Stockholm University

Was Einstein Wrong?

Theoretical and observational constraints on massive gravity

Marcus Högås

Doctoral Thesis in Theoretical Physics at Stockholm University, Sweden 2022

Was Einstein Wrong?

Theoretical and observational constraints on massive gravity

Marcus Högås

Academic dissertation for the Degree of Doctor of Philosophy in Theoretical Physics at Stockholm University to be publicly defended on Friday 22 April 2022 at 10.00 in sal FB42, AlbaNova universitetscentrum, Roslagstullsbacken 21 and online via Zoom, public link is available at the department website.

Abstract

For more than a century, Einstein's theory of general relativity has described gravitational phenomena with astonishing precision. However, for the theory to fit observations we need to add two elusive substances: dark energy and dark matter. Together they add up to 95% of the energy budget of the Universe. Yet, we do not know what these substances are. Another question mark is the expansion rate of the Universe; two incompatible values are obtained depending on the measuring method. These problems (dark energy, dark matter, and the expansion rate) belong to the big questions within gravity today and they may be interpreted as signs that general relativity is not the final theory for gravity. As an alternative, in this thesis we analyze an extended theory of gravity called bimetric gravity.

In general relativity (GR), gravity is massless which means that gravitational waves propagate at the speed of light. Hence, a natural extension is to consider theories where gravity has a mass. This is precisely what bimetric gravity achieves. The theoretical consistency of this theory is firmly established but it is also crucial to test if the theory agrees with observations. In fact, in this theory there are two types of gravitational waves/fields, one massless as in GR but also one massive. When observing gravitational phenomena, we observe a mix of the two. Depending on the mixing and on the mass of the massive field, observational signatures appear for example on cosmological scales, in gravitational wave events or on solar-system scales. Until recently the phenomenology of the full theory was still uncharted, and an important question was if all observational tests could be satisfied at the same time. To address this, we devised a unified framework that enables straightforward comparison between constraints from different probes, without being restricted to a particular region of the parameter space. The result is that bimetric gravity is compatible with observations and even fit data slightly better than GR. Together with the fact that the dark energy can be explained by the interaction between the two gravitational fields, we have shown that the theory is a viable dark energy candidate. At the same time, the observational data provides a substantial restriction on the parameter space that excludes many of the popular models in the literature – an important result in and of itself.

A longstanding issue within this theory has been to predict the growth of structure while avoiding exponential instabilities. Here, we propose a simple model which solves the full, nonlinear equations of motion, which can be used to calculate the growth of structure, without any instabilities. We also describe our work towards a framework for calculating the process of gravitational collapse in this theory where we manage to solve the equations numerically for a short time interval. The results indicate that the gravitational collapse proceeds as in general relativity, assuming that the initial conditions are similar.

Future work is needed to decide whether bimetric gravity can solve any of the other big questions within gravity today, such as the discrepant expansion rate of the Universe. In this thesis, we show that it is an observationally viable dark energy candidate that exhibits novel gravitational features. In short, gravity can be massive.

Keywords: *Gravity, Massive gravity, Bimetric gravity, Cosmology, Gravitational collapse, Gravitational waves, Observational constraints.*

Stockholm 2022
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-202495

ISBN 978-91-7911-806-8
ISBN 978-91-7911-807-5

Stockholm University

Department of Physics

Stockholm University, 106 91 Stockholm

WAS EINSTEIN WRONG?

Marcus Högås

Stockholm
University

Was Einstein Wrong?

Theoretical and observational constraints on massive gravity

Marcus Högås

ISBN print 978-91-7911-806-8
ISBN PDF 978-91-7911-807-5

Cover image: *Yellow Ochre Sky with Black Hole Over Wheat Field*, by Gregory Horndeski

Typeset in pdfLATEX

Printed in Sweden by Universitetsservice US-AB, Stockholm 2022

To my family

The heavens declare the glory of God;
the skies proclaim the work of his hands.

Psalm 19:1, New International Version,
attributed to King David c. 1000 B.C.

Abstract

For more than a century, Einstein's theory of general relativity has described gravitational phenomena with astonishing precision. However, for the theory to fit observations we need to add two elusive substances: dark energy and dark matter. Together they add up to 95 % of the energy budget of the Universe. Yet, we do not know what these substances are. Another question mark is the expansion rate of the Universe; two incompatible values are obtained depending on the measuring method. These problems (dark energy, dark matter, and the expansion rate) belong to the big questions within gravity today and they may be interpreted as signs that general relativity is not the final theory for gravity. As an alternative, in this thesis we analyze an extended theory of gravity called bimetric gravity (a.k.a. massive gravity).

In general relativity (GR), gravity is massless which means that gravitational waves propagate at the speed of light. Hence, a natural extension is to consider theories where gravity has a mass. This is precisely what bimetric gravity achieves. The theoretical consistency of this theory is firmly established but it is also crucial to test if the theory agrees with observations. In fact, in this theory there are two types of gravitational waves/fields, one massless as in GR but also one massive. When observing gravitational phenomena, we observe a mix of the two. Depending on the mixing and on the mass of the massive field, observational signatures appear for example on cosmological scales, in gravitational wave events or on solar-system scales. Until recently the phenomenology of the full theory was still uncharted, and an important question was if all observational tests could be satisfied at the same time. To address this, we devised a unified framework that enables straightforward comparison between constraints from different probes, without being restricted to a particular region of the parameter space. The result is that bimetric gravity is compatible with observations and even fit data slightly better than GR. Together with the fact that the dark energy can be explained by the interaction between the two gravitational fields, we have shown that the theory is a viable dark energy candidate. At the same time, the observational data provides a substantial restriction on the parameter space that excludes many of the popular models in the literature – an important result in and of itself.

A longstanding issue within this theory has been to predict the growth of structure while avoiding exponential instabilities. Here, we propose a simple model which solves the full, nonlinear equations of motion, which can be used to calculate the growth of structure, without any instabilities. We also describe our work towards a framework for calculating the process of gravitational collapse in this theory where we manage to solve the equations numerically for a short time interval. The results indicate that the gravitational collapse proceeds as in general relativity, assuming that the initial conditions are similar.

Future work is needed to decide whether bimetric gravity can solve any of the other big questions within gravity today, such as the discrepant expansion rate of the Universe. In this thesis, we show that it is an observationally viable dark energy candidate that exhibits novel gravitational features. In short, gravity can be massive.

Svensk sammanfattning

I mer än ett sekel har Einsteins allmänna relativitetsteori varit den teori som bäst beskriver gravitationen. Hans teori överensstämmer med observationer till en makalös precision. Detta under förutsättning att vi introducerar två mystiska substanser: mörk energi och mörk materia, som tillsammans utgör 95 % av universums energiinnehåll. För närvarande känner vi till mycket litet om dessa substanser. Ytterligare ett frågetecken är det uppmätta värdet av universums expansionshastighet. Beroende på vilken mätmetod som används, erhålls två olika, och motsägelsefulla, värden. Dessa gåtor (mörk energi, mörk materia och universums expansionshastighet) hör till de stora frågorna inom gravitation idag och kan tyda på att den allmänna relativitetsteorin trots allt inte är den slutgiltiga teorin för gravitationen. Som ett alternativ studerar vi i denna avhandling en generaliserad teori som kallas bimetrisk gravitation.

I Einsteins teori är gravitationen masslös, vilket innebär att gravitationsvågor fortplantas med ljusets hastighet. En naturlig generalisering är att introducera en massa till gravitationsfältet, vilket är precis vad bimetrisk gravitation gör. Att bimetrisk gravitation är teoretiskt konsistent är ett väletablerat faktum men det är också av yttersta vikt att undersöka om teorin stämmer överens med observationer. Faktum är att det finns två typer av gravitationsfält/vågor i denna teori: en masslös, som i Einsteins teori, och en massiv. En observatör ser en blandning av dessa två. Nya fenomen dyker till exempel upp på kosmologiska längdskalor, för gravitationsvågor eller i solsystemstest beroende på hur stor massa gravitationen har och på hur stor del av det totala gravitationsfältet som utgörs av det massiva fältet. Fram till nyligen var det oklart hur olika gravitationsfenomen ter sig för allmänna modeller i denna teori och om det finns någon modell som kan uppfylla alla observationella gravitationstest på en och samma gång. För att klargöra detta konstruerade vi ett enhetligt ramverk som gör det möjligt att, på ett enkelt sätt, jämföra resultat mellan olika typer av test utan att begränsa sig till en specifik bimetrisk modell. Slutsatsen är att bimetrisk gravitation, eller massiv gravitation som vi också kallar det, stämmer väl överens med observationer, till och med något bättre än Einsteins allmänna relativitetsteori. Tillsammans med det faktum att den mörka energin kan förklaras av interaktionen mellan de två

gravitationsfälten så visar vi att denna teori ger en möjlig förklaring till vad mörk energi är. Samtidigt begränsar dessa observationer det möjliga parameterutrymmet, vilket leder till att många populära modeller som har studerats tidigare i litteraturen måste uteslutas – ett viktigt resultat i sig.

En utmaning inom bimetrisk gravitation är att förklara hur universums nuvarande struktur har uppkommit till följd av de ursprungliga krusningarna i det tidiga universum. Problemet är att linjär störningsteori, som används inom den kosmologiska standardmodellen, ofta är oanvändbar på grund av en exponentiell instabilitet. Här presenterar vi en enkel modell som löser de fullständiga ickelinjära rörelseekvationerna och som kan användas för att modellera strukturformation utan någon instabilitet. Vi presenterar också vårt arbete med att analysera gravitationskollaps i denna teori, där vi lyckas lösa ekvationerna numeriskt under ett litet tidsintervall. Resultaten tyder på att gravitationskollaps i denna teori skiljer sig mycket litet från den allmänna relativitetsteorin, förutsatt att initialvillkoren liknar varandra.

Ytterligare insatser krävs för att avgöra huruvida bimetrisk gravitation kan lösa någon av de andra stora frågorna inom gravitation idag, som till exempel universums motsägelsefulla expansionshastighet. Denna avhandling visar att teorin stämmer väl överens med observationer och är en möjlig kandidat till mörk energi. Kort sagt, gravitationen kan vara massiv.

List of papers

The following papers are included in the thesis. They are referred to by their Roman numerals in the text.

I F. Torsello, M. Kocic, **M. Högås**, and E. Mörtsell, *Spacetime symmetries and topology in bimetric relativity, Phys. Rev. D* **97** (2018) 084022, [1710.06434]

II **M. Högås**, F. Torsello, E. Mörtsell, *On the stability of bimetric structure formation, JCAP* **04** (2020) 046, [1910.01651]

III **M. Högås** and E. Mörtsell, *Constraints on bimetric gravity. Part I. Analytical constraints, JCAP* **05** (2021) 001, [2101.08794]

IV **M. Högås** and E. Mörtsell, *Constraints on bimetric gravity. Part II. Observational constraints, JCAP* **05** (2021) 002, [2101.08795]

V **M. Högås** and E. Mörtsell, *Constraints on bimetric gravity from Big Bang nucleosynthesis, JCAP* **11** (2021) 001, [2106.09030]

VI **M. Högås**, M. Kocic, F. Torsello, and E. Mörtsell, *Generalized Vaidya solutions in bimetric gravity, Class. Quant. Grav.* **37** (2020) 145010, [1905.09832]

VII M. Kocic, F. Torsello, **M. Högås**, and E. Mörtsell, *Initial data and first evolutions of dust clouds in bimetric relativity, Class. Quant. Grav.* **37** (2020) 165010, [1904.08617]

The following papers are complementary and not included in the thesis. They are quoted as ordinary references in the main text.

VIII M. Kocic, **M. Högås**, F. Torsello, and E. Mörtsell, *Algebraic Properties of Einstein Solutions in Ghost-Free Bimetric Theory, J. Math. Phys.* **60** (2019) 102501, [1706.00787]

IX M. Kocic, **M. Högås**, F. Torsello, and E. Mörtsell, *On Birkhoff's theorem in ghost-free bimetric theory*, [1708.07833]

X F. Torsello, M. Kocic, **M. Högås**, and E. Mörtsell, *Covariant BSSN formulation in bimetric relativity, Class. Quant. Grav.* **37** (2020) 025013, [1904.07869]. [Erratum: Class.Quant.Grav. 37, 079501 (2020)]

XI S. Castello, **M. Högås**, and E. Mörtsell, *A Cosmological Underdensity Does Not Solve the Hubble Tension*, [2110.04226] (under review)

Contribution to papers

The author order is in decreasing order of contribution, see the list of papers above. Accordingly, the first author made the most significant contribution to the paper and so on.

Paper I. I participated in the discussions and calculations leading to the results. Specifically, I developed the methods of generating the solutions in Section 2.B.3, A.1, and A.2. These sections were also written by me.

Paper II. I had the original idea to use this setup as a model of structure formation in bimetric gravity. I performed all calculations leading to the results and wrote the paper.

Paper III. I initiated the project. Together with the second author, I devised the framework of the physical parameters which is promoted in this paper and used in all succeeding papers (also based on work by previous authors). I did all the calculations and wrote the paper.

Paper IV. This was the natural follow-up paper of Paper III. With guidance from the second author, I confronted the theory (bimetric gravity) with observational tests of gravity. The actual calculations, coding, and statistical data analysis were done by me. I also wrote the paper.

Paper V. While working on this thesis, I got the idea of using Big Bang nucleosynthesis to constrain the theory. I also found out how to use the available data to readily constrain the parameter space. The calculations, coding, and writing the paper is due to me.

Paper VI. I set up the model, derived and analyzed the results and wrote and edited the paper with additional input from the coauthors.

Paper VII. I checked the solutions for the initial data independently and showed that bidiagonality of the metrics cannot be maintained dynamically. I was involved in the discussions of the results and in the writing process.

Acknowledgements

My deepest gratitude goes to Edvard Mörtsell. Thank you for always being available and supportive and for continuously inspiring me to dig deeper into the astounding phenomenon that we call gravity. I cannot think of a better supervisor. I also want to thank all my colleagues. The present work would not have been possible without you. Special thanks to Joakim Flinckman, Sten Hellman, Sören Holst, Marvin Lüben, Edvard Mörtsell, Igor Pikovski, and Francesco Torsello who have commented the thesis. The thesis is typeset in LaTeX and based on a template created by Mikica Kocic

Preface

This doctoral thesis consists of two parts. The first includes Chapters 1-8. Chapters 1-2 introduce the topic. In Chapters 2-4 and Chapter 6-7, we provide a context to and present some of the results obtained during the author's PhD studies, that is, Refs. [1–12]. Chapter 5 is based on a master thesis project by Erik Schildt which the author was co-supervising together with Edvard Mörtsell. Most of the results in this chapter can be found in the literature although the stability analysis in Section 5.2 is, to the author's knowledge, presented in a novel way. Section 5.3 follows Ref. [13] to large extent, although the analysis of the propagation speed of gravitational waves and gravitational wave echoes was not studied in this reference and we parameterize the theory in terms of the physical parameters rather than the usual β-parameters. Towards the end of Chapters 3-7, we provide a summary of the most important results. The busy reader may want to concentrate on these. The second part of this thesis consists of the corresponding scientific papers.

Tensors are denoted by their components, for example the (physical) metric is referred to as $g_{\mu\nu}$. Space-time indices are denoted by Greek letters and run from zero to three. Spatial components are denoted by Latin indices and run from one to three. We use the convention where the Minkowski metric reads $\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$. Quantities constructed from the second metric $f_{\mu\nu}$ are denoted by a tilde. For example, $\widetilde{G}^{\mu}{}_{\nu}$ is the Einstein tensor of the second metric. Unless stated otherwise, we use geometrized units, setting Newton's gravitational constant and the speed of light to unity ($G = c = 1$), in which case length, time, and mass have the same units, $\mathsf{L} = \mathsf{T} = \mathsf{M}$. Occasionally, we state the result in so-called natural units where the reduced Planck constant and the speed of light are set to unity ($\hbar = c = 1$). In this case, $\mathsf{L} = \mathsf{T} = \mathsf{M}^{-1}$ and we give the mass in units of eV.

The main subject of this thesis is a massive gravity theory where both metrics are dynamical. We will refer to this as massive gravity, massive bigravity, bimetric relativity (BR), bimetric gravity, or bigravity. The massive gravity theory where one of the metrics is nondynamical is referred to as dRGT massive gravity which is a special limiting case of BR.

Material from the licentiate thesis. Parts of Section 1.2 are based on Section 1.2 of the licentiate thesis [14]. The "Quantization of gravity" paragraph in Section 1.4.2 is similar to the same paragraph in the licentiate thesis. Section 2.3 is based on material from Chapter 2 of Ref. [14]. The section is rewritten slightly and Fig. 2.1 is new. Section 2.4 is largely copied from the licentiate thesis. Section 7.1 have similarities with Section 2.8 of the licentiate thesis. Section 7.3 and Section 7.4 share most content with Section 2.4 and Section 2.6 of Ref. [14], respectively. Fig 7.5 is new. Appendices B-C are copied from the licentiate thesis with only slight modifications.

New results. Some of the results in the thesis cannot be found in the published literature. This includes the stability analysis in Section 5.2 which is presented in a novel way, the grid scan related to the excision of the gradient instability in Section 6.3 and the application of the nonlinear structure formation model to the best-fit cosmologies in the same section, as well as the analyzes in Section 7.1-7.2.

Abbreviations and symbols

Abbreviations

AdS	anti-de Sitter
AIC	Akaike information criterion
BBN	Big Bang nucleosynthesis
BIC	Bayesian information criterion
BR	bimetric relativity. This is the main subject of the thesis and is also referred to as bimetric gravity, bigravity, massive bigravity, massive gravity
CDM	cold dark matter
CMB	cosmic microwave background (radiation)
COBE	Cosmic Background Explorer
D	deuterium
DE	dark energy
DGP	Dvali–Gabadadze–Porrati
dRGT	de Rham–Gabadadze–Tolley
dS	de Sitter
EEP	Einstein equivalence principle
EH	event horizon
FLRW	Friedmann–Lemaître–Robertson–Walker
FP	Fierz–Pauli
FRB	fast radio burst
GR	general relativity
GW	gravitational wave
h.o.	higher orders
HR	Hassan–Rosen

KVF	Killing vector field
LIGO	Laser Interferometer Gravitational-Wave Observatory
LTB	Lemaître–Tolman–Bondi
MOND	modified Newtonian dynamics
OS	Oppenheimer–Snyder
RNG	radial null geodesic
SAdS	Schwarzschild–anti-de Sitter
SdS	Schwarzschild–de Sitter
SNIa	supernova(e) type Ia
SS	spherically symmetric
SSS	static, spherically symmetric
WEP	weak equivalence principle
ΛCDM	Λ cold dark matter

Units

AU	astronomical unit $= 1.50 \times 10^{11}\,\mathrm{m}$
eV	electronvolt $= 1.60 \times 10^{-19}\,\mathrm{J}$
pc	parsec $= 3.26$ light-years $= 3.09 \times 10^{16}\,\mathrm{m}$

Operators

$\sim$	same order of magnitude
$\simeq$	approximately equal to
$\equiv$	identically equal to (by definition)
$\propto$	proportional to
∇_μ	covariant derivative
∂_μ	partial derivative $\partial/\partial x^\mu$
∇^2	Laplacian $\nabla^i\nabla_i$
$\Box$	wave operator/d'Alembertian $g^{\mu\nu}\nabla_\mu\nabla_\nu$
$\dot{}$	overdot denotes a time derivative, e.g. $\dot{f} \equiv df/dt$
$'$	in Chapters 5-6, prime denotes derivative with respect to conformal time $d/d\tau$ and in Chapter 3, prime denotes the derivative $d/d\ln a$

Symbols

a	scale factor of the physical metric
$\widetilde{a}$	scale factor of the second metric
c_s	sound speed
$\cot\theta$	$\equiv 1/\tan\theta$
$\csc\theta$	$\equiv 1/\sin\theta$
d_A	angular diameter distance
D_A	comoving angular diameter distance
d_L	luminosity distance
$\mathcal{D}_L$	dimensionless luminosity distance $\mathcal{D}_L \equiv H_0 d_L$
$d_{\rm phys}$	proper (physical) size/distance
D_V	volume average distance
$\mathrm{diag}(\cdot,\cdot,\cdot,\cdot)$	components of a diagonal matrix
det	determinant
$d\Omega^2$	line element on the unit 2-sphere, $d\Omega^2 \equiv d\theta^2 + \sin^2\theta\, d\phi^2$
E	normalized expansion rate $E \equiv H/H_0$
f_g	growth rate $f_g \equiv d\,\ln\delta/d\,\ln a$
$f_{\mu\nu}$	second metric
$g_{\mu\nu}$	physical metric
$G_{\mu\nu}$	Einstein tensor
h	$h \equiv H_0/(100\,\mathrm{km/s/Mpc})$
H	Hubble parameter $H \equiv \dot{a}/a = a^{-1}da/dt$
H_0	Hubble constant $H_0 = H(z=0) \simeq 70\,\mathrm{km/s/Mpc}$, or in natural units $H_0 \simeq 10^{-33}\,\mathrm{eV}$
$\mathcal{H}$	conformal Hubble parameter $\mathcal{H} \equiv a^{-1}da/d\tau$
K_{ij}	extrinsic curvature
$\mathcal{L}$	likelihood, $\mathcal{L} \equiv e^{-\chi^2/2}$
$m_{\rm FP}$	in natural units: graviton mass/Fierz–Pauli mass, measured in units of H_0
$\hat{m}_{\rm FP}$	(dimensionful) graviton mass
$M_\odot$	mass of the Sun ($\simeq 2.0 \times 10^{30}\,\mathrm{kg}$)
r_H	Hubble radius $r_H \equiv 1/H_0 \simeq 5\,\mathrm{Gpc} \sim 10^{26}\,\mathrm{m}$
r_s	comoving sound horizon
r_S	Schwarzschild radius, $r_S \equiv 2M$

R	Ricci curvature scalar
$R_{\mu\nu}$	Ricci tensor
$R^{\mu}{}_{\nu\rho\sigma}$	Riemann tensor
Tr	trace
S	modified expansion rate $S \equiv H/H_{\rm GR}$
$\sec\theta$	$\equiv 1/\cos\theta$
x	$x \equiv (d\widetilde{a}/d\tau)/(da/d\tau)$
X	$X = x/y$
y	ratio of scale factors $y \equiv \widetilde{a}/a$
Y_p	^{4}He abundance
z	redshift $z \equiv a_0/a - 1$
z_d	redshift of the baryon drag epoch $z_d \simeq 1060$
$z_{\rm eq}$	redshift at matter-radiation equality
z'	redshift before which there is a gradient instability at the linear level
z_*	redshift of photon decoupling $z_* \simeq 1090$
γ	growth index/gravitational slip, depending on the context
γ_{ij}	induced metric
δ	density contrast $\delta \equiv (\rho_m - \bar{\rho}_m)/\bar{\rho}_m$
δ_{ij}	Kronecker delta
ϵ_{ijk}	Levi-Civita symbol
$\eta_{\mu\nu}$	Minkowski metric, i.e. diag$(-1, 1, 1, 1)$
θ	mixing angle between the massless and massive gravitons
ϑ	volume expansion rate
Θ	list of the physical parameters, $\Theta \equiv (\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta)$
κ	ratio of the gravitational constants $\kappa \equiv \kappa_g/\kappa_f$
κ_f	gravitational constant of the f-metric
κ_g	gravitational constant of the g-metric
λ_g	Compton wavelength of the massive graviton
Λ	(effective) cosmological constant
ρ_c	critical energy density of the Universe today, $\rho_c \equiv 3H_0^2/\kappa_g$
ρ_i	energy density of species i
τ	denotes proper time in Chapter 1 and conformal time in the remaining chapters

Φ	(Newtonian) gravitational potential
Ψ	scalar curvature
Ω_i	energy density in units of the critical energy density today, $\Omega_i \equiv \rho_i/\rho_c$

Contents

Chapter 1

Introduction

1.1 What is gravity?

More than three centuries ago, Newton proposed a grand idea, namely that the same force is responsible for the motion of the celestial bodies as well as the motion of an apple falling to the ground. The signature of this force is that it affects all objects equally. For example, neglecting air resistance, all objects fall towards the Earth at the same rate. This is known as the weak equivalence principle and it has been confirmed experimentally to at least thirteen digits precision [15]. In contrast, the other fundamental forces of nature affect different particles in different ways. For example, a proton and an electron accelerate in opposite directions in the same electric field. Also, the other forces are either short-ranged or cancel out on astronomical length scales, so the net charge of a planet or a star is effectively zero. Gravity is different. The gravitational charge is mass, which is always positive. So, although the gravitational force is weak compared with the other fundamental forces, astronomical objects accumulate a lot of mass. Hence, gravity plays the main role in the large-scale evolution and fate of our universe.

In the beginning of the 20th century, Einstein proposed general relativity (GR) as a theory for gravity, providing important corrections to Newton's law of universal gravitation. Together with quantum field theory (describing the other fundamental forces), GR is the cornerstone of modern physics. A major achievement of these two theories is that they lead to the cosmological concordance model, which successfully describes the evolution of our universe from a few fractions of a second after the Big Bang to present day. But there are some caveats. In the cosmological concordance model, the main content of the Universe consists of largely unknown substances, known as dark matter and dark energy, whose properties we can only infer indirectly. Dark

energy is required to explain the accelerated expansion of the Universe [16,17]. Within general relativity, this effect is attributed to the cosmological constant. However, its value must be fine-tuned to at least 60 digits to almost exactly cancel the contributions from vacuum energy and thus give rise to the observed expansion. Dark matter is normally assumed to be a, yet-to-be discovered, invisible particle whose presence explains why galaxies are not ripped apart despite the rapid motion of their visible components such as stars and gas. The problem is that dark matter has never been observed directly. Moreover, improved cosmological observations within the last decade have revealed an internal crack in the cosmological concordance model. Namely, measuring the current expansion rate of the Universe (the Hubble constant) using probes in the nearby universe or at large distances yields incompatible results [18]. This is called the Hubble tension.

A possible solution to one or several of these issues is that general relativity is not the ultimate theory for gravity, hence the main question of the thesis: Was Einstein wrong? In general relativity, gravity is massless, meaning that gravitational waves travel at the speed of light. A natural modification to Einstein's theory is to give gravity a mass. A first step was taken already in 1939 by Fierz and Pauli [19] but the final form of this theory was not found until 2010 by de Rham, Gabadadze, Tolley, Hassan, and Rosen [20–23]. This class of massive gravity theories is the main subject of this thesis.

Any new theory for gravity must pass elementary observational tests. In this thesis we expose massive gravity to observations probing the theory on cosmological scales as well as on galactic and solar-system scales. We also analyze the behavior of gravitational waves and gravitational collapse in this theory. We find that the theory is viable and has interesting implications for the nature of dark energy and dark matter as well as the propagation of gravitational waves, and, possibly, the Hubble tension.

1.2 Gravity as geometry

The inertial mass, $m_{\rm in}$, appears in Newton's second law and is the resistance of a body towards changing its velocity, as a certain force is applied. The gravitational mass, $m_{\rm gr}$, appears in Newton's law of universal gravitation and is the gravitational charge of a body, determining how strongly it couples to gravity. An empirical fact is that the gravitational mass and the inertial mass take the same value, $m_{\rm gr} = m_{\rm in}$, for every single particle in the whole Universe, which is by no means obvious *a priori*. A consequence of this is that freely falling particles follow the same trajectories (given the same initial conditions). This is the weak equivalence principle (WEP). The principle may

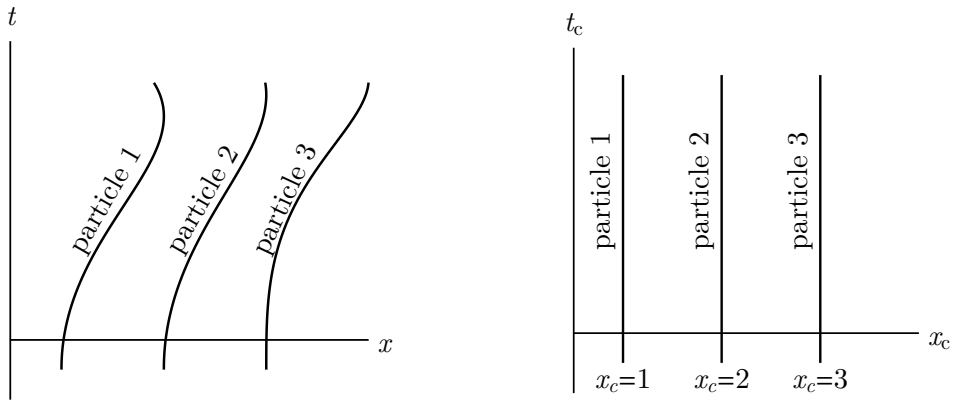


Figure 1.1: *Left*: Freely falling particles in a generic coordinate system. Each particle trajectory is labeled by a number. Since the trajectories are non-intersecting (at least locally), we can define a new coordinate system where the time coordinate is the proper time along the trajectories $t_c = \tau$ and space is labeled by the trajectories. *Right*: Freely falling particles in comoving coordinates.

appear miraculous when treating gravity as a force, but changing perspective and treating gravity as a manifestation of space-time geometry makes the principle inevitable. This is what we will do now.

Let us imagine a (small) region filled with hypothetical, freely falling particles. In a general coordinate system, the particles follow curved paths as illustrated in the left panel of Fig. 1.1. However, we can define a comoving coordinate system $x_{\rm c}^\mu$ where time is labeled by the proper time τ along each trajectory and space is labeled by the trajectories as in the right panel of Fig. 1.1. Which coordinates that are comoving depends on the gravitational field at hand, hence $x_{\rm c}^\mu$ contains information about gravity. In these coordinates, any freely falling body tracing out the trajectory $x_{\rm c}^\mu(\tau)$, remains at the same spatial coordinate. Hence,

$$\frac{d^2 x_{\rm c}^\mu(\tau)}{d\tau^2} = 0. \tag{1.1}$$

This equation can be transformed to some other (generic) coordinate system, x^μ, using the chain rule,

$$\frac{\partial x_{\rm c}^\mu}{\partial x^\nu}\frac{d^2 x^\nu}{d\tau^2} + \frac{\partial^2 x_{\rm c}^\mu}{\partial x^\nu \partial x^\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} = 0. \tag{1.2}$$

With some manipulations we get the geodesic equation [24],

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu{}_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} = 0, \tag{1.3}$$

where the connection $\Gamma^\mu{}_{\nu\rho}$ is defined by,

$$\Gamma^\mu{}_{\nu\rho} \equiv \frac{\partial x^\mu}{\partial x^\sigma_{\rm c}}\frac{\partial^2 x^\sigma_{\rm c}}{\partial x^\nu \partial x^\rho}. \tag{1.4}$$

Note that the form of the geodesic equation depends on the connection only. So, eq. (1.3) is independent of the internal decomposition of the freely falling particles. In this perspective, the uniform motion of freely falling particles (the WEP) is due to them moving in the same space-time geometry, as encoded in the connection.

Note that $\Gamma^\mu{}_{\nu\rho}$ contains information both about the coordinates x^μ and about gravity (via the comoving coordinates $x^\mu_{\rm c}$). See below for two examples. Using $\Gamma^\mu{}_{\nu\rho}$, we can define the Riemann tensor, which is a geometric object quantifying the space-time curvature,

$$R^\mu{}_{\nu\rho\sigma} \equiv \partial_\rho \Gamma^\mu{}_{\sigma\nu} - \partial_\sigma \Gamma^\mu{}_{\rho\nu} + \Gamma^\mu{}_{\rho\lambda}\Gamma^\lambda{}_{\sigma\nu} - \Gamma^\mu{}_{\sigma\lambda}\Gamma^\lambda{}_{\rho\nu}. \tag{1.5}$$

The Ricci tensor is defined as $R_{\mu\nu} \equiv R^\rho{}_{\mu\rho\nu}$ and the Ricci curvature scalar as $R \equiv g^{\mu\nu}R_{\mu\nu}$. The covariant derivative ∇_μ specifies how vectors are parallel transported. A vector V^μ is parallel transported along itself if,

$$V^\mu \nabla_\mu V^\nu \equiv V^\mu \partial_\mu V^\nu + \Gamma^\nu{}_{\mu\rho} V^\mu V^\rho = 0. \tag{1.6}$$

Identifying V^μ as the tangent vector of a curve, $V^\mu = dx^\mu/d\tau$, this defines the notion of a straight line (a geodesic). In fact, $V^\mu \nabla_\mu V^\nu = 0$ is equivalent to the geodesic equation (1.3).

Note that our geometrization of gravity assumed only the WEP and since this principle holds both in Newtonian gravity and in general relativity, they can both be formulated in geometric terms. The geometrical formulation of Newtonian gravity is referred to as Newton–Cartan theory [25–27].

Example 1. In Newtonian gravity the connection can be written in terms of a scalar field Φ [28] (the gravitational potential),

$$\Gamma^0{}_{\nu\rho} = \Gamma^i{}_{jk} = \Gamma^i{}_{j0} = 0, \quad \Gamma^i{}_{00} = \frac{\partial \Phi}{\partial x^i}, \tag{1.7}$$

and the geodesic equation takes the form,

$$\frac{d^2x^i}{d\tau^2} = -\frac{\partial \Phi}{\partial x^i}, \tag{1.8}$$

which we recognize as Newton's second law for a particle in a gravitational potential Φ. Here, the Riemann tensor $R^{\mu}{}_{\nu\rho\sigma}(\Gamma)$ does not vanish, hence describing a curved geometry.

□

Example 2. Assuming that there is no gravity present ($R^{\mu}{}_{\nu\rho\sigma} = 0$), the comoving coordinates coincide with the inertial coordinates. Let us write the geodesic equation (1.3) in rotating coordinates, (x, y),

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\omega t & \sin\omega t \\ -\sin\omega t & \cos\omega t \end{pmatrix} \begin{pmatrix} x_{\rm c} \\ y_{\rm c} \end{pmatrix}. \tag{1.9}$$

Calculating $\Gamma^{\mu}{}_{\nu\rho}$ using eq. (1.4), we find,

$$\Gamma^{0}{}_{\nu\rho} = 0, \quad \Gamma^{i}{}_{00} = -\omega^2 x^i, \quad \Gamma^{i}{}_{j0} = \omega\epsilon_{ij3}, \quad \Gamma^{i}{}_{jk} = 0, \tag{1.10}$$

and the geodesic equation reads,

$$\ddot{x}^i = \omega^2 x^i + 2\epsilon_{ijk}\omega^j \dot{x}^k \quad \Leftrightarrow \quad m\mathbf{a} = m\omega^2\mathbf{x} + 2m\,\boldsymbol{\omega}\times\mathbf{v}, \tag{1.11}$$

where we defined $\omega_i = (0, 0, \omega)$ and ϵ_{ijk} denotes the Levi-Civita symbol. This is Newton's second law in rotating coordinates. The Riemann tensor of this connection is zero, so there is no curvature, as expected. However, since we used curvilinear coordinates, inertial (fictitious) forces appear on the right-hand sides of eq. (1.11) as encoded in $\Gamma^{\mu}{}_{\nu\rho}$. The first of these is the centrifugal force and the second is the Coriolis force.

□

Based on the WEP, we introduced the connection and related geometric objects like the Riemann tensor and the covariant derivative. To introduce the metric, we invoke the Einstein equivalence principle (EEP) stating that an observer in free fall measures the same physics, locally, as in the absence of gravity (i.e., as in special relativity). According to Einstein himself, this was his most fortunate thought [29]. The EEP implies that, at an arbitrarily chosen point, p, we can construct a coordinate system x^{μ}_{p} such that special relativity is recovered in the close vicinity of p. That is, space-time appears locally flat to an observer so that $\nabla_{\mu}|_{p} = \partial_{\mu}$ and the proper time $d\tau$ between two infinitesimally close events, is given by,

$$d\tau^2 = -\eta_{\mu\nu} dx^{\mu}_{p} dx^{\nu}_{p}, \tag{1.12}$$

where $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ is the Minkowski metric and dx^{μ}_{p} is the coordinate distance between the points. Transforming to some generic coordinates x^{μ}, we

obtain,

$$d\tau^2 = -g_{\mu\nu}dx^\mu dx^\nu, \quad g_{\mu\nu} \equiv \frac{\partial x_p^\rho}{\partial x^\mu}\frac{\partial x_p^\sigma}{\partial x^\nu}\eta_{\rho\sigma}, \tag{1.13}$$

where we have defined the metric $g_{\mu\nu}$. At a point p we can use the EEP to identify the comoving coordinates of an observer, x_c^μ, with the locally flat coordinates x_p^μ. From eqs. (1.4) and (1.13) it then follows [24],

$$\Gamma^\mu{}_{\nu\rho} = \frac{1}{2}g^{\mu\sigma}\left(\partial_\nu g_{\rho\sigma} + \partial_\rho g_{\nu\sigma} - \partial_\sigma g_{\nu\rho}\right). \tag{1.14}$$

A freely falling observer making a local measurement finds that the laws of physics conform to special relativity. To get a valid description throughout the whole space-time (not just locally), we replace $\eta_{\mu\nu} \to g_{\mu\nu}$ and $\partial_\mu \to \nabla_\mu$. This recipe is called the minimal coupling prescription. As an example, we obtain the Lagrangian for a massless scalar field ϕ in a curved space-time from flat space-time by making the replacement: $-\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \to -g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi$.

In the words of Wheeler, "Spacetime tells matter how to move; matter tells spacetime how to curve" [30]. We have explored the first part of this statement. The next step is to specify how to determine the metric (and hence the connection), given some matter content.

1.3 General relativity

As shown by Lovelock in 1971, it is possible to derive the Einstein equations (in vacuum) with only a small number of assumptions [31]. Namely, assuming that the equations of motion can be derived from an action principle where the Lagrangian is a scalar density of the form $\mathcal{L}(g_{\mu\nu})$ (i.e., depending on the metric only), the only local second-order equations of motion in four dimensions are,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 0, \tag{1.15}$$

where $G_{\mu\nu} \equiv R_{\mu\nu} - g_{\mu\nu}R/2$ is the Einstein tensor and Λ is the cosmological constant. These are the (vacuum) equations of motion for the metric in general relativity. Expanding eq. (1.15) to linear order around a flat background solution, that is $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $h_{\mu\nu}$ small, the Einstein equations can be written as a massless wave equation,

$$\Box h_{\mu\nu} = 0. \tag{1.16}$$

Hence, in GR gravity is massless, propagating with the speed of light. Without imposing any constraints, $h_{\mu\nu}$ has ten independent components in four

dimensions. However, the freedom of choosing coordinates (plus the associated constraints) can be used to eliminate eight of these, resulting in an effective number of two propagating modes, corresponding to the two polarization modes of a massless spin-2 field.

The action giving rise to the Einstein equations is the Einstein–Hilbert action,

$$\mathcal{S}_{\text{EH}} = \int d^4x \sqrt{-\det g} \left[\frac{1}{2\kappa} (R - 2\Lambda) + \mathcal{L}_m \right], \tag{1.17}$$

where we have included a matter Lagrangian, $\mathcal{L}_m$. Varying eq. (1.17) with respect to the metric, we get the Einstein equations (now with matter),

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}. \tag{1.18}$$

The stress–energy tensor is the variation of the matter Lagrangian with respect to the metric, $T_{\mu\nu} \equiv -2/\sqrt{-\det g}\, \partial(\sqrt{-\det g}\mathcal{L}_m)/\partial g^{\mu\nu}$. To reproduce Newtonian gravity in the weak-field limit, we set $\kappa = 8\pi G/c^4 = 8\pi$ where G is Newton's gravitational constant and c is the speed of light.

General relativity is a remarkably simple theory in the sense that there are only two free parameters, the cosmological constant, Λ, and the Einstein gravitational constant, κ (henceforth called the gravitational constant). However, this does not mean that it is straightforward to solve the Einstein equations, on the contrary, they constitute ten coupled, non-linear, partial differential equations in four independent variables. Accordingly, there are only a few analytical solutions, valid under special symmetries. In general, the equations must be solved numerically, in itself a nontrivial task [32].

1.4 The status of general relativity

1.4.1 Successes

For more than a century, Einstein's theory of general relativity has been remarkably successful in explaining and predicting observations of gravitational phenomena ranging from sub-millimeter scales to the length scale of the observable universe (10^{26} m) [33]. On solar-system scales ($\sim 1\,\text{AU} \sim 10^{11}$ m), we can test the laws of gravity by observing the motion of the planets and the bending of light around massive sources. In general relativity, the Schwarzschild metric describes the gravitational field around a static, spherically symmetric source like the Sun. This solution predicts a small correction to Kepler's law of elliptical orbits. Applied to the motion of Mercury, this results in a shift of the perihelion of about $43''$/century which was observed already in the mid 19th century, before the invention of general relativity. Explaining this fact

was one of the first observational successes of Einstein's theory. Concerning the motion of light, Newtonian gravity and general relativity both predict that light rays bend when passing massive sources. However, there is a factor of two difference between the predicted deflection angle, with GR forecasting $\simeq 1.7''$ for light passing just outside the surface of the Sun. This was subsequently confirmed by Eddington's famous observations during the 1919 solar eclipse. As we will see in Chapter 4, there is a screening mechanism which can make massive gravity agree with general relativity in the solar system.

A consequence of GR is the existence of black holes. These are regions where gravity is so strong that nothing can escape, not even light. When a star runs out of fuel, it cannot support its own weight and undergoes a series of transformations. If the mass of the remnant exceeds the Chandrasekhar limit, $1.4M_\odot$, it can collapse into a black hole. The final black hole is characterized by its mass and angular momentum. (In principle, it can also have electric charge but we expect typically a net charge of zero.) Such a space-time is described by the Kerr metric. In 2019, the Event Horizon Telescope captured the first image of the shadow of a supermassive black hole ($\sim 10^{10}M_\odot$) in the galaxy M87. This observation is consistent with the Kerr metric as predicted by GR [34]. In massive gravity there are both Schwarzschild solutions and Kerr solutions. The question is if and how they can form from reasonable initial conditions, which is the topic of Chapter 7.

In general relativity, binary systems emit gravitational waves (GWs) carrying energy away from the system. Hence, the orbital period decreases, which was confirmed by the observation of the Hulse–Taylor binary system in 1974 where a neutron star and a pulsar orbit their common center of mass. Binary systems can merge and then emit gravitational waves that can be observed directly by gravitational wave interferometers on the Earth. The first such detection was made in 2015 by the LIGO collaboration, measuring a signal from two merging black holes [35]. In 2017, LIGO/Virgo observed the GW signal from the merger of a binary neutron star system and at the same time, an electromagnetic signal from this merger was detected [36, 37]. This counterpart signal was detected with a time delay of 1.7 s, setting a tight limit on the speed of GWs, which can differ from the speed of light by at most one part in 10^{15} [38]. This imposes strong constraints on theories predicting an anomalous speed of GWs. In the general class of massive gravity theories that we study in this thesis, there is not only a massive graviton but also a massless graviton propagating at the speed of light. This makes it possible to satisfy GW observations without severely restricting the parameter space of the theory, see Chapter 5. A theoretical success of general relativity is the possibility of setting up the equations of motion in a way which is suitable for solving numerically. This is called numerical relativity and is used to calculate and

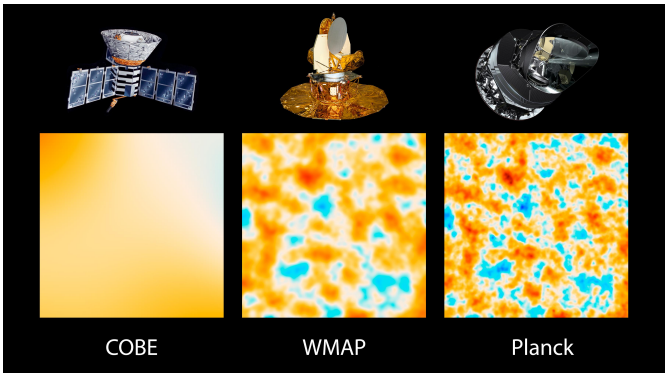


Figure 1.2: Three generations of satellites measuring the temperature distribution of the cosmic microwave background. Each panel shows a 10-square-degree patch of the full sky. The anisotropies are order 10^{-5} perturbations of the background temperature $\simeq 2.7\,\mathrm{K}$ and are the seeds to all structure that we see today. Credit: NASA/JPL-Caltech/ESA.

predict the behavior of for example merger events. See Chapter 7 for more details.

The Λ cold dark matter (ΛCDM) model is the concordance model for the evolution of the Universe. The model is based on general relativity and the standard model of particle physics plus the cosmological principle, that is, that the Universe is homogeneous and isotropic on large enough scales ($\gtrsim$ 100 Mpc). This means that the Universe looks, on average, the same from all points and in all directions. On scales in the range $\sim 1 - 100\,\mathrm{Mpc}$, there are inhomogeneities whose evolution is modeled with linear perturbation theory [39]. On smaller scales ($\lesssim 1\,\mathrm{Mpc}$), linear perturbation theory breaks down and N-body simulations are used to model the evolution of the inhomogeneities. The cosmological concordance model has been very successful in predicting and accounting for observations, with the relative abundance of hydrogen versus helium and the existence of the cosmic microwave background radiation (CMB, Fig. 1.2) as two of the first predictions [40–42]. We can use data from the CMB observations to constrain the parameters of the ΛCDM model, with the result that the dominant contributions to the energy budget of the Universe resides in dark energy ($\simeq 70\,\%$) and dark matter ($\simeq 25\,\%$) [43]. Only $\simeq 5\,\%$ of the energy content is due to "ordinary" (baryonic) matter like neutrons and protons plus an even smaller contribution from radiation (photons), see Fig. 1.3. Comparing the results from the CMB measurements with other cosmological probes, the ΛCDM model is consistent except from (at least) one important case, namely the value of the Hubble constant H_0 (see below).

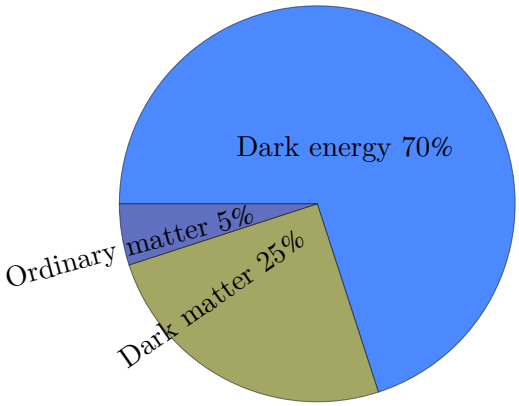


Figure 1.3: Energy budget of the Universe according to the cosmological concordance model.

The list of successes can be extended even further. To summarize, Einstein's theory describes most aspects of gravitation with exquisite precision. Nevertheless, there are reasons to believe that it is not the final theory for gravity, as we explore now.

1.4.2 Challenges

On solar-system scales and below, general relativity gives an excellent description of gravity. It is only on larger scales (or very small length scales close to the Planck length) that problems appear. So, what we are looking for is a theory of gravity that gives significant modifications on cosmological scales while restoring GR results on solar-system scales and below.

Dark energy. From observations on cosmological length scales, we infer that $\simeq 70\,\%$ of the energy content of the Universe is due to dark energy, which is needed to explain the accelerated expansion of the Universe. To give an accelerating contribution, the dark energy must have a negative pressure. The question is, what substance has this unexpected property? In the ΛCDM model, the accelerated expansion is due to a (positive) cosmological constant, Λ, which appears as a geometric term in the Einstein equations (1.18). To give a significant contribution on cosmological scales, we expect its value to be $\Lambda \sim H_0^2 \sim 10^{-52}\,\mathrm{m}^{-2}$. The quantum vacuum energy also gives a contribution to the accelerated expansion of the Universe. However, its value differs

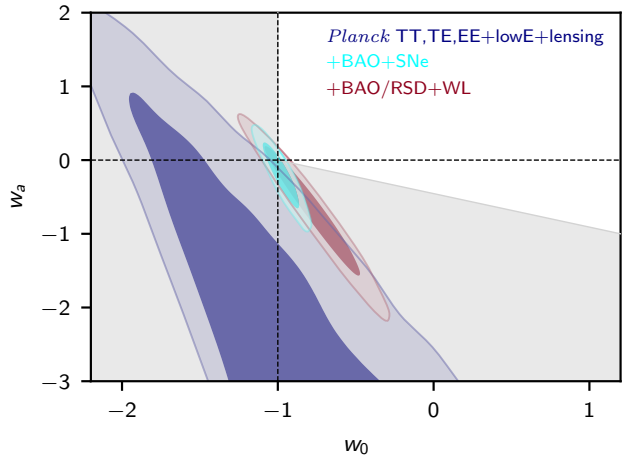


Figure 1.4: Marginalized confidence contours for the dark energy parameters w_0 and w_a, where the equation of state is parameterized as $w(a) = w_0 + (1-a)w_a$. The cosmological concordance model (ΛCDM) corresponds to $(w_0 = -1, w_a = 0)$, that is, the intersection between the dashed lines in the figure. The gray shaded area corresponds to phantom dark energy, and we see that it is preferred when using CMB data only (purple). Including other data sets such as baryon acoustic oscillations (BAO), centers the best-fit value around ΛCDM. Credit: Ref. [43], reproduced with permission ©ESO.

from the observed value by at least 60 orders of magnitude [39]. This is allegedly the worst prediction in the history of physics, so to be consistent with observations, the cosmological constant Λ must be fine-tuned with enormous precision. This sort of fine-tuning is contrary to the philosophical preferences of many physicists although, it should be stressed, there is no mathematical inconsistency in it.

One can go beyond the ΛCDM model and allow for dynamical dark energy where the equation of state changes with time. In fact, CMB data alone prefers a phantom dark energy which means a growing energy density with time [44], see Fig. 1.4. As we will see in Chapter 3, massive gravity provides such a contribution. Although this theory exhibits some novel features, it does not solve the fine-tuning problem. As an example of a new feature, there are self-accelerating cosmologies (i.e., without a cosmological constant) where the accelerated expansion is due to the interaction between the massless and massive gravitons. To date, there is no established solution to the fine-tuning problem [39].

Figure 1.5: Image of the type Ia supernova SN 1994D (bright spot on the bottom left) exploding in the host galaxy NGC 4526. A type Ia supernova is an exploding star with a peak absolute magnitude that can be standardized. Hence, by measuring its apparent magnitude, we can estimate the distance to it. Credit: NASA/ESA, The Hubble Key Project Team, and The High-Z Supernova Search Team.

Hubble tension. As cosmological observations have become increasingly precise, a discrepancy has appeared in the ΛCDM model. There are several ways to determine the Hubble constant, H_0. The nearby probes are based on observations in our cosmological neighborhood and are largely independent of the assumed cosmological model. There are several independent probes of this type and they all give "high" values of H_0. Most significantly, the SH0ES team reports $H_0 = 73.0 \pm 1.0\,\mathrm{km/s/Mpc}$, using a Cepheid-based cosmic distance ladder method measuring distances and redshifts of type Ia supernovae (SNIa, Fig. 1.5) [45]. Most of the distant probes involve CMB data. Here, H_0 is not measured directly but is inferred, assuming the ΛCDM cosmological model. Any CMB data set predicts a "low" value of H_0 in the range $67 - 68.5\,\mathrm{km/s/Mpc}$ [18]. For example, the inferred value from the Planck 2018 data release is $H_0 = 67.3 \pm 0.6\,\mathrm{km/s/Mpc}$ [43]. It is clear this value disagrees with the one from the Cepheid-based distance ladder. In Ref. [18], it is concluded that the discrepancy is at least 4σ. The discrepancy is largely independent of the methods and data sets employed. Thus, explaining it by a systematic error appears to be difficult, although there are some noteworthy exceptions, see for example Refs. [46–49]. This may suggest that there is new physics beyond the ΛCDM model.

