Observational tests and theoretical developments of hybrid MOND dark matter models

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Physik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Tobias Mistele
aus Heilbronn

Frankfurt am Main 2021
(D 30)

vom Fachbereich Physik der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan:
Prof. Dr. Harald Appelshäuser

Gutachter:
Prof. Dr. Dr. h.c. Marcus Bleicher
Prof. Dr. Dr. h.c. mult. Horst Stöcker

Datum der Disputation:

Contents

Zusammenfassung v

1. Introduction 1
- 1.1. Why MOND is wrong . . . 3
- 1.2. Why ΛCDM is wrong . . . 6
- 1.3. Hybrid models . . . 12
- 1.4. Superfluid dark matter . . . 14

2. Observational tests of superfluid dark matter 21
- 2.1. Strong lensing . . . 21
  - 2.1.1. The model . . . 22
  - 2.1.2. Results . . . 24
- 2.2. The Milky Way rotation curve . . . 26
  - 2.2.1. The model . . . 26
  - 2.2.2. Results . . . 28

3. Three problems of superfluid dark matter and their solution 35
- 3.1. The stability problem . . . 35
- 3.2. The MOND limit problem . . . 36
- 3.3. The equilibrium problem . . . 40
- 3.4. A solution: Two-field SFDM . . . 42

4. Cherenkov radiation from stars in hybrid models 51
- 4.1. Introduction . . . 51
- 4.2. Cherenkov radiation in a prototype hybrid model . . . 54
  - 4.2.1. Prototype Lagrangian . . . 55
  - 4.2.2. Cherenkov radiation . . . 56
  - 4.2.3. Orbits of stars emitting Cherenkov radiation . . . 58
  - 4.2.4. Regime of validity . . . 62
  - 4.2.5. Constraints from Cherenkov radiation . . . 65
- 4.3. Application to standard SFDM . . . 66
  - 4.3.1. Perturbations . . . 67
  - 4.3.2. Constraints . . . 68
- 4.4. Application to two-field SFDM . . . 73
- 4.5. Application to the SZ model . . . 76
  - 4.5.1. Suppressed matter coupling . . . 78
  - 4.5.2. Energy loss through Cherenkov radiation . . . 79
- 4.6. Discussion . . . 80
- 4.7. Summary . . . 82

5. Conclusion **83**
5.1. Summary 83
5.2. Outlook 87

Acknowledgements **89**

A. Calculation of energy loss through Cherenkov radiation **91**
A.1. Special case: $\vec{V} \parallel \hat{a}$ 92
A.2. Special case: $\vec{V} \perp \hat{a}$ 93
A.3. Direction of friction force 94
A.4. Two-field SFDM corrections 94

B. Standard SFDM perturbations **97**

C. Standard SFDM Cherenkov radiation **99**
C.1. Reference calculation 99
C.2. Mapping to the reference system 101
C.3. Energy loss without cutoffs 102
C.4. Cutoffs 106
C.5. Direction of friction force 110
C.6. When are perturbations small 113

D. Two-field SFDM perturbations **119**
D.1. Equations of motion 119
D.2. Effective Lagrangian 121

E. SZ model perturbations **123**
E.1. Equations of motion 123
E.2. Effective Lagrangian 127

Bibliography **129**

Zusammenfassung

Sowohl auf kosmologischen als auf galaktischen Skalen gibt es Phänomene, die üblicherweise durch kalte dunkle Materie (CDM) erklärt werden. Das sind zum Beispiel die Anisotropien der kosmischen Hintergrundstrahlung und galaktische Rotationskurven. Ein alternativer Erklärungsansatz ist die sogenannte modifizierte Gravitation, welche diese Phänomene durch ein modifiziertes Gravitationsgesetz erklärt. Insbesondere die sogenannte modifizierte Newtonsche Dynamik (MOND) ist auf galaktischen Skalen sehr erfolgreich.

Tatsächlich sind sowohl dunkle Materie (DM) als auch MOND für sich gesehen nicht zufriedenstellend. Unter anderem gibt es in MOND bisher keine gute Erklärung für die Anisotropien der kosmischen Hintergrundstrahlung. Zumindest nicht, ohne doch eine Form von dunkler Materie zu postulieren. DM dagegen hat unter anderem keine gute Erklärung für beobachtete Regelmäßigkeiten in Galaxien wie zum Beispiel die sogenannte Radial Acceleration Relation (RAR).

Ein möglicher Ausweg sind sogenannte hybride MOND-DM-Modelle. Das sind Modelle, die sowohl ein MOND-artiges Gravitationsgesetz in Galaxien als auch eine DM-Komponente auf kosmologischen Skalen postulieren. Ein Beispiel eines solchen Modells ist die sogenannte superfluide dunkle Materie (SFDM). SFDM postuliert eine spezielle Art von Teilchen, die um Galaxien herum zu einem Superfluid kondensieren. Die Phononen dieses Superfluids können dann eine MOND-artige Kraft vermitteln. Auf größeren Skalen dagegen sind diese Teilchen in einer nicht-superfluiden Phase und verhalten sich wie normale DM-Teilchen. Damit vereint das Modell die Vorteile von MOND auf Galaxienskalen mit denen von DM auf kosmologischen Skalen.

Wir diskutieren SFDM und andere Hybridmodelle wie zum Beispiel ein Modell, das Skordis und Złośnik vor kurzem vorgestellt haben (SZ-Modell), das sowohl auf Bekensteins TeVeS-Modell als auch auf Einstein-Äther-Modellen basiert. Zum einen testen wir diese Modelle mit Daten aus Beobachtungen. Zum anderen verbessern wir das theoretische Verständnis dieser Modelle und entwickeln sie weiter. Wir beginnen mit zwei Tests des SFDM-Modells mit Beobachtungsdaten.

Der erste Test ist, ob SFDM beobachtete starke Gravitationslinseneffekte erklären kann. Das ist nicht-trivial, weil die MOND-artige Phononkraft nicht auf Photonen wirken kann. Denn aus der Beobachtung von GW170817 und des zugehörigen elektromagnetischen Gegenstücks wissen wir, dass die gravitative Tensormode und Licht sich mit der gleichen Geschwindigkeit ausbreiten. Eine MOND-Kraft, die auf Photonen wirkt, führt oft zu einer signifikanten Shapiro-Zeitverzögerung für Photonen. Deshalb kann die MOND-artige Pho-

nonkraft in SFDM nicht zu Gravitationslinseneffekten beitragen. Wir zeigen, dass SFDM trotzdem die beobachteten Gravitationslinsendaten erklären kann. Der Superfluid-Halo um Galaxien hat genügend zusätzliche Masse, um die beobachteten Gravitationslinseneffekte zu erklären.

Dafür muss die gravitative Anziehungskraft des Halos bei großen galaktozentrischen Radien groß sein, denn dort wird ein großer Teil des Gravitationslinsen-Signals erzeugt. Bei kleineren Radien, wo Rotationskurven gemessen werden, muss seine gravitative Anziehungskraft dagegen relativ klein sind. Ansonsten wären die Rotationsgeschwindigkeiten zu groß, weil auf Sterne sowohl die MOND-artige Phononkraft als auch die normale Newtonsche Gravitationskraft des Halos wirkt. Das ist in SFDM prinzipiell möglich, weil der Superfluid-Halo ein sehr flaches Dichteprofil hat. Wir verifizieren, dass SFDM gleichzeitig die Einstein-Radien und die Geschwindigkeitsdispersionen der Gravitationslinsen-Galaxien erklären kann.

Der zweite Test ist die Rotationskurve der Milchstraße. Wir zeigen, dass SFDM die beobachtete Rotationskurve akzeptabel fitten kann, dafür aber ca. 20% weniger baryonische Masse benötigt als Standard-MOND-Modelle. Der Grund dafür ist, dass in SFDM der Übergang zwischen dem Newtonschen und dem MOND-Regime gradueller verläuft, sodass die Gesamtkraft bei nicht sehr großen oder sehr kleinen Newtonschen Beschleunigungen größer ist als in MOND.

Wir schätzen außerdem die Größe des superfluiden Kerns der Milchstraße sowie deren Dunkle-Materie-Gesamtmasse ab. Dazu verallgemeinern wir zuerst die gängigen Methoden, wie man die Größe des superfluiden Kerns abschätzen kann, von sphärischer zu zylindrischer Symmetrie. Das ist relativ unkompliziert, weil wir bei großen Radien sphärisch-symmetrische Randbedingungen annehmen. Diese Randbedingungen haben einen freien Parameter, μ_∞, der die Größe des superfluiden Kerns und die Dunkle-Materie-Gesamtmasse bestimmt. Für viriale Dunkle-Materie-Massen M_{200}^{DM} im Bereich $0.5-3.0\cdot10^{12}\,M_\odot$ variiert der sogenannte NFW-Radius zwischen 65 kpc und 73 kpc, während der sogenannte thermische Radius zwischen 67 kpc und 105 kpc variiert.

Als nächstes betrachten wir SFDM genauer aus einer theoretischen Perspektive. Wir identifizieren drei Probleme, die innerhalb SFDM nicht einfach vermieden werden können. Das erste Problem ist, dass der Gleichgewichtszustand instabil ist. Diese Instabilität wird normalerweise durch Korrektur-Terme für endliche Temperaturen, parametrisiert durch einen Parameter $\bar{\beta}$, vermieden. Allerdings sind sowohl der numerische Werte von $\bar{\beta}$ als auch die Form dieser Korrektur-Terme ad-hoc und könnten sich leicht als unphysikalisch erweisen.

Das zweite Problem ist, dass viele Galaxien nicht das MOND-Limit von SFDM erreichen können, obwohl dieses MOND-Limit eine der Haupt-Motivationen von SFDM ist. Die Größe ε_* kontrolliert das MOND-Limit. Die Phononkraft erfüllt nur im MOND-Limit $\varepsilon_* \ll 1$ eine MOND-artige Gleichung. Allerdings haben isolierte Galaxien, zumindest für $\bar{\beta} \approx 2$, sogar für $\varepsilon_* = \mathcal{O}(1)$ MOND-artige Rotationskurve, obwohl die Phononkraft dann keine MOND-artige Gleichung erfüllt. Wir bezeichnen diesen Fall als Pseudo-MOND-Limit.

Dieses Pseudo-MOND-Limit ist aber kein zufriedenstellender Ersatz für das echte MOND-Limit $\varepsilon_* \ll 1$. Denn das Pseudo-MOND-Limit hängt sensitiv von den Details der oben genannten Korrekturen für endliche Temperaturen ab. Zum Beispiel funktioniert das Pseudo-MOND-Limit nur für $\bar{\beta} \approx 2$. Wie die Korrekturen für endliche Temperaturen könnte sich also auch das Pseudo-MOND-Limit leicht als unphysikalisch erweisen.

Das dritte Problem ist, dass es in SFDM in Galaxien möglicherweise keinen Superfluid-Gleichgewichtszustand gibt, der deutlich länger als galaktische Zeitskalen existiert. Der Grund ist, dass die MOND-artige Phononkraft eine direkte Kopplung der Baryonen zum Phonon-Feld θ ohne Ableitungen benötigt. Das bricht explizit die $U(1)$-Symmetrie von θ, die sonst typisch für Superfluide ist. Heuristisch gesprochen entspricht ein chemisches Potential μ einer Lösung $\theta = \mu \cdot t$. Für solche Lösungen hat die Baryon-Phonon-Kopplung eine explizite Zeitabhängigkeit, weil θ dort ohne Ableitungen auftritt. Wir können diese Zeitabhängigkeit nur dann ignorieren und einen zeitunabhängigen Gleichgewichtszustand in Galaxien annehmen, wenn wir uns auf Zeiten kürzer als eine Zeitskala t_Q beschränken. Wir schätzen diese Zeitskala ab und finden, dass sie vergleichbar mit galaktischen Zeitskalen sein kann. Daher könnten sich Galaxien in SFDM entgegen den üblichen Annahmen nicht in einem Gleichgewichtszustand befinden.

Diese drei Probleme haben alle die gleiche Grundursache, nämlich dass dem Phonon-Feld θ eine Doppelrolle zukommt. Es liefert sowohl die Masse des Superfluids als auch die MOND-artige Phononkraft. Daher schlagen wir ein verbessertes Modell vor, das diese Probleme vermeidet. Dieses verbesserte Modell funktioniert, indem es die zwei Rollen des Phonon-Feldes zwischen zwei Feldern aufteilt. Ein Feld liefert das Superfluid, das andere überträgt eine MOND-artige Kraft. Daher nennen wir dieses Modell Zwei-Felder-SFDM. Wir zeigen, dass die Phänomenologie dieses Modells auf galaktischen Skalen ähnlich ist wie im originalen SFDM-Modell.

Genauer gesagt ist die Phänomenologie des Zwei-Felder-Modells auf galaktischen Skalen bezüglich des superfluiden Kerns ähnlich der des Standard-SFDM-Modells. Außerhalb dieses superfluiden Kerns stimmt das nur, wenn wir den sogenannten NFW-Radius als den Radius annehmen, an dem der superfluide Kern endet. In Standard-SFDM stimmt dieser NFW-Radius einigermaßen mit dem sogenannten thermischen Radius überein. Im Zwei-Felder-Modell dagegen können die beiden Radien weit voneinander abweichen. Das wirft die Frage auf, ob der NFW-Radius oder der thermische Radius im Zwei-Felder-Modell korrekt ist oder ob der Übergang von der superfluiden zur nichtsuperfluiden Phase vielleicht ganz anders behandelt werden muss.

Zumindest prinzipiell gibt es dieses Problem auch schon im originalen SFDM-Modell. Der NFW- und der thermische Radius sind nämlich nur dann ähnlich, wenn die Parameter-Kombination σ/m^5 nah an dem Wert ist, der üblicherweise angenommen wird. Im Allgemeinen liegen diese zwei Radien aber auch in Standard-SFDM weit auseinander. Das ist relevant, weil der Selbstwechselwirkungs-Wirkungsquerschnitt σ in Standard-SFDM normalerweise ad-hoc

gewählt und nicht von einem Lagrangian ausgerechnet wird. Das heißt, es ist nicht klar, ob der üblicherweise verwendet Wert realistisch ist. In jedem Fall muss der Übergang von der superfluiden zu der nicht-superfluiden Phase in Galaxien sowohl in Standard-SFDM als auch in Zwei-Felder-SFDM genauer untersucht werden.

Schließlich schlagen wir einen neuartigen Test von hybriden MOND-DM-Modellen bezüglich Cherenkov-Strahlung vor. Dieser Test betrifft nicht nur SFDM sondern eine Vielzahl von hybriden MOND-DM-Modellen. Genauer gesagt betrifft er Hybrid-Modelle mit einem gemeinsamen Ursprung von galaktischen und kosmologischen Phänomenen. Cherenkov-Strahlung ist ein wohlbekanntes Phänomen in Modellen modifizierter Gravitation. Wenn normale Materie an eine masslose Mode gekoppelt ist, die mit einer Geschwindigkeit c_s propagiert, und ein Objekt aus normaler Materie sich schneller als c_s bewegt, strahlt dieses Objekt Energie in Form der masselosen Mode ab. Normalerweise geben nur relativistische Objekte Cherenkov-Strahlung ab, weil c_s normalerweise relativistisch ist. In Hybrid-Modellen mit einem gemeinsamen Ursprung von galaktischen und kosmologischen Phänomenen ist dies aber nicht so. Auch nicht-relativistische Objekte wie Sterne können Cherenkov-Strahlung abgeben. Der Grund ist folgendermaßen. Für die MOND-Phänomenologie auf galaktischen Skalen gibt es eine masselose Mode, die direkt an normale Materie gekoppelt ist. Für die CDM-Phänomenologie auf kosmologischen Skalen gibt es ein kollisionsloses Fluid mit nicht-relativistischer Schallgeschwindigkeit. Wenn diese Phänomene einen gemeinsamen Ursprung haben, hat die masselose Mode, die direkt an Materie gekoppelt ist, typischerweise ein nicht-relativistisches c_s, da sie in Zusammenhang mit der nicht-relativistischen Schallgeschwindigkeit des kollisionslosen Fluids steht. Daher können in hybriden MOND-DM-Modellen sogar nicht-relativistische Objekte wie Sterne Cherenkov-Strahlung abgeben.

Diese Cherenkov-Strahlung ähnelt auf der einen Seite normaler gravitativer Cherenkov-Strahlung in dem Sinne, dass sie durch eine direkte Kopplung zu normaler Materie zustande kommt. Auf der anderen Seite ähnelt sie normaler dynamischer Reibung in dem Sinne, dass sie es sogar nicht-relativistischen Objekten erlaubt, Energie zu verlieren. In der speziellen Näherung, die wir machen, ähnelt diese Cherenkov-Strahlung dynamischer Reibung noch in einem anderen Punkt. Nämlich gibt es keinen signifikanten Rückstoß. Denn jede einzelne Emission von Cherenkov-Strahlung ist niederenergetisch. Ein Überschall-Objekt verliert nur durch eine große Zahl dieser niederenergetischen Emissionen einen signifikanten Anteil seiner Energie. Das ist eine Folge der strikten Cutoffs, die wir in unserer Rechnung für die Phasenraum-Integrale verwenden. Diese Cutoffs haben einen doppelten Zweck. Zum einen vermeiden sie Komplikationen mit der jedem MOND-Modell inhärenten Nicht-Linearität. Zum anderen stellen sie sicher, das wir im MOND-Regime eines gegebenen Modells bleiben. Außerhalb dieses MOND-Regimes verhalten Modelle sich unter Umständen ganz anders als im MOND-Regime. Als Folge dieser Cutoffs ist der Energie-Verlust, den wir berechnen, eine untere Schranke. Der tatsächliche

Energieverlust kann größer sein, ist aber auch schwerer zu berechnen.

Wir betrachten zuerst ein Prototyp-Modell mit einem für hybride MOND-DM-Modelle typischen Lagrangian. Damit berechnen wir die Zeitskala τ_E, auf der nicht-relativistische Überschall-Objekte mit Geschwindigkeit V einen signifikanten Anteil ihrer Energie durch Cherenkov-Strahlung verlieren. Wir finden, dass τ_E proportional zu $V^3/(g_m^2\, c_s^2\, a_b^{\mathrm{gal}})$ ist. Dabei ist c_s die Propagations-Geschwindigkeit der masselosen Mode, die mit einer Kopplungskonstante proportional zu g_m direkt an normale Materie gekoppelt ist, und a_b^{gal} ist die baryonische Newtonsche Beschleunigung der Host-Galaxie an der Position des Überschall-Objektes. Diese Zeitskala ist unabhängig von der Masse des Überschall-Objektes.

Um nicht ausgeschlossen zu sein, müssen Hybrid-Modelle mindestens eine von zwei Bedingungen erfüllen. Entweder muss die Schallgeschwindigkeit c_s hinreichend groß sein, sodass die meisten nicht-relativistischen Objekte wie Sterne sich langsamer als c_s bewegen. In dem Fall ist Cherenkov-Strahlung kinematisch verboten. Oder die Zeitskala τ_E muss hinreichend groß sein, sodass Überschall-Objekte, die Cherenkov-Strahlung abgeben, dadurch auf galaktischen Zeitskalen nicht viel Energie verlieren. Letzteres kann durch eine kleine Kopplungskonstante g_m oder durch eine kleine Schallgeschwindigkeit c_s erreicht werden.

Wir wenden diese allgemeinen Ergebnisse auf drei konkrete Modelle an: Auf Standard-SFDM, auf Zwei-Felder-SFDM und auf das SZ-Modell. In Standard-SFDM liefert das Phonon-Feld θ eine masselose Mode, die direkt an Materie gekoppelt ist (für die MOND-artige Phononkraft) und die eine nicht-relativistische Schallgeschwindigkeit hat (weil diese Mode im Zusammenhang mit dem nicht-relativistischen Superfluid steht). Damit ist die oben genannte Art von Cherenkov-Strahlung in Standard-SFDM möglich und wir können Standard-SFDM mit der beobachteten Rotationskurve der Milchstraße testen. Wir schließen ein MOND-Limit in der Milchstraße für bestimmte Parameter-Werte aus. Für $\bar{\beta} = 2$ schließen wir zum Beispiel das MOND-Limit von SFDM aus für Werte der Parameter-Kombination $\sqrt{\bar{\alpha}}/m$ im Intervall $0.34\,\mathrm{eV}^{-1} \lesssim \sqrt{\bar{\alpha}}/m \lesssim 3.29\,\mathrm{eV}^{-1}$. Das schließt insbesondere die üblicherweise verwendeten Werte $\bar{\beta} = 2$ und $\sqrt{\bar{\alpha}}/m = 2.4\,\mathrm{eV}^{-1}$ aus. Kugelsternhaufen oder einzelne Sterne mit besonders großer Geschwindigkeit könnten SFDM noch weiter einschränken, was in künftigen Arbeiten genauer untersucht werden kann.

Zwei-Felder-SFDM beinhaltet zwei masselose Moden, die ungefähr den zwei Feldern entsprechen, die die MOND-artige Kraft und das Superfluid liefern. Nur die Mode, die das Superfluid liefert, lässt Cherenkov-Strahlung zu, weil nur diese Mode langsamer als Licht in Vakuum propagiert. Diese Mode ist allerdings nur indirekt durch eine Vermischung mit der anderen Mode an normale Materie gekoppelt. Denn sie ist nicht direkt für die MOND-artige Kraft verantwortlich. Daher können in Zwei-Felder-SFDM zwar nicht-relativistische Objekte wie Sterne Energie in Form von Cherenkov-Strahlung verlieren, genauso wie in Standard-SFDM. Der Energieverlust ist aber viel kleiner, da die Kopplung an Materie unterdrückt ist, sodass die Zeitskala τ_E typischerwei-

se viel größer als galaktische Zeitskalen ist. Dieses Modell entgeht möglichen Einschränkungen also dadurch, dass die Verbindung zwischen galaktischen und kosmologischen Phänomenen geschwächt ist.

Auch im SZ-Modell gibt es eine masselose Mode, durch die Cherenkov-Strahlung möglich ist. Dieses Modell entgeht möglichen Einschränkungen durch einen neuen Mechanismus: Die Kopplung an Materie ist im statischen Limit eine normale gravitative Kopplung. In dynamischen Situationen ist die Kopplung aber unterdrückt. Nicht-relativistische Objekte wie Sterne geben also Cherenkov-Strahlung ab, verlieren aber nur auf Zeitskalen, die länger sind als das Alter des Universums, einen signifikanten Anteil ihrer Energie.

Zusammengefasst kann diese neue Art Cherenkov-Strahlung hybride MOND-DM-Modelle massiv einschränken (wie zum Beispiel Standard-SFDM), es gibt aber auch Mechanismen, um dies zu verhindern (wie zum Beispiel für Zwei-Felder-SFDM oder das SZ-Modell). Alle zukünftigen Hybrid-Modelle werden solche Tests bestehen müssen.

1. Introduction

Traditionally, various observations on cosmological and galactic scales are attributed to dark matter particles. Most notably, the anisotropies in the cosmic microwave background (CMB) and the rotation curves of galaxies. An alternative paradigm is modified gravity which aims to explain these phenomena in terms of a modified gravitational force law. In this section, we will explain why both dark matter and modified gravity on their own are not fully satisfactory. We argue that a combination of both paradigms is needed, specifically what we will refer to as hybrid MOND dark matter models.

The leading dark matter paradigm is called ΛCDM, where Λ denotes the cosmological constant and CDM denotes cold dark matter. It postulates the existence of dark matter particles that are non-relativistic throughout most of the history of the universe. These dark matter particles must interact sufficiently weakly with normal matter to evade direct detection constraints. On cosmological scales, they form a collisionless perfect fluid which gravitationally collapses to Navarro-Frenk-White (NFW) halos around galaxies. The CMB anisotropies are explained by the fluctuations of both the CDM fluid and normal matter. Galactic rotation curves are explained by the gravitational mass of the NFW halos and normal matter. Gravity is assumed to be that of General Relativity (GR). In practice, full GR is needed mainly for the homogeneous background cosmology, i.e. the Friedmann equations. Much of the phenomenology of NFW halos and CMB anisotropies can be derived from purely Newtonian gravity.

This is where an alternative paradigm comes into play. Namely Modified Newtonian Dynamics (MOND), a specific type of modified gravity model [1–4]. Consider a galaxy without dark matter particles. The acceleration due to Newtonian gravity $\vec{a}_N$ is usually dominated by the mass of the stars and the gas in the galaxy, i.e. by the baryonic mass, $\vec{a}_N = \vec{a}_b$. With spherical symmetry, $a_N = GM_b(r)/r^2$ where G is Newton's gravitational constant and $M_b(r)$ is the baryonic mass inside a sphere with radius r. The idea behind MOND is to modify the gravitational acceleration that objects in a galaxy feel, depending on whether a_b is larger or smaller than a critical acceleration scale $a_0 \approx 10^{-10}\,\mathrm{m/s^2}$. In the high-acceleration regime, $a_b \gg a_0$, the gravitational acceleration $\vec{a}_{\mathrm{tot}}$ is simply the Newtonian one,

$$\vec{a}_{\mathrm{tot}} = \vec{a}_b\,. \tag{1.1}$$

But in the low-acceleration regime, $a_b \ll a_0$, the gravitational acceleration is

enhanced,

$$\vec{a}_{\rm tot} = \vec{a}_b \sqrt{\frac{a_0}{|\vec{a}_b|}} \,. \tag{1.2}$$

That is,

$$a_{\rm tot} = \sqrt{a_0 a_b} \,, \tag{1.3}$$

so that the force scales as $\sqrt{M_b}$ rather than M_b and falls off as $1/r$ rather than $1/r^2$. In spherical symmetry,

$$a_{\rm tot} = \frac{\sqrt{a_0 G M_b(r)}}{r} \,. \tag{1.4}$$

This simple prescription captures many phenomena on galactic scales remarkably well, as we will discuss in more detail later. For example, it predicts that rotation curves will be asymptotically flat and that the asymptotic rotation velocity will be $v_\infty = (a_0 G M_b)^{1/4}$. In particular, all galaxies with the same total baryonic mass have the same asymptotic rotation velocity. This is different from ΛCDM where, in principle, various dark matter halo masses are possible for a given M_b. That is, galaxies with the same M_b can have different asymptotic rotation velocities in ΛCDM but not in MOND.

In fact, a similar statement is true at all radii, not only asymptotically. The total gravitational acceleration $\vec{a}_{\rm tot}$ in MOND is a function of the baryonic Newtonian acceleration $\vec{a}_b$, at least approximately. An interpolation function ν interpolates between the high-acceleration (Newtonian) regime and the low-acceleration (MOND) regime,

$$a_{\rm tot} = a_b \, \nu \left(\frac{a_b}{a_0} \right) \,, \tag{1.5}$$

where $\nu(y) \to 1$ for $y \to \infty$ and $\nu(y) \to 1/\sqrt{y}$ for $y \to 0$. As a consequence, galaxies with the same baryonic mass distribution have the same rotation curve at all radii. Again, this is different from ΛCDM where dark matter halos may differ between galaxies that have the same baryonic mass distribution. We will discuss this in more detail later.

This basic MOND phenomenology pertains to the gravitational force in the non-relativistic, weak-field limit. It is an alternative to the Newtonian limit of GR. As mentioned above, this suffices for many phenomenological purposes. But a complete model also needs to specify what happens in relativistic situations. A significant part of the theoretical work on MOND is about finding viable relativistic completions that give the above-mentioned MOND phenomenology in the non-relativistic limit.

In the following, we employ units with $c = \hbar = 1$ and the metric signature $(+,-,-,-)$, unless otherwise stated. Small Greek indices run from 0 to 3 and denote spacetime dimensions.

1.1. Why MOND is wrong

As mentioned above, MOND works remarkably well in galaxies. But this is not the case for groups of galaxies and galaxy clusters. This is shown in Fig. 1.1. Galaxies follow the MOND relation $v_c^4 = a_0 G M_b$ for the asymptotic circular velocity v_c. But groups and clusters of galaxies do not. More precisely, clusters and groups of galaxies still roughly satisfy the scaling relation $v_c^4 \propto M_b$. But the prefactor in this relation is different for galaxies and clusters of galaxies. In a logarithmic plot like Fig. 1.1, this means that clusters and galaxies lie on lines with the same slope but different intercept. This is a problem for MOND. It could mean that a_0 must depend on the scale of the system under consideration. It could mean that there are additional undetected baryons around galaxy clusters. But it could also mean that there is a deeper problem for the concept of MOND.

Another issue on cluster scales is the famous Bullet cluster [7–9]. The Bullet cluster consists of two galaxy clusters that have collided in the past. The visible gas from the two clusters is located relatively close to the collision point. In contrast, gravitational lensing reveals most of the gravitational mass to be located much further away from the collision point. This has a natural explanation in ΛCDM but not in MOND.

In ΛCDM, most of the gravitational mass comes from the dark matter particles. If the dark matter particles have weaker self-interactions than the cluster gas, the dark matter can easily be offset from the gas. This can explain the different locations of the visible gas and the gravitational mass. In contrast, the idea behind MOND is that the gravitational force is produced only by the baryons. So, naively, the lensing signal should come from the same location as the visible gas. This contradicts the Bullet cluster observations.

One loop-hole in this argument is that the inherent non-linearity of MOND may allow the inferred gravitational mass to not be located strictly at the position of the baryonic mass, at least for non-symmetric baryonic mass distributions. Indeed, MOND models can easily produce lensing peaks that are displaced from the peaks of the baryonic mass distribution. However, peaks outside rather than in between the baryonic peaks, as required for the Bullet cluster, were so far not shown to be possible. Of course, gravitational lensing is a relativistic phenomenon, while MOND is concerned only with the non-relativistic limit. Thus, in principle, the Bullet cluster must be considered separately in each fully-relativistic model that reduces to MOND in the non-relativistic limit. For example, such relativistic models often contain multiple fields. Possibly, some of these fields play a non-trivial role in non-equilibrium situations like colliding clusters. Still, the Bullet cluster is potentially a serious challenge to the MOND paradigm [10].

But even in galaxies, there are problems for MOND. One such problem is that dwarf spheroidals with Newtonian baryonic accelerations below about $10^{-12}\,\mathrm{m/s^2}$ do not seem to follow the MOND acceleration $\sqrt{a_0 a_b}$. The measured accelerations seem to be significantly larger [11]. Another problem is

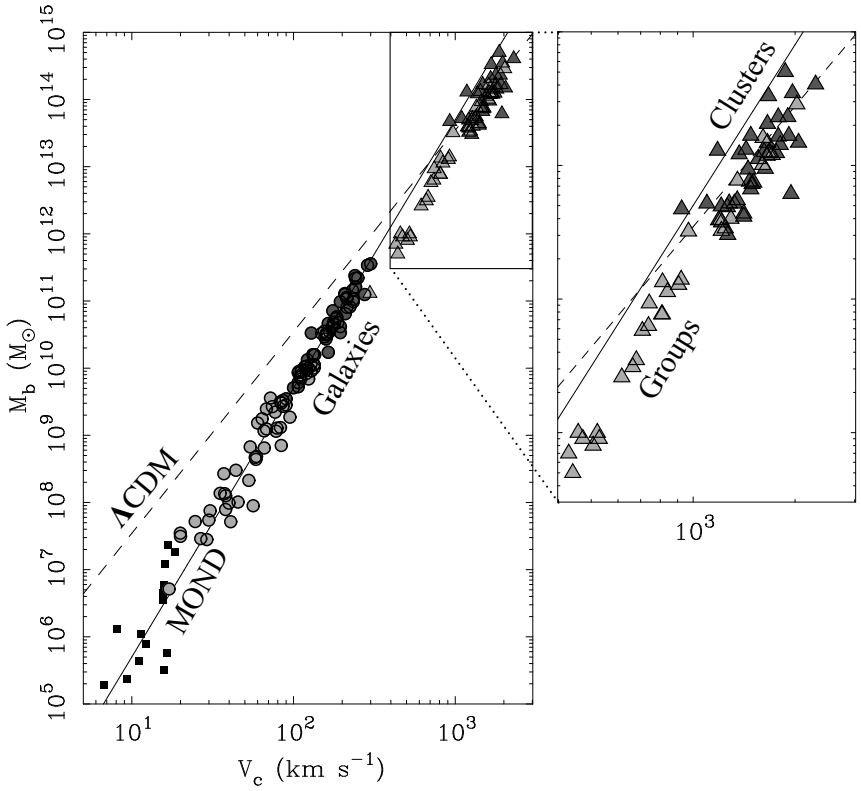

Figure 1.1.: The baryonic mass M_b as a function of the asymptotic circular velocity v_c. The solid black line shows the MOND prediction $v_c^4 = a_0 G M_b$. The dashed black line shows an a priori ΛCDM prediction from Ref. [5]. Measurements are included for dwarf spheroidals (squares), gas rich spiral galaxies (light gray circles), star-dominated spiral galaxies (dark gray circles), galaxy groups (light gray triangles), and galaxy clusters (dark gray triangles). Adapted from Ref. [6].

that the predicted vertical acceleration above the Milky Way disk may be too large in MOND [12, 13]. However, contrary to assumptions in these analyses, the Milky Way disc is not in equilibrium [14]. So this measurement should be taken with a grain of salt.

Another issue are the CMB anisotropies. Predicting the CMB temperature fluctuations again requires a relativistic completion of MOND. But so far no relativistic completion of MOND has achieved a completely satisfactory fit of the angular power spectrum of temperature fluctuations. At least not without introducing some kind of dark matter after all (see Sec. 1.3). The main problem is getting the relative amplitudes of the second and third peaks correct. The basic reason behind this can be understood as follows. If a_0 is constant in time, one may expect that MOND dynamics becomes important only at a redshift of order one [15]. At earlier times, relevant for the CMB, things may be quite similar to standard GR. That is, the CMB can be calculated as in GR but without any DM component. The angular power spectrum of the CMB temperature anisotropies for such a no-CDM model is shown in Fig. 1.2 [16]. The first two observed peaks are straightforward to fit in such a model, but not the third peak. This is because, without DM, baryon damping forces the third peak to be smaller than the second peak. But the measured third peak has about the same amplitude as the second peak. This is why the CMB is hard to get right in relativistic completions of MOND.

It is possible to avoid this argument in specific models. For example, it was shown that Bekenstein's TeVeS model from Ref. [17] can overcome this problem, but only at the cost of too large anisotropies on large angular scales [10, 18, 19].

More recently, gravitational wave events with optical counterparts have proven problematic for MOND. Specifically, GW170817 [21]. In this event, the gravitational tensor mode and its electromagnetic counterpart were detected roughly at the same time on Earth. In contrast, in many relativistic completions of MOND the photons experience a Shapiro time delay due to the MOND force. That is, photons are slowed down significantly compared to the tensor mode. This rules out many relativistic completions of MOND, especially those with a so-called disformal coupling of ordinary matter to the gravitational degrees of freedom [22, 23].

Another constraint for MOND comes from precision tests of the gravitational force in the solar system [10]. The non-relativistic limit of GR, including its post-Newtonian corrections, has so far passed all tests. Deviations from this are well-constrained. The post-Newtonian corrections in relativistic completions of MOND can be different in different models. So, in principle, the solar system constraints need to be checked separately for each relativistic completion. One approach to avoid constraints that is applicable in many relativistic completions of MOND is to introduce specific higher-derivative terms, see Ref. [24]. Another approach is called tracking, see for example Ref. [25]. The sun produces a Newtonian gravitational acceleration a_b that is many orders of magnitude larger than the MOND acceleration scale a_0.

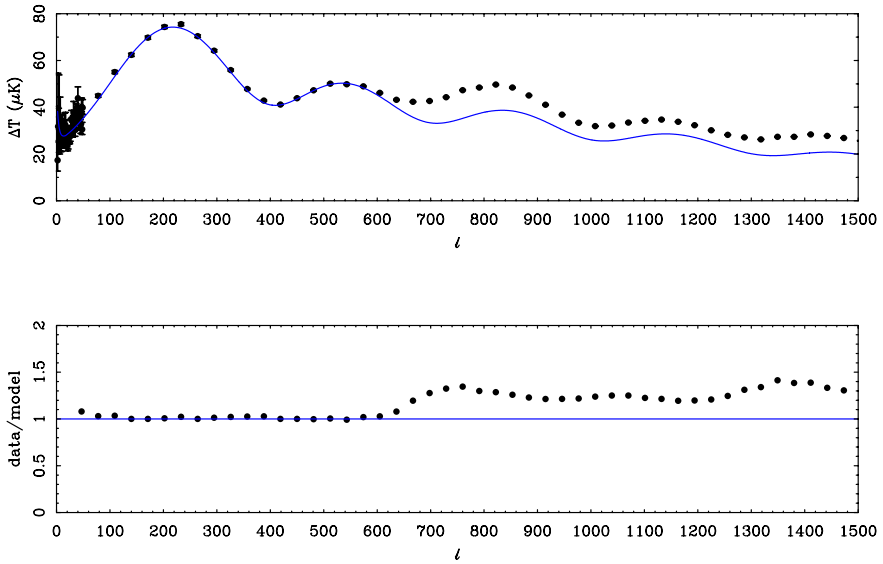

Figure 1.2.: Planck 2013 [20] measurements of the CMB temperature anisotropies (black points) and the no-CDM model from Ref. [16] normalized to fit the first peak (blue line). The no-CDM model serves as a naive toy model of what to expect in a relativistic completion of MOND. Adapted from Ref. [6].

Thus, one usually expects that solar system constraints require at least that the interpolation function ν approaches 1 quite rapidly for large accelerations $a_b \gg a_0$. Otherwise, already the leading-order non-relativistic limit of MOND gives too large deviations from Newtonian gravity in the solar system, even before taking into account post-Newtonian corrections.

1.2. Why ΛCDM is wrong

But not only MOND has problems. As mentioned above, a priori, galaxies with the same baryonic mass distribution can have different dark matter halos in ΛCDM. So one would expect a lot of scatter in any relation between the baryonic and dark matter mass distributions. For example, galaxies with the same baryonic mass distribution may have different rotation curves in ΛCDM, especially since the dark matter halo often dominates compared to the baryonic mass.

This expectation goes well with the visual impression one gets from a plot like Fig. 1.3, namely that rotation curves come in various shapes and sizes. However, it turns out that there is a very tight relation hidden in observed rotation curves. This can be seen if one plots the rotation curve data in the right way, namely not in terms of velocities and radii but in terms of accelerations. This is shown in Fig. 1.4, which shows a tight relation between the Newtonian baryonic acceleration a_b (or g_{bar}) and the observed total acceler-

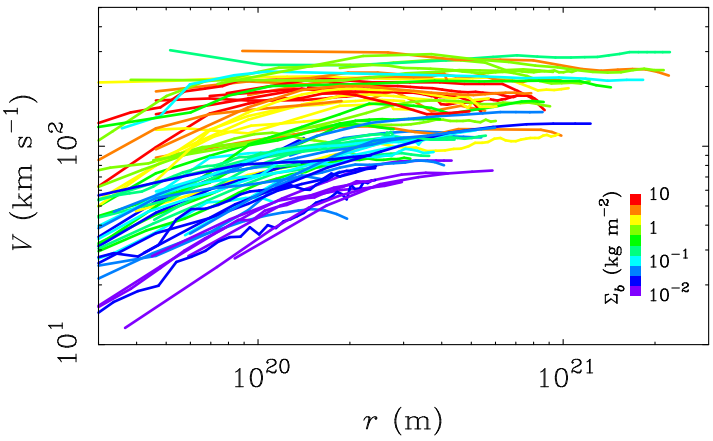

Figure 1.3.: Various observed rotation curves. Color encodes the characteristic baryonic surface density, i.e. $\Sigma_b = 0.75 M_b/R_b^2$ where R_b is the radius where the contribution of the baryons to the rotation curve peaks [10]. Adapted from Ref. [10].

ation a_{tot} (or g_{obs}). This is the Radial Acceleration Relation (RAR) [11]. In ΛCDM and with spherical symmetry, this is $G(M_b(r) + M_{\text{DM}}(r))/r^2$ plotted against $GM_b(r)/r^2$. The scatter in this relation is compatible with zero intrinsic scatter, i.e. only scatter from imperfect measurements [11].

This is exactly what one expects in MOND. But not, a priori, what one expects in ΛCDM. Why should it be possible to know the amount of dark matter at each radius just by knowing the baryonic mass distribution? Why should there be such a tight relation between the two? This is an unsolved question in ΛCDM.

In principle, it is possible that complicated baryonic processes during galaxy formation adjust the baryonic and dark matter in precisely the right way. Modern galaxy formation simulations include a host of empirical mechanisms which try to do this. These are often surprisingly successful. But often at the expense of other related relations, e.g. the so-called central density relation (see also below) [26–28]. One consequence of this is called the diversity problem [29, 30]: Rotation curves at a fixed small radius vary much more between different observed galaxies than between different simulated galaxies. Another concern is that empirical sub-grid models come with a lot of parameters that can be tuned to give various results and it's not clear whether they are actually physical.

Besides the RAR, there are various other scaling relations that are easily understood in MOND but not in ΛCDM. For example, the baryonic Tully-Fisher relation (BTFR) [32] which is the observed relation between the asymptotic rotation velocity and the total baryonic mass from Fig. 1.1. Another example is the central surface density relation (CSDR) mentioned in the previous paragraph. This is a relation between the baryonic central surface density

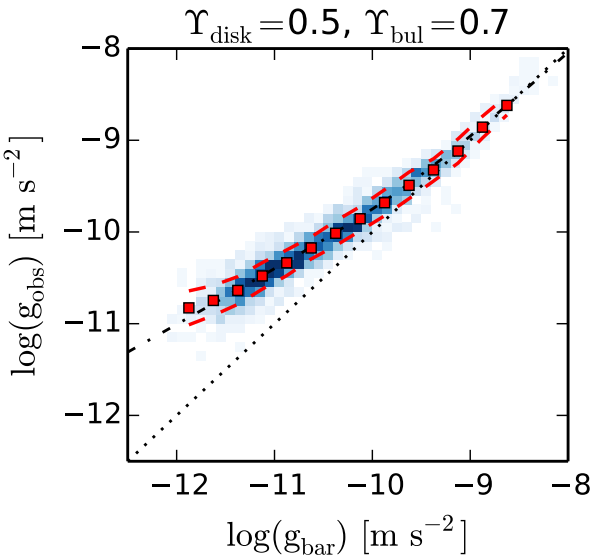

Figure 1.4.: The Radial Acceleration Relation. That is, the Newtonian baryonic acceleration $g_{\rm bar}$ versus the total acceleration inferred from observed rotation curves $g_{\rm obs}$ for SPARC [31] galaxies. This assumes the same stellar mass-to-light ratios for all galaxies. Specifically, $\Upsilon_{\rm disk} = 0.5$ and $\Upsilon_{\rm bulge} = 0.7$ as suggested by stellar population synthesis models. Adapted from Ref. [11].

$\Sigma_b(R = 0)$ and the total central surface density $\Sigma_{\rm dyn}(R = 0)$, including dark matter. This is shown in Fig. 1.5.

More qualitatively, Renzo's rule is difficult to understand in ΛCDM [33]. Renzo's rule corresponds to the observation that features in the rotation curve at a certain radius correspond to features in the baryonic mass distribution at the same radius. As long as the baryons' gravitational pull dominates, this is natural in ΛCDM. But, observationally, Renzo's rule holds even where dark matter dominates over the baryons. This is why Renzo's rule is surprising in ΛCDM.

Another surprising phenomenon are spiral galaxies with fast-rotating bars. Bars are a certain type of non-axisymmetric feature at the centers of galaxies. In ΛCDM, the dark matter halo slows down these bars due to a process called dynamical friction [34] (see also Chapter 4). As a result, bars in ΛCDM tend to rotate slower than in observed galaxies [35–37]. It may be possible to avoid this problem by having a larger stellar-to-dark matter ratio. But only at the cost of violating the widely-used abundance matching relations [38]. Fast bars are more natural in modified gravity models since the dark matter halo is absent [36].

Then there is the plane-of-satellites problem: Satellites of various galaxies, including the Milky Way and Andromeda, co-orbit in thin, planar structures. The problem is that such satellite planes are rare in ΛCDM [39, 40]. Note that

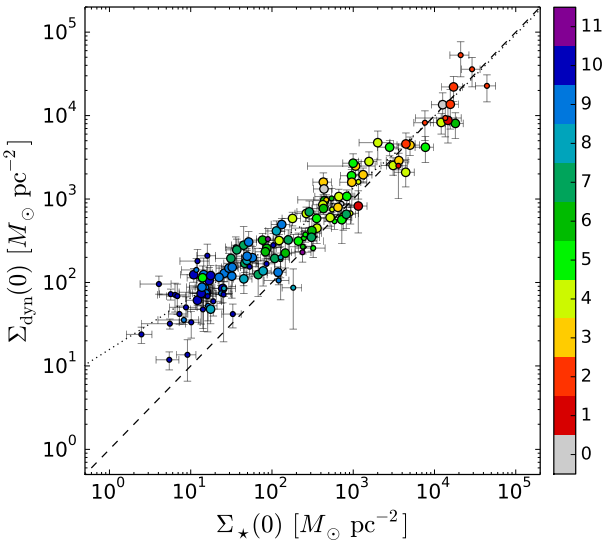

Figure 1.5.: The central surface density relation between the stellar central surface density $\Sigma_*(0)$ and the total central surface density $\Sigma_{\rm dyn}(0)$ for a subset of the SPARC galaxies [27, 31]. The baryonic central surface density $\Sigma_b(0)$ is expected to be close to $\Sigma_*(0)$ [27]. $\Sigma_{\rm dyn}(0)$ is inferred from the rotation curve of each galaxy. Color encodes the numerical Hubble type. Adapted from Ref. [27].

the problem is not just that satellites are spatially aggregated in a flattened structure, but also that they are kinematically correlated [41]. One possible explanation for the planes of satellites is that the satellites in these planes are tidal dwarfs that are correlated because they were created from the tidal tails of interacting galaxies. In this case, these satellite galaxies would not have their own dark matter halo due to their tidal nature. But, in ΛCDM, this contradicts the observed high internal velocity dispersions which imply that these satellites must be dark matter dominated. In contrast, tidal dwarfs can easily have high internal velocity dispersions in MOND, so it may be easier to explain the planes of satellites in MOND [39, 42].

There are also issues regarding structure formation at relatively high redshift, i.e. at relatively early times in the universe. In ΛCDM, structure forms hierarchically. Smaller objects form first. These smaller objects then combine to form bigger objects. There is a limit to how fast this process can form massive clusters. That is, a limit to how massive a cluster observed at high redshift can be in ΛCDM. Observations indicate that massive clusters form faster than what ΛCDM predicts, see e.g. Refs. [43–46]. A striking example is the El Gordo cluster [44]. This is a cluster with a mass of about $3 \cdot 10^{15}\, M_\odot$ at redshift $z = 0.87$. This is a problem for ΛCDM because such massive clusters do not form so early in ΛCDM [47], see also Fig. 1.6.[1] In contrast, structure

[1]Recently, Ref. [48] disputed this tension with ΛCDM. The most important difference

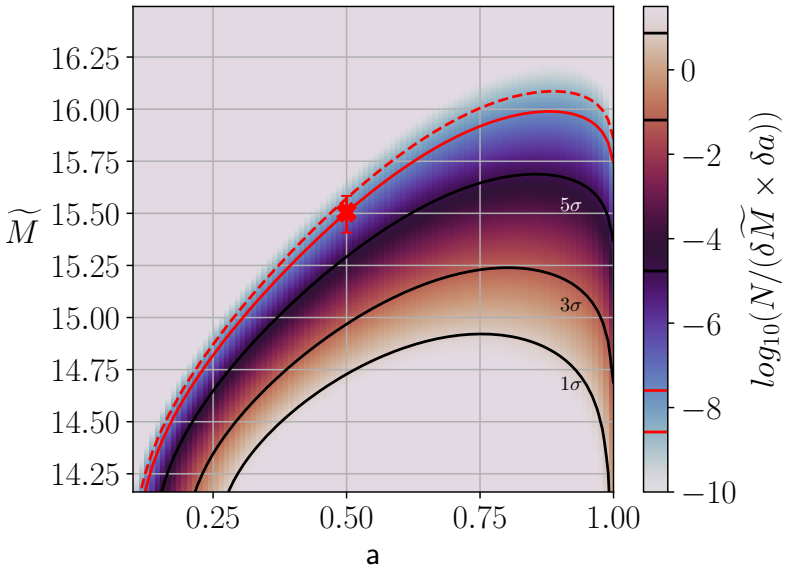

Figure 1.6.: The number density of clusters with a given mass and at a given redshift in a ΛCDM simulation. The mass is shown as the logarithm of the mass in solar units $\tilde{M}$. The redshift is shown as the cosmic scale factor $a = 1/(1+z)$. Contour lines indicate the probability to observe a cluster with given mass and redshift. The red cross shows the observed El Gordo cluster, which is very unlikely in ΛCDM. El Gordo lies on the solid red contour corresponding to $6.16\,\sigma$. Most of the probability of observing El Gordo comes from between the solid and dashed red contours. Adapted from Ref. [47].

is expected to form more rapidly in relativistic completions of MOND [15].

Regarding clusters, it should also be noted that the Bullet cluster cited in Sec. 1.1 as potential evidence against MOND is at best marginally consistent with ΛCDM. This is because the colliding subclusters have a high relative velocity that is rare in ΛCDM [50, 51]. For example, Ref. [51] estimates that, assuming ΛCDM, there is an about 1 in 10 chance to observe a Bullet-like cluster in a redshift-limited survey up to $z = 0.3$.

In ΛCDM, the Strong Equivalence Principle (SEP) holds. This is not true in MOND due to its non-linearity, which gives rise to the so-called external field effect (EFE) [4]. Recently, statistical evidence for an EFE-like violation

compared to Ref. [47] is that Ref. [48] adopts a much smaller value for the pre-merger infall velocity for the two subclusters of El Gordo. Running the code of Ref. [47] with such a small infall velocity makes the tension go away. But, as discussed in Ref. [49], this small infall velocity is derived without taking into account dynamical friction, although dynamical friction is likely to be important for El Gordo. Another issue is that Ref. [48] adopts a scenario where the subclusters are observed when coming into contact a second time. In this case, the pre-merger configuration in question has to be present even earlier than is assumed in Ref. [47], making the scenario more unlikely in ΛCDM.

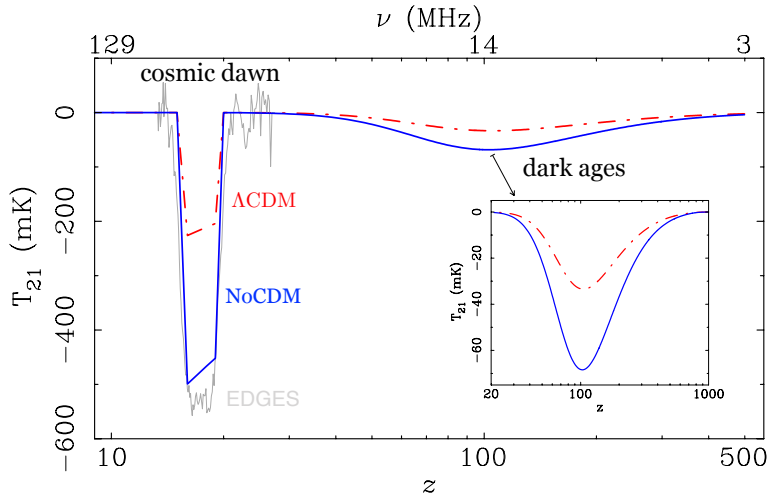

Figure 1.7.: The sky-averaged 21 cm absorption in terms of the brightness temperature T_{21} as a function of redshift. Shown are the ΛCDM prediction (dash-dotted red line), the no-CDM model from Ref. [16] (solid blue line), and the EDGES measurement (solid gray line). ΛCDM underpredicts the signal by roughly a factor of two. Adapted from Ref. [56].

of the SEP was found [52, 53]: There is a correlation between the asymptotic rotation curves in galaxies and the gravitational acceleration of their environment. A larger external acceleration tends to come with a (slightly) declining rotation curve. This is natural in MOND due to the EFE, but it is difficult to understand in ΛCDM (but see Ref. [54] which claims that so far ΛCDM and MOND cannot be observationally differentiated in this way).

Another unexpected result for ΛCDM is the recent measurement of the 21 cm absorption signal from redshift $z \approx 20$ by the EDGES experiment [55]. After recombination, but before most stars form, the universe consists mostly of neutral hydrogen and helium. During this time, CMB photons can be absorbed by the 21 cm spin-flip transition of the neutral hydrogen. The expected absorption signal in ΛCDM is only about half as strong as the signal measured by EDGES. In contrast to ΛCDM, a simple no-CDM model agrees with the measured absorption signal [56]. This is shown in Fig. 1.7.

At even earlier times, there is the Lithium problem, related to Big Bang Nucleosynthesis (BBN) [57]. BBN calculations depend sensitively on the cosmological baryon density parameter $\Omega_b h^2$, where $h = H_0/(100\,\mathrm{km\,s^{-1}Mpc^{-1}})$ with the Hubble constant H_0. The Lithium problem occurs when one imposes the CMB result for $\Omega_b h^2$ in BBN calculations. Then, BBN calculations predict about three times more primordial ^{7}Li than observed. Even without imposing the CMB value of Ω_b, there may be an internal discrepancy within BBN. Measured Lithium and Deuterium abundances require different values of Ω_b in BBN calculations. Deuterium requires the same value as the CMB, while

Lithium needs a lower value. Although it is curious that, before the CMB was measured, Deuterium measurements were in agreement with Lithium. Only after the CMB value of Ω_b became known did Lithium and Deuterium measurements diverge, see Ref. [58]. This issue around ^{7}Li is still unresolved.

Recently, the Hubble tension became pressing. Measurements of the Hubble constant H_0 at low redshifts systematically prefer a larger value than measurements at high redshifts. Most prominently, type-Ia supernovae with a Cepheid-calibrated distance ladder give $H_0 = (73.2 \pm 1.3)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [59], while the CMB gives $H_0 = (67.4 \pm 0.5)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [60]. This result has been confirmed by various other methods of measuring H_0. As just one example, the BTFR mentioned above can be used to independently measure H_0 [61]. Different measurements are quite consistent with each other at both high and low redshifts. This makes it challenging to build theoretical models explaining this tension within the ΛCDM paradigm, see for example Refs. [62, 63]. Thus, the Hubble tension is a major open problem in ΛCDM.

One possible explanation for the H_0 tension is that we live in a giant local void. This explanation is usually discarded because sufficiently large voids are unlikely to form in ΛCDM and because, assuming ΛCDM, supernova data is inconsistent with such a void [64]. This is despite various observational hints that such a local underdensity exists [65–70]. If this evidence does not go away, this may be another problem for ΛCDM. In contrast, both over- and underdensities are expected to grow quicker in MOND compared to ΛCDM.

