able to invoke whatever mass spectrum of PBHs they desire [14, 458, 469–473]. In actuality, this is not quite true. Although there is certainly a lot of freedom to toy with PBH formation models, we still need to consider physically viable and realistic formation mechanisms. Even better would be to motivate a particular PBH mass distribution based on specific physics in the early universe. Much work has been done on this subject, to the extent where so many models exist that one can presumably justify almost any PBH distribution that they feel

like. Still, some models are perhaps more interesting, or physically likely, than others.

There are many mechanisms for PBH formation [14, 469, 474, 475]. PBHs could form from collapse of cosmic strings [476–483] and domain walls [476, 477, 484–488], which will be relevant to the axion models discussed later in this thesis. Other mechanisms include bubble collisions [489–491], phase transitions [492–494], early matter-dominated eras [495–499], or more exotic particle-physics processes including additional scalar fields, inflatons, and supersymmetry [500–506].

However, the 'standard' PBH formation mechanism invokes large overdensities which collapse once they re-enter the horizon in the very early universe [14, 147, 465, 507–511]. In this scenario, the initial PBH mass m is related to the Hubble horizon mass $m_{\rm H}$ by,

$$m = \gamma m_{\rm H} = \frac{4\pi}{3}\gamma\rho R_{\rm H}^{-3} \,, \tag{2.11}$$

where the factor $\gamma \sim 0.2 - 10^{-4}$ accounts for the details of the gravitational collapse [14, 512]. In this scenario we presume that this occurs during radiation-domination, in which case the density ρ of the cosmological fluid and the Hubble horizon radius $R_{\rm H}$ can be written analytically in terms of time t, giving us,

$$m \sim 2 \times 10^5 \gamma \left(\frac{t}{1\,{\rm s}}\right) \,. \tag{2.12}$$

In order for such a perturbation to collapse into a black hole, the initial overdensity needs to be sufficiently large. This is often parametrized with the *density contrast* $\delta \equiv \delta\rho/\rho$ of the perturbation. For a PBH to form, the density contrast needs to be above the critical value $\delta_{\rm c} = c_{\rm s}^2$, where $c_{\rm s} \sim 1/\sqrt{3}$ is the sound speed during radiation domination [14, 513, 514].

The requirement of these large density contrasts naturally leads to a search for a specific theory which would source these kind of perturbations. Indeed for any formation mechanism, we are motivated to find theories which can make predictions about the produced PBH distribution, starting from some assumptions about the state of the early universe. For example, in Ref. [147], it was shown that modifying the slow-roll scenario close to the end of inflation could enhance small-scale perturbations for PBH formation, without changing CMB observations on larger scales.

Another particularly compelling scenario is described in Ref. [458]. The main idea here is that the critical density $\delta_{\rm c}$ is decreased whenever the effective equation of state parameter ω decreases. This is an effect which occurs during the various cosmological

phase transitions, notably the decoupling of Standard Model particles (electrons, pions, weak bosons, heavy quarks, etc.), from the thermal bath. Since ω, and therefore δ_c, decreases temporarily during these transitions, PBH production is enhanced, with a precise spectrum that carefully follows the thermal history of the universe. The resulting PBH spectrum, beginning from a scale-invariant primordial power spectrum, nicely satisfies the constraints on PBHs which we will discuss in detail below. In addition, the resulting PBH spectra naturally explains a few other important observations—namely, an excess in planet-sized and black hole mass-gap sized microlensing events, microlensing of misaligned quasars, unexplained correlations in X-ray and cosmic infrared background fluctuations, the non-observation of small dwarf galaxies which would be disrupted by PBH heating, the relationship of galaxy mass to central black hole mass, and finally (perhaps most importantly), the LIGO/virgo events, of which some appear to involve black holes in the mass-gap.

2.3.2 PBHs as dark matter

Let us now return to arguably the most interesting PBH question—whether or not they could comprise the dark matter [367]. For ease, we will ignore the effects of spin and charge of the black holes, which leaves only the mass of the PBH as a free parameter. Whereas particle dark matter usually has two dimensions for constraint (something like mass and cross-section), PBH dark matter only has the one dimension. This means we can plot on the vertical axis the fraction of PBH which could be the dark matter,

$$f_{\rm pbh} \equiv \frac{\Omega_{\rm PBH}}{\Omega_{\rm DM}} \,. \tag{2.13}$$

Although such a quantity can be defined for all dark matter candidates, the limitations of two-dimensional media means that this fraction is not always shown explicitly when reviewing, e.g., particle dark matter, and the constraints are plotted as if the candidate comprises the entirety of dark matter. It is therefore somewhat of an advantage for the comparatively-simple PBH dark matter that we can plot $f_{\rm PBH}$ on the vertical axis of the constraint plots, especially since I am personally somewhat partial to the possibility that the dark matter is comprised of a number of different candidates.

Since mass is a convenient characterization of PBH dark matter, I will summarize the various constraints across a few loosely-defined mass ranges. Most of these constraints assume a monochromatic mass spectrum for the PBHs.

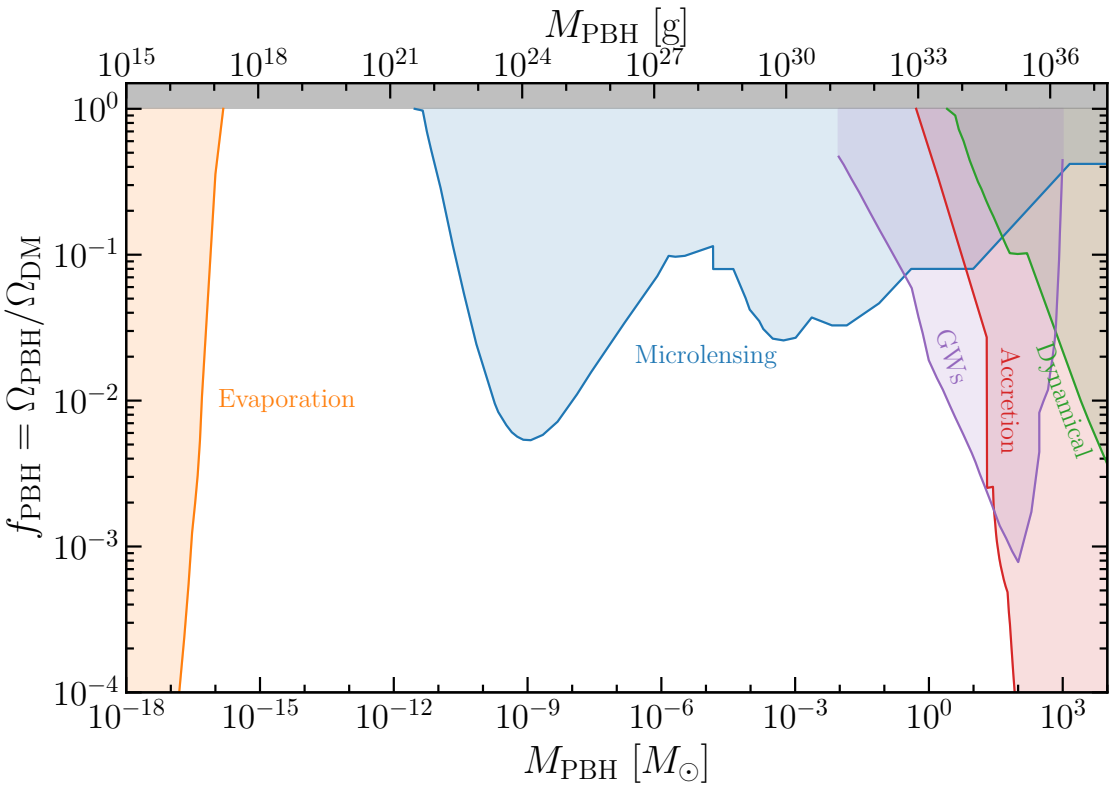

Figure 2.1: Simplified plot of PBH constraints, showing the five 'main' mechanisms of constraint—dynamical effects on stellar populations, electromagnetic signals from accretion, gravitational waves from PBH binary mergers, microlensing, and evaporating black holes. It is particularly interesting to point out the large, unconstrained asteroid-mass region, as well as the ∼LIGO-mass region where the gravitational wave constraints dominate. Notably, in the latter region, the limits on $f_{\rm PBH}$ from microlensing and accretion are both small and subject to significant observational and theoretical uncertainty. This particular PBH constraint plot [515] can be found at `https://github.com/bradkav/PBHbounds`.

Ultramassive black holes

The largest PBHs fall in the mass range of around $10^5 M_\odot \lesssim m \lesssim 10^{19} M_\odot$. Perhaps the most trivial bound is the *cosmic incredulity* limit [516] , which denotes where there is exactly one PBH within the Hubble horizon (if there was less than one, we would have nothing to write about). The other high-mass limit comes from the measurement of the CMB dipole moment, since a very large PBH would attract the Milky Way [516].

Slightly lower in mass are the the *dynamical* limits [517–525]. One such limit is due to dynamical friction, where large PBHs would be dragged to the nuclei of galaxies, forming a central black hole larger than observed. The second is that large PBHs would heat up and

destroy, or distort, galaxies within clusters. In the same range are the large-scale structure limits [468, 526, 527] which require that structure formation does not take place too early in the universe.

Then there are the accretion constraints [388, 389, 528–532] where matter of some kind falls onto the PBHs, generating an electromagnetic signal, in either X-ray or radio. These constraints arise not only from gas in the Milky Way [533–535], but on the effect of these processes in the early universe, impacting the CMB and reionization [536–540]. These constraints bleed also into the intermediate mass black hole category below. Both the dynamical and accretion constraints, however, are quite complicated—there are various loopholes [520, 541, 542] and astrophysical uncertainties [543, 544] for these bounds.

Intermediate mass black holes

Moving to smaller masses, I will colloquially define the intermediate PBH masses as $1M_\odot \lesssim m \lesssim 10^5 M_\odot$. The accretion constraints mentioned above also impact the upper range of these PBHs. Then, there are more dynamical constraints in this range—ultra-faint dwarf galaxies are sensitive to heating from PBHs [520, 521, 523, 544], and wide binary star systems can be disrupted by too many encounters with PBHs [545–549]

Most relevant to this thesis are the LIGO/Virgo constraints in this range [550–555]. These constraints involve an estimation of the PBH binary formation rate in the early universe, finding that by radiation-domination almost all PBHs are a part of a binary system. Then the number of binary mergers from this primordial binary population can be estimated, and compared with the actual rate observed by gravitational wave detectors. It turns out that, for $f_{PBH} \sim 1$, there would be $\sim 10^4$ mergers observable by LIGO per year, which therefore places very strong constraints on the PBH fraction.

The stochastic gravitational wave background also constrains these PBHs [556, 557]. I will return to these binary abundance calculations in more detail later in this thesis, however, when we investigate cosmological black holes and their effects on this constraint. On top of this, the question becomes more complicated when considering potential clustering of PBHs, which could disrupt binaries and loosen the constraints [558–562].

Microlensing-sized black holes

Smaller still are the 'microlensing-range' PBHs, which range from $10^{-12} M_\odot \lesssim m \lesssim 10 M_\odot$. Constraints in this regime are mostly related to gravitational lensing phenomena.

Microlensing refers to the scenario in which a compact object passes between a source and a detector, magnifying the background source [563–566]. Light rays passing near the compact object are deflected, so that more light converges on the detector than would otherwise. By observing a fixed population of sources for a long period (such as the Magellanic Clouds, the galactic bulge, or the Andromeda galaxy), constraints can be placed on the population of compact objects. A number of missions, such as MACHO [567], HSC [568], EROS [569], and OGLE [570] have conducted these observations, allowing relatively tight constraints in this window.

On top of the microlensing surveys, there are a few other lensing-related constraints in this mass region. Some observed lensing events, such as by Icarus [571], require a sufficiently smooth dark matter population, which constrains the somewhat 'lumpier' PBH population [572]. Similarly, the supernova magnification distribution appears to depend on the smoothness of dark matter [573–575]. Finally, extragalactic microlensing of objects such quasars can also constrain PBHs [576–580].

It does not seem like the dust has fully settled regarding microlensing constraints—while recent long-duration results [581] appear to conclusively rule out LIGO-sized PBHs at around the ten percent level, new papers seem to be published faster than I can write this thesis. Indeed, by the time you are reading this, my bounds may be out of date. For example, the recent result Ref. [582] appears to *require* a substantial solar-mass population of black holes to explain quasar microlensing surveys. It surely is an exciting period for primordial black hole dark matter.

Perhaps the volatility in the literature is due to the large number of uncertainties and model-dependent caveats that microlensing constraints must account for [583–585]. In addition, theorists always have new tricks up their sleeves—for example, two of my collaborators in Ref. [586, 587] found an alternative black hole solution in quadratic gravity [78, 79] which is stable and carries a topological charge, disrupting microlensing and so evading the constraints.

Asteroid mass black holes

An interesting range is the so-called 'asteroid range', comprising $10^{-17}M_\odot \lesssim m \lesssim 10^{-12}M_\odot$ [588]. At the moment, this range appears to be entirely unconstrained. PBHs of these masses are too small for traditional microlensing constraints, and too large for Hawking radiation to have been significant in their lifetimes. There were a handful of

mechanisms proposed to constrain this range, but all have been dismissed with various issues. When we examine cosmological black holes, however, we will explore this space again, but in the context of Hawking radiation.

One of the possible constraints in this mass range are the so-called *femtolensing* constraints of gamma-ray bursts [588–592]. These observations are limited by finite source-size and diffraction effects— finite source-size effects refer to the shortening of the microlensing event time, while the diffraction effects occurs because the Schwarzschild radius of the black hole is comparable to the wavelength of the detected optical light.

There are also PBH capture scenarios in this mass range. One constraint involved stellar capture, where the PBH would settle in the interior and destroy the star [593–596]. However, detailed calculations found that the event rate was insufficient to place any constraints [588]. A second involved ignition of white dwarfs [597], but detailed calculation again showed that PBHs in this mass range were probably not able to cause viable explosions, both theoretically and at the necessary rate for constraint [588].

Evaporating black holes

For the smallest masses, it begins to become more convenient to use grams instead of solar-mass units. For black holes with mass $m \lesssim 10^{17} g = 10^{-17} M_\odot$, Hawking evaporation becomes relevant. There are a large number of interesting constraints for these black holes, many of which occur in the early universe.

Firstly, there is the *stability constraint*, which claims that PBHs smaller than the critical mass $m_{\rm crit}$ which completely evaporates with the lifetime of the universe cannot be the dark matter [5, 14, 422]. This constraint must be carefully worded, but I have not seen it layed out explicitly in the literature (partially prompting the writing of Ref. [422] with Markus Mosbech). Clearly PBHs which *form* with masses smaller than $m_{\rm crit}$ cannot comprise the dark matter today. But, if PBHs formed with masses only slightly larger than $m_{\rm crit}$, perhaps they could comprise the dark matter *today*, but with very small masses. However, it can be argued that since the population today is extremely sensitive to the precise initial mass, that we have a kind of fine-tuning or stability problem, which makes it untenable to have such a monochromatic population today.

We can also place constraints on the observation of Hawking radiation itself— whether from positrons in the galactic center [598, 599], electrons and positrons from Voyager 1 [600], or galactic and extragalactic γ-rays [601–603]. In addition, exten-

sive Hawking radiation would imprint on the CMB anisotropies [538, 604], the 21 cm line [605], and if the PBHs were sufficiently small, might also affect BBN processes [601]. As noted above, we will return to the issue of Hawking radiation for cosmological black holes.

Extended mass distributions

Finally, I will address the question of extended mass distributions [458, 470–473]. Perhaps the most generic distribution is the lognormal distribution, which can be used to approximate most physically realistic PBH formation theories [473, 524, 585, 598–600, 606]. Other distributions considered include power-laws of various kinds and critical collapse functions with cutoffs at particular mass scales.

While one might naively think that it would be easy to concoct a distribution which skates under all the monochromatic constraints, the situation is not so straightforward. There is a careful process for 'transferring' monochromatic distribution constraints to arbitrary extended distributions, and in many cases we do not find the constraints to be significantly lifted [470, 473]. However, in the evaporating black hole regime, it is not appropriate to merely recast the monochromatic constraints. This is because the evaporation of PBHs significantly alters the distribution between formation and today, so a more careful analysis must be undertaken for these constraints [422, 473].

The layperson box: primordial black holes Previous — Next

I have been going on about how we are searching for some kind of invisible object to be the dark matter. Are black holes not ideal? Not only are they dark, they are relatively tiny—a black hole with the Earth's mass would be ~ 2 cm wide, while a solar mass black hole would be only ~ 1 km.

And not only that—we have actually, definitely, seen black holes. In the center of galaxies there are supermassive black holes ($\sim 10^8$ solar masses), which we see in our galaxy by the movement of stars around 'nothing', and in other galaxies by the unimaginably energetic beams of radiation emitted when gas falls onto these black holes. We have even taken a 'picture' of two of these black holes (including our own central one in the Milky Way), using radio waves collected from different points across the Earth.

We have also detected smaller black holes, only $\sim 5 - 100$ times the mass of the sun. Just recently, we saw one wandering aimlessly through our galaxy, based on the motion of stars near it. And most incredibly of all, we have detected 'gravitional waves'—literal ripples in the fabric of spacetime—that form when black holes collide. We have detected more than a hundred of these collisions so far, a truly momentous piece of science in only the last seven years.

But black holes produced by stars collapsing *cannot* be dark matter, because dark matter needs to have existed from near the beginning of the universe, and stars only showed up later. How else can you make black holes? We must look to the very early universe, when everything was just a hot, dense soup of particles. Just by random accident, sometimes this soup could fluctuate in density. If it accidentally had a massive enough fluctuation, it might have been possible to suddenly collapse to a black hole. These hypothetical ancient black holes are called 'primordial' black holes.

Primordial black holes were one of the earliest known dark matter candidates and they are still going strong today. We know they cannot be too small, because black holes slowly evaporate—if they were too small, they would not still exist. We also know they cannot be too big, or they either would not fit into galaxies, or otherwise would mess up the motions of stars in them. If they are roughly moon-to-Earth-mass, we also can rule them out, because there would be so many of them that they would pass in front of distant stars, making them briefly brighter, like tiny magnifying glasses. Medium sized black holes ($100 - 1000$ solar masses) we can also probably rule out since gas in outer space would fall onto them, heating up and producing radiation. Finally, black holes that are the right size to seen by gravitational waves ($10 - 100$ solar masses) would be ruled out, because they would merge often, leading to more gravitational wave events than observed.

That is only some of the many constraints clever scientists have devised, but there are still plenty of available masses these black holes can be. Then, there is the (probably likely) possibility that the black holes come in a whole spectrum of different sizes, in which case it becomes yet easier for these black holes to be a good portion of the dark matter, if not all. It is a relatively exciting time to be working on black hole physics...

3
Cosmological black holes

Time is a flat circle.

—Rust Cohle, *True Detective*

In the early universe, primordial black holes are born within the hot dense plasma of the cosmological fluid, where the Hubble horizon may not be actually very far away from the black hole horizon. Meanwhile, the Schwarzschild metric—universally used for modelling black holes (along with the Kerr metric)—is embedded in empty, asymptotically flat space. These conditions are clearly not correct for black holes in the early universe. Two questions arise: how do we properly model black holes in the early universe? And does treating them differently affect their phenomenology, particularly with respect to dark matter?

In fact, these questions go all the way back to the historical origins of primordial black holes. The birth of the field arguably began when the Soviet physicist I.D. Novikov noted in 1964 that objects which are seeded by overdensities in the early universe could in fact be black holes (and not just stars and galaxies) [390]. His idea was expounded on in Ref. [391], by Zel'dovich and Novikov—this paper is often considered the first PBH paper. However, it is not always recorded that this paper attempted to discredit the PBH hypothesis. Using a simple analysis of accretion onto these objects, they found that the black holes grew catastrophically large, clearly in contrast to the observed composition of the universe.

This problem was solved by Carr and Hawking in 1974 [366]. They found that when they explicitly considered an alternative black hole metric which was asymptotically-FLRW, the runaway accretion problem was solved. They did not offer a specific metric, but rather provided a detailed mathematical proof for any such metric. This paper truly began the field of primordial black holes—now that they were established as being physically possible, the study of their properties and cosmological implications could begin.

Interestingly, despite the great importance of Carr and Hawking's work, the study of PBH phenomenology eventually proceeded on with Schwarzschild black holes—perhaps

under the assumption that they could be safely stitched into an FLRW background sufficiently far away from both the black hole and cosmological horizon. As we will see, such a solution is actually far from trivial. Regardless, many studies of PBH dark matter more or less approximate the PBHs as Schwarzschild (or Kerr) objects.

This is not to say that cosmological black hole solutions were not explored. In fact, there is a wealth of both contemporary and historical literature [99, 415, 607–610] on these issues, which I will explore further in the following sections. However, for various reasons, it has not been particularly popular to examine PBH phenomenology using this variety of cosmological black hole metrics. Perhaps the general wisdom is that it would be unlikely to find major consequences, since the black holes will be approximately Schwarzschild-like up close and FLRW-like far away.

It is also the case that it is very hard to find a cosmological metric with no faults—almost all the examples in the literature either:

- contain strange singularities (violating various energy conditions),
- require a specific background (usually, dust-like, as in a matter-dominated universe), or
- have somewhat confusing or ill-defined physical interpretation.

However, it is past time that these things are carefully checked. Primordial black holes are again in the spotlight of the dark matter community, partly due to the wealth of recent black hole observations, and partly due to the distinct lack of wealth of any particle dark matter observations. In this chapter, I will begin by generally describing some details of cosmological black holes, before looking at a selection of the more well-studied possible metrics.

For all our searching, we only found one metric, the Thakurta metric [611], which did not have the first two of the above faults—it contains no singularities, and works perfectly fine during radiation-domination. The most important phenomenological aspect of the Thakurta metric is its time-dependent Misner–Sharp mass, which is roughly proportional to the cosmological scale factor. We found that this feature radically reshapes the landscape of PBH dark matter. In later sections I will show in detail how both the gravitational wave bounds [6] and the Hawking radiation bounds [5] are significantly altered. Spoilers: the former is removed entirely, while the latter is extended across the entire asteroid-mass range.

The third fault, of being generally confusing, the Thakurta metric certainly possesses. The three papers I have written with my collaborators about Thakurta black holes [1, 5, 6] have received quite a bit of attention from the community, for better or worse [4, 612–614]. I would say that this criticism has broadly ranged from relatively unfounded, to hard-to-say-exactly, to fair. I will do my best to elucidate and explain these features and criticisms as they arise. Regardless, there appears to be a bit of inertia in the community against the application of cosmological black holes, and it is hard to judge what proportion of this push-back is scientific or sociological [615].

It very well may be the case, ultimately, that the Thakurta metric is not a correct PBH description—there are certainly many more questions left to answer, and not many people trying to answer them currently. However, I hope that, at the least, our work with the Thakurta metric encourages the PBH-dark matter community to keep pursuing new cosmological metrics, and investigating their impact on PBH phenomenology. It is heartening at least that the last couple years appear to have seen an increase in the number of papers considering such effects [616–618], whether or not that is due to us (or perhaps, the publicity of the various criticism and comment papers related to our work).

3.1 General formulation

Because cosmological black holes are necessarily coupled to an expanding universe, it is perhaps not that surprising that many of them are dynamical, as opposed to stationary, solutions. In fact, almost all of the solutions we will look at in later sections are dynamical in some way—even the solution which merely 'stitches' a Schwarzschild black hole into the universe requires a kind of dynamical stitching [608, 619]. Actually, it is not just for cosmological reasons that we care about dynamical black hole solutions. Black holes which grow by accretion, or shrink by Hawking radiation, are really properly described only by a dynamical solution.

Understanding dynamical black holes adds an extra level of difficulty and ambiguity, compared to static or stationary black hole solutions. In particular, these black holes lack an asymptotically-timelike Killing vector field [620]. Killing fields are vector fields which preserve the metric, so that flow with respect to the Killing vector expresses some symmetry of the spacetime. For static black holes, the time vector is a Killing vector, and points in the forward direction of time unambiguously. For dynamical black holes, which generically lack this Killing vector, however, we lack a 'preferred' time coordinate to understand black hole properties such as surface gravity.

Kodama time

The situation was partially remedied when Kodama [620, 621] wrote down a natural, divergence free vector field, later named the *Kodama vector*. This vector field exists in any spherically symmetric spacetime and is given by,

$$k^a \equiv \epsilon_\perp^{ab} \nabla_b r \ , \tag{3.1}$$

which lies in the $(1+1)$-dimensional radial-temporal plane, where $\epsilon_\perp^{ab}$ is the Levi-Civita tensor in this plane. Although not a particularly intuitive definition, it can be shown that,

$$k^a \nabla_a r = 0 \ , \tag{3.2}$$

so that the Kodama vector does indeed define a natural timelike direction. This does not specify a naturally preferred time coordinate in and of itself, but the argument was taken further in Ref. [621], where the Clebsch decomposition theorem was used to assert that there are two unique factors, α and β, such that the Kodama covector (written somewhat

unusually as $k^\flat$), takes the form,

$$k^\flat = \alpha \mathrm{d}\beta \ . \tag{3.3}$$

In the exterior spacetime region where $\mathrm{d}r$ is spacelike, we know that the Kodama vector and covector are certainly timelike, guaranteeing that $\mathrm{d}\beta$ is timelike itself. The natural conclusion is that β is therefore the preferred time coordinate, called the *Kodama time*, τ. In fact, it turns out that this choice is the unique time coordinate so that integral curves of the Kodama vector coincide with ∂_τ.

Spherically symmetric metrics

Perhaps even more tellingly, using this time coordinate, the cross-term in the metric is naturally zero, so that we are always able to write a spherically symmetric metric as,

$$ds^2 = e^{-2\phi(R,\tau)} F(R,\tau)\, d\tau^2 - \frac{dR^2}{F(R,\tau)} - R^2 d\Omega_2 \ , \tag{3.4}$$

where the Schwarzschild-like factor can be written in terms of the Misner–Sharp mass Eq. 2.2 as,

$$F(R,t) = \left(1 - \frac{2Gm_{MS}}{R}\right) \ . \tag{3.5}$$

The factor $e^{-\phi}$ is related to the Kodama time-translation vector $\mathcal{T}$ and Kodama vector k by $e^{-\phi} = ||\mathcal{T}||/||k||$. It is not exactly a coincidence that this choice of Schwarzschild-like coordinates conveniently uses the Misner–Sharp mass—in fact, the mass quantity above *automatically* gains the interpretation of being the Misner–Sharp mass when written in these coordinates.

The metric in Eq. 3.4 provides a compelling and natural way to describe dynamical black holes. However, it is sometimes useful to work in Painlevé–Gullstrand (PG) coordinates instead, since they are better suited for future horizons. The line element in these coordinates is given by,

$$\mathrm{d}s^2 = \frac{e^{-2\phi}}{\alpha^2} F\, \mathrm{d}\tilde{\tau}^2 - 2\frac{e^{-\phi}}{\alpha}\sqrt{1-F}\mathrm{d}\tilde{\tau}\mathrm{d}R - \mathrm{d}R^2 - R^2 \mathrm{d}\Omega_2 \ , \tag{3.6}$$

where the new time coordinate $\tilde{\tau}$ is defined by,

$$\mathrm{d}\tilde{\tau} = \alpha \mathrm{d}\tau + \alpha e^{\phi} \frac{\sqrt{1-F}}{F} \mathrm{d}R \,. \tag{3.7}$$

In the above, we are free to choose the function $\alpha > 0$ so that the PG time direction coincides with the Kodama time direction. When changing coordinates we have to satisfy the integrability condition—that the order of two partial derivatives does not matter. Ensuring this holds for the radial and temporal derivatives gives the condition,

$$\frac{\partial}{\partial R}(\alpha) = \frac{\partial}{\partial \tau}\left(\alpha e^{\phi} \frac{\sqrt{1-F}}{F}\right) \,. \tag{3.8}$$

For simplicity, I will follow the conventions of Refs. [415, 417], and define $c(R,t) = e^{-\phi}/\alpha$. Then the PG metric can be written as,

$$\mathrm{d}s^2 = c^2 F \, \mathrm{d}\tilde{\tau}^2 - c\sqrt{1-F}\mathrm{d}\tilde{\tau}\mathrm{d}R - \mathrm{d}R^2 - R^2 \mathrm{d}\Omega_2 \,. \tag{3.9}$$

Although this section may seem like technical formalities, understanding the Kodama foliation is essential for studying dynamical black holes. It feels somewhat simple written out here, but it took quite a bit of work and confusion during my research before I began to put some of these pieces together. Still, there is much that is either ambiguously-defined or not well-understood for dynamical black holes—both by me and the wider PBH community. With the exciting prospect of unanswered questions in mind, let us now look more concretely at some cosmological PBH metrics.

3.2 Cosmological PBH metrics

There are a number of cosmological black hole candidates, so I will only give a brief summary of some of the more prominent metrics (see, e.g. [415] for a more in-depth summary).

Einstein–Strauss spacetime

The Einstein–Strauss (Swiss cheese vacuole) solution [608, 619, 622] is perhaps the simplest cosmological black hole spacetime—here, a Schwarzschild solution is matched to an FLRW background along a radially-comoving hypersurface. The matching surface must be dynamical, so that the volume 'cut out' by the black hole spacetime has matching mass with some volume of cosmological fluid. Although this represents a somewhat *ad hoc* solution to the cosmological BH problem, it is not uncommon in physics to describe a locally decoupled system independently to its background, assuming that it is somehow 'stitched' into the background with minimal local effects.

The main drawback of this solution, however, comes from the matching conditions along the hypersurface, which require a specific kind of conservation of stress-energy. Because the interior solution is a vacuum, we have to enforce zero pressure on the exterior of the stitching and so similarly everywhere in the FLRW spacetime [619]. As a result, the Swiss cheese vacuole is restricted to matter-dominated FLRW backgrounds. While this solution is presumably useful in the later universe, it can not be readily applied during radiation-domination, when PBHs form, and we must be careful not to asssume that some other roughly-similar solution even exists when intuiting the behavior of these early-universe cosmological black holes.

Lemaître-Tolman-Bondi spacetimes

The Lemaître-Tolman-Bondi class of metrics [609, 610, 623] is another dynamical black hole candidate. This metric describes a spherical dust cloud which is collapsing under gravity. These spacetimes are a relatively broad class, and so the realization depends on specific model choices, which can include cosmological backgrounds. There are a number of difficulties with this class of metrics—for one, there is again the issue that it applies to dust backgrounds. Secondly, the solutions are plagued with shell-crossing singularities related to the inhomogeneous dust collapse. It turns out that for physically

realistic collapse scenarios, enforcing the conditions which remove these singularities also make the black hole singularity locally visible, violating the cosmic censorship hypothesis discussed in Sec. 2.1.6.

McVittie spacetimes

The McVittie class of metrics [607, 624] is another old and well-studied candidate, and actually contains as a special case the oldest known cosmological black hole, the Schwarzschild-de Sitter/ Kottler metric [415, 625–628]. The McVittie solution is a rather generic solution for a spherically symmetric inhomogeneity embedded in an FLRW universe. An important assumption for this solution is that there is no accretion or energy flow onto the central inhomogeneity. This spacetime contains a stationary singularity at $R = 2m$, where R is the areal (or comoving) radius—however, this singularity is spacelike, in contrast with the standard black hole horizon. Since the metric is sourced by a perfect fluid, this leads to a divergent pressure at the horizon, which we generally would prefer to avoid.

To remedy this, the *generalized McVittie solutions* [624] remove the no-accretion condition, in order to create a more physically reasonable stress-energy source. One prominent example of these metrics is the Sultana–Dyer solution [629, 630], which is found by conformally transforming the Schwarzschild metric like $g_{\mu\nu} \to \Omega^2 g_{\mu\nu}$, with conformal factor [631],

$$\Omega = a(t,r) = \left(t + 2m \ln\left|\frac{r}{2m} - 1\right|\right)^2 , \tag{3.10}$$

so that the timelike Killing field in the Schwarzschild metric becomes a conformal Killing field. Notably, because the metric was originally found by a conformal transformation of the Schwarzschild metric in isotropic coordinates, the conformal factor is both temporally and radially dependent, as seen above [631]. As a result, the stress-energy source of this metric is forced to be a mixture of two non-interacting perfect fluids—a null dust and a massive dust. One of the flaws of the Sultana–Dyer metric is that it is therefore only valid in matter-dominated eras. Another potential issue with the Sultana–Dyer metric is that there is superluminal dust flow near the the apparent horizon, although that may not be a serious flaw and rather a consequence of the over-simplicity of the metric [630].

The generalized McVittie metrics are a rather broad class of metrics, depending on the particular choice of stress-energy source. However, there is one particularly interesting

special case—the late-time attractor of these solutions [410, 415, 632]. This spacetime is known as the Thakurta metric [611], and also happens to be the General Relativistic limit of a class of exact solutions of Brans-Dicke gravity with a cosmological fluid [633], as well as 'cuscuton gravity' [631, 634, 635] and shape dynamics [631, 636]. Because this particular metric is so important to this thesis, it deserves its own section.

3.3 The Thakurta metric

We chose to study the Thakurta metric [611] because it satisfies two of the three previously-mentioned desirable qualities for a cosmological black hole metric—it has no pathologies (spacelike singularities on the horizon, negative energy densities, or naked singularities), and it is perfectly valid in radiation (and matter) domination. The third quality, of being not-too-confusing, it perhaps satisfies less easily, but I am not sure we can hold out much hope for that qualification with any decent spacetime.

For these reasons we chose the Thakurta metric for further investigation, even though it may not be 'perfect'. One might consider it merely a toy model, or justifiably find its flaws insurmountable. However, we have yet to find any other spacetime which could possibly take its place. If you know of one, please let me know. With such warnings behind us, let us then delve deeper into the Thakurta metric.

The non-rotating Thakurta metric was first proposed very similarly to the Sultana-Dyer metric [629]. The spacetime interval for the Thakurta metric is essentially that of a conformal-time Schwarzschild metric, multiplied by the scale factor:

$$\mathrm{d}s^2 = a^2(\eta)\left[f(r)\mathrm{d}\eta^2 - \frac{\mathrm{d}r^2}{f(r)} - r^2\mathrm{d}\Omega^2\right] , \tag{3.11}$$

where,

$$f(r) = 1 - \frac{2Gm}{r} . \tag{3.12}$$

We should carefully note that this is *not* the same as the conformal transformation that leads to the Sultana–Dyer metric, although it is certainly very similar. The differences can be seen clearly in e.g. Refs. [631, 637]. In contrast with the Sultana–Dyer metric, the Thakurta metric conformal factor is only time-dependent. This metric can then be rewritten in terms of *cosmological time* t [1, 638] and areal radius $R = a(t)r$ as,

$$\mathrm{d}s^2 = f(R,t)\left(1 - \frac{H^2R^2}{f(r,t)^2}\right)\mathrm{d}t^2 + \frac{2HR}{f(R,t)}\mathrm{d}t\mathrm{d}R - \frac{\mathrm{d}R^2}{f(R,t)} - R^2\mathrm{d}\Omega^2 , \tag{3.13}$$

where $f(R,t) = f(r)$ is simply,

$$f(R,t) = 1 - \frac{2Gma(t)}{R} . \tag{3.14}$$

The two limits of this spacetime are relatively straightforward to see. As $f(r \to \infty)$, the metric becomes the FLRW metric, while the Schwarzschild metric is recovered when $H \to 0$.

The stress-energy source

The Thakurta metric is sourced by an imperfect fluid with a radial heat flow, but no radial fluid accretion (in other words, the fluid four-velocity is orthogonal to the heat flow four-vector). The stress energy source is generically written as,

$$\begin{aligned} T_{\mu\nu} &= (\rho + P)u_\mu u_\nu + g_{\mu\nu}P + q_{(\mu}u_{\nu)} \ , \\ q_\mu &= (0, q, 0, 0), \\ u_\mu &= (u, 0, 0, 0) \ . \end{aligned} \tag{3.15}$$

That the Thakurta metric has no accretion is a subtle but important point—no matter is actually flowing onto the central imhomogeneity. Rather, the radial energy flow can presumably be thought of as a temperature gradient towards the black hole. As we will see shortly, the Thakurta black hole has a growing Misner–Sharp mass—however, this mass does not grow from any actual matter accretion, but is a purely geometrical effect.

I belabor the no-accretion point as it is a common point of confusion or criticism regarding the Thakurta metric. In fact, the Thakurta metric is derived from the generalized McVittie metric by requiring the central 'mass' parameter m to be constant [415]. The fact that the Misner–Sharp mass grows without accretion of matter does not violate any conservation laws, because we must be very careful with our intuition regarding such laws in General Relativity—the growth of the FLRW universe also violates conservation of energy, because no such conservation law is applicable to the spacetime. See appendix B for further musings on this subject.

Apparent horizons

One relatively remarkable feature of the Thakurta metric (compared to many cosmological black holes), is the analytical tractability of the apparent horizons. The spacetime contains two apparent horizons, where $f(R, t) = HR$, which are interpreted as the cosmological

and black hole horizons, respectively:

$$R_{\rm C} = \frac{1}{2H}\left(1+\sqrt{1-8HGma}\right) \tag{3.16}$$

$$R_{\rm BH} = \frac{1}{2H}\left(1-\sqrt{1-8HGma}\right) \ . \tag{3.17}$$

In the static limit where $H \to 0$, the cosmological horizon approaches infinity and the black hole horizon approaches the Schwarzschild event horizon with $a = 1$. Meanwhile, when $m \to 0$, the black hole horizon vanishes and the cosmological horizon approaches $1/H$ as expected. It will be useful to define the small parameter,

$$\delta \equiv HGma \ < \frac{1}{8} \ , \tag{3.18}$$

since the factor in the square root is always positive after the black holes have formed. To zeroth order in this parameter, in fact, the black hole horizon becomes $R_{\rm BH} \to 2Gma$, while the cosmological horizon approaches $1/H$. It is also important to note that while there is still the spacelike singularity at $f(r) = 0$ which plagues the McVittie metric, this surface is *always* below the apparent horizon and so is not physically consequential [4].

The Misner–Sharp mass

While this analysis follows the original definition of the Thakurta metric in Eq. 3.11, it is more useful to write the metric in the Kodama time foliation [1, 620, 621] of Eq. 3.4:

$$ds^2 = e^{-2\phi(R,\tau)}F(R,\tau)\,d\tau^2 - \frac{dR^2}{F(R,\tau)} - R^2 {\rm d}\Omega^2 \ , \tag{3.19}$$

with,

$$F(R,t) = \left(1 - \frac{2Gm_{MS}}{R}\right) \ . \tag{3.20}$$

For the Thakurta metric, the Misner–Sharp mass in the above is given by,

$$m_{\rm MS} = ma + \frac{H^2R^3}{2Gf(R,t)} \ . \tag{3.21}$$

It is important to note that near the black hole, the first term always dominates, even though $f(R,t)$ blows up at small distances—only once $R < R_{\rm BH}$ does the second term again

dominate. The horizons satisfy,

$$R_{\rm h} = 2Gm_{\rm MS}(R_h) \ . \tag{3.22}$$

We can find the explicit form of the energy density [631,638] using the flow vector $v_\mu \equiv \left(-\sqrt{f(R,t)}, 0, 0, 0\right)$ as follows:

$$\rho(R,t) \equiv T_{\mu\nu} v^\mu v^\nu = \frac{3H^2}{8\pi f(R,t)} \ . \tag{3.23}$$

Then we can rewrite the Misner–Sharp mass as,

$$m_{\rm MS} = ma + \frac{4\pi}{3}\rho(R,t)R^3 \ . \tag{3.24}$$

However, we must be very careful with the proper interpretation of these terms, which appear to describe a central point mass on the left and the mass of matter surrounding it on the right. However, the terms are not neatly separable into these interpretations—for instance, we can see that the definition of the energy density in fact includes the central m parameter. The *full* Misner–Sharp mass is the quasi-local mass of the Thakurta black hole.

The transformation between the 'cosmological time' of Eq. 3.13 and the Kodama time is given by,

$$\mathrm{d}\tau = e^{\phi(R,t)}\mathrm{d}t + e^{\phi(R,t)}\frac{HR}{f(R,t)}\frac{\mathrm{d}R}{F(R,t)} \ , \tag{3.25}$$

and must satisfy the integrability condition,

$$\frac{\partial}{\partial R}\left(e^{\phi(R,t)}\right) = \frac{\partial}{\partial t}\left(e^{\phi(R,t)}\frac{HR}{f(R,t)F(R,t)}\right) \ . \tag{3.26}$$

Unfortunately, this condition is not well-defined at the horizon $F = 0$, so it is not straightforward to evaluate the factor e^ϕ there. Indeed, at the time of writing, I have devoted a significant amount of effort to evaluating this factor, via the Einstein equations in various forms, and have not managed to find a clear answer (often, the factor divides out). The explicit form of this factor is not important for much of this chapter, at least until Sec. 3.5, so I will leave it for the moment. It is useful to note, however, that the Kodama time approximately coincides with the cosmological time for observers which are

distant from both the black hole and the cosmological horizon (who might be considered the natural observers to care about). This fact makes it generally easier to compare, for example, the Hawking radiation seen by a distant observer for stationary and dynamic spacetimes.

3.3.1 The Thakurta horizons

First, we should check that the Thakurta horizons are in fact black hole horizons, as discussed in Sec. 2.1.4. Determining this is somewhat nontrivial, as was pointed out in Ref. [614], where it was found that the Thakurta metric appeared to have a white hole horizon. The issue here is that while the expansion scalars Eq. 2.4 are coordinate independent, the apparent horizon is foliation dependent. As a result, the expansion scalars themselves—and therefore, the nature of the horizons—is foliation dependent. This subtle fact was indeed already known [639] but seems to be sufficiently not well-understood in the community that we decided to write Ref. [1] specifically addressing this for the Thakurta metric.

First let us look at the horizons in the cosmological foliation (i.e., the metric Eq. 3.11). The null radial geodesics, where $\mathrm{d}s^2 = \mathrm{d}\Omega^2 = 0$, are given by,

$$\frac{dr}{dt} = \pm\frac{f}{a} \ , \tag{3.27}$$

so that the tangent vectors of Eq. 2.3 are given by,

$$\begin{aligned} \ell^\mu =& (1, f/a, 0, 0) \ , \\ n^\mu =& \frac{1}{2f}\,(1, -f/a, 0, 0) \ . \end{aligned} \tag{3.28}$$

Then it is relatively straightforward to calculate the expansion scalars as,

$$\begin{aligned} \theta_\ell =& \frac{2}{R}\,(HR + f(R,t)) \ , \\ \theta_n =& \frac{1}{Rf(R,t)}\,(HR - f(R,t)) \ . \end{aligned} \tag{3.29}$$

Then at the Thakurta horizons, where $f(R,t) = HR$, the ingoing expansion scalar $\theta_n = 0$ while the outgoing scalar satisfies $\theta_l > 0$. These observers therefore see past horizons—light rays exit the horizon and never return. We should also check that it is a trapping

horizon:

$$\begin{aligned}\mathcal{L}_\ell\theta_n|_{R=R_h} &= \left(\partial_t + \frac{f}{a}\partial_r\right)\theta_n\Big|_{R=R_h} \\ &= -\frac{1}{HR}\left(\frac{4\pi G}{3}f(\rho-3p) + \frac{H\left(1+HR\right)}{R}\right)\Bigg|_{R=R_h} ,\end{aligned} \tag{3.30}$$

where the second equality is derived using the trace of the Einstein equations,

$$\dot{H} + 2H^2 = -\frac{8\pi G}{3}f(\rho - 3P) , \tag{3.31}$$

in terms of the density ρ and pressure P of the cosmological fluid at the horizon. In radiation- or matter-domination, $\rho > 3P$ so that the Lie derivative above is negative definite as would be required for a trapping horizon. We conclude then, that the cosmological observers view both horizons as white holes or cosmological horizons.

Now let us examine the Kodama observer. Actually, it is generally more suitable to work in PG coordinates, using the metric Eq. 3.9. Although the PG time $\tilde{\tau}$ is not strictly the Kodama time, the coordinate singularity at $F = 0$ is removed in these coordinates, which is more convenient for calculations at future horizons (it is anyways possible to show that the same conclusions hold for the Kodama time metric Eq. 3.4, but the maths is quite a bit uglier). The null radial geodesics in the PG metric satisfy,

$$\frac{dR}{d\tilde{\tau}} = -c\sqrt{1-F} \pm c , \tag{3.32}$$

so that the tangent vectors are,

$$\begin{aligned}\ell^\mu &= \frac{1}{c}\left(1, c - c\sqrt{1-F}, 0, 0\right) , \\ n^\mu &= \frac{1}{2c}\left(1, -c - c\sqrt{1-F}, 0, 0\right) .\end{aligned} \tag{3.33}$$

Again we calculate the expansion scalars, finding,

$$\begin{aligned}\theta_\ell &= \frac{2}{R}\left(1 - \sqrt{1-F}\right) , \\ \theta_n &= -\frac{1}{R}\left(1 + \sqrt{1-F}\right) .\end{aligned} \tag{3.34}$$

In contrast to the cosmological observers, we have $\theta_\ell = 0$ and $\theta_n < 0$ at the Thakurta

horizons $F = 0$. Thus the horizons are future apparent horizons. In addition, we have,

$$\begin{aligned}\mathcal{L}_n\theta_\ell|_{R=R_\mathrm{h}} &= -\frac{1}{2R}\left(2F' - \frac{\dot{F}}{c}\right)\bigg|_{R=R_\mathrm{h}} \\ &= -4\pi G\frac{\dot{R}_\mathrm{h} + 2c}{\dot{R}_\mathrm{h}}\, T_{\mu\nu}\ell^\mu\ell^\nu\bigg|_{R=R_\mathrm{h}} \,. \end{aligned} \tag{3.35}$$

In the above, we have used the fact that $F(R_\mathrm{h}(\tilde{\tau}), \tilde{\tau}) = 0$ is true for any $\tilde{\tau}$ to assert that,

$$\frac{\mathrm{d}F}{\mathrm{d}\tilde{\tau}} = \dot{F} + \dot{R}_\mathrm{h}F' = 0 \,, \tag{3.36}$$

as well as the Einstein equations evaluated at the horizon:

$$-\frac{\dot{F}}{cR_\mathrm{h}} = G_{\mu\nu}\ell^\mu\ell^\nu = 8\pi G T_{\mu\nu}\ell^\mu\ell^\nu \,. \tag{3.37}$$

Finally, the rate of change of the horizon is given by,

$$\begin{aligned}\dot{R}_\mathrm{h} &= \frac{2G\dot{m}_{MS}}{1 - 2Gm'_{MS}}\bigg|_{R=R_\mathrm{h}} \\ &= 2G\frac{4\pi R_\mathrm{h}^2 T_{\mu\nu}\ell^\mu\ell^\nu}{1 - 2Gm'_{MS}}\bigg|_{R=R_\mathrm{h}} \,, \end{aligned} \tag{3.38}$$

using the fact that $\dot{m}_{MS} = 4\pi R_\mathrm{h}^2 T_{\mu\nu}\ell^\mu\ell^\nu|_{R=R_\mathrm{h}}$. As long as we are at distances $R \geq R_\mathrm{h}$, we have that $1 - 2Gm'_\mathrm{MS}$, and we can use the null energy condition,

$$T_{\mu\nu}\ell^\mu\ell^\nu \geq 0 \,, \tag{3.39}$$

to find that the Thakurta horizon is always increasing with respect to the PG time. Since $c > 0$, we therefore satisfy the condition $\mathcal{L}_n\theta_\ell|_{R=R_h} < 0$, so that the Thakurta horizon is indeed seen as a future outer trapped horizon.

Although the maths is clear, it may be helpful to have a physical picture in mind. How can an object appear to have radically different horizons in two different foliations? I believe, however, that physicists are already somewhat comfortable with observational discrepancies in General Relativity—infalling observers on the Schwarzschild horizon observe a radically different black hole to stationary observers at the same location. Perceiving the Thakurta black hole with two different time directions is perhaps, then, not

a radical leap in thinking from perceiving it with two different observer velocities.

I suspect one might even be able to picture it the following way—in the following instant, the cosmological time observer at the horizon will find themselves within the apparent horizon (similarly to how a cosmological horizon works), so necessarily light rays will be seen leaving the horizon. Meanwhile, while the Kodama observer is not exactly comoving, the Kodama time direction 'includes' a radial component so that the observer is not instantaneously behind the horizon. Then they do indeed observe a black hole horizon, as light rays do not return from the apparent horizon. This understanding is probably not technically accurate, but I have found it useful for wrapping my head around these issues.

3.3.2 Thakurta spacetimes as PBHs

Having the correct horizon structure is a necessary prerequisite for being a black hole, but we really need to more carefully model how these objects can be PBHs. Technically, the Thakurta black hole is 'eternal'—it begins with a big bang, and ends at future infinity. However, we are interested in PBHs which form at a particular epoch and eventually become compact objects embedded in local structures like galaxies and clusters. We therefore must include additional physically-motivated assumptions regarding how the Thakurta black holes act as PBHs.

Firstly, we assume that their formation is more-or-less identical to the standard PBH formation from horizon-sized anisotropies in the early universe, where the mass of the overdensity is the same size as the horizon mass. It might be understandable to ask whether we should require the Misner–Sharp mass of the overdensity to match the Misner–Sharp horizon mass, but we should remember that the black hole is (perhaps redundantly) not yet formed, before it is formed—so we cannot use the Thakurta Misner–Sharp mass for this purpose. Then we will assume that an overdensity of mass $\sim m$ collapses to form the Thakurta spacetime Eq. 3.11, giving us a useful physical interpretation of the parameter m. Throughout we will refer to this mass as the *physical* mass of the Thakurta black hole, implying that it represents both the original anisotropy mass as well as the mass of the black hole today.

This leads us to the other end of the Thakurta spacetime's life. The Thakurta black hole cannot live in the background of the imperfect fluid forever—eventually, in the later universe, it will find its local environment dominated by something else. That could be

another black hole (in which case, the two might form a binary which is decoupled from the Hubble flow), or it could be the formation of structure such as galaxies (in which case the PBH will decouple from the Hubble flow and virialize with the galaxy). Regardless, the black hole will no longer be suitably described by the Thakurta spacetime at this point. We therefore assume that the black hole solution at decoupling should be interpolated into a solution more like the Schwarzschild metric, or perhaps more carefully, the Einstein–Strauss spacetime. In this case, we assume the mass of the now-decoupled black hole would merely be the physical mass m (the Misner–Sharp and ADM masses coincide for the Schwarzschild black hole).

The black hole described above, therefore, is not a pure Thakurta black hole—it must transition from an anisotropy in FLRW space, to a Thakurta spacetime, then to a local Schwarzschild spacetime. We do not have the full metric describing such an object, and it would probably be rather complicated if we did. In that sense, our black hole model is something of a toy model, in order to demonstrate the possible phenomenological implications of real cosmological black holes. While it would be great to have a more explicit model, we will for now make do with what we have.