Dark matter. There are observations spanning a wide range of length scales, from galactic scales to the largest cosmological scales, indicating that the gravitational field is stronger than what can be inferred from the amount of visible matter, assuming that GR is the correct theory for gravity. Granted this, there must be some sort of (yet) undetected matter, dubbed dark matter (DM). Observations suggest that this particle is "cold" (i.e., non-relativistic) and interacting very weakly with standard model particles.

One of the most striking observations of this phenomenon is the motion of the stars and gas in galaxies. Instead of the orbital velocities decreasing towards the edge of the galaxies, they remain constant, meaning that the galaxies would be ripped apart unless the gravitational force is much stronger than indicated by the amount of matter that can be observed. This also holds on galaxy cluster scales.

On cosmological scales, observations of the CMB anisotropies lead to the conclusion that $\simeq 25\,\%$ of the energy content of the Universe is in form of dark matter (again, assuming a cosmological concordance model, based on GR). Only $5\,\%$ of the energy budget is in form of ordinary matter (baryons). Within the cosmological concordance model, dark matter is also needed to account for the formation of structure (e.g., galaxies) from the initial perturbations as imprinted in the CMB. Without it, there would not have been deep enough potential wells for the matter to fall into in order to form the structure that we observe today.

Since there is no established direct detection of dark matter particles, one may want to look for alternative explanations. Perhaps the dark matter phenomenology is not due to a new particle but due to gravity working differently compared with GR. However, since dark matter is observed on such a wide range of length scales, it is challenging to find a modified law of gravity taking all these observations into account. This is in contrast to dark energy whose existence is inferred from cosmological scale observations only, allowing for a wide range of possible explanations. Massive gravity can give two important contributions to the dark matter problem. First, the massive graviton can act as a dark matter particle [50–52]. However, to provide a viable dark matter particle, the theory is pushed towards a limit where it becomes indistinguishable from GR. Second, massive gravity exhibits (for some parameter values) a screening mechanism that restores GR far inside some radius, r_V. However, as one approaches this radius, gravity becomes stronger than what GR predicts. Applied to galaxies and galaxy clusters, this can reduce the need of dark matter as gravity becomes stronger towards the edges, compared with the GR prediction. Moreover, r_V adapts to the density and mass of the object, so that the gravitational force is not affected on solar-system scales while giving a significant contribution on galactic scales.

Quantization of gravity. In the Einstein equations (1.18), the right-hand side contains the matter fields in form of the stress–energy tensor and the left-hand side describes the space-time geometry. Since matter (at the microscopic scale) is described by quantum theory, the left-hand side of eq. (1.18), which contains the space-time geometry, should also be quantized. We expect the quantum properties of gravity to be important at least at the Planck scale ($\sim 10^{-35}\,\mathrm{m}$) which is relevant in the very early universe or very close to the singularity of a black hole. Despite persistent efforts during at least half a century, there is no satisfying theory of quantum gravity. We will not say much more about this within the scope of this thesis. An interesting feature of a modified theory of gravity is if singularities can be avoided, possibly evading the need for a quantum theory of gravity. However, the results of Chapter 7 do not indicate that this is the case for massive gravity.

1.4.3 Ways forward

With a new theory of gravity, the hope is that one or several of the above problems could be solved. To construct new theories for gravity, one or several assumptions of Lovelock's theorem must be broken, that is:

- Introducing fields other than (or in addition to) the metric. Examples include adding extra fields of (non-minimally coupled) scalar, vector, or tensor type [53]. Another option is to encode gravity in terms of torsion or non-metricity [54].
- Allowing higher-order equations of motion, for example $f(R)$ theories where the Lagrangian is some function of the Ricci scalar [53].
- Giving up locality. One example is $f(\Box^{-1}R)$ theories.
- Introducing new dimensions, for example Kaluza–Klein theories or Dvali–Gabadadze–Porrati (DGP) gravity [53].
- Abandoning the action principle with a scalar density as the Lagrangian. Modified Newtonian dynamics (MOND) as initially formulated by Milgrom in Ref. [55] is an example of this as the modification of gravity is at the level of the equations of motion.

In this thesis, we study massive gravity where we introduce an additional spin-2 field, hence belonging to the first category.

Chapter 2

Massive gravity

Relativistic fields with spin ≤ 2 come in two versions, massless and massive, see Tab. 2.1. Spin 0 is described by a scalar field, ϕ, spin 1/2 by a spinor, ψ_α, spin 1 by a vector field A^μ, and massless spin-2 by a metric $g_{\mu\nu}$ (i.e., a symmetric nondegenerate rank-2 tensor field). In case spin ≤ 1, the massive version is obtained by simply adding the square of the field to the massless Lagrangian. For example, the Lagrangian of a massive spin-1 field is,

$$\mathcal{L}_{\text{spin 1}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{m^2}{2}A^\mu A_\mu, \quad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{2.1}$$

where the first term is the Maxwell Lagrangian for a massless spin-1 field and the second term provides mass to the field. The free spin ≤ 1 theories are simple in the sense that their equations of motion are linear. A spin-2 field coupled to matter is different. First, the equations of motion are nonlinear (see the Einstein equations (1.18)). Second, adding the square of the field to the Einstein–Hilbert action in eq. (1.17) only amounts to adding a cosmological constant term, since $g^{\mu\nu}g_{\mu\nu} = 4$. To give mass to the spin-2 field while preserving Lorentz invariance, we must introduce an additional metric, $f_{\mu\nu}$, allowing non-trivial terms like $g^{\mu\nu}f_{\mu\nu}$ to be added to the Einstein–Hilbert action.

There are infinitely many ways to add Lorentz invariant combinations of $g_{\mu\nu}$ and $f_{\mu\nu}$, so the question is what combinations (if any) that give sensible theories for a massive spin-2 field. In fact, as shown by Boulware and Deser a generic theory of massive gravity propagates six degrees of freedom including a ghost-like scalar mode with a negative kinetic energy, thus rendering the theory unstable [56]. Hence, the theory must be constructed to avoid this feature.

The first massive gravity theory was proposed in 1939 by Fierz and Pauli who constructed the linear theory for a massive spin-2 field propagating on a flat background [19]. Here, the second metric $f_{\mu\nu}$ was fixed to the flat

	spin 0	spin 1/2	spin 1	spin 2
Massless	Klein–Gordon	Dirac	Maxwell	GR/Massive gravity
Massive	Klein–Gordon	Dirac	Proca	Massive gravity

Table 2.1: Relativistic theories of spin ≤ 2. As we will see later, the massive gravity theory studied in this thesis exhibits both a massless and a massive spin-2 field, which is why it appears in both the massless and massive spin-2 rows.

Minkowski metric and the mass term was carefully designed to exhibit only the five healthy degrees of freedom of a massive spin-2 field. Nevertheless, the linear theory is problematic. In particular, in the zero mass limit, the scalar mode of the spin-2 field remains coupled to the trace of the stress–energy, hence general relativity is not recovered, contrary to expectations. On the other hand, setting the graviton mass identically equal to zero recovers GR. This is referred to the van Dam–Veltman–Zakharov (vDVZ) discontinuity [57, 58]. In particular, the bending of light differs by a factor of 4/3 compared with GR, thus ruling out Fierz–Pauli theory on observational grounds.

Shortly after the discovery of the vDVZ discontinuity, Vainshtein showed that nonlinear terms can become influential in the zero mass limit, thereby invalidating the linear Fierz–Pauli theory in this limit [59]. Thus, we should look for a nonlinear theory of massive gravity. However, Boulware and Deser showed that a generic nonlinear completion of Fierz–Pauli theory reintroduces the pathological ghost mode [56], thereby impeding further advances in the subject until it was reexamined almost forty years later.

The resurrection of massive gravity occurred in 2010 when dRGT (de Rham, Gabadadze, and Tolley) carefully constructed a nonlinear theory for massive gravity that avoids the Boulware–Deser ghost, thus propagating only the five healthy modes of a massive spin-2 field [20–23, 60, 61]. In this theory, the second metric is a fixed reference metric, usually taken to be flat Minkowski. We will refer to this version of massive gravity as dRGT theory, see Ref. [62] for a review. Unfortunately, dRGT massive gravity does not seem to be observationally viable, for example with respect to the expansion history of the Universe [63]. It might also appear unsatisfactory to have a non-dynamical second metric in the theory without any fundamental principle determining its form.

The natural extension of the dRGT theory is to promote the second metric to a dynamical field, as proposed by Hassan and Rosen in 2011 [23]. This is achieved by adding an Einstein–Hilbert term for $f_{\mu\nu}$ to the action, under

which the theory remains ghost-free [23, 60, 64]. We will refer to this theory as bimetric relativity (BR), bimetric gravity, massive gravity, massive bigravity, or simply bigravity and it is the main topic of this thesis. For a review of the subject, see Ref. [63]. In this theory, with both metrics being dynamical, there are 2+5 propagating modes corresponding to the two modes of a massless spin-2 field plus the five modes of a massive spin-2 field. The mass of the spin-2 field is referred to as the Fierz–Pauli mass, $m_{\rm FP}$. Notably, the mass eigenstates are linear superpositions of the metric perturbations, with the amount of mixing parameterized by an angle θ. With these remarks, we can already anticipate that massive bigravity allows for a rich phenomenology compared both with GR and dRGT massive gravity. In fact, GR and dRGT are retrieved at special points in the bimetric parameter space where one of the metrics is completely aligned with the massless mode ($\theta = 0$) or the massive mode ($\theta = \pi/2$), respectively.

It has previously been demonstrated that massive bigravity is observationally viable, with its cosmology being in accordance with data from the cosmic microwave background, baryon acoustic oscillations (BAO), and type Ia supernovae [10, 11, 65–71]. The theory can also satisfy solar-system tests [72–74], observations of velocity dispersion and gravitational lensing on galactic scales and galaxy cluster scales [72, 73, 75, 76] as well as gravitational wave observations [13, 77–80]. These observations set constraints in the bigravity parameter space. However, the previous literature have only explored particular subsets of the parameter or have only taken a subset of the observations into account. In Papers III-V, we treat the constraints in a unified way and explore the full parameter space of the theory. Thereby, we show that BR is compatible with the union of all these data sets, see Chapters 3-5 for details.

2.1 Action and equations of motion

The action for the ghost-free bimetric theory reads [23],

$$\mathcal{S}_{\rm HR} = \int d^4x \Big[\frac{1}{2\kappa_g} \sqrt{-\det g}\, R + \frac{1}{2\kappa_f} \sqrt{-\det f}\, \widetilde{R} - \sqrt{-\det g} \sum_{n=0}^{4} \beta_n e_n(S) + \\ + \sqrt{-\det g}\, \mathcal{L}_m + \sqrt{-\det f}\, \widetilde{\mathcal{L}}_m \Big]. \tag{2.2}$$

R is the Ricci scalar of $g_{\mu\nu}$ and $\widetilde{R}$ is the Ricci scalar of $f_{\mu\nu}$, so the first two terms in the Hassan–Rosen action (2.2) are Einstein–Hilbert terms for each one of the metrics, rendering them both dynamical. This is in contrast to dRGT massive gravity exhibiting an Einstein–Hilbert term for only one of the metrics.

The third term in eq. (2.2) is the interaction potential between the two metrics, which by its special form guarantees the absence of the Boulware–Deser ghost [23]. Here, $e_n(S)$ are the elementary symmetric polynomials of the square root of the two metrics, defined by the equation $S^\mu{}_\rho S^\rho{}_\nu = g^{\mu\rho} f_{\rho\nu}$ [81, 82].[1] The elementary symmetric polynomials are,

$$\begin{aligned} &e_0(S) = 1, \quad e_1(S) = [S], \quad e_2(S) = \frac{1}{2}\left([S]^2 - [S^2]\right), \\ &e_3(S) = \frac{1}{6}\left([S]^3 - 3[S^2][S] + 2[S^3]\right), \quad e_4(S) = \det S, \end{aligned} \tag{2.3}$$

where $[S] \equiv \operatorname{Tr} S$.

There can be two independent matter sectors, with Lagrangians $\mathcal{L}_m$ and $\widetilde{\mathcal{L}}_m$, coupled to $g_{\mu\nu}$ and $f_{\mu\nu}$, respectively [83, 84]. With $\mathcal{L}_m$ containing the standard model particles and fields, we identify $g_{\mu\nu}$ as the physical metric since it determines the motion of freely falling observers. Thus, $g_{\mu\nu}$ defines the observable curvature of the Universe which can be measured by rulers and clocks. For an observer, although $f_{\mu\nu}$ is a metric by definition (i.e., a symmetric nondegenerate rank-2 tensor field), it can only be measured indirectly via its interaction with the physical metric.

In eq. (2.2), κ_g and κ_f are the gravitational constants of $g_{\mu\nu}$ and $f_{\mu\nu}$, respectively. In geometrized units, $\kappa_g = 8\pi$. The ratio of the gravitational constants is denoted,

$$\kappa \equiv \kappa_g/\kappa_f. \tag{2.4}$$

In the limit $\kappa \to 0$, BR approaches GR continuously and in the limit $\kappa \to \infty$, it approaches dRGT theory [85]. In addition to the gravitational constants there are five, constant, free parameters in the Hassan–Rosen action (2.2), referred to as β_n, with dimensions of curvature $1/\mathsf{L}^2$. Since $e_0(S) = 1$, the β_0-term adds a cosmological constant term with respect to the g-metric. Moreover, since $e_4(S) = \det\sqrt{g^{-1}f}$ and $\sqrt{-\det g}\det\sqrt{g^{-1}f} = \sqrt{-\det f}$, it follows that the β_4-term contributes with a cosmological constant to the f-metric. The other β-parameters $(\beta_1, \beta_2, \beta_3)$ determine the interaction between the two metrics.

Varying the action (2.2) with respect to the metrics, we obtain the equations of motion,

$$G^\mu{}_\nu = \kappa_g \mathcal{T}^\mu{}_\nu, \quad \mathcal{T}^\mu{}_\nu \equiv T^\mu{}_\nu + V^\mu{}_\nu, \tag{2.5a}$$

$$\widetilde{G}^\mu{}_\nu = \kappa_f \widetilde{\mathcal{T}}^\mu{}_\nu, \quad \widetilde{\mathcal{T}}^\mu{}_\nu \equiv \widetilde{T}^\mu{}_\nu + \widetilde{V}^\mu{}_\nu, \tag{2.5b}$$

[1] Note that there is no unique solution to this equation. However, as argued in Ref. [81], choosing the principal square root guarantees a sensible space-time interpretation of the theory. The principal square root is unique and defined by the requirement that its eigenvalues lie in the right-half complex plane.

where the bimetric stress–energies $V^{\mu}{}_{\nu}$ and $\widetilde{V}^{\mu}{}_{\nu}$ are obtained by varying the interaction term with respect to $g_{\mu\nu}$ and $f_{\mu\nu}$. The result is,

$$V^{\mu}{}_{\nu} \equiv -\sum_{n=0}^{3} \beta_n \sum_{k=0}^{n} (-1)^{n+k} e_k(S) (S^{n-k})^{\mu}{}_{\nu}, \tag{2.6a}$$

$$\widetilde{V}^{\mu}{}_{\nu} \equiv -\sum_{n=0}^{3} \beta_{4-n} \sum_{k=0}^{n} (-1)^{n+k} e_k(S^{-1}) (S^{-(n-k)})^{\mu}{}_{\nu}. \tag{2.6b}$$

To avoid the Boulware–Deser ghost, each matter field must couple to a single metric [83, 86] (see Ref. [87] for an analysis of the phenomenology of doubly coupled bigravity). Henceforth, we assume that there is no matter sector coupled to $f_{\mu\nu}$, that is, $\widetilde{T}^{\mu}{}_{\nu} = 0$. See Refs. [88, 89] for examples of works analyzing bimetric theories with nonzero $\widetilde{T}^{\mu}{}_{\nu}$. If we assume that the stress–energy is conserved,

$$\nabla_{\mu} T^{\mu}{}_{\nu} = 0, \tag{2.7}$$

the bimetric stress–energy is also conserved,

$$\nabla_{\mu} V^{\mu}{}_{\nu} = 0, \tag{2.8}$$

which follows from eq. (2.5a) and the Bianchi identity $\nabla_{\mu} G^{\mu}{}_{\nu} = 0$.

2.2 The physical parameters

The β-parameters have traditionally been used to parameterize the bimetric interaction. However, there is a problem. Using the identity $e_n(\lambda S) = \lambda^n e_n(S)$ where λ is a scalar, it is straightforward to show that the Hassan–Rosen action is invariant under the constant rescaling,

$$(f_{\mu\nu}, \kappa_f, \beta_n) \to (\omega f_{\mu\nu}, \omega \kappa_f, \omega^{-n/2} \beta_n), \quad \omega = \text{const.} \tag{2.9}$$

This means that the β-parameters are not observables and cannot be constrained by observations. To circumvent this issue, we introduce below a set of physical parameters which is independent of the rescaling.

As we will see in Chapter 3, there are several types of homogeneous and isotropic solutions in massive bigravity. However, the only ones that have proven healthy (in a sense that is specified in Chapter 3) are those for which the metrics approach a proportional de Sitter solution in the infinite future,

$$f_{\mu\nu}|_{t\to\infty} = c^2 g_{\mu\nu}|_{t\to\infty}, \tag{2.10}$$

with $c = \text{const}$. For the local and cosmological solutions to be consistent with each other, we consider local solutions which are proportional asymptotically (i.e., at radial infinity).

In terms of the conformal factor c, the rescaling in eq. (2.9) reads,

$$(c^2, \kappa, \beta_n) \to (\omega c^2, \omega^{-1}\kappa, \omega^{-n/2}\beta_n). \tag{2.11}$$

So, in the bigravity action (2.2) we have the theory parameters κ_g, κ_f, $\beta_0, ..., \beta_4$ of which one is redundant due to the rescaling invariance. Hence, there are five free parameters (or six, if including κ_g). Usually, the rescaling invariance is used to set $c = 1$ or $\kappa_g/\kappa_f = 1$. Here, we introduce a set of parameters which is independent of this choice. First, note that $\beta_n c^n$ is invariant under the rescaling. Thus, we define the dimensionless, rescaling invariant parameters,

$$B_n \equiv \kappa_g \beta_n c^n / H_0^2, \tag{2.12}$$

that is, expressing the β-parameters in units of the curvature scale defined by the Hubble constant H_0. This is convenient for cosmological applications where we expect significant massive gravity contributions to the present-day expansion rate if $\kappa_g \beta_n c^n \sim H_0^2$, that is $B_n \sim 1$.

Instead of the B-parameters, we can use the dimensionless physical parameters $\Theta \equiv (\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta)$, given by,

$$\tan^2\theta = \frac{B_1 + 3B_2 + 3B_3 + B_4}{B_0 + 3B_1 + 3B_2 + B_3}, \tag{2.13a}$$

$$m_{\rm FP}^2 = (B_1 + 2B_2 + B_3) \,/\, \sin^2\theta, \tag{2.13b}$$

$$\Omega_\Lambda = \frac{B_0}{3} + B_1 + B_2 + \frac{B_3}{3}, \tag{2.13c}$$

$$\alpha = -\frac{B_2 + B_3}{B_1 + 2B_2 + B_3}, \tag{2.13d}$$

$$\beta = \frac{B_3}{B_1 + 2B_2 + B_3}. \tag{2.13e}$$

In Refs. [68, 90], a limited parameterization in terms of $(\theta, m_{\rm FP}, \Omega_\Lambda)$ was first proposed, applicable to models with three free parameters, or less. In Paper III, by introducing α and β we generalized the framework to include any bimetric model. The parameters in eq. (2.13) have immediate physical interpretations, as explained further in the following chapters.

- In massive bigravity, the spin-2 mass eigenstates are not aligned with the metrics. Here, $0 \leq \theta \leq \pi/2$ is the mixing angle between the massless and massive gravitons. In the limit $\theta \to 0$, the physical metric is aligned with the massless spin-2 field (GR) and in the limit $\theta \to \pi/2$, it is aligned with the massive spin-2 field (dRGT).

- In so-called natural units, $m_{\rm FP}$ is the mass of the graviton in units of $H_0 \sim 10^{-33}\,\text{eV}$. In geometrized units, $m_{\rm FP}$ is the inverse Compton wavelength of the massive graviton in units of the Hubble radius, that is, $m_{\rm FP} = r_H/\lambda_g$ where $r_H \equiv 1/H_0 \simeq 5\,\text{Gpc}$. We will refer to the Compton wavelength and the graviton mass interchangeably, keeping in mind that they reflect the same physical parameter.
- Ω_Λ is the effective cosmological constant in the final de Sitter phase of the expansion history of the Universe.
- α and β are parameters that determine (among other things) the screening mechanism, restoring GR locally.

The relations in eq. (2.13) can be inverted, resulting in,

$$B_0 = 3\Omega_\Lambda - \sin^2\theta\, m_{\rm FP}^2(3+3\alpha+\beta), \tag{2.14a}$$
$$B_1 = \sin^2\theta\, m_{\rm FP}^2(1+2\alpha+\beta), \tag{2.14b}$$
$$B_2 = -\sin^2\theta\, m_{\rm FP}^2(\alpha+\beta), \tag{2.14c}$$
$$B_3 = \sin^2\theta\, m_{\rm FP}^2\beta, \tag{2.14d}$$
$$B_4 = 3\tan^2\theta\,\Omega_\Lambda + \sin^2\theta\, m_{\rm FP}^2(-1+\alpha-\beta). \tag{2.14e}$$

The Jacobian determinant of eq. (2.14) is $18\sin^6\theta\, m_{\rm FP}^5\Omega_\Lambda$. Hence, there is a one-to-one correspondence between the physical parameters and the B-parameters, except when θ, $m_{\rm FP}$, or Ω_Λ vanishes. In the limit $\theta \to 0$, we have $B_1, B_2, B_3, B_4 \to 0$ and only the cosmological constant of $g_{\mu\nu}$ survives. In this limit, GR is recovered, as expected. From eq. (2.14), one may expect to recover GR also as $m_{\rm FP} \to 0$. However, this limit is problematic since it excites the Higuchi ghost (see Chapter 3 for details).

2.3 Symmetries and topology

Besides the theoretical consistency of BR (avoiding the Boulware–Deser ghost), it should agree with observations. Therefore, it is of paramount importance to understand the spectrum of solutions to the equations of motion. A first step is to analyze the equations under the assumption of certain symmetries. In this way, one can obtain approximate solutions to more realistic setups. In this section, we review the main results of Paper I. Due to the presence of the second metric, "spherical symmetry" for example is an ambiguous notion. Does it mean that $g_{\mu\nu}$ is spherically symmetric or both $g_{\mu\nu}$ and $f_{\mu\nu}$? If the latter, do the metrics look spherically symmetric in the same coordinates?

Since $g_{\mu\nu}$ can be identified as the physical metric due to its coupling to matter, it is sufficient to demand spherical symmetry (SS) with respect to

$g_{\mu\nu}$ in order for an observer to measure a spherically symmetric space-time geometry. The second metric may or may not be spherically symmetric, that is, have an SO(3) isometry, and if it is, the Killing vector fields (KVFs) of $g_{\mu\nu}$ and $f_{\mu\nu}$ may or may not coincide. A KVF generates a set of curves along which a metric stays invariant. For example, if the two metrics are both SS (i.e., they have the same isometry group) but their KVFs do not coincide, it means that the spherical symmetry is manifest in different coordinates for $g_{\mu\nu}$ and $f_{\mu\nu}$, see Example 1 below. For a recap of symmetries and topology, see Appendix B.

Example 1. Let us fix $\beta_0 = 1$, $\beta_1 = 0$, $\beta_2 = -1/2$, $\beta_3 = 1$, and $\beta_4 = -7/4$. In this case,

$$ds_g^2 = -dt^2 + dx^2 + dy^2 + dz^2, \tag{2.15a}$$
$$ds_f^2 = -\lambda_1^2 dt^2 + \lambda_1^2 dx^2 + \lambda_2^2 dy^2 + \lambda_2^2 dz^2, \tag{2.15b}$$

solves the equations of motion, with $\lambda_1 = 2-\sqrt{2}$, $\lambda_2 = 2+\sqrt{2}$. The metrics in eq. (2.15) are written in the Cartesian coordinates of $g_{\mu\nu}$, hence the spherical symmetry of $g_{\mu\nu}$ is manifest. The Cartesian coordinates of $f_{\mu\nu}$ are given by the rescaling,

$$(t', x', y', z') = (\lambda_1 t, \lambda_1 x, \lambda_2 y, \lambda_2 z), \tag{2.16}$$

in which the metrics read,

$$ds_g^2 = -\lambda_1^{-2} dt'^2 + \lambda_1^{-2} dx'^2 + \lambda_2^{-2} dy'^2 + \lambda_2^{-2} dz'^2, \tag{2.17a}$$
$$ds_f^2 = -dt'^2 + dx'^2 + dy'^2 + dz'^2. \tag{2.17b}$$

Both $g_{\mu\nu}$ and $f_{\mu\nu}$ are Minkowski metrics, but they do not share their Cartesian coordinates. Hence, they are both spherically symmetric but with different Killing vector fields,

$$J_x = y\partial_z - z\partial_y, \qquad J_y = -z\partial_x + x\partial_z, \qquad J_z = x\partial_y - y\partial_x, \tag{2.18a}$$
$$\widetilde{J}_{x'} = y'\partial_{z'} - z'\partial_{y'}, \quad \widetilde{J}_{y'} = -z'\partial_{x'} + x'\partial_{z'}, \quad \widetilde{J}_{z'} = x'\partial_{y'} - y'\partial_{x'}. \tag{2.18b}$$

The KVFs of $f_{\mu\nu}$ can be transformed to the Cartesian coordinates of $g_{\mu\nu}$, for example,

$$\widetilde{J}_{z'} = -\frac{\lambda_2}{\lambda_1} y\,\partial_x + \frac{\lambda_1}{\lambda_2} x\,\partial_y, \tag{2.19}$$

and we see that $\widetilde{J}_{z'} \neq J_z$. In other words, $g_{\mu\nu}$ and $f_{\mu\nu}$ look manifestly spherically symmetric in different coordinates. An observer in the Cartesian chart of $g_{\mu\nu}$ sees this metric as spherically symmetric while a transformation on the surface of an ellipsoid represents the symmetry of $f_{\mu\nu}$, see Fig. 2.1.

□

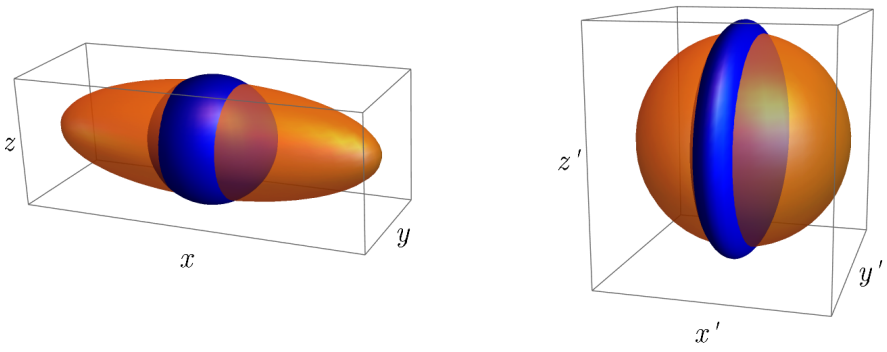


Figure 2.1: Symmetry orbits of the space-time (2.15). The $g_{\mu\nu}$-metric is invariant when transformed along the blue surface while $f_{\mu\nu}$ is invariant when transformed along the orange surface. Both metrics are spherically symmetric but they do not share their KVFs. *Left panel*: Cartesian coordinates of $g_{\mu\nu}$. *Right panel*: Cartesian coordinates of $f_{\mu\nu}$.

Example 2. An example of a solution where the two metrics have different symmetry groups is the case of the non-bidiagonal cosmological solutions, see Chapter 3. Here, the second metric $f_{\mu\nu}$ is maximally symmetric (anti-de Sitter) while the physical metric $g_{\mu\nu}$ is only homogeneous and isotropic.

□

Since the two metrics live on the same manifold, they must be compatible with the same topology. This imposes an additional consistency requirement besides satisfying the equations of motion. See Tab. 4.1 for examples of compatible and incompatible solutions for a set of static, spherically symmetric solutions. Note that the topological constraint applies also to dRGT massive gravity. In this case, the auxiliary metric is a fixed reference metric, usually Minkowski. However, note that the requirement applies only to the full space-time. If we think of our solution as an approximation, valid in some limited region, the requirement does not apply.

2.4 Spherical symmetry

Within this thesis, we make extensive use of spherical symmetry: (1) for the background cosmology, (2) for the local solutions, modeling for example the gravitational field of stars and galaxies, and (3) for gravitational collapse. Here, we assume that the metrics exhibit the same spherical symmetry, that is, that they share their KVFs. With these assumptions, a general ansatz is,

$$ds_g^2 = -A(t,r)dt^2 + B(t,r)dr^2 + 2C(t,r)dtdr + R^2(t,r)d\Omega^2, \tag{2.20a}$$
$$ds_f^2 = -\widetilde{A}(t,r)dt^2 + \widetilde{B}(t,r)dr^2 + 2\widetilde{C}(t,r)dtdr + \widetilde{R}^2(t,r)d\Omega^2, \tag{2.20b}$$

where $d\Omega^2$ is the line element on the unit 2-sphere, $d\Omega^2 \equiv d\theta^2 + \sin^2\theta\, d\phi^2$. Using the coordinate freedom, one can for example choose t and r such that $R = r$ and $C = 0$ (at least locally),

$$ds_g^2 = -A(t,r)dt^2 + B(t,r)dr^2 + r^2d\Omega^2, \tag{2.21a}$$
$$ds_f^2 = -\widetilde{A}(t,r)dt^2 + \widetilde{B}(t,r)dr^2 + 2\widetilde{C}(t,r)dtdr + \widetilde{R}^2(t,r)d\Omega^2. \tag{2.21b}$$

These solutions can then be classified according to whether $\widetilde{C} = 0$ (bidiagonal solutions) or $\widetilde{C} \neq 0$ (non-bidiagonal solutions). If the physical metric is also homogeneous (cosmological background solutions) or static (local solutions, e.g. for a galaxy), the non-bidiagonal configuration reduces to a GR solution, see Refs. [91–93]. Hence, in these cases only the bidiagonal configuration yields interesting deviations from GR.

2.5 Summary

Bimetric gravity is free from the Boulware–Deser ghost and exhibits both a massless and a massive spin-2 field. The theory has a continuous GR limit as well as a continuous limit to dRGT massive gravity, thus embracing these two theories in a unified framework. Bimetric gravity can be conveniently parameterized by five dimensionless (constant) parameters, the physical parameters, which have immediate physical interpretations. Those are the mixing angle between the massless and massive spin-2 field, θ, the graviton mass, m_{FP}, the effective cosmological constant Ω_Λ, and two parameters, denoted α and β, that determine (among other things) the screening mechanism which restores GR locally. Since there are four additional parameters compared with GR (which only has the cosmological constant and the gravitational constant), BR has a very rich phenomenology but it is also more difficult to solve the equations of motion.

Due to the presence of two metrics, a symmetry with respect to one metric does not imply the same symmetry with respect to the other metric. Moreover, even if the metrics have the same symmetry (i.e., the same symmetry group), their Killing vector fields can be different, that is, the symmetries of the two metrics are manifest in different coordinates. Usually, in the literature it is assumed that the metrics exhibit the same symmetries as well as the same Killing vector fields. Our result demonstrates the possibility of finding new solutions which may satisfy our demands on the physical metric but leaving the second metric on a more general form.

Chapter 3

Background cosmology

The cosmological principle states that the Universe is homogeneous and isotropic on large scales ($\gtrsim 100\,\mathrm{Mpc}$), that is, it looks the same from all points and in all directions. This is an observational fact that is supported for example by large-scale structure surveys, observing the distribution of galaxies in space, and by the isotropy of the cosmic microwave background radiation. However, the Universe is not the same at all times. In fact, the Universe expands at an accelerated rate, which was discovered more than two decades ago by observing type Ia supernovae. Tracing the evolution backwards in time hints of a very dense initial state 13.8 billion years ago. However, it is still a matter of dispute what drives the accelerated expansion. The simplest explanation is a cosmological constant but, as we have seen, it is usually regarded as problematic. In this chapter we show that, in BR, the accelerated expansion can be due to the interaction of a massive spin-2 field, in which case no cosmological constant is needed. These expansion histories are called self-accelerating.

3.1 Ansatz and equations of motion

Applying the cosmological principle, we assume that the physical metric is homogeneous and isotropic. For simplicity, we also assume that the second metric exhibits the same spherical symmetry (with the same KVFs). In this case, the general ansatz reads,

$$ds_g^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right), \tag{3.1a}$$

$$ds_f^2 = -\widetilde{A}(t,r)dt^2 + \widetilde{B}(t,r)dr^2 + 2\widetilde{C}(t,r)dtdr + \widetilde{R}^2(t,r)d\Omega^2, \tag{3.1b}$$

where k is a constant determining the spatial curvature of the physical metric. The freedom of choosing coordinates have been used to write the physical

metric on the form in eq. (3.1a). As shown in Ref. [91], for the non-bidiagonal solutions (i.e., with $\widetilde{C} \neq 0$), the two metric sectors decouple and the bimetric interaction contributes with a cosmological constant term to the equations of motion for $g_{\mu\nu}$. In this case, the background cosmology is identical to that of GR. For these solutions, the second metric is anti-de Sitter (AdS). However, the topologies associated with the two metrics are incompatible, which is inconsistent.

Henceforth, we focus on the bidiagonal branch ($\widetilde{C} = 0$). Assuming that also the second metric is homogeneous, the general ansatz is,

$$ds_g^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right), \tag{3.2a}$$

$$ds_f^2 = c^2\left[-x^2(t)dt^2 + \tilde{a}^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right)\right], \tag{3.2b}$$

with $a(t)$ being the scale factor of the physical metric and $\tilde{a}(t)$ the scale factor of the second metric. The constant factor c has been included for convenience but can be absorbed in x and $\tilde{a}$ if desired. From now on, we let the time dependence of the functions be implicit.

The matter content can be split into relativistic and non-relativistic species, with equation of state,

$$p_i = w_i \rho_i, \tag{3.3}$$

where i is a label referring to the species. The relativistic species (also referred to as radiation) have $w_r = 1/3$ and include photons and neutrinos.[1] The non-relativistic species (also referred to simply as matter) have $w_m = 0$ and include baryons (i.e., ordinary matter) and cold dark matter. A cosmological constant has equation of state $w_\Lambda = -1$.

The stress–energy takes the form of a perfect fluid,

$$T^\mu{}_\nu = p_{\rm tot}(\delta^\mu{}_\nu + U^\mu U_\nu) + \rho_{\rm tot} U^\mu U_\nu, \tag{3.4}$$

where $p_{\rm tot} = \sum_i p_i$ and $\rho_{\rm tot} = \sum_i \rho_i$ is the total pressure and energy density, respectively and U^μ is the four-velocity of the fluids. Here, the coordinates are comoving with the flow of the perfect fluid, hence $U^\mu = \delta^\mu{}_t$ and $T^\mu{}_\nu = \mathrm{diag}(-\rho_{\rm tot}, p_{\rm tot}, p_{\rm tot}, p_{\rm tot})$.

[1] The neutrinos are very light ($\lesssim 0.1\,\mathrm{eV}$), hence they are relativistic at least in the early universe. In the late universe (depending on their mass), they can be non-relativistic.

With the ansatz (3.2), the bimetric stress–energies in eq. (2.6) can be written,

$$V^{\mu}{}_{\nu} = p_{\rm DE}(\delta^{\mu}{}_{\nu} + U^{\mu}U_{\nu}) + \rho_{\rm DE}U^{\mu}U_{\nu} = \mathrm{diag}(-\rho_{\rm DE}, p_{\rm DE}, p_{\rm DE}, p_{\rm DE}), \tag{3.5a}$$

$$\widetilde{V}^{\mu}{}_{\nu} = \widetilde{p}_{\rm DE}(\delta^{\mu}{}_{\nu} + U^{\mu}U_{\nu}) + \widetilde{\rho}_{\rm DE}U^{\mu}U_{\nu} = \mathrm{diag}(-\widetilde{\rho}_{\rm DE}, \widetilde{p}_{\rm DE}, \widetilde{p}_{\rm DE}, \widetilde{p}_{\rm DE}), \tag{3.5b}$$

where the explicit expressions for the energy densities and pressures can be found in eq. (D.1) of Appendix D. Notably, the bimetric stress–energies act as perfect fluids, comoving with the matter fluid. We use the subscripts DE (dark energy) for the bimetric energy densities and pressures since, as we will see, it gives a dark energy contribution to the equations of motion, that is, contributing to the accelerated expansion of the Universe.

With the ansatz (3.2), the bimetric conservation law (2.8) reads,

$$\left(x - \dot{\tilde{a}}/\dot{a}\right)\left[1 + 2\alpha(1-y) + \beta(1-y)^2\right] = 0, \tag{3.6}$$

where an overdot denotes a time derivative and y is the ratio of the scales factors,

$$y \equiv \tilde{a}/a. \tag{3.7}$$

Setting the expressions within the square brackets of eq. (3.6) to zero, yields y in terms of the constant parameters α and β. Hence, the scale factors of the two metrics are proportional. In this case, $V^{\mu}{}_{\nu}$ contributes with a cosmological constant term to the equations of motion and the cosmological background solutions are again identical to those of GR [65]. This branch has a tachyonic instability at the linear level in the late universe [94] and exhibits a non-perturbative ghost mode [95]. Therefore, we will focus on the dynamical branch from now on. This branch is obtained by setting the expression within the first parenthesis of eq. (3.6) to zero, thus y becomes dynamical. The equation is solved for x with the result,

$$x = \frac{\dot{\tilde{a}}}{\dot{a}} = y + \frac{dy}{d\,\ln a}. \tag{3.8}$$

The last step can be shown using the chain rule. The Hubble parameter is defined as the expansion rate of the Universe. Since we have two metrics, there are two Hubble parameters,

$$H \equiv \dot{a}/a, \quad \widetilde{H} \equiv \dot{\tilde{a}}/(\tilde{a}x). \tag{3.9}$$

From eq. (3.8), it is straightforward to show the following relation between the Hubble parameters,

$$H = y\widetilde{H}. \tag{3.10}$$

In the following, we write the energy densities in terms of the critical energy density today,[2]

$$\rho_c = 3H_0^2/\kappa_g, \tag{3.11}$$

defining,

$$\Omega_i \equiv \frac{\rho_i}{\rho_c}, \quad \Omega_k \equiv -\frac{k}{H_0^2 a^2}. \tag{3.12}$$

More explicitly, the dark energy density is,

$$\Omega_{\rm DE} = \Omega_\Lambda - \sin^2\theta\, m_{\rm FP}^2(1-y)\left[1+\alpha(1-y)+\frac{\beta}{3}(1-y)^2\right]. \tag{3.13}$$

Note that $\Omega_{\rm DE}$ is a function of y (and the physical parameters) and since y is dynamical, the dark energy $\Omega_{\rm DE}$ is also dynamical.

The tt-component of the $g_{\mu\nu}$-equations of motion yields a Friedmann equation,

$$E^2 = \Omega_{\rm tot} + \Omega_k + \Omega_{\rm DE}, \tag{3.14}$$

where we have defined the normalized expansion rate and the total energy density due to matter and radiation as,

$$E \equiv H/H_0, \quad \Omega_{\rm tot} \equiv \Omega_m + \Omega_r. \tag{3.15}$$

The tt-component of the $f_{\mu\nu}$-equations of motion reads,

$$E^2 = \Omega_k + \widetilde{\Omega}_{\rm DE}, \tag{3.16}$$

where,

$$\widetilde{\Omega}_{\rm DE} = \frac{\cos^2\theta\, m_{\rm FP}^2}{y^2}\left[\frac{1}{3}(1+2\alpha+\beta)y - (\alpha+\beta)y^2 + \beta y^3 + \frac{1}{3}\left((-1+\alpha-\beta) + \frac{3\sec^2\theta\,\Omega_\Lambda}{m_{\rm FP}^2}\right)y^4\right], \tag{3.17}$$

so that $\widetilde{\Omega}_{\rm DE}$ can be expressed as a function of y and the physical parameters only. Using eqs. (3.14) and (3.16), we can eliminate E and Ω_k and write, $\Omega_{\rm tot} = \widetilde{\Omega}_{\rm DE} - \Omega_{\rm DE}$. Expanding the equation gives,

[2] When $\rho_{\rm tot} + \rho_{\rm DE} = \rho_c$, the Universe is spatially flat. For $\rho_{\rm tot} + \rho_{\rm DE} > \rho_c$, the Universe has a positive spatial curvature and for $\rho_{\rm tot} + \rho_{\rm DE} < \rho_c$, the Universe has a negative spatial curvature.

$$
\begin{aligned}
&- \frac{1}{3}\cos^2\theta\, m_{\rm FP}^2(1+2\alpha+\beta) \\
&+ \left[\Omega_{\rm tot} + \Omega_\Lambda + m_{\rm FP}^2\left(\cos^2\theta\,(\alpha+\beta) - \sin^2\theta\,\left(1+\alpha+\frac{\beta}{3}\right)\right)\right] y \\
&+ m_{\rm FP}^2\left[-\cos^2\theta\,\beta + \sin^2\theta\,(1+2\alpha+\beta)\right] y^2 \\
&- \left[\Omega_\Lambda + \frac{1}{3}m_{\rm FP}^2\left(\cos^2\theta\,(-1+\alpha-\beta) + 3\sin^2\theta\,(\alpha+\beta)\right)\right] y^3 \\
&+ \frac{1}{3}\sin^2\theta\, m_{\rm FP}^2\beta y^4 = 0.
\end{aligned}
\tag{3.18}
$$

This equation can be used to solve for y in terms of $\Omega_{\rm tot}$ and the physical parameters. Conversely, $\Omega_{\rm tot}$ can be expressed purely in terms of y and the physical parameters.

To close the system of equations, we write down the conservation equation (2.7) for matter and radiation,

$$
\frac{d\Omega_m}{d\,\ln a} = -3(1+w_m)\Omega_m, \quad \frac{d\Omega_r}{d\,\ln a} = -3(1+w_r)\Omega_r, \tag{3.19}
$$

and solve,

$$
\Omega_m(z) = \Omega_{m,0}(1+z)^3, \quad \Omega_r(z) = \Omega_{r,0}(1+z)^4, \tag{3.20}
$$

where we used $w_m = 0$ and $w_r = 1/3$. Here, we have introduced the redshift, z, defined by,

$$
1 + z = a_0/a, \tag{3.21}
$$

where subscript "0" denotes the present-day value of the quantity. The equation of state parameter $w_{\rm tot}$ is defined by the equation,

$$
\frac{d\Omega_{\rm tot}}{d\,\ln a} = -3(1+w_{\rm tot})\Omega_{\rm tot}. \tag{3.22}
$$

The value of $w_{\rm tot}$ depends on whether matter or radiation is the dominating contribution to $\Omega_{\rm tot}$. During matter-domination, $w_{\rm tot} = w_m = 0$, and during radiation-domination, $w_{\rm tot} = w_r = 1/3$.

Summarizing, the equations of motion for the bimetric background cosmology are given by eq. (3.14) (or, alternatively, eq. (3.16)) together with the evolution of the matter and radiation densities, eq. (3.20), and the quartic polynomial for y, eq. (3.18).

Instead of the quartic polynomial for y, eq. (3.18), we can use eq. (3.22) and write down an evolution equation for y,

$$
\frac{dy}{d\,\ln a} = -3(1+w_{\rm tot})\Omega_{\rm tot}\left(\frac{d\Omega_{\rm tot}}{dy}\right)^{-1}. \tag{3.23}
$$

Here, $d\Omega_{\rm tot}/dy$ can be calculated by solving eq. (3.18) for $\Omega_{\rm tot}$ as a function of y and then differentiate.

Two useful identities are,

$$y\frac{d\Omega_{\rm DE}}{dy} = \left(1+\frac{1}{\tan^2\theta\, y^2}\right)^{-1} m_{\rm eff}^2, \tag{3.24}$$

$$-y\frac{d\Omega_{\rm tot}}{dy} = m_{\rm eff}^2 - 2\widetilde{\Omega}_{\rm DE}. \tag{3.25}$$

They can be shown by plugging in the definitions of $\Omega_{\rm DE}$, $\widetilde{\Omega}_{\rm DE}$, and $m_{\rm eff}$ and simplifying the expressions. From eq. (3.24), it follows that $\Omega_{\rm DE}$ increases if y increases and vice versa. In the equations, we introduced an effective graviton mass,

$$\begin{aligned} m_{\rm eff}^2 &\equiv \left(1+\frac{1}{\tan^2\theta\, y^2}\right)(B_1 y + 2B_2 y^2 + B_3 y^3) \\ &= m_{\rm FP}^2\left(\sin^2\theta\, y + \cos^2\theta\, y^{-1}\right)\left(1+2\alpha(1-y)+\beta(1-y)^2\right). \end{aligned} \tag{3.26}$$

In the final de Sitter phase, we have $y|_{t\to\infty} = 1$ and the effective mass coincides with the Fierz–Pauli mass, $m_{\rm eff}|_{t\to\infty} = m_{\rm FP}$.

Using eq. (3.25), we can rewrite eq. (3.23) as,

$$\frac{dy}{d\,\ln a} = \frac{3(1+w_{\rm tot})y\,\Omega_{\rm tot}}{m_{\rm eff}^2 - 2\widetilde{\Omega}_{\rm DE}}. \tag{3.27}$$

The right-hand side can be expressed as a function of y (and the physical parameters) only. Hence, the phase space of y is one-dimensional, see Fig. 3.1 for some examples. In this figure, we plot $(1+w_{\rm tot})^{-1}dy/d\,\ln a$ against y in the range $0 \le y \le 1$. In these cases, $y = 0$ at the Big Bang and then increases monotonically until reaching the de Sitter point $y = 1$ in the infinite future. In the intermediate region, the evolution of y can take a wide spectrum of different forms depending on the physical parameters.

The equation of state parameter for the bimetric fluid is given by $w_{\rm DE} = p_{\rm DE}/\rho_{\rm DE}$ or, equivalently, by the equation,

$$\frac{d\Omega_{\rm DE}}{d\,\ln a} = -3(1+w_{\rm DE})\Omega_{\rm DE}, \tag{3.28}$$

which can be rewritten as,

$$w_{\rm DE} = -1 - \frac{1}{3\Omega_{\rm DE}}\frac{dy}{d\,\ln a}\frac{d\Omega_{\rm DE}}{dy}. \tag{3.29}$$

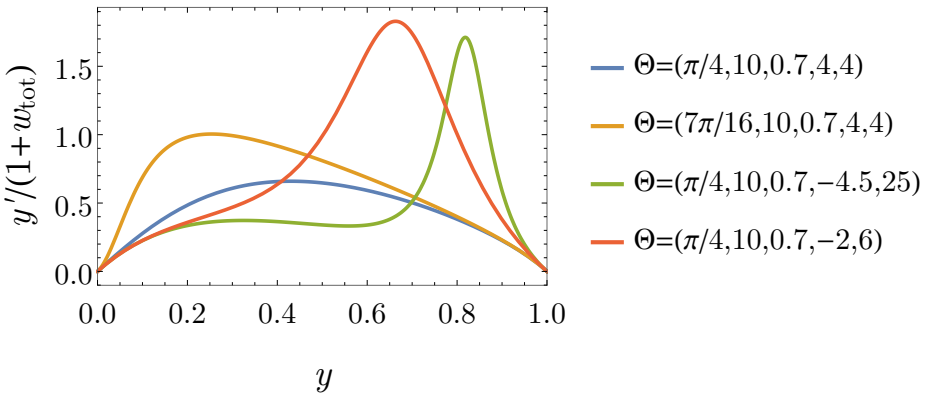


Figure 3.1: Examples of the phase space of y (the ratio of the scale factors). Here, prime denotes $d/d\ln a$. Recall, $\Theta \equiv (\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta)$.

Using eq. (3.24), we get,

$$w_{\rm DE} = -1 - \frac{m_{\rm eff}^2}{3\Omega_{\rm DE}} \left(1 + \frac{1}{\tan^2\theta\, y^2}\right)^{-1} \frac{d\ln y}{d\ln a}. \tag{3.30}$$

As the next step, we solve the cosmological equations of motion (3.14) and (3.18).

3.2 Solutions

The Universe is currently expanding, as shown by observations of type Ia supernovae [16, 17]. Here, we assume that the Universe has always been expanding, that is $H > 0$.[3] This means that the scale factor, a, is increasing with time and accordingly the matter and radiation densities dilute as the Universe evolves.

Equation (3.18) is a quartic polynomial in y, so it has up to four real solutions. We can classify them according to their behavior in the early universe, that is when $\Omega_{\rm tot} \to \infty$ (i.e., $z \to \infty$) [67]. Looking at eq. (3.18), there are two possible solutions, either $y \to 0$ or $y \to \infty$ as $\Omega_{\rm tot} \to \infty$. They are referred to as the finite branch and infinite branch solution, respectively. They can be visualized in the phase space of y as shown in Fig. 3.2. Since the points where $dy/d\ln a = 0$ cannot be crossed dynamically, they set the borders between the branches. This implies that the sign of $dy/d\ln a$ is the same on each branch

[3] In some regions of the bimetric parameter space, there are also exotic branches, for example bouncing cosmologies where the scale factor decreases to then turn around and start increasing, hence violating our assumption of expanding cosmology.

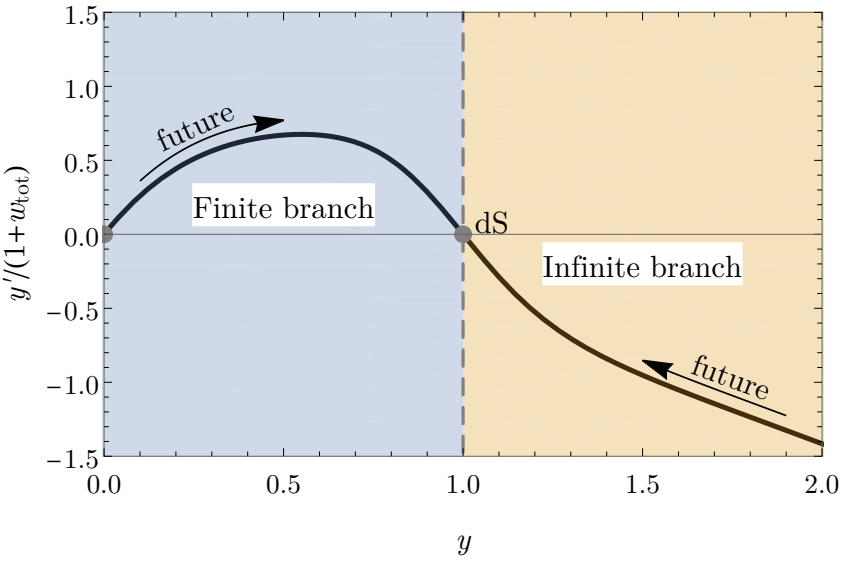


Figure 3.2: Phase space of y (the ratio of the scale factors). Here, $y' \equiv dy/d\ln a$ and $(\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta) = (45°, 10, 0.7, 0, 6)$. The solution branches are marked in different colors. The gray dots indicate when $y' = 0$, separating the branches. The arrows mark the future time direction on each branch, assuming expanding cosmology.