1.3. Hybrid models

Above, we have seen that both MOND and ΛCDM have genuine successes and genuine problems. MOND is successful mostly on galactic scales, while ΛCDM is successful mostly on cosmological scales. A natural idea is then to combine these successes in a single model. We will refer to such models as hybrid MOND dark matter models, or just hybrid models for brevity.

Like ΛCDM, most hybrid models predict a collisionless fluid on cosmological scales, whose perturbations grow gravitationally. Like MOND, most hybrid models predict a gravitational acceleration $\sqrt{a_0 a_b}$ at small Newtonian accelerations a_b in galaxies. This is how hybrid models explain both the CMB and the rotation curves of galaxies, including scaling relations like the BTFR and the RAR.

The collisionless fluid usually forms some type of dark matter halo around galaxies. But this dark matter halo must provide a much smaller gravitational pull than in ΛCDM. Otherwise, rotation curves will be different than in MOND, contrary to one of the main motivations of hybrid models. Therefore, the collisionless fluid must grow much less structure on galactic scales than in ΛCDM. This is one generic feature of hybrid models.

A concrete example of a hybrid model is superfluid dark matter (SFDM) [71, 72]. This model introduces a new type of particle. On cosmological scales, particles of this type form a collisionless fluid as in ΛCDM. In the inner

parts of galaxies, in the so-called superfluid core, these particles condense to a superfluid, whose phonons then mediate a MOND-like force. The superfluid also acts as a dark matter halo around galaxies. But the halo profile is very cored, i.e. it is quite flat as a function of galactocentric radius. As a result, its gravitational pull in the superfluid core is subdominant compared to the MOND-like phonon force. The superfluid breaks down at larger radii so that the superfluid's constituent particles are in their non-condensed, CDM phase without the phonon force. At these larger radii, the gravitational pull of the superfluid becomes significant. We will discuss this model in more detail below. To distinguish the original SFDM model proposed in Ref. [71] from a modified two-field version of this model proposed in Ref. [73], we will sometimes refer to this original SFDM model as standard SFDM.

In SFDM, the cosmological and galactic scale phenomena share a common origin. Both are a consequence of the special type of superfluid. This is not the case in all hybrid models. In some models, the MOND-like force in galaxies and the collisionless fluid on cosmological scales are independent of each other. An example of such a model is the sterile neutrino based model (νHDM) from Ref. [74].

In νHDM, the collisionless fluid on cosmological scales is provided by an 11 eV sterile neutrino. This is essentially a form of hot dark matter (HDM), where the dark matter particles stay relativistic much longer than in CDM. Hot dark matter clusters much less on galactic scales compared to CDM. Just as needed for hybrid models, where the MOND-like force is supposed to dominate on galactic scales. On galaxy cluster scales, on the other hand, the sterile neutrinos do grow structure. This is important since a standard MOND-like force is not sufficient to account for the apparent missing mass on galaxy cluster scales. A significant dark matter halo on cluster scales is another generic feature of hybrid models.

A dark matter halo on cluster scales helps not only with the missing mass problem in clusters. It also helps with the Bullet cluster. Just as in ΛCDM, the lensing signal may be offset from the visible baryons due to the gravitational mass of the dark matter halo.

Hybrid models also have a generic advantage over standard MOND models regarding gravitational lensing. As discussed above, GW170817 requires the speed of light to be the same as that of gravitational tensor modes. This is a problem since photons often experience a Shapiro delay if they are affected by the field carrying the MOND force. A simple way to avoid this is to just not couple photons to this field. The most popular way to do this is to couple all matter to an effective metric $g^{\text{eff}}_{\alpha\beta}$,

$$g^{\text{eff}}_{\alpha\beta} = g_{\alpha\beta} \cdot f(\phi)\,, \tag{1.6}$$

where $g_{\alpha\beta}$ is the standard metric and f is a function of a field ϕ that mediates a MOND force. This type of coupling is called conformal coupling since the field ϕ enters in a conformal factor. Such a coupling does not affect photons since their equations of motion are conformally invariant.

Unfortunately, such a coupling is not possible in many MOND models. The lensing signal is too small if photons are not coupled to the field that mediates the MOND force. This is different in hybrid models, where there is dark matter around galaxies. As mentioned above, there must not be much dark matter in the inner parts of galaxies so that rotation curves are MOND-like. But the lensing signal comes from relatively large radii, where the dark matter halo may dominate. We discuss this in more detail in Sec. 2.1.

In addition to SFDM and νHDM, there are various other proposed hybrid models. Examples are dipolar dark matter [75], the dark fluid approach from Ref. [76], the two-field model from Ref. [77], and the nonminimal scalar-tensor model proposed in Ref. [78] (see also Ref. [79]). Verlinde's Emergent Gravity [80, 81] is another example, at least according to the interpretation of Ref. [82], which also provides a covariant version of Verlinde's model (Covariant Emergent Gravity, CEG).

Most recently, Skordis and Złośnik proposed a hybrid model that is based on both TeVeS and Einstein-Aether models [25, 83]. We will refer to this model as the SZ model. One advantage of this model is that it needs no ad-hoc prescriptions beyond its Lagrangian. This is in contrast to, for example, νHDM and SFDM which currently need such ad-hoc prescriptions for the transition from the MOND regime to the CDM regime. The SZ model has a MOND limit in galaxies, can successfully fit both the CMB and the matter power spectrum, and has a tensor mode that propagates with the speed of light. We will discuss this model in more detail in Sec. 4.5.

In the solar system, hybrid models face similar challenges as MOND. The MOND-like force and any relativistic post-Newtonian corrections are severely constrained by precision tests in the solar system. As in MOND, possible remedies include higher-derivative terms as proposed in Ref. [24]. But there may also be remedies specific to hybrid models. For example, Ref. [71] argues that the superfluid in SFDM may break down close to stars so that there would be no MOND-like phonon force around stars and therefore no solar system constraints on such a force.[2]

1.4. Superfluid dark matter

In the following chapters, we will often investigate properties of SFDM. To prepare for this, we will now introduce SFDM in more detail. As discussed above, the idea of SFDM is that there is a new type of particle which condenses to a superfluid on galactic scales where this superfluid's phonons then mediate a MOND-like force. On cosmological scales, these particles simply behave as cold dark matter.

[2] Though it is not clear how the superfluid can break down around stars while still having a MOND-like force in galaxies. After all, the stars are supposed to source the MOND-like phonon force and their motion is supposed to follow this MOND-like phonon force. How can this work if there are no phonons around stars because the superfluid breaks down?

In practice, the two main tools to make predictions in SFDM are a low-energy effective Lagrangian for the condensed superfluid on galactic scales and a prescription for the transition of this condensed phase to the non-condensed, CDM phase on larger scales. We will first introduce the Lagrangian for the condensed phase and explain its motivation. Then, we will discuss different prescriptions for the transition to the non-condensed phase.

Ref. [71] proposes the following low-energy non-relativistic effective Lagrangian for the phonon field θ,

$$\mathcal{L} = \frac{2\Lambda}{3}(2m)^{3/2}\sqrt{|X-\bar{\beta}Y|}X - \frac{\bar{\alpha}\Lambda}{M_{\rm Pl}}\rho_b\,\theta\,, \tag{1.7}$$

with

$$X = \dot{\theta} + \hat{\mu} - (\vec{\nabla}\theta)^2/(2m)\,, \quad Y = \dot{\theta} + \hat{\mu}\,, \quad \hat{\mu} = \mu_{\rm nr} - m\phi_{\rm N}\,. \tag{1.8}$$

Here, m is the mass of the superfluid constituent's particles, $\mu_{\rm nr}$ is the non-relativistic chemical potential, Λ is a constant with mass-dimension one related to self-interactions, $\bar{\alpha}$ is a dimensionless constant, $\bar{\beta}$ parametrizes finite-temperature effects, ϕ_N is the Newtonian gravitational potential, and ρ_b is the baryonic energy density. The non-relativistic chemical potential $\mu_{\rm nr}$ is related to the relativistic chemical potential μ by $\mu = m + \mu_{\rm nr}$ with $\mu_{\rm nr} \ll m$.

We will now explain how this Lagrangian is related to both MOND and superfluids. For the relation to MOND, the two relevant properties are the coupling to matter, proportional to $\theta\rho_b$, and the fact that spatial derivatives enter roughly as $X^{3/2}$. Indeed, the prototypical MOND-type Lagrangian for a gravitational potential φ reads [4, 79]

$$\mathcal{L} = \frac{2M_{\rm Pl}^2}{3a_0}\sqrt{|K_\varphi|}K_\varphi - \rho_b\,\varphi\,, \tag{1.9}$$

where $K_\varphi = \nabla_\alpha\varphi\nabla^\alpha\varphi$ is the standard kinetic term of φ and a_0 is the MOND acceleration scale. This is the form of the Lagrangian in the deep-MOND regime $K_\varphi \ll a_0^2$. In general, the $K_\varphi^{3/2}$ term is replaced by a function $F(K_\varphi)$ with $F(K) \approx K$ for $K \gg a_0^2$ and $F(K) \approx (2/3a_0)K\sqrt{|K|}$ for $K \ll a_0^2$. The form of F determines the form of the interpolation function ν which interpolates between the MOND and the Newtonian regimes. In the static limit, the equation of motion of φ reads

$$\vec{\nabla}\left(F'\left(|\vec{\nabla}\varphi|^2\right)\vec{\nabla}\varphi\right) = 4\pi G\rho_b\,, \tag{1.10}$$

where $F'(K)$ is the derivative of $F(K)$ with respect to K. The acceleration of matter due to this gravitational potential follows from the coupling term $-\rho_b\varphi$ and is

$$\vec{a}_\varphi = -\vec{\nabla}\varphi\,. \tag{1.11}$$

The right-hand side of Eq. (1.10) can be written in terms of the Newtonian acceleration $\vec{a}_b$ as $-\vec{\nabla}\vec{a}_b$. Thus, up to a term that is the curl of a vector field,

$$F'\left(a_\varphi^2\right)\vec{a}_\varphi = \vec{a}_b\,. \tag{1.12}$$

This is called the no-curl approximation. It is exact in spherical symmetry and is a good approximation in many other situations [84, 85]. In the MOND regime $a_b \ll a_0$, this gives the standard MOND result

$$a_\varphi = \sqrt{a_0 a_b}\,. \tag{1.13}$$

It is now straightforward to see that SFDM has a MOND limit with

$$a_0 = \frac{\bar{\alpha}^3\Lambda^2}{M_{\rm Pl}}\,, \tag{1.14}$$

for[3]

$$(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}\,, \tag{1.15}$$

i.e. when spatial derivatives dominate in the kinetic term. In this case, SFDM reproduces the prototypical MOND Lagrangian Eq. (1.9) up to a rescaling. In SFDM, the phonon force is in addition to that of the Newtonian gravitational potential. Thus, the total acceleration is

$$\vec{a}_{\rm tot} = \vec{a}_\theta + \vec{a}_N\,, \tag{1.16}$$

where $\vec{a}_\theta = -(\bar{\alpha}\Lambda/M_{\rm Pl})\vec{\nabla}\theta$ is the acceleration due to the phonon force and $\vec{a}_N$ is that due to the Newtonian potential. To strictly reproduce MOND, $\vec{a}_N$ should be the same as $\vec{a}_b$. But this is not necessarily true in SFDM since the superfluid itself contributes to the gravitational mass felt by the Newtonian potential,

$$\rho_{\rm SF} = \frac{2\sqrt{2}}{3}m^{5/2}\Lambda\,\frac{3(\bar{\beta}-1)\hat{\mu}+(3-\bar{\beta})\frac{(\vec{\nabla}\theta)^2}{2m}}{\sqrt{(\bar{\beta}-1)\hat{\mu}+\frac{(\vec{\nabla}\theta)^2}{2m}}}\,. \tag{1.17}$$

In a useful MOND limit, the Newtonian gravitational pull due to the superfluid should be subdominant so that $\vec{a}_N \approx \vec{a}_b$, which is usually the case in the inner parts of galaxies. This is illustrated in Fig. 1.8, which shows an example rotation curve and its individual contributions in SFDM. The baryonic Newtonian acceleration and the MOND-like phonon force dominate the superfluid's Newtonian gravitational pull. The superfluid's Newtonian gravitational pull becomes important only at larger radii. At smaller radii, the total acceleration felt by matter is roughly $a_{\rm tot} = a_b + \sqrt{a_0 a_b}$ which corresponds to MOND with an interpolation function

$$\nu(y) = 1 + \frac{1}{\sqrt{y}}\,. \tag{1.18}$$

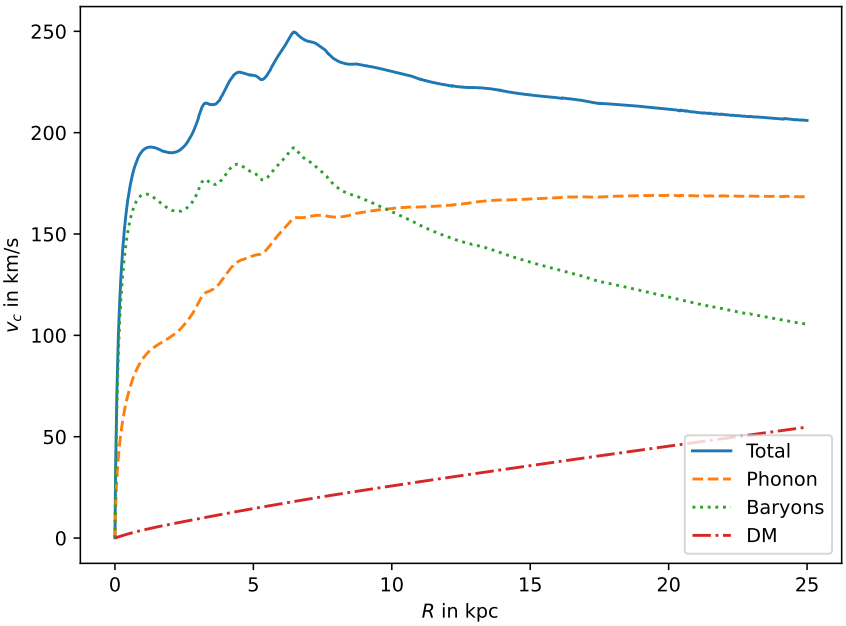

Figure 1.8.: The rotation curve v_c (solid blue line) of the Milky Way model discussed in Sec. 2.2 in SFDM. This rotation curve can be obtained by adding three individual rotation curve components in quadrature. These three components correspond to the Newtonian baryonic acceleration (dotted green line), the MOND-like phonon force (dashed orange line), and the Newtonian acceleration from the superfluid's mass (dash-dotted red line). The superfluid's gravitational pull is subdominant compared to the other contributions. The MOND-like phonon force is responsible for the flat rotation curve at large radii.

This is how SFDM is related to MOND.

The relation of SFDM to superfluids is best illustrated with a toy model. Consider a complex scalar field $\phi = \rho e^{-i\theta}/\sqrt{2}$ with a standard Lagrangian with a sextic interaction

$$\mathcal{L} = \frac{1}{2}\nabla_\alpha\rho\,\nabla^\alpha\rho + \frac{1}{2}\rho^2\,\nabla_\alpha\theta\,\nabla^\alpha\theta - \frac{1}{2}m^2\rho^2 - \frac{1}{6}\lambda_6\rho^6\,, \tag{1.19}$$

with some constant $\lambda_6 > 0$. For simplicity, we assume a flat Minkowski background spacetime. This is a standard superfluid-type $U(1)$-symmetric Lagrangian [86]. The superfluid phase corresponds to the case where this $U(1)$ symmetry is spontaneously broken. In this case, the ground state has non-zero particle number. The Goldstone mode associated with this broken symmetry corresponds to the superfluid's phonons. The linear dispersion relation of these phonons is responsible for the frictionless flow that is typical for superfluids [86].

The $U(1)$ symmetry is associated with a conserved charge Q. Thus, in statistical physics, we would replace the standard Hamiltonian H with the shifted Hamiltonian $H - \mu Q$ with chemical potential μ. On the level of the Lagrangian, this change of the Hamiltonian corresponds to a shift in time derivatives of θ, namely $\dot\theta \to \dot\theta + \mu$ [87–90]. Equivalently, at least in equilibrium at zero temperature, one can directly consider solutions $\theta = \mu \cdot t$. However, introducing μ as described above makes it clear that μ is a chemical potential in the statistical physics sense.

Including the chemical potential μ, the effective potential for ρ is

$$V_{\text{eff}} = \frac{1}{2}\rho^2\left(m^2 + (\vec\nabla\theta)^2 - (\dot\theta+\mu)^2\right) + \frac{1}{6}\lambda_6\rho^6\,. \tag{1.20}$$

This potential has its minimum at a non-zero value of ρ, if $\mu > m$ and if derivatives of θ are not too large. That is, V_{eff} becomes a Mexican-hat-type potential for $\mu > m$. Then, the non-zero minimum of $V_{\text{eff}}(\rho)$ approximately solves the ρ equation of motion. At least as long as derivatives of ρ can be neglected. This solution describes a superfluid condensate that spontaneously breaks the $U(1)$ symmetry,

$$\sqrt{\lambda_6}\,\rho^2 = \sqrt{(\dot\theta+\mu)^2 - (\vec\nabla\theta)^2 - m^2}\,. \tag{1.21}$$

If we plug this solution back into the Lagrangian, we get a low-energy effective Lagrangian for the phonons. Neglecting derivatives of ρ,

$$\mathcal{L}_{\text{eff},\theta} = \frac{1}{3}\frac{1}{\sqrt{\lambda_6}}\left((\dot\theta+\mu)^2 - (\vec\nabla\theta)^2 - m^2\right)^{3/2}\,. \tag{1.22}$$

Up to the baryon coupling, this is very similar to the SFDM Lagrangian from Eq. (1.7). This shows the connection between SFDM and standard superfluid models. The Lagrangian Eq. (1.7) should be understood as the low-energy

[3] This assumes $\bar\beta$ to be of order 1. Otherwise, the MOND limit is reached for $(\vec\nabla\theta)^2 \gg 2m\hat\mu\bar\beta$.

effective theory for the phonons of a more complete superfluid model that is valid also at higher energies.

Our toy model with sextic interactions nicely illustrates the ideas behind SFDM. But it cannot be a realistic SFDM model. The reason is that spatial derivatives are supposed to dominate in SFDM in order to give a useful MOND limit, while the solution Eq. (1.21) exists only when spatial derivatives do not dominate. To fix this, the sign of λ_6 could be reversed. But this leads to an instability, as was first discussed by Bekenstein [91, 92]. Actual relativistic completions of the Lagrangian Eq. (1.7) must be more complicated [71].

This complication in getting the Lagrangian Eq. (1.7) from a fully relativistic completion pertains to its kinetic term. But the coupling term $-(\bar{\alpha}\Lambda/M_{\rm Pl})\theta\rho_b$ also does not immediately follow from any standard superfluid models. This is because it explicitly breaks the $U(1)$ symmetry that is typical of superfluid Lagrangians. We will discuss this coupling term in more detail in Sec. 3.3. For now, we will take the effective low-energy Lagrangian Eq. (1.7) as given. This is sufficient to make predictions on galactic scales in many situations.

In our toy model, condensation happens when the chemical potential is sufficiently large, $\mu > m$. In the non-relativistic limit and including the Newtonian gravitational potential in the standard way, this condition $\mu > m$ becomes $\hat{\mu} > 0$. Here, we define $\hat{\mu} = \mu_{\rm nr} - m\phi_N$ as in SFDM, see Eq. (1.8). What does this condition $\hat{\mu} > 0$ correspond to for SFDM with the Lagrangian Eq. (1.7)? The original SFDM proposal Ref. [71] assumes the condition $\hat{\mu} > 0$ to hold in the superfluid phase. In our toy model, this condition is necessary for equilibrium solutions. Otherwise, the effective Lagrangian Eq. (1.22) becomes imaginary. But this is not necessarily the case for SFDM with the Lagrangian Eq. (1.7). So, in principle, one might allow (moderately) negative values of $\hat{\mu}$. Here, we will not do this and instead follow Ref. [71] in assuming $\hat{\mu} > 0$. The effects of allowing negative $\hat{\mu}$ will be explored elsewhere.

As mentioned above, the second important ingredient for predictions in galaxies in SFDM is a prescription for the transition from the condensed to the non-condensed phase. Often, predictions at small galactic radii will be possible from the effective Lagrangian Eq. (1.7) alone. Modeling the transition to the non-condensed phase is important for predictions at larger radii and for estimating which radii count as small and which as large in this sense.

Refs. [71, 72] make two specific proposals for this transition. Both postulate that the superfluid phase is present only in the inner parts of galaxies and ends at some finite transition radius. The non-condensed phase at radii larger than this transition radius is described by a standard NFW halo. The transition radii from the two specific proposals from Refs. [71, 72] are called the thermal radius R_T and the NFW radius $R_{\rm NFW}$, respectively. Both prescriptions should be understood as rough estimates. Indeed, we argued in Ref. [73] that these estimates likely need to be revisited, see also Sec. 3.4. Ref. [93] refines the arguments of Ref. [72], but in a different context, i.e. a superfluid model without the MOND-like phonon force. In what follows, we adopt the original proposals from Ref. [72] for simplicity.

The thermal radius R_T is determined by [72]

$$\Gamma = t_{\mathrm{dyn}}^{-1}\,, \tag{1.23}$$

where Γ is the local self-interaction rate and t_{dyn} is the galactic dynamical time. The idea is that the superfluid is in equilibrium at radii smaller than R_T but does not interact sufficiently to reach equilibrium beyond R_T. The interaction rate is given by $\Gamma = (\sigma/m)\mathcal{N}v\,\rho_{\mathrm{SF}}$ with the Bose enhancement factor $\mathcal{N} = (\rho_{\mathrm{SF}}/m)(2\pi/mv)^3$, the self-interaction cross-section σ, and the average velocity v of the superfluid constituent particles. A simple estimate for the dynamical time is $t_{\mathrm{dyn}} = r/v$ with the spherical galactocentric radius r. A simple estimate for v is the Newtonian circular velocity $\sqrt{r\,a_N}$. In practice, a weakness of the thermal radius is that the cross-section σ has not yet been calculated from first principles. Instead, Ref. [72] simply assumes a value for σ. This is one reason why these estimates should be revisited [73].

The NFW radius R_{NFW} is defined as the radius where both the superfluid's density and pressure can be matched to that of an NFW halo [72]. These are two conditions for three variables: The NFW radius R_{NFW} and the two free parameters ρ_c and r_s of the NFW profile ρ_{NFW},

$$\rho_{\mathrm{NFW}} = \frac{\rho_c}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}\,. \tag{1.24}$$

One approach to fix the remaining free variable is to match only the $1/r^3$ tail of the NFW profile which depends only on one parameter. Another approach is to require the scale radius r_s to be close to what it would be in ΛCDM given the total halo mass [72, 94].

Both the thermal radius and the NFW radius as defined in Ref. [72] assume spherical symmetry. For many real galaxies, assuming axisymmetry is much more realistic than assuming spherical symmetry. However, we have shown in Ref. [85] that these estimates can be straightforwardly extended to axisymmetric situations. The reason is basically that the transition happens at relatively large radii where the superfluid may be approximately spherically symmetric even if the baryonic mass distribution is not.

2. Observational tests of superfluid dark matter

In this chapter, we present two specific observational tests of SFDM. The first regards strong lensing. Lensing is a challenge for various MOND models because the detection of GW170817 and its electromagnetic counterpart implies that photons do not experience a significant Shapiro time delay compared to gravitational tensor modes. We will demonstrate explicitly that SFDM can simultaneously fit Einstein radii and velocity dispersions for a sample of lensing galaxies. This first test, described in Sec. 2.1, summarizes our work in collaboration with Sabine Hossenfelder from Ref. [95]. The second test is the Milky Way rotation curve. Since this is the first discussion of a rotation curve of an axisymmetric galaxy in SFDM, we will first discuss how to treat axisymmetric systems in SFDM. We will then show that SFDM provides a reasonable fit for the Milky Way rotation curve, requiring about 20% less baryonic mass than standard MOND models. This second test, described in Sec. 2.2, summarizes our work in collaboration with Sabine Hossenfelder from Ref. [85].

2.1. Strong lensing

As discussed above, the detection of GW170817 and its electromagnetic counterpart made it more difficult for various MOND models to explain lensing data. This is because, in many MOND models, photons are affected by the field carrying the MOND force while the gravitational tensor mode is not. In such models, the photons of the electromagnetic counterpart of GW170817 experience a large Shapiro time delay, which is ruled out because the gravitational wave and its electromagnetic counterpart were detected almost simultaneously on Earth.

One way to avoid this problem is to just not couple photons to the field carrying the MOND force. But this reduces the lensing signal. In many models, the lensing signal will then be the same as in GR with just the baryons and no dark matter. This is in conflict with observations.

Hybrid models like SFDM can avoid this conclusion because they have a dark matter halo on galactic scales. This dark matter halo may produce enough lensing to fit observations. One might worry that rotation curve velocities then come out too large since stars feel both the MOND-like phonon force and the superfluid halo's Newtonian gravitational pull. But this is not necessarily a problem since the superfluid halo is very cored. Rotation curves are measured at relatively small radii where the phonon force dominates, while

much of the lensing signal comes from relatively large radii where the dark matter halo dominates. In this section, we quantitatively investigate this mechanism in SFDM.

In ΛCDM, the same gravitational mass M is inferred both from lensing measurements and from kinematic measurements, like velocity dispersion or rotation curve measurements. This is different in SFDM. Kinematic measurements inside the superfluid core receive contributions from the MOND-like phonon force. In contrast, lensing signals are independent of this phonon force. Thus, if SFDM is true and if we infer M in the same way as in ΛCDM, we will, in general, get different results from kinematic measurements and from lensing measurements.

It would be interesting to directly compare the mass M inferred from kinematic measurements to that inferred from lensing measurements. But this is not possible with the data we will be using below. Basically, the reason is that M is always the mass inside a specific volume, for example the mass $M(r)$ inside a sphere with radius r. We cannot get masses for the same volume from both strong lensing and kinematic measurements for two reasons. First, strong lensing measurements give the mass within a cylinder with a given radius along the line of sight, while kinematic measurements depend on the mass within a spherical volume. Second, even if we were to get the mass within a sphere from lensing, we couldn't compare it directly to kinematic measurements. The reason is that, for the galaxies considered here [96, 97], only averaged velocity dispersions are available. We cannot resolve a radial dependence.

Still, we can check whether the strong lensing Einstein radii and the velocity dispersions can be fit at the same time in SFDM. Fitting both ensures that, for example, we do not fit the strong lensing data at the cost of completely unrealistic stellar kinematics.

We consider velocity dispersions and Einstein radii of 65 lenses from the Sloan Lens ACS (SLACS) Survey. These lenses are classified as ellipticals. They have velocity dispersions from SDSS measurements and complete photometric data from Hubble Space Telescope (HST) observations. We take redshifts, Einstein radii, effective radii, and velocity dispersions from Ref. [97] and the aperture radius and seeing from Ref. [96]. This set of lenses was previously studied in Ref. [98] in the context of a phenomenological MOND model. Here, we study these lenses in the context of SFDM.

2.1.1. The model

Following Ref. [98] we assume a spherically symmetric Jaffe mass distribution for the baryons of the lensing galaxies [99],

$$\rho_b(r) = \frac{M_b}{4\pi R_\mathrm{J}} \frac{1}{r^2(1+r/R_\mathrm{J})^2} \,. \tag{2.1}$$

Here, M_b is the total baryonic mass and $R_\mathrm{J} = 1.31\, R_\mathrm{eff}$ is derived from the effective radius $R_\mathrm{eff} = \Theta_\mathrm{eff} \cdot D_\mathrm{l}$ with the measured angular effective radius

$\Theta_{\rm eff}$ and the measured angular distance of the lens $D_{\rm l}$ [99]. Thus, M_b is the only free parameter of the baryonic mass distribution. This is one of two free parameters we will use to fit the data.

To solve for the phonon field θ, we use the no-curl approximation discussed in Sec. 1.4. Since we assume spherical symmetry, this approximation is exact. This gives an algebraic expression for $\vec{\nabla}\theta$ in terms of $\vec{a}_b$ and $\hat{\mu}$. For the SFDM model parameters we adopt the fiducial values from Ref. [72]. That is, $\bar{\alpha} = 5.7$, $m = 1\,\text{eV}$, $\Lambda = 0.05\,\text{meV}$, and $\bar{\beta} = 2$.

The Poisson equation for the Newtonian gravitational potential ϕ_N can be rewritten as an equation for the quantity $\hat{\mu} = \mu_{\rm nr} - m\phi_N$,

$$\Delta\left(-\frac{\hat{\mu}}{m}\right) = 4\pi G\left[\rho_b + \rho_{\rm SF}(\vec{\nabla}\theta, \hat{\mu})\right] . \tag{2.2}$$

Using the no-curl approximation for $\vec{\nabla}\theta$ gives $\rho_{\rm SF}$ as a function of $\vec{a}_b$ and $\hat{\mu}$. We numerically solve this equation using Mathematica [100] with values for $\hat{\mu}(r_0)$ and $\hat{\mu}'(r_0)$ at $r_0 = 0.01\,\text{kpc}$ as boundary conditions. The value $\hat{\mu}(r_0)$ is the second free parameter in our fit. It determines the size of the superfluid halo. To avoid numerical issues at $r = 0$, we do not impose $\hat{\mu}'(0) = 0$ but rather $-\hat{\mu}'(r_0)/m = 4\pi G r_0(\rho_b(r_0) + \rho_{\rm SF}(r_0))$. This can be obtained by expanding $\hat{\mu}$ for small r using the assumed boundary condition value of $\hat{\mu}(r_0)$.

Following Refs. [71, 72] we assume that outside the superfluid phase the dark matter energy density is that of an NFW halo. For simplicity, we adopt the thermal radius R_T for the transition radius and we approximate the NFW profile as $\rho_{\rm NFW} \propto 1/r^3$. The thermal radius depends on the quantity σ/m. We adopt the fiducial value $\sigma/m = 0.01\,\text{cm}^2/\text{g}$ from Ref. [72]. We have checked that our results do not depend strongly on this choice by redoing the calculation also with the NFW radius $R_{\rm FNW}$ as the transition radius and matching to a full NFW profile [95].

The angular Einstein radius $\Theta_{\rm E}$ can be calculated from [101]

$$\Theta_{\rm E}^2 = \frac{D_{\rm ls}}{D_{\rm l} D_{\rm s}}\, 4\, G\, M_{\rm E}(R_{\rm E})\,, \tag{2.3a}$$

$$R_{\rm E} = D_{\rm l}\,\Theta_{\rm E}\,, \tag{2.3b}$$

where $D_{\rm l}$, $D_{\rm s}$, and $D_{\rm ls}$ are the angular distances of the lens, the source, and the angular distance between the source and the lens, respectively. Further, $M_{\rm E}(R_{\rm E})$ denotes the gravitational mass inside the cylinder with radius $R_{\rm E}$ along the line of sight.

The radial velocity dispersion σ_r is usually calculated by [101]

$$\sigma_r^2(r) = \frac{G\int_r^\infty dr' \rho_b(r') M_\sigma(r') (r')^{2\beta-2}}{r^{2\beta}\rho_b(r)}\,, \tag{2.4}$$

with the anisotropy parameter β. Here, GM_σ/r^2 is the acceleration felt by the baryons. In our case, this acceleration includes both the acceleration

from the Newtonian potential and that from the MOND-like phonon force, $GM_\sigma/r^2 = a_N + a_\theta$. The measured velocity dispersion σ_* is then

$$\sigma_*^2 = \frac{\int_0^\infty dR\, R\, w(R) \int_{-\infty}^\infty dz\, \rho_b(r) \left(1 - \beta \frac{R^2}{r^2}\right) \sigma_r^2(r)}{\int_0^\infty dR\, R\, w(R) \int_{-\infty}^\infty dz\, \rho_b(r)}\,, \tag{2.5a}$$

$$w(R) = e^{-R^2/2\tilde{R}_{\rm atm}^2}, \tag{2.5b}$$

$$\tilde{R}_{\rm atm} = \tilde{\sigma}_{\rm atm}\, D_{\rm l}\,, \tag{2.5c}$$

$$\tilde{\sigma}_{\rm atm} = \sigma_{\rm atm}\sqrt{1 + \theta_{\rm ap}^2/4 + \theta_{\rm ap}^4/40}\,, \tag{2.5d}$$

where $r = \sqrt{R^2 + z^2}$, $\sigma_{\rm atm} = 1.4''$ is the seeing, and $\theta_{\rm ap} = 1.5''$ is the spectrometric aperture [96]. For simplicity, we take $\beta = 0$.

Angular distances D are calculated from the measured redshifts $z_{\rm l}$ and $z_{\rm s}$ of the lens and the source, respectively. Specifically, we use

$$D(z_1, z_2) = \frac{1}{H(1+z_2)} \int_{z_1}^{z_2} \frac{dz'}{\sqrt{\Omega_{\rm m}(1+z')^3 + (1-\Omega_{\rm m})}}\,, \tag{2.6}$$

where z_1 and z_2 are the redshifts of the objects whose angular distance is to be calculated. We adopt $H = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_{\rm m} = 0.3$ following Ref. [97].

Our fitting procedure is a relatively simple parameter scan of $\hat{\mu}(r_0)$ and M_b. We force the measured and calculated Einstein radii to agree to at least 0.01 kpc and the measured and calculated velocity dispersions to agree within the measurement error $\sigma_*^{\rm error}$.

2.1.2. Results

With this method, we find a successful fit for 64 of the 65 galaxies. That is, the calculated and measured Einstein radii agree to better than 0.01 kpc and the calculated and measured velocity dispersions $\sigma_*^{\rm meas}$ and $\sigma_*^{\rm calc}$ agree to better than the measurement error $\sigma_*^{\rm error}$. This is shown in Fig. 2.1, left.

Since we adjusted the value of M_b, we must check that our fits do not require unphysical stellar mass-to-light ratios. We find an averaged $M/L_{\rm V}$ of 3.5 ± 1.1 compared to the 4.2 ± 1.0 obtained in Ref. [98]. This is somewhat lower but not unphysical, see also Fig. 2.1, right. Indeed, our fit results fall between those predicted by stellar population synthesis models with Salpeter and Chabrier initial mass functions (IMF). This is shown in Fig. 2.2. Thus, our fitted stellar masses are reasonable.

Strong lensing in SFDM can be qualitatively understood from Fig. 2.3. The left panel shows the baryonic and dark matter mass contained in a sphere as a function of its radius r using the galaxy J0029-0055 as an example. The velocity dispersion σ_* is mainly determined by the baryonic and dark matter masses within small spherical radii of a few kpc. Baryonic matter dominates over the superfluid halo at these radii. This is how SFDM can fit rotation curves and velocity dispersions in a way that is relatively independent of the superfluid halo.

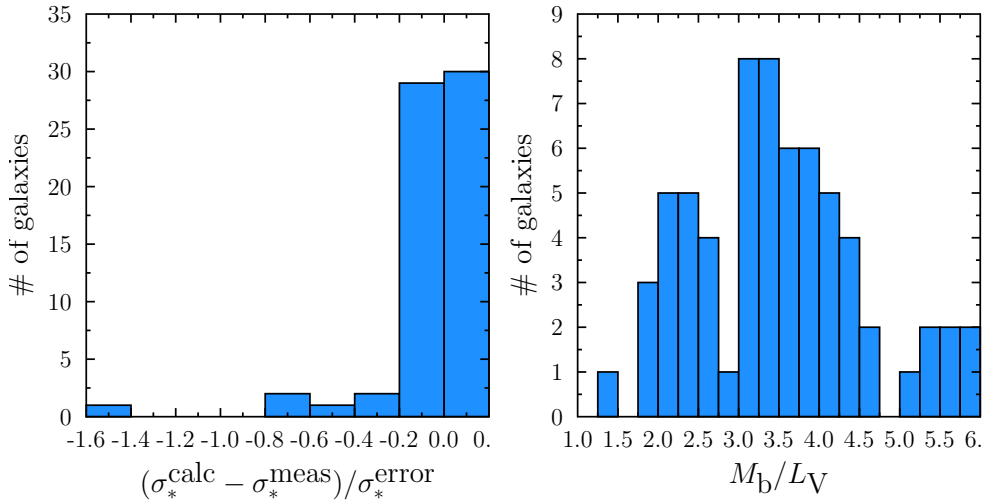

Figure 2.1.: Left: Histogram of the best-fit $(\sigma_*^{\rm calc} - \sigma_*^{\rm meas})/\sigma_*^{\rm error}$ for each galaxy. We find a successful fit for 64 out of 65 galaxies. Right: Histogram of the calculated stellar mass-to-light-ratios for M_b and $\hat{\mu}(r_0)$ such that $|\sigma_*^{\rm calc} - \sigma_*^{\rm meas}|$ is minimal. This shows that our fits do not require unphysical M/L.

In contrast, the lensing signal is sensitive to the mass inside a cylinder with radius of a few kpc, not a sphere. The baryonic mass is still the largest contribution at relatively small cylindrical radii. But now the contribution of the superfluid halo is significantly larger than for a sphere with a comparable radius. This is shown in the right panel of Fig. 2.3. The lensing signal is much more sensitive to the superfluid's mass than the velocity dispersion. Thus, even if we force our parameters to fit the velocity dispersion, there is always room to significantly adjust the lensing signal by adjusting the superfluid halo's mass. This is how SFDM can fit the Einstein radii and velocity dispersions at the same time.

The one galaxy for which we could not obtain a successful fit is J0737+3216. Our best-fit for this galaxy satisfies our requirement that the calculated and measured Einstein radii agree to better than 0.01 kpc. But it does not satisfy the requirement that the measured and calculated velocity dispersions agree within the measurement error $\sigma_*^{\rm error}$. This can be seen in Fig. 2.1, left, where J0737+3216 is the one outlier at $(\sigma_*^{\rm calc} - \sigma_*^{\rm meas})/\sigma_*^{\rm error} \approx -1.5$. That is, the calculated and observed velocity dispersions agree within $1.5 \cdot \sigma_*^{\rm error}$ rather than within $1 \cdot \sigma_*^{\rm error}$. This one outlier is not too interesting. The slight discrepancy in the velocity dispersion can easily be due to any number of uncertainties in the astronomical data or our theoretical modelling.

Thus, the strong lensing and velocity dispersion data does not pose a challenge for SFDM. Both can easily be fit at the same time. A caveat is that

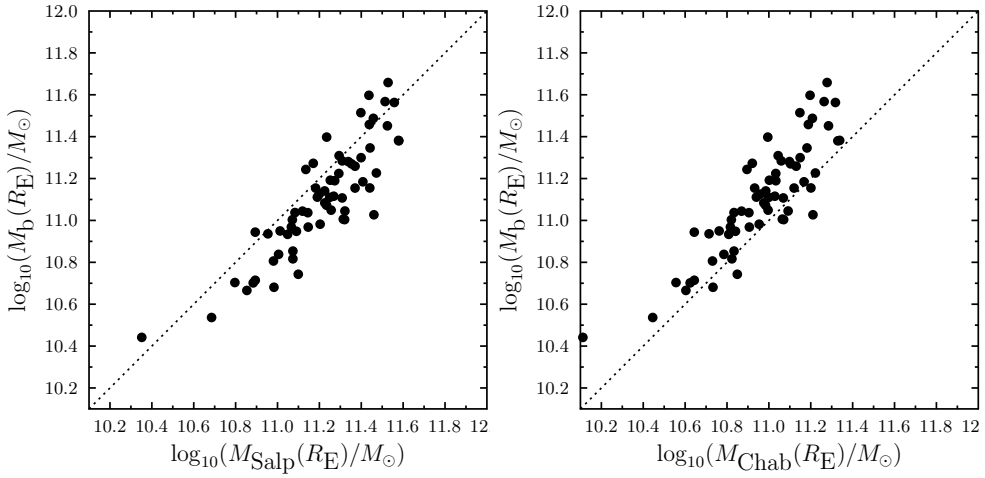

Figure 2.2.: Left: The best-fit baryonic mass M_b for each galaxy compared to that inferred from the Salpeter IMF, in terms of the baryonic mass within a cylinder with radius $R_{\rm E}^{\rm meas}$. Right: The same as left, but with the Chabrier IMF instead of the Salpeter IMF.

a fully satisfactory fit in SFDM needs the inner parts of a galaxy to be in the MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$ (see Sec. 1.4). Otherwise, rotation curves and velocity dispersions are not naturally MOND-like, contrary to one of the main motivations behind SFDM (see also Sec. 3.2). Future work should check whether or not the above fits satisfy this condition in a reasonable spatial region of the lensing galaxies, and if not, whether other acceptable fits that do satisfy it are possible.

2.2. The Milky Way rotation curve

In this section, we will discuss the Milky Way (MW) rotation curve in SFDM. Our aim is to see whether SFDM can reasonably fit this rotation curve and, more generally, to develop the tools required to understand axisymmetric galaxies in SFDM.

2.2.1. The model

Consider the superfluid core of the Milky Way in SFDM. We assume this superfluid core to be in equilibrium. For a fully axisymmetric treatment, we cannot use the no-curl approximation. Instead, we have to solve the coupled

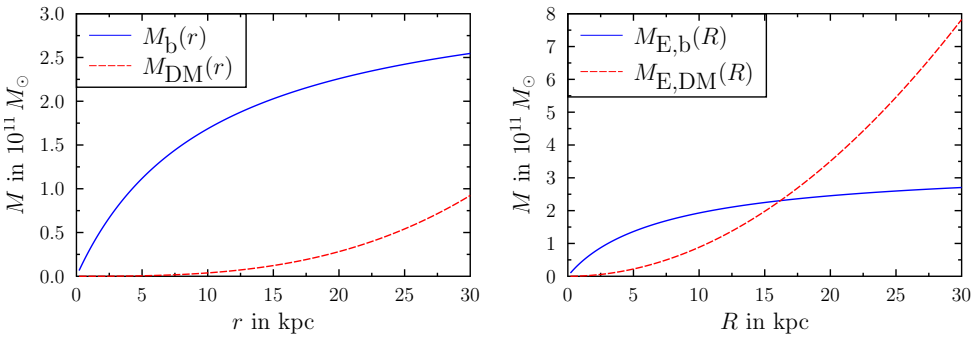

Figure 2.3.: Left: The baryonic and non-baryonic masses $M_b(r)$ and $M_{\rm DM}(r)$ inside a sphere with radius r for the galaxy J0029-0055. Small radii in this panel are relevant for the velocity dispersion. Right: The baryonic and non-baryonic masses $M_{{\rm E},b}(R)$ and $M_{\rm E,DM}(R)$ inside a cylinder with radius R for the same galaxy. Small radii in this panel are relevant for the strong lensing signal.

differential equations

$$\Delta\left(-\frac{\hat{\mu}}{m}\right) = 4\pi G\left(\rho_b + \rho_{\rm SF}\left(\vec{\nabla}\theta, \hat{\mu}\right)\right)\,, \tag{2.7a}$$

$$\vec{\nabla}\left(\frac{(\vec{\nabla}\theta)^2 + 2m(\frac{2\bar{\beta}}{3}-1)\hat{\mu}}{\sqrt{(\vec{\nabla}\theta)^2 + 2m(\bar{\beta}-1)\hat{\mu}}}\,\vec{\nabla}\theta\right) = \frac{\bar{\alpha}}{2M_{\rm Pl}}\,\rho_b\,. \tag{2.7b}$$

In addition to axisymmetry we assume a $z \to -z$ symmetry. Then, it suffices to numerically solve these equations in a region defined by $z > 0$ and $\sqrt{R^2+z^2} < r_\infty$ for some r_∞. Here, R and z are the standard cylindrical coordinates. As boundary conditions, we impose

$$\partial_z\hat{\mu}|_{z=0} = 0\,, \tag{2.8a}$$

$$\partial_z\theta|_{z=0} = 0\,, \tag{2.8b}$$

and

$$\hat{\mu}|_{\sqrt{R^2+z^2}=r_\infty} = \mu_\infty\,, \tag{2.9a}$$

$$\theta|_{\sqrt{R^2+z^2}=r_\infty} = 0\,, \tag{2.9b}$$

for some μ_∞. The value of μ_∞ determines the size of the superfluid halo. It is analogous to $\hat{\mu}(r_0)$ in our spherically symmetric calculation in the previous section. The boundary value of θ is inconsequential for us so we just set it to 0. This form of the boundary conditions enforces spherical symmetry at r_∞. This was shown to be a reasonable approximation in MOND in Ref. [102]. We expect it to be reasonable in SFDM as well. Unless stated otherwise, we choose $r_\infty = 100\,\text{kpc}$ and $\mu_\infty/m = 1.25\cdot 10^{-8}$. As we will see below, the rotation curve at $R < 25\,\text{kpc}$ depends only weakly on this choice. For the

SFDM model parameters we again adopt the fiducial values from Ref. [72], i.e. $\bar{\alpha} = 5.7$, $m = 1\,\mathrm{eV}$, $\Lambda = 0.05\,\mathrm{meV}$, and $\bar{\beta} = 2$. We numerically solve these equations using Mathematica [103].

We adopt the baryonic mass distribution of Ref. [104]. This mass distribution consists of a bulge, a gas disk, and a stellar disk. We take the stellar disk exactly as in Ref. [104]. That is, we take the numerical surface density from Ref. [104] with a scale height of 300 pc. For the gas disk, we similarly take the numerical surface density from Ref. [104]. But we adopt a scale height of 130 pc instead of an infinitely thin disk, following Ref. [105]. A finite scale height is easier to handle with our numerical code. This choice does not significantly affect our results.

For the bulge, we use [106]

$$\rho_{\mathrm{bulge}}(b) = \frac{\rho_{\mathrm{bulge},0}}{\eta\zeta b_m^3} \frac{\exp\left[-(b/b_m)^2\right]}{(1 + b/b_0)^{1.8}} \,, \tag{2.10}$$

where $\eta = 0.5$, $\zeta = 0.6$, $b_m = 1.9\,\mathrm{kpc}$, $b_0 = 0.1\,\mathrm{kpc}$, and $b = r/(\eta\zeta)^{1/3}$ with the spherical radius r. This is the spherically equivalent version of the triaxial bulge model used in Ref. [104]. The constant $\rho_{\mathrm{bulge},0}$ controls the total bulge mass. We choose its value to reproduce the asymptotic Newtonian acceleration due to the bulge from Ref. [104].

We keep the shape of the baryonic mass distribution fixed. We only use the overall normalization of the baryonic mass distribution as a free parameter. Specifically, we introduce a parameter f_b that multiplies the baryonic density $\rho_b(R, z)$. Below, we will adjust f_b to find a reasonable fit to the observed rotation curve.

We also adopt the observed rotation curve from Ref. [104]. This is based on the rotation curve data from Ref. [107] for $R > 5\,\mathrm{kpc}$ and that from Ref. [108] for $R < 2.2\,\mathrm{kpc}$, but adjusted to fit the assumptions about ρ_b used in Ref. [104]. Simply rescaling $\rho_b(R, z) \to f_b \cdot \rho_b(R, z)$ does not require such adjustments. Thus, we do not need to do any adjustments to this data when adjusting f_b.

The numerical parameters used in this section are summarized in Table 2.1.

2.2.2. Results

We first consider the rotation curve at $R < 25\,\mathrm{kpc}$ for different baryonic masses, i.e. different f_b, for a fixed boundary condition μ_∞. This is shown in Fig. 2.4. A reasonable fit to the observed rotation curve is achieved for $f_b = 0.8$. That is, SFDM requires about 20% less baryonic mass compared to the MOND model discussed in Ref. [104].

Specifically, Ref. [104] considered the no-curl approximation of MOND with the interpolation function

$$\nu_e(y) = \frac{1}{1 - e^{-\sqrt{y}}} \,, \tag{2.11}$$

where $y = a_b/a_0$. This interpolation function is commonly used in the context of the RAR [11]. In contrast, in the MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$ and without

Table 2.1.: The numerical parameters used in this section.

Model parameters		
m	$1\,\mathrm{eV}$	from Ref. [72]
$\bar{\alpha}$	5.7	from Ref. [72]
Λ	$0.05\,\mathrm{meV}$	from Ref. [72]
$\bar{\beta}$	2	from Ref. [72]
Baryonic mass		
Bulge	$1.29 \cdot 10^{10}\, M_\odot \times f_b$	see Eq. (2.10)
Stellar disk	$4.94 \cdot 10^{10}\, M_\odot \times f_b$	from Ref. [104]
Gas disk	$1.22 \cdot 10^{10}\, M_\odot \times f_b$	from Ref. [104]
f_b	0.8	unless stated otherwise
Boundary condition		
μ_∞/m	$1.25 \cdot 10^{-8}$	unless stated otherwise
r_∞	$100\,\mathrm{kpc}$	unless stated otherwise

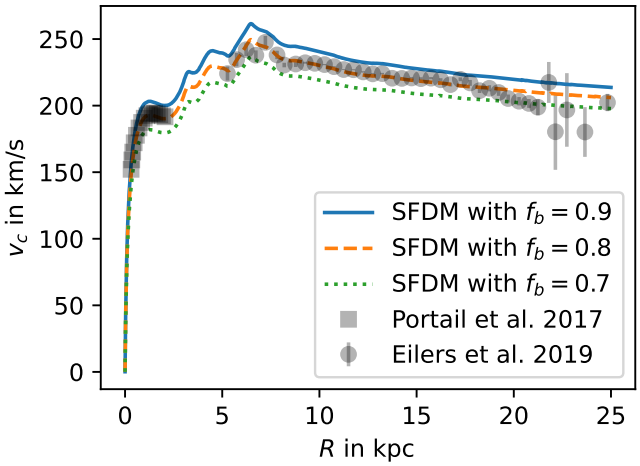

Figure 2.4.: SFDM MW rotation curve for $f_b = 0.9$ (solid blue line), $f_b = 0.8$ (dashed orange line), and $f_b = 0.7$ (dotted green line). Also shown are the observed rotation curve data from Ref. [107] (black dots) and Ref. [108] (black squares), both adjusted to match the assumptions of Ref. [104].

the superfluid halo's gravitational pull, SFDM corresponds to the interpolation function

$$\nu_\theta(y) = 1 + \frac{1}{\sqrt{y}}\,, \tag{2.12}$$

as discussed in Sec. 1.4 (see also Ref. [109]). The main difference between the interpolation functions ν_e and ν_θ is that ν_e approaches its limits, $\nu \to 1$ for $y \to \infty$ and $\nu \to 1/\sqrt{y}$ for $y \to 0$, much faster than ν_θ. At intermediate accelerations a_b, the total acceleration is a bit larger in SFDM compared to standard MOND models. This is the main reason SFDM requires less baryonic mass.

At relatively large radii, there is another potential reason why SFDM might require less baryonic mass. Namely, the superfluid halo's gravitational pull contributes to the rotation curve, so that less baryonic mass is needed to fit the observed rotation curve. However, this is countered by the fact that the value of a_0 is usually chosen a bit smaller in SFDM than in standard MOND models. Indeed, the fiducial parameters from Ref. [72] give $a_0 = 0.87 \cdot 10^{-10}\,\mathrm{m/s^2}$ while Ref. [104] uses $a_0 = 1.2 \cdot 10^{-10}\,\mathrm{m/s^2}$.

This is illustrated in Fig. 2.5. At small radii, SFDM with $f_b = 0.8$ agrees with the ν_e-based model from Ref. [104], irrespective of whether or not we include the superfluid's gravitational pull. At larger radii, SFDM without the superfluid's gravitational pull would require f_b larger than 0.8. If we include the superfluid's gravitational pull, this gravitational pull and the lower value of a_0 roughly compensate each other. Thus, the main reason why SFDM requires less baryonic mass than standard MOND models is indeed that the interpolation function ν_θ produces more additional acceleration than ν_e at intermediate a_b.

As mentioned above, Ref. [72] proposes two estimates for the radius where the superfluid core ends. Both of these implicitly assume spherical symmetry. Thus, in principle, we cannot use them here. But in practice it turns out to be straightforward to extend these to axisymmetric situations. The reason is that the transition usually happens at rather large radii. At these radii, both $\hat{\mu}$ and θ are approximately spherically symmetric. Indeed, we impose spherical symmetry as a boundary condition at r_∞. This is illustrated in Fig. 2.6. For $R \gtrsim 20\mathrm{kpc}$, assuming spherical symmetry for the superfluid's energy density is an excellent approximation.

Thus, we can do the following. Consider the two proposals from Ref. [72] for the thermal radius R_T and the NFW radius R_{NFW}. Then simply set $z = 0$ and use the cylindrical radius R instead of the spherical radius r whenever the original proposal from Ref. [72] uses r. This gives radii $R_{T,R}$ and $R_{\mathrm{NFW},R}$. Similarly, we can set $R = 0$ and use z instead of the spherical radius r. This gives radii $R_{T,z}$ and $R_{\mathrm{NFW},z}$. We expect that $R_{T,R} \approx R_{T,z}$ and $R_{\mathrm{NFW},R} \approx R_{\mathrm{NFW},z}$ since we expect approximate spherical symmetry at the transition radius. Then, we can use either of the two versions of R_T and R_{NFW} as the thermal and NFW radius, respectively. The small difference between the $R = 0$ and the $z = 0$ version will not matter for practical purposes.

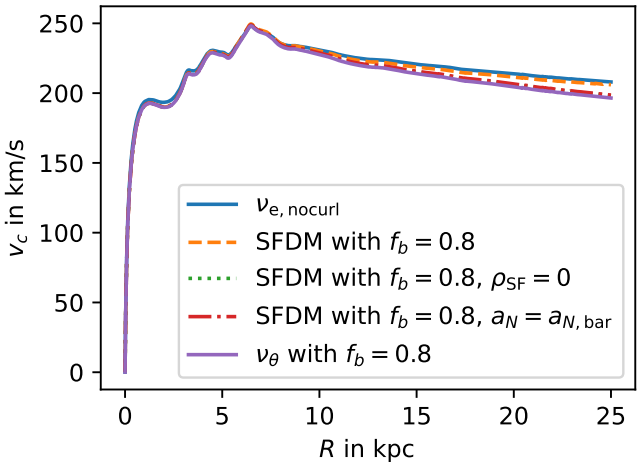

Figure 2.5.: The MW rotation curve from the ν_e-based model from Ref. [104] (solid blue line), from SFDM with $f_b = 0.8$ (dashed orange line), and from SFDM with $f_b = 0.8$ with the superfluid's gravitational pull removed using three different methods. The first method is to set $\rho_{SF} = 0$ in SFDM's equations (dotted green line). The second method is to keep ρ_{SF}, but to use only the baryonic Newtonian acceleration a_b instead of the total Newtonian acceleration a_N when calculating the rotation curve (dash-dotted red line). The third method is to set $\rho_{SF} = 0$ in SFDM's equations and to determine the phonon force from the MOND limit relation Eq. (3.7) (see Sec. 3.2).

