

Axion-Like-Particle Dark Matter Beyond the Standard Paradigm

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und
Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
Philip Sørensen

Hamburg, 2022



CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Gutachter/innen der Dissertation:	Prof. Dr. Geraldine Servant Prof. Dr. Andreas Ringwald
Zusammensetzung der Prüfungskommission:	Prof. Dr. Geraldine Servant Prof. Dr. Andreas Ringwald Prof. Dr. Günter Sigl Prof. Dr. Erika Garutti Dr. Hyungjin Kim
Vorsitzende/r der Prüfungskommission:	Prof. Dr. Günter Sigl
Datum der Disputation:	04.10.2022
Vorsitzender Fach-Promotionsausschuss PHYSIK: Leiter des Fachbereichs PHYSIK: Dekan der Fakultät MIN:	Prof. Dr. Wolfgang J. Parak Prof. Dr. Günter H. W. Sigl Prof. Dr. Heinrich Graener

Abstract

The nature of dark matter remains one of the most important questions in modern physics. Historically, the most investigated dark matter candidates have been thermal in origin. Such candidates are championed by the paradigm of weakly interacting massive particles (WIMPs). However, with the continued non-discovery of WIMPs by the many dedicated direct detection searches as well as the non-observation of supersymmetry and other beyond-Standard-Model theories related to the WIMP paradigm, we are motivated to look elsewhere and explore new scenarios.

An interesting possibility is that dark matter may not be of a thermal origin. Non-thermal production mechanisms of dark matter suggest that dark matter may have a coupling to the Standard Model many orders of magnitude smaller than the range suggested by WIMPs. This allows for the possibility of dark matter being both very light and cosmologically stable. The canonical examples of such light, non-thermally produced dark matter candidates are the axion-like-particles, which in recent years have seen a resurgence of attention from theorists and experimentalists alike.

Axion-like-particles are pseudoscalar particles which are natural ultra-light dark matter candidates because even minimal models have an inherent production mechanism known as the misalignment mechanism. This mechanism relies on the observation that an oscillating scalar field behaves as cold dark matter and that the most generic initial conditions inevitably lead to such oscillations. However, the energy density predicted by the misalignment mechanism strongly depends on the initial condition of the axion field. In much of the parameter space that is experimentally accessible, the standard misalignment mechanism underproduces dark matter. The key motivation for this thesis is to question the standard assumption on the initial conditions for the axion field.

This thesis is devoted to the investigation of novel production mechanisms of axion-like-particle dark matter, which go beyond the standard paradigm by exploring initial conditions involving non-zero kinetic energy of the axion-like-particle field. This family of scenarios goes under the common name of *kinetic misalignment*. Our exploration of kinetic misalignment is divided into two stages: In the first stage, part II of this thesis, we investigate the dynamics of such scenarios independently of how the initial conditions are realized. In particular, we study the role of parametric resonance and the phenomenon of axion fragmentation. We find that this form of parametric resonance is indeed efficient in much of the relevant parameter space and that the produced relic, therefore, has a non-trivial momentum spectrum which could lead to observational prospects in the form of dark matter mini-clusters. In the second stage, part III of this thesis, we investigate how the assumed initial conditions can be implemented in scenarios inspired by Affleck-Dine baryogenesis. This allows us to present the full story of how and under which constraints the standard paradigm can be extended to such scenarios of kinetic misalignment. We find that a KSVZ-like model with a nearly-quadratic potential can support much of the interesting parameter space in the $10^{-6} \text{ eV} \lesssim m_a \lesssim 10 \text{ eV}$ regime if the spectrum of primordial fluctuations can be adequately suppressed. In this case, a period of kination can lead to an associated signature in amplified primordial gravitational waves.

In summary, we show that axion-like-particle dark matter can be motivated in nearly all of the $[m_a, f_a]$ parameter space that is currently unconstrained by experiments. In the coming years, a large number of searches will extend the experimental reach. These experiments have the prospect of not only discovering a QCD axion or an axion-like-particle but also having that discovery constitute all of the observed dark matter.

Zusammenfassung

Die Natur der Dunklen Materie bleibt eine der wichtigsten Fragen der modernen Physik. Die historisch am besten untersuchten Kandidaten für Dunkle Materie waren thermische Kandidaten, wie sie vom Paradigma der schwach wechselwirkenden massiven Teilchen (WIMPs) vertreten werden. Die anhaltende Nichtentdeckung von WIMPs durch die vielen dedizierten Suchen per direkter Detektion, sowie die Nichtbeobachtung von Supersymmetrie und anderen mit dem WIMP-Paradigma verbundenen Theorien, motiviert uns neue Szenarien zu erforschen.

Eine interessante Möglichkeit ist, dass Dunkle Materie keinen thermischen Ursprung hat. Nicht-thermische Produktionsmechanismen eröffnen die Möglichkeit, dass Dunkle Materie sehr leicht und trotzdem kosmologisch stabil sein kann. Das kanonische Beispiel für solche leichten, nicht thermisch erzeugten Kandidaten für Dunkle Materie sind die Axion-ähnlichen-Teilchen, die in den letzten Jahren sowohl von Theoretiker*innen als auch von Experimentator*innen wieder mehr Aufmerksamkeit erfahren haben. Axion-ähnliche-Teilchen sind pseudoskalare Teilchen, die natürliche Kandidaten für ultraleichte Dunkle Materie sind, da selbst minimale Modelle einen inhärenten Produktionsmechanismus haben, der als Fehlausrichtungsmechanismus bekannt ist. Dieser Mechanismus beruht auf der Beobachtung, dass sich ein oszillierendes Skalarfeld wie kalte Dunkle Materie verhält, und dass die allgemeinsten Anfangsbedingungen unweigerlich zu solchen Oszillationen führen. Die durch den Fehlausrichtungsmechanismus vorhergesagte Energiedichte hängt jedoch stark vom Anfangszustand des Axionfeldes ab. Die Hauptmotivation dieser Arbeit ist es, die Standardannahme zu den Anfangsbedingungen für das Axionfeld zu hinterfragen.

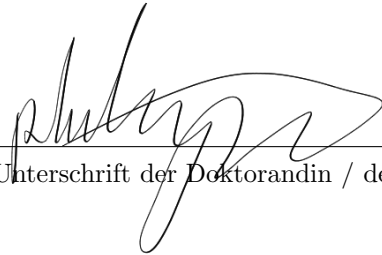
Diese Dissertation widmet sich der Untersuchung neuartiger Produktionsmechanismen Axion-ähnlicher-Teilchen Dunkler Materie, die über das Standardparadigma hinausgehen, indem sie Anfangsbedingungen untersucht, die eine von Null verschiedene kinetische Energie beinhalten. Diese Familie von Szenarien wird allgemein als *kinetische Fehlausrichtung* bezeichnet. Unsere Erforschung der kinetischen Fehlausrichtung ist in zwei Phasen unterteilt: In der ersten Phase untersuchen wir die Dynamik solcher Szenarien unabhängig davon, wie die Anfangsbedingungen realisiert werden. Insbesondere untersuchen wir die Rolle der parametrischen Resonanz und das Phänomen der Axionfragmentierung. Wir stellen fest, dass diese Form der parametrischen Resonanz in einem Großteil des relevanten Parameterraums tatsächlich effizient ist, und dass das erzeugte Relikt daher ein nicht triviales Impulsspektrum aufweist, welches in Form von Miniclustern beobachtbare Effekte in Aussicht stellt. In der zweiten Phase untersuchen wir, wie die angenommenen Anfangsbedingungen in Szenarien, die von der Affleck-Dine-Baryogenese inspiriert sind, implementiert werden können. Dies ermöglicht es uns, umfassend darzustellen, wie und unter welchen Bedingungen das Standardparadigma auf solche Szenarien kinetischer Fehlausrichtung erweitert werden kann. Wir stellen fest, dass ein KSVZ-ähnliches Modell mit einem nahezu-quadratischen Potential in einen Großteil des interessanten Parameterraums im Bereich $10^{-6} \text{ eV} \lesssim m_a \lesssim 10 \text{ eV}$ realisiert werden kann, falls das Spektrum der primordialen Fluktuationen ausreichend unterdrückt werden kann. In diesem Fall kann eine Kinationsperiode zu einer zugehörigen Signatur in verstärkten primordialen Gravitationswellen führen.

Zusammenfassend zeigen wir, dass Dunkle Materie sich durch Axion-ähnliche-Teilchen in fast dem gesamten $[m_a, f_a]$ -Parameterraum erklären lässt. In den kommenden Jahren wird eine Vielzahl von Experimenten diesen Parameterraum weiter erschließen. Diese Experimente stellen die Möglichkeit in Aussicht, nicht nur ein QCD-Axion oder ein Axion-ähnliches-Teilchen zu entdecken, sondern auch, dass diese Entdeckung die gesamte Dunkle Materie ausmacht.

Eidesstattliche Versicherung / Declaration on oath

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Padova, den 18.10.2023



Unterschrift der Doktorandin / des Doktoranden

Acknowledgments

The work behind this thesis would not have been possible without the support of many people. First, I wish to thank my supervisor Geraldine Servant for devoting her time to our project. Her perspective will surely guide me in many years to come. I want to thank Cem Eroncel and Ryosuke Sato for our collaboration, for the many good discussions and for all the things I have learnt from them. In particular, I wish to thank Pablo Quílez for the exceptional amount of time he has spent on discussions and for the many things he has taught me. Also, Hyungjin Kim, Peera Simakachorn and Yann Gouttenoire have dedicated hours of discussion that have helped me understand many topics, for which I am deeply thankful. I thank Matthias Koschnitzke and Lisa Ziegler for their assistance with the German translation of the abstract for this thesis.

Furthermore, I wish to thank Andreas Ringwald and Hyungjin Kim for agreeing to take up the roles of my co-supervisor and mentor, respectively. Also, I am grateful to Günter Sigl and Erika Garutti for agreeing to take part in my examination commission and to the excellence cluster Quantum Universe for funding my studies. More broadly, it has been a privilege to take part in the excellent research environment that is the DESY theory group, and I am grateful for having had this opportunity.

Contributions

The work described in this thesis is the result of a collaboration with Cem Eroncel, Ryosuke Sato and Geraldine Servant. Our work branched out in two main directions: One direction focused primarily on a more detailed analysis of fragmentation itself, which resulted in our first publication [1]. The other direction was focused on implementations of kick mechanisms, which resulted in our second publication [2]. Both works are based on a common analysis of fragmentation carried out jointly by me, Cem Eroncel and Ryosuke Sato.

The first paper, [1], contains both the original, shared analysis of fragmentation as well as developments which were chiefly carried out by Cem Eroncel. The second paper, [2], on which Part III is based, is the conclusion of the kick mechanism study. This paper contains work chiefly carried out by me. All appendices included in this thesis, except for appendix B, are also based on [2], and thus represent work primarily carried out by myself.

Introduction

The standard model (SM) of particle physics and the Λ CDM model of modern cosmology provide a hugely successful description of our universe. However, despite the great success of these theories, there remain a number of experimental observations that are not accounted for. These include open questions such as the nature of dark matter, dark energy and inflation, as well as the origins of the matter-antimatter asymmetry. Furthermore, the standard model itself contains a number of theoretical features which invite a more fundamental explanation. Such features include the strong CP problem, the electroweak hierarchy problem and the origin of the observed flavour structure of the SM fermions. Explanations demand physics beyond the standard model (BSM physics).

The question of dark matter has historically been dominated by the paradigm of weakly-interacting particles (WIMPs) and other such thermal dark matter candidates. However, as was alluded to in the abstract, the physics community is today driven to explore alternative candidates for dark matter by the non-observation of WIMPs and their related BSM theories. In this thesis, I investigate how the class of particles known as axion-like-particles (ALPs) may provide non-thermally produced dark matter in new regimes. The general class of ALPs owe their name to the QCD axion, which provides both an elegant solution to the strong CP problem of Quantum Chromodynamics (QCD) [3, 4] and also constitutes a natural dark matter candidate [5–7]. ALPs form a quite generic class of particles which frequently appear in string theories [8, 9] and other BSM frameworks such as majoron models [10], supersymmetric (SUSY) models [11], and dynamical models of flavour [12, 13]. An ALP is a pseudo-Nambu-Goldstone boson (pNGB) of a global $U(1)$ symmetry that is spontaneously broken at a large scale f_a , which is known as the decay constant. The low-energy properties are characterized by a small mass m_a and couplings that, due to the pNGB nature, are suppressed by f_a and are either derivative or anomalous. QCD axions and ALPs share similar effective theories. Depending on the author, the term *axion* can refer either specifically to the QCD axion or to the more general class of ALPs. In this work, I will be inclusive and use the term *axion* in the broader meaning of any axion-like-particle and refer specifically to the axion that solves the strong CP problem as the QCD axion.

The QCD axion is predicted to acquire a small mass from meson mixing [14] due to non-perturbative effects in QCD [15, 16]. In minimal implementations, this effect generally restricts the axion to a fixed relationship between the axion mass and the decay constant. We will discuss this further in section 2. For a general axion, this relationship need not apply, and such ALPs may populate a large $[m_a, f_a]$ parameter space.

The standard paradigm for the production of axion dark matter is the *misalignment mechanism*. In the standard misalignment mechanism (SMM), the dark matter arises as coherent oscillations of the axion field [5–7]. These oscillations are initiated by an initial misalignment of the axion with respect to the potential minimum. This mechanism is appealing because it is inherent to the axion scenario and requires little model building beyond the initial conditions that would be expected by inflation. However, as we will see in section 3, this relic density is proportional to a positive power of f_a . Therefore, SMM with an $\mathcal{O}(1)$ misalignment angle tends to overproduce DM for very large f_a and underproduce DM for very low values of f_a . Topological defects, i.e. domain walls and cosmic strings, may increase the produced DM abundance in some scenarios [17–19]. The decay of such defects may increase the axion relic abundance by a factor of a few relative to the prediction from the SMM, although the exact impact of axions from topological defects is the topic of ongoing debate, e.g., see [20–26].

Unfortunately, the low- f_a regime, which cannot be accounted for by the standard misalignment mechanism, is exactly the regime which is most experimentally accessible. Due to axions being a particularly well-motivated class of particles, an impressive experimental effort has been directed

towards their discovery. Many of these experiments are specifically trying to detect axions in the dark matter wind. The suppression of $1/f_a$ in axion-SM couplings implies that any direct detection scheme would be most sensitive in the regime where axion DM would be underproduced by SMM if QCD-axion-like interactions are assumed.

This disparity between experimental accessibility and theoretical prediction motivates us to go beyond the standard paradigm. One such potential avenue was independently presented by Co et al. [27] and Chang and Cui [28]. These authors proposed the kinetic misalignment mechanism (KMM), which is the topic of this thesis. The kinetic misalignment mechanism goes beyond the standard paradigm by considering an axion which has initial conditions that endow it with sufficient kinetic energy to overcome the potential barriers and rotate in field space. In the original KMM proposal, such rotation was assumed to persist until the field is eventually trapped when the kinetic energy becomes insufficient to overcome the potential barriers. After trapping, the homogenous field oscillates and behaves as cold dark matter. Because trapping can take place much later than the time where oscillations would start in SMM, kinetic misalignment enhances the relic density relative to SMM by reducing the amount of redshift.

Whereas the original KMM proposal focused on the homogenous component of the field, Fonseca et al. [29] realized that such a rotating field will experience parametric resonance in a phenomenon known as axion fragmentation. In our recent project [1], we studied this effect in the context of KMM, which resulted in the work that is presented in Part II of this thesis. We show that this effect indeed is efficient across most of the KMM parameter space and that it can take place either before or after trapping. We find that fragmentation does not significantly alter the produced DM abundance. However, fragmentation does significantly modify the axion spectrum as the particles are much hotter than in the non-fragmented misalignment scenario. This can lead to observational signatures in the form of axion DM mini-clusters [30]. We also study the potential for signatures in the form of gravitational waves (GWs) generated by the parametric resonance. Furthermore, we identify implementation-independent constraints from structure formation and BBN. Finally, we find that given appropriate initial conditions, KMM, with or without fragmentation, can account for the observed DM relic also in the experimentally accessible low- f_a regime.

With the prospect of KMM providing experiments with the possibility of detecting axion dark matter, it is of paramount importance to investigate whether the necessary initial conditions assumed by KMM can be realized. Part III is devoted to this question, which we investigated in [2]. We re-evaluate implementations inspired by Affleck-Dine baryogenesis [31, 32] which were proposed by Co et al. [27, 33–35]. We impose new constraints identified in our initial paper [1], which significantly impact implementations. We specify solutions directly in terms of model parameters, which allows us to show for the first time how the axion DM parameter space is impacted by implementation constraints. Furthermore, radial oscillations present a significant challenge to KMM implementations. Such oscillations must be damped through some additional mechanism since they would dominate the cosmology if left unchecked. We present a detailed treatment of damping both in general terms and with specific axion-SM couplings. We implement a numerical solution of the Boltzmann equations that allows us to take into account thermal effects and the complex interplay between the axion and the SM sector in a robust way. This solution also allows us to calculate the potential range of damping temperatures. Furthermore, the dynamics of KMM implementations may involve a period where the kinetic energy of the rotating axion dominates the energy density of the universe. Such a period, known as *kination*, may lead to a signal from amplification of the primordial GW spectrum [36, 37]. We map out the region that may be observable by such an amplified GW spectrum. We find that all of the KMM parameter space can be supported, and we identify the model parameters which are required to achieve this.

The structure of the thesis is as follows: In part I, we introduce the strong CP problem and the QCD axion solution. We then generalize to ALPs, discuss their production as dark matter by

the standard misalignment mechanism (SMM) and conclude the section with an overview of the experimental effort to detect axions. In part II, we present our implementation-independent study of fragmentation in KMM. In this part, we develop an analytic description of fragmentation, categorize the parameter space as to the efficiency of fragmentation and present implementation-independent constraints on the axion DM parameter space. In part III, we discuss our re-evaluation of Affleck-Dine-like models of KMM as implemented in two realizations: nearly-quadratic and quartic. For each of the two models, we present constraints and an analysis of the required damping mechanism. In part IIII, we collect results and summarize the thesis. To assist the reader, appendix [A](#) includes a summary of subscripts, symbols and abbreviations used throughout the thesis.

Contents

I	Motivations for axions	12
1	CP violation in QCD and the θ-vacuum	12
1.1	The missing meson problem	12
1.2	Instantons and the θ -vacuum of QCD	13
1.3	CP violation in physical QCD	15
1.4	The strong CP problem	16
2	The Peccei-Quinn solution and the axion	16
2.1	Axion mass and interactions	17
2.2	KSVZ models	18
2.3	Other axion models	19
3	Beyond the QCD Axions: Generic ALPs, dark matter and searches	20
3.1	Standard misalignment and axion dark matter	20
3.2	Axion searches	23
4	Summary of part I	24
II	Axion kinetic misalignment	28
5	Axion dark matter from kinetic misalignment	28
5.1	Derivation of the relic density today from the adiabatic invariant	29
5.2	Trapping temperature	32
6	Analytical theory of parametric resonance in kinetic misalignment	32
6.1	Parametric resonance	32
6.2	An estimate of the back-reaction	37
6.3	Initial conditions for the mode functions	39
7	ALP dark matter from fragmentation	41
7.1	Overview of the fragmentation regions	41
7.1.1	Boundary between regions 1 and 2:	43
7.1.2	Boundary between regions 2 and 3:	43
7.1.3	Boundary between regions 3 and 4:	44
7.1.4	Breakdown of perturbativity for larger m_*/H_*	45
7.2	ALP relic density with fragmentation	46
8	Constraints on the ALP parameter space due to fragmentation	48
8.1	Structure formation	48
8.2	Big Bang Nucleosynthesis (BBN)	48
8.3	ALP parameter space with existing and future experiments	49
9	Gravitational waves	50
10	Summary of part II	55

III	Model implementations of axion kinetic misalignment	57
11	Essential components of a KMM implementation	58
11.1	Definitions for the full complex scalar field	58
11.2	Starting the rotation	58
11.3	Evolution after the kick	59
11.4	Constraints	61
12	Models with nearly-quadratic potentials	64
12.1	Initial conditions	64
12.2	Application of constraints	66
12.3	Early damping	67
12.4	General damping	68
13	Thermal damping with Yukawa interactions	75
13.1	Thermal effects and damping rates	75
13.2	Boltzmann equations and solution approach	76
13.3	Discussion	76
13.4	Results	78
14	Models with quartic potentials	85
14.1	Initial conditions	85
14.2	Application of constraints	86
14.3	Early damping	90
14.4	General damping	92
14.5	Obstacles to thermal damping	95
15	Summary of part III	96
IV	Conclusions and future perspectives	98
	Appendix	101
A	Table of subscripts, abbreviations and symbols	101
B	Detailed discussion of the parametric resonance	104
C	Thermal relics	108
D	Scalar field dynamics in a thermal environment	111
E	Restrictions on the inflaton reheating temperature	114
F	Boltzmann equations and solutions of thermal damping	117
G	Additional details for nearly-quadratic models	123
H	Additional details for quartic models	129

Part I

Motivations for axions

We first review the paradigm of QCD axions and axion-like-particles. We start by introducing the strong CP problem and the Peccei-Quinn solution. We then discuss the properties of the QCD axion in a benchmark KSVZ model. Finally, we motivate the generalization to the broader class of axion-like-particles and discuss how they can play the role of dark matter when produced by the standard misalignment mechanism. We conclude with an overview of the observational searches for axion-like-particles and a summary.

1 CP violation in QCD and the θ -vacuum

In this section, we review the origin of CP violation in QCD and how it follows from the non-trivial vacuum of QCD. We take the historical approach and start by motivating the existence of the problem by first discussing the missing meson problem. This leads us to consider the θ -vacuum and the presence of a θ -term in QCD. Finally, we discuss how the θ -term leads to the strong CP problem.

1.1 The missing meson problem

Let us consider the QCD Lagrangian including the first generation of quarks:

$$\mathcal{L}_{QCD} \supset -\frac{1}{4}G^{a\mu\nu}G_{\mu\nu}^a + \bar{Q}(i\not{D} - M)Q, \quad (1.1)$$

where $Q = (u \ d)^T$ is a vector containing the quark fields, M is the corresponding mass matrix and $G_{\mu\nu}^a$ is the gluon field strength tensor. In the limit where the masses can be neglected, which is a good approximation at least for the first generation of quarks, this Lagrangian possesses a $U(2)_L \times U(2)_R$ symmetry, which corresponds to a unitary rotation of the left- and right-handed components of the quark fields, respectively. At energies below Λ_{QCD} , this classical symmetry is spontaneously broken down to its vectorial subgroup by the QCD condensate with the following breaking pattern:

$$SU(2)_L \times SU(2)_R \times U(1)_V \times U(1)_A \rightarrow SU(2)_V \times U(1)_V, \quad (1.2)$$

where the subscripts A and V refer to vectorial and axial groups, respectively. The axial part of $SU(2)_L \times SU(2)_R$ as well as $U(1)_A$ are spontaneously broken. If this is the true breaking pattern of the theory, then the spontaneous breaking of the three generators of $SU(2)_A$ ¹ as well as the one generator of $U(1)_A$ should lead to a total of four pseudo-Nambu-Goldstone bosons (pNGBs) [38–40]. However, in the observed spectrum of hadrons, the only pNGB candidates are the three pions. The absence of a fourth light pNGB in the meson spectrum constitutes the missing meson problem, which was one of the major challenges for QCD until the mid 70's [41, 42].

This problem was resolved when 't Hooft realized [43–45] that the Adler-Bell-Jackiw (ABJ) anomaly [46–48] explicitly reduces the physical symmetry of QCD. This anomaly affects the approximate $U(1)_A$ symmetry, which is associated with the axial rotation,

$$Q \rightarrow e^{i\alpha\gamma_5}Q. \quad (1.3)$$

¹Note that $SU(2)_A$ is not closed, and as such it does not form a group. With this complication in mind, we will nonetheless associate with $SU(2)_A$ the axial part of $SU(2)_L \times SU(2)_R$.

In the absence of quantum effects, the current related to this transformation would be broken only by the mass term. However, even in the massless limit, the ABJ anomaly gives the current of such an axial transformation a divergence of

$$\partial_\mu J_{U(1)_A}^\mu = \frac{g_s^2}{16\pi^2} G\tilde{G}. \quad (1.4)$$

Therefore, $U(1)_A$ does not correspond to a true approximate symmetry of the physical theory and only the three generators of $SU(2)_A$ correspond to the broken generators of an approximate symmetry. Here $\tilde{G}_{\mu\nu}^a = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}G^{a\rho\sigma}$ is the dual of the field strength tensor, g_s is the QCD gauge coupling, and we have omitted contracted indices. This effect reduces the expected number of pNGBs from four to three, which resolves the missing meson problem.

One might object that $G\tilde{G}$ is itself a total derivative,

$$G^{a\mu\nu}\tilde{G}_{\mu\nu}^a = \partial_\mu K^\mu, \quad (1.5)$$

where K^μ is the Chern-Simons current

$$K^\mu = \epsilon^{\mu\alpha\beta\gamma} \left(A_\alpha^a G_{\beta\gamma}^a - \frac{g_s}{3} f^{abc} A_\alpha^a A_\beta^b A_\gamma^c \right), \quad (1.6)$$

such that one might construct a new conserved current from $J^\mu = J_{U(1)_A}^\mu - K^\mu$. If such a new conserved current could be constructed, then the associated symmetry would perpetuate the missing meson problem. However, in order for the charge to be conserved, the surface integral of the current has to vanish at infinity. As we will see in the next section, it turns out that there exist gauge field configurations for which K^μ does not vanish at infinity. Therefore, $U(1)_A$ is indeed explicitly broken by QCD effects. This solution resolves the missing meson problem, but it also proves that $G\tilde{G}$ has an important physical impact which is related to non-trivial field configurations.

1.2 Instantons and the θ -vacuum of QCD

To see how the $G\tilde{G}$ term impacts QCD, we will now discuss the QCD vacuum. To discuss the vacuum structure of QCD, we can apply semi-classical approximations [49]. Such a semi-classical approach allows one to study the vacuum of gauge field theories by linking it to properties of the corresponding Euclidean action. In order for such an approach to yield non-trivial results, we need to specifically consider configurations of finite Euclidean action, as solutions of infinite action do not contribute in the semi-classical approach [49]. We will simplify the problem by neglecting matter fields for now and studying a pure Yang-mills theory of $SU(3)$ for which the Euclidean action is

$$S_E = \int d^4x \frac{1}{4} GG. \quad (1.7)$$

In order for the action to converge to a finite value, $G^{a\mu\nu}$ must decay at least as fast as $G^{a\mu\nu} \sim \mathcal{O}(1/r^{2+\epsilon})$ when $r \rightarrow \infty$. Therefore, the part of A_μ^a which contributes to $G^{a\mu\nu}$ must decay at least as fast as $A_\mu^a \sim \mathcal{O}(1/r^{2+\epsilon})$. However, by performing a gauge transformation $g(x) \in SU(3)$, we see that there may be an additional component of A_μ^a which need not vanish at the boundary because it does not contribute to $G^{a\mu\nu}$. In terms of $A_\mu = T^a A_\mu^a$, where T^a are the generators of $SU(3)$,

$$\lim_{|x| \rightarrow \infty} A'_\mu = ig(x)\partial_\mu g(x)^{-1} + \mathcal{O}(1/r^2). \quad (1.8)$$

This asymptotic field configuration is referred to as *pure gauge*, because the non-vanishing component is specified entirely by the gauge transformation $g(x)$ and how this maps to the S^3 sphere at infinity. One might wonder if such configurations are all related by a gauge transformation to the trivial solution. It turns out that there are additional classes of solutions which cannot be related

to the trivial solution by any continuous gauge transformation. The field configurations fall into different homotopy classes. It has been shown by Bott [50] that a mapping from S^3 to any simple Lie group G can be continuously deformed to a mapping from S^3 to an $SU(2)$ subgroup of G . Therefore, the field configurations of $SU(3)$ can be classified according to the homotopy classes of $SU(2)$. These classes are identified by $\nu \in \mathbb{Z}$, which is known as the winding number and which counts how many times the configuration in field space wraps around the S^3 hypersphere at infinity [49]. ν is a topological quantity which depends only on the mapping of $g(x) \rightarrow S^3$ at infinity. Furthermore, it can be shown [49] that the winding number is given by the integral

$$\nu = \frac{g_s^2}{32\pi^2} \int d^4x G\tilde{G}. \quad (1.9)$$

These field configurations then have a finite action but are topologically distinct. If we express the Euclidean action in terms of $G\tilde{G}$,

$$S_E = \frac{1}{4} \int d^4x GG = \frac{1}{4} \int d^4x \left[G\tilde{G} - \frac{1}{2}(G - \tilde{G})^2 \right], \quad (1.10)$$

we see that the action Euclidean contains a topological piece,

$$S_E = \nu \frac{8\pi^2}{g_s^2} + \frac{1}{8} \int d^4x (G - \tilde{G})^2. \quad (1.11)$$

Since the second term is positive definite, configurations for which $G = \tilde{G}$ minimize the action for all configurations with a given winding number ν and thus correspond to classical solutions [51]. Additional solutions with $-\nu$ and $G = -\tilde{G}$ can also be found by instead separating out a $(G + \tilde{G})^2$ piece in eq. (1.11). These solutions are known as *instantons* and *anti-instantons* respectively.

Returning to physical Minkowski space, families of classical vacua can be found. Such classical vacua are configurations in which the hamiltonian \mathcal{H} vanishes everywhere in space and are thus pure gauge configurations. As the winding number is a pure topological quantity which does not depend on the metric, it is also defined for gauge field configurations in the Minkowski space, in which it labels distinct classical vacua. We denote a classical vacuum of winding number n as $|n\rangle$. It can be shown [52] that in Minkowski space a ν -instanton corresponds to tunnelling between an $|n\rangle$ classical vacuum and the corresponding $|n + \nu\rangle$ vacuum. As classical vacua are thus subject to time evolution, they cannot correspond to a physical vacuum. Instead, the physical minimum must be a linear superposition of the classical vacua:

$$\sum_{n \in \mathbb{Z}} a_n |n\rangle. \quad (1.12)$$

There exist gauge transformations T_m which change the winding number of all classical vacua by a common integer value m , i.e. $T_m |n\rangle = |n + m\rangle$. Therefore, the difference in winding number $n - m$ between two states $|n\rangle$ and $|m\rangle$ is gauge invariant, but the winding numbers of any individual state are not gauge invariant. We can restrict the form of a_n in eq. (1.12) by demanding that such gauge transformations T_m at most change the vacuum by the complex phase $e^{i\theta}$ up to which a state is defined in projective Hilbert space. This condition yields a recursive relation that restricts the form of a_n such that the physical vacuum $|\theta\rangle$ can be written as [49]

$$|\theta\rangle = \sum_{n \in \mathbb{Z}} e^{i\theta} |n\rangle. \quad (1.13)$$

Therefore, the QCD-vacuum is characterized by a phase θ which determines the superposition of classical vacua of which it is composed. This is the θ -vacuum of QCD.

Interestingly, the θ -vacuum is subject to a super-selection rule, which prevents any physical process from transitioning from a $|\theta\rangle$ state to a $|\theta'\rangle$ state. Therefore, θ is unchangeable in any one realization of QCD, and different θ parameters label physically distinct theories [49].

To see how the θ -vacuum impacts the Lagrangian, we can consider the vacuum expectation value (VEV) of an operator \mathcal{O} . Using the property that gauge transformations can relabel states $|m\rangle, |n\rangle$ to $|k\rangle, |0\rangle$, this can be written as

$$\langle\theta|\mathcal{O}|\theta\rangle = \sum_{n,m} \langle m|e^{i\theta(m-n)}\mathcal{O}|n\rangle = \sum_{n,m} \langle k|e^{i\theta k}\mathcal{O}|0\rangle. \quad (1.14)$$

Expressing the winding number in terms of eq. (1.9) this becomes

$$\langle\theta|\mathcal{O}|\theta\rangle = \frac{1}{Z} \sum_k \int [dA] \mathcal{O} e^{i\left(\int d^4x \mathcal{L} + \theta \frac{g_s^2}{32\pi^2} G\tilde{G}\right)} \delta\left(k - \frac{g_s^2}{32\pi^2} \int d^4x G\tilde{G}\right), \quad (1.15)$$

where δ is a Dirac delta function which restricts the path integral to configurations with winding number k . From this we observe that the phase factor $e^{i\theta}$ in the θ -vacuum extends the Lagrangian with a $\theta G\tilde{G}$ term. The θ -vacuum implies a Lagrangian term of

$$\mathcal{L}_\theta = \theta \frac{g_s^2}{32\pi^2} G\tilde{G}. \quad (1.16)$$

This term is usually referred to as the θ -term. We have now shown² that not only is it possible to add a $\theta G\tilde{G}$ term to the QCD Lagrangian but that such a term is actually implied by the non-trivial vacuum structure of the theory.

1.3 CP violation in physical QCD

In the previous section, we showed that in a pure Yang-Mills theory of $SU(3)$ the non-trivial vacuum structure implies the existence of a $\theta G\tilde{G}$ term. Let us now consider the full QCD Lagrangian, including matter fields:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G^{a\mu\nu} G_{\mu\nu}^a + \tilde{\theta}_{\text{QCD}} \frac{g_s^2}{32\pi^2} G^{a\mu\nu} \tilde{G}_{\mu\nu}^a + \bar{Q}(i\not{D} - M)Q, \quad (1.17)$$

where, as before, the Q is a vector containing the quark fields, and M is the quark mass matrix. Both the θ -term and the global phase of the mass matrix break CP. Furthermore, CP is broken in the weak sector of the SM, such that there is no symmetry to ensure that either term must vanish. Therefore we should expect CP to also be violated by QCD.

Naively, this Lagrangian suggests that there are two independent sources of CP violation in QCD. However, these contributions are not independent. Because the $U(1)_A$ symmetry associated with the axial transformation $Q \rightarrow e^{i\gamma_5\alpha}Q$ is broken by both the ABJ anomaly and by the quark mass term both parameters can be shifted by performing such an axial transformation. Therefore, one of the two CP-violating parameters can be absorbed by performing a suitable axial rotation of the quark fields such that the single, physically observable CP-violating parameter is

$$\theta_{\text{QCD}} = \tilde{\theta}_{\text{QCD}} + \arg(\det M), \quad (1.18)$$

where $\arg(\det M)$ is the complex phase of the determinant of the quark mass matrix. Thus, the ABJ anomaly of $U(1)_A$ reduced what appeared as two parameters to only one CP-violating parameter. We may choose to remove any common phase from the quark mass matrix such that all potential CP violation is carried by the θ -term.

²The argument presented here was based on semiclassical reasoning which breaks down as $g_s \rightarrow \infty$. Although we can learn much about QCD from such reasoning it is not useful to describe QCD at $T < \Lambda_{\text{QCD}}$. A more general description is beyond the scope of this review.

1.4 The strong CP problem

Currently, the most sensitive probe of CP violation in QCD is the neutron electric dipole moment (nEDM). The strength of this moment is measured in terms of d_n as defined by the non-relativistic Hamiltonian

$$H = -d_n \vec{E} \cdot \hat{S}. \quad (1.19)$$

This quantity is currently constrained by experiment to less than $|d_n| \lesssim 10^{-26} e \text{ cm}$ [53]. The contribution of θ_{QCD} to d_n has been performed with a number of theoretical approaches, see [54] for a review. One result, based on QCD sum-rules yields

$$d_n \sim 10^{-16} e \text{ cm} \times \theta_{\text{QCD}}, \quad (1.20)$$

which corresponds to an upper bound on θ_{QCD} of

$$|\theta_{\text{QCD}}| \lesssim 10^{-10}. \quad (1.21)$$

The smallness of this parameter constitutes the strong CP problem.

The strong CP problem is quite different from that of the Higgs mass hierarchy problem in that θ_{QCD} is stable against radiative corrections. One could expect large contributions to θ_{QCD} since CP is violated in the SM, but Ellis and Gaillard[55] showed that renormalization of θ_{QCD} contributes only at the level of $\delta\theta_{\text{QCD}} \sim 10^{-16}$, which is several orders of magnitude below the level probed by neutron EDM. Therefore, running of θ_{QCD} would preserve the smallness of θ_{QCD} if it is set to 0 in the UV. Such solutions where CP is a symmetry of the UV theory and which is then spontaneously broken include the Nelson-Barr models. However, such models may require somewhat exotic model building or may themselves be fine-tuned unless supersymmetry is invoked [56].

Furthermore, a much larger value of θ_{QCD} does not appear to cause any catastrophic issues which could motivate anthropic solutions. Ubaldi found that nuclear physics and BBN tolerates values θ_{QCD} up to about 1% [57], such that any attempt to resolve the CP problem anthropologically would have to link θ_{QCD} to some other small parameter [58, 59]. Therefore, the anthropic principle might not provide the most elegant solution to the strong CP problem.

2 The Peccei-Quinn solution and the axion

In section 1.3 we saw how the ABJ anomaly of the axial rotation allowed us to render one of the two possible sources of CP violation unphysical. Peccei and Quinn exploited this phenomenon by introducing a new global axial $U(1)$ symmetry to render the CP violation in QCD fully unphysical [3, 4]. This is achieved by having the new symmetry, usually denoted $U(1)_{\text{PQ}}$, be an exact symmetry of the classical theory that is only broken by an ABJ anomaly with QCD in the full theory. This QCD anomaly allows one to perform a $U(1)_{\text{PQ}}$ transformation to shift the Lagrangian by

$$\delta\mathcal{L} = \alpha \frac{g_s^2}{32\pi^2} G\tilde{G}, \quad (2.1)$$

where α is an arbitrary angle, such that the θ -term of the QCD vacuum can be fully absorbed. This renders CP violation in QCD unphysical.

A necessary component of this solution is that there exist fermions without a PQ-violating mass term. However, no massless fermions are observed in nature, so the PQ-charged fermions must nonetheless acquire a mass. Therefore, the PQ symmetry is generally assumed to be spontaneously broken at a scale f_a , which generates a PQ-conserving mass term for the relevant fermions.

Weinberg [15] and Wilczek [16] pointed out that this generically leads to the prediction of a pNGB, which was named *the axion* by the latter³. Usually, this axion field is denoted by a .

As a pNGB, the axion in the absence of the QCD anomaly would feature a shift symmetry $a \rightarrow a + \alpha f_a$. This shift symmetry is broken down to a discrete symmetry by the QCD anomaly. Because of the ABJ anomaly of $U(1)_{\text{PQ}}$, the axion necessarily features a QCD-coupling,

$$\mathcal{L} \supset \frac{a}{f_a} \frac{g_s^2}{32\pi^2} G\tilde{G}. \quad (2.2)$$

This is the sole coupling between the axion and the SM that the axion must feature, although specific implementations of the PQ solution may introduce additional interactions.

If this axion has a VEV away from the CP conserving value, then it would regenerate the CP violating effects that it was introduced to solve. Fortunately, it was shown by Vafa and Witten that vector-like theories such as QCD cannot break parity spontaneously [60] such that the minimum of the axion potential must conserve CP.

2.1 Axion mass and interactions

At low energies, the properties of an axion can be described with an EFT specified by the following Lagrangian:

$$\mathcal{L}_a = \frac{1}{2} (\partial_\mu a)^2 + \frac{a}{f_a} \frac{g_s^2}{32\pi^2} G\tilde{G} + \frac{1}{4} g_{a\gamma} a F\tilde{F} + \sum_{N \in \{n,p\}} \frac{1}{2} \frac{\partial_\mu a}{f_a} c_N \bar{N} \gamma^\mu \gamma_5 N, \quad (2.3)$$

where $F^{\mu\nu}$ is the electromagnetic fields strength tensor and N is either the proton p or the neutron n . $g_{a\gamma}$ and c_N respectively parametrize axion-photon and axion-nucleon couplings. We will now discuss how the generic axion-gluon coupling, eq. (2.2), leads to such couplings between the axion and the SM. At energies below Λ_{QCD} it is convenient to absorb the $aG\tilde{G}$ term by redefining the quark fields as

$$Q \rightarrow e^{i\gamma_5 \frac{a}{2f_a} Q_a} Q, \quad (2.4)$$

where Q_a is a matrix specifying the rotation. Due to the anomaly, the axion will mix with the neutral mesons, e.g., see [14]. At energies much lower than Λ_{QCD} , this mixing can be described by chiral perturbation theory. In this limit, specifying Q_a such that the axion-meson mixing matrix is diagonal results in a physical axion state with a zero-temperature mass of [14],

$$m_a^2 \approx \frac{m_u m_d}{(m_u + m_d)^2} \frac{m_\pi^2 f_\pi^2}{f_a^2} \approx \left[(5.70 \pm 0.07) \times 10^{-6} \text{ eV} \left(\frac{10^{12} \text{ GeV}}{f_a} \right) \right]^2. \quad (2.5)$$

Furthermore, the axion-meson mixing induces model-independent contributions to the interactions [14]. The model-independent photon coupling is

$$g_{a\gamma, \text{indep.}} = \frac{\alpha_{\text{EM}}}{2\pi f_a} \left[\frac{2}{3} \frac{4m_d + m_u}{m_d + m_u} \right] \approx -\frac{\alpha_{\text{EM}}}{2\pi} \frac{1.92 \pm 0.04}{f_a} \approx -\frac{(2.227 \pm 0.044) \times 10^{-3}}{f_a}, \quad (2.6)$$

where $\alpha_{\text{EM}} \approx 1/137$ is the electromagnetic fine structure constant. The model-independent proton and neutron couplings are

$$c_{p, \text{indep.}} \approx 0.47 \pm 0.03 \quad \text{and} \quad c_{n, \text{indep.}} \approx -0.02 \pm 0.03. \quad (2.7)$$

³Humorously, Wilczek choose the name "axion" in honour of an American laundry detergent as it "cleans up" the strong CP problem.

Specific axion implementations may introduce additional interactions for the axion, which will modify these model-independent contributions, but these couplings will be shared by all axion models in which the axion couples to the SM exclusively through eq. (2.2).

The QCD axion mass is temperature dependent and only approaches the zero-temperature result eq. (2.5) in the limit $T \ll \Lambda_{\text{QCD}}$. In the regime $T \gg \Lambda_{\text{QCD}}$, where the temperature dependence is significant, chiral perturbation theory is no longer a useful description. Instead, the axion mass can be calculated in the *dilute instanton gas model* [61], the *interacting instanton liquid model* [62] or on the lattice [63–67]. To describe $m_a(T)$ we will fit the lattice result of [67] with

$$m_a^2(T) \approx m_a^2 \times \begin{cases} (T/T_c)^{-\gamma} & \text{if } T > T_c \\ 1 & \text{if } T < T_c \end{cases}, \quad \text{where} \quad \begin{aligned} T_c &= 2.12 \times \sqrt{f_a m_a}, \\ \gamma &= 8.16, \end{aligned} \quad (2.8)$$

which sets the amplitude of a cosine potential for the axion:

$$V_a(T) \approx m_a(T)^2 f_a^2 [1 - \cos(\theta)]. \quad (2.9)$$

This provides a good parametrization of the lattice result of [67].

We define our notation such that m_a when written without T dependence always refers to the zero-temperature result given by eq. (2.5). The full fit given by eq. (2.8) is only referenced when the T dependence is explicitly written in $m_a(T)$.

2.2 KSVZ models

One of the most popular classes of invisible axions are the Kim-Shifman-Vainshtein-Zakharov (KSVZ) models [68, 69]. These realize the PQ mechanism by extending the SM Lagrangian with a complex scalar field P as well as at least one new coloured fermion χ . These fields are then assumed to possess a Yukawa coupling such that the KSVZ Lagrangian takes the form

$$\mathcal{L}_{\text{KSVZ}} = |\partial_\mu P|^2 - V(P) + \bar{\chi} i \not{D} \chi - \sqrt{2} y \bar{\chi}_L \chi_R P + h.c., \quad (2.10)$$

where $V(P)$ is the potential of P . This minimal Lagrangian is invariant under an axial transformation specified by

$$P \rightarrow e^{i\alpha} P, \quad \chi_L \rightarrow e^{i\alpha/2} \chi_L, \quad \text{and} \quad \chi_R \rightarrow e^{-i\alpha/2} \chi_R, \quad (2.11)$$

which plays the role of $U(1)_{\text{PQ}}$. Different representations of these fields are possible. The minimal requirement of the KSVZ solution is that the χ -fermions must be in a non-trivial representation of $SU(3)_{\text{QCD}}$, such that they are subject $U(1)_{\text{PQ}}$ to an ABJ anomaly from QCD. The potential $V(P)$ is chosen such that $U(1)_{\text{PQ}}$ is spontaneously broken when the scalar field acquires a VEV of $|P| = f_a/\sqrt{2}$. The axion arises as the angular mode a ,

$$P = \frac{1}{\sqrt{2}} \phi e^{ia/f_a} \quad (2.12)$$

which becomes relevant when ϕ acquires a VEV of f_a . When ϕ acquires a VEV the Yukawa couplings generates a χ -fermion mass of $m_\chi = y f_a$, which makes them heavy unless y is very small. We typically take $y \sim \mathcal{O}(1)$, such that $m_\chi \sim f_a$ and the χ -fermions decouple from physics below the f_a scale. To integrate out the fields, we have to remove the light angular degree of freedom from the mass terms,

$$\sqrt{2} y \bar{\chi}_L \chi_R P \approx m_\chi \bar{\chi}_L \chi_R e^{ia/f_a}. \quad (2.13)$$

The axion can be decoupled by performing an axial redefinition of the χ -fermion fields:

$$\chi \rightarrow e^{i\gamma_5 \frac{a}{2f_a}} \chi \quad (2.14)$$

Due to the ABJ anomaly, this rotation generates the expected $aG\tilde{G}$ term,

$$\mathcal{L} \supset \frac{g_s^2}{32\pi^2} \frac{a}{f_a} G\tilde{G}. \quad (2.15)$$

If the χ -fermions are chosen to be in the minimal representation $(3, 1, 0)$ of $SU(3)_{\text{QCD}} \times SU(2)_L \times U(1)_Y$, then eq. (2.15) is the only direct coupling between the KSVZ axion and the SM. In such a minimal KSVZ model, the axion-photon and the axion-nucleon couplings have only the model-independent contributions given by eqs. (2.6) and (2.7). As a KSVZ model of this type is a minimal implementation of the PQ solution, we will use this as our benchmark model.

2.3 Other axion models

Historically, the original axion implementation was the Peccei-Quinn-Weinberg-Wilczek (PQWW) model [3, 15, 16]. In the PQWW model, the PQ symmetry is implemented in a two Higgs doublet model. In PQWW, one Higgs gives mass to the d-type quarks, and the other gives mass to the u-type quarks. Like the SM Higgs, this scenario requires the PQ field to have a potential which spontaneously breaks $U(1)_{\text{PQ}}$ at $f_a \sim 250$ GeV. However, $f_a \sim 250$ GeV does not sufficiently suppress the axion-SM couplings such that the PQWW model was excluded by a combination of beam dump experiments and rare π and K meson decays [70].

For axion models to be compatible with current constraints, much larger values of f_a are required. Such models are termed *invisible* axion models because the large decay constants suppress axion-SM couplings such that the axions are hidden from searches. The two most widespread families of invisible axion models are the Kim-Shifman-Vainshtein-Zakharov (KSVZ) models and the Dine-Fischler-Srednicki-Zhitnitsky (DFSZ) models [71, 72], of which the former were discussed above. DFSZ models extend the two-Higgs scenario of the PQWW model with an SM-singlet complex scalar Φ . This allows the model to decouple f_a from the electroweak scale by letting Φ acquire a VEV much larger than the VEV of the two Higgs doublets. These two Higgs fields, $H_u \sim (1, 2, -1/2)$ and $H_d \sim (1, 2, +1/2)$ are then connected to Φ by an interaction term of the form

$$V_{\text{DFSZ}} \supset \xi H_u H_d |\Phi|^2. \quad (2.16)$$

The physical axion state a is a linear superposition between the angular mode of Φ and components of H_d and H_u , which is dominated by the angular mode of Φ . Because the SM quarks in DFSZ models must carry PQ charge, and these are in non-trivial representations of $SU(3)_{\text{QCD}} \times SU(2)_L \times U(1)_Y$, the axion has both color and EM anomalies. Therefore, DFSZ models feature additional couplings which modify the model-independent contributions described in section 2.1. The exact couplings depend on the DFSZ implementation.

Another notable class of axion models are the composite axion models, which were originally introduced by Kim [73] a few years after the KSVZ model was proposed. These composite axion models were originally motivated by the possibility of dynamically generating $f_a \ll m_{\text{Pl}}$ through dimensional transmutation. Here $m_{\text{Pl}} \approx 2.4 \times 10^{18}$ GeV is the reduced Planck mass. This is completely analogous to how Λ_{QCD} is generated by the running of the gauge coupling in QCD. Furthermore, in analogy with the chiral symmetry breaking of QCD, the condensate of the new strong sector spontaneously breaks part of the global symmetry such that the composite axion can be identified with a pNGB state, i.e. a pion of the new confining sector. This class of models is particularly appealing because these models can naturally protect the PQ solution from global symmetry violating effects at the Planck scale. This possibility was originally demonstrated by Randall [74] who proposed a composite axion model which realizes the PQ symmetry as an approximate accidental symmetry. Composite axions with accidental PQ symmetry were more recently studied by, e.g., Redi and Sato [75].

3 Beyond the QCD Axions: Generic ALPs, dark matter and searches

Light pseudoscalars such as the axion play an important role as dark matter candidates [5–7]. Furthermore, the motivation for new light pseudoscalar particles goes beyond the strong CP problem as particles with similar properties are a generic prediction of a broad class of string theories [8, 9] and other BSM frameworks such as majoron models [10], supersymmetric models [11] and dynamical models of flavour [12, 13]. We, therefore, extend our discussion of dark matter to the broader class of axion-like-particles, ALPs, which refers broadly to any pseudoscalar pNGB of a global $U(1)$ spontaneously broken at a scale f_a . Importantly, although such particles are still specified in terms of a mass m_a and a decay constant f_a , they need not conform to the QCD axion mass relation given by eq. (2.5). The ALP parameter space, therefore, corresponds to a larger $[m_a, f_a]$ plane of which the QCD axion is a subset. We will continue to use the symbols m_a, f_a for the ALP parameters. From this point, we will no longer refer to the axion field as a , and we instead use the symbol φ to reflect that our statements apply to ALPs generically. Typically, we will work with the dimensionless field $\theta = \varphi/f_a$. As mentioned earlier, the inclusiveness of the term *axion* depends on the author, and we will define it to refer to the broader class of ALPs. We will specifically refer to an axion which solves the strong CP problem as the QCD axion.

We here discuss how such light pseudoscalars may be produced as dark matter by the misalignment mechanism. Furthermore, we will give a short discussion of the experimental approaches that have been and will be used to search for ALPs.

3.1 Standard misalignment and axion dark matter

To see how axion dark matter might be non-thermally produced in the early universe [76], consider the Lagrangian,

$$\mathcal{L} \supset \frac{1}{2}(\partial_\mu\varphi)^2 - \frac{1}{2}m_a^2(T)\varphi^2. \quad (3.1)$$

Here we used m_a and φ to match the notation to the scenario of an axion, but the argument below applies to any scalar field with a sufficiently slowly changing mass. In an expanding Friedmann–Lemaître–Robertson–Walker (FLRW) universe, such a scalar field will have an equation of motion of

$$\ddot{\varphi} + 3H\dot{\varphi} + m_a^2(T)\varphi = 0, \quad (3.2)$$

where H is the Hubble parameter, overdots indicate derivatives with respect to time, and we have assumed that the field is homogenous so that the gradient term does not enter. The $3H\dot{\varphi}$ term can be traced back to the non-zero Γ_{ij}^0 Christoffel symbols of an FLRW metric and is referred to as the Hubble friction term. The behaviour of this equation can be divided into two distinct limits. In the $m_a(T) \ll H$ limit the field is overdamped such that φ is approximately constant. We in this regime say the field is frozen since any evolution of φ takes place on timescales much longer than the timescale H^{-1} . In the other limit, $m_a(T) \gg H$, the field oscillates. We typically estimate the onset of oscillations to take place at a temperature T_{osc} for which $m_a(T_{\text{osc}}) \approx 3H$. For the case of the QCD axion, the temperature dependence of the axion makes the field nearly massless prior to the QCD phase transition, such that $m_a \sim 3H$ is only satisfied once the axion mass begins to grow near $T \sim \Lambda_{\text{QCD}}$.

Thus, the field undergoes coherent oscillations after T_{osc} . It turns out that such oscillations behave as cold dark matter. To see this, we need to solve the oscillating limit of the EOM. If the change in mass is adiabatic, then this limit of the EOM can be solved by the Wentzel–Kramers–Brillouin (WKB) approximation [77–80]. Specifically, this requires $\dot{m}_a(T)/m_a^2(T) \ll 1$, which is the case of

for a QCD-like scaling of $m_a(T)$. The WKB approximation implies

$$\varphi \propto m_a^{-1/2}(T)a^{-3/2} \cos\left(\int m_a(T) dt\right), \quad (3.3)$$

where a is the scale factor⁴. If we label the initial amplitude φ_{osc} we can then write φ as

$$\begin{aligned} \varphi &\approx \mathcal{A} \cos(\alpha) \\ \text{where } \tilde{\Theta}(a) &= \varphi_{\text{osc}} \left(\frac{m_a(T_{\text{osc}})}{m_a(T)}\right)^{1/2} \left(\frac{a_{\text{osc}}}{a}\right)^{3/2} \quad \text{and} \quad \alpha = \int_{t_{\text{osc}}}^t dt m_a(t). \end{aligned} \quad (3.4)$$

The energy density and pressure of φ is

$$\rho_\theta = \frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}m_a(T)^2\varphi^2 \quad (3.5)$$

$$p_\theta = \frac{1}{2}\dot{\varphi}^2 - \frac{1}{2}m_a(T)^2\varphi^2 \quad (3.6)$$

By expressing these in terms of \mathcal{A} and α and applying double angle identities, we arrive at

$$\rho_\theta = \frac{1}{2}m_a^2(T)\mathcal{A}^2 + \frac{1}{2}\dot{\mathcal{A}}^2 \cos^2(\alpha) - \frac{1}{2}m_a(T)\mathcal{A}\dot{\mathcal{A}} \sin(2\alpha) \quad (3.7)$$

$$p_\theta = -\frac{1}{2}m_a^2(T)\mathcal{A}^2 \cos(2\alpha) + \frac{1}{2}\dot{\mathcal{A}}^2 \cos^2 \alpha - \frac{1}{2}m_a(T)\mathcal{A}\dot{\mathcal{A}} \sin(2\alpha) \quad (3.8)$$

Averaged over many oscillations,

$$\langle \cos^2 \alpha \rangle = \frac{1}{2}, \quad \langle \sin(2\alpha) \rangle = \langle \cos(2\alpha) \rangle = 0, \quad (3.9)$$

such that

$$\langle \rho_\theta \rangle = \frac{1}{2}m_a^2(T)\mathcal{A}^2 + \frac{1}{4}\dot{\mathcal{A}}^2, \quad (3.10)$$

$$\langle p_\theta \rangle = \frac{1}{4}\dot{\mathcal{A}}^2. \quad (3.11)$$

The WKB solution implies that $m_a(T)\mathcal{A} \gg \dot{\mathcal{A}}$. Therefore the equation of state ω is that of a pressureless fluid,

$$\omega = \frac{\langle p_\theta \rangle}{\langle \rho_\theta \rangle} = \frac{\dot{\mathcal{A}}^2}{\dot{\mathcal{A}}^2 + 2m_a^2(T)\mathcal{A}^2} \approx 0, \quad (3.12)$$

which proves that the axion field behaves as cold dark matter in the oscillating regime.

Returning to the dimensionless field $\theta = \varphi/f_a$, which we will mostly use in the rest of this thesis, this implies a cold matter relic

$$\rho_a \approx \frac{1}{2}m_a(T)m_a(T_{\text{osc}})f_a^2\theta_{\text{osc}}^2 \left(\frac{a_{\text{osc}}}{a}\right)^3, \quad (3.13)$$

where ρ_a is the axion energy density. This DM production mechanism is known as the misalignment mechanism. Because the mechanism is inherent to models of light scalars and does not require model building beyond the assumption of rather generic initial conditions, it is one of the most elegant and most often invoked mechanisms for the production of axion dark matter.

⁴Be aware that the symbol a until this point represented the axion field. From this point onwards, a will instead refer to the scale factor unless it is used in a subscript. As the distinction is usually quite clear from context, this should not lead to undue confusion.

The relic produced by misalignment crucially depends on the decay constant f_a and the initial misalignment angle θ_{osc} . Solving the condition $m_a(T_{\text{osc}}) \approx 3H(T_{\text{osc}})$ for T_{osc} predicts a dark matter relic of

$$\rho_{a,\text{today}} \approx 0.22 \rho_{\text{CDM}} \left(\frac{m_a}{\text{1eV}}\right)^{1/2} \left(\frac{m_a}{m_a(T_{\text{osc}})}\right)^{1/2} \left(\frac{f_a}{10^{11} \text{ GeV}}\right)^2 \left(\frac{\theta_{\text{osc}}}{1}\right)^2 \frac{g_*(T_{\text{osc}})^{3/4}}{g_{*s}(T_{\text{osc}})}, \quad (3.14)$$

Where we normalized $\rho_{a,\text{today}}$ to the present day abundance of CDM $\rho_{\text{CDM}} = 1.27 \text{ keV} / (\text{cm}^3 c^2)$ [81]. For a QCD-like temperature dependence of $m_a(T)$ as specified by eq. (2.8) the relic becomes

$$\rho_{a,\text{today}} \approx 40 \rho_{\text{CDM}} \left(\frac{m_a}{\text{1eV}}\right)^{1/2} \left(\frac{f_a}{10^{11} \text{ GeV}}\right)^{1.66} \left(\frac{\theta_{\text{osc}}}{1}\right)^2 \frac{g_*(T_{\text{osc}})^{0.58}}{g_{*s}(T_{\text{osc}})}. \quad (3.15)$$

Pre- and post inflationary scenarios: In the above analysis of the misalignment mechanism, we started the discussion by assuming that the field was homogenous. This is true if the PQ symmetry⁵ of the axion is broken before or during inflation, such that inflation selects a single initial value of θ_{osc} for the entire observable universe. If this angle is chosen randomly, we then expect that

$$\theta_{\text{osc}}(x) \sim \mathcal{O}(1) \quad \text{for pre-inflationary scenarios.} \quad (3.16)$$

Otherwise, if PQ symmetry is broken after inflation, we then expect the universe to consist of many different patches in which θ_{osc} takes different, random values. Nevertheless, we can average over many randomly chosen angles such that the averaged value of θ_{osc} can be described by

$$\langle \theta_{\text{osc}}^2 \rangle \approx \frac{1}{\pi} \int_0^\pi d\theta \theta^2 = \frac{\pi^2}{3} \quad \text{for post-inflationary scenarios.} \quad (3.17)$$

In either case, we can therefore expect θ_{osc} to be an order $\mathcal{O}(1)$ number unless additional dynamics drive a particular choice of initial angles. In such a scenario, the observed dark matter relic is produced by exactly one value of f_a . Therefore, the misalignment mechanism naturally produces the observed dark matter abundance for one specific combination of m_a and f_a . For f_a lower than this value, dark matter tends to be underproduced, while for larger f_a dark matter tends to be overproduced. In the large- f_a regime the overproduction can be compensated by tuning $\theta_{\text{osc}} \rightarrow 0$. To account for axion DM in the low- f_a regime the production must be enhanced relative to the misalignment mechanism.

Topological defects Cosmic strings are topological objects which can form when CP symmetry is broken. Domain walls can later arise when the axion mass further breaks the residual shift symmetry to a discrete subgroup. In pre-inflationary scenarios, such topological defects usually play no significant role because they are inflated away. However, topological defects can play a major role in post-inflationary models where defects are formed after inflation. In particular, decay of such defects may increase the produced DM abundance beyond what is expected from misalignment alone [17–19]. The exact abundance of axions produced by the decay of topological defects is the topic of ongoing debate, but the relic appears to be no larger than a factor of a few times the prediction from misalignment, e.g., see [20–26].

Domain walls can be pathological for theories in which they form because they tend to lead to cosmologies radically different from the one we observe in our universe. We will return to this problem later in this thesis.

⁵We will continue to refer to the $U(1)$ of which the ALP is a pNGB as the *PQ symmetry* even when the ALP may not be the QCD axion.

3.2 Axion searches

Due to their strong theoretical motivation, many experimental approaches have been applied to search for axions. This is a field in continuous growth, and there are many experiments which will constrain the parameter space further in the future. We will here give a brief introduction to the different classes of axion searches. Due to the diverse nature of the field and the little space that can be devoted here, this summary is necessarily incomplete. For a more complete overview, see, e.g., the summary by Irastorza and Redondo [82] or the compilation by Ciaran O’Hare [83]. The summary here serves only to provide the basic orientation within a very active field.

In the following paragraphs, we discuss search strategies as grouped by ALP source:

Haloscopes: If the dark matter of our universe is made up of ALPs, then one of the most straightforward approaches is to attempt to detect this relic directly. As such experiments directly probe the dark matter halo of the Milky Way, they are typically referred to as *haloscopes*. The conventional haloscope strategy was proposed by Sikivie [84] and relies on resonant ALP conversion within microwave cavities. As enhancement is only achieved when the resonant frequency of the cavity closely matches the ALP mass, these experiments must be tuned. As such, haloscopes tend to be very sensitive but only to a narrow range of m_a . A prominent example of such an experiment is ADMX [85–92] which is currently probing QCD axion models in the 3.3–4.2 μeV range.

As the resonant frequencies are intimately tied to the dimensions of the cavities, traditional haloscopes are only practical within a limited ALP mass range. To search for ALPs of higher masses, dielectric haloscopes such as MADMAX [93] or LAMPOST [94] are being developed as well as dish antenna such as BRASS [95]. Lower ALP masses are being explored using LC circuits in experiments such as DM Radio [96, 97] and WISPLC [98].

Helioscopes: It can be challenging to search for DM axions in large part because of the large mass range. Therefore, one of the currently most successful approaches is to search for ALPs produced by Primakoff interactions in the Sun, hence the term *helioscope*. Since the temperatures of the relevant layers of the Sun are much hotter than the relevant ALP masses, these solar axions are relativistic with energy set by the solar dynamics. This property makes helioscope searches largely independent of the ALP mass m_a . The currently most prominent helioscope experiment is CAST [99, 100] which searches for ALP \rightarrow x-ray conversion within a repurposed LHC dipole. IAXO [82, 101, 102] will improve upon CAST as a next-generation helioscope.

Laboratory experiments: Instead of relying on external sources of ALP, another possible approach is to construct experiments which are entirely self-contained. One of the most prominent principles in the class of searches is the *light-shining-through-wall* (LSW) approach. These LSW experiments use powerful lasers and magnetic fields to search for photon \rightarrow ALP \rightarrow photon conversion enabling light to traverse otherwise opaque barriers. This class of experiments includes ALPS [103], CROWS [104] and OSQAR [105], none of which are competitive with the current generation of helioscopes. However, ALPS will soon be followed by its successor ALPSII [106], which expects to outcompete CAST in a large range of m_a . Other types of laboratory experiments include polarization experiments such as PVLAS [107] and fifth-force experiments such as ARIADNE [108].

Astrophysics and cosmology: In addition to the experimental searches described above, a large and diverse class of axion signatures have also been investigated in the realm of astrophysics and cosmology. These principles range from modification of structure formation as probed by the Lyman- α forest (e.g. [109]), conversion of axions in cosmic magnetic fields as observed in

x-rays (e.g. [110–113]), modification of stellar evolution as probed by observations of horizontal branch stars [114] or the impact of ALP thermal relics on CMB and BBN [115].

A more complete overview of experimental searches is shown in figure 1. We show current constraints as filled regions. Projections are distinguished by having outlines only. In order to translate the bounds/projections from $[m_a, \text{coupling}]$ -space to $[m_a, f_a]$ -space, we here and in the rest of this thesis assume a minimal axion coupling, such as the one implemented by a minimal KSVZ model. Such couplings are simply the model-independent contributions from the $\theta G\tilde{G}$ term:

$$g_{\theta\gamma\gamma}^{\text{KSVZ}} = |g_{a\gamma, \text{indep.}}| \approx \frac{2.23 \times 10^{-3}}{f_a}, \quad \text{and} \quad C_n = |C_{n, \text{indep.}}| \approx 0.023 \quad (3.18)$$

To be specific, we assume such coupling both for QCD axions and for non-QCD axions. For non-minimal couplings, such as those found in, e.g., DFSZ models, the constraints and the projections of all the experiments need to be adjusted with the exception of the superradiance bounds. The superradiance bounds depend on the axion mass and the axion decay constant directly, as the cut-off is set by the quartic interaction induced by the axion cosine potential.