(unless it diverges at some point). From eq. (3.27), we see that there are two ways to satisfy $dy/d\ln a = 0$, either $y = 0$ or $\Omega_{\rm tot} = 0$. The latter case corresponds to a de Sitter solution. The finite branch cosmology starts with the Big Bang at $y = 0$. Thereafter, y increases monotonically until it hits the de Sitter point, $y = 1$, in future infinity. The infinite branch solution starts with the Big Bang at $y = \infty$ after which y decreases monotonically and approaches the de Sitter point towards the infinite future.

Infinite branch (inconsistent)

In this branch, y spans an infinite range, starting from infinity at the Big Bang, then decreasing monotonically until reaching the final de Sitter point in the asymptotic future. Hence,

$$\frac{dy}{d\ln a} < 0, \quad \text{(infinite branch)}. \tag{3.31}$$

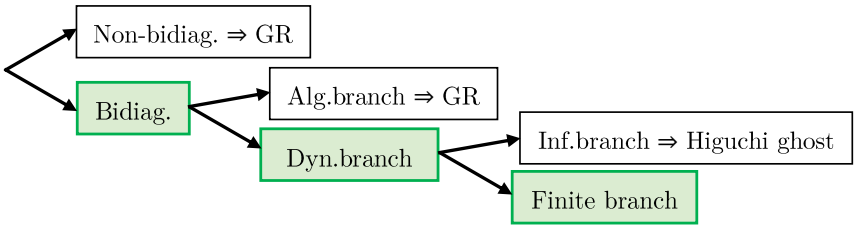


Figure 3.3: Flow chart for the cosmological background solutions, starting from a homogeneous and isotropic ansatz for the physical metric, $g_{\mu\nu}$, and assuming that the second metric has the same spherical symmetry as $g_{\mu\nu}$, according to eq. (3.1). The green branches are viable and provide corrections to the background cosmology of GR.

This implies that the dark energy density, $\Omega_{\rm DE}$, decreases with time on this branch (cf. eq. (3.24)). From eq. (3.18) we read off (assuming $\beta \neq 0$),

$$y \simeq \left(-\frac{3\Omega_{\rm tot}}{\sin^2\theta\, m_{\rm FP}^2 \beta}\right)^{1/3}, \quad \text{(early universe)}. \tag{3.32}$$

Thus, to have a real solution in the early universe, we require $\beta \leq 0$ (corresponding to $B_3 \leq 0$). This implies that there is no Vainshtein screening mechanism, as we will see in Chapter 4. Moreover, as we show below, due to eq. (3.31), this branch excites the Higuchi ghost and is hence not viable. Therefore, we focus on finite branch solutions henceforth. In Fig. 3.3, we show a flow chart of different solution branches for the bimetric background cosmology.

Finite branch (consistent)

In this branch, y spans a finite range, starting at $y = 0$ at the Big Bang, then increasing monotonically with time until the final de Sitter point is reached in the infinite future, that is, when $y = 1$. Thus,

$$\frac{dy}{d\,\ln a} > 0, \quad \text{(finite branch)}, \tag{3.33}$$

and hence the dark energy density, $\Omega_{\rm DE}$, increases with time (cf. eq. (3.24)), see Fig. 3.4 for some examples. In other words, $\Omega_{\rm DE}$ is a phantom dark energy. However, since the equation of state approaches $w_{\rm DE} = -1$ sufficiently fast in the future, the Big Rip is avoided, see Ref. [96] and Paper III.

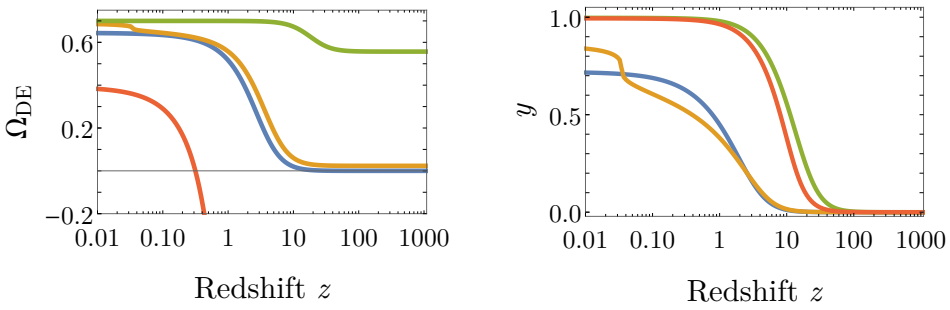


Figure 3.4: Examples of $\Omega_{\rm DE}(z)$ and $y(z)$. The evolution of these quantities varies significantly depending on the physical parameters. $\Omega_{\rm DE}$ approaches Ω_Λ in the late universe. Both $\Omega_{\rm DE}$ and y decrease monotonically with redshift. Remember, $\Theta \equiv (\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta)$. Blue curves: $\Theta = (18°, 1.2, 0.70, 1.0, 10.3)$. This is a self-accelerating model, hence $\Omega_{\rm DE}$ vanishes in the early universe. Yellow curves: $\Theta = (19°, 2.3, 0.74, -2.4, 8.5)$. This is a model with parameters close to the dynamical Higuchi bound, giving rise to a sharp transition in y and $\Omega_{\rm DE}$ at $z \simeq 0.03$. This can be understood from eq. (3.27). Green curves: $\Theta = (0.6°, 10, 0.70, 10, 10)$. Red curves: $\Theta = (45°, 10, 0.70, 10, 10)$. In the early universe, this model approaches $\Omega_{\rm DE}|_{z\to\infty} = -716$.

The cosmological constant in the asymptotic future is Ω_Λ, as seen from eq. (3.13). For the (normalized) expansion rate to be real-valued in the late universe where $E^2 \simeq \Omega_\Lambda$, we demand,

$$\Omega_\Lambda > 0. \tag{3.34}$$

Expanding eq. (3.18) around large $\Omega_{\rm tot}$ (i.e., early universe), we see that,

$$y \simeq \frac{\cos^2\theta\, m_{\rm FP}^2(1+2\alpha+\beta)}{3\Omega_{\rm tot}}, \quad \text{(early universe)}. \tag{3.35}$$

Therefore, for $\Omega_{\rm tot}$ to be positive, we require,

$$1 + 2\alpha + \beta > 0, \tag{3.36}$$

or, equivalently, $B_1 > 0$.

In the early universe (in the limit $y \to 0$), the bimetric dark energy is,

$$\Omega_{\rm DE}|_{z\to\infty} = \Omega_\Lambda - \sin^2\theta\, m_{\rm FP}^2(3+3\alpha+\beta)/3, \tag{3.37}$$

and gives a cosmological constant contribution to the Friedmann equation (3.14) since $w_{\rm DE}|_{z\to\infty} = -1$. See Fig. 3.5 for some examples of the equation of

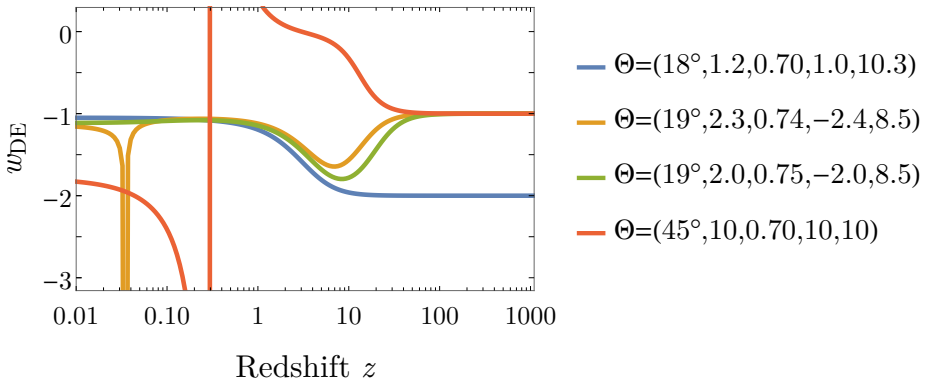


Figure 3.5: Examples of the equation of state for the bigravity fluid. The resulting evolution of $w_{\rm DE}$ varies significantly depending on the value of the physical parameters. However, all models approach -1 in the asymptotic future. In the past, they also approach -1, except from the self-accelerating model (blue) which approaches $-(2+w_{\rm tot})$. For the red curve, $\Omega_{\rm DE}$ is negative in the early universe. Hence, $w_{\rm DE}$ diverges as $\Omega_{\rm DE}$ passes through zero. There is no physical singularity at this point. Both the yellow and green models have a pronounced phantom dark energy phase around $z \simeq 10$. The physical parameters of the yellow model are close to the dynamical Higuchi bound and it also has a very brief phantom phase around $z \simeq 0.03$. The equation of state does not diverge here but has a sharp dip, reaching a minimum value $w_{\rm DE,min} = -12$.

state. From eq. (3.37) we see that the dark energy density in the early universe achieves negative values for large $m_{\rm FP}$, α, or β, see Fig. 3.4 for some examples. The exceptions are the self-accelerating models, defined by a vanishing dark energy density in the early universe (i.e., $\beta_0 = 0$, cf. eqs. (2.14a) and (3.37)). In this case, $w_{\rm DE}|_{z\to\infty} = -(2+w_{\rm tot})$. For these models, there is no cosmological constant term in the action (2.2), so the accelerated expansion of the Universe is only due to the interaction of the massive spin-2 field.

Since $\Omega_{\rm DE}$ approaches a cosmological constant contribution both in the late universe and in the early universe (except for self-accelerating cosmologies), there are two ΛCDM phases in bigravity cosmology. It should be stressed that the redshift ranges of these phases (and hence what is meant by "early" and "late") are determined by the physical parameters. In the transition between these two regions the expansion history can differ considerably from ΛCDM. Due to the richness of the interaction between the spin-2 fields, this transition can take many different forms depending on the physical parameters.

From the equation of state (3.30), we see that $w_{\rm DE} > -1$ when $\Omega_{\rm DE} < 0$ and $w_{\rm DE} < -1$ when $\Omega_{\rm DE} > 0$. This means that the standard energy conditions (i.e., weak, strong, null, and dominant energy conditions) are violated by the bimetric stress–energy. The violation of the dominant energy condition means that the energy density of the bimetric stress–energy $V^{\mu}{}_{\nu}$ flows faster than the speed of light. However, there is no problem with causality since the null cone of the $f_{\mu\nu}$-metric is wider than the $g_{\mu\nu}$-metric (see below), so there is no superluminal propagation with respect to the overall causal cone [97].

Null cones. We calculate the null cones of the two metrics in conformal time, τ, defined by $a\,d\tau = dt$. It is straightforward to show that the slope of the g-null cone is 45° in the τr-plane, that is, $dr_g(\tau)/d\tau = \pm 1$ where $r = r_g(\tau)$ are the radial null geodesics of $g_{\mu\nu}$. In these coordinates, the f-metric reads,

$$ds_f^2 = c^2\widetilde{a}^2\left[-x^2 d\tau^2/y^2 + dr^2 + r^2 d\Omega^2\right], \tag{3.38}$$

hence the slope of the f-null cone is,

$$\frac{dr_f(\tau)}{d\tau} = \pm\frac{x}{y}, \tag{3.39}$$

in the τr-plane. The relative width of the f-null cone, compared with the g-null cone, is,

$$\frac{dr_f}{dr_g} = \frac{x}{y} = 1 + \frac{d\,\ln y}{d\,\ln a}, \tag{3.40}$$

where we used the identity (3.8) to get the last equality. See Fig. 3.6 for some examples. Since y is increasing with a on the finite branch, eq. (3.40) means that the f-null cone is always wider than the g-null cone, hence allowing for superluminal propagation while still being inside the overall causal cone. In the early universe (expanding around $y = 0$),

$$\frac{dr_f}{dr_g} = 4 + 3w_{\rm tot} + \mathcal{O}(y), \quad y \simeq 0, \tag{3.41}$$

while in the future de Sitter phase, $d\,\ln y/d\,\ln a = 0$, and the null cones coincide. Eq. (3.41) implies that the f-null cone is five times as wide as the g-null cone in the early universe where radiation is dominating (i.e., $w_{\rm tot} = w_r = 1/3$). In Chapter 5, we show that some of the gravitational wave modes propagate on the f-null cone in the early universe (i.e., faster than the speed of light).

A wider casual cone has potential consequences for the cosmological horizon problem which is the problem of explaining how CMB photons from seemingly causally disconnected regions of the sky can have the same temperature. Further work is needed to prove the effectiveness of this mechanism, in particular

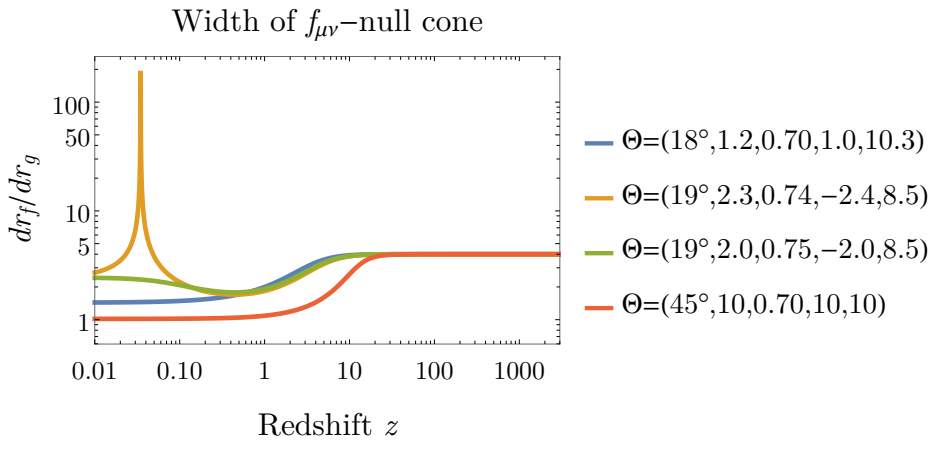


Figure 3.6: Width of the $f_{\mu\nu}$-null cone (compared with the $g_{\mu\nu}$-null cone) as a function of redshift. The physical parameters are the same as in Fig. 3.5. In the plot range, the Universe is matter dominated ($w_{\text{tot}} \simeq 0$), so the $f_{\mu\nu}$-null cones are four times as wide as the $g_{\mu\nu}$-null cone at large redshifts. In the infinite future, the null cones coincide and the function approaches unity, although some of the models are far away from this state at present-day. The yellow curve has a sharp and brief peak around $z \simeq 0.03$ where the ratio of the scale factor increases rapidly, cf. Fig. 3.4.

to show that this faster than light propagation affects the temperature distribution of the cosmic fluid, which is not obvious. In any case, the fivefold increase in the causal cone can only serve to widen the causally connected region of CMB photons from $\simeq 2°$ to $\simeq 10°$, which is not enough to cover the whole sky and thus solve the horizon problem.

3.3 Dynamical Higuchi bound

Bimetric gravity was constructed to avoid the Boulware–Deser ghost, plaguing generic theories of massive gravity. However, this does not preclude the appearance of other types of ghosts or instabilities around certain background solutions. It is well-known that the Higuchi ghost is excited for small enough values of the graviton mass around a de Sitter background [98]. This ghost mode is propagating on a de Sitter background, rendering these solutions unstable. This is opposed to the Boulware–Deser ghost which would make the theory itself unstable, independent on any particular solution. Nevertheless, since our cosmological models approach a de Sitter phase as $t \to \infty$, the

Higuchi ghost should be avoided, imposing a lower bound the on graviton mass,

$$m_{\rm FP}^2 > 2\Omega_\Lambda. \tag{3.42}$$

This can be generalized to dynamical background cosmologies (including spatial curvature) as shown in Refs. [89, 99],

$$m_{\rm eff}^2 > 2\widetilde{\Omega}_{\rm DE}. \tag{3.43}$$

In the final de Sitter phase where $m_{\rm eff} = m_{\rm FP}$ and $\widetilde{\Omega}_{\rm DE} = \Omega_{\rm DE} = \Omega_\Lambda$, it reduces to the ordinary Higuchi bound (3.42). Alternatively, we can use the identity (3.25) to rewrite eq. (3.43) as,

$$\frac{d\Omega_{\rm tot}}{dy} < 0. \tag{3.44}$$

Now, $\Omega_{\rm tot}$ can be written as a function of y and the physical parameters. Analyzing this expression and demanding eq. (3.44) to be satisfied in the range $0 \leq y \leq 1$ (i.e., the range of the finite branch solution), one obtains a set of necessary and sufficient conditions on the physical parameters for eq. (3.44) to be satisfied. The complete expressions can be found in Appendix A and more details of the derivation can be found in Paper III. In Fig. 3.7, we show the region of parameter space excluded by demanding the absence of the Higuchi ghost and how it varies with the value of θ and $m_{\rm FP}^2/(2\Omega_\Lambda)$. In the figure, we also included the constraint from demanding a positive energy density in the early universe, eq. (3.36).

As we have seen, $dy/d\ln a > 0$ is a consistency condition for the finite branch solutions. However, at a generic point in the parameter space, we are not guaranteed that this condition is satisfied at all redshifts, even if it is satisfied at some initial time. As an example, it can happen that the solution is well-behaved at $y = 0$ (i.e., at the Big Bang) and then as the Universe evolves encounters a pole (i.e., y diverging). If we plug the dynamical Higuchi bound (3.43) into the identity (3.27), we see that $dy/d\ln a$ is positive and well-behaved whenever eq. (3.43) is satisfied. In other words, the dynamical Higuchi bound provides a requirement on the physical parameters which guarantees that the finite branch solution exists at all redshifts. The infinite branch solutions violate the dynamical Higuchi bound since in that case, $dy/d\ln a < 0$ (cf. eq. (3.44)).

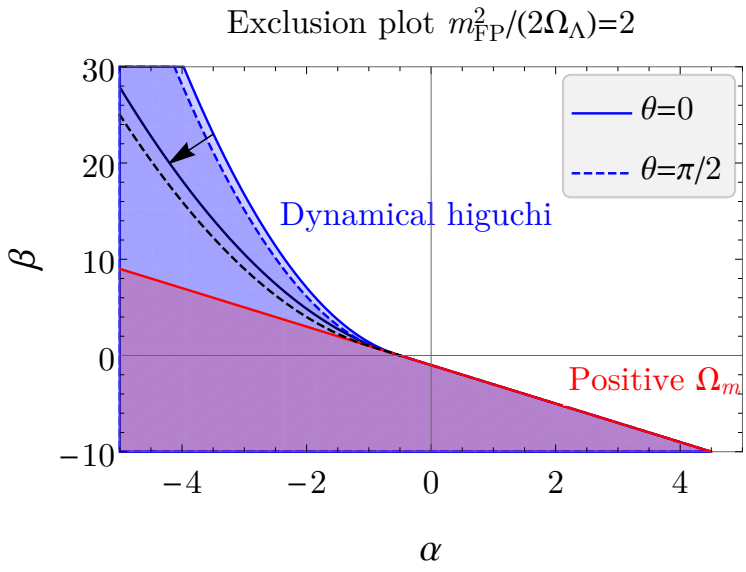


Figure 3.7: Exclusion plot in the $\alpha\beta$-plane due to the cosmological viability conditions, that is, a positive matter density (3.36) and the dynamical Higuchi bound (3.43). The solid blue curve is the boundary in case $\theta = 0$ while the dashed blue curve is the boundary in case $\theta = \pi/2$, showing that the constraint depends only weakly on θ. The black curve shows the boundary in the limit $m^2_{\rm FP}/(2\Omega_\Lambda) \to \infty$ and the arrow shows how the boundary moves as this limit is approached. Figure reprinted from Paper III under the terms of Creative Commons Attribution 4.0 license.[4]

3.4 General relativity limits

In Chapter 2, we saw what,

$$\theta \to 0, \tag{3.45}$$

is the GR limit for the full bigravity theory. Nevertheless, for special solutions there may exist additional limits, where the bimetric solutions approach GR solutions. In the case of background cosmology, this translates to the Friedmann equation (3.14) reducing to the ordinary Friedmann equation,

$$E^2 = \Omega_\Lambda + \Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4, \tag{3.46}$$

for some positive constants $(\Omega_\Lambda, \Omega_{m,0}, \Omega_{r,0})$. To identify such limits in the bimetric parameter space, we start by expanding eq. (3.14) around $\Omega_{\rm tot} = 0$

[4]See https://creativecommons.org/licenses/by/4.0/.

(i.e., around the final de Sitter point),

$$E^2 = \Omega_\Lambda + \frac{\kappa_g^{\text{eff}}}{\kappa_g}\Omega_{\text{tot}} + (1+2\alpha)\frac{\sin^2\theta}{2m_{\text{FP}}^2}\frac{2\Omega_\Lambda/m_{\text{FP}}^2}{(1-2\Omega_\Lambda/m_{\text{FP}}^2)^3}\Omega_{\text{tot}}^2 + \mathcal{O}\left(\Omega_{\text{tot}}^3\right), \quad \Omega_{\text{tot}} \ll 1, \tag{3.47}$$

where we have defined an effective gravitational constant,

$$\kappa_g^{\text{eff}} \equiv \kappa_g\left(1 - \frac{\sin^2\theta}{1-2\Omega_\Lambda/m_{\text{FP}}^2}\right). \tag{3.48}$$

The same expansion can be found for example in Ref. [52]. An expansion up to order Ω_{tot}^4 can be found in Paper III. The idea with the expansion (3.47) is that any GR limit for the cosmological background solutions should, as a necessary condition, approach GR at the final de Sitter phase when $\Omega_{\text{tot}} = 0$. To agree with the GR Friedmann equation (3.46), we must set all quadratic terms and higher in the expression (3.47) equal to zero. The cases $\alpha \to \infty$ and $\beta \to \infty$ must be treated separately since the higher-order terms in eq. (3.47) diverge in these limits. Thus, we identify the possible GR limits,

$$\theta \to 0, \tag{3.49a}$$
$$(\alpha, \beta) \to (-1/2, 0), \tag{3.49b}$$
$$m_{\text{FP}} \to \infty, \tag{3.49c}$$
$$\alpha \to \infty, \tag{3.49d}$$
$$\beta \to \infty, \tag{3.49e}$$

for the cosmological background solutions. In case (3.49b), $B_1 = B_3 = 0$ and the background cosmology reduces to ΛCDM with rescaled matter density and cosmological constant as shown in Ref. [65]. However, this limit saturates the matter positivity constraint eq. (3.36) and there is a singularity in the past, before reaching the Big Bang. Moreover, the requirement of a working screening mechanism excludes also the neighborhood around this point, see Chapter 4.

Large m_{FP}, α, β limits

As we show in Paper III, it is possible to make series expansions around each of $m_{\text{FP}}^{-1}, \alpha^{-1}, \beta^{-1} = 0$. To first order, the result is the same in all cases, namely,

$$E^2 = \Omega_\Lambda + \cos^2\theta\,\Omega_{\text{tot}} + \text{h.o.}, \tag{3.50a}$$
$$\Omega_{\text{tot},0} = (1-\Omega_\Lambda)/\cos^2\theta + \text{h.o.}, \tag{3.50b}$$
$$\Omega_{\text{DE}} = \Omega_\Lambda - \sin^2\theta\,\Omega_{\text{tot}} + \text{h.o.}, \tag{3.50c}$$

where h.o. stands for higher order corrections. The higher order terms in eq. (3.50a) grow faster with redshift than the first order term. Hence, the expansion is valid up to a certain redshift. In fact, the range of applicability of eq. (3.50a) increases with increasing values of $m_{\rm FP}$, α, or β. See Paper III for more details.

Recall that $\Omega_{\rm tot}$ is the sum of matter and radiation, $\Omega_{\rm tot} = \Omega_m + \Omega_r$. The present-day value of the radiation density is set by the CMB temperature and neutrino physics (which we assume is standard). We set [43],

$$\Omega_{r,0}h^2 = 4.4 \times 10^{-5}, \tag{3.51}$$

that is, $\Omega_{r,0} = 9.0{\times}10^{-5}$ (assuming $h = 0.7$), corresponding to matter-radiation equality at $z_{\rm eq} \simeq 3350$ for $\Omega_{m,0} \simeq 0.3$.

From eq. (3.50b), we get,

$$\Omega_{m,0} = -\Omega_{r,0} + (1 - \Omega_\Lambda)/\cos^2\theta + \text{h.o.} \tag{3.52}$$

Since $\Omega_{m,0} \gg \Omega_{r,0}$ for viable models,

$$\Omega_{m,0} \simeq (1 - \Omega_\Lambda)/\cos^2\theta + \text{h.o.} \tag{3.53}$$

So, for a fixed cosmological constant Ω_Λ, the present-day matter density is rescaled with a factor $1/\cos^2\theta$ compared with a ΛCDM model with the same expansion history. Interestingly, from eq. (3.50c), we see that the dark energy imitates the dominating matter component (but with a negative energy density) plus a cosmological constant, see Fig. 3.8. This is the reason for the rescaling between the cosmological constant and the present-day matter density in eq. (3.53). Summarizing, in the limits studied here, the expansion rate is,

$$E^2 = \Omega_\Lambda + (1 - \Omega_\Lambda - \cos^2\theta\,\Omega_{r,0})(1+z)^3 + \cos^2\theta\,\Omega_{r,0}(1+z)^4 + \text{h.o.} \tag{3.54}$$

That is, a Friedmann equation.

However, as mentioned, the redshift range where the expansion (3.54) is applicable is always finite for any finite values of $(m_{\rm FP}, \alpha, \beta)$. But what happens at earlier times when (3.54) is not valid? Remember that there is always a redshift, z_t, before which the dark energy contribution is subdominant, which follows from $w_{\rm DE} \to -1$ as $y \to 0$. That is, $\Omega_{\rm DE}$ approaches a cosmological constant contribution at early enough times and is thus subdominant compared with matter or radiation. At redshifts earlier than z_t (i.e., $z \gg z_t$), the dark energy is subdominant and we get a Friedmann equation again,

$$E^2 = \Omega_\Lambda + \frac{1 - \Omega_\Lambda}{\cos^2\theta}(1+z)^3 + \Omega_{r,0}(1+z)^4, \quad z \gg z_t. \tag{3.55}$$

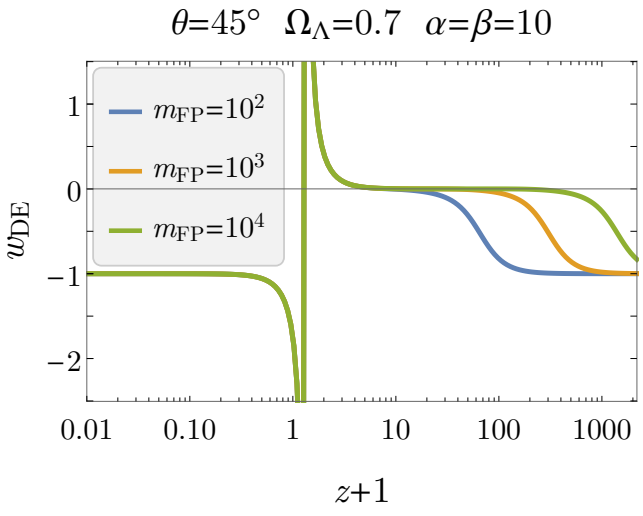


Figure 3.8: Equation of state of the dark energy as a function of redshift. For large values of $m_{\rm FP}$, $\Omega_{\rm DE}$ acts a cosmological constant plus an ordinary matter component (but with negative energy density). This can be seen as $w_{\rm DE} = -1$ in the late universe and $w_{\rm DE} = w_{\rm tot} = 0$ during the matter-dominated era. Increasing $m_{\rm FP}$ extends the $w_{\rm DE} = w_{\rm tot}$ phase to higher redshifts. Accordingly, z_t is also pushed to higher redshifts in this limit. Similar behavior is observed in the large α and β limits. The divergence at $z \sim 1$ is due to $\Omega_{\rm DE}$ crossing zero. There is no physical divergence at this point.

Increasing values of $m_{\rm FP}$, α, or β pushes the transition z_t backwards in time, see Fig. 3.8.

We quantify the difference between the bimetric expansion history and ΛCDM by the difference in their squared expansion rates,

$$\Delta E^2 \equiv E^2 - E^2_{\Lambda\rm CDM}. \tag{3.56}$$

The difference vanishes at present-day due to the normalization, $E|_{z=0} = 1$. At early enough times ($z \gg z_t$), the dark energy density is subdominant and matter or radiation gives the dominant contribution to the expansion rate. In this case, the relative difference in the (squared) expansion rate is,

$$\frac{\Delta E^2}{E^2_{\Lambda\rm CDM}} = \frac{\Omega_{\rm tot} - \Omega^{\Lambda\rm CDM}_{\rm tot}}{\Omega^{\Lambda\rm CDM}_{\rm tot}} = \frac{\Omega_{\rm tot}(z)}{\Omega^{\Lambda\rm CDM}_{\rm tot}(z)} - 1, \quad z \gg z_t. \tag{3.57}$$

We split into two cases depending on whether $\Omega_{\rm DE}$ becomes subdominant be-

fore or after matter-radiation equality, that is, whether $z_t < z_{\rm eq}$ or $z_t > z_{\rm eq}$.[5]

Case $z_t < z_{\rm eq}$. If $(m_{\rm FP}, \alpha, \beta)$ are not too large, $\Omega_{\rm DE}$ becomes subdominant during the matter-dominated era. In Fig. 3.9, this corresponds to the cases $m_{\rm FP} = 10$ and $m_{\rm FP} = 10^3$. In the region $z < z_t$, both Ω_m and $\Omega_{\rm DE}$ contribute as matter components in such a way that the expansion follows a ΛCDM model with cosmological constant Ω_Λ, according to eq. (3.54). Hence,

$$\Delta E^2/E^2_{\Lambda\rm CDM} \simeq 0, \quad z < z_t. \tag{3.58}$$

After the transition but before matter-radiation equality, $z_t < z < z_{\rm eq}$, the dark energy becomes subdominant and only Ω_m contributes. Due to the rescaled present-day matter density in eq. (3.53), there is a temporary phase of increased expansion according to,

$$\frac{\Delta E^2}{E^2_{\Lambda\rm CDM}} \simeq \frac{(1-\Omega_\Lambda)/\cos^2\theta}{1-\Omega_\Lambda} - 1 = \tan^2\theta, \quad z_t < z < z_{\rm eq}. \tag{3.59}$$

The increased expansion continues to matter-radiation equality. During radiation domination, the expansion follows ΛCDM again and,

$$\Delta E^2/E^2_{\Lambda\rm CDM} \simeq 0, \quad z > z_{\rm eq}. \tag{3.60}$$

Recall, $\Omega_{r,0}$ is fixed by the observed CMB temperature according to eq. (3.51). From eqs. (3.58)-(3.60), it follows that $\tan^2\theta$ is a maximum value for how much the expansion rate is increased in the region $z_t < z < z_{\rm eq}$. Increasing $m_{\rm FP}$, α, or β pushes the transition closer to $z_{\rm eq}$, implying that the maximum value is not reached before radiation becomes influential, see Fig. 3.9.

Case $z_t > z_{\rm eq}$. If $m_{\rm FP}$, α, or β are large enough, the dark energy density becomes subdominant after matter-radiation equality. This corresponds to the cases $m_{\rm FP} = 10^8$, $m_{\rm FP} = 10^{12}$, and $m_{\rm FP} = 10^{16}$ in Fig. 3.9. During matter domination, Ω_m and $\Omega_{\rm DE}$ contribute as matter components in such a way that the expansion follows a ΛCDM model with cosmological constant Ω_Λ, hence,

$$\Delta E^2/E^2_{\Lambda\rm CDM} \simeq 0, \quad z < z_{\rm eq}. \tag{3.61}$$

However, when radiation becomes dominant but before the dark energy becomes subdominant, $z_{\rm eq} < z < z_t$, $\Omega_{\rm DE}$ contributes as radiation with negative energy density according to eq. (3.50c) and,

$$\frac{\Delta E^2}{E^2_{\Lambda\rm CDM}} \simeq \frac{\cos^2\theta\,\Omega_{r,0}(1+z)^4}{\Omega_{r,0}(1+z)^4} - 1 = -\sin^2\theta, \quad z_{\rm eq} < z < z_t. \tag{3.62}$$

[5]The redshift at matter-radiation equality, $z_{\rm eq}$, is set by $\Omega_{m,0}$, which depends on θ according to eq. (3.53).

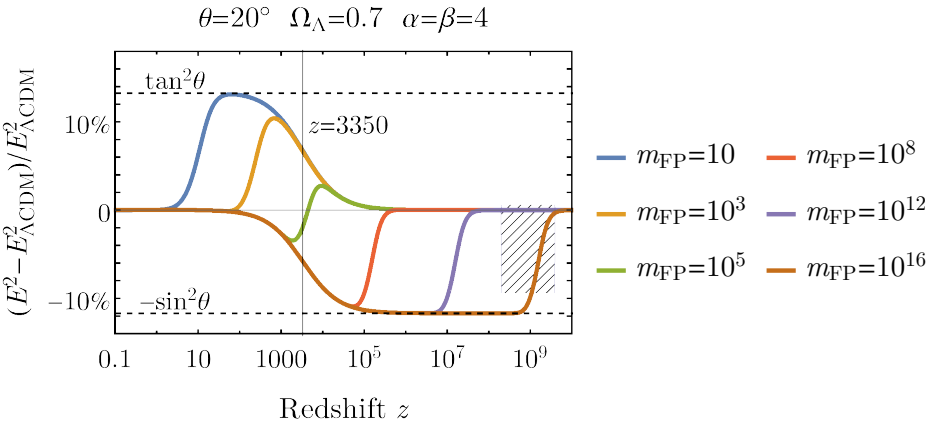


Figure 3.9: Relative difference between the (squared) expansion rate of the bigravity model and the cosmological concordance model. We recognize the late- and early-universe phases where the expansion agrees. The transition between these phases is pushed backwards in time by pushing $m_{\rm FP}$ to larger values. Similar behavior is observed in the $\alpha, \beta \to \infty$ limits. Constraints from Big Bang nucleosynthesis requires the expansion rate to be in the hatched region (see Section 3.6), so the $m_{\rm FP} = 10^{16}$ model violates this.

In this phase, the BR model has a slower expansion rate, set by $-\sin^2\theta$, see Fig. 3.9.

When dark energy becomes subdominant (i.e., $z \gg z_t$), the only remaining contribution to the expansion is due to radiation and thus,

$$\frac{\Delta E^2}{E^2_{\Lambda\rm CDM}} \simeq \frac{\Omega_{r,0}(1+z)^4}{\Omega_{r,0}(1+z)^4} - 1 = 0, \quad z \gg z_t, \tag{3.63}$$

so the bimetric model again follows the cosmological concordance model. Since z_t increases with increasing $m_{\rm FP}$, α, or β, the phase of slower expansion is extended in the limits where these parameters are large, see Fig. 3.9.

3.5 Special models

The bimetric parameter space is five-dimensional, corresponding to the five physical parameters. In the literature, submodels are typically studied, that is, models obtained by setting one or several of the B-parameters (or β-parameters) to zero. This makes the physical parameters depend on each other [68]. We will refer to a model which has n free parameters as an n-parameter model (e.g., a three-parameter model). It seems reasonable to demand that at least the mixing angle, graviton mass, and cosmological constant are independent, meaning that one can set at most two of the B-parameters to zero at the same time. This is confirmed by the combined cosmological constraints (as presented in this chapter) together with the constraints due to the requirement of a working Vainshtein screening mechanism (see Chapter 4 for details). Nevertheless, two-parameter models are common in the literature and we review one example below.

Two-parameter models. As an example, consider a B_1B_2-model (i.e., $B_0 = B_3 = B_4 = 0$). For more examples, see Ref. [68]. From eqs. (2.13)-(2.14), we deduce,

$$m_{\rm FP}^2 = \frac{1}{2}(\csc^2\theta + 3\sec^2\theta)\Omega_\Lambda, \tag{3.64a}$$

$$\alpha = \frac{1 - 3\tan^2\theta}{1 + 3\tan^2\theta}, \tag{3.64b}$$

$$\beta = 0, \tag{3.64c}$$

so the graviton mass depends on the mixing angle and the cosmological constant. In particular, in the GR limit $\theta \to 0$, the graviton mass goes to infinity. Moreover, eq. (3.64a) guarantees that the Higuchi bound $m_{\rm FP}^2 > 2\Omega_\Lambda$ is satisfied. The B-parameters can be expressed in terms of the physical parameters as,

$$B_1 = \frac{3}{2}(1 - \tan^2\theta)\Omega_\Lambda, \tag{3.65a}$$

$$B_2 = -\frac{1}{2}(1 - 3\tan^2\theta)\Omega_\Lambda. \tag{3.65b}$$

To have a positive matter density in the early universe, we need $B_1 > 0$, that is, $\theta < \pi/4$ (see eq. (3.36)). However, since $B_3 = 0$, there is no working Vainshtein screening mechanism for this model, see Chapter 4.

For the two-parameter models, we can predict analytically the confidence contours in the parameter space without making any statistical data analysis, assuming that the best-fit cosmology is close to a ΛCDM model [10]. We set

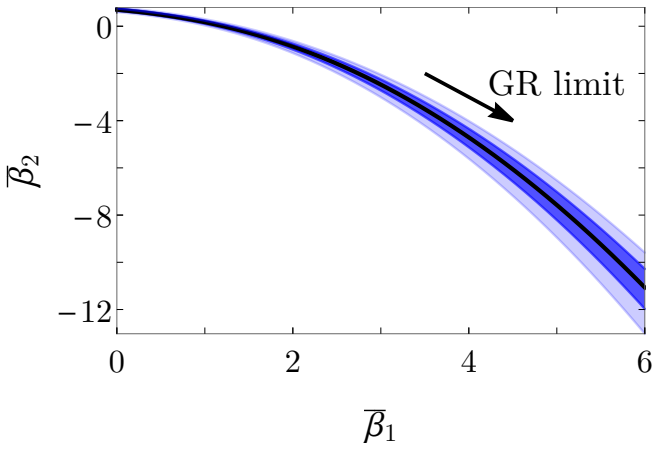


Figure 3.10: Predicted confidence contours of the B_1B_2-model. The black curve is traced out by fixing $\Omega_\Lambda = 0.7$ and letting $\theta \to 0$. The dark blue region is defined by $\Omega_\Lambda = 0.70 \pm 0.05$ and the light blue region by $\Omega_\Lambda = 0.70 \pm 0.10$. To facilitate comparison with Ref. [67], we have introduced the $\bar{\beta}$-parameters which can be identified with the β-parameter of Ref. [67]. They are related to the B-parameters as $\bar{\beta}_n = B_n/\tan^n\theta$. Figure reprinted from Paper IV under the terms of Creative Commons Attribution 4.0 license.

$\Omega_\Lambda = 0.7$ and consider the curve $\theta \to 0$ approaching the GR limit in the parameter space. With our assumptions, the best-fit cosmology should lie in the neighborhood of this curve. An example is shown in Fig. 3.10. The result is in agreement with statistical data analysis of Refs. [66, 67]. In this model, the GR limit corresponds to $(B_1, B_2) \to (1.05, -0.35)$, assuming $\Omega_\Lambda = 0.7$.

Three-parameter models. In Chapter 4, we show that (B_1, B_2, B_3) must all be nonzero to satisfy the cosmological viability constraints presented in this chapter and the constraints guaranteeing a working screening mechanism. Hence, the $B_1B_2B_3$-models (i.e., $B_0 = B_4 = 0$) are the most minimal models compatible with these constraints. From the Hassan–Rosen action (2.2) and the definition of the elementary symmetric polynomials (2.3), we see that B_0 and B_4 are the cosmological constant terms of the $g_{\mu\nu}$- and $f_{\mu\nu}$-metrics, respectively. Setting the cosmological constants to zero gives rise to self-accelerating cosmologies where the accelerated expansion of the Universe is due to the interaction between the metrics.

Setting $B_0 = B_4 = 0$, we can solve for α and β in terms of the other

parameters,

$$\alpha = \frac{-m_{\rm FP}^2/\Omega_\Lambda + 3\cot^2\theta - 3\tan^2\theta}{4m_{\rm FP}^2/\Omega_\Lambda}, \tag{3.66a}$$

$$\beta = 3\frac{-2m_{\rm FP}^2/\Omega_\Lambda + \csc^2\theta + 3\sec^2\theta}{4m_{\rm FP}^2/\Omega_\Lambda}. \tag{3.66b}$$

For the $B_1B_2B_3$-models, an additional symmetry appears in the limit where these parameters vanish. Then, the metrics decouple and two copies of diffeomorphism symmetry appear, one for each metric, which means that the action stays invariant under independent coordinate transformations acting on each metric separately. This means that small values of (B_1, B_2, B_3), and hence also the effective cosmological constant (Ω_Λ), are technically natural in the sense of 't Hooft and the value of Ω_Λ is protected against quantum corrections [100].

Four-parameter models. The four-parameter models with viable cosmology and working screening mechanism (i.e., $B_1, B_2, B_3 \neq 0$) are $B_0B_1B_2B_3$ and $B_1B_2B_3B_4$. The latter deserves extra attention since it presents self-accelerating cosmological solutions without a cosmological constant for the physical metric. Setting $B_0 = 0$, we obtain a linear relation between α and β (see eq. (2.14a)),

$$\alpha + \frac{\beta}{3} = -1 + \frac{1}{\sin^2\theta}\frac{\Omega_\Lambda}{m_{\rm FP}^2}, \quad (B_1B_2B_3B_4 \text{ models}). \tag{3.67}$$

Using this relation and demanding the dynamical Higuchi bound to be satisfied, one can derive the constraint,

$$\sin^2\theta \lesssim 2\Omega_\Lambda/m_{\rm FP}^2, \quad \text{(self-acc.)}, \tag{3.68}$$

see Paper IV. The result is plotted in Fig. 3.11. For the theory to contribute with significant modifications to GR, the mixing angle must be nonzero (which is the GR limit). From Fig. 3.11, we see that this requires $m_{\rm FP} \sim 1$, that is, that the graviton mass is of the order of the present-day Hubble scale, assuming $\Omega_\Lambda \simeq 0.7$.

3.6 Observational constraints

A viable theory of gravity must pass observational tests including cosmological observations of the cosmic microwave background (CMB), baryon acoustic oscillations (BAO), type Ia supernovae (SNIa), and the observed abundance of light elements. These data sets can be used to constrain the free parameters of the theory, thereby focusing in on the viable part of the parameter space.

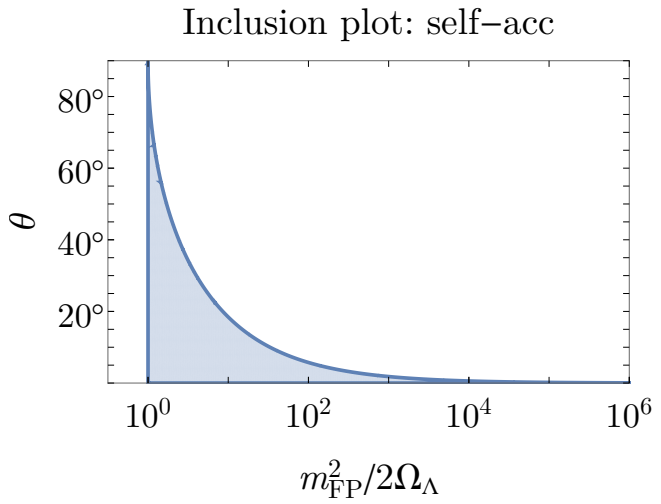


Figure 3.11: Inclusion plot for self-accelerating models due to requiring a consistent cosmology satisfying the dynamical Higuchi bound. For the theory not to be in the GR limit (i.e., $\theta = 0$), the graviton mass must be around $m_{\rm FP} \sim 1$.

3.6.1 Cosmological distances

The expansion of the Universe affects the physical distance scales determining the SNIa, BAO, and CMB observations. In a curved space-time the physical distance to an object does not coincide with the coordinate distance. The proper size $d_{\rm phys}$ of an object at a time t is given by,

$$d_{\rm phys} = a(t)\Delta r, \tag{3.69}$$

where Δr is the extent of the object at time t, measured in units of comoving coordinates. Remember, we have set $k = 0$ so there is no spatial curvature. The luminosity distance d_L is defined such that the energy flux F satisfies the usual inverse-squared law,

$$F \propto 1/d_L^2, \tag{3.70}$$

and can be calculated using,

$$d_L(z) = (1+z)\int_0^z \frac{dz'}{H(z')}. \tag{3.71}$$

Rescaling with a factor H_0 we obtain the dimensionless luminosity distance $\mathcal{D}_L(z)$,

$$\mathcal{D}_L(z) = (1+z)\int_0^z \frac{dz'}{E(z')}. \tag{3.72}$$

For small angles, the angular diameter distance d_A is defined by,

$$d_A = d_{\rm phys}/\delta\theta, \tag{3.73}$$

where $\delta\theta$ is the angle subtended on the sky by an object of proper size $d_{\rm phys}$, and it can be calculated from,

$$d_A(z) = d_L(z)/(1+z)^2. \tag{3.74}$$

The comoving angular diameter distance is defined as,

$$D_A(z) \equiv (1+z)d_A(z). \tag{3.75}$$

The volume average distance is defined as,

$$D_V(z) \equiv \left[zD_A^2(z)/H(z)\right]^{1/3}. \tag{3.76}$$

The comoving sound horizon designates how far a density perturbation have had time to travel in the cosmic fluid at a certain point in time. It is given by,

$$r_s(z) = \int_z^\infty dz' \frac{c_s(z')}{H(z')}, \tag{3.77}$$

where c_s is the sound speed,

$$c_s = \frac{1}{\sqrt{3\left(1+\frac{3\rho_b}{4\rho_\gamma}\right)}}, \tag{3.78}$$

with ρ_b being the baryonic energy density and ρ_γ is the energy density of the photons.

3.6.2 Statistical data analysis

As a measure how well a theoretical model fits data, we compute the χ^2-value and the likelihood $\mathcal{L}$ of the model. The former is given by,

$$\chi^2 = \sum_{ij} (\Pi_i^{\rm model} - \Pi_i^{\rm obs}) C_{ij}^{-1} (\Pi_j^{\rm model} - \Pi_j^{\rm obs}), \tag{3.79}$$

where the indices i, j run over the data points, C is the covariance matrix, $\Pi_i^{\rm model}$ is the theoretical model prediction of data point i, and $\Pi_i^{\rm obs}$ is the observed value of the data point i. A smaller value of χ^2 means a better fit of the model to the data and vice versa. As a rule of thumb, the model is a poor fit if $\chi^2/N_{\rm dof} \gg 1$, where the number of degrees of freedom $N_{\rm dof}$ is the

number of data points minus the number of parameters, $N_{\rm dof} = N_{\rm data} - N_{\rm param}$. When combining independent data sets like SNIa and CMB/BAO, the total chi-squared value is obtained by adding the individual χ^2. The likelihood is,

$$\mathcal{L} = e^{-\chi^2/2}. \tag{3.80}$$

A good model is one which gives as good fit to data as possible with as few free parameters as possible. We will use the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) to indicate how well a model is performing against some reference model (e.g. a bimetric model against a GR-based model) [101–104]. They are defined as,

$$\text{AIC} = 2N_{\rm param} + \chi^2_{\rm min}, \tag{3.81a}$$
$$\text{BIC} = N_{\rm param} \ln N_{\rm data} + \chi^2_{\rm min}, \tag{3.81b}$$

with $\chi^2_{\rm min}$ being the minimal χ^2 value. With a lower value of $\chi^2_{\rm min}$ or $N_{\rm param}$ giving a smaller [A/B]IC, we see that a small value of [A/B]IC signifies a good model. The BIC gives a heavier punishment on a larger number of parameters compared with AIC. We define,

$$\Delta[\text{A/B}]\text{IC} = [\text{A/B}]\text{IC}_{\rm model} - [\text{A/B}]\text{IC}_{\rm ref}. \tag{3.82}$$

A value $\Delta\text{AIC} > +5$ is commonly regarded as strong preference for the reference model [104].

3.6.3 Data sets

Type Ia supernovae. A white dwarf in a binary system can accrete mass from the acompanying star until it reaches a critical mass of $\simeq 1.4 M_\odot$.[6] When it reaches this mass, a thermonuclear explosion can take place that can be detected several billion light-years away—a type Ia supernova. The common origin of these dramatic events make them standardizable candles, meaning that we can infer their luminosity distance from eq. (3.70) by measuring their apparent magnitude (m_B) and redshift. Since the luminosity distance depends on the expansion history as in eq. (3.72), these measurements probe $E(z)$. The SNIa are standardized by including an empirical correction for the color (C) and width (X_1) of the light curve so that,

$$m_B = m_B^{\rm obs} + \alpha X_1 - \beta C + \Delta_{\rm M}. \tag{3.83}$$

Here in a cosmological context, α and β are nuisance parameters and Δ_M is a correction depending on the stellar mass of the host galaxy. A mass

[6] Another possible SNIa progenitor is a merging binary white dwarf system.

$\ll 10^{10} M_\odot$ gives $\Delta_M = 0$ while a mass $\gg 10^{10} M_\odot$ gives $\Delta_M = 0.04$ [105]. We use the binned version of the Pantheon data sample containing 1048 SNIa in the redshift range $0.01 < z < 2.3$ placed in 40 redshift bins.[7] The empirical corrections are included in the tabulated values of m_B.

The theoretical model prediction of the apparent magnitude is,

$$m_B = \mathcal{M} + 5 \log_{10} \mathcal{D}_L, \tag{3.84}$$

where $\mathcal{M}$ is defined as,

$$\mathcal{M} \equiv 25 + M_B - 5 \log_{10}(H_0 \, \mathrm{Mpc}), \tag{3.85}$$

with M_B being the absolute magnitude which is defined as the apparent magnitude that the source would have if it was placed at a distance of 10 pc from the observer. Here, we marginalize over $\mathcal{M}$ with the result,

$$\chi^2_{\mathrm{SNIa}} = \hat{\chi}^2 - B^2/D + \ln(D/2\pi). \tag{3.86}$$

Here,

$$B \equiv \sum_{ij} u_i C^{-1}_{ij} \Delta_j, \qquad D \equiv \sum_{ij} u_i C^{-1}_{ij} u_j, \tag{3.87a}$$

$$\hat{\chi}^2 \equiv \sum_{ij} \Delta_i C^{-1}_{ij} \Delta_j, \quad \Delta_i \equiv 5 \log_{10} \mathcal{D}_L(z_i) - m_B(z_i), \tag{3.87b}$$

with the indices i, j running over the 40 redshift bins and,

$$u_i = 1, \quad \text{for all } i = 1, 2, ..., 40. \tag{3.88}$$

See for example Ref. [106] for a detailed derivation.

The cosmic microwave background. During the first 400 000 years after the Big Bang, the photons, electrons, and baryons interacted very frequently, thus behaving as a single fluid, referred to as the baryon-photon fluid. Due to the frequent interaction, the photons could only travel a very short distance before being scattered by an electron. Hence, the Universe was opaque. However, as the temperature dropped due to the expansion, the electrons were captured by the protons to form electrically neutral hydrogen atoms and the Universe became transparent to the photons. This happened around $z \simeq 1090$ and is referred to as the epoch of photon decoupling. From this last scattering surface, the photons have traveled almost freely, carrying the imprint of the

[7] Available at https://github.com/dscolnic/Pantheon/tree/master/Binned_data, last checked 2020-11-05.

baryon-photon fluid at the time of decoupling. Today, we can observe these photons as the cosmic microwave background.