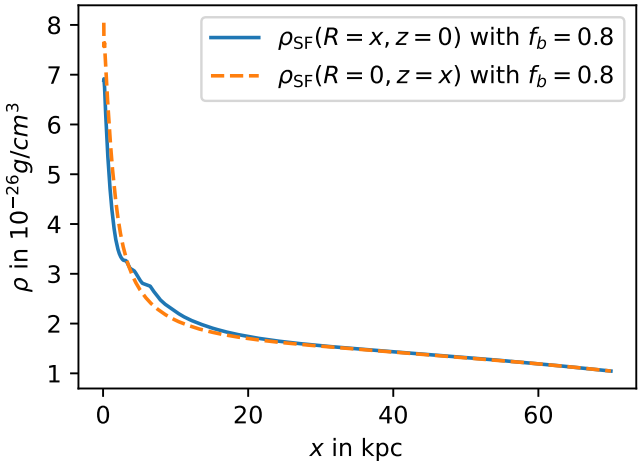

Figure 2.6.: The superfluid's energy density ρ_{SF} within the galactic plane at $z = 0$ (solid blue line) and above the galactic plane at $R = 0$ (dashed orange line). For a spherically symmetric superfluid these two are exactly equal. In our case, assuming spherical symmetry is an excellent approximation at $R \gtrsim 20\,\mathrm{kpc}$.

Indeed, for $f_b = 0.8$, we find $R_{T,R} = 87.5\,\mathrm{kpc}$ and $R_{T,z} = 87.6\,\mathrm{kpc}$. Similarly, $R_{\mathrm{NFW},R} = 66.0\,\mathrm{kpc}$ and $R_{\mathrm{NFW},z} = 65.7\,\mathrm{kpc}$. Thus, our simple procedure seems reasonable. The difference between the NFW and the thermal radius is about 30%. This is comparable to the differences found in spherically symmetric situations in Ref. [72].

For a given μ_∞ we can now calculate the dark matter profile even beyond the transition radius by matching an NFW halo to the superfluid core at the transition radius. For concreteness, we use the NFW radius $R_{\mathrm{NFW},R}$ which, for brevity, we simply call R_{NFW} in the following. Then, we can calculate the total dark matter mass M_{200}^{DM} within the virial radius r_{200}. Here, r_{200} is the spherical radius where the averaged dark matter density drops below 200 times the cosmological density $3H_0^2/(8\pi G)$. We adopt $H_0 = 67.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ in this section. We match the superfluid's energy density and pressure to that of a full NFW profile. So we need one more condition to fix all parameters (the transition radius as well as the two parameters of the NFW profile, as discussed in Sec. 1.4). Here, we fix the NFW concentration parameter to be that of ΛCDM abundance matching expectations [72, 85, 94].

In contrast to the rotation curve at $R < 25\,\mathrm{kpc}$, the total dark matter mass depends strongly on the value of the boundary condition μ_∞. To illustrate this, we show the rotation curves for a number of choices of μ_∞ and r_∞ in Fig. 2.7, see also Table 2.2. These values are chosen to cover the range of measured values of M_{200}^{DM} given in Ref. [110]. But note that some of these measurements may not apply directly in SFDM, since they assume ΛCDM.

The rotation curves in Fig. 2.7 are discontinuous at $R = R_{\mathrm{NFW}}$. This is

Table 2.2.: The NFW radius $R_{\mathrm{NFW}} \equiv R_{\mathrm{NFW},R}$, the total dark matter mass M_{200}^{DM}, the virial radius r_{200}, and the thermal radius $R_T \equiv R_{T,R}$ for $f_b = 0.8$ and various boundary conditions μ_∞ imposed at radii r_∞. Here, M_{200}^{DM} and r_{200} are calculated assuming the NFW radius as the transition radius.

r_∞ kpc	μ_∞/m 10^{-8}	R_{NFW} kpc	M_{200}^{DM} $10^{12}\,M_\odot$	r_{200} kpc	R_T kpc
110	7.80	73	3.0	306	105
100	6.24	70	2.0	265	97
100	1.25	66	1.2	225	87
90	0.25	65	0.7	189	76
80	0.12	69	0.5	163	67

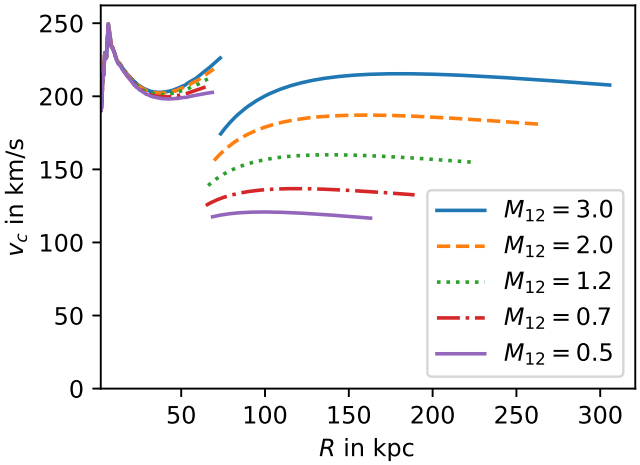

Figure 2.7.: MW rotation curves for different boundary conditions μ_∞ imposed at different radii r_∞, corresponding to different total dark matter masses $M_{200}^{\mathrm{DM}} \equiv M_{12} \cdot 10^{12}\,M_\odot$. Each rotation curve is plotted up to the respective virial radius r_{200}, assuming the NFW radius R_{NFW} as the transition radius. There is a discontinuity at $R = R_{\mathrm{NFW}}$ because we abruptly switch off the phonon force at this radius for simplicity. In a real galaxy, this transition should be more gradual. See also Table 2.2.

because in SFDM the phonon force is assumed to be present only inside the superfluid core [71, 72]. Thus, we include this force only for $R < R_{\mathrm{NFW}}$. In a real galaxy, this transition should be gradual and not abrupt. The details of this transition are unknown and require further theoretical work. Thus, the rotation curves in Fig. 2.7 should not be taken seriously around the discontinuity at $R = R_{\mathrm{NFW}}$. Away from this discontinuity the rotation curves represent the SFDM expectation.

We also see in Fig. 2.7 that the rotation curves bend upwards around $R = R_{\mathrm{NFW}}$. This is due to the combination of the phonon force, which gives flat rotation curves, and the cored superfluid halo which starts to contribute significantly at these radii. This was already observed in Ref. [72]. This may rule out larger values of M_{200}^{DM} for which this effect is more pronounced.

To summarize, we have first demonstrated how to make predictions in SFDM in axisymmetric situations. We have then shown that SFDM can reasonably fit the MW rotation curve and we have explained why SFDM requires less baryonic mass to fit the MW rotation curve than standard MOND models.

3. Three problems of superfluid dark matter and their solution

So far, we have taken the SFDM model as proposed in Refs. [71, 72] as given and tested its predictions against observations. In this chapter, we take a closer look at how this model is constructed. We will discuss three problems regarding its internal consistency: The stability problem, the MOND limit problem, and the equilibrium problem. All these problems have the same root cause, namely that the phonon field θ plays a double role. It carries both the superfluid's energy density and the MOND-like phonon force. As a possible solution we propose a model where these two roles are split between two different fields. This summarizes our work from Refs. [73, 87].

3.1. The stability problem

The stability problem is related to the parameter $\bar{\beta}$ in the low-energy effective Lagrangian Eq. (1.7). As mentioned above, $\bar{\beta}$ is supposed to parametrize finite-temperature corrections. We will now discuss why this parameter was introduced in Ref. [71] and why it is not a satisfactory solution to the problem it is supposed to solve.

To this end, consider the Lagrangian

$$\mathcal{L}|_{\bar{\beta}=0} = f(K_\theta - m^2) - \frac{\bar{\alpha}\Lambda}{M_{\mathrm{Pl}}}\,\theta\,\rho_b\,, \tag{3.1a}$$

$$f(K_\theta - m^2) = \frac{2\Lambda}{3}\sqrt{|K_\theta - m^2|}(K_\theta - m^2)\,, \tag{3.1b}$$

$$K_\theta = \nabla^\alpha\theta\nabla_\alpha\theta\,. \tag{3.1c}$$

This reproduces the Lagrangian Eq. (1.7) with $\bar{\beta} = 0$ after introducing a chemical potential, $\dot{\theta} \to \dot{\theta} + \mu$ with a constant μ, and after taking the non-relativistic limit, $\mu = m + \mu_{\mathrm{nr}}$ with $|\mu_{\mathrm{nr}}| \ll m$ and $|\dot{\theta}| \ll m$.

The problem with this $\bar{\beta} = 0$ Lagrangian is that equilibrium solutions in galaxies are unstable. To see this, consider a static background solution θ_0 with a background chemical potential $\mu_0 = m + \mu_{\mathrm{nr},0}$. For simplicity, we consider a MOND limit solution, i.e. $(\vec{\nabla}\theta_0)^2 \gg 2m\hat{\mu}_0$ with $\hat{\mu}_0 = \mu_{\mathrm{nr},0} - m\phi_N$. The second-order Lagrangian for perturbations δ on top of this background solution θ_0 is

$$\mathcal{L}_{\mathrm{pert}} = \left[f_0' g^{\alpha\beta} + 2f_0''\nabla^\alpha\theta_0\nabla^\beta\theta_0\right]\nabla_\alpha\delta\nabla_\beta\delta\,, \tag{3.2}$$

where

$$f_0' = \Lambda\sqrt{|K_{\theta_0} - m^2|} > 0\,, \tag{3.3a}$$

$$f_0'' = -\frac{\Lambda}{2}\frac{1}{\sqrt{|K_{\theta_0} - m^2|}} < 0\,, \tag{3.3b}$$

and where K_{θ_0} is the background value of K_θ . Here, we used the fact that $K_{\theta_0} - m^2 < 0$ for the type of solution under consideration. Namely,

$$K_{\theta_0} - m^2 \approx 2m\hat{\mu}_0 - (\vec{\nabla}\theta_0)^2\,, \tag{3.4}$$

which is negative when spatial derivatives dominate. In the MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$, this is the case. In the opposite limit, f_0'' becomes positive.

For stability, the prefactor of $\dot{\delta}^2$ in Eq. (3.2) should be positive. But this is not the case. For a galaxy in equilibrium, this prefactor is

$$f_0' g^{00} + 2\left(g^{00}\right)^2 f_0'' \mu_0^2 < 0\,. \tag{3.5}$$

This is negative since f_0'' is negative in the MOND limit and since the $\mu_0^2 f_0''$ term dominates compared to the f_0' term in the non-relativistic limit [71].

This instability is the reason why Ref. [71] introduced the finite-temperature corrections parametrized by $\bar{\beta}$. These corrections do avoid the instability but are not entirely satisfactory. Both the functional form of the finite-temperature corrections and the numerical value of $\bar{\beta}$ are chosen ad-hoc. In principle, the finite-temperature Lagrangian Eq. (1.7) for $\bar{\beta} \neq 0$ should follow from the zero-temperature Lagrangian with $\bar{\beta} = 0$. So far this has not been established. For example, it is unclear which value of $\bar{\beta}$ corresponds to which temperature. It is also unclear whether the finite-temperature corrections actually take the form assumed in Ref. [71]. Thus, the proposed finite-temperature corrections may easily turn out to be unphysical.

The reason behind the instability at $\bar{\beta} = 0$ is that θ carries both the phonon force and the superfluid. Specifically, the instability is due to the fact that the combination $f_0''\mu_0^2$ is negative and dominates in the prefactor of $\dot{\delta}^2$. The fact that this term dominates is due to θ carrying the superfluid's chemical potential μ_0. The fact that his term is negative is due to the fact that spatial derivatives dominate in $K_{\theta_0} - m^2$ which is required for a MOND-like force. Below, we will introduce an improved model that avoids this instability in a more natural way without ad-hoc finite-temperature corrections.

3.2. The MOND limit problem

One of the main motivations of SFDM is that it has a MOND limit for

$$(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}\,. \tag{3.6}$$

In this limit, the acceleration $\vec{a}_\theta$ due to the phonons satisfies a MOND-like equation

$$\vec{\nabla}\left(|\vec{a}_\theta|\vec{a}_\theta\right) = \vec{\nabla}\left(a_0\vec{a}_b\right)\,. \tag{3.7}$$

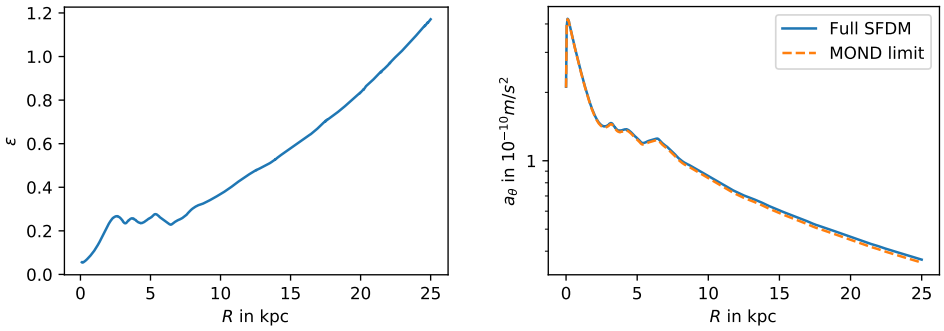

Figure 3.1.: Left: The quantity $\varepsilon = (2m\hat{\mu})/(\vec{\nabla}\theta)^2$ that controls the MOND limit of SFDM for the Milky Way model from Sec. 2.2 at $z = 0$. The Milky Way is not everywhere in the MOND limit $\varepsilon \ll 1$. Right: The phonon acceleration a_θ at $z = 0$ for the full SFDM model (solid blue line) and assuming the MOND limit relation Eq. (3.7) (dashed orange line) for the same Milky Way model as in the left panel. The phonon acceleration is close to its MOND limit value even when ε is not small.

In this section, we will explain why many galaxies cannot reach this proper MOND limit in SFDM and why this cannot easily be fixed by changing the model's parameters.

Consider the quantity $\varepsilon \equiv 2m\hat{\mu}/(\vec{\nabla}\theta)^2$ that controls the MOND limit. In the MOND limit $\varepsilon \ll 1$ and assuming the no-curl approximation, we can write ε as $2m\hat{\mu}/(\bar{\alpha}M_{\rm Pl}a_b)$. A rough lower bound on $\hat{\mu}/m$ is GM_b/r.[1] With $a_b = GM_b/r^2$, we then get in the MOND limit $\varepsilon \ll 1$

$$\varepsilon \gtrsim \frac{2m^2 r}{\bar{\alpha}M_{\rm Pl}} = 0.11 \cdot \left(\frac{r}{5\,\mathrm{kpc}}\right) , \tag{3.8}$$

where we assumed the fiducial numerical values from Ref. [72] for the equality. We see that ε cannot be very small for many galaxies. That is, galaxies may easily fail to reach the MOND limit of SFDM. This is illustrated in Fig. 3.1, left, for the Milky Way model discussed in Sec. 2.2. The quantity ε is larger than 1 for $R \gtrsim 20\,\mathrm{kpc}$. Even at smaller radii, ε is at best moderately small.

In principle, the parameter combination $m^2/\bar{\alpha}$ that determines our lower bound Eq. (3.8) can be made smaller to avoid this conclusion. But this is in tension with the other role θ plays in galaxies, namely that θ carries the superfluid's energy density. For example, in the MOND limit $\varepsilon \ll 1$, this energy density is proportional to $m^2/\bar{\alpha}$,

$$\rho_{\rm SF} \propto \frac{m^2}{\bar{\alpha}} M_{\rm Pl} |\vec{a}_\theta| \, . \tag{3.9}$$

[1] If we allow for negative $\hat{\mu}$, as briefly discussed in Sec. 1.4, this lower bound becomes $\hat{\mu}/m \gtrsim \hat{\mu}_{\rm min}/m + GM_b/r$, where $\hat{\mu}_{\rm min}$ is the minimum allowed value of $\hat{\mu}$. This will not change our qualitative conclusion.

Thus, significantly decreasing $m^2/\bar{\alpha}$ to get a small ε also significantly decreases the superfluid's energy density. But the superfluid density cannot be too small because it must produce a sufficient strong lensing signal, see Sec. 2.1. In general, even when ε is not small, the energy density $\rho_{\rm SF}$ becomes larger with a larger $\hat{\mu}/m$. Thus, the smaller prefactor $m^2/\bar{\alpha}$ in $\rho_{\rm SF}$ can be countered by making $\hat{\mu}/m$ larger. However, a larger $\hat{\mu}/m$ also makes ε larger. So this effort to suppress the lower bound Eq. (3.8) by making $m^2/\bar{\alpha}$ smaller might not work after all. Indeed, in the MOND limit, ε grows linearly with $\hat{\mu}/m$ while $\rho_{\rm SF}$ grows sublinearly with $\hat{\mu}/m$ as can be seen from Eq. (1.17).

Thus, the MOND limit cannot be fixed by simply choosing different numerical parameters. This is what we refer to as the MOND limit problem.

But there is also a puzzle here, which is illustrated in Fig. 3.1, right, for the MW model from Sec. 2.2. Namely, the phonon acceleration a_θ can be close to its MOND limit value (roughly $\sqrt{a_0 a_b}$) even when ε is not small, i.e. even outside the MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$. This can happen as long as ε is not too large, at least for $\bar{\beta} \approx 2$. That is, galaxies may still have a MOND-like rotation curve even if ε is not small. We will refer to this as the pseudo-MOND limit of SFDM. We will now explain where this pseudo-MOND limit comes from and why it is not satisfactory to rely on this pseudo-MOND limit for MOND-like rotation curves.

Consider the no-curl approximation for the phonon acceleration $\vec{a}_\theta$. This gives a_θ as an algebraic function of a_b and $\hat{\mu}$. Specifically,

$$a_\theta = \sqrt{a_0 a_b} \cdot \sqrt{x}\,, \tag{3.10}$$

where $x = x(\varepsilon_*, \bar{\beta})$ is determined by the cubic equation

$$0 = x^3 + 2\left(\frac{2\bar{\beta}}{3} - 1\right)\varepsilon_* \cdot x^2 + \left(\left(\frac{2\bar{\beta}}{3} - 1\right)^2 (\varepsilon_*)^2 - 1\right) x - \left(\bar{\beta} - 1\right)\varepsilon_*\,, \tag{3.11}$$

with

$$\varepsilon_* \equiv \frac{2m^2}{\bar{\alpha} M_{\rm Pl}\, a_b}\frac{\hat{\mu}}{m} = \varepsilon \cdot \frac{a_\theta^2}{a_0 a_b} = \varepsilon \cdot x\,. \tag{3.12}$$

The acceleration a_θ is MOND-like whenever $\sqrt{x}$ is close to 1. This depends on the values of ε_* and $\bar{\beta}$. The quantity ε_* is the same as ε whenever the phonon acceleration a_θ is close to its MOND limit value $\sqrt{a_0 a_b}$. This is the case both in the proper MOND limit $\varepsilon \ll 1$ and in the pseudo-MOND limit where ε is not small. When the phonon acceleration is not close to its MOND limit value, ε and ε_* differ.

The quantity $\sqrt{x} = a_\theta/\sqrt{a_0 a_b}$ as a function of ε_* and $\bar{\beta}$ is illustrated in Fig. 3.2, left. The phonon force a_θ is close to its MOND limit value $\sqrt{a_0 a_b}$ for $\varepsilon_* \ll 1$, independently of the value of $\bar{\beta}$. This is the proper MOND limit $\varepsilon \ll 1$ already discussed above. Indeed, for small ε_*, Eq. (3.11) gives $\sqrt{x} \approx 1$ which also implies $\varepsilon \ll 1$ according to Eq. (3.12).

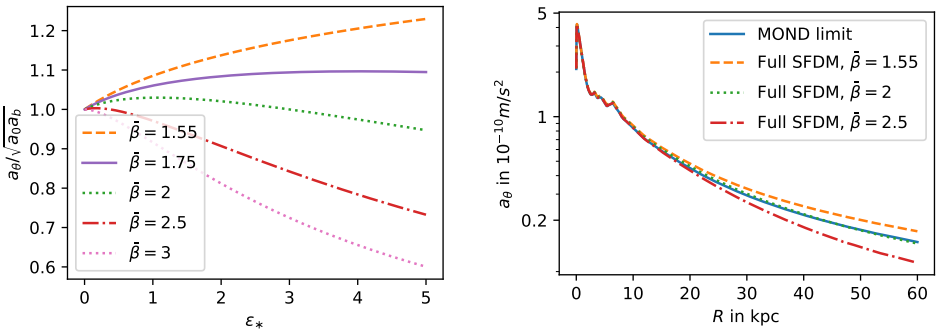

Figure 3.2.: Left: The quantity $\sqrt{x} = |\vec{a}_\theta|/\sqrt{a_0|\vec{a}_b|}$ in the no-curl-approximation of standard SFDM as a function of $\varepsilon_* = (2m\hat{\mu})/(\bar{\alpha} M_{\rm Pl}|\vec{a}_b|)$. For $\sqrt{x} = 1$, the phonon force has a MOND-like form. Right: The phonon acceleration a_θ at $z = 0$ for the Milky Way model from Sec. 2.2 in the full standard SFDM model with $\bar{\beta} = 2.5$ (dash-dotted red line), $\bar{\beta} = 1.55$ (dashed orange line), the fiducial value $\bar{\beta} = 2$ from Ref. [72] (dotted green line), and assuming the MOND limit relation Eq. (3.7) (solid blue line). The baryonic mass distribution and the boundary condition for $\hat{\mu}/m$ are as described in Sec. 2.2.

But a_θ can be close to $\sqrt{a_0 a_b}$ even when ε_* is not small. This is the pseudo-MOND limit, where rotation curves of galaxies may be MOND-like although a_θ does not satisfy a MOND-like equation. From Fig. 3.2, left, we see that this happens for $\bar{\beta} = 2$ and $\varepsilon_* = \mathcal{O}(1)$. It does not work anymore for $\varepsilon_* \gg 1$. This can also be seen by expanding Eq. (3.11) for large ε_*. Numerically, the conditions $\varepsilon_* = \mathcal{O}(1)$ and $\varepsilon = \mathcal{O}(1)$ are roughly equivalent. Thus, the pseudo-MOND limit roughly requires $\bar{\beta} = 2$ and $\varepsilon = \mathcal{O}(1)$.

This explains the puzzle described above: The MW model from Sec. 2.2 has a MOND-like phonon acceleration a_θ although ε is not small because it is in the pseudo-MOND limit with $\bar{\beta} = 2$ and $\varepsilon = \mathcal{O}(1)$. This is illustrated in Fig. 3.2, right.

Thus, in principle, standard SFDM can produce MOND-like rotation curves despite the MOND limit problem. While many galaxies may not be in the proper MOND limit $\varepsilon_* \ll 1$, many more galaxies could be in the pseudo-MOND limit with $\bar{\beta} = 2$ and $\varepsilon_* = \mathcal{O}(1)$. An interesting consequence of this would be that only isolated galaxies would have MOND-like rotation curves. Other MOND phenomenology like the external field effect would be different. This is in contrast to what is usually expected in SFDM and may change the predictions for satellite galaxies and globular clusters [71].

But relying on the pseudo-MOND limit is not a satisfactory solution. At $\varepsilon_* = \mathcal{O}(1)$, the phonon force depends sensitively on the details of the ad-hoc finite temperature corrections such as the value of $\bar{\beta}$ (see Fig. 3.2, left). Thus, just as these finite-temperature corrections, the pseudo-MOND limit

may easily be unphysical.

The improved model we introduce below allows for a proper MOND limit more generally and without relying on the details of ad-hoc finite-temperature corrections.

3.3. The equilibrium problem

In SFDM, it is usually assumed that the superfluid core of galaxies is in an equilibrium with a chemical potential $\mu = \mathcal{O}(m)$. The chemical potential is important, for example, to cancel the m^2 in $K_\theta - m^2$ (see the Lagrangian Eq. (3.1) and Sec. 3.4 below). This allows spatial derivative terms to dominate in the non-relativistic limit as required for a MOND-like phonon force. This chemical potential corresponds to the $U(1)$ shift symmetry of the phonon field θ. But, in SFDM, this symmetry is explicitly broken by the coupling of θ to baryons, $(\bar{\alpha}\Lambda/M_{\rm Pl})\theta\rho_b$. Here, we will explain how this restricts the standard SFDM phenomenology to relatively short timescales.

To see why there might be a problem consider first a heuristic argument. In field-theoretical contexts, a chemical potential μ is sometimes introduced by shifting $\theta \to \mu \cdot t + \theta$ in the Lagrangian (see e.g. Sec. 1.4). In SFDM, this introduces an explicit time-dependence in the baryon-phonon coupling,

$$(\bar{\alpha}\Lambda/M_{\rm Pl})(\mu t)\rho_b\,. \tag{3.13}$$

This is because θ occurs without derivatives in this coupling. Usually, one assumes a time-independent equilibrium in SFDM. So the question is how long one can ignore this time-dependence from the coupling to matter.

A somewhat less heuristic argument can be made based on statistical physics. In statistical physics, a chemical potential μ corresponds to a conserved quantity Q. In SFDM, the $U(1)$ symmetry is explicitly broken. So at most there is an approximately conserved quantity Q associated with this $U(1)$ symmetry, if the symmetry-breaking terms are sufficiently small. Applying the statistical physics formalism to this approximately conserved quantity Q, in equilibrium, one uses a shifted Hamiltonian $H \to H - \mu Q$ for calculations. On a Lagrangian level, this corresponds to a shift $\dot{\theta} \to \dot{\theta} + \mu$ without shifting occurrences of θ without derivatives. This does not generate an explicit time dependence in the baryon-phonon coupling. But since θ does occur without derivatives in this coupling, such a shift is not just a change of variables. For example, it leads to a violation of energy conservation when applied to the continuity equation. Framed in this way, the question is how long one can ignore this energy non-conservation.

In any case, we are interested in the timescale t_Q that tells us how long one can ignore the explicit $U(1)$ symmetry breaking and assume a standard time-independent superfluid equilibrium in galaxies. It is not consistent to just ignore this symmetry breaking, as argued above. But it may be possible to assume a standard equilibrium for sufficiently short timescales.

Specifically, without the baryon coupling, $\bar{\alpha} = 0$, the Lagrangian has a shift symmetry of θ which implies that the current $j^\alpha = \partial\mathcal{L}/\partial(\nabla_\alpha\theta)$ is conserved. The associated conserved charge is $Q = \int dV j^0$. This is not the case if we include the baryon coupling, $\bar{\alpha} \neq 0$, as required for the MOND-like phonon force. Then, we have $\nabla_\alpha j^\alpha = -(\bar{\alpha}\Lambda/M_{\rm Pl})\rho_b$. This allows to estimate how long it takes for the charge Q around a galaxy as a whole to change significantly,

$$\left|\frac{\dot{Q}}{Q}\right| \approx m\frac{\bar{\alpha}\Lambda}{M_{\rm Pl}}\frac{M_b}{M_{\rm DM}} = \frac{1}{t_Q} \approx \frac{1}{10^8\,\mathrm{yr}}\frac{M_b}{M_{\rm DM}}\,. \tag{3.14}$$

Here, we defined $t_Q^{-1} \equiv m(\bar{\alpha}\Lambda/M_{\rm Pl})(M_b/M_{\rm DM})$. For the numerical estimate we used the fiducial parameter values from Ref. [72].

The timescale t_Q is not necessarily much larger than galactic timescales of about $10^8\,\mathrm{yr}$. Thus, this may be relevant for SFDM. Indeed, the quantity t_Q concerns a galaxy is a whole and may be too optimistic. In Ref. [87], we also estimated a local timescale $t_{\rm loc}$ separately for each point in space. The local equilibrium is valid only on timescales shorter than $t_{\rm loc}^{-1} = m(\bar{\alpha}\Lambda/M_{\rm Pl})(\rho_b/\rho_{\rm SF})$. In the inner parts of galaxies, $\rho_{\rm SF}$ is typically smaller than ρ_b so that $t_{\rm loc}$ is much shorter than t_Q.

Thus, the standard SFDM phenomenology on galactic scales cannot be valid on timescales much longer than galactic timescales. This may be a problem for SFDM. In principle, this can be avoided by adjusting the model's parameters so that the combination $m(\bar{\alpha}\Lambda/M_{\rm Pl})/\rho_{\rm SF}$ that determines t_Q^{-1} and $t_{\rm loc}^{-1}$ is much smaller. However, this is not easily possible. We have

$$\left(\frac{m\bar{\alpha}\Lambda}{M_{\rm Pl}\rho_{\rm SF}}\right)^2 \propto \frac{1}{m^2/\bar{\alpha}} \cdot \left(\frac{\rho_{\rm SF}^{\rm MOND}}{\rho_{\rm SF}}\right)^2\,, \tag{3.15}$$

where $\rho_{\rm SF}^{\rm MOND}$ is the MOND-limit expression for $\rho_{\rm SF}$. We don't show factors involving a_b and a_0 since these cannot be adjusted much, at least not without giving up on MOND-like rotation curves. One way to make this quantity smaller is to make $m^2/\bar{\alpha}$ larger. But then fewer galaxies can reach the MOND limit, see our lower bound on ε from Eq. (3.8). Alternatively, $\dot{Q}/Q$ becomes small when $\rho_{\rm SF}$ is much larger than its MOND limit value. But then, by definition, the galaxy is not in the MOND limit either. In any case, a large $\rho_{\rm SF}$ may make the superfluid's contribution to the rotation curve significant. Then rotation curves depend significantly on the boundary condition of the superfluid and will not be automatically MOND-like contrary to one of the main motivations behind SFDM.

This equilibrium problem is again due to the double role of θ. That θ mediates a MOND-like force requires the coupling to baryons. That θ also carries the superfluid's chemical potential is in tension with this. The improved model we introduce below has an equilibrium that is valid on much larger timescales.

3.4. A solution: Two-field SFDM

The three problems discussed above have a common root cause which is that the phonon field θ plays a double role. That θ carries a MOND-like force is in tension with the fact that θ also carries the superfluid. A natural way to avoid these problems is to split these two roles between two different fields. Here, we will introduce a model that does this, but still has phenomenology close to the original SFDM model. We will refer to this improved SFDM model as two-field SFDM and to the original SFDM model as standard SFDM. In two-field SFDM, one field carries the MOND-like force on galactic scales. The other field carries the superfluid.

The Lagrangian of two-field SFDM has two parts,

$$\mathcal{L} = \mathcal{L}_- + \mathcal{L}_+ \,. \tag{3.16}$$

The first part, $\mathcal{L}_-$, is a standard superfluid Lagrangian for a complex scalar field $\phi_- = \rho_- e^{-i\theta_-}/\sqrt{2}$ with quartic interactions,

$$\begin{aligned} \mathcal{L}_- &= (\nabla_\alpha \phi_-)^*(\nabla^\alpha \phi_-) - m^2|\phi_-|^2 - \lambda_4|\phi_-|^4 \\ &= \frac{1}{2}\left(K_{\rho_-} + \rho_-^2(K_- - m^2)\right) - \frac{\lambda_4}{4}\rho_-^4 \,, \end{aligned} \tag{3.17}$$

where m is the mass, λ_4 is a coupling constant, and we have

$$K_{\rho_-} = \nabla^\alpha \rho_- \nabla_\alpha \rho_- \,, \tag{3.18a}$$

$$K_- = \nabla^\alpha \theta_- \nabla_\alpha \theta_- \,. \tag{3.18b}$$

This part of the Lagrangian is responsible for the superfluid. We chose a quartic coupling for simplicity but a sextic interaction, for example, would also work. The superfluid phase has a chemical potential $\mu = \mu_{\rm nr} + m > m$ which we can introduce in the standard way by shifting $\dot\theta_- \to \dot\theta_- \to \mu$. If we neglect derivatives of ρ_-, this gives in the non-relativistic limit

$$\rho_-^2 = \frac{g^{00}\mu^2 - m^2}{\lambda_4} \approx \frac{2m^2}{\lambda_4}\frac{\hat\mu}{m} \,, \tag{3.19}$$

where, as in standard SFDM, $\hat\mu = \mu_{\rm nr} - m\phi_N$. The dominant contribution to the energy density is

$$\rho_{\rm DM-} = m^2\rho_-^2 = \frac{2m^4}{\lambda_4}\frac{\hat\mu}{m} \,. \tag{3.20}$$

The second part of the Lagrangian, $\mathcal{L}_+$, is responsible for the MOND-like force. It is similar to the standard SFDM Lagrangian from Eq. (3.1) with $\bar\beta = 0$. Before taking the non-relativistic limit and before introducing the chemical potential, this standard SFDM Lagrangian reads

$$\mathcal{L}|_{\bar\beta=0} = f(K_\theta - m^2) - \frac{\bar\alpha\Lambda}{M_{\rm Pl}}\,\theta\rho_b \,. \tag{3.21}$$

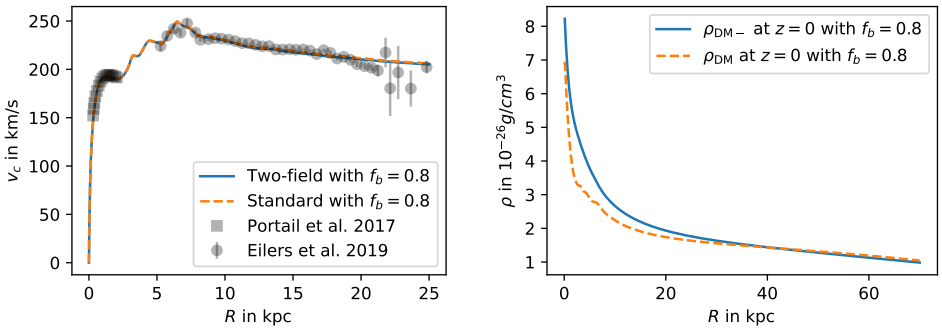

Figure 3.3.: Left: Milky Way rotation curve in two-field SFDM (solid blue line) and standard SFDM (dashed orange line) for the baryonic mass distribution from Sec. 2.2. For standard SFDM, we use the parameters and boundary conditions from Sec. 2.2. For two-field SFDM, we use those from Ref. [73]. Also shown are the measured rotation curves from Ref. [108] (black squares) and Ref. [107] (black dots), both adjusted to match the assumptions of Ref. [104], see also Sec. 2.2. Right: The superfluid's energy density for two-field SFDM (solid blue line) and standard SFDM (dashed orange line) at $z = 0$ for the same Milky Way models as in the left panel.

In $\mathcal{L}_+$, instead we have

$$\mathcal{L}_+ = f(K_+ + K_- - m^2) - \frac{\bar{\alpha}\Lambda}{M_{\rm Pl}}\,\theta_+\rho_b\,, \tag{3.22}$$

where $K_+ = \nabla_\alpha\theta_+\nabla^\alpha\theta_+$. This is just the standard SFDM Lagrangian for $\bar{\beta} = 0$ with K_θ replaced by $K_+ + K_-$ and with θ replaced by θ_+ in the baryon coupling. We couple only θ_+, but not θ_-, to the baryons because only θ_+ carries the MOND-like force.

We will now discuss this two-field model in more detail. As an illustrative example, we consider the Milky Way model from Sec. 2.2. In particular, we numerically solve the two-field SFDM equations of motion for a galaxy in equilibrium with the baryonic mass distribution ρ_b from Sec. 2.2, as discussed in detail in Ref. [73]. This calculation depends on a boundary condition analogous to the boundary condition μ_∞ in standard SFDM. We choose this boundary condition so that the dark matter density of two-field SFDM is roughly comparable to that of standard SFDM with the choice of boundary condition from Sec. 2.2 [73]. The choice of the model parameters $\bar{\alpha}$, m, λ_4, and Λ is described in more detail below. The resulting MW rotation curve for $R < 25\,\text{kpc}$ is shown in Fig. 3.3, left. It is quite close to the rotation curve of standard SFDM discussed in Sec. 2.2. The shape of the superfluid's energy density differs between the two models, but its magnitude is comparable. This is illustrated in Fig. 3.3, right.

In two-field SFDM, the role of ϕ_- is to provide the superfluid. This has two

main effects. First, ϕ_- carries the superfluid's energy density which provides the dark matter halo around galaxies. This superfluid can be in an equilibrium with a chemical potential $\mu = \mathcal{O}(m)$ for a much longer time than in standard SFDM. The reason is that only θ_+ is coupled to the baryons such that the $U(1)$ shift symmetry of θ_- is not explicitly broken. This is how this model avoids the equilibrium problem discussed in Sec. 3.3.

The second effect of ϕ_- is that it induces the MOND regime for the field θ_+, which then carries the MOND-like force. To see this, consider the equation of motion for a static θ_+ in the equilibrium superfluid provided by ϕ_-,

$$\vec{\nabla}\left[2f'\left(K_- - (\vec{\nabla}\theta_+)^2 - m^2\right)\vec{\nabla}\theta_+\right] = \frac{\bar{\alpha}\Lambda}{M_{\rm Pl}}\,\rho_b\,. \tag{3.23}$$

This is a MOND-like equation for θ_+ if the argument of f' is dominated by the $(\vec{\nabla}\theta_+)^2$ term. In the non-relativistic case, we have $(\vec{\nabla}\theta_+)^2 \ll m^2$. Thus, we get a MOND-like equation only if K_- cancels most of the m^2 term. This is exactly what happens for a non-relativistic superfluid with chemical potential $\mu = \mathcal{O}(m)$. The chemical potential is introduced by shifting $\dot{\theta}_- \to \dot{\theta}_- + \mu$. In the non-relativistic case, we have $\mu = m + \mu_{\rm nr}$ with $|\mu_{\rm nr}| \ll m$, so that

$$K_- - m^2 \approx g^{00}\mu^2 - m^2 \approx 2m\hat{\mu}\,. \tag{3.24}$$

Thus, the chemical potential μ in K_- indeed cancels most of the m^2. The argument of f' becomes $2m\hat{\mu} - (\vec{\nabla}\theta_+)^2$. As a result, two-field SFDM has a MOND limit for

$$(\vec{\nabla}\theta_+)^2 \gg 2m\hat{\mu}\,. \tag{3.25}$$

This is analogous to the standard SFDM MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$.

As discussed above, many galaxies are not in this MOND limit in standard SFDM. Basically, the condition $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$ requires $m^2/\bar{\alpha}$ to be sufficiently small. But a small $m^2/\bar{\alpha}$ implies a small superfluid energy density, which is, for example, in tension with strong lensing observations. In two-field SFDM, the MOND limit still requires a small $m^2/\bar{\alpha}$. But this does not imply a small superfluid energy density. A small $m^2/\bar{\alpha}$ still implies a small energy density $\rho_{\rm DM+}$ associated with $\mathcal{L}_+$. But this is not a problem since we can independently choose the model parameters to give a sufficiently large energy density $\rho_{\rm DM-}$ associated with $\mathcal{L}_-$. Then, the superfluid's energy density is mainly due to $\rho_{\rm DM-}$ rather than due to $\rho_{\rm DM+}$. This is illustrated in Fig. 3.5, left. As a consequence, the quantity $(2m\hat{\mu})/(\vec{\nabla}\theta_+)^2$ that controls the MOND limit in two-field SFDM can easily be small. This is illustrated in Fig. 3.4. This is how two-field SFDM avoids the MOND limit problem discussed in Sec. 3.2.

This leaves the stability problem discussed in Sec. 3.1. If we consider only $\mathcal{L}_+$ but not $\mathcal{L}_-$, two-field SFDM has exactly the same problem as standard SFDM. To see this, consider the second-order Lagrangian for perturbations $\delta_\pm$ around background solutions $\theta^0_\pm$ for the non-relativistic equilibrium discussed

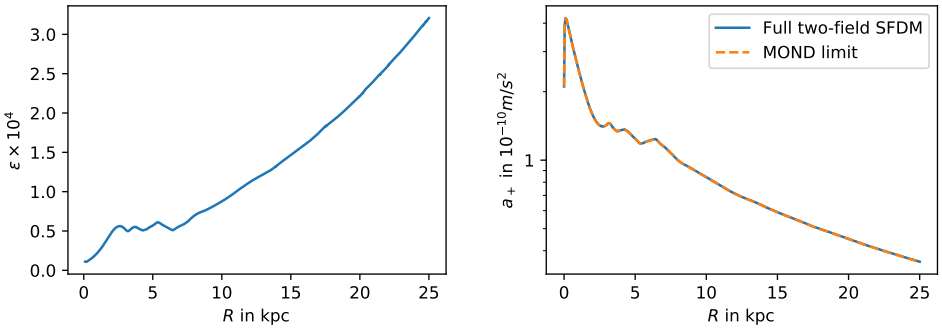

Figure 3.4.: Left: The quantity $\varepsilon = (2m\hat{\mu})/(\vec{\nabla}\theta_+)^2$ that controls the MOND limit in two-field SFDM at $z = 0$ for the same Milky Way model as in Fig. 3.3. Note the different scale of the vertical axis compared to Fig. 3.1, left. The MOND limit is much easier to reach in two-field SFDM compared to standard SFDM. Right: The acceleration a_+ due to the field θ_+ in two-field SFDM at $z = 0$ for the same Milky Way model as in Fig. 3.3. We show a_+ for the full two-field SFDM model (solid blue line) and assuming the MOND limit relation Eq. (3.7) (dashed orange line), see Ref. [73] for more details. Both are almost identical as expected from $\varepsilon \ll 1$.

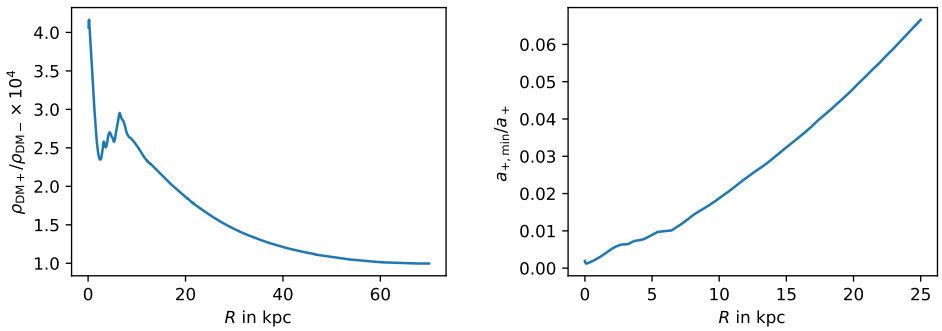

Figure 3.5.: Left: The ratio of the dark matter density $\rho_{\mathrm{DM}+}$ from $\mathcal{L}_+$ and the dark matter density $\rho_{\mathrm{DM}-}$ from $\mathcal{L}_-$ at $z = 0$ for the same Milky Way model as in Fig. 3.3. The energy density $\rho_{\mathrm{DM}-}$ dominates, which helps to avoid the MOND limit problem. Right: The acceleration a_+ relative to the minimum acceleration $a_{+\mathrm{min}}$ at $z = 0$ for the same Milky Way model as in the left panel. The instability problem is avoided for $a_{+\mathrm{min}}/a_+ < 1$ which is easily the case for $R < 25\,\mathrm{kpc}$.

above,

$$\begin{aligned}\mathcal{L}_{+,\mathrm{pert}} =& f_0' g^{00}\dot{\delta}_+^2 - \left(f_0'(\vec{\nabla}\delta_+)^2 - 2f_0''(\vec{\nabla}\theta_+^0\vec{\nabla}\delta_+)^2\right)\\ &+ \left(f_0' g^{00} + 2f_0''\mu_0^2\right)\dot{\delta}_-^2 - f_0'(\vec{\nabla}\delta_-)^2\\ &- 4f_0''\,\mu_0\dot{\delta}_-(\vec{\nabla}\theta_+^0\vec{\nabla}\delta_+)\,.\end{aligned} \tag{3.26}$$

In the MOND limit on galactic scales, spatial derivatives dominate in the argument of f so that $K_+^0 + K_-^0 - m^2 < 0$. Thus,

$$f_0' = \Lambda\sqrt{|K_+^0 + K_-^0 - m^2|} > 0\,, \tag{3.27a}$$

$$f_0'' = -\frac{\Lambda}{2}\frac{1}{\sqrt{|K_+^0 + K_-^0 - m^2|}} < 0\,, \tag{3.27b}$$

where $K_\pm^0$ are $K_\pm$ evaluated at the background solutions $\theta_\pm^0$. The third line of Eq. (3.26) does not enter the Hamiltonian since it is linear in time derivatives. Thus, it does not affect our stability analysis. The remaining terms all have the correct signs except for the $\dot{\delta}_-^2$ term. The prefactor of $\dot{\delta}_-^2$ can be negative since the $\mu_0^2 f_0''$ term dominates and $f_0'' < 0$, which leads to an instability. This is analogous to the stability problem from standard SFDM.

But in two-field SFDM, there is also $\mathcal{L}_-$. This part of the Lagrangian contains a term $\rho_-^2\dot{\theta}_-^2$. This contributes a term of order $(2m\hat{\mu}_0/\lambda_4)\dot{\delta}_-^2$ to the perturbations' Lagrangian. Here, $\hat{\mu}_0$ is the background value of $\hat{\mu}$. This fixes the instability if $2m\hat{\mu}_0/\lambda_4$ is sufficiently large. Specifically, the condition to avoid the instability is

$$a_+ > a_0\frac{\lambda_4}{\bar{\alpha}^2}\frac{m}{\hat{\mu}} \equiv a_{+\min}\,, \tag{3.28}$$

where $\vec{a}_+ = -(\bar{\alpha}\Lambda/M_{\mathrm{Pl}})\vec{\nabla}\theta_+$ is the acceleration due to θ_+. In the MOND limit and using the no-curl approximation, we have $a_+ = \sqrt{a_0 a_b}$ and this becomes

$$a_b\left(10^7\frac{\hat{\mu}}{m}\right)^2 > a_0\left(10^7\frac{\lambda_4}{\bar{\alpha}^2}\right)^2 \equiv \bar{a}\,. \tag{3.29}$$

We have introduced the factor of 10^7 since 10^{-7} is a typical value of $\hat{\mu}/m$. Thus, typically, the instability is avoided for baryonic accelerations larger than $\bar{a}$,

$$a_b \gtrsim \bar{a}\,. \tag{3.30}$$

That is, the stability problem does not occur in the superfluid cores of galaxies if we choose $\lambda_4/\bar{\alpha}^2$ sufficiently small. For our example Milky Way model, the stability is easily avoided for $R \lesssim 25\,\mathrm{kpc}$ as illustrated in Fig. 3.5, right. This is how two-field SFDM avoids the stability problem.

To summarize, two-field SFDM can avoid the problems from standard SFDM discussed above, if we choose $m^2/\bar{\alpha}$ sufficiently small (to avoid the MOND limit

problem) and $\lambda_4/\bar{\alpha}^2$ sufficiently small (to avoid the stability problem). In addition, we must choose m^4/λ_4 sufficiently large to produce a sufficiently large dark matter halo and $a_0 = \bar{\alpha}^3\Lambda^2/M_{\rm Pl}$ on the order of $10^{-10}\,\mathrm{m/s^2}$ for standard MOND phenomenology. It is possible to choose the four parameters $\bar{\alpha}$, m, λ_4, and Λ to satisfy these conditions.

These four parameters can be expressed through four other parameters that are more closely related to phenomenology,

$$\begin{aligned} r_0 &= \sqrt{\lambda_4}\,\frac{M_{\rm Pl}}{m^2}\,, \quad & \frac{\sigma}{m} &= \frac{1}{8\pi}\frac{\lambda_4^2}{m^3}\,, \\ a_0 &= \frac{\bar{\alpha}^3\Lambda^2}{M_{\rm Pl}}\,, & \bar{a} &= a_0\left(10^7\frac{\lambda_4}{\bar{\alpha}^2}\right)^2\,. \end{aligned} \tag{3.31}$$

Here, r_0 is the typical length scale of the dark matter halo, σ is the self-interaction cross-section calculated from $\mathcal{L}_-$, a_0 is the MOND acceleration scale, and $\bar{a}$ is the acceleration scale below which equilibrium solutions typically become unstable. One possible choice, discussed in Ref. [73], is

$$\begin{aligned} a_0 &= 0.87\cdot 10^{-10}\,\frac{\mathrm{m}}{\mathrm{s}^2}\,, & \frac{\sigma}{m} &= 0.01\,\frac{\mathrm{cm}^2}{\mathrm{g}}\,, \\ r_0 &= 50\,\mathrm{kpc}\,, & \bar{a} &= 10^{-12}\,\frac{\mathrm{m}}{\mathrm{s}^2}\,. \end{aligned} \tag{3.32}$$

Two-field SFDM contains two low-energy massless modes, roughly corresponding to the two fields θ_+ and θ_-. The θ_- mode corresponds to the standard superfluid phonons with sound speed

$$c_{s-}^2 = \frac{\hat{\mu}}{m}\,, \tag{3.33}$$

up to small corrections from the $f(K_+ + K_- - m^2)$ term. The θ_+ mode corresponds to the standard superluminal mode of RAQUAL models [4, 79],

$$c_{s+}^2 = 1 + \gamma^2\,, \tag{3.34}$$

where γ is the cosine of the angle between the background MOND force and the wavevector of the perturbation. This is up to small corrections due to the mixing with θ_-. This superluminality does not necessarily imply problems with causality, as discussed in Refs. [79, 111, 112].[2] Still, we can make c_{s+} subluminal by replacing $f(K_+ + K_- - m^2)$ in our Lagrangian with

$$f\left(K_+ + K_- - m^2 + C(\nabla_\alpha\theta_+\nabla^\alpha\theta_-)^2\right)\,, \tag{3.35}$$

with a constant $C \gtrsim 1/m^2$. This does not change the equilibrium equations of motion, because all additional terms are proportional to $\nabla_\alpha\theta_+\nabla^\alpha\theta_-$ which

[2] It could, however, indicate the lack of a conventional Wilsonian UV completion [113, 114]. In this case, a non-Wilsonian UV completion like classicalization may be possible [115–118].

vanishes in equilibrium. But perturbations around equilibrium do change and c_{s+} becomes subluminal, $c_{s+}^2 = (1+\gamma^2)/(1+C\mu^2)$ up to small corrections.

So far, we have discussed two-field SFDM only inside the superfluid cores of galaxies. In principle, the matching to an NFW halo outside the superfluid core can be done as in standard SFDM. But the specific choice of parameters Eq. (3.32) raises questions regarding this procedure. Specifically, the NFW radius $R_{\rm NFW}$ and the thermal radius R_T are wildly different for this choice of parameters. The NFW radius is qualitatively the same as in standard SFDM, but the thermal radius is many orders of magnitude larger.

The reason for this discrepancy between R_T and $R_{\rm NFW}$ is as follows. The energy density and pressure of the superfluid halo are comparable in standard SFDM and two-field SFDM. This is all that is needed to calculate the NFW radius. Therefore, the NFW radius is similar in both models. But the thermal radius is calculated from the self-interaction rate Γ which depends also on a different combination of parameters. Specifically, the combination $(\sigma/m)/m^4$ is much smaller in standard SFDM with the fiducial parameters from Ref. [72] than in two-field SFDM with the choice of parameters from Eq. (3.32). This makes Γ and therefore also the thermal radius much larger.

This could be avoided by choosing the parameters of two-field SFDM such that $(\sigma/m)/m^4$ is closer to its value in standard SFDM. Still, the conceptual question remains whether the NFW radius, the thermal radius, or a completely different radius determines the size of a galaxy's superfluid core. Indeed, this same problem is already present in standard SFDM, at least in principle. It was not noticed so far because R_T and $R_{\rm NFW}$ agree reasonably well with each other for the standard parameter choice from Ref. [72]. But there the self-interaction cross-section σ/m that enters Γ was chosen ad-hoc and not calculated from a Lagrangian, so it may easily be unphysical. For different choices of parameters, these two radii differ wildly also in standard SFDM. Thus, the transition of the condensed to the not-condensed phase should be revisited, both in two-field SFDM and in standard SFDM.

At face value, a large thermal radius implies that there is no transition to an NFW halo on galactic scales. This is in contrast to what is usually expected in SFDM. One possibility is that the superfluid phase extends to cosmological scales. Another possibility is that the superfluid ends at a finite radius where its density and pressure reach zero, i.e. $\hat{\mu}/m = 0$. In this case, the superfluid would resemble a giant non-relativistic boson star [119–122][3].

To sum up, we have identified three problems in the original SFDM model which share the same root cause. Namely that the phonon field θ plays a double role, carrying both the MOND-like force and the superfluid. To avoid these problem we have proposed a two-field model of superfluid dark matter.

[3] Strictly speaking, $\rho_{\rm DM-}$ does not reach zero. Boson stars don't have a surface [120]. This is because derivatives of ρ_- become important before $\rho_{\rm DM-}$ reaches zero. Above, we neglected these derivatives. They become important when $\rho_{\rm DM-}$ drops to about $(m^3/\lambda_4 M_{\rm Pl})^2$. This is much smaller than the cosmological dark matter density in our case. Thus, for practical purposes, $\rho_{\rm DM-}$ does reach zero on galactic scales.

This model works by having one field for the MOND-like force and another field for the superfluid. The galactic scale phenomenology is similar as in standard SFDM. But the transition from the condensed phase to the not-condensed phase must be reconsidered in both two-field SFDM and in standard SFDM.

4. Cherenkov radiation from stars in hybrid models

In this chapter, we will introduce a novel method to constrain hybrid MOND dark matter models with observations. In particular, hybrid models in which CDM phenomena on cosmological scales and MOND phenomena on galactic scales share a common origin, such as SFDM. As we will explain, such models often allow even non-relativistic objects like stars to lose energy through Cherenkov radiation. This is unusual. In standard modified gravity models, only high-energy cosmic rays emit Cherenkov radiation. This allows to constrain such hybrid models in a novel way. Below, we explain the general idea behind these constraints and apply our results to three specific models. For standard SFDM, we can rule out a MOND limit in the Milky Way for part of the parameter space, including the fiducial parameter values from Ref. [72]. Two-field SFDM and the SZ model avoid such constraints using two specific mechanisms that we will discuss. A short summary of these results is available in Ref. [123].

4.1. Introduction

All hybrid MOND-dark-matter models produce both MOND-like and CDM-like phenomena, as discussed in Sec. 1.3. Some hybrid models simply introduce two independent sectors, one of which provides a collisionless fluid on cosmological scales, while the other provides a MOND-like force on galactic scales. Examples are the νHDM model [68, 74] or the two-field model from Ref. [77]. In such models, the MOND and CDM phenomena exist independently of each other. In contrast, other models seek a common origin for the cosmological and galactic phenomena. Examples are the original SFDM model [71, 72] and the recently-proposed SZ model [25].