3.3.3 Criticisms of Thakurta black holes

I have already touched on a number of the common issues found in the literature regarding Thakurta black holes [612–614]. I will do my best to summarize these flaws here, and give my own opinion on them. In some cases, there is a clear solution, or the issue is merely a misconception—in other cases, I can only offer my thoughts and not a clear resolution.

Accretion

The nature of the horizons was already discussed at length, so I will not rehash that here. I also discussed the fact that the Thakurta metric has no mass accretion, and merely a radial heat flow (see also appendix B for further discussion). It would be useful to briefly touch on accretion here, however. Some arguments [612, 613] dismissing the physicality of the Thakurta metric insist that there is no way that the cosmological fluid, composed mainly of Standard Model particles if $f_{\mathrm{PBH}} = 1$, can accrete onto the black holes efficiently enough. On top of that, the merging of these PBHs can also certainly not make up the mass growth. I certainly agree that the Misner–Sharp mass growth does not resemble in any way the usual physics of accretion, and that accretion processes (or mergers) could not explain its

growth—because, again, the Thakurta black hole does not grow by the standard accretion of matter as represented by an ideal fluid.

If we wanted, we *could* add accretion effects to the Thakurta black hole in somewhat the same way that we treat accretion for the Schwarzschild black hole—generally, we do not construct a new, expanding black hole to model accretion, but rather impose accretion in an adiabatic way onto a static Schwarzschild solution, although it would be perhaps preferable to do the former. For Thakurta black holes, then, we could either impose some kind of equivalent Bondi accretion, or construct a new metric which explicitly contains a radial matter flow of cosmological fluid onto the black hole.

Temperature gradients

A related issue regards the radial heat flow onto the black hole, which we interpret as a temperature gradient. It has been brought up in personal correspondence that the existence of such a temperature gradient may not be physically realistic (although I have also heard disagreeing opinions regarding this). I do not have a confident opinion on this issue—it is not obvious to me that these temperature gradients couldn't exist in the cosmological fluid, since temperature anisotropies are well-studied phenomena. This question is presumably worthy of further quantitative examination, or perhaps it merely needs a more well-posed problem in the first place. However, as I will later show, Thakurta black holes are indeed capable of emitting Hawking radiation, despite their mass growth. Presumably this prevents these temperature anisotropies from being considered a kind of violation of thermodynamic laws around the expanding black hole.

Multi-black holes and global energy density

Perhaps the most important question regards the stitching of multiple Thakurta black holes. If we presume these black holes compose some fraction of the dark matter, then there need to be many within the cosmological horizon. The multi-Thakurta black hole spacetime will probably not be trivial, since the stress-energy matching conditions (particularly for the radial heat flow) will greatly complicate the stitching. The reason this question is pressing is because if you treat each black hole somewhat naively as objects with masses $\sim ma(t)$, then you easily will find that the energy density in these objects would scale $\propto a(t)^{-2}$, compared to the $\propto a(t)^{-3}$ required for dark matter [612, 613] (ignoring for the moment that, as we argued above with respect to decoupling, once two Thakurta black holes are

sufficiently close together that they feel the effect of the others' gravity, we can no longer describe the black holes with the Thakurta metric). As I pointed out before, however, it is not correct to treat the Misner–Sharp mass as two disconnected parts, where we have a central mass and a mass of surrounding fluid, so it is not reasonable to assume that we could create a global dark matter density by 'adding up' black holes with masses $\sim ma(t)$. If you wrote the energy density in Thakurta black holes more carefully in terms of the number density n as $\rho_{\rm PBH} = nm_{\rm MS}$, however, you would end up with a recursive relation for the quantity you are attempting to define.

Regardless of the above, the question of the black hole energy density is actually not well-posed in the first place—it is not meaningful to treat the Thakurta black hole as a mass which is 'separate' from its surrounding spacetime. We are accustomed, when studying cosmology, to break up the perfect fluid energy density ρ into its components, like $\rho = \rho_{\rm radiation} + \rho_{\rm curvature} + \rho_{\rm black\ holes} + \cdots$. However, writing it in this way for the Thakurta metric would make the 'black hole energy density' part of the source for the black hole spacetime itself. It is not possible to write some global energy density term that includes the black hole contribution, since the black hole is truly part of the entire spacetime. The Thakurta spacetime, which necessarily includes both the central inhomogeneity and the cosmological background, evolves as determined by the Einstein equations with the imperfect fluid source. Only once the Thakurta black holes decouple, as described earlier, can we attribute to them an energy density component within the FLRW spacetime. Once that occurs, since we assume they interpolate into approximately Schwarzschild objects, they will pick up the standard cold dark matter scaling relation of $\rho_{\rm PBH} = nm$.

The 'Synge' procedure

A common complaint with the Thakurta solution, as well as other conformally-derived solutions such as the Sultana-Dyer black hole, is that they are found by specifying a spacetime in some way, and then deriving the stress-energy conditions which would source it [631]. Often, this procedure leads to unphysical spacetimes, of which there are many famous examples in General Relativity. It is not clear to me to what extent this is an issue or not. Indeed, I have always felt that famous spacetimes like the Schwarzschild or FLRW metric are in fact often derived first from a geometric direction, but they happen to correspond to physical stress-energy conditions. This procedure does not feel so different from using an ansatz to solve differential equations, except that the Einstein equations are

particularly troublesome and so often ‘unassuming’ spacetimes are actually problematic. But if this procedure does lead to an acceptable spacetime, is there actually an issue? Or is this complaint merely a warning to look closer at such solutions?

3.4 Gravitational wave PBH constraints

Now we have suitably introduced the Thakurta metric, let us investigate its phenomenology as a dark matter candidate. As discussed in Sec. 2.3.2, the strongest constraints on LIGO-mass PBHs are, perhaps unsurprisingly, from LIGO/Virgo [87–89, 550–554]. We will reinvestigate these limits for the Thakurta metric, originally published in our paper Ref. [6].

First, I will outline the standard calculation of the LIGO/Virgo PBH constraints. Calculating the formation of Schwarzschild PBH binaries in the early universe takes the following rough trajectory:

We start with an initial PBH population, just after formation. As the universe expands, at some point two neighboring black holes will feel a stronger gravitational attraction than the cosmological 'pull' separating them—we call this *decoupling* from the Hubble flow. It turns out that for large fractions $f_{\rm PBH}$, the density of PBHs is generally high enough that almost all black holes have decoupled before the end of radiation-domination. However, these neighboring PBHs do not just fall into each other, since their next-nearest neighbor also gives the now-formed binary some eccentricity. Using our knowledge of the initial PBH distribution, and the physics of this decoupling, we can calculate the distribution of binary parameters (i.e., semimajor axis $\mathfrak{a}$ and eccentricity e).

With an understanding of the power emitted by a binary system in gravitational waves, we can then write a distribution of the PBH binaries in terms of coalescence time τ and eccentricity. From this distribution we can read off how many binaries we expect to be coalescing today, or more specifically, within the sensitivity and range of LIGO/Virgo observations. Strong constraints can then be placed by the non-observation of the $\sim \mathcal{O}(10^4)$ mergers that is predicted by this calculation.

Although this calculation is somewhat simplistic, more detailed studies [553] of the binary formation processes did not return significantly different results. In addition to this, the binary population leads to a stochastic gravitational wave background, which can be used to place quite similar constraints to the above [556, 557]. I will not explicitly address the stochastic background calculations here, but we will see that re-examining the binary formation for Thakurta black holes would lead to the relaxation of these constraints as well.

Finally, we should note that these calculations do not account for clustered PBH formation, which is possibly more physically realistic [558–562]. In this case, the binaries

are disrupted and the bounds are lowered, which could be seen as a complimentary result to our conclusion.

3.4.1 Decoupling conditions

As it turns out, we will not need to completely redo the standard calculations for these constraints. The decoupling condition for Thakurta black holes is much more difficult to satisfy than for the Schwarzschild black hole, suppressing binary formation and avoiding the LIGO constraints altogether. Specifically, dynamical black holes must satisfy an additional decoupling condition which is satisfied automatically by Schwarzschild black holes. For convenience, I will call the standard decoupling condition the *static* decoupling condition, and the additional condition the *dynamic* condition.

The static decoupling condition

The static condition requires that the inward force towards a PBH exceeds the outward cosmological pull. The most straightforward way to derive this is to examine the radial geodesic condition for a test particle in an FLRW universe augmented by the Newtonian attraction of a Schwarzschild black hole:

$$\ddot{R} = -\frac{Gm}{R^2} + \frac{\ddot{a}}{a} R \,. \tag{3.40}$$

The condition can be set by requiring that the Newtonian force dominate over the cosmological 'drag':

$$\frac{m}{V} \gg \frac{3}{4\pi G} \left| \frac{\ddot{a}}{a} \right| = \rho_{\rm cr} \left| -\Omega_{\rm m}(1+z)^3 - 2\Omega_{\rm r}(1+z)^4 + 2\Omega_\Lambda \right| \,, \tag{3.41}$$

where $V \equiv (4\pi/3)R^3$.

Typically, the static decoupling condition is stated as a requirement that the local mass density around the black hole exceeds the cosmological energy density [550, 553]. This statement leads to a very slightly different condition than the above—the main difference coming from the double (rather than single) derivative of the scale factor. Either is an acceptable choice and the calculations are not extremely sensitive to the difference, but I find that the argument here perhaps has more physical merit.

The static condition is straightforwardly modified when considering cosmological

black holes. Very roughly, The ADM mass m in Eq. 3.40 should be replaced by the Misner–Sharp mass, $m_{\rm MS}$, and similarly in Eq. 3.41. For Thakurta black holes, the Misner–Sharp mass is dominated by the ma term, at least up close (farther away, we do not have to be concerned with binary formation anyway). We can clearly see, then, that the smaller effective mass in the static decoupling condition will make it more difficult to satisfy—the black holes will have to be much closer together than in the Schwarzschild case before a binary is formed (or rather, binaries are formed much later in time, once the right-hand side of Eq. 3.41 is reduced by the expansion of the universe).

The dynamic decoupling condition

However, there is a second, stronger, dynamical decoupling condition which we must introduce. In fact, we originally discovered this condition by accident—I had written a small code to numerically simulate the inspiral of two Thakurta black holes, and we saw that we couldn't get the black holes to actually start inspiralling until this particular condition was met.

There is a good physical explanation for this condition. Binary systems lose energy by gravitational wave emission, resulting in orbital decay [640]. In order for two black holes to form a binary, then, the power emitted in gravitational waves must dominate the cosmological power which separates the black holes. This condition emerges relatively simply, by considering the Newtonian gravitational energy between two Thakurta black holes,

$$E \sim -\frac{GM\mu a^2}{2R}, \tag{3.42}$$

where $M \equiv m_1 + m_2$, $\mu \equiv m_1 m_2/M$ and we have again approximated $m_{\rm MS} \sim ma$ (we will use this approximation throughout this section). By taking the time derivative of this energy, we can find the change in radial separation of these two black holes:

$$\frac{\dot{R}}{R} \sim -\frac{\dot{E}}{E} + 2H\ . \tag{3.43}$$

It is then easy to see the origin of the dynamical condition, since the change in radial

separation can only be negative if,

$$\frac{\dot{E}}{E} \gtrsim 2H \ . \tag{3.44}$$

In the Schwarzschild case, the scale factor a is not present in the gravitational energy, so there is no H term, and this condition is satisfied as long as there are any gravitational waves.

The power $\dot{E}$ is not totally trivial to work out for Thakurta black holes. Since we are trying to figure out when they decouple, we have to examine the gravitational waves produced by two neighboring PBHs before decoupling. To derive the gravitational wave power, we can follow the foundational work of Peters [640, 641]. Although it is somewhat more complicated for the Thakurta metric, we found that the leading order term for the gravitational wave power loss is simply,

$$\dot{E} \sim -\frac{32}{5}\frac{G^4 M^3 \mu^2 a^5}{\mathfrak{a}^5 (1-e^2)^{7/2}}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right) = \dot{E}_{\text{schw.}} a^5 \ . \tag{3.45}$$

This dynamical decoupling condition is quite strict compared to the static condition, and is responsible for much of the 'heavy lifting' in our conclusion.

Now that we have the dynamic decoupling condition in terms of the binary semi-major axis $\mathfrak{a}$ and eccentricity e, we can apply this to the PBH population to see what kind of binaries are able to form.

3.4.2 Thakurta binary formation

For simplicity, we will assume a monochromatic PBH mass function at physical mass m, and a random PBH distribution (i.e., not clustered). In this case, the average separation of these black holes at formation is given by [550, 552],

$$\bar{x}_0 = \left(\frac{m}{f_{\text{PBH}}\rho_{\text{cr}}\Omega_{\text{DM}}}\right)^{1/3} \sim \frac{1.2\ \text{kpc}}{f_{\text{PBH}}^{1/3}}\left(\frac{m}{30 M_\odot}\right)^{1/3} \ . \tag{3.46}$$

To model the binary formation, we consider a pair of neighboring PBHs which are separated by some distance $x = x_0/(1+z)$ at redshift z. I also note that the the quantity $\bar{x}_0$ is often defined in the literature at matter-radiation equality, instead of today, We will be sticking with the otherwise-more-conventional $a_0 = 1$ today, since for Thakurta black

holes, decoupling can continue even into matter-domination.

Following Ref. [550], the semi-major axis $\mathfrak{a}$ is given by,

$$\mathfrak{a} = \alpha x \tag{3.47}$$

where $\alpha \sim 0.4$ is a numerical factor from 3-body simulations for Schwarzschild black holes[551]. The next-nearest PBH to the pair helps form a binary system through tidal effects (and we presumably will assume that the the nearest neighbor of this third black hole is not one of the two currently under consideration, which may not always be a reasonable assumption). The semiminor axis $\mathfrak{b}$ is estimated by considering the product of the tidal force and square of the free fall time, where $y = y_0/(1+z)$ is the distance to the third black hole:

$$\mathfrak{b} = \beta \frac{Gmx}{y^3} \times \frac{x^3}{Gm} = \beta \left(\frac{x_0}{y_0}\right)^3 \mathfrak{a} \,, \tag{3.48}$$

where again $\beta \sim 0.8$ is a numerical factor from simulations [551]. Once both decoupling conditions are met at some redshift $z = z_{\rm dec}$, the binary is considered formed, with semimajor axis $\mathfrak{a}_{\rm dec}$ and eccentricity $e_{\rm dec}$ (calculated in the usual way, $e = \sqrt{1 - b^2/a^2}$). Since we assume that the binary evolution after decoupling proceeds as it would with Schwarzschild black holes, the binary lifetime before merging is given by [640, 641],

$$\tau_b = \frac{3}{85} \frac{\mathfrak{a}_{\rm dec}^4 (1 - e_{\rm dec}^2)^{7/2}}{r_{\rm s}^3} \,, \tag{3.49}$$

where we haved defined the binary Schwarzschild scale $r_s \equiv (G^3 M^2 \mu)^{1/3}$.

We now have all the elements to rederive the binary parameter distribution. However, this calculation is not even necessary. Using the lifetime τ_b, we can rewrite the dynamic decoupling condition Eq. 3.44 in terms of lifetime, eccentricity, and $z_{\rm dec}$, which will allow us to see the lifetime of binaries which are decoupling at any given epoch. Combining these equations gives,

$$(1 + z_{\rm dec})^3 H(z_{\rm dec}) = \frac{1}{\tau_b} \frac{96}{425} \left(1 + \frac{73}{24} e_{\rm dec}^2 + \frac{37}{96} e_{\rm dec}^4\right) \,. \tag{3.50}$$

We can solve this for the lifetime to see the evolution of binaries decoupling at any

particular time. For example,

$$\tau_b(z_{\rm dec} = z_{\rm eq} \sim 3000) = 135 \text{ s} \,, \tag{3.51}$$

$$\tau_b(z_{\rm dec} \sim 1) = 1.0 \text{ Gyr} \,. \tag{3.52}$$

Even extremely late in the universe, at $z \sim 1$, the only binaries that are able to decouple are so close together that they merge within one billion years. At matter-radiation equality, meanwhile, the lifetime of decoupled black holes is only two minutes, meaning these binaries would need to have formed extraordinarily close together. We do not need to carefully recalculate the binary distribution to see that such closely-neighboring PBHs constitute a negligible portion of the distribution, so we can conclude that almost no binaries at all decouple by matter-radiation equality. The relationship between lifetime and decoupling time is shown more explicitly in Fig. 3.1.

The crux then of the argument is as follows. By late times (well before $z = 1$), galaxies have already formed. Since the majority of Thakurta PBHs have not formed binaries in the early universe, they will become virialized into galaxies (or galaxy clusters) as isolated black holes. This process also represents a decoupling from the Hubble flow, so that these PBHs are presumably no longer well-described by the Thakurta metric, and should be interpolated into a more standard Schwarzschild-like solution. The binary formation of virialized black holes proceeds quite differently [553, 555], and appears to be more-or-less compatible with the observed LIGO/Virgo merger rate [426]. As a result, the strong constraints on PBH dark matter are fully avoided. The suppression of binary formation would also lead to an equivalent avoidance of the stochastic gravitational wave background constraints [556, 557].

This would represent a rather exciting shift in the possibilities for PBH dark matter. The LIGO-mass window, one of the few areas where we have been detecting a wealth of black holes (in the pair-instability mass gap, as well), could actually account for a large portion of the dark matter. Microlensing and accretion constraints constrain these black holes to account for maybe 10% of the dark matter—however, there are heavy uncertainties on both sides, as discussed in Sec. 2.3.2.

Regardless, we demonstrate here a dramatic change in the landscape of PBH dark matter when we properly consider a cosmological black hole metric. Even if the Thakurta metric is eventually discarded, the moral of this story is that using the Schwarzschild metric is not necessarily a good approximation, especially in the early universe. It appears that

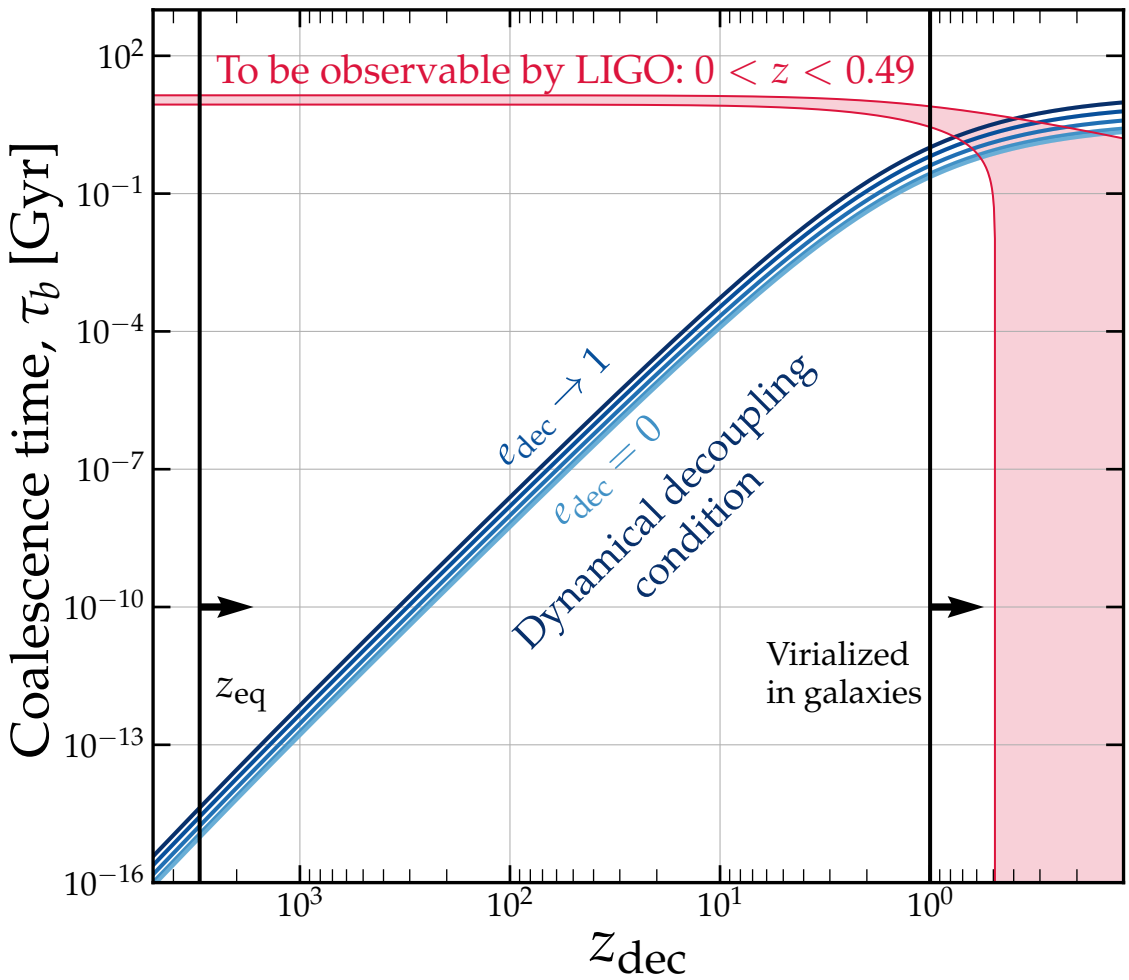

Figure 3.1: PBH coalescence times, as a function of the epoch in which they decouple. The area which would result in mergers visible by LIGO is shaded red. The blue bands represent the different possibilities for the eccentricities of the binary. We can see that even at $z = 1$, well after galaxies have formed, the PBHs merge before they could be observed. The bound on the LIGOwindow is set by the very-distant observation of GW170729 at $z \sim 0.49$ [89, 642].

cosmological black holes have radically different phenomenological consequences, which are in need of urgent consideration, considering the current Renaissance that the PBH-dark matter community is experiencing. We will see another example of this radically different phenomenology in the following section, where we investigate Hawking radiation from Thakurta black holes, and the effect that has on PBH dark matter in the asteroid-mass range.

3.5 Evaporation constraints

This final section is based on my paper Ref. [5]. In the early universe, Thakurta black holes have significantly smaller horizons than their Schwarzschild counterparts. As a result, we may expect that the rate of Hawking radiation would be significantly higher. Then even relatively large black holes may evaporate before today, disqualifying them as dark matter candidates. We are therefore interested in finding the largest black hole which evaporates before today, whose mass is referred to as the critical mass $m_{\rm c}$.

The following calculation is only a rough approximation, following the basic Hawking mass loss derivation that I outlined in Sec. 2.1.5. As we will see, the critical mass is some seven or so orders of magnitude higher than in the Schwarzschild case—even though this calculation is a relatively basic estimate, we can already see the dramatic effects on the landscape of dark matter constraints.

To start, we must define the surface gravity κ at the apparent horizon. Defining this quantity for dynamical black holes is somewhat ambiguous, and there are a few options in the literature [621]. One option is to use the null radial geodesics Eq. 2.3. The geodesic equation satisfied by these can be written in the non-affine form [621],

$$\ell^b \nabla_b \ell^a = \kappa_\ell \ell^a \ , \tag{3.53}$$

and equivalently for n. The scalar κ gains a natural interpretation as surface gravity, so at the horizon, we recover Hayward's original result [416],

$$\kappa_{\rm h} = \frac{1 - 2m'_{\rm MS}}{2r_{\rm h}} \ . \tag{3.54}$$

Unfortunately, this result does not limit to the usual result in the static case. Instead, however, we could parametrize the null vectors with the Kodama time. This leads instead to the similar result [621],

$$\kappa_{\rm h} = e^{-\phi(r_{\rm h}(t),t)} \left(\frac{1 - 2m'_{\rm MS}}{2r_{\rm h}} \right) \ . \tag{3.55}$$

As discussed in Sec. 3.3.1, however, I have not yet been able to work out an explicit form of the factor $e^{-\phi}$ at the horizon. So with this caveat, I will choose the the form Eq. 3.54 for the surface gravity, and proceed on with my calculation.

3.5.1 Thakurta black hole evaporation

The other useful factor for the following derivation is the small parameter δ, defined in Eq. 3.18 as,

$$\delta \equiv HGma < \frac{1}{8} \,. \tag{3.56}$$

Working to lowest order in this factor greatly simplifies the derivation, which follows very closely the standard form outlined in Sec. 2.1.5. To first order in δ, the surface gravity for the Thakurta metric is then,

$$\kappa \sim \frac{1}{2gma} - \frac{3\delta}{Gma} \sim 2\kappa_{\text{Schw.}}/a \,. \tag{3.57}$$

We will then make the assumption that the black hole temperature is related by $T = \kappa/2\pi$, and entropy by $S = A/4$. While it would be an interesting computation to one day carefully rederive the Hawking radiation via quantum field theory on the Thakurta spacetime, it is beyond the scope of this section. Instead, we are making the 'usual' black hole thermodynamic assumptions and assessing the phenomenological consequences. The thermodynamic identity is more involved now, since we cannot *a priori* drop the pressure and volume term:

$$\frac{\mathrm{d}U}{d\tau} = T\frac{\mathrm{d}S}{d\tau} - P\frac{\mathrm{d}V}{d\tau} \,, \tag{3.58}$$

in terms of the Kodama time τ. However, it turns out that the additional pressure term only contributes at higher orders of δ, so we can safely drop it to lowest order.

For Schwarzschild black holes, one usually assumes that the black hole is a static blackbody source. In our case, that is not the case, since the black hole horizon is explicitly growing. We need to first ensure that the surface is not too far from being thermal, so that it is meaningful to assume there is Hawking radiation at all. Indeed, it is possible to check that the rate of temperature change due to this growth is less than the rate of Hawking evaporation:

$$\left|\dot{T}\right| < \left|\dot{U}_H\right| \,, \tag{3.59}$$

where $\dot{U}_H$ is the Hawking power given by the Stefan-Boltzman law. This inequality can

be rewritten for the Thakurta black holes as,

$$240\, Gma \lesssim 1/H\ , \tag{3.60}$$

which is presumably always satisfied since the black holes form when $2Gm \sim 1/H$, and when the scale factor a is significantly smaller than $1/120$. It therefore seems reasonable that despite the rapid Misner–Sharp mass growth, the black hole horizons are instantaneously close to thermal surfaces, and Hawking radiation can occur.

An additional complication is that the change in internal energy on the left-hand side of the thermodynamic identity is not only replaced with the Stefan-Boltzmann blackbody law—we must also include the change in internal energy related to the growth of the horizon in the Thakurta metric. For that, we use the Brown-York quasi-local energy U [643] as defined in Ref. [621]. For the Thakurta metric at the horizon, we have $U = GR_{\rm h}$. To lowest order in δ, then, we must replace the left-hand side with,

$$\frac{\mathrm{d}U}{d\tau} = -\sigma T^4 A + 2\delta\ . \tag{3.61}$$

Just as in Sec. 2.1.5, we use the photon blackbody power given by the Stefan-Boltzman law. Presumably, sufficiently hot Thakurta black holes would actually produce more particles than just photons, making them evaporate quicker, possibly even by an order of magnitude or more. However, in the pursuit of a simpler and more conservative estimate, we will ignore the Page factors related to these processes here [418, 419, 422]. Then the thermodynamic identity can be rearranged to find,

$$\frac{\mathrm{d}m}{\mathrm{d}\tau} \sim -\frac{1}{1920\pi G^2 m^2 a^2} \sim \frac{8}{a^2}\frac{\mathrm{d}m}{\mathrm{d}\tau}\bigg|_{\rm Schw.}\ . \tag{3.62}$$

Specifically, this equation represents the loss of the *physical* mass m, rather than the Misner–Sharp mass $m_{\rm MS}$. While the physical mass m is not useful for determining things like geodesics of nearby particles, it is a useful parameter of the Thakurta metric to track, since it represents the mass of the black hole after decoupling. When $m \to 0$, the Misner–Sharp mass limits to the mass of cosmic fluid within some volume. In that sense, it is presumably more appropriate to track when the physical mass m goes to zero by evaporation, rather than the Misner–Sharp mass, since an identically zero Misner–Sharp mass would correspond to an empty universe and we are merely interested in an evaporated black hole.

3.5.2 The critical mass Thakurta black hole

In the previous section, we argued that by matter-radiation equality, essentially no Thakurta black holes had decoupled from the Hubble flow. This is still the case, even for tiny black holes comprising the same fraction of dark matter. Although they are closer together, the decrease in size has a stronger contribution to the time of decoupling, so that generally smaller Thakurta black holes decouple later than large ones. As a result, it is safe to assume that these black holes remain well-described by the Thakurta metric, at the very least until matter-radiation equality.

To keep the calculations simple, let us then calculate the mass of the black hole which evaporates exactly at the epoch of matter-radiation equality. Since the majority of the mass loss of Thakurta black holes is in the very earliest universe (where the scale factor is tiny), the mass of the black hole which evaporates at matter-radiation equality is only marginally smaller than the 'true' critical mass which would finish evaporating today. In light of this, I will refer now to the mass which evaporates completely at matter-radiation equality as the critical mass, $m_{\rm c}$.

In radiation-domination, the Friedmann equation takes the form $H \sim H_0\sqrt{\Omega_{\rm r}(1+z)^4}$. Of course, these equations are defined only far from the black hole, since strictly they are derived from the FLRW metric—but we can still use them here to approximately model the large scale evolution of the universe. Recalling that the Kodama time and the cosmological time approximately coincide far from the black hole and cosmological horizon, we can integrate Eq. 3.62 to find,

$$m_*^3 \sim \frac{z_{\rm eq}}{640\pi G^2 H_0\sqrt{\Omega_r}}\left(\frac{z_{\rm f}(m_*)}{z_{\rm eq}} - 1\right) , \tag{3.63}$$

where $z_{\rm eq}$ is the redshift at matter-radiation equality, and $z_{\rm f}$ is the redshift when the black holes were formed. We should keep in mind then that the critical mass is now relatively sensitive to this formation time. The model of Thakurta black holes which we assumed defines the formation time as occurring when an overdensity of mass m satisfied $m = \gamma m_{\rm H}$, for some collapse-related factor γ with $m_{\rm H}$ the Hubble mass. Then we can derive the redshift at formation as,

$$z_{\rm f}(m) = \left(\frac{2GmH_0\sqrt{\Omega_{\rm r}}}{\gamma}\right)^{-1/2} . \tag{3.64}$$

Since the formation time is safely earlier than matter-radiation equality, we can then approximately solve for the critical mass,

$$m_* \sim 9.6 \times 10^{-13}\, M_\odot \left(\frac{\gamma}{0.2}\right)^{\frac{1}{7}} \left(\frac{h}{0.67}\right)^{-\frac{3}{7}} \left(\frac{\Omega_{\rm r}}{5.4 \times 10^{-5}}\right)^{-\frac{3}{14}}, \tag{3.65}$$

where we have used the Planck 2015 values [644] and the relation $H_0 = h \times 100$ km/s.Mpc. We have then found that almost the entirety of asteroid-mass Thakurta black holes evaporate before today. Again, we must refer to the *stability constraint* described in Sec. 2.3.2—although an initial mass just above the critical mass could technically seed a population of black holes today which have masses $m < m_{\rm c}$, this initial mass would have to be incredibly fine-tuned to a value significantly more precise than our calculations even allow (see Fig. 3.2). As a result, we consider such a scenario unnatural, and the critical mass can be safely considered the smallest mass that a large population of PBH dark matter could have today, as shown in Fig. 3.3

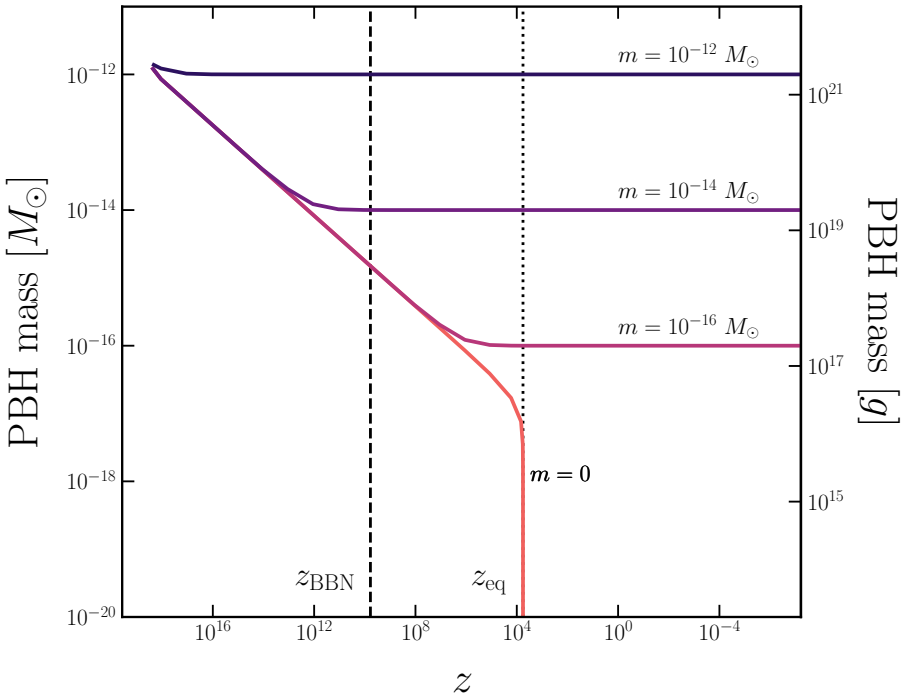

Figure 3.2: Thakurta black holes are evolved numerically following Eq. 3.62, starting from four different initial massses very close to m_* in Eq. 3.65. We can see that we would need incredible fine-tuning to produce a population today consisting of black holes of masses below the critical mass.

Of course, there are also other constraints that could be placed on these black holes,

since they are injecting a significant amount of energy into the early universe. The most relevant is probably the BBN constraints [601], which occur at redshifts $z_{\rm BBN} \sim 10^9$. Constraints which occur later, such as CMB anisotropies [538, 604], occur at significantly smaller redshift and so the effects of the Thakurta phenomenology are lessened.

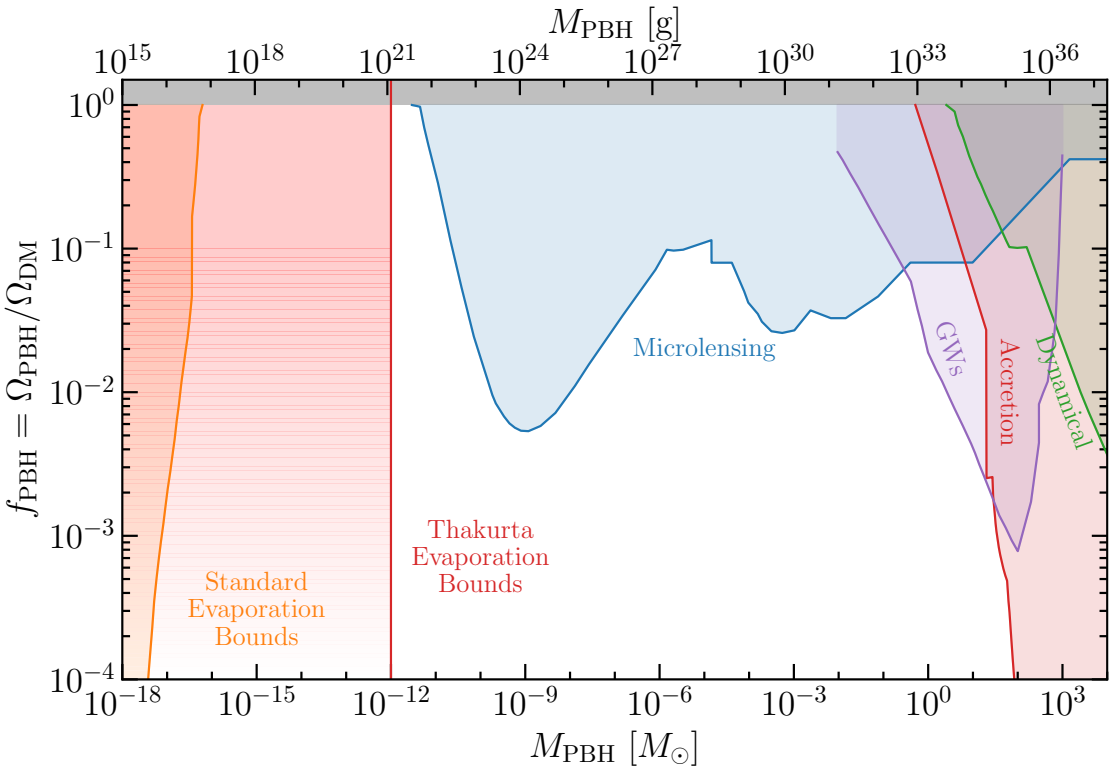

Figure 3.3: The evaporation bounds on the PBH constraint plot [515]. Fig. 2.1 is modified in the Thakurta case. Because larger black holes evaporate rapidly in the early universe, the smallest black hole which can comprise the dark matter has mass $m_* \sim 10^{-12} M_\odot$, effectively closing the asteroid-mass range for PBH dark matter..

This calculation was quite simplistic and did not account for a number of factors, such as the possible normalization $e^{-\phi}$ for the surface gravity, and the Page factors related to the full spectrum of particle production at a given temperature. The latter, however, would only increase the critical mass, strengthening the bounds. While I do not know exactly what the results of the detailed quantum mechanics calculations would give for Thakurta or other dynamical black holes, my hunch is that the results will not be able to depart greatly from the rough approximation here. Still, it would be interesting and valuable to examine these details, and I plan on doing so in the future.

The layperson box: cosmological black holes (my work) Previous — Next

Finally, we arrive at *my* PhD research. Black holes are traditionally described using a very simple model—essentially, they are big 'dips' in spacetime, surrounded by absolutely nothing, all the way to infinity. However, when primordial black holes (page 63) form, they find themselves in an unusual environment. The universe, being much smaller than it is today, is hot and dense with particles of all kinds. Not only are the surroundings of the black holes not nothing, you cannot even go all the way to infinity, since you will rapidly encounter the 'edge' of the universe (it is not a real edge, though, but rather a point at which the black hole would disappear from your sight because its being 'expanded away' from you faster than lightspeed).

It turns out that modelling black holes in this environment is somewhat complicated. Actually, the very first PBH paper, by soviet physicists Igor Novikov and Yakov Zeldovich, attempted to explain that PBHs couldn't exist, because this hot, dense fluid would cause them to grow too rapidly. It was then pointed out by Bernard Carr and Stephen Hawking, however, that the way to solve this problem involved a different model of black holes which explicitly accounts for this *cosmological* background they find themselves in. Thus, the field of primordial black holes was born.

Weirdly enough, while some physicists spent time coming up with these kinds of models, few people bothered to check if these different models had an effect on, say, whether the primordial black holes could be the dark matter—most people just used the old, empty-space model. So my collaborators and I took one of these cosmological black hole models that seemed the best to us and had a look at its physics.

The model we chose, the 'Thakurta' black hole, has a weird property. As the universe grows around it, the black hole grows too. One can imagine a hole in a piece of stretchy fabric—as you pull on the fabric, the hole would stretch as well. Really, this means that the black holes looked tiny in the much-smaller early universe. This has two big effects: 1; black holes evaporate faster, and 2; the black holes are harder to 'find', so they will not orbit each other as easily. Because Thakurta black holes evaporate more rapidly, this first effect rules out a whole chunk of medium-sized (asteroid-mass) black holes that otherwise were totally viable as dark matter. The second effect remarkably rules back *in* the kind of black holes that we see from gravitational waves, since fewer black holes are orbiting each other and eventually colliding. In both cases, what we thought about black hole dark matter is radically altered.

If you know me in real life, you might be aware that our results were the subject of some 'discussion' in the literature. In my opinion, it is totally fair enough to question our model (although I believe our *conclusions* are relatively solid). The Thakurta black hole is not perfect, and I would be very appreciative if someone could suggest a better one. Still, we demonstrated an important lesson—not only is choosing a model not easy, it has serious physics consequences.

4
Axions

Once there were brook trout in the streams in the mountains. You could see them standing in the amber current where the white edges of their fins wimpled softly in the flow. They smelled of moss in your hand. Polished and muscular and torsional. On their backs were vermiculate patterns that were maps of the world in its becoming. Maps and mazes. Of a thing which could not be put back. Not be made right again. In the deep glens where they lived all things were older than man and they hummed of mystery.

—Cormac McCarthy, *the Road*

Now we have finished our exploration of black holes in the early universe, it is time to take a step back into the world of particle physics (before stepping forward again into the world of dark matter). In this chapter I will introduce the *axion*, a hypothetical particle which was introduced to solve a long-standing issue with the theory of the strong force known as the strong-CP problem. As it turns out, the axion is also a rather handy dark matter candidate. This two-for-one property of the axion, coupled with the ever-tightening bounds on WIMP dark matter, has seen the axion become an increasingly attractive dark matter theory.

4.1 The strong-CP problem

Often, introductory texts on axions begin at the end of the story, where we postulate that there may be terms in the Standard Model Lagrangian which violate CP-invariance for the strong force. I have always found these brief explanations more confusing than helpful, and so I would like to take a longer and more historical route to fully understanding this problem, and therefore, the motivation for the axion. This section is partially based on the excellent notes and reviews of Refs. [12, 15, 356, 645], but Refs. [175, 176, 179, 181, 646–652] represents an inexhaustive list of the seminal works summarized here.

4.1.1 The quark flavor symmetry

The road to the axion begins with a particle symmetry known as *strong isospin*, which elucidates an interesting feature of the strong force. This symmetry should not be confused (except in mathematical structure) with *weak isospin*, which is more akin to charge under the weak SU(2) gauge group—for example, in the weak doublet,

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} , \tag{4.1}$$

one would say that the electron neutrino ν_e has weak isospin $+1/2$. In contrast, strong isospin is a symmetry, or a quantum number, belonging to hadrons. Historically, it was noticed that groups of hadrons had very similar properties, and so could be put in multiplets such as the Delta baryon 4-plet:

$$\begin{pmatrix} \Delta^{++} & \mathrm{uuu} \\ \Delta^{+} & \mathrm{uud} \\ \Delta^{0} & \mathrm{udd} \\ \Delta^{-} & \mathrm{ddd} \end{pmatrix} , \tag{4.2}$$

where I have also written the quark content of these hadrons for demonstration. In this case, for example, the Δ^{++} baryon has strong isospin component $+3/2$, while the whole Delta baryon multiplet, as an object, has strong isospin of $3/2$. In fact, as experiments pushed to higher energies, they found that there were two conserved quantities—strong hypercharge Y, and strong isospin I_3 (sometimes denoted T_3), and groups of hadrons could be placed nicely on a two-dimensional plot with these axes, as shown in Fig. 4.1.

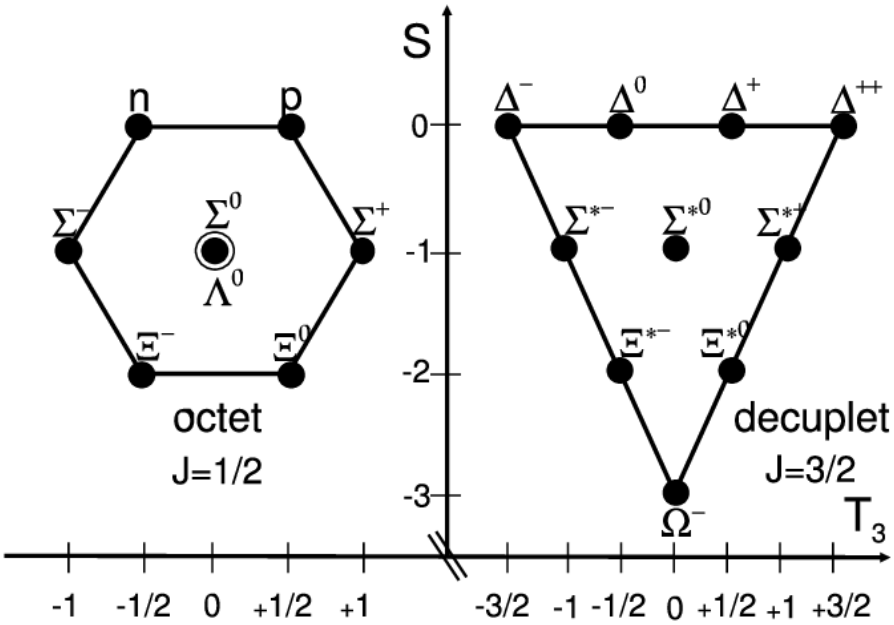

Figure 4.1: In this image, the octets and decouplets of an SU(3) flavor symmetry are plotted [653]. The vertical axis is strangeness S and the horizontal axis is strong isospin (specifically the 'third' component T_3). The Delta baryon four-tet can be seen at the top of the right triangle. More technically, these baryon states are known as the 'weights' of the representation under SU(3). Then, the lines which connect these states can be heuristically thought of as representing operators, analogous to ladder operators, which move between these eigenstates.

With modern eyes, the origin of these symmetries is best understood by examining the quark content of the hadrons. In fact, what we have here is the continuous version of an exchange symmetry between the up, down, and strange quarks—one can (approximately) swap these quarks for each other and the overall properties of the hadrons is conserved, apart from electromagnetic charge. While the discrete exchange symmetry would just swap quarks, however, we are actually free to continuously rotate the three quarks, so that a new state could be a superposition of the three quarks. Specifically, this continuous global symmetry on the three quark states is known as *flavor symmetry*.

The two historical symmetries are actually written out more clearly as,

$$\begin{aligned} I_3 &= 1/2\,(n_u - n_d) \\ Y &= B + S\,(+C + B' + T)\,, \end{aligned} \tag{4.3}$$

where n_u and n_d are the number of up and down quarks respectively, B is baryon number, S is strangeness, and although it is not relevant to our discussion here, C, B', and T are the charm, bottomness, and topness.

This quark exchange symmetry is actually approximate, since these quarks do not have exactly the same masses. However, since their masses m_q are all lower than the QCD scale Λ_{QCD}, we have an *approximate symmetry*. The meaning of approximate symmetry is perhaps best understood through Nöether's theorem—you can still calculate currents corresponding to this symmetry, but there are corrections to the Nöether current that scale proportionately to something like $m_q/\Lambda_{\mathrm{QCD}}$. In addition, in this limit where quark masses vanish, we actually have a chiral symmetry (which is otherwise broken by the quark mass terms).

For simplicity, we will mostly stick to just the up and down quark flavor symmetry from now on. The quark flavor symmetry is represented mathematically like,

$$\begin{pmatrix} u \\ d \end{pmatrix} \to U \begin{pmatrix} u \\ d \end{pmatrix}, \qquad U \in \mathrm{U}(2)\ , \tag{4.4}$$

where $\mathrm{U}(2)$ is the unitary group of rank two. More precisely, U is an element of the fundamental representation of $\mathrm{U}(2)$. If we include strange quarks, we would use $\mathrm{U}(3)$ and a triplet of the quarks. Because we have an approximate chiral symmetry, we can use that $\mathrm{U}(N) = \mathrm{SU}(N) \times \mathrm{U}(1)$ to find that we have the following global symmetry of the Standard Model Lagrangian:

$$\mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{SU}(2)_{\mathrm{R}} \times \mathrm{U}(1)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{R}}\ . \tag{4.5}$$

However, it is sometimes preferable to describe our fields in a vector/axial basis, rather than left/right:

$$\Psi_{\mathrm{V}} = \left(\Psi_{\mathrm{L}} + \Psi_{\mathrm{R}}\right)/2 \qquad \Psi_{\mathrm{A}} = \left(\Psi_{\mathrm{L}} - \Psi_{\mathrm{R}}\right)/2\ . \tag{4.6}$$

Then, our symmetry is written,

$$\mathrm{SU}(2)_{\mathrm{V}} \times \mathrm{SU}(2)_{\mathrm{A}} \times \mathrm{U}(1)_{\mathrm{V}} \times \mathrm{U}(1)_{\mathrm{A}}\ . \tag{4.7}$$

4.1.2 The $\mathrm{U}(1)_\mathrm{A}$ problem

It is all well and good to observe that the Lagrangian seems to have a global symmetry, but we should then actually check if this symmetry is obeyed in real life. Two of these symmetries are easy to identify—$\mathrm{SU}(2)_\mathrm{V}$ actually corresponds to strong isospin, and $\mathrm{U}(1)_\mathrm{V}$ corresponds to baryon number, bringing us full circle to the historical beginning of this discussion. In that case, what do we make of the remaining axial symmetries?

At first, it does not appear that we actually do observe these symmetries in real life. For example, the vacuum expectation values of non-zero quark condensates seem to violate this symmetry:

$$\langle 0|\bar{Q}Q|0\rangle = \langle 0|\bar{Q}_\mathrm{L}Q_\mathrm{R} + \bar{Q}_\mathrm{R}Q_\mathrm{L}|0\rangle \ , \tag{4.8}$$

where Q is a quark state. The above is not invariant under the axial symmetry, since if $Q_\mathrm{L} \to U_\mathrm{L}Q_\mathrm{L}$, and the same for the right quarks, we can see that it is only invariant if $U_\mathrm{L} = U_\mathrm{R}$, which is only the case for vector part of the symmetry.

This situation, for particle physicists, implies that the axial symmetry must be *spontaneously broken*—that is, the vacuum state is not invariant under the symmetry, even though the original Lagrangian is. When this occurs, the particle spectrum is left with what are known as *Nambu–Goldstone bosons*, which can be thought of as excitations of the fields in the directions that the symmetry is broken. In this case, because we actually started only with an approximate symmetry, we expect the Nambu–Goldstone bosons to pick up a small mass. Since it appears that the $\mathrm{U}(2)_\mathrm{A}$ symmetry has been broken, there should be four Nambu–Goldstone bosons, corresponding to each broken generator of the group. Three of these come from the spontaneous breaking of $\mathrm{SU}(2)$, while one comes from $\mathrm{U}(1)$. If we were to consider strange quarks, there would be eight Nambu–Goldstone bosons from $\mathrm{SU}(3)$ instead.

We must then look, in real life, for four pseudoscalar mesons which would be these Nambu–Goldstone bosons. It turns out, we can at least find three such particles—the pions. However, the fourth, which might have been the η-meson, is far too heavy. (For that reason, the $\mathrm{U}(1)_\mathrm{A}$ problem is also sometimes referred to as the η-meson problem.)