In this and in the following figures, we highlight searches which are only sensitive to axion DM in green and depict searches which seek axions of other origins in orange. The reach of these searches is compared to the parameter space in which axion DM is naturally accounted for by SMM. Note that a large population of experiments, chiefly haloscopes, are primarily sensitive to axion DM in the low- f_a regime in which standard misalignment underproduces dark matter.

In table 1 and table 2 the experiments shown on figure 1 are listed along with brief notes on the experimental principles. All experiments rely on the axion-photon coupling unless otherwise indicated. Most of the digitized constraints are sourced from the `AxionLimits` repository [83].

4 Summary of part I

In the last few sections, we have discussed how the QCD axion is motivated as a solution to the strong CP problem and how axions and ALPs are natural dark matter candidates. In section 1, we discussed how the strong CP problem arises from non-perturbative effects and the non-trivial vacuum structure of QCD. In section 2 we discussed how the strong CP problem can be solved by the PQ mechanism and how this leads to the prediction of the QCD axion. We then discussed model-independent properties of the QCD axion as well as a specific implementation in the form of the KSVZ model. In section 3, we generalized to the broader class of ALPs and discussed how a dark matter relic of such axions might be produced by the misalignment mechanism. We also provided an overview of how such axion DM is searched for.

Axions make compelling dark matter candidates. However, as is apparent in figure 1, the standard misalignment mechanism does not account for ALP DM in the low- f_a regime, which is particularly susceptible to experimental scrutiny. To improve upon this unsatisfactory situation, we, in the next parts of this thesis, investigate how ALP DM can be accounted for also in the low- f_a regime by invoking Axion Kinetic Misalignment.

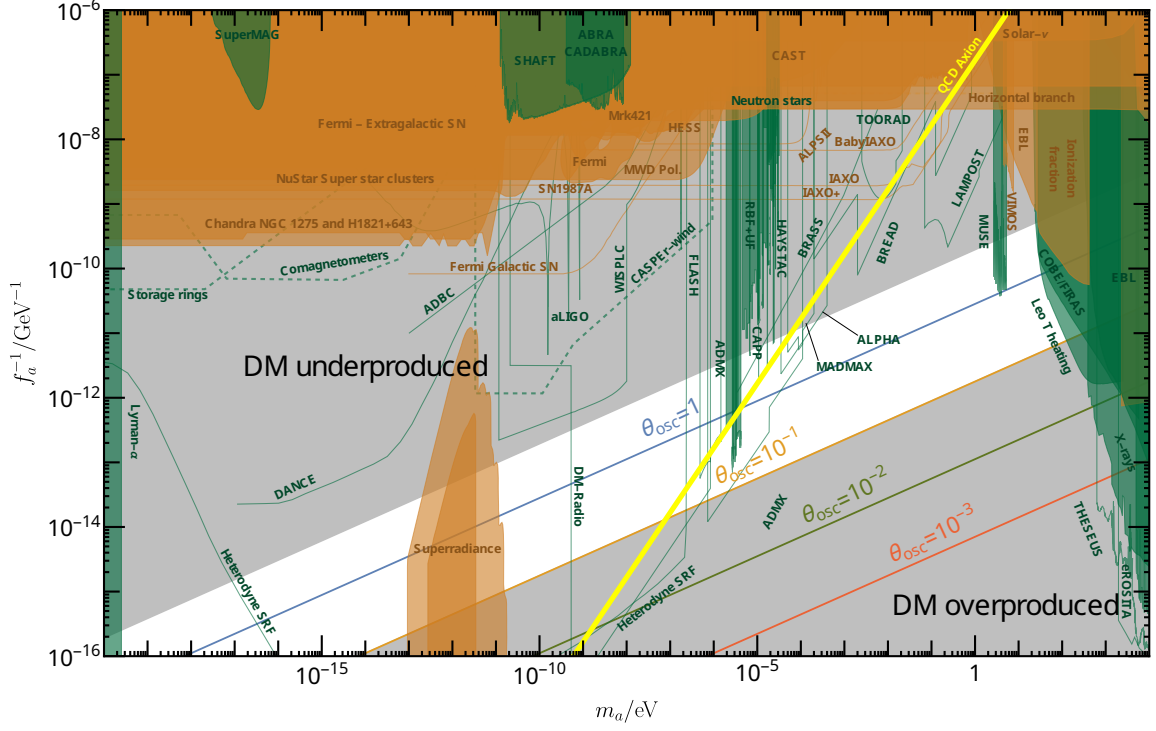


Figure 1. Range of $[m_a, f_a]$ parameter space in which axion dark matter can be realized by the misalignment mechanism. The range in which $\theta_{\text{osc}} \sim \mathcal{O}(1)$ leads to the observed DM relic appears in white, see eq.(3.15). The high- f_a region (in gray) is characterized by DM overproduction unless the misalignment angle θ_{osc} is tuned towards the minimum. The low- f_a regime, also in gray, is characterized by DM underproduction and requires dynamics beyond the standard misalignment mechanism in order to account for axion DM. Experimental axion searches are overlaid in orange and green. Searches coloured green assume the axion to make up all of the observed dark matter.

Experiment:	Principle	DM?	Ref.
<i>Haloscope constraints</i>			
ABRACADABRA-10cm	Haloscope	DM	[116]
ADMX	Haloscope	DM	[85–91]
BASE	Haloscope (Cryogenic Penning Trap)	DM	[117]
CAPP	Haloscope	DM	[118–120]
CAST-RADES	Haloscope	DM	[121]
DANCE	Haloscope (Optical cavity polarization)	DM	[122]
Grenoble Haloscope	Haloscope	DM	[123]
HAYSTAC	Haloscope	DM	[124, 125]
ORGAN	Haloscope	DM	[126]
QUAX	Haloscope	DM	[127, 128]
RBF	Haloscope	DM	[129]
SHAFT	Haloscope	DM	[130]
SuperMAG	Haloscope (Using terrestrial magnetic field)	DM	[131]
UF	Haloscope	DM	[132]
Upload	Haloscope	DM	[133]
<i>Haloscope projections</i>			
ABDC	Haloscope	DM	[134]
ADMX	Haloscope	DM	[92]
aLIGO	Haloscope	DM	[135]
ALPHA	Haloscope (Plasma haloscope)	DM	[136]
BRASS	Haloscope	DM	[95]
BREAD	Haloscope (Parabolic reflector)	DM	[137]
DANCE	Haloscope (Optical cavity polarization)	DM	[138]
DMRadio	Haloscope (All stages: 50L, m^3 and GUT)	DM	[96, 97]
FLASH	Haloscope (Formerly KLASH)	DM	[139, 140]
Heterodyne SRF	Haloscope (Superconduct. Resonant Freq.)	DM	[141, 142]
LAMPOST	Haloscope (Dielectric)	DM	[94]
MADMAX	Haloscope (Dielectric)	DM	[93]
ORGAN	Haloscope	DM	[126]
QUAX	Haloscope	DM	[143]
TOORAD	Haloscope (Topological anti-ferromagnets)	DM	[144, 145]
WISPLC	Haloscope (Tunable LC circuit)	DM	[98]
<i>LSW and optics</i>			
ALPS	Light-shining-through wall	Any	[103]
ALPS II	Light-shining-through wall (projection)	Any	[106]
CROWS	Light-shining-through wall (microwave)	Any	[104]
OSQAR	Light-shining-through wall	Any	[105]
PVLAS	Vacuum magnetic birefringence	Any	[107]
<i>Helioscopes</i>			
CAST	Helioscope	Any	[99, 100]
babyIAXO	Helioscope (projection)	Any	[82, 101, 102]
IAXO	Helioscope (projection)	Any	[82, 101, 102]
IAXO+	Helioscope (projection)	Any	[82, 101, 102]

Table 1. List of experimental searches for axions and ALPs. The table is continued in table 2.

Experiment:	Principle	DM?	Reference
<i>Astrophysical constraints</i>			
Breakthrough Listen	ALP \rightarrow radio γ in neutron star magn. fields	DM	[146]
Bullet Cluster	Radio signal from ALP DM decay	DM	[147]
Chandra	AGN X-ray prod. in cosmic magn. field	Any	[110–113]
BBN + N_{eff}	ALP thermal relic perturbing BBN and N_{eff}	Any	[148]
Chandra MWD	X-rays from Magnetic White Dwarf ALP prod.	Any	[149]
COBE/FIRAS	CMB spectral distortions from DM relic decay	DM	[150]
Distance ladder	ALP \leftrightarrow γ perturbing luminosity distances	Any	[151]
Fermi-LAT	SN ALP product. \rightarrow γ -rays in cosmic magn. field	Any	[152–154]
Fermi-LAT	AGN X-ray production \rightarrow ALP in cosmic magn. field	Any	[155]
Haystack Telescope	ALP DM decay \rightarrow microwave photons	DM	[156]
HAWC TeV Blazars	$\gamma \rightarrow$ ALP \rightarrow γ conversion reducing γ -ray attenuation	Any	[157]
H.E.S.S.	AGN X-ray production \rightarrow ALP in cosmic magn. field	Any	[158]
Horizontal branch stars	Stellar metabolism and evolution	Any	[114]
LeoT dwarf galaxy	Heating of gas-rich dwarf galaxies by ALP decay	DM	[159]
Magnetic white dwarf pol.	$\gamma \rightarrow$ ALP conversion polarizing light from MWD stars	Any	[160]
MUSE	ALP DM decay \rightarrow optical photons	DM	[161]
Mrk 421	Blazar X-ray production \rightarrow γ in cosmic magn. field	Any	[162]
NuStar	Stellar ALP production \rightarrow γ in cosmic magn. fields	Any	[163, 164]
NuStar, Super star clusters	Stellar ALP production \rightarrow γ in cosmic magn. fields	Any	[164]
Solar neutrinos	ALP energy loss \rightarrow changes in neutrino production	Any	[165]
SN1987A ALP decay	SN ALP production \rightarrow γ decay	Any	[166]
SN1987A gamma rays	SN ALP production \rightarrow γ in cosmic magnetic field	Any	[167, 168]
SN1987A neutrinos	SN ALP luminosity less than neutrino flux	Any	[168, 169]
Thermal relic compilation	Decay and BBN constraints from ALP thermal relic	Any	[115]
VIMOS	Thermal relic ALP decay \rightarrow optical photons	Any	[170]
White dwarf mass relation	Stellar ALP production perturbing WD metabolism	Any	[171]
XMM-Newton	Decay of ALP relic	DM	[172]
<i>Astrophysical projections</i>			
eROSITA	X-ray signal from ALP DM decay	DM	[173]
Fermi-LAT	SN ALP production \rightarrow γ in cosmic magnetic field	Any	[174]
IAXO	Helioscope detection of supernova axions	Any	[175]
THESEUS	ALP DM decay \rightarrow x-ray photons	DM	[176]
<i>Neutron coupling:</i>			
CASPER-wind	NMR from oscillating EDM (projection)	DM	[177, 178]
CASPER-ZULF-Comag.	NMR from oscillating EDM	DM	[178, 179]
CASPER-ZULF-Sidechain	NMR (constraint & projection)	DM	[178, 180]
NASDUCK	ALP DM perturbing atomic spins	DM	[181]
nEDM	Spin-precession in ultracold neutrons and Hg	DM	[178, 182]
K-3He	Comagnetometer	DM	[183]
Old comagnetometers	New analysis of old comagnetometers	DM	[184]
Future comagnetometers	Comagnetometers	DM	[184]
SNO	Solar ALP flux from deuterium dissociation	Any	[185]
Proton storage ring	EDM signature from ALP DM	DM	[186]
Neutron Star Cooling	ALP production modifies cooling rate	Any	[187]
SN1987 Cooling	ALP production modifies cooling rate	Any	[188]
<i>Coupling independent:</i>			
Black hole spin	Superradiance	Any	[189]
Lyman- α	Modification of small-scale structure	DM	[109]

Table 2. List of experimental searches for axions and ALPs.

Part II

Axion kinetic misalignment

In part I, we discussed how axions and ALPs are well-motivated extensions of the standard model that play an important role as dark matter candidates. As is summarized in fig. 1, many experiments search for such ALP dark matter across the $[m_a, f_a]$ parameter space. Many of these experiments are primarily sensitive to axion dark matter in the low- f_a regime, as the axion-SM couplings in this regime are less suppressed. However, the production mechanism which is most commonly invoked, the standard misalignment mechanism (SMM), does not produce sufficient dark matter in the low- f_a regime. In this section, which is based on our publication [1], we explore how we can go beyond the standard paradigm and produce dark matter with axion kinetic misalignment.

Axion kinetic misalignment involves axion scenarios with an initial large kinetic energy. In such models, which were pioneered in [27], the axion has sufficient energy to roll over the periodic potential barriers, such that the complex PQ field is rotating in field space at early times rather than being frozen. As this rotation can persist beyond the time of onset of oscillations in SMM of an ALP with similar mass and decay constant, the dark matter relic can be enhanced relative to SMM. This allows axion DM to be accounted for in the low- f_a regime. Furthermore, in kinetically misaligned scenarios, the field can undergo parametric resonance in the form of axion fragmentation [29].

We here in part II study axion kinetic misalignment and the role that fragmentation plays in such scenarios. We develop an analytical description of kinetic misalignment and fragmentation. We remain agnostic to the mechanism which sets the initial conditions prior to kinetic misalignment and develop the description in this part of the thesis in an implementation-independent way. Within this description, we map out precisely which regions of parameter space correspond to fragmentation before trapping, fragmentation after trapping, kinetic misalignment with weak fragmentation and which regions can only be supported by the standard misalignment mechanism. Finally, we point out how the non-trivial momentum spectrum predicted by axion misalignment can give rise to a gravitational wave signature that could be seen in measurements of the μ distortions of the CMB.

5 Axion dark matter from kinetic misalignment

We consider the cosmological evolution of an axion field θ whose Lagrangian is given by

$$\mathcal{L} = -\frac{f_a^2}{2}(\partial_\mu\theta)^2 - V(\theta) = -\frac{f_a^2}{2}(\partial_\mu\theta)^2 - m_a^2(T)f_a^2[1 - \cos(\theta)], \quad (5.1)$$

Here we will either take $m_a(T)$ to be constant or to have a QCD-like temperature dependence as specified by eq. (2.8). We decompose the ALP field $\theta(t, \mathbf{x})$ into a homogeneous mode $\Theta(t)$, and small fluctuations $\delta\theta(t, \mathbf{x})$, where the latter can be expanded into Fourier modes as

$$\delta\theta(t, \mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} \theta_{\mathbf{k}}(t) e^{-i\mathbf{k}\cdot\mathbf{x}}. \quad (5.2)$$

We will assume that the Fourier modes $\theta_{\mathbf{k}}(t)$ evolve independently. We also separate the time evolution of the modes from their statistical properties by defining

$$\theta_{\mathbf{k}}(t) = \hat{\theta}_{\mathbf{k}} \theta_k(t), \quad (5.3)$$

where $\hat{\theta}_{\mathbf{k}}$'s carry the statistical properties, and θ_k 's are c -number functions which depend only on the magnitude of the momentum $k \equiv |\mathbf{k}|$ due to the isotropy of the equations of motions in the linear approximation. We also assume that $\hat{\theta}_{\mathbf{k}}$'s obey statistical homogeneity and isotropy:

$$\langle \hat{\theta}_{\mathbf{k}} \hat{\theta}_{\mathbf{k}'}^* \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{k}') \quad (5.4)$$

The initial conditions for the mode function θ_k will be determined by the initial field power spectrum P_θ which depends on the physical process that creates the initial conditions for the Fourier modes. We study the initial conditions in section 6.3.

If we neglect the back-reaction, the equations of motion of the homogeneous mode Θ and of the fluctuations θ_k are decoupled from each other. These are respectively given by

$$\ddot{\Theta} + 3H\dot{\Theta} + m_a^2(T) \sin \Theta = 0, \quad (5.5)$$

and

$$\ddot{\theta}_k + 3H\dot{\theta}_k + \left[\frac{k^2}{a^2} + m_a^2(T) \cos \Theta \right] \theta_k = 0. \quad (5.6)$$

We start by studying the evolution of the homogeneous mode and derive today's value of the relic density in sections 5.1 and 5.2. Then we discuss the implications of the parametric resonance in section 6.1. We conclude this section by providing a semi-analytical estimate of the back-reaction of the fluctuations on the homogeneous mode. We comment on the implications in section 6.2.

5.1 Derivation of the relic density today from the adiabatic invariant

In the kinetic misalignment mechanism [27] we assume that the ALP field has a large initial velocity such that its initial kinetic energy is much larger than its potential energy. For now, we remain agnostic as to the source of this initial velocity and postpone a discussion of how this kinetic energy can be delivered until Part III of this thesis. Regardless of the source, given sufficient kinetic energy, the field travels over many barriers without being trapped. As the universe expands, the kinetic energy will be drained by Hubble friction. At the same time, the potential energy will grow if the axion mass has a QCD-like temperature dependence as discussed in section 2.1. At some temperature T_* , the energy in the homogeneous mode ρ_Θ becomes smaller than the height of the barriers. After this time, the ALP field cannot continue rolling as it is *trapped* by the barrier. The trapping temperature T_* is defined by the condition

$$\rho_\Theta(T_*) = \frac{1}{2} f_a^2 \dot{\Theta}^2(T_*) + m_a^2(T_*) f_a^2 [1 - \cos(\Theta(T_*))] = 2m_a^2(T_*) f_a^2 [1 - \cos(\Theta(T_*))]. \quad (5.7)$$

From now on the *-subscript denotes the quantities evaluated at $T = T_*$. We introduce the dimensionless parameter $\varrho(t)$ as the ratio between the ALP kinetic energy and its potential energy⁶

$$\varrho(t) \equiv \frac{\rho_\Theta}{2m_a^2(t) f_a^2} = \frac{1}{4} \frac{\dot{\Theta}^2}{m_a^2(t)} + \sin^2 \left(\frac{\Theta}{2} \right). \quad (5.8)$$

The field is trapped at $\varrho(t_*) = 1$. If $\varrho > 1$ it keeps rolling, while for $\varrho < 1$ it is oscillating around its minimum.

The evolution of $\varrho(t)$ can be studied analytically by using the *action-angle formalism* [190, 191], see also [192]. Neglecting the back-reaction, the evolution of the homogeneous mode is governed by the Lagrangian density

$$\mathcal{L}_0 = f_a^2 a^3(t) \left[\frac{1}{2} \dot{\Theta}^2 - m_a^2(t) (1 - \cos \Theta) \right]. \quad (5.9)$$

This Lagrangian has a discrete shift symmetry given by $\Theta \rightarrow \Theta + 2\pi$, so the motion in Θ -space can be considered periodic. In the following, we define the period to be the time it takes for the field to roll from one maximum to the other.

⁶We will use the time and temperature interchangeably depending on the nature of the quantity that we want to calculate. We always assume radiation domination so $t = (2H)^{-1}$, and $H(T) \propto g_*(T)^{1/2} T^2$ where g_* is the effective number of relativistic degrees of freedom. We assume that g_* is constant throughout the parametric resonance so we can change from temperature to time and vice versa by $T^2 \propto t^{-1}$.

Now we assume that the parameters carrying explicit time dependence, namely a and m_a , change *adiabatically*; in other words, they remain approximately constant during a single period. This assumption is justified if the Hubble time scale is larger than the period of the motion. We will justify this assumption later. In the case of periodic motion where the Lagrangian changes adiabatically, the *action variable* J defined by

$$J \equiv \oint \Pi_q dq, \quad (5.10)$$

is a constant of motion, where the integral is over a single period, and

$$q = f_a \Theta \quad \text{and} \quad \Pi_q = \frac{\partial \mathcal{L}_0}{\partial \dot{q}} = a^3(t) f_a \dot{\Theta}. \quad (5.11)$$

Assuming that the scale factor and the axion mass are constant during one period, J becomes

$$J \approx f_a^2 a^3(t) \oint d\Theta \dot{\Theta} = 2m_a(t) f_a^2 a^3 \oint \sqrt{\varrho - \sin^2 \left(\frac{\Theta}{2} \right)}. \quad (5.12)$$

The integral can be performed analytically, and one gets

$$J \approx m_a(t) f_a^2 a^3(t) \times \begin{cases} 8\sqrt{\varrho} E(1/\sqrt{\varrho}), & \varrho > 1 \\ 8[(\varrho - 1)K(\sqrt{\varrho}) + E(\sqrt{\varrho})], & \varrho < 1 \end{cases}, \quad (5.13)$$

where K and E are complete elliptic integrals of the first and second kind, respectively⁷. By using the limiting behaviour of these as $\varrho \rightarrow 1$, we can show that

$$\lim_{\varrho \rightarrow 1^+} J(\varrho) = \lim_{\varrho \rightarrow 1^-} J(\varrho) = 8m_a(t) f_a^2 a^3(t). \quad (5.14)$$

The adiabatic invariant J remains approximately constant throughout the whole evolution. This makes $J(\varrho)$ a perfect quantity to compare the early and late time behaviors. At early times when $\varrho \gg 1$ we have

$$J(\varrho \gg 1) \approx 2\pi f_a^2 \dot{\Theta} a^3, \quad (5.15)$$

whereas at late times we get

$$J(\varrho \ll 1) \approx \pi \frac{\rho_{\Theta} a^3}{m_a}, \quad (5.16)$$

where m_a is the zero-temperature axion mass. Let us introduce the *yield* quantity defined by

$$Y \equiv n_{\text{PQ}}/s, \quad (5.17)$$

where s is the entropy density of the universe, and

$$n_{\text{PQ}} = f_a^2 \dot{\Theta} \quad (5.18)$$

is the Noether charge of the axion shift symmetry. This quantity is conserved during the early evolution when the ALP mass is negligible, so the equation of motion of the ALP field is just

⁷We use the following definitions for the elliptic integrals:

$$K(k) = \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}} \quad \text{and} \quad E(k) = \int_0^{\pi/2} d\varphi \sqrt{1 - k^2 \sin^2 \varphi}$$

Note that most software packages such as `Mathematica` and `scipy` [193] uses $m = k^2$ instead of k as their argument when defining elliptic integrals.

$\ddot{\Theta} + 3H\dot{\Theta} \approx 0$. The conservation of $J(\varrho)$ implies that eq. (5.15) and eq. (5.16) should be equal, which gives

$$\frac{\rho_{\Theta, \text{today}}}{s_{\text{today}}} \approx 2m_a Y. \quad (5.19)$$

where $\rho_{\Theta, \text{today}}$ and s_{today} are respectively the relic density⁸ and entropy density today. This equation is also derived in [27], however, the numerical factor there was obtained numerically. Here we presented an analytical derivation.

From this result, we obtain a simple formula for the fractional energy density of the ALP field today, which is

$$h^2 \Omega_{\Theta, \text{today}} \approx h^2 \Omega_{\text{DM}} \left(\frac{m_a}{5 \times 10^{-3} \text{ eV}} \right) \left(\frac{Y}{40} \right), \quad (5.20)$$

where we took $h^2 \Omega_{\text{DM}} = 0.12$ [194]. Note that the yield only depends on the zero temperature axion mass, not on the high-temperature behaviour. The latter will be important in determining the trapping temperature T_* . With zero-temperature QCD axion mass, eq. (2.5), the above relation implies

$$h^2 \Omega_{\phi, \text{today}}^{\text{QCDaxion}} \approx h^2 \Omega_{\text{DM}} \left(\frac{10^9 \text{ GeV}}{f_a} \right) \left(\frac{Y}{40} \right). \quad (5.21)$$

Before closing this subsection, we mention in which conditions the adiabaticity assumption is justified. To do this, we need to calculate the periods of the motion before and after trapping. By neglecting the energy loss due to the Hubble expansion during a single oscillation, the periods $\mathcal{T}_>$ before and $\mathcal{T}_<$ after trapping can be derived as

$$\mathcal{T}_>(\varrho) = \frac{2}{m_a(T)\sqrt{\varrho}} \text{K}(1/\sqrt{\varrho}) \quad \text{and} \quad \mathcal{T}_<(\varrho) = \frac{2}{m_a(T)} \text{K}(\sqrt{\varrho}). \quad (5.22)$$

At early times when $\varrho \gg 1$ we have

$$\mathcal{T}_>(\varrho \gg 1) \approx \frac{\pi}{m_a(T)} \sqrt{\frac{1}{\varrho}} \approx \frac{2\pi}{\dot{\Theta}}, \quad (5.23)$$

so the adiabaticity condition reads $\dot{\Theta} > 2\pi H$. At early times $\dot{\Theta} \propto a^{-3}$ and $H \propto a^{-2}$ so it is sufficient if this condition is satisfied at trapping $\dot{\Theta} = 2m_a(T_*)$. This yields $m_* > \pi H_*$ as the adiabaticity condition. On the other hand, at late times when $\varrho \ll 1$ we have

$$\mathcal{T}_<(\varrho \ll 1) \approx \frac{\pi}{m_a(T)}, \quad (5.24)$$

so the adiabaticity condition becomes $m > \pi H$. Again, it is sufficient that this condition is satisfied at trapping, which yields

$$\frac{m_a(T_*)}{H_*} \gtrsim \pi, \quad (5.25)$$

as the adiabaticity condition. As we will show explicitly in section 7, the most interesting region of the parameter space is where $m_a(T_*)/H_* \gg 1$ which justifies the adiabaticity assumption. We note that even for $m_a(T_*)/H_* \gg 1$, the adiabaticity is broken for a period around $\varrho = 1$, but this period is shorter for larger $m_a(T_*)/H_*$.

⁸As we in this part of the thesis consider fluctuations, ρ_a is split into the energy density of the homogenous field ρ_{Θ} and the energy density in fluctuations ρ_{fluct} .

5.2 Trapping temperature

The trapping temperature is an important quantity since fragmentation happens around the same time as trapping. We recall that the trapping happens at $\varrho = 1$ for which

$$J(\varrho = 1) = 8m_a(T_*)f_a^2a_*^3. \quad (5.26)$$

By comparing this with eq.(5.16), and using conservation of $J(\varrho)$ we obtain

$$\frac{m_a(T_*)}{m_a} \left(\frac{a_*}{a_{\text{today}}} \right)^3 = \frac{\pi}{8} \frac{\rho_{\Theta, \text{today}}}{\Lambda_{b,0}^4}, \quad (5.27)$$

where $\Lambda_{b,0}^4 = m_a^2 f_a^2$ is the zero-temperature barrier height. By assuming the QCD-like temperature dependence of $m_a(T)$ in eq. (2.8) we can calculate the trapping temperature via (5.27). The result is

$$\boxed{\frac{T_*}{\Lambda_{b,0}} \approx (2.12)^{\frac{\gamma}{6+\gamma}} (2 \times 10^8)^{\frac{2}{6+\gamma}} \left(\frac{g_s(T_*)}{72} \right)^{-\frac{2}{6+\gamma}} \left(\frac{\Lambda_{b,0}}{\text{GeV}} \right)^{\frac{2}{6+\gamma}} \left(\frac{h^2 \Omega_{\phi, \text{today}}}{h^2 \Omega_{\text{DM}}} \right)^{-\frac{2}{6+\gamma}}}. \quad (5.28)$$

We can observe that the trapping temperature depends only on the zero temperature barrier height and on the scaling of the axion mass at high temperatures. For the QCD axion, the trapping temperature does not depend on the axion mass, and it is given by

$$T_*^{\text{QCD}} \approx (1.21 \text{ GeV}) \left(\frac{h^2 \Omega_{\phi, \text{today}}}{h^2 \Omega_{\text{DM}}} \right)^{-0.141}. \quad (5.29)$$

We show a plot of the trapping temperature as a function of the zero-temperature barrier height for various choices of γ in figure 2. Equation (5.28) is a key quantity that determines the different fragmentation regimes.

6 Analytical theory of parametric resonance in kinetic misalignment

6.1 Parametric resonance

After the discussion of the evolution of the homogeneous mode Θ , we now turn our attention to fluctuations. We generalize the results of [29] by including the regime of fragmentation after trapping. We also use different initial conditions for the fluctuations. If we consider time scales much shorter than H^{-1} , we can neglect the expansion of the universe and take a and $m_a(T)$ to be approximately constant. The equation of motion for axion mode functions (5.6) then takes the form

$$\ddot{\theta}_k + \left[\frac{k^2}{a^2} + m_a^2(T) \cos \Theta \right] \theta_k = 0. \quad (6.1)$$

This is a second-order differential equation with periodic coefficients, so it has the form of the *Hill's equation* [195]. According to the Floquet theorem, the solutions should be of the form

$$\theta_k(t) = \theta_+(t; k)e^{\mu_k t} + \theta_-(t; k)e^{-\mu_k t}, \quad (6.2)$$

where $\theta_{\pm}(t; k)$ are periodic functions in time, and μ_k 's are in general complex numbers known as *Floquet exponents*. If $\text{Re}\{\mu_k\} > 0$, then the mode grows exponentially during one oscillation, and one says that the mode experiences instability via parametric resonance. The primary goal of the *Floquet analysis* is to determine the *instability bands* of the system, which are the regions where $\text{Re}\{\mu_k\} > 0$.

In the case of oscillations after trapping, a closed-form expression for the Floquet exponents has been obtained in [196]. We have found that this method can easily be extended to derive an

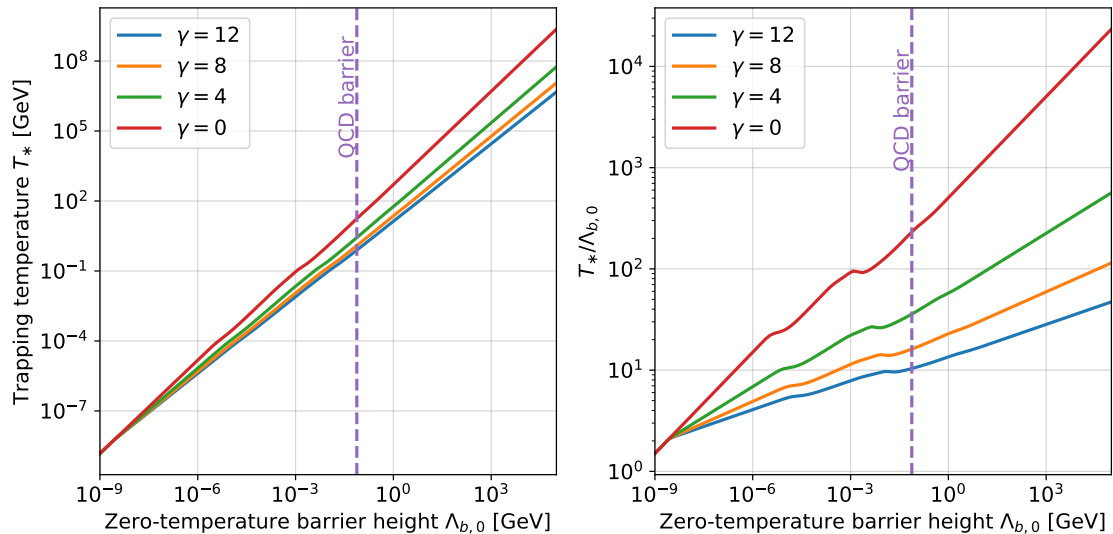


Figure 2. Trapping temperature T_* given in eq. (5.28) (left figure), and the ratio between T_* and the zero-temperature barrier height $\Lambda_{b,0}$ (right plot) as a function of $\Lambda_{b,0}$ for different choices of γ which parametrizes the early-time scaling of the barrier height, and thus of the axion mass, with temperature as defined in (2.8). We see that this ratio increases for larger $\Lambda_{b,0}$, and the trapping temperature becomes more sensitive to the early time behaviour of the axion potential. The dashed vertical purple line shows the zero-temperature QCD barrier height 75.6 MeV.

analogous expression for Floquet exponents during the evolution before trapping. In this section, we will just state the result and discuss the consequences. The detailed derivation is presented in Appendix B.

Let ϱ be the dimensionless energy of the homogeneous mode as defined in (5.8), and let us introduce a dimensionless momentum

$$\kappa \equiv k/(a m_a(T)). \quad (6.3)$$

Then the Floquet exponents have the following analytical solution:

$$\mu_k = \sqrt{8\kappa^2(\varrho - \kappa)(1 - \varrho + \kappa^2)} \times \begin{cases} \mathcal{T}_>^{-1}(\varrho) \mathcal{I}(\arcsin(1/\sqrt{2\varrho - 1})), & \varrho > 1 \\ \mathcal{T}_<^{-1}(\varrho) \mathcal{I}(\pi/2), & \varrho < 1 \end{cases}, \quad (6.4)$$

The first line in (6.4) is new, while the second line was derived in the literature. $\mathcal{T}_>$ and $\mathcal{T}_<$ are the periods before and after trapping, respectively, and $\mathcal{I}(\varphi)$ is the following integral:

$$\mathcal{I}(\varphi) = \int_0^\varphi \frac{d\vartheta'}{1 + (1 - 2\varrho + 2\kappa^2) \sin^2 \vartheta'} \frac{2 \sin^2 \vartheta'}{\sqrt{(1 + \sin^2 \vartheta') [1 + (1 - 2\varrho) \sin^2 \vartheta']}}. \quad (6.5)$$

The periods \mathcal{T} can be derived from the conservation of energy. They read

$$\mathcal{T}_>(\varrho) = \frac{2}{m_a(T)\sqrt{\varrho}} K(1/\sqrt{\varrho}) \quad \text{and} \quad \mathcal{T}_<(\varrho) = \frac{2}{m_a(T)} K(\sqrt{\varrho}). \quad (6.6)$$

From the square root term in (6.4) we can immediately obtain the instability band. Before trapping, it is given by

$$\varrho - 1 < \kappa^2 < \varrho, \quad (6.7)$$

while its form after trapping is

$$0 < \kappa^2 < \varrho. \quad (6.8)$$

If one neglects the Hubble expansion, both ϱ and κ do not change, so these modes keep growing. Therefore, in order to get realistic results, we need to incorporate the Hubble expansion, which we will do next.

The Hubble expansion and the temperature dependence of the axion mass can be incorporated by restoring the time dependence of ϱ and κ in the expression for the Floquet exponent (6.4):

$$\kappa \rightarrow \kappa(t) = \frac{k/a(t)}{m_a(t)} \quad \text{and} \quad \varrho \rightarrow \varrho(t), \quad (6.9)$$

where $\varrho(t)$ is given by (5.8). As a result, the Floquet exponents now become time-dependent:

$$\mu_k \rightarrow \mu_k(\kappa(t), \varrho(t)). \quad (6.10)$$

Then we can obtain the total amplification factor N_k of a given mode by integrating the time-dependent Floquet exponents over time:

$$N_k(t) = \exp\left(\int_{t_i}^t dt' \mu(\kappa(t'), \varrho(t'))\right), \quad (6.11)$$

where t_i is some initial time before which the parametric resonance is not effective. At this point, it is convenient to introduce dimensionless quantities via

$$\kappa_* \equiv \frac{k/a_*}{m_a(T_*)}, \quad \tau \equiv 2H_* t, \quad \tilde{\mu} \equiv \frac{\mu}{m_a(T_*)}, \quad (6.12)$$

where *-subscript implies that the quantities are evaluated at T_* defined by (5.7). In terms of these quantities, (6.11) becomes

$$N_k(\tau) = \exp\left(\frac{m_a(T_*)}{2H_*} \int_{\tau_i}^{\tau} d\tau' \tilde{\mu}(\kappa(\tau'), \varrho(\tau'))\right) \equiv \exp\left(\frac{m_a(T_*)}{2H_*} \mathcal{B}_k(\tau)\right). \quad (6.13)$$

The growth factor \mathcal{B}_k depends on the temperature scaling γ of the axion mass, but not on the model parameters such as $m_a(T_*)$ and H_* . For the modes which amplify most efficiently its value is $\sim \mathcal{O}(0.5)$. Therefore the efficiency of the fragmentation is effectively determined by the hierarchy between the axion mass and Hubble rate at trapping. This result can be understood physically. If the Hubble rate is much smaller than the axion mass at the beginning of oscillations, the redshifting of the homogeneous mode becomes very slow, which allows the axion to probe non-quadratic parts of its potential for a longer time. In the standard misalignment mechanism, $m_a(T_*)/H_* \sim 3$ so the expansion quickly redshifts the amplitude of the oscillations, which makes the parametric resonance ineffective. In a nutshell, efficient parametric resonance requires a mechanism which delays the onset of oscillations. The Kinetic Misalignment Mechanism provides this via large initial kinetic energy. In the Large Misalignment Mechanism [197, 198] this is achieved by tuning the initial angle to the top of the axion potential such that the onset of oscillations is delayed due to the small potential gradient at the top.

It is also instructive to study how the shape of the instability bands changes with time. For this, let us introduce the following quantity:

$$\tilde{J}(t) = \frac{m_a(t_*)}{m_a(t)} \left(\frac{a_*}{a(t)}\right)^3. \quad (6.14)$$

Then we can write $\kappa(\tau)$ as

$$\kappa(\tau) = \kappa_* \tilde{J}(\tau) \tau. \quad (6.15)$$

To obtain $\varrho(\tau)$ we note that the adiabatic invariant (5.13) can be expressed in terms of \tilde{J} by

$$J(\varrho(\tau)) = 8f_a^2 m_a(t) a^3(t) \tilde{J}(\tau), \quad (6.16)$$

so that

$$\tilde{J}(\varrho) = \begin{cases} \sqrt{\varrho} E(1/\sqrt{\varrho}), & \varrho > 1 \\ (\varrho - 1)K(\sqrt{\varrho}) + E(\sqrt{\varrho}), & \varrho < 1 \end{cases}. \quad (6.17)$$

At a given time, this function can be inverted to get $\varrho(\tau)$. Generally, this needs to be done numerically. However, in asymptotic regimes, we can use the following approximations:

$$\tilde{J}(\varrho \gg 1) \approx \frac{\pi}{2} \sqrt{\varrho} \quad \text{and} \quad \tilde{J}(\varrho \ll 1) \approx \frac{\pi}{4} \varrho. \quad (6.18)$$

We are now ready to discuss the behavior of the instability bands. Before trapping, the expression for the instability band (6.7) can be written as

$$\frac{\varrho - 1}{(\tilde{J}(\tau)\tau)^2} < \kappa_*^2 < \frac{\varrho}{(\tilde{J}(\tau)\tau)^2}. \quad (6.19)$$

Initially, ϱ is large, so the width of the instability band is narrow. As the energy of the homogeneous mode decreases, the instability band gradually widens. The behavior after trapping is slightly more involved. The expression for the instability band (6.8) becomes

$$0 < \kappa_*^2 < \frac{\varrho}{(\tilde{J}(\tau)\tau)^2}. \quad (6.20)$$

At late times when $\varrho \ll 1$ we have $\varrho \approx 4\tilde{J}/\pi$. Therefore the upper limit of the instability band can be approximated as

$$\frac{\varrho}{(\tilde{J}(\tau)\tau)^2} \approx \frac{4}{\pi} \tau^{-1/2} \times \begin{cases} \tau^{\gamma/4}, & T > T_c \\ m_a/m_a(T_*), & T < T_c \end{cases}. \quad (6.21)$$

We see that for $\gamma > 2$, the upper limit of the instability band grows with time until the axion reaches its zero-temperature mass and then decreases with time. This means that the modes with $\kappa_* > 1$ which did enter and exit the instability band at the rolling stage will re-enter the instability band after trapping. On the other hand, for $\gamma < 2$, and also for constant axion mass, the width of the instability band will shrink after trapping. We show the evolution of the instability bands, Equations (6.19) and (6.20), as a function of the scale factor for these two cases in figure 3.

Even though the Floquet analysis predicts a very wide instability band, the fragmentation is most efficient around trapping, as can be observed in figure 4. In this plot, we have used the parameters of figure 3 with $\gamma = 8$, but we also show the value of the Floquet exponents and zoom in to the region around trapping. The white lines denote the boundaries of the instability band.

Our approximation so far sizably overpredicts the amplification factor. The reason is that the Floquet analysis gives the amplification factor during one oscillation by neglecting the damping of the mode due to the redshift. More specifically, if the Floquet exponents for a mode are imaginary at all times, our approximation would predict that the amplitude of this mode will stay constant, which of course is not correct. Instead, the amplitude of this mode will get redshifted and becomes smaller. To account for this, we assume that on top of the amplification given by (6.13), all modes redshift like a free particle. This decay factor can be calculated using WKB approximation, which yields

$$A_k(t) \propto \omega_k^{-1/2}(t) a^{-3/2}(t) \quad \text{where} \quad \omega_k(t) = \sqrt{\frac{k^2}{a^2(t)} + m_a^2(T)}. \quad (6.22)$$

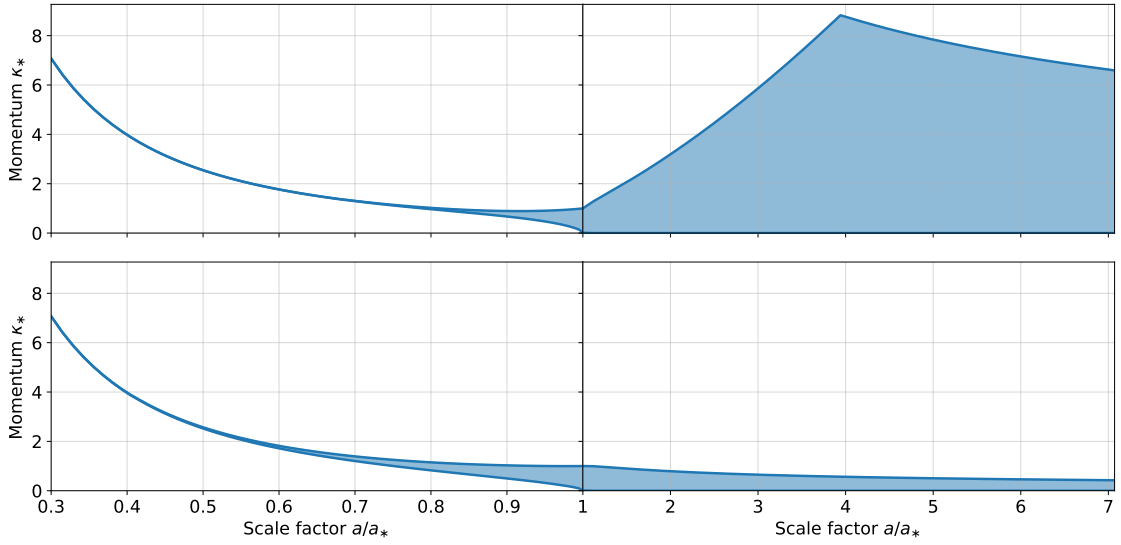


Figure 3. Time evolution of the instability bands, eq. (6.19) and (6.20), as a function of the scale factor for two benchmark points with $\gamma = 8$ (upper plot), and constant mass $\gamma = 0$ (lower plot). Here, $m_* = m_a(T_*)$. In both cases we set $m_a = 10^{-11}$ eV, and choose the decay constant f_a such that $m_a(T_*)/H_* = 10^2$.

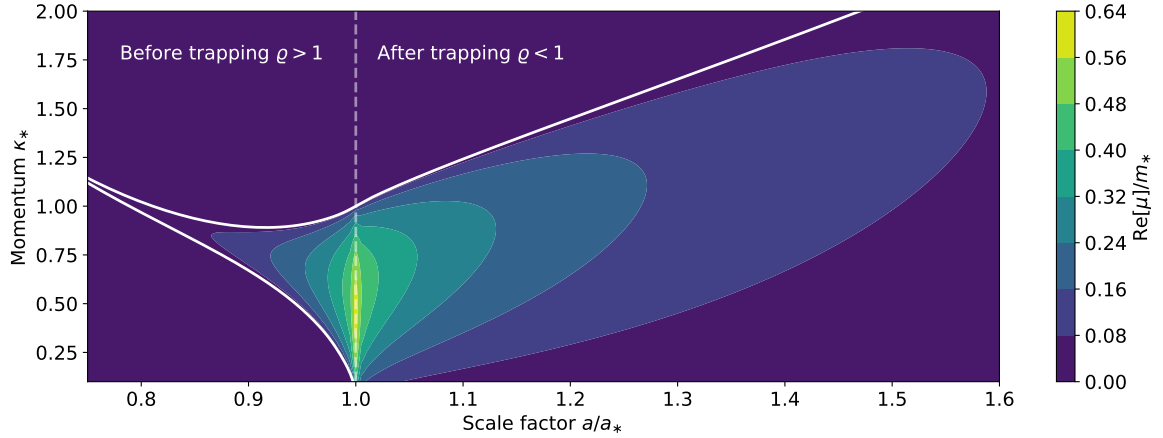


Figure 4. Evolution of the instability bands together with the value of the Floquet exponents by zooming into the region around trapping. The white lines denote the boundaries of the instability band. Fragmentation is most efficient around trapping $a \approx a_*$. The model parameters are identical to the ones used in figure 3 with $\gamma = 8$.

So our final ansatz for the mode functions is

$$|\theta_k(t)| = \theta_{k,i} A_k(t) N_k(t), \quad (6.23)$$

where $\theta_{k,i}$ is related to the initial field power spectrum (5.4) by $\theta_{k,i} = \sqrt{P_\theta(\vec{k})}$, $A_k(t)$ is normalized such that it is unity initially, and we have omitted the oscillatory term. We provide a comparison between this ansatz and the full numerical solutions of the mode functions in figure 5. The parameters are the same as in figures 3 and 4. The thin solid lines show the numerical solutions, while the thick dashed lines are calculated via (6.23). We started the numerical solution at $\tau = 0.1$ with

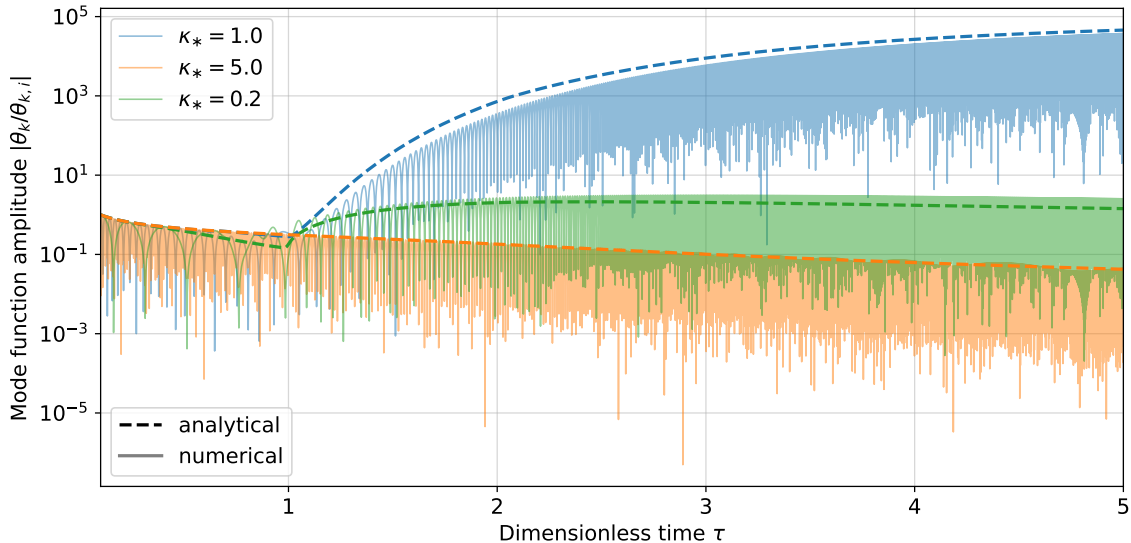


Figure 5. Comparison between the numerical solution (thin solid lines) and our analytical ansatz (6.23) (thick dashed lines) for three benchmark momentum modes, showing a very reasonable agreement. The parameters are the same as in figures 3 and 4.

an initial amplitude $\theta_{k,i} = 1$ for all the modes. We confirm that our ansatz (6.23) provides a very reasonable approximation to the numerical solution.

6.2 An estimate of the back-reaction

In our discussion so far, we have neglected the back-reaction of the fluctuations on the homogeneous mode. However, this approximation breaks down when fragmentation becomes efficient because the fluctuations eventually cannot be considered small compared to the homogeneous mode [29]. If the only thing we want to know is whether at a given point on the ALP parameter space the relic density is dominated by fragmented axions or by the homogeneous mode, then the back-reaction effects might not be important. However, such effects are important for observational consequences of fragmentation such as mini-clusters [30]. To estimate such consequences, we need to calculate the density power spectrum of the ALP field after fragmentation, and in order to get accurate results we need to take the back-reaction into account.

A precise study of the back-reaction should be performed by using non-perturbative methods such as lattice simulations. However, we can obtain a simple estimate semi-analytically. For this, we only need to assume that fragmentation steals energy from the homogeneous mode and gives it to the fluctuations. This implies that the energy gained by the fluctuations during the fragmentation is equal to the energy lost by the homogeneous mode during the same period:

$$\Delta\rho_{\text{fluct}} = -\Delta\rho_{\Theta}, \quad (6.24)$$

where ρ_{Θ} is the energy density in the homogeneous mode. This idea has also been used in [29] for an extensive discussion of the fragmentation before trapping. As we will show later, our calculations reproduce these results under appropriate limits.

The energy density in the fluctuations is

$$\rho_{\text{fluct}} = \frac{f_a^2}{2} \int \frac{d^3k}{(2\pi)^3} \left[\left| \dot{\theta}_k(t) \right|^2 + \left(\frac{k^2}{a^2} + m_a^2(t) \cos \Theta \right) \left| \theta_k(t) \right|^2 \right], \quad (6.25)$$

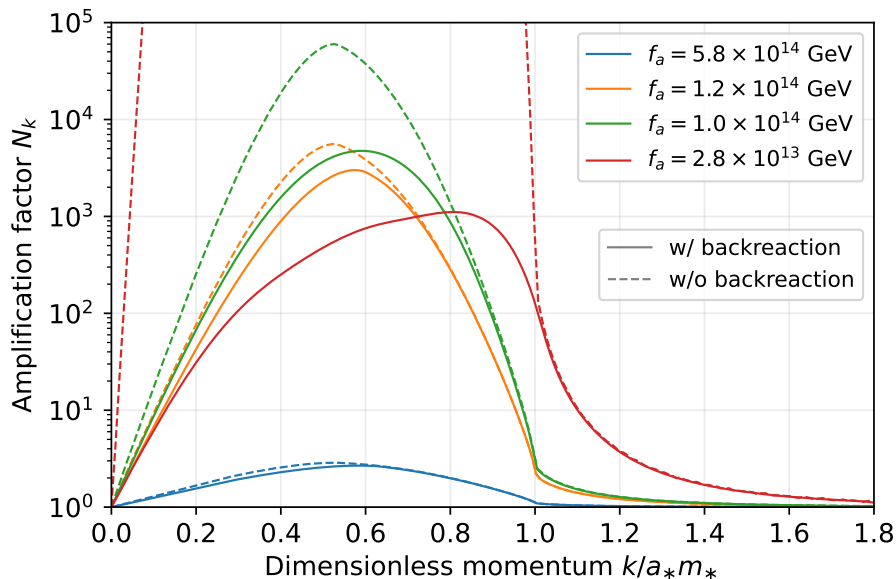


Figure 6. This plot compares the estimates of the amplification factor N_k performed by including/not including the back-reaction effects through the procedure (6.30). It shows the importance of back-reaction effects as f_a is decreased. This plot is made for a constant mass $m_a(T_*) = m_a = 10^{-15}$ eV.

where we have neglected higher-order terms in θ_k . Since the mode functions oscillate very rapidly, we can assume that the kinetic and potential terms contribute equally. We also average over the oscillations of the homogeneous mode. With these assumptions and our ansatz (6.23), we approximate the energy density in fluctuations as

$$\rho_{\text{fluct}}(t) \approx \frac{f_a^2}{4\pi^2} \int dk k^2 \left(\frac{k^2}{a^2} + m_a^2(t) \overline{\cos \Theta} \right) P_\theta(k) A_k^2(t) N_k^2(t), \quad (6.26)$$

where

$$\overline{\cos \Theta} \equiv \frac{1}{\Theta_{\max} - \Theta_{\min}} \int_{\Theta_{\min}}^{\Theta_{\max}} d\Theta \cos \Theta = \begin{cases} 0, & \varrho > 1, \\ \text{sinc}(2 \arcsin \sqrt{\varrho}), & \varrho < 1 \end{cases}. \quad (6.27)$$

The extra 1/2 factor in (6.26) arises from averaging over the mode function oscillations. From this result, we can estimate the energy lost by the homogeneous mode during a short time period Δt as

$$-\frac{\Delta \rho_\Theta}{\Delta t} = \frac{\Delta \rho_{\text{fluct}}}{\Delta t} \approx \frac{f_a^2}{4\pi^2} \int dk k^2 \left(\frac{k^2}{a^2} + m_a^2(t) \overline{\cos \Theta} \right) P_\theta(k) A_k^2(t) \frac{N_k^2(t) - N_k^2(t - \Delta t)}{\Delta t}. \quad (6.28)$$

Note that we have only varied the amplification factor N_k since we are only interested in the change in the energy due to the fragmentation. In the $\Delta t \rightarrow 0$ limit, we find

$$-\frac{d\rho_\Theta}{dt} = \frac{d\rho_{\text{fluct}}}{dt} \approx \frac{f_a^2}{4\pi^2} \int dk k^2 \left(\frac{k^2}{a^2} + m_a^2(t) \overline{\cos \Theta} \right) P_\theta(k) A_k^2(t) [2N_k^2(t) \mu(\kappa(t), \varrho(t))], \quad (6.29)$$

where we have used (6.11) when taking the time derivative.

With this result, we can employ the following procedure to calculate the amplification factors including back-reaction: Choose an initial time τ_i at which $P_\theta(k)$ is known and $N_k(\tau_i) = 1$ for all modes. Pick a sufficiently small time-step $\Delta\tau$, and at each time interval calculate the amplification factors by

$$\ln N_k(\tau + \Delta\tau) = \ln N_k(\tau) + \frac{m_a(T_*)}{2H_*} \int_\tau^{\tau + \Delta\tau} d\tau' \tilde{\mu}_k(\tau') \approx \ln N_k(\tau) + \frac{m_a(T_*)}{2H_*} \Delta\tau \tilde{\mu}_k(\tau). \quad (6.30)$$

For each time step, find ϱ via (6.14) and (6.17) and therefore the time-dependent Floquet exponent μ_k . Take into account the decrease in ϱ due to the back-reaction via (6.29). Then repeat the steps either until fragmentation is no longer efficient or when ϱ drops to zero, indicating complete fragmentation. At this stage, the dynamics of the system become non-perturbative, so our calculation should be considered an estimate. We illustrate the importance of back-reaction effects in figure 6. We will use this procedure in the following, especially to determine the boundary between the ‘after-trapping’ and ‘before-trapping’ fragmentation regimes defined in section 6.

6.3 Initial conditions for the mode functions

In order to study the consequences of fragmentation for ALP dark matter, we need to specify the initial conditions for the mode functions prior to fragmentation. A scalar field, like any component in the universe, can have adiabatic and isocurvature fluctuations. Adiabatic fluctuations arise solely due to the temperature fluctuations in the universe and are unavoidable for any cosmological fluid, including scalar fields. On the other hand, the isocurvature fluctuations are the fluctuations on constant density slices, and they describe any other kind of fluctuation which is not adiabatic.

In the case of standard misalignment mechanism, adiabatic fluctuations of an ALP field are negligible during the early evolution when it is Hubble frozen [199]. They start to grow once the homogeneous mode starts oscillating. If the ALP field is present during inflation, for example, if the Peccei-Quinn (PQ) symmetry is broken during inflation in the case of the QCD axion, then it will pick up quantum fluctuations given by $\delta\phi = H_I/2\pi$, where H_I is the inflation scale. This causes fluctuations in the initial axion angle Θ_i given by $\delta\Theta_i \simeq H_I/(2\pi S_I)$ where S_I is the effective axion decay constant during inflation which can be different from its value at low temperatures [200].

The situation changes drastically in the case of rotating axions. In this case, the ALP field has a much more complicated cosmological history which is highly model-dependent. A common property in all of these histories is that the ALP field starts moving much earlier compared to the standard case, and it has a large velocity long before the ALP potential turns on. As we show below, this velocity, together with curvature perturbations, acts as a source term for adiabatic fluctuations and makes them grow. In addition to these, the kick mechanism can induce isocurvature fluctuations from the quantum fluctuations of the field responsible for producing the kick. These can constrain the parameter space due to their isocurvature nature and can also lead to domain wall problems in some cases [33].

A discussion of model realizations of the kinetic misalignment mechanism, including the possible cosmological histories as well as the implications of isocurvature fluctuations, was presented in the context of kination in [37]. In part III of this thesis, which is based on [2], we will present DM implementations of the kinetic misalignment mechanism in a wider range of cosmological histories. We postpone a discussion of the early cosmological history until part III. For now, we remain agnostic about the model implementation and consider a standard cosmological history. We assume that the modes that are relevant for fragmentation enter the horizon when the ALP field scales as kination. We will present the necessary condition for the validity of this assumption towards the end of this sub-section.

To study the evolution of initial fluctuations, we start by writing the FLRW metric with curvature perturbations included. In conformal time η and Newtonian (conformal) gauge, the metric reads

$$ds^2 = a^2(\eta) \{ -[1 + 2\Psi(\eta, \mathbf{x})] d\eta^2 + [1 + 2\Phi(\eta, \mathbf{x})] \delta_{ij} dx^i dx^j \}. \quad (6.31)$$

Here Ψ and Φ are Bardeen potentials [201], and they represent the curvature perturbations. The mode functions have the following equations of motion at early times: [1]

$$\theta_k'' + 2\mathcal{H}\theta_k' + k^2\theta_k = -4\Phi_k'\Theta', \quad (6.32)$$

where \prime denotes the derivative with respect to conformal time, and $\mathcal{H} = aH$ is the conformal Hubble parameter. The right side of this equation represents the source term due to the curvature fluctuations. This term clearly represents the difference between the standard and kinetic misalignment mechanisms. In the standard case, the ALP field is frozen due to the Hubble friction, so $\Theta' = 0$. Therefore, the source term is absent, and the adiabatic fluctuations remain zero until the oscillations start. However, in the case of kinetic misalignment mechanism, the ALP field receives a kick at a much earlier time, so the right side is non-zero for a much longer time. As a result, the standard and kinetic misalignment mechanisms predict different initial conditions for mode functions. To our knowledge, this is the first time that this has been pointed out. Further implications of this source term will be discussed in upcoming work by Eroncel et al. [202].

By an explicit calculation of the mode functions which we described in Appendix B of [1], see eq. (B.53), we obtained the following result for the power spectrum:

$$\boxed{P_\theta(k; a) = |\theta_k(a)|^2 \approx \frac{2\pi^2}{k^3} \left(\frac{1}{3}\right)^2 A_s \left(\frac{\dot{\Theta}}{H}\right)^2} \quad (6.33)$$

where A_s is the amplitude of the primordial power spectrum. At the pivot scale $k = 0.05 \text{ Mpc}^{-1}$, Planck 2018 (TT,TE,EE+lowE+lensing 68%) [194] has measured $A_{s,\text{Planck}} = 2.1 \times 10^{-9}$. This result is valid for both super- and sub-horizon modes but assumes that the mode is super-horizon when the ALP field starts its kination-like scaling. We denote quantities measured at this onset with the subscript kin . The behavior of the modes which were sub-horizon at a_{kin} cannot be determined without specifying the cosmological history before a_{kin} .

From this result alone, we can put a bound on the duration of the kination-like scaling. For this, we evaluate the variance of the axion velocity:

$$\left\langle (\delta\dot{\theta})^2 \right\rangle = \frac{1}{2\pi^2} \int_0^\infty dk k^2 |\dot{\theta}_k|^2 = \frac{1}{2\pi^2 a^2} \int_0^\infty dk k^2 |\theta'_k(\eta)|^2. \quad (6.34)$$

The integral is dominated by the modes which are sub-horizon but were super-horizon at a_{kin} . Then by approximating $|\theta'_k(\eta)|^2 \approx k^2 |\theta_k(\eta)|^2$, and using (6.33) we find at late times

$$\left\langle (\delta\dot{\theta})^2 \right\rangle \approx \frac{1}{2} \left(\frac{1}{3}\right)^2 A_s \dot{\Theta}^2 \left(\frac{a}{a_{\text{kin}}}\right)^2. \quad (6.35)$$

Then we can estimate the density contrast by

$$\delta_\phi \sim \frac{\delta(\dot{\theta}^2/2)}{\dot{\Theta}^2/2} \sim \frac{2}{\dot{\Theta}} \sqrt{\left\langle (\delta\dot{\theta})^2 \right\rangle} \approx \sqrt{2A_s} \left(\frac{1}{3}\right) \frac{a}{a_{\text{kin}}}. \quad (6.36)$$

This becomes $\mathcal{O}(1)$ when

$$\boxed{a/a_{\text{kin}} \gtrsim 10^5.} \quad (6.37)$$

This implies that if the ALP field scales as kination more than $\ln(10^5) \sim 10$ e-folds, then the ALP field cannot be considered homogeneous anymore. A similar bound applies if the ALP drives a kination era. The bound eq. (6.37) is an important constraint when considering UV completions [2], in particular models where the ALP temporarily drives a kination era in the early universe, enhancing primordial GW signals. Such bound on the total duration of a kination era was overlooked in previous literature on kination, as discussed in [37]. This bound leads to important constraints on the model implementations, which we explore in part III.

As we have stated previously, in this work, we will assume that all the modes relevant for fragmentation are super-horizon at a_{kin} . This assumption requires that

$$\frac{k}{a_{\text{kin}} H_{\text{kin}}} < 1 \quad \Rightarrow \quad \kappa_* \frac{m_a(T_*)}{H_*} \frac{a_{\text{kin}}}{a_*} < 1. \quad (6.38)$$

If we demand that this condition is satisfied for the relevant modes for fragmentation, i.e. $\kappa_* \sim \mathcal{O}(1)$, then we need to demand that

$$\frac{m_a(T_*)}{H_*} \lesssim \frac{a_*}{a_{\text{kin}}}. \quad (6.39)$$

The quantity a_*/a_{kin} depends on the specifics of the UV completion, however we have shown that the homogeneity of the ALP field prior to fragmentation requires $a_*/a_{\text{kin}} \gtrsim 10^5$ which puts the bound $m_a(T_*)/H_* \lesssim 10^5$. In the next section, we will show that this bound is satisfied in the region of the ALP parameter space where our analytical approximation is under control.

With all these assumptions above, we can fix the field power spectrum at the onset of fragmentation by

$$P_\theta(k; a_i) = \frac{2\pi^2}{k^3} \left(\frac{1}{3}\right)^2 A_s \left(\frac{\dot{\Theta}_i^2}{H_i^2}\right). \quad (6.40)$$

As we will see explicitly in the coming sections, the choice of a_i is not relevant for the final result, as long as it is early enough that $\dot{\Theta} \propto a^{-3}$, and any fragmentation prior to a_i can be neglected.

7 ALP dark matter from fragmentation

7.1 Overview of the fragmentation regions

We categorize the ALP parameter space by dividing it into different regions according to whether and when fragmentation takes place. There are four different scenarios:

1. **Standard misalignment:** In this case the onset of oscillations is not delayed from its conventional value $m_a(T_{\text{osc}}) = 3H(T_{\text{osc}})$, and the standard misalignment mechanism is at play.
2. **Kinetic misalignment with weak fragmentation:** The ALP field has a non-zero initial velocity such that the onset of oscillations is delayed, but the particle production is not strong enough so that the energy density in fluctuations is always subdominant to the zero mode.
3. **Fragmentation after trapping:** The ALP field is completely fragmented, but the fragmentation ends after it would have been trapped by the potential in the absence of fragmentation, i.e. $T_{\text{end}} < T_*$.
4. **Fragmentation before trapping:** The ALP field is completely fragmented, and the fragmentation ends before it would have been trapped by the potential in the absence of fragmentation. In other words the fragmentation is complete at $T_{\text{end}} > T_*$ where T_* is given by (5.29).

The boundaries between these regions depend strongly on the hierarchy between the axion mass and the Hubble rate at trapping, and have a very mild dependence on the other model parameters. In the rest of this section, we will calculate these boundaries while giving details on the properties of the fragmentation in each region.

An overview of these regions on the $[m_a, f_a]$ plane along with various model parameters can be seen in figures 7 and 8.