In the CMB temperature map, there is a typical angular scale which exhibits maximal correlation between two points, $\theta_{\rm CMB} \sim 1^\circ$. This defines a standard ruler of physical length,

$$d_{\rm phys}(z_*) = d_A(z_*)\,\theta_{\rm CMB}. \tag{3.89}$$

The multipole number corresponding to this angular scale is,

$$l_A \simeq \pi/\theta_{\rm CMB} = \pi d_A(z_*)/d_{\rm phys}(z_*), \tag{3.90}$$

observationally constrained to [43, 107],

$$l_A = 301.462^{+0.089}_{-0.090}. \tag{3.91}$$

There is much more information in the CMB than the value of l_A. However, since there is no established framework of treating small perturbations on a cosmological background in BR (see Chapter 6), it is difficult to make a theory prediction of the CMB in massive bigravity without introducing questionable assumptions. Accordingly, we will be very conservative and only make use of the CMB standard ruler in combination with BAO in a way that eliminates the dependence on the physics before photon decoupling and which does not require us to specify the physical origin of the standard ruler (see below).

We use the analytical approximation of Ref. [108] and the best-fit values of the baryon and matter densities from CMB and BAO data [43] to set the redshift of photon decoupling, $z_* = 1090$.

Baryon acoustic oscillations. Since the baryons and photons were interacting very frequently in the early universe, the distance between typical density fluctuations in the baryon distribution coincided with the CMB standard ruler. However, due to the excess of photons over baryons ($\sim 10^9 : 1$ at the time of decoupling), the baryons were released from the photon pressure slightly after photon decoupling at $z \simeq 1060$ and we refer to this as the baryon drag epoch. This means that the same standard ruler that can be observed in the CMB is also imprinted in the baryon distribution, only at a slightly later point in time. Today we can observe this distance scale by measuring the typical angular separation between galaxies at some redshift z,

$$\theta_{\rm BAO}(z) = d_{\rm phys}(z_d)/d_A(z). \tag{3.92}$$

Here, we take a conservative approach and combine the CMB and BAO measurements,

$$\Pi_i = \frac{\theta_{\rm CMB}}{\theta_{\rm BAO}(z_i)}\frac{d_{\rm phys}(z_d)}{d_{\rm phys}(z_*)} = \frac{d_A(z_i)}{d_A(z_*)}, \tag{3.93}$$

where i runs between the BAO points. Note that Π_i only depends on the expansion history between present-day and photon decoupling. The dependence on the Hubble constant H_0 disappears in eq. (3.93) so the measured values of Π_i constrain the dimensionless expansion rate $E(z)$. Using eq. (3.79), we get,

$$\chi^2_{\text{CMB/BAO}} = \sum_{ij} (\Pi_i^{\text{model}} - \Pi_i^{\text{obs}}) C_{ij}^{-1} (\Pi_j^{\text{model}} - \Pi_j^{\text{obs}}). \tag{3.94}$$

Here, we use the BAO data sets 6dFGS at $z = 0.106$ [109], SDSS MGS at $z = 0.15$ [110], BOSS DR12 at $z = \{0.38, 0.51, 0.61\}$ [111], BOSS DR14 at $z = 0.72$ [112], and eBOSS QSO at $z = \{0.978, 1.23, 1.526, 1.944\}$ [113]. Note that 6dFGS, SDSS MGS, and BOSS DR14 include observations of the galaxy distribution along the line-of-sight and thus measure $r_{\text{phys}}(z_d)/D_V(z)$ where D_V is the volume average distance (3.76). In this case, we let $\Pi_i = D_V(z_i)/D_A(z_*)$. See Paper IV for details.

In eq. (3.93), we set the ratio of the standard ruler at z_d and z_* equal to the ratio of the ΛCDM values of the sound horizon at these redshifts [43]. Thus,

$$d_{\text{phys}}(z_d)/d_{\text{phys}}(z_*) \simeq r_s(z_d)/r_s(z_*) = 1.019 \pm 0.005. \tag{3.95}$$

We expect this to be a good approximation since the difference in redshift between z_d and z_* is relatively small. In fact, to change the ratio of the sound horizons at the percent level in the interval $z_d < z < z_*$ requires an $\mathcal{O}(1)$ modification of the expansion history, which is a rather extreme deviation from the standard expansion history. The redshift of the drag epoch is calculated using the equations in Ref. [114] with the result, $z_d = 1060$.

Matter density constraints. The ΛCDM cosmological concordance model has $\Omega_\Lambda \simeq 0.7$, $\Omega_{m,0} \simeq 0.3$, and $\Omega_{r,0} \simeq 9 \times 10^{-5}$. In the large $(m_{\text{FP}}, \alpha, \beta)$ limits, if we set $\Omega_\Lambda = 0.7$ the bimetric background expansion mimics a cosmological concordance model during matter domination. However, due to eq. (3.50b), the present-day matter density is $\Omega_{m,0} = 0.3/\cos^2\theta$ rather than 0.3. This phenomenon is referred to as a "dark degeneracy" [115] and makes it impossible in these limits to determine the value of $\Omega_{m,0}$ by only probing the background cosmology during matter domination as we do using SNIa and CMB/BAO data. To break the degeneracy and thus rule out arbitrarily large values of $\Omega_{m,0}$, we invoke a constraint on $\Omega_{m,0}$.

The Chandra satellite measures the X-ray gas fraction in galaxy clusters, $f_{\text{gas}} = \rho_{\text{gas}}/\rho_m$ [116]. It is assumed that the matter content of the largest galaxy clusters resemble that of the Universe as a whole. Hence, the average gas fraction should be the same as the cosmological density fraction,

$$\bar{f}_{\text{gas},0} = \Omega_{\text{gas},0}/\Omega_{m,0}. \tag{3.96}$$

Following Ref. [116], we assume that the energy density of the luminous matter amounts to $0.19h^{0.5}$ of the energy density in the X-ray gas. Here, $h \equiv H_0/(100\,\mathrm{km/s/Mpc})$. Hence, the total amount of baryonic matter can be written $\Omega_b = (1 + 0.19h^{0.5})\Omega_{\mathrm{gas}}$ and,

$$\bar{f}_{\mathrm{gas},0} = (1 + 0.19h^{0.5})^{-1}\Omega_{b,0}/\Omega_{m,0}. \tag{3.97}$$

Thus, we can infer the present-day matter density in the Universe from the observed gas fraction, given a measurement of $\Omega_{b,0}$. However, the analysis is complicated by the fact that the gas fraction is measured at different redshifts, $f_{\mathrm{gas}}(z)$. To obtain the present-day matter density, we need to assume a cosmological model. We are interested in constraining $\Omega_{m,0}$ in the limit of large values of m_{FP}, α, or β. In this case, the background mimics a ΛCDM model in the relevant redshift range.

Here, we use data from Ref. [116] which is based on 26 galaxy clusters in a redshift range $0.07 < z < 0.9$. They assume a prior $\Omega_{b,0}h^2 = 0.0214 \pm 0.0020$ and $h = 0.72 \pm 0.08$ and obtain $\Omega_{m,0} = 0.24 \pm 0.04$. The prior on the baryon density is due to the observed abundance of deuterium compared with hydrogen plus the assumption of standard Big Bang nucleosynthesis [117]. However, since the expansion history in the early universe is modified in models with large values of m_{FP}, α, or β (see Fig. 3.9), we use an independent prior on the baryon density from a sample of fast radio bursts (FRBs).[8] A fast radio burst is a transient pulse in the radio frequency regime. The physical origin of FRBs is not known, but since the photons interact with the free electrons in the extragalactic gas, longer wavelengths from the burst lag behind and the time delay can be measured. The number of free electrons traces the number of baryons, which makes it is possible to obtain a constraint on the baryon density. In Ref. [119], they use five localized FRBs in the redshift range $0.1 \lesssim z \lesssim 0.5$ and obtain $\Omega_{b,0}h = 0.036 \pm 0.011$ [119]. This is larger than the BBN-based prior used in Ref. [116]. Assuming that the present-day matter density scales up with the same factor (cf. eq. (3.97)), we get,

$$\Omega_{m,0} = 0.29 \pm 0.09. \tag{3.98}$$

We refer to this data point as XCL (X-ray cluster) due to its origin from the

[8] Here, we discuss the results of Paper IV. At the time of writing that paper, we had not analyzed Big Bang nucleosynthesis (BBN) in BR. A modified expansion history during BBN can for example be accounted for by introducing an effective number of neutrino species, see e.g. Ref. [118]. This allows us to make a theory prediction for the abundance of the light elements also in the case of a modified cosmology. Together with the measured value of these abundances, we obtain a constraint on $\Omega_{b,0}h^2$ (or, equivalently, η_{10}) also in the case of modified cosmology, see the section below for more details. Note that this is independent of CMB data.

X-ray gas fraction in galaxy clusters. The χ^2 value is calculated as,

$$\chi^2_{\rm XCL} = (\Omega^{\rm model}_{m,0} - \Omega^{\rm obs}_{m,0})^2/\sigma^2, \tag{3.99}$$

where $\Omega^{\rm obs}_{m,0}$ and σ are given by eq. (3.98). Assuming that $\Omega_\Lambda \simeq 0.7$, we can use eqs. (3.98) and (3.53) to estimate an upper limit on the mixing angle in the limit of large $m_{\rm FP}$, α, or β,

$$\theta \lesssim 27^\circ, \quad (\text{large } m_{\rm FP},\ \alpha,\ \text{or } \beta). \tag{3.100}$$

Big Bang nucleosynthesis. During the first few minutes after the Big Bang, the light elements were formed, including deuterium (D) and ^{4}He. The resulting abundances of these elements depend on the expansion in the redshift range $0.2 \times 10^9 \lesssim z \lesssim 3.7 \times 10^9$, corresponding to temperatures $0.06\,{\rm MeV} \lesssim T \lesssim 1\,{\rm MeV}$ (in units where the Boltzmann constant is set to unity $k_B = 1$). This can be used to constrain the physical parameters which determine the expansion history in the early universe. The observed abundance of D can be used to constrain the baryon-to-photon ratio [120],

$$\eta_{10} \equiv 10^{10} n_b/n_\gamma, \tag{3.101}$$

with the result [118, 121],

$$\eta_{10} = 6.1 \pm 0.2. \tag{3.102}$$

Here, n_x denotes the number density of species x. We define the fractional difference in the expansion rate (compared with GR/ΛCDM) as $S \equiv H/H_{\rm GR} = E/E_{\rm GR}$ with $S - 1$ the relative difference in the expansion rate (cf. Fig. 3.9).

The abundance of ^{4}He, Y_p, can be approximated as a linear function in the modified expansion rate S and the baryon-to-photon ratio η_{10} [120, 122],

$$Y_p = 0.2386 + (0.0006)_{\rm num} + [100(S-1) + \eta_{10}]/625. \tag{3.103}$$

The second term on the right-hand side is the numerical error due to the linear approximation. A slower expansion rate (smaller S) results in more time for the neutrons to decay into protons and thus a decreased abundance of ^{4}He (smaller Y_p). In GR, $S = 1$ and the theory prediction for Y_p is,

$$Y_p^{\rm GR} = 0.2484 \pm (0.0006)_{\rm num}, \tag{3.104}$$

which is slightly greater than the observationally inferred value [123],

$$Y_p = 0.2449 \pm 0.0040. \tag{3.105}$$

From eqs. (3.102), (3.103), and (3.105), we can derive a constraint on the modified expansion rate during BBN,

$$S = 0.978 \pm 0.025. \tag{3.106}$$

The allowed region is marked as hatched in Fig. 3.9. For large values of $m_{\rm FP}$, α, or β, the expansion history is modified in the early universe. From eq. (3.62) and assuming that $S \simeq 1$, we can derive,

$$S \simeq 1 - \frac{1}{2}\sin^2\theta, \quad \text{large } (m_{\rm FP}, \alpha, \beta) \tag{3.107}$$

valid in the redshift range $z > z_{\rm eq}$. From eqs. (3.106) and (3.107), we can immediately derive a constraint on the mixing angle θ in the limit of large $m_{\rm FP}$, α, or β,

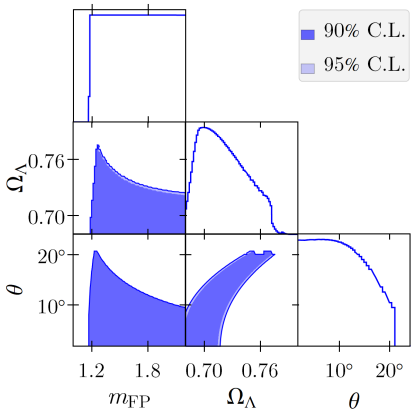
$$\theta \lesssim 18^\circ, \quad \text{large } (m_{\rm FP}, \alpha, \beta). \tag{3.108}$$

3.6.4 Results

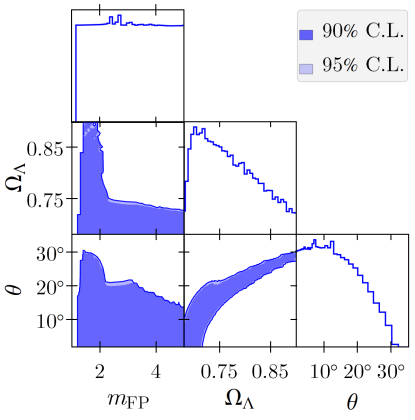
CMB/BAO+SNIa+XCL. Here, we explore the physical parameter space of BR by fitting the cosmological models to data from CMB/BAO, SNIa, and XCL as described above. These data sets probe the dimensionless expansion rate, $E(z) \equiv H(z)/H_0$, between present day and photon decoupling. Evaluating the Friedmann equation (3.14) and the quartic polynomial for y (3.18) at $z = 0$, we get two equations for the two unknowns $\Omega_{\rm tot,0}$ and y_0 that can be solved numerically (recall that the spatial curvature is set to zero). The present-day matter density is then obtained from $\Omega_{\rm tot,0} = \Omega_{m,0} + \Omega_{r,0}$ and eq. (3.51). Hence, the dimensionless expansion rate $E(z)$ is fixed entirely in terms of the physical parameters.

Since the parameter space is only five-dimensional, it is possible to set up a discrete grid and scan each of the grid points. At each point we check if the conditions for a consistent background cosmology are satisfied (i.e., real-valued expansion history with positive matter density, devoid of the Higuchi ghost) and if there exists a Vainshtein screening mechanism that restores GR results on solar-system scales (see Chapter 4 for details). If not, the likelihood is set to zero. This ensures that the model has a consistent cosmology and that it can satisfy observations in the local neighborhood of our universe.

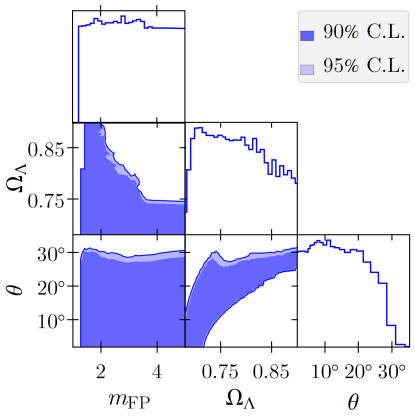
We report the results for three different models: the general bimetric models ($B_0B_1B_2B_3B_4$), the self-accelerating models ($B_1B_2B_3B_4$), and the minimal models ($B_1B_2B_3$). In Fig. 3.12, we plot the two-dimensional (marginalized) confidence contours, showing the results in the parameters θ (mixing angle), $m_{\rm FP}$ (graviton mass), and Ω_Λ (effective cosmological constant). Results for the parameters α and β can be found in Paper IV. In Tab. 3.1, we show the constraints on the physical parameters (90 % confidence). In Fig. 3.13, we plot the best-fit models. The expansion rate follows a ΛCDM concordance model except in the range $0.01 \lesssim z \lesssim 10$ where it is smaller. For the general and self-accelerating models, there is a quick, temporary dip in the equation of



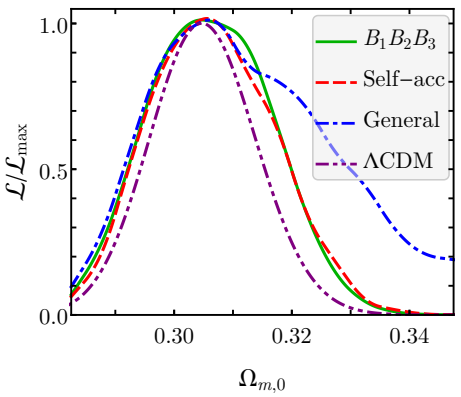
(a) Minimal models ($B_1B_2B_3$).



(b) Self-acc. models ($B_1B_2B_3B_4$).



(c) General models ($B_0B_1B_2B_3B_4$).



(d) Present-day matter density.

Figure 3.12: (a)-(c): Two-dimensional confidence contours from CMB/BAO+SNIa+XCL data. The plots on the diagonals are normalized likelihoods, $\mathcal{L}/\mathcal{L}_{\rm max}$. The graviton mass $m_{\rm FP}$ is expressed in units of $H_0 = 100h\,\mathrm{km/s/Mpc} = 2.1h \times 10^{-33}\,\mathrm{eV}$. Figure reprinted from Paper IV under the terms of Creative Commons Attribution 4.0 license.

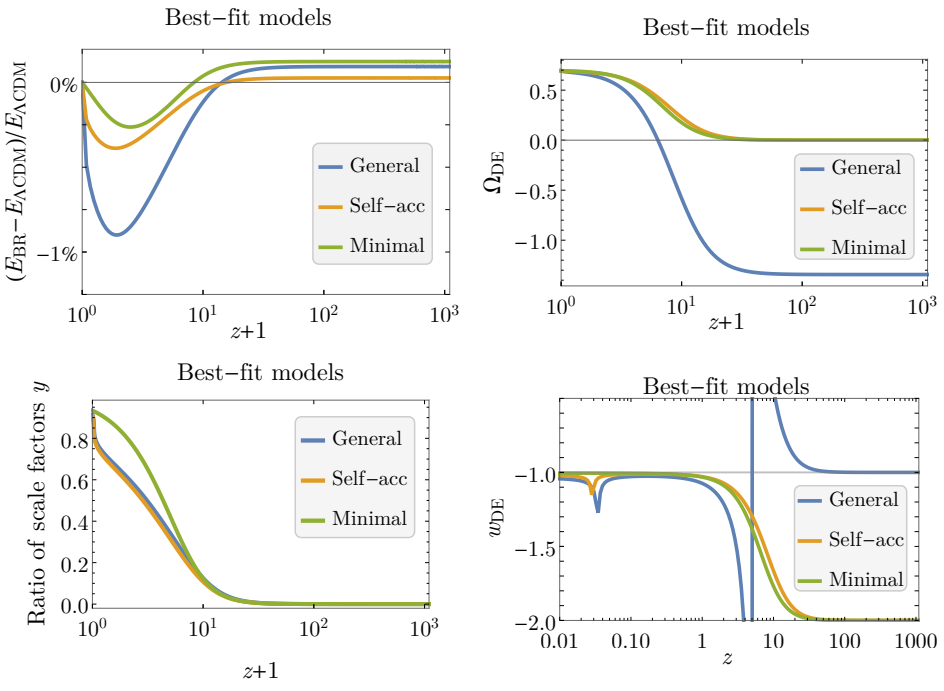


Figure 3.13: Cosmological parameters for the best-fit bimetric models. In the upper left panel, we plot the relative difference in the expansion rate between the best-fit bimetric and the best-fit ΛCDM model. Note the different horizontal axis for the w_{DE} plot.

state around $z \simeq 0.03$. For the general model, the equation of state diverges at $z \simeq 5$. This is due to Ω_{DE} crossing zero there. Hence, this is not a physical divergence.

It is interesting to note that these cosmological data sets do not force θ towards the GR limit, but allow for significant modifications of gravity. For the general and self-accelerating models $\theta < 28°$ and for the minimal models $\theta < 21°$. On the other hand, the GR limit is not excluded by the data. This means that we can set $\theta = 0$ and the model fits data irrespective of the values of m_{FP}, α, and β. Accordingly, the only constraints on these parameters come from the requirements of a consistent cosmology and a working screening mechanism. In Fig. 3.12, we see that there is a degeneracy between θ and Ω_Λ. The reason is that a large value of θ implies that we are far away from the final de Sitter phase today, hence a larger value of Ω_Λ can be allowed.

The BR models fit the cosmological data sets slightly better than ΛCDM, see Tab. 3.1. On the other hand, they exhibit 2-4 additional parameters.

	General model	Self-accelerating	Minimal ($B_1B_2B_3$)	ΛCDM
θ	$11^{\circ}{}^{+17^{\circ}}_{-11^{\circ}}$	$7^{\circ}{}^{+21^{\circ}}_{-7^{\circ}}$	$6^{\circ}{}^{+15^{\circ}}_{-6^{\circ}}$	0°
$m_{\rm FP}$	$2.8^{+\infty}_{-1.6}$	$2.7^{+\infty}_{-1.6}$	$1.6^{+\infty}_{-0.4}$	–
Ω_Λ	$0.71^{+0.25}_{-0.03}$	$0.70^{+0.18}_{-0.02}$	$0.70^{+0.07}_{-0.02}$	$0.70^{+0.01}_{-0.01}$
α	$-5.1^{+\infty}_{-\infty}$	–	–	–
β	$34^{+\infty}_{-32}$	$29^{+\infty}_{-27}$	–	–
$\Omega_{m,0}$	$0.31^{+0.03}_{-0.02}$	$0.31^{+0.02}_{-0.02}$	$0.31^{+0.02}_{-0.02}$	$0.31^{+0.01}_{-0.01}$
$\chi^2_{\rm min}$	52.0	52.0	52.2	52.5
DoF	5	4	3	1

Table 3.1: Constraints on the physical parameters from CMB/BAO+SNIa+XCL data. DoF (degrees of freedom) denotes the number of free parameters.

	ΛCDM	General bimetric	Self-accelerating	Minimal ($B_1B_2B_3$)
AIC	54.5	62.0	60.0	58.2
BIC	56.4	71.7	67.7	64.0
ΔAIC	0	+7.5	+5.5	+3.7
ΔBIC	0	+15.3	+11.3	+7.6

Table 3.2: Information criteria. The reference model is ΛCDM, thus $\Delta[\mathrm{A/B}]\mathrm{IC} = [\mathrm{A/B}]\mathrm{IC}_{\rm model} - [\mathrm{A/B}]\mathrm{IC}_{\Lambda\rm CDM}$.

To assess the performance of these models compared with ΛCDM, we show the AIC and BIC in Tab. 3.2. Interestingly, since the minimal χ^2 value is unchanged when going from the general to the self-accelerating models, the best fit is provided by a self-accelerating model. Since these models exhibit one parameter less than general models, one may argue that the former are preferred. This is clear when comparing their AIC and BIC values, see Tab. 3.2. A $\Delta\mathrm{AIC} > +5$ is commonly regarded as strong preference for the reference (flat ΛCDM) model compared with the (bimetric) model [104]. The BR model performing best with respect to the information criteria is the miminal model. As a final note, even if a bigravity model is performing worse than ΛCDM according to AIC/BIC, the theory may have other beneficial properties which are not quantified by these measures such a self-accelerating cosmologies.

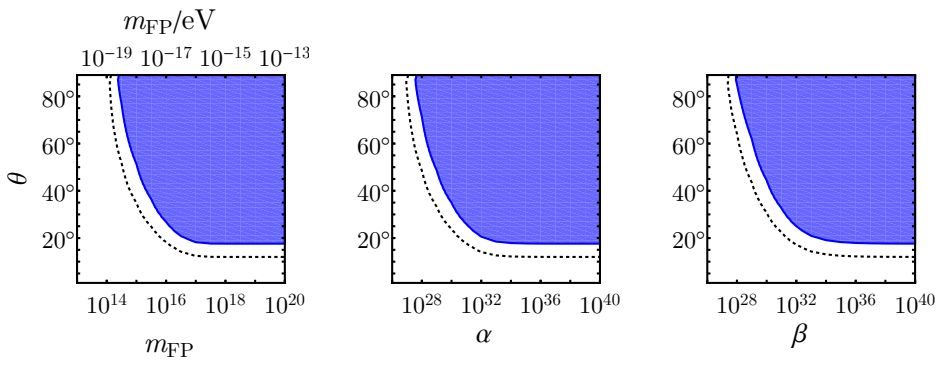


Figure 3.14: Exclusion plot due to BBN constraints. The dashed curves are the best-fit points were $S = 0.978$. Figure reprinted from Paper V under the terms of Creative Commons Attribution 4.0 license.

BBN. As above, the physical parameters plus the observed CMB temperature specify uniquely the cosmological parameters determining the dimensionless expansion rate $E(z)$. In Fig. 3.14, we show the result of imposing the constraint (3.106) due to the measured ^{4}He abundance. The GR limit (vanishing θ) is allowed but the bimetric model gives a slightly better fit with a slower expansion rate at BBN. If $m_{\rm FP} \gtrsim 10^{17}$ (i.e., $m_{\rm FP} \gtrsim 10^{-16}\,$eV), $\alpha \gtrsim 10^{34}$, or $\beta \gtrsim 10^{36}$, the BBN constraint requires $\theta \lesssim 18°$, in accordance with the estimate (3.108). The results are consistent with those in Ref. [124].[9]

3.7 Summary

In bimetric gravity we introduce a massive spin-2 field, which results in an additional energy component, $\Omega_{\rm DE}$, in the Friedmann equation. This species is dynamical, growing, and contributes to the accelerated expansion of the Universe, hence acting as a phantom dark energy, though without a Big Rip. In the late and early universe, $\Omega_{\rm DE}$ approaches constant values and thus the background cosmology is characterized by a late-time ΛCDM phase and an early-time ΛCDM phase with the former exhibiting a greater value of the cosmological constant than the latter. Between these phases, there is a rich spectrum of possible evolution histories. Due to the interaction between the massless and massive spin-2 fields, we can set the cosmological constants of the two metrics to zero and still have an accelerated expansion in the late universe. These solutions are called self-accelerating. Interestingly, these submodels give

[9]This paper appeared after the publication of our results in Paper V.

the best fit to the cosmological data sets that we use. For the theory to contribute with a significant correction to GR, the mixing angle should not be too small. In fact, provided that the graviton mass is of the order of the present-day Hubble scale, the cosmological data sets do not enforce the GR limit.

For a real-valued expansion in the late universe, we require $\Omega_\Lambda > 0$. For a positive matter density in the early universe, we demand $1 + 2\alpha + \beta > 0$. To avoid the Higuchi ghost, which would otherwise introduce an instability in the theory, we set $m_{\rm eff}^2 > 2\widetilde{\Omega}_{\rm DE}$. This can be translated into a condition on the physical parameters, see eqs. (A.3)-(A.5) in Appendix A. Avoiding the Higuchi ghost also guarantees the existence of a healthy finite branch solution at all redshifts. For self-accelerating cosmologies, significant modifications of the expansion history (i.e., $\theta \neq 0$) are only realized when the graviton mass is of the order of the Hubble scale $m_{\rm FP} \sim 10^{-33}\,{\rm eV}$, see Fig. 3.11.

The only limit in the parameter space where the background cosmology mimics a ΛCDM model throughout the whole expansion history is that of vanishing mixing angle, $\theta \to 0$, which is also the GR limit for the full theory. Taking $m_{\rm FP}$, α, or β to large values results in the expansion history mimicking ΛCDM when radiation is negligible but expanding slower during radiation domination.

Bimetric gravity is observationally viable with respect to the cosmological data sets that we have imposed. Interestingly, these data sets do not enforce the GR limit (vanishing θ) but allows for a significant mixing angle between the massless and massive spin-2 fields. This means that the physical metric contains a mix of the massless and massive spin-2 fields. The question is if this is a result of our fitting only the ratio of the CMB and BAO distance scales. To decide upon this issue, we must establish a framework for treating perturbations on cosmological backgrounds. Due to the break down of linear perturbation theory, this is a challenge, see Chapter 6.

Chapter 4

Gravity on local scales

In Chapter 3, we have seen how the presence of the massive spin-2 field affects the evolution of the Universe on the largest scales ($\gtrsim 100\,\mathrm{Mpc}$). In this chapter, we analyze the effects of the theory on galaxy scales ($\sim 1\,\mathrm{kpc}$-$100\,\mathrm{kpc}$, i.e., $10^{19}\,\mathrm{m}$-$10^{21}\,\mathrm{m}$), solar-system scales ($\sim 1\,\mathrm{AU} \sim 10^{11}\,\mathrm{m}$), and below. To model the gravitational force due to the Sun or due to a galaxy, we use static, spherically symmetric (SSS) solutions of the equations of motion and we refer to these as local solutions. In GR, Jebsen–Birkhoff's theorem states that the Schwarzschild space-time is the unique spherically symmetric vacuum solution [125–127]. In BR, the theorem does not hold and there is a spectrum of SSS solution [2]. There are two primary ways of probing gravity on local scales. First, the motion of test bodies in the gravitational field from an object, for example the planetary orbits around the Sun. Second, the bending of light around a massive source, that is, gravitational lensing.

4.1 Ansatz and solutions

A general ansatz for a static, spherically symmetric space-time reads,

$$ds_g^2 = -A(r)dt^2 + B(r)dr^2 + r^2 d\Omega^2, \tag{4.1a}$$
$$ds_f^2 = -\widetilde{A}(r)dt^2 + \widetilde{B}(r)dr^2 + 2\widetilde{C}(r)dtdr + \widetilde{R}^2(r)d\Omega^2. \tag{4.1b}$$

The Einstein tensor of the physical metric is diagonal in these coordinates so that $V^t{}_r = 0$ in vacuum (cf. eq. (2.5a)). There are three possible solutions,

$$\widetilde{C} = 0, \tag{4.2a}$$
$$\widetilde{C} = \pm(A\widetilde{B} - \widetilde{A}B)/\left(2\sqrt{AB}\right), \tag{4.2b}$$
$$0 = \beta_1 + 2\beta_2\widetilde{R} + \beta_3\widetilde{R}^2. \tag{4.2c}$$

Table 4.1: Static, spherically symmetric, non-bidiagonal solutions. M = Minkowski, S = Schwarzschild, AdS = anti-de Sitter, dS = de Sitter, SAdS = Schwarzschild–anti-de Sitter, and SdS = Schwarzschild–de Sitter. All combinations are local solutions to the equations of motion but only those with check marks exhibit metrics that can be defined globally on the same manifold.

$g_{\mu\nu}$ \ $f_{\mu\nu}$		M	S	AdS	dS	SAdS	SdS
$\mathbb{R}^4$	M	✓	×	✓	×	×	×
$\mathbb{R}^2 \times \mathbb{S}^2$	S	×	✓	×	×	✓	×
$\mathbb{S}^1 \times \mathbb{R}^3$ or $\mathbb{R}^4$	AdS	✓	×	✓	×	×	×
$\mathbb{R}^1 \times \mathbb{S}^3$	dS	×	×	×	✓	×	×
$\mathbb{S}^1 \times \mathbb{R}^1 \times \mathbb{S}^2$ or $\mathbb{R}^2 \times \mathbb{S}^2$	SAdS	×	✓	×	×	✓	×
$\mathbb{R}^1 \times \mathbb{S}^1 \times \mathbb{S}^2$	SdS	×	×	×	×	×	✓
Topology	Solution						

The last equation (4.2c) is analyzed in Ref. [2], where we also allow for a time-dependence. In the second solution (4.2b), the metrics are not diagonal in the same coordinates (i.e., non-bidiagonal). In this case, the equations of motion reduce to two decoupled Einstein vacuum equations, with or without cosmological constant depending on the β-parameters, see Ref. [128] and Paper I. Hence, the metrics can be solved analytically. In Tab. 4.1, we present the possible solutions also taking into account whether the two metrics can be compatible with a single topology. Some metrics are compatible with multiple topologies, for example anti-de Sitter which admits both $\mathbb{S}^1 \times \mathbb{R}^3$ and $\mathbb{R}^4$ and Schwarzschild–anti-de Sitter which admits both $\mathbb{S}^1 \times \mathbb{R}^1 \times \mathbb{S}^2$ and $\mathbb{R}^2 \times \mathbb{S}^2$.

Here, we are interested in the first solution (4.2a) to the SSS equations of motion where the off-diagonal term in the f-metric vanishes, so the metrics are bidiagonal. In this case, there are modifications of gravity beyond GR. Generically, the equations must be solved numerically but approximate analytical solutions can be found in two different regimes: the Yukawa (linearized) solution and the Vainshtein (screening) solution. In the bidiagonal case, it is convenient to rewrite the metric ansatz as,

$$ds_g^2 = -e^{2\Phi}dt^2 + e^{-2\Psi}dr^2 + r^2 d\Omega^2, \tag{4.3a}$$

$$ds_f^2 = c^2\left[-e^{2\widetilde{\Phi}}dt^2 + e^{-2\widetilde{\Psi}}\left(\partial_r \widetilde{R}\right)^2 dr^2 + \widetilde{R}^2 d\Omega^2\right], \tag{4.3b}$$

where the radial dependence of the fields $(\Phi, \Psi, \widetilde{\Phi}, \widetilde{\Psi}, \widetilde{R})$ is implicit and c is the conformal constant between the metrics in the asymptotic region $f_{\mu\nu}|_{r\to\infty} = c^2 g_{\mu\nu}|_{r\to\infty}$. Note that $c\widetilde{R}$ is the area radius of the f-metric, which can be

rewritten as,

$$ds_f^2 = c^2\left[-e^{2\widetilde{\Phi}}dt^2 + e^{-2\widetilde{\Psi}}d\widetilde{R}^2 + \widetilde{R}^2 d\Omega^2\right]. \tag{4.4}$$

The coordinate transformation from the area radius of $g_{\mu\nu}$ to the area radius of $f_{\mu\nu}$ is $r \to \widetilde{R}$ and the relative difference between these two coordinates can be parameterized by μ, defined as,

$$\mu \equiv (\widetilde{R} - r)/r. \tag{4.5}$$

Since μ parameterizes the transformation "restoring" $\widetilde{R}$ as the radial coordinate, we refer to it as a Stückelberg field.

Linear regime – Yukawa

To understand when the Yukawa (linearized) solution and the Vainshtein (screening) solution are valid respectively, let us consider a static, spherically symmetric linear perturbation around a flat background $\bar{f}_{\mu\nu} = c^2\bar{g}_{\mu\nu} = c^2\eta_{\mu\nu}$. This amounts to assuming that the metric fields and their radial derivatives are small, that is, $(\Phi, \Psi, \widetilde{\Phi}, \widetilde{\Psi}, \mu) \ll 1$ and $r\partial_r(\Phi, \Psi, \widetilde{\Phi}, \widetilde{\Psi}, \mu) \ll 1$. Outside a source of mass M, the solution is [128],

$$\Phi(r) = -M\cos^2\theta\left(\frac{1}{r} + \frac{4\tan^2\theta}{3}\frac{e^{-r/\lambda_g}}{r}\right), \tag{4.6a}$$

$$\Psi(r) = -M\cos^2\theta\left(\frac{1}{r} + \frac{2\tan^2\theta(1 + r/\lambda_g)}{3}\frac{e^{-r/\lambda_g}}{r}\right), \tag{4.6b}$$

$$\widetilde{\Phi}(r) = -M\cos^2\theta\left(\frac{1}{r} - \frac{4\tan^2\theta}{3}\frac{e^{-r/\lambda_g}}{r}\right), \tag{4.6c}$$

$$\widetilde{\Psi}(r) = -M\cos^2\theta\left(\frac{1}{r} - \frac{2\tan^2\theta(1 + r/\lambda_g)}{3}\frac{e^{-r/\lambda_g}}{r}\right), \tag{4.6d}$$

$$\mu(r) = -M\cos^2\theta\frac{2}{3}\left[1 + (r/\lambda_g)^{-1} + (r/\lambda_g)^{-2}\right]\frac{e^{-r/\lambda_g}}{r}, \tag{4.6e}$$

where $\lambda_g = 1/(H_0 m_{\rm FP})$ is the Compton wavelength of the massive graviton. $\Phi(r)$ is the gravitational potential and $\Psi(r)$ is the spatial curvature. In GR, $\Phi_{\rm GR}(r) = \Psi_{\rm GR}(r) = -M/r$. In bimetric gravity, the potentials acquire an additional exponential Yukawa term. The motion of a massive (non-relativistic) test particle is determined by the gravitational potential Φ while the motion of a massless (relativistic) particle is dictated by the Weyl potential $\varphi \equiv (\Phi + \Psi)/2$.

Here, we assumed flat asymptotics $f_{\mu\nu}|_{r\to\infty} = c^2 g_{\mu\nu}|_{r\to\infty} = c^2\eta_{\mu\nu}$ but the asymptotic structure of the Universe is closer to de Sitter. This can

be accounted for by adding a cosmological constant term $-\Lambda r^2/3$ to the potentials $(\Phi, \Psi, \widetilde{\Phi}, \widetilde{\Psi})$ [129]. Since the cosmological constant is of the order $\Lambda \sim H_0^2 = 1/r_H^2$ and all local systems have a size much less than the Hubble radius, it follows that $\Lambda r^2 \ll 1$. Thus, the effect of a cosmological constant can be neglected at local scales.

The linear solution (4.6) is valid (i.e., the metric fields are small) far away from the Schwarzschild radius $r \gg r_S = 2M$. However, from eq. (4.6e) it follows that the Stückelberg field μ is of the order unity if $r \lesssim r_V$ where r_V is the Vainshtein radius, defined as,

$$r_V \equiv \left(r_S \lambda_g^2\right)^{1/3}, \tag{4.7}$$

or, equivalently,

$$r_V \simeq 0.095 \left[\frac{M}{M_\odot} \frac{1}{h^2 m_{\rm FP}^2}\right]^{1/3} \text{kpc}. \tag{4.8}$$

For the linearized solution to apply, we require $r \gg (r_V, r_S)$. When the linear approximation breaks down at $r \simeq r_V$, non-linear terms must be added. From eq. (4.7), we see that in the limit of vanishing Compton wavelength $\lambda_g \to 0$ (i.e., large graviton mass), the linear solution is applicable down to length scales of the order r_S, as in GR.

If we naively take the zero mass limit at the linear level (i.e., $\lambda_g \to \infty$), the gravitational slip, γ, evaluates to,

$$\gamma|_{m_{\rm FP}\to 0} = \left[5 + \cos(2\theta)\right]/6, \tag{4.9}$$

where $\gamma \equiv \Phi/\varphi$ and $\varphi \equiv (\Phi + \Psi)/2$ is the Weyl potential. Hence, $\gamma < 1$ in the zero mass limit which results in a modified relation between the dynamics of massless and massive particles compared with GR [75]. This is the (in)famous van Dam–Veltman–Zakharov discontinuity [57, 58]. However, when $m_{\rm FP} \to 0$, the Vainshtein radius increases indefinitely and the linear solution breaks down at all length scales. That is, non-linear terms must be included [59]. In fact, these terms can restore GR results, which we will show now.

Non-linear regime – Vainshtein

Inside the Vainshtein radius $r \lesssim r_V$, non-linear terms must be added to the metrics. Since in this case the Stückelberg field μ becomes of order unity while the potentials $(\Phi, \Psi, \widetilde{\Phi}, \widetilde{\Psi})$ are still small, we expand to linear order in the potentials but include all orders in μ. The result outside a source of mass M is [72, 130],

$$r\partial_r\Phi(r) = \frac{M}{r}\left[1 - \sin^2\theta\left(\frac{r}{r_V}\right)^3\left(\mu(r) - \frac{\beta}{3}\mu^3(r)\right)\right], \tag{4.10a}$$

$$\Psi(r) = -\frac{M}{r}\left[1 + \sin^2\theta\left(\frac{r}{r_V}\right)^3\left(\mu(r) - \alpha\mu^2(r) + \frac{\beta}{3}\mu^3(r)\right)\right]. \tag{4.10b}$$

The equation for the Stückelberg field is a seventh-order polynomial, as shown in eq. (D.2) in Appendix D. The physical root is the one satisfying $\mu = -1/\sqrt{\beta}$ at $r = 0$ and $\mu \to 0$ as $r \to \infty$. The latter is implied by requiring a flat asymptotic structure. The equations (4.10) apply outside of a source. On the inside, we must include the radial dependence of the mass and the pressure, see Ref. [72] for details.

In the equation for μ (D.2), the radius only appears in the combination r/r_V as is the case for the correction to the gravitational potentials in eq. (4.10). This means that the radius within which GR is restored is proportional to r_V. In fact, far inside the Vainshtein radius $r \ll r_V$, the bimetric corrections to the GR potentials are proportional to $(r/r_V)^3$ and hence vanish. Thus we have a working screening mechanism with r_V being the minimal radius within which GR is restored. However, the equation determining μ also contains the parameters (θ, α, β) so, possibly, GR can be restored also outside of r_V depending on the value of these parameters. As we discuss in Section 4.3, GR is restored within an increasingly large radius in the limits $\alpha \to \infty$ and $\beta \to \infty$.

The Vainshtein radius can be written,

$$r_V/r_* = (\rho_*/\rho_c m_{\rm FP}^2)^{1/3}, \tag{4.11}$$

where $\rho_* \equiv M_*/(4\pi r_*^3/3)$ is the mean density of the source and ρ_c is the critical density of the Universe today $\rho_c \equiv 3H_0^2/\kappa_g$. So, an object which has a density much greater than $\rho_c m_{\rm FP}^2$ lies far inside its own Vainshtein radius and is thus screened.

4.2 Analytical constraints

To guarantee the existence of the physical root of μ and hence a working Vainshtein screening, we must constrain the parameter space [72]. The full equations are presented in Appendix A. In Fig. 4.1, we plot these constraints in the $\alpha\beta$-plane and indicate the dependence on θ. A significant portion of the parameter space is excluded. From eq. (A.6a), we have $\beta > 1$ which implies $B_3 > 0$. Together with $B_1 > 0$ from cosmology (see eq. (3.36)), one can use eq. (A.6) to show that $B_2 < 0$. Taken together, this means that the

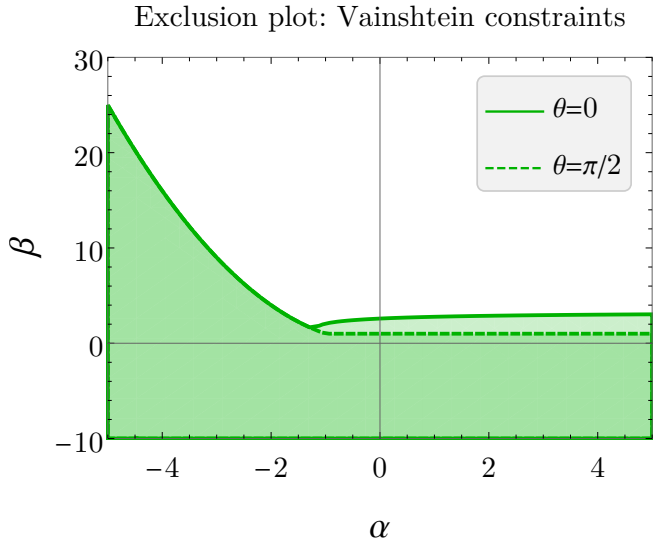


Figure 4.1: Exclusion plot from requiring a working Vainshtein screening mechanism. $\theta = 0$ gives the most restrictive constraint while $\theta = \pi/2$ the least restrictive. Figure reprinted from Paper III under the terms of Creative Commons Attribution 4.0 license.

requirement of a consistent cosmology and a working screening mechanism implies,

$$B_1 > 0, \quad B_2 < 0, \quad B_3 > 0. \tag{4.12}$$

So, the most minimal bimetric model satisfying these demands is a $B_1B_2B_3$-model. The constraint (4.12) excludes many of the popular models in the literature, including all two-parameter models. Self-accelerating cosmologies are still viable since the cosmological constant parameters of each metric sector, B_0 and B_4, are allowed to vanish. In Fig. 4.2, we combine the cosmological and the Vainshtein constraints for two different values of $m_{\rm FP}^2/(2\Omega_\Lambda)$. The Higuchi and Vainshtein constraints are complementary. A significant part of the parameter space is excluded. Self-accelerating solutions are concentrated to a line in the $\alpha\beta$-plane and indicated in the figure and its horizontal position depends on θ and $m_{\rm FP}^2/(2\Omega_\Lambda)$. For large enough θ or $m_{\rm FP}^2/(2\Omega_\Lambda)$, this line lies completely in the excluded region and, in this case, there are no self-accelerating solutions.

For a working Vainsthein mechanism, it is necessary that $\beta \gtrsim 3$, see Paper III. In case of a $B_1B_2B_3$, we can use eq. (3.66b) to derive,

$$\tan^2\theta \lesssim \frac{1}{3}\left(-2 + 3m_{\rm FP}^2/\Omega_\Lambda - \sqrt{1 + 6m_{\rm FP}^2/\Omega_\Lambda(-2 + 3m_{\rm FP}^2/2\Omega_\Lambda)}\right). \tag{4.13}$$

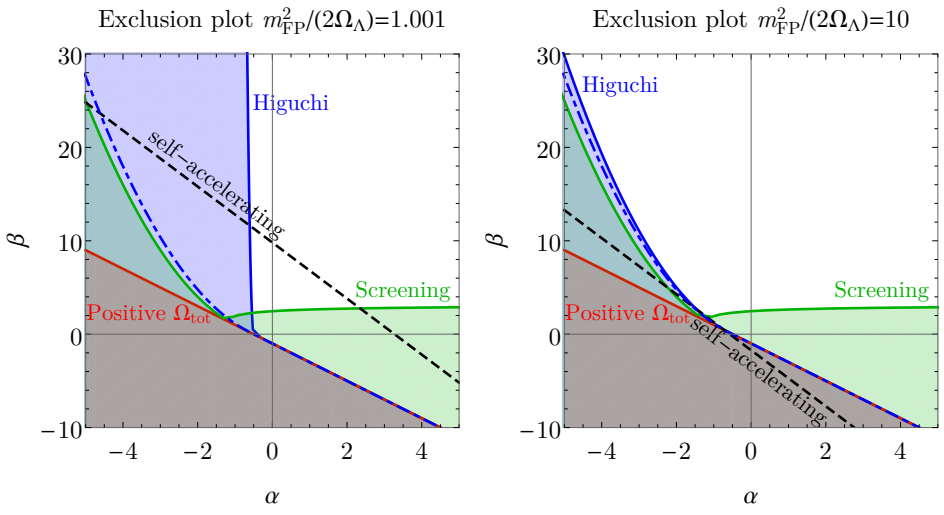


Figure 4.2: Exclusion plot. Combined constraints from demanding an existing Vainshtein screening mechanism (green), a positive matter density in the early universe (red), and absence of the Higuchi ghost (blue). Here, $\theta = 20°$. The dash-dotted curve shows how the boundary of the blue region changes in the limit $m_{\rm FP}^2/2\Omega_\Lambda \to \infty$. The green region is unaffected by the value of $m_{\rm FP}$ and Ω_Λ. The self-accelerating solutions lie along the dashed lines and always have slope -3 (cf. eq. (3.67)). These lines move upwards if θ or $m_{\rm FP}^2/2\Omega_\Lambda$ decreases.

The largest value of θ is achieved at the Higuchi bound when $m_{\rm FP}^2 = 2\Omega_\Lambda$. At this point $\theta = 20°$. From these analytical calculations we can already conclude that,

$$\theta \lesssim 20°, \quad (B_1B_2B_3 \text{ models}), \tag{4.14}$$

for these models.

4.3 General relativity limits

The deviation from GR at local scales can be formulated in terms of two distance-dependent gravitational "constants": $\kappa_g^{\rm massive}(r)$ and $\kappa_g^{\rm massless}(r)$. Here,

$$\kappa_g^{\rm massive}(r) \equiv \kappa_g \Phi(r)/\Phi_{\rm GR}(r), \tag{4.15}$$

parameterizes the modified gravitational force law for massive particles and,

$$\kappa_g^{\rm massless}(r) \equiv \kappa_g \varphi(r)/\varphi_{\rm GR}(r), \tag{4.16}$$

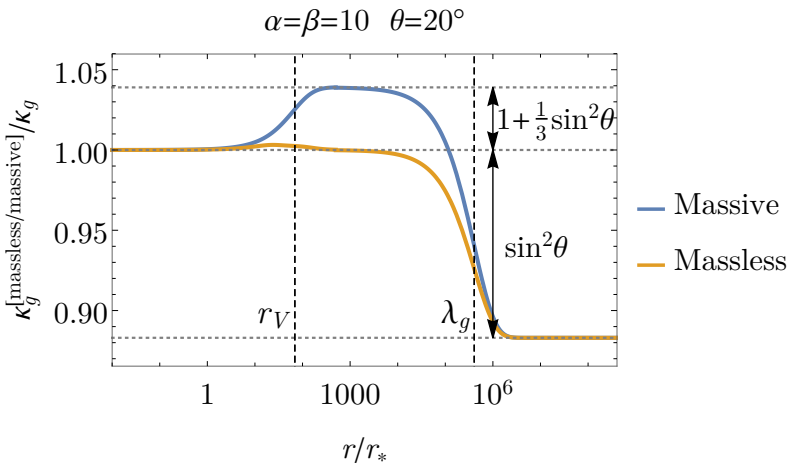


Figure 4.3: Gravitational "constants" for massive (blue) and massless (orange) particles. Here, r_* is the radius of the source and we have set $\lambda_g = r_H$ (i.e., $m_{\rm FP} = 1$). When $\kappa_g^{\rm massive} = \kappa_g^{\rm massless} = \text{const.}$, GR results are recovered. This is the case far inside the Vainshtein radius $r \ll r_V$ and far outside the Compton wavelength $r \gg \lambda_g$.

parameterizes the modified gravity effect on the massless particles (e.g., light). GR results are recovered whenever $\kappa_g^{\rm massive} = \kappa_g^{\rm massless} = \kappa_g$, or more generally when $\kappa_g^{\rm massive} = \kappa_g^{\rm massless} = \text{const.}$, by rescaling of κ_g. See Fig. 4.3 for an example of how the gravitational constants vary with radius.

Linear regime – Yukawa

In the linear regime ($r \gg r_V$), there are two ways of reproducing GR results. First, if the mixing angle θ is small enough, the amplitudes of the Yukawa terms are suppressed in Φ and Ψ (cf. eqs. (4.6a)-(4.6b)). This results in $\kappa_g^{\rm massive} = \kappa_g^{\rm massless} = \kappa_g$. Second, if the Compton wavelength is much smaller than the length scales probed in the system, $\lambda_g \ll r$, the Yukawa term is exponentially suppressed. In this case, $\kappa_g^{\rm massive} = \kappa_g^{\rm massless} = \cos^2\theta\kappa_g$, resulting in a constant rescaling of the gravitational constants.

For self-accelerating solutions, a large graviton mass (which is required for the linearized solution to be valid) implies a small mixing angle via eq. (3.68), see Fig. 3.11. For example, if $m_{\rm FP} = 10^{33}$ (i.e., $m_{\rm FP} = 1\,\text{eV}$), the exponential Yukawa term is suppressed enough to reproduce GR within experimental precision. However, for a self-accelerating solution this implies $\theta \sim 10^{-34}$ so for such a model there are no detectable deviations from GR. In short, to

have self-accelerating solutions in the Yukawa regime requires a mixing angle extremely close to the GR limit.