So if the $\mathrm{U}(1)_\mathrm{A}$ symmetry is not spontaneously broken, the only remaining option is that it is *explicitly broken*—there must be an anomaly at the quantum level. Such a process can certainly be found, as in Fig. 4.2. From this diagram, it was calculated [177, 178] that

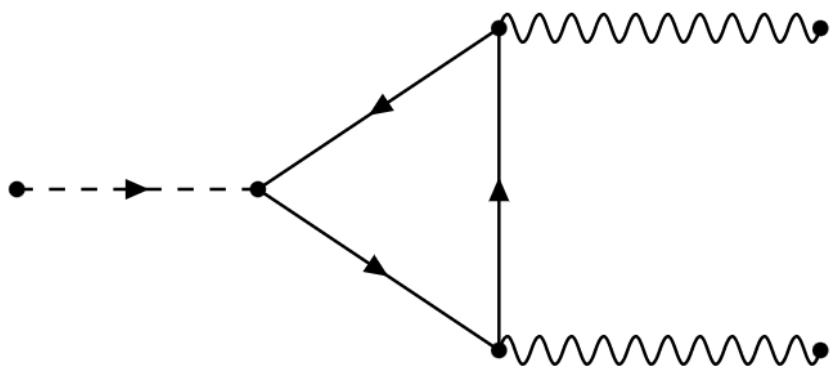

Figure 4.2: A schematic Feynman diagram of the chiral anomaly, also known as the Adler-Bell-Jackiw anomaly [177, 178]. The dashed line represents an ingoing axial current, while the interior lines are quarks and the outgoing lines are photons. In historical contexts, this was analyzed as the decay of a neutral pion to two photons, but the final states could also be gluons.

the flux of axial current goes like,

$$\partial_\mu J^\mu_{\mathrm{A}} \propto G^a_{\mu\nu}\widetilde{G}^{\mu\nu a}, \tag{4.9}$$

where the a index runs over the gluons, and the tilde gives the dual of the gluon tensor. What this means is that applying the $\mathrm{U}(1)_{\mathrm{A}}$ transformation to the Standard Model Lagrangian gives us,

$$\mathcal{L} \to \mathcal{L} + \frac{\alpha}{16\pi^2} G^a_{\mu\nu}\widetilde{G}^{\mu\nu a}\ , \tag{4.10}$$

where α is the parameter of the transformation and the normalization is for convention. However, while I will not show the algebra here, it is possible to write this additional term as a total derivative $\partial_\mu K^\mu$:

$$K^\mu \equiv \epsilon^{\mu\alpha\beta\gamma} A_{a\alpha}\left[F_{a\beta\gamma} - \frac{g}{3} f_{abc} A_{b\beta} A_{c\gamma}\right], \tag{4.11}$$

where f_{abc} is the structure constant of the $\mathfrak{su}(2)$ or $\mathfrak{su}(3)$ lie algebra and g is the strong force coupling constant.

If the gluon fields vanish at infinity, then total derivatives do not contribute to the field equations, and so are not observable. Then we are apparently left with a confusing problem—we are out of ways to easily explain why we do not observe the $\mathrm{U}(1)_{\mathrm{A}}$ symmetry

in real life. And yet, $\mathrm{U}(1)_{\mathrm{A}}$ symmetry appears to be conserved in the Standard Model Lagrangian. This is what was known, historically, as the $\mathrm{U}(1)_{\mathrm{A}}$-problem.

4.1.3 The QCD vacuum

The $\mathrm{U}(1)_{\mathrm{A}}$-problem, and its solution, sets the stage for the strong-CP problem. To answer it, we must look more closely at the QCD vacuum, and in particular, the assumption that the gluon fields vanish at infinity—i.e., $A^{\mu}\left(|x| \to \infty\right) = 0$. Recall that under the action of a member of a gauge group, like $U \in \mathrm{SU}(3)$ for the strong force, the gluon fields transform like,

$$A_{\mu} \to U A \mu U^{\dagger} - \frac{i}{g} U \partial_{\mu} U^{\dagger} \ , \tag{4.12}$$

where g is the coupling constant of the force. Then we see that there are actually a continuous class of field configurations, known as *pure-gauge* states, which are just gauge transformations of the field at infinity:

$$A_{\mu}\left(|x| \to \infty\right) = 0 \ \ \Rightarrow \ \ A_{\mu}^{\mathrm{PG}}\left(|x| \to \infty\right) \equiv -\frac{i}{g} U \partial_{\mu} U^{\dagger} \ . \tag{4.13}$$

What's more, these pure-gauge states are defined uniquely by U. We would like to investigate these states, under three main assumptions. Firstly, we will use the temporal gauge $A_0 = 0$. Secondly, we will demand that $U(x)$ is continuous, or else we do not have a well-defined pure-gauge state. The third assumption is more difficult to motivate before we continue with our analysis, but is still extremely important. We require that:

$$U\left(|x| \to \infty\right) = 1 \ . \tag{4.14}$$

At first glance, this is true for the case $A^{\mu}\left(|x| \to \infty\right) = 0$ (actually, it only needs to equal a constant, but 1 is a perfectly fine choice). If this assumption still feels a bit ad-hoc for the pure-gauge states, however, hopefully the continuing discussion will alleviate some worry.

To gain intuition without becoming bogged down in mathematics, let us back up and investigate the more simple $\mathrm{U}(1)$ gauge group in 1+1 spacetime. The power of the third assumption we made is that the spacetime is compactified, since the points at infinity are identified. That is, we now have a spacetime coordinate $x \in S^1$, rather than $x \in \mathbb{R}$. It

is also prudent to note that $\mathrm{U}(1)$, as a group, is isomorphic to S^1, since we can think of $U(x) \in \mathrm{U}(1)$ as unit-length vectors in the complex plane.

Since the pure-gauge states are purely defined by $U \in \mathrm{U}(1)$, we are therefore interested in the ways that $U(x)$ can be defined on S^1. To illustrate these, Fig. 4.3 shows three ways this can be mapped, keeping $U(x)$ continuous, and representing $U(x)$ with arrows in the two-dimensional complex plane.

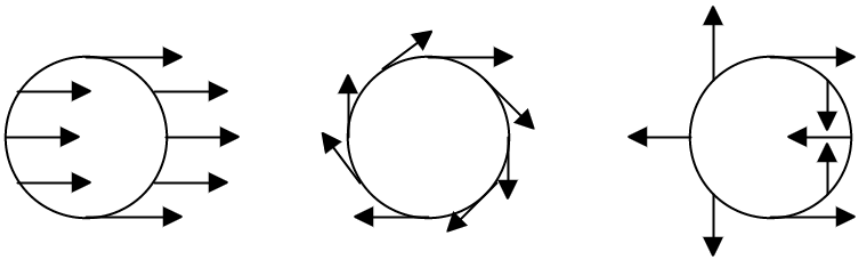

Figure 4.3: The one-dimensional compactified spacetime is represented by the circle, whereas the arrows represent an element $U \in \mathrm{U}(1)$ in the complex plane. Since $\pm\infty$ is identified, the arrow needs to return to the same position at the top of the circle, but it can continuously change across the rest of the circle. In the first situation, the arrows are constant and so wrap zero times around the circle. In the second, the arrows wrap exactly once around the space. In the third, the arrows wrap twice. We can see then that there are integer-labelled windings of the gauge field on the circle, and with a small stretch of the imagination, we can also realize that no continuous deformation will be able to move between these windings without 'breaking' somewhere.

Without delving too far into the mathematical toolbox of topology, we can intuit that there will not be any continuous transformations you can apply to one of these mappings to transform it into another, without breaking $U(x)$ somewhere. What this means is that there is a class of *distinct* pure-gauge states, which we can index by an integer called a *winding number*. We are essentially asking, 'how many times can we wind a circle around a circle?'. With the language of algebraic topology, the answer to this question is written $\pi_1(S^1) \cong \mathbb{Z}$, or in English, the first fundamental group of the circle is isomorphic to the group of integers.

Now we should return to the real world, where our gauge group is $\mathrm{SU}(3)$ and we have three spatial directions, $\mathbb{R}^3$, which is now compactified to the three-sphere S^3. To save some headache, we can first use the fact that $\mathrm{SU}(2)$ is embedded in $\mathrm{SU}(3)$, so we can examine the mapping of $\mathrm{SU}(2)$ onto S^3 instead. Again, we exploit a similar fact, that we have a group isomorphism $S^3 \cong \mathrm{SU}(2)$. We are then asking a similar but harder-to-

imagine question—how many ways can a three-sphere be wrapped around a three-sphere? This answer has long been known by algebraic topologists:

$$\pi_3\left(S^3\right) \cong \mathbb{Z}\,. \tag{4.15}$$

This result is handily similar to our simpler analysis before, and so may perhaps be considered unsurprising to those with excellent topological intuition.

We should stop then to summarize the argument so far. We have found a number of topologically distinct pure-gauge states, for the case when $U(x) = 1$ at spatial infinity. These states are indexed by an integer and cannot be smoothly transformed into each other.

Let us then consider this compactifying assumption again, with the aid of a simple toy problem from Ref. [12]. Say we are looking at the transition between two of these pure-gauge states, written as $A^0 \to A^1$. To allow the states to smoothly make this transition, we will have to pass through pure-gauge states where $U(x) \neq 1$ at infinity. Let us imagine a real parameter β which parametrizes this transition:

$$A^\beta_\mu = \beta A^1_\mu\,. \tag{4.16}$$

To be more general, we could make β a function of x, but the argument would still be similar. Let us look at the field-strength tensor $G_{\mu\nu}$ for $0 < \beta < 1$:

$$\begin{aligned} G_{\mu\nu} &\equiv \partial_\mu A^\beta_\nu - \partial_\nu A^\beta_\mu + \left[A^\beta_\mu, A^\beta_\nu\right] \\ &= \beta\left(\partial_\mu A^1_\nu - \partial_\nu A^1_\mu\right) + \beta^2\left[A^1_\mu, A^1_\nu\right] \\ &= \left(\beta^2 - \beta\right)\left[A^1_\mu, A^1_\nu\right] \\ &\neq 0\,. \end{aligned} \tag{4.17}$$

In the second line, we have used the definition in Eq. 4.16 before rearranging, and in the third line, we have used that A^1 is a pure-gauge state for which $G_{\mu\nu} = 0$. Then we can see that the energy E of this intermediate state goes like,

$$E \propto \int d^3x G^a_{\mu\nu} G_{a\mu\nu} \neq 0\,. \tag{4.18}$$

The 'physics' interpretation of all this is the following—passing from one of these integer-labelled pure-gauge vacuum states to another requires you to tunnel through a potential energy barrier. In some sense, we have now justified the third assumption earlier. We can

consider only the pure-gauge states which satisfy $U(x) = 1$ at infinity, since the states for which this isn't true are actually higher-energy states which must be passed through when moving between these topologically distinct states. And because we would like to include $A^\mu\left(|x| \to \infty\right) = 0$ as our 'starting point', the other relevant vacua also then satisfy the compactifying condition as well.

Instantons

Our task now is to explicitly find these field configurations which tunnel between vacua. For intuition, we can briefly recall some first-year quantum mechanics and imagine the finite potential well problem. The solution inside the well is proportional to e^{-iEt}, whereas the tunneling solution is proportional to e^{-Et}. The lesson here is that we can find tunneling solutions by first transforming our field equations from Minkowski space to Euclidean space, via the transformation $t \to it$. Solutions to field equations in Euclidean space are known as *instantons*.

In our case, we are looking for a solution which changes the winding number of the pure-gauge states by one (we can find other instantons which mediate higher-integer changes, but their impacts are suppressed compared to lower integers). To find the instanton solution, we must minimize the four-dimensional Yang-Mills Lagrangian, $\mathcal{L} = -\frac{1}{4}G^a_{\mu\nu}G_{a\mu\nu}$, in Euclidean space. An explicit solution can be found, and is known as the Belavin–Polyakov–Schwarz–Tyupkin (BPST) instanton [648],

$$A^a_\mu\left(x\right) = \frac{2}{g}\frac{\eta^a_{\mu\nu}\left(x-z\right)_\nu}{\left(x-z\right)^2+\rho^2}\ , \tag{4.19}$$

where z is the center of the instanton and η here is the 't Hooft symbol. The name 'instanton' is derived from the fact that this solution kind of looks like a localized, instantaneous flash as we move from one vacuum configuration to another. Interestingly, the difference in winding number of two particular configuration can be calculated by examining the axial anomaly, using the tellingly-familiar calculation:

$$m - n = \frac{g^2}{32\pi^2}\int \mathrm{d}^4x\, G^{a\mu\nu}\widetilde{G}^a_{\mu\nu}\ , \tag{4.20}$$

and it is easily checked that $m - n = 1$ for the BPST instanton. Sometimes the value $Q \equiv m - n$ is called the topological charge of the instanton.

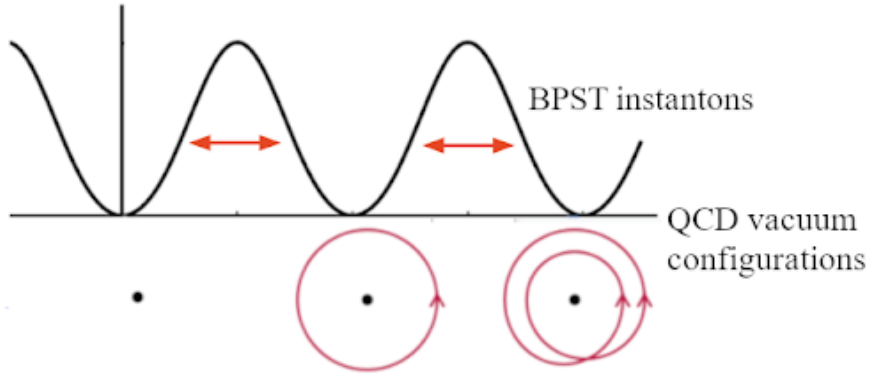

Figure 4.4: This cartoon diagram shows the potential wells which form between the topologically distinct QCD vacuum configurations. The BPST instantons are field solutions which tunnel between these vacua.

The θ-vacuum

With an infinity of vacua and explicit tunneling processes between them, the remaining question is, 'which of these vacua is the true QCD vacuum?' We can at least simply argue that it is no single state, since none of them are gauge-invariant—while the integer-labelled vacua are invariant under so-called 'small' gauge transformations (i.e., which can be smoothly transformed to the identity), they are clearly not invariant under the aptly-named 'large' gauge transformations (i.e. which possess some nonzero winding).

Actually, it was shown in Ref. [646] that we should write the 'true' QCD vacuum as a coherent superposition of all these topologically distinct pure-gauge vacua. One argument for this has to do with the locality-related cluster decomposition theorem for QFTs. This theory proposes that that the vacuum expectation value at infinity of a product of operators, defined at distant spacial points, should equal the product of the expectation values of each of the operators. In other words, we can decouple isolated systems from each other and maintain causality if the separation is spacelike. Perhaps it is unsurprising that this appears in our instanton solutions, since they are heavily spatially-dependent.

The argument of Ref. [646] is relatively simple, but well elaborated in Ref. [645]. The essence of the argument is that we need to sum over all the possible gluon configurations with some unspecified weighting $w(n)$, in terms of the winding number n. Two distant gluon fields will then have $w(n_1 + n_2)$, but the cluster decomposition splits the vacuum expectation value into a product of operators—as a result, we require $w(n_1 + n_2) = w(n_1)w(n_2)$. Thus the weight must be an exponential function and the vacuum can only

be written as,

$$|\theta\rangle \equiv \sum_{n=-\infty}^{\infty} e^{in\theta} |n\rangle \,, \tag{4.21}$$

which is conveniently gauge-invariant, and where we have used the ket $|n\rangle$ to represent the topologically distinct vacua. This vacuum is parametrized by an unspecified parameter θ, giving it the name θ-*vacuum*. This vacuum angle $\theta \in (0, 2\pi)$ is a parameter of the theory, and it is straightforward to show that different θ-vacua have zero overlap. That is, there is a super-selection rule $\langle\theta_1| e^{-H\tau} |\theta_2\rangle \propto \delta(\theta_2 - \theta_1)$. While the $\theta = 0$ vacuum has the lowest energy (an important fact for later), the lack of transitions between these vacua means that in principle the theory could take any value. Beyond this, it is not so straightforward to give a physical interpretation to θ. Often, it is described heuristically as the phase picked up by a field when tunneling—it can be thought of analogously to momentum in field space, since it is conserved when shifting $|n\rangle \to |n+1\rangle$.

Finally—the strong-CP problem

We are very nearly at the heart of the strong-CP problem. What is the consequence of using the θ-vacuum? Here I believe it is more instructive to examine these effects qualitatively, rather than expending significant effort with mathematical details.

In the path integral formalism, we must sum over all paths that could possibly mediate some particle process. This includes, then, paths which move between distinct vacua via instanton processes. These transitions have finite energy and so must be accounted for in the path integral. In essence, we must recalculate the vacuum-vacuum transition $\langle\theta| e^{-H\tau} |\theta\rangle$. It turns out that including these processes leads to a somewhat trivial dependence of the Lagrangian on θ:

$$\Delta\mathcal{L} = \frac{\theta}{16\pi^2} G^a_{\mu\nu} \widetilde{G}^{a\mu\nu} \,. \tag{4.22}$$

Notice how this is exactly the same functional form as the axial anomaly in Eq. 4.10! This arises after a relatively straightforward calculation, and it is essentially the form of Eq. 4.20 that leads to this simple θ-dependence in the Lagrangian.

This finally allows us to answer the $\mathrm{U}(1)_\mathrm{A}$-problem. Recall that it appeared that there was a $\mathrm{U}(1)_\mathrm{A}$ symmetry in the Lagrangian which wasn't conserved in real life. It wasn't

spontaneously broken either, and while it appeared to be explicitly broken, the anomaly term in Eq. 4.10 could not contribute to the equations of motion because it was a full derivative. Now, we finally have our answer—there is no $\mathrm{U}(1)_{\mathrm{A}}$ symmetry in the Standard Model Lagrangian in the first place. Rather, such transformations effectively modify the Lagrangian by shifting the theta parameter,

$$\theta \to \theta + \alpha \,. \tag{4.23}$$

On first glance, this is a handy fact—we are given a degree of freedom here to transform this term to whatever it needs to be to conform to real life. There is one catch, however, and we must examine the quark masses again to see it. It is a curious fact of the weak force that the mass eigenbasis differs from the weak eigenbasis, and so quark masses mix under the weak force. Because the weak force is chiral, if we would like to fix the quark masses, we are forced to use up this bonus degree of freedom, like so:

$$\Psi_{\mathrm{L}} \to e^{-i\mathrm{ArgDet}(M)}\Psi_{\mathrm{L}} \,, \tag{4.24}$$

where M is the quark mass matrix, and similarly for Ψ_{R} with a difference of sign in the exponent. It is handy to define the parameter,

$$\bar{\theta} \equiv \theta + \mathrm{ArgDet}(M), \tag{4.25}$$

which we find is now totally fixed by the two (presumably) entirely unrelated weak and strong forces. The following term is then necessarily included in the Standard Model Lagrangian:

$$\frac{\bar{\theta}}{16\pi^2} G^a_{\mu\nu}\widetilde{G}^{a\mu\nu} \,. \tag{4.26}$$

If this term did not affect any physical observables, perhaps we would not need to worry. However, $\bar{\theta}$ does appear in measurements—it provides a CP-violating effect for the strong force. Most notably, it appears in the calculation of the neutron electric dipole moment [180, 654–660]. Experimental results [661–663], however, find the surprising result:

$$\bar{\theta} \lesssim 10^{-10} \,. \tag{4.27}$$

If the weak and strong forces are unrelated, how is it possible that such an incredible fine tuning has occurred between the two terms in Eq. 4.25? *This* is the strong-CP problem.

The layperson box: Problems with the strong force Previous — Next

Actually, fundamental particles like electrons and quarks can be split up even further. There are really two 'halves' to these particles—left-handed and right-handed versions, and one can (vaguely) imagine this refers to whether they spin clockwise or counter-clockwise. Originally, physicists thought there were no differences in the left- and right-handed particles. However, experiments in the 1950s found that there *were* differences, at least whenever the weak force was acting on the particles. Specifically, they were looking at neutrinos, an elementary particle which only interacts with the weak force. They found that left-right symmetry was *maximally* violated—there did not seem to even be any such thing as right-handed neutrinos.

There was a perhaps stranger-yet discovery in the 1970s. When the strong force—the force that holds the nucleons in atoms together—acts on particles, experiments showed it did not care about leftiness or rightiness (inside the nucleons, it is particles called quarks which feel this strong force). But, exactly oppositely to the situation with the weak force, now theorists were saying there *should* have been a difference!

The reason theorists were saying this is somewhat esoteric in plain English, but interesting nonetheless. Far from any particles, the strong force does not just decrease to zero like the other forces. Instead, it can 'twist' along the way. While the number of twists does not matter, changing the number of twists takes a small amount of energy. You can imagine these tiny twistings and untwistings happening randomly and constantly, like how particles and antiparticles can pop out of nothing then disappear again. The cumulative effect of all this twisting and untwisting would not matter, however, if the twistings were 'in phase' (i.e., starting at the same point) as the left- and right- spinning of the quarks. If they were out of phase, though, these twists and untwists would necessarily affect the left-handed quarks differently to the right-handed ones, which could be observable in experiment.

The kicker, however, is that the phase of the quarks are already set-in-stone by the unusual machinations of the *weak* force, since it seemed to mess with lefty and righty particles differently. There appeared to be no way of aligning the strong force's twisting with the quark's twisting, other than by assuming the twists just 'happened' to be randomly exactly in phase with the *totally unrelated* weak force. And yet, experiments tell us there are definitely no differences in left- and right-handed quarks. Was it all just perfectly aligned by total accident?

This kind of assumption furrows the brows of physicists—when something requires incredible 'fine-tuning' to work out, there is probably a problem in the theory. Fortunately, there is nothing theorists love more than a theoretical problem, and they rather quickly found a solution...

4.2 The Peccei–Quinn solution

Let us recap the main points of the previous discussion. We began by observing the $\mathrm{U}(1)_\mathrm{A}$ problem—the Standard Model seems to have this global axial symmetry, but it is not manifest in nature. It is also not spontaneously broken, and while it is anomalously broken, it is done so with topological terms which appear to vanish. However, by looking closer at the non-trivial structure of the QCD vacuum, we saw that in fact, the anomaly terms do not vanish. Rather, they contribute a source of CP-violation for the strong force, via the $\bar{\theta}$ angle. However, observations tell us that this parameter must be extremely close to zero, which is quite surprising, since the $\bar{\theta}$ parameter is a combination of presumably-unrelated weak and strong force considerations. This fine-tuning problem is known as the strong-CP problem.

Before we examine Roberto Peccei and Helen Quinn's solution [354, 355], we briefly note some alternative solutions to the strong-CP problem. In Ref. [356], Peccei outlines a few possibilities. Of course, the angle could just happen to be so small, but there are not clear anthropic reasons for it to be so, unless you engage in some rather finicky dark matter model-building [664]. It could be that the QCD vacuum structure is a mathematical artifact and not a physical feature of the theory, but then we would need a new solution to the $\mathrm{U}(1)_\mathrm{A}$ problem [665, 666]. It could also be that CP is itself spontaneously broken [667–671]. However, in addition to the complexities in the technical details of such models, observations today indicate that quark masses are described with the Cabibbo–Kobayashi–Maskawa (CKM) Model [163, 672, 673], where CP is broken explicitly rather than spontaneously.

4.2.1 General considerations

The inspiration for the Peccei–Quinn (PQ) solution comes from the following observation. In general, for chiral field theories, axial currents have an anomaly term in the form of a triangle diagram such as Fig. 4.2. We particularly have in mind the case where the quarks are massless, which would make QCD a chiral theory. Then we could use the axial symmetry to transform the quarks with $q \to \exp(i\gamma_5\alpha)q$, effectively sending $\bar{\theta} \to \bar{\theta} - 2\alpha$. The freedom to choose the parameter α as we like means we could avoid the strong-CP problem. Unfortunately, it does not appear that any of the quarks are massless—in other words, $\mathrm{ArgDet}(M) \neq 0$. However, we can use the general idea behind this observation to guide our considerations for a solution.

First, we note that introducing an extra massless quark is generally difficult, as we would need to escape rather stringent observational constraints, although it has certainly been considered [674–676]. Generally it is far easier to introduce new massive particles in the more-unknown ultraviolet regime. Secondly, we will need a *new* global axial symmetry, since we have already disposed of the $\mathrm{U}(1)_{\mathrm{A}}$ portion of the quark flavor symmetry. Third, we require that currents charged under this new symmetry will have an axial anomaly, so that we can recreate the above solution.

That last requirement means that we probably will want the quarks to be charged under this new axial symmetry, so that we can use them in the familiar triangle diagram. However, we need to avoid the fact that quarks do indeed have mass terms, and we are looking for a chiral symmetry. To avoid this roadblock, we look at the origin of these mass terms—Yukawa interaction terms, with the Higgs field. If we ensure that the Higgs field is also charged under our new symmetry, then we can set up our theory so that the chiral transformations of the quarks can be absorbed by an equivalent transformation of the Higgs. This demonstrates explicitly why we need a *new* symmetry—the Higgs field is not charged under the strong force, and so we couldn't exploit this idea previously.

The Higgs mechanism requires that the Higgs field acquires a non-zero vacuum expectation value in order to generate the masses of the weak gauge bosons. However, when examining the quark condensates in Eq. 4.8, we saw that they necessarily could not respect an axial symmetry. Similarly, the fact that the Higgs has a non-zero vacuum expectation value requires this new symmetry to be either spontaneously or explicitly broken. In the case of the $\mathrm{U}(1)_{\mathrm{A}}$ symmetry, we found that it was not spontaneously broken but explicitly broken (after a long discussion on the QCD vacuum structure). For our new symmetry, we cannot rely on such arguments to save the day, and so we have to conclude that it will be spontaneously broken. The necessary result of that, though, is the existence of a pseudo Nambu-Goldstone boson in the particle spectrum. This important consequence was not noticed at first by Peccei and Quinn, and was subsequently pointed out by Steven Weinberg and Frank Wilczek [677, 678].

Regarding the name of this new particle, the folklore is that Wilczek named the particle after a cleaning product he saw in the supermarket called 'axion', because it cleans up the strong-CP problem. I always had trouble believing this story, since the name so strongly evokes a reference to the axial $\mathrm{U}(1)_{\mathrm{A}}$ symmetry which is responsible for the whole problem. It turns out that the folklore is half-correct—Wilczek had indeed found the name 'axion' in a supermarket, but held onto it for several years before realizing that it

fit perfectly for this new axial particle [679]. Weinberg originally called the new particle the 'Higglet' but thankfully deferred to Wilczek's suggestion.

4.2.2 The PQ mechanism

The solution of Peccei and Quinn was to posit a new $\mathrm{U}(1)_{\mathrm{PQ}}$ axial global symmetry, satisfying the above considerations. Before we discuss explicit models for extending the Standard Model to include this, though, let us look at how such a mechanism is able to solve the strong-CP problem, and assume that we have a model satisfying the above considerations. After the $\mathrm{U}(1)_{\mathrm{PQ}}$ symmetry has been spontaneously broken at some scale f_a, we are left with a Nambu-Goldstone boson a corresponding to the angular degree of freedom of the $\mathrm{U}(1)_{\mathrm{PQ}}$ symmetry. This means that a bispinor Ψ (such as a quark) with charge Q under the $\mathrm{U}(1)_{\mathrm{PQ}}$ symmetry transforms like,

$$\Psi \to e^{iQ\gamma^5 a/f_a}\Psi \ , \tag{4.28}$$

where the γ^5 factor ensures we have an axial symmetry. The Lagrangian, now including the axion, is generically written as,

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \frac{g^2}{32\pi^2}\left(\bar{\theta} + N\frac{a}{f_a}\right) G_a^{\mu\nu}\widetilde{G}^{a\mu\nu} - \frac{1}{2}\partial_\mu a \partial^\mu a + \mathcal{L}_{\mathrm{int}}\left[\frac{\partial_\mu a}{f_a}, \Psi\right] \ . \tag{4.29}$$

The axion-gluon interaction term arises from the requirement that there be an axial anomaly, and the *anomaly coefficient* N is a model-dependent parameter which describes how the axion is embedded into the Standard Model. The third term is the axion kinetic term, and the final terms are interaction terms with the standard model—it is a generic result of spontaneous symmetry breaking that we are left with derivative couplings, and it is important to note that the size of these interactions are suppressed by the vacuum scale f_a.

The axion-gluon term sources an effective potential for the axion, although this calculation is somewhat nontrivial. One method uses the dilute-instanton-gas approximation, where n instantons and $\bar{n}$ anti-instantons with $n - \bar{n} = 1$ and widely separate centers are summed up, leading to vaccuum energy $E(\theta) \propto \cos(\theta)$. Thus, the axion potential in this

approximation becomes,

$$V_{\text{eff}} \sim -K \cos\left(\bar{\theta} + N\frac{\langle a\rangle}{f_a}\right) \ . \tag{4.30}$$

The minimum of this potential is clearly realized at,

$$\langle a\rangle = -\frac{f_a}{N}\bar{\theta} \ . \tag{4.31}$$

All we need to do, then, is use our shift symmetry to transform $a \to a - \langle a\rangle$, and the CP-violating $\bar{\theta}$ term in the Lagrangian is driven to zero. Since it is QCD instanton effects which source this potential, it is often said that the axion potential is 'tilted' by QCD, allowing the dynamic cancellation of the $\bar{\theta}$ term and the solution of the strong-CP problem.

4.3 Axion phenomenology and detection

In this section I will briefly examine the phenomenology of the QCD axion. Generally, I will use the phrase 'QCD axion' and 'axion' interchangeably—however, confusion arises when one begins to consider other axion-like-particles (ALPs) which may have similar properties to the QCD axion, but are not invoked specifically to solve the strong-CP problem [680–682].

Even before specifying a specific model, however, we can get a generic understanding of both the QCD axion mass, and its coupling to standard model particles. The mass of the axion can be read off the quadratic term in the expansion of the potential 4.30. The coefficient K comes from QCD instanton contributions, and while I will not explicitly show the calculation here, it can be found to be,

$$K = \frac{m_u m_d}{m_u + m_d} m_\pi^2 f_\pi^2 \,, \tag{4.32}$$

in terms of the masses of the up and down quarks, the pion mass, and the pion decay constant f_π. After expanding the potential and examining the coefficient of the quadratic term, we then find the axion mass,

$$m_a \simeq 5.7 \text{ eV} \left(\frac{10^6 \text{ GeV}}{f_a} \right) \,. \tag{4.33}$$

The linear dependence of mass on the vacuum scale leads to the so-called *QCD-band* for the QCD axion. It is also convenient to parametrize axion models by the way in which they couple to photons. After some calculation, the axion-photon coupling can be found to be,

$$g_{a\gamma\gamma} = -\frac{\alpha_{\text{EM}} N}{2\pi f_a} \zeta, \tag{4.34}$$

where α_{EM} is the electromagnetic fine-structure constant and the factor ζ is given by,

$$\zeta \equiv \frac{E}{N} - \frac{2}{3} \frac{4m_d + m_u}{m_u + m_d} \,. \tag{4.35}$$

The second term in the definition of ζ ensures that there is no axion-pion mixing [683], and E/N is the ratio of the electromagnetic axial anomaly to the color axial anomaly, which is fixed by the particular extension of the standard model under consideration.

4.3.1 Benchmark models

The first and simplest model for the axion was the original Peccei–Quinn–Weinberg–Wilczek (PQWW) model [354, 355, 677, 678]. Here, an additional Higgs doublet field was introduced, so that independent chiral transformations of the quarks could be absorbed. Schematically the Yukawa terms in this model are proportional to,

$$\bar{Q}_{\mathrm{L}}\Phi_1 u_{\mathrm{R}}, \qquad \bar{Q}_{\mathrm{L}}\Phi_2 d_{\mathrm{R}} \ . \tag{4.36}$$

In the simplest model, both the Higgs doublets would have similar symmetry breaking structures, breaking at the usual electroweak scale. After breaking, and some careful basis selection, we are left with four real scalars—a massive Standard Model Higgs boson, the Z boson mass, a second massive Higgs boson (ideally, quite large), and finally an angular degree of freedom to become the axion [684].

This model is a relatively painless extension of the Standard Model, but it was quickly realized that it was flawed. When the axion scale f_a is at the electroweak scale, the couplings to Standard Model particles are too strong and the PQWW model is ruled out by laboratory searches [685–688].

In order to escape observational constraints, we would like our axion to have a vacuum scale much larger than the electroweak scale. To do this, we need to introduce new scalar fields which carry PQ charge, but not electric or weak charge. Such models are known as *invisible axion* models. There are two benchmark examples of this kind of UV extension. The first is the Dine–Fischler–Srednicki–Zhitnitsky (DFSZ) type [689, 690], where Standard Model quarks carry PQ charge. The second is the Kim–Shifman–Vainshtein–Zakharov (KSVZ) type [691, 692], where we introduce new colored fermions.

For the DFSZ model, we still have two Higgs doublets as in the PQWW case, but we introduce an additional PQ-charged scalar field φ which is otherwise a Standard Model singlet. This field breaks at a much higher scale, leading to the axion as its Nambu-Goldstone boson. We are left with axion-Standard Model couplings, because fermions are given PQ charge through terms such as,

$$\lambda_{\mathrm{H}}\varphi^2\Phi_1\Phi_2 \ . \tag{4.37}$$

Notably, for the DFSZ model, the ratio $E/N = 0.36$.

Meanwhile, for the KSVZ model, we introduce new heavy quarks $Q_{\mathrm{L,R}}$ which hold

PQ charge, but are strong triplets and electroweak singlets, as well as a new scalar φ as above. Then we have terms such as,

$$\phi \bar{Q}_{\mathrm{L}} Q_{\mathrm{R}} + \mathrm{h.c.}\ , \tag{4.38}$$

where the coupling constant is a free parameter of the theory. The heavy quarks can be integrated out at some UV cutoff, leaving only the axion-gluon interaction term, and no Standard Model couplings. Here, the ratio $E/N = -0.97$.

4.3.2 Axions as dark matter

It was quickly realized that the axion would also serve as an interesting dark matter candidate [357–360, 684, 693, 694]. One of the main hurdles of new dark matter ideas is developing a mechanism for producing the correct abundance of cold dark matter. In fact, sufficiently large axions can be ruled out immediately, since they would thermalize in the early universe and impact CMB and BBN observations. From a variety of cosmological measurements, the axion mass then has an upper bound $m_a < 0.5\mathrm{eV}$ [695].

However, smaller axions can be produced from non-thermal mechanisms to make up the correct abundance of dark matter. In addition, they can act as a proper cold dark matter candidate, since for small vacuum scale they are coupled only weakly to the Standard Model. There are a number of formation mechanisms [684], but we will focus here on the most well-known (and model-independent) mechanism—the *misalignment mechanism* [357–359].

The basic premise of this mechanism is quite simple. We assume that in the early universe (i.e., at high temperatures), the axion field has some initial angle θ_0 differing from the final CP-conserving value. Once the cosmological fluid cools sufficiently, the axion field will 'roll down' and oscillate around its final value. The larger these oscillations, the more axions are produced—recall that when transitioning from a classical field theory to a quantum field theory, 'field value' is promoted to 'particle number'.

The exact details of the temperature-dependent axion potential are rather complicated, but we can get a basic idea by truncating the full axion potential at the mass term. Then, the axion equation of motion in an expanding universe is,

$$\ddot{a} + 3H\dot{a} - \frac{1}{R^2(t)}\nabla^2 a + m_a(t)^2 a = 0\ , \tag{4.39}$$

where we have used $R(t)$ for the FLRW scale factor to avoid confusion with the axion field. At high temperatures $m_a(t) > 3H$ the oscillations of the field are frozen out. Once $m_a(t) < 3H$, however, the axion field begins to oscillate, with a damping term due to the expansion of the universe. The number of axions produced depends on the size of these oscillations (i.e., the initial misalignment angle), as well as the timescale of these oscillations. The energy density of the universe in axions can be estimated [684] as,

$$\Omega_a h^2 \simeq 0.4 \left(\frac{6\mu\text{eV}}{m_a}\right)^{7/6} \langle\theta_0^2\rangle \ . \tag{4.40}$$

We see that in order to produce the correct dark matter abundance, we must carefully balance the axion mass with the initial misalignment value.

In the scenario where the PQ symmetry breaking occurs before inflation, a single causal patch containing some random value for θ_0 would be expanded to horizon size, so θ_0 would be homogeneous in our universe and we can consider it effectively as a random angle. However, if the PQ symmetry breaking occurs after inflation, there would be many patches in our universe with differing initial angles. In this case, we must take the average field value over these patches, $\langle\theta_0^2\rangle = \pi^2/3$, leading to the so-called 'classic' axion dark matter window at $m_a \sim 80 - 400\mu\text{eV}$.

However, the post-inflationary scenario is complicated by the possibility of topological defects in the axion field [476, 477, 693, 696–702]. When the PQ symmetry is first broken, field configurations known as *strings* can form, where the axion field wraps around values in $[0, 2\pi f_a]$. These strings can either be closed in loops, or open, going across an entire Hubble patch. The dynamics and decay of axionic strings can have extremely significant cosmological effects, and is a much-debated and active area of research today [15].

In addition, when the axion picks up its mass, topological defects known as *domain walls* can form. These arise from the axion effective potential Eq. 4.30, where we can see that there is actually a discrete $\mathbb{Z}_N$ symmetry left over once the PQ symmetry has broken. Uncorrelated neighboring patches may then choose different vacua from the N choices, causing topological sheets known as domain walls to form between them, with strings as their boundaries. Even for $N = 1$, domain walls can form, where they interpolate between identical vacua by winding once around the bottom of the axion potential. Notably, theories with $N > 1$ are stable and the energy density in these domain walls can easily overclose the universe—this is known as the *domain wall problem*.

Another wrinkle arises in the post-inflationary scenario. Since the initial misalignment angle varies across patches, the axion number density varies as well. This can lead to sizeable inhomogenities in the axion distribution. Since the axion free-streaming length is relatively short, these inhomogenities are not erased before matter-radiation equality, and large-scale features known as *miniclusters* could form [15, 703–706]. In addition, since axions can have an attractive self-interaction, it is possible for *axion stars* to form over the timescale of the universe [707, 708] The possible existence of these structures, of course, has significant impacts on detection prospects for axion dark matter. Finally, there is also the interesting possibility that the axion population could form a Bose-Einstein condensate, leading to novel phenomenology such as halo rotation [709, 710].

4.3.3 Observation

Although QCD axions can interact with a variety of Standard Model particles, constraining the coupling of axions to photons is by far the most popular. I will very briefly outline some of these searches here. Since the photon coupling Eq. 4.34 is a function of vacuum scale, it is conventional to plot axion (or more general ALP) constraints as a function of mass and photon coupling, as in Fig 4.5. I will only give a brief summary of these constraints, and more complete reviews can be found in e.g. Ref. [15].

Laboratory experiments

First, we have laboratory experiments which aim to detect the axion. *Helioscopes* [712–715], such as CAST [716–719], Sumico [720] and the future IAXO [721, 722], aim to detect axions from the sun by converting them to x-rays inside large magnetic cavities. Since helioscopes rely on axions produced in the sun, they are independent of dark matter considerations and can exploit well-understood solar physics [723]. In addition, WIMP detectors such as LUX [724], XENON100 [725] and PandaX [726] are also able to function as axion helioscopes.

Another type of laboratory search are the *light-shining-through-wall* experiments, such as ALPS [727] and OSQAR [728]. These experiments use strong magnetic fields to convert lasers to axions, aimed at an opaque barrier, and attempt to detect reconversion to photons on the other side.

Haloscopes, such as ADMX [729], work similarly to helioscopes but aim specifically to detect dark matter axions. The standard technique involves converting axions to photons

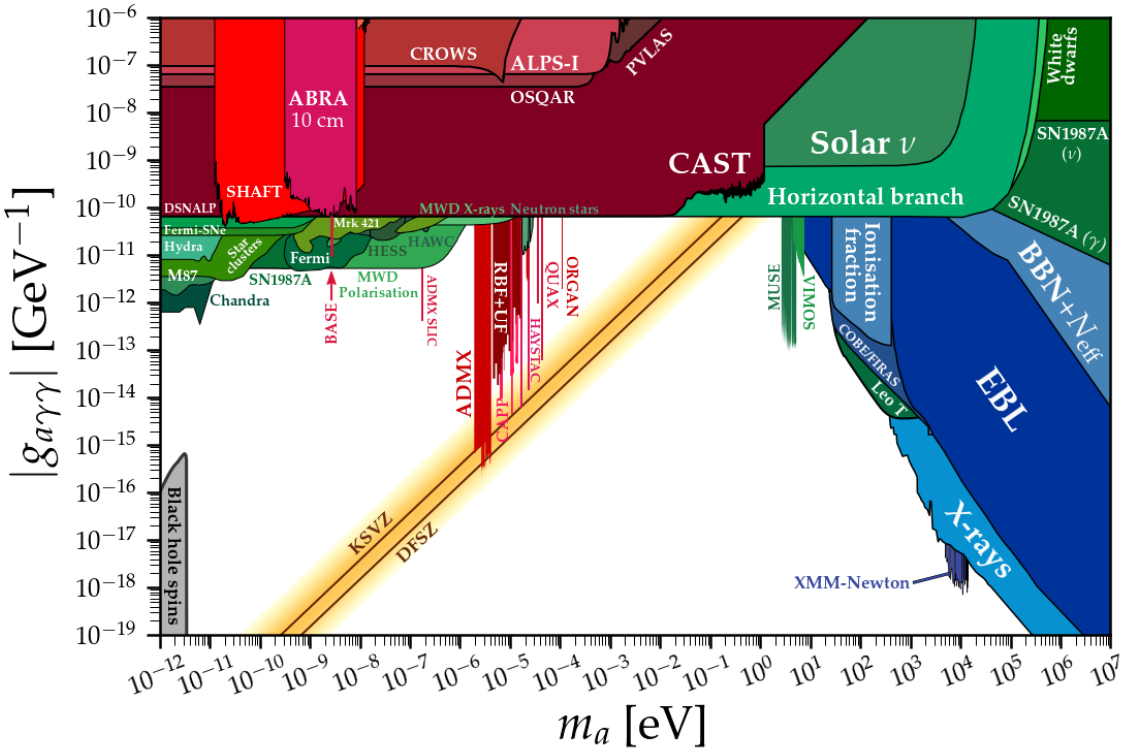

Figure 4.5: Axion-photon limits [711]. Astrophysical limits are shown in green, direct detection bounds in red, and cosmological constraints in blue. The QCD band is in yellow, and shows explicitly the two benchmark models—the width of the band arises from various choices of model-specific parameters.

in large magnetic fields within a resonant cavity, which can be slowly sweeped over a range of frequencies to reach QCD axion sensitives over a narrow band. There are many current or planned haloscopes, with relatively diverse measurement techniques and sensitivity ranges. A short list includes MADMAX [730], HAYSTAC [731], ORGAN [732], DMRadio/ABRACADABRA [733–735] and CASPEr [736] .

Astrophysical bounds

Next there are the astrophysical bounds on axions. Since there is quite a large wealth of these, I will just briefly mention a few of the most important bounds.

Axions which are formed within stellar cores can stream out of the star, depleting the stars' energy and generally leading to shorter stellar lifetimes. By counting the number of horizontal branch stars in globular clusters, we can get an estimate of the time spent burning Helium and so place stringent constraints on the axion-photon coupling [737].

Similarly, production of axions with mass lower than the solar core temperature may heat up the solar interior, producing more neutrinos [738]. Supernovae also place constraints on axion-like particles, since axions produced in them could convert to γ-rays in the magnetic field of the Milky Way [739–741].

A different kind of astrophysical constraint exploits black hole spins [742–754]. The idea here is that light bosonic fields can form bound states around black holes. If the field has Compton wavelength around the size of the black hole ergosphere—an area surrounding spinning black holes in which it is possible to extract energy from the black hole spin—then the bosonic fields can spin down the black holes. The mechanism is known as black hole superradiance, and observations of black hole spins allows the constraint of very light axions, with the caveat that there remains some disagreement in the literature over these bounds [751–755].

Cosmological bounds

Finally, there are cosmological bounds on axions, outside those already discussed for the QCD axion as dark matter. Relatively large axions are ruled out by the fact that they would decay to UV photons and create excess ionization before reionization [756] began at around $z \sim 6 - 15$ [757]. In fact, decay of axions would generally contribute to the the extragalactic microwave-to-γ background [758], or to optical searches in galaxy clusters [759]. Finally, there are the ever-present early universe bounds—axions decaying to photons before recombination are constrained by spectral distortions to the CMB, by constraints on the effective number of neutrino species, and by BBN processes [757].

The layperson box: axions Previous — Next

In the previous box (page 112), I introduced a strange problem in the strong force. The quarks were not aligned with the winding and unwinding of the strong force because the weak force was aligning the quarks in different ways. This was a purely theoretical problem, because no such effects were noticed in nature for the strong force.

At the end of the day, the toolbox of the particle physicist has essentially one tool in it—suggest a new particle. Of course, if you look at the page count of this chapter you would realize that this is not always as trivial a process as I make it seem. But that is exactly what they did here. In a rather clever way, they constructed a hypothetical particle which interacted with the twisting of the strong force fields in a kind of opposite way so that they would have no effect on the quarks anymore. They called this particle the 'axion', related to the word 'axis', about which something spins (and also, famously, a kind of detergent from Texas called Axion).

In order to introduce a new particle, however, you have to be very careful. Humans have been searching for new particles both in particle colliders and from outer space for almost a century, so you need to explain how it has not been seen yet. It has to be invisible, or at least, only rarely interact with any other particles at all. This is again a somewhat tricky task, and as a result there is a wide spectra of axion models with quite similar predictions, but different setups.

Invisible particles, however, are exactly what dark matter searchers are after. Could the axion also be the dark matter then? The answer to that depends on the axion properties—most importantly, the mass of the particle (very light, usually), and how rarely it interacts.

Axions which interact too often are problematic because they will be produced inside the cores of stars, where even rare interactions happen frequently because of the density and heat. This will drain energy from stars, causing them to live shorter than observed. We can also build labs on Earth to detect the axions, using huge 'telescopes' of giant magnetic fields tuned exactly to the axion. These searches either look for axions from the sun, or axions as the ambient dark matter. Finally, the early universe was also hot and dense, so sufficiently massive axions back then would produce too much light (and muck up various other predictions as well)

However, despite all those constraints (and more), the vast majority of axion models are still totally viable, making them a very popular dark matter candidate today. Because the axion has to be very light and rarely-interacting in order to solve the strong force problems, it is very difficult to detect—plus, we haven't been looking quite as long as with other dark matter candidates. I suspect there are a lot of physicists out there who will secretly tell you they think the dark matter will turn out to be axions.

5
Companion axions

i dunno, but man
every time i think those little guys hit the ceiling, they bust through it
almost makes you wish there was extraterrestrial life out here just so they
could peek in and get a load of this shit.
—Jupiter Icy Moons Explorer, *20020: The Future of College Football*

In the previous chapter I explored in detail the landscape of QCD vacua. We saw that the 'true' vacuum was the so-called θ-vacuum, a superposition of the topologically distinct vacua that can be reached by instanton processes. The consequence of this vacuum was to induce a CP-violating term in the Standard Model, necessitating the PQ axion (and accompanying new UV physics), which dynamically sends this term to zero. In this chapter, we will see that 'true' vacuum actually contains more states than the ones we examined previously—when we include gravity in the mix, we actually find that there is an 'independent direction' of states we can reach. As a result, there are two theta angles, spoiling the single axion solution, since the minimum of the potential no longer coincides with the cancellation of the CP-violating terms.

A simple solution to this new strong-gravity-CP problem is to introduce an additional axion. However, simple additions do not always entail simple consequences. The bulk of this chapter explores the phenomenology of the two-axion system that we called the *companion axion* [2, 3, 760].

5.1 Motivation and theory

The background and details of the strong-gravity-CP problem are not actually the focus of my PhD work, so I will discuss them more qualitatively and leave the gory details to the work of my collaborators in Ref. [760]. Nevertheless, I will do my best to explain the rather-complicated process of including gravity in the QCD vacuum.

5.1.1 The strong-gravity-CP problem

Previously, we examined the BPST instanton, a solution which interpolated between vacua which differed by one winding number. Other instanton solutions are possible, such as instantons which change winding number by more than one. The effects of these instantons, however, are suppressed relative to the BPST contribution.

But there are other kinds of instantons we can consider, if we turn our gaze from pure QCD and consider the other fundamental forces of physics. Electromagnetic theta-terms can also be considered, but the QED vacuum is much simpler than for QCD. Since $\mathrm{U}(1)$ can be thought of as a circle, it is not hard to see why any pure-gauge state can be shrunk to a trivial point on the $\mathbb{R}^3$ background—in topological language, the third homotopy group of the circle is trivial. In physics-language, there is no spectrum of distinct vacua to build a theta-vacuum out of. Meanwhile, there will certainly be a weak-force theta angle, just as there was for QCD. However, the weak force acts separately on left- and right-handed fields, allowing the theta term in the Lagrangian to be easily rotated away.

Gravitational instantons

A similar game can be played for gravity. It is not a new revelation that gravitational instantons may exist [761,762], and that they may source CP-violating terms in the Standard Model similar to the QCD theta-term [763]. Hawking and Gibbons first found a family of gravitational multi-instanton solutions in Ref. [762], and they are not so dissimilar to the instantons encountered in the previous chapter. In Sec. 1.2.3 we saw that gravity can be described as a gauge theory. The situation is somewhat simpler after the Wick rotation $t \to i\tau$ into Euclidean space, when searching for instanton solutions, since the gauge group of gravity will be $\mathrm{SO}(4)$ instead of $\mathrm{SO}(3,1)$. In particular, the Lie algebras satisfy $\mathfrak{su}(4) \simeq \mathfrak{su}(2) \times \mathfrak{su}(2)$, leading to a double covering $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$. This means that there will be topologically distinct gravitational vacua, but now labelled

with a *pair* of integers, (n, m) [762].

At the same time, another particular realization (or rather, subclass) of pure-gravity instantons was found by Tohru Eguchi and Andrew Hanson (EH) in Refs. [764, 765]. In fact, it was shortly realized that this metric was equivalent to the multi-instanton solution of Hawking and Gibbons [766]. The EH instanton has metric,

$$ds^2 = \frac{1}{1 - \frac{a^4}{r^4}} \mathrm{d}r^2 + \frac{r^4}{4} \left[\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2 + \left(1 - \frac{a^4}{r^4}\right) (\mathrm{d}\psi + \cos\theta \mathrm{d}\phi)^2 \right], \tag{5.1}$$

written in the slightly unusual basis where r is a four-dimensional radial coordinate and we have three angular coordinates (θ, ϕ, ψ). An arbitrary integration constant, a, can be thought of as the instanton size. This metric has the property that it is asymptotically locally Euclidean, which means it is Ricci flat $(R = 0)$ and has non-negative action $S_{\mathrm{EH}} \geq 0$. These properties mean the EH metric does indeed describe a topologically non-trivial metric which can be included when integrating over all metric manifolds in the path integral.

It can be shown that pure-gravity instantons cannot mediate anomalies with chiral charge violation, so there is no effect on QCD or axion physics. This is because the condition $R = 0$ for the Ricci scalar implies that the Dirac operator in the background of these gravitational instantons has no non-trivial normalizable zero-mode solutions, even for massless fermions [767]. Although the pair of topological numbers (n, m) do change between topologically distinct vacua, there is no accompanying change in chiral charge.

In the presence of another gauge field, however, the situation may be different, and so we should look at combined gravity and Yang-Mills instantons [768, 769]. Electromagnetically charged EH instantons were explored, for example, in Ref. [770], finding that the PQ axion solution was not changed but that the domain wall problem might be relaxed.