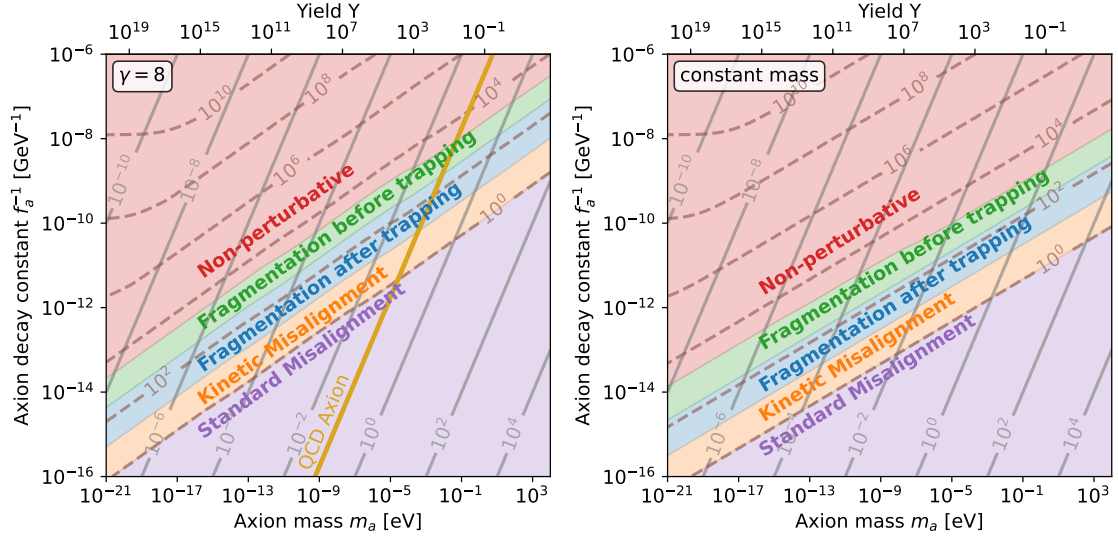


Figure 7. Overview of the fragmentation regions for temperature-dependent axion mass with $\gamma = 8$ (left plot), and for constant axion mass (right plot). The solid contours denote the zero-temperature barrier heights, while the dashed ones are the $m_*/3H_*$ contours, where T_* is given in eq. (5.28).

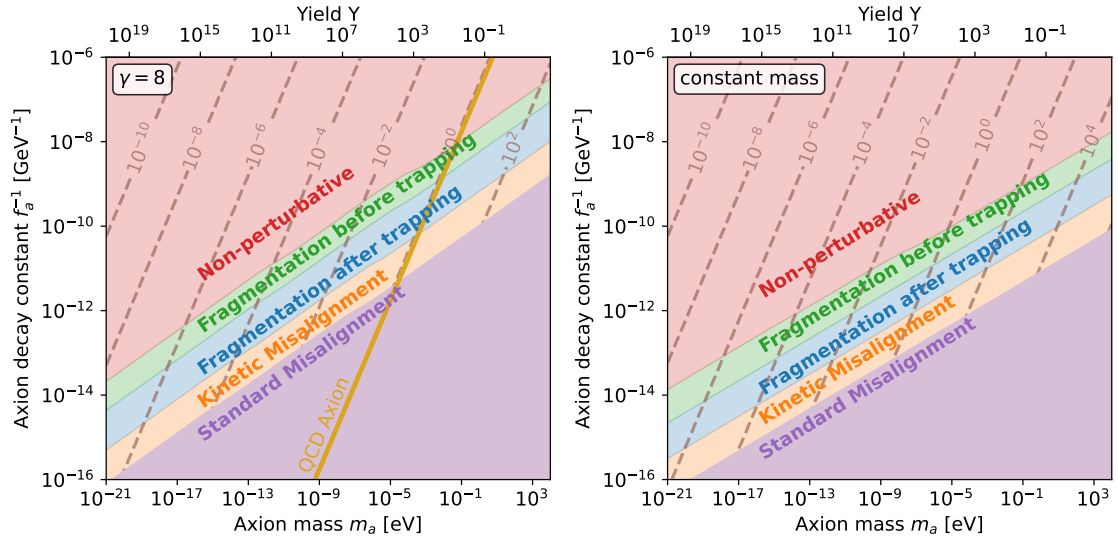


Figure 8. Same as figure 7, but the dashed contours now show the trapping temperature T_* in GeV, as given in eq. (5.28).

7.1.1 Boundary between regions 1 and 2:

In order to have kinetic misalignment, the onset of oscillations needs to be delayed so $m_a(T_*) > 3H(T_*)$. Therefore the boundary between the regions 1 and 2 is given by the condition

$$m_a(T_*) = 3H(T_*). \quad (7.1)$$

In a generic ALP model, the critical value of the axion decay constant $f_{a,\text{crit}}$ for a given zero-temperature axion mass m_a can be calculated via (2.8) and (5.27). It can approximately be expressed as

$$f_{a,\text{crit}}^{1-2} \approx 10^{12} \text{ GeV} \left[(3\pi)^{-12-2\gamma} \left(\frac{g_s(T_*)}{72} \right)^{2+\gamma} \left(\frac{m_a}{\text{eV}} \right)^{-4-\gamma} \left(\frac{h^2 \Omega_{\Theta,\text{today}}}{h^2 \Omega_{\text{DM}}} \right)^{8+2\gamma} \right]^{\frac{1}{16+3\gamma}} \quad (7.2)$$

For the QCD axion this value is

$$f_{a,\text{crit}}^{1-2} \approx 2.15 \times 10^{11} \text{ GeV}, \quad \text{for a QCD axion.} \quad (7.3)$$

This is in agreement with the result obtained in [27].

7.1.2 Boundary between regions 2 and 3:

To calculate this boundary, we need to find the region where the energy density in the fluctuation remains subdominant compared to the energy density in the homogeneous mode. In this region, the back-reaction can be neglected. The energy density in the homogeneous mode is

$$\rho_{\Theta} = 2f_a^2 m_a^2(T) \varrho, \quad (7.4)$$

while the energy density in the fluctuations is (6.26)

$$\rho_{\text{fluct}} = \frac{f^2}{4\pi^2} \int dk k^2 \left(\frac{k^2}{a^2} + m_a^2(t) \overline{\cos \Theta} \right) P_{\theta}(k) A_k^2(t) N_k^2(t). \quad (7.5)$$

Let us define a quantity $\Delta \equiv \lim_{t \rightarrow \infty} \rho_{\text{fluct}} / \rho_{\Theta}$ as the ratio between the two energy densities at late time limit. We refer to this quantity as the ‘‘efficiency’’ of the fragmentation. Then the transition from weak fragmentation to complete fragmentation should occur approximately when Δ reaches unity. This point will determine the boundary between the regions 2 and 3.

We are interested in the late time limit of Δ . In this limit $\overline{\cos \Theta} = 1$, and we can assume that all the modes which did grow during the parametric resonance become non-relativistic so that the momentum term in (7.5) can be ignored. The redshift factor $A_k^2(t)$ at late times can be approximated by

$$A_k^2(t) = \frac{\omega_k(t_i)}{\omega_k(t)} \left(\frac{a_i}{a} \right)^3 \approx \frac{k/a_i}{m_a} \left(\frac{a_i}{a} \right)^3 = \frac{k a_i^2}{m_a a^3}. \quad (7.6)$$

By plugging this result into (7.5) and using the expression for the initial power spectrum, we find that the energy density in the fluctuations approach an asymptotic value given by

$$\rho_{\text{fluct}} \rightarrow \frac{m_a m_a(T_*) f_a^2}{2} \left(\frac{1}{3} \right)^2 A_s \left(\frac{\dot{\Theta}_i a_i}{H_i a_*} \right)^2 \left(\frac{a_*}{a} \right)^3 \int d\kappa_* \exp \left(\frac{m_a(T_*)}{H_*} \mathcal{B}_{\kappa}^{\infty} \right), \quad (7.7)$$

where $\mathcal{B}_{\kappa}^{\infty}$ is the asymptotic value of the growth factor defined in (6.13). The late time limit of the energy density of the homogeneous mode can be found by using the fact that for $\varrho \ll 1$ we have

$$\varrho \approx \frac{4}{\pi} \tilde{I} = \frac{4}{\pi} \frac{m_a(T_*)}{m_a} \left(\frac{a_*}{a} \right)^3. \quad (7.8)$$

With this result, the efficiency factor takes the form

$$\Delta = \frac{\pi}{16} \left(\frac{1}{3}\right)^2 A_s \left(\frac{\dot{\Theta}_i a_i}{H_i a_*}\right)^2 \int d\kappa_* \exp\left(\frac{m_*}{H_*} \mathcal{B}_\kappa^\infty\right). \quad (7.9)$$

Note that at early time $\dot{\Theta} \propto a^{-3}$, therefore the pre-factor in this expression does not depend on the choice of the initial time provided that it is early enough so that $\dot{\Theta} \propto a^{-3}$, and any fragmentation prior to this time can be neglected. By expressing the axion velocity $\dot{\Theta}_i$ in terms of the yield parameter introduced in section 5.1 as $\dot{\Theta}_i = s(T_i)Y/f_a^2$ where $s(T_i)$ is the entropy density at T_i , we can derive the following result:

$$\frac{\dot{\Theta}_i a_i}{H_i a_*} \approx \frac{4\pi}{3} \sqrt{\frac{g_s(T_*)}{10}} \frac{YT_* m_{\text{Pl}}}{f_a^2}. \quad (7.10)$$

When arriving at this, we have neglected the difference between the effective degrees of freedom at T_i and T_* , and also assumed that the effective degrees of freedoms in the entropy and the energy density are the same. So our final expression for the efficiency factor is

$$\Delta = \frac{\pi^3}{10} \left(\frac{1}{9}\right)^2 A_s g_s(T_*) \left(\frac{YT_* m_{\text{Pl}}}{f_a^2}\right)^2 \int d\kappa_* \exp\left(\frac{m_a(T_*)}{H_*} \mathcal{B}_\kappa^\infty\right). \quad (7.11)$$

With this factor, we can calculate the critical point at which the efficiency becomes unity. We find that across the parameter space the transition happens around $m_a(T_*)/H_* \approx 42$ for $\gamma = 8$, and around $m_a(T_*)/H_* \approx 38$ for constant axion mass ($\gamma = 0$). This leads to a critical axion decay constant given by

$$f_{a,\text{crit}}^{2-3} \approx 10^{12} \text{ GeV} \left[(40\pi)^{-12-2\gamma} \left(\frac{g_s(T_*)}{72}\right)^{2+\gamma} \left(\frac{m_a}{\text{eV}}\right)^{-4-\gamma} \left(\frac{h^2 \Omega_{\Theta,\text{today}}}{h^2 \Omega_{\text{DM}}}\right)^{8+2\gamma} \right]^{\frac{1}{16+3\gamma}} \quad (7.12)$$

For the QCD axion, this boundary corresponds to a decay constant of

$$f_{a,\text{crit}}^{2-3} \approx 1.5 \times 10^{10} \text{ GeV}, \quad \text{for a QCD axion.} \quad (7.13)$$

7.1.3 Boundary between regions 3 and 4:

In both regions the parametric resonance is efficient enough for complete fragmentation. The question here is in which regime fragmentation becomes efficient prior to trapping of the field. At times much earlier than trapping, i.e., for $\varrho \gg 1$, we can simplify the fragmentation calculation significantly. In this regime, the analysis we have outlined in section 6 reproduces the results of [29], which we showed in detail in Appendix C of [1]. In this regime, the particle production by fragmentation is efficient when

$$H(T) \lesssim \frac{\pi m_a^4(T)}{4\dot{\Theta}^3(T)}. \quad (7.14)$$

To a good approximation, exponential particle production begins if the inequality (7.14) is satisfied before the trapping. Very quickly after the onset of exponential particles production, the back-reaction becomes relevant. This takes place when [1, 29]

$$H(T) \lesssim \frac{\pi m_a^4(T)}{4\dot{\Theta}^3(T)} \frac{1}{\ln \alpha^{-1}(T)}, \quad (7.15)$$

where for adiabatic initial conditions

$$\alpha(T) = \left(\frac{4\pi}{9}\right)^2 \left(\frac{g_s(T_*)}{80}\right)^2 \left(\frac{YT_* m_{\text{Pl}}}{f_a^2}\right)^2 A_s \left(\frac{T}{T_*}\right)^2. \quad (7.16)$$

If the inequality (7.15) is saturated before trapping, then the energy density of the homogeneous mode is efficiently converted into fluctuations, and the fragmentation is completed in a short amount of time. This criterion is what we use to determine when fragmentation is efficient.

To estimate the temperature at which the back-reaction becomes prominent, we can replace $\dot{\Theta}(T) \approx s(T)Y/f^2$ and $m_a(T) = m_a(T_*)(T_*/T)^{\gamma/2}$ in (7.15) to get an equation in terms of the temperature T and other model parameters. Let T_s denote the solution of this equation. For fragmentation to be completed before trapping, this temperature T_s should be larger than the trapping temperature T_* . Therefore, region 4 can be approximately defined by the condition $T_s \gtrsim T_*$. Across the parameter space, the transition happens around $m_a(T_*)/H_* \sim 9(5) \times 10^2$ for $\gamma = 8(0)$. The expression for the critical decay constant can be approximated by

$$f_{a,\text{crit}}^{3-4} \approx 10^{12} \text{ GeV} \left[(\mathcal{O}(1)\pi \times 10^2)^{-12-2\gamma} \left(\frac{g_s(T_*)}{72} \right)^{2+\gamma} \left(\frac{m_a}{\text{eV}} \right)^{-4-\gamma} \left(\frac{h^2 \Omega_{\Theta, \text{today}}}{h^2 \Omega_{\text{DM}}} \right)^{8+2\gamma} \right]^{\frac{1}{16+3\gamma}} \quad (7.17)$$

where the $\mathcal{O}(1)$ factor is γ -dependent. For the QCD axion, the critical decay constant at this boundary is

$$f_{a,\text{crit}}^{1-2} \approx 7.3 \times 10^8 \text{ GeV}, \quad \text{QCD axion.} \quad (7.18)$$

7.1.4 Breakdown of perturbativity for larger m_*/H_*

Finally we discuss what happens in the parameter space where m_*/H_* is even larger. For this, we revisit the equation of motion for the homogeneous mode (5.5), but we also include the back-reaction at leading order. Then the equation of motion is modified to [29]

$$\ddot{\Theta} + 3H\dot{\Theta} + m_a^2(T) \sin \Theta - \frac{1}{2} m_a^2(T) \sin \Theta \langle (\delta\theta)^2 \rangle = 0. \quad (7.19)$$

The last term is responsible for the back-reaction of fluctuations, and as a result for the completion of the fragmentation. All of our analysis depends on the assumption that this term is negligible before the onset of fragmentation. Therefore we should check whether the adiabatic initial conditions for the mode functions (6.33) do not violate this assumption at the beginning of fragmentation.

By using (6.33), the variance can be calculated by

$$\langle (\delta\theta)^2 \rangle = \int \frac{d^3k}{(2\pi)^3} |\theta_k|^2 \sim \left(\frac{1}{3} \right)^2 A_s \left(\frac{\dot{\Theta}}{H} \right)^2, \quad (7.20)$$

where we have neglected the modes with $k/a_{\text{kin}} H_{\text{kin}} > 1$, and assumed $\ln(a/a_{\text{kin}}) \sim \mathcal{O}(1)$. At early times, the background evolution is dominated by the first two terms in (7.19) since the mass is negligible. However, if the variance becomes much larger than unity at these times, it can affect the slow-roll evolution much earlier than the fragmentation does. To avoid this case, we demand that

$$3H(T)\dot{\Theta}(T) > \frac{m_a^2(T)}{2} \left(\frac{1}{3} \right)^2 A_s \left(\frac{\dot{\Theta}(T)}{H(T)} \right)^2 \quad (7.21)$$

until the onset of fragmentation T_{frag} given by the solution of (7.14). This puts a strong bound on the parameter space which can roughly be approximated by

$$\frac{m_a(T_*)}{H_*} \lesssim \mathcal{O}(1-10) \times 10^3, \quad (7.22)$$

where the factor depends on γ . We want the stress that this result does not exclude the parameter space where this bound is violated. It just implies that the analysis we describe might not be

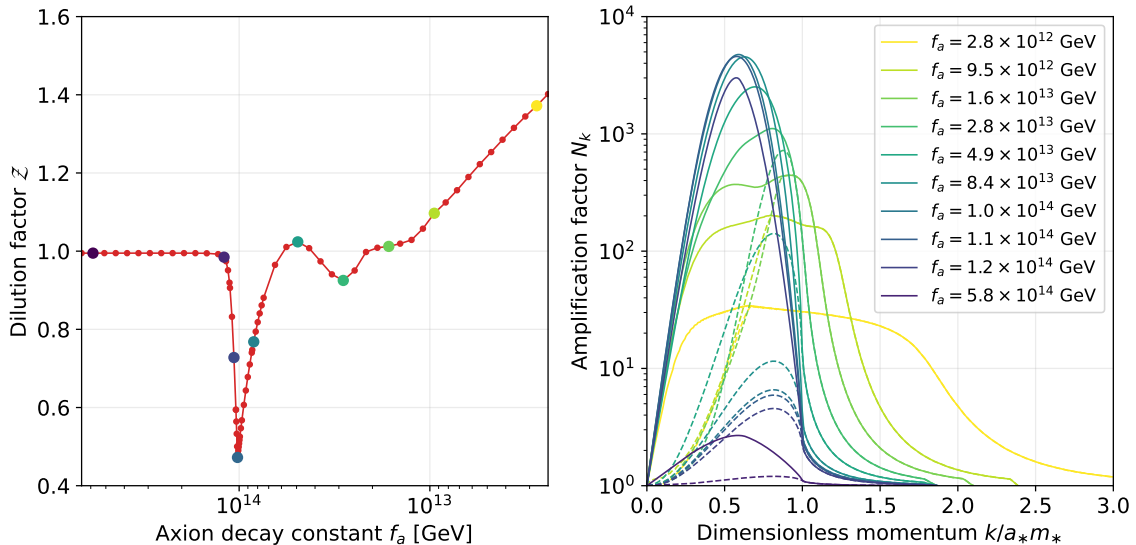


Figure 9. *Left:* Dilution factor (7.24) as a function of the axion decay constant for an ALP with constant mass $m_a(T) = 10^{-15}$ eV. *Right:* Amplification factor N_k (6.13) as a function of the dimensionless momentum $k/a_* m_*$. The solid lines show the spectrum at the end of fragmentation, while the thin dashed lines together with the shaded regions show the spectrum when the ALP is trapped. The line colors match with the colored dots on the left plot.

reliable, and a dedicated study is needed. Here in part II, we concentrate on the region of the parameter space where (7.22) is satisfied⁹.

7.2 ALP relic density with fragmentation

If the ALP field is completely fragmented, then all the energy density in the homogeneous mode gets transferred into the fluctuations. This will have an effect on the relic density today, since the redshift of the ALP energy density is not necessarily the same with and without fragmentation. Naive expectation is that the energy density will be diluted slightly since the modes that are enhanced exponentially are mildly relativistic right after fragmentation. Therefore it is natural to ask whether this effect is significant or not.

At late times, the ALP energy density is given by

$$\rho_\theta = \bar{\rho}_\theta + \frac{f_a^2}{2} \int \frac{d^3k}{(2\pi)^3} \left[\dot{\theta}_k^2 + \left(\frac{k^2}{a^2} + m_a^2 \right) \theta_k^2 \right], \quad (7.23)$$

where $\bar{\rho}_\theta$ is the energy density remaining in the homogeneous mode which is negligible in the case of complete fragmentation. Let ρ_Θ be the energy density without fragmentation. We define the dilution factor as

$$\mathcal{Z} \equiv \frac{\rho_\theta}{\rho_\Theta}, \quad (7.24)$$

which will be relevant for the section on the gravitational-wave signal. Unfortunately, it is not possible to calculate this factor precisely without a proper lattice simulation. In the non-linear regime after the fragmentation, the self-interactions between the enhanced momentum modes can modify the momentum spectrum of the fluctuations which can cause $\mathcal{O}(1)$ modifications in the dilution factor. We leave the careful study of this dilution factor with lattice simulations for future work.

⁹We have also checked whether the variance term can dominate the mass term in (7.19) before the onset of fragmentation. This yields a weaker bound given by $m_a(T_*)/H_* \lesssim \mathcal{O}(1) \times 10^5$.

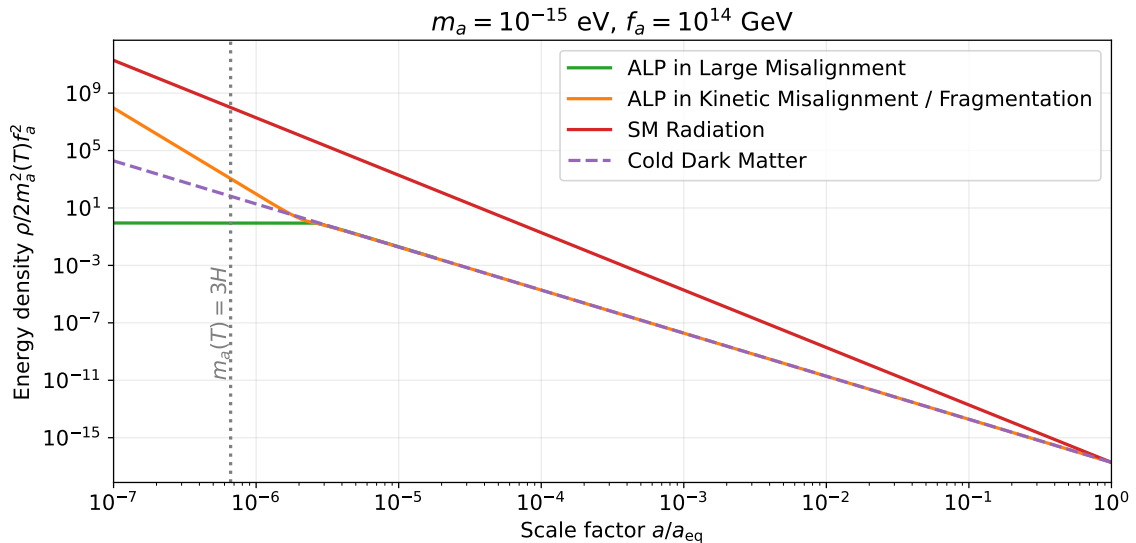


Figure 10. Evolution of the energy density of the axion in the large misalignment case (where the field is initially frozen and then redshifts as matter once oscillations turn on) compared to the kinetic misalignment/fragmentation case (a^{-6} redshifting preceding the matter era, with almost no time for a radiation-like a^{-4} behaviour in between).

At least we show in fig. 9 our estimate for \mathcal{Z} . Here, we assume an ALP with constant mass, fixed it to be $m_a = 10^{-15}$ eV, and show the behaviour of the dilution factor as a function of the decay constant f_a on the left plot. For large decay constants, the fragmentation is inefficient, therefore we have $\mathcal{Z} \approx 1$. As the fragmentation becomes more efficient, i.e. for smaller decay constants, more and more energy is transferred to the modes that are mildly relativistic after fragmentation. As a result, the dilution factor decreases confirming our expectations. However, we observe the opposite trend at even lower decay constants where our calculation predicts dilution factors larger than one, meaning that the relic density is *enhanced* by the fragmentation. This might be seen counter-intuitive at first, but it is not. For small decay constants, most of the fragmentation occurs before trapping as we can see from the right plot of fig. 9. Recall that the energy density of the homogeneous mode scales as $\rho_{\Theta} \propto a^{-6}$ in this regime. During fragmentation, the energy density is transferred to the fluctuations that redshift slower compared to a^{-6} . This yields to an enhancement in the relic density. Even though the behaviour of the dilution factor can be understood for large and small decay constants, we cannot derive conclusions about the behaviour in between with our simplified calculation. The self-interactions between the excited modes can modify the momentum distribution and therefore the dilution factor. This effect cannot be captured by our approximation and require a lattice calculation.

We also show the evolution of the energy density of the axion in figure 10. This clearly shows that there is almost no time for a radiation-like equation of state. The produced axions are not very relativistic when they are produced and very quickly cool down so that we transit quickly from a kination-like equation of state to a matter-like equation of state for the axion.

8 Constraints on the ALP parameter space due to fragmentation

We need to impose a number of constraints to ensure that the fragmentation process does not spoil existing cosmological observations. These are the following¹⁰:

8.1 Structure formation

If all of the dark matter is made up of fragmented axions, they will need to be sufficiently cold at matter-radiation equality in order to be consistent with structure formation. This puts a constraint on the axion velocity v_{eq} at matter-radiation equality a_{eq} . In section 6, we learned that the momentum modes which grow most efficiently due to the parametric resonance are those with $k \sim a_* m_*$. Then we can estimate the axion velocity at matter-radiation equality as

$$v_{\text{eq}} \sim \frac{k/a_{\text{eq}}}{m_a} \sim \left(\frac{a_*}{a_{\text{eq}}} \right) \left(\frac{m_a(T_*)}{m_a} \right). \quad (8.1)$$

To get precise constraints from structure formation would require involved numerical simulations, which is beyond the scope of this work. Instead, we will use the bound $v_{\text{eq}} < 10^{-3}$ which is commonly considered in the literature [204]. So we demand

$$\left(\frac{a_*}{a_{\text{eq}}} \right) \left(\frac{m_a(T_*)}{m_a} \right) \lesssim 10^{-3}. \quad (8.2)$$

It should be noted that in [35] the authors also consider a similar structure formation constraint. The condition used in that reference yields a constraint that is somewhat stronger than the one implied by our condition. However, since either condition is at most comparable with the BBN constraint, we do not devote further attention to the matter.

8.2 Big Bang Nucleosynthesis (BBN)

The presence of an additional energy component at the time of BBN will modify the expansion rate and therefore be subject to constraints from observations of primordial helium-4 and deuterium abundances. Conventionally, the energy densities of new, dark and relativistic particle species are recast in terms of a neutrino density to be constrained through the effective number of neutrino species N_{eff} . However, this recast is only possible when the new energy density redshifts like radiation. Because our axions can either behave as cold dark matter, radiation or be in a phase with a kination-like scaling, N_{eff} does not provide a convenient language to cast our BBN constraints in. Instead, we work directly in terms of energy densities and run a full numerical BBN calculation in the presence of such an additional dark density component. Such a calculation can be done with the built-in routines of the numerical code `AlterBBN 2.2` [205].

The routine `alter_standmod` in `AlterBBN 2.2` permits the addition of an additional energy density component of the form

$$\rho_D(T) = \zeta \rho_\gamma(T_{\text{BBN}}) \left(\frac{T}{T_{\text{BBN}}} \right)^n, \quad (8.3)$$

where $\zeta = \rho_D(T_{\text{BBN}})/\rho_\gamma(T_{\text{BBN}})$ is the ratio of the additional energy density to the photon density as measured at the reference temperature $T_{\text{BBN}} = 1 \text{ MeV}$ and n is the scaling exponent such that $\rho_D \propto a^{-n}$. Expressed in terms of the total radiation density the parameter ζ is

$$\zeta = \frac{\rho_D}{\rho_\gamma} \Big|_{\text{MeV}} = \frac{g_\rho(T_{\text{BBN}})}{2} \frac{\rho_D}{\rho_{\text{rad}}} \Big|_{\text{MeV}} \approx 5.29 \frac{\rho_D}{\rho_{\text{rad}}} \Big|_{\text{MeV}}. \quad (8.4)$$

¹⁰Note that the process of axion fragmentation does not generate a domain wall problem [203]. In part III, we will encounter domain wall problems; these are not generated by fragmentation but rather by the evolution assumed by the UV completion.

If we assume that the ALP energy density scales cleanly with either $n = 3, 4$ or 6 during BBN, we can then use this routine to calculate the produced helium-4 and deuterium abundances. These abundances are then compared to the most current PDG constraints [206], which at 1σ uncertainty are

$$Y_p = 0.245 \pm 0.003, \quad (8.5)$$

$$D/H \Big|_p = (2.547 \pm 0.025) \times 10^{-5}, \quad (8.6)$$

where Y_p and $D/H|_p$ are the helium-4 and deuterium fractions respectively. Using these constraints with the `AlterBBN` routine `alter_standmod` we obtain 2σ constraints on ALP densities. Depending on the scaling of the axion relic at $T \sim 1$ MeV, the constraints are the following:

- $\rho_D \propto a^{-3}$: Cold dark matter is not constrained by BBN.
- $\rho_D \propto a^{-4}$: Constrained by BBN if $\rho_D/\rho_{\text{rad}}|_{1 \text{ MeV}} \gtrsim 3.2 \times 10^{-2}$.
- $\rho_D \propto a^{-6}$: Constrained by BBN if $\rho_D/\rho_{\text{rad}}|_{1 \text{ MeV}} \gtrsim 1.9 \times 10^{-1}$.

Only the last constraint is relevant for us. Taking $\rho_D = \rho_\Theta$, this corresponds to a lower bound of

$$T_* \gtrsim 20 \text{ keV}. \quad (8.7)$$

By using (7.10) and assuming that the field is still rolling at T_{BBN} the last condition above becomes equivalent to

$$\frac{4\pi^2}{135} \left(\frac{g_s(T_*)}{g_s(T_{\text{BBN}})} \right)^{1/3} g_s(T_{\text{BBN}}) \left(\frac{YT_{\text{BBN}}}{f_a} \right)^2 \lesssim 0.19. \quad (8.8)$$

Expressing the yield in terms of the zero-temperature axion mass by using (5.20) we find that the BBN constraint implies a bound on the zero-temperature barrier height:

$$\Lambda_{b,0} \gtrsim 9 \times 10^{-7} \text{ GeV} \times \left(\frac{h^2 \Omega_{\Theta,0}}{h^2 \Omega_{\text{DM}}} \right)^{1/2}. \quad (8.9)$$

This result is independent of all other model parameters including the temperature-dependence of the axion mass. We observe that this bound is always stronger than the constraint from structure formation.

8.3 ALP parameter space with existing and future experiments

We now show the fragmentation regions of section 7.1, and the model constraints of section 8 on the axion parameter space together with the long list of experimental constraints/projections of diverse nature discussion in section 3.2. This results in figures 11 and 12.

The Kinetic Misalignment Mechanism (KMM) is not the only method to get ALP dark matter with lower decay constants. Another possibility is the Large Misalignment Mechanism (LMM) where one chooses an initial angle very close to the top [197, 198]. The distinction between the Standard and Large misalignment is arbitrary. In this work, we will define the LMM region as the region above the orange line in figures 11 and 12, as this region requires tuning in the misalignment angle to be realized by misalignment without initial kinetic energy. The required tuning in the LMM scenario rapidly increases with decreasing f_a . This is illustrated by the $|\pi - \theta_i| \sim 10^{-1}$ and the $|\pi - \theta_i| \sim 10^{-12}$ lines on fig. 12.

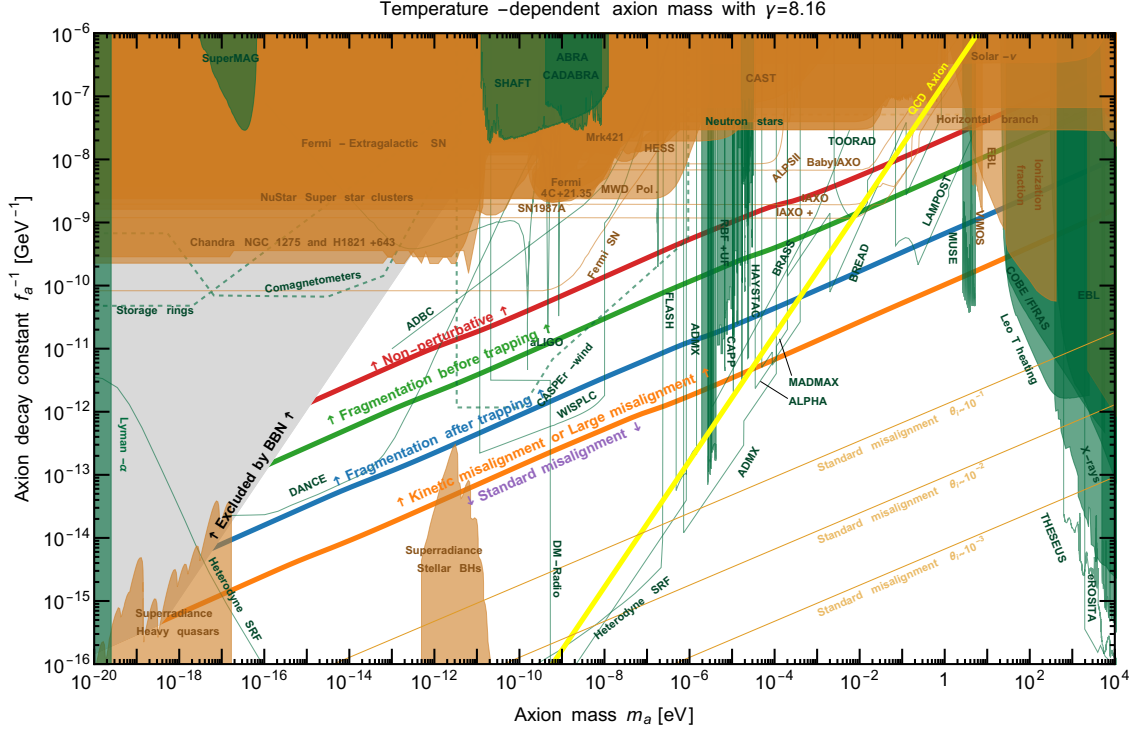


Figure 11. *Parameter space for ALP dark matter. Every point in the white region can have the correct relic abundance to explain DM. We assume a temperature-dependent axion mass with $\gamma = 8$, see (2.8) for definition, consistent with the QCD axion. The orange line separates the Standard and Kinetic Misalignment regions. The region above the orange line can also be reached without the initial kinetic energy, but by choosing the initial angle very close to the top of the potential, i.e. Large Misalignment Mechanism. Above the blue line, the fragmentation is efficient enough so that all the energy density is transferred from the homogeneous mode to the fluctuations. Above the green line, the fragmentation becomes efficient before the ALP gets trapped by the potential; see section 7.1 for the precise definitions. Above the red line, the variance of the ALP angle becomes larger than unity before the onset of fragmentation, so our calculation cannot be trusted. In the gray region, the ALP field is rolling during BBN with a large enough kinetic energy so that it spoils the BBN predictions, see eq. (8.9). Thin lines correspond to experimental projections. The bounds/projections on the axion-photon coupling are translated into bounds on the axion decay constant by assuming a KSVZ-like coupling given in (3.18). Orange constraints apply to any ALP while the green ones assume the ALP is DM.*

9 Gravitational waves

In the case of efficient parametric resonance, the ALP field at the end of fragmentation can have significant anisotropic stress, and this stress can source gravitational waves (GW). A precise evaluation of this process is not the topic of this study, and will be studied in a future work. Here, we will present a very rough estimate based on the method introduced in [207]. A more detailed analysis can be found in [208].

We start by discussing the frequency of the gravitational waves. Today's frequency is related to the comoving momentum k by

$$\nu = \frac{1}{2\pi} \frac{k}{a_{\text{today}}} = \frac{1}{2\pi} \frac{k}{a_{\text{emit}}} \frac{a_{\text{emit}}}{a_{\text{today}}}, \quad (9.1)$$

where a_{emit} is the scale factor at which the GW is emitted. By assuming that emission happens

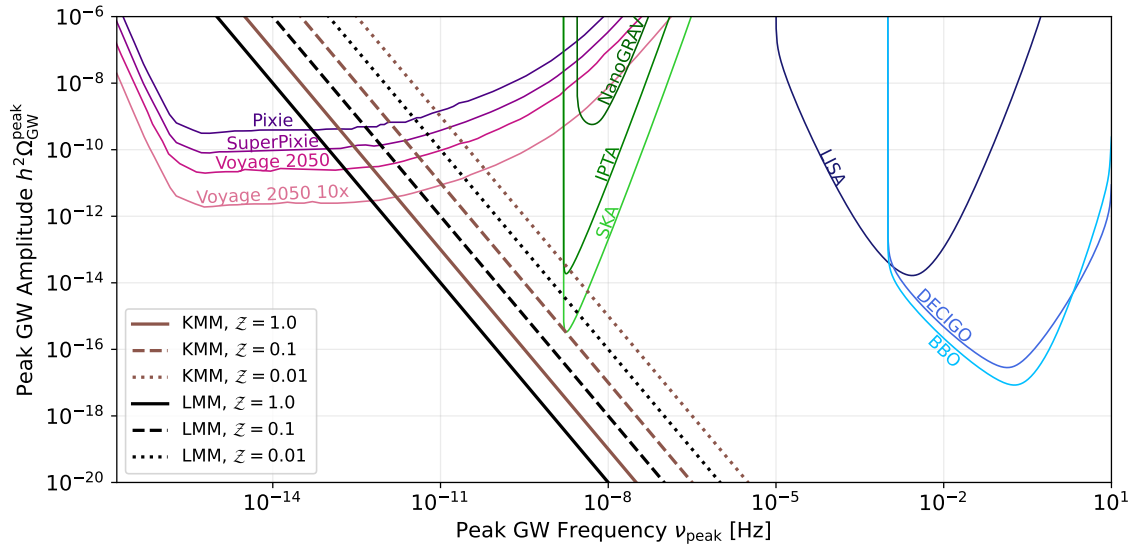


Figure 13. Peak amplitude and frequency of gravitational waves induced by axion fragmentation occurring in kinetic misalignment mechanism (KMM) versus large misalignment mechanism (LMM) according to eq. (9.11). The lines assume constant axion mass and should not be understood as GW spectra. The predictions are compared to the sensitivity of future experiments. The expression of ν_{peak} in terms of the axion mass and decay constant is given by eq. (9.4), see contours in fig. 14 and 15.

After the emission, the energy density in gravitational waves dilutes as radiation. Today's GW amplitude is given in terms of the amplitude at emission as

$$\Omega_{\text{GW},0} = \frac{\rho_{\text{GW},0}}{\rho_{\text{crit},0}} = \frac{\rho_{\text{GW,emit}}}{\rho_{\text{crit,emit}}} \frac{\rho_{\text{GW},0}}{\rho_{\text{GW,emit}}} \frac{\rho_{\text{crit,emit}}}{\rho_{\text{crit},0}} = \Omega_{\text{GW,emit}} \left(\frac{a_{\text{emit}}}{a_0} \right)^4 \left(\frac{H_{\text{emit}}}{H_0} \right)^2. \quad (9.6)$$

Let k_{peak} be the comoving momentum corresponding to the peak frequency ν_{peak} . Ref. [207] gives the following estimate for the peak amplitude at emission:

$$\Omega_{\text{GW,emit}}^{\text{peak}} \sim \frac{64\pi^2}{3m_{\text{Pl}}^4 H_{\text{emit}}^2} \frac{\rho_{\theta,\text{emit}}^2}{(k_{\text{peak}}/a_{\text{emit}})^2} \frac{\alpha^2}{\beta}, \quad (9.7)$$

where $\alpha \lesssim 1$ roughly measures the fraction of the energy stored in the fluctuations, and $\beta \gtrsim 1$ is the typical logarithmic width of the spectrum of fluctuations in momentum space. We set both of them to unity for our estimates. Again we assume that the GW emission takes place at trapping. Then, the energy density of the ALP field at emission is

$$\rho_{\theta,\text{emit}} \approx 2m_a(T_*)^2 f_a^2. \quad (9.8)$$

Also, the peak momentum becomes $k_{\text{peak}} = \kappa_* a_* m_a(T_*) \sim a_* m_a(T_*)$. Then (9.7) is simplified to

$$\Omega_{\text{GW,*}}^{\text{peak}} \sim \frac{256\pi^2}{3} \left(\frac{m_a(T_*)}{H_*} \right)^2 \left(\frac{f}{m_{\text{Pl}}} \right)^4. \quad (9.9)$$

Evolving this amplitude until today by using (9.6) and (9.3) we obtain

$$\Omega_{\text{GW},0}^{\text{peak}} \sim 1.5 \times 10^{-15} \left(\frac{m_a(T_*)}{m_a} \right)^{2/3} \left(\frac{m_a}{10^{-16} \text{ eV}} \right)^{-2/3} \left(\frac{f}{10^{14} \text{ GeV}} \right)^{4/3} \mathcal{Z}^{-4/3}. \quad (9.10)$$

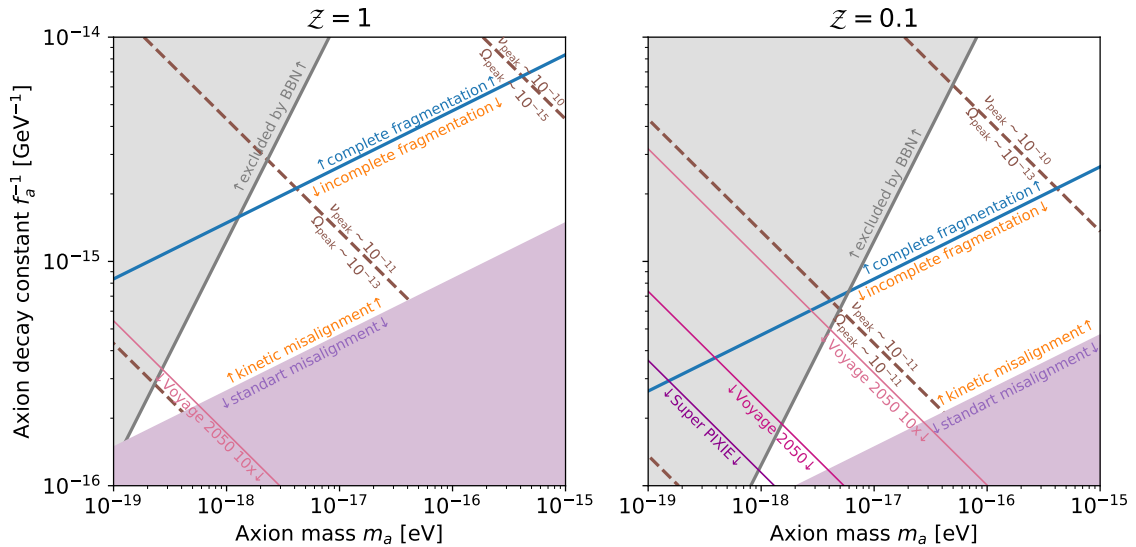


Figure 14. The peak frequency eq. (9.4) and peak amplitude eq. (9.11) contours in the ALP parameter space, and the regions that can be probed by gravitational waves (GW) induced by the fragmentation process by measuring the μ -distortions in the CMB [209]. On the left plot we assume that the fragmentation does not dilute the relic density, while the right plot assumes a factor of 10 dilution. Above the blue lines fragmentation is complete, and it is likely that the efficient GW production only happens in this region. The gray regions are excluded by the BBN constraints studied in section 8. GW prospects of kinetic fragmentation are thus severely constrained by BBN.

By combining this result with (9.4) we can obtain a simple relation between the peak frequency and the peak amplitude:

$$\Omega_{\text{GW},0}^{\text{peak}} \sim 10^{-35} \left(\frac{m_a(T_*)/m_a}{(\nu_{\text{peak}}/\text{Hz})\mathcal{Z}} \right)^2. \quad (9.11)$$

From this, we learn that ALP models with a constant mass have better prospects for an observable gravitational-wave signal. Secondly, the models with a lower peak frequency predict a larger GW amplitude. Finally, if there is an additional dilution in the energy density due to the fragmentation, the gravitational-wave amplitude is also enhanced. We show the contours of the peak GW frequency calculated via (9.4), and the peak GW amplitude calculated via (9.11) in the upper plots of figure 14.

A similar estimate has also been obtained in [208] by a different method, but their estimate is lower than ours by a factor of ~ 20 . This reference also presents a numerical calculation of the GW spectrum at the linearized level, and confirms that their estimate predicts the peak amplitude quite well. However, the GW amplitude can be enhanced significantly during the non-linear phase which cannot be captured by the linearized analysis [210]. Therefore, one can interpret our estimate and the estimate of [208] as optimistic and conservative respectively.

The power-law-integrated-sensitivity for the SKA mission [211] can reach up to $h^2\Omega_{\text{GW}} \sim 3 \times 10^{-16}$ at frequency $\nu \sim 2 \times 10^{-9}$ Hz assuming a signal-to-noise of 1 and an observation time of 20 years [212]. So even considering a constant mass and assuming a suppression of factor of 10, i.e. $\mathcal{Z} = 0.1$, our optimistic estimate tells us that the signal is barely observable. Recently, [209] did point out that the gravitational waves with frequencies much smaller than the ones probed by SKA might be observable by measuring the μ -distortions in the CMB via the experiments such as COBE/FIRAS [213, 214], PIXIE [215], SuperPIXIE [216], and Voyage 2050 [217]. The COBE/FIRAS experiment sets an upper limit on the μ -distortions ($\mu < 9 \times 10^{-5}$ 95%CL), while

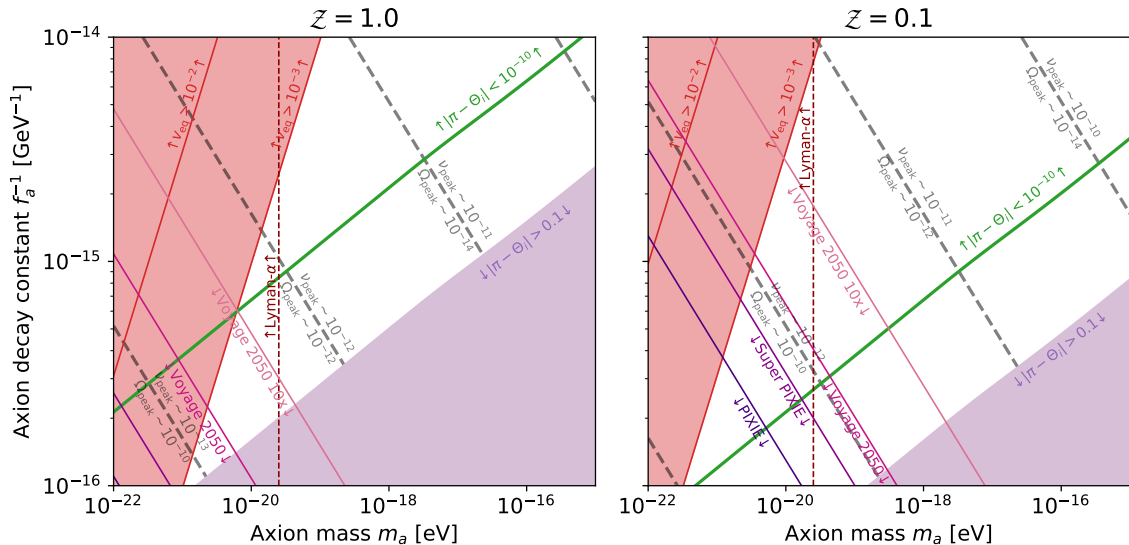


Figure 15. Analog of figure 14 but for the case of large misalignment. Unlike the case of kinetic misalignment there are no BBN constraints, however the requirement that the ALPs are not too warm during the matter-radiation equality gives a similar but weaker bound. In the purple shaded region on the lower right, the initial angle is not close to the top, and the standard misalignment is at play. The green line shows the parameter space where $|\pi - \Theta_i| \approx 10^{-10}$ above which the density contrast of the ALP field becomes $\mathcal{O}(1)$, and it is expected that complete fragmentation will happen above this line. It is likely that efficient GW production will happen only in this region.

the forecasted constraints for the experiments are $\mu < 3 \times 10^{-8}$ for PIXIE, $\mu < 7.7 \times 10^{-9}$ for SuperPIXIE, and $\mu < 1.9 \times 10^{-9}$ for Voyage 2050.

These upper bounds on the observable μ -distortions can be translated into lower bounds on the observable gravitational wave signal for a given frequency [209]. By using these bounds, we can estimate the potentially observable regions in the ALP parameter space which we show in figure 14. The curve labeled “Voyage 2050 10x” assumes an upper limit of $\mu < 1.9 \times 10^{-10}$, see [209] for details. We also show the critical lines separating complete and incomplete fragmentation (blue line), as well as the region which is excluded by the BBN constraints (gray region) discussed in section 8. Likely, the efficient GW production does happen only in the region where the fragmentation is complete. We conclude that the BBN bound severely constrains the parameter space which can be observable by GWs. The prospects do improve if the fragmentation efficiently dilutes the relic density of ALPs, since this will increase the energy budget for the GWs without overclosing the universe. Precise estimations of the dilution factor, peak GW amplitude and the GW frequency require a dedicated lattice analysis.

Since the BBN constraint is the main obstacle to get an observable gravitational wave signal, one might wonder what happens in a model which is immune to the BBN constraint. The Large Misalignment Mechanism is an example of such a model since the energy density prior to the oscillations does not redshifts as a^{-6} . Therefore, we did repeat the above analysis for the Large Misalignment scenario, and obtained a result which is very close to (9.11) except an $\mathcal{O}(1)$ factor which depends very mildly on the value of $|\pi - \Theta_i|$. We show the results in figure 15. Even though the BBN constraints are absent, there is still a similar but weaker constraint that the ALPs should not be too warm at matter-radiation equality in order to be consistent with the structure formation. The green line shows approximately the critical initial angle $|\pi - \Theta_i| \approx 10^{-10}$ at which the density contrast of the ALP field becomes $\mathcal{O}(1)$. Above this line, the initial angle needs to be chosen

closer to the top of the potential, and we expect that complete fragmentation and the efficient GW production does happen in this region. We see that the absence of the BBN constraint opens up a sizable region which might be probed by future CMB surveys.

The ALP masses below $\sim 10^{-20}$ eV are constrained by various probes such as Lyman- α observations [109], and the galactic rotation curves [218–220]. However, it is possible that some of these constraints can be evaded as a consequence of the fragmentation [221].

We close this section by stressing another, and potentially much stronger, signal of GW that can arise in the Kinetic Misalignment Mechanism, and which is not related to fragmentation. While in this work we have assumed a radiation-dominated universe during the period when the equation of state of the axion scales as a^{-6} , it is also possible that the axion energy density dominates temporarily, inducing a kination era inside the radiation era. Such a kination era enhances primordial signals of GW from inflation and cosmic strings leading to striking peak features that are observable by upcoming interferometers such as LISA and the Einstein Telescope [35–37]. The occurrence of a kination era in specific UV completions and the precise parameter space region that lead to observable GWs are worked out in [37]. We return to this signal of amplified primordial gravitational waves in part III, where we compute the parameter space of dark matter implementations of axion kinetic misalignment which can potentially be probed by such a gravitational wave signal.

10 Summary of part II

In summary, the correct axion relic abundance to explain dark matter can be recovered naturally in the majority of the $[m_a, f_a]$ plane, if appropriate initial conditions can be realized. The region which is usually discarded because it leads to insufficient dark matter becomes open in the scenario of kinetic misalignment, in which oscillations are delayed due to the initial velocity of the axion field. This was pointed out in [27, 28] and the underlying framework was extensively discussed in [34, 222, 223] in some specific UV completions.

In the preceding sections, part II of this thesis, we have seen that one can no longer describe this regime in terms of the homogeneous zero-mode. In fact, the axion field entirely fragments in much of the KMM parameter space. We have provided a detailed analytical derivation of the phenomenon. One main implementation-independent consequence of fragmentation is a distinct prediction for dense compact mini-clusters, which was explored by Eroncel and Servant in [30]. There is also a stochastic gravitational wave background generated by axion fragmentation as we discuss in our section 9. However, this signal appears at extremely low frequency and is typically below the sensitivity of Pulsar Timing Arrays. For the very lowest masses of ALPs, the signal might be visible in CMB experiments such as SuperPIXIE [216] or Voyage 2050 [217]. In part III of this thesis we will return to the topic of gravitational waves and explore how another type of signal may be generated by amplification of primordial GW's. Our results apply generically to any ALP, including the QCD axion.

The main results of this part, which are UV completion independent, are summarised in the following expressions and figures:

- The temperature T_* when the field gets trapped by the barrier is given in eq. (5.28). Fig. 8 shows the contours of the trapping temperature.
- The duration of the kination-like scaling underlying the kinetic misalignment mechanism is constrained by eq. (6.37).
- Fig. 7 shows the contours of $m_*/3H_*$ at the time when the field gets trapped, which is orders of magnitude larger compared to the unit value that characterises the time when the field starts oscillating around its minimum in the standard misalignment mechanism.

- The amplification factor of a given mode depends crucially on the value of m_*/H_* as expressed in eq. (6.13). It is also plotted in fig. 9.
- Fig. 11 and 12 show the contours of the different regimes of dark matter production for respectively the constant axion mass and temperature-dependent mass cases. The expressions for these region boundaries are given in eqs. (7.2), (7.12), (7.17), and (7.22). The BBN constraint, $T_* > 20$ keV, translates into a bound on the size of the zero-temperature barrier, see eq. (8.9).
- All phenomenological implications of fragmentation and in particular the determination of the different regimes are controlled by the power spectrum eq. (6.33).
- Figure 9-left indicates that the impact of fragmentation on the relic abundance prediction from kinetic misalignment is relatively weak. It typically differs by a factor of order 1, we do not expect more than one order of magnitude effect. This can be understood as the momentum of the produced axions during fragmentation is of the same order as the axion mass (see fig. 9-right), higher modes are not excited. A precise estimate of the $\mathcal{O}(1)$ factor requires a lattice calculation which we will present in a future work.
- We estimated the GW signal in fig. 13 and the parameter space that can be probed this way in fig. 14. We showed that there are better prospects for GW from large misalignment in fig. 15, something which had not been mentioned in [198].

One key question that was not addressed in this part of the thesis is the origin of the initial axion velocity. We postponed this question so that we could perform an implementation independent analysis of kinetic misalignment, which revealed much interesting phenomenology. With the scenario of kinetic misalignment well motivated, we will now investigate the precise regions of ALP DM parameter space that can be supported by specific implementations of axion kinetic misalignment. This leads us to part III of this thesis.

Part III

Model implementations of axion kinetic misalignment

In part I, we motivated the axion as an attractive DM candidate. We noted that, in the standard paradigm, dark matter is underproduced in the regime most susceptible to experimental scrutiny. In part II, we studied how the kinetic misalignment mechanism accounts for dark matter in this low- f_a regime and discussed both implementation-independent constraints and the interesting phenomenology that arises from axion fragmentation. We observed that the kinetic misalignment mechanism holds much promise for axion dark matter. So far, we have remained agnostic as to how such a scenario can be realized. It is now time to address this question. In part III of this thesis, which is based on our paper [2], we will now investigate model implementations of the axion kinetic misalignment mechanism.

Scenarios of rotating scalar fields cannot be realized without some mechanism to start the rotation. We refer to the event that initiates the rotation as the *kick*. The historically most important family of kick mechanisms relies on higher-dimensional terms initially introduced for the Affleck-Dine mechanism [31, 32, 224]. Such Affleck-Dine-like kick mechanisms were used in the context of rotating axions by Co et al. [27, 33–35] and Gouttenoire et al. [36, 37]. Alternatives to Affleck-Dine-like kick mechanisms include the trapped misalignment mechanism [225, 226].

We here re-examine Affleck-Dine-like implementations of KMM. We map out the $[m_a, f_a]$ axion parameter space which can support axion dark matter produced by these mechanisms. We specify solutions directly in terms of UV model parameters, which allows us to present, for the first time, how $[m_a, f_a]$ regions of axion dark matter from kinetic misalignment is impacted by constraints on the kick mechanisms. This analysis also clarifies which UV parameters are required to account for dark matter through kinetic misalignment. Furthermore, in part II, we identified a new constraint on the rotation axion, which results from the growth of fluctuations during kination-like scaling of the axion. We now show how this condition significantly impacts the parameter space which can be supported by implementations of axion kinetic misalignment.

For simplicity, we in this part of the thesis restrict ourselves to axions with a QCD-like temperature dependence as specified by eq. (2.8) with $\gamma = 8.16$. Furthermore, we assume that non-linear effects do not modify the KMM relic in the regime identified in section 7.1.4, although this regime lies outside the range of validity of the analysis of Part II. Future study is required to determine the validity of this assumption.

Radial oscillations present a significant challenge to Affleck-Dine-like mechanisms. Such radial oscillations will, in general, be generated by the kick, and they will end up dominating the energy density of the universe if they are not damped out. In such a scenario, the saxion would behave as cold dark matter and dominate the axion. While this may itself be an intriguing possibility, such a scenario is outside of the scope of this thesis. We here instead focus on the scenario where these radial oscillations are damped such that the dark matter relic is composed of axions. We perform a detailed treatment of how damping affects KMM implementations, with particular attention to thermal effects. We take into account yield dilution from damping and describe how this affects solutions. This allows us to map out which damping temperatures are required to support the axion DM parameter space. We supplement the analytic treatment of damping, which was carried out in the literature [27, 33–35, 37], with a numerical solution of the Boltzmann equations that allows us to accurately take into account how damping is affected by the interplay between the new ALP sector and the SM plasma. Furthermore, the numerical solution enables us to follow the initial

conditions of the kick mechanisms and robustly map out how all resulting families of solutions populate the axion DM parameter space. We use this numerical solution to investigate realized damping temperatures, constraints, and which regions of parameter space may be detectable by amplified gravitational waves.

Part III is structured as follows: In section 11 we present general aspects of Affleck-Dine-like implementations of axion kinetic misalignment and provide an overview of the subsequent evolution and constraints. Section 12 is devoted to models with nearly-quadratic potentials, which we introduce a damping mechanism for in section 13. Section 14 is devoted to models with quartic potentials. In section 15 we summarize our results.

11 Essential components of a KMM implementation

Before we turn to a full implementation of the axion kinetic misalignment mechanism (KMM), we here first provide an overview of the essential components, phenomena, and constraints that we will encounter.

11.1 Definitions for the full complex scalar field

Until now, we have only been working with just the angular mode θ , which corresponds to the axion. However, the kick mechanisms, which we will later explore, rely crucially on radial dynamics. Therefore, we now consider the full complex scalar field, which we parametrize as

$$P = \frac{1}{\sqrt{2}}\phi e^{i\theta}. \quad (11.1)$$

We will no longer be working with fluctuations, so for the remainder of this thesis, we take ϕ and θ to be homogeneous fields. The radial mode ϕ is often referred to as the *saxion*, a name coined in the context of supersymmetric theories that we will adopt here.

For a non-static radial mode, the Noether charge of the axion shift symmetry, which we previously defined in eq. (5.18), generalizes to

$$n_{\text{PQ}} = \dot{\theta}\phi^2. \quad (11.2)$$

By analogy with a QCD axion, we will refer to n_{PQ} as the PQ charge and the complex scalar P as the PQ field. However, keep in mind that these labels are generally used by analogy only and do not by themselves imply that the axion in question is the QCD axion of the PQ solution.

11.2 Starting the rotation

The Affleck-Dine-inspired mechanisms, which we explore in this work, rely on two key features:

PQ-violating operators: Affleck-Dine-like mechanisms rely on higher-dimensional operators to start the axion rotation. Such higher-dimensional terms can originate from non-renormalizable terms in the super-potential of SUSY models [224]. Such terms lead to an angular potential,

$$\frac{A}{n} \frac{P^n}{M^{n-3}} + \frac{A}{n} \frac{P^{*n}}{M^{n-3}} = \frac{2A}{n2^{n/2}} \cos(n\theta + \delta_\theta) \frac{\phi^n}{M^{n-3}}, \quad (11.3)$$

which can drive a rotation. Here A is a parameter that sets the amplitude of the PQ-violating term, and M is a suppression scale that we typically set to m_{Pl} and δ_θ is the phase of the potential, which in general need not be aligned with the minimum at $\theta = 0$. The key idea is that even though this potential may be felt initially it rapidly becomes negligible as ϕ decreases, which leads the potential to generate a rotation rather than angular oscillations.

Large VEV driver: To probe such higher-dimensional operators, the VEV of the PQ field has to be driven to a large initial value, wherefore a mechanism that ensures such a large VEV is also an essential component. In this thesis, we will consider two mechanisms for driving the large VEV - one based on the Affleck-Dine mechanism [32, 224], and one based on de Sitter fluctuations.

A model with the above two components can start a rotation of the PQ field. The resulting kick can be found by solving the equation of motion [33, 224], which predicts

$$\dot{\theta}_{\text{kick}} = 2^{1-\frac{n}{2}} \frac{A\phi_{\text{kick}}^{n-2} \sin(n\theta_{\text{kick}})}{m_\phi M^{n-3}}, \quad (11.4)$$

where θ_{kick} is the initial angle by which θ is misaligned from the minimum set by eq. (11.3). This result is independent of the mechanism which sets the large VEV.

It is useful to parametrize the angular velocity imparted by the kick with the parameter ϵ [27], which gives the angular velocity as a fraction of the velocity required to maintain a circular orbit at the apoapsis of the field orbit. If we consider the equation of motion for the radial mode,

$$\ddot{\phi} + 3H\dot{\theta} + V' = \dot{\theta}^2 \phi, \quad (11.5)$$

then we see that the radial potential competes with the centrifugal term $\dot{\theta}^2 \phi$. The centrifugal term is exactly balanced against the potential when $\dot{\theta} = \sqrt{\phi^{-1}V'}$, which then corresponds to a perfectly circular orbit. We are working with large field values, so the usual mass defined by V'' is not the relevant parameter as this is defined at the potential minimum. Instead, we define the effective radial mode mass

$$m_\phi \equiv \sqrt{\phi^{-1}V'}, \quad (11.6)$$

as this is the parameter that enters the equation of motion and thus the relevant parameter far from the minimum [34]. The link to the potential parameters will be made in the next sections. We then observe that $\epsilon = \dot{\theta}_{\text{ap}}/m_\phi$ is a useful parametrization, where $\dot{\theta}_{\text{ap}}$ is the angular velocity at the apoapsis of an orbit. This is equivalent to normalizing the PQ charge to $n_\phi = V/m_\phi$, i.e.,

$$\epsilon = \frac{n_{\text{PQ}}/2}{n_\phi} = \frac{n_{\text{PQ}}/2}{V/m_\phi} = \frac{\dot{\theta}_{\text{ap}}}{m_\phi}, \quad (11.7)$$

where we assumed that the radial potential was a pure quadratic potential. Since the centrifugal term in the ϕ -EOM is balanced with the potential exactly for $\epsilon = 1$, this value corresponds to a perfectly circular orbit. $\epsilon = 0$ corresponds to purely radial oscillations and by construction $0 \leq \epsilon \leq 1$. Co et al. [27, 33, 34] note that an insufficiently circular orbit can lead to a parametric resonance in the radial mode. Throughout this thesis, we assume that ϵ is large enough that no such resonance takes place.

11.3 Evolution after the kick

After the kick, the radial mode is not in its ground state. It has a large VEV, $\phi_{\text{kick}} \gg f_a$, and unless the kick is perfectly circular, ϕ will oscillate around the equilibrium position. That is, P will be in an elliptic orbit. For the analysis of fragmentation and kinetic misalignment to apply, this initial condition must be relaxed to a non-oscillating state with $\phi = f_a$.

The need for damping: The energy density of radial oscillations behaves as cold dark matter. The radial energy density will therefore dominate that of the angular mode if left to oscillate. To realize dark matter from kinetic misalignment, it is necessary to dampen these oscillations.

Depending on when such damping takes place, the oscillations may or may not have a significant impact. If damping takes at a time where $\rho_\phi > \rho_r$, then damping significantly impacts the SM plasma. We denote the time where $\rho_\phi = \rho_r$ as T_{dom} , such that yield and cosmology are only significantly impacted if the damping temperature T_{damp} is in the regime $T_{\text{dom}} > T_{\text{damp}}$.

Before we specify the full model, including a specific damping mechanism, we wish to first understand the general case. We, therefore, structure our analysis into the following steps:

- 1) **Early damping**, which is restricted to the simpler case of $T_{\text{damp}} > T_{\text{dom}}$. In this case, the exact value T_{damp} does not impact the cosmology or DM relic directly and thus does not need to be specified beyond the assumption $T_{\text{damp}} > T_{\text{dom}}$.
- 2) **General damping**, which allows for either $T_{\text{damp}} > T_{\text{dom}}$ or $T_{\text{damp}} < T_{\text{dom}}$ but remains agnostic as to the damping mechanism. This contains step 1).
- 3) **Thermal damping**, where damping is implemented by a specific coupling to the SM plasma. This realizes a subset of the possible solutions identified in step 2).

The simplest approach, 1), is to assume that the oscillations are damped by an unspecified early mechanism that acts while the PQ field is subdominant to radiation such that no significant entropy injection takes place. In this case, the dynamics of the damping do not enter into the relic density such that a solution for the DM yield can be found directly from the kick parameters. The more general approach is to consider the damping temperature as a free parameter. To allow for lower damping temperatures, we take yield dilution into account and provide solutions for the DM yield also in the presence of yield dilution. Finally, we implement a specific realization of the damping mechanism, which allows us to properly solve the evolution of the field and map out the supported parameter space.