Non-linear regime – Vainshtein

In Section 4.1, we showed that r_V is the minimal radius within which GR is restored. See Fig. 4.3 for an example. Thus, by taking the graviton mass to zero (i.e., $\lambda_g \to \infty$), the Vainshtein radius diverges and GR is restored at all length scales. That is, $\kappa_g^{\text{massless}} = \kappa_g^{\text{massive}} = \kappa_g$ as $m_{\text{FP}} \to 0$. A caveat is that we assumed a flat asymptotic structure rather than de Sitter and due to the Higuchi bound $m_{\text{FP}}^2 > 2\Omega_\Lambda$, the zero mass limit cannot be taken without exciting the Higuchi ghost. So, $m_{\text{FP}} \sim 1$ (i.e., $m_{\text{FP}} \sim 10^{-33}\,\text{eV}$) is the lower limit for the graviton mass, corresponding to a Compton wavelength which is of the order of the Hubble radius, $\lambda_g \sim r_H$.

In Fig. 4.3, we plot the modified gravitational constants, showing that far inside the Vainshtein radius $r \ll r_V$, we have $\kappa_g^{\text{massive}} = \kappa_g^{\text{massless}} = \kappa_g$ and GR is recovered. Approaching r_V, there is an increase in $\kappa_g^{\text{massive}}$, eventually reaching a (temporary) plateau. This means that the gravitational force is effectively increased with some constant factor as we go further away from the source. $\kappa_g^{\text{massless}}$ has a bump around r_V and then reaches a (temporary) plateau value. Approaching the Compton wavelength λ_g, the gravitational constant begins to decrease to reach a final constant value as $r \gg \lambda_g$ and GR is once again recovered.

Expanding eqs. (4.10) and (D.2) around $1/\alpha = 0$ and $1/\beta = 0$, one can show that the radius within which $\kappa_g^{\text{massless}} = \kappa_g^{\text{massive}} = \kappa_g$ grows indefinitely with increasing α or β, see Fig. 4.4 for an example. That is, $\alpha \to \infty$ and $\beta \to \infty$ are GR limits with respect to the local physics (i.e., SSS solutions). This is for general bimetric models. For self-accelerating solutions however, these limits violate the conditions for a working screening mechanism which is easily seen by following the self-accelerating line in Fig. 4.2 in any of these limits.

4.4 Observational constraints

4.4.1 Solar-system tests

In the solar system, the gravitational force law has been probed to great accuracy and does not deviate more than 10^{-9} from an inverse-square law at length scales of 1 AU [33]. At smaller distances, there are laboratory tests down to at least $\sim 10\,\mu$m. Following Ref. [74], we constrain the deviation from GR to be within 10^{-9} in this range of length scales. This is a very conservative

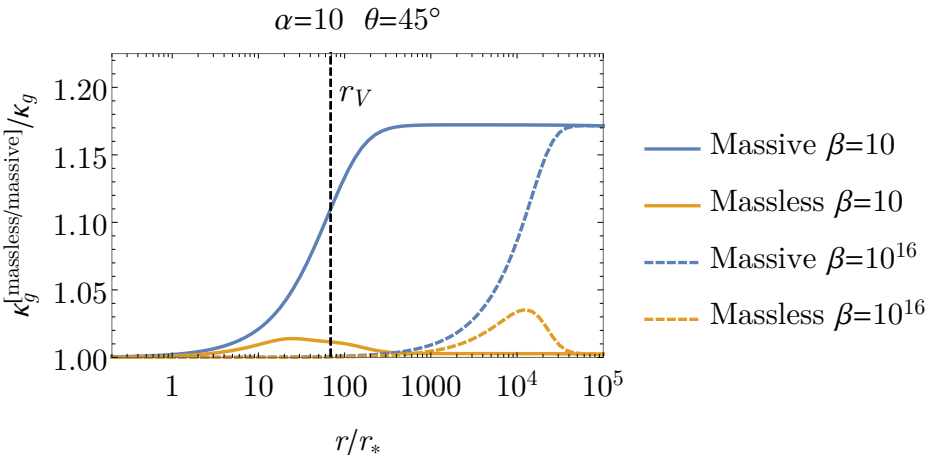


Figure 4.4: Gravitational "constants" for massive (blue) and massless (orange) particles. Increasing β from 10 to 10^{16} recovers GR results (i.e., $\kappa_g^{\text{massive}} = \kappa_g^{\text{massless}} = \kappa_g$) within an increasingly large radius; r_* is the radius of the source. Increasing α has a similar effect.

constraint that can be refined by analyzing the probes one by one. A careful analysis of the individual solar-system tests is presented in Refs. [71, 131].

If we are far inside the Vainshtein radius, GR is restored and solar-system tests are satisfied. This is the case if the graviton mass is small enough. However, due to the Higuchi bound, m_{FP} cannot be set to zero, but $m_{\text{FP}} \gtrsim 1$ is allowed. Setting the graviton mass to its minimal value $m_{\text{FP}} \sim 1$ implies that all systems which have a mean density much greater than the critical density of the Universe, lies far inside its own Vainshtein radius (see eq. (4.11)). This includes the solar system and all galaxies. Thus, if we set the graviton mass to its minimal value $m_{\text{FP}} \sim 1$, GR is restored in all local systems.

Assuming that $\mathcal{O}(\alpha) = \mathcal{O}(\beta) = 1$, one can expand around $r/r_V = 0$ with the result,

$$\Phi(r) \simeq -\frac{M}{r}\left[1 + \sin^2\theta\left(\frac{r}{r_V}\right)^3\right], \quad r \ll r_V, \tag{4.17a}$$

$$\Psi(r) \simeq -\frac{M}{r}\left[1 - \sin^2\theta\left(\frac{r}{r_V}\right)^3\right], \quad r \ll r_V. \tag{4.17b}$$

Hence, the requirement that solar-system tests are satisfied is,

$$\sin^2\theta\left(\frac{r}{r_V}\right)^3 < 10^{-9}, \quad r \ll r_V. \tag{4.18}$$

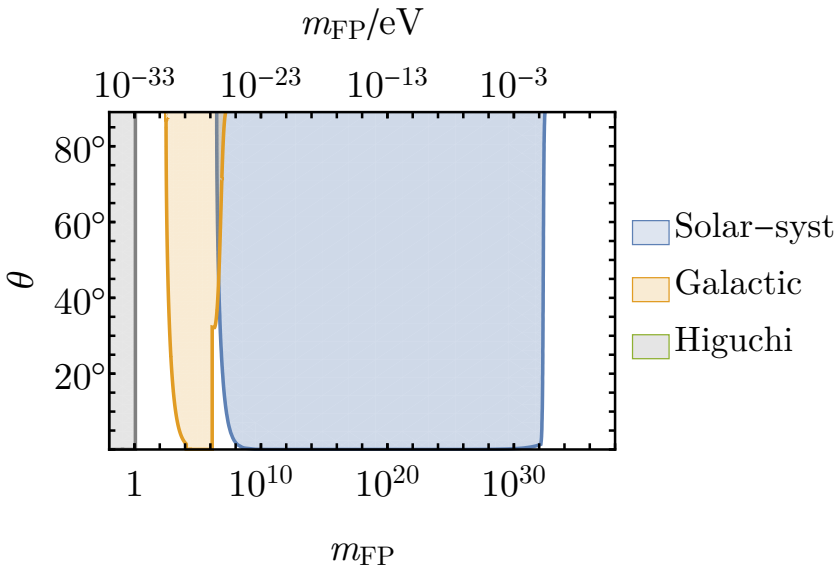


Figure 4.5: Exclusion plot from solar-system tests (blue), gravitational lensing from galaxies (orange), and the Higuchi bound (gray). In the low-mass regime, the Vainshtein mechanism can effectively restore GR results while in the large-mass regime, the Yukawa term is exponentially suppressed and GR is restored. A large portion of the parameter space in the range $10^4 \lesssim m_{\rm FP} \lesssim 10^6$ and $10^8 \lesssim m_{\rm FP} \lesssim 10^{32}$ is excluded. Here, we have assumed $\mathcal{O}(\alpha) = \mathcal{O}(\beta) = 1$. The solar-system and galactic constraints can be eliminated by choosing α or β to take large values.

If we are far outside the Vainshtein radius, that is in the linear regime, we demand,

$$\frac{4\tan^2\theta}{3} e^{-r/\lambda_g} < 10^{-9}, \quad r \gg r_V, \tag{4.19}$$

see eq. (4.6a). Here, r is the length scale probed by the observational tests, which range from $10\,\mu$m to 1 AU. The resulting exclusion plot is presented in Fig. 4.5. For self-accelerating models, the requirement of a consistent cosmology in eq. (3.68) is more strict than eqs. (4.18) and (4.19), so a self-accelerating model with a consistent cosmology automatically satisfies solar-system tests.

The above constraints were derived assuming $\mathcal{O}(\alpha) = \mathcal{O}(\beta) = 1$. This is implicitly assumed in most of the literature, for example Refs. [72–74]. In Fig. 4.6 we plot the confidence contours in the $\alpha\beta$-plane in a case which violates solar-system tests for parameters α and β of order unity. As we have seen, $\alpha \to \infty$ and $\beta \to \infty$ are GR limits of the local solutions, so taking these

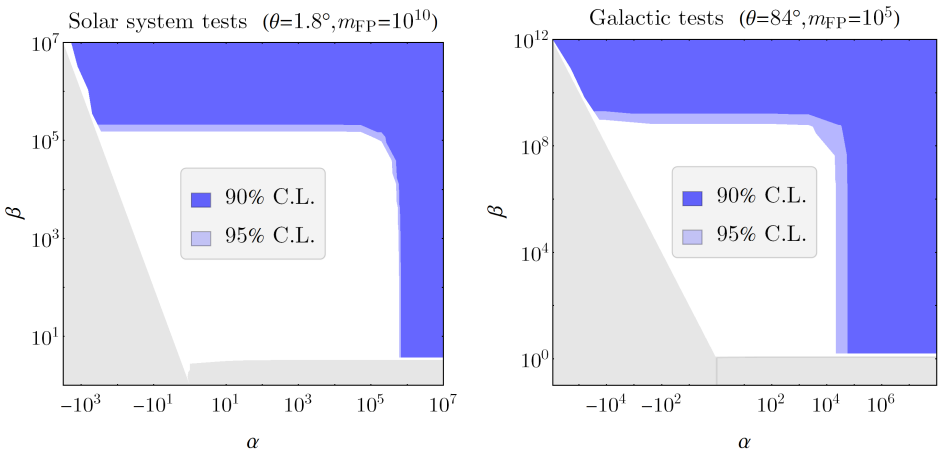


Figure 4.6: Confidence contours in the $\alpha\beta$-plane. The gray region is excluded by the requirement of a working Vainshtein mechanism, cf. Fig. 4.1. *Left panel*: Solar system tests with $(\theta = 1.8°, m_{\rm FP} = 10^{10})$. *Right panel*: Gravitational lensing by a $10^{11}\,M_\odot$ galaxy. Here, $(\theta = 84°, m_{\rm FP} = 10^{5})$. In both panels, if $\mathcal{O}(\alpha) = \mathcal{O}(\beta) = 1$, the observational tests are violated. As can be seen, when $\alpha \to \infty$ or $\beta \to \infty$, GR is restored and the theory is viable. Figure reprinted from Paper IV under the terms of Creative Commons Attribution 4.0 license.

parameters to be large alleviates these constraints.

4.4.2 Gravitational lensing by galaxies

As the light from a distant source passes through the vicinity of a galaxy, the light is deflected due to gravitational lensing. Here, we assume a simple model for the gravitational lensing at a distance $D_{\rm source} \sim 1\,{\rm Gpc}$ by a galaxy of mass $M \sim 10^{11} M_\odot$ at half this distance. We set the lensing radius to $r_{\rm lens} = 8\,{\rm kpc}$ and assume, for simplicity, that the lensing takes place at the edge or outside the galaxy. The value of the gravitational slip γ,

$$\gamma \equiv \Phi/\varphi, \tag{4.20}$$

is obtained by measuring the degree of lensing and the dynamical mass of the galaxy. The GR value is $\gamma_{\rm GR} = 1$. From observations of the gravitational lensing, mass-dynamical structure, and velocity dispersion of 53 massive, early-type galaxies in a redshift range $0.1 \lesssim z \lesssim 0.3$, it was concluded in Ref. [132]

that $\gamma = 1 \pm 0.05$. Hence, we calculate the χ^2-value as,

$$\chi^2_{\text{gal.lens.}} = \left(\frac{\gamma(r_{\text{lens}}) - 1}{0.05} \right)^2 . \tag{4.21}$$

The result is plotted in Fig. 4.5 in the case $\mathcal{O}(\alpha) = \mathcal{O}(\beta) = 1$. As we have seen, by choosing the parameters α or β to be large, GR is restored via the Vainshtein mechanism and the constraints in Fig. 4.5 are alleviated. See Fig. 4.6 for an example.

4.5 Summary

In bimetric gravity, the massive spin-2 field can modify the physics on solar-system scales and galactic scales (collectively referred to as local scales). Here, we analyze the gravitational force by looking at static, spherically symmetric (SSS) setups. There are two types of approximate SSS solutions, applicable in different regimes. First, for large values of the graviton mass, the equations of motion can be solved perturbatively on local scales, resulting in the ordinary Newtonian potential plus an exponentially decaying Yukawa term. Second, if the graviton mass is small enough, the linear approximation breaks down and nonlinear terms must be added. This results in the Vainshtein screening mechanism, restoring GR phenomenology on local scales. To have a working Vainshtein mechanism, the parameter space must be constrained and we show that the $B_1B_2B_3$-models are the most minimal models compatible with this requirement plus a consistent cosmology. This excludes many of the popular models in the literature, including all two-parameter models. Self-accelerating cosmologies (i.e., $B_0 = 0$) are still viable.

For the local (SSS) solutions, there are five different GR limits. The vanishing mixing angle $\theta \to 0$ is the GR limit for the full theory and hence also for the local solutions. In the large graviton mass limit $m_{\text{FP}} \to \infty$, we are in the linear regime and the Yukawa term is exponentially suppressed. In the small mass limit $m_{\text{FP}} \to 0$, the Vainshtein mechanism is effectively restoring GR results at all length scales. However, this limit is forbidden by the Higuchi bound (see Chapter 3). Also in the large parameter limits $\alpha \to \infty$ and $\beta \to \infty$, the Vainshtein screening is effective at all length scales, restoring GR. Due to the several GR limits, it is not possible to constrain individual parameters from solar-system tests or galactic tests. However, in the literature it is usually assumed (implicitly) that $\mathcal{O}(\alpha) = \mathcal{O}(\beta) = 1$, in which case it is possible to rule out a region in the θm_{FP}-plane, see Fig. 4.5.

Chapter 5

Gravitational waves

Gravitational waves (GWs) are ripples in space-time, propagating away from its source according to a wave equation. In GR, Jebsen–Birkhoff's theorem prevents spherically symmetric systems from emitting gravitational waves. This is contrary to BR in which there is no corresponding theorem [2]. However, if a Vainshtein screening mechanism is active close to the source, we expect GW emission from spherically symmetric systems to be heavily suppressed [133]. In GR, gravity is massless meaning that GWs propagate at the speed of light while in bimetric gravity there is both a massless and a massive spin-2 field, giving rise to a rich phenomenology including gravitational wave oscillations and gravitational wave echoes. To date (2022), the LIGO/Virgo collaboration has observed 90 confident gravitational wave events in the redshift range $0.01 \lesssim z \lesssim 1$ with an electromagnetic counterpart signal in one case [134–139]. These observations can be used to place constraints on gravitational theories with modified GW propagation. In this chapter, we analyze the stability and phenomenology of gravitational waves and reproduce selected results in the literature.

5.1 Equations of motion

We consider gravitational waves propagating on a spatially flat cosmological background. There are two polarization modes of the massless graviton and, in principle, five polarization modes of the massive graviton. However, due to the Vainshtein screening mechanism, we expect the scalar and vector modes of the massive graviton to be heavily suppressed at the source of a binary merger system. This can be seen from eq. (4.11) by noting that the merger system is much denser than the solar system and that we require the solar system to lie within its own Vainshtein radius. Thus the merger system should lie far inside

its Vainshtein radius where GR is restored. Therefore, we will only study the tensor modes of $g_{\mu\nu}$ and $f_{\mu\nu}$ as encapsulated in the transverse and traceless perturbations h_{ij} and $\widetilde{h}_{ij}$,[1]

$$ds_g^2 = a^2\left[-d\tau^2 + (\delta_{ij} + h_{ij})\, dx^i dx^j\right], \tag{5.1a}$$
$$ds_f^2 = c^2\widetilde{a}^2\left[-x^2 d\tau^2/y^2 + \left(\delta_{ij} + \widetilde{h}_{ij}\right) dx^i dx^j\right], \tag{5.1b}$$

with,

$$\partial_i h_{ij} = \partial_i \widetilde{h}_{ij} = h^i{}_i = \widetilde{h}^i{}_i = 0, \tag{5.2}$$

and,

$$x \equiv \widetilde{a}'/a', \quad y \equiv \widetilde{a}/a. \tag{5.3}$$

Here, we use conformal time τ which is related to physical time t via $dt = a\, d\tau$ and denote the derivative with respect to τ by a prime. The background quantities a, $\widetilde{a}$, x, and y are functions of τ only while the perturbations depend on all coordinates. For simplicity, we consider a gravitational wave traveling in the z-direction, for which the tensor perturbations take the form [140],

$$h_{ij}dx^i dx^j = h_+(dx^2 - dy^2) + 2h_\times dxdy, \tag{5.4a}$$
$$\widetilde{h}_{ij}dx^i dx^j = \widetilde{h}_+(dx^2 - dy^2) + 2\widetilde{h}_\times dxdy. \tag{5.4b}$$

The equations of motion for the "+" and "×" modes look exactly the same [140], so we skip these subscripts henceforth. The equations are [141],

$$h'' + \gamma_g h' + \left(m_g^2 - c_g^2\nabla^2\right) h = q_g\widetilde{h}, \tag{5.5a}$$
$$\widetilde{h}'' + \gamma_f \widetilde{h}' + \left(m_f^2 - c_f^2\nabla^2\right) \widetilde{h} = q_f h, \tag{5.5b}$$

where $\mathcal{H} \equiv a'/a$ is the conformal Hubble parameter and,

$$\gamma_g = 2\mathcal{H}, \qquad \gamma_f = 2\mathcal{H} + 3y'/y - x'/x, \tag{5.6a}$$
$$m_g^2 = \mathcal{H}^2 yQ/\widetilde{\Omega}_{\rm DE}, \qquad m_f^2 = \mathcal{H}^2 xQ/(\widetilde{\Omega}_{\rm DE} y^2\tan^2\theta), \tag{5.6b}$$
$$c_g^2 = 1, \qquad c_f^2 = x^2/y^2, \tag{5.6c}$$
$$q_g = m_g^2, \qquad q_f = m_f^2, \tag{5.6d}$$

and we have defined,

$$Q \equiv \sin^2\theta\, m_{\rm FP}^2\left[1 + 2\alpha + \beta - (\alpha+\beta)(x+y) + \beta xy\right]. \tag{5.7}$$

[1] As opposed to Chapter 4, here the perturbations are not canonically normalized.

5.2 Stability

On time scales where the background cosmology can be considered static, we can solve the GW equations (5.5) by a plane wave ansatz,

$$h = h_k e^{i(\omega\tau - kz)}, \quad \widetilde{h} = \widetilde{h}_k e^{i(\omega\tau - kz)}, \tag{5.8}$$

where h_k and $\widetilde{h}_k$ are the constant (with respect to time) amplitudes. Real eigenfrequencies (ω) imply oscillating modes and hence stable propagation. A complex ω with positive imaginary part gives a decaying, hence stable, amplitude while a negative imaginary part gives an exponentially growing amplitude corresponding to an unstable mode. Combining h and $\widetilde{h}$ into a column vector $\mathbf{u}^T \equiv (h, \widetilde{h})$, eq. (5.5) can be written on the form,

$$\left(-\omega^2 \mathbf{1} + i\omega \mathbf{A} + \mathbf{B}\right) \mathbf{u} = 0, \tag{5.9}$$

where the matrices $\mathbf{A}$ and $\mathbf{B}$ only depend on the background quantities and the wave number k,

$$\mathbf{A} \equiv \begin{pmatrix} \gamma_g & 0 \\ 0 & \gamma_f \end{pmatrix}, \quad \mathbf{B} \equiv \begin{pmatrix} m_g^2 + c_g^2 k^2 & -q_g \\ -q_f & m_f^2 + c_f^2 k^2 \end{pmatrix}. \tag{5.10}$$

To have solutions other than the trivial $h_k = \widetilde{h}_k = 0$, we demand,

$$\det\left(-\omega^2 \mathbf{1} + i\omega \mathbf{A} + \mathbf{B}\right) = 0. \tag{5.11}$$

Solving the equation gives us the eigenfrequencies.

5.2.1 Early times

Assuming a finite branch background cosmology, $y \to 0$ in the early universe. Expanding each matrix element to lowest order around $y = 0$,

$$\mathbf{A} = 2\mathcal{H} \begin{pmatrix} 1 & 0 \\ 0 & 4 + 3w_{\text{tot}} \end{pmatrix}, \qquad y \simeq 0, \tag{5.12a}$$

$$\mathbf{B} = \mathcal{H}^2 \begin{pmatrix} 3\tan^2\theta\, y^2 + \mathcal{H}^{-2}k^2 & -3\tan^2\theta\, y^2 \\ -3(4 + 3w_{\text{tot}}) & 3(4 + 3w_{\text{tot}}) + (4 + 3w_{\text{tot}})^2 \mathcal{H}^{-2} k^2 \end{pmatrix}, \quad y \simeq 0. \tag{5.12b}$$

Subhorizon scales. Subhorizon scales are defined by $k \gg \mathcal{H}$. Using eq. (5.12) and solving eq. (5.11) in the subhorizon limit yields the eigenfrequencies,

$$\omega_{1/2} = \pm k, \quad \omega_{3/4} = \pm(4 + 3w_{\rm tot})k, \tag{5.13}$$

showing that the tensor perturbations are stable on subhorizon scales in the early universe. The $\omega_{1/2}$ modes propagate at the speed of light while $\omega_{3/4}$ propagate superluminally with the speed $4 + 3w_{\rm tot}$. During radiation domination, the propagation speed is five times the speed of light. Note that it coincides with the f-null cone of the background cosmology, see eq. (3.41), so causality is not violated [97].

Superhorizon scales. In the superhorizon limit $k \ll \mathcal{H}$,

$$\omega_1 = 0, \quad \omega_2 = 2i\mathcal{H}, \quad \omega_{3/4} = \left[(4 + 3w_{\rm tot}) \pm \sqrt{4 + 15w_{\rm tot} + 9w_{\rm tot}^2}\right] i\mathcal{H}. \tag{5.14}$$

The first mode ω_1 gives a constant amplitude. The second mode ω_2 has an exponentially decaying amplitude. Since the half-life is of the order $\mathcal{H}^{-1}$ and the approximation only applies to time scales where $\mathcal{H}$ can be considered constant, this mode will be approximately constant on time scales smaller than $\mathcal{H}^{-1}$. The same goes for the third and fourth modes $\omega_{3/4}$. These modes have a positive imaginary part if $w_{\rm tot} > -4/3$, so they are stable.

5.2.2 Late times

At late times, $y \to 1$, and the Universe is in a de Sitter phase. Expanding the matrix elements around $y = 1$,

$$\mathbf{A} = 2\mathcal{H}\,\mathbf{1}, \qquad y \simeq 1, \tag{5.15a}$$

$$\mathbf{B} = \mathcal{H}^2 \begin{pmatrix} \sin^2\theta\, m_{\rm FP}^2/\Omega_\Lambda + \mathcal{H}^{-2}k^2 & -\sin^2\theta\, m_{\rm FP}^2/\Omega_\Lambda \\ -\cos^2\theta\, m_{\rm FP}^2/\Omega_\Lambda & \cos^2\theta\, m_{\rm FP}^2/\Omega_\Lambda + \mathcal{H}^{-2}k^2 \end{pmatrix}, \quad y \simeq 1. \tag{5.15b}$$

We can now solve eq. (5.11) for the eigenfrequencies with the result,

$$\omega_{1/2} = \left(i \pm \sqrt{(k/\mathcal{H})^2 - 1}\right)\mathcal{H}, \quad \omega_{3/4} = \left(i \pm \sqrt{m_{\rm FP}^2/\Omega_\Lambda + (k/\mathcal{H})^2 - 1}\right)\mathcal{H}. \tag{5.16}$$

Subhorizon scales. In the subhorizon limit $k \gg \mathcal{H}$,

$$\omega_{1/2} = i\mathcal{H} \pm k, \quad \omega_{3/4} = i\mathcal{H} \pm \sqrt{\mathcal{H}^2 m_{\rm FP}^2/\Omega_\Lambda + k^2}, \tag{5.17}$$

and the gravitational wave propagation is stable.

Superhorizon scales. In the superhorizon limit $k \ll \mathcal{H}$,

$$\omega_1 = 0, \quad \omega_2 = 2i\mathcal{H}, \quad \omega_{3/4} = \left(i \pm \sqrt{m_{\rm FP}^2/\Omega_\Lambda - 1}\right)\mathcal{H}. \tag{5.18}$$

Applying the Higuchi bound $m_{\rm FP}^2 > 2\Omega_\Lambda$, the square-root in $\omega_{3/4}$ is real and hence the gravitational wave propagation is stable also on superhorizon scales at late times.

To summarize, gravitational wave perturbations are stable on time scales where the background cosmology can be treated as constant. On time scales where the background expansion cannot be ignored, one can show that the tensor perturbations are stable and the plane wave equations of motion can be solved analytically on sub- and superhorizon scales in the early and late universe, see Ref. [140].

5.3 Binary mergers

Gravitational waves from binary mergers have been observed out to redshifts of $z \simeq 1$. Following Ref. [13], as an approximation, we assume that the background expansion of the Universe is in its late-time de Sitter phase with $y = 1$. The matrices $\mathbf{A}$ and $\mathbf{B}$ are given by eq. (5.15). Since $\mathbf{A}$ is proportional to the identity matrix in this case, $\mathbf{A}$ and $\mathbf{B}$ can be simultaneously diagonalized using,

$$\hat{\mathbf{A}} = \mathbf{J}\mathbf{A}\mathbf{J}^{-1} = \mathbf{A}, \quad \hat{\mathbf{B}} = \mathbf{J}\mathbf{B}\mathbf{J}^{-1} \tag{5.19}$$

where $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are diagonal matrices and the transformation is given by $\mathbf{J}$. In fact,

$$\hat{\mathbf{A}} = 2\mathcal{H}\,\mathbf{1}, \quad \hat{\mathbf{B}} = \begin{pmatrix} k^2 & 0 \\ 0 & k^2 + \mathcal{H}^2 m_{\rm FP}^2/\Omega_\Lambda \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} \cos^2\theta & \sin^2\theta \\ 1 & -1 \end{pmatrix}. \tag{5.20}$$

We define new variables h_1 and h_2 according to,

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \cos^2\theta & \sin^2\theta \\ 1 & -1 \end{pmatrix} \begin{pmatrix} h \\ \widetilde{h} \end{pmatrix}, \tag{5.21}$$

or, equivalently,

$$\begin{pmatrix} h \\ \widetilde{h} \end{pmatrix} = \begin{pmatrix} 1 & \sin^2\theta \\ 1 & -\cos^2\theta \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}. \tag{5.22}$$

With the new variables $h_{1/2}$, eqs. (5.5) decouple and the equations of motion read,

$$h_1'' + 2\mathcal{H}h_1' + k^2 h_1 = 0, \tag{5.23a}$$
$$h_2'' + 2\mathcal{H}h_2' + k^2 h_2 = -\mathcal{H}^2 m_{\rm FP}^2 h_2/\Omega_\Lambda. \tag{5.23b}$$

The first-order time derivatives can be eliminated by yet another choice of variables,

$$u_1 \equiv a h_1, \quad u_2 \equiv a h_2. \tag{5.24}$$

Ignoring the background expansion, we can set $a = 1$ and use $\mathcal{H}^2 = H^2 = H_0^2\Omega_\Lambda$ (which is true in the late-universe de Sitter phase) to obtain,

$$u_1'' + k^2 u_1 = 0, \tag{5.25a}$$
$$u_2'' + \left(k^2 + \hat{m}_{\rm FP}^2\right) u_2 = 0, \tag{5.25b}$$

where $\hat{m}_{\rm FP} \equiv H_0 m_{\rm FP}$ is the dimensionful graviton mass (i.e., the inverse Compton wavelength). Hence, u_1 (h_1) describes the propagation of a massless GW and u_2 (h_2) and describes the propagation of a massive GW. The physical GW, h, is a linear combination of the massless and massive modes, see eq. (5.22). This leads to novel phenomena including gravitational wave oscillations [13] and gravitational wave echoes [80]. This is similar to neutrinos in the standard model of particle physics, having a mixing between the flavor eigenstates and the mass eigenstates, giving rise to neutrino oscillations. At the moment, we just note that the massless part of h propagates with the speed of light and thus always arrives at the same time as any electromagnetic counterpart, provided that they were emitted at the same time. In other words, there are no trivial constraints from the observed speed of gravitational waves [38].

If we solve eq. (5.25) with the initial conditions $h(t = 0) = h_k e^{-ikz}$ and $\widetilde{h}(t = 0) = 0$ and transform back to $h_{1/2}$ we get,

$$h_1 = \cos^2\theta\, h_k e^{i(\omega_1 t - kz)}, \tag{5.26a}$$
$$h_2 = h_k e^{i(\omega_2 t - kz)}, \tag{5.26b}$$

with,

$$\omega_1 \equiv k, \quad \omega_2 \equiv \sqrt{k^2 + \hat{m}_{\rm FP}^2}, \tag{5.27}$$

and we used $a = 1$ to identify $\tau = t$. The group velocities of h_1 and h_2 are given by,

$$v_{g,1} = \frac{\partial\omega_1}{\partial k} = 1, \quad v_{g,2} = \frac{\partial\omega_2}{\partial k} = \frac{1}{\sqrt{1 + \hat{m}_{\rm FP}^2/k^2}} < 1. \tag{5.28}$$

Note that we could have arrived at eq. (5.28) already from eq. (5.17) by using $\mathcal{H}^2 = H^2 = H_0^2\Omega_\Lambda$. As expected, the massless mode, h_1, propagates with

the speed of light while the massive mode, h_2, propagates subluminally. Since the group velocity of the massive mode h_2 depends on the wave number, k, a wave packet (a linear combination of plane waves with different wave numbers) experiences dispersion. For example, a Gaussian wave packet is smeared out as it propagates, see Fig. 5.2. In the limit $\hat{m}_{\rm FP} \ll k$,

$$\omega_2 = k + \Delta\omega, \quad \Delta\omega \equiv \hat{m}_{\rm FP}^2/(2k), \tag{5.29}$$

thus,

$$v_{g,2} = 1 - \Delta v_g, \quad \Delta v_g \equiv \hat{m}_{\rm FP}^2/(2k^2). \tag{5.30}$$

We have presented the plane wave solutions. The propagation of a wave packet is a linear combination of all plane waves of different wave numbers k,

$$h_1(t,z) = \cos^2\theta \int_{-\infty}^{\infty} h_k e^{i(\omega_1 t - kz)} dk, \tag{5.31a}$$

$$h_2(t,z) = \int_{-\infty}^{\infty} h_k e^{i(\omega_2 t - kz)} dk. \tag{5.31b}$$

Using eq. (5.22), we write this in terms of the metric perturbations as,

$$h(t,z) = \int_{-\infty}^{\infty} h_k \left[\cos^2\theta\, e^{i\omega_1 t} + \sin^2\theta\, e^{i\omega_2 t}\right] e^{-ikz} dk, \tag{5.32a}$$

$$\widetilde{h}(t,z) = \int_{-\infty}^{\infty} h_k \cos^2\theta \left[e^{i\omega_1 t} - e^{i\omega_2 t}\right] e^{-ikz} dk. \tag{5.32b}$$

Note that the only bimetric parameters that enter the binary merger systems are the mixing angle, θ, and the graviton mass, $m_{\rm FP}$.

5.3.1 Gravitational wave oscillations

Let us consider a plane wave solution with wave number k, see eq. (5.26). The perturbations of the metrics can be written,

$$h = h_k e^{i(\omega_1 t - kz)} \left[\cos^2\theta + \sin^2\theta\, e^{i\Delta\omega\, t}\right], \tag{5.33a}$$

$$\widetilde{h} = h_k \cos^2\theta\, e^{i(\omega_1 t - kz)} \left[1 - e^{i\Delta\omega\, t}\right], \tag{5.33b}$$

where $\Delta\omega \equiv \omega_2 - \omega_1$. Assuming that the difference $\Delta\omega$ between the eigenfrequencies is small, the perturbations h and $\widetilde{h}$ can be written as plane waves of high frequency $(\omega_1 + \omega_2)/2$ with the amplitude modulated by an envelope of low frequency $(\omega_2 - \omega_1)/2$, see Fig. 5.1. Note that $\Delta\omega$ is guaranteed to be small if $\hat{m}_{\rm FP} \ll k$, which we assume henceforth. With $k \sim \omega \sim 100\,{\rm Hz}$, this assumption amounts to $\hat{m}_{\rm FP} \ll 10^{-14}\,{\rm eV}$ (i.e., $m_{\rm FP} \ll 10^{19}$), where we used $H_0 \sim 10^{-17}\,{\rm Hz}$.

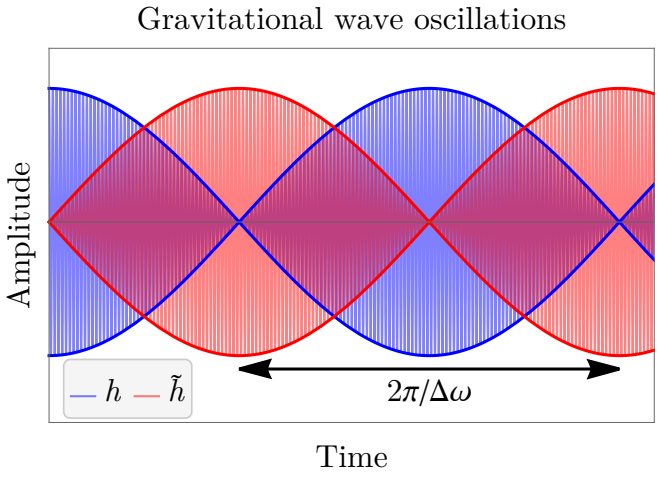


Figure 5.1: Gravitational wave oscillations of a plane wave solution at a fixed point in space. Here, $\Delta\omega = 10^{-2}$ and $\theta = \pi/4$ (i.e., maximal oscillations of the envelope). There are high-frequency oscillations with $\omega = (\omega_1 + \omega_2)/2$ and a low-frequency envelope with $\omega = (\omega_2 - \omega_1)/2$.

Due to the infinite extent of the plane wave, the low-frequency oscillations continue indefinitely. In practice however, the gravitational wave from a binary merger consists of a wave packet of finite extent. In this case, the envelope modulation (i.e., the low-frequency oscillations) continues as long as the wave packets of the massless and massive gravitons have not separated (decohered), see Fig. 5.2. The envelope oscillates on a length scale $L_{\rm osc} \equiv \pi/\Delta\omega \sim 2\pi\omega/\hat{m}_{\rm FP}^2$ or, equivalently,

$$L_{\rm osc} \sim \left(\frac{\omega}{100\,\text{Hz}}\right)\left(\frac{10^{-23}\,\text{eV}}{\hat{m}_{\rm FP}}\right)^2 \text{Gpc}. \tag{5.34}$$

If $L_{\rm osc}$ exceeds the Hubble radius ($\simeq 5\,$Gpc), there is no amplitude modulation. For a binary system with $\omega \sim 100\,$Hz, the graviton mass must satisfy $\hat{m}_{\rm FP} \gtrsim 10^{-23}\,$eV (i.e., $m_{\rm FP} \gtrsim 10^{10}$) for the gravitational wave oscillations to be influential.

Example. Let the Compton wavelength of the massive graviton be of the order of the Hubble radius, that is, $\hat{m}_{\rm FP} = H_0$. This is a common choice that allows for viable cosmological background solutions (see Section 3.6). The frequency of a binary merger is typically $\omega \sim 100\,$Hz. From eq. (5.34), it follows that the oscillation length of the envelope exceeds the radius of the observable universe by many orders of magnitude, hence there is no modulation of the

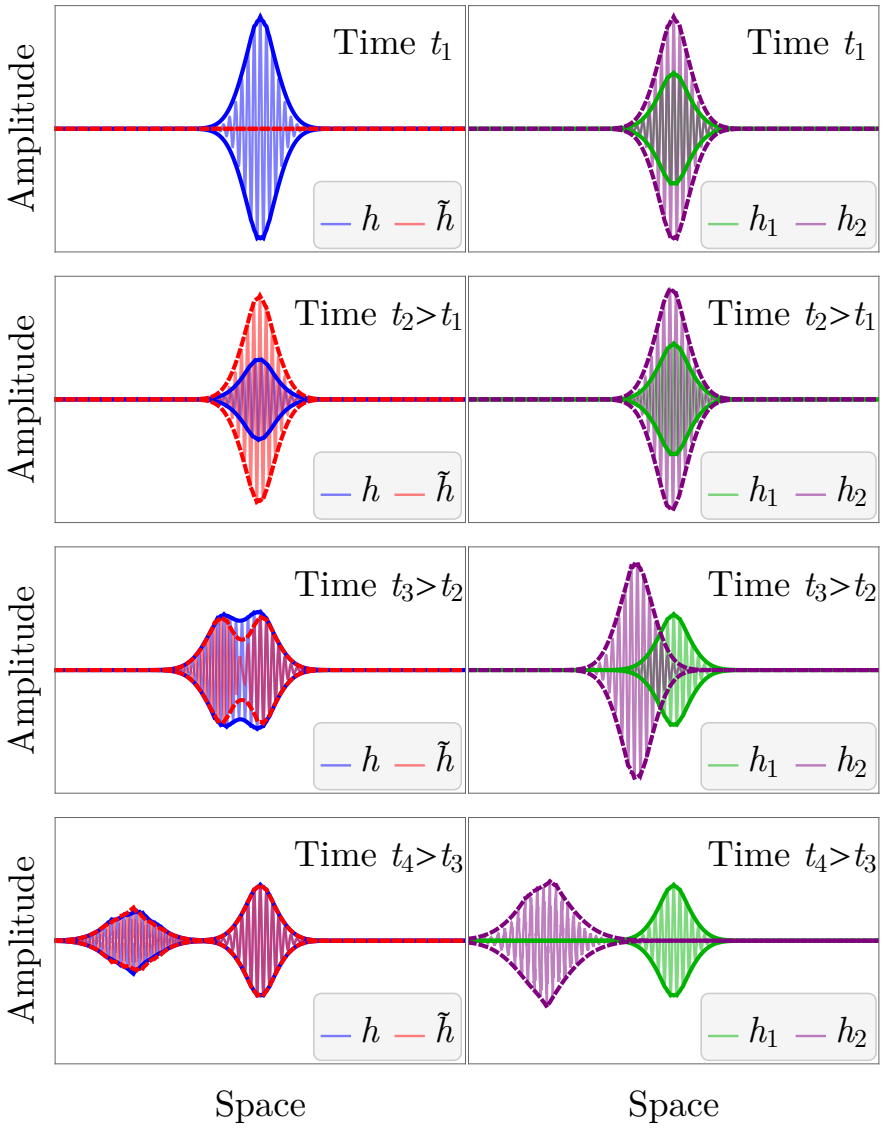


Figure 5.2: Propagation of bimetric gravitational waves. Here, $\theta = \pi/4$. *Left*: The metric perturbations. *Right*: The mass eigenstates. (The spatial coordinates follow the massless wave.)

amplitude. The propagation speed of the massive GW is given by eq. (5.30), $v_{g,2} = 1 - \Delta v_g$. Here, $\Delta v_g \sim 10^{-38}$ which is an undetectably small deviation from the speed of light. Hence, for $\hat{m}_{\rm FP} = H_0$ the bimetric modification to the gravitational wave propagation is insignificant as far as binary mergers are concerned.

□

Following Ref. [13], we derive an approximate formula that allows us to compute the physical waveform $h(t, z)$ in BR, assuming that the initial waveform generated at the source is the same as in GR. This assumption is justified by the fact that the system lies far inside its Vainshtein radius. In GR, a plane wave propagates with constant amplitude and the waveform is preserved (ignoring the evolution of the background), while in massive bigravity, the amplitude oscillates due to the mixing of the mass eigenstates, see Fig. 5.1.

Let us calculate how much the amplitude of a plane wave is suppressed over a time scale, $T_{\rm obs}$, which is much greater than the period of the high frequency oscillations, $2\pi/(\omega_1 + \omega_2)$, but much smaller than the period of the envelope, $2\pi/(\omega_2 - \omega_1)$. That is,

$$2\pi/(\omega_1 + \omega_2) \ll T_{\rm obs} \ll 2\pi/(\omega_2 - \omega_1). \tag{5.35}$$

The existence of such a time scale is guaranteed by the assumption that $\hat{m}_{\rm FP} \ll k$. We integrate out the high frequency oscillations and obtain the envelope function,

$$A^2(k) \equiv \left\langle h^2 \right\rangle = \cos^4\theta \left[1 + \tan^4\theta + 2\tan^2\theta \cos\left(\frac{\Delta\omega(k)}{H_0}\ln(1+z)\right)\right], \tag{5.36}$$

see Ref. [142] for a detailed derivation. Here, z is the redshift of the source and the angle brackets denote the time average. To derive this formula, we used eq. (5.22) with the plane wave solutions (5.26) and our late-time approximation which allows us to identify $t = \ln(1+z)/H_0$. Since a plane wave in GR propagates with constant amplitude (ignoring the background expansion), A is the factor by which the GR wave is suppressed in BR and we also refer to A as the "amplitude suppression factor".

If the gravitational wave can be approximated as a plane wave, we can directly multiply the GR waveform by A to get the physical perturbation h. Hence, for a plane wave of wave number k at a fixed point in space (e.g., the position of an observer),

$$h^{\rm BR}(t) = A(k)h^{\rm GR}(t) = A(k)h_k^{\rm GR}e^{i\omega_{\rm GR}t}, \quad \text{(plane wave)}. \tag{5.37}$$

Note that $\omega_{\rm GR} = k$. On the other hand, a wave packet is composed of plane waves of different wave numbers. In this case, the bimetric waveform is (cf.

eq. (5.32)),

$$h^{\rm BR}(t) = \int_{-\infty}^{\infty} dk\, A(k) h_k^{\rm GR} e^{ikt} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dk dt'\, e^{ik(t-t')} A(k) h^{\rm GR}(t'), \quad (5.38)$$

where we used the Fourier transform to express $h_k^{\rm GR}$ in terms of $h^{\rm GR}(t)$. Hence, given the GR wave, $h^{\rm GR}(t)$, the BR wave, $h^{\rm BR}(t)$, is obtained from eq. (5.38) together with eq. (5.36).

5.3.2 Gravitational wave echoes

The massive wave lags behind the massless one as they propagate. Due to the finite extent of the wave packet from a binary merger, after some time the massless and massive modes separate (decohere), see Fig. 5.2. This happens roughly when they have propagated a distance $L_{\rm dec} \sim W/\Delta v_g$, where W is the width of the wave packet and Δv_g is the difference in speeds between the modes [13]. Using eq. (5.30), this can be expressed as,

$$L_{\rm dec} \sim \left(\frac{W}{0.1\,{\rm s}}\right)\left(\frac{\omega}{100\,{\rm Hz}}\right)^2\left(\frac{10^{-22}\,{\rm eV}}{\hat{m}_{\rm FP}}\right)^2 {\rm Gpc}. \quad (5.39)$$

When the modes decohere, the low-frequency oscillations stop and the physical perturbation, h, consists of a massless wavepacket propagating ahead of a massive wavepacket. If the amplitude of both modes are detectable, there is a possibility of detecting GW echoes (depending on the waveform of the dispersive, i.e. massive, wave and the difference between arrival times compared with the observing time of the detectors). For a signal of width $W \sim 0.1\,{\rm s}$ and frequency $\omega \sim 100\,{\rm Hz}$ from a source at a distance 1 Gpc, the signal is decoherent if $\hat{m}_{\rm FP} \gtrsim 10^{-21}\,{\rm eV}$ (i.e., $m_{\rm FP} \gtrsim 10^{12}$).

If $\theta < 45°$, the amplitude of the massless mode $h_1 \sim \cos^2\theta$ is greater than that of the massive mode $h_2 \sim \sin^2\theta$. If the massless mode is detected with a signal-to-noise ratio ${\rm SNR}_1 \equiv {\rm SNR}$ and the threshold for detectability is $\widehat{\rm SNR}$, the massive mode can be detected provided that $h_1/h_2 = {\rm SNR}_1/{\rm SNR}_2 > {\rm SNR}/\widehat{\rm SNR}$. This condition can be expressed as a constraint on the mixing angle, $\tan^2\theta > \widehat{\rm SNR}/{\rm SNR}$. Similarly, in the case $\theta > 45°$, we get $1/\tan^2\theta > \widehat{\rm SNR}/{\rm SNR}$ for detectability. Summarizing, a GW echo can be detected if,

$$\theta_{\rm min} < \theta < \theta_{\rm max}, \quad \text{with} \begin{cases} \theta_{\rm min} \equiv \arctan\sqrt{\widehat{\rm SNR}/{\rm SNR}} \\ \theta_{\rm max} \equiv \arctan\sqrt{{\rm SNR}/\widehat{\rm SNR}} \end{cases}. \quad (5.40)$$

5.3.3 Observational constraints

Propagation velocity. In Ref. [38], a constraint on the propagation speed of gravitational waves was set, based on the arrival time of the electromagnetic counterpart of the GW170817 signal. The constraint is one part in 10^{15}. In the regime of large mixing angle, θ, only the massive mode propagates, so even if the massless mode arrives "on time", it will not be detected due to its suppressed amplitude. This is the case if $\theta > \theta_{\max}$, see eq. (5.40). For GW170817, the signal-to-noise ratio is SNR = 32 [134]. Here, we set $\widehat{\text{SNR}} = 1$, and thus only the massive mode is detected if $\theta > 80°$. In this case, we must require $v_{g,2} < 10^{-15}$. Using eq. (5.30), this excludes the region,

$$\theta > 80°, \quad \hat{m}_{\text{FP}} \gtrsim 10^{-21}\,\text{eV}, \tag{5.41}$$

see Fig. 5.3. To get eq. (5.41), we assumed $k \sim \omega$ and set ω to the lowest frequency from the signal, that is $\omega = 2\pi \times 24\,\text{Hz}$.

Distance to source. The massless wave propagates with unchanged waveform (no dispersion, see Fig. 5.2) but from eq. (5.32), we see that its amplitude is suppressed by a constant factor of $\cos^2\theta$ compared with GR, which results in an overestimation of the reported distances to the sources. With no independent means of estimating this distance (which is the case for all GW events except GW170817), this phenomenon does not lead to any constraints. For the GW170817 event with an electromagnetic counterpart, the host galaxy is identified as NGC 4993 and a GW-independent estimate of the luminosity distance is estimated from its Hubble flow velocity and two measurements of the Hubble constant [44, 143]. The combined result is $(42.9 \pm 3.2)\,\text{Mpc}$ (i.e., $\simeq 7\,\%$ precision) [38]. This is consistent with the distance determined with GW data alone (assuming GR), which is $40^{+8}_{-14}\,\text{Mpc}$ (corresponding to $z = 0.008^{+0.002}_{-0.003}$) [38].

The GW170817 signal has a temporal width of $W \sim 100\,\text{s}$. Setting the frequency equal to the minimal frequency of the signal, $\omega = 2\pi \times 24\,\text{Hz}$, it follows from eq. (5.39) that the massless and massive modes are decoherent if $\hat{m}_{\text{FP}} \gtrsim 10^{-20}\,\text{eV}$. The amplitude of the perturbation is proportional to $1/d_L$, with d_L the luminosity distance, so h should not be suppressed by more than $7\,\%$ in the decoherent regime. To ensure this, we must exclude the region,

$$15° \lesssim \theta \lesssim 75°, \quad \hat{m}_{\text{FP}} \gtrsim 10^{-20}\,\text{eV}. \tag{5.42}$$

Echoes. If the massless and massive modes are decoherent, the massive wave arrives at the detector as an echo of the massless wave. The amplitude of the massive wave packet is suppressed by a factor of $\sin^2\theta$ compared with a GR signal. Due to dispersion, the amplitude of the massive wave also evolves

as it propagates. Depending on the value of θ, the echo may or may not be detectable. The threshold for detection of both modes is set by the SNR according to eq. (5.40). For the GW150914 event, SNR = 24 [144], so a GW echo can be detected only if $10^\circ \lesssim \theta \lesssim 80^\circ$. If $\theta \gtrsim 80^\circ$, only the massive wave is detected and if $\theta \lesssim 10^\circ$, only the massless wave is detected.

For the GW150914 event, the GR-inferred luminosity distance is $L_{\rm GR} = 440^{+160}_{-180}\,{\rm Mpc}$ [145]. This can be expressed in terms of the physical luminosity distance as $L = L_{\rm GR}/\cos^2\theta$. Setting $\omega = 2\pi \times 100\,{\rm Hz}$ and $W = 0.3\,{\rm s}$, the modes are decoherent if $\hat{m}_{\rm FP} \gtrsim 10^{-21}\,{\rm eV} \times \cos\theta$ (i.e., $m_{\rm FP} > 10^{12} \times \cos\theta$). Provided that no GW echoes have been observed, we exclude the region,

$$10^\circ \lesssim \theta \lesssim 80^\circ, \quad \hat{m}_{\rm FP}/\cos\theta \gtrsim 10^{-21}\,{\rm eV}, \tag{5.43}$$

see Fig. 5.3. However, it should be stressed that the above conditions are only necessary conditions for detectability of an echo. First, the GW interferometers of LIGO/Virgo observe during finite time intervals, potentially missing an echo during a shut down. Second, due to dispersion, the waveform of the massive mode can change substantially as it propagates and since the current detectors are not designed to look for such signals, it is possible that they are not detected even if their amplitude is above the signal-to-noise threshold. For example, the frequency of a binary system increases with time as the system merges, so the signal emitted initially travels at a lower speed than the signal emitted later, see eq. (5.30). As the high-frequency part of the signal catches up on the low-frequency part, this can potentially lead to a gravitational shock wave at some distance from the source. At an even later stage, the low and high frequency signals switch places, possibly leading to an inversion of the initial waveform. For these reasons, we consider the exclusion (5.43) as an indication rather than a strict constraint. Further study is needed to consolidate the constraint.

Waveform. In the regime where there are low-frequency GW oscillations, the massless and massive modes interfere to produce the physical waveform, as explained in Section 5.3.1. Given that the observed GW is compatible with GR, we exclude regions in parameter space where the observed GW deviates more than 10 % from the GR wave (here, GW150914). The bigravity GW can be calculated from the corresponding GR wave using eq. (5.38) and the amplitude suppression factor (5.36). The details are explained in Ref. [146]. The result can be found in Fig. 5.3.