An interesting topological property of the EH metric was explored by my collaborators recently in Ref. [767]. This arises from the singularity at the horizon, $r = a$. The removal of this singularity means that the boundary of the metric at infinity is isomorphic to $S^3/\mathbb{Z}_2 \cong \mathbb{RP}^3$, the real projective space of dimension three. Because the EH background is topologically non-trivial, it is able to support Yang-Mills field configurations which would otherwise vanish in flat spacetime. In Ref. [767], a $\mathrm{U}(1)_{\mathrm{EM}}$ instanton, inspired by the famous *Dirac string* monopole solution, was considered. Notably, there existed non-contractible loops in which fermions picked up a phase proportional to their electro-

magnetic charge, and requiring this phase to vanish implied the quantization of electric charge (a rather remarkable observation!). Since the smallest observed charge is 1/3, this implied the instantons were restricted to charges of multiples of three.

We are interested here, however, in strong force-charged EH instanton solutions. A similar game can be played with colored fields in the EH background, and for analogous reasons to the above, these instantons, called *colored Eguchi Hanson* (CEH) instantons, have topological charge $Q = 3$. The explicit embedding of these gauge potentials in the EH background was explored in Ref. [760].

Including CEH instantons

Accounting simultaneously for both the gravitational and BPST instantons means that a new vacuum angle must be included, leading to its own CP-violating term:

$$\theta_{\text{CEH}} \equiv \frac{3}{2}\theta_{\text{QCD}} + \theta_{\text{EH}} \,. \tag{5.2}$$

The factor $3/2$ comes from the $\mathbb{Z}_2$ identification of the EH spacetime which essentially halves the Euclidean space, combined with the fact that the CEH instantons have topological charge of three. A cartoon of the effect of the CEH instantons is shown in Fig. 5.1.

Calculating the axion potential in the presence of the BPST instantons is already an involved calculation, and it is not made easier by the inclusion of CEH instantons. The CEH instanton zero-modes, needed for calculating transition amplitudes, are presently not well-understood. However, an educated guess can be made based on knowledge of renormalization processes in analogy to the standard case [760], leading to,

$$V(a) = -2K_{\text{QCD}} \cos\left(N\frac{a}{f_a} + \theta_{\text{QCD}}\right) - 2K_{\text{CEH}} \cos\left(N_g\frac{a}{f_a} + \theta_{\text{CEH}}\right) \,. \tag{5.3}$$

The ratio $\kappa \equiv \frac{K_{\text{CEH}}}{K_{\text{QCD}}}$ is estimated as,

$$\kappa \sim \frac{16}{25}\left(\frac{2\pi}{\alpha_S(\Lambda)}\right)^2 e^{-\pi/\alpha_s(\Lambda)} \sim 0.06 - 0.4 \,, \tag{5.4}$$

where $\alpha_s(\Lambda = 1\,\text{GeV})$ is the running strong coupling constant. It is interesting to note that since $\alpha \sim 1$ for the strong force, this contribution comes from large sized instantons which are not exponentially suppressed. In contrast, for electromagnetism, $\alpha \sim 1/137$ and so

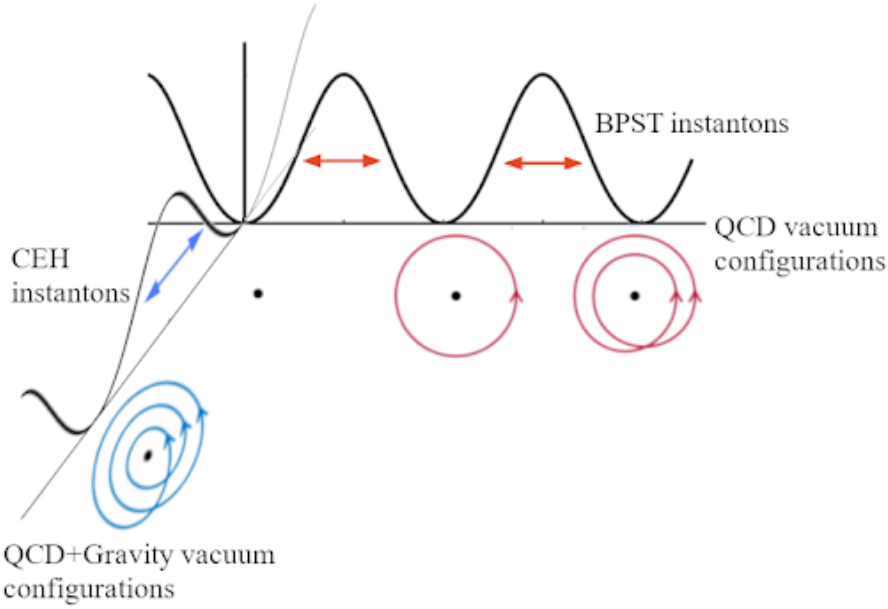

Figure 5.1: This cartoon diagram shows the potential wells which form between the topologically distinct QCD vacuum configurations, and the distinct colored gravitational configurations. I have added the CEH instanton wells on a second orthogonal axis to highlight that the CEH vacuum angle is *a priori* unrelated to the usual theta angle, and so the true vacuum superposition will need to sum over the 'grid' of vacuum states. In addition, the QCD-gravity vacuum configurations are shown to change winding number by 3 in the diagram.

this ratio is significantly suppressed [770] and the axion solution is protected. In the same way as in standard QCD, the new parameter θ_{CEH} would affect the neutron electric dipole moment—to calculate this, one would need to write the new effective pion-nucleon field theory accounting for the addition of the second theta angle in the effective pion-nucleon interaction.

This potential has large implications for the axion solution to the strong-CP problem. We can't merely ignore this new angle, since it will impact the neutron electric dipole moment as well as disrupting the standard axion solution. At the minimum of the potential, we can clearly see that the axion state will no longer cancel out either of the theta angles. Even if the axion state is chosen arbitrarily to cancel one of the angles, the other will be left nonzero. Thus, the standard, single axion case is no longer able to solve what is now the *strong-gravity-CP problem*.

5.1.2 The companion axion solution

We have seen that when gravity is properly incorporated into the QCD instanton background, the standard Peccei-Quinn solution—which proposes a single additional $\mathrm{U}(1)_{\mathrm{PQ}}$ symmetry and so a single resulting axion—is no longer able to resolve the strong-CP problem. The simplest way to remedy this, perhaps unsurprisingly, is to extend this symmetry into a $\mathrm{U}(1)_{\mathrm{PQ}} \times \mathrm{U}(1)'_{\mathrm{PQ}}$ symmetry. We will similarly posit that the new symmetry is spontaneously broken and anomalous, and so we will have two axions, a and a'. In this basis, which we will call the *vacuum basis* for convenience, the effective potential for the axions is mixed:

$$V(a, a') = -2K \cos\left(N\frac{a}{f_a} + N'\frac{a'}{f'_a} + \theta\right) - 2\kappa K \cos\left(N_g\frac{a}{f_a} + N'_g\frac{a'}{f'_a} + \theta_g\right) . \quad (5.5)$$

Here all gravitational contributions are now labelled with a subscript g for convenience. Now that we have an additional degree of freedom, the minimum of this potential will occur at axion field values which are able to cancel both CP-violating terms (so long as we require $NN'_g \neq N_g N'$). Since the axions are mixed in the potential, we need to change basis to the mass eigenbasis, given by,

$$a_1 = a\cos\alpha - a'\sin\alpha \quad (5.6)$$

$$a_2 = a\sin\alpha + a'\cos\alpha, \quad (5.7)$$

where the mixing angle α is defined explicitly as,

$$\tan 2\alpha = \frac{2\epsilon\left(NN' + \kappa N_g N'_g\right)}{\left(N^2 + \kappa N_g^2\right) - \epsilon^2\left(N'^2 + \kappa N_g'^2\right)} . \quad (5.8)$$

In the above we have defined the handy parameter,

$$\epsilon \equiv f_a/f'_a, \quad (5.9)$$

as the ratio of vacuum scales for the two axions. Without loss of generality we can assume that a_1 corresponds to the heavier axion, so that $\epsilon \leq 1$. The quadratic terms in the potential will give us the axion masses, although the exact form of the terms is not exceptionally

pretty, given by,

$$m_1^2 = \frac{\Delta m^2}{2} + \frac{K}{f_a^2}\Big[\left(N^2 + \kappa N_g^2\right) + \epsilon^2\left(N'^2 + \kappa N_g'^2\right)\Big]. \tag{5.10}$$

where the mass-squared difference $\Delta m^2 = m_1^2 - m_2^2$ is given by,

$$\Delta m^2 = \frac{2K}{f_a^2}\left[4\left(NN' + \kappa N_g N_g'\right)^2\epsilon^2 + \Big(\left(N^2 + \kappa N_g^2\right) - \epsilon^2\left(N'^2 + \kappa N_g'^2\right)\Big)^2\right]^{1/2}. \tag{5.11}$$

Since we are all theorists here, it is helpful to make some approximations. First, we assume that all the anomaly terms are roughly the same order. Then we will take terms only to the lowest order in the relative strength of the gravitational instanton contribution κ. In this limit the axion masses are given by,

$$\begin{aligned} m_1^2 &\sim 2K/f_a^2 \\ m_2^2 &\sim \kappa\epsilon^2 m_1^2\,. \end{aligned} \tag{5.12}$$

We see that the first axion has mass roughly equivalent to the mass of the single axion in the standard PQ scenario, whereas the second axion has a smaller mass governed by the difference in scales and the exact value of the relative gravitational contribution.

It is important to note again that we now are investigating a two-parameter solution to the strong-CP problem. Whereas the standard PQ solution existed on a single band given by the mass-vacuum scale relation, we now have two vacuum scales. This means that companion axions which solve the strong-CP problem exist on a two-dimensional plane, rather than a narrow strip on the traditional mass-photon coupling plot. This can be seen in Fig. 5.2—the two axes are given by the first vacuum scale and then the relative scale ϵ. The mass of the first axion roughly depends only on the first vacuum scale, and so can be plotted also on the x-axis. Meanwhile, the mass of the second axion is conserved along diagonal bands in the plane. Although this plot is strikingly different from the usual axion-photon coupling plot, it is much more useful for the phenomenology of the companion axion. We should also note that while the usual axion-photon coupling plot shows constraints on all axion-like particles, not just QCD axions, all areas on our two-dimensional 'QCD-plane' correspond to companion axions which solve the strong-gravity-CP problem.

We are now ready to investigate more carefully the phenomenology of these companion

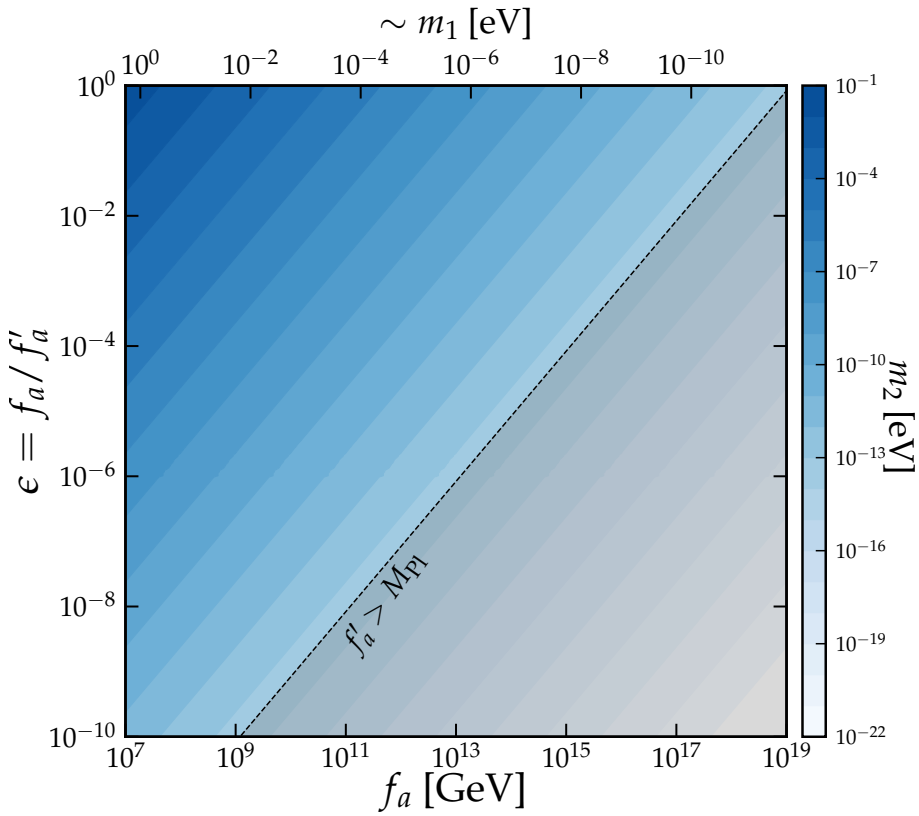

Figure 5.2: This figure shows the parameter space of the companion axion. The vacuum scale of the first axion and the relative vacuum scale are the two axes, meaning that the horizontal axis roughly corresponds to the first axion mass, while the second axion mass is invariant along diagonal lines. In contrast to the familiar axion-photon diagram where only axions in the QCD band solve the strong-CP problem, *all* points on this plot correspond to companion axions which solve the strong-gravity-CP problem. The grey triangle in the bottom-right corner arises from the requirement that the higher vacuum scale be smaller than the Planck scale, $f_a' < M_{\rm Pl}$.

axions. There are two points which demonstrate why such a model may have rich and unique characteristics.

First, the two axions are necessarily coupled together in order to solve the strong-CP problem. That means that if one of the two axions in a particular realization is ruled out by observation, its companion is also necessarily ruled out. In the case of generalized multi-axion models, such as axiverse scenarios, this is not the case. This can be seen most clearly in Fig. 5.3 , where we have fixed the value of ϵ and shown with dotted lines the 'connection' between the two axions which is required in order to successfully solve the new strong-CP problem.

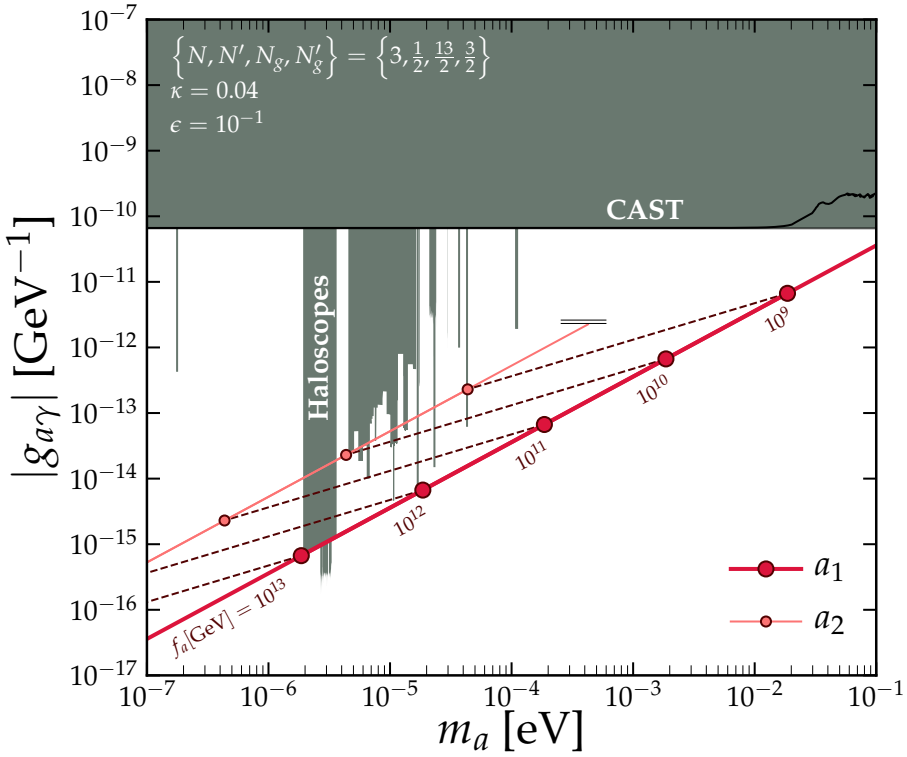

Figure 5.3: The photon-mass relation is shown for two axions with fixed $\epsilon = 10^{-1}$, so that the only free parameter is the mass of the larger axion. The dotted lines indicate the connection between the larger axion and its companion smaller axion required to solve the strong-gravity-CP problem. In areas where haloscopes reach the lighter axion band, the *combined* model with both axions is constrained, and similarly with the stellar cooling/helioscope [771, 772] bounds for the heavier axion. The haloscope bounds are combined from Refs. [719, 729, 732, 773–786]. This figure is intended only to illustrate the unusual coupling characteristics of the two axions—in reality, these bounds must be revised to account for novel phenomenology of the companion axion, which we will do below.

Second, the coupling of the two axions means that there will be novel effects, like oscillations, due to mixing of axions in the vacuum basis. This has important ramifications not just for experimental observation, but for the cosmological and early-universe behavior of companion axions.

5.2 Companion axions as dark matter

In this section, we reinvestigate the companion axion as a dark matter candidate [3]. In the previous chapter, we discussed the standard misalignment mechanism for axions, in particular focusing on the two cases where the PQ symmetry was broken either before or after inflation. When the axion was broken before inflation, the misalignment angle could take any value, but only corresponded to the correct dark matter abundance for a specific axion mass. Meanwhile, if the PQ symmetry was broken after inflation, the average misalignment angle was fixed, forcing a relatively narrow window of axion masses which were able to give the correct dark matter abundance.

It is again the case here that the companion axions can only comprise the correct dark matter abundance for a specific portion of the parameter space. Now that we have two degrees of freedom to tinker with, the physics is more complicated. For the moment, we will again ignore axion production from the string-domain wall network, which requires a numerical treatment even in the standard axion case [787–794], and focus on the classic misalignment mechanism. For the companion axion, however, there are three scenarios for the PQ symmetry breaking:

(I) Both PQ symmetries are broken before the end of inflation (i.e. both axions are *pre-inflationary*). Both angles $\theta_{1,2}$ can take any value from $-\pi$ to π.

(II) The a' symmetry is broken before the end of inflation, while a is post-inflationary. In this case, θ_2 can take any value but $\theta_1 = \pi/\sqrt{3}$ takes the stochastic average, just as in the standard case.

(III) Both axions are post-inflationary (or, there is no inflation). Both misalignment angles take the stochastic average.

In the pre-inflationary case, just as in the standard case, there is always the freedom to take the initial angle sufficiently close to 0 or $\pm\pi$ to either minimize or maximize the dark matter abundance as necessary. However, this might be considered a kind of artificial fine-tuning, especially since the axion was introduced in the first place to remove such a fine-tuning in the CP-violating parameter.

To demonstrate the 'natural' window for the the dark matter abundance, then, we can introduce a small parameter $\delta = 0.1$ and plot the dark matter distribution for initial angles $\theta_{1,2} \in [\delta, \pi - \delta]$. The specific value of δ is arbitrary, and chosen for demonstrative purposes. The overlap of these two 'natural' bands, in each of the three cases above,

comprises the new preferred window for companion axion dark matter.

5.2.1 The misalignment mechanism

We still need to calculate explicitly the abundance of each of the two axion species, since the relative abundances will be extremely important for both detection and cosmological implications. Again, we will follow the standard misalignment mechanism calculations, but with the additional twist that the companion axions are coupled. We start as before with the zero-mode evolution equations for the two axions as a linearized system of coupled oscillators:

$$\begin{aligned} &\partial_t^2 a + \frac{3}{2t}\partial_t a + M_{11} a + M_{12} a' = 0 \\ &\partial_t^2 a' + \frac{3}{2t}\partial_t a' + M_{22} a' + M_{21} a = 0 \,, \end{aligned} \tag{5.13}$$

where M_{ij} are elements of the axion mass matrix, and we have specified a radiation-dominated universe so that the Hubble damping term is given by $H = 1/2t$. It was noted in Ref. [795] that including the full nonlinearities in the potential [796–800] could lead to a resonant transfer of energy between the two particles. In our case, however, it would only apply in the case where the two axions have very similar masses $m_1 \sim m_2$ but with a somewhat small value of $\epsilon \lesssim 0.2$, which would require relatively contrived tinkering with the anomaly coefficients to obtain.

To be more precise, we must use the thermally corrected axion mass matrix. The thermal axion mass calculation is complicated even in the standard case, and will be even more unwieldy when incorporating the effects of colored gravitational instantons. Nevertheless, we can get an approximate idea by reusing the standard thermal axion mass for the single axion [801],

$$m_1^2(T) = \min\left[m_1^2, m_1^2\left(\frac{\widetilde{T}}{T}\right)^n\right] \,, \tag{5.14}$$

with numerical values $n = 6.68$ and $\widetilde{T} = 103\mathrm{MeV}$. We then assume for simplicity that the thermal mass matrix scales just like the non-thermal companion axion mass matrix,

but with the first axion mass now given by Eq. 5.14 so that,

$$M(T) = m_1^2(T) \begin{pmatrix} 1 & -\epsilon^2 \\ -\epsilon^2 & \kappa\epsilon^2 \end{pmatrix} + \mathcal{O}(\epsilon^4) \,. \tag{5.15}$$

This allows us to decouple the oscillator equations in the mass basis:

$$\partial_t^2 a_1 + \frac{3}{2t}\partial_t a_1 + m_1^2(T) a_1 = 0 \,, \tag{5.16}$$

$$\partial_t^2 a_2 + \frac{3}{2t}\partial_t a_2 + \kappa\epsilon^2 m_1^2(T) a_2 = 0 \,. \tag{5.17}$$

As we would in the standard case, we would like to know the temperatures T_i at which each of the axion fields respectively begins oscillating due to the presence of their mass terms. This occurs at $m_i(T_i) = 3H(T_i)$, giving,

$$T_i = \left(\frac{m_i M_{\rm P} \sqrt{90}}{24\pi^2 \sqrt{g_*^i}} \right)^{2/(n+4)} \widetilde{T}^{n/(n+4)} \,, \tag{5.18}$$

where g_*^i is the number of relativistic degrees of freedom at temperature T_i and $M_{\rm Pl} \simeq 1.2 \times 10^{19}$GeV is the Planck mass. When $\epsilon \ll 1$, this means the temperature at onset of oscillations for the lighter axion is smaller by a factor,

$$\frac{T_2}{T_1} \sim (\kappa\epsilon^2)^{1/(n+4)} \,. \tag{5.19}$$

The oscillation equations are again solved as in the standard case, using the Wentzel–Kramers–Brillouin (WKB) approximation and an average over multiple oscillations, giving the energy density in each species as,

$$\rho_i|_{\rm today} = m_i(T_i) m_i \langle a_i^2 \rangle \left(\frac{T_0}{T_i} \right)^3 \frac{g_{*s}^0}{g_{*s}^i}, \tag{5.20}$$

where $T_0 = 2.7$ K, and $g_{*s}^i = g_{*s}(T_i)$ are the entropy degrees of freedom. The total axion abundance is of course the sum of the two contributions, $\Omega_a = \Omega_1 + \Omega_2$. What is perhaps most interesting is the ratio of the lighter to heavier axion abundances:

$$\frac{\Omega_{a_2}}{\Omega_{a_1}} \sim \frac{\theta_2^2}{\theta_1^2} \kappa^{0.41} \epsilon^{-1.19} \,. \tag{5.21}$$

Unless $\theta_2 \ll \theta_1$ (which might occur in cases (I) and (II), but could perhaps be considered unlikely), the lighter axion dominates the relic abundance here. This will be quite important for the haloscope constraints, as we will later see.

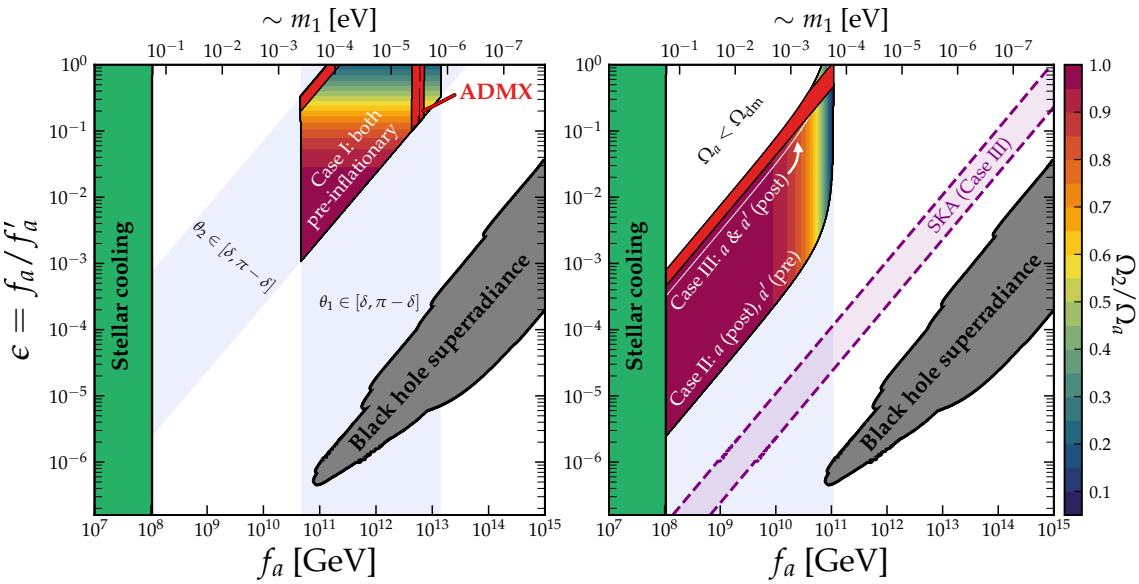

Figure 5.4: These plots show the preferred space for companion axion dark matter produced by the misalignment mechanism. The light blue bands are the 'natural' bands discussed above, and so their intersection represents the preferred dark matter window for companion axions. Within the intersection, the color scale highlights the relative abundances of the two axions. Case I, where both axions are pre-inflationary, is shown in the first plot. Cases II and III are shown in the second plot, although Case III is only a thin line across the diagram since both angles are fixed at the average angle. We can see that allowing δ to take smaller values will allow the preferred regions to grow slightly. A few bounds are shown, which will be discussed shortly—helioscope bounds from ADMX [802] are shown in red, while the projected SKA[803–805] bounds on gravitational waves from axion domain wall collapse are shown in pink. Black hole superradiance bounds are in dark grey.

Finally, it is possible to write the total relic abundance in the scenario $\epsilon \ll 1$ in terms of our 'fundamental' theory parameters:

$$\Omega_a h^2 \sim \Omega_1 h^2 \left(1 + \frac{\theta_2^2}{\theta_1^2} \frac{g_{*s}^1}{g_{*s}^2} \kappa^{\frac{n+2}{2(n+4)}} \epsilon^{-\frac{n+6}{n+4}}\right) \quad (\epsilon \ll 1) \,. \tag{5.22}$$

In the case where $\epsilon \lesssim 1$, the argument proceeds similarly to all of the above, but with the

lighter axion having the thermal mass $m_2(T) \sim \sqrt{\kappa} m_1(T)$. In the same manner we find,

$$\Omega_a h^2 \sim \Omega_1 h^2 \left(1 + \frac{\theta_2^2}{\theta_1^2} \kappa^{\frac{n+2}{2(n+4)}}\right) \quad (\epsilon \lesssim 1) \,. \tag{5.23}$$

Now it is possible for a_1 to dominate the abundance, but it requires both vacuum scales to be similar and the misalignment angles to to be of similar size.

The relative abundances of the two axions are colored in Fig 5.4 within the preferred dark matter windows. Some of the axion constraints which we will derive in the next section are included on these plots for demonstrative value.

5.3 New photon constraints

To get a glimpse into the behavior of the companion axion, we first must derive again the standard photon coupling constraints [2]. In what follows, we take these representative values for the anomaly coefficients:

$$\{N, N', N_g, N'_g\} = \{3, 1/2, 13/2, 3/2\} \ . \tag{5.24}$$

The precise choice of these values is not important for most of the qualitative conclusions below, although the exact shape of some of the plots can shift, and there are some non-trivial cases where interesting cancellations can occur. In addition, to simplify things and minimize the number of parameters, we will assume that the particular UV completion of our model is similar to the KSVZ model [691, 692]. This means that the axions couple only to heavy quarks which carry charges under the $\mathrm{U}(1)_{\mathrm{PQ}} \times \mathrm{U}(1)'_{\mathrm{PQ}}$ symmetry and are otherwise electromagnetically neutral, so that the electromagnetic anomaly is identically $E = 0$ in Eq. 4.34. Finally, we adopt the demonstrative value $\kappa = 0.04$ for the gravitational instanton contribution.

The coupling of axions to photons is extracted in the same way as is done in the standard case [15, 806, 807]:

$$\mathcal{L}_{a\gamma} = \frac{1}{4}\left(a g_{a\gamma} + a' g'_{a\gamma}\right) F_{\mu\nu}\widetilde{F}^{\mu\nu}. \tag{5.25}$$

We find that the second photon coupling is related to the first by,

$$g_{a\gamma} = g'_{a\gamma}\frac{f'_a}{f_a}\frac{N}{N'} = -\frac{\alpha_{\mathrm{em}} N}{2\pi f_a}\zeta, \quad \zeta = \frac{2}{3}\frac{4m_d + m_u}{m_u + m_d} \ , \tag{5.26}$$

where $\zeta \sim 1.92$ again is fixed to avoid axions mixing with the QCD mesons [683]. The first axion-photon coupling constant is the same as that of the standard PQ axion in Eq. 4.34. In terms of the mass eigenstates, the photon couplings are found to be,

$$\begin{aligned} g_1 &= \frac{\alpha_{\mathrm{EM}}\zeta}{2\pi f_a}\left(N\cos\alpha - \epsilon N' \sin\alpha\right) \\ g_2 &= \frac{\alpha_{\mathrm{EM}}\zeta}{2\pi f_a}\left(N\sin\alpha + \epsilon N' \cos\alpha\right) \ . \end{aligned} \tag{5.27}$$

With the model set up, we can now revise the axion-photon constraints for the companion axion. Although we will not redo every one of the many constraints, we have selected

a handful of representative and competitive bounds to reexamine. In each case, we can 'reuse' the observations from the single-axion case, modifying the physics appropriately so that we can recast the bounds into constraints on the companion axion.

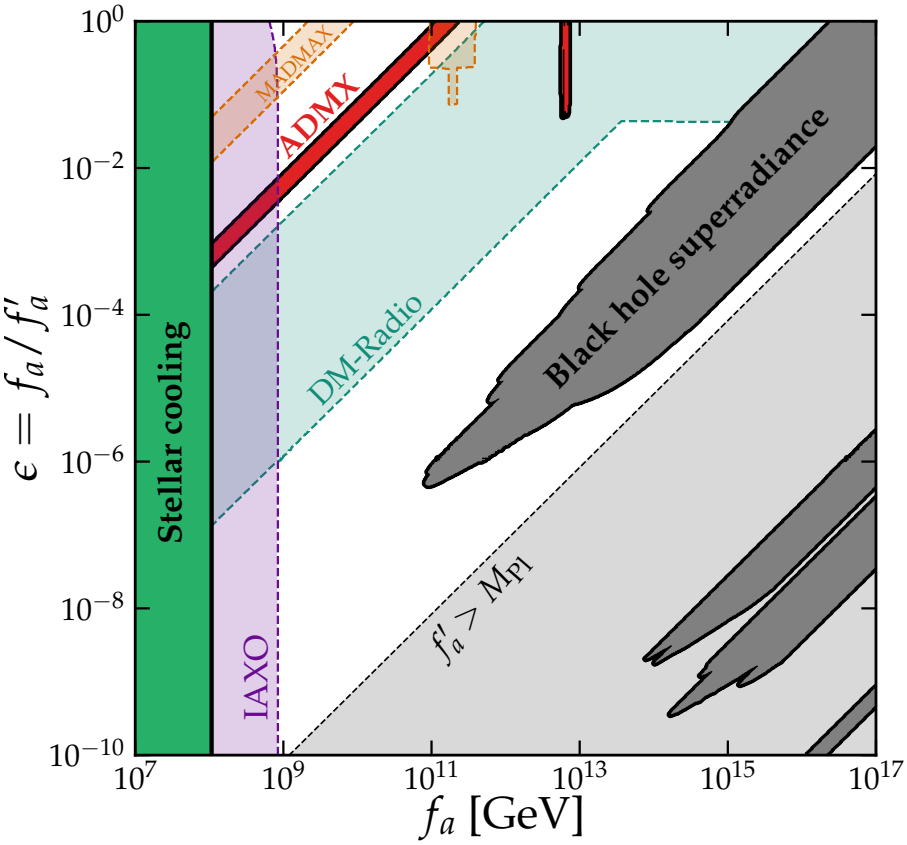

Figure 5.5: Companion axion photon constraints are shown for a representative sample of axion observations. Stellar cooling bounds are shown in green, while the haloscope ADMX [802] is shown in red, and the black hole superradiance bounds [751, 753, 754] are in dark grey. Future haloscopes bounds from MADMAX [730] and DMRadio/ABRACADABRA [734, 735] are shown in dotted orange and green respectively, while the future helioscope IAXO [808] is shown in dashed purple.

5.3.1 Stellar cooling

As discussed briefly in the previous chapter, sufficiently light axions can be readily produced in stellar interiors [739]. There are a number of constraints this places on the axion-photon coupling, but the strongest arise from examining the population of horizontal branch stars in globular clusters. The primary axion production in these stars is the

Primakoff process,

$$\gamma + Ze \to Ze + a \,, \tag{5.28}$$

which works to shorten the lifetimes of these stars. The result of this consideration is the constraint [737],

$$|g_{a\gamma}| < 6.6 \times 10^{-11}\mathrm{GeV}^{-1} \,. \tag{5.29}$$

This constraint can be straightforwardly recast in the companion axion parameter space.

It is convenient when considering axion-photon interactions to move again to a new basis where one of the axions is the combined state which is electromagnetically active, and the second axion is the state which is electromagnetically 'hidden', and so does not interact with the photon [809]. The mixing angle ϕ to this state is given by,

$$\cos\phi = \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \,, \tag{5.30}$$

so that the two axion states are,

$$\begin{aligned} a_{\mathrm{EM}} &= \frac{g_1 a_1 + g_2 a_2}{\sqrt{g_1^2 + g_2^2}} \\ a_{\mathrm{h}} &= \frac{g_2 a_1 - g_1 a_2}{\sqrt{g_1^2 + g_2^2}} \,, \end{aligned} \tag{5.31}$$

and it is easy to see from Eq. 5.25 that the hidden state will have no coupling to the photon field. In this basis, the axion mass matrix has off-diagonal terms,

$$M_{12} = M_{21} = -\Delta m^2 \frac{\cos^2\alpha}{1+\epsilon^2}\Big(\tan\alpha + \epsilon\left(1 - \tan^2\alpha\right) - \epsilon^2 \tan\alpha\Big) \,, \tag{5.32}$$

written as a function of the original vacuum basis to mass basis mixing angle α. As a result, the axion states will oscillate as they travel. This oscillation does not strictly affect the stellar cooling bounds, but as we will see in the following section, will be important for helioscope observations.

Nevertheless, it is only the active axion state a_{EM} that is produced in stellar interiors, which means for practical purposes that the effective coupling is increased from the standard axion case, to $g_{\mathrm{EM}} = \sqrt{g_1^2 + g_2^2}$. In the case where the second axion is significantly

lighter than than the first (i.e., $\epsilon \ll 1$), this means that there will not be significant change on the stellar constraints from the usual case. However, when $f_a \sim f'_a$, the bound is enhanced by a factor,

$$\frac{\sqrt{N^2 + N'^2}}{N} \,. \tag{5.33}$$

The result of this constraint can be seen in Fig 5.5. For most of this parameter space only the larger axion contributes to the constraints, but when $\epsilon \to 1$ there is a very slight strengthening of the bound.

5.3.2 Helioscopes

Helioscopes [712] are Earth-based experiments which measure axions produced in the sun. They consist of a long magnetic bore so that axions passing through the helioscope are converted to photons. When the axions are sufficiently light, $m \lesssim 0.02$ eV, the axion and photon oscillate coherently along the length of the magnet, significantly boosting the ability to detect photons. For larger masses, the characteristic length-scale of the axion-photon conversion is much smaller than the magnetic bore, suppressing the signal since coherence is lost. In this case, the bore can be filled with a buffer gas such as Helium, which gives the photon an effective mass. Then, the momentum mismatch between the photon and axion is made much smaller and coherence can be restored to the oscillation [810, 811].

There are two effects which must be taken into account to convert helioscope constraints into the companion axion space. First, we must account for the axion oscillation described above during the axion's path from the sun to the Earth, which will reduce the flux at the detector [809]. Secondly, once the axions are inside the magnetic field, we now have a three-particle oscillation problem. Although this makes calculations analytically much more difficult, we will argue that the effect is not significant.

For the axions travelling to Earth, the survival probability of the active axion, with energy ω after travelling a distance L is given by the well-known relation,

$$P = 1 - \sin^2\phi \sin^2\left(\frac{\Delta m^2 L}{4\omega}\right), \tag{5.34}$$

in terms of the mixing angle described in Eq. 5.30. In terms of our more 'fundamental'

companion axion parameters, the first sine function can be rewritten as,

$$\sin^2 2\phi = \frac{4\cos^4}{1+\epsilon^2}\alpha\Big(\tan\alpha + \epsilon\left(1-\tan^2\alpha\right) - \epsilon^2\tan\alpha\Big)^2 ,$$

$$\simeq \begin{cases} 4\epsilon^2 & \epsilon \ll 1 \\ \cos^2 2\alpha & \epsilon \simeq 1 \end{cases} . \tag{5.35}$$

For typical values of $L = 1\mathrm{AU}$ and $\omega \sim \mathrm{keV}$, the axion-axion oscillation length is shorter than the Earth-Sun distance for $\Delta m^2 \gtrsim 10^{-12}\mathrm{eV}^2$. In order for the mass difference to be smaller than this, either the companion axions would need to be finely-tuned to the same mass, *or* the mass of the larger axion needs to itself be roughly smaller than this value. The latter occurs for $f_a \lesssim 10^{13}$ GeV, but next-generation helioscopes will still only be sensitive to $f_a \lesssim 10^9\mathrm{GeV}$. So, we can safely say that we are always in the small-oscillation length regime and the second sine in Eq. 5.34 can be averaged to $1/2$. Then Eqs. 5.34, 5.35 together give us the proportion of axions which reach the detector compared to the standard case,

$$P \simeq \begin{cases} 1-\epsilon^2 & \epsilon \ll 1 \\ 1-\frac{1}{4}\cos^2 2\alpha & \epsilon \simeq 1 \end{cases} . \tag{5.36}$$

The number of photons that a helioscope observes actually scales proportionally to $g_{a\gamma}^4$, so we can recast the standard constraints into the companion axion parameter space by multiplying the minimum detectable photon coupling by $P^{-1/4}$.

Next we need to argue that the effects of the three-oscillation problem are not significant. Generally, the coupling between the two axions is much stronger than the coupling between the axion and photon. The axion-photon coupling in a magnetic field is given by $g_{\mathrm{EM}}B\omega$, where B is the magnetic field strength ($\sim 9\ T$ for CAST). For ease we will work in the regime where $\epsilon \ll 1$, but a similar argument holds in the general case. The off-diagonal term Eq. 5.32 parametrizes the axion-axion coupling, and to lowest order in ϵ, can be approximated as $\Delta m^2\epsilon$. Then we can evaluate explicitly the ratio of the axion-axion and axion-photon oscillation terms:

$$\frac{g_{\mathrm{EM}}B\omega}{\Delta m^2\epsilon} \sim \frac{\alpha_{\mathrm{EM}}\zeta B\omega}{4\pi K} f_a' \lesssim 10^{-2} . \tag{5.37}$$

In the final limit, we are using the requirement that the lighter axion has vacuum scale smaller than the Planck scale, $f'_a < M_{\rm Pl}$. Essentially, the oscillation to the photon is suppressed compared to the axion-axion oscillation. Since the companion axions are strongly mixed, then, we can treat the photon conversion just as we would in the standard case, relative to the slightly smaller flux already described above. Although the three-particle oscillation nominally complicates things, these additional effects will be suppressed by the factor in Eq. 5.37.

The result of this calculation are again included on Fig 5.5. The CAST [719] bounds are less sensitive than the stellar cooling bounds, so we show the projected bounds for the future helioscope IAXO [808].

5.3.3 Haloscopes

Haloscopes aim to detect axions which constitute the dark matter halo of the Milky Way. Since we have already calculated the ratio of the misalignment abundance in each of the two dark matter species, it is straightforward to rescale the constraints set by experiments such as ADMX [774] for the standard axion. We also assume here that the axion dark matter does not have additional structure, such as miniclusters, which could impact the detection signal.

The most interesting feature of the companion axion for haloscopes is that for the majority of the parameter space, the dark matter abundance is dominated by the smaller axion. Resonance-based haloscopes are forced to scan slowly over the axion mass range as they search for a signal, and so have only searched so far over a small area near the classic preferred dark matter mass in the standard axion case. For the companion axion, however, the haloscope's optimal region would be searching for the lighter axion, since they comprise the majority of the dark matter. This gives additional motivation to continue haloscope experiments past the classical preferred region, and specifically, to extend the observation to much smaller masses than previously considered.

In addition, the haloscope constraints illustrate well a particular novel feature of the companion axion constraints. Since a detection could not *a priori* know whether it has observed a_1 or a_2, any constraint from non-observation leads to two bands being excluded in the f_a, ϵ parameter space. This is because a particular companion axion solution to the strong-CP problem requires the existence of both axions. These two bands can be seen on the constraint plot Fig 5.5 for the ADMX and future MADMAX haloscopes, where

we have converted the constraints on KSVZ-like axions into a constraint on companion axions, using the relative dark matter abundances Eq. 5.21.

5.3.4 Black hole superradiance

The black hole superradiance bounds [751, 753, 754] can also be straightforwardly adapted to the companion axion case, with some simplifying assumptions, since the calculations are already sufficiently difficult in the standard case. Here we assume that the formation of the axions' bound states around the black holes are not affected by mixing between the two axion species. This assumption is probably safe, except for the region where ϵ is closer to 1, where the coupling between the two axions might make the superradiance process significantly more complicated. A more detailed and dedicated examination of superradiance in the companion axion case would certainly be interesting, but is not the focus of our 'first-pass' over the companion axion constraints. With that in mind, the standard superradiance constraints are easily mapped to the companion axion case, under the assumption that each of the axions form independent bound states which could spin down black holes.

5.3.5 Unique signals

Although the companion axion clearly has rich phenomenology which differentiates it from the standard QCD axion, if some observation were to find a signal, it would not be immediately obvious whether it has detected the single axion, or one of the two companion axions. It is worth showing then, for optimistic future observations, what regions of the parameter space are able to detect *both* axions. This can be done either in one experiment, which could be sensitive enough to find both axions, or in two combined experiments. In Fig. 5.6 we have plotted some of the regions, using relatively ambitious projections for these experiments, including IAXO+ [808, 812, 813], resonant-cavity [732, 814–818] and LC-circuit-based haloscopes [733–735, 776, 780, 819, 820].

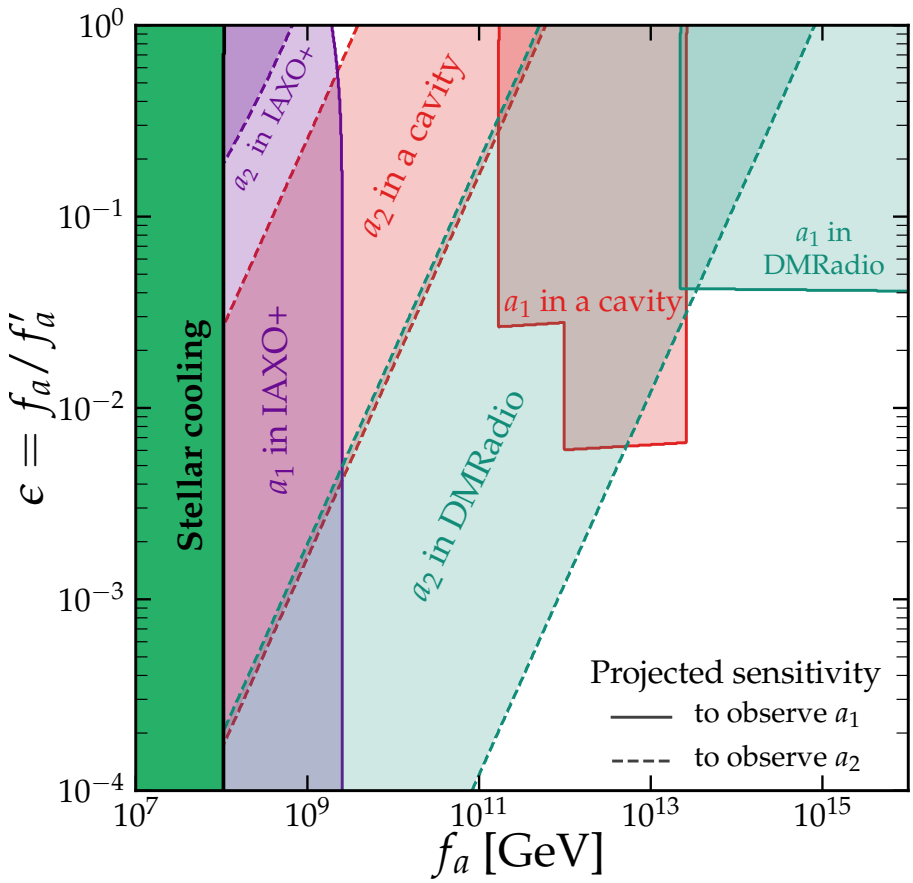

Figure 5.6: Future axion experiments, with some optimistic projections, might allow us to see both the companion axions either within one experiment, or by combining the results of two. This would allow us to more definitely rule in the companion axion model compared to the standard axion.

5.4 New cosmological constraints

We have already shown how the companion axion model can produce the correct dark matter abundance via the misalignment mechanism. However, that is not the end of the story for the early universe—there is a wealth of interesting effects and observations which would be affected by this model [3]. Here we focus on a small handful of particularly interesting or prominent ones, in no particular order.

5.4.1 Isocurvature bounds

In the pre-inflationary case, massless axion fields would undergo large amplitude quantum fluctuations during inflation. However, the energy density of the axion field fluctuations is negligible compared to that of the inflaton field, the dominant energy density in this period. As a result, it is likely not possible to observe the axion field by observing the perturbations in total energy density—otherwise known as *adiabatic* perturbations—which manifest as imprints on large scale structure formation and CMB anisotropies.

However, there is another kind of possible perturbation, known as either *entropy* or *isocurvature* perturbations [821–824]. These perturbations are spatial perturbations in the ratio of axion number density with entropy. In contrast with the total energy density perturbations, it would be possible to see these perturbations in the temperature and polarization of the CMB, since after the isocurvature perturbations are converted to true curvature perturbations, they are uncorrelated with the adiabatic inflaton perturbations. This conversion takes place at the so-called *pivot scale*. If we assume that both the companion axion fluctuations are respectively uncorrelated, the perturbation power spectrum at the pivot scale $k_{\rm low} \simeq 0.002{\rm Mpc}^{-1}$ is given by,

$$\Delta^2_{a_1} = \Delta^2_{a_2}\epsilon^{-2}\frac{\theta_2^2}{\theta_1^2} \simeq \frac{H_I^2}{\pi^2 f_a^2 \theta_1^2} \,, \tag{5.38}$$

where H_I is the Hubble rate during inflation. For Case I, where both axions are pre-inflationary, the ratio of the the total isocurvature power spectrum to the the amplitude of the adiabatic power spectrum is then,

$$\beta = \frac{\Delta^2_{a_1}}{A_s}\left(1 + \epsilon^{-2}\frac{\theta_2^2}{\theta_1^2}\right) \,, \tag{5.39}$$

where the adiabatic amplitude at the pivot scale is $A_s \simeq H_I^2/(\pi^2 M_{\rm Pl}^2 \varepsilon) \sim 2 \times 10^{-9}$, with

ε the usual inflation 'slow-roll' parameter. This gives us the constraint,

$$H_I \lesssim \frac{\sqrt{\beta_\star A_s}\pi f_a \theta_1}{(1+\epsilon^{-2}\theta_2^2/\theta_1^2)^{1/2}} \,, \tag{5.40}$$

where the ratio $\beta < \beta_\star = 0.011$ is constrained by Planck [126] at 95% confidence level. Since this constraint depends on a number of model parameters, it is slightly hard to picture. In Fig. 5.7 we have illustrated the constraint for a particular choice of f_a and ϵ. The band going across the figure gives the correct dark matter abundance, and the color scheme within the band shows the constraint above on the scale of inflation. Of particular note is that when $\theta_1 \sim 0$, inflation must occur at very low scales.

In Case II, the heavier axion is post-inflationary, and the isocurvature bounds on the pre-inflationary axion is the same as in the literature, with $f_a \to f_a'$.

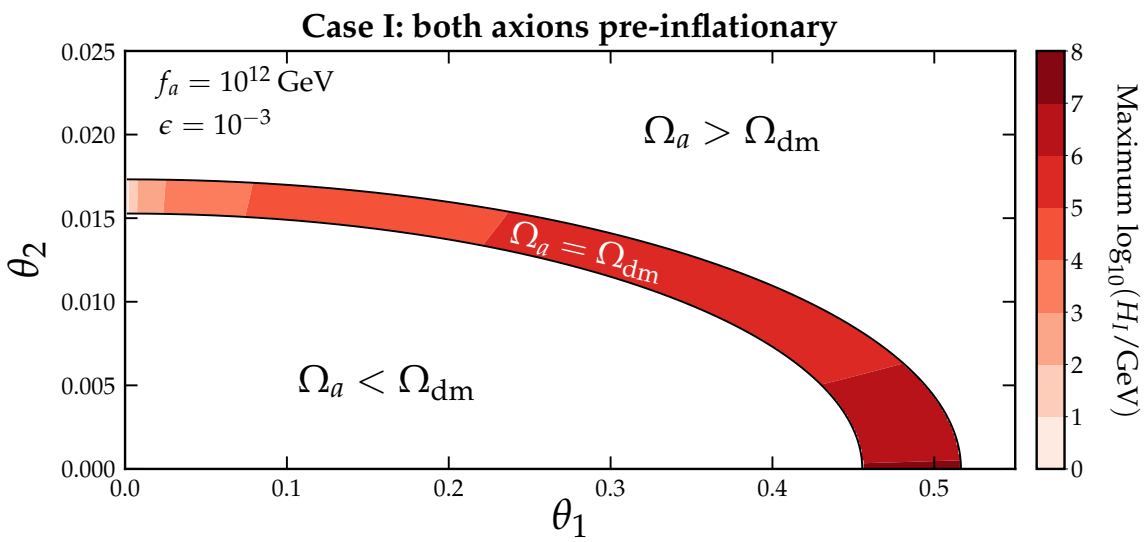

Figure 5.7: Isocurvature bounds for a particular choice of f_a and ϵ are plotted. The red band shows where the dark matter abundance is correct, while the shade illustrates the constraints on the inflationary scale. The position of the bend depends on the choice of vacuum scales—for different values, the red band would occupy a different portion of the (θ_1, θ_2) parameter space.