Evolution of the radial and angular modes: After the kick and the onset of the elliptic orbit of the field, the equilibrium value $\langle \phi \rangle$, about which ϕ oscillates, will be continuously decreasing. This relaxation can be characterized by PQ charge conservation. Once radial oscillations are damped the generally elliptic orbit with $0 \leq \epsilon \leq 1$ is relaxed to a circular orbit with $\epsilon = 1$, such that $\dot{\theta} \approx m_\phi$ after damping. Furthermore, it was shown in [37] that even before damping we have $\langle \dot{\theta} \rangle \approx m_\phi$. Therefore, $\langle n_{\text{PQ}} \rangle = m_\phi \phi^2 \propto a^{-3}$, which ensures that

$$\langle \phi \rangle \propto m_\phi^{-1/2} a^{-3/2}, \quad (11.8)$$

which is the same result we would expect from the WKB approximation. If the radial mass is time-dependent either as a consequence of radial dynamics or thermal effects, then this impacts the relaxation of the radial mode. Expressing such a possible time-dependence as $m_\phi \propto a^{-c_{m_\phi}}$, we can conclude that ϕ reaches the minimum at f_a at a_{kin} given by

$$\frac{\phi_{\text{kick}}}{f_a} \approx \left(\frac{a_{\text{kin}}}{a_{\text{kick}}} \right)^{(3-c_{m_\phi})/2}. \quad (11.9)$$

The analyses of fragmentation and kinetic misalignment [1, 27, 33, 34] assumes that the radial mode is relaxed to $\phi \approx f_a$ before T_* . Therefore, these analyses are only applicable for $T_{\text{kin}} > T_*$. If this condition is violated, then a fragmentation or kinetic misalignment analysis must be carried out in the presence of radial dynamics. Such an analysis is beyond the scope of this work, and we mark any parameter space in which this assumption is violated. However, it must be pointed out that there is no principal reason why fragmentation could not take place before the relaxation of the radial mode.

After ϕ reaches f_a , the angular mode needs no longer follow the equilibrium value $\dot{\theta} \approx m_\phi$. Applying again PQ charge conservation, $n_{\text{PQ}} = \dot{\theta}\phi^2 \propto a^{-3}$ we observe that once $\langle\phi\rangle = f_a$ the angular mode redshifts as $\dot{\theta} \propto a^{-3}$. Therefore,

$$\langle\dot{\theta}\rangle \propto a^{-c_\theta} \quad \text{where} \quad c_\theta = \begin{cases} c_{m_\phi} & \text{for } T > T_{\text{kin}}, \\ 3 & \text{for } T < T_{\text{kin}}. \end{cases} \quad (11.10)$$

As suggested by the subscript, T_{kin} marks the onset of kination-like scaling of the energy density of $\rho_a \propto \dot{\theta}^2 \propto a^{-6}$ such that a period of kination¹¹ is triggered if $\rho_a(T_{\text{kin}})$ dominates the energy density.

11.4 Constraints

Implementation-independent constraints: Any UV completion of kinetic misalignment must at least satisfy the implementation-independent constraints that we identified in part II. In particular, the constraints from structure formation and BBN discussed in section 8 as well as the abundance of axion searches discussed in section 3.2 constrain the parameter space.

Further BBN constraints: BBN also constrains damping further. We do not perform a detailed analysis of how the entropy injection associated with damping interferes with the observed ⁴He production, and we instead simply require damping to take place before BBN. Therefore, we also demand that

$$T_{\text{damp}} \gtrsim \text{MeV}. \quad (11.11)$$

In regimes in which the energy in the radial oscillations is small compared to the energy in radiation, this constraint may be overly strict, as it is conceivable that dilution would be small enough not to disturb BBN. Nevertheless, to be conservative, we impose this constraint regardless of how significant the entropy injection is.

Perturbativity: For the physics to remain perturbative all dimensionless couplings in the Lagrangian must be smaller than 4π . This, in particular, applies to the self-interactions induced by the nearly-quadratic potential, which we will study in section 12. We will return to this constraint after we have specified the model. As the quartic coupling must be very small in a viable quartic model of KMM, constraints from perturbativity will not play a role in section 14.

Homogeneity condition: In part II, we studied the final phases of evolution prior to trapping or fragmentation. During that investigation, we identified a constraint that had not previously been discussed in the literature. In particular, we in section 6.3 observed that primordial density fluctuations in the axion field grow relative to the homogenous mode during the period of a^{-6} kination-like redshift, which the θ must undergo in KMM. If this period of a^{-6} evolution is too long, the fluctuations may come to dominate the homogeneous mode. As we here seek to provide model implementations that lead to a rotating, homogenous axion field, we here treat this condition as a constraint. Nevertheless, such an inhomogeneous state may lead to interesting phenomenology or provide a relic in its own right, which Eroncel et al. will study in upcoming work [202].

The strength of this constraint depends on the amplitude of the primordial power spectrum A_s . From eq. (6.36), we see that growth in scale factor between the onset of kination-like scaling at a_{kin} and trapping at a_* is bounded by

$$A_s(k_{\text{kin}}) \lesssim \left(\frac{a_{\text{kin}}}{a_*}\right)^2. \quad (11.12)$$

¹¹*Kination* is a period of cosmological history characterized by a a^{-6} evolution of dominant component of the energy density of the universe. We use the term *kination-like* when a specific component of the energy density undergoes a^{-6} evolution without regard as to whether that component dominates the energy density.

If A_s is a flat spectrum specified by $A_{s,\text{Planck}} \approx 2.1 \times 10^{-9}$, then homogeneity is lost before axion fragmentation unless $a_*/a_{\text{kin}} \lesssim 10^5$. However, A_s is well-constrained only around the CMB pivot of comoving momenta $k = 0.05 \text{ Mpc}^{-1}$. We expect the comoving momentum scale of fluctuations entering the horizon around the onset of kination-like scaling, k_{kin} , to be many orders of magnitude larger than the CMB pivot scale. Therefore, $A_s(k_{\text{kin}})$ may very well deviate from $A_{s,\text{Planck}}$, which provides a possible way to escape this constraint. We will explore this potential loophole later. Until then, we will make the standard assumption that A_s is approximately $A_{s,\text{Planck}}$ on all scales.

Thermal ϕ relics: An interaction that is efficient enough to dampen the radial oscillations will also bring the radial mode into thermal equilibrium. Then, thermal fluctuations will be present in ϕ . These fluctuations are potentially dangerous for BBN and CMB. The thermal relic will decouple from the plasma if the interaction freezes out at a temperature T_{fo} and $T_{\text{fo}} > m_\phi$. After decoupling from the plasma, the thermal relic will have an energy density of

$$\rho_{\phi,\text{thermal}}(T_{\text{fo}}) \sim \left. \frac{\rho_r}{g_*} \right|_{T_{\text{fo}}}, \quad (11.13)$$

such that it will represent a fraction of the radiation density. It will be further diluted by reheating in the SM sector, i.e. changes in g_* and g_{*s} . Therefore, as long as the relic remains hot, it will not be problematic unless it freezes out below $\sim 200 \text{ MeV}$. However, if the relic becomes cold, it will grow relative to the radiation density and can become problematic.

At sufficiently large m_ϕ , such constraints are absent because the thermal relic decays into axions as soon as it is no longer relativistic, see appendix C. At sufficiently low m_ϕ the constraints are absent because the relic remains relativistic until after T_{eq} such that it both underproduces dark matter and does not disturb the CMB. The intermediate range is excluded by N_{eff} constraints from either BBN or CMB. Therefore, ϕ must satisfy either

$$m_\phi < \mathcal{O}(1) \times T_{\text{eq}} \quad \text{or} \quad \mathcal{O}(1) \times \frac{f_a^2}{m_{\text{Pl}}} < m_\phi, \quad (11.14)$$

where the exact values of the prefactors depend on the exact number of relativistic degrees of freedom. These constraints are discussed in more detail in appendix C.

Constraints from equivalence principle searches: As the radial mode is a scalar, it couples to photons through FF rather than $F\tilde{F}$. We can express the saxion-photon interaction as

$$\mathcal{L}_{\phi\gamma} = \frac{1}{4} g_{\phi\gamma\gamma} \phi FF. \quad (11.15)$$

We generally expect $g_{\phi\gamma\gamma}$ to be of the order of the axion-photon coupling $g_{\theta\gamma\gamma}$. This is important because scalar couplings such as (11.15) are subject to much stronger constraints than those from pseudoscalar interactions. Although such constraints are not relevant for theories in which $m_\phi \sim f_a$, constraints from $g_{\phi\gamma\gamma}$ are significant for theories with light radial modes such as those considered in this work. For a recent review of constraints on light scalar fields, see e.g. the Snowmass 2021 white paper on scalar and vector ultralight dark matter [227].

The reason why scalar-photon interactions give rise to such strong constraints is that the light scalar interactions give rise to long-range forces [228]. Fundamentally, these forces can be understood as scalar exchange. Classically, the effect appears because a background scalar field will perturb the effective EM coupling through terms such as (11.15). Such perturbations will slightly shift the EM contribution to the rest mass of macroscopic objects, which gives rise to long-range interactions and apparent equivalence-principle (EP) violation that can be tested in fifth-force searches such as Eöt-Wash [229] and through lunar ranging [230]. Searches for EP violation are sensitive to EM-coupled scalar fields even if such fields are not DM. If the scalar is dark matter, it is also constrained from atomic clocks through the oscillating EM perturbations.

The most relevant constraints are summarized by Hess et al. [231] from torsion balance data [229, 232, 233] as well as data from the MICROSCOPE space mission [234]. Assuming a KSVZ-like photon coupling, such data constrain the $[m_\phi, f_a]$ parameter space. However, as the f_a cut-offs for such constraints are beyond the range considered here, constraints from EP violation in our context appear simply as a lower bound on saxion mass:

$$m_\phi \gtrsim 10^{-5} \text{ eV}. \quad (11.16)$$

It is possible to reuse data from ALP searches to constrain saxion-photon interactions. This was studied by Flambaum et al. [235], who derived saxion-photon constraints from CAST and ADMX. However, although such a remapping is possible, scalar-photon constraints from EP violation dominate those from ALP searches by many orders of magnitude [235].

Axion-saxion mass hierarchy: PQ symmetry should be spontaneously broken. In addition to that, we have a small explicit PQ breaking for the axion mass. This small explicit PQ symmetry breaking naturally realizes a mass hierarchy $m_a \ll m_\phi$. One may ask if it is possible to have the opposite hierarchy, i.e., $m_a \gg m_\phi$. If the saxion is also light, we can address this question with an effective potential for the axion and the saxion. By integrating out all of the other particles, this potential is of the form

$$V_{\text{eff}}(P) = V_{\text{PQinv}}(\phi) + V_{\text{PQviol}}(\phi, \theta). \quad (11.17)$$

Here V_{PQinv} is invariant under PQ symmetry transformation, but V_{PQviol} is not. In order to realize $m_\phi \ll m_a$, we need to satisfy

$$\frac{\partial^2 V_{\text{PQinv}}}{\partial \phi^2} + \frac{\partial^2 V_{\text{PQviol}}}{\partial \phi^2} \ll \frac{1}{\phi^2} \frac{\partial^2 V_{\text{PQviol}}}{\partial \theta^2} \quad (11.18)$$

at the potential minimum. This requires some fine-tuning. First, it is natural to have $\partial^2 V_{\text{PQinv}}/\partial \phi^2 \gg \partial^2 V_{\text{PQviol}}/\partial \phi^2$, $(1/\phi^2)\partial^2 V_{\text{PQviol}}/\partial \theta^2$ because the PQ-violating terms are subdominant. Also, it is natural to have $\partial^2 V_{\text{PQviol}}/\partial \phi^2 \sim (1/\phi^2)\partial^2 V_{\text{PQviol}}/\partial \theta^2$. In this thesis, we do not pursue the possibility of having $m_\phi \ll m_a$ and focus on the region with $m_a < m_\phi$.

Impact of CP-violating term on the axion quality: In this thesis, we are studying both QCD axions and axions in the more general sense of ALPs. However, the subsets of axions that are intended to solve the strong QCD problem are sensitive to additional PQ-violating effects such as those introduced by the higher dimensional operator eq. (11.3). In particular, the additional CP breaking introduced must not induce an axion VEV of more than $\theta \sim 10^{-10}$, where $\theta = 0$ is the CP-conserving minimum of the QCD potential. To estimate this shift, we can take $\delta_\theta = \pi/2$ to maximally misalign the higher-dimensional potential from the QCD potential. In this case, in the late-time potential, where $\phi = f_a$, is

$$V_{\text{today}} \approx -2^{1-\frac{n}{2}} \frac{A f_a^n}{M^{n-3}} \theta + \frac{1}{2} m_a^2 f_a^2 \theta^2 + \mathcal{O}(\theta^3). \quad (11.19)$$

The CP-violating VEV induced by this potential is then

$$\theta_{\text{CPV}} \approx \frac{2^{1-\frac{n}{2}} \frac{A f_a^n}{M^{n-3}}}{m_a^2 f_a^2}. \quad (11.20)$$

For a QCD axion to successfully solve the strong CP problem, we require $\theta_{\text{CPV}} \ll 10^{-10}$. As this constraint applies only to the subset of axions that are QCD axions, we generally display this condition by shading the otherwise yellow QCD line of $[m_a, f_a]$ parameter space blue in the range where it cannot be satisfied.

Saxion-to-axion decay: Saxions may decay into axions. As discussed in appendix C, this occurs at a rate of

$$\Gamma_{Sa} \sim \frac{1}{64\pi} \frac{m_\phi^3}{\phi^2}. \quad (11.21)$$

Such decay may drain the energy in radial oscillations into a hot axion relic. Such a relic will continue to dominate the plasma if it is efficiently produced in excess of the plasma density because the axion relic never redshifts faster than radiation if we neglect changes in $g_{*(s)}$. Therefore, we impose the following constraint:

$$\Gamma_{Sa} < H \quad \text{while} \quad \rho_{\text{kin}}^\phi > \rho_r \quad (\text{hot axions dominate the plasma}), \quad (11.22)$$

where ρ_{kin}^ϕ is the energy in radial oscillations. Furthermore, even if the hot axion relic does not dominate the plasma, it may still lead to DM overproduction. To ensure that this does not happen, we also impose the following constraint:

$$\Gamma_{Sa} < H \quad \text{while} \quad \rho_{\text{kin}}^\phi \frac{2m_a}{m_\phi} \frac{g_*(T_{\text{eq}})}{g_*(T)} \left(\frac{T_{\text{eq}}}{T} \right)^3 > \rho_{r,\text{eq}} \quad (\text{DM overproduction}). \quad (11.23)$$

Here the subscript $_{\text{eq}}$ refers to quantities measured at radiation-matter equality.

These constraints depend on the damping mechanism because they are avoided if the radial oscillations are damped before saxion-to-axion decay becomes efficient. We, therefore, impose them only in the full model implementation where the damping mechanism is specified.

12 Models with nearly-quadratic potentials

For the radial dynamics to be effective at driving angular motion, the radial mass must be much lighter than the usual value of $m_\phi \sim f_a$. Therefore, the first model we here consider is a SUSY-motivated model in which PQ symmetry is spontaneously broken by radiative corrections. Such a potential was originally proposed by Moxhay and Yamamoto [236], who found that soft SUSY breaking generates a nearly-quadratic potential of the form

$$V_{\text{late}} = m_\phi^2 |P|^2 \left(\frac{1}{2} \ln \left(\frac{2|P|^2}{f_a^2} \right) - \frac{1}{2} \right) + \frac{1}{4} m_\phi^2 f_a^2. \quad (12.1)$$

The radial mass m_ϕ may then be naturally light as it is protected by SUSY. In this framework, PQ is spontaneously broken by the logarithmic running of the potential, which gives the radial mode a VEV at $\phi = f_a$ once the radial mode is fully relaxed. Such models were first studied in the context of kinetic misalignment in [27]. In the following sections, we will investigate how radial dynamics in such a model can provide the initial conditions for fragmentation and kinetic misalignment.

Note that the mass of the radial mode, which around the minimum is m_ϕ , receives a logarithmic correction at high field values. However, this logarithmic correction is at most $\mathcal{O}(1)$. We therefore neglect the correction at large field amplitudes where we then take $V \approx m_\phi^2 |P|^2 = \frac{1}{2} m_\phi^2 \phi^2$. Thus, the fundamental parameter m_ϕ here agrees with $m_\phi \equiv \sqrt{\phi^{-1} V'}$ as defined in eq. (11.6), such that there is no distinction between the effective mass at large field values and the fundamental parameter, as long as we neglect the logarithmic running of the potential.

12.1 Initial conditions

To set up initial conditions for axion kinetic misalignment, we here study how such a rotation can be started and which yield is produced from the resulting kick. A classic example of models that involve a rotating complex scalar field is Affleck-Dine Baryogenesis [31]. Such models can be

realized by utilizing higher-dimensional terms to drive up the VEV and introduce PQ-breaking at early times, thereby spinning up the field. This setup was proposed in the context of Affleck-Dine Baryogenesis in [32] and was used to motivate kinetic misalignment in [27]. The mechanism utilizes operators of the form

$$V_{\text{early}} = (m_\phi^2 - c_H H^2)|P|^2 + \frac{A + c_A H}{n} \frac{P^n}{M^{n-3}} + h.c. + \frac{|P|^{2n-2}}{M^{2n-6}}, \quad (12.2)$$

where we neglected the logarithmic correction to the late-time mass m_ϕ , H is the Hubble parameter, M is a suppression scale which we typically take to be $m_{\text{Pl}} \approx 2.4 \times 10^{18}$ GeV, A is a dimension-full constant which we typically take to be $\mathcal{O}(1) \times m_\phi$ and c_A is an $\mathcal{O}(1)$ constant. The negative Hubble-induced mass can be naturally generated in models where P couples to fields that dominate the energy density of the universe [32]. If at early times $m_\phi \ll H$ then this negative mass contribution drives up the VEV. To stabilize against this negative potential contribution, one can introduce a higher-dimensional term in the superpotential [32] of the form

$$W = \frac{1}{n} \frac{P^n}{M^{n-3}}. \quad (12.3)$$

This superpotential sources a scalar potential corresponding to the last term in eq. (12.2). The A-term, i.e. the middle term of eq. (12.2), is generated from W by soft SUSY breaking. In addition to the full discussion in [32], a summary of the generation of these terms can be found in [37].

At early times where $H \gg m_\phi$ the H^2 term then drives ϕ to a large VEV stabilized by the $|P|^{2n-2}$ term. The VEV found at the minimum between these two terms is

$$\phi_{\text{early}} = \left(\frac{2^{n-2}}{n-1} \right)^{\frac{1}{2n-4}} (HM^{n-3})^{\frac{1}{n-2}}. \quad (12.4)$$

Around this VEV, both the radial and angular modes acquire large masses,

$$m_{\phi,\text{early}} = \sqrt{2(n-1)}H \quad \text{and} \quad m_{a,\text{early}} = \frac{\sqrt{2|c_A|n}}{\sqrt[4]{n-1}}H, \quad (12.5)$$

such that both fields are heavy during inflation and can track the minimum until $m_\phi \propto 3H$, where the late-time mass becomes important. We assume that the higher-dimensional operators become irrelevant once $m_\phi > 3H$ so that we can take the potential to be simply eq. (12.1) after this transition.

During the transition, the PQ violating $P^n + P^{*n}$ terms can drive rotation. This is possible because the angular mode transitions from a minimum set by the $c_A H$ term into a minimum set by the A term. The angular minima of these two contributions are determined by the complex phase of c_A and A , respectively. If the minima of these two contributions are not accidentally aligned, the field experiences a torque as the minimum shifts [224] and eq. (11.4) applies with $\theta_{\text{kick}} \neq 0$. With $A = \mathcal{O}(1) \times m_\phi$ this naturally yields $\epsilon = \mathcal{O}(1)$ such that

$$\dot{\theta}_{\text{kick}} = \mathcal{O}(1) \times m_\phi, \quad (12.6)$$

where m_ϕ is the potential parameter appearing in eq. (12.2) and not the much larger early-time mass given by (12.5). This kick is delivered when $m_\phi \approx 3H$, which implies a kick temperature of

$$T_{\text{kick}} \approx \sqrt[4]{\frac{10}{g_*(T_{\text{kick}})\pi^2} \sqrt{m_\phi m_{\text{Pl}}}}. \quad (12.7)$$

We assume that soon after the kick the PQ-violating potential becomes irrelevant so that we can disregard the higher-dimensional terms after the first rotation. This is supported by numerical evaluation of the equations of motion [32].

In the nearly-quadratic model, the radial mass m_ϕ is approximately constant in time. Therefore, applying eq. (11.9) we see that ϕ reaches the minimum at f_a at a_{kin} given by

$$\frac{\phi_{\text{kick}}}{f_a} \approx \left(\frac{a_{\text{kin}}}{a_{\text{kick}}} \right)^{3/2}. \quad (12.8)$$

If no significant entropy injection takes place, then this corresponds to a temperature of

$$T_{\text{kin}} \approx T_{\text{kick}} \left(\frac{g_{*s}(T_{\text{kick}})}{g_{*s}(T_{\text{kin}})} \right)^{1/3} \left(\frac{f_a}{\phi_{\text{kick}}} \right)^{2/3}. \quad (12.9)$$

After this time, $\dot{\theta}$ transitions to kination-like scaling as described in section 11.3.

Estimation of produced yield: After the kick is complete and the PQ-violating terms have become subdominant, the comoving PQ-charge becomes conserved, so that $n_{\text{PQ}} \propto a^{-3}$. The resulting yield, $Y_{\text{kick}} = n_{\text{PQ}}/s$, is therefore constant in the absence of entropy injection. This yield is

$$Y_{\text{kick}} = \frac{n_{\text{PQ}}}{s} = \epsilon \frac{m_\phi \phi_{\text{kick}}^2}{\frac{2\pi^2}{45} g_{*s} T_{\text{kick}}^3} \approx 0.8 \times \epsilon \left(\frac{M}{m_{\text{Pl}}} \right)^{\frac{2n-6}{n-2}} \left(\frac{m_{\text{Pl}}}{m_\phi} \right)^{\frac{n-6}{2n-4}}, \quad (12.10)$$

where the prefactor has a mild dependence on n , $g_*(T_{\text{kick}})$, and on $g_{*s}(T_{\text{kick}})$.

Dark matter solution in the absence of yield dilution: If no entropy injection takes place and either kinetic misalignment or fragmentation takes place, then the yield eq. (12.10) eventually sources axion dark matter. $2Y_{\text{kick}} \approx Y_{\text{DM}}$ is achieved for

$$m_{\phi, \text{DM}} \approx \xi_n \times g_*(T_{\text{kick}})^{\frac{3(n-2)}{2(n-6)}} \left(\frac{g_{*s}(T_{\text{eq}})}{g_{*s}(T_{\text{kick}})g_*(T_{\text{eq}})} \right)^{\frac{2(n-2)}{n-6}} \times m_{\text{Pl}} \left(\frac{M}{m_{\text{Pl}}} \right)^{\frac{4(n-3)}{n-6}} \left(\frac{m_a \epsilon}{T_{\text{eq}}} \right)^{\frac{2(n-2)}{n-6}}$$

where $\xi_n = (n-1)^{-\frac{2}{n-6}} \pi^{\frac{2-n}{n-6}} 2^{\frac{9(n-2)}{2(n-6)}} 3^{\frac{2(n-4)}{n-6}} 5^{\frac{n-2}{2(n-6)}}$.

(12.11)

For a few representative choices of n ,

$$m_{\phi, \text{DM, not diluted}} \approx \begin{cases} 70 \left(\frac{M}{m_{\text{Pl}}} \right)^{16} \left(\frac{m_a}{10^{-2} \text{ eV}} \right)^{10} \left(\frac{\epsilon}{0.9} \right)^{10} & \text{if } n = 7, \\ 2 \times 10^{12} \left(\frac{M}{m_{\text{Pl}}} \right)^7 \left(\frac{m_a}{10^{-2} \text{ eV}} \right)^4 \left(\frac{\epsilon}{0.9} \right)^4 & \text{if } n = 10, \\ 5 \times 10^{13} \left(\frac{M}{m_{\text{Pl}}} \right)^{\frac{40}{7}} \left(\frac{m_a}{10^{-2} \text{ eV}} \right)^{\frac{22}{7}} \left(\frac{\epsilon}{0.9} \right)^{\frac{22}{7}} & \text{if } n = 13. \end{cases} \quad (12.12)$$

It must be noted that the yield has a somewhat shallow dependence on the radial mass, i.e. $Y_{\text{kick}} \propto m_\phi^l$ where $l \approx 0.1, 0.25$ or 0.3 for $n = 7, 10$ or 13 respectively. Therefore small changes in various parameters require a large change in m_ϕ to compensate, especially for lower values of n . Correspondingly, the numerical prefactor of eq. (12.12) is sensitive to small changes in other parameters.

12.2 Application of constraints

The scenario described above is subject to several constraints of diverse nature. These constraints were described in section 11.4. The most significant of these are the conditions that the kick takes place before T_* , that the field remains homogeneous until T_* , and the condition that a thermal ϕ relic does not spoil CMB or BBN.

The perturbativity constraint and the condition that the kick happens early enough are model dependent, so we in the following paragraphs discuss how they apply to the nearly-quadratic model.

Perturbativity The nearly-quadratic potential implies self-interactions which we require not to lead to non-perturbative behaviour. To extract the coupling parameters from the nearly-quadratic potential, we define the radial excitation around f_a as $S = \phi - f_a$ and expand the potential around $S = 0$ in powers of S/f_a . This yields the series

$$V_{\text{late}} = m_\phi^2 |P|^2 \left(\frac{1}{2} \ln \left(\frac{2|P|^2}{f_a^2} \right) - \frac{1}{2} \right) \approx \frac{1}{2} m_\phi^2 S^2 + \frac{\lambda_{\text{ln}}}{3!} f_a S^3 - \frac{\lambda_{\text{ln}}}{4!} S^4 + \mathcal{O} \left(\lambda_{\text{ln}} \frac{S^5}{f_a} \right), \quad (12.13)$$

where the dimensionless coupling parameter λ_{ln} , which is common to the leading terms, is

$$\lambda_{\text{ln}} = \frac{m_\phi^2}{f_a^2}. \quad (12.14)$$

Therefore, perturbativity of the self-interactions implies a bound of

$$\frac{m_\phi^2}{f_a^2} < 4\pi. \quad (12.15)$$

Early enough kick: It is essential for the consistency of the nearly-quadratic model as a kick mechanism that the kick takes place before the field would have been trapped, i.e., $T_{\text{kick}} > T_*$. Comparing eq. (5.28) with eq. (12.7) we find that this constraint sets a significantly lower bound on the viable axion masses. In particular, demanding $T_{\text{kick}} > T_*$ implies a lower bound of

$$m_a \gtrsim \begin{cases} 9.0 \times 10^{-5} \text{ eV} \left(\frac{f_a}{10^{10} \text{ GeV}} \right)^{0.13} \left(\frac{M}{m_{\text{Pl}}} \right)^{1.80} & \text{for } n = 7 \\ 1.3 \times 10^{-12} \text{ eV} \left(\frac{f_a}{10^{10} \text{ GeV}} \right)^{0.40} \left(\frac{M}{m_{\text{Pl}}} \right)^{2.45} & \text{for } n = 10 \\ 1.6 \times 10^{-17} \text{ eV} \left(\frac{f_a}{10^{10} \text{ GeV}} \right)^{0.57} \left(\frac{M}{m_{\text{Pl}}} \right)^{2.86} & \text{for } n = 13 \end{cases} \quad (12.16)$$

12.3 Early damping

The simplest case is when the radial mode is assumed to damp through some unspecified mechanism early enough that no yield dilution takes place. In this scenario, the yield generated by the kick, as given by eq. (12.10), remains constant until late times, and the above results apply without modifications. Because the kick is controlled by the higher-dimensional terms, the viability of parameter space is sensitive to the choice of n . The constraints from eq. (12.15) and eq. (11.14) can be directly applied to the m_ϕ solution eq. (12.12). The constraint from homogeneity, eq. (11.12), implies an upper bound on m_a :

$$m_a < \mathcal{O}(10 - 100) \times (A_s(k_{\text{kin}}))^{-\frac{177(n-6)}{8(19n+4)}} f_a^{-\frac{17(n-6)}{76n+16}} \left(\frac{M}{m_{\text{Pl}}} \right)^{-\frac{59(n-3)}{19n+4}} m_{\text{Pl}}^{-\frac{59(n-6)}{76n+16}} \left(\frac{T_{\text{eq}}}{\epsilon} \right)^{\frac{59(3n-10)}{76n+16}}, \quad (12.17)$$

where the exact value of the prefactor depends on n and the number of relativistic degrees of freedom. Equivalently, for representative choices of n , the condition is

$$m_a \lesssim \begin{cases} 2 \times 10^{-2} \text{ eV} \left(\frac{A_s(k_{\text{kin}})}{2.1 \times 10^{-9}} \right)^{-0.16} \left(\frac{f_a}{10^9 \text{ GeV}} \right)^{-0.031} \left(\frac{M}{m_{\text{Pl}}} \right)^{-1.7} \epsilon^{-1.2} & \text{if } n = 7 \\ 5 \times 10^{-6} \text{ eV} \left(\frac{A_s(k_{\text{kin}})}{2.1 \times 10^{-9}} \right)^{-0.46} \left(\frac{f_a}{10^9 \text{ GeV}} \right)^{-0.088} \left(\frac{M}{m_{\text{Pl}}} \right)^{-2.1} \epsilon^{-1.5} & \text{if } n = 10 \\ 5 \times 10^{-8} \text{ eV} \left(\frac{A_s(k_{\text{kin}})}{2.1 \times 10^{-9}} \right)^{-0.62} \left(\frac{f_a}{10^9 \text{ GeV}} \right)^{-0.12} \left(\frac{M}{m_{\text{Pl}}} \right)^{-2.3} \epsilon^{-1.7} & \text{if } n = 13 \end{cases} \quad (12.18)$$

Furthermore, for the kick mechanism to provide initial conditions for kinetic misalignment, the kick must take place before T_* . Together with the constraint from the thermal relic, the condition $T_* < T_{\text{kick}}$ most strongly restricts the scenario. The resulting parameter space is displayed for $n = 7, 10$ and 13 in figures 16 and 17. The shallow dependence of Y_{kick} to m_ϕ leaves only a narrow range of parameter space viable for $n = 7$. For larger n , Y_{kick} becomes more sensitive to m_ϕ such that a larger range of parameter space can be supported, although this space will also be shifted to lower m_a .

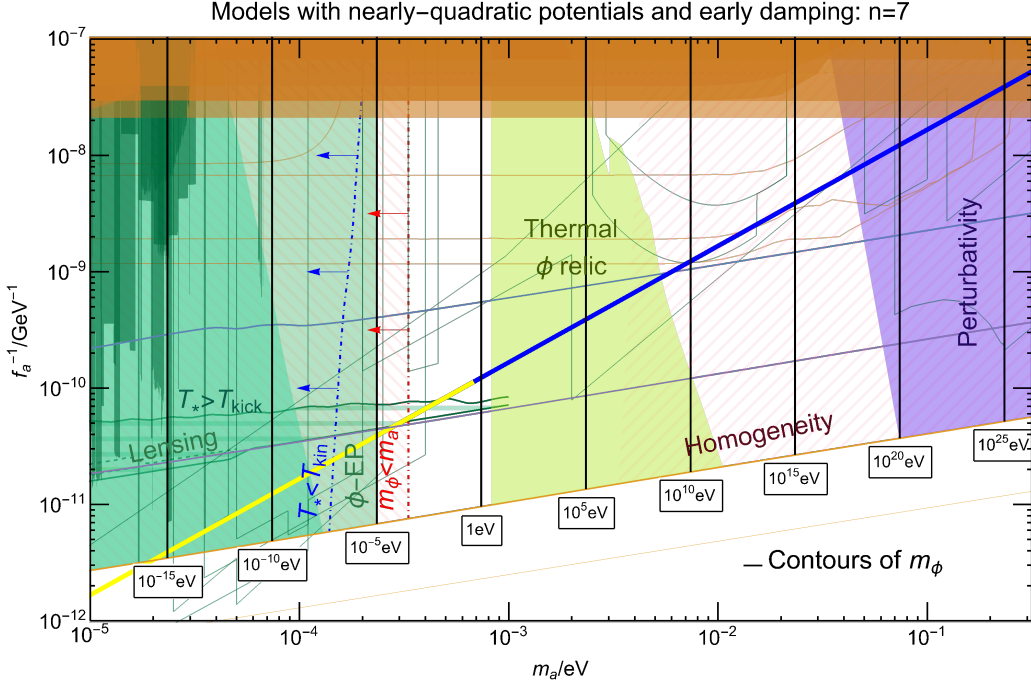


Figure 16. Constraints on the $[m_a, f_a]$ parameter space for $n = 7$ and $M = m_{\text{Pl}}$ in the models with nearly-quadratic potentials, when early damping without entropy dilution is assumed. Here n and M are parameters of the potential eq. (12.2). The constraints from thermal ϕ relics, perturbativity, homogeneity, and EP violation are given by eqs. (11.14), (12.15), (12.18), and (11.16), respectively. The temperatures T_{kick} , T_{kin} , and T_* are given by eqs. (12.7), (12.9) and (5.28), respectively. The vertical contours indicate values of the radial mode mass m_ϕ , which are fixed by eq. (12.12) to obtain the observed DM relic. Note that a few additional experiments have been added relative to the discussion in section 3, see [83] for details.

12.4 General damping

We now generalize to consider damping later than the time when the PQ field becomes dominant. Because the energy density of the radial mode redshifts as a^{-3} , it will eventually dominate the energy density of radiation if not damped sufficiently early. Therefore, we first need to estimate the yield dilution which results from damping after the radial mode dominates the energy density.

Estimation of yield dilution for late damping: The radial mode will begin to dominate radiation at the temperature T_{dom} for which $\rho_\phi \approx \rho_r$. Solving $\rho_\phi(T_{\text{dom}}) \approx \rho_r(T_{\text{dom}})$ for T_{dom} we find

$$T_{\text{dom}} \approx 0.8 \left(\frac{M}{m_{\text{Pl}}} \right)^{\frac{n-3}{2(n-2)}} m_{\text{Pl}}^{\frac{n-6}{8(n-2)}} m_\phi^{\frac{n+2}{8(n-2)}} T_{\text{damp}}^{3/4}, \quad (12.19)$$

where the prefactor has a mild dependence g_* and the exact value of n .

If damping takes place after this time, i.e. if $T_{\text{dom}} > T_{\text{damp}}$, then the associated entropy injection is significant. Therefore, the entropy density deviates from $s \propto a^{-3}$ and the yield fails to be constant in time. The resulting yield dilution must be taken into account before an accurate estimate of the present-day relic can be made. If all energy in the radial oscillations is instantaneously transferred to the SM plasma at some temperature T_{damp} and the entropy injection increases the plasma

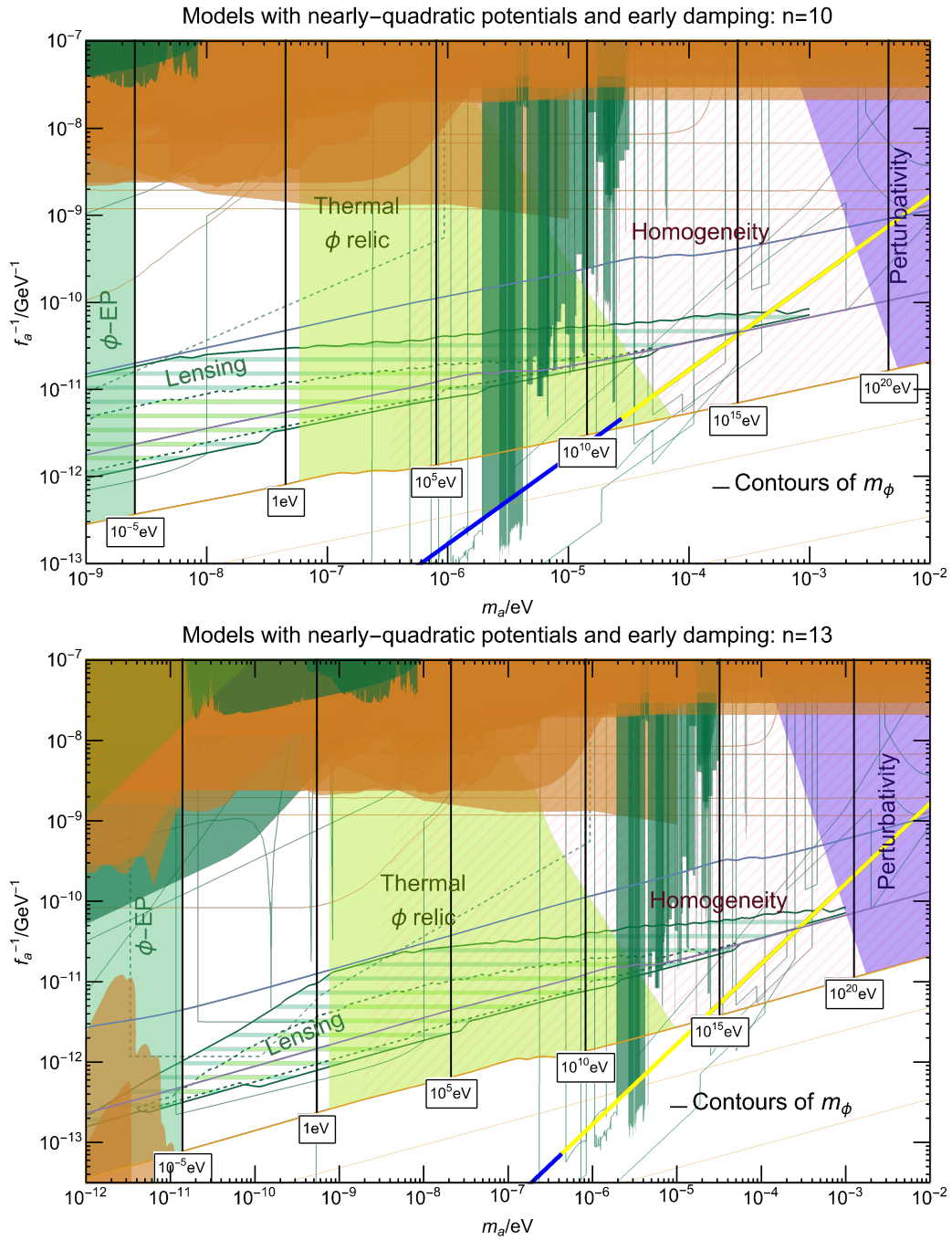


Figure 17. Same as figure 16 but for $n = 10$ and $n = 13$.

temperature to $T_{\text{damp,heated}}$, then the yield is diluted to

$$Y_{\text{diluted}} = Y_{\text{kick}} \frac{g_{*s}(T_{\text{damp}})T_{\text{damp}}^3}{g_{*s}(T_{\text{damp,heated}})T_{\text{damp,heated}}^3}. \quad (12.20)$$

where the reheated temperature $T_{\text{damp,heated}}$ is determined by

$$\frac{1}{2}m_\phi^2\phi_{\text{kick}}^2 \frac{g_{*s}(T_{\text{damp}})T_{\text{damp}}^3}{g_{*s}(T_{\text{kick}})T_{\text{kick}}^3} = \frac{\pi^2}{30}g_*(T_{\text{damp,heated}})T_{\text{damp,heated}}^4. \quad (12.21)$$

This dilution of the yield implies that the observed DM yield is realized for radial masses which differ quite significantly from the scenario without significant entropy injection.

Solution for yield in the presence of yield dilution: If $T_{\text{dom}} > T_{\text{damp}}$, then the initial yield, given by eq. (12.10) is diluted to the result given by eq. (12.20). Correspondingly, in such diluted scenarios, the observed DM relic is realized by

$$m_{\phi,\text{DM,diluted}} \approx 5 \left(\epsilon \frac{m_a}{T_{\text{eq}}} \right)^{\frac{8(n-2)}{7n-18}} \left(m_{\text{Pl}}^{3(n-2)} M^{4(n-3)} T_{\text{damp}}^{6(n-2)} \right)^{\frac{1}{7n-18}}, \quad (12.22)$$

where the exact value of the $\mathcal{O}(1)$ factor depends on n and the number of relativistic degrees of freedom. For specific choices of n , this corresponds to

$$m_{\phi,\text{DM,diluted}} \approx \begin{cases} 2 \times 10^3 \text{ GeV} \left(\frac{m_a}{\text{eV}} \right)^{40/31} \left(\frac{M}{m_{\text{Pl}}} \right)^{16/31} \left(\frac{T_{\text{damp}}}{10^2 \text{ GeV}} \right)^{30/31} \epsilon^{40/31} & \text{if } n = 7, \\ 1 \times 10^4 \text{ GeV} \left(\frac{m_a}{\text{eV}} \right)^{16/13} \left(\frac{M}{m_{\text{Pl}}} \right)^{7/13} \left(\frac{T_{\text{damp}}}{10^2 \text{ GeV}} \right)^{12/13} \epsilon^{16/13} & \text{if } n = 10, \\ 2 \times 10^4 \text{ GeV} \left(\frac{m_a}{\text{eV}} \right)^{88/73} \left(\frac{M}{m_{\text{Pl}}} \right)^{40/73} \left(\frac{T_{\text{damp}}}{10^2 \text{ GeV}} \right)^{66/73} \epsilon^{88/73} & \text{if } n = 13. \end{cases} \quad (12.23)$$

Note in particular that the m_a dependence in m_ϕ is much shallower in the presence of entropy injection than in the undiluted case, see eq. (12.12). This allows the solutions to span over a larger range of axion masses without requiring unreasonable radial masses.

Classification of scenarios It is useful to classify the possible scenarios based on whether significant entropy injection takes place or not, i.e. whether damping takes place before or after radial mode (saxion) domination at T_{dom} . Furthermore, there are interesting phenomenological consequences [36, 37] of scenarios that involve a period of kination¹², wherefore, we also highlight scenarios in which such a period takes place. We classify the scenarios as follows:

Scenario classification:		Kination	No kination
	Yield diluted	Scenario A	Scenario B
Yield not diluted	Scenario C	Scenario D	

Kination takes place when the energy density of the universe is dominated by the kinetic energy of the axion. Therefore, scenarios with significant yield dilution, i.e. with $T_{\text{dom}} > T_{\text{damp}}$, feature kination only if the radial oscillations are damped before ϕ relaxes to f_a , i.e. if $T_{\text{damp}} > T_{\text{kin}}$. If this condition is violated, then radial oscillations would dominate in place of the axion kinetic energy, and no kination era would arise. Scenarios without significant yield dilution, i.e. with $T_{\text{dom}} < T_{\text{damp}}$, feature kination if the kinetic energy of the axion dominates when the radial mode relaxes to the minimum at f_a , i.e. if $\rho_a(T_{\text{kin}}) > \rho_r(T_{\text{kin}})$.

Visualizations of the scenarios classified here are given in figure 18. A map of which scenarios are realized across the $[m_a, f_a]$ parameter space for various choices of T_{damp} is provided in figure 19.

¹²Kination is a period of cosmological history characterized by a a^{-6} evolution of the energy density.

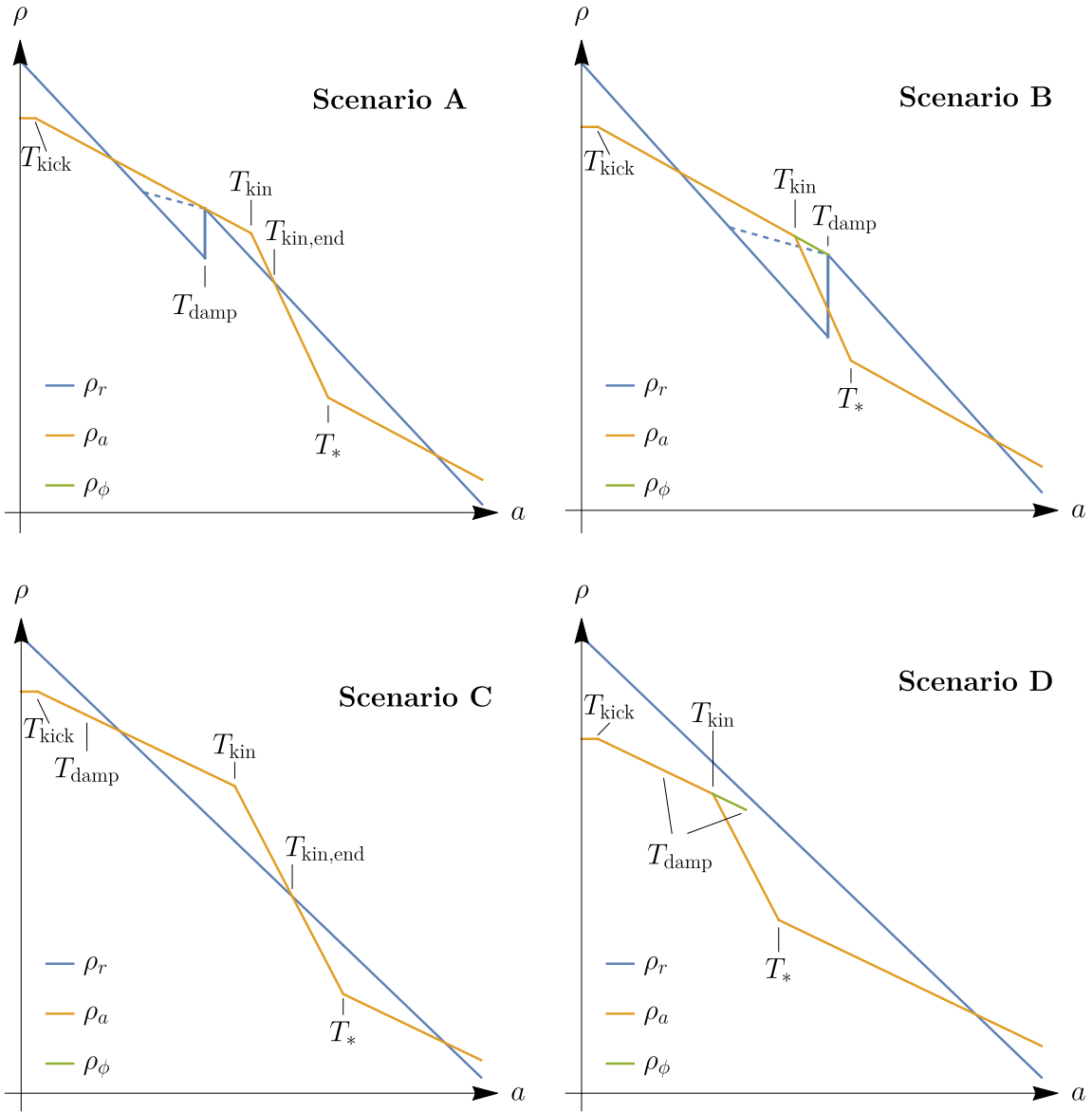


Figure 18. Overview of possible scenarios in models with nearly-quadratic potentials. The energy densities of the angular (axion, orange) and radial (saxion, green) modes are comparable until damping. When the densities coincide, only the angular mode is shown. Scenarios A and B are dominated by the PQ field at T_{damp} while C and D are radiation-dominated at T_{damp} . Scenarios A and C feature kination, while scenarios B and D do not. Regardless of whether a period of kination is triggered or not, the temperature at which the radial mode is relaxed to f_a is labeled T_{kin} . The end of kination is labeled by $T_{\text{kin,end}}$ if it does take place. Physically damping may not be nearly-instantaneous, and the energy density of the plasma could reheat smoothly, as indicated by the blue dashed lines. However, as only the initial and final states matter for yield dilution, this question does not impact the damping-mechanism-agnostic solution presented in this section. The physical evolution of T does impact the more realistic damping rates considered in later sections of this work. The evolutions displayed here are sketches intended to highlight the characteristic differences between scenarios and do not represent actual solutions.

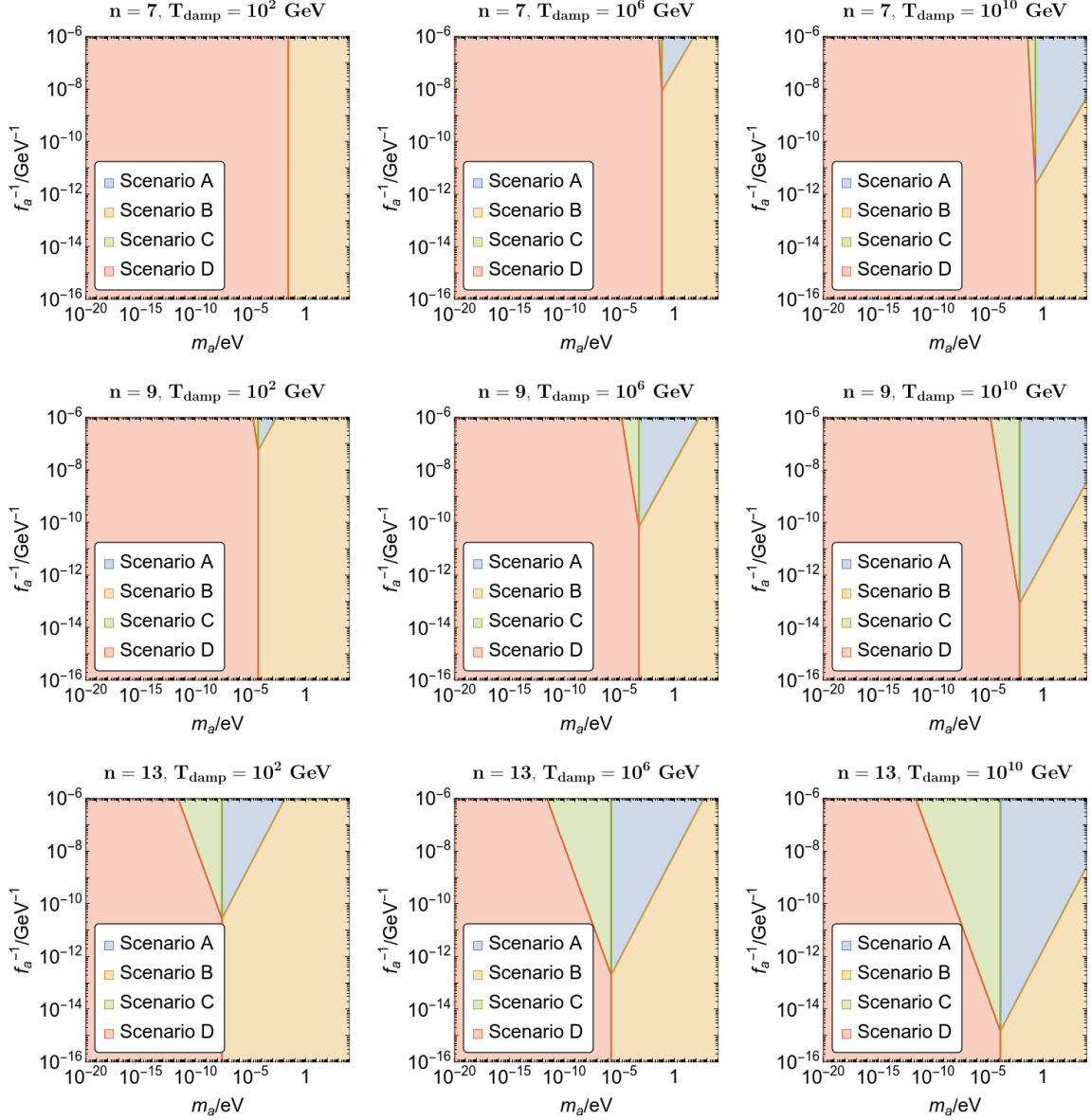


Figure 19. Overview of scenarios in the models with nearly-quadratic potentials. Plots here assume $M = m_{\text{Pl}}$; for lower values of M scenario D (always radiation dominated) is favored. The scenarios referred to here are illustrated in fig. 18. Note that the plots presented here differ from figs. 20 and 21 in that a single, constant value for T_{damp} is assumed for each plot.

Viable range of T_{damp} We here discuss which ranges of parameter space can be supported. While the radial oscillations must be damped before T_* for the analysis of [1] to apply such damping would interfere with the kick if efficient at T_{kick} . Therefore, T_{damp} is ultimately limited by¹³

$$T_{\text{kick}} > T_{\text{damp}} > T_*. \quad (12.24)$$

This condition gives the wedge-shaped envelope of the $[T_{\text{damp}}, m_a]$ parameter space, which is displayed for constant f_a in fig. 20.

¹³It is possible to consider scenarios in which damping takes place after fragmentation. However, the fragmentation analysis would have to be redone in the presence of radial oscillations. We leave such studies to future work.

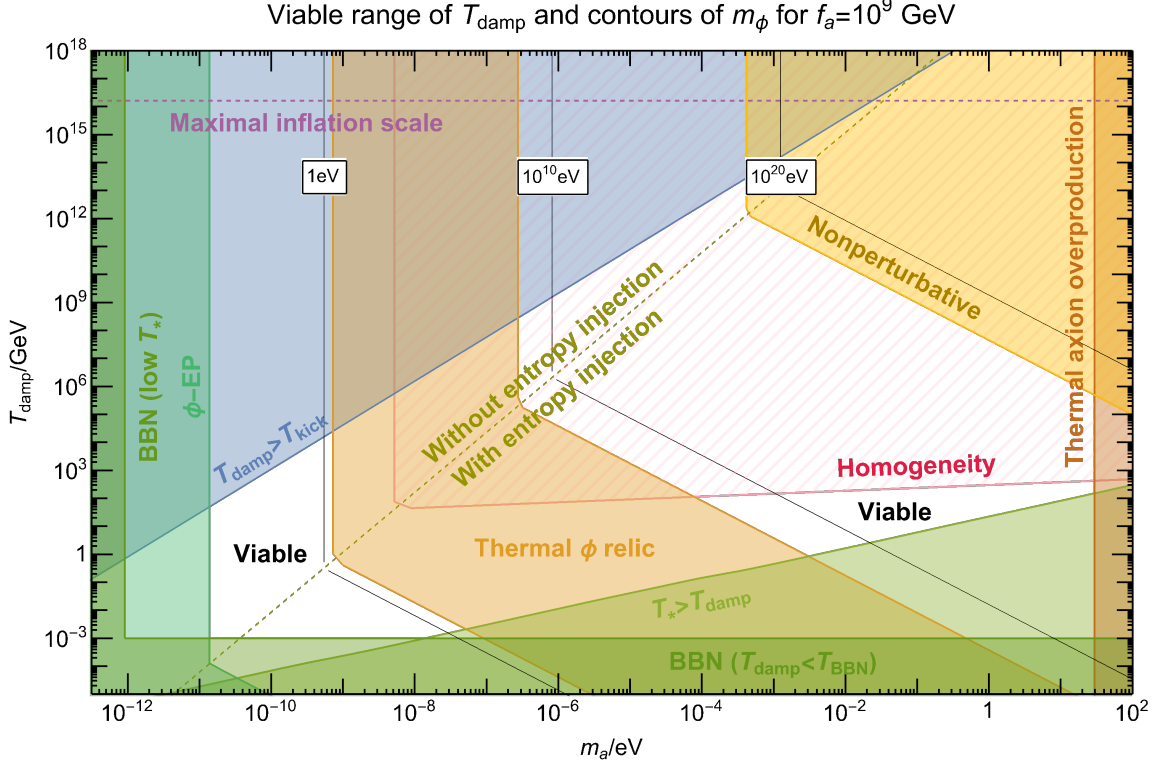


Figure 20. Constraints on the $[T_{\text{damp}}, m_a]$ parameter space for fixed $f_a = 10^9$ GeV, $n = 13$ and $M = m_{\text{Pl}}$. In addition to the constraints presented in figure 16, the damping temperature T_{damp} is here constrained by the BBN constraint (11.11) and the range of validity of the analysis (12.24). The plot corresponds to a horizontal slice of fig. 21.

The damping temperature is further restricted by constraints on the thermal relic of saxions. These constraints arise when m_ϕ , as determined by either eq. (12.12) or eq. (12.23), falls in the range given by eq. (11.14) where ALP thermal relics are constrained by CMB or BBN [206, 237]. The constraint corresponds to the orange band labeled "Thermal ϕ relic" in figure 20 and to the central dark grey constraint in fig. 21.

Yield dilution also modifies the homogeneity condition. Combining the solution eq. (12.23) with the condition eq. (11.12) yields an upper bound on T_{damp} :

$$T_{\text{damp}} \lesssim \begin{cases} 2 \times 10^1 \text{ GeV} \left(\frac{m_a}{\text{eV}}\right)^{0.27} \left(\frac{A_s(k_{\text{kin}})}{2.1 \times 10^{-9}}\right)^{-1.4} \left(\frac{f_a}{10^9 \text{ GeV}}\right)^{-0.27} \left(\frac{M}{m_{\text{Pl}}}\right)^{0.97} \epsilon^{4/3} & \text{if } n = 7 \\ 4 \times 10^2 \text{ GeV} \left(\frac{m_a}{\text{eV}}\right)^{0.15} \left(\frac{A_s(k_{\text{kin}})}{2.1 \times 10^{-9}}\right)^{-1.3} \left(\frac{f_a}{10^9 \text{ GeV}}\right)^{-0.25} \left(\frac{M}{m_{\text{Pl}}}\right)^{0.93} \epsilon^{4/3} & \text{if } n = 10 \\ 1 \times 10^3 \text{ GeV} \left(\frac{m_a}{\text{eV}}\right)^{0.10} \left(\frac{A_s(k_{\text{kin}})}{2.1 \times 10^{-9}}\right)^{-1.3} \left(\frac{f_a}{10^9 \text{ GeV}}\right)^{-0.24} \left(\frac{M}{m_{\text{Pl}}}\right)^{0.92} \epsilon^{4/3} & \text{if } n = 13 \end{cases} \quad (12.25)$$

Interestingly, this condition permits entropy injection to circumvent the homogeneity condition such that a large- m_a window opens up in parameter space. This window is visualized in both fig. 20, which displays the structure of the constraints along a line of constant f_a , and in fig. 21, which displays the viable range of T_{damp} across the $[m_a, f_a]$ parameter space.

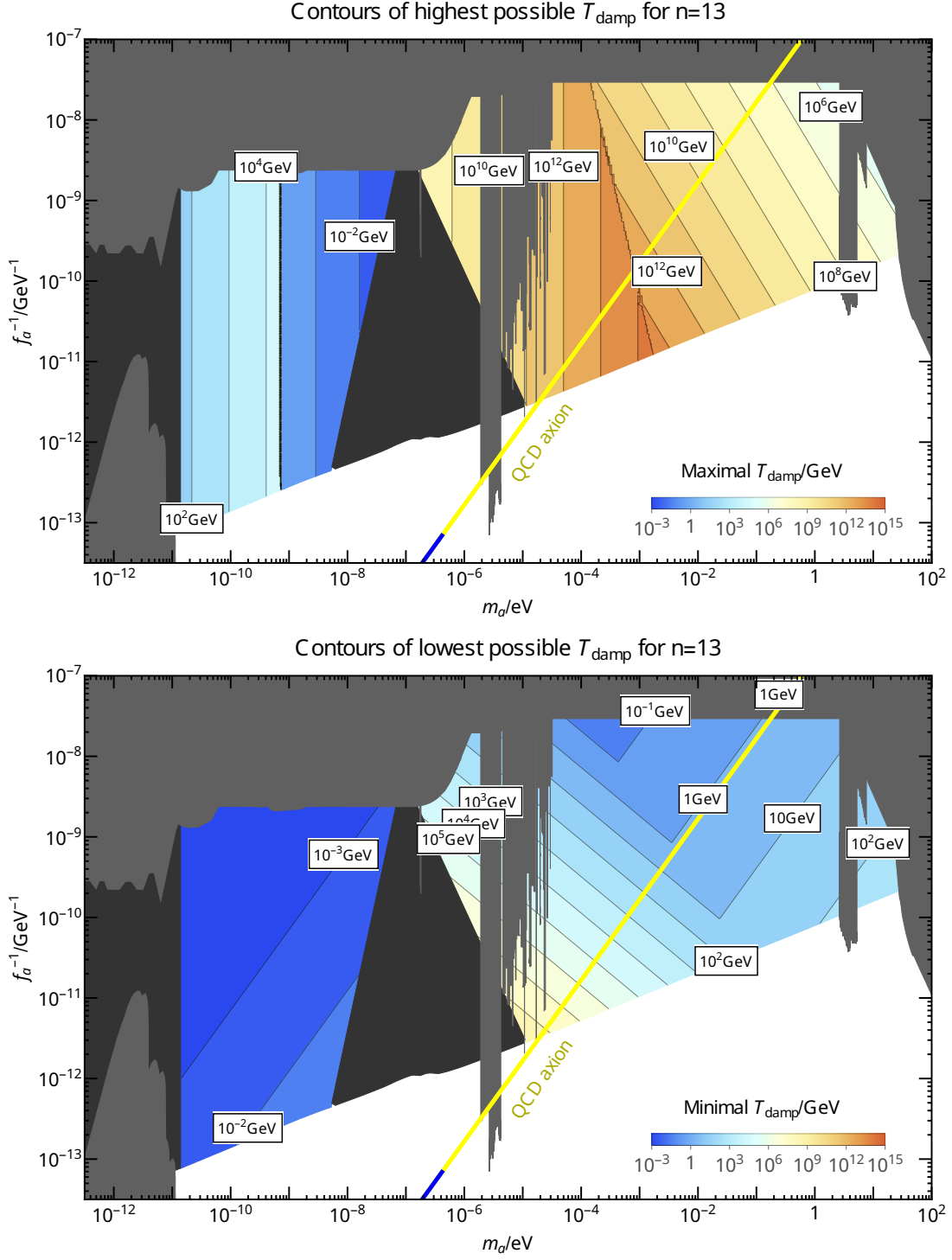


Figure 21. Survey of viable damping temperatures. We here assume that the homogeneity condition can be relaxed such that the dotted region in fig. 4 can be made viable. The temperature range can be extended further in scenarios in which the thermal relic constraint is resolved. **Top:** Contours of the highest possible damping temperature for which the scenario can be realized. The maximum thermalization temperature is reduced around $m_a \sim 10^{-8}$ eV because of N_{eff} constraints imposed by the thermal saxion relic. **Bottom:** Similar plot of the lowest viable damping temperature. Figure 20 serves as a guide to the structure of this plot.

13 Thermal damping with Yukawa interactions

In the previous section, we studied the nearly-quadratic model and discussed how it is necessary to damp radial oscillations to obtain the desired cosmology. Furthermore, we studied the general case in which we stayed agnostic to the damping interaction and specified the process entirely in terms of a damping temperature T_{damp} . We then mapped out the potential range damping temperatures, which we presented in figure 21. Now, we seek to implement specific a realization of the damping mechanism by considering a KSVZ-like fermion Yukawa interaction.

To achieve damping, we here consider a Yukawa interaction of the form

$$\mathcal{L}_{\text{int}} = \sqrt{2}yP\bar{\chi}\chi + h.c., \quad (13.1)$$

where χ is a fermion with interactions that couple it efficiently to the SM thermal bath. The fermion χ can be connected to the SM through either a Yukawa coupling or a gauge coupling [238]. For definiteness, we will here assume that the χ -SM coupling takes the form of a gauge interaction with coupling constant g_χ and gauge bosons γ . Schematically, the setup is then

$$\phi \begin{array}{c} \xleftarrow{y} \\ \text{Yukawa} \end{array} \chi \begin{array}{c} \xleftarrow{g_\chi} \\ \text{Gauge} \end{array} \text{SM plasma with gauge bosons } \gamma. \quad (13.2)$$

Such a Yukawa is interesting because it is an integral part of KSVZ-type axion models. In such a framework, the χ -fermions would be KSVZ fermions, the gauge interaction would be $\text{SU}(3)_{\text{QCD}}$, and damping as described here would be a natural consequence of the KSVZ construction.

The fermion receives a potential contribution through both the ϕ interaction and the SM gauge interaction. We neglect any other sources such that the effective χ mass is

$$m_{\chi,\text{eff}}^2 \approx m_{\chi,\phi}^2 + m_{\chi,T}^2 \approx y^2\phi^2 + g_\chi^2 T^2. \quad (13.3)$$

Additionally, we define the fine-structure constant of the gauge coupling to be $\alpha_\chi \equiv g_\chi^2/4\pi$. We assume $g_\chi < 1$ and that the Yukawa coupling is weak such that $y < \alpha_\chi < g_\chi < 1$.

13.1 Thermal effects and damping rates

Although a scalar field coupled through a Yukawa may at first glance appear to be a simple system, thermal effects lead to surprisingly rich dynamics. For a detailed study of such thermal effects, we refer to Mukaida et al. [238–240]. In particular, a Yukawa coupling of the type considered here generates thermal contributions to the radial potential such that the full effective mass $m_{\phi,\text{eff}}$ is of the form [239]

$$m_{\phi,\text{eff}}^2 \approx \begin{cases} m_{\phi,0}^2 + m_{\phi,\text{th}}^2 & \text{if } T > y\phi, \\ m_{\phi,0}^2 + m_{\phi,\text{ln}}^2 & \text{if } T < y\phi, \end{cases} \quad (13.4)$$

where $m_{\phi,\text{th}}^2 \approx y^2 T^2$ and $m_{\phi,\text{ln}}^2 \approx \alpha^2 \frac{T^4}{\phi^2}$.

Here $m_{\phi,0}$ is the zero-temperature mass considered in the previous section, $m_{\phi,\text{th}}$ is the thermal mass and $m_{\phi,\text{ln}}$ is the thermal-log mass. The thermal-log mass replaces the ordinary thermal mass when the χ -fermions are Boltzmann suppressed, i.e. for $T < y\phi$. If either of the thermal potentials is non-negligible, then it can significantly alter the dynamics of the kick. Furthermore, when the Yukawa interaction damps the radial oscillations, this energy is transferred to the SM plasma. This can further modify the parameter space if the yield dilution is significant.