Gravitational wave oscillations occur when the distance to the source exceeds the length scale $L_{\rm osc}$ (5.34) but is below the decoherence length $L_{\rm dec}$ (5.39). For GW150914 with $\omega = 2\pi \times 100\,{\rm Hz}$, $W = 0.3\,{\rm s}$, and $L_{\rm GR} = 440\,{\rm Mpc}$, this is the case when,

$$10^{-23}\,{\rm eV} \lesssim \hat{m}_{\rm FP}/\cos\theta \lesssim 10^{-21}\,{\rm eV}. \tag{5.44}$$

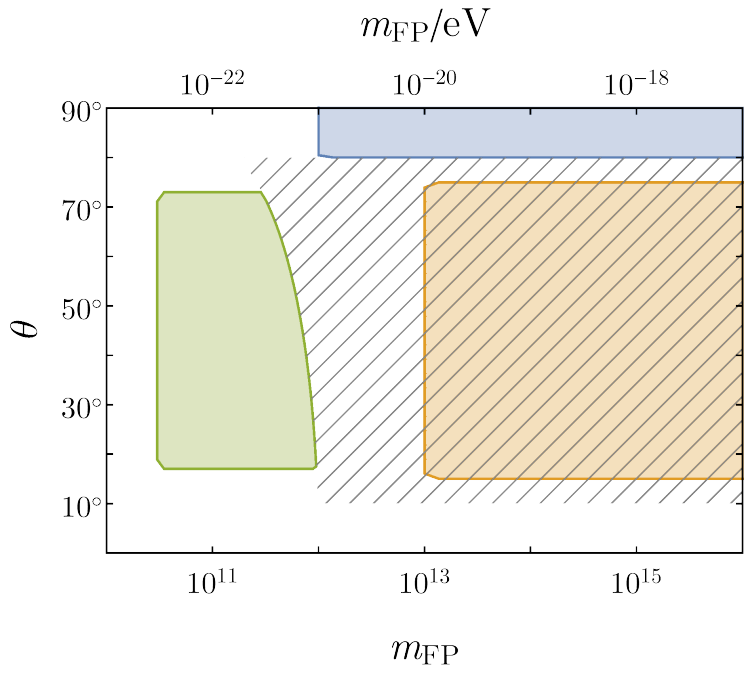


Figure 5.3: Exclusion plot based on gravitation wave observations. Blue: propagation velocity. Orange: distance to source. Green: gravitational waveform. Hatched: GW echoes. NB! Due to the dispersion of the massive wave, it is not clear whether a full analysis excludes the whole hatched region. It is a conservative constraint, that is, guaranteeing the absence of GW echoes.

5.4 Summary

By calculating the eigenfrequencies of the gravitational waves we concluded that the tensor perturbations are stable at the linear level.

In bimetric gravity, a gravitational wave consist of a combination of a massless and a massive mode, see Fig. 5.2. The mixing angle θ and the graviton mass $m_{\rm FP}$ are the only bimetric parameters entering the phenomenology of binary merger systems. In the range $10^{-23}\,{\rm eV} \lesssim \hat{m}_{\rm FP} \lesssim 10^{-20}\,{\rm eV}$ (i.e., $10^{10} \lesssim m_{\rm FP} \lesssim 10^{13}$), the mixing of the mass eigenstates gives rise to low-frequency oscillations modulating the waveform. If $\hat{m}_{\rm FP} \gtrsim 10^{-20}\,{\rm eV}$, the mass eigenstates decohere before reaching the detector, hence giving rise to a gravitational wave echo, with the massless wave arriving ahead of the massive one. However, the echo is only detectable if $10^\circ \lesssim \theta \lesssim 80^\circ$, otherwise the amplitude of the

weakest mode is too suppressed. That is, if $\theta \lesssim 10^\circ$, only the massless mode is observable and if $\theta \gtrsim 80^\circ$, only the massive mode is observable. When only the massive wave is detectable, we must ensure that its propagation speed is the speed of light $\pm 10^{-15}$ [38]. This excludes the region $(\hat{m}_{\rm FP} \gtrsim 10^{-21}\,{\rm eV}, \theta \gtrsim 80^\circ)$. Since there are two tensor perturbations in bimetric gravity, the amplitude of the GW is suppressed compared to GR. This leads to an overestimation of the reported distance to the binary systems. For the GW170817 event with an electromagnetic counterpart, there is an independent estimate of the distance which leads to an exclusion of the region $(\hat{m}_{\rm FP} \gtrsim 10^{-20}\,{\rm eV}, 15^\circ \lesssim \theta \lesssim 75^\circ)$. In the parameter range $10^{-23}\,{\rm eV} \lesssim \hat{m}_{\rm FP}/\cos\theta \lesssim 10^{-21}\,{\rm eV}$, there are GW oscillations modulating the amplitude on large length scales. This leads to a modified waveform compared with GR and to be compatible with observations, we exclude $(15^\circ \lesssim \theta \lesssim 75^\circ, 10^{-23}\,{\rm eV} \lesssim \hat{m}_{\rm FP} \lesssim 10^{-21}\,{\rm eV})$. In Fig. 5.3, we present a combined exclusion plot.

Here, the constraints were based on two events: GW150914 and GW170817. To date, there are 90 additional GW events. Future analyses should implement these to see how much the constraints are improved.

Chapter 6

Structure formation

The cosmological principle states that the Universe is homogeneous and isotropic on large scales. Observations confirm this principle on length scales $\gtrsim 100\,\mathrm{Mpc}$. On smaller scales, there is structure in form of galaxy clusters, galaxies, stars etc. The seeds of this structure is usually attributed to quantum fluctuations amplified by a short period of exponential expansion, called inflation, in an extremely early phase of the Universe ($\sim 10^{-36}\,\mathrm{s}$ after the Big Bang). Around 400 000 years after the Big Bang, the Universe became transparent to the photons and today we observe the imprint of the perturbations at that time in the temperature distribution of the photons constituting the cosmic microwave background radiation. To zeroth order, this radiation is isotropic and has a temperature of $\simeq 2.7\,\mathrm{K}$, but careful observations reveal anisotropies of the order 10^{-5} as first observed by the Cosmic Background Explorer (COBE) [147]. The growth of structure is a competition between the overall expansion of the Universe and the gravitational attraction between nearby objects, so a viable theory of gravity should be able to predict the statistical distribution of matter, given a set of initial conditions. The conventional framework to treat perturbations on length scales $1 - 100\,\mathrm{Mpc}$ is linear perturbation theory [39]. On smaller scales, $\lesssim 1\,\mathrm{Mpc}$, linear perturbation theory breaks down and N-body simulations are typically used. In this chapter, we study perturbations around cosmological background solutions in BR.

6.1 The linear equations

Linear perturbations around bimetric background cosmologies have been treated extensively in the literature, see for example Refs. [88, 89, 99, 140, 148–156]. In this section, we follow (partly) the outline of Ref. [140]. Here, instead of physical time, t, as used in Chapter 3, we use conformal time, τ, defined by $dt = a\, d\tau$. As we are interested in structure formation, we focus on scalar perturbations. Tensor perturbations are studied in Chapter 5. Vector perturbations are studied in Ref. [140] for example, where it is shown that they are decaying with time (assuming a finite branch cosmology).

Assuming no spatial curvature, the general ansatz for the metrics takes the form,

$$ds_g^2 = a^2\left[-(1+2\Phi)d\tau^2 + 2\partial_i B dx^i d\tau + \left((1-2\Psi)\delta_{ij} + 2\partial_{ij}^2 E\right)dx^i dx^j\right], \tag{6.1a}$$

$$ds_f^2 = c^2\widetilde{a}^2\left[-X^2(1+2\widetilde{\Phi})d\tau^2 + 2X\partial_i \widetilde{B} dx^i d\tau + \left((1-2\widetilde{\Psi})\delta_{ij} + 2\partial_{ij}^2 \widetilde{E}\right)dx^i dx^j\right]. \tag{6.1b}$$

Here, we have defined,

$$X \equiv x/y, \tag{6.2}$$

and $\partial_{ij}^2 \equiv \partial^2/\partial x^i \partial x^j$. Following Refs. [140, 157], we write the scalar perturbation of the perfect fluid stress–energy tensor in terms of a scalar field χ,

$$\delta T^0{}_0 = -(\rho + p)(3\Psi - \partial_{ii}^2 E - \partial_{ii}^2 \chi), \tag{6.3a}$$

$$\delta T^i{}_0 = -(\rho + p)\partial_i \chi', \tag{6.3b}$$

$$\delta T^0{}_i = (\rho + p)(\partial_i B + \partial_i \chi'), \tag{6.3c}$$

$$\delta T^i{}_j = w_m(\rho + p)(3\Psi - \partial_{kk}^2 E - \partial_{kk}^2 \chi)\delta^i{}_j. \tag{6.3d}$$

Here, prime denotes the derivative with respect to τ. The perturbed fields $(\Phi, \Psi, B, E, \widetilde{\Phi}, \widetilde{\Psi}, \widetilde{B}, \widetilde{E}, \chi)$ depend on all space-time coordinates while the background fields $(a, \widetilde{a}, X, \rho, p)$ depend on τ only. There are four scalar fields associated with each metric and one scalar field from the perturbed stress–energy tensor, adding up to a total of nine scalar fields. However, BR was constructed to avoid the scalar ghost mode of Boulware and Deser, so in fact there are only two propagating modes, corresponding to the matter perturbation and the scalar mode of the massive spin-2 field. So, seven of these fields must be auxiliary fields without dynamics.

At the linear level, the gauge freedom is the freedom of choosing coordinates of the perturbed space-time. An infinitesimal coordinate transformation

reads $x^\mu \to x'^\mu = x^\mu + \xi^\mu(x)$ with ξ^μ small. One can decompose ξ^μ into its scalar components, (ξ^0, ξ^S), and its vector components, ξ_i^V, where the latter is divergenceless, $\partial_i \xi_i^V = 0$. The decomposition reads, $\xi^\mu = (\xi^0, \partial_i \xi^S + \xi_i^V)$. The freedom of choosing coordinates can be used to eliminate two scalar fields by choosing ξ^0 and ξ^S. Here, we use it to set $\chi = \widetilde{\Psi} = 0$ (see Ref. [140] for details).

Seven fields remain, of which five should be auxiliary. In the equations of motion, only $(E, \widetilde{E}, \Psi)$ enter with second-order time derivatives [140]. Hence, $(\Phi, \widetilde{\Phi}, B, \widetilde{B})$ are auxiliary variables and can be eliminated from the equations of motion. Doing so, also Ψ turns out the be auxiliary and the remaining $(E, \widetilde{E})$ are the propagating scalar modes. With our gauge choice ($\chi = \widetilde{\Psi} = 0$), these modes coincide with the gauge invariant parameters [140],

$$\zeta = \frac{1}{3}\left(\partial_{ii}^2 E + \partial_{ii}^2 \chi\right), \quad \widetilde{\zeta} = \frac{1}{3}\left(\partial_{ii}^2 \widetilde{E} + \partial_{ii}^2 \chi\right). \tag{6.4}$$

Since the equations of motion are linear, it is convenient to study them in Fourier space, using the transformation,

$$f(\tau, x^i) = \frac{1}{(2\pi)^{3/2}} \int d^3k\, \hat{f}(\tau, k^i) e^{ik_i x^i}. \tag{6.5}$$

The equations of motion for the physical modes can be written formally as,

$$\mathbf{E}'' + \mathbf{A}\mathbf{E}' + \mathbf{B}\mathbf{E} = 0, \tag{6.6}$$

where $\mathbf{E}^T = (E, \widetilde{E})$ and $\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices depending on the background fields and the wavenumber k only. See Ref. [140] for the explicit expressions.

Quasistatic subhorizon limit

Here, we follow Ref. [152] and analyze the linear equations in the quasistatic subhorizon limit, defined by $k^2 f \gg \mathcal{H}^2 f \sim \mathcal{H} f' \sim f''$ for any field f. This limit is well-defined assuming that the linear equations are stable. To allow the same limit in the $f_{\mu\nu}$-sector, we also assume that $\mathcal{H} \sim \widetilde{\mathcal{H}}$, with the conformal Hubble parameters $\mathcal{H} \equiv a'/a$ and $\widetilde{\mathcal{H}} \equiv \widetilde{a}'/\widetilde{a}$. We also assume that $(\Phi, \widetilde{\Phi}, \Psi, \widetilde{\Psi})$ are of the same order of magnitude as $(kB, k\widetilde{B}, k^2 E, k^2 \widetilde{E})$. Here, we choose the Newtonian gauge with respect to the physical metric, that is, $B = E = 0$. Now, one can derive the growth parameters Q (not to be confused with the same symbol defined in Chapter 5), η, and γ, see Refs. [152, 158] for details.

Modified gravitational constant. Here, $Q(a,k)$ is the factor rescaling the gravitational constant in the Poisson equation,

$$-2\frac{k^2}{a^2}\Psi = Q(a,k)\kappa_g\bar{\rho}_m\delta, \tag{6.7}$$

where the density contrast, δ, is defined as,

$$\delta \equiv \frac{\rho_m - \bar{\rho}_m}{\bar{\rho}_m}, \tag{6.8}$$

where $\bar{\rho}_m$ is the background matter density. From eq. (6.7), we see that $Q = 1$ corresponds to the GR value.

Anisotropic stress. The bimetric stress–energy introduces an effective anisotropic stress, parameterized by $\eta(a,k)$,

$$\eta(a,k) \equiv \Psi/\Phi. \tag{6.9}$$

In GR, the anisotropic stress vanishes for all practical purposes (after matter-radiation equality), leading to $\Phi = \Psi$, so $\eta = 1$ corresponds to the GR value.

Growth rate. The growth rate, $f_g(a,k)$, is defined as,

$$f_g(a,k) \equiv \frac{d\,\ln\delta}{d\,\ln a}. \tag{6.10}$$

The parameters Q and η can be calculated analytically, see Ref. [152]. To find γ, we need to solve the conservation equation for δ (a differential equation).

Other common parameters are $\mu(a,k)$, modifying the gravitational constant in the Poisson equation for Φ,

$$-2\frac{k^2}{a^2}\Phi = \mu(a,k)\kappa_g\bar{\rho}_m\delta, \tag{6.11}$$

and the parameter $\Sigma(a,k)$, related to the gravitational lensing properties,

$$-\frac{k^2}{a^2}(\Phi+\Psi) = \Sigma(a,k)\kappa_g\bar{\rho}_m\delta. \tag{6.12}$$

6.2 Stability

In the early universe, the ratio of the scale factors approaches $y = 0$, assuming a finite branch background cosmology (see Chapter 3). Expanding around $y = 0$, assuming radiation domination $w_{\rm tot} = 1/3$, and taking the subhorizon limit ($k \gg \mathcal{H}$), the linear equations (6.6) can be approximated as,

$$E'' + \frac{k^2}{3}E = 0, \tag{6.13a}$$

$$\widetilde{E}'' + 6\mathcal{H}(\widetilde{E}' - E') - \frac{5k^2}{3}\widetilde{E} + 2k^2 E = 0, \tag{6.13b}$$

with analytical solutions,

$$E = c_\pm e^{\pm ik\tau/\sqrt{3}}, \quad \widetilde{E} = \widetilde{c}_\pm \frac{e^{\pm\sqrt{15}k\tau/3}}{(k\tau)^3} + E, \tag{6.14}$$

with $c_\pm$ and $\widetilde{c}_\pm$ being integration constants [140]. Hence, E oscillates while $\widetilde{E}$ exhibits an exponentially growing mode, invalidating linear perturbation theory as $\widetilde{E}$ violates the assumption of being a small perturbation on the background space-time. The exponential modes of $\widetilde{E}$ are due to a negative sign k^2-term in eq. (6.13b), hence it is referred to as a gradient instability.

On the other hand, in the late universe where $y \simeq 1$, subhorizon perturbations are stable [140], suggesting the possibility of pushing the unstable regime backwards in time by choosing $m_{\rm FP}$ to be large, under which the transition from $y \simeq 1$ to $y \simeq 0$ is pushed backwards in time, see Fig. 6.1.

Note that the exponentially growing mode does not imply that there is a physical instability, only that linear perturbation theory is invalid at these times. This is not completely surprising since this framework was constructed for GR cosmology, not general theories of gravity [159]. We also recognize the breakdown of linear perturbation theory from the static, spherically symmetric solutions in Chapter 4. In this case, there is a working screening mechanism that restores GR results below a certain radius when taking nonlinear terms into account. In Section 6.3, we present alternative ways of analyzing cosmological perturbation theory.

To determine the time before which the gradient instability is present, we follow Ref. [155] and make an exponential ansatz for the scalar perturbations, that is, $\mathbf{E} \propto e^{\omega \ln a}$, assuming that ω is independent of time. This assumption is valid provided that $|d\omega/d\ln a| \ll |\omega^2|$. The instability is on subhorizon scales, so we also make use of $k \gg \mathcal{H}$. With these assumptions, the scalar equations can be solved for the eigenfrequency, ω, with the result,

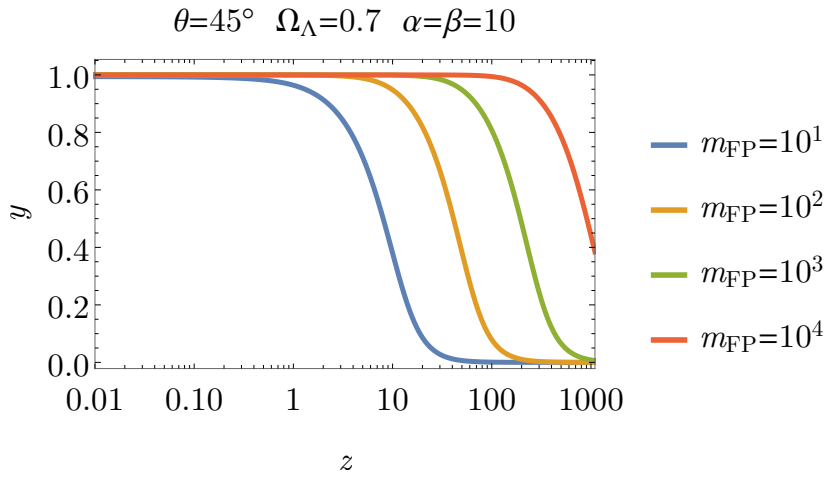


Figure 6.1: Ratio of the scale factors, y, as a function of redshift. For increasing values of the graviton mass, $m_{\rm FP}$, the transition from $y \simeq 1$ to $y \simeq 0$ is pushed backwards in time.

$$\omega^2 = \frac{1}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \left[\left(1+\frac{1}{\tan^2\theta\, y^2}\right)\frac{B_1 y - B_3 y^3}{3(1+w_{\rm tot})\Omega_m}\left(\frac{d\,\ln y}{d\,\ln a}\right)^2 + \right.$$
$$\left. -\frac{B_1 y + 4B_2 y^2 + 3B_3 y^3}{B_1 y + 2B_2 y^2 + B_3 y^3}\frac{d\,\ln y}{d\,\ln a} - 3\right]. \qquad (6.15)$$

If desired, the B-parameters can be written in terms of the physical parameters using eq. (2.14). When ω^2 is negative, the scalar modes are oscillating. When ω^2 is positive, the scalar modes are growing (or decreasing) exponentially. The transition between the stable and the unstable regime happens when ω crosses zero and we denote the redshift at which this happens with z'. Note that the expression within the square parenthesis can be written entirely in terms of y and the physical parameters, using eqs. (3.18) and (3.23). As an example, for the best-fit general model cosmology of Section 3.6 the scalar perturbations are unstable before $z' \simeq 0.03$, that is, relatively late in the history of structure formation. The same holds for the best-fit self-accelerating cosmology. For the best-fit $B_1B_2B_3$-model, the instability is present before $z' \simeq 0.3$.

Expanding around $y = 0$ (early times),

$$\omega^2 = (1+2w_{\rm tot})(k/\mathcal{H})^2 > 0, \quad y \simeq 0. \qquad (6.16)$$

Since $w_{\rm tot} > -1/2$, the scalar perturbations are unstable in the early universe (except during inflation when $w_{\rm tot} \simeq -1$ [124]). Expanding around $y = 1$ (late

times),

$$\omega^2 = -(k/\mathcal{H})^2 < 0, \quad y \simeq 1, \tag{6.17}$$

so the scalar perturbations are stable in the late universe.

We have analyzed perturbations around finite branch background cosmologies and concluded that, at early enough times, a gradient instability occurs when applying standard cosmological perturbation theory. On the other hand, infinite branch solutions do not have any gradient instability at the linear level [155]. However, they are plagued by the Higuchi ghost which renders them unviable. We face the inconvenient fact that bimetric cosmology either exhibits a ghost (infinite branch) or a gradient instability (finite branch). Since the latter is a practical challenge due to the framework used, rather than a theoretical problem, we continue using the finite branch solutions and analyze possible ways forward.

6.3 Ways forward

Exorcising the instability

The redshift at which the instability sets in, z', is determined entirely by the physical parameters and the matter equation of state (cf. eq. (6.15)). As argued in the previous section, we expect a large Fierz–Pauli mass, $m_{\rm FP}$, to push the instability backwards in time. Hence, choosing $m_{\rm FP}$ large enough, the instability can be pushed beyond any desired redshift, in particular, back to the inflationary epoch where $w_{\rm tot} \simeq -1$ and the instability is exorcised [124]. This idea was first discussed in Ref. [160] for the B_1B_2-models. However, for these models, the Fierz–Pauli mass and the mixing angle are correlated and the $m_{\rm FP} \to \infty$ limit coincides with the GR limit $\theta \to 0$ (keeping Ω_Λ fixed), as shown in Section 3.5. Exorcising the instability for these models is equivalent to taking the GR limit, in which case the theory looses most of its novel features. The obvious question is if the same conclusion holds for general BR models for which θ and $m_{\rm FP}$ are independent. To the knowledge of the author, this has not been studied in the literature, so the remaining results in this subsection are novel.

To analyze the onset of the gradient instability for general BR models, we make a grid scan in the physical parameter space, calculating z' for 2.1×10^6 points in the range presented in Tab. 6.1, considering only points compatible with viable background cosmology and an existing Vainshtein mechanism. The result is that the only possibility to exorcise the instability is to let $m_{\rm FP}$ be increasingly large, as shown in Fig. 6.2. This is consistent with the results from the B_1B_2-models although for a general BR model, large graviton mass does

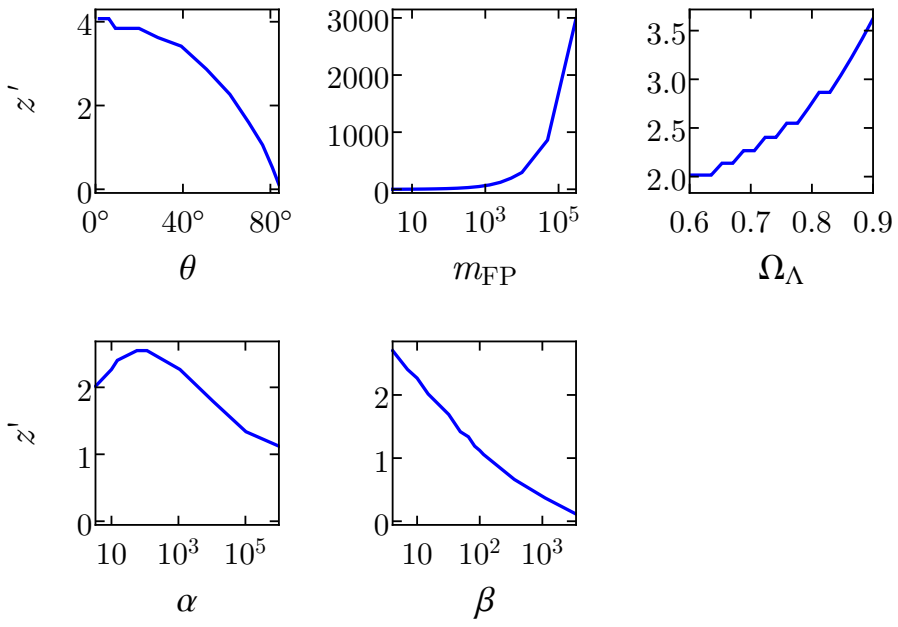


Figure 6.2: Plotting z' as a function of the physical parameters. In each panel, one physical parameter is varied while the others are held fixed at $\mathcal{O}(1)$ values. The qualitative behavior of these plots are general. For example, z' increases with $m_{\rm FP}$ also if the values of the other physical parameters are different from unity.

not imply small mixing angle, so there is still room for significant modifications of gravity in this limit. The effective cosmological constant, Ω_Λ, also has a positive correlation with z' but large values of Ω_Λ are observationally disfavored (see Section 3.6).

As an example, to push the gradient instability beyond photon decoupling, $z_* \simeq 1100$, we need $m_{\rm FP} \gtrsim 10^4$ (i.e., $m_{\rm FP} \gtrsim 10^{-29}\,{\rm eV}$), assuming that the other parameters are of order $\mathcal{O}(1)$. This means that we can use linear perturbation theory to calculate the (large-scale) growth of structure from z_* to today, provided that we know the CMB anisotropies at z_*. However, it is possible that the quadratic (or higher-order) perturbations are of the same order of magnitude as the linear perturbations. In this case, linear perturbation theory breaks down even if the linear perturbations stay small. This is discussed in the next subsection.

Since $m_{\rm FP} \to \infty$ is not a GR limit of bimetric gravity, there are observa-

Parameter	Range
θ	$[2^\circ, 88^\circ]$
$m_{\rm FP}$	$[10^{-1}, 10^4]$
Ω_Λ	$[0.6, 0.9]$
α	$[-580, 10^6]$
β	$[1, 10^6]$

Table 6.1: Scanning range for computing z'.

tional constraints on θ in this limit, coming from cosmological and gravitational wave observations. The latter provides slightly stronger constraints, namely $\theta \lesssim 15^\circ$ in the $m_{\rm FP} \to \infty$ limit, see Fig. 5.3.

Vainshtein mechanism

Consider the gravitational force due to a sub-horizon size spherical inhomogeneity (e.g., an overdense region). Assuming that linear perturbation theory is valid, we can integrate and transform eq. (6.7) back to real space with the result,

$$\Psi(r) = -Q(t,r)\frac{M}{r}, \tag{6.18}$$

for the spatial curvature Ψ, where M is the mass of the inhomogeneity. (Here, r is the physical distance.) On time scales much shorter than the Hubble time, we can ignore the overall expansion of the Universe. Approximating the inhomogeneity as static, the spatial curvature can also be written as in eq. (4.10b), that is,

$$\Psi(r) = -\frac{2M}{r} - \sin^2\theta\left(\frac{r}{\lambda_g}\right)^2\left[\mu(r) - \alpha\mu^2(r) + \frac{\beta}{3}\mu^3(r)\right], \quad r \ll \lambda_g. \tag{6.19}$$

Equations (6.18) and (6.19) describe the same system but the expressions are inconsistent, which is apparent from the fact that Q takes different values at different points in time while eq. (6.19) is always the same. This indicates that the linearized solution is inapplicable and hence we need to solve the nonlinear equations of motion. A convenient work-around was proposed in Ref. [161] and builds on the assumption that the Vainshtein mechanism is applicable on the growth of structure in the same way as it is on static spherically symmetric systems, see also Ref. [50,90] for similar ideas. In this case, GR is restored well inside the Vainshtein radius $r_V = (2M\lambda_g^2)^{1/3}$, or equivalently, if the density is

well above the Vainshtein density $\rho_V = \rho_c m_{\rm FP}^2$, see Chapter 4. In the cosmological context, we identify the density with $\delta\rho_m$ and hence GR is restored if the density contrast is greater than δ_V,

$$\delta_m > \delta_V, \quad \delta_V \equiv \rho_V/\bar{\rho}_m = m_{\rm FP}^2/\Omega_m(z). \tag{6.20}$$

Recall, $\Omega_m(z) \equiv \bar{\rho}_m(z)/\rho_c$. In the linearly unstable regime ($z > z'$), a low-density region with $\delta_m < \delta_V$ grows exponentially until it reaches the value δ_V where GR is restored. Since δ_V is a function of time, it sets a time-dependent lower bound on the density contrast δ_m, boosting any perturbation with an initial value less than δ_V. In Ref. [161], the authors analyzed the consequences for the matter power spectrum of such a cosmological Vainshtein mechanism for a B_1-model. Further investigation is needed to establish the consequences for general bimetric models.

Nonlinear solutions

When linear perturbation theory breaks down due to the gradient instability, it is necessary to solve the nonlinear equations of motion to obtain the evolution of cosmological perturbations. In principle, there are two types of solutions: numerical and analytical. We do not expect the latter to be achievable in general, but there is a particular example where it is, namely the evolution of a homogeneous overdensity, which we analyze below. In general, the equations of motion must be solved numerically. We study this topic in Chapter 7 and conclude that it is a challenging task. Partial results have been obtained in spherical symmetry [4, 5, 8, 162–167].

In case of shared spherical symmetry, a general bidiagonal ansatz is,

$$ds_g^2 = -dt^2 + A^2(t,r)dr^2 + B^2(t,r)d\Omega^2, \tag{6.21a}$$
$$ds_f^2 = -\widetilde{\alpha}^2(t,r)dt^2 + \widetilde{A}^2(t,r)dr^2 + \widetilde{B}^2(t,r)d\Omega^2. \tag{6.21b}$$

Since we are interested in the growth of structure, we let the stress–energy consist of pressureless dust, $T^\mu{}_\nu = \mathrm{diag}(-\rho_m(t,r), 0, 0, 0)$. Plugging into the equations of motion, from the tr-equation it follows,

$$A(t,r) = \frac{\partial_r B(t,r)}{\sqrt{1-k(r)}}, \tag{6.22}$$

where $k(r)$ is a function of the radius only. The conservation of stress–energy ($\nabla_\mu T^\mu{}_\nu = 0$) reads,

$$\partial_t \left[\rho_m(t,r)B^2(t,r)\partial_r B(t,r)\right] = 0. \tag{6.23}$$

That is, the expression within the square brackets is a conserved quantity. Inside a homogeneous overdensity (or underdensity) of radius r_*, an Oppenheimer–Snyder (OS) model [168], the matter density is a function of time only,

$$\rho_m = \rho_m(t), \quad r < r_*. \tag{6.24}$$

Evaluating the conserved quantity at some generic time t and at an initial time $t = 0$ and equating, we get,

$$\rho_m(t,r)B^2(t,r)\partial_r B(t,r) = \rho_{m,0}, \quad r < r_*, \tag{6.25}$$

where $\rho_{m,0} \equiv \rho_m(t)|_{t=0}$ and we used the freedom of choosing coordinates to set $B(t,r)|_{t=0} = r$. Integrating eq. (6.25), results in,

$$B^3(t,r) = \frac{\rho_{m,0}}{\rho_m(t)} r^3 + b(t), \tag{6.26}$$

with $b(t)$ being a function of time only. Demanding that B is regular at the origin, $r = 0$, implies that B is either even or odd as a function of r [169]. Hence, $b = 0$. Identifying the scale factor as, $a^3(t) \equiv \rho_{m,0}/\rho_m(t)$, it follows,

$$B(t,r) = a(t)r. \tag{6.27}$$

Hence, the physical metric is homogeneous and isotropic and the evolution of a homogeneous overdensity (or underdensity) is according to a cosmological background solution, inside the overdense region ($r < r_*$).

To summarize so far, we have shown that a homogeneous density perturbation evolves as a background cosmology and thus we obtain the analytical solution to the nonlinear equations of motion without any approximation. This solution can be used to estimate the growth of structure. In particular, since the background cosmology does not exhibit any exponential divergence (on the finite branch), the nonlinear evolution of these perturbations is regular as opposed to exponential runaway of the linear perturbations.

In Fig. 6.3, we plot the relative difference between the GR and BR density contrasts for the best-fit background cosmological models (see Section 3.6). We start with an initial overdensity $\delta(z_*) = 10^{-5}$ and set the growth rate (6.10) to unity and the scale factors of the background and the overdensity to be equal at $z_* = 1100$. To satisfy the equations of motion for the overdensity, there is a (positive) spatial curvature Ω_k in its Friedmann equation (3.14). For the best-fit background cosmologies, the density contrast predicted by BR cosmology differs at most a few percent from GR.

In GR (ΛCDM), the growth rate can be approximated as a power-law in $\Omega_m(z)/H^2(z)$,

$$f_g(z) \simeq \left[\left(\frac{H_0}{H(z)}\right)^2 \Omega_m(z)\right]^\gamma, \tag{6.28}$$

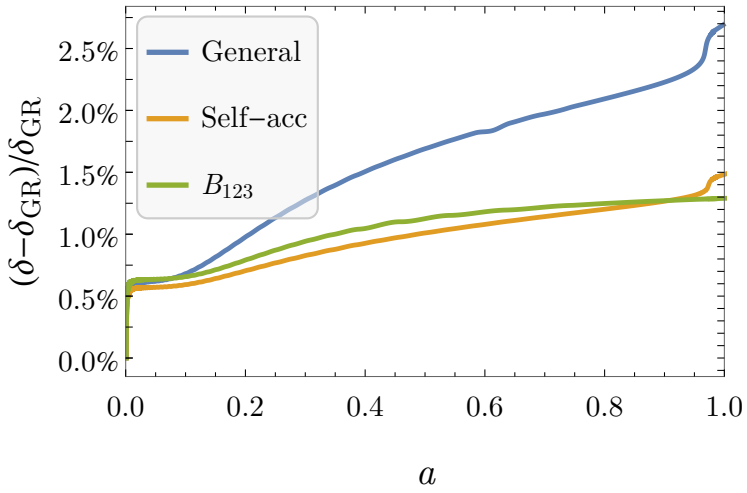


Figure 6.3: Relative difference in the density contrast, comparing BR cosmology with GR (ΛCDM). We plot the models that give the best-fit bimetric background cosmology (see Section 3.6). Blue: best-fit general model $(\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta) = (11°, 2.8, 0.71, -5.1, 34)$. Orange: best-fit self-accelerating model $(\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta) = (7°, 2.7, 0.70, -4.7, 29)$. Green: best-fit B_{123}-model $(\theta, m_{\rm FP}, \Omega_\Lambda, \alpha, \beta) = (6°, 1.6, 0.70, 19, 19)$. For these models, the relative difference in δ grows from zero to a few percent.

where the exponent γ, called the growth index, is a constant $\gamma_{\rm GR} \simeq 0.545$ [170]. The Euclid spacecraft (scheduled to launch in 2023) is forecasted to discriminate between models differing ± 0.03 in the growth index [158, 171, 172]. That is, for models where f_g can be fitted as in eq. (6.28) with a constant γ. This is not the case for modified gravity models in general, and massive bigravity in particular, whose growth index is redshift-dependent. An alternative parameterization of γ that is employed in Ref. [158] is,

$$\gamma(z) = \gamma_0 + \gamma_1 \frac{z}{1+z}, \tag{6.29}$$

where γ_0 and γ_1 are constants. Euclid is forecasted to distinguish (at 1σ) between models differing ± 0.2 in γ_0 and ± 0.4 in γ_1, assuming that the model can be approximated as in eq. (6.28) with eq. (6.29). In Tab. 6.2, we show the theory predictions of γ_0 and γ_1 for the best-fit background cosmologies although, for the general and self-accelerating models, the parameterization (6.29) does not provide a good fit to the growth rate. This is due to the temporary, rapid decrease in the dark energy equation of state at low redshifts (see Fig. 3.13) which leads to a rapid, temporary increase in f_g. This is shown

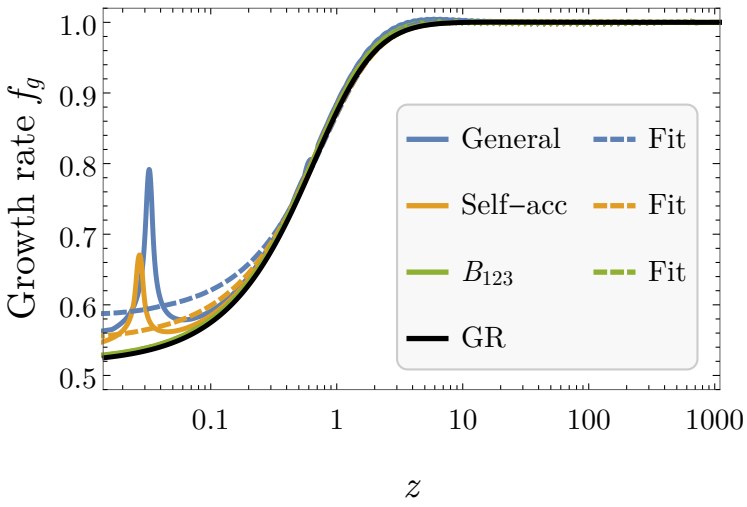


Figure 6.4: Growth rate as a function of redshift. Solid, colored curves: the three best-fit bimetric background cosmologies. Dashed, colored curves: fitting the growth rate with a function of the form $(\Omega_m(z)/H^2(z))^{\gamma(z)}$, with $\gamma(z)$ as in eq. (6.29). For the B_{123}-model, the fit overlaps with the analytical solution while this functional form cannot be used to fit the general and self-accelerating models due to their temporary spikes in f_g. Solid black curve: GR (ΛCDM) model with $\Omega_\Lambda = 0.7$.

in Fig. 6.4, where we plot the growth rate for the three best-fit background cosmologies. For the B_{123}-model, eq. (6.29) provides an excellent fit and its growth index is indistinguishable from GR (with respect to the forecasted Euclid sensitivity).

Alternatively, the growth rate can be parameterized as a step-wise function with a fixed value in each redshift bin,

$$f_g(z) \simeq f_i, \tag{6.30}$$

where i is an index that runs over the bins. In Ref. [158], this parameterization is employed by dividing the redshift range $0.5 \leq z \leq 2.1$ into eight, evenly spaced bins and constraints are derived on f_i in each bin. Unfortunately, the redshift range and the large bin size makes the parameterization ineffective to constrain the general bimetric models.

Model:	Best-fit general(*)	Best-fit self-acc(*)	Best-fit B_{123}	ΛCDM (GR)
$\gamma_0 =$	0.46	0.51	0.55	0.55
$\gamma_1 =$	0.30	0.12	-0.04	0

Table 6.2: Growth index with the parameterization (6.29). For models that can be fitted by eq. (6.29), Euclid will be able to distinguish a modified gravity model from ΛCDM if γ_0 or γ_1 lie outside the range $\gamma_0 = 0.55 \pm 0.20$ or $\gamma_1 = \pm 0.4$. (*)For these models, a growth index of the form (6.29) cannot be used to fit the growth rate.

Here, we have shown that the nonlinear equations of motion for a homogeneous density perturbation can be solved analytically. This is an Oppenheimer–Snyder model. In general relativity, the OS model can be generalized to density profiles with a radial dependence $\rho_m(t, r)$ and the equations can still be solved analytically [173–175]. This is called a Lemaître–Tolman–Bondi (LTB) model. Unfortunately, such an analytical solution is not achievable in BR, as shown in Paper II, and the equations must be solved numerically. This subject is treated in Chapter 7.

6.4 Summary

Standard cosmological perturbation theory was constructed with general relativity in mind and hence we should not be surprised by the fact that it breaks down when applied to extended theories for gravity [159]. In particular, at the level of linear perturbations, bimetric cosmology exhibits an exponentially growing mode at early enough times. Hence, linear perturbation theory breaks down and nonlinear terms become important, which is not to say that there is a physical instability. In other words, the inapplicability of linear, cosmological perturbation theory is a practical problem rather than a theoretical one. To date, finding a formulation that enables calculation of the evolution of small perturbations on cosmological backgrounds is one of the most important problems to solve.

In this chapter, we reviewed three possible ways forward. The first option is to exorcise the instability to the inflationary epoch (where it disappears) by letting the graviton mass $m_{\rm FP}$ be large enough. Generically, $m_{\rm FP} \to \infty$ does not enforce the GR limit of the theory, which is when the mixing angle vanishes, $\theta \to 0$.

The second option to treat the growth of cosmological perturbations is to invoke the Vainshtein mechanism, restoring GR results nonlinearly for pertur-

bations with a density above a certain threshold.

The final option that we analyzed is to solve the nonlinear equations of motion for the cosmological perturbations. In the case of a homogeneous density perturbation, we showed that this can be done analytically which for example enables calculation of the density contrast and growth rate. Using this model, the calculated growth rate of the three best-fit bimetric background cosmologies cannot be distinguished from a ΛCDM (GR) model using the forecasted data from the future Euclid satellite. However, there is a caveat for two of these models due to a sharp, temporary peak in the growth rate around $z \simeq 0.03$, see Fig. 6.4. For a general (inhomogeneous) perturbation, the nonlinear equations of motion must be solved numerically. This topic is treated in Chapter 7, although in a slightly different context.

Chapter 7

Gravitational collapse

In general relativity, gravity is attractive, meaning that two nearby objects attract each other. If there are no other forces at play, a cloud of particles, initially at rest in a static background, collapses under its own weight, possibly resulting in an infinitely dense mass point, a singularity. At this point, spacetime itself breaks down. Gravitational collapse can take place towards the end of a star's lifetime when it runs out of fuel and cannot maintain the pressure gradient which otherwise prevents it from collapse. Arguably, the most extreme end state of this process is a black hole which is a region from which no information can escape, not even light. The 2020 Nobel Prize in physics was (partly) awarded for the observations of a supercompact object at the center of our own galaxy, consistent with the presence of a $\sim 10^6 M_\odot$ supermassive black hole. Moreover, in 2019 the Event Horizon Telescope Collaboration published a direct image consistent with the presence of a supermassive black hole at the center of the galaxy M87. In addition, many of the gravitational wave events observed by the LIGO/Virgo collaboration are consistent with a pair of merging black holes (see Chapter 5). In other words, black holes are not only theoretical constructions but are likely to exist, even in our own galactic neighborhood.

The statement that gravity is attractive assumes that the strong energy condition (see eq. (7.2)) is satisfied, which is a requirement on the allowed matter content in our Universe. In a modified theory of gravity, this condition can be violated. Indeed, as we saw in Chapter 3, it is violated by the bimetric stress–energy. Thus, *a priori* the fate of a collapsing object is completely open. The questions that we seek to answer are: Does gravitational collapse lead to the formation of a black hole? Can a singular end state be avoided? As we will discover, there are several practical obstacles that must be overcome before we can answer these questions conclusively.

7.1 A (non-)singularity theorem

To demonstrate the effect of the bimetric stress–energy on gravitational collapse, consider a region which is filled with massive, non-interacting particles, hence moving along geodesics of our space-time. Consider a volume element flowing along the geodesics. The volume expansion rate along the flow is referred to as ϑ, so $\vartheta < 0$ means that the volume decreases along the curves. If the geodesics are non-intersecting, a decreasing comoving volume implies an increasing matter density and in the limit $\vartheta \to -\infty$, the matter density diverges and we have a singularity. The Raychaudhuri equation describes the evolution of ϑ and reads,

$$\frac{d\vartheta}{d\tau} = -\frac{1}{3}\vartheta^2 - \sigma_{\mu\nu}\sigma^{\mu\nu} + \omega_{\mu\nu}\omega^{\mu\nu} - R_{\mu\nu}U^\mu U^\nu, \tag{7.1}$$

with τ being the proper time along the geodesics, U^μ the tangent vector of the geodesics, $\sigma_{\mu\nu}$ the shear, and $\omega_{\mu\nu}$ the rotation. See Ref. [176] for more details. Negative terms on the right-hand side of eq. (7.1) give a decreasing contribution to the expansion rate, and vice versa for positive terms. It can be shown that the shear term $\sigma_{\mu\nu}\sigma^{\mu\nu}$ must be positive. Moreover, let us assume that there is no rotation, so $\omega_{\mu\nu} = 0$. In this case, all terms on the right-hand side of eq. (7.1) are negative except possibly the last one. The requirement that also the last term is negative reads,

$$\left(\mathcal{T}_{\mu\nu} - \frac{1}{2}\mathcal{T}^\rho{}_\rho g_{\mu\nu}\right) U^\mu U^\nu > 0, \tag{7.2}$$

and is referred to as the strong energy condition. If this is satisfied, the volume expansion rate is always decreasing, $d\vartheta/d\tau < 0$. Starting from rest $\vartheta|_{t=0} = 0$, the volume decreases and the matter density increases continuously. In other words, if eq. (7.2) is satisfied gravity is attractive and the massive, non-interacting particles converge, leading to gravitational collapse.

In GR, the matter stress–energy reads $\mathcal{T}_{\mu\nu} = T_{\mu\nu} = \mathrm{diag}(\rho_m, 0, 0, 0)$ and thus,

$$\left(T_{\mu\nu} - \frac{1}{2}T^\rho{}_\rho g_{\mu\nu}\right) U^\mu U^\nu = \kappa_g \rho_m/2 > 0, \tag{7.3}$$

and the strong energy condition is satisfied and thus $d\vartheta/d\tau < 0$. We now engage the bimetric stress–energy $V_{\mu\nu}$ so that $\mathcal{T}_{\mu\nu} = T_{\mu\nu} + V_{\mu\nu}$. As an example, consider a spherically symmetric, bidiagonal space-time,

$$ds_g^2 = -dt^2 + A^2 dr^2 + B^2 d\Omega^2, \tag{7.4a}$$

$$ds_f^2 = c^2\left[-\tilde{\alpha}^2 dt^2 + \tilde{A}^2 dr^2 + \tilde{B}^2 d\Omega^2\right], \tag{7.4b}$$

where c^2 is the asymptotic conformal factor between the two metrics and we have suppressed the (t, r) dependence in $(A, B, \widetilde{\alpha}, \widetilde{A}, \widetilde{B})$. Then,

$$\left(V_{\mu\nu} - \frac{1}{2}V^{\rho}{}_{\rho}g_{\mu\nu}\right)U^{\mu}U^{\nu} = -H_0^2\left[B_0 + \frac{B_1}{2}\left(y_A + 2y_B + 3y_\alpha\right)\right.$$
$$\left. + B_2\left(y_A y_\alpha + 2y_B y_\alpha\right) + \frac{B_3}{2}\left(y_B^2 y_\alpha - y_A y_B^2 - 2y_A y_B y_\alpha\right)\right]. \tag{7.5}$$

Here, we have defined,

$$y_A \equiv \frac{\widetilde{A}}{A}, \quad y_B \equiv \frac{\widetilde{B}}{B}, \quad y_\alpha \equiv \widetilde{\alpha}. \tag{7.6}$$

The sign of the B_1-term in eq. (7.5) is determined by B_1 and since a consistent cosmology requires $B_1 > 0$ (3.36), this term necessarily gives a repulsive contribution which slows down or possibly avoids the gravitational collapse. On the other hand, the B_2-term gives an attractive contribution if $B_2 < 0$ (which is required by consistent cosmology plus a working Vainshtein mechanism). The contribution from B_3 must be studied on a case by case basis. Due to the bimetric stress–energy, we cannot determine at this level whether gravitational collapse will take place or not. This is contrary to GR where gravitational collapse is inevitable under the conditions analyzed in this section.

7.2 A toy model

As a toy model, we consider the evolution of a homogeneous ball of dust in vacuum. This is the Oppenheimer–Snyder (OS) model [168]. In Chapter 6, we showed that the ball evolves as a homogeneous and isotropic solution (3.2) so we apply our knowledge about the cosmological background solutions from Chapter 3. In this case, the Raychaudhuri equation (7.1) is equivalent to the time derivative of the modified Friedmann equation (3.14),

$$\frac{\ddot{a}}{a} = -\frac{\kappa_g}{6}\left[\rho_m + (1 + 3w_{\rm DE})\rho_{\rm DE}\right], \tag{7.7}$$

where a is the scale factor determining physical size of the dust cloud (the ball). As we saw in Chapter 3, the bimetric stress–energy behaves as a phantom dark energy. In other words, the second term on the right-hand side of eq. (7.7) is negative. Thus, the bimetric stress–energy gives a repulsive contribution in this setup, slowing down the gravitational collapse.

As an example, consider a dust cloud of initial density $\rho_{m,0} = \rho_\odot \sim 10^3\,\mathrm{kg/m^3}$. This is very large compared with the present-day critical density $\rho_{m,0}/\rho_c \sim 10^{29}$. In the context of cosmological background solutions,

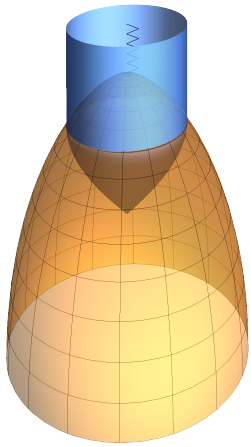


Figure 7.1: The Oppenheimer–Snyder model (one spatial dimensions is suppressed). The figure shows how the area radius of a collapsing dust cloud evolves with time. Time runs upwards. The cloud (orange) collapses under its own weight, eventually resulting in a curvature singularity (serrated line). Already before the singularity is formed, the event horizon (blue) starts growing from the center. This is the surface traced out by the (hypothetical) outgoing light rays that are on the boundary of reaching out to infinity and (eventually) being deflected back towards the center. Due to the strong gravitational field, no information which is sent out inside of the blue surface will escape out of it.

such a density corresponds to very early times, hence negligible dark energy in comparison with the matter density, unless m_{FP}, α, or β are very large. The difference here compared with cosmological solutions is that the dust cloud starts at rest initially ($\dot{a}|_{t=0} = 0$) which requires a positive spatial curvature. This, however, does not change our conclusion about the relation between $\rho_{m,0}$ and $\rho_{\mathrm{DE},0}$ since the equation for the ratio of the scale factors (3.18) is unchanged under the addition of spatial curvature and so is the dark energy ρ_{DE}. We conclude that although the bimetric stress–energy gives a repulsive gravitational contribution, the effect is negligible unless m_{FP}, α, or β are very large. See Fig. 7.1 for an example.

In Paper II, we showed that there is no static outside vacuum solution that can be continuously matched with the inside solution for a homogeneous dust cloud, thus Jebsen–Birkhoff's theorem is violated in BR.

Here, we have analyzed a special model consisting of a homogeneous region of pressureless dust. This enabled us to solve the equations of motion analytically. The natural extension is to allow for an arbitrary radial matter

profile. In GR, this leads to the Lemaître–Tolman–Bondi model which is also an exact solution. As we show in Paper II, unfortunately, in BR there is no exact solution and the equations of motion must be solved numerically which we do in Section 7.4.

7.3 Generalized Vaidya solutions

Besides the Lemaître–Tolman–Bondi solution (of which the Oppenheimer–Snyder model is a special case), the Vaidya space-time is a well-studied analytical solution of gravitational collapse in GR. The latter describes the spherically symmetric inflow (collapse) of null dust from radial infinity towards a center.

General relativity

In ingoing Eddington–Finkelstein coordinates, the Schwarzschild metric reads,

$$ds^2 = -\left(1 - \frac{2m}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2, \tag{7.8}$$

where m is the constant mass of the source. Letting m be a function of v promotes ds^2 to a Vaidya metric [177]. In this case, the Einstein tensor is nonzero and a nonzero stress–energy tensor must be introduced. We employ Synge's method and declare the stress–energy tensor to be whatever the Einstein tensor evaluates to, so that Einstein's equations are satisfied. Doing so, we have to check that the stress–energy arises from a reasonable matter model, for example by imposing energy conditions on $T^{\mu}{}_{\nu}$. This is rarely the case, with the Vaidya metric being an exception.