5.4.2 Domain walls

We now look at an effect relevant only to the cases II and III, where one or both of the axions are post-inflationary. In the companion axion case, there is now a discrete

$Z_N \times Z'_N$ symmetry in the first cosine term of the companion axion potential Eq. 5.5. When the two axion fields spontaneously break their PQ symmetries, a domain wall network [696, 697, 825] could form, where each patch breaks the symmetry spontaneously, choosing one of the states from the $Z_N \times Z'_N$ symmetry. However, for companion axions, there is also the second cosine term in Eq. 5.5 to consider, with its own different residual symmetry. This term explicitly breaks the residual discrete symmetry from the first term, and vice-versa, since we required initially that $NN'_g \neq N'N_g$. This means that the degeneracy of these axion vacuum states is actually lifted, since there is now an energy difference,

$$V_{\text{bias}} \sim \kappa K \tag{5.41}$$

between the vacua. This term drives the annihilation of the domain walls, and is generally known as a *bias* term for the axion potential—alternative models which aim to solve the standard axion domain wall problem commonly introduce such terms [826–834]. For the companion axion, this term comes freely, and is relatively large, allowing the remarkable, automatic solution of the domain wall problem.

We can explicitly estimate the temperature where this annihilation will occur, and show that for the majority of the companion axion parameter space, this temperature is higher than temperature T_i at which the axion mass 'switches on'. The interpretation of this, then, is that the bias term prevents domain walls from forming at all. Explicit domain wall solutions for the $Z_N \times Z'_N$ symmetry will be quite complicated, so we will make some simplifying assumptions to get an order-of-magnitude estimation. We assume here that two independent sets of domain walls form at the QCD phase transition via the Kibble mechanism [476, 698], corresponding to each of the two axions. The width of the domain wall corresponds to the Compton wavelength of the axion:

$$\delta_i \sim 1/m_i \,. \tag{5.42}$$

The requirement that the domain walls not be thicker than the horizon size, $m_i \gtrsim H(T)$ places a lower bound $m_i \gtrsim 10^{-10}\text{eV}$ on the portion of the parameter space in which domain walls can be considered.

We can estimate the surface tension σ_i of the barrier by considering the difference in

energy between the two vacua:

$$\sigma_i \sim K(1+\kappa)\delta_i \ . \tag{5.43}$$

We can see that the walls associated to the lighter axion are wider and more energetic. The energy difference $V_{\rm bias}$ acts as a pressure p_V on the walls, so that a domain of size $r \sim t$ is annihilated when the pressure dominates the surface tension, $p_T \sim \sigma/t$. This means the wall annihilation time can be estimate as,

$$t_{\rm ann} \sim \sigma/t \ . \tag{5.44}$$

During radiation domination, this corresponds to a temperature,

$$T_{\rm ann} \sim 13.5\ {\rm MeV} \left(\frac{m_i}{10^{-12}{\rm eV}}\right)^{1/2} \left(\frac{11\kappa}{1+\kappa}\right)^{1/2} \left(\frac{10}{g_*}\right)^{1/4} \ , \tag{5.45}$$

written in terms of the fundamental companion axion parameters. The domain walls only form after the axions gain mass term at temperatures T_i given in Eq. 5.18. Requiring that $T_{\rm ann} \lesssim T_i$, then, places a lower bound on the axion masses relevant for domain wall formation:

$$10^{-10}{\rm eV} \lesssim m_i \lesssim 10^{-9}{\rm eV}, \tag{5.46}$$

where we have included also the upper bound from the the barrier width consideration. The conclusion, then, is that for the vast majority of the companion axion parameter space, the formation of domain walls is not relevant, and the misalignment production is not significantly impacted. This is quite a significant theoretical improvement over the standard axion scenario.

5.4.3 Gravitational waves

It is natural to wonder whether gravitational waves would be detectable from the annihilation of these domain walls [477]. We will briefly show here that in the small mass range above, there would certainly be a relatively strong source of gravitational waves, which would be detectable in the near-future by experiments such as the Square Kilometre Array (SKA) [803–805] in the context of pulsar timing arrays. However, it is important to note that the portion of the companion axion parameter space which could be constrained by this

observation is already excluded, since there would be a massive overabundance of axions compared to the dark matter in this region (following the misalignment mechanism).

Regardless, the calculation is straightforward, following the results of numerical simulations in Refs. [835–838]. It was found that the gravitational wave power spectrum grows as $\sim k^3$, up to a peak comoving wavenumber $k_{\rm peak}$, after which it falls as $\sim k^{-1}$ until it reaches a cutoff set by the wall thickness. Estimating the peak gravitational wave frequency as $f_{\rm peak} = k_{\rm peak}/2\pi R(t)$, we find in our case,

$$f_{\rm peak} \sim 1.1 \times 10^{-8}\,{\rm Hz} \left(\frac{m_i}{10^{-10}{\rm eV}}\right)^{1/2} \left(\frac{11\kappa}{1+\kappa}\right)^{1/2}, \tag{5.47}$$

using $g_*(T_{\rm ann}) = g_{*S}(T_{\rm ann}) = 10$. Then we can find the relic density at the peak frequency to be,

$$(\Omega_{\rm GW} h^2)_{\rm peak} \sim 3 \times 10^{-10} \left(\frac{10^{-10}{\rm eV}}{m_i}\right)^4 \left(\frac{(1+\kappa)^2}{12.1\kappa}\right)^2 . \tag{5.48}$$

The predicted gravitational wave signal is shown in Fig. 5.8, plotting the relic density as a function of frequency. The two bands are for two values of the mass at either end of the small acceptable range, and the band width accounts for the uncertainty in the gravitational instanton contribution κ. Although current gravitational wave observations are not quite sensitive enough to see such a signal, future pulsar timing arrays would probe this range. The diagonal band in Fig 5.4 corresponds to the case where the second axion has mass $m_2 \sim 10^{-9}$ (and so forms the domain walls), which would occur in Case III where both axions are post-inflationary. The case where the larger axion a_1 forms the domain walls would appear as a vertical stripe to the right of the diagram.

Of course, we can see from Fig 5.8 that the region constrainable by SKA falls outside the correct dark matter abundance, at least according to our misalignment calculation. Of course, a modification of this result, or a different production mechanism altogether, might allow companion axions in this region, so the exercise is not pointless.

5.4.4 Primordial black hole formation

Similarly, it is prudent to wonder if the collapse of domain walls could form primordial black holes, and if so, what their properties would be [484–488]. This calculation follows similar caveats to the above, in that the relevant mass region corresponds to far too large

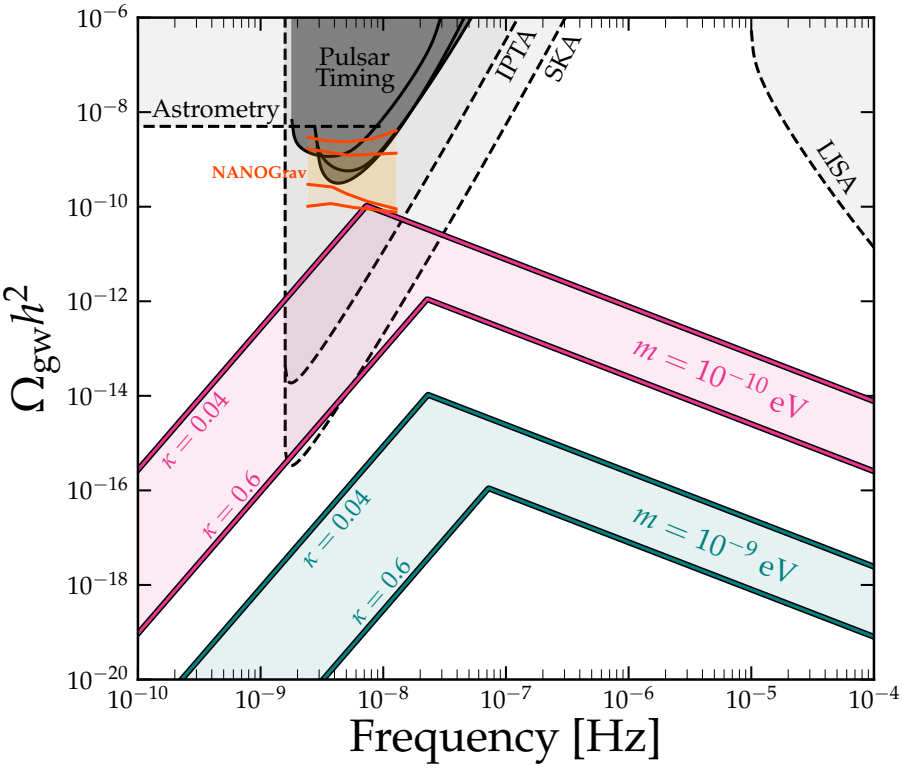

Figure 5.8: The gravitational wave signal from collapsing domain walls for two choices of axion mass are shown. For comparison, the power-law integrated exclusion curves [839, 840] for pulsar timing array searches for a gravitational wave background are shown. This includes NANOGrav [841–844], EPTA [845–847] and PPTA [848, 849], as well as the future sensitivities of LISA [850] and SKA [803–805]. We make use of the power-law integrated curves presented in Ref. [851] with the exception of the sensitivity to gravitational waves with astrometric data, which is taken from Ref. [852]. In addition, we show here the 'hint' of a strong signal of a stochastic background reported recently by NANOGrav [853, 854], although classifying this signal as a gravitational wave background is still somewhat premature.

an abundance of companion axions. Nevertheless, the arguments are still interesting to consider.

PBHs can be formed from the collapse of closed domains containing a false vacuum, which start shrinking once they are smaller than the Hubble horizon. The total energy in the closed domain comes from the wall tension, as well as the interior false vacuum energy. The former is a surface effect, while the latter is volumetric. This means that the mass of the closed domain is,

$$M_i = 4\pi r_i^2 \sigma_i + \frac{4\pi}{3} r_i^3 V_{\rm bias}, \tag{5.49}$$

where r_i is the radius of the closed domain wall for axion a_i. Once the radius is less than the Schwarzschild radius, the closed domain will collapse into a black hole (ignoring, for now, the fact that we might prefer to describe the PBH with a cosmologically-embedded metric). For the small range of relevant masses Eq. 5.46, the bias term dominates, so we can estimate the black hole mass after collapse:

$$M_{\rm PBH} \sim \frac{\sqrt{3}}{4\sqrt{2}} \frac{M_{\rm Pl}^3}{(\pi\kappa K)^{1/2}} \sim 150 M_\odot \left(\frac{\kappa}{0.1}\right)^{-1/2} . \tag{5.50}$$

This mass, interestingly, is somewhere in the range of what LIGO/Virgo would be able to detect. The temperature at collapse is then,

$$T_{\rm coll} \sim 25{\rm MeV} \left(\frac{\kappa}{0.1}\right)^{1/4} \left(\frac{g_*}{10}\right)^{-1/4} . \tag{5.51}$$

The uncertainty on κ allows for a relatively wide range of possible PBH masses. An important difference in our calculation compared to others [486, 487] is that they were forced to assume very small bias terms in order to preserve the standard axion solution to the strong-CP problem—then PBHs could only be formed by having long-lasting $N > 1$ domain wall networks.

In our case, it is not easy to predict with much precision the population of PBHs produced. This is primarily because it is difficult to estimate the survival probability of the domain wall network until the collapse temperature. Numerical simulations [787] exist for the standard axion case, but the large bias term and complexity of the network in our case make these predictions unreliable. We can make a very crude argument, however, by treating this process as the decay of a false vacuum with mean lifetime $\sim t_{\rm ann}$. Since the tunneling processes which destroy false vacuums have exponential time dependence, we can estimate that the fraction of closed domains which survive until the collapse temperature would be,

$$p_{\rm coll} \sim e^{-(T_{\rm ann}/T_{\rm coll})^2} \sim 10^{-22} - 10^{-9}. \tag{5.52}$$

The large uncertainy in the prediction corresponds to $\kappa \in [0.04, 0.6]$, for fixed axion mass $m_i = 10^{-10}$eV. For larger axion masses, the mass dependence of $T_{\rm ann}$ in Eq. 5.45 means that this probability becomes negligibly small. It is straightforward to compute the

present-day PBH energy density from the fraction $p_{\rm coll}$ as,

$$\rho_{\rm PBH} \sim \frac{p_{\rm coll}}{r_i^3} M_{\rm PBH} \left(\frac{T_0}{T_{\rm coll}}\right)^3 \sim p_{\rm coll} \frac{M_{\rm Pl}^6}{M_{\rm PBH}^2} \left(\frac{T_0}{T_{\rm coll}}\right)^3 , \tag{5.53}$$

leading to the fraction of DM density,

$$f_{\rm PBH} \sim 34.9\, p_{\rm coll} \frac{M_{\rm Pl}^4}{H_0^2 M_{\rm PBH}^2} \left(\frac{T_0}{T_{\rm coll}}\right)^3 . \tag{5.54}$$

The large range of uncertainties leads to the dramatic range of the dark matter fraction, from $\mathcal{O}(10^{-13})$ to $\mathcal{O}(1)$. Again, for heavier axions, this fraction is negligibly small. Although this calculation comes with many caveats, it is certainly interesting that it could account for some population of roughly LIGO-sized black holes.

The layperson box: companion axions (my work) Previous — Next

It is common folklore that particle physics and gravity do not get along. This is true in some technical senses related to quantum mechanics, but in some circumstances there are no *fundamental* reasons why one cannot do them both at the same time. The real reason in most cases is that both of those subjects are just quite difficult and it is exponentially harder to do them simultaneously.

It was realized around the '70s and '80s that gravity does the exact same kind of twisty maneuvers that the strong force does. They found, however, that gravity did its twisting and untwisting in the background, and so did not affect any quarks or other particles. But the situation changes when you try to do gravity and one of the other forces at the same time. First, they tried gravity and electromagnetism, and realized there were only minor consequences. Then, they did not look again. (welcome to the frontier of science. Weird, hey?)

My supervisor Archil and one of his students, Zhe, looked at it again in 2021. They found that if you examine the gravity twistings and the strong force twistings at the same time, the single axion which solved the problem before could no longer handle both. The only solution, they suggested, was to introduce a second axion (genius).

Finally, after everything, this is where I came in. If you now have two axions, which we called the 'companion axion' model, the whole landscape of their behavior is different—not only do they have different properties, they can also interact with each other. Along with Zhe, Giovanni (from the University of New South Wales) and our postdoc Ciaran, we took the dozen or so most-important axion results/observations and did the maths again. Fortunately for me, Giovanni and Ciaran were already axion experts so I could kind of bumble along learning from them.

Some things worked out more-or-less the same as the regular axion. I would say our two most important results were: 1; much lighter axions would have to be the dark matter, compared to the kind they are currently searching for, and 2; the 'domain-wall problem' is automatically solved. This is a peculiar problem for regular axions, in which it seems theoretically likely that axions would configure themselves in the early universe in a specific way which contains more energy than the entire universe (you can see that this would be an issue).

And thus concludes my three-or-so years of PhD research. Thanks for following along—it's been a trip. You may as well read the conclusion, while you're here...

6
Conclusion

Our goal is to discover that we have always been where we ought to be.

—Aldous Huxley, *The Doors of Perception*

Sometimes, after the more formal questions at the end of a talk, you are asked, 'so... what do you reckon it is, really?' Usually, this comes with a sly nudge and a wink. After all this—the technical ambiguities of cosmological black holes, or the possible gravitational interactions of a possible new particle—are we closer to knowing dark matter?

My usual answer is somewhat diplomatic: 'I would not be surprised if it was a mixture...fifty percent black holes, fifty percent axions. Something like that'. But it is really just a guess, and any careful physicist would just say that it could be anything that's not yet constrained.

It is hard not to be disappointed with this state of affairs. Colliders must probe increasingly high energies to find new physics, promising diminishing returns as the expenses and technological difficulties balloon. Dark matter detectors must be bigger, more sensitive and more complicated. In the absence of new discoveries, theoretical models become more and more technically challenging and specialized, and new ideas come with ever more caveats and ifs and buts. The search for the answers is slow, bogged down by the politics of large groups of people and the difficulty of communication across increasingly specialized disciplines.

But there is still much to be excited for. We have only been detecting gravitational waves for less than a decade, and they have provided a great wealth of new data. In the following decades, as the next-generation gravitational wave detectors come online (both on Earth and in space), we will have remarkable multi-messenger probes that extend beyond even the surface of the CMB—and who knows what we will be able to learn about black holes. Multidisciplinary efforts with fields such as quantum sensing [855, 856], condensed matter physics [857, 858], and even DNA biotechnology [859] are leading to novel new dark matter detection mechanisms. Small-to-large-sized cracks in the Standard

Model, such as flavor anomalies, neutrino masses, and of course—dark matter—may be pointing the way for exciting discoveries at the next order of magnitude in energy. For a while, it has seemed like theory was 'ahead' of experiment, but there may be cause for a great deal of new astroparticle theory in the coming years.

The biggest unknown frontier left in fundamental physics lies somewhere in the intersection of the large and small scales, in the intersection of gravity and particle physics. In this thesis, I explored two realizations of this tension—black holes embedded in the hot bath of the early universe, and QCD physics modified by gravitational instantons.

After a lengthy introduction, Chapters 2 and 3 discussed black holes and their cosmological embeddings. The former included a relatively detailed introduction to the history and physics of black holes, focusing in particular on primordial black holes and their constraints as a dark matter candidate. In Chapter 3, I introduced the topic of *cosmological black holes*, spacetimes which are embedded properly into the hot, thermal bath of the early universe. Although much has been written on these black holes, they still carry many ongoing questions and ambiguities, despite being increasingly important in the contemporary Renaissance of primordial black hole study.

To emphasize the phenomenological consequences of these spacetimes, we chose the most viable metric we could find—the Thakurta metric—and reconsidered the dark matter bounds with this spacetime. We found two rather drastic results. First, the gravitational wave bounds, which limit the abundance of the important LIGO-mass black holes, were completely evaded. Secondly, these black holes evaporated rapidly in the early universe, significantly increasing the mass of the lightest black hole which could survive until today and closing the otherwise-unconstrained asteroid-mass range for primordial black holes.

The back half of this thesis was concerned with axions. In Chapter 4, I reviewed the strong-CP problem and the QCD vacuum from a relatively pedagogical lens, before summarising the benchmark axion models and their constraints. Then I introduced our new theory of the *companion axion* in Chapter 5. This second axion was introduced to account for the disruption of the axion solution when gravity is included in the QCD vacuum. After summarizing the theory of colored gravitational instantons, we rederived a large sample of axion physics in the context of the new companion axion model. This included various axion-photon constraints, the misalignment production of axion dark matter, and numerous other cosmological concerns. Throughout we found that the companion axion had qualitatively different phenomenology compared to the standard axion, especially

when considering its viability as a dark matter candidate.

Presumably, the tension between gravity and particle physics will leak into more than just these two examples. Maybe, there is an important result hiding in plain sight, in some system where the two forces were assumed to be decoupled. If there is any lesson to be taken from this thesis, it is that such systems certainly exist, and that there may be large qualitative consequences from taking them seriously.

In the introduction of this thesis, I lamented the millennia-or-so of physics that happened before I was born, and so had to catch up with. I will admit, it is sometimes frustrating to find out that a neat idea I had come up with had actually already been researched and published in 1989, only to be slowly forgotten again. I suspect, however, that all of those ancient and less-ancient scientists would gladly agree to swap places with us today, given the chance. As the saying goes, we know the most we have ever known, and the least we will ever know again (hopefully). At the very least, we can get our revenge by making the workload of learning even larger for the generation after us.

If you have genuinely made it this far, or if you are just passing through, thanks for reading. It has been my pleasure—maybe I will catch you around sometime. Goodbye.

A
The layperson appendix: symmetries

The best time to plant a tree is thirty years ago. But the second best time is... twenty-nine years ago. Then... twenty-eight years ago, etc. etc.

—Ancient proverb

Of all the aesthetic revelations in fundamental physics, the importance of symmetries is probably my favorite. I'll do my best here to leave most of the maths out of it, but sometimes departing from actual physics can leave things a bit esoteric—you'll just have to trust that I am not tricking you. This appendix is a continuation of the layperson boxes throughout, but I felt I needed a bit more space to really flesh out my thoughts here.

Back to the particle physics box

Back to the index

A.1 What do we mean by symmetry?

It is impossible to separate the concept of symmetry from that of 'transformations'. When we say that some system or object possesses a kind of symmetry, what we are really saying is that if we modified the object in a particular way, it would still appear the same.

Consider a square. It has two kinds of symmetries—the horizontal and vertical mirror reflections, and the rotations by 90 degrees. Already in the language I use, the relation between a symmetry of the square and how I must transform it is apparent. The mirror reflections involve swapping corners across a line through the center of the square, while the rotations involve cycling each corner around to the next one[1].

The situation is the same, or maybe even easier, when we are discussing mathematical

[1] If you are concerned about the diagonal mirror reflections, it is possible to reproduce them by combining the rotations and horizontal/vertical reflections. Try it yourself—it may be helpful to label the corners with numbers to keep track.

equations. For example, consider the relatively harmless,

$$y = x^2 \ . \tag{A.1}$$

If you know what the plot of this looks like, you might be aware that it possesses some symmetry. You can see this explicitly, though, because if you swap $x \to -x$, the equation stays the same (since $(-x)^2 = x^2$). The transformation in question, there, was swapping negatives for positives, and vice-versa.

The two examples I have shown so far are what is known as 'discrete' symmetries. Not in the sense that they are particularly sneaky, but in the sense that the transformations are big, chunky operations on the whole system. I have to modify the system with a specific singular action in one go. The other kind of symmetries are called 'continuous', and they involve being able to transform the system by any amount I like, big or small. The easiest example is the circle—I can rotate it about its center by any angle and it remains the same.

For an easy maths example of this, let's look at the formula for kinetic energy—the amount of energy an object has due to its speed:

$$\begin{aligned} E_{\mathrm{K}} &= \frac{1}{2}mv^2 \\ &= \frac{1}{2}m\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 \ . \end{aligned} \tag{A.2}$$

Where the second line uses the definition of speed: velocity = change in distance over change in time. In maths we would call that a derivative of position, x, with respect to time, t. Now, what happens under the following transformation:

$$x \to x + c \ , \tag{A.3}$$

where c is just some constant distance? The derivative changes like so:

$$\frac{\mathrm{d}x}{\mathrm{d}t} \to \frac{\mathrm{d}(x+c)}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\mathrm{d}c}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}t} + 0 \ , \tag{A.4}$$

because constants, by definition, do not change with time. As you can probably see, the kinetic energy does not change. I can make c any number I want, and it doesn't matter. If we have a transformation that doesn't change the system, then, we have a symmetry—and specifically, a continuous one, here. In physics language, I would say that the kinetic

energy doesn't actually depend on where the origin of my ruler is, or rather, where I'm measuring it from. The symmetry $x \to x + c$ is known as a 'translation' symmetry.

Although it may seem obvious that physics shouldn't depend on where you measure it, it turns out these kinds of continuous symmetries are extremely important.

A.2 Emmy Nöether's theorem

Ok now, a difficult question for you. What is energy? This is probably an obtuse term you were taught in year ten science, but it may have been never carefully explained. Is it just some magical quality that objects 'possess?' How do we know it is always conserved, and what does that even mean if it is often being transferred between different kinds of energies, which may not seem related to each other at all?

I might not have an entirely satisfying answer for you, but maybe I can slightly elucidate the scene—the answer comes from these continuous symmetries. The intuition here comes from realizing that whenever we have a continuous transformation, there is some quantity that is 'unchanged' during the transformation. For the circle, maybe you would say that the radius of the circle is conserved under these transformations. But what about the translation symmetry?

These were the kind of questions that led to Emmy Nöether's famous and remarkable theorem of 1918—one of the most important and influential pieces of maths and physics in the last century. She showed that *every* every system which has a continuous symmetry has a particular conserved 'quantity' attached to it, which is easily derived.

When you follow Nöether's theorem for the translation symmetry, you find that the conserved quantity is $m \times v$, otherwise known as *momentum*. In other words, 'conservation of momentum' is directly linked to the fact that physics is the same at every location. If, for some odd reason, the laws of physics were different in different spatial points, momentum would no longer be conserved. In that sense, momentum is not just some abstract thing that your teacher told you is conserved—it is a consequence of the more fundamental principle that physics doesn't care where you do it.

There are two more of these kind of 'spacetime' symmetries. One is related to rotations—it doesn't matter how you orient your physics lab, physics should be the same from every position. This leads to conservation of angular momentum. The last is 'time translation', or rather, it doesn't matter *when* you do your physics experiment, since physics

doesn't change over time. The conserved quantity here is... *energy*.

Energy then, is not some magical property that objects seem to possess. It is just a restatement of the fact that physics is the same from one time to another. The order of implication is perhaps important. We give the conserved quantity the name 'energy' after the fact—energy *is* the thing-which-is-conserved-in-time, not some independently existing entity which also happens to be conserved.

The curious thing here, is that if something like time translation was *not* a symmetry, then energy would not be conserved (or, for that matter, definable at all). This is relevant to the history of our universe, however, which used to be much smaller and hotter than it was today (see: page 20). If I did a physics experiment fourteen billion years ago, I *would* get different results to today. The expanding universe does not conserve energy!

A.3 How symmetries make particles interact

The circle-and-radius analogy above is actually relevant to particle physics, as well. The equations for particles like the electron have a kind of circle-looking symmetry to them. As it turns out, this is a very important symmetry, whose conserved quantity is named *electric charge*. This can very roughly be thought of as the radius of this circular symmetry—more carefully, it might be considered a factor which determines 'how much the circle turns' whenever you turn it (which in my opinion kind of feels related to radius).

The fact that particles 'carry' extra conserved quantities besides energy and momentum makes the universe interesting. If these didn't exist, the particles would just float around, never interacting with each other. That's because the fundamental particles aren't really 3D objects, like billiard balls, which can just bang into each other (and recall from page 54—'banging into each other' is really just the electric fields in atoms repelling each other). Rather, particles are infinitely small points with no 'size'. But if they have extra properties, like charges, they all of a sudden can 'see' each other and interact, instead of just floating around boringly. You can change particles into other particles, even, so long as the conserved quantities are conserved.

Particle physicists talk about these 'charges', and the fundamental forces, interchangeably. In some sense, the charges that particles feel are like 'code words' for which of the fundamental forces a particle feels. If it has an electric charge, we say that it feels the electromagnetic force.

Electric charge and electromagnetism

The electric charge of the electron is probably its most important property. It determines the majority of the electron's activities, since the electron's behavior in electric and magnetic fields depends on the charge. This determines, for example, how electrons orbit the nuclei of atoms, and so therefore all of chemistry and ultimately biology and everything else as well. In addition, light can be thought of as oscillating electric and magnetic fields, so the electron is able to interact with light too. This means that, for example, light is absorbed and emitted by atoms—this is handy if you enjoy seeing the world around you.

I am actually simplifying a bit when I said that electrons have a circle kind of symmetry. The reality is somehow even crazier. When you actually look at the equations, it turns out that they *don't* actually have such a symmetry—at least, on their own. It is only when you look at the *combined* equations for electrons and photons (i.e., light) that the circle symmetry actually works. This might be considered the most unusual, and possibly most beautiful, revelation in physics. For some reason (which we have zero clue about), requiring that electrons have such a circle symmetry actually *enforces* the rules of interactions between electrons and photons, and therefore, pretty much every process you care about. The situation is the same for the other forces, too. This is why particle physicists will talk about photons 'carrying' the electromagnetic force—the interactions of photons and charged particles *is* electromagnetism.

Why should the enforcement of these circular symmetries lead to every single rule of nature? If you get to heaven before I do, please ask your favorite deity for me.

Weak charges and the weak force

Actually, every kind of particle has some number of these kind of symmetries, beyond the circular symmetry which leads to electric charge. There are two kinds of so-called 'weak' charges—unsurprisingly, this is because the equations for some particles can be 'rotated' in two directions. The weak charges lead to weird interactions which change particles in the nuclei of atoms into other ones, causing the nuclei to decay (i.e., beta decay). It turns out this is somewhat important for the fusion processes which power stars, which is nice. Accompanying the weak charges are three weak equivalents of the photon, known as the 'weak bosons.'

Color charges and the strong force

The last fundamental force is the strong force. This force has three charges, called 'red', 'green', and 'blue', and you guessed it—three circle-y symmetries. The interactions that these charges allow keep the nuclei of atoms stuck together, which is quite useful. This time, there are eight photon equivalents, known as gluons (if you were wondering—the rule for the number of the photon equivalents is 'number of charges squared, minus one').

Actually, it seems like nature does not 'like' it when there are individual colors floating around. As a result, colored particles (quarks) always get stuck together into bigger particles. For example, the proton, inside the nuclei of atoms, is made up of three quarks, one of each color (two 'up' quarks and one 'down' quark), so that the proton is ultimately 'colorless'. The protons and neutrons inside of the atom's nucleus then stick together because there is a kind of residual glue left over from all these shenanigans.

Gravity

Finally, you may have noticed that I have distinguished the 'fundamental forces' from gravity, which is surely fundamental as well. On page 13 I discussed how gravity can be seen as a purely 'geometric' effect, rather than a particle physics one. This is part of a broader tension in theoretical physics between these two regimes. Actually, it is perfectly possible to reimagine all of particle physics using geometry, just like General Relativity, although the geometry is substantially more abstract. And similarly, it is possible to discuss gravity in particle physics language.

Just as the photon is the particle of light (i.e., electromagnetism, which interacts with electric charges), and the gluons are the particles of the strong force (which interact with color charges), the 'graviton' is the particle of gravity, which interacts with *mass*. Unfortunately, there are some curious technical problems with this synthesis—one reason is that gravity likes to form itself into weird configurations like black holes, and it is not known how to put black holes into particle physics equations. Often it is said that gravity is incompatible with quantum mechanics, but this is presumably more a statement of our current lack of knowledge of the theory which combines both, and not an immutable law of the universe.

A.4 The path integral

Okay, this final point doesn't really have to do with symmetries, but I wanted to put it somewhere, and I figured here was as good as any. This might be considered the *other* craziest thing about particle physics, besides the circle-interactions thing above, for those who have heard before of quantum mechanics.

I am occasionally asked which is my favorite interpretation of quantum mechanics—many-worlds, or the other one? The truth is, I don't think much about it. Classical quantum mechanics, of which those interpretations relate to, is not even the most fundamental quantum theory that we have in physics. Particle physics is really formulated in the language of 'quantum field theory', a daunting, if wonderful, mathematical formalism that combines Special Relativity and quantum mechanics. And it comes with a particularly wild idea called the 'path integral'.

You may recall from your life that objects like to follow trajectories through the world that minimize something like their change in energy (called 'action' in physics). Water follows the path of least resistance, and balls in valleys stay there instead of absorbing energy from the ground to roll up hills, even if that wouldn't violate conservation of energy. This certainly applies to particles and all their interactions, too. The crazy thing, first described by Richard Feynman, is that actually all those other possible-but-unphysical paths are important.

It turns out that if you would like to find the *quantum-mechanical* path that a particle takes, you can actually 'add up' all those extra paths—even the craziest, most convoluted ones—as equally probable possibilities. Once you add them all up, what you are left with is the correct quantum particle process. To me, this is an even more wondrous way to phrase quantum mechanics. It's not a matter of multiverses—particles really take all possible paths and kind of average them out. I just think that's pretty neat.

A.5 Some final philosophical musings

Often, physics as presented to a lay-audience is somewhat sloppy with its ontology (the philosophical branch dealing with what exists), or even its epistemology (the branch dealing with what we know), and I too have been no better than average so far. I think this is understandable from a physicist's perspective—what we have are a series of models, with somewhat-well-defined internal logical systems, which we can use to produce

quantitative predictions about nature. When a physicist says 'the electron field does this' or 'the spacetime is curved', they are (generally) not making a rigorous statement about the true *existence* of these constructs, because they just don't care. The mathematical formalisms get the job done—not only is their possible Platonic existence irrelevant, but it is unanswerable within physics.

What physicists do excellently (and particle physicists in particular) is explicitly laying out the region of validity in which a particular theory models nature up to some known precision. Beyond the borders of a theory's capabilities, there is no longer anything meaningful to be said, at least as far as physics is considered. And so we must be very careful extrapolating lessons about the 'true nature of reality' or whatever from even our most fundamental physics.

Still, many of you, and I, might care about such matters as 'what actually exists'. These questions are usually out of the realm of physics—so long as the answers don't contradict it—and in the realm of philosophers, a truly unruly lot. They have a lot to say about things like categorizing scientific representation, time, the mind, etc., ranging from fascinating to entirely illegible. Generally, as in all good things, I find that more questions are posed than answered, and I am left with a desire to forget about it and have a beer. You are where you ought to be, as they say.

B
The Hamiltonian in General Relativity

As my artist's statement explains, my work is utterly incomprehensible and is therefore full of deep significance.

—Calvin, *Calvin and Hobbes, July 15, 1995*

The following is based on a set of notes that my supervisor Archil distributed to some of his first year students. This derivation is not essential to the arguments in this thesis, but is interesting nonetheless and I have not seen it written so clearly elsewhere. I will show here that in theories which have a time-reparametrization symmetry, the Hamiltonian is identically zero. The point of emphasis here is that 'energy' is somewhat tricky to define in General Relativity, which is built on full diffeomorphism invariance (which of course includes time-reparametrization invariance).

This does not mean that we cannot discuss mass or energy in General Relativity, but rather that we have to put more work into their definitions. For example, in the ADM formalism [412] spacetime is split between spatial slices and time, and the dynamics of fields in these foliations are controlled by the *Hamiltonian constraint*.

The relevance to this thesis has to do with our intuition regarding energy in cosmological contexts. For example, the concept of cosmological redshift is sometimes misunderstood—where does the energy of the photon go as it redshifts due to the expansion of space? Really, it does not 'go' anywhere. There is no conservation of energy because there is no time translation invariance in the first place. Similarly, we must be careful with our intuition regarding the growing Misner–Sharp mass of the Thakurta metric. The growth of the mass does not need to 'come' from somewhere, any more than the energy of the photon goes anywhere.

B.1 The derivation

For simplicity we will just discuss the case where our theory has a full time-reparametrization symmetry. Consider an inhomogenous time translation,

$$t' = t + \epsilon(t) \ . \tag{B.1}$$

Of course, in the limit where ϵ is constant, we would recover the time-translation symmetry which leads to conservation of energy under Nöether's theorem. Let us focus on this case for a moment, since it will be useful when we try to write an action invariant under Eq. B.1. In the homogeneous case,

$$\mathrm{d}t' = \mathrm{d}t \ . \tag{B.2}$$

The Hamiltonian is, by definition, the conserved quantity under homogeneous time translations. What we would like to do then is begin with an action with this symmetry, modify it so that the symmetry is extended to the full inhomogenous translation, and examine the effect on the Hamiltonian of this action. We can do this by defining a new time τ, and introducing a generalized coordinate $Q(t)$,

$$\mathrm{d}\tau = Q(t)\mathrm{d}t \ , \tag{B.3}$$

such that,

$$Q(t')\mathrm{d}t' = Q(t)\mathrm{d}t \ . \tag{B.4}$$

The above condition means that the Langrangian with respect to τ satisfies the homogeneous time translation property, since

$$\mathrm{d}\tau = Q(t)\mathrm{d}t = Q(t')\mathrm{d}t' \equiv \mathrm{d}\tau' \ . \tag{B.5}$$

The point of this preamble was to show the following: if we have an action which is invariant under time-translation symmetry,

$$S = \int \mathrm{d}\tau \ L\left[q, q'\right] \ , \tag{B.6}$$

where $q' \equiv \mathrm{d}q/\mathrm{d}\tau$, then we can use $\mathrm{d}\tau = Q(t)\mathrm{d}t$ to find the action,

$$S = \int Q\mathrm{d}t\; L\left[q(t), \dot{q}/Q\right] \;, \tag{B.7}$$

where $\dot{q}$ now refers to differentiation with respect to t. By inspecting Eq. B.5, however, we can see that this action is now invariant under the more general time-reparametrization invariance given by Eq. B.1. In informal language, the coordinate Q 'carries' the inhomogenous part of the time symmetry. It is the enforcement of the condition in Eq. B.5 which promotes the time-translation invariance of the Lagrangian to the 'stronger' time-reparametrization invariance. Since the new symmetry is stronger, we expect tighter constraints—as we will shortly see, this results in an identically zero Hamiltonian.

Since we defined Q as a generalized coordinate, we can inspect the Euler-Lagrange equations for Q:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial(QL)}{\partial\dot{Q}} = \frac{\partial(QL)}{\partial Q} \;. \tag{B.8}$$

The left-hand side is zero, since there are no $\dot{Q}$ terms, giving,

$$\begin{aligned} 0 &= L + Q\frac{\partial L}{\partial Q} \\ &= L + Q\frac{\partial L}{\partial\left(\dot{q}/Q\right)}\left(-\frac{\dot{q}}{Q^2}\right) \\ &= L - \frac{\dot{q}}{Q}\frac{\partial L}{\partial q'} \\ &= L - q'\frac{\partial L}{\partial q'} \;. \end{aligned} \tag{B.9}$$

By definition, however, the Hamiltonian for the time-translation invariant τ-Lagrangian Eq. B.6 is the standard relation,

$$H \equiv q'\frac{\partial L}{\partial q'} - L \;. \tag{B.10}$$

Then the result from Eq. B.9 implies that,

$$H = 0 \;, \tag{B.11}$$

for any system with time-reparametrization invariance. The specific reason we have such

a strong condition on the Hamiltonian now results from the imposed relation in Eq. B.4. This condition enforces the time-reparametrization invariance, ultimately leading to the result that the Hamiltonian with such a condition must be identically zero.

Bibliography

[1] A. Kobakhidze and Z. S. C. Picker, *Apparent horizons of the Thakurta spacetime and the description of cosmological black holes*, *Eur. Phys. J. C* **82** (2022) 347 [2112.13921]. [Pages v, vi, 66, 73, 75, and 77]

[2] Z. Chen, A. Kobakhidze, C. A. J. O'Hare, Z. S. C. Picker and G. Pierobon, *Phenomenology of the companion-axion model: photon couplings*, 2109.12920. [Pages v, vi, 125, and 139]

[3] Z. Chen, A. Kobakhidze, C. A. J. O'Hare, Z. S. C. Picker and G. Pierobon, *Cosmology of the companion-axion model: dark matter, gravitational waves, and primordial black holes*, 2110.11014. [Pages v, vi, 125, 134, and 147]

[4] C. Boehm, A. Kobakhidze, C. A. J. O'Hare, Z. S. C. Picker and M. Sakellariadou, *Comment on: Cosmological black holes are not described by the Thakurta metric*, 2105.14908. [Pages v, vi, 66, and 75]

[5] Z. S. C. Picker, *Navigating the asteroid field: New evaporation constraints for primordial black holes as dark matter*, 2103.02815. [Pages v, vi, 61, 65, 66, and 92]

[6] C. Boehm, A. Kobakhidze, C. A. J. O'hare, Z. S. C. Picker and M. Sakellariadou, *Eliminating the LIGO bounds on primordial black hole dark matter*, *JCAP* **03** (2021) 078 [2008.10743]. [Pages v, vi, 65, 66, and 85]

[7] H. Georgi, *Lie algebras in particle physics*, vol. 54. Perseus Books, Reading, MA, 2nd ed. ed., 1999. [Page 3]

[8] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995. [Page 3]

[9] S. M. Carroll, *Spacetime and Geometry*. Cambridge University Press, 7, 2019. [Pages 3 and 29]

[10] V. Faraoni, *Cosmological and Black Hole Apparent Horizons*, vol. 907. Lect. Notes Phys., 2015, 10.1007/978-3-319-19240-6. [Page 3]

[11] J. Schwichtenberg, *Physics from Symmetry*, Undergraduate Lecture Notes in Physics. Springer International Publishing, Cham, 2018, 10.1007/978-3-319-66631-0. [Page 3]

[12] J. Schwichtenberg, *Demystifying the qcd vacuum*, http://www.jakobschwichtenberg.com/demystifying-the-qcd-vacuum-part-1/. [Pages 3, 100, and 107]

[13] D. Baumann, *Cosmology*, http://cosmology.amsterdam/education/cosmology/, June, 2018. [Page 3]

[14] B. Carr, K. Kohri, Y. Sendouda and J. Yokoyama, *Constraints on primordial black holes*, 2002.12778. [Pages 3, 40, 55, 56, and 61]

[15] L. Di Luzio, M. Giannotti, E. Nardi and L. Visinelli, *The landscape of QCD axion models*, *Phys. Rept.* **870** (2020) 1 [2003.01100]. [Pages 3, 100, 120, 121, and 139]

[16] D. Hey, *Latex thesis template for the university of sydney honours/masters/phd*, https://github.com/danhey/thesis, 2018. [Page 3]

[17] Aristotle, *Physics*, in *The Works of Aristotle, Volume 1*. The University of Chicago, 1990. [Page 5]

[18] Aristotle, *On the heavens*, in *The Works of Aristotle, Volume 1*. The University of Chicago, 1990. [Page 5]

[19] Aristotle, *On generation and corruption*, in *The Works of Aristotle, Volume 1*. The University of Chicago, 1990. [Page 5]

[20] C. Rovelli, *Aristotle's Physics: a Physicist's Look*, 1312.4057. [Page 5]

[21] Archimedes, *THE SAND-RECKONER*, p. 221–232. Cambridge Library Collection - Mathematics. Cambridge University Press, 2009. 10.1017/CBO9780511695124.017. [Page 5]

[22] N. Copernicus, *De revolutionibus orbium cœlestium*, https://www.reed.edu/math/wieting/mathematics537/DeRevolutionibus.pdf, 1543. [Pages 5 and 14]

[23] J. Kepler and B. Tycho, *Astronomia nova aitiologetos*. G. Voegelinus, 1609, https://doi.org/10.5479/sil.126675.39088002685477. [Page 5]

[24] G. Galilei, *Sidereus nuncius*. Apud Thomam Baglionum, 1610, https://doi.org/10.5479/sil.95438.39088015628597. [Page 5]

[25] I. Newton, *Philosophiæ Naturalis Principia Mathematic*. Jussu Societatis Regiæ ac Typis Joseph Streater, 1687. [Page 5]

[26] M. Heydari-Malayeri, *The Persian-Toledan Astronomical Connection and the European Renaissance*, 0709.1216. [Page 5]

[27] A. al Battānī, *Kitāb az-zīj*, https://gutenberg.beic.it/webclient/DeliveryManager?pid=4783627, 800-900 AD. [Page 5]

[28] K. Plofker, *Fazārī: Muḥammad ibn Ibrāhīm al-Fazārī*, pp. 362–363. Springer New York, 2007. 10.1007/978-0-387-30400-7_443. [Page 5]

[29] J. Samsó, *Biṭrūjī: Nūr al-Dīn Abū Isḥāq [Abū Jacfar]Ibrāhīm ibn Yūsuf al-Biṭrūjī*, pp. 133–134. Springer New York, 2007. 10.1007/978-0-387-30400-7_164. [Page 5]

[30] S. P. Blake, *The observatory in Maragha*, p. 73–87. Edinburgh University Press, 2016. [Page 5]

[31] Aryabhata, *The Aryabhatiya of Aryabhata*, https://en.wikisource.org/wiki/The_Aryabhatiya_of_Aryabhata, 510. [Page 5]

[32] J. P. Montada, *Ibn Rushd's Natural Philosophy*, in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.). Metaphysics Research Lab, Stanford University, Fall 2018 ed., 2018. [Page 5]

[33] C. Huffman, *Philolaus*, in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.). Metaphysics Research Lab, Stanford University, Fall 2020 ed., 2020. [Page 5]

[34] C. Ptolemy, *Almagest*, https://classicalliberalarts.com/resources/PTOLEMY_ALMAGEST_ENGLISH.pdf, 138-161. [Page 5]

[35] J. C. Maxwell, *On physical lines of force*, *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science* **4** (1861) . [Page 5]

[36] J. C. Maxwell, *A dynamical theory of the electromagnetic field*, *Phil. Trans. R. Soc. Lond.* **155** (1865) . [Page 5]

[37] O. Heaviside, *Electrical papers Vol I*. London: Macmillan and Co., 1892. [Page 5]

[38] C. Donaghy-Spargo, *On heaviside's contributions to transmission line theory: wave diffusion and energy flux*, *R. Soc. A.* **376** (2018) . [Page 5]

[39] H. A. Lorentz, *Simplified Theory of Electrical and Optical Phenomena in Moving Systems*, *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences* **1** (1898) 427. [Pages 5 and 6]

[40] H. A. Lorentz, *Electromagnetic phenomena in a system moving with any velocity smaller than that of light*, *Proceedings of the Royal Netherlands Academy of Arts and Sciences* **6** (1904) . [Pages 5 and 6]

[41] H. Poincaré, *La théorie de lorentz et le principe de réaction*, *Archives néerlandaises des sciences exactes et naturelles* **5** (1900) . [Pages 5 and 6]

[42] H. Poincaré, *Sur la dynamique de l'électron*, *Rend. Circ. Matem. Palermo* **21** (1906) . [Pages 5, 6, 12, and 40]

[43] G. Galilei, *Dialogue concering the two chief world systems—ptolemaic and copernican*, 1632. [Page 5]

[44] A. Einstein, *Zur elektrodynamik bewegter köorper*, *Annalen der Physik* **17** (1905) . [Page 6]

[45] H. Minkowski, *Die grundgleichungen für die elektromagnetischen vorgänge in bewegten körpern*, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* (1908) . [Pages 6 and 9]

[46] J. D. Norton, *Einstein, Nordström, and the Early Demise of Scalar, Lorentz Covariant Theories of Gravitation*, pp. 1337–1411. Springer Netherlands, 2007. 10.1007/978-1-4020-4000-9_27. [Pages 8 and 9]

[47] A. Einstein, *On the relativity principle and the conclusions drawn from it*, *Jahrbuch der Radioaktivitat und Elektronik* **4** (1907) . [Page 8]

[48] A. Einstein, *Über den einfluß der schwerkraft auf die ausbreitung des lichtes*, *Annalen der Physik* **340** (1911) 898 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19113401005]. [Page 8]

[49] R. v. Eötvös, *Conference proceedings*, *Verhandlungen der 16 Allgemeinen Konferenz der Internationalen Erdmessung* (1910) . [Pages 9 and 12]

[50] A. Einstein, *Is there a gravitational effect that is analogous to electrodynamic induction?*, *Vierteljahrsschrift fur gerichtliche Medizin und offentliches Sanitatswesen* **44** (1912) . [Page 9]

[51] M. Abraham, *Relativität und gravitation. erwiderung auf eine bemerkung des hrn. a. einstein*, *Annalen der Physik* **343** (1912) 1056 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19123431013]. [Page 9]

[52] M. Abraham, *Nochmals relativität und gravitation. bemerkungen zu a. einsteins erwiderung*, *Annalen der Physik* **344** (1912) 444 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19123441208]. [Page 9]

[53] G. Nordström, *Relativitätsprinzip und gravitation*, *Physikalische Zeitschrift* (1912) . [Page 9]

[54] A. Einstein and M. Grossman, *Outline of a generalized theory of relativity and of a theory of gravitation*, *Zeitschrift für Mathematik und Physik* **62** (1913) . [Page 9]

[55] M. von Laue, *Zur dynamik der relativitätstheorie*, *Annalen der Physik* **35** (1911) . [Page 9]

[56] M. von Laue, *in beispiel zur dynamik der relativitätstheorie*, *erhandlungen der Deutschen Physikalischen Gesellschaft* (1911) . [Page 9]

[57] G. Nordström, *Träge und schwere masse in der relativitätsmechanik*, *Annalen der Physik* (1913) . [Page 10]

[58] G. Nordström, *Zur theorie der gravitation vom standpunkt des relativitätsprinzip*, *Annalen der Physik* **42** (1913) . [Page 10]

[59] A. Einstein and A. D. Fokker, *ie nordströmsche gravitationstheorie vom standpunkt des absoluten differentialkalküls.*, *Annalen der Physik* **44** (1913) . [Page 10]

[60] G. Nordström, *über die möglichkeit, das elektromagnetische feld und das gravitationsfeld zu vereinigen*, *Physikalische Zeitschrift* **15** (1914) . [Page 10]

[61] G. Nordström, *Zur elektrizitäts- und gravitationstheorie*, *Ofversigt af Finska Vetenskaps-Societetens Förhandlingar* **57** (1914) . [Page 10]

[62] G. Nordström, *über eine mögliche grundlage einer theorie der materie*, *Ofversigt af Finska Vetenskaps-Societetens Förhandlingar* **57** (1915) . [Page 10]

[63] T. Kaluza, *Zum Unitätsproblem der Physik*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin* (1921) 966. [Page 10]

[64] O. Klein, *Quantentheorie und fünfdimensionale Relativitätstheorie*, *Zeitschrift fur Physik* **37** (1926) 895. [Page 10]

[65] O. Klein, *The Atomicity of Electricity as a Quantum Theory Law*, Nature **118** (1926) 516. [Page 10]

[66] C. D. Perrine, *Contribution to the history of attempts to test the theory of relativity by means of astronomical observations*, *Astronomische Nachrichten* **219** (1923) 281 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/asna.19232191706]. [Pages 10 and 12]

[67] 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** (1920) 291. [Pages 10 and 12]

[68] 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*, PASP **35** (1923) 158. [Pages 10 and 12]

[69] U. Le Verrier, *Lettre de m. le verrier à m. faye sur la théorie de mercure et sur le mouvement du périhélie de cette planète*, *Comptes rendus hebdomadaires des séances de l'Académie des sciences (Paris)* **49** (1859) . [Pages 10 and 12]

[70] A. Einstein, *Die grundlage der allgemeinen relativitätstheorie*, *Annalen der Physik* **354** (1916) 769 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19163540702]. [Pages 10 and 12]

[71] D. Hilbert, *Die grundlagen der physik*, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen – Mathematisch-Physikalische Klasse* **3** (1915) . [Page 11]

[72] A. A. Starobinsky, *A New Type of Isotropic Cosmological Models Without Singularity*, *Phys. Lett. B* **91** (1980) 99. [Page 12]

[73] A. Vilenkin, *Classical and Quantum Cosmology of the Starobinsky Inflationary Model*, *Phys. Rev. D* **32** (1985) 2511. [Page 12]

[74] C. Skordis, D. F. Mota, P. G. Ferreira and C. Boehm, *Large Scale Structure in Bekenstein's theory of relativistic Modified Newtonian Dynamics*, *Phys. Rev. Lett.* **96** (2006) 011301 [astro-ph/0505519]. [Page 12]

[75] A. Pourtsidou, C. Skordis and E. J. Copeland, *Models of dark matter coupled to dark energy*, *Phys. Rev. D* **88** (2013) 083505 [1307.0458]. [Page 12]

[76] M. Milgrom, *A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis.*, ApJ **270** (1983) 365. [Pages 12 and 38]

[77] J. D. Bekenstein, *Relativistic gravitation theory for the modified newtonian dynamics paradigm*, *Phys. Rev. D* **70** (2004) 083509. [Pages 12 and 38]

[78] R. Utiyama and B. S. DeWitt, *Renormalization of a classical gravitational field interacting with quantized matter fields*, *J. Math. Phys.* **3** (1962) 608. [Pages 12, 38, and 60]