The Yukawa coupling introduced in eq. (13.1) gives rise to thermal damping through a variety of mechanisms. The most relevant rates are the following: Perturbative decay,

$$\Gamma_{\phi\bar{\chi}\chi} \approx \frac{y^2 m_{\phi,0}}{8\pi} \quad \text{if } m_{\phi,0} < gT \quad \text{and} \quad m_{\phi,0} < y\phi, \quad (13.5)$$

$$\Gamma_{\phi\gamma\gamma} \approx \alpha^2 \frac{m_{\phi,0}^3}{\phi^2} \quad \text{if } m_{\phi,0} < gT \quad \text{and} \quad m_{\phi,0} > y\phi \quad (13.6)$$

$$\text{where } b = \frac{T(r) (12\pi)^2}{16\pi^2 \ln \alpha^{-1}},$$

scattering processes with χ -fermions,

$$\Gamma_{\chi\text{-scat.}} \approx \begin{cases} y^2 \alpha T & \text{if } y\phi < \alpha T, \\ y^4 \frac{\phi^2}{\alpha T} & \text{if } \alpha T < y\phi < T, \end{cases} \quad (13.7)$$

and loop-induced scattering processes with the gauge bosons γ

$$\Gamma_{\gamma\text{-scat.}} \approx \frac{b\alpha^2 T^3}{\phi^2}. \quad (13.8)$$

These damping rates enter as a friction term in the equations of motion, such that this becomes

$$\ddot{\phi} + 3H\dot{\phi} + \Gamma_{\text{tot}}\dot{\phi} + m_{\phi,\text{eff}}^2\phi = \dot{\theta}^2\phi, \quad (13.9)$$

where the last term is the centrifugal term connecting the radial dynamics to the angular dynamics and Γ_{tot} is the sum of all damping rates whose assumptions are fulfilled.

13.2 Boltzmann equations and solution approach

The energy lost through damping is transferred to the radiation bath. However, since several of the damping rates depend on ϕ and/or T , the dynamics of this system can be non-trivial. Furthermore, the PQ fields may dominate the energy density and thus control H . Therefore, we derive a set of Boltzmann equations that we solve numerically. The derivation is detailed in appendix F and the system of equations is

$$\dot{\rho}_{\text{kin}}^\phi = -(3H + \Gamma_{\text{tot}} + c_{m_\phi} H)\rho_{\text{kin}}^\phi, \quad (13.10)$$

$$\dot{\rho}_{\text{circ}} = -(3H + c_{m_\phi} H)\rho_{\text{circ}}, \quad (13.11)$$

$$\dot{\rho}_r = -4H\rho_r + \Gamma_{\text{tot}}\rho_{\text{kin}}^\phi, \quad (13.12)$$

$$\dot{\rho}_a = -c_{\rho_a} H\rho_a, \quad (13.13)$$

where ρ_{kin}^ϕ is the kinetic energy of the radial oscillations, ρ_{circ} is the potential energy associated with the circular orbit about which the radial mode is oscillating and ρ_r is the energy density of the SM plasma. We parametrize a possible temperature dependence of the radial mass with a parameter c_{m_ϕ} defined from $m_\phi \propto a^{-c_{m_\phi}}$, such that $c_{m_\phi} = 1$ if thermal effects dominate the potential and $c_{m_\phi} = 0$ otherwise. Similarly, we parametrize the axion redshift with $\rho_a \propto a^{-c_{\rho_a}}$ where $c_{\rho_a} = 3$ when $\phi \gg f_a$ and $c_{\rho_a} = 6$ when $\phi \sim f_a$. The initial conditions for the system, as well as details on the handling of χ -fermion relativistic/non-relativistic transitions, are also discussed in appendix F.

13.3 Discussion

The Yukawa interaction introduces several interesting consequences which we discuss in the following paragraphs.

Thermal potential and the absence of low- m_a solutions: One interesting consequence of the thermal potential is that it precludes the low- m_a class of solutions seen e.g. on the left-hand-side of figure 20 and 21. To understand this, recall that $Y_{\text{kick}} \propto \epsilon \times m_{\phi,\text{eff}}^{-k}$, where k is a constant depending on n . In the absence of thermal effects, the parameter ϵ can be assumed to be of $\mathcal{O}(1)$. This is a result of the balance between $A \sim \mathcal{O}(1)m_{\phi,0}$ in the PQ-breaking angular potential and the radial potential. Therefore, a thermal mass contribution will suppress ϵ . From eq. (11.4) and eq. (11.7) we find

$$\epsilon \approx \mathcal{O}(1) \times \frac{m_{\phi,0}}{m_{\phi,\text{eff}}}. \quad (13.14)$$

This suppression disfavors¹⁴ kicks with $m_{\phi,\text{eff}} \gg m_{\phi,0}$, and introduces a lower bound on $m_{\phi,0}$. Since the large yield corresponds to low m_ϕ , this affects low axion masses. As discussed in the next paragraph, this problem can be ameliorated by considering scenarios with only intermediate relativistic phases. Nevertheless, this precludes solutions with axion/saxion masses below the thermal ϕ relic bound.

Intermediate relativistic phases: Intermediate relativistic phases play an important role in overcoming the thermal mass contribution at the time of the kick. In solutions with an intermediate relativistic phase, fermions are initially non-relativistic because of a large ϕ -induced mass and turn relativistic once the ϕ -induced mass falls below T . This phenomenon allows efficient damping through scattering with relativistic fermions without the problems associated with a large thermal mass at the time of the kick. In particular, if $y\phi_{\text{kick}} > T_{\text{kick}}$ then the lower limit on the $m_{\phi,0}$ is relaxed to $m_{\phi,0} > m_{\phi,\text{ln}}$.

Because $\phi \propto a^{-3/2}$ if ϕ is not dominated by the thermal potential, it is still possible for the fermions to enter an intermediate relativistic phase with efficient damping. We find that this method of evading ϵ -suppression leads to solutions with DM for $m_a \gtrsim 10^{-6}$ eV. An explicit example of this evolution can be found among the example solutions in the appendix F, see fig. 34.

Solutions, where a thermal contribution is evaded by an early non-relativistic phase, are only consistent for some values of the reheating temperature after inflation, T_{reheat} . In particular, some choices of y impose upper bounds on T_{reheat} below which a kick in a non-relativistic phase is inconsistent. See appendix E for a discussion of such potential contradictions. We generally allow T_{reheat} to take any value between the inflation scale and T_{kick} , except for regimes in which a too large T_{reheat} would contradict a kick with non-relativistic fermions. For simplicity, we always assume $T_{\text{reheat}} > T_{\text{kick}}$. Our results have no further dependence on T_{reheat} .

Collider constraints: Since the χ fermions cannot be SM fermions and they have not yet been observed they must not be observable at the colliders available today. We already assume the χ -fermions are efficiently interacting with the SM, so to ensure compatibility with collider data, we impose that

$$m_{\chi,\text{today}} \approx y f_a \gtrsim 1 \text{ TeV}, \quad (13.15)$$

so that the χ -fermions are too heavy to be produced in present-day experiments.

Kination and gravitational waves: Because the energy density of the angular mode transitions from a^{-3} to a^{-6} evolution at a_{kin} , at which time ϕ reaches the minimum at f_a , the scenarios here can involve a period of kination. Such a period can amplify gravitational waves (GWs) from inflation.

¹⁴Note that if ϵ is suppressed then the radial mode may also experience parametric resonance which is not taken into account here. Since we favor solutions with $m_{\phi,\text{eff}} \sim m_{\phi,0}$, this does not have a major impact on our results.

This effect was studied in [37] where it was found that a period of kination leads to GWs with a peak amplitude of

$$\Omega_{\text{GW-peak}}(f_{\text{GW-peak}})h^2 \approx 2.8 \times 10^{-13} \times \left(\frac{g_*(T_{\text{kin,end}})}{106.25}\right) \left(\frac{g_{*s}(T_{\text{kin,end}})}{106.25}\right)^{-4/3} \left(\frac{E_I}{10^{16} \text{ GeV}}\right)^4 \left(\frac{e^{2N_{\text{kin}}}}{e^{10}}\right), \quad (13.16)$$

and a peak frequency of

$$f_{\text{GW-peak}} \approx 1.1 \times 10^{-3} \text{ Hz} \left(\frac{g_*(T_{\text{kin,end}})}{3.37}\right) \left(\frac{g_{*s}(T_{\text{kin,end}})}{3.91}\right)^{-4/3} \left(\frac{E_{\text{kin}}}{10 \text{ TeV}}\right) \left(\frac{e^{N_{\text{kin}}/2}}{10}\right). \quad (13.17)$$

Here $T_{\text{kin,end}}$ is the temperature at the end of kination domination, N_{kin} is the number of e-folds of inflation and E_I is the energy scale of inflation which is constrained to $E_I < 1.6 \times 10^{16}$ GeV by Planck [241]. The peak frequency is controlled by the energy scale of kination,

$$E_{\text{kin}} = \rho_\theta(T_{\text{kin}}) = \frac{1}{2}m_\phi^2 f_a^2. \quad (13.18)$$

Therefore, the frequency of the GW peak depends on the radial mass m_ϕ . We discuss the observability of such a signal in the results section. However, the suppression of $A_s(k_{\text{kin}})$, which is required to avoid homogeneity constraints, might also suppress the stochastic GW background from inflation. We do not take any such impact into account and leave such modifications to future work.

Visualization parameter space: Even when we fix the strength of the gauge interaction to the reference value $\alpha_\chi = 0.1$, the model we are considering has six parameters: m_a, f_a, y, M, n and $m_{\phi,0}$. This makes visualizing the viable parameter space challenging. For simplicity, we only consider $M = m_{\text{Pl}}$ and only show results for $n = 13$, although alternative choices of n can be found in appendix G. The dimensionality is further reduced to 3 by enforcing that the axion relic matches the observed DM density. Because of yield dilution, for any given choice of $[m_a, f_a]$ the correct DM relic can be realized by a family of solutions with varying $m_{\phi,0}$ and y . We visualize the viable parameter in the two following ways:

- a) To visualize the impact of the most important constraints, we show the parameter space that would become viable if a given constraint is *relaxed*. Any such region is indicated by colored hatching. Parameter space that cannot be made viable by the relaxation of any single condition is shown in gray.
- b) To visualize parameter values, we plot contours of the highest and/or lowest values possible in any given point of $[m_a, f_a]$.

In particular, *a)* implies that if two constraints overlap in our visualization, then that point of $[m_a, f_a]$ space becomes viable if *either* constraint is relaxed. This is quite different from usual visualizations of two-dimensional parameter spaces and is a consequence of our three-dimensional parameter space being projected down onto the two-dimensional $[m_a, f_a]$ plane.

13.4 Results

By numerically solving the Boltzmann equations (13.10)-(13.13) we can sample the across the parameter space $[f_a, y, M, n, m_{\phi,0}]$ and robustly identify solutions that correspond to the correct DM relic for a given m_a . Solutions both with and without yield dilution can be found with all solutions featuring (s)axion masses above the thermal ϕ relic bound. We have investigated the choices $n = 7, 10$, and $n = 13$. Of these choices, $n = 13$ supports most parameter space wherefore we focus on this case. Solutions for $n = 7$ and $n = 13$ can be found in appendix G. Interestingly,

there are regions of parameter space in which the radial mass can be large enough to saturate the perturbativity condition $m_{\phi,0}^2 \leq 4\pi f_a^2$. The full range of parameter values that are required to support these solutions are detailed in appendix G. From our results, we highlight the following:

Dominant damping rates: We find that the process most commonly responsible for damping is scattering with relativistic χ -fermions as given by eq. (13.7). Damping tends to take place at temperatures near either the end or the beginning of a phase of relativistic χ -fermions. Damping near the beginning of a relativistic phase is common because of lifting of the χ -fermion Boltzmann suppression rapidly increases the damping rate. Damping near the end of a relativistic phase is common because the $\Gamma_{\chi\text{-scat.}} \propto T^{-1}$ behavior rapidly out-competes $H \propto T^2$ in the short regime, $\alpha T < y\phi < T$, in which the damping rate is relevant. Damping near the end of a relativistic phase can be problematic if the damping does not complete before Boltzmann suppression closes the process down. The ability to explicitly confirm complete damping in such transition scenarios is another advantage of our numerical solution. See appendix F for explicit examples of these scenarios.

Aside from χ -scattering, saxion-to-axion decay can also play an important role in the damping of radial oscillations. As discussed in section 11.4 and appendix C, this mechanism is constrained by the danger of producing a hot axion relic in excess of constraints on dark radiation. This limits saxion-to-axion decay as a stand-alone damping mechanism. Interestingly, we find that when saxion-to-axion decay is considered in conjunction with χ -scattering then the decay can account for the remaining energy density which is left behind after χ -scattering has damped most, but not all, of energy out of radial oscillations. The synergy in such two-step damping lies in the fact that after χ -scattering has reduced the energy in radial oscillations to a fraction of ρ_r saxion-to-axion decay can take place without the resulting hot axion being a significant part of the overall energy density. An explicit example of such two-step damping can be found in appendix F.

Constraints from saxion-to-axion decay: The condition given in eq. (11.22) leads to the exclusion of solutions with large (s)axion masses. This condition ensures that saxion-to-axion decay does not lead to the production of a relic of hot axions dominating the thermal plasma. In some parts of parameter space, this condition is poorly captured by our visualization because high- m_a solutions are shifted by lower- m_a when saxion-to-axion decay reduces yield dilution. This applies in particular to solutions with damping triggered by an intermediate relativistic phase which would have featured large yield dilution if the energy in radial oscillations had been dumped into radiation rather than hot axions. To better highlight parameter space excluded by this effect, we have performed an analytic approximation of the constraint as applied to this specific family of solutions and indicated the upper bound on m_a (and thus implicitly the upper bound on m_ϕ) with a red dashed line. An $\mathcal{O}(1)$ constant was chosen to match the analytic solution to the numerical results. Because this analytic bound only applies to solutions with damping triggered by an intermediate relativistic phase some solutions featuring late damping not triggered by an intermediate relativistic phase can be found at m_a higher than the indicated dashed line.

Necessity of $A_s(\mathbf{k}_{\text{kin}})$ suppression: All solutions found involve an extended duration of a^{-6} evolution of the energy density of the angular mode. This leads to significant constraints from the homogeneity condition eq. (11.12), which applies if the spectrum of density perturbations from inflation A_s is assumed to be a flat spectrum. If the bound on A_s from CMB measurements, i.e. $A_{s,\text{Planck}} \sim 2.1 \times 10^{-9}$ [241], is assumed to apply on all scales then Yukawa damping is not viable. However, the relevant scale k_{kin}^{-1} is many orders of magnitude smaller ($k_{\text{kin}}^{-1} \ll 10^{-10}$ Mpc) than those constrained by CMB measurements ($k \sim 0.05$ Mpc $^{-1}$), wherefore $A_s(k_{\text{kin}})$ is allowed to differ by many orders of magnitude from the value suggested by CMB measurements without tension.

If $A_s(k_{\text{kin}})$ is assumed to be sufficiently smaller than $A_s \sim 2.1 \times 10^{-9}$, then a region around $m_a \sim 10^{-6}$ eV to 10^{-2} eV becomes viable for a benchmark value of $n = 13$. We display this region in figure 22 along with the necessary $A_s(k_{\text{kin}})$ suppression and the surrounding constraints. We display the associated values of k_{kin} in appendix G. Assuming $A_s(k_{\text{kin}}) \ll 2.1 \times 10^{-9}$ corresponds to relaxation of the homogeneity constraint seen e.g. in figure 20.

Realized range of damping temperatures: For scenarios in which χ -scattering is the sole relevant damping rate, we for $n = 13$ find scenarios with damping temperatures that generally lie in the range $T_{\text{damp}} \sim 10^4$ GeV to 10^{12} GeV. When saxion-to-axion decay is also considered, then scenarios can be found in which some energy remains in radial oscillations down to $T_{\text{damp}} \gtrsim$ GeV. This may appear to be in conflict with the results seen in section 12.4. However, because damping through axion-to-saxion decay does not establish thermal contact with the plasma damping through this channel avoids thermal relic constraints given by eq. 11.14. Therefore, lower damping temperatures than those suggested by e.g. 21 can be achieved without conflict.

Furthermore, the definition of T_{damp} is made more complex by the possibility of two-step damping. For the analytic approach of section 12.4 we considered an instantaneous damping in which T_{damp} is unambiguously defined. For the numerical solutions, damping generally does not take place at any singular time, so we must be more careful with the definition of T_{damp} . We here define T_{damp} as the temperature at which the energy in radial oscillations is damped below all other energy densities of the problem, i.e. we demand $\rho_{\text{kin}}^\phi < \rho_r, \rho_a$ for all $T < T_{\text{damp}}$ in addition to Γ_{tot} . The possibility of two-step damping implies that T_{damp} is not necessarily anywhere near the temperature at which the majority of energy is transferred from the PQ field to the plasma. Crucially, the bulk of the energy in two-step solutions is transferred by χ -scattering much earlier than T_{damp} . In such scenarios, the behavior is quite different from the simple scenario of instantaneous damping, and the analysis of section 12.4 should not be expected to apply.

With the knowledge that the very different behavior of T_{damp} in two-step scenarios in some circumstances circumvent some lower bounds discussed in section 12.4, our numerical solution otherwise realizes a subset of the temperature range derived in that section. This range is visualized in figure 26, which can be compared to the general case visualized in figure 21. Again, note that each point in the $[m_a, f_a]$ plane may be covered by a range of solutions with different combinations of m_ϕ and y , such that each point in our $[m_a, f_a]$ parameter space is associated with a range of possible temperatures. However, note that the color gradient between the two plots cannot be compared because the range has been chosen to best fit each plot individually.

The low damping temperatures permitted around $m_a \sim 10^{-4}$ eV and $m_a \sim 10^{-1}$ eV correspond to two-step damping in which the damping is finalized by saxion-to-axion decay without violation of the thermal ϕ relic constraints that enforced a higher lower bound in figure 21.

Gravitational waves: In figure 23 we map out how long a duration of kination can be realized in this setup. Such long durations may lead to an observable GW signal if the necessity of $A_s(k_{\text{kin}})$ suppression does not interfere with the primordial spectrum. Since the frequency and amplitude of any such signal depend on m_ϕ and because each point in $[m_a, f_a]$ parameter space can, in general, be supported by multiple combinations of m_ϕ and Yukawa coupling y , each such point will have a family of solutions of varying peak frequency and amplitude. Within such a family, changes in m_ϕ are compensated by changes in yield dilution such that the present-day yield is constant. Increased yield dilution shortens the duration of kination and also reduces the peak frequency. In fig. 24, we display the region of $[f_{\text{GW-peak}}, \Omega_{\text{GW-peak}} h^2]$ parameter space that is populated by KMM solutions. In this figure, we also visualize the effect of yield dilution by displaying a family of solutions with constant $[m_a, f_a]$. Future gravitational wave observatories such as Big Bang Observer [242, 243] and DECi-hertz Interferometer Gravitational wave Observatory (DECIGO) [244] may be sensitive

enough to detect the stochastic background from inflation if this takes place at the largest energy scales permitted by Planck. For the maximal value of $E_{I,\text{max}} \sim 1.6 \times 10^{16}$ GeV [245], even a small amplification of such background will be observable if the relevant frequency falls within the BBO or DECIGO range. Therefore, it is possible that much of the parameter space supported in this scenario could be observed by BBO or DECIGO. Significant regions of parameter space are also within the potential reach of Einstein Telescope [246] (ET) and Cosmic Explorer [247] (CE). The $[m_a, f_a]$ parameter space which contains solutions in reach of these observatories is presented in fig. 25. However, since $\Omega_{\text{GW-peak}}(f_{\text{GW-peak}}) \propto E_I^4$, the signal is significantly reduced if E_I is not near the maximal value permitted by Planck.

Yukawa: Necessary $A_s(k_{\text{kin}})$ suppression for $M=m_{\text{Pl}}$ and $n=13$

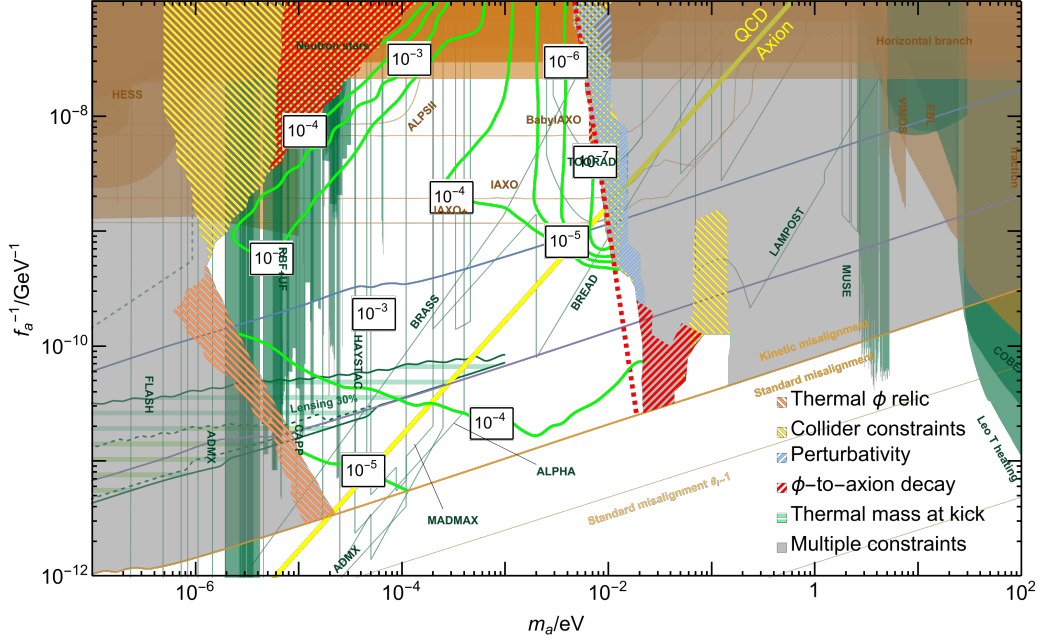


Figure 22. Map of necessary suppression of $A_s(k_{\text{kin}})$ relative to $A_{s,\text{Planck}}$, where A_s is the amplitude of the primordial power spectrum, see eq. (11.12). Constraints on the implementation are indicated by the hatched regions. The dominant constraints are from the thermal ϕ relic (11.14), collider constraints (13.15), and hot axion relics produced by saxion-to-axion decay (11.22). Because each point in $[m_a, f_a]$ space may be supported by a range of solutions, we illustrate the constraints by highlighting regions that would become viable if a given constraint is lifted.

Yukawa: Contours of e-folds of kination for $M=m_{\text{Pl}}$ and $n=13$

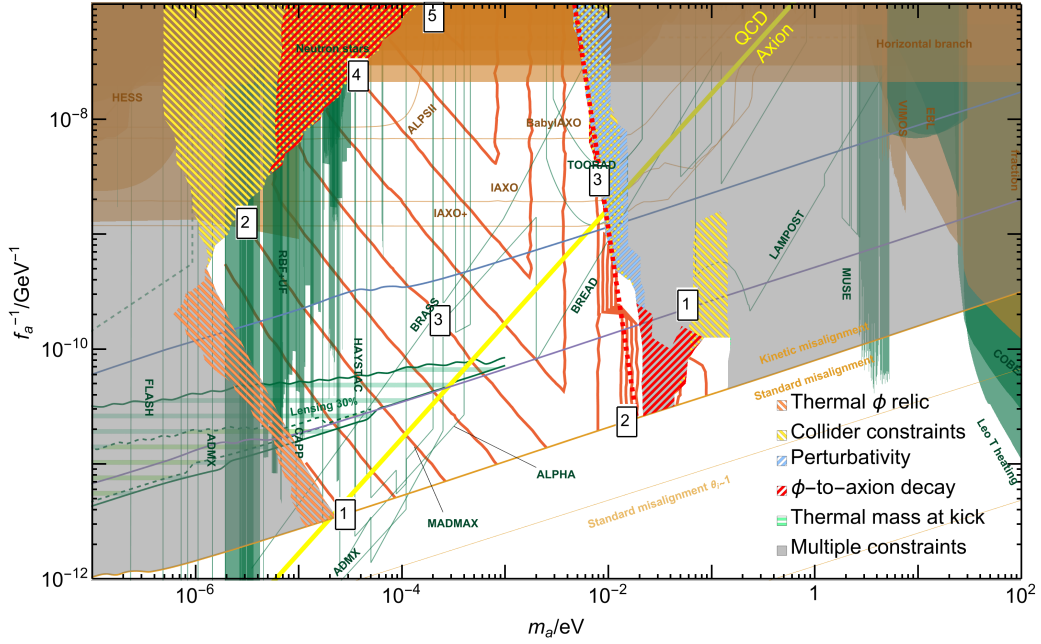


Figure 23. Map of the largest number of e-folds of kination which can be realized if it is assumed that the curvature fluctuations can be sufficiently suppressed. Each point in the $[m_a, f_a]$ parameter space, in general, has a range of solutions with varying radial mode mass m_ϕ and the Yukawa coupling y specified by eq. (13.1). The contours given here indicate the largest possible number of e-folds at each point.

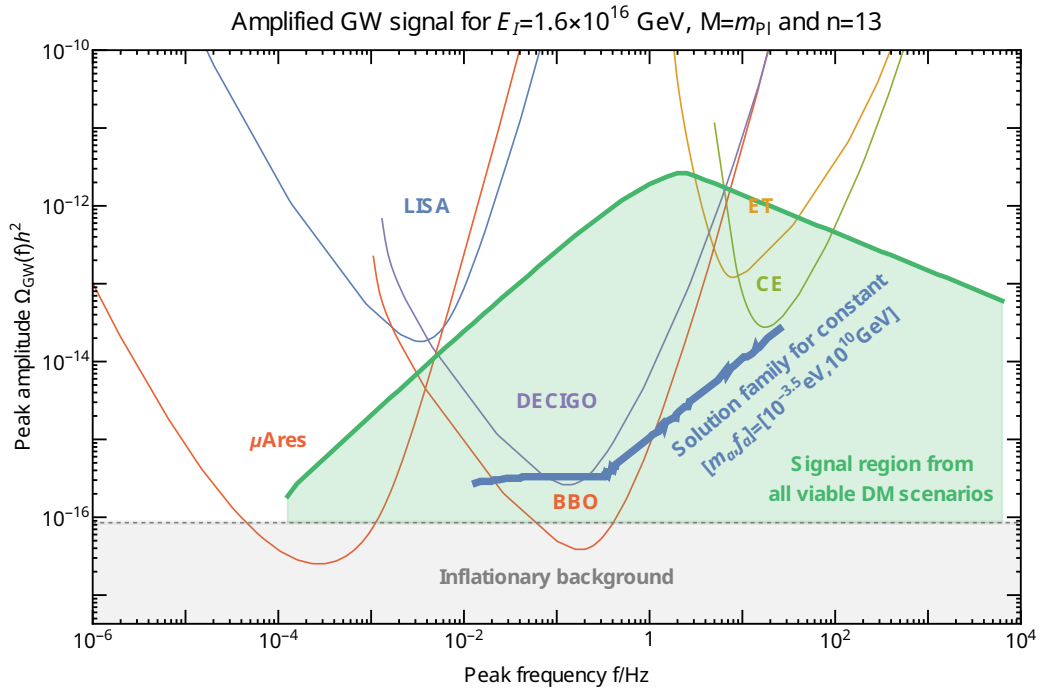


Figure 24. Region in the $[\Omega_{GW}(f)h^2, f]$ plane in which gravitational waves from inflation may be amplified by kination in the nearly-quadratic model. Crucially, the green region is not a frequency spectrum. Instead, the region indicates possible locations of the peak frequency f and amplitude $\Omega_{GW}(f)$ of the GW signal. The effect of entropy injection is visualized by the blue line, which represents a family of solutions with constant $[m_a, f_a]$. Along this line, increased entropy injection reduces both frequency and amplitude of the GW signal.

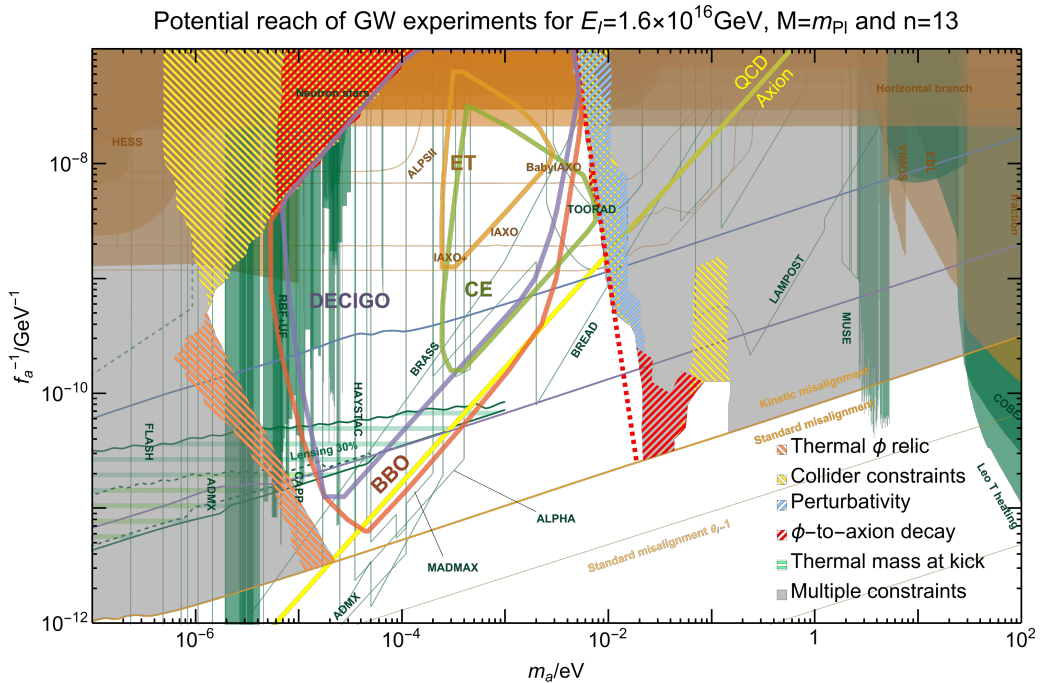
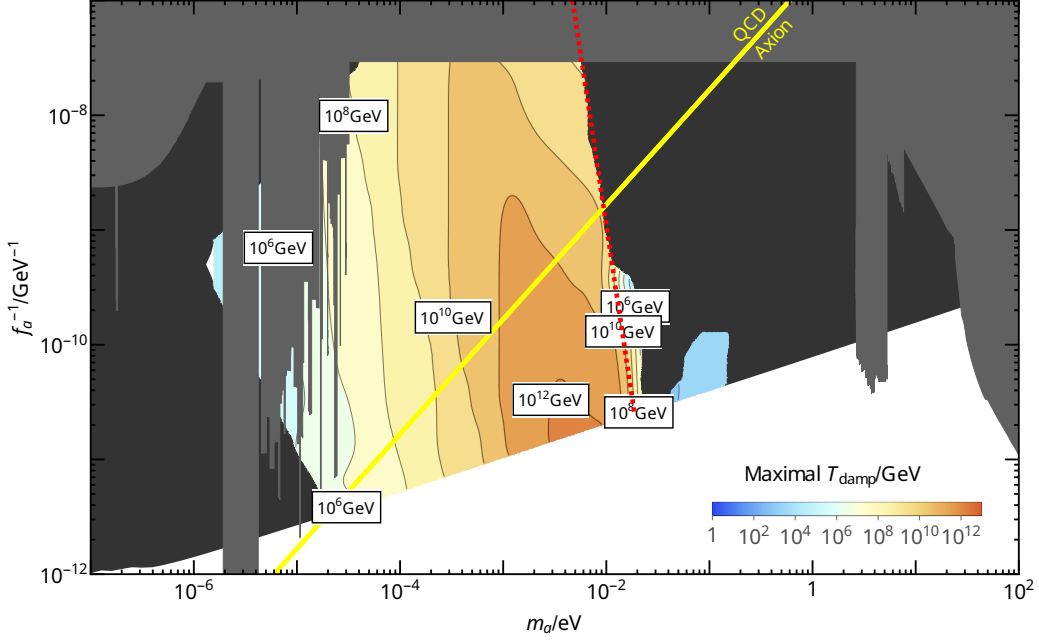


Figure 25. Regions in $[m_a, f_a]$ parameter space in which amplification of the inflationary GW background may be observable by the indicated GW observatories.

Yukawa: Contours of upper bounds on T_{damp} for $M=m_{\text{Pl}}$ and $n=13$.



Yukawa: Contours of lower bounds on T_{damp} for $M=m_{\text{Pl}}$ and $n=13$.

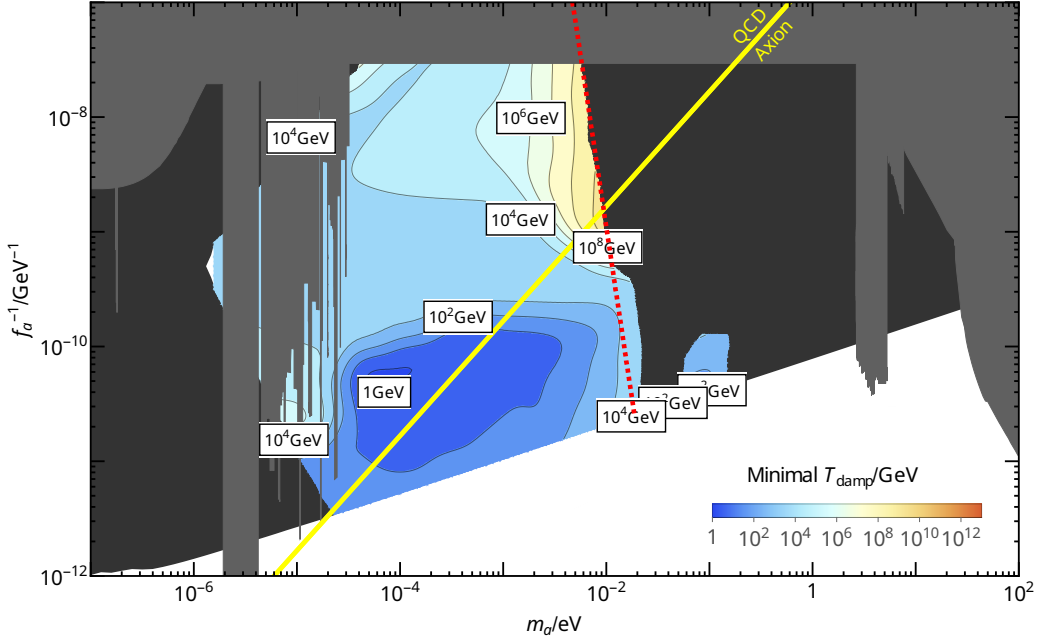


Figure 26. Realized range of damping temperatures. Here the homogeneity condition (11.12) has been relaxed by assuming that $A_s(k_{\text{kin}})$ is sufficiently suppressed. Therefore, the realized temperatures are much higher than those suggested by figure 21. This can be understood by referring to fig. 20, where relaxation of the homogeneity condition unlocks the red-shaded region. **Top:** Contours of the highest possible thermalization temperature. **Bottom:** Contours of the lowest possible thermalization temperature

14 Models with quartic potentials

The next model we consider features a quartic potential. Such a potential is interesting because it is the simplest potential that features spontaneous symmetry breaking. Specifically, we consider the potential of the form

$$V_{\text{late}} = \lambda^2 \left(|P|^2 - \frac{f_a^2}{2} \right)^2. \quad (14.1)$$

The effective radial mass provided by this potential is

$$m_\phi^2 = \frac{1}{\phi} V'_{\text{late}} \approx \lambda^2 \phi^2. \quad (14.2)$$

The quartic model fundamentally differs from the nearly-quadratic model in that the quartic mass depends linearly on ϕ , while the nearly-quadratic model features a mass approximately independent of ϕ . The property that m_ϕ becomes large at large field values presents a challenge for the generation of the VEV via higher-dimensional terms. In appendix H.1 we show that the higher-dimensional terms used in the nearly-quadratic model only lead to a limited amount of parameter space in the quartic model, wherefore we do not consider such terms further as a high-VEV driver in the quartic model. Instead, we assume that both the radial and angular modes are light during inflation such that a large VEV is generated by de Sitter fluctuations. However, note that this inherently makes the quartic model somewhat tuned in that a very low quartic coupling is required, e.g., $\lambda \sim 10^{-20}$.

14.1 Initial conditions

We assume that the kick is driven by the same type of PQ-violating terms as in the nearly-quadratic model so that the PQ-violating part of the potential is

$$V_{\mathcal{PQ}} = \frac{A(P^n + P^{*n})}{nm_{\text{Pl}}^{n-3}}. \quad (14.3)$$

The complete zero-temperature potential at early times is $V_{\text{early}} = V_{\text{late}} + V_{\mathcal{PQ}}$. Once P starts oscillating at $m_\phi \approx 3H$, the kick proceeds as described in section 11 and, as in the nearly-quadratic case, the initial angular velocity is given by eq. (11.4). However, in the present case the SUSY-motivated assumption of $A \approx \mathcal{O}(1) \times m_\phi$ does not yield a kick that naturally results in $\epsilon \sim \mathcal{O}(1)$. Therefore, the quartic model instead assumes that the parameter A can be chosen to an arbitrary value such that $\epsilon \sim \mathcal{O}(1)$ is fixed by hand [33]. This point will be further discussed in section 14.2.

If a field is lighter than H_I during inflation, then fluctuations grow [248, 249]. This growth of de Sitter fluctuations has previously been exploited to generate a large VEV for the radial in kinetic misalignment [33]. In particular, if the radial mode is lighter than the inflation scale H_I , then the expectation value of the VEV grows up to a maximal value of

$$\phi_{\text{kick}}^2 \sim \frac{3}{8\pi^2} \frac{H_I^4}{m_\phi^2} \sim \frac{3}{8\pi^2} \frac{H_I^4}{\phi_{\text{kick}}^2 \lambda^2}, \quad (14.4)$$

which implies an initial VEV of

$$\phi_{\text{kick}} \sim \sqrt[4]{\frac{3}{8\pi^2} \frac{H_I}{\sqrt{\lambda}}}. \quad (14.5)$$

Notably, the dependence on the quartic coupling λ drops out in the potential initial energy such that the energy density of ϕ at the time of the kick only depends on H_I :

$$\rho_{\phi, \text{pot, kick}} \approx \frac{1}{2} (\lambda \phi_{\text{kick}})^2 \phi_{\text{kick}}^2 \approx \frac{3}{16\pi^2} H_I^4. \quad (14.6)$$

To avoid a period of inflation generated by ϕ before the field starts oscillating, we demand that $\rho_{\phi,\text{pot}}$ does not dominate the energy density at the time of the kick. Furthermore, we assume that the kick takes place after the reheating of the inflaton so that radiation dominates the universe at the time of the kick. The kick temperature T_{kick} is fixed from $m_{\phi,\text{kick}} \approx \lambda\phi_{\text{kick}} \approx 3H$, which implies

$$T_{\text{kick}} \approx 0.2 \times \lambda^{1/4} \sqrt{H_I m_{\text{Pl}}}, \quad (14.7)$$

$$\rho_r(T_{\text{kick}}) \approx \frac{1}{2\sqrt{6}\pi} \lambda H_I^2 m_{\text{Pl}}^2. \quad (14.8)$$

Both the saxion and axion energy densities redshift as a^{-4} until the radial mode is relaxed to its minimum at T_{kin} . This can be verified by the WKB approximation according to which $\phi \propto m_\phi^{-1/2} a^{-3/2}$ such that $\phi \propto a^{-1}$ if $m_\phi \propto \phi$. After the radial mode is relaxed to f_a the mass becomes constant $m_\phi \sim \lambda f_a$, wherefore $\rho_\phi \propto a^{-3}$. As in the nearly-quadratic model PQ charge conservation ensures $\rho_a \propto a^{-6}$ after T_{kin} .

If the universe is radiation-dominated at the time of the kick, then radiation will stay dominant at least until T_{kin} . Because of this, kination can never take place in the quartic model and entropy dilution is possible only if there is enough time between T_{kin} and T_{damp} for ρ_ϕ , which redshifts as a^{-3} after T_{kin} , to dominate ρ_r .

Estimation of the produced yield: The yield produced by the kick is

$$Y_{\text{kick}} \approx \left. \frac{n_{PQ}}{s} \right|_{\text{kick}} \approx \epsilon \frac{\lambda \phi_{\text{kick}}^3}{\frac{2\pi^2}{45} g_* s T_{\text{kick}}^3}, \quad (14.9)$$

which in terms of model parameters corresponds to

$$Y_{\text{kick}} \approx 0.2 \times \frac{\epsilon}{\lambda^{5/4}} \left(\frac{H_I}{m_{\text{Pl}}} \right)^{3/2}, \quad (14.10)$$

where the prefactor has mild dependence on $g_*(T_{\text{kick}})$.

Solution for dark matter in the absence of yield dilution: If no yield dilution takes place, then the yield given eq. (14.10) corresponds to the present-day yield. To match the observed DM yield, i.e. $Y_{\text{DM}} \approx 0.64 \frac{T_{\text{eq}}}{m_a}$, we choose λ such that $2Y_{\text{kick}} = Y_{\text{DM}}$. This implies a present-day radial mass of

$$m_{\phi,\text{today,DM}} \approx 0.7 \times f_a \left(\frac{H_I}{m_{\text{Pl}}} \right)^{6/5} \left(\frac{m_a \epsilon}{T_{\text{eq}}} \right)^{4/5}. \quad (14.11)$$

The quartic coupling is also fixed from eq. (14.11) and can be obtained from $\lambda = m_{\phi,\text{today}}/f_a$.

14.2 Application of constraints

We here discuss the constraints on the quartic model. As in the nearly-quadratic case, the quartic model is subject to constraints from the growth of fluctuations (the homogeneity condition) and the constraints from the radial thermal relic.

The nearly-quadratic model was subject to constraints from perturbativity and from the need for the kick to take place before T_* . For the quartic model, the perturbativity constraint is not relevant because the model involves very low values of the quartic coupling λ . Furthermore, for any realistic value of H_I , the kick in the quartic model will always take place before the T_* , so this condition also does not impose a significant constraint.

Because the axion and saxion are assumed to be light during inflation in the quartic model, this scenario is subject to additional constraints. In particular, both modes will contain fluctuations that give rise to isocurvature and domain wall constraints, which we will discuss in the following sections. Finally, we discuss limits on the amplitude of the PQ-violating potential.

Isocurvature constraint: Since both degrees of freedom are assumed to be light during inflation, quantum fluctuations in both the angular mode and the radial mode will grow [248, 249]. After inflation, we expect fluctuations of size

$$\Delta\phi_{\text{kick}} \sim \frac{H_I}{2\pi} \quad \text{and} \quad \Delta\theta_{\text{kick}} \sim \frac{H_I}{2\pi\phi_{\text{kick}}}. \quad (14.12)$$

These fluctuations will source isocurvature fluctuations in the dark matter relic. Such fluctuations are tightly constrained on CMB scales. As was argued in [33], the yield fluctuations can be related to the power spectrum of cold dark matter isocurvature perturbations,

$$\mathcal{P}_\phi(k) = \left\langle \left(\frac{\delta\Omega_{DM}}{\Omega_{DM}} \right)^2 \right\rangle = \left\langle \left(\frac{\delta Y}{Y} \right)^2 \right\rangle. \quad (14.13)$$

Assuming a pivot scale of $k \sim 0.05 \text{ Mpc}^{-1}$ this power spectrum is constrained to

$$\mathcal{P}_\phi(k) < 8.7 \times 10^{-11}. \quad (14.14)$$

Since the yield Y is a function of the angular axion field θ and radial saxion field ϕ , we can express the fluctuations as

$$\mathcal{P}_\phi(k) = \left\langle \left(\frac{\delta Y}{Y} \right)^2 \right\rangle = \left(\frac{1}{Y} \frac{\partial Y}{\partial \theta} \right)^2 \langle \delta\theta_{\text{kick}}^2 \rangle + \left(\frac{1}{Y} \frac{\partial Y}{\partial \phi} \right)^2 \langle \delta\phi_{\text{kick}}^2 \rangle. \quad (14.15)$$

With $Y = Y_{\text{kick}}$, this yields

$$\mathcal{P}_\phi(k) = \frac{n^2}{N_{dw}^2} \cot^2 \left(\frac{n\theta_{\text{kick}}}{N_{dw}} \right)^2 \left(\frac{H_I}{2\pi\phi_{\text{kick}}} \right)^2 + \frac{(5-2n)^2}{4\phi_{\text{kick}}^2} \left(\frac{H_I}{2\pi} \right)^2, \quad (14.16)$$

where we neglected any influence of entropy injection if such takes place. Note that dependence of θ in Y enters through ϵ as given by eqs. (11.4) and (11.7). Since ϕ_{kick} is related to H_I through eq. (14.4) we can eliminate the powers of H_I/ϕ_{kick} from the spectrum:

$$\mathcal{P}_\phi(k) = \frac{\lambda n^2}{\sqrt{6}\pi N_{dw}^2} \cot^2 \left(\frac{n\theta_{\text{kick}}}{N_{dw}} \right)^2 + \frac{\lambda(5-2n)^2}{4\sqrt{6}\pi} \quad (14.17)$$

The angular contribution could be removed by tuning of the cotangent. However, the two contributions are expected to be of similar magnitude, so there is little advantage to tuning in this scenario. If one takes the cot term to be of $\mathcal{O}(1)$, then the contributions are of similar importance, and the constraint is

$$\lambda \lesssim \mathcal{O}(1) \times 10^{-11}, \quad (14.18)$$

where the $\mathcal{O}(1)$ constant depends on n and the exact value of the initial angular misalignment θ_{kick} .

Domain wall constraint: Inflationary quantum fluctuations can also lead to a domain wall problem. This may take place because fluctuations in the initial angular velocity over time will lead to slightly different angular displacements in different Hubble patches. If such angular fluctuations have become $\mathcal{O}(1)$ at the time of trapping, then we expect the field to form a configuration where different patches are separated by domain walls. If this configuration is stable, then the energy in the domain walls will eventually dominate the energy density of the universe, which is in contradiction with observed cosmology.

To estimate whether inflationary perturbations can induce a DW constraint, we integrate the total angular displacement:

$$\theta_* \approx \int_{t_{\text{kick}}}^{t_*} \dot{\theta} dt \quad (14.19)$$

This integration is carried out in appendix H.2, where we find that $\mathcal{O}(1)$ angular fluctuations are generated unless

$$\frac{H_I}{f_a} \lesssim \mathcal{O}(0.1), \quad (14.20)$$

where the exact value of the $\mathcal{O}(0.1)$ factor depends on the choice of n in the PQ-violating potential and on the exact initial condition of θ .

Co et al. [33] addresses this problem by invoking parametric resonance to resolve the domain wall problem by non-thermally restoring PQ-symmetry temporarily and thereby generating cosmic strings. Such strings would render the resulting string-wall network unstable. However, in the absence of a more detailed analysis of the impact of such an event, we do not consider this solution to the DW problem.

Homogeneity condition on a^{-6} evolution: As discussed in section 11.4, fluctuations grow in relative importance during a^{-6} evolution of the axion. This effect implies an upper bound on the amplitude of curvature perturbations $A_s(k_{\text{kin}}) < (a_{\text{kin}}/a_*)^2$. In the present model, we find that

$$\left(\frac{a_{\text{kin}}}{a_*}\right)^2 \approx 2 \times 10^{-7} \left(\frac{f_a}{10^9 \text{ GeV}}\right)^{-.86} \left(\frac{m_a}{\text{eV}}\right)^{-.059} \left(\frac{H_I}{10^9 \text{ GeV}}\right)^{4/5}, \quad (14.21)$$

if we assume the no entropy injection takes place. If the spectrum of curvature perturbations is flat, then we must demand that $A_{s,\text{Planck}} = 2.1 \times 10^{-9} < (a_{\text{kin}}/a_*)^2$ [241]. If this condition is violated, then fluctuations may dominate the relic generated from n_{PQ} unless the spectrum of fluctuations can be suppressed on the relevant scales.

Limits on the PQ-violating potential: Recall that the initial rotation is driven by the higher-dimensional PQ-violating potential:

$$V_{\text{PQ}} = \frac{A(P^n + P^{*n})}{nm_{\text{Pl}}^{n-3}} \quad (14.22)$$

In the SUSY motivated nearly-quadratic model one would expect that $A \sim m_{\phi,0}$, which in that model naturally leads to $\epsilon \sim \mathcal{O}(1)$. However, in a quartic model, $A \sim m_{\phi,0}$ does not naturally lead to $\epsilon \sim \mathcal{O}(1)$. This is a consequence of the initial amplitude, which in the quartic model depends on H_I instead of on $H(T)$. Therefore, factors of m_{Pl} in the kick estimate do not cancel, and A must be chosen to compensate. Specifically, in order to get $\epsilon \sim \mathcal{O}(1)$ the parameter A needs to be

$$A \sim 9. \times 10^{-3} \times e^{1.16n} m_{\text{Pl}} \left(\frac{m_{\text{Pl}}}{H_I}\right)^{n-4} \lambda^{n/2}. \quad (14.23)$$

This result is in agreement with [33]. Depending on H_I and λ this parameter A may need to take super-Planckian values to realize an $\epsilon \sim 1$ kick. We will here stay agnostic to how such a term may be motivated, but we impose that $A < m_{\text{Pl}}$, which implies an upper bound on λ :

$$\lambda < \mathcal{O}(0.1) \times \left(\frac{H_I}{m_{\text{Pl}}}\right)^{2-\frac{8}{n}}. \quad (14.24)$$

If one assumes that no yield dilution takes place, then this corresponds to an upper bound on the axion mass that can account for the observed DM relic:

$$m_a < \mathcal{O}(1) \times T_{\text{eq}} \left(\frac{H_I}{m_{\text{Pl}}} \right)^{1 - \frac{10}{n}}. \quad (14.25)$$

Although this constraint could potentially present a challenge, we find the thermal relics and domain wall constraint impose stronger bounds on the parameter space.

Limits from second inflation: If P dominates the energy density of the universe prior to the kick, then it would drive a second period of inflation which is inconsistent with the KMM scenario [33]. To remain as agnostic as possible, we assume that reheating of the inflaton takes place before the kick, such that the avoidance of a second period of inflation driven by P amounts to the condition that radiation dominates the energy density of the PQ field at the time of the kick. Since λ only appears in the energy density of radiation and not in ρ_ϕ , there is a minimum threshold for λ below which we do not have radiation domination at the time of the kick. Radiation domination is ensured for

$$\lambda > \frac{3}{4\pi} \sqrt{\frac{3}{2}} \left(\frac{H_I}{m_{\text{Pl}}} \right)^2. \quad (14.26)$$

If the quartic coupling is fixed according to eq. (14.11) such that Y_{kick} corresponds to the observed DM relic, then this condition implies a lower bound on m_a of

$$0.6 \frac{H_I T_{\text{eq}}}{m_{\text{Pl}}} \lesssim m_a. \quad (14.27)$$

A related lower bound is found by demanding $\phi_{\text{kick}} < m_{\text{Pl}}$. This condition is somewhat weaker than eq. (14.27) and implies $0.4 H_I T_{\text{eq}} / m_{\text{Pl}} \lesssim m_a$.

Limits from eternal inflation: The model studied here relies on de Sitter fluctuation to establish the initial condition for the kick, see eq. (14.5). However, because the radial masses here are very light, this equilibrium configuration takes a very large number of e-folds of inflation to establish. By solving the Fokker-Planck equation, one finds that the required number of e-folds, N , is on the order of

$$N \gtrsim N_{\text{min}} \approx \frac{3}{2} \frac{H_I^2}{(\lambda \phi_{\text{kick}})^2}. \quad (14.28)$$

If the duration of inflation, as measured by N , is too long, then the theory enters the regime of eternal inflation. The condition to avoid this is

$$N_{\text{min}} \lesssim \frac{m_{\text{Pl}}^2}{H_I^2}. \quad (14.29)$$

We in fig. 27 provide both contours of the minimum number of e-folds given by eq. (14.28) and map out the region in which the equilibrium distribution cannot be reached without violating the upper bound (14.29).

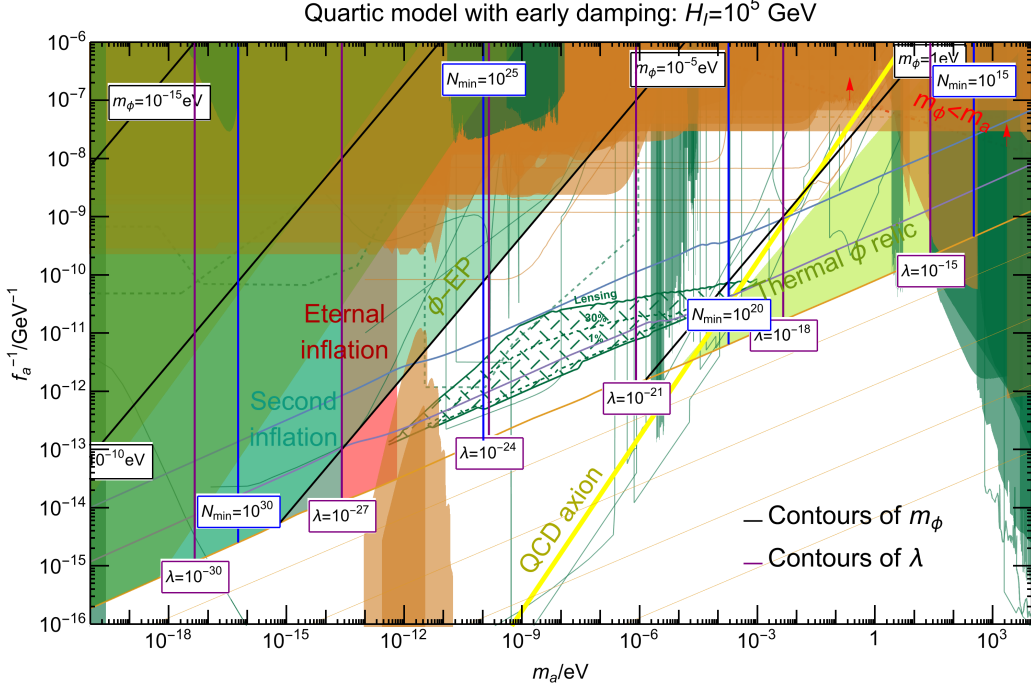


Figure 27. Viable parameter space in the quartic model under the assumption of early damping. The top plot assumes a Hubble scale during inflation of $H_I = 10^5$ GeV. The correct DM yield can be achieved in every unshaded point. The constraints from a second period of inflation, domain walls, and the thermal ϕ relic are given by eqs. (14.27), (14.20) and (11.14), respectively. We also show contours of λ , $m_{\phi, \text{today}} = \lambda f_a$ and N_{min} as defined by eq. (14.28). The thermal relic constraints may depend on the damping mechanism, see appendix C.

14.3 Early damping

As before, we first consider the simplest scenario in which radial damping proceeds through an unspecified mechanism and takes place sufficiently early that any yield dilution is insignificant. This is the case when damping takes place before the energy density of the radial modes dominates the plasma, i.e. when $T_{\text{damp}} > T_{\text{dom}}$, where T_{dom} is specified by eq. (14.31).

In figure 27 and figure 28 we display the above-mentioned constraints on parameter space for various choices of H_I . In general, the combination of DM isocurvature constraints and the thermal relic constraints prefer solutions with lower H_I . In particular, lower values of m_a require lower values of H_I . Solutions for the QCD axion are found only for $H_I \lesssim 10^5$ GeV and no solutions for ALPs of any mass are found for $H_I \gtrsim 10^{10}$ GeV.

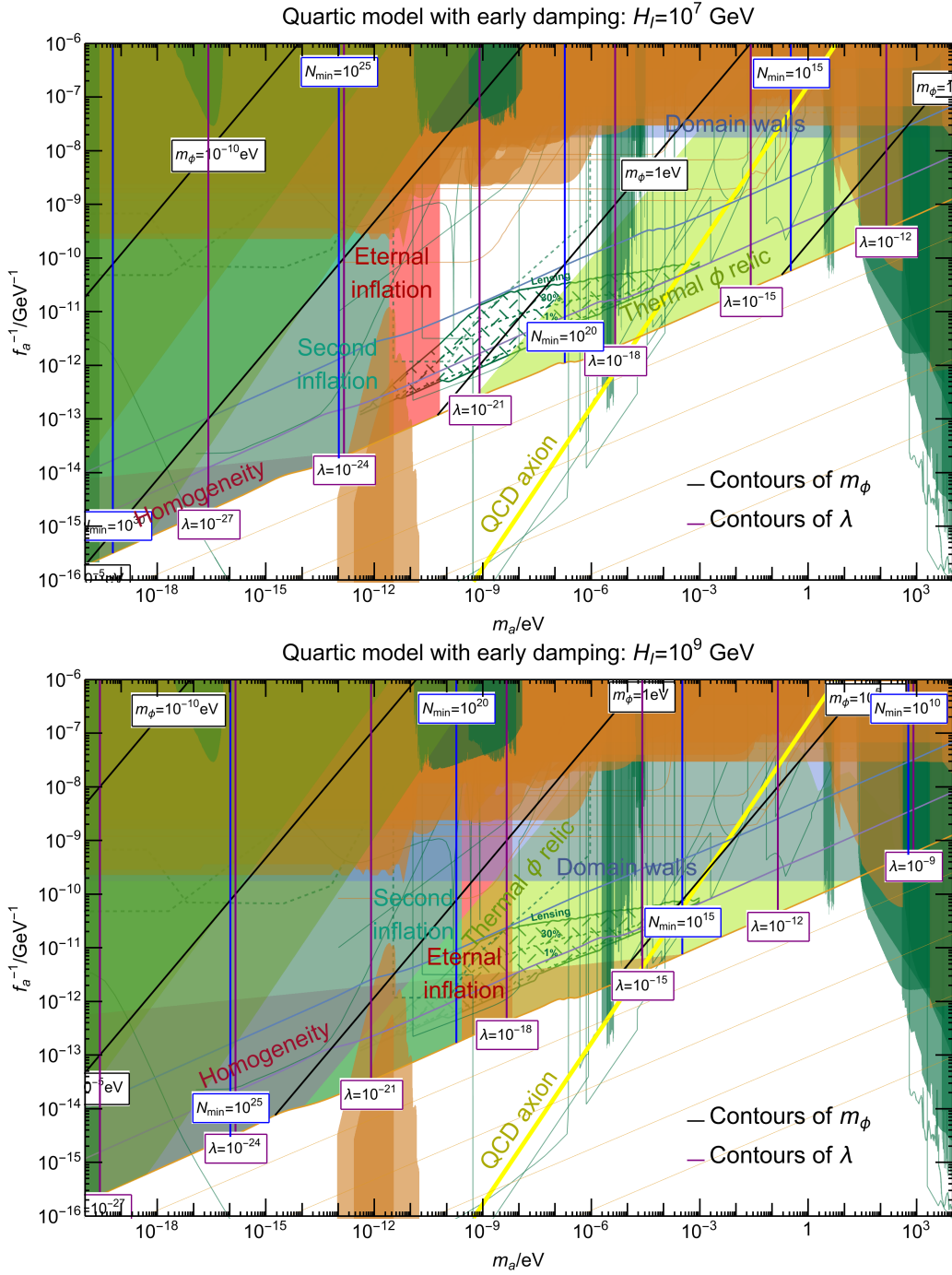


Figure 28. Same as figure 27, but for $H_I = 10^7$ GeV and 10^9 GeV. The constraints close the parameter space completely for values of $H_I \sim 10^{10}$ GeV and above.

14.4 General damping

We now generalize to the scenario in which damping takes place late enough to impact the yield.

Estimation of yield dilution for late damping: If damping of the radial oscillations takes place after ϕ reaches f_a , i.e. if $T_{\text{kin}} > T_{\text{damp}}$, then radial oscillations redshift as cold matter. If this evolution is not checked by radial damping, then ρ_ϕ will dominate the energy density after T_{dom} defined by

$$\frac{1}{2}(\lambda\phi_{\text{kick}})^2\phi_{\text{kick}}^2\left(\frac{a_{\text{kick}}}{a_{\text{kin}}}\right)^4\left(\frac{a_{\text{kin}}}{a_{\text{dom}}}\right)^3 = \frac{\pi^2}{30}g_{*s}(T_{\text{dom}})T_{\text{dom}}^4, \quad (14.30)$$

which implies

$$T_{\text{dom}} \approx \mathcal{O}(0.1) \times \frac{f_a}{\lambda^{1/4}} \left(\frac{H_I}{m_{\text{Pl}}}\right)^{3/2}, \quad (14.31)$$

where the exact value of the prefactor depends on the exact number of relativistic degrees of freedom. If damping takes place after T_{dom} , where ϕ dominates the energy density, then the yield dilution is significant. In the instantaneous approximation, damping is assumed to instantaneously transfer all energy stored in radial fluctuations to the SM plasma such that the plasma is reheated from T_{damp} to $T_{\text{damp,heated}}$, where $T_{\text{damp,heated}}$ is defined from

$$\frac{1}{2}(\lambda\phi_{\text{kick}})^2\phi_{\text{kick}}^2\left(\frac{a_{\text{kick}}}{a_{\text{kin}}}\right)^4\left(\frac{a_{\text{kin}}}{a_{\text{damp}}}\right)^3 = \frac{\pi^2}{30}g_{*s}(T_{\text{damp,heated}})T_{\text{damp,heated}}^4. \quad (14.32)$$

The dilution yield is then

$$Y_{\text{diluted}} \approx 0.9 \epsilon \lambda^{-17/16} \left(\frac{H_I}{m_{\text{Pl}}}\right)^{3/8} \left(\frac{T_{\text{damp}}}{f_a}\right)^{3/4} \quad (14.33)$$

Solution in the presence of yield dilution: If $T_{\text{dom}} > T_{\text{damp}}$ such that the yield is diluted from eq. (14.10) to eq. (14.33) then the observed DM yield is found for

$$m_{\phi,\text{today,DM,diluted}} \approx \mathcal{O}(1) \times \left(\frac{H_I}{m_{\text{Pl}}}\right)^{6/17} \left(\frac{m_a}{\epsilon T_{\text{eq}}}\right)^{16/17} \left(\frac{T_{\text{damp}}}{f_a}\right)^{12/17}, \quad (14.34)$$

which ensures $\rho_a = \rho_{\text{DM}}$ in the presence of yield dilution. As pointed out above, the corresponding quartic coupling is found from $\lambda \approx m_{\phi,\text{today}}/f_a$.

Impact and parameter space Because the ratio ρ_ϕ/ρ_r stays constant until ϕ has relaxed until f_a the phenomena of yield dilution plays a much smaller role in the quartic model. This can be seen in figure 29 where we give constraints on the $[T_{\text{damp}}, m_a]$ parameter space. Yield dilution takes place when T_{damp} is lower than the temperature where ϕ dominates the energy density, which corresponds to the region below the green dashed line in figure 29. As can be seen, only a small amount of parameter space is impacted by this dilution, and contrary to the nearly-quadratic model, the structure of the constraints is such that dilution does not open up any disconnected regions of parameter space.

Viable range of damping temperatures As in the nearly-quadratic model, the viable range of damping temperatures is ultimately limited by T_{kick} and T_* . Depending on the choice of parameters, the lower bound may be further restricted by the homogeneity condition or the condition that radiation dominates at the time of the kick. The structure of the constraints on T_{damp} is visualized in figure 29, where the individual constraints are shown for constant f_a and H_I . The viable range of T_{damp} is shown on the $[m_a, f_a]$ plane in figure 30 for an exemplary choice of H_I .