Such models with a common origin for the cosmological and galactic phenomena typically contain a component that plays a double role. That is, a component that is involved both in providing a significant energy density (for the CDM phenomenology) and in providing a modified force law (for the MOND phenomenology). Such a double role may lead to tensions between these two different roles. This happens for example in SFDM, as discussed in Chapter 3. This is why we proposed two-field SFDM, which weakens the link between the cosmological and galactic phenomena. Similarly, for the SZ model, a MOND-like force on galactic scales tends to induce a too large pressure of the DM-like fluid in the early universe if one is not careful [25]. Here,

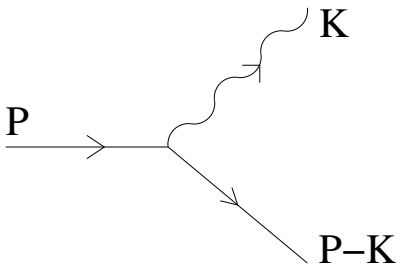

Figure 4.1.: The Feynman diagram for Cherenkov radiation. The straight lines denote matter, the wiggly line denotes the radiation mode, e.g. a massless modified gravity mode. The incoming matter particle has four-momentum P and radiates away energy and momentum K. This process is kinematically allowed only if the matter particle moves faster than the massless mode propagates. Adapted from Ref. [124].

we will discuss another consequence of components that play such a double role, namely a novel type of Cherenkov radiation.

Modified gravity theories often contain a massless mode that is directly coupled to matter in order to provide a modified force law. If this massless mode propagates with a speed c_s that is smaller than the speed of light, matter can have a velocity V that is larger than c_s. In this case, matter loses energy by emitting radiation in the form of the massless modified gravity mode. This radiation is called Cherenkov radiation. In general, Cherenkov radiation is emitted whenever a massless mode is directly coupled to matter and the matter velocity V is larger than c_s. For brevity, we will refer to the speed c_s as the speed of sound, but it is not in general necessary that there is a hydrodynamical description. In particle physics language, Cherenkov radiation corresponds to the Feynman diagram shown in Fig. 4.1. A direct coupling to matter implies that the vertex in this Feynman diagram exists. When V is larger than c_s, the external legs can go on-shell. For smaller V, this process is kinematically forbidden by energy-momentum conservation.

Usually, gravitational Cherenkov radiation is studied for highly relativistic matter [79, 124–128]. Indeed, if the sound speed is close to the speed of light, $c_s \approx 1$, only relativistic objects can emit Cherenkov radiation, because only relativistic objects can satisfy $V > c_s$. But if c_s is much smaller than the speed of light, $c_s \ll 1$, even non-relativistic matter can emit Cherenkov radiation. An example of this phenomenon is dynamical friction in collisional fluids [129–131]: If such a fluid constitutes dark matter in galaxies, then stars or globular clusters may experience dynamical friction. That is, if normal matter moves faster than the dark matter sound speed, it loses energy through Cherenkov radiation, i.e. it radiates away sound waves.[1]

[1] The standard Chandrasekhar dynamical friction for collisionless fluids [34] is related but different. E.g. no sound waves are emitted and the friction is due to the (instantaneous) Newtonian gravitational force. Also, this friction force does not drop off sharply below

The timescale on which a non-relativistic object loses a significant fraction of its energy due to this dynamical friction is [129]

$$\tau_{\rm df} = |\dot{E}|/E_{\rm kin} = \frac{V^3}{MG^2\rho_{\rm DM}}\,, \tag{4.1}$$

where we neglected $\mathcal{O}(1)$ factors, $E_{\rm kin}$ is the object's kinetic energy, M is its mass, and $\rho_{\rm DM}$ is the dark matter density. This timescale is dynamically relevant for heavy objects since $\tau_{\rm df} \propto 1/M$. For example, globular clusters in the Fornax dwarf spheroidal may have a timescale $\tau_{\rm df}$ smaller than the age of the Universe [132, 133]. In comparison, individual stars have much smaller masses so that the timescale $\tau_{\rm df}$ is much larger and not phenomenologically relevant. For example, $\tau_{\rm df} \sim 10^{20}\,{\rm yr}$ for the Sun in a dark matter fluid with $\rho_{\rm DM} = 0.3\,{\rm GeV/cm^3}$.

A standard dark matter fluid is not directly coupled to matter, only indirectly through Newtonian gravity. This is why $\tau_{\rm df}$ scales as $1/G^2$. If, instead, matter were directly gravitationally coupled to the dark matter fluid, this timescale would scale as $1/G$. This is typical for gravitational Cherenkov radiation in modified gravity models. In this case, the relevant timescale may be much shorter so that even comparably light objects can lose energy on a dynamically relevant timescale.

Hybrid MOND-dark-matter models with a common origin for MOND and CDM phenomena typically allow for Cherenkov radiation. The reason is as follows. These models have a massless mode that couples to normal matter, namely the mode that carries the MOND-like force. If there is a common origin of this MOND-like force and the cosmological DM-like fluid, this massless mode naturally propagates with a non-relativistic speed. A non-relativistic propagation speed is natural because the dark matter fluid must be pressureless on cosmological scales.

For example, in SFDM, the collisionless fluid on cosmological scales is a non-relativistic superfluid, which also forms a cored halo on galactic scales. The MOND-like force is carried by the phonons of this superfluid. Thus, these phonons are non-relativistic and directly coupled to normal matter. This allows for Cherenkov radiation. Since the sound speed is non-relativistic, this allows Cherenkov radiation even from non-relativistic objects. That is, even stars may lose energy through Cherenkov radiation, i.e. to phonons.

This is a novel type of Cherenkov radiation. It comes from a direct coupling to matter as in standard modified gravity models. But it also allows for non-relativistic objects to emit Cherenkov radiation like dynamical friction in collisional fluids.

We can constrain models allowing for this type of Cherenkov radiation by requiring that stars do not lose a significant amount of their energy on galactic timescales. Specifically, either the sound speed c_s must be large enough so that most stars are subsonic and Cherenkov radiation is avoided kinematically. Or

a critical velocity. Still, the order of magnitude of the friction force can be the same in the collisional and the collisionless case [129].

the timescale τ_E on which stars lose a significant amount of energy due to Cherenkov radiation must be much larger than galactic timescales. We will show that the timescale τ_E scales as

$$\tau_E \propto \frac{1}{c_s^2\, g_m^2 G}\,, \tag{4.2}$$

where c_s is the non-relativistic sound speed of the massless mode and $g_m\sqrt{G}$ its coupling to matter. Thus, to avoid Cherenkov radiation by having a large energy loss timescale τ_E one can make either c_s or g_m small. This is different from the relativistic limit considered in standard Cherenkov radiation constraints, where τ_E cannot be made large by making c_s small.

Our results for τ_E will be a conservative lower bound, because we assume strict cuts in the phase space integrals in our calculation. We do this for two reasons. First, to avoid technical difficulties due to the non-linearities inherent in any MOND model. Second, to rely only on the MOND regime of each model, i.e. to be independent of the behavior for accelerations larger than the MOND acceleration scale a_0. This makes our calculation simple and robust. For example, we do not need to consider potential higher-derivative terms [24] which may become important in the solar system, as has been considered in both SFDM and the SZ model [25, 72].

These strict cuts also imply that there is no significant recoil. In a quasi-particle picture: Any single emission carries very little energy and momentum. A macroscopic change happens only after a large number of these soft emissions. In contrast, for standard gravitational Cherenkov radiation, the emissions are not soft. Supersonic objects lose a significant amount of energy and momentum with each emission, i.e. recoil is important [128].

We first demonstrate the general idea and calculations behind our Cherenkov constraints for a simple prototype model in Sec. 4.2. In Secs. 4.3, 4.4, and 4.5 we apply these results, with appropriate modifications, to the original SFDM model, two-field SFDM, and the SZ model, respectively. For the original SFDM model, we can rule out a MOND limit in the Milky Way for a certain part of the parameter space. In contrast, two-field SFDM and the SZ model evade our constraints due to two different mechanisms that we explain in detail. After a short discussion in Sec. 4.6, we summarize in Sec. 4.7.

4.2. Cherenkov radiation in a prototype hybrid model

In this section, we will introduce the general idea behind our Cherenkov radiation constraint using a prototype model. Specifically, we consider a prototype model for a massless mode with non-relativistic sound speed $c_s \ll 1$ that directly couples to matter with a standard gravitational coupling constant. As discussed above, this is a natural setup for hybrid MOND dark matter models with a common origin for the cosmological and galactic scale phenomena. The three models we explicitly discuss below in Secs. 4.3, 4.4, and 4.5 are not

exactly of the form discussed here. However, we prefer to keep the discussion simple for now and later adapt our results as needed.

We consider a matter object, for example a star, with mass M and velocity V at a distance R_p from the center of its host galaxy. We denote the host galaxy's baryonic mass inside a radius R_p by $M_{\rm gal}$. In general, we use a capital R for distances to the center of the host galaxy and a lower case r for distances to the matter object, also referred to as the perturber.

Our main result will be an estimate for the timescale τ_E on which non-relativistic objects like stars lose a significant fraction of their energy, if their velocity V is larger than a critical velocity $V_{\rm crit} = \mathcal{O}(\bar{c})$, i.e. if they are supersonic,

$$\tau_E = \frac{2\cdot 10^8\,\mathrm{yr}}{f_a f_p^2\, g_m^2}\cdot\left(\frac{V/\bar{c}}{2}\right)^2\cdot\left(\frac{a_0}{a_b^{\rm gal}}\right)\cdot\left(\frac{V}{200\,\mathrm{km/s}}\right)\cdot\left(\frac{1.2\cdot 10^{-10}\,\mathrm{m/s^2}}{a_0}\right). \tag{4.3}$$

Here, $\bar{c}$ sets the scale of the sound speed c_s, g_m is a model-dependent coupling constant, a_0 is the MOND acceleration scale, and $a_b^{\rm gal}$ is the Newtonian acceleration due to baryons in the host galaxy at the position of the matter object. The factor f_a depends on the relative orientation of $\vec{V}$ and the MOND force of the background galaxy. The factor f_p is determined by the radius where the perturber's field becomes smaller than the galaxy's background field. Both f_a and f_p are typically of order 1. For a standard gravitational coupling, g_m is also of order 1. Details are explained below.

Thus, stars with $V > V_{\rm crit}$ may lose a significant fraction of their energy on galactic timescales. This constrains hybrid MOND-DM models. Such models either need a sound speed large enough such that most stars are subsonic, or a timescale τ_E that is much larger than galactic timescales.

4.2.1. Prototype Lagrangian

Consider a model with a real scalar field φ that satisfies a MOND-type equation on galactic scales in the static limit. Then, roughly, $\varphi \propto \sqrt{GM_{\rm gal}a_0}\ln(R)$. Typically, the Lagrangian for perturbations δ around a galaxy's static background field φ_0 can be written as

$$\mathcal{L} = \frac{1}{2}\frac{1}{\bar{c}^2}(\partial_t\delta)^2 - \frac{1}{2}\left((\vec{\nabla}\delta)^2 + (\hat{a}\vec{\nabla}\delta)^2\right) - \frac{g_m}{\sqrt{2}M_{\rm Pl}}\delta\,\delta_b\,, \tag{4.4}$$

after an appropriate definition of δ. We will discuss concrete examples of this later. Here, g_m and $\bar{c}$ are constants that may depend on the background field φ_0, and δ_b is a perturbation of the baryonic density ρ_b. Further, $\hat{a}$ is a unit vector that points into the direction of the background $\vec{\nabla}\varphi_0$, i.e. into the direction of the background MOND force or in the opposite direction, depending on the signs. The dispersion relation is $\omega = c_s|\vec{k}|$ with the sound speed

$$c_s^2 = \bar{c}^2(1+\gamma^2)\,, \tag{4.5}$$

where γ is the cosine of the angle between the perturbation's wavevector $\vec{k}$ and $\hat{a}$.

As mentioned in Sec. 1.4, a prototypical MOND Lagrangian is

$$\mathcal{L} = \frac{2M_{\rm Pl}^2}{3a_0}\sqrt{|K_\varphi|}K_\varphi - \rho_b\,\varphi\,, \tag{4.6}$$

where $K_\varphi = \nabla_\alpha\varphi\nabla^\alpha\varphi$ is a standard kinetic term and a_0 is the MOND acceleration scale. After rescaling, this gives the Lagrangian Eq. (4.4) with

$$\bar{c} = 1\,, \quad g_m = \sqrt{\frac{a_0}{|\vec{\nabla}\varphi_0|}}\,, \tag{4.7}$$

as the second-order Lagrangian for perturbations δ. Thus, the sound speed is relativistic (even superluminal [79, 111, 112]) and g_m is roughly of order 1 on galactic scales since $|\vec{\nabla}\varphi_0|$ is roughly of order a_0 there.

In many hybrid MOND-DM models, we expect this to be still qualitatively true, except that $\bar{c}$ will give a non-relativistic sound speed

$$\bar{c} \ll 1\,. \tag{4.8}$$

For example, in standard SFDM, δ corresponds to the phonons of a non-relativistic superfluid. These have a non-relativistic sound speed so that $\bar{c} \ll 1$. We will discuss this and other examples in more detail below.

As already mentioned above, the models we discuss below do not exactly have the form of our prototype Lagrangian. For example, standard SFDM has an additional term that mixes spatial and time derivatives of δ. Two-field SFDM has a more complicated coupling g_m and no $\hat{a}$ term for the relevant mode. The SZ model also has a more complicated coupling g_m. We will discuss how the calculation needs to be adjusted in each case in Secs. 4.3, 4.4, and 4.5, respectively. Nevertheless, the general considerations for our prototype Lagrangian give a useful qualitative picture.

4.2.2. Cherenkov radiation

Consider a non-relativistic perturber, e.g. a star, that moves in a galaxy. Following Refs. [130, 131], we model the perturber as a real scalar field χ with mass M. In the non-relativistic limit, this scalar field's energy density is

$$\rho_\chi \approx \frac{1}{2}(\dot{\chi}^2 + M^2\chi^2) \approx M^2\chi^2\,. \tag{4.9}$$

We take this as the baryon density's perturbation δ_b in our prototype Lagrangian from Eq. (4.4). We can then calculate the perturber's energy loss rate

$$\dot{E} = -\int \omega d\Gamma\,, \tag{4.10}$$

where $d\Gamma$ is the differential decay rate and $\omega = c_s|\vec{k}|$ is the energy of the quasi-particle that is radiated away. It is important to impose cutoffs in the integral in Eq. (4.10) that reflect the regime in which our prototype Lagrangian is valid. For now, we keep the calculation general and assume cutoffs $k_{\min}$ and $k_{\max}$ with $k_{\min} < k_{\max}$ in the momentum integral. We will discuss the numerical values of these cutoffs below. We find (see Appendix A),

$$|\dot{E}| = f_a \frac{\bar{c}^2}{16\pi V} \frac{g_m^2 M^2}{M_{\rm Pl}^2} (k_{\max}^2 - k_{\min}^2) \cdot \Theta(V - V_{\rm crit}) \,. \tag{4.11}$$

Here, f_a is a factor that depends on the direction of $\hat{a}$ relative to $\vec{V}$. The critical velocity $V_{\rm crit}$ also depends on the direction of $\vec{V}$ and denotes the velocity below which no Cherenkov radiation is emitted. For $\vec{V} \parallel \hat{a}$:

$$f_a^{\parallel} = \frac{1}{1 - (\bar{c}/V)^2} \,, \tag{4.12a}$$

$$V_{\rm crit}^{\parallel} = \sqrt{2}\bar{c} \,. \tag{4.12b}$$

And for $\vec{V} \perp \hat{a}$:

$$f_a^{\perp} = \frac{1}{\sqrt{1 + (\bar{c}/V)^2}} \,, \tag{4.13a}$$

$$V_{\rm crit}^{\perp} = \bar{c} \,. \tag{4.13b}$$

The factor f_a does not change the order of magnitude of the energy loss. Numerically, for $V > V_{\rm crit}$, it varies between 1 and 2 for $\vec{V} \parallel \hat{a}$ and between 1 and $1/\sqrt{2}$ for $\vec{V} \perp \hat{a}$.

Note that $\dot{E}$ scales as $\bar{c}^2$. This is because the correct normalization for using the standard QFT formalism with the standard Feynman rules is not the one shown in Eq. (4.4). Instead, the correct normalization is to scale the $1/\bar{c}^2$ in front of the time derivatives away [130]. This scaling makes the coupling proportional to $\bar{c}$ and the amplitude squared proportional to $\bar{c}^2$.

The critical velocity $V_{\rm crit}$ reflects the fact that Cherenkov radiation is allowed only for supersonic perturbers, $V > c_s$. Here, the sound speed c_s depends on the direction of the wave vector $\vec{k}$ of the perturbation relative to that of the background field $\hat{a}$, i.e. $c_s = \bar{c}\sqrt{1+\gamma^2}$ depends on γ. This is why f_a and $V_{\rm crit}$ depend on the relative orientation of $\vec{V}$ and $\hat{a}$. Without the $\hat{a}$-term in our prototype Lagrangian, this would just be $f_a = 1$ and $V_{\rm crit} = \bar{c}$, independently of the direction of $\vec{V}$.

Our result for the energy loss $|\dot{E}|$ scales differently from what Refs. [124–126, 134] have found for standard gravitational Cherenkov radiation. This is because these consider the limit of a relativistic perturber and sound speed, $V \approx 1$ and $c_s \approx 1$, while we consider a non-relativistic perturber and sound speed, $V \ll 1$ and $c_s \ll 1$.

The timescale on which perturbers like stars lose a significant amount of their energy E is roughly $E/|\dot{E}|$. Thus, assuming $V > V_{\rm crit}$, we define the

timescale τ_E as

$$\tau_E \equiv \frac{E_{\rm kin}}{|\dot{E}|} \equiv \frac{\frac{1}{2}MV^2}{|\dot{E}|} = \frac{8\pi V^3 M_{\rm Pl}^2}{f_a \bar{c}^2 g_m^2 M k_{\rm max}^2} \frac{1}{1-(k_{\rm min}/k_{\rm max})^2}\,. \tag{4.14}$$

As alluded to in the introduction, this scales as $1/G \propto M_{\rm Pl}^2$, which is different from the standard dynamical friction timescale that scales as $1/G^2 \propto M_{\rm Pl}^4$ [34, 129, 130]. Stars have both kinetic and potential energy. For simplicity, our definition of τ_E includes only the kinetic energy. Including the gravitational energy does not change the order of magnitude. Indeed, we will now discuss a concrete example to see that τ_E is a useful quantity.

For a star in a galaxy, we have

$$\partial_t \left(E_{\rm kin} + E_{\rm grav}\right) = \dot{E}\,, \tag{4.15}$$

where $\dot{E}$ is the energy loss through Cherenkov radiation, and $E_{\rm grav}$ is the gravitational energy. If $E_{\rm grav}$ depends only on the star's position R_p in the galaxy, we have

$$\partial_t \left(\frac{1}{2}MV^2\right) + M\dot{R}_p\, a_{\rm grav} = \dot{E}\,, \tag{4.16}$$

where $a_{\rm grav}$ is the gravitational acceleration produced by the host galaxy at the star's position. In the MOND-dominated part of the galaxy with a flat rotation curve, we have $a_{\rm grav} = \sqrt{GM_{\rm gal}a_0}/R_p$. If we further assume that the star is always approximately on a circular orbit, we have $V^2 = \sqrt{GM_{\rm gal}a_0}$, which is constant. This gives

$$\frac{\dot{R}_p}{R_p} = -\frac{1}{2\tau_E}\,. \tag{4.17}$$

As long as τ_E depends only weakly on R_p, this means that stars transition to smaller galactic radii as $\exp(-t/2\tau_E)$ due to Cherenkov radiation.

If this calculation is correct, stars in the flat part of the rotation curve lose energy by transitioning to smaller galactic radii, not by losing their velocity. This is due to the assumption that the star's velocity is always the circular velocity, even as it transitions to smaller radii. In the next subsection, we show numerically that this assumption is justified.

4.2.3. Orbits of stars emitting Cherenkov radiation

Here, we numerically investigate how stars orbit around a galaxy under the influence of the non-relativistic Cherenkov radiation derived above. We model the effect of the Cherenkov radiation as a friction force that reproduces the energy loss $\dot{E}$ calculated above. Modelling the Cherenkov radiation as a friction force is justified because the energy loss happens through a large number of emissions, each of which carries away only a very small fraction of the star's energy. This is different from the high-energy Cherenkov radiation emitted

from ultra-relativistic protons discussed in Refs. [124–126, 134], where a single emission carries away a significant fraction of the proton's energy and thus produces a significant recoil [128].

For simplicity, we consider the deep-MOND regime of a galaxy and we consider orbits $\vec{X}(t)$ that would be circular without the energy loss due to Cherenkov radiation. Concretely, we consider initial conditions $Y(t_0) = R_0$, $X(t_0) = Z(t_0) = 0$, $\dot{X}(t_0) = V_0$, and $\dot{Y}(t_0) = \dot{Z}(t_0) = 0$, with equation of motion

$$\ddot{\vec{X}} = \vec{a}_{\rm grav} - \left(R_\eta \cdot \dot{\vec{X}}\right) \eta \,. \tag{4.18}$$

Here, V_0 is the MOND circular velocity $(GM_{\rm gal}a_0)^{1/4}$, $\vec{a}_{\rm grav}$ is the MOND gravitational acceleration, and η is the coefficient of the friction force due to Cherenkov radiation. The direction of this friction force is that of $R_\eta \dot{\vec{X}}$, where R_η is a rotation matrix.

For our prototype Lagrangian, the friction force points in the direction of $\vec{V}$ for $\vec{V} \perp \hat{a}$ and $\vec{V} \parallel \hat{a}$, i.e. R_η is just the identity matrix. For a general orientation of $\vec{V}$, the force is rotated within the X-Y plane. This is discussed in Appendix A.3. For standard SFDM, the force is rotated in the X-Y plane even for the special case $\vec{V} \perp \hat{a}$. This is because spatial and time derivatives are mixed there, see Sec. 4.3 and Appendix C.5.

We can find the coefficient η by contracting Eq.(4.18) with $\dot{\vec{X}} = \vec{V}$ and then comparing to the energy balance equation

$$\partial_t \left(\frac{1}{2} M V^2 + E_{\rm grav}\right) = \dot{E}_{\rm Cherenkov} \,, \tag{4.19}$$

where $E_{\rm grav}$ is the gravitational energy and $\dot{E}_{\rm Cherenkov}$ is the energy loss rate due to Cherenkov radiation. This gives

$$\eta = \frac{|\dot{E}_{\rm Cherenkov}|}{M(\vec{V}^T R_\eta \vec{V})} = \frac{1}{2\tau_E \cos\theta_\eta} \,, \tag{4.20}$$

where θ_η is the angle by which R_η rotates the vector $\vec{V}$ in the X-Y plane. Since τ_E scales as V^3, we choose to write

$$\ddot{\vec{X}} = \vec{a}_{\rm grav} - \frac{R_\eta \cdot \dot{\vec{X}}}{2\cos\theta_\eta V^3} \cdot \frac{V^3}{\tau_E} \,, \tag{4.21}$$

where the factor V^3/τ_E is independent of V and is time-dependent only through factors like $a_0/a_b^{\rm gal}$ that occur in τ_E in some models and depend on the star's position in the galaxy. For simplicity, we assume such factors to be roughly constant, but we have numerically verified that the same conclusions hold also if we include R_p-dependent factors in τ_E.

We first consider the case where $\theta_\eta = 0$ that is relevant for our prototype Lagrangian for circular orbits. The resulting orbits will not be exactly circular,

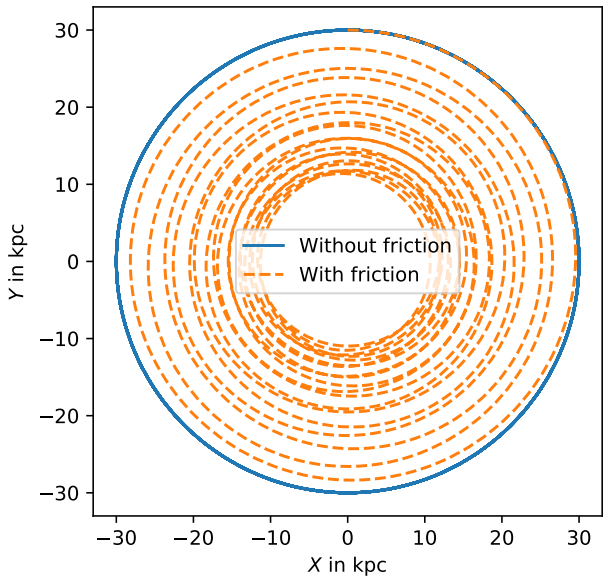

Figure 4.2.: Orbit of a perturber in the MOND regime of a galaxy with (dashed orange line) and without (solid blue line) friction due to Cherenkov radiation. The friction corresponds to a timescale $\tau_E = 5 \cdot 10^9\,\mathrm{yr}$ and acts in the direction of the velocity, i.e. $\theta_\eta = 0$. The initial conditions are $X(0) = 30\,\mathrm{kpc}$, $Y(0) = 0$, and $V_0 = 200\,\mathrm{km/s}$. The galaxy mass is chosen such that these initial conditions give a circular orbit without friction. The orbits are integrated for a total time of $10^{10}\,\mathrm{yr}$.

that is we do not have $\vec{V} \perp \hat{a}$ exactly. But we expect this to be a reasonable approximation for the numerical values we consider. For very large τ_E, it is clear that the friction term does not have any effect. Similarly, it is clear that stars will quickly fall in to the center of the galaxy for very small τ_E. The most interesting case is when τ_E is roughly of the order of galactic timescales. Thus, here we choose as numerical values

$$R_0 = 30\,\mathrm{kpc}\,, \quad V_0 = 200\,\mathrm{km/s}\,, \quad \frac{\tau_E}{V^3} = \frac{5 \cdot 10^9\,\mathrm{yr}}{(200\,\mathrm{km/s})^3}\,. \tag{4.22}$$

We do not have to choose a mass for the star since the equations are independent of this mass. Our choice of V_0 and R_0 also fixes the host galaxy's mass, since we require orbits to be circular in the absence of Cherenkov radiation. Specifically, $\sqrt{GM_{\mathrm{gal}}a_0}/R_p = V_0^2/R_p$.

We numerically integrate the equation of motion with these parameters for $10^{10}\,\mathrm{yr}$ in Mathematica [103]. The resulting orbit in the X-Y plane is shown in Fig. 4.2. We see that the star transitions to lower galactic radii due to the friction force. Fig. 4.3, left, shows that the mean velocity of the star stays

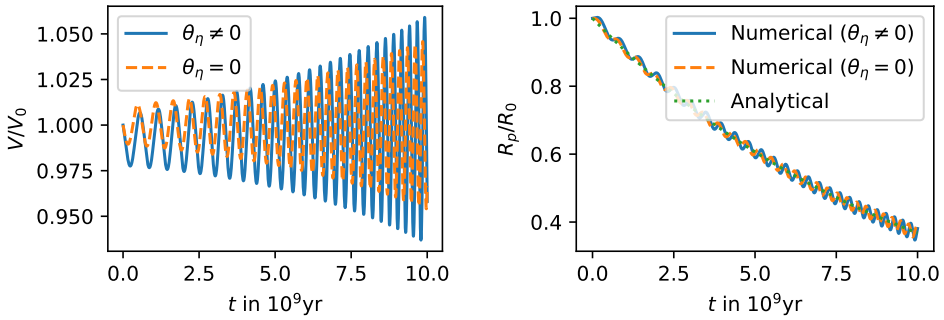

Figure 4.3.: Left: Velocity of a perturber losing energy through Cherenkov radiation relative to its initial velocity. The dashed orange line corresponds to the orbit with friction from Fig. 4.2 ($\theta_\eta = 0$). The solid blue line shows an orbit with the same parameters but with a force that is rotated according to Eq. (4.24) ($\theta_\eta \neq 0$). Right: The perturber's position R_p in the host galaxy relative to its initial position for the same orbits as in the left panel (solid blue line for $\theta_\eta = 0$ and dashed orange line for $\theta_\eta \neq 0$), and additionally for the analytical estimate Eq. (4.23) (dotted green line).

roughly constant during this transition, but there are small oscillations on the percentage level. Fig 4.3, right, shows the star's distance to the host galaxy's center $R_p = |\vec{X}|$ for both the numerical calculation and for our analytical estimate from Sec. 4.2.2,

$$\frac{\dot{R}_p}{R_p} = -\frac{1}{2\tau_E}\,. \tag{4.23}$$

Up to small oscillations, the numerical and analytical results agree with each other. This justifies the analytical estimate from Sec. 4.2.2.

Consider now the case $\theta_\eta \neq 0$. For concreteness we take θ_η from standard SFDM, where $\theta_\eta = 0$ for $\vec{V} \parallel \hat{a}$, but $\theta_\eta \neq 0$ for $\vec{V} \perp \hat{a}$, as shown in Appendix C.5. In both cases, the force stays within the X-Y plane, so that we do not need to include a Z component. For $\vec{V} \perp \hat{a}$, we have

$$\tan\theta_\eta = -\frac{1}{2}\frac{f_{\bar{\beta}} V}{\bar{c}}\,, \tag{4.24}$$

where $f_{\bar{\beta}}$ is a constant that depends on the model parameter $\bar{\beta}$, see Sec. 4.3. The orbit we consider is mostly, but not exactly, circular. Still, we expect that using Eq. (4.24) is a reasonable approximation. We further adopt

$$f_{\bar{\beta}} = 1/\sqrt{3}\,, \quad \bar{c} = 50\,\mathrm{km/s}\,. \tag{4.25}$$

The value of $f_{\bar{\beta}}$ is the largest possible value for the parameter range we consider, see Sec. 4.3. The value for $\bar{c}$ is unusually small. This gives an unusually

large $\tan\theta_\eta$. Despite this, the resulting orbit is very similar to that for $\theta_\eta = 0$. This can be seen from the dashed orange lines in Fig. 4.3. The oscillations differ slightly, but the qualitative behavior is the same.

As mentioned above, there may be R_p-dependent factors in τ_E that we have so far neglected. We have numerically verified that our conclusions hold even with an R_p-dependent τ_E. That is, Eq. (4.23) accurately captures the transition to smaller radii even in this more complicated case. Of course, it may not be possible to integrate Eq. (4.23) analytically depending on its specific R_p-dependence.

4.2.4. Regime of validity

We now come back to the question of how to choose the cutoffs $k_{\min}$ and $k_{\max}$ in our calculation of the energy loss due to Cherenkov radiation. We will argue that reasonable values are

$$k_{\min} \sim 1/\text{kpc} \sim 10^{-26}\,\text{eV}\,, \tag{4.26a}$$

$$k_{\max} \sim f_p \cdot \sqrt{\frac{a_0}{GM}}\sqrt{\frac{a_b^{\text{gal}}}{a_0}} \sim 10^{-22}\,\text{eV}\cdot f_p \cdot \sqrt{\frac{a_b^{\text{gal}}}{a_0}}\,. \tag{4.26b}$$

Here, $k_{\min}$ is determined by the scale on which the background field φ_0 varies. In galaxies, this is typically $\mathcal{O}(\text{kpc})$. The upper cutoff $k_{\max}$ is determined by the distance from the perturber where its field drops below the galaxy's background field, i.e. the distance from where on we can actually treat the perturber's field as a perturbation. This is related to the perturber's MOND radius $r_{\text{MOND}} = \sqrt{GM/a_0}$ outside of which the MOND regime begins. It's also related to the galaxy's Newtonian baryonic acceleration $a_b^{\text{gal}} = GM_{\text{gal}}/R_p^2$ at the perturber's position. For $M = M_\odot$, we have $r_{\text{MOND}} \approx 3\cdot 10^{-5}\,\text{kpc}$, corresponding to $k_{\max} \approx 10^{-22}\,\text{eV}$. Since we assume the galaxy to be in the MOND regime for our calculations, the factor $\sqrt{a_b^{\text{gal}}/a_0}$ is smaller than 1, but large enough such that $k_{\max}$ is still much larger than $k_{\min}$. The factor f_p is a model-dependent correction factor.

Our method of calculation gives a nonzero energy loss only for $k_{\max} > k_{\min}$. This is because these are the limits of a momentum integral, see Appendix A. Thus, our method gives useful results only for $M \lesssim 10^8\,M_\odot$. This means we can apply our results to stars and also to many globular clusters. But most dwarf galaxies are too heavy. In the following, we will mainly consider stars.

We choose the cutoffs $k_{\min}$ and $k_{\max}$ to ensure that our prototype Lagrangian from Eq. (4.4) is valid. The conditions for this are:

- Our calculation assumes that g_m and $\bar{c}$ do not vary on the scales under consideration. Also, in deriving a Lagrangian like Eq. (4.4) for perturbations, one usually neglects derivatives of the background fields against derivatives of the perturbations. This restricts us to scales smaller than $\sim$ kpc since the galactic background varies on this scale. This means we

need to choose

$$k_{\min} \gtrsim 1/\mathrm{kpc}\,. \tag{4.27}$$

- Close to a star, the baryonic acceleration is larger than the MOND acceleration scale a_0 and we are in the Newtonian regime. Depending on the model, other terms in the Lagrangian may become important there. For example, higher derivative terms might become important [24]. Thus, our prototype Lagrangian may not be valid in this regime. Or it may be valid but with different parameters g_m and $\bar{c}$. This Newtonian regime ends at a radius of about

$$r_{\mathrm{MOND}} = \sqrt{GM/a_0}\,. \tag{4.28}$$

  This means we should restrict our calculation to larger scales, i.e. we should choose

$$k_{\max} \lesssim \frac{1}{r_{\mathrm{MOND}}}\,. \tag{4.29}$$

- Close to a star, the field due to the star dominates compared to the background field of the host galaxy. Thus, we are not allowed to expand around the galaxy's background field. This regime ends roughly at a radius

$$r_{\mathrm{pert}} = R_p\sqrt{M/M_{\mathrm{gal}}}\,. \tag{4.30}$$

  We can write this as

$$r_{\mathrm{pert}} = r_{\mathrm{MOND}}\sqrt{\frac{a_0}{a_b^{\mathrm{gal}}}}\,. \tag{4.31}$$

  Thus, we should choose

$$k_{\max} \lesssim \frac{1}{r_{\mathrm{MOND}}}\sqrt{\frac{a_b^{\mathrm{gal}}}{a_0}}\,. \tag{4.32}$$

- The estimate Eq. (4.30) gives the radius where the static field of a static perturber becomes smaller than the host galaxy's background field. However, we are interested in a dynamical situation where Cherenkov radiation is emitted. It is possible that a different condition controls when the dynamical field due to the perturber is small compared to the galaxy's static background field. In general, this cannot be estimated in a model-independent way. Thus, we introduce a factor $f_p \lesssim 1$ to allow for a smaller cutoff

$$k_{\max} \lesssim f_p\frac{1}{r_{\mathrm{MOND}}}\sqrt{\frac{a_b^{\mathrm{gal}}}{a_0}}\,. \tag{4.33}$$

For standard SFDM, we have explicitly checked which values of f_p are allowed, see Appendix C.6. We find that we can set $f_p = 1$ for $\vec{V} \perp \hat{a}$,

$$f_p^{\perp} = 1\,. \tag{4.34}$$

For $\vec{V} \parallel \hat{a}$, we may need to decrease f_p a bit. The worst case is for parallel (rather than anti-parallel) $\vec{V}$ and $\hat{a}$ with $\bar{\beta} = 3/2$ and a velocity V that is just barely larger than the critical velocity. In this case, we estimate

$$\left(f_p^{\parallel}\right)^{\text{worst case}} \approx 1/\sqrt{2.8}\,. \tag{4.35}$$

For larger values of $\bar{\beta}$, larger velocities, or for anti-parallel $\vec{V}$ and $\hat{a}$, we can choose $f_p^{\parallel}$ much closer to 1. As explained in Appendix C.4, introducing cutoffs in the way we do means we cannot take $V \to V_{\text{crit}}$ without perturbations becoming large. Thus, here we restrict ourselves to V at least 1% away from V_{crit}. This does not significantly affect any of our conclusions.

- We need to choose $k_{\text{max}} < 2M(V - \bar{c}\sqrt{2})$. This is the standard cutoff in Cherenkov radiation calculations, see Appendix A. However, numerically, this is not relevant in our case because the other cutoffs we considered are much more restricting. Concretely, we have $2MV \sim 10^{63}\,\text{eV} \cdot (M/M_\odot) \cdot (V/200\text{km/s})$. This is the reason our energy loss behaves like a friction force without a significant recoil. We restrict k to be much smaller than the perturber's momentum MV.

In principle, there can be additional model-dependent cuts, i.e. further restrictions on the validity of our prototype Lagrangian. Here, we assume that there are no such cuts or, at least, that they are less restricting than the cuts we already considered above.

As long as $k_{\text{min}} \ll k_{\text{max}}$, the result Eq. (4.14) for the timescale τ_E is dominated by k_{max}. Then, with our particular choice of k_{min} and k_{max},

$$\tau_E = \frac{V^3}{f_a f_p^2 \bar{c}^2 g_m^2 a_0} \cdot \frac{a_0}{a_b^{\text{gal}}}\,. \tag{4.36}$$

This is independent of the perturber's mass, but depends on the position of the perturber in the galaxy. Numerically,

$$\tau_E = \frac{2 \cdot 10^8\,\text{yr}}{f_a f_p^2\, g_m^2} \cdot \left(\frac{V/\bar{c}}{2}\right)^2 \cdot \left(\frac{a_0}{a_b^{\text{gal}}}\right) \cdot \left(\frac{V}{200\,\text{km/s}}\right) \cdot \left(\frac{1.2 \cdot 10^{-10}\,\text{m/s}^2}{a_0}\right)\,. \tag{4.37}$$

Thus, for g_m of order 1 and $V > V_{\text{crit}}$, stars in galaxies lose a significant fraction of their energy on timescales τ_E which are not much larger than typical galactic timescales.

As already mentioned above, our result for $|\dot{E}|$ is a conservative lower bound. The actual energy loss may be higher. In particular, modes with $k > k_{\text{max}}$ and $k < k_{\text{min}}$ may contribute to the energy loss, but are not considered here.

4.2.5. Constraints from Cherenkov radiation

The energy loss from Cherenkov radiation rules out models unless either their $V_{\rm crit}$ is large enough to make most stars subsonic or their τ_E is large enough so that supersonic stars do not lose much energy on galactic timescales. That is, we need for most stars

$$V > V_{\rm crit} \quad \text{or} \quad \tau_E > \tau_{\rm min}\,, \tag{4.38}$$

for some reasonable minimum timescale $\tau_{\rm min}$. We will now make this more concrete.

Consider first the rotation curve of a disk galaxy with approximately circular orbits. In this case, we can assume $\vec{V} \perp \hat{a}$ since $\hat{a}$ (or $-\hat{a}$) points to the center of the galaxy. Then,

$$V_{\rm crit} = \bar{c}\,, \quad \tau_E = \frac{V^3}{f_a^\perp f_p^2 \bar{c}^2 g_m^2 a_0} \frac{a_0}{a_b^{\rm gal}}\,, \tag{4.39}$$

with $f_a^\perp = \sqrt{1+(\bar{c}/V)}^{-1}$. Below, it will be useful to allow a more general form of $V_{\rm crit}$ and $f_a^\perp$,

$$V_{\rm crit} = f_V\,\bar{c}\,, \quad f_a^\perp = \frac{1}{A\sqrt{1+B(\bar{c}/V)^2}}\,, \tag{4.40}$$

for some constants f_V, A, and B. The condition $\tau_E > \tau_{\rm min}$ can then be written as

$$\frac{\sqrt{1+B(\bar{c}/V)^2}}{(\bar{c}/V)^2} > \tau_{\rm min} \frac{f_p^2 g_m^2 a_0}{V} \frac{a_b^{\rm gal}}{a_0} \frac{1}{A} \equiv X\,. \tag{4.41}$$

In terms of $\bar{c}$,

$$\bar{c} < V \cdot \sqrt{\frac{B+\sqrt{B^2+4X^2}}{2X^2}} \equiv V \cdot f_X\,. \tag{4.42}$$

Thus, for rotation curves, the two conditions from Eq. (4.38) become

$$\bar{c} > V/f_V \quad \text{or} \quad \bar{c} < V \cdot f_X\,. \tag{4.43}$$

A set of model parameters is ruled out if the measured rotational velocity V satisfies neither of these two conditions. Specifically, each measured point in a rotation curve rules out $\bar{c}$ in the interval

$$(f_X V, V/f_V)\,, \tag{4.44}$$

because if $\bar{c}$ lies in this interval, neither $\bar{c} > V/f_V$ nor $\bar{c} < V f_X$ can be satisfied. This constraint applies only for $f_X < 1/f_V$, of course. That is, we get a constraint only for

$$X > f_V \sqrt{B+f_V^2}\,, \tag{4.45}$$

which translates to a maximum velocity,

$$V < f_p^2 g_m^2 \frac{\tau_{\min} a_0}{A f_V \sqrt{B + f_V^2}} \frac{a_b^{\text{gal}}}{a_0} \,. \tag{4.46}$$

For $A = B = f_V = 1$, this is,

$$V < 2700 \, \frac{\text{km}}{\text{s}} \cdot f_p^2 g_m^2 \cdot \left(\frac{a_b^{\text{gal}}}{a_0} \right) \cdot \left(\frac{\tau_{\min}}{10^9 \, \text{yr}} \right) \cdot \left(\frac{a_0}{1.2 \cdot 10^{10} \, \frac{\text{m}}{\text{s}^2}} \right) \,. \tag{4.47}$$

Thus, for g_m of order 1, rotation curves usually do give a constraint.

For general orbits, we can assume neither $\vec{V} \perp \hat{a}$ nor $\vec{V} \parallel \hat{a}$. We can still obtain simple constraints if we take the most conservative values of V_{crit} and τ_E. That is, we can take $V_{\text{crit}} = V_{\text{crit}}^{\parallel}$ and $f_a = f_a^{\perp}$. Then, we can rule out an interval

$$(f_X V, V/f_V) \,, \tag{4.48}$$

with $f_V = \sqrt{2}$.

4.3. Application to standard SFDM

We will now use the results of Sec. 4.2 to constrain standard SFDM. We assume the MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$. In Sec. 3.2 we estimated that many galaxies cannot reach this MOND limit. Here, we assume the MOND limit anyway for two reasons. First, this simplifies the equations so that our results for standard SFDM serve as an illustrative example of how one can constrain hybrid models using Cherenkov radiation from non-relativistic objects. Second, our results provide an independent method to rule out that a given galaxy is in the MOND limit. That is, compared to our estimates from Sec. 3.2, our Cherenkov radiation constraints can rule out a MOND limit for a different set of galaxies and using a different physical mechanism. Of course, our constraints can be avoided by simply allowing galaxies to go outside the MOND limit. But the MOND limit is one of the main motivations behind SFDM. Knowing under which circumstances such a MOND limit can exist is important.

We will use the observed Milky Way rotation curve to rule out a MOND limit in the Milky Way for certain values of the SFDM model parameters. Specifically, assuming a MOND limit, we find constraints on the parameter combination $\sqrt{\bar{\alpha}}/m$ for each fixed value of the parameter $\bar{\beta}$, which parametrizes finite-temperature corrections. For the fiducial value $\bar{\beta} = 2$ from Ref. [72], we will rule out roughly the interval

$$0.34 \, \text{eV}^{-1} \lesssim \frac{\sqrt{\bar{\alpha}}}{m} \lesssim 3.29 \, \text{eV}^{-1} \,. \tag{4.49}$$

This includes the fiducial value $\sqrt{\bar{\alpha}}/m = 2.4 \, \text{eV}^{-1}$ from Ref. [72].

4.3.1. Perturbations

We consider perturbations on top of a background galaxy in the MOND limit $(\vec{\nabla}\theta)^2 \ll 2m\hat{\mu}$. We further consider the MOND regime $a_b < a_0$, so that potential higher-order terms are negligible, as discussed in Sec. 4.2.4. Then, the perturbed Lagrangian is (see Appendix B)

$$\mathcal{L} = \frac{1}{2}\dot{\delta}^2\bar{c}^{-2} - \frac{1}{2}(\vec{\nabla}\delta)^2 - \frac{1}{2}(\hat{a}\vec{\nabla}\delta)^2 + \bar{c}^{-1}(3-\bar{\beta})\bar{f}_{\bar{\beta}}\hat{a}\vec{\nabla}\delta\dot{\delta} - \frac{g_m}{\sqrt{2}M_{\rm Pl}}\delta\,\delta_b\,, \quad (4.50)$$

with

$$\bar{c} = 3\bar{f}_{\bar{\beta}}\,\frac{|\vec{a}_{\theta_0}|}{a_0}\frac{\bar{\alpha}^2\Lambda}{m}\,, \quad g_m = \sqrt{\frac{a_0}{|a_{\theta_0}|}}\,, \quad \hat{a} = \frac{\vec{\nabla}\theta_0}{|\vec{\nabla}\theta_0|} = -\frac{\vec{a}_{\theta_0}}{|a_{\theta_0}|}\,, \quad (4.51)$$

and

$$\bar{f}_{\bar{\beta}} = \frac{1}{\sqrt{3\left(\bar{\beta}-1\right)\left(\bar{\beta}+3\right)}}\,. \quad (4.52)$$

This has the form of our prototype Lagrangian from Eq. (4.4) up to the term that mixes spatial and time derivatives.

The quantity $\bar{c}$ determines the sound speed up to order 1 corrections, which we discuss below. This gives a non-relativistic sound speed, as expected for a non-relativistic superfluid. For the fiducial parameters from Ref. [72] ($m = 1\,\mathrm{eV}$, $\bar{\alpha} = 5.7$, $\Lambda = 0.05\,\mathrm{meV}$, $\bar{\beta} = 2$), we have

$$\bar{c} = 1.25\cdot 10^{-3}\frac{|\vec{a}_{\theta_0}|}{a_0} = 375\,\mathrm{km/s}\cdot\frac{|\vec{a}_{\theta_0}|}{a_0}\,. \quad (4.53)$$

Since we consider the MOND regime $a_b < a_0$, we have $|\vec{a}_{\theta_0}|/a_0 \lesssim 1$. Thus, indeed $\bar{c} \ll 1$. The full dispersion relation is, see Appendix B,

$$\omega = \bar{c}|\vec{k}|\left(\sqrt{1+\gamma^2(1+f_{\bar{\beta}}^2)} + f_{\bar{\beta}}\gamma\right)\,, \quad (4.54a)$$

with $f_{\bar{\beta}} \equiv (3-\bar{\beta})\bar{f}_{\bar{\beta}}$. This is the standard result from our prototype Lagrangian up to corrections from $f_{\bar{\beta}}$, i.e. up to corrections from the mixing of spatial and time derivatives.

Unfortunately, due to this mixing of spatial and time derivatives, our standard calculation of the energy loss $\dot{E}$ from Appendix A based on the standard QFT formalism does not apply here. Instead of adjusting the QFT formalism for our case at hand, we instead choose to do a classical calculation for standard SFDM. This calculation is done in Appendix C and follows that of the standard electromagnetic Cherenkov radiation from Ref. [135]. The result from the classical calculation has the same form as that from our previous QFT calculation, but with slightly adjusted critical velocities $V_{\rm crit}$ and $\hat{a}$-dependent factors f_a to account for the mixing of spatial and time derivatives. We also use this classical calculation to explicitly determine the factor f_p. That is,

we check when perturbations are small compared to the galaxy's background field. We find that we can use $f_p = 1$ for our purposes, see Appendix C.6 and Sec. 4.2.4.

For $\vec{V} \parallel \hat{a}$, the critical velocity $V_{\rm crit}$ is

$$V_{\rm crit}^{\parallel} = \bar{c}\left(\sqrt{2 + f_{\bar{\beta}}^2} \pm f_{\bar{\beta}}\right) . \tag{4.55}$$

The plus and minus signs are for parallel and antiparallel orientations, respectively. For the parallel case, $V_{\rm crit}^{\parallel}$ can be significantly larger compared to our prototype Lagrangian where $V_{\rm crit}^{\parallel} = \sqrt{2}\bar{c}$. For the antiparallel case, $V_{\rm crit}^{\parallel}$ can be significantly smaller. The timescale τ_E has the same form as for our prototype Lagrangian but with a modified factor f_a,

$$f_a^{\parallel} = \frac{1}{1 - (\bar{c}/V)^2 \mp 2 f_{\bar{\beta}} (\bar{c}/V)} , \tag{4.56}$$

where the minus sign is for the parallel and the plus sign for the antiparallel case. Depending on this sign, $f_a^{\parallel}$ is smaller or larger than for our prototype Lagrangian where $f_a^{\parallel} = 1/(1 - (\bar{c}/V)^2)$. Both $V_{\rm crit}^{\parallel}$ and $f_a^{\parallel}$ agree with our standard calculation in the $f_{\bar{\beta}} = 0$ case.

For $\vec{V} \perp \hat{a}$, the critical velocity $V_{\rm crit}$ is

$$V_{\rm crit}^{\perp} = \bar{c}\sqrt{\frac{2}{2 + f_{\bar{\beta}}^2}} . \tag{4.57}$$

In this case, some angular integrals are difficult to calculate analytically, see Appendix C.4. This makes it difficult to analytically obtain the timescale τ_E. But we can do a somewhat more conservative estimate and find as a lower limit on $f_a^{\perp}$,

$$f_a^{\perp} = \frac{1}{\sqrt{2}} \frac{1}{1 + f_{\bar{\beta}}^2} . \tag{4.58}$$

The critical velocity $V_{\rm crit}^{\perp}$ reproduces our standard result for the $f_{\bar{\beta}} = 0$ case. But $f_a^{\perp}$ does not due to the more conservative estimate, see Appendix C.4.

4.3.2. Constraints

We can now put quantitative constraints on standard SFDM using the general results from Sec. 4.2.5. But first we give a qualitative discussion.

In the MOND limit, a good approximation is $|\vec{a}_{\theta_0}| = \sqrt{a_0 a_b^{\rm gal}}$ with $a_b^{\rm gal} = GM_{\rm gal}/R_p^2$. This gives

$$\bar{c} = \frac{3\bar{f}_{\bar{\beta}}}{\sqrt{8\pi}} \frac{\sqrt{\bar{\alpha}}}{m} \sqrt{\frac{M_{\rm gal}}{M_{\rm Pl}}} \frac{1}{R_p} . \tag{4.59}$$

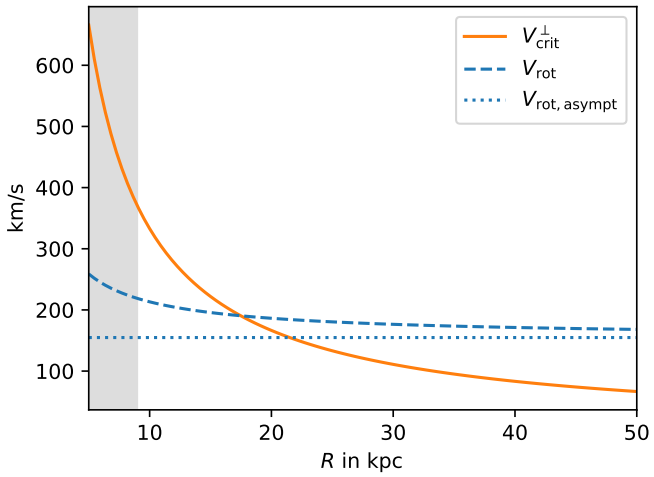

Figure 4.4.: Critical velocity $V_{\mathrm{crit}}^{\perp}$ (solid orange line) and rotation curve V_{rot} (dashed blue line) of a galaxy with mass $M_{\mathrm{gal}} = 5 \cdot 10^{10}\, M_{\odot}$ concentrated at its center in standard SFDM with the fiducial parameter values from Ref. [72]. The shaded region is not in the MOND regime since $a_b^{\mathrm{gal}} > a_0$. The dotted blue line shows the asymptotic rotation velocity $V_{\mathrm{rot,asympt}}$. Stars emit Cherenkov radiation when they move faster than the critical velocity.

For circular orbits, the critical velocity is $V_{\mathrm{crit}}^{\perp} = \bar{c}\sqrt{2/(2+f_{\bar{\beta}}^2)}$. This critical velocity scales as $1/R_p$ inside a given galaxy. In contrast, rotation curves are flat at large R_p. Thus, there is a galactocentric radius where V_{crit} drops below the rotation curve velocity V_{rot}. Beyond this radius, stars with velocity V_{rot} lose energy on timescales τ_E. This is illustrated in Fig. 4.4.

For standard SFDM, the timescale τ_E is

$$\tau_E = \frac{V^3}{f_a f_p^2 \bar{c}^2 a_0} \cdot \sqrt{\frac{a_0}{a_b^{\mathrm{gal}}}}\,. \tag{4.60}$$

This depends on parameters of the model mainly through the factor $\bar{c}^{-2}$. In principle, $a_0 = \bar{\alpha}^3\Lambda^2/M_{\mathrm{Pl}}$ also depends these parameters. But a_0 is mostly fixed by requiring standard MOND-like rotation curves. Thus, τ_E depends on the model parameters as $\tau_E \propto 1/\bar{c}^2$ so that our Cherenkov constraints can be avoided in two ways: Either, $\bar{c}$ is sufficiently large so that most stars are subsonic and therefore don't emit Cherenkov radiation since Cherenkov radiation is kinematically forbidden. Or $\bar{c}$ is sufficiently small so that stars that are supersonic lose only little energy through Cherenkov radiation.

We will now make this more quantitative using the results from Sec. 4.2.5. For simplicity, we consider three different values of $\bar{\beta}$

$$\bar{\beta} \in \{3/2, 2, 3\}\,. \tag{4.61}$$

This covers the range of plausible values for $\bar{\beta}$. Namely, $\bar{\beta}$ must be larger than 3/2 to fix an instability[2] and it must be smaller than 3 so that the superfluid energy density is positive [71]. The fiducial value from Ref. [72] is $\bar{\beta} = 2$. For a fixed value of $\bar{\beta}$, we have $\bar{c} \propto \sqrt{\bar{\alpha}}/m$. This allows us to constrain $\sqrt{\bar{\alpha}}/m$ given a value of $\bar{\beta}$.

Consider first rotation curves with approximately circular orbits. As discussed in Sec. 4.2.4 and Appendix C.6, this means we can set $f_p = 1$. In terms of the notation from Sec. 4.2.5, we have

$$B = 0\,, \quad A = \sqrt{2}(1 + f_{\bar{\beta}}^2)\,, \quad f_V(\bar{\beta}) = \sqrt{\frac{2}{2 + f_{\bar{\beta}}^2}}\,. \tag{4.62}$$

Then, we can rule out $\bar{c}$ in the interval $(f_X V, V/f_V)$ given that V is smaller than a certain maximum velocity, as discussed in Sec. 4.2.5. For standard SFDM, this maximum velocity requires

$$V < \frac{\tau_{\min}\sqrt{a_0 G M_{\rm gal}}}{\sqrt{2} R_p} \frac{1 + \frac{1}{2} f_{\bar{\beta}}^2}{1 + f_{\bar{\beta}}^2}\,. \tag{4.63}$$

As discussed in Sec. 4.2.5, this is usually satisfied for rotation curves. Then, we can rule out $\sqrt{\bar{\alpha}}/m$ in the interval

$$\frac{V R_p}{\bar{f}_{\bar{\beta}}} \frac{\sqrt{8\pi}}{3} \sqrt{\frac{M_{\rm Pl}}{M_{\rm gal}}} \cdot \left(f_X, 1/f_V(\bar{\beta})\right)\,, \tag{4.64}$$

where we used the notation $a \cdot (x_1, x_2) \equiv (a x_1, a x_2)$.