[79] Y. Kim, A. Kobakhidze and Z. S. C. Picker, *Probing Quadratic Gravity with Binary Inspirals*, *Eur. Phys. J. C* **81** (2021) 362 [1906.12034]. [Pages 12 and 60]

[80] E. Landau, *A total solar eclipse 100 years ago proved einstein's general relativity*, May, 2019. [Pages 12 and 13]

[81] D. Powell, *How the 1919 solar eclipse made einstein the world's most famous scientist*, May, 2019. [Pages 12 and 13]

[82] T. Wagner, S. Schlamminger, J. Gundlach and E. Adelberger, *Torsion-balance tests of the weak equivalence principle*, *Class. Quant. Grav.* **29** (2012) 184002 [1207.2442]. [Page 12]

[83] P. Touboul, G. Métris, M. Rodrigues, Y. André, Q. Baghi, J. Bergé, D. Boulanger, S. Bremer, P. Carle, R. Chhun, B. Christophe, V. Cipolla, T. Damour, P. Danto, H. Dittus, P. Fayet, B. Foulon, C. Gageant, P.-Y. Guidotti, D. Hagedorn, E. Hardy, P.-A. Huynh, H. Inchauspe, P. Kayser, S. Lala, C. Lämmerzahl, V. Lebat, P. Leseur, F. m. c. Liorzou, M. List, F. Löffler, I. Panet, B. Pouilloux, P. Prieur, A. Rebray, S. Reynaud, B. Rievers, A. Robert, H. Selig, L. Serron, T. Sumner, N. Tanguy and P. Visser, *Microscope mission: First results of a space test of the equivalence principle*, *Phys. Rev. Lett.* **119** (2017) 231101. [Page 12]

[84] N. Ashby, *Relativity in the Global Positioning System*, *Living Rev. Rel.* **6** (2003) 1. [Page 12]

[85] A. Einstein, *Näherungsweise Integration der Feldgleichungen der Gravitation*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin* (1916) 688. [Pages 12 and 40]

[86] A. Einstein, *Über Gravitationswellen*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin* (1918) 154. [Pages 12 and 40]

[87] LIGO Scientific Collaboration, B. P. Abbott et al., *LIGO: The Laser interferometer gravitational-wave observatory*, *Rept. Prog. Phys.* **72** (2009) 076901 [0711.3041]. [Pages 12, 40, 49, 50, 51, and 85]

[88] VIRGO Collaboration, F. Acernese et al., *Advanced Virgo: a second-generation interferometric gravitational wave detector*, *Class. Quant. Grav.* **32** (2015) 024001 [1408.3978]. [Pages 12, 40, 49, 50, 51, and 85]

[89] LIGO Scientific, Virgo Collaboration, B. 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*, *Phys. Rev. X* **9** (2019) 031040 [1811.12907]. [Pages 12, 40, 49, 50, 51, 85, and 91]

[90] LIGO Scientific, Virgo Collaboration, B. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016) 061102 [1602.03837]. [Pages 12, 42, 51, and 55]

[91] G. M. Bernstein, J. A. Tyson and C. S. Kochanek, *A Large Arc in the Gravitational Lens System 0957+561*, AJ **105** (1993) 816. [Pages 13 and 34]

[92] F. De Marchi and G. Cascioli, *Testing General Relativity in the Solar System: present and future perspectives*, *Class. Quant. Grav.* **37** (2020) 095007 [1911.05561]. [Page 13]

[93] K. Nordtvedt, *Testing relativity with laser ranging to the moon*, *Phys. Rev.* **170** (1968) 1186. [Page 13]

[94] J. G. Williams, S. G. Turyshev and D. H. Boggs, *Progress in lunar laser ranging tests of relativistic gravity*, *Phys. Rev. Lett.* **93** (2004) 261101. [Page 13]

[95] H. Bondi, *Cosmology.* 1952. [Page 14]

[96] A. Friedmann, *Über die Krümmung des Raumes*, *Zeitschrift fur Physik* **10** (1922) 377. [Pages 14, 15, and 16]

[97] A. Friedmann, *Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes*, *Zeitschrift fur Physik* **21** (1924) 326. [Pages 14 and 15]

[98] G. Lemaître, *Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques*, *Annales de la Société Scientifique de Bruxelles* **47** (1927) 49. [Pages 14 and 16]

[99] G. Lemaitre, *The expanding universe*, *Annales Soc. Sci* **53** (1933) 51. [Pages 14, 41, and 65]

[100] H. P. Robertson, *Kinematics and World-Structure*, ApJ **82** (1935) 284. [Page 14]

[101] H. P. Robertson, *Kinematics and World-Structure III.*, ApJ **83** (1936) 257. [Page 14]

[102] H. P. Robertson, *Kinematics and World-Structure II.*, ApJ **83** (1936) 187. [Page 14]

[103] A. G. Walker, *On milne's theory of world-structure**, *Proceedings of the London Mathematical Society* **s2-42** (1937) 90 [https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/plms/s2-42.1.90]. [Page 14]

[104] E. Hubble, *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*, *Proceedings of the National Academy of Science* **15** (1929) 168. [Pages 15 and 16]

[105] V. M. Slipher, *The radial velocity of the Andromeda Nebula*, *Lowell Observatory Bulletin* **1** (1913) 56. [Page 16]

[106] A. G. Lemaître, *Contributions to a British Association Discussion on the Evolution of the Universe.*, Nature **128** (1931) 704. [Page 16]

[107] R. A. Alpher, H. Bethe and G. Gamow, *The origin of chemical elements*, *Phys. Rev.* **73** (1948) 803. [Pages 16, 18, and 35]

[108] R. A. Alpher and R. Herman, *Evolution of the Universe*, Nature **162** (1948) 774. [Page 16]

[109] H. Georgi and A. Pais, *Calculability and naturalness in gauge theories*, *Phys. Rev. D* **10** (1974) 539. [Pages 16 and 21]

[110] A. Buras, J. Ellis, M. Gaillard and D. Nanopoulos, *Aspects of the grand unification of strong, weak and electromagnetic interactions*, *Nuclear Physics B* **135** (1978) 66. [Pages 16 and 21]

[111] A. Salam and J. Ward, *Electromagnetic and weak interactions*, *Physics Letters* **13** (1964) 168. [Pages 16, 21, and 28]

[112] S. Weinberg, *A model of leptons*, *Phys. Rev. Lett.* **19** (1967) 1264. [Pages 16, 21, and 28]

[113] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry*, . [Pages 16 and 21]

[114] M. Gell-Mann, *A schematic model of baryons and mesons*, *Physics Letters* **8** (1964) 214. [Pages 16 and 21]

[115] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the color octet gluon picture*, *Physics Letters B* **47** (1973) 365. [Pages 16 and 21]

[116] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, *Phys. Rev. Lett.* **30** (1973) 1343. [Pages 16 and 21]

[117] H. D. Politzer, *Reliable perturbative results for strong interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346. [Pages 16 and 21]

[118] J. Yang, M. S. Turner, G. Steigman, D. N. Schramm and K. A. Olive, *Primordial nucleosynthesis : a critical comparison of theory and observation.*, ApJ **281** (1984) 493. [Pages 16 and 35]

[119] T. P. Walker, G. Steigman, D. N. Schramm, K. A. Olive and H.-S. Kang, *Primordial Nucleosynthesis Redux*, ApJ **376** (1991) 51. [Pages 16 and 35]

[120] B. D. Fields, P. Molaro and S. Sarkar, *Big-Bang Nucleosynthesis*, *Chin. Phys. C* **38** (2014) 339 [1412.1408]. [Pages 16 and 35]

[121] A. G. Doroshkevich and I. D. Novikov, *Republication of: Mean density of radiation in the Metagalaxy and certain problems in relativistic cosmology*, *General Relativity and Gravitation* **50** (2018) 138. [Page 16]

[122] R. H. Dicke, *The Measurement of Thermal Radiation at Microwave Frequencies*, *Review of Scientific Instruments* **17** (1946) 268. [Page 16]

[123] A. A. Penzias and R. W. Wilson, *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*, ApJ **142** (1965) 419. [Page 16]

[124] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss and D. T. Wilkinson, *Structure in the COBE Differential Microwave Radiometer First-Year Maps*, ApJ **396** (1992) L1. [Pages 16, 17, and 35]

[125] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss and D. T. Wilkinson, *Structure in the COBE Differential Microwave Radiometer First-Year Maps*, ApJ **396** (1992) L1. [Pages 16, 17, and 18]

[126] Planck Collaboration, Y. Akrami et al., *Planck 2018 results. X. Constraints on inflation*, *Astron. Astrophys.* **641** (2020) A10 [1807.06211]. [Pages 16, 17, 18, 35, and 148]

[127] T. Giannantonio, R. Scranton, R. G. Crittenden, R. C. Nichol, S. P. Boughn, A. D. Myers and G. T. Richards, *Combined analysis of the integrated Sachs-Wolfe effect and cosmological implications*, *Phys. Rev. D* **77** (2008) 123520 [0801.4380]. [Page 16]

[128] P. J. E. Peebles and J. T. Yu, *Primeval adiabatic perturbation in an expanding universe*, *Astrophys. J.* **162** (1970) 815. [Pages 17 and 35]

[129] J. E. Gunn and B. A. Peterson, *On the Density of Neutral Hydrogen in Intergalactic Space.*, ApJ **142** (1965) 1633. [Page 17]

[130] SDSS Collaboration, R. H. Becker et al., *Evidence for Reionization at Z ~ 6: Detection of a Gunn-Peterson trough in a Z = 6.28 Quasar*, *Astron. J.* **122** (2001) 2850 [astro-ph/0108097]. [Page 17]

[131] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff and J. Tonry, *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, AJ **116** (1998) 1009 [astro-ph/9805201]. [Pages 17 and 18]

[132] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch and T. S. C. Project, *Measurements of* Ω *and* Λ *from 42 High-Redshift Supernovae*, ApJ **517** (1999) 565 [astro-ph/9812133]. [Pages 17 and 18]

[133] J.-P. Dai, Y. Yang and J.-Q. Xia, *Reconstruction of the Dark Energy Equation of State from the Latest Observations*, ApJ **857** (2018) 9. [Page 17]

[134] W. Rindler, *Visual horizons in world models*, MNRAS **116** (1956) 662. [Page 18]

[135] A. A. Starobinsky, *Spectrum of relict gravitational radiation and the early state of the universe*, *ZhETF Pisma Redaktsiiu* **30** (1979) 719. [Pages 18 and 38]

[136] A. H. Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, *Phys. Rev. D* **23** (1981) 347. [Page 18]

[137] D. Kazanas, *Dynamics of the universe and spontaneous symmetry breaking*, ApJ **241** (1980) L59. [Page 18]

[138] K. Sato, *Cosmological Baryon Number Domain Structure and the First Order Phase Transition of a Vacuum*, *Phys. Lett. B* **99** (1981) 66. [Page 18]

[139] M. B. Einhorn and K. Sato, *Monopole Production in the Very Early Universe in a First Order Phase Transition*, *Nucl. Phys. B* **180** (1981) 385. [Page 18]

[140] L. Kofman, A. D. Linde and A. A. Starobinsky, *Reheating after inflation*, *Phys. Rev. Lett.* **73** (1994) 3195 [hep-th/9405187]. [Page 18]

[141] V. F. Mukhanov and G. V. Chibisov, *Quantum fluctuations and a nonsingular universe*, *Soviet Journal of Experimental and Theoretical Physics Letters* **33** (1981) 532. [Page 18]

[142] V. F. Mukhanov and G. V. Chibisov, *Vacuum energy and large-scale structure of the Universe*, *Soviet Journal of Experimental and Theoretical Physics* **56** (1982) 258. [Page 18]

[143] S. Hawking, *The development of irregularities in a single bubble inflationary universe*, *Physics Letters B* **115** (1982) 295. [Page 18]

[144] A. Starobinsky, *Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations*, *Physics Letters B* **117** (1982) 175. [Page 18]

[145] A. H. Guth and S.-Y. Pi, *Fluctuations in the new inflationary universe*, *Phys. Rev. Lett.* **49** (1982) 1110. [Page 18]

[146] J. M. Bardeen, P. J. Steinhardt and M. S. Turner, *Spontaneous creation of almost scale-free density perturbations in an inflationary universe*, *Phys. Rev. D* **28** (1983) 679. [Page 18]

[147] P. Ivanov, P. Naselsky and I. Novikov, *Inflation and primordial black holes as dark matter*, *Phys. Rev. D* **50** (1994) 7173. [Pages 18 and 56]

[148] WMAP Collaboration, E. Komatsu et al., *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *Astrophys. J. Suppl.* **192** (2011) 18 [1001.4538]. [Page 18]

[149] A. et al, *The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data*, ApJS **259** (2022) 35 [2112.02026]. [Pages 18, 34, and 35]

[150] D. Collaboration, *The DESI Experiment Part I: Science,Targeting, and Survey Design*, *arXiv e-prints* (2016) arXiv:1611.00036 [1611.00036]. [Pages 18, 34, and 35]

[151] 2ᴅFGRS Collaboration, S. Cole et al., *The 2dF Galaxy Redshift Survey: Power-spectrum analysis of the final dataset and cosmological implications*, *Mon. Not. Roy. Astron. Soc.* **362** (2005) 505 [astro-ph/0501174]. [Pages 18, 34, and 35]

[152] EUCLID Collaboration, R. Laureijs et al., *Euclid Definition Study Report*, 1110.3193. [Page 18]

[153] LSST Collaboration, v. Ivezić et al., *LSST: from Science Drivers to Reference Design and Anticipated Data Products*, *Astrophys. J.* **873** (2019) 111 [0805.2366]. [Pages 18, 34, and 35]

[154] DES Collaboration, E. Macaulay et al., *First Cosmological Results using Type Ia Supernovae from the Dark Energy Survey: Measurement of the Hubble Constant*, *Mon. Not. Roy. Astron. Soc.* **486** (2019) 2184 [1811.02376]. [Page 18]

[155] P. Gondolo and G. Gelmini, *Cosmic abundances of stable particles: Improved analysis*, *Nucl. Phys. B* **360** (1991) 145. [Page 18]

[156] G. Steigman, *Primordial Nucleosynthesis in the Precision Cosmology Era*, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 463 [0712.1100]. [Page 18]

[157] L. Perivolaropoulos and F. Skara, *Challenges for ΛCDM: An update*, 2105.05208. [Pages 19 and 35]

[158] W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, *Zeitschrift fur Physik* **43** (1927) 172. [Page 21]

[159] M. Planck, *Ueber das gesetz der energieverteilung im normalspectrum*, *Annalen der Physik* **309** (1901) 553 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19013090310]. [Page 21]

[160] A. Einstein, *Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt*, *Annalen der Physik* **322** (1905) 132 [https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19053220607]. [Page 21]

[161] N. Bohr, *On the Constitution of Atoms and Molecules. Pt.II*, *Philosophical Magazine* **26** (1913) 476. [Page 21]

[162] R. A. Millikan, *A direct photoelectric determination of planck's "h"*, *Phys. Rev.* **7** (1916) 355. [Page 21]

[163] Particle Data Group Collaboration, P. A. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01. [Pages 21 and 113]

[164] A. Pais and S. B. Treiman, *How many charm quantum numbers are there?*, *Phys. Rev. Lett.* **35** (1975) 1556. [Page 21]

[165] E. Gildener, *Gauge-symmetry hierarchies*, *Phys. Rev. D* **14** (1976) 1667. [Page 21]

[166] L. Susskind, *Dynamics of spontaneous symmetry breaking in the weinberg-salam theory*, *Phys. Rev. D* **20** (1979) 2619. [Page 21]

[167] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59** (1980) 135. [Page 21]

[168] M. J. G. Veltman, *The Infrared - Ultraviolet Connection*, *Acta Phys. Polon. B* **12** (1981) 437. [Page 21]

[169] H. Harari, *Three generations of quarks and leptons*, *V International Conference on Experimental Meson Spectroscopy* (1974) . [Page 21]

[170] B. Holdom, W. S. Hou, T. Hurth, M. L. Mangano, S. Sultansoy and G. Unel, *Four Statements about the Fourth Generation*, *PMC Phys. A* **3** (2009) 4 [0904.4698]. [Page 21]

[171] B. Pontecorvo, *Mesonium and Antimesonium*, *Soviet Journal of Experimental and Theoretical Physics* **6** (1958) 429. [Page 21]

[172] R. Davis, D. S. Harmer and K. C. Hoffman, *Search for neutrinos from the sun*, *Phys. Rev. Lett.* **20** (1968) 1205. [Page 21]

[173] SNO Collaboration Collaboration, Q. R. Ahmad et al., *Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8b solar neutrinos at the sudbury neutrino observatory*, *Phys. Rev. Lett.* **87** (2001) 071301. [Page 21]

[174] F. Capozzi, E. Lisi, A. Marrone, D. Montanino and A. Palazzo, *Neutrino masses and mixings: Status of known and unknown 3ν parameters*, *Nuclear Physics B* **908** (2016) 218. Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015. [Page 21]

[175] S. Weinberg, *The u(1) problem*, *Phys. Rev. D* **11** (1975) 3583. [Pages 21 and 100]

[176] G. 't Hooft, *Symmetry breaking through bell-jackiw anomalies*, *Phys. Rev. Lett.* **37** (1976) 8. [Pages 21 and 100]

[177] S. L. Adler, *Axial-vector vertex in spinor electrodynamics*, *Phys. Rev.* **177** (1969) 2426. [Pages 21, 103, and 104]

[178] J. S. Bell and R. Jackiw, *A PCAC puzzle: $\pi 0 \to \gamma\gamma$ in the σ-model*, *Nuovo Cimento A Serie* **60** (1969) 47. [Pages 21, 103, and 104]

[179] W. A. Bardeen, *Anomalous Currents in Gauge Field Theories*, *Nucl. Phys. B* **75** (1974) 246. [Pages 21 and 100]

[180] V. Baluni, CP-*nonconserving effects in quantum chromodynamics*, *Phys. Rev. D* **19** (1979) 2227. [Pages 21 and 111]

[181] R. Jackiw and C. Rebbi, *Vacuum periodicity in a yang-mills quantum theory*, *Phys. Rev. Lett.* **37** (1976) 172. [Pages 21 and 100]

[182] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32. [Page 21]

[183] J. D. Barrow and M. S. Turner, *Baryosynthesis and the Origin of Galaxies*, *Nature* **291** (1981) 469. [Page 21]

[184] M. S. Turner, *Big band baryosynthesis and grand unification*, in *Weak Interactions as Probes of Unification*, vol. 72 of *American Institute of Physics Conference Series*, pp. 224–243, May, 1981, DOI. [Page 21]

[185] Muon g-2 Collaboration, B. Abi et al., *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (2021) 141801 [2104.03281]. [Page 21]

[186] LHCb Collaboration, R. Aaij et al., *Test of lepton universality in beauty-quark decays*, *Nature Phys.* **18** (2022) 277 [2103.11769]. [Page 21]

[187] H. Miyazawa, *Baryon Number Changing Currents*, *Prog. Theor. Phys.* **36** (1966) 1266. [Pages 21 and 37]

[188] H. Miyazawa, *Spinor currents and symmetries of baryons and mesons*, *Phys. Rev.* **170** (1968) 1586. [Pages 21 and 37]

[189] D. V. Volkov and V. P. Akulov, *Possible universal neutrino interaction*, *JETP Lett.* **16** (1972) 438. [Pages 21 and 37]

[190] J. A. Wheeler, *On the mathematical description of light nuclei by the method of resonating group structure*, *Phys. Rev.* **52** (1937) 1107. [Page 21]

[191] W. Heisenberg, *Die beobachtbaren Größen in der Theorie der Elementarteilchen. II*, *Zeitschrift fur Physik* **120** (1943) 673. [Page 21]

[192] J. Scherk and J. H. Schwarz, *Dual Models for Nonhadrons*, *Nucl. Phys. B* **81** (1974) 118. [Page 21]

[193] T. Yoneya, *Connection of Dual Models to Electrodynamics and Gravidynamics*, *Progress of Theoretical Physics* **51** (1974) 1907 [https://academic.oup.com/ptp/article-pdf/51/6/1907/5275734/51-6-1907.pdf]. [Page 21]

[194] M. B. Green and J. H. Schwarz, *Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory*, *Phys. Lett. B* **149** (1984) 117. [Page 21]

[195] D. J. Gross, J. A. Harvey, E. J. Martinec and R. Rohm, *The Heterotic String*, *Phys. Rev. Lett.* **54** (1985) 502. [Page 21]

[196] W. Nahm, *Supersymmetries and their Representations*, *Nucl. Phys. B* **135** (1978) 149. [Page 21]

[197] E. Witten, *String theory dynamics in various dimensions*, *Nucl. Phys. B* **443** (1995) 85 [hep-th/9503124]. [Page 21]

[198] P. Jordan, *Über eine neue Begründung der Quantenmechanik*, *Zeitschrift fur Physik* **40** (1927) 809. [Page 22]

[199] P. A. M. Dirac and N. H. D. 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** (1927) 243 [https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1927.0039]. [Pages 22 and 26]

[200] P. Jordan and W. Pauli, *Zur Quantenelektrodynamik ladungsfreier Felder*, *Zeitschrift fur Physik* **47** (1928) 151. [Page 22]

[201] E. Fermi, *Versuch einer Theorie der β-Strahlen. I*, *Zeitschrift fur Physik* **88** (1934) 161. [Page 22]

[202] H. A. Bethe, *The electromagnetic shift of energy levels*, *Phys. Rev.* **72** (1947) 339. [Pages 22, 26, and 30]

[203] J. Schwinger, *On quantum-electrodynamics and the magnetic moment of the electron*, *Phys. Rev.* **73** (1948) 416. [Pages 22, 26, and 30]

[204] R. P. Feynman, *Space-time approach to non-relativistic quantum mechanics*, *Rev. Mod. Phys.* **20** (1948) 367. [Pages 22, 26, and 30]

[205] R. P. Feynman, *A relativistic cut-off for classical electrodynamics*, *Phys. Rev.* **74** (1948) 939. [Pages 22, 26, and 30]

[206] S. Tomonaga, *On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields**, *Progress of Theoretical Physics* **1** (1946) 27 [https://academic.oup.com/ptp/article-pdf/1/2/27/24027031/1-2-27.pdf]. [Pages 22 and 26]

[207] F. J. Dyson, *The radiation theories of tomonaga, schwinger, and feynman*, *Phys. Rev.* **75** (1949) 486. [Pages 22 and 26]

[208] V. Fock, *Konfigurationsraum und zweite Quantelung*, *Zeitschrift fur Physik* **75** (1932) 622. [Page 22]

[209] J. H. V. Vleck, *The correspondence principle in the statistical interpretation of quantum mechanics*, *Proceedings of the National Academy of Sciences* **14** (1928) 178 [https://www.pnas.org/doi/pdf/10.1073/pnas.14.2.178]. [Page 22]

[210] P. A. M. Dirac, *The Lagrangian in Quantum Mechanics*, in *Feynman's Thesis - A New Approach to Quantum Theory*, pp. 111–119. World Scientific Publishing Co. Pte. Ltd, 2005. DOI. [Page 22]

[211] D. Kaiser, *Physics and feynman's diagrams*, *American Scientist* **93** (2005) . [Page 22]

[212] E. Noether, *Invariante variationsprobleme*, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918** (1918) 235. [Page 23]

[213] W. Pauli, *Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren*, *Zeitschrift fur Physik* **31** (1925) 765. [Page 24]

[214] H. Weyl, *Gravitation und Elektrizität*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin* (1918) 465. [Page 24]

[215] H. Weyl, *Elektron und Gravitation. I*, *Zeitschrift fur Physik* **56** (1929) 330. [Pages 24 and 29]

[216] W. Pauli, *Relativistic Field Theories of Elementary Particles*, *Rev. Mod. Phys.* **13** (1941) 203. [Page 24]

[217] C. N. Yang and R. L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, *Phys. Rev.* **96** (1954) 191. [Page 24]

[218] E. Fermi, *Quantum theory of radiation*, *Rev. Mod. Phys.* **4** (1932) 87. [Page 26]

[219] R. P. Feynman, *Mathematical formulation of the quantum theory of electromagnetic interaction*, *Phys. Rev.* **80** (1950) 440. [Page 26]

[220] L. Morel, Z. Yao, P. Cladé and S. Guellati-Khélifa, *Determination of the fine-structure constant with an accuracy of 81 parts per trillion*, *Nature* **588** (2020) 61. [Page 26]

[221] L. Faddeev and V. Popov, *Feynman diagrams for the yang-mills field*, *Physics Letters B* **25** (1967) 29. [Page 27]

[222] C. Becchi, A. Rouet and R. Stora, *The Abelian Higgs-Kibble Model. Unitarity of the S Operator*, *Phys. Lett. B* **52** (1974) 344. [Page 27]

[223] I. V. Tyutin, *Gauge Invariance in Field Theory and Statistical Physics in Operator Formalism*, 0812.0580. [Page 27]

[224] P. Teller, *The gauge argument*, *Philosophy of Science* **67** (2000) S466. [Page 27]

[225] A. Guay, *The arbitrariness of local gauge symmetry*, December, 2004. In this version, some references have been corrected. [Page 27]

[226] M. Redhead, *The interpretation of gauge symmetry*, in *Symmetries in Physics: Philosophical Reflections* (K. A. Brading and E. Castellani, eds.), pp. 124–139. Cambridge University Press, 2003. [Page 27]

[227] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321. [Page 28]

[228] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508. [Page 28]

[229] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585. [Page 28]

[230] A. Polyakov and A. Migdal, *Spontaneous breakdown of strong interaction symmetry and absence of massless particles in soviet physics*, *JETP* **24** (1967) . [Page 28]

[231] Y. Nambu, *Quasi-particles and gauge invariance in the theory of superconductivity*, *Phys. Rev.* **117** (1960) 648. [Page 28]

[232] J. Goldstone, *Field Theories with Superconductor Solutions*, *Nuovo Cim.* **19** (1961) 154. [Page 28]

[233] J. Goldstone, A. Salam and S. Weinberg, *Broken Symmetries*, *Phys. Rev.* **127** (1962) 965. [Page 28]

[234] H. Yukawa, *On the Interaction of Elementary Particles I*, *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48. [Page 28]

[235] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Experimental test of parity conservation in beta decay*, *Phys. Rev.* **105** (1957) 1413. [Page 29]

[236] M. Goldhaber, L. Grodzins and A. W. Sunyar, *Helicity of Neutrinos*, *Phys. Rev.* **109** (1958) 1015. [Page 29]

[237] L. D. Landau, *On the conservation laws for weak interactions*, *Nucl. Phys.* **3** (1957) 127. [Page 29]

[238] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Evidence for the 2π decay of the k_2^0 meson*, *Phys. Rev. Lett.* **13** (1964) 138. [Page 29]

[239] J. Schwinger, *The theory of quantized fields. i*, *Phys. Rev.* **82** (1951) 914. [Page 29]

[240] G. Luders, *On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories*, *Kong. Dan. Vid. Sel. Mat. Fys. Med.* **28N5** (1954) 1. [Page 29]

[241] J. Yepez, *Einstein's vierbein field theory of curved space*, 1106.2037. [Page 29]

[242] M. Blagojevic and F. W. Hehl, *Gauge Theories of Gravitation*, 1210.3775. [Page 30]

[243] F. W. Hehl, Y. N. Obukhov and D. Puetzfeld, *On Poincaré gauge theory of gravity, its equations of motion, and Gravity Probe B*, *Phys. Lett. A* **377** (2013) 1775 [1304.2769]. [Page 30]

[244] R. P. Feynman, *Quantum theory of gravitation*, *Acta Phys. Polon.* **24** (1963) 697. [Page 30]

[245] B. S. DeWitt, *Quantum Theory of Gravity. 2. The Manifestly Covariant Theory*, *Phys. Rev.* **162** (1967) 1195. [Page 30]

[246] H. van Dam and M. J. G. Veltman, *Massive and massless Yang-Mills and gravitational fields*, *Nucl. Phys. B* **22** (1970) 397. [Page 30]

[247] E. C. G. Stueckelberg de Breidenbach and A. Petermann, *Normalization of constants in the quanta theory*, *Helv. Phys. Acta* **26** (1953) 499. [Page 30]

[248] L. P. Kadanoff, *Scaling laws for ising models near T_c*, *Physics Physique Fizika* **2** (1966) 263. [Page 30]

[249] M. Gell-Mann and F. E. Low, *Quantum electrodynamics at small distances*, *Phys. Rev.* **95** (1954) 1300. [Page 30]

[250] C. G. Callan, *Broken scale invariance in scalar field theory*, *Phys. Rev. D* **2** (1970) 1541. [Page 30]

[251] K. Symanzik, *Small distance behavior in field theory and power counting*, *Commun. Math. Phys.* **18** (1970) 227. [Page 30]

[252] K. G. Wilson, *Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture*, *Phys. Rev. B* **4** (1971) 3174. [Page 30]

[253] A. Shomer, *A Pedagogical explanation for the non-renormalizability of gravity*, 0709.3555. [Page 31]

[254] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helvetica Physica Acta* **6** (1933) 110. [Pages 33 and 34]

[255] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, ApJ **86** (1937) 217. [Pages 33 and 34]

[256] W. Thomson, Baron Kelvin, *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, Cambridge Library Collection - Physical Sciences. Cambridge University Press, 2010, 10.1017/CBO9780511694523. [Page 33]

[257] J. C. Kapteyn, *First Attempt at a Theory of the Arrangement and Motion of the Sidereal System*, ApJ **55** (1922) 302. [Page 33]

[258] K. Lundmark, *Über die Bestimmung der Entfernungen, Dimensionen, Massen und Dichtigkeit fur die nächstgelegenen anagalacktischen Sternsysteme.*, *Meddelanden fran Lunds Astronomiska Observatorium Serie I* **125** (1930) 1. [Page 33]

[259] J. H. Oort, *The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems*, Bull. Astron. Inst. Netherlands **6** (1932) 249. [Page 33]

[260] D. H. Rogstad and G. S. Shostak, *Gross Properties of Five Scd Galaxies as Determined from 21-CENTIMETER Observations*, ApJ **176** (1972) 315. [Page 33]

[261] K. C. Freeman, *On the Disks of Spiral and S0 Galaxies*, ApJ **160** (1970) 811. [Page 33]

[262] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, ApJ **159** (1970) 379. [Page 33]

[263] V. C. Rubin, J. Ford, W. K. and N. Thonnard, *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc).*, ApJ **238** (1980) 471. [Page 33]

[264] M. S. Roberts, *A High-Resolution 21-CM Hydrogen-Line Survey of the Andromeda Nebula*, ApJ **144** (1966) 639. [Page 33]

[265] M. S. Roberts and R. N. Whitehurst, *The rotation curve and geometry of M31 at large galactocentric distances.*, ApJ **201** (1975) 327. [Page 33]

[266] N. P. Pitjev and E. V. Pitjeva, *Constraints on dark matter in the solar system*, *Astronomy Letters* **39** (2013) 141 [1306.5534]. [Page 33]

[267] E. Belbruno and J. Green, *When leaving the Solar system: Dark matter makes a difference*, MNRAS **510** (2022) 5154 [2201.06575]. [Page 33]

[268] J. I. Read, *The Local Dark Matter Density*, *J. Phys. G* **41** (2014) 063101 [1404.1938]. [Page 33]

[269] P. F. de Salas and A. Widmark, *Dark matter local density determination: recent observations and future prospects*, *Rept. Prog. Phys.* **84** (2021) 104901 [2012.11477]. [Pages 33 and 35]

[270] S. M. Faber and R. E. Jackson, *Velocity dispersions and mass-to-light ratios for elliptical galaxies.*, ApJ **204** (1976) 668. [Page 33]

[271] M. Schmidt, *The distribution of mass in M 31*, Bull. Astron. Inst. Netherlands **14** (1957) 17. [Page 33]

[272] L. M. J. S. Volders, *Neutral hydrogen in M 33 and M 101*, Bull. Astron. Inst. Netherlands **14** (1959) 323. [Page 33]

[273] V. C. Rubin, J. Ford, W. K. and N. Thonnard, *Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa -> Sc.*, ApJ **225** (1978) L107. [Page 33]

[274] A. Bosma, *21-cm line studies of spiral galaxies. 2. The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types.*, *Astron. J.* **86** (1981) 1825. [Page 33]

[275] F. Lelli, S. S. McGaugh and J. M. Schombert, *SPARC: Mass Models for 175 Disk Galaxies with Spitzer Photometry and Accurate Rotation Curves*, *Astron. J.* **152** (2016) 157 [1606.09251]. [Page 33]

[276] R. Carlberg, H. K. C. Yee, E. Ellingson, R. Abraham, P. Gravel, S. Morris and C. J. Pritchet, *Galaxy cluster virial masses and Omega*, *Astrophys. J.* **462** (1996) 32 [astro-ph/9509034]. [Page 34]

[277] S. W. Allen, A. E. Evrard and A. B. Mantz, *Cosmological Parameters from Observations of Galaxy Clusters*, ARA&A **49** (2011) 409 [1103.4829]. [Page 34]

[278] G. L. Bryan and M. L. Norman, *Statistical properties of x-ray clusters: Analytic and numerical comparisons*, *Astrophys. J.* **495** (1998) 80 [astro-ph/9710107]. [Page 34]

[279] A. Vikhlinin, A. Kravtsov, W. Forman, C. Jones, M. Markevitch, S. S. Murray and L. Van Speybroeck, *Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation*, *Astrophys. J.* **640** (2006) 691 [astro-ph/0507092]. [Page 34]

[280] S. Ettori, A. Donnarumma, E. Pointecouteau, T. H. Reiprich, S. Giodini, L. Lovisari and R. W. Schmidt, *Mass profiles of Galaxy Clusters from X-ray analysis*, *Space Sci. Rev.* **177** (2013) 119 [1303.3530]. [Page 34]

[281] A. N. Taylor, S. Dye, T. J. Broadhurst, N. Benitez and E. van Kampen, *Gravitational lens magnification and the mass of abell 1689*, *The Astrophysical Journal* **501** (1998) 539. [Page 34]

[282] M. Oguri, C. E. Rusu and E. E. Falco, *The stellar and dark matter distributions in elliptical galaxies from the ensemble of strong gravitational lenses*, *Mon. Not. Roy. Astron. Soc.* **439** (2014) 2494 [1309.5408]. [Page 34]

[283] D. Harvey, R. Massey, T. Kitching, A. Taylor and E. Tittley, *The non-gravitational interactions of dark matter in colliding galaxy clusters*, *Science* **347** (2015) 1462 [1503.07675]. [Page 34]

[284] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones and D. Zaritsky, *A direct empirical proof of the existence of dark matter*, *Astrophys. J. Lett.* **648** (2006) L109 [astro-ph/0608407]. [Page 34]

[285] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, L. David, W. Forman, C. Jones, S. Murray and W. Tucker, *Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56*, *Astrophys. J.* **606** (2004) 819 [astro-ph/0309303]. [Page 34]

[286] W. A. Dawson, D. Wittman, M. J. Jee, P. Gee, J. P. Hughes, J. A. Tyson, S. Schmidt, P. Thorman, M. Bradač, S. Miyazaki, B. Lemaux, Y. Utsumi and V. E. Margoniner, *Discovery of a Dissociative Galaxy Cluster Merger with Large Physical Separation*, ApJ **747** (2012) L42 [1110.4391]. [Page 34]

[287] G. R. Blumenthal, S. M. Faber, J. R. Primack and M. J. Rees, *Formation of Galaxies and Large Scale Structure with Cold Dark Matter*, *Nature* **311** (1984) 517. [Page 34]

[288] M. Davis, G. Efstathiou, C. S. Frenk and S. D. M. White, *The evolution of large-scale structure in a universe dominated by cold dark matter*, ApJ **292** (1985) 371. [Page 34]

[289] V. Springel et al., *Simulating the joint evolution of quasars, galaxies and their large-scale distribution*, *Nature* **435** (2005) 629 [astro-ph/0504097]. [Page 34]

[290] T. Sawala et al., *The APOSTLE simulations: solutions to the Local Group's cosmic puzzles*, *Mon. Not. Roy. Astron. Soc.* **457** (2016) 1931 [1511.01098]. [Page 34]

[291] J. Schaye et al., *The EAGLE project: Simulating the evolution and assembly of galaxies and their environments*, *Mon. Not. Roy. Astron. Soc.* **446** (2015) 521 [1407.7040]. [Page 34]

[292] R. E. Angulo and O. Hahn, *Large-scale dark matter simulations*, 2112.05165. [Page 34]

[293] B. Moore, *Evidence against dissipationless dark matter from observations of galaxy haloes*, *Nature* **370** (1994) 629. [Page 34]

[294] J. F. Navarro, C. S. Frenk and S. D. M. White, *The Structure of cold dark matter halos*, *Astrophys. J.* **462** (1996) 563 [astro-ph/9508025]. [Page 34]

[295] J. F. Navarro, V. R. Eke and C. S. Frenk, *The cores of dwarf galaxy halos*, *Mon. Not. Roy. Astron. Soc.* **283** (1996) L72 [astro-ph/9610187]. [Page 34]

[296] M. Boylan-Kolchin, J. S. Bullock and M. Kaplinghat, *Too big to fail? The puzzling darkness of massive Milky Way subhaloes*, MNRAS **415** (2011) L40 [1103.0007]. [Page 35]

[297] A. A. Klypin, A. V. Kravtsov, O. Valenzuela and F. Prada, *Where are the missing Galactic satellites?*, *Astrophys. J.* **522** (1999) 82 [astro-ph/9901240]. [Page 35]

[298] A. Pontzen and F. Governato, *How supernova feedback turns dark matter cusps into cores*, MNRAS **421** (2012) 3464 [1106.0499]. [Page 35]

[299] F. Governato, A. Zolotov, A. Pontzen, C. Christensen, S. H. Oh, A. M. Brooks, T. Quinn, S. Shen and J. Wadsley, *Cuspy no more: how outflows affect the central dark matter and baryon distribution in Λ cold dark matter galaxies*, MNRAS **422** (2012) 1231 [1202.0554]. [Page 35]

[300] J. D. Simon and M. Geha, *The Kinematics of the Ultra-Faint Milky Way Satellites: Solving the Missing Satellite Problem*, *Astrophys. J.* **670** (2007) 313 [0706.0516]. [Page 35]

[301] K. Freese, B. Fields and D. Graff, *Limits on stellar objects as the dark matter of our halo: nonbaryonic dark matter seems to be required*, *Nucl. Phys. B Proc. Suppl.* **80** (2000) 0305 [astro-ph/9904401]. [Page 35]

[302] S. Y. Kim, A. H. G. Peter and J. R. Hargis, *Missing Satellites Problem: Completeness Corrections to the Number of Satellite Galaxies in the Milky Way are Consistent with Cold Dark Matter Predictions*, *Phys. Rev. Lett.* **121** (2018) 211302 [1711.06267]. [Page 35]

[303] L. V. Sales, A. Wetzel and A. Fattahi, *Baryonic solutions and challenges for cosmological models of dwarf galaxies*, 2206.05295. [Page 35]

[304] R. A. Sunyaev and Y. B. Zeldovich, *Small scale fluctuations of relic radiation*, *Astrophys. Space Sci.* **7** (1970) 3. [Page 35]

[305] E. Komatsu and U. Seljak, *The Sunyaev-Zel'dovich angular power spectrum as a probe of cosmological parameters*, *Mon. Not. Roy. Astron. Soc.* **336** (2002) 1256 [astro-ph/0205468]. [Page 35]

[306] SDSS Collaboration, D. J. Eisenstein et al., *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*, *Astrophys. J.* **633** (2005) 560 [astro-ph/0501171]. [Page 35]

[307] M. Viel, J. S. Bolton and M. G. Haehnelt, *Cosmological and astrophysical constraints from the Lyman-alpha forest flux probability distribution function*, *Mon. Not. Roy. Astron. Soc.* **399** (2009) L39 [0907.2927]. [Page 35]

[308] N. Vittorio and J. Silk, *Fine-scale anisotropy of the cosmic microwave background in a universe dominated by cold dark matter*, ApJ **285** (1984) L39. [Page 35]

[309] C. S. Frenk and S. D. M. White, *Dark matter and cosmic structure*, *Annalen Phys.* **524** (2012) 507 [1210.0544]. [Page 35]

[310] D. Boyanovsky, H. J. de Vega and N. Sanchez, *Constraints on dark matter particles from theory, galaxy observations and N-body simulations*, *Phys. Rev. D* **77** (2008) 043518 [0710.5180]. [Page 35]

[311] C. Boehm, P. Fayet and R. Schaeffer, *Constraining dark matter candidates from structure formation*, *Phys. Lett. B* **518** (2001) 8 [astro-ph/0012504]. [Page 35]

[312] J. R. Bond, G. Efstathiou and J. Silk, *Massive Neutrinos and the Large-Scale Structure of the Universe*, Phys. Rev. Lett. **45** (1980) 1980. [Pages 35 and 36]

[313] D. N. Spergel, D. H. Weinberg and I. Gott, J. Richard, *Can neutrinos be the galactic missing mass?*, Phys. Rev. D **38** (1988) 2014. [Pages 35 and 36]

[314] S. Hannestad, A. Mirizzi, G. G. Raffelt and Y. Y. Y. Wong, *Neutrino and axion hot dark matter bounds after WMAP-7*, *JCAP* **08** (2010) 001 [1004.0695]. [Pages 35 and 36]

[315] S. D. McDermott, H.-B. Yu and K. M. Zurek, *Turning off the Lights: How Dark is Dark Matter?*, *Phys. Rev. D* **83** (2011) 063509 [1011.2907]. [Page 36]

[316] M. R. Buckley and A. DiFranzo, *Collapsed dark matter structures*, *Phys. Rev. Lett.* **120** (2018) 051102. [Page 36]

[317] C. Boehm, A. Riazuelo, S. H. Hansen and R. Schaeffer, *Interacting dark matter disguised as warm dark matter*, *Phys. Rev. D* **66** (2002) 083505 [astro-ph/0112522]. [Page 36]

[318] C. Boehm and R. Schaeffer, *Constraints on dark matter interactions from structure formation: Damping lengths*, *Astron. Astrophys.* **438** (2005) 419 [astro-ph/0410591]. [Page 36]

[319] C. Boehm, J. A. Schewtschenko, R. J. Wilkinson, C. M. Baugh and S. Pascoli, *Using the Milky Way satellites to study interactions between cold dark matter and radiation*, *Mon. Not. Roy. Astron. Soc.* **445** (2014) L31 [1404.7012]. [Page 36]

[320] R. J. Wilkinson, J. Lesgourgues and C. Boehm, *Using the CMB angular power spectrum to study Dark Matter-photon interactions*, *JCAP* **04** (2014) 026 [1309.7588]. [Page 36]

[321] R. J. Wilkinson, C. Boehm and J. Lesgourgues, *Constraining Dark Matter-Neutrino Interactions using the CMB and Large-Scale Structure*, *JCAP* **05** (2014) 011 [1401.7597]. [Page 36]

[322] E. Di Valentino, C. Bøehm, E. Hivon and F. R. Bouchet, *Reducing the H_0 and σ_8 tensions with Dark Matter-neutrino interactions*, *Phys. Rev. D* **97** (2018) 043513 [1710.02559]. [Page 36]

[323] C. Boehm. personal correspondence. [Page 36]

[324] C. Boehm, P. Fayet and J. Silk, *Light and heavy dark matter particles*, *Phys. Rev. D* **69** (2004) 101302 [hep-ph/0311143]. [Page 36]

[325] C. Boehm and P. Fayet, *Scalar dark matter candidates*, *Nucl. Phys. B* **683** (2004) 219 [hep-ph/0305261]. [Page 36]

[326] S. Profumo, L. Giani and O. F. Piattella, *An Introduction to Particle Dark Matter*, *Universe* **5** (2019) 213 [1910.05610]. [Page 36]

[327] R. Essig et al., *Working Group Report: New Light Weakly Coupled Particles*, in *Community Summer Study 2013: Snowmass on the Mississippi*, 10, 2013, 1311.0029. [Page 36]

[328] R. Foot, *Mirror dark matter: Cosmology, galaxy structure and direct detection*, *Int. J. Mod. Phys. A* **29** (2014) 1430013 [1401.3965]. [Page 36]

[329] B. W. Lee and S. Weinberg, *Cosmological Lower Bound on Heavy Neutrino Masses*, *Phys. Rev. Lett.* **39** (1977) 165. [Page 36]

[330] E. W. Kolb and K. A. Olive, *The Lee-Weinberg Bound Revisited*, *Phys. Rev. D* **33** (1986) 1202. [Erratum: Phys.Rev.D 34, 2531 (1986)]. [Page 36]

[331] S. Dodelson and L. M. Widrow, *Sterile neutrinos as dark matter*, *Phys. Rev. Lett.* **72** (1994) 17. [Page 37]

[332] A. Kusenko, *Detecting sterile dark matter in space*, *Int. J. Mod. Phys. D* **16** (2008) 2325 [astro-ph/0608096]. [Page 37]

[333] K. N. Abazajian et al., *Light Sterile Neutrinos: A White Paper*, 1204.5379. [Page 37]

[334] R. A. Battye and A. Moss, *Evidence for Massive Neutrinos from Cosmic Microwave Background and Lensing Observations*, *Phys. Rev. Lett.* **112** (2014) 051303 [1308.5870]. [Page 37]

[335] T. Yanagida, *Horizontal gauge symmetry and masses of neutrinos*, *Conf. Proc. C* **7902131** (1979) 95. [Page 37]

[336] G. Arcadi, M. Dutra, P. Ghosh, M. Lindner, Y. Mambrini, M. Pierre, S. Profumo and F. S. Queiroz, *The waning of the WIMP? A review of models, searches, and constraints*, *Eur. Phys. J. C* **78** (2018) 203 [1703.07364]. [Page 37]

[337] L. Roszkowski, E. M. Sessolo and S. Trojanowski, *WIMP dark matter candidates and searches—current status and future prospects*, *Rept. Prog. Phys.* **81** (2018) 066201 [1707.06277]. [Page 37]

[338] S. Wolfram, *Abundances of Stable Particles Produced in the Early Universe*, *Phys. Lett. B* **82** (1979) 65. [Page 37]

[339] M. Srednicki, R. Watkins and K. A. Olive, *Calculations of Relic Densities in the Early Universe*, *Nucl. Phys. B* **310** (1988) 693. [Page 37]

[340] R. J. Scherrer and M. S. Turner, *On the Relic, Cosmic Abundance of Stable Weakly Interacting Massive Particles*, *Phys. Rev. D* **33** (1986) 1585. [Erratum: Phys.Rev.D 34, 3263 (1986)]. [Page 37]

[341] P. Gondolo and G. Gelmini, *Cosmic abundances of stable particles: Improved analysis*, *Nuclear Physics B* **360** (1991) 145. [Page 37]

[342] G. Steigman, B. Dasgupta and J. F. Beacom, *Precise relic wimp abundance and its impact on searches for dark matter annihilation*, *Phys. Rev. D* **86** (2012) 023506. [Page 37]

[343] G. Jungman, M. Kamionkowski and K. Griest, *Supersymmetric dark matter*, *Physics Reports* **267** (1996) 195. [Page 37]

[344] S. Dimopoulos and H. Georgi, *Softly broken supersymmetry and su(5)*, *Nuclear Physics B* **193** (1981) 150. [Page 37]

[345] N. Craig, *The State of Supersymmetry after Run I of the LHC*, in *Beyond the Standard Model after the first run of the LHC*, 9, 2013, 1309.0528. [Page 37]

[346] C. A. J. O'Hare, *New Definition of the Neutrino Floor for Direct Dark Matter Searches*, *Phys. Rev. Lett.* **127** (2021) 251802 [2109.03116]. [Page 37]

[347] F. Mayet et al., *A review of the discovery reach of directional Dark Matter detection*, *Phys. Rept.* **627** (2016) 1 [1602.03781]. [Page 37]

[348] C. A. J. O'Hare, A. M. Green, J. Billard, E. Figueroa-Feliciano and L. E. Strigari, *Readout strategies for directional dark matter detection beyond the neutrino background*, *Phys. Rev. D* **92** (2015) 063518 [1505.08061]. [Page 37]

[349] C. A. J. O'Hare and A. M. Green, *Directional detection of dark matter streams*, *Phys. Rev. D* **90** (2014) 123511 [1410.2749]. [Page 37]

[350] C. Boehm, D. Hooper, J. Silk, M. Casse and J. Paul, *MeV dark matter: Has it been detected?*, *Phys. Rev. Lett.* **92** (2004) 101301 [astro-ph/0309686]. [Page 37]

[351] G. B. Gelmini, *Light weakly interacting massive particles*, *Rept. Prog. Phys.* **80** (2017) 082201 [1612.09137]. [Page 37]

[352] R. K. Leane, T. R. Slatyer, J. F. Beacom and K. C. Y. Ng, *GeV-scale thermal WIMPs: Not even slightly ruled out*, *Phys. Rev. D* **98** (2018) 023016 [1805.10305]. [Page 37]

[353] B. S. Acharya, G. Kane, S. Watson and P. Kumar, *A Non-thermal WIMP Miracle*, *Phys. Rev. D* **80** (2009) 083529 [0908.2430]. [Page 37]

[354] R. D. Peccei and H. R. Quinn, *CP conservation in the presence of instantons*, *Phys. Rev. Lett.* **38** (1977) 1440. [,328(1977)]. [Pages 37, 113, and 118]

[355] R. D. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791. [Pages 37, 113, and 118]

[356] R. D. Peccei, *The Strong CP problem and axions*, *Lect. Notes Phys.* **741** (2008) 3 [hep-ph/0607268]. [Pages 37, 100, and 113]

[357] J. Preskill, M. B. Wise and F. Wilczek, *Cosmology of the Invisible Axion*, *Phys. Lett. B* **120** (1983) 127. [Pages 37 and 119]

[358] L. F. Abbott and P. Sikivie, *A Cosmological Bound on the Invisible Axion*, *Phys. Lett. B* **120** (1983) 133. [Pages 37 and 119]

[359] M. Dine and W. Fischler, *The Not So Harmless Axion*, *Phys. Lett. B* **120** (1983) 137. [Pages 37 and 119]

[360] J. Ipser and P. Sikivie, *Are Galactic Halos Made of Axions?*, *Phys. Rev. Lett.* **50** (1983) 925. [Pages 37 and 119]

[361] L. Hui, *Wave Dark Matter*, *Ann. Rev. Astron. Astrophys.* **59** (2021) 247 [2101.11735]. [Page 37]

[362] E. Witten, *Cosmic Separation of Phases*, *Phys. Rev. D* **30** (1984) 272. [Page 38]

[363] Y. Bai, A. J. Long and S. Lu, *Dark Quark Nuggets*, *Phys. Rev. D* **99** (2019) 055047 [1810.04360]. [Page 38]

[364] Y. B. Zel'dovich and I. D. Novikov, *The hypothesis of cores retarded during expansion and the hot cosmological model*, *Sov. Astron.* **10** (1966) 602. [Pages 38, 40, and 42]

[365] S. Hawking, *Gravitationally collapsed objects of very low mass*, *Mon. Not. Roy. Astron. Soc.* **152** (1971) 75. [Pages 38, 40, 42, and 55]