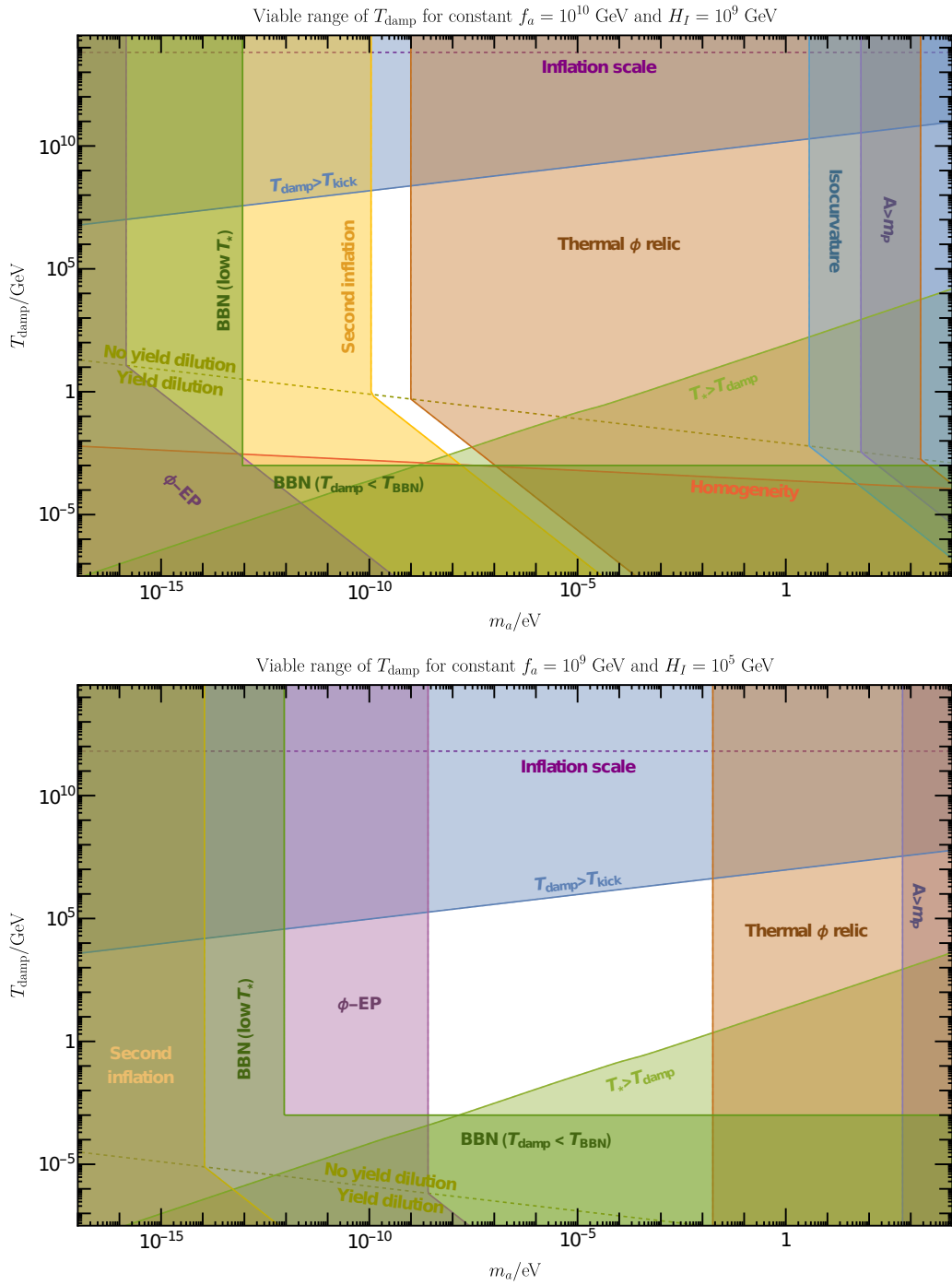


Figure 29. Constraints on the $[T_{\text{damp}}, m_a]$ parameter space for fixed f_a and H_I . The correct DM yield can be achieved in every unshaded point. The lower plot corresponds to a horizontal slice of fig. 30.

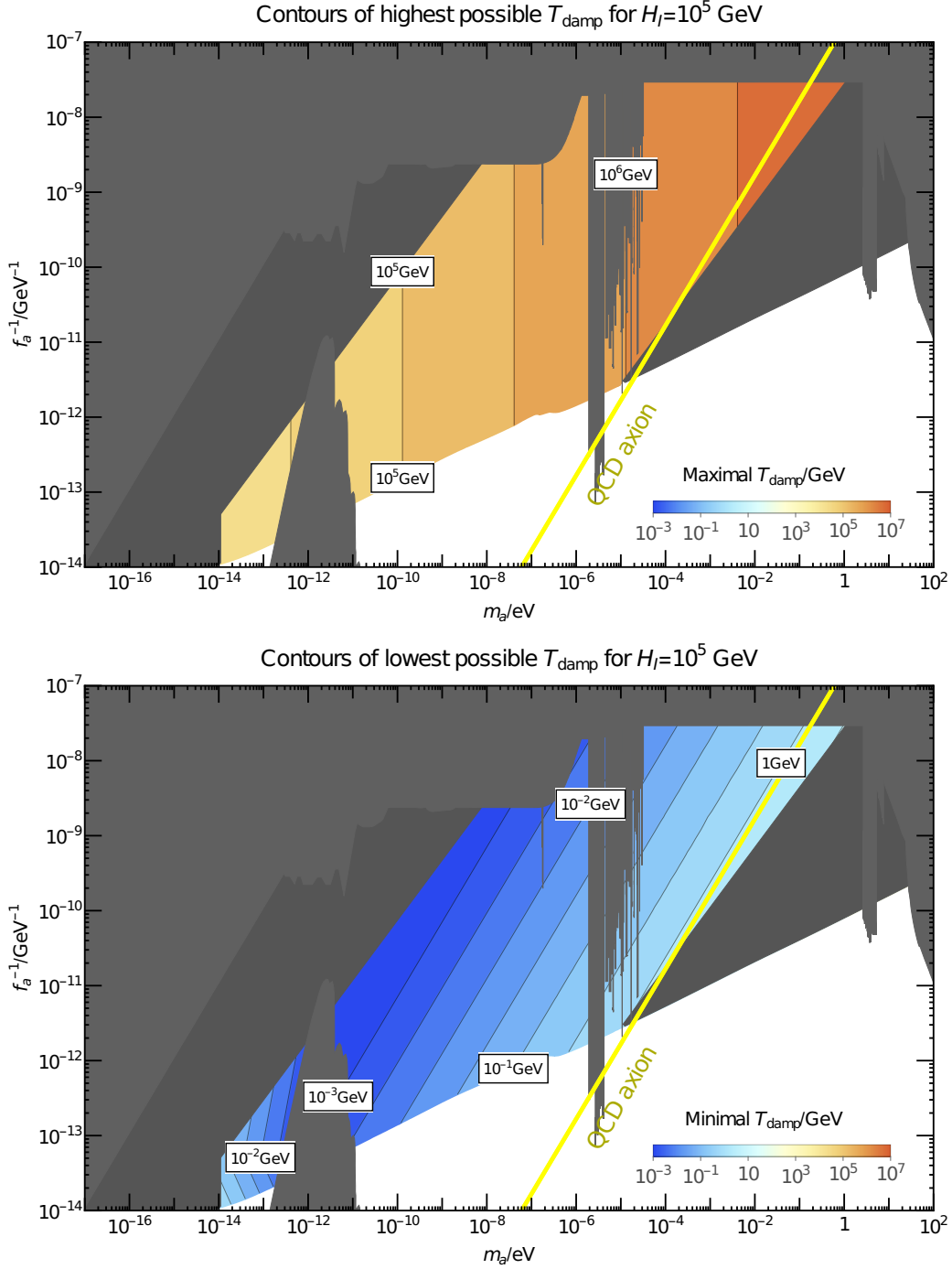


Figure 30. Viable range of damping temperatures in the quartic model for $H_I = 10^5$ GeV. Colored regions are viable, and gray regions are excluded for all possible damping temperatures. See figure 29 for a visualization of the constraints that shape these limits. **Top:** Maximally viable damping temperature. **Bottom:** Minimally viable damping temperature.

14.5 Obstacles to thermal damping

Without a solution to the domain wall problem, we do not find any viable solutions for models with quartic potentials. To understand this, consider the yield produced in this class of models, which is

$$Y_{\text{kick}} \approx 0.2 \times \frac{\epsilon}{\lambda^{5/4}} \left(\frac{H_I}{m_{\text{Pl}}} \right)^{3/2}, \quad (14.35)$$

where the prefactor has a mild dependence on $g_*(T_{\text{kick}})$. To realize solutions corresponding to sub-eV axions masses it is therefore necessary to consider either large values of H_I or low values of λ . Setting aside the concerns of naturalness for the moment, we can attempt to avoid domain wall problems by considering low values of λ . Specifically, we can consider values of λ which go below the range constrained by thermal ϕ relics, eq. (C.15), which corresponds to

$$\lambda < \mathcal{O}(1) \frac{T_{\text{eq}}}{f_a}. \quad (14.36)$$

However, this range of λ implies low radial masses. These low radial masses cannot compete with the thermal potential and avoid thermal suppression of the kick once a damping interaction is introduced. This holds true both for the Yukawa considered in section 13 and for Higgs portal interactions like $\mathcal{L} \supset \xi^2 \phi^2 H^\dagger H$. Such a Higgs portal interaction leads to a thermal potential of a similar form as eq. 13.4. If we demand $m_{\phi,0} > m_{\phi,\text{th}}$ at the time of the kick, then this corresponds to the condition

$$\xi \lesssim 2 \sqrt{\frac{H_I}{m_{\text{Pl}}}} \lambda^{1/4}, \quad (14.37)$$

where we wrote the Higgs-portal coupling ξ although the same condition applies for Yukawa interactions upon substitution $\xi \rightarrow y$. Such low values of the coupling constant do not permit damping. For instance, even with the most efficient Higgs portal interaction, [240]

$$\Gamma_\phi \approx 2 \frac{\xi^4 \phi^2}{\pi^2 \alpha T}, \quad (14.38)$$

taken at the most competitive conditions, $\phi = f_a$ and $T \sim m_{\text{H},0}$, efficient damping as defined by $\Gamma > H$ requires

$$H_I \gtrsim 10^{12} \text{ GeV} \times \left(\frac{10^8 \text{ GeV}}{f_a} \right), \quad (14.39)$$

which is in contradiction with the simple domain wall constraint eq. (14.20). If we consider instead solutions above the range constrained by thermal ϕ relics, i.e.

$$\mathcal{O}(1) \times \frac{f_a}{m_{\text{Pl}}} < \lambda, \quad (14.40)$$

then the large values λ can only correspond to scenarios of eV-scale axions with values of H_I that are also constrained by the domain wall problem.

We have implemented a numerical solution of the Boltzmann equations and we find no solutions without DW problems for either Yukawa-type interactions or for Higgs portal interactions. If a solution to the DW problem is assumed, then some parameter space may become viable. We do not peruse the possibility further here.

15 Summary of part III

We have re-examined Affleck-Dine-like implementations of axion kinetic misalignment. We have shown for the first time how the $[m_a, f_a]$ parameter space is impacted by constraints on the mechanism that starts the rotation, i.e. the *kick* mechanism. Given suitable parameters, we find that the kick mechanisms studied here can generate axion-like-particles which account for all observed dark matter in much of the experimentally unconstrained regions of $[m_a, f_a]$ space in which the standard misalignment mechanism underproduces axion-like-particles dark matter. We have mapped which subsets of this parameter space can be supported by implementations based on either nearly-quadratic or quartic potentials, and for models with nearly-quadratic potentials, we completed the implementation with a detailed study of the radial damping mechanism that is required for a consistent cosmology.

In this study, we directly calculated the evolution starting from the UV potential which improves on previous work where the dependence on UV parameters was at implicit in parameters such as the kick temperature T_{kick} and initial radial VEV ϕ_{kick} . Furthermore, we studied which values of the UV parameters can lead to consistent implementations of the kinetic misalignment mechanism and which constraints arise directly from these assumptions. These parameter values required for successful implementation are reported in appendix G.

The large initial values of the field required for a successful implementation of kinetic misalignment generally require radial masses much lighter than the conventional scenario $m_\phi \sim f_a$. The rich radial dynamics made possible by $m_\phi \ll f_a$ can provide the initial conditions of $\dot{\theta} \gg \Lambda_b$ required for kinetic misalignment, but this avenue also gives rise to several constraints associated with the radial mode ϕ . Firstly, oscillations in the radial mode can overproduce dark matter if they are not damped out. Secondly, if the ϕ field is brought into thermal contact with the SM plasma to dampen said oscillations, then the resulting thermal ϕ relic can be pathological for the cosmology. We find in eq. (11.14) that this relic constrains ϕ masses in the range to $\text{eV} \lesssim m_\phi < f_a^2/m_{\text{Pl}}$. Thirdly, the lower bound on the radial mass is ultimately set by tests of the equivalence principle which demands $m_\phi \gtrsim 10^{-5}$ eV, see eq. 11.16. Fourthly, the decay of radial oscillations may source a hot axion relic that can be problematic for the cosmology. The dynamics of kinetic misalignment also imply a constraint from the loss of homogeneity associated with a prolonged a^{-6} *kiniation-like* scaling of the energy in the radial mode. This constraint may be relaxed if it is assumed that the primordial spectrum of curvature fluctuation $A_s(k)$ is smaller on the relevant scales. With these general conditions in mind, we conclude the following:

For models with nearly-quadratic potentials, we conclude the following:

- In the absence of any form of dilution or ameliorating assumptions on $A_s(k)$, the combination of thermal ϕ relics and the homogeneity condition leave only implementations with sub-eV values of the radial mass. This can be summarized from figs. 16 to 17.
- Once the possibility of entropy dilution from damping of the radial mode is taken into account or the homogeneity bound is relaxed, solutions with radial masses above the thermal ϕ bound, i.e. $f_a^2/m_{\text{Pl}} < m_\phi$, become possible. We summarize the constraints in fig. 20 and map the compatible range of damping temperatures in 26.
- Once damping through a KSVZ-like Yukawa coupling is assumed, thermal effects rule out solutions with sub-eV radial masses because such low masses cannot compete with the thermal potential. This is discussed in section 13.3. Furthermore, higher damping temperatures are required than those allowed by the homogeneity condition. Therefore, $A_s(k)$ must be at least a few orders of magnitude lower than the value suggested by CMB measurements for thermal damping to work. The required degree of suppression is presented in 23.

- Given that assumptions on $A_s(k)$ are met, thermal damping supports ALP and QCD axion dark matter in the range 10^{-6} eV to 10^{-2} eV. This includes parameter space in reach of experiments such as MADMAX, ALPHA, BRASS, BREAD, IAXO, and ALPSII. The range of damping temperatures compatible with thermal damping is mapped out in fig. 26.
- If this $A_s(k)$ suppression does not interfere with the primordial spectrum of gravitational waves and the inflation scale E_I is near 10^{16} GeV, then these may be observable in future gravitational wave observatories. See figs. 23-25.

For the models with quartic potentials, we conclude the following:

- We map out the parameter space that can be supported in the absence of thermal effects in figs. 27 and 28. From those figures it is apparent that the combination of thermal ϕ relics and domain wall problems force extremely small values of the quartic coupling λ . If these problems are not addressed¹⁵ then only scenarios with $m_\phi \lesssim$ eV and $\lambda = m_\phi/f_a \lesssim 10^{-15}$ are possible¹⁶.
- In the absence of a solution to the domain wall problem we do not find implementations with thermal damping for models with quartic potentials. This problem can be traced back to $Y_{\text{kick}} \propto \lambda^{-5/4} (H_I/m_{\text{Pl}})^{3/2}$, which implies either very light masses in tension with thermal effects or large H_I in tension with the domain wall problem. See section 14.5 or discussion in appendix H.3.
- Yield dilution plays only a minor role for the models with quartic potentials. This is because the energy density of the PQ field does not grow relative to the radiation density for much of the cosmological evolution, such that the PQ field tends not to become dominant. This was discussed in section 14.4.
- The parameter A in the PQ violating potential has to be chosen by hand to a value that has no other justification than being necessary for the kick to be of $\epsilon \sim \mathcal{O}(1)$, see discussion in section 14.2. Furthermore, in [34] the authors find that even with an appropriate choice of A it may not be possible to achieve ϵ large enough to avoid parametric resonance in the radial mode, which is not taken into account here. This makes roughly circular kicks in models with quartic potentials somewhat unnatural.

In summary, the models based on nearly-quadratic potentials support a large range of parameter space, including the QCD axion, but these scenarios generally require suppression of $A_s(k_{\text{kin}})$. Such A_s suppression is observationally unconstrained because k_{kin} is many orders of magnitude larger than the comoving momentum scale probed by Planck. If $A_s(k_{\text{kin}})$ suppression can be implemented, QCD axion dark matter as well as general ALP dark matter in reach of experiments such as IAXO, ALPSII, and MADMAX can be realized by extending a KSVZ-type model with higher-dimensional terms and a light, nearly-quadratic radial potential.

Models based on quartic potentials appear less appealing. Not only do such models suffer from unnaturally small quartic couplings and unmotivated choices of A , but they also face difficulties with thermal damping.

¹⁵In [33] this problem is avoided by using parametric resonance to resolve the domain wall problem. However, in the absence of a more detailed analysis, we do not consider this solution to the domain wall problem.

¹⁶Note that definition of λ here differs from the convention used some other places. We follow Co et al. [27, 33, 34] in our definition. One may consider λ^2 the fundamental parameter.

Part IV

Conclusions and future perspectives

We have, in this thesis, investigated novel production mechanisms for axion dark matter. In particular, we have studied how the axion kinetic misalignment mechanism can account for axion dark matter in the regime that is most accessible to experimental observations.

In part II, we made an implementation-independent study and performed a detailed analytical treatment of both kinetic misalignment and fragmentation in such scenarios of rotating axion fields. We show that the homogenous mode of the axion field in most of the relevant parameter space completely fragments into fluctuations with a non-trivial momentum spectrum. Specifically, we found that the efficiency of fragmentation depends crucially on $m_a(T_*)/3H_*$, such that fragmentation becomes more efficient as trapping is delayed to temperatures further below the oscillation temperature expected by the standard misalignment mechanism. Because this temperature cannot be lowered indefinitely, the axion masses that can be realized as dark matter are ultimately limited from below by BBN. We found that the $[m_a, f_a]$ parameter space of kinetic misalignment can be classified according to the efficiency of fragmentation: Fragmentation either becomes efficient before trapping, becomes efficient after trapping, or remains sufficiently inefficient for kinetic misalignment to be realized with a predominantly homogeneous field.

The fragmentation process itself can generate a gravitational wave signal. We discussed this gravitational wave signal in section 9, where we concluded that the signal was hard to observe as they are of too low frequency and amplitude. In scenarios where the axion relic is relativistic for a significant period following fragmentation, this amplitude may be boosted enough that it becomes observable for CMB experiments such as Voyage or SuperPIXIE, but our analysis suggests that any such relativistic period may be too short for realistic observational prospects. The axion relic will be generated with a non-trivial momentum spectrum if fragmentation becomes efficient. This non-trivial momentum spectrum can lead to the formation of axion mini-clusters. Such axion mini-clusters, which were investigated by Eroncel and Servant [30], may provide a stronger signature of axion fragmentation within scenarios of kinetic misalignment. Furthermore, we identified implementation-independent constraints from structure formation and BBN. Crucially, we also pointed out that if the period of kination-like scaling of the axion field is too long, then primordial fluctuations in the field dominate the homogenous mode of the field, which leads to constraints on the implementations in part III.

In part III, we studied Affleck-Dine-like implementations of kinetic misalignment and mapped out precisely which regions the $[m_a, f_a]$ plane can be supported by various choices of UV parameters. This direct link between regions in the $[m_a, f_a]$ plane and the associated UV parameters had not been made in previous literature on KMM. We investigated in detail the two model implementations proposed by Co et al. [27] and, in particular, did a careful analysis of the mechanism required for damping of radial oscillations. This analysis found that much of the parameter space above $m_a \sim 10^{-6}$ eV can be realized in a nearly-quadratic model with higher dimensional terms of $n = 13$. This scenario relies on a Yukawa coupling such as the one inherent to KSVZ-type axion models. One key limitation of such a KSVZ-like nearly-quadratic scenario is that primordial fluctuations tend to ruin homogeneity long before kinetic misalignment or axion fragmentation of the type studied in part II can take place. To avoid this constraint, it is necessary to postulate that the spectrum of primordial fluctuations is suppressed relative to the value suggested by CMB measurements, i.e. $A_s(k_{\text{kin}}) \lesssim 10^{-4} A_{s,\text{Planck}} \sim 10^{-4} \times 2.1 \times 10^{-9}$. This postulate is compatible with observations because Planck [194] only constrains A_s around the pivot scale 0.05 Mpc^{-1} which is many orders magnitude of smaller than the relevant comoving momentum scale k_{kin} . The relevant

scale generally lies in the range $10^{14} \text{ Mpc}^{-1} \gtrsim k_{\text{kin}} \gtrsim 10^{10} \text{ Mpc}^{-1}$. If this suppression of $A_s(k_{\text{kin}})$ can be accommodated, then the KSVZ-like nearly-quadratic models provide UV completions with interesting phenomenology. In particular, some scenarios lead to kination. If the scale of inflation is near the maximal value allowed for by Planck measurements, i.e. not much lower than $E_I \sim 10^{16}$ GeV, then such kination can lead to an observable gravitational wave signal by amplifying the primordial GW spectrum [36, 37]. As the energy density of the PQ field may dominate the energy density in this family of scenarios and the damping rates depend intimately on the temperature of the SM plasma, these solutions involve a complex interplay of thermal effects between the SM sector and the new scalar sector. This interplay is accurately captured by our numerical solution of the Boltzmann equations.

We also studied models based on quartic potentials. However, we find that such models suffer from a number of problems including unnaturally small quartic couplings, somewhat less natural kick dynamics, and difficulties with thermal damping arising from a domain wall problem. If the domain wall problem can be resolved such models may be viable, but we did not pursue that possibility further here.

Regardless of which of these options is chosen as a UV completion, our analysis reaffirms that the initial conditions assumed by kinetic misalignment can be found. This is fantastic news for the many upcoming axion experiments since it motivates the possibility of axion dark matter in the regime of low values of the decay constant f_a , where direct experimental searches are most sensitive. It is particularly exciting that experiments such as ALPSII, which is about to start running at DESY, might be sensitive to axion dark matter. It is also quite remarkable that IAXO can probe the QCD axion as DM.

In view of the many exciting possibilities presented by kinetic misalignment, much future study is warranted. Interesting avenues for future work include the following:

- The analytic description of fragmentation which we developed in part II, breaks down for the very lowest values of f_a . To facilitate our later analysis, we made the working assumption that the relic from fragmentation in this regime was unchanged despite this. Future study is warranted to check this working assumption and evaluate whether there are large modifications of the abundance and the spectrum of the axion relic in this regime. This could impact predictions for, e.g., ALPSII, MADMAX, or IAXO.
- It would be interesting to identify other observational prospects from the non-trivial momentum spectrum, which we calculated in this work.
- The KSVZ-like nearly-quadratic model required suppression of A_s to be viable. This solicits a concrete realization of inflation that realizes this suppression. Furthermore, it is important to check if gravitational wave predictions are modified by this suppression.
- More studies of non-Affleck-Dine mechanisms such as the trapped misalignment mechanism can provide an alternative to the mechanisms studied above.
- An important constraint for the KSVZ-like model was that primordial fluctuations do not fragment the field prematurely. We here treated this condition as a constraint because it diverges from the misalignment scenario. However, the axions produced from primordial fluctuations in a rotating axion field may constitute a DM relic in its own right. This scenario is being investigated in upcoming work by Eroncel et al. [202].

To conclude, this thesis has once again demonstrated that axion-like-particles hold many interesting possibilities for non-thermally produced dark matter. We have shown that these possibilities extended beyond the standard assumption on the initial conditions of the axion field and that rich

dynamics can be found by investigating scenarios of non-trivial initial velocity. This is especially interesting given the many upcoming experiments that will probe deeper into the axion parameter space in the near future. In particular, we now know that axions can provide non-thermal dark matter candidates in nearly all of the experimentally allowed ALP parameter space. The conclusion further strengthens the science case for further research into the search for axion-like-particles.

A Table of subscripts, abbreviations and symbols

In this appendix, we provide a list of subscripts, abbreviations and symbols that are used throughout part II and part III of the thesis. As part I is a review, quantities only used there are left out to reduce the length of the list. We here use q as a placeholder for other quantities.

List of subscripts:

q_* : Quantity measured at the time where the rotation of the PQ field ends, i.e. the time of trapping. Around this time, the field may either fragment, begin oscillating or become frozen.

q_0 : Quantity measured at zero-temperature (i.e., not necessarily today).

q_{ap} : Quantity measured at the apoapsis (i.e., the top) of the orbit. Only used with $\dot{\theta}_{\text{ap}}$.

q_{damp} : Quantity measured at the time of damping.

q_{diluted} : Quantity which has been diluted by the entropy injection associated with radial damping.

q_{dom} : Quantity measured at the time when the energy density of the PQ field starts to dominate the energy density of the universe if such domination takes place.

q_{eq} : Quantity measured at the time of radiation-matter equality.

q_{heated} : Quantity which has been heated by the entropy injection associated with radial damping.

q_I : Quantity measured during inflation.

q_{kick} : Quantity measured at the time of the kick, i.e. at the onset of rotation of the PQ field.

q_{kin} : Quantity measured at the time where ϕ reaches the minimum at f_a , which corresponds to the onset of kination-like scaling of the axion.

q_{today} : Quantity measured today.

List of symbols:

ϱ : The ratio between the ALP kinetic energy density and its potential energy, see eq. (5.8).

α : Fine structure constant $\alpha \equiv g^2/4\pi$. Used for an unspecified gauge interaction in the context of damping mechanisms.

α_{EM} : The electroweak fine structure constant, $\alpha_{\text{EM}} \approx 1/137$.

γ : Symbol both used for the photon and for the parameter which quantifies the temperature dependence of the axion mass $m_a(T)$ (see eq. (2.8)).

Γ : Damping rate or interaction rate, usually specified by a subscript.

ϵ : Parameter which measures how elliptical the initial rotation corresponds. By construction, $0 \leq \epsilon \leq 1$ where $\epsilon = 0$ corresponds to perfectly radial oscillation and $\epsilon = 1$ corresponds to a perfectly circular rotation, see eq. (11.7).

θ : The angular mode of the PQ field P , also referred to as the axion field. Can refer to either a QCD axion or a general axion-like-particle. This field is assumed to be homogenous in part I and part III. In part II, θ is decomposed into a homogeneous mode Θ and fluctuations θ_k .

κ : Dimensionless momentum parameter defined by eq. (6.3).

λ : Quartic coupling, see eq. (14.1).

- μ_k : Floquet exponents of mode k , see eq.(6.2).
- $\tilde{\mu}$: Dimensionless Floquet exponents defined by eq. (6.12).
- ρ_q : Energy density of q . Note that ρ_a is the axion energy density, which in part II is split into the energy density of the homogenous field ρ_Θ and the energy density in fluctuations ρ_{fluct} .
- τ : Dimensionless time variable defined by eq. (6.12).
- ϕ : The radial mode of the PQ field P , also referred to as the saxion field, see eq. (11.1).
- Ω_q : Critical energy density of the universe of component q , i.e. ρ_q/ρ_{tot} , where ρ_{tot} is the total energy density of the universe.
- $\Omega_{\text{GW}}(f)$: Fraction of energy density in gravitational waves per logarithmic frequency, see eq. (9.5).
- a : The scale factor of the universe; also used for the axion field in Part I.
- A : Dimension-full scale of the PQ-violating higher-dimensional terms, see eq. (11.3).
- A_k : Redshift factor defined by eq. (6.22).
- $A_s(k)$: Amplitude of the primordial power spectrum measured on comoving momentum scale k .
- $A_{s,\text{Planck}}$: The value of $A_s(k)$ measured by Planck [241] at the comoving momentum scale $k = 0.05 \text{ Mpc}^{-1}$, which is referred to as the pivot scale. $A_{s,\text{Planck}} \approx 2.1 \times 10^{-9}$.
- c_q : Redshift parameter which indicates that quantity q redshifts as $q \propto a^{-c_q}$.
- f_a : Decay constant of the axion. Also the VEV of the radial mode / saxion.
- g : Gauge coupling, the interaction depends on context.
- g_* : Number of relativistic degrees of freedom in the SM plasma.
- g_{*s} : Number of effective degrees of freedom in entropy.
- h : Scaling factor for Hubble expansion rate, $h \approx 0.674$.
- H : Hubble parameter.
- m_ϕ : Mass of the saxion / radial mode. As we are concerned with dynamics far from the minimum, we define this as $m_\phi \equiv \sqrt{V'\phi^{-1}}$ so that the EOM takes the standard form.
- $m_{\phi,DM}$: Solution for the saxion / radial mode mass which ensures that the observed DM relic is produced by KMM.
- $m_{\phi,\text{eff}}, m_{\phi,0}, m_{\phi,\text{th}}$, **and** $m_{\phi,\text{ln}}$: The contributions to the saxion mass are specified when thermal effects are involved. These indicate the effective (total), zero-temperature, thermal, and thermal-log masses, respectively.
- $m_a(T)$: Mass of the axion / angular mode as measured at temperature T .
- m_a : Zero-temperature mass of the axion / radial mode.
- m_{Pl} : The reduced Planck mass, $m_{\text{Pl}} \approx 2.4 \times 10^{18} \text{ GeV}$.
- M : Suppression scale of higher dimensional terms. In this work, we usually assume $M = m_{\text{Pl}}$ unless otherwise stated.
- n : Dimension of the higher dimensional terms.

n_a : Axion number density.

n_ϕ : Saxion number density.

n_{PQ} : PQ-charge. The conserved charge related to the axion shift symmetry.

N_{eff} : Effective number of neutrinos. See appendix C.

N_k : The total amplification factor of a given mode, see eq. (6.11).

P : The full complex PQ field, which is parametrized as $P = \frac{1}{\sqrt{2}}\phi e^{i\theta}$.

P_θ : The power spectrum of initial fluctuations, see eq. (6.33).

y : Yukawa coupling. See, e.g., eq. (13.1).

Y : The yield variable $Y = n_{\text{PQ}}/s$, where s is the entropy density.

Y_{DM} : The observed DM yield defined from n_{DM}/s where n_{DM} is the number density of DM particles and s is the entropy density. It is conveniently parametrized as $Y_{\text{DM}} \approx 0.64T_{\text{eq}}/m_a$.

List of abbreviations:

ALP Axion-like-particle. In this work, the word *axion* is used synonymously with ALP.

BBN Big Bang Nucleosynthesis

CMB Cosmic Microwave Background

DW Domain wall

EOM Equation of motion

FLRW Friedmann–Lemaître–Robertson–Walker

GW Gravitational wave

KMM Kinetic misalignment mechanism

KSVZ Kim-Shifman-Vainshtein-Zakharov. Our benchmark QCD axion model, see section 2.2.

LMM Large Misalignment Mechanism. See discussion in section 8.3.

PQ Peccei-Quinn, as related to the PQ solution of the strong CP problem. Here used to refer to the PQ field P or the PQ charge n_{PQ} related to the axion shift symmetry. Note that as we are working with general ALP scenarios, the reference to Peccei and Quinn here is often by analogy only and need not imply a QCD axion scenario.

QCD Quantum Chromodynamics. Often used in conjunction with the term "QCD axion", which refers to the axion which solves the strong CP problem of QCD.

SM Standard model

SMM Standard misalignment mechanism

SUSY Super symmetry

WKB Wentzel–Kramers–Brillouin

B Detailed discussion of the parametric resonance

In this appendix, we discuss in detail the solutions of (5.6). If we neglect the expansion of the universe, then (5.6) becomes

$$\ddot{\theta}_k + \left(\frac{k^2}{a^2} + m^2 \cos \Theta \right) \theta_k = 0, \quad (\text{B.1})$$

where both the scale factor a and the axion mass m are constant. We define a dimensionless time $t_m \equiv mt$, and dimensionless momentum variable $\kappa \equiv k/am$ so that (B.1) takes a simpler form:

$$\theta''_{\kappa} + (\kappa^2 + \cos \Theta) \theta_{\kappa} = 0, \quad (\text{B.2})$$

where primes denote derivatives with respect to t_m . The solutions of this equation have been studied in [196] in the case of oscillations after trapping, i.e. $\varrho < 1$. We here review the method of getting these solutions. Furthermore, we generalize to obtain the solutions for the rolling axion, $\varrho > 1$.

We start by defining a new time variable z by

$$z(t_m) \equiv \cos \Theta(t_m), \quad (\text{B.3})$$

where Θ is the solution of the homogeneous mode. Without expansion, the energy density ϱ becomes a constant of motion. By taking the derivative of $z(t_m)$ by using the fact that ϱ is conserved, we obtain

$$\frac{dz}{dt_m} = \pm \sqrt{2(1-z^2)(2\varrho-1+z)}, \quad (\text{B.4})$$

where the sign depends on the sign of $\sin \Theta$ and Θ' . Plugging this result into (B.2) gives

$$2(1-z^2)(2\varrho-1+z)\theta''_{\kappa} + [1+2(1-2\varrho)z-3z^2]\theta'_{\kappa} + (\kappa^2+z)\theta_{\kappa} = 0, \quad (\text{B.5})$$

where primes are now denoting derivatives with respect to z . Note that z is not single-valued for the whole oscillation. Thus, this equation can describe the solutions only for a quarter of the period when $\varrho < 1$ and half of the period when $\varrho > 1$. Nevertheless, we can still use this equation to get the solutions in each patch, and then glue them together to get the full solution.

Let $\theta_{\pm}(z)$ denote the two linearly independent solutions to (B.5), where we have omitted the momentum subscript for cleaner notation. It turns out that the linear combinations θ_+^2 , θ_-^2 , and $\theta_+\theta_-$ obey a third order equation:

$$2(z^2-1)(2\varrho-1+z)M''' + [9z^2-6(1-2\varrho)z-3]M'' + 2(z-1+2\varrho-2\kappa^2)M' - 2M = 0. \quad (\text{B.6})$$

This equation has a polynomial solution given by

$$M(z) = z - 1 + 2\varrho - 2\kappa^2. \quad (\text{B.7})$$

The original equation we are trying to solve (B.2) is a *Hill differential equation*, so according to *Floquet's theorem*, the solutions must be of the form

$$\theta_{\pm}(t_m; \kappa) = \psi_{\pm}(t_m; \kappa) e^{\pm i\mu_{\kappa} t_m}. \quad (\text{B.8})$$

Therefore the polynomial solution (B.7) should correspond to $\theta_+\theta_-$ which implies

$$\theta_+(z)\theta_-(z) = \mathcal{N}^2(z-1+2\varrho-2\kappa^2), \quad (\text{B.9})$$

where \mathcal{N} is a normalization constant. To obtain the individual solutions, we note that the Wronskian of the system $W \equiv \theta'_+\theta_- - \theta_+\theta'_-$ obeys

$$W'(z) = -\frac{1+2(1-2\varrho)z-3z^2}{2(1-z^2)(2\varrho-1+z)}W(z) = -\frac{d}{dz} \ln \sqrt{(1-z^2)(2\varrho-1+z)}W(z). \quad (\text{B.10})$$

This can easily be solved as

$$W(z) = \theta'_+(z)\theta_-(z) - \theta_+(z)\theta'_-(z) = \frac{c_\kappa \mathcal{N}^2}{(1-z^2)(2\varrho-1+z)}, \quad (\text{B.11})$$

where c_κ is an integration constant coming from the integration of (B.10) which we will determine shortly. By combining (B.8) and (B.11) we obtain the following differential equations for θ_\pm :

$$2M(z)\frac{d\ln\theta_\pm}{dz} = M'(z) \pm \frac{c_\kappa}{\sqrt{(1-z^2)(2\varrho-1+z)}}. \quad (\text{B.12})$$

The solutions are¹⁷

$$\theta_\pm(z) = \mathcal{N}\sqrt{|M(z)|} \exp\left(\pm \frac{c_\kappa}{2} \int^z \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}\right). \quad (\text{B.13})$$

The integration constant c_κ can be obtained by plugging this solution into (B.5). The result is

$$c_\kappa^2 = 8\kappa^2(\varrho - \kappa^2)(1 - \varrho + \kappa^2). \quad (\text{B.14})$$

These coefficients determine the instability bands in the parametric resonance. If $c_\kappa^2 > 0$, the Floquet exponent will be real, and parametric resonance happens. Otherwise, they will be imaginary, and mode functions will only have oscillatory solutions. Therefore by using (B.14) we can directly find the modes which are inside the instability bands. They are given by

$$\varrho - 1 < \kappa^2 < \varrho, \quad \text{for } \varrho > 1, \quad (\text{B.15})$$

$$0 < \kappa^2 < \varrho, \quad \text{for } \varrho < 1. \quad (\text{B.16})$$

To find the value of the Floquet exponent, we need to do a little bit more work. We work out the cases before and after trapping separately.

Before trapping: Without loss of generality, we can assume that the homogeneous mode travels from $\Theta = 0$ to $\Theta = 2\pi$ during one period. In the first half of the period, Θ moves from 0 to π , and z decreases from $z = 1$ to $z = -1$. In this patch, the exponentially growing solution is θ_+ . By choosing the normalization factor \mathcal{N} such that the solution is unity initially, its value after half an oscillation is given by

$$\theta_\kappa^{(1/2)} = \sqrt{\frac{|M(-1)|}{|M(1)|}} \exp\left(\frac{c_\kappa}{2} \int_1^{-1} \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}\right) \quad (\text{B.17})$$

In the second half of the period, Θ moves from π to 2π while z increases from $z = -1$ to $z = 1$. Now the exponentially growing solution is θ_- . In order to glue the solutions, the normalization \mathcal{N} should be chosen such that the full solution is continuous. Then we obtain the solution after a full oscillation as

$$\begin{aligned} \theta_\kappa^{(1)} &= \theta_\kappa^{(1/2)} \sqrt{\frac{|M(1)|}{|M(-1)|}} \exp\left(-\frac{c_\kappa}{2} \int_{-1}^1 \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}\right) \\ &= \exp\left(c_\kappa \int_1^{-1} \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}\right). \end{aligned} \quad (\text{B.18})$$

¹⁷Note that while $(1-z^2)(2\varrho-1+z)$ is always non-negative, $M(z)$ does cross zero at $z = 1 - 2\varrho + 2\kappa^2$. If this point lies in the range of the integral, then the integral is understood as its Cauchy principal value.

The Floquet exponent μ_κ can be obtained by matching this solution to (B.8) and using the fact that ψ_\pm 's are periodic functions. We then obtain

$$\mu_\kappa^{(\varrho>1)} = \frac{c_\kappa}{\mathcal{T}_>} \int_1^{-1} \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}}, \quad (\text{B.19})$$

where $\mathcal{T}_>$ is the period of oscillation of the homogeneous mode in physical time t when $\varrho > 1$. This can be derived from the conservation of energy as

$$\mathcal{T}_> = \frac{2}{m\sqrt{\varrho}} \text{K}(1/\sqrt{\varrho}). \quad (\text{B.20})$$

The integral in (B.19) can be expressed in a more useful form by changing the integral path using Cauchy's integral theorem:

$$- \int_{-\infty}^{1-2\varrho} \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}}. \quad (\text{B.21})$$

By defining ϑ such that $z = -1/\sin^2 \vartheta$ we obtain our final result.

$$\mu_\kappa^{(\varrho>1)} = \frac{c_\kappa}{\mathcal{T}_>} \int_0^{\arcsin(1/\sqrt{2\varrho-1})} \frac{d\vartheta}{1 + (1-2\varrho+2\kappa^2)\sin^2 \vartheta} \frac{2\sin^2 \vartheta}{\sqrt{(1+\sin^2 \vartheta)[1+(1-2\varrho)\sin^2 \vartheta]}}}. \quad (\text{B.22})$$

After trapping: Here we start the oscillation at the minimum $\Theta = 0$. In the first half of the oscillation Θ travels from 0 to $2\arcsin(\sqrt{\varrho})$ while z decreases from $z = 1$ to $z = 1-2\varrho$. The growing solution is the θ_+ solution. Again normalizing the mode functions to unity initial amplitude, we find the solution after the first half of the oscillation as

$$\theta_\kappa^{(1/2)} = \sqrt{\frac{|M(1-2\varrho)|}{|M(1)|}} \exp\left(\frac{c_\kappa}{2} \int_1^{1-2\varrho} \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}}\right). \quad (\text{B.23})$$

In the second half of the oscillation Θ travels from $2\arcsin(\sqrt{\varrho})$ back to 0 while z increases from $z = 1-2\varrho$ to 1. Now the growing solution is θ_- . So after a full period, the solution is

$$\theta_\kappa = \exp\left(c_\kappa \int_1^{1-2\varrho} \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}}\right). \quad (\text{B.24})$$

The Floquet exponent can directly be read from this result as

$$\mu_\kappa^{(\varrho<1)} = \frac{c_\kappa}{\mathcal{T}_<} \int_1^{1-2\varrho} \frac{dz'}{M(z')\sqrt{(1-z'^2)(2\varrho-1+z')}}}, \quad (\text{B.25})$$

where the period $\mathcal{T}_<$ in this case is

$$\mathcal{T}_< = \frac{2}{m} \text{K}(\sqrt{\varrho}). \quad (\text{B.26})$$

Applying the integral transformations that we did in deriving (B.22) yields the result

$$\mu_\kappa^{(\varrho<1)} = \frac{c_\kappa}{\mathcal{T}_<} \int_0^{\pi/2} \frac{d\vartheta}{1 + (1-2\varrho+2\kappa^2)\sin^2 \vartheta} \frac{2\sin^2 \vartheta}{\sqrt{(1+\sin^2 \vartheta)[1+(1-2\varrho)\sin^2 \vartheta]}}}. \quad (\text{B.27})$$

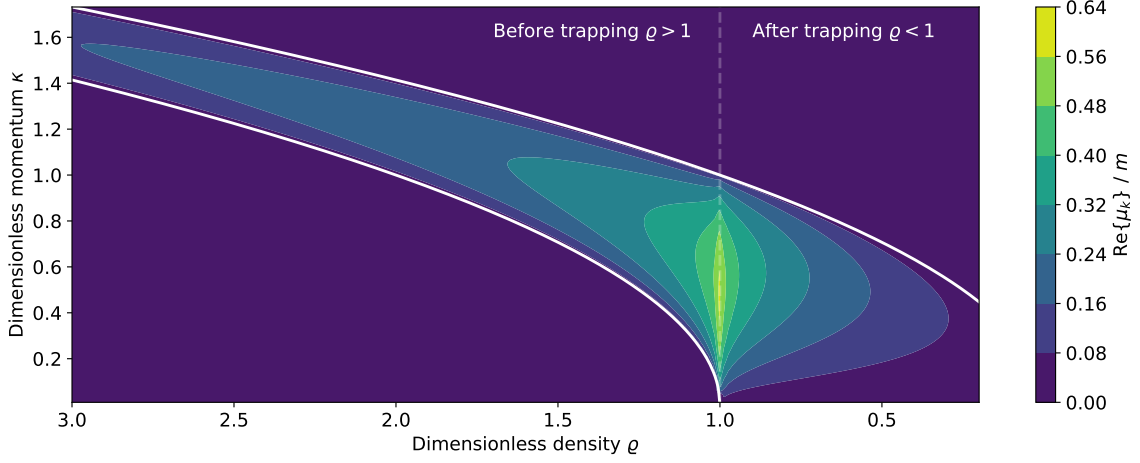


Figure 31. Floquet exponents as a function of the dimensionless energy density ϱ , and the dimensionless momentum κ , using the analytical result ((B.28)). The white lines denote the boundaries of the instability bands as given in ((B.15)) and ((B.16)). We see that the parametric resonance is most efficient around trapping $\varrho \approx 1$, and for the modes $\kappa \sim 0.5$.

Final result: Our final result for the Floquet exponents can be summarized as

$$\mu_k = \sqrt{8\kappa^2(\varrho - \kappa)(1 - \varrho + \kappa^2)} \times \begin{cases} \mathcal{T}_>^{-1}(\varrho) \mathcal{I}(\arcsin(1/\sqrt{2\varrho - 1})), & \varrho > 1 \\ \mathcal{T}_<^{-1}(\varrho) \mathcal{I}(\pi/2), & \varrho < 1 \end{cases}, \quad (\text{B.28})$$

where

$$\mathcal{I}(\varphi) = \int_0^\varphi \frac{d\vartheta'}{1 + (1 - 2\varrho + 2\kappa^2) \sin^2 \vartheta'} \frac{2 \sin^2 \vartheta'}{\sqrt{(1 + \sin^2 \vartheta') [1 + (1 - 2\varrho) \sin^2 \vartheta']}}. \quad (\text{B.29})$$

We show a plot of the Floquet exponents together with the instability bands in figure 31. The boundaries of the instability bands are shown in white lines. We can observe that the parametric resonance is most efficient around trapping $\varrho \approx 1$, and for the modes $\kappa \sim 0.5$.

C Thermal relics

Generally, the orbit produced by the kick mechanisms considered in this thesis will be elliptic rather than perfectly circular. Therefore, there will be oscillations of the radial mode in addition to the rotation of the angular mode. These oscillations will interfere with the cosmology if they are not damped by some interaction. Any such interaction, parametrized by a total interaction rate Γ_{damp} , will have to be efficient in order to damp the radial oscillations, i.e. it must satisfy

$$\Gamma_{\text{damp}} > H. \quad (\text{C.1})$$

An inevitable consequence of such an interaction is that the radial mode is brought into thermal equilibrium and that a thermal population of saxions will be generated. This population may itself be problematic, wherefore we in this appendix explore the phenomenology of the thermal relic.

The thermal population of ϕ will be generated as long as ϕ is light compared to T . The population will have a number density of

$$n_\phi = \frac{\zeta(3)}{\pi^2} T^3, \quad (\text{C.2})$$

which is simply the relativistic limit predicted by equilibrium thermodynamics. Here ζ is the Riemann-Zeta function. If subsequently the interactions freeze out, i.e. if $\Gamma_{\text{damp}} < H$, then n_ϕ becomes conserved such that $n_\phi \propto a^{-3}$ after this time. Therefore, assuming entropy conservation and that the interactions remain inefficient, the corresponding yield,

$$Y_{\phi,\text{thermal}} = \frac{n_\phi(T_{\text{fo}})}{s(T_{\text{fo}})}, \quad (\text{C.3})$$

becomes constant in time. The freeze-out temperature T_{fo} is the last temperature for which $\Gamma_{\text{damp}} \gtrsim H$. Interestingly, up to changes in g_{*s} , the yield becomes independent of T_{fo} ,

$$Y_{\phi,\text{thermal}} = \frac{n_\phi}{s} = \frac{\zeta(3)}{\pi^2} \frac{45}{2\pi^2} \frac{1}{g_{*s}(T_{\text{fo}})}, \quad (\text{C.4})$$

which implies that the late-time energy density is not sensitive to temperature T_{fo} except for g_{*s} dependence. This relic density will apply at sufficiently late times when the relic is cold.

While the relic is still hot, its energy density will redshift as hot matter, i.e. a^{-4} . Thus, the energy density of the thermal relic will remain a constant fraction of SM plasma as long as no reheating of the latter takes place. When reheating within the SM plasma does take place, this will slow down the redshift of the SM plasma and therefore dilute the thermal relic. Specifically, while the decoupled relic is still hot, its energy density can be described by

$$\rho_{\phi,\text{thermal,hot}}(T) \approx \frac{1}{g_*(T)} \left(\frac{g_{*s}(T)}{g_{*s}(T_{\text{fo}})} \right)^{4/3} \rho_r(T), \quad (\text{C.5})$$

so that at T_{fo} the relic is $\rho_{\phi,\text{thermal}} \sim \rho_r/g_*$.

Cosmological expansion cools the relic so that it will inevitably become non-relativistic. This transition takes place when the saxions reach a temperature of m_ϕ . As the saxions are decoupled from the thermal bath, this does not take place exactly at $T \sim m_\phi$, where T is the SM thermal bath temperature. Rather, the transition takes place at

$$T_{\text{nr}} \approx \left(\frac{g_{*s}(T_{\text{fo}})}{g_{*s}(T_{\text{nr}})} \right)^{1/3} m_\phi. \quad (\text{C.6})$$

After T_{nr} , the relic density will instead be described by

$$\rho_{\phi,\text{thermal,cold}}(T) = m_\phi Y_{\phi,\text{thermal}} s(T) = m_\phi \frac{\zeta(3)}{\pi^2} \frac{g_{*s}(T)}{g_{*s}(T_{\text{fo}})} T^3. \quad (\text{C.7})$$

Approximately, the thermal relic density can be expressed as

$$\rho_{\phi,\text{thermal}}(T) \approx \max(\rho_{\phi,\text{thermal,hot}}(T), \rho_{\phi,\text{thermal,cold}}(T)) \quad \text{for } T \leq T_{\text{fo}}, \quad (\text{C.8})$$

which applies as long as the relic remains decoupled from thermal equilibrium. Note again that this relic is present whenever $m_\phi < T_{\text{fo}}$ regardless of the mechanism through which equilibrium was held and of which temperature the interaction froze out (up to changes in g_{*s}).

The thermal ϕ relic does not present problems for BBN or CMB as long as the relic decouples sufficiently early, $T_{\text{fo}} \gtrsim 200$ MeV, and it remains in the hot phase. However, the relic can become problematic if it experiences a cold phase since the relic density would grow relative to the SM plasma density in such a cold phase.

Saxion to axion decay: The radial and angular modes, i.e. the saxions and the axions, are coupled. Therefore, a cold ϕ relic may be avoided if the saxions decay into axions, which remain relativistic much longer. At the very least, the Lagrangian must contain the kinetic term

$$|\partial_\mu P|^2 = \frac{1}{2}(\partial_\mu S)^2 + \frac{1}{2}(\partial_\mu \varphi)^2 + \frac{1}{2} \frac{S^2}{f_a^2} (\partial_\mu \varphi)^2 - \frac{S}{f_a} (\partial_\mu \varphi)^2, \quad (\text{C.9})$$

where $S = \phi - f_a$ and $\varphi = \theta f_a$ are the saxion and axion fields respectively. The latter of these terms corresponds to an $S \rightarrow \varphi\varphi$ decay which, following dimensional arguments, takes place with a rate of

$$\Gamma_{S\varphi} \sim \frac{1}{64\pi} \frac{m_\phi^3}{f_a^2}. \quad (\text{C.10})$$

Although the saxion does also experience EM couplings, this rate is, in general, more efficient than EM decay. Therefore, the saxion will decay into axions if the damping interactions do not dominate this decay. We denote the temperature for which $\Gamma_{S\varphi} \approx H$ as $T_{S\varphi}$.

If the saxions do decay into axions, then the relic will end up as dark radiation at BBN. This relic density might be boosted relative to eq. (C.5) if the saxion relic underwent a cold period prior to the decay. Therefore, the relic density is

$$\rho_{\phi/a,\text{thermal}} \approx \xi(T) \times \frac{1}{g_*(T)} \left(\frac{g_{*s}(T)}{g_{*s}(T_{\text{fo}})} \right)^{4/3} \rho_r(T), \quad (\text{C.11})$$

where $\xi(T)$ accounts for enhancement during cold saxion evolution and is

$$\xi(T) = \begin{cases} 1 & \text{if } T > T_{\text{nr}} \text{ or } T_{S\varphi} > T_{\text{nr}}, \\ \frac{g_{*s}^{1/3}(T_{\text{nr}})}{g_{*s}^{1/3}(T)} \frac{T_{\text{nr}}}{T} & \text{if } T_{\text{nr}} > T > T_{S\varphi}, \\ \frac{g_{*s}^{1/3}(T_{\text{nr}})}{g_{*s}^{1/3}(T_{S\varphi})} \frac{T_{\text{nr}}}{T_{S\varphi}} & \text{if } T_{\text{nr}} > T_{S\varphi} > T. \end{cases} \quad (\text{C.12})$$

N_{eff} constraints: If the thermal relic is too large at late times, then the relic will disturb BBN and CMB. Constraints on such hot relics are usually parametrized in terms of change in the effective number of neutrinos, i.e. ΔN_{eff} . Assuming that the relic is not in a cold phase $\rho_{\phi/a,\text{thermal}}$ can be expressed in terms a change in N_{eff} as

$$\Delta N_{\text{eff}} = \frac{g_*}{2} \frac{8}{7} \left(\frac{11}{4} \right)^{4/3} \frac{\rho_{\phi/a,\text{thermal}}}{\rho_r} \quad (\text{C.13})$$

For the scenario where the relic undergoes a cold period ended by axion decay prior to BBN, i.e. for $T_{\text{nr}} > T_{S\varphi} > T_{\text{BBN}}$, this constraint is

$$\Delta N_{\text{eff}}(T) \approx 2 \frac{f_a}{\sqrt{m_{\text{Pl}} m_\phi}} \frac{g_{*s}^{4/3}(T)}{g_{*s}(T_{\text{fo}}) g_*^{1/12}(T_{S\varphi})}. \quad (\text{C.14})$$

The $\Delta N_{\text{eff}} \propto m_\phi^{-1/2}$ dependence arises because large m_ϕ enhances $\Gamma_{\phi a}$. This more efficient $\phi \rightarrow aa$ decay shortens the cold period, which would otherwise boost the relic density relative to ρ_r .

The change in N_{eff} is constrained directly at BBN to $\Delta N_{\text{eff}} \lesssim 0.4$ [237]. The change in the effective number of neutrinos is also constrained from CMB to $\Delta N_{\text{eff}} \lesssim 0.3$ [206]. For simplicity, we here demand that ΔN_{eff} does not exceed 0.3 at either T_{BBN} or at around the time of CMB, i.e. $T \sim \text{eV} \sim T_{\text{eq}}$.

If the saxion is very heavy, then the $\phi \rightarrow aa$ decay rate is efficient at the temperature where the saxion becomes non-relativistic. Therefore, the thermal relic density will not be enhanced by the cold period, and N_{eff} constraints are avoided. Conversely, if the saxion is very light, then the thermal relic will not enter a cold period until after T_{eq} , in which case N_{eff} constraints can also be avoided. Therefore, $\Delta N_{\text{eff}} > 0.3$ excludes a range of m_ϕ corresponding to

$$\mathcal{O}(1) \times T_{\text{eq}} < m_\phi < \mathcal{O}(1) \times \frac{f_a^2}{m_{\text{Pl}}} \quad (\text{excluded by } \Delta N_{\text{eff}} > 0.3). \quad (\text{C.15})$$

The exact value of the $\mathcal{O}(1)$ factors depends on the exact number of relativistic degrees of freedom. The constraint applies only if the thermal relic freezes out at T_{fo} such that $T_{\text{damp}} > m_\phi$.

If the thermal relic satisfies ΔN_{eff} constraints at the time of CMB, then it must also underproduce dark matter. Therefore, constraints from DM overproduction do not need to be imposed in addition to ΔN_{eff} constraints.

Avoidance of thermal relic constraints by SM decay: The thermal relics constraints can be avoided if the saxions are sufficiently unstable to decay into the SM thermal bath before BBN. This happens if the damping interaction remains efficient until $T \sim m_\phi$, i.e. the constraint is avoided for $T_{\text{fo}} < m_\phi$. Alternatively, if other interactions become efficient after freeze-out of the damping interaction, then such may mediate $\phi \rightarrow \text{SM}$ decay. In order to avoid N_{eff} and DM overproduction constraints such interactions have to become efficient before T_{BBN} and must dominate $S \rightarrow \phi\phi$ decay. Note that saxion-to-axion decay dominates saxion EM decay so that EM decay cannot destroy the saxion thermal relic unless the axion is also sufficiently heavy to decay before BBN.

D Scalar field dynamics in a thermal environment

The following review is based largely on work by Mukaida and Nakayama [238], which was also extensively applied in, e.g., [37]. We consider the scenario of the Yukawa coupling given by eq. (13.1) and discuss the impact of the thermal environment on both the saxion ϕ , the fermion χ and the gauge bosons γ . Note that here γ denotes the gauge bosons of the χ -SM gauge interaction, such that γ signifies the gluon if the χ -fermions are KSVZ-fermions.

Thermal and effective masses: The mediating fermion χ receives mass contributions from both its coupling to ϕ and its coupling to the SM plasma. The contribution from ϕ , which we denote $m_{\chi,\phi}$, is generated by the VEV of ϕ and is therefore simply $m_{\chi,\phi} \sim y\phi$ in our notation. We denote the thermal mass arising from the gauge coupling to the SM plasma as $m_{\chi,th}$. Under the assumption that the dynamics are sufficiently non-violent that the fermions remain in the thermal distribution, this contribution was found to be $m_{\chi,th} \sim yT$ by Mukaida and Nakayama [238]. We assume that χ is dominated by these mass contributions such that the complete mass-squared is

$$m_{\chi,tot}^2 = m_{\chi,\phi}^2 + m_{\chi,th}^2 = y^2\phi^2 + g^2T^2. \quad (\text{D.1})$$

Since we assume $g < 1$, we will always have $m_{\chi,th} \sim gT < T$ and the fermions will therefore always be relativistic if dominated by the thermal mass. Therefore,

$$\chi \text{ are relativistic} \implies m_{\chi,\phi} \sim y\phi < T. \quad (\text{D.2})$$

If this condition is not satisfied, then the χ 's become Boltzmann suppressed and decouple from the plasma. This will be important for the scattering processes described below.

Since the gauge boson γ also has the gauge coupling g , it receives thermal mass-contribution similar to that of χ . We assume that this is the only mass contribution to γ such that the total γ mass is identical to the χ thermal mass:

$$m_\gamma \sim m_{\chi,th} \sim gT. \quad (\text{D.3})$$

Again, since we, by assumption, always have $gT < T$, the gauge bosons will always be relativistic.

The radial mode ϕ also receives thermal contributions, which we here denote $m_{\phi,th}$. We denote the zero-temperature mass $m_{\phi,0}$. Because ϕ only couples to the plasma through χ , the thermal contributions depend on whether or not these χ -fermions are relativistic or not. ϕ acquires the usual Yukawa thermal mass if the χ -fermions are relativistic and participating in the SM plasma:

$$m_{\phi,th} \approx yT \quad \text{if} \quad T > y\phi \sim m_{\chi,\phi}. \quad (\text{D.4})$$

However, if the χ -fermions are Boltzmann suppressed, i.e. if $T < y\phi$, then ϕ cannot interact with the thermal plasma at tree-level, and the above contribution will therefore be absent. Because the SM plasma still depends on ϕ at loop order the ϕ field will still receive a so-called thermal-log mass, which we here denote $m_{\phi,ln}$. By integrating out the heavy χ fields, this loop-induced contribution is estimated in [238] to be

$$m_{\phi,ln} \sim \alpha \frac{T^2}{\langle \phi \rangle} \quad \text{if} \quad \frac{m_{\phi,0}}{\alpha} < T < y\phi \quad (\text{D.5})$$

The calculation of $m_{\phi,ln}$ in [238] assumed that the dynamics of ϕ are slow compared to the equilibration time of the plasma, which is of order $\sim \alpha T$, which limits the range of validity to the above. The scenario for which $T < \frac{m_{\phi,0}}{\alpha} < y\phi$ is not investigated in [238] and we will here simply extrapolate the range of validity such that

$$m_{\phi,tot}^2 \approx \begin{cases} m_{\phi,0}^2 + y^2T^2 & \text{if } T > y\phi, \\ m_{\phi,0}^2 + \alpha \frac{T^2}{\langle \phi \rangle} & \text{if } T < y\phi. \end{cases} \quad (\text{D.6})$$

Mukaida and Nakayama [238] note that oscillons might be formed if the saxion oscillates with the thermal-log mass. Because we later require the field to be homogeneous, we cannot allow such oscillons to form, wherefore we must require that $m_{\phi,0} > m_{\phi,ln}$ if the thermal log mass is generated. We, therefore, never consider the case where $m_{\phi,ln}$ is dominant.

Perturbative decay: If the saxion mass exceeds that of either the gauge boson or the χ fermions, then it will be possible for the saxion oscillations to be damped via perturbative decay. Note that since both the gauge bosons and the χ receive thermal mass contributions of order gT , and we assume throughout that $y < g$ we can conclude that perturbative decay is never kinematically allowed if the saxion is dominated by its thermal mass, i.e.

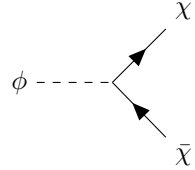
$$m_{\phi,tot} < m_{\chi,tot} \quad \text{and} \quad m_{\phi,tot} < m_{\gamma,tot} \quad \text{if} \quad m_{\phi,0} < yT \sim m_{\phi,th}. \quad (\text{D.7})$$

If ϕ is dominated by the zero-temperature mass, perturbative decay can be kinematically allowed if $m_{\phi,0}$ is larger than the $m_{\chi,tot}$ or $m_{\gamma,tot}$. This leads to the following decay conditions:

$$\chi\text{-decay:} \quad m_{\phi,0}^2 > y^2\phi^2 + g^2T^2, \quad (\text{D.8})$$

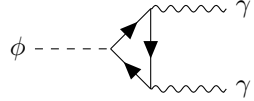
$$\gamma\text{-decay:} \quad m_{\phi,0}^2 > g^2T^2. \quad (\text{D.9})$$

Perturbative χ -decay takes place directly through the Yukawa vertex:



$$\Rightarrow \quad \Gamma_{\phi\bar{\chi}\chi} = \frac{y^2 m_{\phi,0}}{8\pi} \left(1 - 4\frac{m_{\chi}}{m_{\phi}}\right)^{3/2} \approx \frac{y^2 m_{\phi,0}}{8\pi}. \quad (\text{D.10})$$

Perturbative γ -decay takes place through the χ -1-loop. If the χ is heavy, i.e. if $y\phi > gT$ such that χ 's become heavy compared to the temperature of the plasma, then the diagram can be evaluated by integrating out the χ -loop:



$$\Rightarrow \quad \Gamma_{\phi\gamma\gamma} \sim b \left(\frac{yg^2}{m_{\chi,\phi}}\right)^2 m_{\phi,0}^3 \sim b\alpha^2 \frac{m_{\phi,0}^3}{\phi^2}, \quad (\text{D.11})$$

where

$$b = \frac{T(r) (12\pi)^2}{16\pi^2 \ln \alpha^{-1}}. \quad (\text{D.12})$$

Here $T(r)$ is the index of χ 's representation r . For reference, Co et al. [33] takes the values $b \sim 1/100$ and $\alpha^2 \sim 10^{-3}$ for a $SU(3)_{\text{QCD}}$ gauge interaction. Note that [239] includes additional $\mathcal{O}(1)$ factors to these results, which we neglect here.

Because of the loop suppression, we assume that direct χ -decay dominates when it is kinematically allowed, such that that the perturbative decay rate can be summarized as

$$\Gamma_{\text{decay}} \approx \begin{cases} \Gamma_{\phi\bar{\chi}\chi} \approx y^2 m_{\phi,0}/(8\pi) & \text{if } m_{\phi,0} < gT \quad \text{and} \quad m_{\phi,0} < y\phi, \\ \Gamma_{\phi\gamma\gamma} \approx \alpha^2 m_{\phi,0}^3/\phi^2 & \text{if } m_{\phi,0} < gT \quad \text{and} \quad m_{\phi,0} > y\phi. \end{cases} \quad (\text{D.13})$$

Scattering processes: Even if direct perturbative decay is not permitted, ϕ may still dissipate energy by scattering with the thermal plasma, which allows for stimulated decay. This can occur directly with the χ fermions if they participate in the plasma (i.e. if they are relativistic, $T > y\phi$), or through χ -mediated γ -interactions if the χ are too heavy to avoid Boltzmann suppression. The

dissipation rate for scattering was calculated under the assumption of adiabatic ϕ -dynamics by Mukaida and Nakayama [238], i.e. under the assumption $m_{\phi,tot} < \alpha T$. As a rough approximation, which is also made by Mukaida and Nakayama [238], we here extrapolate the range of validity to include the otherwise not covered range of $\alpha T < m_{\phi,tot} < gT$. Thereby we consider scattering processes valid until perturbative decay becomes allowed.

For the case of relativistic fermions, i.e. $T > y\phi$, the dissipation rate depends on whether the χ are dominated by their thermal mass or by the ϕ -induced mass. The decay rates are estimated by Mukaida and Nakayama [238, 239] to be:

$$\Gamma_{\chi\text{-scat.}} \approx \begin{cases} y^2 \alpha T & \text{if } y\phi < \alpha T, \\ y^4 \frac{\phi^2}{\alpha T} & \text{if } \alpha T < y\phi < T. \end{cases} \quad (\text{D.14})$$

If the fermions are Boltzmann suppressed, then the scattering rate can instead be calculated by integrating out the χ fermions. This results in a gauge boson scattering rate of

$$\Gamma_{\gamma\text{-scat.}} \approx \frac{b\alpha^2 T^3}{\phi^2}. \quad (\text{D.15})$$

Redshift: The redshift of ϕ is modified in the presence of a thermal mass. If we assume that the mass changes adiabatically, then we can apply the usual WKB approximation:

$$\langle \phi \rangle \propto m_{\phi,\text{eff}}^{-1/2} a^{-3/2}. \quad (\text{D.16})$$

With the properties

$$m_{\phi,0} \propto T^0 \propto a^0, \quad m_{\phi,th} \propto T \propto a^{-1} \quad \text{and} \quad m_{\phi,ln} \propto \frac{T^2}{\phi} \propto \frac{a^{-2}}{\phi}, \quad (\text{D.17})$$

we can then infer that

$$\phi \propto \begin{cases} a^{-3/2} & \text{for } m_{\phi,tot} \sim m_{\phi,0}, \\ a^{-1} & \text{for } m_{\phi,tot} \sim m_{\phi,th}, \\ a^{-1} & \text{for } m_{\phi,tot} \sim m_{\phi,ln}. \end{cases} \quad (\text{D.18})$$

E Restrictions on the inflaton reheating temperature

In this work, we generally avoid model-dependent references to the reheating temperature of the inflaton. Thus, we generally allow for any reheating temperature between T_{kick} and the energy scale of inflation E_I . However, there is a regime in which the choice of reheating temperatures is restricted by the consistency of the scenario, which we explore in this appendix. This discussion is specific to the nearly-quadratic models.