Here, the quantity X is

$$X = \frac{\tau_{\min}\sqrt{a_0 G M_{\rm gal}}}{R_p V \sqrt{2}(1 + f_{\bar{\beta}}^2)}\,. \tag{4.65}$$

In standard SFDM, a_0 is given by $\bar{\alpha}^3 \Lambda^2 / M_{\rm Pl}$. Thus, strictly speaking, the lower boundary of the interval in Eq. (4.64) depends on a different combination of model parameters than the upper boundary, which depends only on $\sqrt{\bar{\alpha}}/m$. However, in practice this is not important. Namely, f_X depends on a_0 only very mildly, as $a_0^{-1/4}$. For a useful MOND regime, we must choose a_0 close to $10^{-10}\,\mathrm{m/s^2}$ [11, 72]. Thus, to get a conservative estimate, we simply take the unusually small but fixed value

$$\bar{a}_0 \equiv 0.5 \cdot 10^{-10}\,\mathrm{m/s^2}\,, \tag{4.66}$$

in f_X instead of $a_0 = \bar{\alpha}^3 \Lambda^2 / M_{\rm Pl}$. That is, we take $\sqrt{\bar{\alpha}}/m$ in the following interval to be ruled out

$$\frac{V R_p}{\bar{f}_{\bar{\beta}}} \frac{\sqrt{8\pi}}{3} \sqrt{\frac{M_{\rm Pl}}{M_{\rm gal}}} \cdot \left(f_X(\bar{a}_0), 1/f_V(\bar{\beta})\right)\,. \tag{4.67}$$

[2] In the MOND limit, $\bar{\beta} > 1$ is sufficient, but more generally $\bar{\beta} > 3/2$ is needed.

Here, $f_X(\bar{a}_0)$ means f_X evaluated at $a_0 = \bar{a}_0$.

We can now rule out various intervals of $\sqrt{\bar{\alpha}}/m$ using the observed Milky Way rotation curve. We assume a Milky Way baryonic mass of $M_{\rm gal} = 6 \cdot 10^{10}\, M_\odot$ following Refs. [73, 85], and the rotation curve data from Refs. [104, 107]. We choose $\tau_{\rm min} = 10^{10}\,{\rm yr}$, i.e. stars should not lose a significant fraction of their energy in $10^{10}\,{\rm yr}$. For $\bar{\beta} = 3/2$, 2, and 3, we list the excluded intervals of $\sqrt{\bar{\alpha}}/m$ from three different galactic radii in Table 4.1. These radii are, roughly, $R_p = 15\,{\rm kpc}$, $20\,{\rm kpc}$, and $25\,{\rm kpc}$. The ruled out intervals from these three different radii overlap. Thus, taken together they rule out $\sqrt{\bar{\alpha}}/m$ in a larger interval

$$\sqrt{\frac{6 \cdot 10^{10}\, M_\odot}{M_{\rm gal}}} \left(\sqrt{\frac{10^{10}\,{\rm yr}}{\tau_{\rm min}}} \left(\frac{6 \cdot 10^{10}\, M_\odot}{M_{\rm gal}} \right)^{1/4} q_l, q_h \right) \cdot {\rm eV}^{-1}\,, \tag{4.68}$$

for some dimensionless numbers q_l and q_h. We have

$$\begin{aligned} q_l &= 0.25\,, q_h = 2.34\,, \quad \text{for } \bar{\beta} = 3/2\,, \\ q_l &= 0.34\,, q_h = 3.29\,, \quad \text{for } \bar{\beta} = 2\,, \\ q_l &= 0.51\,, q_h = 5.01\,, \quad \text{for } \bar{\beta} = 3\,. \end{aligned} \tag{4.69}$$

This also rules out the fiducial value $\sqrt{\bar{\alpha}}/m \approx 2.4\,{\rm eV}^{-1}$ for $\bar{\beta} = 2$ from Ref. [72].

The precise ruled out values of $\sqrt{\bar{\alpha}}/m$ depend on our choice for $M_{\rm gal}$ and $\tau_{\rm min}$, as shown in Eq. (4.68). For example, Ref. [104] adopts the somewhat higher baryonic Milky Way mass $M_{\rm gal} = 7.4 \cdot 10^{10}\, M_\odot$. This would decrease both the upper and lower boundaries of the excluded $\sqrt{\bar{\alpha}}/m$ interval by about 10%. The value of $\tau_{\rm min}$ enters our result as the inverse square root, but affects only the lower interval boundary. Uncertainties in the measured rotation curve affect both the upper and lower boundaries of the excluded $\sqrt{\bar{\alpha}}/m$ interval, see Eq. (4.68). However, the formal errors in the rotation curve from Refs. [104, 107] are less than 5% for $R_p \lesssim 25\,{\rm kpc}$. Thus, the uncertainties in $M_{\rm gal}$ and $\tau_{\rm min}$ are the dominant sources of uncertainty in our constraints.

It may be possible to push the upper boundary of our excluded interval of $\sqrt{\bar{\alpha}}/m$ even higher by considering hypervelocity stars (HVS) or globular clusters (GCs). The reason is that HVS and GCs can be at much larger galactic radii and have much higher velocities than the Milky Way stellar rotation curve. Thus, the quantity VR_p can be much larger, which allows to reach much higher values of $\sqrt{\bar{\alpha}}/m$. However, in practice, this is more complicated, because HVS and GCs cannot be assumed to be on circular orbits. Thus, we are forced to make various worst-case assumptions if we know only the distance to the host galaxy's center and total velocity and nothing more about the orbit. That is, in general, we are forced to assume the worst-case critical velocity $V_{\rm crit}$ and direction-dependent factors f_a and f_p. We must also choose a very conservative timescale $\tau_{\rm min}$ since we cannot, in general, assume the HVS or GC to have been at its current position for most of the host galaxy's lifetime.

R_p kpc	V km/s	(q_l, q_h) for $\bar{\beta} = 3/2$	(q_l, q_h) for $\bar{\beta} = 2$	(q_l, q_h) for $\bar{\beta} = 3$
15.2	220^{+1}_{-1}	$(0.25, 1.56)$	$(0.34, 2.19)$	$(0.51, 3.34)$
20.3	203^{+3}_{-3}	$(0.35, 1.92)$	$(0.46, 2.70)$	$(0.69, 4.11)$
24.8	202^{+6}_{-6}	$(0.47, 2.34)$	$(0.62, 3.29)$	$(0.93, 5.01)$

Table 4.1.: Assuming a MOND limit in the Milky Way, we rule out $\sqrt{\bar{\alpha}}/m$ intervals $(q_l, q_h)\cdot\text{eV}^{-1}$ from the observed Milky Way rotation curve at different radii. We assume $M_{\text{gal}} = 6\cdot10^{10}\,M_\odot$ and $\tau_{\text{min}} = 10^{10}\,\text{yr}$. See Eq. (4.68) for how the exclusion interval depends on these quantities. The rotation curve data is that of Ref. [107] adopted to the assumptions of Ref. [104], as also used in Sec. 2.2. The different radii taken together exclude a larger total interval, see Eq. (4.69).

As a result, we cannot get good constraints from HVS or GCs with the simple procedure we use for the Milky Way rotation curve. Doing better is certainly possible but requires a more detailed modelling of their orbits. We leave such a more involved analysis for future work.

The parameter combination $\sqrt{\bar{\alpha}}/m$ that we have constrained is phenomenologically important since the superfluid energy density is directly proportional to powers of it. For example, in the MOND limit and using the no-curl approximation,

$$\rho_{\text{SF}} = 2\left(1 - \frac{\bar{\beta}}{3}\right)\left(\frac{m}{\sqrt{\bar{\alpha}}}\right)^2 M_{\text{Pl}}\sqrt{a_0\, a_b^{\text{gal}}}\,, \tag{4.70}$$

where we used $\Lambda m^3 = (m/\sqrt{\bar{\alpha}})^3\sqrt{a_0 M_{\text{Pl}}}$. Thus, our constraints likely have further implications for situations where the superfluid energy density is important, for example for strong lensing. Investigating this is left for future work.

Our results rule out that the Milky Way is in the MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$ for a certain range of parameters. As mentioned above, this does not completely rule out this range of parameters. Being in the MOND limit is one of the main motivations behind SFDM. Still, we might just accept that the Milky Way is outside the MOND limit[3] and hope that this makes it one of only few outliers, while most galaxies actually are in the MOND limit. This should be checked in future work by doing a similar analysis for a comprehensive

[3] Indeed, the Milky Way model discussed in Sec. 2.2 goes outside the MOND limit $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$, see Sec. 3.2. But we have verified that one can adjust the $\hat{\mu}$ boundary condition in order to obtain other solutions with $\varepsilon_* \lesssim 0.15$ at the radii $15\,\text{kpc} \lesssim R \lesssim 25\,\text{kpc}$ that are relevant here. The condition $\varepsilon_* \ll 1$ is equivalent to the condition $(\vec{\nabla}\theta)^2 \gg 2m\hat{\mu}$, see Sec. 3.2. So equilibrium solutions in the MOND limit may, barely, be possible. The price to pay for this is that the superfluid core ends shortly after the last measured rotation curve data point. The associated dark matter mass is probably too small to be plausible, see Sec. 2.2. Our Cherenkov radiation constraint rules out such solutions without relying on knowing which dark matter masses are or are not plausible in SFDM.

sample of galaxies with measured rotation curves. But see the discussion in Sec. 4.6 for why this may be more complicated than one might naively expect.

4.4. Application to two-field SFDM

Standard SFDM runs into problems due to the double role played by the phonon field θ, as discussed in Chapter 3. To avoid these problems, we proposed two-field SFDM which splits the two roles of the phonon field θ between two separate fields, see Sec. 3.4. We will now check whether or not this model is affected by our Cherenkov radiation constraint for hybrid models.

Two-field SFDM contains two gapless low-energy modes, roughly corresponding to the two fields θ_+ and θ_-. Of these, only the mode corresponding to θ_- has a non-relativistic sound speed

$$c_s = \sqrt{\frac{\hat{\mu}}{m}}\,, \tag{4.71}$$

so that only this mode can potentially be radiated away by stars as Cherenkov radiation. The coupling of this mode to normal matter is suppressed because θ_- does not directly couple to normal matter. This is because only θ_+ carries the MOND-like force in this model, not θ_-. Cherenkov radiation is possible only through a mixing of θ_- and θ_+ from the $f(K_+ + K_- - m^2)$ term of the Lagrangian, see Eq. (3.22).

We can write down a low-energy effective Lagrangian for this non-relativistic mode. This gives our prototype Lagrangian Eq. (4.4) without the $\hat{a}$ term and with (see Appendix D)

$$\bar{c} = \sqrt{\frac{\hat{\mu}}{m}}\,, \quad g_m = -\frac{\sqrt{\lambda_4}}{\bar{\alpha}}\frac{\gamma}{1+\gamma^2}\frac{a_0}{|\vec{a}_{\theta_+^0}|}\,. \tag{4.72}$$

This is different from our prototype model in two ways. First, since there is no $\hat{a}$ term, we have $c_s = \bar{c}$ without the $\sqrt{1+\gamma^2}$ factor. As a result, the critical velocity V_{crit} is $V_{\text{crit}} = \bar{c}$ independently of the orientation of $\vec{V}$. Similarly, we have $f_a = 1$ independently of the orientation of $\vec{V}$. Otherwise, the energy-loss time scale τ_E is not affected by the missing $\hat{a}$ term. Second, the coupling g_m depends on γ. This requires a more careful evaluation of the integrals in the $\dot{E}$ calculation, as we will discuss below.

Numerically, the combination $\sqrt{\lambda_4}/\bar{\alpha}$ that controls the coupling g_m is small. Specifically,

$$\frac{\sqrt{\lambda_4}}{\bar{\alpha}} = \frac{1}{10^{7/2}}\left(\frac{\bar{a}}{a_0}\right)^{1/4}, \tag{4.73}$$

where $\bar{a} \ll a_0$ is the minimum baryonic acceleration below which the equilibrium on galactic scales typically becomes unstable, see Sec. 3.4. Thus, this non-relativistic mode does not have a standard gravitational coupling constant

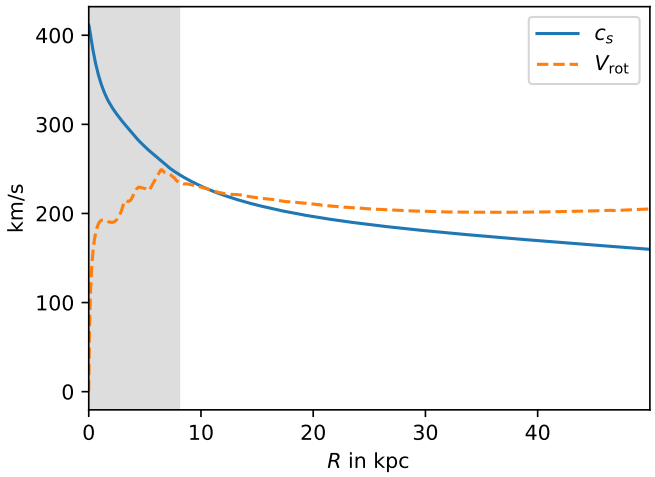

Figure 4.5.: Sound speed c_s and rotation curve v_{rot} for the two-field SFDM Milky Way model from Sec. 3.4. The shaded region is not in the MOND regime since $a_b^{\mathrm{gal}} > a_0$.

of order $\sqrt{G}$ to normal matter. This is because this mode couples to normal matter only indirectly through a mixing between θ_+ and θ_-. This enhances the timescale τ_E.

In addition to the cutoffs discussed in Sec. 4.2, this model requires another model-dependent cutoff, see Appendix D.1 or Ref. [131]. Namely, the dispersion relation $\omega = c_s k$ of the non-relativistic mode receives corrections from higher orders of k for $k \gtrsim mc_s$. Thus, our standard calculation applies only if we introduce a cutoff of order mc_s. However, in practice this is not relevant since the other cutoffs introduced in Sec. 4.2 are much more restrictive. Roughly, the mc_s cutoff becomes relevant only if $mc_s \lesssim 10^{-22}\,\mathrm{eV}$. Typical values of c_s are of order $100\,\mathrm{km/s}$. Thus, the mc_s cutoff is relevant only if $m \lesssim 10^{-18}\,\mathrm{eV}$ which is much smaller than typical masses in two-field SFDM [73] (see also Sec. 3.4).

In many galaxies, at least some stars will be supersonic and lose energy through Cherenkov radiation, since typically $c_s = \sqrt{\hat{\mu}/m} \sim \mathcal{O}(100\,\mathrm{km/s})$, as mentioned above. The precise value of c_s depends on the boundary condition μ_∞ of $\hat{\mu}$ which controls how much dark matter a galaxy contains [73]. Typically, c_s falls off as a function of galactocentric radius since $\hat{\mu}/m$ does. Indeed, in spherical symmetry $\hat{\mu}'(r)/m = -G(M_b + M_{\mathrm{SF}})/r^2 < 0$. In contrast, rotation curves become constant at large radii. Thus, as in standard SFDM, there is typically a critical radius R_c beyond which the rotational velocity is supersonic and stars emit Cherenkov radiation. This is illustrated in Fig. 4.5, which shows the sound speed and rotation curve of the Milky Way model discussed in Sec. 3.4.

We can analytically estimate the critical radius R_c as a function of how much

the superfluid's gravitational pull contributes to the rotation curve velocity. This is because both the superfluid energy density $\rho_{\rm DM}$ and the sound speed c_s^2 are proportional to $\hat{\mu}/m$. For given model parameters, a smaller superfluid mass implies a smaller sound speed which means R_c becomes smaller. Another way to say this is that more MOND-like rotation curves lead to more stars emitting Cherenkov radiation. Quantitatively, we have

$$c_s^2 = \frac{\lambda_4}{2m^4}\rho_{\rm DM} = \frac{1}{2}\frac{r_0^2}{M_{\rm Pl}^2}\rho_{\rm DM}\,. \tag{4.74}$$

For simplicity, we assume a spherically symmetric superfluid. Since $\rho_{\rm DM}$ decreases with distance to the host galaxy's center, we have

$$\rho_{\rm DM} \equiv f_\rho \bar{\rho}_{\rm DM} < \bar{\rho}_{\rm DM}\,, \tag{4.75}$$

where $\bar{\rho}_{\rm DM}$ is the dark matter density averaged over a sphere of radius R. Then,

$$c_s^2 = \frac{3 f_\rho r_0^2}{R}\cdot a_{\rm DM}\,, \tag{4.76}$$

where $a_{\rm DM} = GM_{\rm DM}(R)/R^2$ is the Newtonian gravitational acceleration due to the superfluid. The condition for a supersonic rotation curve, $V^2 > c_s^2$, then becomes

$$\frac{3 f_\rho\, r_0^2}{R^2} - 1 < \frac{a_b + a_+}{a_{\rm DM}}\,, \tag{4.77}$$

where we used $V^2/R = a_b + a_+ + a_{\rm DM}$ with the MOND-like acceleration a_+ due to the field θ_+ and the Newtonian baryonic acceleration a_b. Rotation curves in two-field SFDM are MOND-like if the contribution from $a_{\rm DM}$ is small. If we allow $a_{\rm DM}$ to change V by at most a fraction ε_V, the rotation curve velocity is supersonic beyond a critical radius R_c given by

$$\frac{R_c}{r_0} = \sqrt{6 f_\rho}\,\frac{\sqrt{\varepsilon_V + \frac{1}{2}\varepsilon_V^2}}{1+\varepsilon_V}\,. \tag{4.78}$$

For small ε_V, this is

$$R_c = 39\,{\rm kpc}\cdot f_\rho \cdot \left(\frac{r_0}{50\,{\rm kpc}}\right)\cdot\sqrt{\frac{\varepsilon_V}{10\%}}\,. \tag{4.79}$$

For the Milky Way model discussed in Sec. 3.4, the contribution from $a_{\rm DM}$ changes the rotation curve by about $\varepsilon_V = 6\%$ at $R = 30\,{\rm kpc}$ where the dark matter density is smaller than its average by about a factor $f_\rho = 0.86$. Thus, $R_c \sim 39\,{\rm kpc}$ is likely a conservative estimate. Indeed, in this Milky Way model the rotation curve is supersonic already at $R \approx 11\,{\rm kpc}$, see Fig. 4.5.

This confirms that stars in many galaxies will likely be supersonic and lose energy on timescales τ_E. In principle, this might be avoided by choosing model

parameters such as r_0 differently. But since r_0 also controls the superfluid energy density, one must be careful not to produce too much or too little dark matter mass to account for strong lensing, see Sec. 2.1.

In any case, two-field SFDM can avoid our Cherenkov radiation constraints if τ_E is larger than galactic timescales. As mentioned above, we need to revisit our calculation to include the factor of $\gamma/(1+\gamma^2)$ in g_m. If, for simplicity, we first set $\gamma = 1$ in g_m, we can use our standard calculation and find

$$\tau_E^{\gamma=1} = \frac{V^3}{f_p^2 g_m^2 \bar{c}^2 a_0} \frac{a_0}{a_b^{\mathrm{gal}}} = 4 \frac{V^3}{f_p^2 \bar{c}^2 a_0} \frac{a_0}{a_b^{\mathrm{gal}}} \frac{\bar{\alpha}^2}{\lambda_4} \left(\frac{|\vec{a}_{\theta_+^0}|}{a_0} \right)^2 = 4 \cdot 10^7 \frac{V^3}{f_p^2 \bar{c}^2 a_0} \sqrt{\frac{a_0}{\bar{a}}} \,. \tag{4.80}$$

Here, we assumed the MOND limit value $a_{\theta_+^0} = \sqrt{a_0 a_b^{\mathrm{gal}}}$ so that the factor a_0/a_b^{gal} cancels. Assuming the MOND limit is usually justified in two-field SFDM, see Sec. 3.4. In the special cases $\vec{V} \perp \hat{a}$ and $\vec{V} \parallel \hat{a}$, the corrections due to the factor $\gamma/(1+\gamma^2)$ in g_m are calculated in Appendix A.4,

$$\tau_E^{\parallel} = \frac{\tau_E^{\gamma=1}}{4} \cdot \left(\frac{c_s}{V} + \frac{V}{c_s} \right)^2 , \tag{4.81}$$

$$\tau_E^{\perp} = \frac{\tau_E^{\gamma=1}}{4} \cdot \frac{2^{5/2} \left(1 - \frac{1}{2} \left(\frac{c_s}{V} \right)^2 \right)^{3/2}}{1 - \left(\frac{c_s}{V} \right)^2} \,. \tag{4.82}$$

Thus, the timescale τ_E is significantly enhanced for fast stars with $\vec{V} \parallel \hat{a}$ and for barely supersonic stars with $\vec{V} \perp \hat{a}$. Otherwise, the order of magnitude is that of $\tau_E^{\gamma=1}$,

$$\tau_E^{\gamma=1} = \frac{8 \cdot 10^{16}\,\mathrm{yr}}{f_p^2} \cdot \left(\frac{V/\bar{c}}{2} \right)^2 \cdot \left(\frac{V}{200\,\mathrm{km/s}} \right) \cdot \left(\frac{1.2 \cdot 10^{-10}\,\mathrm{m/s^2}}{a_0} \right) \cdot \sqrt{\frac{10^{-2} a_0}{\bar{a}}} \,. \tag{4.83}$$

As discussed in Sec. 3.4, $\bar{a}$ must be much smaller than a_0 for a useful MOND limit. Thus, for reasonable values of $\bar{a}$, the timescale τ_E is much larger than galactic timescales. That is, our method does not constrain two-field SFDM. The reason is that the non-relativistic gapless mode couples to normal matter only indirectly through a mixing.

4.5. Application to the SZ model

Recently, Skordis and Złośnik proposed a novel type of hybrid MOND dark matter model [25]. This model has a cosmological limit that reproduces the CMB, a static limit that reproduces MOND, and has gravitational tensor modes that propagate at the speed of light. In the cosmological limit, a scalar field $\phi(t)$ is responsible for a dark fluid that plays the role of dark matter. In

the late-time static limit relevant for galaxies, this model has $\phi = Q_0 \cdot t + \varphi$ where φ carries the MOND-like force and Q_0 is a constant. That is, the field ϕ plays a double role, analogously to the phonon field θ in standard SFDM. Cosmological and galactic phenomena share a common origin. Therefore, our method is, in principle, able to constrain this model.

This SZ model is complex and we will not attempt a full calculation of perturbations on top of a galactic background. Instead, we follow Ref. [25] and consider perturbations on top of the late-time Minkowski background with $\phi = Q_0 t$. We assume that this gives results that are at least qualitatively valid in galaxies. Concretely, we consider $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$ with the Minkowski metric $\eta_{\alpha\beta}$, the vector field $A_\alpha = (-1 + \frac{1}{2}h_{00}, \vec{A})$, and the scalar field $\phi = Q_0 \cdot t + \varphi$. Then, the second-order Lagrangian reads [25]

$$\begin{aligned}\mathcal{L} =& K_\mathrm{B}|\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}|^2 - 2K_\mathrm{B}\vec{\nabla}_{[i}A_{j]}\vec{\nabla}^{[i}A^{j]} \\ &+ 2\,(2-K_\mathrm{B})\left(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}\right)\cdot\vec{\nabla}\varphi + 2(2-K_\mathrm{B})Q_0\vec{A}_i\left(-\frac{1}{2}\partial_i h_{00}\right) \\ &- (2-K_\mathrm{B})(1+\lambda_s)\left(\vec{A}^2Q_0^2 + (\vec{\nabla}\varphi)^2 + 2Q_0A^i(\partial_i\varphi)\right) \\ &+ 2\mathcal{K}_2\left(\dot{\varphi} + \frac{1}{2}h_{00}Q_0\right)^2 + \frac{1}{\tilde{M}_\mathrm{Pl}^2}T_{\alpha\beta}h^{\alpha\beta}\,, \end{aligned} \tag{4.84}$$

where we left out the standard metric perturbations from the Ricci scalar R for simplicity and we set $h_{0i} = 0$. Here, $\lambda_s = 2\mathcal{J}'(0)/(2-K_\mathrm{B})$ where the function $\mathcal{J}$ determines the MOND interpolation function [25]. The quantities K_B, Q_0, $\mathcal{K}_2$, and $\tilde{M}_\mathrm{Pl}$ are constants.

There is a peculiar term proportional to $\mathcal{K}_2Q_0^2h_{00}$ in this Lagrangian. This acts as a kind of mass term for h_{00} in the static limit, as discussed in Ref. [25]. A useful static limit with a long-range gravitational force on galactic scales then requires

$$\sqrt{\mathcal{K}_2}Q_0 \lesssim 1/\mathrm{Mpc}\,. \tag{4.85}$$

For simplicity, we further assume

$$2 - K_\mathrm{B} = \mathcal{O}(1)\,, \quad 1 + \lambda_s = \mathcal{O}(1)\,. \tag{4.86}$$

The first condition is fulfilled as long as K_B is not too close to 2. This is the case for the concrete numerical examples considered in Ref. [25]. The second condition is natural in our case, since for an actual galactic background $\mathcal{J}'(0)$ would be replaced by $\mathcal{J}'((\vec{\nabla}\varphi)^2) \sim |\vec{\nabla}\varphi|/a_0 \lesssim 1$.

There is one scalar mode relevant at low energies [25] (see also Appendix E). This mode has sound speed

$$c_s = \frac{c'}{\sqrt{\mathcal{K}_2K_\mathrm{B}}}\,, \qquad c'^2 = (2-K_\mathrm{B})\left(1 + \frac{1}{2}\lambda_sK_\mathrm{B}\right)\,. \tag{4.87}$$

Since we did not include the galaxy's background field, there is no term $(\hat{a}\vec{\nabla}\varphi)^2$ in the Lagrangian and the sound speed is $c_s = \bar{c}$ without a factor $\sqrt{1+\gamma^2}$. In a full calculation including the galaxy's background field, we expect such a term to be present. Here, we assume that such corrections do not change the order of magnitude of our result.

Since $c' = \mathcal{O}(1)$, stars with velocity V are supersonic if, roughly,

$$\sqrt{\mathcal{K}_2 K_\mathrm{B}} \gtrsim 10^3 \cdot \left(\frac{200\,\mathrm{km/s}}{V}\right) . \tag{4.88}$$

In this case, stars may lose energy to Cherenkov radiation, if this mode is coupled to matter. One possibility to avoid constraints is to choose the combination $\mathcal{K}_2 K_\mathrm{B}$ large enough to make stars subsonic. But it turns out that this is not necessary. The reason is that the coupling to matter is sufficiently suppressed to make the energy-loss time scale τ_E much larger than galactic timescales, as we will now explain.

The field φ mediates a MOND-like force in galaxies. Therefore, in the static limit $\omega = 0$, this field is coupled to normal matter by a standard gravitational coupling. Since the low-energy scalar mode discussed above contains the field φ [25], one might expect this coupling to survive also for dynamical modes, $\omega > 0$. However, it turns out that this coupling is suppressed by powers of $Q_0/(\sqrt{K_\mathrm{B}}\omega)$ for dynamical modes with $\omega \gg Q_0/\sqrt{K_\mathrm{B}}$. This can be seen in an explicit calculation in Appendix E and we give a heuristic argument in Sec. 4.5.1. On-shell and ignoring order 1 factors, the condition $\omega \gg Q_0/\sqrt{K_\mathrm{B}}$ can be written as

$$\frac{Q_0}{\sqrt{K_\mathrm{B}}\omega} \sim \frac{Q_0\sqrt{\mathcal{K}_2}}{|\vec{k}|} \ll 1 . \tag{4.89}$$

For our Cherenkov radiation constraints, we consider only $|\vec{k}| > k_\mathrm{min} \sim 1/\mathrm{kpc}$. Together with the condition $\sqrt{\mathcal{K}_2}Q_0 \lesssim 1/\mathrm{Mpc}$ from Eq. (4.85) this implies that the condition Eq. (4.89) is satisfied. That is, the coupling to matter is suppressed in our case. Thus, the energy loss through Cherenkov radiation is reduced and the timescale τ_E is enhanced so that our constraints are weakened. We will now discuss this in more detail.

4.5.1. Suppressed matter coupling

The coupling of φ to matter in the static limit comes from the $(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00})\vec{\nabla}\varphi$ term in the Lagrangian. Specifically, in the static limit, the φ equation of motion is

$$0 = \cdots - \frac{1}{2}\vec{\nabla}^2 h_{00} , \tag{4.90}$$

where we only show the contributions from the $(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00})\vec{\nabla}\varphi$ term. Here, $\vec{\nabla}^2 h_{00}$ contains a term that is proportional to the matter density, just as in

standard general relativity. That is, the matter coupling of φ is due to a mixing of φ and h_{00}. Consider now a dynamical situation where time derivatives may be important. Then, in the φ equation of motion, we have

$$0 = \cdots + \vec{\nabla}\left(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}\right) . \tag{4.91}$$

The combination $\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}$ also occurs in the $\vec{A}$ equation of motion, from the kinetic term of $\vec{A}$ proportional to $(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00})^2$. Roughly,

$$0 = \cdots + \partial_t\left(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}\right) . \tag{4.92}$$

For scalar perturbations, we may choose $\vec{A}$ as the gradient of a scalar field [25]. As a result, spatial derivatives from the kinetic term of $\vec{A}$ drop out. Then, the $\partial_t^2\vec{A}$ term may dominate over other terms proportional to $\vec{A}$ in Eq. (4.92), even though the dispersion relation $\omega = c_s k$ with non-relativistic sound speed $c_s \ll 1$ forces time derivatives to be much smaller than spatial derivatives. In this case, the solution of Eq. (4.92) becomes $\partial_t\vec{A} = \frac{1}{2}\vec{\nabla}h_{00} + \ldots$. This cancels the $\vec{\nabla}^2 h_{00}$ term in the φ equation of motion Eq. (4.91). Thus, for sufficiently large ω, the standard coupling of φ to matter is absent.

We expect this to hold also for perturbations on top of a galaxy instead of empty Minkowski space. Roughly, the galaxy introduces a new scale of order $a_0 \sim 10^{-34}\,\mathrm{eV}$. This is much smaller than $\omega = c_s k$ since we typically have a lower cutoff of order $k_{\min} \sim 10^{-26}\,\mathrm{eV}$. Except possibly for an extremely low sound speed, i.e. $c_s \lesssim 10^{-8} \sim \mathrm{m/s}$. We leave a detailed study of this regime for future work.

Thus, for $\omega \to 0$, the field φ has a standard gravitational coupling to matter which is crucial for the MOND force in the static limit. However, for $\omega \to \infty$, this coupling vanishes due to the specific form of the $\vec{A}$ equation of motion. A more detailed calculation shows that the relevant scale to compare ω to is $Q_0/\sqrt{K_\mathrm{B}}$. The coupling to matter is suppressed for $\omega \gg Q_0/\sqrt{K_\mathrm{B}}$ but not for $\omega \ll Q_0/\sqrt{K_\mathrm{B}}$. This is calculated explicitly in Appendix E.

4.5.2. Energy loss through Cherenkov radiation

In the language of our prototype Lagrangian from Eq. (4.4), the matter coupling of the non-relativistic mode of the SZ model is suppressed by

$$g_m \propto \left(\frac{\sqrt{\mathcal{K}_2}Q_0}{k}\right)^3 , \tag{4.93}$$

see Appendix E. Thus, the matrix element squared for Cherenkov radiation has the order of magnitude of

$$|\mathcal{M}|^2 \approx c_s^2 \frac{M^4}{M_\mathrm{Pl}^2}\left(\frac{\sqrt{\mathcal{K}_2}Q_0}{|\vec{k}|}\right)^6 , \tag{4.94}$$

where we assumed $\tilde{M}_{\rm Pl}$ to be of the same order of magnitude as $M_{\rm Pl}$. Due to the k-dependent factor $(\sqrt{\mathcal{K}_2}Q_0/k)^6$ in $|\mathcal{M}|^2$, the k integral in $\dot{E}$ does not give $k_{\rm max}^2$ as usual but instead

$$\frac{(\sqrt{\mathcal{K}_2}Q_0)^6}{k_{\rm min}^4} = k_{\rm max}^2 \left(\frac{k_{\rm min}}{k_{\rm max}}\right)^2 \left(\frac{\sqrt{\mathcal{K}_2}Q_0}{k_{\rm min}}\right)^6 , \tag{4.95}$$

where we again neglected $\mathcal{O}(1)$ constants. Thus, as a rough order of magnitude estimate,

$$\tau_E \sim 10^8\,{\rm yr} \cdot \left(\frac{V}{c_s}\right)^2 \cdot \left(\frac{k_{\rm max}}{k_{\rm min}}\right)^2 \cdot \left(\frac{k_{\rm min}}{\sqrt{\mathcal{K}_2}Q_0}\right)^6 \cdot \left(\frac{V}{200\,{\rm km/s}}\right) . \tag{4.96}$$

We typically have $k_{\rm max} \sim 10^4 k_{\rm min}$ and $k_{\rm min} \sim 1/{\rm kpc}$. Further, for a useful static limit we need $\sqrt{\mathcal{K}_2}Q_0 \ll 1/{\rm kpc}$. Thus, τ_E is much larger than galactic timescales and does not constrain the model. The reason is that the matter coupling is suppressed in dynamical situations.

4.6. Discussion

Usually, Cherenkov radiation is discussed for highly relativistic objects like cosmic rays. Above, we have considered non-relativistic perturbers like stars. That these can emit Cherenkov radiation is a novel feature typical of certain hybrid MOND dark matter models, as discussed above. Of course, relativistic perturbers can also emit Cherenkov radiation in these models. We have not discussed them here since our strict cutoffs do not allow useful constraints in this case. For example, for a proton with mass m_p, the upper cutoff $k_{\rm max}$ that we used in Sec. 4.2.4 would be

$$k_{\rm max} \sim \sqrt{\frac{M_\odot}{m_p}} 10^{-22}\,{\rm eV} \sim {\rm MeV} . \tag{4.97}$$

This is very high compared to the scales where the models discussed above are usually probed, see e.g. Ref. [71]. Thus, the models we discuss may not be valid at such scales. But even if we ignore this, we do not find a useful constraint. Concretely, the timescale τ_E would scale as[4]

$$\tau_E \sim \frac{E M_{\rm Pl}^2}{c_s^2 m_p^2 k_{\rm max}^2} \sim 10^{29}\,{\rm yr} , \tag{4.98}$$

where we used $c_s = 100\,{\rm km/s}$ and assumed a proton energy $E \sim 10^{11}\,{\rm GeV}$. This is much larger than the age of the universe even for much smaller proton energies.[5]

[4] Our calculation in Appendix A assumes a non-relativistic perturber. Very roughly, we expect $d\dot{E} \sim \omega d\Gamma \to (\omega\gamma)(d\Gamma/\gamma)$ in the relativistic case, where the Lorentz factor γ cancels. So we can reuse our formulas for $\dot{E}$ to get an order-of-magnitude estimate even for the relativistic case.

[5] This estimate assumes that the Cherenkov radiation mode couples to the proton's rest mass. Depending on the details of the model under consideration in the relativistic limit,

Above, we have used the Milky Way rotation curve to constrain standard SFDM. We could not go to radii larger than about 25 kpc because the stellar rotation curve is not measured at larger radii. The rotation curves of other galaxies are sometimes measured very precisely to much larger radii, see e.g. Ref. [31]. However, we cannot readily use the rotation curve data from Ref. [31] since these rotation curves are not obtained from stars but from gas. It may be possible to apply our results to gas clouds as a whole, depending on their specific size and mass. Alternatively, we could try to apply our results to individual neutral hydrogen atoms. Naively, our minimum radius $r_{\min} = 1/k_{\max}$ is roughly 10 fm in this case. This corresponds to $k_{\max} \sim$ MeV which may be too large for the models discussed here, as already mentioned above. But even if we ignore this, our naive $r_{\min}$ of ~ 10 fm is much smaller than the Bohr radius $\sim 10^5$ fm. Thus, we must anyway decrease our usual cutoff by a factor of about 10^4 if we want to treat the hydrogen atoms as point particles. This enhances τ_E by a factor of about 10^8 so that we do not get useful constraints. A potential way around these issues is to argue that, from astrophysical considerations, stellar rotation curves trace those obtained from gas and restrict the analysis to sufficiently small radii where the stellar number density is still significant. Investigating this in more detail is left for future work.

A possible limitation of our results is that our point-particle approximation for stars breaks down if stars are closer to each other than their effective size $r_{\min} = 1/k_{\max}$. Numerically,

$$r_{\min} = 3 \cdot 10^{-5}\,\mathrm{kpc} \cdot \sqrt{\frac{M}{M_\odot}} \cdot \frac{1}{f_p} \cdot \sqrt{\frac{a_0}{a_b^{\mathrm{gal}}}}\,, \tag{4.99}$$

for $a_0 = 1.2 \cdot 10^{-10}\,\mathrm{m/s^2}$. This is much smaller than the typical distance between stars of about 10^{-3} kpc in the Milky Way. Thus, this does not affect our results. Another possible effect is that stars could absorb some of the Cherenkov radiation emitted by other stars. In this way, stars may regain some of the energy lost to Cherenkov radiation. Of course, purely geometrically, some of the Cherenkov radiation should be able to escape the galaxy, since Cherenkov radiation is dominated by wavelengths of order $1/k_{\max}$, while stars are typically much further away from each other than $1/k_{\max}$. In addition, the coupling of Cherenkov radiation to matter is suppressed by $1/M_{\mathrm{Pl}}$ so that reabsorption happens only rarely. Thus, we expect this to be only a small effect. We leave a more detailed investigation for future work.

Another interesting effect was discovered by Ref. [130]. Namely, when $r_{\max}$ is comparable to the Jeans length r_J of a superfluid, the dispersion relation may be modified so that Cherenkov radiation is possible for arbitrarily low

the proton rest mass might have to be replaced by its energy, see for example Ref. [124]. In this case, the cutoff would be smaller but the energy loss rate might be enhanced due to the larger coupling. Which effect wins out needs to be checked separately for each specific model. In some models, the coupling might even be suppressed compared to the non-relativistic limit. For example for scalar fields that are coupled to the trace of the energy-momentum tensor, see Ref. [136].

velocities V. This concerns perturbations with large wavelengths. Thus, this would not affect our result for the energy loss which is dominated by short wavelengths. However, it means there may not be a hard cutoff of the energy loss at a critical velocity. That is, evading our constraints by increasing the sound speed, as we have done in standard SFDM, would only decrease the energy loss but not shut it down entirely. Of course, this applies only if our cutoff $r_{\max} \sim$ kpc is larger than the Jeans length r_J. So possibly our strict cutoffs do not allow to directly see this effect in our case.

4.7. Summary

In modified gravity theories, one usually tries to avoid superluminal sound speeds for theoretical reasons. But, for empirical reasons, one must also be careful with subluminal sound speeds. Specifically, a subluminal sound speed often leads to gravitational Cherenkov radiation which allows various astrophysical objects to lose energy.

In this chapter, we have discussed a new type of Cherenkov radiation that is often possible in hybrid MOND dark matter models with a common origin for the cosmological and galactic scale phenomena. Such models typically contain a massless mode that couples directly to normal matter for the MOND phenomenology. This same mode often has a non-relativistic sound speed to account for the CDM phenomenology. This allows even non-relativistic objects like stars to lose energy.

In our calculation, we use a controlled approximation that relies only on the MOND regime of each model, not on the Newtonian, high-acceleration regime. This avoids technical issues with the non-linearities of MOND and the possibly non-trivial high-acceleration behavior of each model. The price to pay is that we obtain only a conservative lower bound on the energy loss through Cherenkov radiation. The actual energy loss may be much larger but is also much more difficult to calculate.

We have first discussed the idea behind this new type of Cherenkov radiation for a prototype model and then applied the results to the original SFDM model, two-field SFDM, and the SZ model. For standard SFDM, we could rule out a MOND limit in the Milky Way for part of the parameter space. Two-field SFDM avoids these constraints by weakening the link between the cosmological and galactic phenomena. The relevant massless mode is coupled to normal matter only indirectly through a mixing. The SZ model avoids our constraints by having a matter coupling that is a standard gravitational coupling only in the static limit. In dynamical situations, the matter coupling is suppressed.

These results do not completely rule out any of the models considered. Still, we have demonstrated that our method is powerful and can severely constrain hybrid MOND dark matter models. This method can be improved upon in the future and it can be applied to more models. Any hybrid MOND dark matter model must satisfy our constraints.

5. Conclusion

We now summarize our results and give a brief outlook.

5.1. Summary

In this thesis, we discussed various aspects of hybrid MOND dark matter models, both from an observational and a theoretical perspective. We started with two observational tests of the SFDM model.

The first test was to check whether or not SFDM can explain strong lensing observations. Lensing is non-trivial to get right in MOND-like models because often a MOND-like force cannot act on photons. The reason is that the observation of GW170817 and its electromagnetic counterpart require that the gravitational tensor mode and light propagate at the same speed. A MOND-like force acting on photons often introduces a significant Shapiro time delay for the photons. Thus, the MOND-like phonon force in SFDM cannot contribute to the lensing signal. Still, we find that SFDM can fit the strong lensing data. The superfluid halo provides sufficient additional mass that generates the strong lensing signal even without the MOND-like phonon force.

For this to work, the superfluid halo's gravitational pull must be large at large galactocentric radii, where most of the lensing signal is generated. But it must be small at smaller galactocentric radii, where rotation curves are measured. Otherwise rotation curve velocities come out too large since stars are affected by both the MOND-like phonon force and the superfluid's standard Newtonian gravitational pull. This may be possible in SFDM because the superfluid has a very cored density profile. Indeed, we have checked that SFDM can simultaneously fit the Einstein radii and the velocity dispersions of the lensing galaxies.

The second test was the Milky Way rotation curve. We have shown that SFDM gives a reasonable fit of the observed rotation curve at $R \lesssim 25\,\mathrm{kpc}$, but needs about 20% less baryonic mass compared to standard MOND models. The reason is that, compared to SFDM, the total acceleration in standard MOND models approaches its limits for small ($a_b \ll a_0$) and large ($a_b \gg a_0$) accelerations much faster.

We have further estimated the size of the Milky Way's superfluid core and its total dark matter mass. For this, we have first demonstrated that the standard estimates for the size of the superfluid core for spherically symmetric situations can be straightforwardly extended to axisymmetric situations. This works because imposing spherically symmetric boundary conditions at large radii is reasonable. The boundary conditions then have one free parameter, μ_∞, which

determines the size of the superfluid halo and the total dark matter mass. We find that for virial dark matter masses M_{200}^{DM} in the range $0.5 - 3.0 \cdot 10^{12}\, M_\odot$ the so-called NFW radius varies between 65 kpc and 73 kpc, while the so-called thermal radius varies between 67 kpc and 105 kpc.

Next we had a closer look at SFDM from a theoretical perspective. We have identified three problems that cannot be easily avoided within SFDM. The first problem is that the equilibrium is unstable. This instability is usually avoided by introducing finite-temperature corrections parametrized by a parameter $\bar{\beta}$. But both the form of these finite-temperature corrections and the numerical value of $\bar{\beta}$ are completely ad-hoc and may easily be unphysical.

The second problem is that many galaxies cannot reach the MOND limit of SFDM, although this MOND limit is one of the main motivations behind SFDM. The MOND limit is controlled the quantity ε_*. The phonon force satisfies a MOND-like equation only in the MOND limit $\varepsilon_* \ll 1$. However, for the choice $\bar{\beta} = 2$, isolated galaxies can have a MOND-like rotation curve even when $\varepsilon_* = \mathcal{O}(1)$, although the phonon force does not satisfy a MOND-like equation. This is what we call the pseudo-MOND limit of SFDM. But this pseudo-MOND limit is not a satisfactory substitute for a proper MOND limit because it depends sensitively on the details of the ad-hoc finite-temperature corrections discussed in the previous paragraph. For example, the pseudo-MOND limit works only for $\bar{\beta} \approx 2$. Thus, like the finite-temperature corrections, this pseudo-MOND limit may easily be unphysical.

The third problem is that the equilibrium of galaxies in SFDM may not be valid on timescales much longer than galactic timescales. The reason is that the MOND-like phonon force requires a direct coupling of baryons to the phonon field θ without derivatives. This explicitly breaks the $U(1)$ shift symmetry of θ that is typical for superfluids. Heuristically, a chemical potential μ corresponds to a solution $\theta = \mu \cdot t$. Such solutions give an explicit time-dependence in the baryon coupling where θ appears without derivatives. We can ignore this time-dependence and assume a time-independent equilibrium in galaxies only on timescales much shorter than a timescale t_Q. We have estimated this timescale and found that it may be comparable to galactic timescales. Thus, galaxies in SFDM may not be in equilibrium, in contrast to what is usually assumed.

All three of these problems have the same root cause. Namely that the phonon field θ plays a double role. It carries both the superfluid and the MOND-like phonon force. We have proposed an improved model which avoids these problem. This improved model works by splitting the two roles between two fields. One field carries the superfluid, the other field carries the MOND-like force. We refer to this model as two-field SFDM. We have shown that the phenomenology on galactic scales is similar to that of standard SFDM, i.e. the original SFDM model.

More precisely, galactic scale phenomenology of two-field SFDM is close to that of standard SFDM regarding the superfluid core. Outside the superfluid core, this is true only if we use the so-called NFW matching procedure to

estimate the superfluid core's size. In standard SFDM, this matching procedure agrees reasonably well with the so-called thermal matching procedure. But in two-field SFDM, thermal matching gives a wildly different answer than NFW matching. This raises the question whether the NFW, the thermal, or a completely different matching procedure is correct in two-field SFDM.

At least in principle, this problem is already present in standard SFDM. The NFW and thermal matching procedures give similar answers only if the combination σ/m^5 has the value usually assumed in standard SFDM. In general, these two procedures give wildly different answers also in standard SFDM. This is important because the value of the self-interaction cross-section σ is usually chosen ad-hoc and not calculated from any Lagrangian in standard SFDM. Thus, it is not clear whether or not the assumed value is realistic. In any case, the transition from the superfluid core at smaller radii to the NFW halo at larger radii needs to be investigated in more detail.

Finally, we have introduced a novel observational test regarding Cherenkov radiation. This test applies not only to SFDM but to a wide variety of hybrid MOND dark matter models. Specifically, hybrid models with a common origin for galactic and cosmological phenomena. Cherenkov radiation is a well-known phenomenon in modified gravity models. If matter is coupled to a massless mode that propagates with speed c_s and a matter object moves faster than c_s, the object radiates away energy in the form of the massless mode. Usually, only highly relativistic objects emit Cherenkov radiation since $c_s \approx 1$. In hybrid models with a common origin for galactic and cosmological phenomena, this is different. Even non-relativistic objects like stars can emit Cherenkov radiation. The reason is as follows. For the MOND phenomenology on galactic scales, there is a massless mode that is coupled directly to normal matter. For the CDM phenomenology on cosmological scales, there is a collisionless fluid with non-relativistic sound speed. If these phenomena share a common origin, the massless mode that is directly coupled to matter typically has $c_s \ll 1$, since it is related to the non-relativistic sound speed of the collisionless fluid. Thus, even non-relativistic objects like stars may emit Cherenkov radiation in such hybrid models.

On the one hand, this Cherenkov radiation is similar to standard gravitational Cherenkov radiation in that it comes from a direct coupling to matter. On the other hand, it is similar to standard dynamical friction in collisional fluids in that it allows even non-relativistic objects to lose energy. For our specific calculation, it is similar to dynamical friction also in that there is no significant recoil. This is because each individual Cherenkov emission is soft. A supersonic object loses a significant amount of energy only through a large number of these soft emissions. This is a consequence of the strict cuts we impose on the phase space integrals in our calculation. The purpose of these cuts is twofold. First, they avoid problems with the inherent non-linearity in any MOND model. Second, they ensure we rely only on the MOND regime of a given model. Outside this MOND regime, models may behave very differently. These cuts also imply that our calculated energy loss is a lower bound.

The actual energy loss may be larger, but is much harder to calculate.

We have first considered a prototype model with a Lagrangian typical for hybrid MOND dark matter models. With this, we have calculated the timescale τ_E on which a non-relativistic, supersonic object with velocity V loses a significant amount of its energy through Cherenkov radiation. We found that τ_E is proportional to $V^3/(g_m^2\, c_s^2\, a_b^{\mathrm{gal}})$ where c_s is the sound speed of the massless mode that is coupled to normal matter with a coupling constant $g_m\sqrt{G}$ and a_b^{gal} is the baryonic Newtonian acceleration of the host galaxy at the object's position. This is independent of the mass of the object under consideration.

Hybrid models are ruled out unless they satisfy at least one of two conditions. Either their sound speed is large enough so that most non-relativistic objects like stars are subsonic. That is, Cherenkov radiation is kinematically forbidden. Or the timescale τ_E is sufficiently large so that supersonic objects emitting Cherenkov radiation don't lose much energy on galactic timescales. This latter case can be achieved by having a small coupling constant g_m or a small sound speed c_s.

We have applied our general results to three concrete models: To standard SFDM, to two-field SFDM, and to the SZ model. For standard SFDM, the phonon field θ gives a massless mode that is directly coupled to matter (for the MOND-like phonon force) and has a non-relativistic sound speed (since it is related to the non-relativistic superfluid around galaxies). This allows for the type of Cherenkov radiation we have discussed above. In particular, we can use the observed Milky Way rotation curve to constrain the MOND limit of standard SFDM. We have ruled out a MOND limit in the Milky Way for values of $\sqrt{\bar{\alpha}}/m$ in an interval that depends on the parameter $\bar{\beta}$. For example, for $\bar{\beta} = 2$, we have ruled out a MOND limit with $0.34\,\mathrm{eV}^{-1} \lesssim \sqrt{\bar{\alpha}}/m \lesssim 3.29\,\mathrm{eV}^{-1}$. This excludes the fiducial values $\bar{\beta} = 2$ and $\sqrt{\bar{\alpha}}/m = 2.4\,\mathrm{eV}^{-1}$ that are widely used in standard SFDM. Globular clusters and hypervelocity stars may give even more stringent constraints, but require a more sophisticated modelling of their trajectories. This is left for future work.

For two-field SFDM, there are two massless modes, roughly corresponding to the two fields that carry the MOND-like force and the superfluid, respectively. Of these, only the mode corresponding to the superfluid allows for Cherenkov radiation, since only this mode is subluminal. However, this mode is coupled to normal matter only indirectly through a mixing. This is because this mode is not directly responsible for the MOND-like force. As a result, as in standard SFDM, non-relativistic objects like stars do lose energy through Cherenkov radiation in two-field SFDM. But the associated energy loss is much smaller because the coupling to matter is suppressed. Indeed, typically $\tau_E \sim 10^{17}\,\mathrm{yr}$. Thus, this model evades our constraints by weakening the link between galactic and cosmological phenomena.

For the SZ model, there is again a massless mode that allows for Cherenkov radiation. This model evades our constraints by a new mechanism: In the static limit, the coupling to normal matter is a standard gravitational coupling. But in dynamical situations with $\omega \gg Q_0/\sqrt{K_{\mathrm{B}}}$, the coupling is suppressed

by powers of $Q_0/(\omega\sqrt{K_\mathrm{B}})$. As a result, non-relativistic objects like stars may emit Cherenkov radiation, but lose a significant amount of energy only after a time that is much longer than the age of the Universe.

To sum up, this novel type of Cherenkov radiation may severely constrain hybrid models (like standard SFDM), but there are also mechanisms to avoid such constraints (like in two-field SFDM and the SZ model). Any future hybrid models, or improvements of existing models, will have to satisfy our constraints.

5.2. Outlook

The evidence for both a collisionless fluid on cosmological scales and MOND-like scaling relations on galactic scales has only become more convincing in recent years. As a result, finding good explanations to account for both of these at the same time is increasingly important. Hybrid MOND-DM models are one promising approach to attack this. Future work must continue to observationally test and theoretically develop such models.

One important area of study is the interplay between the MOND and dark matter components in hybrid models. For example, we have shown that SFDM can simultaneously fit Einstein radii and velocity dispersions of observed lensing galaxies. But an open question is whether SFDM can have sufficient superfluid mass for strong lensing at the same time as naturally MOND-like rotation curves and velocity dispersions. This is nontrivial for two reasons.

First, in SFDM specifically, the phonon force is naturally MOND-like only in a certain limit, $\varepsilon_* \ll 1$, which may not be compatible with having a sufficiently large superfluid mass for strong lensing. For example, it needs to be checked whether or not our strong lensing fits satisfy the condition $\varepsilon_* \ll 1$ in the inner parts of the lensing galaxies where the velocity dispersions are measured. Second, more generally, if the dark matter mass around a galaxy becomes too large, the total acceleration will not be dominated by the MOND-like force, but will have significant contributions from the Newtonian acceleration of the dark matter component. Then, the total acceleration will not be naturally MOND-like. This is a potential issue in all hybrid MOND-DM models.

Another test of the interplay between the MOND-like force and the dark matter component in galaxies is suggested by a recent measurement of the RAR. In Ref. [137], the authors extend the observed RAR to smaller accelerations (larger radii) using weak lensing. This extended RAR follows the predictions of pure MOND. Hybrid models naturally follow the MOND predictions only if the dark matter component remains subdominant. But this is often not true at sufficiently large radii. So it may be non-trivial for hybrid models to naturally explain this new observed RAR at larger radii.

Hybrid MOND-DM models allow for some new phenomena like Cherenkov radiation emitted from non-relativistic sources like stars, as we have discussed in Chapter 4. In future work, constraints from this phenomenon should be applied to more models and to more general situations, e.g. not just to stars

on circular orbits but also to hypervelocity stars or dwarf galaxies. Similarly, a more complete calculation of this Cherenkov radiation from non-relativistic sources in the SZ model is warranted, including a proper galaxy background to see whether or not this constrains the model.

But even standard observational tests like rotation curves remain to be fully explored. For example, we have seen in Sec. 2.2 that SFDM requires about 20% less baryonic mass than standard MOND models to fit the Milky Way rotation curve. It will be important to see whether or not rotation curves in SFDM require less baryonic mass more generally and whether or not this is compatible with astrophysical constraints, for example from stellar population synthesis models.

On the purely theoretical side, understanding existing and developing new hybrid models is another important area of study. For example, in both standard SFDM and two-field SFDM the transition between the superfluid and the non-superfluid phases is not well-understood. In the fully-relativistic SZ model, so far little is known about the dark matter component in galaxies or galaxy clusters. More generally, other novel theoretical approaches to MOND (e.g. based on non-metricity [138, 139]) remain to be explored in more detail and in the context of hybrid MOND-DM models.