[366] B. J. Carr and S. Hawking, *Black holes in the early Universe*, *Mon. Not. Roy. Astron. Soc.* **168** (1974) 399. [Pages 38, 40, 42, 55, and 64]

[367] G. F. Chapline, *Cosmological effects of primordial black holes*, *Nature* **253** (1975) 251. [Pages 38, 40, 42, 55, and 57]

[368] D. M. Jacobs, G. D. Starkman and B. W. Lynn, *Macro Dark Matter*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 3418 [1410.2236]. [Page 38]

[369] J. Singh Sidhu, R. J. Scherrer and G. Starkman, *Death and serious injury from dark matter*, *Phys. Lett. B* **803** (2020) 135300 [1907.06674]. [Page 38]

[370] J. W. Moffat, *Modified Gravity or Dark Matter?*, 1101.1935. [Page 38]

[371] J. D. Bekenstein, *Alternatives to Dark Matter: Modified Gravity as an Alternative to dark Matter*, 1001.3876. [Page 38]

[372] N. Rosen, *General relativity and flat space. i*, *Phys. Rev.* **57** (1940) 147. [Page 38]

[373] N. Rosen, *A bi-metric theory of gravitation*, *General Relativity and Gravitation* **4** (1973) 435. [Page 38]

[374] S. F. Hassan and R. A. Rosen, *Bimetric Gravity from Ghost-free Massive Gravity*, *JHEP* **02** (2012) 126 [1109.3515]. [Page 38]

[375] K. Bamba and S. D. Odintsov, *Inflationary cosmology in modified gravity theories*, *Symmetry* **7** (2015) 220 [1503.00442]. [Page 38]

[376] S. Nojiri, S. D. Odintsov and V. K. Oikonomou, *Modified Gravity Theories on a Nutshell: Inflation, Bounce and Late-time Evolution*, *Phys. Rept.* **692** (2017) 1 [1705.11098]. [Page 38]

[377] A. Joyce, L. Lombriser and F. Schmidt, *Dark Energy Versus Modified Gravity*, *Ann. Rev. Nucl. Part. Sci.* **66** (2016) 95 [1601.06133]. [Page 38]

[378] C. Skordis, *Generalizing tensor-vector-scalar cosmology*, *Phys. Rev. D* **77** (2008) 123502 [0801.1985]. [Page 38]

[379] X.-d. Xu, B. Wang and P. Zhang, *Testing the tensor-vector-scalar Theory with the latest cosmological observations*, *Phys. Rev. D* **92** (2015) 083505 [1412.4073]. [Page 38]

[380] F. Schmidt, M. Liguori and S. Dodelson, *Galaxy-CMB Cross-Correlation as a Probe of Alternative Models of Gravity*, *Phys. Rev. D* **76** (2007) 083518 [0706.1775]. [Page 38]

[381] P. Brax and A.-C. Davis, *Modified Gravity and the CMB*, *Phys. Rev. D* **85** (2012) 023513 [1109.5862]. [Page 38]

[382] M. Feix, C. Fedeli and M. Bartelmann, *Asymmetric Gravitational Lenses in TeVeS and Application to the Bullet Cluster*, *Astron. Astrophys.* **480** (2008) 313 [0707.0790]. [Page 38]

[383] J. R. Brownstein and J. W. Moffat, *The Bullet Cluster 1E0657-558 evidence shows Modified Gravity in the absence of Dark Matter*, *Mon. Not. Roy. Astron. Soc.* **382** (2007) 29 [astro-ph/0702146]. [Page 38]

[384] N. S. Israel and J. W. Moffat, *The Train Wreck Cluster Abell 520 and the Bullet Cluster 1E0657-558 in a Generalized Theory of Gravitation*, *arXiv e-prints* (2016) arXiv:1606.09128 [1606.09128]. [Page 38]

[385] J. D. Bekenstein and R. H. Sanders, *TeVeS/MOND is in harmony with gravitational redshifts in galaxy clusters*, *Mon. Not. Roy. Astron. Soc.* **421** (2012) L59 [1110.5048]. [Page 38]

[386] S. W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199. [Erratum: Commun.Math.Phys. 46, 206 (1976)]. [Pages 40, 42, 46, and 47]

[387] S. W. Hawking, *Black hole explosions*, *Nature* **248** (1974) 30. [Pages 40, 42, 46, and 47]

[388] H. Bondi, *On spherically symmetrical accretion*, *Mon. Not. Roy. Astron. Soc.* **112** (1952) 195. [Pages 40, 52, and 59]

[389] E. E. Salpeter, *Accretion of Interstellar Matter by Massive Objects.*, ApJ **140** (1964) 796. [Pages 40, 52, and 59]

[390] Y. B. Zel'dovich, *The Fate of a Star and the Evolution of Gravitational Energy Upon Accretion*, *Soviet Physics Doklady* **9** (1964) 195. [Pages 40, 52, and 64]

[391] Y. B. . N. Zel'dovich, I. D., *The Hypothesis of Cores Retarded during Expansion and the Hot Cosmological Model*, *Soviet Astron. AJ (Engl. Transl.)*, **10** (1967) 602. [Pages 40, 55, and 64]

[392] J. 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** (1784) 35. [Page 41]

[393] *John michell anticipates black holes*, https://www.aps.org/publications/apsnews/200911/physicshistory.cfm. [Page 41]

[394] A. Einstein, *Zur allgemeinen Relativitätstheorie. (German) [Toward a General Theory of Relativity]*, *j.S.B.PREUSS.AKAD.WISS.2* (1915) 778. [Page 41]

[395] A. Einstein, *Die Feldgleichungen der Gravitation. (German) [The field equations of gravitation]*, *j.S.B.PREUSS.AKAD.WISS.2* (1915) 844. [Page 41]

[396] K. Schwarzschild, *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)* (1916) 189. [Pages 41 and 42]

[397] A. S. Eddington, *A Comparison of Whitehead's and Einstein's Formulæ*, Nature **113** (1924) 192. [Page 41]

[398] S. Chandrasekhar, *The Maximum Mass of Ideal White Dwarfs*, ApJ **74** (1931) 81. [Page 41]

[399] J. R. Oppenheimer and G. M. Volkoff, *On massive neutron cores*, *Phys. Rev.* **55** (1939) 374. [Page 41]

[400] R. C. Tolman, *Static solutions of einstein's field equations for spheres of fluid*, *Phys. Rev.* **55** (1939) 364. [Page 41]

[401] D. Finkelstein, *Past-future asymmetry of the gravitational field of a point particle*, *Phys. Rev.* **110** (1958) 965. [Page 41]

[402] M. D. Kruskal, *Maximal extension of schwarzschild metric*, *Phys. Rev.* **119** (1960) 1743. [Page 42]

[403] R. P. Kerr, *Gravitational field of a spinning mass as an example of algebraically special metrics*, *Phys. Rev. Lett.* **11** (1963) 237. [Pages 42 and 52]

[404] R. Penrose, *Gravitational collapse and space-time singularities*, *Phys. Rev. Lett.* **14** (1965) 57. [Pages 42 and 49]

[405] J. D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333. [Pages 42, 46, and 47]

[406] J. M. Bardeen, B. Carter and S. W. Hawking, *The Four laws of black hole mechanics*, *Commun. Math. Phys.* **31** (1973) 161. [Pages 42, 46, and 47]

[407] A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott and R. A. Collins, *Observation of a Rapidly Pulsating Radio Source*, Nature **217** (1968) 709. [Page 42]

[408] S. Bowyer, E. T. Byram, T. A. Chubb and H. Friedman, *Cosmic X-ray Sources*, *Science* **147** (1965) 394. [Page 42]

[409] H. L. Shipman, *The Implausible History of Triple Star Models for Cygnus X-1: Evidence for a Black Hole*, *aplett* **16** (1975) 9. [Page 42]

[410] J. T. Jebsen, *On the general spherically symmetric solutions of Einstein's gravitational equations in vacuo*, *General Relativity and Gravitation* **37** (2005) 2253. [Pages 42 and 72]

[411] G. D. Birkhoff and R. E. Langer, *Relativity and modern physics*. 1923. [Page 42]

[412] R. L. Arnowitt, S. Deser and C. W. Misner, *Dynamical Structure and Definition of Energy in General Relativity*, *Phys. Rev.* **116** (1959) 1322. [Pages 43 and 167]

[413] C. W. Misner and D. H. Sharp, *Relativistic equations for adiabatic, spherically symmetric gravitational collapse*, *Phys. Rev* **136** (1964) B571. [Page 44]

[414] W. C. Hernandez and C. W. Misner, *Observer Time as a Coordinate in Relativistic Spherical Hydrodynamics*, *Astrophys. J.* **143** (1966) 452. [Page 44]

[415] V. Faraoni, *Embedding black holes and other inhomogeneities in the universe in various theories of gravity: a short review*, *Universe* **4** (2018) 109 [1810.04667]. [Pages 44, 46, 65, 69, 70, 71, 72, and 74]

[416] S. A. Hayward, *General laws of black hole dynamics*, *Phys. Rev. D* **49** (1994) 6467. [Pages 45 and 92]

[417] A. B. Nielsen and M. Visser, *Production and decay of evolving horizons*, *Class. Quant. Grav.* **23** (2006) 4637 [gr-qc/0510083]. [Pages 46 and 69]

[418] D. N. Page, *Hawking radiation and black hole thermodynamics*, *New J. Phys.* **7** (2005) 203 [hep-th/0409024]. [Pages 47 and 94]

[419] D. N. Page, *Particle Emission Rates from a Black Hole: Massless Particles from an Uncharged, Nonrotating Hole*, *Phys. Rev. D* **13** (1976) 198. [Pages 48, 55, and 94]

[420] A. Arbey and J. Auffinger, *BlackHawk: A public code for calculating the Hawking evaporation spectra of any black hole distribution*, *Eur. Phys. J. C* **79** (2019) 693 [1905.04268]. [Page 48]

[421] A. Arbey and J. Auffinger, *Physics Beyond the Standard Model with BlackHawk v2.0*, *Eur. Phys. J. C* **81** (2021) 910 [2108.02737]. [Page 48]

[422] M. R. Mosbech and Z. S. C. Picker, *Effects of Hawking evaporation on PBH distributions*, 2203.05743. [Pages 48, 61, 62, and 94]

[423] R. Penrose, *Gravitational collapse: The role of general relativity*, *Riv. Nuovo Cim.* **1** (1969) 252. [Page 49]

[424] R. M. Wald and V. Iyer, *Trapped surfaces in the schwarzschild geometry and cosmic censorship*, *Phys. Rev. D* **44** (1991) R3719. [Page 49]

[425] Y. C. Ong, *Space–time singularities and cosmic censorship conjecture: A Review with some thoughts*, *Int. J. Mod. Phys. A* **35** (2020) 14 [2005.07032]. [Page 49]

[426] LIGO Scientific, Virgo Collaboration, B. P. Abbott et al., *Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo*, *Astrophys. J. Lett.* **882** (2019) L24 [1811.12940]. [Pages 49, 50, and 90]

[427] C. D. Bailyn, R. K. Jain, P. Coppi and J. A. Orosz, *The Mass distribution of stellar black holes*, *Astrophys. J.* **499** (1998) 367 [astro-ph/9708032]. [Page 50]

[428] F. Özel, D. Psaltis, R. Narayan and J. E. McClintock, *The Black Hole Mass Distribution in the Galaxy*, ApJ **725** (2010) 1918 [1006.2834]. [Page 50]

[429] W. M. Farr, N. Sravan, A. Cantrell, L. Kreidberg, C. D. Bailyn, I. Mandel and V. Kalogera, *The Mass Distribution of Stellar-mass Black Holes*, *APJ* **741** (2011) 103 [1011.1459]. [Page 50]

[430] G. Rakavy and G. Shaviv, *Instabilities in Highly Evolved Stellar Models*, ApJ **148** (1967) 803. [Page 50]

[431] R. Farmer, M. Renzo, S. E. de Mink, P. Marchant and S. Justham, *Mind the gap: The location of the lower edge of the pair instability supernovae black hole mass gap*, 1910.12874. [Page 50]

[432] R. M. Cutri, M. F. Skrutskie, S. van Dyk, C. A. Beichman, J. M. Carpenter, T. Chester, L. Cambresy, T. Evans, J. Fowler, J. Gizis, E. Howard, J. Huchra, T. Jarrett, E. L. Kopan, J. D. Kirkpatrick, R. M. Light, K. A. Marsh, H. McCallon, S. Schneider, R. Stiening, M. Sykes, M. Weinberg, W. A. Wheaton, S. Wheelock and N. Zacarias, *VizieR Online Data Catalog: 2MASS All-Sky Catalog of Point Sources (Cutri+ 2003)*, *VizieR Online Data Catalog* (2003) II/246. [Page 50]

[433] S.-W. Wu, A. Bik, J. M. Bestenlehner, T. Henning, A. Pasquali, W. Brandner and A. Stolte, *The massive stellar population of W49: A spectroscopic survey*, A&A **589** (2016) A16 [1602.05190]. [Page 50]

[434] LIGO Scientific, Virgo Collaboration, R. Abbott et al., *GW190521: A Binary Black Hole Merger with a Total Mass of* $150M_\odot$, *Phys. Rev. Lett.* **125** (2020) 101102 [2009.01075]. [Pages 50 and 55]

[435] B. Liu and D. Lai, *Hierarchical Black-Hole Mergers in Multiple Systems: Constrain the Formation of GW190412, GW190814 and GW190521-like events*, *Mon. Not. Roy. Astron. Soc.* **502** (2021) 2049 [2009.10068]. [Page 50]

[436] G. Fragione, A. Loeb and F. A. Rasio, *On the Origin of GW190521-like events from repeated black hole mergers in star clusters*, *Astrophys. J. Lett.* **902** (2020) L26 [2009.05065]. [Page 50]

[437] K. Belczynski, *The most ordinary formation of the most unusual double black hole merger*, *Astrophys. J. Lett.* **905** (2020) L15 [2009.13526]. [Page 50]

[438] G. Franciolini, V. Baibhav, V. De Luca, K. K. Y. Ng, K. W. K. Wong, E. Berti, P. Pani, A. Riotto and S. Vitale, *Quantifying the evidence for primordial black holes in LIGO/Virgo gravitational-wave data*, 2105.03349. [Pages 50 and 55]

[439] V. De Luca, V. Desjacques, G. Franciolini, P. Pani and A. Riotto, *GW190521 Mass Gap Event and the Primordial Black Hole Scenario*, *Phys. Rev. Lett.* **126** (2021) 051101 [2009.01728]. [Pages 50 and 55]

[440] K. Kritos, V. De Luca, G. Franciolini, A. Kehagias and A. Riotto, *The Astro-Primordial Black Hole Merger Rates: a Reappraisal*, *JCAP* **05** (2021) 039 [2012.03585]. [Pages 50 and 55]

[441] LIGO Scientific, Virgo Collaboration, B. P. Abbott et al., *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, *Phys. Rev. Lett.* **119** (2017) 161101 [1710.05832]. [Page 51]

[442] LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South

Africa/MeerKAT Collaboration, B. P. Abbott et al., *Multi-messenger Observations of a Binary Neutron Star Merger*, *Astrophys. J. Lett.* **848** (2017) L12 [1710.05833]. [Page 51]

[443] Event Horizon Telescope Collaboration, K. Akiyama et al., *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, *Astrophys. J. Lett.* **875** (2019) L1 [1906.11238]. [Page 51]

[444] Event Horizon Telescope Collaboration, K. Akiyama et al., *First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way*, *Astrophys. J. Lett.* **930** (2022) L12. [Page 51]

[445] OGLE, MOA, PLANET, microFUN, MiNDSTEp Consortium, RoboNet Collaboration, K. C. Sahu et al., *An Isolated Stellar-Mass Black Hole Detected Through Astrometric Microlensing*, 2201.13296. [Page 52]

[446] B. Balick and R. L. Brown, *Intense sub-arcsecond structure in the galactic center.*, ApJ **194** (1974) 265. [Page 52]

[447] R. Genzel, D. Hollenbach and C. H. Townes, *The nucleus of our Galaxy*, *Reports on Progress in Physics* **57** (1994) 417. [Page 52]

[448] W. M. Goss and R. X. McGee, *The Discovery of the Radio Source Sagittarius A (Sgr A)*, in *The Galactic Center* (R. Gredel, ed.), vol. 102 of *Astronomical Society of the Pacific Conference Series*, p. 369, Jan., 1996. [Page 52]

[449] R. Schodel et al., *A Star in a 15.2 year orbit around the supermassive black hole at the center of the Milky Way*, *Nature* **419** (2002) 694 [astro-ph/0210426]. [Page 52]

[450] S. Gillessen, F. Eisenhauer, S. Trippe, T. Alexander, R. Genzel, F. Martins and T. Ott, *Monitoring stellar orbits around the Massive Black Hole in the Galactic Center*, *Astrophys. J.* **692** (2009) 1075 [0810.4674]. [Page 52]

[451] T. A. Matthews and A. R. Sandage, *Optical Identification of 3C 48, 3C 196, and 3C 286 with Stellar Objects.*, ApJ **138** (1963) 30. [Page 52]

[452] M. Schmidt, *3C 273 : A Star-Like Object with Large Red-Shift*, Nature **197** (1963) 1040. [Page 52]

[453] D. Lynden-Bell, *Galactic Nuclei as Collapsed Old Quasars*, Nature **223** (1969) 690. [Page 52]

[454] L. Ferrarese and H. Ford, *Supermassive Black Holes in Galactic Nuclei: Past, Present and Future Research*, Space Sci. Rev. **116** (2005) 523 [astro-ph/0411247]. [Page 52]

[455] E. L. Schucking, *The first texas symposium on relativistic astrophysics*, *Physics Today* **42** (1989) 46. [Page 52]

[456] R. Giacconi, H. Gursky, F. R. Paolini and B. B. Rossi, *Evidence for x Rays From Sources Outside the Solar System*, Phys. Rev. Lett. **9** (1962) 439. [Page 52]

[457] S. E. Motta et al., *The INTEGRAL view on black hole X-ray binaries*, *New Astron. Rev.* **93** (2021) 101618 [2105.05547]. [Page 52]

[458] B. Carr, S. Clesse, J. Garcia-Bellido and F. Kuhnel, *Cosmic Conundra Explained by Thermal History and Primordial Black Holes*, 1906.08217. [Pages 55, 56, and 62]

[459] R. H. Cyburt, B. D. Fields and K. A. Olive, *Primordial nucleosynthesis in light of WMAP*, *Phys. Lett. B* **567** (2003) 227 [astro-ph/0302431]. [Page 55]

[460] R. Lehoucq, M. Cassé , J.-M. Casandjian and I. Grenier, *New constraints on the primordial black hole number density from galactic γ-ray astronomy*, *Astronomy and Astrophysics* **502** (2009) 37. [Page 55]

[461] A. Barrau, *Primordial black holes as a source of extremely high-energy cosmic rays*, *Astropart. Phys.* **12** (2000) 269 [astro-ph/9907347]. [Page 55]

[462] C. Bambi, A. D. Dolgov and A. A. Petrov, *Primordial black holes and the observed Galactic 511-keV line*, *Phys. Lett. B* **670** (2008) 174 [0801.2786]. [Erratum: Phys.Lett.B 681, 504–504 (2009)]. [Page 55]

[463] K. M. Belotsky and A. A. Kirillov, *Primordial black holes with mass $10^{16} - 10^{17}$ g and reionization of the Universe*, *JCAP* **01** (2015) 041 [1409.8601]. [Page 55]

[464] D. B. Cline, D. A. Sanders and W. Hong, *Further evidence for gamma-ray bursts consistent with primordial black hole evaporation*, *Astrophys. J.* **486** (1997) 169. [Page 55]

[465] P. Meszaros, *Primeval black holes and galaxy formation*, *Astron. Astrophys.* **38** (1975) 5. [Pages 55 and 56]

[466] B. J. Carr and M. J. Rees, *Can pregalactic objects generate galaxies?*, MNRAS **206** (1984) 801. [Page 55]

[467] R. Bean and J. Magueijo, *Could supermassive black holes be quintessential primordial black holes?*, *Phys. Rev. D* **66** (2002) 063505 [astro-ph/0204486]. [Page 55]

[468] N. Afshordi, P. McDonald and D. N. Spergel, *Primordial black holes as dark matter: The Power spectrum and evaporation of early structures*, *Astrophys. J. Lett.* **594** (2003) L71 [astro-ph/0302035]. [Pages 55 and 59]

[469] A. M. Green and B. J. Kavanagh, *Primordial Black Holes as a dark matter candidate*, 2007.10722. [Pages 55 and 56]

[470] B. Carr, M. Raidal, T. Tenkanen, V. Vaskonen and H. Veermäe, *Primordial black hole constraints for extended mass functions*, *Phys. Rev. D* **96** (2017) 023514 [1705.05567]. [Pages 55 and 62]

[471] F. Kühnel and K. Freese, *Constraints on Primordial Black Holes with Extended Mass Functions*, *Phys. Rev. D* **95** (2017) 083508 [1701.07223]. [Pages 55 and 62]

[472] N. Bellomo, J. L. Bernal, A. Raccanelli and L. Verde, *Primordial Black Holes as Dark Matter: Converting Constraints from Monochromatic to Extended Mass Distributions*, *JCAP* **01** (2018) 004 [1709.07467]. [Pages 55 and 62]

[473] A. Arbey, J. Auffinger and J. Silk, *Constraining primordial black hole masses with the isotropic gamma ray background*, *Phys. Rev. D* **101** (2020) 023010 [1906.04750]. [Pages 55 and 62]

[474] M. Y. Khlopov, *Primordial Black Holes*, *Res. Astron. Astrophys.* **10** (2010) 495 [0801.0116]. [Page 56]

[475] G. Franciolini, *Primordial Black Holes: from Theory to Gravitational Wave Observations*, Ph.D. thesis, Geneva U., Dept. Theor. Phys., 2021. 2110.06815. 10.13097/archive-ouverte/unige:156136. [Page 56]

[476] T. W. B. Kibble, *Topology of Cosmic Domains and Strings*, *J. Phys. A* **9** (1976) 1387. [Pages 56, 120, and 149]

[477] A. Vilenkin, *Gravitational Field of Vacuum Domain Walls and Strings*, *Phys. Rev. D* **23** (1981) 852. [Pages 56, 120, and 150]

[478] S. W. Hawking, *Black Holes From Cosmic Strings*, *Phys. Lett. B* **231** (1989) 237. [Page 56]

[479] A. Polnarev and R. Zembowicz, *Formation of Primordial Black Holes by Cosmic Strings*, *Phys. Rev. D* **43** (1991) 1106. [Page 56]

[480] J. Garriga and M. Sakellariadou, *Effects of friction on cosmic strings*, *Phys. Rev. D* **48** (1993) 2502 [hep-th/9303024]. [Page 56]

[481] R. R. Caldwell and P. Casper, *Formation of black holes from collapsed cosmic string loops*, *Phys. Rev. D* **53** (1996) 3002 [gr-qc/9509012]. [Page 56]

[482] J. H. MacGibbon, R. H. Brandenberger and U. F. Wichoski, *Limits on black hole formation from cosmic string loops*, *Phys. Rev. D* **57** (1998) 2158 [astro-ph/9707146]. [Page 56]

[483] A. C. Jenkins and M. Sakellariadou, *Primordial black holes from cusp collapse on cosmic strings*, 2006.16249. [Page 56]

[484] S. G. Rubin, M. Y. Khlopov and A. S. Sakharov, *Primordial black holes from nonequilibrium second order phase transition*, *Grav. Cosmol.* **6** (2000) 51 [hep-ph/0005271]. [Pages 56 and 151]

[485] H. Deng and A. Vilenkin, *Primordial black hole formation by vacuum bubbles*, *JCAP* **12** (2017) 044 [1710.02865]. [Pages 56 and 151]

[486] T. Vachaspati, *Lunar Mass Black Holes from QCD Axion Cosmology*, 1706.03868. [Pages 56, 151, and 153]

[487] F. Ferrer, E. Masso, G. Panico, O. Pujolas and F. Rompineve, *Primordial Black Holes from the QCD axion*, *Phys. Rev. Lett.* **122** (2019) 101301 [1807.01707]. [Pages 56, 151, and 153]

[488] J. Liu, Z.-K. Guo and R.-G. Cai, *Primordial Black Holes from Cosmic Domain Walls*, *Phys. Rev. D* **101** (2020) 023513 [1908.02662]. [Pages 56 and 151]

[489] S. W. Hawking, I. G. Moss and J. M. Stewart, *Bubble Collisions in the Very Early Universe*, *Phys. Rev. D* **26** (1982) 2681. [Page 56]

[490] I. Musco and J. C. Miller, *Primordial black hole formation in the early universe: critical behaviour and self-similarity*, *Class. Quant. Grav.* **30** (2013) 145009 [1201.2379]. [Page 56]

[491] A. Kusenko, M. Sasaki, S. Sugiyama, M. Takada, V. Takhistov and E. Vitagliano, *Exploring Primordial Black Holes from the Multiverse with Optical Telescopes*, *Phys. Rev. Lett.* **125** (2020) 181304 [2001.09160]. [Page 56]

[492] M. Crawford and D. N. Schramm, *Spontaneous Generation of Density Perturbations in the Early Universe*, *Nature* **298** (1982) 538. [Page 56]

[493] H. Kodama, M. Sasaki and K. Sato, *Abundance of Primordial Holes Produced by Cosmological First Order Phase Transition*, *Prog. Theor. Phys.* **68** (1982) 1979. [Page 56]

[494] K. Jedamzik, *Primordial black hole formation during the QCD epoch*, *Phys. Rev. D* **55** (1997) 5871 [astro-ph/9605152]. [Page 56]

[495] M. Y. Khlopov and A. G. Polnarev, *PRIMORDIAL BLACK HOLES AS A COSMOLOGICAL TEST OF GRAND UNIFICATION*, *Phys. Lett. B* **97** (1980) 383. [Page 56]

[496] A. G. Polnarev and M. Y. Khlopov, *COSMOLOGY, PRIMORDIAL BLACK HOLES, AND SUPERMASSIVE PARTICLES*, *Sov. Phys. Usp.* **28** (1985) 213. [Page 56]

[497] A. M. Green, A. R. Liddle and A. Riotto, *Primordial black hole constraints in cosmologies with early matter domination*, *Phys. Rev. D* **56** (1997) 7559 [astro-ph/9705166]. [Page 56]

[498] T. Harada, C.-M. Yoo, K. Kohri, K.-i. Nakao and S. Jhingan, *Primordial black hole formation in the matter-dominated phase of the Universe*, *Astrophys. J.* **833** (2016) 61 [1609.01588]. [Page 56]

[499] B. Carr, T. Tenkanen and V. Vaskonen, *Primordial black holes from inflaton and spectator field perturbations in a matter-dominated era*, *Phys. Rev. D* **96** (2017) 063507 [1706.03746]. [Page 56]

[500] M. Khlopov, B. A. Malomed and I. B. Zeldovich, *Gravitational instability of scalar fields and formation of primordial black holes*, *Mon. Not. Roy. Astron. Soc.* **215** (1985) 575. [Page 56]

[501] E. Cotner and A. Kusenko, *Primordial black holes from supersymmetry in the early universe*, *Phys. Rev. Lett.* **119** (2017) 031103 [1612.02529]. [Page 56]

[502] E. Cotner and A. Kusenko, *Primordial black holes from scalar field evolution in the early universe*, *Phys. Rev. D* **96** (2017) 103002 [1706.09003]. [Page 56]

[503] E. Cotner, A. Kusenko and V. Takhistov, *Primordial Black Holes from Inflaton Fragmentation into Oscillons*, *Phys. Rev. D* **98** (2018) 083513 [1801.03321]. [Page 56]

[504] E. Cotner, A. Kusenko, M. Sasaki and V. Takhistov, *Analytic Description of Primordial Black Hole Formation from Scalar Field Fragmentation*, *JCAP* **10** (2019) 077 [1907.10613]. [Page 56]

[505] M. M. Flores and A. Kusenko, *Primordial Black Holes from Long-Range Scalar Forces and Scalar Radiative Cooling*, *Phys. Rev. Lett.* **126** (2021) 041101 [2008.12456]. [Page 56]

[506] M. M. Flores and A. Kusenko, *Primordial black holes as a dark matter candidate in theories with supersymmetry and inflation*, 2108.08416. [Page 56]

[507] D. K. Nadezhin, I. D. Novikov and A. G. Polnarev, *The hydrodynamics of primordial black hole formation*, Soviet Ast. **22** (1978) 129. [Page 56]

[508] G. V. Bicknell and R. N. Henriksen, *Formation of primordial black holes.*, ApJ **232** (1979) 670. [Page 56]

[509] A. Dolgov and J. Silk, *Baryon isocurvature fluctuations at small scales and baryonic dark matter*, *Phys. Rev. D* **47** (1993) 4244. [Page 56]

[510] B. J. Carr and J. E. Lidsey, *Primordial black holes and generalized constraints on chaotic inflation*, *Phys. Rev. D* **48** (1993) 543. [Page 56]

[511] J. Yokoyama, *Formation of MACHO primordial black holes in inflationary cosmology*, *Astron. Astrophys.* **318** (1997) 673 [astro-ph/9509027]. [Page 56]

[512] B. J. Carr, *The primordial black hole mass spectrum.*, ApJ **201** (1975) 1. [Page 56]

[513] T. Harada, C.-M. Yoo and K. Kohri, *Threshold of primordial black hole formation*, *Phys. Rev. D* **88** (2013) 084051 [1309.4201]. [Erratum: Phys.Rev.D 89, 029903 (2014)]. [Page 56]

[514] I. Musco, *Threshold for primordial black holes: Dependence on the shape of the cosmological perturbations*, *Phys. Rev. D* **100** (2019) 123524 [1809.02127]. [Page 56]

[515] B. J. Kavanagh, *bradkav/pbhbounds: Release version*, Nov., 2019. 10.5281/zenodo.3538999. [Pages 58 and 97]

[516] B. Carr, F. Kuhnel and L. Visinelli, *Constraints on Stupendously Large Black Holes*, *Mon. Not. Roy. Astron. Soc.* **501** (2021) 2029 [2008.08077]. [Page 58]

[517] G. Xu and J. P. Ostriker, *Dynamics of Massive Black Holes as a Possible Candidate of Galactic Dark Matter*, ApJ **437** (1994) 184. [Page 58]

[518] B. J. Carr and M. Sakellariadou, *Dynamical constraints on dark compact objects*, *Astrophys. J.* **516** (1999) 195. [Page 58]

[519] M. A. Monroy-Rodríguez and C. Allen, *The end of the MACHO era- revisited: new limits on MACHO masses from halo wide binaries*, *Astrophys. J.* **790** (2014) 159 [1406.5169]. [Page 58]

[520] T. D. Brandt, *Constraints on MACHO Dark Matter from Compact Stellar Systems in Ultra-Faint Dwarf Galaxies*, *Astrophys. J. Lett.* **824** (2016) L31 [1605.03665]. [Pages 58 and 59]

[521] S. M. Koushiappas and A. Loeb, *Dynamics of Dwarf Galaxies Disfavor Stellar-Mass Black Holes as Dark Matter*, *Phys. Rev. Lett.* **119** (2017) 041102 [1704.01668]. [Pages 58 and 59]

[522] P. Lu, V. Takhistov, G. B. Gelmini, K. Hayashi, Y. Inoue and A. Kusenko, *Constraining Primordial Black Holes with Dwarf Galaxy Heating*, 2007.02213. [Page 58]

[523] J. Stegmann, P. R. Capelo, E. Bortolas and L. Mayer, *Improved constraints from ultra-faint dwarf galaxies on primordial black holes as dark matter*, *Mon. Not. Roy. Astron. Soc.* **492** (2020) 5247 [1910.04793]. [Pages 58 and 59]

[524] A. M. Green, *Microlensing and dynamical constraints on primordial black hole dark matter with an extended mass function*, *Phys. Rev. D* **94** (2016) 063530 [1609.01143]. [Pages 58 and 62]

[525] DES Collaboration, T. Li et al., *Farthest Neighbor: The Distant Milky Way Satellite Eridanus II*, *Astrophys. J.* **838** (2017) 8 [1611.05052]. [Page 58]

[526] B. Carr and J. Silk, *Primordial Black Holes as Generators of Cosmic Structures*, *Mon. Not. Roy. Astron. Soc.* **478** (2018) 3756 [1801.00672]. [Page 59]

[527] R. Murgia, G. Scelfo, M. Viel and A. Raccanelli, *Lyman-α Forest Constraints on Primordial Black Holes as Dark Matter*, *Phys. Rev. Lett.* **123** (2019) 071102 [1903.10509]. [Page 59]

[528] B. J. Carr, *The contribution of accreting black holes to the background radiation density.*, MNRAS **189** (1979) 123. [Page 59]

[529] B. J. Carr, *Pregalactic black hole accretion and the thermal history of the universe*, MNRAS **194** (1981) 639. [Page 59]

[530] A. Soltan, *Masses of quasars.*, MNRAS **200** (1982) 115. [Page 59]

[531] E. Bertschinger, *Self-similar secondary infall and accretion in an Einstein-de Sitter universe*, ApJS **58** (1985) 39. [Page 59]

[532] M. Ricotti, *Bondi accretion in the early universe*, *Astrophys. J.* **662** (2007) 53 [0706.0864]. [Page 59]

[533] J. Manshanden, D. Gaggero, G. Bertone, R. M. Connors and M. Ricotti, *Multi-wavelength astronomical searches for primordial black holes*, *JCAP* **06** (2019) 026 [1812.07967]. [Page 59]

[534] D. Gaggero, G. Bertone, F. Calore, R. M. T. Connors, M. Lovell, S. Markoff and E. Storm, *Searching for Primordial Black Holes in the radio and X-ray sky*, *Phys. Rev. Lett.* **118** (2017) 241101 [1612.00457]. [Page 59]

[535] Y. Inoue and A. Kusenko, *New X-ray bound on density of primordial black holes*, *JCAP* **10** (2017) 034 [1705.00791]. [Page 59]

[536] M. Ricotti, J. P. Ostriker and K. J. Mack, *Effect of Primordial Black Holes on the Cosmic Microwave Background and Cosmological Parameter Estimates*, *Astrophys. J.* **680** (2008) 829 [0709.0524]. [Page 59]

[537] A. Hektor, G. Hütsi, L. Marzola, M. Raidal, V. Vaskonen and H. Veermäe, *Constraining Primordial Black Holes with the EDGES 21-cm Absorption Signal*, *Phys. Rev. D* **98** (2018) 023503 [1803.09697]. [Page 59]

[538] V. Poulin, P. D. Serpico, F. Calore, S. Clesse and K. Kohri, *CMB bounds on disk-accreting massive primordial black holes*, *Phys. Rev. D* **96** (2017) 083524 [1707.04206]. [Pages 59, 62, and 97]

[539] B. Horowitz, *Revisiting Primordial Black Holes Constraints from Ionization History*, 1612.07264. [Page 59]

[540] P. D. Serpico, V. Poulin, D. Inman and K. Kohri, *Cosmic microwave background bounds on primordial black holes including dark matter halo accretion*, *Phys. Rev. Res.* **2** (2020) 023204 [2002.10771]. [Page 59]

[541] Y. Ali-Haïmoud and M. Kamionkowski, *Cosmic microwave background limits on accreting primordial black holes*, *Phys. Rev. D* **95** (2017) 043534 [1612.05644]. [Page 59]

[542] V. Bosch-Ramon and N. Bellomo, *Mechanical feedback effects on primordial black hole accretion*, *Astron. Astrophys.* **638** (2020) A132 [2004.11224]. [Page 59]

[543] D. P. Quinn, M. I. Wilkinson, M. J. Irwin, J. Marshall, A. Koch and V. Belokurov, *On the reported death of the MACHO era*, *Mon. Not. Roy. Astron. Soc.* **396** (2009) L11 [0903.1644]. [Page 59]

[544] Q. Zhu, E. Vasiliev, Y. Li and Y. Jing, *Primordial Black Holes as Dark Matter: Constraints From Compact Ultra-Faint Dwarfs*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 2 [1710.05032]. [Page 59]

[545] J. N. Bahcall, P. Hut and S. Tremaine, *Maximum mass of objects that constitute unseen disk material*, ApJ **290** (1985) 15. [Page 59]

[546] J. Yoo, J. Chaname and A. Gould, *The end of the MACHO era: limits on halo dark matter from stellar halo wide binaries*, *Astrophys. J.* **601** (2004) 311 [astro-ph/0307437]. [Page 59]

[547] D. P. Quinn, M. I. Wilkinson, M. J. Irwin, J. Marshall, A. Koch and V. Belokurov, *On the reported death of the MACHO era*, MNRAS **396** (2009) L11 [0903.1644]. [Page 59]

[548] C. Allen and M. A. Monroy-Rodríguez, *An Improved Catalog of Halo Wide Binary Candidates*, ApJ **790** (2014) 158 [1406.5164]. [Page 59]

[549] A. M. Price-Whelan, S. Oh and D. N. Spergel, *Spectroscopic confirmation of very-wide stellar binaries and large-separation comoving pairs from Gaia DR1*, *arXiv e-prints* (2017) arXiv:1709.03532 [1709.03532]. [Page 59]

[550] T. Nakamura, M. Sasaki, T. Tanaka and K. S. Thorne, *Gravitational waves from coalescing black hole MACHO binaries*, *Astrophys. J. Lett.* **487** (1997) L139 [astro-ph/9708060]. [Pages 59, 85, 86, 88, and 89]

[551] K. Ioka, T. Chiba, T. Tanaka and T. Nakamura, *Black hole binary formation in the expanding universe: Three body problem approximation*, *Phys. Rev. D* **58** (1998) 063003 [astro-ph/9807018]. [Pages 59, 85, and 89]

[552] M. Sasaki, T. Suyama, T. Tanaka and S. Yokoyama, *Primordial Black Hole Scenario for the Gravitational-Wave Event GW150914*, *Phys. Rev. Lett.* **117** (2016) 061101 [1603.08338]. [Erratum: Phys.Rev.Lett. 121, 059901 (2018)]. [Pages 59, 85, and 88]

[553] Y. Ali-Haïmoud, E. D. Kovetz and M. Kamionkowski, *Merger rate of primordial black-hole binaries*, *Phys. Rev. D* **96** (2017) 123523 [1709.06576]. [Pages 59, 85, 86, and 90]

[554] B. Kocsis, T. Suyama, T. Tanaka and S. Yokoyama, *Hidden universality in the merger rate distribution in the primordial black hole scenario*, *Astrophys. J.* **854** (2018) 41 [1709.09007]. [Pages 59 and 85]

[555] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli and A. G. Riess, *Did LIGO detect dark matter?*, *Phys. Rev. Lett.* **116** (2016) 201301 [1603.00464]. [Pages 59 and 90]

[556] S. Wang, Y.-F. Wang, Q.-G. Huang and T. G. F. Li, *Constraints on the Primordial Black Hole Abundance from the First Advanced LIGO Observation Run Using the Stochastic Gravitational-Wave Background*, *Phys. Rev. Lett.* **120** (2018) 191102 [1610.08725]. [Pages 59, 85, and 90]

[557] M. Raidal, V. Vaskonen and H. Veermäe, *Gravitational Waves from Primordial Black Hole Mergers*, *JCAP* **09** (2017) 037 [1707.01480]. [Pages 59, 85, and 90]

[558] V. Vaskonen and H. Veermäe, *Lower bound on the primordial black hole merger rate*, *Phys. Rev. D* **101** (2020) 043015 [1908.09752]. [Pages 59 and 85]

[559] M. Raidal, C. Spethmann, V. Vaskonen and H. Veermäe, *Formation and Evolution of Primordial Black Hole Binaries in the Early Universe*, *JCAP* **02** (2019) 018 [1812.01930]. [Pages 59 and 85]

[560] K. Jedamzik, *Primordial Black Hole Dark Matter and the LIGO/Virgo observations*, 2006.11172. [Pages 59 and 85]

[561] K. Jedamzik, *Evidence for primordial black hole dark matter from LIGO/Virgo merger rates*, 2007.03565. [Pages 59 and 85]

[562] V. Atal, A. Sanglas and N. Triantafyllou, *LIGO/Virgo black holes and dark matter: The effect of spatial clustering*, *JCAP* **11** (2020) 036 [2007.07212]. [Pages 59 and 85]

[563] B. Paczynski, *Gravitational Microlensing by the Galactic Halo*, ApJ **304** (1986) 1. [Page 60]

[564] Macho Collaboration, R. Allsman et al., *MACHO project limits on black hole dark matter in the 1-30 solar mass range*, *Astrophys. J. Lett.* **550** (2001) L169 [astro-ph/0011506]. [Page 60]

[565] EROS-2 Collaboration, P. Tisserand et al., *Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds*, *Astron. Astrophys.* **469** (2007) 387 [astro-ph/0607207]. [Page 60]

[566] M. Oguri, J. M. Diego, N. Kaiser, P. L. Kelly and T. Broadhurst, *Understanding caustic crossings in giant arcs: Characteristic scales, event rates, and constraints on compact dark matter*, *Phys. Rev. D* **97** (2018) 023518 [1710.00148]. [Page 60]

[567] Macho Collaboration, R. A. Allsman et al., *MACHO project limits on black hole dark matter in the 1-30 solar mass range*, *Astrophys. J. Lett.* **550** (2001) L169 [astro-ph/0011506]. [Page 60]

[568] H. Niikura et al., *Microlensing constraints on primordial black holes with Subaru/HSC Andromeda observations*, *Nature Astron.* **3** (2019) 524 [1701.02151]. [Page 60]

[569] EROS-2 Collaboration, P. Tisserand et al., *Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds*, *Astron. Astrophys.* **469** (2007) 387 [astro-ph/0607207]. [Page 60]

[570] H. Niikura, M. Takada, S. Yokoyama, T. Sumi and S. Masaki, *Constraints on Earth-mass primordial black holes from OGLE 5-year microlensing events*, *Phys. Rev. D* **99** (2019) 083503 [1901.07120]. [Page 60]

[571] P. L. Kelly et al., *Extreme magnification of an individual star at redshift 1.5 by a galaxy-cluster lens*, *Nature Astron.* **2** (2018) 334 [1706.10279]. [Page 60]

[572] J. M. Diego et al., *Dark Matter under the Microscope: Constraining Compact Dark Matter with Caustic Crossing Events*, *Astrophys. J.* **857** (2018) 25 [1706.10281]. [Page 60]

[573] R. B. Metcalf and J. Silk, *New Constraints on Macroscopic Compact Objects as a Dark Matter Candidate from Gravitational Lensing of Type Ia Supernovae*, *Phys. Rev. Lett.* **98** (2007) 071302 [astro-ph/0612253]. [Page 60]

[574] J. Garcia-Bellido, S. Clesse and P. Fleury, *Primordial black holes survive SN lensing constraints*, *Phys. Dark Univ.* **20** (2018) 95 [1712.06574]. [Page 60]

[575] P. L. Schechter, *The Contribution of Faint, Failed, and Defunct Stars to the "Stellar" Masses of Galaxies*, in *Stellar Populations and the Distance Scale* (J. Jensen, R. M. Rich and R. de Grijs, eds.), vol. 514 of *Astronomical Society of the Pacific Conference Series*, p. 79, Nov., 2018, 1712.00393. [Page 60]

[576] E. Mediavilla, J. A. Muñoz, E. Falco, V. Motta, E. Guerras, H. Canovas, C. Jean, A. Oscoz and A. M. Mosquera, *Microlensing-based Estimate of the Mass Fraction in Compact Objects in Lens Galaxies*, *Astrophys. J.* **706** (2009) 1451 [0910.3645]. [Page 60]

[577] E. Mediavilla, J. Jiménez-Vicente, J. Muñoz, H. Vives-Arias and J. Calderón-Infante, *Limits on the Mass and Abundance of Primordial Black Holes from Quasar Gravitational Microlensing*, *Astrophys. J. Lett.* **836** (2017) L18 [1702.00947]. [Page 60]

[578] J. B. Muñoz, E. D. Kovetz, L. Dai and M. Kamionkowski, *Lensing of Fast Radio Bursts as a Probe of Compact Dark Matter*, *Phys. Rev. Lett.* **117** (2016) 091301 [1605.00008]. [Page 60]

[579] K. Liao, S.-B. Zhang, Z. Li and H. Gao, *Constraints on compact dark matter with fast radio burst observations*, *Astrophys. J. Lett.* **896** (2020) L11 [2003.13349]. [Page 60]

[580] M. Zumalacarregui and U. Seljak, *Limits on stellar-mass compact objects as dark matter from gravitational lensing of type Ia supernovae*, *Phys. Rev. Lett.* **121** (2018) 141101 [1712.02240]. [Page 60]

[581] T. Blaineau et al., *New limits from microlensing on Galactic Black Holes in the mass range 10xMsun < M < 1000xMsun*, 2202.13819. [Page 60]

[582] M. R. S. Hawkins, *New evidence for a cosmological distribution of stellar mass primordial black holes*, *Mon. Not. Roy. Astron. Soc.* **512** (2022) 5706 [2204.09143]. [Page 60]

[583] M. R. S. Hawkins, *A new look at microlensing limits on dark matter in the Galactic halo*, A&A **575** (2015) A107 [1503.01935]. [Page 60]

[584] A. M. Green, *Astrophysical uncertainties on stellar microlensing constraints on multi-Solar mass primordial black hole dark matter*, *Phys. Rev. D* **96** (2017) 043020 [1705.10818]. [Page 60]

[585] J. Calcino, J. Garcia-Bellido and T. M. Davis, *Updating the MACHO fraction of the Milky Way dark halowith improved mass models*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 2889 [1803.09205]. [Pages 60 and 62]

[586] Y. Kim and A. Kobakhidze, *Topologically induced black hole charge and its astrophysical manifestations*, 2008.04506. [Page 60]

[587] Y. Kim and A. Kobakhidze, *Non-Schwarzschild Primordial Black Holes as Dark Matter in Quadratic Gravity*, 2106.11764. [Page 60]

[588] P. Montero-Camacho, X. Fang, G. Vasquez, M. Silva and C. M. Hirata, *Revisiting constraints on asteroid-mass primordial black holes as dark matter candidates*, *Journal of Cosmology and Astroparticle Physics* **2019** (2019) 031. [Pages 60 and 61]

[589] A. Gould, *Femtolensing of Gamma-Ray Bursters*, ApJ **386** (1992) L5. [Page 61]

[590] A. Barnacka, J. F. Glicenstein and R. Moderski, *New constraints on primordial black holes abundance from femtolensing of gamma-ray bursts*, Phys. Rev. D **86** (2012) 043001 [1204.2056]. [Page 61]

[591] A. Katz, J. Kopp, S. Sibiryakov and W. Xue, *Femtolensing by Dark Matter Revisited*, *JCAP* **12** (2018) 005 [1807.11495]. [Page 61]

[592] A. Ulmer and J. Goodman, *Femtolensing: Beyond the semiclassical approximation*, *Astrophys. J.* **442** (1995) 67 [astro-ph/9406042]. [Page 61]

[593] F. Capela, M. Pshirkov and P. Tinyakov, *Constraints on Primordial Black Holes as Dark Matter Candidates from Star Formation*, *Phys. Rev. D* **87** (2013) 023507 [1209.6021]. [Page 61]

[594] F. Capela, M. Pshirkov and P. Tinyakov, *Constraints on primordial black holes as dark matter candidates from capture by neutron stars*, *Phys. Rev. D* **87** (2013) 123524 [1301.4984]. [Page 61]

[595] P. Pani and A. Loeb, *Tidal capture of a primordial black hole by a neutron star: implications for constraints on dark matter*, *JCAP* **06** (2014) 026 [1401.3025]. [Page 61]

[596] F. Capela, M. Pshirkov and P. Tinyakov, *Adiabatic contraction revisited: implications for primordial black holes*, *Phys. Rev. D* **90** (2014) 083507 [1403.7098]. [Page 61]

[597] P. W. Graham, S. Rajendran and J. Varela, *Dark Matter Triggers of Supernovae*, *Phys. Rev. D* **92** (2015) 063007 [1505.04444]. [Page 61]

[598] W. DeRocco and P. W. Graham, *Constraining Primordial Black Hole Abundance with the Galactic 511 keV Line*, *Phys. Rev. Lett.* **123** (2019) 251102 [1906.07740]. [Pages 61 and 62]

[599] R. Laha, *Primordial Black Holes as a Dark Matter Candidate Are Severely Constrained by the Galactic Center 511 keV γ -Ray Line*, *Phys. Rev. Lett.* **123** (2019) 251101 [1906.09994]. [Pages 61 and 62]

[600] M. Boudaud and M. Cirelli, *Voyager 1 $e^{\pm}$ Further Constrain Primordial Black Holes as Dark Matter*, *Phys. Rev. Lett.* **122** (2019) 041104 [1807.03075]. [Pages 61 and 62]

[601] B. J. Carr, K. Kohri, Y. Sendouda and J. Yokoyama, *New cosmological constraints on primordial black holes*, *Phys. Rev. D* **81** (2010) 104019 [0912.5297]. [Pages 61, 62, and 97]

[602] S. Clark, B. Dutta, Y. Gao, L. E. Strigari and S. Watson, *Planck Constraint on Relic Primordial Black Holes*, *Phys. Rev. D* **95** (2017) 083006 [1612.07738]. [Page 61]

[603] R. Laha, J. B. Muñoz and T. R. Slatyer, *INTEGRAL constraints on primordial black holes and particle dark matter*, *Phys. Rev. D* **101** (2020) 123514 [2004.00627]. [Page 61]

[604] V. Poulin, J. Lesgourgues and P. D. Serpico, *Cosmological constraints on exotic injection of electromagnetic energy*, *JCAP* **03** (2017) 043 [1610.10051]. [Pages 62 and 97]

[605] S. Clark, B. Dutta, Y. Gao, Y.-Z. Ma and L. E. Strigari, *21 cm limits on decaying dark matter and primordial black holes*, *Phys. Rev. D* **98** (2018) 043006 [1803.09390]. [Page 62]

[606] K. Kannike, L. Marzola, M. Raidal and H. Veermäe, *Single Field Double Inflation and Primordial Black Holes*, *JCAP* **09** (2017) 020 [1705.06225]. [Page 62]

[607] G. C. McVittie, *The mass-particle in an expanding universe*, *Mon* **93** (1933) 325. [Pages 65 and 71]

[608] A. Einstein and E. G. Straus, *The influence of the expansion of space on the gravitation fields surrounding the individual stars*, *Rev. Mod. Phys.* **17** (1945) 120. [Pages 65, 67, and 70]

[609] R. C. Tolman, *Effect of imhomogeneity on cosmological models*, *Proc. Nat. Acad. Sci.* **20** (1934) 169. [Pages 65 and 70]

[610] H. Bondi, *Spherically symmetrical models in general relativity*, *Mon* **107** (1947) 410. [Pages 65 and 70]

[611] G. S. N. Thakurta, *Kerr metric in an expanding universe*, *Indian J.Phys.B* **55** (1981) 304. [Pages 65, 72, and 73]