The stress–energy takes the form,

$$T_{\mu\nu} = 2\frac{\partial_v m(v)}{r^2} l_\mu l_\nu, \quad l_\mu \equiv -\partial_\mu v. \tag{7.9}$$

Here l_μ is an ingoing radial null covector. The weak, strong, dominant, and null energy conditions are satisfied if $\partial_v m(v) \geq 0$. The stress–energy (7.9) represents an inflow of massless particles towards the center $r = 0$. For example, starting with $m = 0$, the metric is flat Minkowski initially. When m increases there is an intermediate Vaidya region and when m reaches its final value, a Schwarzschild black hole is obtained, see Fig. 7.2.

The Vaidya metric can be generalized by including an r dependence of m [178, 179]. With,

$$m(v, r) = m_{\text{mass}}(v) - \frac{Q^2(v)}{2r} + \frac{\Lambda}{6} r^3, \tag{7.10}$$

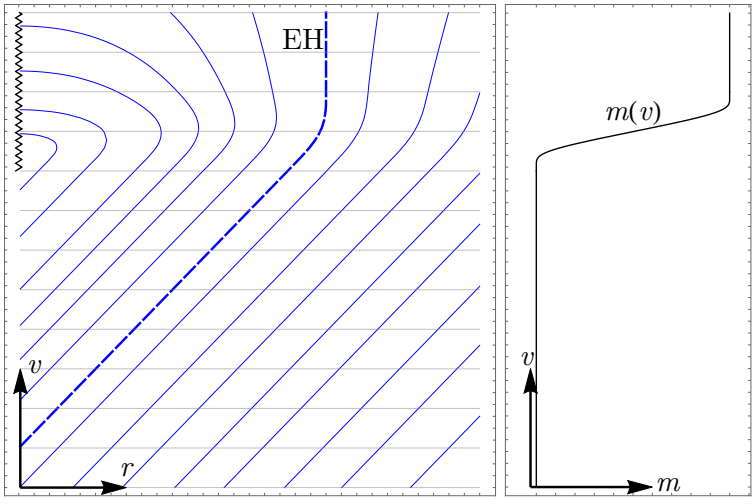


Figure 7.2: A Vaidya space-time. (a) Radial null geodesics (RNG). Outgoing RNG in blue and ingoing RNG in gray. The serrated line is the singularity and the event horizon (EH) is dashed. Inside the EH, all light rays end up in the singularity while the outgoing light rays outside the EH reach out to infinity. (b) The black hole mass as a function of time v. Initially, the mass is zero and we have a Minkowski space-time. Then m increases for some time until reaching the final constant black hole mass, in which case we have a Schwarzschild space-time. Panel (a) and (b) have the same vertical axis.

$m_{\text{mass}}(v)$ is the mass of the black hole, $Q(v)$ enters as an electric charge [180, 181], and Λ a cosmological constant. The stress–energy can be decomposed as,

$$T_{\mu\nu} = \mu l_\mu l_\nu + 2(\rho + p) l_{(\mu} n_{\nu)} + p g_{\mu\nu}, \tag{7.11}$$

with ingoing and outgoing null covectors,

$$l_\mu \equiv -\partial_\mu v, \quad n_\mu \equiv \frac{1}{2}\left(1 - \frac{2m}{r}\right)\partial_\mu v + \partial_\mu r, \tag{7.12}$$

respectively. Introducing an orthonormal frame with basis vectors $E^\mu{}_a$ [179],

$$E^\mu_{(0)} = \frac{l^\mu + n^\mu}{\sqrt{2}}, \quad E^\mu_{(1)} = \frac{l^\mu - n^\mu}{\sqrt{2}}, \quad E^\mu_{(2)} = \frac{1}{r}\delta^\mu{}_\theta, \quad E^\mu_{(3)} = \frac{1}{r\sin\theta}\delta^\mu{}_\phi, \tag{7.13}$$

the stress–energy can be projected onto this frame with the result,

$$T_{ab} = T_{\mu\nu} E^{\mu}{}_{a} E^{\nu}{}_{b} = \begin{pmatrix} \mu/2 + \rho & \mu/2 & 0 & 0 \\ \mu/2 & \mu/2 - \rho & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \tag{7.14}$$

Thus, the stress–energy is of Type II in the classification of Hawking and Ellis [182] and μ can be interpreted as the energy flux in the direction $E^{\mu}{}_{(0)}$, p is the tangential pressure, and ρ is related to the energy density (as measured by the observer associated with this orthonormal frame).

The dominant energy condition reads,

$$\mu \geq 0, \quad \rho \geq P \geq 0. \tag{7.15}$$

If it is respected, the weak and strong energy conditions are also satisfied.

Bimetric gravity

The ansatz is,

$$ds_g^2 = -\left(1 - \frac{2m(v,r)}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2, \tag{7.16a}$$

$$ds_f^2 = c^2 \left[-\left(1 - \frac{2\widetilde{m}(v,r)}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2 \right], \tag{7.16b}$$

with c constant. With eq. (7.16), the bimetric stress–energies read,

$$\kappa_g V^{\mu}{}_{\nu} = -\begin{pmatrix} \Lambda & 0 & 0 & 0 \\ \sin^2\theta\, \hat{m}_{\mathrm{FP}}^2 \dfrac{m - \widetilde{m}}{r} & \Lambda & 0 & 0 \\ 0 & 0 & \Lambda & 0 \\ 0 & 0 & 0 & \Lambda \end{pmatrix}, \tag{7.17a}$$

$$\kappa_f \widetilde{V}^{\mu}{}_{\nu} = -\frac{1}{c^2}\begin{pmatrix} \widetilde{\Lambda} & 0 & 0 & 0 \\ \cos^2\theta\, \hat{m}_{\mathrm{FP}}^2 \dfrac{\widetilde{m} - m}{r} & \widetilde{\Lambda} & 0 & 0 \\ 0 & 0 & \widetilde{\Lambda} & 0 \\ 0 & 0 & 0 & \widetilde{\Lambda} \end{pmatrix}, \tag{7.17b}$$

where $\hat{m}_{\mathrm{FP}}$ is the dimensionful Fierz–Pauli mass and the cosmological constants of the two sectors are defined according to,

$$\Lambda \equiv 3H_0^2\Omega_{\rm DE}|_{y=1} = H_0^2(B_0 + 3B_1 + 3B_2 + B_3), \tag{7.18a}$$

$$\widetilde{\Lambda} \equiv 3H_0^2\widetilde{\Omega}_{\rm DE}|_{y=1} = \frac{H_0^2}{\tan^2\theta}(B_1 + 3B_2 + 3B_3 + B_4). \tag{7.18b}$$

If the metrics approach a proportional de Sitter solution asymptotically, the cosmological constants coincide, $\Lambda = \widetilde{\Lambda}$. With eqs. (7.16) and (7.17), the bimetric conservation law (2.8) reads,

$$(B_1 + 2B_2 + B_3)\left[(m - \widetilde{m}) - r\partial_r(m - \widetilde{m})\right] = 0. \tag{7.19}$$

There are two branches of solutions, corresponding to setting each of the parentheses to zero.

Branch I. The first branch is obtained by setting,

$$B_1 + 2B_2 + B_3 = 0. \tag{7.20}$$

Naively, this corresponds to setting $m_{\rm FP} = 0$ (cf. eq. (2.13b)) but note that due to the vanishing Jacobian determinant the physical parameterization is not applicable in this special case and we have to stick with the B-parameters. In the rest of the thesis, we have always assumed $B_1+2B_2+B_3 > 0$, corresponding to a positive graviton mass. With eq. (7.20), the bimetric stress-energies (7.17) reduce to cosmological constant contributions so that,

$$G^\mu{}_\nu + \Lambda\delta^\mu{}_\nu = \kappa_g T^\mu{}_\nu, \tag{7.21a}$$

$$\widetilde{G}^\mu{}_\nu + \widetilde{\Lambda}\delta^\mu{}_\nu/c^2 = 0. \tag{7.21b}$$

Integrating the f-sector equations of motion results in,

$$\widetilde{m}(v, r) = \widetilde{m}_{\rm mass} + \frac{\widetilde{\Lambda}}{6}r^3, \tag{7.22}$$

where $\widetilde{m}_{\rm mass}$ is an integration constant. Thus, the second metric is Minkowski, dS, AdS, Schwarzschild, SdS, or SAdS depending on the signs of $\widetilde{m}_{\rm mass}$ and $\widetilde{\Lambda}$. If $\widetilde{m}_{\rm mass} \neq 0$, the Kretschmann scalar ($\widetilde{R}_{\mu\nu\rho\sigma}\widetilde{R}^{\mu\nu\rho\sigma}$) reveals that there is a curvature singularity at the center $r = 0$. This is covered by the event horizon of $f_{\mu\nu}$.

The equations in the g-sector are solved using Synge's method and we can set,

$$m(v, r) = m_{\rm mass}(v) - \frac{Q(v)}{2r} + \frac{\Lambda}{6}r^3. \tag{7.23}$$

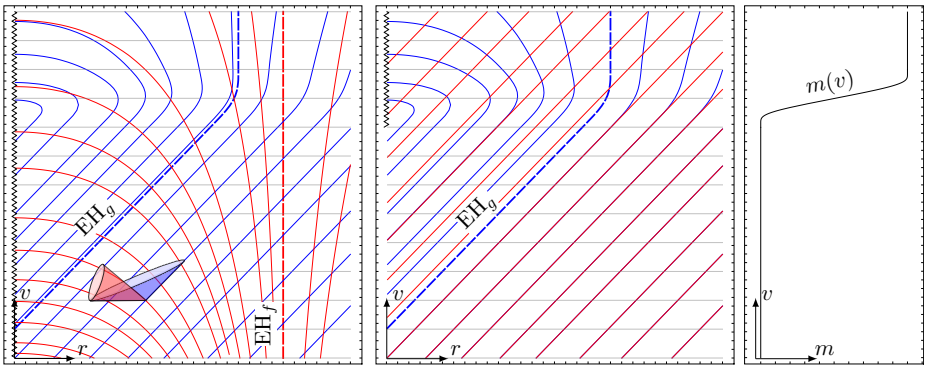


Figure 7.3: Bimetric Vaidya space-times, Branch I. (a)-(b) Outgoing RNG of $g_{\mu\nu}$ (blue) and of $f_{\mu\nu}$ (red). The ingoing RNG (gray) are shared between the metrics. The event horizons are dashed. The null cones of the metrics are drawn at a common point. Here $Q = 0$ and $\Lambda = \widetilde{\Lambda} = 0$. (a) $\widetilde{m}_{\text{mass}} = 1.2m_{\text{mass}}$. At $r = 0$ there is a curvature singularity with respect to $f_{\mu\nu}$, stretching from $v = -\infty$ to $v = +\infty$. This is only partly covered by the $g_{\mu\nu}$ event horizon. (b) $\widetilde{m}_{\text{mass}} = 0$. The singularity is covered by the $g_{\mu\nu}$ event horizon but there is no event horizon with respect to $f_{\mu\nu}$. (c) Mass $m_{\text{mass}}(v)$ as a function of advanced time v. The plots have the same vertical axis. Figure reprinted from Paper VI under the terms of Creative Commons Attribution 4.0 license.

Since the bimetric stress–energy gives a cosmological constant contribution, the matter stress–energy generated by eq. (7.23) has the same physical interpretation as in GR. The end state of gravitational collapse is reached when m_{mass} and Q reach constant values. If $Q = 0$, the end state corresponds to a static, spherically symmetric, non-bidiagonal black hole [128]. If $Q \neq 0$, the end state is a non-bidiagonal, charged black hole [183].

A notable property of the solution is that the curvature singularity of $f_{\mu\nu}$ at $r = 0$ is naked with respect to $g_{\mu\nu}$. The event horizon of $g_{\mu\nu}$ covers the singularity only partly as can be seen in Fig. 7.3(a). An observer approaching $r = 0$ at early times does not note anything special, but when reaching the center, the worldline ends suddenly. On the other hand, with $\widetilde{m}_{\text{mass}} = 0$, the singularity is covered with the event horizon of $g_{\mu\nu}$ but is naked with respect to $f_{\mu\nu}$, see Fig. 7.3(b). This raises the question of cosmic censorship. Are naked singularities ubiquitous in this theory or just a feature of this particular solution?

Branch II. Here, we solve eq. (7.19) by setting the second parenthesis to zero. The equations of motion in the f-sector can be integrated with the result,

$$m(v,r) = m_{\rm mass}(v) + \frac{\Lambda}{6}r^3 + \frac{2}{\cos^2\theta\,\hat{m}_{\rm FP}^2}\frac{\partial_v m_{\rm mass}(v)}{r}, \tag{7.24a}$$

$$\widetilde{m}(v,r) = m_{\rm mass}(v) + \frac{\Lambda}{6}r^3. \tag{7.24b}$$

The metrics approach a proportional de Sitter solution as $r \to \infty$, hence the cosmological constants of the two sectors can be identified, $\Lambda = \widetilde{\Lambda}$. If desired, a flat asympotic structure can be obtained by setting $\Lambda = 0$. The end state of collapse is reached when $m_{\rm mass} = \text{const}$. From eq. (7.24) we see that the end state is a proportional SdS black hole. In this regime, the event horizons of the two metrics coincide, as required by the theorem of Deffayet and Jacobson [184].

The stress–energy is declared to be,

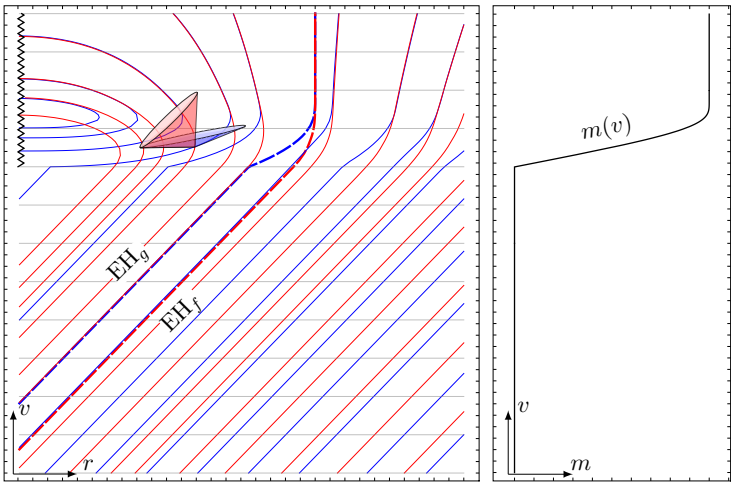
$$T^{\mu}{}_{\nu} = \kappa_g^{-1} G^{\mu}{}_{\nu} + V^{\mu}{}_{\nu}, \tag{7.25}$$

and the resulting $T^{\mu}{}_{\nu}$ takes a slightly complicated form involving first and second derivatives of $m_{\rm mass}(v)$ as well as an explicit r dependence, see Paper VI for details. The stress–energy is of Type II in the classification of Hawking and Ellis [182] and has an equation of state parameter $w = p/\rho = 1$, describing an ultrastiff fluid [185]. In this sense, these solutions require a very particular fluid.

It is not obvious how to assess whether this stress–energy can be regarded as physical. In bimetric theory there are several stress–energies ($T^{\mu}{}_{\nu}$, $V^{\mu}{}_{\nu}$, $\mathcal{T}^{\mu}{}_{\nu}$, etc.) and it is not clear what energy conditions should be applied to each one of these. There is only a small amount of literature on the subject [186, 187]. Imposing the weak, strong, and dominant energy conditions on the matter stress–energy (7.15), the result is,

$$B_1 + 2B_2 + B_3 < 0, \quad \partial_v m_{\rm mass}(v) \geq 0, \quad \partial_v^2 m_{\rm mass}(v) \leq 0, \tag{7.26}$$

for an accreting black hole (i.e., $\partial_v m_{\rm mass}(v) \geq 0$). This implies that $\partial_v m_{\rm mass}(v)$ is discontinuous at the point where it starts to increase from zero. For an example, see Fig. 7.4. For the energy conditions to apply, we need to require that the massive spin-2 field is tachyonic around proportional background solutions ($m_{\rm FP}^2 < 0$). Interestingly, a similar condition is found empirically in Section 7.4 for the numerical solutions of gravitational collapse. Since the Branch II metrics are proportional when they have reached their end states (that is when $m_{\rm mass} = \text{const.}$), this signals an instability for this solution branch [188].



(a) (b)

Figure 7.4: A bimetric Vaidya space-time, Branch II. (a) Outgoing RNG of $g_{\mu\nu}$ (blue) and $f_{\mu\nu}$ (red). The ingoing RNG (gray) are shared. The event horizons are dashed. Here, the EH of $g_{\mu\nu}$ lies inside that of $f_{\mu\nu}$ until they join as the mass reach the final (constant) value. The null cones of the metrics are drawn at a common point. At $(v \geq 0, r = 0)$ there is a curvature singularity. (b) Mass $m_{\text{mass}}(v)$ as a function of advanced time v. Initially when $v < 0$, $g_{\mu\nu}$ and $f_{\mu\nu}$ are flat Minkowski metrics. Then m_{mass} starts increasing until reaching the constant mass end state when $g_{\mu\nu}$ and $f_{\mu\nu}$ are Schwarzschild metrics. Panel (a) and (b) have the same vertical axis. Figure reprinted from Paper VI under the terms of Creative Commons Attribution 4.0 license.

7.4 Numerical solutions

To solve the equations of motion numerically, we make a 3+1 decomposition of our space-time into space+time. Here, we give a rough picture of the process. For details, see for example Ref. [32, 162, 167]. We start by defining a time function, t, whose level surfaces are three-dimensional spacelike hypersurfaces that foliate our space-time, see Fig. 7.5. The induced metric γ_{ij} is the geometry measured by an observer living on one of the time slices, Σ_t say. It is defined by,

$$\gamma_{ij} \equiv \frac{\partial x^\mu}{\partial y^i} \frac{\partial x^\nu}{\partial y^j} g_{\mu\nu}, \tag{7.27}$$

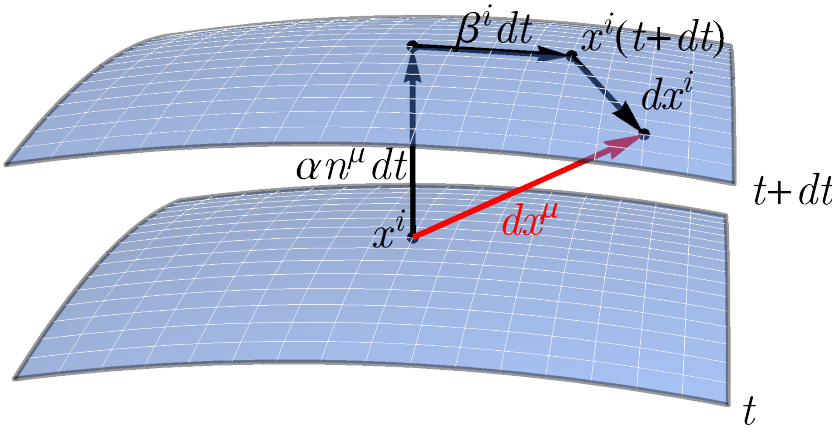


Figure 7.5: Schematic representation of the 3+1 decomposition. x^i and $x^i(t+dt)$ denote two points with the same spatial coordinate at two instants of time, t and $t+dt$ respectively. The coordinate increment dx^μ can be written $dx^\mu = \alpha n^\mu dt + (\beta^i dt + dx^i)\partial x^\mu/\partial x^i$.

with y^i being the coordinates imposed on Σ_t. Taking a small step along the normal direction n^μ from the point x^μ on the t-slice to the $(t+dt)$-slice, we end up at the coordinate $x^\mu + \alpha n^\mu dt$. Here, we have introduced the lapse α. Generically, the spatial coordinate x^i shifts when taking such a step. To return to the same spatial coordinate $x^i(t+dt)$ we have to take an additional step $\beta^i dt$ where β^i is referred to as the shift. The proper distance between two points separated by a coordinate distance dx^μ is,

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu. \tag{7.28}$$

On the other hand, from Fig. 7.5 we see that we can also write this as,

$$\begin{aligned} ds^2 &= g_{\mu\nu}(\alpha n^\mu dt)(\alpha n^\nu dt) + \gamma_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt) = \\ &= -\alpha^2 dt^2 + \gamma_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt), \end{aligned} \tag{7.29}$$

where we have used γ_{ij} to calculate the proper distance between points on the hypersurface and used the fact that $n^\mu n_\mu = -1$. This is the 3+1 decomposition of the metric into lapse α, shift β^i, and the induced metric γ_{ij}. From this construction, one can see that the lapse and the shift are related to our choice of coordinates. Setting the shift to zero, the lapse encodes the proper time between points of the same spatial coordinates at the different time slices,

$$d\tau = \alpha(t, x^i)dt|_{\beta^i=0}. \tag{7.30}$$

Besides the induced metric describing the intrinsic geometry on each time slice, the extrinsic curvature K_{ij} encodes the way in which the hypersurfaces are embedded in the four-dimensional manifold, and is defined as,

$$K_{ij} \equiv \frac{\partial x^\mu}{\partial y^i}\frac{\partial x^\nu}{\partial y^j}\nabla_\nu n_\mu. \tag{7.31}$$

In fact, K_{ij} is related to the first time derivative of γ_{ij}.

Writing the Einstein equations in these variables, the lapse and shift enter as nondynamical variables and are determined by our coordinate choice. The equations split into 4 constraint equations (one Hamiltonian constraint and three momentum constraints) and 6+6 evolution equations for γ_{ij} and K_{ij}. The constraint equations do not involve any time derivatives and do not contain the lapse or the shift. The evolution equations are first order in the time derivatives of γ_{ij} and K_{ij}. In the free evolution scheme, we solve the constraint equations on the initial hypersurface only and evolve the fields using the evolution equations. The constraints are preserved during the evolution [189].

In BR, the first thing to note is that although we have two metrics, we still have the freedom to choose only one set of coordinates. This is a source of additional complications in this theory. First, we must ensure that there is a common timelike direction and a common spacelike hypersurface with respect to the two metrics. This is indeed the case [81]. There are two induced metrics γ_{ij} and $\widetilde{\gamma}_{ij}$ and two extrinsic curvatures K_{ij} and $\widetilde{K}_{ij}$, one for each metric. The number of evolution equations is doubled as well as the number of Hamiltonian and momentum constraints. There are two lapses and two shifts, α, $\widetilde{\alpha}$, β^i, and $\widetilde{\beta}^i$. As in GR, these are nondynamical and the coordinate freedom is sufficient to fix one set of lapse and shift. The projection of the bimetric conservation law (2.8) results in an additional constraint and its preservation in time yields an additional constraint in form of a relation between the lapses of the two metrics. These remove the Boulware–Deser ghost and thus have no correspondence in GR. For details see Refs. [8, 162–164, 167].

In spherical symmetry, the metrics can be decomposed as,

$$ds_g^2 = -\alpha^2 dt^2 + A^2\left(dr + \beta dt\right)^2 + B^2 d\Omega^2, \tag{7.32a}$$

$$ds_f^2 = -\widetilde{\alpha}^2 dt^2 + \widetilde{A}^2(dr + \widetilde{\beta}dt)^2 + \widetilde{B}^2 d\Omega^2, \tag{7.32b}$$

with the metric components depending on t and r. The components of the extrinsic curvatures are denoted,

$$K_1 \equiv K^r{}_r, \quad K_2 \equiv K^\theta{}_\theta = K^\phi{}_\phi, \quad \widetilde{K}_1 \equiv \widetilde{K}^r{}_r, \quad \widetilde{K}_2 \equiv \widetilde{K}^\theta{}_\theta = \widetilde{K}^\phi{}_\phi. \tag{7.33}$$

In this setup, we start with 8 dynamical variables, $(A, B, K_1, K_2, \widetilde{A}, \widetilde{B}, \widetilde{K}_1, \widetilde{K}_2)$ and 4 nondynamical variables, $(\alpha, \beta, \widetilde{\alpha}, \widetilde{\beta})$. There are 2 coordinate freedoms

(leaving the angular part untouched), reducing the number of true dynamical variables by two and determining one set of lapse and shift, for example α and β. Further, there are six constraint equations of which two determine the other set of lapse and shift, in this case $\widetilde{\alpha}$ and $\widetilde{\beta}$, and the other four impose constraints on the dynamical variables. In the end, we have $8-2-4=2$ true dynamical degrees of freedom corresponding to one propagating mode. In GR, the counting ends up with zero. This is yet another evidence of the violation of Jebsen–Birkhoff's theorem and the existence of a propagating scalar mode in massive gravity.

Initial data. Here, we measure the radius and time coordinates in units of the final black hole mass (up to some numerical factor of order unity). In Paper VII, instead of the two shifts β and $\widetilde{\beta}$, we use the mean shift q and relative shift p between the metrics, see Appendix C for the definitions. The first step is to solve the constraint equations on the initial hypersurface, see Fig. 7.6 for examples and Appendix C for the equations. For simplicity, we do this at a moment of time symmetry, that is, when $K_1 = K_2 = \widetilde{K}_1 = \widetilde{K}_2 = 0$. This implies that $p = 0$, hence the metrics are bidiagonal at $t = 0$. To simplify the constraints, we also assume that the metrics are conformally flat initially and that they approach $g_{\mu\nu} = c^{-2} f_{\mu\nu} = \eta_{\mu\nu}$ asymptotically as $r \to \infty$. In Paper VII, we use the rescaling invariance to set the conformal factor c to unity. Here, we restore the factor and report the results in terms of rescaling invariant parameters. For the matter content, we take a pressureless perfect fluid (dust) with a profile similar to a Gaussian distribution (see Paper VII). As seen in Fig. 7.7, 95 % of the total rest mass of the object is contained within a few Schwarzschild radii, similar to the case of a neutron star. In GR such a dust cloud would collapse to a black hole within less than $\sim 1\,\mathrm{ms}$.

The B-parameters used here range from $\sim 10^{-4}(r_H/r_S)^2$ to $\sim 10(r_H/r_S)^2$ where r_H is the Hubble radius and r_S is the Schwarzschild radius. Since the Hubble radius is much greater than the Schwarzschild radius even for a supermassive black hole, the B-parameters are much greater than unity. In fact, $B_n \gtrsim 10^{30}$ in our examples. For parameters closer to $B_n \sim 1$, which is typical for cosmological models with modifications to the late-time accelerated expansion, the bimetric contribution to the constraint equation scales down with a factor of at least 10^{30} compared with the examples considered here. Thus, we expect gravitational collapse to proceed according to GR with negligible corrections. This is in line with the conclusion of Section 7.2.

Remarkably, we found empirically that it is necessary to require,

$$B_1 + 2B_2 + B_3 < 0, \tag{7.34}$$

to have regular initial data in this setup. If we set the expression in eq. (7.34)

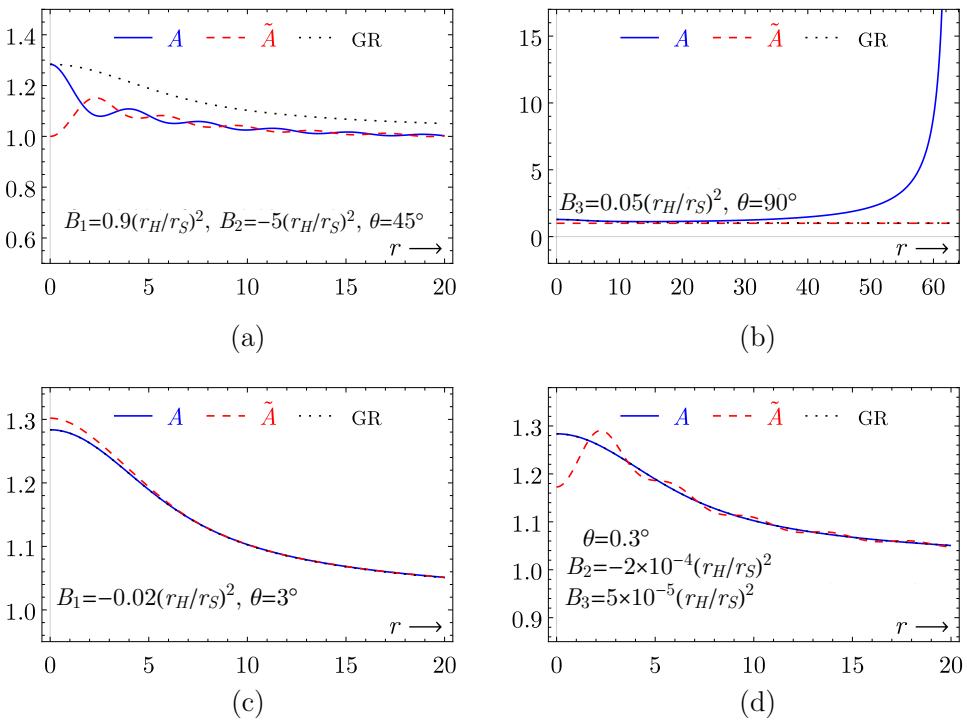


Figure 7.6: Initial data for A and $\tilde{A}$. The oscillations in r are generic for the initial data with $m_{\rm FP}^2 < 0$ and increase in frequency with increasing $|m_{\rm FP}|$. In the GR limit $\theta \to 0$, GR is approached as expected, see (c)-(d) which exhibit small values of the mixing angle. In panel (b), we show initial data with $m_{\rm FP}^2 > 0$. In this case, A diverges at some radius. In each panel, the nonzero B-parameters are stated. Reprinted from Paper VII under the terms of Creative Commons Attribution 4.0 license.

to be positive, the metric fields diverge at a finite radius. See Fig. 7.6(b) for an example. The condition (7.34) is equivalent to demanding that the Fierz–Pauli mass is tachyonic (i.e., imaginary), see eq. (2.13b). However, this does not imply that the theory itself is pathological but signals an instability of the background solutions around which the combination $B_1 + 2B_3 + B_3$ can be identified with the sign of the squared graviton mass [188] (in this case, the proportional solutions $f_{\mu\nu} = c^2 g_{\mu\nu}$). Discarding the proportional solutions forces us to look for other types of flat solutions. Although such solutions exist (see Paper I), they are basically unexplored and should be investigated further.

The nonlinear initial data equations are solved numerically. Nevertheless,

far away from the source, the solution approaches flat Minkowski at $t = 0$. In this weak-field regime, we can study the solution by analyzing the linear perturbations around this background (assuming that linear perturbation theory is valid). The perturbative ansatz is (7.32) with $\alpha = 1+\Phi$, $B/r = A = 1-\Psi$, $\widetilde{\alpha} = 1+\widetilde{\Phi}$, $\widetilde{A} = 1-\widetilde{\Psi}$, $\widetilde{B} = r^2(1+\delta\widetilde{B})$, and $\beta = \widetilde{\beta} = 0$ and $(\Phi, \Psi, \widetilde{\Phi}, \widetilde{\Psi}, \delta\widetilde{B}) \ll 1$ by assumption. Due to the requirement of a tachyonic graviton mass (7.34), instead of the ordinary Yukawa decay (4.6), the potentials oscillate,

$$\Phi = -M\cos^2\theta\left(\frac{1}{r} + \frac{4\tan^2\theta}{3}\frac{\cos\left(\omega r/r_S\right)}{r}\right), \tag{7.35a}$$

$$\Psi = -M\cos^2\theta\left(\frac{1}{r} + \frac{2\tan^2\theta}{3}\frac{\cos\left(\omega r/r_S\right)}{r}\right), \tag{7.35b}$$

with,

$$\omega \equiv \sqrt{-\left(\frac{r_S}{r_H}\right)^2 \frac{B_1 + 2B_2 + B_3}{\sin^2\theta}} = \frac{r_S}{r_H}|m_{\rm FP}|. \tag{7.36}$$

Hence, the Newtonian $1/r$ fall-off of the gravitational potentials is modulated by a cosine function which explains the radial oscillations in the initial data. The oscillation frequency ω scales with the square root of the B-parameters. This behavior is observed also for the nonlinear solutions (see Paper VII), confirming the validity of the linear perturbations in this regime.

As discussed in Chapter 4, the deviation from an inverse square law of gravity should not exceed one part in 10^9 on solar-system scales. Hence, we may demand that the amplitude of the oscillations should comply to,

$$\tan^2\theta < 10^{-9}. \tag{7.37}$$

Naively, we could also demand that the oscillation period is much greater than the size of the system so that the potentials are rescaled with constant factors over the length scale which is probed (L). That is,

$$\frac{1}{2}\left(\frac{\omega L}{r_S}\right)^2 \lesssim 10^{-9}. \tag{7.38}$$

However, with a small ω, the vDVZ discontinuity reappears. In the small ω limit,

$$\frac{\Phi}{\varphi} = \frac{1 + \frac{4\tan^2\theta}{3}}{1+\tan^2\theta}, \tag{7.39}$$

so that light bending observations are violated, see Chapter 4.

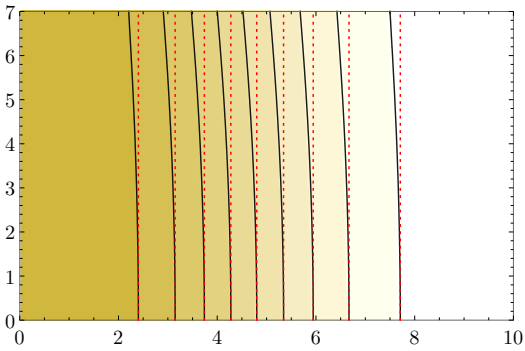


Figure 7.7: Lagrange shells of collapsing matter (solid curves) with initial data according to Fig. 7.6(c). The vertical, dotted lines are the initial positions of each shell. A Lagrange shell is the surface within which the total rest mass takes a particular value. Here, 95 % of the rest mass is contained within the outermost shell. Reprinted from Paper VII under the terms of Creative Commons Attribution 4.0 license.

Evolution. After solving the constraint equations, the next step is to evolve the fields in time. However, it has proved difficult to obtain a long term evolution. For the system to relax to a supposed final black hole, the simulation is expected to run until $t \sim 100$ whereas only the region $t \lesssim 10$ is covered by the simulations. One example can be seen in Fig. 7.7. At larger t, there are significant numerical errors. A main contribution to these come from round-off errors when calculating the ratio of the lapses and its radial derivatives. The former expression is derived from the preservation of the so-called secondary constraint and can be found in Ref. [164]. Nevertheless, by making many different runs with different parameters, we observed that as long as the numerical errors are small, the evolution follows GR closely if we start with initial data close to GR.

We can also note that bidiagonality is not preserved during time evolution. That is, even if we start with bidiagonal metrics so that $\beta|_{t=0} = \widetilde{\beta}|_{t=0}$, at the next step in time the shifts do not coincide $\beta|_{t=\Delta t} \neq \widetilde{\beta}|_{t=\Delta t}$ and hence the metrics are not bidiagonal. This can be seen by evaluating the constraint (C.8).

7.5 Summary

In bimetric gravity, due to the interaction between the two metrics, the standard energy conditions are broken. This implies that the standard singularity theorems do not apply and that there can be repulsive gravitational contributions, possibly preventing gravitational collapse from taking place. This can be shown using the Raychaudhuri equation. The ultimate question we want to answer is: What is the end state of a cloud of pressureless (i.e., dust) particles? Does it collapse to form a black hole? Does a singularity appear?

To answer these questions, we first studied a setup due to Oppenheimer and Snyder, modeling the evolution of a homogeneous, spherically symmetric dust cloud. The conclusion is that although the bimetric contribution is necessarily repulsive, preventing collapse, the effect is negligible unless the theory parameters $m_{\mathrm{FP}}, \alpha, \beta \gg 1$.

We also studied generalized Vaidya solutions in BR, describing the formation of a black hole by spherically symmetric accretion of massless dust particles. The solutions require a tachyonic or vanishing graviton mass (i.e., $m_{\mathrm{FP}}^2 \leq 0$). In the case of vanishing graviton mass, the bimetric stress–energy yields a cosmological constant contribution and the gravitational collapse proceeds as in GR. In the case of tachyonic mass, a very particular type of fluid is required that accretes in a very specific way, so we must be careful about drawing any general conclusions from this solution branch.

For more realistic setups, the equations of motion must be solved numerically (which is the case also in GR). This is done by splitting space-time into space+time, solving the constraint equations at an initial time slice and then evolving the fields using a set of evolution equations. Doing so, it is important to choose coordinates that allow for a numerically stable solution scheme. A fundamental complication in BR compared with GR is that there are two metrics but still only one set of coordinate freedom. We solved the constraint equations for the initial data in spherical symmetry. Unfortunately, when evolving the solution in time numerical errors accumulate quickly and invalidate the solution before reaching the end state of gravitational collapse. Additional work is needed to establish a framework for (numerically) stable long-term evolution. However, for initial data close to GR, the solution stays close to GR as long as the numerical error is small. Remarkably, to have regular initial data in this setup, it is required that $B_1 + 2B_2 + B_3 < 0$, corresponding to a tachyonic graviton mass around proportional background solutions where $f_{\mu\nu} = c^2 g_{\mu\nu}$. This signals an instability for these types of background solutions that are traditionally used to model for example flat space-times. Discarding these solutions forces us to study a previously unexplored regime, including flat space-times where the metrics are not proportional. Alternatively, it is

possible to relieve some of the assumptions that we make when solving the initial data equations for gravitational collapse. Possibly, that would allow for stable initial data with $B_1 + 2B_2 + B_3 > 0$. More work is needed in this area.

Chapter 8

Summary and outlook

Let us take a step back and reflect upon the results presented in this thesis. Bimetric gravity is a ghost-free theory of massive gravity, exhibiting both a massive and a massless spin-2 field. The theory has a continuous GR limit as well as a continuous limit to dRGT massive gravity, thus embracing these two theories in a unified framework. The theory has five, free, constant parameters: the mixing angle θ between the two spin-2 fields, the mass of the massive graviton $m_{\rm FP}$, the effective cosmological constant Ω_Λ, and two parameters, α and β, that are (among other things) related to the screening mechanism which restores GR results on solar-system scales. GR is retained in the limit $\theta \to 0$ and dRGT massive gravity is approached as $\theta \to \pi/2$. As the title of this thesis states, the main question is if Einstein's theory of general relativity is the final theory of gravity. In this context, it can be formulated as the question whether $\theta = 0$ or not. To answer this, we have confronted the theory with a wide range of observational tests and the result is summarized in Fig. 8.1.

Fig. 8.1 is a combined exclusion plot in the plane of the mixing angle and the graviton mass, compiling all constraints that have been discussed in the thesis.

- The constraints from local observations is due to solar-system tests and gravitational lensing on galactic scales. The excluded region in Fig. 8.1 is very conservative since it assumes that α and β are not large. If $\alpha \gg 1$ or $\beta \gg 1$, the constraint disappears. If the graviton mass is above $m_{\rm FP} \gtrsim 0.1\,\text{eV}$, the Compton wavelength of the massive graviton is smaller than $\lambda_g \lesssim 1\,\mu\text{m}$ which is below the smallest distance probed by gravitational experiments. In this case the Yukawa term is exponentially suppressed and GR is retained.

- Cosmological data gives an upper bound on the mixing angle $\theta \lesssim 30°$. Here, we used data from type Ia supernovae (SNIa), a constraint on the

present-day matter density due to measurements of the X-ray gas fraction in galaxy clusters (XCL), and data points from the cosmic microwave background (CMB) and baryon acoustic oscillations (BAO). Note that we used CMB and BAO data in a very conservative way, fitting only the ratio of the CMB and BAO distance scales. The reason being that there is no established framework for treating cosmological perturbations in bimetric gravity (see Chapter 6) and hence we use the data in a combination which eliminates the dependence on this as much as possible.

- For large values of the graviton mass ($m_{\rm FP} \gg 1$), the expansion history of the early universe is modified which affects the predicted abundance of the light elements as produced during Big Bang nucleosynthesis (BBN). To be compatible with the observed abundances, the mixing angle is constrained to be $\theta \lesssim 18^\circ$ in this region.

- Concerning gravitational waves (GWs), remember that there are two spin-2 fields in the theory, one massless and one massive. Hence, the theory can be compatible with GW observations either if the mixing angle is small enough so that the physical metric is aligned with the massless mode or by letting the graviton mass be small enough so that it does not affect the GW propagation over the relevant distances ($\sim$ Gpc).

- The Higuchi bound is necessary to avoid a pathological instability in our cosmological solutions. This yields a lower bound on the graviton mass.

- If we are interested in self-accelerating cosmological solutions we have to impose the constraint $\sin^2\theta \lesssim 2\Omega_\Lambda/m_{\rm FP}^2$ (3.68).

For self-accelerating solutions to contribute with significant modifications of the cosmological expansion history (i.e., $\theta \neq 0$), the graviton mass must be of the order of the Hubble scale, $m_{\rm FP} \sim 10^{-33}\,\text{eV}$. In this regime, the mixing angle can take values all the way up to $\theta \simeq 30^\circ$, representing a substantial deviation from the GR limit. For general bimetric models, there is also a viable region for $m_{\rm FP} \gtrsim 0.1\,\text{eV}$. Here, gravitational wave observations set an upper bound $\theta \lesssim 15^\circ$. There is also a viable region between $10^{-27}\,\text{eV} \lesssim m_{\rm FP} \lesssim 10^{-25}\,\text{eV}$ where the Vainshtein screening mechanism restores GR on solar-system scales but the linearized solution is applicable on galactic scales, restoring GR by virtue of suppressing the Yukawa term. Here, again $\theta \lesssim 30^\circ$.

As seen in Fig. 8.1, if we adopt all of these constraints, a large portion of the parameter space is excluded. On the other hand, the mixing angle can still reach values up to 30° when the graviton mass is of the order of the Hubble scale, which is precisely the region where there are modifications of the late-time expansion history of the Universe. We should also remember

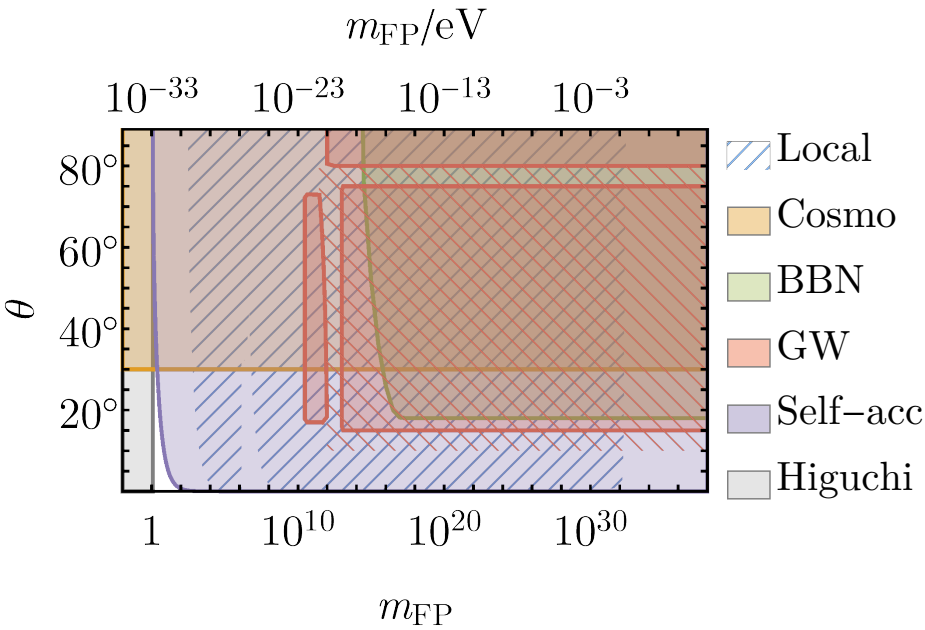


Figure 8.1: Exclusion plot in the $\theta m_{\rm FP}$-plane. The colored regions can be excluded. Blue hatched: Constraints from local tests of gravity, that is, due to solar-system tests and strong lensing in galaxies. The region is hatched since the constraint can be alleviated by choosing α or β to be large enough. Orange: cosmological tests due to CMB/BAO+SNIe+XCL. Green: constraint due to the observed abundance of the light elements as predicted by BBN. Red: gravitational wave observations. The hatched region is due to absence of observed GW echoes. However, it is not clear that such events would be detected by the current GW searches. Gray: demanding absence of the Higuchi ghost. Purple: excluded region by requiring existence of self-accelerating cosmological solutions. For the self-accelerating models, there is a viable region around $m_{\rm FP} \sim 10^{-33}\,\text{eV}$. For general bimetric models (to which the purple constraint does not apply), there are additional viable regions for $10^{-27}\,\text{eV} \lesssim m_{\rm FP} \lesssim 10^{-25}\,\text{eV}$ and $m_{\rm FP} \gtrsim 0.1\,\text{eV}$.

that although the allowed region appears narrow, there are three additional dimensions in the parameters space, not shown in the figure. So, there is in fact a rich spectrum of cosmological models in the viable region. An important question is if this rich spectrum of allowed cosmologies is a result of the conservative use of CMB/BAO data. To make use of the full information contained in CMB and BAO, we must be able to treat perturbations on cosmological backgrounds. In Chapter 6, we derived the full, nonlinear solution for a homogeneous and isotropic overdensity on a cosmological background. This gives us a hint of what can be expected in the general case but a complete framework that allows modeling of arbitrary perturbations is still lacking. If we want to make further progress in this theory, this is arguably one of the most pressing issues to solve.

In Fig. 8.1, we show only two of the five dimensions in the parameter space. The effective cosmological constant must take values $\Omega_\Lambda = 0.71^{+0.25}_{-0.03}$ to be compatible with cosmological observations. The parameters α and β are constrained by the absence of the Higuchi ghost (see Fig. 3.7). An additional constraint is added if we also want to require a working Vainshtein screening mechanism (see Fig. 4.1). These two constraints are complementary, basically excluding negative values of β and some negative values of α, see Fig. 4.2 for details.

Arguably, the greatest challenge to BR at present is computational: Given the breakdown of linear perturbation theory, how can we model the evolution of perturbations on cosmological backgrounds? What is a suitable framework for numerical bimetric relativity that would enable for example the calculation of gravitational collapse or the merger of a binary system? As we saw briefly in Chapter 7, even the spherically symmetric case presents a formidable challenge and a numerically stable evolution could only be obtained for a short amount of time. These are questions that must be answered in the future to fully understand the phenomenology of bimetric gravity.

Bimetric gravity passes the observational tests that we have implemented, without being pushed towards the GR limit and it even fits data slightly better. In short, gravity can be massive. On the other hand, GR (i.e., $\theta = 0$) was not excluded, which is of course expected since GR fits data very well with only a few exceptions (including the discrepant expansion rate of the Universe). So, given the increased complexity of the theory, what did we gain? Does it solve any of the problems that motivated us to look for new theories of gravity in the first place?

- The theory does not solve the fine-tuning problem related to vacuum energy and the accelerated expansion of the Universe. However, we gained a dynamical dark energy component which gives rise to an accelerated expansion of the Universe, without a cosmological constant.

- We obtained a much more flexible set of cosmological models, not captured by standard parameterizations. This might help in alleviating current and future observational anomalies. The prime example of a current anomaly is the discrepancy between the measured value of the Hubble constant using probes in the nearby versus the distant Universe. Two main tracks that have been pursued in the literature is to modify the expansion history in the late universe or in the early universe. Bimetric gravity can provide both types of modifications within a unified framework. There is some work in the literature on this topic, but only analyzing special bimetric submodels [70, 90, 190]. An exploration of the full parameter space with respect to the Hubble tension is still lacking.

- Bimetric gravity can have two important effects on the dark matter problem. Due to the Vainshtein screening mechanism, the gravitational force is effectively increased in strength towards the edge of a galaxy, thus reducing the need to dark matter. However, the effect is not enough to completely get rid of dark matter [9,72]. Another feature is that the massive graviton can constitute a dark matter particle candidate. However, to provide a viable dark matter particle, the theory is pushed towards a limit where it becomes indistinguishable from GR [52].

So, was Einstein wrong? The answer depends on your standards of right and wrong. The answer is apparently "no" in the sense that his theory of general relativity gives an incredibly precise description of gravitational phenomena on a very wide range of length scales. However, taken all together, there are indications that general relativity is not the definitive theory of gravity. In that sense, the answer is "yes". The question is then: What is a better theory of gravity that can explain some (or all) of the gravitational anomalies? We have shown that gravity can indeed be massive which is an interesting result in and of itself. Although the theory gives us a dark energy candidate, we did not manage to fully solve any of the grand puzzles in modern gravity. The big questions remain to be answered in the future.

Appendix A

List of analytical constraints

Cosmological constraints. A real-valued expansion rate in the late universe requires,

$$\Omega_\Lambda > 0. \tag{A.1}$$

For the matter density to be positive in the early universe, we must demand,

$$1 + 2\alpha + \beta > 0. \tag{A.2}$$

To exorcise the Higuchi ghost, the necessary and sufficient conditions are,

$$\frac{m_{\rm FP}^2}{2\Omega_\Lambda} > \frac{1}{2}\frac{3\sec^2\theta\, S_0}{\sqrt{\beta}(3\tan^2\theta\, S_0^2 - \beta) + S_0(1 - \alpha + \beta - 3\tan^2\theta[\alpha + \beta])}, \tag{A.3}$$

and,

$$\begin{cases} \alpha > -1/2 \\ \beta > -1 - 2\alpha \end{cases}, \tag{A.4a}$$

$$\begin{cases} \alpha < -1/2 \\ \beta > \frac{1}{12}\left(S_1 + 2\alpha\left[\frac{6}{1-3\tan^2\theta} - 6 + 3\alpha - S_1\right] + \frac{3S_2}{1-3\tan^2\theta} - 6\right) \end{cases}, \tag{A.4b}$$

with the definitions,

$$S_0 \equiv \sqrt{1 + 2\alpha + \beta}, \tag{A.5a}$$

$$S_1 \equiv \sqrt{9(1+\alpha)^2 - 12(1+2\alpha)\cos^2\theta}, \tag{A.5b}$$

$$S_2 \equiv \sqrt{3\sec^2\theta\,[-1 - 2\alpha + 3\alpha^2 + 3\tan^2\theta(1+\alpha)^2]}. \tag{A.5c}$$

Vainshtein constraints. Here, we list the necessary and sufficient conditions for a working Vainshtein mechanism to exist,[1]

$$\beta > 1, \tag{A.6a}$$

$$\alpha > -\sqrt{\beta}, \tag{A.6b}$$

$$(\alpha + \sqrt{\beta})\,(\alpha d_1 + d_2) > 0, \tag{A.6c}$$

where,

$$d_1 \equiv 1 + 3\tan^2\theta - 6\sqrt{\beta}\sec^2\theta + 3\beta\sec^2\theta, \tag{A.7a}$$

$$d_2 \equiv -1 + 6\sqrt{\beta}(1+\beta)\sec^2\theta - \beta(13 + 12\tan^2\theta). \tag{A.7b}$$

[1]There are two typos in Ref. [72]. In the sentence before their eq. (4.15), the constraint should read $\alpha > -\sqrt{\beta}$. In their eq. (4.15), the definitions of d_1 and d_2 should be interchanged.

Appendix B

Symmetries and topology

Topology. Topology provides a notion of continuity. Heuristically, it tells us what points that lie close to each other. The topology of a space can be diagnosed by considering closed curves and how they can (or cannot) be continuously deformed into each other. The two-dimensional torus provides a nice example, see Fig. B.1. For a rigorous treatment, see for example [191]. Common topologies are $\mathbb{R}^n$ and $\mathbb{S}^n$, with n being a positive integer. $\mathbb{R}^n$ is the topology of the n-dimensional plane and $\mathbb{S}^n$ is the topology of an n-sphere.