We expect hybrid MOND dark matter models to play an increasingly important role in the future. It will be exciting to learn what these models have to teach us about the universe.

Acknowledgements

I would like to thank everyone who helped and supported me during my years in Frankfurt.

First of all thank you to Dr. Sabine Hossenfelder for advising me, for countless insightful discussions about physics and many other topics, and for plenty of opportunities to present and talk about my work. Thank you to Prof. Dr. Marcus Bleicher for accepting me as a PhD student and for offering help whenever needed. Thank you to Prof. Dr. Horst Stöcker for agreeing to read and review this thesis. I would also like to thank everyone at FIAS for their support and for creating an institute with a diverse set of research areas. On this note, thank you to Prof. Dr. Christoph von der Malsburg, I always enjoyed our discussions over lunch. Thank you to Prof. Dr. Stacy McGaugh for many patient explanations, these were invaluable to me.

A huge thank you to Lisa for being always there for me and for being who you are. Thank you to Konne and Patrick for all the time on your couch and the analytical continuation of that in more recent times. Thank you to Sebastian and Markus for your friendship – it means a lot to me. And last but not least thank you to my parents for the life-long support which made all of this possible.

A. Calculation of energy loss through Cherenkov radiation

Consider the prototype model from Sec. 4.2. Consider further an incoming, non-relativistic matter particle with initial four-momentum $p_i^\alpha = (M + \frac{1}{2}MV^2, M\vec{V})$. For concreteness, we assume $\vec{V} = (0, 0, V)$. When this matter particle emits Cherenkov radiation, the Cherenkov radiation carries away four-momentum $k^\alpha = (\omega, \vec{k})$ which leaves the matter particle with a final four-momentum $p_f^\alpha = (E_f, \vec{p}_f) = (M + \frac{1}{2}MV'^2, M\vec{V}')$. We have [140]

$$d\Gamma = \frac{1}{2M}\frac{d^3\vec{p}_f}{(2\pi)^3\,2E_f}\frac{d^3\vec{k}}{(2\pi)^3\,2\omega}|\mathcal{M}|^2(2\pi)^4\delta^{(4)}(p_i - p_f - k)\,. \tag{A.1}$$

And we can calculate the energy loss per time from this,

$$\frac{dE}{dt} = -\int \omega d\Gamma\,. \tag{A.2}$$

The coupling term in the Lagrangian Eq. (4.4) is $-M^2\chi^2\, g_m\delta/(\sqrt{2}M_{\rm Pl})$. The Feynman rules give a factor of 2! for the two factors of χ. Thus, we have

$$|\mathcal{M}| = \frac{\sqrt{2}\bar{c}g_m}{M_{\rm Pl}}M^2\,. \tag{A.3}$$

The factor $\bar{c}$ is because we need to canonically normalize our field, $\delta \to \bar{c}\cdot\delta$, so that we can apply the standard Feynman diagram formalism. Here, we assume that g_m does not depend on $\vec{k}$. This is not true in two-field SFDM and in the SZ model. Corrections due to this are discussed in Appendix A.4 and in Sec. 4.5.2.

Then, using $\omega = c_s|\vec{k}|$,

$$\begin{aligned} d\Gamma &= \frac{1}{(2M)^2}\frac{1}{(2\pi)^2}\frac{d^3\vec{k}}{2\omega}|\mathcal{M}|^2\delta\left(\frac{1}{2}MV^2 - \frac{1}{2}M(\vec{V} - \vec{k}/M)^2 - c_s|\vec{k}|\right) \\ &= \frac{1}{(2M)^2}\frac{1}{(2\pi)^2}\frac{d^3\vec{k}}{2\omega}|\mathcal{M}|^2\delta\left(V|\vec{k}|\cos\theta - \left(\frac{|\vec{k}|^2}{2M} + c_s|\vec{k}|\right)\right) \\ &= \frac{1}{(2M)^2}\frac{1}{(2\pi)^2}\frac{d|\vec{k}||\vec{k}|\,d\varphi\,d\cos\theta}{2\omega V}|\mathcal{M}|^2\delta\left(\cos\theta - \left(\frac{|\vec{k}|}{2MV} + \frac{c_s}{V}\right)\right)\,, \end{aligned} \tag{A.4}$$

where θ is the angle between $\vec{k}$ and $\vec{V}$. We see that the delta function requires $|\vec{k}|$ to be not too large. In particular, we need to cut off the $|\vec{k}|$ integral at

$$|\vec{k}| = 2M(V - c_s) = 2M\left(V - \bar{c}\sqrt{1+\gamma^2}\right)\,, \tag{A.5}$$

where γ is the cosine of the angle between $\vec{k}$ and $\hat{a}$. This condition implies that there will be energy loss due to Cherenkov radiation only for supersonic velocities $V > c_s$.

Numerically, we are interested in $M \sim M_\odot$ and we impose an additional cutoff $k_{\max} \sim 10^{-22}\,\text{eV}$ to stay in the regime of validity of our calculation, see Sec. 4.2.4. In this case, the c_s/V term dominates over the $k/(2MV)$ term in the delta function. Thus, this delta function becomes

$$\delta\left(\cos\theta - \frac{\bar{c}}{V}\sqrt{1+\gamma^2}\right) . \tag{A.6}$$

In general, $\gamma = \vec{k}\cdot\hat{a}/k$ depends on $\cos\theta$ in a complicated way. Thus, analytically evaluating this delta function is nontrivial. Here, we first give the result for the simpler case $\hat{a} = 0$ where the $\sqrt{1+\gamma^2}$ factor in c_s is absent. In the following sections, we then give the result including $\hat{a}$ for the special cases $\vec{V} \perp \hat{a}$ and $\vec{V} \parallel \hat{a}$. For $\hat{a} = 0$, we have

$$\begin{aligned} |\dot{E}|^{\hat{a}=0} &= \int_{k_{\min}}^{k_{\max}} d|\vec{k}| \frac{1}{(2M)^2}\frac{1}{(2\pi)}\frac{|\vec{k}|}{2V}|\mathcal{M}|^2 \\ &= \frac{1}{(2M)^2}\frac{1}{(2\pi)}\frac{1}{4V}(k_{\max}^2 - k_{\min}^2)\frac{2\bar{c}^2 g_m^2 M^4}{M_{\text{Pl}}^2} \\ &= \frac{\bar{c}^2}{16\pi V}\frac{g_m^2 M^2}{M_{\text{Pl}}^2}(k_{\max}^2 - k_{\min}^2) . \end{aligned} \tag{A.7}$$

A.1. Special case: $\vec{V} \parallel \hat{a}$

In this case, we have

$$\gamma^2 = \frac{(\vec{k}\hat{a})^2}{|\vec{k}|^2} = \cos^2\theta , \tag{A.8}$$

since θ is the angle between $\vec{V}$ and $\vec{k}$ and $\vec{V} \parallel \hat{a}$. Then, the delta function Eq. (A.6) in $d\Gamma$ can be written as

$$\delta\left(\cos\theta - \frac{\bar{c}}{V}\sqrt{1+\cos^2\theta}\right) . \tag{A.9}$$

This is solved by

$$\cos\theta = \frac{1}{\sqrt{(V/\bar{c})^2 - 1}} . \tag{A.10}$$

Thus, Cherenkov radiation requires

$$V > \sqrt{2}\bar{c} , \tag{A.11}$$

otherwise $\cos\theta$ would be larger than 1. Evaluated at this solution, the delta function becomes

$$\frac{1}{1-(\bar{c}/V)^2}\,\delta\left(\cos\theta - \frac{1}{\sqrt{(V/\bar{c})^2-1}}\right)\,, \tag{A.12}$$

which implies

$$|\dot{E}| = \frac{1}{1-(\bar{c}/V)^2}\cdot|\dot{E}|^{\hat{a}=0}\,. \tag{A.13}$$

A.2. Special case: $\vec{V} \perp \hat{a}$

For simplicity, we choose our coordinate system such that $\hat{a}$ points into the positive x direction. This is perpendicular to $\vec{V}$ which points into the positive z direction. Then,

$$\gamma^2 = \frac{(\vec{k}\hat{a})^2}{|\vec{k}|^2} = \sin^2\theta\,\cos^2\varphi\,. \tag{A.14}$$

That is, the delta function Eq. (A.6) in $d\Gamma$ becomes

$$\delta\left(\cos\theta - \frac{\bar{c}}{V}\sqrt{1+(1-\cos^2\theta)\cos^2\varphi}\right)\,. \tag{A.15}$$

This is solved by

$$\cos\theta = \frac{\bar{c}}{V}\sqrt{\frac{1+\cos^2\varphi}{1+(\bar{c}/V)^2\,\cos^2\varphi}}\,. \tag{A.16}$$

As a result, Cherenkov radiation requires

$$V > \bar{c}\,, \tag{A.17}$$

otherwise $\cos\theta$ would be larger than 1. Evaluated at this solution, the delta function becomes

$$\frac{1}{1+(\bar{c}/V)^2\cos^2\varphi}\,\delta\left(\cos\theta - \frac{\bar{c}}{V}\sqrt{\frac{1+\cos^2\varphi}{1+(\bar{c}/V)^2\cos^2\varphi}}\right)\,. \tag{A.18}$$

After doing the φ integral, this implies

$$|\dot{E}| = \frac{1}{\sqrt{1+(\bar{c}/V)^2}}\cdot|\dot{E}|^{\hat{a}=0}\,. \tag{A.19}$$

A.3. Direction of friction force

Above, we calculated the total energy loss, but not the direction in which the Cherenkov emission happens, i.e. the direction in which the resulting effective friction force acts. This direction is proportional to the direction of the following integral

$$\int \vec{k}\, d\Gamma \propto \int_{-1}^{+1} d\cos\theta \int_0^{2\pi} d\varphi \frac{1}{\sqrt{1+\gamma^2}} \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix} \delta\left(\cos\theta - \bar{c}/V\sqrt{1+\gamma^2}\right) . \tag{A.20}$$

Here, as in the previous section, we choose y to be the direction that is perpendicular to both $\vec{V}$ and $\hat{a}$. For the two special cases $\vec{V} \parallel \hat{a}$ and $\vec{V} \perp \hat{a}$, the friction force points exactly in the z direction, i.e. in the direction of $\vec{V}$. This follows by using the delta function to do the $\cos\theta$ integral and then noticing that the integrand depends on φ only through $\cos^2\varphi$, except for the factors $\cos\varphi$ and $\sin\varphi$ in the x and y components of $\vec{k}$. That is, the x and y components vanish for these two special cases.

For a general orientation of $\vec{V}$ relative to $\hat{a}$, it is still true that the force in the y direction vanishes. That is, the force stays in the x-z plane spanned by $\vec{V}$ and $\hat{a}$. This can be seen analytically by using the delta function to do the φ integral. This gives two solutions for $\cos\varphi$,

$$\cos\varphi = \frac{1}{a_\perp \sin\theta}\left(\pm\sqrt{\left(\frac{V}{\bar{c}}\right)^2 \cos^2\theta - 1} - a_\parallel \cos\theta\right) , \tag{A.21}$$

where $a_\perp$ is the component of $\hat{a}$ in the x direction and $a_\parallel$ that in the z direction. Each of these solutions for $\cos\varphi$ corresponds to two different values of φ in the interval $(0, 2\pi)$. These two values of φ give the same $\cos\varphi$, but $\pm\sin\varphi$. The integrand in Eq. (A.20) depends on the sign of $\sin\varphi$ only through the factor $\sin\varphi$ in the y component. Thus, the y component of the force vanishes, since the contributions from the two different values of φ cancel each other.

In contrast, the x component does not vanish in general. The integrals are hard to evaluate analytically, but we have numerically verified that the x component can be as large as a few 10% of the z component. Thus, in general, the friction force does not act in the direction of $\vec{V}$.

A.4. Two-field SFDM corrections

As discussed in Sec. 4.4 (see also Appendix D below), two-field SFDM differs from our prototype model in two ways. First, the sound speed c_s does not have the $\sqrt{1+\gamma^2}$ factor, so that $c_s = \bar{c} = \sqrt{\hat{\mu}/m}$ is independent of the direction of $\vec{k}$. Second, the matter coupling g_m depends on this direction through the γ-dependent factor $\gamma/(1+\gamma^2)$.

That c_s has no factor $\sqrt{1+\gamma^2}$ means that the energy-conserving delta function gives $\cos\theta = c_s/V$. Thus, there is Cherenkov radiation as long as $V > c_s = \sqrt{\hat{\mu}/m}$ and the correction factor f_a is absent.

The additional factor $\gamma^2/(1+\gamma^2)^2$ in $|\mathcal{M}|^2$ modifies the result for $|\dot{E}|$. We again orient our coordinate system such that $\vec{V}$ points in the z direction and $\hat{a}$ has zero component in the y direction. Then

$$\gamma^2 = \frac{(\vec{k}\hat{a})^2}{|\vec{k}|^2} = (a_z \cos\theta + a_x \sin\theta\, \cos\varphi)^2 \,. \tag{A.22}$$

If, for simplicity, we set $\gamma = 1$, we get from the φ integration in $|\dot{E}|$

$$\int d\varphi \, \frac{\gamma^2}{(1+\gamma^2)^2}\bigg|_{\gamma=1} = \frac{2\pi}{4} \,. \tag{A.23}$$

For general γ, the φ integral can be done analytically for the special cases $\vec{V} \parallel \hat{a}$ and $\vec{V} \perp \hat{a}$. For $\hat{a} \parallel \vec{V}$, we have $a_x = 0$ and $a_z^2 = 1$, so that

$$\int d\varphi \, \frac{\gamma^2}{(1+\gamma^2)^2}\bigg|_{\vec{V}\parallel\hat{a}} = 2\pi \frac{\cos^2\theta}{(1+\cos^2\theta)^2} = 2\pi \left(\frac{c_s}{V} + \frac{V}{c_s}\right)^{-2} \,. \tag{A.24}$$

For $\hat{a} \perp \vec{V}$, we have $a_z = 0$ and $a_x^2 = 1$, so that

$$\begin{aligned} \int d\varphi \, \frac{\gamma^2}{(1+\gamma^2)^2}\bigg|_{\vec{V}\perp\hat{a}} &= \int d\varphi \frac{\sin^2\theta\cos^2\varphi}{(1+\sin^2\theta\cos^2\varphi)^2} \\ &= 2\pi\,\frac{1}{2}\frac{\sin^2\theta}{(1+\sin^2\theta)^{3/2}} \\ &= 2\pi\,\frac{1}{2}\frac{1-\left(\frac{c_s}{V}\right)^2}{\left(2-\left(\frac{c_s}{V}\right)^2\right)^{3/2}} \,. \end{aligned} \tag{A.25}$$

Thus,

$$|\dot{E}| = 4|\dot{E}|^{\gamma=1} \cdot \begin{cases} \left(\frac{c_s}{V} + \frac{V}{c_s}\right)^{-2} \,, & \vec{V} \perp \hat{a} \\ \frac{1}{2}\frac{1-\left(\frac{c_s}{V}\right)^2}{\left(2-\left(\frac{c_s}{V}\right)^2\right)^{3/2}} \,, & \vec{V} \parallel \hat{a} \end{cases} \,. \tag{A.26}$$

For $\vec{V} \parallel \hat{a}$ and $\vec{V} \perp \hat{a}$, the direction of the effective friction force stays the same as in the case of our prototype Lagrangian. However, for a general orientation of $\vec{V}$, the additional γ-dependence may give corrections also to the direction of the force. Here, we do not calculate these since they are not important for our results.

B. Standard SFDM perturbations

Consider the SFDM Lagrangian Eq. (1.7). If we define

$$\begin{aligned} \bar{f}(K) &\equiv \frac{2\Lambda}{3}(2m)^{3/2}\sqrt{|K|}\,, \\ f(K) &\equiv \bar{f}(K)K\,, \end{aligned} \tag{B.1}$$

the kinetic term in this Lagrangian can be written as $\bar{f}(X-\bar{\beta}Y)X$. If $X-\bar{\beta}Y \approx X$, this becomes approximately $f(X)$. We now consider a background equilibrium solution θ_0 in the MOND limit, i.e. $(\vec{\nabla}\theta_0)^2 \gg 2m\hat{\mu}$. We expand X and Y in perturbations δ of the phonon field θ and introduce a formal expansion parameter ϵ,

$$\begin{aligned} X &= X_0 + \epsilon X_1 + \epsilon^2 X_2 + \dots\,, \\ Y &= Y_0 + \epsilon Y_1\,, \end{aligned} \tag{B.2}$$

with

$$\begin{aligned} X_1 &= \dot{\delta} - \frac{\vec{\nabla}\theta_0\vec{\nabla}\delta}{m}\,, \\ X_2 &= -\frac{(\vec{\nabla}\delta)^2}{2m}\,, \\ Y_1 &= \dot{\delta}\,, \end{aligned} \tag{B.3}$$

where we left out a possible ϵ^2 term in Y since Y is linear in the fields. We further define

$$\begin{aligned} f_0' &= f'(X_0) = (2m)\Lambda|\vec{\nabla}\theta_0|\,, \\ f_0'' &= f''(X_0) = -(2m)^2\frac{\Lambda}{2}\frac{1}{|\vec{\nabla}\theta_0|}\,, \\ \bar{f}_0' &= \bar{f}'(X_0) = -(2m)^2\frac{\Lambda}{3}\frac{1}{|\vec{\nabla}\theta_0|}\,, \\ \bar{f}_0'' &= \bar{f}''(X_0) = -(2m)^3\frac{\Lambda}{6}\frac{1}{|\vec{\nabla}\theta_0|^3}\,. \end{aligned} \tag{B.4}$$

For simplicity, we also write

$$\lambda \equiv \frac{\bar{\alpha}\Lambda}{M_{\rm Pl}}\,. \tag{B.5}$$

For a background galaxy in the MOND limit, spatial derivatives dominate so that $X_0 - \bar{\beta}Y_0 \approx X_0$. Then, the perturbed Lagrangian is

$$\begin{aligned}\mathcal{L} &= X_2 f_0' + \frac{1}{2}X_1^2 f_0'' - \bar{\beta}X_1 Y_1(\bar{f}_0' + X_0\bar{f}_0'') + \frac{1}{2}\bar{\beta}^2 Y_1^2 X_0 \bar{f}_0'' - \lambda\delta\,\delta_b \\ &= \Lambda|\vec{\nabla}\theta_0|\left(-(\vec{\nabla}\delta)^2 - \frac{m^2}{|\vec{\nabla}\theta_0|^2}\left(\dot{\delta} - \frac{\vec{\nabla}\theta_0\vec{\nabla}\delta}{m}\right)^2\right. \\ &\qquad \left. + \frac{2}{3}\frac{m^2}{|\vec{\nabla}\theta|^2}\bar{\beta}\dot{\delta}\left(\dot{\delta} - \frac{\vec{\nabla}\theta\vec{\nabla}\delta}{m}\right) + \frac{1}{3}\bar{\beta}^2\dot{\delta}^2\frac{m^2}{|\vec{\nabla}\theta_0|^2}\right) - \lambda\delta\,\delta_b \\ &= 2\Lambda|\vec{\nabla}\theta_0|\left(\dot{\delta}^2\frac{m^2}{|\vec{\nabla}\theta_0|^2}\frac{1}{6}(\bar{\beta}-1)(\bar{\beta}+3) - \frac{1}{2}(\vec{\nabla}\delta)^2 - \frac{1}{2}(\hat{a}\vec{\nabla}\delta)^2\right. \\ &\qquad \left. - \frac{1}{3}\frac{m}{|\vec{\nabla}\theta_0|}\hat{a}\vec{\nabla}\delta\dot{\delta}(\bar{\beta}-3)\right) - \lambda\delta\,\delta_b\,,\end{aligned} \tag{B.6}$$

where we set $\hat{a} = \vec{\nabla}\theta_0/|\vec{\nabla}\theta_0|$. With the definition

$$\bar{c}^{-1} = \frac{m}{|\vec{\nabla}\theta_0|}\sqrt{\frac{(\bar{\beta}-1)(\bar{\beta}+3)}{3}}\,, \tag{B.7}$$

this can be written as

$$\mathcal{L} = 2\Lambda|\vec{\nabla}\theta_0|\left(\frac{1}{2}\dot{\delta}^2\bar{c}^{-2} - \frac{1}{2}(\vec{\nabla}\delta)^2 - \frac{1}{2}(\hat{a}\vec{\nabla}\delta)^2 + \bar{c}^{-1}f_{\bar{\beta}}\hat{a}\vec{\nabla}\delta\dot{\delta}\right) - \lambda\delta\,\delta_b\,, \tag{B.8}$$

where, following Ref. [71, 72], we assumed $\bar{\beta} < 3$ and defined

$$f_{\bar{\beta}} = \sqrt{\frac{(3-\bar{\beta})^2}{3(\bar{\beta}-1)(\bar{\beta}+3)}}\,. \tag{B.9}$$

With an appropriate rescaling we finally get

$$\mathcal{L} = \frac{1}{2}\dot{\delta}^2\bar{c}^{-2} - \frac{1}{2}(\vec{\nabla}\delta)^2 - \frac{1}{2}(\hat{a}\vec{\nabla}\delta)^2 + \bar{c}^{-1}f_{\bar{\beta}}\hat{a}\vec{\nabla}\delta\dot{\delta} - \frac{g_m}{\sqrt{2}M_{\rm Pl}}\delta\,\delta_b\,, \tag{B.10}$$

with

$$g_m = \sqrt{\frac{a_0}{|a_{\theta_0}|}}\,, \tag{B.11}$$

where we used $a_0 = \bar{\alpha}^3\Lambda^2/M_{\rm Pl}$ and $\vec{a}_\theta = -\lambda\vec{\nabla}\theta$ [71].

C. Standard SFDM Cherenkov radiation

For standard SFDM, our simple QFT calculation from Appendix A does not apply because spatial and time derivatives are mixed in the Lagrangian for the perturbations. Thus, we do a classical calculation instead, along the lines of the standard calculation of electromagnetic Cherenkov radiation [135].

C.1. Reference calculation

For reference, we first consider a field $\bar{\phi}(\bar{t}, \vec{\bar{x}})$ with Lagrangian

$$\bar{\mathcal{L}} = \frac{1}{2}\frac{1}{c_{\text{eff}}^2}(\partial_{\bar{t}}\bar{\phi})^2 - \frac{1}{2}(\vec{\bar{\nabla}}\bar{\phi})^2 - \frac{g_m}{\sqrt{2}M_{\text{Pl}}}\bar{\phi}\,\rho_{\text{b,eff}}\,, \tag{C.1}$$

with some constant c_{eff}. This is our prototype Lagrangian from Sec. 4.2 but in the simpler case with $\hat{a} = 0$ and with $\bar{c}$ replaced by c_{eff}. We also introduced a modified baryonic density

$$\rho_{\text{b,eff}} = M_{\text{eff}}\delta(\bar{z} - V_{\text{eff}}\bar{t})\delta(\bar{x})\delta(\bar{y})\,, \tag{C.2}$$

with constants M_{eff} and V_{eff}. This is the usual form for the perturber's density used in classical Cherenkov radiation calculations [129, 130, 135]. Below, we will express solutions of the equation of motion of the standard SFDM perturbation field δ in terms of solutions of the equation of motion of $\bar{\phi}$.

We use the following convention for the Fourier transform and inverse Fourier transform of a function f

$$f(\omega, \vec{k}) = \frac{1}{(2\pi)^2}\int d^4x\, e^{-i\omega t + i\vec{k}\vec{x}} f(t, \vec{x})\,, \tag{C.3a}$$

$$f(t, \vec{x}) = \frac{1}{(2\pi)^2}\int d^4k\, e^{+i\omega t - i\vec{k}\vec{x}} f(\omega, \vec{k})\,. \tag{C.3b}$$

The equation of motion of $\bar{\phi}$ is

$$0 = -\frac{1}{c_{\text{eff}}^2}\partial_{\bar{t}}^2\bar{\phi} + \vec{\bar{\nabla}}^2\bar{\phi} - \frac{g_m}{\sqrt{2}M_{\text{Pl}}}M_{\text{eff}}\delta(\bar{z} - V_{\text{eff}}\bar{t})\delta(\bar{x})\delta(\bar{y})\,. \tag{C.4}$$

In Fourier space

$$0 = \left(\frac{\bar{\omega}^2}{c_{\text{eff}}^2} - \vec{\bar{k}}^2\right)\bar{\phi}(\bar{\omega}, \vec{\bar{k}}) - \frac{g_m}{\sqrt{2}M_{\text{Pl}}}\frac{M_{\text{eff}}}{2\pi}\delta(\bar{\omega} - V_{\text{eff}}\bar{k}_{\bar{z}})\,, \tag{C.5}$$

where we used $\int dk e^{ikx} = (2\pi)\delta(x)$ and introduced the notation $\bar{\omega}$ and $\vec{\bar{k}}$ to indicate that this is a Fourier transform with respect to $\bar{t}$ and $\vec{\bar{x}}$. This gives

$$\begin{aligned}\bar{\phi}(\bar{\omega}, \vec{\bar{k}}) &= \frac{g_m}{\sqrt{2}M_{\rm Pl}} \frac{M_{\rm eff}}{2\pi} \frac{\delta(\bar{\omega} - V_{\rm eff}\bar{k}_{\bar{z}})}{\frac{\bar{\omega}^2}{c_{\rm eff}^2} - \bar{k}^2} \\ &= \frac{1}{|V_{\rm eff}|} \frac{g_m}{\sqrt{2}M_{\rm Pl}} \frac{M_{\rm eff}}{2\pi} \frac{\delta\left(\frac{\bar{\omega}}{V_{\rm eff}} - \bar{k}_{\bar{z}}\right)}{\bar{\lambda}^2 - \bar{k}_\perp^2}\,,\end{aligned} \tag{C.6}$$

where

$$\bar{\lambda}^2 = \bar{\omega}^2 \left(\frac{1}{c_{\rm eff}^2} - \frac{1}{V_{\rm eff}^2}\right)\,, \tag{C.7}$$

and $\vec{\bar{k}}_\perp$ denotes the components of $\vec{\bar{k}}$ in the $\bar{x}$ and $\bar{y}$ directions. On-shell radiation (that is Cherenkov radiation) is possible only for $\bar{\lambda}^2 > 0$. Otherwise, the propagator has no pole that could be picked up. This gives the standard condition for Cherenkov radiation

$$|V_{\rm eff}| > c_{\rm eff}\,. \tag{C.8}$$

In this case, we have [135]

$$\begin{aligned}\bar{\phi}(\bar{t}, \vec{\bar{x}}) &= \frac{1}{(2\pi)^2} \int d^4\bar{k} e^{i\bar{\omega}\bar{t} - i\vec{\bar{k}}\vec{\bar{x}}} \frac{1}{|V_{\rm eff}|} \frac{g_m}{\sqrt{2}M_{\rm Pl}} \frac{M_{\rm eff}}{2\pi} \frac{\delta\left(\frac{\bar{\omega}}{V_{\rm eff}} - \bar{k}_{\bar{z}}\right)}{\bar{\lambda}^2 - \bar{k}_\perp^2} \\ &= \frac{1}{(2\pi)^2} \frac{1}{|V_{\rm eff}|} \frac{g_m}{\sqrt{2}M_{\rm Pl}} \frac{M_{\rm eff}}{2\pi} \int d\bar{\omega} e^{i\bar{\omega}\left(\bar{t} - \frac{1}{V_{\rm eff}}\bar{z}\right)} \\ &\qquad \times \int d^2\vec{\bar{k}}_\perp e^{-i\vec{\bar{k}}_\perp \vec{\bar{x}}_\perp} \frac{1}{\bar{\lambda}^2 - \bar{k}_\perp^2} \\ &= -\frac{M_{\rm eff}}{(2\pi)^2} \frac{1}{|V_{\rm eff}|} \frac{g_m}{\sqrt{2}M_{\rm Pl}} \int d\bar{\omega} e^{i\bar{\omega}\left(\bar{t} - \frac{1}{V_{\rm eff}}\bar{z}\right)} K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)\,,\end{aligned} \tag{C.9}$$

where K_0 is the zeroth modified Bessel function of the second kind and $\vec{\bar{x}}_\perp$ denotes components of $\vec{\bar{x}}$ in the $\bar{x}$ and $\bar{y}$ directions. Here, $\bar{\lambda}$ is given by

$$\bar{\lambda} = \bar{\omega}\sqrt{\frac{1}{c_{\rm eff}^2} - \frac{1}{V_{\rm eff}^2}} \equiv \frac{\bar{\omega}}{c_{\rm eff}} f_{\rm crit}\,. \tag{C.10}$$

When doing the $\vec{\bar{k}}_\perp$ integral, we implicitly assumed the retarded propagator. For our convention of the Fourier transform, this gives the $K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)$. Ref. [135] uses the opposite sign in the $e^{\pm i\omega t}$ factors of the Fourier transform and therefore gets $\bar{\lambda} \to -\bar{\lambda}$ in the Bessel function's argument. We see that $\bar{\phi}$ depends on $\vec{\bar{x}}$ only through the combinations

$$\bar{z} - V_{\rm eff}\bar{t} \quad \text{and} \quad |\vec{\bar{x}}_\perp|\,. \tag{C.11}$$

Below, we are interested in the limit $|\bar{\lambda}\vec{\bar{x}}_\perp| \gg 1$. In this limit,

$$\bar{\phi}(\bar{t}, \vec{\bar{x}}) = -\sqrt{\frac{\pi}{2}} \frac{M_{\rm eff}}{(2\pi)^2} \frac{1}{|V_{\rm eff}|} \frac{g_m}{\sqrt{2}M_{\rm Pl}} \int d\bar{\omega} e^{i\bar{\omega}\left(\bar{t} - \frac{1}{V_{\rm eff}}\bar{z}\right)} \frac{e^{-i\bar{\lambda}|\vec{\bar{x}}_\perp|}}{\sqrt{i\bar{\lambda}|\vec{\bar{x}}_\perp|}}\,. \tag{C.12}$$

C.2. Mapping to the reference system

To calculate Cherenkov radiation for standard SFDM, we use the standard perturber density

$$\delta_b(t,\vec{x}) = M\delta(z-Vt)\delta(x)\delta(y)\,. \tag{C.13}$$

This assumes that the emitted Cherenkov radiation is soft. Otherwise, we could not assume the perturber to travel on a straight line with constant velocity. This is justified because we set very conservative limits on the integration boundaries. This is discussed in Appendix A and Sec. 4.2.3. The equation of motion for the perturbation δ is then

$$\begin{aligned} 0 = -\frac{1}{\bar{c}^2}\partial_t^2\delta + \vec{\nabla}^2\delta + \vec{\nabla}\left((\hat{a}\cdot\vec{\nabla}\delta)\hat{a}\right) - \frac{2f_{\bar{\beta}}}{\bar{c}}\hat{a}\vec{\nabla}\dot{\delta} \\ - \frac{g_m}{\sqrt{2}M_{\mathrm{Pl}}}M\delta(z-Vt)\delta(x)\delta(y)\,. \end{aligned} \tag{C.14}$$

To facilitate an analytical treatment, we discuss only the special cases $\vec{V} \parallel \hat{a}$ and $\vec{V} \perp \hat{a}$. Here, $\vec{V}$ points in the positive z direction.

For $\vec{V} \parallel \hat{a}$, we have $\hat{a} = a_\parallel \hat{e}_z$ where $\hat{e}_z$ is the unit vector in the z direction and $a_\parallel = \pm 1$. The solutions for $\delta(t,\vec{x})$ can be obtained from those for $\bar{\phi}(\bar{t},\vec{\bar{x}})$ by a coordinate transformation. Specifically, we take

$$\begin{aligned} \bar{t} &= t + a_\parallel \frac{f_{\bar{\beta}}}{2\bar{c}} z\,, \\ \bar{z} &= \frac{z}{\sqrt{2}}\,, \\ \vec{\bar{x}}_\perp &= \vec{x}_\perp\,. \end{aligned} \tag{C.15}$$

Then, δ defined as

$$\delta(t,\vec{x}) = \bar{\phi}(\bar{t},\vec{\bar{x}})\,, \tag{C.16}$$

solves the δ equation of motion if we take

$$c_{\mathrm{eff}} = \bar{c}\frac{1}{\sqrt{1+\frac{1}{2}f_{\bar{\beta}}^2}}\,, \tag{C.17a}$$

$$V_{\mathrm{eff}} = V\frac{\sqrt{2}}{2 + a_\parallel f_{\bar{\beta}}(V/\bar{c})}\,, \tag{C.17b}$$

$$M_{\mathrm{eff}} = M\frac{\sqrt{2}}{|2 + a_\parallel f_{\bar{\beta}}(V/\bar{c})|}\,, \tag{C.17c}$$

in the $\bar{\phi}$ solution. Note that c_{eff}, V_{eff}, and M_{eff} are constants. They do not depend on spacetime coordinates or fields. This immediately allows to find

the condition for Cherenkov radiation. As discussed above, $\delta(t, \vec{x})$ contains a pole only if $|V_{\text{eff}}| > c_{\text{eff}}$. This gives the critical velocity

$$V_{\text{crit}}^{\parallel} = \bar{c}\left(\sqrt{2 + f_{\bar{\beta}}^2} + a_{\parallel} f_{\bar{\beta}}\right) . \tag{C.18}$$

For $\vec{V} \perp \hat{a}$, we choose our coordinates such that $\hat{a}$ points in the positive x direction, $\hat{a} = \hat{e}_x$. We further choose

$$\begin{aligned} \bar{t} &= t + \frac{f_{\bar{\beta}}}{2\bar{c}} x \,, \\ \bar{x} &= \frac{x}{\sqrt{2}} \,, \\ \bar{y} &= y \,, \\ \bar{z} &= z \,. \end{aligned} \tag{C.19}$$

Then, δ defined as

$$\delta(t, \vec{x}) = \bar{\phi}(\bar{t}, \vec{\bar{x}}) \,, \tag{C.20}$$

solves the δ equation of motion if we take

$$c_{\text{eff}} = \bar{c} \frac{1}{\sqrt{1 + \frac{1}{2} f_{\bar{\beta}}^2}} \,, \tag{C.21}$$

$$V_{\text{eff}} = V \,, \tag{C.22}$$

$$M_{\text{eff}} = \frac{M}{\sqrt{2}} \,, \tag{C.23}$$

in the $\bar{\phi}$ solution. This gives the critical velocity

$$V_{\text{crit}}^{\perp} = \bar{c} \frac{1}{\sqrt{1 + \frac{1}{2} f_{\bar{\beta}}^2}} \,. \tag{C.24}$$

C.3. Energy loss without cutoffs

To calculate the energy loss through Cherenkov radiation following Ref. [135], we first need to know the energy flux in different spatial directions. The canonical energy density of the standard SFDM perturbations is

$$\mathcal{H} = \frac{1}{2} \frac{1}{\bar{c}^2} \dot{\delta}^2 + \frac{1}{2} (\vec{\nabla} \delta)^2 + \frac{1}{2} (\vec{\nabla} \delta \hat{a})^2 \,. \tag{C.25}$$

The $f_{\bar\beta}$ term that mixes spatial and time derivatives does not contribute since it is linear in time derivatives. From this we find, using the equation of motion

$$\begin{aligned}\partial_t\mathcal{H} &= \frac{\dot\delta\ddot\delta}{\bar c^2} + \vec\nabla\delta\vec\nabla\dot\delta + (\vec\nabla\delta\hat a)(\vec\nabla\dot\delta\hat a)\\ &= \dot\delta\left[\vec\nabla^2\delta + \vec\nabla\left((\vec\nabla\delta\cdot\hat a)\hat a\right) - \frac{2f_{\bar\beta}}{\bar c}\hat a\vec\nabla\dot\delta\right] + \vec\nabla\delta\vec\nabla\dot\delta + (\vec\nabla\delta\hat a)(\vec\nabla\dot\delta\hat a)\\ &= \left[\dot\delta\vec\nabla^2\delta + \vec\nabla\delta\vec\nabla\dot\delta\right] + \left[\dot\delta\vec\nabla\left((\vec\nabla\delta\cdot\hat a)\hat a\right) + (\vec\nabla\delta\hat a)(\vec\nabla\dot\delta\hat a)\right]\\ &\qquad - \vec\nabla\left(\frac{f_{\bar\beta}}{\bar c}\hat a\dot\delta^2\right)\\ &= \vec\nabla\left[\dot\delta\vec\nabla\delta + \dot\delta\hat a\,(\hat a\cdot\vec\nabla\delta) - \hat a\frac{f_{\bar\beta}}{\bar c}\dot\delta^2\right].\end{aligned} \tag{C.26}$$

That is, we have $\partial_t\mathcal{H} - \vec\nabla\vec J = 0$ with

$$\vec J = \dot\delta\left(\vec\nabla\delta + \hat a(\vec\nabla\delta\cdot\hat a) - \hat a\frac{f_{\bar\beta}}{\bar c}\dot\delta\right). \tag{C.27}$$

To derive this, we neglected the coupling of δ to the baryonic density perturbation δ_b. Corrections from this coupling can only be non-zero at the position of the perturber, as long as we model the perturber as a point particle. Below, we integrate $\vec J$ over a surface at a finite distance from the perturber. For this purpose, such corrections are irrelevant.

For $\vec V \parallel \hat a$, we have

$$J_x(t,\vec x) = \dot\delta(t,\vec x)\,(\partial_x\phi(t,\vec x)) = (\partial_{\bar t}\bar\phi)(\bar t,\vec{\bar x})(\partial_{\bar x}\bar\phi)(\bar t,\vec{\bar x}), \tag{C.28a}$$

$$J_y(t,\vec x) = \dot\delta\,\partial_y\phi = (\partial_{\bar t}\bar\phi)(\partial_{\bar y}\bar\phi), \tag{C.28b}$$

$$J_z(t,\vec x) = \dot\delta\left(2\partial_z\delta - a_\parallel\frac{f_{\bar\beta}}{\bar c}\dot\delta\right) = \sqrt{2}(\partial_{\bar t}\bar\phi)(\partial_{\bar z}\bar\phi), \tag{C.28c}$$

where, for brevity, we explicitly show the arguments only in the first line. And for $\vec V \perp \hat a$,

$$J_x(t,\vec x) = \dot\delta\left(2\partial_x\delta - \frac{f_{\bar\beta}}{\bar c}\dot\delta\right) = \sqrt{2}(\partial_{\bar x}\bar\phi)(\partial_{\bar t}\bar\phi), \tag{C.29a}$$

$$J_y(t,\vec x) = \dot\delta(\partial_y\delta) = (\partial_{\bar t}\bar\phi)(\partial_{\bar y}\bar\phi), \tag{C.29b}$$

$$J_z(t,\vec x) = \dot\delta(\partial_z\delta) = (\partial_{\bar t}\bar\phi)(\partial_{\bar z}\bar\phi). \tag{C.29c}$$

Note that the $f_{\bar\beta}/\bar c$ terms drop out when written in terms of $\bar\phi$.

To streamline the calculation, we write derivatives with respect to $\bar t$ and $\vec{\bar x}$ of $\bar\phi$ as

$$(\partial_{\bar\mu}\bar\phi) = -\sqrt{\frac{\pi}{2}}\frac{M_{\rm eff}}{(2\pi)^2}\frac{1}{|V_{\rm eff}|}\frac{g_m}{\sqrt{2}M_{\rm Pl}}\int d\bar\omega(i\bar\omega)f_{\bar\mu}e^{i\bar\omega\left(\bar t - \frac{1}{V_{\rm eff}}\bar z\right)}\frac{e^{-i\bar\lambda|\vec{\bar x}_\perp|}}{\sqrt{i\bar\lambda|\vec{\bar x}_\perp|}}, \tag{C.30}$$

with

$$f_{\bar{\mu}} = \begin{cases} 1\,, & \bar{\mu} = \bar{t} \\ -\frac{1}{V_{\rm eff}}\,, & \bar{\mu} = \bar{z} \\ -\frac{f_{\rm crit}}{c_{\rm eff}}\frac{\bar{x}}{|\vec{\bar{x}}_\perp|}\,, & \bar{\mu} = \bar{x} \\ -\frac{f_{\rm crit}}{c_{\rm eff}}\frac{\bar{y}}{|\vec{\bar{x}}_\perp|}\,, & \bar{\mu} = \bar{y} \end{cases}\,, \tag{C.31}$$

where we assumed the $|\bar{\lambda}\vec{\bar{x}}_\perp| \gg 1$ limit. Consider then the integral over the surface of a cylinder oriented along the z axis with radius $a \to \infty$

$$I^g_{\bar{\mu}\bar{\nu}} \equiv a \int d\varphi \int dz\, g(\varphi)\, (\partial_{\bar{\mu}}\bar{\phi})(\bar{t}, \vec{\bar{x}})(\partial_{\bar{\nu}}\bar{\phi})(\bar{t}, \vec{\bar{x}})\,, \tag{C.32}$$

with $x = a\cos\varphi$ and $y = a\sin\varphi$. Here, g is a function of φ. Later, we will consider $g = \cos\varphi$ and $g = \sin\varphi$. We can now calculate (assuming $\bar{z}/z$ to be constant)

$$\begin{aligned} I^g_{\bar{\mu}\bar{\nu}} &= a\frac{1}{64\pi^3}\frac{M^2_{\rm eff}}{V^2_{\rm eff}}\frac{g^2_m}{M^2_{\rm Pl}} \int d\varphi g(\varphi) \int dz \int d\bar{\omega}(i\bar{\omega}) f_{\bar{\mu}} \int d\bar{\omega}'(-i\bar{\omega}') f^{\prime\dagger}_{\bar{\nu}} \\ &\qquad \times e^{i\bar{t}(\bar{\omega}-\bar{\omega}')} e^{-i\bar{z}\frac{1}{V_{\rm eff}}(\bar{\omega}-\bar{\omega}')} \frac{e^{-i|\vec{\bar{x}}_\perp|(\bar{\lambda}-\bar{\lambda}')}}{\sqrt{\bar{\lambda}\bar{\lambda}'}|\vec{\bar{x}}_\perp|} \\ &= a\frac{1}{64\pi^3}\frac{M^2_{\rm eff}}{V^2_{\rm eff}}\frac{g^2_m}{M^2_{\rm Pl}} \left(2\pi \frac{1}{\left|\frac{\bar{z}}{z}\frac{1}{V_{\rm eff}} - a_\parallel \frac{f_{\bar{\beta}}}{2\bar{c}}\right|}\right) \\ &\qquad \times \int d\varphi g(\varphi) \frac{1}{|\vec{\bar{x}}_\perp|} \int d\bar{\omega} \frac{\bar{\omega}^2}{\bar{\lambda}} f_{\bar{\mu}} f^\dagger_{\bar{\nu}} \\ &= \frac{1}{16\pi^2}\frac{M^2_{\rm eff}}{|V_{\rm eff}|}\frac{g^2_m}{M^2_{\rm Pl}} \left(\frac{c_{\rm eff}}{f_{\rm crit}} \frac{1}{|V_{\rm eff}|\left|\frac{\bar{z}}{z}\frac{1}{V_{\rm eff}} - a_\parallel \frac{f_{\bar{\beta}}}{2\bar{c}}\right|}\right) \\ &\qquad \times \int d\varphi g(\varphi) \frac{a}{|\vec{\bar{x}}_\perp|} \int_+ d\bar{\omega}\bar{\omega}\, f_{\bar{\mu}} f^\dagger_{\bar{\nu}}\,. \end{aligned} \tag{C.33}$$

Here, $\int_+$ means we integrate only over positive values. This result holds for both $\vec{V} \parallel \hat{a}$ and $\vec{V} \perp \hat{a}$. For $\vec{V} \perp \hat{a}$, we have $a_\parallel = 0$ in Eq. (C.33). For $\vec{V} \parallel \hat{a}$, we have $a_\parallel = \pm 1$.

The energy loss due to Cherenkov radiation is an integral over such a cylinder surface [135]. We take this cylinder to be infinitely long along the z direction with a large but finite radius a. We consider only the side of the cylinder, since the top and bottom bases do not contribute, as we show below,

$$\dot{E} = \int d\vec{S}\vec{J} = a\int d\varphi \int dz(\cos\varphi J_x + \sin\varphi J_y)\,. \tag{C.34}$$

For $\vec{V} \parallel \hat{a}$, we have $\vec{\bar{x}}_\perp = \vec{x}_\perp$ such that

$$\begin{aligned}\dot{E} &= I_{\bar{t}\bar{x}}^{\cos\varphi} + I_{\bar{t}\bar{y}}^{\sin\varphi} \\ &= -\frac{1}{16\pi^2}\frac{M_{\text{eff}}^2}{|V_{\text{eff}}|}\frac{g_m^2}{M_{\text{Pl}}^2}\frac{1}{|V_{\text{eff}}|\left|\frac{1}{\sqrt{2}}\frac{1}{V_{\text{eff}}} - a_\parallel\frac{f_{\bar{\beta}}}{2\bar{c}}\right|} \\ &\qquad\qquad \times \int d\varphi(\cos^2\varphi + \sin^2\varphi)\int_+ d\bar{\omega}\bar{\omega} \\ &= -\frac{1}{8\pi}\frac{M^2}{V}\frac{g_m^2}{M_{\text{Pl}}^2}\int_+ d\bar{\omega}\bar{\omega}\,.\end{aligned} \tag{C.35}$$

For $\vec{V} \perp \hat{a}$, we have

$$|\vec{\bar{x}}_\perp| = a\sqrt{\frac{1+\sin^2\varphi}{2}}\,. \tag{C.36}$$

This gives

$$\begin{aligned}\dot{E} &= \sqrt{2}I_{\bar{t}\bar{x}}^{\cos\varphi} + I_{\bar{t}\bar{y}}^{\sin\varphi} \\ &= -\frac{1}{16\pi^2}\frac{M_{\text{eff}}^2}{|V_{\text{eff}}|}\frac{g_m^2}{M_{\text{Pl}}^2} \\ &\qquad \times \int d\varphi\sqrt{\frac{2}{1+\sin^2\varphi}}^2\left(\sqrt{2}\frac{1}{\sqrt{2}}\cos^2\varphi + \sin^2\varphi\right)\int_+ d\bar{\omega}\bar{\omega} \\ &= -\frac{1}{\sqrt{2}\,8\pi}\frac{M^2}{V}\frac{g_m^2}{M_{\text{Pl}}^2}\int_+ d\bar{\omega}\bar{\omega}\,.\end{aligned} \tag{C.37}$$

The contributions of the top and bottom cylinder bases are proportional to

$$\int_{|\vec{x}_\perp|<a} d^2\vec{x}_\perp(\partial_{\bar{t}}\bar{\phi})(\partial_{\bar{z}}\bar{\phi})\Big|_{z\to\pm\infty}\,. \tag{C.38}$$

This holds for both $\vec{V} \parallel \hat{a}$ and $\vec{V} \perp \hat{a}$. In position space [135], we have

$$\bar{\phi} \propto \frac{1}{\sqrt{(\bar{z} - V_{\text{eff}}\bar{t})^2 - \frac{f_{\text{crit}}^2}{1-f_{\text{crit}}^2}|\vec{\bar{x}}_\perp|^2}}\,, \tag{C.39}$$

whenever the argument of the square root is positive. Otherwise, $\bar{\phi}$ vanishes,

$$\bar{\phi} = 0 \quad \text{for} \quad (\bar{z} - V_{\text{eff}}\bar{t})^2 - \frac{f_{\text{crit}}^2}{1-f_{\text{crit}}^2}|\vec{\bar{x}}_\perp|^2 < 0\,. \tag{C.40}$$

For both $\vec{V} \parallel \hat{a}$ and $\vec{V} \perp \hat{a}$, the limit $z \to \pm\infty$ at fixed t and $\vec{x}_\perp$ implies $|\bar{z} - V_{\text{eff}}\bar{t}| \to \infty$. Thus, the integrand in Eq. (C.38) vanishes at least as $1/z^4$. Therefore, the cylinder bases do not contribute to $\dot{E}$.

C.4. Cutoffs

The above calculations were without any kind of cutoff. But, as discussed in Sec. 4.2, we need to impose cutoffs. In our QFT calculation in Appendix A, we imposed cutoffs $k_{\min}$ and $k_{\max}$ on the spatial wavevector $\vec{k}$ of the emitted radiation. In our classical computation, $\vec{k}$ corresponds to the prefactor of $i\vec{x}$ in the $e^{i\vec{k}\vec{x}}$ factor in the Fourier transform of δ. We express $\delta(t,\vec{x})$ through $\bar{\phi}(\bar{t},\vec{\bar{x}})$ which depends on the coordinates as $e^{i(\bar{\omega}\bar{t}-\vec{\bar{k}}\vec{\bar{x}})}$. We can write the dependence of this exponential on $\vec{x}$ as $e^{-i\vec{k}\vec{x}}$ for some $\vec{k}$ using the transformation between $\bar{t}$, $\vec{\bar{x}}$ and t, $\vec{x}$. Then, we can try to impose our standard cutoffs on the absolute value of $\vec{k}$. The calculation in the previous section naturally leaves an $\bar{\omega}$ integral in the end. Thus, we will try to express $|\vec{k}|$ through $\bar{\omega}$. Then, our standard cutoffs on $|\vec{k}|$ translate to cutoffs on $\bar{\omega}$.

Consider first $\vec{V} \parallel \hat{a}$. Then,

$$\begin{aligned}\bar{\omega}\bar{t} - \vec{\bar{k}}\vec{\bar{x}} &= \bar{\omega}\bar{t} - \vec{\bar{k}}_\perp\vec{\bar{x}}_\perp - \bar{k}_{\bar{z}}\bar{z} \\ &= \bar{\omega}t - \vec{\bar{k}}_\perp\vec{x}_\perp - z\left(\frac{1}{\sqrt{2}}\bar{k}_{\bar{z}} - a_\parallel\frac{f_{\bar{\beta}}}{2\bar{c}}\bar{\omega}\right).\end{aligned} \tag{C.41}$$

Thus

$$|\vec{k}|^2 = \bar{k}_\perp^2 + \left(\frac{1}{\sqrt{2}}\bar{k}_{\bar{z}} - a_\parallel\frac{f_{\bar{\beta}}}{2\bar{c}}\bar{\omega}\right)^2. \tag{C.42}$$

The source term $\delta_{b,\text{eff}} \propto \delta(\bar{z} - V_{\text{eff}}\bar{t})$ gives in Fourier space $\bar{\omega} = V_{\text{eff}}\bar{k}_{\bar{z}}$. Further, we are interested only in the on-shell contribution from the pole $\bar{\omega} = c_{\text{eff}}\bar{k}$. Thus, using $\bar{k}_\perp^2 = \bar{k}^2 - \bar{k}_{\bar{z}}^2$, we have

$$\begin{aligned}|\vec{k}|^2 &= \bar{\omega}^2\left[\left(\frac{1}{c_{\text{eff}}^2} - \frac{1}{V_{\text{eff}}^2}\right) + \left(\frac{1}{\sqrt{2}}\frac{1}{V_{\text{eff}}} - a_\parallel\frac{f}{2\bar{c}}\right)^2\right] \\ &= \frac{\bar{\omega}^2}{\bar{c}^2}\left(\frac{1}{1-(\bar{c}/V)^2 - 2a_\parallel f_{\bar{\beta}}(\bar{c}/V)}\right)^{-1} \equiv \frac{\bar{\omega}^2}{\bar{c}^2}\frac{1}{\left(f_{\max}^\parallel\right)^2}.\end{aligned} \tag{C.43}$$

This is a relation between k and $\bar{\omega}$. We can use this to impose our standard cutoffs on k, i.e. $k < k_{\max}$ and $k > k_{\min}$, in the $\bar{\omega}$ integrals in the result for $\dot{E}$ found in the previous section. This gives the final result for the energy loss

$$\dot{E} = -\frac{\bar{c}^2}{16\pi}\frac{M^2}{V}\frac{g_m^2}{M_{\text{Pl}}^2}\frac{1}{1-(\bar{c}/V^2) - 2a_\parallel f_{\bar{\beta}}(\bar{c}/V)}(k_{\max}^2 - k_{\min}^2). \tag{C.44}$$

Consider next $\vec{V} \perp \hat{a}$. Then,

$$\begin{aligned}\bar{\omega}\bar{t} - \vec{\bar{k}}\vec{\bar{x}} &= \bar{\omega}\bar{t} - \bar{k}_{\bar{x}}\bar{x} - \bar{k}_{\bar{y}}\bar{y} - \bar{k}_{\bar{z}}\bar{z} \\ &= \bar{\omega}t - x\left(\frac{1}{\sqrt{2}}\bar{k}_{\bar{x}} - \frac{f_{\bar{\beta}}}{2\bar{c}}\bar{\omega}\right) - \bar{k}_{\bar{y}}y - \bar{k}_{\bar{z}}z.\end{aligned} \tag{C.45}$$

That is

$$|\vec{\bar{k}}|^2 = \left(\frac{1}{\sqrt{2}}\bar{k}_{\bar{x}} - \frac{f_{\bar{\beta}}}{2\bar{c}}\bar{\omega}\right)^2 + \bar{k}_{\bar{y}}^2 + \bar{k}_{\bar{z}}^2 \,. \tag{C.46}$$

We can again use the two relations $\bar{\omega} = V_{\text{eff}}\bar{k}_{\bar{z}}$ and $\bar{\omega} = c_{\text{eff}}\bar{k}$. But, in contrast to the $\vec{V} \parallel \hat{a}$ case, this does not uniquely specify $|\vec{\bar{k}}|$ in terms of $\bar{\omega}$. Instead, $|\vec{\bar{k}}|$ depends on both $\bar{\omega}$ and the orientation of $\vec{\bar{k}}$, i.e. the spherical angles $\bar{\theta}$ and $\bar{\varphi}$. Concretely

$$\begin{aligned}
|\vec{\bar{k}}|^2 &= \frac{1}{2}\bar{k}_{\bar{x}}^2 + \bar{k}_{\bar{y}}^2 + \bar{k}_{\bar{z}}^2 + \left(\frac{f_{\bar{\beta}}}{2\bar{c}}\right)^2\bar{\omega}^2 - \frac{1}{\sqrt{2}}\frac{f_{\bar{\beta}}}{\bar{c}}\bar{k}_{\bar{x}}\bar{\omega} \\
&= \frac{1}{2}\bar{k}_{\bar{x}}^2 + (\bar{k}^2 - \bar{k}_{\bar{x}}^2 - \bar{k}_{\bar{z}}^2) + \bar{k}_{\bar{z}}^2 + \left(\frac{f_{\bar{\beta}}}{2\bar{c}}\right)^2\bar{\omega}^2 - \frac{1}{\sqrt{2}}\frac{f_{\bar{\beta}}}{\bar{c}}\bar{k}_{\bar{x}}\bar{\omega} \\
&= \bar{k}^2 - \frac{1}{2}\bar{k}^2 s_{\bar{\theta}}^2 c_{\bar{\varphi}}^2 - \frac{1}{\sqrt{2}}\frac{f_{\bar{\beta}}}{\bar{c}}\bar{k}\bar{\omega}s_{\bar{\theta}}c_{\bar{\varphi}} + \left(\frac{f_{\bar{\beta}}}{2\bar{c}}\right)^2\bar{\omega}^2 \\
&= \bar{k}^2 + \left(\frac{f_{\bar{\beta}}}{2\bar{c}}\right)^2\bar{\omega}^2 - \frac{1}{2}\left(\bar{k}s_{\bar{\theta}}c_{\bar{\varphi}} + \frac{1}{\sqrt{2}}\frac{f_{\bar{\beta}}}{\bar{c}}\bar{\omega}\right)^2 + \frac{1}{4}\frac{f_{\bar{\beta}}^2}{\bar{c}^2}\bar{\omega}^2 \,,
\end{aligned} \tag{C.47}$$

where $s_{\bar{\theta}} = \sin\bar{\theta}$ and $c_{\bar{\varphi}} = \cos\bar{\varphi}$. To get a relation between k and $\bar{\omega}$, we could use the relations $\bar{\omega} = V_{\text{eff}}\bar{k}_{\bar{z}}$ and $\bar{\omega} = c_{\text{eff}}\bar{k}$ to fix $\cos\bar{\theta} = c_{\text{eff}}/V_{\text{eff}}$. But this still leaves $\bar{\varphi}$ undetermined.