[612] G. Hütsi, T. Koivisto, M. Raidal, V. Vaskonen and H. Veermäe, *Cosmological black holes are not described by the thakurta metric: Ligo-virgo bounds on pbhs remain unchanged*, 2105.09328. [Pages 66, 81, and 82]

[613] G. Hütsi, T. Koivisto, M. Raidal, V. Vaskonen and H. Veermäe, *Reply to "comment on: Cosmological black holes are not described by the thakurta metric"*, 2106.02007. [Pages 66, 81, and 82]

[614] T. Harada, H. Maeda and T. Sato, *Thakurta metric does not describe a cosmological black hole*, 2106.06651. [Pages 66, 77, and 81]

[615] A. Kobakhidze. personal correspondence. [Page 66]

[616] S. Xavier, A. Sunny and S. Shankaranarayanan, *An exact model for evaporating primordial black holes in cosmological space-time*, 2110.14379. [Page 66]

[617] K. S. Croker, M. J. Zevin, D. Farrah, K. A. Nishimura and G. Tarle, *Cosmologically Coupled Compact Objects: A Single-parameter Model for LIGO–Virgo Mass and Redshift Distributions*, *Astrophys. J. Lett.* **921** (2021) L22 [2109.08146]. [Page 66]

[618] A. Barrau, K. Martineau and C. Renevey, *The catastrophic fate of Schwarzschild black holes in a thermal bath*, 2203.13297. [Page 66]

[619] M. Carrera and D. Giulini, *Influence of global cosmological expansion on local dynamics and kinematics*, *Rev. Mod. Phys.* **82** (2010) 169. [Pages 67 and 70]

[620] H. Kodama, *Conserved energy flux for the spherically symmetric system and the back reaction problem in the black hole evaporation*, *Prog. Theor. Phys* **63** (1980) 1217. [Pages 67 and 75]

[621] G. Abreu and M. Visser, *Kodama time: Geometrically preferred foliations of spherically symmetric spacetimes*, *Phys. Rev* **82** (2010) 044027 [1004.1456]. [Pages 67, 75, 92, and 94]

[622] A. Einstein and E. G. Straus, *Corrections and additional remarks to our paper: The influence of the expansion of space on the gravitation fields surrounding the individual stars*, *Rev. Mod. Phys.* **18** (1946) 148. [Page 70]

[623] P. S. Joshi and D. Malafarina, *All black holes in Lemaître–Tolman–Bondi inhomogeneous dust collapse*, *Class. Quant. Grav.* **32** (2015) 145004 [1405.1146]. [Page 70]

[624] V. Faraoni and A. Jacques, *Cosmological expansion and local physics*, *Phys. Rev. D* **76** (2007) 063510 [0707.1350]. [Page 71]

[625] W. de Sitter, *Einstein's theory of gravitation and its astronomical consequences. Third paper*, MNRAS **78** (1917) 3. [Page 71]

[626] F. Kottler, *Über die physikalischen Grundlagen der Einsteinschen Gravitationstheorie*, *Annalen der Physik* **361** (1918) 401. [Page 71]

[627] H. Nariai, *On Some Static Solutions of Einstein's Gravitational Field Equations in a Spherically Symmetric Case*, *General Relativity and Gravitation* **31** (1999) 945. [Page 71]

[628] H. Nariai, *On a New Cosmological Solution of Einstein's Field Equations of Gravitation*, *General Relativity and Gravitation* **31** (1999) 963. [Page 71]

[629] J. Sultana and C. C. Dyer, *Cosmological black holes: A black hole in the Einstein-de Sitter universe*, *Gen. Rel. Grav.* **37** (2005) 1347. [Pages 71 and 73]

[630] V. Faraoni, *An Analysis of the Sultana-Dyer cosmological black hole solution of the Einstein equations*, *Phys. Rev. D* **80** (2009) 044013 [0907.4473]. [Page 71]

[631] F. Hammad, D. K. Çiftci and V. Faraoni, *Conformal cosmological black holes: Towards restoring determinism to Einstein theory*, *Eur. Phys. J. Plus* **134** (2019) 480 [1805.09422]. [Pages 71, 72, 73, 76, and 83]

[632] C. Gao, X. Chen, V. Faraoni and Y.-G. Shen, *Does the mass of a black hole decrease due to the accretion of phantom energy*, *Phys. Rev. D* **78** (2008) 024008 [0802.1298]. [Page 72]

[633] T. Clifton, D. F. Mota and J. D. Barrow, *Inhomogeneous gravity*, *Monthly Notices of the Royal Astronomical Society* **358** (2005) 601. [Page 72]

[634] N. Afshordi, D. J. H. Chung, M. Doran and G. Geshnizjani, *Cuscuton Cosmology: Dark Energy meets Modified Gravity*, *Phys. Rev. D* **75** (2007) 123509 [astro-ph/0702002]. [Page 72]

[635] N. Afshordi, G. Geshnizjani and J. Khoury, *Do observations offer evidence for cosmological-scale extra dimensions?*, *JCAP* **08** (2009) 030 [0812.2244]. [Page 72]

[636] E. Abdalla, N. Afshordi, M. Fontanini, D. C. Guariento and E. Papantonopoulos, *Cosmological black holes from self-gravitating fields*, *Phys. Rev. D* **89** (2014) 104018 [1312.3682]. [Page 72]

[637] B. R. Majhi, *Thermodynamics of Sultana-Dyer Black Hole*, *JCAP* **05** (2014) 014 [1403.4058]. [Page 73]

[638] M. M. C. Mello, A. Maciel and V. T. Zanchin, *Evolving black holes from conformal transformations of static solutions*, *Phys. Rev* **95** (2017) 084031 [1611.05077]. [Pages 73 and 76]

[639] V. Faraoni, G. F. R. Ellis, J. T. Firouzjaee, A. Helou and I. Musco, *Foliation dependence of black hole apparent horizons in spherical symmetry*, *Phys. Rev. D* **95** (2017) 024008 [1610.05822]. [Page 77]

[640] P. Peters, *Gravitational Radiation and the Motion of Two Point Masses*, *Phys. Rev.* **136** (1964) B1224. [Pages 87, 88, and 89]

[641] P. C. Peters and J. Mathews, *Gravitational radiation from point masses in a keplerian orbit*, *Phys. Rev.* **131** (1963) 435. [Pages 88 and 89]

[642] K. Chatziioannou et al., *On the properties of the massive binary black hole merger GW170729*, *Phys. Rev. D* **100** (2019) 104015 [1903.06742]. [Page 91]

[643] J. D. Brown and J. W. York, Jr., *Quasilocal energy and conserved charges derived from the gravitational action*, *Phys. Rev. D* **47** (1993) 1407 [gr-qc/9209012]. [Page 94]

[644] Planck Collaboration, P. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13 [1502.01589]. [Page 96]

[645] H. Forkel, *A Primer on instantons in QCD*, hep-ph/0009136. [Pages 100 and 109]

[646] C. Callan, R. Dashen and D. Gross, *The structure of the gauge theory vacuum*, *Physics Letters B* **63** (1976) 334. [Pages 100 and 109]

[647] F. Wilczek, *Problem of strong p and t invariance in the presence of instantons*, *Phys. Rev. Lett.* **40** (1978) 279. [Page 100]

[648] A. Belavin, A. Polyakov, A. Schwartz and Y. Tyupkin, *Pseudoparticle solutions of the yang-mills equations*, *Physics Letters B* **59** (1975) 85. [Pages 100 and 108]

[649] D. V. Nanopoulos, *Towards a renormalizable unified gauge theory of strong, electromagnetic and weak interactions*, *Lett. Nuovo Cim.* **8S2** (1973) 873. [Page 100]

[650] S. Weinberg, *Non-abelian gauge theories of the strong interactions*, *Phys. Rev. Lett.* **31** (1973) 494. [Page 100]

[651] G. 't Hooft, *Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle*, *Phys. Rev. D* **14** (1976) 3432. [Erratum: Phys.Rev.D 18, 2199 (1978)]. [Page 100]

[652] G. 't Hooft, *How instantons solve the u(1) problem*, *Physics Reports* **142** (1986) 357. [Page 100]

[653] A. J. Buchmann, *Structure of strange baryons*, in *9th International Conference on Hypernuclear and Strange Particle Physics*, pp. 329–333, 2007, 0712.4383, DOI. [Page 101]

[654] E. M. Purcell and N. F. Ramsey, *On the possibility of electric dipole moments for elementary particles and nuclei*, *Phys. Rev.* **78** (1950) 807. [Page 111]

[655] R. J. Crewther, P. Di Vecchia, G. Veneziano and E. Witten, *Chiral estimate of the electric dipole moment of the neutron in quantum chromodynamics*, *Physics Letters B* **91** (1980) 487. [Page 111]

[656] A. Pich and E. de Rafael, *Strong cp-violation in an effective chiral lagrangian approach*, *Nuclear Physics B* **367** (1991) 313. [Page 111]

[657] M. Pospelov and A. Ritz, *Theta vacua, qcd sum rules, and the neutron electric dipole moment*, *Nuclear Physics B* **573** (2000) 177. [Page 111]

[658] L. Bartolini, F. Bigazzi, S. Bolognesi, A. L. Cotrone and A. Manenti, *Neutron electric dipole moment from gauge-string duality*, *Phys. Rev. Lett.* **118** (2017) 091601. [Page 111]

[659] M. Abramczyk, S. Aoki, T. Blum, T. Izubuchi, H. Ohki and S. Syritsyn, *Lattice calculation of electric dipole moments and form factors of the nucleon*, *Phys. Rev. D* **96** (2017) 014501 [1701.07792]. [Page 111]

[660] J. Dragos, T. Luu, A. Shindler, J. de Vries and A. Yousif, *Confirming the Existence of the strong CP Problem in Lattice QCD with the Gradient Flow*, 1902.03254. [Page 111]

[661] J. H. Smith, E. M. Purcell and N. F. Ramsey, *Experimental limit to the electric dipole moment of the neutron*, *Phys. Rev.* **108** (1957) 120. [Page 111]

[662] C. A. Baker et al., *An Improved experimental limit on the electric dipole moment of the neutron*, *Phys. Rev. Lett.* **97** (2006) 131801 [hep-ex/0602020]. [Page 111]

[663] J. M. Pendlebury, S. Afach, N. J. Ayres, C. A. Baker, G. Ban, G. Bison, K. Bodek, M. Burghoff, P. Geltenbort, K. Green, W. C. Griffith, M. van der Grinten, Z. D. Grujić, P. G. Harris, V. Hélaine, P. Iaydjiev, S. N. Ivanov, M. Kasprzak, Y. Kermaidic, K. Kirch, H.-C. Koch, S. Komposch, A. Kozela, J. Krempel, B. Lauss, T. Lefort, Y. Lemière, D. J. R. May, M. Musgrave, O. Naviliat-Cuncic, F. M. Piegsa, G. Pignol, P. N. Prashanth, G. Quéméner, M. Rawlik, D. Rebreyend, J. D. Richardson, D. Ries, S. Roccia, D. Rozpedzik, A. Schnabel, P. Schmidt-Wellenburg, N. Severijns, D. Shiers, J. A. Thorne, A. Weis, O. J. Winston, E. Wursten, J. Zejma and G. Zsigmond, *Revised experimental upper limit on the electric dipole moment of the neutron*, *Phys. Rev. D* **92** (2015) 092003. [Page 111]

[664] M. Dine, L. Stephenson Haskins, L. Ubaldi and D. Xu, *Some Remarks on Anthropic Approaches to the Strong CP Problem*, *JHEP* **05** (2018) 171 [1801.03466]. [Page 113]

[665] S. Y. Khlebnikov and M. E. Shaposhnikov, *Extra Space-time Dimensions: Towards a Solution to the Strong CP Problem*, *Phys. Lett. B* **203** (1988) 121. [Page 113]

[666] G. Schierholz, *Towards a dynamical solution of the strong CP problem*, *Nucl. Phys. B Proc. Suppl.* **37** (1994) 203 [hep-lat/9403012]. [Page 113]

[667] K. S. Babu and R. N. Mohapatra, *A Solution to the Strong CP Problem Without an Axion*, *Phys. Rev. D* **41** (1990) 1286. [Page 113]

[668] A. E. Nelson, *Calculation of θ Barr*, *Phys. Lett. B* **143** (1984) 165. [Page 113]

[669] A. E. Nelson, *Naturally Weak CP Violation*, *Phys. Lett. B* **136** (1984) 387. [Page 113]

[670] S. M. Barr, *Solving the Strong CP Problem Without the Peccei-Quinn Symmetry*, *Phys. Rev. Lett.* **53** (1984) 329. [Page 113]

[671] S. M. Barr, *A Natural Class of Nonpeccei-quinn Models*, *Phys. Rev. D* **30** (1984) 1805. [Page 113]

[672] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531. [Page 113]

[673] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652. [Page 113]

[674] A. G. Cohen, D. B. Kaplan and A. E. Nelson, *Testing $m_u = 0$ on the lattice*, *Journal of High Energy Physics* **1999** (1999) 027. [Page 114]

[675] M. Dine, P. Draper and G. Festuccia, *Instanton effects in three flavor qcd*, *Phys. Rev. D* **92** (2015) 054004. [Page 114]

[676] C. Alexandrou, J. Finkenrath, L. Funcke, K. Jansen, B. Kostrzewa, F. Pittler and C. Urbach, *Ruling Out the Massless Up-Quark Solution to the Strong $\boldsymbol{CP}$ Problem by Computing the Topological Mass Contribution with Lattice QCD*, *Phys. Rev. Lett.* **125** (2020) 232001 [2002.07802]. [Page 114]

[677] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223. [Pages 114 and 118]

[678] F. Wilczek, *Problem of Strong p and t Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279. [Pages 114 and 118]

[679] F. Wilczek, *Time's (almost) reversible arrow*, https://www.quantamagazine.org/how-axions-may-explain-times-arrow-20160107. [Page 115]

[680] P. Svrcek and E. Witten, *Axions in string theory*, *JHEP* **06** (2006) 051 [hep-th/0605206]. [Page 117]

[681] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper and J. March-Russell, *String axiverse*, *Phys. Rev. D* **81** (2010) 123530 [0905.4720]. [Page 117]

[682] M. Cicoli, M. Goodsell and A. Ringwald, *The type IIB string axiverse and its low-energy phenomenology*, *JHEP* **10** (2012) 146 [1206.0819]. [Page 117]

[683] W. A. Bardeen and S. H. H. Tye, *Current Algebra Applied to Properties of the Light Higgs Boson*, *Phys. Lett. B* **74** (1978) 229. [Pages 117 and 139]

[684] D. J. E. Marsh, *Axion cosmology*, *Phys. Rept.* **643** (2016) 1 [1510.07633]. [Pages 118, 119, and 120]

[685] Y. Asano, E. Kikutani, S. Kurokawa, T. Miyachi, M. Miyajima, Y. Nagashima, T. Shinkawa, S. Sugimoto and Y. Yoshimura, *Search for a Rare Decay Mode K+ —> pi+ Neutrino anti-neutrino and Axion*, *Phys. Lett. B* **107** (1981) 159. [Page 118]

[686] W. A. Bardeen, R. D. Peccei and T. Yanagida, *CONSTRAINTS ON VARIANT AXION MODELS*, *Nucl. Phys. B* **279** (1987) 401. [Page 118]

[687] J. E. Kim, *Light pseudoscalars, particle physics and cosmology*, *Phys. Rept.* **150** (1987) 1. [Page 118]

[688] K. Mimasu and V. Sanz, *ALPs at Colliders*, *JHEP* **06** (2015) 173 [1409.4792]. [Page 118]

[689] M. Dine, W. Fischler and M. Srednicki, *A simple solution to the Strong CP Problem with a harmless axion*, *Phys. Lett. B* **104** (1981) 199. [Page 118]

[690] A. R. Zhitnitsky, *On Possible Suppression of the Axion Hadron Interactions. (In Russian)*, *Sov. J. Nucl. Phys.* **31** (1980) 260. [Yad. Fiz.31,497(1980)]. [Page 118]

[691] J. E. Kim, *Weak interaction singlet and Strong CP invariance*, *Phys. Rev. Lett.* **43** (1979) 103. [Pages 118 and 139]

[692] M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, *Can Confinement Ensure Natural CP Invariance of Strong Interactions?*, *Nucl. Phys. B* **166** (1980) 493. [Pages 118 and 139]

[693] R. L. Davis, *Cosmic Axions from Cosmic Strings*, *Phys. Lett. B* **180** (1986) 225. [Pages 119 and 120]

[694] C. A. J. O'Hare, *Detecting WIMPs, neutrinos and axions in the next generation of dark matter experiment*, Ph.D. thesis, Nottingham U., Nottingham U., 2017. [Page 119]

[695] E. Di Valentino, E. Giusarma, M. Lattanzi, O. Mena, A. Melchiorri and J. Silk, *Cosmological Axion and neutrino mass constraints from Planck 2015 temperature and polarization data*, *Phys. Lett. B* **752** (2016) 182 [1507.08665]. [Page 119]

[696] A. Vilenkin and A. E. Everett, *Cosmic Strings and Domain Walls in Models with Goldstone and PseudoGoldstone Bosons*, *Phys. Rev. Lett.* **48** (1982) 1867. [Pages 120 and 149]

[697] P. Sikivie, *Of Axions, Domain Walls and the Early Universe*, *Phys. Rev. Lett.* **48** (1982) 1156. [Pages 120 and 149]

[698] T. W. B. Kibble, G. Lazarides and Q. Shafi, *Walls Bounded by Strings*, *Phys. Rev. D* **26** (1982) 435. [Pages 120 and 149]

[699] T. Vachaspati and A. Vilenkin, *Formation and Evolution of Cosmic Strings*, *Phys. Rev. D* **30** (1984) 2036. [Page 120]

[700] R. L. Davis, *Goldstone Bosons in String Models of Galaxy Formation*, *Phys. Rev. D* **32** (1985) 3172. [Page 120]

[701] E. P. S. Shellard, *Cosmic String Interactions*, *Nucl. Phys. B* **283** (1987) 624. [Page 120]

[702] R. L. Davis and E. P. S. Shellard, *DO AXIONS NEED INFLATION?*, *Nucl. Phys. B* **324** (1989) 167. [Page 120]

[703] C. J. Hogan and M. J. Rees, *AXION MINICLUSTERS*, *Phys. Lett. B* **205** (1988) 228. [Page 121]

[704] E. W. Kolb and I. I. Tkachev, *Axion miniclusters and Bose stars*, *Phys. Rev. Lett.* **71** (1993) 3051 [hep-ph/9303313]. [Page 121]

[705] E. W. Kolb and I. I. Tkachev, *Nonlinear axion dynamics and formation of cosmological pseudosolitons*, *Phys. Rev. D* **49** (1994) 5040 [astro-ph/9311037]. [Page 121]

[706] E. W. Kolb and I. I. Tkachev, *Femtolensing and picolensing by axion miniclusters*, *Astrophys. J. Lett.* **460** (1996) L25 [astro-ph/9510043]. [Page 121]

[707] I. I. Tkachev, *On the possibility of Bose star formation*, *Phys. Lett. B* **261** (1991) 289. [Page 121]

[708] B. Eggemeier, J. Redondo, K. Dolag, J. C. Niemeyer and A. Vaquero, *First Simulations of Axion Minicluster Halos*, *Phys. Rev. Lett.* **125** (2020) 041301 [1911.09417]. [Page 121]

[709] P. Sikivie and Q. Yang, *Bose-Einstein Condensation of Dark Matter Axions*, *Phys. Rev. Lett.* **103** (2009) 111301 [0901.1106]. [Page 121]

[710] O. Erken, P. Sikivie, H. Tam and Q. Yang, *Axion BEC Dark Matter*, in *7th Patras Workshop on Axions, WIMPs and WISPs*, pp. 43–46, 11, 2011, 1111.3976, DOI. [Page 121]

[711] C. O'Hare, *cajohare/axionlimits: Axionlimits*, July, 2020. 10.5281/zenodo.3932430. [Page 122]

[712] P. Sikivie, *Experimental tests of the "invisible" axion*, *Phys. Rev. Lett.* **51** (1983) 1415. [Pages 121 and 142]

[713] E. A. Paschos and K. Zioutas, *A Proposal for solar axion detection via Bragg scattering*, *Phys. Lett. B* **323** (1994) 367. [Page 121]

[714] A. V. Derbin, A. S. Kayunov, V. N. Muratova, D. A. Semenov and E. V. Unzhakov, *Constraints on the axion-electron coupling for solar axions produced by Compton process and bremsstrahlung*, *Phys. Rev. D* **83** (2011) 023505 [1101.2290]. [Page 121]

[715] K. Arisaka, P. Beltrame, C. Ghag, J. Kaidi, K. Lung, A. Lyashenko, R. D. Peccei, P. Smith and K. Ye, *Expected Sensitivity to Galactic/Solar Axions and Bosonic Super-WIMPs based on the Axio-electric Effect in Liquid Xenon Dark Matter Detectors*, *Astropart. Phys.* **44** (2013) 59 [1209.3810]. [Page 121]

[716] K. van Bibber, P. M. McIntyre, D. E. Morris and G. G. Raffelt, *A Practical Laboratory Detector for Solar Axions*, *Phys. Rev. D* **39** (1989) 2089. [Page 121]

[717] K. Zioutas et al., *A Decommissioned LHC model magnet as an axion telescope*, *Nucl. Instrum. Meth. A* **425** (1999) 480 [astro-ph/9801176]. [Page 121]

[718] CAST Collaboration, K. Zioutas et al., *First results from the CERN Axion Solar Telescope (CAST)*, *Phys. Rev. Lett.* **94** (2005) 121301 [hep-ex/0411033]. [Page 121]

[719] CAST Collaboration, A. A. Melcón et al., *First results of the CAST-RADES haloscope search for axions at 34.67 μeV*, 2104.13798. [Pages 121, 133, and 144]

[720] Y. Inoue, T. Namba, S. Moriyama, M. Minowa, Y. Takasu, T. Horiuchi and A. Yamamoto, *Search for sub-electronvolt solar axions using coherent conversion of axions into photons in magnetic field and gas helium*, *Phys. Lett. B* **536** (2002) 18 [astro-ph/0204388]. [Page 121]

[721] I. G. Irastorza et al., *Towards a new generation axion helioscope*, *JCAP* **06** (2011) 013 [1103.5334]. [Page 121]

[722] IAXO Collaboration, I. Irastorza et al., *The International Axion Observatory IAXO. Letter of Intent to the CERN SPS committee*, . [Page 121]

[723] J. Redondo, *Solar axion flux from the axion-electron coupling*, *JCAP* **1312** (2013) 008 [1310.0823]. [Page 121]

[724] LUX Collaboration, D. S. Akerib et al., *First Searches for Axions and Axionlike Particles with the LUX Experiment*, *Phys. Rev. Lett.* **118** (2017) 261301 [1704.02297]. [Page 121]

[725] XENON100 Collaboration, E. Aprile et al., *First Axion Results from the XENON100 Experiment*, *Phys. Rev. D* **90** (2014) 062009 [1404.1455]. [Erratum: Phys.Rev.D 95, 029904 (2017)]. [Page 121]

[726] PandaX Collaboration, C. Fu et al., *Limits on Axion Couplings from the First 80 Days of Data of the PandaX-II Experiment*, *Phys. Rev. Lett.* **119** (2017) 181806 [1707.07921]. [Page 121]

[727] R. Bähre et al., *Any light particle search II -Technical Design Report*, *JINST* **8** (2013) T09001 [1302.5647]. [Page 121]

[728] OSQAR Collaboration, R. Ballou et al., *New exclusion limits on scalar and pseudoscalar axionlike particles from light shining through a wall*, *Phys. Rev. D* **92** (2015) 092002 [1506.08082]. [Page 121]

[729] ADMX Collaboration, N. Du et al., *A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment*, *Phys. Rev. Lett.* **120** (2018) 151301 [1804.05750]. [Pages 121 and 133]

[730] S. Beurthey et al., *MADMAX Status Report*, 2003.10894. [Pages 122 and 140]

[731] B. M. Brubaker et al., *First results from a microwave cavity axion search at 24 μeV*, *Phys. Rev. Lett.* **118** (2017) 061302 [1610.02580]. [Page 122]

[732] B. T. McAllister, G. Flower, E. N. Ivanov, M. Goryachev, J. Bourhill and M. E. Tobar, *The ORGAN Experiment: An axion haloscope above 15 GHz*, *Phys. Dark Univ.* **18** (2017) 67 [1706.00209]. [Pages 122, 133, and 145]

[733] J. L. Ouellet et al., *First Results from ABRACADABRA-10 cm: A Search for Sub-μeV Axion Dark Matter*, *Phys. Rev. Lett.* **122** (2019) 121802 [1810.12257]. [Pages 122 and 145]

[734] https://irwinlab.sites.stanford.edu/dark-matter-radio-dmradio. [Pages 122, 140, and 145]

[735] Y. Kahn, B. R. Safdi and J. Thaler, *Broadband and Resonant Approaches to Axion Dark Matter Detection*, *Phys. Rev. Lett.* **117** (2016) 141801 [1602.01086]. [Pages 122, 140, and 145]

[736] D. Jackson Kimball et al., *Overview of the Cosmic Axion Spin Precession Experiment (CASPEr)*, *Springer Proc. Phys.* **245** (2020) 105 [1711.08999]. [Page 122]

[737] A. Ayala, I. Domínguez, M. Giannotti, A. Mirizzi and O. Straniero, *Revisiting the bound on axion-photon coupling from Globular Clusters*, *Phys. Rev. Lett.* **113** (2014) 191302 [1406.6053]. [Pages 122 and 141]

[738] N. Vinyoles, A. Serenelli, F. L. Villante, S. Basu, J. Redondo and J. Isern, *New axion and hidden photon constraints from a solar data global fit*, *JCAP* **1510** (2015) 015 [1501.01639]. [Page 123]

[739] G. G. Raffelt, *Astrophysical axion bounds*, *Lect. Notes Phys.* **741** (2008) 51 [hep-ph/0611350]. [Pages 123 and 140]

[740] J. Jaeckel, P. C. Malta and J. Redondo, *Decay photons from the axionlike particles burst of type II supernovae*, *Phys. Rev. D* **98** (2018) 055032 [1702.02964]. [Page 123]

[741] A. Payez, C. Evoli, T. Fischer, M. Giannotti, A. Mirizzi and A. Ringwald, *Revisiting the SN1987A gamma-ray limit on ultralight axion-like particles*, *JCAP* **1502** (2015) 006 [1410.3747]. [Page 123]

[742] R. Penrose and R. Floyd, *Extraction of rotational energy from a black hole*, *Nature* **229** (1971) 177. [Page 123]

[743] S. R. Dolan, *Instability of the massive Klein-Gordon field on the Kerr spacetime*, *Phys. Rev. D* **76** (2007) 084001 [0705.2880]. [Page 123]

[744] P. Pani, V. Cardoso, L. Gualtieri, E. Berti and A. Ishibashi, *Black hole bombs and photon mass bounds*, *Phys. Rev. Lett.* **109** (2012) 131102 [1209.0465]. [Page 123]

[745] A. Arvanitaki and S. Dubovsky, *Exploring the String Axiverse with Precision Black Hole Physics*, *Phys. Rev. D* **83** (2011) 044026 [1004.3558]. [Page 123]

[746] R. Brito, V. Cardoso and P. Pani, *Superradiance: Energy Extraction, Black-Hole Bombs and Implications for Astrophysics and Particle Physics*, vol. 906. Springer, 2015, 10.1007/978-3-319-19000-6, [1501.06570]. [Page 123]

[747] A. Arvanitaki, M. Baryakhtar, S. Dimopoulos, S. Dubovsky and R. Lasenby, *Black Hole Mergers and the QCD Axion at Advanced LIGO*, *Phys. Rev. D* **95** (2017) 043001 [1604.03958]. [Page 123]

[748] C. Herdeiro, E. Radu and H. Rúnarsson, *Kerr black holes with Proca hair*, *Class. Quant. Grav.* **33** (2016) 154001 [1603.02687]. [Page 123]

[749] V. Cardoso, O. J. C. Dias, G. S. Hartnett, M. Middleton, P. Pani and J. E. Santos, *Constraining the mass of dark photons and axion-like particles through black-hole superradiance*, *JCAP* **03** (2018) 043 [1801.01420]. [Page 123]

[750] M. J. Stott and D. J. E. Marsh, *Black hole spin constraints on the mass spectrum and number of axionlike fields*, *Phys. Rev. D* **98** (2018) 083006 [1805.02016]. [Page 123]

[751] M. J. Stott, *Ultralight Bosonic Field Mass Bounds from Astrophysical Black Hole Spin*, 2009.07206. [Pages 123, 140, and 145]

[752] M. Baryakhtar, M. Galanis, R. Lasenby and O. Simon, *Black hole superradiance of self-interacting scalar fields*, *Phys. Rev. D* **103** (2021) 095019 [2011.11646]. [Page 123]

[753] V. M. Mehta, M. Demirtas, C. Long, D. J. E. Marsh, L. Mcallister and M. J. Stott, *Superradiance Exclusions in the Landscape of Type IIB String Theory*, 2011.08693. [Pages 123, 140, and 145]

[754] V. M. Mehta, M. Demirtas, C. Long, D. J. E. Marsh, L. McAllister and M. J. Stott, *Superradiance in string theory*, *JCAP* **07** (2021) 033 [2103.06812]. [Pages 123, 140, and 145]

[755] M. Baryakhtar, R. Lasenby and M. Teo, *Black Hole Superradiance Signatures of Ultralight Vectors*, *Phys. Rev. D* **96** (2017) 035019 [1704.05081]. [Page 123]

[756] S. Zaroubi, *The Epoch of Reionization*, in *The First Galaxies* (T. Wiklind, B. Mobasher and V. Bromm, eds.), vol. 396 of *Astrophysics and Space Science Library*, p. 45, Jan., 2013, 1206.0267, DOI. [Page 123]

[757] D. Cadamuro and J. Redondo, *Cosmological bounds on pseudo Nambu-Goldstone bosons*, *JCAP* **02** (2012) 032 [1110.2895]. [Page 123]

[758] J. M. Overduin and P. S. Wesson, *Dark matter and background light*, *Phys. Rept.* **402** (2004) 267 [astro-ph/0407207]. [Page 123]

[759] D. Grin, G. Covone, J.-P. Kneib, M. Kamionkowski, A. Blain and E. Jullo, *A Telescope Search for Decaying Relic Axions*, *Phys. Rev. D* **75** (2007) 105018 [astro-ph/0611502]. [Page 123]

[760] Z. Chen and A. Kobakhidze, *Coloured gravitational instantons, the strong CP problem and the companion axion solution*, 2108.05549. [Pages 125, 126, and 128]

[761] S. W. Hawking, *Gravitational Instantons*, *Phys. Lett. A* **60** (1977) 81. [Page 126]

[762] G. W. Gibbons and S. W. Hawking, *Gravitational Multi - Instantons*, *Phys. Lett. B* **78** (1978) 430. [Pages 126 and 127]

[763] S. Deser, M. J. Duff and C. J. Isham, *Gravitationally Induced CP Effects*, *Phys. Lett. B* **93** (1980) 419. [Page 126]

[764] T. Eguchi and A. J. Hanson, *Asymptotically Flat Selfdual Solutions to Euclidean Gravity*, *Phys. Lett. B* **74** (1978) 249. [Page 127]

[765] T. Eguchi and A. J. Hanson, *Selfdual Solutions to Euclidean Gravity*, *Annals Phys.* **120** (1979) 82. [Page 127]

[766] M. K. Prasad, *Equivalence of Eguchi-Hanson metric to two-center Gibbons-Hawking metric*, *Phys. Lett. B* **83** (1979) 310. [Page 127]

[767] S. Arunasalam and A. Kobakhidze, *Charged gravitational instantons: extra CP violation and charge quantisation in the Standard Model*, *Eur. Phys. J. C* **79** (2019) 49 [1808.01796]. [Page 127]

[768] H. Boutaleb-Joutei, A. Chakrabarti and A. Comtet, *Gauge field configurations in curved space-times. iv. self-dual su(2) fields in multicenter spaces*, *Phys. Rev. D* **21** (1980) 2280. [Page 127]

[769] A. Chakrabarti, *Classical Solutions of Yang-Mills Fields. (Selected Topics)*, *Fortsch. Phys.* **35** (1987) 1. [Page 127]

[770] R. Holman, T. W. Kephart and S.-J. Rey, *Semiclassical gravity and invisible axions*, *Phys. Rev. Lett.* **71** (1993) 320 [hep-ph/9207208]. [Pages 127 and 129]

[771] CAST Collaboration, S. Andriamonje et al., *An Improved limit on the axion-photon coupling from the CAST experiment*, *JCAP* **04** (2007) 010 [hep-ex/0702006]. [Page 133]

[772] CAST Collaboration, V. Anastassopoulos et al., *New CAST Limit on the Axion-Photon Interaction*, *Nature Phys.* **13** (2017) 584 [1705.02290]. [Page 133]

[773] S. J. Asztalos, G. Carosi, C. Hagmann, D. Kinion, K. van Bibber, M. Hotz, L. J. Rosenberg, G. Rybka, J. Hoskins, J. Hwang, P. Sikivie, D. B. Tanner, R. Bradley, J. Clarke and ADMX Collaboration, *SQUID-Based Microwave Cavity Search for Dark-Matter Axions*, *Phys. Rev. Lett.* **104** (2010) 041301 [0910.5914]. [Page 133]

[774] ADMX Collaboration, T. Braine et al., *Extended Search for the Invisible Axion with the Axion Dark Matter Experiment*, *Phys. Rev. Lett.* **124** (2020) 101303 [1910.08638]. [Pages 133 and 144]

[775] ADMX Collaboration, C. Boutan et al., *Piezoelectrically Tuned Multimode Cavity Search for Axion Dark Matter*, *Phys. Rev. Lett.* **121** (2018) 261302 [1901.00920]. [Page 133]

[776] N. Crisosto, P. Sikivie, N. S. Sullivan, D. B. Tanner, J. Yang and G. Rybka, *ADMX SLIC: Results from a Superconducting LC Circuit Investigating Cold Axions*, *Phys. Rev. Lett.* **124** (2020) 241101 [1911.05772]. [Pages 133 and 145]

[777] S. Lee, S. Ahn, J. Choi, B. Ko and Y. Semertzidis, *Axion Dark Matter Search around 6.7 μeV*, *Phys. Rev. Lett.* **124** (2020) 101802 [2001.05102]. [Page 133]

[778] J. Jeong, S. Youn, S. Bae, J. Kim, T. Seong, J. E. Kim and Y. K. Semertzidis, *Search for Invisible Axion Dark Matter with a Multiple-Cell Haloscope*, *Phys. Rev. Lett.* **125** (2020) 221302 [2008.10141]. [Page 133]

[779] CAPP Collaboration, O. Kwon et al., *First Results from an Axion Haloscope at CAPP around 10.7 μeV*, *Phys. Rev. Lett.* **126** (2021) 191802 [2012.10764]. [Page 133]

[780] J. A. Devlin et al., *Constraints on the Coupling between Axionlike Dark Matter and Photons Using an Antiproton Superconducting Tuned Detection Circuit in a Cryogenic Penning Trap*, *Phys. Rev. Lett.* **126** (2021) 041301 [2101.11290]. [Pages 133 and 145]

[781] HAYSTAC Collaboration, L. Zhong et al., *Results from phase 1 of the HAYSTAC microwave cavity axion experiment*, *Phys. Rev. D* **97** (2018) 092001 [1803.03690]. [Page 133]

[782] HAYSTAC Collaboration, K. M. Backes et al., *A quantum-enhanced search for dark matter axions*, *Nature* **590** (2021) 238 [2008.01853]. [Page 133]

[783] D. Alesini et al., *Galactic axions search with a superconducting resonant cavity*, *Phys. Rev. D* **99** (2019) 101101 [1903.06547]. [Page 133]

[784] D. Alesini et al., *Search for invisible axion dark matter of mass $m_a = 43$ μeV with the QUAX–$a\gamma$ experiment*, *Phys. Rev. D* **103** (2021) 102004 [2012.09498]. [Page 133]

[785] S. DePanfilis, A. C. Melissinos, B. E. Moskowitz, J. T. Rogers, Y. K. Semertzidis, W. U. Wuensch, H. J. Halama, A. G. Prodell, W. B. Fowler and F. A. Nezrick, *Limits on the abundance and coupling of cosmic axions at 4.5< m_a <5.0 μev*, *Phys. Rev. Lett.* **59** (1987) 839. [Page 133]

[786] C. Hagmann, P. Sikivie, N. S. Sullivan and D. B. Tanner, *Results from a search for cosmic axions*, *Phys. Rev. D* **42** (1990) 1297. [Page 133]

[787] M. Kawasaki, K. Saikawa and T. Sekiguchi, *Axion dark matter from topological defects*, *Phys. Rev. D* **91** (2015) 065014 [1412.0789]. [Pages 134 and 153]

[788] L. Fleury and G. D. Moore, *Axion dark matter: strings and their cores*, *JCAP* **01** (2016) 004 [1509.00026]. [Page 134]

[789] V. B. Klaer and G. D. Moore, *The dark-matter axion mass*, *JCAP* **1711** (2017) 049 [1708.07521]. [Page 134]

[790] M. Buschmann, J. W. Foster and B. R. Safdi, *Early-Universe Simulations of the Cosmological Axion*, *Phys. Rev. Lett.* **124** (2020) 161103 [1906.00967]. [Page 134]

[791] A. Vaquero, J. Redondo and J. Stadler, *Early seeds of axion miniclusters*, *JCAP* **04** (2019) 012 [1809.09241]. [Page 134]

[792] M. Gorghetto, E. Hardy and G. Villadoro, *Axions from Strings: the Attractive Solution*, *JHEP* **07** (2018) 151 [1806.04677]. [Page 134]

[793] M. Gorghetto, E. Hardy and G. Villadoro, *More Axions from Strings*, *SciPost Phys.* **10** (2021) 050 [2007.04990]. [Page 134]

[794] M. Buschmann, J. W. Foster, A. Hook, A. Peterson, D. E. Willcox, W. Zhang and B. R. Safdi, *Dark Matter from Axion Strings with Adaptive Mesh Refinement*, 2108.05368. [Page 134]

[795] D. Cyncynates, T. Giurgica-Tiron, O. Simon and J. O. Thompson, *Friendship in the Axiverse: Late-time direct and astrophysical signatures of early-time nonlinear axion dynamics*, 2109.09755. [Page 135]

[796] M. S. Turner, *Cosmic and Local Mass Density of Invisible Axions*, *Phys. Rev. D* **33** (1986) 889. [Page 135]

[797] D. H. Lyth, *Axions and inflation: Sitting in the vacuum*, *Phys. Rev. D* **45** (1992) 3394. [Page 135]

[798] K. Strobl and T. J. Weiler, *Anharmonic evolution of the cosmic axion density spectrum*, *Phys. Rev. D* **50** (1994) 7690 [astro-ph/9405028]. [Page 135]

[799] K. J. Bae, J.-H. Huh and J. E. Kim, *Update of axion CDM energy*, *JCAP* **09** (2008) 005 [0806.0497]. [Page 135]

[800] L. Visinelli and P. Gondolo, *Dark Matter Axions Revisited*, *Phys. Rev. D* **80** (2009) 035024 [0903.4377]. [Page 135]

[801] O. Wantz and E. P. S. Shellard, *Axion Cosmology Revisited*, *Phys. Rev. D* **82** (2010) 123508 [0910.1066]. [Page 135]

[802] ADMX Collaboration, C. Bartram et al., *Search for "Invisible" Axion Dark Matter in the 3.3–4.2 μeV Mass Range*, 2110.06096. [Pages 137 and 140]

[803] C. L. Carilli and S. Rawlings, *Science with the Square Kilometer Array: Motivation, key science projects, standards and assumptions*, *New Astron. Rev.* **48** (2004) 979 [astro-ph/0409274]. [Pages 137, 150, and 152]

[804] G. Janssen et al., *Gravitational wave astronomy with the SKA*, *PoS* **AASKA14** (2015) 037 [1501.00127]. [Pages 137, 150, and 152]

[805] A. Weltman et al., *Fundamental physics with the Square Kilometre Array*, *Publ. Astron. Soc. Austral.* **37** (2020) e002 [1810.02680]. [Pages 137, 150, and 152]

[806] M. Srednicki, *Axion Couplings to Matter. 1. CP Conserving Parts*, *Nucl. Phys. B* **260** (1985) 689. [Page 139]

[807] G. Grilli di Cortona, E. Hardy, J. Pardo Vega and G. Villadoro, *The QCD axion, precisely*, *JHEP* **01** (2016) 034 [1511.02867]. [Page 139]

[808] IAXO Collaboration, E. Armengaud et al., *Physics potential of the International Axion Observatory (IAXO)*, *JCAP* **06** (2019) 047 [1904.09155]. [Pages 140, 144, and 145]

[809] F. Chadha-Day, *Axion-like particle oscillations*, 2107.12813. [Pages 141 and 142]

[810] CAST Collaboration, M. Arik et al., *Search for Solar Axions by the CERN Axion Solar Telescope with ^{3}He Buffer Gas: Closing the Hot Dark Matter Gap*, *Phys. Rev. Lett.* **112** (2014) 091302 [1307.1985]. [Page 142]

[811] CAST Collaboration, M. Arik et al., *New solar axion search using the CERN Axion Solar Telescope with ^{4}He filling*, *Phys. Rev. D* **92** (2015) 021101 [1503.00610]. [Page 142]

[812] E. Armengaud et al., *Conceptual design of the international axion observatory (IAXO)*, *JINST* **9** (2014) T05002 [1401.3233]. [Page 145]

[813] IAXO Collaboration, A. Abeln et al., *Conceptual design of BabyIAXO, the intermediate stage towards the International Axion Observatory*, *JHEP* **05** (2021) 137 [2010.12076]. [Page 145]

[814] I. Stern, *ADMX Status*, *PoS* **ICHEP2016** (2016) 198 [1612.08296]. [Page 145]

[815] A. A. Melcón et al., *Axion Searches with Microwave Filters: the RADES project*, *JCAP* **05** (2018) 040 [1803.01243]. [Page 145]

[816] A. Álvarez Melcón et al., *Scalable haloscopes for axion dark matter detection in the 30μeV range with RADES*, *JHEP* **07** (2020) 084 [2002.07639]. [Page 145]

[817] D. Alesini, D. Babusci, D. Di Gioacchino, C. Gatti, G. Lamanna and C. Ligi, *The KLASH Proposal*, 1707.06010. [Page 145]

[818] J. Jeong, S. Youn, S. Ahn, J. E. Kim and Y. K. Semertzidis, *Concept of multiple-cell cavity for axion dark matter search*, *Phys. Lett. B* **777** (2018) 412 [1710.06969]. [Page 145]

[819] A. V. Gramolin, D. Aybas, D. Johnson, J. Adam and A. O. Sushkov, *Search for axion-like dark matter with ferromagnets*, *Nature Phys.* **17** (2021) 79 [2003.03348]. [Page 145]

[820] C. P. Salemi et al., *Search for Low-Mass Axion Dark Matter with ABRACADABRA-10 cm*, *Phys. Rev. Lett.* **127** (2021) 081801 [2102.06722]. [Page 145]

[821] M. Beltran, J. Garcia-Bellido and J. Lesgourgues, *Isocurvature bounds on axions revisited*, *Phys. Rev. D* **75** (2007) 103507 [hep-ph/0606107]. [Page 147]

[822] M. Beltran, J. Garcia-Bellido, J. Lesgourgues, A. R. Liddle and A. Slosar, *Bayesian model selection and isocurvature perturbations*, *Phys. Rev. D* **71** (2005) 063532 [astro-ph/0501477]. [Page 147]

[823] P. Crotty, J. Garcia-Bellido, J. Lesgourgues and A. Riazuelo, *Bounds on isocurvature perturbations from CMB and LSS data*, *Phys. Rev. Lett.* **91** (2003) 171301 [astro-ph/0306286]. [Page 147]

[824] T. Kobayashi, R. Kurematsu and F. Takahashi, *Isocurvature Constraints and Anharmonic Effects on QCD Axion Dark Matter*, *JCAP* **09** (2013) 032 [1304.0922]. [Page 147]

[825] Y. B. Zeldovich, I. Y. Kobzarev and L. B. Okun, *Cosmological Consequences of the Spontaneous Breakdown of Discrete Symmetry*, *Zh. Eksp. Teor. Fiz.* **67** (1974) 3. [Page 149]

[826] S. M. Barr, X. C. Gao and D. Reiss, *Peccei-Quinn symmetries without domains*, *Phys. Rev. D* **26** (1982) 2176. [Page 149]

[827] G. Lazarides and Q. Shafi, *Axion Models with No Domain Wall Problem*, *Phys. Lett. B* **115** (1982) 21. [Page 149]

[828] G. R. Dvali, Z. Tavartkiladze and J. Nanobashvili, *Biased discrete symmetry and domain wall problem*, *Phys. Lett. B* **352** (1995) 214 [hep-ph/9411387]. [Page 149]

[829] S. Chang, C. Hagmann and P. Sikivie, *Axions from wall decay*, *Nucl. Phys. B Proc. Suppl.* **72** (1999) 99 [hep-ph/9808302]. [Page 149]

[830] B. Rai and G. Senjanovic, *Gravity and domain wall problem*, *Phys. Rev. D* **49** (1994) 2729 [hep-ph/9301240]. [Page 149]

[831] S. M. Barr and J. E. Kim, *New Confining Force Solution of the QCD Axion Domain-Wall Problem*, *Phys. Rev. Lett.* **113** (2014) 241301 [1407.4311]. [Page 149]

[832] M. Reig, *On the high-scale instanton interference effect: axion models without domain wall problem*, *JHEP* **08** (2019) 167 [1901.00203]. [Page 149]

[833] A. Caputo and M. Reig, *Cosmic implications of a low-scale solution to the axion domain wall problem*, *Phys. Rev. D* **100** (2019) 063530 [1905.13116]. [Page 149]

[834] G. B. Gelmini, M. Gleiser and E. W. Kolb, *Cosmology of Biased Discrete Symmetry Breaking*, *Phys. Rev. D* **39** (1989) 1558. [Page 149]

[835] M. Kawasaki and K. Saikawa, *Study of gravitational radiation from cosmic domain walls*, *JCAP* **09** (2011) 008 [1102.5628]. [Page 151]

[836] T. Hiramatsu, M. Kawasaki and K. Saikawa, *Gravitational Waves from Collapsing Domain Walls*, *JCAP* **05** (2010) 032 [1002.1555]. [Page 151]

[837] T. Hiramatsu, M. Kawasaki and K. Saikawa, *On the estimation of gravitational wave spectrum from cosmic domain walls*, *JCAP* **02** (2014) 031 [1309.5001]. [Page 151]

[838] K. Saikawa, *A review of gravitational waves from cosmic domain walls*, *Universe* **3** (2017) 40 [1703.02576]. [Page 151]

[839] E. Thrane and J. D. Romano, *Sensitivity curves for searches for gravitational-wave backgrounds*, *Phys. Rev. D* **88** (2013) 124032 [1310.5300]. [Page 152]

[840] C. J. Moore, R. H. Cole and C. P. L. Berry, *Gravitational-wave sensitivity curves*, *Class. Quant. Grav.* **32** (2015) 015014 [1408.0740]. [Page 152]

[841] M. A. McLaughlin, *The North American Nanohertz Observatory for Gravitational Waves*, *Class. Quant. Grav.* **30** (2013) 224008 [1310.0758]. [Page 152]

[842] NANOGRAV Collaboration, Z. Arzoumanian et al., *The NANOGrav 11-year Data Set: Pulsar-timing Constraints On The Stochastic Gravitational-wave Background*, *Astrophys. J.* **859** (2018) 47 [1801.02617]. [Page 152]

[843] K. Aggarwal et al., *The NANOGrav 11-Year Data Set: Limits on Gravitational Waves from Individual Supermassive Black Hole Binaries*, *Astrophys. J.* **880** (2019) 2 [1812.11585]. [Page 152]

[844] A. Brazier et al., *The NANOGrav Program for Gravitational Waves and Fundamental Physics*, 1908.05356. [Page 152]

[845] M. Kramer and D. J. Champion, *The European Pulsar Timing Array and the Large European Array for Pulsars*, *Class. Quant. Grav.* **30** (2013) 224009. [Page 152]

[846] L. Lentati et al., *European Pulsar Timing Array Limits On An Isotropic Stochastic Gravitational-Wave Background*, *Mon. Not. Roy. Astron. Soc.* **453** (2015) 2576 [1504.03692]. [Page 152]

[847] S. Babak et al., *European Pulsar Timing Array Limits on Continuous Gravitational Waves from Individual Supermassive Black Hole Binaries*, *Mon. Not. Roy. Astron. Soc.* **455** (2016) 1665 [1509.02165]. [Page 152]

[848] R. N. Manchester et al., *The Parkes Pulsar Timing Array Project*, *Publications of the Astronomical Society of Australia (PASA)* **30** (2013) e017 [1210.6130]. [Page 152]

[849] R. M. Shannon et al., *Gravitational waves from binary supermassive black holes missing in pulsar observations*, *Science* **349** (2015) 1522 [1509.07320]. [Page 152]

[850] P. Amaro-Seoane et al., *Laser Interferometer Space Antenna*, 1702.00786. [Page 152]

[851] K. Schmitz, *New Sensitivity Curves for Gravitational-Wave Signals from Cosmological Phase Transitions*, *JHEP* **01** (2021) 097 [2002.04615]. [Page 152]

[852] J. Darling, A. E. Truebenbach and J. Paine, *Astrometric Limits on the Stochastic Gravitational Wave Background*, *Astrophys. J.* **861** (2018) 113 [1804.06986]. [Page 152]

[853] NANOGrav Collaboration, Z. Arzoumanian et al., *The NANOGrav 12.5 yr Data Set: Search for an Isotropic Stochastic Gravitational-wave Background*, *Astrophys. J. Lett.* **905** (2020) L34 [2009.04496]. [Page 152]

[854] K. Kohri and T. Terada, *Solar-Mass Primordial Black Holes Explain NANOGrav Hint of Gravitational Waves*, *Phys. Lett. B* **813** (2021) 136040 [2009.11853]. [Page 152]

[855] D. Carney et al., *Mechanical Quantum Sensing in the Search for Dark Matter*, *Quantum Sci. Technol.* **6** (2021) 024002 [2008.06074]. [Page 156]

[856] D. Carney et al., *Snowmass 2021: Quantum Sensors for HEP Science – Interferometers, Mechanics, Traps, and Clocks*, in *2022 Snowmass Summer Study*, 3, 2022, 2203.07250. [Page 156]

[857] Y. Kahn and T. Lin, *Searches for light dark matter using condensed matter systems*, 2108.03239. [Page 156]

[858] A. Mitridate, T. Trickle, Z. Zhang and K. M. Zurek, *Snowmass White Paper: Light Dark Matter Direct Detection at the Interface With Condensed Matter Physics*, in *2022 Snowmass Summer Study*, 3, 2022, 2203.07492. [Page 156]

[859] C. A. J. O'Hare et al., *Particle detection and tracking with DNA*, *Eur. Phys. J. C* **82** (2022) 306 [2105.11949]. [Page 156]