Lie derivative. Given a vector field ξ^μ, the Lie derivative of the tensor $T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots}$ along the vector field ($\pounds_\xi T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots}$) tells us how the tensor changes when dragged along the integral curves of ξ^μ. If,

$$\pounds_\xi T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots} = 0, \tag{B.1}$$

the tensor is invariant with respect to displacements along ξ^μ. Aligning one coordinate along ξ^μ, the components of the tensor are constant along the curves. In particular, if

$$\pounds_\xi g_{\mu\nu} = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0, \tag{B.2}$$

we say that ξ^μ is an isometry of the metric $g_{\mu\nu}$ and that ξ^μ is a Killing vector field (KVF).

Symmetry group. The orbits of the action of a Lie group can be identified with the integral curves of the KVFs, say $\xi^\mu_{(1)}$, $\xi^\mu_{(2)}$, $\xi^\mu_{(3)}$. If,

$$\pounds_{\xi_{(1)}} T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots} = \pounds_{\xi_{(2)}} T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots} = \pounds_{\xi_{(3)}} T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots} = 0, \tag{B.3}$$

it is a symmetry group of $T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots}$ and in the case where $T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\ldots}$ is a metric, it is called an isometry group.

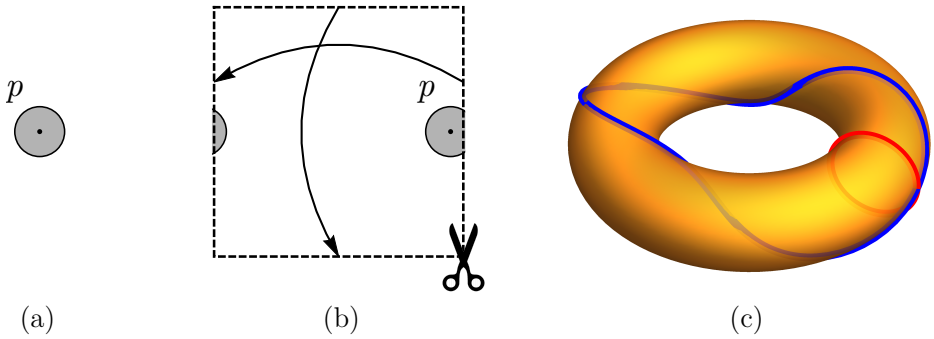


Figure B.1: (a) Topology of the plane, $\mathbb{R}^2$. The shaded points lie close to p. (b) Topology of a two-dimensional torus, $\mathbb{S}^1 \times \mathbb{S}^1$. Such a space can be constructed by cutting along the perforation and gluing together the sides as shown by the arrows. (c) A two-dimensional torus embedded in a three-dimensional Euclidean space. The red curve is wound around the torus once and the blue curve twice. They cannot be continuously deformed into each other.

Algebra. To every Lie group, there is an associated Lie algebra which is the tangent space at the identity element of the group manifold, together with a Lie bracket. The Killing vector fields satisfy the same algebra as the generators of the Lie algebra. As an example, take spherical symmetry. The group, SO(3), can be identified with the rotations of the 2-sphere. If the sphere is round, there are three independent KVFs. In spherical coordinates,

$$\xi_{(1)} = \partial_\phi, \quad \xi_{(2)} = \cos\phi\,\partial_\theta - \cot\theta\sin\phi\,\partial_\phi, \quad \xi_{(3)} = -\sin\phi\,\partial_\theta - \cot\theta\cos\phi\,\partial_\phi. \tag{B.4}$$

These satisfy the properties of the $\mathfrak{so}(3)$ Lie algebra,

$$\left[\xi_{(a)}, \xi_{(b)}\right] = \epsilon_{abc}\xi_{(c)}, \quad a, b, c \in \{1, 2, 3\}, \tag{B.5}$$

and ϵ_{abc} is the Levi-Civita symbol.

Appendix C

3+1 equations in spherical symmetry

Here, the 3+1 equations of bimetric gravity are presented as in Paper VII. The evolution equations take the same form as in GR with the addition of bimetric source terms. For the components of the induced metrics they are,

$$\partial_t A = -\alpha A K_1 + \partial_r(qA + \alpha v), \qquad \partial_t B = -\alpha B K_2 + \Big(q + \alpha A^{-1} v\Big)\partial_r B, \tag{C.1a}$$

$$\partial_t \widetilde{A} = -\widetilde{\alpha}\widetilde{A}\widetilde{K_1} + \partial_r(q\widetilde{A} - \widetilde{\alpha} v), \qquad \partial_t \widetilde{B} = -\widetilde{\alpha}\widetilde{B}\widetilde{K_2} + \Big(q - \widetilde{\alpha}\widetilde{A}^{-1} v\Big)\partial_r \widetilde{B}, \tag{C.1b}$$

where the shifts β and $\widetilde{\beta}$ are reparametrized by q and v (or p),

$$\beta = q + \alpha A^{-1} v, \quad \widetilde{\beta} = q + \widetilde{\alpha}\widetilde{A}^{-1} v, \quad v = p/\lambda, \quad \lambda = \sqrt{1 + p^2}, \tag{C.2}$$

see [162]. The evolution equations for the extrinsic curvatures are,

$$\begin{aligned}\partial_t K_1 = {} & \Big(q + \alpha A^{-1} v\Big)\partial_r K_1 + \alpha K_1\Big(K_1 + 2K_2\Big) - \alpha\kappa_g\Big\{ J_1 - \frac{1}{2}(J - \rho)\Big\} \\ & + \left(\frac{\partial_r \alpha}{A^2}\frac{\partial_r A}{A} - \frac{\partial_r^2 \alpha}{A^2} + 2\frac{\alpha}{A^2}\frac{\partial_r A}{A}\frac{\partial_r B}{B} - 2\frac{\alpha}{A^2}\frac{\partial_r^2 B}{B}\right),\end{aligned} \tag{C.3a}$$

$$\begin{aligned}\partial_t \widetilde{K_1} = {} & \Big(q - \widetilde{\alpha}\widetilde{A}^{-1} v\Big)\partial_r \widetilde{K_1} + \widetilde{\alpha}\widetilde{K_1}\Big(\widetilde{K_1} + 2\widetilde{K_2}\Big) - \widetilde{\alpha}\widetilde{\kappa}_g\Big\{ \widetilde{J}_1 - \frac{1}{2}(\widetilde{J} - \widetilde{\rho})\Big\} \\ & + \left(\frac{\partial_r \widetilde{\alpha}}{\widetilde{A}^2}\frac{\partial_r \widetilde{A}}{\widetilde{A}} - \frac{\partial_r^2 \widetilde{\alpha}}{\widetilde{A}^2} + 2\frac{\widetilde{\alpha}}{\widetilde{A}^2}\frac{\partial_r \widetilde{A}}{\widetilde{A}}\frac{\partial_r \widetilde{B}}{\widetilde{B}} - 2\frac{\widetilde{\alpha}}{\widetilde{A}^2}\frac{\partial_r^2 \widetilde{B}}{\widetilde{B}}\right),\end{aligned} \tag{C.3b}$$

$$\partial_t K_2 = \Big(q + \alpha A^{-1} v\Big)\partial_r K_2 + \alpha K_2\Big(K_1 + 2K_2\Big) - \alpha\kappa_g\Big\{ J_2 - \frac{1}{2}(J-\rho)\Big\} + \left(\frac{\alpha}{B^2} - \frac{\partial_r\alpha}{A^2}\frac{\partial_r B}{B} + \frac{\alpha}{A^2}\frac{\partial_r A}{A}\frac{\partial_r B}{B} - \frac{\alpha}{A^2}\frac{(\partial_r B)^2}{B^2} - \frac{\alpha}{A^2}\frac{\partial_r^2 B}{B}\right), \quad \text{(C.3c)}$$

$$\partial_t \widetilde{K}_2 = \Big(q - \widetilde{\alpha}\widetilde{A}^{-1} v\Big)\partial_r \widetilde{K}_2 + \widetilde{\alpha}\widetilde{K}_2\Big(\widetilde{K}_1 + 2\widetilde{K}_2\Big) - \widetilde{\alpha}\widetilde{\kappa}_g\Big\{ \widetilde{J}_2 - \frac{1}{2}(\widetilde{J}-\widetilde{\rho})\Big\} + \left(\frac{\widetilde{\alpha}}{\widetilde{B}^2} - \frac{\partial_r\widetilde{\alpha}}{\widetilde{A}^2}\frac{\partial_r \widetilde{B}}{\widetilde{B}} + \frac{\widetilde{\alpha}}{\widetilde{A}^2}\frac{\partial_r \widetilde{A}}{\widetilde{A}}\frac{\partial_r \widetilde{B}}{\widetilde{B}} - \frac{\widetilde{\alpha}}{\widetilde{A}^2}\frac{(\partial_r \widetilde{B})^2}{\widetilde{B}^2} - \frac{\widetilde{\alpha}}{\widetilde{A}^2}\frac{\partial_r^2 \widetilde{B}}{\widetilde{B}}\right). \quad \text{(C.3d)}$$

The source terms (denoted ρ and J) derive from the effective stress–energy and have contributions from both matter and bimetric stress–energy. For example, $\rho = \rho^{\rm m} + \rho^{\rm b}$, with "m" for matter and "b" for bimetric. The bimetric source functions are defined as,

$$\rho^{\rm b} = -\left[\langle y\rangle_0^2 + \lambda\frac{\widetilde{A}}{A}\langle y\rangle_1^2\right], \qquad j_r^{\rm b} = -p\widetilde{A}\langle y\rangle_1^2, \quad \text{(C.4a)}$$

$$J_1^{\rm b} = \langle y\rangle_0^2 + \left[\frac{1}{\lambda}\left(\frac{\widetilde{\alpha}}{\alpha} + \frac{\widetilde{A}}{A}\right) - \lambda\frac{\widetilde{A}}{A}\right]\langle y\rangle_1^2, \quad \text{(C.4b)}$$

$$J_2^{\rm b} = \langle y\rangle_0^1 + \frac{\widetilde{\alpha}\widetilde{A}}{\alpha A}\langle y\rangle_1^2 + \frac{1}{\lambda}\left(\frac{\widetilde{\alpha}}{\alpha} + \frac{\widetilde{A}}{A}\right)\langle y\rangle_1^1, \quad \text{(C.4c)}$$

$$\widetilde{\rho}^{\rm b} = -\left[\langle y\rangle_2^2 + \lambda\frac{A}{\widetilde{A}}\langle y\rangle_1^2\right]\frac{1}{y^2}, \qquad \widetilde{j}_r^{\rm b} = pA\langle y\rangle_1^2\frac{1}{y^2}, \quad \text{(C.4d)}$$

$$\widetilde{J}_1^{\rm b} = \left\{\langle y\rangle_2^2 + \left[\frac{1}{\lambda}\left(\frac{\alpha}{\widetilde{\alpha}} + \frac{A}{\widetilde{A}}\right) - \lambda\frac{A}{\widetilde{A}}\right]\langle y\rangle_1^2\right\}\frac{1}{y^2}, \quad \text{(C.4e)}$$

$$\widetilde{J}_2^{\rm b} = \left\{\langle y\rangle_3^1 + \frac{\alpha A}{\widetilde{\alpha}\widetilde{A}}\langle y\rangle_1^1 + \frac{1}{\lambda}\left(\frac{\alpha}{\widetilde{\alpha}} + \frac{A}{\widetilde{A}}\right)\langle y\rangle_2^1\right\}\frac{1}{y}, \quad \text{(C.4f)}$$

and,

$$J \equiv J_1 + 2J_2, \quad \widetilde{J} \equiv \widetilde{J}_1 + 2\widetilde{J}_2, \quad y \equiv \widetilde{B}/B. \quad \text{(C.5)}$$

The Hamiltonian and momentum constraints take the same form as in GR with the bimetric stress–energy added to the sources. The Hamiltonian constraints are,

$$(2K_1+K_2)K_2+\frac{1}{A^2}\left(\frac{A^2}{B^2}+2\frac{\partial_r A}{A}\frac{\partial_r B}{B}-\frac{(\partial_r B)^2}{B^2}-2\frac{\partial_r^2 B}{B}\right)-\kappa_g(\rho^{\mathrm{b}}+\rho^{\mathrm{m}})=0, \tag{C.6a}$$

$$(2\widetilde{K}_1+\widetilde{K}_2)\widetilde{K}_2+\frac{1}{\widetilde{A}^2}\left(\frac{\widetilde{A}^2}{\widetilde{B}^2}+2\frac{\partial_r \widetilde{A}}{\widetilde{A}}\frac{\partial_r \widetilde{B}}{\widetilde{B}}-\frac{(\partial_r \widetilde{B})^2}{\widetilde{B}^2}-2\frac{\partial_r^2 \widetilde{B}}{\widetilde{B}}\right)-\widetilde{\kappa}_g\widetilde{\rho}^{\mathrm{b}}=0, \tag{C.6b}$$

and the momentum constraints,

$$(K_1-K_2)\frac{\partial_r B}{B}-\partial_r K_2-\frac{1}{2}\kappa_g(j_r^{\mathrm{b}}+j_r^{\mathrm{m}})=0, \tag{C.7a}$$

$$(\widetilde{K}_1-\widetilde{K}_2)\frac{\partial_r \widetilde{B}}{\widetilde{B}}-\partial_r\widetilde{K}_2-\frac{1}{2}\widetilde{\kappa}_g\widetilde{j}_r^{\mathrm{b}}=0. \tag{C.7b}$$

The latter two equations can be combined in an algebraic equation for p,

$$p=-\frac{2}{\kappa_g\widetilde{A}B\langle y\rangle_1^2}\Big(K_1\partial_r B-K_2\partial_r B-B\partial_r K_2-Bj_r^{\mathrm{m}}\Big). \tag{C.8}$$

The bimetric constraint, originating from the bimetric conservation law, reads,

$$\begin{aligned}&\widetilde{A}\Big(\widetilde{K}_1\langle y\rangle_1^2+2\widetilde{K}_2 y\langle y\rangle_2^1\Big)+2A\widetilde{K}_2\lambda y\langle y\rangle_1^1\\&\quad-A\Big(K_1\langle y\rangle_1^2+2K_2\langle y\rangle_1^1\Big)-2\widetilde{A}K_2\lambda\langle y\rangle_2^1\\&\quad+2p\left(\langle y\rangle_1^1\frac{A}{\widetilde{A}}\frac{\partial_r\widetilde{B}}{B}+\langle y\rangle_2^1\frac{\widetilde{A}}{A}\frac{\partial_r B}{B}\right)+\lambda^{-1}\langle y\rangle_1^2\partial_r p=0.\end{aligned} \tag{C.9}$$

Finally, there is a constraint which gives the ratio of the lapses α and $\widetilde{\alpha}$ in terms of the other variables. Symbolically, we write,

$$\alpha/\widetilde{\alpha}=W, \tag{C.10}$$

where W is some algebraic expression of the dynamical fields and their spatial derivatives. The full expression can be found in Ref. [164].

Appendix D

Miscellaneous equations

For the background cosmology, the bimetric stress–energies can be written as perfect fluids, see eq. (3.5), where the energy densities and pressures are given by,

$$\kappa_g \rho_{\rm DE}/H_0^2 = 3\Omega_\Lambda - \sin^2\theta\, m_{\rm FP}^2 \left[3(1-y) + 3\alpha(1-y)^2 + \beta(1-y)^3\right], \tag{D.1a}$$

$$\kappa_g p_{\rm DE}/H_0^2 = -3\Omega_\Lambda + \sin^2\theta\, m_{\rm FP}^2 \left[3 + 3\alpha + \beta - 2(1+2\alpha+\beta)y + (\alpha+\beta)y^2\right] - x\sin^2\theta\, m_{\rm FP}^2 \left[1 + 2\alpha(1-y) + \beta(1-y)^2\right], \tag{D.1b}$$

$$\kappa_g \widetilde{\rho}_{\rm DE}/H_0^2 = \frac{1}{y^4 c^4}\left[\sin^2\theta\, m_{\rm FP}^2(1+2\alpha+\beta)y - 3\sin^2\theta\, m_{\rm FP}^2(\alpha+\beta)y^2 + 3\sin^2\theta\, m_{\rm FP}^2\beta y^3 + \left(3\tan^2\theta\,\Omega_\Lambda - \sin^2\theta\, m_{\rm FP}^2(1-\alpha+\beta)\right)y^4\right], \tag{D.1c}$$

$$\kappa_g \widetilde{p}_{\rm DE}/H_0^2 = -\frac{1}{xy^3c^4}\left[\sin^2\theta\, m_{\rm FP}^2 y\left(1 + 2\alpha(1-y) + \beta(1-y)^2\right) + x\left(3\tan^2\theta\,\Omega_\Lambda y^3 - \sin^2\theta\, m_{\rm FP}^2 y(\alpha+\beta-2\beta y + (1-\alpha+\beta)y^2)\right)\right]. \tag{D.1d}$$

The equation for the Stückelberg field $\mu(r)$ outside a source of mass M is,

$$\begin{aligned}
&3\sec^2\theta\,\mu + 6\sec^2\theta\,(1-\alpha)\mu^2 + \\
&\frac{1}{3}\left[6\sec^2\theta\,\alpha^2 - 2(17+18\tan^2\theta)\alpha + 4\sec^2\theta\,\beta + 10 + 9\tan^2\theta\right]\mu^3 + \\
&\frac{2}{3}\left[6\sec^2\theta\,\alpha^2 - (7+9\tan^2\theta)\alpha + 4\sec^2\theta\,\beta + 1\right]\mu^4 + \\
&\frac{1}{3}\left[2(1+3\tan^2\theta)\alpha^2 - \sec^2\theta\,\beta^2 + 2(1+2\tan^2\theta)\beta - 4\alpha\beta - 2\alpha\right]\mu^5 + \\
&-\frac{2}{3}\tan^2\theta\,\beta^2\mu^6 - \frac{1}{3}\tan^2\theta\,\beta^2\mu^7 \\
&= -\sec^2\theta\,(1+\mu)^2\left(\frac{r_V}{r}\right)^3(1-\beta\mu^2). \qquad \text{(D.2)}
\end{aligned}$$

References

[1] M. Kocic, M. Högås, F. Torsello and E. Mortsell, *Algebraic Properties of Einstein Solutions in Ghost-Free Bimetric Theory*, *J. Math. Phys.* **60** (2019) 102501, [1706.00787].

[2] M. Kocic, M. Högås, F. Torsello and E. Mortsell, *On Birkhoff's theorem in ghost-free bimetric theory*, 1708.07833.

[3] F. Torsello, M. Kocic, M. Högås and E. Mortsell, *Spacetime symmetries and topology in bimetric relativity*, *Phys. Rev. D* **97** (2018) 084022, [1710.06434].

[4] F. Torsello, M. Kocic, M. Högås and E. Mörtsell, *Covariant BSSN formulation in bimetric relativity*, *Class. Quant. Grav.* **37** (2020) 025013, [1904.07869]. [Erratum: Class.Quant.Grav. 37, 079501 (2020)].

[5] M. Kocic, F. Torsello, M. Högås and E. Mortsell, *Spherical dust collapse in bimetric relativity: Bimetric polytropes*, 1904.08617.

[6] M. Högås, M. Kocic, F. Torsello and E. Mörtsell, *Generalized Vaidya solutions in bimetric gravity*, *Class. Quant. Grav.* **37** (2020) 145010, [1905.09832].

[7] M. Högås, F. Torsello and E. Mörtsell, *On the stability of bimetric structure formation*, *JCAP* **04** (2020) 046, [1910.01651].

[8] M. Kocic, F. Torsello, M. Högås and E. Mörtsell, *Initial data and first evolutions of dust clouds in bimetric relativity*, *Class. Quant. Grav.* **37** (2020) 165010.

[9] M. Högås and E. Mörtsell, *Constraints on bimetric gravity. Part I. Analytical constraints*, *JCAP* **05** (2021) 001, [2101.08794].

[10] M. Högås and E. Mörtsell, *Constraints on bimetric gravity. Part II. Observational constraints*, *JCAP* **05** (2021) 002, [2101.08795].

[11] M. Högås and E. Mörtsell, *Constraints on bimetric gravity from Big Bang nucleosynthesis*, *JCAP* **11** (2021) 001, [2106.09030].

[12] S. Castello, M. Högås and E. Mörtsell, *A Cosmological Underdensity Does Not Solve the Hubble Tension*, 2110.04226.

[13] K. Max, M. Platscher and J. Smirnov, *Gravitational Wave Oscillations in Bigravity*, *Phys. Rev. Lett.* **119** (2017) 111101, [1703.07785].

[14] M. Högås, *Gravitational collapse in bimetric gravity*. Licentiate thesis (unpublished), Stockholm University, 2019.

[15] S. Schlamminger, K. Y. Choi, T. A. Wagner, J. H. Gundlach and E. G. Adelberger, *Test of the equivalence principle using a rotating torsion balance*, *Phys. Rev. Lett.* **100** (2008) 041101, [0712.0607].

[16] Supernova Search Team collaboration, A. G. Riess et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116** (1998) 1009–1038, [astro-ph/9805201].

[17] Supernova Cosmology Project collaboration, S. Perlmutter et al., *Measurements of* Ω *and* Λ *from 42 high redshift supernovae*, *Astrophys. J.* **517** (1999) 565–586, [astro-ph/9812133].

[18] L. Verde, T. Treu and A. G. Riess, *Tensions between the Early and the Late Universe*, *Nature Astron.* **3** (7, 2019) 891, [1907.10625].

[19] M. Fierz and W. Pauli, *On relativistic wave equations for particles of arbitrary spin in an electromagnetic field*, *Proc. Roy. Soc. Lond.* **A173** (1939) 211–232.

[20] C. de Rham and G. Gabadadze, *Generalization of the Fierz-Pauli Action*, *Phys. Rev.* **D82** (2010) 044020, [1007.0443].

[21] C. de Rham, G. Gabadadze and A. J. Tolley, *Resummation of Massive Gravity*, *Phys. Rev. Lett.* **106** (2011) 231101, [1011.1232].

[22] S. F. Hassan and R. A. Rosen, *Resolving the Ghost Problem in non-Linear Massive Gravity*, *Phys. Rev. Lett.* **108** (2012) 041101, [1106.3344].

[23] S. F. Hassan and R. A. Rosen, *Bimetric Gravity from Ghost-free Massive Gravity*, *JHEP* **02** (2012) 126, [1109.3515].

[24] S. Weinberg, *Gravitation and Cosmology*. John Wiley & Sons, Inc., 1972.

[25] E. Cartan, *Sur les variétés à connexion affine et la théorie de la relativité généralisée. (première partie)*, *Annales Sci. Ecole Norm. Sup.* **40** (1923) 325–412.

[26] E. Cartan, *Sur les variétés à connexion affine et la théorie de la relativité généralisée. (première partie) (Suite).*, *Annales Sci. Ecole Norm. Sup.* **41** (1924) 1–25.

[27] K. O. Friedrichs, *Eine Invariante Formulierung des Newtonschen Gravitationsgesetzes und der Grenzüberganges vom Einsteinschen zum Newtonschen Gesetz*, *Mathematische Annalen* **98** (1927) 566–575.

[28] A. Trautman, F. A. E. Pirani and H. Bondi, *Lectures on general relativity*. Brandeis Summer Institute in Theoretical Physics 1964, Vol. 1. Prentice-Hall, 1965.

[29] A. Einstein, "Fundamental ideas and methods of the theory of relativity, presented in their development." January, 1920 (Reprinted as Vol. 7, Doc. 31 CPAE).

[30] J. A. Wheeler and K. Ford, *Geons, black holes, and quantum foam: A life in physics*. 1998.

[31] D. Lovelock, *The Einstein tensor and its generalizations*, *J. Math. Phys.* **12** (1971) 498–501.

[32] T. W. Baumgarte and S. L. Shapiro, *Numerical Relativity: Solving Einstein's Equations on the Computer*. Cambridge University Press, 2010, 10.1017/CBO9781139193344.

[33] C. M. Will, *The Confrontation between General Relativity and Experiment*, *Living Rev. Rel.* **17** (2014) 4, [1403.7377].

[34] Event Horizon Telescope collaboration, K. Akiyama et al., *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, *Astrophys. J.* **875** (2019) L1, [1906.11238].

[35] LIGO Scientific, Virgo collaboration, B. P. Abbott et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016) 061102, [1602.03837].

[36] 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].

[37] 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].

[38] LIGO Scientific, Virgo, Fermi-GBM, INTEGRAL collaboration, B. P. Abbott et al., *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, *Astrophys. J. Lett.* **848** (2017) L13, [1710.05834].

[39] P. Bull et al., *Beyond ΛCDM: Problems, solutions, and the road ahead*, *Phys. Dark Univ.* **12** (2016) 56–99, [1512.05356].

[40] R. A. Alpher, H. Bethe and G. Gamow, *The origin of chemical elements*, *Phys. Rev.* **73** (1948) 803–804.

[41] R. A. Alpher and R. Herman, *Evolution of the Universe*, *Nature* **162** (1948) 774–775.

[42] R. A. Alpher and R. C. Herman, *On the Relative Abundance of the Elements*, *Phys. Rev.* **74** (1948) 1737–1742.

[43] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [1807.06209].

[44] Planck collaboration, P. A. R. Ade et al., *Planck 2015 results. XIV. Dark energy and modified gravity*, *Astron. Astrophys.* **594** (2016) A14, [1502.01590].

[45] A. G. Riess et al., *A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km/s/Mpc Uncertainty from the Hubble Space Telescope and the SH0ES Team*, 2112.04510.

[46] H. Desmond, B. Jain and J. Sakstein, *Local resolution of the Hubble tension: The impact of screened fifth forces on the cosmic distance ladder*, *Phys. Rev. D* **100** (2019) 043537, [1907.03778]. [Erratum: Phys.Rev.D 101, 069904 (2020), Erratum: Phys.Rev.D 101, 129901 (2020)].

[47] H. Desmond and J. Sakstein, *Screened fifth forces lower the TRGB-calibrated Hubble constant too*, *Phys. Rev. D* **102** (2020) 023007, [2003.12876].

[48] E. Mortsell, A. Goobar, J. Johansson and S. Dhawan, *The Hubble Tension Bites the Dust: Sensitivity of the Hubble Constant Determination to Cepheid Color Calibration*, 2105.11461.

[49] E. Mortsell, A. Goobar, J. Johansson and S. Dhawan, *The Hubble Tension Revisited: Additional Local Distance Ladder Uncertainties*, 2106.09400.

[50] K. Aoki, K.-i. Maeda and R. Namba, *Stability of the Early Universe in Bigravity Theory*, *Phys. Rev.* **D92** (2015) 044054, [1506.04543].

[51] E. Babichev, L. Marzola, M. Raidal, A. Schmidt-May, F. Urban, H. Veermäe et al., *Bigravitational origin of dark matter*, *Phys. Rev. D* **94** (2016) 084055, [1604.08564].

[52] E. Babichev, L. Marzola, M. Raidal, A. Schmidt-May, F. Urban, H. Veermäe et al., *Heavy spin-2 Dark Matter*, *JCAP* **09** (2016) 016, [1607.03497].

[53] T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, *Modified Gravity and Cosmology*, *Phys. Rept.* **513** (2012) 1–189, [1106.2476].

[54] J. B. Jiménez, L. Heisenberg and T. S. Koivisto, *The Geometrical Trinity of Gravity*, *Universe* **5** (2019) 173, [1903.06830].

[55] M. Milgrom, *A Modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis*, *Astrophys. J.* **270** (1983) 365–370.

[56] D. G. Boulware and S. Deser, *Can gravitation have a finite range?*, *Phys. Rev.* **D6** (1972) 3368–3382.

[57] H. van Dam and M. Veltman, *Massive and mass-less yang-mills and gravitational fields*, *Nuclear Physics B* **22** (1970) 397 – 411.

[58] V. I. Zakharov, *Linearized gravitation theory and the graviton mass*, *JETP Lett.* **12** (1970) 312. [Pisma Zh. Eksp. Teor. Fiz.12,447(1970)].

[59] A. I. Vainshtein, *To the problem of nonvanishing gravitation mass*, *Phys. Lett. B* **39** (1972) 393–394.

[60] S. F. Hassan and R. A. Rosen, *Confirmation of the Secondary Constraint and Absence of Ghost in Massive Gravity and Bimetric Gravity*, *JHEP* **04** (2012) 123, [1111.2070].

[61] S. F. Hassan, R. A. Rosen and A. Schmidt-May, *Ghost-free Massive Gravity with a General Reference Metric*, *JHEP* **02** (2012) 026, [1109.3230].

[62] C. de Rham, *Massive Gravity*, *Living Rev. Rel.* **17** (2014) 7, [1401.4173].

[63] A. Schmidt-May and M. von Strauss, *Recent developments in bimetric theory*, *J. Phys. A* **49** (2016) 183001, [1512.00021].

[64] S. F. Hassan and A. Lundkvist, *Analysis of constraints and their algebra in bimetric theory*, *JHEP* **08** (2018) 182, [1802.07267].

[65] M. von Strauss, A. Schmidt-May, J. Enander, E. Mortsell and S. F. Hassan, *Cosmological Solutions in Bimetric Gravity and their Observational Tests*, *JCAP* **03** (2012) 042, [1111.1655].

[66] Y. Akrami, T. S. Koivisto and M. Sandstad, *Accelerated expansion from ghost-free bigravity: a statistical analysis with improved generality*, *JHEP* **03** (2013) 099, [1209.0457].

[67] F. Koennig, A. Patil and L. Amendola, *Viable cosmological solutions in massive bimetric gravity*, *JCAP* **03** (2014) 029, [1312.3208].

[68] M. Lüben, A. Schmidt-May and J. Weller, *Physical parameter space of bimetric theory and SN1a constraints*, *JCAP* **09** (2020) 024, [2003.03382].

[69] S. Dhawan, A. Goobar, E. Mörtsell, R. Amanullah and U. Feindt, *Narrowing down the possible explanations of cosmic acceleration with geometric probes*, *JCAP* **07** (2017) 040, [1705.05768].

[70] M. Lindner, K. Max, M. Platscher and J. Rezacek, *Probing alternative cosmologies through the inverse distance ladder*, *JCAP* **10** (2020) 040, [2002.01487].

[71] A. Caravano, M. Lüben and J. Weller, *Combining cosmological and local bounds on bimetric theory*, 2101.08791.

[72] J. Enander and E. Mortsell, *On stars, galaxies and black holes in massive bigravity*, *JCAP* **11** (2015) 023, [1507.00912].

[73] M. Platscher, J. Smirnov, S. Meyer and M. Bartelmann, *Long Range Effects in Gravity Theories with Vainshtein Screening*, *JCAP* **12** (2018) 009, [1809.05318].

[74] M. Lüben, E. Mörtsell and A. Schmidt-May, *Bimetric cosmology is compatible with local tests of gravity*, *Class. Quant. Grav.* **37** (2020) 047001, [1812.08686].

[75] S. Sjors and E. Mortsell, *Spherically Symmetric Solutions in Massive Gravity and Constraints from Galaxies*, *JHEP* **02** (2013) 080, [1111.5961].

[76] J. Enander and E. Mörtsell, *Strong lensing constraints on bimetric massive gravity*, *JHEP* **10** (2013) 031, [1306.1086].

[77] A. De Felice, T. Nakamura and T. Tanaka, *Possible existence of viable models of bi-gravity with detectable graviton oscillations by gravitational wave detectors*, *PTEP* **2014** (2014) 043E01, [1304.3920].

[78] M. Fasiello and R. H. Ribeiro, *Mild bounds on bigravity from primordial gravitational waves*, *JCAP* **07** (2015) 027, [1505.00404].

[79] G. Cusin, R. Durrer, P. Guarato and M. Motta, *Inflationary perturbations in bimetric gravity*, *JCAP* **09** (2015) 043, [1505.01091].

[80] K. Max, M. Platscher and J. Smirnov, *Decoherence of Gravitational Wave Oscillations in Bigravity*, *Phys. Rev. D* **97** (2018) 064009, [1712.06601].

[81] S. F. Hassan and M. Kocic, *On the local structure of spacetime in ghost-free bimetric theory and massive gravity*, *JHEP* **05** (2018) 099, [1706.07806].

[82] N. Higham, *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2008.

[83] C. de Rham, L. Heisenberg and R. H. Ribeiro, *On couplings to matter in massive (bi-)gravity*, *Class. Quant. Grav.* **32** (2015) 035022, [1408.1678].

[84] C. de Rham, L. Heisenberg and R. H. Ribeiro, *Ghosts and matter couplings in massive gravity, bigravity and multigravity*, *Phys. Rev. D* **90** (2014) 124042, [1409.3834].

[85] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Particular Solutions in Bimetric Theory and Their Implications*, *Int. J. Mod. Phys. D* **23** (2014) 1443002, [1407.2772].

[86] Y. Yamashita, A. De Felice and T. Tanaka, *Appearance of Boulware–Deser ghost in bigravity with doubly coupled matter*, *Int. J. Mod. Phys. D* **23** (2014) 1443003, [1408.0487].

[87] Y. Akrami, T. S. Koivisto, D. F. Mota and M. Sandstad, *Bimetric gravity doubly coupled to matter: theory and cosmological implications*, *JCAP* **10** (2013) 046, [1306.0004].

[88] D. Comelli, M. Crisostomi and L. Pilo, *FRW Cosmological Perturbations in Massive Bigravity*, *Phys. Rev.* **D90** (2014) 084003, [1403.5679].

[89] A. De Felice, A. E. Gümrükçüoğlu, S. Mukohyama, N. Tanahashi and T. Tanaka, *Viable cosmology in bimetric theory*, *JCAP* **1406** (2014) 037, [1404.0008].

[90] M. Lüben, A. Schmidt-May and J. Smirnov, *Vainshtein Screening in Bimetric Cosmology*, *Phys. Rev. D* **102** (2020) 123529, [1912.09449].

[91] M. S. Volkov, *Cosmological solutions with massive gravitons in the bigravity theory*, *JHEP* **01** (2012) 035, [1110.6153].

[92] M. S. Volkov, *Exact self-accelerating cosmologies in the ghost-free bigravity and massive gravity*, *Phys. Rev. D* **86** (2012) 061502, [1205.5713].

[93] M. S. Volkov, *Self-accelerating cosmologies and hairy black holes in ghost-free bigravity and massive gravity*, *Class. Quant. Grav.* **30** (2013) 184009, [1304.0238].

[94] G. Cusin, R. Durrer, P. Guarato and M. Motta, *A general mass term for bigravity*, *JCAP* **04** (2016) 051, [1512.02131].

[95] A. De Felice, A. E. Gumrukcuoglu and S. Mukohyama, *Massive gravity: nonlinear instability of the homogeneous and isotropic universe*, *Phys. Rev. Lett.* **109** (2012) 171101, [1206.2080].

[96] E. Mortsell, *Cosmological histories in bimetric gravity: A graphical approach*, *JCAP* **02** (2017) 051, [1701.00710].

[97] R. Geroch, *Faster Than Light?*, *AMS/IP Stud. Adv. Math.* **49** (2011) 59–70, [1005.1614].

[98] A. Higuchi, *Forbidden Mass Range for Spin-2 Field Theory in De Sitter Space-time*, *Nucl. Phys. B* **282** (1987) 397–436.

[99] M. Fasiello and A. J. Tolley, *Cosmological Stability Bound in Massive Gravity and Bigravity*, *JCAP* **1312** (2013) 002, [1308.1647].

[100] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59** (1980) 135–157.

[101] H. Akaike, *A new look at the statistical model identification*, *IEEE Transactions on Automatic Control* **19** (1974) 716–723.

[102] G. Schwarz, *Estimating the Dimension of a Model*, *Annals Statist.* **6** (1978) 461–464.

[103] A. R. Liddle, *How many cosmological parameters?*, *Mon. Not. Roy. Astron. Soc.* **351** (2004) L49–L53, [astro-ph/0401198].

[104] A. R. Liddle, *Information criteria for astrophysical model selection*, *Mon. Not. Roy. Astron. Soc.* **377** (2007) L74–L78, [astro-ph/0701113].

[105] Pan-STARRS1 collaboration, D. M. Scolnic et al., *The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample*, *Astrophys. J.* **859** (2018) 101, [1710.00845].

[106] M. Goliath, R. Amanullah, P. Astier, A. Goobar and R. Pain, *Supernovae and the nature of the dark energy*, *Astron. Astrophys.* **380** (2001) 6–18, [astro-ph/0104009].

[107] L. Chen, Q.-G. Huang and K. Wang, *Distance Priors from Planck Final Release*, *JCAP* **02** (2019) 028, [1808.05724].

[108] W. Hu and N. Sugiyama, *Small scale cosmological perturbations: An Analytic approach*, *Astrophys. J.* **471** (1996) 542–570, [astro-ph/9510117].

[109] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell et al., *The 6dF Galaxy Survey: Baryon Acoustic Oscillations and the Local Hubble Constant*, *Mon. Not. Roy. Astron. Soc.* **416** (2011) 3017–3032, [1106.3366].

[110] A. J. Ross, L. Samushia, C. Howlett, W. J. Percival, A. Burden and M. Manera, *The clustering of the SDSS DR7 main Galaxy sample – I. A 4 per cent distance measure at* $z = 0.15$, *Mon. Not. Roy. Astron. Soc.* **449** (2015) 835–847, [1409.3242].

[111] BOSS collaboration, S. Alam et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 2617–2652, [1607.03155].

[112] J. E. Bautista et al., *The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations at redshift of 0.72 with the DR14 Luminous Red Galaxy Sample*, *Astrophys. J.* **863** (2018) 110, [1712.08064].

[113] G.-B. Zhao et al., *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: a tomographic measurement of cosmic structure growth and expansion rate based on optimal redshift weights*, *Mon. Not. Roy. Astron. Soc.* **482** (2019) 3497–3513, [1801.03043].

[114] D. J. Eisenstein and W. Hu, *Baryonic features in the matter transfer function*, *Astrophys. J.* **496** (1998) 605, [astro-ph/9709112].

[115] M. Kunz, *The dark degeneracy: On the number and nature of dark components*, *Phys. Rev. D* **80** (2009) 123001, [astro-ph/0702615].

[116] S. W. Allen, R. W. Schmidt, H. Ebeling, A. C. Fabian and L. van Speybroeck, *Constraints on dark energy from Chandra observations of the largest relaxed galaxy clusters*, *Mon. Not. Roy. Astron. Soc.* **353** (2004) 457, [astro-ph/0405340].

[117] D. Kirkman, D. Tytler, N. Suzuki, J. M. O'Meara and D. Lubin, *The Cosmological baryon density from the deuterium to hydrogen ratio towards QSO absorption systems: D/H towards Q1243+3047*, *Astrophys. J. Suppl.* **149** (2003) 1, [astro-ph/0302006].

[118] R. H. Cyburt, B. D. Fields, K. A. Olive and T.-H. Yeh, *Big Bang Nucleosynthesis: 2015*, *Rev. Mod. Phys.* **88** (2016) 015004, [1505.01076].

[119] J. Macquart, J. Prochaska, M. McQuinn et al., *A census of baryons in the Universe from localized fast radio bursts*, *Nature* **581** (2020) 391–395.

[120] J. P. Kneller and G. Steigman, *BBN for pedestrians*, *New J. Phys.* **6** (2004) 117, [astro-ph/0406320].

[121] B. D. Fields, K. A. Olive, T.-H. Yeh and C. Young, *Big-Bang Nucleosynthesis after Planck*, *JCAP* **03** (2020) 010, [1912.01132]. [Erratum: JCAP 11, E02 (2020)].

[122] G. Steigman, *Primordial Nucleosynthesis in the Precision Cosmology Era*, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 463–491, [0712.1100].

[123] E. Aver, K. A. Olive and E. D. Skillman, *The effects of He I λ10830 on helium abundance determinations*, *JCAP* **07** (2015) 011, [1503.08146].

[124] E. Dimastrogiovanni, M. Fasiello and A. E. Gümrükcüoglü, *Spinning guest fields during inflation: leftover signatures*, *JCAP* **11** (2021) 047, [2108.06722].

[125] J. T. Jebsen, *Über die allgemeinen kugelsymmetrischen Lösungen der Einsteinschen Gravitationsgleichungen im Vakuum*, *Ark. Mat. Astr. Fys.* **15** (1921) .

[126] J. T. Jebsen, *On the general spherically symmetric solutions of einstein's gravitational equations in vacuo*, *General Relativity and Gravitation* **37** (Dec, 2005) 2253–2259.

[127] G. Birkhoff and R. Langer, *Relativity and Modern Physics*. Harvard University Press, 1923.

[128] D. Comelli, M. Crisostomi, F. Nesti and L. Pilo, *Spherically Symmetric Solutions in Ghost-Free Massive Gravity*, *Phys. Rev. D* **85** (2012) 024044, [1110.4967].

[129] M. Platscher and J. Smirnov, *Degravitation of the Cosmological Constant in Bigravity*, *JCAP* **03** (2017) 051, [1611.09385].

[130] E. Babichev and M. Crisostomi, *Restoring general relativity in massive bigravity theory*, *Phys. Rev. D* **88** (2013) 084002, [1307.3640].

[131] M. Lüben, *Phenomenological aspects of bimetric theory*, Ph.D. thesis, Munich U., 4, 2021. 10.5282/edoc.28191.

[132] J. Schwab, A. S. Bolton and S. A. Rappaport, *Galaxy-Scale Strong Lensing Tests of Gravity and Geometric Cosmology: Constraints and Systematic Limitations*, *Astrophys. J.* **708** (2010) 750–757, [0907.4992].

[133] F. Dar, C. De Rham, J. T. Deskins, J. T. Giblin and A. J. Tolley, *Scalar Gravitational Radiation from Binaries: Vainshtein Mechanism in Time-dependent Systems*, *Class. Quant. Grav.* **36** (2019) 025008, [1808.02165].

[134] 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].

[135] LIGO Scientific, Virgo, Fermi-GBM, INTEGRAL collaboration, B. P. Abbott et al., *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, *Astrophys. J. Lett.* **848** (2017) L13, [1710.05834].

[136] LIGO Scientific, Virgo collaboration, B. P. 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].

[137] LIGO Scientific, Virgo collaboration, R. Abbott et al., *GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run*, *Phys. Rev. X* **11** (2021) 021053, [2010.14527].

[138] LIGO Scientific, VIRGO collaboration, R. Abbott et al., *GWTC-2.1: Deep Extended Catalog of Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run*, 2108.01045.

[139] LIGO Scientific, VIRGO, KAGRA collaboration, R. Abbott et al., *GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run*, 2111.03606.

[140] M. Lagos and P. G. Ferreira, *Cosmological perturbations in massive bigravity*, *JCAP* **12** (2014) 026, [1410.0207].

[141] L. Amendola, F. Könnig, M. Martinelli, V. Pettorino and M. Zumalacarregui, *Surfing gravitational waves: can bigravity survive growing tensor modes?*, *JCAP* **05** (2015) 052, [1503.02490].

[142] M. E. L. Platscher, *Phenomenology of Massive Spin-2 Fields*, Ph.D. thesis, U. Heidelberg (main), 2019. 10.11588/heidok.00025876.

[143] A. G. Riess et al., *A 2.4% Determination of the Local Value of the Hubble Constant*, *Astrophys. J.* **826** (2016) 56, [1604.01424].

[144] LIGO Scientific, Virgo collaboration, B. P. Abbott et al., *GW150914: First results from the search for binary black hole coalescence with Advanced LIGO*, *Phys. Rev. D* **93** (2016) 122003, [1602.03839].

[145] LIGO Scientific, Virgo collaboration, B. P. Abbott et al., *Properties of the Binary Black Hole Merger GW150914*, *Phys. Rev. Lett.* **116** (2016) 241102, [1602.03840].

[146] E. Schildt, *Gravitational waves in bimetric gravity*, Master's thesis, Uppsala University, 2021.

[147] COBE collaboration, G. F. Smoot et al., *Structure in the COBE differential microwave radiometer first year maps*, *Astrophys. J. Lett.* **396** (1992) L1–L5.

[148] D. Comelli, M. Crisostomi and L. Pilo, *Perturbations in Massive Gravity Cosmology*, *JHEP* **06** (2012) 085, [1202.1986].

[149] N. Khosravi, H. R. Sepangi and S. Shahidi, *Massive cosmological scalar perturbations*, *Phys. Rev.* **D86** (2012) 043517, [1202.2767].

[150] M. Berg, I. Buchberger, J. Enander, E. Mortsell and S. Sjors, *Growth Histories in Bimetric Massive Gravity*, *JCAP* **1212** (2012) 021, [1206.3496].

[151] F. Könnig and L. Amendola, *Instability in a minimal bimetric gravity model*, *Phys. Rev.* **D90** (2014) 044030, [1402.1988].

[152] A. R. Solomon, Y. Akrami and T. S. Koivisto, *Linear growth of structure in massive bigravity*, *JCAP* **1410** (2014) 066, [1404.4061].

[153] F. Koennig, Y. Akrami, L. Amendola, M. Motta and A. R. Solomon, *Stable and unstable cosmological models in bimetric massive gravity*, *Phys. Rev.* **D90** (2014) 124014, [1407.4331].

[154] J. Enander, Y. Akrami, E. Mörtsell, M. Renneby and A. R. Solomon, *Integrated Sachs-Wolfe effect in massive bigravity*, *Phys. Rev.* **D91** (2015) 084046, [1501.02140].

[155] F. Könnig, *Higuchi Ghosts and Gradient Instabilities in Bimetric Gravity*, *Phys. Rev.* **D91** (2015) 104019, [1503.07436].

[156] M. Kenna-Allison, A. E. Gümrükçüoğlu and K. Koyama, *Viability of bigravity cosmology*, *Phys. Rev.* **D99** (2019) 104032, [1812.05496].

[157] V. Mukhanov, H. Feldman and R. Brandenberger, *Theory of cosmological perturbations*, *Physics Reports* **215** (1992) 203–333.

[158] Euclid Theory Working Group collaboration, L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **16** (2013) 6, [1206.1225].

[159] A. Ijjas, F. Pretorius and P. J. Steinhardt, *Stability and the Gauge Problem in Non-Perturbative Cosmology*, *JCAP* **1901** (2019) 015, [1809.07010].

[160] Y. Akrami, S. F. Hassan, F. Könnig, A. Schmidt-May and A. R. Solomon, *Bimetric gravity is cosmologically viable*, *Phys. Lett.* **B748** (2015) 37–44, [1503.07521].

[161] E. Mortsell and J. Enander, *Scalar instabilities in bimetric gravity: The Vainshtein mechanism and structure formation*, *JCAP* **1510** (2015) 044, [1506.04977].

[162] M. Kocic, *Geometric mean of bimetric spacetimes*, *Class. Quant. Grav.* **38** (2021) 075023, [1803.09752].

[163] M. Kocic, *Causal propagation of constraints in bimetric relativity in standard 3+1 form*, *JHEP* **10** (2019) 219, [1804.03659].

[164] M. Kocic, A. Lundkvist and F. Torsello, *On the ratio of lapses in bimetric relativity*, *Class. Quant. Grav.* **36** (2019) 225013, [1903.09646].

[165] F. Torsello, *The mean gauges in bimetric relativity*, *Class. Quant. Grav.* **36** (2019) 235010, [1904.09297].

[166] F. Torsello, `bimEX`*: A Mathematica package for exact computations in* $3+1$ *bimetric relativity*, *Comput. Phys. Commun.* **247** (2020) 106948, [1904.10464].

[167] F. Torsello, *Theoretical and numerical bimetric relativity*, Ph.D. thesis, Stockholm University, Faculty of Science, Department of Physics., 2020.

[168] J. R. Oppenheimer and H. Snyder, *On Continued gravitational contraction*, *Phys. Rev.* **56** (1939) 455–459.

[169] M. Ruiz, M. Alcubierre and D. Nunez, *Regularization of spherical and axisymmetric evolution codes in numerical relativity*, *Gen. Rel. Grav.* **40** (2008) 159–182, [0706.0923].

[170] E. V. Linder, *Cosmic growth history and expansion history*, *Phys. Rev. D* **72** (2005) 043529, [astro-ph/0507263].

[171] Euclid Imaging collaboration, A. Refregier, A. Amara, T. D. Kitching, A. Rassat, R. Scaramella and J. Weller, *Euclid Imaging Consortium Science Book*, 1001.0061.

[172] L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **21** (2018) 2, [1606.00180].

[173] G. Lemaître, *L'Univers en expansion*, *Annales de la Societe Scietique de Bruxelles* **53** (1933) .

[174] R. C. Tolman, *Effect of Inhomogeneity on Cosmological Models*, *Proceedings of the National Academy of Science* **20** (1934) 169–176.

[175] H. Bondi, *Spherically symmetrical models in general relativity*, *Mon. Not. Roy. Astron. Soc.* **107** (1947) 410–425.

[176] E. Poisson, *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics*. Cambridge University Press, 12, 2009, 10.1017/CBO9780511606601.

[177] P. Vaidya, *The Gravitational Field of a Radiating Star*, *Proc. Natl. Inst. Sci. India* **A33** (1951) 264.

[178] V. Husain, *Exact solutions for null fluid collapse*, *Phys. Rev.* **D53** (1996) 1759–1762, [gr-qc/9511011].

[179] A. Wang and Y. Wu, *Generalized Vaidya solutions*, *Gen. Rel. Grav.* **31** (1999) 107, [gr-qc/9803038].

[180] W. B. Bonnor and P. C. Vaidya, *Spherically symmetric radiation of charge in Einstein-Maxwell theory*, *Gen. Rel. Grav.* **1** (1970) 127–130.

[181] A. Ori, *Charged null fluid and the weak energy condition*, *Class.Quant.Grav.* **8** (1991) 1159.

[182] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2011, 10.1017/CBO9780511524646.

[183] E. Babichev and A. Fabbri, *A class of charged black hole solutions in massive (bi)gravity*, *JHEP* **07** (2014) 016, [1405.0581].

[184] C. Deffayet and T. Jacobson, *On horizon structure of bimetric spacetimes*, *Class. Quant. Grav.* **29** (2012) 065009, [1107.4978].

[185] L. Rezzolla and O. Zanotti, *Relativistic Hydrodynamics*. EBSCO ebook academic collection. OUP Oxford, 2013.

[186] V. Baccetti, P. Martin-Moruno and M. Visser, *Null Energy Condition violations in bimetric gravity*, *JHEP* **08** (2012) 148, [1206.3814].

[187] S. Capozziello, F. S. N. Lobo and J. P. Mimoso, *Generalized energy conditions in Extended Theories of Gravity*, *Phys. Rev.* **D91** (2015) 124019, [1407.7293].

[188] Y. Aharonov, A. Komar and L. Susskind, *Superluminal behavior, causality, and instability*, *Phys. Rev.* **182** (1969) 1400–1403.

[189] S. Frittelli, *Note on the propagation of the constraints in standard 3+1 general relativity*, *Phys. Rev. D* **55** (May, 1997) 5992–5996.

[190] E. Mörtsell and S. Dhawan, *Does the Hubble constant tension call for new physics?*, *JCAP* **09** (2018) 025, [1801.07260].

[191] R. M. Wald, *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984, 10.7208/chicago/9780226870373.001.0001.