To avoid complications from this, we do something simpler. We set $\bar{\varphi}$ to the value that gives the largest value $(|\vec{\bar{k}}|^2)_{\text{largest}}$ of $|\vec{\bar{k}}|$. Using this largest value of $|\vec{\bar{k}}|$ to set the cutoff on $\bar{\omega}$ gives a more conservative cutoff than the cutoffs we used previously. This largest possible value of $|\vec{\bar{k}}|$ is

$$\left(|\vec{\bar{k}}|^2\right)_{\text{largest}} = \bar{k}^2 + \frac{f_{\bar{\beta}}^2}{2}\left(\frac{\bar{\omega}}{\bar{c}}\right)^2 \,. \tag{C.48}$$

With $\bar{\omega} = c_{\text{eff}}\bar{k}$, this gives

$$\begin{aligned}
\left(|\vec{\bar{k}}|^2\right)_{\text{largest}} &= \left(\frac{\bar{\omega}}{\bar{c}}\right)^2\left[\left(\frac{\bar{c}}{c_{\text{eff}}}\right)^2 + \frac{1}{2}f_{\bar{\beta}}^2\right] \\
&= \left(\frac{\bar{\omega}}{\bar{c}}\right)^2(1 + f_{\bar{\beta}}^2) \equiv \left(\frac{\bar{\omega}}{\bar{c}}\right)^2\frac{1}{(f_{\text{max}}^{\perp})^2} \,.
\end{aligned} \tag{C.49}$$

We will use this relation between k and $\bar{\omega}$ to set the upper cutoff in the $\bar{\omega}$ integral. For the lower cutoff, we should in principle use the value of $\bar{\varphi}$ that gives the smallest possible $|\vec{\bar{k}}|^2$. However, our result is dominated by the upper cutoff. Thus, for simplicity, we take the same value of $\bar{\varphi}$ for both the upper and the lower cutoff. Imposing our standard cutoffs on k then gives the final result for the energy loss

$$\dot{E} = -\frac{\bar{c}^2}{\sqrt{2}\,16\pi}\frac{M^2}{V}\frac{g_m^2}{M_{\text{Pl}}^2}\frac{1}{1 + f_{\bar{\beta}}^2}(k_{\text{max}}^2 - k_{\text{min}}^2) \,. \tag{C.50}$$

Above, we have ignored three subtleties regarding cutoffs that we discuss now. First, we have calculated the energy loss through Cherenkov radiation by integrating the energy flux over the surface of an infinitely large cylinder. But we justified our cutoffs by arguing that the perturbation's Lagrangian is valid only between a distance $r_{\min}$ and $r_{\max}$ away from the perturber. So maybe we should use a cylinder with radius and length at most $r_{\max}$ instead of an infinitely large cylinder. Here, we do not do this. Instead, we keep the infinitely large cylinder but use the field $\bar{\phi}$ as calculated with a momentum-space cutoff. One reason is that energy conservation guarantees that the integral over an infinitely large cylinder gives the same result as that over a finite cylinder. Of course, the infinitely large cylinder is unphysical. But it suffices that, mathematically, the result will be the same in both cases. More precisely, this holds only up to small corrections because, with the cutoffs, the perturber is no longer a strictly localized point particle. Instead, the perturber is smeared out over a distance $r_{\min}$ so that some of the Cherenkov radiation is produced outside any finite volume. Corrections from this are negligible as long as $r_{\min} \ll r_{\max}$, which is true in our case. Another reason why we keep the infinitely large cylinder is that this is what we already did implicitly when we employed the standard QFT formalism with momentum-space cutoffs in our standard calculation in Appendix A.

The second subtlety is that, in our calculation of $\bar{\phi}(\bar{t}, \vec{\bar{x}})$, we used the following result from Ref. [135]

$$I \equiv \int d^2\bar{k}_\perp e^{-i\vec{\bar{k}}_\perp \vec{\bar{x}}_\perp} \frac{1}{\bar{\lambda}^2 - \bar{k}_\perp^2} = -2\pi K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)\,. \tag{C.51}$$

This, in principle, includes wavevectors $\bar{k}_\perp$ that are arbitrarily large. Instead we should use a version of this integral with appropriate cuts on $\bar{k}_\perp$. However, the $e^{-i\vec{\bar{k}}_\perp \vec{\bar{x}}_\perp}$ factor already effectively cuts off the integral at

$$\bar{k}_\perp \sim 1/|\vec{\bar{x}}_\perp|\,. \tag{C.52}$$

Thus, our previous calculation is valid as long as $|\vec{\bar{x}}_\perp|$ is larger than $1/\bar{k}_{\perp,\max}$ where $\bar{k}_{\perp,\max}$ is the cutoff we should impose on $\bar{k}_\perp$. For smaller values of $|\vec{\bar{x}}_\perp|$ we should use a version of the integral with cutoffs.

Specifically, we have for the upper cutoff on $\bar{k}_\perp$,

$$\left(\bar{k}^{\parallel}_{\perp,\max}\right)^2 = k_{\max}^2 - \left(\frac{\bar{\omega}}{\bar{c}}\right)^2 \left(\frac{1}{\sqrt{2}}\frac{\bar{c}}{V_{\text{eff}}} - a_\parallel \frac{f_{\bar{\beta}}}{2}\right)^2 , \tag{C.53a}$$

$$\left(\bar{k}^{\perp}_{\perp,\max}\right)^2 = k_{\max}^2 - \left(\frac{\bar{\omega}}{\bar{c}}\right)^2 \left(\left(\frac{\bar{c}}{V_{\text{eff}}}\right)^2 + \frac{1}{2} f_{\bar{\beta}}^2\right) , \tag{C.53b}$$

where $\bar{k}^{\parallel}_{\perp,\max}$ applies for $\vec{V} \parallel \hat{a}$ and $\bar{k}^{\perp}_{\perp,\max}$ applies for $\vec{V} \perp \hat{a}$. This follows from Eq. (C.42) and Eq. (C.48) using $\bar{k}_{\bar{z}} = \bar{\omega}/V_{\text{eff}}$. Both cases can be written

as

$$\begin{aligned}(\bar{k}_{\perp,\max})^2 &= k_{\max}^2 - \bar{\omega}^2 \left(\left(\frac{1}{\bar{c} f_{\max}} \right)^2 - \left(\frac{\bar{\lambda}}{\bar{\omega}} \right)^2 \right) \\ &= k_{\max}^2 \left(1 - \left(\frac{\bar{\omega}}{\bar{\omega}_{\max}} \right)^2 \left(1 - f_{\text{crit}}^2 f_{\max}^2 \left(\frac{\bar{c}}{c_{\text{eff}}} \right)^2 \right) \right) .\end{aligned} \tag{C.54}$$

The smallest possible value of $\bar{k}_{\perp,\max}^2$ is exactly $\bar{\lambda}^2$ evaluated at $\bar{\omega} = \bar{\omega}_{\max}$. Thus, our previous calculation is valid if

$$|\vec{\bar{x}}_\perp| \gtrsim \frac{1}{\bar{\lambda}|_{\bar{\omega}=\bar{\omega}_{\max}}} = r_{\min} \cdot \frac{c_{\text{eff}}}{\bar{c}} \frac{1}{f_{\max} f_{\text{crit}}} . \tag{C.55}$$

The right-hand side is of the same of order of magnitude as $r_{\min}$ except if $c_{\text{eff}}/V_{\text{eff}}$ is close to 1, i.e. for barely supersonic perturbers. In this case, f_{crit} becomes small. But, as we will discuss below, we consider only perturbers that are at most 1% away from the critical velocity. So this is never much larger than $r_{\min}$ for our purposes. Thus, for our energy loss calculation we can keep our previous result.

In principle, there is also a lower cutoff on $\bar{k}_\perp$. However, this cutoff is 0 except for $\bar{\omega}$ close to its lower cutoff. As already mentioned above, our results are dominated by the upper cutoff of $\bar{\omega}$. Thus, we leave out this lower cutoff on $\bar{k}_\perp$ for simplicity.

The third subtlety is regarding stars that are barely supersonic, $V \to V_{\text{crit}}$. One may suspect a logarithmic divergence in $\bar{\phi}$ at $V \to V_{\text{crit}}$, i.e. at $V_{\text{eff}} \to c_{\text{eff}}$. The reason is that $\bar{\lambda}^2 \propto f_{\text{crit}}^2 = 1 - (c_{\text{eff}}/V_{\text{eff}})^2$ and $K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)$ diverges logarithmically for small arguments. Indeed, such a divergence exists in our case and we will restrict ourselves to perturbers at least 1% away from V_{crit} in order to stay away from this divergence.

This logarithmic divergence does not affect the calculated energy loss. The energy loss is calculated from an expansion of $K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)$ for large arguments, while the divergence comes from an expansion of this function for small arguments. However, the logarithmic divergence is relevant when judging whether or not perturbations are small at various locations, for example at $\vec{\bar{x}}_\perp = 0$ where the argument of K_0 is definitely small. This is discussed in more detail in Appendix C.6. In electrodynamics, such logarithms occur as well (see for example Eq. (13.39) in Ref. [135] and the discussion below this equation).

For later reference, we explicitly give the result of the $\bar{k}_\perp$ integral at $|\vec{\bar{x}}_\perp| = 0$

including cutoffs,

$$\begin{aligned}
&\frac{1}{2\pi}\int d^2\bar{k}_\perp e^{-i\vec{\bar{k}}_\perp\vec{\bar{x}}_\perp}\frac{\Theta(\bar{k}_{\perp,\max}-\bar{k}_\perp)}{\bar{\lambda}^2-k_\perp^2}\\
&=-K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)-\int_{\bar{k}_{\perp,\max}}^{\infty}d\bar{k}_\perp\frac{\bar{k}_\perp J_0(\bar{k}_\perp|\vec{\bar{x}}_\perp|)}{\bar{\lambda}^2-\bar{k}_\perp^2}\\
&\approx -K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)-\int_{\bar{k}_{\perp,\max}}^{N|\bar{\lambda}|}d\bar{k}_\perp\frac{\bar{k}_\perp}{\bar{\lambda}^2-\bar{k}_\perp^2}-\int_{N|\bar{\lambda}|}^{\infty}d\bar{k}_\perp\frac{J_0(\bar{k}_\perp|\vec{\bar{x}}_\perp|)}{-\bar{k}_\perp}\\
&\approx \left(\gamma_E+\ln\left(\frac{|\bar{\lambda}\vec{\bar{x}}_\perp|}{2}\right)+i\sigma(\bar{\lambda})\frac{\pi}{2}\right)\\
&\quad-\left(-\frac{1}{2}\left[\ln\left(\bar{k}_\perp^2-\bar{\lambda}^2\right)\right]_{\bar{k}_{\perp,\max}}^{N|\bar{\lambda}|}+\gamma_E+\ln\left(\frac{N|\vec{\bar{x}}_\perp||\bar{\lambda}|}{2}\right)\right)\\
&\approx -\frac{1}{2}\ln\left(\frac{\bar{k}_{\perp,\max}^2}{\bar{\lambda}^2}-1\right)+i\sigma(\bar{\lambda})\frac{\pi}{2}\,,
\end{aligned}\tag{C.56}$$

where we expanded in $|\vec{\bar{x}}_\perp|$, N is a large integer, and we used the fact that $\bar{k}_{\perp,\max}^2$ is always larger than $\bar{\lambda}^2$. If we define $\bar{f}_{\rm crit}\equiv f_{\rm crit}\cdot f_{\max}\cdot(\bar{c}/c_{\rm eff})$, we can write the remaining logarithm as

$$\ln\left(\frac{1-(1-\bar{f}_{\rm crit}^2)(\bar{\omega}/\bar{\omega}_{\max})^2}{\bar{f}_{\rm crit}^2(\bar{\omega}/\bar{\omega}_{\max})^2}-1\right)=\ln\left(\frac{1-(\bar{\omega}/\bar{\omega}_{\max})^2}{(\bar{\omega}/\bar{\omega}_{\max})^2}\right)-\ln\left(\bar{f}_{\rm crit}^2\right)\,.\tag{C.57}$$

C.5. Direction of friction force

Above, we calculated the total energy loss of a perturber due to Cherenkov radiation. But in which direction does this energy loss push the perturber? To answer this question we can calculate the linear momentum loss in different directions. Or, a bit simpler, we can calculate the relative amount of the momentum loss in different directions.

The canonical linear momentum of perturbations is

$$\vec{P}=\frac{\partial\mathcal{L}}{\partial\dot{\delta}}(\vec{\nabla}\delta)=\left(\frac{\dot{\delta}}{\bar{c}^2}+\frac{f_{\bar{\beta}}}{\bar{c}}\hat{a}\vec{\nabla}\delta\right)(\vec{\nabla}\delta)\,.\tag{C.58}$$

Each component P_i satisfies a continuity equation $\partial_t P_i-\partial_j(J_i^P)^j=0$. If we know $(J_i^P)^j$ we can calculate the momentum loss in a way similar to how we calculated the energy loss, i.e. by evaluating a surface integral.

We have, using the δ equation of motion

$$\begin{aligned}\partial_t P_i &= \left(\frac{\ddot{\delta}}{\bar{c}^2} + \frac{f_{\bar{\beta}}}{\bar{c}}\hat{a}\vec{\nabla}\dot{\delta}\right)(\partial_i\delta) + \left(\frac{\dot{\delta}}{\bar{c}^2} + \frac{f_{\bar{\beta}}}{\bar{c}}\hat{a}\vec{\nabla}\delta\right)(\partial_i\dot{\delta}) \\ &= \left(\vec{\nabla}^2\delta + \vec{\nabla}\left((\vec{\nabla}\delta\hat{a})\hat{a}\right) - 2\hat{a}\frac{f_{\bar{\beta}}}{\bar{c}}\vec{\nabla}\dot{\delta} + \frac{f_{\bar{\beta}}}{\bar{c}}\hat{a}\vec{\nabla}\dot{\delta}\right)(\partial_i\delta) \\ &\qquad + \left(\frac{\dot{\delta}}{\bar{c}^2} + \frac{f_{\bar{\beta}}}{\bar{c}}\hat{a}\vec{\nabla}\delta\right)(\partial_i\dot{\delta}) \qquad \text{(C.59)} \\ &= \partial_j\left(\partial_i\delta\partial_j\delta + a_j(a_k\partial_k\delta)(\partial_i\delta) - \frac{f_{\bar{\beta}}}{\bar{c}}a_j\dot{\delta}(\partial_i\delta)\right) \\ &\quad + \partial_i\left(\frac{1}{2}\frac{\dot{\delta}^2}{\bar{c}^2} - \frac{1}{2}(\partial_j\delta)^2 - \frac{1}{2}(a_j\partial_j\delta)^2 + \frac{f_{\bar{\beta}}}{\bar{c}}(a_j\partial_j\delta)\dot{\delta}\right).\end{aligned}$$

Thus

$$\begin{aligned}(J_i^P)^j &= \partial_i\delta\partial_j\delta + a_j(a_k\partial_k\delta)(\partial_i\delta) - \frac{f_{\bar{\beta}}}{\bar{c}}a_j\dot{\delta}(\partial_i\delta) \\ &\quad + \delta_i^j\left(\frac{1}{2}\frac{\dot{\delta}^2}{\bar{c}^2} - \frac{1}{2}(\partial_k\delta)^2 - \frac{1}{2}(a_k\partial_k\delta)^2 + \frac{f_{\bar{\beta}}}{\bar{c}}(a_k\partial_k\delta)\dot{\delta}\right). \qquad \text{(C.60)}\end{aligned}$$

The linear momentum loss of the i-th component through the surface of a cylinder with radius $a \to \infty$ is then

$$\dot{P}_i = a\int dz\int d\varphi\left(\cos\varphi(J_i^P)^x + \sin\varphi(J_i^P)^y\right). \qquad \text{(C.61)}$$

As in the case of the energy loss, we can express $(J_i^P)^j$ as a linear combination of $(\partial_{\bar{\mu}}\bar{\phi})(\partial_{\bar{\nu}}\bar{\phi})$ using the transformation between $\bar{t}$, $\vec{\bar{x}}$ and t, $\vec{x}$. Then, $\dot{P}_i$ is just a linear combination of the integrals $I^{\cos\varphi}_{\bar{\mu}\bar{\nu}}$ and $I^{\sin\varphi}_{\bar{\mu}\bar{\nu}}$ defined in Appendix C.3. Many of these $I^g_{\bar{\mu}\bar{\nu}}$ terms vanish after doing the φ integral. The only non-zero integrals for $g = \cos\varphi$ are

$$I^{\cos\varphi}_{\bar{x}\bar{t}}\,, \quad I^{\cos\varphi}_{\bar{x}\bar{z}}\,, \qquad \text{(C.62)}$$

and the only non-zero integrals for $g = \sin\varphi$ are

$$I^{\sin\varphi}_{\bar{y}\bar{t}}\,, \quad I^{\sin\varphi}_{\bar{x}\bar{z}}\,. \qquad \text{(C.63)}$$

This can be seen from the fact that the φ integral is of the form

$$\int d\varphi\,\tilde{g}(\sin^2\varphi)\,g(\varphi)\,f_{\bar{\mu}}f_{\bar{\nu}}\,, \qquad \text{(C.64)}$$

where $\tilde{g}$ is some function of $\sin^2\varphi$ and then using the specific form of $f_{\bar{\mu}}$ and $f_{\bar{\nu}}$ given by Eq. (C.31).

For $\vec{V} \parallel \hat{a}$, we have

$$(J_x^P)^x = (\partial_x\delta)^2 + \frac{1}{2}\frac{\dot{\delta}^2}{\bar{c}^2} - \frac{1}{2}(\vec{\nabla}\delta)^2 - \frac{1}{2}(\partial_z\delta)^2 + \frac{f_{\bar{\beta}}}{\bar{c}}(a_\parallel\partial_z\delta)\dot{\delta}\,. \tag{C.65}$$

When written in terms of $\bar{\phi}$, this contains no terms proportional to $(\partial_{\bar{x}}\bar{\phi})(\partial_{\bar{t}}\bar{\phi})$ or $(\partial_{\bar{x}}\bar{\phi})(\partial_{\bar{z}}\bar{\phi})$. The same is true for $(J_x^P)^y$. Thus, $\dot{P}_x = 0$. An analogous argument shows that $\dot{P}_y = 0$. Thus, the friction force points in the z direction.

For $\vec{V} \perp \hat{a}$, we have for J_x^P

$$\begin{aligned}(J_x^P)^x &= (\partial_x\delta)^2 + (\partial_x\delta)^2 - \frac{f_{\bar{\beta}}}{\bar{c}}\dot{\delta}(\partial_x\delta) + \frac{1}{2}\frac{\dot{\delta}}{\bar{c}^2} - \frac{1}{2}(\vec{\nabla}\delta)^2 \\ &\qquad - \frac{1}{2}(\partial_x\delta)^2 + \frac{f_{\bar{\beta}}}{\bar{c}}(\partial_x\delta)\dot{\delta} \\ &= (\partial_x\delta)^2 + \cdots = \frac{1}{\sqrt{2}}\frac{f_{\bar{\beta}}}{\bar{c}}(\partial_{\bar{t}}\bar{\phi})(\partial_{\bar{x}})\bar{\phi} + \dots\,,\end{aligned} \tag{C.66}$$

where "…" denotes terms that cannot contribute to $\dot{P}_x$. Similarly,

$$\begin{aligned}(J_x^P)^y &= (\partial_x\delta)(\partial_y\delta) = \left(\frac{1}{\sqrt{2}}(\partial_{\bar{x}}\bar{\phi}) + \frac{f_{\bar{\beta}}}{2\bar{c}}(\partial_{\bar{t}}\bar{\phi})\right)(\partial_{\bar{y}}\bar{\phi}) \\ &= \frac{f_{\bar{\beta}}}{2\bar{c}}(\partial_{\bar{t}}\bar{\phi})(\partial_{\bar{y}}\bar{\phi}) + \dots\,.\end{aligned} \tag{C.67}$$

For (J_y^P), we find

$$\begin{aligned}(J_y^P)^x &= (\partial_y\delta)(\partial_x\delta) + (\partial_x\delta)(\partial_y\delta) - \frac{f_{\bar{\beta}}}{\bar{c}}\dot{\delta}(\partial_y\delta) \\ &= (\partial_y\delta)\left(2\partial_x\delta - \frac{f_{\bar{\beta}}}{\bar{c}}\dot{\delta}\right) \\ &= (\partial_{\bar{y}}\bar{\phi})\left(2\frac{1}{\sqrt{2}}(\partial_{\bar{x}}\bar{\phi}) + 2\frac{f_{\bar{\beta}}}{2\bar{c}}(\partial_{\bar{t}}\bar{\phi}) - \frac{f_{\bar{\beta}}}{\bar{c}}(\partial_{\bar{t}}\bar{\phi})\right) = 0 + \dots\,,\end{aligned} \tag{C.68}$$

and

$$\begin{aligned}(J_y^P)^y &= (\partial_y\delta)^2 + \frac{1}{2}\frac{\dot{\delta}^2}{\bar{c}^2} - \frac{1}{2}(\vec{\nabla}\delta)^2 - \frac{1}{2}(\partial_x\delta)^2 + \frac{f_{\bar{\beta}}}{\bar{c}}(\partial_x\delta)\dot{\delta} \\ &= (\partial_x\delta)\left(-(\partial_x\delta) + \frac{f_{\bar{\beta}}}{\bar{c}}\dot{\delta}\right) \\ &= \left(\frac{1}{\sqrt{2}}(\partial_{\bar{x}}\bar{\phi}) + \frac{f_{\bar{\beta}}}{2\bar{c}}(\partial_{\bar{t}}\bar{\phi})\right)\left(-\frac{1}{\sqrt{2}}(\partial_{\bar{x}}\bar{\phi}) + \frac{f_{\bar{\beta}}}{2\bar{c}}(\partial_{\bar{t}}\bar{\phi})\right) = 0 + \dots\,.\end{aligned} \tag{C.69}$$

Thus, $\dot{P}_y = 0$, but $\dot{P}_x \neq 0$. That is, the direction of the friction force lies in the x-z plane, i.e. the plane spanned by $\vec{V}$ and $\hat{a}$.

To find the direction within this plane, we can compare $\dot{P}_x$ to $\dot{P}_z$. We have

$$\begin{aligned}(J_z^P)^x &= (\partial_z\delta)(\partial_x\delta) + (\partial_x\delta)(\partial_z\delta) - \frac{f_{\bar{\beta}}}{\bar{c}}\dot{\delta}(\partial_z\delta) \\ &= 2(\partial_z\delta)(\partial_x\delta) + \cdots = \sqrt{2}(\partial_{\bar{z}}\bar{\phi})(\partial_{\bar{x}}\bar{\phi}) + \dots\,,\end{aligned} \tag{C.70}$$

and

$$(J_z^P)^y = (\partial_z\delta)(\partial_y\delta) = (\partial_{\bar{z}}\bar{\phi})(\partial_{\bar{y}}\bar{\delta})\,. \tag{C.71}$$

Using $(\partial_{\bar{z}}\bar{\phi}) = (-1/V_{\text{eff}})(\partial_{\bar{t}}\bar{\phi})$, we then have

$$\frac{\dot{P}_x}{\dot{P}_z} = -V_{\text{eff}}\frac{f_{\bar{\beta}}}{\bar{c}}\frac{\frac{1}{\sqrt{2}}I_{\bar{t}\bar{x}}^{\cos\varphi} + \frac{1}{2}I_{\bar{t}\bar{y}}^{\sin\varphi}}{\sqrt{2}I_{\bar{t}\bar{x}}^{\cos\varphi} + I_{\bar{t}\bar{y}}^{\sin\varphi}}\,. \tag{C.72}$$

The relevant φ integral in $I_{\bar{t}\bar{x}}^{\cos\varphi}$ is

$$\begin{aligned}\int d\varphi\frac{a\cos\varphi}{|\vec{\bar{x}}_\perp|}f_{\bar{x}} &= \int d\varphi\frac{a\cos\varphi}{|\vec{\bar{x}}_\perp|}\left(-\frac{\bar{x}}{|\vec{\bar{x}}_\perp|}\right)\\ &= -\sqrt{2}\int d\varphi\frac{\cos^2\varphi}{1+\sin^2\varphi} = 2\pi(\sqrt{2}-2)\,.\end{aligned} \tag{C.73}$$

The corresponding φ integral in $I_{\bar{t}\bar{y}}^{\sin\varphi}$ is

$$\begin{aligned}\int d\varphi\frac{a\sin\varphi}{|\vec{\bar{x}}_\perp|}f_{\bar{y}} &= \int d\varphi\frac{a\sin\varphi}{|\vec{\bar{y}}_\perp|}\left(-\frac{\bar{y}}{|\vec{\bar{x}}_\perp|}\right)\\ &= -2\int d\varphi\frac{\sin^2\varphi}{1+\sin^2\varphi} = 2\pi\cdot 2(\sqrt{2}-2)\,.\end{aligned} \tag{C.74}$$

Thus,

$$\begin{aligned}\frac{\dot{P}_x}{\dot{P}_z} &= -\frac{f_{\bar{\beta}}V}{\bar{c}}\frac{\frac{1}{\sqrt{2}}+\frac{1}{2}\cdot 2}{\sqrt{2}+2}\\ &= -\frac{1}{2}\frac{f_{\bar{\beta}}V}{\bar{c}}\,,\end{aligned} \tag{C.75}$$

where we used that $V_{\text{eff}} = V$ for $\vec{V}\perp\hat{a}$. Usually, a friction force points into the negative $\vec{V}$ direction, i.e. in the negative z direction. Our result shows that there is also a component into the positive x direction in our case. For standard SFDM, $\hat{a} = \hat{e}_x$ points into the opposite direction of the background MOND force. Thus, this additional component of the friction force pushes the perturber away from the galactic center.

C.6. When are perturbations small

Let us now check explicitly when perturbations are smaller than the background fields. This determines when we can expand in them so that the perturbations' Lagrangian we derived above is valid. In terms of our prototype calculation in Sec. 4.2.4, this determines the factor f_p. The standard SFDM Lagrangian has a coupling to matter that is linear in the phonon field θ. Thus, the coupling term can always be expanded, no matter how large the

perturbations are. The kinetic term is more complicated. In the MOND limit, it is a function of

$$X - \bar{\beta}Y = (1-\bar{\beta})\dot{\delta}_\theta + \frac{(\vec{\nabla}\theta_0 + \vec{\nabla}\delta_\theta)^2}{2m}\,, \tag{C.76}$$

where $\theta = \theta_0 + \delta_\theta$ is the full phonon field with background field θ_0 and perturbation δ_θ. The kinetic term also depends on X without the $-\bar{\beta}Y$ term. This has the same form as Eq. (C.76) but with 1 instead of $1-\bar{\beta}$. We can expand in δ_θ as long as $\vec{\nabla}\delta_\theta$ is smaller than $\vec{\nabla}\theta_0$ and both $(1-\bar{\beta})\dot{\delta}_\theta$ and $\dot{\delta}_\theta$ are smaller than $(\vec{\nabla}\theta_0)^2/2m$.

We can estimate in the MOND limit

$$|\vec{\nabla}\theta_0| \approx a_\theta/\lambda \approx \sqrt{a_0 a_b^{\mathrm{gal}}}/\lambda = \bar{\alpha}^2\Lambda\sqrt{\frac{a_b^{\mathrm{gal}}}{a_0}}\,, \tag{C.77}$$

where $\lambda = \bar{\alpha}\Lambda/M_{\mathrm{Pl}}$. The variables δ and δ_θ are related by the rescaling $\delta^2 = 2\Lambda|\vec{\nabla}\theta_0|\delta_\theta^2$, see Appendix B. This gives

$$\delta_\theta = \frac{1}{\sqrt{2\Lambda|\vec{\nabla}\theta_0|}}\delta \approx \frac{1}{\sqrt{2}\bar{\alpha}\Lambda}\left(\frac{a_0}{a_b^{\mathrm{gal}}}\right)^{1/4}\delta\,. \tag{C.78}$$

Our goal is to check if δ_θ is sufficiently small at a distance $r_{\min}$ away from the perturber, see Sec. 4.2.4. For simplicity, here we check explicitly at a few specific points: At a distance $r_{\min}$ purely in the direction of the perturber's trajectory, purely in the perpendicular direction, and on the shockwave front. That is, we consider $(z-Vt)^2+|\vec{x}_\perp|^2 = r_{\min}^2$ together with one of the following conditions

$$|\vec{x}_\perp| = 0\,, \tag{C.79a}$$

$$z - Vt = 0\,, \tag{C.79b}$$

$$\bar{t} - \bar{z}/V_{\mathrm{eff}} = |\vec{x}_\perp|\frac{f_{\mathrm{crit}}}{c_{\mathrm{eff}}}\,. \tag{C.79c}$$

Without our cutoffs, we could skip some of these cases. E.g. the field at $|\vec{x}_\perp| = 0$ is exactly zero for $\bar{z} - V_{\mathrm{eff}}\bar{t} < 0$, since we used the retarded propagator. However, with our cutoffs, we effectively impose a finite spatial resolution and the shockwave front cannot be precisely resolved, see Appendix C.4. Thus, the field can be nonzero even outside the shockwave front.

To check the smallness of perturbations, we will again express δ in terms of $\bar{\phi}$. For $\vec{V} \parallel \hat{a}$ we have

$$\bar{t} - \frac{\bar{z}}{V_{\mathrm{eff}}} = t - \frac{z}{V}\,, \quad |\vec{\bar{x}}_\perp| = |\vec{x}_\perp|\,. \tag{C.80}$$

And for $\vec{V} \perp \hat{a}$,

$$\bar{t} - \frac{\bar{z}}{V_{\mathrm{eff}}} = t - \frac{z}{V} + \frac{f_{\bar{\beta}}}{2\bar{c}}|\vec{x}_\perp|c_\varphi\,, \quad |\vec{\bar{x}}_\perp| = |\vec{x}_\perp|\left(1 - \frac{1}{2}c_\varphi^2\right)\,, \tag{C.81}$$

where φ is the azimuthal angle in the $\vec{x}_\perp$ plane. Including our cutoffs, the coordinate-dependence of the field $\bar{\phi}$ is given by the integral

$$\bar{I} = 2\Re\left[\frac{1}{2\pi}\int_{\bar{\omega}_{\min}}^{\bar{\omega}_{\max}} d\bar{\omega} e^{i\bar{\omega}\left(\bar{t}-\frac{\bar{z}}{V_{\text{eff}}}\right)} \int d^2\bar{k}_\perp e^{-i\vec{\bar{k}}_\perp \vec{\bar{x}}_\perp} \frac{\Theta(\bar{k}_{\perp,\max}-\bar{k}_\perp)}{\bar{\lambda}^2-\bar{k}_\perp^2}\right], \quad \text{(C.82)}$$

where Θ denotes the Heaviside theta function. For large $|\vec{\bar{x}}_\perp|$, the $\bar{k}_{\perp,\max}$ cutoff can be neglected, as discussed in Appendix C.4. For $|\vec{\bar{x}}_\perp| = 0$, it is required. For $|\vec{\bar{x}}_\perp| \sim r_{\min}$, it may or may not be necessary depending on the precise numerics. As a conservative estimate, we impose the cutoff for $|\vec{\bar{x}}_\perp| = 0$, but leave it out at $|\vec{\bar{x}}_\perp| \sim r_{\min}$. The $\bar{k}_\perp$ integral then gives the standard $-2\pi K_0(i\bar{\lambda}|\vec{\bar{x}}_\perp|)$ for $|\vec{\bar{x}}_\perp| \sim r_{\min}$. For $|\vec{\bar{x}}_\perp| = 0$, we can use Eq. (C.56). That is, we have

$$\bar{I} = 2\Re\left[\int_{\bar{\omega}_{\min}}^{\bar{\omega}_{\max}} d\bar{\omega} e^{i\bar{\omega}\left(\bar{t}-\frac{\bar{z}}{V_{\text{eff}}}\right)} \times \begin{cases} -K_0(i(\bar{\lambda}/\bar{\omega})\bar{\omega}|\vec{\bar{x}}_\perp|)\,, & |\vec{\bar{x}}_\perp| \sim r_{\min}\,, \\ -\frac{1}{2}\ln\left(\frac{\bar{k}_{\perp,\max}^2}{(\bar{\lambda}/\bar{\omega})^2\bar{\omega}^2}-1\right)+i\frac{\pi}{2}\,, & |\vec{\bar{x}}_\perp| = 0 \end{cases}\right]. \quad \text{(C.83)}$$

At the shockwave front, Eq. (C.79c), both $|\vec{\bar{x}}_\perp| \sim r_{\min}$ and $|\vec{\bar{x}}_\perp| \ll r_{\min}$ are possible, depending on whether or not V is close to V_{crit}. For V close to V_{crit}, we have $|\vec{\bar{x}}_\perp| \sim r_{\min}$. For $V \gg V_{\text{crit}}$, we have $|\vec{\bar{x}}_\perp| \ll r_{\min}$. Thus, we should leave out the cutoff only for $V \sim V_{\text{crit}}$. Here, we leave out the cutoff in both cases as a conservative estimate.

Below, we are interested in derivatives of the integral $\bar{I}$. After a rescaling

$$\bar{\omega} = \bar{\omega}_{\max}\hat{\omega} = \bar{c} f_{\max} k_{\max}\hat{\omega} = \frac{\bar{c} f_{\max}}{r_{\min}}\hat{\omega}\,, \quad \text{(C.84)}$$

we have

$$\begin{aligned}\partial_{\bar{\mu}}\bar{I} &= \hat{f}_{\bar{\mu}}\left(\frac{\bar{c} f_{\max}}{r_{\min}}\right)^2 2\Re\left[\int_{\frac{\bar{\omega}_{\min}}{\bar{\omega}_{\max}}}^{1} d\hat{\omega}(i\hat{\omega}) e^{i\hat{\omega} f_{\text{exp}}} \times \begin{cases} -K_{\hat{n}}(i\hat{\omega} f_K)\,, & |\vec{\bar{x}}_\perp| \sim r_{\min}\,, \\ -\frac{1}{2}\ln\left(\frac{1}{f_K'^2\hat{\omega}^2}-1\right)+i\frac{\pi}{2}\,, & |\vec{\bar{x}}_\perp| = 0 \end{cases}\right] \\ &\equiv \hat{f}_{\bar{\mu}}\left(\frac{\bar{c} f_{\max}}{r_{\min}}\right)^2 \hat{I}_{\bar{\mu}}\,,\end{aligned} \quad \text{(C.85)}$$

where

$$f_{\text{exp}} = f_{\max}\left(\frac{\bar{c}\bar{t}}{r_{\min}} - \frac{\bar{c}}{V_{\text{eff}}}\frac{\bar{z}}{r_{\min}}\right), \quad \text{(C.86a)}$$

$$f_K = f_{\max}\frac{\bar{c}\bar{\lambda}}{\bar{\omega}}\frac{|\vec{\bar{x}}_\perp|}{r_{\min}}\,, \quad \text{(C.86b)}$$

$$f_K' = f_{\max}\frac{\bar{c}\bar{\lambda}}{\bar{\omega}}\frac{k_{\max}}{\bar{k}_{\perp,\max}}\,, \quad \text{(C.86c)}$$

and

$$\hat{f}_{\bar{\mu}} = \begin{cases} 1\,, & \bar{\mu} = \bar{t}\,, \\ -\frac{1}{V_{\rm eff}}\,, & \bar{\mu} = \bar{z}\,, \\ -\frac{\bar{\mu}}{|\vec{\bar{x}}_\perp|}\frac{\bar{\lambda}}{\bar{\omega}}\,, & \bar{\mu} = \bar{x}, \bar{y}\,, |\vec{\bar{x}}_\perp| \sim r_{\rm min}\,, \\ 0\,, & \bar{\mu} = \bar{x}, \bar{y}\,, |\vec{\bar{x}}_\perp| = 0\,, \end{cases} \tag{C.87}$$

and

$$\hat{n} = \begin{cases} 0\,, & \bar{\mu} = \bar{t}, \bar{z}\,, \\ 1\,, & \bar{\mu} = \bar{x}, \bar{y}\,. \end{cases} \tag{C.88}$$

Note that f'_K depends on $\hat{\omega}$ through $\bar{k}_{\perp,\rm max}$. In contrast, $f_{\rm exp}$ and f_K are independent of $\hat{\omega}$.

We can now calculate various ratios of $\partial_{\bar{\mu}}\bar{\phi}$ and $\vec{\nabla}\theta_0$. For example,

$$\begin{aligned} \bar{r}_{\bar{\mu}} &\equiv \frac{\frac{1}{\sqrt{2\Lambda|\vec{\nabla}\theta_0|}}\partial_{\bar{\mu}}\bar{\phi}}{|\vec{\nabla}\theta_0|} = \frac{1}{\sqrt{2}a_0 M_{\rm Pl}}\left(\frac{a_0}{a_b^{\rm gal}}\right)^{3/4}\partial_{\bar{\mu}}\bar{\phi} \\ &= \frac{1}{\sqrt{2}a_0 M_{\rm Pl}}\left(\frac{a_0}{a_b^{\rm gal}}\right)^{3/4}\frac{M_{\rm eff}}{(2\pi)^2}\frac{1}{|V_{\rm eff}|}\frac{g_m}{\sqrt{2}M_{\rm Pl}}\hat{f}_{\bar{\mu}}\left(\frac{\bar{c}f_{\rm max}}{r_{\rm min}}\right)^2\hat{I}_{\bar{\mu}}\,. \end{aligned} \tag{C.89}$$

We can simplify this using

$$g_m = \left(\frac{a_0}{a_b^{\rm gal}}\right)^{1/4}\,, \quad r_{\rm min}^{-2} = f_p^2\frac{8\pi M_{\rm Pl}^2 a_0}{M}\frac{a_b^{\rm gal}}{a_0}\,. \tag{C.90}$$

We find

$$\bar{r}_{\bar{\mu}} = \frac{f_p^2 f_{\rm max}^2}{\pi}\frac{\hat{f}_{\bar{\mu}}\bar{c}^2}{|V_{\rm eff}|}\frac{M_{\rm eff}}{M}\hat{I}_{\bar{\mu}}\,. \tag{C.91}$$

We also need

$$\bar{r}'_{\bar{\mu}} \equiv \frac{\frac{1}{\sqrt{2\Lambda|\vec{\nabla}\theta_0|}}\partial_{\bar{\mu}}\bar{\phi}}{|\vec{\nabla}\theta_0|^2/(2m)} = \frac{\bar{r}_{\bar{\mu}}}{\bar{c}}\frac{2m\bar{c}}{|\vec{\nabla}\theta_0|}\,. \tag{C.92}$$

We can simplify this using

$$\bar{c} = 3\bar{f}_{\bar{\beta}}\sqrt{\frac{a_b^{\rm gal}}{a_0}}\frac{\bar{\alpha}^2\Lambda}{m}\,. \tag{C.93}$$

We find

$$\bar{r}'_{\bar{\mu}} = 6\bar{f}_{\bar{\beta}}\frac{\bar{r}_{\bar{\mu}}}{\bar{c}}\,. \tag{C.94}$$

As discussed above, we are interested in the following ratios that tell us whether or not we can expand in the perturbation δ_θ,

$$r_t \equiv \frac{\max(|1-\bar{\beta}|,1)\dot{\delta}_\theta}{(\vec{\nabla}\theta_0)^2/(2m)}\,, \quad r_i \equiv \frac{\partial_i \delta_\theta}{|\vec{\nabla}\theta_0|}\,, \tag{C.95}$$

where $i = x, y, z$. We can express these ratios in terms of the ratios $\bar{r}_{\bar{\mu}}$ and $\bar{r}'_{\bar{\mu}}$ by expressing δ_θ through $\bar{\phi}$. We find

$$r_t = \max(|1-\bar{\beta}|,1)\bar{r}'_{\bar{t}}\,, \quad r_y = \bar{r}_{\bar{y}}\,. \tag{C.96}$$

Further, for $\vec{V} \parallel \hat{a}$

$$r_x = \bar{r}_{\bar{x}}\,, \quad r_z = \frac{1}{\sqrt{2}}\bar{r}_{\bar{z}} + a_\parallel \frac{f_{\bar{\beta}}}{2\bar{c}}\bar{r}_{\bar{t}}\,. \tag{C.97}$$

And for $\vec{V} \perp \hat{a}$

$$r_x = \frac{1}{\sqrt{2}}\bar{r}_{\bar{x}} + \frac{f_{\bar{\beta}}}{2\bar{c}}\bar{r}_{\bar{t}}\,, \quad r_z = \bar{r}_{\bar{z}}\,. \tag{C.98}$$

These should all be smaller than 1 in magnitude when evaluated at a distance $r_{\min}$ from the perturber. Otherwise, we must adjust our choice of $r_{\min}$, i.e. the factor f_p.

It can easily be seen that $\bar{r}_{\bar{\vec{x}}}$, $\bar{r}'_{\bar{t}}$, and $(f_{\bar{\beta}}/2\bar{c})\bar{r}_{\bar{t}}$ cannot be much larger than 1 for supersonic velocities $V_{\text{eff}} > c_{\text{eff}}$, if the integral $\hat{I}_{\bar{\mu}}$ is not much larger than 1. Indeed, this integral cannot be much larger than 1 at a distance $r_{\min}$ from the perturber. Thus, at a distance $r_{\min}$ from the perturber, the background field and the perturbation are of the same order of magnitude. This roughly confirms our choice of $r_{\min}$ from Sec. 4.2.4. As discussed in Appendix C.4, one caveat to this is that our integral diverges logarithmically for $V \to V_{\text{crit}}$ at $|\vec{\bar{x}}_\perp| = 0$, corresponding to the fact that $f'_K \to 0$ for $V \to V_{\text{crit}}$. Thus, we restrict ourselves to velocities at least 1% above the critical velocity, as also discussed above. There is, in principle, a similar divergence also for $|\vec{\bar{x}}_\perp| \sim r_{\min}$ from the integral within $\hat{I}_{\bar{\mu}}$, corresponding to the fact that $f_K \to 0$ for $V \to V_{\text{crit}}$. But it affects the imaginary part, while only the real part enters our result.

We can make this more precise by numerically evaluating these ratios in Mathematica [103]. We find that they are smaller than 1 for many cases. In these cases, we can safely keep $r_{\min}$ as it is, i.e. set $f_p = 1$. In some cases, r_t/f_p^2 is larger than 1. In particular, this can happen for $\vec{V} \parallel \hat{a}$ with $a_\parallel = +1$ and at velocities that are only slightly larger than the critical velocity. The ratio is larger for smaller values of $\bar{\beta}$. We show the two worst cases in Fig. C.1. These occur for $\bar{\beta} = 3/2$. At $V = 1.01 V_{\text{crit}}$, we have $|r_t| \approx 2.3 f_p^2$ for $|\vec{\bar{x}}_\perp| = r_{\min}$ and $z - Vt = 0$. Similarly, at the same V, we have $|r_t| \approx 2.8 f_p^2$ for $z - Vt = r_{\min}$ and $|\vec{\bar{x}}_\perp| = 0$. In this case, we may need to make f_p smaller by a factor of about $1/\sqrt{2.8}$.

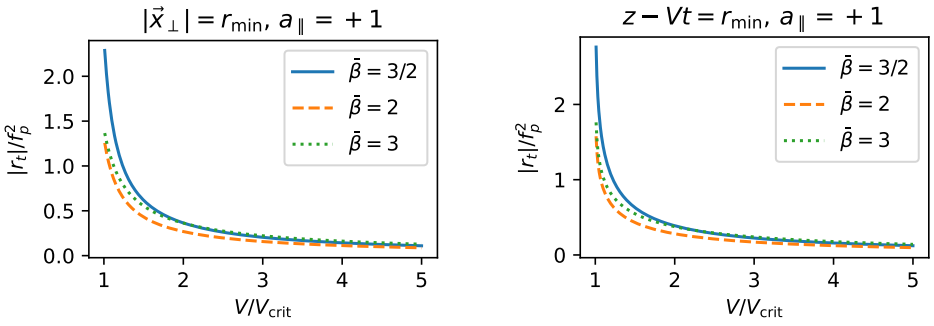

Figure C.1.: The absolute value of the ratio $r_t = \frac{\dot{\delta}_\theta \max(|1-\bar{\beta}|,1)}{(\vec{\nabla}\theta_0)^2/2m}$ as a function of the perturber's velocity V relative to the critical velocity $V_{\rm crit}$. The ratio r_t determines whether or not the perturbation's time derivative is smaller than the relevant combination of background fields. Left: For $\vec{V} \parallel \hat{a}$ with $a_\parallel = +1$ at $|\vec{x}_\perp| = r_{\min}$ and $z - Vt = 0$. Right: For $\vec{V} \parallel \hat{a}$ at $|\vec{x}_\perp| = 0$ and $z - Vt = r_{\min}$. These are the two worst cases, i.e. the cases with the largest $|r_t|/f_p^2$.

D. Two-field SFDM perturbations

We consider perturbations on top of a background galaxy in equilibrium in two-field SFDM. Ref. [73] derived the second-order Lagrangian for such perturbations. With the definition $\lambda \equiv \bar{\alpha}\Lambda/M_{\rm Pl}$ and including the $-\lambda\delta_+\delta_b$ coupling term,

$$\begin{aligned}\mathcal{L} =& + \frac{1}{2}\left(\dot{\delta}_\rho^2 - (\vec{\nabla}\delta_\rho)^2\right) - (\mu_0^2 g^{00} - m^2)\delta_\rho^2 \\ &+ \left(f_0'\dot{\delta}_+^2 - (f_0' - 2f_0''\gamma^2|\vec{\nabla}\theta_+|^2)(\vec{\nabla}\delta_+)^2\right) \\ &+ \left((f_0' + 2f_0''\mu_0^2 + A)\dot{\delta}_-^2 - (f_0' + A)(\vec{\nabla}\delta_-)^2\right) \\ &+ \left(-4f_0''\mu_0|\vec{\nabla}\theta_+^0|\gamma\right)|\vec{\nabla}\delta_+|\dot{\delta}_- \\ &+ 2\mu_0\rho_0\delta_\rho\dot{\delta}_- \\ &- \lambda\delta_+\delta_b\,.\end{aligned} \tag{D.1}$$

Here, δ_ρ is the perturbation on top of the background solution ρ_0 of the field ρ_- and $\delta_\pm$ are the perturbations on top of the background solutions $\theta_\pm^0$ of the fields $\theta_\pm$. Further, μ_0 is the background chemical potential and $A = \rho_0^2/2$.

D.1. Equations of motion

The δ_ρ equation of motion in Fourier space gives

$$\delta_\rho = -\frac{2i\omega\mu_0\rho_0}{\omega^2 - \vec{k}^2 - 2(\mu^2 g^{00} - m^2)}\delta_-\,. \tag{D.2}$$

Plugging this into the δ_- equation, we find

$$\begin{aligned}0 =& (\delta_- e^{ikx} + \delta_-^* e^{-ikx})\left[\left(2(f_0' + 2f_0''\mu^2 + A)\omega^2 - 2(f_0' + A)\vec{k}^2\right)\right. \\ &\left. - \frac{(2\mu_0\rho_0\omega)^2}{\omega^2 - \vec{k}^2 - 2(\mu_0^2 g^{00} - m^2)}\right] \\ &+ \left(4f_0''\mu_0|\vec{\nabla}\theta_+^0|\gamma\right)|\vec{k}|\omega(\delta_+ e^{ikx} + \delta_+^* e^{-ikx})\,.\end{aligned} \tag{D.3}$$

The δ_+ equation reads

$$\begin{aligned}0 =& \left(2f_0'\omega^2 - 2(f_0' - 2f_0''\gamma^2|\vec{\nabla}\theta_+^0|^2)\vec{k}^2\right)(\delta_+ e^{ikx} + \delta_+^* e^{-ikx}) \\ &+ \left(4f_0''\mu_0|\vec{\nabla}\theta_+^0|\gamma\right)|\vec{k}|\omega(\delta_- e^{ikx} + \delta_-^* e^{-ikx}) \\ &+ \lambda(\delta_b e^{ikx} + cc)\,.\end{aligned} \tag{D.4}$$

Solving the δ_+ equation for δ_+ gives

$$\delta_+ = -\frac{\delta_- 4 f_0'' \mu_0 |\vec{\nabla}\theta_+^0| \gamma |\vec{k}| \omega + \lambda \delta_b}{2 f_0' \omega^2 - 2(f_0' - 2 f_0'' \gamma^2 |\vec{\nabla}\theta_+^0|^2)\vec{k}^2} \,. \tag{D.5}$$

Plugging this result into the δ_- equation and dividing by A yields

$$\begin{aligned} 0 = \Bigg[& \left(2 \left(\frac{f_0' + 2 f_0'' \mu_0^2}{A} + 1 \right) \omega^2 - 2 \left(\frac{f_0'}{A} + 1 \right) \vec{k}^2 \right) \\ & - \frac{8 \mu_0^2 \omega^2}{\omega^2 - \vec{k}^2 - 2(\mu_0^2 g^{00} - m^2)} \\ & - \frac{2 f_0'' \mu_0^2}{A} \frac{4 \cdot 2 f_0'' |\vec{\nabla}\theta_+^0|^2 \gamma^2 |\vec{k}|^2 \omega^2}{2 f_0' \omega^2 - 2(f_0' - 2 f_0'' \gamma^2 |\vec{\nabla}\theta_+^0|^2)\vec{k}^2} \Bigg] (\delta_- e^{ikx} + cc) \\ & - \frac{1}{A} \frac{4 \mu_0 f_0'' |\vec{\nabla}\theta_+^0| \gamma |\vec{k}| \omega}{2 f_0' \omega^2 - 2(f_0' - 2 f_0'' \gamma^2 |\vec{\nabla}\theta_+^0|^2)\vec{k}^2} \lambda (\delta_b e^{ikx} + cc) \,. \end{aligned} \tag{D.6}$$

Here,

$$A = \frac{\rho_0^2}{2} = \frac{m^2}{\lambda_4} \frac{\hat{\mu}_0}{m} \,, \tag{D.7}$$

where $2m\hat{\mu}_0 \approx \mu_0^2 g^{00} - m^2$. Numerically, $2 f_0'' |\vec{\nabla}\theta_+^0|^2 \approx -f_0'$. So we can write

$$\begin{aligned} 0 = \Bigg[& \left(2 \left(\frac{f_0' + 2 f_0'' \mu_0^2}{A} + 1 \right) \omega^2 - 2 \left(\frac{f_0'}{A} + 1 \right) \vec{k}^2 \right) \\ & - \frac{8 \mu_0^2 \omega^2}{\omega^2 - \vec{k}^2 - 2(\mu_0^2 g^{00} - m^2)} \\ & - \left(- \frac{2 f_0'' \mu_0^2}{A} \right) \frac{2 \cdot \gamma^2 |\vec{k}|^2 \omega^2}{\omega^2 - (1 + \gamma^2)\vec{k}^2} \Bigg] (\delta_- e^{ikx} + cc) \\ & + \frac{\mu_0}{A |\vec{\nabla}\theta_+^0|} \frac{\omega}{|\vec{k}|} \frac{\gamma}{\omega^2/\vec{k}^2 - (1 + \gamma^2)} \lambda (\delta_b e^{ikx} + cc) \,. \end{aligned} \tag{D.8}$$

As discussed in Ref. [73], at low energies with $\omega = c_s |\vec{k}|$ and for a non-relativistic background, this equation takes the form

$$\begin{aligned} 0 = & \left(-2\vec{k}^2 + \frac{2 m \omega^2}{\hat{\mu}_0} \right) (\delta_- e^{ikx} + cc) \\ & + \frac{\mu_0}{A |\vec{\nabla}\theta_+^0|} \frac{\omega}{|\vec{k}|} \frac{\gamma}{\omega^2/\vec{k}^2 - (1 + \gamma^2)} \lambda (\delta_b e^{ikx} + cc) \,. \end{aligned} \tag{D.9}$$

Note that this not only assumes $k \ll m$ but also $k \ll m\sqrt{\hat{\mu}_0/m}$ [131]. This introduces an additional cutoff on k. However, numerically this is not relevant compared to our cutoff $k_{\max} \sim 10^{-22}\,\mathrm{eV}$ that we impose anyway for independent reasons.