Because thermal effects only enhance the radial potential and not the angular potential, ϵ and yield are suppressed in scenarios where the radial mode is dominated by thermal mass contributions at the time of the kick. We, therefore, favour solutions where $m_{\phi,0}$ dominates at the time of the kick. Such $m_{\phi,0}$ dominated scenarios can be realized in the regime where the χ -fermions are Boltzmann suppressed, and the thermal potential is reduced from $m_{\phi,\text{th}}$ to the much smaller $m_{\phi,\text{ln}}$ at the time of the kick. However, this scenario is only consistent if there is no earlier time in which an early phase of relativistic χ -fermions contribute a thermal mass large enough to start a kick.

To see how a would-be $m_{\phi,0}$ -dominated kick can be spoiled by an earlier phase of relativistic χ -fermions, note that prior to the kick ϕ will track the early potential minimum given by eq. (12.4), which evolves as

$$\phi \propto T^{\frac{2}{n-2}} \quad \text{before } T_{\text{kick}}. \quad (\text{E.1})$$

Thus, ϕ redshifts slower than T for any choice of $n > 4$. Therefore, any era in which $y\phi > T$ will always be preceded by an era in which $y\phi < T$ if the T can be extrapolated to arbitrarily high temperatures, i.e. if T_{reheat} is arbitrarily high. In such an earlier phase of relativistic χ -fermions, ϕ acquires the usual thermal mass $m_{\phi,\text{th}}$, which can be much larger than $m_{\phi,\text{ln}}$ and which can therefore lead to an early kick. For some choices of parameters, this phenomenon makes solutions in which the kick is given in the non-relativistic phase, i.e. with $y\phi_{\text{kick}} > T_{\text{kick}}$, inconsistent unless reheating after inflation takes place in the window between the end of any would-be early relativistic phase with $m_{\phi,\text{th}} > 3H$ and T_{kick} itself. A specific example of how different choices T_{reheat} can choose between two different families of solutions is illustrated in figure 32. An exemplary map of the $[T_{\text{reheat}}, y]$ parameter space is given in figure 33, where the blue region corresponds to parameter space in which the kick takes place with $m_{\phi,\text{th}}$ and the triangular region around $y \sim 10^{-5}$ corresponds to region in which an early $m_{\phi,\text{th}}$ kick can be avoided by postulating a sufficiently low T_{reheat} .

In the regime for which either a $m_{\phi,0}$ -dominated kick or a $m_{\phi,\text{th}}$ -dominated kick is possible depending on T_{reheat} we in this work assume that T_{reheat} is chosen such that the $m_{\phi,0}$ kick is realized. We do not make any further assumptions on T_{reheat} other than $E_I > T_{\text{reheat}} > T_{\text{kick}}$. In most of our parameter space, our solutions are compatible with any $E_I > T_{\text{reheat}} > T_{\text{kick}}$.

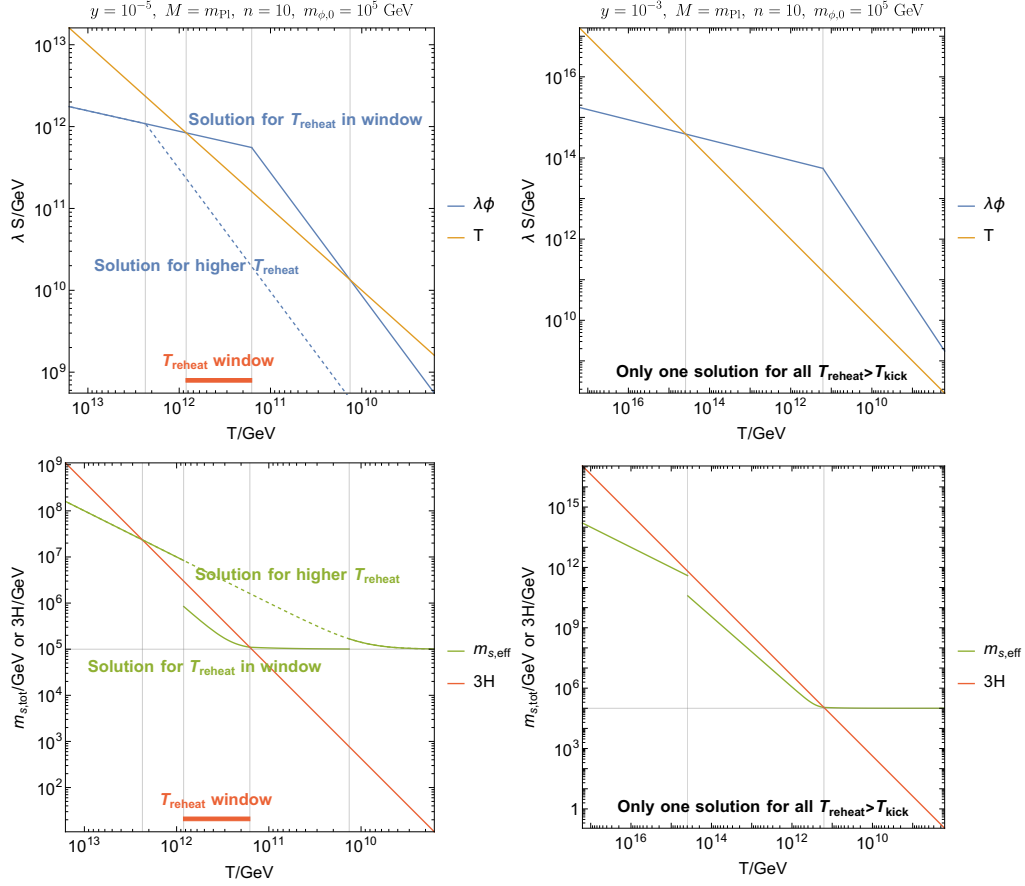


Figure 32. This figure illustrates how T_{reheat} has an impact on which type of kick is realized. **Left:** If T_{reheat} falls in the indicated window then a kick with non-relativistic fermions and $m_{\phi,\text{eff}} \approx m_{\phi,0}$ will be realized. If T_{reheat} is larger than this window, then a kick will be given with $m_{\phi,\text{th}} \gg m_{\phi,0}$ instead. The evolution corresponding to this early kick is indicated with the dashed lines. **Right:** For this choice of parameters the kick is given with $m_{\phi,\text{eff}} \approx m_{\phi,0}$ for any $T_{\text{reheat}} > T_{\text{kick}}$ because the fermions become non-relativistic before a $m_{\phi,\text{th}}$ -dominated kick can take place, even if the T_{reheat} is assumed to be arbitrarily high. The transition between scenarios in which either type of kick is possible depending on T_{reheat} and scenarios in which a $m_{\phi,0}$ -dominated kick is realized regardless of T_{reheat} leads to the discontinuity is seen around $y \approx 10^{-3.5}$ in fig. 33.

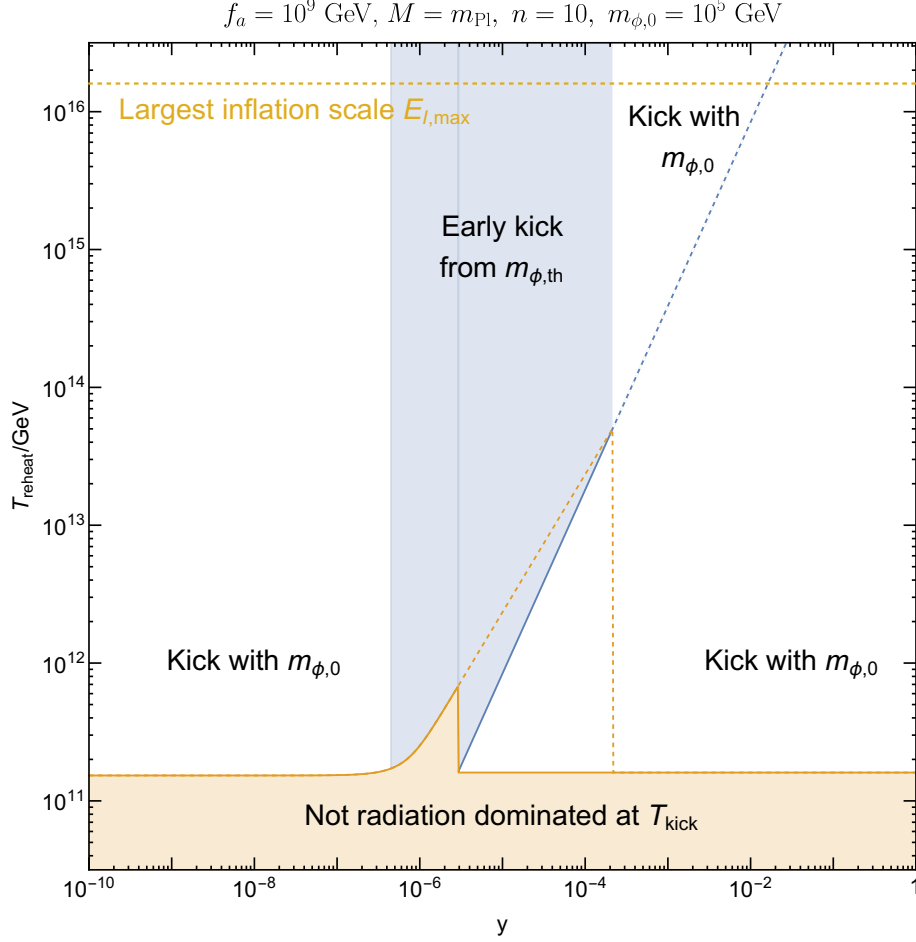


Figure 33. *Impact of T_{reheat} on the choice of kick type. In the white region the kick is realised with $m_{\phi,\text{eff}} \approx m_{\phi,0}$ while in the blue region the kick is realised with $m_{\phi,\text{eff}} \approx m_{\phi,\text{th}}$. The orange range corresponds to $T_{\text{reheat}} < T_{\text{kick}}$, which we do not consider in this work. **In the central range of y** , approximately $10^{-5.5} \lesssim y \lesssim 10^{-3.5}$, the choice of T_{reheat} can determine which type of kick is realized. This T_{reheat} -dependent region corresponds to the left column of fig. 32. **The high- y region**, approximately $10^{-3.5} \lesssim y$, in which only the late $m_{\phi,\text{eff}} \approx m_{\phi,0}$ kick can be realized corresponds to the right column of fig. 32. **In the low- y region**, i.e. $y \lesssim 10^{-6.5}$, the kick can only take place with $m_{\phi,0}$ because a $m_{\phi,\text{th}}$ -dominated kick fails to take place before the zero-temperature mass becomes dominant. **Lines:** The orange dashed line corresponds to T_{kick} in the $m_{\phi,\text{th}}$ -dominated solution and lowest temperature for which $T > y\phi$ corresponds to blue line. **Other remarks:** Note that the parameters here are fixed such that the correct DM-yield is realized with the $m_{\phi,0}$ -dominated kick, i.e. in the white parameter space. For simplicity, we have not imposed any other constraints on this plot. Therefore, the white parameter space needs not correspond to physically viable scenarios.*

F Boltzmann equations and solutions of thermal damping

In this appendix, we derive the Boltzmann equations, which describe thermal damping. We then describe the numerical scheme we use to solve these systems of equations. We first consider the nearly-quadratic potential and then generalize to the quadratic potential. Finally, in subsection F.2 we present examples of solutions.

F.1 Derivation

We first consider the case of the nearly-quadratic potential. We will, as usual, neglect the logarithmic correction and restrict ourselves to the regime in which higher-dimensional terms are suppressed such that the potential is

$$V \sim \frac{1}{2} m_{\phi, \text{eff}}^2 \phi^2, \quad (\text{F.1})$$

where we have included the thermal correction in $m_{\phi, \text{eff}}$. The equations of motion for the radial and angular degrees of freedom of the PQ field are

$$\ddot{\phi} + 3H\dot{\phi} + \Gamma\dot{\phi} + m_{\phi, \text{eff}}^2 \phi = \dot{\theta}^2 \phi, \quad (\text{F.2})$$

$$\ddot{\theta} + 3H\dot{\theta} = -2\frac{\dot{\theta}\dot{\phi}}{\phi}, \quad (\text{F.3})$$

where Γ is the total dissipation rate. The total energy density of the field is

$$\rho_{\text{tot}} = \rho_{\text{kin}}^{\phi} + \rho_{\text{kin}}^{\theta} + \rho_{\text{pot}} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}\phi^2\dot{\theta}^2 + \frac{1}{2}m_{\phi, \text{eff}}^2\phi^2. \quad (\text{F.4})$$

To express the EOM in terms of Boltzmann equations, we consider the derivative of the energy density. If we allow for a time-dependence in the mass term from, e.g., thermal corrections, then we find

$$\dot{\rho}_{\text{tot}} = \dot{\phi}\ddot{\phi} + \phi\dot{\phi}\ddot{\theta}^2 + \phi^2\dot{\theta}\ddot{\theta} + m_{\phi, \text{eff}}\dot{m}_{\phi, \text{eff}}\phi^2 + m_{\phi, \text{eff}}^2\phi\dot{\phi}. \quad (\text{F.5})$$

We then apply the EOM and reduce:

$$\dot{\rho}_{\text{tot}} = -(3H + \Gamma)\dot{\theta}^2 - 3H\dot{\theta}^2\phi^2 + m_{\phi, \text{eff}}\dot{m}_{\phi, \text{eff}}\phi^2. \quad (\text{F.6})$$

It is convenient to parametrise the time dependence of $m_{\phi, \text{eff}}$ in terms of the parameter $c_{m_{\phi}}$,

$$m_{\phi, \text{eff}} \propto a^{-c_{m_{\phi}}}. \quad (\text{F.7})$$

In terms of this parameter $c_{m_{\phi}}$ the last term of eq. (F.6) becomes

$$m_{\phi, \text{eff}}\dot{m}_{\phi, \text{eff}}\phi^2 = -c_{m_{\phi}}Hm_{\phi, \text{eff}}^2\phi^2, \quad (\text{F.8})$$

so that $\dot{\rho}_{\text{tot}}$ can be written as

$$\dot{\rho}_{\text{tot}} = -(6H + 2\Gamma)\rho_{\text{kin}}^{\phi} - 6H\rho_{\text{kin}}^{\theta} + -2c_{m_{\phi}}H\rho_{\text{pot}}. \quad (\text{F.9})$$

Virial theorem: To understand the relation between the potential and kinetic energy of the system, we apply the virial theorem, which for a monomial potential $\propto \phi^n$ states that

$$\frac{n}{2}\langle V \rangle = \langle T \rangle \quad \text{where } n = 2 \text{ for a quadratic potential,} \quad (\text{F.10})$$

where V is the total potential energy, and T is the total kinetic energy. We then have

$$\langle \rho_{\text{pot}} \rangle = \langle \rho_{\text{kin}}^\theta + \rho_{\text{kin}}^\phi \rangle. \quad (\text{F.11})$$

We henceforth drop the brackets and understand all energy densities as time-averaged quantities. Combining eq. (F.11) with eq. (F.4) we conclude that

$$\rho_{\text{tot}} = 2\rho_{\text{kin}}^\phi + 2\rho_{\text{kin}}^\theta, \quad (\text{F.12})$$

so that

$$\dot{\rho}_{\text{tot}} = 2\dot{\rho}_{\text{kin}}^\phi + 2\dot{\rho}_{\text{kin}}^\theta = -(6H + 2\Gamma + 2c_{m_\phi}H)\rho_{\text{kin}}^\phi - (6H + 2c_{m_\phi}H)\rho_{\text{kin}}^\theta. \quad (\text{F.13})$$

If we consider ρ_{kin}^ϕ and ρ_{kin}^θ to be independent we can separate this equation into two Boltzmann equations:

$$\dot{\rho}_{\text{kin}}^\phi = -(3H + \Gamma + c_{m_\phi}H)\rho_{\text{kin}}^\phi, \quad (\text{F.14})$$

$$\dot{\rho}_{\text{kin}}^\theta = -(3H + c_{m_\phi}H)\rho_{\text{kin}}^\theta. \quad (\text{F.15})$$

Let us denote the energy of the circular orbit corresponding to a given angular velocity $\dot{\theta}$ as ρ_{circ} . Since for a circular orbit we have $\dot{\theta} = m_{\phi,\text{eff}}$ it follows that $\rho_{\text{circ}} = \rho_{\text{kin}}^\phi$. We can therefore express the Boltzmann equations as a Boltzmann equation for the circular orbit and a Boltzmann equation for the oscillations about this orbit:

$$\dot{\rho}_{\text{kin}}^\phi = -(3H + \Gamma + c_{m_\phi}H)\rho_{\text{kin}}^\phi, \quad (\text{F.16})$$

$$\dot{\rho}_{\text{circ}} = -(3H + c_{m_\phi}H)\rho_{\text{circ}}. \quad (\text{F.17})$$

With a total energy density of $\rho_{\text{tot}} = 2\rho_{\text{kin}}^\phi + 2\rho_{\text{circ}}$. By conservation of energy this can be linked to the energy density of the plasma:

$$\dot{\rho}_r = -4H\rho_r + \Gamma\rho_{\text{kin}}^\phi. \quad (\text{F.18})$$

The set of equations eq. (F.16-F.18) are then the equations we will solve numerically to describe the thermal damping of the PQ field. To complete the numerical setup, we also need to determine the initial conditions and specify how to describe the transition between phases of relativistic/non-relativistic fermions, where we do not have a description of the mass change of the form $m \propto a^{-c_m}$.

Initial conditions: After the kick, we assume that the energy density is composed of angular kinetic and potential energy, i.e.

$$\rho_{\text{ini}} = \frac{1}{2}m_{\phi,\text{eff}}^2\phi_{\text{kick}}^2 + \frac{1}{2}\dot{\theta}_{\text{kick}}^2\phi_{\text{kick}}^2 = \frac{1+\epsilon^2}{2}m_{\phi,\text{eff}}^2\phi_{\text{kick}}^2. \quad (\text{F.19})$$

After the kick this energy is distributed into ρ_{kin}^ϕ , ρ_{kin}^θ and ρ_{pot} . To identify the distribution, note that in [37] it was found that after the kick

$$\rho_{\text{kin}}^\theta = \epsilon\rho_{\text{pot}}, \quad (\text{F.20})$$

so that in combination with the virial theorem, this implies

$$\rho_{\text{pot}} = \frac{1}{2}\rho_{\text{ini}}, \quad (\text{F.21})$$

$$\rho_{\text{kin}}^\phi = \frac{1-\epsilon}{2}\rho_{\text{ini}} \quad (\text{F.22})$$

$$\rho_{\text{kin}}^\theta = \frac{\epsilon}{2}\rho_{\text{ini}}. \quad (\text{F.23})$$

These equations then specify the initial conditions for the Boltzmann equations.

Relativistic/non-relativistic fermion transition: The parameter k takes into account continuous and gradual changes in the mass as with the temperature-dependence of a thermal mass. However, we generally approximate the transition between phases with relativistic and non-relativistic fermions as instantaneous because it is non-trivial to express in greater detail. Therefore we cannot encode the change in radial mass across such a transition in the parameter k as described above. To take the transition into account in a physical way, we will instead assume that the transition is gradual enough to be adiabatic such that the WKB approximation $\phi \propto m^{-1/2}$ applies. Across a transition where the mass changes from $m \rightarrow m^*$ the WKB approximation implies that ϕ and an energy density $\rho \propto m^2 \phi^2$ is modified as

$$\phi \xrightarrow{m \rightarrow m^*} \left(\frac{m}{m^*} \right)^{\frac{1}{2}} \phi, \quad (\text{F.24})$$

$$\rho \xrightarrow{m \rightarrow m^*} \left(\frac{m^*}{m} \right) \rho. \quad (\text{F.25})$$

We use this solution to account for a change in $m_{\phi, \text{eff}}$ across a relativistic/non-relativistic fermion transition. We assume that ρ_{pot} , ρ_{kin}^{ϕ} and $\rho_{\text{kin}}^{\theta}$ are all shifted by the same factor and that the energy required for this change is transferred to or from the radiation bath.

In the numerical Boltzmann solver, we then apply this transformation to ϕ and the energy densities whenever the fermions transition between their relativistic and non-relativistic phases.

Quartic potential For a quartic potential,

$$V \sim \frac{1}{4} \lambda^2 \phi^4, \quad (\text{F.26})$$

we can follow the same steps as above to arrive again at

$$\dot{\rho}_{\text{tot}} = -6H\rho_{\text{kin}}^{\phi} - 2\Gamma\rho_{\text{kin}}^{\phi} - 6H\rho_{\text{kin}}^{\theta}. \quad (\text{F.27})$$

The virial theorem for ϕ^n with $n = 4$ now states that $2\langle V \rangle = \langle T \rangle$, so that

$$\rho_{\text{pot}} = \frac{1}{2}\rho_{\text{kin}}^{\phi} + \frac{1}{2}\rho_{\text{kin}}^{\theta}. \quad (\text{F.28})$$

The combined Boltzmann equation is then

$$\frac{3}{2}\dot{\rho}_{\text{kin}}^{\phi} + \frac{3}{2}\dot{\rho}_{\text{kin}}^{\theta} = -6H\rho_{\text{kin}}^{\phi} - 2\Gamma\rho_{\text{kin}}^{\phi} - 6H\rho_{\text{kin}}^{\theta}. \quad (\text{F.29})$$

If we again assume that ρ_{kin}^{ϕ} and $\rho_{\text{kin}}^{\theta}$ are independent, then we arrive at the following two Boltzmann equations:

$$\dot{\rho}_{\text{kin}}^{\phi} = -4H\rho_{\text{kin}}^{\phi} - \frac{4}{3}\Gamma\rho_{\text{kin}}^{\phi}, \quad (\text{F.30})$$

$$\dot{\rho}_{\text{kin}}^{\theta} = -4H\rho_{\text{kin}}^{\theta}. \quad (\text{F.31})$$

If a thermal potential becomes dominant, then the evolution is instead determined by [F.14](#) and [F.15](#) as in the quadratic case. The initial conditions are modified to

$$\rho_{\text{pot}} = \frac{1}{3}\rho_{\text{ini}}, \quad (\text{F.32})$$

$$\rho_{\text{kin}}^{\phi} = (1 - \epsilon)\frac{2}{3}\rho_{\text{ini}} \quad (\text{F.33})$$

$$\rho_{\text{kin}}^{\theta} = \epsilon\frac{2}{3}\rho_{\text{ini}}. \quad (\text{F.34})$$

F.2 Examples of solutions

We here provide two examples of solutions produced with the Boltzmann prescription detailed above:

Fig. 34: Solution from a model with a nearly-quadratic potential which features damping through χ -fermion scattering triggered by the onset of an intermediate relativistic phase.

Fig. 35: Solution from a model with a nearly-quadratic potential which features two-step damping through χ -fermion scattering near the end of a relativistic phase and saxion-to-axion decay.

In both examples, the top two frame shows the evolution of the damping rates, the middle frame shows the evolution of the energy densities and the third frame shows the χ -fermion mass compared to the temperature.

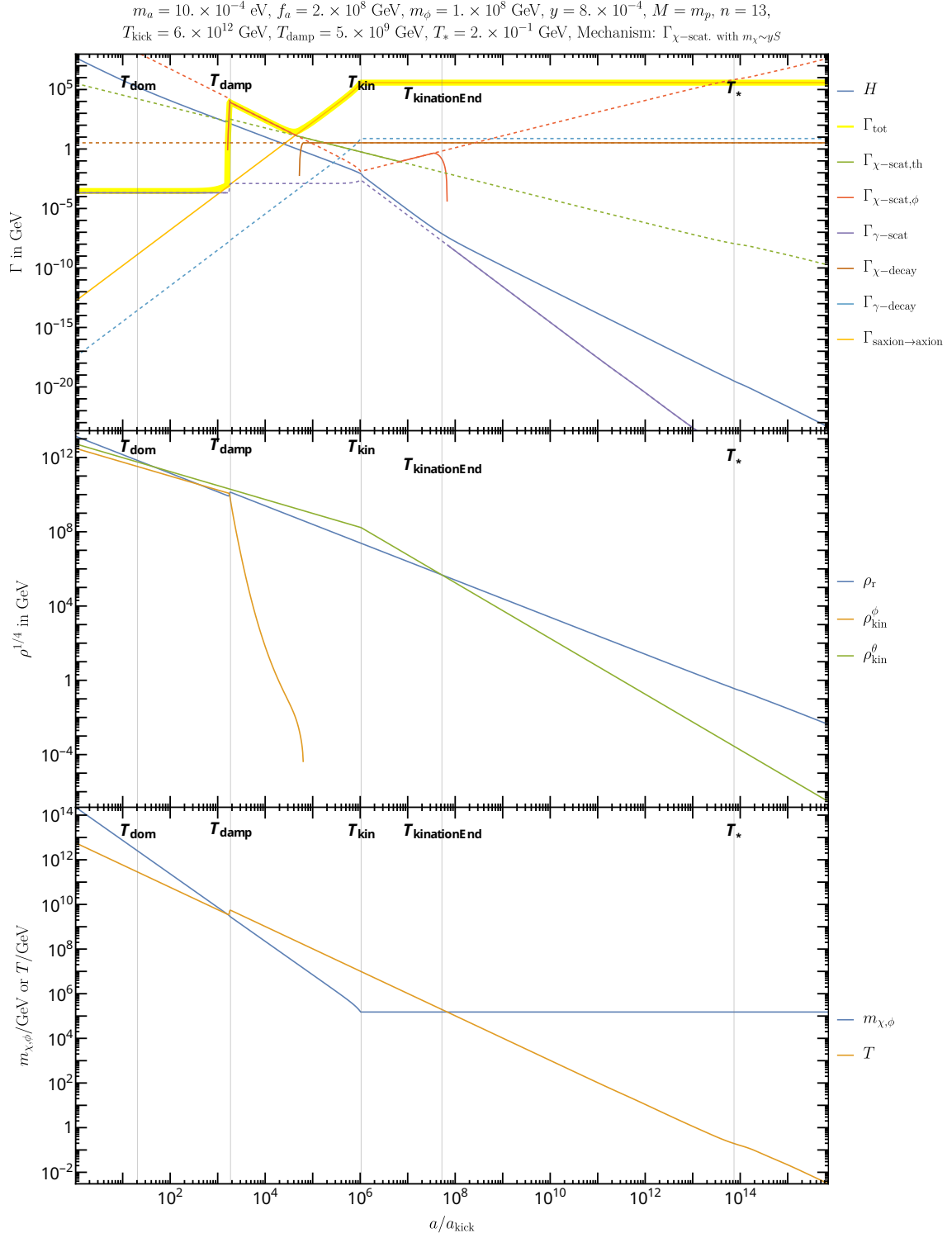


Figure 34. Example of thermal damping within the KSVZ-like nearly-quadratic model, which assumes Yukawa damping. The solution is found by numerical solution of eqs. (13.10-13.13) and damping is triggered by an intermediate relativistic phase. Various events of interest are tagged with vertical lines. **Top:** Damping rates compared to the Hubble parameter. Dashed lines indicate that the conditions for a given effect are not met. **Middle:** Evolution of the energy densities. Note the transfer of kinetic energy from ρ_ϕ to ρ_r . **Bottom:** Evolution of temperature as compared to the fermion mass.

$$m_a = 7. \times 10^{-4} \text{ eV}, f_a = 3. \times 10^{10} \text{ GeV}, m_\phi = 2. \times 10^3 \text{ GeV}, y = 6. \times 10^{-7}, M = m_p, n = 13,$$

$$T_{\text{kick}} = 2. \times 10^{10} \text{ GeV}, T_{\text{damp}} = 20. \text{ GeV}, T_* = 3. \text{ GeV}, \text{Mechanism: } \Gamma_{\text{axion-decay}}$$

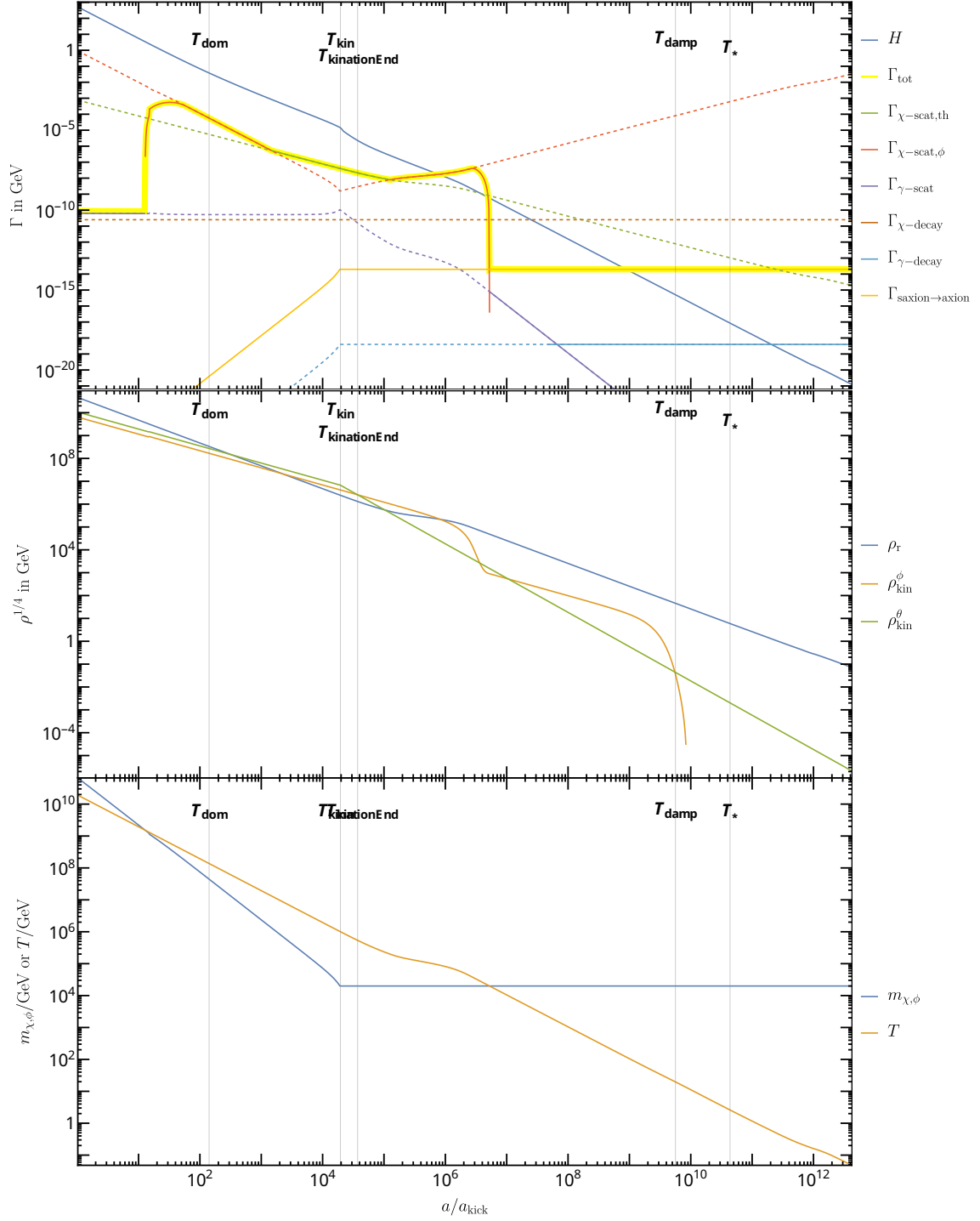


Figure 35. Similar to figure 34, but for another choice of parameters for which damping is not triggered by an intermediate relativistic phase. Instead, late damping is partially realized towards the end of a relativistic phase. Here, χ -scattering is insufficient to completely damp ρ_{kin}^ϕ . Damping is completed by saxion-to-axion decay. The initial, incomplete, damping through χ -scattering prevents the production of a cosmologically dangerous hot axion relic.

G Additional details for nearly-quadratic models

We in this appendix provide additional details for the nearly-quadratic model.

Scale of $A_s(k_{\text{kin}})$ suppression: It is essential to the nearly-quadratic model that $A_s(k_{\text{kin}})$ can be suppressed. Planck [241] constrains A_s to $A_s(k_{\text{pivot}}) = A_{s,\text{Planck}} \approx 2.1 \times 10^{-9}$ at the pivot scale $k_{\text{pivot}} = 0.05 \text{ Mpc}^{-1}$. To demonstrate that this pivot scale is far removed from the scales which are relevant for axion kinetic misalignment we in figure 36 display the comoving momentum scale k_{kin} relevant for $A_s(k_{\text{kin}})$. It is clear that this pivot scale is many orders of magnitude removed from k_{kin} , such that $A_s(k_{\text{kin}})$ may very well deviate from the value constrained by Planck.

Alternative choices of n : Furthermore, we provide an overview of the viable parameter space for $n = 7$ and $n = 10$ in figure 37. In both cases, we find that less parameter space is supported than in the $n = 13$ scenario. $A_s(k_{\text{kin}})$ suppression is also required in these cases. For $n = 10$, the lower bound on m_a is set by thermal-log contributions to m_ϕ . We do not consider parameter space with thermal-log domination as such may feature oscillon formation [238]. We do not show contours of k_{kin} for $n = 10$ and $n = 7$. For these, the relevant scales are either similar or one to two orders of magnitude further from the Planck pivot scale.

Range of viable parameters We in this appendix connect the fundamental model parameters and the viable range of parameter space in the $[m_a, f_a]$ plane. These model parameters are the following six variables: m_a, f_a, y, M, n and $m_{\phi,0}$. As stated before, we assume that the axion relic makes up all of the dark matter which reduces the number of free parameters to five. To further reduce the number of free parameters we restrict ourselves to the well-motivated case of $M = m_{\text{Pl}}$. Thus, when plotting across the $[m_a, f_a]$ plane, each point may be supported by a family of solutions with a range of viable values of $y, m_{\phi,0}$, the values of which we here detail for $n = 13$ and $n = 10$.

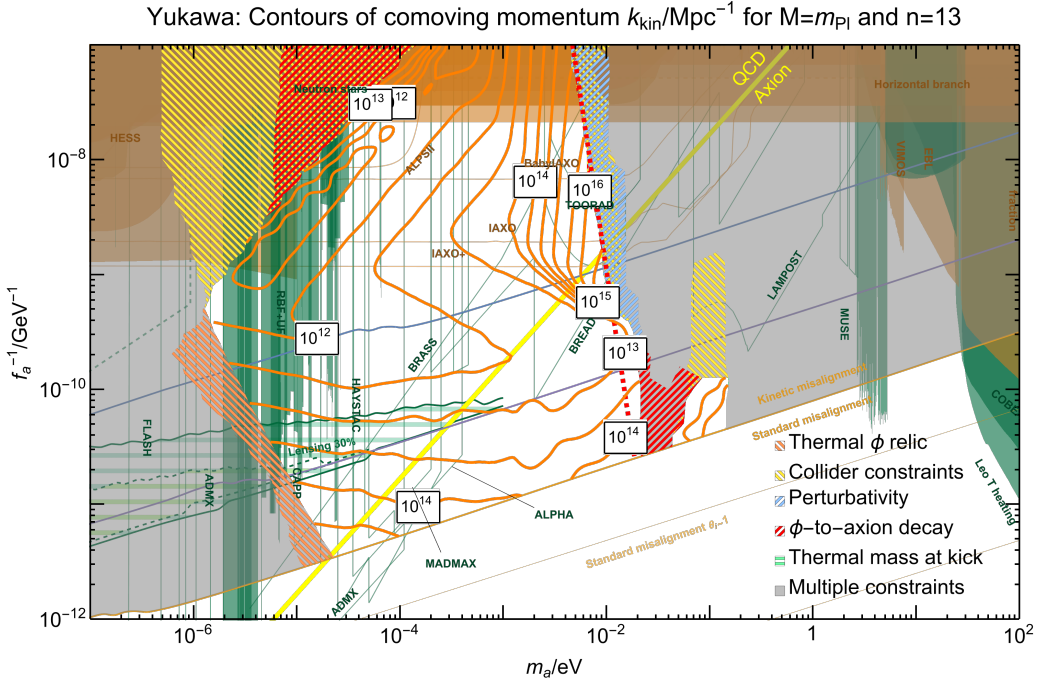
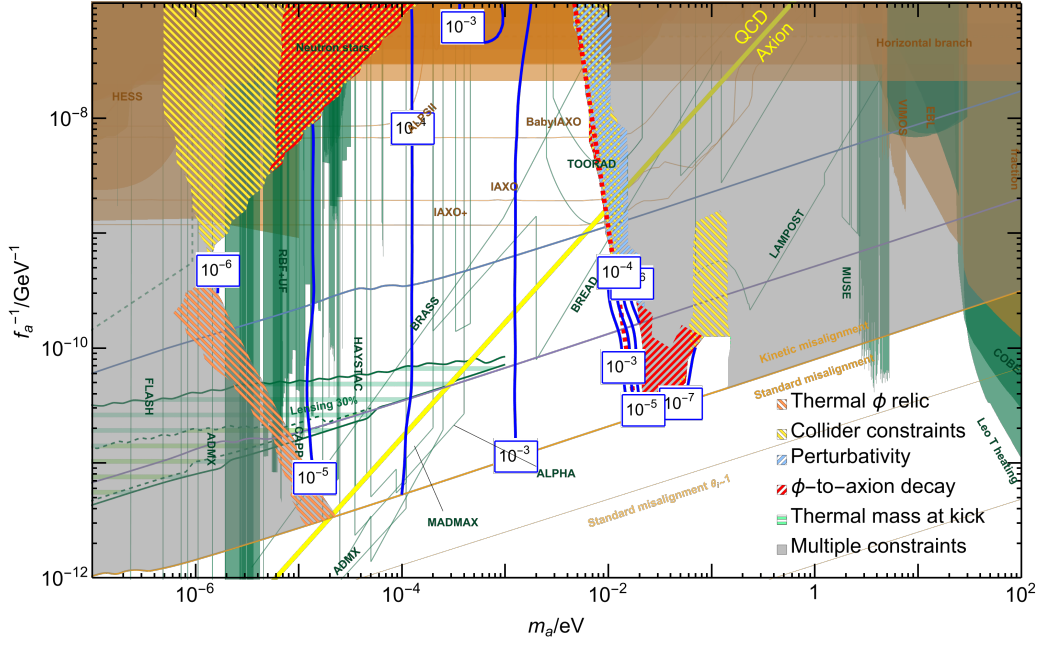


Figure 36. Map of the comoving momentum scale k_{kin} on which $A_s(k_{\text{kin}})$ should be suppressed to avoid the homogeneity constraint. The solutions shown here correspond to those shown in fig. 22. This clearly shows that $k_{\text{kin}} \gg 0.05 \text{ Mpc}^{-1}$, which is the scale probed by Planck.

Yukawa: Upper bound on y for $M=m_{\text{Pl}}$ and $n=13$



Yukawa: Lower bound on y for $M=m_{\text{Pl}}$ and $n=13$

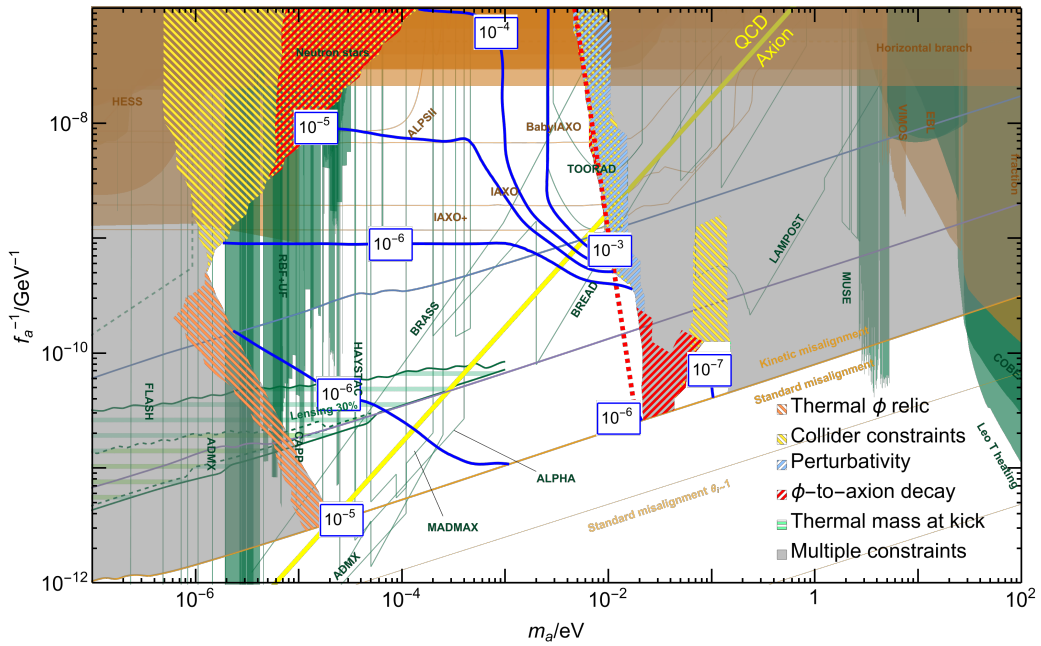
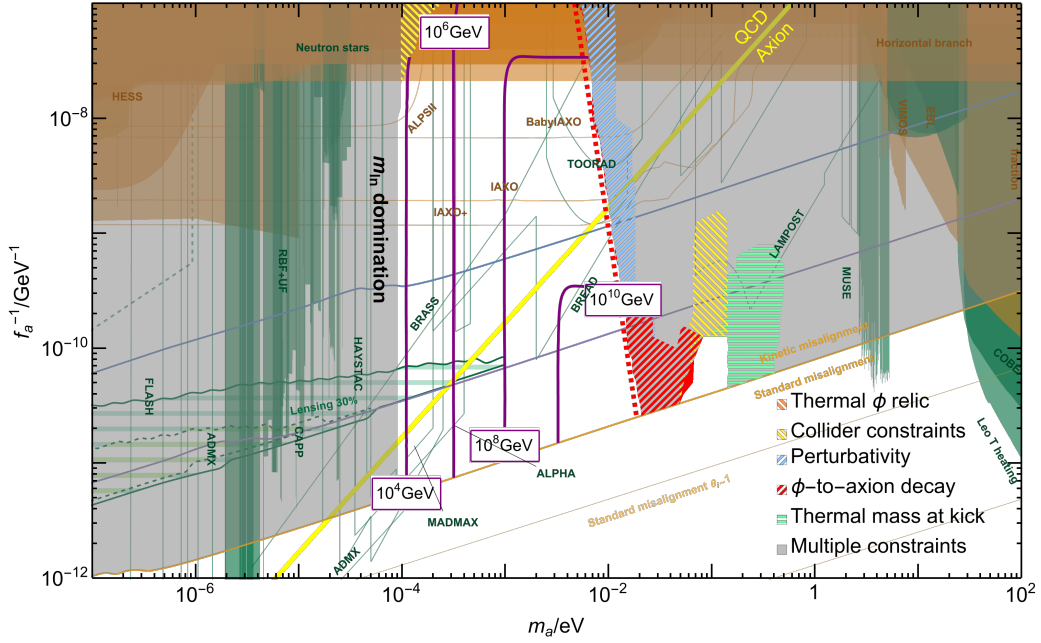


Figure 39. Upper and lower bounds on the Yukawa coupling y in Yukawa damped models with $n = 13$.

Yukawa: Upper bound on $m_{\phi,0}$ for $M=m_{\text{Pl}}$ and $n=10$



Yukawa: Lower bound on $m_{\phi,0}$ for $M=m_{\text{Pl}}$ and $n=10$

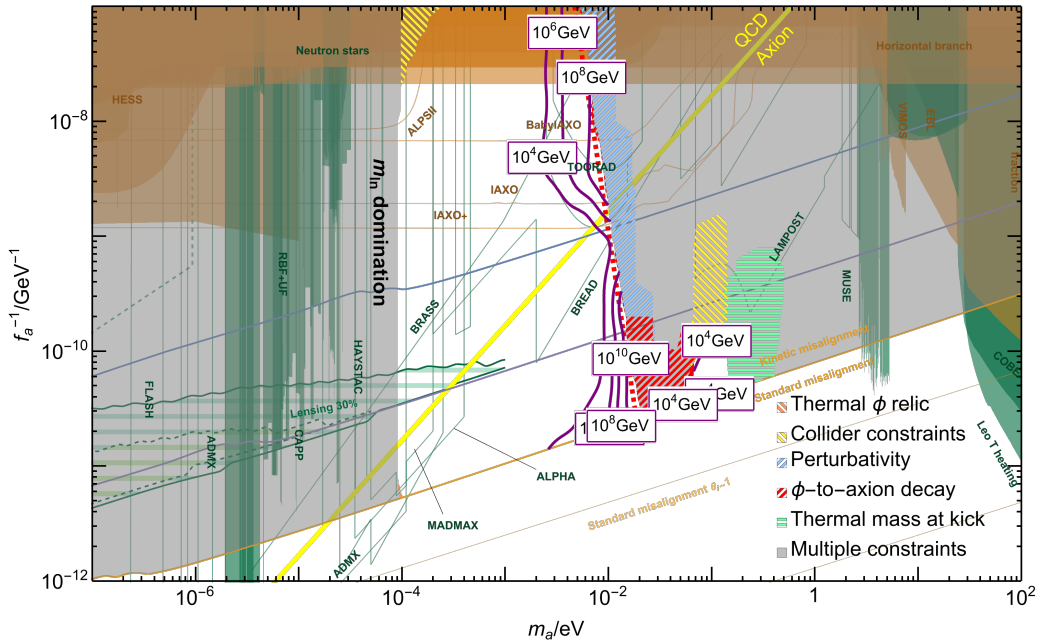
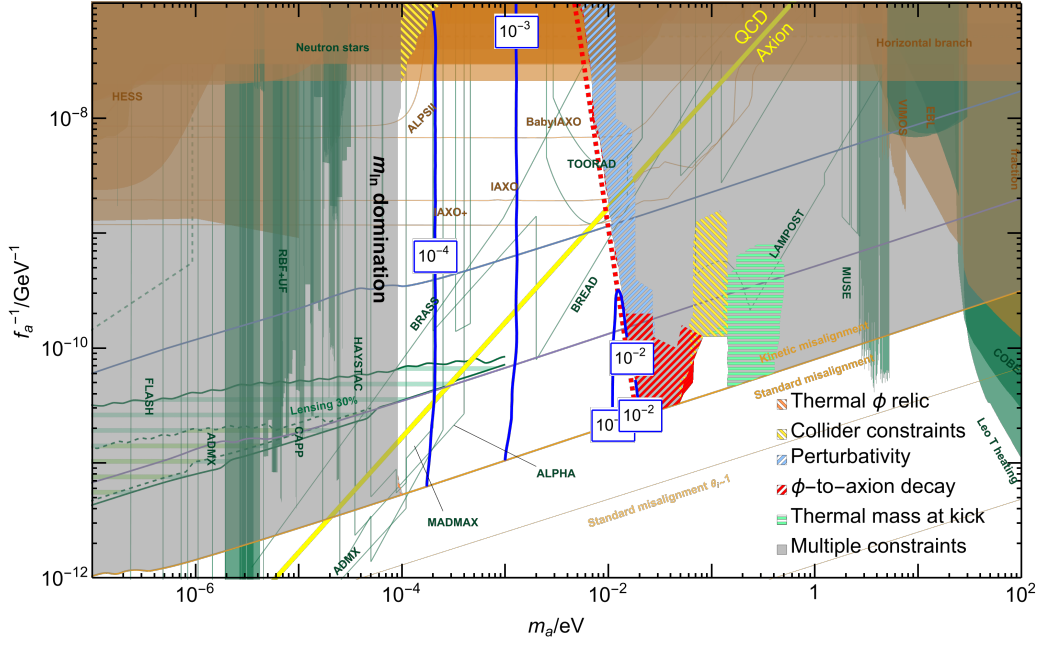


Figure 40. Upper and lower bounds on the radial mass $m_{\phi,0}$ in Yukawa damped models with $n = 10$.

Yukawa: Upper bound on y for $M=m_{\text{Pl}}$ and $n=10$



Yukawa: Lower bound on y for $M=m_{\text{Pl}}$ and $n=10$

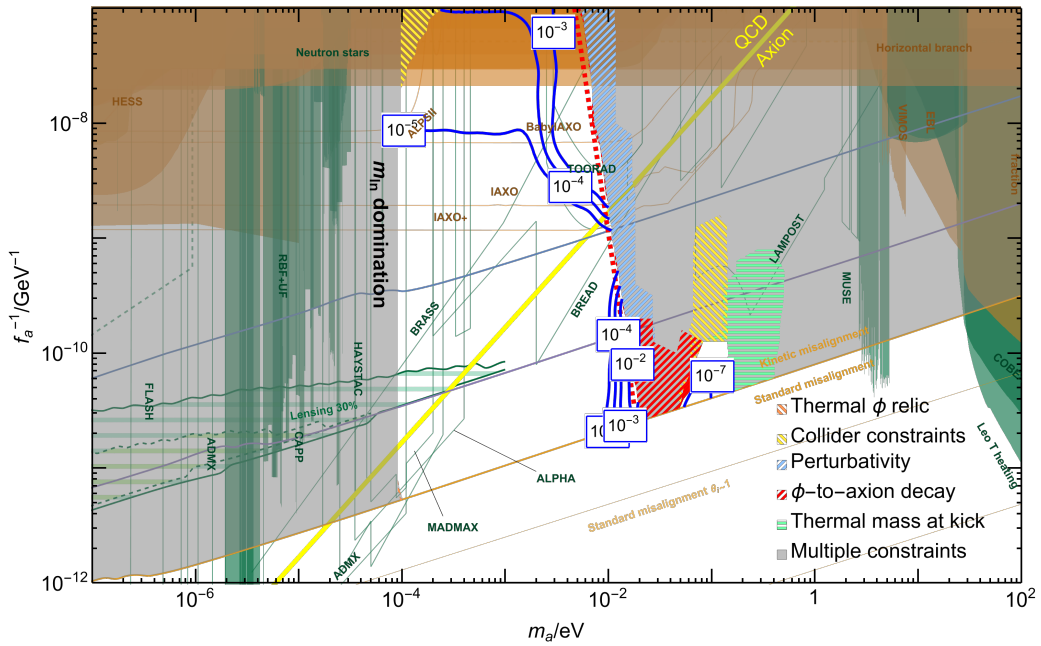


Figure 41. Upper and lower bounds on the Yukawa coupling y in Yukawa damped models with $n = 10$.

H Additional details for quartic models

We in this appendix provide additional material on the quartic models. In particular, we in the following sections discuss the use of higher-dimensional terms, derive domain wall constraints and discuss why Yukawa damping is not viable in quartic models.

H.1 Higher-dimensional terms in the quartic model

We here investigate how higher-dimensional terms can be used to drive a large VEV in the quartic model. The early-time potential for this setup is

$$V_{\text{early}} = \lambda^2 \phi^4 - c_H H^2 \phi^2 + \frac{A + aH}{n} \frac{P^n}{M^{n-3}} + h.c. + \frac{|P|^{2n-2}}{M^{2n-6}}. \quad (\text{H.1})$$

As in the nearly-quadratic model, the Hubble-induced terms generate an early large mass around a VEV of

$$\phi_{\text{early}} = \left(\frac{2^{n-2}}{n-1} \right)^{\frac{1}{2n-4}} (HM^{n-3})^{\frac{1}{n-2}}. \quad (\text{H.2})$$

This time-dependent VEV is the same as in the nearly-quadratic model since the VEV is set purely by the higher dimensional terms. What does change is the VEV and temperature at the kick. The kick takes place at $m_\phi \sim \lambda\phi \sim 3H$ which implies

$$T_{\text{kick}} \approx 0.3 \times \sqrt{M m_{\text{Pl}} \lambda^{\frac{n-2}{2n-6}}}, \quad (\text{H.3})$$

$$\phi_{\text{kick}} \approx 1. \times M \lambda^{\frac{1}{n-3}} \quad (\text{H.4})$$

where the exact value of the prefactors depends on n and $g_*(T_{\text{kick}})$. Such a kick implies a yield of

$$Y_{\text{kick}} \approx 0.8 \times \epsilon \left(\frac{M}{m_{\text{Pl}}} \right)^{3/2} \lambda^{\frac{n-6}{2n-6}}. \quad (\text{H.5})$$

The observed DM yield is achieved for

$$\lambda \sim \mathcal{O}(1) \times \left(\frac{M}{m_{\text{Pl}}} \right)^{\frac{3(n-3)}{n-6}} \left(\frac{\epsilon m_a}{T_{\text{eq}}} \right)^{\frac{2(n-3)}{n-6}}, \quad (\text{H.6})$$

which determines the late-time radial mass $m_{\phi\text{-late}} \approx \lambda f_a$.

Similar to the nearly-quadratic models, this setup requires a light radial mode in order to satisfy the various constraints. However, while the nearly-quadratic models could motivate a light radial mass by appealing to SUSY, there is no comparable motivation for a very light radial mass in the quartic model. In particular, large regions of parameter space require $\lambda < 10^{-20}$ or $\lambda < 10^{-30}$ in order to produce a variable phenomenology. This tuning problem is made worse by the large m_a sensitivity in λ and m_ϕ . By comparing eq. (H.6) with eq. (14.11) we observe that

$$\text{Higher-Dim.: } m_\phi \propto m_a^k \quad \text{where } 2.8 \leq k \leq 8 \quad \text{for } 7 \leq n \leq 13, \quad (\text{H.7})$$

$$\text{de Sitter: } m_\phi \propto m_a^k \quad \text{where } k = 4/5. \quad (\text{H.8})$$

We conclude that although both mechanisms require significant tuning of λ , then a setup with higher dimensional terms requires significantly more tuning than a setup with de Sitter fluctuations. We, therefore, primarily consider de Sitter fluctuations as the driver of the large VEV in this work.

The viable parameter space for the quartic model with higher-dimensional terms under the assumption of early damping is displayed in figures 42 to 44. Regions with $\lambda < 10^{-20}$ is indicated by hatched filling. Note that significantly less parameter space is available than in the model with de Sitter fluctuations, as displayed in figures 27 and 28.

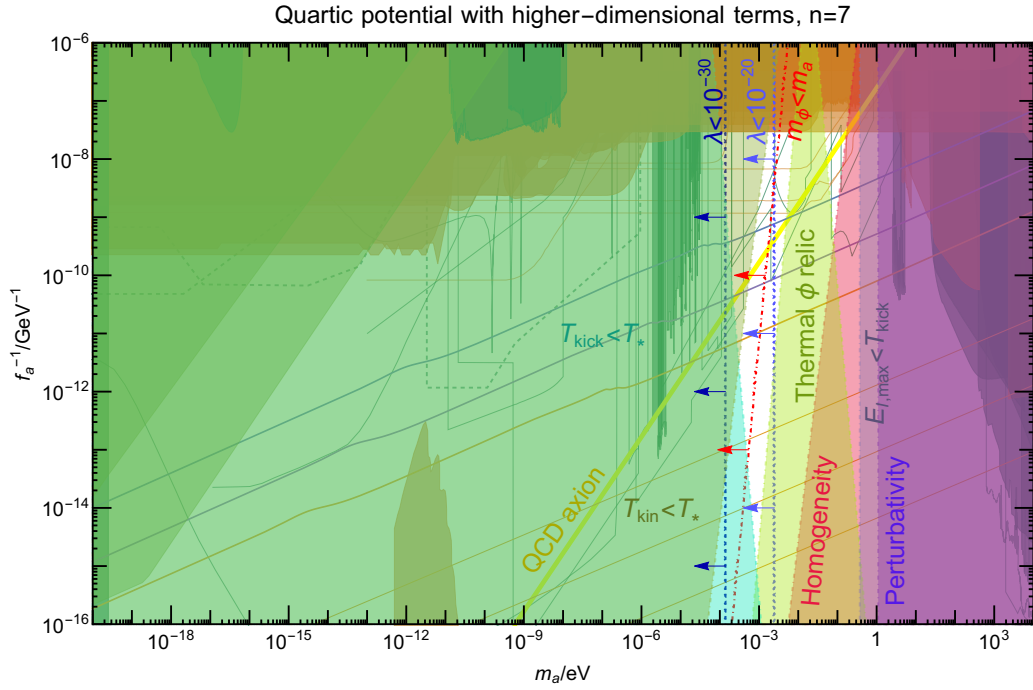


Figure 42. Constraints on the $[m_a, f_a]$ parameter space for $n = 7$ in the quartic model when the large VEV is driven by higher-dimensional operators rather than de Sitter fluctuations, see eq. (H.1). We here assume early damping without entropy dilution and $M = m_{\text{Pl}}$. Regions with quartic couplings lower than $\lambda < 10^{-20}$ and $\lambda < 10^{-30}$ are indicated by the blue dashed lines.

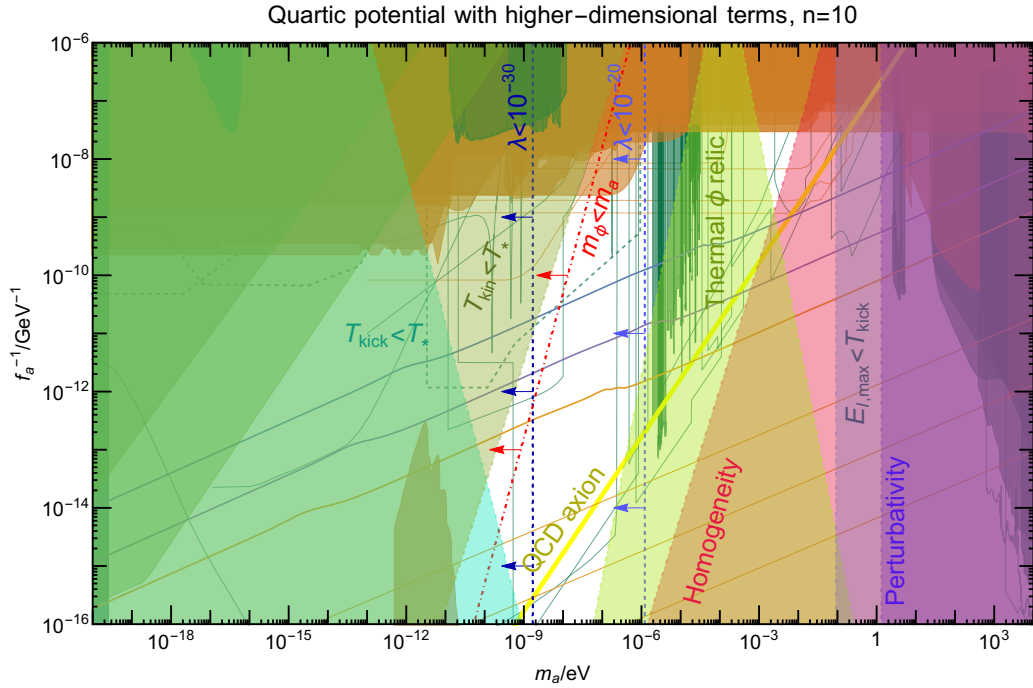


Figure 43. Same as figure 42 but for $n = 10$.

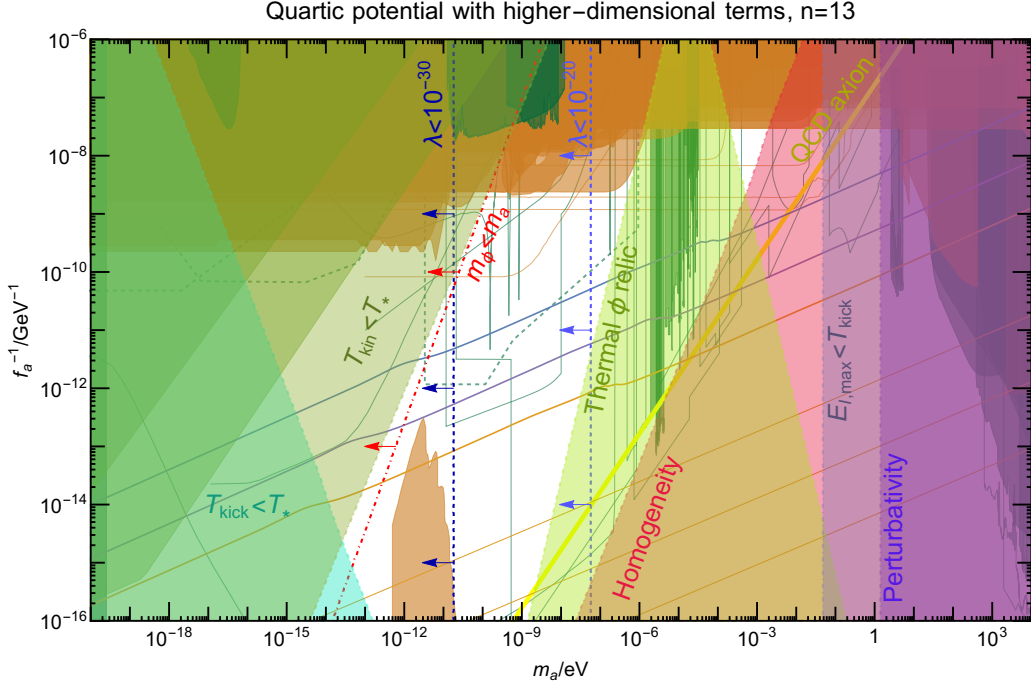


Figure 44. Same as figure 42 but for $n = 13$.

H.2 Domain wall constraints on the quartic model

We here derive the domain wall constraint on the quartic model. This constraint appears when fluctuations in the kick velocity evolve into $\mathcal{O}(1)$ angular fluctuations at the time of trapping. To estimate when this happens, we carry out the integration which was specified in eq. (14.19):

$$\theta_* \approx \int_{t_{\text{kick}}}^{t_*} \dot{\theta} dt \quad (\text{H.9})$$

We carry out the integration using physical time t as the integration variable. This physical time t is related to the scale factor a as

$$a \propto \begin{cases} t^{2/3} & \text{if matter dominated,} \\ t^{1/2} & \text{if radiation dominated,} \\ t^{1/3} & \text{if kination dominated.} \end{cases} \quad (\text{H.10})$$

In a quartic model, the universe is radiation dominated unless damping takes place sufficiently after a_{kin} that the energy of radial oscillations can come to dominate. For simplicity, we neglect such a period of matter domination, such that we can always take $a \propto t^{1/2}$. Applying in addition that $\dot{\theta} \propto a^{-1}$ we find for the quartic model:

$$\theta_* \approx \frac{3}{2} \dot{\theta}_{\text{kick}} t_{\text{kick}} \left(\frac{t_{\text{kin}}}{t_{\text{kick}}} \right)^{1/2}. \quad (\text{H.11})$$

Applying $\phi \propto a^{-1} \propto t^{-1/2}$ this becomes

$$\theta_* \approx \frac{3}{2} \dot{\theta}_{\text{kick}} t_{\text{kick}} \left(\frac{\phi_{\text{kick}}}{\phi_{\text{kin}}} \right) + \text{const.} \quad (\text{H.12})$$

The variation of the angular displacement is then

$$\Delta\theta_* \approx \theta_* \left[\frac{\Delta\dot{\theta}_{\text{kick}}}{\dot{\theta}_{\text{kick}}} + \frac{\Delta t_{\text{kick}}}{t_{\text{kick}}} + \frac{\Delta\phi_{\text{kick}}}{\phi_{\text{kick}}} \right]. \quad (\text{H.13})$$

From $\dot{\theta}_{\text{kick}} \propto \phi_{\text{kick}}^{n-2} \sin(n\theta_{\text{kick}})$ we can express the variation of the kick velocity in terms of variation of the angular and radial modes:

$$\frac{\Delta\dot{\theta}_{\text{kick}}}{\dot{\theta}_{\text{kick}}} \approx (n-2) \frac{\Delta\phi_{\text{kick}}}{\phi_{\text{kick}}} + n \cot(n\theta_{\text{kick}}) \frac{\Delta\theta_{\text{kick}}}{\theta_{\text{kick}}}. \quad (\text{H.14})$$

The fluctuations in the timing of the kick in quartic models can be inferred from $m_{\phi, \text{UV}} \approx \lambda\phi$ and $m_{\phi, \text{UV}} \approx 3H_{\text{kick}}$. We conclude that $t_{\text{kick}} \sim 3/(2m_{\phi, \text{UV}})$ such that

$$\frac{\Delta t_{\text{kick}}}{t_{\text{kick}}} \sim -\frac{\Delta m_{\phi, \text{UV}}}{m_{\phi, \text{UV}}} \sim -\frac{\Delta\phi_{\text{kick}}}{\phi_{\text{kick}}}. \quad (\text{H.15})$$

The complete variation is then

$$\Delta\theta_* \approx \theta_* \left[(n-2) \frac{\Delta\phi_{\text{kick}}}{\phi_{\text{kick}}} + n \cot(n\theta_{\text{kick}}) \frac{\Delta\theta_{\text{kick}}}{\theta_{\text{kick}}} \right]. \quad (\text{H.16})$$

Evaluating θ_* with $\dot{\theta}_{\text{kick}} t_{\text{kick}} \sim 1$ and $\phi_{\text{kin}} \sim f_a$ and applying the inflationary fluctuations $\Delta\theta_{\text{kick}} \approx H_I/(2\pi)$ and $\Delta\theta_{\text{kick}} \approx H_I/(2\pi\phi_{\text{kick}})$ we find that

$$\Delta\theta_* \approx \frac{3}{\pi} (n-2) \frac{H_I}{f_a} + \frac{3n \cot(n\theta_{\text{kick}})}{\pi} \frac{H_I}{\theta_{\text{kick}}} \frac{H_I}{f_a}, \quad (\text{H.17})$$

which must satisfy $\Delta\theta_* < 1$ to avoid domain wall formation. The exact constraint depends on both the choice of n and potentially on the initial θ_{kick} , but if the cot term is not tuned to a large value, we generically expect a constraint of

$$\frac{H_I}{f_a} \lesssim \mathcal{O}(0.1) \quad (\text{H.18})$$

If this condition is violated, then large angular fluctuations would lead to the formation of domain walls at the time of trapping. Since we are considering a scenario with PQ breaking during inflation, we do not expect cosmic strings to be present, and the domain walls would therefore be stable and pathologic to the cosmology.

Co et al. [33] propose to resolve this domain wall problem by invoking parametric resonance to non-thermally restore PQ-symmetry, generate cosmic strings and thereby render the string-wall network unstable. However, in the absence of a more detailed analysis, we do not consider this solution.

H.3 Problems with Yukawa damping in the quartic model

In the case of the nearly-quadratic model, we provided a damping mechanism in the form of a fermion Yukawa coupling as would be realized, e.g., in a KSVZ-type axion model. However, this mechanism does not lead to viable damping in the quartic model. We in this section explore why.

To estimate the efficiency of a damping mechanism, the relevant quantity is the largest value of Γ/H which can be reached within the limits of constraints. Therefore, we in the following investigate how Γ/H can be maximized under various constraints and show that damping cannot be efficient. We demonstrate this in two classes of scenarios where damping takes place in relativistic and non-relativistic phases, respectively.

Relativistic scenario: If we introduce a Yukawa coupling of the form $y\phi\bar{\chi}\chi$, as was used in section 13 and was described in detail in appendix D, then the most efficient damping rate is scattering with relativistic fermions. We, therefore, first consider the scenario in which damping takes place with a damping rate of

$$\Gamma_{\chi\text{-scat}} \approx y^2 \alpha T \quad (\text{H.19})$$

We have assumed the universe to be radiation dominated such that $H \propto T^2$. Therefore, $\Gamma_{\chi\text{-scat}}/H$ grows linearly with a . The rate is valid until the fermions become non-relativistic, which for a near maximal value of y takes place shortly after T_{kin} where the radial mode relaxes to the minimum at f_a . The relevant quantity is, therefore,

$$\frac{\Gamma_{\chi\text{scat}}(T_{\text{kick}})}{H(T_{\text{kick}})} \frac{T_{\text{kick}}}{T_{\text{kin}}} \approx 0.6 \frac{\alpha y^2}{\lambda^{3/4}} \frac{\sqrt{H_I m_P}}{f_a}. \quad (\text{H.20})$$

To achieve efficient damping, we then seek to maximize the Yukawa y . One upper bound on y results from the assumption that the fermions are relativistic at the kick, i.e. that $y\phi_{\text{kick}} < T_{\text{kick}}$. Another upper bound on y exists from the requirement that ϕ is dominated by its zero-temperature mass at the kick, i.e. that $\lambda\phi_{\text{kick}} > yT_{\text{kick}}$. Both upper bounds on y are maximized for

$$H_I \approx 0.2 \times m_P \sqrt{\lambda}, \quad (\text{H.21})$$

for which y can take the maximal value of

$$y \lesssim \sqrt{\lambda}. \quad (\text{H.22})$$

With these optimal choices for H_I and y , the damping rate becomes

$$\frac{\Gamma_{\chi\text{scat}}(T_{\text{kin}})}{H(T_{\text{kin}})} \approx \frac{\Gamma_{\chi\text{scat}}(T_{\text{kick}})}{H(T_{\text{kick}})} \frac{T_{\text{kick}}}{T_{\text{kin}}} \Bigg|_{\max y} \approx 0.3 \times \alpha \frac{m_P \sqrt{\lambda}}{f_a} \quad (\text{H.23})$$

We observe that this upper bound on Γ/H depends on λ to a positive power. The quartic coupling λ must then also be maximized. λ is ultimately limited from above by the DM isocurvature constraint which demanded $\lambda < 10^{-11}$. However, such low values of the quartic coupling limits m_ϕ , which makes it difficult for ϕ to decay before BBN¹⁸. Therefore, the radial mode must instead be sufficiently light to escape N_{eff} constraints by remaining relativistic until CMB temperatures. This requires $m_{\phi,\text{late}} \lesssim \text{eV}$, which corresponds to the upper bound

$$\lambda \lesssim 10^{-19} \left(\frac{10^{10} \text{ GeV}}{f_a} \right). \quad (\text{H.24})$$

Evidently, this upper bound is much stronger than the bound from isocurvature fluctuations. Finally, the χ -fermions are assumed to couple efficiently with the visible sector. Therefore they have to be heavy enough not to be produced at colliders. We, therefore, impose that $m_{\chi,\text{today}} \gtrsim \text{TeV}$, which with the choices above requires

$$f_a \gtrsim 10^{15} \text{ GeV}. \quad (\text{H.25})$$

As is evident from fig. 11, $f_a \gtrsim 10^{15} \text{ GeV}$ is outside the KMM regime as only standard misalignment would be possible with decay constants of that magnitude.

¹⁸For decay constants lower than $f_a \lesssim \mathcal{O}(1) \times 10^7 \text{ GeV}$ it may be possible for the saxion to decay before BBN. However, such scenarios are ruled out by collider constraints because the saxion mass is limited to $m_{\phi,\text{today}} \lesssim \mathcal{O}(10 \text{ GeV})$.

Further, even if we ignore that values of f_a outside the KMM regime are required to be compatible with constraints, the damping is still not efficient. The optimal conditions for damping through χ -scattering are found by saturating the bounds given by eqs. (H.22), (H.24) and (H.25) as well as assuming the optimal value of H_I given by eq. (H.21). However, even under these optimal conditions, the damping rate is still only

$$\left. \frac{\Gamma_{\chi\text{-scat.}}}{H} \right|_{\text{max}} \approx 10^{-11}. \quad (\text{H.26})$$

Any other choice will only lead to the less efficient damping, wherefore we conclude scattering processes with fully relativistic fermions is not viable.

After T_{kin} the χ -fermions will quickly become dominated by the ϕ -mass. In the intermediate regime $\alpha T < y\phi < T$ where the fermions become affected by the ϕ -contribution to their mass but are still (barely) relativistic, the interaction rate is momentarily boosted. During this period, Γ/H grows with T^2 . However, maximizing y also limits the length of this period because the χ -fermions are only just light enough to be relativistic. This tension keeps the intermediate period $\alpha T < y\phi < T$ from significantly altering the conclusions above.

Non-relativistic scenario: Although the damping rate in the regime with non-relativistic χ -fermions is in general much lower than the rate in the relativistic scenario one may wonder if there is nonetheless a possibility to damp with non-relativistic fermions. In this case, the most efficient damping mechanism is the loop-induced scattering with a damping rate of

$$\Gamma_{\gamma\text{-scat.}} \approx \frac{b\alpha^2 T^3}{\phi^2}. \quad (\text{H.27})$$

Interestingly, this rate is independent of y , wherefore the exact value of y does not play a role in the damping. This rate drops faster than the Hubble rate after the radial mode becomes relaxed to f_a , wherefore the largest ratio of $\Gamma_{\gamma\text{-scat.}}/H$ is reached at T_{kin} . At T_{kin} the ratio is

$$\frac{\Gamma_{\gamma\text{-scat.}}(T_{\text{kin}})}{H(T_{\text{kin}})} \approx 0.1b\alpha^2 \frac{m_P^{3/2} \lambda^{3/4}}{f_a \sqrt{H_I}} \quad (\text{H.28})$$

With the largest possible value of H_I (limited by the need for the zero-temperature mass to dominate the thermal-log mass at T_{kick}) and the largest possible λ (limited by the thermal relic constraint), the maximal rate is

$$\left. \frac{\Gamma_{\gamma\text{-scat.}}(T_{\text{kin}})}{H(T_{\text{kin}})} \right|_{\text{max}} \approx \frac{7 \times 10^5 \text{ GeV}}{f_a} \quad (\text{H.29})$$

To achieve efficient damping, the decay constant would therefore have to be less than

$$f_a \lesssim 10^4 \text{ GeV}, \quad (\text{H.30})$$

which is strongly constrained by astrophysics and helioscopes. Thus, damping with non-relativistic χ -fermions is also not viable.

Mixed scenarios: An attractive solution in the nearly-quadratic model is the scenario in which the radial mode avoids thermal corrections at T_{kick} by having the χ -fermions be non-relativistic at T_{kick} and in which damping is achieved through efficient scattering in an intermediate phase of relativistic fermions. This is possible in the nearly-quadratic model because ϕ redshifts faster than T in those models. However, in the quartic model ϕ never redshifts faster than T (neglecting reheating), wherefore the χ -fermions will remain non-relativistic if first they become non-relativistic. This absence of intermediate relativistic phases rules out such mixed scenarios in quartic models.

Conclusion - Problems with Yukawa damping in the quartic model: We find that a Yukawa interaction cannot be used to damp the radial oscillations in the quartic model. This conclusion results from the tension between the various constraints discussed above (isocurvature, thermal relic, collider, scenario consistency, efficient damping) rather than from any individual constraint. Another type of interaction is needed for the quartic scenario to be viable. In the previous paragraphs, we motivated this conclusion by optimizing Γ/H in analytic approximations of both relativistic and non-relativistic scenarios. This conclusion is confirmed by numerical solutions of the Boltzmann equation described in appendix F, which affirm that no viable scenario exists given the constraints imposed.

References

- [1] C. Eröncel, R. Sato, G. Servant, and P. Sørensen, *ALP Dark Matter from Kinetic Fragmentation: Opening up the Parameter Window*, [arXiv:2206.14259](#).
- [2] C. Eröncel, R. Sato, G. Servant, and P. Sørensen, *Model implementations of axion kinetic fragmentation*, (2023) to appear.
- [3] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440–1443.
- [4] R. D. Peccei and H. R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791–1797.
- [5] J. Preskill, M. B. Wise, and F. Wilczek, *Cosmology of the Invisible Axion*, *Phys. Lett. B* **120** (1983) 127–132.
- [6] L. F. Abbott and P. Sikivie, *A Cosmological Bound on the Invisible Axion*, *Phys. Lett. B* **120** (1983) 133–136.
- [7] M. Dine and W. Fischler, *The Not So Harmless Axion*, *Phys. Lett. B* **120** (1983) 137–141.
- [8] J. E. Kim, *Light Pseudoscalars, Particle Physics and Cosmology*, *Phys. Rept.* **150** (1987) 1–177.
- [9] P. Svrcek and E. Witten, *Axions In String Theory*, *JHEP* **06** (2006) 051, [[hep-th/0605206](#)].
- [10] G. B. Gelmini and M. Roncadelli, *Left-Handed Neutrino Mass Scale and Spontaneously Broken Lepton Number*, *Phys. Lett. B* **99** (1981) 411–415.
- [11] B. Bellazzini, A. Mariotti, D. Redigolo, F. Sala, and J. Serra, *R-axion at colliders*, *Phys. Rev. Lett.* **119** (2017), no. 14 141804, [[arXiv:1702.02152](#)].
- [12] Y. Ema, K. Hamaguchi, T. Moroi, and K. Nakayama, *Flaxion: a minimal extension to solve puzzles in the standard model*, *JHEP* **01** (2017) 096, [[arXiv:1612.05492](#)].
- [13] L. Calibbi, F. Goertz, D. Redigolo, R. Ziegler, and J. Zupan, *Minimal axion model from flavor*, *Phys. Rev. D* **95** (2017), no. 9 095009, [[arXiv:1612.08040](#)].
- [14] G. Grilli di Cortona, E. Hardy, J. Pardo Vega, and G. Villadoro, *The QCD axion, precisely*, *JHEP* **01** (2016) 034, [[arXiv:1511.02867](#)].
- [15] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223–226.
- [16] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279–282.
- [17] R. L. Davis, *Goldstone Bosons in String Models of Galaxy Formation*, *Phys. Rev. D* **32** (1985) 3172.
- [18] R. L. Davis, *Cosmic Axions from Cosmic Strings*, *Phys. Lett. B* **180** (1986) 225–230.
- [19] D. Harari and P. Sikivie, *On the Evolution of Global Strings in the Early Universe*, *Phys. Lett. B* **195** (1987) 361–365.
- [20] A. Vaquero, J. Redondo, and J. Stadler, *Early seeds of axion miniclusters*, *JCAP* **04** (2019) 012, [[arXiv:1809.09241](#)].
- [21] M. Buschmann, J. W. Foster, and B. R. Safdi, *Early-Universe Simulations of the Cosmological Axion*, *Phys. Rev. Lett.* **124** (2020), no. 16 161103, [[arXiv:1906.00967](#)].
- [22] M. Hindmarsh, J. Lizarraga, A. Lopez-Eiguren, and J. Urrestilla, *Scaling Density of Axion Strings*, *Phys. Rev. Lett.* **124** (2020), no. 2 021301, [[arXiv:1908.03522](#)].
- [23] M. Gorghetto, E. Hardy, and G. Villadoro, *More axions from strings*, *SciPost Phys.* **10** (2021), no. 2 050, [[arXiv:2007.04990](#)].
- [24] M. Dine, N. Fernandez, A. Ghalsasi, and H. H. Patel, *Comments on axions, domain walls, and cosmic strings*, *JCAP* **11** (2021) 041, [[arXiv:2012.13065](#)].

- [25] M. Buschmann, J. W. Foster, A. Hook, A. Peterson, D. E. Willcox, W. Zhang, and B. R. Safdi, *Dark matter from axion strings with adaptive mesh refinement*, *Nature Commun.* **13** (2022), no. 1 1049, [[arXiv:2108.05368](#)].
- [26] M. Hindmarsh, J. Lizarraga, A. Lopez-Eiguren, and J. Urrestilla, *Comment on "More Axions from Strings"*, [arXiv:2109.09679](#).
- [27] R. T. Co, L. J. Hall, and K. Harigaya, *Axion Kinetic Misalignment Mechanism*, *Phys. Rev. Lett.* **124** (2020), no. 25 251802, [[arXiv:1910.14152](#)].
- [28] C.-F. Chang and Y. Cui, *New Perspectives on Axion Misalignment Mechanism*, *Phys. Rev. D* **102** (2020), no. 1 015003, [[arXiv:1911.11885](#)].
- [29] N. Fonseca, E. Morgante, R. Sato, and G. Servant, *Axion fragmentation*, *JHEP* **04** (2020) 010, [[arXiv:1911.08472](#)].
- [30] C. Eröncel and G. Servant, *ALP Dark Matter Mini-Clusters from Kinetic Fragmentation*, [arXiv:2207.10111](#).
- [31] I. Affleck and M. Dine, *A New Mechanism for Baryogenesis*, *Nucl. Phys. B* **249** (1985) 361–380.
- [32] M. Dine, L. Randall, and S. D. Thomas, *Baryogenesis from flat directions of the supersymmetric standard model*, *Nucl. Phys. B* **458** (1996) 291–326, [[hep-ph/9507453](#)].
- [33] R. T. Co, L. J. Hall, K. Harigaya, K. A. Olive, and S. Verner, *Axion Kinetic Misalignment and Parametric Resonance from Inflation*, *JCAP* **08** (2020) 036, [[arXiv:2004.00629](#)].
- [34] R. T. Co, N. Fernandez, A. Ghalsasi, L. J. Hall, and K. Harigaya, *Lepto-Axiogenesis*, [arXiv:2006.05687](#).
- [35] R. T. Co, D. Dunsky, N. Fernandez, A. Ghalsasi, L. J. Hall, K. Harigaya, and J. Shelton, *Gravitational Wave and CMB Probes of Axion Kination*, [arXiv:2108.09299](#).
- [36] Y. Gouttenoire, G. Servant, and P. Simakachorn, *Revealing the Primordial Irreducible Inflationary Gravitational-Wave Background with a Spinning Peccei-Quinn Axion*, [arXiv:2108.10328](#).
- [37] Y. Gouttenoire, G. Servant, and P. Simakachorn, *Kination cosmology from scalar fields and gravitational-wave signatures*, [arXiv:2111.01150](#).
- [38] Y. Nambu, *Axial vector current conservation in weak interactions*, *Phys. Rev. Lett.* **4** (1960) 380–382.
- [39] J. Goldstone, *Field Theories with Superconductor Solutions*, *Nuovo Cim.* **19** (1961) 154–164.
- [40] J. Goldstone, A. Salam, and S. Weinberg, *Broken Symmetries*, *Phys. Rev.* **127** (1962) 965–970.
- [41] S. Weinberg, *The U(1) Problem*, *Phys. Rev. D* **11** (1975) 3583–3593.
- [42] L. Di Luzio, M. Giannotti, E. Nardi, and L. Visinelli, *The landscape of QCD axion models*, *Phys. Rept.* **870** (2020) 1–117, [[arXiv:2003.01100](#)].
- [43] G. 't Hooft, *Symmetry Breaking Through Bell-Jackiw Anomalies*, *Phys. Rev. Lett.* **37** (1976) 8–11.
- [44] G. 't Hooft, *Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle*, *Phys. Rev. D* **14** (1976) 3432–3450. [Erratum: *Phys.Rev.D* **18**, 2199 (1978)].
- [45] G. 't Hooft, *How Instantons Solve the U(1) Problem*, *Phys. Rept.* **142** (1986) 357–387.
- [46] S. L. Adler, *Axial vector vertex in spinor electrodynamics*, *Phys. Rev.* **177** (1969) 2426–2438.
- [47] J. S. Bell and R. Jackiw, *A PCAC puzzle: $\pi^0 \rightarrow \gamma\gamma$ in the σ model*, *Nuovo Cim. A* **60** (1969) 47–61.
- [48] W. A. Bardeen, *Anomalous Ward identities in spinor field theories*, *Phys. Rev.* **184** (1969) 1848–1857.
- [49] S. Coleman, *Aspects of Symmetry: Selected Erice Lectures*. Cambridge University Press, Cambridge, U.K., 1985.

- [50] R. Bott, *An Application of Morse theory to the topology of Lie groups*, *Bull. Soc. Math. Fr.* **84** (1956) 251–281.
- [51] E. B. Bogomolny, *Stability of Classical Solutions*, *Sov. J. Nucl. Phys.* **24** (1976) 449.
- [52] K. M. Bitar and S.-J. Chang, *Vacuum Tunneling of Gauge Theory in Minkowski Space*, *Phys. Rev. D* **17** (1978) 486.
- [53] C. Abel et al., *Measurement of the Permanent Electric Dipole Moment of the Neutron*, *Phys. Rev. Lett.* **124** (2020), no. 8 081803, [[arXiv:2001.11966](#)].
- [54] M. Pospelov and A. Ritz, *Electric dipole moments as probes of new physics*, *Annals Phys.* **318** (2005) 119–169, [[hep-ph/0504231](#)].
- [55] J. R. Ellis and M. K. Gaillard, *Strong and Weak CP Violation*, *Nucl. Phys. B* **150** (1979) 141–162.
- [56] M. Dine and P. Draper, *Challenges for the Nelson-Barr Mechanism*, *JHEP* **08** (2015) 132, [[arXiv:1506.05433](#)].
- [57] L. Ubaldi, *Effects of theta on the deuteron binding energy and the triple-alpha process*, *Phys. Rev. D* **81** (2010) 025011, [[arXiv:0811.1599](#)].
- [58] F. Takahashi, *A possible solution to the strong CP problem*, *Prog. Theor. Phys.* **121** (2009) 711–725, [[arXiv:0804.2478](#)].
- [59] N. Kaloper and J. Terning, *Landscaping the Strong CP Problem*, *JHEP* **03** (2019) 032, [[arXiv:1710.01740](#)].
- [60] C. Vafa and E. Witten, *Restrictions on Symmetry Breaking in Vector-Like Gauge Theories*, *Nucl. Phys. B* **234** (1984) 173–188.
- [61] D. J. Gross, R. D. Pisarski, and L. G. Yaffe, *QCD and Instantons at Finite Temperature*, *Rev. Mod. Phys.* **53** (1981) 43.
- [62] O. Wantz and E. P. S. Shellard, *Axion Cosmology Revisited*, *Phys. Rev. D* **82** (2010) 123508, [[arXiv:0910.1066](#)].
- [63] E. Berkowitz, M. I. Buchoff, and E. Rinaldi, *Lattice QCD input for axion cosmology*, *Phys. Rev. D* **92** (2015), no. 3 034507, [[arXiv:1505.07455](#)].
- [64] E. Berkowitz, *Lattice QCD and Axion Cosmology*, *PoS LATTICE2015* (2016) 236, [[arXiv:1509.02976](#)].
- [65] R. Kitano and N. Yamada, *Topology in QCD and the axion abundance*, *JHEP* **10** (2015) 136, [[arXiv:1506.00370](#)].
- [66] S. Borsanyi, M. Dierigl, Z. Fodor, S. D. Katz, S. W. Mages, D. Nogradi, J. Redondo, A. Ringwald, and K. K. Szabo, *Axion cosmology, lattice QCD and the dilute instanton gas*, *Phys. Lett. B* **752** (2016) 175–181, [[arXiv:1508.06917](#)].
- [67] S. Borsanyi et al., *Calculation of the axion mass based on high-temperature lattice quantum chromodynamics*, *Nature* **539** (2016), no. 7627 69–71, [[arXiv:1606.07494](#)].
- [68] J. E. Kim, *Weak Interaction Singlet and Strong CP Invariance*, *Phys. Rev. Lett.* **43** (1979) 103.
- [69] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Can Confinement Ensure Natural CP Invariance of Strong Interactions?*, *Nucl. Phys. B* **166** (1980) 493–506.
- [70] W. A. Bardeen, R. D. Peccei, and T. Yanagida, *CONSTRAINTS ON VARIANT AXION MODELS*, *Nucl. Phys. B* **279** (1987) 401–428.
- [71] A. R. Zhitnitsky, *On Possible Suppression of the Axion Hadron Interactions. (In Russian)*, *Sov. J. Nucl. Phys.* **31** (1980) 260.
- [72] M. Dine, W. Fischler, and M. Srednicki, *A Simple Solution to the Strong CP Problem with a Harmless Axion*, *Phys. Lett. B* **104** (1981) 199–202.

- [73] J. E. Kim, *A COMPOSITE INVISIBLE AXION*, *Phys. Rev. D* **31** (1985) 1733.
- [74] L. Randall, *Composite axion models and Planck scale physics*, *Phys. Lett. B* **284** (1992) 77–80.
- [75] M. Redi and R. Sato, *Composite Accidental Axions*, *JHEP* **05** (2016) 104, [[arXiv:1602.05427](https://arxiv.org/abs/1602.05427)].
- [76] P. Arias, D. Cadamuro, M. Goodsell, J. Jaeckel, J. Redondo, and A. Ringwald, *WISPy Cold Dark Matter*, *JCAP* **06** (2012) 013, [[arXiv:1201.5902](https://arxiv.org/abs/1201.5902)].
- [77] H. Jeffreys, *On certain approximate solutions of lineae differential equations of the second order**, *Proceedings of the London Mathematical Society* **s2-23** (1925), no. 1 428–436, [<https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/plms/s2-23.1.428>].
- [78] G. Wentzel, *Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik*, *Z. Phys.* **38** (1926), no. 6 518–529.
- [79] H. A. Kramers, *Wellenmechanik und halbzahlige Quantisierung*, *Z. Phys.* **39** (1926), no. 10 828–840.
- [80] L. Brillouin, *La mécanique ondulatoire de Schrödinger; une méthode générale de resolution par approximations successives*, *Compt. Rend. Hebd. Seances Acad. Sci.* **183** (1926), no. 1 24–26.
- [81] **Particle Data Group** Collaboration, R. Workman et al., *Review of Particle Physics*, . to be published (2022).
- [82] I. G. Irastorza and J. Redondo, *New experimental approaches in the search for axion-like particles*, *Prog. Part. Nucl. Phys.* **102** (2018) 89–159, [[arXiv:1801.08127](https://arxiv.org/abs/1801.08127)].
- [83] C. O’Hare, “cajohare/axionlimits: Axionlimits.” <https://cajohare.github.io/AxionLimits/>, July, 2020.
- [84] P. Sikivie, *Experimental Tests of the Invisible Axion*, *Phys. Rev. Lett.* **51** (1983) 1415–1417. [Erratum: *Phys.Rev.Lett.* 52, 695 (1984)].
- [85] **ADMX** Collaboration, S. J. Asztalos et al., *A SQUID-based microwave cavity search for dark-matter axions*, *Phys. Rev. Lett.* **104** (2010) 041301, [[arXiv:0910.5914](https://arxiv.org/abs/0910.5914)].
- [86] **ADMX** Collaboration, N. Du et al., *A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment*, *Phys. Rev. Lett.* **120** (2018), no. 15 151301, [[arXiv:1804.05750](https://arxiv.org/abs/1804.05750)].
- [87] **ADMX** Collaboration, T. Braine et al., *Extended Search for the Invisible Axion with the Axion Dark Matter Experiment*, *Phys. Rev. Lett.* **124** (2020), no. 10 101303, [[arXiv:1910.08638](https://arxiv.org/abs/1910.08638)].
- [88] **ADMX** Collaboration, C. Boutan et al., *Piezoelectrically Tuned Multimode Cavity Search for Axion Dark Matter*, *Phys. Rev. Lett.* **121** (2018), no. 26 261302, [[arXiv:1901.00920](https://arxiv.org/abs/1901.00920)].
- [89] N. Crisosto, P. Sikivie, N. S. Sullivan, D. B. Tanner, J. Yang, and G. Rybka, *ADMX SLIC: Results from a Superconducting LC Circuit Investigating Cold Axions*, *Phys. Rev. Lett.* **124** (2020), no. 24 241101, [[arXiv:1911.05772](https://arxiv.org/abs/1911.05772)].
- [90] **ADMX** Collaboration, C. Bartram et al., *Search for Invisible Axion Dark Matter in the 3.3–4.2 μeV Mass Range*, *Phys. Rev. Lett.* **127** (2021), no. 26 261803, [[arXiv:2110.06096](https://arxiv.org/abs/2110.06096)].
- [91] C. Bartram et al., *Dark Matter Axion Search Using a Josephson Traveling Wave Parametric Amplifier*, [[arXiv:2110.10262](https://arxiv.org/abs/2110.10262)].
- [92] I. Stern, *ADMX Status*, *PoS ICHEP2016* (2016) 198, [[arXiv:1612.08296](https://arxiv.org/abs/1612.08296)].
- [93] S. Beurthey et al., *MADMAX Status Report*, [[arXiv:2003.10894](https://arxiv.org/abs/2003.10894)].
- [94] M. Baryakhtar, J. Huang, and R. Lasenby, *Axion and hidden photon dark matter detection with multilayer optical haloscopes*, *Phys. Rev. D* **98** (2018), no. 3 035006, [[arXiv:1803.11455](https://arxiv.org/abs/1803.11455)].
- [95] “Webpage: Broadband radiometric axion searches.” <https://www.physik.uni-hamburg.de/iexp/gruppe-horns/forschung/brass.html>.
- [96] **DMRadio** Collaboration, L. Brouwer et al., *DMRadio- m^3 : A Search for the QCD Axion Below 1 μeV* , [[arXiv:2204.13781](https://arxiv.org/abs/2204.13781)].

- [97] L. Brouwer et al., *Introducing DMRadio-GUT, a search for GUT-scale QCD axions*, [arXiv:2203.11246](#).
- [98] Z. Zhang, O. Ghosh, and D. Horns, *WISPLC: Search for Dark Matter with LC Circuit*, [arXiv:2111.04541](#).
- [99] **CAST** Collaboration, S. Andriamonje et al., *An Improved limit on the axion-photon coupling from the CAST experiment*, *JCAP* **04** (2007) 010, [[hep-ex/0702006](#)].
- [100] **CAST** Collaboration, V. Anastassopoulos et al., *New CAST Limit on the Axion-Photon Interaction*, *Nature Phys.* **13** (2017) 584–590, [[arXiv:1705.02290](#)].
- [101] I. Shilon, A. Dudarev, H. Silva, and H. H. J. ten Kate, *Conceptual Design of a New Large Superconducting Toroid for IAXO, the New International AXion Observatory*, *IEEE Trans. Appl. Supercond.* **23** (2013), no. 3 4500604, [[arXiv:1212.4633](#)].
- [102] E. Armengaud et al., *Conceptual Design of the International Axion Observatory (IAXO)*, *JINST* **9** (2014) T05002, [[arXiv:1401.3233](#)].
- [103] K. Ehret et al., *New ALPS Results on Hidden-Sector Lightweights*, *Phys. Lett. B* **689** (2010) 149–155, [[arXiv:1004.1313](#)].
- [104] M. Betz, F. Caspers, M. Gasior, M. Thumm, and S. W. Rieger, *First results of the CERN Resonant Weakly Interacting sub-eV Particle Search (CROWS)*, *Phys. Rev. D* **88** (2013), no. 7 075014, [[arXiv:1310.8098](#)].
- [105] **OSQAR** Collaboration, R. Ballou et al., *New exclusion limits on scalar and pseudoscalar axionlike particles from light shining through a wall*, *Phys. Rev. D* **92** (2015), no. 9 092002, [[arXiv:1506.08082](#)].
- [106] M. D. Ortiz et al., *Design of the ALPS II optical system*, *Phys. Dark Univ.* **35** (2022) 100968, [[arXiv:2009.14294](#)].
- [107] F. Della Valle, A. Ejlli, U. Gastaldi, G. Messineo, E. Milotti, R. Pengo, G. Ruoso, and G. Zavattini, *The PVLAS experiment: measuring vacuum magnetic birefringence and dichroism with a birefringent Fabry–Perot cavity*, *Eur. Phys. J. C* **76** (2016), no. 1 24, [[arXiv:1510.08052](#)].
- [108] **ARIADNE** Collaboration, A. A. Geraci et al., *Progress on the ARIADNE axion experiment*, *Springer Proc. Phys.* **211** (2018) 151–161, [[arXiv:1710.05413](#)].
- [109] K. K. Rogers and H. V. Peiris, *Strong Bound on Canonical Ultralight Axion Dark Matter from the Lyman-Alpha Forest*, *Phys. Rev. Lett.* **126** (2021), no. 7 071302, [[arXiv:2007.12705](#)].
- [110] D. Wouters and P. Brun, *Constraints on Axion-like Particles from X-Ray Observations of the Hydra Galaxy Cluster*, *Astrophys. J.* **772** (2013) 44, [[arXiv:1304.0989](#)].
- [111] M. C. D. Marsh, H. R. Russell, A. C. Fabian, B. P. McNamara, P. Nulsen, and C. S. Reynolds, *A New Bound on Axion-Like Particles*, *JCAP* **12** (2017) 036, [[arXiv:1703.07354](#)].
- [112] C. S. Reynolds, M. C. D. Marsh, H. R. Russell, A. C. Fabian, R. Smith, F. Tombesi, and S. Veilleux, *Astrophysical limits on very light axion-like particles from Chandra grating spectroscopy of NGC 1275*, *Astrophys. J.* **890** (2020) 59, [[arXiv:1907.05475](#)].
- [113] J. S. Reynés, J. H. Matthews, C. S. Reynolds, H. R. Russell, R. N. Smith, and M. C. D. Marsh, *New constraints on light axion-like particles using Chandra transmission grating spectroscopy of the powerful cluster-hosted quasar H1821+643*, *Mon. Not. Roy. Astron. Soc.* **510** (2021), no. 1 1264–1277, [[arXiv:2109.03261](#)].
- [114] A. Ayala, I. Domínguez, M. Giannotti, A. Mirizzi, and O. Straniero, *Revisiting the bound on axion-photon coupling from Globular Clusters*, *Phys. Rev. Lett.* **113** (2014), no. 19 191302, [[arXiv:1406.6053](#)].
- [115] D. Cadamuro and J. Redondo, *Cosmological bounds on pseudo Nambu-Goldstone bosons*, *JCAP* **02** (2012) 032, [[arXiv:1110.2895](#)].

- [116] C. P. Salemi et al., *Search for Low-Mass Axion Dark Matter with ABRACADABRA-10 cm*, *Phys. Rev. Lett.* **127** (2021), no. 8 081801, [[arXiv:2102.06722](#)].
- [117] J. A. Devlin et al., *Constraints on the Coupling between Axionlike Dark Matter and Photons Using an Antiproton Superconducting Tuned Detection Circuit in a Cryogenic Penning Trap*, *Phys. Rev. Lett.* **126** (2021), no. 4 041301, [[arXiv:2101.11290](#)].
- [118] S. Lee, S. Ahn, J. Choi, B. R. Ko, and Y. K. Semertzidis, *Axion Dark Matter Search around $6.7 \mu\text{eV}$* , *Phys. Rev. Lett.* **124** (2020), no. 10 101802, [[arXiv:2001.05102](#)].
- [119] J. Jeong, S. Youn, S. Bae, J. Kim, T. Seong, J. E. Kim, and Y. K. Semertzidis, *Search for Invisible Axion Dark Matter with a Multiple-Cell Haloscope*, *Phys. Rev. Lett.* **125** (2020), no. 22 221302, [[arXiv:2008.10141](#)].
- [120] **CAPP** Collaboration, O. Kwon et al., *First Results from an Axion Haloscope at CAPP around $10.7 \mu\text{eV}$* , *Phys. Rev. Lett.* **126** (2021), no. 19 191802, [[arXiv:2012.10764](#)].
- [121] **CAST** Collaboration, A. A. Melcón et al., *First results of the CAST-RADES haloscope search for axions at $34.67 \mu\text{eV}$* , *JHEP* **21** (2020) 075, [[arXiv:2104.13798](#)].
- [122] Y. Oshima, H. Fujimoto, M. Ando, T. Fujita, Y. Michimura, K. Nagano, I. Obata, and T. Watanabe, *Dark matter Axion search with riNg Cavity Experiment DANCE: Current sensitivity*, in *55th Rencontres de Moriond on Gravitation*, 5, 2021. [[arXiv:2105.06252](#)].
- [123] T. Grenet, R. Ballou, Q. Basto, K. Martineau, P. Perrier, P. Pignat, J. Quevillon, N. Roch, and C. Smith, *The Grenoble Axion Haloscope platform (GrAHal): development plan and first results*, [[arXiv:2110.14406](#)].
- [124] **HAYSTAC** Collaboration, L. Zhong et al., *Results from phase 1 of the HAYSTAC microwave cavity axion experiment*, *Phys. Rev. D* **97** (2018), no. 9 092001, [[arXiv:1803.03690](#)].
- [125] **HAYSTAC** Collaboration, K. M. Backes et al., *A quantum-enhanced search for dark matter axions*, *Nature* **590** (2021), no. 7845 238–242, [[arXiv:2008.01853](#)].
- [126] B. T. McAllister, G. Flower, J. Kruger, E. N. Ivanov, M. Goryachev, J. Bourhill, and M. E. Tobar, *The ORGAN Experiment: An axion haloscope above 15 GHz*, *Phys. Dark Univ.* **18** (2017) 67–72, [[arXiv:1706.00209](#)].
- [127] D. Alesini et al., *Galactic axions search with a superconducting resonant cavity*, *Phys. Rev. D* **99** (2019), no. 10 101101, [[arXiv:1903.06547](#)].
- [128] D. Alesini et al., *Search for invisible axion dark matter of mass $m_a = 43 \mu\text{eV}$ with the QUAX- $\alpha\gamma$ experiment*, *Phys. Rev. D* **103** (2021), no. 10 102004, [[arXiv:2012.09498](#)].
- [129] S. DePanfilis, A. C. Melissinos, B. E. Moskowitz, J. T. Rogers, Y. K. Semertzidis, W. U. Wuensch, H. J. Halama, A. G. Prodell, W. B. Fowler, and F. A. Nezrick, *Limits on the abundance and coupling of cosmic axions at $4.5 < m_a < 5.0 \mu\text{eV}$* , *Phys. Rev. Lett.* **59** (Aug, 1987) 839–842.
- [130] A. V. Gramolin, D. Aybas, D. Johnson, J. Adam, and A. O. Sushkov, *Search for axion-like dark matter with ferromagnets*, *Nature Phys.* **17** (2021), no. 1 79–84, [[arXiv:2003.03348](#)].
- [131] A. Arza, M. A. Fedderke, P. W. Graham, D. F. J. Kimball, and S. Kalia, *Earth as a transducer for axion dark-matter detection*, *Phys. Rev. D* **105** (2022), no. 9 095007, [[arXiv:2112.09620](#)].
- [132] C. Hagmann, P. Sikivie, N. S. Sullivan, and D. B. Tanner, *Results from a search for cosmic axions*, *Phys. Rev. D* **42** (Aug, 1990) 1297–1300.
- [133] C. A. Thomson, B. T. McAllister, M. Goryachev, E. N. Ivanov, and M. E. Tobar, *Upconversion Loop Oscillator Axion Detection Experiment: A Precision Frequency Interferometric Axion Dark Matter Search with a Cylindrical Microwave Cavity*, *Phys. Rev. Lett.* **126** (2021), no. 8 081803, [[arXiv:1912.07751](#)]. [Erratum: *Phys.Rev.Lett.* 127, 019901 (2021)].
- [134] H. Liu, B. D. Elwood, M. Evans, and J. Thaler, *Searching for Axion Dark Matter with Birefringent Cavities*, *Phys. Rev. D* **100** (2019), no. 2 023548, [[arXiv:1809.01656](#)].

- [135] K. Nagano, T. Fujita, Y. Michimura, and I. Obata, *Axion Dark Matter Search with Interferometric Gravitational Wave Detectors*, *Phys. Rev. Lett.* **123** (2019), no. 11 111301, [[arXiv:1903.02017](#)].
- [136] M. Lawson, A. J. Millar, M. Pancaldi, E. Vitagliano, and F. Wilczek, *Tunable axion plasma haloscopes*, *Phys. Rev. Lett.* **123** (2019), no. 14 141802, [[arXiv:1904.11872](#)].
- [137] **BREAD** Collaboration, J. Liu et al., *Broadband Solenoidal Haloscope for Terahertz Axion Detection*, *Phys. Rev. Lett.* **128** (2022), no. 13 131801, [[arXiv:2111.12103](#)].
- [138] Y. Michimura, Y. Oshima, T. Watanabe, T. Kawasaki, H. Takeda, M. Ando, K. Nagano, I. Obata, and T. Fujita, *DANCE: Dark matter Axion search with riNg Cavity Experiment*, *J. Phys. Conf. Ser.* **1468** (2020), no. 1 012032, [[arXiv:1911.05196](#)].
- [139] D. Alesini et al., *KLASH Conceptual Design Report*, [[arXiv:1911.02427](#)].
- [140] “Conference: Physics opportunities at 100-500 mhz haloscopes.” <https://indico.cern.ch/event/1115163/>.
- [141] A. Berlin, R. T. D’Agnolo, S. A. R. Ellis, and K. Zhou, *Heterodyne broadband detection of axion dark matter*, *Phys. Rev. D* **104** (2021), no. 11 L111701, [[arXiv:2007.15656](#)].
- [142] “Private communication with raffaele tito d’agnolo.” https://agenda.infn.it/event/20431/contributions/137687/attachments/82511/108428/Rettaroli_Patras2021_compressed.pdf.
- [143] “Conference presentation: Probing the axion-photon interaction with quax experiment: status and perspectives.” https://agenda.infn.it/event/20431/contributions/137687/attachments/82511/108428/Rettaroli_Patras2021_compressed.pdf.
- [144] D. J. E. Marsh, K.-C. Fong, E. W. Lentz, L. Smejkal, and M. N. Ali, *Proposal to Detect Dark Matter using Axionic Topological Antiferromagnets*, *Phys. Rev. Lett.* **123** (2019), no. 12 121601, [[arXiv:1807.08810](#)].
- [145] J. Schütte-Engel, D. J. E. Marsh, A. J. Millar, A. Sekine, F. Chadha-Day, S. Hoof, M. N. Ali, K.-C. Fong, E. Hardy, and L. Šmejkal, *Axion quasiparticles for axion dark matter detection*, *JCAP* **08** (2021) 066, [[arXiv:2102.05366](#)].
- [146] J. W. Foster, S. J. Witte, M. Lawson, T. Linden, V. Gajjar, C. Weniger, and B. R. Safdi, *Extraterrestrial Axion Search with the Breakthrough Listen Galactic Center Survey*, [[arXiv:2202.08274](#)].
- [147] M. H. Chan, *Constraining the axion–photon coupling using radio data of the Bullet cluster*, *Sci. Rep.* **11** (2021), no. 1 20087, [[arXiv:2109.11734](#)].
- [148] P. F. Depta, M. Hufnagel, and K. Schmidt-Hoberg, *Robust cosmological constraints on axion-like particles*, *JCAP* **05** (2020) 009, [[arXiv:2002.08370](#)].
- [149] C. Dessert, A. J. Long, and B. R. Safdi, *No Evidence for Axions from Chandra Observation of the Magnetic White Dwarf RE J0317-853*, *Phys. Rev. Lett.* **128** (2022), no. 7 071102, [[arXiv:2104.12772](#)].
- [150] B. Bolliet, J. Chluba, and R. Battye, *Spectral distortion constraints on photon injection from low-mass decaying particles*, *Mon. Not. Roy. Astron. Soc.* **507** (2021), no. 3 3148–3178, [[arXiv:2012.07292](#)].
- [151] M. A. Buen-Abad, J. Fan, and C. Sun, *Constraints on axions from cosmic distance measurements*, *JHEP* **02** (2022) 103, [[arXiv:2011.05993](#)].
- [152] F. Calore, P. Carenza, M. Giannotti, J. Jaeckel, and A. Mirizzi, *Bounds on axionlike particles from the diffuse supernova flux*, *Phys. Rev. D* **102** (2020), no. 12 123005, [[arXiv:2008.11741](#)].
- [153] F. Calore, P. Carenza, C. Eckner, T. Fischer, M. Giannotti, J. Jaeckel, K. Kotake, T. Kuroda, A. Mirizzi, and F. Sivo, *3D template-based Fermi-LAT constraints on the diffuse supernova axion-like particle background*, *Phys. Rev. D* **105** (2022), no. 6 063028, [[arXiv:2110.03679](#)].

- [154] M. Meyer and T. Petrushevska, *Search for Axionlike-Particle-Induced Prompt γ -Ray Emission from Extragalactic Core-Collapse Supernovae with the Fermi Large Area Telescope*, *Phys. Rev. Lett.* **124** (2020), no. 23 231101, [[arXiv:2006.06722](#)]. [Erratum: *Phys.Rev.Lett.* 125, 119901 (2020)].
- [155] **Fermi-LAT** Collaboration, M. Ajello et al., *Search for Spectral Irregularities due to Photon-Axionlike-Particle Oscillations with the Fermi Large Area Telescope*, *Phys. Rev. Lett.* **116** (2016), no. 16 161101, [[arXiv:1603.06978](#)].
- [156] B. D. Blout, E. J. Daw, M. P. Decowski, P. T. P. Ho, L. J. Rosenberg, and D. B. Yu, *A Radio telescope search for axions*, *Astrophys. J.* **546** (2001) 825–828, [[astro-ph/0006310](#)].
- [157] S. Jacobsen, T. Linden, and K. Freese, *Constraining Axion-Like Particles with HAWC Observations of TeV Blazars*, [arXiv:2203.04332](#).
- [158] **H.E.S.S.** Collaboration, A. Abramowski et al., *Constraints on axionlike particles with H.E.S.S. from the irregularity of the PKS 2155-304 energy spectrum*, *Phys. Rev. D* **88** (2013), no. 10 102003, [[arXiv:1311.3148](#)].
- [159] D. Wadekar and Z. Wang, *Strong constraints on decay and annihilation of dark matter from heating of gas-rich dwarf galaxies*, [arXiv:2111.08025](#).
- [160] C. Dessert, D. Dunskey, and B. R. Safdi, *Upper limit on the axion-photon coupling from magnetic white dwarf polarization*, *Phys. Rev. D* **105** (2022), no. 10 103034, [[arXiv:2203.04319](#)].
- [161] M. Regis, M. Taoso, D. Vaz, J. Brinchmann, S. L. Zoutendijk, N. F. Bouché, and M. Steinmetz, *Searching for light in the darkness: Bounds on ALP dark matter with the optical MUSE-faint survey*, *Phys. Lett. B* **814** (2021) 136075, [[arXiv:2009.01310](#)].
- [162] H.-J. Li, J.-G. Guo, X.-J. Bi, S.-J. Lin, and P.-F. Yin, *Limits on axion-like particles from Mrk 421 with 4.5-year period observations by ARGO-YBJ and Fermi-LAT*, *Phys. Rev. D* **103** (2021), no. 8 083003, [[arXiv:2008.09464](#)].
- [163] M. Xiao, K. M. Perez, M. Giannotti, O. Straniero, A. Mirizzi, B. W. Grefenstette, B. M. Roach, and M. Nynka, *Constraints on Axionlike Particles from a Hard X-Ray Observation of Betelgeuse*, *Phys. Rev. Lett.* **126** (2021), no. 3 031101, [[arXiv:2009.09059](#)].
- [164] C. Dessert, J. W. Foster, and B. R. Safdi, *X-ray Searches for Axions from Super Star Clusters*, *Phys. Rev. Lett.* **125** (2020), no. 26 261102, [[arXiv:2008.03305](#)].
- [165] N. Vinyoles, A. Serenelli, F. L. Villante, S. Basu, J. Redondo, and J. Isern, *New axion and hidden photon constraints from a solar data global fit*, *JCAP* **10** (2015) 015, [[arXiv:1501.01639](#)].
- [166] J. Jaeckel, P. C. Malta, and J. Redondo, *Decay photons from the axionlike particles burst of type II supernovae*, *Phys. Rev. D* **98** (2018), no. 5 055032, [[arXiv:1702.02964](#)].
- [167] A. Payez, C. Evoli, T. Fischer, M. Giannotti, A. Mirizzi, and A. Ringwald, *Revisiting the SN1987A gamma-ray limit on ultralight axion-like particles*, *JCAP* **02** (2015) 006, [[arXiv:1410.3747](#)].
- [168] A. Caputo, G. Raffelt, and E. Vitagliano, *Muonic boson limits: Supernova redux*, *Phys. Rev. D* **105** (2022), no. 3 035022, [[arXiv:2109.03244](#)].
- [169] J. S. Lee, *Revisiting Supernova 1987A Limits on Axion-Like-Particles*, [arXiv:1808.10136](#).
- [170] D. Grin, G. Covone, J.-P. Kneib, M. Kamionkowski, A. Blain, and E. Jullo, *A Telescope Search for Decaying Relic Axions*, *Phys. Rev. D* **75** (2007) 105018, [[astro-ph/0611502](#)].
- [171] M. J. Dolan, F. J. Hiskens, and R. R. Volkas, *Constraining axion-like particles using the white dwarf initial-final mass relation*, *JCAP* **09** (2021) 010, [[arXiv:2102.00379](#)].
- [172] J. W. Foster, M. Kongsore, C. Dessert, Y. Park, N. L. Rodd, K. Cranmer, and B. R. Safdi, *Deep Search for Decaying Dark Matter with XMM-Newton Blank-Sky Observations*, *Phys. Rev. Lett.* **127** (2021), no. 5 051101, [[arXiv:2102.02207](#)].

- [173] A. Dekker, E. Peerbooms, F. Zimmer, K. C. Y. Ng, and S. Ando, *Searches for sterile neutrinos and axionlike particles from the Galactic halo with eROSITA*, *Phys. Rev. D* **104** (2021), no. 2 023021, [[arXiv:2103.13241](#)].
- [174] M. Meyer, M. Giannotti, A. Mirizzi, J. Conrad, and M. A. Sánchez-Conde, *Fermi Large Area Telescope as a Galactic Supernovae Axionscope*, *Phys. Rev. Lett.* **118** (2017), no. 1 011103, [[arXiv:1609.02350](#)].
- [175] S.-F. Ge, K. Hamaguchi, K. Ichimura, K. Ishidoshiro, Y. Kanazawa, Y. Kishimoto, N. Nagata, and J. Zheng, *Supernova-scope for the Direct Search of Supernova Axions*, *JCAP* **11** (2020) 059, [[arXiv:2008.03924](#)].
- [176] C. Thorpe-Morgan, D. Malyshev, A. Santangelo, J. Jochum, B. Jäger, M. Sasaki, and S. Saeedi, *THESEUS insights into axionlike particles, dark photon, and sterile neutrino dark matter*, *Phys. Rev. D* **102** (2020), no. 12 123003, [[arXiv:2008.08306](#)].
- [177] D. F. Jackson Kimball et al., *Overview of the Cosmic Axion Spin Precession Experiment (CASPER)*, *Springer Proc. Phys.* **245** (2020) 105–121, [[arXiv:1711.08999](#)].
- [178] G. P. Centers et al., *Stochastic fluctuations of bosonic dark matter*, *Nature Commun.* **12** (2021), no. 1 7321, [[arXiv:1905.13650](#)].
- [179] T. Wu et al., *Search for Axionlike Dark Matter with a Liquid-State Nuclear Spin Comagnetometer*, *Phys. Rev. Lett.* **122** (2019), no. 19 191302, [[arXiv:1901.10843](#)].
- [180] A. Garcon et al., *Constraints on bosonic dark matter from ultralow-field nuclear magnetic resonance*, *Sci. Adv.* **5** (2019), no. 10 eaax4539, [[arXiv:1902.04644](#)].
- [181] **NASDUCK** Collaboration, I. M. Bloch, G. Ronen, R. Shaham, O. Katz, T. Volansky, and O. Katz, *New constraints on axion-like dark matter using a Floquet quantum detector*, *Sci. Adv.* **8** (2022), no. 5 ab18919, [[arXiv:2105.04603](#)].
- [182] C. Abel et al., *Search for Axionlike Dark Matter through Nuclear Spin Precession in Electric and Magnetic Fields*, *Phys. Rev. X* **7** (2017), no. 4 041034, [[arXiv:1708.06367](#)].
- [183] G. Vasilakis, J. M. Brown, T. W. Kornack, and M. V. Romalis, *Limits on new long range nuclear spin-dependent forces set with a K - He-3 co-magnetometer*, *Phys. Rev. Lett.* **103** (2009) 261801, [[arXiv:0809.4700](#)].
- [184] I. M. Bloch, Y. Hochberg, E. Kufflik, and T. Volansky, *Axion-like Relics: New Constraints from Old Comagnetometer Data*, *JHEP* **01** (2020) 167, [[arXiv:1907.03767](#)].
- [185] A. Bhusal, N. Houston, and T. Li, *Searching for Solar Axions Using Data from the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **126** (2021), no. 9 091601, [[arXiv:2004.02733](#)].
- [186] P. W. Graham, S. Hacıömeroğlu, D. E. Kaplan, Z. Omarov, S. Rajendran, and Y. K. Semertzidis, *Storage ring probes of dark matter and dark energy*, *Phys. Rev. D* **103** (2021), no. 5 055010, [[arXiv:2005.11867](#)].
- [187] M. Buschmann, C. Dessert, J. W. Foster, A. J. Long, and B. R. Safdi, *Upper Limit on the QCD Axion Mass from Isolated Neutron Star Cooling*, *Phys. Rev. Lett.* **128** (2022), no. 9 091102, [[arXiv:2111.09892](#)].
- [188] P. Carena, T. Fischer, M. Giannotti, G. Guo, G. Martínez-Pinedo, and A. Mirizzi, *Improved axion emissivity from a supernova via nucleon-nucleon bremsstrahlung*, *JCAP* **10** (2019), no. 10 016, [[arXiv:1906.11844](#)]. [Erratum: *JCAP* 05, E01 (2020)].
- [189] A. Arvanitaki, M. Baryakhtar, and X. Huang, *Discovering the QCD Axion with Black Holes and Gravitational Waves*, *Phys. Rev. D* **91** (2015), no. 8 084011, [[arXiv:1411.2263](#)].
- [190] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*. Pearson, San Francisco Munich, 3rd edition ed., June, 2001.

- [191] L. D. Landau and E. M. Lifshitz, *Mechanics: Volume 1*. Butterworth-Heinemann, Oxford New York, 3rd edition ed., Jan., 1976.
- [192] E. Masso, F. Rota, and G. Zsembinszki, *Scalar field oscillations contributing to dark energy*, *Phys. Rev. D* **72** (2005) 084007, [[astro-ph/0501381](#)].
- [193] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, *Nature Methods* **17** (2020) 261–272.
- [194] **Planck** Collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [[arXiv:1807.06209](#)].
- [195] W. Magnus and S. Winkler, *Hill's Equation*. John Wiley & Sons, New York, 1966.
- [196] P. B. Greene, L. Kofman, and A. A. Starobinsky, *Sine-Gordon parametric resonance*, *Nucl. Phys. B* **543** (1999) 423–443, [[hep-ph/9808477](#)].
- [197] U.-H. Zhang and T. Chiueh, *Cosmological Perturbations of Extreme Axion in the Radiation Era*, *Phys. Rev. D* **96** (2017), no. 6 063522, [[arXiv:1705.01439](#)].
- [198] A. Arvanitaki, S. Dimopoulos, M. Galanis, L. Lehner, J. O. Thompson, and K. Van Tilburg, *Large-misalignment mechanism for the formation of compact axion structures: Signatures from the QCD axion to fuzzy dark matter*, *Phys. Rev. D* **101** (2020), no. 8 083014, [[arXiv:1909.11665](#)].
- [199] D. J. E. Marsh, *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1–79, [[arXiv:1510.07633](#)].
- [200] A. D. Linde, *Axions in inflationary cosmology*, *Phys. Lett. B* **259** (1991) 38–47.
- [201] J. M. Bardeen, *Gauge Invariant Cosmological Perturbations*, *Phys. Rev. D* **22** (1980) 1882–1905.
- [202] C. Eröncel, R. Sato, G. Servant, and P. Simakachorn, *Work in progress*, (2022) to appear.
- [203] E. Morgante, W. Ratzinger, R. Sato, and B. A. Stefanek, *Axion fragmentation on the lattice*, *JHEP* **12** (2021) 037, [[arXiv:2109.13823](#)].
- [204] J. Berges, A. Chatrchyan, and J. Jaeckel, *Foamy Dark Matter from Monodromies*, *JCAP* **08** (2019) 020, [[arXiv:1903.03116](#)].
- [205] A. Arbey, J. Auffinger, K. P. Hickerson, and E. S. Jansson, *AlterBBN v2: A public code for calculating Big-Bang nucleosynthesis constraints in alternative cosmologies*, *Comput. Phys. Commun.* **248** (2020) 106982, [[arXiv:1806.11095](#)].
- [206] **Particle Data Group** Collaboration, P. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (2020), no. 8 083C01.
- [207] A. Chatrchyan and J. Jaeckel, *Gravitational waves from the fragmentation of axion-like particle dark matter*, *JCAP* **02** (2021) 003, [[arXiv:2004.07844](#)].
- [208] E. Madge, W. Ratzinger, D. Schmitt, and P. Schwaller, *Audible Axions with a Booster: Stochastic Gravitational Waves from Rotating ALPs*, *SciPost Phys.* **12** (2022) 171, [[arXiv:2111.12730](#)].
- [209] T. Kite, A. Ravenni, S. P. Patil, and J. Chluba, *Bridging the gap: spectral distortions meet gravitational waves*, *Mon. Not. Roy. Astron. Soc.* **505** (2021), no. 3 4396–4405, [[arXiv:2010.00040](#)].
- [210] J. F. Dufaux, A. Bergman, G. N. Felder, L. Kofman, and J.-P. Uzan, *Theory and Numerics of Gravitational Waves from Preheating after Inflation*, *Phys. Rev. D* **76** (2007) 123517, [[arXiv:0707.0875](#)].
- [211] G. Janssen et al., *Gravitational wave astronomy with the SKA*, *PoS AASKA14* (2015) 037, [[arXiv:1501.00127](#)].

- [212] K. Schmitz, *New Sensitivity Curves for Gravitational-Wave Signals from Cosmological Phase Transitions*, *JHEP* **01** (2021) 097, [[arXiv:2002.04615](#)].
- [213] J. C. Mather, E. S. Cheng, D. A. Cottingham, J. Eplee, R. E., D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jensen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, J. Moseley, S. H., R. F. Silverberg, G. F. Smoot, R. Weiss, and D. T. Wilkinson, *Measurement of the Cosmic Microwave Background Spectrum by the COBE FIRAS Instrument*, *Astrophysical Journal* **420** (Jan., 1994) 439.
- [214] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, *The Cosmic Microwave Background spectrum from the full COBE FIRAS data set*, *Astrophys. J.* **473** (1996) 576, [[astro-ph/9605054](#)].
- [215] A. Kogut, D. J. Fixsen, D. T. Chuss, J. Dotson, E. Dwek, M. Halpern, G. F. Hinshaw, S. M. Meyer, S. H. Moseley, M. D. Seiffert, D. N. Spergel, and E. J. Wollack, *The Primordial Inflation Explorer (PIXIE): a nulling polarimeter for cosmic microwave background observations*, *JCAP* **2011** (July, 2011) 025, [[arXiv:1105.2044](#)].
- [216] A. Kogut, M. H. Abitbol, J. Chluba, J. Delabrouille, D. Fixsen, J. Hill, S. Patil, and A. Rotti, *Cmb spectral distortions: Status and prospects*, *Bulletin of the AAS* **51** (9, 2019).
<https://baas.aas.org/pub/2020n7i113>.
- [217] J. Chluba et al., *Spectral Distortions of the CMB as a Probe of Inflation, Recombination, Structure Formation and Particle Physics: Astro2020 Science White Paper*, *Bull. Am. Astron. Soc.* **51** (2019), no. 3 184, [[arXiv:1903.04218](#)].
- [218] N. Bar, D. Blas, K. Blum, and S. Sibiryakov, *Galactic rotation curves versus ultralight dark matter: Implications of the soliton-host halo relation*, *Phys. Rev. D* **98** (2018), no. 8 083027, [[arXiv:1805.00122](#)].
- [219] N. Bar, K. Blum, J. Eby, and R. Sato, *Ultralight dark matter in disk galaxies*, *Phys. Rev. D* **99** (2019), no. 10 103020, [[arXiv:1903.03402](#)].
- [220] N. Bar, K. Blum, and C. Sun, *Galactic rotation curves versus ultralight dark matter: A systematic comparison with SPARC data*, *Phys. Rev. D* **105** (2022), no. 8 083015, [[arXiv:2111.03070](#)].
- [221] K.-H. Leong, H.-Y. Schive, U.-H. Zhang, and T. Chiueh, *Testing extreme-axion wave-like dark matter using the BOSS Lyman-alpha forest data*, *Mon. Not. Roy. Astron. Soc.* **484** (2019), no. 3 4273–4286, [[arXiv:1810.05930](#)].
- [222] R. T. Co and K. Harigaya, *Axiogenesis*, *Phys. Rev. Lett.* **124** (2020), no. 11 111602, [[arXiv:1910.02080](#)].
- [223] R. T. Co, L. J. Hall, and K. Harigaya, *Predictions for Axion Couplings from ALP Cogenesis*, *JHEP* **01** (2021) 172, [[arXiv:2006.04809](#)].
- [224] M. Dine, L. Randall, and S. D. Thomas, *Supersymmetry breaking in the early universe*, *Phys. Rev. Lett.* **75** (1995) 398–401, [[hep-ph/9503303](#)].
- [225] L. Di Luzio, B. Gavela, P. Quilez, and A. Ringwald, *An even lighter QCD axion*, *JHEP* **05** (2021) 184, [[arXiv:2102.00012](#)].
- [226] L. Di Luzio, B. Gavela, P. Quilez, and A. Ringwald, *Dark matter from an even lighter QCD axion: trapped misalignment*, [[arXiv:2102.01082](#)].
- [227] D. Antypas et al., *New Horizons: Scalar and Vector Ultralight Dark Matter*, [[arXiv:2203.14915](#)].
- [228] T. Damour and J. F. Donoghue, *Equivalence Principle Violations and Couplings of a Light Dilaton*, *Phys. Rev. D* **82** (2010) 084033, [[arXiv:1007.2792](#)].
- [229] S. Schlamminger, K. Y. Choi, T. A. Wagner, J. H. Gundlach, and E. G. Adelberger, *Test of the equivalence principle using a rotating torsion balance*, *Phys. Rev. Lett.* **100** (2008) 041101, [[arXiv:0712.0607](#)].

- [230] J. G. Williams, S. G. Turyshev, and D. H. Boggs, *Progress in lunar laser ranging tests of relativistic gravity*, *Phys. Rev. Lett.* **93** (2004) 261101, [[gr-qc/0411113](#)].
- [231] A. Hees, O. Minazzoli, E. Savalle, Y. V. Stadnik, and P. Wolf, *Violation of the equivalence principle from light scalar dark matter*, *Phys. Rev. D* **98** (2018), no. 6 064051, [[arXiv:1807.04512](#)].
- [232] G. L. Smith, C. D. Hoyle, J. H. Gundlach, E. G. Adelberger, B. R. Heckel, and H. E. Swanson, *Short range tests of the equivalence principle*, *Phys. Rev. D* **61** (2000) 022001.
- [233] T. A. Wagner, S. Schlamminger, J. H. Gundlach, and E. G. Adelberger, *Torsion-balance tests of the weak equivalence principle*, *Class. Quant. Grav.* **29** (2012) 184002, [[arXiv:1207.2442](#)].
- [234] P. Touboul et al., *MICROSCOPE Mission: First Results of a Space Test of the Equivalence Principle*, *Phys. Rev. Lett.* **119** (2017), no. 23 231101, [[arXiv:1712.01176](#)].
- [235] V. V. Flambaum, B. T. McAllister, I. B. Samsonov, and M. E. Tobar, *Searching for Scalar Field Dark Matter using Cavity Resonators and Capacitors*, [[arXiv:2207.14437](#)].
- [236] P. Moxhay and K. Yamamoto, *Peccei-Quinn Symmetry Breaking by Radiative Corrections in Supergravity*, *Phys. Lett. B* **151** (1985) 363–366.
- [237] B. D. Fields, K. A. Olive, T.-H. Yeh, and C. Young, *Big-Bang Nucleosynthesis after Planck*, *JCAP* **03** (2020) 010, [[arXiv:1912.01132](#)]. [Erratum: *JCAP* 11, E02 (2020)].
- [238] K. Mukaida and K. Nakayama, *Dynamics of oscillating scalar field in thermal environment*, *JCAP* **01** (2013) 017, [[arXiv:1208.3399](#)].
- [239] K. Mukaida and K. Nakayama, *Dissipative Effects on Reheating after Inflation*, *JCAP* **03** (2013) 002, [[arXiv:1212.4985](#)].
- [240] K. Mukaida, K. Nakayama, and M. Takimoto, *Fate of Z_2 Symmetric Scalar Field*, *JHEP* **12** (2013) 053, [[arXiv:1308.4394](#)].
- [241] **Planck** Collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [[arXiv:1807.06209](#)]. [Erratum: *Astron. Astrophys.* 652, C4 (2021)].
- [242] J. Crowder and N. J. Cornish, *Beyond LISA: Exploring future gravitational wave missions*, *Phys. Rev. D* **72** (2005) 083005, [[gr-qc/0506015](#)].
- [243] V. Corbin and N. J. Cornish, *Detecting the cosmic gravitational wave background with the big bang observer*, *Class. Quant. Grav.* **23** (2006) 2435–2446, [[gr-qc/0512039](#)].
- [244] S. Sato et al., *The status of DECIGO*, *J. Phys. Conf. Ser.* **840** (2017), no. 1 012010.
- [245] **Planck** Collaboration, Y. Akrami et al., *Planck 2018 results. X. Constraints on inflation*, *Astron. Astrophys.* **641** (2020) A10, [[arXiv:1807.06211](#)].
- [246] M. Punturo et al., *The Einstein Telescope: A third-generation gravitational wave observatory*, *Class. Quant. Grav.* **27** (2010) 194002.
- [247] D. Reitze et al., *Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO*, *Bull. Am. Astron. Soc.* **51** (2019), no. 7 035, [[arXiv:1907.04833](#)].
- [248] A. Vilenkin, *Quantum Fluctuations in the New Inflationary Universe*, *Nucl. Phys. B* **226** (1983) 527–546.
- [249] K. Enqvist, K. W. Ng, and K. A. Olive, *Scalar Field Fluctuations in the Early Universe*, *Nucl. Phys. B* **303** (1988) 713.