We will now write this in terms of

$$\bar{\delta}_- \equiv \sqrt{2A}\delta_- = \rho_0 \delta_- \,. \tag{D.10}$$

This is a useful normalization as we will see in the next section. This gives the dispersion relation

$$\omega = c_s|\vec{k}| = \sqrt{\frac{\hat{\mu}_0}{m}}|\vec{k}| \,. \tag{D.11}$$

Since $\hat{\mu}_0 \ll m$, we will have $\omega \ll k$ for on-shell radiation. We can then write

$$\begin{aligned} 0 = &\left(-\vec{k}^2 + \frac{m\omega^2}{\hat{\mu}_0}\right)(\bar{\delta}_- e^{ikx} + cc) \\ &- \frac{m}{\sqrt{2A}|\vec{\nabla}\theta_+^0|}\frac{\omega}{|\vec{k}|}\frac{\gamma}{1+\gamma^2}\lambda(\delta_b e^{ikx} + cc) \,, \end{aligned} \tag{D.12}$$

which gives

$$\begin{aligned} 0 = &\left(-\vec{k}^2 + \frac{m\omega^2}{\hat{\mu}}\right)(\bar{\delta}_- e^{ikx} + cc) \\ &- \frac{1}{\sqrt{2}M_{\rm Pl}}\frac{\sqrt{\lambda_4}\gamma}{\bar{\alpha}}\frac{a_0}{|\vec{a}_{\theta_+^0}|}\frac{\gamma}{1+\gamma^2}(\delta_b e^{ikx} + cc) \,. \end{aligned} \tag{D.13}$$

D.2. Effective Lagrangian

Our calculation so far tells us how the non-relativistic mode of two-field SFDM reacts to a given perturber. But not how much energy such a mode carries. For this, we calculate the associated effective Lagrangian. I.e., we simply put our results for δ_ρ and δ_+ back into the original Lagrangian $\mathcal{L}$ from Eq. (D.1). In our case, we have $\omega \ll k$ and $k^2 \ll m^2 c_s^2$. This gives

$$\delta_\rho \approx \frac{1}{2}(i\omega)\frac{\sqrt{2A}}{\hat{\mu}}\delta_- \,, \tag{D.14a}$$

$$\begin{aligned} \delta_+ &\approx \frac{2f_0''\mu_0\vec{\nabla}\theta_+^0\vec{k}\omega}{f_0'(1+\gamma^2)k^2}\delta_- + \frac{\lambda\delta_b}{2f_0'k^2(1+\gamma^2)} \\ &\approx -c_s\frac{\bar{\alpha}m\Lambda}{M_{\rm Pl}a_{\theta_0}}\frac{\gamma}{1+\gamma^2}\delta_- + \frac{\lambda\delta_b}{2f_0'k^2(1+\gamma^2)} \equiv (\delta_+)_- + (\delta_+)_b \,. \end{aligned} \tag{D.14b}$$

Numerically,

$$(\delta_+)_- = -\frac{c_s}{10^{7/4}}\frac{\gamma}{1+\gamma^2}\frac{a_0}{a_{\theta_0}}\frac{1}{\sqrt{a_0 r_0}}\left(\frac{\bar{a}}{a_0}\right)^{1/8}\delta_- \,, \tag{D.15}$$

where the prefactor of δ_- is typically not much larger than 1.

The non-negligible terms in the Lagrangian are

$$\begin{aligned}\mathcal{L} \approx &- (\mu_0^2 g^{00} - m^2)\delta_\rho^2 - (f_0' - 2f_0''\gamma^2|\vec{\nabla}\theta_+|^2)(\vec{\nabla}\delta_+)^2 \\ &- (f_0' + A)(\vec{\nabla}\delta_-)^2 + \left(-4f_0''\mu_0|\vec{\nabla}\theta_+^0|\gamma\right)|\vec{\nabla}\delta_+|\dot{\delta}_- \\ &+ 2\mu_0\rho_0\delta_\rho\dot{\delta}_- - \lambda\delta_+\delta_b\,.\end{aligned} \tag{D.16}$$

A coupling of δ_- to the matter density δ_b can in principle come from three terms. From the $\delta_+\delta_b$ term, from the $(\vec{\nabla}\delta_+)^2$ term, and from the $\vec{\nabla}\delta_+\dot{\delta}_-$ term. However, the contributions from the latter two terms cancel,

$$\begin{aligned}&- f_0'(1+\gamma^2)2(-i\vec{k})^2(\delta_+)_b(\delta_+)_- - 4f_0''\mu_0\vec{\nabla}\theta_+^0(-i\vec{k})(i\omega)(\delta_+)_b\delta_- \\ =&2(\delta_+)_b\left(f_0'(1+\gamma^2)k^2(\delta_+)_- - 2f_0''\mu_0\vec{k}\omega\vec{\nabla}\theta_+^0\delta_-\right) = 0\,.\end{aligned} \tag{D.17}$$

Thus, the coupling to matter comes only from the $\delta_+\delta_b$ term. We further have

$$2\mu_0\rho_0\delta_\rho\dot{\delta}_- \approx \rho_0^2\frac{1}{c_s^2}\dot{\delta}_-^2\,, \tag{D.18a}$$

$$-(\mu_0^2 g^{00} - m^2)\delta_\rho^2 \approx -\frac{1}{2}\rho_0^2\frac{1}{c_s^2}\dot{\delta}_-^2\,. \tag{D.18b}$$

Both terms together give the $\dot{\delta}_-^2$ term. The mixing term $\vec{\nabla}\delta_+\dot{\delta}_-$ in principle also contributes to the $\dot{\delta}_-^2$ term. However, its size relative to the other contributions is of order

$$\frac{m^2}{|\vec{\nabla}\theta_+^0|^2}\frac{f_0'}{A}c_s^2 = \frac{1}{10^7}\left(\frac{a_0}{a_{\theta_0}}\right)\left(\frac{\bar{a}}{a_0}\right)^{1/2}\,, \tag{D.19}$$

which is typically much smaller than 1.

This finally gives, using $A \gg f_0'$,

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\rho_0^2\frac{1}{c_s^2}\dot{\delta}_-^2 - \frac{1}{2}\rho_0^2(\vec{\nabla}\delta_-)^2 - \lambda\delta_+\delta_b \\ &= \frac{1}{2}\frac{1}{c_s^2}\dot{\bar{\delta}}_-^2 - \frac{1}{2}(\vec{\nabla}\bar{\delta}_-)^2 - \frac{\lambda}{\rho_0}\frac{\delta_+}{\delta_-}\bar{\delta}_-\delta_b \\ &= \frac{1}{2}\frac{1}{c_s^2}\dot{\bar{\delta}}_-^2 - \frac{1}{2}(\vec{\nabla}\bar{\delta}_-)^2 + \frac{1}{\sqrt{2}M_{\rm Pl}}\frac{\sqrt{\lambda_4}}{\bar{\alpha}}\frac{a_0}{a_{\theta_0}}\frac{\gamma}{1+\gamma^2}\bar{\delta}_-\delta_b\,.\end{aligned} \tag{D.20}$$

This is our prototype Lagrangian without the $\hat{a}$ term and with

$$\bar{c} = \sqrt{\frac{\hat{\mu}}{m}}\,, \quad g_m = -\frac{\sqrt{\lambda_4}}{\bar{\alpha}}\frac{a_0}{a_{\theta_0}}\frac{\gamma}{1+\gamma^2}\,. \tag{D.21}$$

To calculate the energy loss through Cherenkov radiation, we can reuse our standard calculation of $\dot{E}$ from Appendix A after adjustments for the missing $\hat{a}$ term and the factor $\gamma/(1+\gamma^2)$ in the coupling g_m. This is discussed in Sec. 4.4 and Appendix A.4.

E. SZ model perturbations

In this Appendix, we consider perturbations of the SZ model on top of a Minkowski background. Here, we use the metric signature $(-,+,+,+)$ to facilitate easier comparison to Ref. [25]. Following Ref. [25] we multiply the Lagrangian of the SZ model by $16\pi\tilde{G} = 2/\tilde{M}_{\rm Pl}^2$. Then, the Lagrangian for the perturbations is [25], see also Sec. 4.5,

$$\begin{aligned} S = \int d^4x \Big\{ & (\text{standard perturbations from } R) + \frac{1}{\tilde{M}_{\rm Pl}^2} T_{\alpha\beta} h^{\alpha\beta} \\ & + K_{\rm B} |\dot{\vec{A}} - \frac{1}{2}\vec{\nabla} h_{00}|^2 - 2K_{\rm B}\vec{\nabla}_{[i}A_{j]}\vec{\nabla}^{[i}A^{j]} \\ & + 2\,(2-K_{\rm B})\left(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}\right)\cdot\vec{\nabla}\varphi + 2(2-K_{\rm B})Q_0\vec{A}_i\left(-\frac{1}{2}\partial_i h_{00}\right) \\ & - (2-K_{\rm B})(1+\lambda_s)\Big((\vec{A}^2 + (h^{0i})^2)Q_0^2 + (\vec{\nabla}\varphi)^2 \\ & \qquad\qquad + 2Q_0(-h^{0i} + A^i)(\partial_i\varphi)\Big) \\ & + 2\mathcal{K}_2\left(\dot{\varphi} + \frac{1}{2}h_{00}Q_0\right)^2 \Big\}\,, \end{aligned} \tag{E.1}$$

where "(standard perturbations from R)" denotes the metric perturbations from the Ricci scalar as in standard General Relativity. In the following, we set $h^{0i} = 0$.

E.1. Equations of motion

The $h_{\alpha\beta}$ equation is

$$\begin{aligned} (\text{standard})_{\alpha\beta} = & \left(K_{\rm B}\vec{\nabla}(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}) + (2-K_{\rm B})(\vec{\nabla}^2\varphi + Q_0\vec{\nabla}\vec{A}) \right. \\ & \left. + 2\mathcal{K}_2 Q_0\left(\dot{\varphi} + \frac{1}{2}h_{00}Q_0\right)\right)\delta^0_\alpha\delta^0_\beta + \frac{1}{\tilde{M}_{\rm Pl}^2}T_{\alpha\beta} \\ \equiv & \frac{1}{M_{\rm Pl}^2}\bar{T}_{\alpha\beta}\,, \end{aligned} \tag{E.2}$$

where "$(\text{standard})_{\mu\nu}$" denotes terms from the metric perturbations from the Ricci scalar as in standard General Relativity. In the harmonic gauge [141],

$$\frac{1}{2}\Box h_{\alpha\beta} = -\frac{1}{\tilde{M}_{\rm Pl}^2}\left(\bar{T}_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta}\bar{T}^\rho_\rho\right)\,. \tag{E.3}$$

For the 00 component, this is

$$(\vec{\nabla}^2 - \partial_t^2)h_{00} = -\frac{2}{\tilde{M}_{\rm Pl}^2}S_{00} - K_{\rm B}\,\vec{\nabla}(\dot{\vec{A}} - \frac{1}{2}\vec{\nabla}h_{00}) - (2-K_{\rm B})(\vec{\nabla}^2\varphi + Q_0\vec{\nabla}\vec{A}) \\ - 2\mathcal{K}_2 Q_0\left(\dot{\varphi} + \frac{1}{2}h_{00}Q_0\right), \quad \text{(E.4)}$$

where $S_{\alpha\beta} = T_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta}T^\rho_\rho$. We can write this as

$$\left((1-K_{\rm B}/2)\vec{\nabla}^2 - \partial_t^2 + \mathcal{K}_2 Q_0^2\right)h_{00} \\ = -\frac{2}{\tilde{M}_{\rm Pl}^2}S_{00} - K_{\rm B}\,\vec{\nabla}\dot{\vec{A}} - (2-K_{\rm B})(\vec{\nabla}^2\varphi + Q_0\vec{\nabla}\vec{A}) - 2\mathcal{K}_2 Q_0\dot{\varphi}\,. \quad \text{(E.5)}$$

In Fourier space this gives

$$h_{00} = \frac{-\frac{2}{\tilde{M}_{\rm Pl}^2}S_{00} - K_{\rm B}\omega\vec{k}\vec{A} + (2-K_{\rm B})(\vec{k}^2\varphi + iQ_0\vec{k}\vec{A}) - 2\mathcal{K}_2 Q_0 i\omega\varphi}{\omega^2 - \vec{k}^2(1-K_{\rm B}/2) + \mathcal{K}_2 Q_0^2}\,. \quad \text{(E.6)}$$

The $\vec{A}$ equation reads

$$0 = -2K_{\rm B}\partial_t\left(\partial_t\vec{A} - \frac{1}{2}\vec{\nabla}h_{00}\right) + 2K_{\rm B}\left(\vec{\nabla}^2\vec{A} - \vec{\nabla}(\vec{\nabla}\cdot\vec{A})\right) \\ - 2(2-K_{\rm B})\partial_t\vec{\nabla}\varphi - 2(2-K_{\rm B})(1+\lambda_s)\left(Q_0^2\vec{A} + Q_0\vec{\nabla}\varphi\right) \\ + 2(2-K_{\rm B})Q_0\left(-\frac{1}{2}\vec{\nabla}h_{00}\right). \quad \text{(E.7)}$$

In Fourier space, we can write this as

$$\left(2K_{\rm B}\omega^2 - 2(2-K_{\rm B})(1+\lambda_s)Q_0^2\right)\vec{A} - 2K_{\rm B}\left(\vec{k}^2\vec{A} - \vec{k}(\vec{k}\cdot\vec{A})\right) \\ = -\omega\vec{k}\,(K_{\rm B}h_{00} - 2(2-K_{\rm B})\varphi) - 2(2-K_{\rm B})iQ_0\vec{k}\left((1+\lambda_s)\varphi + \frac{1}{2}h_{00}\right). \quad \text{(E.8)}$$

The only vector in this equation is $\vec{k}$, so we can assume $\vec{A} \parallel \vec{k}$. Then,

$$\vec{A} = -\frac{\vec{k}(2-K_{\rm B})}{2K_{\rm B}}\frac{h_{00}\left(\omega\frac{K_{\rm B}}{2-K_{\rm B}} + iQ_0\right) + 2\left((1+\lambda_s)iQ_0 - \omega\right)\varphi}{\omega^2 - (2-K_{\rm B})(1+\lambda_s)Q_0^2/K_{\rm B}} \\ = -\frac{\vec{k}(2-K_{\rm B})}{2K_{\rm B}\omega}\frac{h_{00}\left(\frac{K_{\rm B}}{2-K_{\rm B}} + iQ_0/\omega\right) - 2\left(1-(1+\lambda_s)iQ_0/\omega\right)\varphi}{1-(2-K_{\rm B})(1+\lambda_s)Q_0^2/(K_{\rm B}\omega^2)}\,. \quad \text{(E.9)}$$

The φ equation of motion is

$$0 = -4\mathcal{K}_2\ddot{\varphi} + 2(2-K_{\rm B})(1+\lambda_s)\left(\vec{\nabla}^2\varphi + Q_0\vec{\nabla}\vec{A}\right) \\ + (2-K_{\rm B})\vec{\nabla}^2 h_{00} - 2(2-K_{\rm B})\partial_t\vec{\nabla}\vec{A} - 4\mathcal{K}_2\left(\frac{1}{2}\dot{h}_{00}Q_0\right). \quad \text{(E.10)}$$

In Fourier space, this is

$$\begin{aligned} 0 =& 4\mathcal{K}_2\omega^2\varphi - 2(2-K_\mathrm{B})(1+\lambda_s)\vec{k}^2\left(\varphi + \frac{iQ_0\vec{k}}{\vec{k}^2}\vec{A}\right) \\ & - (2-K_\mathrm{B})\vec{k}^2 h_{00}\left(1 + \frac{1}{2-K_\mathrm{B}}\frac{i\omega}{Q_0}\frac{2\mathcal{K}_2Q_0^2}{\vec{k}^2}\right) - 2(2-K_\mathrm{B})\omega\vec{k}\vec{A}\,. \end{aligned} \tag{E.11}$$

For a useful static limit we need

$$\sqrt{\mathcal{K}_2}Q_0 \ll |\vec{k}|\,, \tag{E.12}$$

as discussed in Ref. [25]. We are interested in non-relativistic modes $\omega = c_s k$ with $c_s \ll 1$. Thus, we assume

$$\omega \ll |\vec{k}|\,. \tag{E.13}$$

As discussed in Sec. 4.5, we also assume

$$2 - K_\mathrm{B} = \mathcal{O}(1)\,, \quad 1 + \lambda_s = \mathcal{O}(1)\,. \tag{E.14}$$

Then, we have

$$(2-K_\mathrm{B})\vec{k}^2 h_{00} = 2f_{00}\left(\frac{2}{\tilde{M}_\mathrm{Pl}^2}S_{00} + K_\mathrm{B}f_A'\omega\vec{k}\vec{A} - (2-K_\mathrm{B})f_\varphi\vec{k}^2\varphi\right)\,, \tag{E.15}$$

with

$$\begin{aligned} f_\varphi &\equiv 1 - \frac{1}{2-K_\mathrm{B}}\frac{2\mathcal{K}_2Q_0^2}{k^2}\frac{i\omega}{Q_0}\,, \\ f_{00}^{-1} &\equiv 1 - \omega^2/(\vec{k}^2(1-K_\mathrm{B}/2)) - \mathcal{K}_2Q_0^2/(k^2(1-K_\mathrm{B}/2))\,, \\ f_A^{-1} &\equiv 1 - (2-K_\mathrm{B})(1+\lambda_s)Q_0^2/(\omega^2 K_\mathrm{B})\,, \\ f_{Q1} &\equiv 1 + (iQ_0/\omega)(2-K_\mathrm{B})/K_\mathrm{B}\,, \\ f_{Q2} &\equiv 1 - (1+\lambda_s)(iQ_0/\omega)\,, \\ f_A' &\equiv 1 - (2-K_\mathrm{B})\frac{iQ_0}{K_\mathrm{B}\omega}\,, \end{aligned} \tag{E.16}$$

which are all typically close to 1 due to our assumptions, except if K_B is very small. This is because below we will find $\omega \sim k/\sqrt{K_\mathrm{B}\mathcal{K}_2}$. Thus, f_A', f_{Q1}, and f_φ can be large if K_B is sufficiently small. For $\vec{A}$, we have

$$\vec{A} = -f_A\frac{\vec{k}}{\omega}\left(\frac{1}{2}h_{00}f_{Q1} - \frac{2-K_\mathrm{B}}{K_\mathrm{B}}f_{Q2}\varphi\right)\,. \tag{E.17}$$

In the φ equation of motion, this gives

$$\begin{aligned} 0 =& 4\mathcal{K}_2\omega^2\varphi \\ & - 2(2-K_\mathrm{B})(1+\lambda_s)\vec{k}^2\varphi\left(1 + \frac{2-K_\mathrm{B}}{K_\mathrm{B}}f_A f_{Q2}\left(\frac{iQ_0}{\omega} + \frac{1}{1+\lambda_s}\right)\right) \\ & - (2-K_\mathrm{B})\vec{k}^2 h_{00}\left((1-f_A f_{Q1}) + \frac{1}{2-K_\mathrm{B}}\frac{i\omega}{Q_0}\frac{2\mathcal{K}_2Q_0^2}{\vec{k}^2}\right. \\ & \qquad \left. - \frac{iQ_0}{\omega}(1+\lambda_s)f_A f_{Q1}\right)\,. \end{aligned} \tag{E.18}$$

In the static limit in galaxies, φ is coupled to matter due to a $k^2 h_{00}$ term with constant prefactor. Here, the only such term cancels if f_{Q1} and f_A are close to 1, i.e. if K_B is not too small. We will further discuss this coupling to matter below. If we plug the $\vec{A}$ solution into the h_{00} equation, we find

$$\begin{aligned}(2-K_\mathrm{B})\vec{k}^2 h_{00} = 2f_{00}\Bigg(&\frac{2}{\tilde{M}_\mathrm{Pl}^2}S_{00} - K_\mathrm{B}\frac{1}{2}h_{00}f_{Q1}f_A f_A'\vec{k}^2 \\ &+(2-K_\mathrm{B})f_{Q2}f_A f_A'\vec{k}^2\varphi - (2-K_\mathrm{B})f_\varphi\vec{k}^2\varphi\Bigg)\,. \end{aligned} \tag{E.19}$$

This gives

$$\begin{aligned}&(2-K_\mathrm{B}+K_\mathrm{B}f_{00}f_{Q1}f_A f_A')\vec{k}^2 h_{00} \\ &\qquad = 2f_{00}\left(\frac{2}{\tilde{M}_\mathrm{Pl}^2}S_{00} - (2-K_\mathrm{B})(f_\varphi - f_{Q2}f_A f_A')\vec{k}^2\varphi\right)\,. \end{aligned} \tag{E.20}$$

As mentioned above, f_{00} and f_A are always close to 1 since we will find $\omega = kc'/\sqrt{\mathcal{K}_2 K_\mathrm{B}}$ with c' of order 1. Then, we have

$$\begin{aligned}&2-K_\mathrm{B}+K_\mathrm{B}f_{00}f_{Q1}f_A f_A' \approx 2-K_\mathrm{B}+K_\mathrm{B}f_{Q1}f_A' \\ &= 2-K_\mathrm{B}+K_\mathrm{B}\left(1+(Q_0/(\omega K_\mathrm{B}))^2(2-K_\mathrm{B})^2\right) \\ &= 2+\left(\frac{1}{c'}\frac{Q_0\sqrt{\mathcal{K}_2}}{k}\right)^2(2-K_\mathrm{B})^2 \\ &\approx 2\,. \end{aligned} \tag{E.21}$$

Further,

$$f_\varphi - f_{Q2}f_A f_A' = \mathcal{O}\left(\frac{Q_0}{\omega K_\mathrm{B}}\right)\,. \tag{E.22}$$

In the non-relativistic limit,

$$S_{00} = T_{00} - \frac{1}{2}\eta_{00}T_0^0 = \frac{1}{2}T_{00} = \frac{1}{2}\rho_m\,, \tag{E.23}$$

where ρ_m is the matter density. This finally gives

$$k^2 h_{00} = \frac{\rho_m}{\tilde{M}_\mathrm{Pl}^2} + \mathcal{O}\left(k^2\varphi\frac{Q_0}{\omega K_\mathrm{B}}\right)\,. \tag{E.24}$$

The φ equation becomes

$$\begin{aligned}0 =&4\mathcal{K}_2\omega^2\varphi - 4\frac{2-K_\mathrm{B}}{K_\mathrm{B}}\vec{k}^2\left(1+\frac{1}{2}\lambda_s K_\mathrm{B}\right)\varphi \\ &-(2-K_\mathrm{B})\left(\frac{\rho_m}{\tilde{M}_\mathrm{Pl}^2}+\mathcal{O}\left(k^2\varphi\frac{Q_0}{\omega K_\mathrm{B}}\right)\right)X\,, \end{aligned} \tag{E.25}$$

where X will be defined shortly. Below, we will see that the $Xk^2\varphi(Q_0/\omega K_\mathrm{B})$ term does not affect the dispersion relation. The dispersion relation is $\omega = c_s|\vec{k}|$ with

$$c_s^2 K_\mathrm{B}\mathcal{K}_2 = c'^2 = (2-K_\mathrm{B})\left(1+\frac{1}{2}\lambda_s K_\mathrm{B}\right). \tag{E.26}$$

We can now evaluate the quantity X using this dispersion relation

$$X \equiv -i\frac{(2-K_\mathrm{B})(2+K_\mathrm{B}\lambda_s)(1+\lambda_s)}{\sqrt{K_\mathrm{B}}}\cdot\bar{\epsilon}^3\cdot\frac{1}{1-\bar{\epsilon}^2(2-K_\mathrm{B})(1+\lambda_s)}, \tag{E.27}$$

where $\bar{\epsilon}\equiv\sqrt{\mathcal{K}_2}Q_0/(c'k)\ll 1$. That is, we have

$$X \approx -i\frac{(2-K_\mathrm{B})(2+K_\mathrm{B}\lambda_s)(1+\lambda_s)}{\sqrt{K_\mathrm{B}}}\cdot\left(\frac{\sqrt{\mathcal{K}_2}Q_0}{c'k}\right)^3. \tag{E.28}$$

We can now check explicitly that the $Xk^2\varphi(Q_0/\omega K_\mathrm{B})$ term indeed does not change the dispersion relation, as claimed above. Corrections from this term scale as

$$k^2\varphi\left(\frac{Q_0}{\omega K_\mathrm{B}}\right)^4 K_\mathrm{B} \sim k^2\varphi\left(\frac{\sqrt{\mathcal{K}_2}Q_0}{k}\right)^4\frac{1}{K_\mathrm{B}}, \tag{E.29}$$

which is much smaller than the already-present term $k^2\varphi/K_\mathrm{B}$ since $\sqrt{\mathcal{K}_2}Q_0\ll k$. Finally, the φ equation becomes

$$\begin{aligned}0 = {}& \frac{\mathcal{K}_2 K_\mathrm{B}}{c'^2}\omega^2\bar{\varphi} - \vec{k}^2\bar{\varphi} \\ & + i\frac{\rho_m}{\sqrt{2}\tilde{M}_\mathrm{Pl}}\left(\frac{\sqrt{\mathcal{K}_2}Q_0}{|\vec{k}|}\right)^3\frac{1+\lambda_s}{\sqrt{2}\sqrt{2-K_\mathrm{B}}\left(1+\frac{1}{2}K_\mathrm{B}\lambda_s\right)^{3/2}},\end{aligned} \tag{E.30}$$

where we introduced

$$\bar{\varphi}\equiv\frac{\tilde{M}_\mathrm{Pl}}{\sqrt{K_\mathrm{B}}}\varphi. \tag{E.31}$$

E.2. Effective Lagrangian

So far, our calculation tells us how the non-relativistic mode of the SZ model reacts to a perturber density δ_b. But, as in Appendix D.2 for two-field SFDM, this does not tell us how much energy this mode carries. This requires the normalization of the effective Lagrangian. We have $\omega\ll k$, $\sqrt{\mathcal{K}_2}Q_0\ll k$, $2-K_\mathrm{B}=\mathcal{O}(1)$, and $1+\lambda_s=\mathcal{O}(1)$. This implies

$$h_{00} = \mathcal{O}\left(\varphi\frac{Q_0}{\omega K_\mathrm{B}}\right), \tag{E.32}$$

$$\vec{A} = \mathcal{O}\left(\frac{\vec{k}}{\omega}\frac{\varphi}{K_\mathrm{B}}\right). \tag{E.33}$$

Then, the spatial derivative terms in the effective Lagrangian scale as

$$\frac{\tilde{M}_{\rm Pl}^2}{K_{\rm B}}(\vec{\nabla}\varphi)^2 \sim (\vec{\nabla}\bar{\varphi})^2 \,. \tag{E.34}$$

Here, we are only interested in the order of magnitude of the terms in the effective Lagrangian since that is sufficient for our purposes, as discussed in Sec. 4.5.2. Thus, Eq. (E.34) together with the equation of motion for $\bar{\varphi}$ suffices to fix the effective Lagrangian. We have in terms of our prototype Lagrangian

$$\bar{c} = \mathcal{O}\left(\frac{1}{\sqrt{\mathcal{K}_2 K_{\rm B}}}\right) , \quad g_m = \mathcal{O}\left(\left(\frac{\sqrt{\mathcal{K}_2}Q_0}{k}\right)^3\right) . \tag{E.35}$$

There is no $\hat{a}$ term for the non-relativistic mode of the SZ model because $\hat{a}$ corresponds to the background galaxy, while we assume a Minkowski background for simplicity.

Bibliography

[1] M. Milgrom, "A Modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis", Astrophys. J. **270**, 365 (1983).

[2] M. Milgrom, "A modification of the Newtonian dynamics - Implications for galaxies.", Astrophys. J. **270**, 371 (1983).

[3] M. Milgrom, "A modification of the Newtonian dynamics: implications for galaxy systems", Astrophys. J. **270**, 384 (1983).

[4] J. Bekenstein and M. Milgrom, "Does the missing mass problem signal the breakdown of Newtonian gravity?", Astrophys. J. **286**, 7 (1984).

[5] A. E. Evrard, C. A. Metzler, and J. F. Navarro, "Mass Estimates of X-Ray Clusters", Astrophys. J **469**, 494 (1996), arXiv:astro-ph/9510058.

[6] S. S. McGaugh, "A tale of two paradigms: the mutual incommensurability of ΛCDM and MOND", Canadian Journal of Physics **93**, 250 (2015), arXiv:1404.7525.

[7] D. Clowe, A. Gonzalez, and M. Markevitch, "Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter", Astrophys. J. **604**, 596 (2004), arXiv:astro-ph/0312273.

[8] G. W. Angus, B. Famaey, and H. S. Zhao, "Can MOND take a bullet? Analytical comparisons of three versions of MOND beyond spherical symmetry", Mon. Not. Roy. Astron. Soc. **371**, 138 (2006), arXiv:astro-ph/0606216.

[9] G. W. Angus, H. Y. Shan, H. S. Zhao, and B. Famaey, "On the Proof of Dark Matter, the Law of Gravity, and the Mass of Neutrinos", Astrophys. J. Lett. **654**, L13 (2007), arXiv:astro-ph/0609125.

[10] B. Famaey and S. S. McGaugh, "Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions", Living Reviews in Relativity **15**, 10 (2012), arXiv:1112.3960.

[11] F. Lelli, S. S. McGaugh, J. M. Schombert, and M. S. Pawlowski, "One Law to Rule Them All: The Radial Acceleration Relation of Galaxies", Astrophys. J. **836**, 152 (2017), arXiv:1610.08981.

[12] M. Lisanti, M. Moschella, N. J. Outmezguine, and O. Slone, "Testing dark matter and modifications to gravity using local Milky Way observables", Phys. Rev. D **100**, 083009 (2019), arXiv:1812.08169.

[13] M. Lisanti, M. Moschella, N. J. Outmezguine, and O. Slone, "The Inconsistency of Superfluid Dark Matter with Milky Way Dynamics", arXiv e-prints (2019), arXiv:1911.12365.

[14] M. Bennett and J. Bovy, "Vertical waves in the solar neighbourhood in Gaia DR2", Mon. Not. Roy. Astron. Soc. **482**, 1417 (2019), arXiv:1809.03507.

[15] R. H. Sanders, "Cosmology with modified Newtonian dynamics (MOND)", Mon. Not. Roy. Astron. Soc. **296**, 1009 (1998), arXiv:astro-ph/9710335.

[16] S. S. McGaugh, "Distinguishing between Cold Dark Matter and Modified Newtonian Dynamics: Predictions for the Microwave Background", Astrophys. J. **523**, L99 (1999), arXiv:astro-ph/9907409.

[17] J. D. Bekenstein, "Relativistic gravitation theory for the modified Newtonian dynamics paradigm", Phys. Rev. D **70**, 083509 (2004), arXiv:astro-ph/0403694.

[18] P. G. Ferreira, "CMB and LSS Test of TeVeS", Talk at the Alternative Gravities & Dark Matter Workshop, Edinburgh, Apr. 2006.

[19] C. Skordis, D. F. Mota, P. G. Ferreira, and C. Bœhm, "Large Scale Structure in Bekenstein's Theory of Relativistic Modified Newtonian Dynamics", Phys. Rev. Lett. **96**, 011301 (2006), arXiv:astro-ph/0505519.

[20] P. A. R. Ade et al. (Planck Collaboration), "Planck 2013 results. XV. CMB power spectra and likelihood", Astronomy & Astrophysics **571**, A15 (2014), arXiv:1303.5075.

[21] B. P. Abbott et al., "Multi-messenger Observations of a Binary Neutron Star Merger", Astrophys. J. **848**, L12 (2017), arXiv:1710.05833.

[22] R. H. Sanders, "Does GW170817 falsify MOND?", Int. J. Mod. Phys. **D27**, 14 (2018), arXiv:1805.06804.

[23] S. Boran, S. Desai, E. O. Kahya, and R. P. Woodard, "GW170817 Falsifies Dark Matter Emulators", Phys. Rev. **D97**, 041501 (2018), arXiv:1710.06168.

[24] E. Babichev, C. Deffayet, and G. Esposito-Farèse, "Improving relativistic modified Newtonian dynamics with Galileon k-mouflage", Phys. Rev. D **84**, 061502 (2011), arXiv:1106.2538.

[25] C. Skordis and T. Złosnik, "New Relativistic Theory for Modified Newtonian Dynamics", Phys. Rev. Lett. **127**, 161302 (2021), arXiv:2007.00082.

[26] M. Milgrom, "Universal Modified Newtonian Dynamics Relation between the Baryonic and "Dynamical" Central Surface Densities of Disc Galaxies", Phys. Rev. Lett. **117**, 141101 (2016), arXiv:1607.05103.

[27] F. Lelli, S. S. McGaugh, J. M. Schombert, and M. S. Pawlowski, "The Relation between Stellar and Dynamical Surface Densities in the Central Regions of Disk Galaxies", Astrophys. J. Lett. **827**, L19 (2016), arXiv:1607.02145.

[28] S. McGaugh, "Testing galaxy formation and dark matter with low surface brightness galaxies", Studies in History and Philosophy of Science Part A **88**, 220 (2021), arXiv:2103.05003.

[29] K. A. Oman et al., "The unexpected diversity of dwarf galaxy rotation curves", Mon. Not. Roy. Astron. Soc. **452**, 3650 (2015), arXiv:1504.01437.

[30] S. S. McGaugh and W. J. G. de Blok, "Testing the Dark Matter Hypothesis with Low Surface Brightness Galaxies and Other Evidence", Astrophys. J. **499**, 41 (1998), arXiv:astro-ph/9801123.

[31] F. Lelli, S. S. McGaugh, and J. M. Schombert, "SPARC: Mass Models for 175 Disk Galaxies with Spitzer Photometry and Accurate Rotation Curves", The Astronomical Journal **152**, 157 (2016), arXiv:1606.09251.

[32] S. S. McGaugh, J. M. Schombert, G. D. Bothun, and W. J. G. de Blok, "The Baryonic Tully-Fisher Relation", Astrophys. J. Lett. **533**, L99 (2000), arXiv:astro-ph/0003001.

[33] R. Sancisi, "The visible matter – dark matter coupling", in Dark Matter in Galaxies, Vol. 220, edited by S. Ryder, D. Pisano, M. Walker, and K. Freeman (July 2004), p. 233, arXiv:astro-ph/0311348.

[34] S. Chandrasekhar, "Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction.", Astrophys. J. **97**, 255 (1943).

[35] N. Peschken and E. L. Łokas, "Tidally induced bars in Illustris galaxies", Mon. Not. Roy. Astron. Soc. **483**, 2721 (2019), arXiv:1804.06241.

[36] M. Roshan, I. Banik, N. Ghafourian, I. Thies, B. Famaey, E. Asencio, and P. Kroupa, "Barred spiral galaxies in modified gravity theories", Mon. Not. Roy. Astron. Soc. **503**, 2833 (2021), arXiv:2103.01794.

[37] M. Roshan, N. Ghafourian, T. Kashfi, I. Banik, M. Haslbauer, V. Cuomo, B. Famaey, and P. Kroupa, "Fast galaxy bars continue to challenge standard cosmology", Mon. Not. Roy. Astron. Soc. **508**, 926 (2021), arXiv:2106.10304.

[38] F. Fragkoudi, R. J. J. Grand, R. Pakmor, V. Springel, S. D. M. White, F. Marinacci, F. A. Gomez, and J. F. Navarro, "Revisiting the tension between fast bars and the ΛCDM paradigm", Astronomy & Astrophysics **650**, L16 (2021), arXiv:2011.13942.

[39] M. S. Pawlowski, "The planes of satellite galaxies problem, suggested solutions, and open questions", Modern Physics Letters A **33**, 1830004 (2018), arXiv:1802.02579.

[40] O. Müller, M. S. Pawlowski, F. Lelli, K. Fahrion, M. Rejkuba, M. Hilker, J. Kanehisa, N. Libeskind, and H. Jerjen, "The coherent motion of Cen A dwarf satellite galaxies remains a challenge for ΛCDM cosmology", Astronomy & Astrophysics **645**, L5 (2021), arXiv:2012.08138.

[41] I. Santos-Santos, R. Domínguez-Tenreiro, H. Artal, S. E. Pedrosa, L. Bignone, F. Martínez-Serrano, M. Á. Gómez-Flechoso, P. B. Tissera, and A. Serna, "Planes of Satellites around Simulated Disk Galaxies. I. Finding High-quality Planar Configurations from Positional Information and Their Comparison to MW/M31 Data", Astrophys. J. **897**, 71 (2020), arXiv:2004.11585.

[42] M. Bílek, I. Thies, P. Kroupa, and B. Famaey, "MOND simulation suggests an origin for some peculiarities in the Local Group", Astronomy & Astrophysics **614**, A59 (2018), arXiv:1712.04938.

[43] C. R. Mullis, P. Rosati, G. Lamer, H. Böhringer, A. Schwope, P. Schuecker, and R. Fassbender, "Discovery of an X-Ray-luminous Galaxy Cluster at z=1.4", Astrophys. J. Lett. **623**, L85 (2005), arXiv:astro-ph/0503004.

[44] F. Menanteau et al., "The Atacama Cosmology Telescope: ACT-CL J0102-4915 "El Gordo," a Massive Merging Cluster at Redshift 0.87", Astrophys. J. **748**, 7 (2012), arXiv:1109.0953.

[45] J. R. Franck, S. S. McGaugh, and J. M. Schombert, "Three Candidate Clusters Around High Redshift Radio-loud Sources: MG1 J04426+0202, 3C 068.2, and MS 1426.9+1052", Astronom. J. **150**, 46 (2015), arXiv:1505.07864.

[46] K. Finner, M. J. Jee, N. Golovich, D. Wittman, W. Dawson, D. Gruen, A. M. Koekemoer, B. C. Lemaux, and S. Seitz, "MC^2: Subaru and Hubble Space Telescope Weak-lensing Analysis of the Double Radio Relic Galaxy Cluster PLCK G287.0+32.9", Astrophys. J. **851**, 46 (2017), arXiv:1710.02527.

[47] E. Asencio, I. Banik, and P. Kroupa, "A massive blow for ΛCDM - the high redshift, mass, and collision velocity of the interacting galaxy cluster El Gordo contradicts concordance cosmology", Mon. Not. Roy. Astron. Soc. **500**, 5249 (2021), arXiv:2012.03950.

[48] J. Kim, M. J. Jee, J. P. Hughes, M. Yoon, K. HyeongHan, F. Menanteau, C. Sifón, L. Hovey, and P. Arunachalam, "Head-to-Toe Measurement of El Gordo: Improved Analysis of the Galaxy Cluster ACT-CL J0102-4915 with New Wide-field Hubble Space Telescope Imaging Data", Astrophys. J. **923**, 101 (2021), arXiv:2106.00031.

[49] I. Banik and H. Zhao, "From galactic bars to the Hubble tension – weighing up the astrophysical evidence for Milgromian gravity", arXiv e-prints (2021), arXiv:2110.06936.

[50] R. Thompson, R. Davé, and K. Nagamine, "The rise and fall of a challenger: the Bullet Cluster in Λ cold dark matter simulations", Mon. Not. Roy. Astron. Soc. **452**, 3030 (2015), arXiv:1410.7438.

[51] D. Kraljic and S. Sarkar, "How rare is the Bullet Cluster (in a ΛCDM universe)?", JCAP **2015**, 050 (2015), arXiv:1412.7719.

[52] K.-H. Chae, F. Lelli, H. Desmond, S. S. McGaugh, P. Li, and J. M. Schombert, "Testing the Strong Equivalence Principle: Detection of the External Field Effect in Rotationally Supported Galaxies", Astrophys. J. **904**, 51 (2020), arXiv:2009.11525.

[53] K.-H. Chae, H. Desmond, F. Lelli, S. S. McGaugh, and J. M. Schombert, "Testing the Strong Equivalence Principle. II. Relating the External Field Effect in Galaxy Rotation Curves to the Large-scale Structure of the Universe", Astrophys. J. **921**, 104 (2021), arXiv:2109.04745.

[54] A. Paranjape and R. K. Sheth, "The external field effect in cold dark matter models", arXiv e-prints (2021), arXiv:2112.00026.

[55] J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen, and N. Mahesh, "An absorption profile centred at 78 megahertz in the sky-averaged spectrum", Nature **555**, 67 (2018), arXiv:1810.05912.

[56] S. S. McGaugh, "Predictions for the Sky-Averaged Depth of the 21 cm Absorption Signal at High Redshift in Cosmologies with and without Nonbaryonic Cold Dark Matter", Phys. Rev. Lett. **121**, 081305 (2018), arXiv:1808.02532.

[57] B. D. Fields, "The Primordial Lithium Problem", Annual Review of Nuclear and Particle Science **61**, 47 (2011), arXiv:1203.3551.

[58] D. Merritt, "Theory Variant T2: A Relativistic Theory", in *A Philosophical Approach to MOND: Assessing the Milgromian Research Program in Cosmology* (Cambridge University Press, 2020), pp. 117–180.

[59] A. G. Riess, S. Casertano, W. Yuan, J. B. Bowers, L. Macri, J. C. Zinn, and D. Scolnic, "Cosmic Distances Calibrated to 1% Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with ΛCDM", Astrophys. J. Lett. **908**, L6 (2021), arXiv:2012.08534.

[60] N. Aghanim et al. (Planck Collaboration), "Planck 2018 results. VI. Cosmological parameters", Astronomy & Astrophysics **641**, A6 (2020), arXiv:1807.06209.

[61] J. Schombert, S. McGaugh, and F. Lelli, "Using the Baryonic Tully-Fisher Relation to Measure H_o", Astronom. J. **160**, 71 (2020), arXiv:2006.08615.

[62] K. Jedamzik, L. Pogosian, and G.-B. Zhao, "Why reducing the cosmic sound horizon alone can not fully resolve the Hubble tension", Communications Physics **4**, 123 (2021), arXiv:2010.04158.

[63] W. Lin, X. Chen, and K. J. Mack, "Early Universe Physics Insensitive and Uncalibrated Cosmic Standards: Constraints on Ω_m and Implications for the Hubble Tension", Astrophys. J. **920**, 159 (2021), arXiv:2102.05701.

[64] W. D. Kenworthy, D. Scolnic, and A. Riess, "The Local Perspective on the Hubble Tension: Local Structure Does Not Impact Measurement of the Hubble Constant", Astrophys. J. **875**, 145 (2019), arXiv:1901.08681.

[65] J. Colin, R. Mohayaee, M. Rameez, and S. Sarkar, "Evidence for anisotropy of cosmic acceleration", Astronomy & Astrophysics **631**, L13 (2019), arXiv:1808.04597.

[66] M. Rameez and S. Sarkar, "Is there really a Hubble tension?", Classical and Quantum Gravity **38**, 154005 (2021), arXiv:1911.06456.

[67] K. Migkas, G. Schellenberger, T. H. Reiprich, F. Pacaud, M. E. Ramos-Ceja, and L. Lovisari, "Probing cosmic isotropy with a new X-ray galaxy cluster sample through the L_X-T scaling relation", Astronomy & Astrophysics **636**, A15 (2020), arXiv:2004.03305.

[68] M. Haslbauer, I. Banik, and P. Kroupa, "The KBC void and Hubble tension contradict ΛCDM on a Gpc scale - Milgromian dynamics as a possible solution", Mon. Not. Roy. Astron. Soc. **499**, 2845 (2020), arXiv:2009.11292.

[69] N. J. Secrest, S. von Hausegger, M. Rameez, R. Mohayaee, S. Sarkar, and J. Colin, "A Test of the Cosmological Principle with Quasars", Astrophys. J. Lett. **908**, L51 (2021), arXiv:2009.14826.

[70] J. H. W. Wong, T. Shanks, and N. Metcalfe, "The Local Hole: a galaxy under-density covering 90% of sky to ∼200 Mpc", arXiv e-prints (2021), arXiv:2107.08505.

[71] L. Berezhiani and J. Khoury, "Theory of dark matter superfluidity", Phys. Rev. **D92**, 103510 (2015), arXiv:1507.01019.

[72] L. Berezhiani, B. Famaey, and J. Khoury, "Phenomenological consequences of superfluid dark matter with baryon-phonon coupling", JCAP **1809**, 021 (2018), arXiv:1711.05748.

[73] T. Mistele, "Three problems of superfluid dark matter and their solution", JCAP **2021**, 025 (2021), arXiv:2009.03003.

[74] G. W. Angus, "Is an 11eV sterile neutrino consistent with clusters, the cosmic microwave background and modified Newtonian dynamics?", Mon. Not. Roy. Astron. Soc. **394**, 527 (2009), arXiv:0805.4014.

[75] L. Blanchet and L. Heisenberg, "Dipolar dark matter with massive bigravity", JCAP **2015**, 026 (2015), arXiv:1505.05146.

[76] H. Zhao and B. Li, "Dark Fluid: A Unified Framework for Modified Newtonian Dynamics, Dark Matter, and Dark Energy", Astrophys. J. **712**, 130 (2010), arXiv:0804.1588.

[77] J. Khoury, "Alternative to particle dark matter", Phys. Rev. D **91**, 024022 (2015), arXiv:1409.0012.

[78] J.-P. Bruneton, S. Liberati, L. Sindoni, and B. Famaey, "Reconciling MOND and dark matter?", JCAP **2009**, 021 (2009), arXiv:0811.3143.

[79] J.-P. Bruneton and G. Esposito-Farèse, "Field-theoretical formulations of MOND-like gravity", Phys. Rev. D **76**, 124012 (2007), arXiv:0705.4043.

[80] E. P. Verlinde, "Emergent Gravity and the Dark Universe", SciPost Phys. **2**, 016 (2017), arXiv:1611.02269.

[81] E. Verlinde, "On the origin of gravity and the laws of Newton", JHEP **2011**, 29 (2011), arXiv:1001.0785.

[82] S. Hossenfelder, "Covariant version of Verlinde's emergent gravity", Phys. Rev. D **95**, 124018 (2017), arXiv:1703.01415.

[83] C. Skordis and T. Zlosnik, "Linear stability of the new relativistic theory of modified Newtonian dynamics", arXiv e-prints (2021), arXiv:2109.13287.

[84] R. Brada and M. Milgrom, "Exact solutions and approximations of MOND fields of disc galaxies", Mon. Not. Roy. Astron. Soc. **276**, 453 (1995).

[85] S. Hossenfelder and T. Mistele, "The Milky Way's rotation curve with superfluid dark matter", Mon. Not. Roy. Astron. Soc. **498**, 3484 (2020), arXiv:2003.07324.

[86] A. Schmitt, "Introduction to Superfluidity", Lect. Notes Phys. **888**, pp.1 (2015), arXiv:1404.1284.

[87] T. Mistele, "The role of the chemical potential in coupling superfluid dark matter to baryons", JCAP **1911**, 039 (2019), arXiv:1909.05710.

[88] J. I. Kapusta, "Bose-Einstein Condensation, Spontaneous Symmetry Breaking, and Gauge Theories", Phys. Rev. **D24**, 426 (1981).

[89] H. E. Haber and H. A. Weldon, "Finite Temperature Symmetry Breaking as Bose-Einstein Condensation", Phys. Rev. **D25**, 502 (1982).

[90] N. Bilic, "Thermodynamics of k-essence", Phys. Rev. **D78**, 105012 (2008), arXiv:0806.0642.

[91] J. D. Bekenstein, "The missing light puzzle: a hint about gravitation?", in Proceedings of the 2nd Canadian Conference on General Relativity and Relativistic Astrophysics (Jan. 1988), pp. 68–104.

[92] J. D. Bekenstein, "Phase coupling gravitation: Symmetries and gauge fields", Physics Letters B **202**, 497 (1988).

[93] L. Berezhiani, G. Cintia, and M. Warkentin, "Core Fragmentation in Simplest Superfluid Dark Matter Scenario", Physics Letters B **819**, 136422 (2021), arXiv:2101.08117.

[94] A. A. Dutton and A. V. Macciò, "Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles", Mon. Not. Roy. Astron. Soc. **441**, 3359 (2014), arXiv:1402.7073.

[95] S. Hossenfelder and T. Mistele, "Strong lensing with superfluid dark matter", JCAP **1902**, 001 (2019), arXiv:1809.00840.

[96] A. S. Bolton, S. Burles, L. V. E. Koopmans, T. Treu, R. Gavazzi, L. A. Moustakas, R. Wayth, and D. J. Schlegel, "The Sloan Lens ACS Survey. V. The Full ACS Strong-Lens Sample", Astrophys. J. **682**, 964 (2008), arXiv:0805.1931.

[97] M. W. Auger, T. Treu, A. S. Bolton, R. Gavazzi, L. V. E. Koopmans, P. J. Marshall, K. Bundy, and L. A. Moustakas, "The Sloan Lens ACS Survey. IX. Colors, Lensing and Stellar Masses of Early-type Galaxies", Astrophys. J. **705**, 1099 (2009), arXiv:0911.2471.

[98] R. H. Sanders, "A dearth of dark matter in strong gravitational lenses", Mon. Not. Roy. Astron. Soc. **439**, 1781 (2014), arXiv:1310.6148.

[99] W. Jaffe, "A Simple model for the distribution of light in spherical galaxies", Mon. Not. Roy. Astron. Soc. **202**, 995 (1983).

[100] Wolfram Research, Inc., *Mathematica, Version 11.*

[101] 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**, 750 (2010), arXiv:0907.4992.

[102] M. Milgrom, "Solutions for the Modified Newtonian Dynamics Field Equation", Astrophys. J. **302**, 617 (1986).

[103] Wolfram Research, Inc., *Mathematica, Version 12.*

[104] S. McGaugh, "The Imprint of Spiral Arms on the Galactic Rotation Curve", Astrophys. J. **885**, 87 (2019), arXiv:1909.11158.

[105] J. Bovy and H.-W. Rix, "A Direct Dynamical Measurement of the Milky Way's Disk Surface Density Profile, Disk Scale Length, and Dark Matter Profile at 4 kpc $\lesssim$R$\lesssim$ 9 kpc", Astrophys. J. **779**, 115 (2013), arXiv:1309.0809.

[106] S. S. McGaugh, "Milky Way Mass Models and MOND", Astrophys. J. **683**, 137 (2008), arXiv:0804.1314.

[107] A.-C. Eilers, D. W. Hogg, H.-W. Rix, and M. K. Ness, "The Circular Velocity Curve of the Milky Way from 5 to 25 kpc", Astrophys. J. **871**, 120 (2019), arXiv:1810.09466.

[108] M. Portail, O. Gerhard, C. Wegg, and M. Ness, "Dynamical modelling of the galactic bulge and bar: the Milky Way's pattern speed, stellar and dark matter mass distribution", Mon. Not. Roy. Astron. Soc. **465**, 1621 (2017), arXiv:1608.07954.

[109] S. Hossenfelder and T. Mistele, "The redshift-dependence of radial acceleration: Modified gravity versus particle dark matter", International Journal of Modern Physics D **27**, 1847010 (2018), arXiv:1803.08683.

[110] J. Bland-Hawthorn and O. Gerhard, "The Galaxy in Context: Structural, Kinematic, and Integrated Properties", Annu. Rev. Astron. Astrophys. **54**, 529 (2016), arXiv:1602.07702.

[111] J.-P. Bruneton, "Causality and superluminal behavior in classical field theories: Applications to k-essence theories and modified-Newtonian-dynamics-like theories of gravity", Phys. Rev. D **75**, 085013 (2007), arXiv:gr-qc/0607055.

[112] E. Babichev, V. Mukhanov, and A. Vikman, "k-Essence, superluminal propagation, causality and emergent geometry", Journal of High Energy Physics **2008**, 101 (2008), arXiv:0708.0561.

[113] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, "Causality, analyticity and an IR obstruction to UV completion", Journal of High Energy Physics **2006**, 014 (2006), arXiv:hep-th/0602178.

[114] E. Babichev, S. Ramazanov, and A. Vikman, "Recovering P(X) from a canonical complex field", JCAP **2018**, 023 (2018), arXiv:1807.10281.

[115] G. Dvali, G. F. Giudice, C. Gomez, and A. Kehagias, "UV-completion by classicalization", Journal of High Energy Physics **2011**, 108 (2011), arXiv:1010.1415.

[116] G. Dvali, A. Franca, and C. Gomez, "Road Signs for UV-Completion", arXiv e-prints (2012), arXiv:1204.6388.

[117] A. Vikman, "Suppressing quantum fluctuations in classicalization", EPL (Europhysics Letters) **101**, 34001 (2013), arXiv:1208.3647.

[118] A. Addazi and A. Marcianò, "UV self-completion of a theory of superfluid dark matter", European Physical Journal C **79**, 354 (2019), arXiv:1801.04083.

[119] I. I. Tkachev, "Coherent Scalar-Field Oscillations Forming Compact Astrophysical Object", Soviet Astronomy Letters **12**, 305 (1986).

[120] M. Colpi, S. L. Shapiro, and I. Wasserman, "Boson stars: Gravitational equilibria of self-interacting scalar fields", Phys. Rev. Lett. **57**, 2485 (1986).

[121] J.-W. Lee and I.-G. Koh, "Galactic halos as boson stars", Phys. Rev. D **53**, 2236 (1996), arXiv:hep-ph/9507385.

[122] R. Sharma, S. Karmakar, and S. Mukherjee, "Boson star and dark matter", (2008), arXiv:0812.3470.

[123] T. Mistele, "Cherenkov radiation from stars constrains hybrid MOND-dark-matter models", arXiv e-prints (2021), arXiv:2103.16954.

[124] G. D. Moore and A. E. Nelson, "Lower bound on the propagation speed of gravity from gravitational Cherenkov radiation", JHEP **2001**, 023 (2001), arXiv:hep-ph/0106220.

[125] C. M. Caves, "Gravitational radiation and the ultimate speed in Rosen's bimetric theory of gravity.", Annals of Physics **125**, 35 (1980).

[126] J. W. Elliott, G. D. Moore, and H. Stoica, "Constraining the New Aether: gravitational Cherenkov radiation", JHEP **2005**, 066 (2005), arXiv:hep-ph/0505211.

[127] M. Milgrom, "Gravitational Cherenkov Losses in Theories Based on Modified Newtonian Dynamics", Phys. Rev. Lett. **106**, 111101 (2011), arXiv:1102.1818.

[128] P. M. Chesler and A. Loeb, "Constraining Relativistic Generalizations of Modified Newtonian Dynamics with Gravitational Waves", Phys. Rev. Lett. **119**, 031102 (2017), arXiv:1704.05116.

[129] E. C. Ostriker, "Dynamical Friction in a Gaseous Medium", Astrophys. J. **513**, 252 (1999), arXiv:astro-ph/9810324.

[130] L. Berezhiani, B. Elder, and J. Khoury, "Dynamical friction in superfluids", JCAP **2019**, 074 (2019), arXiv:1905.09297.

[131] L. Berezhiani, "On effective theory of superfluid phonons", Physics Letters B **805**, 135451 (2020), arXiv:2001.08696.

[132] S. D. Tremaine, "The formation of the nuclei of galaxies. II. The local group.", Astrophys. J. **203**, 345 (1976).

[133] F. J. Sánchez-Salcedo, J. Reyes-Iturbide, and X. Hernandez, "An extensive study of dynamical friction in dwarf galaxies: the role of stars, dark matter, halo profiles and MOND", Mon. Not. Roy. Astron. Soc. **370**, 1829 (2006), arXiv:astro-ph/0601490.

[134] R. Kimura and K. Yamamoto, "Constraints on general second-order scalar-tensor models from gravitational Cherenkov radiation", JCAP **2012**, 050 (2012), arXiv:1112.4284.

[135] J. D. Jackson, *Classical Electrodynamics, 3rd Edition* (1998).

[136] C. Dalang, P. Fleury, and L. Lombriser, "Scalar Cherenkov radiation from high-energy cosmic rays", arXiv e-prints (2021), arXiv:2109.10812.

[137] M. M. Brouwer et al., "The weak lensing radial acceleration relation: Constraining modified gravity and cold dark matter theories with KiDS-1000", Astronomy & Astrophysics **650**, A113 (2021), arXiv:2106.11677.

[138] M. Milgrom, "Noncovariance at low accelerations as a route to MOND", Phys. Rev. D **100**, 084039 (2019), arXiv:1908.01691.

[139] F. D'Ambrosio, M. Garg, and L. Heisenberg, "Non-linear extension of non-metricity scalar for MOND", Physics Letters B **811**, 135970 (2020), arXiv:2004.00888.

[140] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (1995).

[141] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (1972).